Stage-specific germ cell marker genes function in establishment and germ cell lineage commitment of pluripotent stem cells

DISSERTATION

zur Erlangung des

Mathematisch-Naturwissenschaftlichen Doktorgrades

„Doctor rerum naturalium“

an der Georg-August-Universität Göttingen

vorgelegt von

Xingbo Xu

aus

Nei Mongol, China

Göttingen, 2012

D7

Referent: Prof. Dr. med. Dr. h. c. Wolfgang Engel

Korreferent: Prof. Dr. Sigrid Hoyer-Fender

Tag der mündlichen Prüfung: 20.09.2012

Dedicated to my family

All I am, or can be, I owe to my supportive parents

There is no royal road to science, and only those who do not dread the fatiguing climb of its steep paths have a chance of gaining its luminous summits

- Karl Marx

Table of contents

Table of contents i
List of figures and tables iii
1. Zusammenfassung 1
1. Summary 3
2. Introduction 5
2.1. Mouse germ cell development 5
2.2. Developmental roles of selected stage-specific germ cell marker genes 7
2.2.1. *Blimp1* 7
2.2.2. *Ifitm3* 7
2.2.3. *Dppa3* 7
2.2.4. *Mvh* 8
2.2.5. *Dazl* 8
2.3. Pluripotent stem cells and their origins 8
2.3.1. Induced pluripotent stem cells (iPSCs) 10
2.4. Objectives 11
3. Results 12
3.1. Stage-specific germ cell marker genes are expressed in all pluripotent cell types and emerges early during induced pluripotency 13
3.2. Dppa3 binds to the IG-DMR of the *Dlk1-Dio3* imprinting cluster and prevents its imprinting loss during iPS cell generation 25
3.3. Mouse Dazl and its novel splice variant function in translational repression of target mRNAs in embryonic stem cells 47
4. Discussion 70
4.1. Spermatogenesis reconstituted from pluripotent stem cells *in vitro* 70
4.2. Stage-specific germ cell marker genes are expressed in all mouse pluripotent cell types and emerge early during induced pluripotency 72
4.2.1. Expansion of PGC fate model to human somatic cell reprogramming 73
4.2.2. Transdifferentiation of somatic cells into induced germline stem cells 75
4.3. Dppa3 binds to the IG-DMR of the *Dlk1-Dio3* imprinting cluster and prevents its imprinting loss during iPS cell generation 78
4.3.1. Functional analysis of Dppa3 during somatic cell reprogramming by inducible transgenic mouse model 80

4.4. Mouse Dazl and its novel splice variant functions in translational repression of target mRNAs in mouse embryonic stem cells 81
4.5. Future endeavors and perspectives 84
5. References 85
6. Abbreviations 94
7. Acknowledgments 97
8. Curriculum vitae 98
9. List of publications 99
10. List of manuscripts in submission stage 99

List of figures and tables

Figure 2.1. A schematic illustration of mouse primordial germ cell specification and migration during embryonic development 6
Figure 2.2. Developmental origins of pluripotent stem cells 9
Figure 2.3. The derivation and applications of iPS cells 10
Figure 4.1. Expression pattern of GC/PrM- and pluripotency-marker genes during the course of hiPS generation 74
Figure 4.2. Schematic diagram depicting the generation of iGSCs 75
Figure 4.3. Generation and characterization of iGSC. 76
Figure 4.4. Characterization of iGSCs by immunocytochemistry 77
Figure 4.5. Schematic representation of Tet-on 3G system and Dppa3 transgenic mouse model generation 80
Figure 4.6. Schematic representation of induced Dppa3 expression during somatic cell reprogramming 81
Figure 4.7. Dazl-Tat recombinant protein purification and functional test 82

Table 4.1 Pathological anomalies in human harbouring genetic defects at the *Dlk1-Dio3* locus 79

1. Zusammenfassung

Keimzellen transferieren genetische und epigenetische Informationen von Generation zu Generation. Ihre Eigenschaften haben Forscher angespornt, die molekularen Eigenschaften, die diese Besonderheit regulieren oder aufrechterhalten, zu untersuchen. Die jüngsten Entwicklungen haben gezeigt, dass eine Umwandlung von Stammzellen der Keimbahn zu pluripotenten Stammzellen möglich ist, was eine reziproke Beziehung dieser Zellen impliziert. Übereinstimmend mit diesen Ergebnissen haben mehrere Studien die Expression einiger Keimzellmarker (germ stem cell marker, GC marker) sowie prämeiotischer Marker (premeiotic marker, PrM marker) in verschiedenen pluripotenten Zelltypen aufgezeigt. Die Funktion und Relevanz der Expression dieser Markergene für die Etablierung sowie Erhaltung der Pluripotenz sind jedoch weitestgehend unbekannt. Im Rahmen dieser Studie sollte die Rolle von GC- und PrM Markergenen bei der Etablierung und Erhaltung sowohl der Pluripotenz als auch der Differenzierung aufgeklärt werden.

Im ersten Teil der vorliegenden Arbeit konnten wir zeigen, dass ausgewählte GC- und PrM Markergene in allen analysierten pluripotenten Stammzellen exprimiert werden, was auf einen gemeinsamen Keimzellursprung von pluripotenten Stammzellen schließen lässt. Weiterhin werden GC Markergene vor der Aktivierung von Pluripotenzgenen während der Reprogrammierung somatischer Zellen zu induzierten, pluripotenten Zellen (iPS-Zellen) aktiviert. Zusammen mit den vorliegenden molekularen Beweisen für die *in vivo* Keimzellspezifizierung lassen diese Ergebnisse vermuten, dass während der frühen Phasen der Reprogrammierung somatischer Zellen eine temporäre Keimzellorientierung existiert.

Im zweiten Teil dieser Arbeit haben wir mittels Reprogrammierungsversuchen mit somatischen Zellen die Funktion ausgewählter GC Markergene bei der Etablierung und Erhaltung der authentischen Pluripotenz untersucht. Kürzlich wurde gezeigt, dass fehlerhaftes Imprinting am *Dlk1-Dio3* Lokus hauptsächlich während der Etablierung von induzierten pluripotenten Stammzellen (iPSCs) auftritt. Interessanterweise wiesen alle iPS Zelllinien, die mit dem *Dppa3* als ein GC Markergen generiert wurden, ein korrektes Imprintingmuster am *Dlk1-Dio3* Lokus auf. Diese Ergebnisse lassen vermuten, dass GC Markergene, insbesonders *Dppa3*, wichtig für die Etablierung authentischer Pluripotenz sind.

Im dritten Teil der vorliegenden Arbeit wurde die Rolle von *Dazl* als ein PrM Markergen in post-transktiptionellen Regulationsmechanismen sowohl in pluripotenten Stammzellen als auch in Keimzellen untersucht. Unsere Ergebnisse deuten daraufhin, dass Dazl die Keimzellspezifizierung befördert und möglicherweise als ein translationeller Repressor fungiert, der die Balance zwischen dem Beitrag zur Keimbahn und zur Pluripotenz aufrechterhält.

Zusammenfassend betonen die Ergebnisse unserer Studien die Bedeutung von GC Markergenen in der Etablierung der Pluripotenz *in vivo* und *in vitro* und bekräftigen die Hypothese eines Keimzellursprungs aller pluripotenten Stammzellen. Somit konnten unsere Arbeiten einen Beitrag zur Analyse der Rolle von GC und PrM Markergenen in pluripotenten Stammzellen geben.

1. Summary

Germ cells are firmly committed to fulfil the mandate of transmitting genetic and epigenetic information from one generation to the next. Their mysterious characteristics have fascinated biological researchers to study the molecular mechanisms which regulate or maintain this speciality. The recent advances have shown the possibility of conversion between germline stem cells and pluripotent stem cells, implicating their reciprocal relationship. In line with these observations, several studies have shown the expression of some germ stem cell (GC) and premeiotic (PrM) marker genes in various pluripotent cell types. However, the function and relevance of this marker genes expression for pluripotency establishment and maintenance is largely unknown. The aim of this thesis was to elucidate the role of GC and PrM marker genes in establishment/maintenance of pluripotency as well as in differentiation.

In the first part of this thesis, we demonstrated the expression of selected GC and PrM marker genes in all analysed pluripotent stem cells, suggesting a common germ cell origin of pluripotent stem cells. Further, the GC marker genes were found to be activated prior to the activation of pluripotency-related genes during somatic cell reprogramming towards induced pluripotency. These results together with the available molecular evidence for *in vivo* germ cell specification led us to suggest a possible existence of a temporary germ cell fate during early stages of somatic cell reprogramming.

In the second part of this study, we examined the function of selected GC marker genes in establishment and maintenance of authentic pluripotency using somatic cell reprogramming studies. Recently, defects in imprinting at the *Dlk-Dio3* locus were shown to occur mostly during establishment of induced pluripotent stem cells (iPSCs). Interestingly, all iPS cell lines generated in the presence of *Dppa3*, a GC marker gene, led to the proper imprinting maintenance at the *Dlk1-Dio3* locus. These results suggest that *Dppa3* is important for the establishment of authentic pluripotency.

In the third part of this thesis, a PrM marker gene, *Dazl* was studied for its role in post-transcriptional regulation mechanisms in pluripotent stem cells as well as in germ cells. Our results suggest that Dazl, a germ cell lineage mediator, might function as a translational repressor to maintain the balance between germ cell lineage commitment and pluripotency.

Collectively, the results of our studies emphasize the importance of GC marker genes in establishment of pluripotency *in vivo* and *in vitro*, and strengthen the hypothesis

of germ cell origin of all pluripotent stem cells. Our studies shed light on the undiscovered role of GC/PrM marker genes in pluripotent cells.

2. Introduction

"Germ cells are the most fascinating cells of all cell types and are still deeply mysterious" (McLaren, 1994). The biological significance of germ cells is to perpetuate the parental genetic information to the next generation. To this end, germ cells undergo extensive genetic and epigenetic reprogramming through complex biological processes such as chromatin condensation, meiosis and morphological transformation into mature gametes. They are specified during embryonic development and can be distinguished from other somatic cells by their unique gene expression profiles. Intriguingly, the expression of germ cell marker genes is not only restricted to germ cells, but are also highly expressed in pluripotent inner cell mass (ICM) and various pluripotent cell types. This may indicate that all pluripotent cells might have the same germ cell origin or probably the expression of germ cell marker genes is indispensable for establishment and maintenance of pluripotency. The goal of this thesis was to investigate the importance of these germ cell marker genes in pluripotent stem cells as well as during the induction of pluripotency in somatic cells. Hence, the role of selected germ cell marker genes during germ cell development will be highlighted along with a brief introduction to origins of pluripotent stem cells.

2.1. Mouse germ cell development

During mouse early embryogenesis, primordial germ cells (PGCs), the founder population of germ cells, are specified at proximal epiblast, undergo further development and finally localize to the genital ridge at midgestation (Ginsburg et al., 1990). After reaching puberty, germ cells are developed into mature haploid spermatozoa or oocytes. Then a new cycle will start again upon fertilization, representing the beginning of a new individual.

PGCs specification is induced by instructive signal, Bone Morphogenetic Protein (BMP), which is secreted from extra embryonic ectoderm prior to gastrulation at embryonic day E6.25. A small cluster of PGC precursors (approximately 8 cells) appear at the proximal epiblast at this stage (Ohinata et al., 2005). During this time, B-lymphocyte-induced maturation protein-1 (*Blimp1*) and Interferon-induced transmembrane protein 3 (*Ifitm3*, also known as *Fragilis*), two typical early germ stem cell marker genes are necessarily expressed and function in differentiation suppression thereby maintaining the PGC congregation (Vincent et al., 2005; Tanaka et al., 2005). At

E7.5, about 50 PGCs are completely separated in the extra- embryonic mesoderm at the base of the allantois (Ginsburg et al., 1990). These cells can be easily distinguished by the expression of Alkaline Phosphatase (*AP*), a surface marker for germ cells as well as pluripotent stem cells, and Developmental pluripotency-associated protein 3 (*Dppa3*), another important germ stem cell marker gene (Saitou et al., 2002). From this region, PGCs migrate and proliferate while passing through primitive streak, embryonic endoderm to reach hind gut with around 250 cells at E9.5 (Tam and Snow 1981). Later on, PGCs change direction and move towards dorsal mesentery and colonize in the genital ridges at E12.5, where about 4000 PGCs start to undergo sex-specific differentiation to form mature gametes (De Felici et al., 1992). The pre-meiotic specific genes Mouse vasa homologue (*Mvh*) and Deleted in azoospermia-like (*Dazl*) start to express at E10.5 and E11.5, respectively. They are believed to have essential roles in fate determination event of the germ cell and in fundamental events of post-migratory germ cell development (Fujiwara et al., 1994; Haston et al., 2009). The process of germ cell specification and development is illustrated in **figure 2.1**.

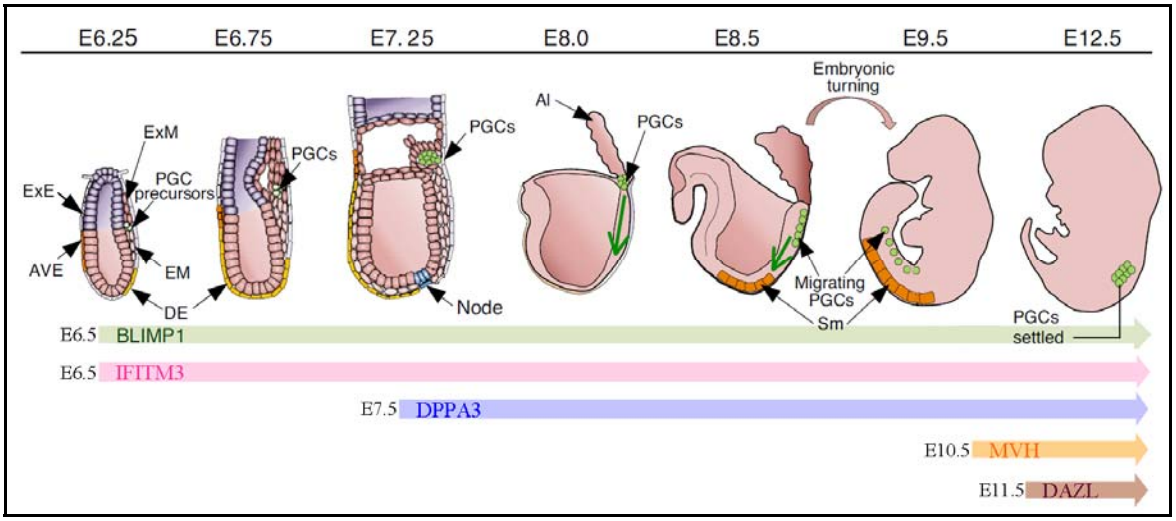


Figure 2.1. A schematic illustration of mouse primordial germ cell specification and migration during embryonic development. Primordial germ cell (PGC) precursors (E6.25) and PGCs are presented as green circles during embryonic development from E6.25 to E12.5, and the PGC migration route is indicated by a green arrow. The timing of germ cell specific marker genes *Blimp1*, *Ifitm3*, *Dppa3*, *Mvh,* and *Dazl* expression is shown below. Al, allantois; AVE, anterior visceral endoderm; DE, distal endoderm; DVE, distal visceral endoderm; EM, embryonic mesoderm; Epi, epiblast; ExE, Extra-embryonic ectoderm; ExM, extra-embryonic mesoderm; PGC, primordial germ cells; VE, visceral endoderm (Figure adapted and modified from: Saitou et al., 2012).

2.2. Developmental roles of selected stage-specific germ cell marker genes

2.2.1. *Blimp1* is expressed as early as in the founder PGCs at E6.25 and functions as a determinant of PGC specification (Ohinata et al., 2005). *Blimp1* encodes a transcriptional repressor protein which contains an N-terminally located positive regulatory domain functioning in chromatin remodeling; a proline-rich region responsible for the recruitment of transcriptional co-repressor and a C-terminally located five krüppel-type zinc finger motifs for DNA binding (Ren et al., 1999; Yu et al., 2000). Strikingly, *Blimp1* knock-out studies showed a complete depletion of PGCs at E7.25 when germ cell specification initiates (Vincent et al., 2005). A recent study reported the successful establishment of embryonic stem cells (ESCs) and epiblast stem cells (EpiSCs) from *Blimp1*-deficient blastocyst and epiblast, respectively (Bao et al., 2012). These results suggest that *Blimp1* expression is only obligatory for the PGCs specification but not for the establishment and maintenance of pluripotency.

2.2.2. *Ifitm3* is a member of interferon-induced transmembrane protein gene family. The encoded protein contains two transmembrane domains and localizes on the cell surface (Lange et al., 2003). Upon the BMP signaling stimulation, the expression of *Ifitm3* is activated as early as in PGC precursors (Lange et al., 2003). It is believed that *Ifitm3* expression is important for the congregation as well as PGCs migration towards the genital ridge (Tanaka et al., 2005; Saitou et al., 2002). Contrary to this hypothesis, genetic depletion of *Ifitm3* does not result in any detectable effect on embryonic development, viability and fertility, possibly due to the existence of compensatory mechanisms (Lange et al., 2008).

2.2.3. *Dppa3* was initially identified as a PGC marker gene expressed at E7.25 embryo, therefore it is also named as *PGC7* (Sato et al., 2002). This gene encodes a protein containing both nuclear localization signal and nuclear export signal allowing it to shuttle between nucleus and cytoplasm (Saiti et al., 2007). Dppa3 is known to harbor a DNA-binding domain responsible for chromosomal organization at the N-terminus of the protein and a splicing factor like motif involved in RNA processing at the C-terminus (Aravind et al., 2000). Experimentally it was also proven that Dppa3 is able to bind both on DNA and RNA (Aravind et al., 2000). *Dppa3* is inherited as a maternal factor since its expression is detected in oocytes and in zygotes, later on the inherited expression is replaced by the onset of zygotic expression, which persists until preimplantation embryo

development (Payer et al., 2003). *Dppa3* expression is silenced until E7.25, there after, is restricted and mark PGCs during migration till their colonization at genital ridges (Sato et al., 2002). Even though *Dppa3* is used as a germ stem cell marker gene, it seems to be not important for PGC emergence (Bortvin et al., 2004; Payer et al., 2003). *Dppa3* knockout studies showed that the first generation mutant mice have no significant defects in germ cells development and differentiation (Bortvin et al., 2004). However, the second generation mutant embryos die at early cleavage stages emphasizing the significance of *Dppa3* as a maternal factor indispensible for pre-implantation embryo development (Bortvin et al., 2004).

2.2.4. *Mvh* encodes an ATP-dependent RNA helicase of the DEAD-box protein family (Hay et al., 1990). In PGCs, the expression of *Mvh* is detectable after PGC colonized in the genital ridge at around E10.5. It is continuously expressed until the sex-specific germ cell differentiation to generate mature gametes (Toyooka et al., 2000). Furthermore, it is also believed that the *Mvh* expression is in response to a germ-soma interaction (Toyooka et al., 2000). Genetic ablation studies revealed that *Mvh*-null male mice are sterile, while females are fertile (Tanaka et al., 2000).

2.2.5. *Dazl* codes for an RNA binding protein that is important for germ cell differentiation in vertebrates (Peng et al., 2009). Dazl protein contains an RNA recognition motif (RRM) that mediates the interaction with its target mRNAs, as well as a DAZ motif and a proline-rich region, which might be involved in protein-protein interactions (Tsui et al., 2000a). *Dazl* expression is first detectable in post-migratory PGCs at E11.5 and persists till gametogenesis (Cauffman et al., 2005). It is localized in both nucleus and cytoplasm of fetal gonocytes, but relocates to the cytoplasm during male meiosis (Reijo et al., 2000). The deletion of mouse *Dazl* leads to the loss of germ cells in both male and female gonads (Ruggiu et al., 1997). Additionally, it has also been reported that *Dazl* deficiency leads to a reduction in post-migratory PGCs number, failure to re-establish genomic imprints, and defects in meiotic progression (Haston et al., 2009).

2.3. Pluripotent stem cells and their origins

Pluripotent stem cells have the capacity of prolonged self-renewal and the ability to differentiate into all three germ layers (ectoderm, endoderm and mesoderm). However, alone they can not give rise to a fetus or entire organism because of lacking the ability of

differentiating into extra-embryonic tissues such as placenta (Byrne, 2011). Based on the origin of the cells which are used for the establishment of pluripotent cells, the pluripotent stem cell lines can be currently defined as follows: [1] Embryonic Stem Cells (ESCs) derived from the Inner Cell Mass (ICM) of E3.5 day mouse blastocysts (Evans et al., 1987; Martin, 1981), [2] Epiblast Stem Cells (EpiSCs) derived by the culture of proximal region of E5.5-6.5 embryo epiblast (Brons et al., 2007; Tesar et al., 2007), [3] Embryonic Germ Cells (EGCs) derived from the *in vitro* culture of PGCs, which have colonized the genital ridge at E12.5 (Matsui et al., 1992; Resnick et al., 1992), [4] Embryonic Carcinoma Cells (ECCs) derived from either testicular or ovarian teratomas of embryos (Kleinsmith and Pierce 1964; Finch and Ephrussi 1967; Kahan and Ephrussi 1970), [5] multipotent Germ line Stem Cells (mGSCs) and multipotent adult Germ line Stem Cells (maGSCs) derived from neonatal and adult mouse testis, respectively (Kanatsu-shinohara et al., 2004, Guan et al., 2006), and [6] induced Pluripotent Stem Cells (iPSCs), generated by reprogramming of somatic cells by enforced expression of several ectopic transcriptional factors such as *Oct4*, *Sox2*, *Klf4*, *c-Myc*, *Nanog,* and *Lin28* (Takahashi and Yamanaka 2006; Takahashi et al., 2007). The developmental origin of fore-mentioned pluripotent stem cells is depicted in **figure 2.2**.

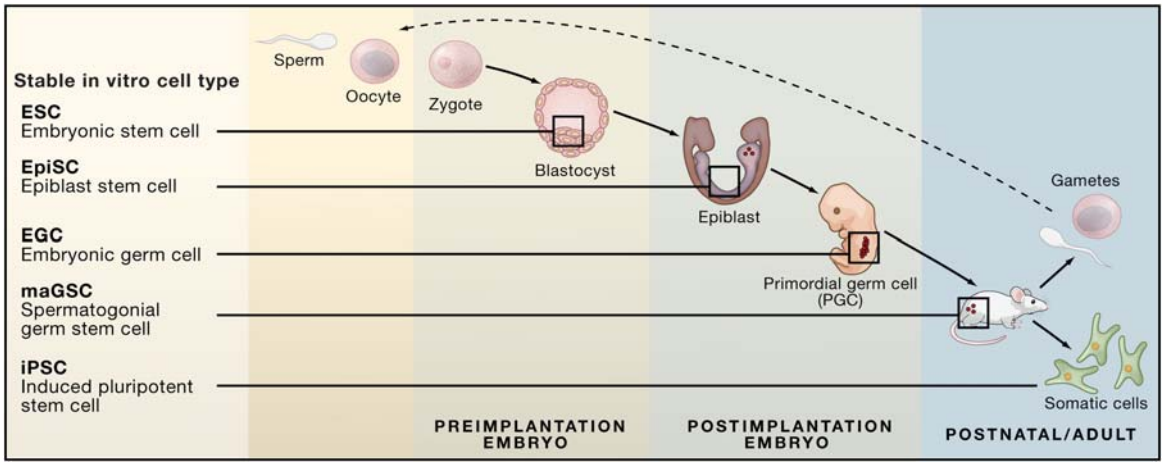


Figure 2.2. Developmental origins of pluripotent stem cells. Different types of pluripotent cells can be derived by explanting cells at various stages of mouse embryonic development or from neonatal and adult mouse testis. Induced pluripotent stem cells (iPSCs) are derived by direct reprogramming of somatic cells *in vitro* by forced expression of certain transcriptional factor cocktails. (Figure adapted from: Hanna et al., 2010).

Tracing back to the origin of above mentioned pluripotent cell lines highlights that EpiSCs, EGCs, ECCs, GSCs and maGSCs, all have the germ line origin, except ESCs and iPSCs whose origins or transitions are currently not clearly understood. Due to the

developmental stage of isolation, they have their unique gene expression marks to distinguish from each other. However, they all share the expression of germ cell stage specific marker genes such as *Blimp1*, *Ifitm3*, *Dppa3*, *Mvh* and *Dazl*. This may indicate that all pluripotent cell types could have a germ cell origin or might transit through the germ cell fate and these marker genes are all necessary for maintenance of pluripotency.

2.3.1. Induced pluripotent stem cells (iPSCs)

iPSCs are being considered as the most promising tool which can be used in disease modeling, regenerative medicine applications, and drug screening because of their high similarities to ESCs (**Fig. 2.3**).

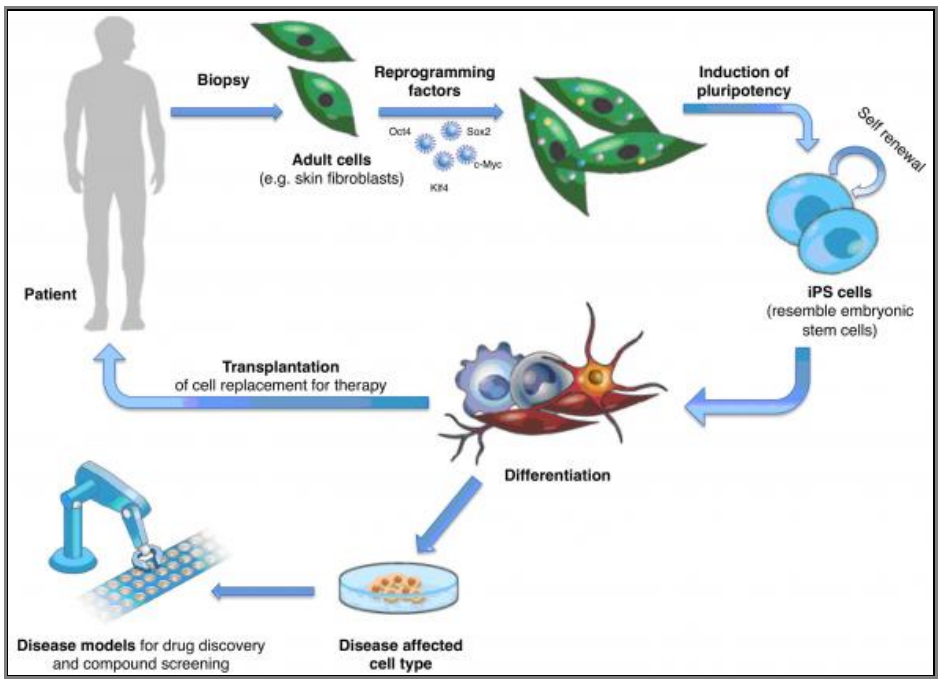


Figure 2.3. The derivation and applications of iPS cells. The iPS cells can be generated from patient biopsies. Later on the patient tailored iPS cells are cell-type-specifically differentiated and used for transplantation in cell replacement therapy. Disease-affected cells differentiated from iPS cells can be used for disease modeling and drug discovery. (Figure adapted from: http://www.eurostemcell.org).

Importantly, the iPS technology allows researchers to obtain and to study pluripotent cells without the controversial use of embryos and also to transplant patient-tailored cells to avoid any immunogenic responses (Nishikawa et al., 2008). Due to their capacity to differentiate into all kinds of cell types in the body, the use of iPSCs in regenerative medicine is very alluring. Although promising, several studies revealed the existence of minor differences between mouse iPSCs and ESCs at both genetic and epigenetic levels, demanding a need to optimize current protocols (Gore et al., 2011;

Hussein et al., 2011; Zhao et al., 2011). Of particular interest is the imprinting defect observed at the *Dlk1-Dio3* imprinting cluster in majority of iPS cell lines (Stadtfeld et al., 2010). The iPSCs with imprinting defects are developmentally compromised and show reduced chimerism with loss of germ-line transmission capacity (Stadtfeld et al., 2010; Liu et al., 2010). Taken together, these results suggest that there are still many unknown flaws with iPSCs and a complete characterization should be performed before consideration of any clinical applications.

2.4. Objectives

To better understand the role of germ cell marker genes in establishment and maintenance of pluripotency, the current study was designed and performed with following objectives:

1. Expression analysis of selected stage-specific germ cell marker genes in various pluripotent cell types at the genetic and epigenetic level.
2. Analysis of selected stage-specific germ cell marker genes expression during the process of iPS cells generation.
3. Role of germ stem cell marker gene, *Dppa3*, in establishment of pluripotency during somatic cell reprogramming.
4. Elucidating the function of mouse Dazl and its novel splice variant in pluripotent stem cells.

3. Results

The present thesis systematically analyzed the expression of selected stage-specific germ cell marker genes in pluripotent cells and studied their functions in establishment and induction of pluripotency. To gain more insights into the molecular mechanism, two representative marker genes, *Dppa3* (germ stem cell marker) and *Dazl* (pre-meiotic maker) were selected and their roles in pluripotency were elucidated separately. The results of this thesis are summarized in the three following manuscripts, where the first one has been published, the second and the third have been submitted to peer-reviewed journals.

3.1. Stage-specific germ-cell marker genes are expressed in all pluripotent cell types and emerge early during induced pluripotency.

3.2. Dppa3 binds to the IG-DMR of the *Dlk1-Dio3* imprinting cluster and prevents its imprinting loss during iPS cell generation.

3.3. Mouse Dazl and its novel splice variant function in translational repression of target mRNAs in embryonic stem cells.

3.1. Stage-specific germ cell marker genes are expressed in all pluripotent cell types and emerge early during induced pluripotency

In the first part of the thesis, using mouse as a model system, we systematically studied RNA and protein expression of germ stem cell/pre-meiotic (GC/PrM) markers in ESCs, iPSCs and germ line derived pluripotent stem cell lines such as EGCs, ECCs and maGSCs. Furthermore, by taking the advantage of gene silencing technique, we showed that the GC/PrM networks are independent from pluripotency network comprising of Oct3/4, Sox2 and Nanog in ESCs. Strikingly, the chromatin immunoprecipitation (ChIP) assay revealed an active chromatin structure for GC marker genes and a bivalent chromatin states at PrM marker genes. Additionally, we studied the gene expression pattern during the time course of induced pluripotent stem cells generation from fibroblast cells and identified that GC marker genes are expressed earlier than that of pluripotency marker genes. To sum up, based on our results, we speculate that the expression of GC markers might indicate the germ cell origin of ESCs whereas the PrM marker genes expression might indicate the poised state of pluripotent stem cells for germ line commitment.

Authors: Xingbo Xu, D.V.Krishna Pantakani, Sandra Lührig, Xiaoying Tan, Tatjana Khromov, Jessica Nolte, Ralf Dressel, Ulrich Zechner and Wolfgang Engel

Status: Published in PLoS ONE 6(7): e22413. doi:10.1371/journal.pone.0022413

Impact Factor: 4.351

Author contributions to the work:

1. Xingbo Xu: Conception and design of experiments, data collection, data analysis, data interpretation, helped in drafting the manuscript.
2. D.V. Krishna Pantakani: Conception and design of experiments, data analysis, data interpretation, drafted the manuscript.
3. Sandra Lührig: Performed Western blot and RT-PCR assay, data analysis.
4. Xiaoying Tan: Data collection, data analysis.
5. Tatjana Khromov: Performed Chip assay, data analysis.
6. Jessica Nolte: Data collection, data analysis.
5. Ralf Dressel: Performed teratoma formation assay, data analysis.
6. Ulrich Zechner: Performed DNA methylation assay, data analysis, helped in drafting the manuscript.
7. Wolfgang Engel: Conception and design of experiments, financial support.

OPEN ACCESS Freely available online

PLoS one

Stage-Specific Germ-Cell Marker Genes Are Expressed in All Mouse Pluripotent Cell Types and Emerge Early during Induced Pluripotency

Xingbo Xu[1], D. V. Krishna Pantakani[1]*, Sandra Lührig[1], Xiaoying Tan[1], Tatjana Khromov[1], Jessica Nolte[1], Ralf Dressel[2], Ulrich Zechner[3], Wolfgang Engel[1]*

1 Institute of Human Genetics, University of Goettingen, Goettingen, Germany, **2** Department of Cellular and Molecular Immunology, University of Goettingen, Goettingen, Germany, **3** Institute of Human Genetics, University of Mainz, Mainz, Germany

Abstract

Embryonic stem cells (ESCs) generated from the *in-vitro* culture of blastocyst stage embryos are known as equivalent to blastocyst inner cell mass (ICM) *in-vivo*. Though several reports have shown the expression of germ cell/pre-meiotic (GC/PrM) markers in ESCs, their functional relevance for the pluripotency and germ line commitment are largely unknown. In the present study, we used mouse as a model system and systematically analyzed the RNA and protein expression of GC/PrM markers in ESCs and found them to be comparable to the expression of cultured pluripotent cells originated from the germ line. Further, siRNA knockdown experiments have demonstrated the parallel maintenance and independence of pluripotent and GC/PrM networks in ESCs. Through chromatin immunoprecipitation experiments, we observed that pluripotent cells exhibit active chromatin states at GC marker genes and a bivalent chromatin structure at PrM marker genes. Moreover, gene expression analysis during the time course of iPS cells generation revealed that the expression of GC markers precedes pluripotency markers. Collectively, through our observations we hypothesize that the chromatin state and the expression of GC/PrM markers might indicate molecular parallels between *in-vivo* germ cell specification and pluripotent stem cell generation.

Citation: Xu X, Pantakani DVK, Lührig S, Tan X, Khromov T, et al. (2011) Stage-Specific Germ-Cell Marker Genes Are Expressed in All Mouse Pluripotent Cell Types and Emerge Early during Induced Pluripotency. PLoS ONE 6(7): e22413. doi:10.1371/journal.pone.0022413

Editor: Eliana Saul Furquim Werneck Abdelhay, Instituto Nacional de Câncer, Brazil

Received April 28, 2011; **Accepted** June 21, 2011; **Published** July 25, 2011

Funding: This work was partly supported by the German Research Foundation (Deutsche Forschungsgemeinschaft) DFG SPP1356 (EN 84/22-1; ZE 442/4-1) and DFG FOR 1041 (EN 84/23-1; ZE 442/5-1). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: krishna.if1@gmail.com (DVKP); wengel@gwdg.de (WE)

Introduction

The capacity of long-term self-renewal and the unique ability to differentiate into all three germ layer cell types (ectoderm, endoderm and mesoderm) define pluripotency. According to the source of cells used for the establishment of pluripotent cells, one can currently distinguish between the following pluripotent cell lines: (1) embryonic stem cells (ESCs) derived from the inner cell mass of mouse blastocysts at embryonic day 3.5 (E3.5) [1,2], (2) embryonic germ cells (EGCs) cultured from primordial germ cells (PGCs) that colonize the genital ridge at E12.5 [3,4], (3) embryonal carcinoma cells (ECCs) isolated from germ-cell tumors of either testis or ovary [5], (4) germ line stem cells isolated from mouse neonatal and adult testis (GSCs and maGSCs, respectively) [6,7] and (5) induced pluripotent stem cells (iPSCs), derived from reprogramming somatic cells by ectopic expression of defined transcription factors [8]. All the above mentioned cultured pluripotent cell lines (EGCs, ECCs, GSCs, and maGSCs) have a germ-cell origin, except ESCs, whose origin is not clearly understood. Although these cell lines have different molecular profiles mostly due to their developmental stage of isolation, they share the expression of germ cell/pre-meiotic (GC/PrM) markers that may indicate a germ-cell origin [9].

