University of Navarra

Faculty of Science

IDENTIFICATION AND CHARACTERIZATION OF PLURIPOTENCY ASSOCIATED LNCRNAS IN HUMAN IPS CELLS

Doctoral Thesis

Natalia María Zapata Linares

Pamplona, June 2015

University of Navarra

Faculty of Science

A dissertation presented by **Natalia María Zapata Linares** to aim for the degree of Doctor of Philosophy by the University of Navarra.

This work has been conducted under our supervision at the Cell Therapy Area of the Center for Applied Medical Research (CIMA), University of Navarra. We hereby authorize its presentation to the Thesis Committee for evaluation.

Pamplona, April 23rd, 2015

Dr. Felipe Prosper
Director

Dr. Juan Roberto Rodríguez Madoz
Co-director

This thesis has been supported by the “Formación de personal de investigación” program, Foundation for the Applied Medical Research (FIMA), “COLCIENCIAS” Scholarship for Graduate Studies Abroad, LASPAU (Harvard University), Escuela de Ingeniería de Antioquia, and the “EnlazaMundos” fellowship for Graduate Studies Abroad, Alcaldía de Medellín.

ACKNOWLEDGEMENTS

Estos años han dejado huella a nivel profesional, pero sobre todo a nivel personal, debido a todas las personas que de cerca o de lejos han estado conmigo, a quienes debo este trabajo y han hecho de éste un periodo fructífero. Quisiera empezar agradeciendo a mis padres, Lileth y Miguel y a mi hermana Carolina quienes son mi motor, mi motivación y mi apoyo; y a Alberto y su familia por abrirme sus corazones.

Gracias a mis directores, a Felipe Prósper por la confianza que puso en mí desde el primer día que entré a su despacho y JuanRo porque más que un jefe ha sido un verdadero tutor y maestro, un incondicional compañero de viaje.

Gracias a Saray, quien me ha prestado sus manos para llegar donde no podía; a Manu M y Angelo P por las casi interminables conversaciones filosofo-científicas, siempre enriquecedoras. Al grupo CUN, Quique, Miguel, y Gloria por amenizarme las tardes. Gracias a Tania por enseñarme esa verdadera amistad pamplonica que no termina nunca y a Adri y Emma por sus consejos siempre oportunos y las buenas conversaciones. A mi *Inc-partner* Andoni por estar siempre disponible a ayudar y a Esti A por su calidez y cercanía. Gracias a todos los miembros del Labo 1.01 pasado y presente, por hacer el día a día más llevadero y darle un nuevo y mejor significado a la palabra *trabajo*.

Gracias a aquellos amigos que se convirtieron en mi familia en este lado del atlántico: Dani, Edwin, Juan Alejandro y Andrés Felipe, Lau, Nico, Carmen D, Hernán, María B, Patri, Kenia, Carmen M, Yira, Ana y Luciana. A mis compañeras de camino y no solo del de Santiago: AnaLau, Blanca, Merche, Marts y María. Y no me puedo olvidar a quienes comenzaron conmigo esta formación en el 2009 y siguen estando ahí: Antonia, Irene y Marina.

Quiero agradecer a todos aquellos que me han demostrado que la verdadera amistad sobrepasa las barreras del tiempo y la distancia, especialmente a La Coste, Susi, Apple, Lina, Adri, Maffe y sus respectivas familias; Pipe, Naty A, Lula, Cata V,Ana F, Thelma, Caro D y July R.

Finalmente, gracias a las instituciones que creyeron en mí y apostaron por mi formación: a la Fundación para la Investigación Médica Aplicada (FIMA), al departamento de Innovación Ciencia y Tecnología de Colombia (COLCIENCIAS), a LASPAU (Harvard University), al programa de Ingeniería Biomédica (EIA-CES) y al programa “EnlazaMundos” de la Alcaldía de Medellín.

Este viaje termina pero sé que seguiré contando con cada uno de ustedes en las aventuras que me esperan después, ¡MUCHAS GRACIAS A TODOS!

"Es una gran escuela de humildad tener que hacer continuamente cosas que solo con gran esfuerzo consigue una llevar a cabo imperfectamente"
(Edith Stein, Carta244, p.292).

To Mom, Dad and Caro

ABBREVIATIONS

3F	Three factors
4F	Four factor
α-MEM	Alpha minimum essential medium
ADSC-hiPSC	ADSC derived human induced pluripotent stem cells
ADSCs	Adipose-derived stem cells
bFGF	Basic fibroblast growth factor
BSA	Bovine serum albumin
cDNA	Complementary DNA
CGH	Comparative genomic hybridization microarray
ChIP	Chromatin immunoprecipitation
dCt	Delta Ct
cMYC	V-myc avian myelocytomatosis viral oncogene homolog
DMEM	Dulbecco's modified eagle medium
DMEM-KO	Knockout Dulbecco's Modified Eagle Medium
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
EBs	Embryoid bodies
EC	Ectoderm
EMT	Epithelial to mesenchymal transition
EN	Endoderm
EpiSCs	Epithelized embryonic stem cells
ESCs	Embryonic stem cells
FACS	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
FC	Fold change
Fib-hiPSC	Fibroblast derived human induced pluripotent stem cells
FOXA2	Forkhead box A2
GFP	Green fluorescent protein
H3K27me3	Histone 3 lysine 27 trimethylation
H3K4me3	Histone 3 lysine 4 trimethylation
HE	Hematoxylin/eosin
hESCs	Human embryonic stem cells
hFibs	Human fibroblasts
HSCs	Hematopoietic stem cells
iPSCs	Induced pluripotent stem cells
KLF (No)	Krüppel-like factor s (Number)
LIF	Leukemia inhibitory factor
LincRNA	Long intergenic non coding RNA
LincRNA-RoR	LincRNA-Regulator of reprogramming
lncRNAs	Long-non coding RNAs
ME	Mesoderm

Abbreviations

MEF	Mouse embryonic fibroblasts
mESCs	Mouse ESCs
MET	Mesenchymal-to-epithelial transition
miPSC	Mouse induced pluripotent stem cells
miRNA	micro RNA
MMLV	Moloney murine leukemic virus
mRNA	Messenger RNA
MSCs	Mesenchymal stem cells
MTS	3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium
NCCIT	Human germ cell tumor-derived cell line
ncRNA	Non-coding RNA
NEAA	Non-essential amino acids
nt	Nucleotides
NTERA	Human teratocarcinoma cell line
OCT (No)	Octamer-binding transcription factor (Number)
PAL	Pluripotency associated lncRNA
PcG	Polycomb group
PGCs	Primordial germ cells
piRNA	PiwiRNA
PL	Pluripotency
pMXs	Retroviral vector system based on MMV
POU5F1	POU domain, class 5, transcription factor 1
PSCs	Pluripotent stem cells
qPCRs	Quantitative polymerase chain reaction
RIP	RNA immunoprecipitation
RNA	Ribonucleic acid
RT	Room temperature
SCID	Severe combined immunodeficient
SCNT	Somatic cell nuclear transfer
shRNA	Short-hairpin RNA
siRNA	Small interfering RNA
snoRNA	Nucleolar RNA
snRNA	Small nuclear
SOX (No)	SRY (sex determining region Y)-box (Number)
SVF	Stromal vascular fraction
TAT	HIV trans-activator of transcription
TF	Transcription Factor
TrxG	Trithorax group
VPA	Valproic acid
WNT	Wingless
XIST	X-inactive-specific transcript

TABLE OF CONTENTS

TABLE OF CONTENTS

1. INTRODUCTION 1

1.1. PLURIPOTENCY AND INDUCED PLURIPOTENCY. 3

1.1.1. Stem cell basics. 3

1.1.2. Pluripotent stem cells. 4

1.1.2.1. Embryonic stem cells (ESCs). 5

1.1.2.2. Germ cells cultured with extrinsic factors. 6

1.1.2.3. Induced pluripotent stem cells. 6

1.1.3. The molecular basis of pluripotency. 6

1.1.4. Cellular reprogramming. 8

1.1.4.1. Somatic cell nuclear transfer (SCNT) or Cloning. 9

1.1.4.2. Cell Fusion. 9

1.1.4.3. Transduction of transcription factors. 9

1.2. INDUCED PLURIPOTENT STEM CELLS (IPSCS). 11

1.2.1. Reprogramming mechanisms. 11

1.2.2. Factors controlling cell reprogramming. 12

1.2.2.1. Reprogramming cocktail. 12

1.2.2.2. Donor Cell type 13

1.2.2.3. Cell culture method. 14

1.2.2.4. Reprogramming methods. 14

1.2.2.4.1. Integrative methods. 14

1.2.2.4.2. Non-integrative methods. 15

1.2.2.5. Conclusions of cell reprogramming. 18

1.2.3. ESCs vs. iPSC. 19

1.2.4. iPSC applications. 19

1.3. LNCRNAS AND CELL FATE REGULATION. 21

1.3.1. LncRNAs: definition 21

1.3.2. lncRNAs: Mechanisms of action. 21

1.3.3. LncRNAs and tissue specificity. 23

1.3.4. LncRNAs in pluripotency and reprogramming. 23

1.3.5. Concluding remarks. 24

2. HYPOTHESIS & OBJECTIVES 27

3. METHODS 31

3.1. CELL CULTURE. 33

3.1.1. Media preparation. 33

3.1.2. Preparation of gelatin- and matrigel-coated dishes. 33

3.1.3. Preparation of ADSC's conditioned media. 33

3.1.4. Freezing/Thawing of cells. 34

3.1.5. Isolation of mouse embryonic fibroblasts (MEFs) and preparation of feeder layers. 34

3.1.6. Isolation of human fibroblasts (hFibs) and adipose derived stem cells (ADSCs). 34

3.2. GENERATION OF INDUCED PLURIPOTENT STEM CELLS. 35

3.2.1. Retroviral generation of human induced pluripotent stem cells. 35

3.2.2. Lentiviral generation of human induced pluripotent stem cells. 36

3.2.3. Colony observation. 37

3.2.4. Picking-up 37

3.2.5. Culture of iPSCs. 37

Table of contents

3.3. CHARACTERIZATION OF INDUCED PLURIPOTENT STEM CELLS. 37
3.3.1. Trizol reagent total RNA extraction. 37
3.3.2. cDNA synthesis, primer design and quantitative qPCR. 38
3.3.3. qPCR for retroviral silencing. 39
3.3.4. Immunostaining. 39
3.3.5. Alkaline phosphatase staining 39
3.3.6. Bisulfite treatment and pyrosequencing. 40
3.3.7. Spontaneous differentiation by embryoid body formation. 40
3.3.8. Teratoma formation assay. 41
3.3.9. Genomic stability: CGH analysis 41
3.4. NON CODING-RNA PROFILING. 41
3.4.1. Array hybridization and bioinformatics analysis. 41
3.4.2. Array validation and LncRNA candidate selection criteria. 42
3.5. OVER-EXPRESSION OF LNCRNAS USING RETROVIRAL VECTORS. 42
3.5.1. Cloning pMXs-LncRNAs. 42
3.5.2. Retroviral production of pMXs-LncRNAs. 42
3.6. LOSS OF FUNCTION EXPERIMENTS. 45
3.6.1. Small interfering RNA (Si-RNA) design. 45
3.6.2. Short-hairpin RNA (shRNA) cloning in lentiviral vector. 45
3.6.3. Lentiviral production and establishment sh-pluripotent cell lines. 45
3.6.4. Proliferation assay. 45
3.7. STATISTICS. 46

4. RESULTS 47
4.1. GENERATION AND CHARACTERIZATION OF HUMAN IPSC. 49
4.1.1. Generation of hiPSC from fibroblasts and ADSCs. 49
4.1.2. Characterization of hiPS cells. 50
4.1.2.1. Expression of endogenous pluripotency markers. 50
4.1.2.2. Silencing of the retroviral transgenes. 52
4.1.2.3. In vitro spontaneous differentiation by embryoid bodies formation. 53
4.1.2.4. In vivo teratoma generation. 55
4.1.2.5. NANOG promoter hypomethylation. 56
4.1.2.6. Genomic stability: CGH analysis. 58
4.1.2.7. Summary of hiPSC characterization. 59
4.2. LNCRNA PROFILING OF INDUCED PLURIPOTENT STEM CELLS. 61
4.2.1. Gene expression profile analysis. 61
4.2.2. LncRNAs expression profile in pluripotent cells. 63
4.2.3. LncRNAs characterization. 64
4.2.3.1. lncRNAs candidates are expressed in diverse pluripotent cell lines. 64
4.2.3.2. LncRNAs candidates are not induced early in the reprogramming process. 66
4.2.3.3. LncRNAs candidates' expression upon spontaneous differentiation. 67
4.2.3.4. LncRNAs characterization summary. 68
4.3. PAL 20 AND PAL21 FUNCTIONAL ANALYSIS. 70
4.3.1. PAL20 and PAL21 genetic features. 70
4.3.2. Overexpression of PAL candidates. 71
4.3.2.1. PAL20 and PAL21 have no effect as reprogramming factors. 71
4.3.2.2. PAL20 and PAL21 overexpression during EBs. 73
4.3.3. Inhibition of PAL candidates: Loss-of-function experiments. 73
4.3.3.1. shRNAs design for PAL20 and PAL21 inhibition. 73
4.3.3.2. PAL20 and PAL21 are required for cell reprogramming. 74

4.3.3.3. Silencing of PAL20 and PAL21 expression in human iPSCs. 75
4.3.3.4. Silencing of PAL20 and PAL21 expression in cell lines induce changes in cell proliferation. 76
4.3.4. Summary of PAL20 and PAL21 functions. 78

5. DISCUSSION 79
5.1. DISCUSSION: GENERAL ASPECTS. 81
5.2. GENERATION OF HUMAN INDUCED PLURIPOTENT STEM CELLS. 82
5.2.1. Somatic cell reprogramming. 82
5.2.2. hiPSC characterization. 84
5.2.3. Gene expression profile. 86
5.3. LNCRNA PROFILING. 88
5.3.1. Identification of PALs. 88
5.3.2. Mechanism of function of PAL20 and PAL21. 89
5.4. REMARKABLE CONCLUSIVE ASPECTS. 92

6. CONCLUSIONS 93

7. BIBLIOGRAPHY 97

8. ANNEX 119

1. INTRODUCTION

1.1. PLURIPOTENCY AND INDUCED PLURIPOTENCY.

1.1.1. Stem cell basics.

Stem cells are natural cells that possess a unique capacity of self-renewal and differentiation to specialized tissues, making them essential for biological organization, natural selection, homeostasis maintenance and are responsible for the regeneration of somatic systems after disease or injury (1,2). Self-renewal stands for the capacity of a stem cell to maintain an undifferentiated state despite a high number of cell divisions. A subset of stem cells is capable of indefinite self-renew, generating two similar daughter cells after symmetrical division and expanding the stem cell pool. Another group of stem cells divides asymmetrically for both self-renew and to generate differentiated daughter cells (Figure 1) (2,3).

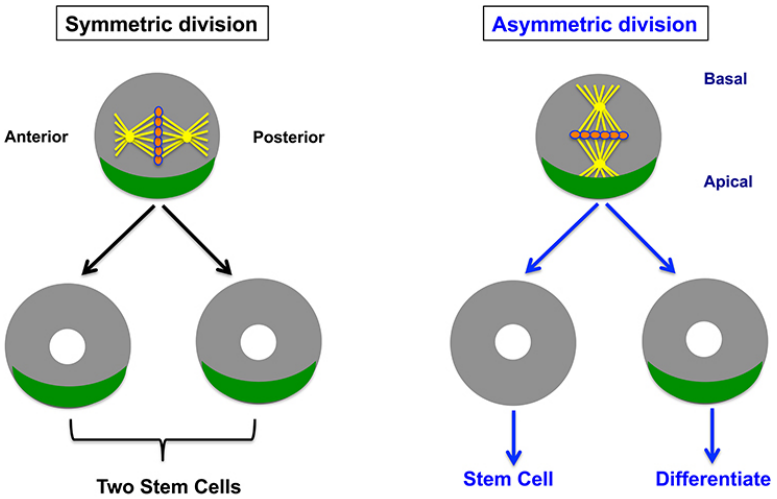


Figure 1. Schematic representation of the two cell division models. During symmetrical division, determinants segregate equally, giving rise to two equal stem cells. During asymmetric division, determinant protein localization are coordinated, giving rise to a differentiating cell and a stem cell.

In terms of the plasticity and the differentiation potential, stem cells can be classified into different groups: i) **unipotential** stem cells, that give rise to only one mature cell type, (i.e. keratinocyte stem cell of the dermis); ii) **multipotent** stem cells, with the capacity to generate at least two or more differentiated cell types (i.e. neural crest stem cells (NCSCs) of the nervous system, hematopoietic stem cells (HSCs) and mesenchymal stem cells (MSCs) of the bone marrow); iii) **pluripotent** stem cells (PSCs), which yield to a vast array of mature cell types belonging to the three germ layers, endoderm, mesoderm and ectoderm (i.e. Embryonic stem cells, (ESCs)) and induced pluripotent stem cells (iPSCs); and iv) **totipotent** stem cells, capable to generate both embryonic and extra-embryonic tissues (1,2) (Figure 2).

The study of stem cells has increased within the past decade due to their repercussions in the understanding of cell biology and cell development, in addition to their potential for therapeutic applications (i.e. regenerative medicine) and revealing disease mechanisms. Regenerative medicine involves growing new cells, tissues and/or organs to replace/repair damaged ones due to injury, disease or aging (4). The first report of translational medicine dates from 1959, when the French oncologist Georges Mathé performed the first bone marrow transplant to a group of five nuclear workers whose own marrow had been damaged by irradiation (5).

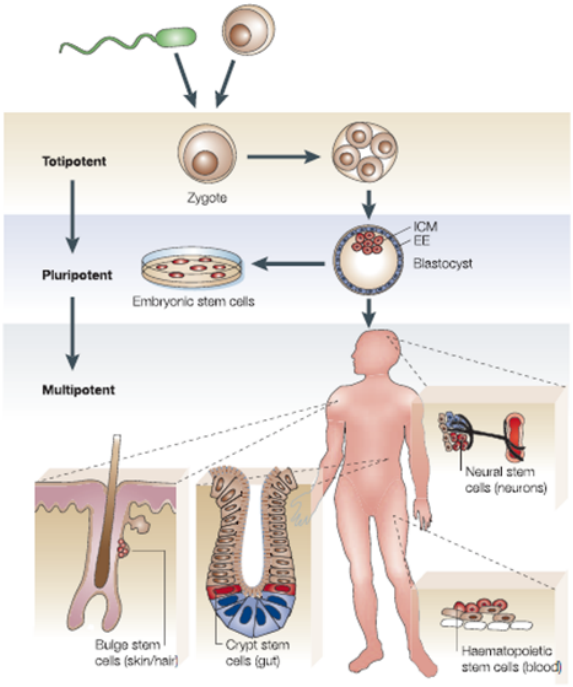


Figure 2. Classification of stem cells according to their differentiation status. Zygote is the most un-differentiated state capable of giving rise both extra-embryonic and embryonary tissues. The un-differentiated status is lost through development, passing for different stages, pluripotency, able to form cells from the three germ layers, multipotent, committed to one germ layer and somatic cells specific for one function within a specific tissue. Image extracted from (6)

Even though their promising applications, some aspects of stem cell therapy still remain poorly understood and are currently under exhaustive study (7,8). Additionally, regenerative medicine using ESCs may undergo both ethical issues (due to human embryos destruction implicated in their isolation) and safety issues (like the possible teratoma formation or the immune rejection upon ESCs transplantation) that should be solved before clinical applications (9). As discussed below the use of iPSCs could overcome some of those issues.

1.1.2. Pluripotent stem cells.

Pluripotent state is characterized by an indefinite self-renewal capacity and a differentiation potential towards cells from the three germ layers. PSC characteristics include i) endogenous expression of pluripotent genes; ii) alkaline phosphatase enzymatic activity; iii) presence of specific surface markers, like stage-specific-embryonic-antigen (SSEAs) members and glycoproteins TRA1-60 and TRA1-81; iv) telomerase activity, to prevent cell senescence protecting telomeres from degradation; v) *in vivo* capacity for chimeras formation with germ line contribution after blastocyst injection; and vi) the capacity to generate an entire new organism through tetraploid complementation (10).

Recently, it has been proposed that pluripotency is highly dynamic with many different states (11), which seem to depend directly from the stoichiometry of core pluripotent transcription factors (TFs) (10). The Ground or Naïve pluripotent state belongs to the cell population able to self-renew and gives rise to chimeras. On the other hand, PSCs in a primed state are committed towards a

specific lineage. These cells in primed state are not able to form chimeras but they express pluripotent markers and form teratomas *in vivo* (12,13), which led to the conclusion that primed state is disposed to differentiate whereas the naïve state correspond to a more immature state of pluripotency. Introduction of molecules that block differentiation signals, like the cytokine leukemia inhibitory factor (LIF), into the cell culture help to maintain the naïve state of PSCs *in vitro* besides promoting growth and viability (14). The PSCs can be obtained from different sources, including:

1.1.2.1. Embryonic stem cells (ESCs).

ESCs are pluripotent stem cells derived from the inner cell mass of a blastocyst, an early-stage pre-implantation embryo. They have been successfully generated from different species, including mouse and humans.

Mouse ESCs (mESCs) constitute the best-characterized pluripotent stem cells. Martin *et al.* were the first in isolating mESCs from the inner cell masses of late blastocysts establishing a pluripotent stem cell line (15). mESCs are maintained in an undifferentiated state by specific culture conditions in the presence of leukemia inhibitory factor (LIF) (15). Formation of embryoid bodies (EBs) is a common and easy test to evaluate their differentiation capacity *in vitro*. By this method, that resembles the early stages of development, mESCs showed the expression of specific genetic markers and the consequent presence of cells representative of all three embryonic germ cell layers. Moreover, mESCs can give rise to a new organism by tetraploid complementation when implanted into a blastocyst, technique that has helped the study of mouse embryo developing (16).

Mouse epiblast-derived stem cells (mEpiSC) are pluripotent stem cells isolated form inner cell masses of post-implantation embryo (12). EpiSC have several different cellular and molecular characteristics with mESC (17) and pluripotency in EpiSC is essentially different from that of embryonic stem cells due to its primed state.

Human embryonic stem cells (hESCs) were first obtained from human embryos produced by *in vitro* fertilization. When culturing on feeder layers these cells formed ES-like colonies with flat morphology and low tolerance to grow in single cell, a key differentiator from mESCs culture. hESCs express cell surface markers that characterize undifferentiated state, like SSEA-3, SSEA-4, TRA1-60 and TRA1-81. Moreover when injected in severe combined immunodeficient (SCID) mice produced teratomas containing representative cells or tissues from all three embryonic germ layers (18). It has been proposed that hESCs are similar to mEpiSCs, being in a primed pluripotent state, probably due to the obtaining method and their posterior *in vitro* culture that includes basic fibroblast growth factor (bFGF), a known lineage specific inducer (19). It is assumed that hESCs can differentiate *in vitro* into any cell lineage, however their *in vivo* chimera contribution is still under study. Additionally, the ethical concerns due to the destruction of human embryos during the obtaining process of hESCs, the possible genetic abnormalities due to *in vitro* expansion and the safety issues, maintain the therapeutic potential of these cells in question (20,21). However, motor neuron progenitor cells and cardiac progenitor cells derived from hESCs are being tested in two different clinical trials for spinal cord injury and severe heart failure respectively (22,23).

1.1.2.2. Germ cells cultured with extrinsic factors.

Germ cells, the cells that give rise to the gametes, are often set-aside during embryonic cleavage. During development, these cells will differentiate into primordial germ cells (PGCs), migrate to the location of the gonad and form the germ line of the animal. PGCs from mouse can be reprogrammed into pluripotent ES-like cells by culturing them with bFGF, steel factor (SL) and LIF. These PGCs were able i) to grow and proliferate in colonies with ESC morphology; ii) to express pluripotency-associated markers like SSEA-1; iii) to differentiate *in vitro* (i.e. EBs formation) and *in vivo* by developing teratomas when injected in nude mice; and iv) to give rise to chimeras when introducing into blastocyst (24). Recently, Kimura *et al*., have reported that the treatment of PGCs with an inhibitor of the transforming growth factor beta receptor (TGFβR) and/or Kempaullone induced pluripotency on germ cells. This could be due to the similarity in function of these molecules with the pluripotency associated TFs, Sox2 and Klf4 (25), and with the more un-differentiated stated of PGCs, allowing to overcome the reprogramming barriers more easily by changing culture conditions in order to reach a pluripotent state.

1.1.2.3. Induced pluripotent stem cells.

Pluripotency can also be induced in somatic cells by the ectopic expression of specific TFs described to be markers of pluripotent stem cells. Different cocktails have been described to induce pluripotency in mouse and human cells. The group of Shinya Yamanaka was the first in generating miPSCs and hiPSCs with a reprogramming cocktail that included POU5F1, SOX2, KLF4 and c-MYC (26). In general these factors activate the endogenous pluripotency genes generating an ES-like population. Since iPSCs are the basis of this work a detailed review is described in section 1.2.

1.1.3. The molecular basis of pluripotency.

Pluripotency is characterized by indefinite self-renewal and the differentiation potential to the three germ layers (i.e. generation of chimeras). Cells that reach these criteria are considered to be pluripotent and are governed by both genetic and epigenetic features. Indeed, the expression of several TFs has been associated to pluripotent state.

POU5F1 (POU domain, class 5, transcription factor 1), also known as OCT4 (octamer-binding transcription factor 4), is a member of the POU-domain transcription factor family, it is expressed throughout early mammalian development and it is consider as the master gene of pluripotency (27). Loss of POU5F1 function induces ESCs differentiation and its overexpression induces primitive endoderm and mesoderm lineages (28).

SOX2 (SRY (sex determining region Y)-box 2) is a member of the SRY-related homeobox genes and although its expression is not exclusive of ESCs, it has a tight relationship with POU5F1. Both molecules heterodimerize forming a complex targeting other pluripotency markers (29). POU5F1 and SOX2 deletion lead to impairment on developmental potential (30).

NANOG is a homeobox protein that is critically involved with self-renewal of undifferentiated embryonic stem cells. Its expression is activated early, at the 8-cell stage, and later to cooperate in X chromosome reactivation (31). The activated state of the two X chromosomes (XaXa) in female cells is an essential feature of naive ESCs (32,33). Cells lacking NANOG expression have an

increased tendency to differentiate (34,35) and the forced expression of this transcription factor is sufficient to prevent ES cell differentiation and render self-renewal. Individual ES cells exhibit variable expression of key factors such as NANOG, have distinct probabilities of self-renewal, and may have differing developmental potential (10).

KLF4 (Krüppel-like factor-4) along with other KLF factors were first related with tumor suppression and oncogenesis, however depletion of KLF2, KLF4 and KLF5 lead to ES cell differentiation. KLF4 is required for self-renewal probably via NANOG with which shares many targets and assisting the action of OCT4 and SOX2 to initiate the expression of ESCs-specific genes (36).

c-MYC (V-myc avian myelocytomatosis viral oncogene homolog) is a multifunctional transcription factor that regulates the expression of a multiple genes in normal cells playing a role in cell cycle progression, apoptosis and cellular transformation. Its indirectly involved in the establishment of pluripotency by facilitating the action of the other reprogramming factor although is dispensable for the induction of pluripotency (37). c-MYC expression in regulated by microRNAs (miRNAs) and WNT signaling pathway, that promotes naive pluripotency (38,39).

It has been reported that the expression of three TFs (POU5F1, SOX2 and NANOG) confers pluripotency to a cell. Those factors establish an auto-regulatory loop with a key role in the suppression of differentiation signals (30,40,41) (Figure 3). POU5F1-SOX2-NANOG constitutes a core transcriptional circuit that also co-activates redundant target genes and secondary TFs like estrogen-related receptor beta (ESRRB), Zinc finger protein 42 (ZFP42) or REX1, KLF4, KLF2, T-box 3 (TBX3), transcription factor CP2-like 1 (TFCP2L1), c-MYC, LIN28, STAT3 pathway; that provide further stability to the pluripotent state (42,43)(Figure 3).

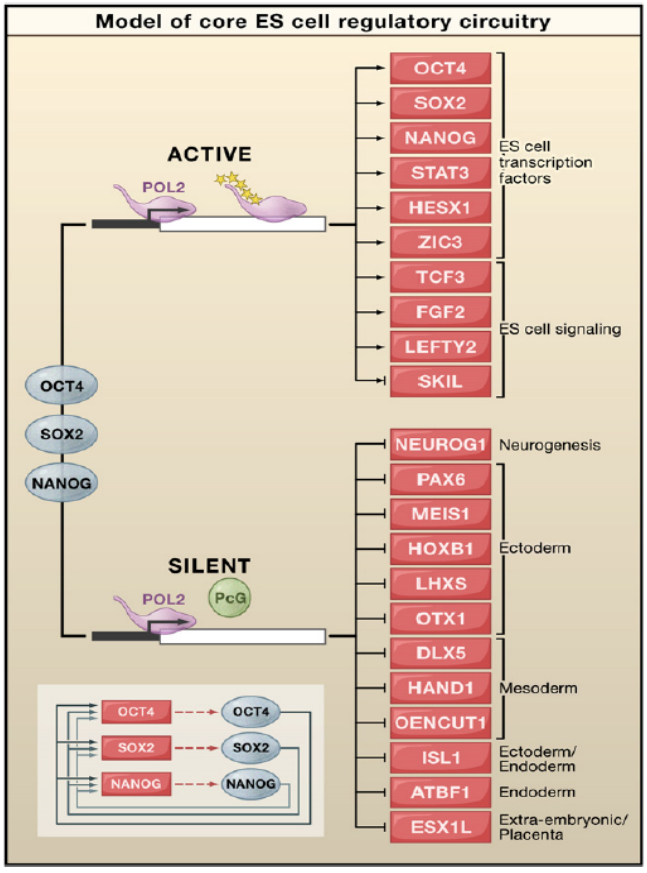


Figure 3. Molecular basis of pluripotency. POU5F1 (OCT4), SOX2 and NANOG are the core pluripotency TFs. They activate downstream pluripotency factors, contributing all to the maintenance of ESCs. POU5F1, SOX2 and NANOG participate in a positive feedback loop as well as occupying promoters of differentiation programs inhibiting their expression. Image extracted from (44).

Regulation of pluripotency is a complex and not well-understood process where epigenetic factors and non-coding RNAs are essential. **Epigenetic factors** include DNA methyltransferases, like Dnmt3a and Dnmt3b (45), the histone deacetylase NuRD complex (46), SWI/SNF chromatin remodeling complexes (47), the Polycomb (PcG) and Trithorax group (TrxG) (48,49) that regulates the demethylation of pluripotency gene promoters, the silencing of lineage specific genes or the presence of bivalent promoters in PSCs. Bivalent promoters are characterized by presence of both activating and repressive histone marks histone 3 lysine 4 trimethylation (H3K4me3) and histone 3 lysine 27 trimethylation (H3K27me3) indicating their repressed but "poised" transcriptional state characteristic of the plasticity of these cells. Pluripotency regulation is also associated with specific signature of **non-coding genes**. Several miRNA families are up-regulated in hPSCs including miR-302, miR-106, miR-372, miR-17, miR-520, miR-195 and miR-200 families. Members of these families share similar seed sequences, suggesting that these miRNAs may have similar messenger RNAs (mRNAs) targets and regulatory functions (50,51). miRNA families play roles in different pathways in PSCs, for example Lin28 inhibits the processing of let-7 a distinctive miRNA of differentiated cell state (52). hsa-miR-372 family is related to cell-cycle pathway of PSCs (53) and miR-200 family has been shown to inhibit the epithelial-to-mesenchymal transition (EMT), a typical somatic cell-process (54). Pluripotent stem cells express a distinctive set of miRNAs that help to maintain their unique characteristics.

1.1.4. Cellular reprogramming.

Cellular reprogramming constitutes the mechanisms by which a differentiated cell is forced to change its gene expression and epigenetic signature in order to become a phenotypically different cell type. This process can naturally occur in the adult in response to disease, aging or injury, and it is due to the cellular plasticity that is present in both the stem and the adult cells. This process was demonstrated for the first time *in vivo* by Wolff and colleagues in 1985 (55). Those findings encouraged researchers to develop methods to transform terminal differentiated cell into pluripotent stem cells to obtain an indefinite source of healthy tissues and even to give rise to a new living organism. So far, three main mechanisms to achieve cellular reprogramming have been described: nuclear transfer, cellular fusion and transduction of TFs (Figure 4).

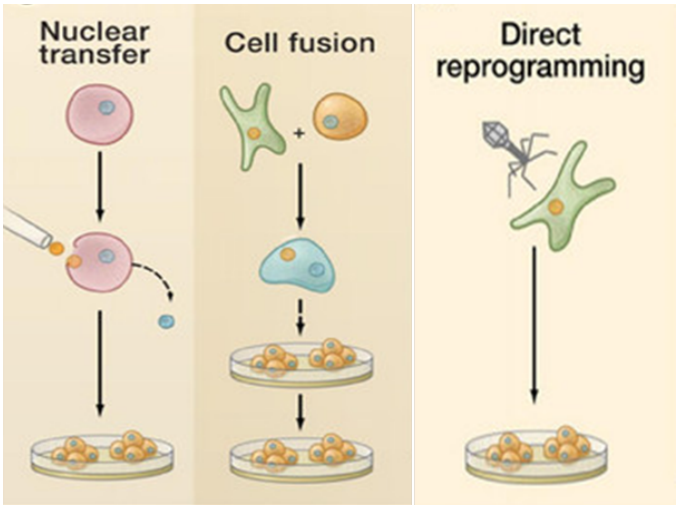


Figure 4. Cellular reprogramming can be achieved by three different approaches: nuclear transfer, cellular fusion and transduction of transcription factors. Extracted from (56).

1.1.4.1. Somatic cell nuclear transfer (SCNT) or Cloning.

This process consist in transferring the nuclei of a fully differentiated cell, like an epithelial cell from the intestine, into an enucleated "younger cell", like an oocyte, in order to generate a new organism that is an identical clone of the somatic cell donor. This concept was first proposed by Hans Spemann (57), but Robert Briggs and Thomas King developed the methods to transplant the nuclei of blastula cells into enucleated frog oocytes (58). John Gurdon extended this work transplanting nuclei from frog intestine into oocytes and producing clonal somatically derived adult frogs (59). It took three decades to transfer Gurdon´s method to mammals; work developed by Wilmut *et al.* cloning Dolly the Sheep (60) and a year later, the introduction of a piezoelectric pulse delivery pipette, allowed Wakayama *et al.* to clone mice (61). Recently, Tachibana *et al.* optimized SNCT conditions to achieve for the first time the derivation of human embryonic stem cells using this methodology (62).

Besides the great impact that SNCT could suppose for cell therapy, there are still main challenges that need to be overcome, like the low efficiency of the process, the genetic abnormalities caused in the derived organism, impaired immune systems, increased cancer susceptibility and premature death (63). Scientists accuse a remnant epigenetic memory of donor cells in the derived organism for these impasses (64). Even though, development of SNCT allowed the conclusion that all the genes needed to generate a pluripotent cell are present in the somatic cell, but specific cues are required to turn them on, like the TFs present in the oocyte and its environment, paving the way for further reprogramming approaches.

1.1.4.2. Cell Fusion.

This method involves the fusion of two or more cellular types to form one single identity, leading to nuclear reprogramming towards a specific cell type. This process can generate: i) hybrids, able to proliferate and contain one nuclei or ii) heterokaryons with multiple nuclei incapable of dividing (65). Cell fusion experiments have helped to conclude that: i) a particular cellular state, like pluripotency, can dominate a differentiated one (66); ii) gene expression is regulated not only by *cis*-acting DNA elements but also by trans-acting repressors (67); and iii) silent genes in a somatic cell can be activated (68). Tada *et al*., showed that somatic cells can acquire a pluripotent state after being fused with ES cells in mice (69) and posteriorly, Cowan *et al*., achieved nuclear reprogramming of human somatic cells by fusing them with human ES cells in tetraploid hybrids (70). Nevertheless, when cell fusion studies are based on replicative hybrids, they usually have chromosome abnormalities.

1.1.4.3. Transduction of transcription factors.

The overexpression of specific TFs can change the cellular fate of a somatic cell. Two different approaches can be defined, the pluripotency induction by defined factors and the cell trans-differentiation. The great breakthrough in the pluripotency induction was the generation of induced pluripotent stem cells (iPSCs) by simultaneous introduction of four TFs (*Pou5f1 (Oct3/4), Sox2, Klf4,* and *c-Myc*) transforming mouse fibroblasts into cells that were nearly identical to mESCs (26). This discovery changed the classical biology dogma where differentiated cell was considered as a terminal state. Furthermore, Shinya Yamanaka was awarded, together with John

Gurdon, with the Nobel Prize in Medicine for the discovery that mature cells can be reprogrammed to become pluripotent (71).

On the other hand, cellular trans-differentiation aims to convert a terminal differentiated cell into another specialized one by the ectopic expression of tissue specific factors without reaching a pluripotent state. Although trans-differentiation has been described in cells from totally different origins, this process reaches the highest efficiency on closely related cell types (72,73). Several groups have used the factor-mediated conversion to obtain insulin producer beta cells from pancreatic exocrine cells (74) and neurons (75), hepatocytes (76) or cardiac myocytes (77) from fibroblasts.

Induction of pluripotency by defined factors and iPSC generation is reviewed in more detail in the following section.

1.2. INDUCED PLURIPOTENT STEM CELLS (iPSCs).

There are more than 100 genes related to pluripotency, however, in the first reported reprogramming experiment towards pluripotent state, 24 TFs were selected and overexpressed using retroviral vectors. Using a reporter gene for pluripotency screening, the list was narrowed down and only four factors (4F), Pou5f1, Sox2, Klf4, and cMyc, were sufficient to generate iPSCs from mouse fibroblasts. Those miPSCs were similar to mES cells in morphology, gene profile and differentiation capacity (26). Moreover miPSCs were able to generate all-iPS mice by tetraploid complementation contributing also to germ layer cells (78,79). Subsequently, human iPSCs (hiPSCs) were generated using the same methodology with identical combination of factors (80) or by lentiviral induction of POU5F1, SOX2, NANOG, and LIN28 factors (81).

1.2.1. Reprogramming mechanisms.

Within the last years it has been shown that reprogramming to the pluripotent state using 4F cocktail is a complex process characterized for poorly defined events. However, the most recent model of reprogramming establishes it like a process that comprehends two phases: i) a stochastic phase of gene activation and ii) a hierarchical phase that begins with SOX2 locus activation (82).

In general, once somatic cells are induced with 4F a stochastic turn on of gene expression is carried out, leading the somatic cell to several possible fates, including, apoptosis, senescence, transdifferentiation or reprogramming. Those random cells that will reprogram start to proliferate, initiate mesenchymal-to-epithelial transition (MET), undergo changes in histone modifications at somatic genes and activate DNA repair and RNA processing. In this phase, cells undergo a stochastic activation of pluripotency markers (83), a temporary onset of developmental regulators (84) and activation of glycolysis (85), constituting an intermediate state with an unknown rate-limiting step that delays the conversion to iPSCs (Figure 5). The stochastic model of reprogramming ensures that almost all somatic cells are prone to reprogram and those that fail to it, is because the events needed to defeat the epigenetic status of somatic cells in order to turn on a pluripotent state, are not carried out properly (86,87). Recently, different groups have suggested that there are multiple states due to the different steps present during cell reprogramming. Transitions between them are determined by chromatin modification, genome methylation, specific miRNA expression and cellular proteome (88-92) (Figure 5).

Later on in the reprogramming process probabilistic events decrease and hierarchical events increase as the cell progresses from a somatic to an iPSC and leads to the activation of *Sox2* locus. This last, initiates the second phase of reprogramming triggering a series of deterministic events like the activation of ESCs markers, such as undifferentiated embryonic cell transcription factor 1 (*Utf1*), *Esrrb*, developmental pluripotency associated 2 (*Dppa2*) and *Lin28* that will finally lead to an iPSC formation. Eventually, the pluripotent state is stabilized by silencing of exogenous transgenes, acquiring an ESC-like phenotype, resetting of the epigenome and the activation of the core pluripotency circuitry (93-95). In human iPSCs, for example, TRA-1-60, DNA methyltranferase 3B (DNMT3B), and REX1 have been proposed as the more stringent markers of fully reprogrammed clones (96) (Figure 5).

