THE UNIVERSITY
of EDINBURGH

Esrrb is a prominent target of Nanog that substitutes for Nanog function in ES cell self-renewal, reprogramming and germline development

Nicola Festuccia

Thesis presented for the degree of Doctor of Philosophy
The University of Edinurgh - 30/11/2012

I declare that the work described in this thesis is my own, except where otherwise stated.

Nicola Festuccia

Abstract

Embryonic stem (ES) cell pluripotency is sustained by a network of transcription factors centred on Oct4, Sox2 and Nanog. Whilst Oct4 and Sox2 expression is relatively uniform, ES cells fluctuate between states of high Nanog expression possessing high self-renewal efficiency, and low Nanog expression exhibiting increased differentiation propensity. Moreover, modulation in the level of Nanog expression determines the efficiency of ES cell self-renewal.

To identify genes regulated by Nanog, genome-wide transcriptional profiling was performed on ES cells expressing different Nanog levels and Nanog-null ES cells expressing a Nanog-ER^{T2} fusion protein in which nuclear Nanog activity can be regulated by tamoxifen. Surprisingly, only a minor fraction of the genes to which Nanog binds showed significant changes in response to Nanog induction. Prominent amongst Nanog-responsive genes is Estrogen-related receptor b (Esrrb). Nanog binds directly to *Esrrb,* enhances binding and pause-release of RNAPolII from the *Esrrb* promoter and stimulates *Esrrb* transcription. Consistent with these findings, elevation of Nanog produces a cell population that expresses uniformly high Esrrb levels. Moreover, double fluorescent reporter lines show that Esrrb and Nanog levels are strongly correlated in individual cells. Loss of Nanog is required for downregulation of Esrrb, which coincides with commitment to differentiate.

Esrrb overexpression results in LIF independent self-renewal, and blocks neural differentiation, even in the absence of Nanog. Cell fusion experiments between ES and neural stem (NS) cells show that elevated Esrrb levels allow the reprogramming of the NS cell genome in the absence of Nanog. Esrrb can rescue stalled reprogramming during the derivation of $Nanog^{-/-}$ induced pluripotent stem (iPS) cells. Moreover, targeted knock-in of Esrrb at the *Nanog* locus rescues the ability of Nanog null ES cells to maintain germ cell development beyond E12. Finally, Esrrb deletion abolishes the defining ability of Nanog to confer LIF-independent self-renewal to ES cells.

Together these data identify Esrrb as a critical downstream mediator of Nanog function.

To my mother, for all she did and I am never going to forget

Acknowledgments

I am greatly indebted to my supervisor Ian Chambers for giving me the opportunity to carry out the research described in this thesis. I am especially grateful for the freedom I enjoyed during the years spent working in his group. Without having the opportunity to pursue my ideas and the constant encouragement, material support and scientific guidance he gave me none of this work would have been possible. I am also grateful to him for the possibility to travel and present my work abroad. In particular, I wish to thank him for introducing me to Japan, its people and culture.

A special thank goes to Pablo Navarro for encouraging me during the difficult moments of my PhD, for helping me focus my research, for the inspiring scientific discussions we have almost daily, and, finally, for becoming a good friend.

My gratitude goes to Alessia Gagliardi for her tender support during these years and for sharing with me both the joyful and the hard moments we had in Edinburgh.

Thanks to Rodrigo Osorno for being open to share results and reagents, for his help with my work and for being an inspirational co-author.

Thanks to Douglas Colby for all his support and for the long hours spent in tissue culture to help the progress of my research. A special thank to him also for his incredible patience.

I am grateful to Violetta Karwacki for giving me the opportunity to base my work on her results, and for helping me when I needed her advice.

I am indebted to Nick Mullin for his daily help, especially during the first year I spent in the group, and for introducing me to molecular biology.

Thanks also to Frederick Wong and Jing Chao Zhang for being enthusiastic colleagues, and for always taking time to help me with my experiments.

I am grateful to my second supervisors Sally Lowell and Tilo Kunath, and to Val Wilson and Keisuke Kaji, for their patience in discussing my work and for the enthusiasm they always show towards my results.

My gratitude also goes to Florian Halbritter and Simon Tomlinson. Their collaboration, support and expertise in meaningfully analysing the data generated in my experiments has been instrumental in drawing some of the key conclusions of this thesis.

My work would have been impossible without the support from Simon Monard, Olivia Rodriguez, Joe Mee, Lynsey Robertson, Sally Inverarity, Carolyn Manson and the animal house staff. Equally, I am indebted to the Medical Research Council for generously supporting my studies.

Thanks to Eleni Chantzoura for being there every time I needed to relax or someone to talk to.

Thanks to James O'Malley for his encouragement and for being a friend from the first days of our PhD.

I have no words to express my love and gratitude to the person that shared emotions and joy, anguish and everyday little miseries, fun and boredom with me over these years in Edinburgh. It is difficult to explain how different and empty my life here would have been without you, Roberta.

Thanks to Giuseppe for being a good friend, for his patience and for bringing his good mood to our home almost every day since he arrived.

Thanks to Sophie for her love, tender affection and for all the happy hours we spent together. You are one of the reasons why I learned to love this country.

I want to express my gratitude and affection to Cristina for her friendship and for sharing my interests and passions.

I also want to thank Simona for always spoiling me and making me feel at home. We all miss you.

Thanks to David, Carolina, Francesca, Claudia and Giulia for arguing, dreaming, discussing, celebrating, or just eating with me so often. You made me feel part of a new family here in Edinburgh.

I wish to thank Davide, Andrea, Laura, Giulia, Mario, Glenda and all my friends back home for growing up with me and for the warm welcome they keep giving me every time I return to Milan.

Finally I wish to thank my mother for her unconditional love and for teaching me how to be curious and restless, my father for being an example of determination and supporting me throughout my education, and my brother for being so often an example. I want to express my love and gratitude also to my grandmother and my uncle for taking care of me when I needed them and encouraging me when I decided to study biology.

Table of contents

Index of tables

Index of Figures

Chapter 1:Introduction

1.1: Pluripotency and development.

Specification of pluripotent cell identity is a fundamental trait of mammalian development. Pluripotency is established during the formation of the inner cell mass (ICM) in pre-implantation embryos. The resulting epiblast cells of the inner cell mass are the cells from which all specialised cells that make up the developing embryo, and indeed all tissues of the adult organism, trace their origins to. A state that shares common traits with pluripotency is established again around E8.5, when primordial germ cells (PGCs) begin transcriptional and epigenetic resetting before initiating gametogenesis, setting the stage for a new cycle of development.

1.1.1: From fertilization to implantation: specification of pluripotency during pre-implantation development.

Fertilisation triggers completion of meiosis in the mouse oocyte and results in the formation of a 1 cell zygote in which the maternal and paternal genomes are still organised in two independent pronuclei. The first cleavage leads to the formation of two diploid cells which undergo a series of consecutive divisions to generate increasing number of progressively smaller cells, the blastomeres. The zygote is transcriptionally silent and initially relies on maternally inherited mRNAs and proteins (Bachvarova, 1985). Transcriptional activation of the zygotic genome occurs in a minor and major wave at the 1 and 2 cell stage respectively (Aoki et al., 1997; Hamatani et al., 2004) and is accompanied by active degradation of maternal transcripts (Bachvarova and Moy, 1985; Paynton et al., 1988; Piko and Clegg, 1982). After the third division the 8 cell embryo, now called morula, is constituted by relatively similar blastomeres. Before dividing further, the 8 cell morula undergoes a process, defined compaction, in which adherens and, subsequently, tight junctions

are established between the blastomeres. The process of compaction leads to polarisation of the blastomeres (Johnson and Ziomek, 1981), and sets the conditions for future lineage decisions.

During the transition from the compacted morula to the late blastocyst stages two fundamental lineage decisions occur in the developing embryo: first the specification of the trophectoderm (TE) and the inner cell mass and subsequently the segregation of the primitive endoderm (PE) from the pluripotent epiblast (reviewed in Cockburn 2010). The specification of the TE lineage is dependent on the function of two crucial transcription factors, Tead4 (Nishioka et al., 2008; Yagi et al., 2007) and Cdx2 (Strumpf et al., 2005), while three transcription factors are required for the formation of the pluripotent ICM: Oct4 (Nichols et al., 1998), Sox2 (Avilion et al., 2003) and Nanog (Chambers et al., 2003; Mitsui et al., 2003). The blastomeres of the compacted 8 cell morula undergo both symmetric and asymmetric divisions that establish two distinct populations of polarised outer cells and apolar inner cells (Johnson and Ziomek, 1981). It was recently proposed that the extensive cell-cell contacts that characterise inner cells lead to activation of Hippo signalling in this population (Nishioka et al., 2009). Activated Hippo receptors promote phosphorylation of Yap proteins by Lat kinases, in turn resulting in their exclusion from the nucleus. Yap proteins are transcriptional coactivators essential for the function of Tead transcription factors. Yap becomes progressively confined to the nucleus of outer cells after the 8 cell stage, leading to transcriptional activation of Cdx2 by Tead4. In line with the crucial role of Yap in driving the initial events in TE specification, genetic ablation of Yap proteins results in embryonic lethality before the morula stage (Nishioka et al., 2009). Cdx2 activates expression of genes that are crucial for TE development and is involved in a reciprocal inhibitory transcriptional circuit with the master pluripotency regulator Oct4 (Niwa et al., 2005). At the 8 cell morula stage, Cdx2 starts being expressed and Oct4 is already present in all blastomeres (Dietrich and Hiiragi, 2007; Nishioka et al., 2009). Cdx2 is then downregulated in inner cells, so that at the 32 cell stage it is almost exclusively confined to the TE (Nishioka et al., 2009). Oct4 expression persists longer in outer cells and some Oct4 positive TE cells are still observed in the 96 cell blastocyst

(Dietrich and Hiiragi, 2007). As a consequence, it appears that the fully committed population of outer TE cells established in the E3.5 blastocyst (Pedersen et al., 1986) is defined by high Cdx2 expression rather than lack of Oct4. The notion that coexpression of Cdx2 and Oct4 in ES cells leads to the repression of Oct4 transcriptional activity and TE differentiation (Niwa et al., 2005) may explain how outer $Oct4^{+}/Cdx2^{+}$ cells commit to trophectoderm fate.

In the 32 cell morula, functional tight junctions are formed between TE cells and a water influx actively driven by the trophectoderm layer starts forming a large internal cavity known as the blastocoel. The ICM is localised on one side of this cavity, making contacts with both the TE and the blastocoel. As observed for Oct4 and Cdx2 during trophectoderm specification, segregation of the PE and the epiblast is driven by the contrasting action of two key transcription factors, Nanog and Gata6 (Chambers et al., 2003; Mitsui et al., 2003; Singh et al., 2007). Gata6 activity, in conjunction with another Gata factor, Gata4, is essential for endoderm specification: In Gata6 and Gata4 mutant embryos the PE is formed but fails to further differentiate into visceral endoderm (Koutsourakis et al., 1999; Morrisey et al., 1998; Soudais et al., 1995). Nanog and Gata6 are homogeneously expressed in the majority of the blastomeres up to the early blastocyst stage (Plusa et al., 2008). Instructive ERK signals (Chazaud et al., 2006), principally activated by fibroblast growth factor 4 (FGF4), result in Nanog downregulation in some ICM cells after the 32 cell stage (Plusa et al., 2008), leading to almost mutually exclusive Nanog and Gata6 expression in the 64 cell blastocyst (Chazaud et al., 2006; Dietrich and Hiiragi, 2007). Gata6 positive cells become localised to the prospective PE layer lining the blastocoel at E4.5, through mechanisms that involve relocalisation of inner Gata6 positive cells and apoptosis (Plusa et al., 2008). The pivotal role of FGF/ERK signalling in instructing PE specification is highlighted by defective endoderm formation in knockout embryos lacking FGF receptor 2 (Arman et al., 1998), FGF4 (Feldman et al., 1995), and Grb2 (Chazaud et al., 2006; Cheng et al., 1998), an adaptor protein involved in signal transduction downstream of the FGF receptor. In addition, recent in vitro culture experiments showed that treatment of 8 cell stage

embryos with ERK1/2 inhibitors results in the complete absence of PE at the blastocyst stage (Nichols et al., 2009).

The fast developmental events occurring during the first days after fertilisation culminate in the formation of the E4.5 blastocyst. At the time of implantation, the blastocyst is composed of an external trophectoderm layer that will contribute to the formation of the placenta, a monolayer of primitive endoderm cells lining the surface of the epiblast that faces the blastocoel and destined to generate the visceral and parietal endoderm, and, enclosed by these, a central mass of cells, the epiblast. The cells of the epiblast are characterised by high levels of Oct4, Sox2 and Nanog expression (Dietrich and Hiiragi, 2007; Plusa et al., 2008; Silva et al., 2009), have completed the process of X inactivation (Mak et al., 2004; Okamoto et al., 2004) and re-established levels of DNA methylation (Santos et al., 2002), H3K9 methylation (Santos et al., 2003) and H3K27 trimethylation (Erhardt et al., 2003) higher than those observed in the trophectoderm. These changes underlie the attainment of the pluripotent state that is required for this population to give rise to all tissues constituting the embryo proper and, later, the adult animal.

1.1.2: Pluripotency after implantation and primordial germ cells.

Primordial germ cells derive from the few cells of the proximal epiblast that activate expression of Prdm1 at embryonic day 6.25 (Ohinata et al., 2005) in response to instructive BMP4 and WNT signals originating from the extraembryonic ectoderm and the epiblast respectively (Ohinata et al., 2009). PGC precursors sequentially acquire expression of Prdm14 (Yamaji et al., 2008), relocate to the posterior extraembryonic mesoderm, upregulate Stella and acquire high levels of alkaline phosphatase expression (Sato et al., 2002). Prdm1 and Prdm14 drive PGC specification by repressing the transcriptional somatic programme activated in the epiblast after implantation and initiating epigenetic reprogramming (Kurimoto et al., 2008; Ohinata et al., 2005; Yamaji et al., 2008). Around E8.0 PGCs start migrating to the genital ridges and undergo profound epigenetic changes (Seki et al., 2005; Seki

et al., 2007), that will be completed between E11.5 and E13.5 in the developing gonads (Hajkova et al., 2008; Hajkova et al., 2002; Hajkova et al., 2010).

Intriguingly, similarities exist between the transcriptional and epigenetic state characterising PGCs during their migration to the genital ridges and the pluripotent population of cells constituting the epiblast of pre-implantation embryos. Both cell populations display similar expression of key pluripotency regulators. In the post-implantation epiblast, Oct4 is progressively downregulated in an anterior to posterior fashion at the onset of gastrulation (Yeom et al., 1996), but expression is maintained in the newly specified PGCs and there increases after E7.25 (Yoshimizu et al., 1999). Sox2 expression also appears to be regained in primordial germ cells after becoming restricted to the neuroectoderm around E7.0 (Avilion et al., 2003; Kurimoto et al., 2008). Declining Nanog expression (Osorno et al., 2012) is upregulated again around E7.75 in PGCs accumulating in the posterior extraembryonic mesoderm region (Yamaguchi et al., 2005). In addition, these pluripotency factors, necessary for the specification of pluripotent cells in the early embryo ICM (Avilion et al., 2003; Mitsui et al., 2003; Nichols et al., 1998), are also crucial for PGC development. Oct4 is essential for the development of the germ line (Kehler et al., 2004) and Nanog is required for the survival of PGCs beyond E11.5 (Chambers et al., 2007; Yamaguchi et al., 2009). In line with their common expression of core pluripotency regulators, pluripotent lines can be derived from both ICM cells and developing PGCs after day E8.5 (Evans and Kaufman, 1981; Matsui et al., 1992). Furthermore, PGCs undergo epigenetic changes that have similarities with early development. In both developmental transitions, establishment of pluripotency is accompanied by reactivation of the inactive X chromosome in females (de Napoles et al., 2007; Mak et al., 2004; Okamoto et al., 2004; Sugimoto and Abe, 2007), preceded or accompanied by a wave of DNA demethylation (Hajkova et al., 2002; Hajkova et al., 2010; Mayer et al., 2000; Santos et al., 2002) and linked to establishment of high levels of H3K27me3 (Erhardt et al., 2003; Seki et al., 2007). Nonetheless, differences also exist between these two cell populations. Notably, repressive H3K9me2 marks are erased in migratory PGCs (Seki et al., 2005) and DNA methylation, which is re-established in the pluripotent ICM (Santos et al., 2002), is

more radically reset in PGCs, involving progressive erasure of maternal and paternal imprints (Hajkova et al., 2002; Seki et al., 2005) which are protected from demethylation in early embryos (Nakamura et al., 2007; Nakamura et al., 2012). Possibly, migratory PGCs, which have not yet completed erasure of methylation marks at imprinted loci (Hajkova et al., 2002) and still display detectable levels of genomewide DNA methylation (Seki et al., 2005), are more closely related to the pluripotent cells of the epiblast than E11.5 germ cells, which embark in a second wave of radical epigenetic changes after their entry into the gonads (Hajkova et al., 2008; Hajkova et al., 2010).

1.2: In vitro culture of pluripotent stem cells.

1.2.1: From teratocarcinomas to embryonic stem cells

The field of pluripotent stem cell biology traces its origins to a period, spanning more than two decades, of vibrant interest in the study of teratocarcinomas (Andrews, 2002; Solter, 2006). Teratocarcinomas are tumours composed of disorganised aggregates of somatic tissues derived from all three embryonic germ layers that in humans normally occur in the ovaries and, more rarely, in the testis of young males. Teratocarcinomas are distinguished from the closely related teratomas by the presence of a component of histologically undifferentiated cells, termed embryonic carcinoma (EC) cells (Pierce and Verney, 1961), responsible for their malignant character. The study of teratocarcinomas, previously limited by the rare occurrence of such tumours in humans and mice, was the subject of a renewed interest after the report that the inbred 129 mice strain spontaneously develops these testicular tumours with an incidence of 1% (Stevens and Little, 1954). The embryonic origin of teratocarcinomas is suggested by the fact that they consist of a mixture of differentiated and more immature cell types, and in extreme cases present structures, called embryoid bodies, that closely resemble embryos at early stages of development (Pierce and Verney, 1961). In addition, upon intraperitoneal injection, teratocarcinoma cells also readily differentiate in embryoid bodies (Pierce and

Dixon, 1959) and these tumours can be observed in mice as early as at day 15 of development (Stevens, 1962). This led to speculation that EC cells are the stem cell component of embryonic origin of these tumours. It was proposed that EC cells can give rise to all differentiated cell types composing the teratocarcinoma by progressive differentiation, while retaining their ability to self-renew (Pierce and Dixon, 1959; Pierce et al., 1960). The stem cell origin of teratocarcinomas was formally demonstrated by the observation that single tumours cells could reconstitute the entire variety of somatic cell types observed in the parental population and give rise to EC cells after transplantation into syngeneic animals (Kleinsmith and Pierce, 1964). The hypothesis that teratocarcinomas were derived from embryonic cells that fortuitously maintained their undifferentiated character was further supported by the finding that such tumours could be derived by ectopic transplantation of both pre- and post-implantation embryos (Solter et al., 1970; Stevens, 1968). Subsequent studies were able to show that teratocarcinoma cells injected into a blastocyst can contribute to the developing embryo and give rise to adult chimaeras (Brinster, 1974), even though the observed contribution was low and germline transmission was not reproducibly achieved.

A fundamental contribution to the successful derivation of ES cells came from the first attempts to culture in vitro teratocarcinoma derived EC lines. The use of conditioned medium and embryonic feeders allowed the derivation and cloning of cell lines that could maintain their pluripotency after indefinite propagation in vitro (Evans, 1972; Kahan and Ephrussi, 1970; Martin and Evans, 1974; Rosenthal et al., 1970). Such studies provided the technical background for the direct derivation of pluripotent cell lines from pre-implantation mouse embryos in 1981 (Evans and Kaufman, 1981; Martin, 1981). These cells, thereafter called embryonic stem (ES) cells, proved to be markedly similar to EC lines, but show greater in-vitro differentiation potential and the absence of gross genomic abnormalities. Their full potential was unveiled by the demonstration that ES cells can extensively contribute to in vivo development and colonise the germline, and thus be transmitted to the offspring of chimaeric animals (Bradley et al., 1984). The uncompromised ability of these cells to successfully resume development was further stressed by the generation

of adult animals entirely originated from cultured ES cells in tetraploid complementation experiments (Nagy et al., 1993).

1.2.2: Other pluripotent cell lines.

Embryonic germ (EG) cells are pluripotent cell lines derived from E8.5-12.5 PGCs (Durcova-Hills et al., 2001; Matsui et al., 1992; Resnick et al., 1992). EG cells can be kept in culture by stimulation with LIF, basic fibroblast growth factor (bFGF) and stem cell factor (SCF) (Matsui et al., 1992; Resnick et al., 1992). Although they express markers characteristic of pluripotent cells and show very limited transcriptional differences with ES cells (Mise et al., 2008; Sharova et al., 2007), EG cells differ from ES cells in that they present signs of erasure of paternal and maternal imprints (Tada et al., 1998), consequent to the epigenetic changes occurring in the PGC population of origin. EG cells can contribute to embryonic development after blastocyst injection and occasionally show germline transmission, but developmental abnormalities are often observed in chimaeric animals (Tada et al., 1998). Recently it has been demonstrated that EG lines can be successfully derived from E8.5 PGCs by direct culture in 2i/LIF (see chapter 1.3.3) (Leitch et al., 2010).

Human ES cells (hESC) were first derived on a fibroblast feeder layer in medium containing basic FGF (Thomson et al., 1998). The development of feeder free culture conditions (Klimanskaya et al., 2005) revealed the existence of growth factor dependence differences between mouse ESC (mESC) and hESC. hESC are not responsive to LIF signalling (Thomson et al., 1998) but are maintained by conjunct stimulation with FGF and Activin (Vallier et al., 2005). A crucial difference between mESC and hESC resides in the X chromosome inactivation status. Female hESC show variable degrees of non-random X chromosome inactivation (Hall et al., 2008; Shen et al., 2008; Silva et al., 2008b). It has recently become clear that it is possible to derive hESC lines that maintain two active X chromosome by culture under physiological oxygen concentrations (Lengner et al., 2010) and that, like mESC, these hESC lines are competent to undergo random X inactivation upon

differentiation. It was thus proposed that a clonal selection mechanism underlies the generation of lines carrying one inactive X chromosome after culture in atmospheric oxygen (Lengner et al., 2010) and long term passaging (Silva et al., 2008a).

Epiblast stem cells (EpiSC) are cell lines derived from the post-implantation blastocyst (Brons et al., 2007; Tesar et al., 2007). Despite being pluripotent, as shown by their differentiation potential in teratoma forming assays and after transplantation into post-implantation cultured embryos (Huang et al., 2012), EpiSC cells differ from ES cells in that they have lost the ability to contribute to chimaeric animals when introduced into pre-implantation embryos (Tesar et al., 2007). EpiSC also present a distinct growth factor dependence, and are maintained in culture by conjunct exposure to Activin and basic FGF. In accordance with their in-vivo origin, one of the two X chromosomes is inactive in EpiSC lines derived from female embryos (Guo et al., 2009). Murine EpiSC present similarities with hESC or human induced pluripotent stem (iPS) cells, including growth factor dependence, X inactivation status, and a transcriptional profile reminiscent of the post-implantation epiblast.

1.3: Extrinsic regulation of ES cell self-renewal

1.3.1: LIF signalling.

ES cell derivation initially required the use of both feeder layers and EC conditioned media (Evans and Kaufman, 1981; Martin, 1981). It was subsequently shown that rat liver buffalo cells secrete a factor, named differentiation inhibitory activity (DIA), that promotes ES cell self-renewal in the absence of a feeder layer (Smith and Hooper, 1987). Two studies independently identified DIA as the glycoprotein Leukemia Inhibitory Factor (LIF) (Smith et al., 1988; Williams et al., 1988). The importance of LIF signalling in maintaining ES cell self-renewal was confirmed by the fact that $LIF^{-/-}$ MEFs are almost unable to sustain ES cell cultures (Stewart et al., 1992). LIF is part of the IL-6 family of cytokines. In this family binding of the

ligand triggers homo or heterodimerisation of the common chain gp130 receptor and leads to transphosphorylation of the receptor associated Jak kinases. Activated Jak kinases in turn phosphorylate the receptor intracellular domains, providing docking sites for STATs and other SH2 domain containing proteins. Phosphorylated STATs can then dimerise, migrate to the nucleus and activate transcription (Kishimoto et al., 1994; O'Shea et al., 2002). In ES cells, signalling by LIF leads to the heterodimerisation of LIF receptor (LIFR) and gp130 (Gearing et al., 1992; Gearing et al., 1991). It was subsequently shown that signalling mediated by the gp130 intracellular domain in sufficient to sustain ES self-renewal. Stimulation with IL-6 and soluble IL6R maintains ES cell undifferentiated (Yoshida et al., 1994) even in the absence of LIFR (Chambers et al., 2003), and allows the derivation of ES lines (Nichols et al., 1994). Similarly, fusion of the intracellular domains of either LIFR or gp130 to the extracellular portion of the G-CSF receptor, and subsequent stimulation with G-CSF, showed that signalling through gp130, but not LIFR, is sufficient to sustain ES cell self-renewal (Niwa et al., 1998). Although different STAT proteins are expressed in ES cells, it was shown that only STAT3 is activated in response to LIF signalling (Boeuf et al., 1997). Gp130 dimerisation triggers STAT3 phosphorylation (Matsuda et al., 1999), that in turn is required for DNA binding and activation of transcription (Boeuf et al., 1997; Niwa et al., 1998). These studies also showed that the transcriptional activity of STAT3 is essential for ES cell self-renewal, since overexpression of a dominant negative form of this factor causes differentiation (Boeuf et al., 1997; Niwa et al., 1998). Conversely activation of a STAT3-ER fusion protein by 4OH-tamoxifen is sufficient to sustain ES cells cultures at high density (Matsuda et al., 1999). Activation of the LIF receptor in ES cells also triggers the ERK/MAPK cascade via phosphorylation of JAK, Grb2 and SHP-2 proteins (Burdon et al., 1999; Dance et al., 2008; Ernst et al., 1996; Matsuda et al., 1999; Niwa et al., 1998). Since ERK activity is instrumental in promoting ES cell differentiation (Kunath et al., 2007; Stavridis et al., 2007), a balance must exist so that the self-renewal function of LIF signalling dominates its ability to trigger differentiation.

The requirement of LIF for ES cell self-renewal is in striking contrast with the notion that disruption of gp130 mediated signalling does not compromise pre-implantation development. Neither deletion of LIF, LIFR, gp130 nor STAT3 results in a phenotype at the blastocyst stage (Li et al., 1995; Stewart et al., 1992; Takeda et al., 1997; Yoshida et al., 1996). The earliest phenotype manifests in STAT3 knock-out embryos, that show a reduction in the size of the epiblast at E6.5 and reduced dimensions of outgrowths after ICM explant (Takeda et al., 1997). Interestingly, the reciprocal pattern of LIF and LIFR/gp130 expression during development would suggest that LIF signalling plays an important role in the development of the ICM. Soluble LIF is expressed in the trophectoderm from the morulae to the egg cylinder stage, and both LIFR and gp130 are expressed in the ICM from the blastocyst stage until implantation (Nichols et al., 1996). This apparent contradiction might have been solved by a later study showing that LIF signalling protects cells in the ICM from differentiation during diapause (Nichols et al., 2001). In addition, LIF seems to play an independent role in the endometrial tissue and is required for the correct implantation of the embryo (Stewart et al., 1992).

Recent studies showed that LIF stimulation has a profound transcriptional consequence in ES cells (Sekkai et al., 2005; Trouillas et al., 2009) and individuated in Klf4 one of the pluripotency factors most sensitive to LIF signalling (Niwa et al., 2009).

1.3.2: BMP signalling.

When cultured in serum free media, ES cells differentiate into neurons. In 2003, it was demonstrated that addition of BMP4 and LIF to chemically defined N2B27 medium allows the efficient propagation of ES cells in the absence of serum (Ying et al., 2003). Signalling from BMPs via ALK receptors triggers the phosphorylation of SMAD transcription factors, which in turn activate *ID* genes transcription (Shi and Massague, 2003; Ying et al., 2003). ID proteins bind to ubiquitously expressed basic helix-loop-helix transcription factors, such as E47 and E2-2, impair their ability to bind DNA (Benezra et al., 1990) and prevent activation of a neural differentiation

programme (Ying et al., 2003). Activation of *ID* genes is a crucial component of BMP activity in ES cells, since forced expression of ID1 maintains ES cells undifferentiated in the absence of BMP4 (Ying et al., 2003).

1.3.3: FGF/ERK, WNT signalling and 2i culture conditions.

Two othe signalling pathways, the FGF/ERK and the Wnt/glycogen synthase kinase-3 (GSK-3) cascades, play an important role in balancing ES cell self-renewal and differentiation.

Autocrine FGF signalling is required to prime ES cells for differentiation (Kunath et al., 2007) through activation of ERK phosphorylation (Stavridis et al., 2007). Conversely, suppression of ERK activity has been shown to promote ES cell self-renewal (Burdon et al., 1999; Kunath et al., 2007; Stavridis et al., 2007; Ying et al., 2008).

A role for Wnt signalling in promoting self-renewal had been suggested by the observation that ES cells treated with small molecule inhibitors of GSK-3 can be maintained undifferentiated for a short period of time in the absence of LIF (Sato et al., 2004). Similarly, Wnt proteins were shown to synergise with LIF in promoting self-renewal of mouse ES cells (Ogawa et al., 2006). In addition, genetic ablation of GSK activity in ES cells compromises their ability to differentiate (Doble et al., 2007).

Combining these two lines of evidence, a new chemically defined medium based on the concomitant inhibition of FGF/ERK signalling and GSK-3 activity was derived that allows the indefinite propagation of completely undifferentiated ES cells in the absence of BMP and LIF signals. Under these conditions, ES cells can be maintained even after ablation of STAT3 (Ying et al., 2008). Nonetheless, GSK-3 inhibition, ERK1/2 inhibition and LIF signalling act in an additive way, and in combination allow optimal ES cell self-renewal (2i/LIF) (Wray et al.; Ying et al., 2008). In ES

cells, GSK inhibition acts primarily by stabilising β-catenin and elevating its nuclear levels. Direct transcriptional activation by β-catenin is not required and this protein seems to act by relieving Tcf3 repression of the pluripotency network (Cole et al., 2008; Wray et al.). Recently, Esrrb was identified as the principal target of the GSK-3/Tcf3 axis in ES cells, and it was shown that GSK-3 inhibition cannot sustain the self-renewal of Esrrb$^{-/-}$ cells (Martello et al., 2012).

1.4: Intrinsic determinants of ES cell self-renewal.

ES cell self-renewal is regulated by the activity of a complex transcriptional network centred around the activity of three core pluripotency factors: Oct4, Sox2 and Nanog (reviewed in (Chambers and Tomlinson, 2009; Jaenisch and Young, 2008)). The activity of a number of additional proteins (Ivanova et al., 2006; Jiang et al., 2008; Niwa et al., 2009) confers robustness to this network and ensures its ability to respond to external stimuli (Niwa et al., 2009). Among these, and of central relevance to this thesis, is Esrrb. The structure, the mechanistic basis of DNA recognition and transcriptional activation, the expression pattern during development and the function in vivo and in ES cells are outlined here for Oct4, Sox2, Nanog and Esrrb.

1.4.1: Oct4 and Sox2: cooperative DNA binding and transcriptional activation.

Oct4 is part of the Octamer class of transcription factors that binds the 8bp consensus sequence ATGCAAAT, initially identified in the promoter region of the immunoglobulin light chain genes (Clerc et al., 1988; Falkner and Zachau, 1984; Mason et al., 1985; Parslow et al., 1984; Sturm et al., 1988). Oct proteins, together with Unc and Pit proteins, form the POU (Pit Unc Oct) class of transcription factors (Sturm et al., 1988), hence the name *Pou5f1* for the Oct4 gene.

Sox proteins are part of the High Mobility Group (HMG) superfamily. The HMG family is composed of proteins that either make sequence specific or aspecific contacts with the DNA. The first group, including TCF-like proteins, yeast mating type products and Sox proteins (Laudet et al., 1993), recognises a (A/T)(A/T)CAAAG consensus sequence on the DNA (Giese et al., 1991; Harley et al., 1992; Sugimoto et al., 1991; van de Wetering et al., 1991). The Sox (SRY-box) family is defined by the high levels of conservation of the HMG domain with that of the testis determining protein SRY (Denny et al., 1992; Gubbay et al., 1990). Sox2 is part of the subgroup b1 of the Sox family, which also includes Sox1 and Sox3 (Bowles et al., 2000).

1.4.1.1: Oct4 and Sox2 proteins.

The Oct4 protein presents both N-terminal and C-terminal transactivation domains flanking a central DNA binding POU domain (Ambrosetti et al., 2000). Oct factors DNA binding domain is bipartite (Sturm and Herr, 1988), including a low affinity POU specific (POU_S) domain and a high affinity homeodomain (POU_{HD}) (Klemm and Pabo, 1996) (**Figure 1.1**). The POU_S domain makes contact with the ATGC sequence in the 5' half of the octamer site, and the POU_{HD} binds the adjacent AAAT sequence, with both domains contacting primarily the major groove of the DNA and binding to opposite sides of the double helix (Klemm et al., 1994). Despite the presence of an apparently unstructured linker between the POU_{HD} and the POU_S domains, the length of which varies between Oct proteins and can be altered without severely compromising binding to the DNA (Sturm and Herr, 1988), and despite the absence of direct contacts between the two domains bound to DNA (Klemm et al., 1994), no spacing can be introduced between the two halves of the octamer site (Klemm and Pabo, 1996). The affinity of the POU_S domain for DNA is significantly augmented by the presence of an already bound POU_{HD} and introduction of as little as 2 base pairs between the ATGC and AAAT sequences disrupts cooperative DNA recognition (Klemm and Pabo, 1996).

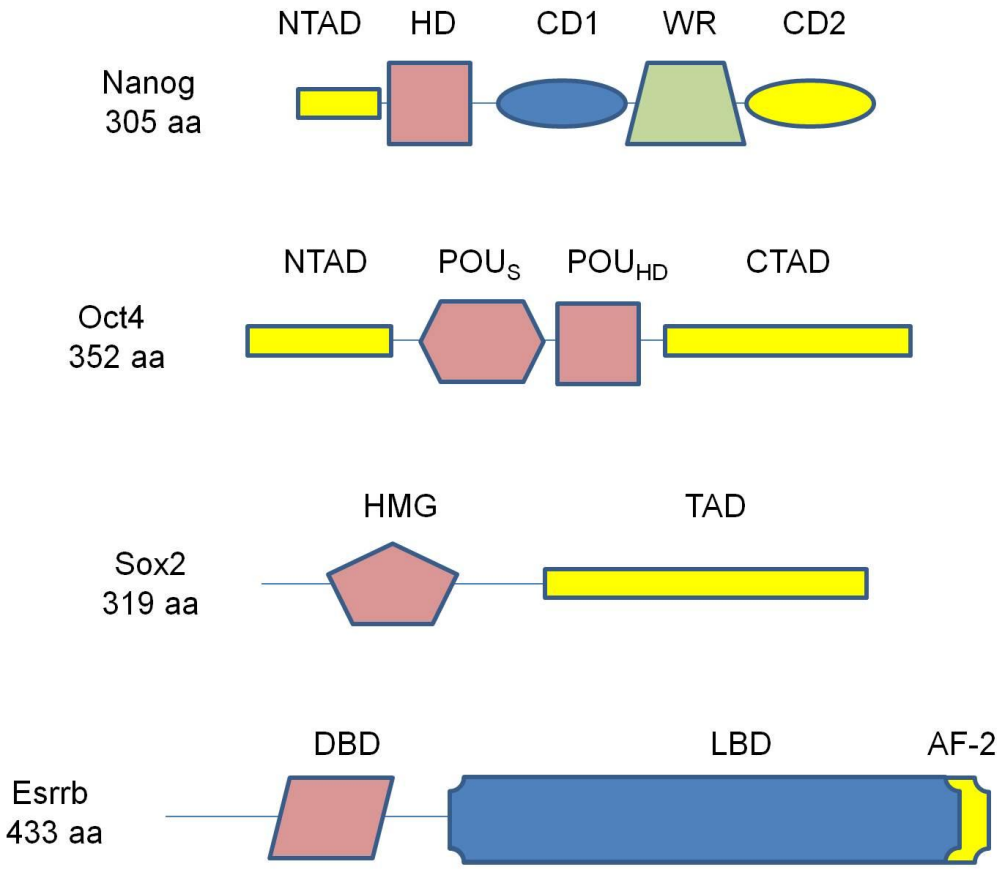


Figure 1.1: Nanog, Oct4, Sox2 and Esrrb protein structure.

Schematic representation of Nanog, Oct4, Sox2 and Esrrb dividing each protein into structural, functional or putative domains. TAD: Transactivation domains; HD: Homeodomain; WR: Tryptophan repeat; POU_S: POU specific DNA binding domain; POU_{HD}: POU homeodomain; HMG: High Mobility Group domain; DBD: DNA Binding Domain; LBD: Ligand Binding Domain. Transactivation domains are in yellow, DNA binding domains in pink. Adapted from (Chambers and Tomlinson, 2009).

The Sox2 protein is constituted by a short N terminal domain, followed by a DNA binding HMG domain and a C-terminal transactivation domain (Ambrosetti et al., 2000) (**Figure 1.1**). Sox2 DNA binding domain, as observed for other HMG class proteins, is composed by three α helixes and a β strand pack that arrange orthogonally in a L-shaped structure of which helixes 1 and 2 form the major wing and helix3 and the b strand form the minor wing (Weiss, 2001). The major wing makes base specific contacts with the minor groove of the DNA, expanding it and causing bending of the double helix with an angle of 90 degrees (Remenyi et al., 2003).

Oct4 and Sox2 binding sites are found in close proximity to one another at the promoter of many important pluripotency genes, including UTF1 (Nishimoto et al., 1999), FGF4 (Yuan et al., 1995), Nanog (Kuroda et al., 2005; Rodda et al., 2005), Oct4 (Chew et al., 2005; Okumura-Nakanishi et al., 2005) and Sox2 (Chew et al., 2005; Tomioka et al., 2002). In addition, searching both Oct4 and Sox2 bound regions identified by genomewide chromatin immunoprecipitation (ChIP) studies for overrepresented sequences led to the identification of very similar Oct and Sox binding motifs (Chen et al., 2008). These observations can be rationalised in light of the notion that binding to the DNA and activation of transcription by Oct factors and Sox2 is highly cooperative (Ambrosetti et al., 1997; Ambrosetti et al., 2000). Structural studies revealed that Sox2 binding to the minor grove of the DNA places its HMG domain on the same side of the double helix as the POU_S domain of the Oct protein partner. Extensive contacts between POU_S and Sox2 result in a more than ten-fold increase of the low affinity POU domain for the DNA (Williams et al., 2004), orders the relatively unstructured C terminal β strand of the Sox2 HMG (Remenyi et al., 2003), and stabilises the contact of both domains with the double helix (Williams et al., 2004). The extent of this cooperative interaction is strongly dependent on the distance between the octamer site and the Sox2 binding motif, deteriorating with an increase in the spacing, and might explain the selectivity shown by Oct/Sox elements for different Octamer proteins (Williams et al., 2004).

1.4.1.2: Oct4 and Sox2 expression during development and consequences of loss of function.

Oct4 and Sox2 messenger RNAs are maternally inherited and detected in all blastomeres throughout early cleavage. Oct4 is homogeneously expressed in all cell of the inner cells mass but downregulated in the trophectoderm after E3.5, and is then highly expressed in the epiblast at E4.5 (Avilion et al., 2003; Dietrich and Hiiragi, 2007; Rosner et al., 1990; Scholer et al., 1990; Yeom et al., 1996). Seminal genetic studies indicated that Oct4 and Sox2 are required for early development. Inner cells from Oct4 or Sox2 null embryos fail to form a pluripotent ICM and are diverted to a trophoblast fate (Avilion et al., 2003; Nichols et al., 1998). In the epiblast of post-implantation embryos, Oct4 is progressively downregulated in an anterior to posterior fashion at the onset of gastrulation, but its expression is maintained in the newly specified PGCs and there increases after E7.25 (Rosner et al., 1990; Scholer et al., 1990; Yeom et al., 1996; Yoshimizu et al., 1999). Sox2 expression becomes restricted to the neuroectoderm around E7.0, is lost in germ cells around E7.0 and regained after E7.25. (Avilion et al., 2003; Kurimoto et al., 2008). In accordance with its role during pre-implantation development, Oct4 is essential for the development of the germ line: Oct4 null PGCs are lost between E9.5 and E10.5 as a consequence of a pronounced wave of premature apoptosis (Kehler et al., 2004).

1.4.1.2: Oct4 and Sox2 expression and function in ES cells.

In accordance with the notion that both factors are strictly required for self-renewal, Oct4 and Sox2 expression is homogeneous in ES cells (Avilion et al., 2003; Chambers et al., 2007). Oct4 levels are strictly controlled in the cells: an increase of more that 50% leads to differentiation into extraembryonic endoderm and mesoderm, whereas a reduction to some level below the 50% of wildtype expression triggers differentiation into trophoblast (Niwa et al., 2000). The effects of Sox2 deletion, which also results in trophoblast differentiation, closely mirror the consequences of

loss of Oct4, further highlighting the intimate connection between the activities of these two factors in ES cells (Masui et al., 2007). It was recently proposed that Sox2 is required in ES cells uniquely to sustain Oct4 expression, since forced expression of Oct4 is able to maintain ES cells undifferentiated in the absence of Sox2. Other Sox proteins are expressed in ES cells and might exert a certain level of functional compensation in Sox2 knockout lines (Masui et al., 2007).

In line with the requirement of these two factor for self-renewal, acute ablation of Oct4 and Sox2 triggers rapid and profound transcriptional changes in ES cells (Hall et al., 2009; Masui et al., 2007), placing these proteins at the core of the pluripotency network.

1.4.2: Nanog.

Nanog was conjunctly identified by two studies aiming at the identification of novel regulators of ES cell self-renewal. The first study took a bioinformatic approach to identify genes specifically expressed in ES cells, and subsequently tested the effects of overexpressing the identified candidates for which a function in pluripotent cells was not reported (Mitsui et al., 2003). The second was based on the functional screen of a cDNA library for genes that were able to confer LIF independence to ES cells (Chambers et al., 2003).

1.4.2.1: Nanog protein.

Nanog is a 305 amino acid long protein that can be classified as a divergent homeodomain transcription factor. The low levels of homology with other known homeodomain proteins make it impossible to assign Nanog to a specific family. Sequence homology with members of the closest family, Nk-2, is confined within the homeodomain and does not exceed 50% (Chambers et al., 2003; Mitsui et al., 2003). In addition, Nanog lacks the TN and NK2-SD domains characteristic of Nk-2 family proteins. Specific amino acid variants in the homeodomain, that are conserved in

other NK-2 proteins at positions involved in contacts with the DNA and which define sequence specificity, are also missing in Nanog (Jauch et al., 2008; Mitsui et al., 2003). Nanog orthologs have been identified and studied in rat, human (Chambers et al., 2003; Mitsui et al., 2003), primates (Hart et al., 2004) and in other vertebrates, including urodele amphibians (Dixon et al.), birds (Lavial et al., 2007) and fish (Camp et al., 2009).

The Nanog protein can be roughly divided into three distinct domains: an N-terminal domain, containing an acidic transactivation region, a homeodomain, and a C-terminal domain, which is in turn divided in two parts, CD1 and CD2, by an intervening region characterised by 10 repetitions of a sequences in which every fifth residue is a tryptophan (WR repeat) (Chambers et al., 2003; Mitsui et al., 2003) (**Figure 1.1**). The N terminal and the CD2 domains of mouse Nanog have been shown to possess a weak and a strong transactivatory activity respectively (Pan and Pei, 2005; Pan and Pei, 2003), while only the CD2 domain of the human protein can activate transcription (Oh et al., 2005). The homeodomain is responsible for the ability of Nanog to bind DNA. Despite substitutions in residues generally involved in conserving the architecture of the homeodomain in other members of the family, superimposition studies revealed that the overall structure of the core domain and the positioning of Cα atoms in the α helix backbone are strongly conserved in Nanog homeodomain (Jauch et al., 2008). Like other homeodomain proteins, Nanog presents two α helixes (1 and 3) running in an antiparallel fashion and a third helix (2) positioned almost orthogonally in a manner that facilitates its contacts with the major grove of the DNA (Jauch et al., 2008). The consensus binding motif emerged from SELEX and EMSA studies includes the canonical ATTA homeodomain recognition sequence and preference for G or C at the following two positions (Jauch et al., 2008; Mitsui et al., 2003). Interestingly, search for overrepresented sequences at Nanog bound genomic regions identified in ChIP-seq studies defined an Oct4/Sox2 binding motif as the Nanog binding consensus (Chen et al., 2008). This is of interest, since it highlights the diffuse clustering of these three transcription factors at the regulatory elements of many target genes of the pluripotency network, and might even reveal a certain level of dependency of Nanog binding on Oct4 and Sox2.

Finally, the WR region has been shown to mediate homo-dimerisation of mouse Nanog and is required for Nanog ability to sustain LIF independent self-renewal (Mullin et al., 2008).

Nanog protein is subject to different forms of post-translational modification that control its levels in the cells, as recently highlighted for human Nanog (Van Hoof et al., 2009). Nanog has been reported to be phosphorylated, and this could affect the stability of the protein (Moretto-Zita et al., 2010; Yates and Chambers, 2005). In addition, Nanog protein has been shown to be the target of caspase mediated proteolysis upon differentiation (Fujita et al., 2008).

Nanog engages in a complex network of protein-protein interactions. Initial studies reported Nanog interaction with Smad1 (Suzuki et al., 2006) and Sall4 (Wu et al., 2006). Immunoprecipitation of Nanog followed by mass spectrometry subsequently revealed interaction with a number of other mediators of ES cell self-renewal, such as Oct4, Nac1, Dax1 and Zfp281, and chromatin remodelers or repressor complexes, such as members of the SWI/SNF and NuRD complexes (Liang et al., 2008; Wang et al., 2006a). Among the identified interactors is Esrrb ((Wang et al., 2006a), see also supplementary material in (van den Berg et al., 2010)).

1.4.2.2: Nanog expression during development and in vivo consequences of loss of Nanog.

In contrast to Sox2 and Oct4, Nanog is absent in unfertilised oocytes and becomes expressed from the 8 cell stage onward, shortly before specification of the ICM (Chambers et al., 2003; Dietrich and Hiiragi, 2007; Guo et al.; Hart et al., 2004; Plusa et al., 2008). Nanog expression, initially uniform in all blastomeres, becomes heterogeneous in the ICM around E3.5 (Chazaud et al., 2006; Dietrich and Hiiragi, 2007). Maintenance or loss of Nanog expression is a key event in the choice between pluripotency and differentiation into primitive endoderm (Plusa et al., 2008). Cells that maintain Nanog expression will complete transcriptional and epigenetic

resetting, form the epiblast of the post-implantation embryo and hence give rise to all tissues constituting the adult body. Further stressing an intimate connection between Nanog expression and pluripotency in the pre-implantation embryo, ICM cells gradually switch from mono to bi-allelic Nanog expression during the transition from early to late blastocyst (Miyanari and Torres-Padilla, 2012). Nanog expression is then transiently down-regulated in the epiblast immediately before implantation (Chambers et al., 2003) and becomes detectable again in the posterior epiblast at the egg cylinder stage before the onset of gastrulation. Nanog expression is lost in the entire epiblast around E8.0, and before then in the cells that migrate through the primitive streak and differentiate into mesoderm (Hart et al., 2004; Osorno et al., 2012). Declining Nanog expression is upregulated again around E7.75 in PGCs accumulating posteriorly in the extraembryonic mesoderm region (Yamaguchi et al., 2005), is detectable in germ cells throughout migration to the gonads, and remains expressed by the majority of PGCs until E12.5, declining thereafter and becoming almost undetectable at E15.5 and E14.5 in male and female embryos respectively (Chambers et al., 2007; Yamaguchi et al., 2005) (**Figure 1.2**).

Nanog activity is essential at two different stages of development: in the pre-implantation embryo and, later on, in the germline. Initially, genetic ablation studies revealed that Nanog is required in vivo at the peri-implantation stage (Mitsui et al., 2003). Subsequently, detailed analysis indicated that in *Nanog*$^{-/-}$ E4.5 embryos an aberrant epiblast is formed that includes few cells showing very low Oct4 expression and, in female embryos, defective reactivation of the X chromosome (Silva et al., 2009). The finding that Nanog can be deleted in ES cells allowed the study of Nanog role in later stages of development. Nanog$^{-/-}$ cells contribute normally to development in chimaeric embryos generated by blastocyst injection (Chambers et al., 2007). In such embryos, PGC derived from the injected ES cells can be detected in the genital ridges until day E11.5, but contribution to the germline is undetectable in embryos dissected at days E12.5 or E13.5 of development. Knock-down studies seem to indicate that elevated levels of apoptosis and reduced proliferation during PGC migration to the genital ridges might contribute to this phenotype (Yamaguchi et al., 2009).

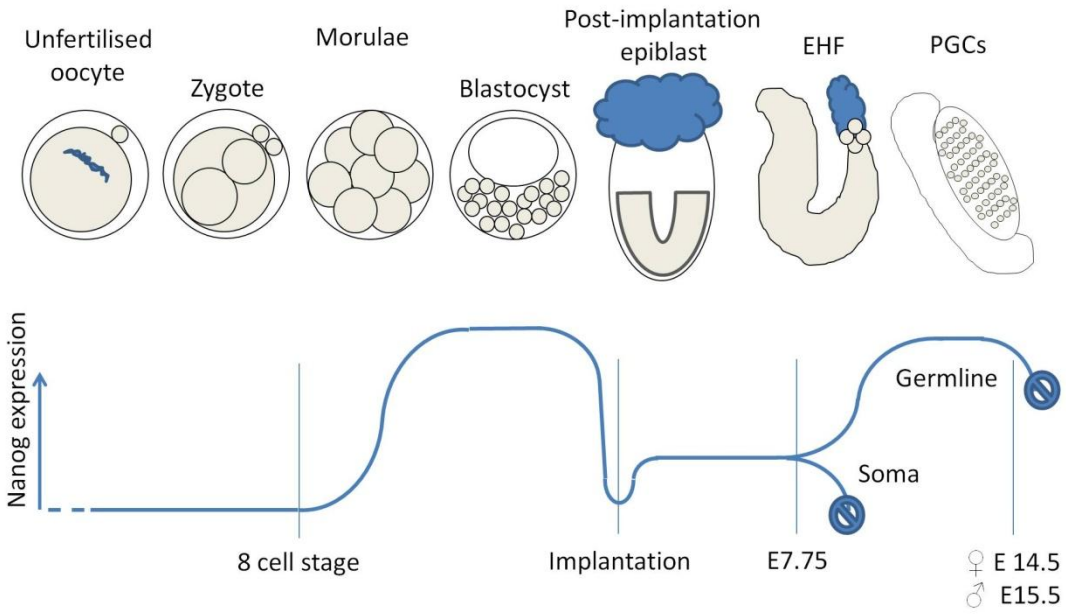


Figure 1.2: Nanog expression dynamics during murine development.

Nanog protein is first detected in all blastomeres of the 8 cell stage embryo. Around E3.5 Nanog expression becomes heterogeneous in the ICM and segregation of Nanog positive and negative cells leads to the formation of the epiblast and hypoblast in the E4.5 blastocyst. Around implantation Nanog is downregulated before its expression is detected again in the posterior part of the epiblast. From E6.5 until the onset of somitogenesis Nanog expression remains regionalised and progressively declines. At E7.75 Nanog expression, confined to the proximal posterior part of the epiblast, is upregulated again in PGC. Nanog remains expressed by the majority of PGC until E12.5 and declines thereafter becoming almost undetectable at E15.5 and E14.5 in males and females respectively. Neither Nanog transcripts nor proteins are present in the unfertilised oocyte.

The timing of Nanog expression and the phenotype of Nanog null embryos suggest that Nanog plays an essential role in vivo in driving acquisition of pluripotency. Nonetheless, there is currently no mechanistic insight on how Nanog promotes this transition in the developing embryo.

1.4.2.3: Nanog in vitro function in the acquisition and maintenance of pluripotency

Nanog is required for the successful derivation of ES cell lines (Mitsui et al., 2003). When the inner cell mass from a Nanog null blastocyst is isolated by immunosurgery and put in culture it fails to produce an outgrowth and differentiates in cells morphologically resembling trophoblast (Silva et al., 2009). The initial observations made by Mitsui and colleagues analysing knockout embryos and Nanog null ES cells led to the speculation that pluripotent cells are dependent on Nanog expression. Intriguingly, subsequent studies showed that it is possible to derive ES cells in which both Nanog alleles have been ablated by homologous recombination (Chambers et al., 2007). Combined with previous findings, this raised the hypothesis that the central role of Nanog in development is promoting the establishment of pluripotency, and that this factor is not strictly required once pluripotency is attained (Silva et al., 2009). Despite this, in pluripotent cells Nanog regulates the self-renewal efficiency. Nanog overexpression abolishes ES cell ability to differentiate in response to the withdrawal of the otherwise essential signals provided by LIF and BMP (Chambers et al., 2003; Ying et al., 2003) and loss of function genetic studies showed that Nanog null ES cells are more prone to differentiate (Chambers et al., 2007). More importantly, ES cells spontaneously fluctuate in culture between states of high Nanog expression, characterised by efficient self renewal, and low Nanog expression, with increased propensity to differentiate (Chambers et al., 2007). Conversely, in culture conditions that lock ES cells in a completely undifferentiated state (Ying et al., 2008), Nanog fluctuations are lost.

Nanog is able to promote reprogramming of hybrids generated by fusion between neural stem cells (NS) and ES cells (Silva et al., 2006). Overexpression of Nanog accompanied by a change in culture conditions has also been shown to drive reprogramming of EpiSC to ES-like cells (Osorno and Chambers, 2011; Silva et al., 2009). Even if it is not included in the quartet of canonical reprogramming factors, recent studies showed that Nanog expression in conjunction with other pluripotency factors, among which notably is Esrrb, allows the generation of induced pluripotent stem cells (iPS) from mouse embryonic fibroblasts (MEFs) (Buganim et al., 2012). Important indications about the role that Nanog plays in promoting the acquisition of pluripotency came from attempts to reprogram Nanog null somatic cells. Nanog is required for completion of the reprogramming process (Silva et al., 2009). Nanog null cells can transit to a state in which they acquire morphology and growth factor dependency characteristics of ES cells but fail to complete epigenetic resetting of their genome and do not turn on transcription of pluripotency factors from the endogenous alleles (Silva et al., 2009). In this state, resembling the partially reprogrammed intermediates observed during reprogramming of *Nanog* wild-type cells (Mikkelsen et al., 2008) "pre-iPS" cells do not silence expression of the transgenes encoding for the reprogramming factors. Enforcing Nanog expression in these lines allows transition to a fully reprogrammed state (Silva et al., 2009). Nanog was thus proposed to be exquisitely required for the acquisition of pluripotency in vitro, substantiating the current interpretation of its role during in vivo development (Theunissen and Silva, 2011).

Interestingly, the role of Nanog in driving attainment of naive pluripotency could be conserved throughout vertebrate evolution. Indications in support of a conserved role for Nanog in other species came from in vitro studies reporting that human, rat, chick and zebrafish Nanog orthologs are able to trigger completion of the reprogramming process in pre-iPS lines derived from Nanog null mouse ES cells (Theunissen et al., 2011a). Strikingly, mouse Nanog homeodomain (HD) was sufficient to complete reprogramming in this experimental setting. Thus, it was proposed that the homeodomain is the region of the Nanog protein principally responsible for Nanog function in early development.

1.4.3: Esrrb.

Estrogen related receptors (ERRs) belong to the NR3B subgroup of nuclear receptors, which is part of the wider NR3 class including estrogens, androgen, progesterone, aldosterone and cortisol receptors (reviewed in (Tremblay and Giguere, 2007)). The NR3B subgroup comprises three highly related members: Esrra, Esrrb and Esrrg (also named ERR1-3, NR3B1-3 or ERRα,β,γ). The first two member of this class, Esrra and Esrrb, were identified by screening a human testes cDNA library with a probe designed against the DNA binding domain of the estrogen receptor α (ERα) (Giguere et al., 1988). Ten years later, the presence of a third member of this family, Esrrg, was independently unveiled by biochemical and bioinformatic approaches (Eudy et al., 1998; Heard et al., 2000; Hong et al., 1999). In both mouse and human, a high degree of homology exist between these three proteins, with the DNA binding domain and the hormone binding domain being the most conserved regions, as commonly observed for other nuclear receptors. Esrrb and Esrrg are the most related proteins, with 99% homology in the DNA binding domains (93% with Esrra) and higher than 70% homology in the hormone binding domains (around 60% with Esrra), in both mice and humans (Heard et al., 2000; Hong et al., 1999). Not surprisingly, given the method used for their original identification, estrogen related receptors show the highest homology outside their subgroup to ERα (69% and 36% homology exists between the DNA and hormone binding domains of hEsrrg and hERa) (Giguere et al., 1988; Heard et al., 2000).

1.4.3.1: Esrrb protein.

Estrogen related receptors are characterised by a common nuclear receptor domain organisation. The N-terminal region of nuclear receptors is the domain that presents the highest sequence variation among the member of this family. This region confers to nuclear receptors weak transactivation activity on a fraction of their target genes. The central part of the protein includes a zinc finger domain responsible for DNA binding. The C-terminal part constitutes the hormone binding domain (HBD or LBD)

and is also responsible for the ligand dependent ability of nuclear receptors to activate transcription (Hollenberg and Evans, 1988; Kumar et al., 1987). In hormone regulated receptors, an α helix at the C-terminal end of the LBD, the AF-2 helix, is relocated after ligand recognition, becoming able to interact with coactivators and trigger transcriptional activation (Wurtz et al., 1996). In contrast, ERRs are able to mediate transcriptional activation in the absence of any bound ligand (Greschik et al., 2002; Heard et al., 2000; Hong et al., 1999; Vanacker et al., 1999a; Vanacker et al., 1999b; Yang et al., 1996). Domain deletion studies demonstrated that the LBDs of Esrra and Esrrg, and specifically residues in the AF-2 helix, are required for the transactivation ability of these proteins (Heard et al., 2000; Hong et al., 1999; Vanacker et al., 1999a; Vanacker et al., 1999b). A very weak transactivation activity is also conferred by the N-terminal domain of Esrrg in luciferase assays (Vanacker et al., 1999b).

X ray crystallography showed that the ligand binding domain of Esrrg is composed of 12 α helixes and, very similarly to ERα, adopts a canonical three layered α helix sandwich structure. The lack of ligand binding to ERRs is due to the reduced dimension of the ligand binding pockets in these receptors compared to ERα (Greschik et al., 2004; Greschik et al., 2002). The major difference between ERRs and ERα resides in the substitution of Leu-525 (ERα) with a bulkier residue, Phe-435 (Esrrg), that invades the ligand binding pocket and prevents docking of common steroidal estrogens (Wang et al., 2006b). The conservation of the residues lining the ligand binding cavity of Esrrb and Esrrg is striking, with only two amino acid changes (V313 to I, N346 to Y), making it possible to extrapolate the conclusions made analysing the structure of Esrrg to Esrrb (Greschik et al., 2002). In the Esrrg ligand binding domain the AF-2 helix is in an active conformation, able to interact with coactivators, in the absence of any ligand (Greschik et al., 2002).

Estrogen receptors bind to inverted repeats of the sequence AGGTCA spaced by three nucleotides (ERE). In contrast, ERRs recognise a conserved motif comprising a 5'-extended AGGTCA half-site. In vitro electrophoretic mobility shift assays (EMSA) and in vivo reporter assays have shown that the three base pairs preceding

the ERE half-site are crucial for DNA binding by this class of receptors, and identified a TCAAGGTCA consensus sequence (Johnston et al., 1997; Sladek et al., 1997; Vanacker et al., 1999a; Vanacker et al., 1999b; Yang et al., 1996). Given the high conservation observed among the DNA binding domains of the three ERRs, it is not surprising that EMSA assays (Vanacker et al., 1999b) and genome-wide ChIP-seq studies (Chen et al., 2008; Dufour et al., 2007) revealed an identical binding consensus for Esrra, Esrrb and Esrrg, and showed that ERRs share target genes. ERRs can bind to the DNA as monomers (Barry et al., 2006; Gearhart et al., 2003; Johnston et al., 1997; Pettersson et al., 1996), homodimers (Barry et al., 2006; Vanacker et al., 1999a) or heterodimers (Dufour et al., 2007), further extending the potential for cross regulation of transcription by these receptors. In addition, at least in vitro, ERRs can bind to palindromic ERE sites (Pettersson et al., 1996; Vanacker et al., 1999b), and form heterodimers with ERα (Johnston et al., 1997).

Resolution of the structure of the DNA binding domain of Esrrb bound to DNA (Gearhart et al., 2003) revealed a characteristic nuclear receptor zinc finger fold, in which an α helix, surrounded by the zinc binding cysteine residues, makes extensive contacts with the major groove of the DNA. In addition, immediately following the core zinc finger domain, the Esrrb DBD possesses a C-terminal extension, common to other nuclear receptor that bind extended ERE half-sites as monomers, which crosses the DNA phosphate backbone and makes additional contacts with the DNA minor groove. The C-terminal extension has been shown to be determinant for the interaction with the 3 nucleotide 5'-extension of the ERE half-site (Gearhart et al., 2003).

Despite the fact that no natural ligand is reported for ERRs, a variety of synthetic hormones has been shown to bind to these receptors and either inactivate (Greschik et al., 2004; Greschik et al., 2002; Wang et al., 2006b) or potentiate (Wang et al., 2006a) their transactivation ability. In particular, 4-OH tamoxifen and diethylstilbestrol (DES) have been shown to contact the ligand binding pocket of Esrrg and force rearrangements in the structure of the protein that lead to AF-2 helix

dislocation and inhibition of transactivation (Tremblay et al., 2001a; Tremblay et al., 2001b; Wang et al., 2006a).

Esrrb protein has been shown to interact with numerous other pluripotency factors, including Nanog, Oct4, Sall4 and Dax1, in mouse ES cells (van den Berg et al., 2010; Wang et al., 2006a). In addition, Esrrb contacts common estrogen receptors coactivators, like Ncoa-3, and components of the basal transcriptional machinery, including the mediator complex, TFIID and RNAPolII (van den Berg et al., 2010). It is thus possible that Esrrb may function as a bridge between pluripotency factors and component of the transcriptional initiation complex at the promoter of target genes of the pluripotency network.

1.4.3.2: Esrrb expression in vivo and phenotype of Esrrb loss of function.

Esrrb expression is detected at extremely high levels in 1 cell embryos (Guo et al., 2010), before transcription from the zygotic genome is activated (Schultz, 2002). Expression is maintained throughout pre-implantation development, becoming restricted to the ICM at the 32 cell stage. At the 64 cell stage, no expression is detected in cells expressing markers characteristic of the trophectoderm or primitive endoderm (Guo et al., 2010). After implantation, Esrrb expression is not detected in the epiblast and is restricted to the extra-embryonic ectoderm at E5.5. Later, expression is confined to the ectodermally derived regions of the amniotic fold at E6.5, and to the chorion in E7.5 embryos. Esrrb expression diminishes after fusion of the chorion with the ectoplacental cone, becoming restricted to the free margin of the chorion at E8.5 (Luo et al., 1997; Pettersson et al., 1996). Esrrb expression is reactivated in E11.5 post-migratory PGCs, and persists in germ cells until E14.5 and E15.5 in females and males embryos respectively (Mitsunaga et al., 2004). At E13.5 Esrrb expression is also detected in the developing brain (Mitsunaga et al., 2004). After birth, Esrrb expression is broad, with high levels detected in the testes (Mitsunaga et al., 2004), the inner ear (Chen and Nathans, 2007), the eye, the heart,

the kidneys and the thyroid (Bookout et al., 2005; Bookout et al., 2006). Lower levels of expression are detected in other organs, especially in the brain (Bookout et al., 2005; Bookout et al., 2006). The observation that high levels of Esrrb are detected in embryos at the 1 cell stage further suggests that Esrrb is expressed in the maturing oocyte (Guo et al., 2010).

Esrrb$^{-/-}$ embryos generated by mating heterozygous animals revealed that Esrrb expression is required for normal placental development. The chorion is nearly absent in Esrrb knockout E7.5 embryos, and abnormalities in the early placenta and the ectoplacental cone are evident at E8.5 (Luo et al., 1997). As a consequence, *Esrrb*$^{-/-}$ E9.5 embryo show evident signs of growth retardation and no viable *Esrrb*$^{-/-}$ embryos are observed after 10.5. The increased numbers of secondary giant cells and the almost complete absence of diploid trophoblast cells observed at E8.5 (Luo et al., 1997), coupled with the notion that DES treatment of trophoblast stem (TS) cells induces differentiation in ploypoid giant cells (Tremblay et al., 2001b), suggests that Esrrb is required for self-renewal of trophoblast cells during early placental development. Tetraploid rescue experiments showed that *Esrrb*$^{-/-}$ embryo defects are restricted to extraembryonic tissues at this time of development (Luo et al., 1997; Mitsunaga et al., 2004) and allowed the generation of viable adult knockout animals (Mitsunaga et al., 2004). Absence of Esrrb results in defective germline development, with reduced numbers of PGCs detected in the gonads of both male and female knockout embryos at E13.5-15.5 (Mitsunaga et al., 2004). Finally, epiblast specific Cre mediate deletion of Esrrb, demonstrated that Esrrb expression is required for the correct development of endolymph producing cells in the inner ear, leading to hearing and balance defects in null animals (Chen and Nathans, 2007).

1.4.3.3: Esrrb function in vitro.

Esrrb is heterogeneously expressed in ES cells (van den Berg et al., 2008), while no expression is detected in EpiSC or human ES cells (Xie et al., 2009). Genome-wide ChIP-seq studies revealed that Esrrb colocalises with other pluripotency factor at the regulatory elements of many transcriptional targets of the pluripotency network in ES

cells (Chen and Nathans, 2007; Marson et al., 2008). Furthermore, Esrrb has been shown to activate Nanog expression in conjunction with Oct4 and Sox2 (van den Berg et al., 2008). Published work has also suggested that Esrrb can drive LIF independent self-renewal, but did not incontrovertibly demonstrate that Esrrb overexpression sustains ES cells in the complete absence of LIF signalling (Zhang et al., 2008).

Knockdown studies concluded that Esrrb expression is required for ES cell-self-renewal (Ivanova et al., 2006). Nonetheless, it was recently demonstrated that it is possible to genetically delete Esrrb in ES cells. The resulting knockout lines show increased propensity to differentiate and a strict dependency on LIF stimulation, but retain pluripotency in culture and can contribute to the developing embryo (this thesis and (Festuccia et al., 2012; Martello et al., 2012)).

Esrrb is able to drive highly efficient reprogramming of EpiSC to ES cell pluripotency (Festuccia et al., 2012) and can substitute for Klf4 during iPS cell generation from MEFs (Feng et al., 2009). Furthermore, it was recently shown that Esrrb can promote reprogramming independently of the four canonical reprogramming factors, and is particularly efficient in this context when combined with Nanog (Buganim et al., 2012).

The three estrogen related receptors are able to activate transcription by interacting with a number of coactivators, that are common partners to other nuclear receptors (Tremblay and Giguere, 2007). Among these Ncoa3 is expressed at high levels in ES cells (Percharde et al., 2012; Wu et al., 2012), while PRC, PRNC-2, Ncoa1, and Tle1 show lower expression ((Percharde et al., 2012; Wu et al., 2012) and **Table3.1**). Recent studies demonstrated that Ncoa3 interacts with Esrrb in ES cells (Percharde et al., 2012; van den Berg et al., 2010), mediates its transcriptional activity, and is essential for efficient reprogramming and ES cell self-renewal (Percharde et al., 2012; Wu et al., 2012).

1.5: Aims of the thesis.

The aim of this thesis is the identification of Nanog target genes in ES cells. Despite the fact that genomewide ChIP studies have recently identified an extensive list of binding sites for the principal pluripotency factors, little is known regarding which of the bound genes are responsive to fluctuations of single transcriptional regulators. Here a list of genes that show prompt response to the perturbation of Nanog levels is identified. Based on this novel information, the present work aims at exploring the molecular mechanisms of transcriptional regulation by Nanog, using single responsive genes as a paradigm for understanding the wider effects of Nanog on gene expression.

Nanog heterogeneous expression has been recognised as a crucial determinant of cell fate decision between self-renewal and differentiation. This thesis aims at describing how fluctuations in Nanog level influence the expression of it target genes in single ES cells, and which are the functional consequences of such regulation.

The role of specific Nanog targets in sustaining ES cell self-renewal is also assessed in the following chapters. In particular this work compares the effects of elevating or ablating expression of a prominent Nanog target, Esrrb, in wild-type ES cells or cells genetically devoid of Nanog, with the objective of understanding whether Esrrb mediates part of Nanog activity and can functionally complement for loss of Nanog. Special attention is devoted to assessing the ability of Esrrb to rescue the reprogramming and germline development defects of Nanog null ES cells.

In a wider context, the data presented in this thesis aims at identifying differences and similarities among the multiple roles that Nanog plays at various stages during development.

Chapter 2:Materials and methods.

2.1 ES cell culture

2.1.1 ES culture materials

Phosphate-buffered saline (PBS; Sigma, cat. D8537)

GMEMβ/FCS/LIF

500 ml Glasgow minimum essential medium (GMEM; Sigma, cat. G5154)

51 ml fetal bovine serum

11 ml of 100× glutamine/pyruvate solution (see recipe)

5.5 ml of 100×MEMnon-essential amino acids (Invitrogen, cat. 11140-036)

570 μl of 0.1 M 2-mercaptoethanol solution (see recipe)

Supplemented with LIF (made in-house) at a final concentration of 100 U/ml

Glutamine/pyruvate stock solution

5.5ml of 100mMsodium pyruvate (Invitrogen, cat. 11360-039)

5.5 ml of 200 mM L-glutamine (Invitrogen, cat. 25030-024)

2-Mercaptoethanol stock solution, 1000×, 0.1M

200 μl of 2-mercaptoethanol (14.3 M; Sigma, cat. M6250)

28.2 ml ultra-high purity (UHP) water

Gelatin, 0.1%

Autoclaved 1% (w/v) solution of gelatin (Sigma, cat. G1890) in ultra-high purity (UHP).

Before use, 11 ml of 1% gelatin were added to 100 ml PBS.

Trypsin, 0.025%

Filter sterilized solution of 0.186 g EDTA (Sigma, cat. E5134) in 500 ml PBS.

5 ml chick serum (Sigma, cat. C5405)

5 ml of the concentrated trypsin stock (2.5%; Invitrogen, cat. 15090-046)

N2B27 and 2i/LIF media

50ml of DMEM:F12(1:1v/v, GIBCO cat. 12634010)

50ml Neurobasal (GIBCO cat. 21103049)

1ml of 100× glutamine (Invitrogen cat. 25030024)

1 ml of 100×MEMnon-essential amino acids (Invitrogen, cat. 11140-036)

100 µl of 0.1 M 2-mercaptoethanol solution (see recipe)

1ml of 100x N2 supplement (GIBCO cat. 17502048)

2ml of 50x B27 supplement(GIBCOcat. 17504044).

Supplemented where indicated with PD0325901 (1µM) (Axon, cat. 1408), CHIR99021 (3µM) (Axon,cat. 1386) and 100U/ml LIF.

2.1.2: ES cell passaging

ES cells were cultured on gelatin coated 25, 75 or 150 cm^2 flasks (IWAKI cat. 3100-025, 3110-075, 3120-150) in GMEMβ/FCS/LIF and incubated in a 37°C/ 7% CO2 incubator. Cells were routinely passaged when they reached 70-80% confluence. Medium was changed every day or every two days depending on the cell line. Flasks were treated with 0.1% gelatin in PBS for 10 min before use. ES cells were washed with pre-warmed PBS and incubated with 0.025% (v/v) trypsin in PBS at 37°C for 2 min or until cells started detaching. ES cells were detached by vigorously tapping the flask and a volume of GMEMβ/FCS/LIF equivalent to 5 times the volume of trypsin was added to block the reaction. Cells were transferred to a universal 30-ml tube and centrifuged for 3 min at 290g. The pellet was resuspended in GMEMβ/FCS/LIF and cells were split 1:5 to1:15 at each passage.

2.1.3: ES cell freezing

1-2 vials were frozen from an 80% confluent T25 cm^2 flask. Before freezing ES cells were collected by trypsinisation as described for cell passaging. Cell pellets were

resuspended in 1-2 ml of GMEMβ/FCS/LIF supplemented with 10% v/v DMSO. 1ml of suspension was then dispensed into each cryotube (NUNC cat. 377224), and tubes transferred immediately to -80°C. The following day tubes were transferred to a $N2_{(lq)}$ tank for long term storage.

2.1.4: ES cell thawing

1 vial of frozen cells was warmed up to 37°C in a waterbath and the ES cell suspension transferred to a universal tube containing 10ml of pre-warmed GMEMβ/FCS/LIF. Cells were collected by spinning at 200g for 3 min and resuspended in 10ml of GMEMβ/FCS/LIF before transfer to a gelatinised T25 flask.

2.1.5: Colony forming assay (in +/- LIF)

ES cells were collected by trypsinisation as described for passaging and resuspended in 5ml of GMEMβ/FCS/LIF. Cells were counted and 600 or 3,600 cells respectively were replated in a gelatinised well of a six well plate (IWAKI cat. 3810-006) or in a 10cm diameter dish (surface area~ 63cm^2) (IWAKI cat. 3020-100). Cells were cultured for 7 days in GMEMβ/FCS/LIF with or without LIF, washed in PBS, and incubated for 1 minute in fixative solution made by mixing 25ml of citrate solution (18mM citric acid, 9mM sodium citrate, 12mM NaCl), 8ml of formaldehyde solution (37% v/v in water) and 65 ml of acetone. Fixed plates were washed in distilled water and stained for alkaline phosphatase (AP) expression using a leukocyte alkaline phosphatase kit (Sigma cat. 86R-1KT).

When required LIF antagonist (hLIF-05) was added to the medium of cells cultured in the absence of LIF. hLIF-05 (Vernallis et al., 1997) was obtained by preparing conditioned medium from COS-7 cells transiently transfected with a hLIF-05 expression plasmid. Briefly, COS-7 cells were grown in a large T150 cm^2 flask (IWAKI cat. 3120-150) in GMEMβ/FCS until confluent. Cells were collected by trypsinisation, counted and 1.5 x 10^6 cells/dish were replated in 10 cm diameter dishes (IWAKI cat. 3020-100) and incubated overnight. The next day, 30 μl of

Fugene-6 transfection reagent (Roche cat. 11815091001) were diluted in 570μl of GMEMβ in a 1.5 ml tube. After 5 min at RT, 18 μg of circular plasmid DNA (AGS 627) were added to the Fugene solution, mixed by flicking the tube and incubated at RT for 15 min. The DNA/Fugene mixture was added dropwise to the COS-7 plates. After 24 hours the medium was replaced and cells cultured for additional 4 days before harvesting and filtering the supernatant. 5,000 CPI ES cells were plated in 6 separate wells of a 24 well plate (IWAKI cat. 3820-024) in ES cell medium containing 10U/ml LIF and incubated for 2 hours. Conditioned medium was added to each well by 1:2 serial dilutions so that a range of 1:5 to 1:320 (v/v) was tested. 4 days later plates were stained with Leishman's stain and the lowest concentration of conditioned medium that caused total inhibition of ES cells self-renewal was determined and used in further experiments.

2.1.7: Neural differentiation assays and laminin coating

5 x 10^4 cells were plated in 6 well plates coated overnight with Poly-L-ornithine 0.01% (Sigma cat. P4957), washed and coated 2 hours with laminin (Millipore cat. CC095) 5μg/ml in PBS. Cells were culture in N2B27 alone (or supplemented with BMP, LIF, or BMP/LIF; made in-house) for 9 days, fixed and stained as described.

2.1.8: Stable transfection of DNA into ES cells

70%-80% confluent ES cells were collected by trypsinisation as described (Chapter 2.1.2), resuspended in 10ml PBS and counted. 10^7 cells were collected by centrifugation for 3 min at 290g and resuspended in 10ml PBS. Cells were re-centrifuged and resuspended in 700μl PBS. The cell suspension was mixed with a solution of 25-50 μg linearised DNA in 100ul of PBS. Cells were transferred to a 0.4 mm gap electroporation cuvette (Biorad cat. 1652088) and electroporation was performed at 0.8 kV and 3 μF using a Gene Pulser machine (Biorad). Cells were collected with a plugged Pasteur pipette, transferred to 10ml of pre-warmed GMEMβ/FCS/LIF and replated at a density of 10^6 cells/10cm diameter dish (IWAKI

cat. 3020-100). Relevant drug selection was started 1 or 2 days after replating for random transgene integration and targeting of specific loci respectively. Cells were cultured until macroscopic colonies appeared, single colonies picked using wide-bore plastic pipette tips loaded with 10μl 0.025% trypsin solution and transferred to single wells of a 96 well plate containing 20μl of 0.025% trypsin solution. After dissociating cell clumps by pipetting, 200μl of GMEMβ/FCS/LIF were added to each well. Cells were subsequently expanded in GMEMβ/FCS/LIF with appropriate selection before freezing and proceeding to further characterisation.

2.1.9: Transient episomal transfection of E14/T ES cells

10^6 supertransfectable E14/T ES cells expressing the Polyoma Large T antigen (Chambers et al., 2003) were replated in a gelatinised well of a 6 well plate (IWAKI cat. 3810-006) 1 hour before transfection. For each transfection 3μl of lipofectamine 2000 (Invitrogen cat. 11668-027) and 3μg of circular plasmid DNA were separately diluted into 250 μl of GMEMβ. After 5 minute incubation at RT the DNA and lipofectamine solutions were combined and mixed by flicking the tubes. After 20 min incubation at RT with occasional mixing the DNA/lipofectamine mixture was added dropwise to each well of E14/T cells and plates returned to the incubator. Antibiotic selection was started 1 day after transfection.

2.2: DNA manipulation

2.2.1: DNA isolation from bacterial cells

A bacterial colony propagating the plasmid of interest was inoculated in LB broth supplemented with the appropriate antibiotics and grown overnight in a shaker (37°C/225rpm). Bacterial cells were collected by centrifugation (5000g, 10 min) and plasmid DNA isolated using a MINIprep or MAXIprep kit (Qiagen cat. 27014, 12162) following the manufacturers instructions. Quantity and purity (A_{260}/A_{280} and

A_{260}/A_{230}) of the DNA were determined using a ND-1000 spectrophotometer (Nanodrop).

2.2.2: Restriction endonuclease digestion

DNA restriction digestions performed for vector construction or diagnostic digestions were performed in a total volume of 20µl on 1-2 µg DNA using an excess of restriction enzymes provided by New England Biolabs or Roche at the temperature and in the buffer suggested by the manufacturer. Digestion performed to linearise circular DNA used in ES cell stable transfection experiments were performed as above with the difference that 25-50µg of DNA were digested in a total volume of 400µl using an excess (typically~50-100 units) of each restriction enzyme.