During embryonic development, the specification of PGCs is crucial for the development of the germ line, which is finally destined to give rise to the totipotent zygote upon fertilization. Prior to gastrulation, the precursors of primordial germ cells arise in the E6.25 proximal epiblast from 4–8 cells positive for the transcriptional repressor Blimp1 [10,11]. These Blimp1-positive cells continuously proliferate and start to express other PGC markers such as Fragilis and Stella by E7.5. Thereafter, PGCs initiate migration and colonization of the genital ridge and increase their number to approximately 4000 by E12.5 [12,13]. Further development of PGC/germ cells to mature spermatozoa or oocytes depends on the coordinated genetic and epigenetic events [14]. Interestingly, several studies have demonstrated the expression of some of the GC/PrM markers like Blimp1, Stella, Fragilis, Piwil2, Dazl and MVH in ES cells at the RNA level [15,16,17], raising the possibility that ES cells might originate from the germ line [9].

In the present study, using mouse as a model system, we have systematically analyzed the expression of GC/PrM markers in ES cells compared to germ line origin cultured pluripotent cells like EGCs, ECCs, GSCs and maGSCs and found comparable expression at the RNA and protein level. Moreover, we show the expression of Stella, Dazl and MVH in preimplantation

embryos and, the independence of pluripotency-specific networks from germ cell-specific networks in ES cells. Interestingly, chromatin immunoprecipitation (ChIP) analysis revealed that ES cells exhibit active chromatin states at GC marker genes and a bivalent chromatin structure at PrM marker genes. Further, gene expression analysis during iPSC generation revealed that the expression of GC markers precedes pluripotency markers. Collectively, our data indicates the possible link between *in-vivo* germ cells specification and *in-vitro* pluripotent stem cells generation.

Materials and Methods

Cell culture

Derivation and maintenance of male mouse ESC and maGSC lines from different genetic backgrounds (129Sv and C57BL/6) were described previously [18]. The female ESC line ES Rosa26 was generated from Rosa26-LacZ transgenic mouse line as described for MPI-VI, a previously generated female ESC line [19]. iPS cells (O18) were a kind gift from Prof. Rudolf Jaenisch [20]. All above cell lines including EGC line (EG 1) and parthenogenetic cells were maintained in standard ESC culture conditions. ECC cell line (F9) protein extract was provided by Mr. Peter Christalla, Goettingen. For knockdown experiments, ES cells were seeded in KO-DMEM supplemented with KO-serum replacement (Invitrogen) at a density of 2×10^5 cells/ml on feeder layer. After 5 h of plating, the cells were transfected with either *Dazl* siRNA (NM_010021.4_stealth_199, _726, _1056, Invitrogen) or *MVH* siRNA (NM_001145885.1_stealth_83, _922, _1599, Invitrogen) or *Oct3/4* siRNA (NM_013633_Stealth_356, _463, _727, Invitrogen) or scrambled siRNA (Control siRNA, Invitrogen) using Lipofectamine-2000 (Invitrogen) according to manufacturer's instructions. After 3 h of transfection, the medium was changed to standard ES culture medium and allowed to grow for 24 h. The next day, transfection was repeated and cells were harvested after additional 24 h of culture.

All animal experimentations were reviewed and approved by the Institutional Animal Care and Use committee of the University of Goettingen (Approval ID: 33.9.42502-097/06).

Generation of iPS cells

We used Yamanaka factors (retroviral expression vectors for Oct3/4, Sox2, Klf4, c-Myc) to generate iPS cells [8]. For reprogramming studies, embryonic fibroblasts isolated from transgenic *Nanog*-EGFP mice [21] were transduced with retroviral particles as previously described [8]. To establish iPS cell lines, colonies which appeared after 10 days of virus infection were picked manually and cultured in 24-well plates containing feeder layer with standard ES medium and were monitored for the *Nanog*-EGFP activation using invert microscope (Olympus). For time course experiments, all cells from culture plates were collected on days 0 (Control), 5-10, 12, 14, 16, 18, 20 and 22 after viral infection for total RNA isolation.

RNA extraction, RT-PCR and qPCR

Total RNA was extracted from cells using Trizol Reagent (Invitrogen) and from ~50 blastocysts using Picopure Kit (Analytical Technology) following the manufacturers' protocols. Total RNA from blastocysts or 5 µg of total RNA from cells was digested with *DNaseI* (Sigma) and used for cDNA synthesis using the SuperScriptII system (Invitrogen). For qPCR analysis, diluted cDNA (1/20) was used as a template in a Platinum SYBR Green qPCR SuperMix-UDG with ROX (Invitrogen) and run in ABI 7900HT Real-Time PCR System (Applied Biosystems). Primers used in RT-PCR and qPCR are listed in supplementary tables (Tables S1, S2, S3).

Protein extraction and Western blotting

Proteins were extracted from cells and tissues using lysis buffer (10 mM Tris–HCl (pH 8.0), 1 mM EDTA, 2.5% SDS, 100 mM PMSF) containing protease inhibitor cocktail (Roche). Protein samples were resolved on 4–12% SDS-PAGE and transferred onto nitrocellulose membrane (Amersham Biosciences). Membranes were processed using standard Western blot protocols, and signals detected using a chemiluminescent kit (Santa Cruz Biotechnology). Antibody sources are listed in supplementary tables (Table S5).

Early-stage-embryo collection, immunocytochemistry, and alkaline phosphatase staining

Early stage embryos collected from the oviduct or uterus of pregnant CD-1 mice were used for whole-mount immunostaining as described previously [22]. Immunostained embryos were mounted on cavity microscope slides using Vectashield DAPI mounting medium (Vector Laboratories) and images were acquired using an Olympus confocal microscope. Immunostaining on iPS cells using SSEA1 antibody was performed as previously described [23]. Alkaline phosphatase (AP) staining was performed according to manufacturer's protocol (Sigma).

Chromatin immunoprecipitation

The chromatin immunoprecipitation (ChIP) was performed essentially as previously described [23]. Briefly, cultured wild-type ESCs were cross-linked with 1% formaldehyde and briefly incubated in lysis buffer followed by sonication (Branson Sonifier 250). Soluble chromatin was pre-cleared with protein-A sepharose beads and incubated with and without (negative-control) antisera (3–5 µg) directed against H3K4me3, H3K9ac, H3K9me3, and H3K27me3. Enriched DNA was analyzed by qPCR using SYBR green (Invitrogen)-based PCR amplification with primers listed in supplementary tables (Table S4). qPCR data from two or more biological replicates were calculated and expressed as percentage of input DNA precipitation. Antibodies used in ChIP are listed in supplementary tables (Table S5).

Teratoma formation assay

Teratoma formation assay was performed as previously described [24,25]. Briefly, iPS cells (1×10^6 cells) were injected subcutaneously into 8 to 10 weeks old female severe combined immunodeficient $RAG2^{-/-}c\gamma c^{-/-}$ mice lacking T, B, and natural killer (NK) cells. Tumor growth was monitored weekly by palpation and size was recorded using linear calipers. Animals were sacrificed when a tumor diameter of 1 cm was reached. Autopsies were performed and tumor tissue was placed in phosphate-buffered 4% formalin for 16 h and then embedded in paraffin. For histological analysis, the specimens were stained with haematoxylin and eosin (HE).

Statistical Analysis

All qPCR data for RNA expression analysis (two or more biological replicates) were calculated using standard curve method. For statistical significance calculations, 2way ANOVA (GraphPad Prism 4.0) test was used.

Results

Pluripotent stem cells express GC/PrM genes

To investigate whether GC/PrM gene expression is characteristic of all known mouse pluripotent cells, we systematically analyzed several pluripotent cells. Firstly, we examined GC/PrM gene expression in male maGSCs and ESCs from different genetic

backgrounds and in iPSCs, EGCs, and F9 cells by Western blotting (Fig. 1A). GC markers Stella and Fragilis were readily detected in all cell types (Fig. 1A), including parthenogenetic cells (Fig. 1B). Further, PrM markers Piwil2, Dazl, and MVH were found to be expressed in all pluripotent cells, except ECCs (Fig. 1A). Protein levels of GC/PrM markers were reduced or absent in ESCs and maGSCs upon spontaneous differentiation with retinoic acid for 20 days (Fig. 1B). Overmore, we analyzed multipotent mesenchymal stem cells (MSCs) and could not detect any expression of GC/PrM markers (Fig. 1B). We also performed RT-PCR analysis for other PrM, meiotic, and post-meiotic markers (Fig. 1C). Expression of PrM markers *Stra8*, *Rnf17*, *Rnh2*, and *Piwil2* was detected in all cells (Fig. 1C). Surprisingly, meiotic markers *Sycp3*, *Pgk2*, and *Creb3/4* were also detected in all pluripotent cell lines (Fig. 1C). However, expression of several other developmental markers such as post-meiotically expressed *Tp2*, *Theg*, *Gpx4*, *Prm1*, and mature spermatozoan marker *Cylc1* was undetectable as expected (Fig. 1C).

To determine whether the expression of GC/PrM markers is specific to male pluripotent cells, we studied two female ES cell lines, namely, MPI VI and ES Rosa26. Pluripotency of these cell lines was confirmed by detecting the expression of the key pluripotency markers Oct3/4 and Sox2 (data not shown). Both female pluripotent cell lines were found to express all analyzed GC/PrM markers with levels similar to those of male pluripotent cells (Fig. 1D).

GC/PrM genes are also expressed in early embryogenesis

Next, we studied the expression of GC marker (Stella) and PrM markers (Dazl and MVH) in early stages of mouse embryogenesis (2-, 4-, 8-cell stages) by immunocytochemistry (ICC) (Fig. 2A). Interestingly, we found Stella, Dazl and MVH to be expressed throughout all stages of embryogenesis (Fig. 2A). To determine the expression levels of GC/PrM markers at the blastocyst stage, we performed qPCR on blastocyst stage embryos (Fig. 2B). In agreement with our ICC results, all analyzed GC/PrM markers (*Fragilis*, *Dazl*, *MVH*, and *Stra8*) were detected at the blastocyst stage with transcript levels, that are, however, markedly lower than those of pluripotency markers such as *Oct3/4*, *Nanog*, *Lin28* (Fig. 2B).

Independency of pluripotent and GC/PrM networks in ESCs

The widespread expression of GC/PrM markers in pluripotent cells led us to study their influence on other GC/PrM and key pluripotency markers (Fig. 3). Firstly, we down-regulated *Dazl* in ES cells using siRNA and found an ~80–90% decrease at both the RNA and protein level (Fig. 3A, B). In contrast, control siRNA treated cells did not exhibit altered Dazl expression levels (Fig. 3A, B). Then, we performed a qPCR-based analysis of expression levels of key pluripotency markers and detected no significant differences among control siRNA treated and Dazl siRNA treated cells (Fig. 3C). Similarly, the expression of PrM markers *MVH* and *Stra8* did not change significantly, whereas GC markers (*Stella* and *Fragilis*) showed significant up-regulation in Dazl down-regulated cells (Fig. 3D). Similarly, ~70% down-regulation of *MVH* expression (Fig. 3E, F) did not influence expression of key pluripotency markers (Fig. 3G). Consistent with Dazl down-regulation, MVH depletion had no effect on *Dazl* and *Stra8*, but *Stella* and *Fragilis* were significantly up-regulated (Fig. 3H). Conversely, we down-regulated *Oct3/4* and studied the expression of GC/PrM and pluripotency markers. The down-regulation of *Oct3/4* resulted in significant down-regulation of *Klf4* expression, whereas the expression of other pluripotent markers such as *Nanog*, *Zfp206*, and *Lin28* did not alter (Supplementary Fig. S1). Furthermore, the down-regulation of *Oct3/4* had no statistically significant effect on the expression of GC/PrM markers (Supplementary Fig. S1).

Active chromatin at GC marker gene promoter regions and bivalent chromatin at PrM marker gene promoters in ESCs and iPSCs

We hypothesized that the chromatin state at the promoter regions of GC/PrM markers might elucidate their role in the establishment/maintenance of pluripotency or lineage specification in ESCs. We analyzed the ChIP sequencing data of mouse ES cells, which is freely available [26] and found that the promoter regions of GC markers *Blimp1*, *Stella* and *Fragilis* were enriched for H3K4me3 indicating the transcriptionally active chromatin state, as seen for *Oct3/4* (Supplementary Fig. S2A). In contrast the promoter regions of *Dazl* and *MVH* were decorated with both H3K4me3 and H3K27me3, highlighting the bivalent chromatin state, which is a hall mark of lineage specification genes, such as *Hoxa11* and *Pax5* (Supplementary Fig. S2B). To further validate these observations, gene specific histone modification profiles (active: H3K4me3, H3K9ac; and repressive: H3K9me3, H3K27me3) were analyzed by ChIP at the promoter regions of GC markers *Fragilis* and *Blimp1*, and PrM markers *Dazl* and *MVH*, and compared to the promoter regions of *Oct3/4* (transcriptionally active chromatin) and *Hoxa11* and *Pax5* (bivalent chromatin) in ES cells (Fig. 4A). qPCR quantification of ChIP DNA showed that the promoter regions of GC markers *Fragilis* and *Blimp1* were enriched for the activating modifications H3K4me3 and H3K9ac, but depleted for the repressive modifications H3K9me3 and H3K27me3, indicating a transcriptionally active chromatin similar to key pluripotency *Oct3/4* gene promoter (Fig. 4A). In contrast, the promoters of PrM genes *Dazl* and *MVH* were enriched for both active (H3K4me3 and H3K9ac) and repressive (H3K27me3) modifications, representing the bivalent chromatin domain similar to lineage specific genes (*Hoxa11* and *Pax5*) (Fig. 4A). Moreover, we also performed gene specific histone modification profiling in established iPS cells [20] and found similar results like ES cells (Fig. 4B).

GC markers emerge during early reprogramming of MEFs into iPSCs

To further understand the role of GC/PrM markers during the establishment and maintenance of pluripotency, we used ectopic expression of the four Yamanaka factors (Oct4, c-Myc, Klf4, and Sox2) for reprogramming of somatic cells to induced pluripotency. Firstly, we reprogrammed MEFs isolated from Nanog-EGFP mice and established four iPS lines (xu1, 2, 5, and 6), which are morphologically similar to ES cells, activate the *Nanog* promoter-driven EGFP expression, positive for AP staining, SSEA1 immunostaining and express endogenous Oct3/4 and Sox2 (Fig. 5A–D). The iPSC lines were further characterized by histone modification, and DNA methylation profiling of key pluripotent marker genes (data not shown). Finally, all the iPSC lines (xu1, 2, 5 and 6) were injected subcutaneously into immunodeficient mice. Two recipients were used per cell line. In all mice, tumors were observed and the mice had to be sacrificed between day 18 and day 33 after injection. The tumors were identified as teratomas by histological examination as exemplified for iPS cell lines xu2 and 6 (Supplementary Fig. S3). Then, we setup a time course experiment and analyzed the expression levels of key pluripotency genes and GC/PrM genes during the course of reprogramming as outlined in figure 5E. Expression analysis using real time qPCR revealed

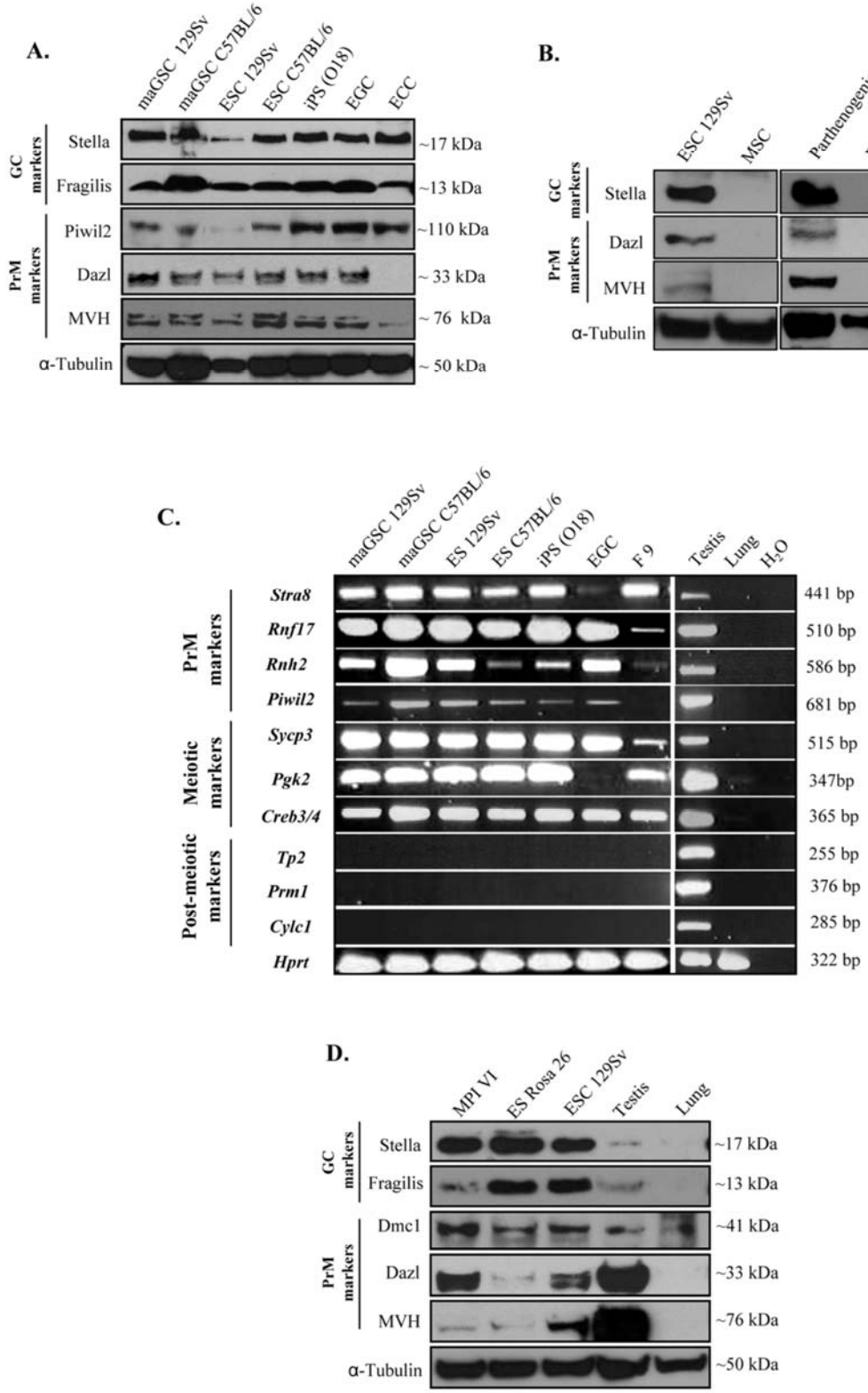

A.
maGSC 129Sv
maGSC C57BL/6
ESC 129Sv
ESC C57BL/6
iPS (O18)
EGC
ECC
GC markers
Stella
Fragilis
PrM markers
Piwil2
Dazl
MVH
α-Tubulin
~17 kDa
~13 kDa
~110 kDa
~ 33 kDa
~ 76 kDa
~ 50 kDa
B.
ESC 129Sv
MSC
Parthenogenic
ESC 129Sv-20d RA
maGSC 129Sv-20d RA
GC markers
Stella
PrM markers
Dazl
MVH
α-Tubulin
~17 KDa
~33 KDa
~76 KDa
~50 KDa
C.
maGSC 129Sv
maGSC C57BL/6
ES 129Sv
ES C57BL/6
iPS (O18)
EGC
F 9
Testis
Lung
H2O
PrM markers
Meiotic markers
Post-meiotic markers
Stra8
Rnf17
Rnh2
Piwil2
Sycp3
Pgk2
Creb3/4
Tp2
Prm1
Cylc1
Hprt
441 bp
510 bp
586 bp
681 bp
515 bp
347bp
365 bp
255 bp
376 bp
285 bp
322 bp
D.
MPI VI
ES Rosa 26
ESC 129Sv
Testis
Lung
GC markers
Stella
Fragilis
PrM markers
Dmc1
Dazl
MVH
α-Tubulin
~17 kDa
~13 kDa
~41 kDa
~33 kDa
~76 kDa
~50 kDa

Figure 1. Expression analysis of GC/PrM genes in pluripotent cell lines. (**A**) Expression of GC markers (Stella, Fragilis) and PrM markers (Piwil2, Dazl, MVH) in maGSCs and ESCs (of 129Sv and C57BL/6 genetic backgrounds), iPSCs, EGCs and ECCs**.** (**B**) Western blot of GC marker (Stella) and PrM markers (Dazl and MVH) in mouse mesenchymal stromal cells (MSC), parthenogenetic cells, ES and maGSC treated with retinoic acid for 20 days. Protein extract from ESCs was used as a positive control. (**C**) RT-PCR analysis showing the expression of PrM (*Stra8*, *Rnf17*, *Rnh2* and *Piwil2*), meiotic (*Sycp3*, *Pgk2* and *Creb3/4*) and post-meiotic (*Tp2*, *Prm1* and *Cylc1*) markers in maGSC and ESC (from different genetic backgrounds), as well as iPSC, EGC and ECC cell lines. cDNA from wild type mouse testis and lung were used as positive and negative controls, respectively. (**D**) Western blot analysis showing the expression of GC markers (Stella, Fragilis), PrM markers (Dmc1, Dazl, MVH) in two different female ESC lines (MPI VI, ES Rosa26). α-tubulin was used as loading control in **A, B and D**.
doi:10.1371/journal.pone.0022413.g001

significant expression levels of key germ cell markers (*Blimp1* and *Fragilis*) at day 6 and a gradual increase to the levels seen in ES cells by day 22 (Fig. 5F, G). Transcripts of *Stella*, another germ cell marker, were significantly detectable at day 10 of reprogramming and reached levels similar to those in ES cells by day 22. (Fig. 5F, G). In contrast, significant endogenous expression levels of the key pluripotency markers *Oct3/4* and *Sox2* occurred only on day 12 of reprogramming and showed expression levels typical for ES cells by

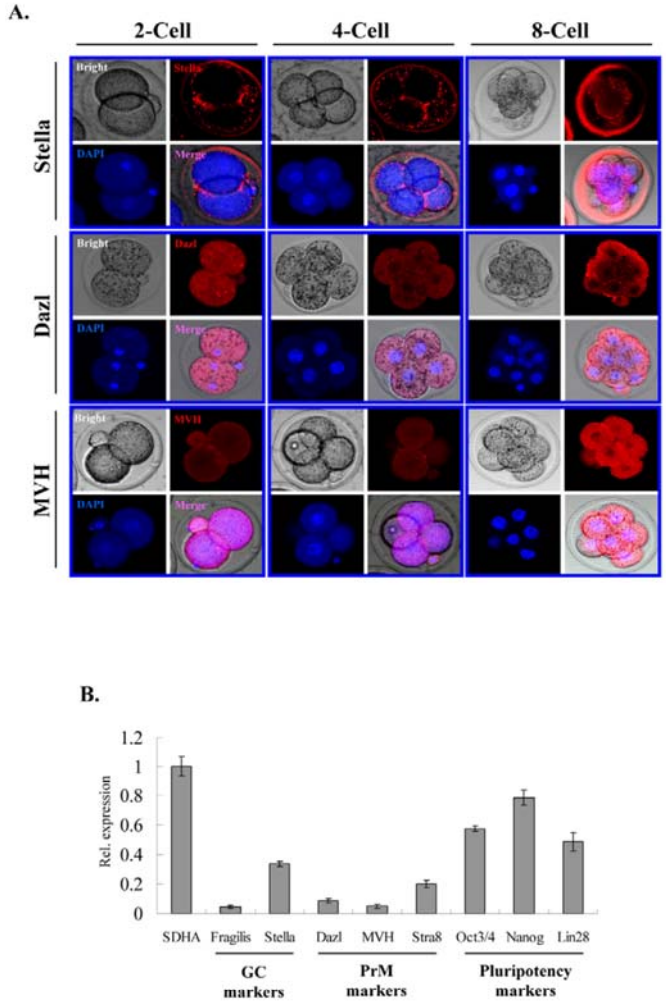


Figure 2. Expression analysis of GC/PrM genes in early embryogenesis. (**A**) Immunocytochemistry showing the expression of GC marker (Stella) and PrM markers (Dazl and MVH) in pre-implantation embryos (2-cell, 4-cell, and 8-cell stage). DAPI was used as a counter stain to visualize the nucleus. (**B**) Quantitative real time PCR analysis was used to evaluate the expression levels of GC (Fragilis), PrM (Dazl, MVH, Stra8) and pluripotency markes (Oct3/4, Nanog, Lin28) in blastocysts. Expression levels were normalized to *Sdha* (a house keeping gene). The qPCR data of three biological replicates (including three technical replicates each) were calculated and represented as a mean ±SD.
doi:10.1371/journal.pone.0022413.g002

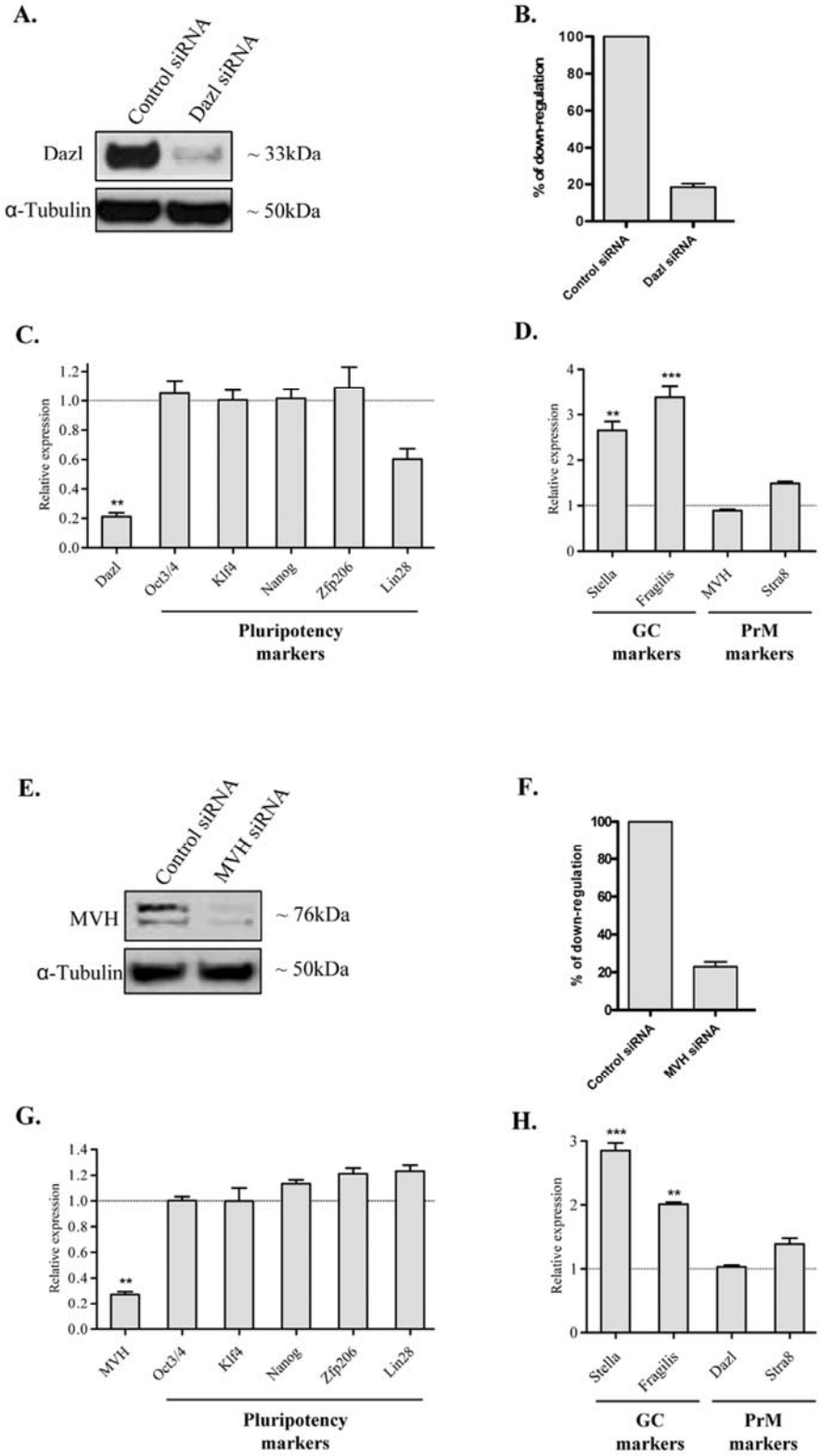

A.
Control siRNA
Dazl siRNA
Dazl
~ 33kDa
α-Tubulin
~ 50kDa
B.
% of down-regulation
Control siRNA
Dazl siRNA
C.
Relative expression
Dazl
Oct3/4
Klf4
Nanog
Zfp206
Lin28
Pluripotency markers
D.
Relative expression
Stella
Fragilis
MVH
Stra8
GC markers
PrM markers
E.
Control siRNA
MVH siRNA
MVH
~ 76kDa
α-Tubulin
~ 50kDa
F.
% of down-regulation
Control siRNA
MVH siRNA
G.
Relative expression
MVH
Oct3/4
Klf4
Nanog
Zfp206
Lin28
Pluripotency markers
H.
Relative expression
Stella
Fragilis
Dazl
Stra8
GC markers
PrM markers

Figure 3. Effect of the downregulation of PrM genes in ESCs. (**A and B**) Efficiency of Dazl down-regulation as shown by Western blot analysis and subsequent densitometric qunatification**.** (**C**) Real time qPCR analysis showing *Dazl* downregulation at the RNA level and the expression profile of the core pluripotency network (*Oct3/4*, *Klf4*, *Nanog*, *Zfp206* and *Lin28*). (**D**) Expression of GC (*Stella* and *Fragilis*) and PrM (*MVH* and *Stra8*) markers in *Dazl* down-regulated ESCs. (**E and F**) Western blot showing the efficiency of MVH downregulation and the densitometric quantification, respectively. (**G**) Real time qPCR analysis showing *MVH* downregulation at the RNA level and the expression profile of the core pluripotency network (*Oct3/4*, *Klf4*, *Nanog*, *Zfp206* and *Lin28*). (**H**) Expression of GC (*Stella* and *Fragilis*) and PrM (*Dazl* and *Stra8*) markers in *MVH* down-regulated ESCs. The dotted lines in **C**, **D**, **G**, **H** indicate the normalized expression levels of analyzed genes in control siRNA treated cells. The qPCR data of two biological replicates (including three technical replicates each) were calculated and represented as a mean ±SD. Expression levels, which are statistically significant, are indicated with asterisks (**p<0.01; ***p<0.001).
doi:10.1371/journal.pone.0022413.g003

day 22 (Fig. 5G). Further pluripotency markers like *Zfp206* and *Nanog* appeared only on day 18 and 20, respectively and increased to levels observed in ES cells only in fully reprogrammed and established iPS cells (Fig. 5G). Surprisingly, we could not detect significant expression of pre-meiotic markers such as *Stra8*, *Dazl* and *MVH* before day 22 of reprogramming. The expression of *Stra8* appeared not until day 22 and the other two markers were only present in established iPS cells (Fig. 5G).

