Maria João de Andrade Terras Arez

Licenciatura em Biologia

Comprehensive Evaluation of Genomic Imprinting Stability in Mouse iPSCs

Dissertação para obtenção do Grau de Mestre em
Genética Molecular e Biomedicina

Orientador: Simão Teixeira da Rocha, Investigador auxiliar, Instituto de Medicina Molecular | João Lobo Antunes

Co-orientador: Bruno Bernardes de Jesus, Investigador auxiliar, Instituto de Medicina Molecular | João Lobo Antunes

Júri:

Presidente: Prof. Doutora Margarida Castro Caldas
Arguente: Prof. Doutor Sérgio Fernandes de Almeida
Vogal: Prof. Doutor Simão José Teixeira da Rocha

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Setembro, 2018

LOMBADA

Comprehensive Evaluation of Genomic Imprinting Stability in Mouse iPSCs
Maria João Arez

2018

Maria João de Andrade Terras Arez

Licenciatura em Biologia

Comprehensive Evaluation of Genomic Imprinting Stability in Mouse iPSCs

Dissertação para obtenção do Grau de Mestre em
Genética Molecular e Biomedicina

Orientador: Simão Teixeira da Rocha, Investigador auxiliar, Instituto de Medicina Molecular | João Lobo Antunes

Co-orientador: Bruno Bernardes de Jesus, Investigador auxiliar, Instituto de Medicina Molecular | João Lobo Antunes

Júri:

Presidente: Prof. Doutora Margarida Castro Caldas
Arguente: Prof. Doutor Sérgio Fernandes de Almeida
Vogal: Prof. Doutor Simão José Teixeira da Rocha

Setembro, 2018

Acknowledgements

I would like to acknowledge all the people directly or indirectly involved in this thesis.

First to Simão Teixeira da Rocha, as my supervisor and as a friend. For all guidance, meetings, laughs and mainly for all the patience and neurons I took from him everytime I appeared/showed up in his office for more work discussions! Also, I would like to acknowledge his importance in my carreer maturation as a scientist, there is no thanking enough for you.

Secondly, to Bruno Bernardes de Jesus for challenging me in the first place to enter in this field and common project between two amazing teams, for always sharing his skills with me and for all the help and participation during this project.

To Maria do Carmo Fonseca and all the teachers for harbouring me and for all the critical suggestions during this thesis.

A special acknowledgement to my beautiful colleagues: to Ana Raposo for being this happy, enthusiastic and drama queen at the same time, which facilitated all my rough days in the lab, for giving me so much support and to have the patience to teach me some techniques applied in this project; To Inês Godinho for being the exactly opposite person and for all the “keep calm we can do it” during our master thesis writing. Thank you girls for being my support. Also to Duarte Brandão for teaching me at the beginning. To Teresinha, Rita Drago and Marta Ribeiro for all the support. To Carolina Pereira for being so helpful and for having always a fun story to tell.

A special acknowledgement also to all the boys in the “party room” for destabilizing my sanity while I wrote my thesis: Rui Luís and Pedro Barbosa for being the sweetest and the most annoying men and for all the love and work advices; Kenny Rebelo for eating everything he sees and he does not see; Pedro Prudêncio for all the noises he explored next to me; To Sérgio Marinho for being the funniest person in the lab.

To all the facilities at Instituto de Medicina Molecular for helping me to carry out this project.

To my family for all the unconditional love and support at all times and circumstance, especially to my sister for all discussions, suggestions for my work and for being always there for me.

Lastly, to Gonçalo Cosme for all the days and nights he stayed next to me during this thesis writing, and for all the love and support during this phase.

Resumo

A reprogramação de células somáticas em células estaminais pluripotentes induzidas (iPSCs) possibilitou a geração de células estaminais a partir de células de pacientes para o desenvolvimento de modelos celulares de doenças e para medicina regenerativa. Este processo requer alterações epigenéticas significativas que, no entanto, levam ainda ao aparecimento de erros dessa natureza. Um fenómeno adequado para analisar estes erros epigenéticos é o "*imprinting"* parental, que consiste numa marcação epigenética diferencial por metilação do DNA entre o alelo materno e o paterno, levando à expressão monoalélica de vários genes. Defeitos no *"imprinting*" foram já reportados em iPSCs humanas e murinas, contudo, a falta de um sistema de reprogramação controlado e de uma metodologia eficaz, não permitiu ainda um estudo sistemático da estabilidade do "*imprinting*" em iPSCs.

Neste projecto, foi usado um sistema de reprogramação controlado no qual a célula dadora é reprogramada de maneira indutível e contem polimorfismos a nível do DNA que permitem distinguir qualquer região de "i*mprinting*" entre os dois alelos parentais. Foram analisadas cinco regiões de "*imprinting*" em iPSCs derivadas de fibroblastos embrionários. Para todas as regiões analisadas observou-se uma tendência para defeitos de hipometilação no alelo parental metilado, tendo estes sido predominantes em células de macho. Além disso, observou-se que a hipometilação levou à expressão bialélica dos genes de "*imprinting".* Isto leva à hipotese de que o "*imprinting*" parental não é protegido durante a desmetilação global que ocorre durante a reprogramação. Concomitantemente foi verificada uma fraca expressão de proteínas envolvidas na proteção do "*imprinting*" no início da reprogramação. Com base nos nossos resultados propõe-se que uma sobreexpressão destas proteínas protectoras, como por exemplo a ZFP57, durante a reprogramação poderá ser uma estratégia promissora para corrigir os erros de "*imprinting*". Esta correcção será crucial para garantir a segurança do uso de iPSCs nas suas mais diversas aplicações.

Abstract

Reprogramming of somatic cells into Induced Pluripotent Stem Cells (iPSCs) opened the prospect to generate patient-derived stem cells for disease modelling and regenerative medicine. This process requires massive epigenetic rewiring which, unfortunately, remains error-prone. An important read-out for epigenetic errors in iPSCs is genomic imprinting, an epigenetic phenomenon where genes become monoallelically due to parental-specific epigenetic marking by DNA methylation. Indeed, imprinting defects have been reported in both mouse and human iPSCs. However, a comprehensive study of imprinting stability in iPSCs has never been performed due to the lack of a controlled reprogramming system and an effective methodology. In this project, we address this issue by using a controlled reprogramming system using murine cells whereby the donor cell containing both an integrated inducible reprogramming cassette and DNA polymorphisms able to distinguish any imprinting region between parental alleles. We specifically assess the imprinting stability at five classical imprinted loci in female and male miPSCs derived from mouse embryonic fibroblasts. For all these imprinted clusters, a clear trend for hypomethylation defects was observed, with the parental methylated allele to partially or totally lose its methylation state. This effect was observed in female cells, but was clearer in male cells. Furthermore, we proved that hypomethylation defects in two particular imprinted loci resulted in biallelic expression of imprinted genes. We postulated that methylation imprints are unprotected from the wave of DNA demethylation accompanying the reprogramming process. Indeed, we verified that proteins involved in protecting methylation imprints are weakly expressed at the onset of reprograming. Based on our results, we propose that overexpression of imprinting protective proteins, such as ZFP57, during reprogramming might be a promising way to correct imprinting defects in iPSCs. Correcting imprinting defects is an important benchmark to guarantee the safe use of iPSCs in their many downstream applications.

Keywords: Genomic Imprinting, Reprogramming, mouse iPSCs, DNA methylation, X-chromosome

Contents

Figure Index

Table Index

Abbreviations

AA	Ascorbic Acid
AS	Angelman Syndrome
BSA	Bovine serum albumin
cDNA	Complementary Deoxyribonucleic Acid
COBRA	Combined Bisulfite Restriction Analysis
DAPI	4',6-diamidino-2-phenylindole
DMR	Differentially DNA-Methylated Region
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic acid
DNMT	De Novo Methyltransferases
DOX	Doxycycline
EBs	Embryoid Bodies
ESCs	Embryonic Stem Cells
EtOH	Ethanol
FBS	Fetal Bovine Serum
hESCs	Human Embryonic Stem Cells
ICM	Inner Cell Mass
ICR	Imprinting Control Region
IG-DMR	Intergenic Differentially DNA-Methylated Region
iPSCs	Induced Pluripotent Stem Cells
LB	Lysogeni Broth
LIF	Leukemia inhibitory factor
lncRNA	Long Non-Coding RNA
MEFs	Mouse Embryonic Fibroblasts
mESC	mouse Embryonic Stem Cell
miPSC	mouse Induced Pluripotent Stem Cell
OSKM	Oct4, Sox2, Klf4 and c-Myc Transcription Factors
PB	Piggy Bac
PBS	Phosphate Buffered Saline
PFA	Paraformaldehyde
PGCs	Primordial Germ Cells
PSCs	Pluripotent Stem Cells
PWS	Prader-Willi Syndrome
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
RT	Room Temperature
RT-PCR	Reverse Transcription Polymerase Chain Reaction
RT-qPCR	Real-Time quantitative Polymerase Chain Reaction
SCs	Stem Cells

SNPs	Single Nucleotide Polymorphisms
TET	Ten-Eleven Translocation
UPD	Uniparental disomy
WT	Wilde-Type
w/ v	Weigh per volume
µL	Microliter
µm	Micrometer
2i	2 inhibitor
3D	Three Dimensional

1. Introduction

For many years was thought that the differentiated state of somatic cells was stable and irreversible. However, seminal work by Takashi and Yamanaka in 2006 proved the contrary. These authors found conditions to reprogram somatic cells into the so-called induced pluripotent stem cells (iPSCs). iPSCs technology is a scientific headway discovery with potential significant medical applications, since it allows to reprogram patient-derived cells for personalized cell-based therapy, human disease modelling, and to reveal new therapeutic drugs (Omole & Fakoya, 2018). However, iPSCs technology has still some drawbacks which renders iPSCs prone to genetic and epigenetic defects (Ma et al., 2014).The inability to fully control the epigenome of iPSCs is indeed a major concern for all their downstream applications and, in this thesis, we address this problem using the epigenetic phenomenon of genomic imprinting as a powerful read-out.

1.1. Stem Cells (SCs):

SCs are defined as cells capable of proliferation, self-renewal and are able to generate other differentiated cell types. They can be divided in three fundamental types in terms of cell potency (i) totipotent SCs, which are able to originate all the types of cells present in the organism, including the extra-embryonic tissues and are only found *in vivo* up to the stage of morula; (ii) pluripotent SCs that comprise embryonic stem cells (ESCs), derived from blastocyst or extra-fetal tissues, and iPSCs, both having the ability to propagate indefinitely *in vitro* and originate a wide variety of cell types, except cells from the extraembryonic lineages, under appropriate culture conditions; (iii) multipotent SCs, which are found within a certain tissue or organ and are able to differentiate only into a limited number of specialized cell types to repopulate these tissues or organs. The best studied examples of multipotent SCs are epidermal, intestinal and hematopoietic SCs (Bacakova et al., 2018; Slack, 2018).

1.1.1. Pluripotent SCs:

Pluripotent SCs, as previously stated, have the potential to proliferate indefinitely and differentiate into any cell type of all three germ layers. Besides that, they can be maintained indefinitely *in vitro,* providing an amenable and useful cellular model system, being the cells with a great impact for new therapies and approaches to disease. Pluripotent SCs include ESCs which are obtained by *in vitro* culture of the inner cell mass (ICM) of preimplantation embryos and the iPSCs which are produced by forcing expression of key pluripotent transcription factors in somatic cells.

1.1.2. ESCs:

Mouse ESCs (mESCs) were derived for the first time by Martin and colleagues in 1981, from ICM of late mouse blastocysts. In this seminal study, the authors proved that mESCs could be maintained *in vitro* and moreover differentiate into a wide variety of cell types when appropriately induced (Martin, 1981). It was only in 1998 when the first human ESCs (hESCs) were derived from *in vitro* fertilized embryos at the blastocyst stage (Thomson, 1998).

The discovery of ESCs has revolutionized the field of stem cell biology and allowed many novel scientific discoveries to be made in recent years. Moreover, hESCs have potentiality for clinical applications. However, the use of hESCs always had strong ethical issues associated with embryo destruction, immune rejection and supply limitations (Omole & Fakoya, 2018). Hence, there was a need to find an alternative source of pluripotent SCs with the same differentiation potential as ESCs, and that overcome these ethical concerns and possible limitations.

1.1.3. iPSCs

iPSCs are generated by reprogramming somatic cells through introduction and forced expression of specific pluripotent-associated genes. The first attempts to reprogram somatic cells were by transplantation of a somatic nuclei into enucleated oocytes, a process termed cloning, leading to formation of an embryo able to develop into a full organism (Gurdon, 1962). A different strategy was tried in mammalian cells and consisted in the fusion of somatic cells with ESCs, leading to generation of cells capable to express pluripotency-related genes (Tada et al., 2001). Both experiments indicated that unfertilized eggs and ESCs probably contain factors that confer totipotency or pluripotency to somatic cells (Tada et al.a, 2001). Inspired by these findings, Takahashi and Yamanaka hypothesized that particular transcription factors could be critical for the induction of pluripotency in somatic cells. Thus, in 2006, they tested 24 candidate transcription factors associated with the maintenance of mESCs identity in mouse embryonic fibroblasts (MEFs) to understand whether they could alter the somatic state into an undifferentiated one. From this pool of 24 genes, a set of four genes - *Oct3/4*, *Sox2, Klf4* and *c-Myc* (OSKM) - were shown to be sufficient to convert somatic cells into iPSCs (Takahashi & Yamanaka, 2006). In the original experiment, reprogramming was achieved with MEFs, but it was also proved that adult human fibroblasts could also be reprogrammed into a pluripotent-like state by two groups in 2007 (Takahashi et al., 2007; Yu et al., 2007), and then successfully translated to other somatic cell types as the initial protocols were improved.

iPSCs technology revolutionized disease modelling and regenerative medicine through the possibility of patient's own cells to become the source of therapeutic tissues, thus overcoming the ethical concerns surrounding the use of hESCs. However, this technology is not flawless and many issues concerning the “first iPSCs generation” have been raised, namely their propensity to several genetic and epigenetic defects, which should be solved for the safe use of these cells for therapeutic means (Nordin, Lai, Veerakumarasivam, & Ramasamy, 2011). A good way to decipher when and why these defects occur, is by understanding the intricacies of the reprogramming process.

1.1.3.1. Reprogramming Methods

iPSCs reprogramming relies on the four reprogramming factors *Oct4*, *Sox2*, *Klf4*, and *c-Myc* (OSKM). OCT4 and SOX2 belong to the core pluripotency network of transcription factors in ESCs, while c-MYC, a proto-oncogene factor, helps OCT4/SOX2 bind to their targets and contain also several downstream targets that enhance proliferation and transformation. KLF4, a tumour suppressor factor, is responsible for the inhibition of c-MYC-induced apoptosis and for anti-proliferation which is inhibited by c-MYC. A balance between c-MYC and KLF4 is important for proper generation of iPSCs (Takahashi &

Yamanaka, 2006). Although, high amounts of these four exogenous factors are required in the initial steps of reprogramming (Takahashi & Yamanaka, 2006), persistent factor expression in the pluripotent state may be detrimental, so it is important that the expression be discontinued as soon as genuine iPSCs become apparent (Maherali & Hochedlinger, 2008).

There are several approaches to introduce the reprogramming factors into somatic cells. These can be subdivided in integrative and non-integrative methods and both of them include the use of viral and non-viral vectors (**Table 1.1**). The "first generation" of iPSCs used by Yamanaka and others was achieved by using retrovirus transduction, where the active retrovirus, containing each of the four factors, was integrated into the genome of the target cell to drive dedifferentiation to a stem-like state (Takahashi & Yamanaka, 2006). Interestingly, these integrated retroviral DNA becomes eventually silenced in the pluripotent state after the endogenous stem program was in place (Maherali & Hochedlinger, 2008). Indeed, this is a sign of an iPSC fully reprogrammed. In any case, spontaneous reactivation of retroviral vectors in iPSCs has been documented which is a disadvantage of this system. For example, the reactivation of c-MYC leads to an increased risk for tumorigenesis, thus being inadequate for clinical applications (Okita, Ichisaka, & Yamanaka, 2007).

Soon after the first retroviral-based reprogramming procedures, the use of a constitutive lentiviral vector expressing the four transcription factors from a single polycistronic transcript became popular (Sommer et al., 2009). This approach also results in the integration of the OSKM cassette into the genome of the target cell. However, unlike retrovirus, lentivirus can integrate into both dividing and non-dividing cells, thus allowing iPSCs generation from most somatic cell types. This strategy is also more efficient due to the existence of only one cassette harbouring all the four OSKM transcription factors and a more robust viral infection (Omole & Fakoya, 2018). However, lentiviral vectors are even less efficiently silenced in pluripotent cells than retroviral vectors, which can lead to transgene reactivation and oncogenic behaviour (Brambrink et al., 2008; Omole & Fakoya, 2018; Stadtfeld & Hochedlinger, 2010).

To overcome these issues, lentivirus expressing OSKM in an inducible fashion were then generated. Induction of the OSKM by the tetracycline analog Doxycycline (DOX) allows temporal control over the expression of the Yamanaka factors (Brambrink et al., 2008; Stadtfeld et al., 2008). The drug is later withdrawn once iPSC colonies emerged. Only cells that continue to grow and maintain stem-like characteristics will be true iPSCs. This approach also ensures the controlled silencing of the transcription factors, resulting in iPSCs with no possibility to reactivate the Yamanaka cassette.

To avoid reactivation of the integrated OSKM cassette (even when inducible), reprogramming strategies based on excisable methods have also been developed. One example of that is an excisable vector based on Cre/loxP system, where a polycistronic vector express all transcription factors from a single promoter and then, is deleted upon activation of the Cre recombinase after recognition of the flanking loxP sites (Chang et al., 2009). Another example takes advantage of non-viral single vector system using piggyBac (PB) transposons that can be removed by the transposase enzyme coded by PB (Kaji et al., 2009; Omole & Fakoya, 2018).

To avoid insertion into the host genome, reprogramming methods have evolved to generate integration-free iPSCs. Many strategies have been attempted based on viral vectors like adenovirus and sendai virus or non-viral vectors such as episomal vectors, like plasmids. All these techniques are very

efficient in transferring the reprogramming factors into many types of somatic cells. However, all require multiple viral infections/transfections and the reprogramming efficiency remains very low (Omole & Fakoya, 2018). Lastly, reprogramming factors can also be directly delivered as recombinant proteins or as synthetic mRNA. These DNA-free methods have a low reprogramming efficiency and require repeated transfections in order to sustain the process. From all the non-integrative systems, the RNA delivery is the method with the best reprogramming efficiency (Schlaeger et al., 2015).

Table 1.1 - Different factor delivery methods and their efficiency [adapted from Omole, A. & Fakoya, A. (2018); Stadtfeld, M. & Hochedlinger, K. (2010)].

	Factor Delivery Strategy		Efficiency
Integrative systems	**Viral vectors**	**Retrovirus**	~0.01% - 0.5%
		Lentivirus	~0.1% - 1%
		Inducible lentivirus	
	Non-viral vectors	Linear/plasmid DNA fragments	~0.1%
		Transposons (Piggybac)	
Non-integrative systems	**Viral vectors**	Adenovirus	~0.001%
		Sendai vírus	
	Non-viral vectors	Episomal vectors	~0.001%
		RNAs	~1%
		Proteins	~0.001%

1.1.3.2. Dynamics of the Reprogramming process

The reprogramming process takes at least 2-3 weeks depending on the reprogramming procedure and the original somatic cell. Of all the somatic cells, fibroblasts have been the cell type mostly used for iPSCs production and, therefore, mostly of what is known about the molecular mechanisms undergoing stem-like conversion have been described in this reprogramming setting.

Stadtfeld et al (2008) provide the first in depth study on the dynamics of reprogramming process. In their study, they used newborn mouse fibroblasts transduced with a lentiviral DOX-inducible OSKM cassette. Their first conclusion was that exogenous OKSM expression is required only during the first ten days for the epigenome of the somatic cell to recede into a pluripotent state. After two days of DOX addition, they already verified morphological changes with the first noticeable molecular event to be the downregulation of Thy1 fibroblast marker, followed modest activation of SSEA-1, a membrane marker present in ESCs. At day 10, cells that are destined to become iPSCs activate the endogenous pluripotency program becoming independent of DOX, while unstable reprogramming intermediates and transformed colonies disappear through differentiation or apoptosis. Lastly, only after cells become independent of DOX and, consequently, of exogenous OSKM expression (around day 11), cells reactivate telomerase and, in the case of female cells, they activate the silenced X chromosome (Stadtfeld et al., 2008) (**Figure 1.1**), which are two known features of mESCs.

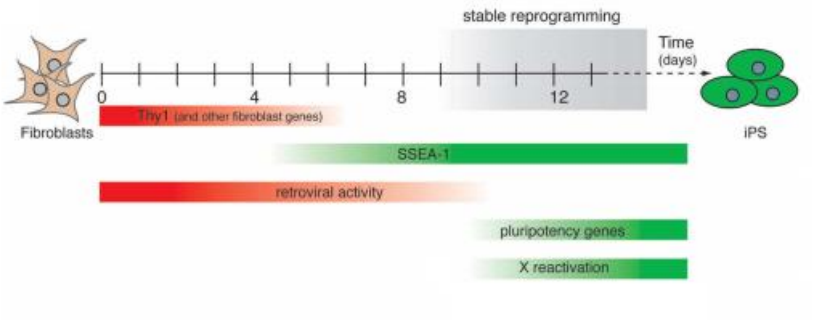

Figure 1.1 - Molecular events during cellular reprogramming (adapted from Stadtfeld, M. et al. (2008)).

1.1.3.3. Characterization of iPSCs Pluripotent Potential

After reprogramming, iPSC lines need to be submitted to several criteria to confirm whether they reached a fully reprogrammed state. This includes morphological characteristics, molecular signature associated with stemness and functional features linked to pluripotency. By morphological criteria, iPSCs must be similar to ESCs and demonstrate the same unlimited self-renewal capacity. In that respect, mouse iPSCs (miPSCs), like mESCs, should have a ‘‘shiny’’ appearance with well-defined borders, while human iPSCs (hiPSCs) should have prominent nucleoli and pronounced individual cell borders identical to hESC. However, non-iPSC colonies also arise during reprogramming and cannot be easily distinguished by their morphology from true iPSCs (Maherali & Hochedlinger, 2008). Therefore, morphological criteria on its own is insufficient.

At the molecular level, iPSCs must have a similar gene expression profile to ESCs. Evaluation of expression of stemness-associated genes such as *Nanog*, *Oct4*, *Sox2* and of surface markers such as SSEA-1, in miPSCs, and SSEA-3/4, in hiPSCs, using immunostaining is a common procedure to characterize iPSCs. A battery of these and other stemness-related genes could and should also be performed by RT-qPCR or RNA-sequencing (RNA-seq). Besides that, it is also important that iPSCs show to have DNA demethylation at the promoters of pluripotency genes and X chromosome reactivation in the case of female miPSCs, which could be examined by DNA methylation analysis (*e.g.,* bisulfite sequencing) (Jaenisch & Young, 2008; Maherali & Hochedlinger, 2008).

At a functional level, iPSCs must demonstrate the ability to differentiate into lineages from all three germ layers. Several pluripotency tests can be used to evaluate this fact (**Table 1.2**). The easiest test to perform is the embryoid body formation *in vitro*: iPSCs are grown in suspension in non-stem conditions to form three-dimensional (3D) aggregates called embryoid bodies (EBs). These structures can then be analysed by immunohistochemistry and expression analysis for differentiation-specific markers (Jaenisch & Young, 2008; Maherali & Hochedlinger, 2008). Bona-fide iPSCs will be able to differentiate into cells of the three embryonic layers - ectoderm, mesoderm and endoderm.