2.2.3: Polymerase chain reaction (PCR)

PCRs used in vector construction or cloning of coding sequences from ES cell derived cDNA preparations were performed using 1U of the PFX high fidelity polymerase enzyme (Invitrogen cat. 11708-039) , dNTPs mixture at a final concentration of 200 nM (Invitrogen, cat. 10297-018) and synthetic DNA oligonucleotides at a final concentration of 300nM (Integrated DNA Technologies). Reactions were performed in a total volume of 50µl according to the manufacturers instructions in a DNA Engine thermal cycler (Biorad). 100ng or 50ng of DNA template were used for amplifications on plasmid DNA or cDNA respectively. Annealing was performed at temperatures between 59- to 61°C for 15 sec. Extension times of 1 minute/ Kb and 10 sec denaturation times were employed. An initial denaturation cycle of 3 min and a final extension cycle of 5 min were performed.

2.2.4: DNA fragment ligation

After restriction digestion DNA fragments were separated by electrophoresis using 1% agarose (Invitrogen cat. 16500-500 or Seachem cat. 50005) 0.5x TBE (Tris Borate EDTA buffer) gels. DNA was purified from the gel using a Gel Extraction Kit (Qiagen cat. 28704). DNA was eluted in 30 µl of nuclease-free water and concentration determined using a ND-1000 spectrophotometer (Nanodrop). Typically, 50 to 100 ng of linearised vector DNA were ligated to DNA fragment present in 3 fold molar excess. Ligation reactions were performed in a total volume of 10µl for 1 hour at RT using T4 DNA ligase (New England Biolabs cat. M0202L) following the manufacturer instructions. Ligated DNA mixtures were transformed into chemically competent DH5a *E.Coli* (transformation efficiency of ~ 10^7 cfu/µg) by a 30 min incubation on ice followed by 45 sec heat-shock at 42°C. Bacteria were then allowed to recover in LB for 1 hour at 37°C and were plated on LB agar plates in presence of the appropriate antibiotic selection. Plates were incubated overnight at 37°C before colony picking.

2.2.5: Cloning of PCR products

Cloning of blunt PCR products was performed using the Zero Blunt cloning Kit (Invitrogen cat. K280040) following the manufacturer's instructions and transforming DNA into One Shot Chemically Competent *E.Coli.* (Invitrogen cat K280040).

2.2.6: Cloning by homologous recombination (recombineering)

E.Coli cells harbouring a bacterial artificial chromosome (BAC) containing the genomic regions of interest and conferring chloramphenicol resistance were obtained from commercial repositories (BioSource). 10 ng of pRedET plasmid (Genebridges cat. K001) was electroporated into BAC harbouring cells before plating on LB agar in the presence of tetracycline and chloramphenicol and incubating at 30°C for 36-48

hours. The pACYC177 plasmid region including the ampicillin resistance cassette and the origin of replication was amplified by PCR using oligonucleotide primers bearing 60 bp long sequences of homology to the 5' and 3' ends of the region to be subcloned from the BAC. PCR products were purified on column using a PCR Cleanup Kit (Qiagen cat. 28104) and digested with *DpnI* to avoid propagation of the template plasmid in the following steps. Digested DNA was ethanol precipitated with the help of 2 μg of glycogen carrier and resuspended in nuclease free water. Single tetracycline resistant colonies were inoculated into LB, grown at 30°C and electroporated with linear PCR DNA. One hour before electroporation, γβα RedET enzymes were induced by adding L-arabinose to a final concentration of 0.2% to the culture medium. Electroporated cells were plated on LB agar in the presence of ampicillin and tetracycline selection and incubated at 30°C for 36-48 hours. Single clones were picked and expanded at 30°C maintaining ampicillin and tetracycline selection. Occurrence of the correct recombination event was assessed by restriction digestion on purified plasmid DNA. Cassettes including the desired fluorescent protein coding region preceded by a 2a peptide, linked to the hygromycin, blasticidin or puromycin resistance gene by an internal ribosome entry site (IRES) sequence and followed by a bacterial EM7 promoter driven kanamycin resistance gene were amplified by PCR using oligonucleotide primers bearing 60 bp long sequences of homology to the region immediately upstream and downstream of the *Esrrb* gene stop codon. PCR products were *DpnI* treated as before and electroporated into *E.Coli* clones harbouring the correctly recombined plasmids. Colonies were picked and inoculated in LB containing ampicillin and tetracycline, expression of γβα RedET enzymes was induced as described and cells were plated on LB agar plates in the presence of kanamycin and incubated at 37°C overnight to induce loss of the pRedET plasmid. Colonies were picked and occurrence of the correct recombination event was assessed by restriction digestion on purified plasmid DNA. Plasmid DNA was digested with a single cutter enzyme, purified on column and self-ligated before transforming into chemically competent *E.Coli* and replating on LB agar plates in the presence of kanamycin. This procedure ensured resolution of concatenated repetition of modified and unmodified plasmid sequences originated by self-recombination of the pACYC177 plasmid during RedET recombination. Re-transformation also

excluded concomitant presence of modified and unmodified plasmid molecules in the *E.Coli* population. Single colonies were picked and achievement of pure preparations of correctly recombined plasmid verified by restriction digestion.

2.2.7: Recombinase mediated excision of resistance cassettes

DH5α *E.Coli* cells harbouring PACYC177 based targeting vectors generated by recombineering as described above were electroporated with 10 ng of p705-FLP plasmid (Genebridges, K001), plated on LB agar in the presence of ampicillin and chloramphenicol and incubated at 30°C for 36-48 hours. Colonies were picked, inoculated in LB containing ampicillin and grown at 30°C until reaching an $OD_{(600)}$ of 0.2. Cells were then transferred at 40°C for 20 min and grown overnight at 37°C to cause loss of the p705FLP plasmid. Plasmid DNA prepared from overnight cultures was transformed into DH5α and cells plated on LB agar in the presence of ampicillin. This procedure allowed the separation of recombined and unrecombined DNA molecules concomitantly present in the *E.Coli* population. After overnight incubation at 37°C colonies were picked, inoculated in LB in ampicillin selection and grown overnight at 37°C. Plasmid DNA was prepared and successful excision of the EM7-kanamycin resistance gene cassette assessed by restriction digestion.

2.2.8: ES cell genomic DNA isolation

ES cells were grown in gelatinised $25cm^2$ flasks (IWAKI cat. 3100-025) until reaching 70-80% confluence. Cells were collected by trypsinisation, washed once in PBS and centrifuged (290g, 3 min). Genomic DNA was prepared from cell pellets using the DNeasy kit (Qiagen cat. 69504) following the manufacturers instructions and performing the elution step twice with 100ul of nuclease free water pre-warmed at 42C. Quantity and purity (A_{260}/A_{280} and A_{260}/A_{230}) of the DNA were determined using a ND-1000 spectrophotometer (Nanodrop).

2.2.9: Southern Blot analysis

4μg of genomic DNA were digested overnight with the indicated restriction endonucleases. Digested DNA was separated by electrophoresis on a 0.8% w/v agarose gel and the gel incubated for 30 min in a water solution of 1μg/ml ethidium bromide. After crosslinking for 2 min by exposure to 254 nm wavelength UV, the gel was incubated twice for 15 min in a 0.5M NaOH, 1M NaCl solution and washed in a 0.5M Tris, 3M NaCl pH 7.4 solution. DNA was wet transferred to a Hybond XL membrane (General Healthcare cat. RPN303-S) soaked in 2X SSC (20X SSC: 3M NaCl, 0.3M Tri-sodium Citrate) by capillarity driven flow of a 20X SSC solution. The membrane was washed in 2X SSC, dried and baked for 2 hours at 80°C. The membrane was blocked with PerfectHyb solution (Sigma cat. H7033) containing 100μg/ml salmon sperm DNA (Sigma cat. D7656) at 68°C in a roller bottle. 25ng of probe DNA were labeled with P^{32} α-dCTP using a Rediprime II Random Prime Labeling Kit (Amersham cat. RPN1633) and cleaned on column. The membrane was hybridised with the labelled probe overnight at 68°C, rinsed, and washed twice in 0.5 SSC 0.1%SDS for 90 and 30 min respectively at 68°C. The membrane was exposed to Hyperfilm (Amersham cat. 28906837) at -80C for 1-7 days depending on the signal intensity.

2.2.10: Detection of *Esrrb* exon 2 excision by quantitative PCR on genomic DNA

Genomic DNA from $Esrrb^{\Delta/\Delta}$ or $Esrrb^{f/fn}$ cells was isolated using DNeasy minikits (Qiagen cat. 69504) as described (Chapter 2.2.8). Real-time RT–PCR reactions were performed in triplicate in 384-wells plates with a 480 LightCycler (Roche) using the LightCycler 480 Probes Master mix (Roche cat. 04707494001). 100ng of genomic DNA diluted in 5μl of water were used per reaction. For each sample two sets of reactions using primers and Universal Probe Library (UPL) (Roche cat. 04683641001) probes binding to *Esrrb* exon2 or exon 6 were performed. For each sample, *Esrrb* exon 2 copy numbers were normalised against exon 6 copy numbers,

to correct for errors in the quantity of genomic DNA used per reaction. Exon 2/exon 6 ratios are shown relative to untreated *Esrrb*$^{f/fn}$ cells. Standard curves of all primers were performed to check for efficient amplification (above 85%). PCR primer sequences are listed in **Table 2.1**

2.2.11: Genotyping of animals by PCR analysis of DNA from ear biopsies

Mice biopsies were incubated overnight at 56°C in 50 µl of a 0.45% Tween 20, 0.45% NP40, Qiagen Taq Buffer (Qiagen cat. 201207), 0.1 mg/ml proteinase K (Sigma cat. P5568-1ML) solution in water. Samples were spun down and incubated at 95°C for 10-20 min, vortexed and spun down at full speed in a benchtop centrifuge for 1 min. 5µl of supernatant were transferred to a PCR tube. 25 µl of PCR mix containing 1U of Taq polymerase enzyme (Qiagen cat. 201207), dNTPs mixture at a final concentration of 200 nM (Invitrogen cat. 10297-018), synthetic DNA oligonucleotides at a final concentration of 300nM (Integrated DNA Technologies), and Qiagen Taq Buffer (Qiagen cat. 201207) in water were added to each PCR tube. Reactions were performed in a DNA Engine termal cycler (Biorad). PCR primer sequences are listed in **Table 2.1**.

2.3: Gene expression quantification

2.3.1: RNA isolation and quantitative real-time RT-PCR

Total RNA from cultured cells was isolated using the RNeasy microkit or minikit (Qiagen cat. 74004, 74104), and performing on-column digestion with DNase I (Qiagen cat. 79254). Reverse transcription reactions were performed on 30ng-2µg of total RNA in a final volume of 20µL with 100U of SuperScriptIII (Invitrogen cat. 18080-093), 200ng random hexamers (Invitrogen cat. N8080127) or 250ng of oligod$T_{(12\text{-}18)}$ (Invitrogen cat. 18418-012) primers at 42°C for 60 min.

Table 2.1: List of PCR primers.

General qPCR primers		
Primer pair	Fw	Rv
TBP	ggggagctgtgatgtgaagt	ccaggaaataattctggctca
28S rRNA	gaaggcaagatgggtcaca	gaacttccgtgggtgactcc
Oct4	gttggagaaggtggaaccaa	ctccttctgcagggctttc
Sox2	gtgtttgcaaaaagggaaaagt	tctttctcccagccctagtct
Nanog	cctccagcagatgcaagaa	gcttgcacttcatcctttgg
Klf4	cgggaagggagaagacact	gagttcctcacgccaacg
Esrrb	cgattcatgaaatgcctcaa	cctcctcgaactcggtca
Gli2	tgaaggattcctgctcgtg	gaagttttccaggacagaacca
Foxd3	ccccaacactgaccaacag	gtttgctccgccagctta
Inhbb	cgagatcatcagctttgcag	ggttgccttcattagagacga
Inhbb 3'UTR	ggaaagaaaaatgttgcaatcg	gccctcccctctaagcataa
Nfib	gggactaagcccaagagacc	tgggtgtcctatttgacacttg
Foxn4	agcatcatggacttcgctct	ggctgaagctgtcctcctt
Pgc	tgaacttctgcagaccatagga	ctgtcgcagctcacaaaatact
Sorl1	cgcggtgtctctttggtag	accacagctgcaacatcagt
Tet2	gaagagtgcggaaagaatgg	agactggatgaaacgcaggt
Manaba	gtggtctccttttacttttcaactg	tgcccttctgagagatgttga
Zic2	caagatccacaaaagaactcataca	tcttcctgtcgctgctgtt
Dax1	accgtgctctttaacccaga	ccggatgtgctcagtaagg
Igf2bp2	gggaaaatcatggaagttgacta	cgggatgttccgaatctg
Tfp1	taacatcgtggttccccagt	tctgctggttgaagacacca
Utf1	gtccctctccgcgttagc	ggggcaggttcgtcattt
Nr5a2	gcatttgggcttttatgcaa	gcttcatttggtcatcaacct
Jarid1b	gactgggttcaggatgtgga	tgtctctaacactggcacacg
Rcor2	cagagcatgaagcagaccaat	agcgggagttgaacttggta
Mov10	ggaccctgtggaccagaa	ctctctgcgacaccaggaa
Xbp1	tgacgaggttccagaggtg	tgcagaggtgcacatagtctg
Hhip	gtgttcggagatcgcaatg	ttttcttgccattgcttggt
Ncoa1	agcaaggaacaatgggaaac	ttgtacactcctggctgtgc
Itga9	catgctactgaacacagagatactga	cgagccatgaactggatga
Sfrp1	atgtgctccagaagcagacc	gtcagagcagccaacatgc
Tcf4	catatttgtggccattgaagg	gtccctaaggcagccattc
c-kit	tttcctctgggagctcttctc	gaagccttccttgatcatctt
A2m	gaccttgttgttgacaaggactta	acctcattggatgaagactgtg
Moyf	acctctctaaaattgctgcctct	cagctcccgaggatgaga
CD38	aagatgttcaccctggagga	actccaatgtgggcaagaga
ID1	gcgagatcagtgccttgg	ctcctgaagggctggagtc

Table 2.1: List of PCR primers (Continued).

Primer pair	Fw	Rv
ID2	gacagaaccaggcgtcca	agctcagaagggaattcagatg
Klf6	tcccacttgaaagcacatca	acttcttgcaaaacgccact
Aff3	taacctggccaaggaaaaca	gtgacaggtcccatgagca
Dnmt3b	atgatcgatgccatcaaggt	gggaagccgaagatcctg
Dnmt3a	aacggaaacgggatgagtg	actgcaattaccttggctttct
Sox17	ttcctatttccccaagaggtc	gcttctctgccaaggtcaac
Hes6	acggatcaacgagagtcttca	ttctctagcttggcctgcac
GFP	aagttcatctgcaccacc	tccttgaagaagatggtggg
Esrrb pre-mRNA 1	gctaacccaccaggcttct	taagggaactgaggcagcac
Esrrb pre-mRNA 2	ggatcagattcctgcctgtg	tggagctgctcttcagacct
Esrrb pre-mRNA 3	tgaatgtgtgtgatgctcttgt	gcccaacacctggaaactaa
Rex1	cgctgtgggcattaggtaag	gcacactcactctattgagagaagaa
Stella	gatgcacaacgatccagattt	tggaaattagaacgtacatactccaa
T/Bra	cagcccacctactggctcta	gagcctggggtgatggta
Foxa2	gagcagcaacatcaccacag	cgtaggccttgaggtccat
Gata6	ggtctctacagcaagatgaatgg	tggcacaggacagtccaag
Lefty1	actcagtatgtggccctgcta	aacctgcctgccacctct
Fgf5	aaaacctggtgcaccctaga	catcacattcccgaattaagc
TdTomato	tgttcctgtacggcatggac	ggattctcctcgacgtcacc
Olig2	agaccgagccaacaccag	aagctctcgaatgatccttcttt
Esrrb exon-exon and exon-intron junctions qPCR primers		
Primer pair	Fw	Rv
Exon 1a- Exon 2	gctggaacacctgagggtaa	tgaaggagccgcaactagag
Exon 1b- Exon 2	ctgatggacgtgtccgaact	tgaaggagccgcaactagag
5’ UTR - Exon 2	actcccgcgttctgtgttc	tgaaggagccgcaactagag
Exon 2 - Exon 3	ggcgttcttcaagagaacca	ctccgtttggtgatctcacat
Exon 3 - Exon 4	cgattcatgaaatgcctcaa	cctcctcgaactcggtca
Exon 4 - Exon 5	gaacagcccctacctgaacc	tgggaggcatagcatacagc
Exon 5 - Exon 6	atcaaggccctgaccactct	tcatctggtccccaagtgtc
Exon 6 - Exon 7	ggtgcgcaggtacaagaaac	cgcctccaggttctcaatgt
Exon 6 – alt. 3’UTR	ggtgcgcaggtacaagaaac	ctcccacagaacccctgaag
Exon 2 – Intron 2	ggcgttcttcaagagaacca	aacccaactagacctgtagaccaa
Exon 3 – Intron 3	cgattcatgaaatgcctcaa	caccaagacaaaccaagtgga
Exon 5 – Intron 5	atcaaggccctgaccactct	ttagcactcttgccctgacc
Exon 6 – Intron 6	ggtgcgcaggtacaagaaac	ctcccacagaacccctgaag

Table 2.1: List of PCR primers (continued).

Esrrb ChIP qPCR primers		
Primer pair	Fw	Rv
Esrrb 5’	ccctatgtcatcacgccttt	cagttgaaaatggggacctg
Promoter 1(RNAPolII)	tagaactgactcccgcgttc	atgaaggagccgcaactaga
Promoter 2 (RNAPolII)	ctgctgaaccgaatgtcgt	gtggctgagggcatcaat
Promoter 3 (Nanog)	agggtcacatgctgatagcc	ctccactcacctgcacagag
Enhancer	cggctggtatcacctgattt	gctttgtcctctttgccaat
Esrrb 3’	actcctccccttacccctgt	ggctgtggtcactgcatcta
Esrrb genomic DNA qPCR primers		
Primer pair	Fw	Rv
Esrrb Exon 2 - Intron 2	ggcgttcttcaagagaacca	aacccaactagacctgtagaccaa
Esrrb Exon 6 - Intron 6	ggtgcgcaggtacaagaaac	ctcccacagaacccctgaag
Esrrb Exon 7 - 3’ UTR	gtgcccatgcacaaactctt	tctccacttggatcgtgtcc
TdTomato	tgttcctgtacggcatggac	ggattctcctcgacgtcacc
NS x ES fusion qPCR primers		
Primer pair	Fw	Rv
Endogenous Nanog	cggctcacttccttctgact	cgagggaagggatttctga
Endogenous Esrrb	actcccgcgttctgtgttc	tgaaggagccgcaactagag
NS x ES fusion RT-PCR primers		
Primer pair	Fw	Rv
Nanog 5’UTR- Hygro^{R}	cggctcacttccttctgact	cgcggccgatgcaaagtgc

Table 2.1: List of PCR primers (continued).

NSΔN-iNanog and NSΔN-iEsrrb viral reprogramming qPCR primers		
Primer pair	Fw	Rv
Endogenous Nanog	cggctcacttccttctgact	cgagggaagggatttctga
Endogenous Esrrb	actcccgcgttctgtgttc	tgaaggagccgcaactagag
Endogenous Oct4	ttccaccaggccccc	ggtgagaaggcgaagtctgaag
Viral Oct4	tcccagtgtggtggtacggg	ggtgagaaggcgaagtctgaag
Viral Klf4	tcccagtgtggtggtacggg	gagcagagcgtcgctgacag
Viral c-Myc	tcccagtgtggtggtacggg	ggtcatagttcctgttggtgaagtt

Real-time RT–PCR reactions were performed in triplicate in 384-wells plates with a 480 LightCycler (Roche) using LightCycler 480 SYBR Green I Master (Roche cat. 04707516001). Five microliters of cDNA or immunoprecipitated chromatin were used per reaction. Standard curves of all primers were performed to check for efficient amplification (above 85%), and all melting curves were generated to verify production of single DNA species with each primer pair. PCR primer sequences are listed in **Table 2.1**. Values for each gene were normalised to expression of TATAbox Binding Protein (TBP) or 28S rRNA (For flavopiridol treatment experiments).

2.3.2: ESΔN-NERT timecourse microarray analysis

1 x 10^6 ESΔN-NERTc3 cells were plated in separate 25cm^2 flasks (IWAKI cat. 3100-025) one day before stimulation. Cells were incubated in GMEMβ/FCS/LIF containing 1μM Tamoxifen for the indicated time and harvested by trypsinisation. RNA was prepared using a RNeasy minikit (Qiagen cat. 74104). 100ng of RNA were reverse transcribed into double stranded cDNA and transcribed/amplified into biotin labelled cRNA using an Illumina TotalPrep RNA Amplification Kit (Ambion cat. AMIL1791). Labelled RNA was submitted to the WTCRF MRC Human Genetics Unit (University of Edinburgh) for further processing. cRNA quality was checked using a Agilent 2100 Bioanalyser and hybridization performed on an MouseWG-6 v2 BeadChip (Illumina cat. BD-201-0202). Raw data was processed in *R* using the *beadarray* (Dunning et al., 2007) and *limma* (Smyth, 2005) packages from the *Bioconductor* suite (Gentleman et al., 2004) Briefly, low-quality probes were removed from the input data and sample-effects normalized by centring individual replicate groups around the 0-hour baseline. The data was subsequently quantile-normalized and log2-transformed before assessing differential expression with the *limma* algorithms. Genes were considered differentially expressed if they showed a FDR-adjusted p-value of at most 0.05 and a fold change of 1.5 or more for at least one time point in comparison to the 0-hour baseline.

2.3.3: DeepSAGE Library Preparing and Sequencing

RNA was submitted for sample preparation and sequencing following the manufacturer's protocol for tag profiling using the *NlaIII* restriction enzyme at the GenePool core facilities of the University of Edinburgh. The RCN(t) (1 library) and RCNβH(t) (2 libraries) samples were sequenced on a first generation Illumina/Solexa Genome Analyzer and the sorted TNG samples were sequenced at a later time point on a Illumina GAII.

2.3.4: Short Read Data Processing

Data analysis of short read sequencing libraries was performed using GeneProf (Halbritter et al., 2012). Briefly, reads were trimmed to the significant 17nt portion corresponding to the tag sequence and extended by the known recognition sequence of the *NlaIII* digestion enzyme (CATG). The libraries were then filtered removing low-quality reads and aligned to the Mouse reference genome (NCBIM37 assembly) using the Bowtie alignment tool (Langmead et al., 2009). Aligned reads were compared to known gene models (GeneProf reference dataset based on the Ensembl 58 database (Flicek et al., 2011)) to calculate read counts per gene, which were then quantile-normalized. We considered as differentially expressed those genes that exhibited a 1.5-fold change between groups and had a minimum (normalized) expression level of at least 10 in at least one sample.

2.3.5: Nanog Target Prioritization

We sought to identify likely *Nanog* candidate genes from the DeepSAGE data by defining an *ad hoc* ranking criterion called the *Nanog Candidate Priority Factor (NCPF)*. Given $E_x(g)$ as the normalized expression intensity of gene g in sample x and

$$fc_1(g) = \log_2\left(\left(E_{RCNbetaH(t).1}(g) + E_{RCNbetaH(t).2}(g)\right)/2 + 1\right) - \log_2\left(E_{RCN(t)}(g) + 1\right),$$

$$fc_2(g) = \log_2(E_{TNG.low}(g)+1) - \log_2(E_{TNG.high}(g)+1),$$

$$avgE(g) = \left(E_{RCNbetaH(t).1}(g) + E_{RCNbetaH(t).2}(g) + E_{RCN(t)}(g) + E_{TNG.low}(g) + E_{TNG.high}(g)\right)/5$$

we define the NCPF as:

$$NCPF(g) = (fc_1(g) + fc_2(g))^2 * sgn(fc_1(g) + fc_2(g)) * avgE(g).$$

2.4: Chromatin immunoprecipitation

2.4.1: Chromatin preparation

2 x 10^7 NERT ES cells were resuspended in 3ml of pre-warmed GMEMβ/FCS/LIF and cross-linked for 10 min at RT with 1% formaldehyde (Sigma cat. F8775-25ML). The cross-linking reaction was stopped by adding 0.125mM glycine for 5 min at RT. Cells were pelleted (3 min, 300rpm, 4°C) and washed twice with cold PBS. Cell pellets were vigorously resuspended in 300µl of swelling buffer (5mM Pipes pH8, 85mM KCl) freshly supplemented with 1X protease inhibitor cocktail (Roche cat. 04 693 116 001) and 0.5% NP-40. After 20 min on ice with occasional shaking, nuclei were centrifuged (1500rpm, 10 min, 4°C) in 15ml conical tubes and resuspended in 1.5ml of TSE150 (0.1% SDS, 1% Triton, 2mM EDTA, 20mM Tris-HCl pH8, 150mM NaCl) buffer, freshly supplemented with 1X protease inhibitor cocktail. Samples were sonicated at 4°C in 15ml conical tubes using a Bioruptor (Diagenode) for 5 cycles of 10 min divided into 30 sec ON-30 sec OFF subcycles at maximum power. The chromatin was then transferred into 1.5ml tubes and microcentrifuged (30 min, 14000rpm, 4°C). Soluble chromatin was divided into 250ul aliquots and stored at -80°C until use. Twenty microlitres were set apart and used to quantify the chromatin concentration and the size of the DNA fragments on a 1.5% w/v TAE agarose gel.

2.4.2: Chromatin ImmunoPrecipitation (ChIP)

For each experiment, the required amount of chromatin was thawed and pre-cleared for 90 min rotating on-wheel at 4°C in 1 ml of TSE150 containing 50µl of pA/pG sepharose beads (Sigma cat. P9424-5ML, P3296-5ML) 50% slurry, previously blocked with 500µg/ml of molecular grade BSA (Roche, cat. 5931665103) and 1µg/ml of yeast tRNA (Invitrogen, cat. AM7119). Pre-cleared chromatin was centrifuged (3000rpm, 1 min) in a benchtop microcentrifuge and the supernatant transferred into fresh tubes so that 20µg of DNA were used per ChIP. In addition, 20ug of diluted chromatin was set apart for input DNA extraction and precipitation. Immunoprecipitation with in-house made anti-Nanog rabbit polyclonal (2 µg /ml) of or anti-RNAPII (Euromedex cat. PB-7C2) (2 µg /ml) antibodies was performed overnight rotating on-wheel at 4°C in a final volume of 500µl of TSE150. Immunocomplexes were recovered with 50µl of blocked pA/pG sepharose beads 50% slurry for 1h and 30 min rotating on wheel at 4°C. Beads were recovered by centrifugation for 1 minute at 3000rpm and washed 5 min rotating on-wheel at RT with 1ml of buffer in the following order: TSE150, TSE500 (0.1% SDS, 1% Triton, 2mM EDTA, 20mM Tris-HCl pH8, 500mM NaCl), washing buffer (10mM Tris-HCl pH8, 0.25M LiCl, 0.5% NP40, 0.5% Na-Deoxycholate, 1mM EDTA), and twice in TE (10mM Tris-HCl pH8, 1mM EDTA). After the last wash, elution was performed in 100µl of elution buffer (1% SDS, 10mM EDTA, 50mM Tris-HCl pH8) for 15 min at 65°C after vigorous vortexing. Eluates were collected after 1 minute centrifugation at 14000rpm in a benchtop microcentrifuge, and the beads rinsed in 150µl of TE-SDS1%. After 1 minute centrifugation at 14000rpm, the supernatant was pooled with the corresponding first eluate. For both immunoprecipitated and input chromatin samples, the crosslinking was reversed by incubation overnight at 65°C, followed by proteinase K treatment (Invitrogen cat. AM2546), phenol/chlorophorm extraction and ethanol precipitation. DNA pellets were resuspended in 200µl of water. Real-time quantitative PCR reactions were performed as described. PCR primer sequences are listed in **Table 2.1**.

2.5: Kidney capsule grafts, recovery and processing

Engraftments were performed as described elsewhere (Tam, 1990). 4 weeks after the graft was performed, the mice were sacrificed and the kidneys removed in PBS. After imaging, the tumour and kidney were fixed in 4% paraformaldehyde (PFA) between 1-7 days depending on tumour size. After fixation, the kidneys were dehydrated through ethanol series (80% EtOH, 2 hours, 4°C; 90% EtOH, 2 hours, 4°C; 96% EtOH, 2 hours, 4°C; 100% EtOH, 2 hours, 4°C), cleared in xylene (15-60 min, 4°C) and embedded in paraffin wax (3x changes: overnight, 2 hours, 2 hours) before being sectioned in a microtome.

2.6: Immunohistochemistry

2.6.1: Immunostaining on cultured ES cells

Cultured cells were fixed in 4% PFA for 10 min at RT. Permeabilisation was done with a solution of 0.1% v/v TritonX100 in PBS (permeabilisation buffer) for 15 min at RT. Blocking was performed for 30 min at RT in permeabilisation buffer supplemented with 3% serum of the same species the secondary antibodies were raised in (blocking buffer). Primary antibodies were diluted in blocking buffer to the working concentrations indicated below and applied for 1-2 hours at RT or overnight at 4 °C. After three washes in permeabilisation buffer, secondary antibodies were diluted to 2μg/ml in blocking buffer and applied for 1 hour at RT. The cells were washed at least three times in permeabilisation buffer and in selected cases nuclei were stained with 4',6-diamidino-2-phenylindole (DAPI).

Primary antibodies were used at the following concentrations: Nanog (Abcam, cat. ab14959), 2.5ug/ml; in-house made anti-Nanog rabbit polyclonal, 1.2-0.25 μg/ml (Chambers et al., 2007); Oct4 (Santa Cruz Biotechnology, cat. sc-5279), 1 μg/ml; Sox2 (Santa Cruz Biotechnology, cat. sc 17320), 1μg/ml; Esrrb (Persaeus Proteomics, PP-H6705-00), 1 μg/ml; TuJ (Covant, MMS-435P), 1 μg/ml.

2.6.2: Whole-mount immunostaining of mouse gonads

Gonads were isolated from dissected day E12.5 or E13.5 embryos and briefly washed in PBS containing 3mg/ml polyvinylpyrrolidone (PBS/PVP) before fixation in a solution of 2.5% paraformaldehyde (PFA) in PBS for 20 min at RT. Fixed gonads were permeabilised for 45 min at RT in 0.3% Triton X-100 PBS/PVP, blocked for 2 hours at RT in a solution of 0.1%BSA, 0.01% Tween 20, 10% goat serum in PBS (blocking buffer). Gonads were then incubated overnight at 4°C with anti-Oct4 C10 monoclonal antibody (Santa Cruz Biotechnology, cat. sc-5279) and rabbit anti mouse Vasa homolog (mvh) (Abcam, cat. Ab13840) diluted 1:200 in blocking buffer, rinsed three times for 15 min in blocking buffer and incubated (2 hours, RT) with goat anti-mouse IgG (H+L) Alexa fluor-568 (Invitrogen, Molecular Probes cat. A-11004) and donkey anti-rabbit IgG (H+L) Alexa fluor-405 (Invitrogen, Molecular Probes cat. A-31556) antibodies diluted 1:500 in blocking buffer. After washing three times for 15 min in blocking buffer, gonads were mounted on glass microscope slides in Vectshiled (Vector Labs cat. H-1000) without DAPI and imaged using a Leica DM IRE2 inverted confocal microscope (Leica Microsystems).

2.7: Immunoblotting

ESΔN-NERT induction: 1.3 x 10^6 ESΔN-NERT cells were plated in separate 25cm^2 flasks (IWAKI cat. 3100-025) two days before lysis. One day before lysis cell were stimulated with or without 1μM tamoxifen. Protein Stability Assays: 10^6 E14Tg2a ES cells were plated in individual wells of a 6-wells plate and 24h later treated with 30mg/ml Cycloheximide for 1 to 6h. For all experiments: Cell were lysed in 300 μl of a 0.5% NP-40, 50mM Tris pH 8,150mM NaCl solution with protease inhibitors (Roche cat. 04 693 116 001). Protein extract treated with 2 μl DNA nuclease Benzonase (Novagen, cat. 70664-3) 1 h at 4°C. 50 μg of protein were denatured in Laemmli buffer at 100°C for 5 min and fractionated on a NuPage-Novex 10% Bis-Tris gels (Invitrogen, cat. NP0302). Proteins were electro-blotted to a nitrocellulose membrane in a 25mM Tris, 0.21 M glycine, 20% methanol solution. The membrane

was blocked in PBS 0.01% Tween (PBST), 10% non-fat dry milk for 2 h and incubated in 5ml of PBST 5% milk containing anti Esrrb (Persaeus Proteomics cat. PP-H6707-00)(1 μg/ml) or anti-Nanog rabbit polyclonal (Chambers, 2005) (2.4 μg /ml) antibodies or anti HDAC-2 (Upstate cat. 05-814)(1 μg /ml) (overnight, 4°C or 2h, RT). Membranes were washed 3 times for 20 min in PBST and incubated with ECL anti-mouse IgG (Amersham cat. NA931) or ECL anti-rabbit IgG (Amersham cat. NA934) HRP conjugated secondary antibodies. Membranes were washed in PBST and developed using a Super-signal West Pico kit (Pierce cat. 34080) for 5 min at RT and exposed to Hyperfilm (Amersham cat. 28906837).

2.8: Imaging

Images were captured using Volocity (Improvision) software in an Olympus IX51 (for cultured cells), or an Olympus BX61 (for teratocarcinoma sections). Image processing was performed using Adobe Photoshop software. ESΔN-NERT cells stained for Nanog were imaged using a Leica DM IRE2 inverted confocal microscope (Leica Microsystems).

2.9: Cell line derivation

2.9.1: Doxycycline inducible expression

E14Tg2a or TβC44Cre6 cells were stably transfected with linearised pPyCAG-rtTA-IRES-BSD^R plasmid. After 10 days of hygromycin selection newly formed colonies were trypsinised and expanded as a mixed population. Cells were then electroporated with linear TetO-TdTomato-2a-Hyg^R-TK plasmid. Clones were screened for high, homogeneous TdTomato expression in doxycycline without continued hygromycin selection and low levels of TdTomato in the absence of doxycycline. An identified cell line was used for FlpE catalysed recombinase mediated cassette exchange

(RCME). 10^6 cells were concomitantly lipofected with 3μg of pShuttle-Nanog-IRES-Puro^R or pShuttle-FlagEsrrb-IRES-Puro^R and 3ug of pPGK-FLPO. Two days after lipofection 10^5 cells were plated in gelatinised T150flasks (IWAKI cat. 3120-150) in in GMEMβ/FCS/LIF containing puromycin. After two weeks colonies were picked and separately expanded before freezing. Successful TdTomato excision was confirmed by flow cytometry in induced cells.

2.9.2: Esrrb fluorescent reporter cell lines

Esrrb-TdTomato reporters: ES cells (TNG, ESΔN-NERT c3 or E14Tg2a) were electroporated with p3'Esrrb-2a-Tdtomato-dest1-IRES-BSD^R or p3'Esrrb-2aX-Tdtomato-IRES-BSD^R vectors, plated at clonal density in GMEM□/FCS/LIF the presence of blasticidin, and after 2 weeks colonies were picked and expanded. Correctly targeted TNG or ESΔN-NERT c3 cells were further electroporated with p3'Esrrb-2a-Tdtomato-dest1-IRES-Hyg^R or p3'Esrrb-2aX-Tdtomato-IRES-Hyg^R vectors to obtain reporter lines harbouring two targeted Esrrb alleles.
Esrrb-GFPdest1 reporter lines: E14Tg2a ES cells were successively targeted with p3'Esrrb-2a-GFP-dest1-IRES-BSD^R and p3'Esrrb-2a-GFP-dest1-IRES-Hyg^R vectors as above.

2.9.3: Derivation of NS cell lines

RCNβH(t) NS cells were derived from the RCNβH(t) ES cell line by neural differentiation in monolayer, as described elsewhere (Pollard et al., 2006). NS cells were propagated in NS cell expansion medium in the presence of FGF and EGF (Conti et al., 2005).

2.9.4: Derivation of *Esrrb*$^{\Delta/\Delta}$ ES cells

Esrrb$^{f/fn}$ ES cells maintained in zeocin to ensure Cre-ER^{T2} expression, were passaged

once in the presence of 1μM PD 0325901 (Axon Medchem, cat. 1408). Cells were replated at clonal density in 10 cm diameter dishes and after 16 hours 1 μM tamoxifen (Sigma, cat. H7904-5MG) was added for 1 hour. Plates were then washed with PBS and cells cultured for additional 2 weeks in GMEMβ/FCS/LIF supplemented with 1μM PD 0325901. Morphologically distinguishable *Esrrb*$^{\Delta/\Delta}$ colonies were picked and expanded in the presence of PD 0325901. After expansion, *Esrrb*$^{\Delta/\Delta}$ cells were further cultured in GMEMβ/FCS/LIF.

2.10: Flow cytometry

2.10.1: Quantification of SSEA-1 expression

4.5 x 10^5 cells were plated in 150 cm^2 flasks (IWAKI cat. 3120-150) and cultured in the absence of any selection for 3 days. Cells were then harvested by trypsinisation, and resuspended at 2 x 10^6 cells/ml in a solution of 10%FCS in PBS (PBS/FCS) containing anti SSEA-1 mouse monoclonal antibody from ascitic fluids (DSHB cat. MC-480) diluted 1:1000. Cells were incubated for 15 min on ice, washed in ice-cold PBS, centrifuged (290rpm, 3 min) and resuspended in PBS/FCS containing goat anti-mouse IgM-Alexa647 secondary antibody (Molecular Probes cat. A-21238). After 15 min on ice, cells were washed in ice-cold PBS, centrifuged as before and resuspended in PBS 10%FCS and analysed using a LSR II flow-cytometer system (Becton, Dickinson).

2.10.2: FACS based cell sorting of fluorescent reporter lines

ES cells plated and stained for SSEA-1 as described (Chapter 2.10.1) were purified using a FacsARIA cell sorter (Becton, Dickinson). After sorting, cell purity was determined using the same instrument. For timecourse culture experiments on sorted populations, cell purity was confirmed with LSR II Fortessa flow-cytometer system (Becton, Dickinson) and cells were replated in GMEMβ/FCS/LIF in separate wells

of 24 well plates and cultured for the indicated time. Cells were harvested every day, stained for SSEA-1 and analysed on a LSRII Fortessa flow-cytometer. Sorted $Esrrb^{+}/Nanog^{+}$ or $Esrrb^{+}/Nanog^{-}$ TNG Esrrb-TdTomato cells were trypsinised and replated at day 1, 2 and 3 for analysis at day 4, 5 and 6 respectively to prevent overgrowth. Data was analysed using the FlowJo software suite (Tree Star). For clonal density plating 600 sorted cells were replated in duplicate in 6 well plates. Cells were cultured in the indicated conditions for six days prior to colony scoring. For cultures in N2B27/2i/LIF, cells were replated in 6 well plates coated overnight with Poly-L-ornithine 0.01% (Sigma cat. P4957), washed and coated 2 hours with laminin (Millipore, cat. CC095) 5 μg/ml in PBS.

2.10.3 Quantification of transcription factor expression in single cells

10^6 cells were resuspended and fixed in 2 ml of 0.25% PFA/PBS (1h, RT) on a tube roller. After 1 wash in PBS (5ml), cells were resuspended in 2 ml of 70% v/v methanol/PBS pre-cooled to 4°C and incubated for 1 hour at 4°C on a rotating wheel. After centrifugation, cells were resuspended in 1ml of PBS containing 5% donkey serum. 100µl (10^5 cells) of cell suspension were transferred into single wells of a 96 well/V-bottom microtitre plate. Cells were centrifuged (450g, 3 min) and resuspended in 100µl of staining buffer (1% BSA, 3mg/ml PVP, 0.1% TritonX-100, and 10% donkey serum in PBS) pre-cooled to 4°C. After incubating (30 min, RT) cells were centrifuged and resuspended in 100µl of pre-cooled staining buffer containing goat polyclonal anti-Oct4 (Santa Cruz Biotechnology, cat. SC-8628) and mouse monoclonal anti-Klf4 (Abcam, cat. Ab75486) antibodies at final concentrations of 1.25µg/ml and 10µg/ml, respectively. Plates were incubated overnight at 4°C and then washed (3x, 20 min, 4°C) in 200µl of pre-cooled staining buffer. After centrifugation, cells were resuspended in 100µl of pre-cooled staining buffer containing donkey anti-mouse IgG (H+L) Alexa fluor-647 (Invitrogen, Molecular Probes cat. A-31571) and donkey anti-goat IgG (H+L) Alexa fluor-568 (Invitrogen, Molecular Probes cat. A-11057) antibodies at a final concentration of 4µg/ml. Cells were incubated (2 hours, RT) in the dark and washed (3x, 20 min, 4°C) in 200µl of pre-cooled staining buffer. Cells were centrifuged, resuspended in 200µl

of PBS containing 5% donkey serum and analysed using a LSR II Fortessa flow-cytometer system (Becton, Dickinson). Data was analysed using the FlowJo software suite (Tree Star). To set the appropriate gates for Klf4, Oct4 and Nanog:GFP negative cells, RA-differentiated E14Tg2a cells were stained in parallel. For all analyses of Klf4 and *Nanog*:GFP heterogeneity in E14Tg2a ES cells, only $Oct4^+$ cells representing the undifferentiated population were taken into account.

2.11: Flavopiridol treatment

For analysis of transcriptional speed, 1.6 x 10^6 EF4 ES cells were plated in separate 10 cm^2 dishes (BD-Falcon, cat. 35-3001) on the day before analysis. Cells were treated for 3 hours with 1μM Flavopiridol (Sigma cat. F3055-1MG) in GMEMβ/FCS/LIF. Cells treated for 3 hours with equivalent amounts of DMSO were used as controls. Cells were quickly washed in warm PBS and cultured for the indicated time in medium without flavopiridol before transfer to ice, washing with ice-cold PBS, and direct lysis on plate with RLT (Qiagen cat. 74104) buffer, followed by RNA purification using a RNeasy minikit (Qiagen cat. 74104). cDNA was prepared as described using 2ug RNA per sample and qPCR reactions performed as described above using primers listed in **Table 2.1**.

2.12: PEG mediated ES x NS cell fusion experiments

For all fusions, 4x10^6 ES cells were fused to 4x10^6 NS cells as previously described (Silva et al., 2006) and plated in GMEMβ/FCS/LIF. For fusions of E14/T NS cells to E14Tg2a, EF4 or EfEsrrb ES cells, puromycin and G418 selections were applied one day after fusion and cell hybrids cultured for 14 days prior to colony scoring. For fusions of RCNβH(t) NS cells to Tβc44c6 or ESΔN-CAGE ES cells, puromycin and hygromycin selection were applied one day and four days after fusion respectively. Cell hybrids were cultured for 14 days prior to colony scoring. For fusions of RCNβH(t) Red NS cells to ESΔN-iNanog or ESΔN-iEsrrb ES cells, blasticidin and hygromycin selection were applied one day after fusion and cell hybrids were

cultured for additional 15 days prior to colony scoring. Doxycycline was added only to the required plates and kept for the duration of the experiment. For plating of sorted hybrids from fusions of RCNβH(t) Red NS cells to ESΔN-iNanog Blue or ESΔN-iEsrrb Blue ES cells 48 h after fusion, cells were trypsinised and stained with 7-AAD (Molecular Probes cat. A1310). 7-AAD negative hybrids (live cells) were purified based on the concomitant expression of EBFP and TdTomato using a FACS-ARIA cell sorter (Becton, Dickinson). After sorting, cell purity was determined using the same FACS-ARIA cell sorter or a LSRII Fortessa flow-cytometer (Becton, Dickinson) and data analysed using the FlowJo software suite (Tree Star). Purified hybrids were replated in GMEMβ/FCS/LIF. Blasticidin and hygromycin selections were applied one day after sorting and cell hybrids were cultured for additional 15 days prior to colony scoring. Doxycycline was added to the culture medium for the duration of the experiment.

2.13: NS cells retroviral transduction for generation of pre-iPS cells

For virus preparation, 10^6 Plat-E cells were seeded into 10mm diameter dishes in GMEMβ/FCS/LIF without LIF on the day before transfection. Separate dishes of Plat-E cells were transfected with 3 μg of either pmX-Oct4, pmX-Klf4, pmX-cMyc or pmX-dsRed (Takahashi and Yamanaka, 2006) using 27μl of Fugene6 (Promega cat. E2693). 24 hours after transfection, medium was replaced and supernatant collected after additional 24 hours. Freshly harvested supernatants were filtered through a 0.45 μm cellulose acetate filter, pooled and supplemented with 4 μg/ml polybrene (Sigma cat. 107689). 2ml of pooled viral supernatants were added to NS cells replated at a density of 10^5 cells per $10cm^2$ dish the previous day. 24 hours after infection viral supernatants were removed and NS cell cultured for 2 additional days in NS cell expansion medium. Infected NS cells were then transferred to GMEMβ/FCS/LIF. When pre-iPS colonies started emerging, cells were replated on STO feeders in GMEMβ/FCS/LIF and expanded.

2.14: Doxycycline timecourse analysis of pre-iPS to iPS conversion

10^5 pre-iPS cells were plated on STO feeders in $10cm^2$ dishes and cultured in GMEMβ/FCS/LIF in the presence or absence of 1 μg/ml doxycycline. After 2 days (timecourse day 0) cells were transferred to GMEMβ/FCS/LIF containing either 1μg/ml doxycycline, 3μM 5-azacytidine, 1μg/ml doxycycline plus 3μM 5-azacytidine or no drugs. Pre-iPS cells were collected every day for 6 days using trypsin and replated in ungelatinised 10mm diameter dishes to allow STO feeders attachment. After 1 hour unattached cells were collected and DsRed and GFP expression quantified by flow cytometry. DAPI staining was used to exclude dead cells from the analysis.

Chapter 3:Identification of Nanog transcriptional targets

In ES cells, Nanog expression levels correlate with self-renewal efficiency (Chambers et al., 2003; Chambers et al., 2007). These functional differences among single ES cells are likely the consequence of differential expression of Nanog target genes. Such genes were therefore identified using a combination of different transcriptional profiling strategies and their contribution to Nanog function was investigated.

3.1: Gene expression profiling of a series of Nanog mutant ES cell lines

Recent publications identified a list of potential Nanog target genes analysing the genome-wide transcriptional alterations that occur in ES cells after Nanog is downregulated by RNA interference (Ivanova et al., 2006; Loh et al., 2006). Despite providing useful indications, studies based on gene knockdown present a number of limitations. In all these studies RNA interference led to the presence of a significant proportion of differentiated cells in the populations analysed. It is therefore difficult to distinguish whether the observed transcriptional alterations are due to variation in Nanog levels or are a consequence of cellular differentiation. In addition, cellular stress and perturbation of the RNA degradation machinery is likely to impact on the observed expression profiles. Similarly, these results do not provide solid ground to investigate the transcriptional differences resulting from the dynamic fluctuations in Nanog expression continuously occurring in ES cells.

To overcome these limitations, this study compared the transcriptional profiles of ES cells expressing unaltered levels of Nanog with profiles generated from Nanog null ES cells. (Respectively RCN(t) and RCNβH(t) ES cell lines (Chambers et al., 2003; Chambers et al., 2007)). In addition, a reporter line, in which a GFP cassette has been knocked-in at one of the Nanog alleles, was used to compare gene expression in

sorted ES cells populations showing high or low levels of Nanog transcription (TNG (Chambers et al., 2007)). By analysing this dataset, it was possible to identify a list of genes the transcript levels of which correlate with Nanog expression in ES cells **(Table 3.1)**. Imposing a 1.5 fold change threshold, differentially expressed genes could be divided in four different groups:

A) 661 genes that that show higher expression in RCN(t) compared to RCNβH(t) and in TNG^{high} compared to TNG^{low} sorted populations.

B) 1002 genes that that show reduced expression in RCN(t) compared to RCNβH(t) and in TNG^{high} compared to TNG^{low} sorted populations.

C) 1948 genes that show higher expression in the TNG^{high} compared to TNG^{low} populations but no correlation with Nanog levels in other cell lines.

D) 1839 genes that show reduced expression in the TNG^{high} compared to the TNG^{low} populations but no correlation with Nanog levels in other cell lines.

Genes belonging to group A and B should be respectively induced and repressed by Nanog. Genes included in group C and D should be enriched for positive and negative regulators of Nanog expression. Since the aim of this project is to identify genes that are transcriptionally regulated by Nanog, subsequent analysis focussed on genes belonging to the first two groups. Differentially expressed genes were ranked according to a Nanog correlation coefficient that weighted the combined fold changes in transcript levels in both RCN(t) compared to RCNβH(t) cells and in TNG^{high} compared to TNG^{low} populations against absolute gene expression levels (see Chapter 2.3.5).

This approach permitted to exclude from the analysis genes that showed high relative fold changes as a consequence of their marginal absolute levels of expression. Of 500 top ranking upregulated or downregulated genes, *Esrrb* was the transcription factor that showed the closest positive correlations with Nanog and consistent

variations in both RCN(t) vs. RCNβH(t) cells and in TNG^{high} vs. TNG^{low} populations (fold change≥1.5, Group A), closely followed by Klf4 (**Figure 3.1**). Amongst downregulated transcription, factors ID1 and ID2 showed the greatest response and consistent variations in both datasets. It is interesting to note that these genes have been reported to be crucial in regulating ES cell self-renewal, stressing the possibility that other less characterised targets identified by this analysis might also be important in this process.

The expression level of 69 genes belonging to group A or B was validated by quantitative PCR in cDNA samples prepared from two different cell lines overexpressing Nanog (RCN, EF4), two ES cell lines with wild-type levels of Nanog expression (E14Tg2a, RCN(t)), two cell lines in which one of the two Nanog alleles has been deleted by homologous recombination (TβC44,RCNβ(t)) and two Nanog null cell lines (TβC44c6, RCNβH(t)) (Chambers et al., 2003; Chambers et al., 2007). Using multiple cell lines should reduce the possibility that the observed variations in expression are due to differences between the cell lines other than the levels of Nanog. In addition, cells were replated at the same density on the day before RNA extraction to minimise the differences resulting from variations in cell-cell interactions or in the concentration of factors secreted in the culture medium. The results from three independent experiments are shown in **Figure 3.2**. 39 out of 69 genes (57%) analysed showed good accordance with the results from high-throughput sequencing.

3.2: Transcriptional dynamics following Nanog nuclear relocalisation in ESΔN-NERT cells

The existence of a direct relation between Nanog transcriptional activity and the expression levels of a subset of validated target genes was then assessed. With this aim, Nanog was linked to a mutated form of the human estrogen receptor ligand binding domain (ER^{T2}) generating a tamoxifen responsive Nanog-ER^{T2} fusion

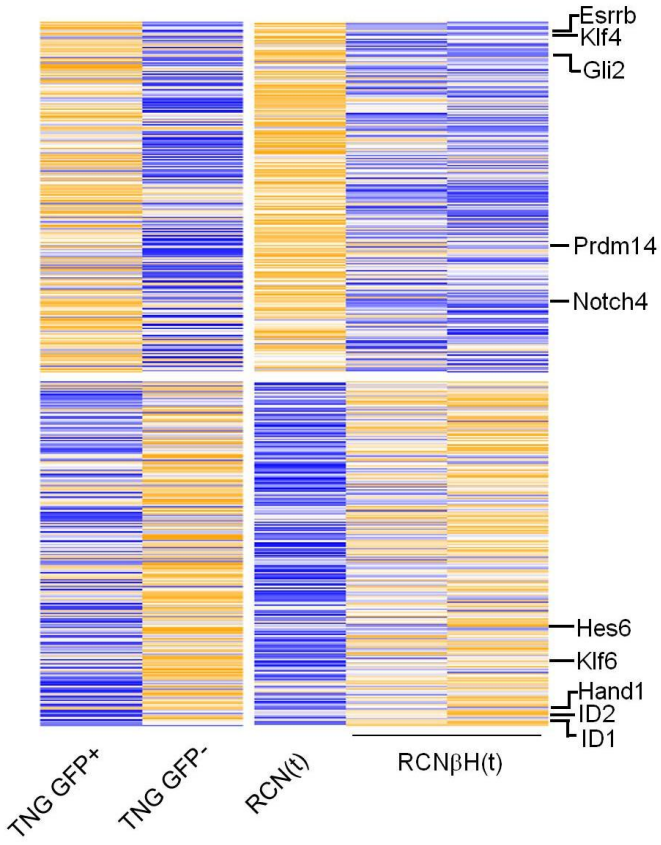
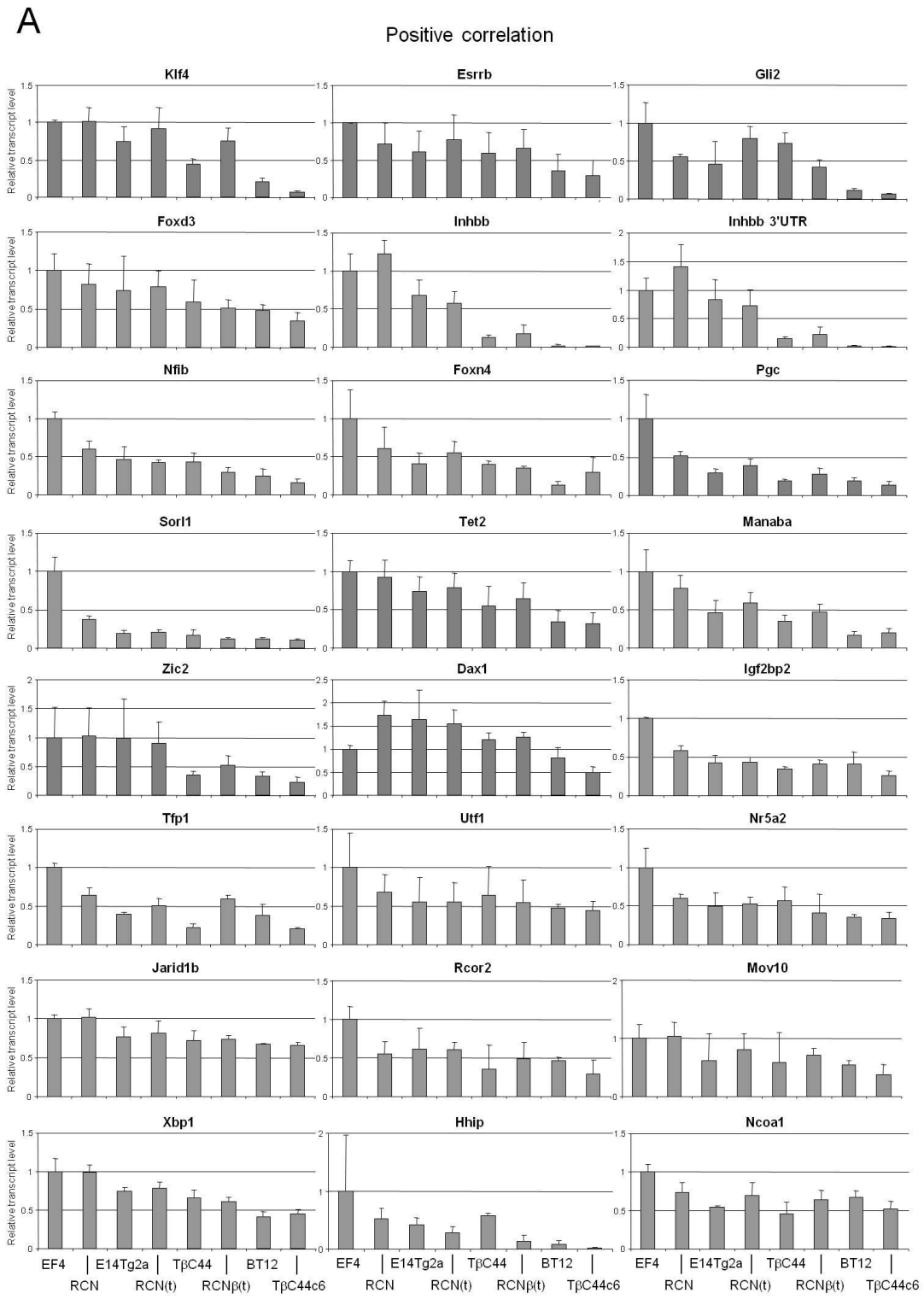


Figure 3.1: Transcriptional changes in response to variations in Nanog levels

Deep-SAGE profile of sorted Nanog positive (GFP+) and Nanog negative (GFP-) TNG cells, ES cells with wild-type levels of Nanog expression (RCN(t)) and *Nanog*$^{-/-}$ ES cells (RCNβH(t)). Genes were ranked according to the expression level and fold difference in expression in TNG+ versus TNG- and RCN(t) versus RCNβH(t); the plot shows the first 250 most upregulated (top) or downregulated (bottom) genes. Names of the top 5 upregulated (group A) or downregulated (Group B) transcription factors are shown on the right. Colours; expression above (yellow) and below average (blue).

The raw data used to produce Figure 1 was kindly provided by Violetta Karwacki-Neisius. All bioinformatic analysis was performed in collaboration with Florian Halbritter using the Geneprof software suite (Halbritter et al., 2012).


A
Positive correlation
Klf4
Esrrb
Gli2
Foxd3
Inhbb
Inhbb 3'UTR
Nfib
Foxn4
Pgc
Sorl1
Tet2
Manaba
Zic2
Dax1
Igf2bp2
Tfp1
Utf1
Nr5a2
Jarid1b
Rcor2
Mov10
Xbp1
Hhip
Ncoa1
Relative transcript level
EF4
E14Tg2a
TβC44
BT12
RCN
RCN(t)
RCNβ(t)
TβC44c6

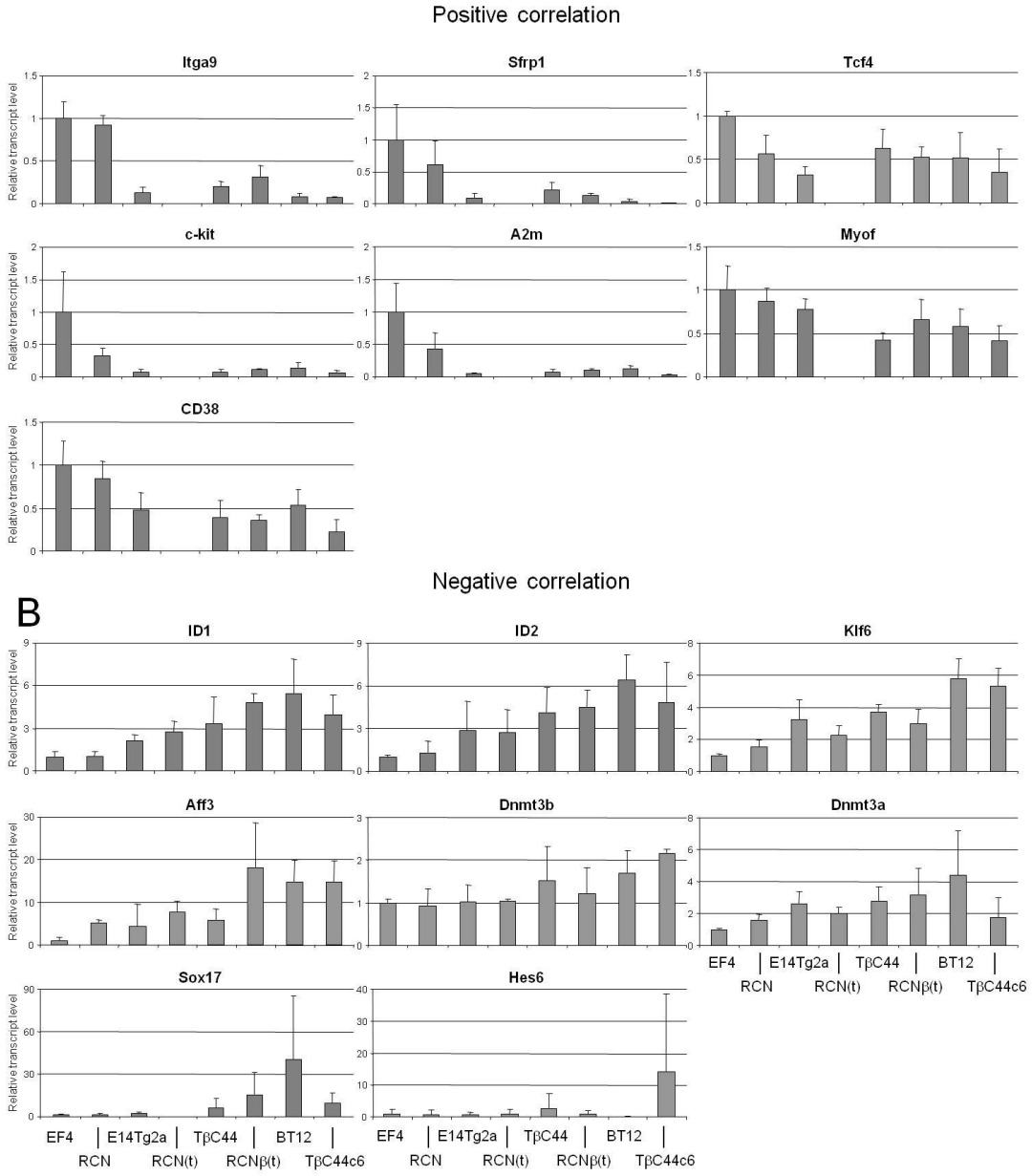


Figure 3.2:Validation of Nanog target gene expression in multiple ES cell lines.

Quantitative PCR analysis of gene expression in two cell lines overexpressing Nanog (RCN, EF4), two lines with wild-type levels of Nanog expression (E14Tg2a, RCN(t)), two cell lines in which one of the two Nanog alleles has been deleted by homologous recombination (TβC44,RCNβ(t)) and two Nanog null cell lines (TβC44c6, RCNβH(t)). Expression values, normalised on TBP transcript levels, and relative to EF4 are plotted for genes that are positively (A) or negatively (B) regulated by Nanog. Error bars: standard deviation of gene expression values measured in at least three independent experiments.

protein. Fusion to ER^{T2} is a well established system used to modulate subcellular protein localisation (Feil et al., 1997). While the fusion protein is confined to the cytoplasm in the absence of tamoxifen, addition of the drug results in prompt nuclear relocalisation. The Nanog-ER^{T2} expression vector was stably transfected in Nanog null TβC44c6 ES cells and three clones showing different transgene expression levels were selected (ESΔN-NERT ES cells (NERT))(**Figure 3.3A, 3.3B**). Nanog localisation in the presence or in the absence of tamoxifen was analysed by immunohistochemistry and confocal microscopy. Since both *Nanog* alleles are deleted in TβC44c6 ES cells, only expression of the fusion protein is detected in these cells. Predominant nuclear localisation of the fusion protein was observed as soon as 15 minutes after addition of tamoxifen. (**Figure 3.3C**).

The transcriptional effects of inducing nuclear translocation of Nanog in ESΔN-NERT cells were then tested by preparing total RNA and protein lysates from cells stimulated for 24 hours with tamoxifen. 7 out of 9 genes analysed showed variations in the level of expression in three independent ESΔN-NERT lines (A2m and Sox17 did not show notable changes, **Figure 3.4**).

Next, the dynamic of Nanog transcriptional control over its target genes were investigated. With this aim, timecourse experiments were performed in ESΔN-NERT cells using the line that showed the highest target gene induction levels in previous experiments (Clone 3, **Figure 3.3B**). ESΔN-NERT c3 cells were stimulated with tamoxifen over a period of 6 hours and total RNA was prepared at 1 hour intervals. The mRNA level of 12 different transcripts was measured over time and compared with the basal levels of expression detected in the parental TβC44c6 cells. In agreement with prior observations, 9 positively regulated genes showed increased transcript levels, and one negative target was downregulated, after tamoxifen addition (**Figure 3.5**). Different kinetics of induction were detected among the responding genes, ranging from changes in transcript levels one hour after tamoxifen stimulation for Klf4,Esrrb or ID1 to a period of 3 hours before detecting induction of Itga9. Moreover, a reduction in GFP transcripts originated from the *Nanog*:GFP

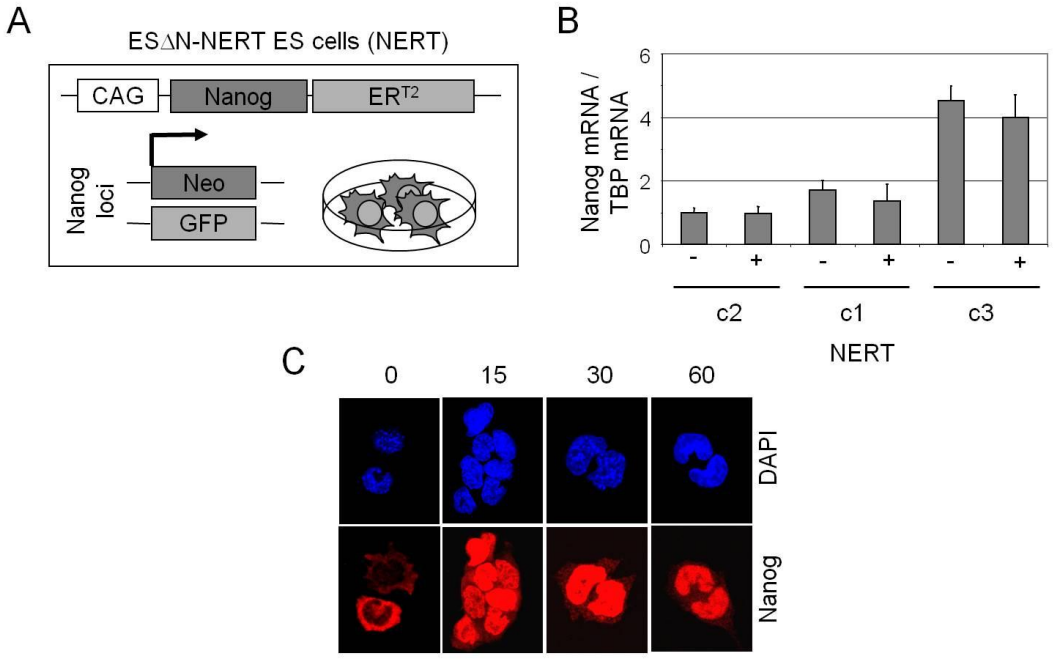


Figure 3.3: Characterisation of ESΔN-NERT ES cells.

A: Schematic representation of the genetic manipulations performed on ESΔN-NERT ES cells. **B**: Transcript level of Nanog in ESΔN-NERT lines cultured in the presence (+) or in the absence (-) of tamoxifen for 24 hours. Relative transcript levels normalised to TBP expression are presented for three independent clones. Error bars: standard deviation of the technical errors in gene expression quantification in one experiment. **C**: Immunohistochemical analysis of the intracellular localisation of Nanog in ESΔN-NERT c3 ES cells treated with 1μM tamoxifen for 0, 15, 30 or 60 minutes.

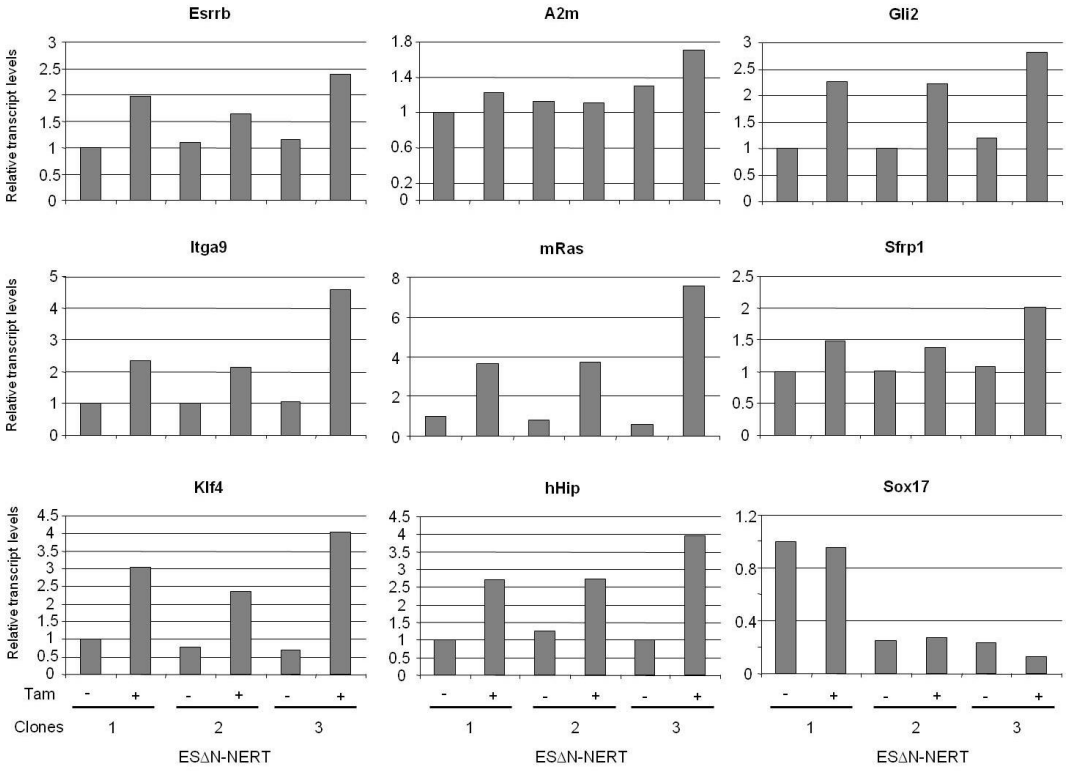


Figure 3.4: The identified target genes respond to Nanog nuclear localisation in ESΔN-NERT ES cells.

Transcript level of 9 Nanog target genes in ESΔN-NERT lines cultured in the presence or in the absence of tamoxifen for 24 hours. Relative transcript levels normalised on TBP expression are presented for three independent clones.

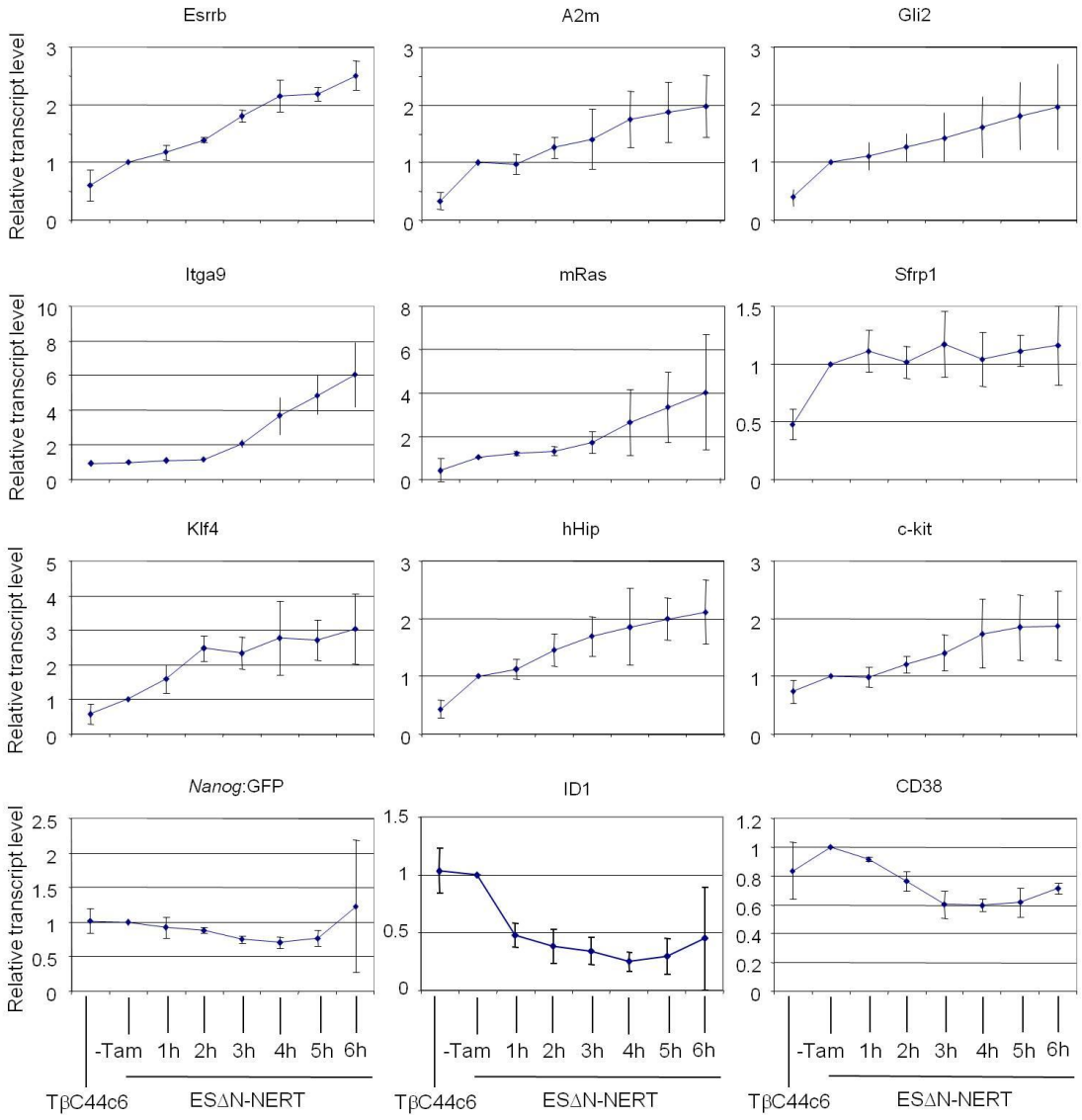


Figure 3.5: Nanog nuclear localisation induces prompt upregulation of its target genes in ESΔN-NERT c3 cells.

Kinetics of transcript upregulation for 12 Nanog target genes in ESΔN-NERT c3 cells stimulated with tamoxifen for 0, 1, 2, 3, 4, 5 or 6 hours. Relative transcript levels normalised on TBP expression. Error bars: standard deviation of gene expression values measured in three independent experiments.

reporter allele present in TβC44c6 ES cells was detected, validating the observation that Nanog represses its own expression (Navarro, 2012). Only one gene, CD38, showed an opposite trend compared to the results obtained in previous experiments (**Figure 3.2**), indicating that the levels of expression observed in Nanog mutant lines could be consequent to indirect regulation. Interestingly, Sfrp1 was strongly upregulated in ESΔN-NERT cells compared to the parental TβC44c6 line even in the absence of tamoxifen. This observation suggests that the Nanog-ER^{T2} fusion protein must not be completely excluded from the nucleus in untreated cells, that very low levels of Nanog activity are sufficient to alter the transcription of some target genes and that *Sfrp1* is exquisitely sensitive to Nanog.

Genes showing longer response time are less likely to be direct targets of Nanog. In an attempt to discriminate more clearly between direct or indirect regulation of gene expression, the transcript levels of 12 target genes were measured in ESΔN-NERT cells stimulated with tamoxifen in the presence or in the absence of cycloheximide. Inhibition of protein synthesis should exclude the possibility that the observed transcriptional modulation is mediated by indirect changes in the levels of a second Nanog-regulated factor. Based on results from previous timecourse experiments, it was possible to select 3 hours as the earliest time-point at which robust transcriptional induction or repression occurs for most target genes. For this reason, RNA was collected from cells cultured for 3 hours in the presence of tamoxifen. Cycloheximide was added to the cells 30 minutes before tamoxifen to allow for complete inhibition of protein synthesis before inducing Nanog translocation to the nucleus. Strikingly, 8 out of 12 Nanog target genes showed induction or repression in response to cycloheximide treatment, independently of tamoxifen stimulation (**Figure 3.6**). Consequently, it was difficult to detect an additional effect of tamoxifen stimulation for the genes which responded to cycloheximide. For example, cycloheximide did not impair tamoxifen induced downregulation of ID1 but resulted in an overall increase in the transcript levels of this gene. 4 (*Nanog*:GFP, CD38, Itg9a, Sfrp1) out of 12 Nanog responsive genes were not affected by treatment with cycloheximide alone (**Figure 3.6**). Among these, downregulation of *Nanog*:GFP

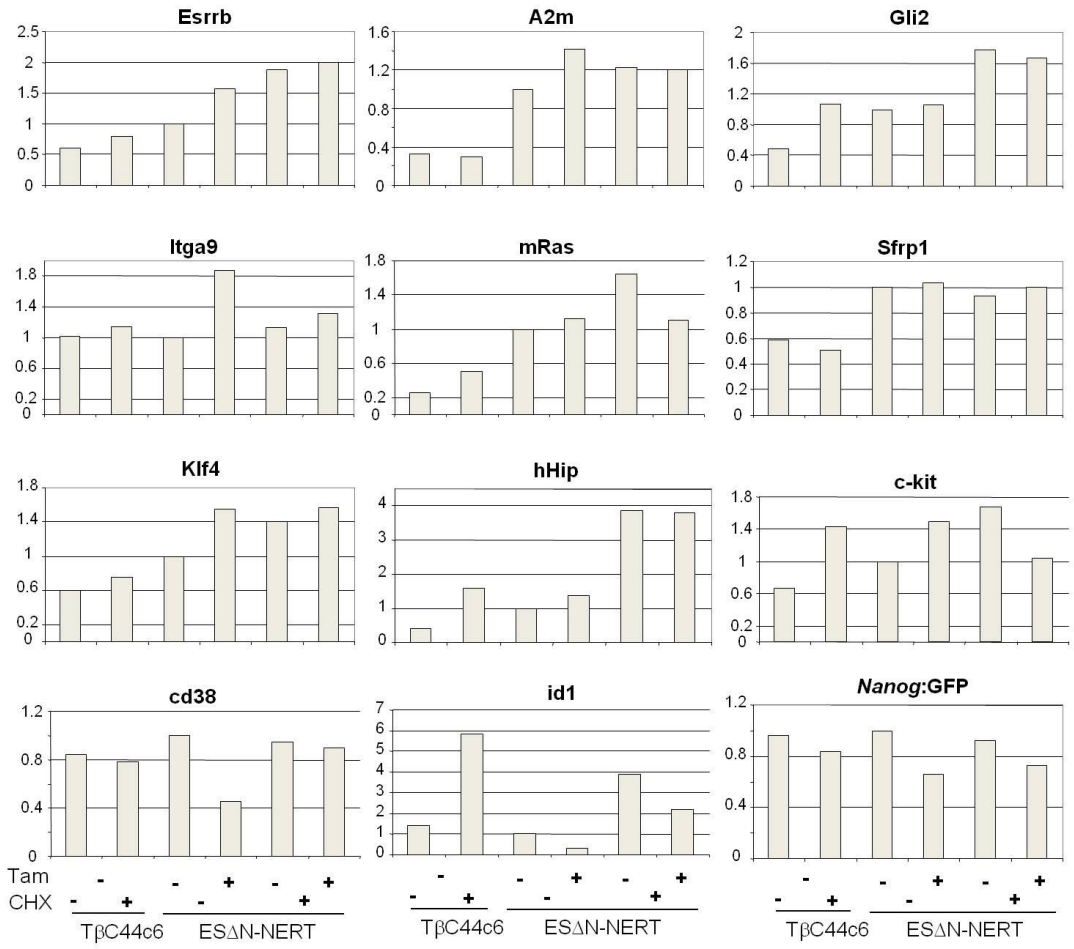


Figure 3.6: Identification of Nanog direct transcriptional targets.

Relative transcript levels of 12 Nanog target genes in TβC44c6 and ESΔN-NERT c3 cells cultured for 3 hours in the presence or in the absence of tamoxifen and 30 μg/ml of cycloheximide. Cycloheximide was added to the culture medium 30 minutes before addition of tamoxifen. Relative transcript levels normalised on TBP expression.

could not be suppressed by cycloheximide while protein synthesis inhibition could suppress tamoxifen induced repression of CD38 and induction of Itga9. Considering that Itg9a also shows a long response time after Nanog nuclear relocalisation, it is probable that this gene is not directly activated by Nanog. Overall, the complex response that cycloheximide treatment elicits in ES cells complicates the interpretation of these results. Treatment of the cells for shorter time periods could help reducing the general perturbation of transcription observed in these experiments. As an alternative to the use of protein synthesis inhibition, claims of direct regulation should be based on the analysis of transcriptional output after short times of tamoxifen stimulation in ESΔN-NERT cells (1 hour or less).

