UNIVERSITÀ DEGLI STUDI DI NAPOLI

"FEDERICO II"

Scuola di Dottorato in Medicina Molecolare

Dottorato di Ricerca in Genetica e Medicina Molecolare
XX ciclo

"Klf5 transcription factor is required to maintain mouse embryonic stem cell pluripotency"

Coordinatore:
Prof. Carmelo Bruno Bruni

Candidato:
Dott. Fabiana Passaro

Anno 2007

INDEX

1. INTRODUCTION

1.1 EMBRYO-DERIVED PLURIPOTENT STEM CELLS

Multicellular organisms need to control and maintain tissue homeostasis by replacing terminally differentiated, aged or injured cells. This mechamism require the existence and precise control of stem cells. During the last years it has been shown that many adult tissues possess groups of precursor cells capable to differentiate into mature tissue-specific cell types. These cells are named somatic stem-cells.

In addition to somatic stem-cell systems, stem cells can be obtained from peri- and early post-implantation-stage embryos. Embryo-derived pluripotent stem cells have been categorized using different names: embryonic stem cells (ESCs), embryonal carcinoma cells (ECCs) and embryonic germ cells (EGCs) (Fig.1). In spite of some differences among them, they have in common some advantages compared to adult stem cells. First, they are easier to identify and to mantain in culture, basically indefinitely, serving as a reservoir of precursor cells. Second, they are pluripotent, possessing the capacity to give rise to all cells of the organism.

Mouse ESC lines were first established in 1981 (Evans and Kaufman, 1981) and they represented a major advance in biology, providing the basis for establishing an in vitro model of early mammalian development, and a putative new source of differentiated cell types for cell replacement therapy.

Mouse ESCs are derived from the inner cell mass (ICM) of E3,5 blastocyst-stage embryos and, like their progenitors, they preserve the ability to generate any derivative of the three primary embryonic germ layers (Fig.2).

Pluripotency of ESCs can be judged by their ability to integrate, after transplantation, into the ICM of E3,5 blastocysts and, similar to the ICM, to express alkaline phosphatase, E-caderin, stage-specific embryonic antigen-1 (SSEA1) and octamer-binding transcription factor-4 (OCT4). ESCs are integrated into the developing embryo and produce a high rate of chimaerism in tissue of developing fetus. Very importantly, they can also give rise to the germline (Bradley et al. 1984).

ESCs have a limited contribution also in extra-embryonic tissues (Bradley et al. 1984; Beddington et al. 1989). In fact, they can contribute to extra-embryonic mesoderm and therefore to composit annexes such as the amnion, the allantois and the yolk sac.

In addition to their developmental potential *in vivo*, ESCs display a remarkable capacity to form differentiated cell types in culture (Smith, 2001). Studies during the past 20 years have led to the development of many different protocols for the generation of a broad spectrum of lineages.

When the factors that maintain them as stem cells are removed, ESCs differentiate and, under appropriate conditions, generate a progeny consisting of derivatives of the three embryonic germ layers: mesoderm, endoderm and ectoderm (Keller, 1995; Smith, 2001). Mesoderm-derived lineages, including the hematopoietic, vascular and cardiac, are among the easiest to generate from ESCs and have been studied in detail.

Ectoderm differentiation is well established, as numerous studies have documented and characterized neuroectoderm commitment and neural differentiation. Each of the three major neural cell types of the central nervous system - neurons, astrocytes, oligodendrocytes - can be generated, and relatively pure populations of each can be isolated when cultured under appropriate conditions (Okabe et al. 1996; Barberi et al. 2003).

Different approaches are used to initiate ESCs differentiation *in vitro* and each one has specific advantages for specific subtype of lineages. ESCs can be allowed to aggregate and

form three dimensional colonies known as embryo bodies (EBs) (Keller, 1995), or they can be cultured directly on stromal cells, and differentiation takes place in contact with these cells (Nakano et al. 1994). Another possibility involves differentiating ESCs in a monolayer on extracellular matrix proteins (Nishikawa et al. 1998).

The ability to derive multiple lineages from ESCs, plus the possibility to manipulate their genome, open exciting new opportunities to model embryonic development *in vitro* for studing the events regulating the earliest stages of lineage induction and specification and/or to study the role of specific genes in specific cell populations.

In addition, the ESCs differentiation system is viewed as a novel and unlimited source of cells and tissues for transplantation for the treatment of many deseases. Type I diabetes, cardiovascular desease, Parkinson's desease, blood cell deseases and certain types of liver desease are considered candidates for cell replacement therapy.

Finally, the ESCs model has widespread applications in the areas of drug discovery and drug development, providing ideal populations of cells for predictive toxicology.

So, a great deal of effort is focused on improvement of culture conditions, generating new ESCs lines suitable for clinical purpose and founding efficient differentiation methods to produce cells suitable for transplantation.

In any case, a detailed understanding of the mechanisms that enable propagation of ESCs in a pluripotent state is essential to realize their therapeutic potential. Infact, understanding the molecular mechanisms by which pluripotency is mantained is the basis for developping new methods to derive and culture ESCs, and understanding the changes which take place in such signaling networks as a cause or consequence of differentiation will help to develop strategies to direct differentiation into defined population suitable for therapeutich uses.

1.2 REGULATORY NETWORKS RESPONSIBLE FOR ESCs SELF-RENEWAL

The precise mechanism that regulates ESCs self-renewal and pluripotency remains largely unknown. Recently, *in vitro* and *in vivo* studies have identified several genetic regulators that may play important roles in these processes, such as extra cellular signaling factors, transcription factors, cell-cycle regulators and microRNA (Fig.3).

Mouse ESCs can be expanded continously in culture if the culture medium is supplemented with the leukemia inhibitori factor (LIF), a member of the IL-6 cytokine family (Smith et al. 1988).

LIF stimulate mouse ESCs through the glycoprotein 130 receptor (gp130), which works as a heterodimer together with LIF receptor (LIFR). Activation of gp130 leads to the activation of the Janus-associated tyrosine kinase (JAK) and of the signal transducer and activatior of transcription (STAT) (Niwa et al. 1998).

STAT3 binds to phosphotyrosine residues on activated LIFR-gp130 heterodimer and undergoes phosphorilation and dimerization itself. Then it can translocate to the nucleus, where it works as a transcription factor (Niwa et al. 1998).

Although activation of STAT3 by LIF is sufficient to prevent mouse ESCs differentiation in the presence of serum (Matsuda et al. 1999), *in vivo* studies demonstrate that LIF cascade is not required for pre-gastrulation mouse development, suggesting that alternative pathways might be involved in maintaining pluripotency *in vivo* and *in vitro*.

In addition to the activation of STAT3, LIF also stimulates the activation of the mitogen-activated protein kinase (MAPK), that promotes differentiation, while suppression of ERKs signal can promote mouse ESCs self-renewal (Burdon et al. 1999). So, MAPK signaling may be a negative regulator for STAT3 and ESCs could maintain their property in the presence of LIF due to the balance of STAT3 activation and MAPK effect.

Unfortunately, the downstream targets of activated STAT3 in mouse ESCs have remained elusive. Recently, in order to isolate these genes, researchers performed a microarray-based kinetic comparison of LIF-stimulated ESCs versus ESCs induced to differentiate by shutting down STAT3 activity throgh either LIF deprivation or expression of a STAT3 dominant-negative mutant (Sekkai et al. 2005). With this experiment some growth factors as Lefty1 or transcriptional regulators as Id1 and Id2 were found to be related to STAT3.

A key target of LIF-STAT3 signaling pathway is the oncogene c-myc. The expression of c-myc is rapidly downregulated within the first 36h after LIF withdrawal. It has been also demostrated that expression of c-myc renders self-renewal indipendent of LIF, while expression of a dominant negative form antagonizes self-renewal and promotes differentiation (Cartwright et al. 2005).

Although forced expression of STAT3 in the absence of LIF is sufficient for mouse ESCs self-renewal, it becomes not sufficient in serum-free medium, indicating that there are other factors in the serum which are required to co-affect with LIF/STAT3.

This serum factor is likely to be bone morphogenetic protein-4 (BMP4), a member of the TGFβ superfamily, acting via activation of Smad1/5/8. Addition of BMP4 to the media enables LIF to maintain mouse ESCs in serum-free culture (Ying et al. 2003).

BMP4 have been shown to phosphorylate Smad1/5 and this results in the expression of inhibitor of differentiation (Id) protein that blocks the neural differentiation. Furthermore, BMP4 can maintain mouse ESCs in the presence of LIF by blocking the MAPK signaling cascade (Qi et al. 2004).

So, it seems that BMP4 contributes to mouse ESCs self-renewal, but this effect need the presence of LIF, and vice versa. The balance between LIF and BMP4 and how this balance of signals is sensed at cell surface are at least in part responsible for maintaining the undifferentiated state of mouse ESCs.

Besides LIF/STAT3 and BMP pathways, evidence has been presented that also the Wnt pathway could be involved in the maintenance of pluripotency. Infact, Wnt signaling is endogenously activated in ESCs and is downregulated upon differentiation (Sato et al. 2004). Its activation by 6-bromoindirubin-3-oxime (BIO), a specific inhibitor of GSK-3, can maintain the undifferentiated phenotype of ESCs and sustain expression of the ESCs specific markers.

Wnt signaling activation can upregulate STAT3 expression, suggesting that it has synergistic effect with LIF/STAT3 (Hao et al. 2006). Furthermore, the Wnt pathway has been shown to elevate the level of c-myc. Thus, Wnt and LIF/STAT3 may converge on c-myc (Cartwright et al. 2005).

So, it is very clear that all this signaling pathways are necessary but not sufficient and that ESCs phenotype need a perfect balance of all them.

Of course, signaling pathways eventually have the nucleus as terminal station and, in the case of ESCs, it results in the transcriptional induction or repression of genes that are responsible for implementing stem-cell plurypotency.

Recent papers have elucidated transcriptional regulatory circuitry responsible for ESCs self-renewal and differentiation, which involves the transcription factors Oct4, Sox2, Nanog etc. Some of these factors are expressed specifically in pluripotent cells, such as Oct4 and Nanog. These transcription factors are switched on/off by the signals described above and they are also regulated by themselves (Fig.4).

Oct4, encoded by Pou5f1 gene, ia a POU domain-containing transcription factor that binds to an octamer sequence (ATGCAAAT). During mouse pre-implantation development, Oct4 expression is activated at the four-cell stage and is later restricted to the pluripotent cells of the ICM and germ cells. Its expression diminishes when these cells differentiate and lose pluripotency.

Oct4-deficient mouse embryos only develop to a stage that looks like a blastocyst, but actually is composed only of trophoectoderm cells. As these structures lack an ICM, they cannot be used to produced ESCs lines (Nichols et al 1998). However, the manipulation of Oct4 expression in mouse ESCs, through inducible or repressible Oct4 transgenes, indicate that precise levels of this gene are required to maintain the ESCs state. Infact, also in culture, loss of Oct4 causes inappropriate differentiation of ESCs into trophoectoderm, whereas it overexpression results in differentiation into primitive endoderm and mesoderm (Niwa 2001).

Several target genes of Oct4 in ESCs have been identified, including Fgf4, Utf1, Opn, Rex1, Nanog and Sox2, an HMG-family protein that occupies many gene targets with Oct4 and is also required for ESCs pluripotency.

LIF does not appear to regulate Oct4 and Oct4 does not appear to regulate Jak-STAT signaling, suggesting that the Oct4 pathway is parallel for maintain self-renewal (Fig.3). Many of Oct4 target genes also contain STAT-binding sites, suggesting a cooperation between these two factors (Tanaka et al. 2002).

In contrast with its target genes, little is known about upstream regulators. *Oct4* promoter contains conserved distal and proximal enhancers that can either repress or activate its expression depending on the binding factors occupying these sites (Pan et al. 2002).

Its expression can be regulated by itself (Chew et al. 2005). Furthermore, FoxD3 and Nanog also can activate Oct4 expression (Pan et al. 2006) (Fig.4).

Another transcriptional factor with an essential role in maintaining the pluripotent phenotype of the ICM is Nanog, an homeobox-containing transcription factor expressed in pluripotent cells, but absent from differentiated cells.

Disruption of Nanog in ESCs results in differentiation into endoderm lineages, while its overexpression renders self-renewal of ESCs indipendent from LIF, although the self-renewal

capability is reduced, suggesting that persistence of Nanog seems to delay, rather than block, differentiation (Chambers et al. 2003; Mitsui et al. 2003).

Nanog overexpression can maintain Id expression (Ying et al. 2003). Moreover, based on the differences in gene expression between wild-type and Nanog-null cells, it has been proposed that Nanog regulates pluripotency mainly as a trancription repressor for downstream genes that are important for cell differentiation, such as Gata4 and Gata6 (Chambers et al. 2003; Mitsui et al. 2003). However, Nanog can also activate the genes necessary for self-renewal such as Rex1 and Oct4 (Shi et al. 2006; Pan et al. 2006).

Concerning Nanog upstream regulators, it seems to be that Nanog is regulated by STAT3 and Brachyury (Brachyury as target of Wnt/β-catenin signaling pathway); then it directly binds to Smad1 and blocks the transcriptional activation by interfering with the recruiment of coactivators (Suzuki et al. 2006).

So, Nanog seems to interact with Wnt and BMP4 pathways and these could also explain the phenomenon that mouse ESCs can maintain self-renewal when overexpressing Nanog, but the capability is reduced (Chambers et al. 2003; Mitsui et al. 2003).

Furthermore, *Nanog* promoter is a direct target of the Oct4/Sox2 complex (Boyer et al. 2005; Kuroda et al. 2005; Rodda et al. 2005).

It has been proposed that Oct4, Sox2 and Nanog may be the main transcriptional factors involved in the regulation of ESCs pluripotency. Recent studies have enabled the construction of transcriptional regulatory networks that provide a foundation for understanding how these factors control self-renewal and influence subsequente differentiation events (Fig.4). Many target genes bound by Nanog, Oct4 and Sox have been identified using RNA interference method, microarray analysis and genome wide chromatin immunoprecipitation experiments, and it seems that these factors form a tight transcriptional regulatory circuitry (Boyer et al. 2005; Ivanova et al 2006; Loh et al. 2006).

Oct4 activity is regulated by interactions with other factors highly expressed in ESCs. In particular, Sox2 has an expression pattern similar to Oct4 and the majority of the Oct4 target genes, included Oct4, Sox2 and Nanog themselves, have octamer and sox heptamer elements separated by either 0 o 3 bp. Oct4/Sox2 complex works often through autoregulatory and multicomponent loop network motifs (Kuroda et al 2005; Rodda et al. 2005).

Oct4/Sox2 complex can regulate *Nanog* expression by directly binding its promoter, maintaining Nanog activity when expressed below steady state, yet represses it at par with or above steady state and this can explain why overexpression of Oct4 induces differentiation (Pan et al. 2006).

In addition to Oct4/Sox2 cooperative binding, interaction between Oct4 and other factors has also been suggested and reported, as in the case of FoxD3.

FoxD3 (forkhead familt member D3) appears to be nonessential. Isolated as a transcription factor with restrictive expression in ESCs, FoxD3 has been described as a novel activator of Nanog and counteracts the repressive activity of Oct4 on Nanog promoter at the steady state (Pan et al 2006).

So, in summary, this is only a small part of the story, not jet completely understood, but it gives a strong evidence that this regulatory circuit works sustaining transcription factors expression at precise levels.

1.3 shRNA INTERFERENCE TO INVESTIGATE MECHANISMS OF PLURIPOTENCY

In order to understad the connections and biological functions of all these biochemical patways, many research groups decided to exploit a systematic approach, using microarray tecnology as well as genome wide chromatin immunoprecipitation experiments (Boyer et al.

2005; Loh et al. 2006). RNA interference method is also useful for this scope (Ivanova et al.2006).

Gene silencing by RNA interference in mammalian cells using short hairpin RNAs (shRNA) has become an excellent tool to investigate gene function by inducing efficient and specific gene knockdown. It is a sequence-specific gene-silencing mechanism that is based on dsRNA molecules. dsRNAs are processed into 21-25 bp dsRNAs called small interfering RNAs (siRNAs) by Dicer, a ribonuclease III family (RNaseIII) enzyme (Zamore et al 2000). The resulting small RNAs enter the RNA-induced silencing complex (RISC), which uses a single stranded version of the small RNA as a guide to select the substrates (Zamore et al 2000; Hammond et al. 2000). Perfect complementarity between the substrate and the small RNA leads to target-RNA cleavage by RNAse H – like active site within an Argonaute protein that forms the core of RISC (Liu et al. 2004) .

miRNAs are a class of endogenus dsRNAs that exert their effects through the RNAi patway. They are important in a wide variety of biological processes including cell-cycle regulation, apoptosis, cell differentiation and, also, maintenance of stemness (Ambros 2004).

miRNAs are transcribed by RNA polymerase II as long primary polyadenylated transcripts (pri-miRNA). The pri-miRNA is recognized and cleaved at a specific processing site by the RNase III enzyme, Drosha, in the context of the Microprocessor complex, to produce an miRNA precursor (pre-miRNA) of approximately 70-90 nucleotides (Han et al. 2006)

The pre-miRNA has a 2-nt 3’ overhang at one end. This distinctive structure is recognized by the Exportin-5-Ran-GTP heterodimer, and the pre-miRNA is shuttled to the cytoplasm. Only then the miRNA precursor is recognized and processed by Dicer into a mature miRNA, using the 3’overhang as a guide for site-specific cleavage at the second processing site (Siolas et al. 2005) (Fig.5).

The discovery of endogenous triggers of RNAi suggested that RNAi might be induced in mammalian cells by synthetic genes that mimic the activity of these naturally occurring regulatory molecules, providing an excellent tool to investigate gene functions in a wide variety of phisiological and pathological precesses.

1.3 AIMS OF THE PROJECT

Recurring questions relate to the existence of other, not yet identified, stemness genes and how to find them. As a matter of fact, although the stemness status applies to the genes described above, neither Oct4 nor Nanog can be considered really the unique "master genes" of pluripotency, considering that on LIF withdrawal Oct4 cannot prevent differentiation of ESCs on its own and Nanog is ineffective without Oct4. Are there other genes that mediate pluripotency? Are there any "master genes"?