Besides this *in vitro* method, there are also *in vivo* pluripotency tests such as the teratoma formation, which is the most stringent assay that can be used for hiPSCs (Maherali & Hochedlinger, 2008). Teratomas are tumours formed by multiple differentiated somatic tissue types derived from all three germ layers, resembling structures identified in the embryo and adult organism. This test consists in injection of iPSCs into an immune-deficient host, usually mice, that results in the formation of a tumour possessing cells of ectodermal, mesodermal and endodermal lineages that grow for 6 - 8 weeks. During

this period, it is checked a tumour mass growth in mice. The mass is then explanted upon mice's sacrifice and examined by histological and immunohistochemical analysis for identification of tissues from the three germ layers (Przyborski, 2005).

For miPSCs, other *in vivo* pluripotency tests can be used, such as mouse chimaera formation to test the competence of iPSCs to contribute to all lineages. iPSCs are transplanted in one of two murine pre implantation embryonic stages - cleavage-stage or blastocysts – and if iPSCs developed into tissues of different germ layers on the chimeric mice, confirms that transplanted iPSCs are indeed pluripotent. Furthermore, if the chimaera can give rise to fertile offspring totally constituted by donor iPSCs, the cells can be capable of germline transmission and, therefore, might not possess significant genomic defects (Maherali & Hochedlinger, 2008; Mascetti & Pedersen, 2016).

There is also an even more stringent pluripotency test for miPSCs known as tetraploid complementation. Here, miPSCs (diploid cells, 2N) are transplanted in a tetraploid blastocyst (4N). The tetraploid cells are only able to originate extraembryonic lineages like trophectoderm, while the fetus is exclusively originated from the diploid miPSCs, which means that only the true pluripotent SCs will be able to support complete fetal development (Mascetti & Pedersen, 2016).

All methods described above may be done at different time points after reprogramming and this differ between miPSCs and hiPSCs (see **Figure 1.2** and **Table 1.2**). As soon as these tests are done the better, to obtain a faster detection of true iPSCs.

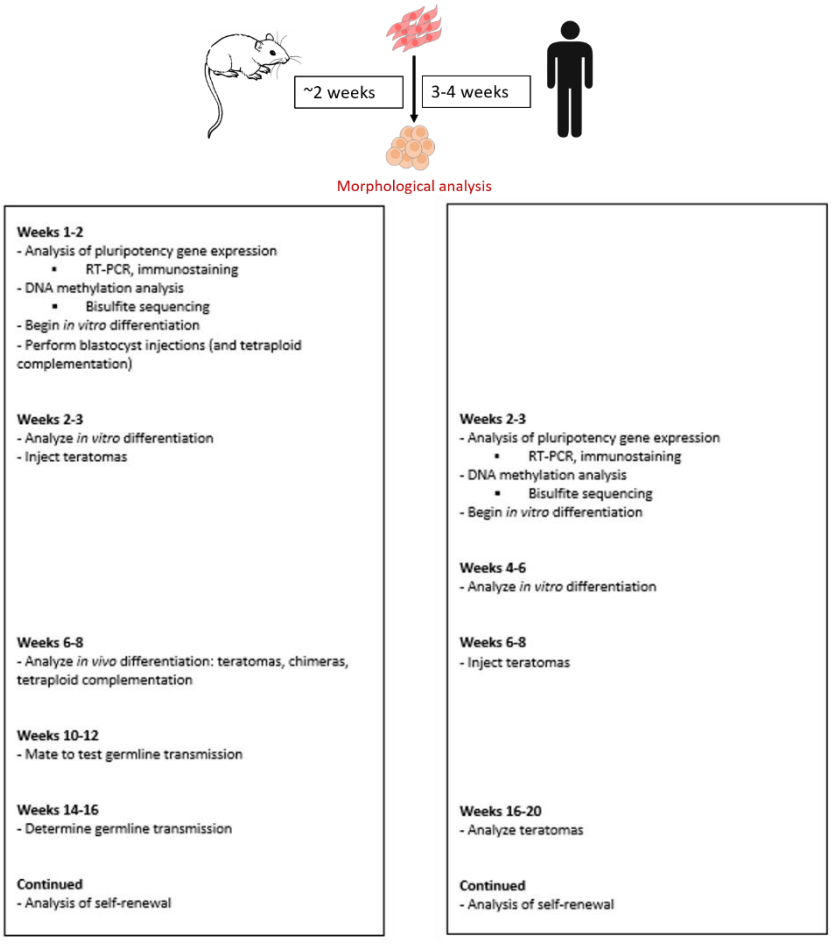

Figure 1.2 - Time points of different characterization methods of novel hiPSCs and miPSCs (adapted from Maherali, N. & Hochedlinger, K. (2008)).

Table 1.2 - Functional pluripotency tests (adapted from Jaenisch, R. & Young, R. (2008)).

Assay	Experimental approach	miPSCs	hiPSCs
Embryoid body formation	**Differentiation induced of suspended cells which are assayed for the expression of cell type specific markers**	**Applicable**	**Applicable**
Teratoma formation	Induction of tumours in immune-compromised mice to demonstrate the potential to generate differentiated cell types of various lineages	Applicable	Applicable
Chimera formation	Contribution of cells to normal development following injection into host blastocyst	Applicable	Not Applicable
Tetraploid complementation	Injection of test cells into 4n host blastocyst.	Applicable	Not Applicable

1.2. Epigenetics and iPSC Reprogramming

Epigenetics is a field of research focused on changes in gene expression independent of changes in the DNA sequence, being important for the regulation of several processes such as cell fate determination, differentiation and aging (Godini, Lafta, & Fallahi, 2018). There are several epigenetic mechanisms, such as DNA methylation, histone post-translational modifications and chromatin remodelling, which affects gene function (Godini et al., 2018). From those, DNA methylation is one of the best studied epigenetic marks, being critical for genomic imprinting, which is the main subject of this study.

Reprogramming of somatic cells into iPSCs requires major epigenetic rewiring. Indeed, a somatic cell must erase its differentiated and aged epigenetic signature in order to adopt a stem cell-like epigenome. Some of those epigenetic changes are well documented such as DNA demethylation of promoter regions of pluripotency genes and concomitant activation (*e.g.*, *Oct4, Nanog*) or the reactivation of the inactive X-chromosome in female miPSCs (Omole & Fakoya, 2018). However, epigenetic memory of the donor somatic cell and its age are not always completely erased, which is one of the big differences between iPSCs and ESCs and that lead to low quality iPSCs (Godini et al., 2018; Lo Sardo et al., 2016). Additionally, since iPSC generation requires forced expression of the Yamanaka transcription factors in a somatic background, it can lead to errors in sensible epigenetically regulated loci. An example of this type of loci are imprinted regions, characterized by parental-specific methylation marks, which drives monoallelic expression of a subset of genes. In fact, imprinted defects have been detected in both mouse and human iPSCs (Godini et al., 2018), as discussed below. However, imprinted defects remains poorly documented and studied. Since these genes are involved in many fundamental developmental processes, assessment of the number and nature of imprinted errors in iPSCs remains an important issue to report and solve.

1.2.1. DNA Methylation

DNA methylation is a major epigenetic mechanism that regulates crucial aspects of genome function, having a significant impact on its expression and stability (Reik, 2001). DNA methylation is required for fundamental physiological processes, such as gene regulation, embryonic development, developmental potential of SCs, X chromosome inactivation, suppression of transposable elements, and genomic imprinting (Heo et al., 2017). In mammals, DNA methylation occurs predominantly at the 5-methylcytosine (5mC) of CpG dinucleotides in repetitive regions and at certain promoters of silenced genes, by the action of methyltransferase enzymes (E. Li & Zhang, 2014). The first enzyme described was the DNA methyltransferase 1 (DNMT1) (E Li, Bestor, & Jaenisch, 1992). Its preferred substrate is the hemymethylated DNA (DNA methylated at CpG on one of the two DNA strands) and its principal function is to maintain methylation through cell divisions (Reik, 2001). Later, two other DNA methyltransferases, DNMT3A and DNMT3B, were shown to establish new methylation patterns in early development and they are known as *de novo* methyltransferases (Okano, Bell, Haber, & Li, 1999). In addition, DNMT3A associates with a co-factor, known as DNMT3-like (DNMT3L) which is catalytically inactive methyltransferase. DNMT3L has been shown to be necessary for the establishment of distinct DNA methylation patterns, namely at imprinted genes in the germline (Hata, Okano, Lei, & Li, 2002).

Besides the DNA methylation machinery, there is also a machinery that demethylate the DNA. This consists in a subset of ten-eleven translocation (TET) proteins, namely TET1, TET2 and TET3, that in association with its co-factors, 2-oxoglutarate and Fe2+, oxidizes the 5mC into 5- hydroxymethylcytosine (5hmC) (Pastor, Aravind, & Rao, 2013). This is the first step of a cascade leading to a completely loss of CpG methylation.

In the mammalian life cycle, the genome undergoes dynamic changes in DNA methylation during early development (**Figure 1.3**). Briefly, immediately after fertilization occurs a genome-wide demethylation wave, where the paternal pronucleus (represented in blue) is subjected to rapid demethylation in the zygote, while the maternal genome (in red) suffers a passive loss of DNA methylation during the subsequent cell divisions. Then, *de novo* DNA methylation occurs around the time of implantation from the ICM stage of the developing embryo, with specific patterns in each cell lineage. Another wave of genome-wide demethylation occurs during the establishment of the primordial germ cells (PGCs), which are the direct progenitors of sperm and oocytes. Lastly, following the demethylation in early PGCs, the genome undergoes again *de novo* methylation to achieve levels of methylation in mature gametes, which interestingly, differ in male and female germ cells, for example, at the regions regulated by genomic imprinting (Seisenberger et al., 2012).

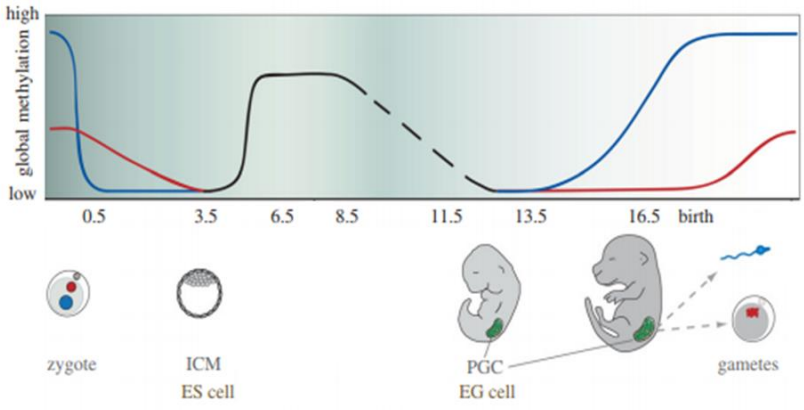

Figure 1.3 - DNA methylation levels during the mammalian life cycle (adapted from Seisenberger, S. et al., (2012)).

1.2.2. Genomic Imprinting

Genomic imprinting is an epigenetic process of gene regulation causing a subset of genes to be monoallelically expressed according to their parental origin (Pólvora Brandão & da Rocha, 2018)(S.̃o T. da Rocha & Heard, 2001). This phenomenon occurs in placental mammals, marsupials and also in some plants (Elhamamsy, 2017).

Mammalian cells carry two matched set of chromosomes, one inherited from the mother and other from the father. Therefore, humans have two copies of every gene (except for most of the genes on X and Y chromosomes in males) and both copies can be expressed. However, a subset of genes, which we call imprinted genes, defy this common rule being always silenced in one of the parental copies and expressed only from the other. Incorrect dosage of imprinted genes are associated with defects in metabolism, brain function, postnatal growth and development leading to several human disorders (Pólvora Brandão & da Rocha, 2018).

The first hint about parental-specific effects on gene regulation was given by the discovery that female marsupial cells always inactivate the paternal X-chromosome (Cooper, VandeBerg, Sharman, & Poole, 1971). The same observation was then made in the extraembryonic tissues of the mouse (Takagi & Sasaki, 1975) and this event was described as “chromosomal imprinting” by Crouse *et al.* (Crouse, Brown, & Mumford, 1971). Later, experiments with mouse carrying chromosomal defects, specifically duplications of one parental chromosome in the absence of the other parental chromosome (known as uniparental disomies), suggested that “imprinting” effects could be extended to autosomes (Barlow & Bartolomei, 2014; Searle & Beechey, 1978).

But it was only in 1984 that Barton *et al.* (Barton, Surani, & Norris, 1984) and McGrath & Solter (McGrath & Solter, 1984) demonstrated that both paternal and maternal genomes were necessary to complete development. Indeed, in these two parallel seminal studies, the authors tried to develop embryos harbouring two male or female pronuclei, without success. These studies suggested that the two parental genomes are not equivalent and that they must have suffered some epigenetic modifications during gametogenesis when they are apart (Barton et al., 1984).

In 1990s, through the use of novel technologies in mouse genetics, the first imprinted genes, *Igf2r* (Barlow, Stöger, Herrmann, Saito, & Schweifer, 1991), *Igf2* (DeChiara, Robertson, & Efstratiadis, 1991)

and *H19* (Bartolomei, Zemel, & Tilghman, 1991), were discovered in the mouse genome. In addition, the discovery that *Igf2* and *H19* were in the vicinity of each other in the genome made an important prediction that imprinted genes could cluster in the genome and their imprinting to be co-regulated (Barlow & Bartolomei, 2014). Today, it is confirmed the existence of 100 and 124 imprinted genes, in human and mouse genomes respectively, however this number might not be final (http://www.geneimprint.com/site/genes-by-species).

1.2.2.1. Genetic and epigenetic features of imprinted clusters

As it was suggested since the discovery of the first imprinted genes, the majority of them, indeed, localize at the same genomic region. These regions are known as imprinted clusters, and there are around 20-25 of them in the mammalian genome. In these clusters, we find both maternally and paternally expressed imprinted genes (Ideraabdullah, Vigneau, & Bartolomei, 2008). It is also common the presence of, at least, one long non coding RNA (lncRNA) that usually is expressed from the opposite parental chromosome to their protein-coding genes (Kanduri, 2016). Additionally, non-imprinted genes can also be present. This organization in clusters suggests that a regulatory element controls the imprinting expression of multiple genes in *cis* (Barlow & Bartolomei, 2014). Indeed, each cluster contains a *cis*-acting element that regulates imprinting, termed Imprinting Control Region (ICR) (Edwards & Ferguson-Smith, 2007). This region is normally rich in CpG dinucleotides and is submitted to a differential DNA methylation between the two parental alleles during gametogenesis (Edwards & Ferguson-Smith, 2007). For this reason, this element is also called germline differential methylated region (gDMR). Indeed, the ICR carries the "imprint" established in the germline asymmetrically and is inherited by the embryo and maintained throughout life. A deletion in the ICR leads to a loss of imprinting expression, which confirms the importance of the presence of this control element (Barlow & Bartolomei, 2014). There are also other sequences that are differentially methylated called somatic DMRs (sDMRs) which can be promoters, silencers or tissue-specific regulatory elements that will affect imprinted expression of specific genes within a cluster (**Figure 1.4**). However, they depend hierarchically on the methylation status of ICRs (Pólvora Brandão & da Rocha, 2018).

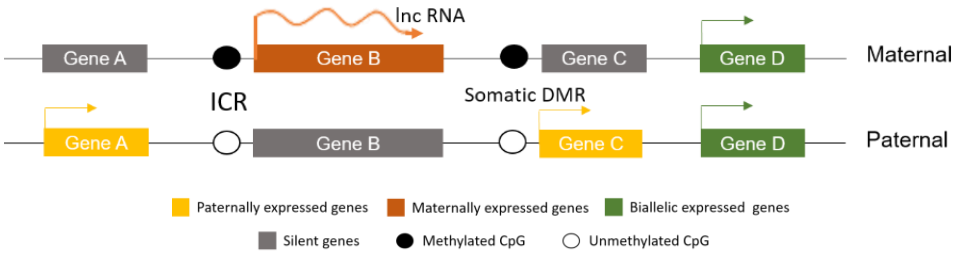

Figure 1.4 - Schematic representation of typical imprinted cluster.

As mentioned before, many imprinted clusters have an imprinted lncRNA which is thought to play a role in the regulation of imprinting of the neighbouring genes in *cis* (Kalish, Jiang, & Bartolomei, 2014). A good example of this can be seen in the *Kcnq1-Kcnq1ot1* cluster on mouse chromosome 7 (**Figure 1.5**). Briefly, this cluster contains the *Kcnq1ot1* lncRNA that is the only gene expressed from the paternal

allele. The promoter of *Kcnq1ot1* is an ICR (known as *Kv*DMR), which is unmethylated in the paternal allele, allowing *Kcnq1ot1* expression, while is methylated on the maternal allele silencing this gene. It is believed that *Kcnq1ot1* establishes a nuclear compartment, rich in repressive histone marks, in the entire cluster on the paternal chromosome, leading to the silencing of the neighbouring protein-coding genes, which remained expressed only from the maternal allele (**Figure 1.5**) (Choufani, Shuman, & Weksberg, 2010; Green et al., 2007).

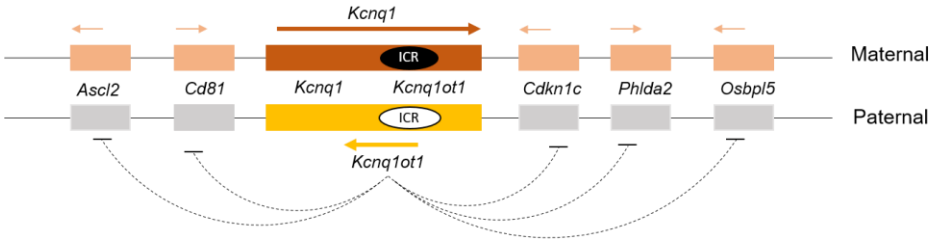

Figure 1.5 - Regulation of imprinting at the *Kcnq1-Kcnq1ot1* cluster on mouse chromosome 7.

1.2.2.2. Imprinting life-cycle

Imprints need to be established, maintained and erased at crucial developmental stages to guarantee proper inheritance of imprinting states.

DNA methylation imprints are initially established later in the germline, when both parental genomes are separated. Most of the ICRs are methylated during oogenesis and not during spermatogenesis, while only a few are specifically methylated during spermatogenesis. After fertilization, methylation imprints are resilient and survive the pre-implantation global DNA demethylation wave (discussed above) (**Figure 1.3**). Likewise, the unmethylated allele remains protected from regaining methylation during the re-methylation wave after implantation. Imprints are then maintained in all the tissues throughout life. However, during the genome-wide demethylation occurring in PGCs, imprints are erased so that they can be established *de novo* during gametogenesis in a differential way (Barlow & Bartolomei, 2014) (See **Figure 1.6**).

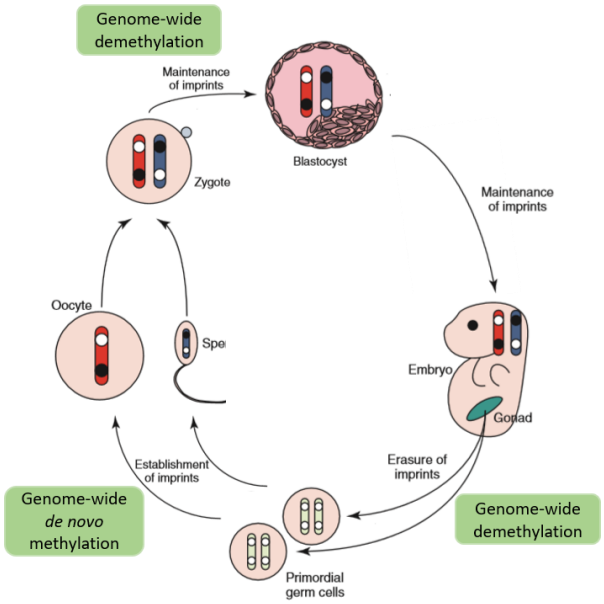

Figure 1.6 - Imprinting life cycle (adapted from Brandão & da Rocha, 2018).

An interesting aspect of the imprinting cycle is to know how methylation imprints are maintained during the massive genome-wide demethylation wave during pre-implantation. It is known that resistance of methylation imprints depends on the presence, at low levels, of the maintenance methyltransferase DNMT1 (En Li, Beard, & Jaenisch, 1993). However, on its own, low levels of DNMT1 cannot ensure protection from genome-wide DNA methylation. Indeed, there are some proteins found to protect imprints at this developmental stage: the primordial germ cell 7 protein (PGC7) also named STELLA/DPPA3, the KRAB zinc finger protein ZFP57 and its cofactor KRAB-ZFP-interacting protein KAP1 also termed TRIM28 (Pólvora Brandão & da Rocha, 2018) (**Figure 1.7**).

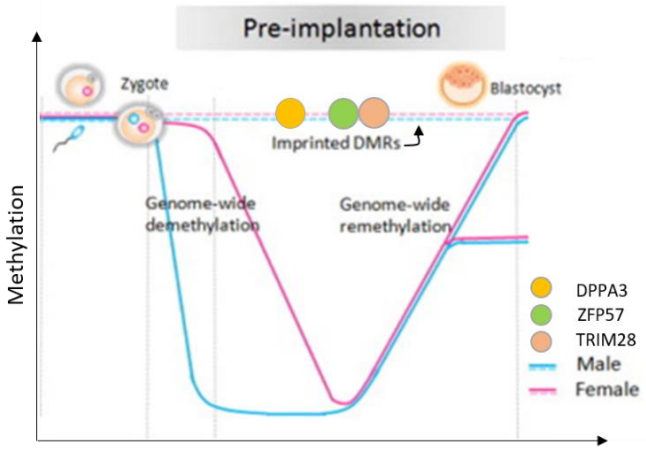

Figure 1.7 - Protection of methylation imprints during genome-wide demethylation at pre-implantation stage of mice development (adapted from Ishida, M. & Moore, G. (2012)).

The PGC7/DPPA3 protein works by protecting methylated DNA and inhibiting the activity of TET3, an enzyme involved in DNA demethylation, through the binding to chromatin regions containing the H3K9me2 histone mark. This effect might not be, however, specific to methylation imprints (Nakamura et al., 2012)

KRAB zinc finger ZFP57 has a critical role in the maintenance of both paternal and maternal methylation imprints after fertilization at multiple imprinted gDMRs in mice (X. Li et al., 2008), but also in humans (Mackay et al., 2008). ZFP57 binds directly to the DNA by a methylated TGCCGCN motif and functions in association with its interacting cofactor KAP1/TRIM28 through a binding of KAP1 to a conversed functional KRAB box that ZFP57 contains. After this interaction, ZFP57 is able to recruit the H3K9me3-catalyzing histone methyltransferase SETDB1, the nucleosome remodelling and histone deacetylation (NuRD) complex, the DNA methyltransferases DNMT1, DNMT3A and DNMT3B, as well as heterochromatin protein 1 (HP1) in order to protect methylated ICRs from demethylating (Luo et al., 2017; Strogantsev et al., 2015). ZFP57 has a remarkable specificity for imprinting regions, but also binds to a few other genomic regions (Strogantsev et al., 2015).

Since reprogramming of somatic cells into iPSCs is associated with changes in the DNA methylation status along the genome (Milagre et al., 2017; Pasque et al., 2018), which might affect imprinted regions, we should bear in mind whether these protecting proteins are expressed during this process.

1.2.2.3. Imprinted genes and Human diseases

Disruption in expression of imprinted genes leads to several disorders, mostly associated with growth and developmental defects, neurological abnormalities and metabolic and hormonal dysfunctions. Besides these, imprinted genes have also been related to cell transformation and cancer (Plasschaert & Bartolomei, 2014). For example, humans with imprinting defects in *DLK1- DIO3* locus suffer skeletal malformations, developmental delay/mental retardation, tumour development, and even postnatal death (Ogata & Kagami, 2016).