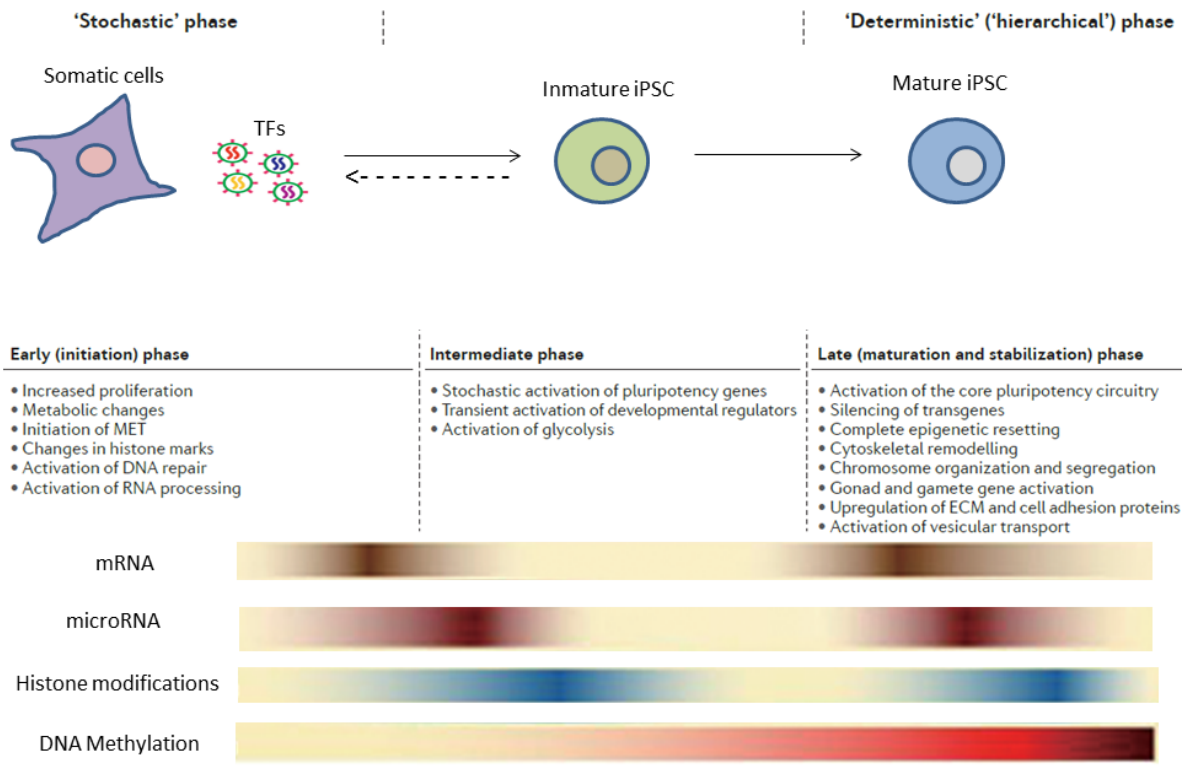


Figure 5. Three phases of reprogramming of somatic cells upon exogenous induction of TFs. First phase: stochastic where genetic manipulation induces changes in phenotype. An intermediated state is present within reprogramming comprising a population of no fully-reprogrammed cells that could be reversed to a differentiated status. Second phase or deterministic: characterized by maturation and stabilization of the pluripotent phenotype. Regulators of transcription like miRNAs, histone modifications and DNA methylation also vary along reprogramming. Image modified from (82,97).

1.2.2. Factors controlling cell reprogramming.

There are some aspects that need to be controlled when reprogramming somatic cells towards pluripotent state, like selecting the accurate **cocktail**, the **donor cell type**, the **culture conditions** or the **delivery method** of the reprogramming factors. All of them can affect significantly the efficiency of the process and the safety of generated iPSCs, determining their future clinical application. Reprogramming efficiency, measured as the number of iPSC colonies generated from a specific number of initial somatic cells, is an important factor when clinical applications are pursued, since the access to clinical samples is usually limited and the number of cells present in those samples is reduced. Here some of those aspects will be discussed with more detail.

1.2.2.1. Reprogramming cocktail.

As mentioned before, iPSCs were first generated from mouse fibroblasts by the overexpression of 24 TFs that were reduced to only 4, Pou5f1, Sox2, Klf4, and cMyc, known as Yamanaka´s factors (OSKM). Those cells resemble mESCs in their capacity to self-renew and differentiate. Later on human somatic cells were reprogrammed using viral particles for the overexpression of both

Yamanaka´s (POU5F1, SOX2, KLF4 and cMYC) (80) and Thomson´s factors (POU5F1, SOX2, NANOG and LIN28) (81). Nevertheless, the system used is not highly efficient and safe when facing a clinical application. Thus, recent studies have been dedicated to increase iPSC safety as well as to increase reprogramming efficiency, including the substitution of some factors, the overexpression of new ones, their stoichiometry or the discovery of small molecules.

As example, KLF4 and c-MYC TFs have an oncogenic nature, thus their presence in the reprogramming cocktail is undesirable. Nevertheless, iPSCs can be generated without c-MYC, (98) or even including just two factors, POU5F1 and SOX2 (99), allowing the generation of safer iPSCs. Overexpression of E-cadherin, the most relevant marker of the Mesenchymal to Epithelial transition (MET), was able to replace Pou5f1 during reprogramming (100).

Regarding the improved reprogramming efficiency, the presence of vitamin C or a soluble Wnt protein during the reprogramming process, as well as the incubation in a hypoxic environment has been shown to increase the frequency of iPSC generation (101,102). By disrupting the signaling pathways mediated by the tumor-suppressor protein p53 or the cell-cycle regulator INK4A, the iPSCs can be generated more efficiently (103,104). Since epigenome plays an important role in somatic cell reprogramming, molecules that modify chromatin status or DNA methylome, like BIX-01294 (a G9a histone methyltransferase inhibitor) (105), valproic acid (VPA) (106) or 5'-azacytidine (107), have been shown to improve iPSCs generation. Moreover reprogramming of mouse somatic cells by using only small molecules has been recently described (108,109).

1.2.2.2. Donor Cell type.

When thinking in the generation of hiPSCs for clinical applications, some characteristics regarding donor cell type are important. In general, parental cells for reprogramming must be easily accessible with minimal risk procedures, available in large quantities, have a relatively high reprogramming efficiency, and fast iPS cell derivation.

Initially iPSCs were generated from MEFs and adult fibroblasts. However, up to date pluripotency by defined factors has been induced in many different cell types, including pancreatic β cells, keratinocytes, hepatocytes and stomach cells (110). iPSCs have been also generated from different species, including human (80), mouse (26), dog (111), rat (112), pig (102), sheep (113) and monkey (114). Nevertheless, it has been proved that donor cell type can influence the reprogramming capacity as well as the differentiation potential of iPSCs. This was demonstrated when neonatal or juvenile human keratinocytes showed an approximately 100-fold improvement in reprogramming efficiency using retroviral OSKM compared to cells from adult donors (115).

Researchers agree that the cell type also affects the iPSC capacity to differentiate. George Daley´s group tested the potential of the cord-blood (CB-iPSCs) and keratinocytes derived iPSCs (K-iPSCs) to differentiate into keratinocytes. Interestingly K-iPSCs showed enhanced keratinocyte potential relative to CB-iPSCs. On the other hand, CB-iPSCs consistently yielded a greater frequency of hematopoietic colonies than multiple clones of K-iPSCs. These differences were due to the inability of iPSCs to erase their epigenetic memory during reprogramming (116). It has been described that early-passaged iPSCs prefer to differentiate into the cell lineage of origin (84).

1.2.2.3. Cell culture method.

Cell culture method is other important aspect to be considered when generating iPSCs. Initially iPSCs cultivation utilized almost the same culturing conditions as those for ES cells (18). Nowadays, there is a wide offer of culture media and reagents for iPSC culture and the decisions of the researchers are based principally on economy and most importantly in the final application of the iPSCs generated.

In terms of the **growth medium** for future clinical applications, the most accurate choice is a serum-free, xeno-free, and chemically defined medium. Knock-Out Serum Replacement (KSR) is widely used with Fibroblast Growth factor-2 (FGF-2) for feeder-based cultures (117). On the other hand for feeder-free conditions defined culture medium, TeSR1 (118) and defined E8 medium (E8), were developed (119). Other parameter to consider is the **environmental cues**. It has been proved that changes in temperature, humidity, osmosity, acidity, rigidity of growth surfaces, cell density or gas diffusion exchanges can influence growth of PSCs and influence the maintenance of the pluripotent state. (120,121). A third aspect to consider is the **type of culture**. hPSCs can be grown in colonies, monolayer or suspension. Depending of the application, each type of culture needs specific extracellular components (i.e. animal cells, hydrogel, individual matrix proteins and synthetic surfaces) and presents pros and cons. Colony type cultures need a substrate that can be synthetic or biological based. Feeder layers using MEFs are the best characterized and are economical, although for clinical applications, the presence of mouse factors is undesired. Matrigel is a highly used covering for feeder-free cultures, whose components are partially defined. Monolayer culture or single cell plating is useful for increasing number of cells and for transfection or infection of hPSCs (i.e. gene editing applications). This type of culture is usually carried out on Matrigel covers and requires the presence of various small molecules (e.g., ROCK inhibitor (ROCKi) and JAK inhibitor (JAKi)), which could increase the cost (122,123). Lastly, hPSCs can be grown on suspension. This type of culture does not require any substrate, the growth conditions can be monitored and allows the use of bioreactors (123-125).

1.2.2.4. Reprogramming methods.

The delivery method of the TFs is an important aspect that must be considered for iPSC generation since it is directly related to the future application. Delivery methods can be divided into two classes, those involving the integration of exogenous genetic material and those involving no genetic modification of the donor cells. Each of them has advantages and disadvantages in terms of reprogramming efficiency and safety for future clinical applications of iPSCs (Figure 6).

1.2.2.4.1. Integrative methods.

The first iPSCs were generated by using integrative methods, in particular with retroviral vectors expressing the different reprogramming factors. In general integrative methods have been shown to be more efficient generating iPSCs and their use is widely established. These methods can be subdivided into:

Retroviral and Lentiviral delivery systems.

Both retroviral and lentiviral systems have been successfully used to reprogram somatic cells. Moloney murine leukemic virus (MMLV) is the commonly used retroviral vector for inducing

pluripotency factors and HIV derived virus constitute lentiviral system (81,126). The efficiency of iPSCs generation with viral systems is ~0.1% in mouse embryonic fibroblasts and ~0.01% in human fibroblasts (127). Retroviral systems can only be used in actively dividing cells when lentiviral particles infect both dividing and non-dividing cells. The greatest bottleneck of using this method is the lack of safety due to randomly integration of the viral genome within the host genome and consequent possibility of tumor generation (79). Tet-inducible (128) and cre-deletable (129) systems have been developed to avoid this complication. Although these approaches can reduce the oncogenic risk and increase the differentiation capacity by the deletion of pluripotent transgenes (130), the reprogramming efficiency is dramatically reduced using inducible vectors and in both cases these advances are not enough to ensure a safe generation of iPSCs for clinical applications.

PiggyBac transposon.

This system consists on the transfection of two plasmids, the transposon, containing the sequences of interest (e.g. reprogramming factors), and the transposase, which is in charge of cut-and-paste the transgenes along the donor cells genome. However some transposases allow the transgene jumping from site to site into the genome of the somatic cell, activity that could lead to genetic abnormalities (131,132) and making critical the length of the transposase expression. PiggyBac transposon system has been reported to be efficient for reprogramming murine and human cells due to the increased number of transduced cells. Even though there has been reported a 100 fold increased in efficiency compared to classical viral induction, it is considered low (133,134).

Transfection of DNA plasmids.

Transfection of mammalian expression plasmids, coding for the correspondent pluripotency factors, using chemical or physical systems avoids the use of viral particles. The main advantage of this methodology is the reduction of genome integrations. However primary cells are often difficult to transfect requiring high number of donor cells to achieve reprogramming, which is not always available. Additionally to the decrement in efficiency of reprogramming there have been reported a high rate of not fully reprogrammed clones generated with this system (135).

1.2.2.4.2. Non-integrative methods.

Non-integrative vectors represent a good alternative when facing clinical applications because these systems avoid genomic modifications allowing the generation of safer iPSCs in terms of therapy. Nevertheless, researchers are still struggling with their intrinsic inefficiency and their low reproducible capacity. Below we mentioned the strategies developed in this field, ranging from viral delivery systems to protein and RNA mediated-induction of pluripotency.

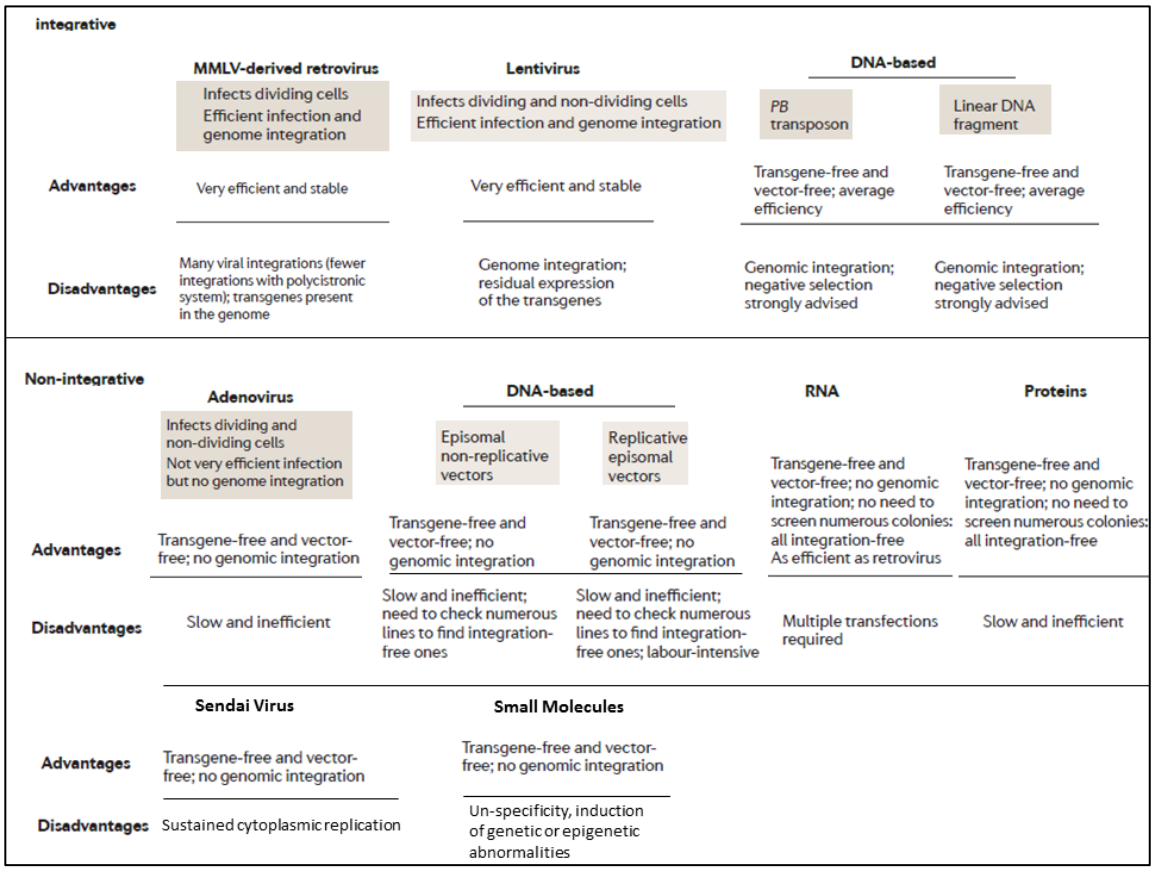


Figure 6. Advantages and disadvantages of the delivery methods described so far to induce pluripotency in somatic cells. Integrative methods constitute the most efficiency alternative described so far, when non-integrative are safer for clinical applications. Table extracted from (127).

Adenoviral vector delivery.

Human and mouse iPSCs can be generated using non-integrative viral systems like adenovirus. However, the reprogramming efficiency (between 0.0001% and 0.0018% respectively) is lower than integrative methods Even though adenoviral vectors have been described as non-integrative, random integration of the viral DNA genome can occur in a very low percentage, requiring further analysis to select integration-free iPSCs (136,137).

Episomal delivery.

Episomal vectors are non-integrative DNA plasmids that once introduced into the cell could replicate as episomes. Yu and colleagues successfully generated iPSC colonies from human somatic cells with the optimized replicative mixture that contains an episomal vector (oriP/EBNA-1) derived from the Epstein-Barr virus backbone (138). These experiments showed low efficiency due to the inherent poor transfection of large plasmids. In order to avoid these adverse effects, non-replicative mini-circle vectors were developed and the generation of iPSCs from human adipose derived stem cells (ADSCs) with an efficiency of 0,005% was reported (139).

RNA delivery.

Warren *et al.*, described a 17-day transfection protocol for delivering synthetic mRNA of POU5F1, SOX2, NANOG, LIN28, KLF4 and c-MYC TFs. Even though this is the most efficient method described so far (approximately 2%) it entails a high oncogenic risk due to the elevated gene dosage of the method (140). Other approach was performed by Miyoshi *et al.*, who were able to reprogram mouse and human cells to pluripotency by direct transfection of mature double-stranded miRNAs. They used a combination of mir-200c, mir-302s and mir-369s families previously described as pluripotency miRNAs (141). The major challenges in these methodologies still the low efficiency and reproducibility in a large number of cellular types and species.

Protein delivery.

Two different groups have reported the use of recombinant proteins to induce pluripotency in somatic cells. The first group generated POU5F1, SOX2, KLF4 and c-MYC proteins in *E. coli* (142) and the second one produced four stable cell lines for protein production previous to apply the supernatant to reprogram somatic cells (143). Protein with the HIV trans-activator of transcription (Tat) or containing poly-arginine tags can be easily introduced to almost any cellular type (144,145). Although this method avoids integration of any transgene, its efficiency has been reported as the lowest one. Furthermore the generation of recombinant proteins is an expensive and complex process, which affects significantly the reproducibility of the technique.

Sendai Virus (SeV) vectors

SeV, a member of the *Paramyxovirdae* family, is an enveloped virus with a single-stranded, negative-sense RNA genome of ~15 kb. SeV vectors replicate only in the cytoplasm of infected cells and do not go through a DNA phase or integrate into the host genome representing a promising integration-free reprogramming system. SeV vectors have been proved to efficiently generate human iPSCs from fibroblasts (146) and blood cells (147). Although these vectors are slowly diluted during iPSCs division the issue of the sustained cytoplasmic replication remained in question. In order to solve this problem temperature-sensitive (TS) SeV vectors have been generated recently. Temperature-shift treatment of these vectors allows the elimination of the remaining SeV constructs in reprogrammed cells (148). Currently, SeV vectors constitute the best alternative to generate integration-free iPSCs for clinical and basic applications.

Small Molecules delivery

The chemical induction of reprogramming has emerged as a new strategy with a potential use in iPSCs generation for clinical applications. The main advantage of this system is that small molecules can target different cellular pathways allowing the control of cell fate, state and function (149). Usually the small molecules used in reprogramming protocols are unspecific modulators of DNA status and chromatin modificators, which could lead to genomic abnormalities. Additionally, the efficiency on the number of iPSC clones generated still very low (150). Despite these facts, the combination of seven small-molecule compounds allowed reprogramming of mouse somatic cells in a frequency up to 0.2% (108). This methodology brought to reprogramming field the controversy that "master genes" are dispensable for cell fate reprogramming. Small molecules could be considered as a powerful reprogramming tool because

they can be restricted temporally and spatially but ensuring chromosomal normality in iPSCs obtained still a challenge in this area.

1.2.2.5. Conclusions of cell reprogramming.

In order to generate harmless iPSCs and to improve the efficiency of the reprogramming process for further clinical applications, a large number of strategies have been developed. Those strategies differ in the transcription factor (or molecule) cocktail used for reprogramming, the donor cell type, the culture conditions and the delivery methods of the selected reprogramming cocktail. Precise selection and management of all of these strategies may vary largely the final outcome of iPSC generation, resolving the different issues related to cell reprogramming.

Regarding the **reprogramming cocktail**, the use several factors ensure high efficiency only when the balance among the individual reprogramming factors is reached. When choosing a less-factor method a delay in the reprogramming process and the presence of not-fully reprogrammed clones is possible. The inclusion of small molecules to the cocktail at certain points of the process could increase the efficiency if needed.

Donor cells should be easily accessible and available in large quantities. Cell type can influence the efficiency of the reprogramming process. Depending on the delivery method a high proliferation rate is desirable and endogenous expression of MET markers or reprogramming factors could enhance the efficiency and iPSCs derivation. When a cell therapy application is pursued, the obtaining of donor cells from the same target germ layer is as a good strategy. This is due to the epigenetic memory maintained even after reprogramming which would facilitate differentiation process.

Additionally, there is a wide offer of iPSC **culturing media** varying from the classical colony growing to maintain the un-differentiated state, to suspension growing for expansion and differentiation. Monolayer culture has been also developed and in general for clinical applications serum-free, xeno-free, and chemically defined medium is the best option.

In terms of the **delivery method**, the most efficient technologies like *integrative systems* produced un-safe iPSCs in terms of their oncogenic potential *in vivo*, low differentiation capacity and culture heterogeneity. On the other hand some of the non-integrative approaches are difficult to apply due to poor transfection efficiencies, poor cell survival or the need of long reprogramming kinetics. Also, *non-integrative* DNA vectors seem to have a low but present integrative ratio into de host genome. In general, analysis of the mutation load of iPSCs generated using any of these approaches is required.

Despite on the advances in technology and innovative proposals, the field of cellular reprogramming (both iPSC generation and cell trans-differentiation) continues in development and the choice of the appropriate methodology seems to depend on the ultimate application, always looking for a balance between efficiency and safety.

1.2.3. ESCs vs. iPSC.

A fundamental question for therapeutic applications of iPSCs is: which are the relevant differences between iPSCs and their reference cell type, ESCs? Both cell types share basic properties of self-renewal and potential of differentiation, however, recently studies suggest considerable dissimilarities between these two pluripotent cell types.

Researchers have suggested that human iPSCs obtained with the overexpression of Yamanaka´s factors, belong to primed pluripotent stem cells (10). Even though the first generation of mouse iPSCs had pluripotent features, they were not able to give rise all-iPS mice by tetraploid complementation (26), suggesting a primed pluripotent state. Further experiments led to generation of miPSC in what is now called Naïve or ground state pluripotent stem cells, those who are able to give rise to a full organism when implanted into a blastocyst (78). Those murine iPSCs share all defining features with naïve mouse ESCs, while human iPSCs resemble more the primed state of EpiESCs (151). Recently, Gafni *et al.* defined conditions to maintain *in vitro* the hiPSCs in a naïve state, which could resemble naïve mESCs (152). Thus, the pluripotent state of iPSC also seems to be dynamic because cells can move between naïve and primed states when changing culture conditions (see 1.1.2).

Differences in gene expression between iPSC and ESC have been examined by microarray and RNA-SEQ technologies. iPSC clones from four distinct laboratories showed to cluster according to the lab of origin rather than to their identity or method of derivation (153). Although iPSC and ESC transcriptomes are practically identical, the expression profile of iPSCs is characterized by some residual expression of somatic genes of donor cells as well as by a weaker induction of pluripotency genes in comparison to ESCs (154-156). These features generally affect the differentiation capacity of pluripotent cells.

Some researchers believe that an extended time in culture could help to fade away the differences with the risk of causing chromosomal abnormalities (157). Some studies have indicated that the residual effect of the donor cells arises from epigenetic memories of the parental cells that persist during reprogramming. Several differentially methylated regions and histone modifications among iPSCs and ESCs have been identified. Moreover ESCs and iPSCs are not identical in the differentiation potential due to genetic and epigenetic features apparently determined by donor cells and culture conditions (155,156,158).

1.2.4. iPSC applications.

iPSCs are an attractive source in biomedical field due to an important spectrum of possible clinical applications including regenerative medicine, disease modeling, tissue engineering and/or drug (small molecule) testing/development.

In **regenerative medicine**, a large number of iPSCs from a patient could be differentiated to almost any lineage in order to be transplanted back. For example, iPSCs could be generated from fibroblasts derived from a Parkinson´s patient, then could be differentiated into healthy neurons using specific protocols and finally transplanted back into the patient to treat the disease.

iPSCs are very useful for **modeling *in vitro* a disease**, including the possibility to screen new drugs using those models (Figure 7). iPSCs derivation from specific genetic (159,160), metabolic (161) and neurodegenerative diseases (162,163) has been achieved by several groups. Thus, in vitro disease modeling could help to a better understanding of disease mechanisms like Alzheimer (164)

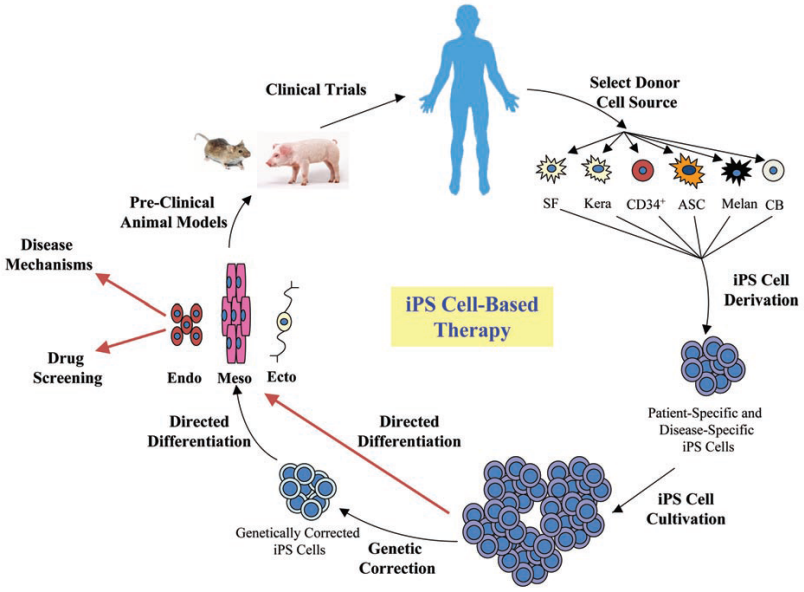


Figure 7. iPSCs applications include cell therapy, genetic corrections for regenerative medicine, drug screening and disease modeling. Defining the iPSCs application will determine specific aspects like donor cell type, culture media and delivery method. SF, skin fibroblasts; Kera, keratinocytes; CD34+, CD34+ cells from peripheral blood; ASC, adipose stem cell; CB, cord blood cell; Endo, endoderm; Meso, mesoderm; Ecto, ectoderm. Image taken from (165).

PSCs have brought closer the ***in vitro* tissue generation** by the development of three-dimensional structures of the retina (166) pituitary gland (167), some parts of the brain (168) as wells as vascularized human livers (169). Recently gastric organoids have been developed from PSCs (170).

iPSC are also very useful for screen chemicals and natural derivatives for the **development of effective drugs** (171). On the other hand, some groups have tested the efficacy of drugs or small molecules for correcting or improving the disease phenotype, as shown for spinal muscular atrophy (172) and Rett syndrome (173).

The first clinical trial using human cells derived from iPSCs is in progress. Skin cells from patients suffering age-related macular degeneration (AMD), a degenerative disease that affects at least 25–30 million people worldwide, constituting the third cause of blindness, (174) were reprogrammed to iPSCs and posteriorly differentiated into retinal pigment epithelial cells. These cells have been re-introduced into one patient suffering AMD, and the efficacy of the treatment is under evaluation (175,176). The application of iPSCs is always on the quest for a balance between safety and efficiency. This balance sways among the main barriers of this methodology that are: the use of oncogenes as reprogramming factors, the possible generation of genetic abnormalities due to delivery system and the lack of reproducible and efficient differentiation protocols.

1.3. LNCRNAS AND CELL FATE REGULATION.

1.3.1. LncRNAs: definition

Even though 70-90% of the human genome is transcribed only a 2% corresponds to coding genes, observation that has been a breakthrough on the classical dogma of molecular biology (177). Several non-coding RNAs are well known, as transfer RNAs (tRNA) and ribosomal RNAs (rRNA) required for translation, or small nuclear (snRNAs) and nucleolar RNAs (snoRNAs) needed in splicing and RNA maturation respectively. Recently, other small RNAs, as miRNAs, piwiRNAs (piRNAs) and small interfering RNA (siRNAs) have been brought to light as regulatory transcripts and transposon silencers (178), but the vast majority of non-coding RNA do not have a described function yet.

Long-non coding RNAs (lncRNAs) are defined as transcripts longer than 200 nucleotides (nt) with non-coding function. Since most of them are RNA polymerase II products, nearly always they are capped at the 5' end, frequently spliced and polyadenylated (approximately 40%) (179). LncRNAs are similar to protein coding genes at several levels including their transcriptional and post-transcriptional regulation, chromatin modification patterns, and splicing signals (180-185).

LncRNAs are often defined by their relative position to neighboring protein-coding genes as: i) **Antisense lncRNAs**, transcribed in the opposite direction of protein-coding genes, ii) **Intronic lncRNAs**, that initiate inside of an intron of a protein-coding gene in either direction and terminate without overlapping exons, iii) **Bidirectional lncRNAs** that initiate in a divergent manner from the promoter of a protein-coding gene; generally within a few hundred base pairs and iv) **Intergenic lncRNAs (lincRNAs)**, with separate transcriptional units at least 5 kb away from protein-coding genes (186). In this work we will be mainly focused in the identification and characterization of new lincRNAs associated to pluripotency.

In order to differentiate a lincRNA from a mRNA the following criteria have been defined: (1) lincRNAs do not overlap with known protein coding genes, which tend to be much longer; (2) usually ORFs of lincRNAs are <50 aminoacids and ORFs have a random codon usage; (3) nucleotide substitutions in lincRNAS are not highly monitored trough evolution (184); (4) lincRNAs does not contain known protein domains and (5) lincRNA regions do not have sequence similarities to known proteins when evaluated in databases. Different combinations of these five points could help to identify potential lincRNAs from mRNAs (179,187).

1.3.2. lncRNAs: Mechanisms of action.

LncRNAs have been described as key regulators of development and may play relevant roles in cell homeostasis and proliferation. LncRNAs exert these functions by physical interacting with DNA, other RNAs and proteins either by nucleotide base-pairing or by forming scaffolds generated by RNA folding (188). LncRNA functions are not exclusive of a specific cellular process because one lncRNAs can performed different functions depending on the cellular type or tissue and even of their developmental stage (178). Herein we review some of the general and possible mechanisms of action of lncRNA, even though each lncRNA function must be proven for the tissue and conditions of interest (Figure 8.)

LncRNAs have been proven to **regulate epigenetics** by recruiting chromatin repressive complexes to neighboring genes (*cis*). Inactivation of one of the chromosomes X in females is performed by this mechanism by lncRNA Xist (189). Conversely, LncRNA HOTAIR uses the same mechanism to regulate transcription of HOXD gene, located in another chromosome, being the first report of the action of a lncRNA in *trans* (190).

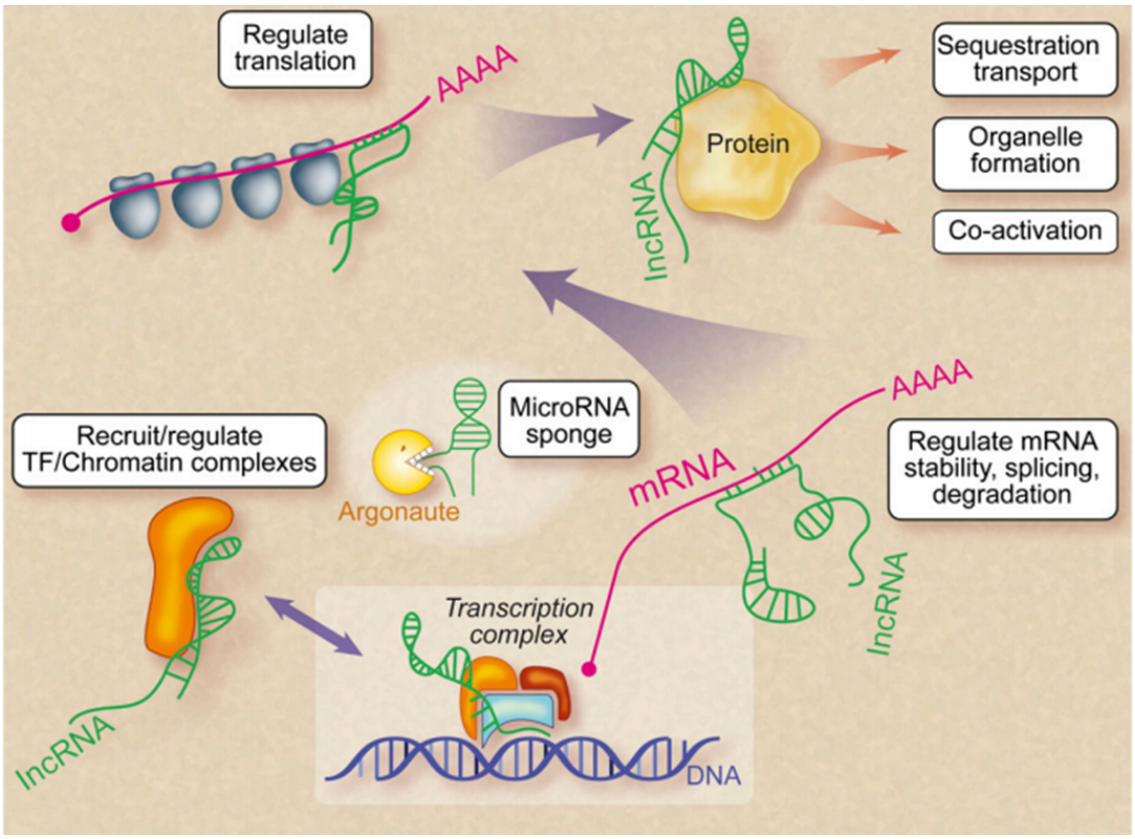


Figure 8. Mechanisms of action of lncRNAs. LncRNAs (in green) are able to directly interact with DNA, RNA or proteins. This feature confers to lncRNA their function as regulator of gene expression at multiple levels such as chromatin, transcription, mRNA, translation, protein and post-transcription. Extracted from (188)

In terms of interaction with other RNAs, lncRNAs can form RNA-RNA hybrids by **antisense RNA base pairing.** Those hybrids can both repress expression of specific genes by inhibiting the interactions required for gene expression or activate gene expression by blocking inhibitory interactions (i.e. lncRNA-ROR) (191).

In addition, lncRNAs can serve as **scaffolds** to bring two or more factors together to regulate gene expression (i.e. lncRNAs TERC, HOTAIR and NEAT-1) (192) or to form complex protein structures to modulate enzymatic activity (193).

A large number of lncRNAs can exert a ‘‘sponge’’ function to **decoy the activity of a protein** activity. For example, lncRNA B2 binds directly to RNA polymerase II to repress transcription in response to heat shock (194). LncRNAs implicated in nuclear arrangement like organization of chromatin domains and chromosomes looping have been also reported and their number increases with the years (195,196).

1.3.3. LncRNAs and tissue specificity.

LncRNA can bind to regulatory proteins giving rise to cell-type-specific RNA–protein complexes and unique transcriptional programs. Additionally, specific expression programs on cells could be onset by interactions between lncRNAs and cell-specific TFs. LncRNAs have been studied in numerous cellular processes revealing their functional importance in epigenetic activation of gene expression during vertebrate cell-fate determination and disease playing crucial biological roles during embryonic development despite limited sequence conservation.

Several groups have identified lncRNAs whose expression or repression can lead the specification of PSCs towards a specific germ layer like mesoderm (197) or ectoderm (198) or even govern PSCs differentiation towards a specific cell type like, skeletal muscle cells (199), cardiomyocytes (200) or cells from the central nervous system (201). Also, expression of certain lncRNA protects the heart from hypertrophy and failure (202), can be implicated in Huntington´s (201) or Duchenne´s Muscular Dystrophy diseases (199) (Table 1).

Differentiation of PSCs towards a desired lineage is a complex and poorly understood process. The mechanisms involved within and all the molecules implicated like lncRNAs must be revealed in order to achieve reproducible protocols for future therapeutic applications of PSCs.

1.3.4. LncRNAs in pluripotency and reprogramming.

LncRNAs have been described to play a role in the modulation of the pluripotent state of stem cells and the maintenance of this feature by different mechanisms, including the direct control of the pluripotent TFs and repressing specific differentiation lineages. Recently their implication in the reprogramming of somatic cells has been also reported (203) (Table 2).

In terms of the control of pluripotency-related TFs, Guttman *et al*., pointed out in 2009, the first experimental validation of a lncRNA directly regulated by the binding of Pou5f1 and Sox2 to its promoter in mESC. This group also reported an unknown lncRNA which expression was clearly diminished during the differentiation process of mESCs using retinoic acid (204).

Within the last years several groups have reported other lncRNAs upregulated in PSCs of multiple species which mechanism of action in those cells involves a relationship with pluripotency TFs. The interaction could be direct, because those TFs bind to the promoter of the lncRNA (205-207), or indirect either because the lncRNA form a complex with other molecules that facilitates the interaction with pluripotency factors (201) or because they act like a sponge for antagonist molecules that inhibit the expression of TFs (208-210). Loss of function experiments of PSC-lncRNAs lead to evident phenotypical changes like loss of ES-morphology, impairment in cell growth and upregulation of lineage specific markers, indicatives of the onset of a differentiation process (201,205,207,209). In some case, overexpression of PSC-lncRNAs lead to the recovery of the pluripotent state (201) (Table 2).

So far these groups have proven the role of lncRNAs in the maintenance of the pluripotent state which is characterized by a unique capacity of the cell to self-renew, proliferate and differentiate to all the adult tissues. Reprogramming process, aims to reach this state by forcing the expression of the silent pluripotent machinery of an adult cell. Since lncRNAs seem to have a function in this

cellular stage, it would make sense to wonder whether lncRNAs also could be playing an important part in the reprogramming of somatic cells.

The questions of a possible use of lncRNAs as reprogramming factors and their roles within the changes at genetic and epigenetic levels of reprogramming process remain unanswered. Rinn *et al.*, performed the first study directed to know if the large-scale transcriptional changes induced by reprogramming apply to lncRNAs in human cells. They designed a probe based-microarray to analyze the expression of several lncRNAs in human fibroblasts (hFib), hESCs and hiPSCs and found a large number of differentially expressed transcripts between pluripotent and somatic human cells. Finally 3 of them shared binding sites with pluripotency associated TFs and were chosen for further experiments. LincRNA-RoR, just one of the three candidates, was proven to increase the efficiency of the reprogramming process when added in the cocktail. Lnc-RoR, or regulator of reprogramming seems to play an important role on not just the maintenance of pluripotent cells of both human ESCs and iPSCs but in the establishment of these lasts through a mechanism that involves the inactivation of cellular stress pathway (207).

Silencing of the whole X chromosome in females and the opposite process, its activation, is performed by the lncRNA ''X-inactive-specific transcript'' (XIST/Xist) and its antagonist, Txist respectively (211,212). The inactivation of the X chromosome in females is a characteristic of the differentiated state; the X chromosome reactivation is specific for the naïve pluripotency state. Expression of both transcripts Xist and Txist can be used as indicator of the quality of the cell clones obtained with the reprogramming protocol (213).

1.3.5. Concluding remarks.

Apparently in nature, the differentiated state of cells and in many organisms is not terminal but a stable state that under diverse environmental conditions such as injury, disease or even aging can be reprogrammed.