So, several questions regarding the mechanisms of pluripotency remains open.

This project set the goal to be a part of the effort in understanding the biology of ESCs.

For this reason, we have performed an high-throughput screening by using RNA interference to identify the factors that regulate the balance between stemness and differentiation commitment.

To this end, we have generated an ES reporter cell line which renders possible to follow neuronal commitment and differentiation by EGFP expression.

The readout of this systematic analysis is that downregulation of crucial genes by shRNAs involved in the fine control of ESC pluripotency network or neuronal commitment will suppress neural differentiation, as well as GFP expression (Fig.6).

Among the shRNAs efficently suppressing ESCs differentiation, one shRNA targeting the Kruppel-like transcription factor 5 (KLF5) has been choosen for further investigations.

We have demonstrate that KLF5 is an essential factor of the core regulatory network responsible for maintaining ESCs pluripotency.

2. MATERIALS AND METHODS

2.1 EXPERIMENTAL DESIGN OF THE SCREENING

2.1.1 GENERATION AND CHARACTERIZATION OF EGFPN1 STABLE CELL LINES

It has been developped an *in vitro* system to screen a collection of 76,896 shRNA designed to target almost the whole mouse trascriptome (Chang et al. 2006).

This system is based on the generation of an ESC line stably expressing EGFP under the control of the neuronal-specific promoter *α1-tubulin* (Schmandt et al. 2005). *α1-tubulin* can be considered a structural marker for neuronal precursor cells, as well as post mitotic neurons. In neural-differentiation conditions, these cells generate neurons that express EGFP, according with the expression pattern of α1-tubulin. Other cell types obtained from this differentiation protocol, as astrocytes, olgodendrocytes ad epitelial cells do not express α1-tubulin.

In order to generate the α1-tubulin-EGFP vector, a fragment of about 1kb (-1022 / +4 from the start site) of the rat *α1-tubulin* promoter (Schmandt et al. 2005) was amplified from rat genomic DNA template using as forward primer an oligonucleotide spanning a SmaI restriction site (5'-TCCCCCGGGCCGTATTAGAAGGGATGGCTC-3') and as reverse primer an oligonucleotide bearing a BamHI site (5'-CGCGGATCCGACTCTTAAGCGGTCGATGT-3'). The fragment was amplified by PCR using 2,5U of TaKaRa Ex Taq^{TM} DNA polymerase (TAKARA Bio Inc.) and these conditions: 35 cycles and denaturation at 98°C for 15 sec, annealing at 58°C for 35 sec, elongation at 72°C for 2 min.

The amplified fragment was digested SmaI/BamHI (NEB) and cloned Blunt/BamHI into pEGFP N1 vector (Clontech) replacing the CMV promoter.

The α1-tubulin-EGFP vector was electroporated with Gene Pulse II Electroporator (BioRad) at 250V and 400μF into E14Tg2a (BayGenomics) using 50μg of DNA every two 100mm tissue colture plate at high cell confluency. Recombinant clones were selected with 500 mg/ml of G418 (Invitrogen), starting 20 hours after electroporation.

About 30 positive clones for α1-tubulin-EGFP vector were collected from plates and seeded in 96 well plates. The presence of GFP cDNA was confirmed by PCR. Positive clones were then pooled together and characterized by immunofluorescence analysis for their ability to differentiate towards cardiomyocites (mesoderm) and neurons (neuroectoderm). Of course, with the neural differentiation protocol (see below), cells begin to express GFP accordingly with α1-tubulin expression (Fig.7 A). Co-immunofluorescence with α1-tubulin and β3-tubulin antibodies demonstrates co-expression of EGFP signals with neuronal differentiation markers at day 13 (Fig.7 B).

In order to better determine the fate of cells in which gene silencing brought to perturbation of α1-tubulin EGFP expression, another cell line that could be used to monitoring the undifferentiated phenotype of ESCs after transfection with shRNAs was generated.

The Nanog-EGFPN1 stable cell line was generated by cloning a 382 bp fragment (-332/+50) form mouse DNA genomic template, representing the minimal region upstream of the *Nanog* transcription start site, able to ricapitulate the right Nanog gene expression in ESCs, into the AseI /NheI sites of pEGFPN1 vector (Chambers et al. 2003). A forward primer bearing an AseI restriction site (5'-TAGTCAATTAATATCGCCAGGGTCTGGAGG-3') and a reverse primer in which an AvrII restriction site, compatible with NheI present into the vector, has been added (5'-TAGTCACCTAGGCGCAGCCTTCCCACAGAAA-3') were used to amplified this fragment by PCR using 2,5U of TaKaRa Ex Taq and these conditions: 35

cycles and denaturation at 98°C for 10 sec, annealing at 64°C for 30 sec, elongation at 72°C for 1 min. Electroporation, selection and PCR screening of positive clones were the same as α1-tubulin EGFPN1 cells.

In the case of Nanog-EGFPN1, cells express EGFP only in undifferentiated state, while the reporter is shut down early during differentiation protocol. So, this system provides a useful tool to reveal the persistence of stemness under differentiation conditions (Fig.8 A).

Last, to evaluate the possibility that perturbation of pluripotency by RNAi could led to aberrant formation of mesodermal derivates, even in neural differentiation conditions, another stable cell line, in which the expression of EGFP was under the control of the rat *Miosin Light Chain-2v* (MLC-2v) promoter, structural marker of cardiomyocites, was generated. To give more strenght to the expression of EGFP, we decided to insert the CMV enhancer upstream the *MLC-2V* promoter. Firstly, a 400bp fragment from pEGFPN1 template consisting of the CMV enhancer was amplified by PCR, using 1,5U of KOD HIFI DNA Polymerase (Novagen), a forward primer that include an AseI restriction site (5'-TAGTCAATTAATCCGCGTTACATAACTTACGG-3') and a reverse primer in which NheI and EcoRI sites were added at 5' terminus of the oligonucleotide (5'-CTAGCTAGCGAATTCCAAAACAAACTCCCATTG-3') with these conditions: 20 cycles and denaturation at 98°C for 15 sec, annealing at 62°C for 7 sec, elongation at 72°C for 20 sec.

This fragment was digested and inserted into the AseI/NheI sites of EGFPN1, now called ECMV-GFPN1. Then a 300bp fragment of the *MLC-2v* rat promoter was amplified from rat genomic DNA template with an EcoR1 forward primer (5'-CCGGAATTCCAGGACCCAGAGCACAGAGC-3') and an NheI reverse one (5'-CTAGCTAGCAAGGAGCCTGCTGGCCGGC-3') using these conditions: 25 cycles and denaturation at 98°C for 15 sec, annealing at 68°C for 30 sec, elongation at 72°C for 26 sec.

1,5U of KOD HIFI DNA Polymerase were used. The *MLC-2v* promoter was then digested and cloned into the EcoR1 and NheI sites of the ECMV-GFPN1. Again, it has been followed the same protocol of electroporation, selection and PCR screening to obtain a stable cell line.

The MLC2v-EGFPN1 cell line express EGFP only if mesodermal differentiation is promoted, using Embryo Body formation (see cell culture, transfection and differentiation) or after silencing of those genes involved in control of balance between neuroectoderm and mesoderm derivates. So, this cell line apperared a good system to evaluate a possible shift of cell fate (Fig8 B).

While α1-tubulin-EGFPN1 cell line was used to perform the screening, Nanog-EGFPN1 and MLC-2v EGFPN1 were used during subsequent characterization of the phenotype obtained from gene silencing of candidate genes.

2.1.2 shRNA COLLECTION AND TRANSFECTION IN E14Tg2A

Gregory Hannon and his co-workers constructed a large scale library of artificial miRNAs that cover the majority of genes in the mouse and human genomes (shRNA-mir library). These libraries are based upon modified primary miRNA transcripts (Chang et al. 2006) .

shRNA-mir collection is from Open Biosystems and consists of about 76,896 shRNAs construct organized in 750 96 well plates. Each well of the plate contains a bacterial stub transformed with a single plasmid vector pSM2 coding for a single shRNA targeting a single mouse mRNA. The following table summarizes the gene coverage distribution, with an average of 2-3 shRNA per target gene.

19% of target genes (5,740) covered by exactly 1 shRNA
31% (9,359) covered by 2 shRNAs
36% (10,983) covered by 3 shRNAs
15% (4,547) covered by ≥ 4 shRNAs

The pSM2 vector was designed to contain the shRNA in an miR-30 cassette that can be packaged in a self-inactivating murine stem cell virus (MSCV). The expression of the small RNA is driven by the U6 promoter along with the U6 snRNA leader sequence. The U6 snRNA leader sequence lies between the promoter and the 5'end of the miR-30 flanking region, and a pol III terminator signal is inserted immediatly after the miR-30 cassette (Chang et al.2006).

The pSM2 vector contains a puromycin resistance that enable selection of stable shRNAs expressing mammalian cell lines.

Before starting the screening, the shRNA library was duplicated an the plasmid DNA recovered. To this aim, the BIOMEK FX robot (Beckman) was used to amplifiy each multiwell in automation. Each well of the library has been duplicated in order to preserve a perfect copy for future experiments. To do this, there have been developped custom programs to fill five 96 deep well plates with 1,2ml of 2X LB (10g Peptone, 20g yeast extract, 5g NaCl per liter) simultaneusly by the robot and to inoculate two microliters of each stub in these five 96 deep well plates. Each deep well has been incubated over night at 37°C with shaking. The day after, five microliters of saturated LB were used to generate the second stub of the original 96 well plate. All the rest of the bacterial cells were pelletted and processed for DNA extraction and the same platform was used to perform high-throughput mini preparations of plasmid DNA using Montage Plasmid Mini Prep kit (Millipore) that enables to process an entire 96well plate mini prep in about 45 minuts, recovering for each sample an high purified plasmid DNA ready to be transfected. The quantification of DNA yield was performed in 96 well plate using Quanti-IT PicoGreen (Invitrogen), according with manufacturer instructions. Fluorescence signal was detected by EnVision 2102 Multilabel Reader (PerkinElmer).

Trasfection of shRNAs into α1-tubulin EGFPN1 ESCs, was performed in 96 well plates (Falcon) to preserve the same order of the library. Cells were plated in gelatin-coated 96 well plates 12 hours before transfection at a confluence of about $3x10^4$ cells/well and then transfected with 500ng of each shRNA plasmid by using 2µl of ArrestIn (Open Biosystems). Transfected cells were selected in puromycin 2,3µg/ml (Sigma) for three days.

Selected cells in each well were trypsinised in order to pull the recombinant clones and replated with differentiation medium at low confluence to induce neural differentiation.

2.1.3 ESC DIFFERENTIATION AND GFP DETECTION

To induce neural differentiation, we have introduced a novel efficient system for *in vitro* neural differentiation of mouse ESC by plating the cells at low density in a chemically-defined medium. ESCs were trypsinised into a single cell suspension, collected by centrifugation and resuspended in the following medium: Knockout Dulbecco's minimal essential medium supplemented with 10% Knockout Serum Replacement (both from Invitrogen), 0,1 mM β-mercaptoethanol (Sigma), 2 mM glutamine, 100 U/ml penicillin/streptomycin (Gibco).

Then cells were plated at low density (1-5 x 10^3 cells/cm^2) on gelatine coated 96 well plates and differentiation medium was changed on alternated days.

After 4-5 days, α1-tubulin EGFP-positive cells appear. Starting from day 7 of differentiation onward, a population of cells showing a strong EGFP expression was observed, the neuronal subtypes population. We analysed the presence/absence of EGFP by observation at fluorescence-inverted microscope (DMI4000, Leica Microsystems) at different time points (4, 7 and 10 days of differentiation).

shRNAs, whose transfection into ESCs disrupted normal α1-tubulin expression pattern, were transfected at least 3 times to confirm the observed phenotype. To be sure that the absence of

GFP was due to misregulation of differentiation and to exclude false positives results, an immunofluorescence assay was performed for each "positives" directly in 96 well plates, using anti α1-tubulin antibody (1:400, SIGMA-Aldrich), that enables the detection of endogenus α1-tubulin. Cells were fixed in 4% paraformaldehyde and permeabilised with 0,2% TX-100 in 10% normal goat serum (Dako Cytomation) 1% BSA in 1X PBS for 15 min at room temperature. Then the samples were incubated with primary antibody over night at 4°C. After 3x5min washes with PBS, texas red-conjugated anti mouse antibody was used as secondary antibody and incubated for one hour at room temperature. Images were captured with an inverted microscope (DMI4000, Leica Microsystems). Strong reduction or absence of GFP reporter signal, confirmed by the evaluation of endogenus α1-tubulin signal, was considered as indicative of an alteration of ESC differentiation.

2.2 CELL CULTURE, TRANSFECTIONS AND DIFFERENTIATION

E14Tg2a (BayGenomics) mouse ESCs were maintained on feeder-free, gelatine coated plates in the following medium: GMEM (Sigma) supplemented with 2 mM glutamine, 100 U/ml penicillin/streptomycin (both from Invitrogen), 0,1 mM β-mercaptoethanol (Sigma), 10% FBS (Hyclone), 1 mM sodium pyruvate, 1X non essential amino acid (both from Invitrogen) and 1,000 U/ml leukemia inhibitory factor (LIF) (Chemicon). NIH3T3 were cultured in DMEM (Invitrogen) supplemented with 10% FBS, 2 mM glutamine, 100 U/ml penicillin/streptomycin.

For transfection experiments, ESCs were plated at $7x10^4$ cells/cm^2 12 hours before transfection. Transfections of plasmids, shRNAs and siRNA smart pool (Dharmacon) were performed using Lipofectamine 2000 (Invitrogen).

ESCs neural differentiation has been described above. Embryo Body differentiation of ESCs was obtained as described in Maltsev et al.(1993).

2.3 PLASMID CONSTRUCTION

Full-lenght *Klf5* was obtained from NIH Mammalian Gene Collection (Invitrogen; Clone ID:4208633) and ligated into the HindIII and EcoRI sites of p3xFlag-CMV^{TM} 7.1 expression vector (Sigma). The 3xFlag-Klf5 cDNA was then amplified by PCR and subcloned under the control of chicken β-actin promoter into an expression vector kind gift of A. Simeone. PCR was performed using 1,5U of KOD HIFI DNA Polymerase and a forward primer containing a BamHI site (5'-ATGGACTACAAAGACCATGACGGTG-3') and a reverse primer in which a Not1 site was inserted (5'-TCAGTTCTGGTGGCGCTTCATG-3'). These conditions were used: 30 cycles and denaturation at 98°C for 15 sec, annealing at 60°C for 5 sec, elongation at 72°C for 14 sec.

To generate the Nanog luciferase reporter vector a fragment of 379 bp of the mouse *Nanog* promoter (-329/+50 from start site) was amplified from mouse genomic DNA, using 1,5U of KOD HIFI DNA Polymerase with a reverse primer encompassing a BglII restriction site (5'-AGATCTCGCAGCCTTCCCACAGAAA-3') and a forward primer into which a KpnI restriction site was introduced (5'-GGTACCATCGCCAGGGTCTGGA-3') with these conditions: 25 cycles and denaturation at 98°C for 15 sec, annealing at 56°C for 5 sec, elongation at 72°C for 8 sec. The amplified fragment was then digested and cloned into KpnI and BglII sites of pGL3-basic vector (Promega) (Fig.10).

For the Oct4 luciferase reporter vector, a fragment of 2,2kb of the mouse *Oct4* promoter (-2200/+16 from start site) was amplified using a forward primer carrying the KpnI restriction site (5'-GGTACCCAAAAGAGAAATCACAATCCATAAGACAAGGTTGG-3') and a reverse primer where a BglII site was introduced (5'-AGATCTTGGAAAGACGGCTCACCTAGGGAC-3'). PCR was performed with these conditions: 35 cycles and denaturation at 95°C for 10 sec, annealing at 58°C for 30 sec,

elongation at 72°C for 3 min. 2,5U of TaKaRa Ex Taq™ DNA polymerase (TAKARA Bio Inc.) were used.

The amplified fragment was digested and ligated to the KpnI/BglII sites of pGL3-Basic (Fig.10).

2.4 RNA ISOLATION, REVERSE TRANSCRIPTION AND REAL-TIME PCR

Total RNAs were extracted using Trizol ultra pure reagent (Invitrogen) according to the manufacturer's instructions and subsequently incubated with RNase-free DNase (Ambion). 2µg/reaction of RNA were reverse transcribed with M-MuLV reverse transcriptase (Biolabs) using random hexamers (Amersham) with these conditions: 25°C for 10 min, 48°C for 60 min, 98°C for 5 min. cDNA samples were subjected to PCR amplification using specific primers for undifferentiated and differentiated ESC markers. The PCR was carried out using standard protocols with EuroTaq DNA polymerase (0,2U/reaction) and dNTPs 0,2 mM (Euroclone). Cycling parameters were as follows: denaturation at 94°C for 45sec, annealing at 56°C for 45 sec and elongation at 72°C for 45 sec. The primer sequences, number of cycles and the size of amplified fragments in semiquantitative RT-PCR are reported in Table 1.

Table 1

Gene	Forward primer	Reverse primer	Size (bp)	Number of cycles
Klf5	TCGGAGGAGCTGGTCCAGAC	CCGTCGACTCGCTCTGGTG	286	24
Klf4	CACACAGGCGAGAAACCTTACC	TCCTTTCTCCTGATTATCCATTCAC	352	30
Nanog	ATGAAGTGCAAGCGGTGGCAGAAA	CCTGGTGGAGTCACAGAGTAGTTC	463	24
Oct3/4	GCAGGAGCACGAGTGGAAAGCAAC	CAAGGCCTCGAAGCGACAGATG	296	24
Gata	CCGAGCAGGAATTTGAAGAGG	GCCTGTATGTAATGCCTGCG	479	35
Cdx2	GGCGAAACCTGTGCGAGTGGATGCGGAA	GATTGCTGTGCCGCCGCCGCTTCAGACC	492	35
Hnf4	GAGGTCCATGGTGTTTAAGGAC	CTGCAGCAGGTTGTCAATCTTGG	410	35
Sox1	CCTCGGATCTCTGGTCAAGT	TACAGAGCCGGCAGTCATAC	593	35
Afp	TCGTATTCCAACAGGAGG	AGGCTTTTGCTTCACCAG	174	35
Brachyury	CTCTAATGTCCTCCCTTGTTGCC	TGCAGATTGTCTTTGGCTACTTTG	194	35
Meox	TGGCCTATGCAGAATCCATTCC	TGAGATCTGAGCTGCGCATGTG	346	35
Eomes	TGCAAGAGAAAGCGCCTGTCTC	CAATCCAGCACCTTGAACGACC	398	35
Gapdh	GTATGACTCCACTCACGGCAAA	CTAAGCAGTTGGTGGTGCAG	326	24

The expression of the GAPDH housekeeping gene was used to normalize PCR reactions. Real Time PCR was carried out by using ABI PRISM 7900HT Sequence Detection System (Applied Biosystems) and Power SYBR Green PCR master mix (Applied Biosystems). Gene specific primers used for amplification are listed in Table 2.