The genome-wide analysis presented in **Figure 3.1** identified a list of genes whose transcript levels respond to stable alternations in Nanog levels or to fluctuations in Nanog expression occurring in ES cells over a period of days (ref chambers 2007). As discussed, such analysis does not distinguish between direct or indirect transcriptional regulation. To solve this limitation, the scope of the observations made on a limited number of target genes was extended by performing microarray analysis of Nanog induction timecourse experiments in ESΔN-NERT cells. Such experiments, analysed in conjunctions with the available transcriptional profiling data, consolidated the identified list of targets of Nanog (**Table 3.1**) and shed light on the dynamics of Nanog control over its target genes at the genome-wide level.

Microarray analyses were performed in triplicate at 1 hour time intervals over a 6 hour period following Nanog nuclear re-localisation in ESΔN-NERT c3 cells. Only 64 genes showed a differential gene expression pattern (fold change$\geq$1.5, p$\leq$0.05) during the time course (**Figure 3.7A, Table 3.1**). This was surprising, since 2 independent ChIP-Seq studies (Chen et al., 2008; Marson et al., 2008) identified more than 5000 common Nanog biding sites in the ES cell genome. Using the publically available GeneProf software (Halbritter et al., 2012), 49 out of 64 Nanog-sensitive genes were determined to be reproducibly bound by Nanog (Chen et al., 2008; Marson et al., 2008). Conversely, 99% of the genes identified as putative

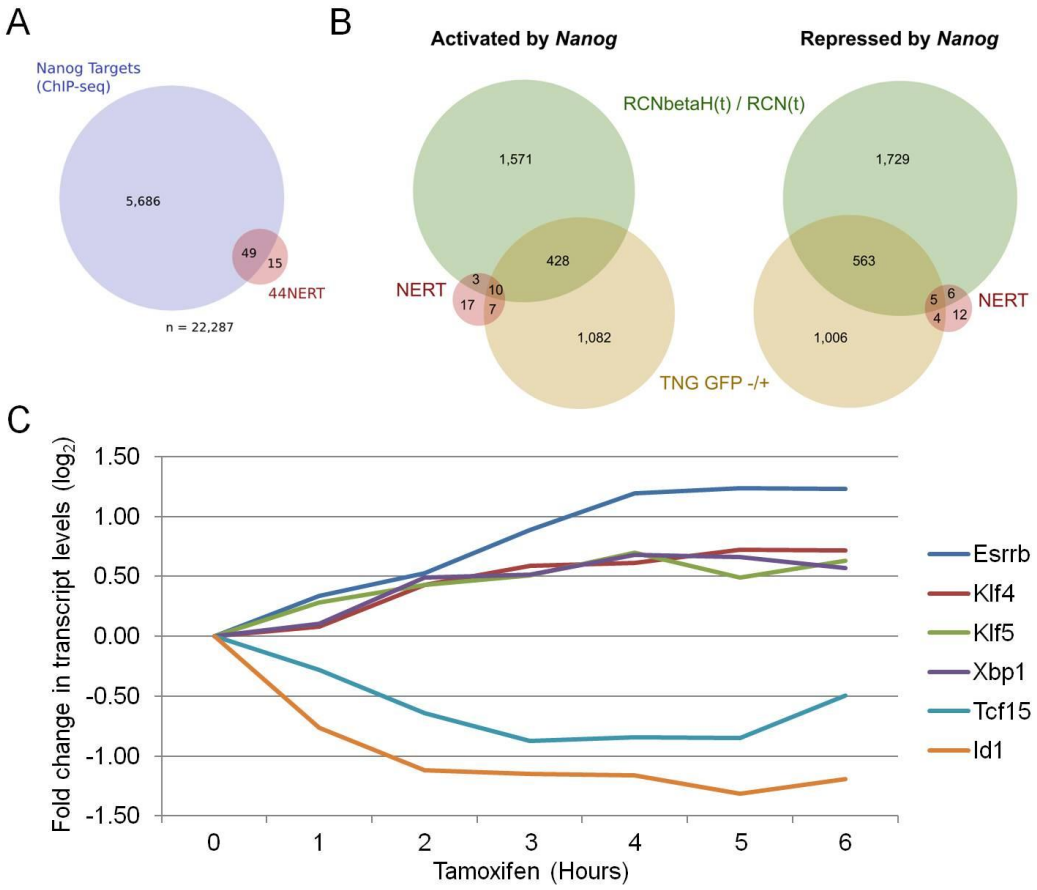


Figure 3.7: Genome-wide analysis of the transcriptional response to Nanog nuclear re-localisation

A: Venn Diagram showing the intersection of significantly up- or downregulated genes (fold change ≥ 1.5; p≤0.05) identified in the ESΔN-NERT timecourse analysis compared to genes bound by Nanog according to two independent genome-wide ChIP studies. **B**: Venn Diagram showing the intersection of significantly up- or downregulated genes (fold change ≥ 1.5; p≤0.05) identified in the ESΔN-NERT timecourse analysis, comparing lines stably expressing different Nanog levels (RCN(t) vs. RCNβH(t)) and analysing the effects of fluctuations in Nanog expression (TNG^{high} vs. TNG^{low}). **C**:Plot showing expression dynamics of significantly up- or downregulated genes (Fold change≥1.5, p≤0.05) after ESΔN-NERT stimulation with tamoxifen as indicated; only transcription factors showing non-contradicting changes in all the datasets and significant expression levels (r.p.m≥50) were selected. Mean expression levels in three independent experiments.

All bioinformatic analysis was performed in collaboration with Florian Halbritter using the Geneprof software suite (Halbritter et al., 2012).

Nanog targets by ChIP are insensitive to changes in Nanog over the time course of the analysis (**Figure 3.7A**). These results show that most of the genes to which Nanog is bound are impervious to changes in the levels of Nanog protein in the cell, possibly because the presence of other transcriptional factors at the regulatory regions of these genes is sufficient to sustain their expression.

It was then decided to determine the overlap between differentially regulated genes identified by **1)** comparing lines stably expressing different Nanog levels (RCN(t) vs. RCNβH(t)), **2)** analysing the effects of fluctuations in Nanog expression (TNG^{high} vs. TNG^{low}) or **3)** monitoring rapid kinetics of transcriptional activation after Nanog induction (NERT timecourse). Interestingly, a progressive reduction in the total number of differentially expressed genes was observed comparing datasets that are increasingly stringent in excluding the effects of indirect regulation (**1:** 2012 ; **2:** 1527 ; **3:** 37 positive targets; **1:** 2303 ; **2:** 1578 ; **3:** 27 negative targets respectively) (**Figure 3.7B**). The low number of targets identified in the timecourse experiments is partially attributable to the stringent 1.5 fold threshold imposed during the analysis (Fold change>=1.5, P<=0.05), since a more relaxed criteria (Fold change>=1.25, P<=0.05) yields 185 differentially expressed genes. Nonetheless, this comparison suggests that the analysis of the long term effects of manipulating the levels of a transcription factor has a poor ability to identify its direct transcriptional targets.

Strikingly, Esrrb was the transcript showing the most pronounced induction in this timecourse analysis (**Figure 3.7C**). Of the other 63 targets, 10 (Klf4, Klf5, Xbp1, Ets2, Lmo4, Sox21, Tcf15, Otx2, Hmgxb4, ID1) are transcription factors expressed at significant levels (reads per million >50 in at least one deep-sequencing dataset). Of these, 6 show changes in expression consistent with variations in RCN(t) vs. RCNβH(t) and TNG^{high} vs. TNG^{low} datasets (**Figure 3.7C, Table 3.1**). Among these, the gene showing the greatest positive change at 6 hours after Esrrb is Klf4. Confirming the results obtained in prior deep sequencing studies, ID1 is the transcript showing the most pronounced decrease after Nanog induction.

3.3: Transcriptional dynamics following Nanog nuclear relocalisation in ESΔN-iNanog cells

The reduced number of Nanog transcriptional targets identified by tamoxifen induction of ESΔN-NERT cells raised the possibility that the Nanog-ER^{2} fusion protein expressed in this line might not be fully functional, and could therefore fail to activate part of wild-type Nanog target genes.

To investigate this possibility, it was decided to test the effects of overexpressing Nanog-ER^{T2} or wild-type Nanog proteins in an identical Nanog$^{-/-}$ background. Tβc44c6 derived ESΔN-NERT or ESΔN-iNanog ES cells (which carry a doxycycline inducible Nanog transgene; See **Figure 5.6**) were replated at clonal density in the presence or absence of LIF, inducing or not inducing Nanog function. After 7 days plates were stained for alkaline phosphatase (AP) expression and colony morphology scored. Nanog-ER^{T2} induction was able to sustain LIF independent self-renewal to an extent comparable to wild-type Nanog protein (**Figure 3.8A**), arguing in favour of an unaffected functionality of the fusion protein.

To directly compare the transcriptional functionality of Nanog-ER^{T2} and wild-type Nanog, timecourse microarray analyses were performed following induction of wild-type Nanog in Nanog-/- ESΔN-iNanog ES cells. Full transcript induction in ESΔN-iNanog cells is achieved by 6 hours of doxycycline treatment (**Figure 3.8B**). Gene expression was therefore analysed after inducing ESΔN-iNanog cells for 0, 6 or 12 hours. In this system, only 31 genes showed ≥1.5-fold change in expression after 12 hours of induction ($p \leq 0.05$) (**Figure 3.8B**). The lower number of genes identified is likely to result from the slower induction of nuclear Nanog in ESΔN-iNanog compared to ESΔN-NERT cells. The vast majority of targets (21/31) were also identified in ESΔN-NERT cells and 8/10 of the remaining genes are differentially expressed in ESΔN-NERT cells but with ≤1.5-fold change.

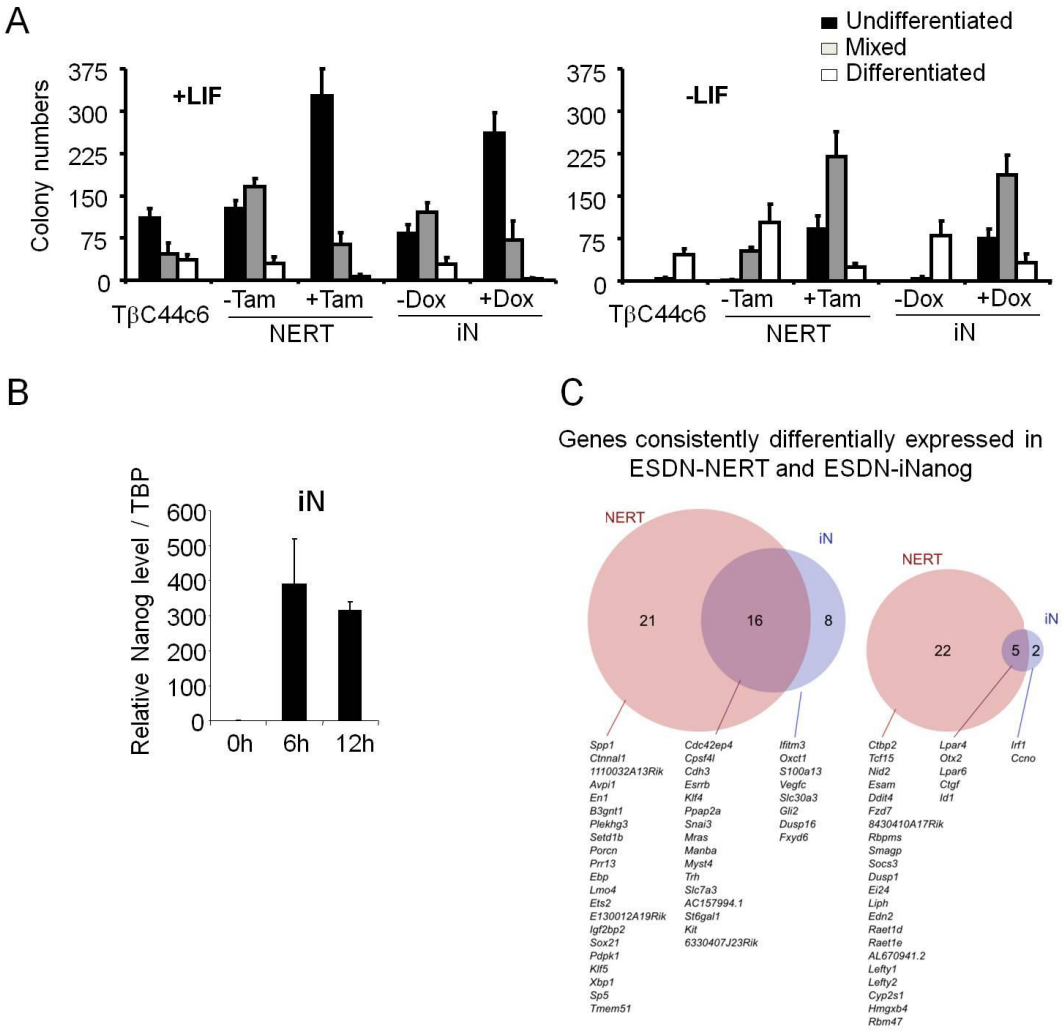


Figure 3.8: Timecourse microarray analysis of doxycycline induction in ESΔN-iNanog and ESΔN-iEsrrb cells

A: Number and type of colonies formed 7 days after clonal density plating of TβC44c6, ESΔN-NERT or ESΔN-iNanog cells in the presence or absence of LIF, with or without doxycycline or tamoxifen. Error bars: standard deviation (n=3). **B**: Nanog mRNA expression in ESΔN-iNanog (iN) cells cultured in the presence of doxycycline for the indicated times. Error bars: standard deviations of 3 independent experiments. **C**: Intersection of significantly up- or downregulated genes identified by doxycycline stimulation of ESΔN-iNanog (iN) for 12 h and tamoxifen stimulation of ESΔN-NERT (NERT) for 6h.

Taken together these results demonstrate the full functionality of the Nanog-ER^{T2} fusion protein used in this analysis and identify a reliable list of Nanog responsive genes with which to explore the mechanisms of Nanog activity in ES cells.

3.4: Nanog control over Esrrb expression

Among the target genes identified by this study is Esrrb, an orphan nuclear receptor that participates in the pluripotency transcriptional network (Chen et al., 2008; Ivanova et al., 2006; Kim et al., 2008; Loh et al., 2006; van den Berg et al., 2010; van den Berg et al., 2008; Wang et al., 2006a; Zhang et al., 2008). Esrrb expression sustains ES cell self-renewal (Ivanova et al., 2006; Loh et al., 2006; Zhang et al., 2008) and promotes reprogramming of mouse embryonic fibroblasts (Feng et al., 2009). Since Esrrb function in ES cells has not been studied extensively, the transcriptional regulation of Esrrb was investigated in greater detail.

The murine *Esrrb* gene is more than 100 kb long, is constituted by 8 different exons, and can produce 4 different described transcript variants by alternative splicing (**Figure 3.9A**). To determine which of these transcript variants are expressed in ES cells, primers were generated that amplify all exon-exon junctions across the gene, the 3 alternative transcript start sites and the two alternative 3' untranslated regions (UTR) of the different messenger RNAs (**Figure 3.9A**). In ES cells the most abundant transcript starts from exon 2 and ends at exon 7, including a 2.5 kb long 3' UTR sequence (**Figure 3.9B**).

Esrrb transcript levels correlate with Nanog expression in different ES cell lines in a Nanog mutant series (**Figure 3.10A**). A robust correlation of Esrrb and Nanog expression in the cells could also be observed at the protein level (**Figure 3.10B**). Pre-mRNA levels, measured with primer pairs designed to cover the entire transcript length, accurately recapitulated the differences in expression observed for mature Esrrb mRNA (**Figure 3.10C,D**), indicating that the observed variations in Esrrb mRNA levels resulted from a direct control of Nanog over Esrrb transcription and were not attributable to the stabilisation of Esrrb messenger RNA.

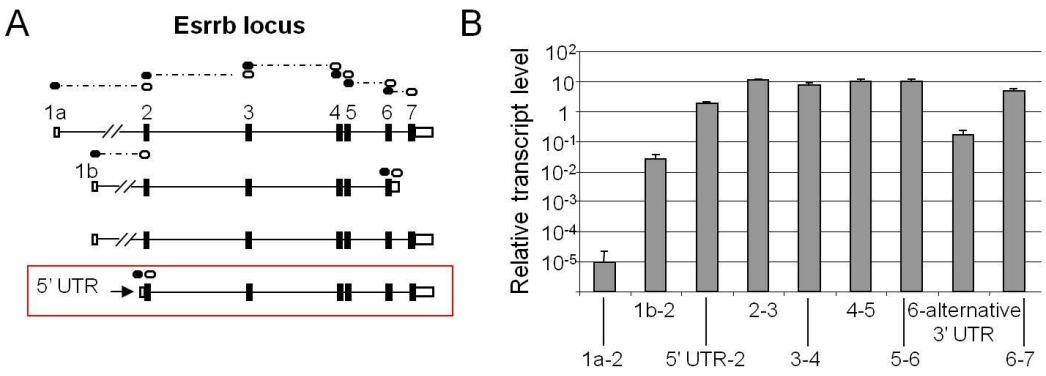


Figure 3.9: Determination of Esrrb transcript variant expression in ES cells.

A: Schematic representation of the structure of four described murine Esrrb transcript variants. Black boxes: exons. White boxes: 5' and 3' untranslated regions (UTR). Solid lines: introns. Black and white circles connected by dashed lines: forward and reverse primer pairs used in panel B. **B**: Relative transcript levels measured using primer pairs that amplify different exon-exon junctions, the 3 alternative transcript start sites and the two alternative 3' UTR sequences of Esrrb.

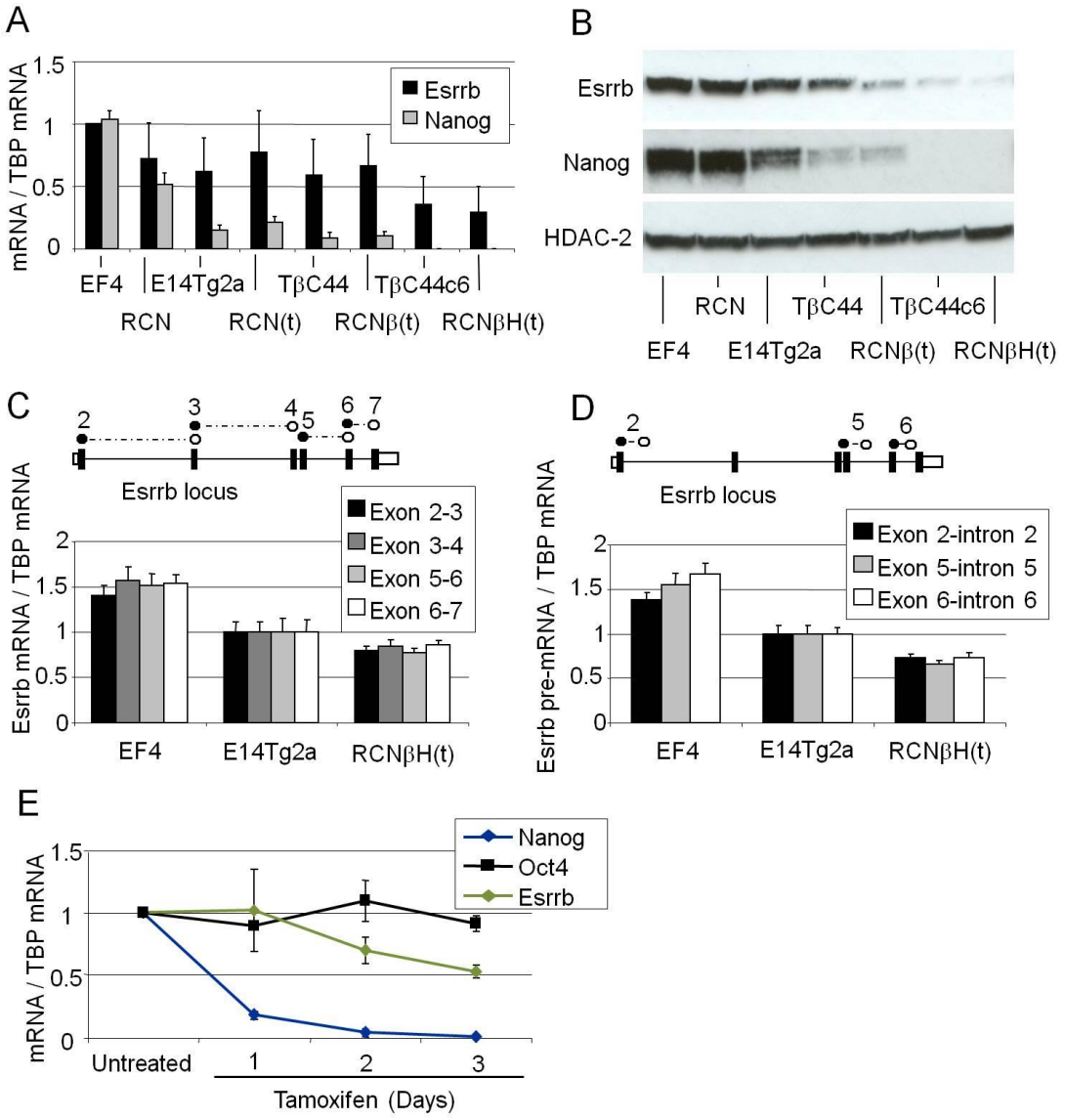


Figure 3.10: Esrrb and Nanog mRNA levels correlate in ES cells

A: Relative Esrrb and Nanog transcript levels in two cell lines overexpressing Nanog (RCN, EF4), two lines with wild-type levels of Nanog expression (E14Tg2a, RCN(t)), two cell lines in which one of the two Nanog alleles has been deleted by homologous recombination (TβC44,RCNβ(t)) and two Nanog null cell lines(TβC44c6, RCNβH(t)). Error bars: standard deviation of gene expression values measured in 4 independent experiments. **B**: Western blot analysis of Esrrb and Nanog protein levels in different ES cell lines. **C,D**: Esrrb mRNA and pre-mRNA levels in EF4, E14Tg2a and RCNβH(t) ES cells. Relative expression levels measured using primer pairs amplifying successive positions along the transcript. Error bars: standard deviation of the mean for gene expression levels measured in 5 independent experiments. **E**: Relative Esrrb, Nanog and Oct4 mRNA levels in RCNβH cells treated for 1, 2 or 3 days with tamoxifen to induce loss of Nanog expression driven by a floxed transgene present in this cell line. Error bars: standard deviation of gene expression values measured in 3 independent experiments.

In order to show that loss of Nanog activity directly results in decreased Esrrb expression, RCNβH cells were treated with tamoxifen to induce excision of the floxed Nanog transgene present in this cell line. An almost complete loss of Nanog transcript was detected as soon as 1 day after treatment, with no residual mRNA detectable at day 2 or day 3. Esrrb expression was significantly reduced after 2 days of treatment, and further decreased on day 3 (**Figure 3.10E**). Induced cell differentiation could not explain these effects, since Oct4 levels were stable over the course of the experiment.

ESΔN-NERT cells were then employed to investigate whether Nanog exerts direct transcriptional control over Esrrb expression. As previously shown, a 1.5 to 2 fold induction in Esrrb mRNA was detected when ESΔN-NERT cells were treated with tamoxifen for 24 hours (**Figure 3.4**). The fold change in Esrrb levels correlated well with the expression of the Nanog-ER^{T2} transgene in the three different lines analysed (**Figure 3.3B**). Similarly, a significant increase in Esrrb protein levels was detected by western blot (**Figure 3.11A**). Crucially, Esrrb pre-mRNA increases as early as 20 minutes after tamoxifen treatment of ESΔN-NERT cells **(Figure 3.11B)**, with mature transcript levels showing an increase after 1 hour (**Figure 3.5**). Finally, ChIP assays showed that Nanog binds to both *Esrrb* promoter and enhancer (Chen et al., 2008). Restoring Nanog binding in ESΔN-NERT cells resulted in a 2 fold increase of RNApolII recruitment at the Esrrb promoter (**Figure 3.11C**).

Taken together these results establish *Esrrb* as a major positive target of direct transcriptional activation by Nanog in ES cells.

3.5: Nanog control over Esrrb transcriptional elongation and pause release

Esrrb was then used as a model to characterise in greater detail the dynamics and mechanisms of Nanog transcriptional control over its target genes.

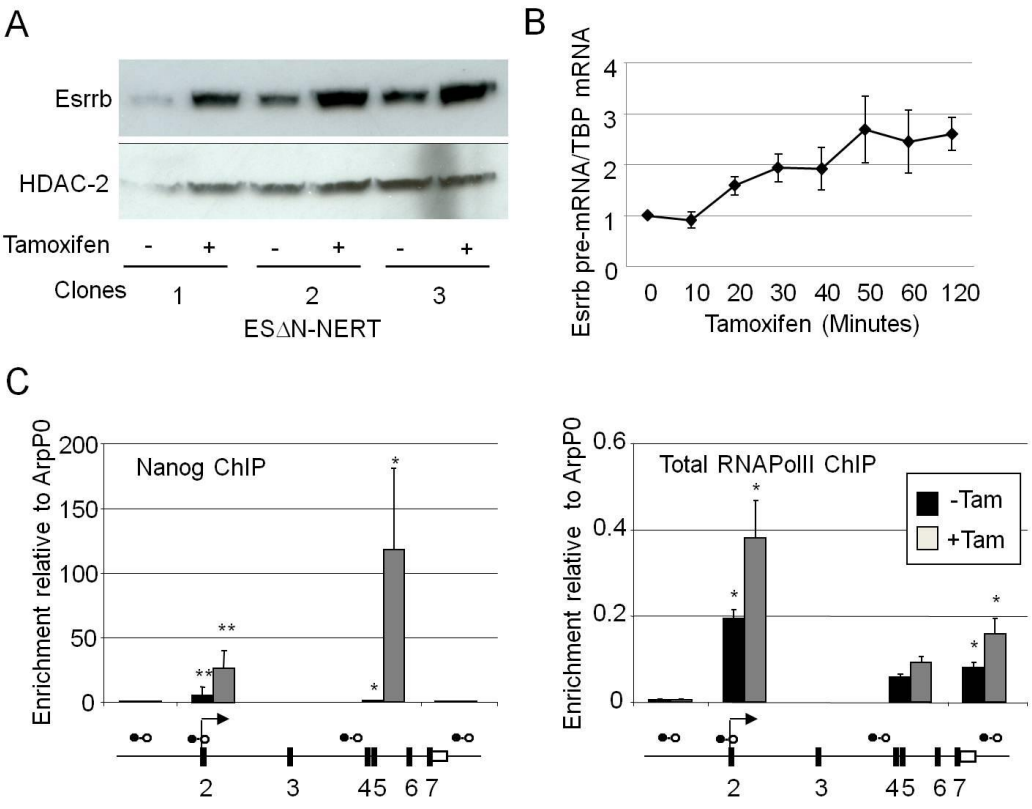


Figure 3.11: Nanog directly regulates Esrrb transcription

A: Western blot analysis of Esrrb and HDAC-2 protein expression in three ESΔN-NERT lines cultured in the presence or in the absence of tamoxifen for 24 hours. **B**: Kinetic of Esrrb pre-mRNA upregulation in ESΔN-NERT cells stimulated with tamoxifen for 0, 10, 20, 30, 40, 50, 60 or 120 minutes. Relative transcript levels normalised on TBP expression. Error bars: standard deviation of gene expression values measured in three independent ESΔN-NERT clones. **C**: Chromatin from ESΔN-NERT cells treated with 1μM tamoxifen for 0 or 24 hours was immunoprecipitated with Nanog or total RNAPolII antibodies. Enrichment relative to the ArpP0 promoter is measured using the primers indicated at *Esrrb*. Error bars: standard deviation (n=3) *(p≤0.05), **(p≤0.01).

The length of the *Esrrb* transcription unit expressed in ES cells is 51.3kb (**Figure 3.9**). This is relatively long and could be a point of controlling Esrrb levels by influencing RNAPII progression across the gene. In an attempt to test this hypothesis, the time required from initiation of transcription to the production of a mature messenger RNA was measured in Nanog overexpressing (EF4) or Nanog null (RCNβH(t)) ES cells. Cells were treated for 3 hours with flavopiridol (Chao et al., 2000; Chao and Price, 2001), an inhibitor of CDK9 (P-Tefb active subunit) which blocks RNA synthesis by RNA Polymerase II (RNAPolII). Under these conditions, recruitment of RNA PolII to gene promoters is unaltered, but release of the polymerase is blocked, resulting in widespread accumulation of paused enzyme at active loci throughout the genome (Rahl et al., 2010). When flavopiridol is removed, productive initiation of transcription and elongation can take place. By measuring the level of nascent messenger RNA, using primers amplifying successive position from 5' to 3' along the gene (**Figure 3.12a**), it is possible to measure the transcriptional speed of elongating RNAPolII (Singh and Padgett, 2009). Esrrb pre-mRNA levels were measured every 5 minutes in cells released from flavopiridol. As shown in **Figure 3.12b**, complete transcription of the 50 kb long Esrrb pre-messenger RNA in EF4 requires approximately 15 minutes, at an average speed of more than 3 kb/min. This is in broad agreement with the rate measured for 9 large human genes in a previous study (average speed 3.8 kb/min)(Singh and Padgett, 2009), suggesting that in the presence of Nanog Esrrb elongation is not subject to substantial delays.

Interestingly, comparison of the *Esrrb* transcription rate in *Nanog*$^{-/-}$ ES cells (**Figure 3.12b**) showed that in the absence of Nanog, RNAPII initiation of productive elongation (Exon 2-Intron 2) and consequently, progression across *Esrrb* were inhibited. Given that RNAPII is detected (**Figure 3.10c**) and would be expected to accumulate at the *Esrrb* promoter during Flavopiridol treatment in both Nanog overexpressing cells and *Nanog*$^{-/-}$ cells, the reduced progression of RNAPII across *Esrrb* in *Nanog*$^{-/-}$ ES cells following wash-out of Flavopiridol suggests that, in addition to enhancing recruitment, Nanog stimulates release of RNAPII from the *Esrrb* promoter.

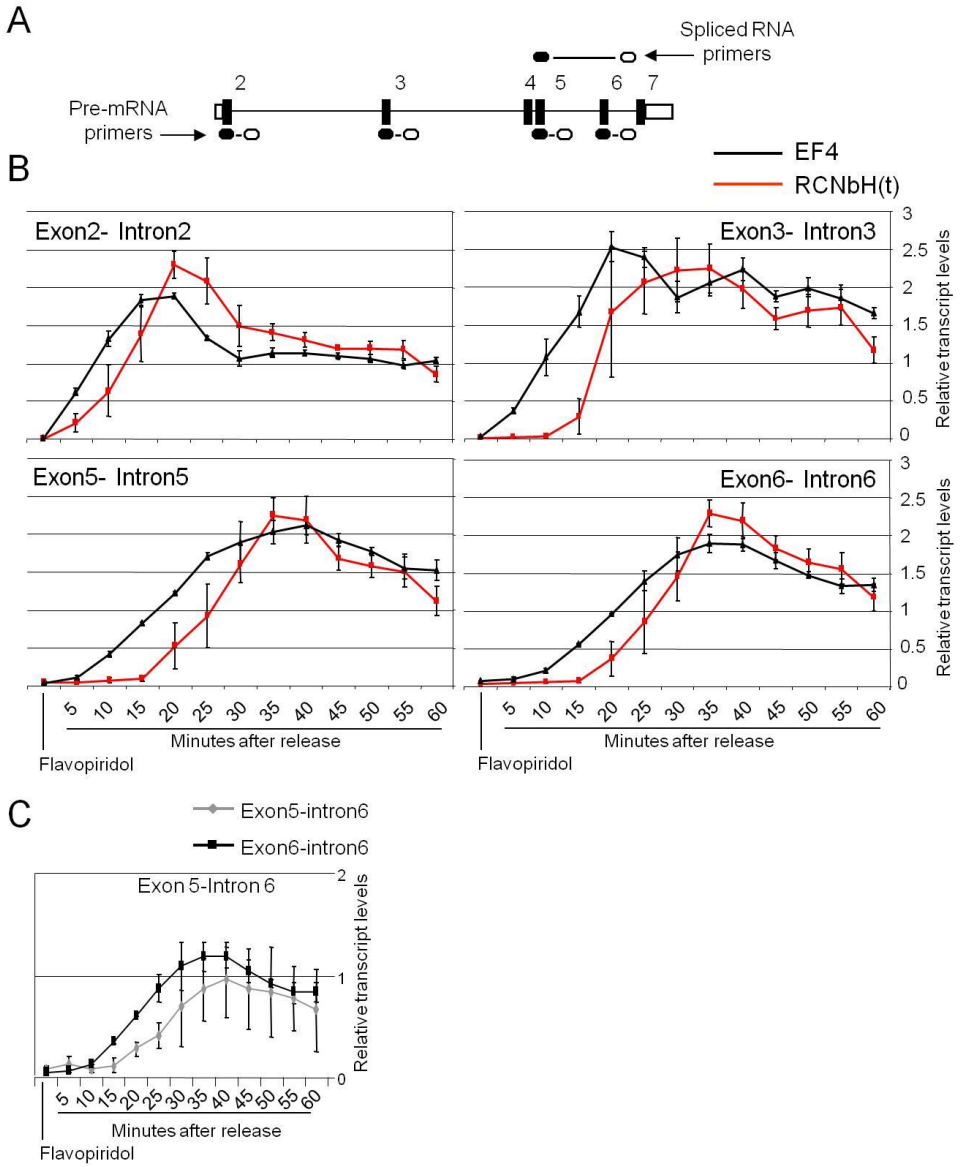


Figure 3.12: Kinetics of Esrrb mRNA transcription and splicing in ES cells

A: Schematic representation of the Esrrb transcript variant expressed in ES cells. Black boxes: exons. White boxes: 5' and 3' UTR regions. Solid lines: introns. Black and white circles connected by lines: forward and reverse primer pairs used in panels B (bottom) and panel C (top). **B**: EF4 or RCNβH(t) cells were released from flavopiridol after 3 hours of treatment and relative levels of Esrrb pre-mRNA measured at the indicated times using primers amplifying successive exon-intron junctions along the gene (see A). Transcript levels normalised to ribosomal RNAs. Error bars: standard deviation of the mean in 3 (EF4) and 4 (RCNβH(t)) independent experiments. **C**: Relative levels of spliced or unspliced Esrrb pre-mRNA measured using primer pairs amplifying the exon 5-intron 5 or exon 5-intron 6 junctions after release of EF4 cells from flavopiridol. Relative transcript levels normalised on ribosomal RNAs. Error bars: standard deviation of the mean for gene expression values measured in 3 independent experiments.

Finally, the time required for completion of the splicing events that will generate a mature mRNA from the nascent transcript was determined using the primers schematised in **Figure 3.12A**. As shown for the exon 5-intron 6 junction, approximately 5 minutes are required to first detect the products of the splicing events that join successive exons along the gene (**Figure 3.12C**). This observation, together with the finding that Esrrb transcription requires 15 minutes, suggests that splicing of Esrrb messenger RNA occurs co-transcriptionally in the nucleus of ES cells.

Taken together the data presented in this chapter suggest that Nanog controls expression of its target genes by influencing RNApolII recruitment at promoters. In addition, Nanog might be involved in modulating RNApolII initiation of transcription and progression of productive elongation of its targets. Further studies are now required to extend the preliminary observations made on Esrrb to other Nanog responsive genes and to fully understand the causes of the observed delayed transcriptional initiation in RCNβH(t) cells. In particular, it will be crucial to determine whether RCNβH(t) cells show reduced levels of RNApolII accumulation at the promoters of Nanog target genes after treatment with flavopiridol, since this could explain the extended time required to detect initiation of transcription in these cells.

3.6: Discussion

By comparing *Nanog*$^{+/+}$ with *Nanog*$^{-/-}$ and *Nanog*:GFP^{hi} with *Nanog*:GFP^{low} ES cells, and by taking advantage of Nanog$^{-/-}$ cells with distinct inducible Nanog function, a reliable list of the direct targets of Nanog transcriptional activity in ES cells was identified. The described approach presents three main advantages compared to previous attempts to characterise the transcriptional response to Nanog (Ivanova et al., 2006; Loh et al., 2006; Lu et al., 2009). First, the use of genetically null lines ensures complete abrogation of Nanog expression in the cells. Previous studies relied on short hairpin RNA (shRNA) expression driven by randomly integrated viral

vectors. In knock-down experiments the possibility that heterogeneous attenuation of Nanog expression occurs in the ES cell population cannot be excluded. Thus, a second advantage of the approach used in this thesis is that the use of inducible Nanog expression in a null background allows early responses to be characterised in a relatively uniform responding population. A third advantage is that, in all knockdown studies, Nanog downregulation coincides with widespread differentiation in the cultures, due partially to the long timescales employed. Many of the differentially expressed genes identified in the prior experiments are therefore likely to be modulated in response to the general dismantling of the pluripotency gene regulatory network (GRN) triggered by differentiation and not specifically responsive to Nanog activity in the cells.

The fact that knock-down of Nanog, and other pluripotency factors such as Esrrb, results in ES cell differentiation is a point of interest, given that stable knock-out ES cell lines for both genes are able to self-renew ((Chambers et al., 2007), and see Chapter 8). The observation that increased differentiation was observed upon derivation of Nanog (Chambers et al., 2007) and Esrrb (see Chapter 8) stable knock-out lines suggest that ES cells are able to adapt to the deprivation of key components of the pluripotency GNR. This adaptation is not consequent to genetic abnormalities acquired during the derivation of the lines, since both Nanog and Esrrb cells retain the ability to contribute to chimaeric embryos (Chambers et al., 2007; Martello et al.).

The analysis presented in this chapter was able to discriminate between direct and indirect transcriptional regulation. This study characterised the transcriptional alterations occurring as a consequence of the spontaneous fluctuation in Nanog protein levels observed in ES cells. Nanog fluctuations occur over a period of days (Chambers et al., 2007). This time frame ensures that the observed differences are not consequent to long-term adaptations to elevated or reduced Nanog levels in the cells. Importantly, analysis of Nanog:GFP^{+} and Nanog:GFP^{-} datasets identified a reduced number of target genes relative to the Nanog$^{+/+}$ vs Nanog$^{-/-}$ comparison (around 4200 and 3000 respectively, **Figure 3.7B**). Many of the transcripts showing differential expression in Nanog:GFP^{+} and Nanog:GFP^{-} cells are likely to originate

from genes that share common regulators with Nanog or are Nanog regulators rather than being transcriptional targets of Nanog. Conversely, a fraction of the genes identified comparing $Nanog^{+/+}$ with $Nanog^{-/-}$ cells may be indirectly activated or repressed in the cells as a consequence of Nanog activity on other transcriptional or epigenetic modulators. The intersection of both $Nanog^{+/+}$ vs $Nanog^{-/-}$ and Nanog:GFP^{+} and Nanog:GFP^{-} datasets permitted to overcome the reciprocal limits of these two approaches and identify ~1000 potential direct Nanog targets, in contrast with the 1531 genes identified by knockdown studies (Ivanova et al., 2006). Most importantly, the use of $Nanog^{-/-}$ ES cell lines in which Nanog function can be re-instated in a rapid and controlled way allowed the investigation of the kinetics of Nanog control over transcription. It was possible to identify transcripts showing induction or repression as soon as 1 hour (**Figure 3.7C**) after Nanog nuclear re-localisation in ESΔN-NERT ES cells. Such rapid kinetics of transcriptional regulation strongly suggest a direct activity of Nanog on the identified target genes. Strikingly, only 64 genes showed a greater than 1.5 fold change 6 hours after induction of Nanog in this cell line. The ability to detect transcriptional alterations by genome-wide microarray analysis depends on mature messenger RNA accumulation or degradation. The short timescale of the analysis in ESΔN-NERT cells might preclude the detection of part of the Nanog regulated network as a consequence of the long time required for these processes to achieve a detectable effect. In particular, since the average half-life of transcripts in mammalian cells is 6.9 hours (Tani et al., 2012), it is possible that a significant proportion of repressed genes was not identified in these experiments. Further experiments should directly measure the stability of ID1 and Tcf15 mRNA in flavopiridol treated ES cells, and compare it to the half-life of other mRNAs showing a slower decrease in response to Nanog induction. It is notable in this respect that lowering to 1.25 the differential expression threshold used to analyse the ESΔN-NERT timecourse data led to the identification of 185 target genes, with a vast proportion of the newly included genes showing less pronounced but steady and consistent variations over the 6 hour period of sampling. Overall this study revealed that transcriptional profiling studies based uniquely on the comparison of cell lines stably overexpressing or lacking expression of one given factor may lead to overestimation of the number of its direct targets.

Recent genome-wide ChIP studies have shown that multiple transcription factors bind in proximity to one another forming clusters at the regulatory elements of the genes controlled by the pluripotency GRN (Chen et al., 2008; Kim et al., 2008; Marson et al., 2008). This may imply that pluripotency factors exert combinatorial control over the expression of their target genes, a concept previously proposed in a more widespread context by Ptashne and Gann (2001). However, it is unclear to what extent changes in the level of a single factor influence the expression of the pluripotency GRN targets (Chambers and Tomlinson, 2009). Analysis of published data with the publically available GeneProf software suite (Halbritter et al., 2012) showed that Nanog is reported to bind to more than 5000 genes in at least two independent studies (Chen et al., 2008; Marson et al., 2008). Combined with the observation that only 64 genes show a greater than 1.5 fold change in expression after Nanog induction in ESΔN-NERT cells (**Figure 3.7A**), this implies that gain of Nanog binding is not sufficient to alter the transcription rate of most of these putative Nanog targets. It is possible that Nanog binding is not required for the efficient recruitment of other pluripotency factors to many of these genes. This might be due to the fact that these pluripotency factors recognise DNA sequences present at the regulatory elements of Nanog bound genes with sufficient affinity to ensure their independent recruitment to these regions. Alternatively, the presence of neighbouring binding sites for multiple pluripotency factors might result in the stabilisation of their association with the DNA at these regulatory regions, so that the additional presence of Nanog becomes irrelevant (Khalil et al., 2012). In any case, binding of these factors is able to drive activation of the transcriptional machinery in the absence of Nanog.

In order to further understand the dynamics of transcriptional regulation in ES cells and the robustness of the pluripotency GRN, future studies should determine whether combinatorial control may also limit the transcriptional response to changes in the level of pluripotency factors other than Nanog. Esrrb modulation also results in transcriptional changes for a limited number of genes (See Chapter 8.3). It is possible that the ability to elicit a transcriptional response in limited numbers of genes is a common feature of transcription factors that are heterogeneously expressed in ES

cells. In this respect, it is relevant that acute Oct4 depletion have a much wider effect on transcription, with 2714 genes showing a ≥1.5 fold change in expression only 5 hours after Oct4 ablation (Hall et al., 2009). The radical differences observed between the modulation of Oct4 and Nanog is supported by genetic evidence showing that tight control of Oct4 levels is necessary to maintain the pluripotent state (Niwa et al., 2000), while fluctuations in Nanog confer flexibility to the network (Chambers et al., 2007). It might therefore be that some pluripotency factors, like Oct4, lie at the heart of the housekeeping functions performed by the pluripotency GRN. If this is the case, binding of Oct4 to the regulatory elements of most GRN targets should prove to be necessary to allow the recruitment of additional pluripotency factors. Other factors, such as Nanog, and possibly Esrrb, might be required to precisely tune the expression of a limited number of genes that set the conditions for cell fate decisions. Combinatorial control would then ensure that depletion of these factors does not result in widespread alterations in the transcriptional output from the pluripotency GRN. Such differences could be explained by an hypothetical higher affinity of factors like Oct4 for their DNA binding elements, or by their different ability to serve as protein interaction hubs (Ding et al., 2012; Pardo et al., 2010; van den Berg et al., 2010; Wang et al., 2006a) in the process of stabilising additional binding at the regulatory elements of the pluripotency GRN.

Nanog has been cloned because it can confer LIF independence to ES cells (Chambers et al., 2003). This analysis characterised the transcriptional response to Nanog in cells cultured in the presence of LIF. Since STAT3 and Nanog bind to an overlapping set of sites throughout the genome (Chen et al., 2008), it is possible that the low number of Nanog responsive genes identified is due to a certain degree of redundancy in the transcriptional regulation exerted by Nanog and STAT3. Re-instating Nanog activity in ESΔN-NERT cells after LIF starvation, could lead to the identification of additional important mediators of Nanog function and shed light on the mechanism through which Nanog elevation renders LIF dispensable for ES cell self-renewal. Conversely, it might be that some of the target genes identified will fail to respond to Nanog in the absence of LIF.

The results presented establish that the use of knock-out cell lines and fusion between transcription factor proteins and the ER^{T2} domain (Feil et al., 1997) is a powerful tool for studying regulation of gene expression in mammalian cells. It would be interesting to apply this system to the investigation of transcriptional control by other pluripotency factors. The rapid translocation time achieved in this system (**Figure 3.3C**) made it possible to study the early transcriptional events that follow restoration of Nanog function in ES cells (**Figure 3.7C, 3.10B**). This work directly compared the potential of -ER^{T2} fusion to the use of doxycycline inducible expression strategies by applying both systems to reinstate Nanog function in an identical $Nanog^{-/-}$ background. Analysis of Nanog induction in ESΔN-NERT and ESΔN-iNanog cells led to the identification of 64 and 31 directly responsive genes respectively (**Figure 3.8C**), highlighting the superior sensitivity of ER^{T2} based approaches. One of the main limitations of fusion to ER^{T2} is the unpredictability of whether the resulting protein will be functional. Importantly, it was possible to show that the Nanog-ER^{T2} protein used in this study is comparable to wild-type Nanog in functional and transcriptional assays (**Figure 3.8A-C**). Potential confounding effects deriving from the biological activity of 4-0H-tamoxifen limit the use of this system to the study of nuclear receptors. In particular, tamoxifen has been shown to interfere with the transcriptional activation mediated by Errg, an orphan nuclear receptor closely related to Esrrb, inhibiting its interaction with the co-activator Ncoa-1 (Coward et al., 2001; Tremblay et al., 2001a) at doses lower than 1μM. Despite the fact that 4-OH-tamoxifen does not exert a comparable activity on Esrrb (Coward et al., 2001) and its effects have not been tested in ES cells, where Ncoa-1 is not expressed and Esrrb seems to partner with a different co-activator (Percharde et al., 2012), the potential antagonistic activity of 4-OH-tamoxifen discourages the use of an Esrrb-ER^{T2} fusion protein. Promising alternatives to ER^{T2} fusion could be the use of the recently described anchor-away (Haruki et al., 2008) or the ligand induced destabilisation (LID) (Banaszynski et al., 2006) systems.

Among the identified Nanog targets, Esrrb is the gene that shows the most pronounced transcriptional activation in tamoxifen treated ESΔN-NERT cells (**Figure 3.7C**). This work demonstrated that direct Nanog binding to the *Esrrb*

promoter and enhancer increases RNAPolII recruitment to the *Esrrb* promoter (**Figure 3.10B**) and elevates *Esrrb* pre-mRNA levels within 20 minutes (**Figure 3.10C**). Notably the second transcription factor that shows consistent correlation with Nanog in all the datasets is Klf4 (**Figure 3.1**, **3.7C, 3.8C**). Esrrb has been reported to be involved in the control of pluripotency and to promote reprogramming ((Chen et al., 2008; Ivanova et al., 2006; Kim et al., 2008; Loh et al., 2006; van den Berg et al., 2010; van den Berg et al., 2008; Wang et al., 2006a; Zhang et al., 2008) and this thesis). Klf4 has also been shown to sustain ES cell self-renewal in response to LIF signalling (Jiang et al., 2008; Niwa et al., 2009) and is a member of the quartet of Yamanaka's reprogramming factors (Takahashi and Yamanaka, 2006). This opens the possibility that other target genes also play a role in promoting ES cell self-renew. In particular, Gli2, the third most responsive transcription factor identified by the analysis of Nanog mutant ES cell lines (**Figure 3.1**), shows prompt upregulation in ESΔN-iNanog and ESΔN-NERT (1.48 fold change in the latter dataset). Gli proteins are zinc-finger transcription factors that mediate the transcriptional response triggered by activation of the hedgehog pathway (Hui and Angers, 2011). Interestingly, using luciferase based assays, recent studies have shown that Gli2 can bind to the *Nanog* promoter and activate its transcription (Po et al., 2010). Although performed in mouse cerebellar neurospheres, in which Nanog is not expressed, these finding might suggest that in ES cells Gli2 is engaged in a positive feedback loop sustaining Nanog expression.

Unexpectedly, the most downregulated genes in both the ESΔN-NERT and ESΔN-iNanog datasets was the bHLH factor ID1 (**Figure 3.7C, 3.8C**).The closely related protein ID2 was also detected among negatively regulated genes (**Figure 3.1**). At first glance, this finding may seem to be at odds with the notion that ID proteins are induced by BMP signalling and play a fundamental role in preventing ES cell neural differentiation (Ying et al., 2003). Since ES cells fluctuate between states of high and low Nanog expression (Chambers et al., 2007), it is possible that in Nanog positive cells *ID* genes expression is significantly diminished. As a consequence, despite the fact that all cells are exposed to similar BMP concentrations, Nanog negative cells could be the only fraction of the ES cell population able to effectively respond to

BMP signals. Elevation of ID levels in response to BMP may help Nanog negative cells to remain undifferentiated and eventually reacquire Nanog expression. Conversely, downregulation of ID protein expression in Nanog positive cells could allow downregulation of Nanog expression . BMP signalling could therefore be a driver of dynamic Nanog heterogeneity in ES cells.

In addition to repressing neural fate BMP signalling has been shown to promote mesodermal differentiation of ES cells (Finley et al., 1999; Johansson and Wiles, 1995) and is known to be essential for mesoderm specification during early mouse development (Winnier et al., 1995). An explanation to the observed repression of ID1 by Nanog could come from studies indicating that Nanog inhibits BMP induced activation of *Brachyury*, and ID1, by binding Smad1 and preventing its interaction with the p300 co-activator. The authors propose an additional function for Nanog in preventing BMP induced mesodermal differentiation (Suzuki et al., 2006). Nanog repression of ID1 could therefore be a consequence of its role in repressing mesodermal fate. Interestingly Nanog is expressed in the primitive streak region of the post-implantation mouse embryo (Osorno et al., 2012), a region in which BMP signalling is also detected (Winnier et al., 1995). In this context, Nanog could contribute to preventing premature mesodermal specification of epiblast cells before their migration through the primitive streak.

The results presented in this chapter establish that Nanog binding enhances the recruitment of RNAPolII at the *Esrrb* promoter and activates its transcription. Transcriptional activation does not depend uniquely on recruitment but also on efficient RNAPolII release from active promoters. RNAPolII is initially recruited at promoters in a “paused” state in which it engages in multiple rounds of abortive transcriptional initiation. Phosphorylation of the second serine in tandem repeats of a conserved sequence in the RNAPolII C-terminal domain augments the enzyme processivity and allows productive elongation of nascent RNAs (Core and Lis, 2008; Margaritis and Holstege, 2008). Pause release is a fundamental point of gene expression regulation in human cells (Core et al., 2008; Guenther et al., 2007). c-Myc, another important mediator of self-renewal, was recently shown to regulate

gene expression by enhancing release of paused RNAPolII in mouse ES cells (Rahl et al., 2010). Nanog might exert a similar role, since flavopiridol treatment experiments revealed that RNApolII pause release from the Esrrb promoter is significantly delayed in Nanog$^{-/-}$ ES cells. Interestingly, c-Myc was reported to bind preferentially to genes that are already active in ES cells, and to act as an amplificator of transcription (Lin et al.; Nie et al.). In this respect it is important to note that, albeit at low levels, *Esrrb* is actively transcribed in Nanog$^{-/-}$ ES cells, and Nanog activation increases its transcriptional output. Further studies should explore the possibility that Nanog might act both as a conventional transcription factor, by binding to enhancers and promoting RNApolII recruitment, and as a transcriptional amplifier, by interacting with promoters and facilitating pause release.

Chapter 4: Nanog and Esrrb fluctuations in single cells. Functional relevance of heterogeneous Nanog expression in ES cells

4.1: Creation of Nanog and Esrrb double reporter lines

The functional differences existing between cells expressing high or low Nanog levels are likely to be determined by differential expression of key target genes involved in promoting ES cell self-renewal. In this context, it is of particular interest to study the regulation that Nanog exerts on such factors in single live cells. Since Esrrb plays an important role in sustaining pluripotency and is the gene that shows the most pronounced transcriptional change in response to Nanog induction in ES cells, a series of fluorescence reporter lines were designed to permit concomitant tracking of Nanog and Esrrb expression in single cells.

As the complex structure of the gene, its length and the presence of a long 3'UTR region could be crucial factors in determining Esrrb transcriptional regulation, it was decided to minimise any alteration to the *Esrrb* gene upon introduction of exogenous sequences. Different fluorescent reporter cassettes were knocked-in at the 3' end of the last exon of *Esrrb*. In a first set of reporter constructs a TdTomato cassette was fused in frame to the end of the Esrrb coding sequence, connected through a 13aa long peptide linker (**Figure 4.1A**). The purpose of using a protein fusion was to generate a reporter line that would accurately reflect Esrrb protein levels in the cells, and could effectively detect any modulation of Esrrb expression exerted at the protein level. In addition a second set of reporters was generated in which a TdTomato protein was linked to the end of the Esrrb coding sequence by a T2a peptide, allowing for cleavage of the fluorescent reporter nascent polypeptide chain by ribosome skipping (Donnelly et al., 2001). In all targeting constructs, expression

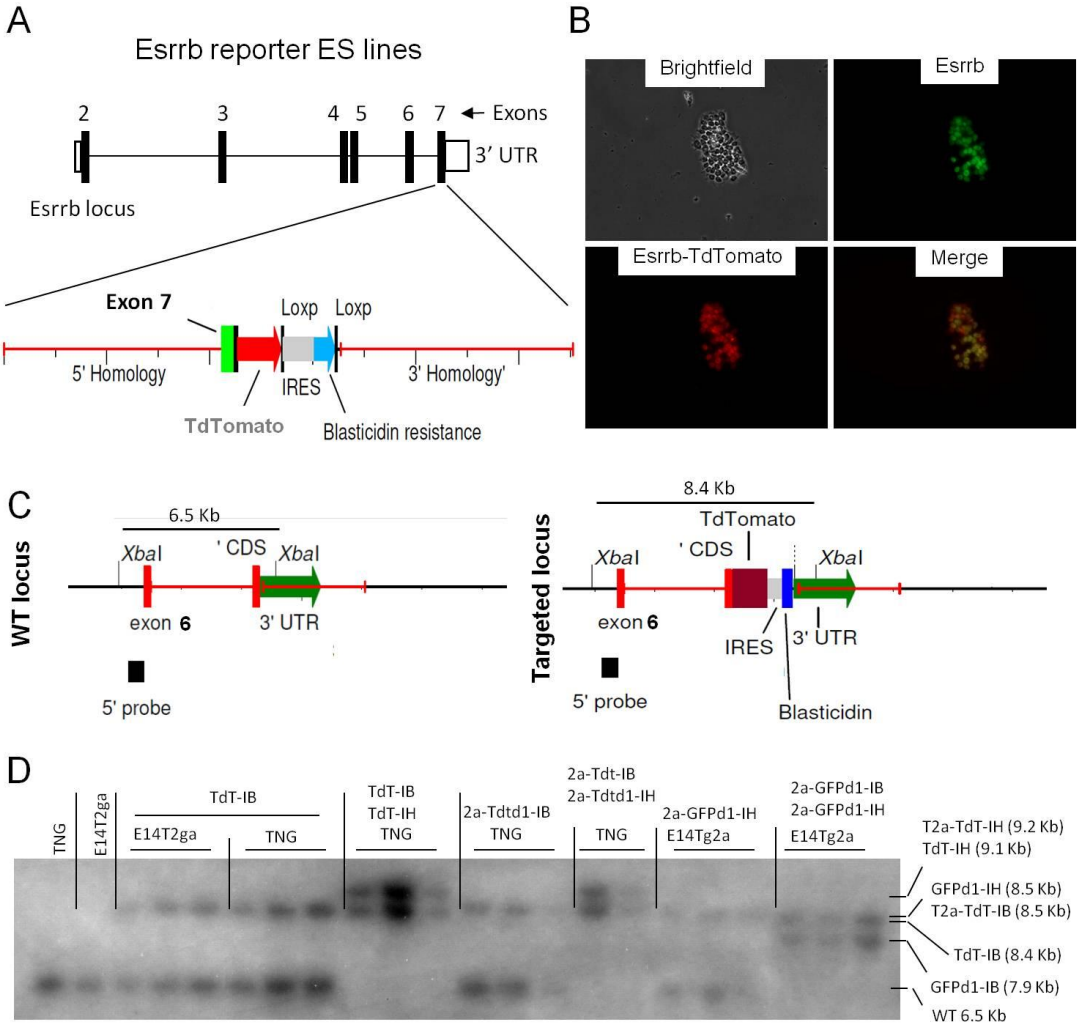


Figure 4.1: Derivation of Esrrb reporter ES cell lines.

A: Schematic representation of the Esrrb locus structure and targeting vectors used to derive E14Tg2a Esrrb-TdTomato cells. **B**: Immunohistochemical analysis of Esrrb protein expression in E14Tg2a Esrrb-TdTomato cells. **C**: Schematic representation of *Esrrb* before and after targeting. 5' and 3' homology arms are red lines, exons are red boxes, 3'UTR is in green. Restriction sites and the probe used for Southern blot analysis are indicated, as are the expected sizes of the DNA fragments obtained after digestion of wild-type and TdTomato-IRES-BSD^{R} targeted alleles. **D**: Southern blot analysis performed on DNA samples prepared from E14Tg2a, TNG, E14Tg2a Esrrb-TdTomato, TNG Esrrb-TdTomato, TNG Esrrb-2a-TdTomato and E14Tg2a Esrrb-2a-GFPdest1 ES cell lines. DNA was digested with *XbaI* and the blot hybridised to a probe binding 5' to the targeted genomic position (See panel C). Expected fragment sizes relative to the targeting of –TdTomato, 2a-TdTomato and 2a-GFPDest1 cassettes in combination with Hygromycin or Blasticidin resistance genes are reported.

of a blasticidin or hygromycin resistance gene was coupled to Esrrb expression by an IRES sequence. No polyadenylation signal was included in the vectors so that expression of functional messenger RNAs would rely on the endogenous Esrrb 3' UTR sequence (**Figure 4.1A**).

E14Tg2a and TNG ES cells were electroporated with Esrrb-TdTomato fusion protein or Esrrb-T2a-TdTomato targeting constructs including an IRES-blasticidin resistance cassette. Single clones were picked, expanded and the presence of a correctly recombined allele was assessed by Southern blot. Digestion of genomic DNA from targeted lines with *XbaI* and hybridisation with a probe annealing 5' to *Esrrb* Exon 7 should lead to the detection of a 6.5Kb band corresponding to the WT allele and an 8.4-8.5Kb band originated by insertion of the TdTomato-IRES-blasticidin cassette at the *Esrrb* stop codon (**Figure 4.1C)**. Homologous recombination occurs with high efficiency using these promoterless vectors, since all clones analysed underwent the desired recombination event (**Figure 4.1D** shows the results obtained for three clones of each line generated).

E14Tg2a Esrrb-TdTomato ES cells were plated at low density and three days later the expression of Esrrb was determined by immunohistochemistry. Under these conditions Esrrb expression is mosaic in ES cell colonies (**Figure 4.1B**). The fluorescent signal detected from the Esrrb-TdTomato fusion protein correctly reported Esrrb expression in the colonies analysed (**Figure 4.1B**).

An important point to consider while designing a reporter system able to accurately track rapid fluctuations in protein levels is whether the half-life of the fluorescent protein used matches the half-life of the endogenous protein of interest. With this aim, the half-life of Esrrb was first determined in E14Tg2a ES cells treated with cycloheximide for 1 to 12 hours, collecting total protein extracts at 2 hour timepoints. As shown in **Figure 4.2A** Esrrb protein half-life is greater than 2 hours but shorter than 4 hours in ES cells. Cycloheximide treatment experiments also showed that the

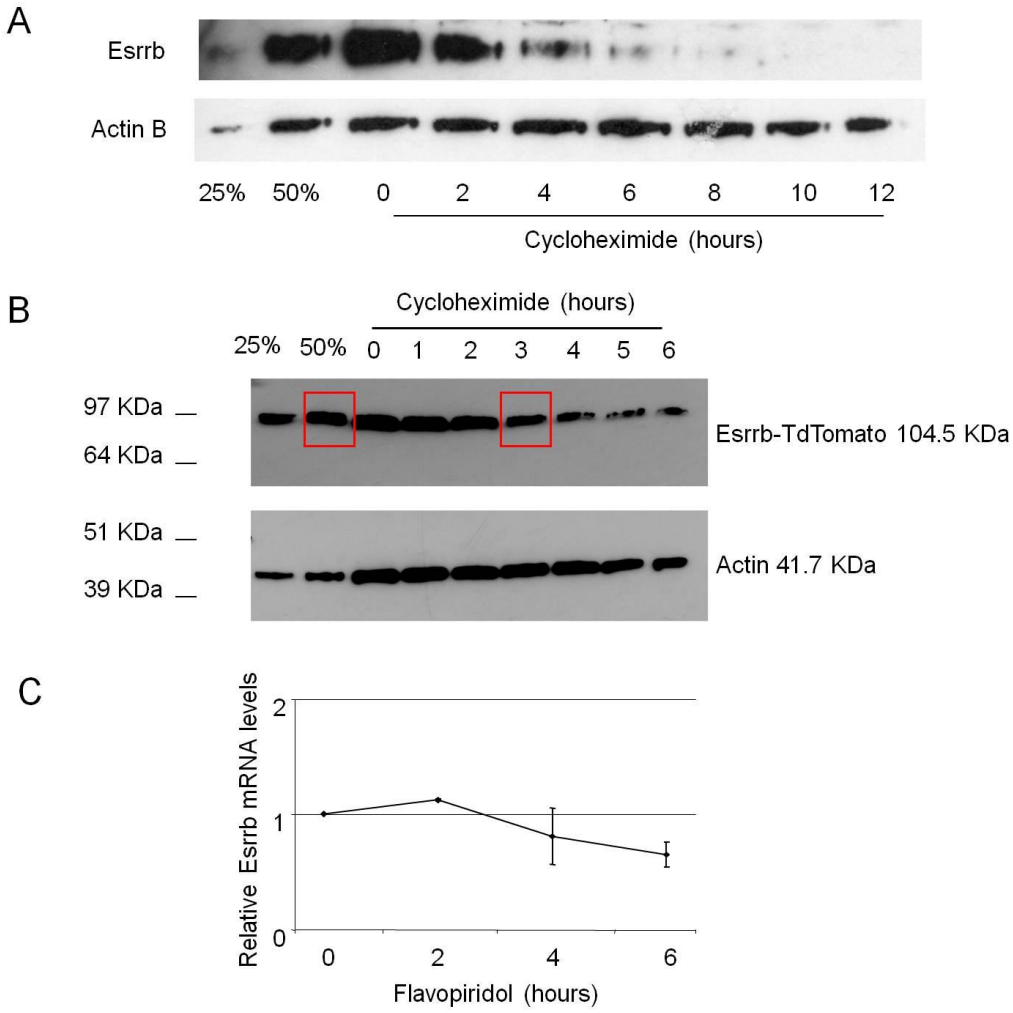


Figure 4.2: Esrrb-TdTomato fusion protein stability in ES cells.

A, B: Esrrb (A) and Esrrb-TdTomato (B) protein half-life in ES cells. E14Tg2a or E14Tg2a-TdTomato ES cells were treated with cycloheximide and total protein extracts were collected at the indicated times. The 50% and 25% of the total amount of protein used for other samples were loaded for untreated cells, and serve as an internal reference. Actin B protein levels are shown as a loading control. **F**: Esrrb pre-mRNA levels in E14Tg2a cells treated for 0, 2, 4 or 6 hours with flavopiridol. Relative transcript levels normalised to ribosomal RNAs. Error bars: standard deviation of the mean for gene expression values measured in 3 independent experiments.

half-life of Esrrb-TdTomato is shorter than 3 hours in E14Tg2a Esrrb-TdTomato cells (**Figure 2B**). It was thus possible to verify that Esrrb and the Esrrb-TdTomato fusion protein have a relatively similar half-life.

Reporter lines can also be used to detect variations in the transcriptional output from an active gene. Supposing that the introduction of exogenous sequences does not result in major alterations of the transcript, the time required to detect reductions in transcriptional activity will not only depend on the half-life of the fluorescent protein chosen but also on the stability of the messenger RNA of the gene of interest. To determine Esrrb mRNA stability, ES cells were treated with flavopiridol for 1 to 6 hours and RNA samples collected every hour. The inferred mRNA half-life in ES cells is greater than 6 hours (**Figure 4.2C**). The stability of Esrrb messenger RNA is therefore likely to affect more significantly than the Esrrb-TdTomato protein half-life any delay in the decay of the signal detected from expression of these reporters. Flavopiridol treatment experiments should be now repeated in targeted reporter lines to exclude major differences between the Esrrb and Esrrb-TdTomato mRNA stabilities.

Treatment of TNG Esrrb-TdTomato cells with blasticidin results in the selection of cells expressing uniformly high Esrrb levels. When TNG Esrrb-TdTomato ES cells were released for multiple passages from blasticidin selection, it was possible to detect a population of Esrrb-TdTomato^{-}/Nanog^{+} cells that progressively increased in frequency. When sorted, these cells never reactivated Esrrb-TdTomato irrespectively of maintaining or losing Nanog expression. Surprisingly, TdTomato^{-} cells were found positive for Esrrb protein by immunostaining (**Figure 4.3A**). Southern blot analysis determined that the progressive increase of this population was due to spontaneous recombination events that led to the repair of the TdTomato targeted allele (**Figure 4.3B**). It was therefore decided to target both Esrrb alleles in these reporter lines. TNG Esrrb-TdTomato and TNG Esrrb-T2a-Tdtomato cells were subjected to a second round of electroporation with Esrrb-TdTomato fusion protein or Esrrb-T2a-TdTomato targeting constructs respectively. In these vectors, the

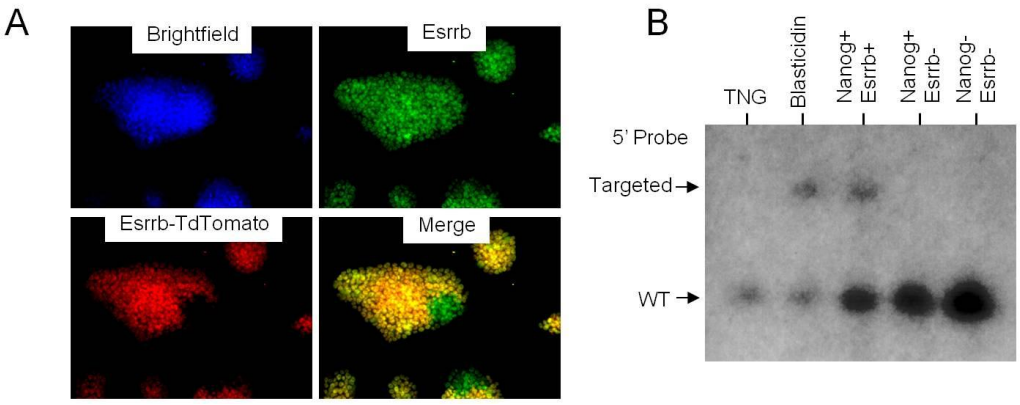


Figure 4.3: Progressive loss of the targeted allele in Esrrb-TdTomato reporter lines.

A: Immunohistochemical analysis of Esrrb expression in TNG Esrrb-TdTomato cells after long term passaging in the absence of blasticidin selection. **B:** $Esrrb^{+}/Nanog^{+}$ TNG Esrrb-TdTomato cells were released from blasticidin selection for 36 days. Southern blot analysis was performed on DNA samples prepared from $Nanog^{+}/Esrrb^{+}$ or $Nanog^{+}/Esrrb^{-}$ and $Nanog^{-}/Esrrb^{-}$ sorted populations or blasticidin selected TNG Esrrb-TdTomato cells (as a control). DNA was digested with *XbaI* and the blot hybridised to a probe binding 5' to the targeted locus (See Figure 4.2c).

IRES-blasticidin resistance cassette present in the original targeting constructs was substituted for an IRES-hygromycin resistance cassette. After electroporation, cells were selected in the presence of blasticidin and hygromycin to isolate clones that retained the previously targeted allele and recombined the second wildtype *Esrrb* allele. As previously noticed, the targeting efficiency was relatively high, since the majority of the clones analysed underwent the desired recombination event (**Figure 4.1D** shows the positive results obtained for three and two clonal lines of TNG Esrrb-TdTomato and TNG Esrrb-T2a-TdTomato cells respectively). When TNG Esrrb-TdTomato cells were selected with blasticidin and hygromycin for two passages and released in unsupplemented medium, Esrrb-TdTomato^{-}/Nanog^{+} cells were never detected, even after extensive passaging. These results suggest that the use of double targeted lines excluded complications arising from spontaneous repair of the reporter alleles. In addition, the potential confounding effects of asymmetric allelic expression are avoided in these lines.

4.2: Nanog and Esrrb expression correlates in single cells

Next, it was decided to determine how mosaic Nanog and Esrrb expression arises in ES cell colonies. TNG Esrrb-TdTomato cells were plated at low density in the presence or absence of puromycin or blasticidin and 3 days later GFP and Esrrb-TdTomato expression was measured by fluorescence microscopy. Nanog and Esrrb expression seems to be concomitantly downregulated in the population of cells residing at the margins of the colonies (**Figure 4.4A**). In parallel, cells were collected, stained for SSEA-1 and analysed by flow cytometry. Differentiated SSEA-1 negative cells were excluded from the analysis. In the absence of selection, Esrrb and Nanog expression levels directly correlate in the fraction of cells that expresses both Nanog and Esrrb (**Figure 4.4B**). Cells expressing Nanog at low levels but retaining expression of Esrrb, as well as cells that had lost expression of both markers, were also detected. Importantly, cells showing high Nanog levels were always positive for Esrrb expression.

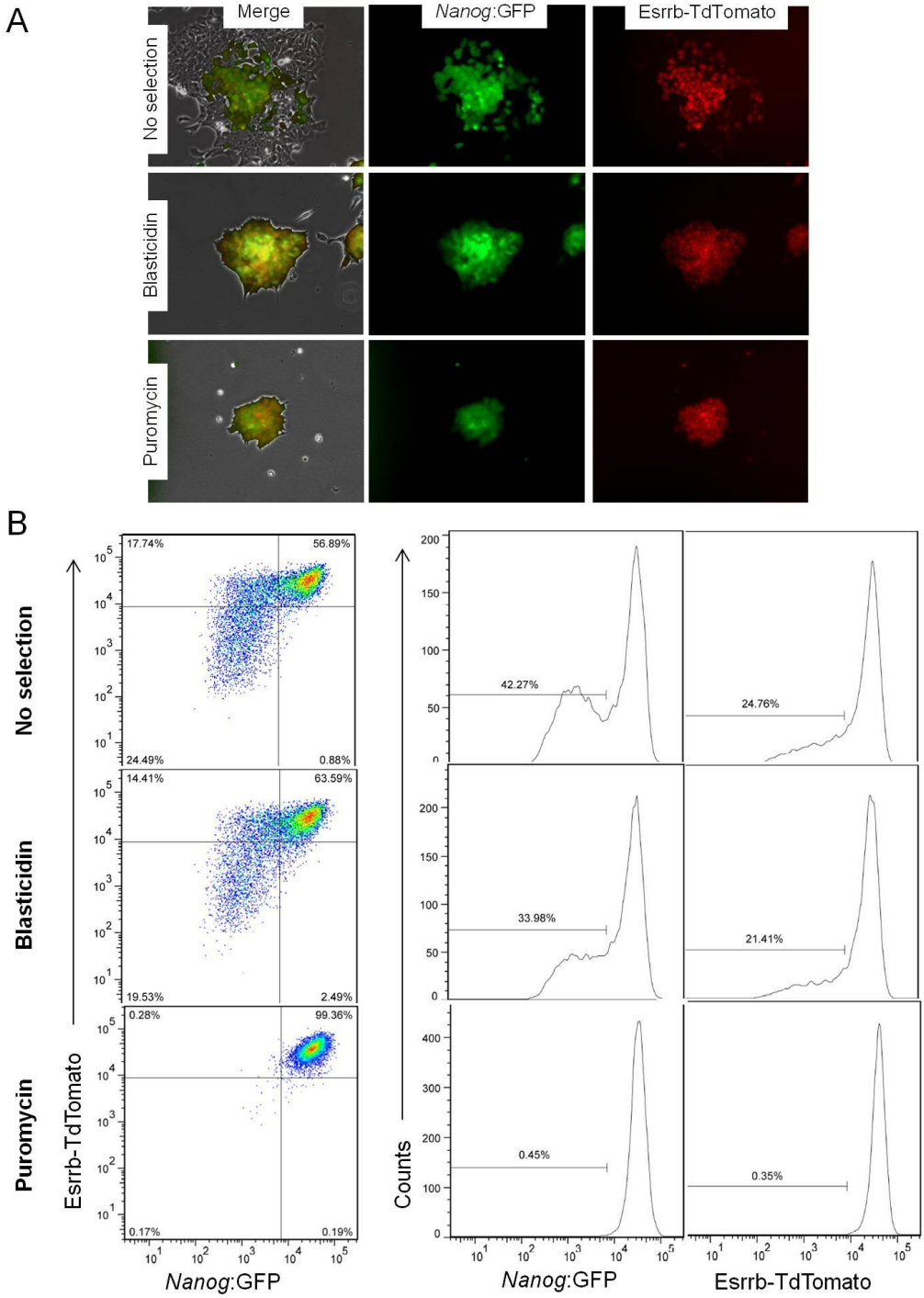


Figure 4.4: Esrrb and Nanog expression correlate in single ES cells.

A: Esrrb and Nanog fluorescent reporter expression in TNG Esrrb-TdTomato cells cultured in the presence or in the absence of blasticidin or puromycin. **D**: Dot plots and histograms showing Esrrb and Nanog fluorescent reporter expression in SSEA-1^{+} TNG Esrrb-TdTomato cells plated at low density and cultured for 3 days in the absence or in the presence of blasticidin or puromycin selection.

When cells were plated in similar conditions and blasticidin was added to the culture medium, it was not possible to stringently select for Esrrb expression ($Esrrb^+$:~80%). Nonetheless, cells grown in selection tended to form higher numbers of tight colonies completely lacking morphological signs of differentiation (**Figure 4.4A**). Selection for Esrrb expression also determined a small increase in the number of cells expressing Nanog (66% against 58%, **Figure 4.4B**). These results are probably due to the relatively long time required for blasticidin to achieve effective selection. In striking contrast, selection for Nanog transcription completely abolished heterogeneity in Esrrb expression (**Figure 4.4A-B**). This result is supported by the observation that Nanog overexpression imposes homogeneous Esrrb expression in ES cells (**Figure 4.5**). Therefore, in the transcriptional hierarchy of ES cells Esrrb seems to be placed downstream of Nanog, since heterogeneity in Nanog expression is not lost in reciprocal Esrrb overexpression experiments (**Figure 4.5**).

4.3: Esrrb downregulation requires loss of Nanog and coincides with commitment to differentiation

It was then decided to determine whether ES cells could fluctuate between states of positive and negative Esrrb expression, how fluctuation related to Nanog expression; the kinetics of interconversion between the different populations were also analysed. TNG Esrrb-TdTomato ES cells were stained for SSEA-1 and sorted into four distinct populations: SSEA-1^+ $Nanog^+/Esrrb^+$, $Nanog^-/Esrrb^+$, $Nanog^-/Esrrb^{low}$ or $Nanog^-/Esrrb^-$. Sorted populations were replated, cultured for 6 days and fluorescent reporter expression determined each day by flow cytometry (**Figure 4.6**). A fraction of $Esrrb^+/Nanog^+$ cells gradually lost expression of both markers, and reconstituted the SSEA-$1^+/Nanog^-/Esrrb^+$ and $Nanog^-/Esrrb^-$ populations. Cells could not downregulate Esrrb without concomitantly losing expression of Nanog. Similarly $Esrrb^+/Nanog^-$ cells could either regain Nanog expression and maintain Esrrb levels or lose expression of both markers. In contrast, $Nanog^-/Esrrb^-$ or $Nanog^-/Esrrb^{low}$ cells progressively differentiated over the first three days of the experiment (**Figure 4.7A**).

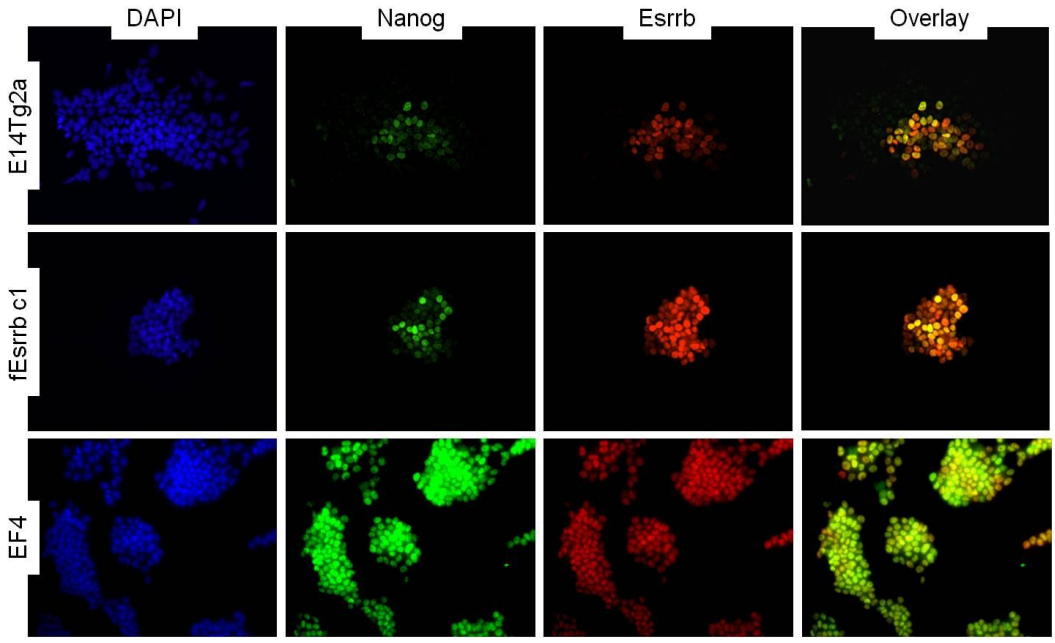


Figure 4.5: Nanog overexpression abolishes Esrrb heterogeneity.

Immunohistochemical analysis of Nanog and Esrrb expression in E14Tg2a, fEsrrb (Esrrb overexpressing) and EF4 (Nanog overexpressing) ES cells. Cells were plated at low density and cultured for three days prior to fixation and staining.