These disorders can affect both males and females, but the defect pass to the offspring in a parent-of-origin-specific manner. Imprinted disorders can occur due to mutations in imprinted genes or regions, methylation defects at ICRs or other regulatory regions, and uniparental disomies (UPD) (Kalish et al., 2014). Independent disorders can also be associated to the same genomic region, as in the case of Angelman Syndrome (AS) and Prader-Willi Syndrome (PWS), both caused by defects on maternal and paternal 15q11-q13 region, respectively. Independent disorders can also have opposite phenotypes as in Silver-Russel Syndrome, characterized by growth retardation, and Beckwith-Weiedemann Syndrome, associated with tissue overgrowth (Pólvora Brandão & da Rocha, 2018). **Table 1.3** describe some examples of imprinted regions and their respective human genetic disorders.

Table 1.3 - Human genetic diseases caused by imprinting defects (adapted from Plasschaert, R. & Bartolomei, M. (2014)).

Imprinted genes	Disorder	Clinical Presentation
H19-IGF2 *CDKN1C*	Beckwith–Wiedemann Syndrome	Overgrowth, increased risk for embryonic tumours
H19-IGF2 *GRB10* *PEG1* *PEG3*	Silver–Russell Syndrome	Undergrowth and asymmetry
SNRPN-UBE3A	Prader–Willi Syndrome	Neonatal feeding difficulty, hypothalamic dysfunction, intellectual delay, obesity
	Angelman Syndrome	Developmental delay, speech impairment, poor motor control, seizures
DLK1-DIO3	Kagami-Ogata Syndrome	Birth weight, facial abnormalities, abdominal wall defects, developmental delay
	Temple Syndrome	Postnatal growth retardation, premature puberty, truncal obesity, small hands and short stature

1.2.2.4. Imprinting Defects in Pluripotency and Reprogramming

Despite many successful studies using ESCs and iPSCs, we still lack the ability to fully control their epigenetic state. Genomic imprinting, which is very sensitive to fluctuations in the DNA methylation machinery, provides an extremely good read-out for the epigenetic stability of stem cells in culture and upon reprogramming. Indeed, ESCs are known to have imprinting defects, which manifest as changes

of the normal differential methylation at ICRs due to either hypermethylation or hypomethylation with consequences on the imprinting expression of the genes. As mentioned (see Imprinting life-cycle) mESCs *in vivo* retain their methylation imprints intact. *In vitro*, the same does not happen. Although the maintenance in classical serum conditions seems not to cause global imprinting instability, mESCs cultured in LIF/2i naïve conditions have a clear trend for hypomehtylation at ICRs, which accompanies the global reduction in DNA methylation when extensively cultured in these conditions (Greenberg & Bourc'his, 2015). Concerning imprinting stability, serum-conditions are better than 2i conditions (Pólvora Brandão & da Rocha, 2018).

iPSCs are reprogrammed from somatic cells which have correct imprinting status. However, during the reprogramming of a somatic cell into an iPSC a drastic rewiring of all somatic epigenome is required. In fact, both female and male cells undergo hypomethylation of the genome to remove the epigenetic memory (Milagre et al., 2017). However, the DNA methylation dynamic during reprogramming differs between female and male cells with the female cells suffering a more pronounced global demethylation than male cells (**Figure 1.8**). This seems to be associated with the reactivation of the inactive X-chromosome during reprogramming of female cells at late stages of the process (Milagre et al., 2017; Pasque et al., 2018). This might be, at least partially, explained by the upregulation of the X-linked *Dusp9* gene previously associated with hypomethylation of female mESCs (Choi et al., 2017a).

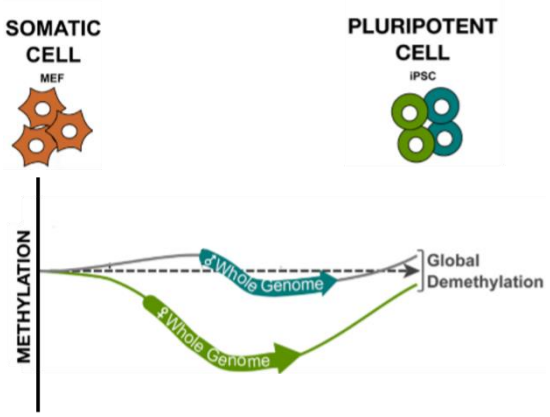

Figure 1.8 -Demethylation events during female and male iPSCs reprogramming (adapted from Milagre, I. et al. (2017)).

How do imprinting status behave during reprogramming and whether they differ between female and male cells? It will be important that stable genomic imprinting resists reprogramming in order to obtain non defective iPSCs that will be useful for cell therapy. Several reports have pointed for the presence of imprinting defects in both miPSCs and hiPSCs. Actually, the frequency of imprinting errors in iPSCs seems higher than in ESCs in classical medium conditions (Ma et al., 2014; Sun et al., 2012). Interestingly, the number and nature of the defects seems to vary among iPSCs lines (Greenberg & Bourc'his, 2015).

The first study highlighting the existence of imprinting defects was from Stadtfeld, M. et al. (2010) which showed a propensity for the *Dlk1-Dio3* locus to gain abnormal methylation on the maternal allele (Stadtfeld et al., 2010). This hypermethylation was associated with a decrease in the pluripotent

properties of miPSCs, which highlighted the important role for the correct dosage of genes of the *Dlk1-Dio3* locus in the pluripotency of miPSCs. Carey et al. (2011) suggested that this hypermethylation was caused by different expression levels of the Yamanaka transcription factors used in iPSCs derivation (Carey et al., 2011). This, highlights for the importance of dosage of reprogramming factors in variations of DNA methylation at the *Dlk1-Dio3* region. Several strategies have been attempted to reduce imprinting defects at this locus, such as the supplement of reprogramming cultures with ascorbic acid (AA) (Stadtfeld et al., 2012) and the expression of *Dppa3* gene (Xu et al., 2015). AA (also known as vitamin C), has been shown to influence DNA methylation patterns in hESCs and, therefore the effect on the preservation of normal imprinting status at the *Dlk1-Dio3* cluster during miPSCs derivation, (Stadtfeld et al., 2012) might not be specific. *Dppa3,* a gene known to protect imprints during the wave of DNA demethylation (see imprinting life cycle), usually emerges early during somatic cell reprogramming, enhancing reprogramming kinetics and also defeats the binding of the DNA methyltransferase DNMT3A enzyme to the ICR region of the *Dlk1-Dio3* cluster (Xu et al., 2015). Of note, similar hypermethylation defects in this locus were monitored in hiPSCs (Ma et al., 2014; Nazor et al., 2012).

Apart from the *Dlk1-Dio3* locus, imprinted errors have also been reported, although not systematically, in other imprinted loci in both miPSCs (Sun et al., 2012; Takikawa et al., 2013) and hiPSCs (Ma et al., 2014; Nazor et al., 2012), with a degree of variability between iPSC lines. A difficulty to address imprinting defects in iPSCs is that for human cells and in many mouse studies is not possible to distinguish the two parental alleles, so imprinting errors might exist, but cannot be easily noticed. In order to address this question, two previous studies using miPSCs from hybrid strains had a look at some imprinted locus based on the ability to distinguish parental alleles through the existence of Single Nucleotide Polymorphisms (SNPs) (Sun et al., 2012; Takikawa et al., 2013). Sun B, et al. (2012) used a reciprocal cross between C57BL/6J and *Mus Musculus Castaneus* strains and VSV-pseudotype retroviruses to reprogramming MEFs (Sun et al., 2012). Takikawa, S et al. used a cross between a female DBA/2 mouse and a transgenic mouse harbouring a DOX-inducible Yamanaka cassette inserted on its genome to generate MEFs and reprogram them simply by addition of DOX (Takikawa et al., 2013). The first interesting conclusion was that, in contrast to previously reported (Stadtfeld et al., 2010), both authors never observed hypermethylation defects at the *Dlk1-Dio3* locus. Instead, they rather verified either stable maintenance or a trend for hypomethylation at this imprinted locus. The other common conclusion was that, except for the *Peg1* locus, a high frequency of hypomethylation defects was found (**Table 1.4**). For some loci (*e.g., Peg10*, *Rasgrf1* and *Peg3*) hypomethylation defects were found in almost all iPSCs. However, for other loci (*e.g.*, *Dlk1-Dio3*, *Igf2-H19*, *Kcnq1-Kcnq1ot1*, *Snrpn* and *Zac1*) there were hypomethylation defects in some, but not all miPSCs lines. From these studies, it is possible to affirm that during reprogramming process the imprinted regions become very sensitive to demethylation, thus resulting in a high frequency of hypomethylation defects.

Table 1.4 - Imprinting defects in miPSCs. n.d – not done

Loci	Imprinting defects in miPSCs	
	Takikawa, S. et al., (2013)	**Sun, B. et al., (2012)**
Dlk1-Dio3	Stable/Hypomethylation	Stable/Hypomethylation
Igf2-H19	Stable/Hypomethylation	Hypomethylation
Kcnq1-Kcnq1ot1	n.d	Variable levels of methylation (trend to hypomethylation)
Peg1	Hypermethylation	n.d
Peg3	Hypomethylation	Hypomethylation
Peg10	Hypomethylation	n.d
Rasgrf1	Hypomethylation	n.d
Snrpn	Variable levels of methylation (trend to hypomethylation)	Stable/Hypomethylation
Zac1	Variable levels of methylation (trend to hypomethylation)	n.d

In conclusion, imprinting defects can occur and are frequent in iPSCs. The origin of such abnormalities and why they vary from line to line remain to be elucidated. The case of *Dlk1-Dio3 locus* illustrates how imprinting defects can have a direct impact on iPSCs pluripotency. Therefore, it is important to study imprinting status during reprogramming, in order to redesign protocols or target epigenetic actors that can avoid defects in methylation and genomic imprinting for iPSCs to become secure and useful in disease modelling and cell-replacing regenerative therapies.

1.2.3. Aging, reprogramming efficiency and genomic imprinting

The majority of mouse reprogramming studies have been based on the use of embryonic somatic cells, especially MEFs. These cells were the first cell type to be reprogrammed and has been heavily used since. However, the main objective of the iPSCs technology is to produce adult patient-specific iPSCs. Since aging is associated with an increase of a variety of diseases such as cancer, cardiovascular dysfunction, metabolic disorders, and neurodegeneration (Mahmoudi & Brunet, 2012), it is important that somatic cells of different ages could be efficiently reprogrammed for the final iPSCs to be useful for cell therapy in all patients, included the oldest ones.

Studies had already shown that the nature of donor cell has an impact in the reprogramming efficiency and in iPSCs properties (Watanabe et al., 2011). It was also reported, by several studies, that aged cells have a decline in reprogramming efficiency in murine (**Figure 1.9**) (Bernardes de Jesus et al., 2018) and human cells (H. Li et al., 2009), presumably due to the fact that old tissues store senescent and genetically unsteady cells, which are resistant to reprograming (reviewed by Mahmoudi, S. and Brunet, A., 2012). In addition, old donor cells have been found to be resistant to normal DNA

demethylation during reprogramming into hiPSCs, which can lead to the appearance of stochastic epigenetic errors during the process, which could impact on imprinting (Lo Sardo et al., 2016).

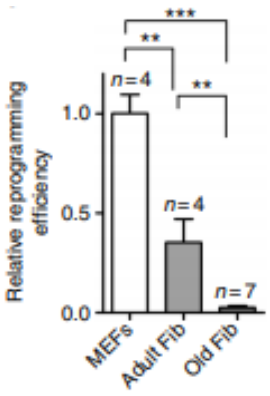

Figure 1.9 - Inefficient reprograming efficiency of old fibroblasts (Bernardes de Jesus, B. et al. (2018)).

Old cells are senescent and have an epigenetic signature of aging based on CpG DNA methylation (Hannum et al., 2013), which is believed to be erased during reprogramming. However certain epigenetic characteristics might not be completely gone (Lo Sardo et al., 2016). Besides this, aged cells have a slow demethylation kinetics during reprogramming. Because elderly people are more prone to develop degenerative diseases, being thus the primary beneficiaries of personalized regenerative therapies, a correction of the genetic and epigenetic alterations during aged-derived-iPSCs is crucial for their future applications.

1.3. Aims of the project

Reprogramming of somatic cells into iPSCs requires a transition from a differentiated to a pluripotency state. Such transition involves major and complex epigenetic rewiring, which can lead to an accumulation of epigenetic errors. It is believed that this effect can be exacerbated during iPSC reprogramming from aged donor cells, which is normally a longer and less efficient process (Bernardes de Jesus et al., 2018; Lo Sardo et al., 2016). A good read-out for epigenetic defects during reprogramming is genomic imprinting. Ideally, iPSCs should maintain the methylation pattern inherited from the original donor cells, but it is known to be very susceptible to defects in both miPSCs and hiPSCs (Sun et al., 2012; Takikawa et al., 2013), impairing their pluripotency potential (Stadtfeld et al., 2010). Yet, the extent, nature and causes of imprinting defects during iPSCs reprogramming remain to be systematically characterized due to the lack of adequate reprogramming systems and technology allowing the simultaneous analysis of all imprinted regions. Furthermore, the impact of the variable age of donor cells on imprinting defects remains to be elucidated. This knowledge is necessary to develop and improve new reprogramming strategies to target potential epigenetic actors that can avoid defects in DNA methylation of imprinted regions in order to guarantee the epigenetic safety of iPSCs in disease modelling and regenerative medicine.

Therefore, the motivation underlying this project is, on one hand, to understand the full extent of imprinted defects in iPSCs, and, on the other hand, to understand the impact of the donor cell age on these defects. For that, we thought of establishing a controlled reprogramming system using donor cells with a genetic setup for which both parental alleles could be distinguished. Basically, the donor cells were obtained from reciprocal crossings between a reprogrammable mouse, containing a DOX-inducible cassette n its genome, and a phylogenetically distant mouse strain. Thus, the reprogramming process can be tightly controlled by DOX-mediated reprogramming with the parental alleles to be distinguished due to the existence of SNPs in both imprinted genes and ICRs for all imprinted clusters. Also, this system based on mice crossings allows the generation of isogenic miPSCs from donor cells with different ages with exactly the same genetic background, allowing to isolate epigenetic effects from genetic ones. All in all, we believe that using this system it will be possible to report the occurrence of imprinting errors in miPSCs, differentiate these defects from locus to locus, line to line and between female and male cells and to understand their consequences in allelic gene expression. To summarize our specific aims are:

1- Generation and characterization of miPSCs from mouse embryonic and adult fibroblasts.

2- Assessment of methylation status at ICRs of imprinted regions in miPSCs.

3- Evaluation of the consequences of methylation imprinting defects in miPSCs on allelic-specific expression of imprinted genes.

Understanding the causes and effects of imprinted defects in iPSCs to a full extent will allow to improve current reprogramming protocols to overcome these defects, which are essential for the safe use of iPSCs in disease modelling and personalized regenerative medicine.

2. Materials and methods

2.1. Animal models

Two different strains of mice were used in this study:

i) the i4F-BL6 transgenic "reprogrammable mice", containing a DOX-inducible polycistronic cassette harbouring the four Yamanaka factors inserted in one of the two copies of the *Pparγ* gene and carrying the transcriptional activator (rtTA) within the ubiquitously expressed *Rosa26* locus in a C57BL6 strain background (Abad et al., 2013);

ii) *Mus Musculus castaneus* (Cast) strain, a wild-type strain.

In order to produce polymorphic and heterozygous F1 mice (harbouring SNPs to distinguish the two parental alleles and containing the DOX inducible cassette), breedings were established using i4F-Bl6 females and Cast males and vice-versa. While F1 mice from the first cross generated several litters, the reciprocal cross (i4F-BL6 males and Cast females) were not able to generate any offspring by unforeseen reasons.

2.1.1. Genotyping for the Yamanaka cassette and rtTA insertion

DNA was extracted from mice ears using a standard extraction protocol. Briefly, the procedure started by an initial incubation with Lysis Buffer (500 µL per sample: 100 mM NaCl; 10 mM Tris, pH 8.0; 25 mM EDTA, pH 8.0; 0.5 % SDS; and 0.2 µg/µL of Proteinase K) at 56 ºC overnight. On the next day, 55 µl of 3 M sodium acetate pH 5.2 (Merck Millipore) and 500 µl of UltraPure™ Phenol: Chloroform: Isoamyl Alcohol were added. After mixing and centrifugation at maximum speed for 10 min at room temperature (RT), the top phase was transferred into new eppendorfs and 500 µL of Chloroform were added to remove Phenol residues. The samples were mixed and centrifuged and the resultant top phase was again transferred into new eppendorfs. After addition of isopropanol (Merck Millipore) to precipitate DNA, samples were incubated at -20 ºC for at least 1h. After centrifugation and removal of the supernatant, DNA pellets were washed with freshly made 70 % Ethanol (Merck Millipore) and then open air dried for around 20 min at RT. Finally, the dried pellets were resuspended in DNAse/RNAse free H_2O and DNA concentrations were measured using a Nanodrop 2000c (Thermo Fisher Scientific). The assessment of the presence of Yamanaka's cassette was done with a PCR reaction (**Table 2.1**) using a flanking primer (underlined in blue) and an internal lentiviral primer (see **Figure 2.1** and **Table 2.2**). A similar strategy was used to monitor rtTA cassette at the *Rosa26* locus using primers in the **Table 2.4**. As a control, a PCR for the normal *Rosa26* locus was also performed using the primers described in the **Table 2.3**.

Figure 2.1 - Yamanaka cassette insertion within Pparγ gene and flanking (in blue) and internal (in red) primers for PCR (adapted from Abad, M. et al. (2013)).

Table 2.1 - PCR reaction mix used for genotyping and *Rosa26* locus alleles assessment.

PCR Reaction Mix	
Reagent	**Volume per reaction (µL)**
DNAse/RNAse Free H_2O	14.68
Nzylong Reaction Buffer (10x) (NZYTech)	2
dNTPs 10mM (NZYTech)	0.6
Forward Primer 10µM (Sigma)	0.7
Reverse Primer 10µM (Sigma)	0.7
Nzylong 5 U/µL (NZYTech)	0.32
DNA	1
Total Volume	20

Table 2.2 - Primers and conditions used for genotyping assessment.

Primer	**Primer sequence**	**PCR Steps**	
Pparγ Fw	5'- CAGCATCAAATGGCTCGGTA -3'	94 ºC/2min	
		94 ºC/45s	40 Cycles
Lenti Rw	5'- GCACCATCCAAAGGTCAGTG -3'	59 ºC/45s	
		68 ºC/1min	
Source	Abad et al. 2013	68 ºC/5min	
Fragment Expected Size	350 bp		

Table 2.3 - Primers and conditions used for assessment of Rosa26 wild type allele.

Primer	**Primer sequence**	**PCR Steps**	
Rosa26 wt Fw	5'- AAAGTCGCTCTGAGTTGTTAT -3'	94 ºC/2min	
		94 ºC/45s	40 Cycles
Rosa26 wt Rw	5'- GGAGCGGGAGAAATGGAATG -3'	59 ºC/45s	
		68 ºC/1min	
Source	Abad et al. 2013	68 ºC/5min	
Fragment Expected Size	500 bp		

Table 2.4 - Primers and conditions used for assessment of rtTA cassette at the *Rosa26* locus.

Primer	Primer sequence	PCR Steps	
Rosa26 wt Fw	5'- AAAGTCGCTCTGAGTTGTTAT -3'	94 ºC/2min	
		94 ºC/45s	40 Cycles
Rosa26 Mut Rw	5'- GCGAAGAGTTTGTCCTCAACC -3'	59 ºC/45s	
		68 ºC/1min	
Source	Abad et al. 2013	68 ºC/5min	
Fragment Expected Size	300 bp		

2.2. Cell Sources

- **MEF-derived miPSCs**

MEF-derived miPSCs used in this study were generared previously in the lab by a Master student, João Paulo von Gilsa Lopes and Dr. Simão José Teixeira da Rocha. Briefly, MEFs were collected from polymorphic and heterozygous F1 e13.5 embryos and then reprogrammed into miPSCs by addition of DOX. In this study we used 22 independent polymorphic miPSCs lines, 11 female lines and 11 male lines, which 11 of them were generated with extra addition of 500 ng/mL AA to the culture media. Pellets from all miPSCs lines were collected by centrifugation at 1000 rpm for 3 min at RT. For DNA extraction the pellets were stored at -80 ºC whereas for RNA extraction the pellets were snap-freezed in liquid nitrogen and stored at -80 ºC.

- **Adult F1 fibroblasts and reprogramming**

Adult F1 fibroblasts were collected from ears of adult polymorphic and heterozygous F1 mice (3-4 weeks from the same breedings) by primary culture (check below).

- **Adult miPSCs**

The adult miPSCs were obtained from adult fibroblasts by the same reprogramming strategy.

- **mESCs**

Commercially available Jmj8-F6 ESCs, grown in both standard ESC medium and in 2i (two inhibitor) (2i) ESC medium conditions, and the Tx1072 female ESC (Schulz et al., 2014) grown only in 2i conditions.

2.2.1.Primary Culture

Skin fibroblasts were collected from ears of two female and two male adult mice (2 months). Ears were then placed on Petri dishes (60.1 cm^2) and sliced several times with a scalpel to 1 mm pieces in the presence of 500 µL of 0.25 % trypsin-EDTA (1x). The pieces were incubated at 37 ºC for about 20 minutes and then re-sliced and re-incubated for another 20 minutes. 5 mL of fibroblast medium (**Table 2.5**) were then added to the Petri dishes and kept at 37 ºC until next day. On the following day, these medium was recovered and transferred into falcons. Petri dishes were further washed with PBS 1X and these volume was then recovered and added to the same falcon tubes. Falcon tubes were centrifuged at 1000 rpm for 5 minutes and the supernatant were discarded. The pellets were then ressuspended in 5 mL of fibroblast medium and transferred to new pre-labelled dishes and incubated at 37 ºC. Medium

was changed every three/four days for 15 days, after which the cells in each dish were passaged, for the first time, to three/four wells, coated with 0.1 % porcine gelatine, of a P6 well plate. For that, the medium of each dish was discarded and a wash with PBS 1X was done. It was added 1 mL of trypsin to each dish which were incubated at 37 ºC for 5 minutes. Then, fibroblasts medium was added to each dish containing trypsin, in order to distribute 2 mL of cells per well. Three days later the cells were passaged again (P2), each sample to three new wells and to T25 or T75 flasks coated with 0.1 % porcine gelatine. The cells in the wells were used to reprogramming and the others to testing for the presence of mycoplasma and then to freeze, in cryovials containing 1 ml of fibroblasts freezing medium (**Table 2.6**), at -80 ºC.

Table 2.5 - Fibroblasts Medium Composition.

Product	Brand	Volumes
DMEM Dulbecco's Modified Eagle Medium	Gibco™ DMEM, Thermo Fisher Scientific	45 mL
Fetal Bovine Serum (FBS)	Gibco™ DMEM, Thermo Fisher Scientific	5 mL
5000 U/mL Pen/Strep	Gibco™ DMEM, Thermo Fisher Scientific	250 µL
200 mM L-Glutamine	Gibco™ DMEM, Thermo Fisher Scientific	500 µL

Table 2.6 - Fibroblasts Freezing Medium Composition.

Product	Brand	Volumes
Fibroblasts Medium		9 mL
DMSO	Sigma-Aldrich	1 mL

2.2.2. Adult miPSCs reprogramming

Adult fibroblasts (passage 2 – P2) were used for iPSC reprogramming. When fibroblasts (P2) grown until reached 40 % of confluency, the medium was changed to miPSCs medium containing 1 µg/mL of Dox (BDClontech) (**Table 2.7**) to start reprogramming. Medium was changed every two/three days until formation of iPSCs-like colonies.

Table 2.7 - miPSCs medium + DOX Composition.