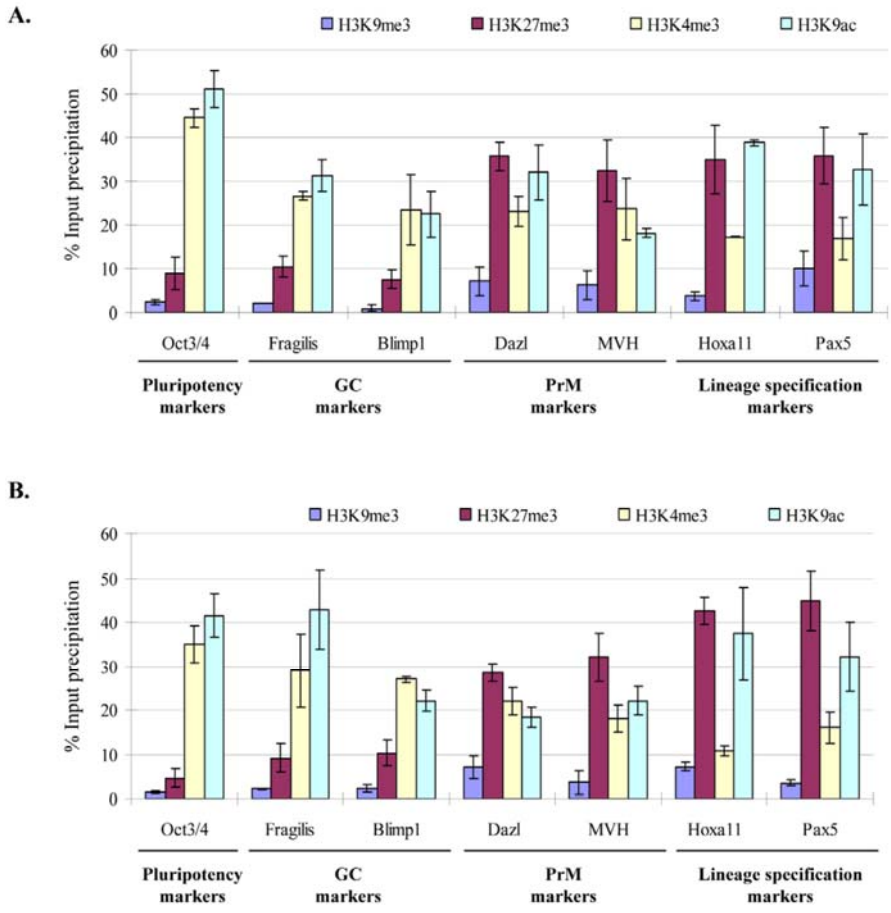
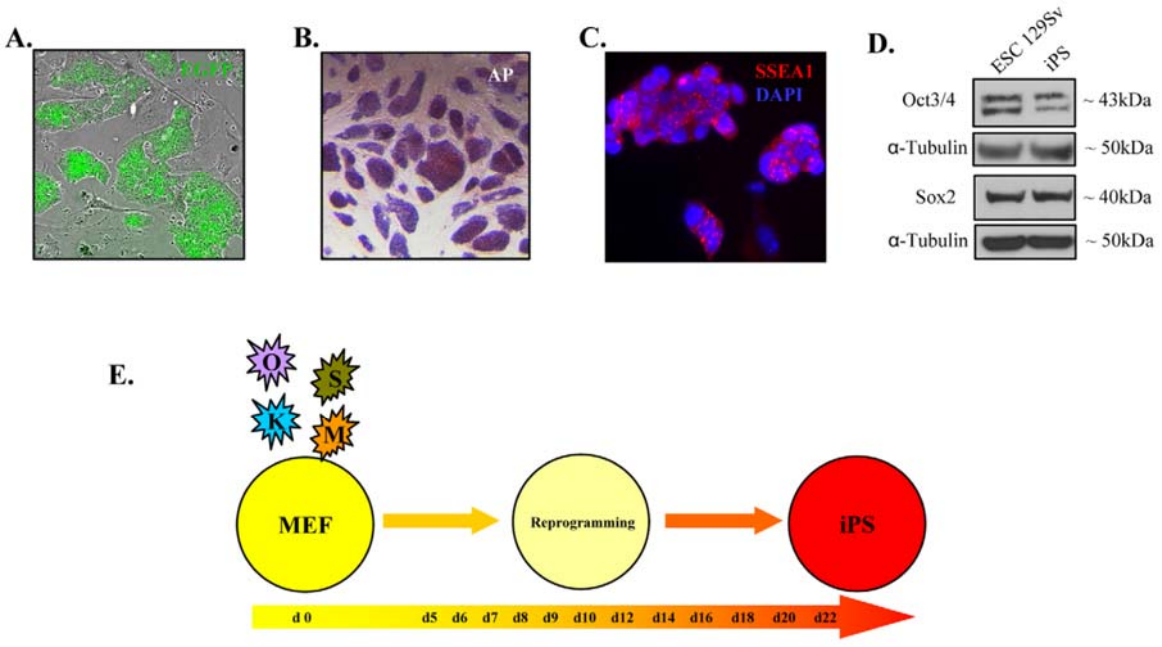
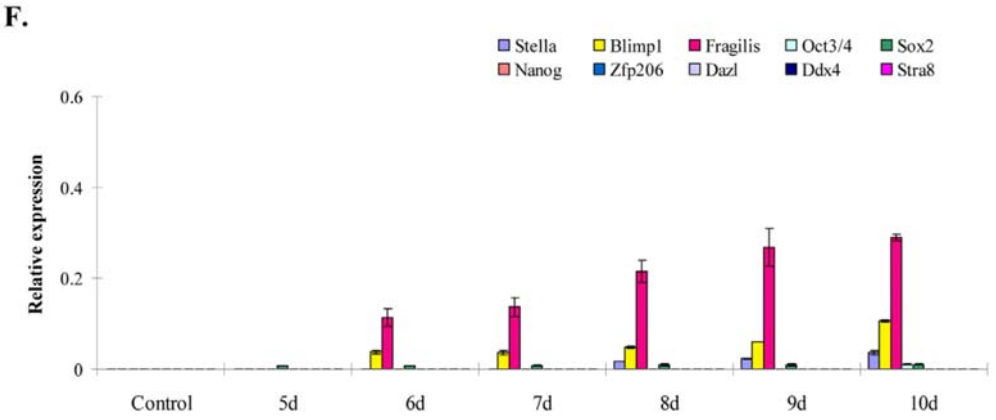


Figure 4. Epigenetic signature of GC/PrM genes in ESCs and iPSCs. ChIP and subsequent real time qPCR for various histone modifications (active: H3K4me3 and H3K9ac; repressive: H3K9me3 and H3K27me3) at the promoter regions of GC (*Fragilis* and *Blimp1*) and PrM markers (*Dazl* and *MVH*) in ESCs (**A**) and iPSCs (**B**). The promoter regions of GC markers are similar to the *Oct3/4* promoter and are enriched for active histone modifications, while the promoters of PrM markers are similar to the promoters of the lineage specific genes *Hoxa11*, *Pax5* and were enriched for both active and repressive marks indicating their bivalent chromatin structure. The qPCR data of two biological replicates (including three technical replicates each) were calculated and represented as a mean ±SD.
doi:10.1371/journal.pone.0022413.g004


A.
B.
AP
C.
SSEA1
DAPI
D.
ESC 129Sv
iPS
Oct3/4
~ 43kDa
α-Tubulin
~ 50kDa
Sox2
~ 40kDa
α-Tubulin
~ 50kDa
E.
O
S
K
M
MEF
Reprogramming
iPS
d 0
d5
d6
d7
d8
d9
d10
d12
d14
d16
d18
d20
d22


F.
Stella
Blimp1
Fragilis
Oct3/4
Sox2
Nanog
Zfp206
Dazl
Ddx4
Stra8
Relative expression
0.6
0.4
0.2
0
Control
5d
6d
7d
8d
9d
10d

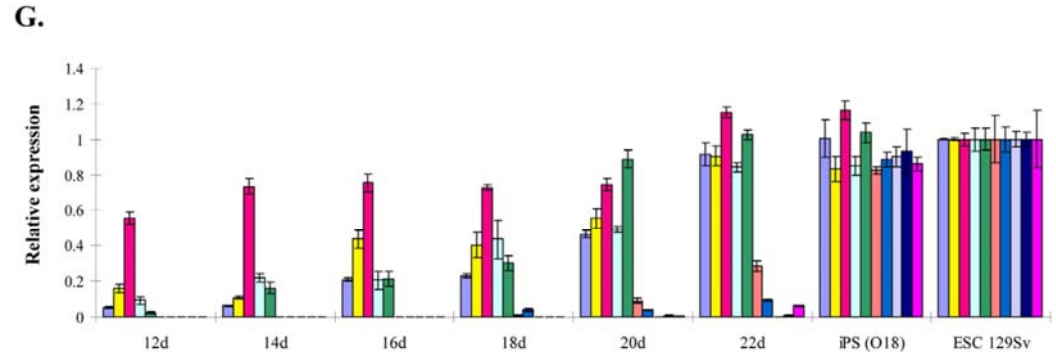

G.
Relative expression
1.4
1.2
1
0.8
0.6
0.4
0.2
0
12d
14d
16d
18d
20d
22d
iPS (O18)
ESC 129Sv

Figure 5. Expression pattern of GC/PrM and pluripotency genes during the course of iPS generation. Fully reprogrammed and established iPS cells express *Nanog* promoter-driven EGFP (**A**), positive for AP staining (**B**), and express SSEA1 (**C**). (**D**) Expression of key pluripotency markers Oct3/4 and Sox2 in iPS cells. Protein extract from ESCs was used as a positive control. (**E**) Schematic diagram depicting the generation of iPS cells using the four Yamanaka factors and the time course of sample collection (starting from the day of retrovirus transduction, day0 (d0) till day22 (d22)) for gene expression analysis. (**F**) Real time qPCR analysis of pluripotency (*Oct3/4*, *Sox2*, *Nanog* and *Zfp206*), germ cell (*Stella*, *Blimp1* and *Fragilis*) and pre-meiotic genes (*Dazl*, *MVH* and *Stra8*) during the time course of iPS cell generation from day5 (5d) to day 10 (10d) after virus infection. *Nanog*-EGFP MEFs were used as a control. **(G)** Real time qPCR analysis for the above mentioned genes during the time course of iPS cell generation from day12 (12d) to day 22 (22d) after virus infection. The qPCR data of two or more biological replicates (including three technical replicates each) were calculated and represented as a mean ±SD.
doi:10.1371/journal.pone.0022413.g005

Discussion

Though traditionally pluripotent ES cells are regarded as *in-vitro* counterpart of the inner cell mass (ICM), their origin is not yet clearly defined. Recently, it has been hypothesized that ESCs may have a germ-cell origin based on common molecular properties with other pluripotent cells of germ-cell origin [9]. Previously, ES cells were shown to express several GC/PrM markers [15,16,17]. In agreement with these results, that demonstrated the expression of several GC/PrM markers at the transcript level, our Western blot analysis detected the expression of GC/PrM markers in all analyzed pluripotent cell types available including iPS cells and female ES cells of mouse origin. Consistent with an earlier report [27], expression of GC/PrM genes was not detectable in bone marrow-derived multipotent stem cells, thus indicating the unique expression of GC/PrM genes only in pluripotent cells. Our immunocytochemistry expression analysis of GC/PrM markers in preimplantation embryos revealed the expression of Stella, Dazl and MVH in all analyzed preimplantation embryo stages (2-, 4-, 8-cell stage). Further, down-regulation of PrM genes in ESCs did not influence the expression levels of pluripotency network genes, but rather increase expression of GC genes. Conversely, down-regulation of pluripotency marker Oct3/4 showed no significant effect on GC/PrM marker genes, thus highlighting the maintenance of parallel but independent networks.

The genome-wide expression profiling of ES cells revealed the expression of a large number of genes at low levels due to the open chromatin state of ES cells leading to leaky expression [28,29,30,31]. To elucidate leaky versus essential expression of GC/PrM markers in ES cells, we analyzed the global ChIP-Seq data of ES cells and found an active chromatin state at PGC/germ cell markers and a bivalent chromatin structure at pre-meiotic markers [26]. In support of global ChIP-seq data, our gene-specific chromatin state of GC/PrM markers in ES cells confirmed the active chromatin state with enrichment for the activating histone modifications H3K4me3 and H3K9ac at the promoter regions of PGC markers *Blimp1* and *Fragilis*, which demonstrates the fundamental expression of these genes. In contrast, the promoter regions of *Dazl* and *MVH* were marked with bivalent chromatin state, i.e. enrichment for the two activating (H3K4me3

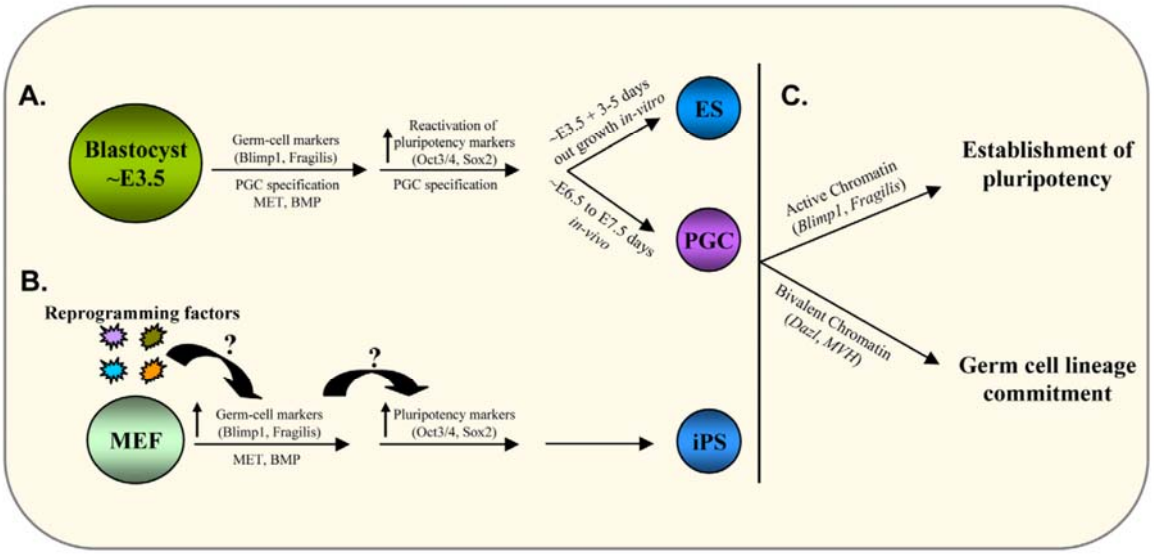


Figure 6. Hypothetical model for the germ cell origin of pluripotent ESCs. (**A**) The inner cell mass cells of the blastocyst are positive for Oct3/4 and Sox2 expression. During further development, primordial germ cell (PGC) specification in mouse implantation embryos (blastocyst (~E3.5) stage onwards) is marked by the expression of germ cell markers Blimp1 and Fragilis followed by reactivation of Oct3/4 and Sox2 and is completed by ~E6.5-E7.5 *in vivo*. This period of *in vivo* PGC specification parallels with the *in vitro* ESCs generation from pre-implantation blastocysts (~E3.5) together with 3–5 days of ESCs outgrowth. (**B**) On the other hand, reprogramming of somatic cells to iPS cells using the four Yamanaka factors (Oct3/4, Sox2, Klf4 and c-Myc) leads to the early appearance of germ cell markers followed by the activation of endogenous Oct3/4 and Sox2 and subsequent establishment of pluripotent state. This pattern of gene activation is equivalent to that of PGC specification and ESC establishment as discussed above. (**C**) The active chromatin state of germ cell markers in ESCs might indicate the developmental origin of ESCs from PGCs; the presence of bivalent chromatin at the promoter regions of pre-meiotic genes indicate the poised state for germ line commitment.
doi:10.1371/journal.pone.0022413.g006

and H3K9ac) and the repressive (H3K27me3) histone modifications, which is a hallmark of key developmental regulation/lineage specific genes [32,33]. The observed active chromatin state at GC marker genes might indicate the possible early germ cell specification epigenetic marks in pluripotent cells. Conversely, the bivalent chromatin state at PrM marker genes might represent the poised germ cell lineage specification or the heterogeneous expression of these genes in pluripotent cells.

Recent advances in direct reprogramming of somatic cells to induced pluripotency opened new avenues not only for tailor-made patient-specific cells for future regenerative medicine in addition to advancing our knowledge of the basic biology of establishment and maintenance of pluripotency [34]. Of particular interest is the role of GC/PrM markers during iPS cell generation using the four Yamanaka's factors. We analyzed the activation of GC/PrM markers along with the endogenous activation of core pluripotency markers during somatic reprogramming and found the activation of the PGC specification markers *Blimp1*, *Stella* and *Fragilis* to occur much earlier (between day 6 and day 9 of reprogramming) than activation of the endogenous pluripotency markers *Oct3/4* and *Sox2* (by day12 of reprogramming). In contrast, the expression of the PrM markers *Dazl*, *MVH* and *Stra8* was only detectable by day 22 and in established iPS cell lines, respectively. Recent studies of the molecular mechanisms underlying somatic reprogramming revealed that somatic cells undergo mesenchymal to epithelial transition (MET) during early reprogramming to acquire pluripotency through BMP signaling and vital expression of E-cadherin [35,36]. Interestingly, during embryonic development, PGC precursors rely on inductive BMP signals followed by MET activation and Fragilis, Blimp1, Stella and E-cadherin expression [37,38]. Loss of BMP signals, Blimp1 and E-cadherin expression results in the depletion or a reduced number of PGCs [38,39,40,41]. Taken together, we assume that even somatic cells acquire a "temporary" PGC/GC fate and finally establish pluripotency during reprogramming.

Based on our study and earlier reports [9], we propose a working model for the germ-cell origin of ESCs and the possible acquisition of PGC/GC fate by somatic cells during iPSCs generation (Fig. 6). According to our model, the ICM of blastocyst stage embryos (~E3.5) expresses key pluripotency markers Oct3/4, Sox2, and c-Myc. Following embryonic development, PGC specification *in-vivo* is marked by the expression of key PGC genes, where Blimp1 is activated by BMP signaling [10], facilitates the activation of Stella and E-cadherin, initiates the repression of the somatic program, and reactivates the pluripotency network before PGCs acquire migratory properties [38] (Fig. 6A). Considering the GC fate and lineage commitment of PGCs, key germ-cell markers may have active chromatin, whereas PrM genes may show bivalent chromatin (Fig. 6C). Similarly, ESC generation also starts with isolation of ~E3.5 blastocysts followed by culture to obtain outgrowth from the ICM. It is more likely that during the *in-vitro* ICM outgrowth, ICM cells proceed with the pre-programmed developmental program of PGC specification via BMP signals, initiate MET, begin expressing Fragilis, Blimp1, and Stella, re-activate pluripotency genes, and acquire the unique self-renewal property (Fig. 6A). The observed active chromatin state of *Blimp1*, *Stella*, and *Fragilis* thus might indicate the unique expression or PGC/GC origin of ESC and the bivalent chromatin state of *Dazl* and *MVH* confers the germ-cell lineage commitment, as has been observed for other lineages (Fig. 6C). Similarly, during somatic reprogramming, addition of Oct3/4, Sox2, c-Myc, and Klf4 to somatic cells might mimic the *in-vivo* ~E3.5 blastocyst ICM cells and follows the induction of BMP signaling and hence the activation of Fragilis, Blimp1, Stella, and E-cadherin, and MET (Fig. 6B). Further, activation of the endogenous pluripotency network from the host cell genome finally establishes pluripotent cell characteristics (Fig. 6B). Finally, the chromatin state of GC/PrM markers may also reflect their transition through germ-cell fate (Fig. 6C).

In summary, we show the expression of GC/PrM markers in all analyzed pluripotent cell types and show parallel but independent maintenance of GC/PrM networks from pluripotent networks. Through our data, we propose a hypothetical model for possible germ-cell origin of ESCs and suggest the plausible transition of somatic cells through germ-cell fate to achieve pluripotency. Further genetic and epigenetic studies aimed at PGC specification during ICM outgrowth may resolve and increase our knowledge of pluripotency.

Supporting Information

Figure S1 Effect of the downregulation of Oct3/4 in ES cells. (**A**) Real time qPCR demonstrating the down-regulation of *Oct3/4* and the expression profile of other pluripotency markers (*Klf4*, *Nanog*, *Zfp206* and *Lin28*). (**B**) Expression profile of germ cell (*Stella* and *Fragilis*) and pre-meiotic (*Dazl*, *MVH* and *Stra8*) markers in *Oct3/4* down-regulated ESCs. The dotted lines indicate the normalized expression levels of analyzed genes in control siRNA treated cells. The qPCR data of two biological replicates (including three technical replicates each) were calculated and represented as a mean ±SD. Expression levels, which are statistically significant, are indicated with asterisks (***$p<0.001$).
(TIF)

Figure S2 Epigenetic signature of pluripotency and GC/PrM genes in ES cells. (**A**) Chip-seq data from the database showing that the promoter regions (red box) of pluripotency marker gene *Oct3/4* and germ cell markers *Blimp1* (*Prdm1*), *Stella* (*Dppa3*) and *Fragilis* (*Ifitm3*) representing open chromatin with abundance of active histone modification H3K4me3 (green peaks) and are depleted of repressive marks like H3K27me3 and H3K9me3 (highlighted with red and brown peaks respectively). In contrast, the promoter regions of pre-meiotic genes *MVH* (*Ddx4*), *Dazl*, *Hoxa11* and *Pax5* were marked with both active and repressive histone modification marks, signifying their bivalent chromatin structure (**B**).
(TIF)

Figure S3 Histopathological analysis identifies iPS cell-derived tumors as teratomas. Tumors grown in RAG2$^{-/-}$cγc$^{-/-}$ mice after injection of iPS cell lines #xu2 and #xu6 were HE stained. The tumors are teratomas showing ectodermal mesodermal and endodermal differentiations (* skin epithelium, # cartilage, → muscle, ▶ gut epithelium). The scale bar represents 100 µm.
(TIF)

Table S1 Primers used in RT-PCR.
(DOC)

Table S2 Quantitative real-time PCR primers for siRNA down regulation study.
(DOC)

Table S3 Quantitative real-time PCR primers used to test endogenous gene expression.
(DOC)

Table S4 List of antibodies used in Western blotting.
(DOC)

Table S5 Quantitative real-time PCR primers used in Chip assay.
(DOC)

Acknowledgments

The authors would like to thank Lukasz Smorag and Marieke Wolf for their support and valuable discussions. The authors would like to thank MSC distribution center, Texas A&M health science center for providing mouse MSCs, Prof. Shinya Yamanaka for kindly providing retroviral constructs for iPS cells generation and Dr. Michael Engelke for helping in retroviral transduction experiments.

Author Contributions

Conceived and designed the experiments: XX DVKP WE. Performed the experiments: XX SL XT TK JN RD. Analyzed the data: DVKP RD UZ. Contributed reagents/materials/analysis tools: WE. Wrote the paper: XX DVKP UZ.

References

1. Evans MJ, Kaufman MH (1981) Establishment in culture of pluripotential cells from mouse embryos. Nature 292: 154–156.
2. Martin GR (1981) Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci U S A 78: 7634–7638.
3. Matsui Y, Zsebo K, Hogan BL (1992) Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. Cell 70: 841–847.
4. Resnick JL, Bixler LS, Cheng L, Donovan PJ (1992) Long-term proliferation of mouse primordial germ cells in culture. Nature 359: 550–551.
5. Kahan BW, Ephrussi B (1970) Developmental potentialities of clonal in vitro cultures of mouse testicular teratoma. J Natl Cancer Inst 44: 1015–1036.
6. Guan K, Nayernia K, Maier LS, Wagner S, Dressel R, et al. (2006) Pluripotency of spermatogonial stem cells from adult mouse testis. Nature 440: 1199–1203.
7. Kanatsu-Shinohara M, Inoue K, Lee J, Yoshimoto M, Ogonuki N, et al. (2004) Generation of pluripotent stem cells from neonatal mouse testis. Cell 119: 1001–1012.
8. Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126: 663–676.
9. Zwaka TP, Thomson JA (2005) A germ cell origin of embryonic stem cells? Development 132: 227–233.
10. Ohinata Y, Payer B, O'Carroll D, Ancelin K, Ono Y, et al. (2005) Blimp1 is a critical determinant of the germ cell lineage in mice. Nature 436: 207–213.
11. Saiti D, Lacham-Kaplan O (2007) Mouse Germ Cell Development in-vivo and in-vitro. Biomark Insights 2: 241–252.
12. Saitou M, Barton SC, Surani MA (2002) A molecular programme for the specification of germ cell fate in mice. Nature 418: 293–300.
13. Sato M, Kimura T, Kurokawa K, Fujita Y, Abe K, et al. (2002) Identification of PGC7, a new gene expressed specifically in preimplantation embryos and germ cells. Mech Dev 113: 91–94.
14. Bowles J, Koopman P (2007) Retinoic acid, meiosis and germ cell fate in mammals. Development 134: 3401–3411.
15. Geijsen N, Horoschak M, Kim K, Gribnau J, Eggan K, et al. (2004) Derivation of embryonic germ cells and male gametes from embryonic stem cells. Nature 427: 148–154.
16. Mise N, Fuchikami T, Sugimoto M, Kobayakawa S, Ike F, et al. (2008) Differences and similarities in the developmental status of embryo-derived stem cells and primordial germ cells revealed by global expression profiling. Genes Cells 13: 863–877.
17. Qing T, Shi Y, Qin H, Ye X, Wei W, et al. (2007) Induction of oocyte-like cells from mouse embryonic stem cells by co-culture with ovarian granulosa cells. Differentiation 75: 902–911.
18. Zechner U, Nolte J, Wolf M, Shirneshan K, Hajj NE, et al. (2009) Comparative methylation profiles and telomerase biology of mouse multipotent adult germline stem cells and embryonic stem cells. Mol Hum Reprod 15: 345–353.
19. Voss AK, Thomas T, Gruss P (1997) Germ line chimeras from female ES cells. Exp Cell Res 230: 45–49.
20. Wernig M, Meissner A, Foreman R, Brambrink T, Ku M, et al. (2007) In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature 448: 318–324.
21. Okita K, Ichisaka T, Yamanaka S (2007) Generation of germline-competent induced pluripotent stem cells. Nature 448: 313–317.
22. Chazaud C, Yamanaka Y, Pawson T, Rossant J (2006) Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. Dev Cell 10: 615–624.
23. Khromov T, Krishna Pantakani DV, Nolte J, Wolf M, Dressel R, et al. (2010) Global and gene-specific histone modification profiles of mouse multipotent adult germline stem cells. Mol Hum Reprod.
24. Dressel R, Guan K, Nolte J, Elsner L, Monecke S, et al. (2009) Multipotent adult germ-line stem cells, like other pluripotent stem cells, can be killed by cytotoxic T lymphocytes despite low expression of major histocompatibility complex class I molecules. Biol Direct 4: 31.
25. Dressel R, Nolte J, Elsner L, Novota P, Guan K, et al. (2010) Pluripotent stem cells are highly susceptible targets for syngeneic, allogeneic, and xenogeneic natural killer cells. Faseb J 24: 2164–2177.
26. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, et al. (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448: 553–560.
27. Nayernia K, Lee JH, Drusenheimer N, Nolte J, Wulf G, et al. (2006) Derivation of male germ cells from bone marrow stem cells. Lab Invest 86: 654–663.
28. Carter MG, Sharov AA, VanBuren V, Dudekula DB, Carmack CE, et al. (2005) Transcript copy number estimation using a mouse whole-genome oligonucleotide microarray. Genome Biol 6: R61.
29. Efroni S, Duttagupta R, Cheng J, Dehghani H, Hoeppner DJ, et al. (2008) Global transcription in pluripotent embryonic stem cells. Cell Stem Cell 2: 437–447.
30. Meshorer E, Misteli T (2006) Chromatin in pluripotent embryonic stem cells and differentiation. Nat Rev Mol Cell Biol 7: 540–546.
31. Roeder RG (2005) Transcriptional regulation and the role of diverse coactivators in animal cells. FEBS Lett 579: 909–915.
32. Azuara V, Perry P, Sauer S, Spivakov M, Jorgensen HF, et al. (2006) Chromatin signatures of pluripotent cell lines. Nat Cell Biol 8: 532–538.
33. Bernstein E, Duncan EM, Masui O, Gil J, Heard E, et al. (2006) Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. Mol Cell Biol 26: 2560–2569.
34. Cox JL, Rizzino A (2010) Induced pluripotent stem cells: what lies beyond the paradigm shift. Exp Biol Med (Maywood) 235: 148–158.
35. Li R, Liang J, Ni S, Zhou T, Qing X, et al. (2010) A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. Cell Stem Cell 7: 51–63.
36. Samavarchi-Tehrani P, Golipour A, David L, Sung HK, Beyer TA, et al. (2010) Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. Cell Stem Cell 7: 64–77.
37. Matsui Y (2009) The molecular mechanisms regulating germ cell development and potential. J Androl 31: 61–65.
38. Saitou M (2009) Germ cell specification in mice. Curr Opin Genet Dev 19: 386–395.
39. Lawson KA, Dunn NR, Roelen BA, Zeinstra LM, Davis AM, et al. (1999) Bmp4 is required for the generation of primordial germ cells in the mouse embryo. Genes Dev 13: 424–436.
40. Okamura D, Kimura T, Nakano T, Matsui Y (2003) Cadherin-mediated cell interaction regulates germ cell determination in mice. Development 130: 6423–6430.
41. Ying Y, Liu XM, Marble A, Lawson KA, Zhao GQ (2000) Requirement of Bmp8b for the generation of primordial germ cells in the mouse. Mol Endocrinol 14: 1053–1063.

3.2. Dppa3 binds to the IG-DMR of the *Dlk1-Dio3* imprinting cluster and prevents its imprinting loss during iPS cell generation.

The observation that the expression of GC marker genes precedes the expression of pluripotency maker genes during generation of iPSCs prompted us to study their role in generation of authentic iPSCs with greater efficiencies. iPSCs are generated by the forced expression of a set of pluripotency-related transcription factors. However, the majority of mouse iPS cell lines displayed an imprinting defect at the *Dlk1-Dio3* imprinting cluster and fails to show germ line transmission. The normal expression level of *Gtl2*, a non-coding RNA from the *Dlk1-Dio3* cluster, is considered as a *bona fide* marker for authentic induced pluripotency. In the second part of this thesis, we studied the role of GC marker genes (*Blimp1*, *Ifitm3* and *Dppa3*) in establishment of authentic iPSCs and found that supplementation of *Dppa3* to the classical reprogramming factor cocktail can enhance the reprogramming and prevent the *Dlk1-Dio3* imprinting defect. At the molecular level, we show that Dppa3, the epigenetic regulator, binds to a specific region within the intergenic-differentially methylated region (IG-DMR), the imprinting control region of the *Dlk1-Dio3* cluster. Further studies showed that the binding of Dppa3 can lead to a reduced binding of Dnmt3a, a de novo DNA methyltransferase, to the IG-DMR locus thereby preventing the abnormal methylation observed in iPSCs. Taken together, our results demonstrate that *Dppa3*-mediated epigenetic regulation is essential for generating authentic iPSCs.

Authors: Xingbo Xu, Toshinobu Nakamura, Xiaoying Tan, Lukasz Smorag, Ralf Dressel, Antje Fitzner, Matthias Linke, Jessica Nolte, Ulrich Zechner, Wolfgang Engel and D.V. Krishna Pantakani

Status: Submitted to Nature Genetics journal

Author contributions to the work:

1. Xingbo Xu: Conception and design of experiments, data collection, data analysis, data interpretation, helped in drafting the manuscript.
2. Toshinobu Nakamura: Data collection, data analysis, data interpretation.
3. Lukasz Smorag: Conception, data collection, data analysis.
4. Ralf Dressel: Performed teratoma formation assay, data analysis.
5. Antje Fitzner: Data collection, data analysis.
6. Matthias Linke: Data collection, data analysis.
7. Jessica Nolte: Conception and design of experiments, data analysis.

8. Ulrich Zechner: Performed DNA methylation assay, data analysis, data interpretation, helped in drafting the manuscript.

9. Wolfgang Engel: Conception and design of experiments, financial support, helped in drafting the manuscript.

10. D.V. Krishna Pantakani: Conception and design of experiments, data analysis, data interpretation, experiment coordination, drafted the manuscript.

Dppa3 binds to the IG-DMR of the *Dlk1-Dio3* imprinting cluster and prevents its imprinting loss during iPS cell generation

Xingbo Xu[1], Toshinobu Nakamura[2], Lukasz Smorag[1], Ralf Dressel[3], Antje Fitzner[4], Matthias Linke[4], Jessica Nolte[1], Ulrich Zechner[4], Wolfgang Engel[1] and D.V. Krishna Pantakani[1‡]

[1]Institute of Human Genetics, University of Goettingen, Goettingen, Germany. [2]Department of Pathology, Graduate School of Medicine, Osaka University, Osaka, Japan. [3]Department of Cellular and Molecular Immunology, University of Goettingen, Goettingen, Germany. [4]Institute of Human Genetics, Johannes Gutenberg-University Mainz, Mainz, Germany

‡Correspondence: D. V. Krishna Pantakani, Tel: +49-551-399302; Fax: +49-551-399303; E-mail: krishna.if1@gmail.com

Abstract

Reprogramming of somatic cells into induced pluripotent stem cells (iPSCs) often results in defects at the genetic and epigenetic level. Of particular interest is the loss of imprinting at the paternally imprinted Dlk1-Dio3 cluster in a majority of the iPSC clones, which ultimately fail to show full pluripotent cell characteristics. In the present study, we analyzed the role of germ cell (GC) marker genes (Blimp1, Fragilis, and Dppa3), that have been described to become activated early during somatic cell reprogramming, in establishment of authentic pluripotency. We found that Dppa3, the guardian of the maternal genome and imprinted loci that protects them against DNA demethylation during early embryonic development, is present only in iPSC lines showing expression of Gtl2 (Gtl2on), an indicator of normal imprinting at Dlk1-Dio3 cluster, but not in iPSCs with loss of imprinting (Gtl2off). Subsequently, we found that exogenous Dppa3 together with classical reprogramming factors can efficiently reprogram somatic cells into iPSC clones that all display normal imprinting at Dlk1-Dio3 region. To dissect the mechanism by which Dppa3 can establish/maintain imprints at the Dlk1-Dio3 cluster, we performed chromatin immunoprecipitation assays with a Dppa3 antibody on embryonic stem cell (ESC) chromatin and found Dppa3 to be associated with a specific region within the intergenic differentially methylated region (IG-DMR) of Dlk1-Dio3. Further, the forced expression of Dppa3 during reprogramming lead to the maintenance of imprinting at this cluster by counteracting and reducing the binding of Dnmt3a, a de novo DNA methyltransferase that establishes DNA methylation. Interestingly, these findings are in parallel with a recent report which describes that the addition of vitamin C can enhance the reprogramming and correct the Dlk1-Dio3 imprinting defect through the inhibition of Dnmt3a binding to this region. Collectively, our results show that Dppa3 is a genetic factor necessary for preventing abnormal imprinting at the Dlk1-Dio3 region during somatic cell reprogramming.

Keywords: iPSCs, Dppa3, Dnmt3a, Gtl2, IG-DMR, Dlk1-Dio3, imprinting, Vitamin C.

Introduction

Pluripotent stem cells boast the differentiation potential virtually into any cell type of the body, and hence hold a great promise for regenerative medicine applications[1]. Recent advances in reprogramming strategies unveiled the induction of pluripotency in somatic cells by using few transcription factors resulting in the generation of induced pluripotent stem cells (iPSCs)[2,3]. These iPSCs are morphologically and functionally similar to ESCs,

although they frequently show minor differences in gene expression, DNA methylation, imprinting of the *Dlk1-Dio3* cluster and gene copy number[4-7]. The transcriptome analysis of iPSCs in comparison to ESCs revealed that the expression of coding and non-coding genes encoded by the *Dlk1-Dio3* imprinting cluster is misregulated due to the aberrant acquisition of DNA methylation at the maternal allele along with the normally methylated paternal allele[7]. The iPSCs with defects in genomic imprinting are developmentally compromised and show reduced chimerism after injection into blastocysts and no germline transmission[7,8]. Recently, addition of vitamin C (Vc) during reprogramming was shown to result in iPSCs with normal *Dlk1-Dio3* imprinting, yet the factor expressed in a Vc-dependent manner was not identified[9].

Genomic imprinting is an epigenetic phenomenon established during gametogenesis and involves differential DNA methylation and post-translational histone modifications. Short DNA sequences called imprinting control regions (ICRs) are methylated on either the maternal or paternal allele to regulate expression of the imprinted gene in *cis*[10]. These marks lead to exclusive or preferential parent-specific monoallelic expression of imprinted genes[11,12]. During embryonic development a pool of primordial germ cells (PGCs) gives rise to progenitors of adult gametes. After their specification and upon their arrival at the genital ridge, these PGCs undergo demethylation of the whole genome including an erasure of parent-specific methylation marks of imprinted genes, the so-called imprints[13]. The reestablishment of genomic imprinting in germ cells according to the sex of the embryo is initiated after the entry of PGCs into gonads and continues through germ cell differentiation[14,15].

The imprint acquisition according to the sex of the embryo is regulated by several components such as primary sequence specificity, chromatin configuration, non-histone proteins and transcriptional events (reviewed by Arnaud, 2010[16]). Specifically, the *de novo* DNA methyltransferase Dnmt3a and its related protein Dnmt3l, which has no methyltransferase activity, were shown to be essential for imprint establishment at several imprinted loci[17-19]. In addition to the DNA methylation machinery, several other DNA binding proteins such as Zfp57, Nlrp2, Nlrp7, Ctcfl and Prmt7 are implicated in the establishment of imprints in a sex-specific manner[20-24]. Once established in germ cells, several factors are known to faithfully maintain and transmit the imprints during the early stages of embryogenesis to all somatic lineages (reviewed by Arnaud, 2010[16]). *Dppa3*/*PGC7* is one such factor expressed mainly in germ cells and known to protect some of the maternal as well as paternal imprints during the wave of DNA demethylation occurring in early embryogenesis[25].

Previously, we have shown that germ cell (GC) marker genes such as *Blimp1*, *Fragilis* and *Dppa3* are expressed in all pluripotent cell types and emerge early during somatic cell reprogramming into iPSCs[26]. In the present study, we evaluated whether these GC marker genes, in particular *Dppa3*, play any role in the maintenance of *Dlk1-Dio3* imprinting during the generation of iPSCs. Interestingly, we found the expression of *Blimp1* and *Fragilis* in iPSCs with normal as well as defective *Dlk1-Dio3* imprinting, whereas *Dppa3* was expressed only in cells having normal *Dlk1-Dio3* imprinting. We then reprogrammed MEFs in the presence of *Dppa3*, in addition to Yamanaka's reprogramming factors (OSKM), and found that all iPSC colonies show normal imprinting at the *Dlk1-Dio3* region. These iPSCs were morphologically and functionally similar to ESCs and displayed all characteristics of pluripotent cells including germline transmission. At the molecular level, we observed that Dppa3 is associated with the intergenic differentially methylated region (IG-DMR) at the *Dlk1-Dio3* region which counteracts the binding of Dnmt3a binding to this region during reprogramming. Further, we found that vitamin C can activate the *Dppa3*

expression, which lets us suggest that *Dppa3* is the genetic component essential for the proper establishment/maintenance of *Dlk1-Dio3* imprinting during iPSC generation.

Materials and methods

Cell culture

Mouse ESCs and iPSCs were cultured in Dulbecco's Modified Eagle Medium (DMEM) (PAN, Germany) supplemented with 15% fetal calf serum (FCS) (PAN, Germany), 2mM L-glutamine (PAN, Germany), 50µM β-mercaptoethanol (Gibco/Life Technology, Germany), 1% non-essential amino acids (NEAA) (Gibco/Life Technologies, Germany), 1% sodium pyruvate (Gibco/Life Technologies, Germany), 1% penicillin/streptomycin (PAN, Germany) and 1000U/ml leukemia inhibitory factor (LIF) (Millipore, Germany). For reprogramming studies with low serum medium, the StemPro medium (Invitrogen) supplemented with StemPro supplement (Invitrogen), 1% FCS, 2mM L-Glutamine, 1mM Sodium pyruvate, 1% NEAA, 100µM β-mercaptoethanol, and 1000U/ml LIF was used either in the presence or absence of 50µg/ml vitamin C (Sigma).