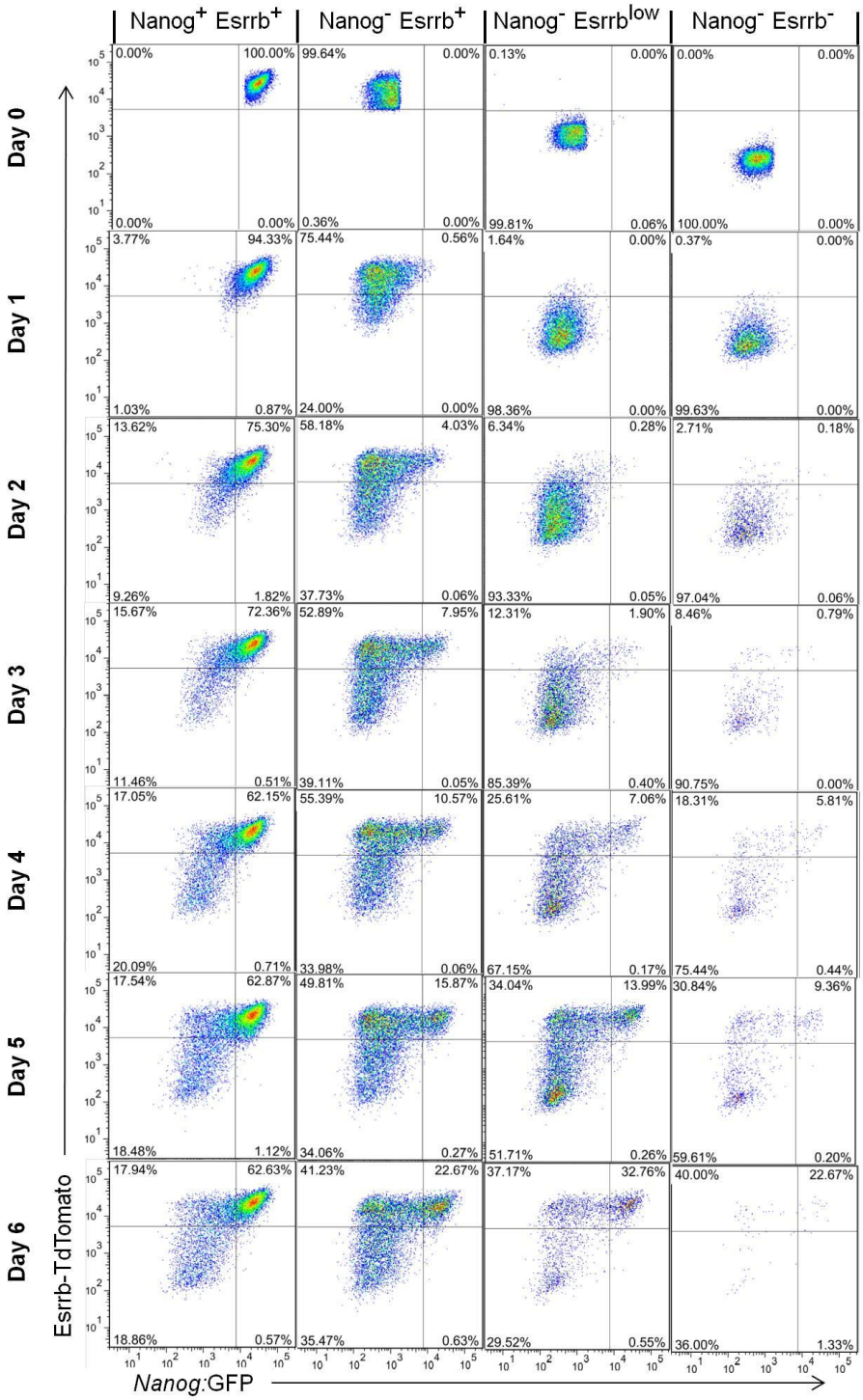


Figure 4.6: Esrrb downregulation requires loss of Nanog.

A: Dot plots showing Esrrb-TdTomato and *Nanog*:GFP expression of TNG Esrrb-TdTomato ES cells after sorting SSEA-1$^+$ Nanog$^+$/Esrrb$^+$, Nanog$^-$/Esrrb$^+$, Nanog$^-$/Esrrb^{low} or Nanog$^-$/Esrrb$^-$ cells and culture for 0 to 6 days in complete ES cell medium. Only SSEA-1$^+$ cells were selected for the analysis.

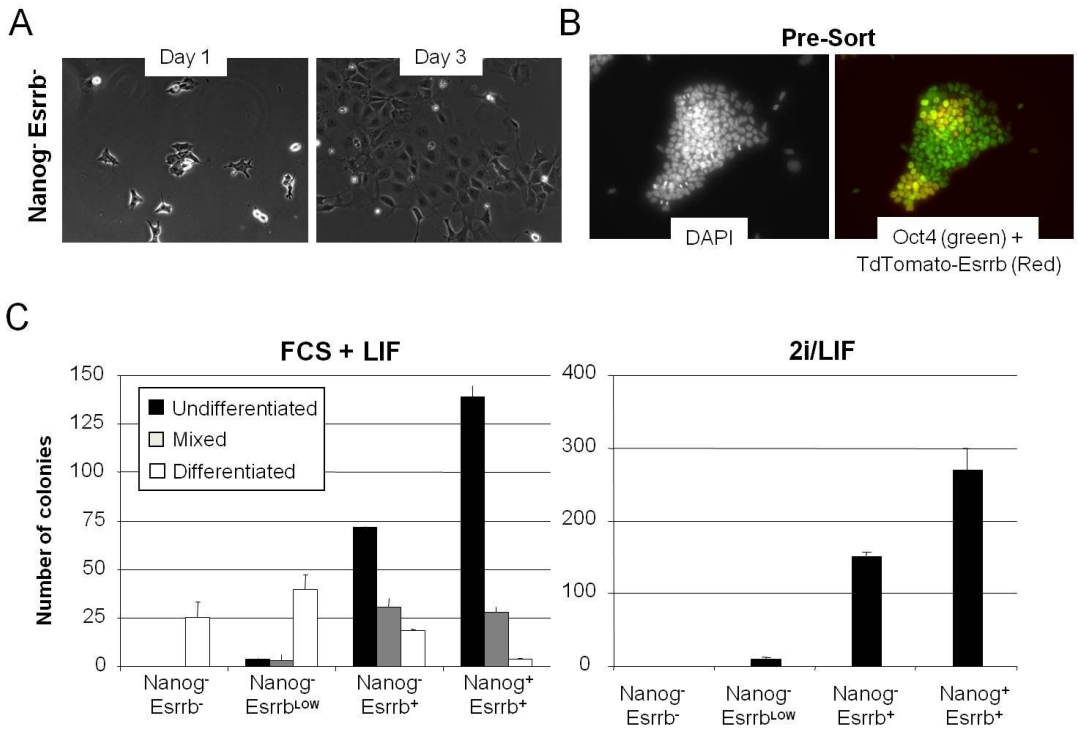


Figure 4.7: Esrrb downregulation coincides with commitment to differentiation

A: Brightfield images showing widespread differentiation of sorted Nanog$^-$/Esrrb$^-$ TNG Esrrb-TdTomato cells 3 days after replating in GMEMβ/FCS/LIF. **B**: Immunohistochemical analysis of Oct4 expression in TNG Esrrb-TdTomato cells cultured for 3 days in GMEMβ/FCS/LIF in the absence of puromycin or blasticidin selection. **C**: Numbers of alkaline phosphatase positive colonies observed 6 days after plating SSEA-1$^+$ Nanog$^+$/Esrrb$^+$, Nanog$^-$/Esrrb$^+$, Nanog$^-$/Esrrb^{low} or Nanog$^-$ /Esrrb$^-$ sorted TNG Esrrb-TdTomato cells at clonal density in GMEMβ/FCS/LIF or in 2i/LIF.

A fraction of the cells from the Nanog^{-}/Esrrb^{low} population were able to regain expression of Esrrb whereas almost none of the Nanog^{-}/Esrrb^{-} cells upregulated Esrrb, with the majority losing SSEA-1 and differentiating. To rule out the possibility that Esrrb^{-} cells were already differentiated at the time of sorting, TNG Esrrb-TdTomato cells cultures were stained for Oct4 expression. In addition to expressing high levels of SSEA-1, the majority of Esrrb^{-} cells expressed Oct4 at similar levels to Esrrb^{+} cells (**Figure 4.7B**).

To quantify the self-renewal ability of the different populations, sorted cells were plated at clonal density in GMEMβ/FCS/LIF (FCS/LIF) or in N2B27/2i/LIF (2i/LIF). Strikingly, cells lacking Esrrb expression were unable to form undifferentiated colonies in FCS/LIF and completely lost their ability to grow in 2i/LIF (**Figure 4.7C**). Taken together these data suggest that loss of Esrrb expression marks the fraction of ES cells that have committed to differentiation.

4.4: Tracking Esrrb transcription in destabilised GFP reporter lines

In order to investigate in greater detail the early commitment events following loss of Esrrb expression, it was next decided to derive an ES cell line in which expression of a destabilised GFP protein with a half-life of 1 hour reports Esrrb transcription. E14Tg2a ES cells were targeted with vectors derived from the constructs employed to generate TNG Esrrb-2a-Tdomato cells by substituting TdTomato for a destabilised form of the GFP protein (Li et al., 1998). As before, both *Esrrb* alleles were targeted with vectors harbouring blasticidin or hygromycin resistance genes to prevent loss of the reporter alleles.

When E14Tg2a Esrrb-2a-GFPdest1 cells are cultured in FCS/LIF it is possible to identify a fraction of the population that has lost Esrrb expression (**Figure 4.8A**). Confirming a connection between loss of Esrrb expression and differentiation, culture in conditions permissive only for undifferentiated cells (2i/LIF) resulted in

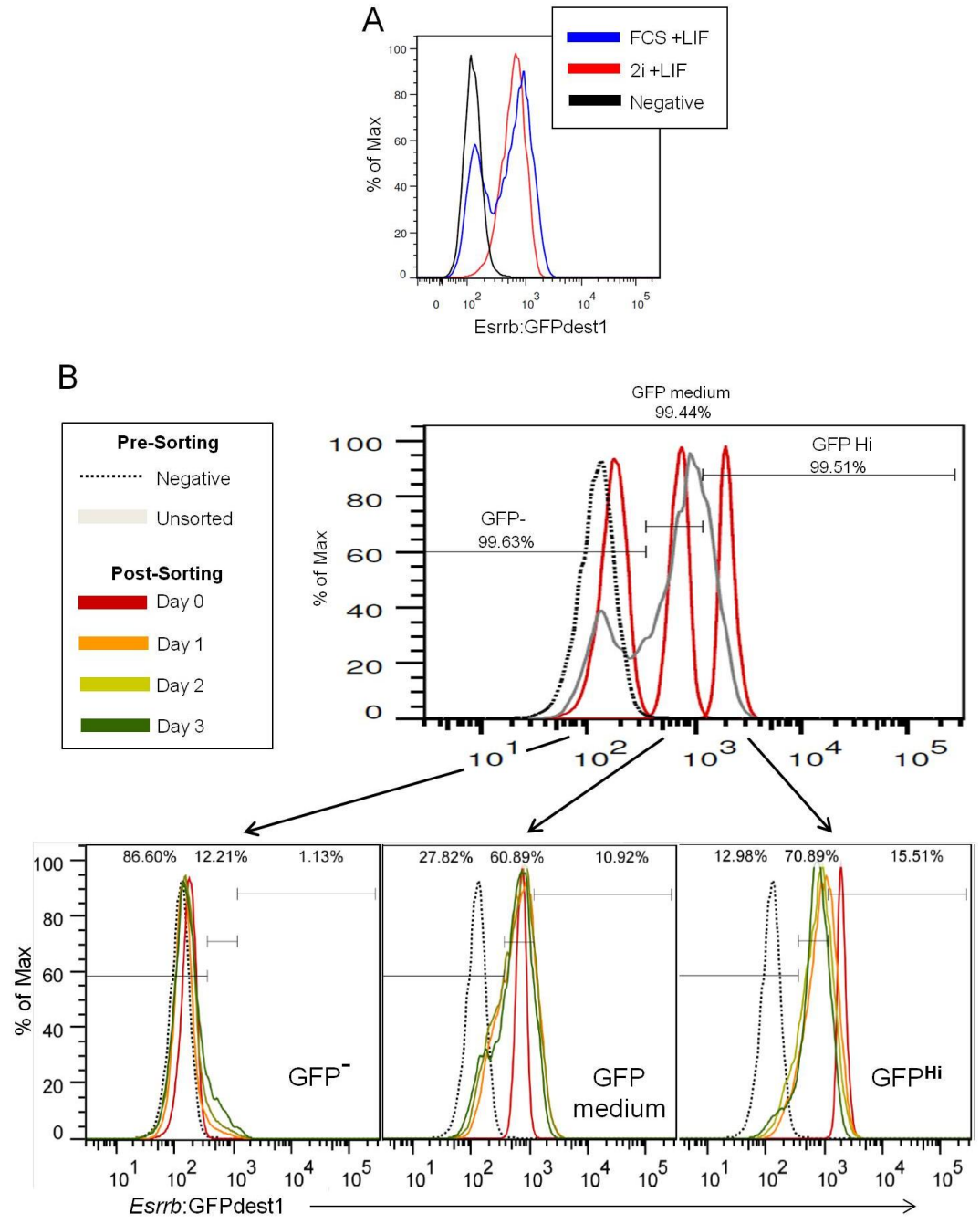


Figure 4.8: Esrrb downregulation is not irreversible.

A: Histogram showing GFP expression in SSEA-1^{+} E14Tg2a Esrrb-2a-GFPdest1 ES cells cultured in 2i/LIF or GMEMβ/FCS/LIF. **B**: Histogrmas showing GFP expression detected 0, 1, 2 or 3 days after plating sorted SSEA-1^{+} GFP^{hi}, GFP^{medium} or GFP^{-} E14Tg2a Esrrb-2a-GFPDest1 cells in GMEMβ/FCS/LIF. The purity of the three sorted population is shown in the top plot along with the GFP expression profile of unsorted cells. The percentage of cells falling in the GFP^{hi}, GFP^{medium} or GFP^{-} gates 3 days after sorting is shown for each population in the bottom plots.

homogeneous expression of GFP in the ES cell population (**Figure 4.8A**). Conversely, release of the cells from blasticidin or hygromycin selection leads to the redistribution of GFP expression and allows the identification of cells that have recently downregulated Esrrb expression. E14Tg2a Esrrb-2a-GFPdest1 ES cells released from selection for 3 days were sorted in three distinct populations (SSEA-1^{+} and GFP^{hi}, GFP^{medium} or GFP^{-}) and cultured for 3 days in FCS/LIF (**Figure 4.8B**). GFP^{neg} cells showed increased signs of differentiation compared to the other populations, but maintained a fraction of morphologically undifferentiated cells that slowly regained Esrrb expression (12% of the population at Day3). GFP downregulation was faster in the GFP^{medium} compared to the GFP^{hi} population, but both populations reconstituted a sizable negative fraction after 1 day in culture. In agreement with previous results, sorted GFP^{neg} cells showed a strongly reduced ability to form undifferentiated colonies in complete ES cell medium, while both GFP^{medium} and GFP^{hi} cells showed similar self-renewal potential. (**Figure 4.9A**). GFP^{neg} cells completely lost their ability to form colonies in 2i/LIF, and a correlation between the number of colonies formed and the levels of GFP expression was observed in GFP^{medium} and GFP^{hi} cells plated in these conditions (**Figure 4.9A**).

In parallel, it was decided to analyse the gene expression profile of the three sorted populations. All populations retained expression of Oct4, and Sox2 expression was only slightly reduced in GFP^{neg} cells, in line with their undifferentiated state at the time of sorting (**Figure 4.9B**). Analysis of Esrrb mRNA levels confirmed the ability of the reporter cell line to accurately track Esrrb expression. Interestingly, a correlation between the expression of markers characteristic of a pristine undifferentiated state and Esrrb expression was observed, with complete loss of ES cell specific markers such as Rex1, Klf4, Dax1 and Nr5a2 in the GFP^{neg} population (**Figure 4.9B**). Nanog was the gene that matched more closely Esrrb levels in the cells. Conversely, loss of Esrrb coincided with upregulation of genes characteristic of cell lines derived from embryos at later stages of development (EpiSC) (Brons et al., 2007; Tesar et al., 2007) (**Figure 4.9B**).

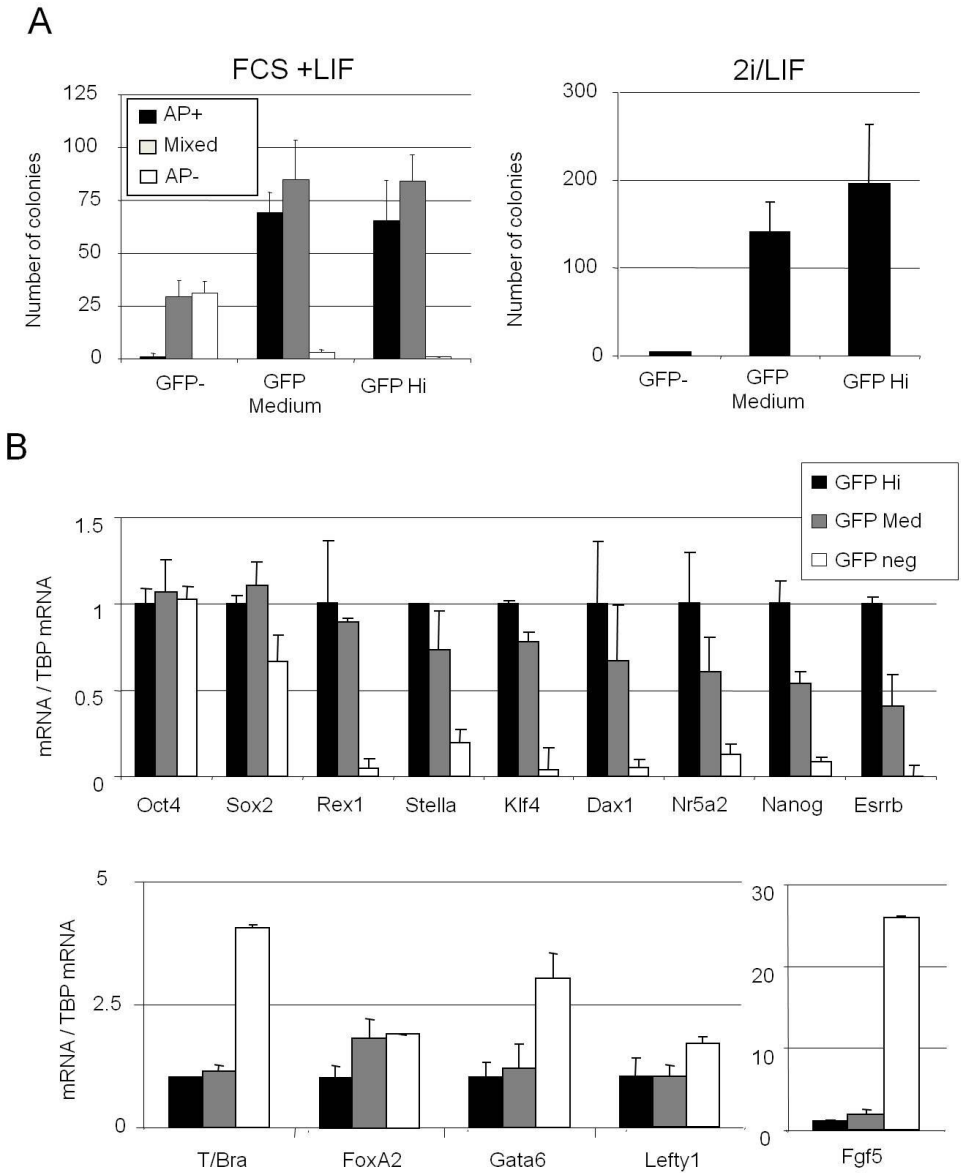


Figure 4.9: Esrrb downregulation marks progressive loss of naïve pluripotency.

A: Numbers of alkaline phosphatase positive colonies observed 6 days after plating SSEA-1^{+} GFP^{hi}, GFP^{medium} or GFP^{-} sorted E14Tg2a Esrrb-2a-GFPdest1 cells at clonal density in GMEMβ/FCS/LIF or in 2i/LIF. **B**: Gene expression profile of sorted SSEA-1^{+} and GFP^{hi}, GFP^{medium} or GFP^{-} E14Tg2a Esrrb-2a-GFPdest1 ES cells. Error bars: standard deviation of gene expression values measured in three independent experiments.

Taken together these results indicate that loss of Esrrb expression in the Nanog$^-$ population of self-renewing ES cells marks commitment to differentiation. Cell fate is progressively locked after downregulation of Esrrb. Cells immediately lose their ability to be cultured in non-permissive conditions (2i/LIF) but retain for a short time the capacity to regain Esrrb expression and inefficiently self-renew in permissive medium (FCS/LIF). The loss of clonogenic ability in non-permissive culture conditions mirrors the dowregulation of ground state pluripotency markers and the upregulation of early differentiation markers that is then completed in cells that have completely silenced Esrrb expression.

4.5: Nanog null Esrrb reporter lines

The previous results showed that Nanog and Esrrb expression correlate in single cells. It was also possible to determine that forced Nanog expression abolishes Esrrb heterogeneity, and that loss of Esrrb requires the downregulation of Nanog. Based on these observations, it was important to explore the effects of Nanog deletion on the dynamics of Esrrb expression at the single cell level.

With this aim, T2a-TdTomato reporter cassettes were knocked in at the stop codon of both *Esrrb* alleles in ESΔN-NERT cells (See chapter 3 for a description of the line). Successful recombination was verified by quantitative PCR on genomic DNA using primer pairs amplifying exclusively exon 7 of WT *Esrrb* alleles. In these reactions, the forward primer anneals to the last portion of the *Esrrb* coding sequence and the reverse primer anneals at short distance in the 3'UTR of the gene. In targeted lines insertion of a 3Kb long sequence between *Esrrb* exon 7 and the 3'UTR prevents productive amplification. Amplification using primer pairs binding *Esrrb* exon6, a position not modified by the targeting event, serves to normalise for variations in the amount of genomic DNA loaded in each PCR reaction, and gives consistent positive amplification in all samples analysed. This system was first validated using genomic DNA from different TNG Esrrb-TdTomato reporter lines for which successful

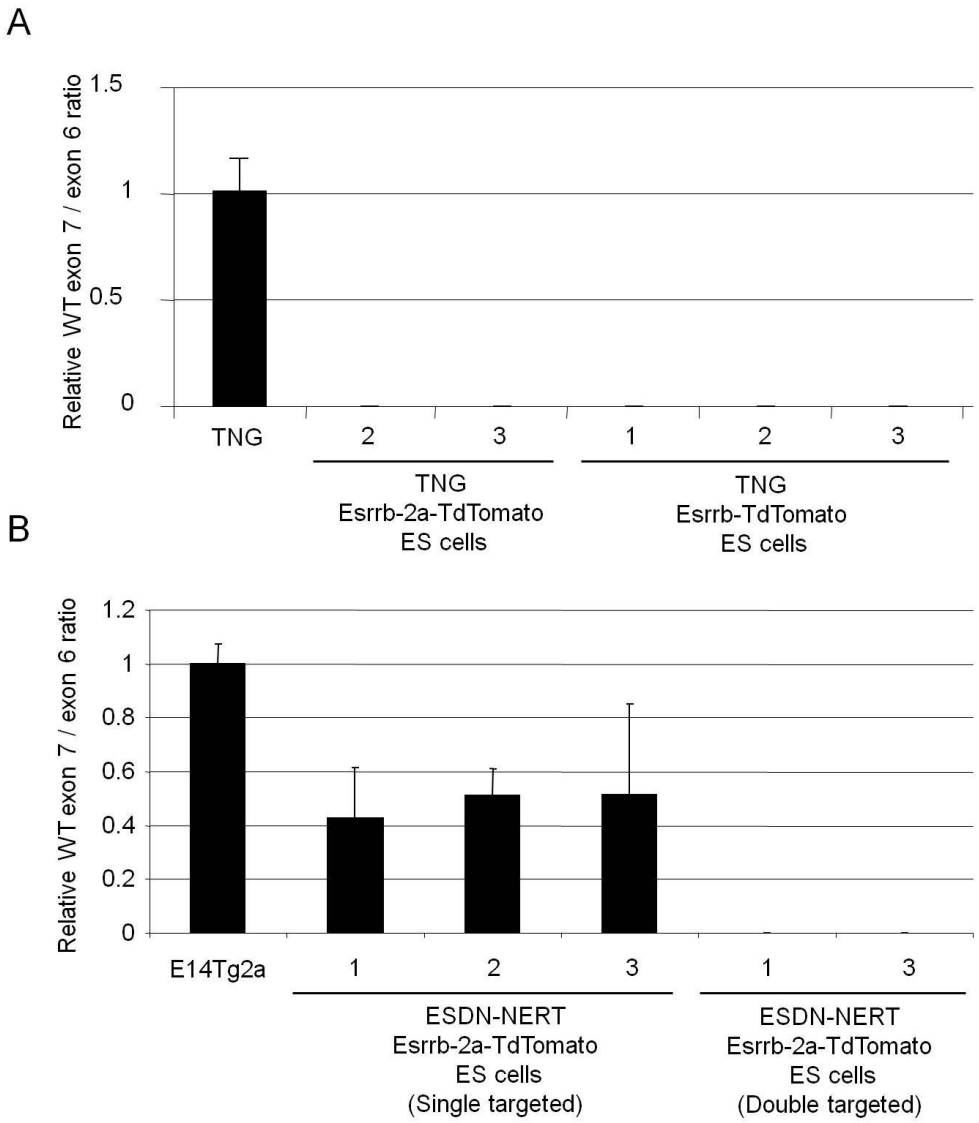


Figure 4.10: Genomic qPCR confirms correct recombination in ESDN-NERT Esrrb-2a-TdTomato ES cells.

A: Quantitative PCR analysis of the relative levels of WT *Esrrb* exon 7 to *Esrrb* exon 6 in TNG, TNG Esrrb-TdTomato and TNG Esrrb-2a-TdTomato ES cells. Error bars are the standard deviation of the 3 technical replicates of one experiment. **B**: Quantitative PCR analysis of the relative levels of WT *Esrrb* exon 7 to *Esrrb* exon 6 in E14Tg2a, ESΔN-NERT Esrrb-2a-Tdtomato (single targeted) and ESΔN-NERT Esrrb-2a-Tdtomato (Double targeted) ES cells. Error bars are the standard deviation of the 3 technical replicates of one experiment.

recombination had been confirmed by Southern blot. Amplification from WT *Esrrb* exon 7 could be readily detected for TNG parental cells, whereas no signal was observed in three independent double targeted lines (**Figure 4.10A**). Next, successful recombination was assessed in ESΔN-NERT targeted lines. Amplification of WT *Esrrb* exon 7 could be detected in E14Tg2a ES cells, was reduced in ESΔN-NERT cells after the first targeting event and was absent in double targeted ESΔN-NERT Esrrb-2a-Tdtomato ES cells (**Figure 4.10B**).

Once correct targeting had been confirmed, TNG Esrrb-2a-TdTomato and ESΔN-NERT Esrrb-2a-TdTomato cells were replated in parallel at low density in the absence of any selection and three days after cells were collected, stained for SSEA-1 expression and analysed by flow cytometry. Since both TNG Esrrb-T2a-TdTomato and ESΔN-NERT Esrrb-2a-TdTomato cells harbour identical Nanog:GFP and Esrrb-2a-TdTomato reporter alleles, it was possible to compare how transcription from *Nanog* and Esrrb levels correlate in the presence or absence of functional Nanog protein in the cells. ESΔN-NERT cells expressed overall reduced levels of Esrrb protein compared to TNG cells (**Figure 4.11A-B**). The downregulation of *Nanog*:GFP reporter expression, readily occurring in TNG cells released from selection, was also compromised in this cell line, confirming the finding that Nanog repression of its own transcription plays an important role in establishing ES cells heterogeneity ((Navarro, 2012) and see chapter 3). Strikingly, in the *Nanog*:GFP^{+}/$Esrrb^{+}$ fraction of ESΔN-NERT Esrrb-2a-TdTomato cells the correlation between Nanog transcriptional output and Esrrb expression was almost completely lost (**Figure 4.11A**). In addition, opposite to what observed in TNG cells, Esrrb downregulation occurred in ESΔN-NERT cells irrespectively of the levels of *Nanog*:GFP expression (**Figure 4.11A**). These results indicate that the correlation between Esrrb and Nanog expression observed in TNG Esrrb-2a-TdTomato cells is driven by the different levels of functional Nanog protein present in single ES cells.

The effects of inducing Nanog nuclear relocalisation was then tested in ESΔN-NERT Esrrb-2a-TdTomato cells. Confirming the importance of Nanog in regulating Esrrb

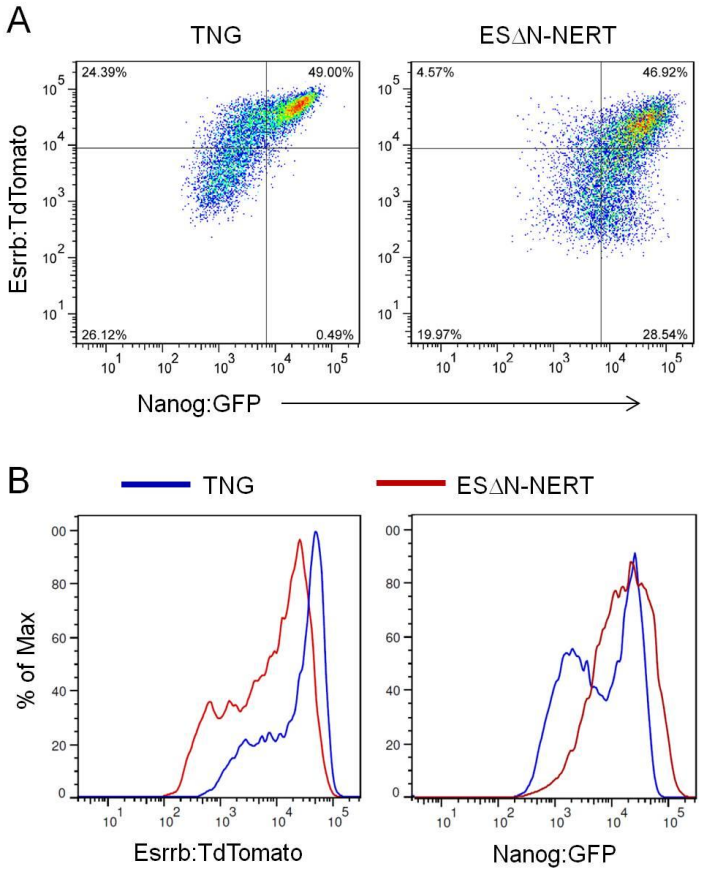


Figure 4.11: Correlation between Nanog transcription and Esrrb levels is lost in Nanog null ES cells.

A,B: Dot plots (A) and histograms (B) showing expression of Esrrb-2a-TdTomato and *Nanog*:GFP in undifferentiated SSEA-1^{+} TNG or ESDN-NERT Esrrb-2a-Tdtomato ES cells released for three days from puromycin and blasticidin (TNG) or G418 and blasticidin (ESΔN-NERT) selections.

expression, fluorescent output from the Esrrb-TdTomato alleles was significantly increased in ESΔN-NERT cells treated with tamoxifen. Esrrb expression in tamoxifen treated cells was also significantly more homogeneous (**Figure 4.12A**).

Next, the amount of time required for fluctuations in Nanog expression to results in alteration of Esrrb protein levels was determined. A precise quantification of this delay can significantly contribute to the understanding of how fluctuations in one pluripotency factor influence the levels of other ES cells regulators and result in functional differences among single cells in a genetically identical population. With this aim, the time required to elevate Esrrb protein levels was determined in ESΔN-NERT Esrrb-2a-TdTomato cells by monitoring the increase in fluorescence signal from the TdTomato protein in tamoxifen treatment timecourse experiments. Since the T2a linked Esrrb and TdTomato proteins are synthesised in equimolar amounts, any increase in TdTomato fluorescence should be linked to an equivalent increase in Esrrb protein in this cell line. The first increase in TdTomato protein levels was detected 3-4 hours after Nanog induction (**Figure 4.12B**). TdTomato accumulated steadily until 20 hours after addition of tamoxifen and reached plateau levels by 24 hours. A 2 fold difference in fluorescence levels was observed in treated versus untreated ESΔN-NERT Esrrb-2a-TdTomato cells (**Figure 4.12B**). These data suggest that any elevation in Nanog levels spontaneously occurring in a single ES cell would require around 22 hours to result in full induction of Esrrb. In this system, the time required to detect the first increase in TdTomato fluorescence is likely to report precisely when Esrrb protein levels are first increasing. In contrast, it is possible that the different stabilities of TdTomato and Esrrb result in discrepancies in the time at which increasing Esrrb and TdTomato protein levels reach plateau. Immunoblot analysis of protein extracts from ESΔN-NERT Esrrb-2a-TdTomato cells treated with tamoxifen for increasing time periods could solve this limitation. In any case, it would be interesting to perform time-lapse fluorescence microscopy on ESΔN-NERT Esrrb-2a-TdTomato cells treated with tamoxifen to confirm that a 3-4 hour delay in Esrrb induction is detected after Nanog induction in single ES cells.

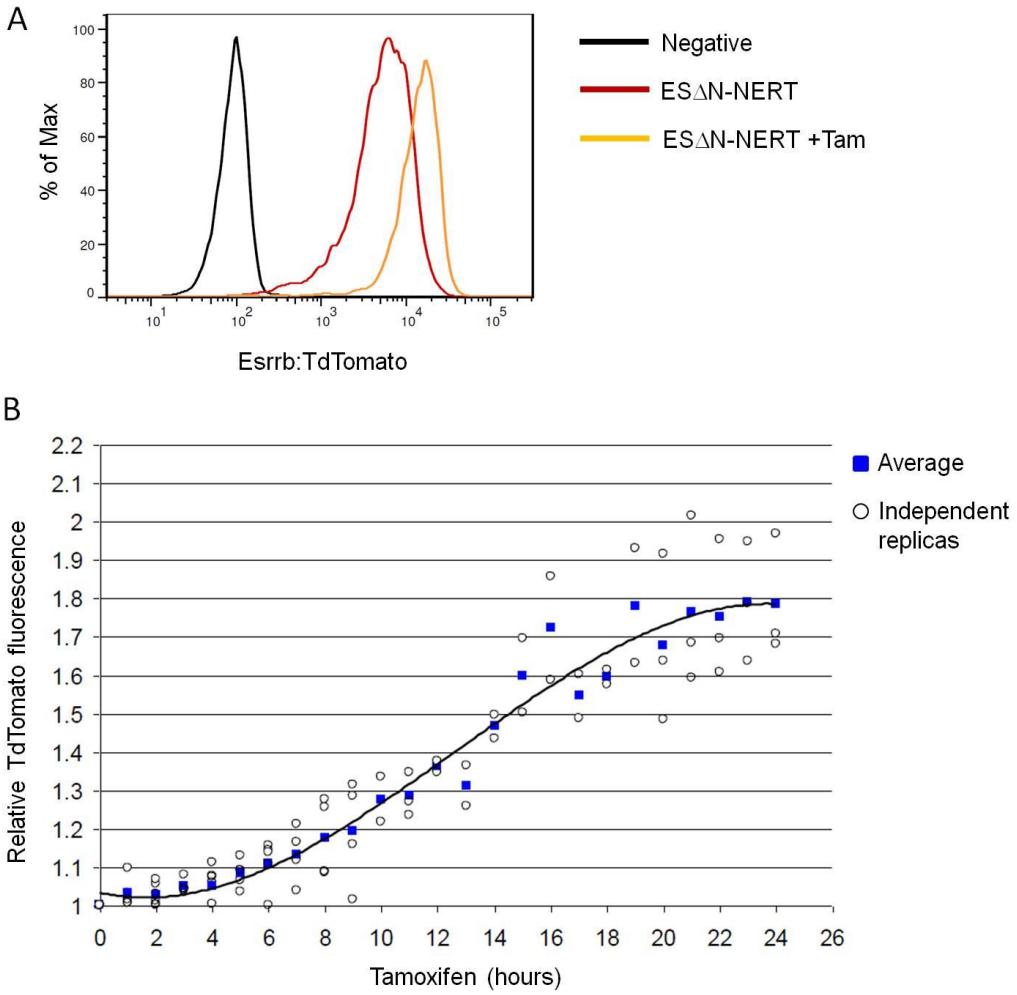


Figure 4.12: Nanog nuclear relocalisation in NERT cells results in the steady increase of Esrrb protein levels.

A: Histogram showing the expression of Esrrb-2a-TdTomato and Nanog:GFP in undifferentiated SSEA-1^{+} ESΔN-NERT Esrrb-2a-Tdtomato ES cells maintained stimulated or not stimulated with tamoxifen for 24 hours. **B**: Graph showing the mean Esrrb-TdTomato fluorescence intensity detected after stimulating ESΔN-NERT Esrrb-2a-Tdtomato ES cells with tamoxifen for the indicated time. Values are relative to untreated cells. Blue squares are the average of 2-3 independent replicas (open circles).

4.6: Esrrb transcriptional response to 2i/LIF

It was possible to show that culture of E14Tg2a Esrrb-2a-GFPdest1 cells in condition permissive only for completely undifferentiated cells (2i/LIF) results in homogeneous Esrrb expression (**Figure 4.8A**). Since Nanog expression is also upregulated by culture in 2i/LIF, it was decided to investigate whether the observed effect on Esrrb expression was mediated through upregulation of Nanog.

ESΔN-NERT or TNG Esrrb-2a-TdTomato cells were replated and passaged four times in 2i/LIF before staining for SSEA-1 expression and analysis by flow cytometry. ESΔN-NERT Esrrb-2a-TdTomato cells kept in FCS/LIF and released from selection were analysed in parallel. Unexpectedly, SSEA-1 expression declined pronouncedly in cells cultured in 2i/LIF. Nonetheless, a fraction of the cells remained SSEA-1^{+} and was selected for further analysis. The levels of *Nanog*:GFP expression detected in ESΔN-NERT Esrrb-2a-TdTomato cells cultured in 2i/LIF were higher than in TNG Esrrb-2a-TdTomato cells (**Figure 4.13B**), indicating that autorepression limits Nanog transcriptional induction under these conditions. Culture of Nanog null ES cells in 2i/LIF resulted in homogeneous and elevated levels of Esrrb expression compared to cells maintained in FCS/LIF (**Figure 4.11A and 4.13A**), suggesting that inhibition of ERK and GSK signalling promotes Esrrb expression independently of Nanog. Nonetheless, TdTomato fluorescence levels in ESΔN-NERT Esrrb-2a-TdTomato cells maintained in 2i/LIF were not identical to those observed in TNG Esrrb-2a-TdTomato cells cultured in identical conditions (**Figure 4.13B**). This raised the hypothesis that the effects of 2i/LIF could be mediated by both a Nanog dependent and independent mechanism. To test this hypothesis, ESΔN-NERT Esrrb-2a-TdTomato cells cultured in 2i/LIF were passaged in the presence of tamoxifen to induce Nanog function. Indeed, treatment of ESΔN-NERT Esrrb-2a-TdTomato cells with tamoxifen resulted in a drastic increase in the levels of TdTomato expression, surpassing the expression levels detected in TNG Esrrb-2a-TdTomato cells (**Figure 4.13A-B**). This indicates that the transcriptional output of *Esrrb* does not reach saturating levels in response to 2i/LIF and that Nanog is required to achieve full activation.

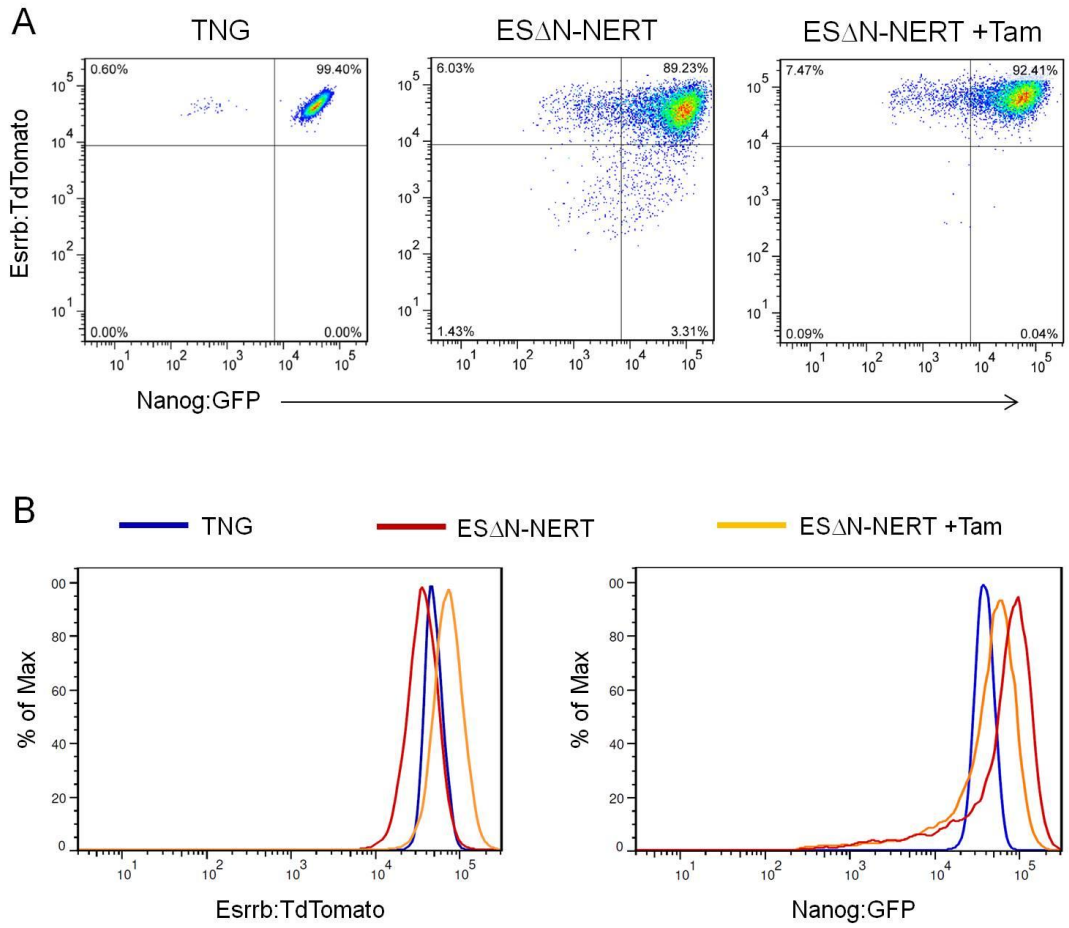


Figure 4.13: Functional Nanog is required for full efficacy of 2i/LIF culture conditions.

A,B: Dot plots (A) and histograms (B) showing expression of Esrrb-2a-TdTomato and *Nanog*:GFP in undifferentiated SSEA-1^{+} TNG or ESΔN-NERT Esrrb-2a-Tdtomato ES cells maintained in 2i/LIF in the presence or absence of tamoxifen.

Understanding whether this observation holds true for other targets of Nanog activity will help clarifying the importance of Nanog in mediating the response to ERK and GSK inhibition. In this respect, Nanog activation resulted in a pronounced reduction of *Nanog*:GFP expression in ESΔN-NERT cultured in 2i/LIF (**Figure 4.13B**), indicating that Nanog might be important in preventing inappropriate expression of its negative targets in ground state culture conditions.

4.7: Determination of pluripotency transcription factor expression by flow cytometry.

In order to be able to extend the scope of these observations to other important mediator of pluripotency for which reporter lines are not available, it was decided to adapt immunofluorescence techniques for use in flow cytometry.

The method employed (Festuccia and Chambers, 2011) is based on a fixation step that couples incubation in a diluted 0.25% PFA PBS solution with transfer to a 70% methanol solution in PBS at 4°C (See also Chapter 2 for a detailed description of the method). This fixation procedure permits to maximise the signal to background ratio and ensures consistency in the results generated from independent samples. First, it was necessary to test whether Nanog and Oct4 expression could be reliably detected with this technique. With this aim, Nanog overexpressing (EF4) and Nanog null (RCNβH(t)) cells were stained in parallel, and the distributions of fluorescence values detected by flow cytometry were compared. A homogeneous population of cells with similar size and granularity was readily distinguishable from cellular debris after fixation, indicating that cellular integrity is preserved during the permeabilisation and staining procedure (**Figure 4.14A**). Indeed, it was possible to distinguish background fluorescence detected in RCNβH(t) cells from the staining signal observed in EF4 cells (**Figure 4.14B**). Staining of the two lines gave rise to completely non-overlapping fluorescence distributions, showing that this methodology has the potential to detect Nanog expression in ES cells.

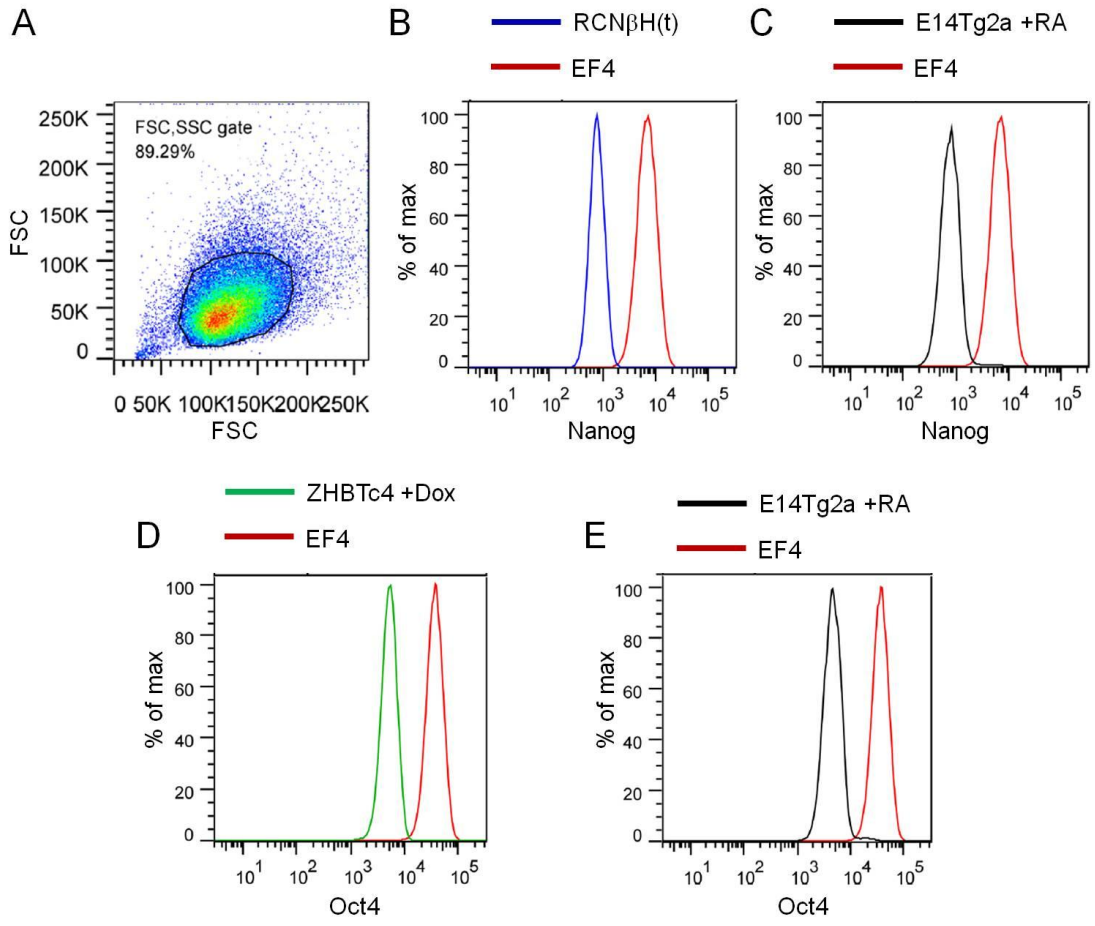


Figure 4.14: Detection of Nanog and Oct4 expression in ES cells by flow cytometry.

A: Dot plot showing FSC (size) and SSC (granularity) parameters of the populations analysed. **B**: Histogram showing the distributions of Nanog staining fluorescence values in Nanog null (RCNβH(t)) and Nanog overexpressing (EF4) ES cells. **C**: Histogram showing the distributions of Nanog staining fluorescence values in EF4 and in E14Tg2a cells differentiated by culture in the presence of 10^{-6}M retinoic acid and absence of LIF for 3 days (E14Tg2a +RA). **D**: Histogram showing the distributions of Oct4 staining fluorescence values in EF4 cells and ZHBTc4.1 cells depleted of Oct4 by culture in 1ug/ml doxycycline for 24h (ZHBTc4.1). **E**: Histogram showing the distributions of Oct4 staining fluorescence values in EF4 and differentiated E14Tg2a cells (E14Tg2a +RA).

The validity of this staining method was confirmed by comparison of EF4 and E14Tg2a differentiated by culture in the presence of retinoic acid (**Figure 4.14C**).

Next, ZHBTc4 ES cells treated with tamoxifen for 24 hours were stained for Oct4 expression and compared to untreated cells. Since tamoxifen treatment of ZHBTc4 cells results in complete loss of Oct4 protein in as short as 12 hours (Niwa et al., 2000; van den Berg et al., 2008), these cells were found ideal to test the ability of this technique to detect variations in Oct4 expression in two otherwise identical cell populations. Treatment of ZHBTc4 cells with tamoxifen resulted in a radical shift in fluorescence levels detected after staining for Oct4 (**Figure 4.14D**). Loss of Oct4 expression was also detected in differentiated E14Tg2a cells (**Figure 4.14E**).

It was then decided to assess whether the sensitivity of this technique permitted the detection of the fine variations in Nanog levels occurring in an ES cell population. E14Tg2a were concomitantly stained for Oct4 and Nanog expression. EF4, RCNβH(t), and E14Tg2a differentiated by exposure to retinoic acid provided reliable controls (**Figure 4.15A-B-D**). Staining of E14Tg2a cells identified a fraction of the population that downregulated Oct4 expression and lost Nanog signal (**Figure 4.15C**). It was thus possible to exclude differentiated cells and select $Oct4^{+}$ cell for the further analysis of Nanog levels. Consistent with its heterogeneous expression, a broad range of fluorescence values was detected after Nanog staining of E14Tg2a cells (**Figure 4.15D**). Importantly the Nanog levels detected fell between the positive and negative boundaries set by EF4 and RCNβH(t) cells. These results establish that transcription factor expression can be accurately quantified by coupling immunostaining and flow cytometry.

It was next determined whether the described technique can be used in cell expressing fluorescence reporters of gene expression. TNG cells were stained for Oct4 and Nanog expression (**Figure 4.16A**) and Nanog staining and GFP signals compared in $Oct4^{+}$ undifferentiated cells. Nanog protein levels detected by staining correlated well with *Nanog*:GFP reporter expression (**Figure 4.16B**), with 61 and 62% of Nanog positive cells detected respectively (**Figure 4.16C-D**).

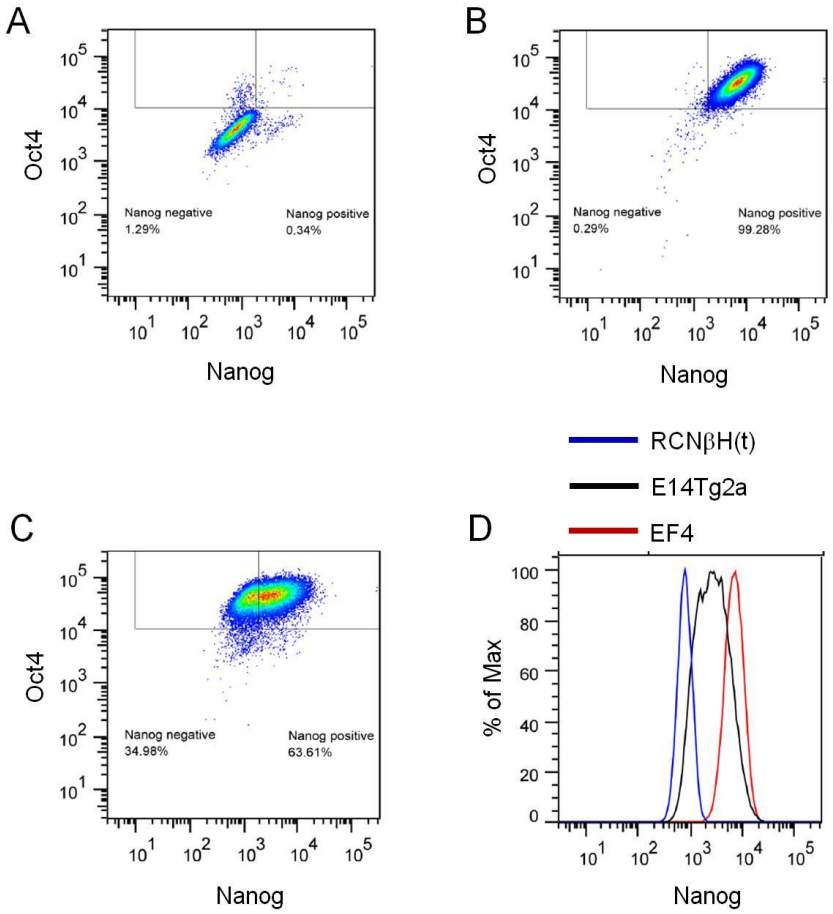


Figure 4.15: Quantification of Nanog and Oct4 expression in E14Tg2a cells.

A,B: Dot plots showing Nanog and Oct4 expression in E14Tg2a cells differentiated by culture in the presence of 10^{-6}M retinoic acid and absence of LIF for 3 days (A) and in EF4 cells (B) . **C**: Dot plots showing Nanog and Oct4 expression in E14Tg2a ES cells. **D**: Histogram plot comparing the distribution of Nanog protein levels E14Tg2a, RCNβH(t) and EF4 cells.

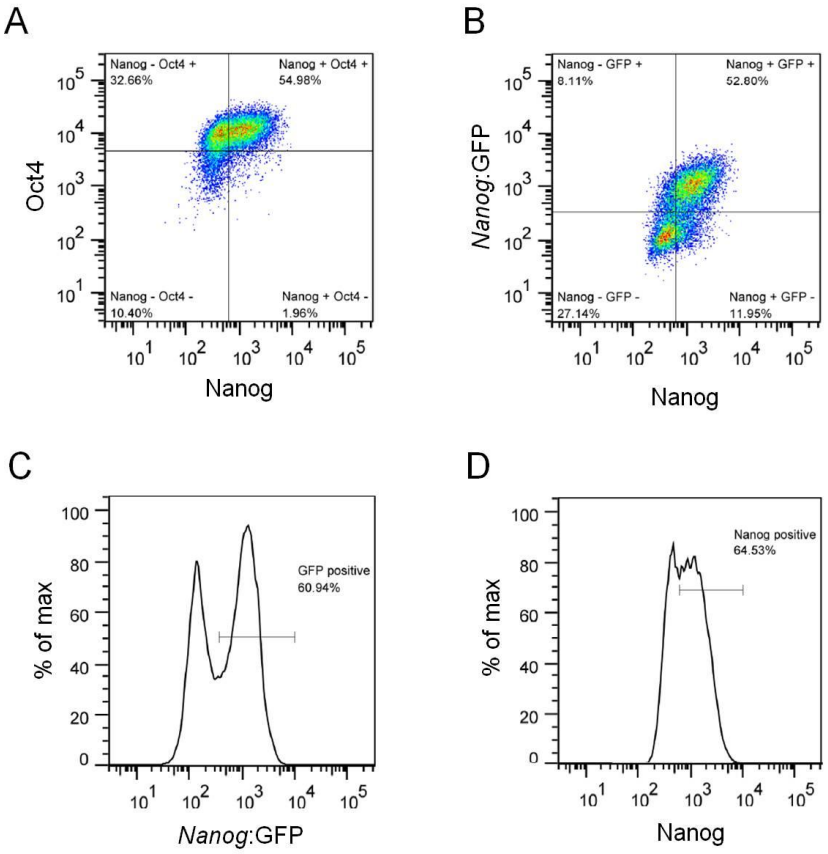


Figure 4.16: Quantification of Nanog and Oct4 levels coupled with use of fluorescent reporters.

A: Dot plot showing Nanog and Oct4 expression in TNG cells. **B**: Dot plot showing the correlation between Nanog protein and *Nanog*:GFP expression in TNG cells. **C,D**: Histograms showing the distribution of *Nanog*:GFP (C) and Nanog protein (B) levels in TNG cells.

These results provided a solid basis for applying flow cytometry to the study of heterogeneous transcription factor expression in ES cells. It was thus decided to extend the results generated analysing Nanog control over Esrrb expression to another important target of Nanog activity, Klf4. In particular, it was assessed whether the correlation between Nanog and Esrrb expression observed in single cells is further linked to variations in Klf4 levels.

TNG Esrrb-TdTomato cells cultured in the presence or in the absence of puromycin were stained for Klf4 and Oct4 expression, allowing for concomitant quantification of the levels of 4 pluripotency regulators in single ES cells. Oct4 and Klf4 negative/positive thresholds were set based on the fluorescence levels detected in Nanog overexpressing cells and E14Tg2a differentiated by retinoic acid, as shown in **Figure 4.15**. $Oct4^{+}$ cells were readily distinguishable from a minority of differentiated cells in the population and were selected for further analysis. The fixation procedure did not compromise detection of the Esrrb-TdTomato and *Nanog*:GFP reporters (**Figure 4.17A**, left panel). It was possible to observe the expected correlation between *Nanog*:GFP and Esrrb levels in fixed cells, with Esrrb negative cells found only in the *Nanog*:GFP^{-} fraction of the population (**Figure 4.17A**, left panel). As previously observed, selection for high levels of Nanog expression with puromycin completely abolished Esrrb heterogeneity (**Figure 4.17B**, left panel). Strikingly, Nanog and Klf4 levels in single cells showed a correlation pattern that was very similar to that observed for Esrrb. Klf4 negative cells were mainly found in the *Nanog*:GFP^{-} fraction of the population, and Klf4 downregulation was not observed in cells retaining high expression of Nanog (**Figure 4.17A**, middle panel). In addition, puromycin selection abolished heterogeneous expression of Klf4 (**Figure 4.17B**, middle panel). It was next assessed whether the fraction of $Nanog^{-}$ cells that downregulated Esrrb coincided with cell that had also lost Klf4 expression. Indeed, Esrrb and Klf4 levels showed a tight correlation in single cells (**Figure 4.17A**, right panel), with the majority of $Esrrb^{-}$ cells also showing reduced Klf4 expression. In addition, whereas a significant proportion of $Esrrb^{-}$ $Klf4^{+}$ cells was not detected, it was possible to observe cells that retained Esrrb expression but had

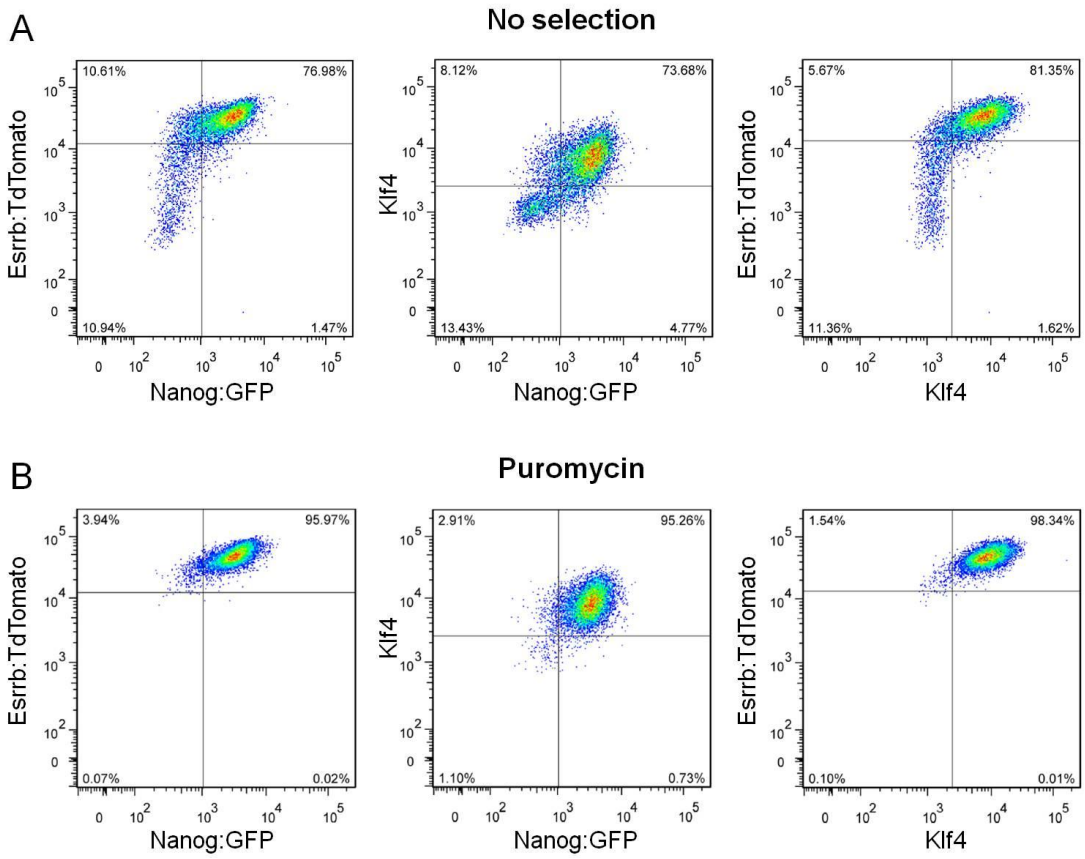


Figure 4.17: Nanog controls Esrrb and Klf4 fluctuations in ES cells.

A,B: Dot plots showing *Nanog*:GFP, Esrrb-TdTomato and Klf4 protein levels in Oct4^{+} TNG Esrrb-TdTomato ES cells cultured in the absence (A) or presence (B) of puromycin.

already downregulated Klf4, suggesting that Klf4 downregulation might follow a faster kinetic that loss of Esrrb (**Figure 4.17A**, right panel).

Taken together these results show that Nanog fluctuations control the levels of two other important pluripotency regulators, Esrrb and Klf4, in single ES cells. Nanog^{+} cells retain Klf4 and Esrrb expression and show high self-renewal efficiency whereas Nanog^{-} cells are susceptible to Esrrb and Klf4 downregulation, an event that marks commitment to differentiation.

4.8: Nanog and Oct4 fusion protein reporter ES cell lines

The work presented in this chapter explored the functional consequences of gain or loss of Nanog expression in single ES cells. The experiments discussed relied on fluorescent reporter lines for the quantification of Nanog and Esrrb expression.While a variety of reporter systems were developed to track Esrrb protein levels or transcriptional output reliably, the quantification of Nanog was based on a GFP knock-in allele. Replacing the Nanog coding sequence with a GFP cassette does not allow the detection of any regulatory events exerted on Nanog protein or could lead to delays in detection of Nanog downregulation due to the long half-life of GFP. To overcome these limitations, it was decided to develop Nanog-GFP or Nanog-RFP fusion protein reporter lines.

Fluorescent reporters were knocked-in at the 3' end of the last exon of *Nanog* without introducing an exogenous polyadenylation signal or altering the 3' UTR of the gene. Tag-RFP or GFP were fused in frame to the end of the Nanog coding sequence, connected through a five-glycine linker. The presence of an IRES sequence ensured expression of either blasticidin or hygromycin resistance genes (**Figure 4.18A**). Correct recombination was assessed by Southern blot (not shown).

The fluorescence output from both fusion protein reporters was low, but could be

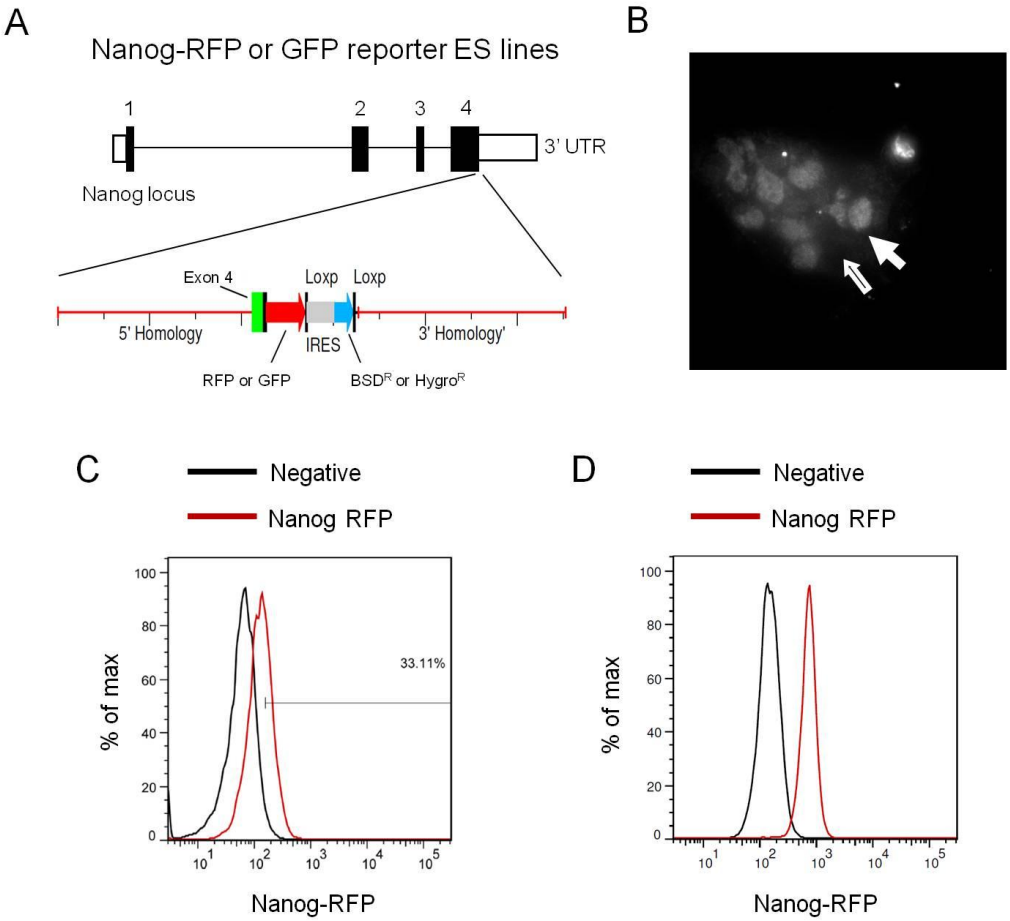


Figure 4.18: Nanog-RFP and Nanog-GFP ES cell reporter lines.

A: Schematic representation of the *Nanog* locus structure and targeting vectors employed to generate Nanog-RFP or Nanog-GFP reporter lines. **B**: Nanog-RFP detection by fluorescence microscopy in live cells. Open arrows indicate Nanog$^-$ cells and solid arrows Nanog$^+$ cells (Cells were identified based on brightfield images). **C, D**: Histograms showing Nanog-RFP fluorescence distributions in EF4 (That serve as negative) or E14Tg2a Nanog-RFP cells cultured in GMEMβ/FCS/LIF (C) or N2B27/2i/LIF (D).

detected as a clear nuclear signal by live cell microscopy (**Figure 4.18B**). Flow-cytometry analysis of Nanog-RFP or Nanog-GFP cells detected a reduced fraction of $Nanog^+$ cells when compared to what generally observed in TNG cells cultured in similar conditions (**Figure 4.18C**). This could be due to an improved ability to track downregulation of Nanog expression or degradation of Nanog protein. Alternatively, the observed reduction of $Nanog^+$ cells may be attributable to an insufficient sensitivity in fluorescence detection by flow cytometry.

Nonetheless, E14Tg2a Nanog-RFP cells maintained in 2i/LIF, a culture condition that results in homogeneous and elevated levels of Nanog expression, showed fluorescence levels completely distinct from negative cells that do not express Nanog-RFP (**Figure 4.18D**). This indicates that flow cytometry can be used to isolate E14Tg2a Nanog-RFP cells expressing high Nanog levels.

Overall, Nanog-RFP or Nanog-GFP fusion protein reporter lines are useful tools to track variations of Nanog protein levels in live imaging experiments, but present some limitations for use in flow cytometry.

4.9: Discussion

The results presented in this chapter show that Nanog fluctuations in single ES cells influence the expression of its target genes, as determined for Esrrb (**Figure 4.4, 4.5**) and Klf4 (**Figure 4.17**). A reciprocal regulation is not observed in overexpression experiments (**Fig 4.5,** see also **Figure 5.9**). This is in contrast with the notion that Esrrb can activate Nanog expression (van den Berg et al., 2008). In light of this observation, it is possible to imagine that Esrrb generally supports Nanog expression but it is not able to dramatically alter Nanog transcriptional status when overexpressed. In accordance with these results, timecourse experiments analysing the transcriptional response to Esrrb induction failed to detect a strong effect on Nanog expression (see Chapter 8.3). Klf4 is also tightly controlled by Nanog (See Chapter 3 and **Figure 4.17**) but knockdown experiments did not result in any effect of Klf4 depletion on Nanog expression (Jiang et al., 2008). It could therefore be that

Nanog lies at the top of a transcriptional cascade that modulates the fluctuations of heterogeneously expressed transcription factors in ES cells. Relevant to this point, heterogeneous expression has been reported not only for Esrrb ((van den Berg et al., 2008) and **Figure 4.4**) and Klf4 (Niwa et al., 2009) but also for the pluripotency modulator Rex1 (Toyooka et al., 2008) and Tbx3 (Niwa et al., 2009). Rex1 was shown to be promptly modulated by Nanog in tamoxifen treatment timecourse experiment, although it presented high variability among samples (1.52 fold change 6 hours after tamoxifen induction of ESΔN-NERT cells, p=0.122, see **Table 3.1**). Tbx3 also responded to Nanog induction (1.27 fold change at 6 hours, p=0.005, see **Table 3.1**). Intriguingly, Stella, another gene heterogeneously expressed in ES cells (Hayashi et al., 2008), did not show any variation in these experiments. Since Esrrb, Klf4 ((Jiang et al., 2008; Niwa et al., 2009) and see **Figure 5.13**) and Tbx3 (Niwa et al., 2009) all have a functional role in ES cells, it is possible that Nanog specifically orchestrates expression of factors involved in sustaining the undifferentiated state.

According to this hypothesis, changes in Nanog levels should be followed by changes in expression of a cohort of its target genes in a synchronous, coordinated way. Indeed, it was possible to observe that Esrrb and Klf4 fluctuation is single cells are linked and both factors show correlation with Nanog expression (**Figure 4.17**). Nanog fluctuations are not rapid, happening over a period of days ((Chambers et al., 2007) and **Figure 4.6**). Experiments exploiting inducible Nanog function in ESΔN-NERT ES cells carrying an Esrrb-T2a-TdTomato fluorescent reporter show that increase in Esrrb protein levels may follow Nanog upregulation with a delay of almost 24 hours (**Figure 4.12**). It would be interesting to follow the modulation of Esrrb, possibly in conjunction with Klf4, Rex1 or Tbx3, in time-lapse microscopy experiments, to ascertain whether changes in the protein levels of these genes present a delayed response to Nanog oscillations and whether reproducible patterns in the kinetics of this response can be identified.

The results presented here highlighted a clear correlation between Nanog and Esrrb protein expression in single cells (**Figure 4.4**). In addition, the data presented in chapter 3 show that elevated Nanog levels impact on Esrrb transcription in whole cell

populations. Whereas Nanog control on the accumulation or reduction of its targets at the protein level seems to achieve its effects over prolonged periods of time, it is reasonable to expect that Nanog control of transcription is characterised by much faster dynamics. Analysis of the transcriptional output from Nanog target genes in live cells might possibly reveal an even tighter correlation between Nanog expression and target gene activation. Analysis of Esrrb expression employing destabilised reporters of transcription detected the appearance of cells completely negative for *Esrrb*:GFPdest1 expression 24 hours after replating sorted Esrrb^{medium} populations (**Figure 4.8**). This shows that silencing of Esrrb transcription follows rapid kinetics compared to modulation of Esrrb protein levels (2 days required to see full downregulation of Esrrb in Nanog negative cells, **Figure 4.6**).

Recent studies uncovered that Nanog is monoallelically expressed in ES cells (Miyanari and Torres-Padilla, 2012). Observations by this group (Pablo Navarro, unpublished data) suggest that Esrrb also presents a predominantly monoallelic expression pattern. It would be interesting to determine whether Nanog levels shift the proportion of cells transcribing both *Esrrb* alleles in the ES cell population. *Esrrb*:GFPdest1 reporter lines engineered to express a distinct unstable fluorescent protein under the control of the second *Esrrb* allele could be used in these experiments. In this respect, it should be noted that, despite destabilised fluorescent reporter systems allow accurate monitoring of transcriptional silencing, detection of gene activation probably requires prolonged, continuous transcriptional output from these reporter alleles. This could lead to gross underestimation of the fraction of ES cells actively transcribing *Esrrb*. Cells that have recently started transcribing *Esrrb* or undergo less frequent or sporadic rounds of transcriptional firing would be detected as negative. In an alternative approach, the relation between Esrrb expression and Nanog protein levels in single cells could be explored coupling immunohistochemistry and RNA-FISH, but such an approach would preclude continuous measures in live cells. The use of two novel imaging techniques that allow the visualisation of foci of active transcription in live cells could circumvent these limitations and permit to directly correlate *Esrrb* transcription and Nanog levels

in ES cells (unpublished method developed by Luke Lee at UCSC Berkley and (Daigle and Ellenberg, 2007; Janicki et al., 2004)(see Chapter 9.2).

The experiments discussed in this chapter determined that loss of Esrrb expression marks ES cells primed for differentiation. Similarly, Nanog expression was reported to correlate with the ability of ES cells to self-renew (Chambers et al., 2007). Taken together these results provide a framework to understand the sequential molecular events that lead from Nanog downregulation to exit from the pluripotent state. Building on the results from transcriptional studies, it was possible to show that Nanog influences expression of a number of pluripotency factors in single cells. Sorting experiments demonstrate that Nanog downregulation is strictly required for ES cells to silence expression of Esrrb (**Figure 4.4**) and that Nanog^{low} cells are in a state in which they are more prone to downregulate expression of this and possibly other pluripotency factors (**Figure 4.6**). Nonetheless, Nanog^{low} cells retain their clonogenic potential in condition permissive only for undifferentiated cells (2i/LIF) and are still able to re-acquire expression of Nanog (**Figure 4.6, 4.7**). Complete loss of Esrrb in this subpopulation requires 2 days and accompanies progressive commitment to differentiation. Esrrb^{low} cells sequentially lose their ability to self-renew in 2i/LIF and FCS/LIF, and progressively downregulate markers characteristic of pristine pluripotency (**Figure 4.8**).

Esrrb and Nanog double reporter lines capture the progressive nature of exit from pluripotency with unprecedented detail and constitute a useful tool for the future characterisation of the early transcriptional and epigenetic events underlying this process. Interestingly, the results presented indicate that differentiating Esrrb^{low} cells acquire a gene expression profile reminiscent of EpiSC (**Figure 4.9**). This observation is in agreement with previous reports showing that Rex1 (Toyooka et al., 2008) and Stella (Hayashi et al., 2008) expression is heterogeneous in ES cells and downregulation of these markers leads to the acquisition of a gene expression pattern characteristic of the post-implantation epiblast. In particular it was shown that Rex1^{-} cells, like the post-implantation epiblast, have lost the ability to differentiate into primitive endoderm and, similarly to EpiSC, cannot contribute to development when

introduced in pre-implantation embryos (Toyooka et al., 2008). These reports and the results presented in this chapter seem to indicate that the early steps of in vitro differentiation recapitulate in vivo development. Interestingly, Esrrb in not expressed in the post-implantation epiblast (Luo et al., 1997) and transient Nanog downregulation is observed upon implantation ((Chambers et al., 2003; Osorno et al., 2012) and Fredrik Wong, unpublished information). It might be that Nanog expression is silenced upon implantation to allow downregulation of a number of ICM specific factors, including Esrrb, that would otherwise prevent normal progression of development. Upon implantation, epiblast cells also embark on the process of random X inactivation (Lyon, 1961; Rastan, 1982). It would be interesting to determine the X inactivation status in the distinct *Esrrb*:Tdtomato and *Nanog*:GFP subpopulations identified in this study. The analogies in the results obtained employing Esrrb, Rex1 and Stella reporters suggest that Esrrb is not a unique marker of early differentiation. The observation that heterogeneously expressed Nanog targets fluctuate in a coordinated way corroborates this interpretation. Despite this, the observation that Esrrb has a strong functional role in sustaining pluripotency (see Chapter 5 and 8) might indicate that loss of Esrrb has a causal role in triggering the differentiation process, rather than merely marking it.

Finally, the results presented highlight the strong influence of both intrinsic and extrinsic regulatory inputs on Esrrb expression. Nanog null cells show increasingly heterogeneous and overall reduced levels of Esrrb expression (**Figure 4.12**). In addition, 2i/LIF culture conditions strongly enhance Esrrb expression (compare **Figure 4.11** and **4.12**). Augmented Esrrb expression is also observed in Nanog null cells grown in 2i/LIF, indicating that ERK and GSK inhibition does not require Nanog to exert its effect. This is in agreement with recent reports that the GSK/β−catenin axis directly controls Esrrb expression in 2i conditions through modulation of Tcf3 activity (Martello et al., 2012). Nonetheless, the fact that Esrrb expression in Nanog null ES cells cultured in 2i/LF is reduced and more heterogeneous compared to wildtype cells indicates that Nanog is required for full Esrrb induction under these conditions. Nanog elevation could therefore be important in achieving the homogeneous gene expression profile observed in 2i/LIF. This

observation also connects to a point of broader relevance. Nanog induction in ESΔN-NERT cells cultured in 2i results in strong upregulation of Esrrb and downregulation of GFP (**Figure 4.13**). This proves that pluripotency factor expression is still strongly dependent of intrinsic regulatory inputs in "ground state" conditions, indicating that emerging proprieties of the pluripotency GRN that drive ES cell heterogeneity could be still operating in 2i/LIF, albeit at reduced levels.

Chapter 5: Effects of Esrrb overexpression on ES cell self-renewal

The data discussed in Chapter 3 identified a reliable list of Nanog target genes. Based on these results, the analysis presented in this chapter aims at getting a better understanding of the role that specific genes play in mediating Nanog activity in ES cells. In particular, experiments were performed to determine whether the overexpression of Esrrb is able to functionally complement the loss of Nanog in ES cells, and how this relates to the effects of elevating of other Nanog target genes.

5.1: Esrrb overexpression confers LIF independent self-renewal

Nanog was isolated by its ability to drive LIF independent self-renewal of ES cells (Chambers et al., 2003). Thus, experiments were performed to determine whether Esrrb overexpression results in a similar phenotype and whether Nanog activity is essential in this context.

Lifr$^{-/-}$ cells (Chambers et al., 2003) were transfected with an episomal Esrrb expression vector and plated at clonal density in the presence or absence of IL6/sIL6R. Transfection with Nanog or an empty vector served as controls. Forced Esrrb expression resulted in self-renewal in the absence of gp130 signalling (**Figure 5.1A**), but seemed to be less efficient than Nanog in episomal experiments.