Table 2

Gene	Forward primer	Reverse primer
Klf5	GGTCCAGACAAGATGTGAAATGG	TTTATGCTCTGAAATTATCGGAACTG
Klf4	ACTCACACAGGCGAGAAACCTTAC	TCAGTTCATCGGAGCGGG
Nanog	TCAGAAGGGCTCAGCACCA	GCGTTCACCAGATAGCCCTG
Oct3/4	AACCTTCAGGAGATATGCAAATCG	TTCTCAATGCTAGTTCGCTTTCTCT
Sox2	CTGCAGTACAACTCCATGACCAG	GGACTTGACCACAGAGCCCAT
Gapdh	GTATGACTCCACTCACGGCAAA	TTCCCATTCTCGGCCTTG

2.5 PROTEIN EXTRACTION AND WESTERN BLOTTING

To obtain protein extracts, cells were lysed in Laemli lysis buffer. Total lysates were separated by SDS-PAGE (Biorad Laboratories) and trasferred to PVDF membrane (Biorad) using the Trans-Blot Semi-dry System (Biorad Laboratories). The membrane were probed with the following antibodies for: Klf5 (KM1784), Oct3/4 (monoclonal, Santa Cruz), Gapdh (Santa Cruz). Antibody-protein complexes were detected by HRP-coniugated secondary antibodies (Santa Cruz and Amersham Bioscience) and ECL (Amersham Bioscience).

2.6 IMMUNOFLUORESCENCE OF CELLS AND EMBRYOS

ESCs and blastocysts were fixed in 4% paraformaldehyde and permeabilised with 0,2% TX-100 in 10% normal goat serum (Dako Cytomation) 1% BSA in 1X PBS for 15 min at room temperature. Then the samples were incubated with primary antibodies at the following working diluitions: βIII-tubulin (1:400, SIGMA-Aldrich), α1-tubulin (1:400, SIGMA-

Aldrich), TROMA-1 (1:50, obtained from the Developmental Studies of Hybridoma Bank, University of Iowa), Oct3/4 (1:100, Santa Cruz), Nanog (1:100, R&D Systems), Klf5 (1:300), KM1784). Following primary antibodies incubation, either cells or blastocysts were incubated with appropriate secondary antibodies detecting mouse, rabbit, goat and rat IgG conjugated with Alexa Fluor 594 or 488 (Molecular Probes). Images were captured with an inverted microscope (DMI4000, Leica Microsystems) and in case of blastocysts with a confocal microscope (LSM 510 META, Zeiss).

2.7 ALKALINE PHOSPHATASE STAINING

For alkaline phosphatase staining, cells were fixed in 10% cold Neutral Formaline Buffer (100 mL formaline; 16g Na_2HPO_4; 4g $NaH_2PO_4.H_2O$) for 15 min and then rinsed once with distilled water for 15 min.

After washing, cells were incubated at room temperature for 45 min with freshly prepared substrate (0,005g Naphthol AS Mx-PO_4; 200µL N,N-Dimethylformamide; 25 mL Tris-HCl 0,2M pH8,3; 0,03g Red Violet LB salt), rinsed in distilled water 4 times and then leaved in water.

2.8 LUCIFERASE ASSAY

Luciferase assay was carried out in ESCs and NIH3T3 cells. Cells were seeded in 12 well plates (Falcon) at $3x10^4$ cells per well 12 hours before transfection. Oct4-Luc, Nanog-Luc and Renilla reporter vectors, Klf5-3xFlag, siRNAs smart pool were transfected with Lipofectamine 2000 (Invitrogen) in a ratio of 1,6 µg reporter vector for 0,8 µg Klf5-3xFlag or siRNAs and 20 ng Renilla. Firefly and Renilla luciferase activities were mesured with the Dual-Luciferase Assay System (Promega), according with manufacturers instructions, 24

hours after transfection by Sirius Luminometer (Berthold Detection Systems). The data generated were normalized to Renilla luciferase reading.

2.9 ChIP ASSAY

ChIP assays were carried out in wt or Klf5 overexpressing ESCs. Briefly, cells were cross-linked with 1% formaldehyde for 10 min at room temperature and formaldehyde was then inactivated by the addition of 125 mM glycine. Cells were washed twice with PBS, then scraped and pelletted in order to obtain about $1x10^7$ cells per sample. Lysis was performed in 500µl of lysis buffer (1%SDS, 10mM EDTA pH8, 50mM Tris-HCl pH8, Protease inhibitory cocktail 1X). The chromatin was then sonicated to an average DNA fragment length of 200 to 1000bp (3 pulse of 10 sec, power 6), clarified by centrifugation at 13.000 rpm for 20 min at 4°C and diluited 10 fold with diluition buffer (1%Triton, 2mM EDTA, 150mM NaCl, 20mM Tris-HCl pH8, protease inhibitory cocktail 1X) (about $2x10^6$ cells per ml).

Soluble chromatin extracts were immunoprecipitated using mouse monoclonal anti-Flag (1:200, Sigma-Aldrich) or goat anti-Nanog (1µg, R&D Systems) antibodies over night at 4°C. Then, 30µl of Protein A/G plus (Santa Cruz) were added to immunoprecipitations for 1 hour. Supernatants obtained without antibody were used as an input control.

The immunoprecipitated samples were washed to eliminate the unbound chromatin with these buffers: 1x5min wash with Low Salt Buffer (0,1% SDS, 1% Triton, 2mM EDTA, 20mM Tris.HCl pH8, 150mM NaCl); 1x5min wash with High Salt Buffer (like low salt but with 500mM NaCl); 1x5min wash with LiCl Buffer (0,25M LiCl, 0,5% NP40, 0,5% Sodium Deoxicolated, 1mM EDTA, 10mM Tris-HCl pH8); 2x5min wash with TE, pelletting beads between washes. For eluition, 2x250 µl of fresh prepared Eluition Buffer (1% SDS, 0,1M NaHCO3) were used. Decrosslinking was performed over night at 65°C and the morning after PK was added to a final concentration of 0,5mg/ml. Recovery of DNA was obtained by

purification with Phenol/chlorophorm and subsequent precipitation in ethanol. For all ChIP experiments, PCR analysis were performed in real time using the ABI PRISM 7900HT sequence detection system and SYBR Green Master Mix (Applied Biosystems). The amount of precipitated DNA was calculated relative to the total input chromatin and expressed as % of total chromatin, according to the following formula: $2^{\{Delta\}Ct}$ x 10, where {Delta}Ct = Ct(input)-Ct(immunoprecipitation). (Ct) Cycle threshold. Oligonucleotide sequences are summarize in Table 3.

Table 3

Klf5	**Forward primer**	**Reverse primer**
K1	GTGATCGGCACGCTTTCTAGT	TTGGATCAATGGAAACCCAGA
K2	CGGATCTGTCCAGAGGAGCTAAAG	CGAGAGGAAATTTACGCGCACACT
NANOG		
N1	GAGGATGCCCCCTAAGCTTTCCCTCCC	CCTCCTACCCTACCCACCCCCTATTCTCCC
POU5f1		
CR1	AGCAACTGGTTTGTGAGGTGTCCGGTGAC	CTCCCCAATCCCACCCTCTAGCCTTGAC
CR2	TGCTCTGGGCTTTTTGAGGCTGTGTGATT	TGGCGGAAAGACACTAAGGAGACGGGATT
CR4	GGAACTGGGTGTGGGGAGGTTGTA	AGCAGATTAAGGAAGGGCTAGGACGAGAG

3. RESULTS

3.1 SYSTEMATIC SCREENING OF GENES THAT REGULATE ESCs FATE

In this study, a systematic analysis to identify genes that regulate the balance between stemness and differentiation commitment of mouse ESCs has been performed.

The shRNA tecnology has been choosen to stably downregulate gene expression and, at the same time, an *in vitro* system has been developped in order to analyze the phenotype obtained after gene silencing (see Metods and Fig.6).

To this aims, three ES cell lines stably expressing the EGFP reporter gene under the control of differentiation-specific promoters were generated: the *Nanog* promoter, marker of pluripotent cells, the *α1-tubulin* gene promoter, structural marker of neuronal precursor as well as post-mytotic neurons and *Myosin light chain-2v* (MLC2v) promoter, marker of cardiomyocites. In this way three reporter systems were obtained to monitor three different cell types: stem cells, neurons (neuroectoderm commitment) and cardiomyocites (mesoderm commitment). *Nanog* promoter drives expression of the reporter gene only in plurypotent cells, infact GFP signal disappear after induction of differentiation. α1-tubulin-GFP cells starts to express GFP at day 4 of neural differentiation, when ESCs became neuronal-precursor cells, and continues to be fluorescent after day 13, when precursor cells turn into post-mytotic neurons. Last, the GFP expression under the control of *MLC2v* is detectable only in mature cardiomyocites generated after 8 days of mesoderm differentiation. So it is possible to monitor differentiation not only by cell morphology, but easly by GFP detection.

Concerning neuroectoderm commitment and neural differentiation, we have introduced a novel *in vitro* protocol to induce neural differentiation, based on a chemically defined medium that enables the generation of different subpopulation of neurons. Furthermore, this protocol has been adapted to work in 96 well plates.

10,850 short hairpin RNAs, about 1/3 of the entire library, were firstly screened in the α1-tubulin-GFP cell line.

To minimize the background due to non-transfected cells, selection with puromycine has been intoduced to obtain stable interferred clones. Furthermore, because neuronal differentiation occurs in ten days, stable clones could be a valid tool to investigate phenotypes depending on downregulation of genes activated in later sages of differentiation.

From this first set of transfection several genes were collected, summarized in table 4, whose silencing was able to stop neuronal differentiation (absence of GFP signal).

Table 4

KNOWN GENES	**Olfr1238**	**Klf5**	**Pou2f1**	**Pdcd6**	**Epha3**	**Scy1L2**
	Olfr 22	**Cnga3**	**Clstn1**	**Melk**	**Map2k6**	
	Olfr 355	**Gap43**	**Keg1**	**D10Ertd802e**	**Saa3**	
	Olfr1188	**Gna13**	**Adfp**	**Epha2**	**Casp3**	

UNKNOWN GENES	**RIKEN cDNA 4632428N05 gene**	**RIKEN cDNA E130310K16 gene**
	RIKEN cDNA 4632404H12 gene	**RIKEN cDNA 4921522P10 gene**
	RIKEN cDNA 4933412D19 gene	**RIKEN cDNA 1810032O08 gene**
	RIKEN cDNA 4833418A01 gene	

Interestingly, some of these “candidate” genes were just annotated as RIKEN cDNAs with unknown functions. The rest of the genes were already described and significant coherence with literature has been founded. For example, the Growth-Associated protein (GAP)-43 is normally expressed at high levels in neuronal growth cones during development and axonal

rigeneration. Knockout mice borne at low frequency and only 10% of mice with a full gene deletion survive, suggesting that basic neural functions are disrupted.

Pou2f1 gene encoded for Oct1 protein that regulates the expression of GnRH gene in neurons and Oct1-deficient embryos died during gestation.

Melk, or maternal embryonic leucine zipper kinase, seems to be necessary for proliferation of embryonic and postnatal MNP and it should regulate the transition from GFAP-expressing progenitors to rapid-amplifying progenitors in the postnatal brain.

It is also interesting to notice the presence of a significant number of genes belonging to the same families as olfactory receptors, Eph receptors and G protein coupled receptors.

For each candidate gene, I started with a preliminary characterization in order to check the normal expression pattern in pluripotent ESCs and during neural differentiation, performing RT PCRs on total mRNAs obtained from wild type undifferentiated and differentiated cells. So it has been found that some genes were constitutively expressed, while some others were upregulated or downregulated during differentiation.

Although the same phenotype for all candidate genes was observed, that is the absence of GFP signal and so the absence of neurons, many different causes could explain this phenotype.

One possibility is that silencing caused the arrest of differentiation because the silenced gene has a key role in promoting neuronal differentiation. Another possibility is that knockdown lead to a shift of cell fate toward different lineages.

The third explanation is that the delicate equilibrium of factors that regulate stem cell pluripotency has been perturbed, rendering these cells unable to differentiate.

All these possibilities required different approaches and a huge work, so I decided to focus my attention on a single gene.

I choose the Kruppel-like transcription factor Klf5, member of the Sp1-like/Krüppel-like factors (KLFs) that are highly related zinc-finger proteins that are important components of the eukaryotic cellular transcriptional machinery (Kaczynski et al. 2003). By regulating the expression of a large number of genes that have GC-rich promoters, Sp1-like/KLF transcription regulators may take part in virtually all facets of cellular function, including cell proliferation, apoptosis, differentiation, and neoplastic transformation.

3.2 KLF5 KNOCKDOWN SUPPRESS NEURONAL DIFFERENTIATION

Among the shRNAs efficiently suppressing ESC differentiation into neuronal lineages, one shRNA targets the Klf5 mRNA resulted in a very strong decrease of EGFP expression (Fig.9). Compared to Non Silencing shRNA (CRL shRNA) transfected cells, showing a normal expression of the reporter gene even after 10 days of differentiation, downregulation of KLF5 resulted in dramatic decrease of GFP signal, already at day 7, demonstrating a severe impact on normal differentiation process.

Klf5, also called intestinal-enriched factor, IKLF, or basic transcription element binding protein 2, BTEB2, is a Zn-finger transcription factor that in the adult is expressed in the proliferating crypt cells of the intestinal epithelium and at low levels in the testis, uterus, placenta, lung and in the proliferating basal layer of the epidermis (Ohnishi et al., 2000). The knockout of *Klf5* causes early embryonic lethality (Shindo et al., 2002), thus indicating a key role of this factor during early development. Together with these data, our observation that Klf5 knockdown prevents normal ESC differentiation was compatible with the possibility that this factor is involved in the control of ESC pluripotency.

3.3 KLF5 IS EXPRESSED IN ESCs AND IN EARLY STAGES OF MOUSE EMBRYO DEVELOPMENT

Klf5 expression pattern was analyzed at various times after the induction of ESC differentiation. Western blot analysis and RT PCRs demonstrated a dramatic decrease of its mRNA and protein and a similar trend was observed when ESCs were differentiated through a different procedure, involving the generation of Embryoid Bodies (Fig. 10, A). Immunostaining of ESCs shows that Klf5 is expressed at various levels almost ubiquitously in undifferentiated ESCs and it co-localizes with Oct4 and Nanog. After 3 days of differentiation, Klf5 signal disappeared from most cells, similarly to that observed for Oct3/4 and Nanog (Fig. 10, B).

The expression of Klf5 mRNA *in vivo* is in agreement with that observed during *in vitro* differentiation of ESCs. In fact, we found Klf5 transcript in the blastocysts at day E3.5 but not in the E6.5 embryos (Fig.11, A). Immunostaining of pre-implantantion embryos showed that Klf5 is present in the nuclei of many cells of the Theiler Stage 4, when the blastocoelic cavity is formed but the layer of trophectoderm cells is still absent. At this stage all the cells expressed Oct4. Similarly, most of the cells expressed both Klf5 and Nanog (Fig.11, B).

Thus, *in vitro* as well as *in vivo*, Klf5 seems to be expressed only in pluripotent cells, while it is rapidly downregulated during differentiation commitment, accordingly with expression pattern of two master genes as Oct4 and Nanog.

3.4 KLF5 KNOCKDOWN CAUSES THE PERTURBATION OF ESC STEMNESS

The observation that Klf5 expression was restricted to undifferentiated ESCs and was tightly regulated when differentiation occurs suggested that it might be required for maintaining the ESC undifferentiated state. To study this hypothesis, we explored the effects of Klf5 knockdown in undifferentiated ESCs. To this aim, ESCs were transfected with the previously

used Klf5 shRNA or with a mixture of four siRNAs, all targeting different regions of the Klf5 mRNA to compare Klf5 knockdown cells with cells transfected with Non Silencing control shRNA or siRNA.

While shNS or siNS transfected cells were undistinguishable from untransfected cells, Klf5 knockdown, by shRNA as well as by siRNAs, resulted in a significant change in morphology of the cells, with the appearance of clusters of enlarged flattened cells (Fig12, A).

To confirm that these morphological changes were due to the loss of stemness, alkaline phosphatase activity was assayed, as a marker of undifferentiated ESCs. Although grown in the presence of LIF and serum for 5 days, Klf5 knockdown cells lost their undifferentiated phenotype as demonstrated by the disappearance of alkaline phosphatase staining (Fig.12, B). According to this behaviour, a significant decrease of Oct4, Nanog and Sox2 mRNA levels was observed. The same effect was observed in ESCs transfected with the Klf5 siRNAs (Fig.12, C).

To confirm that the expression of Klf5 is restricted to undifferentiated cells, an shRNA targeting Oct4 was transfected. As discussed in Introduction, misregulation in Oct4 expression lead to differentiation toward extraembryonic cell lineages. As expected, silencing of Oct4 resulted in a >50% decrease of the cognate mRNA and protein and was accompanied by a significant decrease of Sox2 and Nanog mRNA levels. In these conditions, Klf5 mRNA is significantly decreased (Fig.12, D), thus further suggesting that the expression of Klf5 is restricted to ESC undifferentiated state.

These data strongly support the evidence that Klf5 is necessary to preserve stem cell identity, but do not exclude the possibility that downregulation of Klf5 cause a shif in fate commitment.