Construction of expression vectors

For generating a retroviral expression vector of Fragilis, the ORF was PCR-amplified from cDNA of mouse ESCs and cloned into the pGEM-T Easy vector (Life Technologies, Germany). After verification by sequencing, the Fragilis ORF was cloned into *Eco*RI restriction sites of pMXs-hOct3/4[3] by replacing hOct3/4. For generating the pMXs-Blimp1 expression vector, the *Eco*RI and *Xho*I fragment of HA-tagged Blimp1 from pMSCV-HA-Blimp1-IRES-GFP[27] was subcloned into *Eco*RI and *Xho*I restriction sites of pMXs-hOct3/4[3] by replacing hOct3/4. The expression of each construct was verified by transfection into Plat-E cells followed by Western blot analysis. The constructs pMXs-hOct3/4 and pMXs-Dppa3 were obtained from Addgene.

Generation of iPS cells

We used retroviral expression vectors for Oct3/4, Sox2, Klf4, and c-Myc[2] together with that of either Fragilis, Dppa3, or Blimp1 to reprogram MEFs into iPSCs. Briefly, MEFs isolated from transgenic *Nanog*-EGFP mice[28] were transduced with retroviral particles as described[2]. To establish iPS cell lines, colonies which appeared after 8-12 days of transduction were picked and cultured in 24-well plates with standard ESC culture conditions and were monitored for the ESC-like morphology using an Olympus-IX71 inverted microscope (Olympus, Germany). Further, several independent iPSC lines were established from each combination to examine the expression of *Gtl2* as well as for further characterization. Alternatively, iPSCs were generated in low serum medium as described above.

Immunostaining and alkaline phosphatase staining

Immunostaining was performed as described previously[29] using mouse monoclonal antibodies to SSEA1 (ab16285, Abcam). Cytochemical staining for alkaline phosphatase activity was performed using Leukocyte Alkaline Phosphatase Kit (Sigma-Aldrich).

Protein extraction and Western blotting

Total protein extracts were prepared using lysis buffer (10mM Tris–HCl (pH 8.0), 1mM EDTA, 2.5% SDS, 100mM PMSF) containing protease inhibitor cocktail (Roche). Protein samples were resolved on 4–12% SDS-PAGE (Life Technologies, Germany) and transferred onto nitrocellulose membrane (Amersham Biosciences). Membranes were processed using standard Western blot protocols, and signals were detected using a chemiluminescent kit (Santa Cruz Biotechnology).

Antibody sources are listed in supplementary tables (Supplementary Table S1).

Total RNA extraction and quantitative real time PCR analysis

Total RNA was extracted from cells using NucleoSpin miRNA kit (Machery-Nagel, Germany) following the manufacturer's protocols. For the mRNA quantification experiment, 5µg of total RNA was converted into cDNA using the SuperScript II system (Life sciences, Germany). For the miRNA quantification assay, 1µg of total RNA was used for cDNA synthesis using miScript II RT Kit (Qiagen, Germany). For qPCR analysis, diluted cDNA (1/10) was used as a template in QuantiFast SYBR Green (Qiagen, Germany) based method and run on ABI7900HT Real-Time PCR System (Applied Biosystems, Germany). The qRT-PCR data were normalized to housekeeping genes (*Hprt* and *Gapdh*) and represented as percentage of expression of housekeeping genes. Primers used in RT-PCR and qPCR are listed in Supplementary Table S2.

Genomic DNA isolation and methylation analysis of the *Gtl2* IG-DMR

Genomic DNA isolation from cells and bisulfite pyrosequencing were performed as previously described[30]. Primers used for analysis of the *Gtl2* IG-DMR are listed in Supplementary Table S2.

Chromatin immunoprecipitation assay

Chromatin immunoprecipitation (ChIP) assays were performed on ESCs and MEFs undergoing reprogramming using the Diagenode OneDay ChIP kit. Briefly, the cells (3×106 cell per pull down) were cross-linked using formaldehyde and lysed with the shearing kit (Diagenode, Belgium) followed by sonication with a bioruptor (Diagenode, UCD-200 TM) to obtain an average chromatin size of 400 bp. Then, the sheared chromatin was immunoprecipitated with 5µg of antibodies as indicated and the Diagenode OneDay ChIP protocol was used to extract the chromatin bound to each specific antibody. Rabbit IgG was used as a control in mock ChIP experiments. Quantitative analysis was performed using real-time qPCR and the primers listed in supplementary table S2. Ct values were determined for both immunoprecipitated DNA and a known amount of DNA from the input sample for different primer pairs. The amount of immunoprecipitated chromatin corresponding to a particular genomic locus was then calculated from the qPCR data using the delta-delta Ct method (ddCt) and reported as fold enrichment relative to the background (IgG). All fold enrichment values are means of three independent experiments.

Teratoma formation assay

The teratoma formation assay was performed as previously described[31,32]. Briefly, iPS cells (1×10^6 cells) were injected subcutaneously into $RAG2^{-/-}c\gamma c^{-/-}$ mice lacking T, B, and natural killer (NK) cells. Tumor growth was monitored weekly by palpation and size was recorded using linear calipers. Animals were sacrificed when a tumor diameter of 1 cm was reached. Autopsies were performed and tumor tissue was placed in phosphate-buffered 4% formalin for 16 h and then embedded in paraffin. For histological analysis, the specimens were stained with hematoxylin and eosin (HE).

Statistical Analysis

All qPCR data for RNA expression analysis (two or more biological replicates) were calculated using the standard curve method. The GraphPad Prism 4.0 software was used to calculate the statistical significance.

Results

Dppa3 expression positively correlates with the activation of *Gtl2* in iPS cells

The loss of *Dlk1-Dio3* imprinting was reported to be a frequently occurring phenomenon during

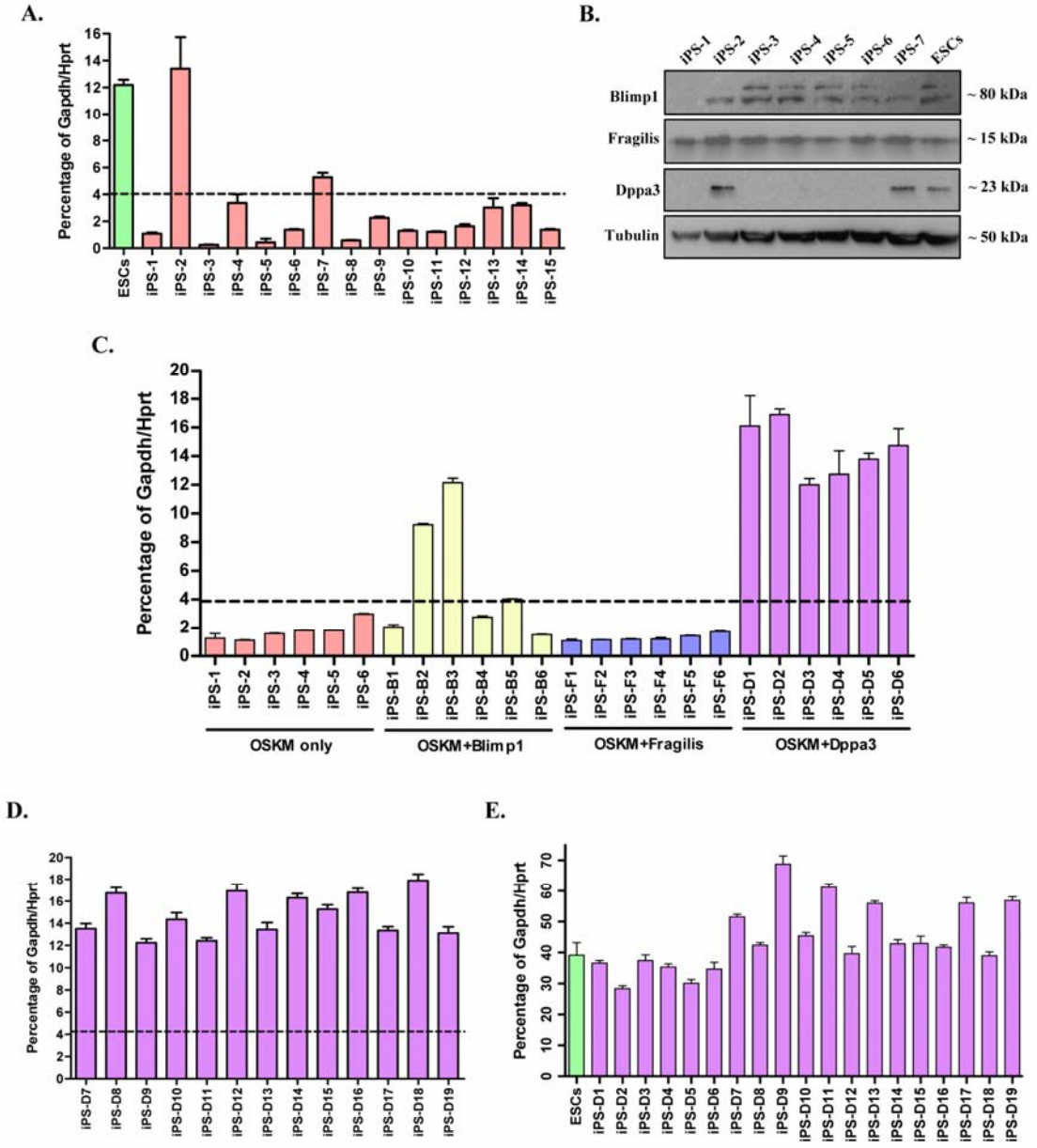


Figure 1. Expression of *Dppa3* positively correlates with proper expression of imprinted genes of from the *Dlk1-Dio3* cluster in iPSCs. (A) qRT-PCR data showing the expression of *Gtl2* above the threshold (dotted line[9]) only in 2/15 iPSCs generated using classical Yamanaka factors (OSKM). The expression of *Gtl2* in ESCs was used as a control. **(B)** Western blot analysis for Blimp1, Fragilis, and Dppa3 expression in iPSCs generated using OSKM. The protein lysate derived from ESCs was used as a control. **(C)** *Gtl2* expression analysis in iPSCs generated using either OSKM or OSKM in combination with indicated germ cell marker genes (Blimp1 (B), Fragilis (F), and Dppa3 (D)). **(D)** qRT-PCR data showing the expression of *Gtl2* in additional iPSC clones generated using OSKM+D. **(D)** qRT-PCR data showing the expression of *Dlk1* in iPSC clones generated using OSKM+D. The expression of *Dlk1* in ESCs was used as a control.

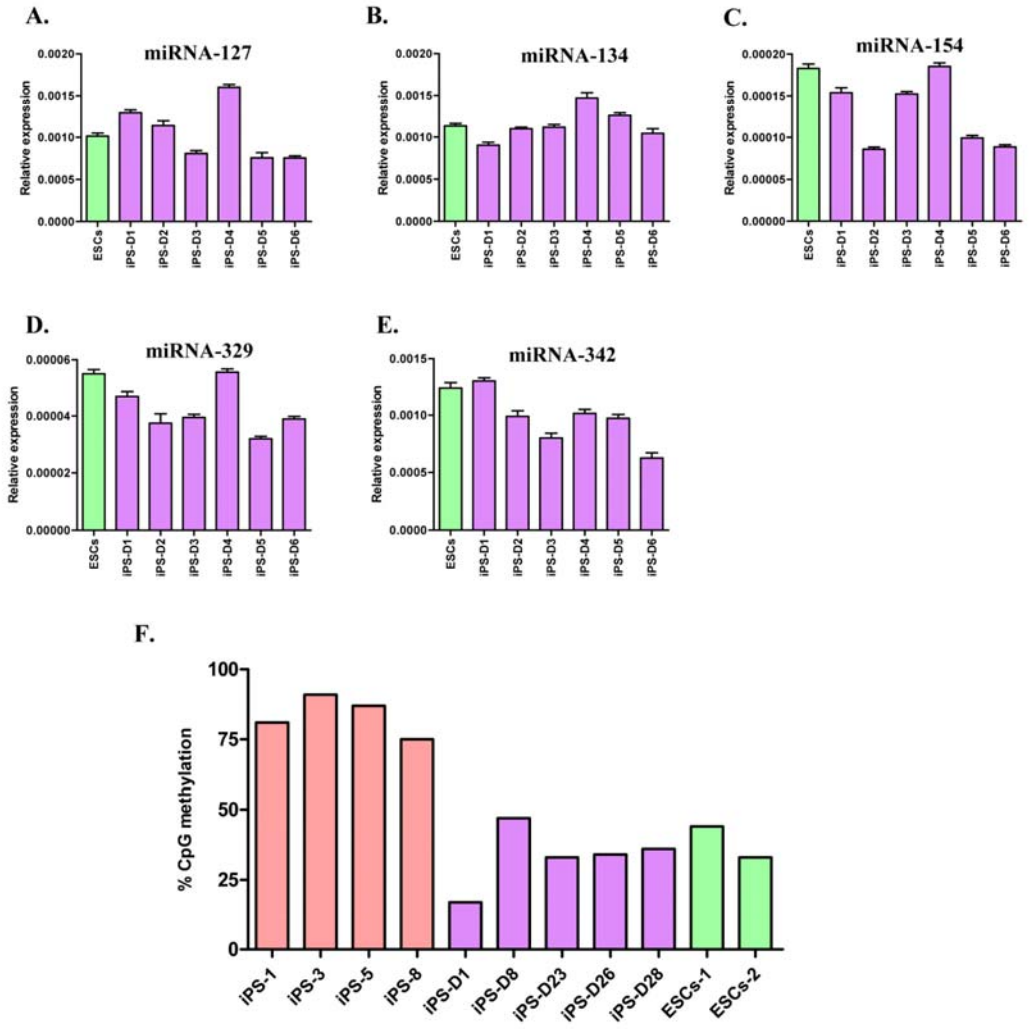


Figure 2. Expression of miRNAs encoded by the *Dlk1-Dio3* cluster and methylation of the *Gtl2* IG-DMR in iPSCs generated using *Dppa3*. **(A-C),** qRT-PCR data showing the expression of miRNA-127 **(A),** miRNA-134 **(B),** miRNA-154 **(C),** and miRNA-329 **(D)** in OSKM+D iPSCs. cDNA derived from ESCs was used as a control for miRNA expression. **(E)** DNA methylation analysis of the *Gtl2* IG-DMR (4 CpG sites) in OSKM+D iPSCs and OSKM only *Gtl2*^{off} iPSCs; genomic DNA from ESCs was used as a control.

mouse iPSC generation[7]. The iPSCs with aberrant imprinting at the *Dlk1-Dio3* cluster failed to generate chimeras with germline transmission, one of the central characteristic of pluripotent cells[7,8]. Since then, the activation/expression of *Gtl2*, a maternally expressed non-coding RNA of this imprinting cluster is used as a marker to identify the authentic iPSC colonies. To decipher the underlying cause of aberrant imprinting, we established several iPSCs lines from mouse embryonic fibroblasts (MEFs) using the classical Yamanaka approach. We found only 2/15 clones (iPS-2 and iPS-7) having *Gtl2* expression (*Gtl2*^{on}) above the threshold level[9] (Fig. 1A), further indicating that the defect at *Dlk1-Dio3* imprinting cluster is a common outcome during somatic cell reprogramming. Recently, we have

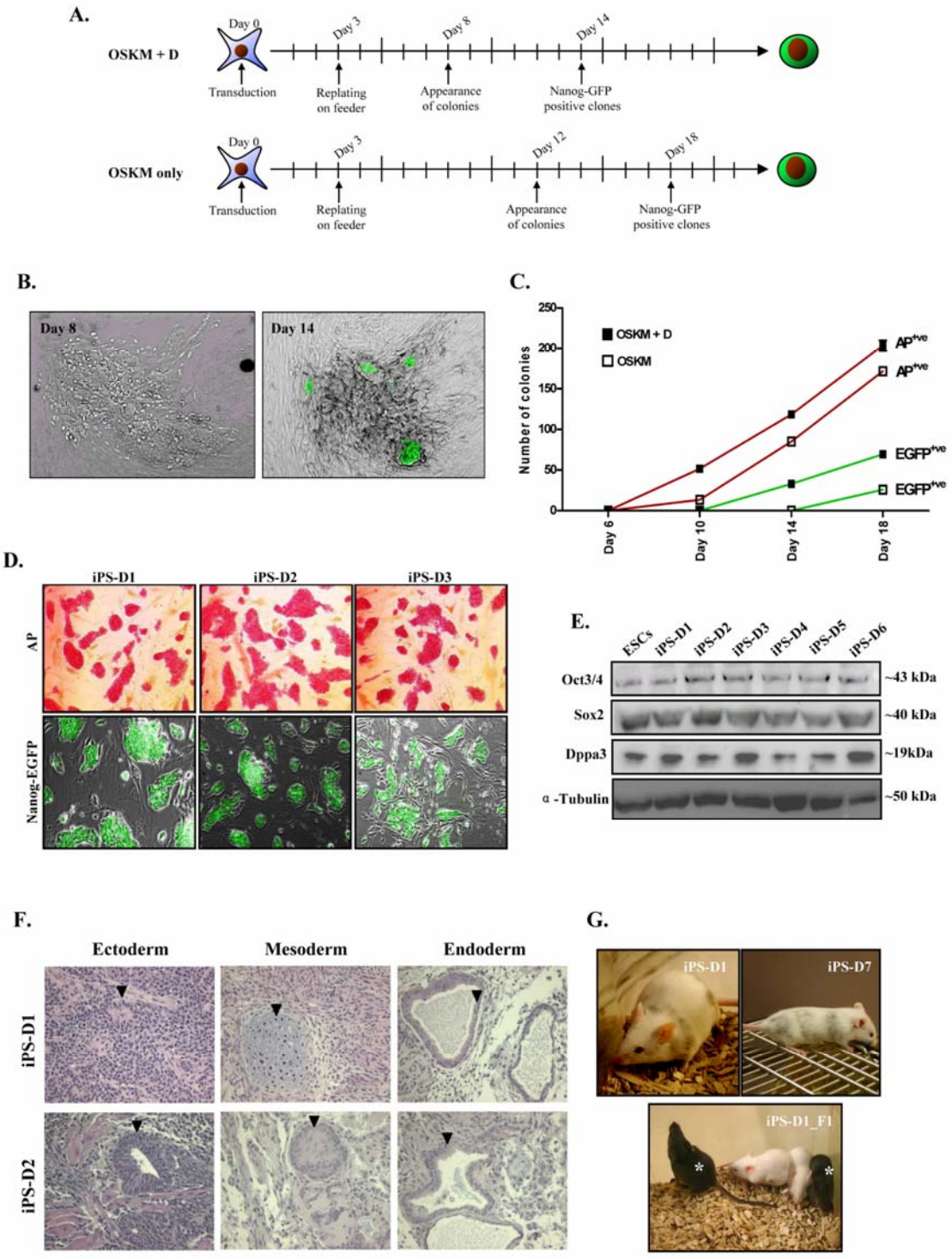


Figure 3. Dppa3 enhances the reprogramming and generates fully pluripotential iPSCs. (A) Time scale showing the start of viral transduction followed by appearance of ESC-like and *Nanog*-EGFP-positive colonies during iPSCs generation using either OSKM or OSKM in combination with Dppa3 (OSKM+D). **(B)** Bright-field image showing the appearance of ESC-like colonies already by day 8 of reprogramming using

OSKM+D (left panel). The same colony was positive for *Nanog*-driven EGFP at day 14 (right panel). **(C)** Line graph showing the number of AP-positive (AP^{+ve}) and EGFP-positive ($EGFP^{+ve}$) colonies during the reprogramming time-course in the presence of OSKM (open box) or OSKM+D (closed box). **(D)** iPSC colonies (iPS-D) generated using OSKM+D show AP activity (top panel) and activate the pluripotency reporter *Nanog*-EGFP (bottom panel). **(E)** Western blots showing the expression of pluripotency marker genes Oct3/4 and Sox2 as well as Dppa3 in OSKM+D iPSCs (iPS-D). Protein extracts from ESCs were used as a control. **(F)** Images showing the presence of cell type derivatives of all three germ layers in teratomas obtained from OSKM+D iPSCs. Arrowheads indicate the presence of the representative tissue/cell type from the respective germ layer. **(G)** Images showing the chimeras obtained from OSKM+D iPSCs (upper panel) and their germline transmission (lower panel). Asterisks indicate the pups derived from the germline competent OSKM+D iPSCs.

shown that the expression of germ cell (GC) marker genes such as *Blimp1*, *Fragilis* and *Dppa3* precedes the expression of endogenous pluripotency-related genes, such as *Oct3/4*, *Sox2*, *Nanog* and *Klf4*, during somatic cell reprogramming[26]. To identify whether these GC marker genes have any role in establishment/maintenance of *Dlk-Dio3* imprinting, we analyzed several iPSC colonies for the expression of GC marker genes (Fig. 1B). The Western blot analysis for the expression of Blimp1 and Fragilis did not show any differences between $Gtl2^{on}$ and $Gtl2^{off}$ clones (Fig. 1A, B). However, the expression of Dppa3 was observed only in $Gtl2^{on}$ clones but not in $Gtl2^{off}$ clones (Fig. 1A, B), indicating that the expression of Dppa3 might be crucial for the establishment/maintenance of *Dlk1-Dio3* imprinting, at least during reprogramming.

To comparatively analyze the effect of GC marker gene expression on maintenance of *Dlk1-Dio3* imprinting, we generated a *Fragilis* retroviral construct and verified its expression (Supplementary Fig. S1A), while the generation of *Blimp1* and *Dppa3* constructs was described previously[2,27]. The transduction of MEFs with each of the GC markers (*Blimp1*, *Fragilis*, and *Dppa3* abbreviated as B, F, and D, respectively) along with Yamanaka factors (*Oct3/4*, *Sox2*, *Klf4*, and *c-Myc* abbreviated as OSKM, respectively) resulted in the establishment of iPS colonies, which are morphologically similar to ESCs and positive for alkaline phosphatase (ALP) activity and expressed *Nanog* promoter-driven EGFP (data not shown). The *Gtl2* expression analysis in 6 individual iPSC colonies from each combination of transductions revealed that the addition of *Dppa3* to the OSKM results in all $Gtl2^{on}$ colonies (6/6), while *Blimp1* gave rise to only 3/6 $Gtl2^{on}$ colonies (Fig. 1C). Further, 6/6 iPS colonies generated with either OSKM alone or OSKM supplemented with *Fragilis* showed *Gtl2* expression levels below the threshold limit (Fig. 1C).

To confirm the faithful maintenance of imprinting at *Dlk1-Dio3* cluster by *Dppa3* (D) supplementation, we next generated additional iPS clones with OSKM+D and found all of them to express *Gtl2* above the threshold level (Fig. 1D). We checked the expression of *Dlk1*, the paternally expressed gene in this imprinting cluster, and found its levels comparable to those in ESCs (Fig. 1E). The *Dlk1-Dio3* imprinting cluster is also known to harbor several maternally expressed miRNAs which are underexpressed in $Gtl2^{off}$ clones[7]. Hence, we analyzed the expression of *Dlk1-Dio3* cluster-encoded miR-127, miR-134, miR-154, miR-329, and miR-342 in OSKM+D iPSCs and found them to be expressed in all iPSCs (Fig. 2A-E). Then, we verified if the presence and absence of *Gtl2* mRNA in OSKM+D iPSCs and OSKM $Gtl2^{off}$ iPSCs, respectively, is associated with differences in DNA methylation levels at the *Gtl2* IG-DMR, the ICR of the *Dlk1-Dio3* imprinting cluster. Indeed, we detected a moderate hypomethylation of the IG-DMR

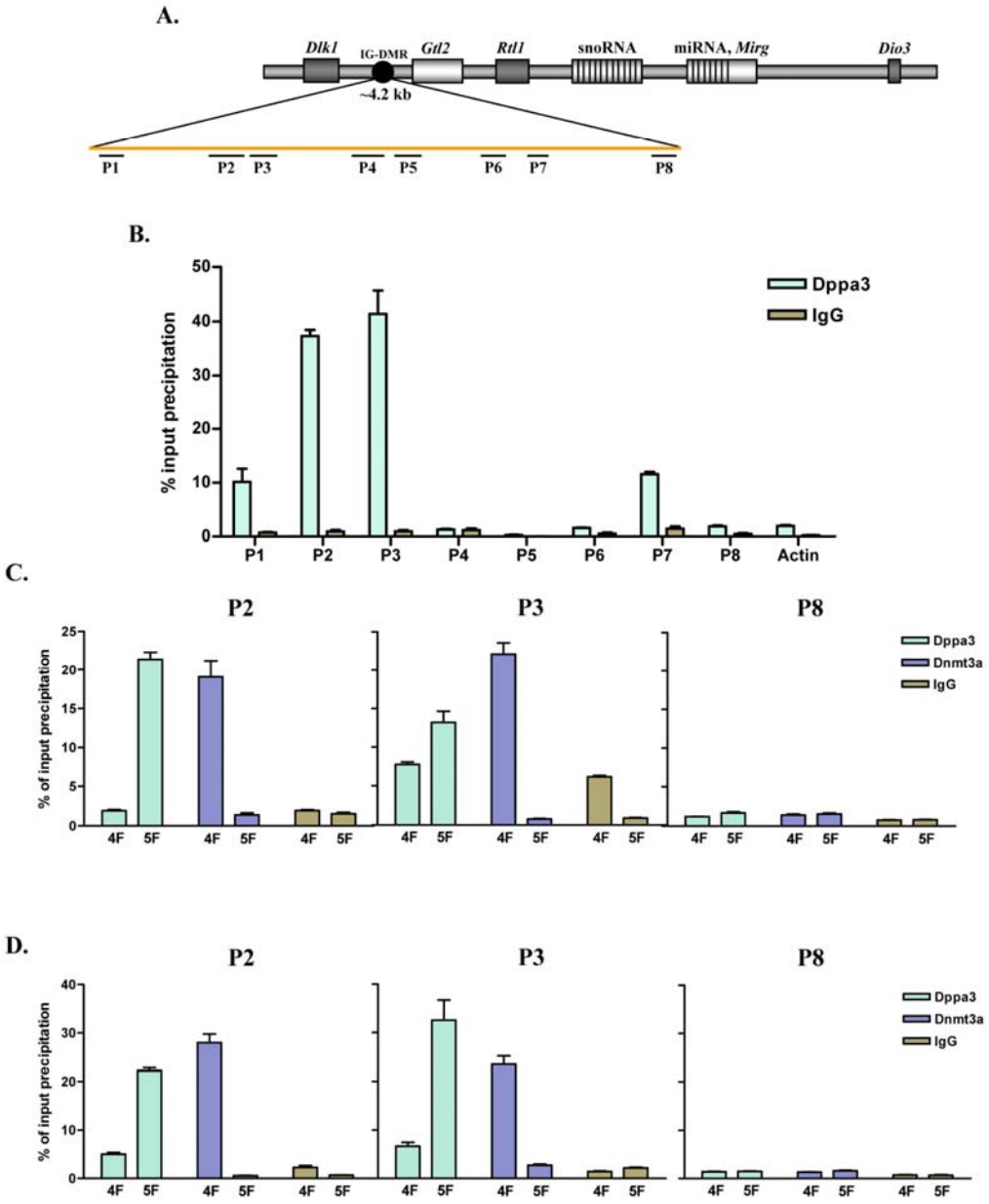


Figure 4. Dppa3 binds to a specific region within the IG-DMR of the *Dlk1-Dio3* imprinting cluster. (A) Simplified schematic diagram showing the *Dlk1-Dio3* imprinting cluster along with paternally expressed genes (dark grey boxes) and maternally expressed genes (light grey boxes). The imprinting control region, intergenic differentially methylated region (IG-DMR) (black circle), that spans ~4.2 kb, and the 8 PCR amplicons (P1-P8) covering the entire IG-DMR are indicated. **(B)** Chromatin immunoprecipitation (ChIP) followed by qPCR analysis for Dppa3 (cyan) binding sites across the 8 PCR amplicons of the IG-DMR in ESCs. ChIP with IgG (olive) was used as a control. **(C)** ChIP followed by qPCR analysis for Dppa3 (cyan) and Dnmt3a (blue) binding at the P2, P3 and P8 amplicons of the IG-DMR in OSKM (4F) and OSKM+D (5F) at 9 days of reprogramming. ChIP with IgG (olive) was used as a control. **(D)** ChIP followed by qPCR analysis for Dppa3 (cyan) and

Dnmt3a (blue) binding at the P2, P3 and P8 amplicons of the IG-DMR in OSKM (4F) and OSKM+D (5F) at 12 days of reprogramming. ChIP with IgG (olive) was used as a control.

in all analyzed OSKM+D iPSCs compared to a hypermethylation of the IG-DMR in all analyzed OSKM *Gtl2*^{off} iPSCs (Fig. 2F). Collectively, these results suggest that *Dppa3* expression can prevent aberrant hypermethylation of the IG-DMR and *Dlk1-Dio3* imprinting defects during MEF reprogramming.

Dppa3 enhances the reprogramming and contributes to the generation of all authentic iPSC colonies

During the course of reprogramming, the OSKM+D-transduced MEFs appeared as early colonies already by day-8 in contrast to the OSKM-only-transduced cells, which showed colony morphology only by day 12 (Fig. 3A, B). Moreover, OSKM+D-reprogrammed cells activated the *Nanog*-EGFP by day 14; whereas OSKM-treated cells were EGFP-positive only at around day 18 of the reprogramming (Fig. 3A, B). Addition of Dppa3 to the reprogramming factors greatly enhanced the number of AP- and EGFP-positive iPSC colonies (Fig. 3C). We characterized several iPSC clones derived from Dppa3-treated cells (iPS-D) for the expression of ALP, *Nanog*-EGFP and pluripotency marker genes such as Oct3/4 and Sox2 as well as Dppa3. The studied iPS-D clones were positive for all analyzed pluripotent cell characteristics (Fig. 3D, E). To test the teratoma formation ability of iPS-D cells, we injected two different iPSC clones (iPS-D1 and iPS-D2) into immunodeficient mice and found teratoma consisting of all three germ-layer derivatives, confirming the pluripotency of the two iPSC clones (Fig. 3F). Furthermore, the injection of iPS-D clones into blastocysts resulted in chimeras (Fig. 3G) and chimeras generated from iPS-D1 were tested for their germline transmission ability. Offspring from iPS-D1 were obtained (Fig. 3G) and a genotyping PCR confirmed the presence of OSKM+D constructs (Supplementary Fig. S2A).

Next, we investigated whether Dppa3 can replace any of the reprogramming factors (OSKM) during somatic cell reprogramming. MEFs reprogrammed in the absence of each of the reprogramming factors but instead supplemented with Dppa3 failed to generate any AP-positive colonies by day 18 of reprogramming (data not shown). Although we obtained few smaller colonies in the absence of c-Myc, it could not be due to Dppa3 as it was already reported that c-Myc is dispensable for the reprogramming process[33]. These results suggest that Dppa3 works in concert with other reprogramming factors but might not be able to drive the reprogramming process without any of these factors.

Dppa3 physically interacts with the IG-DMR of the *Dlk1-Dio3* cluster

Previously, Dppa3 was shown to be essential for protecting maternal chromosomes against DNA demethylation during early post-fertilization development[25]. It was also reported that Dppa3 protects against DNA demethylation at several but not all imprinted genes/regions[25]. To test the hypothesis that Dppa3 is physically associated with the IG-DMR of the *Dlk1-Dio3* region, we performed a chromatin immunoprecipitation (ChIP) assay with an antibody against Dppa3 on ESC chromatin. Since the IG-DMR of the *Dlk1-Dio3* cluster spans around 4.2kb, we designed 8 PCR amplicons (P1-P8) covering the entire IG-DMR region (Fig. 4A). The qPCR analysis of immunoprecipitated chromatin revealed that Dppa3 binds specifically to the IG-DMR region represented by PCR amplicons P2 and P3 PCR (Fig. 4B). Recent data showed that Dnmt3a is highly enriched at the IG-DMR during somatic cell reprogramming, probably leading to the

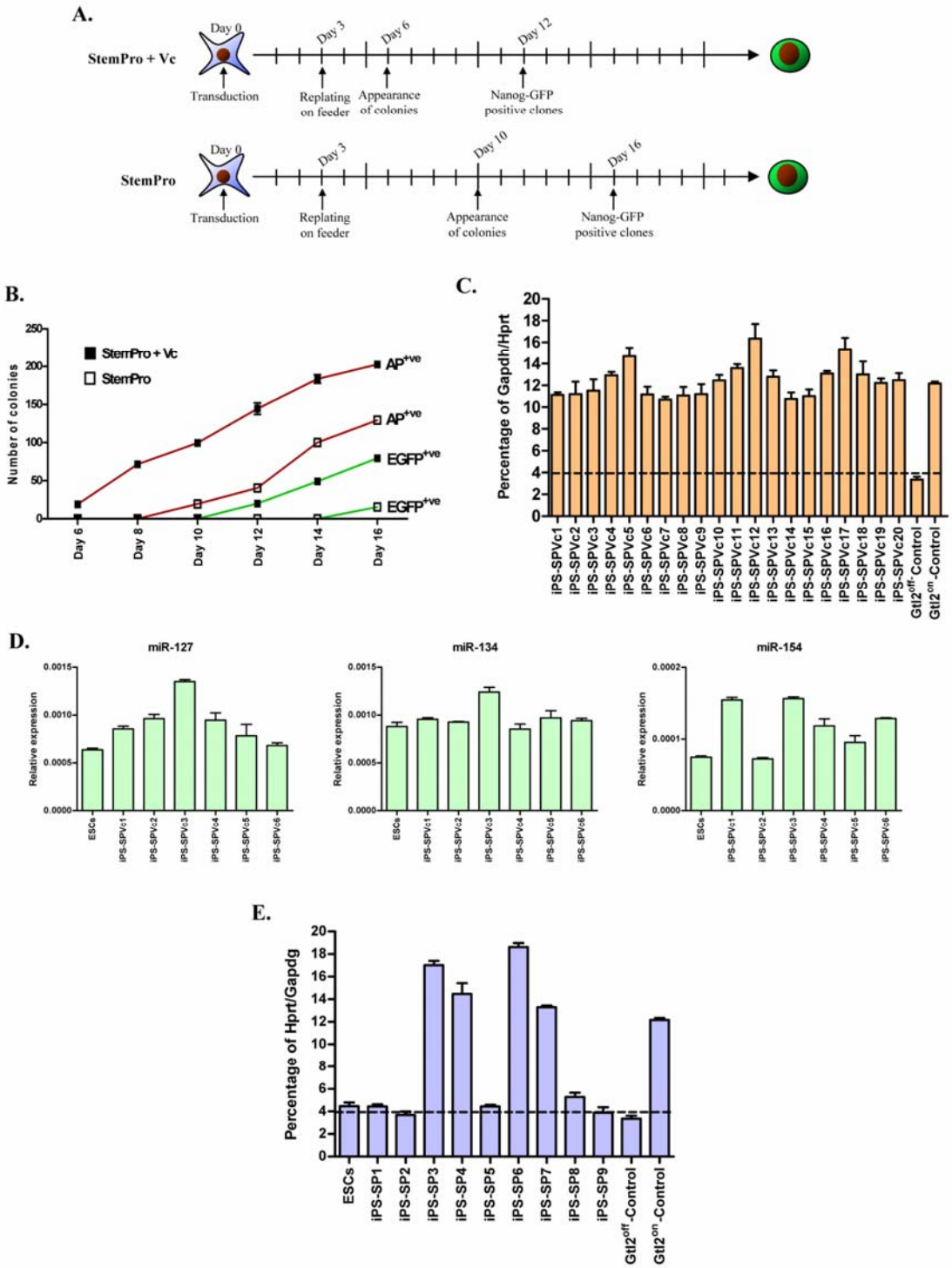


Figure 5. Vitamin C enhances the reprogramming and generates fully pluripotential iPSCs in low serum medium. (A) Time scale showing the start of viral transduction followed by appearance of ESC-like and *Nanog*-EGFP-positive colonies during iPSC generation using OSKM in StemPro (low serum medium) supplemented with Vc (SPVc) or StemPro without Vc (SP). **(B)** Line graph showing the number of AP^{+ve} and

EGFP^{+ve} colonies during the reprogramming time course in the presence of SP (open box) or SPVc (closed box). **(C)** qPCR data showing the expression of *Gtl2* in iPSC clones generated using StemPro+Vc (iPS-SPVc). Previously generated *Gtl2*^{off} and *Gtl2*^{on} iPSC clones were used as controls. **(D)** qRT-PCR data showing the expression of miRNA-127 (left panel), miRNA-134 (middle panel), and miRNA-154 (right panel) in iPS-SPVc clones. ESCs were used as a control for miRNA expression. **(E)** qPCR data showing the expression of *Gtl2* in iPSC clones generated using SP only (iPS-SP). Previously generated *Gtl2*^{off} and *Gtl2*^{on} iPSC clones with *Gtl2*^{off} and *Gtl2*^{on} were again used as controls.

silencing of the maternal allele and that this effect can be counteracted by the addition of ascorbic acid (vitamin C, Vc)[9]. This led us to analyze whether exogenous Dppa3 can also counteract Dnmt3a binding at the IG-DMR. Interestingly, the association of Dnmt3a with the IG-DMR PCR amplicons P2 and P3 was significantly reduced in OSKM+D-transduced cells at day-9 of reprogramming, whereas OSKM only-transduced cells showed enrichment for Dnmt3a at this region (Fig. 4C). Similar results were obtained at day-12 of reprogramming (Fig. 4D), confirming that Dppa3 protects the IG-DMR of the *Dlk1-Dio3* locus against DNA hypermethylation by counteracting Dnmt3a binding.