Next, it was determined whether the observed effect could be recapitulated in clonal lines of stable integrants. A loxP-flanked Esrrb transgene was integrated in E14Tg2a cells (EfEsrrb cells) (**Figure 5.1B**). EfEsrrb cells showed robust Esrrb overexpression (**Figure 5.1C**), a phenotype reversed by transient transfection with Cre. When plated at clonal density in the presence of LIF, EfEsrrb cells formed a higher proportion of undifferentiated alkaline phosphatase (AP) positive colonies

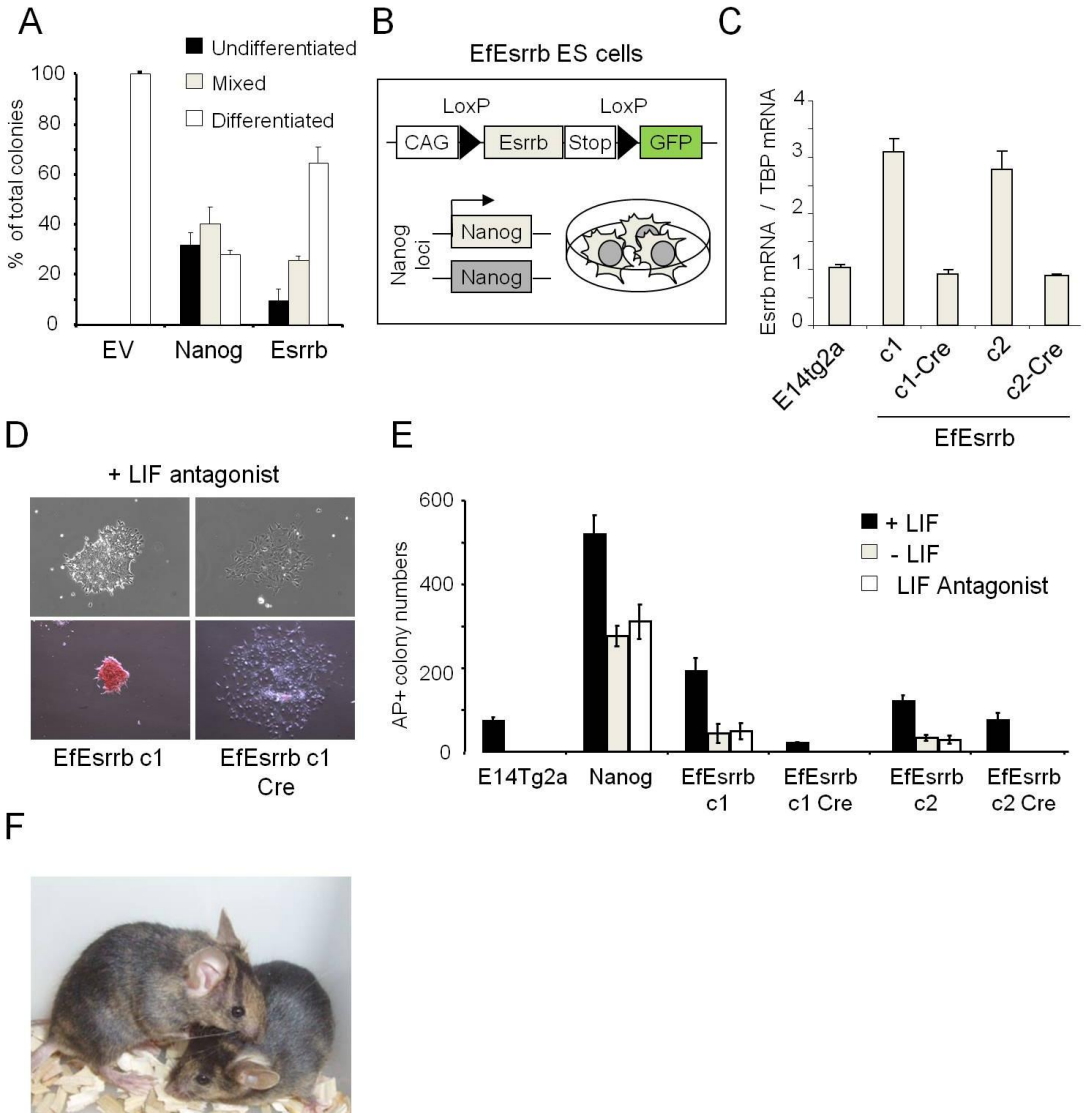


Figure 5.1: Esrrb promotes LIF-independent self-renewal.

A: lifr$^{-/-}$:PyLT+ LRK1 cells were transfected with episomal plasmids encoding Nanog or Esrrb (EV; empty vector is shown as a control) and the number of AP-positive colonies determined after clonal density plating in the absence of IL-6/sIL6R. Error bars: standard deviation (n=3). **B**: Schematic representation of the genetic manipulations performed on EfEsrrb ES cells. **C**: Relative Esrrb transcript level in E14tg2a, EfEsrrb and EfEsrrb-Cre cells.Error bars: standard deviation of the technical errors. **D**: Colony morphology and AP staining of EfEsrrb c1 cultured in the presence of hLIF-05.**E**: Number of AP^{+} colonies scored after plating at clonal density E14Tg2a or Nanog (EF4) and Esrrb overexpressing cells (before and after Cre reversion) and culture in the presence or absence of LIF or with hLIF-05 for 7 days. Error bars: standard deviation (n=3). **F**: Chimaeras generated after blastocyst injection of EfEsrrb cells passaged twice at clonal density in the absence of LIF and transfected with a Cre expression vector to excise the Esrrb transgene.

The data presented in this figures, with the exception of panel C and D, were kindly provided by Adam Yates and Douglas Colby.

compared to parental E14Tg2a cells. Importantly, in the presence of the LIF antagonist hLIF-05 (Vernallis et al., 1997) that completely abolishes LIF signalling, EfEsrrb cells were also able to form a sizable number of undifferentiated colonies (**Figure 5.1D-E**). This ability was completely lost after deletion of the Esrrb transgene, incontrovertibly indicating that forced Esrrb expression was responsible for the observed phenotype (**Figure 5.1D-E**). As previously observed in episomal experiments, EfEsrrb cells self-renewed less efficiently compared to EF4 (**Figure 5.1E**), a cell line expressing 6-8 fold wild-type levels of Nanog (**Figure 3.10A** and (Yates and Chambers, 2005)).

To rigorously determine whether Esrrb overexpression is sufficient to maintain pluripotency through clonal expansion in the absence of LIF signalling, EfEsrrb cells were plated at clonal density in the presence of LIF antagonist and passaged twice at clonal density. At this point, control parental cells had completely differentiated and could not be passaged further. In contrast, EfEsrrb clones continued to self-renew. These cells were treated with Cre and GFP expressing cells that had deleted the Esrrb open reading frame (ORF) were expanded in LIF. Injection of these cells into C57BL/6 blastocysts gave rise to adult chimaeras (**Figure 5.1F**), indicating that Esrrb overexpressing cells remained pluripotent after repetitive passaging in the absence of LIF.

5.2: Esrrb overexpression promotes LIF independent self-renewal independently of Nanog

It was then determined whether Esrrb is able to functionally substitute for Nanog in sustaining gp130-independent self-renewal. With this aim, the Esrrb coding sequence was cloned in a CAG promoter driven expression vector downstream of a TdTomato reporter, linked through a 2a peptide, and the vector was electroporated into $Nanog^{-/-}$ TβC44c6 ES cells (ESΔN-CAGE ES cells)(**Figure 5.2A**). The 6 clones isolated showed robust transgene expression when compared to parental TβC44c6 ES cells

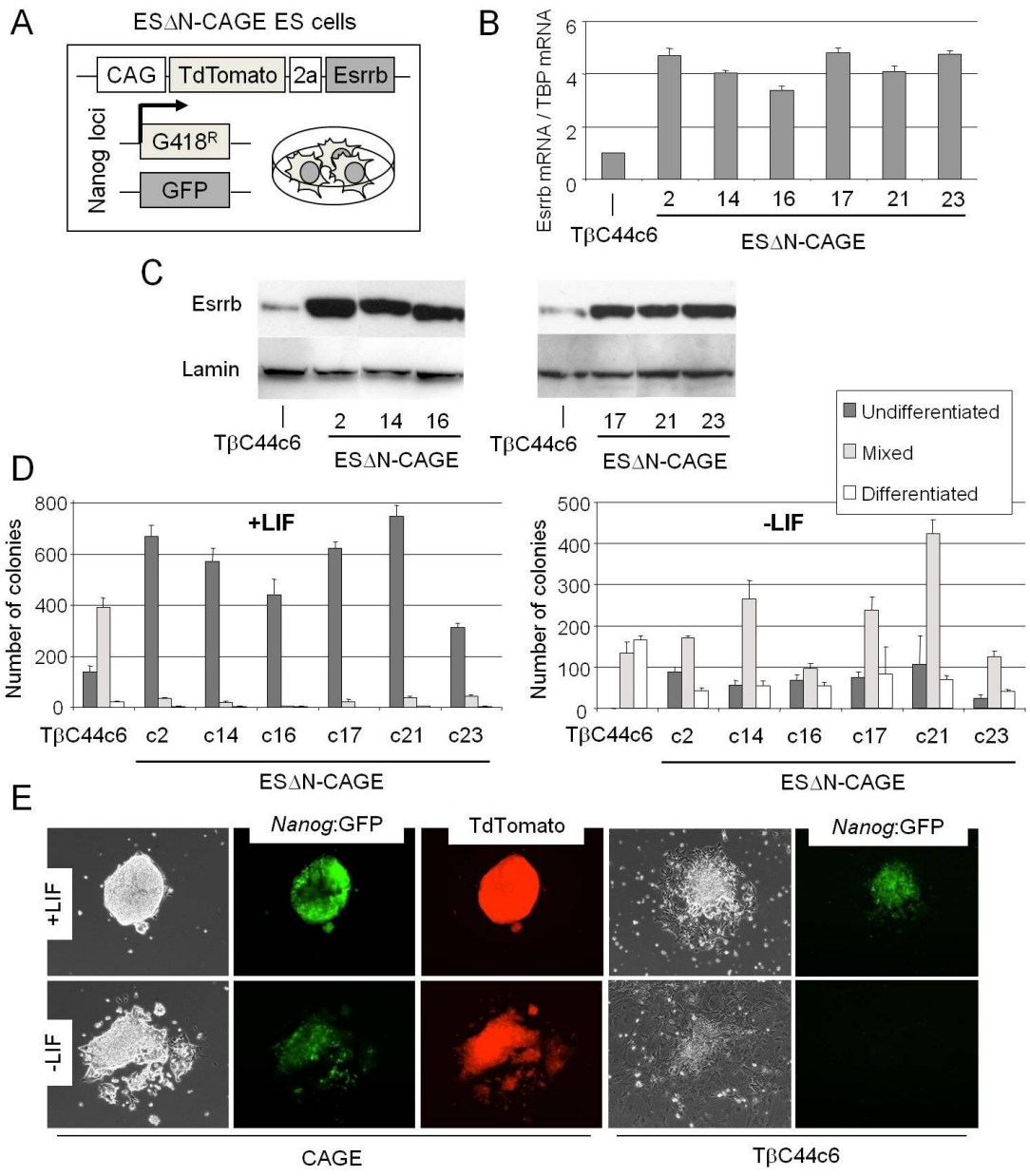


Figure 5.2: Esrrb promotes LIF-independent self-renewal in the absence of Nanog.

A: Schematic representation of the genetic manipulations performed on ESΔN-CAGE ES cells. **B**: Esrrb transcript levels in TβC44c6 and ESΔN-CAGE ES cells. Error bars: standard deviation of the technical errors in gene expression quantification in one experiment. **C**: Esrrb protein expression in parental TβC44c6 and ESΔN-CAGE ES cells. **D**: Number of AP^+ colonies scored after plating at clonal density TβC44c6 or ESΔN-CAGE cells and culture in the presence or absence of LIF for 8 days. Error bars: standard deviation of the number of colonies counted in three different plates. **E**: TdTomato and *Nanog*:GFP reporter expression detected in the cells used for the experiment presented in D.

(**Figure 5.2B-C**). Each line was plated at clonal density in the presence or in the absence of LIF and the number of alkaline phosphatise (AP) positive colonies was scored 8 days after plating. When LIF was added to the culture medium, Esrrb expression in ESΔN-CAGE cells resulted in an increased numbers of AP positive colonies compared to the parental line, showing that Esrrb can complement the self-renewal defects observed in Nanog null ES cells. In the absence of LIF, parental TβC44c6 ES cells could not form any undifferentiated colony (**Figure 5.2D**). In contrast, Esrrb overexpression led to the formation of tight, morphologically undifferentiated AP^+ colonies, indicating that Esrrb activity is not dependent on Nanog (**Figure 5.2D**).

Prior to staining, the activity of the *Nanog*:GFP reporter present in TβC44c6 cells was assessed by fluorescence microscopy. Active transcription from the Nanog locus was detected after 8 days of culture in the absence of LIF. Intriguingly, expression of GFP was not homogeneous in the culture. It was possible to observe colonies showing a tight morphology, generally associated with an undifferentiated state, in which part of the cells lacked GFP expression (**Figure 5.2E**). All the colonies observed retained high expression of the Esrrb transgene, as judged by the strong TdTomato fluorescence. These data further strengthens the notion that Nanog activation is not a crucial component of Esrrb activity in promoting LIF independence in ES cells.

5.3: Design and construction of doxycycline inducible expression systems

To conclusively prove that forced Esrrb expression results in LIF independent self-renewal in cells lacking Nanog, it was necessary to determine whether $Nanog^{-/-}$ Esrrb overexpressing cells could maintain their pluripotent state after long-term culture in the absence of LIF, showing uncompromised differentiation potential after silencing of the Esrrb transgene expression.

It was thus decided to derive a doxycycline inducible system that would allow for tuneable and reversible transgene expression in ES cells. A modified version of the pPyPCAG vectors was obtained by substitution of the constitutive CAG promoter for a doxycycline responsive element (**Figure 5.3A**). This element is based on the fusion of a minimal CMV promoter with six tandem repeats of the tetracycline operator sequence (Courtesy of Keisuke Kaji).

5.3.1: Rosa26 based rtTA expression

A TdTomato-2a-Esrrb cassette was cloned into this vector and the construct randomly integrated into the genome of *Rosa*:rtTA ES cells, a cell line that expresses an improved form of the rtTA transactivator (rtTAMs2) from the *Rosa26* locus ((Urlinger et al., 2000), courtesy of Andrew Smith). Clones were cultured in doxycycline and selected for high expression of the TdTomato fluorescent reporter by flow cytometry. In conjunction, the basal levels of transgene expression were evaluated in untreated cells. After two consecutive screening rounds on a total of 34 clones, the three lines showing highest transgene induction and lowest basal expression were selected for further analysis (**Figure 5.3B**). A three (clone 17) to fourfold (clone 15) induction in the levels of Esrrb transcript was observed in *Rosa*:rtTA TdTomato-Esrrb lines cultured for 24 hours in the presence of doxycycline compared to the parental line (**Figure 5.3C**). Conversely, basal levels of transgene expression varied from being almost undetectable (clone 17) to resulting in almost a twofold induction in Esrrb transcript (clone 7). An induction in Esrrb protein could be detected by western blot (**Figure 5.3D**). Little or no difference in the basal level of Esrrb protein compared to *Rosa*:rtTA cells was observed for clone 17, while increased levels were detected for clones 15 and 7. TdTomato reporter expression was evident by fluorescence microscopy in all three lines after 24 hours of induction, while little or no signal was observed in untreated cells (**Figure 5.3E**). In addition, it should be noted that TdTomato levels appeared to be heterogeneous in these lines (**Figure 5.3B**), and that transgene expression was confined to the undifferentiated fraction of the ES cells in culture (**Figure 5.3E**).

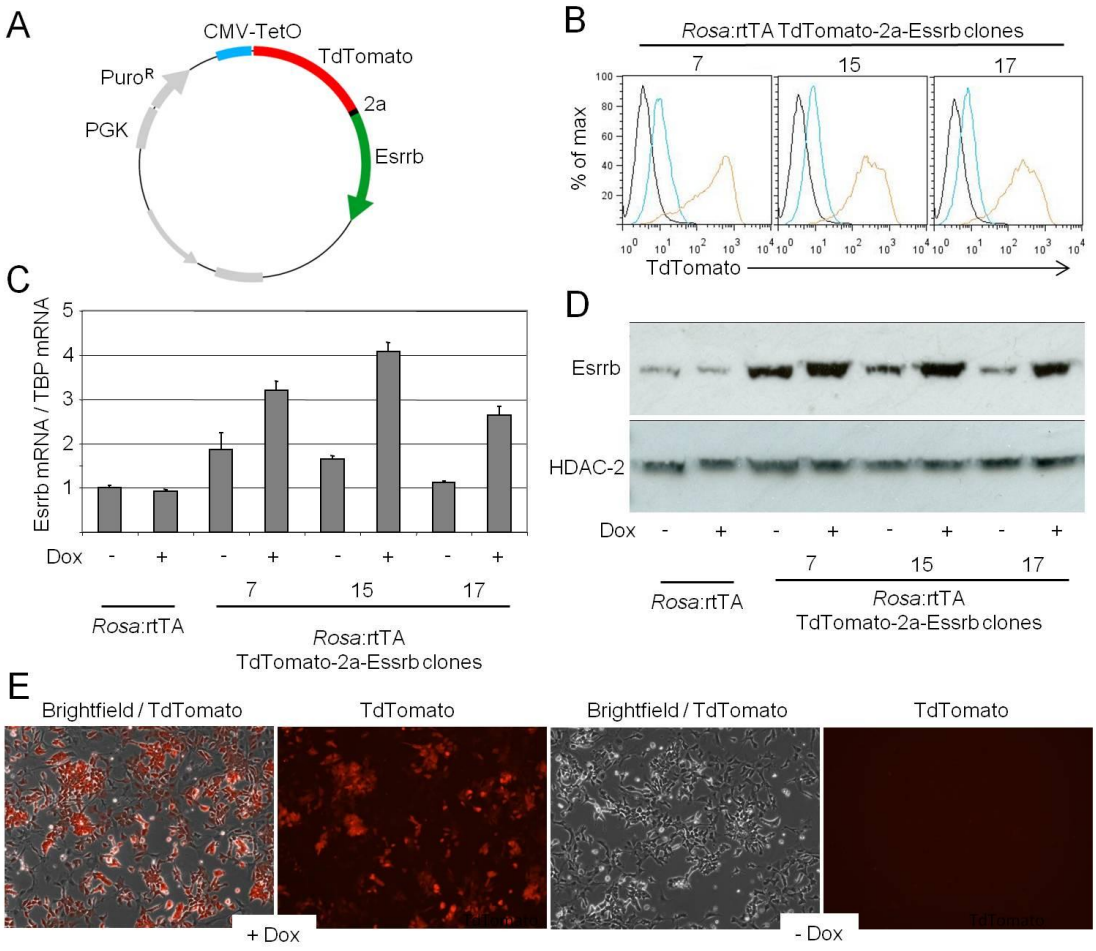


Figure 5.3: Characterisation of *Rosa*:rtTA Esrrb-2a-Tdtomato ES cells.

A: Map of the doxycycline inducible Esrrb-2a-TdTomato expression vector. B: Histograms showing the distributions of TdTomato fluorescence in three *Rosa*:rtTA TdTomato-2a-Esrrb lines cultured in the presence (brown) or in the absence (blue) of 1 μg/ml doxycycline for 4 days. The fluorescence profile of cells not expressing TdTomato is shown as a negative control (black). **C,D**: Esrrb transcript (C) and protein (D) levels in *Rosa*:rtTA and three *Rosa*:rtTA TdTomato-2a-Esrrb ES cell lines cultured in the presence or absence of 1 μg/ml doxycycline for 4 days. Error bars in panel C: standard deviation of the technical errors in gene expression quantification in one experiment. **E**: TdTomato expression in *Rosa*:rtTA TdTomato-2a-Esrrb cells cultured in the absence or presence of 1 μg/ml doxycycline.

The potential to finely tune transgene induction was then assessed by culturing Rosa:rtTA TdTomato-2a-Esrrb ES cells in the presence of increasing concentration of doxycycline. TdTomato and Esrrb transcripts levels in the cells showed a sigmoidal response to the concentration of doxycycline, with a nearly linear response on a logarithmic scale for doxycycline concentrations ranging from 10ng/ml to 1000ng/ml (**Figure 5.4A**). A similar response was observed by flow cytometry for the fluorescence levels of the TdTomato reporter (**Figure 5.4B**). Importantly, the distribution of TdTomato fluorescence values became more heterogeneous when cells were cultured in the presence of 100ng/ml of doxycyline, suggesting that special attention should be used when such inducible lines are used to test the effects of low levels of Esrrb expression in ES cells (**Figure 5.4C**). Finally, the transgene expression achieved in these doxycycline inducible lines was compared with the levels of Esrrb observed in other overexpressing cells. Total Esrrb levels were measured by quantitative PCR in ESΔN-CAGE and Rosa:rtTA TdTomato-2a-Esrrb cells. A three to fourfold induction over TβC44c6 levels was observed for doxycycline inducible lines while a four to five fold induction was achieved with CAG promoter based expression vectors (**Figure 5.4D**).

The functionality of *Rosa*:rtTA TdTomato-2a-Esrrb cells was then tested in a LIF independence assay. Cells were plated at clonal density in the presence or absence of LIF and after 7 days the number of AP positive clones was scored. Culture of inducible lines in the presence of LIF and doxycycline resulted in an increased numbers of AP positive colonies compared to uninduced or parental *Rosa*:rtTA cells (**Figure 5.5A**). Nonetheless, in sharp contrast to what observed for CAG promoter driven overexpression of Esrrb, addition of doxycycline to cells cultured in the absence of LIF did not result in the presence of undifferentiated colonies in the plates (**Figure 5.5A**). Only limited numbers of mixed colonies in which self-renewing ES cells were surrounded by differentiating cells could be observed (**Figure 5.5B**). This result may be explained by the previous observation that transgene expression is heterogeneous in *Rosa*:rtTA TdTomato-2a-Esrrb cells, in particular in differentiated cells. In *Rosa*:rtTA ES cells a PGK-rtTA cassette is targeted to the Rosa26 locus.

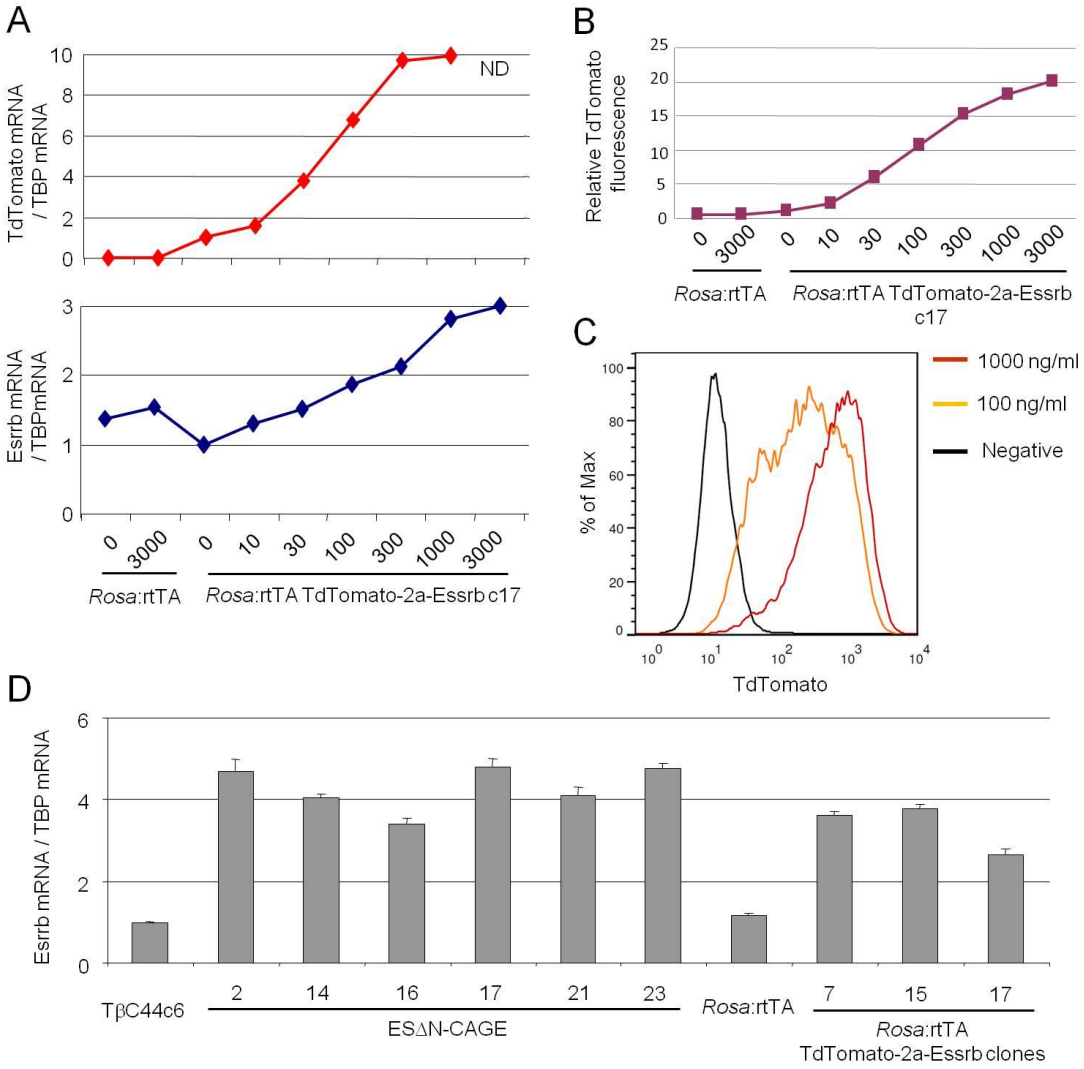


Figure 5.4: Tunable transgene induction in *Rosa*:rtTA TdTomato-2a-Esrrb ES cells.

A: Esrrb and TdTomato transcript levels in *Rosa*:rtTA and *Rosa*:rtTA TdTomato-2a-Esrrb ES cells cultured in the presence of the indicated concentrations of doxycycline (ng/ml) for 4 days. Error bars: standard deviation of the technical errors in gene expression quantification in one experiment. **B**: Mean TdTomato fluorescence values detected in *Rosa*:rtTA and *Rosa*:rtTA TdTomato-2a-Esrrb cells cultured in the presence of the indicated concentrations of doxycycline (ng/ml) for 4 days. **C**: Histograms showing the distributions of TdTomato fluorescence in *Rosa*:rtTA TdTomato-2a-Esrrb lines cultured in the presence of the indicated doxycycline concentration for 4 days. The fluorescence profile of cells not expressing TdTomato is shown as a negative control. **D**: Esrrb transcript level in TβC44c6 and 6 ESΔN-CAGE lines compared to *Rosa*:rtTA and 3 *Rosa*:rtTA TdTomato-2a-Esrrb lines cultured in the presence of 1 μg/ml doxycycline for 4 days. Error bars: standard deviation of the technical errors in gene expression quantification in one experiment.

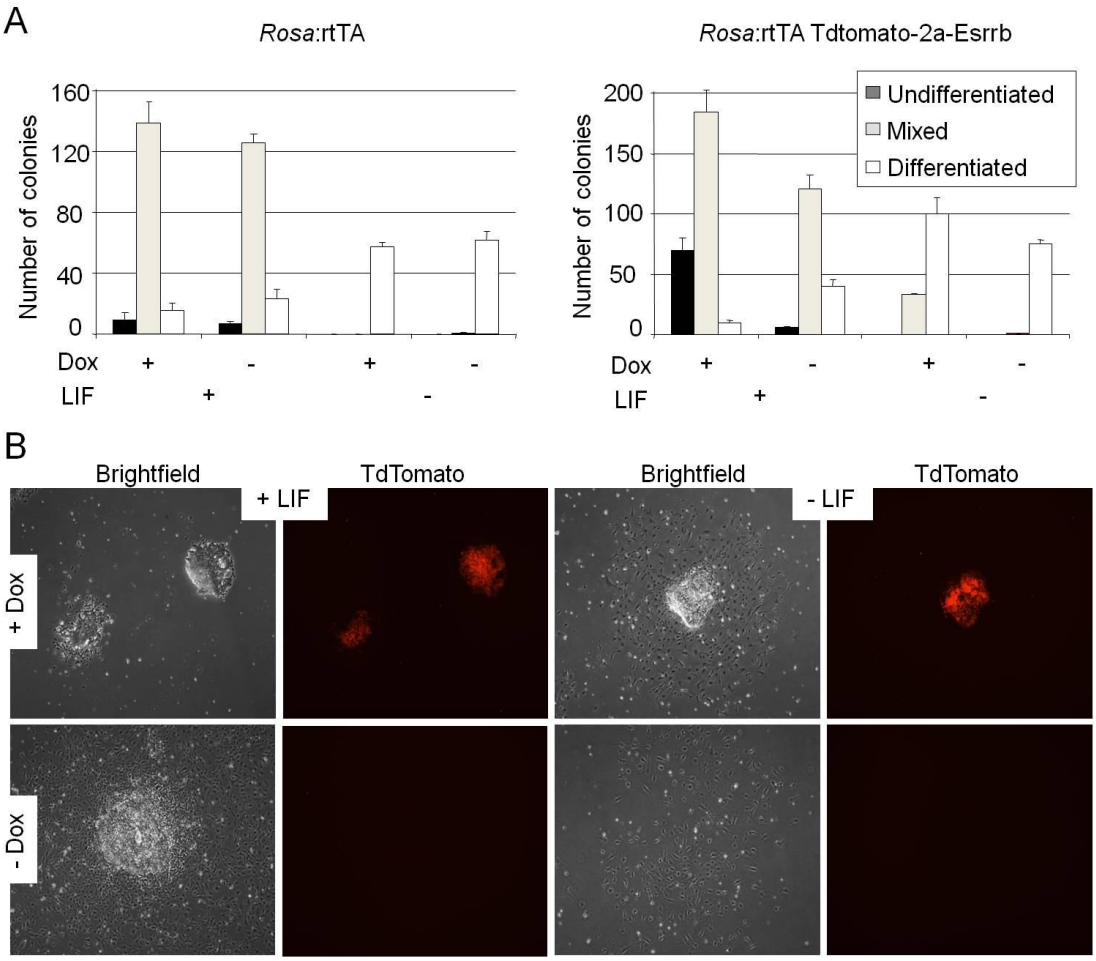


Figure 5.5: Doxycycline inducible Esrrb expression fails to promote LIF-independent self-renewal in *Rosa*:rtTA TdTomato-2a-Esrrb ES cells.

A: Number of AP+ colonies scored after clonal density plating of *Rosa*:rtTA TdTomato-2a-Esrrb cells and culture in the presence or absence of LIF and 1 μg/ml doxycycline for 7 days. Error bars: standard deviation of the number of colonies counted in three different plates. **B**: Colony morphology and TdTomato expression in *Rosa*:rtTA TdTomato-2a-Esrrb cells plated at clonal density and cultured in the presence or absence of LIF and 1 μg/ml doxycycline for 7 days.

This ensures controlled genomic alteration and ubiquitous expression of rtTA during development, making this cell line suitable for in-vitro and in-vivo studies. However, optimal levels of rtTA expression could not be achieved in this system.

5.3.2: Randomly integrated rtTA system

To test whether variations in the rtTA expression levels might result in higher transgene induction or lower background expression in inducible lines, a CAG-rtTA cassette was randomly integrated into the genome of TβC44c6 Nanog null ES cells. Stably transfected clones were pooled and electroporated with a version of the previously described doxycycline inducible vectors modified to allow for recombinase mediated cassette exchange (RMCE) (ESΔN-iTdT cells, **Figure 5.6A**). In this system cells are initially transfected with a vector driving expression of TdTomato and a hygromycin resistance - thymidine kinase (Hygro-TK) fusion protein. Once clones showing optimal expression are identified, negative selection ensured by thymidine kinase can be exploited to insert different sequences at the integration site by RMCE, using Flp recombinase and heterotypical FRT sites (Raymond and Soriano, 2007) (**Figure 5.6A**). This system allows the insertion of sequences encoding any gene of interest in substitution or downstream of the TdTomato fluorescent reporter present in the integrated construct. It is therefore possible to compare the effects of forced expression of different genes in ES cell lines with identical genetic background. This excludes phenotypic differences resulting from alterations caused by the random insertion of the transgene into the ES cell genome. In addition, similar levels of expression should be achieved for all genes, facilitating direct comparison of the results.

After electroporation of the TβC44c6 line, ESΔN-iTdT cells were selected with hygromycin in the presence of doxycycline and blasticidin for 8 days. At this time hygromycin and blasticidin selection were withdrawn and cell cultured for additional 4 days in order to identify clones in which transgene expression was maintained in

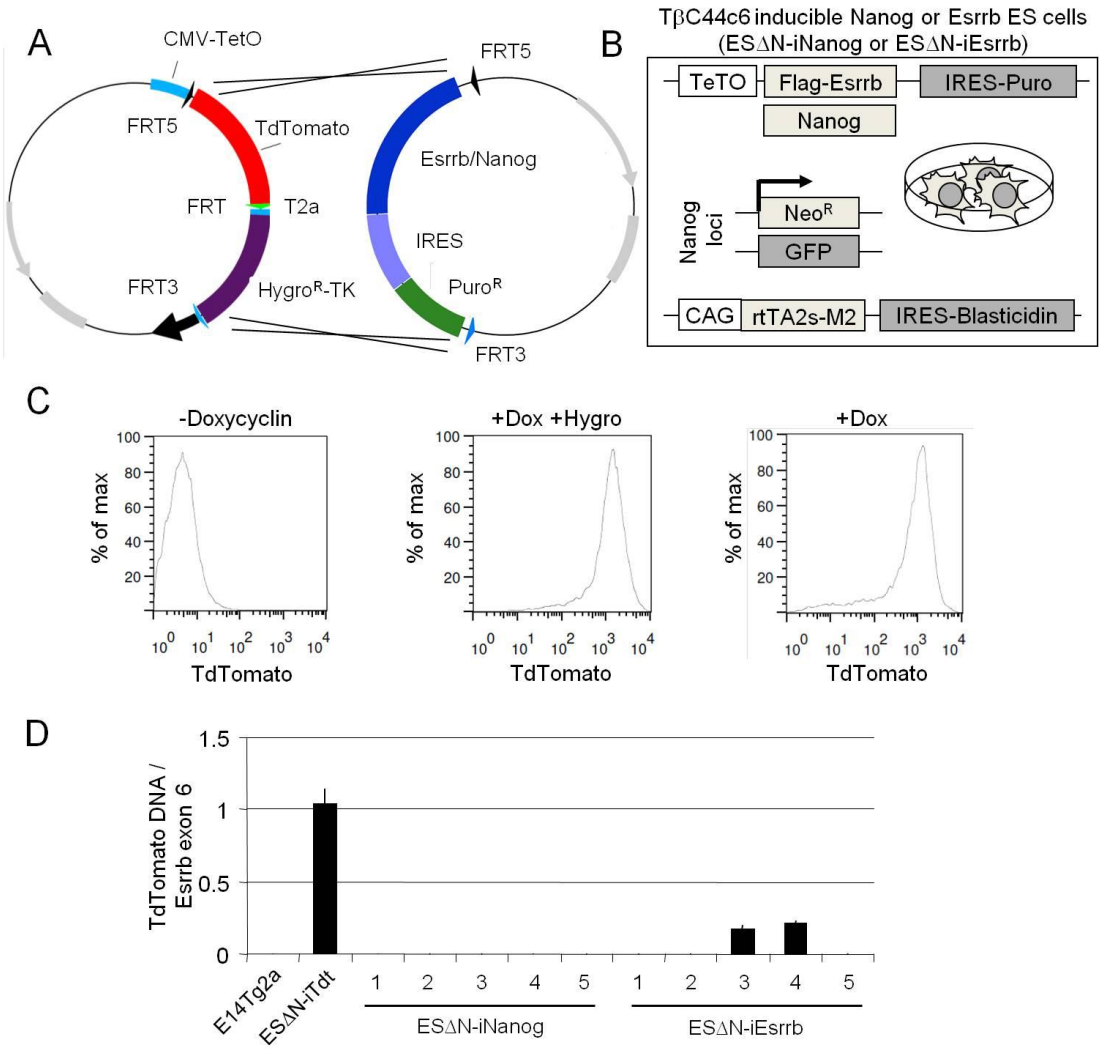


Figure 5.6: Doxycycline inducible ESΔN-iNanog or ESΔN-iEsrrb lines.

A: Map of the inducible TdTomato-2a-Hygro^R-TK and Nanog or Esrrb shuttle vectors. **B**: Schematic representation of the genetic manipulations performed on ESΔN-Nanog or ESΔN-Esrrb ES cells. **C**: Histograms comparing the distribution of TdTomato fluorescence in ESΔN-iTdT cells cultured in the presence of doxycycline and hygromycin, or in the presence of doxycycline after releasing from hygromycin selection for 2 weeks. Fluorescence levels of untreated cells are shown in the left panel. **D**: Quantitative PCR analysis of the relative levels of TdTomato to *Esrrb* exon 6 performed on genomic DNA from E14Tg2a, ESΔN-iTdT or 5 ESΔN-iNanog and ESΔN-iEsrrb clones. Error bars are the standard deviation of the technical error (n=3).

the absence of selection. In addition, since neomycin selects for active transcription from the *Nanog* locus in TβC44c6 ES cells, nascent colonies were cultured in the absence of neomycin to allow for low levels of spontaneous differentiation to occur in the ES cell population, (**Figure 5.5B**). 35 colonies showing intense fluorescence and sustained expression in differentiated cells at the rim of the colony were selected. After picking, clonal lines were expanded and screened for TdTomato expression by flow cytometry. Neomycin selection was then restored and cells were expanded in the absence of blasticidin and hygromycin for 14 additional days. Each clone was then replated in the presence or in the absence of doxycycline and after 4 days TdTomato fluorescence was assessed again by flow cytometry. The distribution of fluorescence intensities varied extensively in different ESΔN-iTdT clones. While complete loss of transgene expression was rarely observed, most of the clones showed greatly reduced TdTomato levels and a broad distribution of fluorescence values. Only two clones maintained satisfactorily high levels of transgene expression in the absence of continuous selection. One clone that showed the highest levels of transgene expression and presented a tight distribution of fluorescence values (**Figure 5.6C**) was selected for RCME. ESΔN-iTdT cells were transiently co-transfected with a plasmid driving expression of the Flp recombinase and a shuttle vector carrying an Esrrb or Nanog-IRES-puromycin resistance gene cassette flanked by heterotypical FRT3 and FRT5 sites. Cells were replated at low density and after 48 hours puromycin selection was applied. 4 days after replating ganciclovir was added to the culture medium. Clones showing loss of TdTomato expression formed viable colonies and 10 days after replating could be picked and expanded. Successful recombination was assessed by qPCR on genomic DNA using primer pairs binding to the TdTomato cassette present in parental ESΔN-iTdT cells. Amplification of an unrelated genomic locus (Esrrb exon 6) served as a control to normalise for the amount of template DNA loaded in each reaction. Esrrb exon 6 was detected in all samples analysed and amplification of TdTomato was robust in genomic DNA samples from parental ESΔN-iTdT cells. No TdTomato amplification was detected in inducible Nanog clones (**Figure 5.6D**). Interestingly, two out of 5 inducible Esrrb clones analysed, retained the TdTomato cassette. Since a greatly reduced number of colonies were observed after RCME for the derivation of inducible Esrrb compared

to inducible Nanog lines, these results might indicate that the Esrrb expression levels achieved in this system are detrimental to optimal self-renewal of ES cells. Nonetheless, 3 inducible Esrrb clones that had undergone correct exchange of the TdTomato cassette could be identified (**Figure 5.6D**).

Expression of Nanog and Esrrb transgenes was evaluated in two independent clones for both inducible Nanog and Esrrb lines after culture in the presence or absence of doxycycline. A three to four fold induction in Esrrb or Nanog levels over E14Tg2a ES cells was detected for all clones analysed. As expected, levels of expression were similar in the two clones selected for each line (averaged expression values are presented in **Figure 5.7A**). Background levels of expression in untreated cells were almost undetectable for both ESΔN-iNAnog and ESΔN-iEsrrb clones. Since TβC44c6 ES cells lack functional *Nanog* alleles, the low levels of Nanog transcript detected in ESΔN-iNanog lines confirm the absence of leaky transgene expression. Finally, induction of Nanog was accompanied by uniform morphological changes, characteristic of Nanog overexpression, in ESΔN-iNanog cells (**Figure 5.7B**), indicating that homogeneous expression is achieved in this line.

5.4: Esrrb expression sustains Nanog null ES cell self-renewal after long-term passaging in the absence of LIF

It was then tested whether the sustained and homogenous transgene expression achieved in ESΔN- inducible lines would allow for their use in functional assays.
With this aim ESΔN-iNanog or ESΔN-iEsrrb cells were plated at clonal density in the presence of LIF or LIF antagonist, with or without doxycycline.
After 7 days in the absence of LIF, only ESΔN-iEsrrb or ESΔN-iNanog cells cultured in the presence of doxycycline formed undifferentiated colonies expressing *Nanog*:GFP reporter. Esrrb overexpression resulted in the presence of tight self-renewing colonies (**Figure 5.8A**), although in lower numbers compared to Nanog (**Figure 5.8B**). Intriguingly, while GFP expression seemed to be homogeneous in ES

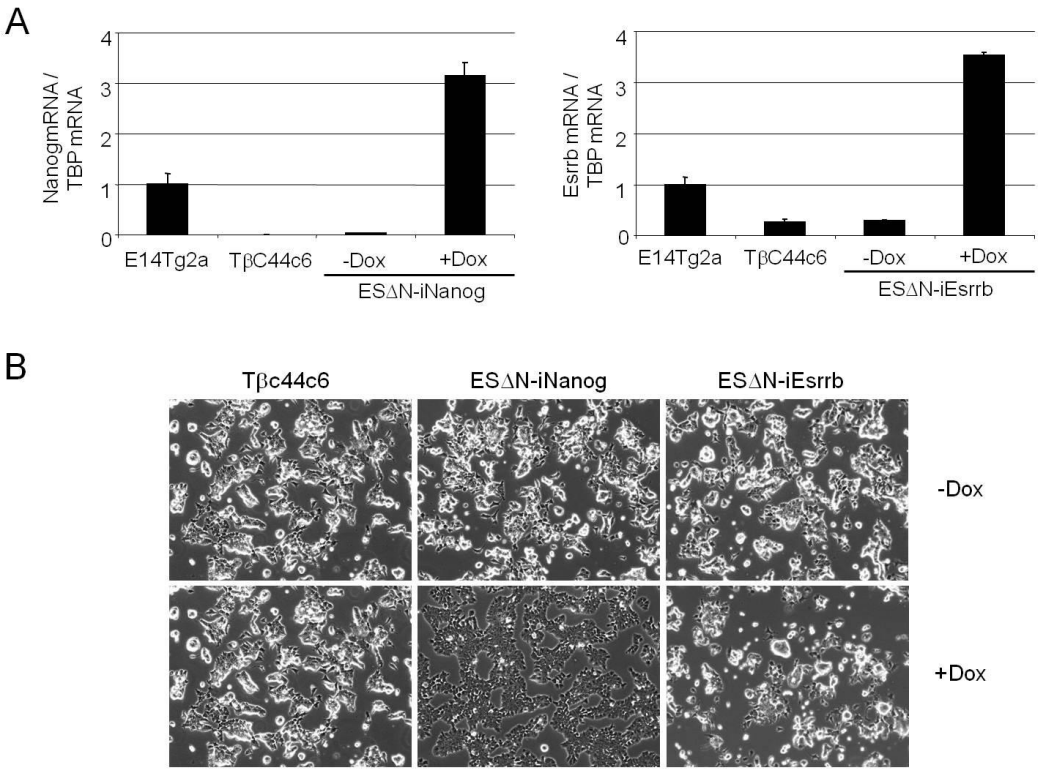


Figure 5.7: Doxycycline inducible ESΔN-iNanog or ESΔN-iEsrrb lines.

A: Esrrb and Nanog transcript levels in E14Tg2a, TβC44c6 and ESΔN-iNanog or ESΔN-iEsrrb ES cells cultured in the presence or in the absence of 1μg/ml of doxycycline. Error bars: standard deviation of gene expression measured in two independent clones (ESΔN-iNanog or ESΔN-iEsrrb) or three independent experiments (E14Tg2a, TβC44c6).

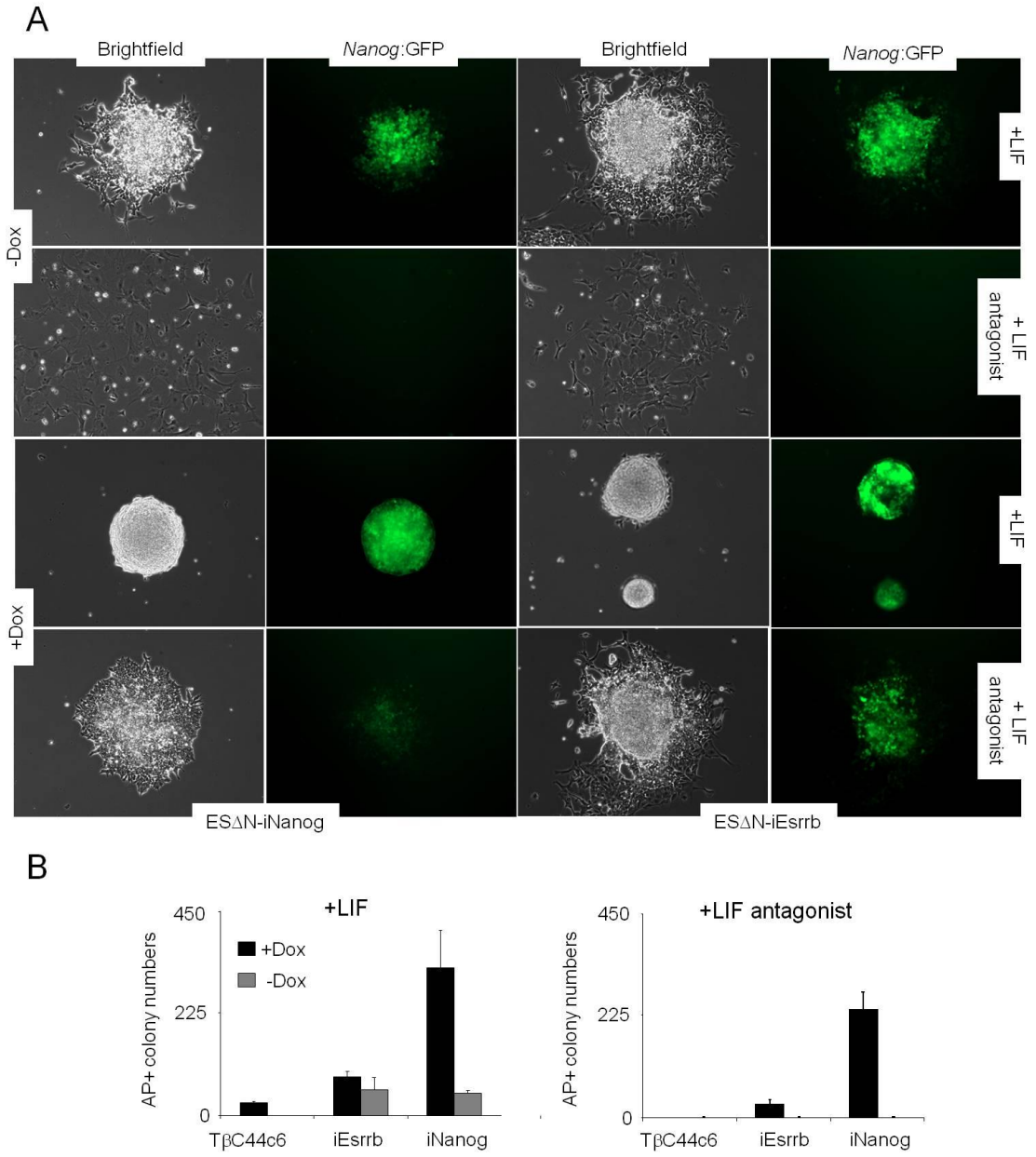


Figure 5.8: Esrrb expression induces LIF independent self-renewal in Nanog null ES cells.

A: Colony morphology and *Nanog*:GFP expression in ESΔN-iNanog or ESΔN-iEsrrb cells plated at clonal density and cultured for 8 days in the presence of LIF or LIF antagonist, with or without 1μg/ml of doxycycline. **B**: Number of AP^+ colonies scored after clonal density plating of ESΔN-iNanog or ESΔN-iEsrrb cells and culture for 8 days in the presence of LIF or LIF antagonist, with or without 1μg/ml of doxycycline.

cells overexpressing Nanog, ESΔN-iEsrrb cells showed heterogeneous *Nanog:*GFP expression. To substantiate the above observations, the experiment was repeated and after 7 days of culture at clonal density cells were harvested to quantify GFP expression by flow cytometry. As shown in **Figure 5.9B**, TβC44c6, ESΔN-iNanog and ESΔN-iEsrrb cultured in the presence of LIF and absence of doxycycline were indistinguishable. In the presence of LIF, Nanog induction resulted in decreased but homogeneous levels of *Nanog*:GFP expression in undifferentiated SSEA-1 positive cells (**Figure 5.9C**). Conversely a clear GFP^- population appeared when Esrrb expression was induced (**Figure 5.9C**). This difference was maintained when cells were cultured in the absence of LIF, with Esrrb induction resulting in the concomitant presence of bright GFP^+ cells and cells that had lost GFP expression in the same population. Nanog expression resulted in homogenously reduced GFP expression in the same conditions (**Figure 5.9D**).

Since these experiments were performed in Nanog null ES cells, it is possible to conclude that the ability of Esrrb to confer LIF independent self-renewal to ES cells is not mediated by sustained activation of Nanog expression and does not rely on the presence of functional Nanog protein in the cells.

ESΔN-iNanog and ESΔN-iEsrrb cells could be cultured in the presence of LIF antagonist for more than one month. While Nanog induction completely suppressed differentiation under these conditions, Esrrb expression resulted in the maintenance of self-renewing ES cells accompanied by the presence of differentiating cells (**Figure 5.10A**). After more than one month in the continuous absence of LIF signalling cells were transferred to the kidney capsule of congenic mice to assess the presence of pluripotent cells in the population. A total of ten animals were used for both Esrrb and Nanog inducible lines. 5 animal of each group were administered 10μg/ml doxycycline in the drinking water, while the remaining 5 were not treated. Exactly one month after kidney capsule injection, animals were sacrificed to score tumour size and evaluate tissue contribution. 8 and 9 animals survived injections with ESΔN-iNanog cells (4 +doxycycline/ 4 –doxycycline) and ESΔN-iEsrrb cells

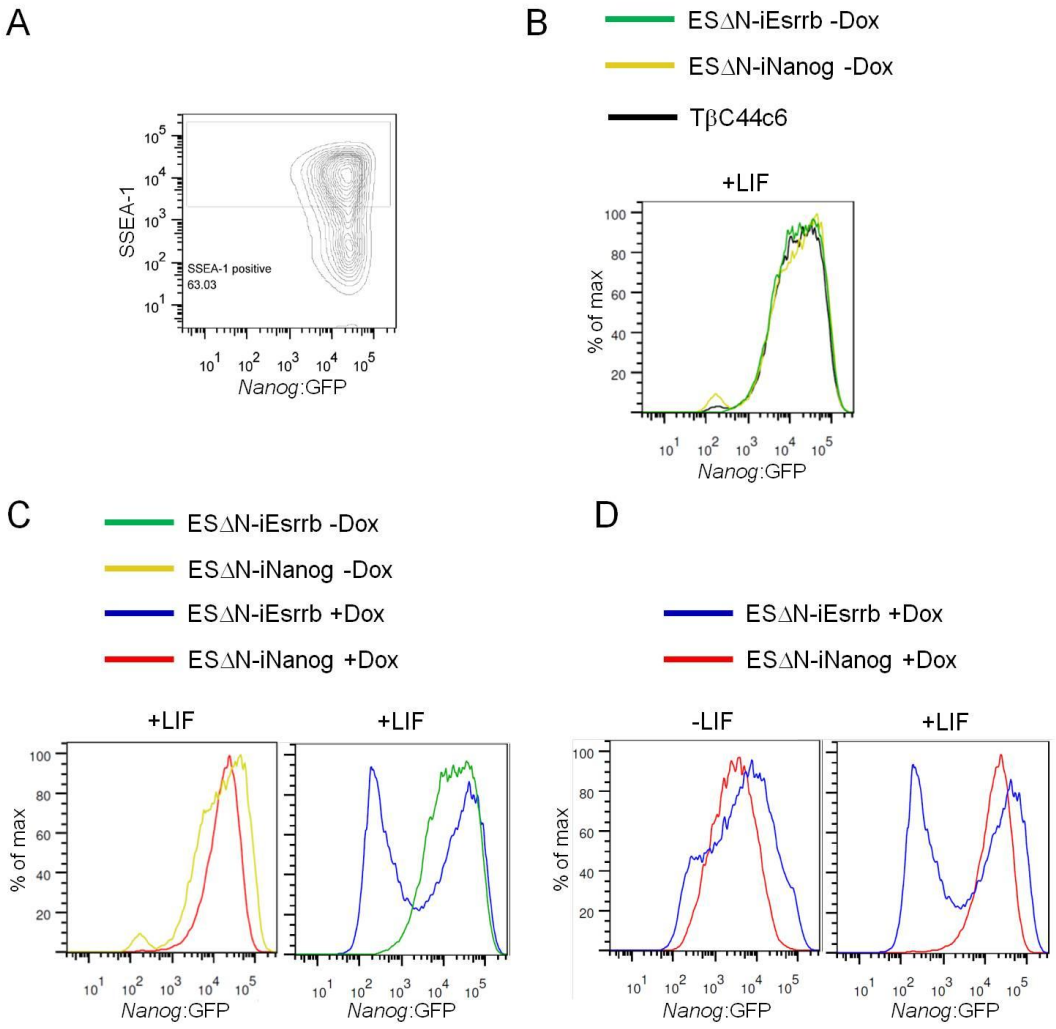


Figure 5.9: Esrrb elevation does not result in homogeneous Nanog expression.

A: Representative contour plot showing the distinct SSEA-1$^+$ and SSEA-1$^-$ populations indentified in this analysis. **B,C,D**: Histogram plots comparing the distribution of *Nanog*:GFP reporter expression in SSEA-1$^+$ TβC44c6, ESΔN-iNanog or ESΔN-iEsrrb cells plated at clonal density and cultured in the presence of LIF or LIF antagonist, with or without 1μg/ml of doxycycline for 8 days.

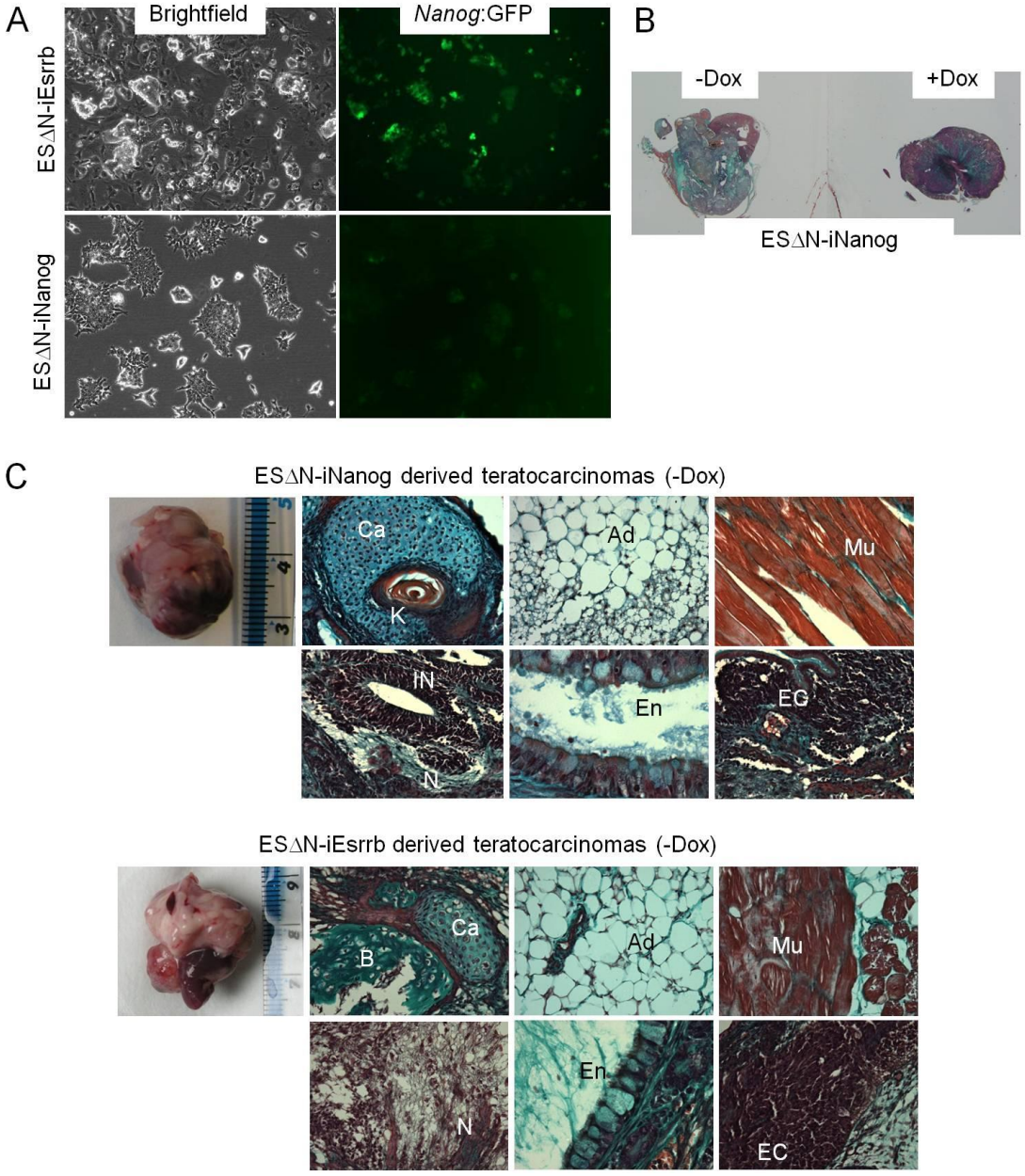


Figure 5.10: ESΔN-iNanog and ESΔN-iEsrrb maintain teratocarcinoma forming potential after long term passaging in the absence of LIF.

A: ESΔN-iEsrrb or ESΔN-iNanog cells cultured in the presence of LIF antagonist and doxycycline for more than one month. **B**: Morphology of the teratocarcinomas recovered from animals injected with ESΔN-iNanog cells and administered or not administered 10μg/ml of doxycycline in the drinking water. **C**: Size of representative teratocarcinomas recovered from animals injected with ESΔN-iNanog or ESΔN-iEsrrb cells and examples of the tissue types used for scoring germ layer contribution. Ad: adipose tissue; B: bone; C: cartilage; Mu: muscle; K: keratinised epithelium; En: Endoderm; N: mature neural tissue; IN: Immature neural tissue.

(4 +doxycycline/ 5 –doxycycline) respectively. All the animals that survived showed evident teratomas, except for one animal injected with ESΔN-iEsrrb cells (-doxycycline group) (**Table 5.1**). There was no marked difference in size between tumours generated by injection of Esrrb or Nanog inducible cells. Similarly, treatment with doxycycline did not affect the tumour size significantly (**Table 5.1**).All tumours recovered from animals that were not administered doxycycline in the drinking water showed the presence of tissues derived from the three germ layers (**Figure 5.10C, Table 5.1**). At a first visual inspection, tumours recovered from animals injected with ESΔN-iNanog cells and administered doxycycline presented a clearly distinct morphology and tissue composition compared to untreated controls (**Figure 5.10B**). These teratomas were mainly composed of embryonal carcinoma (EC) cells and necrotic tissue. A similar difference in cell type composition could not be observed after Esrrb induction (**Table 5.1**).

Taken together these data demonstrate that ES cells overexpressing Esrrb can indefinitely self-renew in the absence of LIF, retaining the ability to differentiate into tissues derived from the three germ layers. Furthermore, they support the notion that forced Esrrb expression counteracts differentiation of ES cells, but to a lesser extent than Nanog.

5.5: Esrrb overexpression blocks neural differentiation in Nanog null ES cells

Nanog overexpression affects the ability of ES cells to undergo neural differentiation in vitro (Chambers et al., 2003). To verify whether expression of high levels of Esrrb could result in a similar phenotype, two ESΔN-CAGE cell lines that showed robust LIF independence were differentiated for 9 days in defined N2B27 medium in the presence or in the absence of BMP and LIF, fixed and stained for Tubulin III (TuJ) expression. As expected, addition of BMP to the culture medium suppressed neural differentiation of TβC44c6 cells. In contrast, a strong TuJ signal was observed for

ESΔN-iNanog

Animal number	Doxycycline	Tumour size (mm x mm)	Ectoderm	Mesoderm	Endoderm	EC
2	+	11 x 13				√
3	+	18 x 24	√			√
4	+	7 x 5				√
5	+	10 x 15	√	√		√
2	-	9 x 6	√	√	√	
3	-	17 x 17	√	√		√
4	-	18 x 20	√	√	√	√
5	-	1 x 1	√	√		

ESΔN-iEsrrb

Animal number	Doxycycline	Tumour size (mm xmm)	Ectoderm	Mesoderm	Endoderm	EC
1	+	20 x 22	√	√		√
2	+	15 x 16	√	√	√	√
3	+	11 x 16	√	√	√	√
4	+	3 x 3	√	√	√	
1	-	16 x 17	√	√	√	√
2	-	10 x 12	√	√	√	
3	-	19 x 22	√	√	√	
4	-	13 x 15	√	√	√	
5	-	No tumour				

Table 5.1: Tissue composition of ESΔN-iNanog and ESΔN-iEsrrb derived teratomas.

Table showing tumour size and tissue contribution observed after kidney capsule injection of ESΔN-iEsrrb or ESΔN-iNanog cells cultured for more than one month in the presence of LIF antagonist and 1μg/ml of doxycycline. Animals were administered or not administered 10μg/ml of doxycycline in the drinking water and tumours dissected exactly one month after injection.

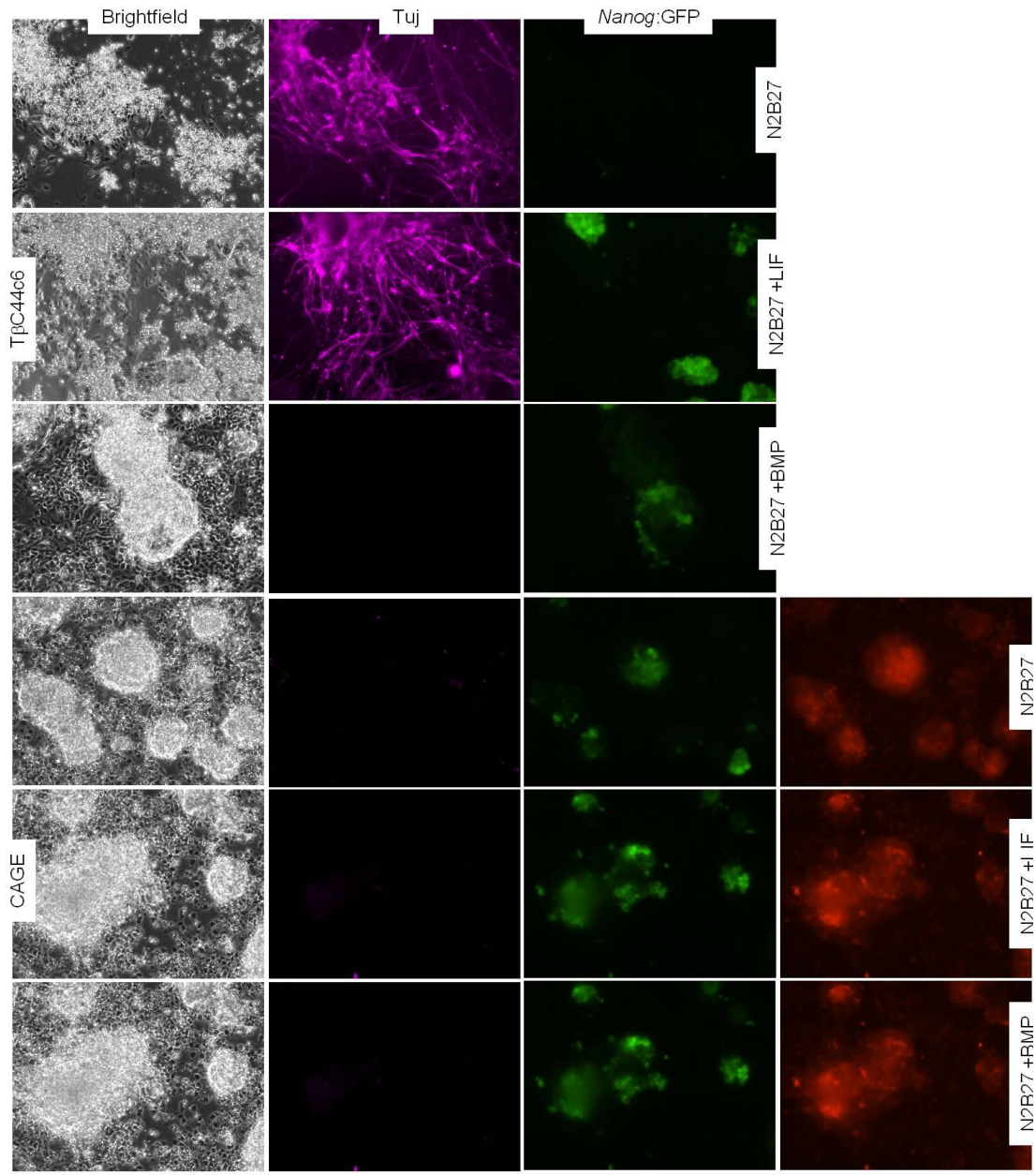
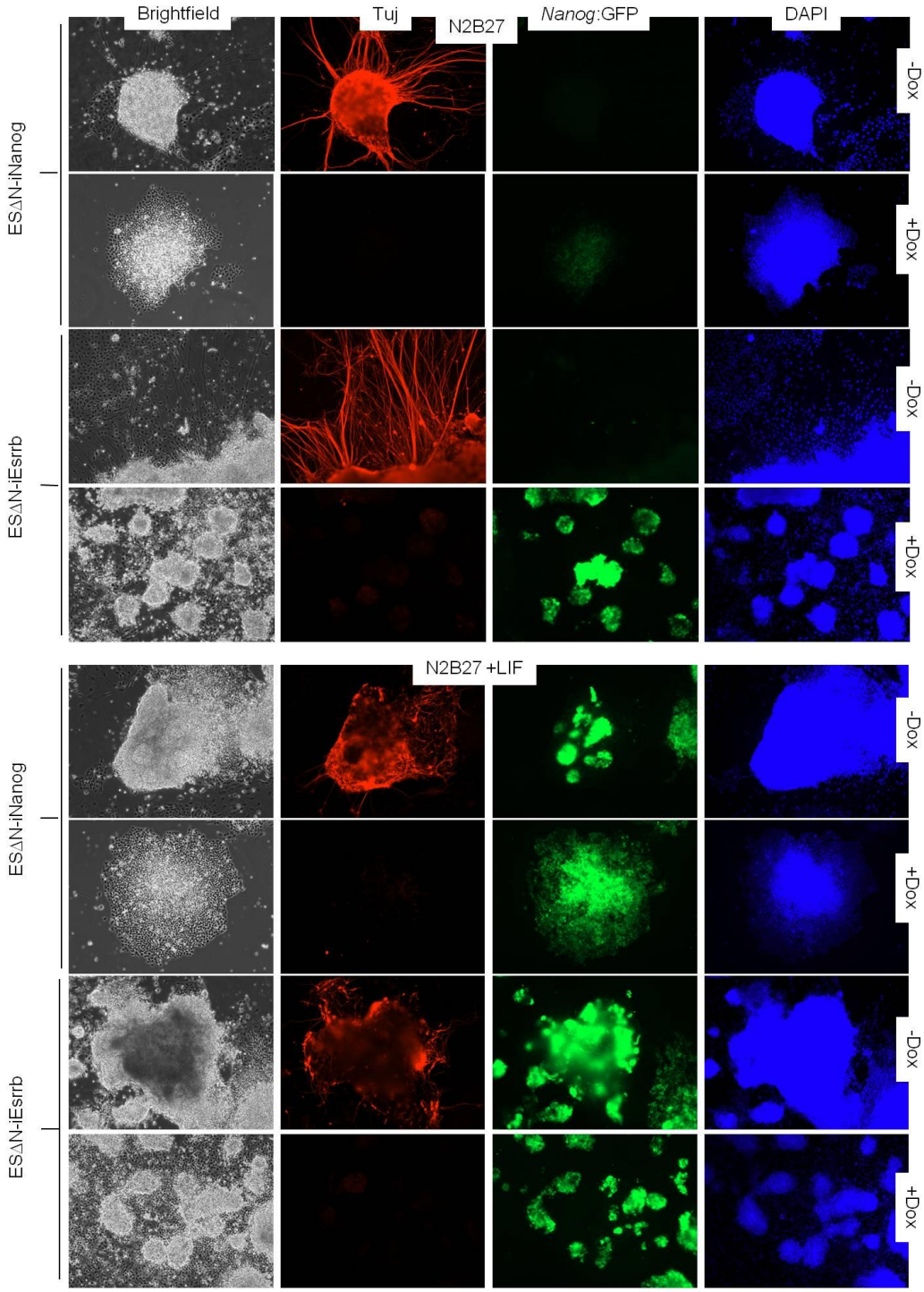


Figure 5.11: Esrrb overexpression block neural differentiation of Nanog$^{-/-}$ ES cells.

Tubulin III (TuJ), *Nanog*:GFP and TdTomato expression in TβC44c6 and ESΔN-CAGE cells after 9 days in N2B27 in the presence or absence of BMP and LIF.

parental TβC44c6 ES cells cultured in N2BB27 alone or in the presence of LIF (**Figure 5.11**). Importantly, Esrrb overexpression prevented neural differentiation of ESΔN-CAGE cells cultured in unsupplemented N2B27 (**Figure 5.11**). In accordance with its role in sustaining the transcription of Nanog (Chambers et al., 2003), LIF addition to the culture medium resulted in increased expression of the *Nanog:*GFP reporter in both Esrrb overexpressing cells and controls (**Figure 5.11**).

In order to confirm these results and compare the efficiency of Esrrb and Nanog in preventing neural differentiation, the experiment was repeated on ESΔN-iNanog and ESΔN-iEsrrb lines. Cells were plated in N2B27 with or without LIF and BMP at the density of 5 x 10^4 cells/well in 6 well plates, with the exception of ESΔN-iNanog cells cultured in N2B27 +LIF and BMP that had to be plated at 5 x10^3 cells/well to prevent overgrowth. After 9 days in culture in the presence or absence of doxycycline, cells were stained for Tubulin III (TuJ) expression and analysed by fluorescence microscopy. Overt neural differentiation was observed for both ESΔN-iNanog and ESΔN-iEsrrb cells cultured in N2B27 without doxycycline. Conversely, culture of inducible Esrrb or Nanog lines in N2B27 supplemented with doxycycline blocked neural differentiation (**Figure 5.12**). Sparse cells showing very low levels of TuJ staining could be detected for the ESΔN-iEsrrb line under these conditions, but it was not possible to identify cells which underwent productive neural differentiation. In addition, Esrrb induction resulted in robust but heterogeneous expression of the *Nanog*:GFP reporter (**Figure 5.12**). Similarly, culture in N2B27 +LIF in the absence of doxycycline resulted in neural differentiation of all lines, although with reduced efficiency. In these conditions expression of the *Nanog*:GFP reporter was retained by a substantial proportion of the cells. As observed for culture in N2B27 alone, addition of doxycycline to Esrrb and Nanog inducible lines blocked neural differentiation (**Figure 5.12**). Non-neural differentiation was observed when cells were cultured in the presence of N2B27 supplemented with BMP in the absence of doxycycline. Induction of Esrrb and Nanog expression under these condition resulted in the presence of self-renewing colonies actively expressing the *Nanog*:GFP reporter. Finally, all cell lines tested efficiently self-renewed in medium containing


Brightfield
Tuj
N2B27
Nanog:GFP
DAPI
ESΔN-iNanog
ESΔN-iEsrrb
-Dox
+Dox
-Dox
+Dox
N2B27 +LIF
ESΔN-iNanog
ESΔN-iEsrrb
-Dox
+Dox
-Dox
+Dox

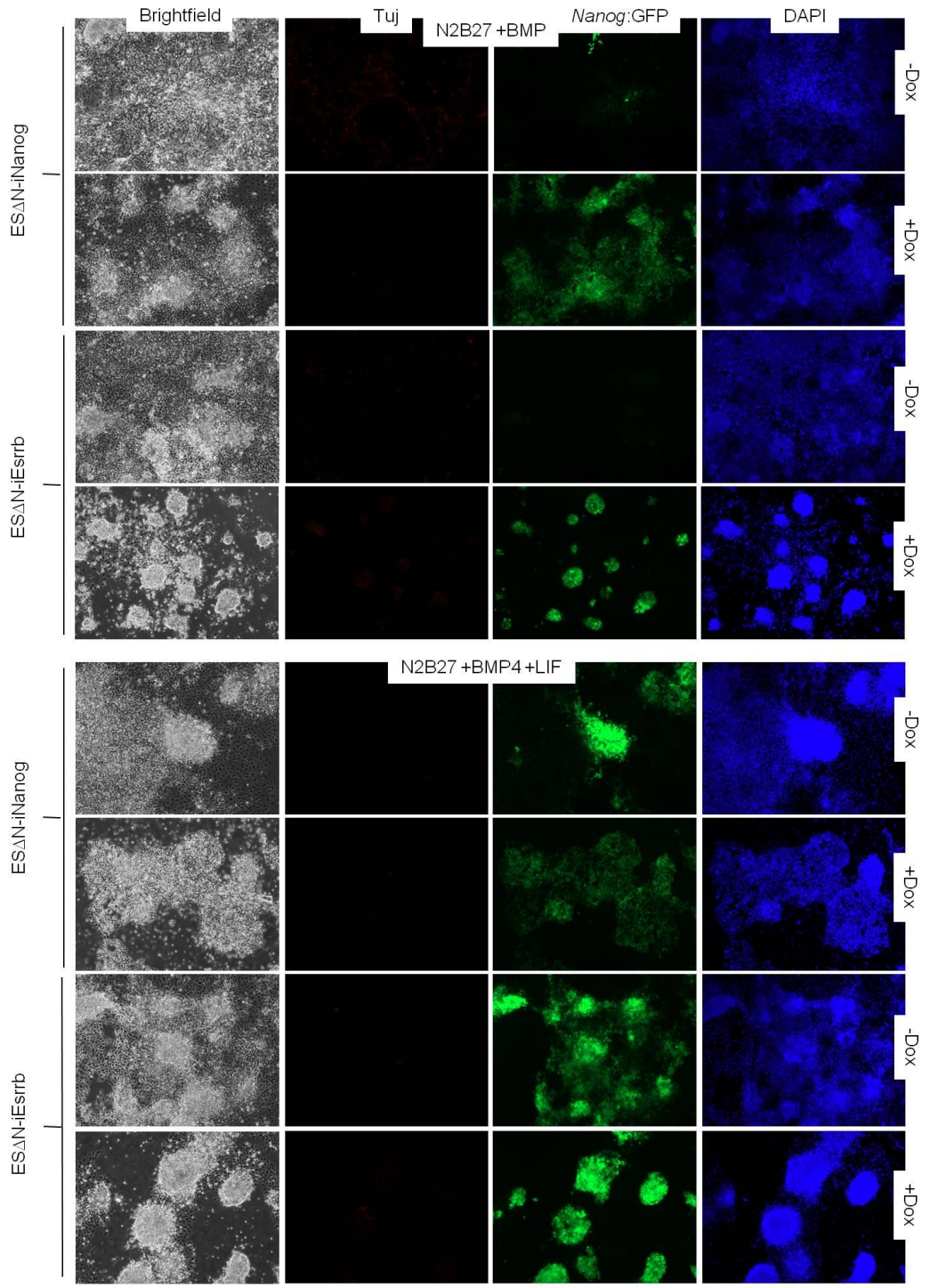


Figure 5.12: Both Nanog and Esrrb overexpression block neural differentiation of Nanog null ES cells.

Tubulin III (TuJ) and *Nanog*:GFP expression in ESΔN-iNanog and ESΔN-iEsrrb cells after 9 days in N2B27 in the presence or absence of BMP and LIF and with or without 1 μg/ml doxycycline.

LIF and BMP (**Figure 5.12**). These results show that Esrrb is able to block neural differentiation of ES cells and sustains ES cell self-renewal in the absence of BMP signalling. As observed in LIF withdrawal experiments, Esrrb ability to block ES cell differentiation is less pronounced than that observed for Nanog.

Taken together with the results from previous experiments, these observations suggest a general role for Esrrb in counteracting ES cells differentiation, a function that seems to be exerted independently of Nanog.

5.6: Esrrb sustains LIF independent self-renewal at lower doses than Nanog

The work heretofore presented identified two transcription factors that that show a prompt transcriptional activation in response to Nanog elevation: Esrrb and Klf4. Esrrb was shown to confer LIF independent self-renewal upon overexpression in ES cells. A similar ability was described for Klf4 (Niwa et al., 2009).

It was thus decided to compare the efficiency with which Nanog, Esrrb and Klf4 promote LIF independence in ES cells. To this end, RMCE was used to introduce doxycycline-inducible transgenes into the same locus of E14Tg2a cells, as already described for the derivation of ESΔN-inducible lines (**Figure5.13A**; details in experimental procedures). ES-iNanog, ES-iEsrrb or ES-iKlf4 cells were plated at clonal density with or without LIF, in increasing doxycycline concentrations. Maximal self-renewal efficiency was observed at 3μg/ml doxycycline for Nanog and Klf4, but at 1μg/ml for Esrrb. In fact, excessive Esrrb induction resulted in widespread differentiation (**Figure 5.13 B-C**). At optimal levels of induction, overexpression of Esrrb and Nanog showed comparable ability to drive LIF independence, and both surpassed Klf4.

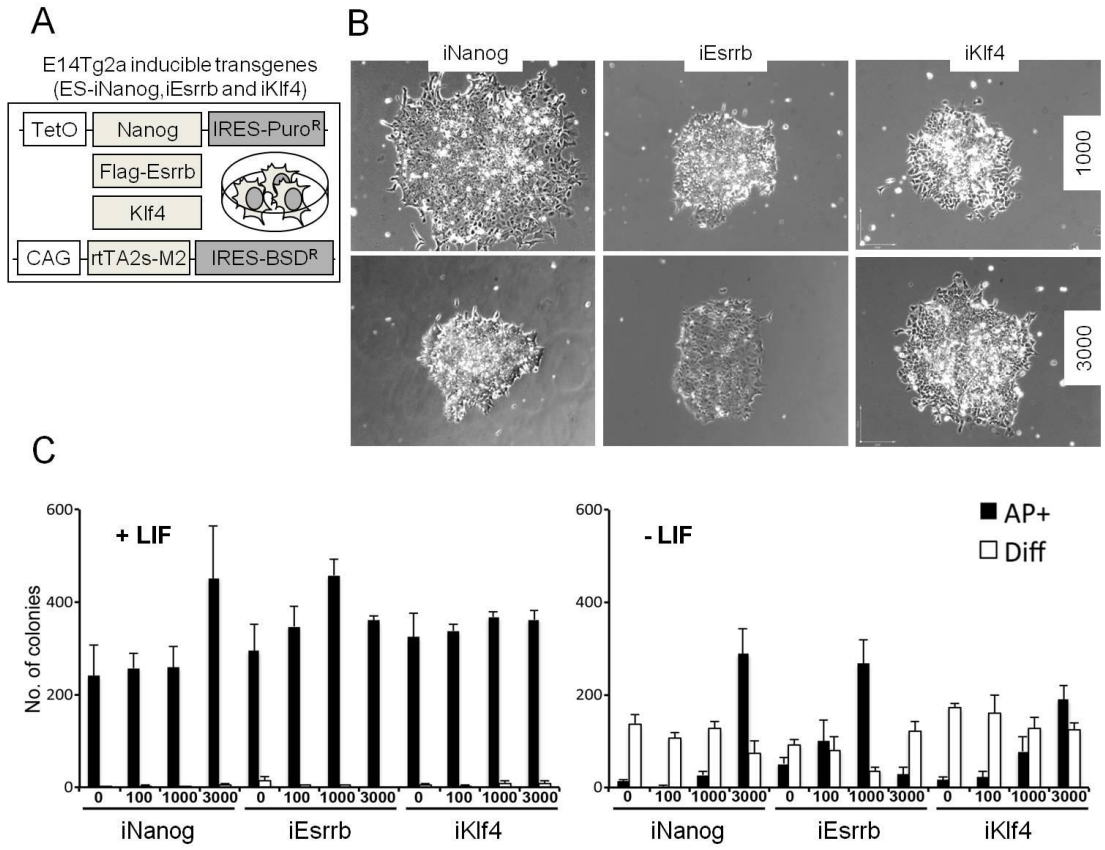


Figure 5.13: Esrrb promotes LIF independent self-renewal at lower doses than Nanog.