Therefore, phenotypic changes induced by Klf5 knockdown were explored, analyzing by RT PCR several markers of cell fate. As shown in Figure 13 (A), endoderm (Gata4 and Hnf4),

ectoderm (Sox1) and visceral endoderm (Afp) markers were undetectable in both Klf5 and Non Silencing transfected cells. Instead mRNAs for Brachyury and Meox, markers of mesoderm, and Cdx2 and Eomes, trophoblast markers, were detected only in Klf5 knockdown cells. Brachyury expression suggests the differentiation of part of these cells towards mesoderm lineage, but these cells were unable to generate terminally differentiated mesoderm-derived cells even after several days in differentiation medium, suggesting that silencing of Klf5 causes a general misregulation of gene expression, rather than a shift of cell fate toward mesoderm lineage.

Concerning Cdx2 and Eomes expression, it is consistent with differentiation of Klf5 knockdown cells towards trophoblast, as confirmed by the expression of cytokeratin 8 (Troma1) in large flattened cells and in giant cells, with swollen nuclei and large cytoplasmic spreading, very similar to those observed in Oct4 knockdown cells (Fig.13, B).

3.5 KLF5 CONSTITUTIVE EXPRESSION CONFERS A LIF-INDEPENDENT ABILITY FOR PLURIPOTENCY

Maintaining pluripotency in ESCs requires a tight control of the amount of master genes, as demonstrated for example by the induction of differentiation into extraembryonic phenotypes by both suppression and overexpression of Oct4 (Niwa et al., 2000). Therefore, we explored the effects of Klf5 ectopic expression in ESCs. To this aim, ESCs were transfected with a Klf5 expression vector under the control of the *β-actin* gene promoter. These cells did not show any major difference from those mock transfected.

It is well known that mouse ESCs should be grown *in vitro* in the presence of LIF and serum, which are required to preseve the self-renewing and pluripotent phenotype. In the absence of one or both factors, ESCs loose their pluripotency. So it has been decided to evaluate the behaviour of Klf5 constitutive-expressing ESCs in the absence of LIF.

Accordingly, six days after the removal of LIF, mock-transfected ESCs lost alkaline phosphatase staining. On the contrary, despite LIF withdrawal Klf5-transfected cells showed numerous colonies highly positive for alkaline phosphatase staining (Fig.14, A).

Moreover, real time PCR experiments showed that Klf5 constitutive expression in cells grown without LIF for 6 days prevents the decrease of Nanog mRNA observed in mock-transfected cells (Fig.14,B), while Oct4 expression was not affected.

These data are very interesting, because, by now, the only transcription factor able to prevent ESC differentiation after LIF withdrawal, when overexpressed, was Nanog (see Introduction). Maybe Klf5 overexpression leads to a forced expression of Nanog which, in turn, preserve ESC pluripotency even without LIF. If it is true, Klf5 should directly regulate Nanog expression.

A number of observations supports the hypothesis that a transcription factor network is responsible for continuous ESC self-renewal (see Introduction). This network consists in both positive and negative feedback loops, where a crucial role is played by Oct4 and Nanog. The effects of Klf5 overexpression or silencing on the expression levels of these genes raise the possibility that this transcription factor has a direct role in this regulatory network.

3.6 KLF5 DIRECTLY REGULATES NANOG AND OCT4 TRANSCRIPTION

Considering that Klf5 overexpression sustains the ESC undifferentiated phenotype even in the absence of LIF and prevents in these conditions the decline of Nanog expression, the possibility that Klf5 may be involved in the regulation of *Nanog* and/or *Oct4* transcription was analyzed. To this aim, Klf5 was co-transfected with reporter vectors driving the expression of luciferase under the control of *Nanog* or *Oct4* gene promoters (Fig.15). First it has been addressed whether Klf5 was able to activate transcription from *Oct3/4* and *Nanog* promoters in a cellular environment where the transcription factor network responsible for

ESC pluripotency is turned off. As shown in Fig.15, in NIH3T3 *Oct4* and *Nanog* promoters drove low levels of transcription, but Klf5 overexpression induced a strong increase of *Nanog* promoter activity, thus suggesting a direct effect of Klf5 on this promoter. A lower, but still significant, increase was also detected for *Oct4* promoter.

As expected, also in ESCs Klf5 overexpression was accompanied by a significant induction of *Nanog*, while again *Oct4* promoter activity appeared weaker, but still detectable.

Next, *Nanog* and *Oct4* promoter expression was monitored in Klf5 knockdown cells. To this aim, a pool of three siRNAs targeting different regions of mouse Klf5 mRNA was transfected in ESCs. Luciferase assay showed that the transcription from both *Oct4* and *Nanog* promoters was reduced by about 50%, thus confirming the dependency of these two genes on Klf5.

In order to address whether Klf5 is regulating these genes by direct interaction with their promoters, we performed ChIP experiments (Fig. 16), transfecting a 3xFlag-tagged full-lenght Klf5 in ESCs in order to immunoprecipitate chromatin samples using the anti-Flag antibody, because the antibody used for western blot and immunostaining assays doesn't work for ChIPs.

The immunoprecipitated chromatin was analyzed in real time PCR and the results demonstrated the direct interaction of Klf5 with *Nanog* promoter. In the case of the *Oct4* promoter, oligonucleotide pairs targeting control regions CR4 and CR2, but not that targeting CR1, showed a modest enrichment of chromatin coimmunoprecipitated with Klf5. These results demonstrate the direct interaction of Klf5 with these two gene promoters resulting in a positive regulation.

Systematic analysis of Oct4 and Nanog target genes led to the identification of numerous candidate genes, including Klf5 (see supplementary information in Boyer et al., 2005 and in Loh et al., 2006). Thus we analyzed the mouse Klf5 genomic region and found that it contains a Nanog candidate cis-element. Alignment of this region with all available genomic

sequences, let us to observe a 100% conservation of these cis-elements in Klf5 orthologs in all mammalian species and also in the chicken ortholog (Fig. 17, A). On this basis ChIP experiments were designed to explore the possible direct interaction of Nanog with *Klf5* promoter. Figure 17 (B) shows a significant enrichment for Nanog binding to Klf5 chromatin. These results indicate that Klf5 is part of the transcription factor network which regulates ESC pluripotent phenotype.

4. DISCUSSION AND FUTURE AIMS

This work reports on the generation of an *in vitro* system for the systematic screening of genes involved in regulation of ESC pluripotency and differentiation.

By using this approach, I was able to show for the first time that Klf5 is an essential factor of the core regulatory network responsible for maintaining ESC pluripotency. Its regulation and function are closely related to those of Nanog and Oct4. In particular a tight functional interaction appears between Klf5 and Nanog. Both factors, when overexpressed, are able to sustain ESC undifferentiated state in the absence of LIF, and the phenotype induced by downregulation of Klf5 seems quite similar to those obtained by Nanog depletion (Hatano et al. 2005).

Furthermore, this observation suggests the hypothesis that Klf5 is necessary for *Nanog* expression, as indicated by the fact that Nanog promoter seems to be particulary sensitive to Klf5.

On the other hand, many questions remain unsolved. First of all, it will be necessary to deepin the role of Klf5 in regulation of *Nanog* expression. To this aim, we are programming band shift experiments in order to identify the binding site of Klf5 on *Nanog* promoter. As a member of the Sp1-like/Kruppel-like family, Klf5 recognize and bind GC-rich elements, GT-type elements and CACCC boxes by its C-terminal DNA-binding domain, made by three zinc-fingers (Wolfe et al. 1999).

A sequence analysis of the last 380bps in the 5' flanking region of the mouse *Nanog* promoter shows that there are at least three GC boxes that has been demonstrated to interact with Sp1 (Da Yong et al.2006). It could be possible that these sites represent a possible target of Klf5.

It has been demonstrated that the zinc-finger domain of the KLFs can also function as protein-protein interaction domain for the interaction with other transcription factors (Song et al. 2002; Zhang et al. 2001)

In 2005 Kuroda et al. compared by EMSA assay nuclear extracts obtained from different pluripotent stem cells and found that in Embryonic Stem cells the Octamer and Sox element on the *Nanog* promoter (see Fig. 15, B) was preferentially bound by Oct4 in complex with an undefined factor that they named Pluripotential-cell specific Sox element-Binding Protein (PSBP). They suggested that *Nanog* expression could be regulated through an interaction between Oct4 and PSBP, which should be prominent in ESCs rather than Sox2.

We are interested in identify this unknown factor, that could be Klf5 itself. Moreover, our ChIP experiments shows localization of Klf5 putative binding sites on *Nanog* and *Oct4* promoters near Octamer elements. So the possibility that Klf5 coincides with PSBP is not so weaky.

Another point to deepin is the regulation of *Klf5* gene expression. As demonstrated by ChIP experiments, *Klf5* is target of Nanog, and our data are in agreement with the ChIP-PET assay results produced by Lho and colleagues that demonstrated a direct interaction between Nanog and *Klf5* promoter. Through *in silico* analysis of *Klf5* promoter sequences of different species I have found a putative Nanog binding site at position – 2600bp from the start site. ChIP assays yet demonstrate that Nanog binds *Klf5* promoter in a region containing also this sequence, and we have in program the generation of reporter plasmids in which the luciferase gene is under the control of *Klf5* promoter with or without the Nanog binding site. This approach should be usefull to identify a possible feedback loop between these factors. Besides

luciferase experiments, band shift assays will be performed to identify the precise Nanog binding site.

Besides the molecular mechanism, we would like to further explore the biological role of Klf5 in ESCs. Individual members of the Sp1-like/KLF family, in spite of similar binding sequences, can function as activators or repressor depending on the cellular context or which promoter they bind and the coregulators with which they interact. The specificity of their activities is determined by different amino termini and/or by tissue-specific expression. It has been demonstrated that some KLFs are simultaneusly expressed in the same cellular enviroment and that they act in combinatorial manner on the same target genes, providing an efficent system to control gene expression.

Klf5 was firstly described by Lingrel and Co. (Conkright et al. 1999) in intestinal epithelium, where it acts in combination with another KLF named Klf4. Klf4 is expressed in terminally differentiated epithelial cells at the lumenal surface of the intestinal mucosa, whereas Klf5 is expressed in actively dividing cells at the base of the intestinal crypts. This contrast between the two proteins carries over into their transcriptional activities, as KLF4 and KLF5 often exhibit opposing effects on shared transcriptional targets and carry out distinctive biological activities (Ghaleb et al.2005).

In recent years, fusion experiments with mouse ESCs have shown the dominance of the ESC phenotype over that of somatic cells, implying that proteins in the nucleus of ESCs are able to reprogram more differentiated cells to an embryo-like state (Cowan et al., 2005). In ESC-NSC (Neural Stem Cells) fusion, the frequency of reprogramming was markedly enhanced by modestly increasing of Nanog expression in ESCs, that is consistent with the view that Nanog is a major driver of pluripotency (Rao and Orkin et al., 2006) (Silva et al.2006). Recently, further studies found that forcing the expression of ESC specific genes, particulary transcription factors Oct4, Sox2, c-myc and Klf4, in somatic cells might induce them to take

on a more embryonic character (Takahashi and Yamanaka, 2006). Unexpectedly, Nanog was dispensable. However, c-Myc and Klf4 are essential in this experiment. The c-Myc may be associated with histone acetyltransferase (HAT) complex, and induces global histone acetilation, thus allowing Oct4 and Sox2 to bind to their specific target loci. The autors suggest that Klf4 might contribute to activation of Nanog and other ESC specific genes through p53 repression (Takahashi and Yamanaka, 2006).

Data not shown in this manuscript provides evidence that Klf4 is not involved in the transcription network of Klf5 in ESCs, in fact Klf4 mRNA remains unmodified in Oct4 knockdown and in Klf5 knockdown or overexpressing cells. Moreover, changing in Klf4 expression levels does not significantly modify the levels of Oct4, Klf5, Sox2 or Nanog mRNAs.

Because Klf4 and Klf5 are strictly correlated and are able to bind the same target sequences, and considering the data produced in this project, we want to investigate the possible role of Klf5 in cell reprogramming as real partner of Oct4, Sox2 and c-Myc in this process. To this aim we will reproduce the fibroblast reprogramming replacing Klf4 with Klf5.

5. REFERENCES

Ambros V.2004. The functions of animal microRNAs. *Nature* **431**: 350-355.

Barberi T, Klivenyi P, Calingasan, NY, Lee H, Kawamata H, Loonam K, Terrier AL, Bruses J, Rubio ME, Topf N. 2003. Neural subtype specification of fertilization and nuclear transfer embryonic stem cells and application in Parkinsonian mice. *Nat. Biotechnol***. 21**: 1200-1207.

Beddington RS, Robertson EJ. 1989. An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo. *Development* **105**: 733-737.

Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, Giffored DK, Melton DA, Jaenisch R, Young RA. 2005. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**: 947-956.

Bradley A, Evans MJ, Kaufman MH, Robertson E. 1984. Formation of germline chimaeras from embryo-derived teratocarcinoma cell lines. *Nature* **309**: 255-256.

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

Cartwright P, McLean C, Sheppard A, Rivett D, Jones K, Dal ton S. 2005 LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism. *Development* **132**: 885-896.

Chambers I, Colby D, Robertson M, Nichols J, Lee S, Tweedie S, Smith A. 2003. Functional expressing cloning Nanog, a pluripotency sustaining factos in embryonic stem cells. *Cell* **113**: 643-655.

Chang K, Elledge SJ, Hannog GJ. 2006. Lesson from Nature: microRNA-based shRNA libraries. *Nature Methods* **9**: 707-714.

Chew JL, Loh YH, Zhang W, Chen X, Tam WL, Yeap LS, Li P, Ang YS, Lim B, Robson P, Ng HH. 2005. Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Mol. Cell Biol*. **25**: 6031-6046.

Conkright MD, Wani MA, Anderson KP, Lingrel JB. 1999. A gene encoding an intestinal-enriched member of the Kruppel-like factor family expressed in intestinal epithelial cells. *Nucl. Acids Res.* **27**:1263–1270.

Cowan CA, Atienza J, Melton DA, Eggan K. 2005. Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. *Science* **309**: 1369-1373.

Da Yong Wu1, Zhen Yao. 2006. Functional analysis of two Sp1/Sp3 binding sites in murine Nanog gene promoter. *Cell Research* **16**: 319-322.

Datano SY, et al. 2005. Pluripotential competence of cells associated with Nanog activity. *Mech.Dev*. **122**: 67-79.

Evans MJ, Kaufman MH. 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**: 154-156.

Ghaleb AM, Nandan MO, Chanchevalap S, Dalton WB, Hisamuddin IM, et al. 2005. Kruppel-like factors 4 and 5: the yin and yang regulators of cellular proliferation. *Cell Res* **15**: 92–96.

Hammond SM, Boettcher S, Caudy AA, Kobayashi R, Hannon GJ. 2000. Argonaute2, a link between genetic and biochemical analyses of iRNA. *Science* **293**: 1146-1150.

Han J et al. 2006. Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* **125:** 887-901.

Hao J, Li TG, Qi X, Zhao DF, Zhao GQ. 2006. Wnt/beta-catenin pathway up-regulates Stat3 and converges on LIF to prevent differentiation of mouse embryonic stem cells. *Dev. Biol.* **290**: 81-91.

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

Kaczynski J, Cook T, Urrutia R. 2003. Sp1- and Krüppel-like transcription factors. *Genome Biology* **4**:206.

Keller G. 1995. In vitro differentiation of embryonic stem cells. *Curr. Opin. Cell Biol.* **7**: 862-869.

Kuroda T, Tada M, Kubota H, Rimura H, Datano SY, Smemori H, Nakatsuji N, Tada T. 2005. Octamer and sox elements are required for transcriptional cis regulation of Nanog gene expression. *Mol. Cell Biol*. **25**: 2475-2485.

Liu J, et al. 2004. Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305:** 1437-1441.

Loh YH, Wu Q, Chew JL, Vega VB, Zhang W, Chen X, Bourque G, Gorge J, Leong B, Liu J, Wong KY, Sung KW, Lee CW, Zhao XD, ChiuKP, Lipovich L, Kuznetsov VA, et al. 2006. The Oct4 and Nanog transcription network regulates plurypotency in mouse embryonic stem cells. *Nat. Genet*. **38**: 431-440.

Maltsev VA, Rohwedel J, Hescheler J, Wobus AM. 1993. Embryonic stem cells differentiate in vitro into cardiomyocites representing sinusoidal, atrial and ventricular cell types. *Mech. Dev*. **44**: 41-50.

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

Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Muratami M, Takahashi K, Maruyama M, Maeda M, Yamanaka S. 2003 The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**: 631-642.

Nakano T, Kodama H, Honjo T. 1994. Generation of lynphohematopoietic cells from embryonic stem cells in culture. *Science* **265**: 1098-1101.

Nichols J et al. 1998. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95**: 379-391.

Nishikawa S, Hirashima M, Matsuyoshi N, Kodama H. 1998. Progressive lineage analysis by cell sorting and culture identifies FLK+VE-caderin cells at a diverging point of endothelialand hemopoietic lineages. *Development* **125**: 1747-1757.

Niwa H. 2001. Molecular mechanism to maintain stem cell renewal of ES cells. *Cell Struct. Funct.* **26**: 137-148.

Niwa H, Burdon T, Chambers I, Smith A. 1998. Self-renewal of pluripotent embryonic stem cellsis mediated via activation of STAT3. *Genes Dev.* **12**: 2048-2060.

Nordhoff V, Hubner K, Bauer A, Orlova I, Malapetsa A, Scholer HR. 2001. Comparative analysis of human, bovine, and murine *Oct-4* upstream promoter sequences. *Mammalian Genome* **12**: 309–317

Ohnishi S, Laub F, Matsumoto N, Asaka M, Ramirez F, Yoshida T, Terada M. 2000. Developmental expression of the mouse gene coding for the Kruppel-like transcription factor KLF5. *Dev. Dyn.* **217**: 421-429.

Okabe S, Forsberg-Nilsson K, Spiro AC, Segal M, McKay RD. 1996. Development of neuronal precursor cells and functional postmitotic neurons from embryonic stem cells in vitro. *Mech. Dev.* **59**: 89-102.