Product	Brand	Volumes used
DMEM Dulbecco's Modified Eagle Medium	Gibco™ DMEM, Thermo Fisher Scientific	42 mL
Knockout-Serum Replacement	Gibco™ DMEM, Thermo Fisher Scientific	7.5 mL
5000 U/mL Pen/Strep	Gibco™ DMEM, Thermo Fisher Scientific	250 µL
200 mM L-Glutamine	Gibco™ DMEM, Thermo Fisher Scientific	500 µL
MEM Non-Essential Amino Acids 100X	Gibco™ DMEM, Thermo Fisher Scientific	500 µL
2-Mercaptoethanol 50 mM	Thermo Fisher Scientific	100 µL
LIF	Merk Millipore	5 µL
1 µg/mL Doxycycline	BDClontech	50 µL

2.2.3. Cell maintenance

MEF-derived miPSCs were maintained *in vitro* in miPSCs medium (**Table 2.8**) containing or not AA, being the medium changed every other day. In order to supplement the medium with AA, 10mg of AA powder (Sigma-Aldrich) was dissolved in 20 mL of MiliQ water and filtered using 0.45 µm syringe filter units (VWR). From this solution, 1 µL per mL was added to the culture medium. The AA solution was made fresh every time, due to the stability issues of prolonged storage of soluble AA. Finally, mESCs lines, were maintained *in vitro* using either ESC standard medium - JmJ6-F8 cell line – or the same medium supplemented with 2i – JmJ6-F8 and Tx1072 cell lines (see **Table 2.9** and **2.10**).

Table 2.8 - miPSCs Medium Composition (KSR+LIF).

Product	Brand	Volumes used
DMEM Dulbecco's Modified Eagle Medium	Gibco™ DMEM, Thermo Fisher Scientific	42 mL
Knockout-Serum Replacement	Gibco™ DMEM, Thermo Fisher Scientific	7.5 mL
5000 U/mL Pen/Strep	Gibco™ DMEM, Thermo Fisher Scientific	250 µL
200 mM L-Glutamine	Gibco™ DMEM, Thermo Fisher Scientific	500 µL
MEM Non-Essential Amino Acids 100X	Gibco™ DMEM, Thermo Fisher Scientific	500 µL
2-Mercaptoethanol 50 mM	Thermo Fisher Scientific	100 µL
LIF	Merk Millipore	5 µL

Table 2.9 - ES cell Standard Medium Composition.

Product	Brand	Volumes used
DMEM Dulbecco's Modified Eagle Medium	Gibco™ DMEM, Thermo Fisher Scientific	42 mL
ES Cell FBS	Gibco™ DMEM, Thermo Fisher Scientific	7.5 mL
5000 U/mL Pen/Strep	Gibco™ DMEM, Thermo Fisher Scientific	250 µL
200 mM L-Glutamine	Gibco™ DMEM, Thermo Fisher Scientific	500 µL
2-Mercaptoethanol 50 mM	Thermo Fisher Scientific	100 µL
LIF	Merk Millipore	5 µL

Table 2.10 - Two inhibitor (2i) ES cell Medium Composition.

Product	Brand	Volumes used
DMEM Dulbecco's Modified Eagle Medium	Gibco™ DMEM, Thermo Fisher Scientific	42 mL
ES Cell FBS	Gibco™ DMEM, Thermo Fisher Scientific	7.5 mL
5000 U/mL Pen/Strep	Gibco™ DMEM, Thermo Fisher Scientific	250 µL
200 mM L-Glutamine	Gibco™ DMEM, Thermo Fisher Scientific	500 µL
2-Mercaptoethanol 50 mM	Thermo Fisher Scientific	100 µL
LIF	Merk Millipore	5 µL
CHIR99021	Sigma	50 µL
PD0325901	Sigma	50 µL

2.3. iPSCs Characterization

2.3.1.Immunofluorescence and microscopy

Two lines of MEF-derived miPSCs – 2C C1AA and 9C E2AA - were analysed by immunofluorescence using the antibodies listed in **Table 2.11**. miPSCs were plated on gelatine-coated coverslips for 24/48 hours. Then the cells were fixed in 3 % paraformaldehyde (PFA) in PBS for 10 minutes at RT, washed with PBS 1X freshly prepared and permeabilized in 0.5 % (w/v) Triton X-100 in PBS for 4 minutes on ice. A blocking step was performed by incubation with 1 % Bovine Serum Albumin (BSA; Sigma-Aldrich) in PBS for 15 minutes and subsequently the cells were incubated with primary antibodies for OCT4 and SSEA1 (**Table 2.11**) for 45 minutes and then with the secondary antibody (**Table 2.11**) for 45 minutes in a dark humid chamber. After three washes with PBS, nuclei were counterstained with DAPI for 2 minutes and then coverslips were mounted in Vectashied Mount Medium (Vector Catalog# H-1000) and stored at 4 ºC or 20 ºC until visualization in a Widefield Fluorescence Microscope Zeiss Axio Observer (Carl Zeiss MicroImaging). Images were processed using ImageJ software (version 1.51h).

Table 2.11 - List of primary and secondary antibodies used for immunofluorescence microscopy

	Antibody	Host	Supplier	Dilution
Primary Antibody	Anti-OCT4	Mouse	Millipore	1:200
	Anti-SSEA1	Mouse	Millipore	1:100
Secondary Antibody	Goat Anti-mouse CY3	Goat	Millipore	1:100

2.3.2. Embryoid body formation

In order to evaluate miPSCs differentiation in a variety of cell types, embryoid body formation was performed using the hanging drop culture method, as described by Kramer. J et al. (2006) (Kramer et al., 2006). Briefly, 1000 cells were plated in 20 µL-drops of differentiation medium (**Table 2.12**) on the bottom of a bacterial petri dishes. Then, the drops were maintained in suspension in a petri dish lid filled with medium and incubated at 37 ºC for 4 days. Afterwards the dish was turned up and medium was

added upon drops. Subsequently EBs were incubated at 37 ºC for 8 extra days and at D12, EBs were pelleted for RNA extraction. To pellet the EBs, all the content of the dish was transferred to a falcon which was centrifuged at 1000 rpm for 3 minutes. The supernatant was discarded and the pellet was first washed in PBS, centrifuged, and finally ressuspended in 1 mL of NZYOL ™ RNA Isolation Reagent (Nzytech) and then stored at -80 ºC for RNA extraction.

Table 2.12 - Differentiation Medium Composition.

Product	Brand	Volumes
DMEM Dulbecco's Modified Eagle Medium	Gibco™ DMEM, Thermo Fisher Scientific	45 mL
Fetal Bovine Serum (FBS)	Gibco™ DMEM, Thermo Fisher Scientific	5 mL
5000 U/mL Pen/Strep	Gibco™ DMEM, Thermo Fisher Scientific	250 µL
200 mM L-Glutamine	Gibco™ DMEM, Thermo Fisher Scientific	500 µL
2-Mercaptoethanol 50 mM	Thermo Fisher Scientific	100 µL

2.3.3. RT-qPCR

Total RNA was extracted by a standard protocol using NZYOL™ RNA Isolation Reagent (Nzytech) and follow supplier's guidelines. Subsequently, DNase I treatment (Roche) and 100 % ethanol precipitation were performed on 5 µg of RNA sample to remove traces of DNA. The resultant RNA was quantified by NanoDrop 2000c (Thermo Fisher Scientific) and used directly for cDNA synthesis. For cDNA synthesis, 500 ng of DNAse I-treated RNA samples were used and submitted to a Transcriptor High Fidelity cDNA Synthesis Kit (Roche) according to the manufacturer's instructions. Gene expression analysis was performed on these cDNA products in duplicated by RT-qPCR on RealTime PCR System 7000 (Applied Biosystems, Foster City, CA, USA) using iTaqTM SYBR® Green Supermix (Bio-Rad). Reaction mix and cDNAs were loaded in white 384 well plates, always maintaining reagents and plates on ice. Cycle threshold (Ct) and melting curves were determined using QuantStudio™ Real-Time PCR Software version 1.3 (Thermo Fisher Scientific). All data was normalized using the *Gapdh* housekeeping gene and analyzed through the $2^{-\Delta\Delta Ct}$ method. The primers used in RT-qPCRs are described on **Table 2.13**.

Table 2.13 - List of primers used for RT-qPCR analysis.

Gene	Primer Forward (F) Primer Reverse (R)
m*Gapdh*	**F-** AACTTTGGCATTGTGGAAGG **R-** ACACATTGGGGGTAGGAACA
m*Esrrb*	**F-** TCTCATCTTGGGCATCGTGT **R-** AGTTTCTTGTACCTGCGCAC
m*Nanog*	**F-** CCAGTCCCAAACAAAAGCTC **R-** ATCTGCTGGAGGCTGAGGTA
m*Oct4*	**F-** CCGGAAGAGAAAGCGAACTA **R-** CGCCGGTTACAGAACCATAC
m*Fgf5*	**F-** TGAAAAGACAGGCCGAGAGT **R-** TCTGTACTTCACTGGGCTGG
m*Nestin*	**F-** CTCTCCCTGACTCTACTCCCT **R-** CATCTTCTTCCTCTCCCTCTT
m*AFP*	**F-** GCTCACACCAAAGAGTCAAC **R-** CCTGTGAACTCTGGTATCAG
m*T(Brachyury)*	**F-** TGAACCTCGGATTCACATCG **R-** CTGGTAGGCAGTCACAGCTA
m*Gata1*	**F-** TTGGACACCTTGAAGACGG **R-** GCATAAGATGGCTGACAGGC
m*Tnnt2*	**F-** AGAGAAGGCCAAGGAGCTGT **R-** AGCTCCTTGGCCTTCTCTCT

2.3.4. Teratoma assay

To evaluate the capacity for teratoma formation, miPSCs were trypsinized and two different cell amount - 2×10^6 and 1×10^6 cells - were subcutaneously injected by Bruno de Jesus into the flanks of 3 months old immunocompromised mice (NOD-SCID mice from Charles River). Animals were killed by anesthetic overdose and a necropsy was performed. Subcutaneous tumors were harvested and fixed in furmol. Tumors were sent to histology facility to be examined by a pathologist blinded to experimental groups in a Leica DM2500 microscope coupled to a Leica MC170 HD microscope camera. All this animal experimentation was conducted strictly within the rules of the Portuguese official veterinary directorate, which complies with the European Guideline 86/609/EC concerning laboratory animal welfare, according to a protocol approved by the institute's Animal Ethics Committee.

2.4. Imprinting Assessment

2.4.1.DNA extraction

In order to analyse methylation status, DNA from the different miPSCs cell lines was obtained. Genomic DNA extraction was performed as previously described in session 2.1.1 from cell pellets collected during from in vitro culture of MEFs, miPSCs and mESCs.

2.4.2. DNA Methylation Analysis

I. Bisulfite Treatment

Genomic DNA from all lines of miPSCs was treated with sodium bisulfite, which converts all unmethylated cytosines in the DNA to uracil, leaving the methylated cytosine residues intact (**Figure**

2.2). Bisulfite treatment was performed using the EZ DNA Methylation-Gold™ kit (Zymo Research), according to the manufacturers' instructions. Briefly, 500 ng of genomic DNA were diluted in a final volume of 20 µL. Then a conversion step was performed in a thermal cycler. Subsequently, DNA was cleaned and desulphonated in spin columns. Finally, 12 µL of Bisulfite treated DNA per sample were resuspended in DNAse/RNAse free H_2O. The treated DNA samples were then amplified by PCR approaches designed for the different regions.

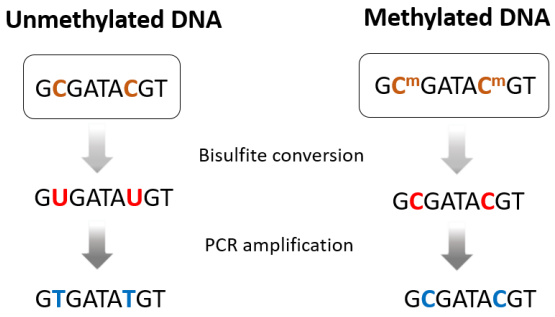

Figure 2.2 - Bisulfite treatment process.

II. PCR approaches

- **Nested PCR**

After bisulfite treatment, DNA gets rich in thymine, which could compromise specificity of downstream PCRs. So a Nested PCR approach was done for the majority of regions. The primers for 1st and 2nd PCRs were chosen according to the regions under study and in order to allow SNP detection. The nested PCRs were performed using NZYLong DNA Polymerase (NZYTech) for *Peg3* and *Igf2-H19* loci. A PCR mix for all reactions was done as described in **Table 2.14**. For the 1st PCR, 1 µL of bisulfite treated DNA was used, while for the 2nd PCR, 2 µL of the 1st PCR product were used.

Table 2.14 - PCR reaction mix used for Nzylong DNA Polymerase.

Reagent	Volume per reaction (µL)
DNAse/RNAse Free H_2O	38.5
Nzylong Reaction Buffer (10x) (NZYTech)	5
dNTPs 10 mM (NZYTech)	1.5
Forward Primer 10 µM (Sigma)	1.75
Reverse Primer 10 µM (Sigma)	1.75
NZYlong 5 U/µL (NZYTech)	0.5
Sample	1 or 2
Total Volume	50

- **PCR**

For *Dlk1-Dio3* locus and *KvDMR* locus only one PCR was done. Both PCRs were performed using KAPA HiFi HotStart Uracil (Roche), which is a DNA polymerase engineered to work better on bisulfite-treated DNA since is rich in uracils. A PCR mix for all reactions was done as described in **Table 2.15**.

Table 2.15 - PCR reaction mix used for KAPA HI FI HotStart Uracil 2x.

Reagent	Volume per reaction (in µL)
DNAse/RNAse Free H_2O	7.75
Forward Primer 10 µM (Sigma)	1.875
Reverse Primer 10 µM (Sigma)	1.875
KAPA HI FI 2X (Roche)	12.5
Sample	1
Total Volume	25

Conditions and primers used for all PCR approaches can be observed from **Table 2.16** to **Table 2.19**.

Table 2.16 - Primers and conditions used for nested PCR of *Peg3* locus.

Primer	Primer sequence	PCR Steps	
1st PCR			
mPeg3DMR Nested F1	5'-GGTTTTGGATTGGTTAGAGAGGAAGTT-3'	95 ºC/5 min	
		95 ºC/30s	35 Cycles
mPeg3DMR Nested R1	5'-TCCCTATCACCTAAATAACATCCCT-3'	56 ºC/30s	
		72 ºC/45s	
Source	Customized	72 ºC/10 min	
2nd PCR			
m*Peg3*DMR Nested F2	5'-TTTTGTAGAGGATTTTGATAAGGAG-3'	95 ºC/5 min	
		95 ºC/30s	
m*Peg3*DMR Nested R2	5'- CAACCTTATCAATTACCCTTAAAAA-3'	56 ºC/30s	35 Cycles
		72 ºC/45s	
Source	Customized	72 ºC/10 min	
Fragment Expected Size	352 bp		

Table 2.17 - Primers and conditions used for nested PCR of *Igf2-H19* locus.

Primer	Primer sequence	PCR Steps	
1st PCR			
m*H19* DMR outF	5'- GAGTATTTAGGAGGTATAAGAATT -3'	95 ºC/5 min	
		95 ºC/4 min	2 Cycles
m*H19* DMR outR	5'-ATCAAAAACTAACATAAACCCCT -3'	55 ºC/2 min	
		72 ºC/2 min	
Source	Nakamura et al., 2007	95 ºC/1 min	35 Cycles
		55 ºC/2 min	
		72 ºC/2 min	
		72 ºC/10 min	
2nd PCR			
m*H19* DMR inF	5'- GTAAGGAGATTATGTTTTATTTTTGG -3'	95 ºC/5 min	
		95 ºC/30 s	35 Cycles
m*H19* DMR inR	5'- CCTCATAAAACCCATAACTAT -3'	56 ºC/30 s	
		72 ºC/45 s	
Source	Nakamura et al., 2007	72 ºC/10 min	
Fragment Expected Size	422 bp		

Table 2.18 - Primers and conditions used for PCR of *KvDMR* locus.

Primer	Primer sequence	PCR Steps	
	1st PCR		
***Kv*DMR-IF**	5'- TAAGGTGAGTGGTTTAGGAT -3'	95 ºC/5 min	
		98 ºC/20 s	35 Cycles
***Kv*DMR-OR**	5'- AATCCCCCACACCTAAATTC -3'	55 ºC/15 s	
		72 ºC/1 min	
Source	Shin et al. 2009	72 ºC/10 min	
Fragment Expected Size	420 bp		

Table 2.19 - Primers and conditions used for PCR of *Dlk1-Dio3* locus.

Primer	Primer sequence	PCR Steps	
	1st PCR		
m*IG*-DMR F	5'- GTGGTTTGTTATGGGTAAGTTT - 3'	95 ºC/5 min	
		98 ºC/20 s	35 Cycles
m*IG*-DMR R	5'- CCCTTCCCTCACTCCAAAAATTAA -3'	57 ºC/15 s	
		72 ºC/1 min	
Source	Nakamura et al., 2007	72 ºC/10 min	
Fragment Expected Size	318 bp		

After PCR, the amplified products were separated by electrophoresis in 1 % TAE agarose gels, stained with 3 µL of Xpert Green DNA Stain (Grisp) per 100 mL of solution. Subsequently, the resultant bands were cut out with clean scalpels and gel-purified using NZYGelpure (NZYTech) kit. Following purification of the different PCR amplicons of the regions of interest, assessment of methylation status was carried out either by Combined Bisulfite Restriction Analysis (COBRA) or Bisulfite Sequencing.

III. Combined Bisulfite Restriction Analysis (COBRA)

Combined Bisulfite Restriction Analysis (COBRA) method relies on the use of restriction enzymes to digest the purified DNA, harbouring the regions of interest, derived from bisulfite-treated DNA. The action of such enzymes is influenced by the original methylation status of the sequence. For example, Bsh1236I is a restriction enzyme that cuts at CG/CG, which means that, since methylated cytosines remain as cytosines after bisulfite treatment, the enzyme will only cut sequences in which original genomic DNA was methylated (see **Figure 2.3**). The restriction enzymes were chosen according to their recognition sequences in order to cut in cytosines in the context of CpG and taking into account that at least one restriction site is present within the methylated bisulfite converted strand and absent in the unmethylated bisulfite-converted strand. Conditions and restriction enzymes used for each locus can be observed in **Table 2.20**. After digestions, the samples were separated by electrophoresis in 2 % TAE agarose gels, stained with 3 µL of Xpert Green DNA Stain (Grisp) per 100 mL of solution for 40 minutes. Images were obtained through the Chemidoc XRS+ system (BioRad) and analysed using the Image Lab 5.2 software (BioRad).

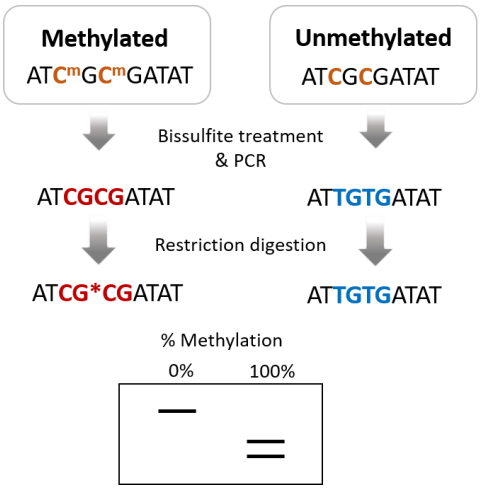

Figure 2.3 - Outline of COBRA protocol.

Table 2.20 - Restriction enzymes and conditions used for COBRA of each region of interest.

		***Peg3* locus**	***Igf2-H19* locus**	***KvDMR* locus**	***Dlk1-Dio3* locus**
Resctriction enzyme		Taq I	Bsh 1236I	Bsh 1236I	MluI
Buffer					
Cut sequence		T/CGA	CG/CG	CG/CG	A/CGCGT
Conditions		65 ºC 1h30min	37 ºC 1h30min	37 ºC 1h30min	37 ºC 1h30min
Fragments size expected	Total	352 bp	422 bp	420 bp	318 bp
	1st fragment	165 bp	87 bp	211 bp	206 bp
	2nd fragment	187 bp	335 bp	209 bp	112 bp

IV. Bisulfite Sequencing

Purified PCR fragments of *Peg3* DMR were ligated to a plasmid using pGEM®-T Easy Vector System (Promega) according to manufacturer's instructions and incubated overnight at 4 ºC. On the following day, DH5-α competent *E. coli* were transformed with the ligation products, using 100 µL of competent cells and 5 µL of ligations. The transformation was performed using heat shock at 42 ºC for 45 seconds, followed by 2 min on ice. Transformed *E. coli* were then allowed to recover in liquid Lysogeni Broth (LB) for 1 hour and then, transformed cells were platted into LB containing Agar with 100 µg/mL of Ampicillin for selection. IPTG and X-gal were used to distinguish between recombinant plasmids (white colonies) from auto-ligated ones (blue colonies). IPTG is a chemical analogue of galactose which functions as an inductor agent, and X-gal, a component hydrolysed by β –galactosidase when present and functional, resulting in a blue colour product. This is an easy way to differentiate cells containing the insert of interest correctly ligated - white colonies - from the ones that do not - blue colonies. After an overnight incubation at 37 ºC, isolated white colonies were picked (~8 for each transformation) and incubated overnight in 3 mL of LB with 100 µg/mL of ampicillin at 37 ºC with shaking. On the next day, plasmid DNA was purified using NZYMiniprep Kit (NZYTech) following manufacturer's instructions. Briefly, bacteria were lysed, and the plasmid DNA was purified using spin columns. Finally,

samples were sent for Sanger sequencing at GATC Biotech, using 5 µL of purified plasmid DNA and 2.5 µL of 5 mM pUC/M13 Reverse Primer (as recommended by the pGEM®-T Easy Vector System protocol). Sequences were then visualized by Chromas Software (version 2.6.4) and analysed by BiQ Analyzer Software (version 2.00) which allows both assessment of DNA methylation status from bisulfite sequencing and SNP detection.

V. Allelic-specific expression analysis

To verify expression of *Peg3* gene with allelic-specific information, a PCR was performed. The amplified products were separated by electrophoresis in 1 % TAE agarose gels, stained with 3 µL of Xpert Green DNA Stain (Grisp) per 100 mL of solution. Subsequently, the resultant bands were cut out with clean scalpels and purified using NZYGelpure (NZYTech) kit according to the manufacturer's instructions. Finally, the purified products were sent for Sanger Sequencing with GATC BioTech. The PCR mix was done as described on **Table 2.21**. Primers and conditions used can be observed in **Table 2.22**.

Table 2.21 - PCR reaction mix used for allelic-specific expression.

Reagent	Volume per reaction (µL)
DNAse/RNAse Free H_2O	34.6
NH4 Reaction Buffer (10x) (Bioline)	5
MgCl2 (50 mM) (Bioline)	3
dNTPs 10 mM (NZYTech)	1
Forward Primer 10 µM (Sigma)	2
Reverse Primer 10 µM (Sigma)	2
BioTaq 2500 U/µl (Bioline)	0.4
cDNA sample	2
Total Volume	50

Table 2.22 - Primers and conditions used for PCR amplification for allelic-specific expression analysis.