These results prompted us to evaluate the expression of *Gtl2* and *Dlk1* in *Dppa3*-deficient ESCs[25]. qRT-PCR results indicated no obvious differences in expression levels of *Glt2* and *Dlk1* between *Dppa3*-deficient and control ESCs (data not shown). Further, DNA methylation analysis of the *Dlk1-Dio3* IG-DMR revealed no significant differences between *Dppa3*-deficient and control ESCs, whereas the DNA methylation of *Rasgrf1*, another paternally methylated locus, showed hypomethylation in *Dppa3*-knockout cells (Supplementary Fig. S3).

Modified-StemPro medium enhances the reprogramming and results in the generation of exclusively *Gtl2*^{on} iPSCs

During the course of our study, we also investigated whether StemPro-based low-serum medium (SPVc), which we regularly use for the culture of spermatogonial stem cells (SSCs)[34], can generate all *Gtl2*^{on} iPSCs with greater efficiency. To this end, we reprogrammed MEFs with OSKM in SPVc medium and found ESC-like colonies already by day 6 of reprogramming which finally resulted in a 2-fold increase in AP- and *Nanog*-EGFP-positive colonies (Fig. 5A and B). Subsequently, we checked the expression of *Gtl2* in several iPSC lines, which were generated using SPVc medium (iPSC-SPVc), and found all of them to be *Gtl2*^{on} iPSCs (Fig. 5C). Further, to validate the activation of *Dlk1-Dio3* locus, we have analyzed the expression of three maternally expressed miRNAs (miR-127, -134, -154) and identified them to be expressed in all analyzed iPSC-SPVc (Fig. 5D). These results led us to hypothesize that the components of SPVc help in proper maintenance of imprinting at *Dlk1-Dio3* locus during reprogramming. Our SPVc medium inherently contains vitamin C (Vc) that was reported to protect against the los of *Dlk1-Dio3* imprinting during reprogramming[9]. Hence, we asked whether StemPro-based low-serum medium depleted from Vc (SP) can also generate exclusively *Gtl2*^{on} iPSCs with similar reprogramming kinetics like SPVc. However, the use of SP medium resulted in slower reprogramming kinetics and only 4 of the 9 generated iPSC clones had the Gtl2 expression markedly above the threshold (Fig. 5B and 5E). Our results are in line with observations of Stadtfeld et al. (2012)[9], indicating that Vc can not only enhance the reprogramming but also prevent the loss of *Dlk1-Dio3* imprinting.

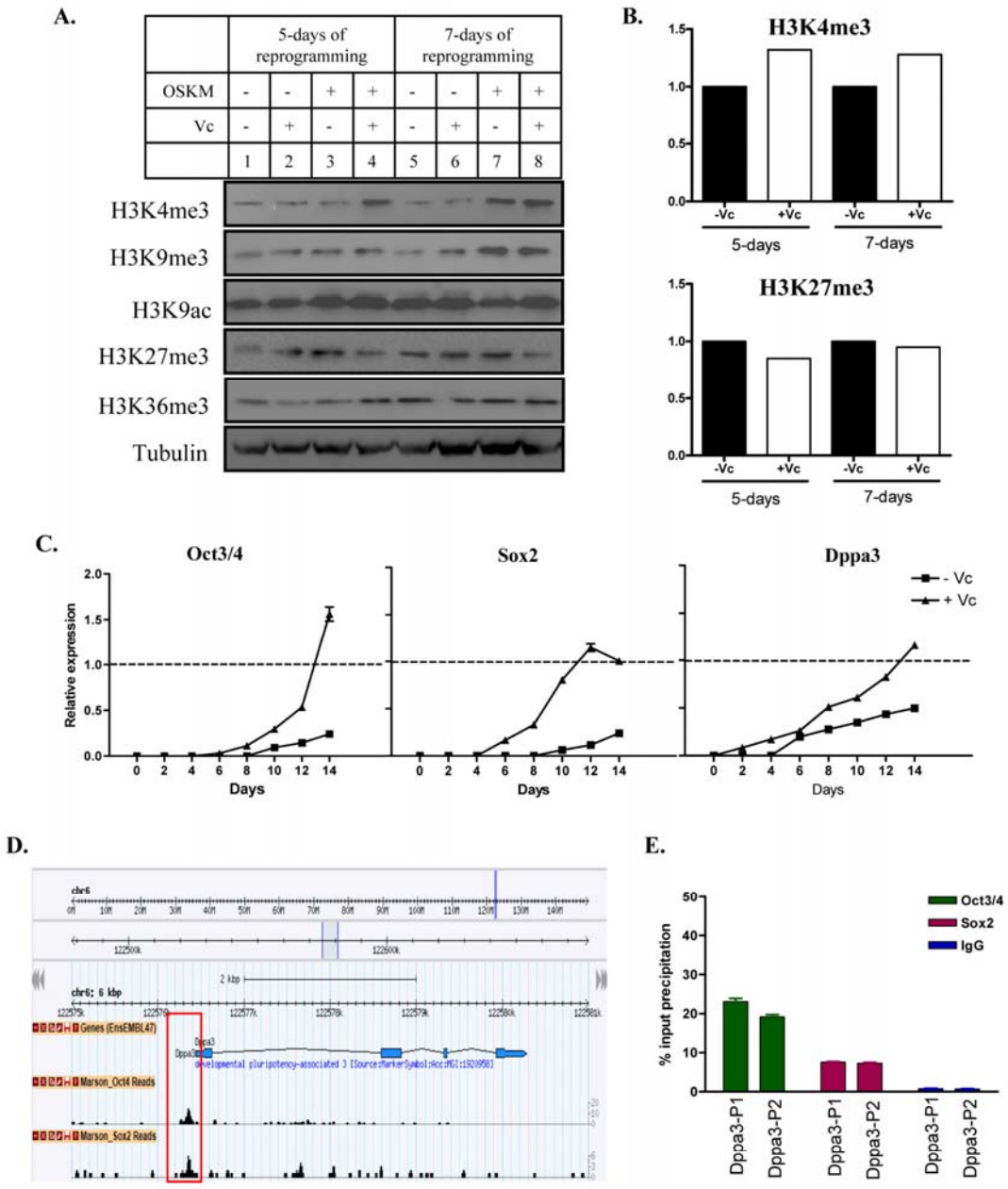


Figure 6. Vitamin C mediates the global chromatin relaxation and early-activation of pluripotency-related genes. (A) Western blots showing the expression of various histone modifications during OSKM reprogramming (5 days and 7 days) in standard ESC culture medium and either presence or absence of Vc (+ or -, respectively). MEFs supplemented with no OSKM and Vc (lanes 1 and 5) or either Vc only (lanes 2 and 6) or OSKM only (lanes 3 and 7) were used as controls. All blots were reprobed with an anti-tubulin antibody and one representative blot is shown as a loading control. **(B)** Densitometric quantification of H3K4me3 and H3K27me3 band intensities at 5 days of reprogramming (lanes 3 and 4 in **A**) and 7 days of reprogramming (lanes 7 and 8 in **A**). **(C)** qPCR data showing the expression of *Oct3/4* (left panel), *Sox2* (middle panel), and *Dppa3* (right panel) during the OSKM reprogramming time course in standard ESC culture medium

and either presence (+Vc, closed box) or absence (–Vc, closed triangle) of Vc. **(D)** Publicly available ChIP-Seq data showing the Oct4 and Sox2 binding (black peaks) at the *Dppa3* promoter region (red box). **(E)** ChIP followed by qPCR analysis for Oct3/4 (green) and Sox2 (maroon) binding at the *Dppa3* promoter region (Dppa3-P1 and Dppa3-P2) in ESCs. ChIP with IgG (blue) was used as a control.

Vitamin C (Vc) mediates global chromatin relaxation leading to early-activation of Dppa3

Recently, it was reported that the Vc-treatment results in global chromatin changes mostly in methylation levels of histone H3 lysine 36 (H3K36)[35]. The observed molecular parallels between Dppa3 and Vc in protecting against the imprinting loss at the *Dlk1-Dio3* cluster prompted us to investigate whether Vc treatment leads to the activation of Dppa3 during early stages of reprogramming. To answer this question we reprogrammed MEFs using OSKM in standard ESC medium supplemented either with or without Vc (+Vc and –Vc, respectively) and analyzed the cells for changes at various histone modification marks and expression levels of *Oct3/4*, *Sox2*, and *Dppa3*. After 5 days of reprogramming, MEFs treated with Vc (+Vc) showed enrichment for H3K4me3, a histone modification mark associated with active gene transcription, whereas MEFs treated without Vc (-Vc) showed basal levels of H3K4me3 (Fig. 6A, B). In contrast to H3K4me3, the levels of H3K27me3, a repressive histone mark, were greatly reduced in +Vc cells but not in –Vc cells (Fig. 6A, B). These changes were also evident at day 7 of reprogramming (Fig. 6A, B). Expression of endogenous *Oct3/4* and *Sox2* during reprogramming was activated at least two days earlier in +Vc cells compared to –Vc cells (Fig. 6C). Further, by day 12 of reprogramming, we observed much higher expression levels of endogenous *Oct3/4* and *Sox2* in +Vc cells compared to –Vc cells (Fig. 6C). Surprisingly, expression of *Dppa3* was already detected by day 2 of reprogramming in +Vc treated MEFs, but only by day 6 of reprogramming in control (-Vc) cells (Fig. 6C).

To correlate the early activation of *Dppa3* with chromatin changes induced by Vc, we hypothesized that exogenous reprogramming factor(s) might gain access to the *Dppa3* promoter and activate its expression early in Vc-treated cells. To verify this assumption, we checked the publicly available ChIP-Seq data of *Oct3/4* and *Sox2* binding sites/targets across the ESC genome and found *Oct3/4* and *Sox2* binding sites at the *Dppa3* promoter region (Fig. 6D). To validate these global ChIP-Seq data, we performed ChIP assays with Oct3/4 and Sox2 antibodies and subsequently qRT-PCR for the *Dppa3* promoter region using two PCR amplicons. The qRT-PCR data clearly indicate that both Oct3/4 and Sox2 bind to *Dppa3*, although with varying affinities (Fig. 6E). Collectively, these results suggest that Vc facilitate the transcription of *Dppa3* during early stages of reprogramming and this in turn leads to the binding of Dppa3 to the *Dlk1-Dio3* cluster and restricts the access of Dnmt3a to this locus.

Discussion

Although the generation of iPSCs from somatic cells has became a routine and reproducible method, the fully pluripotential capacity of each iPSC line generated is in question. In support of this view, two independent research groups have identified an imprinting defect at the *Dlk1-Dio3* cluster in the majority of mouse iPSC cell lines, which ultimately leads to the loss of germline transmission in chimeras[7,8]. These results strongly suggested that not all generated iPSC lines are faithfully reprogrammed. Recently, we have shown that GC marker genes such as *Blimp1*, *Fragilis*, and *Dppa3* are expressed earlier than the endogenous pluripotency-related genes and suggested that they might play a crucial role in establishment of pluripotency[26]. In the current study, we endeavored to find out the role of these GC marker genes in the establishment of authentic iPSC

lines with proper imprinting at the *Dlk1-Dio3* cluster. In our experiments, we made the interesting discovery that all iPSC lines with imprinting defects at the *Dlk1-Dio3* cluster did not express the Dppa3 protein, whereas normally imprinted iPSC lines showed Dppa3 expression. These results indicate that the expression of Dppa3 might be essential for the imprint establishment/maintenance at the *Dlk1-Dio3* cluster. In line with this hypothesis, all iPSC lines generated in the presence of exogenous *Dppa3* showed no *Dlk1-Dio3* imprinting defects. In detail, these iPSC lines displayed normal DNA methylation at the *Gtl2* IG-DMR,, expression of *Gtl2* and miRNAs encoded within the *Dlk1-Dio3* cluster as well as all pluripotent cell characteristics.

Dppa3 is highly expressed during embryonic development as well as in germ cells and pluripotent stem cells[36,37]. Gene ablation studies of *Dppa3* in mice revealed that it functions during early embryonic development to protect the maternal genome against the first wave of active DNA demethylation[25]. Moreover, *Dppa3* was also shown to be essential for protecting several but not all imprinted loci and transposons against DNA demethylation during post-fertilization events[25]. Interestingly, the paternally methylated *Dlk1-Dio3* region was not affected in *Dppa3*-null fertilized oocytes, whereas *H19* and *Rasgrf1*, the two other paternally methylated imprinted loci, showed partial loss of imprints[25]. Similarly, our current analysis of *Dppa3*-deficient ESCs also revealed no obvious defects in *Dlk1-Dio3* imprinting. In contrast to these observations, the expression of Dppa3 during somatic cell reprogramming was found to be positively correlated with the activation of *Gtl2*, thus resulting in normal imprint establishment or maintenance at the *Dlk1-Dio3* cluster in iPSCs. This discrepancy could be mainly due to artificial reprogramming of somatic cells versus normal early embryonic development. Another explanation would be that the imprinting of *Dlk1-Dio3* is stably maintained on the sperm-derived paternal allele in *Dppa3*-null embryos. In support of this view, the imprinting of the paternally methylated *H19* and *Rasgrf1* genes shows fluctuations in ESCs, whereas *Dlk1-Dio3* displays stable imprinting in these cells.

Dppa3 is regarded as a DNA binding protein and indeed *in-vitro* studies have shown that it binds non-specifically to DNA[25]. A recent report described that Dppa3 binds to nucleosomes containing a dimethylation mark at lysine 9 of histone 3 (H3K9me2)[38]. Our ChIP analysis showed that Dppa3 binds to a specific region in the *Dlk1-Dio3* IG-DMR, which also suggests that the interaction might involve modified histones. Unlike many imprinted loci, the IG-DMR of the *Dlk1-Dio3* cluster was reported to contain less or no H3K9me2[38]. Thus, it can be hypothesized that the association of Dppa3 with the IG-DMR is mediated by histone modifications other than H3K9me2 or some binding partner.

Recently, Stadtfeld et al., (2012)[9] showed that Vc can prevent the *Dlk1-Dio3* imprinting defect during reprogramming mainly by preserving the active chromatin structure at this locus. The resulting active chromatin was suggested to counteract the recruitment of Dnmt3a, a de novo DNA methyltransferase, to this locus thus maintaining the normal imprinting in Vc-treated cells[9]. The identified role of Dppa3 in preventing the *Dlk1-Dio3* imprinting defect suggests that Vc might function through the activation of *Dppa3* or that both Vc and *Dppa3* may function in a similar mechanistic pathway. This hypothesis is supported by the observation that *Dppa3* transcripts are upregulated Vc-treated cells. Moreover, the acceleration and efficiency of reprogramming seen after addition of Vc[9,35,39] can also be observed after addition of exogenous Dppa3, indicating that Dppa3 might be the genetic factor responsible for the Vc-mediated maintenance of proper imprinting at the *Dlk1-Dio3* cluster.

Collectively, our data show that Dppa3 is a genetic component essential for establishing

exclusively *Gtl2*^{on} iPSCs, which show full pluripotential characteristics. Based on these results, we propose that Dppa3 protects the *Dlk1-Dio3* imprinting during somatic cells reprogramming. The early expression of Dppa3 either by exogenous expression or by addition of Vc leads to the binding of Dppa3 to the *Dlk1-Dio3* IG-DMR and probably also to other known or unknown genomic regions. Upon binding, Dppa3 counteracts the recruitment of Dnmt3a to this region thereby maintaining the expression of imprinted genes from the *Dlk1-Dio3* cluster.

Acknowledgments

The authors would like to thank Prof. Masatsugu Ohora for providing pMXs-Blimp1 construct. This work was partly supported by the German Research Foundation (Deutsche Forschungsgemeinschaft) DFG SPP_1356 (EN 84/22-1 and ZE 442/4-2), DFG FOR 1041 (EN 84/23-1 and ZE 442/5-2).

References

1. Grabel, L. Prospects for pluripotent stem cell therapies: into the clinic and back to the bench. *J Cell Biochem* **113**, 381-7 (2012).
2. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-76 (2006).
3. Takahashi, K. et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861-72 (2007).
4. Chin, M.H. et al. Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell* **5**, 111-23 (2009).
5. Doi, A. et al. Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat Genet* **41**, 1350-3 (2009).
6. Hussein, S.M. et al. Copy number variation and selection during reprogramming to pluripotency. *Nature* **471**, 58-62 (2011).
7. Stadtfeld, M. et al. Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature* **465**, 175-81 (2010).
8. Liu, L. et al. Activation of the imprinted Dlk1-Dio3 region correlates with pluripotency levels of mouse stem cells. *J Biol Chem* **285**, 19483-90 (2010).
9. Stadtfeld, M. et al. Ascorbic acid prevents loss of Dlk1-Dio3 imprinting and facilitates generation of all-iPS cell mice from terminally differentiated B cells. *Nat Genet* **44**, 398-405, S1-2 (2012).
10. Spahn, L. & Barlow, D.P. An ICE pattern crystallizes. *Nat Genet* **35**, 11-2 (2003).
11. Ferguson-Smith, A.C. & Surani, M.A. Imprinting and the epigenetic asymmetry between parental genomes. *Science* **293**, 1086-9 (2001).
12. Tycko, B. & Efstratiadis, A. Genomic imprinting: piece of cake. *Nature* **417**, 913-4 (2002).
13. Brandeis, M. et al. The ontogeny of allele-specific methylation associated with imprinted genes in the mouse. *EMBO J* **12**, 3669-77 (1993).
14. Allegrucci, C., Thurston, A., Lucas, E. & Young, L. Epigenetics and the germline. *Reproduction* **129**, 137-49 (2005).
15. Lucifero, D., Mertineit, C., Clarke, H.J., Bestor, T.H. & Trasler, J.M. Methylation dynamics of imprinted genes in mouse germ cells. *Genomics* **79**, 530-8 (2002).
16. Arnaud, P. Genomic imprinting in germ cells: imprints are under control. *Reproduction* **140**, 411-23 (2010).
17. Bourc'his, D., Xu, G.L., Lin, C.S., Bollman, B. & Bestor, T.H. Dnmt3L and the establishment of

maternal genomic imprints. *Science* **294**, 2536-9 (2001).

18. Hata, K., Okano, M., Lei, H. & Li, E. Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development* **129**, 1983-93 (2002).
19. Kaneda, M. et al. Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature* **429**, 900-3 (2004).
20. Jelinic, P., Stehle, J.C. & Shaw, P. The testis-specific factor CTCFL cooperates with the protein methyltransferase PRMT7 in H19 imprinting control region methylation. *PLoS Biol* **4**, e355 (2006).
21. Kou, Y.C. et al. A recurrent intragenic genomic duplication, other novel mutations in NLRP7 and imprinting defects in recurrent biparental hydatidiform moles. *Mol Hum Reprod* **14**, 33-40 (2008).
22. Li, X. et al. A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. *Dev Cell* **15**, 547-57 (2008).
23. Meyer, E. et al. Germline mutation in NLRP2 (NALP2) in a familial imprinting disorder (Beckwith-Wiedemann Syndrome). *PLoS Genet* **5**, e1000423 (2009).
24. Murdoch, S. et al. Mutations in NALP7 cause recurrent hydatidiform moles and reproductive wastage in humans. *Nat Genet* **38**, 300-2 (2006).
25. Nakamura, T. et al. PGC7/Stella protects against DNA demethylation in early embryogenesis. *Nat Cell Biol* **9**, 64-71 (2007).
26. Xu, X. et al. Stage-specific germ-cell marker genes are expressed in all mouse pluripotent cell types and emerge early during induced pluripotency. *PLoS One* **6**, e22413 (2011).
27. Nishikawa, K. et al. Blimp1-mediated repression of negative regulators is required for osteoclast differentiation. *Proc Natl Acad Sci U S A* **107**, 3117-22 (2010).
28. Okita, K., Ichisaka, T. & Yamanaka, S. Generation of germline-competent induced pluripotent stem cells. *Nature* **448**, 313-7 (2007).
29. Khromov, T. et al. Global and gene-specific histone modification profiles of mouse multipotent adult germline stem cells. *Mol Hum Reprod* **17**, 166-74 (2011).
30. Zechner, U. et al. Comparative methylation profiles and telomerase biology of mouse multipotent adult germline stem cells and embryonic stem cells. *Mol Hum Reprod* **15**, 345-53 (2009).
31. Dressel, R. et al. Multipotent adult germ-line stem cells, like other pluripotent stem cells, can be killed by cytotoxic T lymphocytes despite low expression of major histocompatibility complex class I molecules. *Biol Direct* **4**, 31 (2009).
32. Dressel, R. et al. Pluripotent stem cells are highly susceptible targets for syngeneic, allogeneic, and xenogeneic natural killer cells. *Faseb J* **24**, 2164-77 (2010).
33. Nakagawa, M. et al. Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat Biotechnol* **26**, 101-6 (2008).
34. Smorag, L. et al. MicroRNA signature in various cell types of mouse spermatogenesis: Evidence for stage-specifically expressed miRNA-221, -203, and -34b-5p mediated spermatogenesis regulation. *Biol Cell* (2012).
35. Wang, T. et al. The histone demethylases Jhdm1a/1b enhance somatic cell reprogramming in a vitamin-C-dependent manner. *Cell Stem Cell* **9**, 575-87 (2011).
36. Payer, B. et al. Stella is a maternal effect gene required for normal early development in mice. *Curr Biol* **13**, 2110-7 (2003).
37. Sato, M. et al. Identification of PGC7, a new gene expressed specifically in preimplantation embryos and germ cells. *Mech Dev* **113**, 91-4 (2002).

38. Nakamura, T. et al. PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. *Nature* **486**, 415-9 (2012).

39. Esteban, M.A. et al. Vitamin C enhances the generation of mouse and human induced pluripotent stem cells. *Cell Stem Cell* **6**, 71-9 (2010).

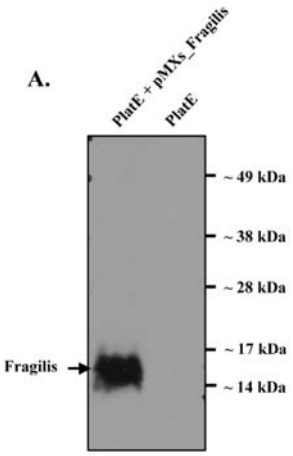


Figure S1. Verification of pMXs-Fragilis construct. Western blot showing the expression of Fragilis in pMXs-Fragilis construct-transfected PlatE cells, whereas PlatE cells alone (control) served as a control.

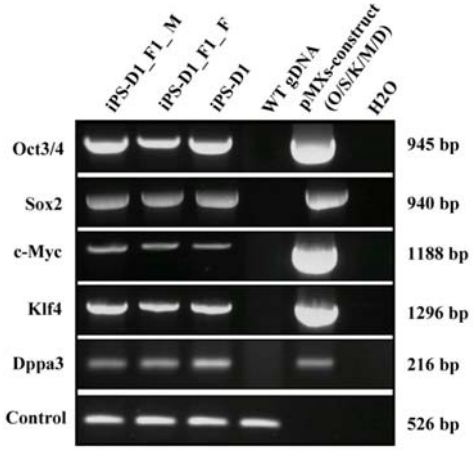


Figure S2. Confirmation of iPS-D germline transmission. Genotyping PCR showing the presence of OSKM+D constructs in two offspring (iPS-D1_F1_M and iPS-D1_F1_F) derived from the iPS-D1 chimera. Genomic DNA from iPS-D1 and wild-type ESCs (WT gDNA) served as positive and negative controls, respectively. Similarly, plasmid DNAs of OSKMD constructs (pMXs-construct) and water served as positive and negative controls, respectively. The sizes of the PCR fragments are indicated on the right side.

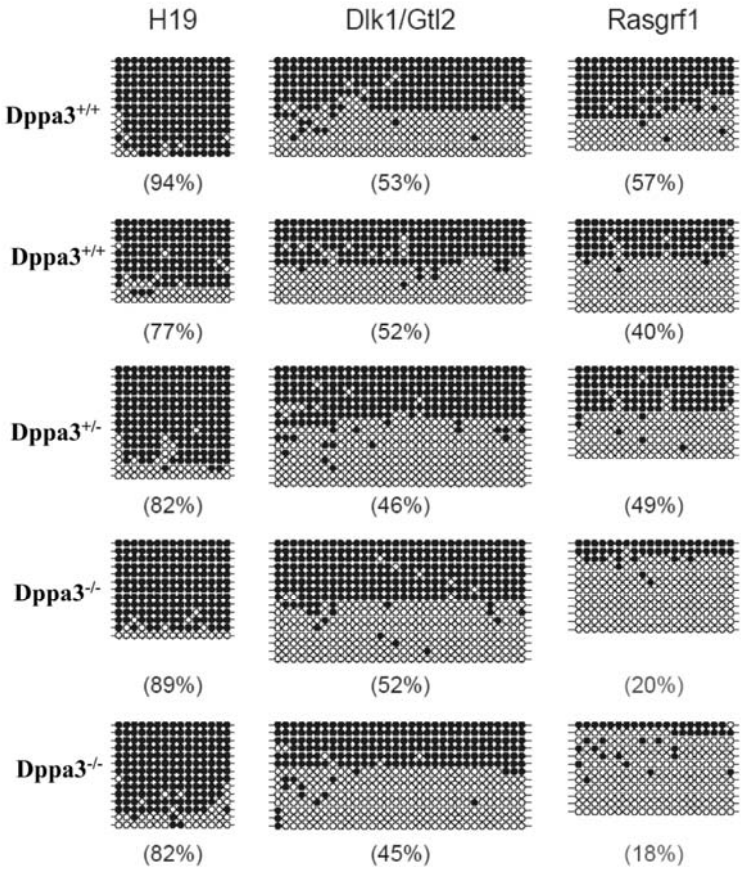


Figure S3. Analysis of the *Dlk1-Dio3* imprinting status in *Dppa3*-deficient ESCs. DNA methylation analysis for the paternally imprinted *H19*, *Dlk1-Dio3*, and *Rasgrf1* loci in wild-type (*Dppa3*$^{+/+}$), heterozygous (*Dppa3*$^{+/-}$), and homozygous (*Dppa3*$^{-/-}$) ESCs.

3.3. Mouse Dazl and its novel splice variant function in translational repression of target mRNAs in embryonic stem cells

Although pluripotent cells are known to express several PrM marker genes, their functional relevance to pluripotency is largely unknown. To unveil the function of PrM marker genes in pluripotent cells, in the third part of this thesis, we choose *Dazl*, a PrM marker gene, as a candidate and elucidated its function in ESCs. During the course of this study we identified a novel *Dazl* splice variant (*Dazl_Δ8*) in mouse ESCs and subsequently found it to be expressed in various pluripotent cell types as well as at different stages of testicular development. Protein sub-cellular localization studies revealed a diffused cytoplasmic and large granular localization pattern for Dazl_Δ8, which is similar to the pattern observed for Dazl full length (Dazl_FL) protein. Moreover, the studies using alteration of Dazl expression in ESCs indicated a role for Dazl in the negative translation regulation of both known and unknown target genes. The translation repressive role of Dazl in ESCs but not in germ line derived GC1 cells was further confirmed in luciferase reporter assays. Furthermore, we identified several specific as well as common targets for both Dazl isoforms in ESCs through RNA-binding immunoprecipitation followed by whole genome transcriptome analysis. Collectively, our results highlight that Dazl might function as a translational repressor at physiological level to maintain the balance between germ cell lineage commitment and pluripotency of ESCs.

Authors: Xingbo Xu, Xiaoying Tan, Qiong Lin, Bernhard Schmidt, Wolfgang Engel and D.V. Krishna Pantakani

Status: Submitted to BBA-Gene Regulatory Mechanisms journal.

Author contributions to the work:

1. Xingbo Xu: Conception and design of experiments, data collection, data analysis, data interpretation, helped in drafting the manuscript.
2. Xiaoying Tan: Data collection, data analysis.
3. Qiong Lin: Microarray data analysis.
4. Bernhard Schmidt: Design of experiments, data analysis.
4. Wolfgang Engel: Conception and design of experiments, financial support, helped in drafting the manuscript.
5. D.V. Krishna Pantakani: Conception and design of experiments, data interpretation, drafted the manuscript.

Mouse Dazl and its novel splice variant functions in translational repression of target mRNAs in embryonic stem cells

Xingbo Xu[1], Xiaoying Tan[1], Qiong Lin[2], Bernhard Schmidt[3], Wolfgang Engel[1] and D.V. Krishna Pantakani[1,‡]

[1]Institute of Human Genetics, University of Goettingen, Goettingen, Germany. [2]Institute for Biomedical Engineering, RWTH Aachen University Medical School, Aachen, Germany. [3]Department of Biochemistry II, University of Goettingen, Goettingen, Germany.

‡Correspondence: D. V. Krishna Pantakani, Tel: +49-551-399302; Fax: +49-551-399303; E-mail: krishna.if1@gmail.com

Abstract

Dazl (deleted in azoospermia-like) is an RNA binding protein that is important for germ cell differentiation in vertebrates. In the present study, we report the identification of a novel *Dazl* isoform (*Dazl_Δ8*) that results from alternative splicing of exon8 of mouse *Dazl*. We observed the expression of *Dazl_Δ8* in various pluripotent cell types, but not in somatic cells. Furthermore, the *Dazl_Δ8* splice variant was expressed along with the full-length isoform of *Dazl* (*Dazl_FL*) throughout male germ-cell development and in the ovary. Sub-cellular localization studies of Dazl_Δ8 revealed a diffused cytoplasmic and large granular pattern, which is similar to the localization patterns of Dazl_FL protein. In contrast to the well documented translation stimulation function in germ cells, overexpression and downregulation studies of Dazl isoforms (*Dazl_FL* and *Dazl_Δ8*) revealed a role for Dazl in the negative translational regulation of *Mvh*, a known target of Dazl, as well as *Oct3/4* and *Sox2* in embryonic stem cells (ESCs). In line with these observations, a luciferase reporter assay with the 3′UTR's of *Oct3/4* and *Mvh* confirmed the translational repressive role of Dazl isoforms in ESCs but not in germ cells derived cell line GC-1. Further, we identified several putative target mRNAs of Dazl_FL and Dazl_Δ8 in ESCs through RNA-binding immunoprecipitation followed by whole genome transcriptome analysis. Collectively, our results show a translation repression function of Dazl in pluripotent stem cells.

Keywords: ESCs, Dazl, splice variant, translation repression, RNA binding

Abbreviations: maGSCs-multipotent adult germline stem cells; EGCs-embryonic germ cells; iPSCs-induced pluripotent stem cells; ECCs- embryonic carcinoma cells; GC-1- a spermatogonia cell line; RIP- RNA-binding protein immunoprecipitation

Introduction

The DAZ (Deleted in Azoospermia) family of proteins, DAZ, DAZL (DAZ-Like), and BOULE play a crucial role in the gametogenesis of metazoans by regulating the translation of target messenger RNAs (mRNAs). These proteins are known to contain a RNA recognition motif (RRM) that mediates the interaction with target mRNAs, as well as characteristic single or multiple repeats of the DAZ motif and a proline rich (PR) region. BOULE is evolutionarily conserved and found in all metazoans; whereas, DAZL is only present in vertebrates [1, 2]. Interestingly, DAZ originated from a duplication transposition of the DAZL gene to the distal long arm of the Y chromosome in multiple copies and is found only in Old World monkeys and humans [3, 4]. In humans, the deletion of the DAZ cluster is believed to cause azoospermia and oligospermia in 5-10% of infertile men [3, 5]. Loss of function

studies of autosomal members of Daz lead to a failure to produce mature gametes in several model organisms [6-9].

The mouse Dazl codes for a germ cell-specific protein present in both the nucleus and cytoplasm of fetal gonocytes [10]. However, in spermatogonial stem cells (SSCs), Dazl is localized mainly to the nucleus and relocates into the cytoplasm during meiosis [10]. Concomitant with the specific expression in germ cells, Dazl has been shown to be important for germ cell development in diverse species, including humans [6, 7, 9, 11-14]. The deletion of mouse Dazl (mDazl) results in loss of germ cells in both male and female gonads, where male germ cells are lost partly due to increased apoptosis, reduced expression of germ cell markers, and aberrant chromatin structure [9, 11, 15]. In addition, it has been reported that Dazl deficiency leads to the reduction of post-migratory PGCs (primordial germ cells), and a failure to erase and re-establish genomic imprints, as well as defects in meiotic progression in mice[16].

Apart from germ cells, several reports have documented the cytoplasmic expression of Dazl in pluripotent embryonic stem cells (ESCs) [17-19]. Furthermore, Dazl depleted PGCc show aberrant expression of pluripotency-related genes and failed to establish pluripotent embryonic germ cells (EGCs), thus highlighting the role of Dazl in the establishment of pluripotency [16]. Consistent with the role of Dazl in germ cells development, Yu et al., (2009) have shown that the forced expression of Dazl leads to the differentiation of mouse ESCs into cells resembling sperm cells and oocytes, in vitro [20]. Strikingly, these in-vitro derived sperm-like cells can fertilize wild-type oocytes, which further underwent embryonic development until the 8-cells stage [20]. Collectively, these results reinforce the idea that Dazl might function as a master regulator for germ cell specification and differentiation.