Pan GJ, Chang ZY, Scholer HR, Pei D. 2002. Stem cell pluripotency and transcription factor Oct4. *Cell Res*. **12**: 321-329.

Pan G, Li J, Zhou Y, Zheng H, Pei D. 2006. A negative feedback loop of transcription factors that controls stem cell pluripotency and self-renewal. *FASEB J*. **20**: 1730-1732.

Qi X, Li TG, Hao J, Hu J, Wang J, Simmons H, Miura S, Missina Y, Zhao GQ. 2004. BMP4 support self-renewal of embryonic stem cells by inhibiting mitogen-activated protein kinase pathways. *Proc. Natl. Acad. USA* **101**: 6027-6032.

Rao S. and Orkin SH. 2006. Unraveling the transcriptional network controllino ES cells pluripotency. *Genome Biol.* **7**:230.

Rodda DJ, Chew JL, Lim LH, Loh YH, Wang B, Ng HH, Robson P. 2005. Transcriptional regulation of Nanog by Oct4 and Sox2. *J. Biol. Chem.* **280**: 24731-24737.

Sato N, Meijer L, Skaltsounis L, Greengard P, Brivenlou AH. 2004. Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by pharmacological GSK-3-specific inhibitor. *Nat. Med* **10**: 55-63.

Schmandt T, et al. 2005. High-purity lineare selection of embryonic stem cell-derived neurons. *Stem Cells Dev.* **14**: 55-64.

Sekkai D, Gruel G, Herry M, Moucadel V, Constantinescu SN, Albagli O, Tronik-Le Roux D, Vainchenker W, Bennaceur-Griscelli A. 2005. Microarray analysis of LIF/Stat3 transcriptional targets in embryonic stem cells. *Stem Cell* **23**: 1634-16-42.

Shi W, Whang H, Pan G, Geng Y, Guo Y, Pei D. 2006. Regulation of the pluripotency marker rex-1 by Nanog and Sox2. *J. Biol. Chem*. **281:** 23319-23325.

Shindo T, Manabe I, Fukushima Y, Tobe K, Aizawa K, Miyamoto S,nKawai-Kowase K, Moriyama N, Imai Y, Kawakami H, Nishimatsu H, Ishikawa T, Suzuki T, Morita H, Maemura K, Sata M, Hirata Y, Komukai M, Kagechika H, Kadowaki T, M Kurabayashi, Nagai R. 2002. Kruppel-like zinc-finger transcription factor KLF5/BTEB2 is a target for angiotensin II signaling and an essential regulator of cardiovascular remodeling. *Nat. Med.* **8**: 856-863.

Silva J., Chambers I, Pollard S., Smith A. 2006. Nanog promotes transfer of pluripotency after cell fusion. *Nature* **441**: 997-1001.

Siolas D, et al. 2005. Syntetic shRNAs as potent RNAi triggers. *Nat. Biotechnol.* **23**: 227-231.

Smith AG. 2001. Embryo-derived stem cells: Of mice and men. *Annu. Rev. Cell Dev. Biol.* **17**: 435-462.

Smith AG, Health JK, Donaldson DD, Wong GG, Moreau J, Stahl M, Rogers D. 1988. Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature* **336**: 688-690.

Song CZ, Keller K, Murata K, Asano H, Stamatoyannopoulos G. 2002. Functional interaction between coactivators CBP/p300, PCAF, and transcription factor FKLF2. *J Biol Chem.* **277**: 7029-7036.

Suzuki A, Raya A, Kawatami Y, Morita M, Matsui T, Nakashima K, Gage FH, Rodriguez-Esteban C, Izpisua Belmonte JC. 2006. Nanog binds to Smad1 and blocks bone morphogenic protein-induced differentiation of embryonic stem cells. *Proc. Nat. Acad. USA* **103**: 10294-10299.

Takahashi K, Yamanaka S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**: 663-676.

Tanaka TS, Kunath T, Kimber WL, Jaradat SA, Stagg CA, Usuda M, Yokota T, Niwa H, Rossant J, Ko MS. 2002. Gene expression profiling of embryo-derived stem cells reveals candidate genes associated with pluripotency and lineage specificity. *Genome Res.* **12**: 1921-1928.

Wolfe SA, Nekludova L, Pabo CO. 2000. DNA recognition by Cys2His2 zinc finger proteins. *Annu Rev Biophys Biomol Struct.* **29**: 183-212.

Ying QL, Nichols J, Chambers I, 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.

Zamore PD, Tuschl T, Sharp PA, Bartel DP. 2000. RNAi: double-stranded RNA directs the ATP-dependente cleavage of mRNA at 21-23 nucleotide intervals. *Cell* **101**:25-23.

Zhang W, Kadam S, Emerson BM, Bieker JJ. 2001. Site-specific acetylation by p300 or CREB binding protein regulates erythroid Kruppel-like factor transcriptional activity via its interaction with the SWI-SNF complex. *Mol Cell Biol* . **21**: 2413-2422.

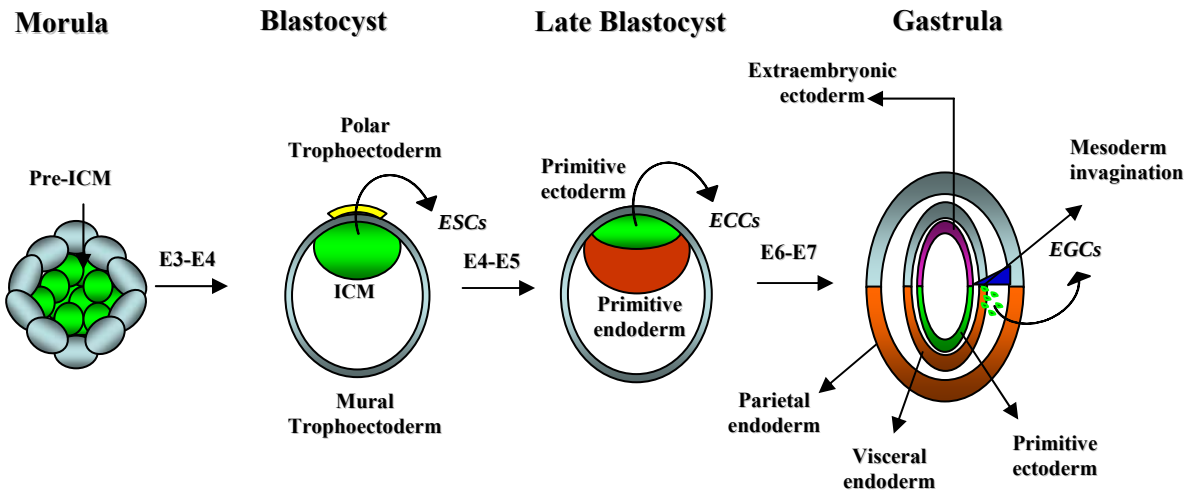


Figure 1| Scheme of early mouse development
Schematic representation of embryonic stem cell's origin in the mouse embryo. The morula-stage embryo (E2.5) holds a primordial ICM (Inner Cell Mass) population, that turn into ICM cells at cavitation/blastulation (E3-E4). At this stage, ESCs (Embrionyc Stem Cells) can be derived *in vitro* and implantation occurs *in vivo*. As the blastocyst undergoes implantation *in vivo*, the ICM give rise to a primitive endoderm and a primitive ectoderm, also known as epiblast, from which a pluripotent cell lines called ECCs (Embryonal Carcinoma Cells) can be derived. At later stages, the capability to derive ESCs and ECCs is progressively lost and the embryo start gastrulation (E6). This process involves the generation of a mesoderm layer between ectoderm and endoderm, and at this stage another pluripotent cell population, known as EGCs (Embryonic Germ Cells), can be derived.

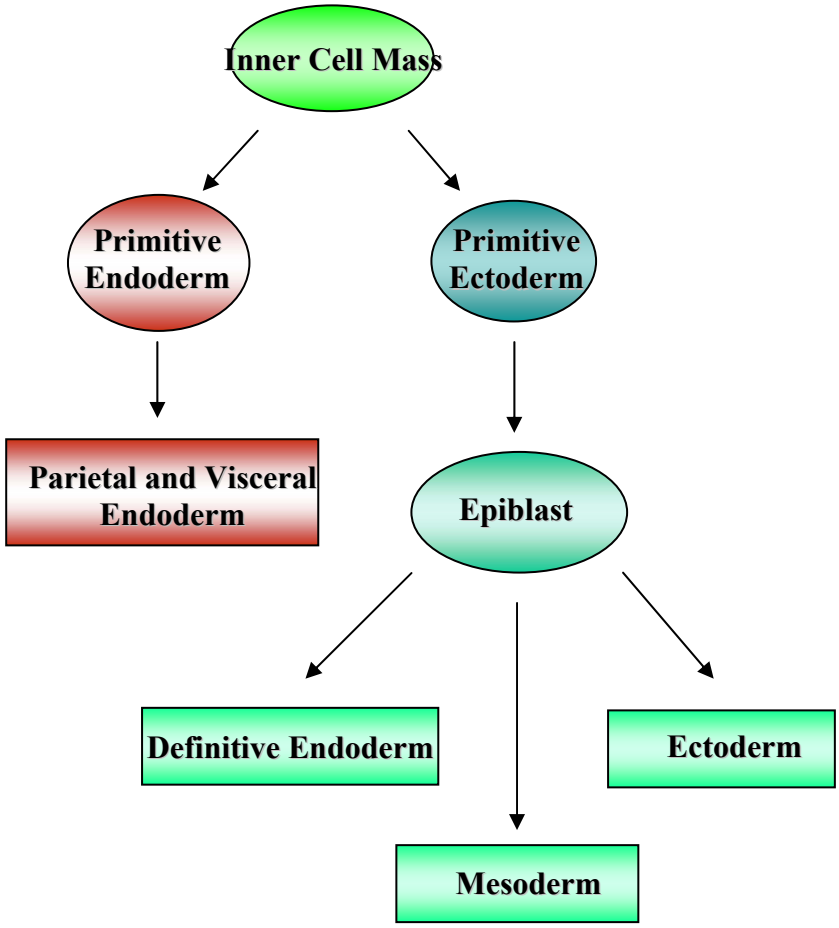


Figure 2| Relationship of early cell population to the primary germ layers
Starting from the ICM, a population of pluripotent cells named primitive ectoderm is generated. Shortly after implantation, the innermost cells of the primitive ectoderm cell mass undergo apoptosis and form a cavity. The surviving primitive ectoderm cells that surround the cavity differentiate to form a pseudostratified epithelium, also known as epiblast, that will give rise to the primary germ layers during gastrulation. Epiblast cells that undergo an epithelial to mesenchymal transition will form mesoderm and definitive endoderm, while cells in the most anterior region of the epiblast will form ectoderm.

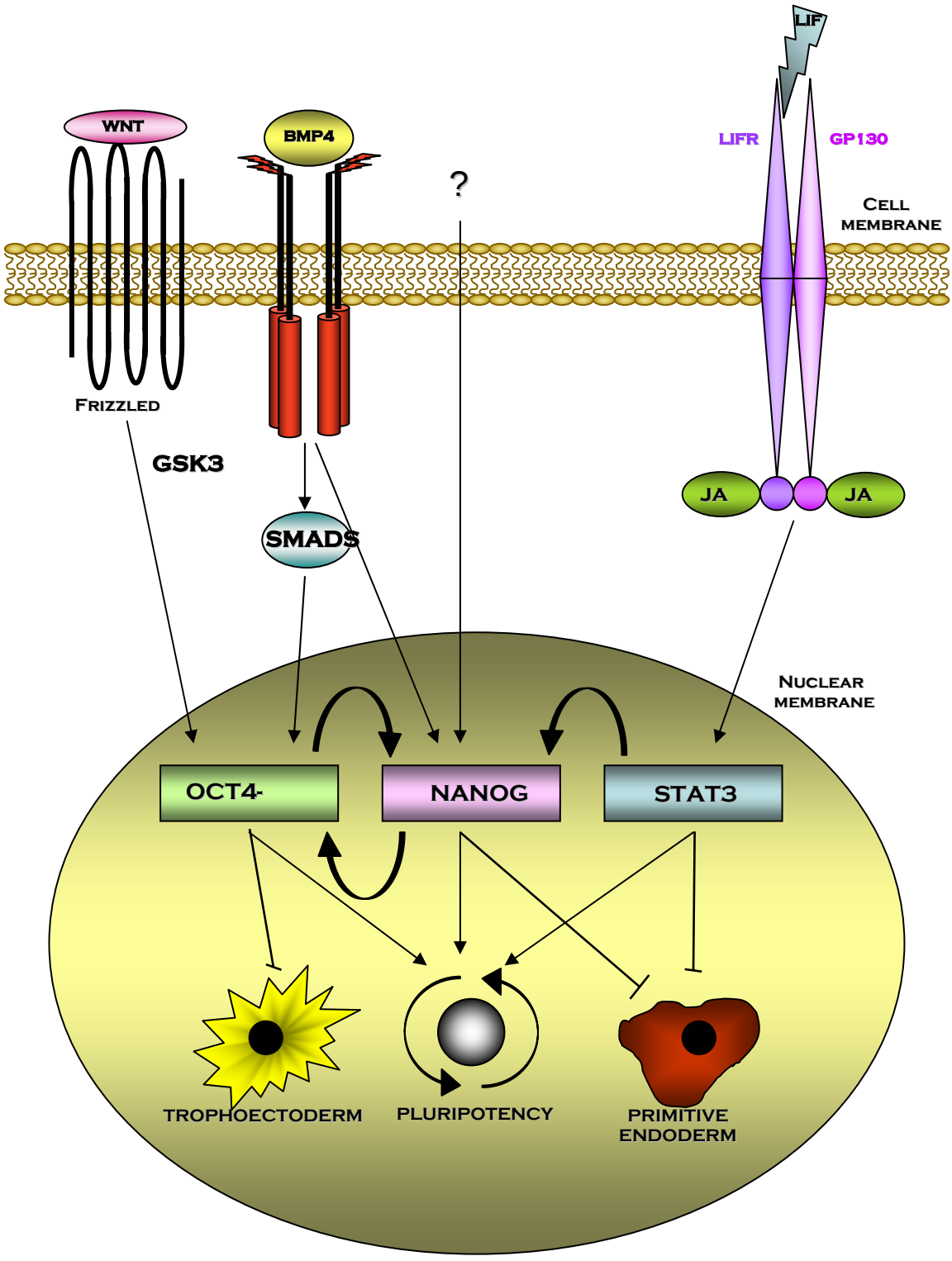

LIF
WNT
BMP4
?
LIFR
GP130
CELL MEMBRANE
FRIZZLED
GSK3
SMADS
JA
JA
NUCLEAR MEMBRANE
OCT4-
NANOG
STAT3
TROPHOECTODERM
PLURIPOTENCY
PRIMITIVE ENDODERM

Figure 3| Signalling pathways involved in self-renewal and pluripotency
Cell-surface receptors initiate signals that are converted to the nucleus and affect key pluripotency transcription factors. By now, only the LIF/STAT pathway has been defined in detail. Binding of LIF to the LIFR at the cell surface causes its heterodimerisation with gp130. The intracellular domain of LIFR-gp130 heterodimer can activate the non-receptor Janus tyrosine kinase (JAK) and the signal transducer and activatior of transcription (STAT). STAT3 binds to phosphotyrosine residues on activated LIFR-gp130 heterodimer and undergoes phosphorilation and dimerization itself. Then it can translocate to the nucleus, where it works as a transcription factor. BMP4 have been shown to phosphorylate Smad1/5, and it results in the expression of inhibitor of differentiation (Id) protein that blocks the neural differentiation. Also the Wnt pathway could be involved in the maintenance of pluripotency. Infact, Wnt signaling is endogenously activated in ESCs and is downregulated upon differentiation.