Primer	Primer sequence	PCR Steps	
m*Peg3* SNP F	5'- AAGGCTCTGGTTGACAGTCGTG -3'	95 ºC/5 min	
		95 ºC/1 min	33 Cycles
m*Peg3* SNP R	5'- TTCTCCTTGGTCTCACGGGC -3'	60 ºC/30 s	
		72 ºC/1 min	
Source	Mann et al., 2004	72 ºC/10 min	
Fragment Expected Size	60 bp		

VI. RNA Fluorescence in situ hybridization (RNA-FISH)

RNA FISH probes for *Xist* (a 19 Kb genomic λ clone 510) and *Meg3* (WT-2686H19 BAC - BACPAC Resources Center) were prepared using the Nick Translation Kit (Abbot) with red (*Xist*) or green (*Meg3*) dUTPs (Enzo Life Sciences). The lines of MEF-derived miPSCs – 2C B5, 2C A5AA, 2C C1, 2C C1AA and 9C E2AA - were analysed using the *Xist* probe, while, 2C B5, 2C A5AA, 2C C1AA and 9C A8 lines were analysed for the *Meg3* probe. RNA FISH was done accordingly to established protocols (Chaumeil et al., 2008). Briefly, cells were dissociated with trypsin (Gibco), centrifuged for 3 minutes at 1000 rpm

and ressuspended in miPSC medium. The protocol was performed using 10 mm coverslips previously incubated in Poly-L-Lysine (SIGMA) during 5 minutes and dried at open air. So the cells were pipetted in the centre of the coverslips and incubated for 10 minutes at RT. Cells were then fixed in 3 % PFA in PBS for 10 min at RT and permeabilized with 0.5 % Triton X-100 diluted in PBS with 2 mM Vanadyl-ribonucleoside complex (New England Biolabs) for 5min in ice. Coverslips were then washed twice in ethanol (EtOH) 70 % for 5 min and then dehydrated through an ethanol series (80 %, 95 % and 100 %) and air-dried quickly before hybridization with the fluorescent labelled probes. Probes were ethanol precipitated with sonicated salmon sperm DNA (and Cot1 DNA for *Meg3* probe) denatured at 75 ºC for 7 min (in the case of *Meg3* probe, it was incubated at 37 ºC after denaturation to allow Cot1 to bind to the repetitive DNA present in these BAC probe to prevent unspecific hybridization). *Xist* (in red) and *Meg3* (in green) probes were co-hybridized in FISH hybridization solution (50% formamide, 20% dextran sulfate, 2x SSC, 1 µg/µl BSA, 10mM Vanadyl-ribonucleoside) overnight. Washes were carried out with 50 % Formamide/2x SSC (three times for 7 min at 42 ºC) and then 2xSCC (three times for 5 min at 42 ºC). After the RNA FISH procedure, nuclei were stained with DAPI (Sigma-Aldrich), and mounted with Vectashield Mounting Medium (Vectorlabs). Cells were observed with a Widefield Fluorescence Microscope Zeiss Axio Observer (Carl Zeiss MicroImaging) and the obtained images were processed using ImageJ software (version 1.51 h).

2.5. Methylation Imprinting Correction

2.5.1. RT-PCR

Total RNA was extracted and processed exactly as described above. cDNA products were then used as template for RT-PCR in a 25 µl reaction volume (**Table 2.23**). The primer pairs and conditions used are described in **Table 2.24**.

Table 2.23 - RT-PCR reaction mix.

Reagent	Volume per reaction (in µL)
DNAse/RNAse Free H_2O	18.79
Nzylong Reaction Buffer (10x) (NZYTech)	2.5
dNTPs 10 mM (NZYTech)	0.75
Forward Primer 10 µM (Sigma)	0.88
Reverse Primer 10 µM (Sigma)	0.88
Nzylong 5 U/µL (NZYTech)	0.2
cDNA Sample	1
Total Volume	25

Table 2.24 - Primers and conditions used for RT-PCR.

Primer	Primer sequence	PCR Steps	
Zfp57 Fw	5'- CACAAATCCACAAAGCCGCAA -3'	94 ºC/2 min	
		94 ºC/30 s	30 Cycles
Zfp57 Rw	5'- TGAACGGGGCCTATAACCTAAA -3'	60 ºC/30 s	
		68 ºC/30 s	
Source	Customized	68 ºC/10 min	
Fragment Expected Size	151 bp		
Trim28 Fw	5'- CGCATGTATCAGGCATGAAG -3'	94 ºC/2 min	
Trim28 Rw	5'- CTTCCAGGAAAGACCTTGAAGA -3'		
Source	Kumar et al., 2016		
Fragment Expected Size	148 bp		
Dppa3 Fw	5'- GACCCAATGAAGGACCCTGAA -3'	94 ºC/30 s	33 Cycles
Dppa3 Rw	5'- GCTTGACACCGGGGTTTAG -3'	60 ºC/30 s	
Source	Customized		
Fragment Expected Size	130 bp	68 ºC/30 s	
Gapdh Fw	5'- TTCTCCTTGGTCTCACGGGC -3'	68 ºC/10 min	
Gapdh Rw	5'- TTCTCCTTGGTCTCACGGGC -3'		
Fragment Expected Size	223 bp		

3. Results and Discussion

To evaluate the stability of imprints in miPSCs, we used a system where all the donor cells harbour both an inducible Yamanaka cassette for controlled reprogramming induction and SNPs to distinguish between the two parental alleles. This system was already implemented in our lab by João Von Gilsa and Dr. Simão Teixeira da Rocha who obtained F1 polymorphic MEFs from a cross between a female mouse (i4F-BL6) harbouring a DOX-inducible Yamanaka cassette inserted on its genome (Abad et al., 2013) and a male mouse from the *Mus Musculus Castaneus* (*Cast*) strain. Polymorphic F1 MEFs obtained from both female and male embryos were reprogrammed into several independent iPSCs by DOX administration. We expect this system to finally determine and report all the imprinting defects that occur in miPSCs upon reprogramming. The use of both female and male iPSCs on this study was motivated by the findings for gender differences in global demethylation during reprogramming, which is more pronounced in female than male cells (Milagre et al., 2017), with possible consequences for imprinting stability. Female cells are designated by 2C and the male cells by 9C. In addition, some iPSCs lines were generated and maintained with AA supplementation, since it has been described as a strategy to reduce imprinting defects at the mouse *Dlk1-Dio3* region (Stadtfeld et al., 2012). These lines are designated using the abbreviation AA (*e.g.* 2C C1AA or 9C E2AA). Besides miPSCs, their parental F1 MEFs from which these miPSCs were generated (2C Fib and 9C Fib), and the following mESC lines: the TX1072 female ESC line (an ESC derived from a similar cross between a male *Cast* and a female C57BL/6 grown in 2i/LIF conditions) (Schulz et al., 2014) and the male Jmj8-F6 ESC line (grown in both standard ESC medium and in 2i/LIF conditions) were also used on this study as controls for imprinting errors in mESCs.

3.1. Characterization of MEF-derived miPSCs

MEF-derived iPSCs previously generated in the lab presented normal morphology and proliferation rate, as expected for mouse SCs *in vitro* (data not shown). However, a comprehensive characterization of their stemness and pluripotent potential was not previously performed. For that reason, the first task was to phenotypically characterize these MEF-derived iPSCs. A first approach consisted in analysing the presence/expression of typical mESC markers both by immunofluorescence (IF) and RT-qPCR. Then, we performed different pluripotency tests to assess their capacity to generate other cell types.

3.1.1. MEF-derived miPSCs express pluripotent markers

In order to characterize the phenotypic identity of miPSCs, two stem cell markers were assessed by IF: a nuclear marker – OCT4 – and a membrane marker – SSEA-1. This assessment was performed in two miPSC lines, the female 2C C1AA and the male 9C E2AA (**Figure 3.1**). As depicted in **Figure 3.1**, both iPSC lines strongly express these stem cell markers, which is consistent with its stem-like state.

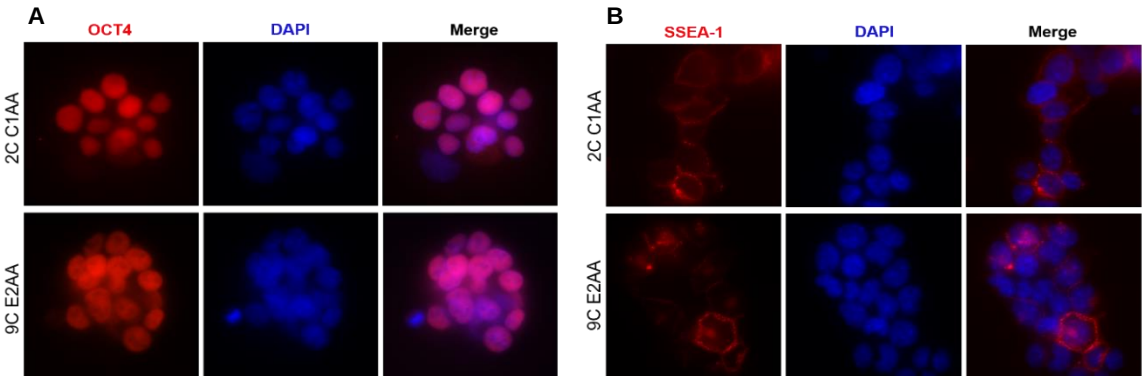

Figure 3.1 - IF for OCT4 and SSEA-1 in MEF-derived miPSCs. (A) Detection of OCT-4 (red) in 2C C1AA and 9C E2AA miPSCs lines. **(B)** Detection of SSEA-1 (red) in 2C C1AA and 9C E2AA miSPCs lines. Nuclei were counterstained with DAPI (blue).

Then, to quantify and confirm the expression of pluripotency markers, a RT-qPCR was performed for all the miPSCs lines, using their respective parental MEFs and the mESCs as negative and positive controls, respectively. This analysis was performed by our lab manager Ana Cláudia Raposo and the pluripotency markers assessed were *Oct4*, *Nanog* and *Esrrb* (**Figure 3.2**).

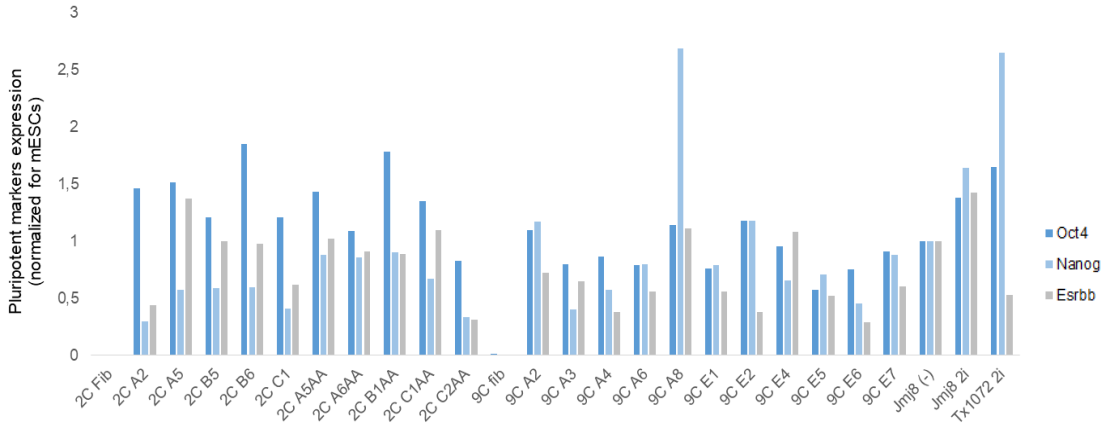

Figure 3.2 - RT-qPCR analysis of *Oct4*, *Nanog* and *Esrrb* markers expression in all miPSCs lines, parental MEFs and mESCs lines. Graph represents the relative expression of *Oct4*, *Nanog* and *Esrrb* normalized to *Gapdh* housekeeping gene from only one independent experiment; Relative expression was set to 1 for the Jmj8 (-) sample.

As expected, gene expression analysis supported and extended our IF data, confirming that all miPSC lines express the three pluripotent markers at equivalent levels to mESCs lines. In contrast, for the parental MEFs (2C Fib and 9C Fib), no expression was observed. In conclusion, all 2C and 9C-derived iPSCs express the tested stem cell markers at the IF and RT-qPCR levels, and therefore, they seem to be true SC lines. This was expected as this system was already used in the past to successfully and easily generate iPSCs (Abad et al., 2013; Bernardes de Jesus et al., 2018). A full characterization of stem cell markers could be extended in the future by looking to additional pluripotent markers or by high-throughput RNA-sequencing and compare to the normal mESCs transcriptional signature. To better characterize these cells, we decided then to test the potential of our miPSCs by pluripotency tests.

3.1.2. miPSCs lines are able to differentiate into different cell types

The next step was to understand whether our lines were truly pluripotent. For that, we decided to generate EBs for the lines: 2C C1AA and 9C E2AA (**Figure 3.3**). There are several methods to perform this assay (Kurosawa, 2007). We chose the hanging drop culture method to generate EBs to better control the number of iPSCs *per* drop to allow formation of homogeneous EBs.

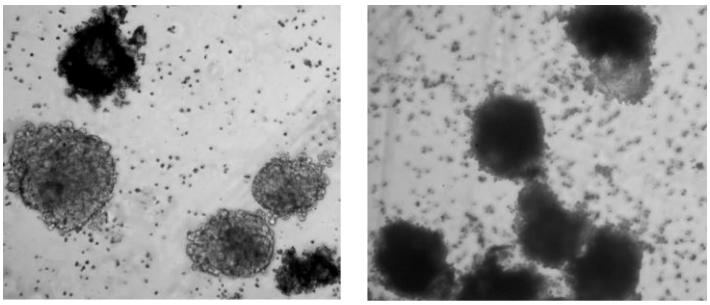

Figure 3.3 - Embryoid body formation assay for 2C C1AA (left) and 9C E2AA (right) miPSCs.

After 12 days of culture in suspension, spherical EBs were observed *in vitro* for both iPSC lines (**Figure 3.3**). To confirm their differentiation capacity into cells of the three germ layers, a gene expression analysis based on RT-qPCR was performed. The following set of primers was used: stem cell markers (*Esrrb*, *Oct4* and *Nanog*), an ectoderm marker (*Nestin*), mesoderm markers (*T-Branchyury*, *Gata1* and *Tnnt2*), an endoderm marker (*AFP*) and also a marker for early stages of embryogenesis (*Fgf5*). For this analysis, we concentrated only in one round of EBs formed from a female (2C C1AA) and a male (9C E2AA) iPSC lines (**Figure 3.4**).

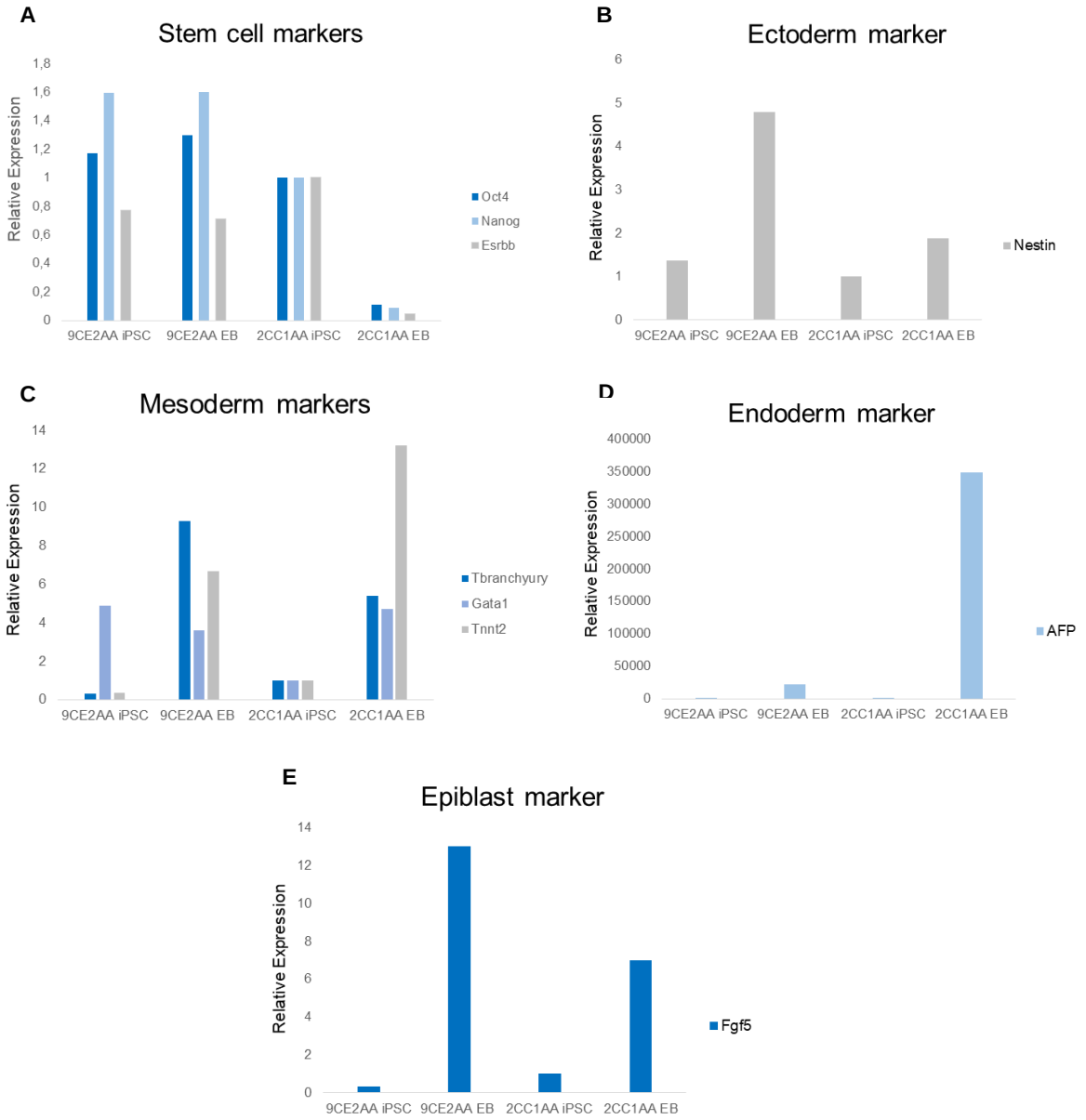

Figure 3.4 - RT-qPCR analysis of stem cell and differentiation markers expression in both miPSCs and EBs samples formed from 2C C1AA and 9C E2AA miPSC lines. (A) Expression of stem cell markers *Oct4*, *Nanog* and *Esrrb*. **(B)** Expression of the ectoderm marker *Nestin*. **(C)** Expression of mesoderm markers *T-branchyury*, *Gata1* and *Tnnt2*. **(D)** Expression of the endoderm marker *AFP*. **(E)** Expression of the epiblast (early stage of embryogenesis) marker *Fgf5*. Graphs represent the relative expression normalized to *Gapdh* housekeeping gene from two independent experiments; Relative expression was set to 1 for the 2CC1AA miPSC line.

For the 2C C1AA line, results were in line with the expected since high expression levels of stemness markers were only observed for the miPSC samples, while high expression of the differentiation markers was seen only in the EB samples. Surprisingly, for the line 9C E2AA, high expression of stemness markers remained in the EBs, which was not expected. Despite this, the differentiation markers also become upregulated in EB samples.

In conclusion, 2C C1AA iPSC line is able to downregulate the stemness program and differentiate into the three germ layers, which is consistent with a bona-fide iPSC line. The 9C E2AA line, on the contrary, did not completely shut down the stemness program, despite being also able to differentiate

into cells from the different germ lineages. The reasons why the stemness program was not properly downegulated upon EB formation in this line is not understood. In any case, we could see that EB formation was quite heterogeneous from experiment to experiment (data not shown) and therefore re-assessement should be done in the future in replicate experiments. Of note, EBs were also successfully formed from other lines such as 9C A8 and 2C B5 (data not shown), however no expression analysis was performed for these lines.

Since we were concerned about the heterogeneity of EB differentiation, we decided to perform teratoma assays, which is a more stringent method to evaluate pluripotency capacity. In this assay, two subcutaneously injections with a total of 1-2 x 10^6 miPSCs of the following lines: 2C B5, 2C A5AA, 9C A8 and 9C E2AA were performed in immune-deficient mice (NOD-SCID mice from Charles River).

All the lines showed capacity to form teratomas upon injection, with the exception of the 2C B5 line (**Figure 3.5**). In this line, no teratoma was observed after approximately one month after injection. Additionally, fast-growing teratomas were noticed in mice injected with the lines 9C A8 and 9C E2AA (both 2 x 10^6 cells) and 2C A5AA, since it was already possible to see a considerable tumor-like mass at day 12 (D12) after infection. **Figure 3.5 (B)** describes the weight, measures and the day when each teratoma was harvested. In all mice injected were observed two teratomas, except in the mouse injected with 1 x 10^6 cells from 9C A8 line.

A

9C A8 (D21) 9C E2AA (D21)

B

	Line	Weight (g)	Measures (cm)	Dissection upon injection
1 x 10^6 cells	2C A5AA	0.54	1 x 1.2 x 0.5	Day 19
		0.75	1.5 x 1.3 x 1	
	9C A8	1.25	1.8 x 1.5 x 0.7	Day 21
	9CE2AA	0.78	2.1 x 1.2 x 0.7	
		0.84	1.2 x 1.2 x 0.7	
2 × 10^6 cells	9C A8	1.21	2 x 1 x 1	Day 19
		0.17	1 x 0.4 x 0.2	
	9C E2AA	0.30	1 x 1 x 0.2	
		0.27	1 x 1 x 0.5	

Figure 3.5 - Teratoma assay for 2C A5AA, 9C A8 and 9C E2AA miPSCs lines. (A) Tumor-like masses at day 21 upon injection of 1 x 10^6 cells from 9C A8 (left) and 1 x 10^6 cells from 9C E2AA (right) in immune-deficient mice. **(B)** Weight and measures of tumor-like masses of each line and day at tumor-like masses were harvested upon injection.

It is reassuring that tumours were formed by all the lines except the 2C B5 line. For this line, further attempts will be necessary to understand whether they could or not generate tumour masses. Of note, for this line we only used 1 x 10^6 cells instead of the standard 2 x 10^6 cells, which could partially explain the failure. However, this was not a problem for the other three lines. Presently, the results to assess the differentiation into endodermal, mesodermal, and ectodermal derivatives are being analysed by the iMM histology facility. Therefore, we cannot yet comment on the pluripotency status of these iPSCs, other than saying that they can make tumour masses. This analysis might confirm or not our hypothesis that 9C E2AA line might still have many single cells that maintain a stem-like state even being in a differentiation setting.

The successful capacity to form 3D aggregates observed from our miPSCs lines accompanied by RT-qPCR data with different markers sthat our miPSCs lines have the capacity to differentiate into other cell types. The teratoma assay, which is an *in vivo* and more stringent pluripotency test, will confirm the true pluripotency potential of our miPSCs. Therefore, from the results of the previous pluripotent markers assessment and from both *in vitro* and *in vivo* pluripotent assays so far, we believed that the miPSCs generated by a well-defined system (Abad et al., 2013) are indeed true iPSCs.

3.1.3. Female MEF-derived miPSCs lines are mostly XO

Gender specific differences in DNA methylation dynamics are known to occur during reprogramming (Milagre et al., 2017). This is thought to be caused by the fact that reprogramming of female somatic cells to iPSCs induces the reactivation of the inactive X chromosome (Maherali et al., 2007) as a late event in the reprogramming process (Pasque et al., 2014). However, the presence of two X chromosomes seems to correlate with low passage number, and over time in culture the cells tend to lose one X chromosome and regain DNA methylation levels similarly to male iPSCs (Pasque et al., 2018). This is a phenomenon that also has long been noticed for female mESCs (Choi et al., 2017a; Choi et al., 2017b). To see if our female miPSCs contained the two X chromosomes or already lost one, we performed a RNA FISH for a probe detecting both *Xist/Tsix antisense pair***,** two X-linked genes essential for regulation of X-chromosome inactivation in female cells (Gendrel & Heard, 2011) (**Figure 3.6 (A)**). In undifferentiated mESCs, this RNA FISH probe (Chaumeil et al., 2008) gives one clear fluorescent dot corresponding to the site of transcript *per* X-chromosome being a good read-out for the copies of X-chromosomes as used previously (Pasque et al., 2018).

We used female miPSCs lines with relatively low passages (2C B5 (P11), 2C A5AA (P14) and 2C C1 (P10)). In the case of 2C C1AA, we looked at both low (P12) and a high passages (P24). As a control, we used the male 9C E2AA miPSCs line, which is XY harbouring only one X chromosome (**Figure 3.6**).