LncRNAs first declared as junk in the genome are known to play important roles in cellular fate of stem cells. Mechanisms of action of lncRNAs are diverse and one lncRNAs can execute different functions in one cellular type, but the vast majority of the functions of these transcripts are unknown. Specifically, there have been reported some lncRNAs related to the maintenance of the pluripotent state and reprogramming but how they do it still unidentified.

Mechanisms involved in reprogramming comprise a vast field of research that need to be explored due to the lack of a detailed and complete molecular characterization of the reprogramming process. This approach could help to refined methods of reprogramming that allow generating iPSCs that more closely approximate to the transcriptome, proteome and epigenome of the embryo-derived stem cells, strengthening the way to more reliable therapeutic applications with iPSCs.

Table **1** **LncRNAs and tissue specificity**

LncRNA	Specie	Described function	Mechanism of action	Reference
MISTRAL (Mira)	Mouse	Specification towards mesoderm derived tissues and organs	Hox genes activation of transcription by the recruitment the epigenetic activator MLL1 to chromatin	(197)
Anti-differentiation lncRNA (ANCR)	Human	Inhibition upregulates a global gene expression program associated with epidermal differentiation	Unknown	(198)
H19	Mammals	Regulation of cell growth and tumor repression	Inhibition of insulin-growth factor 2 (IGF2) and its receptor Igfr by miRNA-675	(214)
Muscle-specific long noncoding RNA (MD1)	Mouse and Human	Governs the time of muscle differentiation in myoblasts and its expression is altered in Duchene´s disease	Expression of MAML1 and MEF2C TFs that control numerous muscle-specific genes by inhibiting miR-133 and miR-135 antagonists of the process	(199)
Braveheart (Bvht).	Mouse	Cardiomyocytes differentiation and organization of a cardiac functional tissue, activation of both EMT gene network and core cardiovascular gene network	Interaction with SUZ12, a component of polycomb repressive complex 2 (PRC2)	(200)
Myosin heavy-chain-associated RNA transcript (Myheart or Mhrt)	Mouse and Human	Protects the heart from hypertrophy and failure	Mhrt is antagonist of aberrant gene expression and cardiac myopathy	(202)
Tcl1 upstream factor associated (TUNA)	Several species	Neurogenesis and its expression is altered Huntington´s disease	Forming a complex with RNA binding proteins to interact with Sox2	(201)
Pnky	Mouse and Human	Neurogenesis	Interacts with splicing regulator PTBP1 to regulate TFs related to neuronal differentiation	(215)

Table 2 **Pluripotency LncRNA**

LnRNA	Specie	Described function	Mechanism of action	Reference
Guttman1	Mouse	Related to pluripotency	Pou5f1 and Sox2 bind to lncRNA promoter	**(204)**
Guttman2	Mouse	Related to pluripotency	Unknown	(204)
ES1 and ES2	Human	Maintaining pluripotency of embryonic stem cells	Repressing SOX2 neural targets	(210)
AK028326 (Human orthologous: MIAT)	Mouse	Maintenance of the pluripotent state	Regulation of POU5F1 mRNA levels and mesodermal and tropho-ectodermal markers	(205)
AK141205	Mouse	Differentiation towards three germ layers	NANOG repression	(205)
Tcl1 upstream factor associated (TUNA)	Several species	Cell proliferation, upregulation of pluripotency markers and increment of reprogramming efficiency	Form a complex with RNA binding proteins to interact with pluripotency TFs	(201)
Uc.283-plus	Human	Embryonic development and maintenance of pluripotency	Sponge for miRNA related to differentiation and oncogenesis	(209)
SFMBT2	Human	Related to pluripotent state	Probably by interactions with pluripontency associated TFs	(207)
VLDLR	Human	Related to pluripotent state	Probably by interactions with pluripontency associated TFs	(207)
RoR	Human	Related to pluripotent state and regulator of reprogramming	inactivation of cellular stress pathway and sponge for miRNA directly related to inhibition of pluripotency TFs	(191,207)
Xist	Several species	Inactivation of X chromosomes in differentiated cells. Downregulated in reprogramming process	Xsist is controlled by the expression of lncRNA Txist and pluripotency TFs	(212) a
Txist	Several species	Antagonist of Xsist, expressed by naive pluripotent cells	Antisense partner of Xist	(213)

2. HYPOTHESIS & OBJECTIVES

HYPOTHESIS AND OBJECTIVES.

Induced pluripotent stem cells (iPSCs) are an attractive source of stem cells for clinical applications like patient-specific cell therapy, disease modeling and drug screening. Generation of iPSCs by cellular reprogramming is a complex and highly inefficient process, that involves a high number of genetic and epigenetic modifications most of them unknown. Those genetic modifications include both coding and non-coding genes like long non-coding RNAs (lncRNAs). These lasts have been found to be involved in a wide variety of cellular processes like embryonic development and disease. Recently, lncRNAs have been described as markers of the pluripotent state in addition to enhance the efficiency of reprogramming of somatic cells.

A detailed knowledge of the factors that make a cell to return to its own origin would help to a deeper understanding of the reprogramming process and to address new strategies that solve the current problems within this field, like low efficiency and the "poor" quality of hiPSCs generation. We hypothesized that there must be un-described human lncRNAs markers of the pluripotent state with a possible role in reprogramming of human somatic cells. Thus, at the present work we aimed to further our knowledge in the reprogramming mechanisms mediated by lncRNAs.

In order to elucidate this hypothesis we established three principal objectives:

1) To generate and characterize iPSCs from human fibroblasts and ADSCs,

2) To analyze lncRNA expression profile in generated ADSC-hiPSCs for revealing novel lncRNAs markers of the pluripotent state with a possible role in somatic cell reprogramming and,

3) To elucidate the mechanism of action of those novel lncRNAs associated to pluripotent state and reprogramming.

3. METHODS

3.1. CELL CULTURE.

3.1.1. Media preparation.

Cell lines and somatic cells.(MEFs, primary human fibroblasts, human fibroblast cell line BJ (ATCC, CRL-2522), NTERA-2 cl.D1 [NT2/D1] pluripotent human testicular embryonic carcinoma cell line (ATCC CRL-1973) and 293T cells (kindly provided by Dr. Boris Greber) were maintained in Dulbecco's Modified Eagle Medium (DMEM) high glucose (4.5 g/l) supplemented with 10% fetal bovine serum (FBS), 0.1 mM non-essential amino acids (NEAA), 2 mM L-glutamine, and 100 UI/ml X penicillin/streptomycin. NCCIT pluripotent embryonic carcinoma cell line (ATCC CRL-2073) were grown on Roswell Park Memorial Institute (RPMI) medium supplemented with 10% FBS, 2mM L-glutamine, and 100UI/ml X penicillin/streptomycin. ADSCs were maintained in Alpha minimum essential medium (α-MEM) supplemented with 10% FBS, 2mM L-glutamine, 100UI/ml penicillin/streptomycin and 1ng/mL of bFGF (Peprotech).

Human induced pluripotent cells (hiPSCs). hiPSCs were maintained in Knockout Dulbecco's Modified Eagle Medium (DMEM-KO) containing 4.5 g/L glucose and supplemented with 20% Knockout serum replacement (KSR), 0.1mM NEAA, 2 mM L-glutamine, 100 UI/ml P/S, 0.1 mM b-mercaptoethanol, and 5 ng/mL bFGF (Peprotech). Cells were expanded on irradiated MEFs feeder layers in the presence of 10 μM of ROCKi compound GSK269962A (AxonMedChem). For some experiments iPSCs were picked and cultured on feeder free conditions using matrigel-coated culture dishes and conditioned media from ADSCs. Cells were routinely passaged at a splitting ratio of 1:3 or 1:6 every week when cells reached confluence.

Embryoid bodies (EBs) were cultured on DMEM-KO containing 4.5 g/L glucose and supplemented with 20% FBS (Biochrom), 2 mM L-glutamine, 100 UI/ml P/S and 0.1 mM b-mercaptoethanol.

All cell culture reagents were purchased from Gibco except those specifically indicated.

3.1.2. Preparation of gelatin- and matrigel-coated dishes.

Embryomax Ultrapure water with 0,1% gelatin (Millipore) was added to the culture dish in a volume sufficient to cover the whole surface. Plates were incubated at 37ºC for 20 minutes. After removing the gelatin, cells were readily plated. Matrigel™ Basement Membrane Matrix (BD Biosciences) was thawed and diluted 1:1 on cold DMEM-KO (stock solution). A volume of 750 ul of the stock solution was diluted into 7,5 mL of DMEM-KO (working solution). The working solution was added to culture dishes in a volume sufficient to cover the surface and incubated 1h at room temperature or overnight at 4ºC. After removing the matrigel, cells were readily plated.

3.1.3. Preparation of ADSC's conditioned media.

ADSCs were expanded until confluence, irradiated at 50Gys and cultured again on hiPS cells media for 24 hours whereupon media was collected and frozen. This process was repeated every 24 hours for a maximum of one week. Prior using, conditioned media was filtered using a 0.2µm membrane (Corning) and 5 ng/mL of bFGF (Peprotech) was added.

3.1.4. Freezing/Thawing of cells.

iPSCs were frozen in a solution containing 45% KSR, 45% FBS and 10% dimethylsulfoxide (DMSO) (Sigma). Other cells were frozen in their respective expansion media with 10% DMSO. Cryovials were immediately stored at -80ºC in freezing containers (Nalgene) and transferred to liquid nitrogen for long-term storage. For thawing, the cryovial containing the cells was quickly transferred from the liquid nitrogen tank to the water bath at 37°C. Before the cells were completely thawed they were mixed with 12 mL of the specific culture media and spun down for 7 minutes at 1200 rpm.

3.1.5. Isolation of mouse embryonic fibroblasts (MEFs) and preparation of feeder layers.

Mouse embryos of 12.5-14.5 days post coitum were dissected and limbs, gut, internal organs and brain were removed while leaving the lower head intact. The embryos were minced in sterile PBS and incubated in 0.05% trypsin solution (Gibco) in 15-ml conical tubes for 10 minutes in a 37ºC water bath with occasional mixing by tube inversion. Neutralized supernatants were collected and the remaining tissue re-trypsinized for three or four more rounds, until all tissues had been sufficiently digested. Pooled supernatants were filtered using a 40 μm nylon cell strainer (BD Biosciences) and spun down at 600g for 8 minutes. The pellet was resuspended in fibroblast medium containing 1 ng/ml bFGF (Peprotech) for the first few passages. Cells were routinely grown in normal fibroblast medium and passaged for a maximum of five times before use (Figure 9a). MEFs within five passages were irradiated at 50 Grays and washed with PBS before replating or freezing. Feeder layers were plated at $5x10^4$ cells/cm^2 on gelatin-coated dishes one day before use. Cells were frozen at $2x10^6$ cells/vial, sufficient to cover a 6-well plate upon thawing. All animal procedures were approved by the University of Navarra Institutional Committee on Care and Use of Laboratory Animals (Ethic protocol # 089-13).

3.1.6. Isolation of human fibroblasts (hFibs) and adipose derived stem cells (ADSCs).

Human fibroblasts were obtained from skin biopsies of adult donors. The skin was cut into small pieces, and digested with 2 mg/mL of collagenase type I (Gibco). The tissues were subjected to two or three rounds of collagenase treatment until the tissue was completely digested. Cells were pelleted down by centrifugation at 600 x g for 10 minutes, seeded in gelatin-coated culture plates in fibroblast medium and expanded for a maximum of five passages before use (Figure 9b).

Adipose derived stem cells (ADSCs) were obtained from the stromal vascular phases of adipose tissue from adult donors. Adipose tissue was carefully separated from skin and vessels, minced until getting a semi-solid paste and digested with 2mg/mL of collagenase type I (Gibco) until getting a two phases. The lower phase containing the stem cells was filtered through a 100 μm and a 40 μm meshes (Falcon) and centrifuged at 600 x g for 10 minutes. The pellet was seeded in gelatin-coated culture plates in ADSCs media and expanded for a maximum of five passages before use (Figure 9c).

Human sample collection and further experimental procedures followed the ethical requirements of the Ethics Committee of the Clínica Universidad de Navarra that specifically issued a favorable report for the realization of the project (ref# 029/2011).

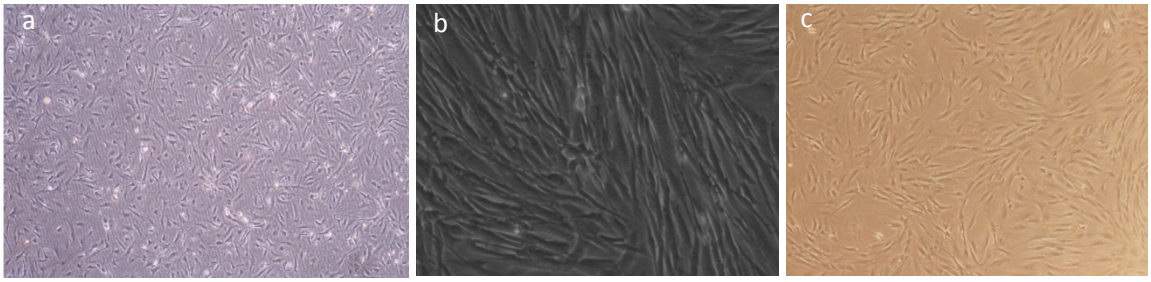


Figure 9. Primary cell culture | a) Mouse embryonic fibroblasts (MEFs) at optimal confluence for hiPSCs culture 10X, b) Isolated human fibroblasts, 20X c) Isolated human ADSCs, 10X.

3.2. GENERATION OF INDUCED PLURIPOTENT STEM CELLS.

All the procedures related to the generation, characterization and use of human iPSCs were supervised and authorized by the "Comisión de Seguimiento y Control de la Donación y Utilización de Células y Tejidos Humanos" (resolution #189-158-1) from the "Instituto de Salud Carlos III".

3.2.1. Retroviral generation of human induced pluripotent stem cells.

VSVG-coated MMLV retroviral vectors coding for the human reprogramming factors POU5F1, SOX2, c-MYC, and KLF4 (kindly provided by Dr. Boris Greber) were generated in 293T cells. For each viral production 293T cells were seeded one day before transfection in one well of a 6-well plate at $4.2x10^4$ cells/cm^2 in fibroblast medium. The following day 1 μg of gagpol plasmid and 0.1 μg of VSV-G plasmid (both plasmids kindly provided by Dr. Boris Greber) were mixed with 1 μg of the pMXs plasmid and transfected using Lipofoectamine2000 transfection reagent (Invitrogen) according to the manufacturer's instructions. Green fluorescent protein (GFP) cloned into the pMXs vector (kindly provided by Dr. Boris Greber) was used as control of transfection. One day before infection, hFibs and/or ADSCs were seeded at $5x10^4$ cells/well of a 6-well plate (Falcon). To infect, fresh retrovirus-containing supernatants were collected 48h after transfection and filtered through ultra-low protein binding 0.45 μm PVDF filters (Millipore). Equal parts of SOX2, c-MYC and KLF4 factors and double of POU5F1 were combined and used to infect ADSCs and/or hFibs cells in the presence of 4 μg/ml polybrene (Sigma). Cells were incubated overnight at 37ºC, 5% CO2. This procedure was repeated for three consecutive days on the same cells, to increase cell infection and 3 days after the last infection, media were replaced with fresh fibroblast media. The following day, $1x10^5$ infected cells were detached; using 0.05% trypsin (Lonza), plated onto previously seeded MEF feeder layers and 24h later culture medium was switched to iPSCs medium supplemented with 1 mM of VPA (Sigma) for one week. Cell culture media was changed every other day. Figure 10 shows a scheme of the reprogramming method used in this study.

3.2.2. Lentiviral generation of human induced pluripotent stem cells.

VSVG-coated doxycycline inducible lentiviral vectors (FUW-Tet-O based vectors) coding for the human POU5F1 (ref#20726), SOX2 (ref#20724), c-MYC (ref#20723), and KLF4 (ref#20725) TFs were generated in 293T cells as described (133). Briefly, for each viral production $5x10^6$ 293T cells were plated in 150 mm culture dishes two days before transduction. Cells were transfected with 9μg of FUW-TetO lentiviral vectors along with 6μg and 3μg of the packaging plasmids psPAX2 (ref#12260) and pMD2.G (ref#12259) respectively, using Lipofoectamine2000 transfection reagent (Invitrogen) according to the manufacturer's instructions. All lentiviral plasmids were obtained from Addgene repository kindly deposited by Dr. Rudolf Jaenisch (133). Virus-containing supernatants were collected 72 h post-transfection, filtered through ultra-low protein binding 0.45 μm PVDF filters (Millipore), pooled for 4 factor infections and supplemented with FUW-M2rtTA virus (ref#20342) in a ratio 2:2:2:2:1 (O:S:K:M:rtTA) in fresh culture medium. BJ human fibroblasts were seeded 24 h before transduction in T75 flasks (Falcon). Two consecutive infections were performed over a period of 48 h in the presence of 4 μg/ml of polybrene (Sigma). Culture medium was changed 12 h after the last infection. Five days after the first infection, fibroblasts were passaged using 0.05% trypsin solution (Lonza) and replated at different densities ($5x10^4$, 10^5 and $5x10^5$ cells/plate) on MEF feeder layers. To induce cell reprogramming, culture medium was replaced 48 h later by iPSC medium supplemented with 1 μg/ml doxycycline (Clontech) and with 1 mM of VPA (Sigma) was added for one week. Cell culture media was changed every other day.

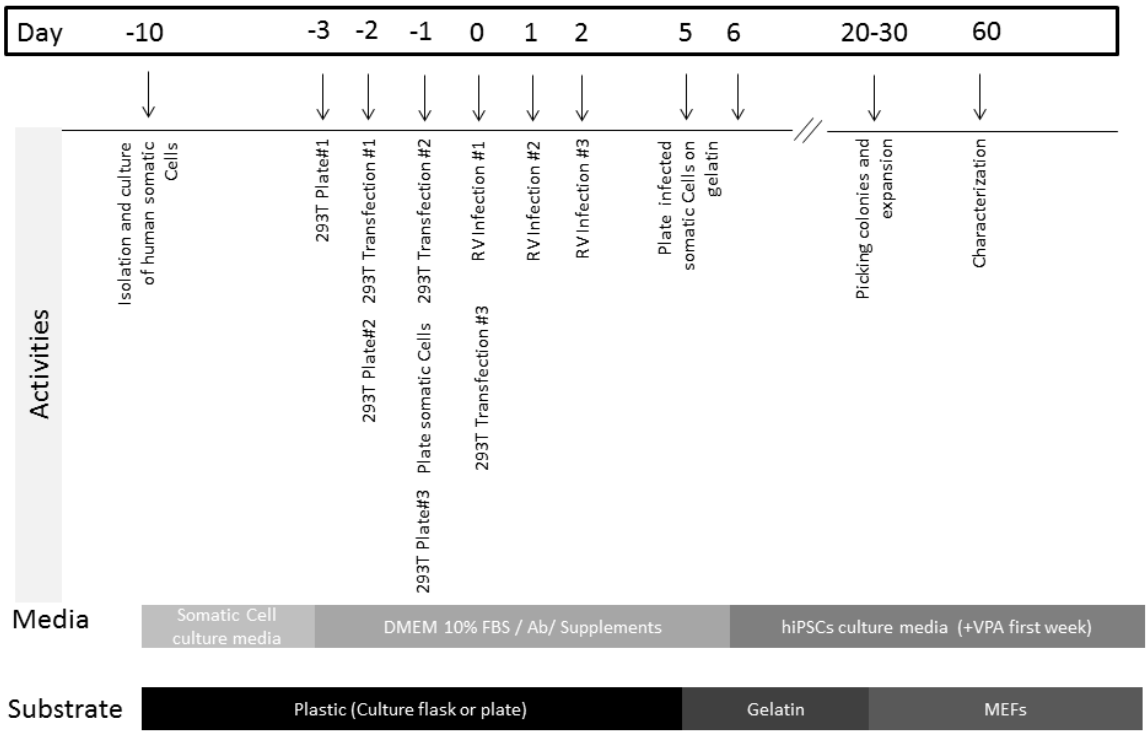


Figure 10. Reprogramming protocol of human somatic cells towards hiPSCs using over–expression of the four factors (4F).

3.2.3. Colony observation.

Cultures were carefully inspected under the microscope for the appearance of ES-like colonies that usually emerged around day 20-30 after infection. True iPSC colonies were identified by their small, round, and compact morphology with refractive or shiny appearance and tight, well-defined borders.

3.2.4. Picking-up.

By day 20-30, iPSC colonies were big enough and were ready for pick-up. Under an inverted microscope, in sterile conditions, colonies were picked-up using a 0.5-10 μl micropipette tip, scratching the feeders surrounding them, and aspirating each individual colony that was plated into previously seeded MEF feeder layers in a 12-well plate (Falcon). Once clones reached confluence were expanded to 6-well plate with MEFs previously incubated with 10uM ROCKi (#GSK269962A, AxonMedChem).

3.2.5. Culture of iPSCs.

iPSCs were maintained in standard culture conditions (section3.1.1). When cells reached 80% confluence or when they formed sufficiently large colonies, were treated with 1mg/mL Collagenase IV (Gibco) for 5 min and then switched to the iPSCs medium again. Colonies were cut into two to four pieces and passaged once a week at 1:3 or 1:6, either onto 2% Matrigel (BD) coated 6-well plates or with MEF-feeders plates previously treated with ROCKi. When density of cultures was low or a high number of spontaneously differentiated colonies were present, undifferentiated iPSC colonies were individually picked-up in the same manner as described above.

3.3. CHARACTERIZATION OF INDUCED PLURIPOTENT STEM CELLS.

3.3.1. Trizol reagent total RNA extraction.

Total RNA was extracted using Trizol reagent (Life Tecnologies) according to the manufacturer's instructions. Cells were lysed using 1 ml of Trizol pipetting up and down vigorously followed by 5-minute incubation at room temperature (RT). After adding 200 μl of chloroform, lysate was mixed by inversion and incubated at RT for 5 minutes before centrifuging it at 12.000 rpm at 4ºC for 15 minutes. After the spin, three layers could be observed. The upper aqueous supernatant, containing the RNA, was carefully aspirated into a new tube, to which an equal volume of isopropanol was added to precipitate the RNA. The mixture was incubated for over-night at -20C. When the amount of cells was low precipitation mix was incubated with 10 μg Glycogen (Life Technologies). Samples were next centrifuged at 12,000 rpm at 4ºC for 15 minutes. The translucent pellet was washed with 1mL cold 75% ethanol by vortex and centrifuged at 12,000 rpm at 4ºC for 10 minutes. The total RNA pellet was air-dried and subsequently resuspended in RNAse-free water. RNA yield was measured using the NanoDrop spectrophotometer (Thermo Scientific).RNA quality was tested using Bioanalyzer (Agilent) and RNA samples were stored at -80ºC.

3.3.2. cDNA synthesis, primer design and quantitative qPCR.

Two micrograms of RNA samples were treated with DNAse (Fermentas) (37°C for 30 minutes) and were retro-transcribed into complementary DNA (cDNA) using the Superscript II Reverse Transcription Kit (Invitrogen) according to the manufacturer's instructions. cDNA solutions were diluted to a final volume of 200 μl with distilled water.

The primers used for quantitative PCRs (qPCRs) (Table 3), unless otherwise specified, were designed using Primer3 input software (216) with annealing temperatures around 60ºC and with amplicons spanning or encompassing large introns if possible. Specificity of primers was verified by BLAST, corroborating no homology with other transcripts. For each qPCR reaction primer concentration was optimized to achieve a reaction efficiency >95%. GAPDH was used as housekeeping gene.

Two microliters of the diluted cDNA was used for each qPCR reaction using SYBR Green Master Mix (Applied Biosystems). The qPCRs were carried out in the 7300 Real-time PCR system (Applied Biosystems) using the following program: Stage I (50ºC for 2 minutes), Stage II (95ºC for 10 minutes), Stage III for 40 cycles (95ºC for 15 seconds and 60ºC for 1 minute), followed by a dissociation stage (95ºC for 15 seconds, 60ºC for 1 minute, 95ºC for 15 seconds, and 60ºC for 15 seconds). Data were analyzed and presented in terms of relative expression using the $2^{(-dCt)}$ method (217).

Table 3 **qPCR primers.**

Gene	Forward	Reverse
cMyc (endogenous)	ACAGAAATGTCCTGAGCAATCACCT	GCCAAGGTTGTGAGGTTGCAT
DPPA4	TGGTGTCAGGTGGTGTGTGG	CCAGGCTTGACCAGCATGAA
KLF4 (endogenous)	ACAGTCTGTTATGCACTGTGGTTTCA	CATTTGTTCTGCTTAAGGCATACTTGG
LIN28a	GGAGGCCAAGAAAGGGAATATGA	AACAATCTTGTGGCCACTTTGACA
NANOG	CCAACATCCTGAACCTCAGC	TGCGTCACACCATTGCTATT
POU5F1 (endogenous)	GGAAGGAATTGGGAACACAAAGG	AACTTCACCTTCCCTCCAACCA
REX1	TGGAGCCTGTGTGAACAGAA	CCACCTCCAGGCAGTAGTGA
SOX2 (endogenous)	TGGCGAACCATCTCTGTGGT	CCAACGGTGTCAACCTGCAT
Ecad (CDH1)	GAGTGCCAACTGGACCATTC	ACCCACCTCTAAGGCCATCT
SNAI1	CTCTAGGCCCTGGCTGCTAC	TGACATCTGAGTGGGTCTGG
TUBB3	GGCCTTTGGACATCTCTTCA	GACCGAATCCACCAGCTC
NES	CGTTGGAACAGAGGTTGGAG	GAGCGATCTGGCTCTGTAGG
GFAP	TGGAGGTTGAGAGGGACAAT	TAGGCAGCCAGGTTGTTCTC
MIXL1	GGTACCCCGACATCCACTT	GCCAAAGGTTGGAAGGATTT
BRACHURY (T)	ACTCACCTGCATGTTTATCCA	CCGTTGCTCACAGACCACAG
MESP1	GAAGGGCAGGCGATGGAG	CAGGCAGCCACTCCAGAG
SOX17	GAATCCAGACCTGCACAACG	CTCTGCCTCCTCCACGAAG
FOXA2(isoform 1)	CACTCGGCTTCCAGTATGCT	GTTCATGTTGCTCACGGAGG
O_pMX_silencing	GGCTCTCCCATGCATTCAAAC	CATGGCCTGCCCGGTTATTA
S_pMX_silencing	GCACACTGCCCCTCTCACAC	CACCAGACCAACTGGTAATGGTAGC
K_pMX_silencing	CCTCGCCTTACACATGAAGAGACA	CACCAGACCAACTGGTAATGGTAGC
M_pMX_silencing	GCTACGGAACTCTTGTGCGTGA	CACCAGACCAACTGGTAATGGTAGC

3.3.3. qPCR for retroviral silencing.

To test for silencing of retroviral transgenes used in reprogramming, qPCR was carried out as described above using primers that distinguish between endogenous transcripts and viral transcripts (Table 3). Transcripts expression was evaluated at both early and late passes of human iPS cells.

3.3.4. Immunostaining.

Human iPS cells adhered on MEF-coated slide chambers (Nunc) were fixed with 4% paraformaldehyde (Sigma) for 30 minutes at room temperature and washed thrice with PBS. At this stage, cells could be stored at 4ºC in PBS. For the detection of SOX2, POU5F1, NANOG and Tra1-81, cells were permeabilized for 10 minutes with 1% TritonX-100 (Sigma) in PBS. To detect Tra1-60, SSEA4 and negative control SSEA1 cells were permeabilized for 10 minutes with 1% Tween (Sigma). After fixing cells were washed five times with PBS for 2 minutes each and blocked with 5% bovine serum albumin (BSA) for 30 minutes at RT. Primary antibodies (Table 4), diluted in PBS/TBS with 1% BSA, were added and incubated for one hour at room temperature or overnight at 4ºC. Cells were again washed with PBS three times, for 5 minutes each. Fluorophore-labeled secondary antibodies (Table 4), diluted in the same solution as the primary antibodies, were added and incubated for 1-1.5 hours at room temperature in dark. Cells were washed with PBS (PBST) thrice, for 5 minutes each. Nuclei were counterstained with 1:4 dilution of DAPI mounting medium (Vector Labs). Samples were visualized under an inverted fluorescence microscope (Nikon Eclipse Ti-S).

Table 4 **Antibodies used in immunostaining for pluripotency markers.**

Type	Antibody	Host	Manufacturer	Catalog No.	Usage
Primary	POU5F1	Rabbit	Santa Cruz	sc-9081	1:100
Primary	Nanog	Rabbit	Abcam	ab21603	1:20
Primary	Sox2	Mouse	R&D	MAB2018	1:500
Primary	SSEA-1	Mouse	Chemicon	MAB4301	1:100
Primary	SSEA-4	Mouse	Chemicon	MAB4304	1:100
Primary	TRA1-81	Mouse	Chemicon	MAB4381	1:100
Primary	TRA1-60	Mouse	Chemicon	MAB4360	1:50
Secondary	Anti-mouse IgM FICTC-conjugated	Sheep	Sigma	C2306	1:200
Secondary	Anti-mouse IgG Cy3-conjugated	Sheep	Sigma	C2181	1:200

3.3.5. Alkaline phosphatase staining

Cells were fixed with Formaline (Merck) for 10 minutes and immediately washed with distilled deionized water. The staining was performed with a freshly prepared Alkaline Phosphatase Blue Membrane Substrate Solution (Sigma) for 30 minutes in dark according to the manufacturer's instructions. Cells were further washed and viewed for violet staining indicating a positive result.

3.3.6. Bisulfite treatment and pyrosequencing.

Genomic DNA was isolated using QIAamp DNA Mini kit (Qiagen). One microgram of DNA was treated with sodium bisulfite using the CpGenome modification kit (Chemicon). NANOG Promoter region (367 bp) was amplified by conventional PCR using High Fidelity Platinum Taq polymerase (Invitrogen) with the reverse primers biotinylated at its 5' end (Table 6). Amplicons were analyzed by pyrosequencing (Biotage/Qiagen) within sequences of 100 bp in length specified by a set of nested forward primers (Table 6). Universal methylated DNA was used as control (Intergen company).

3.3.7. Spontaneous differentiation by embryoid body formation.

Cells were grown on 1:6 Matrigel dilution until reach confluence in iPSC medium. EBs were formed by bending the cells over using a scrapper and growing them in suspension in a 6-well ultralow-attachment plate (Falcon) with the EBs medium described above. First medium change was performed at day 5 and after that, medium was changed every 2-3 days by pooling cells and media in a 15-ml conical tube, allowing the EBs to settle to the bottom of the tube, removing the supernatant, and replacing it with fresh medium. EBs were collected at day 2, 5, 8, 15, 20 and 30 of differentiation. The presence of markers representative of the three germ layers and non-coding RNAs was checked at these different time points by qPCR (Table 3). Figure 11 shows the protocol for the spontaneous differentiation of PSCs with EBs.

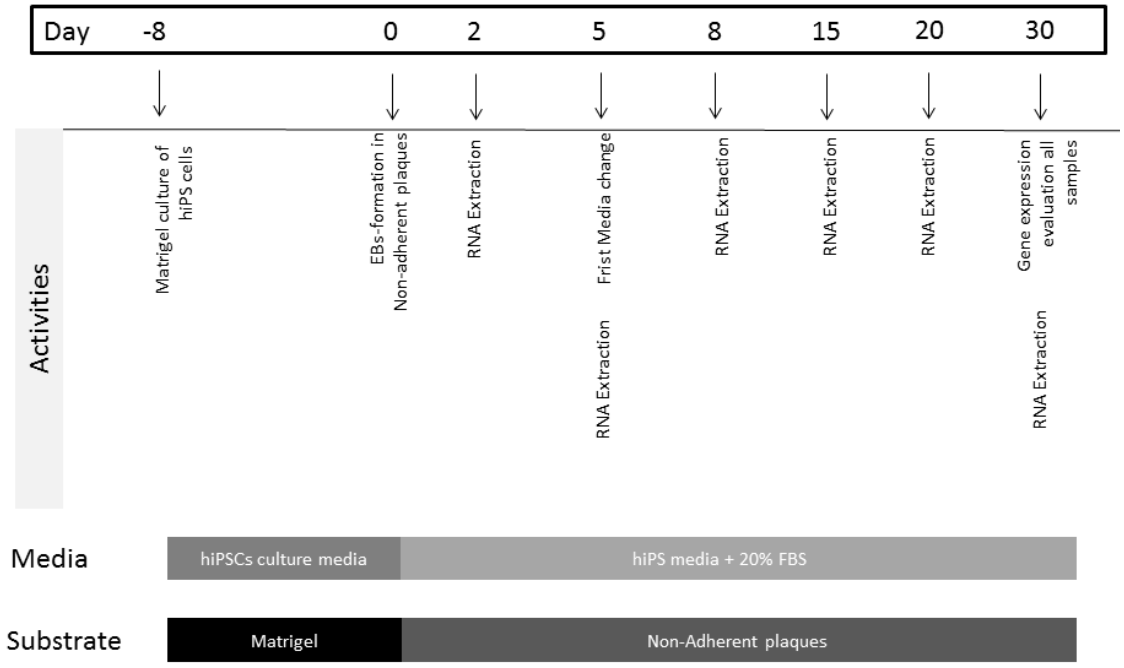


Figure 11. Protocol for the spontaneous differentiation of hiPSCs with embryoid bodies.

3.3.8. Teratoma formation assay.

Teratomas were generated by subcutaneous injection of one confluent T25 flask (Falcon) of iPSCs previously cultured on Matrigel into the dorsal flanks of 4-6 week-old male immune-deficient $Rag2^{-/-}\gamma c^{-/-}$ mice according to the ethical guidelines observed by the University of Navarra (Ethic protocol #049-14). About 4-6 weeks after injection, tumors were dissected, fixed in 10% formalin (Sigma) and stored at 4ºC until shipment to the Department of Animal Pathology, University of Zaragoza for histological analysis. Paraffin-embedded tumors were sectioned, stained with hematoxylin/eosin, and the presence of differentiated cells or tissues representative of the three embryonic germ layers was indicative of teratoma formation. MEFs were injected as negative control and fibroblast-derived MSUH001 hiPSC clone (kindly provided by Dr. Pablo Menéndez) was used as positive control.

3.3.9. Genomic stability: CGH analysis

Cellular pellet of three iPSCs samples (G15.AO, G15.D and G15.L) (passage number>40) were sent to NimGenetics (Madrid, Spain) for genomic stability evaluation. Specifically, Genomic DNA from iPSCs samples and a male control (Promega Biotech) were hybridized into array-CGH platform, composed by 6000 probes distributed across all the human genome (Agilent Technologies). ADAM-2 (window 0,5Mb, A=6) statistical was used as the parameter to measure the number of DNA copies per probe. A minimum of five consecutive probes per DNA sample is considered as an alteration in the number of DNA copies. The resolution of the analysis was 200kb for the majority of the regions. Genomic coordinates correspond to NCBI37 genomic construct (Version hg19).

3.4. NON CODING-RNA PROFILING.

3.4.1. Array hybridization and bioinformatics analysis.

High quality RNA samples (RQ1>9) from 5 ADSCs-hiPSCs (G15.D, G15.L, G15.AE, G15.AI and G15.AO) and 3 ADSCs samples (G10, G15 and G17) were hybridized into SurePrint G3 Human Gene Expression 8x60K Microarray (Agilent). This platform targeted 27.958 coding genes and 7.419 lincRNAs. Analysis was started with a filtering process to remove undetermined detectors, LIMMA (Linear Models for Microarray Data) (218) was used to determine the genes that showed significant differential expression between ADSCs-hiPSCs and control groups. Using this model, iPSCs samples were compared to parental cells to evaluate those genes specifically expressed in reprogramming, maintenance of pluripotency and terminally differentiated cells. R/Bioconductor was used for preprocessing and statistical analysis (219).

Clustering analysis including our results and public data from two different research groups with accession numbers GSE24182 (207) and GSE25970 (220) was performed. LIMMA was used to evaluate differential expression of coding and non-coding genes of twenty-two hiPS lines, twenty-two human ESC lines and sixteen somatic cells in total.

3.4.2. Array validation and LncRNA candidate selection criteria.

Differential expression of lncRNAs was defined as significant using a cut-off value of B > 2 and fold change (FC) > 2. Those lncRNAs who reached these criteria and were described in at least one of these databases: Ensamble, Genecode or Refseq were chosen for further qPCR validation experiments. A minimum of one set of validation primers was designed as described above, for the exonic region of the candidates where hybridizes the array probe (Table 5). Those candidates where the probes located within intronic genome regions were validated using primers specific for the probe. qPCR was performed as described above 3.3.2.

3.5. OVER-EXPRESSION OF LNCRNAS USING RETROVIRAL VECTORS.

3.5.1. Cloning pMXs-LncRNAs.

Retroviral pMXs-OCT4 vector, kindly provided by Dr. Boris Greber, was digested with *EcoRI* enzyme (Takara) at 37°C overnight. Digested pMX vector was then de-phosphorylated using Shrimp Alkaline Phosphatase (SAP) (USB-Affymatrix). The product of the reaction was run in a 1% agarose gel and the vector band was purified using NucleoSpin Gel and PCR clean-up (Macherey-Nagel). The resulted DNA was eluted in 30 μl of distilled water and quantified using the NanoDrop. LncRNAs candidates were amplified by PCR using cloning primers (Table 6) with *EcoRI* sites at the ends in order to facilitate cloning into pMXs vector. hiPS cDNA was used as template for the PCR reaction performed using High Fidelity Platinum Taq polymerase (Invitrogen) according to the manufacturer's instructions. PCR products were cloned into pGEMT vector (Promega) under the manufacturer´s conditions. Briefly, PCR products were ligated into pGEMT using T4 DNA ligase incubating one hour at room temperature. Subcloning efficient DH5α competent cells (Life Technologies) were transformed with 2 ul of the product of the ligations by a heat shock at 42°C for 50 seconds and incubation on ice for 2 minutes. After adding SOC medium (Life Technologies), the transformed DH5α were incubated for one and a half hours at 37°C with shaking before being plated into LB Agar plates (Pronoadisa) supplemented with ampicillin (Sigma), IPTG (Sigma) and XGal (Sigma). White colonies were selected and grown on 3 mL of LB BROTH (Pronodisa) supplemented with ampicillin at 37°C. DNA was isolated using the NucleoSpin Plasmid kit (Macherey-Nagel) according to the manufacturer's protocol. DNA was sequenced with T7 and SP6 primers to check the quality of the inserts resulted of the PCR amplification. The clones containing correct inserts were digested with *EcoRI* and purified inserts were ligated into the already digested and de-phosphorylated pMXs vector. Enzymatic digestion and sequencing (Table 6) were performed to check the clones with the inserts in the correct orientation.

3.5.2. Retroviral production of pMXs-LncRNAs.

Retroviral particles for pMXs-LncRNAs were produced as described above in section 3.2.1.