The ability of Dazl to control germ cell development is attributed to its potential role in the translational regulation of known and unknown target mRNAs [21-26]. Molecular analysis of the human and mouse genomes over the last decade revealed alternative splicing as one of the most essential and common mechanisms of generating a large repertoire of mRNA and protein isoforms, which is occurring in >80% of genes [27, 28]. The protein products of these alternative isoforms have changed structures and functions, leading to a more complex proteome. Unfortunately, nothing is known about the Dazl splice variants and their possible functions in translation regulation. In the present study, we report the alternative splicing of exon8 of mDazl that leads to the generation of the novel splice variant, Dazl_Δ8. This novel transcript was detectable in all analyzed pluripotent cell types, adult mouse testis, as well as in ovary, albeit at low levels compared to Dazl_FL expression. Furthermore, we observed the expression of Dazl_Δ8 throughout mouse testicular development, specifically in germ cells, not in testicular somatic cells. Interestingly, overexpression and downregulation studies of Dazl isoforms (Dazl_FL and Dazl_Δ8), and luciferase reporter assays using the 3′UTRs of Oct3/4 and Mvh revealed a negative translational role for Dazl in ESCs. In addition, we performed RNA-binding protein immunoprecipitation (RIP) assays on ESCs transiently overexpressing E2-epitope tagged Dazl_Δ8 or Dazl_FL and could identify several specific as well as common mRNA targets of Dazl isoforms.

Materials and methods

Animals and cell culture

The mouse mutants *W/W^{V}* [29], *Tfm/y* [30], *olt/olt* [31] and *qk/qk* [32] were purchased from the Jackson Laboratory, whereas *leyI-I$^{-/-}$* was generated in-house [33]. All animal experiments were performed according to protocols approved by the Institutional Ethical Committee of University of Goettingen. Derivation and cultivation of mouse ESC,

maGSC, EGC, iPSC, ECC and GC-1 cell lines have been described previously [19, 34]. For transient overexpression experiments, ESCs were seeded in KO-DMEM supplemented with KO-serum replacement (Life technologies) at a density of 2×10^5 cells/well of a 6-well plate containing feeder layer. Five hours after plating, the cells were transfected with indicated constructs using Lipofectamine-2000 (Life technologies), according to manufacturer's instructions. After 3 h of transfection, the medium was changed back to standard ESCs culture medium and the cells were harvested after 48 h of culture. Dazl downregulation studies were performed as described [19]. Briefly, ESCs were seeded at a density of 2×10^5 cells/ well of a 6-well plate containing feeder. After 5 h of plating, the cells were transfected with either of two independent *Dazl* siRNA, Dazl_siRNA_exon7 (NM_010021.4_stealth_726, Life technologies) and Dazl_siRNA_exon11 (NM_010021.4_stealth_1056, Life technologies), or scrambled siRNA (Stealth_control, Life technologies) using Lipofectamine-2000 (Life technologies). After 3 h of transfection, the medium was changed to standard ESCs culture medium and allowed to grow for 24 h. The next day, transfection was repeated and cells were harvested after an additional 24 h of culture. The sequences of Stealth siRNA duplex oligoribonucleotides are listed in Table S2.

Construction of the Dazl overexpression and luciferase vectors

The mouse Dazl open reading frame was amplified by Takara LA Taq Polymerase (Clontech) using mESC cDNA as the template. The PCR products were cloned into pGEM-T Easy vector (Promega). After sequencing verification, fragments were sub-cloned into the phEF1α expression vector to generate phEF1α-Dazl_FL (lacking an epitope tag, Dazl_F: GGATCCGATGTCTGCCACAACTTCTGAG and Dazl_R: GCGGCCGCTTAGCAGAGATGATCAGATTTAAG). In addition, PCR and subsequent cloning was performed using primers containing C-terminal E2-epitope tag (Dazl-E2_F: GGATCCGATGTCTGCCACAACTTCTGAG and Dazl-E2_R: GCGGCCGCTTAGCGATCTCTAAAATCAGAAGAAGTACTTGAACCACCGCAGAGATGATCAGATTTAAG) to generate phEF1α-Dazl_FL-E2 and phEF1α-Dazl_Δ8-E2.

To generate Luciferase reporter constructs for *Oct3/4* and *Mvh* 3′ UTRs, PCR and subsequent cloning was performed using cDNA (1μg) from ESCs as a template to amplify *Oct3/4* and *Mvh* 3′ UTR sequences. Primers used for PCR amplification of 3′ UTRs are provided in the supplementary Table S3. The amplified PCR product was gel-purified and subcloned into pGEM-T Easy vector (Promega). The inserted sequences were first sequence verified and then excised with *Sac*I and *Hind*III restriction enzymes (Life technologies), gel-purified, and cloned into the pMIR-REPORT miRNA luciferase reporter vector (Ambion), referred to as pMIR-*Oct3/4*-3′UTR and pMIR-*Mvh*-3′UTR.

Protein extraction and Western blotting

Proteins were extracted from cells and tissues using lysis buffer (10mM Tris–HCl (pH 8.0), 1mM EDTA, 2.5% SDS, 100mM PMSF) containing protease inhibitor cocktail (Roche). Protein samples were resolved on 4-12% SDS-PAGE and transferred onto nitrocellulose membrane (Amersham Biosciences). Membranes were processed using standard Western blot protocols, and signals were detected using a chemiluminescent kit (Santa Cruz Biotechnology). Antibody sources are listed in supplementary Table S4.

RNA extraction, RT-PCR and qPCR

Total RNA was extracted from cells and mouse testis samples using Trizol Reagent (Life technologies) following the manufacturers' protocols.

Total RNA was digested with *DNaseI* (Sigma) and used for cDNA synthesis using the SuperScriptII system (Life technologies). For qPCR analysis, diluted cDNA (1/10) was used as a template in QuantiFast SYBR Green (Qiagen) reaction and run in ABI 7900HT Real-Time PCR System (Applied Biosystems). Primers used in RT-PCR and qPCR are listed in supplementary Tables S5 and S6, respectively.

Immunofluorescence staining

ESCs, NIH-3T3, and GC-1 cells grown on round cover slips were transfected with indicated constructs using Lipofectamine 2000 (Life technologies). After 24 h of transfection, the cells were processed for immunofluorescence staining, as described [35]. Microscopic images were acquired using Olympus BX60 fluorescent microscope equipped with UplanFI 20X/0.50 objective lens and Cell^F program. The acquired images were processed using the same Cell^F software.

Luciferase reporter assay

Luciferase reporter assay was performed as described previously [36]. Briefly, one day before transfection, ESCs or GC-1 were trypsinized and transferred to a fresh well of a 6 well plate at a density of 1×10^5 cells/well. Next day, the cells were transiently transfected in duplicates with indicated luciferase vector (pMIR-*Oct3/4*-3′UTR or pMIR-*Mvh*-3′UTR), pMIR-Report β-Gal (control plasmid for use in normalization of the results) (Life technologies) and with either phEF1α-Dazl_FL-E2 or phEF1α-Dazl_Δ8-E2 using Lipofectamine 2000 (Life technologies). The optimized DNA transfection amount per each vector was 2.5µg. After 48h of transfection, the cells were lysed and assayed for both luciferase and β-Gal activity using Luciferase Assay System (Promega) and Galacto-Light Plus (Applied Biosystems), respectively. Normalization was performed with luciferase/β-gal activity in cells transfected with negative control plasmid.

RIP, microarray analysis and data manipulation

RNA-binding protein immunoprecipitation (RIP) assays were performed using EZ-Magna RIP Kit (Millipore) according to manufacturer's instructions. Briefly, ESCs (~5 x 10^6 cells) were transfected with E2-tagged isoform specific constructs (*Dazl*_FL-E2 or *Dazl*_Δ8-E2). After 48h of transfection, the cells were washed two times with ice-cold PBS and transferred into a falcon tube by scrapping. After centrifugation, the pellet was re-suspended with an equal volume of RIP lysis buffer and mixed properly until the cells appeared homogeneous. After 5 min of incubation, the protein-RNA complex was obtained by centrifugation at 11,000 × g for 10 min at 4 °C. The protein-RNA complex (100µl) was used for each immunoprecipitation and pulled down by a beads-antibody complex consisting of 6µl of E2-antibody coupled to magnetic beads. After 3 h of incubation at 4°C, the protein was digested by using proteinase-K at 55 °C for 30 min. Then, RNA purification was carried out using 400µl phenol:chloroform:isoamyl alcohol (125:24:1 pH 4.3) (Sigma-Aldrich) and precipitated with Salt Solution I, II, and precipitate enhancer. The RNA pellet was washed with 80% ethanol and dissolved in 20µl of RNase-free water and quantified using the NanoDrop ND-1000 (Thermo Scientific). The precipitated RNA (RIPed-RNA) samples were used in cDNA microarray as described previously [37].

The microarray raw data was processed as previously described [38]. Briefly, the raw data was quantile normalized and each gene's intensity value was given a percentile rank based on the ratio of the gene's intensity compared to the intensities for the entire array (the gene with the greatest intensity value has the highest rank). This rank was transformed to a percent relative to the other genes on the array, and each gene given a median percentile rank across each set of RIP experiments. Top ranked genes (>0.98) were selected. As a result, 116 genes were identified as Dazl_Δ8 targets and 626 genes as Dazl_FL targets. Out of the 116 Dazl_Δ8 and 626 Dazl_FL targets, 85

genes were identified as common targets to both Dazl isoforms.

Mass spectrometry analysis

The ESC protein extracts were immunoprecipitated using anti-Dazl antibodies and the resulting protein complex was resolved on 4-12% SDS –PAGE and stained using coomassie blue staining solution. The indicated protein bands were excised out of the gel, the proteins reduced and carbamidomethylated, in-gel digested with trypsin and analyzed by LC-MALDI mass spectrometry as described [39]. Combined mass lists from MS/MS spectra were used for a database search in the NCBInr Protein database using Mascot 2.4 (Matrix Science).

BNPS-Skatole cleavage of protein extracts

BNPS-Skatole cleavage experiments were carried out as previously described [40]. Briefly, ESCs grown to a confluence in a well of six-well plate were harvested and the resulting cell pellet was resuspended in 75µl of PBS and subjected to five rounds of freeze-thaw cycles using dry ice/ethanol. The cell insoluble fraction was removed by centrifugation at 20,000 × g for 5 min at 4 °C. The soluble protein fraction was denatured with 75µl of 8M urea (dissolved in PBS). Then, 250µl of freshly prepared BNPS-Skatole (1.6 mg/ml) in glacial acetic acid was added to the lysates. The reactions were kept overnight at room temperature (RT) in dark. Proteins were precipitated by adding 1.2 ml of cold acetone and incubation at −20 °C for 2 h, followed by centrifugation at 20,000 × g for 10 min at 4 °C. The supernatant was aspirated and the pellet was completely dried at RT. Then the samples were dissolved in 35µl of 8M urea in PBS and inactivated at 95 °C for 10 min. Proteins samples were prepared by adding equal volume of 2 × non-reducing Laemmli buffer and subsequent incubation at 95 °C for 5 min and analyzed by SDS-PAGE followed by Western blot.

Statistical Analysis

All qPCR data for RNA expression analysis (two or more biological replicates with three technical replicates each) were calculated using the standard curve method and are expressed as mean±SD. For statistical significance calculations, 2way ANOVA or Student's T-test (GraphPad Prism 4.0) were used.

Results

Identification of new mouse *Dazl* isoform, *Dazl_Δ8*

The genomic architecture of mouse *Dazl* and the details of the encoded protein and its functional domains are depicted in Fig. 1A. The mouse *Dazl* gene spans ~14 kb and consists of 11 exons, which encodes a 298 aa polypeptide, hereafter referred as Dazl_FL (Fig. 1A). The Dazl protein contains an RNA recognition motif (RRM), a proline rich region (PR) in addition to the characteristic DAZ (deleted in azoospermia) domain (Fig. 1B). In Western blot analysis of protein extracts from testis and pluripotent ESCs, a second band of ~30 kDa (marked with asterisk) was detected along with the known ~32 kDa band of Dazl_FL (Fig. 1C). To analyze whether Dazl exists as an alternatively spliced isoform, we performed RT-PCR on cDNA derived from ESCs using open reading frame (ORF) specific primers and detected a second fragment (~850 bp) along with the expected ~900 bp band corresponding to the *Dazl_FL* (Fig. 1 D). We then cloned these two fragments into pGEM-T Easy vector and had them sequenced. The sequencing results revealed an alternatively spliced isoform of *Dazl* with a deletion of exon8 (51 bp), referred to as *Dazl_Δ8* (Fig. 1A, E). Analysis of the protein sequence revealed that Dazl_Δ8 isoform is specifically missing a 17 aa region in the PR domain (Fig. 1B). To determine whether this isoform is expressed in pluripotent cell types other than ESCs,

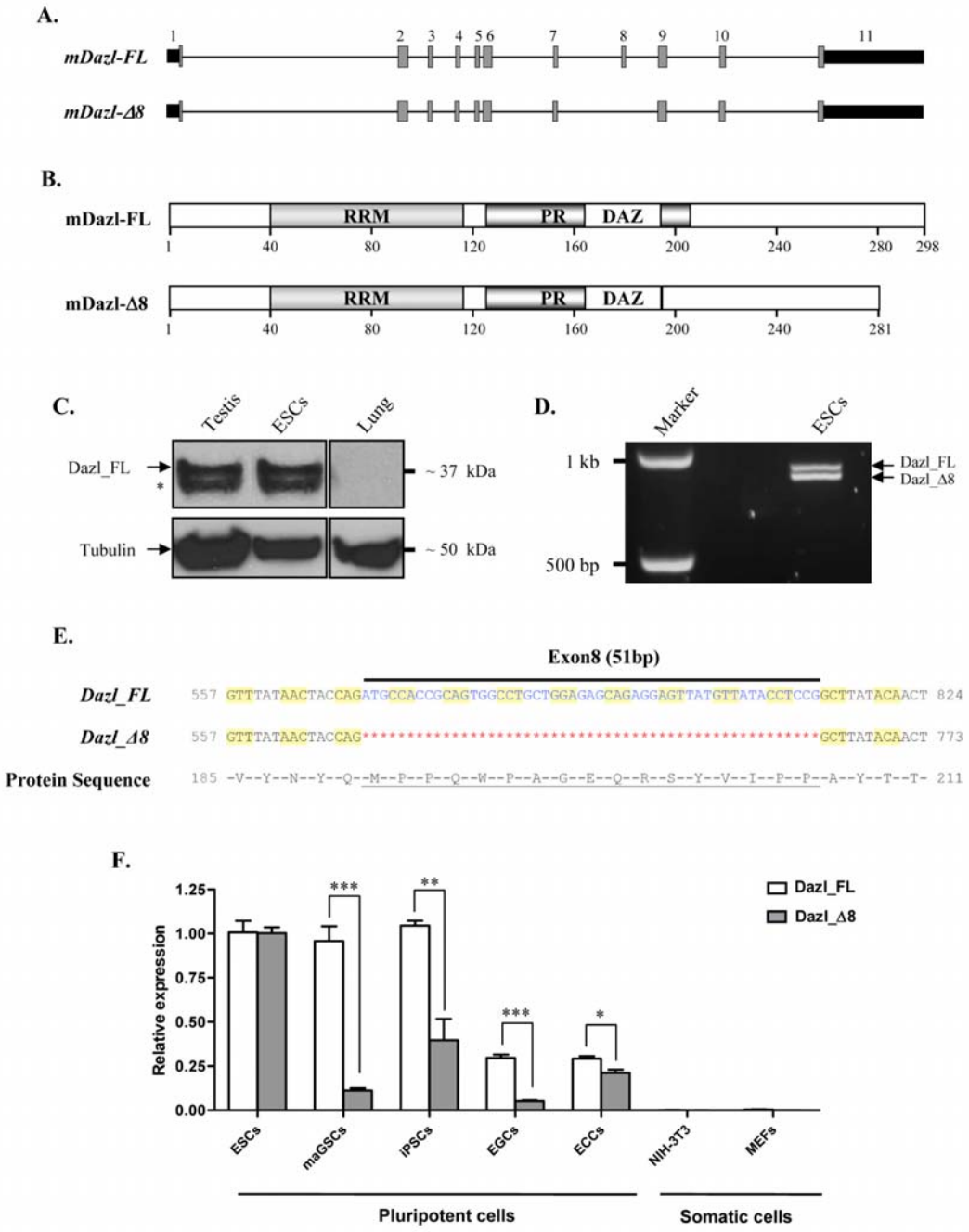


Figure 1. Identification and expression analysis of novel Dazl splice variant. (A) Schematic diagram showing the genomic architecture of mouse *Dazl* and its novel splice variant. Numerics on the top indicate the exon number, black rectangles – untranslated regions, grey – coding exons. **(B)** Protein structural domains of Dazl_FL and Dazl_Δ8. Amino acid positions along the protein are indicated below. RRM – RNA recognition motif, PR – Proline-rich region, and DAZL – DAZ like domain. **(C)** Western blot for Dazl expression in testis, embryonic stem cell (ESCs), and lung. The Dazl_FL isoform was indicated with arrow. Asterisk denotes the additional band and is most likely the Dazl_Δ8. The blot was re-probed with tubulin antibody to verify the

protein quantities. **(D)** RT-PCR analysis and subsequent agarose gel electrophoresis showing two PCR products, *Dazl_FL,* and *Dazl_Δ8*. **(E)** Pairwise alignment showing the nucleotide sequences of Dazl splice variants along with the Dazl_FL protein sequence. The nucleotide as well as amino acid numbering is indicated. **(F)** qRT-PCR showing the expression of *Dazl_Δ8* in comparison to *Dazl_FL* in various pluripotent cell types and somatic cells. The error bar represents the mean±SD and the values that are statistically significant (* $p<0.05$, ** $p<0.001$) are indicated.

we analyzed maGSCs, induced pluripotent stem cells (iPSCs), EGCs and embryonal carcinoma cells (ECCs) in comparison with somatic cells, such as NIH-3T3 and MEFs (Fig. 1F). qRT-PCR analysis using primers specific for each *Dazl* isoform revealed the presence of both *Dazl* isoforms in all pluripotent cell types, but not in somatic cells (Fig. 1F). The expression level of *Dazl_Δ8* was similar to that of *Dazl_FL* in ESCs, while though the *Dazl_Δ8* was detectable in all other pluripotent cell types, its levels were significantly lower than the *Dazl_FL* expression (Fig. 1F). To rule out that the differences in expression levels is due to differences in qPCR primer efficiencies, we verified both Dazl_FL and Dazl_Δ8-specific primers using standard curve method and found that the PCR amplification efficiencies are similar for both primer sets (Supplementary Fig. S1).

In order to characterize the second band seen in Western blot analysis, we performed immunoprecipitation on ESCs total protein extracts using anti-Dazl antibodies followed by LC-MALDI mass spectrometry analysis. The SDS-PAGE analysis and subsequent coomassie blue staining revealed the precipitation of two putative Dazl protein isoforms along with several other co-precipitated proteins (Supplementary Fig. S2A). When we analyzed these two putative Dazl protein isoforms using mass spectrometry we could identify only the N- terminal region of Dazl, indicating that the N-terminal region is identical for both Dazl protein isoforms (Supplementary Fig. S2B). To prove that the two Dazl protein isoforms differ in the internal sequence aa 191-207 (present in Dazl_FL and missing in Dazl_Δ8), we performed a chemical cleavage on total protein extracts of ESCs using 2-(2-nitrophenylsulfenyl)-3-methyl-3-bromoindolenine (BNPS-Skatole), which is known to cleave proteins at the C-terminal peptide bond following the tryptophan residue [40]. The amino acid sequence analysis of Dazl contains two tryptophan residues at position 143 and 195, the latter one inside the sequence deleted in Dazl_Δ8. After cleavage with BNPS-Skatole followed by Western blot analysis using an antibody against the C-terminus of Dazl, two low molecular weight bands differing by about 4 kDa were detected, indicating that Trp195 is missing in the Dazl fragment of the upper band, as predicted for the Dazl_Δ8 protein (Supplementary Fig. S2C and D).

Dazl_Δ8 isoform is expressed in various testicular developmental stages

To analyze the expression of *Dazl_Δ8* during testicular development, we performed qRT-PCR on cDNA samples prepared from various prenatal [15.5dpc (days post coitum), 18.5dpc], new born, postnatal [5dpp (days post partum), 10dpp and 15dpp] and adult testis samples of wild-type animals (Fig. 2A). The qRT-PCR analysis on adult testis samples revealed the presence of the *Dazl_Δ8* isoform, albeit at ~50% lower levels compared to *Dazl_FL* expression levels (Fig. 2A). The *Dazl* isoform expression analysis in prenatal and new born testis revealed very low expression of both isoforms, in accordance with the absence of meiosis at these stages of testicular development (Fig. 2A). Interestingly, the expression levels of the *Dazl_Δ8* isoform increased drastically along with *Dazl_FL* expression in postnatal testis developmental stages

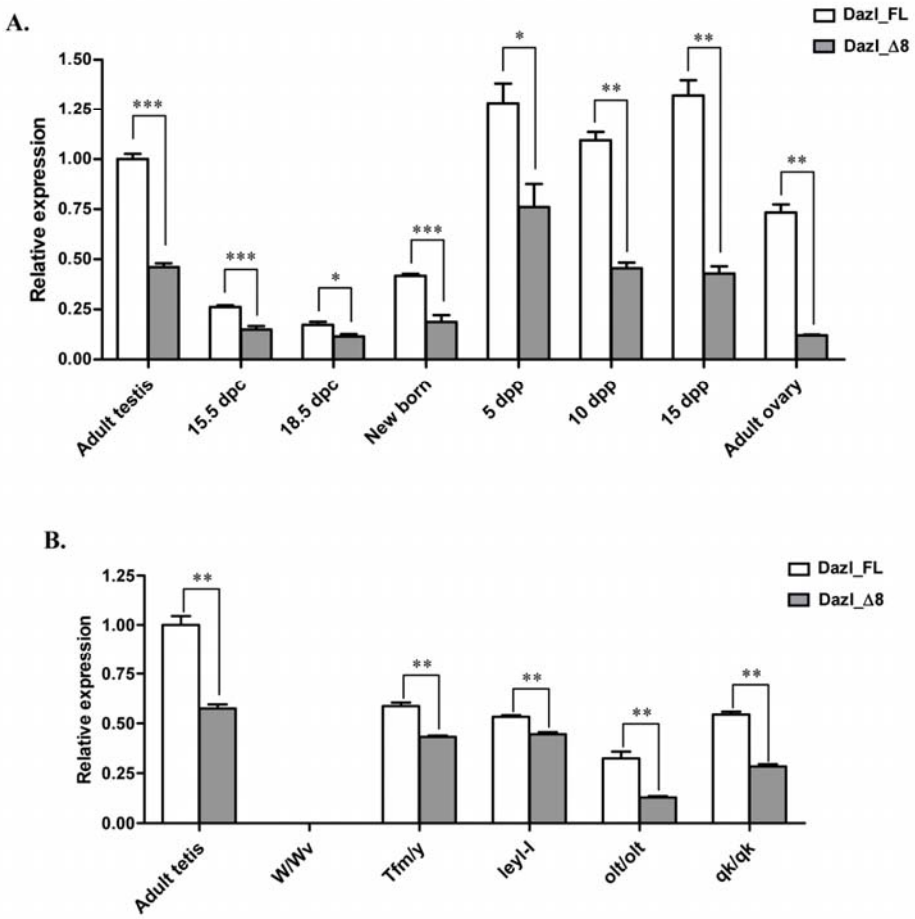


Figure 2. Expression analysis of Dazl isoforms in mouse testis and ovary. (A) qRT-PCR showing the expression of *Dazl_FL* and *Dazl_Δ8* in neonatal testis (15.5 dpc and 18.5 dpc), postnatal testis (newborn, 5 dpp, 10 dpp, 15 dpp), adult testis, and ovary. dpc- days post coitum, dpp – days post partum. **(B)** Expression analysis of *Dazl_Δ8* and *Dazl_FL* in spermatogenesis defective adult mouse mutants. White bars – *Dazl_FL* and grey bars – *Dazl_Δ8*, expression in both **(A)** and **(B).** The error bar represents the mean±SD and the values that are statistically significant (* $p<0.05$, ** $p<0.001$) are indicated.

(5-, 10-, 15-dpp), which coincided well with the initiation of meiosis at these stages (Fig. 2A). Again, *Dazl_Δ8* expression was significantly lower than *Dazl_FL* expression at all testicular stages (Fig. 2A). We also analyzed the expression of the *Dazl_Δ8* isoform in adult ovary and detected very low levels of expression compared to *Dazl_FL* (Fig. 2A). To characterize cell-type specific expression, we analyzed cDNA samples from various well characterized spermatogenesis adult mutants, in which spermatogenesis is disrupted at different stages, such as W/W^V (which lack all germ cells), *Tfm/y* & $leyI\text{-}I^{-/-}$ (in which spermatogenesis is arrested at the primary spermatocyte stage), and, *olt/olt* & *qk/qk* (in which spermatogenesis is blocked at the spermatid stage) mouse mutant models. qRT-PCR analysis of mutant testes samples indicated that both *Dazl* isoforms are undetectable in W/W^V mutant testis, but are expressed in all of the other analyzed testis mutants (Fig. 2B). These

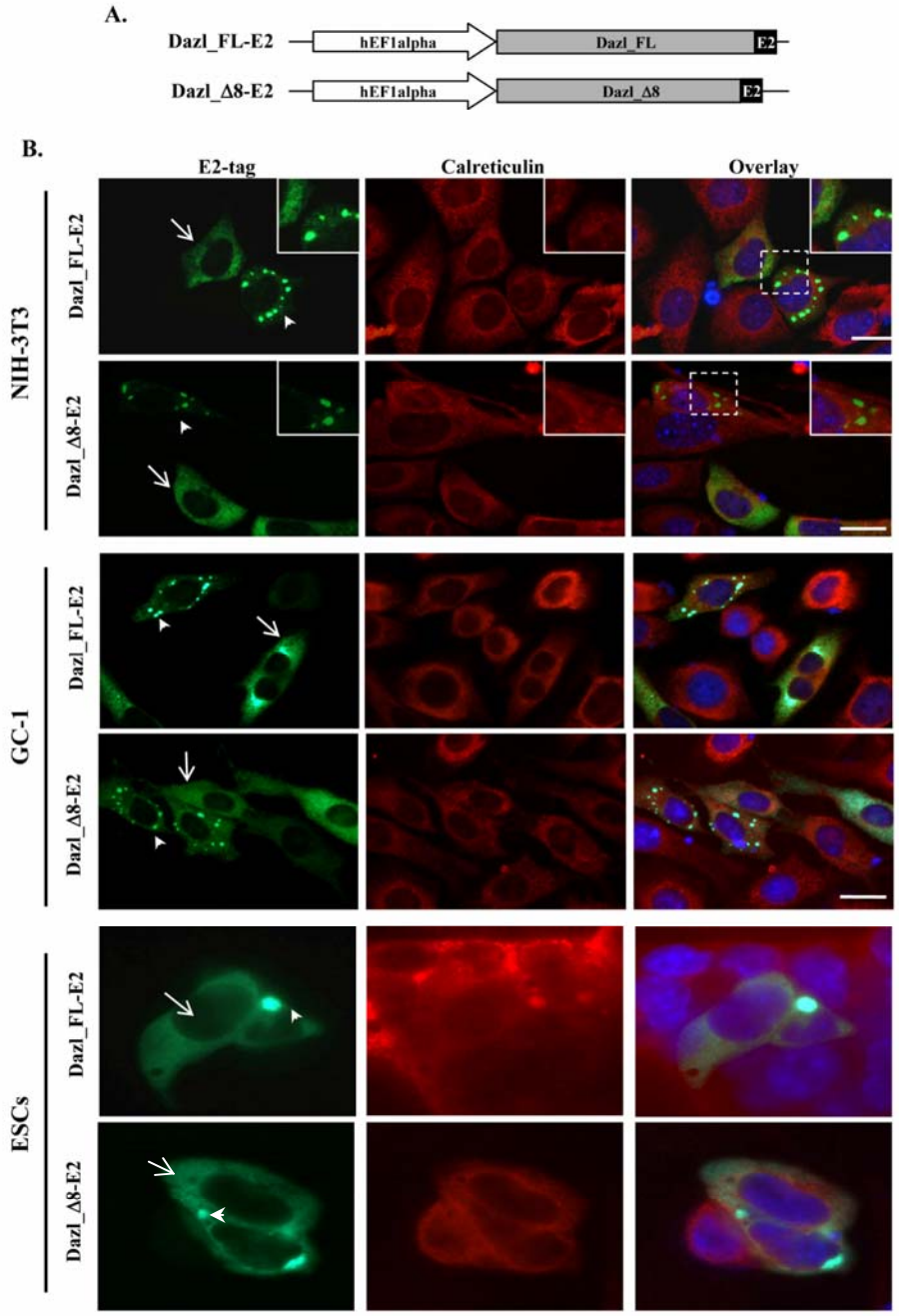


Figure 3. Sub-cellular localization of Dazl isoforms in NIH-3T3 cells. (A) Schematic representation of C-terminally tagged Dazl isoforms. **(B)** IFC images showing the localization of Dazl_FL-E2 and Dazl_Δ8-E2 protein isoforms (green, E2-tag) in NIH-3T3, GC-1 and ES cells. Cells were co-stained with calreticulin (red), an endoplasmic reticulum (ER) marker to visualize the cell morphology. Cells showing diffused cytoplasmic localization are indicated with an arrow, while cells showing large granular localization are highlighted with an arrow head. Insets in NIH-3T3 panels showing the enlarged image of a region indicated with checked line in overlay.

results suggest the specific expression of *Dazl* isoforms in either germ stem cells or in different spermatogenic cells or both, but not in somatic cells of testis, such as Sertoli and Leydig cells (Fig. 2B). As was observed in adult testis, the expression levels of *Dazl_Δ8* isoform were significantly lower than the level of *Dazl_FL* expression in each of the various mutants (Fig. 2B).

Dazl_Δ8 protein localized to the cytoplasm and granular structures, like Dazl_FL protein

To analyze the sub-cellular localization of Dazl_Δ8 protein relative to Dazl_FL, we generated mammalian expression constructs with a C-terminal E2-epitope tag (Dazl_FL-E2 and Dazl_Δ8-E2) (Fig. 3A). Expression of Dazl_FL-E2 in NIH-3T3 cells revealed the localization of Dazl_FL throughout the cytoplasm in ~75% of transfected cells, while ~20% of the cells showed large granular localization most probably to the stress granules (SGs) (Fig. 3B). These results are in consistence with previous studies describing the localization of Dazl as being both diffused through the cytoplasm and also in a granular pattern [41] Interestingly, the expression of Dazl_Δ8-E2 revealed a diffused cytoplasmic localization in ~84% of cells, while ~12% of cells showed granular expression pattern (Fig. 3B). We co-stained the cells with calreticulin, an endoplasmic reticulum (ER) marker, to visualize the cellular morphology (Fig. 3B). To rule out any cell type specific localization patterns, we performed Dazl isoforms localization studies in ESCs and germ cells derived GC-1 cell line. Similar to the results observed in NIH-3T3 cells, both ESCs and GC-1 cells revealed both cytoplasmic and granular expression pattern for both Dazl isoforms (Fig. 3B).

Dazl functions in the translational repression of target mRNAs in ESCs

Since Dazl is a known RNA binding protein and can regulate the translation of specific target mRNAs, we overexpressed Dazl_FL and Dazl_Δ8 (Dazl_Δ8-E2) in ESCs and measured the expression levels of target gene *Mvh* as well as pluripotency-related genes at the RNA and protein levels (Fig. 4A-C). Unexpectedly, overexpression of Dazl_FL resulted in reduced protein expression levels of target gene Mvh as well as Sox2 and Oct3/4 compared to control cells, while the protein levels of Klf4 was not affected (Fig. 4A). Interestingly, overexpression of Dazl_Δ8-E2 also leads to reduced protein levels of Sox2, Oct3/4, and Mvh, but not Klf4 (Fig. 4A), indicating that both Dazl isoforms might function in the translational repression of known as well as unknown target mRNA in ESCs, as opposed to the translation stimulation role they have in germ cells. To rule out that the reduction in protein levels of Mvh, Sox2 and Oct3/4 was not due to the transcriptional silencing of these genes, we checked the transcript levels of these genes and found no significant differences between the control cells and those overexpressing the Dazl isoforms (Supplementary Fig. S3A). We also performed downregulation studies for *Dazl* in ESCs using two independent Dazl siRNA oligos (Dazl_siRNA_exon7 and Dazl_siRNA_exon11), which target the Dazl transcript in exon7 and exon11, respectively. Since exon7 and exon11 are present in both *Dazl_FL* and *Dazl_Δ8* isoforms (Fig. 1A), we expected the simultaneous downregulation of both Dazl protein isoforms. The treatment of ESCs with Dazl_siRNA resulted in a dramatic decrease (up to 80%) in the protein levels of both Dazl isoforms, as expected (Fig. 4B). Cells treated with the control_siRNA (scrambled siRNA) did not have a reduction in Dazl protein levels (Fig. 4B). Consistently, the protein expression levels of Sox2, Oct3/4 and Mvh were dramatically increased in Dazl_siRNA treated cells, but not in control_siRNA treated cells (Fig. 4B). The protein expression of Klf4 did not differ between control_siRNA and Dazl_siRNA treated cells, as expected (Fig. 4B). We then analyzed the transcript levels of Sox2, Oct3/4 and Mvh in

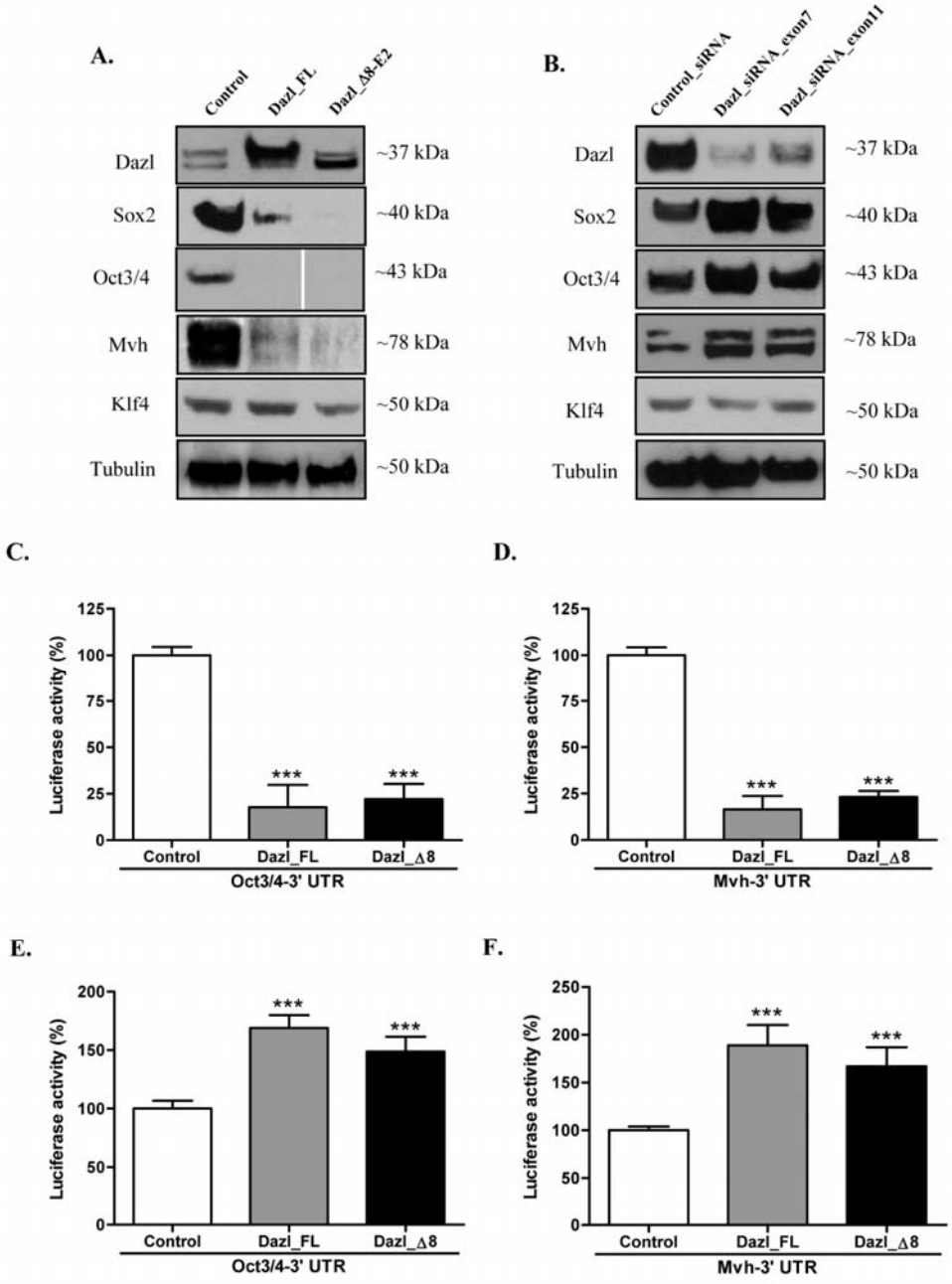


Figure 4. Dazl represses the translation of known target mRNAs in embryonic stem cells (ESCs). (A) Western blots showing the expression of Dazl, Sox2, Oct3/4, and Mvh upon either Dazl_FL-E2 or Dazl_Δ8-E2 overexpression in ESCs. Klf4 was served as a control. All the blots were re-probed with tubulin antibody to verify equal protein loading and one representative blot is shown. **(B)** Western blots showing the protein expression of Dazl, Sox2, Oct3/4, and Mvh upon Dazl down-regulation using two independent siRNA oligonucleotides (Dazl_siRNA_exon7 and Dazl_siRNA_exon11). Transfection of control_siRNA (scrambled siRNA) into ESCs was used as a control. Klf4 was served as a control. All the blots were re-probed with tubulin and one representative blot is shown. **(C)** Bar graph showing the luciferase reporter activity of *Oct3/4*-3′UTR

construct (pMIR-*Oct3/4*-3′UTR) in control (empty vector) and either Dazl_FL-E2 or Dazl_Δ8_E2 overexpressing ESCs. **(D)** Bar graph showing the luciferase reporter activity of the *Mvh*-3′UTR construct (pMIR-*Mvh*-3′UTR) in control (empty vector) and either Dazl_FL-E2 or Dazl_Δ8_E2 overexpressing ESCs. **(E)** Bar graph showing the luciferase reporter activity of the *Oct3/4*-3′UTR construct (pMIR-*Oct3/4*-3′UTR) in control (empty vector) and either Dazl_FL-E2 or Dazl_Δ8_E2 overexpressing GC-1 cells. **(F)** Bar graph showing the luciferase reporter activity of the *Mvh*-3′UTR construct (pMIR-*Mvh*-3′UTR) in control (empty vector) and either Dazl_FL-E2 or Dazl_Δ8_E2 overexpressing GC-1 cells. The error bar represents the mean±SD and the values that are statistically significant (***p<0.001) are indicated.