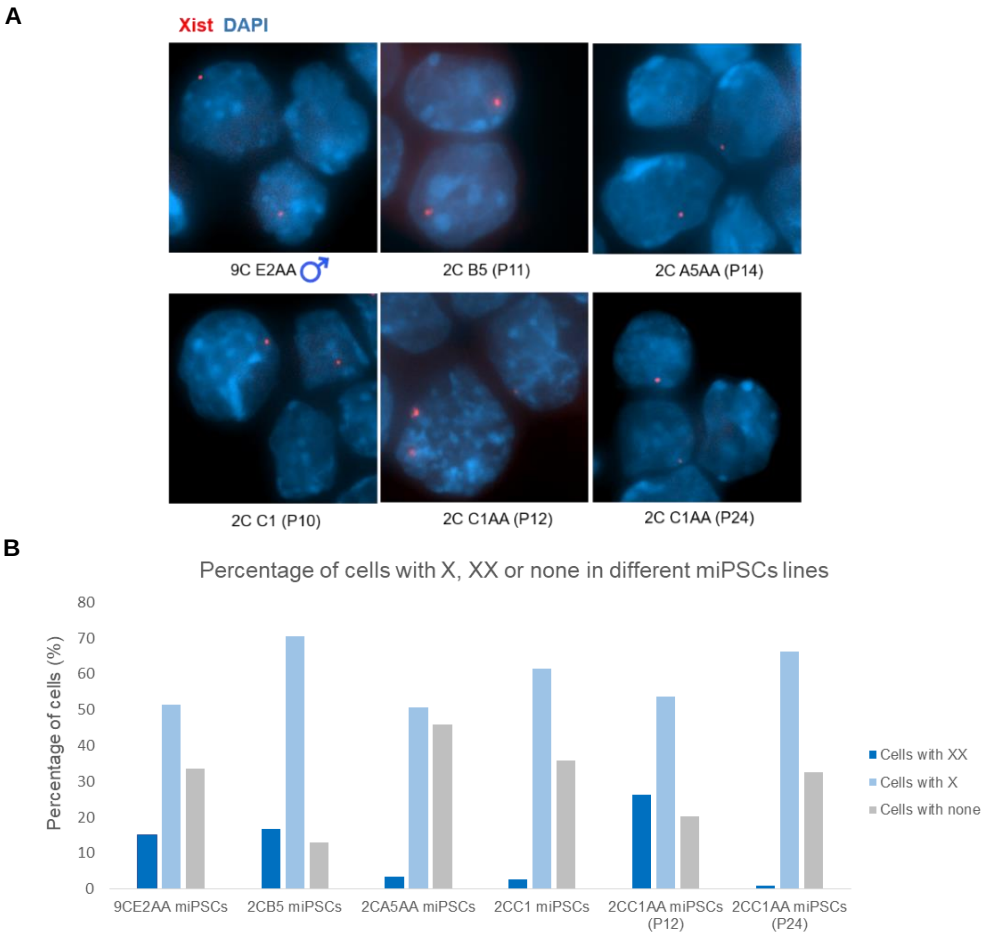

Figure 3.6 - RNA FISH analysis of *Xist/Tsix antisense* pair expression in male 9C E2AA and female 2C B5, 2C A5AA, 2C C1 and 2C C1AA miPSCs lines with different cell passages. (A) RNA-FISH analysis of *Xist/Tsix* (red dot). Nuclei were counterstained with DAPI (blue). **(B)** Quantification of the percentage of cells with two dots, one dot or no dot for *Xist/Tsix*. The female miPSC line 2C C1AA was analyzed at both passage 12 and passage 24. A total >130 of cells was counted *per* sample. Data are from only one independent experiment.

For the male 9C E2AA line, one dot was detected in about 50 % of cells. Interestingly, cells with two dots were observed in 15 % of cells. This result was unexpected. This might be a counting artefact, since two dots in close proximity were scored as two, but they could be, in fact, from the same X-chromosome already with two chromatids. Alternatively, we cannot rule out cases of tetraploidy, which are rare, but could happen in ESCs. Importantly, in around 30 % of cells no *Xist/Tsix* signal was observed setting up the limits of this technique.

When we look to the female iPSCs, we found that the majority of them had cells with one dot comparable to the 9C E2AA XY line. This suggests that female iPSCs already lost one of the two X chromosomes. Interestingly, for the 2C C1AA line at low passage was counted a considerable number of cells with two FISH dots comparable to the same line at higher passages. This suggests that these miPSCs lost one X chromosome over time in culture as previously reported by Pasque, V et al. (Pasque et al., 2018).

In conclusion, our female iPSCs already lost one X-chromosome even though they were from a reasonable low passage. To confirm that female iPSCs have indeed lost one of the two chromosomes, we could assess SNPs on the X-chromosome distinguishing the *Cast* or the C57BL/6 X-chromosomes.

As a consequence, we expect overall DNA methylation levels being equivalent between female and male iPSCs.

3.2. Reprogramming of adult fibroblasts

After characterizing the MEF-derived miPSCs, we tried to generate adult miPSCs by reprogramming adult fibroblasts using the same F1 hybrid cross (i4F-B female vs *Cast* male) used to obtain F1 MEFs. **Figure 3.7** summarizes the schematic process of a reprogramming process.

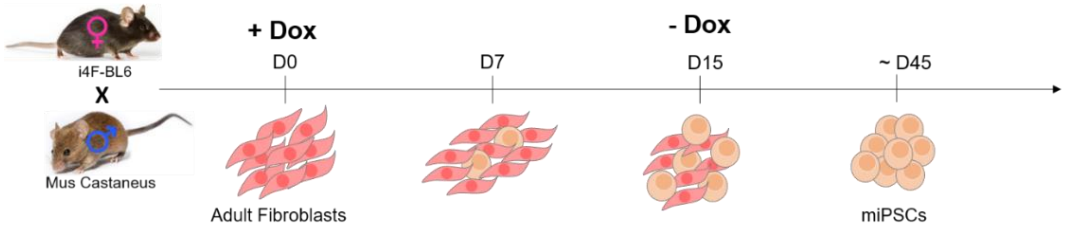

Figure 3.7 - Schematic process of reprogramming of adult fibroblasts into miPSCs.

3.2.1. Adult fibroblasts did not reprogram into iPSCs

In order to generate miPSCs from adult cells for further DNA methylation analysis, adult fibroblasts were firstly obtained from ears of two adult heterozygous and polymorphic F1 mice (a male, 731, and a female, 736). The fibroblasts were successfully obtained by primary culture (**Figure 3.8**).

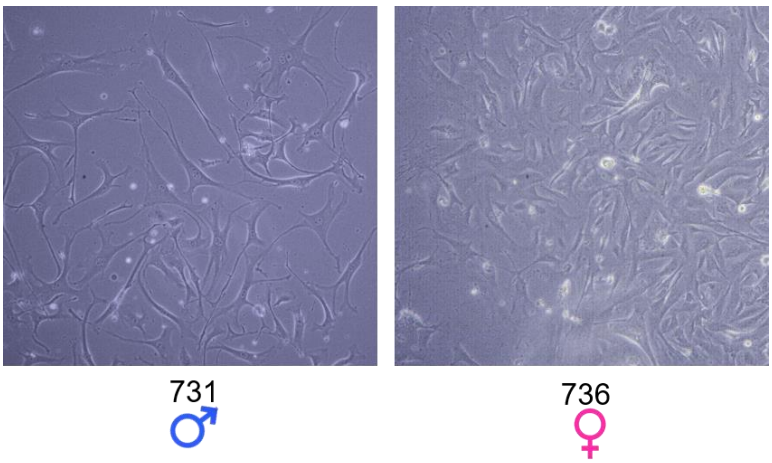

Figure 3.8 - Adult fibroblasts obtained by primary culture from ears of adult heterozygous and polymorphic F1 male (left) and female (right) mice.

Then, reprogramming of the two fibroblast populations was initiated at 40 % confluence by addition of miPSCs medium supplemented with DOX. During this period, some signs of morphological changes in the original fibroblast culture were noticed, however, they never really expand as it would be normal for an iPSC colony (**Figure 3.9**). Therefore, iPSCs-like colonies never developed/formed into true iPSCs, even when letting DOX induction past 15 days. Therefore, they were not picked due to their small size and absence of correct ESC-like morphology.

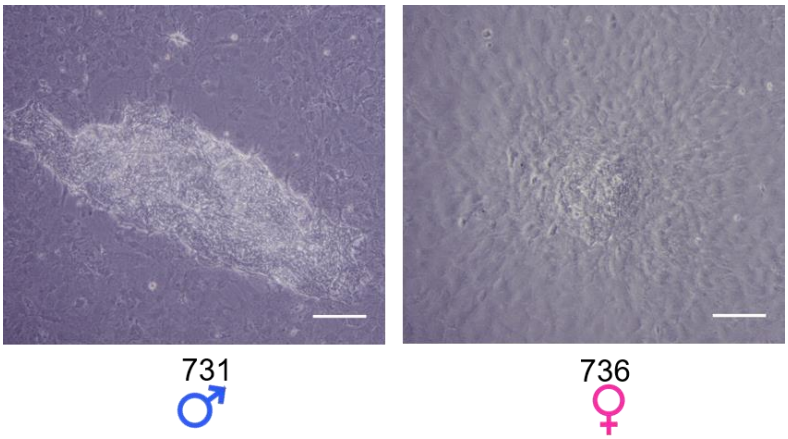

Figure 3.9 - miPSC-like structure, in miPSCs medium supplemented with DOX at day 15 (right) and day 22 (left) of reprogramming- Scale bars represent 300 µm.

Given this unexpected outcome, mice were genotyped to confirm the presence of the Yamanaka cassette. We confirmed that all the mice have the Yamanaka cassette in their genome (**Figure 3.10 (A)**). We used a positive (with the Yamanaka cassette) and a negative sample (without the Yamanaka cassette) to confirm these results. We then wondered if the problem was due to the loss of rtTA cassette inserted at the *Rosa26* locus, which is a DOX-inducible transcriptional activator within the *Rosa*26 locus (Gossen & Bujard, 1992). To confirm whether rtTA was present, we performed another PCR (called *Rosa26* mutant allele) and alongside a control PCR for the normal *Rosa26* locus (*Rosa26* wt allele) (**Figure 3.10 (B)**). All mice showed a positive result for the *Rosa26* mutant allele, indicating the presence of the *Rosa26* mutant allele, which was not present in the negative samples (an unspecific band, indicated by an asterisk, of slight bigger size, normally arises in this protocol). As expected all the mice also contained the *Rosa26* WT allele, since rtTA was integrated only in one allele.

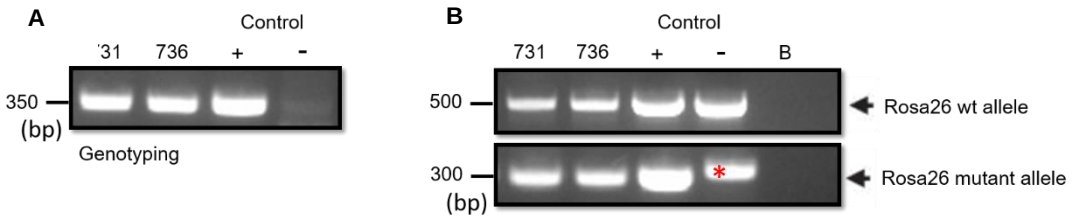

Figure 3.10 - Monitoring of Yamanaka cassette and rtTA insertion in female and male adult heterozygous and polymorphic F1 fibroblasts by PCR. (A) PCR for Yamanaka cassette detection. **(B)** PCR for *Rosa26* locus (up) and for rtTA (down) detection. The red asterisk correspond to a typical unspecific band in the protocol.

These results suggested that fibroblasts contained all the necessary machinery for reprogramming, therefore, we cannot understand why adult fibroblasts did not reprogram, as this has been achieved before (Bernardes de Jesus et al., 2018). We hypothesise that an unexpected and undetected technical problem might have arisen during these experiments: for example, the quantity of initial fibroblasts which was measured by confluence might not have been ideal or perhaps the DOX used was not effective enough to induce expression of Yamanaka transcription factors. Other possibility is that probably the hybrid F1 mice used for generation of donor cells might lead to the inactivation of OSKM cassette with

aging, given the different genetic background compared to the previously used (Abad et al., 2013; Bernardes de Jesus et al., 2018). In the future, it will be important to re-try this experiment and evaluate whether the Yamanaka cassette was or not properly induced. However, given the time constrains for my Master thesis, we decided not to investigate this further. We decided instead, to concentrate on studying imprinting stability on the MEF-derived miPSC lines we characterized.

3.3. Methylation Imprinting Assessment

To evaluate the extent to which imprinting defects occur in miPSCs, we assessed the methylation status of the ICR of some imprinted clusters. Five described imprinted clusters were chosen to evaluate the stability of imprints in both female and male miPSCs, with or without AA supplementation: *Peg3*, *Igf2-H19*, *Kcnq1-Kcnq1ot1*, *Dlk1-Dio3* and the *PWS/AS* loci. In addition to miPSCs, the parental MEFs from which these miPSCs were generated, and mESCs controls - the TX1072 female ESC line (grown in 2i/LIF conditions) and Jmj8-F6 male ESCs lines (grown in both 2i/LIF and classical medium conditions) – were also analysed.

3.3.1. Imprinted clusters showed a trend for hypomethylation defects in MEF-derived miPSCs

The initial methylation assessment was performed by COBRA, which allows a quick detection of both methylated and unmethylated alleles in a non-allelic specific manner. The data is illustrated for a few analysed lines and loci in **Figure 3.11** and the full analysis is summarized in **Table 3.1**.

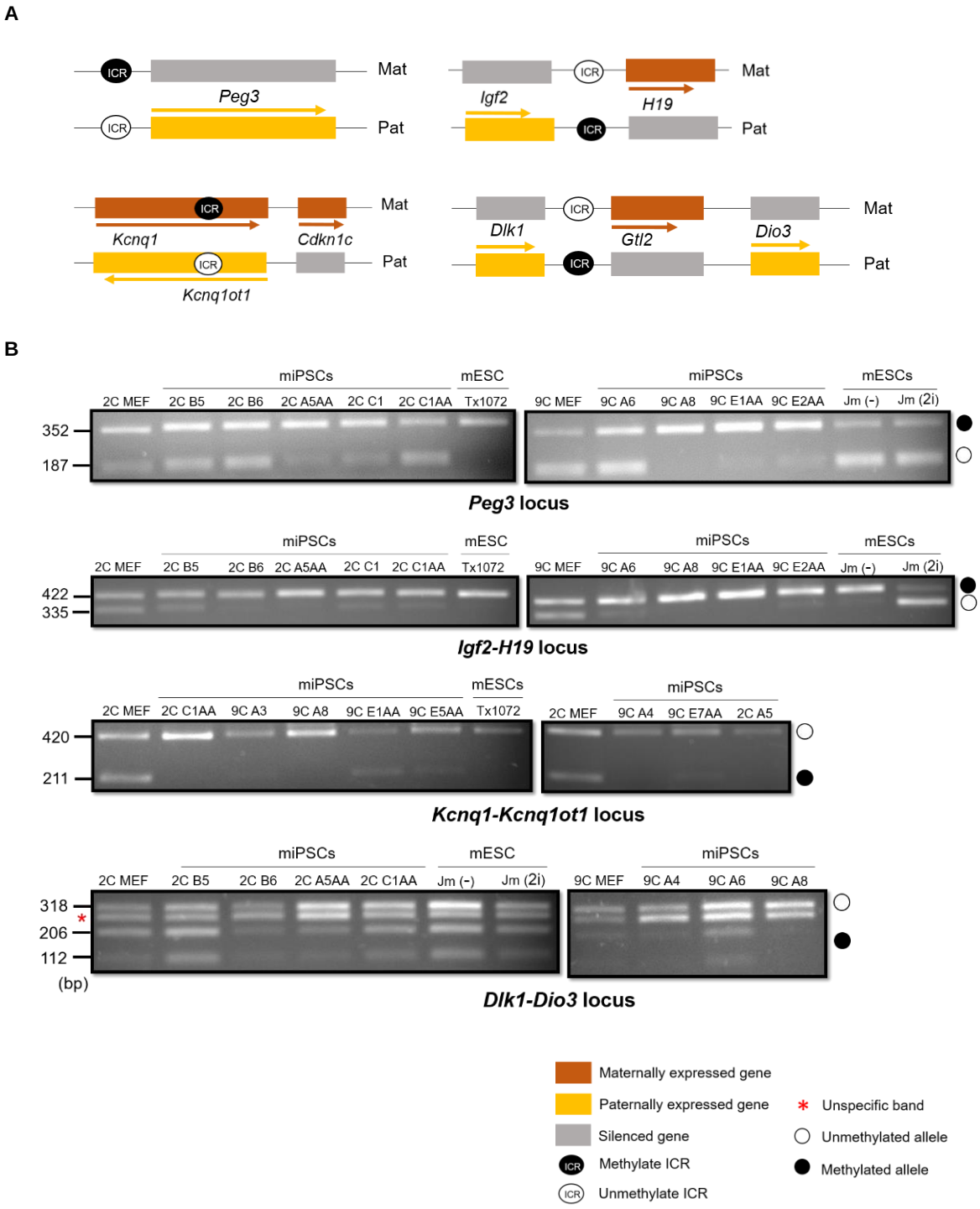

Figure 3.11 - Methylation status at *Peg3*, *Igf2-H19*, *Kcnq1-Kcnq1ot1* and *Dlk1-Dio3* ICRs in parental MEFs, female and male miPSCs and mESCs by COBRA. (A) Schematic representation of the imprinted expression landscape at imprinted *Peg3*, *Igf2-H19*, *Kcnq1-Kcnq1ot1* and *Dlk1-Dio3 loci.* **(B)** COBRA for *Peg3*, *Igf2-H19*, *Kcnq1-Kcnq1ot1* and *Dlk1-Dio3 loci.*

Table 3.1- Summary of the methylation imprinting defects found in all miPSCs, parental MEFs and mESCs lines obtained by COBRA. n.d – not done.

Sex	Line	*Peg3* locus	*Igf2-H19* locus	*KvDMR* locus	*Dlk1-Dio3* locus	*PWS/AS* locus
Female	2C Fib	○/●	○/●	○/●	○/●	○/●
Female	2C A2	○/●	○/●	○/●	n.d	○/◑
Female	2C A4	○/●	○/●	○/●	n.d	○/●
Female	2C A5	○/●	○/◑	○/○	n.d	n.d
Female	2C B5	○/●	○/◑	○/○	○/●	○/●
Female	2C B6	○/●	○/●	n.d	○/◑	○/●
Female	2C C1	○/◑	○/●	○/●	n.d	○/◑
Female	2C A5 AA	○/◑	○/◑	○/●	○/◑	○/◑
Female	2C A6 AA	●/●	n.d	○/○	n.d	n.d
Female	2C B1 AA	n.d	n.d	○/○	n.d	○/●
Female	2C C1 AA	○/●	○/◑	○/○	○/●	○/◑
Female	2C C2 AA	○/●	○/●	○/●	n.d	○/●
Male	9C Fib	○/●	○/●	n.d	○/◑	○/●
Male	9C A2	n.d	○/◑	n.d	n.d	○/◑
Male	9C A3	○/◑	○/◑	○/◑	n.d	○/◑
Male	9C A4	○/◑	○/○	○/○	○/◑	○/◑
Male	9C A6	○/●	○/◑	n.d	○/◑	○/◑
Male	9C A8	○/○	○/○	○/○	○/○	○/○
Male	9C E1 AA	○/◑	○/○	○/●	n.d	○/○
Male	9C E2 AA	○/◑	○/◑	n.d	n.d	○/●
Male	9C E4 AA	○/◑	○/◑	n.d	n.d	○/◑
Male	9C E5 AA	○/◑	○/◑	○/◑	n.d	○/◑
Male	9C E6 AA	○/○	○/○	n.d	n.d	○/◑
Male	9C E7 AA	○/◑	○/◑	○/◑	n.d	○/◑
Female	Tx 1072 ESC (2i)	○/○	○/◑	○/○	n.d	n.d
Male	JmJ6-F8 (-)	●/◑	○/◑	n.d	○/◑	n.d
Male	JmJ6-F8 (2i)	●/◑	●/◑	n.d	○/●	n.d

○ Unmethylated allele

● Methylated allele

◑ Partially methylated allele

Figure 3.11 (A) illustrates schematically the imprinted expression within four clusters analysed in this figure. At both *Peg3* and *Kcnq1ot1-Kcnq1* loci, the ICRs (*Peg3* DMR and *KV*DMR, respectively) are methylated on the maternal allele (Green et al., 2007; He & Kim, 2014). In contrast, at the *Igf2-H19* and *Dlk1-Dio3* loci the ICRs (*H19* DMR and *IG*-DMR, respectively) are methylated on the paternal allele (Bell & Felsenfeld, 2000; Stadtfeld et al., 2010).

The ICR methylation status of all imprinted regions was successfully obtained by COBRA. This assay relies on the use of restriction enzymes that will cleave the DNA, harbouring the regions of interest, depending on the original methylation status of the sequence (Materials and Methods). Therefore, the non-cleaved products represent the unmethylated allele (white circles) and the cleaved PCR products represent the methylated allele (black circles) (**Figure 3.11 (B)**). As expected for all imprinted regions analysed, both methylated and unmethylated alleles were detected in parental MEFs, from both female (2C) and male (9C), demonstrating the normal ICR imprinting pattern. An exception was *Dlk1-Dio3* locus in 9C MEF, which was biased towards the unmethylated allele, a result which needs to be confirmed in the future. In striking contrast to the parental MEFs, most iPSCs do not have the normal pattern. As showed in **Figure 3.11 (B)** and **Table 3.1**, for the *Peg3* locus, only a minority of miPSCs maintained the right methylation profile, while the majority have a clear prominence for the unmethylated allele, which suggests a partial loss of methylation on the methylated allele. In the case

of the 9C A8 line and 9C E6AA, no methylated allele could even be detected, suggesting a complete reversal of the maternal allele to the unmethylated state. For both *H19* DMR and *Kv*DMR, very few lines of miPSCs have the expected methylation profile on the ICRs, with methylation almost completely missing in most miPSC lines. These is suggestive that both ICRs are prone to lose methylation during iPSCs reprogramming. A tendency for hypomethylation was also observed for many of the miPSCs in the *PWS/AS* locus (this data was obtained by a colleague of mine, Inês Godinho, who is also a Master student). Also, for *Dlk1-Dio3* locus, although this COBRA protocol gives an unspecific band (identified with a red asterisk in **Figure 3.11 (B)**), it was observed that only the 2C B5 and 2C C1AA miPSCs lines remain with the expected methylation profile. All the others have a tendency for certain degree of hypomethylation, which is in contrast to previous findings describing this locus as being the most defective due to hypermethylation imprinting defects (Stadtfeld et al., 2010). However, it is in line with two previous reports which also found hypomethylation at *Dlk1-Dio3* using non-allelic approaches (COBRA and pyrosequencing) in some iPSC lines (Sun et al., 2012; Takikawa et al., 2013).

Finally, for mESCs, our results also confirmed that 2i/LIF conditions have a major impact on imprinting. Indeed, the TX1072 mESCs showed clear hypomethylation for all the loci analysed, corroborating published results (Greenberg & Bourc'his, 2015). The Jmj6-F8, which was cultured on classical medium or recently adapted to 2i/LIF medium showed variable effects on imprinting, exhibiting either hypomethylation, hypermethylation or no imprinting defects depending on the imprinted loci.

An important aspect of our results is the variability of imprinting defects on the different generated isogenic iPSCs. For example, the female 2C A4 and 2C C2AA miPSC lines presented normal imprinting pattern for all ICRs analysed, while 9C A8 and 9C E5AA lines presented, complete and partial loss of methylation for all of the ICRs, respectively. Also some imprinted loci are more prone to imprinted errors than others (Greenberg & Bourc'his, 2015). In our study, this is demonstrated by the *Igf2-H19* locus, which was affected in all miPSCs, while the *Peg3* locus was affected only in some of them.