Table 5 **qPCR primers for lncRNAs validation**

LncRNA	Forward	Reverse	LncRNA	Forward	Reverse	LncRNA	Forward	Reverse
PAL1.1	TTCCAGCTACCAGGATGTGC	ATGGGGGAAGAAGTGAAATC	**PAL9.7**	GTTGAGTTGGTGCCATTCCT	AAGGGAAACAGTCACCATGC	**PAL26**	TGTCTATGCCGTCACCAAAA	GTGTTCTCAGGGCAAATGGT
PAL1.2	GGCAGAGCGAGTTTCATGTT	TGGGTGGAATTGGTTTTGTT	**PAL10**	ACAAGCTCAAGAGGCTGCAT	CGGCTTTTTAAATCCATCCA	**PAL28**	TGAAAACCCATTTCACAGCA	GTGCCTAAATCATGCCACCT
PAL2	ACCATGTGCACCTGTTGTGT	GCCAATCCAAGGTGAAAGAA	**PAL11**	GGGATGATGCAGCAAGAAGG	GCTGCCATGTGTCTCCATTC	**PAL30**	CAAAGGCCCAAGTTAGGTGA	ACAGTGCTTGGAACATGCAA
PAL3	GCCACACAAAGAAAGGAGGA	GGTCCCCATACCCTCTCCTA	**PAL12**	TTCAAGAGCTGCTCAGTCCA	GCACTCCTTTTAGGGAGCATC	**PAL31**	GCATCCATCCTGACTGGTCT	TTCAAAGCAGGGATTGCTCT
PAL4.1	TCAGCCTCCAACAGGACTCT	CCCTCTGACACCCAACCTTA	**PAL13**	ACGTTTGAAACTATTGCTTTGC	ATCCTAGCCTGGCAGTCAGA	**PAL32**	CTTCCTGCTCTTCGTTGTCC	GTCCAACTCTGCCCTGTCAT
PAL4.2	CTGACGTGAAGGGACCACAT	AGACCAAGGGGCAGGAGT	**PAL14**	CAGTGAGGAAGGGGACTCAG	CAGACTTGGCGTGAAGATGA	**PAL33**	TTCACTGTCGAAGTCTTCTTGA	TTAGTGGCTCAAAAGCAGCA
PAL5	TGCAGAAGCCACTGTGGTAG	TTTCCCTTCAAATGGCAAAC	**PAL15**	ATCCAACACCCGTTCAAGTC	TGTCTCCCATGAGAAGAGCAT	**PAL34**	CCCTGCCTGATAGTCATTGG	CGATGCAAACCGATAATGAG
PAL6	TGAAGGTGCTGTCAGAAGAGT	ATTGTTGAGTCCCGAGGTCT	**PAL16**	CAGGTGGAAAATGAAGCACA	TGCCCTTGCTTCCTCATAGT	**PAL35**	CTGATTGATCACGCAGCATT	AGGTTCCCAGGAGAAGTGGT
PAL7	TCATGGTGCCTGAGGATTCA	GTTGCGTTAGACAGCTTCCA	**PAL17**	ACCTTCCATGTACAGCAGCA	GGCTCGAATCATAGTGTCTGA	**PAL36**	ACTCGTGACCTCAGCTATGG	ACTTGGGCAGGAAGTCAGAG
PAL8	ATGGGACAGCCCTCTGAAAT	CAAAAACCCCGATGATAGGA	**PAL19**	ACATCTTCCCCATGCTTCCA	ACCACAGTTTCTCCATTCATCTG	**PAL37**	CGGCTAGCAGAACGAGGAG	CGGGTTCAGTCGTTCGATCT
PAL9.1	TTCCAAGAAGGAACATTTTATCG	GCTGAGGGAACGAAGAACAC	**PAL20**	GGTGGAAGGATTTGTTTTGC	TGAAGCACTCCACAATCTGC	**PAL38**	CTGAGCAGCATTCCCACG	GAAGACATTCCCCAGGATCCA
PAL9.2	CTCTGCTTCTACCCCACAGC	GTATAAACTTGGCCCCAGCA	**PAL21**	TGAATCAGAAGCCTGGGTCT	TTGGCACCTGGATCAAAAGT	**PAL39**	TACTGCCCAATCTCCACCAG	TGGAATTGTCTCAGCACAGC
PAL9.3	GTGACCTGCTTTTTCGCAAG	AACGTATCCATGGGGTTGAA	**PAL22**	TGGGATAGCTCTGTCTTTTCG	CTACAGTGGGAGGGGGAAGT	**PAL40**	CAAACATCGAGCAGGTCGTA	GGCAGTAGTGCTTGGGACTC
PAL9.4	CACGTCAACTCCCATGATGA	TTTCTTAAGACCACAGAAAGACAGT	**PAL23**	CATGGATGAAGCTGGAAGGC	GTGCCCGGCTTCATGTATTA	**PAL41**	AACAAGAAACTTCCCCAAGG	TGAGGTGTGGTGTGTGTCTG
PAL9.5	GAAGCAAAAGAATGAATTAAAGAAAA	GGGTGTCCAAGAGACCACAG	**PAL24**	ACACCTCTGTTTGCATCTGT	CCATCTTCCAATATCATGCAGCA	**PAL41.1**	TGAAGACAATGCTGCCTCTG	ATAAAGGTGGCCACAACTGC
PAL9.6	CTTCCTGGGAACTCTTCTGG	AGTGGGTGGTCTAGGCAATG	**PAL25**	TGACACCAAGAACCAATGAGC	GAAATGAGAAACCGCAGATAAGC	**PAL42.2**	AGGTCAGGGCATCTTTTGAC	AATTACAGGCCGGGTGTCAT

Table 6 **Cloning, sequencing and methylation primers**

Primer name	Sequence	Use	Primer name	Sequence	Use
PAL20.cloning1	aattGAATTCgctgactctcttttcagactc	Cloning	pMX-S1811-seq	GACGGCATCGCAGCTTGGATACAC	Sequencing
PAL20.cloning1_R	aattAAGCTTattaatacaagactc	Cloning	pMXs-Forward	ACCAGACTAAGAACCTAG	Sequencing
PAL20.cloning2_F	aattAAGCTTcaaaatattttattg	Cloning	pMXs-Reverse	CATTCCCCCCTTTTTCTG	Sequencing
PAL20.cloning2_R	aattGAATTCcttcctaattgaactgctac	Cloning	pLKO.1-seq	caaggctgttagagagataattgga	Sequencing
PAL21.cloning_F	aattGAATTCggaacatctcaccaatttta	Cloning	Nanog-MET-seq2	tttttgttgaattatattattgatttaaaagttgg	Methylation
PAL21.cloning_R	aattGAATTCccgtaaacactcccttcgtt	Cloning	Nanog-MET-seq1	ggttttaaatttttgattttaggtgatt	Methylation
shRNA.PAL20.21.1_F	ccggAACCAGATGCTAAAAATTCAActcgagTTGAATTTTTAGCATCTGGTTttttttg	Cloning	MET-Nanog-F	gattttgtaggtgggattaattgtgaattt	Methylation
shRNA.PAL20.21.1_R	aattcaaaaaAACCAGATGCTAAAAATTCAActcgagTTGAATTTTTAGCATCTGGTT	Cloning	MET-Nanog-F	gattttgtaggtgggattaattgtgaattt	Methylation
shRNA.PAL20.21.2_F	ccggAAGAACAtCtCACCAAttttActcgagTaaaaTTGGTGaGaTGTTCTTttttttg	Cloning	MET-Nanog-Rbiotin	bio-accaaaaaaacccacactcatatcaatata	Pyrosequencing
shRNA.PAL20.21.2_R	aattcaaaaaAAGAACAtCtCACCAAttttActcgagTaaaaTTGGTGaGaTGTTCTT	Cloning	Nanog-seq1B	ttattgggattataggggtgggtta	Pyrosequencing
shRNA.PAL20.1_F	ccggAAGGACAAttCAtGtAAtCtActcgagTaGaTTaCaTGaaTTGTCCTTttttttg	Cloning	Pyro-Nanog-R	bio-taattgagaatttataagggtg	Pyrosequencing
shRNA.PAL20.1_R	aattcaaaaaAAGGACAAttCAtGtAAtCtActcgagTaGaTTaCaTGaaTTGTCCTT	Cloning	Pyro-Nanog-F	tggttaggttggttttaaattttt	Pyrosequencing
shRNA.PAL21.1_F	ccggAACAACtAGCttGGAAAtttActcgagTaaaTTTCCaaGCTaGTTGTTttttttg	Cloning			
shRNA.PAL21.1_R	aattcaaaaaAACAACtAGCttGGAAAtttActcgagTaaaTTTCCaaGCTaGTTGTT	Cloning			

3.6. LOSS OF FUNCTION EXPERIMENTS.

3.6.1. Small interfering RNA (Si-RNA) design.

RNA sequences of lncRNAs candidates were introduced on i-Score siRNA design web tool (221). Those si-RNAs with a Reynolds value > 6 and DSIR value between 85 and 90 were considered as accurate inhibitors of the lncRNA candidates. The results were confirmed by introducing the same RNA sequence into si-RNA Scales web site (222) and looking for the same siRNAs obtained with i-Score and a Scales value < 2. At least three siRNAs sequences for each lincRNA candidate were selected.

3.6.2. Short-hairpin RNA (shRNA) cloning in lentiviral vector.

pLKO.1 puro plasmid (Addgene, ref# 8453) was used as vector for shRNA cloning following manufacturer´s instructions. Designed siRNAs described above were used as sense and antisense sequences for shRNA oligos (Table 6). Briefly, 5 μl of 20 μM shRNA forward and reverse oligos were mixed and incubated at 95°C in a water bath and then slowly cooled down at room temperature over a period of several hours letting the oligos to anneal. pLKO.1 cloning vector was digested with EcoRI (Takara) and AgeI (New England Biolabs) enzymes and digested DNA was run in a 1% agarose gel and purified using NucleoSpin Gel and PCR clean-up (Macherey-Nagel). Annealed oligos were ligated into the digested pLKO.1 cloning vector using T4 DNA ligase, transformed into Subcloning efficient DH5α competent cells (Life Technologies) and plated into LB Agar plates (Pronoadisa) with ampicillin (Sigma). Colonies were selected, grown on 3 mL of LB BROTH (Pronodisa) with ampicillin and DNA was isolated using NucleoSpin Plasmid kit (Macherey-Nagel). Proper ligation and the correct sequence of the oligos were confirmed by sequencing (Table 6). Final plasmids were transformed into subcloning efficiency DH5α competent cells, grown in 450mL of LB BROTH with ampicillin, and plasmid DNA was isolated using NucleoBond Xtra Maxi EF (Macherey-Nagel).

3.6.3. Lentiviral production and establishment sh-pluripotent cell lines.

Lentiviral particles for pLKO.1-shRNAs were produced as described above in section 3.2.2. Teratocarcinoma cell line NCCIT was infected with pLKO.1-puro lentiviral particles expressing shRNAs specific against PAL20 and PAL21 lncRNAs (shRNA20, shRNA21, shRNA20.21.1 and shRNA20.21.2.) pLKO.1-shRNAscramble (negative control containing a scrambled shRNA sequence) was used as control. 48h post-infection cells were treated with puromycin (5mg/ml) in order to establish cell lines that constitutively have a silenced expression of lncRNAs candidates.

3.6.4. Proliferation assay.

Proliferation analysis was carried out with proliferation kit CellTiter 96® AQueous Non-Radiactive Cell Proliferation Assay (Promega, Madison, WI) which is based on a colorimetric method for determining the number of viable cells. It consists of a tetrazolium solution (MTS) and a coupling reagent. Dehydrogenase enzymes present in metabolically active cells are able to reduced MTS to formazan producing a color change, from yellow to purple. The amount of formazan crystals measured directly by the absorbance at 490 nm in a spectrophotometer plate, is directly proportional to the amount of viable cells. For cell proliferation analysis 1 x 10^5 cells per well were

plated in a final volume of 100 ul of culture medium in 96-weell plate. To prevent evaporation of the medium, a humid chamber was built by adding 100 ul of PBS to all wells of the periphery of the plate. The proliferation of established NCCIT-shRNA20.1 and NCCIT shRNA21.1 cell lines was evaluated two and four days after being plated in 96-well plates. Plates were centrifuged at 800 g for 10 minutes, the supernatant was removed and the cell pellet was resuspended in 100 ul of culture medium containing 20 ul of Cell Titer 96®AQueuos MTS reagent (Promega, Madison, WI). The plate was incubated at 37 °C for 2-5 hours and the absorbance at 490 nm was measured in the plate reader Magellan Plate Reader.

3.7. STATISTICS.

For the microarray profiling analysis, differential expression of coding genes and lncRNAs was defined as significant using a cut-off of B>0 or B>2 respectively when comparing ADCS-iPSCs vs. ADSCs. All other numerical data were analyzed using the Prism GraphPad software version 6 (www.graphpad.com) expressed as mean fold-change or expression ± SEM. Statistical analysis was performed using two-way analysis of variance (ANOVA) and multiple comparison tests. Statistical significance is indicated as follows: * $p<0.05$; ** $p<0.01$; *** $p<0.001$**** $p<0.0001$ ns: no-significative.

4. RESULTS

4.1. GENERATION AND CHARACTERIZATION OF HUMAN IPSC.

Nowadays the generation of human induced pluripotent stem cells (hiPSCs) by the overexpression of core pluripotency TFs has been performed within a great number of parental adult healthy tissues and diseased ones. Additionally, both integrative and non-integrative methods have been used on the quest for a balance between efficiency and safety (127). In this work, for hiPSCs generation we chose a retroviral vector system for the transduction of the four core TFs (4F): POU5F1, SOX2, KLF4 and c-MYC. This system allows efficient hiPSC generation from a different cell types and it is suitable for the purposes of this work. Once generated, hiPSCs were characterized according to standard techniques.

4.1.1. Generation of hiPSC from fibroblasts and ADSCs.

To generate iPSCs, terminally differentiated hFibs and ADSCs from five different donors, either male or female, and at different culture passages (Table 7) were infected with retroviral particles containing the 4F TFs following the protocol described in section 0 (Figure 10). Briefly, retroviral particles coding for the human reprogramming factors POU5F1, SOX2, KLF4 and cMYC were generated and used to infect either hFibs or ADSCs. After 48 hours, infected cells were cultured in iPSCs medium in MEFs feeder layer until the emergence of the ES-like colonies. Colony formation, their time of appearance as well as the cell morphology were evaluated.

Three weeks after hFibs and ADSCs infection the first colonies with the typical morphology of pluripotent stem cells were observed. iPSCs colonies were flat, tight, bright, rounded with well-defined borders and a high rate of proliferation (Figure 12). No differences in morphology or colony formation related to the age and/or the sex of the donor were found. Even though ADSCs, compared to fibroblasts, are considered to be in a more undifferentiated state no difference in the time of colony appearance was observed between those cell types. We observed that the cells that failed to form colonies were those with a culture passage greater than five, fact that has been previously described (115), and those that did not have a normal proliferation rate prior infection (sample SVF10 (Table 7)).

In order to expand the generated hiPSCs colonies, four to five weeks after infection a minimum of 24 colonies were picked from each reprogramming experiment, and individual colonies were cultured on 12 well plates coated with irradiated MEFs (Figure 12d and e). Each clone was identified using the sample name followed by a number or a letter (i.e. M45a9 or G15D). In general, selected clones reached confluence in the 12 well plates after 8 days of culture and were passaged by carefully picking up the ES-like colonies with a pipette tip to a MEF-coated 6 well plate (one clone per well) and expanded in hiPSCs media until having a confluent 6 well plate per clone. Once a confluent 6-well plate was obtained for each clone, three wells were used for further expansion and characterization, and the rest of the cells were frozen and stored in liquid nitrogen (one vial per well). One non-confluent P6 well in an early passage of each clone was selected for alkaline phosphatase staining as described in section 3.3.5. This test was used as an indicative of pluripotency program on-setting of the established clones, since all of them were positive further characterization analysis needed to be done, in order to finally conclude that our clones were pluripotent. (Figure 12 g and h). Those clones that did not grow properly and/or lost the typical

iPSCs morphology were discarded. We considered that hiPSCs clones were established when they can be routinely passaged once a week in a dilution not less than 1:2 and are able to maintain the un-differentiated state for more than 5 passages.

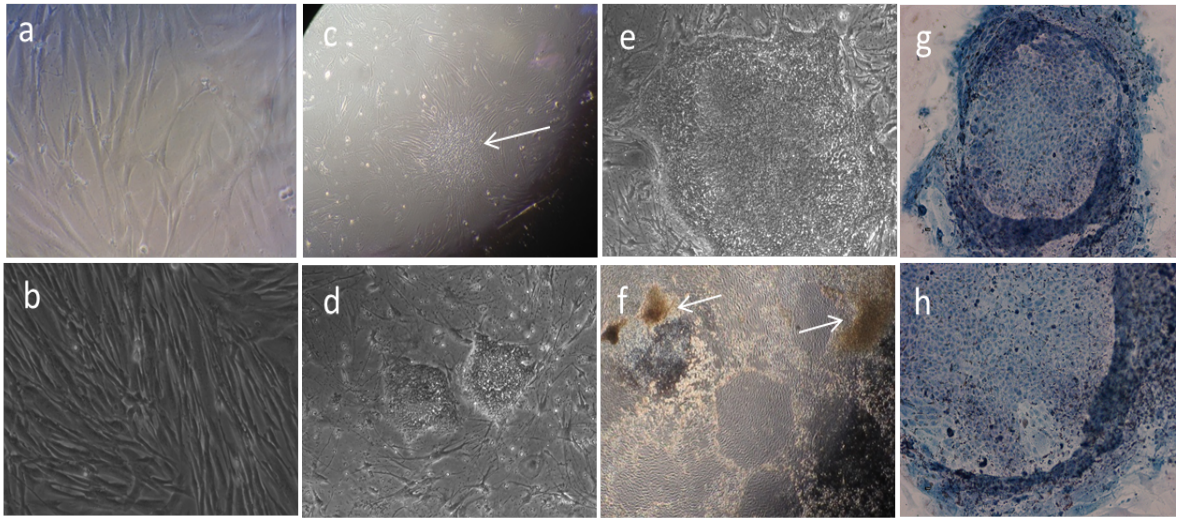


Figure 12. Timeline evolution of reprogramming of human somatic cell towards hiSPCs | a) Human ADSCs in culture 10X b) Confluent culture of human fibroblast 10X. c) Human somatic cells two weeks after being infected with 4F. White arrow shows a group of cells that suffered typical changes in morphology due to reprogramming process and likely to form an iPSC colony 10X. d) Two hiPSC colonies, each of them will give rise to two independent hiPSC clones 20X. e) Typical human ES-like colonies with flat morphology growing on MEFs of M45a9 clone and G15D clone respectively 20X and f) Over-confluent culture of hiPSCs, white arrows show spontaneously differentiated areas within culture due to confluence 10X, g and h) show AP staining of hFib-hiPSC clone 20X and 40X respectively. All these photos are representative examples of all the clones generated in different reprogramming experiments performed for this study.

In summary, we were able to generate 92 hiPSCs clones; 25 clones were derived from 5 human fibroblasts samples and 57 clones from 5 different ADSCs samples. Some samples failed to generate clones suggesting a relationship between the passages of somatic cells previous to reprogramming and their capacity to generate clones.

4.1.2. Characterization of hiPS cells.

Fully reprogrammed hiPSCs are defined by the expression of pluripotency markers at gene and protein level, silencing of the transgenes used in the reprogramming process, loss of somatic cell epigenetic marks and *in vitro* and *in vivo* differentiation capacity. Standard procedures were followed to assess the pluripotent stem cells characteristics in the established hiPSC clones from human fibroblasts and ADSCs.

4.1.2.1. Expression of endogenous pluripotency markers.

The expressions of markers that are indicative of a pluripotent state were analyzed at gene and protein level, in different established hiPSCs clones at early passage (not greater than 10). We used quantitative PCR (qPCR) to evaluate the gene expression of pluripotency markers not included in the reprogramming cocktail (i.e. NANOG and LIN28) as well as the activation of the endogenous POU5F1, SOX2, KLF4 and c-MYC genes. Fibroblast-derived hiPS clone MSUH001,

kindly provided by Dr. Pablo Menendez (Andalusian Stem Cell Bank), was used as positive control in all the experiments.

Table 7 **Human cells samples used for hiPS generation**

Sample Name	Sex	Age	Source	Passage	Infection date	Colony formation	Number of clones	Characterized
M38a	F	38	Fibroblast	0	14/10/2010	Yes	5	No
H59a	M	59	Fibroblast	0	14/10/2010	Yes	6	No
M76a	F	76	Fibroblast	1	27/01/2011	Yes	5	No
M76a	F	76	Fibroblast	5	17/03/2011	No	NA	NA
H3a	M	3	Fibroblast	6	14/04/2011	Yes	3	No
M45a	F	45	Fibroblast	5	03/09/2011	Yes	6	Yes
G6	ND	ND	ADSCs	7	27/01/2011	No	NA	NA
G6	ND	ND	ADSCs	9	10/02/2011	No	NA	NA
G8	ND	ND	ADCs	4	22/10/2010	Yes	9	No
SVF2	ND	ND	ADCs	3	14/10/2010	Yes	1	No
SVF10	ND	ND	ADSCs	1	24/02/2011	No	NA	NA
SVF10	ND	ND	ADSCs	3	17/03/2011	No	NA	NA
G10	F	39	ADSCs	0	14/04/2011	Yes	5	No
G15	M	ND	ADSCs	2	03/09/2011	Yes	23	Yes
G15	M	ND	ADSCs	4	13/09/2011	Yes	15	Yes
G16	M	ND	ADSCs	2	03/09/2011	Yes	3	Yes
G16	M	ND	ADSCs	4	13/09/2011	Yes	1	No

M: Male, F: Female, ADSCs: Adipose Derived Stem Cells, SVF: Stromal vascular fraction, G: isolated ADSCs, ND: Not determined, NA: Not applicable.

First, NANOG induction was measured in 20 hiPS clones derived from both hFibs and ADSCs and high levels of expression were observed in all the clones compared to parental cell lines, where gene expression was undetectable (Figure 13a). However variable levels of NANOG expression was observed, a reported common feature among PSCs (10). Thus, ten clones with different NANOG expression patterns, were randomly chosen for further characterization procedures, five Fib-hiPSCs (M45a.9, M45a.22, M45a.28, M45a.29 and M45a.30) and five ADSCs-hiPSCs (G15.D, G15.L, G15.AO, G15.AE, G15.AI). The endogenous expression of the 4F and LIN28 was evaluated and we could observe a clear induction in all the evaluated clones, nevertheless the gene expression of NANOG, LIN28 and endogenous 4F was very heterogeneous among different clones and in most of the cases higher than the levels observed in the MSUH001 hiPSC used as control (Figure 13b).

The presence of the pluripotency factors POU5F1 and SOX2, included in the reprogramming cocktail, as well as NANOG, a well-known pluripotency marker (10), were also observed at protein level in all fibroblast and ADSCs-derived hiPSC clones by immunofluorescence. Other pluripotency associated membrane markers, such are TRA1-81, TRA 1-60 and SSEA-4, were also detected

(Figure 13c). SSEA-1, used as negative control, since it has been described to be marker specific for mouse pluripotent cells, was not detected in any clone (Data not shown).

These results clearly indicate that all the clones generated up-regulated the expression of markers indicative of a pluripotent state in comparison to donor cells.

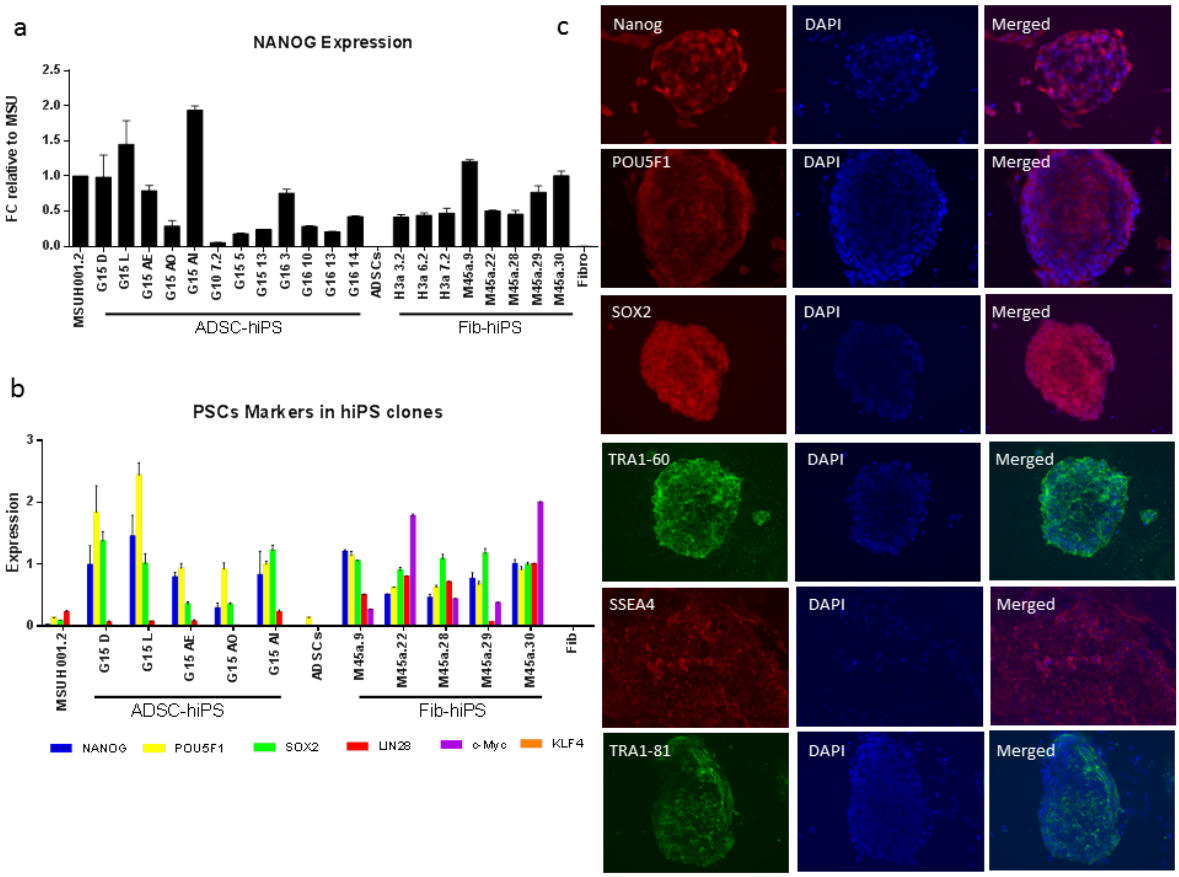


Figure 13. Characterization of hiPSCs clones | a) NANOG expression in 12 clones of ADSC-hiPSCs and 8 clones Fib-hiPSCs relative to control MSUHOO1.2. Levels of NANOG vary highly from clone to clone, cells were analyzed at passage 5. b) 5 clones of ADCS-hiPSCs and Fib-hiPSCs show robust induction of endogenous pluripotency markers NANOG, POU5F1, SOX2, LIN28, c-MYC and KLF4 compared to donor cells Fibroblasts and ADSCs. c) Immunofluorescence staining for pluripotency markers NANOG, POU5F1, SOX2, TRA1-60, TRA1-81 (20X), SSEA4 (40X), pictures correspond to a Fibroblast derived hiPS clone M45a9.

4.1.2.2. Silencing of the retroviral transgenes.

The following step performed during the characterization process was to verify the silencing of retroviral transgenes. This is an important feature in order to obtain good quality hiPSC, since continuous expression of exogenous reprogramming factors could affect the differentiation capacity of the hiPSCs. cDNAs samples, at passage 5 and passage 30 of culture, from selected Fib-hiPSC and ADSCs-hiPSC clones were amplify using qPCR with primers specific for exogenous POU5F1, SOX2, KLF4 and c-MYC (Table 6). We compared relative expressions of exogenous transgenes in the same clone at the different passages.

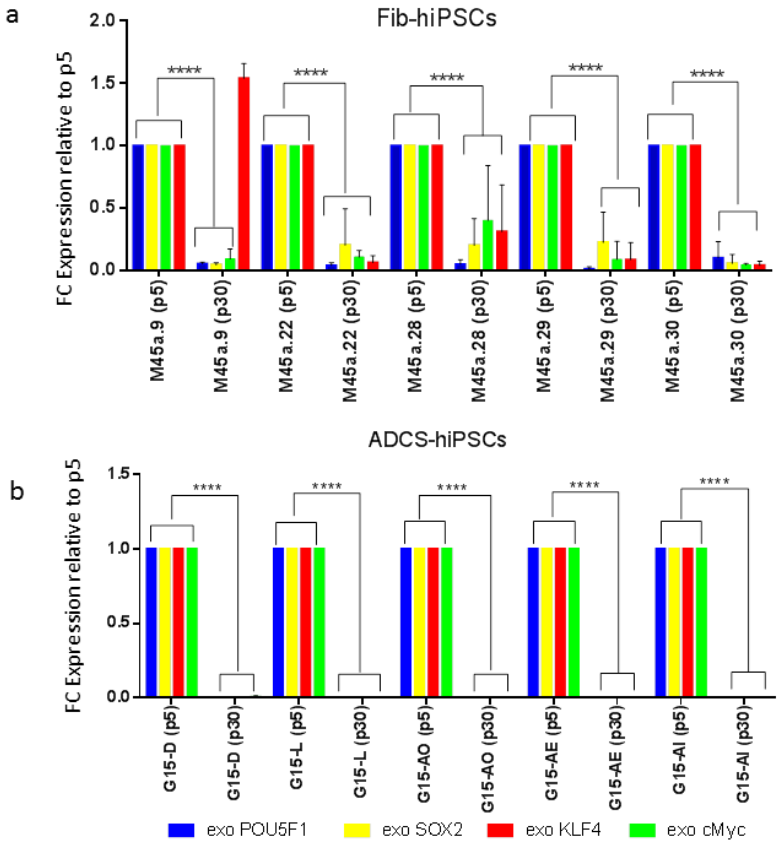


Figure 14. Silencing of exogenous transgenes expression in clones of hiPSCs at passage 30 relative to passage 5 | a) Most of the Fib-hiPSCs clones significantly silenced the exogenous transgenes. Clone M49a9 showed an increment on expression of KLF4. b) All ASDCs derived hiPSC clones silenced the exogenous transgenes.

In general iPSC clones at early passage from both Fib-hiPSC and ADSCs-hiPSC showed high levels of exogenous 4F that were silenced at later passages. The pattern of silencing varied between the different fib-hiPSC, specifically in clone M45a9 where high levels of KLF4 transgene expression were still detected at passage 30. The silencing of the exogenous 4F in ADSCs-hiPSCs was clearly more homogeneous, compared to Fib-hiPSC. ADSCs-iPSCs showed a more effective silencing of transgenes being almost undetectable at passage 30 (Figure 14). These results, together with the up-regulated expression of the endogenous pluripotency markers, clearly indicate the activation of the hiPSCs own pluripotency program in all the clones evaluated.

4.1.2.3. In vitro spontaneous differentiation by embryoid bodies formation.

By definition pluripotent cells must give rise to cells from the three germ layers when submitted to differentiation process. Thus the differentiation capacity of generated hiPSC clones was analyzed using an *in vitro* model of early lineage differentiation (223) by EBs formation. All clones evaluated from both origins, Fib-hiPSCs and ADSC-hiPSCs, successfully formed EBs and no differences were observed when compared to the control EBs generated from MSUH001 (Figure 15a).

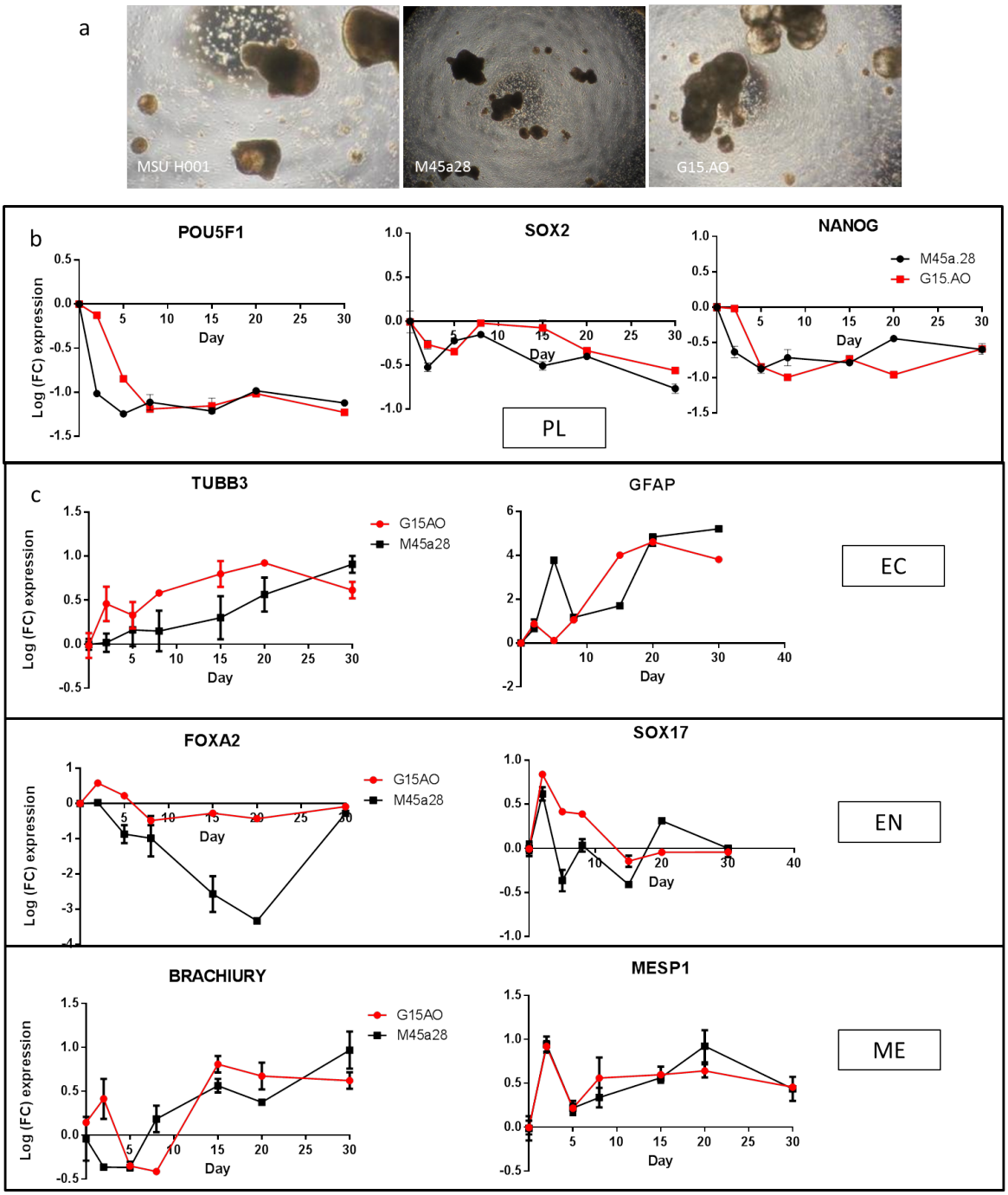


Figure 15. ***In vitro*** **spontaneous differentiation with embryoid bodies (EBs) after 8 days in culture | a) EBs generated by control hiPSCs MSUH001 , generated by human fibroblasts derived hiPSCs (M45a28) and by human ADSCs derived hiPSCs (G15.AO), 10X b) Loss of pluripotent core (POU5F1, SOX2 and NANOG) due to differentiation, was maintained through time. c) Early induction of differentiation markers representative from three germ layers was maintain trough time except for endodermal markers. Data from clones: M45a.28 derived from fibroblasts and G15.AO derived from ADSCs, are representative examples from all the clones evaluated. EN: Endoderm, ME: Mesoderm, EC: Ectoderm, PL: Pluripotency.**

Differentiation process in PSCs should lead to disappearance of pluripotent core factors and the up-regulation of tissue-specific genes. To further characterize the spontaneous differentiation process in our generated hiPSCs, RNA was extracted at different time points for 30 days after EBs formation and qPCR was performed to evaluate the expression of POU5F1, SOX2 and NANOG genes. Figure 15b, shows the loss of expression of pluripotency markers as soon as day 2, feature that was maintained through 30 days. This reduction of the pluripotency-associated genes was observed in all the clones analyzed from both fib-hiPSCs and ADSCs-hiPSCs. In general POU5F1 and NANOG showed the most dramatically down-regulation upon differentiation, whereas SOX2 reach the lowest values of expression nearly to the end of the differentiation process (Figure 15b). This could be due to the presence of a strong barrier applied by SOX2 loci against differentiation or the onset of ectodermal differentiation, evidenced by the expression of TUBB3 and GFAP markers (Figure 15c).

The induction of markers representative of the three germ layers, like SOX17 and FOXA2 (endoderm), TUBB3 and GFAP (ectoderm), and Brachyury and MESP1 (mesoderm) was also confirmed by qPCR. The expression of those differentiation markers was significantly increased at different time points during the spontaneous differentiation The expression of the different markers was very variable among the different iPSCs clones. As an example, during the first days of spontaneous differentiation, Brachyury and FOXA2 expression was reduced in Fib-hiPSCs whereas their expression was induced in ADSCs-hiPSCs. Moreover, at day 30 of differentiation, the induction of endodermal markers was completely lost (Figure 15c). These differences are probably due to the nature of the process of the EB formation; where usually cells differentiate randomly to any cell what could cause the onset of a specific germ layer and the silencing of others.

The induction of markers from the three germ layers together with the reduction of the pluripotency-associated markers implies that iPSCs are susceptible to differentiate spontaneously (Figure 15c).

4.1.2.4. In vivo teratoma generation.

To test the pluripotency of iPSCs *in vivo*, Fib-hiPSC and ADSC-hiPSC clones were subcutaneously injected into the dorsal flanks of immunodeficient $Rag2^{-/-}\gamma c^{-/-}$ mice. Four to six weeks after injection, tumor nodules appeared at the sites of the iPSCs injection whereas no tumor growth was observed when control fibroblast was used. Those tumor nodules were removed, fixed and embedded in paraffin for histological analysis. Hematoxylin/eosin method (HE) and Masson´s staining were used as common procedures for histopathological evaluation. Tissues of all three germ layers were observed: keratinocytes, nervous tissue, blood cells, gut and glandular epithelia, cartilage, bone, skeletal and smooth muscle, among others, indicating that the both Fib-hiPSC and ADSC-hiPSCs clones were pluripotent (Figure 16 and Table 8). No significative differences were observed between de different clones analyzed.

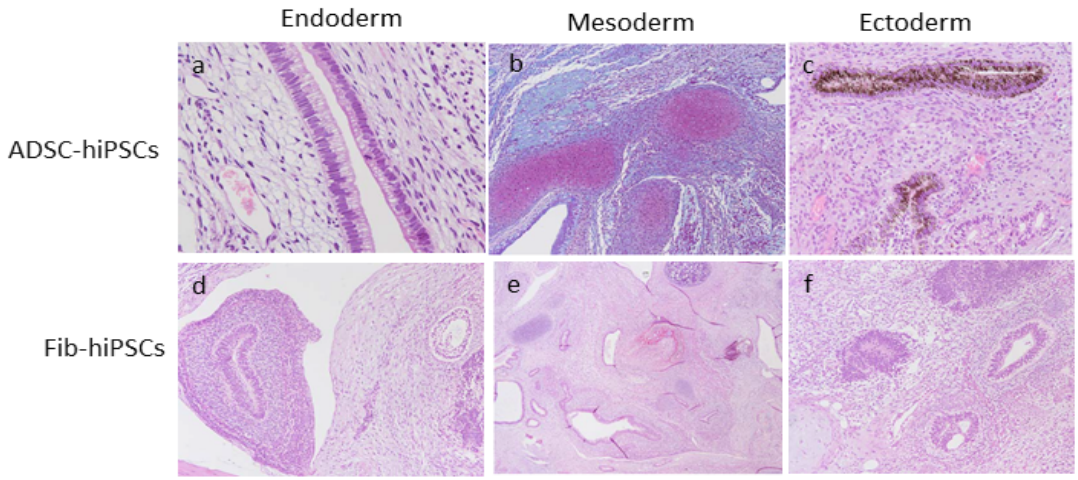


Figure 16. hiPSCs derived from fibroblasts and ADSCs have *in vivo* differentiation potential │ Teratoma formation for the evaluation of the differentiation potential *in vivo*, towards the three germ layers: a), b) and c) correspond to representative ADSC-hiPSCs teratomas (clon G15.D) : a) Pseudo-stratified epithelia corresponding to the trachea 400X (EN). b) cartilage tissue embedded in smooth muscle fibers 100X (ME). c) presence of melanin in epithelial tissue 200X (EC). d), e) and f) correspond to representative Fib-hiPSCs teratomas (clon M45a28): d) digestive epithelium 100X (EN). e) cartilage and bone tissues 40X (ME). f) neural tissue 100X (EC)

Table 8 **Presence of tissues from the three germ layers in teratomas**

Tissue	M45a.28	M45a.29	G15.D	G15.AO
Glandular epithelium	+	+	++	+++
Digestive and respiratory epithelium	+++++	++++	++++	++++
Neuronal rosettes	++	-	+++	+++++
Conjunctive tissue	+++	++	++	+++
Myxoid tissue	++	++	++	++
Cartilage	+++++	++++	++++	+++
Bone	++	+	+	+
Skeletal muscle	-	-	-	-
Smooth muscle	+++	++	++	+++
Cysts	+++++	+++++	+++++	++++
Melanin	+	-	+	+++

Valuation code: -: not observed, +: rarely observed or intense, ++: infrequent or non-intense, +++: moderately observed or intense, ++++: quite common or intense, +++++: very frequent or intense.