Dazl_siRNA treated cells compared to control_siRNA treated cells and found no significant differences (Supplementary Fig. S3B). To confirm the translational repressive role of Dazl, we performed luciferase reporter assays using constructs containing the 3′UTR of *Oct3/4* (pMIR-*Oct3/4*-3′UTR) and the 3′UTR of *Mvh* (pMIR-*Mvh*-3′UTR), which are cloned downstream of luciferase ORF into pMIR vector system. The co-expression of Dazl_FL-E2 and pMIR-*Oct3/4*-3′UTR in ESCs resulted in a significant decrease in the luciferase activity compared to empty vector and pMIR-*Oct3/4*-3′UTR transfected cells (Fig. 4C). Similarly, Dazl_Δ8-E2 expression also produced a significant reduction in luciferase reporter activity (Fig. 4C). Next, we co-expressed either Dazl_FL-E2 or Dazl_Δ8-E2 together with the pMIR-*Mvh*-3′UTR construct in ESCs and found a very significant reduction in luciferase reporter activity, as observed for the pMIR-*Oct3/4*-3′UTR reporter (Fig. 4D). In addition, we performed luciferase reporter assays using GC-1 cell line and found dramatic increase in the luciferase activity of both *Oct3/4* and *Mvh* luciferase reporter constructs (Fig. 4E, F), confirming the translational stimulation role of Dazl in germ cells but not in ESCs.

Identification of mRNAs associated with Dazl protein isoforms in ESCs

Our studies on Dazl protein isoforms expression analysis and the translational regulation mechanisms in ESCs revealed high similarities between both Dazl_FL and Dazl_Δ8 protein functions. To determine whether the Dazl_Δ8 protein has any preferential mRNA targets in ESCs or rather binds and regulates only a subset of Dazl_FL mRNA targets, we transiently overexpressed each Dazl isoform (Dazl_FL-E2 and Dazl_Δ8-E2) separately in ESCs and performed RIP using E2-tag specific antibodies. Simultaneously, we also performed RIP with E2-tag antibodies using untransfected ESCs extract as a negative control, but failed to detect any RNA as measured on a Nanodrop spectrophotometer. Two independent RIPs were performed for each of Dazl isoform and the resulting RNA was independently hybridized to microarrays. The mRNA candidates enriched in each experiment were identified by intensity value in the array and the subsequent mean percentile rank assignment for two independent replicates of each Dazl isoform RIPs. The top ranked candidates of each Dazl isoform were selected based on >0.98 median percentile rank (Fig. 5A, B). Comparison of top ranked (>0.98) candidates for Dazl_FL and Dazl_Δ8 revealed both isoform specific as well as common targets (Fig. 5C). Altogether, we identified 626 probes precipitated by the Dazl_FL isoform, of which 541 probes corresponding to 487 independent genes and 170 probes corresponding to unknown (NA) genes were specific for Dazl_FL (Fig. 5C). The remaining 85 probes corresponding to 66 individual genes and 16 probes corresponding to unknown were also precipitated by the Dazl_Δ8 isoform, which were named common target mRNAs (Fig. 5C). Some of the known Dazl target mRNAs

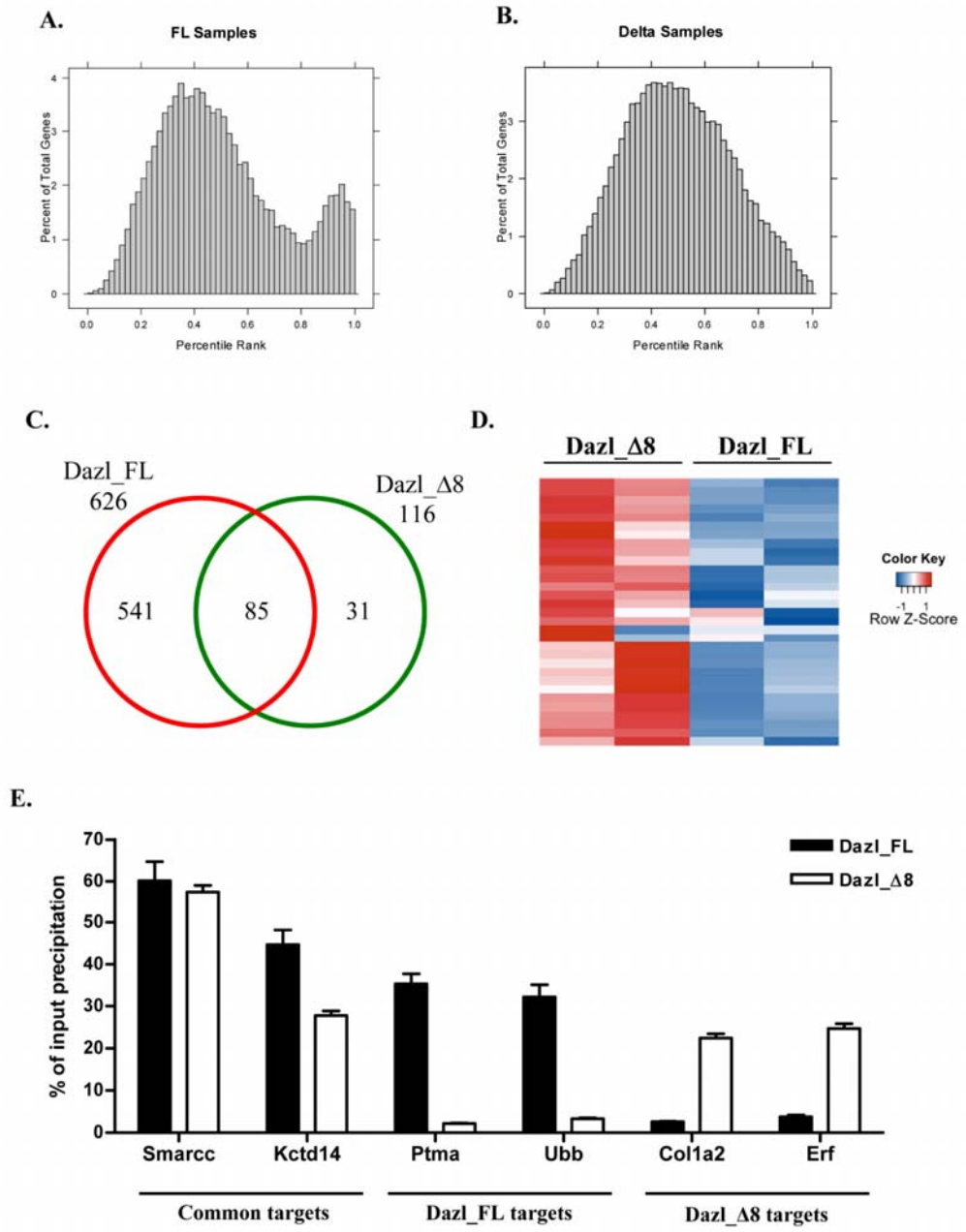


Figure 5. Identification and validation of Dazl-isoforms target mRNAs in embryonic stem cells (ESCs). Transcriptome analysis and subsequent percentile rank assignment image showing the distribution of the median percentile ranks for Dazl_FL **(A)** and Dazl_Δ8 **(B)**. **(C)** Venn diagram showing the number of common and isoform-specific target mRNAs between Dazl_FL and Dazl_Δ8. **(D)** Heat map showing Dazl_Δ8-specific target mRNAs (red-highly enriched) compared to Dazl_FL (blue-lower enriched). **(E)** Bar graph showing the qPCR validation of Dazl common (Smarcc and Kctd14) as well as Dazl_FL-specific (Ptma and Ubb) and Dazl_Δ8-specific (Col1a2 and Erf) targets mRNAs.

such as *Actg1*, *Calm2*, and *Slc2a3* (Reynolds et al., 2005) were readily detected in our array as Dazl_FL specific targets, confirming the specificity of our array (Table S1). For the Dazl_Δ8 isoform, we identified 116 probes, of which 31 probes corresponding to 24 genes and 7 corresponding to unknown genes were specific for Dazl_Δ8, while the remaining 85 probes/genes were common targets of both Dazl isoforms (Fig. 5C, D). A list of both common and isoform-specific target mRNAs is listed in supplementary Table S1. The transcriptome results were further confirmed by real-time PCR analysis for three selected candidates from each group (common, Dazl_FL-specific and Dazl_Δ8-specific) (Fig. 5E)

The Gene Ontology (GO) classification of common targets of Dazl isoforms revealed them as structural constituent of ribosome and translation regulators (Supplementary Table S7). Similarly, many Dazl_FL targets were classified as being involved in ribosome assembly and translation regulation. In contrast to these observations, the Dazl_Δ8 isoform targets were identified as being involved in homeostasis processes such as cellular chemical, and ion homeostasis (Supplementary Table S7).

Discussion

Several lines of evidence suggest that the germ cell specific translational regulator Dazl is expressed in all pluripotent cell types and may play a key role in germ cell specification and differentiation [19, 20]. From Western blot analysis with a Dazl specific antibody on lysates from pluripotent cell types, we identified an additional band apart from the expected *Dazl_FL*. Our subsequent RT-PCR studies identified a novel splice variant of mouse *Dazl* (*Dazl_Δ8*), which was expressed in the adult mouse testis, ovary, and also in all analyzed pluripotent cell types, albeit at low levels compared to *Dazl_FL*. Alternative splicing of *Dazl* exon8 was found to lead to the synthesis of protein lacking 17aa in the PR domain, while the rest of the protein architecture is intact and in-frame relative to *Dazl_FL*. We suggest that the additional band that we identified in Western blots is the Dazl_Δ8 protein isoform, as we could also observe a reduced expression of this protein during downregulation experiments. Further support for this suggestion comes from our BNPS-Skatole cleavage experiments, where we could see two Dazl antibody reactive bands with an expected molecular weight difference. Generally, the PR regions are known for their role in mediating protein-protein interactions [42] and, indeed, the PR region consisting of a DAZ domain in the Dazl protein has been shown to mediate the interaction with several known interaction partners [43]. Recently it was proposed that the interrelated functions of Poly (A) binding family proteins, whose direct binding to Dazl regulate the translation [22], could be attributable to splice variants with specific deletion of proline-rich region [44]. However, the functional significance of such splice variants is currently unknown. Based on the above observations, we hypothesize that the specific deletion in the PR region of Dazl might mediate or impede specific protein interactions, thereby altering the target mRNA specificity.

In ESCs, the transcript levels of *Dazl_Δ8* were almost equal to those of Dazl_FL, whereas other pluripotent cell types had significantly lower levels of *Dazl_Δ8.* For instance, maGSCs had significantly lower transcript levels of *Dazl_Δ8* compared to *Dazl_FL*. Interestingly, all pluripotent cell types with low levels of *Dazl_Δ8* originate from the germ-line stem cells, except for iPSCs; therefore, this might have some functional relevance for germ cell lineage. *Dazl* is expressed in oocytes and in the early stages of embryonic development [2, 17], as well as in the early and late germ-cell developmental stages [10, 16]. To delineate the expression of *Dazl_Δ8* in mouse testis development, as well as in the ovary, we performed qRT-PCR and detected the presence of the *Dazl_Δ8* transcript in all analyzed testis developmental stages and also the adult ovary. The

transcript levels of *Dazl_Δ8* were almost half of those of *Dazl_FL* in both adult mouse testis and at any given developmental stage. These observations, together with the *Dazl_Δ8* transcript levels in pluripotent cells of germ-line origin further supports the possibility that there is functional relevance for these differences. The most widely used mouse testis mutants with defects at various stages of spermatogenesis such as W/W^{V}, *Tfm/y*, *leyI-l(-/-)*, *olt/olt* and *qk/qk* offer the advantage of detecting the expression of a specific gene during various stages of spermatogenesis. Expression analysis of *Dazl_Δ8* in these mouse mutants revealed specific expression in all germ stem cells and their descending populations, but not in somatic cells of the testis. Collectively, these results argue for a conserved role for the Dazl_Δ8 isoform along with the Dazl_FL protein, in both male and female mouse gametogenesis.

Dazl is a cytoplasmic protein with diffuse cytoplasmic protein localization, as well as being localized to stress granules, under both native and ectopically expressed conditions [41, 45]. Consistent with these reports, the ectopic expression of C-terminally tagged Dazl_FL protein had both a diffuse cytoplasmic and a large granular expression pattern. Intriguingly, the novel Dazl isoform, Dazl_Δ8 also displayed both cytoplasmic and large granular localization, indicating that both Dazl isoforms might have overlapping functions.

The first evidence for the role of DAZ family proteins in translation regulation comes from genetic interaction studies in *Drosophila*, where *twine* was identified as a target of post-transcriptional regulation by *boule* [46]. Subsequent studies in mouse and Zebrafish identified the Dazl protein on actively translating polysomes and uncovered a role in translational stimulation of target mRNAs [47, 48]. Contrary to the known translational stimulatory role for Dazl in germ cells, we observed a translational repression role for Dazl in ESCs, and indeed, the Dazl known target *Mvh* was negatively regulated by both Dazl isoforms. Surprisingly, we also observed a negative translation regulation for Oct3/4 and Sox2, two well-characterized pluripotency marker genes. Our findings are in agreement with a previous study, which reported the translational repression of Tex19.1 mRNA by Dazl in Zebrafish embryos [49]. It is interesting to note that human hDAZ-associated protein 1 (DAZAP-1), a protein thought to function in translational repression, interacts with hDAZ, thus, competing for the interaction with translation promoting factor, PABP [50]. These observations led to the suggestion that the binding of DAZAP-1 to DAZ might modulate or repress translation [51]. Moreover, hDAZL is known to interact with Pumilio-2, a well-known translational repressor, and suggested to function as translational regulator in ESCs and germ cells [18]. In light of these observations it is possible that the competition between Dazl interacting translation activator and repressor proteins in ESCs might render both translational activation and repression functions of Dazl, respectively. Alternatively, Dazl mediated mRNA stability might be important for changes in the target mRNA translation in a cell-type dependent manner. Recently, it was shown that Dazl plays a crucial role in the formation of SGs during heat induced stress conditions in male germ cells [45]. These observations led to a suggestion that the formation and the function of SGs during stress conditions are to repress translation initiation and to activate protective mechanisms in male germ cells [41, 45]. In this context, the association of Dazl with SGs is interesting and we assume that Dazl might be important for the observed translational repression in SGs. Interestingly, the persistent expression of pluripotency marker genes such as Oct3/4 and Nanog were reported in *Dazl*-deficient germ cells, suggesting that Dazl might regulate the stem cell characteristics and differentiation [52]. Taken together, it is interesting to speculate that the physiological role of Dazl is to suppress the translation of germ cell specific transcripts in pluripotent cells.

Research over the last decade has led to the identification of several mRNA targets and conserved sequence elements in the 3′ UTRs that are bound by Dazl protein [24, 25, 46, 47, 49, 53, 54]. The discovery of the Dazl_Δ8 isoform prompted us to investigate the specific mRNA targets of this isoform, as well as to identify Dazl_FL target mRNAs in ESCs. Our immunoprecipitation-coupled transcriptome analysis revealed several target mRNAs that are bound by the Dazl isoforms, either selectively or in common. Interestingly, several of these target mRNAs (*Actg1*, *Calm2*, and *Slc2a3*) have already been described as putative targets of Dazl in germ cells [25], confirming the reliability of the method employed in the current study. Moreover, we identified specific target mRNAs for both Dazl isoforms in ESCs, suggesting functional divergence of these two isoforms. However, further studies are necessary to validate and dissect the interaction of Dazl isoforms with their target mRNAs and the functional significance of Dazl mediated translational regulation in ESCs. Unfortunately, our current RIP based transcriptome analysis failed to detect *Mvh*, a known target of Dazl, for which we observed the translational repression by Dazl. Hence, it is possible that our present study could not detect all Dazl target mRNAs, this might be due to the sensitivity of RIP assay and the strength between mRNA and protein interaction. Alternatively, the observed translational repression by Dazl, at least for *Oct3/4* and *Sox2*, might be an indirect effect, as some of Dazl interaction proteins can bind RNA and function as translational repressors. The GO analysis clearly indicated that both common and Dazl_FL targets as translation regulators, in line with the known function of Dazl in translation regulation. The identification of Dazl_Δ8 putative targets as cellular homeostasis regulators is quite unexpected and requires further validation of these targets to delineate Dazl_Δ8 function in these processes.

Conclusions

In summary, our current study describes the identification of a novel splice variant of *Dazl* (*Dazl_Δ8*) and its expression in pluripotent cells, as well as in both testis and ovary. Further, we uncovered a translational repressive role for Dazl in ESCs, as opposed to the well-known translation stimulation function in germ cells. Finally, we have identified several putative target mRNAs that bound either specifically or by both Dazl isoforms in ESCs. Further studies aimed at understanding the role of Dazl_Δ8 and their target mRNAs in pluripotent cells and in gametogenesis might uncover the functional significance of this novel protein isoform.

Acknowledgments

The authors would like to thank Dr. Gabriela Salinas-Riester for her great support with microarray experiments, Mr. Olaf Bernhard for his technical assistance and Dr. Jessica Nolte for valuable discussions. This work was partly supported by the German Research Foundation (Deutsche Forschungsgemeinschaft) DFG SPP1356 (EN 84/22-1) and DFG FOR 1041 (EN 84/23-1).

References

[1] E.S. Haag, Rolling back to BOULE, Proc Natl Acad Sci U S A, 98 (2001) 6983-6985.

[2] J.X. Peng, J.L. Xie, L. Zhou, Y.H. Hong, J.F. Gui, Evolutionary conservation of Dazl genomic organization and its continuous and dynamic distribution throughout germline development in gynogenetic gibel carp, J Exp Zool B Mol Dev Evol, 312 (2009) 855-871.

[3] R. Reijo, T.Y. Lee, P. Salo, R. Alagappan, L.G. Brown, M. Rosenberg, S. Rozen, T. Jaffe, D. Straus, O. Hovatta, et al., Diverse spermatogenic defects in humans caused by Y chromosome deletions encompassing a novel RNA-binding protein gene, Nat Genet, 10 (1995) 383-393.

[4] R. Saxena, L.G. Brown, T. Hawkins, R.K. Alagappan, H. Skaletsky, M.P. Reeve, R. Reijo, S. Rozen, M.B. Dinulos, C.M. Disteche, D.C. Page, The

DAZ gene cluster on the human Y chromosome arose from an autosomal gene that was transposed, repeatedly amplified and pruned, Nat Genet, 14 (1996) 292-299.

[5] N. Reynolds, H.J. Cooke, Role of the DAZ genes in male fertility, Reprod Biomed Online, 10 (2005) 72-80.

[6] C.G. Eberhart, J.Z. Maines, S.A. Wasserman, Meiotic cell cycle requirement for a fly homologue of human Deleted in Azoospermia, Nature, 381 (1996) 783-785.

[7] D.W. Houston, M.L. King, A critical role for Xdazl, a germ plasm-localized RNA, in the differentiation of primordial germ cells in Xenopus, Development, 127 (2000) 447-456.

[8] T. Karashima, A. Sugimoto, M. Yamamoto, Caenorhabditis elegans homologue of the human azoospermia factor DAZ is required for oogenesis but not for spermatogenesis, Development, 127 (2000) 1069-1079.

[9] M. Ruggiu, R. Speed, M. Taggart, S.J. McKay, F. Kilanowski, P. Saunders, J. Dorin, H.J. Cooke, The mouse Dazla gene encodes a cytoplasmic protein essential for gametogenesis, Nature, 389 (1997) 73-77.

[10] R.A. Reijo, D.M. Dorfman, R. Slee, A.A. Renshaw, K.R. Loughlin, H. Cooke, D.C. Page, DAZ family proteins exist throughout male germ cell development and transit from nucleus to cytoplasm at meiosis in humans and mice, Biol Reprod, 63 (2000) 1490-1496.

[11] Y. Lin, D.C. Page, Dazl deficiency leads to embryonic arrest of germ cell development in XY C57BL/6 mice, Dev Biol, 288 (2005) 309-316.

[12] P.T. Saunders, J.M. Turner, M. Ruggiu, M. Taggart, P.S. Burgoyne, D. Elliott, H.J. Cooke, Absence of mDazl produces a final block on germ cell development at meiosis, Reproduction, 126 (2003) 589-597.

[13] J.Y. Tung, M.P. Rosen, L.M. Nelson, P.J. Turek, J.S. Witte, D.W. Cramer, M.I. Cedars, R.A. Pera, Variants in Deleted in AZoospermia-Like (DAZL) are correlated with reproductive parameters in men and women, Hum Genet, 118 (2006) 730-740.

[14] J.Y. Tung, M.P. Rosen, L.M. Nelson, P.J. Turek, J.S. Witte, D.W. Cramer, M.I. Cedars, R.A. Reijo-Pera, Novel missense mutations of the Deleted-in-AZoospermia-Like (DAZL) gene in infertile women and men, Reprod Biol Endocrinol, 4 (2006) 40.

[15] B.H. Schrans-Stassen, P.T. Saunders, H.J. Cooke, D.G. de Rooij, Nature of the spermatogenic arrest in Dazl -/- mice, Biol Reprod, 65 (2001) 771-776.

[16] K.M. Haston, J.Y. Tung, R.A. Reijo Pera, Dazl functions in maintenance of pluripotency and genetic and epigenetic programs of differentiation in mouse primordial germ cells in vivo and in vitro, PLoS One, 4 (2009) e5654.

[17] G. Cauffman, H. Van de Velde, I. Liebaers, A. Van Steirteghem, DAZL expression in human oocytes, preimplantation embryos and embryonic stem cells, Mol Hum Reprod, 11 (2005) 405-411.

[18] F.L. Moore, J. Jaruzelska, D.M. Dorfman, R.A. Reijo-Pera, Identification of a novel gene, DZIP (DAZ-interacting protein), that encodes a protein that interacts with DAZ (deleted in azoospermia) and is expressed in embryonic stem cells and germ cells, Genomics, 83 (2004) 834-843.

[19] X. Xu, D.V. Pantakani, S. Luhrig, X. Tan, T. Khromov, J. Nolte, R. Dressel, U. Zechner, W. Engel, Stage-specific germ-cell marker genes are expressed in all mouse pluripotent cell types and emerge early during induced pluripotency, PLoS One, 6 (2011) e22413.

[20] Z. Yu, P. Ji, J. Cao, S. Zhu, Y. Li, L. Zheng, X. Chen, L. Feng, Dazl promotes germ cell differentiation from embryonic stem cells, J Mol Cell Biol, 1 (2009) 93-103.

[21] J. Chen, C. Melton, N. Suh, J.S. Oh, K. Horner, F. Xie, C. Sette, R. Blelloch, M. Conti, Genome-wide analysis of translation reveals a critical role for deleted in azoospermia-like (Dazl) at the

oocyte-to-zygote transition, Genes Dev, 25 (2011) 755-766.

[22] B. Collier, B. Gorgoni, C. Loveridge, H.J. Cooke, N.K. Gray, The DAZL family proteins are PABP-binding proteins that regulate translation in germ cells, Embo J, 24 (2005) 2656-2666.

[23] X. Jiao, P. Trifillis, M. Kiledjian, Identification of target messenger RNA substrates for the murine deleted in azoospermia-like RNA-binding protein, Biol Reprod, 66 (2002) 475-485.

[24] N. Reynolds, B. Collier, V. Bingham, N.K. Gray, H.J. Cooke, Translation of the synaptonemal complex component Sycp3 is enhanced in vivo by the germ cell specific regulator Dazl, Rna, 13 (2007) 974-981.

[25] N. Reynolds, B. Collier, K. Maratou, V. Bingham, R.M. Speed, M. Taggart, C.A. Semple, N.K. Gray, H.J. Cooke, Dazl binds in vivo to specific transcripts and can regulate the pre-meiotic translation of Mvh in germ cells, Hum Mol Genet, 14 (2005) 3899-3909.

[26] M. Zeng, W. Deng, X. Wang, W. Qiu, Y. Liu, H. Sun, D. Tao, S. Zhang, Y. Ma, DAZL binds to the transcripts of several Tssk genes in germ cells, BMB Rep, 41 (2008) 300-304.

[27] A.J. Matlin, F. Clark, C.W. Smith, Understanding alternative splicing: towards a cellular code, Nat Rev Mol Cell Biol, 6 (2005) 386-398.

[28] S. Stamm, S. Ben-Ari, I. Rafalska, Y. Tang, Z. Zhang, D. Toiber, T.A. Thanaraj, H. Soreq, Function of alternative splicing, Gene, 344 (2005) 1-20.

[29] J.L. Coulombre, E.S. Russel, Analysis of the pleiotropism at the W-locus in the mouse; the effect of W and W substitution upon postnatal development of germ cells, Journal of Experimental Zoology, 126 (1954) 277-296.

[30] M.F. Lyon, S.G. Hawkes, X-linked gene for testicular feminization in the mouse, Nature, 227 (1970) 1217-1219.

[31] R. Moutier, New mutation causing sterility restricted to the male in rats and mice. In The Laboratory Animal in the Study of Reproduction, The Laboratory Animal in the Study of Reproduction, (1976) 115-117.

[32] W.I. Bennett, A.M. Gall, J.L. Southard, R.L. Sidman, Abnormal spermiogenesis in quaking, a myelin-deficient mutant mouse, Biol Reprod, 5 (1971) 30-58.

[33] S. Zimmermann, G. Steding, J.M. Emmen, A.O. Brinkmann, K. Nayernia, A.F. Holstein, W. Engel, I.M. Adham, Targeted disruption of the Insl3 gene causes bilateral cryptorchidism, Mol Endocrinol, 13 (1999) 681-691.

[34] J.H. Lee, D. Schutte, G. Wulf, L. Fuzesi, H.J. Radzun, S. Schweyer, W. Engel, K. Nayernia, Stem-cell protein Piwil2 is widely expressed in tumors and inhibits apoptosis through activation of Stat3/Bcl-XL pathway, Hum Mol Genet, 15 (2006) 201-211.

[35] T. Khromov, D.V. Pantakani, J. Nolte, M. Wolf, R. Dressel, W. Engel, U. Zechner, Global and gene-specific histone modification profiles of mouse multipotent adult germline stem cells, Mol Hum Reprod, 17 (2011) 166-174.

[36] L. Smorag, Y. Zheng, J. Nolte, U. Zechner, W. Engel, D.V. Pantakani, MicroRNA signature in various cell types of mouse spermatogenesis: Evidence for stage-specifically expressed miRNA-221, -203, and -34b-5p mediated spermatogenesis regulation, Biol Cell, (2012).

[37] S. Meyer, J. Nolte, L. Opitz, G. Salinas-Riester, W. Engel, Pluripotent embryonic stem cells and multipotent adult germline stem cells reveal similar transcriptomes including pluripotency-related genes, Mol Hum Reprod, 16 (2010) 846-855.

[38] W.H. Townley-Tilson, S.A. Pendergrass, W.F. Marzluff, M.L. Whitfield, Genome-wide analysis of mRNAs bound to the histone stem-loop binding protein, Rna, 12 (2006) 1853-1867.

[39] C. Schulz, O. Lytovchenko, J. Melin, A. Chacinska, B. Guiard, P. Neumann, R. Ficner, O. Jahn, B. Schmidt, P. Rehling, Tim50's presequence receptor domain is essential for signal driven

transport across the TIM23 complex, J Cell Biol, 195 643-656.

[40] D.L. Crimmins, S.M. Mische, N.D. Denslow, Chemical cleavage of proteins in solution, Curr Protoc Protein Sci, Chapter 11 (2005) Unit 11 14.

[41] K.H. Lee, S. Lee, B. Kim, S. Chang, S.W. Kim, J.S. Paick, K. Rhee, Dazl can bind to dynein motor complex and may play a role in transport of specific mRNAs, Embo J, 25 (2006) 4263-4270.

[42] M.P. Williamson, The structure and function of proline-rich regions in proteins, Biochem J, 297 (Pt 2) (1994) 249-260.

[43] S. Tsui, T. Dai, S. Roettger, W. Schempp, E.C. Salido, P.H. Yen, Identification of two novel proteins that interact with germ-cell-specific RNA-binding proteins DAZ and DAZL1, Genomics, 65 (2000) 266-273.

[44] B. Gorgoni, N.K. Gray, The roles of cytoplasmic poly(A)-binding proteins in regulating gene expression: a developmental perspective, Brief Funct Genomic Proteomic, 3 (2004) 125-141.

[45] B. Kim, H.J. Cooke, K. Rhee, DAZL is essential for stress granule formation implicated in germ cell survival upon heat stress, Development, 139 (2012) 568-578.

[46] J.Z. Maines, S.A. Wasserman, Post-transcriptional regulation of the meiotic Cdc25 protein Twine by the Dazl orthologue Boule, Nat Cell Biol, 1 (1999) 171-174.

[47] S. Maegawa, M. Yamashita, K. Yasuda, K. Inoue, Zebrafish DAZ-like protein controls translation via the sequence 'GUUC', Genes Cells, 7 (2002) 971-984.

[48] S. Tsui, T. Dai, S.T. Warren, E.C. Salido, P.H. Yen, Association of the mouse infertility factor DAZL1 with actively translating polyribosomes, Biol Reprod, 62 (2000) 1655-1660.

[49] M. Zeng, Y. Lu, X. Liao, D. Li, H. Sun, S. Liang, S. Zhang, Y. Ma, Z. Yang, DAZL binds to 3'UTR of Tex19.1 mRNAs and regulates Tex19.1 expression, Mol Biol Rep, 36 (2009) 2399-2403.

[50] S. Morton, H.T. Yang, N. Moleleki, D.G. Campbell, P. Cohen, S. Rousseau, Phosphorylation of the ARE-binding protein DAZAP1 by ERK2 induces its dissociation from DAZ, Biochem J, 399 (2006) 265-273.

[51] M. Brook, J.W. Smith, N.K. Gray, The DAZL and PABP families: RNA-binding proteins with interrelated roles in translational control in oocytes, Reproduction, 137 (2009) 595-617.

[52] M.E. Gill, Y.C. Hu, Y. Lin, D.C. Page, Licensing of gametogenesis, dependent on RNA binding protein DAZL, as a gateway to sexual differentiation of fetal germ cells, Proc Natl Acad Sci U S A, 108 (2011) 7443-7448.

[53] M. Fox, J. Urano, R.A. Reijo Pera, Identification and characterization of RNA sequences to which human PUMILIO-2 (PUM2) and deleted in Azoospermia-like (DAZL) bind, Genomics, 85 (2005) 92-105.

[54] J.P. Venables, M. Ruggiu, H.J. Cooke, The RNA-binding specificity of the mouse Dazl protein, Nucleic Acids Res, 29 (2001) 2479-2483.

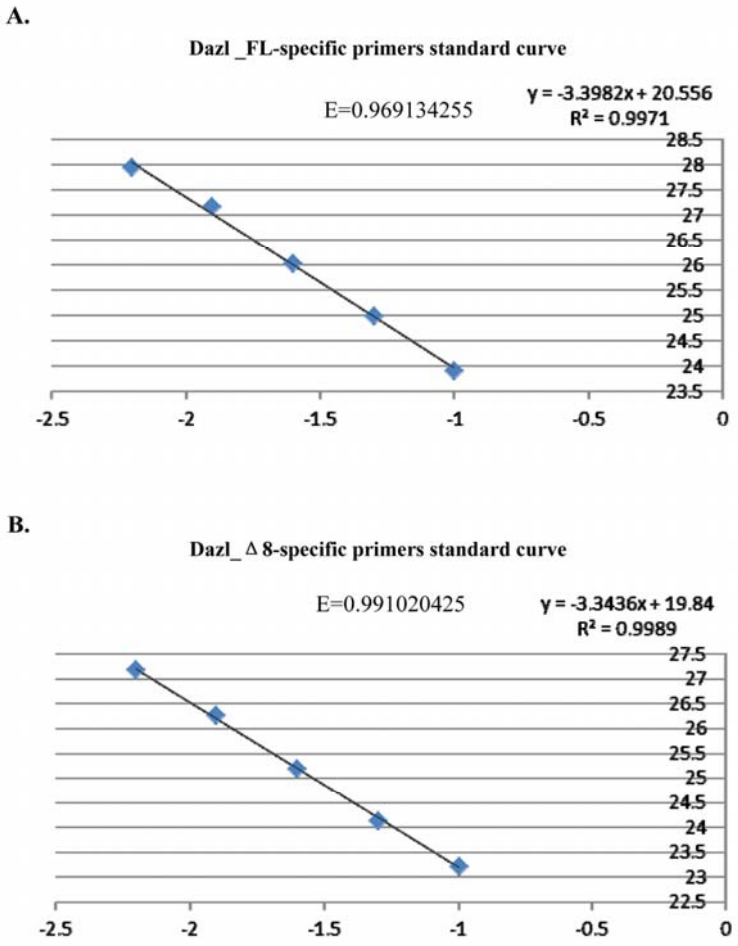


Figure S1. Amplification efficiency of Dazl isoform-specific primers. The standard curve graph showing the amplification efficiency of Dazl_FL-specific (A) and Dazl_Δ8-specific (B) primer pairs.