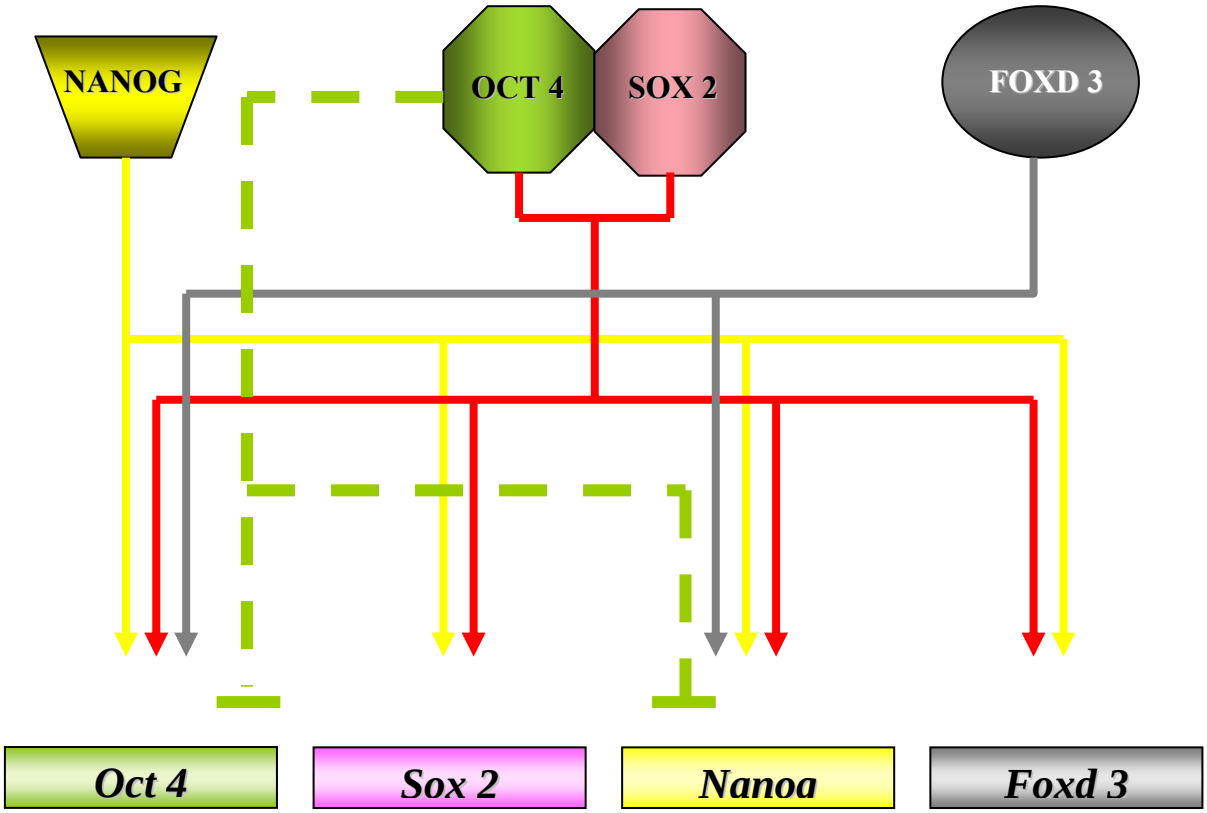
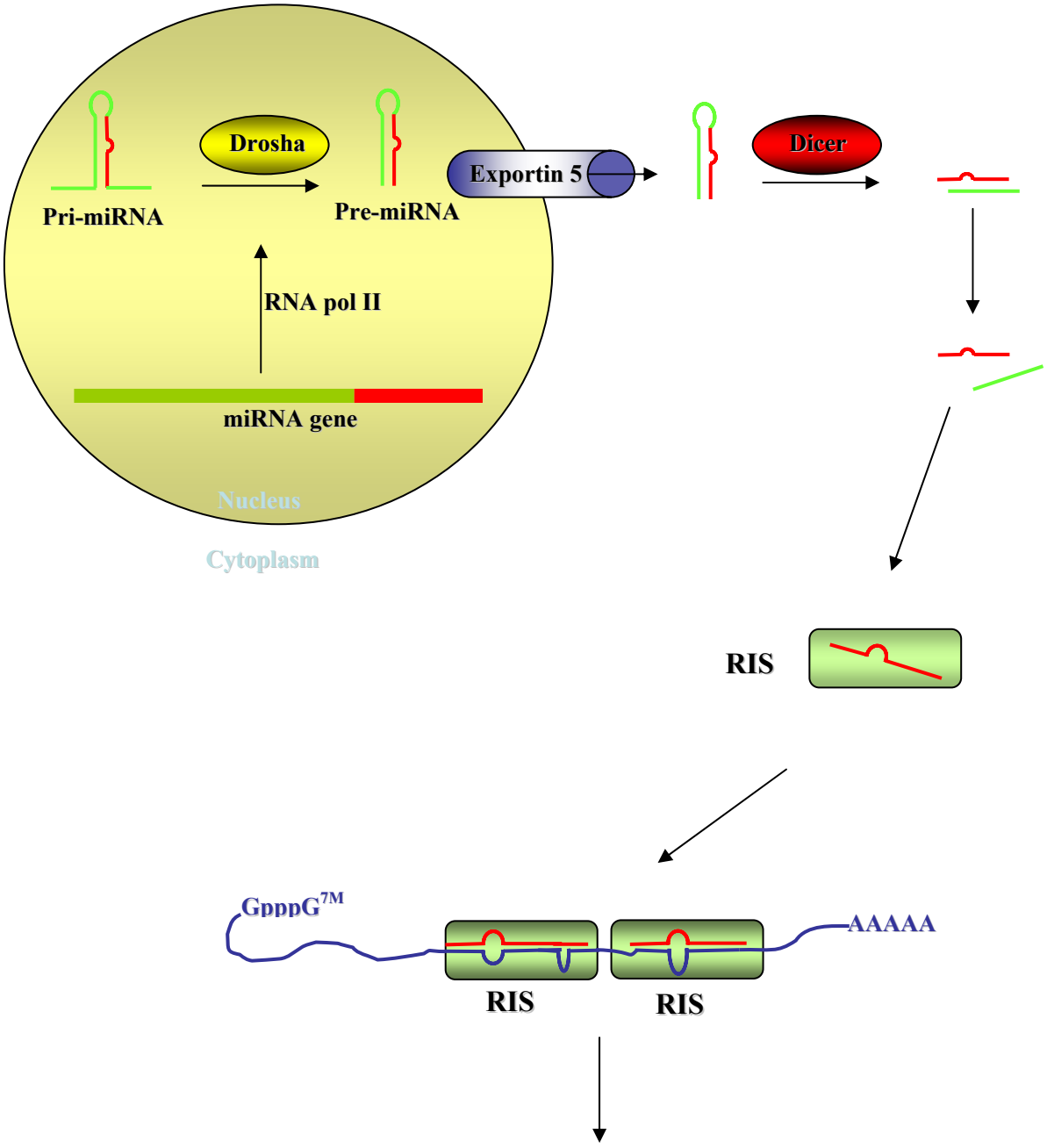


Figure 4| Regulatory network of key transcription factors in maintaining ES cell pluripotency

The levels of key factors that maintain pluripotency need to be strictly controlled to balance the maintenance of undifferentiated state and the ability of lineage commitment.

Regulators such as Oct4, Nanog, Sox2 and FoxD3 bind to each other's promoter, and support or limit each other's expression, forming an interconnected autoregulatory network to maintain ES cell pluripotency and self-renewal. Oct4 maintains Nanog expression by directly binding to the *Nanog* promoter when present at a sub-steady level, but represses it when Oct4 is above the normal level. On the other hand, FoxD3 positively regulates *Nanog* to counter the repression effect of excess Oct4. Conversely, Nanog and FoxD3 function as activators for *Oct4* expression. When the expression level of Oct4 rises above a steady level, it represses its own promoter as well as *Nanog*, thus exerting a negative feedback regulation loop to limit its own expression. Arrows connected to factors by solid lines indicate positive regulation of a promoter by the factors. Broken lines linking to Oct4 indicate negative regulation.



Translational repression and/or mRNA degradation

Figure 5| The biogenesis and mechanism of action of miRNAs
miRNA genes are transcribed by RNA polymerase II to generate the primary miRNA transcripts (pri-miRNA), which are processed into the 70 nt hairpin-structured miRNA precursor (pre-miRNA) by Drosha in the nucleus. After being transported to the cytoplasm by exportin 5, the pre-miRNA is furthed processed by Dicer, a ribonuclease III family (RNaseIII) enzyme, to generate a small duplex. The resulting small RNAs enter the RNA-induced silencing complex (RISC), which uses a single stranded version of the small RNA as a guide to substrate selection. Perfect complementarity between the substrate and the small RNA leads to target-RNA cleavage by RNAse H – like active site within an Argonaute protein that forms the core of RISC.

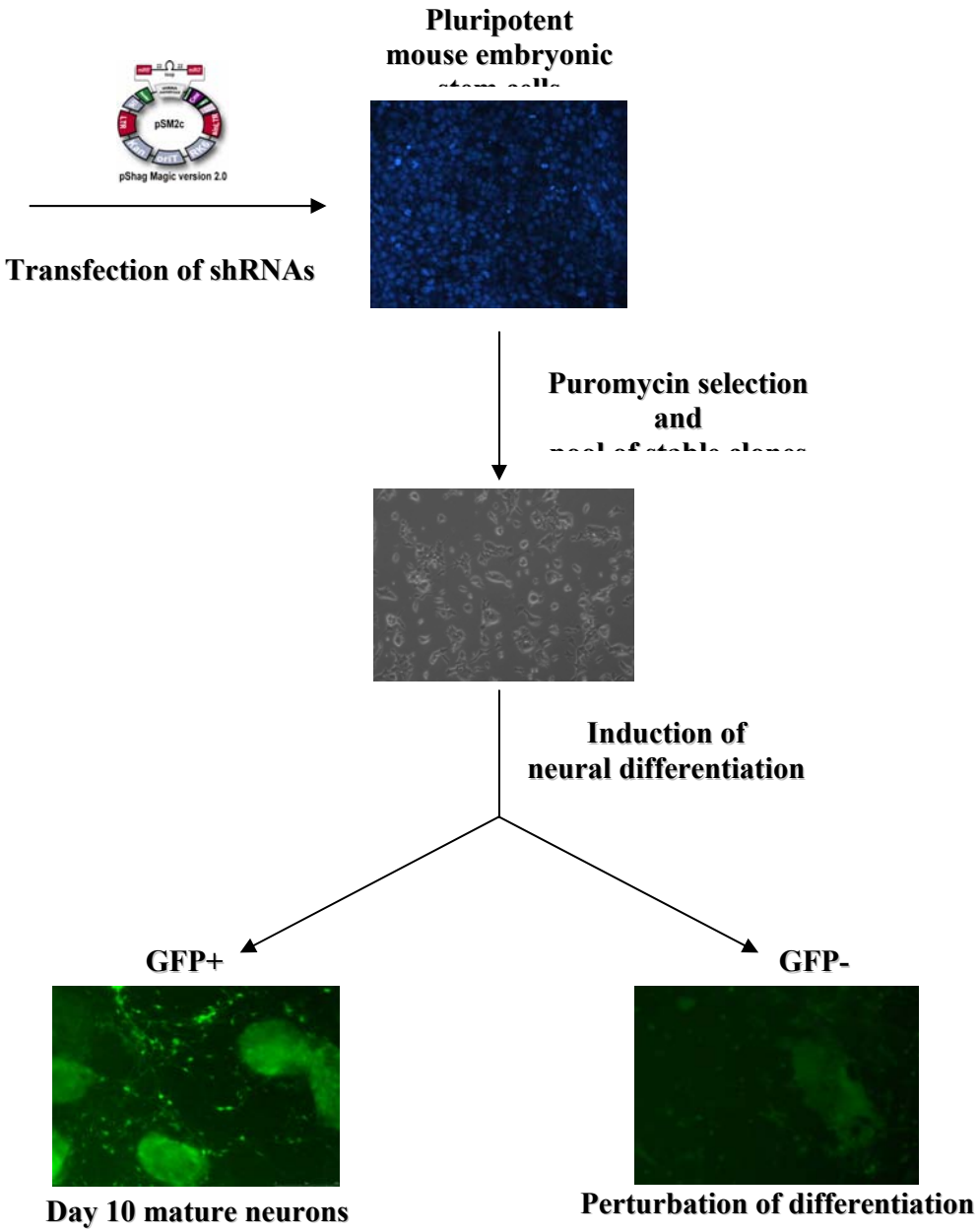


Figure 6| General representation of the high-throughput screening
We focused on identification and characterization of genes that regulate the balance between stemness and differentiation commitment, using a hightroughput screening system based on an shRNA collection designed on the sequence of murine mRNAs. We have generated an ES cell line stably expressing the GFP reporter gene under the control of the neuronal-specific alpha1- tubulin promoter, which directs GFP expression specifically in neuronal precursors and post-mytotic neurons. Cells were transfected with shRNAs and selected in puromycin to obtain stable interferred clones. Resulting clones were pooled together and seeded at low confluence in a chemically defined medium to induce neural differentiation. Genes whose suppression by specific shRNAs modifies neuronal differentiation, will change GFP expression pattern. We are interested in genes that can block normal differentiation showing a decrease or a complete absence of GFP signal.

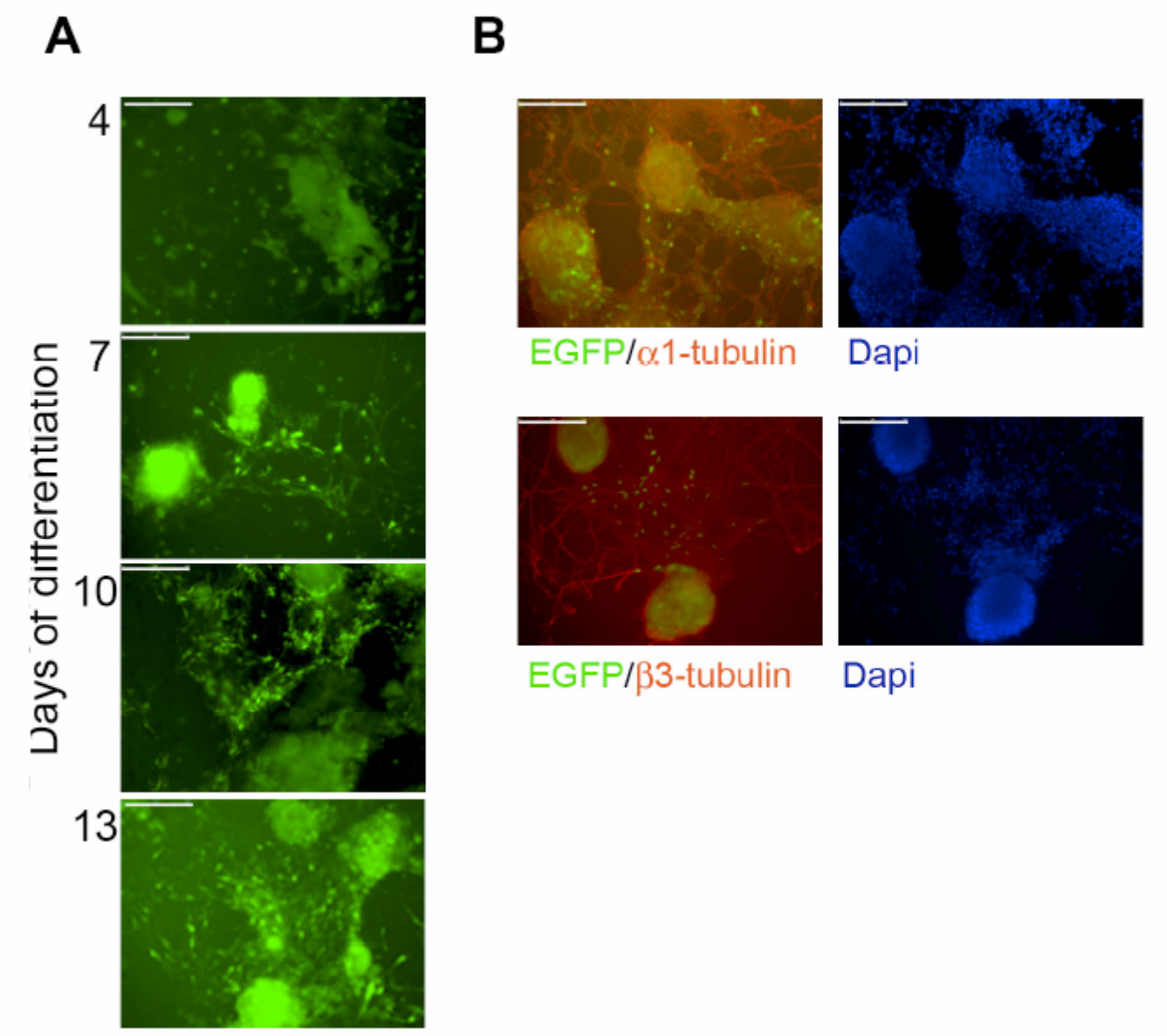


Figure 7| GFP expression pattern of α1-tubulin cells
The GFP reporter signal appears about four days after induction of neural differentiation in chemically defined medium, when ESCs differentiate into neuronal precursor cells. As maturation of neurons occurs, fluorescence becomes stronger and a change in cell morphology is evident. When differentiation is complete, neurons of different cell subpopulations are totally marked by reporter, while astrocytes, oligodendrocytes and ephitelial cells are not (**A**).
Neuron-specific immunostaining with α1-tubulin and β3-tubulin antibodies demonstrates co-expression of EGFP signals with neuronal differentiation markers at day 13 (**B**).

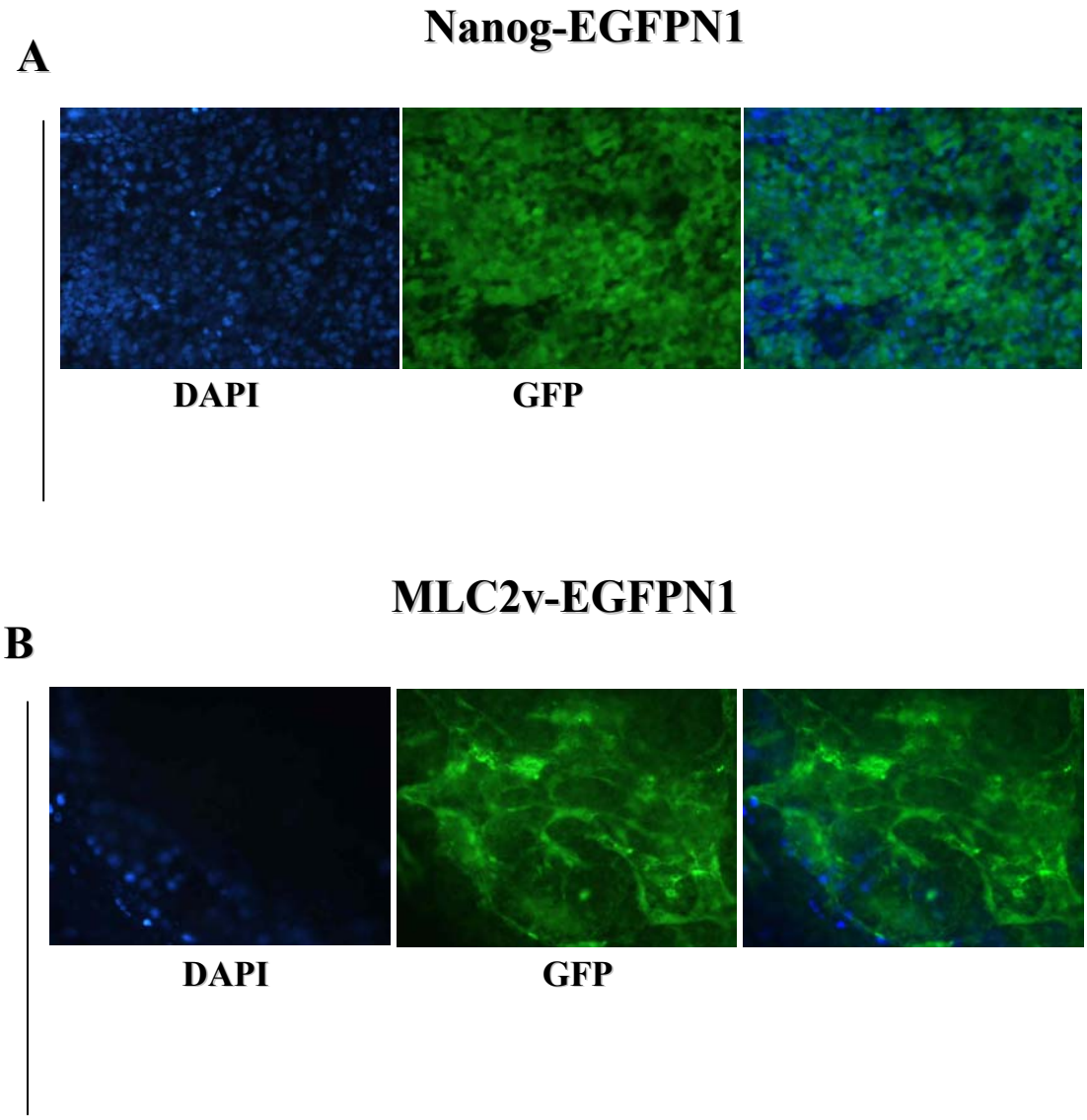


Figure 8| Nanog-EGFPN1 and MLC-2vEGFPN1 cell lines
Nanog-EGFPN1 and MLC-2v EGFPN1 cell lines where used during subsequent characterization of the phenotypes obtained from gene silencing of interesting genes.
Expression of GFP reporter signal driven by *Nanog* promoter is detectable in all undifferentiated ESCs, in presence of LIF and serum. Soon after induction of differentiation, the signal becomes faint untill it desappears completely at day three of neural differentiation (**A**).
MLC-2v promoter upstream the GFP reporter gene drives the fluorescence signal only in mature cardiomyocites, obtained by Embryo Bodies (EBs) formation and in presence of batch selected serum that drives the formation of mesoderm derivates. In this cell system, the GFP is completely absent in undifferentiated cells and in cells differentiating into neurons (**B**).