4.1.2.5. NANOG promoter hypomethylation.

One of the conditions that the "Comisión de Seguimiento y Control de la Donación y Utilización de Células y Tejidos Humanos" impose in order to authorize the generation and use of human iPSCs is the deposit of the generated hiPSCs clones to the Spanish Stem Cell Bank, making them available for research by other groups. Since other groups have been already deposited several Fib-hiPSCs, we chose ADSCs-derived hiPS as a novel contribution to this field. Moreover, our final objective is the identification and characterization of novel lncRNAs related to pluripotency and reprogramming and the non-coding RNA profile of Fib-hiPSCs has been reported before (207,220).

Thus, for those reasons we used only ADSCs-hiPSCs clones to perform the following characterization procedures: the methylation status of promoters and their genomic stability.

In the case of methylation status, PSCs possess specific epigenetic marks like demethylation of promoters associated with pluripotent genes like POU5F1 or NANOG. Herein we evaluated the methylation status of the NANOG promoter in three ADSC-hiPS clones by bisulfite treatment followed by individual sequencing and pyrosequencing, a technique that allows the quantification of the methylation level at each CpG. Briefly, genomic DNA was extracted from the three ADCS-hiPSCs clones and treated with bisulfite previous to amplification with conventional PCR using biotinylated primers (Table 6). Amplified genomic DNA regions were subsequently analyzed by manual sequencing and pyrosequencing.

By both methods a hypomethylation of NANOG promoter was observed in the three ADSCs-hiPSCs analyzed. Pyrosequencing allowed us to quantify the methylation levels at each CpG being lower than ≤40% in hiPSCs compared to the levels observed in the parental cell lines (≥40% and ≤80%) and in the methylated control (≥80%) (Figure 17). One of the distal CpG seemed to be hypomethylated in both ADSC and ADSC-hiPS this need additional evaluation. Demethylation of the NANOG promoter is a specific characteristic of pluripotent cells, which confirmed once more that the clones evaluated had acquired pluripotent state.

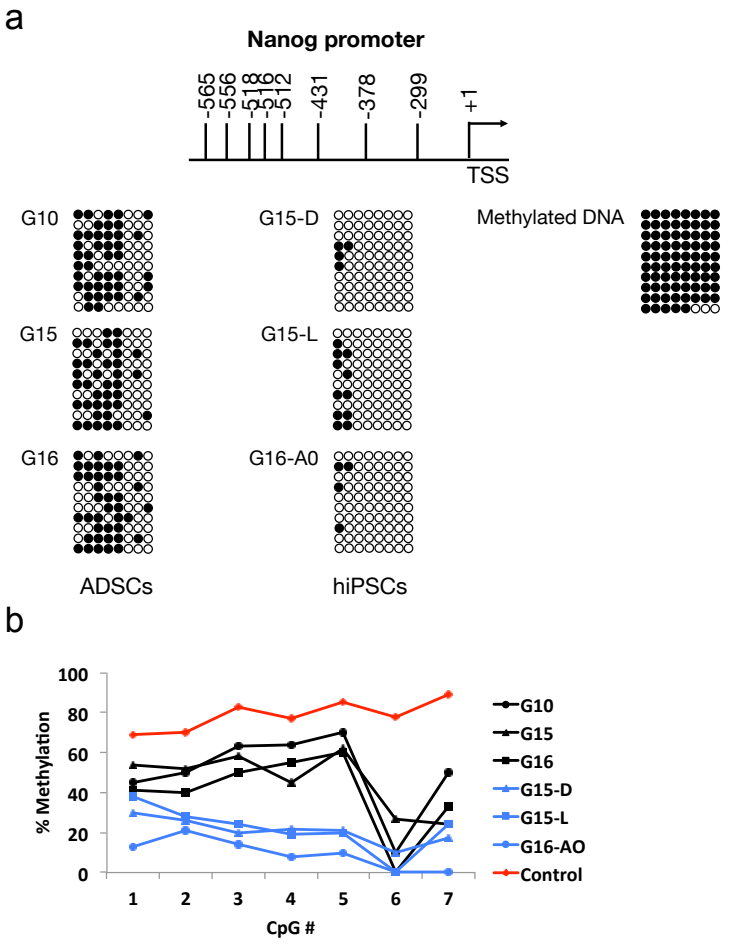


Figure 17. NANOG promoter methylation analysis | a) Manual sequencing and b) pyrosequencing. ASDCs-hiPSCs clones G15-D, G15-L and G16-AO are hypomethylated at the NANOG promoter compared to donor ADSCs cells, G10, G15 and G16. CpGs #6 are hypotemethylated in both parental and hiPS cells. Top panel shows the promoter region of NANOG with base pair numbers in relation to the transcription start site and the most proximal analyzed CpG sites. Universal methylated DNA was used as control.

4.1.2.6. Genomic stability: CGH analysis.

Since the reprogramming process and the maintenance of hiPSC in culture is associated to genomic instability leading to chromosomal aberrations, we test the genetic stability of three of our ADSCs-hiPSC lines that would be included in further experiments of lncRNA profiling. Briefly, genomic DNA from G15.AO, G15.D, G15.L clones at passage greater than 40 were hybridized into a comparative genomic hybridization microarray (CGH) that allows genome-wide DNA copy number variation profiling, focusing on known genes, promoters, miRNAs, pseudoautosomal and telomeric regions.

Clone G15.AO showed a deletion of 520kb corresponding to two genes FRMD6 and GNG2 and G15.D clone presented as well, a deletion of 1,78Mb consequent with the loss of 13 genes (CLMP, SCN3B, ZNF202, VWA5A, PANX3, TBRG1, SIAE, SPA17, NRGN, ESAM, ROBO3, ROBO4) (Figure 18 and Table 9). Both deletions were reported of unknown relevance and the removed genes are not directly associated to proliferation and/or survival, important features of PSCs. The biggest abnormality was found on clone G15.L which suffered a complete loss of chromosome *Y* (Y chromosome nullisomy), corresponding to 28,34Mb and 50 genes (Figure 18 and Table 9). This loss in males is not clinical relevant and is a common feature of highly proliferative cells, like hiPSCs. All these results indicate that ADSC-hiPSCs can acquire genetic abnormalities probably due to culture conditions and reprogramming process as has been reported before (224).

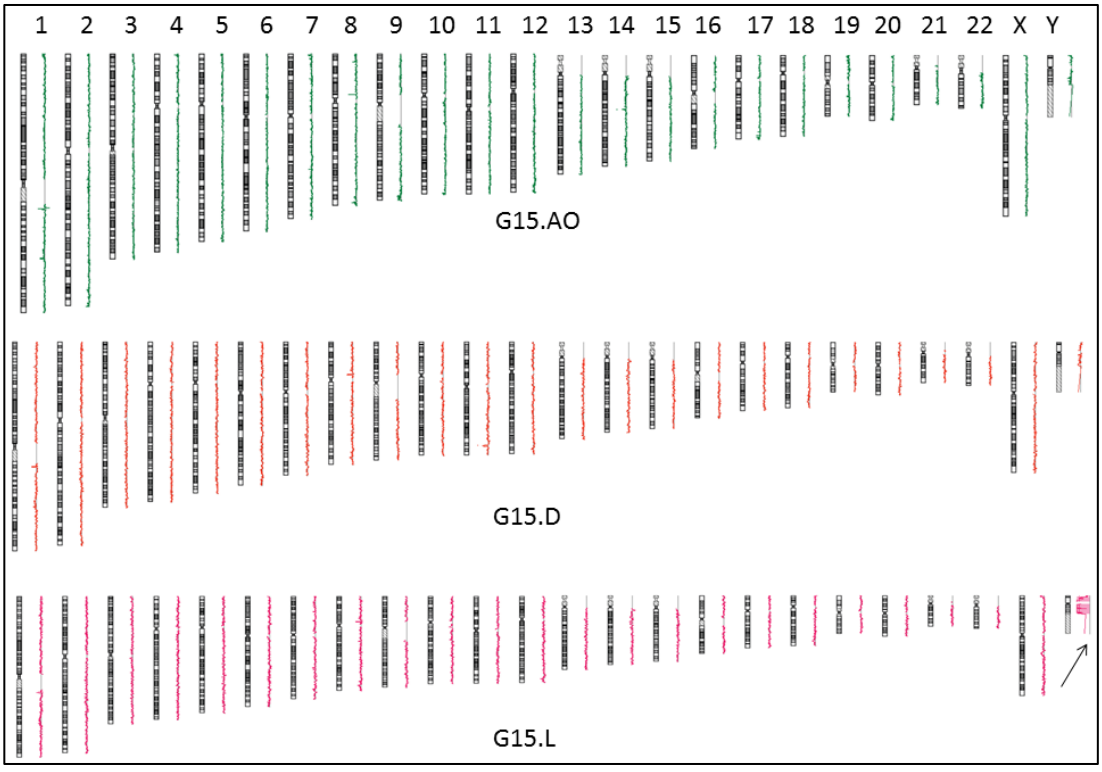


Figure 18. Representative image of the genome of the samples evaluated G15.AO , G15.D and G15.L. Arrow shows chromosome Y nullisomy. Genomic coordinates correspond to NCBI37 genomic construct (Version hg19).

Table 9 **Copy number variation detailed for G15.AO, G15.D and G15.L clones**

Chromosome	Cytoband	Size	Type	Clone	Genes
14	q22.1	520 Kb	Deletion	G15.AO	FRMD6 , GNG2
11	q24.1 q24.2	1,78 Mb	Deletion	G15.D	CLMP, SCN3B, ZNF202, VWA5A, PANX3, TBRG1, SIAE, SPA17, NRGN, ESAM, ROBO3, ROBO4
Y	-	28,34Mb	Total loss	G15.L	50 genes

4.1.2.7. Summary of hiPSC characterization.

In summary, we were able to generate a great number of hiPSCs clones derived from both human fibroblasts and ADSCs samples. All hiPSCs clones evaluated showed the typical human ES-like morphology characterized by flat, tight, bright, rounded colonies with well-defined borders and high rate of proliferation. Some human samples failed to generate clones suggesting a relationship between the passages of somatic cells previous to reprogramming and their capacity to generate hiPSC clones. In general all hiPSCs clones were positive for AP staining, expressed high levels of endogenous pluripotency markers at gene and protein levels, showed time culture-depending silencing of retroviral transgenes, demethylation of NANOG promoter and differentiation capacity both *in vitro* and *in vivo* towards the three germ layers, even though some genetic instability was found (Table 10). These results demonstrate that we were able to reprogram human fibroblasts and ADSCs to an undifferentiated state generating fully characterized human induced pluripotent stem cells clones.

Table 10 **Summary of hiPSC characterization**

iPSC	Clone	Pluripotency gene expression	Immuno Fluorescence	Exogenous gene silencing	EBs formation	Reduction of pluripotency genes (EBs)	Induction of genes from 3 germ layers (EBs)	Teratoma formation	DNA methylation	CGH
Fib-hiPSC	**M45a.9**	**Yes**	**Yes**	**Yes**	**Yes**	**No**	**No**	**Yes**	**No**	**No**
	M45a.22	**Yes**	**Yes**	**Yes**	**Yes**	**No**	**No**	**Yes**	**No**	**No**
	M45a.28	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**No**	**No**
	M45a.29	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**No**	**No**
	M45a.30	**Yes**	**Yes**	**Yes**	**Yes**	**No**	**No**	**No**	**No**	**No**
ADSCs-hiPSC	**G15.D**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
	G15.L	**Yes**	**Yes**	**Yes**	**Yes**	**No**	**No**	**Yes**	**Yes**	**Yes**
	G15.AO	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
	G15.AI	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**No**	**No**
	G15.AE	**Yes**	**Yes**	**Yes**	**Yes**	**No**	**No**	**Yes**	**No**	**No**

Yes: performed. No: not-performed

4.2. LncRNA PROFILING OF INDUCED PLURIPOTENT STEM CELLS.

Both pluripotent and differentiated cells have a specific profile of coding and non-coding genes that define their phenotype. LncRNAs are transcripts of 200 nt or more that have been related to a great number of cellular processes including development, reprogramming and disease (179). In order to search for novel lncRNAs with a possible role in pluripotency maintenance and/or somatic cells reprogramming, the differential expression profile of coding and non-coding genes was evaluated. Since lncRNA profile of human fibroblast iPSCs has been reported before (207,220), we chose to evaluate the lncRNA profile of five ADSCs-derived hiPS clones and three parental ADSC cell lines as a novel contribution to this field.

4.2.1. Gene expression profile analysis.

As commented before, ADSCs-derived hiPS clones G15.D, G15.L, G15.AE, G15.AI and G15.AO were selected to hybridize into SurePrint G3 Human Gene Expression 8x60K Microarray (Agilent); because lncRNA profiling using ADSC derived hiPSCs has not been reported before, conferring an innovative aspect to this study. The parental ADSC line G15 together with other two ADSCs lines, G10 and G17, were also analyzed. This platform covers 27.958 coding genes and 7.419 lncRNAs.

An un-supervised clustering analysis revealed that ADSCs and ADSC-hiPSCs separated into two different cell populations, the first cluster contained all parental cells and the second one all the ADSC-derived hiPS clones (Figure 19a). Those results allowed us to conclude that generated ADSCs-hiPSC were different from parental ADSCs at genomic level. Furthermore, the differential expression profile between ADSCs and their derived hiPSCs clones was analyzed, resulting in 3347 differentially expressed coding genes, being 2559 upregulated and 788 downregulated in hiPS cell lines (Figure 19b).

Then, we further analyze the similarities of our ADSCs-hiPSC with a bunch of hESC and hiPSC generated by different groups. Thus, a clustering analysis including public data from two different research groups (207,220) was performed. The first study (207) included five Fib-hiPS cell lines, five human fibroblast lines as parental cells, and two hESC lines. The second one evaluated twenty hESC lines, twelve Fib-hiPSCs lines and six human fibroblasts. In total, the gene expression profile of twenty-two hiPSCs lines, twenty-two hESC lines and sixteen somatic cells was compared to determine the genes that were differentially expressed between parental cells and pluripotent cell lines. Samples were divided into two different subgroups, one containing the differentiated cell types (i.e. fibroblasts and ADSCs) and the other one with all the pluripotent stem cells. ADSC-hiPSCs lines generated in this study were dispersed between hiPS cells and hES cells obtained by other groups, indicating a high similarity at genomic level among them (Figure 19e). Thus, we can conclude that our generated ADSCs-hiPSCs are as similar to hESCs as other hiPSCs generated by different groups, corroborating the proper induction of a pluripotent state.

Although there was some variation in global gene expression between the different ADSCs-hiPSC lines generated in this study, all of them overexpressed a set of well-known pluripotency markers (Figure 19c). The overexpression of NANOG, SOX2 or POU5F1, was already corroborated by qPCR (Figure 13). On the other hand most of the analyzed ADSCs markers including SNAI2 and CD44 that

are highly expressed in parental ADSCs, were not silenced in G15.D and G15.AO clones, results that were corroborated by qPCR; probably to the epigenetic memory maintained even after the reprogramming process (Figure 19 c and d). Even though the presence of clones G15.D and G15.AO in the analysis could mask interesting data, those samples were not excluded in the next experiments due to their pluripotent features (evidenced in section 4.1.2) and because they were comparable and similar to other pluripotent cell lines when analyzing public data (Figure 19e). Finally, these results ratify once more the differences observed between those cell populations and the effect of the reprogramming process allowing us to confirm the expression of pluripotent markers on ADSC derived hiPSCs clones and the absence in most cases of linage specific markers.

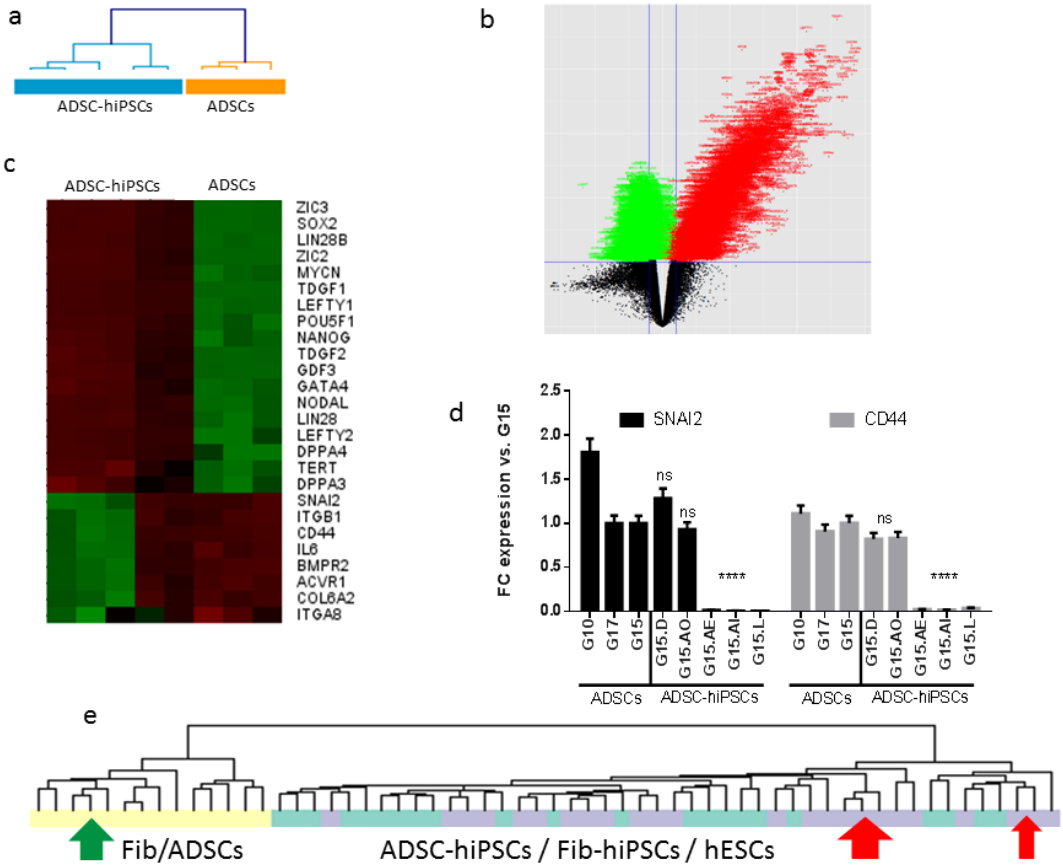
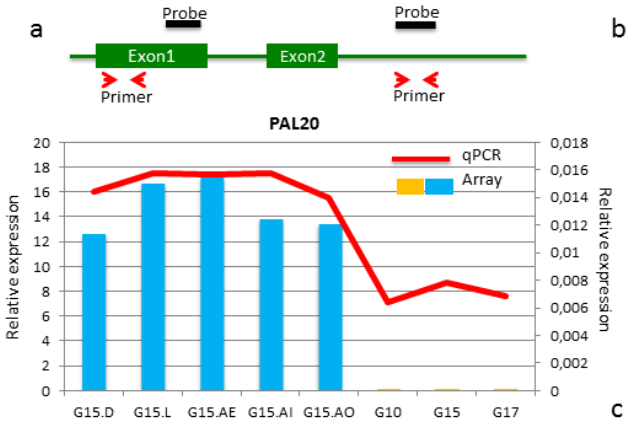


Figure 19. Somatic cells and pluripotent cells have different gene expression profiles | a) Un-supervised clustering analysis showed that parental ADSCs are different from the hiPSCs clones generated from them at genetic level b) Volcano plot showing differentially expressed coding genes between ADSCs and ADSCs derived hiPSCs, red(upregulated) green (downregulated), B>6 c) Heat map shows principal markers of pluripotency and cellular commitment are differentially expressed between ADSCs and ADSCs derived hiPSCs, B>0. d) ADSCs markers SNAI2 and CD44 that are highly expressed in parental ADSCs, down-regulated in 3 out of 5 ADSCs-hiPSC. Clones G15.AO and G15.D did not silenced those genes e) Un-supervised clustering analysis samples formed two groups, one contained 44 pluripotent cell lines (hiPSCs, hESCs) and the other one the 16 somatic cell lines. ADSC-hiPSCs generated by our group are contained in the pluripotent group. Parental ADSC cell lines: G10, G17 and G15 (green arrow). ADSCs-hiPSC cell lines: G15.AE, G15.AI, G15.L (first red arrow), G15.AO, G15.D (last red arrow).

4.2.2. LncRNAs expression profile in pluripotent cells.

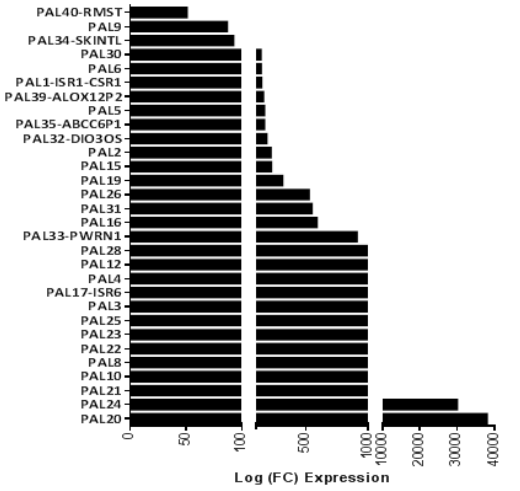
In order to search for novel lncRNAs with a possible role in maintenance of pluripotency and somatic cells reprogramming, the hybridization into the platform SurePrint G3 Human Gene Expression 8x60K Microarray allowed us to evaluate the expression of 7.419 lncRNAs. LIMMA analysis showed a total of 169 differentially expressed lncRNAs (105 up and 64 down in iPSCs with B>0) between ADSC lines and their derived iPSC clones.

In order to identify possible lncRNAS candidates a selection criterion was established. Thus we focus on those lncRNAs that i) were differentially expressed with a B value > 2; ii) had an expression fold change (FC) > 2 among parental and pluripotent samples; and iii) were annotated in at least one of these largely known databases: ENSAMBLE, RefSeq, UCSC and Genecode (http://genome.ucsc.edu/) (225).



b

PAL	Name GB	PAL	Name GB	PAL	Name GB
1	ISR1	14	CSR12	30	RP11-308N19.1
2	RP11-5N23.3	15	AC144521.1	31	RP11-218I7.2
3	RP11-554I8.2	16	RP11-44F21.5	32	DIO3OS
4	miR4697	17	ISR6	33	PWRN1
5	LINC00458	19	RP11-115D19.1	34	SKINTL
6	CTD-3006G17.2	20	EF565118	35	ABCC6P1
7	RP11-129M6.1	21	RP11-562F9.2	36	KRT8P12
8	RP6-65G23.3	22	CTC-254B4.1	37	ECEL1P2
9	SNHG14	23	CTD-2306M5.1	39	ALOX12P2
10	RP11-499F3.2	24	RP3-399L15.3	40	RMST
11	RP11-456O19.2	25	RP11-359N11.2	42	PRINS
12	AC074363.1	26	RP1-40E16.9		
13	AC007092.1	28	AC005062.2		



d

ID	GB Name
PAL2	RP11-5N23.3
PAL3	RP11-554I8.2
PAL8	RP6-65G23.3
PAL10	RP11-499F3.2
PAL12	AC074363.1
PAL15	AC144521.1
PAL20	EF565118
PAL21	RP11-562F9.2
PAL22	CTC-254B4.1
PAL23	CTD-2306M5.1
PAL40	RMST

Figure 20. Parental ADSCs and ADSC-hiPCs have different lncRNAs profiles | a) Primer design for validation of lncRNAs candidates and representative graph of the relative expression level of lncRNA-candidate in the array (blue bars) and validation by qPCR (red line). The validation of the differential expression between ADSCs and ADSCs-hiPSCs of PAL20 is indicated as example. b) List of the 37 differential expressed lncRNAs transcripts following the criteria: B>2, FC>2 and described in Genome Browser as lncRNAs, green: validated red: non-validated c) Validation of 30 lncRNAs in ADSC-hiPSCs lines by qPCR, FC: Fold change d) List of the final 11 lncRNAs candidates with a possible role in reprogramming and pluripotency, selected by a more stringent criteria. Parental ADSC cell lines: G10, G17 and G15. ADSCs-hiPSC cell lines: G15.AO, G15.D, G15.AE, G15.AI, G15.L

This search was done in order to confirm if candidates have been reported previously as lncRNAs, to know their structure, their tissue specificity, a possible described function and to obtain their sequence in order to design at least one pair of qPCR validation primers for each lncRNA candidate. As a result 37 lncRNAs were selected (36 up regulated in pluripotent cells and just one down regulated) (Figure 20b). In order to validate the microarray results, at least one set of validation primers at the exon detected by the array probe for each one of the 37 selected transcripts was designed (Figure 20a). Those candidates where the probes located within intronic genome regions were validated using primers specific for the probe (Figure 20a and Table 5).

Thirty-one lncRNAs candidates were validated by qPCR analysis, 30 of them up regulated and one downregulated in hiPS cells lines (Figure 20b, green and Figure 20c). Validated candidates were named as Pluripotency Associated LncRNA (PAL) and an arbitrary number, i.e. PAL2.

In order to select those candidates with a possible stronger role in pluripotency and/or cell reprogramming, a more stringent selection criteria was applied based on public date available through ENCODE project. Those new criteria included: i) the presence of activation marks in pluripotent cells, reflected by H3K4me3 and H3K36me3 in promoter and gene body regions respectively; ii) evidence of pluripotency core transcription factor binding sites; iii) the presence of neighboring genes and their possible function and iv) the existence of CpG islands in the promoter regions of the gene. All these criteria would help to elucidate both the mechanism of function and the regulation of selected PALs in pluripotent cells. Lastly, after this filtering process, **a group of 11 lncRNAs candidates**, all of them up regulated in hiPSCs, were selected for further characterization (Figure 20d).

4.2.3. LncRNAs characterization.

The goal of this study is to unveiled lncRNAs with a possible role on pluripotency and/or cell reprogramming, for this reason we wondered if the selected 11 lncRNA candidates were a specific feature of the undifferentiated state. Expression of the candidates among different pluripotent cell lines that were not included in the microarray study was evaluated by qPCR. Additionally, we intended to elucidate when along the reprogramming process they became active and if their expression is affected when pluripotent cells are submitted to differentiation cues.

4.2.3.1. lncRNAs candidates are expressed in diverse pluripotent cell lines.

In order to evaluate if the 11 PAL candidates expression was associated with the pluripotent state, their expression on different pluripotent cell lines was evaluated. In addition to the ADSCs-hiPSCs clones previously used in microarray platform (G15.AO, G15.D, G15.L, G15.AI and G15.AE), four fibroblast-derived hiPS cell lines (three obtained by our group and characterized in this study (M45a9, M45a30 and M45a29), and one kindly provided by Dr. Pablo Menendez (MSUH001)), two CD34-derived hiPSCs (CD34iPS1 and CD34iPS2, kindly provided by Dr. Pablo Menendez) and the NCCIT and NTERA embryonic carcinoma cell lines were included. Parental ADSCs (G10, G17 and G15), fibroblasts (H3a and H59a) and $CD34^+$ cell (CD3417 and CD34N9) cells were used as controls.

Total RNA was extracted from all samples, and the expression of the 11 PAL candidates was evaluated by qPCR with specific validation primers (Table 5). The overexpression of all the PAL

candidates was observed in all the hiPSCs analyzed (Figure 21). As expected, there was some heterogeneity in gene expression among the different clones.

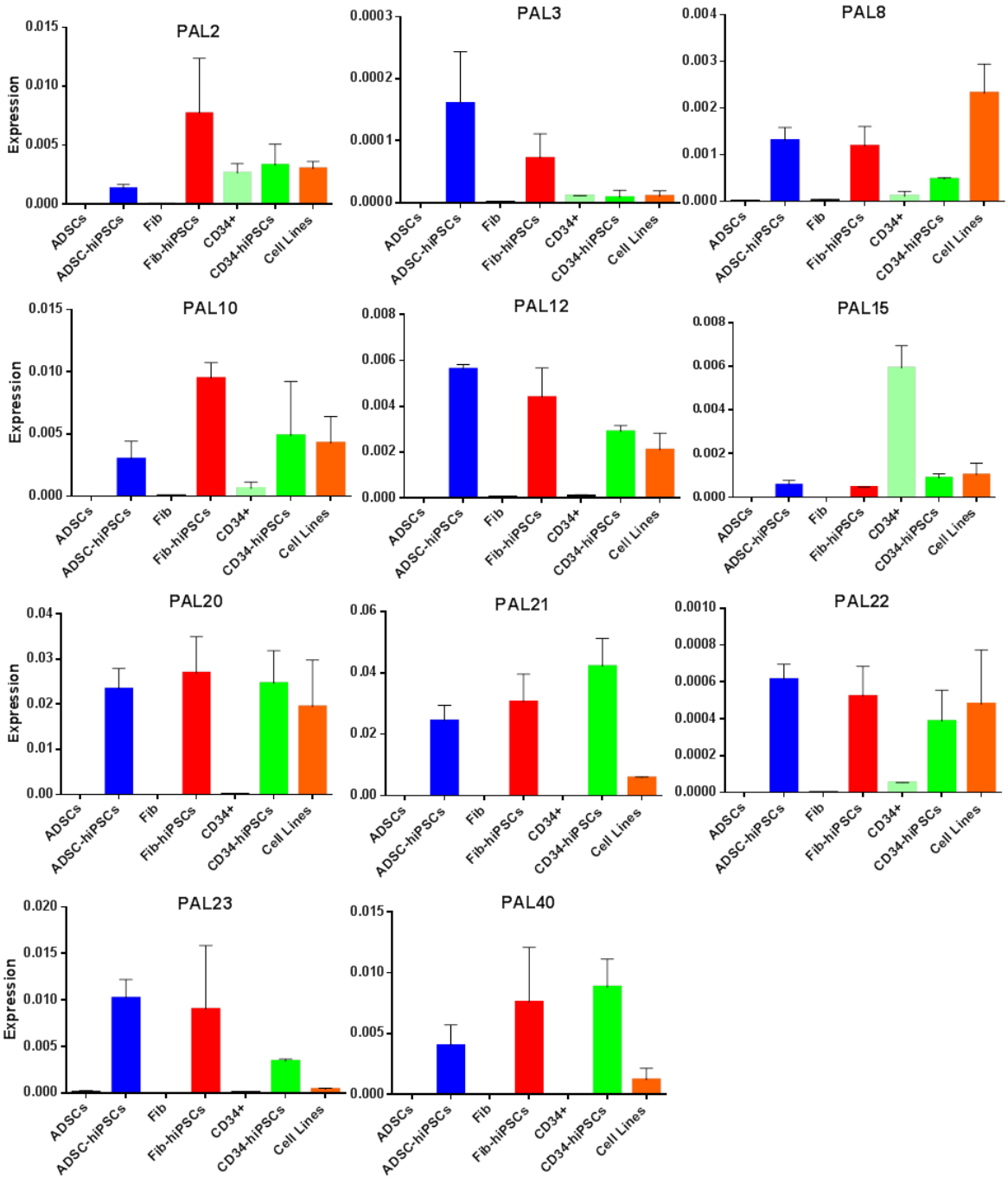


Figure 21. Expression analysis of 11 PAL candidates in pluripotent and somatic cell lines | Relative expression of 11 PAL candidates analyzed by qPCR in different samples. Mean ± SD of samples: ADSCs (G10, G17and G15); ADSCs-hiPSC (G15.AO, G15.D, G15.L, G15.AI and G15.AE); Fibroblasts (H3a and H59a); Fib-hiPSC (MSUH001, M45a9, M45a30 and M45a29); CD34+ cell lines (CD3417 and CD34N9); CD34-hiPSCs (CD34iPS1 and CD34iPS2); and human teratocarcinoma cell lines (NCCIT and NTERA).

Some candidates, like PAL10, 12, 20, 21, 23 and 40, were in general equally overexpressed in all hiPSC lines analyzed. Other candidates, like PAL8, were particularly overexpressed in cell lines (Figure 21), others, like PAL2 and 3, showed a stronger expression in Fib-hiPSCs and ADSCs-hiPSCs respectively (Figure 21). Unexpectedly, PAL15 was highly expressed in CD34+ cells; however further experiments are required to clarify this point. These results clearly demonstrate that most of our 11 PAL candidates were associated with the pluripotent state, independently of the hiPSC line origin.

4.2.3.2. LncRNAs candidates are not induced early in the reprogramming process.

Previous results showed that the 11 PAL candidates were clearly expressed in hiPSCs cells; suggesting that those transcripts must be turned on at some point in the reprogramming process. We next wonder to determine if the presence of our candidates was induced during the early stages of the reprogramming process, first 15 days, and if they could be contributing to the onset of pluripotent state.

To assess this question we used a doxycycline-inducible reprogramming system described in section 3.2.2, which allowed us to control the induction of the reprogramming factors when reprograming somatic cells. Briefly, FUW-Tet-O based lentiviral particles for the core 4F were generated and BJ cells (human fibroblasts) were infected for two consecutive days. Five days after the first infection, BJ cells were seeded at different densities on MEFs-coated plates and cultured in pluripotent cell medium with doxycycline to induce cell reprogramming. RNA was extracted at different time points after the 4F induction and the expression of the PAL candidates as well as markers of cell reprogramming was evaluated by qPCR with specific primers (Table 5). The induction of reprogramming was evidenced by the induction of NANOG (pluripotency marker) and CDH1 (MET marker) expression, with the consequent downregulation of SNAI1 (EMT marker) (Figure 22a). Nevertheless the analysis of the PAL expression did not show an early induction of any of our candidates, whereas their expression was clearly detected in established ADSC-hiPSC line G15.AO (Figure 21). As an example the expression of PAL20 and PAL21 during the reprogramming process is shown.

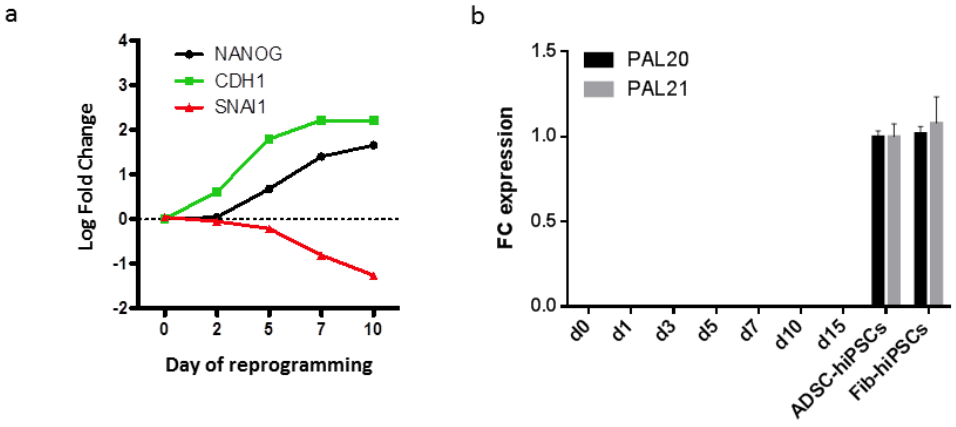


Figure 22. Expression analysis of PAL candidates during reprogramming of BJ cells. | a) Early induction of reprogramming markers NANOG and CDH1 and downregulation of SNAI1. b) A representative graph with the expression of PAL20 and PAL21 during the early reprogramming process. There was not evidenced of PAL expression at early times (Day 0 to Day15) of reprogramming. PAL expression was detected in hiPSCs. D: day. FC: Fold change.

4.2.3.3. LncRNAs candidates' expression upon spontaneous differentiation.

Since our candidates are overexpressed in pluripotent cells we wondered what would happen to their expression when cells are submitted to a differentiation process. Fib-hiPSCs line M45a28 and ADSC-hiPSCs line G15.AO were spontaneously differentiated through EBs formation, following the protocol described in Figure 11. Briefly, hiPSC clones were cultured on matrigel for five days and then cells were aggregated to form EBs and cultured on non-adherent plates. RNA was collected at day 2, 5, 8, 15, 20 and 30 of differentiation (section 3.3.7).

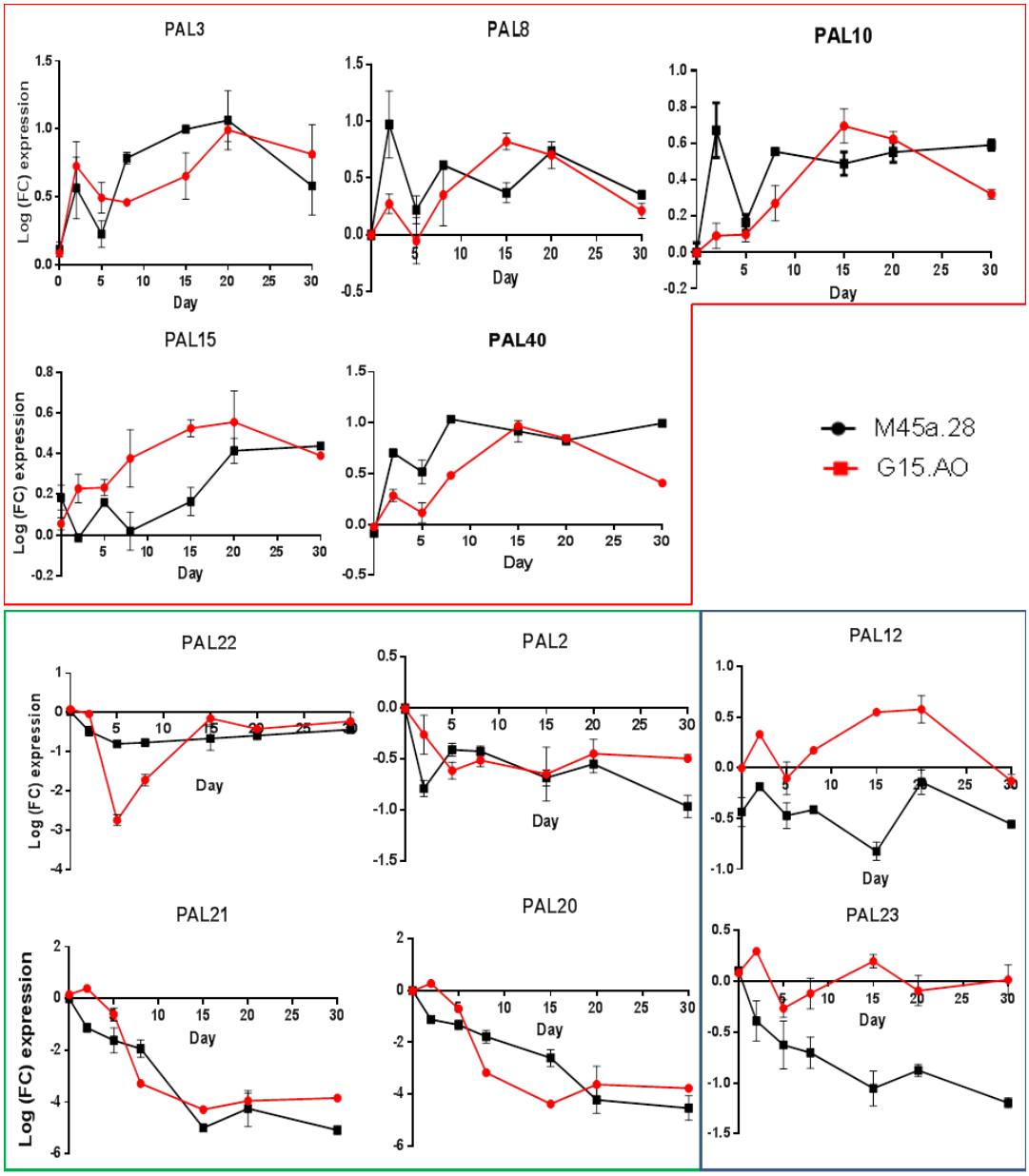


Figure 23. Expression of PAL candidates at day 20 of spontaneous differentiation induction of two hiPSCs clones | Expression of PAL3 PAL8, PAL10, PAL15 and PAL40 increased compared to day 0 of differentiation, while PAL22 and PAL2 were down-regulated. PAL12 and PAL23 were downregulated only in Fib-hiPSCs: M45a28 while transcripts PAL20 and PAL21 suffered the greatest depletion on expression, 10.000 times and maintained through differentiation process in all clones evaluated. ADSC-hiPSCs: G15.AO, Fib-hiPSCs: M45a28. FC: Fold change vs. Day0.