A: Schematic representation of the genetic manipulations used to make ES-iNanog, ES-iEsrrb or ES-iKlf4 cells. **B**: Colony morphology of ES-iNanog (iNanog), ES-iEsrrb (iEsrrb) or ES-iKlf4 (iKlf4) cells plated at clonal density and cultured in 1000 or 3000 ng/ml doxycycline and in the absence of LIF for 8 days. **C**: Numbers of colonies positive or negative for alkaline phosphatase scored after plating ES-iNanog (iNanog), ES-iEsrrb (iEsrrb) or ES-iKlf4 (iKlf4) cells at clonal density in the presence or absence of LIF and with the indicated concentration of doxycycline for 7 days. Error bars: standard deviation (n=3).

These results indicate that Esrrb exerts its function at lower doses than Nanog and that comparable effects are observed after overexpression of these two factors at appropriate levels.

5.7: Discussion

In this chapter evidence is presented that Esrrb overexpression confers to ES cells complete independence from gp130 signalling. Esrrb overexpressing cells can be passaged at clonal density in the presence of hLIF-05 and maintain the ability to contribute to adult chimaeras, a function first described for Nanog (Chambers et al., 2003).

Another target identified in this analysis is Klf4, which, like Klf2 and Tbx3, has also been reported to sustain pluripotency in the absence of LIF, although in conditions in which gp310 signalling was not completely abrogated (Hall et al., 2009; Niwa et al., 2009). The finding that *Esrrb*, *Klf4* and, less pronouncedly, *Tbx3* (**Table 3.1**) are controlled by Nanog, coupled with the notion that Esrrb and Tbx3 can positively regulate Nanog (Jiang et al., 2008; van den Berg et al., 2008) and that Klf4 is a crucial mediator of LIF signalling (Niwa et al., 2009) identifies a circuit composed of Nanog, Esrrb, Klf4 and Tbx3 that stabilizes ES cell self-renewal through positive feedback (Davidson, 2010; Oliveri et al., 2008). This circuit can integrate intrinsic cues, possibly through modulation of the prominent regulator Nanog, and extrinsic inputs from LIF signalling, through activation of Klf4, to balance self-renewal and differentiation in mouse ES cells.

Interestingly, irrespective of the presence of LIF in the cultures, Esrrb overexpression results in the accumulation of a sizable population of Nanog negative cells (**Figure 5.8A, 5.9**).This proves that Esrrb is not a strong activator of Nanog transcription and might suggest that Esrrb elevation is be able to promote self-renewal of the subpopulation of cells that have silenced Nanog expression or that Esrrb is involved in transcriptional switching of Nanog.

The results presented further show that Esrrb function in ES cells is not mediated by Nanog, since Esrrb overexpression induces robust self-renewal in *Nanog*$^{-/-}$ cells. LIF independence of *Nanog*$^{-/-}$ cells is also induced by Klf2 (Hall et al., 2009). In addition, Klf2 overexpression was reported to promote self-renewal of *Nanog*$^{+/+}$ cells in unsupplemented N2B27 (Hall et al., 2009). The present analysis indicates that, like Klf2, Esrrb can suppress differentiation in serum-free medium but remarkably this ability is conserved in cells lacking Nanog. Taken together these observations define Esrrb, together with Nanog, as the two factors for which the ability to intrinsically promote self-renewal in ES cells is best characterised. In support of this conclusion, Esrrb induces LIF independence with efficiency comparable to Nanog, and both surpass Klf4 (**Figure 5.13C**). Nonetheless, in all the experiments Esrrb overexpressing cells cultured in the absence of LIF gave rise to a fraction of colonies that had differentiated margins. This was never observed for Nanog. In addition, the teratocarcinomas produced by injecting ESΔN-iNanog cells into the kidney capsule of mice that were treated with doxycycline were predominantly composed of EC cells (**Table 5.1**). Under the same conditions, ESΔN-iEsrrb cells differentiated into tissues derived from the three germ layers. Taken together these observations suggest that Nanog is a stronger suppressor of differentiation than Esrrb, confirming Nanog at the top of the hierarchy of factors able to sustain the undifferentiated state in ES cells.

Analysing data across the entire study, it was noticed that the system employed to achieve Esrrb and Nanog overexpression strongly influenced the experimental outcome. Whereas Nanog performed better than Esrrb in episomal overexpression experiments (**Figure 5.1A**), comparison of cells in which both factor are induced at similar levels (**Figure 5.13C**), showed that Esrrb confers LIF independence with similar efficiency and at lower doses than Nanog. Stable integration experiments confirmed that any difference in the effect of Nanog or Esrrb elevation is strictly dependent on the level of expression achieved. EF4, an ES cell line in which Nanog mRNA and protein levels are increased 6 folds (See **Figure 3.9A** and (Yates and Chambers, 2005)), showed more robust LIF independence than EfEsrrb cells (**Figure 5.1E**), overexpressing Esrrb 3-fold over wild-type levels (**Figure 5.1C**). Conversely,

a comparable ability to self-renew in the absence of LIF (see **Figure 8.5C**) was noticed in 6 stably transfected lines overexpressing either Nanog or Esrrb transgenes to similar levels (average 6 fold against 4 fold, see **Figure 8.4A-B**). Interestingly, in all the lines derived Esrrb mRNA overexpression was never higher than 4-5 fold over WT levels. It seems therefore that ES cells can tolerate higher levels of Nanog overexpression than Esrrb, as confirmed by experiments performed in doxycycline inducible lines (**Figure 5.13C**). Preliminary data investigating the effects of elevation of other Nanog targets (Jing Chao Zhang, unpublished data) revealed that a limited tolerance for overexpression might be a common characteristic of a significant proportion of these genes. Given the toxicity observed after excessive Esrrb elevation and the differences in the expression levels required to achieve maximum effect in overexpression experiments, it would be interesting to directly compare the absolute levels of Nanog and Esrrb mRNAs and proteins in ES cells.

The work presented in this chapter is based on the derivation of a doxycycline inducible system that permits comparison of the effects of overexpressing different genes at the same level and from an identical genomic location (**Figure 5.6A**). This system was designed after observing that the vast majority of the clones analysed during the derivation of doxycycline inducible lines that expressed rtTA from *Rosa26* presented a vastly heterogeneous expression pattern. This phenomenon could be due the low levels of rtTA expression achieved in this system. This observation is relevant to the interpretation of the results generated in reprogramming studies that employed inducible transgenes also based on expression of rtTA from *Rosa26* (Hanna et al., 2008; Woltjen et al., 2009). In alternative heterogeneous expression might be entirely dependent on the genomic location in which integration of the doxycycline inducible transgenes occurred. The use of RMCE to introduce the ORFs of any gene of interest into a previously identified genomic location that ensures homogeneous expression circumvented this problem. High and homogeneous expression is achieved in ESΔN-iTdT lines (**Figure 5.6C**) and ESΔN-iNanog derivative clones (**Figure 5.7B**). Relevant to this point, when investigating transgene induction levels in doxycycline titration experiments on *Rosa*:rtTA TdTomato-2a-Esrrb ES cells (**Figure 5.4A-B**), it was noticed that expression at low concentrations

of doxycycline tended to be more heterogeneous (**Figure 5.4C**). This should be considered when applying this system to genes that require low levels of overexpression. Nonetheless, overall transgene expression in these lines can be finely tuned (**Figure 5.4A-B**), presenting undeniable advantages when compared to constitutive expression systems. Nanog was originally isolated by episomal overexpression of an ES cell derived cDNA library (Chambers et al., 2003). It would be interesting to apply the described RMCE transgene exchange system to the screening of similar cDNA libraries, avoiding the toxicity effects observed after episomal transgene expression and possibly broadening the scope for identification novel pluripotency factors. The feasibility of such an approach will depend principally on the efficiency of RMCE in these lines. Possibly this system should be employed to screen libraries of limited size including a reduced number of previously selected candidate genes.

Chapter 6: Esrrb drives reprogramming in the absence of Nanog

Nanog exerts a fundamental role in driving the specification and the maintenance of different pluripotent cell populations during embryonic development. Nanog null embryos fail to develop a structured epiblast, lose expression of Oct4 and arrest development before implantation (Mitsui et al., 2003; Silva et al., 2009). A clue to understand how Nanog promotes the transition to pluripotency may come from in vitro reprogramming experiments. Nanog is able to increase the efficiency of reprogramming (Silva et al., 2006). Strikingly, cells lacking Nanog expression cannot be reprogrammed to a pluripotent state by fusion with ES cells or overexpression of specific transcription factors. The reprogramming process seems to stall in an intermediate state, in which cells are dependent of feeders for propagation and fail to reactivate transcription from endogenous pluripotency alleles, such as Oct4, Nanog, Rex1 and Klf2 (Silva et al., 2009). Using in vitro reprogramming as a model, it is possible to test whether Nanog exerts a unique role in driving attainment of pluripotency or if other factors can compensate for its absence.

6.1: Esrrb enhances reprogramming by cell fusion

Esrrb has been shown to promote in vitro reprogramming of mouse embryonic fibroblasts (MEFs) to induced pluripotent stem cells (iPS) in conjunction with Oct4 (Feng et al., 2009). Recent studies showed that Esrrb can even drive reprogramming in the absence of the canonical reprogramming factors Oct4, c-Myc, Klf4 and Sox2. The results presented in chapter 3 and 4 indicate that Nanog regulates Esrrb expression. Thus, it was important to test whether Nanog enhances reprogramming efficiency through activation of Esrrb and whether Esrrb elevation renders Nanog function dispensable in this process.

Polyethylene glycol mediated cell fusion has been employed to test the ability of Nanog null neural stem (NS) cells to successfully undergo reprogramming, and Nanog levels in the ES cell fusion partner have been shown to correlate with the efficiency of this process (Silva et al., 2006; Silva et al., 2009). It was thus decided to ascertain if Esrrb overexpression could promote reprogramming of NS cells by cell fusion. The efficiency of formation of pluripotent hybrid colonies was compared following fusion of E14/T NS cells with E14Tg2a, EF4 or EfEsrrb ES cells. Interestingly, Esrrb overexpression stimulated the formation of pluripotent hybrid colonies with a similar efficiency as that observed for Nanog (**Figure 6.1A-B**).

6.2: Esrrb overexpression allows the generation of pluripotent hybrids from Nanog null NS cells.

It was then determined whether Nanog activity was strictly required for the successful generation of pluripotent hybrids by developing an experimental system in which Nanog null NS cells are fused to Nanog null ES cells overexpressing Esrrb. NS cells were derived from the RCNβH(t) ES cell line by neural differentiation in monolayer (Pollard et al., 2006). RCNβH(t) NS cells show the characteristic morphology of NS cells, can be readily propagated in NS cell expansion medium in the presence of FGF and EGF (Conti et al., 2005) and show positive staining for Nestin by immunohistochemistry (**Figure 6.2A**).

During the reprogramming process the transcription of pluripotency genes is activated from the NS cell genome as soon as 1 day after cellular fusion (Bhutani et al., 2010; Han et al., 2008). In RCNβH(t) cells a hygromycin resistance cassette is placed under the control of the endogenous *Nanog* promoter (**Figure 6.2B**). Thus, the use of RCNβH(t) NS cells in fusion experiments permitted the selection of cellular hybrids that reactivated transcription of Nanog from the NS cell genome by culture in the presence of hygromycin. Nanog null RCNβH(t) NS cells were fused to two different ESΔN-CAGE lines (clone 17 and clone 21). As a control, RCNβH(t) NS cells were fused to Nanog and Esrrb overexpressing ES cells (EF4 and EfEsrrb), or

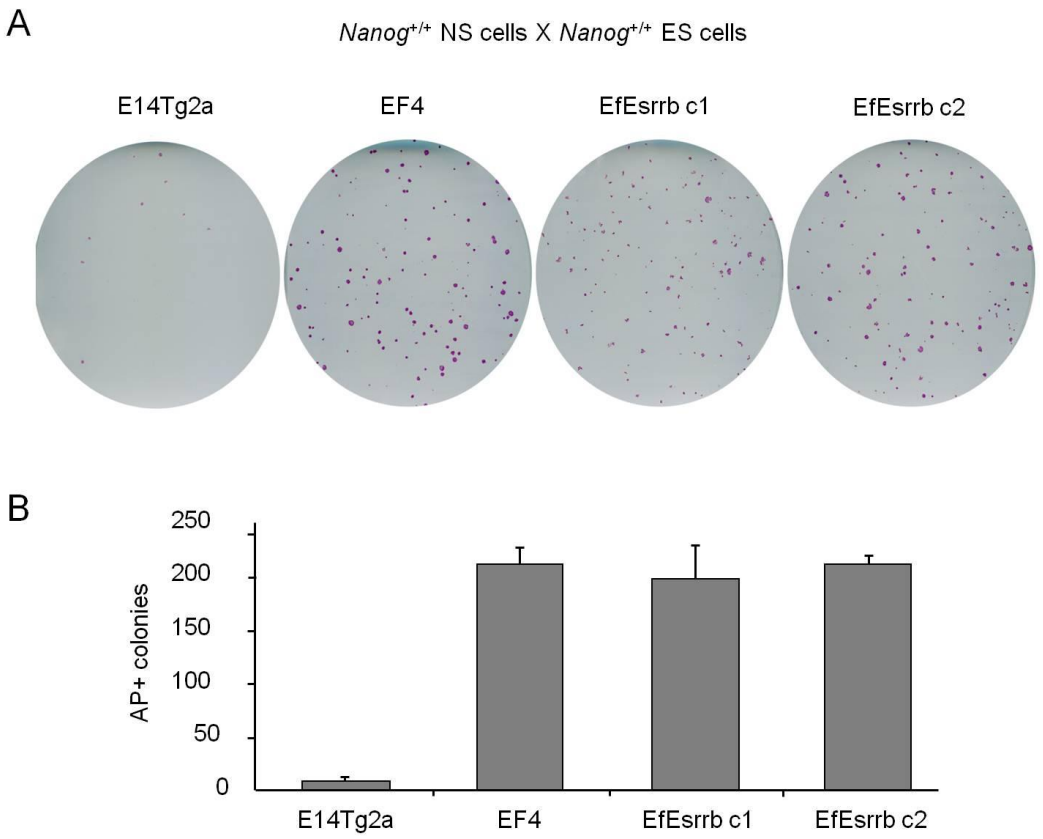


Figure 6.1: Esrrb promotes reprogramming by cell fusion.

A: Representative plates from ES x NS fusion experiments performed using E14/T NS cells and E14Tg2a, EF4 (Nanog overexpressing), or EfEsrrb (Esrrb overexpressing) ES cells. ES x NS hybrids were cultured for 12 days in puromycin and neomycin selection prior to staining for alkaline phosphatase. **B**: AP positive colony numbers/million NS fused observed in the above experiment; error bars: standard deviation (n=4).

The data presented in this figure were kindly provided by Adam Yates and Douglas Colby.

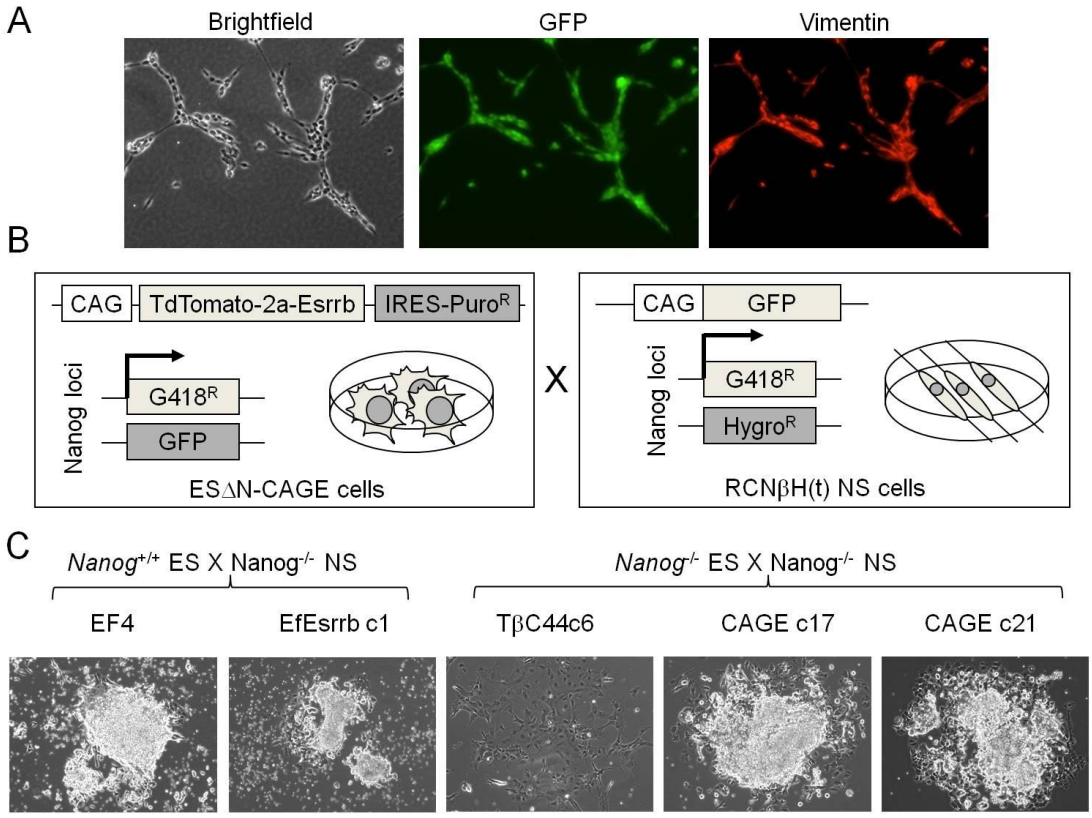


Figure 6.2: Esrrb overexpression allows the generation of pluripotent hybrids from Nanog null NS cells.

A: Vimentin and GFP expression in RCNβH(t) NS cells. **B**: Schematic representation of two *Nanog*$^{-/-}$ lines used for fusion experiments: ESΔN-CAGE cells and RCNβH(t) NS cells. **C**: Hybrid colonies formed by fusion of RCNβH(t) NS cells with the indicated *Nanog*$^{+/+}$ or *Nanog*$^{-/-}$ ES cell lines. Hybrids were cultured for 14 days; puromycin was applied after 1 day; hygromycin application was delayed until day 4 to allow reprogramming of the NS genome.

to E14Tg2a and TβC44c6 ES cells expressing a GFP-IRES-Puromycin transgene. One day after fusion, puromycin selection was applied to eliminate unfused NS cells from the culture. Four days after fusion, hygromycin selection was applied to negatively select ES cells which had not undergone fusion and cell hybrids were cultured for additional 10 days prior to colony scoring. As expected, cell fusion in the presence of functional Nanog protein gave rise to a sizable numbers of morphologically undifferentiated, alkaline phosphatise positive hybrid colonies (**Table 6.1, Figure 6.2C**). No undifferentiated colonies were observed after fusion of RCNβH(t) NS cells with TβC44c6 ES cells. The few Nanog null hybrid colonies observed after selection differentiated during the course of the experiment (**Figure 6.2C**). In contrast, appearance of undifferentiated colonies could be observed for fusions between ESΔN-CAGE ES and RCNβH(t) NS cells (**Figure 6.2C**, **Table 6.1**). Nanog null Esrrb overexpressing hybrids could be continuously passaged under standard ES cell culture conditions without showing signs of differentiation.

Since marked cellular death was observed after the combined addition of puromycin and hygromycin to the culture medium, the experiment was repeated without puromycin selection. No colonies were observed after RCNβH(t) NS x RCNβH(t) NS control fusions, showing that puromycin selection against unfused NS cells was not required. In accordance with the previous experiment, the only Nanog null hybrid colony originating from fusion between TβC44c6 ES x and RCNβH(t) NS cells underwent differentiation. The exclusion of puromycin selection resulted in an increased number of hybrids colonies from RCNβH(t) NS x ESΔN-CAGE ES cell fusions, compared to the previous experiment (**Table 6.1**). Nonetheless, reprogramming by cell fusion in the absence of Nanog showed a significantly compromised efficiency. As observed before, the hybrid colonies could be passaged in ES cell medium for several passages without apparent signs of differentiation.

Taken together, these results support the idea that Nanog expression strongly favours, but is not strictly required, during NS cell reprogramming by cell fusion. Expression of Esrrb can compensate for the lack of Nanog in supporting the transition of hybrid cells to pluripotency.

Experiment 1

RCNβH(t) NS X	Hybrids/10^6 cells fused
Nanog (EF4)	453
EfEsrrb c1	133
TβC44c6	0
ESΔN-CAGE c17	1.6
ESΔN-CAGE c21	2.5

Experiment 2

RCNβH(t) NS X	Hybrids/10^6 cells fused
NS RCNβH(t)	0
TβC44c6	0
TβC44	1
ESΔN-CAGE c17	4.25
ESΔN-CAGE c21	32.5

Table 6.1: Efficiency of Nanog null NS cells reprogramming.

Number of AP positive colonies formed after fusion of RCNβH(t) NS cells with the indicated ES cell lines in two independent experiments (left panel - right panel). Hybrids were cultured for 14 days; puromycin and hygromycin selections were applied from day 1 and day 4 respectively.

It was then necessary to establish that the reprogramming of Nanog null cells driven by Esrrb overexpression was complete and stable. With this aim, fusion experiments were repeated employing ESΔN-iNanog and ESΔN-iEsrrb ES lines (**Figure 6.3A**). RCNβH(t) NS cells were transfected with a CAG driven TdTomato-IRES-hygromycin^R transgene and clones selected for high and homogeneous expression of the fluorescent protein by flow cytometry. In contrast with previous experiments, presence of a constitutively expressed hygromycin resistance gene allowed for immediate selection of primary hybrids independently of reactivation of Nanog transcription from the NS cell genome. RCNβH(t) Red NS cells were fused with ESΔN-iNanog or ESΔN-iEsrrb ES cells (**Figure 6.3A**) and primary hybrids replated in GMEM/FCS/LIF. Two days after fusion blasticidin and hygromycin selections were applied and cell hybrids cultured for additional 14 days prior to colony scoring. Fused cells cultured in the absence of doxycycline failed to form self-renewing colonies. The few colonies initially observed, progressively differentiated over the course of the experiment and could not be expanded (**Table 6.2**). In contrast Nanog and Esrrb induction resulted in the formation of self-renewing colonies (**Figure 6.3B**) that stained strongly for alkaline phosphatase and expressed TdTomato, which marks the NS cell genome (**Figure 6.3D**). While restoring Nanog expression resulted in the formation of 300 reprogrammed hybrid colonies/10^6 cells fused, Esrrb induction in ESΔN-iEsrrb ES cells could reprogram NS cells with a 12 fold lower efficiency (25 hybrid colonies/10^6 cells fused: **Table 6.2**). Nonetheless, Esrrb reprogrammed hybrid lines could be expanded and cultured over multiple passages, when neomycin selection for active transcription from *Nanog* was applied (**Figure 6.1C**). It was then assessed whether constant expression of the Nanog or Esrrb transgenes was necessary to maintain the reprogrammed lines undifferentiated. Cells were released from doxycycline and passaged 3 times in the presence or in the absence of neomycin. ESΔN-iEsrrb derived hybrid lines could be serially passaged in the absence of doxycycline when neomycin selection was maintained (**Figure 6.1C**). Hybrid cells showed increasing signs of differentiation after withdrawal of neomycin selection, as observed for Nanog$^{-/-}$ ES cells. This propensity was eliminated by Esrrb induction, underlining the ability of Esrrb to compensate for the loss of Nanog.

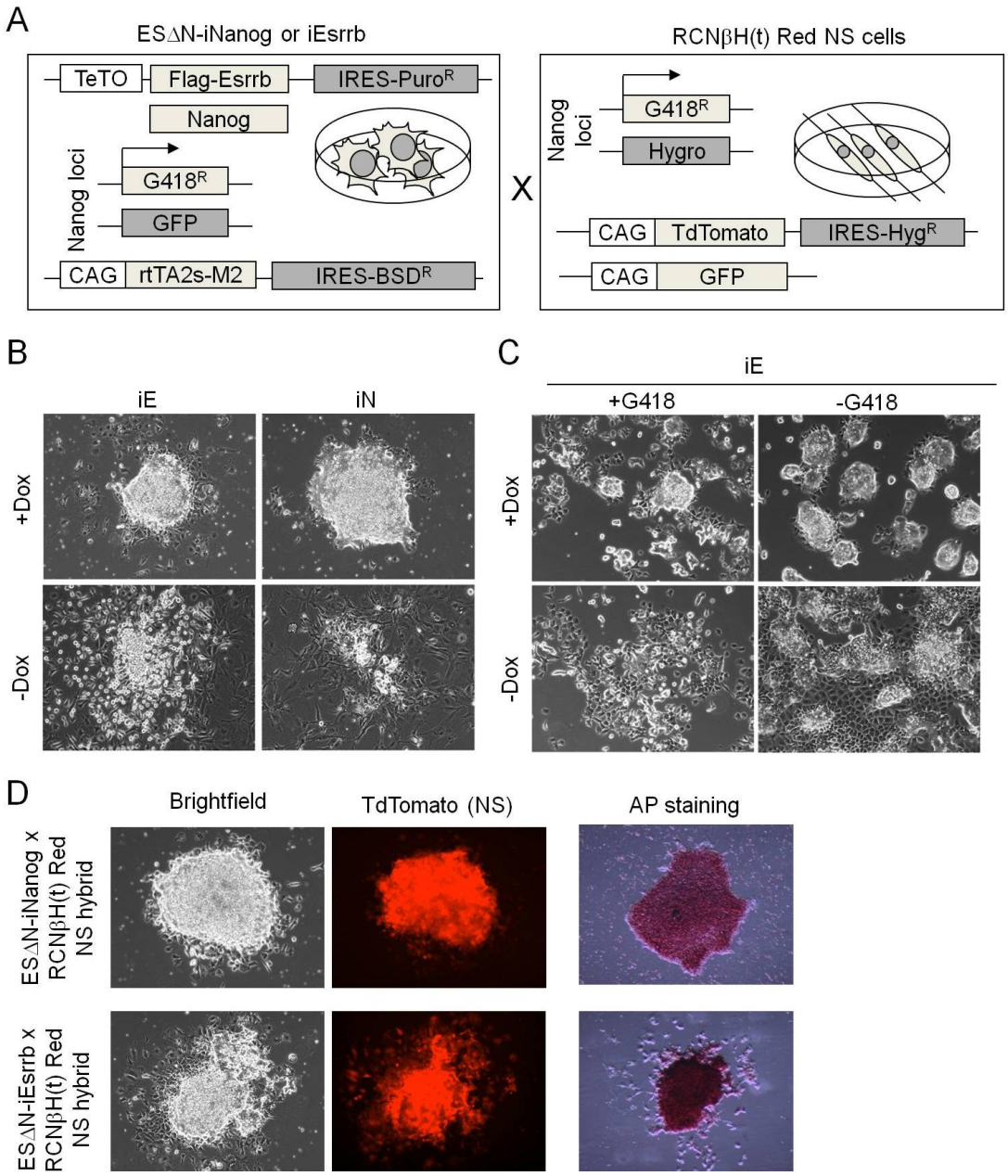


Figure 6.3: Reprogramming by cell fusion in the absence of Nanog generates stable lines.

A: Schematic representation of the ES and NS cell lines used in the fusion experiments: ESΔN-iNanog and ESΔN-iEsrrb ES cells and RCNβH(t) Red NS cells. **B**: Morphology of RCNβH(t) Red NS cells X ESΔN-iNanog or ESΔN-iEsrrb ES cells primary hybrid colonies. Blasticidin and hygromycin selections were applied two days after fusion and cells cultured for additional 14 days in the presence or absence of doxycycline. **C**: RCNβH(t) Red NS cells X ESΔN-iEsrrb ES cells hybrid lines cultured in the presence or absence of doxycycline and G418 for 3 passages. **D**: TdTomato expression (marking the NS cell genome) and alkaline phosphatase staining of ESΔN-iNanog or ESΔN-iEsrrb cells x RCNβH(t) Red NS hybrid colonies confirms productive fusion.

ESΔN-iNanog or ESΔN-iEsrrb
x
RCNβH(t) Red NS plating

Fusion	Number of NS cells fused	AP+ Colonies (± st dev.)	AP+ colonies/10^6 cells fused (± st dev.)
ESΔN-iNanog +Dox	4,000,000	**1215** (± 588)	**303.7** (± 147.1)
ESΔN-iNanog -Dox	4,000,000	6 (± 8)	1.6 (± 2.1)
ESΔN-iEsrrb +Dox	4,000,000	**99** (± 76)	**24.8** (± 19.1)
ESΔN-iEsrrb -Dox	4,000,000	17 (± 27)	4.2 (± 6.7)

Table 6.2: Efficiency of Nanog null NS cells reprogramming by Nanog or Esrrb induction.

Number of alkaline phosphatase positive colonies scored after fusion of to RCNβH(t) Red NS cells to ESΔN-iNanog or ESΔN-iEsrrb cells and culture in the presence or absence of doxycycline for 16 days. Hygromicin and blasticidin were applied 2 days after fusion.

The stability of RCNβH(t) NS cells reprogramming was further confirmed by characterising the gene expression profile of hybrid lines cultured in the presence or absence of doxycycline or G418. Reprogrammed hybrid had lost NS cell specific markers and reactivated endogenous pluripotency genes in all lines analysed (**Figure 6.4A**). Activation of transcription of *Nanog* from the NS cell genome was confirmed by PCR analysis on cDNA prepared from ESΔN-iNanog and ESΔN-iEsrrb derived hybrid lines using primer pairs binding specifically to the hygromycin targeted allele present in RCNβH(t) cells (**Figure 6.4B**). Importantly, release from doxycycline did not lead to NS cell gene reactivation or silencing of pluripotency genes (**Figure 6.4A**). In addition, both Nanog and Esrrb reprogrammed hybrids could be serially passaged without doxycycline in 2i/LIF, a condition permissive only for completely undifferentiated cells (**Figure 6.4C**). Hybrid lines released from doxycycline and G418 showed elevated GATA6 expression, in line with the acquisition of primitive endoderm morphology (**Figure 6.4A, 6.3C**).

Taken together these results suggest that the reprogramming of hybrids generated by Esrrb elevation in the absence of Nanog is stable, since their self-renewal does not dependent on continued transgene expression.

The experiments presented showed that reprogramming of $Nanog^{-/-}$ cells by Esrrb proceeds with a reduced efficiency compared to that observed after Nanog induction. It was then necessary to exclude that the observed differences were attributable to variations in the fusion efficiency among the different lines. ESΔN-iNanog or ESΔN-iEsrrb ES cells were stably transfected with a CAG -EBFP expression vector. Fusion experiments were performed as before but, 24 hours after replating, primary hybrids were purified by FACS sorting on the basis of their concomitant EBFP and TdTomato expression (**Figure 6.5**). After measuring the purity of the sorted population, primary hybrids were replated in complete ES cell medium and two days after blasticidin and hygromycin selections were applied. Cells were cultured for additional 14 days prior to scoring the number of AP positive colonies. Forced Nanog expression resulted in 1.2% of primary hybrids successfully undergoing

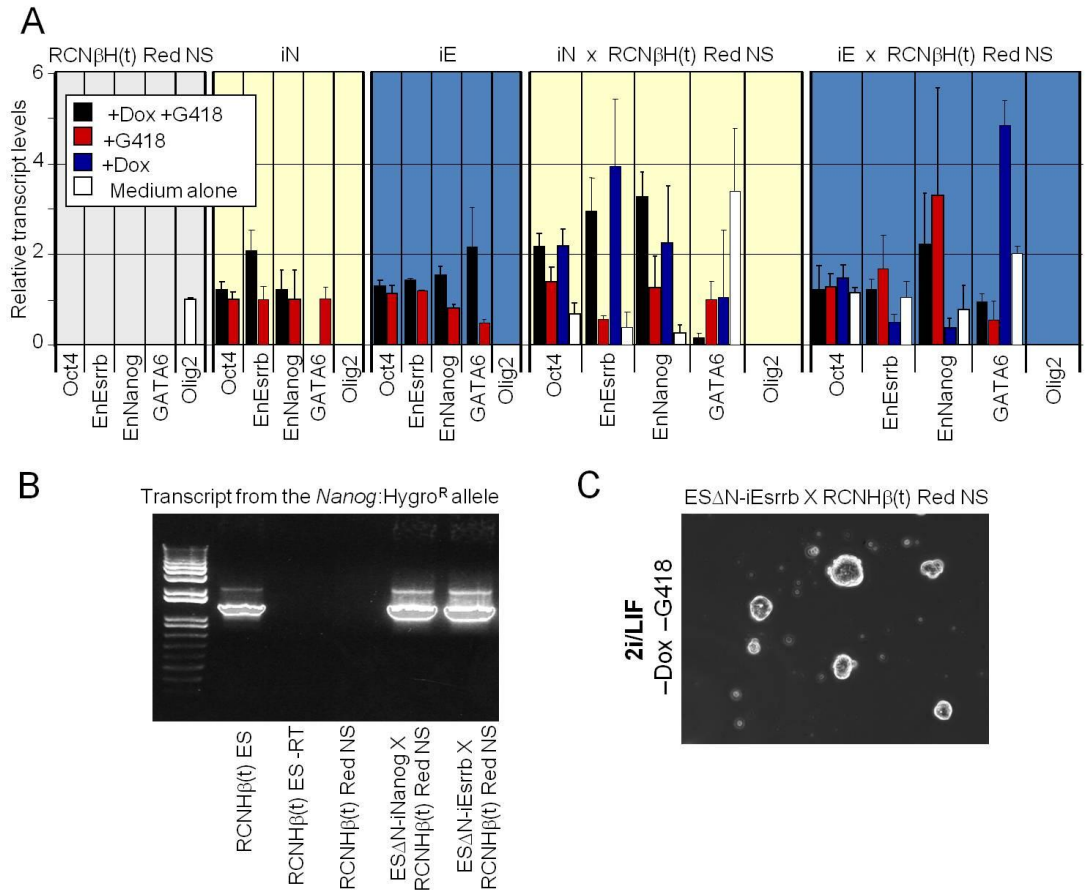


Figure 6.4: Transcriptional resetting of the NS cell genome in Nanog null NS X ES hybrid lines.

A: Gene expression profiles of RCNβH(t) Red NS cells, ESΔN-iNanog (iN) or ESΔN-iEsrrb (iE) cells, and NS X ES hybrid lines after 3 passages in the indicated conditions. Primers do not detect doxycycline inducible transgenes. Nanog primers bind to the *Nanog* 5'UTR which remains in all targeted alleles. Transcript levels normalised to TBP and relative to expression in RCNβH(t) Red NS (for Olig2) or ESΔN-iNanog cells cultured in G418 (for all other genes). Error bars: ES x NS hybrids: standard deviation of gene expression in three independent lines. ES and NS lines: standard deviation of gene expression in two independent experiments. **B**: RT-PCR amplifying transcript from the *Nanog*:hygromycin^R targeted allele on cDNA prepared from RCNβH(t) ES, RCNβH(t) NS cells and ESΔN-iNanog or ESΔN-iEsrrb derived hybrids. **C**: RCNβH(t) NS cells X ESΔN-iEsrrb ES cells hybrids cultured in N2B27/2i/LIF in the absence of doxycycline and G418.

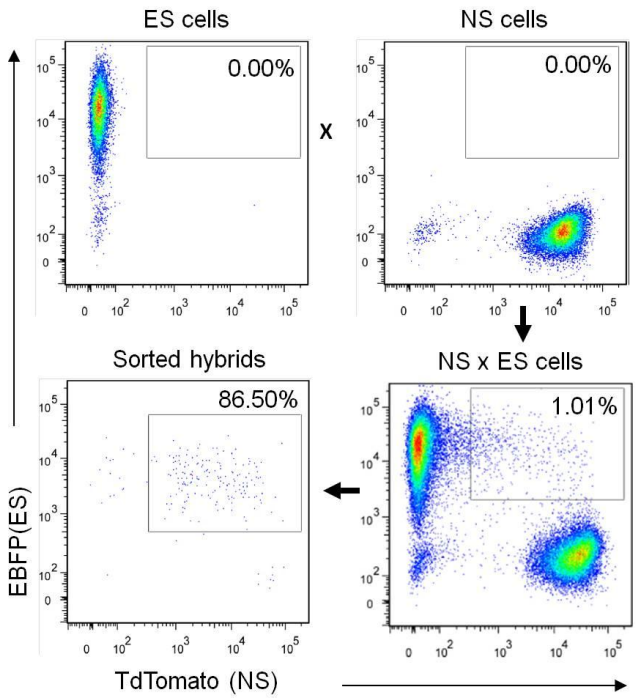


Figure 6.5: Sorting of RCNβH(t) Red NS cells X ESΔN-iNanog Blue or ESΔN-iEsrrb Blue ES cells primary hybrids.

ESΔN-iNanog or ESΔN-iEsrrb cells were transfected with a CAG-EBFP expression vector and clones showing high EBFP expression were selected. RCNβH(t) Red NS X ESΔN-iNanog Blue or ESΔN-iEsrrb Blue ES cells primary hybrids were purified by FACS sorting 24h after fusion on the basis of EBFP and TdTomato expression. Dot plots showing TdTomato and EBFP expression in ES cells, NS cells, in the cultures after fusion and in primary hybrids after sorting. The gating strategy used for sorting is shown along with an example of the typical fusion efficiency and sorting purity achieved in the experiments.

reprogramming (**Table 6.3**). As previously observed, Esrrb overexpression resulted in a 15 times lower reprogramming efficiency (0.07%, **Table 6.3**).

Taken together these data show that Esrrb can functionally compensate for Nanog in promoting NS cell reprogramming.

6.3: Esrrb rescues stalled reprogramming of Nanog null pre-iPS

Reprogramming by cell fusion proceeds with high efficiency and provides a useful model to investigate the potential of individual factors in promoting the transcriptional and epigenetic changes that accompany the transition to pluripotency. Nonetheless, this system generates tetraploid hybrids, precluding the possibility of rigorously testing the developmental potential of reprogrammed cells and conclusively prove acquisition of pluripotency.

To overcome these limitations, the ability of Esrrb to substitute for Nanog was tested during reprogramming by transcription factor overexpression (Takahashi and Yamanaka, 2006). Nanog is strictly required for completion of this process with *Nanog*$^{-/-}$ cells stalling in an intermediate state, termed pre-iPS, in which they acquire the morphology and growth factor dependence of ES cells but do not express endogenous pluripotency genes nor silence retroviral transgene expression (Silva et al., 2009). NS cells were generated in vitro from ESΔN-iNanog and ESΔN-iEsrrb ES cells and passaged 10 times in NS medium to ensure complete differentiation. These lines express the NS cell marker Olig2 and Sox2 but have completely silenced expression of other pluripotency factors (**Figure 6.7A**). Next, ESΔN-iNanog and ESΔN-iEsrrb cells were infected with retroviral vectors encoding Oct4, Klf4, c-Myc and dsRed. The inclusion of a dsRed fluorescent reporter under the control of viral LTRs permitted monitoring of the silencing of pMX transgenes upon completion of reprogramming (**Figure 6.6A**). Five days after infection, numerous colonies resembling pre-iPS cells emerged in the plates. These colonies could be passaged

ESΔN-iNanog or ESΔN-iEsrrb
x
RCNβH(t) Red NS

Sorted hybrids plating

Fusion	Number of primary Hybrids	AP+ colonies	Purity	Reprogramming frequency
ESΔN-iNanog replica 1	24,000	360	85%	1.76%
ESΔN-iNanog replica 2	14,000	222	82%	1.93%
Total				**1.82%**
Fusion	**Number of primary Hybrids**	**AP+ colonies**	**Purity**	**Reprogramming frequency**
ESΔN-iEsrrb replica 1	7659	6	62%	0.126%
ESΔN-iEsrrb replica 2	8763	3	45%	0.076%
ESΔN-iEsrrb replica 3	4872	2	62%	0.066%
ESΔN-iEsrrb replica 4	8042	1	66%	0.019%
ESΔN-iEsrrb replica 5	7359	4	51%	0.107%
ESΔN-iEsrrb replica 6	9946	3	50%	0.06%
Total				**0.074%**

Table 6.3: Efficiency of reprogramming in sorted RCNβH(t) Red NS cells X ESΔN-iNanog Blue or ESΔN-iEsrrb Blue ES cells primary hybrids.

A: Number of alkaline phosphatase positive colonies scored after culturing sorted RCNβH(t) Red NS cells X ESΔN-iNanog Blue or ESΔN-iEsrrb Blue ES cells primary hybrids for 16 days in blasticidin, hygromycin and doxycycline. The table shows the sorting purity, the numbers of primary hybrids plated and AP positive colonies counted in each independent experiment, along with the efficiency of NS cells reprogramming.

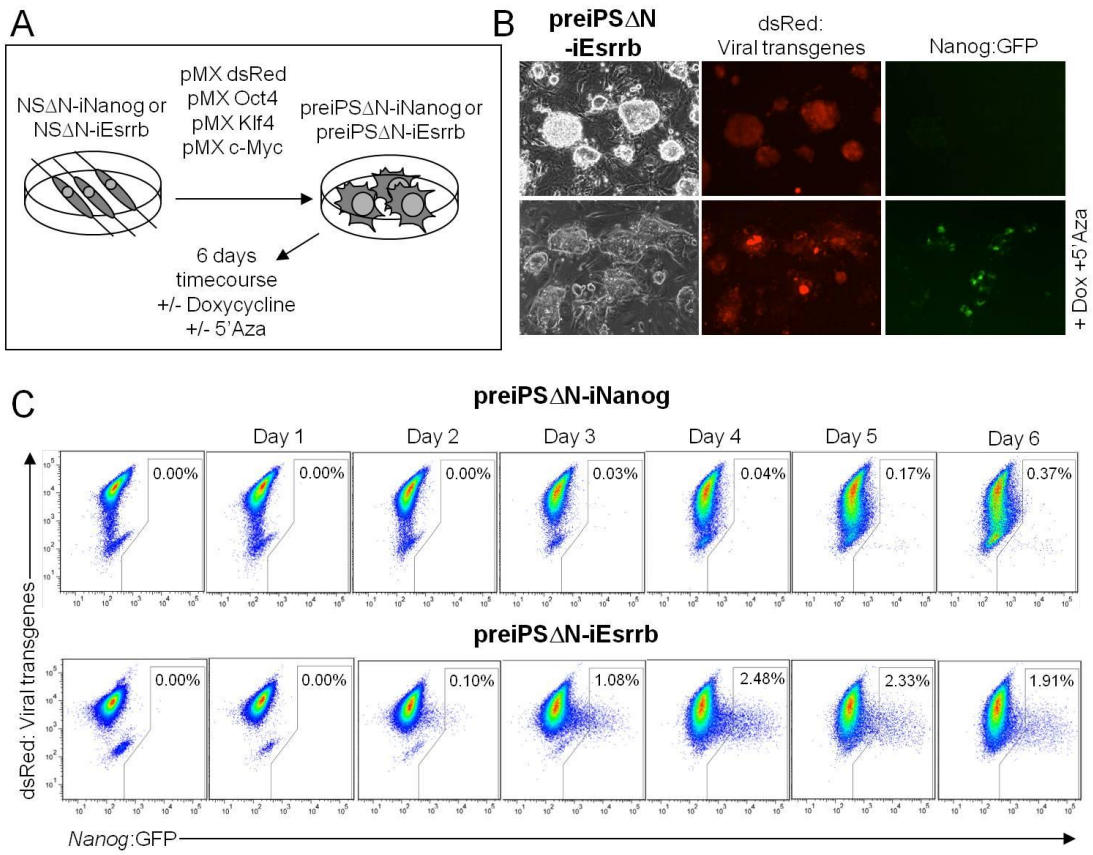


Figure 6.6: Esrrb can rescue the stunted reprogramming of *Nanog*$^{-/-}$ pre-iPS cells.

A: Experimental scheme used to derive pre-iPSΔN cells and to induce completion of reprogramming upon doxycycline induction. **B**: Morphology and *Nanog*:GFP expression after culture of pre-iPSΔN-iEsrrb cells in the absence or presence of doxycycline and 5'-azacytidine for 3 days. **C**: Dot plots showing dsRed (viral transgenes) and *Nanog*:GFP expression in pre-iPSΔN-iNanog (iN) or pre-iPSΔN-iEsrrb (iE) cells treated with doxycycline and 5'-azacytidine for the indicated number of days. The percentages of *Nanog*:GFP^{+} cells are shown for each plot.

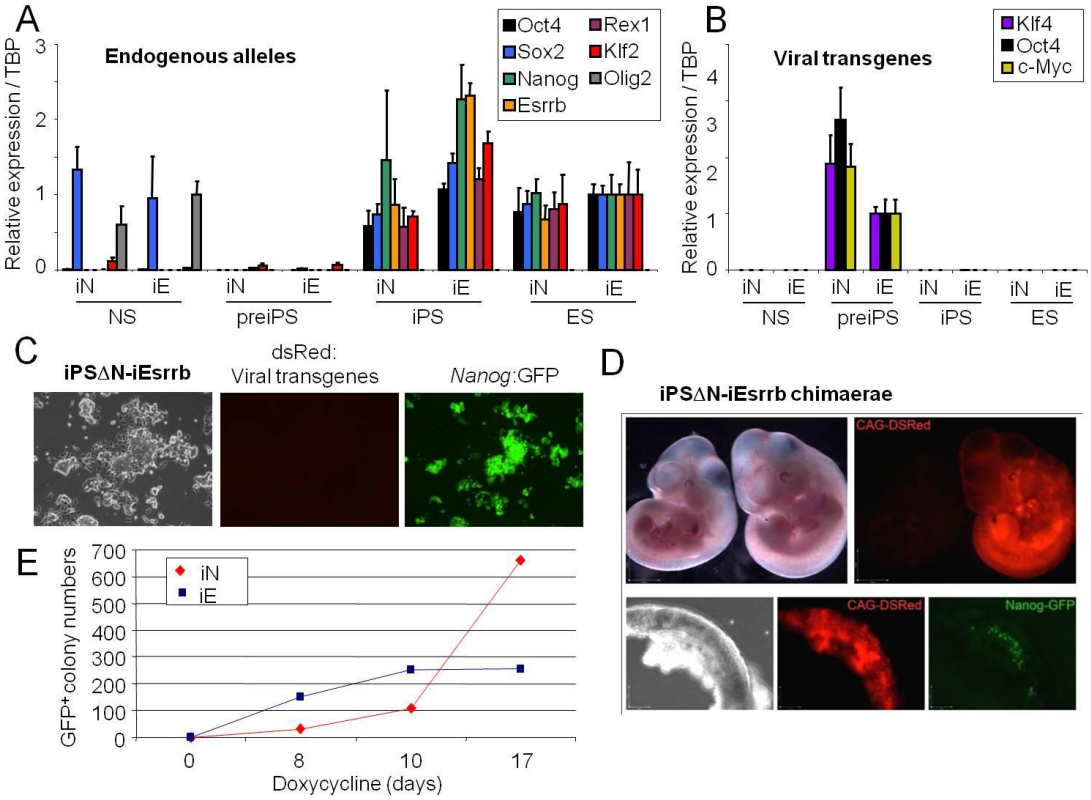


Figure 6.7: Esrrb can reprogramme *Nanog*$^{-/-}$ somatic cells to naïve pluripotency.

A: Endogenous pluripotency or NS cell specific gene expression in ESΔN-iNanog (iN) or ESΔN-iEsrrb (iE) cells and derivative NS, pre-iPS, and iPS cells cultured without doxycycline for at least 3 passages. The primers do not detect viral transgenes. Nanog primers bind to the *Nanog* 5'UTR which is present in all targeted alleles. Transcript levels are relative to expression in NSΔN-iEsrrb cells (for Olig2) or ESΔN-iEsrrb cells (for all other genes). Error bars: iPS cells: standard deviation in three independent clones. ES, pre-iPS and NS lines: standard deviation in three independent experiments. **B**: Expression of retroviral transgenes in ESΔN-iNanog (iN) or ESΔN-iEsrrb (iE) cells and derivative NS, pre-iPS, and iPS cells cultured without doxycycline for at least 3 passages. The primers do not detect endogenous transcripts. Transcript levels relative to pre-iPSΔN-iEsrrb cells. Error bars: standard deviation in three independent experiments. **C**: DsRed and *Nanog*:GFP expression in iPSΔN-iEsrrb cells cultured in the absence of STO feeders and without doxycycline for 3 passages. **D**: Representative midgestation (E11.5) embryo obtained from blastocyst injection of iPSΔN-iEsrrb cells transfected with a ubiquitously expressed TdTomato transgene (right); control embryo (left). Bottom panels show *Nanog*:GFP expression in iPSΔN-iEsrrb derived primordial germ cells observed in dissected gonads at E11.5. **E**: Number of *Nanog*:GFP colonies scored after clonal density plating (5000 cells/10 cm diameter dish) of pre-iPSΔN-iNanog or pre-iPSΔN-iEsrrb cells and culture in the presence of doxycycline for the indicated number of days. Pre-iPS cells were left to form macroscopic colonies for 6 days before addition of doxycycline and 5'-Aza. 5'-Aza was withdrawn from the culture medium 4 days after doxycycline induction.

indefinitely on feeders without reactivating *Nanog:*GFP (**Figure 6.6B**). Pluripotency genes other than Sox2 remained silenced in pre-iPS cells and robust expression of all viral transgenes could be detected (**Figure 6.7A-B**). Pre-iPSΔN-iNanog and pre-iPSΔN-iEsrrb cells were then treated with doxycycline to activate the Nanog or Esrrb transgenes present in these lines. Induction was performed in the presence or absence of the DNA methyltransferase inhibitor 5'-azacytidine, since this drug was shown to promote reprogramming (Huangfu et al., 2008) and facilitate the pre-iPS to iPS transition (Mikkelsen et al., 2008; Theunissen et al., 2011b). Nanog induction in pre-iPSΔN-iNanog cells led to the emergence of *Nanog*:GFP$^+$ cells by day 5 , irrespective of the presence of 5'-azacytidine (**Figure 6.6C**). Strikingly, Esrrb induction combined with 5'-azacytidine treatment resulted in a faster and more pronounced reactivation of the *Nanog:*GFP allele. In contrast, Esrrb induction in the absence of 5'-azacytidine failed to trigger completion of the reprogramming process even after prolonged doxycycline treatment, indicating that an additional component of Nanog activity might directly or indirectly involve erasure of DNA methylation marks. Despite this difference, G418-resistant *Nanog*:GFP$^+$ colonies could be picked and expanded for both pre-iPSΔN-iNanog and pre-iPSΔN-iEsrrb cells. The resulting iPSΔN-iNanog and iPSΔN-iEsrrb cells could be expanded without the need for feeder co-culture and were not dependent on continued doxycycline treatment. These cells morphologically resembled the parental ES cell lines, maintained *Nanog:*GFP$^+$ expression and completely silenced the dsRed reporter (**Figure 6.7C**). Gene expression analysis confirmed that both iPSΔN-iNanog and iPSΔN-iEsrrb expressed endogenous pluripotency genes and had silenced the viral transgenes (**Figure 6.7A-B**).

The developmental potential of Nanog null iPS cells could now be rigorously tested. Blastocyst injection of iPSΔN-iEsrrb cells resulted in extensive contribution to midgestation embryos (**Figure 6.7D; Table 6.4**). Furthermore, the presence of GFP expressed from *Nanog* in the genital ridges of E11.5 embryos indicated that injected iPS cells were able to colonise the germline (8/19 chimaeras showed germline contribution, **Table 6.4**, **Figure 6.7D**).

Cell line	No. of embryos	Midgestation Chimaeras (E11.5)	% of chimaeras	Somatic chimaerism (E11.5)	Germline contribution (E11.5)
iPSΔN-iEsrrb C1 (Injection 1)	18	11	61.1%	3 high; 2 medium; 3 low (3 high contribution abnormal chimaeras)	3/11
iPSΔN-iEsrrb C1 (Injection 2)	19	8	42.1%	2 high; 3 medium; 2 low (1 high contribution abnormal chimaera)	5/8

Table 6.4: Esrrb induces NS derived iPSΔN-iEssrb chimaera forming capacity.

Quantification of midgestation chimaeras (E11.5) obtained from two blastocyst injections of iPSΔN-iEsrrb c1 cells. Contribution was judged by the proportion of TdTomato positive cells present in the chimaeras (High>66.6%; 66.6%<Medium>33.3%; Low<33.3%). Germline contribution was judged by the presence of GFP expressed from the *Nanog* locus in iPSΔN-iEsrrb cells colonising the genital ridges of E11.5 embryos

Doxycycline treatment timecourse experiments indicated that Nanog and Esrrb trigger completion of stalled reprogramming with different kinetics. Nanog:GFP reactivation was detected as soon as 3 days after Esrrb expression in pre-iPSΔN-iEsrrb cells. Interestingly, not all the GFP positive cells observed at this stage showed complete silencing of dsRed expression (**Figure 6.6C**). In contrast, Nanog:GFP expression was evident in a minority of iPSΔN-iNanog cells only 6 days after doxycycline induction, but all GFP^+ cells detected had completely silenced transgene expression (**Figure 6.6C**). This could indicate that Nanog reactivation precedes completion of reprogramming in induced pre-iPSΔN-iEsrrb cells and that not all GFP expressing cells eventually give rise to iPS clones.

In order to rigorously compare the reprogramming efficiency after Nanog or Esrrb induction, pre-iPSΔN-iNanog and pre-iPSΔN-iEsrrb cells were plated at clonal density (5000 cells) in 10cm dishes and cultured for 6 days in the absence of doxycycline. When macroscopic colonies appeared, doxycycline was added to the culture medium and the number of GFP positive colonies monitored over the course of 3 weeks. 5'-azacytidine was added to the culture during the first 4 days of induction. In line with the observed reprogramming kinetics, GFP positive colonies appeared in pre-iPSΔN-iEsrrb plates soon after doxycycline addition (**Figure 6.7E**). The GFP^+ colony number reached plateau 10 days after induction, remaining stable afterwards. In contrast, Nanog induction resulted in a delayed appearance of GFP^+ colonies, but their number steadily increased over the 3 weeks of the experiment. Approximately 650 and 250 GFP^+ colonies were detected after 17 days of continued Nanog and Esrrb induction respectively.

These results demonstrate that Esrrb can drive completion of reprogramming in the absence of Nanog, indicating that Esrrb can substitute for *Nanog* in the acquisition of pluripotency. Preliminary results also suggest that Esrrb and Nanog trigger reprogramming with different kinetics, an observation that deserves further attention.

6.4: Discussion

Extending the characterisation of the functional similarities between Nanog and Esrrb presented before, this chapter shows that Esrrb is able to drive reprogramming of NS cells in cell fusion experiments to an extent comparable to Nanog (**Figure 6.1**). Fusion to ES cells is a fast and efficient system to drive reprogramming of differentiated cells (Bhutani et al., 2010; Han et al., 2008; Tada et al., 2001). NS x ES cell fusion drives reprogramming of the NS cell genome without the need for expression of additional reprogramming factor (Silva et al., 2006), making it an ideal tool to study the role of specific genes in increasing reprogramming efficiency in a system in which any additional contribution is easily detected. In addition, the role of Nanog in promoting NS cell reprogramming in this context has been extensively characterised: Nanog increases dramatically the reprogramming efficiency (Silva et al., 2006) and the presence of Nanog is a prerequisite for reprogramming to proceed (Silva et al., 2009). Fusion of Nanog null ES cells to Nanog null NS cells leads to the formation of unstable intermediates that cannot be expanded. Thus, this system is also useful to identify factors that are able to complement loss of Nanog function. As discussed before, Esrrb can rescue the impaired self-renewal of $Nanog^{-/-}$ cells and is able to substitute for Nanog in conferring LIF independence to ES cells. The data presented in this chapter shows that Esrrb can also rescue the reprogramming defects of $Nanog^{-/-}$ cells in NS x ES cell fusion experiments. Since Esrrb is a target of Nanog, it is possible that part of the activity of Nanog in driving reprogramming is mediated through upregulation of Esrrb. Nonetheless, Esrrb and Nanog reprogramming efficiencies are not comparable, indicating that additional Nanog targets might play a role in this context. Alternatively, Nanog protein might be important in recruiting other transcription factors or chromatin remodelers to the promoters or enhancers of crucial targets of the pluripotency GRN. To address these questions, it will be necessary to test the effects of overexpressing selected Nanog targets in NS x ES cell fusion experiments. In addition, it would be interesting to determine whether Nanog is able to bind to the regulatory elements of any of the pluripotency GRN targets during the early stages of reprogramming, possibly exploiting allelic differences in heterokaryons generated by fusion of cells from

different species. The data presented also shows that *Nanog*$^{-/-}$ reprogrammed hybrid lines generated by overexpression of Esrrb correctly reactivate expression of Nanog from the NS cell genome (**Figure 6.4B**), indicating that self renewal of these hybrids is not exclusively sustained by expression of pluripotency genes from the already active ES cell genome. This suggests that reversion of epigenetic silencing and reactivation of expression from the endogenous pluripotency alleles in differentiated cells is not strictly dependent on Nanog activity.

Additional cues to the role that Nanog exerts during reprogramming might come from experiments demonstrating that Esrrb can substitute for Nanog function in completing stalled reprogramming of Nanog$^{-/-}$ pre-IPS cells. Activation of Esrrb in pre-iPSΔN-iEsrrb leads to the reactivation of pluripotency genes (**Figure 6.7A**), triggers silencing of the viral transgenes (**Figure 6.7B**) and allows reprogramming of NS cells to chimaera competency (**Figure 6.7D**). Importantly, Esrrb requires the addition of 5'-azacytidine to complete this process whereas restoring Nanog function is independently able to generate fully reprogrammed iPS cells. 5'-azacytidine inhibits DNA methyltransferases, leading to gradual genome-wide reduction of 5mCpG levels after successive cell divisions (Creusot et al., 1982), reversing repressive hypermethylation at pluripotency loci and possibly making these genes responsive to Esrrb activation. Recent work analysing the transcriptional and epigenetic state of different pre-iPS lines, showed that they are trapped in an intermediate state between differentiated cells (MEFs) and ES cells. Almost all high-CpG promoters (HCP) are marked by H3K4me3 in ES cells. A subfraction of these is also marked by the repressive H3K27me3 modification. These domains are defined as bivalent (Mikkelsen et al., 2007) and mark lineage specific genes poised for activation during the differentiation process. In MEFs HCP resolve to become monovalently marked by H3K4me3 or H3K27me3. In addition, some of these promoters, including pluripotency genes, lose both marks and become transcriptionally silent and hypermethylated (Mikkelsen et al., 2007). During the reprogramming process bivalent marks are re-established and methylation of pluiripotency genes is reversed (Maherali et al., 2007; Mikkelsen et al., 2008). In pre-iPS bivalency is not completely restored, and pluripotency genes remain devoid

of both H3K4me3 and H3K27me3, showing elevated CpG methylation. Interestingly, 5'-azacytidine treatment of *Nanog*$^{+/+}$ pre-iPS reverses CpG methylation and leads to the spontaneous reactivation of pluripotency gene transcription (Mikkelsen et al., 2008). In contrast, in the absence of concomitant transgene induction, the appearance of *Nanog*:GFP^{+} cells was not detected even after long term culture of Nanog$^{-/-}$ pre-iPSΔN-iNanog or pre-iPSΔN-iEsrrb cells with 5'-azacytidine. This difference, and the fact that Nanog restoration in pre-iPSΔN-iNanog lines can complete reprogramming without 5'-azacytidine, indicates that Nanog is directly or indirectly involved in reversing silencing at highly methylated pluripotency gene promoters. Nanog could perform a pioneering role (Zaret and Carroll, 2011) in binding to these loci in pre-iPS cells. An alternative model is suggested by recent reports revealing that the most differentially expressed genes between pre-iPS and iPS cells, including pluripotency genes, lack clear binding by Oct4 Sox2 and Klf4. Nonetheless, in reprogramming intermediates Oct4 and Sox2 often show weak levels of binding to these loci (Sridharan et al., 2009). This might indicate that Oct4 and Sox2 are the first factors accessing the chromatin of repressed pluripotency genes, yet their binding might not be sufficient to activate transcription. Binding of additional factors, possibly Nanog, might be required to complete reactivation of these genes. Intriguingly, Nanog is not expressed in pre-iPS cells or during the initial phases of reprogramming ((Buganim et al., 2012; Mikkelsen et al., 2008; Silva et al., 2008a; Silva et al., 2009; Sridharan et al., 2009) and **Figure 6.7A**). It is possible that stochastic activation of Nanog expression and its consequent binding to the regulatory elements of pluripotency genes in pre-iPS lines or reprogramming intermediates leads to the erasure of DNA methylation and contributes to the completion of their transcriptional activation. Lack of Nanog could impair this process and result in the stalled reprogramming observed in Nanog-/- cells. It would be interesting to determine the genome-wide pattern of Nanog binding after Nanog re-induction in pre-iPSΔN-iNanog cells, and determine whether restored Nanog binding correlates with reversion of DNA hypermethylation at its target genes. In light of the data presented, it is possible that Esrrb is able to bind to a similar set of targets in pre-iPS cells but fails to trigger complete epigenetic resetting. To test this hypothesis, it will be necessary to compare the 5mCpG levels at the promoter of

selected pluripotency genes in doxycycline treated pre-iPSΔN-iNanog and pre-iPSΔN-iEsrrb cells. In this regard, it is interesting to note that whereas Esrrb requires 5'-azacytidine for triggering completion of pre-iPS reprogramming, its overexpression in NS x ES cell fusion experiments independently leads to reprogramming of *Nanog*$^{-/-}$ cells. It is reasonable to imagine that among the number of factors contributed by the undifferentiated ES cells in fusion hybrids must be elements driving the reversion of repressive epigenetic marks at the regulatory elements of pluripotency genes. Interestingly in this respect, recent reports indicate a role for activation-induced cytidine deaminase (AID) in mediating DNA demethylation at the *Oct4* and *Nanog* loci in ES x B cell heterokaryons (Bhutani et al., 2010).

The need for additional 5'Aza treatment during Esrrb driven reprogramming of *Nanog*-/- cells could also be a consequence of the fact that pre-iPS might represent abnormal intermediates of the reprogramming process. In pre-iPS cells endogenous pluripotency genes are not expressed, the transcription of genes characteristic of differentiated cells is repressed, viral transgenes are not yet silenced (**Figure 6.7A-B** and (Mikkelsen et al., 2007; Silva et al., 2008a; Silva et al., 2009; Sridharan et al., 2009)) and X chromosome inactivation persists in female cells (Silva et al., 2008a; Sridharan et al., 2009). Analysis of the transcriptional and epigenetic events accompanying the generation of iPS cells seems to indicate that reprogramming intermediates progress through a similar state. Early events during reprogramming are the silencing of genes expressed in differentiated cells, an increase in the proliferation rate and the completion of mesenchymal to epithelial transition (Li et al., 2010; Samavarchi-Tehrani et al., 2010). Acquisition of alkaline phosphatase and SSEA-1 expression precedes reactivation of pluripotency markers (Brambrink et al., 2008), which is accompanied by X chromosome reactivation and re-expression of telomerase (Stadtfeld et al., 2008). However, whereas most of the cells that activate SSEA-1 expression during reprogramming have already silenced expression of the viral transgenes, and SSEA-1 expression in these cells correlates well with the potential to form colonies of fully reprogrammed cells (Stadtfeld et al., 2008), in stable pre-iPS lines SSEA-1 expression is variable and does not correlate with viral

transgene silencing, DNA methylation status or potential to complete reprogramming (Mikkelsen et al., 2008). Similarly, whereas the sporadic activation of endogenous Oct4 transcription detected in pre-iPS lines is abortive (Silva et al., 2008a; Theunissen et al., 2011b), and does not imply epigenetic resetting nor completion of reprogramming (Theunissen et al., 2011b), Oct4 reactivation correlates well with the ability to achieve full reprogramming during the derivation of iPS cells (Buganim et al., 2012; Stadtfeld et al., 2008). It might therefore be that stable pre-iPS lines are derived from cells trapped in states that are particularly unfavourable for the progression of reprogramming. Even if pre-iPS cells are still able to be rescued by ERK and GSK inhibition (Silva et al., 2008a) or overexpression of single reprogramming factors (Theunissen et al., 2011b), they might not represent likely transition states for the rare cells that successfully originate iPS colonies. In light of these observations, it would be interesting to perform experiments inducing Esrrb elevation from the initial phases of the reprogramming of *Nanog*$^{-/-}$ cells. Esrrb elevation may be able to avoid trapping of these cells in a refractory state and lead to successful reprogramming without need for 5'-azacytidine addition.

Another indication that Esrrb and Nanog might not be exerting completely overlapping roles during reprogramming comes from the observation that these two factors trigger the pre-iPS to iPS transition with radically different kinetics. Whereas Esrrb leads to fast reactivation of *Nanog*:GFP expression in pre-iPS lines but its action seems to plateau quickly, Nanog requires longer time to act, but results in efficient reprogramming in an exponentially increasing number of cells over time (**Figure 6.7E**). The delayed activation of Nanog transcription observed in doxycycline induced pre-iPSΔN-iNanog cells could be explained by the notion that Nanog represses its own transcription. In these cells *Nanog*:GFP expression might not be directly driven by Nanog binding to its own regulatory elements, but could require activation of other pluripotency factors not expressed in pre-iPS cells. *Nanog* expression in this line would therefore be a stringent reporter of full reprogramming. Conversely, the ability of Esrrb to directly activate Nanog transcription (van den Berg et al., 2008) might result in early expression of *Nanog*:GFP in pre-iPSΔN-iEsrrb cells, but might not always coincide with completion of reprogramming.

Reprogramming driven by Esrrb proceeds in *Nanog*$^{-/-}$ NS cells with significantly reduced efficiency compared to WT. It was recently shown that Esrrb can reprogram EpiSC to naïve pluripotency with higher efficiency than Nanog (Festuccia et al., 2012). Intriguingly, Esrrb is able to compensate for the absence of Nanog also in this setting, but reprogramming is inefficient in *Nanog*$^{-/-}$ EpiSC. This indicates that Esrrb and Nanog act cooperatively to induce pluripotency, similarly to what observed in LIF independence assays (See Chapter 5). Such cooperativity may be important for the generation of iPS cells, since Nanog and Esrrb expression are upregulated at similar times during reprogramming (Samavarchi-Tehrani et al., 2010). The data presented in this thesis also show that the ability of Nanog overexpression to confer LIF independence is abrogated in the absence of Esrrb (See Chapter 8). It would be now interesting to determine whether Nanog ability to promote reprogramming is similarly dependent on Esrrb expression. In addition, given that Nanog is required for the progression of reprogramming, it would be important to test whether a similar requirement exists for Esrrb. Esrrb might be required for establishment of pluripotency, but dispensable once pluripotency is attained (see Chapter 8), further underscoring its functional similarity with Nanog (Silva et al., 2009).

Finally, the current model postulates that Nanog exerts a function in the last stages of the reprogramming process (Silva et al., 2009; Theunissen and Silva, 2011), implying that similarities might exist between Nanog function in the establishment of pluripotency in the ICM and its role in completing reprogramming. Although this is an intriguing possibility, there is no definitive evidence that reprogramming in vitro proceeds through the same routes as acquisition of pluripotency in the early embryo.

Chapter 7:Esrrb complements Nanog function in germline development

ES cells lacking Nanog fail to contribute to the germline; in chimaeric embryos, primordial germ cells (PGC) derived from Nanog null ES cells are lost between 11.5 and 12.5 days of development (Chambers et al., 2007). Interestingly, Nanog is not required for the initial specification of the founder PGC population that occurs in the proximal epiblast between E6.25 and E7.25 (Ohinata et al., 2009; Ohinata et al., 2005; Sato et al., 2002; Yamaji et al., 2008). Epiblast cells lacking Nanog are able to maintain the expression of both transcriptional regulators associated with pluripotency, like Oct4, and acquire expression of germ cell markers, like Mvh (Chambers et al., 2007). Instead, Nanog activity seems to be crucial in the maintenance of the already specified PGC population (Chambers et al., 2007) and in the promotion of its survival (Yamaguchi et al., 2009). Results in previous chapters have shown that Esrrb can substitute for Nanog in sustaining the acquisition and maintenance of pluripotency. Importantly, developing PGCs undergo complex transformations that specifically involve epigenetic resetting and reactivation of the pluripotency network, with parallels to those observed during the specification of the pluripotent ICM in pre-implantation embryos (See chapter 1.1.2), a process for which Nanog is required (Mitsui et al., 2003; Silva et al., 2009). In addition, Esrrb has also been shown to be important for PGC development. Reduced numbers of PGCs are detected in the genital ridges of Esrrb null embryos after E11.5 (Mitsunaga et al., 2004). Thus, it was decided to determine whether Esrrb expression in Nanog null ES cells could rescue their defective contribution to the developing germline.

7.1: Placement of Esrrb at the *Nanog* locus in *Nanog*$^{-/-}$ ES cells.

A construct encoding for $(Flag)_3$Esrrb was inserted at the *Nanog* locus in RCNβH(t) ES cells (**Figure 7.1A**). In the resulting cell line (RCNβH(t) Esrrb Knock-In (Esrrb KI)), the $(Flag)_3$Esrrb transgene is expressed under the control of the endogenous *Nanog* regulatory elements; since this is an insertion and not a replacement, no additional loss of genetic information from the already disrupted *Nanog* locus occurs. In chimaeric embryos generated by blastocyst injection of Esrrb KI cells, the Esrrb transgene should be expressed in PGC precursors derived from injected cells when *Nanog* transcription is reactivated around E7.75 and expression should persist beyond migration to the genital ridges until E14.5 in female and E15.5 in male embryos respectively (Yamaguchi et al., 2005). Thus, in this cell line, Esrrb will be continuously expressed from the *Nanog* locus during the whole period before and beyond the time when the Nanog null phenotype is manifest at E11.5-12.5.

Correctly targeted Esrrb KI clones were identified by Southern blot analysis (**Figure 7.1A-B**). Two clones that had undergone correct recombination at either the βGeo (clone 4) or the Hygro^R (clone 5) targeted alleles were selected for further analysis. In Esrrb KI ES cells, Esrrb expression is higher in comparison to the parental RCNβH(t) cell line and wild-type E14Tg2a ES cells (**Figure 7.2A**). This is explained by the fact that expression of Esrrb driven from *Nanog* in this cell line adds to the endogenous Esrrb levels. The $(Flag)_3$Esrrb protein encoded by the KI transgene can be distinguished from the endogenous Esrrb protein by virtue of its increased molecular weight. As expected, expression of both wild-type and $(Flag)_3$Esrrb protein was detected in Esrrb KI cells, with no evidence of negative feedback by Esrrb protein on *Esrrb* (**Figure 7.2B**).

Five sets of blastocyst injections were performed for Esrrb KI clone 4 cells and one

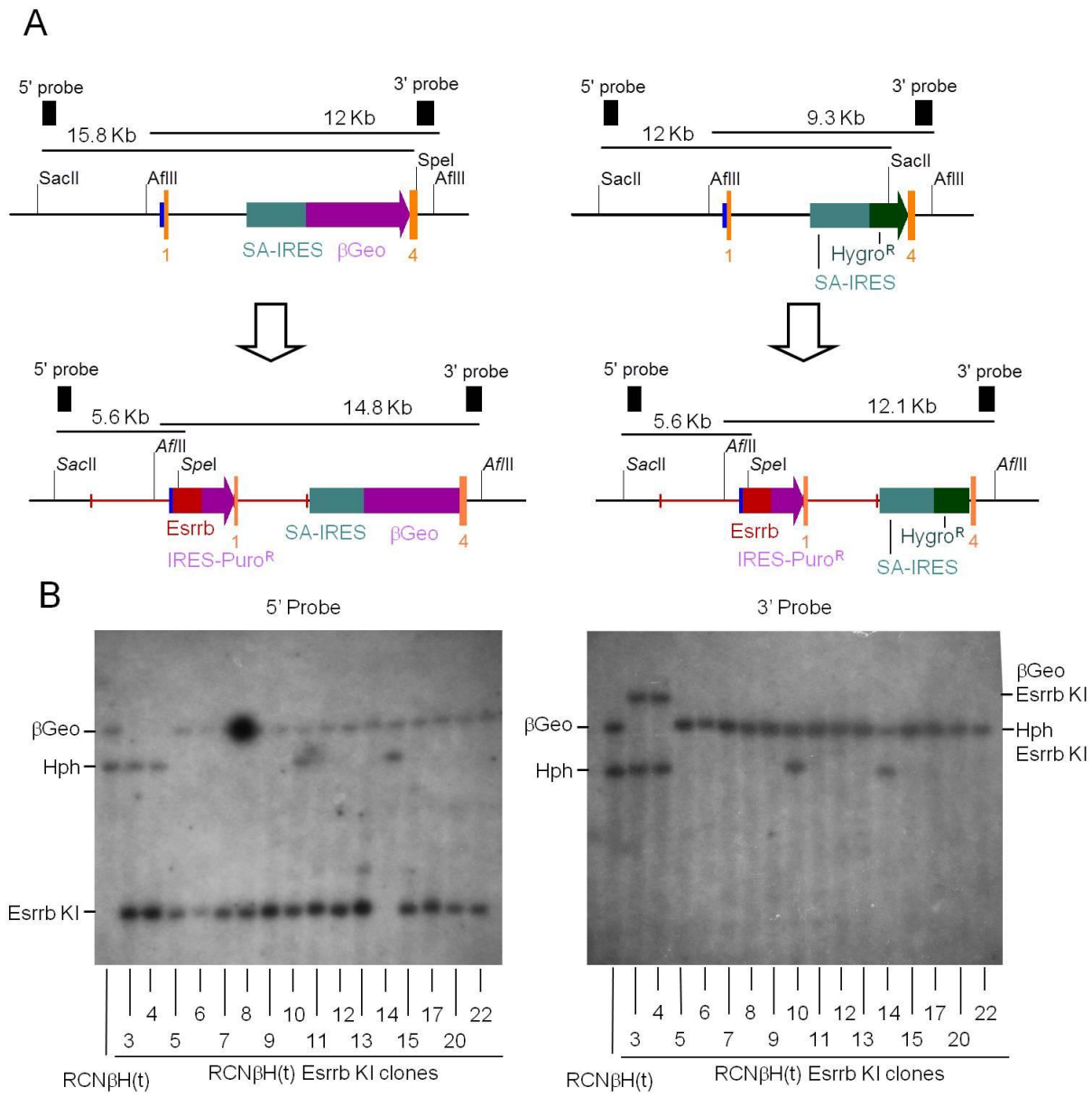


Figure 7.1: Derivation of RCNβH(t) Esrrb KI ES cells.

A: Schematic representation of the *Nanog* locus structure in RCNβH(t) Esrrb KI ES cells, showing the hygromycin and β-geo targeted alleles in the parental RCNβH(t) lines and the respective Esrrb knock-in alleles, along with the restriction sites and DNA probes used for Southern blot analysis. The expected sizes of the DNA fragments obtained after digestion are shown on top of each diagram. The homology arms of the targeting vector are shown in red. *Nanog* exons are shown in orange. The *Nanog* 5'UTR is in blue. **B**: Southern blot analysis performed on DNA samples prepared from RCNβH(t) and RCNβH(t) Esrrb KI ES cell lines.

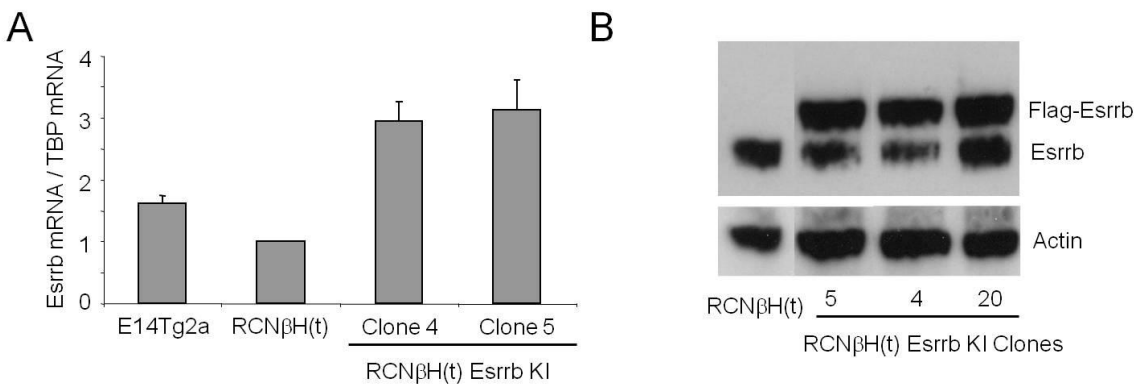


Figure 7.2: Esrrb expression in RCNβH(t) Esrrb KI cells.

A: Esrrb transcript levels in E14Tg2a, RCNβH(t) and RCNβH(t) Esrrb KI ES cells. Error bars: standard deviation of the technical errors in gene expression quantification in one experiment. **D**: Esrrb protein levels in RCNβH(t) and RCNβH(t) Esrrb KI ES cells. Actin B protein levels are shown as loading control.

for RCNβH(t) parental *Nanog*$^{-/-}$ cells. In RCNβH(t) ES cells, GFP is expressed from a constitutively active CAG driven transgene, allowing for accurate tracing of the injected cells in chimaeric embryos (Chambers et al., 2007). Thus, dissected E13.5 chimaeric embryos could be easily distinguished from their wild-type littermates on the basis of GFP expression (**Figure 7.3A**). Pictures of every embryo were taken before dissection to ensure that a correct staging could be subsequently established. After dissection, pictures were also taken for all genital ridges isolated (**Figure 7.3A**). The levels of somatic contribution were determined by flow cytometry on cell suspensions obtained by mechanical dissociation and trypsinisation of the embryo bodies, after removal of the genital ridges (**Figure 7.3B, Table 7.1**). Next, the genital ridges were stained for Oct4 and Mvh and contribution to the germline of GFP positive cells was assessed by confocal fluorescence microscopy. No GFP^{+} cell expressing Mvh or Oct4 could be detected after control injections of the RCNβH(t) parental line (**Figure 7.4B**). In contrast and unexpectedly, Esrrb KI cells positive for Mvh and Oct4 expression, and displaying a round morphology typical of PGCs, could be readily identified in the genital ridges of E13.5 chimaeric embryos (solid arrows, **Figure 7.4A-B**), along with PGCs derived from the host embryo (open arrows).

A summary of the results from all injection performed is shown in **Tables 7.1-7.2**. It is important to note that parental RCNβH(t) cells showed consistently higher somatic contribution than Esrrb KI cells. This observation may be due to intrinsic differences between the clonal lines employed (Esrrb KI clone 4 and RCNβH(t) c1one 12 (Chambers et al., 2007)) or by the hypothesis that elevated levels of Esrrb protein may partially block differentiation of Esrrb KI cells during early embryonic development. 10 out of the 26 embryos analysed after injection of Esrrb KI cells showed contribution to the germline at E13.5 (38%). Despite this elevated frequency, the levels of germline contribution were consistently low in all sets of injections performed (<5%-20% of Esrrb KI derived PGCs).