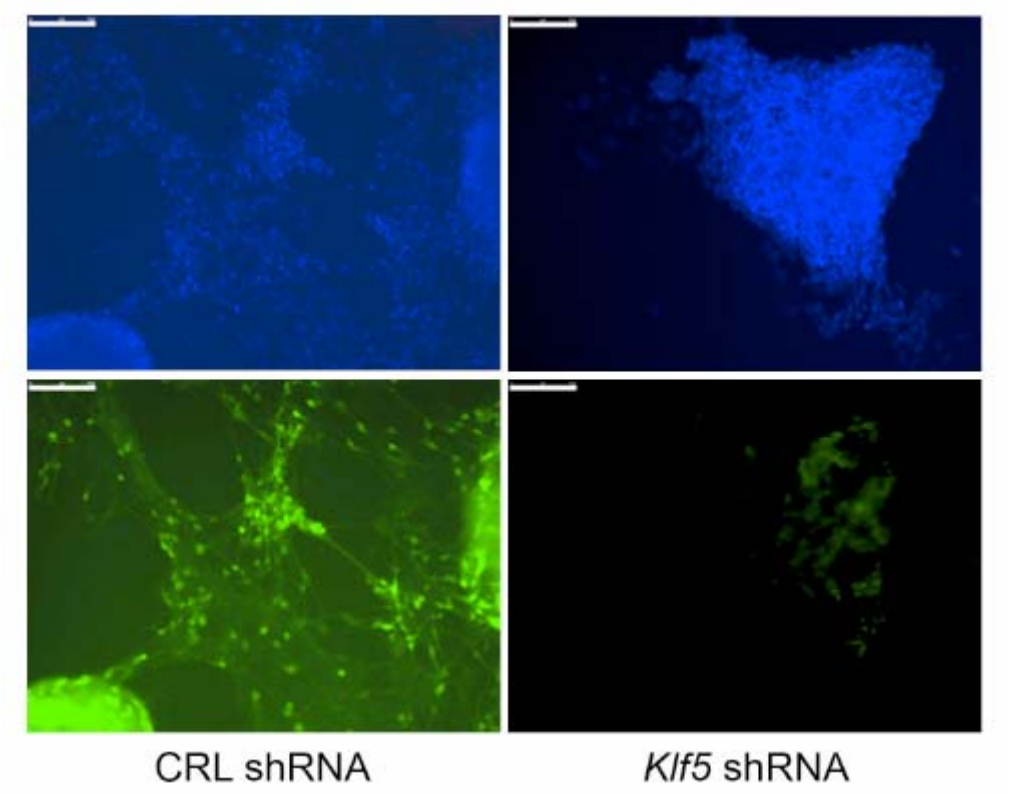


Figure 9| Klf5 knockdown suppress neuronal differentiation
Klf5 knockdown by RNA interference suppresses differentiation of EGFP-tubulin ESCs. *Klf5* knockdown by RNA interference results in the abolishment of ESC differentiation into neurons. E14Tg2a cell clones stably expressing EGFP under the control of the α1-tubulin neuron-specific promoter were used to screen a collection of shRNAs. The *Klf5*-targeting shRNA prevents the appearance of EGFP-positive cells after the induction of ES cell differentiation. Upper panels: DAPI staining; lower panels: EGFP fluorescence observed 7 days after the induction of differentiation. Scale bar: 100 μm. *Klf5* silencing results in the decrease of the protein down to 20% of the basal levels.

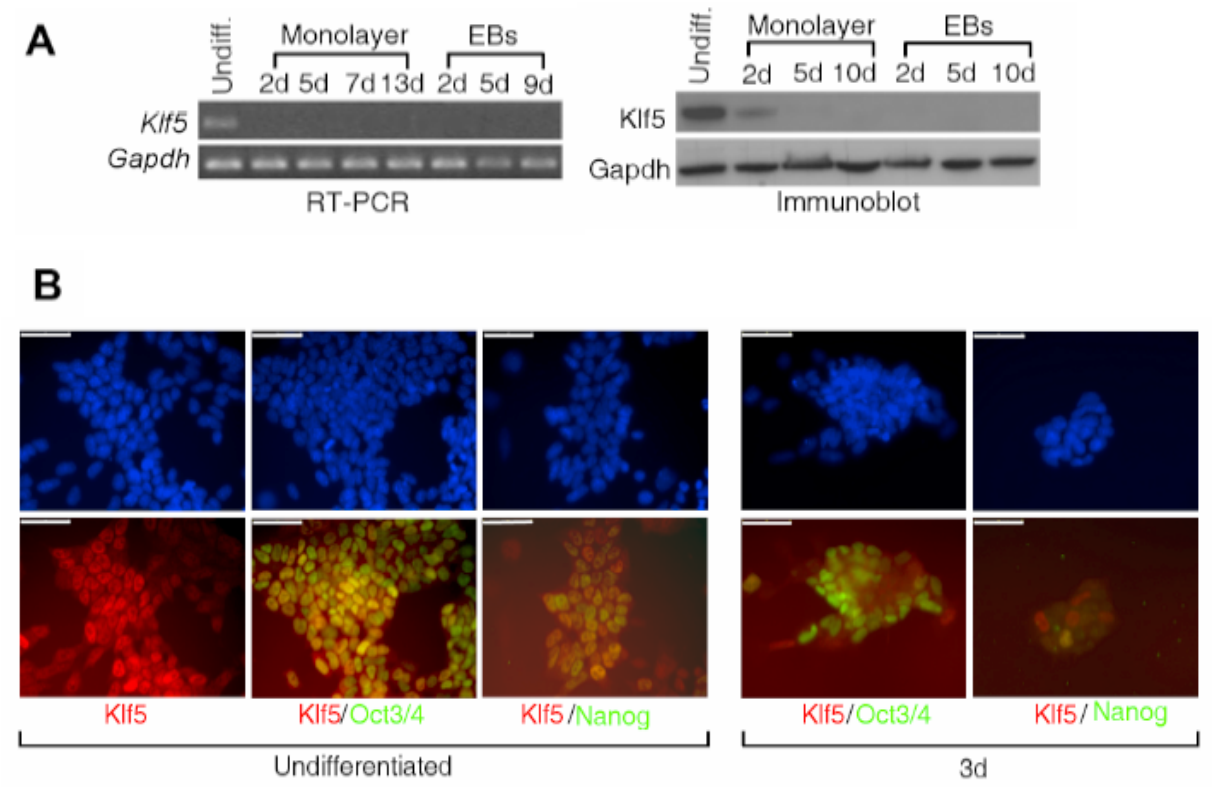


Figure 10| Klf5 expression in ESCs

We analyze Klf5 expression pattern at various times after the induction of ESCs differentiation. Klf5 mRNA and protein levels decrease soon after the induction of ESC differentiation. ESCs were induced to differentiate through two approaches, i.e. allowing them to form Embryoid Bodies (EBs) or in monolayer in serum-free medium (**A**).

Klf5 is present in the nuclei of most of the undifferentiated ESCs and co-localizes with Oct3/4 and Nanog. The Klf5 signal was variable from cell to cell. Three days after the induction of differentiation (3d), Klf5 levels decrease together with those of Oct3/4 and Nanog (**B**). Upper panels: DAPI staining; lower panel: Klf5, red; Oct3/4 or Nanog, green. Scale bar: 50 µm.

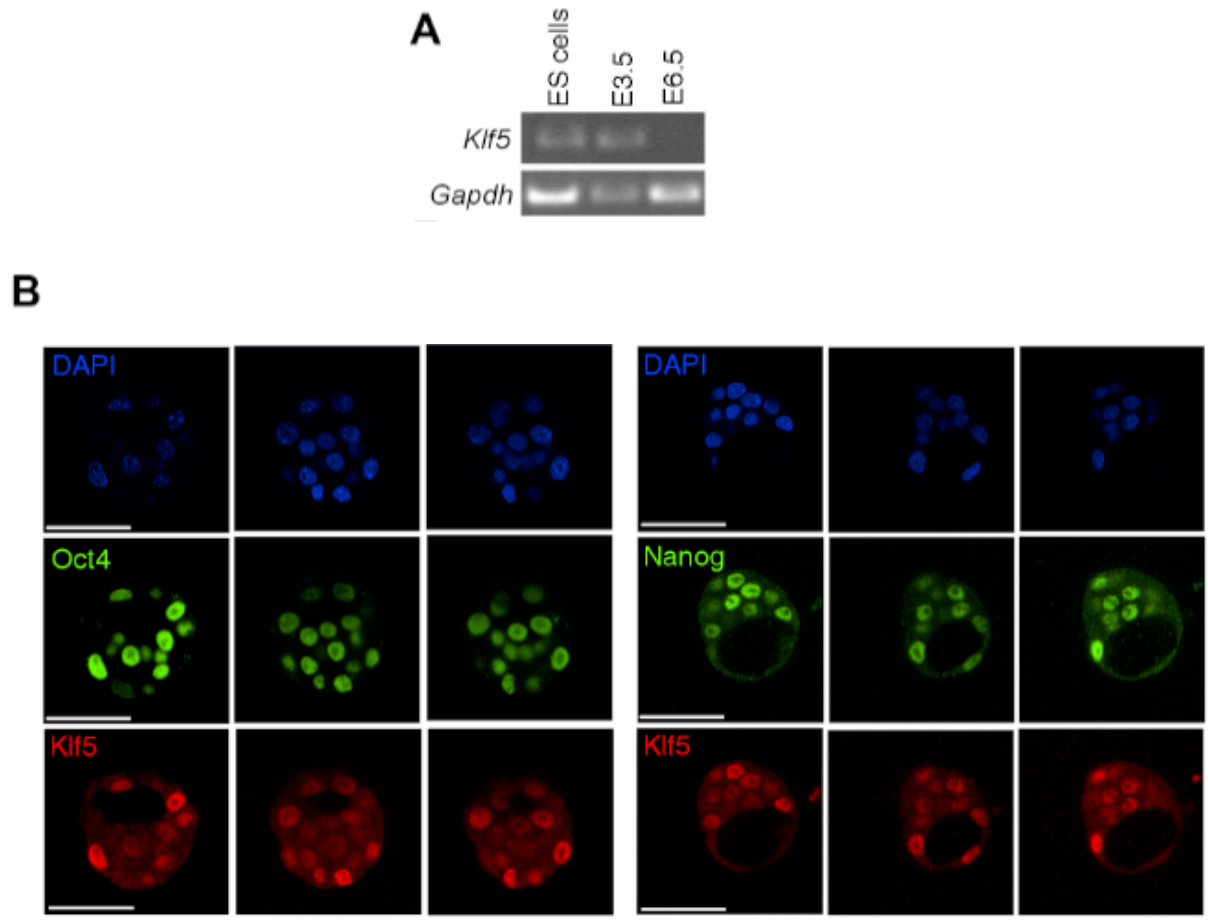


Figure 11| Klf5 is expressed in early stages of mouse embryo development
Klf5 is expressed in the blastocyst. RT-PCR of *Klf5* mRNA demonstrates that it is expressed in E3.5 but not in E6.5 embryos (**A**).
E3.5 mouse embryos were stained with Klf5 (red) and Oct3/4 or Nanog (green) antibodies. Confocal microscopy demonstrated that most of the nuclei are stained by Klf5 antibody with a wide range of intensities and co-localize with Oct3/4 and Nanog (**B**). Scale bar: 50 µm.

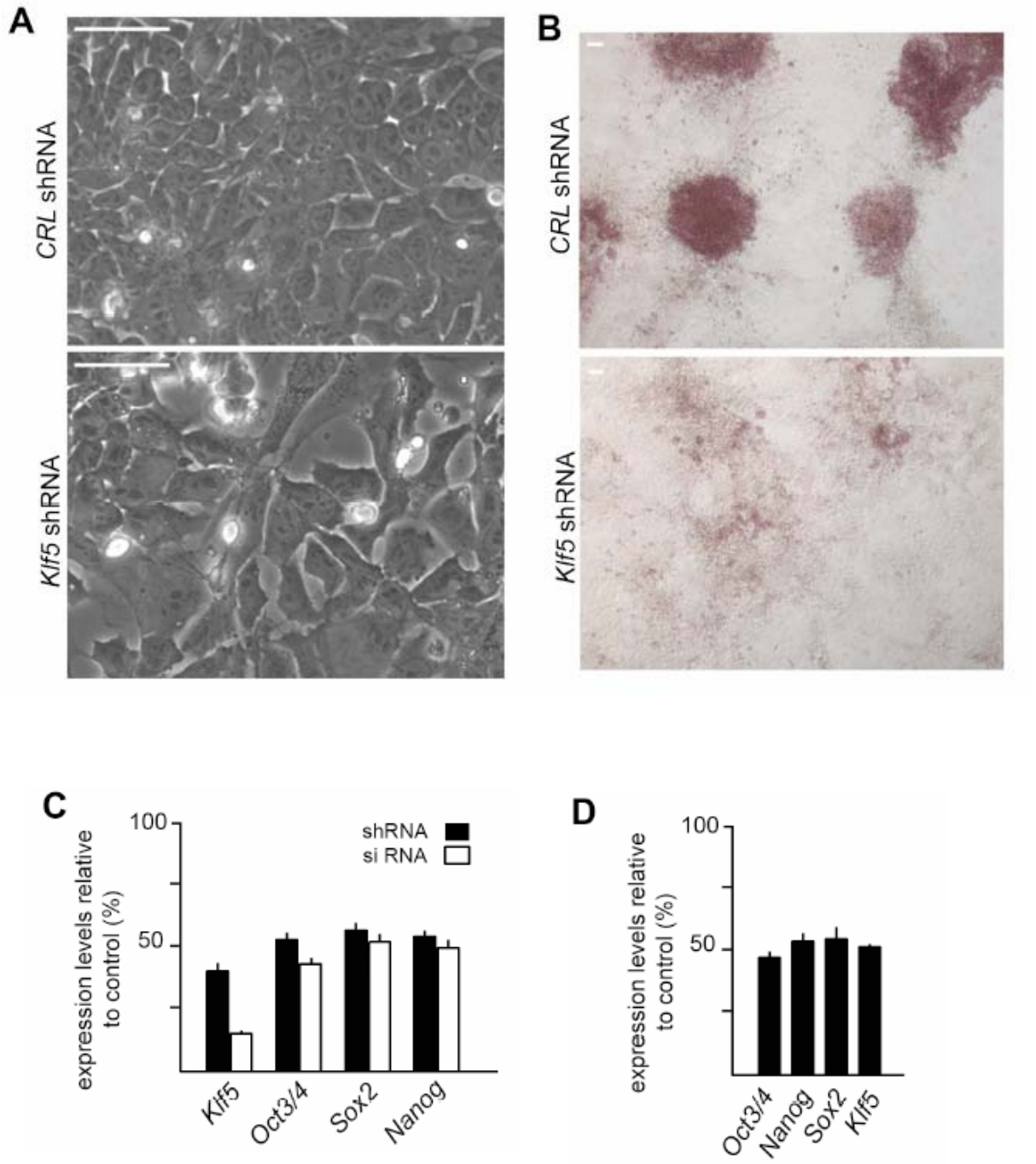

A
CRL shRNA
Klf5 shRNA
B
CRL shRNA
Klf5 shRNA
C
expression levels relative to control (%)
100
50
shRNA
si RNA
Klf5
Oct3/4
Sox2
Nanog
D
expression levels relative to control (%)
100
50
Oct3/4
Nanog
Sox2
Klf5

Figure 12| Klf5 knockdown cells lost their undifferentiated phenotype
Klf5 is required for maintaining ESCs in undifferentiated state.
ESCs transfected with *Klf5* shRNA show a phenotype different from that of ESCs transfected with non-silencing CRL shRNA. Lower panel (*Klf5* shRNA) shows an example of groups of enlarged cells characteristic of the *Klf5* knockdown (**A**). Scale bar: 50 μm.
Klf5 shRNA transfection induced the loss of colony formation and the almost complete loss of alkaline phosphatase staining. The upper panel shows ESCs transfected with the non-silencing CRL shRNA, whose microscopic phenotype is not distinguishable from that of untransfected cells (round colonies homogeneously positive for alkaline phosphatase). In the lower panel, *Klf5* shRNA transfected cells have completely different phenotype characterized by the absence of colonies, with only few cells weakly positive for alkaline phosphatase (**B**). Scale bar: 100 μm.
Klf5 knockdown, resulting from the transfection of the shRNA (black bars) or of a mixture of four siRNAs targeting different regions of the *Klf5* mRNA (white bars), reduced the expression of the gene to 45% and 12.5%, respectively, compared to basal levels measured in cells transfected with non silencing sh or siRNAs. This suppression is accompanied by a significant ($p<0.01$) decrease of *Oct3/4*, *Sox2* and *Nanog* mRNAs of about 50% compared to basal levels. Standard errors of three independent experiments are reported (**C**).
Silencing of *Oct3/4* downregulates Klf5 expression. ESCs were transfected with *Oct3/4* shRNA; four days after transfection, the mRNA levels of *Oct3/4*, *Sox2*, *Nanog* and *Klf5* were measured: *Klf5* mRNA levels were downregulated of about 50% (**D**).