In order to understand, what could cause this variability between miPSCs, we looked into whether imprinting errors were more or less frequent according to AA supplementation or gender. We observed that AA supplementation seems not to ameliorate imprinted defects in miPSCs. With the exception of the 2C C2AA line, all the other iPSCs showed imprinting defects at similar proportions to miPSCs not supplemented with AA during reprogramming and maintenance. The protective effect on imprinting maintenance by AA at the *Dlk1-Dio3* locus (Stadtfeld et al., 2012) seems also not to be observed, as in our system, *Dlk1-Dio3* is actually hypomethylated, rather than hypermethylated, even without AA. In any case, since only a few iPSC lines were analysed for *Dlk1-Dio3* locus, we cannot affirm that the presence of AA was not protective for epigenetic disturbances at this locus. However, for the other loci, we concluded that the presence of AA did not make a difference in our experiment.

We then analyse the effect of gender. Both female and male iPSC lines showed hypomethylation defects. Surprisingly, male iPSCs, in general, displayed more hypomethylation defects than female iPSCs. This result was not expected since female cells suffer more pronounced global demethylation during reprogramming than male cells (Milagre et al., 2017). Our iPSCs are already XO, as demonstrated by *Xist/Tsix* RNA-FISH (**Figure 3.6**). Therefore, it is expected that they have equivalent levels of DNA methylation to male iPSCs (Pasque et al., 2018). However, as reported by Pasque et al,

depletion in ICR methylation would not be fully re-established after X chromosome loss. Our data does not completely fit with these previous findings. It rather suggests that both female and male cells undergo a similar level of hypomethylation during reprogramming, which affect imprinting in both genders. However, female cells, by losing one X-chromosome, might recover certain level of methylation at ICRs which justifies an overall milder imprinting defects when analysed by COBRA (discussed further in the Conclusion and Future Perspectives).

3.3.2. Hypomethylation of the maternal *Peg3* DMR causes loss of imprinting of *Peg3* gene.

COBRA allows for quick methylation assessment at ICRs, however, it has some limitations, namely the fact that only one CpG is being analysed at the time and no allelic-specific information is provided. As an alternative, other methods to study methylation can be used, such as bisulfite sequencing. This assay relies on the ligation, cloning and sequencing of post-PCR single amplicons to obtain its methylation information of several CpG dinucleotides. Furthermore, taking advantage of our polymorphic miPSCs, this methodology distinguishes methylation profiles on the two parentally inherited alleles. We decided to look into the 9C A8 miPSC line, which lost completely methylation for all ICRs analysed and the parental 9C MEF, with normal differential methylation. We used the *Peg3* locus as a read-out.

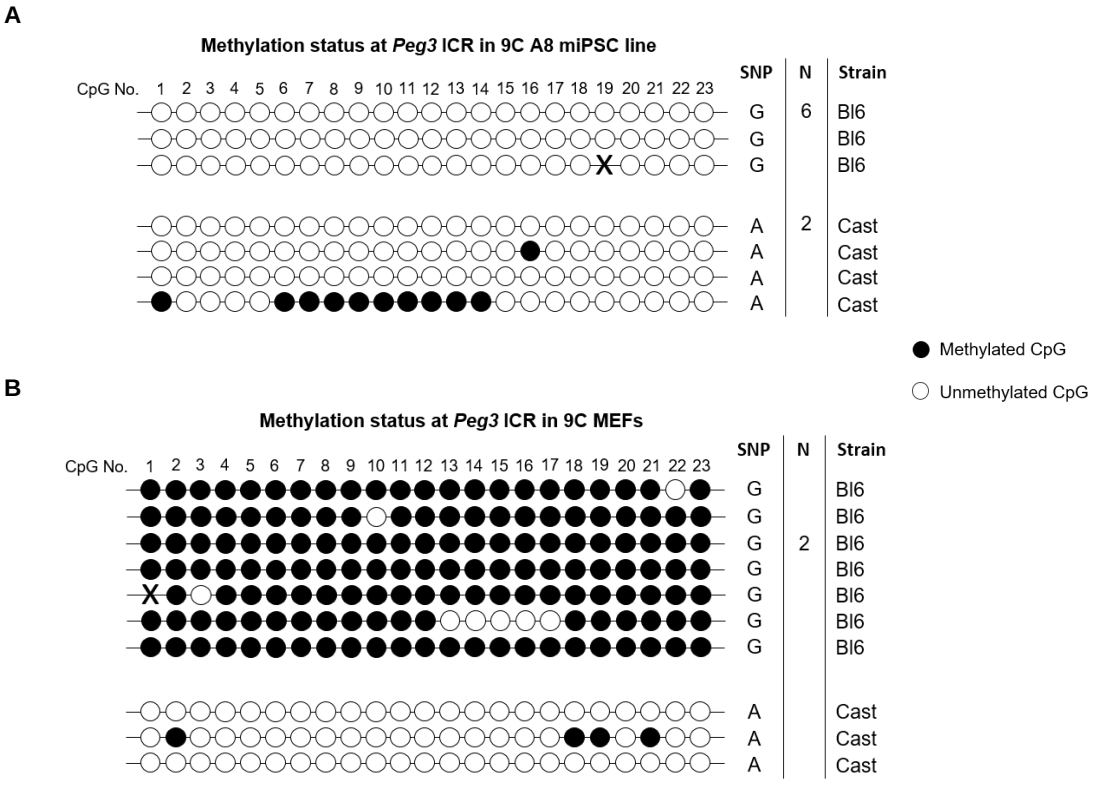

Figure 3.12 - Bisulfite sequencing for the methylation status of *Peg3* ICR in 9C MEFs and 9C A8 miPSCs. (A) Bisulfite sequencing results for the *Peg3* ICR in 9C A8 miPSCs line. **(B)** Bisulfite sequencing results for the *Peg3* ICR in MEFs.

Figure 3.12 shows the individual sequenced PCR amplicons of bisulfite-treated DNA from 9C A8 miPSC line and 9C MEFs, for the *Peg3* DMR region, the ICR at this locus. In 9C MEFs, the normal methylation imprinting pattern was kept, with the maternal allele (Bl6 origin) being mostly methylated, and the paternal allele (Cast) being unmethylated. In contrast, the 9C A8 miPSC line presented a loss of methylation, with both paternal (Cast) and maternal (BL6) alleles being unmethylated in basically all the CpGs. These results demonstrated that, in fact, the maternal allele at the *Peg3* domain of the 9C A8 line lost the original methylation status present in the donor 9C cell line. These results are concordant with COBRA results, and are also in line with previous studies that reported hypomethylation at the *Peg3* ICR (Sun et al., 2012; Takikawa et al., 2013).

Then, to understand the impact of this methylation change in the imprinted expression of *Peg3* gene, we performed allelic-specific expression analysis for *Peg3* thanks to the existence of 2 SNPs between C57BL6 and Cast genomes in the *Peg3* gene in our polymorphic miPSCs (**Figure 3.13**).

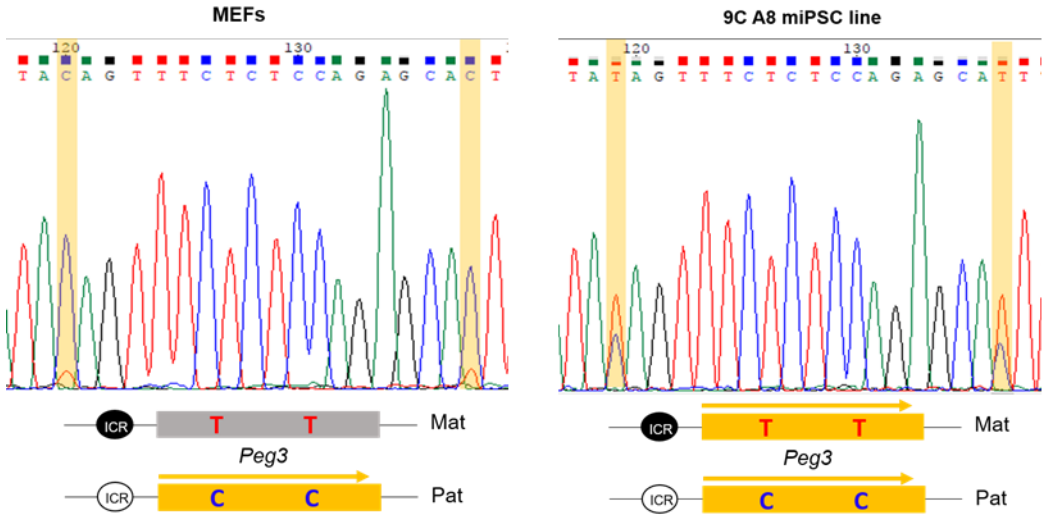

Figure 3.13 - *Peg3* allelic-specific expression analysis. The two SNPs are highlighted in the chromatogram in yellow; Maternal SNPs corresponds to two Ts, while paternal SNPs corresponds to two Cs.

In MEFs, we verified that *Peg3* was expressed almost exclusively from the paternal allele as expected for normal imprinting pattern. In contrast, in 9C A8 miPSCs line, both the paternal and maternal SNPs were detected, indicating biallelic expression of the *Peg3* gene. This result suggested that hypomethylation of the maternal *Peg3* DMR has functional consequences leading to loss of imprinting of *Peg3* gene. This result also suggests that our COBRA results, despite the lack of allelic-specific information, might be a reliable approach to assess the methylation imprinting status in a quick fashion.

3.3.3. Hypomethylation of the *IG*-DMR causes loss of imprinting of *Meg3* gene at the *Dlk1-Dio3* cluster

Dlk1-Dio3 cluster was the first region in which imprinting defects were found in miPSCs (Stadtfeld et al., 2010). However, the results obtained by COBRA were opposite to the original hypermethylation result previously described (Stadtfeld et al., 2010). Therefore, we decided to address the consequences of the DNA methylation alterations at this ICR on the allelic expression of the maternally expressed

Meg3 gene, which encodes for a long non-coding RNA located in *Dlk1-Dio3* locus. To get single-cell resolution, we performed nascent-transcript RNA-FISH using a probe detecting *Meg3* RNA (**Figure 3.14**). It is expected that hypomethylation will cause *Meg3* to become biallelically expressed according to what is known about the regulation of this locus (da Rocha et al., 2008).

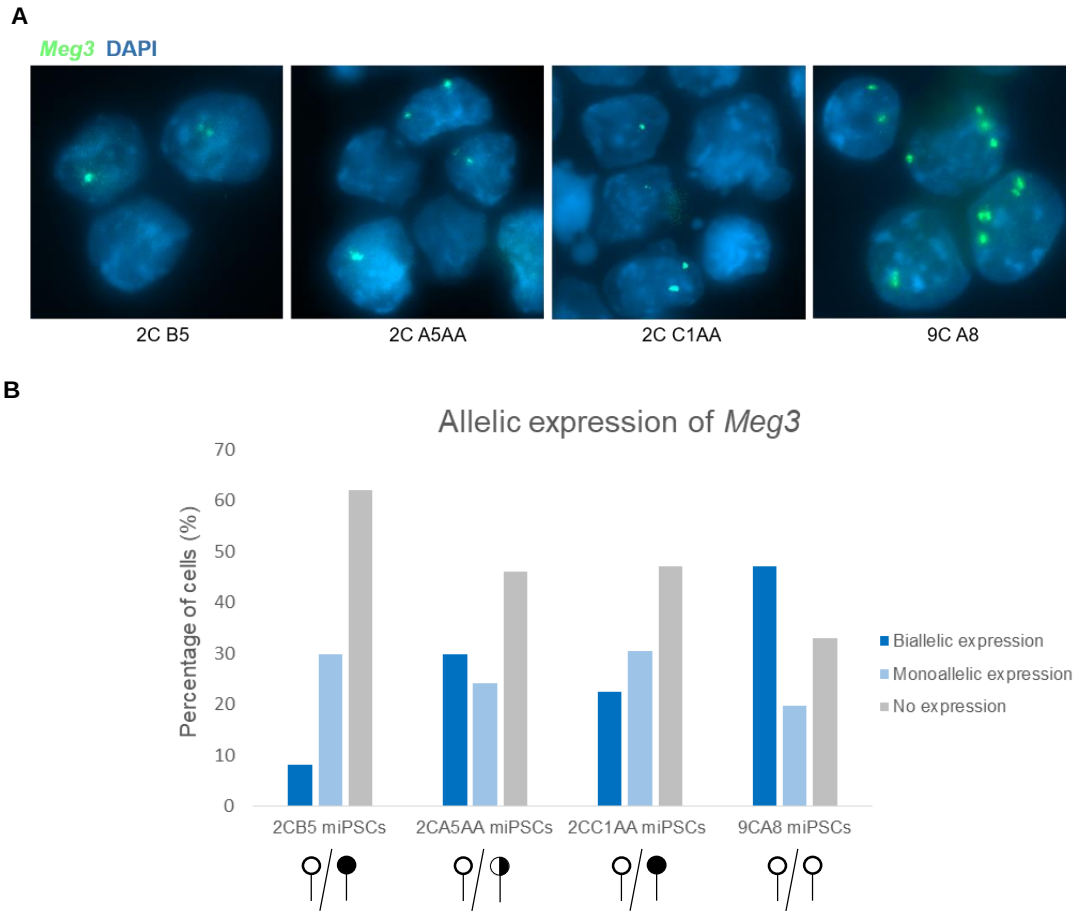

Figure 3.14 - RNA FISH analysis of *Meg3* expression in 2C B5, 2C A5AA, 2C C1 AA and 9C A8 miPSCs lines. **(A)** RNA-FISH analysis of *Meg3* (green dot). Nuclei were counterstained with DAPI (blue). **(B)** Quantification of the percentage of cells with two dots, one dot or no dot for *Meg3* probe. The symbols down each line correspond to the COBRA result (white lollipop – unmethylated *IG*-DMR; half black lollipop – partially methylated *IG*-DMR; black lollipop – methylated *IG*-DMR). A total >105 of cells was counted per sample. Data are from only one independent experiment.

In this analysis, we observed that the percentage of cells presenting two RNA FISH signals for *Meg3* correlate with the degree of hypomethylation of the paternal ICR of this imprinted region. Indeed, 2C B5 and 2C C1AA, which displays correct methylation profile based on COBRA (**Figure 3.14 (B)**) have respectively 8 % and 20 % of cells with biallelic *Meg3* signal. This increased to around 30 % in the 2C A5AA, displaying a trend for hypometyhlation and even more for the 9C A8 line, with only unmethylated alleles, with over 50 % of cells with two *Meg3* dots. Therefore, there was a correlation between the degree of hypomethylation with the amount of cells displaying *Meg3* biallelic signal, which demonstrates, like in the case of *Peg3* gene, that alterations at the *Dlk1-Dio3* locus also have direct consequences for the imprinting regulation of the genes in this locus. In fact, these results seem to suggest that

hypomethylation at the paternal *IG*-DMR, the ICR in this cluster, leads to loss of imprinting at the *Dlk1-Dio3* locus, resulting in biallelic expression of *Meg3*.

The results from both *Peg3* and *Dlk1-Dio3* loci showed their sensibility to demethylation defects in miPSCs, which resulted in imprinting instability, with imprinted genes expressed in a wrong dosage. Therefore, it is important to correct imprinting in iPSCs so that they can be safely used for their many downstream applications such as disease modelling or even regenerative medicine. Also, we do know that miPSCs with imprinting defects cause decrease of pluripotency of miPSCs (Stadtfeld et al., 2010). Since we identified trend for hypomethylation in all the imprinted loci as a major deficiency in miPSCs, we then started wondering on the causes driving this problem.

3.4. Lack of expression of protective proteins of methylation imprints at the original donor cells

From the results of imprinting methylation assessment obtained by COBRA, was already evident that imprinting defects are very frequent in MEF-derived miPSCs. Moreover, a clear trend toward hypomethylation from the originally methylated parental allele was evident. Given such an impressive effect, which we were not initially expected, we decided to start investigating the reasons behind this problem. We reasoned that methylation imprints could be prone to the decrease in overall levels of DNA methylation that accompanies reprogramming (Milagre et al., 2017; Pasque et al., 2018). Thus, we decided to investigate the expression profile of proteins involved in imprinting protection during early development, such as *DPPA3*, *ZFP57* and *TRIM28* (co-factor of *ZFP57*) during reprogramming. For this we used the lines 2C B5, 2C C1AA, 9C A8 and 9C E2AA of miPSCs, the parental MEFs and also the TX1072 mESC line (**Figure 3.15**).

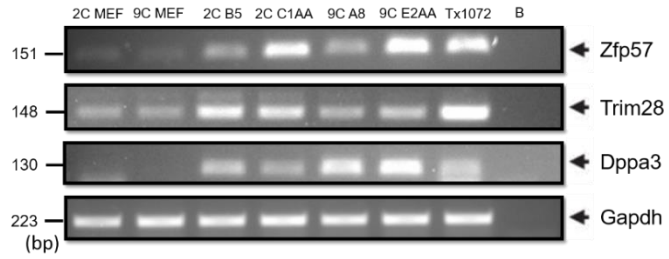

Figure 3.15 - Expression profile of imprinting protective proteins ZFP57, TRIM28 and DPPA3 in miPSCs lines, parental MEFs and Tx1072 mESC line by RT-PCR.

We observed that all these proteins showed low expression levels in the initial MEFs (both 2C and 9C) and they become higher in the final iPSCs. This was clearer for the *Zfp57* and *Dppa3* genes and, but no so clear, for the *Trim28*. Since imprints are well established and maintained in somatic cells, there is no need for the presence of these proteins in the MEFs. However, their absence at the early stages of the reprogramming process, which is artificially induced in these cells, might render methylation imprints unprotected from the overall decrease in DNA methylation accompanying this process. In the future will be necessary to determine exactly, when these proteins regain the levels seen in miPSCs. In any case, we postulate that ectopic expression of these kind of proteins might be an interesting future strategy to resolve the problems with imprinting defects which harness iPSC potential.

4. Conclusions and Future Perspectives

Using our implemented system of controlled reprogramming through a DOX-inducible Yamanaka cassette inserted at a single site in all donor cells and with distinction of parental alleles based on SNPs, we managed to study imprinted defects in both female and male miPSCs. These miPSCs expressed stem cell markers and were able to generate teratomas, which are properties of bona-fide iPSCs.

In this study, we started by looking at particular imprinted clusters such as *Peg3*, *Igf2-H19*, *Kcnq1-Kcnq1ot1*, *Dlk1-Dio3* and *PWS/AS,* for which imprinted defects have already been reported (Sun et al., 2012; Takikawa et al., 2013). Simply by looking at these imprinted regions by COBRA, which is a technique that allows visualization of only one or two CpGs at most, without distinguishing the two parental alleles, we found a high tendency for hypomethylation defects in all clusters analysed. This is consistent with previous findings, which addressed imprinted defects in miPSCs (Sun et al., 2012; Takikawa et al., 2013). In addition, hypomethylation defects in two independent imprinted loci resulted in loss of imprinting of certain genes. This is the case of both *Peg3* and *Meg3* genes that lost their normal paternal and maternal allele-specific expression pattern, respectively, becoming biallelicaly expressed.

Besides the 5 diagnostic imprinted loci we have studied mainly by COBRA, the mouse genome harbours at least 23 distinct imprinted regions, which from the results we obtained so far are unlikely to be exempt of imprinting defects. To complete the full analysis of imprinted defects in all imprinted regions, we initiated a novel technique called Implicon in collaboration with Dr. M. Eckersley-Maslin (Prof. W. Reik's lab, Babraham Institute, Cambridge). This novel approach combines targeted bisulfite treatment of the 23 imprinted regions followed by Illumina amplicon sequencing to quantify methylation levels with a 1000-fold coverage (Masser, Stanford, & Freeman, 2015). This ultra-deep sequencing approach combined with our biological system that discriminates the two parental alleles, allows the identification of any methylation imprinting errors even when present in a minority of cells. Although, we have initiated the Implicon protocol within the scope of my Master thesis, the sequencing results unfortunately did not arrive before the conclusion of this thesis. However, we expect indeed the confirmation of the hypomethylation phenotype across all the loci, as we have seen using COBRA.

Although imprinted defects were evident in both male and female miPSCs, the frequency of defects was higher in male than female iPSCs, which was surprising given the fact that female iPSCs are known to undergo a more accentuated global hypomethylation during reprogramming (Milagre paper). Indeed, Milagre et al. showed that female cells undergo a major global hypomethylation during reprogramming, but reaching at the end the same high methylation state as male iPSCs (Milagre et al., 2017). Later on Pasque et al. complemented this finding demonstrating that the global hypomethylation levels are associated with the reactivation of the inactive X-chromosome, which occurs during reprogramming, and that the regain of methylation is associated with loss of one of the two X-chromosomes in female cells (Pasque et al., 2018). Concerning imprinted regions, Pasque also noticed that imprints are depleted upon X-chromosome reactivation, but not re-established upon X-chromosome loss, despite the regain of genome methylation.

Given the fact that we see that hypomethylation defects in imprinting loci are stronger in male than in female iPSCs, we postulate that in our reprogramming system, male reprogramming cells suffer a higher degree of global hypomethylation, possibly more severe than other systems, and closer to the

ones seen in female iPSCs. This could be due to the use of KNOCkout serum instead of normal fetal bovine serum. Like that, in both male and female iPSCs, the great majority of imprinting regions will lose the methylation alleles. However, since female miPSCs lost one X-chromosome, perhaps some regain of methylation will be felt in imprinted loci (**Figure 4.1**), which by COBRA will look as having less imprinted defects. If this is the case, methylation regain will be expected to occur randomly in both parental alleles, since they both have lost initially the majority of methylation imprinting marks. The results from Implicon will be crucial to test this hypothesis as parental-specific information will be disclosed.

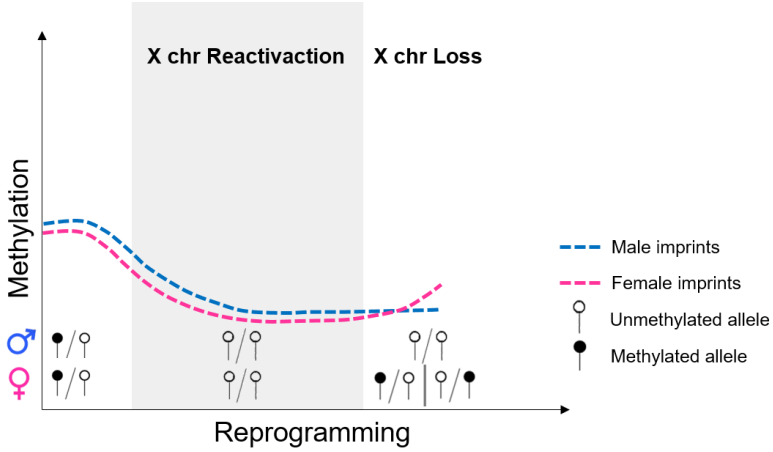

Figure 4.1 - Hypothesis for the less imprinting defects in female miPSCs than in male miPSCs.

Our study feeds on previous studies also documented imprinted defects in miPSCs due to hypomethylation (Sun et al., 2012; Takikawa et al., 2013), that need to be corrected for the safe use of iPSCs in disease modelling and personalized regenerative medicine. These results suggest that methylation imprints are not protected from the global wave of DNA demethylation upon reprogramming. Interestingly, we found out that several imprinting protective proteins such as ZFP57, DPPA3 and TRIM 28 crucial for methylation imprinting maintenance during preimplantation are lowly expressed in the donor cell and perhaps in the early stages of reprogramming. Therefore, we postulate that insufficient expression of these proteins may be the cause of the occurrence of so many imprinting defects due to hypomethylation. In order to eliminate imprinting defects in miPSCs, we propose to improve the reprogramming process through the ectopic expression of these protective proteins. To do so our next step is to reprogram MEFs with a simultaneous overexpression of these proteins. We expect that this overexpression during all the reprogramming process will be able to protect imprinted regions from demethylation (**Figure 4.2**).