These results indicate that, at least up until E13.5, Esrrb can rescue the defective development of *Nanog*$^{-/-}$ PGCs.

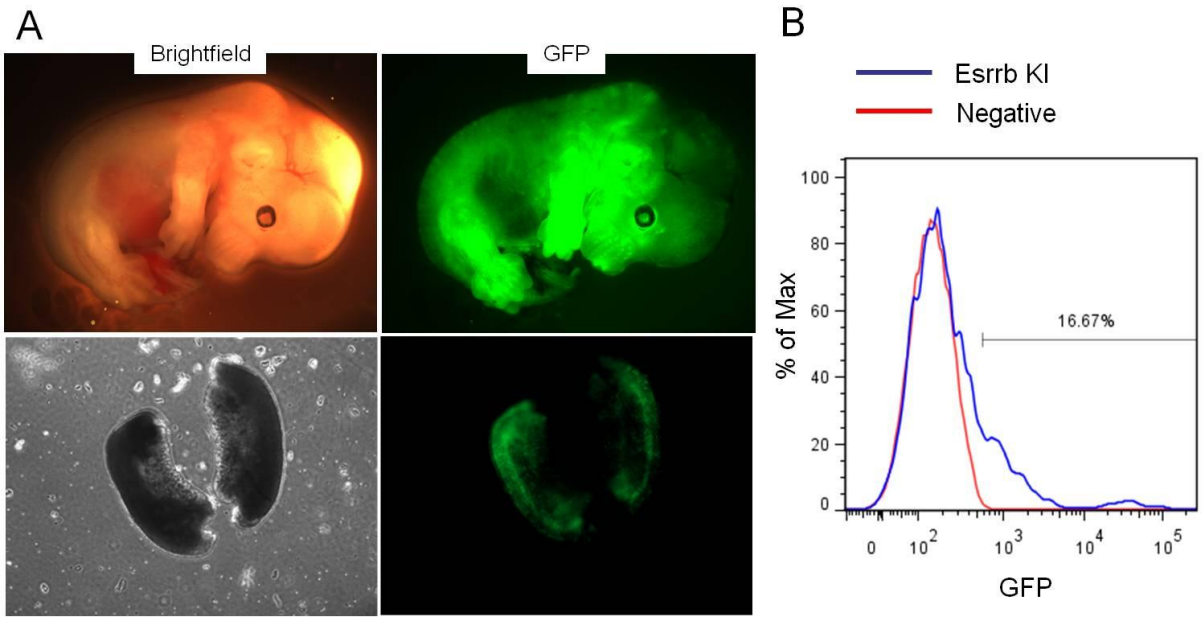


Figure 7.3: Blastocyst injection of RCNbH(t) Esrrb KI ES cells.

A: GFP expression in E13.5 chimaeric embryos derived from blastocyst injections of RCNβH(t) Esrrb KI ES cells and in the dissected gonads. GFP marks Esrrb KI ES cells **B**: Histogram plot showing the distribution of GFP expression in cell suspensions obtained by mechanical dissociation and trypsinisation of chimaeric embryos after removal of the gonads. Cells from non-chimaeric embryos served as controls and are shown in red. Somatic contribution was determined by analysing the percentage of GFP^+ cells.

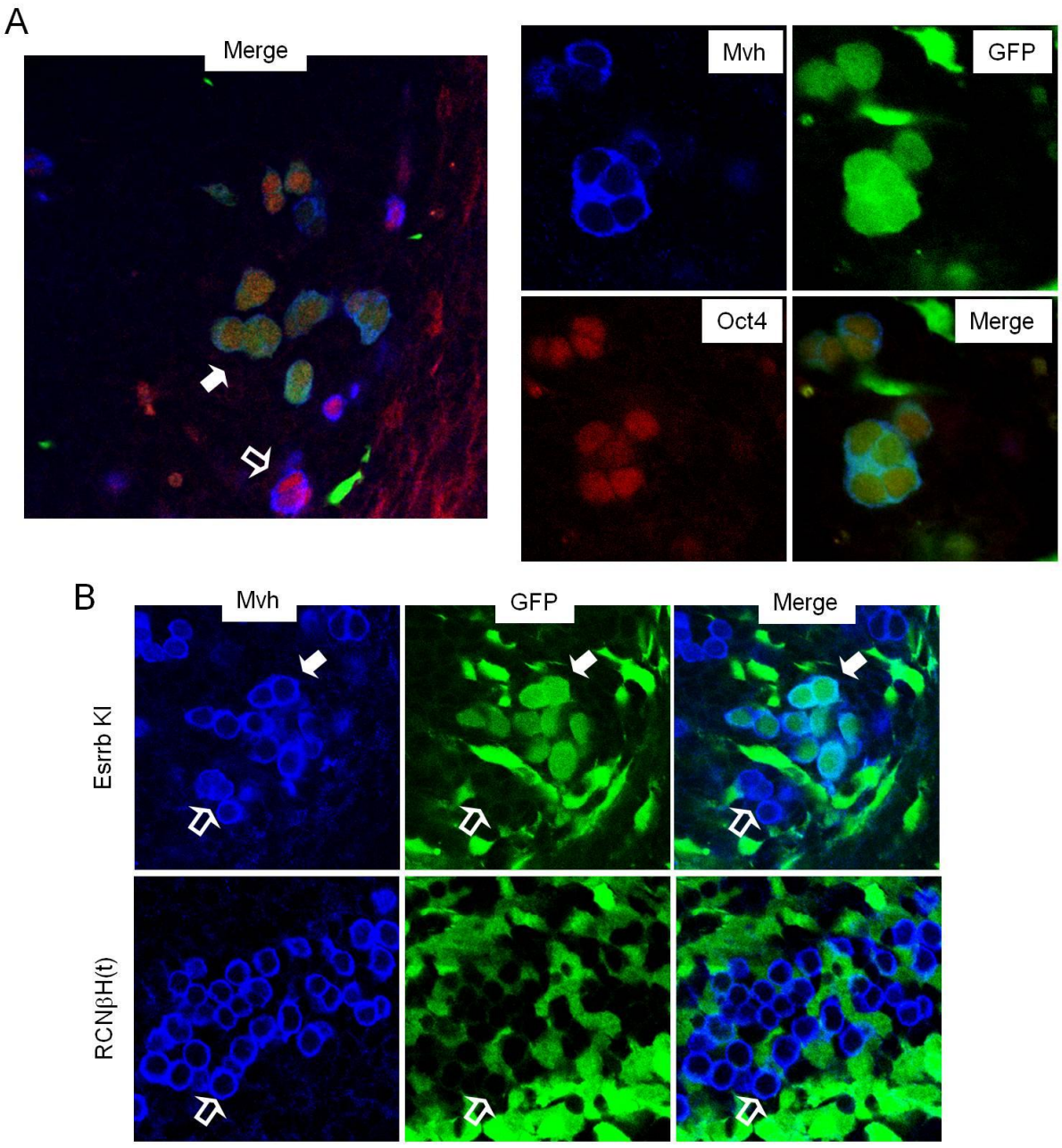


Figure 7.4: Esrrb knock-in extends the development of Nanog null PGCs beyond E11.5.

A: Oct4 and Mvh expression in the genital ridges of E13.5 embryos derived from blastocyst injections of RCNβH(t) Esrrb KI cells. **B**: Mvh expression in the genital ridges of E13.5 embryos derived from blastocyst injections of RCNβH(t) Esrrb KI cells (top) or RCNβH(t) cells. Esrrb KI and RCNβH(t) cells are GFP^+. Solid arrows: GFP^+ PGCs derived from RCNβH(t) Esrrb KI ES cells. GFP^- PGCs derived from the host embryo. Right panels in A show a higher magnification image of PGCs in the same gonad presented in the left panel

RCNβH(t) Esrrb KI injections

Stained for: Mvh Oct4; Stage E13.5

Embryo	Somatic contribution	Germline contribution
1	12.13	No
2	5.12	5-10%
3	10.33	No
4	5.24	No
5	1.49	No
6	2.71	No
7	0.53	No
8	0.67	No
9	4.3	10-20%

Stained for: Mvh; Stage E13.5

Embryo	Somatic contribution	Germline contribution
1	5.4	<5%
2	1.44	No
3	16.94	10-20%
4	17.16	No
5	6.4	No
6	2	No
7	2.38	No
8	0.74	No

Stained for: Mvh; Stage E13.5

Embryo	Somatic contribution	Germline contribution
1	12.65	No
4	11.21	5-10%

Stained for: Mvh and Oct4; Stage E12.5

Embryo	Somatic contribution	Germline contribution
12	3.22	5-10%

Stained for: Mvh; Stage E13.5

Embryo	Somatic contribution	Germline contribution
1	9.12	10-20%
5	11.91	No
8	1.39	10-20%
10	47.60	<5%
12	29.19	<5%
13	1.89	No

Table 7.1: Germline contribution potential of *Nanog*$^{-/-}$ RCNβH(t) ES cells.

Somatic and germline contribution observed in E13.5 (or E12.5) embryos after 5 independent sets of blastocyst injections using RCNβH(t) Esrrb KI ES cells. Contribution to the germline was scored comparing the numbers of GFP^{+}/Mvh^{+}/(and Oct4^{+}, as specified above each table) PGCs and GFP^{-} PGCs in multiple confocal image stacks from dissected gonads.

RCNβH(t) ES cells control injections

Staining for: Mvh and Oct4; Stage: E13.5

Embryo	Somatic contribution	Germline contribution
1	34.92	No
2	27.57	No
3	35.64	No
4	36.27	No
5	40.39	No
6	27.66	No
7	26.58	No
8	18.82	No

Table 7.2: Germline contribution potential of Nanog$^{-/-}$ RCNβH(t) ES cells.

Somatic and germline contribution observed in E13.5 embryos after blastocyst injection of RCNβH(t) ES cells. Contribution to the germline was scored comparing the numbers of GFP^{+}/Mvh^{+}/Oct4^{+} PGCs and GFP^{-} PGCs in multiple confocal image stacks from dissected gonads.

7.2: Generation of Esrrb KI mice

From the previous results, it is possible to hypothesise that precocious and elevated Esrrb expression is able to compensate for the lack of Nanog during PGC development. Nonetheless, the low levels of germline contribution observed for rescued *Nanog*$^{-/-}$ ES cells might indicate that the observed PGCs are not correctly developing, and that Esrrb KI may only result in a delayed loss of Nanog$^{-/-}$ PGCs in chimaeric embryos.

It was therefore decided to determine whether *Nanog*$^{-/-}$ Esrrb KI PGCs could develop normally to later embryonic stages, and give rise to functional gametes in adult animals. With this aim, Esrrb knock-in at the *Nanog* locus was repeated in E14Tg2a ES cells (**Figure 7.5A-B**) and clones which underwent correct recombination were identified by Southern blot. Two clones, clone 16 and clone 18, were selected for blastocyst injection and chimaeric mice derived from this line. Germline transmission was achieved in all chimeras generated by both clones and the Esrrb KI mouse line established by crossing these animals to females of the 129/Ola strain.

Nanog $^{+/-}$ (Mitsui et al., 2003) or *Nanog*$^{+/\text{Esrrb KI}}$ animals were then mated to mice in which a Cre transgene is put under the control of the Prdm1 regulatory elements, ensuring that recombinase expression is confined to the PGC compartment (Ohinata et al., 2005). Male or female *Nanog*$^{+/-}$ / Prdm1:Cre or *Nanog*$^{+/\text{Esrrb KI}}$ / Prdm1:Cre animals have now been crossed to mice carrying two floxed *Nanog* alleles (*Nanog*$^{\text{fl-STOP-GFP}}$). In these animals, *Nanog* deletion by Cre recombinase is accompanied by activation of GFP expression from the same locus (Chambers et al., 2007) (**Figure 7.6A-B**). Some progeny of *Nanog*$^{+/-}$ / Prdm1:Cre animals will inherit one *Nanog*$^{\text{fl-STOP-GFP}}$ allele, one *Nanog*$^{-}$ allele and the Prdm1:Cre transgene. Likewise, some progeny of Nanog$^{+/\text{Esrrb KI}}$ / Prdm1:Cre animals will inherit one *Nanog*$^{\text{fl-STOP-GFP}}$ allele, one *Nanog*$^{\text{Esrrb KI}}$ allele and the Prdm1:Cre transgene. PGCs in these embryos should recombine the *Nanog*$^{\text{fl-STOP-GFP}}$ allele soon after Blimp1 transcription is first activated around day E6.25 (Ohinata et al., 2005), making it

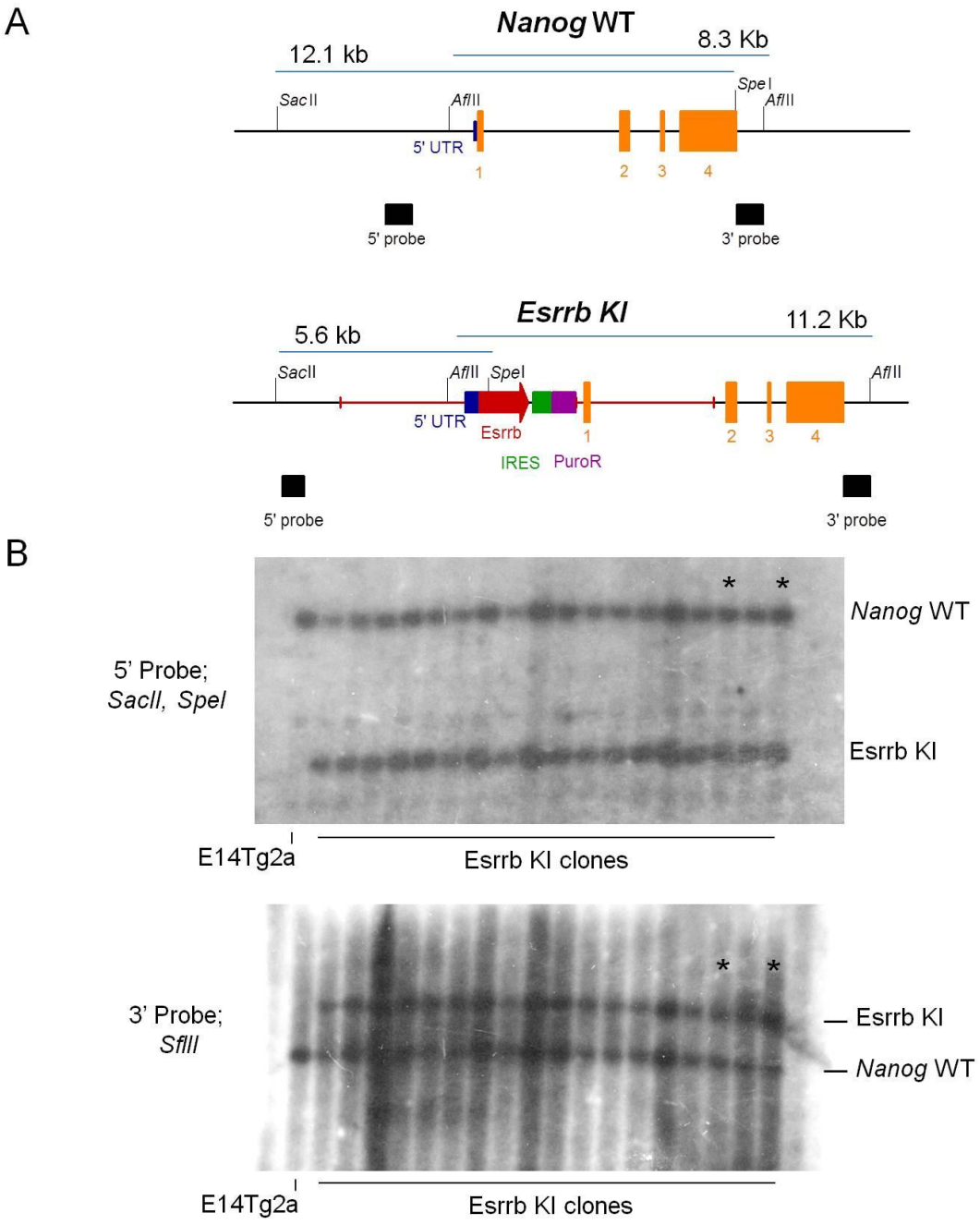


Figure 7.5: Derivation of E14Tg2a Esrrb KI ES cells.

A: Schematic representation of the *Nanog* locus structure in E14Tg2a and E14Tg2a Esrrb KI ES cells, showing wild-type and targeted *Esrrb* alleles, along with the restriction sites and DNA probes used for Southern blot analysis. The expected sizes of the DNA fragments obtained after digestion are shown on top of each diagram. The homology arms of the targeting vector are shown in red. **B**: Southern blot analysis performed on DNA samples prepared from E14Tg2a and E14Tg2a Esrrb KI ES cell lines.*: Clones 16 and 18 were selected for blastocyst injection.

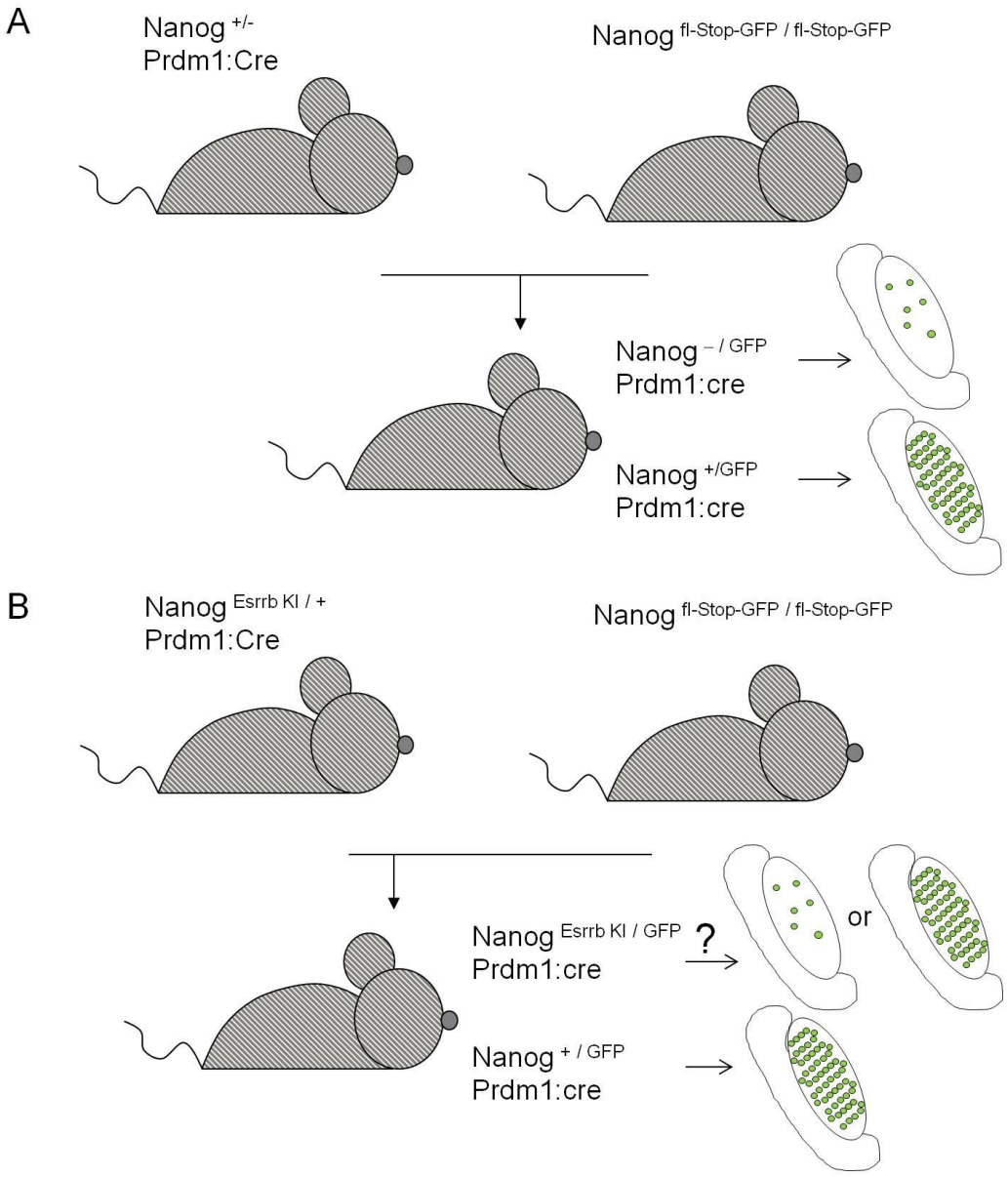


Figure 7.6: Experimental scheme for Esrrb KI mice crosses.

A: Scheme showing the planned crosses between *Nanog* $^{+/-}$ Prdm1:Cre and *Nanog* $^{\text{fl-STOP-GFP / fl-STOP-GFP}}$ animals, along with two of the possible genotypes expected in the progeny. **B**: Scheme showing the planned crosses between *Nanog* $^{\text{EsrrbKI / +}}$ Blimp1:Cre and *Nanog* $^{\text{fl-STOP-GFP / fl-STOP-GFP}}$ animals, along with two of the possible genotypes expected in the progeny. The *Nanog* $^{\text{fl-STOP-GFP}}$ allele is a conditional knockout *Nanog* allele in which recombination leads to excision of *Nanog* and expression of GFP under the control of the *Nanog* regulatory elements (shown as *Nanog* $^{\text{GFP}}$). Prdm1:Cre is expressed from a random integrated BAC transgene present in one copy per genome.

possible to track PGCs in which recombination has occurred. Littermates inheriting a *Nanog*$^{+}$, a *Nanog*$^{\text{fl-STOP-GFP}}$ allele and the Prdm1:Cre transgene will serve as controls (**Figure 7.6A-B**). It will be possible to follow the development of Nanog null PGCs avoiding the limitation imposed by chimaera experiments. In addition, the presence of a *Nanog* fluorescent reporter will potentially permit sorting and characterisation of the developing PGC population.

Adult animals derived from the discussed crosses will be genotyped and *Nanog*$^{\text{Esrrb KI / GFP}}$ mice mated to wild-type animals. Potential transmission of a recombined floxed *Nanog* allele to the progeny of Nanog $^{\text{Esrrb KI/GFP}}$ animals would strongly suggest that fully functional PGC developed in these animals in the absence of Nanog.

7.3: Discussion

During their migration to the genital ridges, PGCs undergo a series of transcriptional and epigenetic changes that culminates in major wave of epigenetic resetting after their entry into the gonads (Hajkova et al., 2008; Hajkova et al., 2002; Hajkova et al., 2010; Seki et al., 2005; Seki et al., 2007). Migratory PGCs, that have not yet completed erasure of methylation marks at imprinted loci (Hajkova et al., 2002), still display detectable levels of genomewide DNA methylation (Seki et al., 2005), but have already started reactivation of the inactive X chromosome (de Napoles et al., 2007; Sugimoto and Abe, 2007) and established high levels of H3K27me3 (Seki et al., 2007), are in a state that present analogies to that observed in the pluripotent cells of the pre-implantation epiblast. Furthermore, throughout migration, PGCs show expression the key pluripotency regulators Oct4, Sox2 and Nanog (Avilion et al., 2003; Chambers et al., 2003; Chambers et al., 2007; Kurimoto et al., 2008; Yamaguchi et al., 2005; Yeom et al., 1996; Yoshimizu et al., 1999), which are also required for correct germline development. (Chambers et al., 2007; Kehler et al., 2004; Yamaguchi et al., 2009) (see chapter 1.1.2).

The function of this transitory acquisition of traits characteristic of pluripotent cells still remains unclear, and little is known about the role that single pluripotency factors play in sustaining the migratory PGC population. Possibly, reactivation of the pluripotency network is required in PGCs to set the stage for further epigenetic changes occurring after day E10.5 (Hajkova et al., 2008; Hajkova et al., 2010). In alternative, it might simply be important for PGCs to escape somatic differentiation during their migration to the genital ridges. In both instances, Nanog absence could result in defective maintenance of pluripotency. The results presented in chapters 5 and 6 have shown that Esrrb is able to substitute for Nanog in driving the attainment and sustaining the maintenance of pluripotency. Esrrb expression is normally lost in the epiblast after implantation and is reactivated when the founder PGC population reaches the developing gonads at day E11.5 (Mitsunaga et al., 2004). Conversely, Nanog expression is reactivated around day E7.75 (Yamaguchi et al., 2005), soon after PGCs are specified. As a consequence, in PGCs derived from Esrrb KI cells, Esrrb should be expressed for a prolonged period of time, throughout migration to the genital ridges. It is possible that Esrrb expression during this time could help maintaining pluripotency genes expression in migratory PGCs.

Nanog could also be instrumental in driving reactivation of the pluripotency GRN in the founder PGC population. Although Nanog null PGCs express Oct4 at day E 11.5 (Chambers et al., 2007), this factor is never downregulated in PGC precursors. In contrast, Sox2 is transiently lost in PGCs accumulating in the posterior extraembryonic mesoderm region (Kurimoto et al., 2008; Yabuta et al., 2006). It would be interesting to determine whether Nanog null PGCs show perturbed reactivation of Sox2 and fail to upregulate other pluripotency or germline markers, such as Klf2 and Prdm14, that are not expressed in the post-implantation epiblast and have been shown to be instrumental in promoting the reversion of EpiSC to naïve pluripotency (Gillich et al., 2012). In particular it should be ascertained whether Nanog null PGCs are able to sustain expression of Prdm14 after its activation at E6.5 (Yamaji et al., 2008), since this gene is crucial in driving germ cell specification (Yamaji et al., 2008) and, although expressed at low levels in ES cells, shows good correlation with Nanog in deep sequencing experiments and prompt upregulation in

response to Nanog induction in ESΔN-NERT and ESΔN-iNanog cells (**Figure 3.1, Table 3.1**).

As mentioned, during their migration to the genital ridges, PGCs gradually reverse random X inactivation (Sugimoto and Abe, 2007). This process initiates in migratory PGCs soon after reactivation of Nanog expression (Yamaguchi et al., 2005). This is in agreement with the notion that Nanog plays a direct role in suppressing X chromosome inactivation by repressing Xist transcription in conjunction with Oct4 and Sox2 (Navarro et al., 2008). It would be interesting to check the extent of X chromosome inactivation in Nanog null PCCs before their disappearance at E12.5. Since Esrrb binds to *Xist* (Pablo Navarro, unpublished observation), the X inactivation status should be also determined in Esrrb KI PGCs at similar stages of development.

Intriguingly, onset of Esrrb expression at E11.5 (Mitsunaga et al., 2004) coincides with the time at which the phenotype of Nanog null cells manifests in the germline (Chambers et al., 2007), opening the additional possibility that the observed defects of *Nanog*$^{-/-}$ PGCs might arise as a consequence of defective Esrrb reactivation. Ablation of Esrrb results in reduced number of PGCs after E11.5 (Mitsunaga et al., 2004) but does not determine their complete loss, arguing against the hypothesis that failure to reactivate of Esrrb can explain the total absence of *Nanog*$^{-/-}$ PGCs at E12.5 (Chambers et al., 2007). Nonetheless, it is worth noting that lack of Esrrb expression in Nanog null PGCs would result in the concomitant absence of both Nanog and Esrrb in the developing germline. Since both ablation of Esrrb and Nanog results in a severely compromised ability to self-renew in ES cells ((Chambers et al., 2007) and see Chapter 8), it is possible that a severe phenotype could arise from the concomitant absence of these two factors in germ cells. In this case, Esrrb KI cells could partially or fully overcome this phenotype by ensuring Esrrb expression in *Nanog*$^{-/-}$ PGCs. To test this possibility, it will be first necessary to determine the dynamics of Esrrb expression in developing Nanog null PGCs soon after their entry into the genital ridges.

Finally, in vitro models could help explore the mechanisms underlying the ability of Esrrb to rescue the phenotype of *Nanog*$^{-/-}$ PGCs during development. Recent work demonstrated that ES cells derived epiblast-like cells can be converted into cells resembling developing PGCs (PGCLCs) by in vitro culture in the presence of a combination of instructive signals including BMP4/8, LIF, SCF and EGF (Hayashi et al., 2011). Since EpiSC can also be reverted to an ES state in vitro and Nanog is required for this transition (Silva et al., 2009), it would be interesting to determine whether Nanog is dispensable for the generation of PGCLC cells. The generation of relatively high numbers of cells acquiring PGC character through transitions that mirror in vivo development could help dissecting the contribution of Esrrb in alleviating the consequences of loss of Nanog. In addition, in such in vitro settings it would be possible to induce Esrrb elevation to levels that are not achieved in Esrrb KI cells (**Figure 7.2a**). If Esrrb expression levels prove to be correlated with the extent of the rescue observed, it could be useful to derive cell lines that will allow PGC restricted overexpression of Esrrb under the control of a strong inducible promoter. A possible strategy could involve targeting of the rtTA transactivator to the *Nanog* locus coupled to stable integration of a doxycycline inducible Esrrb transgene.

Chapter 8:Effects of Esrrb ablation and identification of Esrrb target genes

The results presented so far define Esrrb as an important factor controlling ES cell self-renewal and driving acquisition of pluripotency. Next, the effects of Esrrb ablation in ES cells were investigated. Since Esrrb is a prominent transcriptional target of Nanog, to conclusively prove that part of Nanog function in ES cells is mediated through activation of Esrrb, the effects of loss of Esrrb were studied in Nanog overexpressing cells.

8.1: Self-renewal in Esrrb knockout cells

Esrrb knockdown has been shown to affect the self-renewal of ES cells (Ivanova et al., 2006). However, these studies did not explore the consequences of complete loss of Esrrb, since they were limited by the potential for residual heterogeneous expression in the ES cell population. It was thus decided to incontrovertibly determine whether Esrrb expression is required for ES cell self-renewal. With this aim, cells homozygous for a conditional Esrrb knockout allele (*Esrrb*$^{f/fn}$) were generated by two successive rounds of homologous recombination using a previously described targeting construct in which loxP sites are placed at either side of *Esrrb* exon 2 (**Figure 8.1A-B**)(Chen and Nathans, 2007). ES cell that showed correct recombination were identified by Southern blot (**Figure 8.1B-C**) and engineered to stably express a tamoxifen inducible Cre-ER^{T2} transgene (**Figure 8.1A**).

The effects of Esrrb deletion were then assessed in Esrrb$^{f/fn}$ ES cells. Tamoxifen treatment of *Esrrb*$^{f/fn}$ cells plated at clonal density severely reduces the number of AP^{+} undifferentiated colonies observed after 7 days of culture in GMEMβ/FCS/LIF (**Figure 8.2A-B**).

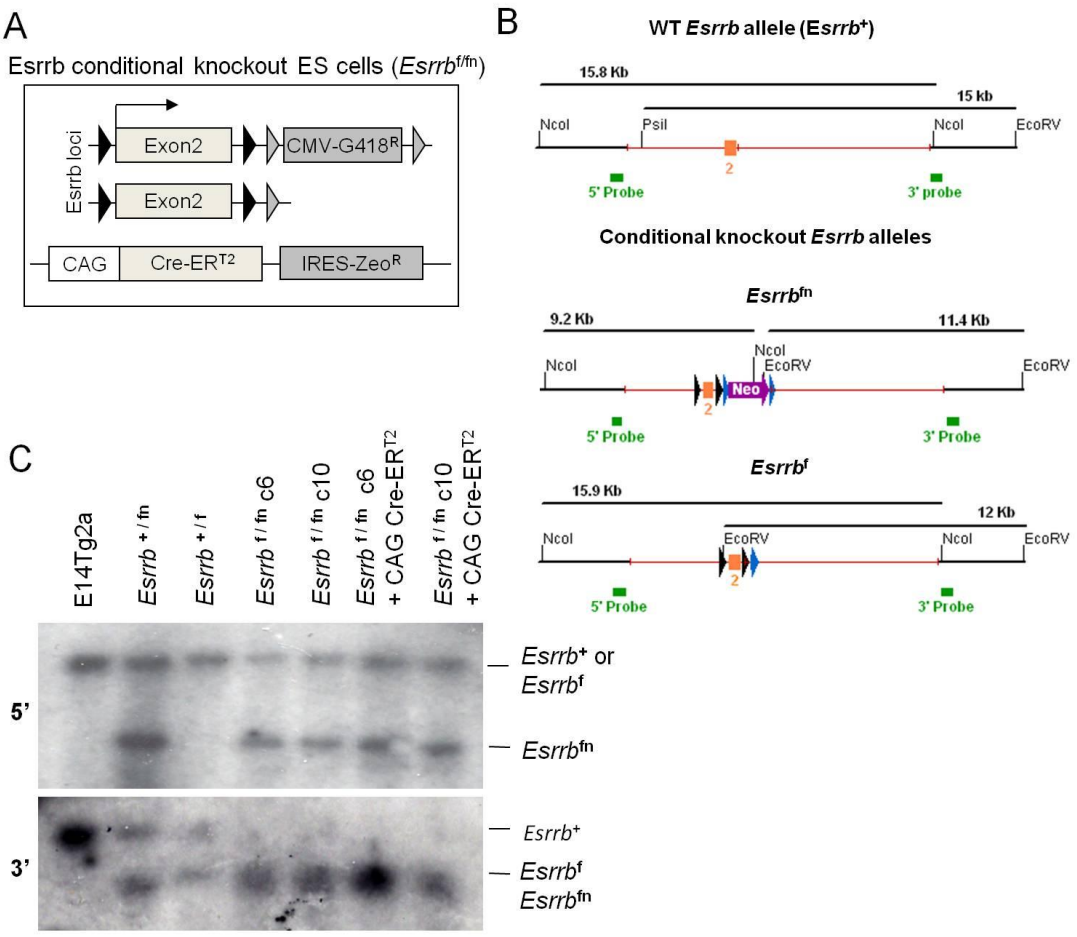


Figure 8.1: Derivation of Esrrb conditional knock-out ES cells.

A: Schematic representation of the genetic manipulations performed to derive conditional knockout (*Esrrb*$^{f/fn}$) ES cells. *Esrrb*$^{f/fn}$ cells have two floxed *Esrrb* alleles and constitutively express a tamoxifen inducible Cre-ER^{T2} transgene. **B:** Schematic representation of wild type and conditional knockout *Esrrb* alleles. The homology arms of the targeting vector are shown in red. The restriction enzyme sites used for Southern blot analysis are shown above each diagram, along with the expected size of the genomic DNA fragment generated by digestion with *NcoI* or *EcoRV* and *PsiI*. LoxP sites are shown as black triangles and FRT sites as blue triangles. **C**: Southern blot analysis of *Esrrb*$^{+/fn}$, *Esrrb*$^{+/f}$ and *Esrrb*$^{f/fn}$ lines generated after the first and second targeting events from E14Tg2a ES cells. Two separate double targeted *Esrrb*$^{f/fn}$ lines are shown before and after stable integration of a CRE-ER^{T2} transgene. Retention of *Esrrb*^{f} and *Esrrb*^{fn} alleles demonstrates that excision does not occur without tamoxifen treatment.

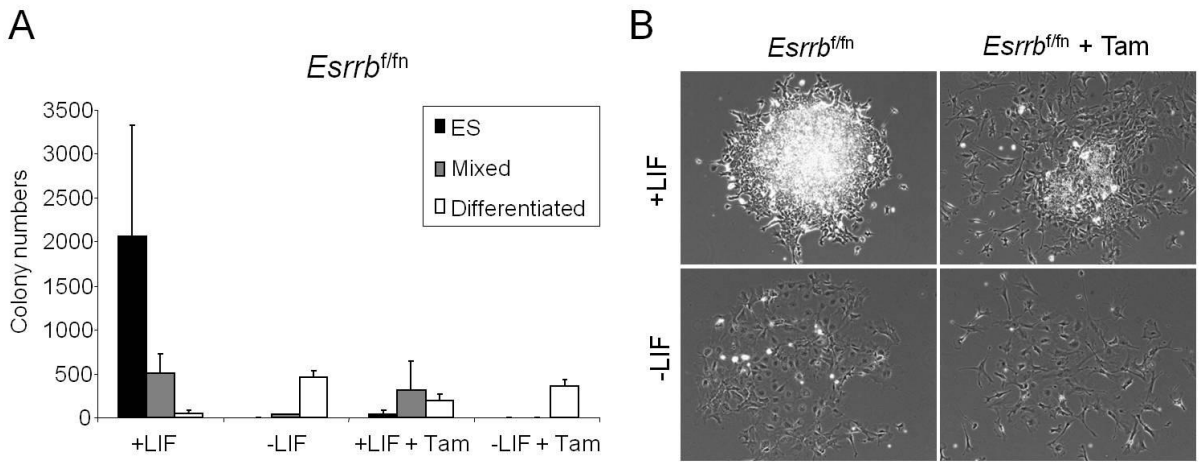


Figure 8.2: Loss of Esrrb compromises ES cell clonogenic potential.

A: Number of colonies formed by *Esrrb*$^{f/fn}$ cells plated at clonal density, treated or not treated with tamoxifen for 1h on the following day, and cultured in the presence or absence of LIF for 7 days. **B**: Morphology of the colonies observed in the same experiment (see A) 7 days after plating.

Since almost exclusively partially differentiated colonies formed after Esrrb excision, it was next important to test whether it was possible to expand these clones and isolate ES cells genetically devoid of Esrrb. To this end, *Esrrb*$^{f/fn}$ cells were plated at clonal density in GMEMβ/FCS/LIF supplemented with 1mM PD0325. In the presence of this inhibitor, cells that failed to recombine the *Esrrb* locus formed tight colonies that completely lacked morphological signs of differentiation. In contrast, clones that had excised *Esrrb* showed substantial differentiation, but could be selectively picked and expanded. Stable *Esrrb*$^{\Delta/\Delta}$ cell lines that can be propagated in FCS/LIF/GMEMβ were isolated. Productive recombination of both alleles was confirmed by quantitative PCR on genomic DNA samples using primer pairs binding to the targeted exon 2 of *Esrrb*. Amplification of *Esrrb* exon 6 served as a control to normalise for the amount of template used in each reaction. Whereas amplification could be detected in *Esrrb*$^{f/fn}$ cells, no *Esrrb* exon 2 was present in the genomic DNA prepared from *Esrrb*$^{\Delta/\Delta}$ ES cells (**Figure 8.3A**). As expected, neither Esrrb transcript (**Figure 8.3B**) nor protein (**Figure 8.3C**) could be detected in these cells. *Esrrb*$^{\Delta/\Delta}$ cells have a markedly increased propensity to differentiate but the self-renewing fraction of the cultures maintains Oct4 expression (**Figure 8.3C**).

The self-renewal ability of *Esrrb*$^{\Delta/\Delta}$ cells was then tested in clonal assays. Whereas Esrrb$^{f/fn}$ ES cells efficiently formed undifferentiated colonies when plated in the presence of LIF, *Esrrb*$^{\Delta/\Delta}$ cells formed reduced numbers of colonies, composed of a majority of AP negative cells (**Figure 8.3D**).

Taken together these results establish that Esrrb is formally dispensable for ES cell self-renewal. Nonetheless, in the absence of Esrrb self-renewal is inefficient and cells show a severely impaired clonogenic potential.

8.2: Effects of Esrrb knock-out on Nanog overexpressing ES cells.

Nanog was originally identified on the basis that its overexpression in ES cells

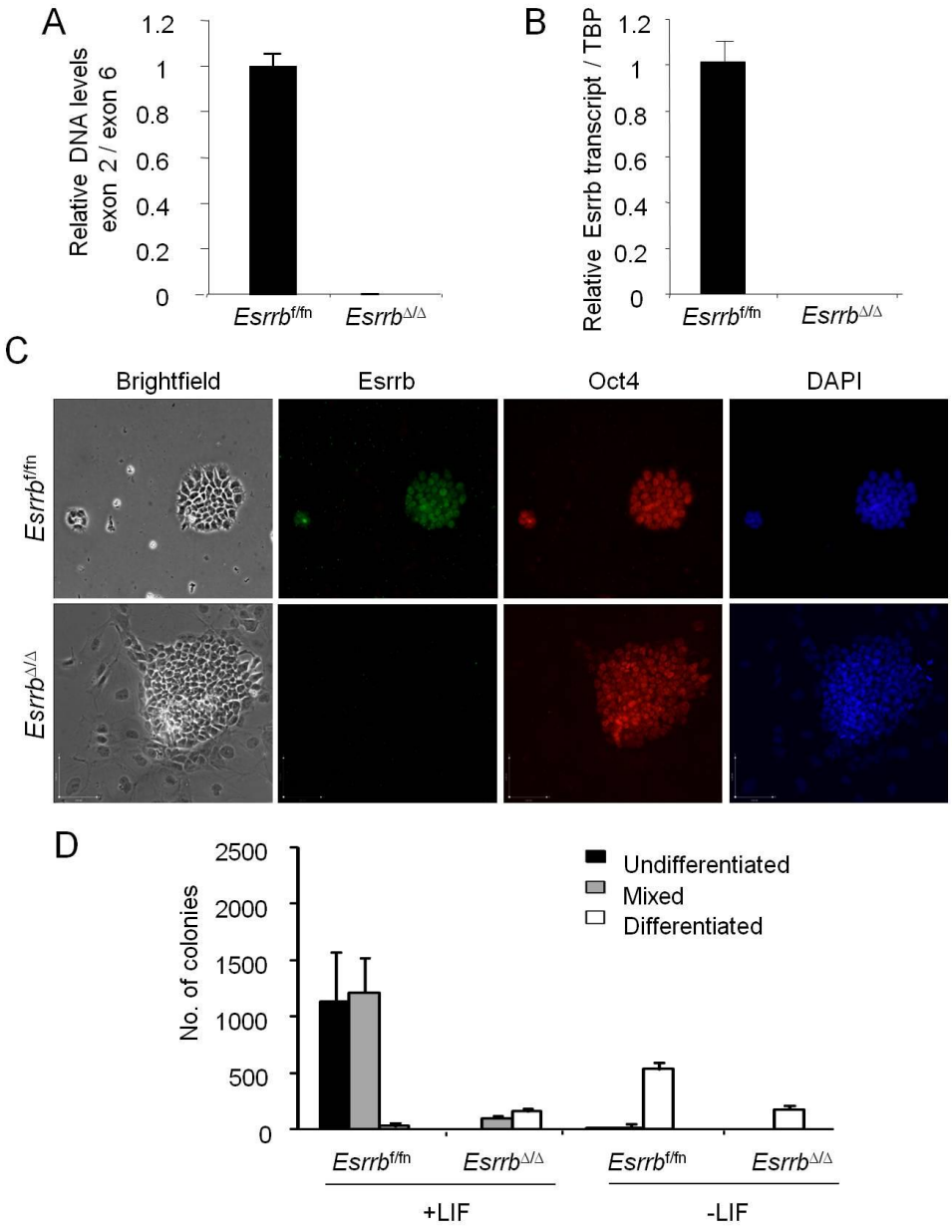


Figure 8.3: Esrrb is formally dispensable for ES cell self-renewal.

A: Quantitative PCR analysis of the levels of *Esrrb* exon 2 relative to *Esrrb* exon 6 in *Esrrb*$^{f/fn}$ and *Esrrb*$^{\Delta/\Delta}$ ES cells. Error bars are the standard deviation of the 3 technical replicates of one experiment. **B**: Esrrb transcript levels in *Esrrb*$^{f/fn}$ and *Esrrb*$^{\Delta/\Delta}$ cells. Error bars are the standard deviation of the 3 technical replicates of one experiment. **C**: Morphology and expression of Oct4 and Esrrb in *Esrrb*$^{f/fn}$ and *Esrrb*$^{\Delta/\Delta}$ cells. **D**: Number of the colonies formed by *Esrrb*$^{f/fn}$ or Esrrb$^{\Delta/\Delta}$ cells plated at clonal density and cultured in the presence or absence of LIF for 7 days.

confers LIF-independent self-renewal (Chambers et al., 2003). If the hypothesis that Nanog mediates a significant part of its function via Esrrb is correct, then Esrrb knockout should reduce the self-renewal efficiency conferred by Nanog overexpression.

Esrrb$^{f/fn}$ cells were transfected with a Nanog transgene or a control Esrrb transgene. 6 clonal lines (*Esrrb*$^{f/fn}$+Nanog and *Esrrb*$^{f/fn}$+Esrrb) were isolated that overexpressed Nanog and Esrrb at different levels (**Figure 8.4A-B**). *Esrrb*$^{f/fn}$+Nanog and *Esrrb*$^{f/fn}$+Esrrb cells were plated at clonal density and treated with tamoxifen as described to induce deletion of *Esrrb.* As expected, tamoxifen treatment of *Esrrb*$^{f/fn}$+Esrrb cells did not result in any apparent change in the morphology and the degree of differentiation of the colonies formed. Colonies originated from *Esrrb*$^{f/fn}$+Nanog treated with tamoxifen and lost the tight undifferentiated morphology characteristic of Nanog overexpressing cells. Nonetheless, tamoxifen treatment of *Esrrb*$^{f/fn}$+Nanog cells never resulted in the evident loss of self-renewal ability observed after treatment of parental Esrrb$^{f/fn}$ lines, indicating that Nanog overexpression facilitates the self-renewal of Esrrb$^{\Delta/\Delta}$ ES cells (**Figure 8.4 C**).

Colonies originated form *Esrrb*$^{f/fn}$+Nanog and *Esrrb*$^{f/fn}$+Esrrb tamoxifen treated cells were picked and expanded and 6 lines that had excised Esrrb exon2 were identified by quantitative PCR on genomic DNA (**Figure 8.5A**). The absence of Esrrb transcript was also confirmed in *Esrrb*$^{f/fn}$+Nanog cells (**Figure 8.5B**).

It was then tested whether loss of endogenous Esrrb expression in *Esrrb*$^{f/fn}$+Nanog cells would impair Nanog driven LIF independent self-renewal. Six Nanog and six Esrrb overexpressing *Esrrb*$^{f/fn}$ and derivative *Esrrb*$^{\Delta/\Delta}$ lines were plated at clonal density in the presence or absence of LIF and the number of AP^{+} colonies scored after 7 days. *Esrrb*$^{f/fn}$+Nanog cells showed robust self-renewal both in the presence or absence of LIF (**Figure 8.5C-D**). Enforced Esrrb expression gave comparable numbers of undifferentiated colonies in these conditions, (**Figure 8.5C-D**) and this

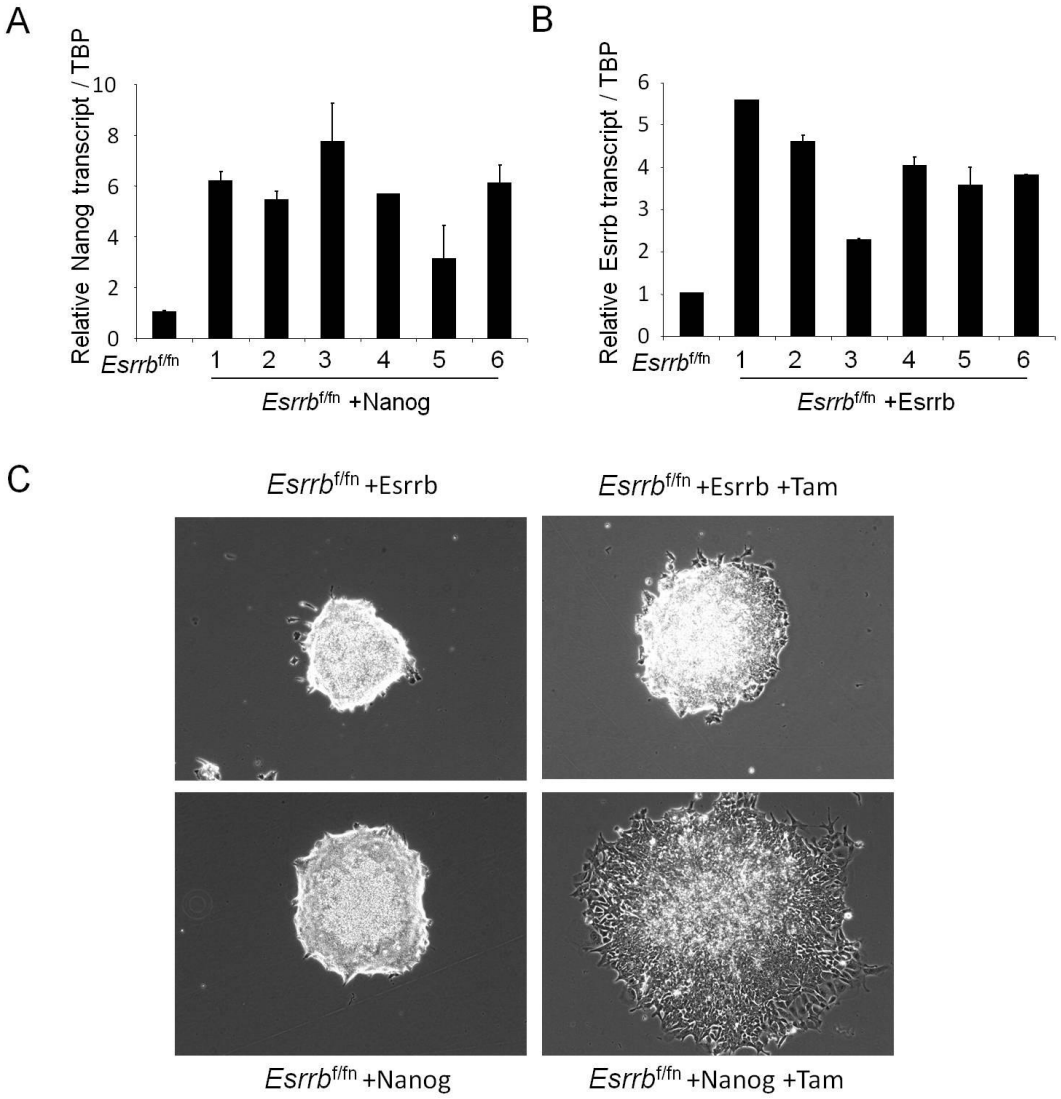


Figure 8.4: Loss of Esrrb affects the phenotype conferred to ES cells by Nanog overexpression.

A: Nanog expression in *Esrrb*$^{f/fn}$ and *Esrrb*$^{f/fn}$ + Nanog ES cells. Error bars are the standard deviation of the 3 technical replicates of one experiment. **B**: Esrrb expression in *Esrrb*$^{f/fn}$ and *Esrrb*$^{f/fn}$ +Esrrb ES cells. Error bars are the standard deviation of the 3 technical replicates of one experiment. **C**: Morphology of the colonies formed by *Esrrb*$^{f/fn}$ +Esrrb and *Esrrb*$^{f/fn}$ +Nanog ES cells plated at clonal density, treated with tamoxifen for 1h on the following day, and cultured in the presence of LIF for 7 days.

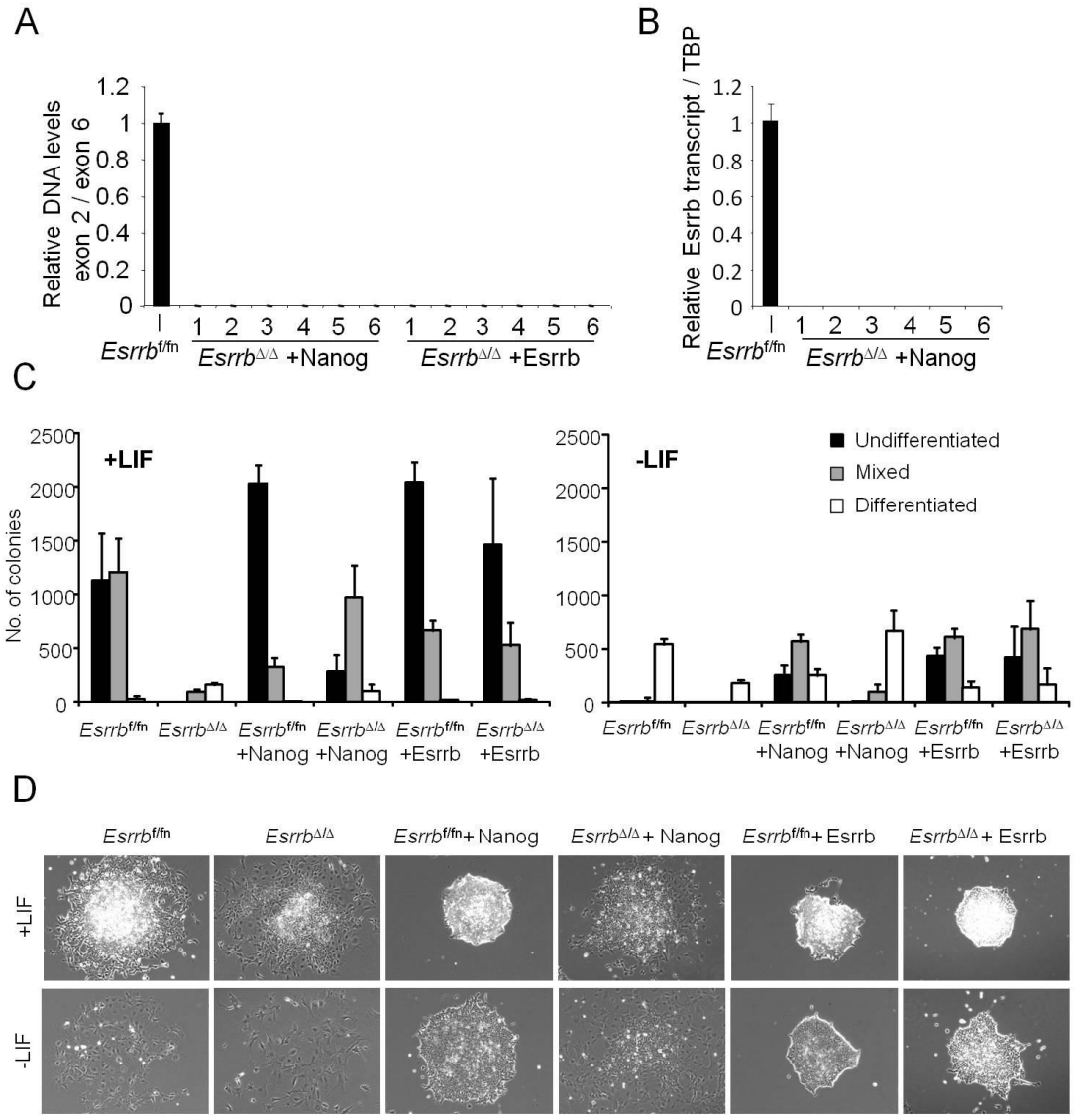


Figure 8.5: Esrrb demolishes Nanog driven LIF independent self-renewal.

A: Quantitative PCR analysis of the levels of *Esrrb* exon 2 relative to *Esrrb* exon 6 in *Esrrb*$^{f/fn}$, *Esrrb*$^{\Delta/\Delta}$ +Nanog and *Esrrb*$^{\Delta/\Delta}$ +Esrrb ES cells. Error bars are the standard deviation of the 3 technical replicates of one experiment. **B**: Esrrb expression in *Esrrb*$^{f/fn}$ and *Esrrb*$^{\Delta/\Delta}$ +Nanog cells. Error bars are the standard deviation of the 3 technical replicates of one experiment. **C**: Average number of colonies scored after clonal density plating of 6 independent *Esrrb*$^{\Delta/\Delta}$ +Nanog and *Esrrb*$^{\Delta/\Delta}$ +Esrrb lines and culture for 7 days in the presence or absence of LIF. Error bars: standard deviation of the results obtained from 6 lines each analysed in triplicate. **D**: Morphology of the colonies formed by the indicated lines after 7 days of culture in the presence or absence of LIF.

number was unaltered in *Esrrb*$^{\Delta/\Delta}$+Esrrb ES cells. As observed during the derivation of these lines, *Esrrb*$^{\Delta/\Delta}$+Nanog ES cells showed decreased self-renewal efficiency when plated in the presence of LIF (**Figure 8.5C-D**). Strikingly, the defining ability of Nanog to promote LIF independence in ES cells was completely demolished by loss of Esrrb. *Esrrb*$^{\Delta/\Delta}$+Nanog ES cells primarily formed differentiated colonies in the absence of LIF (**Figure 8.5c-d**). These observations establish that activation of Esrrb expression mediates an important component of Nanog function in ES cells.

8.3: Esrrb and Nanog share target genes

Taken together the results presented in the previous chapters argue in favour of the existence of a degree of functional overlap between Esrrb and Nanog activity in pluripotent cells. This could be due to similarities in the transcriptional programme activated by Nanog and Esrrb expression. Thus, the transcriptional response to Nanog and Esrrb elevation was compared by microarray analysis of doxycycline treated ESΔN-iNanog and ESΔN-iEsrrb cells. Since ESΔN-iEsrrb cells lack functional Nanog protein, it is possible to exclude that the observed effects of Esrrb induction are indirectly due to the elevation of Nanog expression (van den Berg et al., 2008). Strikingly, an overall similar transcriptional response was detected after Esrrb or Nanog induction (**Figure 8.6A**). Next, the overlap between the 50 genes most upregulated in response to Nanog or Esrrb activation was determined. In order to focus the analysis on potential direct targets of transcriptional activation by both factors, only genes for which binding of Nanog and Esrrb had been reported in two independent ChIP-seq studies were considered (Chen et al., 2008; Marson et al., 2008). 20% of the top 50 upregulated genes were common between ESΔN-iNanog and ESΔN-iEsrrb cells datasets (**Figure 8.6B**). Klf4 was the only transcription factor in this group and, interestingly, Nanog induced Klf4 more effectively than Esrrb (**Figure 8.6C**).

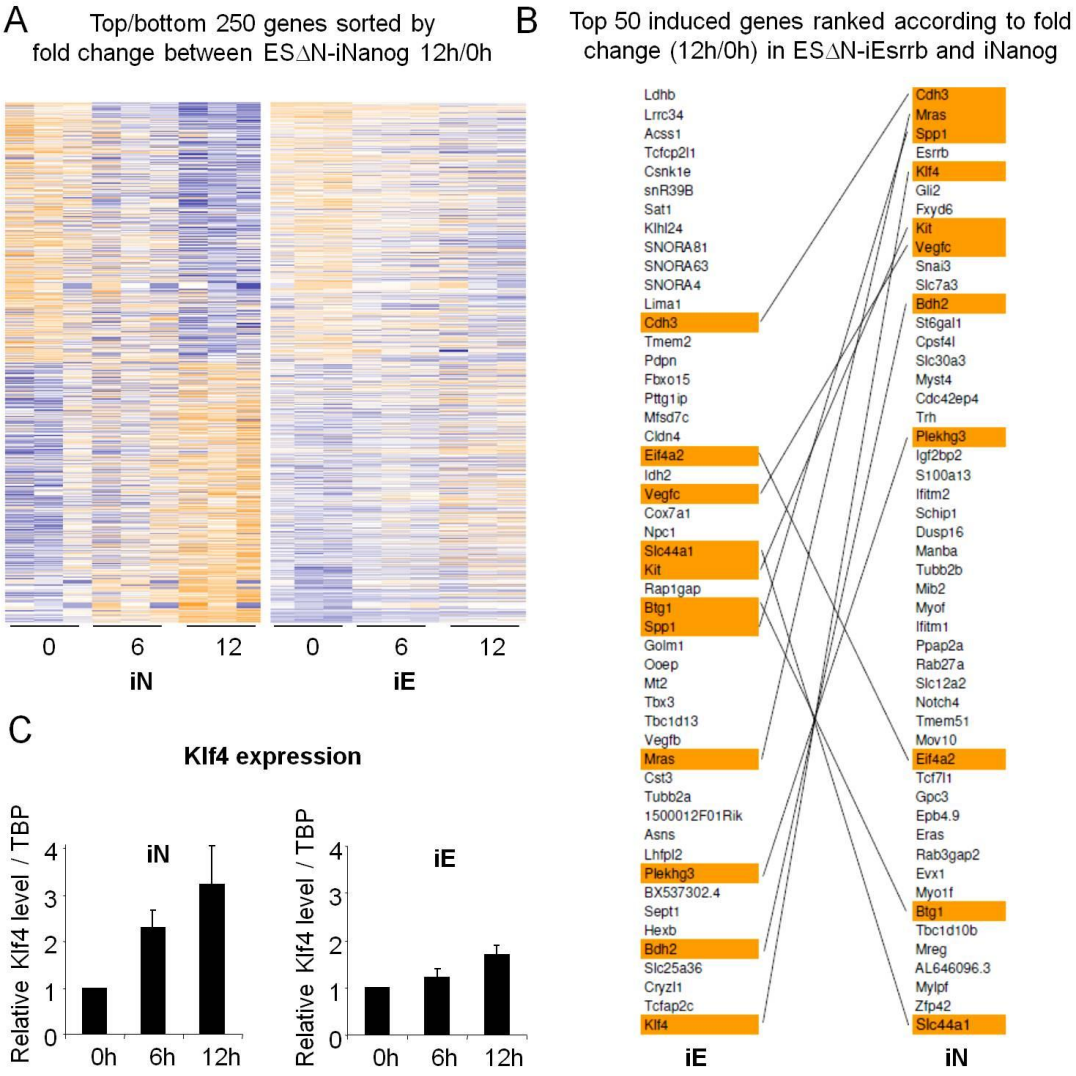


Figure 8.6: Esrrb demolishes Nanog driven LIF independent self-renewal.

A: Heatmap showing expression (in ESΔN-iNanog (iN) or ESΔN-iEsrrb (iE) cells) of the top 250 genes that were induced or repressed in ESΔN-iNanog (0/12h). Colours: expression above average (yellow) and below average (blue). **B**: Comparison of the 50 genes most upregulated in ESΔN-iNanog (iN) and ESΔN-iEsrrb (iE) after a 12h doxycycline treatment. Only Nanog and Esrrb bound genes were considered and ranked according to fold induction in ESΔN-iNanog or ESΔN-iEsrrb respectively. Shared targets are highlighted (orange) and connected by black lines. **C**: Expression of Klf4 in ESΔN-iNanog (iN) and ESΔN-iEsrrb (iE) cells after doxycycline treatment for the indicated time. Error bars: standard deviations of 3 independent experiments.

These results suggest that the observed similarities in the activity of Esrrb and Nanog in ES cells might arise from the modulation of a common set of transcriptional targets.

8.4: Discussion

The data presented in this chapter shows that deletion of Esrrb severely impairs the ability of ES cells to self-renew, reminiscent of the consequences of loss of Nanog expression (Chambers et al., 2007). Nonetheless, both *Esrrb*$^{-/-}$ and *Nanog*$^{-/-}$ ES cell can be derived. It would be now intriguing to determine whether ES cells are able to self-renew in the absence of both Nanog and Esrrb. The results presented in this and in the previous chapters also reveal that a high degree of mutual dependence exists between Nanog and Esrrb function in ES cells and that these two factors conjunctly regulate expression of many target genes (**Figure 8.6**). It is thus crucial to understand which pluripotency GRN targets are sensitive to the combined loss of both factors. A promising candidate is Klf4. Klf4 is activated by both Nanog and Esrrb (**Figure 8.6C**), but its transcriptional elevation in response to Esrrb induction is less pronounced than in response to Nanog. Since transcriptional profiling experiments were performed in *Nanog*$^{-/-}$ cells, it is possible that the absence of Nanog binding to *Klf4* impairs its activation by Esrrb. To start dissecting the mechanisms of Esrrb and Nanog cooperativity, it would be interesting to determine whether Esrrb elevation in cells expressing wild-type levels of Nanog results in enhanced Klf4 upregulation. Conversely, the response of Klf4, and other target genes, to the induction of Nanog should be analysed in Esrrb$^{\Delta/\Delta}$ lines. The extent to which Nanog and Esrrb depletion reciprocally affect their binding to known target sequences also remains unexplored. Few genes radically alter their expression in response to Nanog or Esrrb induction (see chapter 3 and **Figure 8.6**). If Nanog and Esrrb conjunctly regulate transcription, the number of responsive genes should increase dramatically in response to concomitant restoration of Nanog and Esrrb binding in double knock-out cell lines.

While both Nanog and Esrrb can be deleted, Oct4 and Sox2 expression is absolutely required in ES cells (Avilion et al., 2003; Masui et al., 2007; Nichols et al., 1998;

Niwa et al., 2000). In addition, Nanog (see Chapter3) or Oct4 (Hall et al., 2009) manipulation have profoundly different transcriptional consequences. As discussed in chapter 3, this suggests that Oct4 activity is essential for a crucial part of the housekeeping functions performed by the pluripotency GRN in ES cells. In contrast, other factors, such as Nanog and possibly Esrrb, may be involved in precisely tuning the expression of a limited number of genes to confer flexibility to the pluripotency network. Deletion of this second class of pluripotency factors could therefore be possible as a consequence of the limited transcriptional alterations that ensue in knock-out ES cell lines. Alternatively, the deletion of Nanog and Esrrb could be achievable in virtue of the ability of one factor to functionally compensate for absence of the other. A similar functional complementation has already been reported for Klf4, whose depletion does not result in ES cell differentiation (Jiang et al., 2008). Knock down studies have demonstrated that Klf4, Klf2 and Klf5 have redundant functions and that two but not all three factors can be depleted in ES cells. Klf2, Klf4 and Klf5 are highly related proteins and have almost completely overlapping binding sites throughout the genome (Jiang et al., 2008). Esrrb and Nanog are structurally distinct proteins (Chambers and Tomlinson, 2009; Gearhart et al., 2003) and show less extensive binding overlap (Chen et al., 2008). This would explain why depletion of single Klf factors has very limited consequences on ES cell self-renewal while Esrrb and Nanog null lines show much more severe phenotypes.

Tbx3, another heterogeneously expressed factor, has been shown to be required for ES cell self-renewal in knockdown experiments (Ivanova et al., 2006). Despite this, and similarly to what observed for Esrrb and Nanog, it could be possible to derive Tbx3 null lines. Combinatorial ablation of Nanog, Esrrb, Klfs and Tbx3 and analysis of the degree of functional compensation between these factors could expand our understanding of the pluripotency GRN topology. Functional studies assessing the transcriptional effect of the concomitant elevation or ablation of more than one factor will shed light on the extent and the mechanisms of combinatorial control sustaining gene expression in pluripotent ES cells.

Chapter 9:Concluding remarks and future directions

9.1: Nanog and Esrrb: overlapping and distinct functions in vitro and at different stages of development

Two essential functions have been reported for Nanog during development: Nanog is required for the establishment of pluripotency in pre-implantation embryos (Mitsui et al., 2003; Silva et al., 2009) and for the progression of germline development beyond E12 (Chambers et al., 2007). In vitro, Nanog is required for efficient ES cell self-renewal (Chambers et al., 2007) and for completion of reprogramming of differentiated cells to pluripotency (Silva et al., 2009). A comprehensive analysis of the data presented in this thesis may help clarify similarities and differences amongst the multiple roles that Nanog plays in acquisition and loss of pluripotency in these different contexts. The results presented in the previous chapters also reveal a high degree of mutual dependence between the functions of Nanog and Esrrb in pluripotent cells. Therefore, an emergent question is whether the observed overlap in function is conserved in different settings.

A central part of this thesis has focussed on the activity that Nanog and Esrrb exert in vitro in sustaining pluripotency. Elevation of either Nanog (Chambers et al., 2003) or Esrrb (see chapter 4) maintains ES cell pluripotency independently of external stimuli whereas genetic ablation of Nanog (Chambers et al., 2007) and Esrrb (see chapter 8) results in compromised self-renewal efficiency. In the previous chapter, it was shown that the ability of Nanog to relieve LIF dependence in ES cells is abrogated in the absence of Esrrb (see **Figure 8.5C-D**). Conversely, it was possible to observe a reduced effect of Esrrb overexpression in LIF independence assays performed in *Nanog*$^{-/-}$ compared to wild-type ES cells (see Chapter 5). Finally, concomitant Nanog and Esrrb downregulation coincides with loss of self-renewal ability (see Chapter 4). Thus, Nanog and Esrrb activity seem to be pronouncedly cooperative. These observations could be rationalised proposing that Nanog and

Esrrb function primarily in pluripotent cells by synergistically regulating the expression of a common set of target genes (see **Figure 8.6)**. Relevant to this point, since Nanog and Esrrb proteins interact (van den Berg et al., 2010; Wang et al., 2006a), these two factors may reciprocally enhance their recruitment to common target genes. In addition, one factor could bridge the interaction of its partner to important transcriptional activators or repressors. Intriguingly, Esrrb has been shown to interact with multiple components of the basal transcriptional machinery (Percharde et al., 2012; van den Berg et al., 2010) and could be required to mediate transcriptional activation by Nanog. Alternatively, both Nanog and Esrrb could independently bind and additively activate a common set of targets important for ES cell self-renewal, so that maximal induction levels are achieved only when both factors are expressed. In any case, Esrrb and Nanog activity in ES cells might not merely involve direct activation or repression of transcription, since both factors take part in extensive protein-protein interactions (Liang et al., 2008; van den Berg et al., 2010; Wang et al., 2006a) and might be required to recruit epigenetic modifiers and chromatin remodelling complexes to crucial locations in the genome.

After establishment of pluripotency in the pre-implantation blastocyst, Nanog expression seems to be dispensable for further progression of somatic development (Chambers et al., 2007). Nonetheless, a role for Nanog in regulating the transition between pre-implantation and post-implantation pluripotency cannot be excluded. Nanog expression is transiently down-regulated in the epiblast immediately before implantation (Chambers et al., 2003), becoming detectable again in the posterior epiblast before the onset of gastrulation ((Osorno et al., 2012) and see chapter 1). In the epiblast, implantation coincides with downregulation of many transcription factors that are highly expressed in the Nanog positive pluripotent ICM of the early embryo, such as Rex1 (Pelton et al., 2002), Klf5 (Ema et al., 2008), Tbx3 (Bollag et al., 1994; Chapman et al., 1996), and Esrrb (Guo et al., 2010; Luo et al., 1997). In chapter 3, evidence was presented that some of these pre-implantation markers (Esrrb, Rex1, Klf4, Klf5 and Tbx3) are highly responsive to Nanog activity in ES cells (see **Figure 3.7**, and **Table 3.1**). This implies a hypothesis that Loss of Nanog

expression in the epiblast might be required to allow downregulation of these genes, thereby facilitating progression of development.

An important tool for understanding the contribution of Nanog and expression of its target genes in regulating the transition from the pre- to post-implantation epiblast comes from reprogramming studies of EpiSCs. In parallel with the changes occurring during development, the transition between ES cells and EpiSCs is accompanied by the silencing of a number of pre-implantation pluripotency markers such as Esrrb, Klf4, Rex1, Stella, Nr0b1 and Tbx3 (Festuccia et al., 2012; Guo et al., 2009). Among the core pluripotency factors, Oct4 and Sox2 expression is maintained, while, notably, Nanog expression is strongly reduced (Guo et al., 2009; Osorno and Chambers, 2011). Intriguingly, overexpression of Nanog accompanied by a change in culture conditions was shown to drive reprogramming of EpiSCs to ES-like cells (Osorno and Chambers, 2011; Silva et al., 2009) and this conversion is accompanied by acquisition of a gene expression profile characteristic of "naïve", pre-implantation, pluripotency (Silva et al., 2009). A similar reprogramming capacity has been reported for c-Myc (Hanna et al., 2009), Nr5a2 (Guo and Smith, 2010), Stat3 (Yang et al., 2010) and, notably, Klf factors (Gillich et al., 2012; Guo et al., 2009; Hall et al., 2009; Hanna et al., 2009) as well as Esrrb (Festuccia et al., 2012). It is therefore possible to hypothesise that Nanog action in directing the conversion of EpiSC to ES cells is mediated through upregulation of a limited number of its target genes, as shown for Klf4, and Esrrb. In support of this interpretation, although Nanog expression is limiting for EpiSC reprogramming (Silva et al., 2009), Esrrb elevation can bypass this requirement (Festuccia et al., 2012). Notably, Esrrb drives reprogramming with compromised efficiency in the absence of Nanog (Festuccia et al., 2012). Esrrb is not expressed in EpiSC (Festuccia et al., 2012) and Nanog levels are reduced in this population (Guo et al., 2009; Osorno and Chambers, 2011). Given the high degree of cooperativity observed between Nanog and Esrrb (see chapter 5, 8 and (Festuccia et al., 2012)), it is possible that the ability of Nanog to activate a set of genes crucial for the maintenance of naïve pluripotency is strongly reduced in EpiSC due to the low expression level and the lack of synergistic activity with Esrrb.

In contrast to mouse ES cells, EpiSC are not responsive to LIF signalling (Tesar et al., 2007; Yang et al., 2010). Interestingly, Nanog, Esrrb and Klf4, all factors able to drive reprogramming of EpiSC, can confer LIF independence to mouse ESC ((Chambers et al., 2003; Niwa et al., 2009) and see chapter 4). It is thus possible that these three factors (and possibly other Nanog targets) as well as the cascade activated by LIF converge on a similar set of targets to impose naïve pluripotency. In support of this interpretation, LIF/STAT3 signalling has been shown to be limiting for EpiSC reprogramming (Yang et al., 2010), but Nanog and Esrrb expression can bypass this requirement (Festuccia et al., 2012; Theunissen et al., 2011b).

Similar to what is observed for murine EpiSC, in human cells Nanog is expressed, but Esrrb is absent (Xie et al., 2009). Overexpression of Nanog has been used in combination with other canonical reprogramming factors to allow generation of human ES cells that present traits of naive pluripotency characteristic of mouse ESC (Buecker et al., 2010; Li et al., 2009; Pomp et al., 2011). Naive hESC can be established from hESC lines (Hanna et al., 2010) or during direct reprogramming from fibroblasts (Buecker et al., 2010; Hanna et al., 2010; Li et al., 2009; Pomp et al., 2011). These cells show robust expression of Nanog and other preimplantation markers, albeit they remain dependent on transgene expression or complex combinations of signal inhibitors (Hanna et al., 2010). In particular, addition of Nanog to the canonical reprogramming cocktail OKSM was reported to be strictly required for the generation of naive hESC (Buecker et al., 2010). Other investigators claimed that conversion of hESC can be achieved in the absence of Nanog using reprogramming cocktails that notably always include Nanog targets able to reprogramme EpiSC (Hanna et al., 2010). It would be interesting to test the effects of ectopic Esrrb expression in hESC lines, and whether naïve hESC can be generated and maintained solely by Esrrb elevation.

Extending the conclusions drawn from these in vitro studies to in vivo development, it is thus possible that Nanog plays a crucial role in maintaining the unrestricted pluripotent state characteristic of the pre-implantation ICM and that loss of Nanog is instrumental in ensuring rapid and coordinated transcriptional changes in epiblast

cells upon implantation. Since the overexpression of single Nanog targets is sufficient in vitro to drive establishment of naïve pluripotency, it is possible to imagine that Nanog might influence the pre- to post-implantation transition mainly through its ability to elevate expression of key pluripotency factors. In line with the existence of an overlap between Nanog and Esrrb targets, Esrrb is highly efficient in complementing Nanog function in this context (Festuccia et al., 2012).

Whereas Esrrb seems to be able to compensate for the absence of Nanog in promoting the transition between two pluripotent populations, such as EpiSC and ESC, Esrrb driven reprogramming of Nanog null somatic cells requires the additional activity of 5'azacytidine (see chapter 6). This compound has been shown to reverse methylation marks at the promoter and regulatory elements of pluripotency genes and enhance reprogramming (Huangfu et al., 2008; Mikkelsen et al., 2007), suggesting that Nanog plays a specific function in driving epigenetic resetting during the acquisition of pluripotency. Importantly, 5'-azacytidine alone is not sufficient to complete reprogramming of stalled Nanog null pre-iPS cells. These observations seem to highlight the existence of two components in the activity that Nanog exerts in this context. The first, transcriptional activation of crucial target genes, can be functionally complemented by elevation of Esrrb expression. The second, which involves triggering epigenetic rearrangement and reversion of methylation marks, might be uniquely performed by Nanog. This activity could be mediated through protein-protein interactions with chromatin remodelling complexes (Liang et al., 2008; Wang et al., 2006a) or by inducing expression of epigenetic modifiers. Curiously, the levels of Tet2, a member of a class of enzymes involved in the hydroxylation of 5meC (Ito et al., 2010; Tahiliani et al., 2009) and erasure of DNA methylation, are severely reduced in Nanog null ES cells (see **Figure 3.2** and **Table 3.1**). Further studies should determine whether defective Tet2 upregulation is also noticed during reprogramming of Nanog null cells, and whether reduced Tet2 expression has functional consequences in this context. In particular, it would be interesting to test whether concomitant elevation of Esrrb and Tet2 might result in successful reprogramming of Nanog null pre-iPS without need for 5'-azacytidine addition.

The data presented in chapter 7 shows that precocious and elevated Esrrb expression is able to at least partially compensate for the absence of Nanog during PGC development. It remains to be established however whether Nanog and Esrrb exert their function in this context primarily by sustaining pluripotent gene expression and repressing somatic fate during PGCs migration to the genital ridges or whether these two factors have an active role in promoting the epigenetic changes that accompany germline development (Hajkova et al., 2008; Hajkova et al., 2002; Hajkova et al., 2010; Seki et al., 2005; Seki et al., 2007). Since broad similarities exist between the molecular events underlying germline and pre-implantation development (see chapter 1 and 7), it is possible that elevated Esrrb expression might also result in the functional complementation of the defects observed in Nanog null blastocysts. This possibility should be tested by crossing heterozygous Esrrb KI animals (see chapter 7.2) and analysing progression of development in homozygous embryos. Relevant to the role of Esrrb in pre-implantation development, Esrrb transcripts, and possibly Esrrb protein, are likely expressed in maturing oocytes, since high levels of Esrrb are detected in 1 cell embryos before activation of the zygotic genome (Guo et al., 2010). In addition, Esrrb expression is also detected in the testes of newborn animals (Mitsunaga et al., 2004). The function of Esrrb in oogenesis and spermatogenesis has never been addressed and, similarly, the role of Esrrb in sustaining development in early cleavage embryos remains unexplored. Since conditional knockout alleles are available for Esrrb ((Chen and Nathans, 2007) and see chapter 8), future studies should assess the consequence of genetic ablation of this factor in postmigratory primordial germ cells of male and female animals, or its removal from the egg, by tissue specific expression of Cre recombinase (Gallardo et al., 2007; Lewandoski et al., 1997).

In summary, in the different in vivo and in vitro contexts examined, Nanog seems to exert its activity either through direct transcriptional activation or repression of specific targets, a relatively well characterised function, or through modulation of the epigenetic state in the cells, a process that remains uncharacterised in its specificity and mechanisms of action. Of interest, reactivation of the inactive X chromosome seems to be a common step in all the transitions in which Nanog is involved in vivo

and in vitro. Reversion of imprinted paternal X chromosome inactivation is specifically completed in the ICM cells of the late blastocyst that maintain Nanog expression (Mak et al., 2004; Okamoto et al., 2004). Coincident with downregulation of Nanog, random X inactivation also starts in female embryos after implantation (Rastan, 1982). Conversely, in vitro reprogramming of EpiSC, a process efficiently driven by Nanog, is marked by reactivation of the silent X chromosome in female lines (Silva et al., 2009). In addition, X reactivation and expression of endogenous Nanog are detected at similar stages during transcription factor based reprogramming of somatic cells (Stadtfeld et al., 2008). Finally, during their migration to the genital ridges, PGCs gradually reverse random X inactivation (Sugimoto and Abe, 2007) and this process initiates in migratory PGCs soon after reactivation of Nanog expression (Yamaguchi et al., 2005). This is in agreement with the observation that Nanog plays a direct role repressing *Xist* transcription in conjunction with Oct4 and Sox2 (Navarro et al., 2008). Apart from the fact that Esrrb binds to *Xist* in ES cells (Pablo Navarro, unpublished information), the ability of Esrrb to complement Nanog activity in repressing X inactivation, and the possible mechanism through which such compensation is achieved, remain completely unexplored.