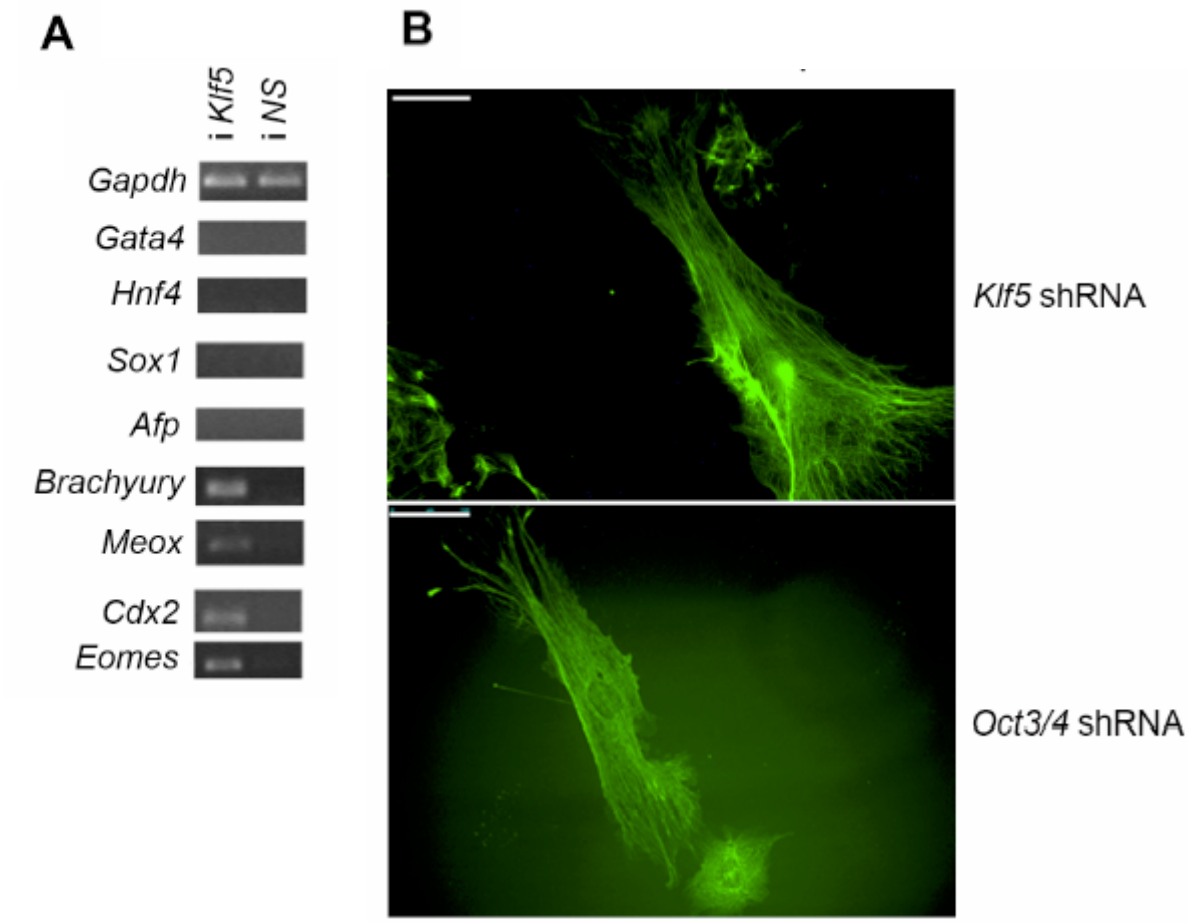
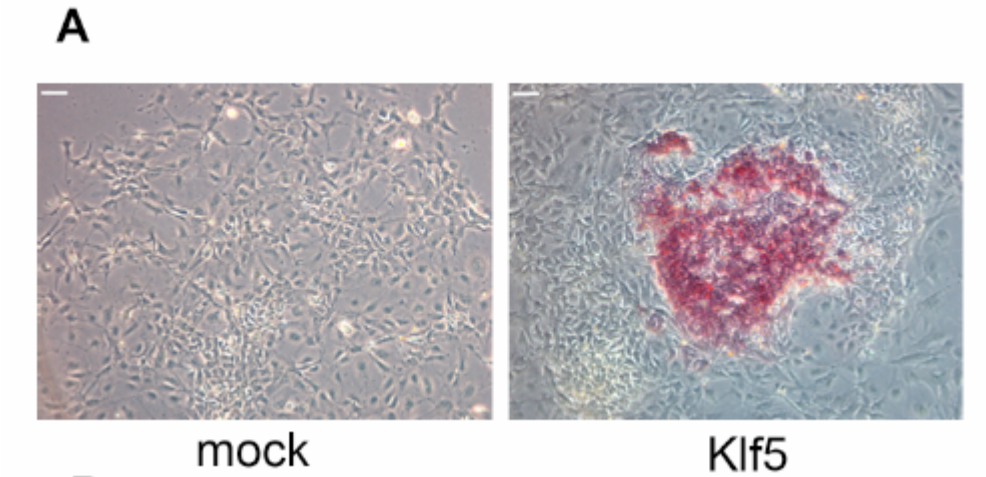


Figure 13| Phenotypic changes induced by Klf5 knockdown
RT-PCR analysis of mRNAs demonstrated that *Gata4* and *Hnf4* (endoderm), *Sox1* (ectoderm) and *Afp* (visceral endoderm) are not expressed both in *Klf5* and CRL shRNA transfected cells. On the contrary, *Klf5* knockdown causes an illegitimate expression of *Brachyury* and *Meox* (mesoderm) and *Cdx2* and *Eomes* (trophoblast), (**A**).
As a consequence of *Klf5* knockdown numerous cytokeratin 8-positive cells appear (upper panel). Very large and flattened cytokeratin 8-positive cells are similar to those observed in Oct3/4 knockdown cells (lower panel), (**B**). Scale bar: 100 μm.



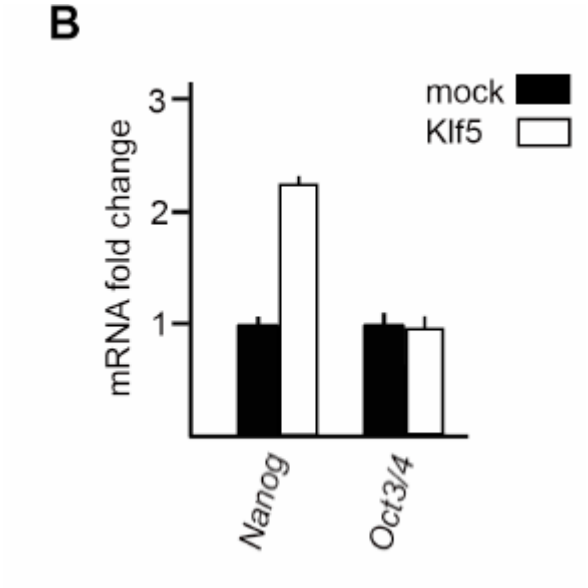
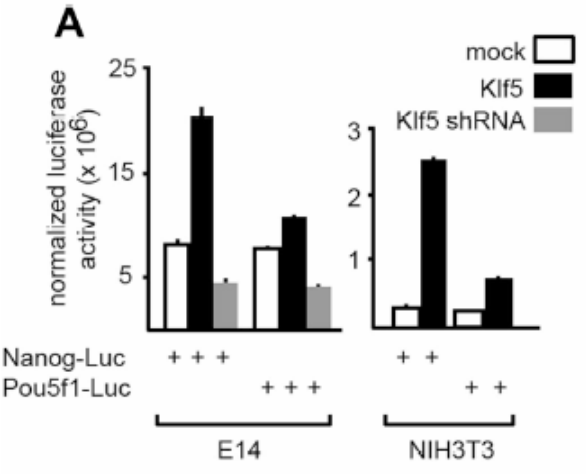
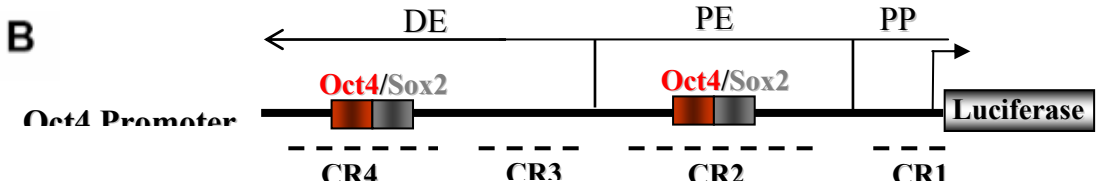


Figure 14| Klf5 constitutive expression confers a LIF-independent ability for pluripotency

Klf5 overexpression maintains ESC undifferentiated phenotype upon the withdrawal of LIF. ESCs were mock transfected (left panel) or transfected with Klf5 cDNA under the control of β-*actin* promoter (right panel) and grown for 6 days in the absence of LIF. In these conditions mock transfected cells completely lost alkaline phosphatase staining, while Klf5 overexpressing cells are still able to form colonies strongly positive for alkaline phosphatase staining (**A**). Scale bar: 100 μm.

Nanog and Oct4 mRNAs were measured by real time PCR six days after the removal of LIF. Nanog levels are significantly higher (2.3 folds) in the cells transfected with Klf5 than in mock-transfected cells ($p<0.01$), (**B**).


A
mock
Klf5
Klf5 shRNA
normalized luciferase activity (x 10^6)
25
15
5
3
2
1
Nanog-Luc
Pou5f1-Luc
+ + +
+ + +
+ +
+ +
E14
NIH3T3


B
DE
PE
PP
Oct4/Sox2
Oct4/Sox2
Luciferase
Oct4 Promoter
CR4
CR3
CR2
CR1

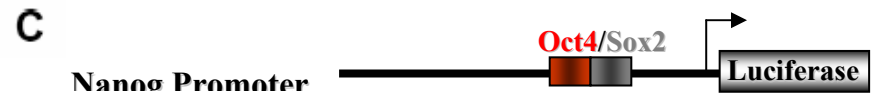

C
Oct4/Sox2
Luciferase
Nanog Promoter

Figure 15| Klf5 activates the transcription from *Nanog* and *Oct4* promoters
Nanog and *Oct4* promoters were cloned upstream of *luciferase* gene and co-transfected with Klf5 expression vector in ESCs or NIH3T3 cells. Although in ESCs both promoters direct a very strong transcription of the reporter, Klf5 overexpression is still able to increase the transcription from the *Nanog* promoter by more than two folds; the increase of Oct4 reporter transcription was small but significant ($p < 0.01$). On the contrary, Klf5 knockdown results in a strong decrease of transcription efficiency of both promoters. In the NIH3T3 cells the basal level of the reporter transcription is very low, according to the lack of competence of these cells to transcribe ESC-specific genes. In these conditions Klf5 overexpression causes a strong induction of *Nanog* promoter and a lower but significant induction of the *Oct4* one (**A**).
Oct4 promoter (-2200+16) and *Nanog* promoter (-329+50) schematic representation. *Oct4* promoter is a TATA-less promoter containing two elements named proximal enhancer (PE) and distal enhancer (DE), based on their position with respect to the transcription initiation site, regulate the stem cell specific activity of Oct-4. The DE is active in ES cells whereas the PE is active in EC. Four conserved regions (CR1-4) lie among these enhancers that share high sequence homology among different species (Nordhoff et al. 2001), (**B**).
The 380bps of Nanog proximal promoter recapitulates appropriate Nanog expression in pluripotent and non-pluripotent cells, and this motif is well conserved between mouse, rat and human. It contains the Oct/Sox motif essential for Nanog ESC specific expression (Kuroda et al. 2005), (**C**).

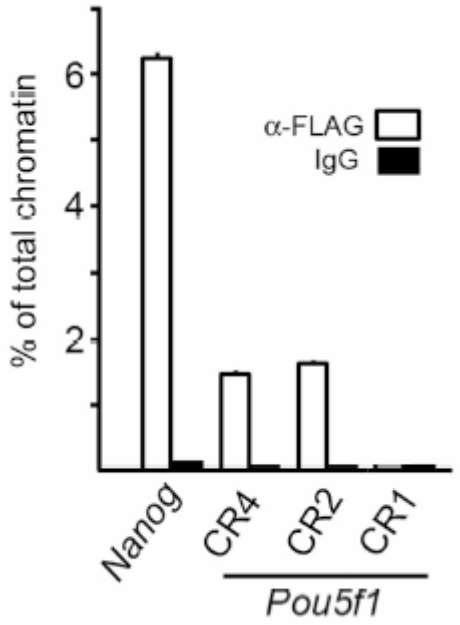


Figure 16| Klf5 directly interacts with *Nanog* and *Oct4* promoters
ESCs were processed for ChIP with antibodies specific for flag epitope (white bars) or IgG as control (black bars). Unprecipitated chromatin preparations were used as "input" control. The amount of precipitated DNA was calculated relative to the total input chromatin, and expressed as the percentage of the total according to the formula reported under Methods. The results demonstrate that Klf5 is associated with the chromatin at the *Nanog* and *Oct3/4* loci. In the last case the enrichment was observed at both CR4 and CR2 regions of the promoter (Chew et al.2005).

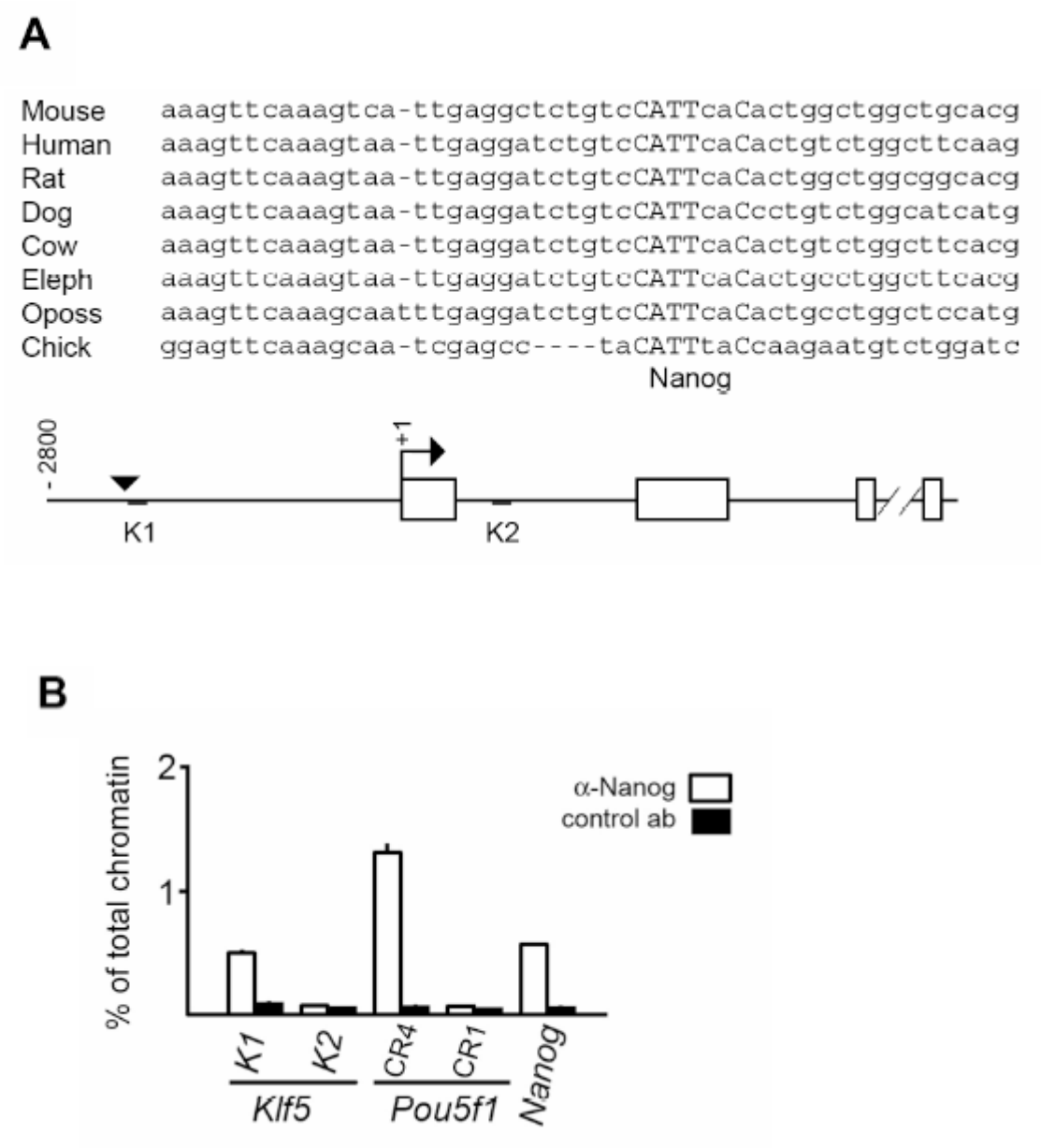


Figure 17| Nanog interacts with *Klf5* promoter

A conserved putative Nanog *cis*-element is present in the *Klf5* promoter region.

Alignment of sequences in the region upstream of the *Klf5* transcription start site (+1) is shown. Mouse sequence is from chromosome 14 starting from nucleotide 98,179,221. Upper case letters indicate conserved nucleotides in Nanog *cis*-element. The lower panel shows a schematic representation of *Klf5* and reports the position of the putative Nanog binding site (arrowhead) and those of the two regions (K1, K2) amplified in the ChIP experiments (**A**).

Nanog directly interacts with *Klf5* promoter. ChIP experiments demonstrate that chromatinised Nanog is present on the *Klf5* promoter. The differences with control antibody and control DNA were significant with $p < 0.001$ (**B**).