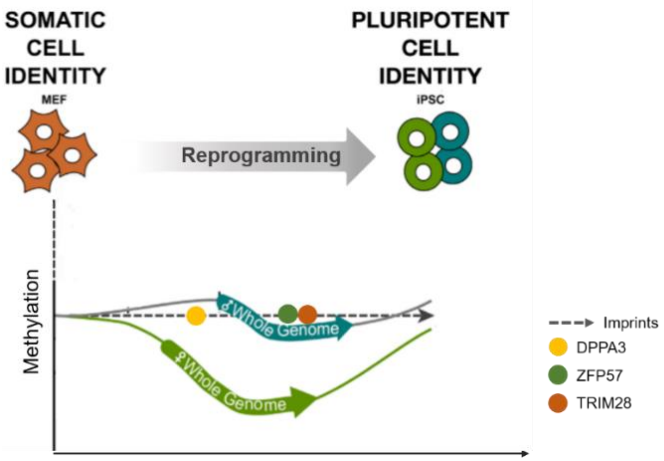

Figure 4.2 - Expected result of overexpression of DPPA3, ZFP57 and TRIM28 during MEFs reprogramming.

5. References

Abad, M., Mosteiro, L., Pantoja, C., Cañamero, M., Rayon, T., Ors, I., … Serrano, M. (2013). Reprogramming in vivo produces teratomas and iPS cells with totipotency features. *Nature*, *502*(7471), 340–345. https://doi.org/10.1038/nature12586

Bacakova, L., Zarubova, J., Travnickova, M., Musilkova, J., Pajorova, J., Slepicka, P., … Molitor, M. (n.d.). Stem cells: their source, potency and use in regenerative therapies with focus on adipose-derived stem cells - a review. *Biotechnology Advances*, *36*(4), 1111–1126. https://doi.org/10.1016/j.biotechadv.2018.03.011

Barlow, D. P., & Bartolomei, M. S. (2014). Genomic Imprinting in Mammals. *Cold Spring Harbor Perspectives in Biology*, *6*(2), a018382–a018382. https://doi.org/10.1101/cshperspect.a018382

Barlow, D. P., Stöger, R., Herrmann, B. G., Saito, K., & Schweifer, N. (1991). The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the Tme locus. *Nature*, *349*(6304), 84–87. https://doi.org/10.1038/349084a0

Bartolomei, M. S., Zemel, S., & Tilghman, S. M. (1991). Parental imprinting of the mouse H19 gene. *Nature*, *351*(6322), 153–155. https://doi.org/10.1038/351153a0

Barton, S. C., Surani, M. A., & Norris, M. L. (n.d.). Role of paternal and maternal genomes in mouse development. *Nature*, *311*(5984), 374–376. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6482961

Bell, A. C., & Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature*, *405*(6785), 482–485. https://doi.org/10.1038/35013100

Bernardes de Jesus, B., Marinho, S. P., Barros, S., Sousa-Franco, A., Alves-Vale, C., Carvalho, T., & Carmo-Fonseca, M. (2018). Silencing of the lncRNA Zeb2-NAT facilitates reprogramming of aged fibroblasts and safeguards stem cell pluripotency. *Nature Communications*, *9*(1), 94. https://doi.org/10.1038/s41467-017-01921-6

Brambrink, T., Foreman, R., Welstead, G. G., Lengner, C. J., Wernig, M., Suh, H., & Jaenisch, R. (2008). Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell*, *2*(2), 151–159. https://doi.org/10.1016/j.stem.2008.01.004

Carey, B. W., Markoulaki, S., Hanna, J. H., Faddah, D. A., Buganim, Y., Kim, J., … Jaenisch, R. (2011). Reprogramming Factor Stoichiometry Influences the Epigenetic State and Biological Properties of Induced Pluripotent Stem Cells. *Cell Stem Cell*, *9*(6), 588–598. https://doi.org/10.1016/j.stem.2011.11.003

Chang, C.-W., Lai, Y.-S., Pawlik, K. M., Liu, K., Sun, C.-W., Li, C., … Townes, T. M. (2009). Polycistronic Lentiviral Vector for "Hit and Run" Reprogramming of Adult Skin Fibroblasts to Induced Pluripotent Stem Cells. *STEM CELLS*, *27*(5), 1042–1049. https://doi.org/10.1002/stem.39

Chaumeil, J., Augui, S., Chow, J. C., & Heard, E. (2008). Combined Immunofluorescence, RNA Fluorescent In Situ Hybridization, and DNA Fluorescent In Situ Hybridization to Study Chromatin Changes, Transcriptional Activity, Nuclear Organization, and X-Chromosome Inactivation (pp. 297–308). https://doi.org/10.1007/978-1-59745-406-3_18

Choi, J., Clement, K., Huebner, A. J., Webster, J., Rose, C. M., Brumbaugh, J., … Hochedlinger, K. (2017a). DUSP9 Modulates DNA Hypomethylation in Female Mouse Pluripotent Stem Cells. *Cell Stem Cell*, *20*(5), 706–719.e7. https://doi.org/10.1016/j.stem.2017.03.002

Choi, J., Huebner, A. J., Clement, K., Walsh, R. M., Savol, A., Lin, K., … Hochedlinger, K. (2017b). Prolonged Mek1/2 suppression impairs the developmental potential of embryonic stem cells.

Nature. https://doi.org/10.1038/nature23274

Choufani, S., Shuman, C., & Weksberg, R. (2010). Beckwith-Wiedemann syndrome. *American Journal of Medical Genetics. Part C, Seminars in Medical Genetics*, *154C*(3), 343–354. https://doi.org/10.1002/ajmg.c.30267

Cooper, D. W., VandeBerg, J. L., Sharman, G. B., & Poole, W. E. (1971). Phosphoglycerate kinase polymorphism in kangaroos provides further evidence for paternal X inactivation. *Nature: New Biology*, *230*(13), 155–157. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5279474

Crouse, H. V, Brown, A., & Mumford, B. C. (1971). --chromosome inheritance and the porblem of chromosome "imprinting" in Sciara (Sciaridae, Diptera). *Chromosoma*, *34*(3), 324–39 8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5112135

DeChiara, T. M., Robertson, E. J., & Efstratiadis, A. (1991). Parental imprinting of the mouse insulin-like growth factor II gene. *Cell*, *64*(4), 849–859. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1997210

Edwards, C. A., & Ferguson-Smith, A. C. (2007). Mechanisms regulating imprinted genes in clusters. *Current Opinion in Cell Biology*, *19*(3), 281–289. https://doi.org/10.1016/j.ceb.2007.04.013

Elhamamsy, A. R. (2017). Role of DNA methylation in imprinting disorders: an updated review. *Journal of Assisted Reproduction and Genetics*, *34*(5), 549–562. https://doi.org/10.1007/s10815-017-0895-5

Gendrel, A.-V., & Heard, E. (2011). Fifty years of X-inactivation research. *Development*, *138*(23), 5049–5055. https://doi.org/10.1242/dev.068320

Godini, R., Lafta, H. Y., & Fallahi, H. (2018). Epigenetic modifications in the embryonic and induced pluripotent stem cells. *Gene Expression Patterns*, *29*, 1–9. https://doi.org/10.1016/j.gep.2018.04.001

Gossen, M., & Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(12), 5547–5551. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1319065

Green, K., Lewis, A., Dawson, C., Dean, W., Reinhart, B., Chaillet, J. R., & Reik, W. (2007). A developmental window of opportunity for imprinted gene silencing mediated by DNA methylation and the Kcnq1ot1 noncoding RNA. *Mammalian Genome*, *18*(1), 32–42. https://doi.org/10.1007/s00335-006-0092-9

Greenberg, M. V. C., & Bourc'his, D. (2015). Cultural relativism: Maintenance of genomic imprints in pluripotent stem cell culture systems. *Current Opinion in Genetics and Development*, *31*, 42–49. https://doi.org/10.1016/j.gde.2015.04.005

GURDON, J. B. (1962). Adult frogs derived from the nuclei of single somatic cells. *Developmental Biology*, *4*, 256–273. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/13903027

Hannum, G., Guinney, J., Zhao, L., Zhang, L., Hughes, G., Sadda, S., … Zhang, K. (2013). Genome-wide Methylation Profiles Reveal Quantitative Views of Human Aging Rates. *Molecular Cell*, *49*(2), 359–367. https://doi.org/10.1016/j.molcel.2012.10.016

Hata, K., Okano, M., Lei, H., & Li, E. (2002). Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development (Cambridge, England)*, *129*(8), 1983–1993. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11934864

He, H., & Kim, J. (2014). Regulation and Function of the Peg3 Imprinted Domain. *Genomics & Informatics*, *12*(3), 105. https://doi.org/10.5808/GI.2014.12.3.105

Heo, J., Lim, J., Lee, S., Jeong, J., Kang, H., Kim, Y., … Shin, D.-M. (2017). Sirt1 Regulates DNA Methylation and Differentiation Potential of Embryonic Stem Cells by Antagonizing Dnmt3l. *Cell Reports*, *18*(8), 1930–1945. https://doi.org/10.1016/j.celrep.2017.01.074

Ideraabdullah, F. Y., Vigneau, S., & Bartolomei, M. S. (2008). Genomic imprinting mechanisms in mammals. *Mutation Research*, *647*(1–2), 77–85. https://doi.org/10.1016/j.mrfmmm.2008.08.008

Jaenisch, R., & Young, R. (2008). Stem Cells, the Molecular Circuitry of Pluripotency and Nuclear Reprogramming. *Cell*, *132*(4), 567–582. https://doi.org/10.1016/j.cell.2008.01.015

Kaji, K., Norrby, K., Paca, A., Mileikovsky, M., Mohseni, P., & Woltjen, K. (2009). Virus-free induction of pluripotency and subsequent excision of reprogramming factors. *Nature*, *458*(7239), 771–775. https://doi.org/10.1038/nature07864

Kalish, J. M., Jiang, C., & Bartolomei, M. S. (2014). Epigenetics and imprinting in human disease. *The International Journal of Developmental Biology*, *58*(2-3–4), 291–298. https://doi.org/10.1387/ijdb.140077mb

Kanduri, C. (2016). Long noncoding RNAs: Lessons from genomic imprinting. *Biochimica et Biophysica Acta*, *1859*(1), 102–111. https://doi.org/10.1016/j.bbagrm.2015.05.006

Kramer, J., Steinhoff, J., Klinger, M., Fricke, L., & Rohwedel, J. (2006). Cells differentiated from mouse embryonic stem cells via embryoid bodies express renal marker molecules. *Differentiation*, *74*(2–3), 91–104. https://doi.org/10.1111/j.1432-0436.2006.00062.x

Kurosawa, H. (2007). Methods for inducing embryoid body formation: in vitro differentiation system of embryonic stem cells. *Journal of Bioscience and Bioengineering*, *103*(5), 389–398. https://doi.org/10.1263/jbb.103.389

Li, E., Beard, C., & Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. *Nature*, *366*(6453), 362–365. https://doi.org/10.1038/366362a0

Li, E., Bestor, T. H., & Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell*, *69*(6), 915–926. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1606615

Li, E., & Zhang, Y. (2014). DNA Methylation in Mammals. *Cold Spring Harbor Perspectives in Biology*, *6*(5), a019133–a019133. https://doi.org/10.1101/cshperspect.a019133

Li, H., Collado, M., Villasante, A., Strati, K., Ortega, S., Cañamero, M., … Serrano, M. (2009). The Ink4/Arf locus is a barrier for iPS cell reprogramming. *Nature*, *460*(7259), 1136–1139. https://doi.org/10.1038/nature08290

Li, X., Ito, M., Zhou, F., Youngson, N., Zuo, X., Leder, P., & Ferguson-Smith, A. C. (2008). A Maternal-Zygotic Effect Gene, Zfp57, Maintains Both Maternal and Paternal Imprints. *Developmental Cell*, *15*(4), 547–557. https://doi.org/10.1016/j.devcel.2008.08.014

Lo Sardo, V., Ferguson, W., Erikson, G. A., Topol, E. J., Baldwin, K. K., & Torkamani, A. (2016). Influence of donor age on induced pluripotent stem cells. *Nature Biotechnology*, *35*(1), 69–74. https://doi.org/10.1038/nbt.3749

Luo, J., Zhang, Y., Guo, Y., Tang, H., Wei, H., Liu, S., … Zhou, P. (2017). TRIM28 regulates Igf2-H19 and Dlk1-Gtl2 imprinting by distinct mechanisms during sheep fibroblast proliferation. *Gene*, *637*, 152–160. https://doi.org/10.1016/j.gene.2017.09.048

Ma, H., Morey, R., O'Neil, R. C., He, Y., Daughtry, B., Schultz, M. D., … Mitalipov, S. (2014). Abnormalities in human pluripotent cells due to reprogramming mechanisms. *Nature*,

511(7508), 177–183. https://doi.org/10.1038/nature13551

Mackay, D. J. G., Callaway, J. L. A., Marks, S. M., White, H. E., Acerini, C. L., Boonen, S. E., … Temple, I. K. (2008). Hypomethylation of multiple imprinted loci in individuals with transient neonatal diabetes is associated with mutations in ZFP57. *Nature Genetics*, *40*(8), 949–951. https://doi.org/10.1038/ng.187

Maherali, N., & Hochedlinger, K. (2008). Guidelines and Techniques for the Generation of Induced Pluripotent Stem Cells. *Cell Stem Cell*, *3*(6), 595–605. https://doi.org/10.1016/j.stem.2008.11.008

Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., … Hochedlinger, K. (2007). Directly Reprogrammed Fibroblasts Show Global Epigenetic Remodeling and Widespread Tissue Contribution. *Cell Stem Cell*, *1*(1), 55–70. https://doi.org/10.1016/j.stem.2007.05.014

Mahmoudi, S., & Brunet, A. (2012). Aging and reprogramming: a two-way street. *Current Opinion in Cell Biology*, *24*(6), 744–756. https://doi.org/10.1016/j.ceb.2012.10.004

Martin, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, *78*(12), 7634–7638. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6950406

Mascetti, V. L., & Pedersen, R. A. (2016). Contributions of Mammalian Chimeras to Pluripotent Stem Cell Research. *Cell Stem Cell*, *19*(2), 163–175. https://doi.org/10.1016/j.stem.2016.07.018

Masser, D. R., Stanford, D. R., & Freeman, W. M. (2015). Targeted DNA Methylation Analysis by Next-generation Sequencing. *Journal of Visualized Experiments*, (96). https://doi.org/10.3791/52488

McGrath, J., & Solter, D. (1984). Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell*, *37*(1), 179–183. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6722870

Milagre, I., Stubbs, T. M., King, M. R., Spindel, J., Santos, F., Krueger, F., … Reik, W. (2017). Gender Differences in Global but Not Targeted Demethylation in iPSC Reprogramming. *Cell Reports*, *18*(5), 1079–1089. https://doi.org/10.1016/j.celrep.2017.01.008

Nazor, K. L., Altun, G., Lynch, C., Tran, H., Harness, J. V., Slavin, I., … Laurent, L. C. (2012). Recurrent Variations in DNA Methylation in Human Pluripotent Stem Cells and Their Differentiated Derivatives. *Cell Stem Cell*, *10*(5), 620–634. https://doi.org/10.1016/j.stem.2012.02.013

Nordin, N., Lai, M. I., Veerakumarasivam, a, & Ramasamy, R. (2011). Induced Pluripotent Stem Cells : History , Properties and. *Medical Journal Of Malaysia*, *66*(1), 4–9.

Ogata, T., & Kagami, M. (2016). Kagami–Ogata syndrome: a clinically recognizable upd(14)pat and related disorder affecting the chromosome 14q32.2 imprinted region. *Journal of Human Genetics*, *61*(2), 87–94. https://doi.org/10.1038/jhg.2015.113

Okano, M., Bell, D. W., Haber, D. A., & Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*, *99*(3), 247–257. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10555141

Okita, K., Ichisaka, T., & Yamanaka, S. (2007). Generation of germline-competent induced pluripotent stem cells. *Nature*, *448*(7151), 313–317. https://doi.org/10.1038/nature05934

Omole, A. E., & Fakoya, A. O. J. (2018). Ten years of progress and promise of induced pluripotent stem cells: historical origins, characteristics, mechanisms, limitations, and potential applications. *PeerJ*, *6*, e4370. https://doi.org/10.7717/peerj.4370

Pasque, V., Karnik, R., Chronis, C., Petrella, P., Langerman, J., Bonora, G., … Plath, K. (2018). X Chromosome Dosage Influences DNA Methylation Dynamics during Reprogramming to Mouse iPSCs. *Stem Cell Reports*, *10*(5), 1537–1550. https://doi.org/10.1016/j.stemcr.2018.03.019

Pasque, V., Tchieu, J., Karnik, R., Uyeda, M., Sadhu Dimashkie, A., Case, D., … Plath, K. (2014). X Chromosome Reactivation Dynamics Reveal Stages of Reprogramming to Pluripotency. *Cell*, *159*(7), 1681–1697. https://doi.org/10.1016/j.cell.2014.11.040

Pastor, W. A., Aravind, L., & Rao, A. (2013). TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nature Reviews Molecular Cell Biology*, *14*(6), 341–356. https://doi.org/10.1038/nrm3589

Plasschaert, R. N., & Bartolomei, M. S. (2014). Genomic imprinting in development, growth, behavior and stem cells. *Development (Cambridge, England)*, *141*(9), 1805–1813. https://doi.org/10.1242/dev.101428

Pólvora Brandão, D., & da Rocha, S. T. (2018). Genomic Imprinting at the Transcriptional Level. In *eLS* (pp. 1–11). Chichester, UK: John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470015902.a0005686.pub3

Przyborski, S. A. (2005). Differentiation of human embryonic stem cells after transplantation in immune-deficient mice. *Stem Cells (Dayton, Ohio)*, *23*(9), 1242–1250. https://doi.org/10.1634/stemcells.2005-0014

Reik, W. (2001). Epigenetic Reprogramming in Mammalian Development. *Science*, *293*(5532), 1089–1093. https://doi.org/10.1126/science.1063443

Rocha, S˜o T. da, & Heard, E. (2001). Genomic Imprinting at the Transcriptional Level. *Encyclopedia of Life Sciences*, (April), 1–13. https://doi.org/DOI: 10.1002/9780470015902.a0005686.pub2

Rocha, S. T. da, Edwards, C. A., Ito, M., Ogata, T., & Ferguson-Smith, A. C. (2008). Genomic imprinting at the mammalian Dlk1-Dio3 domain. *Trends in Genetics*, *24*(6), 306–316. https://doi.org/10.1016/j.tig.2008.03.011

Schlaeger, T. M., Daheron, L., Brickler, T. R., Entwisle, S., Chan, K., Cianci, A., … Daley, G. Q. (2015). A comparison of non-integrating reprogramming methods. *Nature Biotechnology*, *33*(1), 58–63. https://doi.org/10.1038/nbt.3070

Schulz, E. G., Meisig, J., Nakamura, T., Okamoto, I., Sieber, A., Picard, C., … Heard, E. (2014). The Two Active X Chromosomes in Female ESCs Block Exit from the Pluripotent State by Modulating the ESC Signaling Network. *Cell Stem Cell*, *14*(2), 203–216. https://doi.org/10.1016/j.stem.2013.11.022

Searle, A. G., & Beechey, C. V. (1978). Complementation studies with mouse translocations. *Cytogenetic and Genome Research*, *20*(1–6), 282–303. https://doi.org/10.1159/000130859

Seisenberger, S., Peat, J. R., Hore, T. A., Santos, F., Dean, W., & Reik, W. (2012). Reprogramming DNA methylation in the mammalian life cycle: building and breaking epigenetic barriers. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*(1609), 20110330–20110330. https://doi.org/10.1098/rstb.2011.0330

Slack, J. M. W. (2018). *The Science of Stem Cells*. https://doi.org/10.1002/9781119235293

Sommer, C. A., Stadtfeld, M., Murphy, G. J., Hochedlinger, K., Kotton, D. N., & Mostoslavsky, G. (2009). Induced Pluripotent Stem Cell Generation Using a Single Lentiviral Stem Cell Cassette. *Stem Cells*, *27*(3), 543–549. https://doi.org/10.1634/stemcells.2008-1075

Stadtfeld, M., Apostolou, E., Akutsu, H., Fukuda, A., Follett, P., Natesan, S., … Hochedlinger, K. (2010).

Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature*, *465*(7295), 175–181. https://doi.org/10.1038/nature09017

Stadtfeld, M., Apostolou, E., Ferrari, F., Choi, J., Walsh, R. M., Chen, T., … Hochedlinger, K. (2012). Ascorbic acid prevents loss of Dlk1-Dio3 imprinting and facilitates generation of all–iPS cell mice from terminally differentiated B cells. *Nature Genetics*, *44*(4), 398–405. https://doi.org/10.1038/ng.1110

Stadtfeld, M., Maherali, N., Breault, D. T., & Hochedlinger, K. (2008). Defining Molecular Cornerstones during Fibroblast to iPS Cell Reprogramming in Mouse. *Cell Stem Cell*, *2*(3), 230–240. https://doi.org/10.1016/j.stem.2008.02.001

Strogantsev, R., Krueger, F., Yamazawa, K., Shi, H., Gould, P., Goldman-Roberts, M., … Ferguson-Smith, A. C. (2015). Allele-specific binding of ZFP57 in the epigenetic regulation of imprinted and non-imprinted monoallelic expression. *Genome Biology*, *16*(1), 112. https://doi.org/10.1186/s13059-015-0672-7

Sun, B., Ito, M., Mendjan, S., Ito, Y., Brons, I. G. M., Murrell, A., … Pedersen, R. A. (2012). Status of Genomic Imprinting in Epigenetically Distinct Pluripotent Stem Cells. *STEM CELLS*, *30*(2), 161–168. https://doi.org/10.1002/stem.793

Tada, M., Takahama, Y., Abe, K., Nakatsuji, N., & Tada, T. (2001). Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. *Current Biology : CB*, *11*(19), 1553–1558. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11591326

Takagi, N., & Sasaki, M. (1975). Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature*, *256*(5519), 640–642. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1152998

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., & Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell*, *131*(5), 861–872. https://doi.org/10.1016/j.cell.2007.11.019

Takahashi, K., & Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*, *126*(4), 663–676. https://doi.org/10.1016/j.cell.2006.07.024

Takikawa, S., Ray, C., Wang, X., Shamis, Y., Wu, T.-Y., & Li, X. (2013). Genomic imprinting is variably lost during reprogramming of mouse iPS cells. *Stem Cell Research*, *11*(2), 861–873. https://doi.org/10.1016/j.scr.2013.05.011

Thomson, J. A. (1998). Embryonic Stem Cell Lines Derived from Human Blastocysts. *Science*, *282*(5391), 1145–1147. https://doi.org/10.1126/science.282.5391.1145

Watanabe, S., Hirai, H., Asakura, Y., Tastad, C., Verma, M., Keller, C., … Asakura, A. (2011). MyoD Gene Suppression by Oct4 Is Required for Reprogramming in Myoblasts to Produce Induced Pluripotent Stem Cells. *STEM CELLS*, *29*(3), 505–516. https://doi.org/10.1002/stem.598

Xu, X., Smorag, L., Nakamura, T., Kimura, T., Dressel, R., Fitzner, A., … Krishna Pantakani, D. V. (2015). Dppa3 expression is critical for generation of fully reprogrammed iPS cells and maintenance of Dlk1-Dio3 imprinting. *Nature Communications*, *6*(1), 6008. https://doi.org/10.1038/ncomms7008

Yu, J., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J. L., Tian, S., … Thomson, J. A. (2007). Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells. *Science*, *318*(5858), 1917–1920. https://doi.org/10.1126/science.1151526