Irrespective of their functional overlap in pluripotent cells, Esrrb has an essential role in additional processes in which Nanog is not involved, consistent with its much broader expression pattern ((Bookout et al., 2005; Bookout et al., 2006) and see chapter 1). In particular, Esrrb is highly expressed in the extraembryonic ectoderm and is required for placental development (Luo et al., 1997). Esrrb function in such context seems to be related to the maintenance of the proliferating population of trophoblast cells responsible for most of the embryonic contribution to placental tissues (Luo et al., 1997). Trophoblast stem (TS) cell lines can be can be derived from post-implantation extraembryonic ectoderm explants and the trophectoderm of pre-implantation embryos (Tanaka et al., 1998). Esrrb is highly expressed in TS cells and inhibition of estrogen related receptors by diethylstilbestrol results in TS cell differentiation into polyploid trophoblast giant cells (Tremblay et al., 2001b). Taking advantage of the available inducible knockout and reporter lines, it would be

interesting to explore the functional role of Esrrb in sustaining TS cell self-renewal and the dynamic control of its expression in this population.

9.2: Nanog control over transcription: live imaging of transcriptional dynamics

The results presented in chapter 3 show that Nanog elevation affects the transcriptional output of a limited number of genes, among which are important regulators of self-renewal, such as Esrrb and Klf4. In chapter 4, the dynamics of the control of Esrrb expression by Nanog were analysed using fluorescent reporter lines: Nanog and Esrrb protein levels show a clear correlation in single ES cells. If similar expression patterns could be observed for other genes, it would be important to determine whether fluctuations in Nanog levels result in the synchronous and coordinated transcriptional modulation of a cohort of its targets in individual cells. Indeed, Esrrb and Klf4 fluctuations are linked in ES cells and both factors show correlation with Nanog expression (**Figure 4.17**). Since such correlation was determined at the protein level, and given that Nanog influences expression of its targets at the transcriptional level, it could be particularly informative to extend our analysis to the study of active transcription from both *Esrrb* and *Klf4* alleles in single cells. qPCR could be employed to quantify expression of Nanog, Esrrb, Klf4 and other Nanog target genes in single ES cells. Nonetheless, such an approach would only detect eventual correlations in the levels of mature transcripts of these genes. Since elevation of mRNA levels results from their accumulation over prolonged periods of transcription, and any decrease is influenced by the stability of the mRNAs, this technique does not have the potential to establish a direct link between Nanog protein levels and transcriptional activity of its target genes. Initially, FISH and immunohistochemistry should be employed to determine whether a correlation exists between Nanog protein levels and the transcriptional output of *Klf4*, *Esrrb* and possibly other Nanog targets. Such experiments could reveal: 1) A correlation in the numbers of actively transcribing *Klf4* and *Esrrb* alleles and Nanog levels across the population of ES cells. 2) A correlation between Nanog levels and the frequency of

biallelic transcription of *Esrrb* and *Klf4*. 3) Synchronicity in the expression of *Esrrb* and *Klf4* in individual cells. 4) Colocalisation of transcribing *Esrrb* and *Klf4* alleles to specific nuclear regions. This last point is of particular relevance to the understanding of the mechanisms through which transcription is regulated in ES cells. Recent reports highlighted the existence of long range interactions between different chromatin domains in mammalian (Li et al., 2012) and, more specifically, ES (Dixon et al., 2012) cells. Such interactions often involve promoters and enhancers of genes located on distant regions on the same chromosome or on different chromosomes, and it has been shown that such multigene complexes tend to localise to foci of high RNApolII density in the nucleus (Li et al., 2012). The model proposing that active alleles segregate to transcription "factories", first proposed by Cook and co-workers (Iborra et al., 1996), refuses the idea that freely diffusing RNApolII is recruited to the promoter of actively transcribed genes. This model rather hypothesises that the alleles of different genes colocalise to structured, possibly pre-existent domains of RNApolII, transcription factor, and splicing factor accumulation that show reduced mobility (reviewed in (Edelman and Fraser, 2012; Sutherland and Bickmore, 2009)). Recruitment of enhancers and promoters of different genes to common regions of the nucleus could provide a basis for the coordinated regulation of gene expression (Li et al., 2012) in ES cells. It would be interesting to understand whether such long range interactions are bridged by transcriptional regulators, like Nanog, and whether Nanog promotes recruitment of its target genes to these transcription factories. FISH experiments should be repeated in Nanog null ES cells, to determine whether the localisation of active *Esrrb* and *Klf4* alleles, or other Nanog targets, to particular nuclear domains is lost in the absence of Nanog. Since Nanog and Esrrb seem to share a consistent number of target genes, including Klf4, similar experiment could be also performed comparing wild-type and Esrrb knockout ES cells.

Discussing the results presented in chapter 4, it was proposed to monitor the effects of Nanog on Esrrb expression using imaging techniques that allow the visualisation of foci of active transcription in live cells. Two alternative techniques could be employed to this end. The first of these is based on the detection of gold nanoparticle

tagged probes (See (Anikeeva and Deisseroth, 2012) for a general review) that can be readily introduced in living cells employing common transfection reagents. After excitation with a near infrared light source, isolated nanoparticles emit at a wavelength that is proportional to their aspect ratio. The proximity of two gold particles causes a shift in their emission wavelength that can be used to detect hybridisation of a pair of tagged probes to contiguous sequences on an mRNA molecule (unpublished method developed by Luke Lee at UCSC Berkley). Probes designed to hybridise immediately adjacent to one another on intronic gene regions could be employed to exclusively detect accumulation of nascent transcripts originating from the active alleles of any gene of interest. A second approach is based on the detection of the accumulation of GFP tagged RNA binding proteins, such as MS2 (Janicki et al., 2004) or λ_N–GFP (Daigle and Ellenberg, 2007), in confined regions of the nuclear volume. Targeting of tandem repetitions of the recognition sequences for MS2 or λ_N–GFP to the introns of one or both alleles of a gene of interest will generate ES cell lines in which GFP is specifically recruited to the nascent transcripts accumulating at the sites of active transcription. These cell lines should be engineered to expressed RFP fusion reporters for Nanog protein expression, as discussed in chapter 4.8. Being able to image transcriptional events in live cells could make it possible to address questions otherwise precluded to investigation by conventional FISH techniques. In particular, the list of proposed experiments could be extended by determining: 1) the temporal pattern and the duration of transcriptional firing from *Esrrb* and *Klf4* alleles in individual cells. 2) whether *Esrrb* and *Klf4* are transcribed more frequently or for longer time in cells expressing high Nanog levels. Such experiments could assess whether the detection of active transcription in one cell at a given time is informative to predict the likelihood to observe a second transcriptional event in the same cell at a later time point. This would ultimately clarify whether active alleles that cease transcription remain in a poised state for subsequent rounds of transcriptional firing.

The data presented in Chapter 3 suggest that Nanog influences the kinetics of release of paused RNApolII from the promoter of its target genes, as exemplified by *Esrrb* (**Figure 3.12**). In vivo imaging of transcription could be also fruitfully applied in this

context. It could be possible to monitor the appearance of nascent transcripts using nanoparticle tagged oligos annealing to successive intron-exon junctions across the *Esrrb* gene in microscopy timelapse experiments performed on ES cells released form flavopiridol block. Such experiments would permit direct comparison of the kinetics of RNApolII pause-release in populations of wildtype or Nanog null ES cells.

9.3: Transcriptional consequences of loss of Nanog and Esrrb and commitment to differentiation

Loss of Nanog determines a state in which ES cell are still able to self-renew and regain Nanog expression (Chambers et al., 2007). It is possible to imagine that in the transient absence of Nanog, ES cells are more susceptible to downregulate the expression of critical Nanog target genes and that this downregulation represents a commitment event. Indeed, results presented in chapter 4 suggest that loss of Nanog is required for Esrrb downregulation (see **Figure 4.4**), which in turn marks commitment to differentiate. Quantitative PCR analysis indicates that differentiating Esrrb^{low} cells acquire a gene expression profile reminiscent of EpiSC (see **Figure 4.9**). It will be now necessary to extend this analysis and determine the genome-wide transcriptional changes occurring in cells spontaneously losing expression of either Nanog alone, or both Nanog and Esrrb. The identified set of differentially expressed genes will help characterise the successive steps that accompany loss of pluripotency. It would also be informative to compare the differentially expressed genes identified in cells that spontaneously lose Esrrb or Nanog expression with the list of Esrrb and Nanog targets identified from the analysis of knockout and inducible lines. Such a comparison would clarify the extent to which loss of these two factors has a causal role in triggering differentiation and may help identifying the crucial targets of Esrrb and Nanog in this process. The results presented in chapter 8.3 lead to an initial characterisation of Esrrb responsive genes. It would now be useful to derive knockout lines that possess inducible Esrrb function to characterise Esrrb transcriptional targets to a similar level of detail as that presented for Nanog (see

chapter 3). In particular, transcriptional profiling of Esrrb null ES cells could clarify whether loss of Esrrb is responsible for the downregulation of pre-implantation markers observed in Esrrb-2a-GFPdest1 reporter lines (see **Figure 4.9**) and whether Esrrb exerts a direct function in maintaining naïve pluripotency.

9.4: Estrogen related receptor genes: unique or redundant functions?

The results presented in chapter 8 show that Esrrb is not strictly required for ES cell self-renewal. Although with reduced self-renewal efficiency, and almost completely compromised clonogenic potential, Esrrb null ES cells can be maintained in culture indefinitely and retain expression of Oct4 (see **Figure 8.3**), and Nanog.

As highlighted in chapter 1.4.3, Esrrb, Esrra and Esrrg proteins present a high degree of homology, particularly in their DNA binding domain. (Heard et al., 2000; Hong et al., 1999). Furthermore, estrogen related receptors show pronounced structural and sequence similarities to ERα (Giguere et al., 1988; Heard et al., 2000). As a consequence of the high conservation observed among their DNA binding domains, ERRs bind to a common TCAAGGTCA consensus motifs and share a consistent number of target genes in identical cell types (Chen et al., 2008; Dufour et al., 2007; Vanacker et al., 1999b). In addition, different ERRs can form heterodimers (Dufour et al., 2007) and interact with ERα (Johnston et al., 1997).

In mouse ES cells, Esrrg is absent but Esrra and ERα are expressed, although at reduced levels compared to Esrrb ((Xie et al., 2009) and **Table 3.1**). Mass spectrometry studies have also shown that Esrra is indeed interacting with Esrrb (van den Berg et al., 2010). It is thus possible that Esrra, and less likely ERα, can at least partially compensate for the absence of Esrrb in knockout cells, as already observed for Klf factors (Jiang et al., 2008). Relevant to this point, diethylstilbestrol has been shown to bind to estrogen related receptors (Greschik et al., 2004; Tremblay et al., 2001b) and, similarly to Esrrb knockdown, induce differentiation of ES cells (Percharde et al., 2012). It could be that the effect of diethylstilbestrol is mediated by

the concomitant inhibition of both Esrrb and Esrra activity in the cells. It would be thus interesting to determine whether further deletion in Esrra in Esrrb null ES cells is compatible with the maintenance of a pluripotent state. Conversely, it should also be assessed whether overexpression of Esrra, Esrrg or ERα is able to rescue the self-renewal defects observed in Esrrb null ES cells. Such experiments would directly test whether ERRs play redundant or unique roles in pluripotent cells. In a broader perspective, results generated in ES cells could help understanding whether the overlapping pattern of ERRs expression observed in different tissues (Bookout et al., 2005; Bookout et al., 2006) is functional to ensuring robustness of the transcriptional programmes controlled by this class of nuclear receptors.

References

Ambrosetti, D.C., Basilico, C., and Dailey, L. (1997). Synergistic activation of the fibroblast growth factor 4 enhancer by Sox2 and Oct-3 depends on protein-protein interactions facilitated by a specific spatial arrangement of factor binding sites. Mol Cell Biol *17*, 6321-6329.

Ambrosetti, D.C., Scholer, H.R., Dailey, L., and Basilico, C. (2000). Modulation of the activity of multiple transcriptional activation domains by the DNA binding domains mediates the synergistic action of Sox2 and Oct-3 on the fibroblast growth factor-4 enhancer. J Biol Chem *275*, 23387-23397.

Andrews, P.W. (2002). From teratocarcinomas to embryonic stem cells. Philos Trans R Soc Lond B Biol Sci *357*, 405-417.

Anikeeva, P., and Deisseroth, K. (2012). Photothermal genetic engineering. ACS Nano *6*, 7548-7552.

Aoki, F., Worrad, D.M., and Schultz, R.M. (1997). Regulation of transcriptional activity during the first and second cell cycles in the preimplantation mouse embryo. Dev Biol *181*, 296-307.

Arman, E., Haffner-Krausz, R., Chen, Y., Heath, J.K., and Lonai, P. (1998). Targeted disruption of fibroblast growth factor (FGF) receptor 2 suggests a role for FGF signaling in pregastrulation mammalian development. Proc Natl Acad Sci U S A *95*, 5082-5087.

Avilion, A.A., Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N., and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev *17*, 126-140.

Bachvarova, R. (1985). Gene expression during oogenesis and oocyte development in mammals. Dev Biol (N Y 1985) *1*, 453-524.

Bachvarova, R., and Moy, K. (1985). Autoradiographic studies on the distribution of labeled maternal RNA in early mouse embryos. J Exp Zool *233*, 397-403.

Banaszynski, L.A., Chen, L.C., Maynard-Smith, L.A., Ooi, A.G., and Wandless, T.J. (2006). A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. Cell *126*, 995-1004.

Barry, J.B., Laganiere, J., and Giguere, V. (2006). A single nucleotide in an estrogen-related receptor alpha site can dictate mode of binding and peroxisome proliferator-activated receptor gamma coactivator 1alpha activation of target promoters. Mol Endocrinol *20*, 302-310.

Benezra, R., Davis, R.L., Lockshon, D., Turner, D.L., and Weintraub, H. (1990). The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. Cell *61*, 49-59.

Bhutani, N., Brady, J.J., Damian, M., Sacco, A., Corbel, S.Y., and Blau, H.M. (2010). Reprogramming towards pluripotency requires AID-dependent DNA demethylation. Nature *463*, 1042-1047.

Boeuf, H., Hauss, C., Graeve, F.D., Baran, N., and Kedinger, C. (1997). Leukemia inhibitory factor-dependent transcriptional activation in embryonic stem cells. J Cell Biol *138*, 1207-1217.

Bollag, R.J., Siegfried, Z., Cebra-Thomas, J.A., Garvey, N., Davison, E.M., and Silver, L.M. (1994). An ancient family of embryonically expressed mouse genes sharing a conserved protein motif with the T locus. Nat Genet *7*, 383-389.

Bookout, A.L., Jeong, Y., Downes, M., R., Y., R.M., E., and D.J., a.M. (2005). NURSA Dataset: Tissue-specific expression patterns of nuclear receptors. http://wwwnursaorg/101621/datasets02001.

Bookout, A.L., Jeong, Y., Downes, M., Yu, R.T., Evans, R.M., and Mangelsdorf, D.J. (2006). Anatomical profiling of nuclear receptor expression reveals a hierarchical transcriptional network. Cell *126*, 789-799.

Bowles, J., Schepers, G., and Koopman, P. (2000). Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. Dev Biol *227*, 239-255.

Bradley, A., Evans, M., Kaufman, M.H., and Robertson, E. (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. Nature *309*, 255-256.

Brambrink, T., Foreman, R., Welstead, G.G., Lengner, C.J., Wernig, M., Suh, H., and Jaenisch, R. (2008). Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. Cell Stem Cell *2*, 151-159.

Brinster, R.L. (1974). The effect of cells transferred into the mouse blastocyst on subsequent development. J Exp Med *140*, 1049-1056.

Brons, I.G., Smithers, L.E., Trotter, M.W., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S.M., Howlett, S.K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R.A., *et al.* (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature *448*, 191-195.

Buecker, C., Chen, H.H., Polo, J.M., Daheron, L., Bu, L., Barakat, T.S., Okwieka, P., Porter, A., Gribnau, J., Hochedlinger, K., *et al.* (2010). A murine ESC-like state facilitates transgenesis and homologous recombination in human pluripotent stem cells. Cell Stem Cell *6*, 535-546.

Buganim, Y., Faddah, D.A., Cheng, A.W., Itskovich, E., Markoulaki, S., Ganz, K., Klemm, S.L., van Oudenaarden, A., and Jaenisch, R. (2012). Single-Cell Expression Analyses during Cellular Reprogramming Reveal an Early Stochastic and a Late Hierarchic Phase. Cell *150*, 1209-1222.

Burdon, T., Stracey, C., Chambers, I., Nichols, J., and Smith, A. (1999). Suppression of SHP-2 and ERK signalling promotes self-renewal of mouse embryonic stem cells. Dev Biol *210*, 30-43.

Camp, E., Sanchez-Sanchez, A.V., Garcia-Espana, A., Desalle, R., Odqvist, L., Enrique O'Connor, J., and Mullor, J.L. (2009). Nanog regulates proliferation during early fish development. Stem Cells *27*, 2081-2091.

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*, 643-655.

Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. Nature *450*, 1230-1234.

Chambers, I., and Tomlinson, S.R. (2009). The transcriptional foundation of pluripotency. Development *136*, 2311-2322.

Chao, S.H., Fujinaga, K., Marion, J.E., Taube, R., Sausville, E.A., Senderowicz, A.M., Peterlin, B.M., and Price, D.H. (2000). Flavopiridol inhibits P-TEFb and blocks HIV-1 replication. J Biol Chem *275*, 28345-28348.

Chao, S.H., and Price, D.H. (2001). Flavopiridol inactivates P-TEFb and blocks most RNA polymerase II transcription in vivo. J Biol Chem *276*, 31793-31799.

Chapman, D.L., Garvey, N., Hancock, S., Alexiou, M., Agulnik, S.I., Gibson-Brown, J.J., Cebra-Thomas, J., Bollag, R.J., Silver, L.M., and Papaioannou, V.E. (1996). Expression of the T-box family genes, Tbx1-Tbx5, during early mouse development. Dev Dyn *206*, 379-390.

Chazaud, C., Yamanaka, Y., Pawson, T., and Rossant, J. (2006). Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. Dev Cell *10*, 615-624.

Chen, J., and Nathans, J. (2007). Estrogen-related receptor beta/NR3B2 controls epithelial cell fate and endolymph production by the stria vascularis. Dev Cell *13*, 325-337.

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J., *et al.* (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell *133*, 1106-1117.

Cheng, A.M., Saxton, T.M., Sakai, R., Kulkarni, S., Mbamalu, G., Vogel, W., Tortorice, C.G., Cardiff, R.D., Cross, J.C., Muller, W.J., *et al.* (1998). Mammalian

Grb2 regulates multiple steps in embryonic development and malignant transformation. Cell *95*, 793-803.

Chew, J.L., Loh, Y.H., Zhang, W., Chen, X., Tam, W.L., Yeap, L.S., Li, P., Ang, Y.S., Lim, B., Robson, P., *et al.* (2005). Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. Mol Cell Biol *25*, 6031-6046.

Clerc, R.G., Corcoran, L.M., LeBowitz, J.H., Baltimore, D., and Sharp, P.A. (1988). The B-cell-specific Oct-2 protein contains POU box- and homeo box-type domains. Genes Dev *2*, 1570-1581.

Cole, M.F., Johnstone, S.E., Newman, J.J., Kagey, M.H., and Young, R.A. (2008). Tcf3 is an integral component of the core regulatory circuitry of embryonic stem cells. Genes Dev *22*, 746-755.

Conti, L., Pollard, S.M., Gorba, T., Reitano, E., Toselli, M., Biella, G., Sun, Y., Sanzone, S., Ying, Q.L., Cattaneo, E., *et al.* (2005). Niche-independent symmetrical self-renewal of a mammalian tissue stem cell. PLoS Biol *3*, e283.

Core, L.J., and Lis, J.T. (2008). Transcription regulation through promoter-proximal pausing of RNA polymerase II. Science *319*, 1791-1792.

Core, L.J., Waterfall, J.J., and Lis, J.T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science *322*, 1845-1848.

Coward, P., Lee, D., Hull, M.V., and Lehmann, J.M. (2001). 4-Hydroxytamoxifen binds to and deactivates the estrogen-related receptor gamma. Proc Natl Acad Sci U S A *98*, 8880-8884.

Creusot, F., Acs, G., and Christman, J.K. (1982). Inhibition of DNA methyltransferase and induction of Friend erythroleukemia cell differentiation by 5-azacytidine and 5-aza-2'-deoxycytidine. J Biol Chem *257*, 2041-2048.

Daigle, N., and Ellenberg, J. (2007). LambdaN-GFP: an RNA reporter system for live-cell imaging. Nat Methods *4*, 633-636.

Dance, M., Montagner, A., Salles, J.P., Yart, A., and Raynal, P. (2008). The molecular functions of Shp2 in the Ras/Mitogen-activated protein kinase (ERK1/2) pathway. Cell Signal *20*, 453-459.

Davidson, E.H. (2010). Emerging properties of animal gene regulatory networks. Nature *468*, 911-920.

de Napoles, M., Nesterova, T., and Brockdorff, N. (2007). Early loss of Xist RNA expression and inactive X chromosome associated chromatin modification in developing primordial germ cells. PLoS One *2*, e860.

Denny, P., Swift, S., Brand, N., Dabhade, N., Barton, P., and Ashworth, A. (1992). A conserved family of genes related to the testis determining gene, SRY. Nucleic Acids Res *20*, 2887.

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

Ding, J., Xu, H., Faiola, F., Ma'ayan, A., and Wang, J. (2012). Oct4 links multiple epigenetic pathways to the pluripotency network. Cell Res *22*, 155-167.

Dixon, J.E., Allegrucci, C., Redwood, C., Kump, K., Bian, Y., Chatfield, J., Chen, Y.H., Sottile, V., Voss, S.R., Alberio, R., *et al.* Axolotl Nanog activity in mouse embryonic stem cells demonstrates that ground state pluripotency is conserved from urodele amphibians to mammals. Development *137*, 2973-2980.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376-380.

Doble, B.W., Patel, S., Wood, G.A., Kockeritz, L.K., and Woodgett, J.R. (2007). Functional redundancy of GSK-3alpha and GSK-3beta in Wnt/beta-catenin signaling shown by using an allelic series of embryonic stem cell lines. Dev Cell *12*, 957-971.

Donnelly, M.L., Luke, G., Mehrotra, A., Li, X., Hughes, L.E., Gani, D., and Ryan, M.D. (2001). Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'. J Gen Virol *82*, 1013-1025.

Dufour, C.R., Wilson, B.J., Huss, J.M., Kelly, D.P., Alaynick, W.A., Downes, M., Evans, R.M., Blanchette, M., and Giguere, V. (2007). Genome-wide orchestration of cardiac functions by the orphan nuclear receptors ERRalpha and gamma. Cell Metab *5*, 345-356.

Dunning, M.J., Smith, M.L., Ritchie, M.E., and Tavare, S. (2007). beadarray: R classes and methods for Illumina bead-based data. Bioinformatics *23*, 2183-2184.

Durcova-Hills, G., Ainscough, J., and McLaren, A. (2001). Pluripotential stem cells derived from migrating primordial germ cells. Differentiation *68*, 220-226.

Edelman, L.B., and Fraser, P. (2012). Transcription factories: genetic programming in three dimensions. Curr Opin Genet Dev *22*, 110-114.

Ema, M., Mori, D., Niwa, H., Hasegawa, Y., Yamanaka, Y., Hitoshi, S., Mimura, J., Kawabe, Y., Hosoya, T., Morita, M., *et al.* (2008). Kruppel-like factor 5 is essential for blastocyst development and the normal self-renewal of mouse ESCs. Cell Stem Cell *3*, 555-567.

Erhardt, S., Su, I.H., Schneider, R., Barton, S., Bannister, A.J., Perez-Burgos, L., Jenuwein, T., Kouzarides, T., Tarakhovsky, A., and Surani, M.A. (2003).

Consequences of the depletion of zygotic and embryonic enhancer of zeste 2 during preimplantation mouse development. Development *130*, 4235-4248.

Ernst, M., Oates, A., and Dunn, A.R. (1996). Gp130-mediated signal transduction in embryonic stem cells involves activation of Jak and Ras/mitogen-activated protein kinase pathways. J Biol Chem *271*, 30136-30143.

Eudy, J.D., Yao, S., Weston, M.D., Ma-Edmonds, M., Talmadge, C.B., Cheng, J.J., Kimberling, W.J., and Sumegi, J. (1998). Isolation of a gene encoding a novel member of the nuclear receptor superfamily from the critical region of Usher syndrome type IIa at 1q41. Genomics *50*, 382-384.

Evans, M.J. (1972). The isolation and properties of a clonal tissue culture strain of pluripotent mouse teratoma cells. J Embryol Exp Morphol *28*, 163-176.

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

Falkner, F.G., and Zachau, H.G. (1984). Correct transcription of an immunoglobulin kappa gene requires an upstream fragment containing conserved sequence elements. Nature *310*, 71-74.

Feil, R., Wagner, J., Metzger, D., and Chambon, P. (1997). Regulation of Cre recombinase activity by mutated estrogen receptor ligand-binding domains. Biochem Biophys Res Commun *237*, 752-757.

Feldman, B., Poueymirou, W., Papaioannou, V.E., DeChiara, T.M., and Goldfarb, M. (1995). Requirement of FGF-4 for postimplantation mouse development. Science *267*, 246-249.

Feng, B., Jiang, J., Kraus, P., Ng, J.H., Heng, J.C., Chan, Y.S., Yaw, L.P., Zhang, W., Loh, Y.H., Han, J., *et al.* (2009). Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. Nat Cell Biol *11*, 197-203.

Festuccia, N., and Chambers, I. (2011). Quantification of pluripotency transcription factor levels in embryonic stem cells by flow cytometry. Curr Protoc Stem Cell Biol *Chapter 1*, Unit 1B 9.

Festuccia, N., Osorno, R., Halbritter, F., Karwacki-Neisius, V., Navarro, P., Colby, D., Wong, F., Yates, A., Tomlinson, S.R., and Chambers, I. (2012). Esrrb Is a Direct Nanog Target Gene that Can Substitute for Nanog Function in Pluripotent Cells. Cell Stem Cell *11*, 477-490.

Finley, M.F., Devata, S., and Huettner, J.E. (1999). BMP-4 inhibits neural differentiation of murine embryonic stem cells. J Neurobiol *40*, 271-287.

Flicek, P., Amode, M.R., Barrell, D., Beal, K., Brent, S., Chen, Y., Clapham, P., Coates, G., Fairley, S., Fitzgerald, S., *et al.* (2011). Ensembl 2011. Nucleic Acids Res *39*, D800-806.

Fujita, J., Crane, A.M., Souza, M.K., Dejosez, M., Kyba, M., Flavell, R.A., Thomson, J.A., and Zwaka, T.P. (2008). Caspase activity mediates the differentiation of embryonic stem cells. Cell Stem Cell *2*, 595-601.

Gallardo, T., Shirley, L., John, G.B., and Castrillon, D.H. (2007). Generation of a germ cell-specific mouse transgenic Cre line, Vasa-Cre. Genesis *45*, 413-417.

Gearhart, M.D., Holmbeck, S.M., Evans, R.M., Dyson, H.J., and Wright, P.E. (2003). Monomeric complex of human orphan estrogen related receptor-2 with DNA: a pseudo-dimer interface mediates extended half-site recognition. J Mol Biol *327*, 819-832.

Gearing, D.P., Comeau, M.R., Friend, D.J., Gimpel, S.D., Thut, C.J., McGourty, J., Brasher, K.K., King, J.A., Gillis, S., Mosley, B., *et al.* (1992). The IL-6 signal transducer, gp130: an oncostatin M receptor and affinity converter for the LIF receptor. Science *255*, 1434-1437.

Gearing, D.P., Thut, C.J., VandeBos, T., Gimpel, S.D., Delaney, P.B., King, J., Price, V., Cosman, D., and Beckmann, M.P. (1991). Leukemia inhibitory factor receptor is structurally related to the IL-6 signal transducer, gp130. EMBO J *10*, 2839-2848.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., *et al.* (2004). Bioconductor: open software development for computational biology and bioinformatics. Genome Biol *5*, R80.

Giese, K., Amsterdam, A., and Grosschedl, R. (1991). DNA-binding properties of the HMG domain of the lymphoid-specific transcriptional regulator LEF-1. Genes Dev *5*, 2567-2578.

Giguere, V., Yang, N., Segui, P., and Evans, R.M. (1988). Identification of a new class of steroid hormone receptors. Nature *331*, 91-94.

Gillich, A., Bao, S., Grabole, N., Hayashi, K., Trotter, M.W., Pasque, V., Magnusdottir, E., and Surani, M.A. (2012). Epiblast stem cell-based system reveals reprogramming synergy of germline factors. Cell Stem Cell *10*, 425-439.

Greschik, H., Flaig, R., Renaud, J.P., and Moras, D. (2004). Structural basis for the deactivation of the estrogen-related receptor gamma by diethylstilbestrol or 4-hydroxytamoxifen and determinants of selectivity. J Biol Chem *279*, 33639-33646.

Greschik, H., Wurtz, J.M., Sanglier, S., Bourguet, W., van Dorsselaer, A., Moras, D., and Renaud, J.P. (2002). Structural and functional evidence for ligand-independent transcriptional activation by the estrogen-related receptor 3. Mol Cell *9*, 303-313.

Gubbay, J., Collignon, J., Koopman, P., Capel, B., Economou, A., Munsterberg, A., Vivian, N., Goodfellow, P., and Lovell-Badge, R. (1990). A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. Nature *346*, 245-250.

Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. Cell *130*, 77-88.

Guo, G., Huss, M., Tong, G.Q., Wang, C., Li Sun, L., Clarke, N.D., and Robson, P. (2010). Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. Dev Cell *18*, 675-685.

Guo, G., and Smith, A. (2010). A genome-wide screen in EpiSCs identifies Nr5a nuclear receptors as potent inducers of ground state pluripotency. Development *137*, 3185-3192.

Guo, G., Yang, J., Nichols, J., Hall, J.S., Eyres, I., Mansfield, W., and Smith, A. (2009). Klf4 reverts developmentally programmed restriction of ground state pluripotency. Development *136*, 1063-1069.

Hajkova, P., Ancelin, K., Waldmann, T., Lacoste, N., Lange, U.C., Cesari, F., Lee, C., Almouzni, G., Schneider, R., and Surani, M.A. (2008). Chromatin dynamics during epigenetic reprogramming in the mouse germ line. Nature *452*, 877-881.

Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., Walter, J., and Surani, M.A. (2002). Epigenetic reprogramming in mouse primordial germ cells. Mech Dev *117*, 15-23.

Hajkova, P., Jeffries, S.J., Lee, C., Miller, N., Jackson, S.P., and Surani, M.A. (2010). Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. Science *329*, 78-82.

Halbritter, F., Vaidya, H.J., and Tomlinson, S.R. (2012). GeneProf: analysis of high-throughput sequencing experiments. Nat Methods *9*, 7-8.

Hall, J., Guo, G., Wray, J., Eyres, I., Nichols, J., Grotewold, L., Morfopoulou, S., Humphreys, P., Mansfield, W., Walker, R., *et al.* (2009). Oct4 and LIF/Stat3 additively induce Kruppel factors to sustain embryonic stem cell self-renewal. Cell Stem Cell *5*, 597-609.

Hall, L.L., Byron, M., Butler, J., Becker, K.A., Nelson, A., Amit, M., Itskovitz-Eldor, J., Stein, J., Stein, G., Ware, C., *et al.* (2008). X-inactivation reveals epigenetic anomalies in most hESC but identifies sublines that initiate as expected. J Cell Physiol *216*, 445-452.

Hamatani, T., Carter, M.G., Sharov, A.A., and Ko, M.S. (2004). Dynamics of global gene expression changes during mouse preimplantation development. Dev Cell *6*, 117-131.

Han, D.W., Do, J.T., Gentile, L., Stehling, M., Lee, H.T., and Scholer, H.R. (2008). Pluripotential reprogramming of the somatic genome in hybrid cells occurs with the first cell cycle. Stem Cells *26*, 445-454.

Hanna, J., Cheng, A.W., Saha, K., Kim, J., Lengner, C.J., Soldner, F., Cassady, J.P., Muffat, J., Carey, B.W., and Jaenisch, R. (2010). Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. Proc Natl Acad Sci U S A *107*, 9222-9227.

Hanna, J., Markoulaki, S., Mitalipova, M., Cheng, A.W., Cassady, J.P., Staerk, J., Carey, B.W., Lengner, C.J., Foreman, R., Love, J., *et al.* (2009). Metastable pluripotent states in NOD-mouse-derived ESCs. Cell Stem Cell *4*, 513-524.

Hanna, J., Markoulaki, S., Schorderet, P., Carey, B.W., Beard, C., Wernig, M., Creyghton, M.P., Steine, E.J., Cassady, J.P., Foreman, R., *et al.* (2008). Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency. Cell *133*, 250-264.

Harley, V.R., Jackson, D.I., Hextall, P.J., Hawkins, J.R., Berkovitz, G.D., Sockanathan, S., Lovell-Badge, R., and Goodfellow, P.N. (1992). DNA binding activity of recombinant SRY from normal males and XY females. Science *255*, 453-456.

Hart, A.H., Hartley, L., Ibrahim, M., and Robb, L. (2004). Identification, cloning and expression analysis of the pluripotency promoting Nanog genes in mouse and human. Dev Dyn *230*, 187-198.

Haruki, H., Nishikawa, J., and Laemmli, U.K. (2008). The anchor-away technique: rapid, conditional establishment of yeast mutant phenotypes. Mol Cell *31*, 925-932.

Hayashi, K., Lopes, S.M., Tang, F., and Surani, M.A. (2008). Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. Cell Stem Cell *3*, 391-401.

Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S., and Saitou, M. (2011). Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. Cell *146*, 519-532.

Heard, D.J., Norby, P.L., Holloway, J., and Vissing, H. (2000). Human ERRgamma, a third member of the estrogen receptor-related receptor (ERR) subfamily of orphan nuclear receptors: tissue-specific isoforms are expressed during development and in the adult. Mol Endocrinol *14*, 382-392.

Hollenberg, S.M., and Evans, R.M. (1988). Multiple and cooperative trans-activation domains of the human glucocorticoid receptor. Cell *55*, 899-906.

Hong, H., Yang, L., and Stallcup, M.R. (1999). Hormone-independent transcriptional activation and coactivator binding by novel orphan nuclear receptor ERR3. J Biol Chem *274*, 22618-22626.

Huang, Y., Osorno, R., Tsakiridis, A., and Wilson, V. (2012). In vivo differentiation potential of epiblast stem cells revealed by chimeric embryo formation. Cell reports *In press*.

Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A.E., and Melton, D.A. (2008). Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. Nat Biotechnol *26*, 795-797.

Hui, C.C., and Angers, S. (2011). Gli proteins in development and disease. Annu Rev Cell Dev Biol *27*, 513-537.

Iborra, F.J., Pombo, A., Jackson, D.A., and Cook, P.R. (1996). Active RNA polymerases are localized within discrete transcription "factories' in human nuclei. J Cell Sci *109 (Pt 6)*, 1427-1436.

Ito, S., D'Alessio, A.C., Taranova, O.V., Hong, K., Sowers, L.C., and Zhang, Y. (2010). Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. Nature *466*, 1129-1133.

Ivanova, N., Dobrin, R., Lu, R., Kotenko, I., Levorse, J., DeCoste, C., Schafer, X., Lun, Y., and Lemischka, I.R. (2006). Dissecting self-renewal in stem cells with RNA interference. Nature *442*, 533-538.

Jaenisch, R., and Young, R. (2008). Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. Cell *132*, 567-582.

Janicki, S.M., Tsukamoto, T., Salghetti, S.E., Tansey, W.P., Sachidanandam, R., Prasanth, K.V., Ried, T., Shav-Tal, Y., Bertrand, E., Singer, R.H., *et al.* (2004). From silencing to gene expression: real-time analysis in single cells. Cell *116*, 683-698.

Jauch, R., Ng, C.K., Saikatendu, K.S., Stevens, R.C., and Kolatkar, P.R. (2008). Crystal structure and DNA binding of the homeodomain of the stem cell transcription factor Nanog. J Mol Biol *376*, 758-770.

Jiang, J., Chan, Y.S., Loh, Y.H., Cai, J., Tong, G.Q., Lim, C.A., Robson, P., Zhong, S., and Ng, H.H. (2008). A core Klf circuitry regulates self-renewal of embryonic stem cells. Nat Cell Biol *10*, 353-360.

Johansson, B.M., and Wiles, M.V. (1995). Evidence for involvement of activin A and bone morphogenetic protein 4 in mammalian mesoderm and hematopoietic development. Mol Cell Biol *15*, 141-151.

Johnson, M.H., and Ziomek, C.A. (1981). The foundation of two distinct cell lineages within the mouse morula. Cell *24*, 71-80.

Johnston, S.D., Liu, X., Zuo, F., Eisenbraun, T.L., Wiley, S.R., Kraus, R.J., and Mertz, J.E. (1997). Estrogen-related receptor alpha 1 functionally binds as a monomer to extended half-site sequences including ones contained within estrogen-response elements. Mol Endocrinol *11*, 342-352.

Kahan, B.W., and Ephrussi, B. (1970). Developmental potentialities of clonal in vitro cultures of mouse testicular teratoma. J Natl Cancer Inst *44*, 1015-1036.

Kehler, J., Tolkunova, E., Koschorz, B., Pesce, M., Gentile, L., Boiani, M., Lomeli, H., Nagy, A., McLaughlin, K.J., Scholer, H.R., *et al.* (2004). Oct4 is required for primordial germ cell survival. EMBO Rep *5*, 1078-1083.

Khalil, A.S., Lu, T.K., Bashor, C.J., Ramirez, C.L., Pyenson, N.C., Joung, J.K., and Collins, J.J. (2012). A synthetic biology framework for programming eukaryotic transcription functions. Cell *150*, 647-658.

Kim, J., Chu, J., Shen, X., Wang, J., and Orkin, S.H. (2008). An extended transcriptional network for pluripotency of embryonic stem cells. Cell *132*, 1049-1061.

Kishimoto, T., Taga, T., and Akira, S. (1994). Cytokine signal transduction. Cell *76*, 253-262.

Kleinsmith, L.J., and Pierce, G.B., Jr. (1964). Multipotentiality of Single Embryonal Carcinoma Cells. Cancer Res *24*, 1544-1551.

Klemm, J.D., and Pabo, C.O. (1996). Oct-1 POU domain-DNA interactions: cooperative binding of isolated subdomains and effects of covalent linkage. Genes Dev *10*, 27-36.

Klemm, J.D., Rould, M.A., Aurora, R., Herr, W., and Pabo, C.O. (1994). Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules. Cell *77*, 21-32.

Klimanskaya, I., Chung, Y., Meisner, L., Johnson, J., West, M.D., and Lanza, R. (2005). Human embryonic stem cells derived without feeder cells. Lancet *365*, 1636-1641.

Koutsourakis, M., Langeveld, A., Patient, R., Beddington, R., and Grosveld, F. (1999). The transcription factor GATA6 is essential for early extraembryonic development. Development *126*, 723-732.

Kumar, V., Green, S., Stack, G., Berry, M., Jin, J.R., and Chambon, P. (1987). Functional domains of the human estrogen receptor. Cell *51*, 941-951.

Kunath, T., Saba-El-Leil, M.K., Almousailleakh, M., Wray, J., Meloche, S., and Smith, A. (2007). FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. Development *134*, 2895-2902.

Kurimoto, K., Yabuta, Y., Ohinata, Y., Shigeta, M., Yamanaka, K., and Saitou, M. (2008). Complex genome-wide transcription dynamics orchestrated by Blimp1 for the specification of the germ cell lineage in mice. Genes Dev *22*, 1617-1635.

Kuroda, T., Tada, M., Kubota, H., Kimura, H., Hatano, S.Y., Suemori, H., Nakatsuji, N., and Tada, T. (2005). Octamer and Sox elements are required for transcriptional cis regulation of Nanog gene expression. Mol Cell Biol *25*, 2475-2485.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.

Laudet, V., Stehelin, D., and Clevers, H. (1993). Ancestry and diversity of the HMG box superfamily. Nucleic Acids Res *21*, 2493-2501.

Lavial, F., Acloque, H., Bertocchini, F., Macleod, D.J., Boast, S., Bachelard, E., Montillet, G., Thenot, S., Sang, H.M., Stern, C.D., *et al.* (2007). The Oct4 homologue PouV and Nanog regulate pluripotency in chicken embryonic stem cells. Development *134*, 3549-3563.

Leitch, H.G., Blair, K., Mansfield, W., Ayetey, H., Humphreys, P., Nichols, J., Surani, M.A., and Smith, A. (2010). Embryonic germ cells from mice and rats exhibit properties consistent with a generic pluripotent ground state. Development *137*, 2279-2287.

Lengner, C.J., Gimelbrant, A.A., Erwin, J.A., Cheng, A.W., Guenther, M.G., Welstead, G.G., Alagappan, R., Frampton, G.M., Xu, P., Muffat, J., *et al.* (2010). Derivation of pre-X inactivation human embryonic stem cells under physiological oxygen concentrations. Cell *141*, 872-883.

Lewandoski, M., Wassarman, K.M., and Martin, G.R. (1997). Zp3-cre, a transgenic mouse line for the activation or inactivation of loxP-flanked target genes specifically in the female germ line. Curr Biol *7*, 148-151.

Li, G., Ruan, X., Auerbach, R.K., Sandhu, K.S., Zheng, M., Wang, P., Poh, H.M., Goh, Y., Lim, J., Zhang, J., *et al.* (2012). Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. Cell *148*, 84-98.

Li, M., Sendtner, M., and Smith, A. (1995). Essential function of LIF receptor in motor neurons. Nature *378*, 724-727.

Li, R., Liang, J., Ni, S., Zhou, T., Qing, X., Li, H., He, W., Chen, J., Li, F., Zhuang, Q., *et al.* (2010). A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. Cell Stem Cell *7*, 51-63.

Li, W., Wei, W., Zhu, S., Zhu, J., Shi, Y., Lin, T., Hao, E., Hayek, A., Deng, H., and Ding, S. (2009). Generation of rat and human induced pluripotent stem cells by combining genetic reprogramming and chemical inhibitors. Cell Stem Cell *4*, 16-19.

Li, X., Zhao, X., Fang, Y., Jiang, X., Duong, T., Fan, C., Huang, C.C., and Kain, S.R. (1998). Generation of destabilized green fluorescent protein as a transcription reporter. J Biol Chem *273*, 34970-34975.

Liang, J., Wan, M., Zhang, Y., Gu, P., Xin, H., Jung, S.Y., Qin, J., Wong, J., Cooney, A.J., Liu, D., *et al.* (2008). Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. Nat Cell Biol *10*, 731-739.

Lin, C.Y., Loven, J., Rahl, P.B., Paranal, R.M., Burge, C.B., Bradner, J.E., Lee, T.I., and Young, R.A. (2012). Transcriptional Amplification in Tumor Cells with Elevated c-Myc. Cell *151*, 56-67.

Loh, Y.H., Wu, Q., Chew, J.L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., *et al.* (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat Genet *38*, 431-440.

Lu, R., Markowetz, F., Unwin, R.D., Leek, J.T., Airoldi, E.M., MacArthur, B.D., Lachmann, A., Rozov, R., Ma'ayan, A., Boyer, L.A., *et al.* (2009). Systems-level dynamic analyses of fate change in murine embryonic stem cells. Nature *462*, 358-362.

Luo, J., Sladek, R., Bader, J.A., Matthyssen, A., Rossant, J., and Giguere, V. (1997). Placental abnormalities in mouse embryos lacking the orphan nuclear receptor ERR-beta. Nature *388*, 778-782.

Lyon, M.F. (1961). Gene action in the X-chromosome of the mouse (Mus musculus L.). Nature *190*, 372-373.

Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, R., Tchieu, J., Jaenisch, R., *et al.* (2007). Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. Cell Stem Cell *1*, 55-70.

Mak, W., Nesterova, T.B., de Napoles, M., Appanah, R., Yamanaka, S., Otte, A.P., and Brockdorff, N. (2004). Reactivation of the paternal X chromosome in early mouse embryos. Science *303*, 666-669.

Margaritis, T., and Holstege, F.C. (2008). Poised RNA polymerase II gives pause for thought. Cell *133*, 581-584.

Marson, A., Levine, S.S., Cole, M.F., Frampton, G.M., Brambrink, T., Johnstone, S., Guenther, M.G., Johnston, W.K., Wernig, M., Newman, J., *et al.* (2008). Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell *134*, 521-533.

Martello, G., Sugimoto, T., Diamanti, E., Joshi, A., Hannah, R., Ohtsuka, S., Gottgens, B., Niwa, H., and Smith, A. (2012). Esrrb is a pivotal target of the gsk3/tcf3 axis regulating embryonic stem cell self-renewal. Cell Stem Cell *11*, 491-504.

Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci U S A *78*, 7634-7638.

Martin, G.R., and Evans, M.J. (1974). The morphology and growth of a pluripotent teratocarcinoma cell line and its derivatives in tissue culture. Cell *2*, 163-172.

Mason, J.O., Williams, G.T., and Neuberger, M.S. (1985). Transcription cell type specificity is conferred by an immunoglobulin VH gene promoter that includes a functional consensus sequence. Cell *41*, 479-487.

Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A.A., *et al.* (2007). Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. Nat Cell Biol *9*, 625-635.

Matsuda, T., Nakamura, T., Nakao, K., Arai, T., Katsuki, M., Heike, T., and Yokota, T. (1999). STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. EMBO J *18*, 4261-4269.

Matsui, Y., Zsebo, K., and Hogan, B.L. (1992). Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. Cell *70*, 841-847.

Mayer, W., Niveleau, A., Walter, J., Fundele, R., and Haaf, T. (2000). Demethylation of the zygotic paternal genome. Nature *403*, 501-502.

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*, 49-55.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P., *et al.* (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553-560.

Mise, N., Fuchikami, T., Sugimoto, M., Kobayakawa, S., Ike, F., Ogawa, T., Tada, T., Kanaya, S., Noce, T., and Abe, K. (2008). Differences and similarities in the developmental status of embryo-derived stem cells and primordial germ cells revealed by global expression profiling. Genes Cells *13*, 863-877.

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*, 631-642.

Mitsunaga, K., Araki, K., Mizusaki, H., Morohashi, K., Haruna, K., Nakagata, N., Giguere, V., Yamamura, K., and Abe, K. (2004). Loss of PGC-specific expression of the orphan nuclear receptor ERR-beta results in reduction of germ cell number in mouse embryos. Mech Dev *121*, 237-246.

Miyanari, Y., and Torres-Padilla, M.E. (2012). Control of ground-state pluripotency by allelic regulation of Nanog. Nature *483*, 470-473.

Moretto-Zita, M., Jin, H., Shen, Z., Zhao, T., Briggs, S.P., and Xu, Y. (2010). Phosphorylation stabilizes Nanog by promoting its interaction with Pin1. Proc Natl Acad Sci U S A *107*, 13312-13317.

Morrisey, E.E., Tang, Z., Sigrist, K., Lu, M.M., Jiang, F., Ip, H.S., and Parmacek, M.S. (1998). GATA6 regulates HNF4 and is required for differentiation of visceral endoderm in the mouse embryo. Genes Dev *12*, 3579-3590.

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

Nagy, A., Rossant, J., Nagy, R., Abramow-Newerly, W., and Roder, J.C. (1993). Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. Proc Natl Acad Sci U S A *90*, 8424-8428.

Nakamura, T., Arai, Y., Umehara, H., Masuhara, M., Kimura, T., Taniguchi, H., Sekimoto, T., Ikawa, M., Yoneda, Y., Okabe, M., *et al.* (2007). PGC7/Stella protects against DNA demethylation in early embryogenesis. Nat Cell Biol *9*, 64-71.

Nakamura, T., Liu, Y.J., Nakashima, H., Umehara, H., Inoue, K., Matoba, S., Tachibana, M., Ogura, A., Shinkai, Y., and Nakano, T. (2012). PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. Nature *486*, 415-419.

Navarro, P., Chambers, I., Karwacki-Neisius, V., Chureau, C., Morey, C., Rougeulle, C., and Avner, P. (2008). Molecular coupling of Xist regulation and pluripotency. Science *321*, 1693-1695.

Navarro, P., Festuccia, N., Colby, D., Gagliardi, A., Mullin, N., Zhang, W., Karwacki-Neisius, V., Osorno, R., Kelly, D., Robertson, M. and Chambers, I. (2012). OCT4/SOX2-independent Nanog autorepression modulates heterogeneous Nanog gene expression in mouse ES cells. EMBO J *in press*.

Nichols, J., Chambers, I., and Smith, A. (1994). Derivation of germline competent embryonic stem cells with a combination of interleukin-6 and soluble interleukin-6 receptor. Exp Cell Res *215*, 237-239.

Nichols, J., Chambers, I., Taga, T., and Smith, A. (2001). Physiological rationale for responsiveness of mouse embryonic stem cells to gp130 cytokines. Development *128*, 2333-2339.

Nichols, J., Davidson, D., Taga, T., Yoshida, K., Chambers, I., and Smith, A. (1996). Complementary tissue-specific expression of LIF and LIF-receptor mRNAs in early mouse embryogenesis. Mech Dev *57*, 123-131.

Nichols, J., Silva, J., Roode, M., and Smith, A. (2009). Suppression of Erk signalling promotes ground state pluripotency in the mouse embryo. Development *136*, 3215-3222.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell *95*, 379-391.

Nie, Z., Hu, G., Wei, G., Cui, K., Yamane, A., Resch, W., Wang, R., Green, D.R., Tessarollo, L., Casellas, R., *et al.* c-Myc Is a Universal Amplifier of Expressed Genes in Lymphocytes and Embryonic Stem Cells. Cell *151*, 68-79.

Nishimoto, M., Fukushima, A., Okuda, A., and Muramatsu, M. (1999). The gene for the embryonic stem cell coactivator UTF1 carries a regulatory element which selectively interacts with a complex composed of Oct-3/4 and Sox-2. Mol Cell Biol *19*, 5453-5465.

Nishioka, N., Inoue, K., Adachi, K., Kiyonari, H., Ota, M., Ralston, A., Yabuta, N., Hirahara, S., Stephenson, R.O., Ogonuki, N., *et al.* (2009). The Hippo signaling pathway components Lats and Yap pattern Tead4 activity to distinguish mouse trophectoderm from inner cell mass. Dev Cell *16*, 398-410.

Nishioka, N., Yamamoto, S., Kiyonari, H., Sato, H., Sawada, A., Ota, M., Nakao, K., and Sasaki, H. (2008). Tead4 is required for specification of trophectoderm in pre-implantation mouse embryos. Mech Dev *125*, 270-283.

Niwa, H., Burdon, T., Chambers, I., and Smith, A. (1998). Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. Genes Dev *12*, 2048-2060.

Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet *24*, 372-376.

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*, 118-122.

Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R., and Rossant, J. (2005). Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. Cell *123*, 917-929.

O'Shea, J.J., Gadina, M., and Schreiber, R.D. (2002). Cytokine signaling in 2002: new surprises in the Jak/Stat pathway. Cell *109 Suppl*, S121-131.

Ogawa, K., Nishinakamura, R., Iwamatsu, Y., Shimosato, D., and Niwa, H. (2006). Synergistic action of Wnt and LIF in maintaining pluripotency of mouse ES cells. Biochem Biophys Res Commun *343*, 159-166.

Oh, J.H., Do, H.J., Yang, H.M., Moon, S.Y., Cha, K.Y., Chung, H.M., and Kim, J.H. (2005). Identification of a putative transactivation domain in human Nanog. Exp Mol Med *37*, 250-254.

Ohinata, Y., Ohta, H., Shigeta, M., Yamanaka, K., Wakayama, T., and Saitou, M. (2009). A signaling principle for the specification of the germ cell lineage in mice. Cell *137*, 571-584.

Ohinata, Y., Payer, B., O'Carroll, D., Ancelin, K., Ono, Y., Sano, M., Barton, S.C., Obukhanych, T., Nussenzweig, M., Tarakhovsky, A., *et al.* (2005). Blimp1 is a critical determinant of the germ cell lineage in mice. Nature *436*, 207-213.

Okamoto, I., Otte, A.P., Allis, C.D., Reinberg, D., and Heard, E. (2004). Epigenetic dynamics of imprinted X inactivation during early mouse development. Science *303*, 644-649.

Okumura-Nakanishi, S., Saito, M., Niwa, H., and Ishikawa, F. (2005). Oct-3/4 and Sox2 regulate Oct-3/4 gene in embryonic stem cells. J Biol Chem *280*, 5307-5317.

Oliveri, P., Tu, Q., and Davidson, E.H. (2008). Global regulatory logic for specification of an embryonic cell lineage. Proc Natl Acad Sci U S A *105*, 5955-5962.

Osorno, R., and Chambers, I. (2011). Transcription factor heterogeneity and epiblast pluripotency. Philos Trans R Soc Lond B Biol Sci *366*, 2230-2237.

Osorno, R., Tsakiridis, A., Wong, F., Cambray, N., Economou, C., Wilkie, R., Blin, G., Scotting, P.J., Chambers, I., and Wilson, V. (2012). The developmental dismantling of pluripotency is reversed by ectopic Oct4 expression. Development *139*, 2288-2298.

Pan, G., and Pei, D. (2005). The stem cell pluripotency factor NANOG activates transcription with two unusually potent subdomains at its C terminus. J Biol Chem *280*, 1401-1407.

Pan, G.J., and Pei, D.Q. (2003). Identification of two distinct transactivation domains in the pluripotency sustaining factor nanog. Cell Res *13*, 499-502.

Pardo, M., Lang, B., Yu, L., Prosser, H., Bradley, A., Babu, M.M., and Choudhary, J. (2010). An expanded Oct4 interaction network: implications for stem cell biology, development, and disease. Cell Stem Cell *6*, 382-395.

Parslow, T.G., Blair, D.L., Murphy, W.J., and Granner, D.K. (1984). Structure of the 5' ends of immunoglobulin genes: a novel conserved sequence. Proc Natl Acad Sci U S A *81*, 2650-2654.

Paynton, B.V., Rempel, R., and Bachvarova, R. (1988). Changes in state of adenylation and time course of degradation of maternal mRNAs during oocyte maturation and early embryonic development in the mouse. Dev Biol *129*, 304-314.

Pedersen, R.A., Wu, K., and Balakier, H. (1986). Origin of the inner cell mass in mouse embryos: cell lineage analysis by microinjection. Dev Biol *117*, 581-595.

Pelton, T.A., Sharma, S., Schulz, T.C., Rathjen, J., and Rathjen, P.D. (2002). Transient pluripotent cell populations during primitive ectoderm formation: correlation of in vivo and in vitro pluripotent cell development. J Cell Sci *115*, 329-339.

Percharde, M., Lavial, F., Ng, J.H., Kumar, V., Tomaz, R.A., Martin, N., Yeo, J.C., Gil, J., Prabhakar, S., Ng, H.H., *et al.* (2012). Ncoa3 functions as an essential Esrrb coactivator to sustain embryonic stem cell self-renewal and reprogramming. Genes Dev.

Pettersson, K., Svensson, K., Mattsson, R., Carlsson, B., Ohlsson, R., and Berkenstam, A. (1996). Expression of a novel member of estrogen response element-binding nuclear receptors is restricted to the early stages of chorion formation during mouse embryogenesis. Mech Dev *54*, 211-223.

Pierce, G.B., and Dixon, F.J., Jr. (1959). Testicular teratomas. I. Demonstration of teratogenesis by metamorphosis of multipotential cells. Cancer *12*, 573-583.

Pierce, G.B., Jr., Dixon, F.J., Jr., and Verney, E.L. (1960). Teratocarcinogenic and tissue-forming potentials of the cell types comprising neoplastic embryoid bodies. Lab Invest *9*, 583-602.

Pierce, G.B., Jr., and Verney, E.L. (1961). An in vitro and in vivo study of differentiation in teratocarcinomas. Cancer *14*, 1017-1029.

Piko, L., and Clegg, K.B. (1982). Quantitative changes in total RNA, total poly(A), and ribosomes in early mouse embryos. Dev Biol *89*, 362-378.

Plusa, B., Piliszek, A., Frankenberg, S., Artus, J., and Hadjantonakis, A.K. (2008). Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. Development *135*, 3081-3091.

Po, A., Ferretti, E., Miele, E., De Smaele, E., Paganelli, A., Canettieri, G., Coni, S., Di Marcotullio, L., Biffoni, M., Massimi, L., *et al.* (2010). Hedgehog controls neural stem cells through p53-independent regulation of Nanog. EMBO J *29*, 2646-2658.

Pollard, S.M., Conti, L., Sun, Y., Goffredo, D., and Smith, A. (2006). Adherent neural stem (NS) cells from fetal and adult forebrain. Cereb Cortex *16 Suppl 1*, i112-120.

Pomp, O., Dreesen, O., Leong, D.F., Meller-Pomp, O., Tan, T.T., Zhou, F., and Colman, A. (2011). Unexpected X chromosome skewing during culture and reprogramming of human somatic cells can be alleviated by exogenous telomerase. Cell Stem Cell *9*, 156-165.

Ptashne, M., and Gann, A. (2001). Transcription initiation: imposing specificity by localization. Essays Biochem *37*, 1-15.

Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A., and Young, R.A. (2010). c-Myc regulates transcriptional pause release. Cell *141*, 432-445.

Rastan, S. (1982). Timing of X-chromosome inactivation in postimplantation mouse embryos. J Embryol Exp Morphol *71*, 11-24.

Raymond, C.S., and Soriano, P. (2007). High-efficiency FLP and PhiC31 site-specific recombination in mammalian cells. PLoS One *2*, e162.

Remenyi, A., Lins, K., Nissen, L.J., Reinbold, R., Scholer, H.R., and Wilmanns, M. (2003). Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of Oct4 and Sox2 on two enhancers. Genes Dev *17*, 2048-2059.

Resnick, J.L., Bixler, L.S., Cheng, L., and Donovan, P.J. (1992). Long-term proliferation of mouse primordial germ cells in culture. Nature *359*, 550-551.

Rodda, D.J., Chew, J.L., Lim, L.H., Loh, Y.H., Wang, B., Ng, H.H., and Robson, P. (2005). Transcriptional regulation of nanog by OCT4 and SOX2. J Biol Chem *280*, 24731-24737.

Rosenthal, M.D., Wishnow, R.M., and Sato, G.H. (1970). In vitro growth and differetiation of clonal populations of multipotential mouse clls derived from a transplantable testicular teratocarcinoma. J Natl Cancer Inst *44*, 1001-1014.

Rosner, M.H., Vigano, M.A., Ozato, K., Timmons, P.M., Poirier, F., Rigby, P.W., and Staudt, L.M. (1990). A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. Nature *345*, 686-692.

Samavarchi-Tehrani, P., Golipour, A., David, L., Sung, H.K., Beyer, T.A., Datti, A., Woltjen, K., Nagy, A., and Wrana, J.L. (2010). Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. Cell Stem Cell *7*, 64-77.

Santos, F., Hendrich, B., Reik, W., and Dean, W. (2002). Dynamic reprogramming of DNA methylation in the early mouse embryo. Dev Biol *241*, 172-182.

Santos, F., Zakhartchenko, V., Stojkovic, M., Peters, A., Jenuwein, T., Wolf, E., Reik, W., and Dean, W. (2003). Epigenetic marking correlates with developmental potential in cloned bovine preimplantation embryos. Curr Biol *13*, 1116-1121.

Sato, M., Kimura, T., Kurokawa, K., Fujita, Y., Abe, K., Masuhara, M., Yasunaga, T., Ryo, A., Yamamoto, M., and Nakano, T. (2002). Identification of PGC7, a new gene expressed specifically in preimplantation embryos and germ cells. Mech Dev *113*, 91-94.

Sato, N., Meijer, L., Skaltsounis, L., Greengard, P., and Brivanlou, A.H. (2004). Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. Nat Med *10*, 55-63.

Scholer, H.R., Dressler, G.R., Balling, R., Rohdewohld, H., and Gruss, P. (1990). Oct-4: a germline-specific transcription factor mapping to the mouse t-complex. EMBO J *9*, 2185-2195.

Schultz, R.M. (2002). The molecular foundations of the maternal to zygotic transition in the preimplantation embryo. Hum Reprod Update *8*, 323-331.

Seki, Y., Hayashi, K., Itoh, K., Mizugaki, M., Saitou, M., and Matsui, Y. (2005). Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice. Dev Biol *278*, 440-458.

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

Sekkai, D., Gruel, G., Herry, M., Moucadel, V., Constantinescu, S.N., Albagli, O., Tronik-Le Roux, D., Vainchenker, W., and Bennaceur-Griscelli, A. (2005). Microarray analysis of LIF/Stat3 transcriptional targets in embryonic stem cells. Stem Cells *23*, 1634-1642.

Sharova, L.V., Sharov, A.A., Piao, Y., Shaik, N., Sullivan, T., Stewart, C.L., Hogan, B.L., and Ko, M.S. (2007). Global gene expression profiling reveals similarities and differences among mouse pluripotent stem cells of different origins and strains. Dev Biol *307*, 446-459.

Shen, Y., Matsuno, Y., Fouse, S.D., Rao, N., Root, S., Xu, R., Pellegrini, M., Riggs, A.D., and Fan, G. (2008). X-inactivation in female human embryonic stem cells is in a nonrandom pattern and prone to epigenetic alterations. Proc Natl Acad Sci U S A *105*, 4709-4714.

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

Silva, J., Barrandon, O., Nichols, J., Kawaguchi, J., Theunissen, T.W., and Smith, A. (2008a). Promotion of reprogramming to ground state pluripotency by signal inhibition. PLoS Biol *6*, e253.

Silva, J., Chambers, I., Pollard, S., and Smith, A. (2006). Nanog promotes transfer of pluripotency after cell fusion. Nature *441*, 997-1001.

Silva, J., Nichols, J., Theunissen, T.W., Guo, G., van Oosten, A.L., Barrandon, O., Wray, J., Yamanaka, S., Chambers, I., and Smith, A. (2009). Nanog is the gateway to the pluripotent ground state. Cell *138*, 722-737.

Silva, S.S., Rowntree, R.K., Mekhoubad, S., and Lee, J.T. (2008b). X-chromosome inactivation and epigenetic fluidity in human embryonic stem cells. Proc Natl Acad Sci U S A *105*, 4820-4825.

Singh, A.M., Hamazaki, T., Hankowski, K.E., and Terada, N. (2007). A heterogeneous expression pattern for Nanog in embryonic stem cells. Stem Cells *25*, 2534-2542.

Singh, J., and Padgett, R.A. (2009). Rates of in situ transcription and splicing in large human genes. Nat Struct Mol Biol *16*, 1128-1133.

Sladek, R., Bader, J.A., and Giguere, V. (1997). The orphan nuclear receptor estrogen-related receptor alpha is a transcriptional regulator of the human medium-chain acyl coenzyme A dehydrogenase gene. Mol Cell Biol *17*, 5400-5409.

Smith, A.G., Heath, J.K., Donaldson, D.D., Wong, G.G., Moreau, J., Stahl, M., and Rogers, D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. Nature *336*, 688-690.

Smith, A.G., and Hooper, M.L. (1987). Buffalo rat liver cells produce a diffusible activity which inhibits the differentiation of murine embryonal carcinoma and embryonic stem cells. Dev Biol *121*, 1-9.

Smyth, G.K. (2005). Limma: linear models for microarray data. In Bioinformatics and Computational Biology Solutions using R and Bioconductor, R Gentleman, V Carey, S Dudoit, R Irizarry, W Huber, eds (New York, USA: Springer Verlag), 397-420.

Solter, D. (2006). From teratocarcinomas to embryonic stem cells and beyond: a history of embryonic stem cell research. Nat Rev Genet *7*, 319-327.

Solter, D., Skreb, N., and Damjanov, I. (1970). Extrauterine growth of mouse egg-cylinders results in malignant teratoma. Nature *227*, 503-504.

Soudais, C., Bielinska, M., Heikinheimo, M., MacArthur, C.A., Narita, N., Saffitz, J.E., Simon, M.C., Leiden, J.M., and Wilson, D.B. (1995). Targeted mutagenesis of the transcription factor GATA-4 gene in mouse embryonic stem cells disrupts visceral endoderm differentiation in vitro. Development *121*, 3877-3888.

Sridharan, R., Tchieu, J., Mason, M.J., Yachechko, R., Kuoy, E., Horvath, S., Zhou, Q., and Plath, K. (2009). Role of the murine reprogramming factors in the induction of pluripotency. Cell *136*, 364-377.

Stadtfeld, M., Maherali, N., Breault, D.T., and Hochedlinger, K. (2008). Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. Cell Stem Cell *2*, 230-240.

Stavridis, M.P., Lunn, J.S., Collins, B.J., and Storey, K.G. (2007). A discrete period of FGF-induced Erk1/2 signalling is required for vertebrate neural specification. Development *134*, 2889-2894.

Stevens, L.C. (1962). Testicular teratomas in fetal mice. J Natl Cancer Inst *28*, 247-267.

Stevens, L.C. (1968). The development of teratomas from intratesticular grafts of tubal mouse eggs. J Embryol Exp Morphol *20*, 329-341.

Stevens, L.C., and Little, C.C. (1954). Spontaneous Testicular Teratomas in an Inbred Strain of Mice. Proc Natl Acad Sci U S A *40*, 1080-1087.

Stewart, C.L., Kaspar, P., Brunet, L.J., Bhatt, H., Gadi, I., Kontgen, F., and Abbondanzo, S.J. (1992). Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor. Nature *359*, 76-79.

Strumpf, D., Mao, C.A., Yamanaka, Y., Ralston, A., Chawengsaksophak, K., Beck, F., and Rossant, J. (2005). Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. Development *132*, 2093-2102.

Sturm, R.A., Das, G., and Herr, W. (1988). The ubiquitous octamer-binding protein Oct-1 contains a POU domain with a homeo box subdomain. Genes Dev *2*, 1582-1599.

Sturm, R.A., and Herr, W. (1988). The POU domain is a bipartite DNA-binding structure. Nature *336*, 601-604.

Sugimoto, A., Iino, Y., Maeda, T., Watanabe, Y., and Yamamoto, M. (1991). Schizosaccharomyces pombe ste11+ encodes a transcription factor with an HMG motif that is a critical regulator of sexual development. Genes Dev *5*, 1990-1999.

Sugimoto, M., and Abe, K. (2007). X chromosome reactivation initiates in nascent primordial germ cells in mice. PLoS Genet *3*, e116.

Sutherland, H., and Bickmore, W.A. (2009). Transcription factories: gene expression in unions? Nat Rev Genet *10*, 457-466.

Suzuki, A., Raya, A., Kawakami, Y., Morita, M., Matsui, T., Nakashima, K., Gage, F.H., Rodriguez-Esteban, C., and Izpisua Belmonte, J.C. (2006). Nanog binds to Smad1 and blocks bone morphogenetic protein-induced differentiation of embryonic stem cells. Proc Natl Acad Sci U S A *103*, 10294-10299.

Tada, M., Takahama, Y., Abe, K., Nakatsuji, N., and Tada, T. (2001). Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. Curr Biol *11*, 1553-1558.

Tada, T., Tada, M., Hilton, K., Barton, S.C., Sado, T., Takagi, N., and Surani, M.A. (1998). Epigenotype switching of imprintable loci in embryonic germ cells. Dev Genes Evol *207*, 551-561.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., *et al.* (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science *324*, 930-935.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Takeda, K., Noguchi, K., Shi, W., Tanaka, T., Matsumoto, M., Yoshida, N., Kishimoto, T., and Akira, S. (1997). Targeted disruption of the mouse Stat3 gene leads to early embryonic lethality. Proc Natl Acad Sci U S A *94*, 3801-3804.

Tanaka, S., Kunath, T., Hadjantonakis, A.K., Nagy, A., and Rossant, J. (1998). Promotion of trophoblast stem cell proliferation by FGF4. Science *282*, 2072-2075.

Tani, H., Mizutani, R., Salam, K.A., Tano, K., Ijiri, K., Wakamatsu, A., Isogai, T., Suzuki, Y., and Akimitsu, N. (2012). Genome-wide determination of RNA stability reveals hundreds of short-lived noncoding transcripts in mammals. Genome Res *22*, 947-956.

Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and McKay, R.D. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature *448*, 196-199.

Theunissen, T.W., Costa, Y., Radzisheuskaya, A., van Oosten, A.L., Lavial, F., Pain, B., Castro, L.F., and Silva, J.C. (2011a). Reprogramming capacity of Nanog is functionally conserved in vertebrates and resides in a unique homeodomain. Development *138*, 4853-4865.

Theunissen, T.W., and Silva, J.C. (2011). Switching on pluripotency: a perspective on the biological requirement of Nanog. Philos Trans R Soc Lond B Biol Sci *366*, 2222-2229.

Theunissen, T.W., van Oosten, A.L., Castelo-Branco, G., Hall, J., Smith, A., and Silva, J.C. (2011b). Nanog overcomes reprogramming barriers and induces pluripotency in minimal conditions. Curr Biol *21*, 65-71.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145-1147.

Tomioka, M., Nishimoto, M., Miyagi, S., Katayanagi, T., Fukui, N., Niwa, H., Muramatsu, M., and Okuda, A. (2002). Identification of Sox-2 regulatory region which is under the control of Oct-3/4-Sox-2 complex. Nucleic Acids Res *30*, 3202-3213.

Toyooka, Y., Shimosato, D., Murakami, K., Takahashi, K., and Niwa, H. (2008). Identification and characterization of subpopulations in undifferentiated ES cell culture. Development *135*, 909-918.

Tremblay, A.M., and Giguere, V. (2007). The NR3B subgroup: an ovERRview. Nucl Recept Signal *5*, e009.

Tremblay, G.B., Bergeron, D., and Giguere, V. (2001a). 4-Hydroxytamoxifen is an isoform-specific inhibitor of orphan estrogen-receptor-related (ERR) nuclear receptors beta and gamma. Endocrinology *142*, 4572-4575.

Tremblay, G.B., Kunath, T., Bergeron, D., Lapointe, L., Champigny, C., Bader, J.A., Rossant, J., and Giguere, V. (2001b). Diethylstilbestrol regulates trophoblast stem cell differentiation as a ligand of orphan nuclear receptor ERR beta. Genes Dev *15*, 833-838.

Trouillas, M., Saucourt, C., Guillotin, B., Gauthereau, X., Ding, L., Buchholz, F., Doss, M.X., Sachinidis, A., Hescheler, J., Hummel, O., *et al.* (2009). Three LIF-dependent signatures and gene clusters with atypical expression profiles, identified by transcriptome studies in mouse ES cells and early derivatives. BMC Genomics *10*, 73.

Urlinger, S., Baron, U., Thellmann, M., Hasan, M.T., Bujard, H., and Hillen, W. (2000). Exploring the sequence space for tetracycline-dependent transcriptional activators: novel mutations yield expanded range and sensitivity. Proc Natl Acad Sci U S A *97*, 7963-7968.

Vallier, L., Alexander, M., and Pedersen, R.A. (2005). Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. J Cell Sci *118*, 4495-4509.

van de Wetering, M., Oosterwegel, M., Dooijes, D., and Clevers, H. (1991). Identification and cloning of TCF-1, a T lymphocyte-specific transcription factor containing a sequence-specific HMG box. EMBO J *10*, 123-132.

van den Berg, D.L., Snoek, T., Mullin, N.P., Yates, A., Bezstarosti, K., Demmers, J., Chambers, I., and Poot, R.A. (2010). An Oct4-centered protein interaction network in embryonic stem cells. Cell Stem Cell *6*, 369-381.

van den Berg, D.L., Zhang, W., Yates, A., Engelen, E., Takacs, K., Bezstarosti, K., Demmers, J., Chambers, I., and Poot, R.A. (2008). Estrogen-related receptor beta interacts with Oct4 to positively regulate Nanog gene expression. Mol Cell Biol *28*, 5986-5995.

Van Hoof, D., Munoz, J., Braam, S.R., Pinkse, M.W., Linding, R., Heck, A.J., Mummery, C.L., and Krijgsveld, J. (2009). Phosphorylation dynamics during early differentiation of human embryonic stem cells. Cell Stem Cell *5*, 214-226.

Vanacker, J.M., Bonnelye, E., Chopin-Delannoy, S., Delmarre, C., Cavailles, V., and Laudet, V. (1999a). Transcriptional activities of the orphan nuclear receptor ERR alpha (estrogen receptor-related receptor-alpha). Mol Endocrinol *13*, 764-773.

Vanacker, J.M., Pettersson, K., Gustafsson, J.A., and Laudet, V. (1999b). Transcriptional targets shared by estrogen receptor- related receptors (ERRs) and estrogen receptor (ER) alpha, but not by ERbeta. EMBO J *18*, 4270-4279.

Vernallis, A.B., Hudson, K.R., and Heath, J.K. (1997). An antagonist for the leukemia inhibitory factor receptor inhibits leukemia inhibitory factor, cardiotrophin-1, ciliary neurotrophic factor, and oncostatin M. J Biol Chem *272*, 26947-26952.

Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D.N., Theunissen, T.W., and Orkin, S.H. (2006a). A protein interaction network for pluripotency of embryonic stem cells. Nature *444*, 364-368.

Wang, L., Zuercher, W.J., Consler, T.G., Lambert, M.H., Miller, A.B., Orband-Miller, L.A., McKee, D.D., Willson, T.M., and Nolte, R.T. (2006b). X-ray crystal

structures of the estrogen-related receptor-gamma ligand binding domain in three functional states reveal the molecular basis of small molecule regulation. J Biol Chem *281*, 37773-37781.

Weiss, M.A. (2001). Floppy SOX: mutual induced fit in hmg (high-mobility group) box-DNA recognition. Mol Endocrinol *15*, 353-362.

Williams, D.C., Jr., Cai, M., and Clore, G.M. (2004). Molecular basis for synergistic transcriptional activation by Oct1 and Sox2 revealed from the solution structure of the 42-kDa Oct1.Sox2.Hoxb1-DNA ternary transcription factor complex. J Biol Chem *279*, 1449-1457.

Williams, R.L., Hilton, D.J., Pease, S., Willson, T.A., Stewart, C.L., Gearing, D.P., Wagner, E.F., Metcalf, D., Nicola, N.A., and Gough, N.M. (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. Nature *336*, 684-687.

Winnier, G., Blessing, M., Labosky, P.A., and Hogan, B.L. (1995). Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse. Genes Dev *9*, 2105-2116.

Woltjen, K., Michael, I.P., Mohseni, P., Desai, R., Mileikovsky, M., Hamalainen, R., Cowling, R., Wang, W., Liu, P., Gertsenstein, M., *et al.* (2009). piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. Nature *458*, 766-770.

Wray, J., Kalkan, T., Gomez-Lopez, S., Eckardt, D., Cook, A., Kemler, R., and Smith, A. Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. Nat Cell Biol *13*, 838-845.

Wu, Q., Chen, X., Zhang, J., Loh, Y.H., Low, T.Y., Zhang, W., Sze, S.K., Lim, B., and Ng, H.H. (2006). Sall4 interacts with Nanog and co-occupies Nanog genomic sites in embryonic stem cells. J Biol Chem *281*, 24090-24094.

Wu, Z., Yang, M., Liu, H., Guo, H., Wang, Y., Cheng, H., and Chen, L. (2012). Role of nuclear receptor coactivator 3 (ncoa3) in pluripotency maintenance. J Biol Chem *287*, 38295-38304.

Wurtz, J.M., Bourguet, W., Renaud, J.P., Vivat, V., Chambon, P., Moras, D., and Gronemeyer, H. (1996). A canonical structure for the ligand-binding domain of nuclear receptors. Nat Struct Biol *3*, 206.

Xie, C.Q., Jeong, Y., Fu, M., Bookout, A.L., Garcia-Barrio, M.T., Sun, T., Kim, B.H., Xie, Y., Root, S., Zhang, J., *et al.* (2009). Expression profiling of nuclear receptors in human and mouse embryonic stem cells. Mol Endocrinol *23*, 724-733.

Yabuta, Y., Kurimoto, K., Ohinata, Y., Seki, Y., and Saitou, M. (2006). Gene expression dynamics during germline specification in mice identified by quantitative single-cell gene expression profiling. Biol Reprod *75*, 705-716.

Yagi, R., Kohn, M.J., Karavanova, I., Kaneko, K.J., Vullhorst, D., DePamphilis, M.L., and Buonanno, A. (2007). Transcription factor TEAD4 specifies the trophectoderm lineage at the beginning of mammalian development. Development *134*, 3827-3836.

Yamaguchi, S., Kimura, H., Tada, M., Nakatsuji, N., and Tada, T. (2005). Nanog expression in mouse germ cell development. Gene Expr Patterns *5*, 639-646.

Yamaguchi, S., Kurimoto, K., Yabuta, Y., Sasaki, H., Nakatsuji, N., Saitou, M., and Tada, T. (2009). Conditional knockdown of Nanog induces apoptotic cell death in mouse migrating primordial germ cells. Development *136*, 4011-4020.

Yamaji, M., Seki, Y., Kurimoto, K., Yabuta, Y., Yuasa, M., Shigeta, M., Yamanaka, K., Ohinata, Y., and Saitou, M. (2008). Critical function of Prdm14 for the establishment of the germ cell lineage in mice. Nat Genet *40*, 1016-1022.

Yang, J., van Oosten, A.L., Theunissen, T.W., Guo, G., Silva, J.C., and Smith, A. (2010). Stat3 activation is limiting for reprogramming to ground state pluripotency. Cell Stem Cell *7*, 319-328.

Yang, N., Shigeta, H., Shi, H., and Teng, C.T. (1996). Estrogen-related receptor, hERR1, modulates estrogen receptor-mediated response of human lactoferrin gene promoter. J Biol Chem *271*, 5795-5804.

Yates, A., and Chambers, I. (2005). The homeodomain protein Nanog and pluripotency in mouse embryonic stem cells. Biochem Soc Trans *33*, 1518-1521.

Yeom, Y.I., Fuhrmann, G., Ovitt, C.E., Brehm, A., Ohbo, K., Gross, M., Hubner, K., and Scholer, H.R. (1996). Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. Development *122*, 881-894.

Ying, Q.L., Nichols, J., Chambers, I., and Smith, A. (2003). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell *115*, 281-292.

Ying, Q.L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature *453*, 519-523.

Yoshida, K., Chambers, I., Nichols, J., Smith, A., Saito, M., Yasukawa, K., Shoyab, M., Taga, T., and Kishimoto, T. (1994). Maintenance of the pluripotential phenotype of embryonic stem cells through direct activation of gp130 signalling pathways. Mech Dev *45*, 163-171.

Yoshida, K., Taga, T., Saito, M., Suematsu, S., Kumanogoh, A., Tanaka, T., Fujiwara, H., Hirata, M., Yamagami, T., Nakahata, T., *et al.* (1996). Targeted disruption of gp130, a common signal transducer for the interleukin 6 family of cytokines, leads to myocardial and hematological disorders. Proc Natl Acad Sci U S A *93*, 407-411.

Yoshimizu, T., Sugiyama, N., De Felice, M., Yeom, Y.I., Ohbo, K., Masuko, K., Obinata, M., Abe, K., Scholer, H.R., and Matsui, Y. (1999). Germline-specific expression of the Oct-4/green fluorescent protein (GFP) transgene in mice. Dev Growth Differ *41*, 675-684.

Yuan, H., Corbi, N., Basilico, C., and Dailey, L. (1995). Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. Genes Dev *9*, 2635-2645.

Zaret, K.S., and Carroll, J.S. (2011). Pioneer transcription factors: establishing competence for gene expression. Genes Dev *25*, 2227-2241.

Zhang, X., Zhang, J., Wang, T., Esteban, M.A., and Pei, D. (2008). Esrrb activates Oct4 transcription and sustains self-renewal and pluripotency in embryonic stem cells. J Biol Chem *283*, 35825-35833.

Appendix: Relevant Publications