First we checked that the differentiation processes had been carried out properly evaluating the expression of both pluripotent and early differentiation markers. Pluripotency markers NANOG, POU5F1 and SOX2, present at day 0 of differentiation (hiPSC status), decreased their expression as early as day 2 after the induction of differentiation (Figure 15b and c). On the other hand expression of MESP1, TUBB3 and SOX17, belonging to the three germ layers, was induced early upon differentiation (Figure 15b and c). These results evidenced the loss of the un-differentiated state of hiPSCs clones.

Once we corroborated the proper spontaneous differentiation of hiPSCs clones, the expression of the 11 PAL candidates was evaluated by qPCR. Surprisingly, a diverse expression profiles among candidates along differentiation, were observed. In some candidates (PAL3 8, 10, 15 and 40) an increment on their expression was observed in both hiPSC lines evaluated. Other candidates (PAL2, 20, 21 and 12) decreased their expression in both hiPSCs, whereas PAL23 and PAL12 showed a differential expression between Fib-hiPSCs and ADSCs-hiPSCs. Interestingly PAL12 and 23 decreased their expression only in Fib-hiPSCs where they showed stronger expression (Figure 23). We focused on **PAL20 and PAL21 candidates**, since those PAL exhibited the highest reduction on gene expression among all 11 PAL candidates (>10.000 times) during the differentiation process (Figure 23). Moreover, these candidates were expressed at very high levels on hiPSCs (Figure 21) and the loss of their expression started very early in the differentiation process (Figure 23), pointing them as robust candidates in this study.

4.2.3.4. LncRNAs characterization summary.

The goal of this study is to unveiled lncRNAs with a possible role in the maintenance of pluripotent state and/or cell reprogramming. After qPCR validation, we ended up with a list of 11 promising Pluripotent associated lncRNAs (PALs), all of them up-regulated in ADSCs derived hiPSCs (Figure 20d). In order to deepen into their possible mechanism of action, we proceeded to characterize their expression among different pluripotent cell lines, during reprogramming process and finally during spontaneous differentiation of hiPSCs.

Specifically, all 11 PALs were overexpressed in all the pluripotent cell lines evaluated, even though heterogeneous expression among cell lines was observed. On the contrary, no expression of PALs was observed during the first two weeks of reprogramming process, using fibroblasts as parental cell line and a dox-inducible system of four Yamanaka´s factors. This could indicate that the presence of PALs is an exclusive feature of mature hiPS cell lines.

Finally, diverse expression profiles among candidates were observed upon differentiation. At day 20 of EBs formation some PALs increment their expression and others suffered a dramatic decrease in expression. In consequence, PAL20 and PAL21 were chosen for further functional essays since they are highly express by hiPSCs and are lost soon in the differentiation process. This result hopefully would allow us to see clear changes in phenotype when manipulating expression of PAL20 and PAL21 in both somatic and PSCs (section 4.3). Table 11 shows a summary of the characterization of 11 PAL candidates, indicating their expression in PSC lines, during EBs formation and if they have binding sites for the four reprogramming factors

Table 11 **Summary of characterization of top lncRNAs candidates**

		hiPS-cells Expression				EB´s Expression (day 20)		TF-Binding sites			
PAL	Name	hiPS-ADSCs	hiPS-Fibs	hiPS-CD34+	Cell lines	hiPS-ADSCs	hiPS-Fibs	POU5F1	SOX2	KLF4	cMYC
2	RP11-5N23.3	+	+++	+++	++	++	++	No	No	No	No
3	RP11-554I8.2	+++	++	+	+	++	+	No	No	No	No
8	RP6-65G23.3	++	++	+	++	++	++	No	No	No	No
10	RP11-499F3.2	+	++	++	++	++	+	Yes	Yes	Yes	Yes
12	AC074363.1	+++	+++	++	++	+	+	No	No	No	No
15	AC144521.1	+	+	+	+	+	+	No	No	No	No
20*	EF565118	+++	+++	+++	+++	+++	+++	No	No	No	No
21*	RP11-562F9.2	++	+++	+++	+	+++	+++	No	No	No	No
22	CTC-254B4.1	+	+	+	+	++	++	No	No	No	No
23	CTD-2306M5.1	+	+++	+	+	-	++	No	No	No	No
40	RMST	++	+++	+++	+	++	++	No	No	No	No

ADSCs cell lines: G10, G15 and G17; ADSCs-hiPSCs: G15.AO, G15.D, G15.L, G15.AI and G15.AE; Fibroblast cell lines: H3a H59a; Fib-hiPSCs: MSUH001, hiPSM45a9, hiPSM45a30, hiPSM45a29; CD34+ cell lines: CD3417 and CD34N9. CD34-hiPSCs: CD34iPS1 and CD34iPS2; Human teratocarcinoma cell lines: NCCIT, NTERA. PAL: Pluripotency associated lncRNAs. Name obtained from Human Genome Browser hg19 assembly, *http://genome.ucsc.edu.* TF: Transcription factor. EBs: Embryoid bodies. Green: upregulated; Red: downregulated; Blue: not change; Valuation code: +++ very high, ++ high, + mild; - not change. Yes: presence of TF-binding sites. No: Absence of TF-binding sites. *: Most robust candidates.

4.3. PAL 20 AND PAL21 FUNCTIONAL ANALYSIS.

To this point, the validation of the gene expression in the pluripotency state of the PAL candidates and their behavior during the differentiation process, pointed PAL20 and PAL21 as the most promising candidates with a possible important role in both the reprogramming process and the maintenance of pluripotency state. We selected those candidates for overexpression and loss of function experiments in order to elucidate their mechanism of action in pluripotent stem cells.

4.3.1. PAL20 and PAL21 genetic features.

In order to have a complete overview of our selected candidates we compile the information available in public databases as well as our own findings from PAL20 and PAL21. According to the Human Genome Browser hg19 assembly (http://genome.ucsc.edu) (225), PAL20 transcript is known as EF565118 and LOC101929194 by two databases, USC and RefSeq respectively (Figure 24). This transcript is located at chr4:93183828-93183769 in the human genome and corresponds to a non-coding RNA of approximately 1200 bp that contains three exons. PAL21 corresponds to a shorter version of PAL20 (approx. 600 bp) with two exons and it is reported as RP11-562F9.2 by different databases, Gencode and Ensemble. PAL21 is located at chr4:93190158-93190099 overlapping to PAL20 and it has been also described as lncRNA but its function still unknown. By RNA-seq a specific expression of this genomic region has been described in hESCs line H1 (Figure 24).

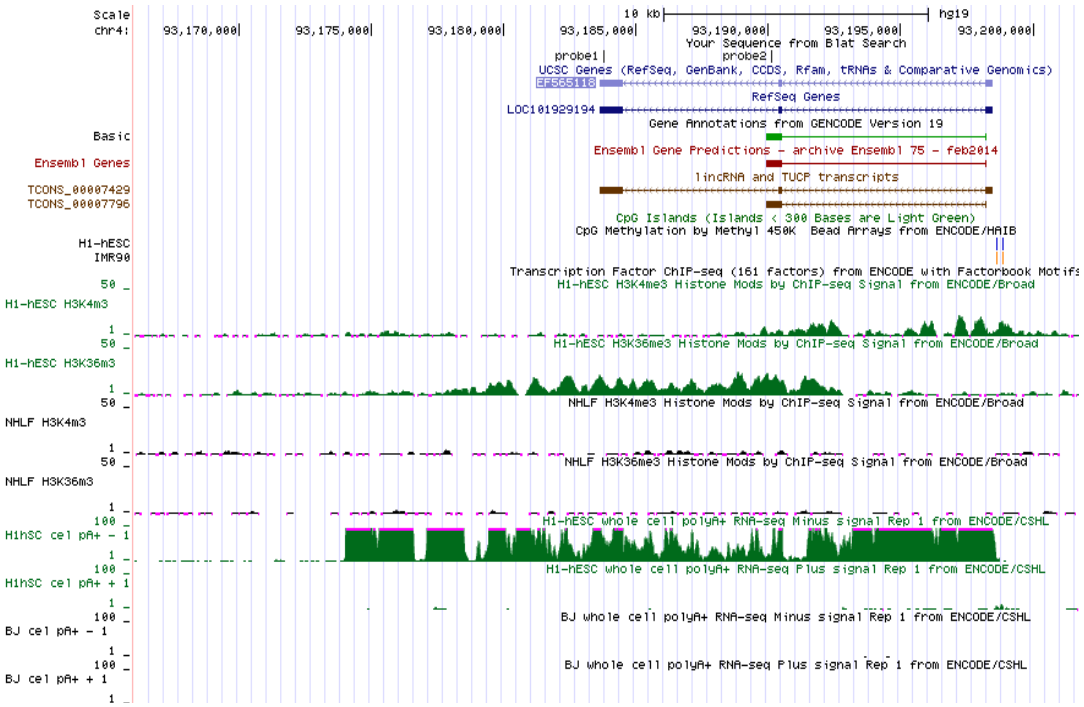


Figure 24. Genome browser screen shot from PAL20 and PAL21 candidates. PAL20 (1200 bp) is reported by USC (EF565118) and RefSeq (LOC101929194) databases, contains 3 exons. PAL21 (600bp) is reported by GENCOCE and Ensemble databases (RP11-562F9.2) and contains 2 exons. Both are reported as lncRNA by TCONs, their promoter is demethylated in hESCs, and gene body has active marks (H3K4me3 and polyA) also in hESCs compared to fibroblast lines (IMR90 and BJ). Probe1: used by the array to detect PAL20. Probe2: used to detect PAL21. Human Genome Browser hg19 assembly, *http://genome.ucsc.edu* (225).

Both transcripts share a common promoter region and the evaluation of the chromatin state of that region showed differences between hESCs and somatic cells. In particular activation histone marks H3K4me3 and H3K36me3 are present in hESC line H1 according to ChIP-seq data from ENCODE, feature that it is not observed in fibroblast cell lines (Figure 24). Moreover, although there is no CpG islands associated to this promoter region, a differential CpG methylation has been described using methyl 450K arrays between hESCs and fibroblast, being hypomethylated in hESCs. These data strongly suggest a specific PAL expression in pluripotent cells. However, ChIP-seq data from 161 different TFs did not show binding of TFs associated to pluripotency that could indicate the regulation of PAL20 and PAL21 expression by those TFs.

In our array, candidates PAL20 and PAL21 were identified by two different probes and the public data does not clarify if they correspond to two different isoforms of one transcript, for this reason, we will refer to them as different lncRNAs for further experiments.

4.3.2. Overexpression of PAL candidates.

Since our PAL candidates were strongly expressed in pluripotent cell lines, we wondered if they were important during cell reprogramming as well as if their overexpression would affect cell differentiation. Thus we studied if they could be used as reprograming factors in order to increase the efficiency of the reprogramming process and their role in the differentiation process through EBs formation.

4.3.2.1. PAL20 and PAL21 have no effect as reprogramming factors.

First we checked the effect of the inclusion of the PAL candidates in the cocktail of reprograming factors. An improvement of the reprogramming efficiency has been previously described for other hiPSC-associated lncRNAs like lincRNA-RoR when expressed together with the canonical reprogramming factors (207). For this aim, pMXs retroviral vectors overexpressing PAL candidates were generated. Thus, PAL20 (~ 1200 bp) and PAL21 (~ 600 bp) sequences were amplified by conventional PCR with *EcoRI* restriction enzyme sites to facilitate cloning in pMXs vectors. Unexpectedly, after a sequencing step, an extra exon on PAL21 fragment that has not been described before was detected and we named this new transcript as PAL21b (Figure 25).

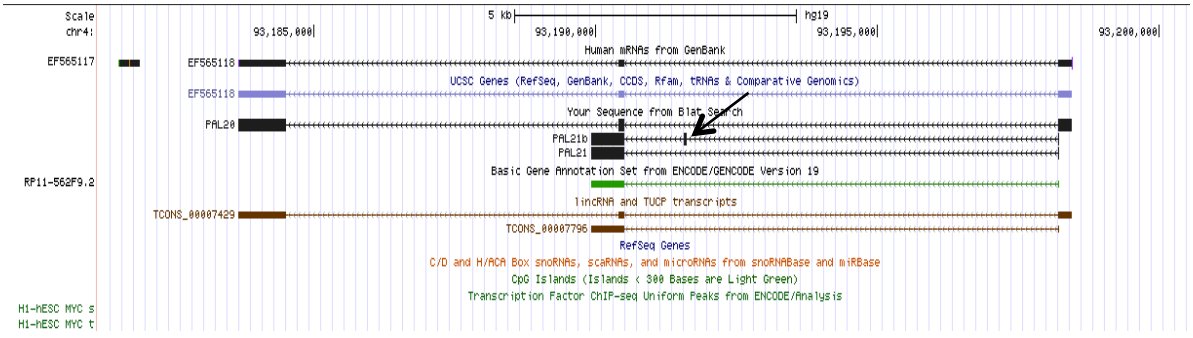


Figure 25. Genome browser screen shot from PAL20 and PAL21 candidates. Black arrow shows the newly identified exon in PAL21 after sequencing. This lncRNA candidate PAL21 with 3 exons was named PAL21b. Human Genome Browser hg19 assembly, *http://genome.ucsc.edu* (225).

Finally we generated pMXs-PAL20 and pMXs-PAL21b retroviral vectors coding for the different PAL candidates (Figure 26a). A retroviral vector coding for GFP protein (pMXs-GFP) was used as control. Concisely, retroviral particles containing PAL20 and PAL21b and GFP were generated in 293T cells and used to infect BJ cells in the presence of polybrene. Properly PAL expression from generated vectors was analyzed as described in section 0. Compared to uninfected and GFP-infected BJ cells, pMXs-PAL20 and pMXs-PAL21b significantly overexpressed PAL20 and PAL21b respectively (Figure 26b).

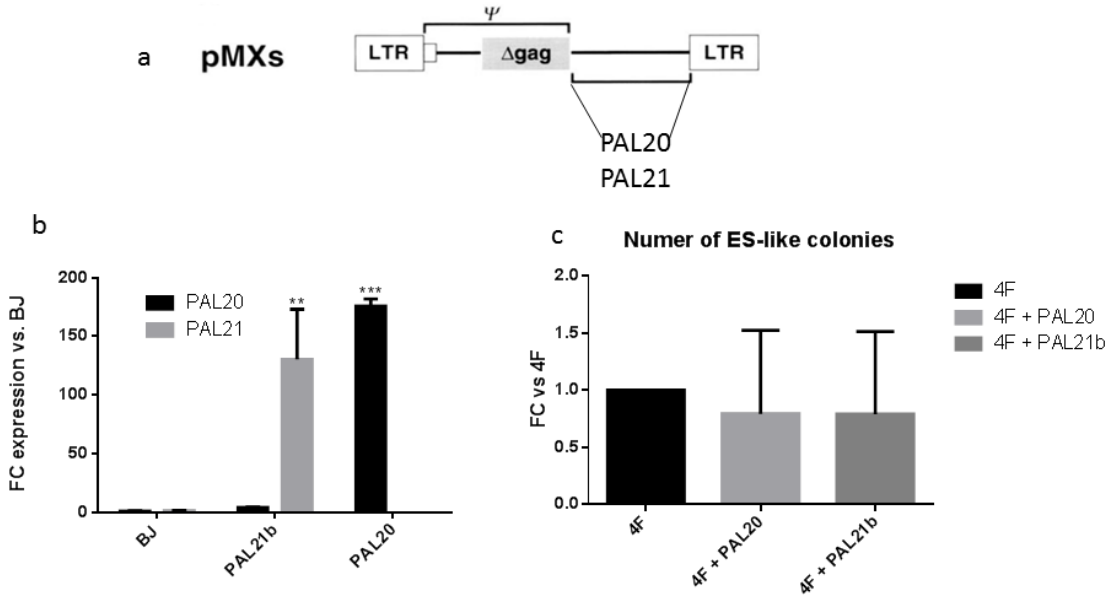


Figure 26. Overexpression of PAL20 and PAL21 lncRNAs │ a) Cloning design of PAL20 and PAL21 in pMX retroviral vectors b) pMXPAL20 and pMXPAL21b plasmids are able to overexpress PAL20 and PAL21 transcripts respectively c) Inclusion of PAL20 and PAL21b in the reprogramming cocktail do not have any effect on number of ES-like colonies compared to 4F transduction alone.

In order to elucidate if PAL20 and PAL21b expression was able to increase the efficiency of the reprogramming process, BJ cells were reprogrammed using retroviral particles for the four Yamanaka´s reprogramming factors (4F) along with both of our candidates PAL20 and PAL21b. Reprogramming protocol was performed as described in section 0 and BJ cells were infected with the 4F as positive control, and with the 4F plus PAL20, PAL21b or a combination of both. Four weeks after the first infection the number of ES-like colonies was quantified and we could not observed any difference due to the addition of PAL20 or PAL21b to the reprogramming cocktail (Figure 26c). The effect due to the combination of PAL20 and PAL21b was not corroborated since we were not able to generate colonies in the control (4F alone) when those experiments were performed. Overexpression of PAL20 and PAL21b with the reprogramming factors generated approximately 100 ES-like colonies in each treatment, with no-significant difference among them (Figure 26). These results suggest that these candidates are not enhancers of the cell reprogramming process or at least that their overexpression is not essential for reprogramming.

4.3.2.2. PAL20 and PAL21 overexpression during EBs.

Since PAL20 and PAL21 expression is really high in hiPSCs and their expression is dramatically reduced during spontaneous differentiation, we reasoned that their overexpression would impair or alter the differentiation process. Briefly, retroviral particles for the overexpression of PAL20 and PAL21b lncRNAs were generated with 293T cells as described before (section 3.5.2). Five days before infection two clones of hiPSCs (MSUH001 and G15.AO) were cultured on a thick layer of matrigel and 48 hours after infection, cells were aggregated using a scrapper, placed into non-adherent plaques and maintained on EBs media (section 3.3.7). hiPSC clones were infected with PAL20 and PAL21b viral vectors and a GPF expressing vector or untreated cells (mock condition) were used as control of infection.

Even though this experiment was performed several times, the presence of GPF+ hiPSCs in the control samples was very low, indicating that probably overexpression was not reached. Due to the lack of an effective infection protocol we could not conclude anything about the maintenance of pluripotent phenotype due to the overexpression of PAL20 and PAL21 in hiPSCs submitted to differentiation cues.

4.3.3. Inhibition of PAL candidates: Loss-of-function experiments.

We have demonstrated that PAL20 and PAL21 were highly expressed in pluripotent cells and their expression was significant depleted when cells lost the undifferentiated state. In our hands the use of PAL20 and PAL21 candidates are not enhancers of the reprogramming process since their use as reprograming factors do not increase the efficiency of the method. Nevertheless, due to their strong expression in all pluripotent cell lines evaluated in this study, loss of function analysis could help to elucidate if they are essential for cell reprogramming as well as their possible role in the maintenance of pluripotency.

4.3.3.1. shRNAs design for PAL20 and PAL21 inhibition.

Since PAL20 and PAL21b overexpression did not improve the reprogramming efficiency we aimed to know if their depletion has a role in pluripotency or during cell reprogramming. To inhibit PAL candidates, shRNAs for both transcripts were designed using two bioinformatics tools: i-Score siRNA design and siRNA Scales web site as described in section 3.6.1. Eight shRNAs sequences were designed: four specific for PAL20 (shRNA-20.1, shRNA-20.2, shRNA-20.3, and shRNA-20.4), two specific for PAL21 (shRNA-21.1 and shRNA-21.2) and two against the common regions of both lncRNAs (shRNA-2021.1 and shRNA-2021.2) (Figure 27a). shRNA sequences were cloned into pLKO.1-puro lentiviral vector (226) that allows, due to puromycin selection, the establishment of cell lines silencing the expression of target transcripts. Finally shRNA-20.1, shRNA-20.2, shRNA-20.3, shRNA-20.4, shRNA-21.1, shRNA-21.2, shRNA2021.1, shRNA2021.2 pLKO.1 plasmids were obtained and a pLKO.1 vector with a shRNAScramble sequence was used as negative control.

The specificity of shRNAs was evaluated in NCCIT human cell line since PAL20 and PAL21 expression had been previously observed and lentiviral vectors efficiently infect those cells. PAL20 and PAL21 expression analysis was performed by qPCR 72h after infection with the different pLKO.1-shRNAs vectors. A significantly silencing of PAL20 expression was observed with shRNA20.1, shRNA20.2, shRNA20.3 compared to control, while PAL 21 expression was reduced by

shRNA21.1. Some cross-reactivity in the inhibition was also observed since shRNA20.1, shRNA20.2, shRNA20.3 and chRNA21.1 showed an effect on both transcripts. However the specific inhibitory effect was always higher than the observed on the other transcript. Since both transcripts, PAL20 and PAL21, are transcribed from the same promoter, the inhibition of one of them could affect the expression of the other one. However more experiments should be performed in order to clarify this observation. On the other hand, shRNA2021.1 and shRNA2021.2, shRNA20.4, did not show any effect on silencing neither of the two candidates compared to the control. Finally shRNA20.1 and shRNA21.1 were chosen for further experiments since they both were more target-specific and achieved higher silencing levels (Figure 27b).

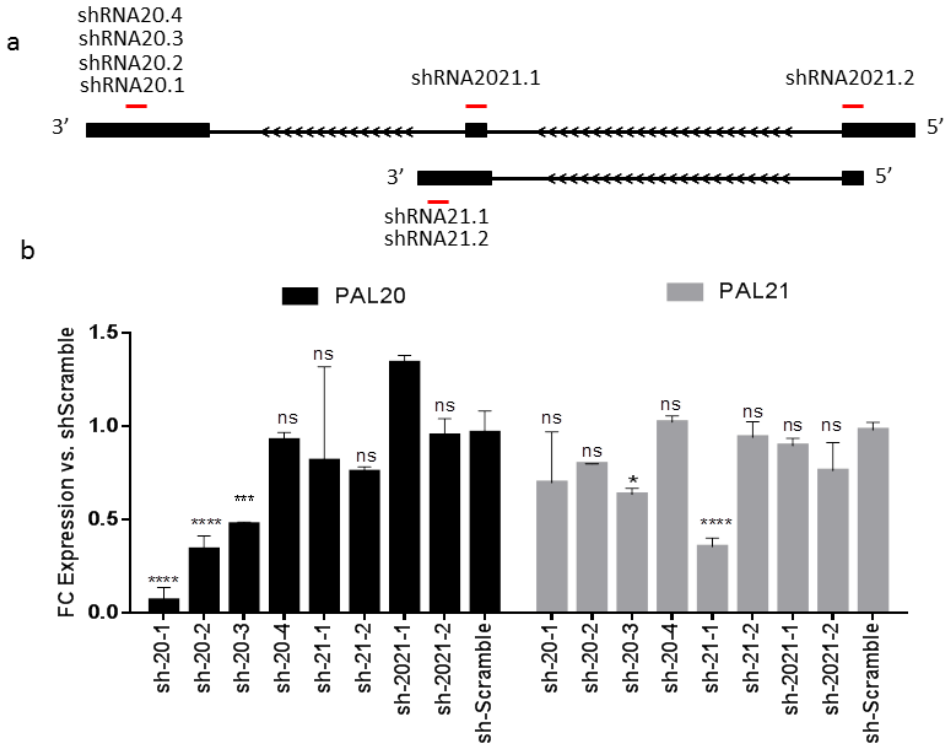


Figure 27. Loss of function experiments with shRNAs for PAL20 and PAL21 lncRNAs | a) Design of shRNAs specific for PAL20 and PAL21. b) shRNA20.1 and shRNA21.1 successfully silenced PAL20 and PAL21 expression respectively in NCCIT cell lines 72h after infection.

4.3.3.2. PAL20 and PAL21 are required for cell reprogramming.

Thus, to elucidate the effect of PAL candidate inhibition during the reprogramming process BJ cells were infected with retroviral vectors expressing the 4F, in order to induce pluripotency, along with the different combinations of pLKO.1-shRNAs.

First, the proliferation of infected BJ cell lines was measured since active proliferation is required for retroviral infection (Figure 28a). Only a slight reduced proliferation rate was observed after 7 days of cell culture, including in BJ-shRNAscramble cell line, indicating that this effect is probably due to the infection itself.

After following a standard reprogramming protocol (section 0) the number of ES-like colonies was quantified 4 weeks after infection (Figure 28b). A reduced number of ES-like colonies was observed when reprogramming was performed in the presence of shRNA20.1, shRNA21.1 or a combination of both shRNAs, indicating that PAL20 and/or PAL21 induction is required for successful cell reprogramming. However reprogramming process was not totally impaired since some ES-like colonies were still generated in the presence of PAL inhibitors. This might be explained if PAL20/PAL21 expression is not totally required for cell reprogramming or if the shRNA mediated inhibition is not totally efficient.

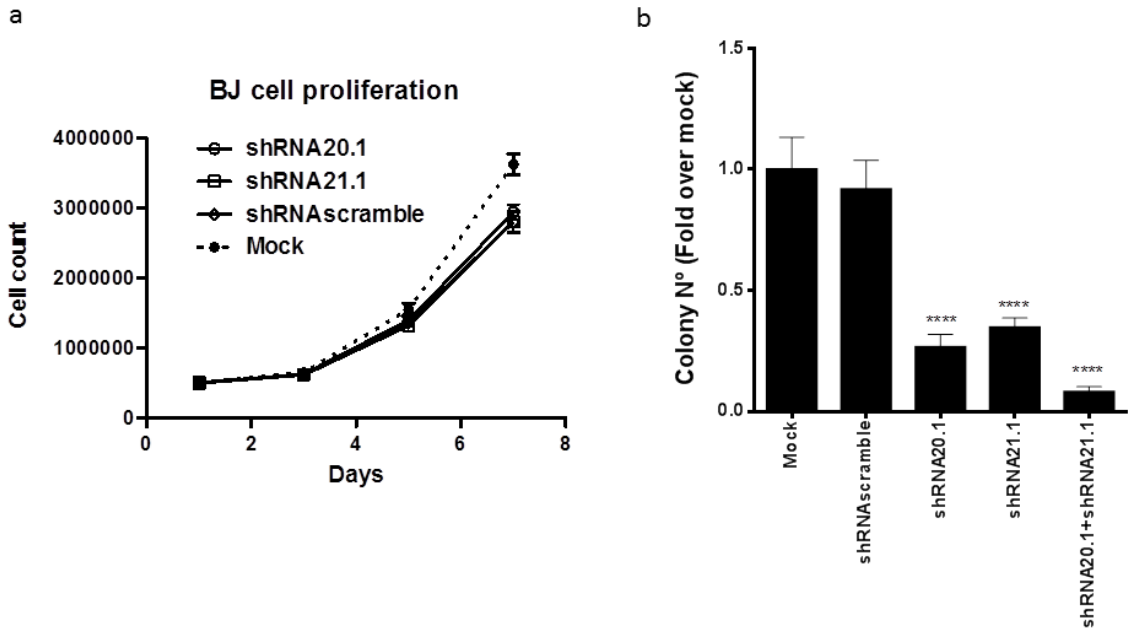


Figure 28. PAL20 and PAL21 expression is required for efficient cell reprogramming │ a) Quantification of ES-like colonies after cell reprogramming with the 4F in the presence of PAL20 and PAL21 shRNAs inhibitors. b) Proliferation of BJ cell lines with stable expression of shRNAs.

4.3.3.3. Silencing of PAL20 and PAL21 expression in human iPSCs.

On the other hand, due to PAL candidates' expression characteristics (high in hiPSCs), we reasoned that their inhibition in pluripotent cells would affect pluripotency characteristics. Initially loss-of-function experiments were performed using both Fib-hiPSC (MSUH001) and ADSCs-hiPSC (G15.AO) clones. hiPSCs clones were cultured on matrigel coated 12-well plates (section 3.1.2) and infected with lentiviral particles expression the different shRNAs (of GFP as control) within 5 days, in order to avoid the formation of large colonies, which commonly prevent viral particles to reach the major quantity of cells. Forty-eight hours after the infection, ES media was changed and GFP presence was assessed by fluorescence microscopy.

The most intense GFP expression was observed in the cells located on the periphery of the ES-colonies and few green cells inside the colonies were observed. These results suggested that shRNA viral particles could follow a similar distribution, situation that would clearly reduce the effect of the shRNAs.

RNA from MSUH001 and G15.AO hiPSCs, infected with shRNA20.1, shRNA21.1, shRNAScramble and GFP viral vectors, was obtained (section 3.3.2). A slight decrease of PAL20 and PAL21 expression was observed after infection with shRNA20.1 and shRNA21.1 respectively that was not consistently observed (Figure 29). In particular, the expression of PAL20 was significantly reduced only in one out of five experiments, although in other experiments reduced levels were also observed. This inhibition was no sufficient to induce morphological changes in infected hiPSCs. On the other hand PAL21 expression was only slightly reduced with shRNA21.1 (no significative reduction). This variability could be explained by the relative inefficient infection of the iPSCs with the lentiviral vectors.

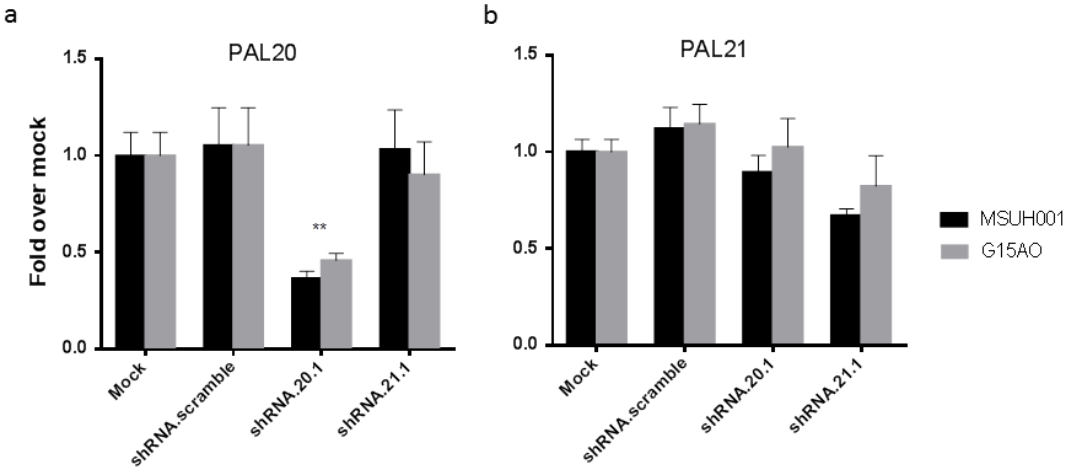


Figure 29. shRNA-mediated PAL inhibition in iPSCs. | a) shRNA20.1 reduced the expression of PAL20 in two clones of hiPSCs just one time. b) shRNA21.1 did not cause an effect on PAL21 expression in either of the clones. Clones: MSUH001 and G15AO.

Thus, we cannot make conclusions about the effect of PAL20/PAL21 inhibition in hiPSCs. A more efficient method to infect hiPSC clones must be implemented in order to achieve efficient shRNA mediated PAL inhibition. Culturing of PSCs in monolayer has been suggested as a good alternative (122,123). Currently, improvements on the protocol of infection of hiPSCs are being implementing in our Lab for further experiments.

4.3.3.4. Silencing of PAL20 and PAL21 expression in cell lines induce changes in cell proliferation.

Since no clear inhibition of PAL candidates could be observed in hiPSCs, silencing experiments were performed in the teratocarcinoma cell line NCCIT because they overexpress both PAL candidates, the efficacy of the shRNA-mediated inhibition was already observed and their culture conditions and infectivity are very reproducible. Thus, generated pLKO.1-puro vectors expressing specific PAL20 and PAL21 shRNAs (Figure 27) were used to infect NCCIT cells and NCCIT-shRNA20.1, NCCIT-shRNA21.1 and NCCIT-shRNAscramble cells lines were generated by puromycin selection. We wondered if the depletion of these candidates could affect the expression of pluripotency associated markers and/or other lncRNAs associated with pluripotency (PALs from our candidate list). Evaluation by qPCR using specific primers for NANOG, LIN28, POU5F1, PAL12 and PAL2 was performed. An overall significant decrease on gene expression was observed among

pluripotent related genes and pluripotency associated lncRNAs in NCCIT cells treated with shRNA20.1 (Figure 30a).

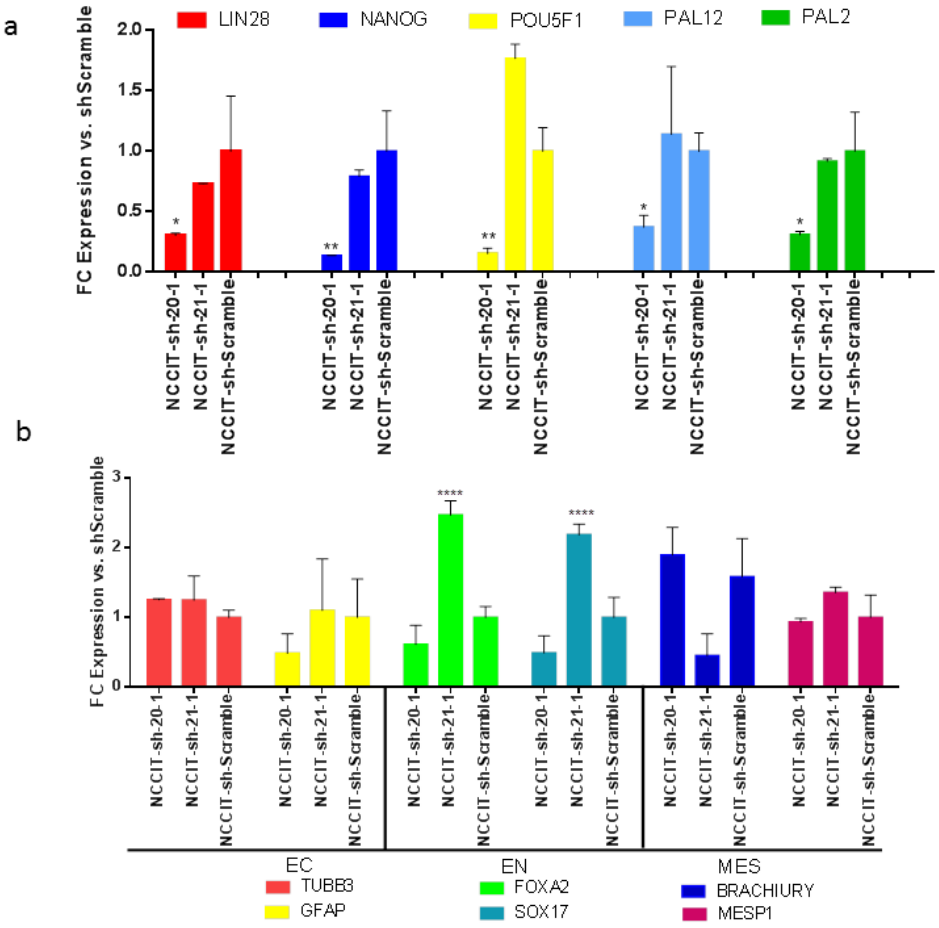


Figure 30. Effect of shRNA-mediated PAL silencing on human pluripotent cells | a) Silencing of PAL20 produced an overall significant decrease on expression of pluripotent genes NANOG, POU5F1 and LIN28 as well as PAL12 and PAL2 candidates, indicating a possible loss of pluripotency of NCCIT cell lines b) Silencing of PAL21 caused a mild upregulation of SOX17 and FOXA2 markers for endodermal differentiation. EN: Endoderm EC: Ectoderm MES: Mesoderm.

The reduction of the expression of pluripotent related genes could be indicative of a loss of the pluripotent state and made us wonder if this could be reflected on the induction of early differentiation markers. In order to answer this question the expression of early differentiation markers belonging to three germ layers was evaluated by qPCR, using specific primers as described before. However, silencing of PAL20 or PAL21 in NCCIT cell line did not lead to significant changes in expression in most of the differentiation markers analyzed. We only observed a slight increase on SOX17 and FOXA2 expression in NCCIT-shRNA21.1 cell line; well-known differentiation markers towards endoderm germ layer (Figure 30b). Further experiments to clarify a possible antagonist relationship between our candidates and those genes, are required.

Additionally to these findings, a low proliferation rate between NCCIT-shRNA20.1, NCCIT-shRNA21.1 and the controls was observed. In order to quantify this phenomenon a MTS assay was performed. Concisely NCCIT-shRNA20.1, NCCIT-shRNA21.1 and NCCIT-shRNAscramble cells were plated into 96-well plates and their proliferation was evaluated after 2, 4 and 6 days in culture. A significant decrease on cell proliferation between NCCIT-shRNA20.1 and NCCIT-shRNA21 cell lines in comparison to control cell lines was observed. Finally, both the decreased expression of pluripotent markers (after PAL20 inhibition) and the reduction of cell proliferation upon silencing of PAL20 and PAL21 constitute the first phenotypic evidence of a possible implication of those novel transcripts in the pluripotency maintenance and proliferation in pluripotent stem cells (Figure 31).

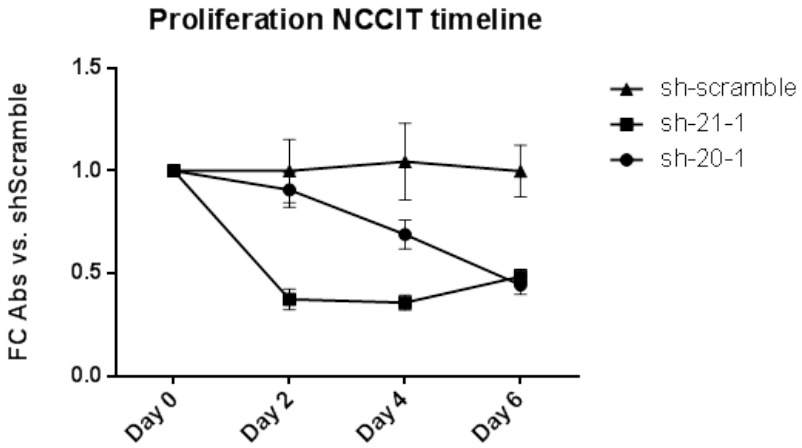


Figure 31. Effect of shRNA-mediated PAL inhibition on proliferation of NCCIT cells | MTS reflected a decrease on proliferation rate on NCCIT cell line treated with shRNAPAL20.1 and shRNAPAL21.1 at day 2, 4 and 6 of culture compared to treatment with shRNAscramble control. Those changes in proliferation remained during all the culture time. FC Abs: Fold change of absorbance

In summary, the results obtained in NCCIT cell lines upon PAL20 and PAL21 silencing (Figure 30 and Figure 31) could make us think that an evident change in the phenotype of hiPSCs would be seen after efficient inhibition of those transcripts.

4.3.4. Summary of PAL20 and PAL21 functions.

In summary, in this study we have identified and validated 11 lncRNAs transcripts up-regulated in hPSCs that we have named Pluripotency Associated LncRNAs (PAL). Two of them, PAL20 and PAL21, were dramatically down-regulated upon hiSPCs differentiation which indicates a possible role in maintenance of pluripotency and/or reprogramming. PAL20 and PAL21 overexpression in hiPSCs during EBs formation and somatic cell reprogramming did not either impair the differentiation or increase the efficiency of the process in terms of number of ES-like colonies. Silencing of PAL20 and PAL21 transcripts in NCCIT cell line caused an overall decrease of pluripotency markers. Additionally, the silencing of these transcripts in NCCIT cells induced a significant decrement on cell proliferation rate, the first phenotypic evidence of the absence of PAL20 and PAL21 in pluripotent cells. A more efficient infection methodology needs to be implemented in order to successfully silence PAL20 and PAL21 transcripts in hiPSCs.

5. DISCUSSION

5.1. DISCUSSION: GENERAL ASPECTS.

Human iPSCs have emerged as a new source of autologous therapeutic pluripotent stem cell that offers all the advantages of hESCs, like the unlimited capacity of self-renewal and the differentiation potential towards almost any cell type. Moreover iPSCs overcomes the ethical issues associated with the use of hESCs, the technical complexities of human embryo procurement and manipulation, and the possible immune rejection of hESCs due to their allogeneic origin. In addition to cell therapy applications, where iPSCs generated from a donor are differentiated to cells of a specific tissue that is injured or aged; iPS technology allows other innovative applications including the study of the molecular mechanisms involving cell reprogramming or the *in vitro* disease modeling.

The field of inducing pluripotency in somatic cells, or cellular reprogramming has come a long way since the results of the experiments of Gurdon (59) and Yamanaka (26). At present, researchers have a wide offer in culture media, delivery methods and reprogramming cocktails for cell reprogramming. Nevertheless, cellular reprogramming is still a complex and highly inefficient process that involves a high number of genetic and epigenetic modifications (87), some of them unknown. A better understanding of the factors implicated in cell reprogramming would help to address new strategies to increase the efficiency.

The control of cell fate regulated by non-coding RNA constitutes a great revolution in molecular biology. Non-coding RNAs, once considered as junk in the genome, have been proved to be master regulators of numerous cellular processes in multiple species (177,183), including embryonic development and disease (188,227). LncRNAs can performed different functions either in the nucleus, like genome organization and molecule´s trafficking leading in epigenetic and genetic changes, or in the cytoplasm by the interaction with proteins or RNAs, to inhibit or induce phenotypic features.

At the present work we aimed to further our knowledge in the reprogramming mechanisms mediated by lncRNAs. Thus, we 1) obtained hiPSCs from human fibroblasts and ADSCs and 2) studied lncRNAs signatures in order to describe novel human lncRNAs markers of the pluripotent state with a role in reprogramming of somatic cells. This second aim could lead to the further manipulation of lncRNAs in order to increase the efficiency of the process.

5.2. GENERATION OF HUMAN INDUCED PLURIPOTENT STEM CELLS.

5.2.1. Somatic cell reprogramming.

Reprogramming is a complex process that involves a number of genetic and epigenetic modifications in order to accomplish the return of a cell to its own origins. In this work, we chose a retroviral vector system for the transduction of the four Yamanaka´s factors (4Fs: POU5F1, SOX2, KLF4 and c-MYC) (26) in human fibroblasts and ADSCs. This methodology allowed us the successful generation of ES-like colonies from ten primary human lines, five fibroblasts lines and five ADSCs (Table 7), as well as from one commercial fibroblast cell line (BJ).

Retroviral vectors have been proved to be efficient delivery system for stable gene transduction into cells that are not easily transfected, such as primary cells. The efficiency of this type of vectors is based on two aspects: first, because they use receptors present in almost all cell types for the infection and second, because after one round of infection stable integration is generated and the DNA from the vector is inserted into host DNA leading to stable and constant expression of foreign transgenes. Nevertheless, active mitosis is required for the entry of the viral integration complex into the nucleus, and in consequence, this constitutes a basic feature of the cell type in order to achieve reprogramming (228). This last aspect was confirmed in our experiments where some samples that failed to form ES-like colonies corresponded to those with a low proliferation rate (i.e. SVF10) when infected with retroviral factors coding for the 4Fs. Although SVF10 was infected in a low passage, it failed to formed colonies, probably to its low proliferation rate after the isolation procedure, in comparison to other ADSC isolated in this study (data not shown). Additionally, it has been suggested the use of human somatic cells in passages between one and five for better reprogramming efficiency, since primary cells start to lose the proliferation capacity soon after *in vitro* culture (115). Thus, in our hands, the other samples that did not generate ES-like colonies (M76a and G6) were those infected at passage higher than five. However, a specific and more detailed evaluation of the relationship between passage number and reprogramming efficiency should be done in order to have a stronger correlation.

Apparently to ensure mitotic activity is not enough to obtain ES-like colonies from 4F-transduced human somatic cells since only a few of them finally are reprogrammed into an iPSC. There are other endogenous factors that can contribute to enhance cell reprogramming. In general progenitor cells or cells with some differentiation potential have been described to be easier to reprogram. As an example, progenitor hematopoietic cells seem to be more readily reprogrammed to become iPSC than more specialized blood cells (229). Following this approach, Guo *et al.*, have recently proposed the existence of a privilege state among somatic cells that make them easier to reprogram proving this hypothesis in granulocyte monocyte progenitors (230). The epigenetic status of the parental cell line have also been reported critical for the success of reprogramming protocol, not only using ectopic expression of TFs but also in cell fusion experiments, where is crucial to the outcome of the heterokaryon (65). Despite all these facts and although ADSCs present a higher differentiation potential than fibroblast, the results presented in this study do not suggest any dependence between the differentiation status of the parental cell

lines and their reprogramming capacity, at least in terms of faster colony formation or number of colonies.

Other factors that could affect ES-like colonies generation are the proper expression of the TFs and the requirements to overcome epigenetic barriers. This is supported by the stochastic and hierarchical reprogramming model (Figure 5) where the stochastic events in gene expression correspond to early stages and the hierarchic gene expression, turned on by SOX2 loci (231), distinguishes the mature iPSCs. The transduction of the four Yamanaka´s factors (POU5F1, SOX2, KLF4 and c-MYC) has been proved to generate ES-like colonies after four weeks in a great number of human cell types (127). The use of this cocktail where only two of the core pluripotency TFs (POU5F1, SOX2 and NANOG) are overexpressed, allowed us to use the internal expression of NANOG as a marker of pluripotent state in the ES-like colonies generated.

Inducible systems for the onset of pluripotency TFs have been used in order to control when the reprogramming process starts (133,232). An additional advantage of this system is the possibility of the infection of post-mitotic samples, since lentiviral vectors infect both dividing and non-dividing cells. In our study we used a standardized lentiviral dox-inducible system to reprogram human fibroblast cell line BJ (133) in order to elucidate the mechanism of action of selected lncRNAs candidates (discussed below). Nevertheless, by the time this methodology was implemented in our group we had already established iPSC clones from human fibroblast and ADSCs samples, reason why we did not use this system to generate iPSCs from SVF10, M76a or G6 samples.

Different small molecules were included in the reprogramming protocol to obtain ES-like colonies as well as during the proper establishment of iPSC clones from them. As expected ES-like colonies from human fibroblasts and ADSCs primary samples appeared after four weeks of the first infection using VPA as a supplement to the culture media during the first week. VPA has been proved to accelerate the emergence of colonies during reprogramming (99). A minimum of thirty colonies per sample were individually picked and seeded on irradiated MEFs and serum free media to establish iPSC clones, the classical method described for the colony-expansion of human PSCs (18). The clones were passaged once a week onto freshly unfrozen MEFs in the presence of ROCKi identified as enhancing survival of hESCs upon single-cell dissociation, as well as enhancing recovery from cryopreservation or passaging (233). We were able to establish multiple clones from each sample and those clones that survived after five passages were selected to evaluate they pluripotent status.

In general, to choose an appropriate reprogramming method in accordance to the specific application of the iPSC is crucial. Since the principal interest of the induced reprogramming is the future clinical applications, some aspects need to be taken into account. Viral systems are efficient but also imply random insertion of transgenes and c-MYC is a potent oncogene. Induction of pluripotency using only 3F (c-MYC free) as well as viral-free delivery systems and DNA-free methodologies, described deeply in section 1.2.2.4.2, are promising options. Nevertheless these innovative and safer methodologies are more expensive, time-consuming, methodologically complex and most of them less efficient than the one use in this work. However the retroviral

transduction of the four Yamanaka´s factors fulfilled the requirement of the present study designed to efficiently generate ES-like colonies from human somatic samples to study cell-reprograming mechanisms.

5.2.2. hiPSC characterization.

Standard procedures were followed in this study to assess the pluripotent status of the clones generated from human fibroblasts and ADSCs Fully reprogrammed hiPSCs are defined by: a) ESC-like morphological and growth properties, b) endogenous expression of pluripotency markers at gene and protein level, c) silencing of the transgenes used in the reprogramming process, d) loss of somatic cell epigenetic marks and e) *in vitro* and *in vivo* differentiation capacity.

The first evidence of the onset of the pluripotent status is the change on the morphology of the parental cells used, fibroblasts and ADSCs. Mensenchymal phenotype was lost and reprogrammed cells acquired an epithelial morphology, a process known as Mesenchymal to epithelial transition (MET). MET have been proven to be a crucial during the early phase of reprogramming and it has been demonstrated by molecular and functional evidence (234,235). Those cells start to group together and posteriorly form a colony and finally a clone. ES-like colonies are characterized by their tight, bright and rounded shape, features that are evident to the microscope, as well as for their high proliferation rate. Those characteristics were observed in all the iPSC clones generated in this work.

It is widely known that pluripotency state, both *in vivo* and *in vitro,* is governed by three principal factors POU5F1, SOX2 and NANOG, and their expression is exclusive to PSCs, being an indicative of the reprogramming efficacy in somatic cells (10). qPCR expression analysis revealed the up-regulation of endogenous POU5F1, SOX2, KLF4, c-MYC, NANOG and LIN28 genes in twenty selected clones generated from both fibroblasts and ADSCs in comparison to their parental cells. Nevertheless, we observed high heterogeneity among clones in terms on the expression of pluripotent markers like NANOG, whose expression showed different patterns. This phenomenon has been demonstrated to occur among different human ES cell lines (10) and due to their similarities to iPSCs, this behavior was also expected among the clones generated herein. The gene expression profile of a specific clone could have an effect on the phenotype of the cells in terms of self-renewal and differentiation capacities and could help to explain why some clones tend to differentiate better to certain lineages. Nevertheless, specialized studies dedicated to evaluate the effect of gene expression profiles will help to delimitate different pluripotent states and identify several reported hierarchies among PSCs. Repercussions of pluripotent gene expression on the iPSCs phenotype is an important aspect to consider when a specific application is searched. However, since the goal of this study was to generate iPSCs from human fibroblasts and ADSCs for lncRNA profiling, we chose clones with different NANOG expression levels through all the following experiments. Other evidence of the switching on the pluripotent state is the presence of specific cell surface markers like SSEA-4, TRA-1-60 and TRA-1-81 (236). In our hands, by immuno-fluorescence the presence of these representative pluripotent markers was detected in all the clones evaluated. Combination of those surface markers with FACS sorting can successfully be used to isolate iPSCs and their differentiated progeny for future cell therapy applications, since

this technique allow cultures purification from feeders or spontaneously differentiating cells, selecting only the cell subpopulations of interest (237).

Afterwards we have demonstrated the silencing of retroviral transgenes in a late passage for the majority of the clones established. In our hands ADSCs derived hiPSCs showed a more effective silencing of transgenes. This silencing of the exogenous transcription factors is required for efficient cell differentiation, however no differences were observed between hiPSC clones in the ability to form EBs or teratomas. Only one fibroblast-derived clone, M45a9 failed to silence KLF4 transgene at passage 30. Failure in silencing exogenous transgenes has been attributed to partially reprogrammed cells as well as the retention of methylated promoters of pluripotency related factors (238), further analysis in those matters need to be performed on that clone.

In addition, we corroborated the demethylation of the NANOG promoter as a specific characteristic of pluripotent cells; this experiment confirmed once more that the clones had acquired pluripotent state in comparison to the parental cells that exhibited methylated promoters. Nevertheless, one of the distal CpG seemed to be hypomethylated in both ADSC and ADSC-hiPS what could be a common feature among stem cells, but this need additional evaluation. Demethylation of POU5F1 promoter is also characteristic of ESCs (238) and in consequence of successfully reprogrammed cells. This evaluation remains to be performed in our clones in order to have a complete overview of the methylation status present in the promoters of the core pluripotency factors after cell reprogramming. On the other hand, the heterogeneity observed in the expression of NANOG and POU5F1 could be due to different demethylation patterns. Partially reprogrammed cells can express low levels of endogenous NANOG or POU5F1 in consequence of residual methylated promoters (238).

By definition pluripotent cells must give rise to cells from the three germ layers when submitted to differentiation process (1). This capacity was tested both *in vitro* and *in vivo* in the generated iPSC clones. Fib-hiPSCs and ADSC-hiPSCs evidenced changes in morphology when aggregated together *in vitro* and successfully formed EBs. As soon as day two of spontaneous differentiation, EBs from Fib-hiPSCs and ADSC-hiPSCs had a reduced expression of the pluripotent core factors and up-regulated known tissue-specific genes. *In vivo* our established clones of iPSCs gave rise to tissues of all three germ layers as expected. Differentiation potential of a hiPS cell line could define its future applications in cell therapy, drug screening and disease modeling.

Extended passage of iPSC clones was proposed to improve their epigenetic resemblance to embryonic stem cells but also has been associated to chromosomal abnormalities and genome instability (157). Reprogramming process directly affects the genome of somatic cells, as an example, reactivation of X chromosome in females is a necessary to achieve a pluripotent state in somatic cells (32). Genome stability was analyzed by copy variation number in three ADSC-hiPSCs evidencing some deletions of genetic material of variable size. The most remarkable deletion was the nullysomi of chromosome Y in one of the clones but according to the experts the aberrations observed in other clones are not associated to pluripotent phenotype and would not influence the outcome of further experiments (156).

Thus in summary, the overexpression of the four Yamanaka´s factors, using retroviral vectors, induced pluripotency in human primary fibroblasts and ADSCs isolated by our group. The clones established resemble all the characteristics of hESCs in terms of morphology, expression of pluripotent markers at genetic and protein level, silencing of the exogenous factors, demethylation of NANOG promoter and *in vitro* and *in vivo* differentiation capacity.

5.2.3. Gene expression profile.

Large-scale expression studies of the genome, epigenome and proteome of PSCs and somatic cells have revealed their differences, providing robust data. Further validation of that data has defined pluripotent exclusive markers at each of those levels. Large-scale gene expression analysis of ESCs directed the identification of both Yamanaka´s and Thomson´s factors for the later generation of induce pluripotent stem cells (26,81). Many other pluripotent markers have been discovered after that, defining new reprogramming cocktails, as has been described above (1.2.2.1).

Later on, researchers used large-scale epigenetic studies to discover that pluripotent state establishment in somatic cells requires epigenetic changes. The presence of repressive marks on certain chromatin domains has been described in differentiated cell types that confer loss of plasticity, a special characteristic of stem cells. Reprogramming process must remodel those domains in order to reflect the ESCs epigenome. However, there is a natural impairment of chromatin remodeling, usually leading to the generation of partially reprogrammed iPSCs. Finally, the evaluation of these marks can be used as discriminatory test between fully and partially reprogrammed clones (239). Additionally, this type of analysis has identified differentially methylated regions proximal to genes involved in pluripotency and differentiation between ESCs and fibroblasts (240).

Most large-scale studies have used fibroblasts as somatic cell type; which is the most expectable choice when using iPSCs derived from fibroblasts. In the present study we have chosen to hybridize into the Agilent microarray platform, five ADSC-hiPSCs and three ADSCs (including the parental ADSC sample use for hiPSC generation) for further characterization to add novelty to our study and provide new bioinformatics data that could be useful for other groups within this field. In our hands this constitutes the first study in using microarray technology to evaluate differentially expressed coding and non-coding genes in ADSCs derived iPSCs using retroviral transduction and their parental cell line. Bioinformatics analysis of coding genes demonstrated the expected differences between ADSC-hiPSCs and ADSCs, splitting them in two different groups. The expression of pluripotent makers in hiPSCs and their absence in ADSCs as well as higher expression of several ADSCs markers in somatic cells was confirmed. We further compared our array data with public data containing the coding profile of more than 40 different PSC lines (iPSC and ESCs) and the corresponding parental cells. Our samples (ADSCs-hiPSCs) were dispersed and mixed within the different pluripotent cell lines from other laboratories, certifying once more that our clones were pluripotent and that our reprogramming method was able to induce the un-differentiate state in ADSCs. In general we could conclude that our ADSCs-iPSC were as equal as other PSCs in terms of gene expression.

Surprisingly, we observed that two ADSCs-clones (G15.AO and G15.D) maintained the expression of some genes related to mesenchymal stem cells. This result can be explained by the well-known residual 'epigenetic memory' that some iPSCs can retain upon reprogramming which can lead to the generation of partially reprogrammed clones (84,239,241). However, the expression of this particular genes did not compromised their pluripotent potential since they were comparable and similar to other pluripotent cell lines when analyzing public data and they pass all the standard procedures for pluripotent stem cells characterization. The only remarkable difference is the capacity of G15.AO hiPSC clone to generate beating areas when submitted to specific differentiation experiments towards cardiac lineage that were developed by other members of our Lab. This fact was not observed with other clones, suggesting a primed state of this particular clone towards mesoderm lineages. Nevertheless, before the microarray analysis, we did not notice any differences in terms on expression of core pluripotent factors, proliferation rate or differentiation capacity (EBs formation and teratoma) among the five clones included into the array; fact that have been reported previously (238).

Since we have observed some genetic aberrations that in principle are not affecting pluripotent capacity, the hiPSC clones selected for further non-coding expression experiments were treated in the same way without any distinction. However more detailed studies need to be performed in this field (84,242).

5.3. LNCRNA PROFILING.

Somatic and pluripotent cells besides presenting different genetic and epigenetic features, they also have different non-coding gene expression profiles. Pluripotent stem cells (PSCs) express a distinctive set of miRNAs that help to maintain their unique characteristics (243); recently there has been described the implication of lncRNAs in development and disease (244,245) and even more, as markers of pluripotent state and reprogramming factors (227,239).

5.3.1. Identification of PALs.

Murine and human fibroblasts derived iPSCs have been used to unveiled lncRNAs related to pluripotency and reprogramming process (201,204,207). The present study constitutes the first using ADSCs-hiPSCs samples to discover novel pluripotency associated lncRNAs (PALs). Large scale lncRNA evaluation done by different groups correspond also to different microarrays platforms with different probes that are in most cases incompatibles due to a lack of a universal probe system. This aspect highly difficult the comparison among lncRNAs profiling obtained by different groups or the validation of a particular lncRNA into other PSCs. In our hands, specific validation of three lncRNAs upregulated in PSCs and reported by Loewer *et al.*, could not be performed designing specific primers for the probe reported in their study (207). In order to unveiled novel genes involved in specific cellular process, starting from large-scale expression data, researchers follow similar procedures: i) choose the samples, ii) choose the platform, iii) perform bioinformatics analysis, iv) select the filtering criteria, v) validate the selected candidates and vi) choose the final candidates for functional analysis.

First of all, the most representative samples from the specific process and the appropriate controls must be chosen. Since our goal is to deepen into pluripotency and reprogramming in a human tissue that has not been evaluated before, we selected iPSCs derived from ADSCs and their parental cell lines. Second, the ideal large-scale platform need to be selected, nowadays the offer is wide either for the evaluation of the genome, epigenome or proteome. We chose SurePrint G3 Human Gene Expression 8 x 60K microarray from Agilent, because we could evaluate the expression of numerous coding and non-coding genes. Thirdly, since large-scale expression platforms in general provide researchers with robust data, researchers must guide bioinformatics towards the appropriate analysis that would answer the question of interest. Herein, we first compared lncRNA profiles of the two cellular populations (ADSC-hiPSCs vs. ADSCs), probing their evident differences at this level between them, and obtaining a great number of differentially expressed lncRNAs from this first analysis (204,207). Fourthly, researchers establish restringing criteria to select from the list the most likely candidates to play a role in the cellular process of interest (section 4.2.2). This step must be followed by validation experiments to confirm the results of the array. Usually qPCR analysis is chosen and a list of less than 20 candidates is left (Figure 20d).

In our hands, the eleven up-regulated lncRNAs candidates can be considered as novel markers of the pluripotent state (PALs), since they all were validated in a large number of PSCs. Heterogeneity in expression of PALs among of PSCs and differentiated cells was expected. The universal markers of pluripotent state present different levels of expression among ESCs and iPSCs (10),

phenomenon that could be easily transposed to non-coding pluripotent markers like lncRNAs. Expression of specific factors like NANOG fluctuates along the differentiation process of PSCs towards specific lineages or EBs formation (246). PAL transcription can be driven or not by pluripotency associated TFs; in those who are, pluripotency TFs bind to promoters of lncRNAs to drive their expression and in consequence they are up-regulated in PSCs (247). In our hands, PAL10 was the only candidate with TF-biding sites for all 4 Yamanaka´s factors and one of the PALs that increased its expression during EBs formation. If transcription of PALs is dependent on pluripotent TFs and expression of those TFs is variable among PSC lines in culture and along their differentiation, is reasonable to think that the PALs whose expression is driven by those TFs can be also variable among PSCs and along differentiation.

Heterogeneity is an inherent characteristic of pluripotent stem cells that is found among the expression of different TFs, among different ESC lines, among their differentiation capacities, some clones differentiate better to certain cell lineages and even donor cell type determines future applications of iPSCs (248) and even among non-coding genes (10). Guttman *et al*., found dynamic variation on expression of some of their lncRNAs candidates, showing temporal changes across the time course of differentiation (227); additionally Rinn *et al.*, observed variation on expression when validating the lncRNAs candidates probably due to non-establishing a starting minimal quantity of RNA. As it can be seen in this study and others heterogeneity can also be found in the expression of lncRNAs transcripts related to pluripotency.

5.3.2. Mechanism of function of PAL20 and PAL21.

Lastly, overexpression and loss of function experiments of the final candidates will provide information on the effects on the phenotype and the possible mechanism of function of the candidates in the particular cellular process. Due to the high heterogeneity inherent to PSCs, those candidates, whose expression is highly evident in the more representative samples for the specific phenotype of study, are ideal for this type of experiments. High levels of expression would mean that the presence of the candidates is indispensable to maintain that specific feature and the effect of loss of function experiments is also more evident in those cases.

In the present study transcripts PAL20 and PAL21 covered those requirements since they were highly expressed in different types of PSCs and they experienced a dramatic depletion on expression levels as soon as day 2 of spontaneous differentiation. The fact that they were not induced early during the reprogramming process can be explained by the possibility that these transcripts are probably not required to initiate cell reprogramming events since we evaluated their expression just until day 15 of infection and no during the following days. This was also suggested by reprogramming experiments using PAL20 and PAL21 within the cocktail, because their presence did not resulted in a difference in the time of appearance or the number of ES-like colonies. In addition to this, expression of PALs in early reprogramming stages could not appropriately assess, probably due to the low presence of reprogramed cells within the cultures, leading to a dilution of their expression. Evaluation of expression in reprogrammed cells selected by FACS technology would help to solve this problem (237). Finally, what makes a cell pluripotent is the presence and expression of 3 TFs (SOX2, POU5F1 and NANOG). Thus, all the strategies that

lead to enhance these TFs pathways, like the presence of specific lncRNAs, will pave the way to obtain pluripotent stem cell.

In general, siRNAs have been used to silence lncRNAs but it is believed that their function is obstructed by lncRNAs secondary structure. Lentiviral particles coding for shRNAs have been proposed like an alternative for silencing lncRNAs transcripts, allowing the generation of stable cell lines that constitutively are silencing the transcripts of interest; pLKO.1-puro has been designed for this purpose (226). Recently, CRISPR-Cas9 technology has emerged as an innovative and efficient technology for gene editing, allowing the rapid generation of cellular and animal models (249). shRNA methodology was available by the time these experiments were performed and also was easily standardized in our group. Since PAL candidates were selected by their high expression in PSCs, their expression may be required for a mature pluripotent state. This hypothesis is supported by the fact that the efficiency of hiPSC generation is reduced when cells undergo reprogramming in the presence of PAL inhibitors.

As it has been emphasized before, PAL20 and PAL21 were chosen because of their high expression levels in iPSCs and their dramatic depletion upon spontaneous differentiation. These two features make the effects of the loss of function experiments easy to track down. During this study infection of hiPSCs to reach silencing of PAL20 and PAL21 were tried several times with a very limited reproducibility. It is widely known that ESCs and hiPSCs are not easily transfected or infected, probably due to their classical way of growing as colonies; which difficult the access of the desired molecules into a great number of cells within the colony and in consequence to observe an effect is unlikely. This, in addition to the percentage of failure inherent to each technique make almost impossible to make any conclusion. It is also important to notice that hiPSCs in this study were obtained by retroviral induction of 4Fs, and cells once they are infected tend to turn on programs to avoid future infections (i.e. interferon system). This last aspect is important when thinking to use viral systems for gene editing of hiPSCs, like our case.

In our hands we were able to infect hiPSCs with both retroviral and lentiviral particles coding for GFP, but the greatest intensity was obtained on the periphery of the colonies and less than 10% of cells within the colony were green. These results made us think that there was a problem with the system of culture as colonies. We tried to infect iPSCs right after passage, when they are still in smaller colonies, but the results were similar than previous ones. To solve this problem, recently new methods of culturing ESCs and hiPSCs that permit an easier transfection or infection for different applications have been proposed.

Monolayer culture of PSCs has been reported to allow high efficiencies of transfection and infection, and proceeds through four phases that include high-density single-cell plating, multicellular association, monolayer fusion and cellular condensation. This type of culture is feeder-free which allows efficient targeting of molecules to the cell of interest, it is carried out on matrigel coated dishes and conditioned media supplemented with ROCKi or JAKi (122,123,250). It has to be taken into account that due to heterogeneity among ESCs and iPSCs not all the clones are suitable for this type of growth, even though it has been ensured that works for a wide range of cell lines, each laboratory need to standardize the technique for their own clones. Currently,

improvements on the culture protocol being performed in order to adapt our generated hiPSC clones to these culture conditions for an efficient infection.

On the other hand, in this work has been proved that NCCIT cell line express the core pluripotent markers as well as PALs candidates. Contrary to hiPSCs, NCCIT grow in monolayer, which make them easy to infect and large amounts of cells are always available for the experiments, because their culture conditions and maintenance are simple. All these characteristics make them an ideal model to standardized PAL20 and PAL21 loss of function experiments in PSCs and posteriorly hypothesized the effects on the phenotype of hiPSCs due to absence of those transcripts. PAL20 and PAL21 expression were depleted by the presence of shRNAs and NCCIT stable cell lines silencing those transcripts were generated. In our hands, silencing of PAL20 and PAL21 in NCCIT cells produced two main effects respectively: a) an overall depletion of pluripotency-related markers (PAL20 inhibition) and b) a slight induction of two well-known endodermal differentiation markers, SOX7 and FOXA2 (251) (PAL21 inhibition). A third effect was also observed after PAL20 or PAL21 inhibition. In this case in both cases a significant reduction of the proliferation rate of NCCIT cells was observed.

In relation to the first aspect, lncRNAs in pluripotent cells can inhibit tissue specific programs in order to maintain the un-differentiated state by forming DNA-RNA, RNA-RNA or RNA-protein complexes. As an example, Loewer *et al.*, proved that the knocking-down of 30 lincRNAs induced gene expression programs associated with specific early differentiation lineages (207). The main difference between that study and ours is that they did not observe any changes in POU5F1 or NANOG expression upon depletion of those transcripts.

For this reason, complementary experiments are required to prove a possible action of PAL21 as an antagonist of specific differentiation TFs that prevents their binding to specific targets turning off differentiation programs and maintaining pluripotent state. Other groups have proven similar functions for LncRNA-ROR and lncMD1 who act as “sponges” for certain miRNAs-related to differentiation pathways (191,199). Moreover it has been recently described that *Pnky* lncRNA expression is downregulated during neurogenesis by the regulation of TFs related to neuronal differentiation (215). In addition, assays guided to know if the cells lacking PAL21 expression are prone to differentiate more easily towards mesendoderm germ layer could complement those results.

It must be notice that one lncRNA could perform many different functions in one cellular type, which could be the case for PAL20/PAL21 in PSCs; nevertheless in order to prove the function of PAL20/PAL21 as direct regulator of pluripotency and/or mesendoderm lineage in PSCs, some experiments need to be perform: i) lncRNA localization and ii) RNA immunoprecipitaion assays. Recently, novel and simple methodologies have been developed to answer these questions. Commercial kits are now available to perform fractionation of nuclear and cytoplasmic RNA, useful to know the localization of lncRNA, providing information of the most likely place in the cell where they are performing their function (252). In addition, RNA immunoprecipitation (RIP) technique allows evaluating the proteins and/or molecules associated to a specific RNA (i.e. mRNA, miRNA or lncRNA); using specific antibodies for the protein of interest (253).

On the other hand, we observed a decrement of proliferation rate when silencing PAL20 and PAL21, which could indicate a different possible mechanism of action. It is well known that both pluripotent and cancer cells share some cellular and molecular phenotypes. This includes high proliferation rate, plasticity in terms of differentiation, a tendency to genomic instability, a similar epigenetic status and expression pattern of non-coding RNAs, like micro-RNAs (254). Moreover, reprogramming process includes the overexpression of two known oncogenes, c-MYC and KLF4 (36,37) and characterization process of PSCs includes their capacity to generate teratomas, defined as a big mass/tumor consisting of tissues from all three embryonic germ layers (ectoderm, mesoderm and endoderm) (255). Finally, the development of strategies to avoid the tumorigenicity of PSCs for therapeutic applications constitutes a field of study within this area.

It could be likely that PSCs and cancer cells could have similar lncRNA signatures, especially if those lncRNAs are related to any of the common phenotypic characteristics like high proliferation rate. In our hands PAL20 and PAL21 shRNA mediated silencing, showed an effect on proliferation tested with a luminometric method. In addition, some methodologies commonly used to test inhibition of proliferation, like cell cycle arrest or apoptosis need to be performed to give robustness to this data. Due to PAL20 and PAL21 probable role in cell proliferation could be expected they were overexpressed also in cancer cell lines, but this experiments need to be perform in the near future.

Finally, overexpression experiments of those transcripts in cell lines silencing PAL20 and PAL21 guided to rescue the pluripotent phenotype will help to confirm the requirement of their presence to maintain this un-differentiated state.

All these data together suggest that we have identified two transcripts governed by a common promoter, which are upregulated in different cell lines of human iPSCs. Loss of function experiments made us intuit their possible mechanism of action. Specifically, the use of shRNA21-1 could lead to the onset of endodermal differentiation through SOX17 and FOXA2, on the other hand PAL20 could help to maintain the levels of pluripotency related factors. Together, PAL20 and PAL21 also could be implicated on proliferation of PSCs and their presence is required for reprogramming of human fibroblasts.

5.4. REMARKABLE CONCLUSIVE ASPECTS.

In conclusion, in this work we were able to generate and establish different lines of human iPSCs from fibroblasts and ADSCs. We probed by standard methods their pluripotency and used them as base product to unveiled novel pluripotency associated lncRNAs (PALs). We further characterized those PALs during reprogramming and differentiation process leading to the identification of two robust candidates, PAL20 and PAL21. These PALs are actively involved in the reprogramming process and apparently help to maintain pluripotent phenotype like high proliferation rate and expression of pluripotency-associated markers.

6. CONCLUSIONS

CONCLUSIONS.

1. Human iPSC can be generated from freshly isolated human fibroblasts and ADSCs by the retroviral induction of POU5F1, SOX2, KLF4 and c-MYC. Formation of ES-like colonies depends directly on two specific aspects of the start cell population: i) high proliferation rate and ii) low passage number (<5) at the time of infection. In total we have generated 25 and 57 iPSC clones from 10 different samples, 5 derived from fibroblast and 5 from ADSCs respectively.

2. Five Fib-hiPSCs and five ADSCs-hiPSCs clones generated in this work have been fully characterized being pluripotent and matching the characteristics required for iPSCs. This aspect has been confirmed by standard and well-known procedures like their characteristic morphology, the expression of specific genetic markers associated with pluripotency, the silencing of retroviral transgenes and the differentiation capacity both *in vitro* and *in vivo*.

3. The gene expression profile of ADSCs-hiPSCs has been analyzed confirming that reprogramming process changes the expression pattern of both the coding and the non-coding genes. All ADSCs-hiPSC clustered together and differed from parental cells in 3347 and 169 differentially expressed genes and lncRNAs respectively (B>0). The presence of some ADSCs markers after reprogramming in two ADSCs-hiPSC clones did not have effect on the pluripotent features analyzed like the proliferation rate, the expression of genetic markers or the differentiation capacity.

4. We have identified 37 lncRNAs associated to the pluripotent state using the selection criteria of B>2, FC>2 and their presence in at least one public database. The expression of 31 lncRNAs has been validated (30 upregulated and 1 downregulated). Further characterization of candidates using public data allowed the selection of 11 robust pluripotency associated lncRNAs (PALs).

5. The expression of the 11 selected PALs is characteristic of the un-differentiated stem cell state being all of them upregulated in a considerable number of different pluripotent cell lines, including iPSCs from different origins and teratocarcinoma cell lines.

Conclusions

6. After spontaneous differentiation 4 out of 11 PALs (PAL2, PAL20, PAL21 and PAL22) decrease their expression in both Fib-hiPSCs and ADSCs-hiPSCs. PAL20 and PAL21 have been identified as the most robust candidates due to their high expression in all pluripotent stem cells analyzed and their dramatic and promptly depletion upon spontaneous differentiation.

7. Induction of PAL20 and PAL21 expression is required for the establishment of ES-like colonies, nevertheless their use as reprogramming factors do not increase the number of iPSC colonies or accelerate their appearance.

8. Silencing of PAL20 in NCCIT cells reduces the expression of specific genetic markers like POU5F1, NANOG, LIN28, PAL2 and PAL12. In addition, PAL21 silencing lead to mild increase of two well-known specific lineage markers, SOX17 and FOXA2. The silencing of both transcripts decreases the proliferation rate of NCCIT cells.

7. BIBLIOGRAPHY

BIBLIOGRAPHY.

(1) Denham M, Conley B, Olsson F, Cole TJ, Mollard R. Stem cells: an overview. Curr Protoc Cell Biol 2005 Oct;Chapter 23:Unit 23.1.

(2) Weissman IL. Stem cells: units of development, units of regeneration, and units in evolution. Cell 2000 Jan 7;100(1):157-168.

(3) Şimşek G, Şimşek E. Stem Cells, Challenging Era Beyond Debate. 2012.

(4) Mason C, Dunnill P. A brief definition of regenerative medicine. 2008.

(5) Jansen J. The first successful allogeneic bone-marrow transplant: Georges Mathe. Transfus Med Rev 2005;19(3):246-248.

(6) Eckfeldt CE, Mendenhall EM, Verfaillie CM. The molecular repertoire of the'almighty'stem cell. Nature Reviews Molecular Cell Biology 2005;6(9):726-737.

(7) Närvä E, Autio R, Rahkonen N, Kong L, Harrison N, Kitsberg D, et al. High-resolution DNA analysis of human embryonic stem cell lines reveals culture-induced copy number changes and loss of heterozygosity. Nat Biotechnol 2010;28(4):371-377.

(8) Erdö F, Bührle C, Blunk J, Hoehn M, Xia Y, Fleischmann B, et al. Host-dependent tumorigenesis of embryonic stem cell transplantation in experimental stroke. Journal of Cerebral Blood Flow & Metabolism 2003;23(7):780-785.

(9) Shih C, Forman SJ, Chu P, Slovak M. Human embryonic stem cells are prone to generate primitive, undifferentiated tumors in engrafted human fetal tissues in severe combined immunodeficient mice. Stem cells and development 2007;16(6):893-902.

(10) Silva J, Smith A. Capturing pluripotency. Cell 2008;132(4):532-536.

(11) MacArthur BD, Lemischka IR. Statistical mechanics of pluripotency. Cell 2013;154(3):484-489.

(12) Brons IGM, Smithers LE, Trotter MW, Rugg-Gunn P, Sun B, de Sousa Lopes, Susana M Chuva, et al. Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature 2007;448(7150):191-195.

(13) Tesar PJ, Chenoweth JG, Brook FA, Davies TJ, Evans EP, Mack DL, et al. New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature 2007;448(7150):196-199.

(14) Ying Q, Nichols J, Chambers I, Smith A. BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. Cell 2003;115(3):281-292.

(15) Martin GR. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci U S A 1981 Dec;78(12):7634-7638.

(16) Nagy A, Rossant J, Nagy R, Abramow-Newerly W, Roder JC. Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. Proc Natl Acad Sci U S A 1993 Sep 15;90(18):8424-8428.

(17) De-Miguel MP, Arnalich-Montiel F, Lopez-Iglesias P, Blazquez-Martinez A, Nistal M. Epiblast-derived stem cells in embryonic and adult tissues. Int J Dev Biol 2009;53(8):1529.

(18) Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, et al. Embryonic stem cell lines derived from human blastocysts. Science 1998 Nov 6;282(5391):1145-1147.

(19) Reubinoff BE, Pera MF, Fong C, Trounson A, Bongso A. Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. Nat Biotechnol 2000;18(4):399-404.

(20) Cunningham JJ, Ulbright TM, Pera MF, Looijenga LH. Lessons from human teratomas to guide development of safe stem cell therapies. Nat Biotechnol 2012;30(9):849-857.

(21) Murry CE, Keller G. Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. Cell 2008;132(4):661-680.

(22) Assistance Publique - Hôpitaux de Paris. **Transplantation of Human Embryonic Stem Cell-derived Progenitors in Severe Heart Failure (ESCORT)**. 2014; Available at: http://clinicaltrials.gov/show/NCT02057900. Accessed 03/17, 2015.

(23) Pfizer/University College L. A Study Of Implantation Of **Human Embryonic Stem Cell** Derived Retinal Pigment Epithelium In Subjects With Acute Wet Age Related Macular Degeneration And Recent Rapid Vision Decline. 2015; Available at: http://clinicaltrials.gov/show/NCT01691261. Accessed 03/17, 2015.

(24) Matsui Y, Zsebo K, Hogan BL. Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. Cell 1992;70(5):841-847.

(25) Kimura T, Kaga Y, Sekita Y, Fujikawa K, Nakatani T, Odamoto M, et al. Pluripotent stem cells derived from mouse primordial germ cells by small molecule compounds. Stem Cells 2015;33(1):45-55.

(26) Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 2006;126(4):663-676.

(27) Nichols J, Zevnik B, Anastassiadis K, Niwa H, Klewe-Nebenius D, Chambers I, et al. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell 1998;95(3):379-391.

(28) Niwa H, Miyazaki J, Smith AG. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet 2000;24(4):372-376.

(29) Masui S, Nakatake Y, Toyooka Y, Shimosato D, Yagi R, Takahashi K, et al. Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. Nat Cell Biol 2007;9(6):625-635.

(30) Niwa H. How is pluripotency determined and maintained? Development 2007 Feb;134(4):635-646.

(31) Silva J, Nichols J, Theunissen TW, Guo G, van Oosten AL, Barrandon O, et al. Nanog is the gateway to the pluripotent ground state. Cell 2009;138(4):722-737.

(32) Maherali N, Sridharan R, Xie W, Utikal J, Eminli S, Arnold K, et al. Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. Cell stem cell 2007;1(1):55-70.

(33) Tada M, Takahama Y, Abe K, Nakatsuji N, Tada T. Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. Current Biology 2001;11(19):1553-1558.

(34) Chambers I, Silva J, Colby D, Nichols J, Nijmeijer B, Robertson M, et al. Nanog safeguards pluripotency and mediates germline development. Nature 2007;450(7173):1230-1234.

(35) Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M, Takahashi K, et al. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell 2003;113(5):631-642.

(36) Jiang J, Chan Y, Loh Y, Cai J, Tong G, Lim C, et al. A core Klf circuitry regulates self-renewal of embryonic stem cells. Nat Cell Biol 2008;10(3):353-360.

(37) Smith KN, Singh AM, Dalton S. Myc represses primitive endoderm differentiation in pluripotent stem cells. Cell stem cell 2010;7(3):343-354.

(38) Marson A, Foreman R, Chevalier B, Bilodeau S, Kahn M, Young RA, et al. Wnt signaling promotes reprogramming of somatic cells to pluripotency. Cell Stem Cell 2008 Aug 7;3(2):132-135.

(39) O'Donnell KA, Wentzel EA, Zeller KI, Dang CV, Mendell JT. c-Myc-regulated microRNAs modulate E2F1 expression. Nature 2005;435(7043):839-843.

(40) Chambers I, Smith A. Self-renewal of teratocarcinoma and embryonic stem cells. Oncogene 2004;23(43):7150-7160.

(41) Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, et al. Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 2005;122(6):947-956.

(42) Niwa H, Burdon T, Chambers I, Smith A. Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. Genes Dev 1998;12(13):2048-2060.

(43) Ben-Shushan E, Thompson JR, Gudas LJ, Bergman Y. Rex-1, a gene encoding a transcription factor expressed in the early embryo, is regulated via Oct-3/4 and Oct-6 binding to an octamer site and a novel protein, Rox-1, binding to an adjacent site. Mol Cell Biol 1998 Apr;18(4):1866-1878.

(44) Jaenisch R, Young R. Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. Cell 2008;132(4):567-582.

(45) Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 1999;99(3):247-257.

(46) Kaji K, Caballero IM, MacLeod R, Nichols J, Wilson VA, Hendrich B. The NuRD component Mbd3 is required for pluripotency of embryonic stem cells. Nat Cell Biol 2006;8(3):285-292.

(47) Gao X, Tate P, Hu P, Tjian R, Skarnes WC, Wang Z. ES cell pluripotency and germ-layer formation require the SWI/SNF chromatin remodeling component BAF250a. Proc Natl Acad Sci U S A 2008 May 6;105(18):6656-6661.

(48) Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G. Genome regulation by polycomb and trithorax proteins. Cell 2007;128(4):735-745.

(49) Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 2006;125(2):315-326.

(50) Bar M, Wyman SK, Fritz BR, Qi J, Garg KS, Parkin RK, et al. MicroRNA discovery and profiling in human embryonic stem cells by deep sequencing of small RNA libraries. Stem Cells 2008;26(10):2496-2505.

(51) Suh M, Lee Y, Kim JY, Kim S, Moon S, Lee JY, et al. Human embryonic stem cells express a unique set of microRNAs. Dev Biol 2004;270(2):488-498.

(52) Melton C, Judson RL, Blelloch R. Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. Nature 2010;463(7281):621-626.

(53) Wang Y, Baskerville S, Shenoy A, Babiarz JE, Baehner L, Blelloch R. Embryonic stem cell–specific microRNAs regulate the G1-S transition and promote rapid proliferation. Nat Genet 2008;40(12):1478-1483.

(54) Gill JG, Langer EM, Lindsley RC, Cai M, Murphy TL, Kyba M, et al. Snail and the microRNA-200 Family Act in Opposition to Regulate Epithelial-to-Mesenchymal Transition and Germ Layer Fate Restriction in Differentiating ESCs. Stem Cells 2011;29(5):764-776.

(55) Wolff G. Entwicklungsphysiologische Studien. 1. Die Regeneration der Urodelenlinse. Arch. Entw.-Mech. Or 1895;1:380-390.

(56) Hanna J, Saha K, Jaenisch R. Pluripotency and cellular reprogramming: facts, hypotheses, unresolved issues. Cell 2010;143(4):508-525.

(57) Spemann H. Embryonic development and induction. : Hafner Publishing Company; 1962.

(58) Briggs R, King TJ. Transplantation of Living Nuclei From Blastula Cells into Enucleated Frogs' Eggs. Proc Natl Acad Sci U S A 1952 May;38(5):455-463.

(59) Gurdon JB, Elsdale TR, Fischberg M. Sexually mature individuals of Xenopus laevis from the transplantation of single somatic nuclei. 1958.

(60) Wilmut,I., Schnieke A,E., McWhir,J., Kind A,J., Campbell KH,S. Viable offspring derived from fetal and adult mammalian cells. Nature 1997;385:810-813.

(61) Wakayama T, Perry AC, Zuccotti M, Johnson KR, Yanagimachi R. Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei. Nature 1998;394(6691):369-374.

(62) Tachibana M, Amato P, Sparman M, Gutierrez NM, Tippner-Hedges R, Ma H, et al. Human embryonic stem cells derived by somatic cell nuclear transfer. Cell 2013;153(6):1228-1238.

(63) Thuan NV, Kishigami S, Wakayama T. How to improve the success rate of mouse cloning technology. Journal of Reproduction and Development 2010;56(1):20-30.

(64) Simonsson S, Gurdon J. DNA demethylation is necessary for the epigenetic reprogramming of somatic cell nuclei. Nat Cell Biol 2004;6(10):984-990.

(65) Yamanaka S, Blau HM. Nuclear reprogramming to a pluripotent state by three approaches. Nature 2010;465(7299):704-712.

(66) Blau HM, Chiu C, Webster C. Cytoplasmic activation of human nuclear genes in stable heterocaryons. Cell 1983;32(4):1171-1180.

(67) Harris H, Miller O. Suppression of malignancy by cell fusion. Nature 1969;223:363-368.

(68) Peterson JA, Weiss MC. Expression of differentiated functions in hepatoma cell hybrids: induction of mouse albumin production in rat hepatoma-mouse fibroblast hybrids. Proc Natl Acad Sci U S A 1972 Mar;69(3):571-575.

(69) Tada M, Takahama Y, Abe K, Nakatsuji N, Tada T. Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. Current Biology 2001;11(19):1553-1558.

(70) Cowan CA, Atienza J, Melton DA, Eggan K. Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. Science 2005 Aug 26;309(5739):1369-1373.

(71) Jaenisch R. Nuclear cloning and direct reprogramming: the long and the short path to Stockholm. Cell Stem Cell 2012;11(6):744-747.

(72) Farah MH, Olson JM, Sucic HB, Hume RI, Tapscott SJ, Turner DL. Generation of neurons by transient expression of neural bHLH proteins in mammalian cells. Development 2000 Feb;127(4):693-702.

(73) Schäfer BW, Blakely BT, Darlington GJ, Blau HM. Effect of cell history on response to helix loop helix family of myogenic regulators. 1990.

(74) Zhou Q, Brown J, Kanarek A, Rajagopal J, Melton DA. In vivo reprogramming of adult pancreatic exocrine cells to &bgr;-cells. Nature 2008;455(7213):627-632.

(75) Vierbuchen T, Ostermeier A, Pang ZP, Kokubu Y, Südhof TC, Wernig M. Direct conversion of fibroblasts to functional neurons by defined factors. Nature 2010;463(7284):1035-1041.

(76) Sekiya S, Suzuki A. Direct conversion of mouse fibroblasts to hepatocyte-like cells by defined factors. Nature 2011;475(7356):390-393.

(77) Ieda M, Fu J, Delgado-Olguin P, Vedantham V, Hayashi Y, Bruneau BG, et al. Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. Cell 2010;142(3):375-386.

(78) Wernig M, Meissner A, Foreman R, Brambrink T, Ku M, Hochedlinger K, et al. In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature 2007;448(7151):318-324.

(79) Okita K, Ichisaka T, Yamanaka S. Generation of germline-competent induced pluripotent stem cells. Nature 2007;448(7151):313-317.

(80) Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 2007;131(5):861-872.

(81) Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, et al. Induced pluripotent stem cell lines derived from human somatic cells. Science 2007 Dec 21;318(5858):1917-1920.

(82) Buganim Y, Faddah DA, Jaenisch R. Mechanisms and models of somatic cell reprogramming. Nature Reviews Genetics 2013;14(6):427-439.

(83) Buganim Y, Faddah DA, Cheng AW, Itskovich E, Markoulaki S, Ganz K, et al. Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. Cell 2012;150(6):1209-1222.

(84) Polo JM, Liu S, Figueroa ME, Kulalert W, Eminli S, Tan KY, et al. Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. Nat Biotechnol 2010;28(8):848-855.

(85) Hansson J, Rafiee MR, Reiland S, Polo JM, Gehring J, Okawa S, et al. Highly coordinated proteome dynamics during reprogramming of somatic cells to pluripotency. Cell reports 2012;2(6):1579-1592.

(86) Waddington CH. The strategy of the genes. A discussion of some aspects of theoretical biology. With an appendix by H. Kacser. The strategy of the genes.A discussion of some aspects of theoretical biology.With an appendix by H.Kacser. 1957:ix.

(87) Yamanaka S. Elite and stochastic models for induced pluripotent stem cell generation. Nature 2009;460(7251):49-52.

(88) Hussein SM, Puri MC, Tonge PD, Benevento M, Corso AJ, Clancy JL, et al. Genome-wide characterization of the routes to pluripotency. Nature 2014;516(7530):198-206.

(89) Tonge PD, Corso AJ, Monetti C, Hussein SM, Puri MC, Michael IP, et al. Divergent reprogramming routes lead to alternative stem-cell states. Nature 2014;516(7530):192-197.

(90) Lee D, Shin J, Tonge PD, Puri MC, Lee S, Park H, et al. An epigenomic roadmap to induced pluripotency reveals DNA methylation as a reprogramming modulator. Nature communications 2014;5.

(91) Benevento M, Tonge PD, Puri MC, Hussein SM, Cloonan N, Wood DL, et al. Proteome adaptation in cell reprogramming proceeds via distinct transcriptional networks. Nature communications 2014;5.

(92) Clancy JL, Patel HR, Hussein SM, Tonge PD, Cloonan N, Corso AJ, et al. Small RNA changes en route to distinct cellular states of induced pluripotency. Nature communications 2014;5.

(93) Zhang J, Nuebel E, Daley GQ, Koehler CM, Teitell MA. Metabolic regulation in pluripotent stem cells during reprogramming and self-renewal. Cell stem cell 2012;11(5):589-595.

(94) Sancho-Martinez I, Belmonte JCI. Stem cells: Surf the waves of reprogramming. Nature 2013;493(7432):310-311.

(95) Smith ZD, Nachman I, Regev A, Meissner A. Dynamic single-cell imaging of direct reprogramming reveals an early specifying event. Nat Biotechnol 2010;28(5):521-526.

(96) Chan EM, Ratanasirintrawoot S, Park I, Manos PD, Loh Y, Huo H, et al. Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells. Nat Biotechnol 2009;27(11):1033-1037.

(97) Zhao W, Ning B, Qian C. Regulatory factors of induced pluripotency: current status. Stem Cell Investigation 2014;1:15.

(98) Nakagawa M, Koyanagi M, Tanabe K, Takahashi K, Ichisaka T, Aoi T, et al. Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nat Biotechnol 2008;26(1):101-106.

(99) Huangfu D, Osafune K, Maehr R, Guo W, Eijkelenboom A, Chen S, et al. Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. Nat Biotechnol 2008;26(11):1269-1275.

(100) Redmer T, Diecke S, Grigoryan T, Quiroga-Negreira A, Birchmeier W, Besser D. E-cadherin is crucial for embryonic stem cell pluripotency and can replace OCT4 during somatic cell reprogramming. EMBO Rep 2011 Jul 1;12(7):720-726.

(101) Yoshida Y, Takahashi K, Okita K, Ichisaka T, Yamanaka S. Hypoxia enhances the generation of induced pluripotent stem cells. Cell stem cell 2009;5(3):237-241.

(102) Esteban MA, Xu J, Yang J, Peng M, Qin D, Li W, et al. Generation of induced pluripotent stem cell lines from Tibetan miniature pig. J Biol Chem 2009 Jun 26;284(26):17634-17640.

(103) Marión RM, Strati K, Li H, Murga M, Blanco R, Ortega S, et al. A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. Nature 2009;460(7259):1149-1153.

(104) Li H, Collado M, Villasante A, Strati K, Ortega S, Cañamero M, et al. The Ink4/Arf locus is a barrier for iPS cell reprogramming. Nature 2009;460(7259):1136-1139.

(105) Shi Y, Desponts C, Do JT, Hahm HS, Schöler HR, Ding S. Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. Cell stem cell 2008;3(5):568-574.

(106) Huangfu D, Maehr R, Guo W, Eijkelenboom A, Snitow M, Chen AE, et al. Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. Nat Biotechnol 2008;26(7):795-797.

(107) Mali P, Chou B, Yen J, Ye Z, Zou J, Dowey S, et al. Butyrate Greatly Enhances Derivation of Human Induced Pluripotent Stem Cells by Promoting Epigenetic Remodeling and the Expression of Pluripotency-Associated Genes. Stem Cells 2010;28(4):713-720.

(108) Hou P, Li Y, Zhang X, Liu C, Guan J, Li H, et al. Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. Science 2013 Aug 9;341(6146):651-654.

(109) De Los Angeles A, Daley GQ. A chemical logic for reprogramming to pluripotency. Cell Res 2013;23(12):1337-1338.

(110) Yamanaka S. A fresh look at iPS cells. Cell 2009;137(1):13-17.

(111) Luo J, Suhr S, Chang E, Wang K, Ross P, Nelson L, et al. Generation of LIF and bFGF-dependent induced pluripotent stem cells from canine adult somatic cells. Stem Cells Dev 2011;20:1669-1678.

(112) Li W, Wei W, Zhu S, Zhu J, Shi Y, Lin T, et al. Generation of rat and human induced pluripotent stem cells by combining genetic reprogramming and chemical inhibitors. Cell stem cell 2009;4(1):16-19.

(113) Li Y, Cang M, Lee AS, Zhang K, Liu D. Reprogramming of sheep fibroblasts into pluripotency under a drug-inducible expression of mouse-derived defined factors. PloS one 2011;6(1):e15947.

(114) Liu H, Zhu F, Yong J, Zhang P, Hou P, Li H, et al. Generation of induced pluripotent stem cells from adult rhesus monkey fibroblasts. Cell stem cell 2008;3(6):587-590.

(115) Aasen T, Raya A, Barrero MJ, Garreta E, Consiglio A, Gonzalez F, et al. Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. Nat Biotechnol 2008;26(11):1276-1284.

(116) Kim K, Zhao R, Doi A, Ng K, Unternaehrer J, Cahan P, et al. Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. Nat Biotechnol 2011;29(12):1117-1119.

(117) Vallier L, Alexander M, Pedersen RA. Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. J Cell Sci 2005 Oct 1;118(Pt 19):4495-4509.

(118) Ludwig TE, Bergendahl V, Levenstein ME, Yu J, Probasco MD, Thomson JA. Feeder-independent culture of human embryonic stem cells. Nature methods 2006;3(8):637-646.

(119) Chen G, Gulbranson DR, Hou Z, Bolin JM, Ruotti V, Probasco MD, et al. Chemically defined conditions for human iPSC derivation and culture. Nature methods 2011;8(5):424-429.

(120) Ezashi T, Das P, Roberts RM. Low O2 tensions and the prevention of differentiation of hES cells. Proc Natl Acad Sci U S A 2005 Mar 29;102(13):4783-4788.

(121) Forsyth NR, Musio A, Vezzoni P, Simpson AHR, Noble BS, McWhir J. Physiologic oxygen enhances human embryonic stem cell clonal recovery and reduces chromosomal abnormalities. Cloning and stem cells 2006;8(1):16-23.

(122) Kunova M, Matulka K, Eiselleova L, Salykin A, Kubikova I, Kyrylenko S, et al. Adaptation to robust monolayer expansion produces human pluripotent stem cells with improved viability. Stem Cells Transl Med 2013 Apr;2(4):246-254.

(123) Chen KG, Mallon BS, Hamilton RS, Kozhich OA, Park K, Hoeppner DJ, et al. Non-colony type monolayer culture of human embryonic stem cells. Stem cell research 2012;9(3):237-248.

(124) Gerecht-Nir S, Cohen S, Itskovitz-Eldor J. Bioreactor cultivation enhances the efficiency of human embryoid body (hEB) formation and differentiation. Biotechnol Bioeng 2004;86(5):493-502.

(125) Krawetz R, Taiani JT, Liu S, Meng G, Li X, Kallos MS, et al. Large-scale expansion of pluripotent human embryonic stem cells in stirred-suspension bioreactors. Tissue Engineering Part C: Methods 2009;16(4):573-582.

(126) Kitamura T, Koshino Y, Shibata F, Oki T, Nakajima H, Nosaka T, et al. Retrovirus-mediated gene transfer and expression cloning: powerful tools in functional genomics. Exp Hematol 2003;31(11):1007-1014.

(127) González F, Boué S, Belmonte JCI. Methods for making induced pluripotent stem cells: reprogramming a la carte. Nature Reviews Genetics 2011;12(4):231-242.

(128) Brambrink T, Foreman R, Welstead GG, Lengner CJ, Wernig M, Suh H, et al. Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. Cell stem cell 2008;2(2):151-159.

(129) Hanna J, Wernig M, Markoulaki S, Sun CW, Meissner A, Cassady JP, et al. Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin. Science 2007 Dec 21;318(5858):1920-1923.

(130) Sommer CA, Sommer AG, Longmire TA, Christodoulou C, Thomas DD, Gostissa M, et al. Excision of reprogramming transgenes improves the differentiation potential of iPS cells generated with a single excisable vector. Stem Cells 2010;28(1):64-74.

(131) Cary LC, Goebel M, Corsaro BG, Wang H, Rosen E, Fraser M. Transposon mutagenesis of baculoviruses: analysis of Trichoplusia ni transposon IFP2 insertions within the FP-locus of nuclear polyhedrosis viruses. Virology 1989;172(1):156-169.

(132) Lacoste A, Berenshteyn F, Brivanlou AH. An efficient and reversible transposable system for gene delivery and lineage-specific differentiation in human embryonic stem cells. Cell Stem Cell 2009;5(3):332-342.

(133) Wernig M, Lengner CJ, Hanna J, Lodato MA, Steine E, Foreman R, et al. A drug-inducible transgenic system for direct reprogramming of multiple somatic cell types. Nat Biotechnol 2008;26(8):916-924.

(134) Woltjen K, Michael I, Mohseni P, Desai R, Mileikovsky M, Hmlinen R, et al. piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. Nature 2009;458(7239):766-770.

(135) Kaji K, Norrby K, Paca A, Mileikovsky M, Mohseni P, Woltjen K. Virus-free induction of pluripotency and subsequent excision of reprogramming factors. Nature 2009;458(7239):771-775.

(136) Stadtfeld M, Nagaya M, Utikal J, Weir G, Hochedlinger K. Induced pluripotent stem cells generated without viral integration. Science 2008;322(5903):945.

(137) Zhou W, Freed CR. Adenoviral gene delivery can reprogram human fibroblasts to induced pluripotent stem cells. Stem Cells 2009;27(11):2667-2674.

(138) Yu J, Hu K, Smuga-Otto K, Tian S, Stewart R, Slukvin II, et al. Human induced pluripotent stem cells free of vector and transgene sequences. Science 2009 May 8;324(5928):797-801.

(139) Jia F, Wilson KD, Sun N, Gupta DM, Huang M, Li Z, et al. A nonviral minicircle vector for deriving human iPS cells. Nature methods 2010;7(3):197-199.

(140) Warren L, Manos PD, Ahfeldt T, Loh Y, Li H, Lau F, et al. Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. Cell stem cell 2010;7(5):618-630.

(141) Miyoshi N, Ishii H, Nagano H, Haraguchi N, Dewi DL, Kano Y, et al. Reprogramming of mouse and human cells to pluripotency using mature microRNAs. Cell stem cell 2011;8(6):633-638.

(142) Zhou H, Wu S, Joo JY, Zhu S, Han DW, Lin T, et al. Generation of induced pluripotent stem cells using recombinant proteins. Cell stem cell 2009;4(5):381-384.

(143) Kim D, Kim CH, Moon JI, Chung YG, Chang MY, Han BS, et al. Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. Cell Stem Cell 2009 Jun 5;4(6):472-476.

(144) Inoue M, Tomizawa K, Matsushita M, Lu Y, Yokoyama T, Yanai H, et al. p53 protein transduction therapy: successful targeting and inhibition of the growth of the bladder cancer cells. Eur Urol 2006;49(1):161-168.

(145) Wadia JS, Dowdy SF. Protein transduction technology. Curr Opin Biotechnol 2002;13(1):52-56.

(146) Fusaki N, Ban H, Nishiyama A, Saeki K, Hasegawa M. Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. Proc Japan Acad , Ser B 2009;85(8):348-362.

(147) Seki T, Yuasa S, Oda M, Egashira T, Yae K, Kusumoto D, et al. Generation of induced pluripotent stem cells from human terminally differentiated circulating T cells. Cell stem cell 2010;7(1):11-14.

(148) Ban H, Nishishita N, Fusaki N, Tabata T, Saeki K, Shikamura M, et al. Efficient generation of transgene-free human induced pluripotent stem cells (iPSCs) by temperature-sensitive Sendai virus vectors. Proc Natl Acad Sci U S A 2011 Aug 23;108(34):14234-14239.

(149) Zhang Y, Li W, Laurent T, Ding S. Small molecules, big roles -- the chemical manipulation of stem cell fate and somatic cell reprogramming. J Cell Sci 2012 Dec 1;125(Pt 23):5609-5620.

(150) Federation AJ, Bradner JE, Meissner A. The use of small molecules in somatic-cell reprogramming. Trends Cell Biol 2014;24(3):179-187.

(151) Hanna JH, Saha K, Jaenisch R. Pluripotency and cellular reprogramming: facts, hypotheses, unresolved issues. Cell 2010;143(4):508-525.

(152) Gafni O, Weinberger L, Mansour AA, Manor YS, Chomsky E, Ben-Yosef D, et al. Derivation of novel human ground state naive pluripotent stem cells. Nature 2013;504(7479):282-286.

(153) Newman AM, Cooper JB. Lab-specific gene expression signatures in pluripotent stem cells. Cell stem cell 2010;7(2):258-262.

(154) Chin MH, Mason MJ, Xie W, Volinia S, Singer M, Peterson C, et al. Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. Cell stem cell 2009;5(1):111-123.

(155) Bock C, Kiskinis E, Verstappen G, Gu H, Boulting G, Smith ZD, et al. Reference maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. Cell 2011;144(3):439-452.

(156) Ma H, Morey R, O'Neil RC, He Y, Daughtry B, Schultz MD, et al. Abnormalities in human pluripotent cells due to reprogramming mechanisms. Nature 2014;511(7508):177-183.

(157) Chin MH, Pellegrini M, Plath K, Lowry WE. Molecular analyses of human induced pluripotent stem cells and embryonic stem cells. Cell stem cell 2010;7(2):263-269.

(158) Lister R, Pelizzola M, Kida YS, Hawkins RD, Nery JR, Hon G, et al. Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. Nature 2011;471(7336):68-73.

(159) Park I, Arora N, Huo H, Maherali N, Ahfeldt T, Shimamura A, et al. Disease-specific induced pluripotent stem cells. Cell 2008;134(5):877-886.

(160) Carvajal-Vergara X, Sevilla A, D'Souza SL, Ang Y, Schaniel C, Lee D, et al. Patient-specific induced pluripotent stem-cell-derived models of LEOPARD syndrome. Nature 2010;465(7299):808-812.

(161) Rashid ST, Corbineau S, Hannan N, Marciniak SJ, Miranda E, Alexander G, et al. Modeling inherited metabolic disorders of the liver using human induced pluripotent stem cells. J Clin Invest 2010 Sep;120(9):3127-3136.

(162) Brennand KJ, Simone A, Jou J, Gelboin-Burkhart C, Tran N, Sangar S, et al. Modelling schizophrenia using human induced pluripotent stem cells. Nature 2011;473(7346):221-225.

(163) Soldner F, Hockemeyer D, Beard C, Gao Q, Bell GW, Cook EG, et al. Parkinson's disease patient-derived induced pluripotent stem cells free of viral reprogramming factors. Cell 2009;136(5):964-977.

(164) Kondo T, Asai M, Tsukita K, Kutoku Y, Ohsawa Y, Sunada Y, et al. Modeling Alzheimer's disease with iPSCs reveals stress phenotypes associated with intracellular Aβ and differential drug responsiveness. Cell stem cell 2013;12(4):487-496.

(165) Sun N, Longaker MT, Wu JC. Human iPS cell-based therapy: considerations before clinical applications. Cell Cycle 2010;9(5):880-885.

(166) Eiraku M, Sasai Y. Self-formation of layered neural structures in three-dimensional culture of ES cells. Curr Opin Neurobiol 2012;22(5):768-777.

(167) Suga H, Kadoshima T, Minaguchi M, Ohgushi M, Soen M, Nakano T, et al. Self-formation of functional adenohypophysis in three-dimensional culture. Nature 2011;480(7375):57-62.

(168) Lancaster MA, Renner M, Martin C, Wenzel D, Bicknell LS, Hurles ME, et al. Cerebral organoids model human brain development and microcephaly. Nature 2013;501(7467):373-379.

(169) Takebe T, Sekine K, Enomura M, Koike H, Kimura M, Ogaeri T, et al. Vascularized and functional human liver from an iPSC-derived organ bud transplant. Nature 2013;499(7459):481-484.

(170) McCracken KW, Catá EM, Crawford CM, Sinagoga KL, Schumacher M, Rockich BE, et al. Modelling human development and disease in pluripotent stem-cell-derived gastric organoids. Nature 2014;516(7531):400-404.

(171) Egawa N, Kitaoka S, Tsukita K, Naitoh M, Takahashi K, Yamamoto T, et al. Drug screening for ALS using patient-specific induced pluripotent stem cells. Sci Transl Med 2012 Aug 1;4(145):145ra104.

(172) Ebert AD, Yu J, Rose FF, Mattis VB, Lorson CL, Thomson JA, et al. Induced pluripotent stem cells from a spinal muscular atrophy patient. Nature 2009;457(7227):277-280.

(173) Marchetto MC, Carromeu C, Acab A, Yu D, Yeo GW, Mu Y, et al. A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. Cell 2010;143(4):527-539.

(174) World Health Organization (WHO). **Priority eye diseases/Age-related macular degeneration**
 . 2012; Available at: http://www.who.int/blindness/causes/priority/en/index7.html. Accessed 03/17, 2015.

(175) Cyranoski D, Hoag H, Callaway E, Hayden EC, Witze A, Lloyd R, et al. News Japanese woman is first recipient of next-generation stem cells. Nature 2014;513:294-295.

(176) Reardon S, Cyranoski D. Japan stem-cell trial stirs envy. Nature 2014 Sep 18;513(7518):287-288.

(177) Mattick JS. Non-coding RNAs: the architects of eukaryotic complexity. EMBO Rep 2001;2(11):986-991.

(178) Cech TR, Steitz JA. The noncoding RNA revolution—trashing old rules to forge new ones. Cell 2014;157(1):77-94.

(179) Ulitsky I, Bartel DP. lincRNAs: genomics, evolution, and mechanisms. Cell 2013;154(1):26-46.

(180) Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev 2011 Sep 15;25(18):1915-1927.

(181) Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res 2012 Sep;22(9):1775-1789.

(182) Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, et al. Circular RNAs are a large class of animal RNAs with regulatory potency. Nature 2013;495(7441):333-338.

(183) Pauli A, Valen E, Lin MF, Garber M, Vastenhouw NL, Levin JZ, et al. Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis. Genome Res 2012 Mar;22(3):577-591.

(184) Ponjavic J, Oliver PL, Lunter G, Ponting CP. Genomic and transcriptional co-localization of protein-coding and long non-coding RNA pairs in the developing brain. PLoS genetics 2009;5(8):e1000617.

(185) Ulitsky I, Shkumatava A, Jan CH, Sive H, Bartel DP. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. Cell 2011;147(7):1537-1550.

(186) Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. Annu Rev Biochem 2012;81:145-166.

(187) Dinger ME, Pang KC, Mercer TR, Mattick JS. Differentiating protein-coding and noncoding RNA: challenges and ambiguities. PLoS computational biology 2008;4(11):e1000176.

(188) Paralkar VR, Weiss MJ. Long noncoding RNAs in biology and hematopoiesis. Blood 2013 Jun 13;121(24):4842-4846.

(189) Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science 2008 Oct 31;322(5902):750-756.

(190) Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell 2007;129(7):1311-1323.

(191) Wang Y, Xu Z, Jiang J, Xu C, Kang J, Xiao L, et al. Endogenous miRNA sponge lincRNA-RoR regulates Oct4, Nanog, and Sox2 in human embryonic stem cell self-renewal. Developmental cell 2013;25(1):69-80.

(192) Zappulla DC, Cech TR. RNA as a flexible scaffold for proteins: yeast telomerase and beyond. Cold Spring Harb Symp Quant Biol 2006;71:217-224.

(193) Spitale RC, Tsai M, Chang HY. RNA templating the epigenome: long noncoding RNAs as molecular scaffolds. Epigenetics 2011;6(5):539-543.

(194) Espinoza CA, Allen TA, Hieb AR, Kugel JF, Goodrich JA. B2 RNA binds directly to RNA polymerase II to repress transcript synthesis. Nature structural & molecular biology 2004;11(9):822-829.

(195) Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. Mol Cell 2011;43(6):904-914.

(196) Yang L, Lin C, Jin C, Yang JC, Tanasa B, Li W, et al. lncRNA-dependent mechanisms of androgen-receptor-regulated gene activation programs. Nature 2013;500(7464):598-602.

(197) Bertani S, Sauer S, Bolotin E, Sauer F. The noncoding RNA Mistral activates Hoxa6 and Hoxa7 expression and stem cell differentiation by recruiting MLL1 to chromatin. Mol Cell 2011;43(6):1040-1046.

(198) Kretz M, Siprashvili Z, Chu C, Webster DE, Zehnder A, Qu K, et al. Control of somatic tissue differentiation by the long non-coding RNA TINCR. Nature 2013;493(7431):231-235.

(199) Cesana M, Cacchiarelli D, Legnini I, Santini T, Sthandier O, Chinappi M, et al. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. Cell 2011;147(2):358-369.

(200) Klattenhoff CA, Scheuermann JC, Surface LE, Bradley RK, Fields PA, Steinhauser ML, et al. Braveheart, a long noncoding RNA required for cardiovascular lineage commitment. Cell 2013;152(3):570-583.

(201) Lin N, Chang K, Li Z, Gates K, Rana ZA, Dang J, et al. An evolutionarily conserved long noncoding RNA TUNA controls pluripotency and neural lineage commitment. Mol Cell 2014;53(6):1005-1019.

(202) Han P, Li W, Lin C, Yang J, Shang C, Nuernberg ST, et al. A long noncoding RNA protects the heart from pathological hypertrophy. Nature 2014.

(203) Guttman M, Garber M, Levin JZ, Donaghey J, Robinson J, Adiconis X, et al. Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. Nat Biotechnol 2010;28(5):503-510.

(204) Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 2009;458(7235):223-227.

(205) Sheik Mohamed J, Gaughwin PM, Lim B, Robson P, Lipovich L. Conserved long noncoding RNAs transcriptionally regulated by Oct4 and Nanog modulate pluripotency in mouse embryonic stem cells. RNA 2010 Feb;16(2):324-337.

(206) Guttman M, Donaghey J, Carey BW, Garber M, Grenier JK, Munson G, et al. lincRNAs act in the circuitry controlling pluripotency and differentiation. Nature 2011;477(7364):295-300.

(207) Loewer S, Cabili MN, Guttman M, Loh Y, Thomas K, Park IH, et al. Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. Nat Genet 2010;42(12):1113-1117.

(208) Wang Y, Xu Z, Jiang J, Xu C, Kang J, Xiao L, et al. Endogenous miRNA sponge lincRNA-RoR regulates Oct4, Nanog, and Sox2 in human embryonic stem cell self-renewal. Developmental cell 2013;25(1):69-80.

(209) Galasso M, Dama P, Previati M, Sandhu S, Palatini J, Coppola V, et al. A large scale expression study associates uc.283-plus lncRNA with pluripotent stem cells and human glioma. Genome Medicine 2014;6:76.

(210) Ng S, Johnson R, Stanton LW. Human long non-coding RNAs promote pluripotency and neuronal differentiation by association with chromatin modifiers and transcription factors. EMBO J 2012;31(3):522-533.

(211) Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science 2008 Oct 31;322(5902):750-756.

(212) Lee J, Davidow LS, Warshawsky D. Tsix, a gene antisense to Xist at the X-inactivation centre. Nat Genet 1999;21(4):400-404.

(213) Payer B, Rosenberg M, Yamaji M, Yabuta Y, Koyanagi-Aoi M, Hayashi K, et al. Tsix RNA and the germline factor, PRDM14, link X reactivation and stem cell reprogramming. Mol Cell 2013;52(6):805-818.

(214) Keniry A, Oxley D, Monnier P, Kyba M, Dandolo L, Smits G, et al. The H19 lincRNA is a developmental reservoir of miR-675 that suppresses growth and Igf1r. Nat Cell Biol 2012;14(7):659-665.

(215) Ramos AD, Andersen RE, Liu SJ, Nowakowski TJ, Hong SJ, Gertz CC, et al. The Long Noncoding RNA Pnky Regulates Neuronal Differentiation of Embryonic and Postnatal Neural Stem Cells. Cell stem cell 2015.

(216) Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. Bioinformatics methods and protocols: Springer; 1999. p. 365-386.

(217) Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative CT method. Nature protocols 2008;3(6):1101-1108.

(218) Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Statistical applications in genetics and molecular biology 2004;3(1):1-25.

(219) Gentleman R, Carey V, Huber W, Irizarry R, Dudoit S. Bioinformatics and computational biology solutions using R and Bioconductor. : Springer Science & Business Media; 2006.

(220) Bock C, Kiskinis E, Verstappen G, Gu H, Boulting G, Smith ZD, et al. Reference maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. Cell 2011;144(3):439-452.

(221) Ichihara M, Murakumo Y, Masuda A, Matsuura T, Asai N, Jijiwa M, et al. Thermodynamic instability of siRNA duplex is a prerequisite for dependable prediction of siRNA activities. Nucleic Acids Res 2007;35(18):e123.

(222) Matveeva O, Nechipurenko Y, Rossi L, Moore B, Saetrom P, Ogurtsov AY, et al. Comparison of approaches for rational siRNA design leading to a new efficient and transparent method. Nucleic Acids Res 2007;35(8):e63.

(223) Keller GM. In vitro differentiation of embryonic stem cells. Curr Opin Cell Biol 1995;7(6):862-869.

(224) Nisler BS, Taapken SM, Leonhard K, McIntire E, Montgomery KD. Validation of karyotypic abnormalities in human pluripotent stem cells: cytogenetic progression in WA09 mirrors cancer. Cancer Genetics 2012;205(7):423.

(225) Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. Genome Res 2002 Jun;12(6):996-1006.

(226) Stewart SA, Dykxhoorn DM, Palliser D, Mizuno H, Yu EY, An DS, et al. Lentivirus-delivered stable gene silencing by RNAi in primary cells. RNA 2003 Apr;9(4):493-501.

(227) Guttman M, Donaghey J, Carey BW, Garber M, Grenier JK, Munson G, et al. lincRNAs act in the circuitry controlling pluripotency and differentiation. Nature 2011;477(7364):295-300.

(228) Cepko C, Pear W. Overview of the retrovirus transduction system. Current Protocols in Molecular Biology 2001:9.9. 1-9.9. 16.

(229) Eminli S, Foudi A, Stadtfeld M, Maherali N, Ahfeldt T, Mostoslavsky G, et al. Differentiation stage determines potential of hematopoietic cells for reprogramming into induced pluripotent stem cells. Nat Genet 2009;41(9):968-976.

(230) Guo S, Zi X, Schulz VP, Cheng J, Zhong M, Koochaki SH, et al. Nonstochastic reprogramming from a privileged somatic cell state. Cell 2014;156(4):649-662.

(231) Hanna J, Saha K, Pando B, Van Zon J, Lengner CJ, Creyghton MP, et al. Direct cell reprogramming is a stochastic process amenable to acceleration. Nature 2009;462(7273):595-601.

(232) Somers A, Jean J, Sommer C, Omari A, Ford C, Mills J, et al. Generation of transgene-free lung disease-specific human induced pluripotent stem cells using a single excisable lentiviral stem cell cassette. Stem Cells 2010;28(10):1728-1740.

(233) Emre N, Vidal JG, Elia J, O'Connor ED, Paramban RI, Hefferan MP, et al. The ROCK inhibitor Y-27632 improves recovery of human embryonic stem cells after fluorescence-activated cell sorting with multiple cell surface markers. PLoS One 2010;5(8):e12148.

(234) Li R, Liang J, Ni S, Zhou T, Qing X, Li H, et al. A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. Cell stem cell 2010;7(1):51-63.

(235) Samavarchi-Tehrani P, Golipour A, David L, Sung H, Beyer TA, Datti A, et al. Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. Cell stem cell 2010;7(1):64-77.

(236) Adewumi O, Aflatoonian B, Ahrlund-Richter L, Amit M, Andrews PW, Beighton G, et al. Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. Nat Biotechnol 2007;25(7):803-816.

(237) Bilousova G, Chen J, Roop DR. Differentiation of mouse induced pluripotent stem cells into a multipotent keratinocyte lineage. J Invest Dermatol 2011;131(4):857-864.

(238) Mikkelsen TS, Hanna J, Zhang X, Ku M, Wernig M, Schorderet P, et al. Dissecting direct reprogramming through integrative genomic analysis. Nature 2008;454(7200):49-55.

(239) Hawkins RD, Hon GC, Lee LK, Ngo Q, Lister R, Pelizzola M, et al. Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. Cell stem cell 2010;6(5):479-491.

(240) Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 2009;462(7271):315-322.

(241) Shao K, Koch C, Gupta MK, Lin Q, Lenz M, Laufs S, et al. Induced pluripotent mesenchymal stromal cell clones retain donor-derived differences in DNA methylation profiles. Molecular Therapy 2013;21(1):240-250.

(242) Kim K, Zhao R, Doi A, Ng K, Unternaehrer J, Cahan P, et al. Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. Nat Biotechnol 2011;29(12):1117-1119.

(243) Wilson KD, Venkatasubrahmanyam S, Jia F, Sun N, Butte AJ, Wu JC. MicroRNA profiling of human-induced pluripotent stem cells. Stem cells and development 2009;18(5):749-757.

(244) Pauli A, Rinn J, Schier A. Non-coding RNAs as regulators of embryogenesis. Nature reviews.Genetics 2011;12(2):136-149.

(245) Batista P, Chang H. Long noncoding RNAs: cellular address codes in development and disease. Cell 2013;152(6):1298-1307.

(246) Murry CE, Keller G. Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. Cell 2008;132(4):661-680.

(247) Sheik Mohamed J, Gaughwin PM, Lim B, Robson P, Lipovich L. Conserved long noncoding RNAs transcriptionally regulated by Oct4 and Nanog modulate pluripotency in mouse embryonic stem cells. RNA 2010 Feb;16(2):324-337.

(248) Kim K, Zhao R, Doi A, Ng K, Unternaehrer J, Cahan P, et al. Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. Nat Biotechnol 2011;29(12):1117-1119.

(249) Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. Cell 2014;157(6):1262-1278.

(250) Braam SR, Denning C, Van Den Brink S, Kats P, Hochstenbach R, Passier R, et al. Improved genetic manipulation of human embryonic stem cells. Nature methods 2008;5(5):389-392.

(251) Peter IS, Davidson EH. A gene regulatory network controlling the embryonic specification of endoderm. Nature 2011;474(7353):635-639.

(252) Thermo Fisher Scientific Inc. PARIS™ Kit. 2015; Available at: http://www.lifetechnologies.com/order/catalog/product/AM1921. Accessed 03/27, 2015.

(253) Tan LP, Seinen E, Duns G, de Jong D, Sibon OC, Poppema S, et al. A high throughput experimental approach to identify miRNA targets in human cells. Nucleic Acids Res 2009 Nov;37(20):e137.

(254) Ben-David U, Benvenisty N. The tumorigenicity of human embryonic and induced pluripotent stem cells. Nature Reviews Cancer 2011;11(4):268-277.

(255) Cunningham JJ, Ulbright TM, Pera MF, Looijenga LH. Lessons from human teratomas to guide development of safe stem cell therapies. Nat Biotechnol 2012;30(9):849-857.

8. ANNEX

MicroRNA signatures of iPSCs and endoderm-derived tissues

Gene Expression Patterns

Volume 13, Issues 1–2, January–February 2013, Pages 12–20

http://www.sciencedirect.com/science/article/pii/S1567133X12000749