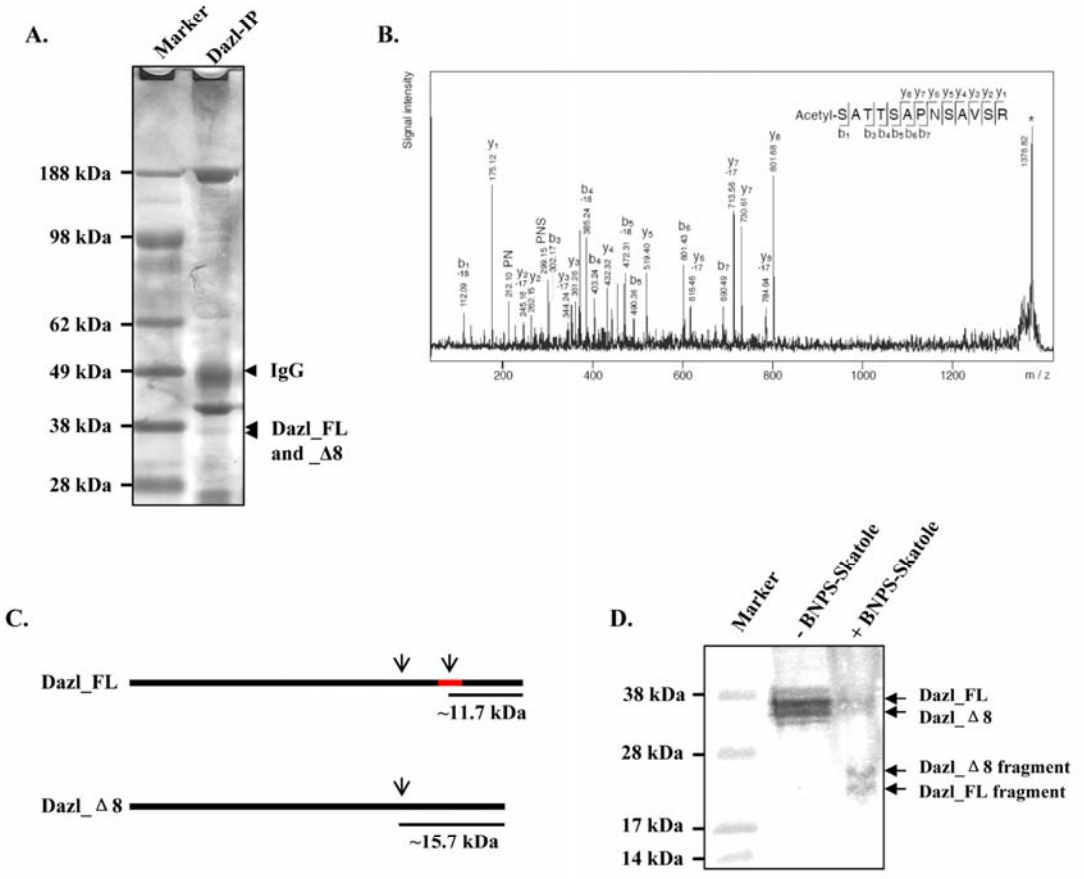


Figure S2. Biochemical characterization of Dazl protein isoforms. (A) Coomassie blue stained SDS-PAGE gel showing the Dazl-IPed proteins. The Dazl_FL and Dazl_Δ8 protein isoforms as well as IgG-heavy chain are indicated with arrowheads. **(B)** The isoforms of Dazl (Dazl_FL + Dazl_Δ8) purified by SDS-PAGE were subjected to tryptic in-gel digestion and LC-MALDI mass spectrometry. The fragment ion spectrum of the N-terminal peptide (precursor mass 1419.67) is shown, which indicates that the N-terminal amino acid sequence starts from serine at position 2 carrying an N-terminal acetyl group. The b- and y-ion series and internal fragments are indicated, the asterisk marks a fragment generated by the neutral loss of the acetyl group. (C) Schematic diagram showing the Dazl_FL and Dazl_Δ8 protein isoforms along with the BNPS-Skatole cleavage sites (arrows). Exon8 coding region is highlighted as red in Dazl_FL protein. The expected C-terminal protein fragments along with their predicted molecular weights are indicated below the each protein. (D) Western blot showing the detection of Dazl protein isoforms either in absence or presence of BNPS-Skatole.

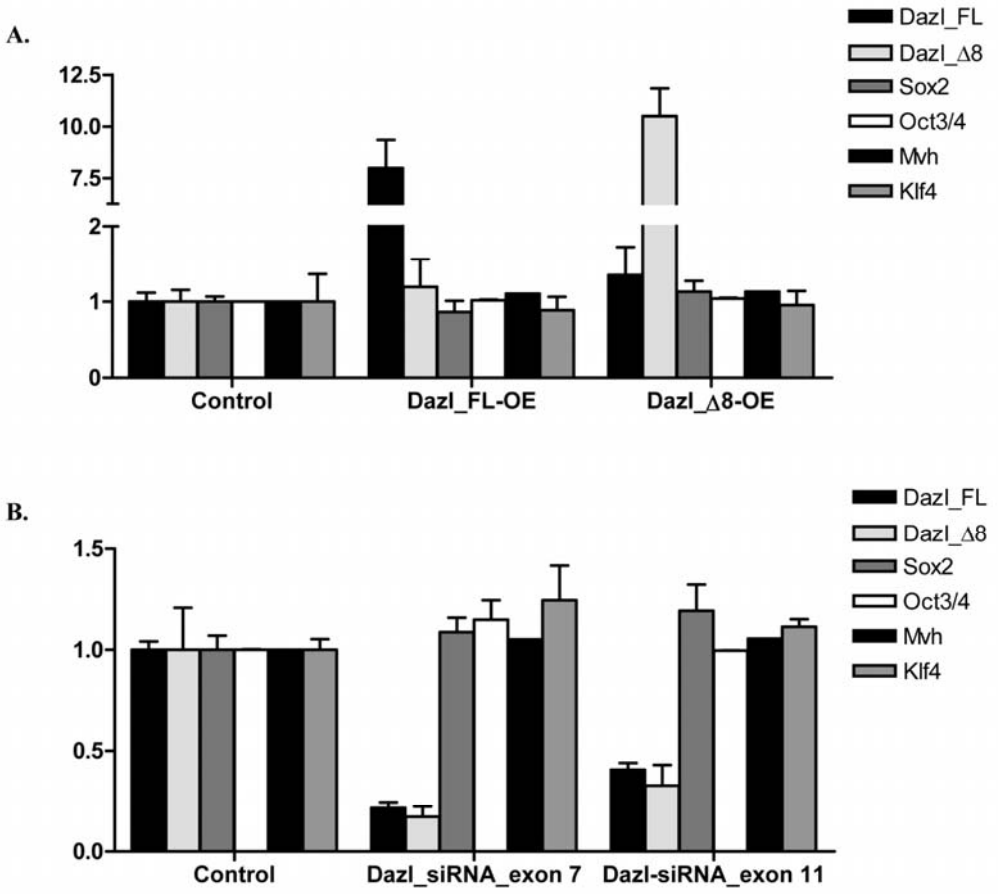


Figure S3. Alteration of Dazl expression has no effect on selected mRNA expression levels. qRT-PCR data showing the transcript levels of Dazl isoforms and target (*Sox2*, *Oct3/4* and *Mvh*) and control (*Klf4*) mRNAs in control ESCs and Dazl splice variants overexpressing (Dazl_FL and Dazl_Δ8_E2) ESCs **(A)** and Dazl downregulated (Dazl_siRNA_exon7 and Dazl_siRNA_exon 11) ESCs **(B)** .

4. Discussion

Embryonic stem cells (ESCs) derived from the inner cell mass (ICM) of blastocysts are capable of differentiating into all cell types of the adult organism and are regarded as the "gold standard" of pluripotency (Smith et al., 2009; Bilic et al., 2012). These unique characteristics imply that the production of lineage-specific cell types differentiated from ESCs might be possibly applied in cell replacement therapies (Lerou and Daley, 2005). Therefore, they have become the focal point of a burgeoning field of biomedical research and regenerative medicine. Germ cells are the truly and unique immortal cells which are able to surpass the genetic information from one generation to the next, indefinitely (McLaren, 2001). Interestingly, germ stem cells, the unipotent stem cells, isolated at various stages of germ cell development have a unique ability to transform into pluripotent stem cells (Kahan et al., 1970; Matsui et al., 1992; Resnick et al., 1992; Guan et al., 2006; Kanatsu-Shinohara et al., 2004). Moreover, several studies have shown the expression of several germ cell marker genes in pluripotent stem cells including ESCs (Geijsen et al., 2004; Qing et al., 2007; Mise et al., 2008) and also reported their spontaneous differentiation towards germ cells (Hübner et al., 2003; Geijsen et al., 2004; Lacham-Kaplan et al., 2006). These results highlight the possible reciprocal relationship between germ cells and pluripotent stem cells and even caste the doubt on the origin of ESCs.

In the present study, we systematically analyzed the expression of selected germ stem cell (GC) and premeiotic (PrM) marker genes in several pluripotent stem cells (Xu et al., 2011). Our results indicate that the expression of GC and PrM marker genes are important for the establishment of pluripotency and germ line commitment of pluripotent stem cells, respectively. Moreover, our findings implied that the expression of GC marker genes triggered by BMP signals parallels between *in vivo* germ cell specification and *in vitro* somatic cell reprogramming to induced pluripotency, indicating the establishment of pluripotency probably via the germ cell fate/origin. In line with these assumptions, our further results indicate that the expression of *Dppa3*, a GC marker gene, is important for the establishment of authentic pluripotency during somatic cell reprogramming, while the expression of *Dazl*, a PrM marker gene, can lead to the generation of putative germ cells.

4.1. Spermatogenesis reconstituted from pluripotent stem cells *in vitro*

Over 10% of couples in the world are suffering from infertility and the males are account for approximately 50% of all cases (Schlegel, 2009). Male infertility can be

attributed by spermatogenic arrest at different stages of meiosis. Therefore, the derivation of haploid gametes *ex vivo* is being considered as an alternative approach to treat male infertility. Recent advances in germ cell research have shown the possibility to generate haploid germ cells using *in vitro* system (Hübner et al., 2003; Nayernia et al., 2006; Yu et al., 2009).

The derivation of haploid germ cells from various pluripotent stem cells has been established almost a decade ago. The first study reported the spontaneous differentiation of ESCs towards germ cell lineage under LIF and feeder depleted culture conditions (Hübner et al., 2003). Thereafter, several distinct strategies were also proved the possibility to obtain PGCs and even haploid gametes from ESCs by mimicking the *in vivo* environment in embryoid bodies (Toyooka et al., 2003; Geijsen et al., 2004; Lacham-Kaplan et al., 2006), by the activation of germ cell differentiation pathways using growth factors (Nayernia et al., 2006) or through overexpression of PrM marker genes (Yu et al., 2009). This differentiation property towards germ cell lineage is not only restricted to ESCs but also demonstrated for other pluripotent stem cells such as embryonal carcinoma cells (ECCs) (Nayernia et al., 2004), embryonic germ cells (EGCs) (Eguizabal et al., 2009), multipotent adult germ line stem cells (maGSCs) (Nolte et al., 2010) and induced pluripotent stem cells (iPSCs) (Imamura et al., 2010; Zhu et al., 2012; Yang et al., 2012). In order to facilitate human reproductive medicine, these protocols were applied to human model and successfully produced primordial germ cells (PGCs) from human ESCs (Tilgner et al., 2008; West et al., 2008; Bucay et al., 2009; Aflatoonian et al., 2009) and human iPSCs (Park et al., 2009). Additionally, several groups have reported that human pluripotent stem cells can enter meiosis and produce haploid cells (Eguizabal et al., 2011; Panula et al., 2011). Thus, many reproducible results have highlighted the feasibility to produce germ cell or haploid cells from pluripotent cells. Hereby this *in vivo* system can be used as an excellent tool to study germ cell development and signalling processes in an easy to access model. Moreover, it holds the therapeutic potential for treatment of male infertility.

Strikingly, several of the above mentioned studies have noted the expression of some of the GC and PrM marker genes already in undifferentiated pluripotent cells (Geijsen et al., 2004; Qing et al., 2007; Mise et al., 2008). These observations led to a suggestion that ESCs, apart from germ line derived pluripotent stem cells, might have a germ cell origin or the expression of GC marker genes might be important for the establishment/maintenance of pluripotency.

4.2. Stage-specific germ cell marker genes are expressed in all mouse pluripotent cell types and emerge early during induced pluripotency

Since the techniques for successful generation of ECCs and ESCs were established in the mouse, a series of pluripotent stem cell lines were established. Historically, the pluripotent EGCs, ECCs and maGSCs have been documented, all with germ line origins (Kahan et al., 1970; Matsui et al., 1992; Resnick et al., 1992; Guan et al., 2006; Kanatsu-Shinohara et al., 2004). However, till now there is no convincing result to show either the ICM or the germ cell origin of ESCs. High similarities between germ cells and ESCs made stem cell researchers to speculate the possible germ cell origin of ESCs (Zwaka et al., 2005; Nichols and Smith 2011). Traditionally, ESCs are regarded as an *in vitro* equivalent of the ICM, but it was shown that ESCs more closely resemble cells derived from the primitive ectoderm (Gardner and Brook, 1997). Recently, two independent groups have shown that only a small subset of cells in the ICM, which are positive for GC marker genes, are able to give rise to ESCs *in vitro* (Nichols et al., 2009; Chu et al., 2011). The authors constructed a reporter system by fusing the promoter region of *Blimp1,* the GC determinant, to a fluorescent protein to monitor the signal during the process of ESC establishment. Notably, the authors observed the positive signal from 2 days after blastocyst culture; moreover, a clear outgrowth resembling ESC-like colony appeared only 4 days after the culture (Chu et al., 2011). As expected, the Blimp1 positive cells also showed expression of other GC marker genes such as *Dppa3*, *Ifitm3*, *Prdm14* and *Lin28* (Chu et al., 2011). This outcome reflected that there is an intermediate state of PGC-like property acquisition during the early derivation process of ESCs. In agreement with these results, our study (Xu et al., 2011) could show that the GC marker genes possess an active chromatin signature at their promoter regions in ESCs, which could be a reminiscent of PGC fate determination during early events of ESC establishment.

These observations raised the question whether there is also a temporary PGC fate during iPS cell generation. In our study (Xu et al., 2011), we attempted to answer this question by direct reprogramming of somatic cells into iPSCs and by analysing the activation of selected GC/PrM markers along with the activation of core pluripotency marker genes during this process. Intriguingly, the activation of GC markers *Blimp1, Dppa3* and *Ifitm3* was found to precede the activation of the endogenous pluripotency markers such as *Oct3/4, Sox2* and *Nanog*. In contrast, the expression of PrM marker genes such as *Dazl, Mvh* and *Stra8* was detected only in established iPS cell lines. Meanwhile, two elegant studies revealed that Mesenchymal-to-Epithelial Transition (MET) is necessary

during the initiation stage of reprogramming and this key event is driven by the BMP signalling pathway (Li et al., 2010; Samavarchi-Tehrani et al., 2010). Furthermore, all ESC-like characteristics are acquired by sequential expression of cell adherence marker genes *Cdh1* and *β-catenin*, followed by concomitant upregulation of early pluripotency marker genes such as *AP* and *SSEA1* (Samavarchi-Tehrani et al., 2010). By a coincidence, at early gastrulation of embryo development, PGC precursors are induced by an instructive BMP signalling followed by MET and the expression of GC marker genes *Blimp1, Dppa3, Ifitm3* along with *Cdh1* (Matsui, 2009; Saitou, 2009). Notably, loss of BMP signal leads to lack of PGCs *in vivo* and inactivation of BMP results in failure to generate iPSCs from somatic cells (Lawson et al., 1999; Saito, 2009). Collectively, through our expression analysis and observation of molecular parallels between PGC specification and iPSC generation, we speculate that even somatic cells have to pass through a temporary PGC fate to achieve pluripotency during reprogramming.

4.2.1. Expansion of PGC fate model to human somatic cell reprogramming

To assess whether the temporary PGC fate is also recapitulated during human somatic cell reprogramming, we decided to analyze the GC, PrM and pluripotency marker genes expression during reprogramming of adult human fibroblast cells. In the first pilot screening, we decided to collect one sample from each three distinct phases (initiation, maturation and stabilization) of reprogramming, i.e., at day 14, day 20, and day 30 as well as wild type fibroblast cells and established human induced pluripotent stem cells (hiPSCs) as negative and positive controls, respectively. Quantitative real time PCR analysis revealed a very high significant expression levels of GC markers (*Blimp1* and *Ifitm3*) already in wild type fibroblasts and these expression levels were gradually reduced and reached to levels seen in established iPSCs by day 30 (**Fig. 4.1**). In contrast, the expression of *Dppa3,* the other GC marker gene, was only detectable at day 20 and significantly increased to 100-fold than the expression levels of iPSCs (**Fig. 4.1**). The endogenous expression of key pluripotency markers *Oct4*, *Sox2* and *Nanog* appeared by day 20 and showed increased expression by day 30 (**Fig. 4.1**). Surprisingly, we could detect a significant expression of PrM marker gene *Dazl* on day 14 of reprogramming with an increased trend till day 30 (**Fig. 4.1**). Collectively, these preliminary results indicate that the expression pattern of GC and PrM marker genes before and during somatic reprogramming differ between mouse and human fibroblasts. In order to rule out that the expression of GC marker genes is specific to this particular fibroblast cell line which we

used for reprogramming studies, we analyzed and found the expression of GC marker genes in 4 different human fibroblast cell lines (data not shown).

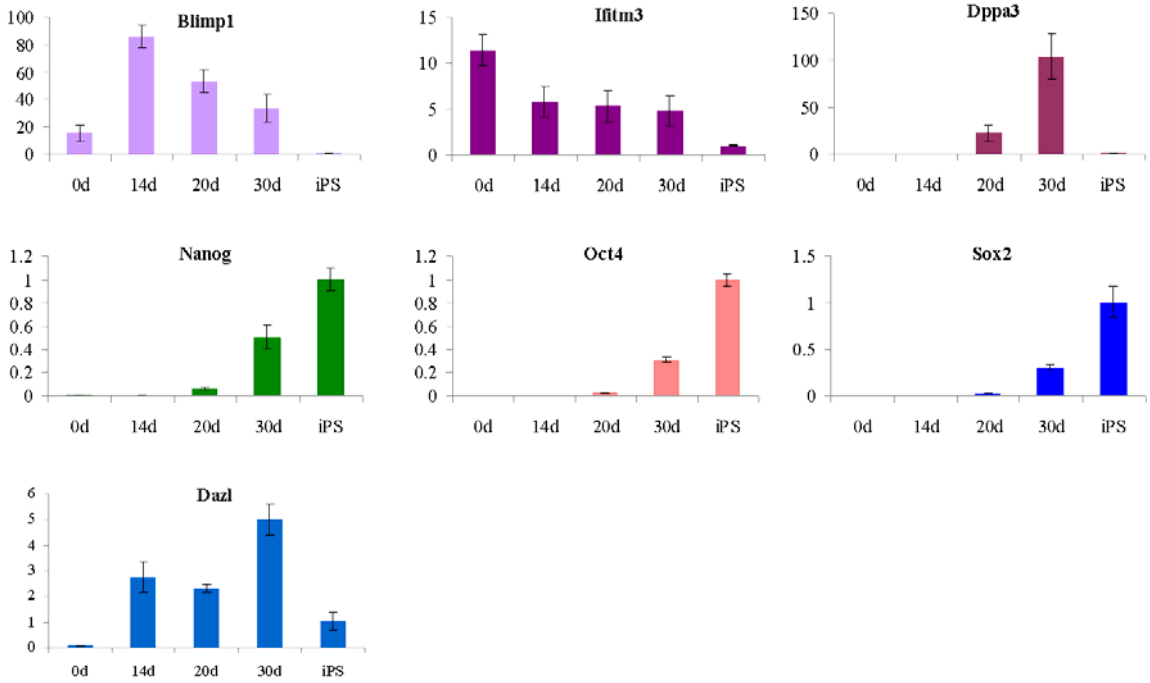


Figure 4.1. Expression pattern of GC/PrM- and pluripotency-marker genes during the course of hiPS generation. Quantitative real time qPCR analysis of germ stem cell (*Blimp1*, *Ifitm3*, and *Dppa3*), pluripotency *(Oct3/4*, *Sox2*, and *Nanog*), and pre-meiotic (*Dazl*) genes during the time course of iPS cell generation. 0d- wild type human fibroblasts; 14d-, 20d-, 30d- : day 14, day 20, and day30 after virus infection; iPS- established human iPS cells. The qPCR data of three technical replicates were calculated and represented as a mean ±SD.

Although human ESCs are broadly used as a model for biomedical research, they are more difficult to manipulate than their counter parts in mouse (Nichols and Smith, 2011). Two distinct groups of pluripotent cells are derived from mouse early embryonic development: (1) ESCs are established from E3.5- E4.5 blastocysts and (2) epiblast stem cells (EpiSCs) are derived from E5.5- E6.5 post implantation embryos (Evans and Kaufman, 1981; Brons et al., 2007; Tesar et al., 2007). ESCs are regarded to bear a ground state (naïve) of pluripotency; however, EpiSCs hold a "primed" state of pluripotency. Unlike ESCs, EpiSCs are not able to contribute to chimeras after their injection into blastocysts (Nichols and Smith, 2009). Interestingly, the high molecular similarities between human ESCs and mouse EpiSCs led to propose that human ESCs, in contrast to mouse ECS, are in a "primed" state of pluripotency (Nichols et al., 2009). Taken together, we speculate that there are fundamental differences between mouse and human ESCs and somatic reprogramming and this could be partly attributable to the cell intrinsic and species

specific mechanisms. However, further studies are necessary to identify the PGC fate during human iPSC generation.

4.2.2. Transdifferentiation of somatic cells into induced germline stem cells

The observation that GC marker genes are expressed earlier than that of pluripotency marker genes and the possible temporary PGC fate during somatic cell reprogramming prompted us to evaluate the feasibility to derive germ line stem cells (GSCs) directly from somatic cells. At the meantime, it was shown that mouse ESCs and iPSCs are able to generate PGC-like cells through an epiblast-like state, and these cells have been demonstrated for their capability to initiate spermatogenesis (Hayashi et al., 2011). Based on these observations, we explored a condition in which fibroblast cells can transdifferentiate into germline stem cells.

In order to screen the candidate GC factor(s) that induces transdifferentiation of fibroblast cells to a germ cell state, we used OSKM (*Oct3, Sox2, Klf4* and *c-Myc*) in combination with GC marker genes (*Blimp1, Ifitm3, Dppa3*) for reprogramming studies as outlined in **figure 4.2**. Approximately after two weeks of transduction, few round and clustered cells were formed and resembled PGC/SSC-like morphology in the combination of OSKM+Blimp1 (OSKMB) (**Fig. 4.3 A**). In contrast, OSKM cells showed compacted ESC like colonies (**Fig. 4.3 A**). The induced GSCs (iGSCs) obtained in OSKMB combination were able to proliferate and showed typical grape-like morphology of germ cells and are positive for alkaline phosphatase (*AP*) staining (**Fig. 4.3 A**).

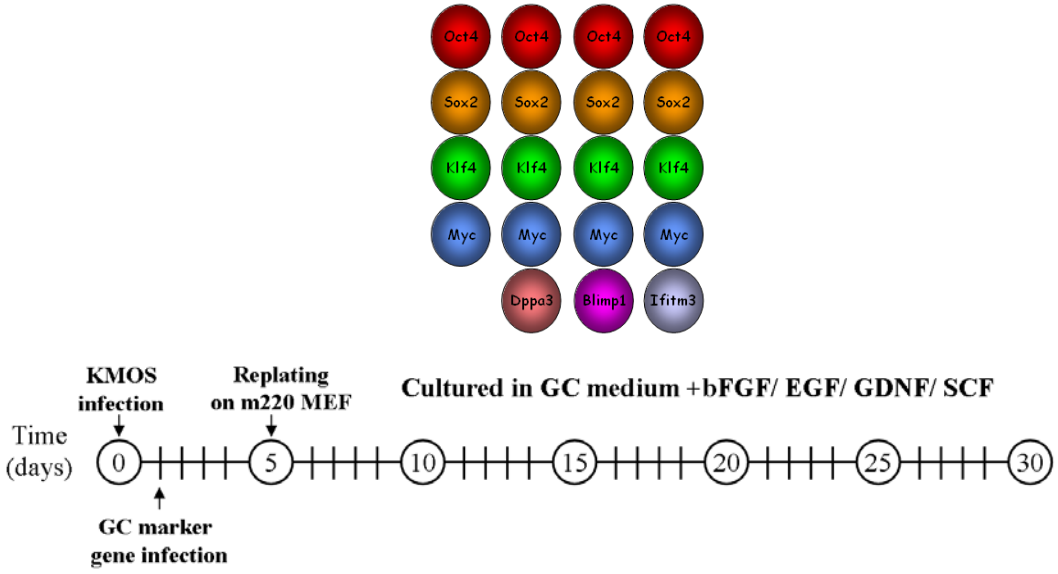


Figure 4.2. Schematic diagram depicting the generation of iGSCs. The iGSCs were generated by using different combinations of reprogramming factors along with GC factors and growth factors, which are supplemented to the medium to facilitate the generation of iGSCs.

To further characterize the iGSCs, we performed RT-PCR analysis for the expression of GC marker genes *Oct4*, *Cdh1*, *Plzf*, and *Gfrα-1*. These studies revealed that the expression of typical GC marker genes Plzf and *Gfrα-1*are highly expressed in iGSCs and spermatogonial stem cells (SSCs), an established SSC line or testis, but not in ESCs (**Fig. 4.3 B**). In addition, these iGSCs were positive for Plzf and Gfrα-1 expression as judged by immunocytochemistry (**Fig. 4.4 A**), however, the expression of pluripotent cell marker SSEA1 was visible only in ESCs but not in iGSCs and SSCs (**Fig. 4.4 B**).

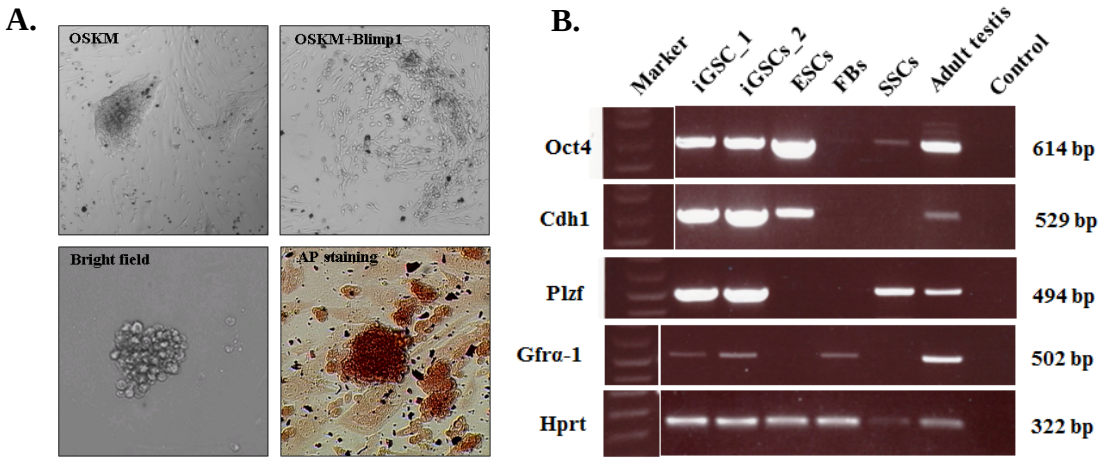


Figure 4.3. Generation and characterization of iGSCs. (**A**) Colony morphology of iPSCs generated with OSKM or iGSCs generated in the presence of OSKMB. (**B**) RT-PCR analysis for the expression of GC genes in two iGSC lines, ESCs, fibroblasts (FBs), SSCs, and WT- testis. Water (control) was used as a negative control.

To rule out that iGSCs are not a pluripotent cell type, we injected these cells into immunodeficient mice and observed a malignant tumour but not the teratoma, suggesting that these cells are not pluripotent. Next, to assess the functional properties of iGSCs in reconstitution of spermatogenesis *in vivo*, we transplanted these cells into the testis of busulfan treated mice. Unfortunately, these mice developed testicular tumours suggesting that the derivation of iGSCs protocol needs to be optimized. Collectively, in this part of our study, we developed an *in vitro* method that achieves three significant endpoints. First, the generation of PGC- or SSC-like cells from somatic cells is possible and it gives a strong evidence for our hypothesis that there is a temporary germ cell fate before somatic cells enter into the pluripotent state. Second, this is the first ever report that successfully generates putative germ cells directly from somatic cells without entering any intermediate cell stage(s). The mechanism of PGC development is difficult to study mainly because of the inaccessibility of PGC during development. Our PGC-like cell induction system thus might facilitate the understanding of this complicated process. Third, our results agree with

those in a previous report, which showed that Blimp1 expression mediates the germ cell fate commitment (Ohinata et al., 2005). These data may therefore also have ramifications for reproductive technology as well as for disease modelling approaches.

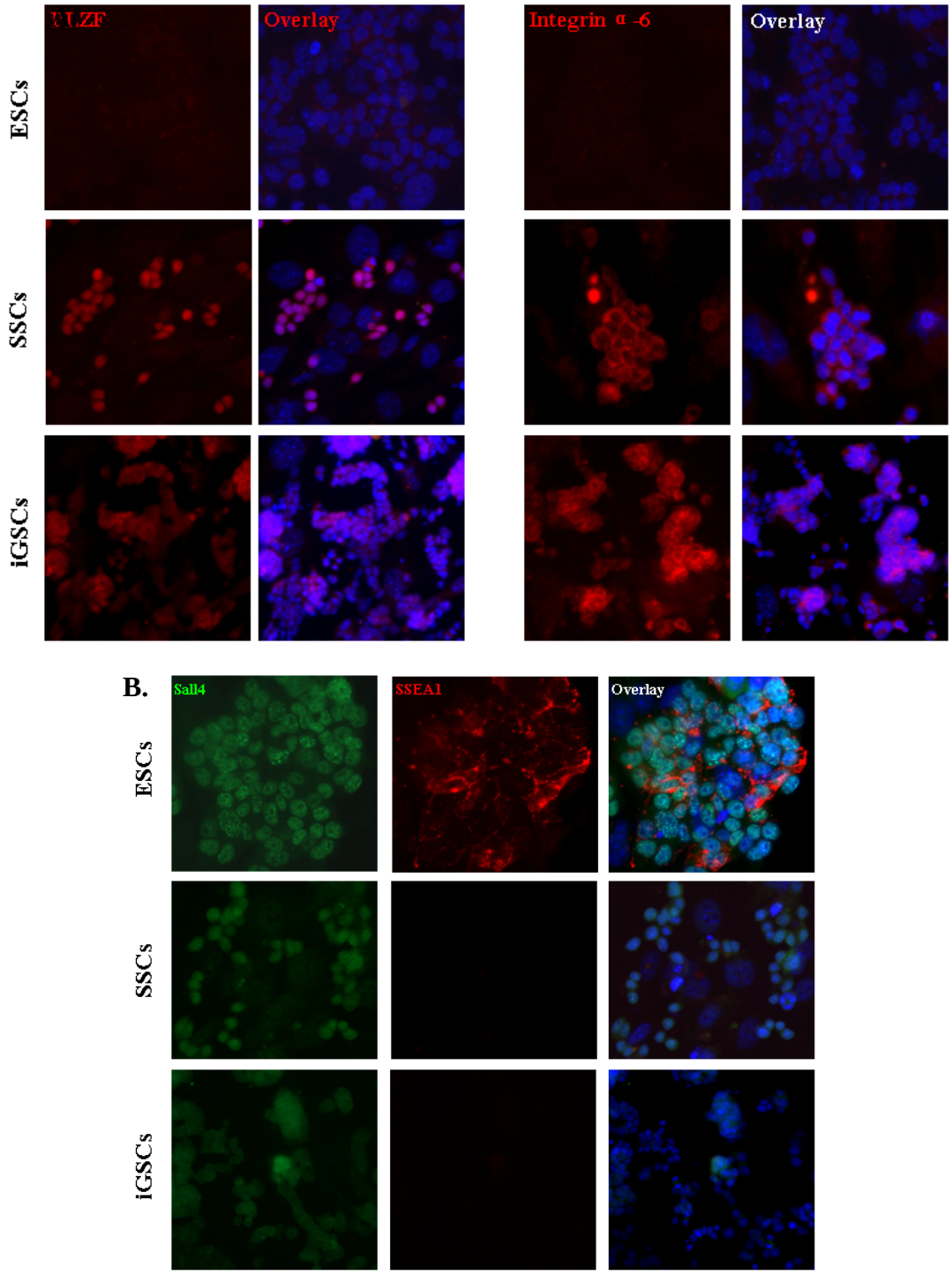


Figure 4.4. Characterization of iGSCs by immunocytochemistry. Immunocytochemical analysis for Plzf and integrin α-6 (**A**) Sall4 and SSEA1 (**B**) in ESCs, SSCs and iGSCs.

4.3. Dppa3 binds to the IG-DMR of the *Dlk1-Dio3* imprinting cluster and prevents its imprinting loss during iPS cell generation

In the past few years, a variety of reprogramming methods have been established to generate iPS cells from somatic cells. Several standards for the evaluation of iPSCs pluripotency and their functional equivalence to ESCs have been accepted including molecular tests and *in vivo* developmental analysis. Among all these criteria, germ-line transmission was proposed to be one of the most stringent assessments reflecting the authenticity of iPS cells. Recently, two independent studies reported that the majority of mouse iPS cell lines show abnormal imprinting at the *Dlk1-Dio3* locus and this leads to the loss of germline transmission of chimeras (Stadtfeld et al., 2010, Liu et al., 2010). The expression of *Gtl2,* a non-coding RNA of the *Dlk1-Dio3* cluster, is being regarded as a *bona fide* marker for assessment of authentic pluripotent cell lines and thereby the germ line transmission competency.

Genomic imprinting is an epigenetic process of gene regulation during gametogenesis in mammals. Imprinted genes are characterized by differential DNA methylation and monoallelic expression according to the parent of origin-dependent manner (Spahn and Barlow, 2003). Imprinting control regions (ICRs) mediates the expression pattern of imprinted genes. Interestingly, many of the imprinted genes are grouped in domains, and the expression is regulated by *cis*-acting control elements which can repress multiple imprinted genes of the same domain. One such particular imprinting domain is at the *Dlk1-Dio3* locus (delta-like 1 homolog-type III iodothyronine deiodinase) which is located on distal mouse chromosome 12 (12qF1) and human chromosome 14q32. Abnormal imprinting pattern at the *Dlk1-Dio3* locus in humans is known to disrupt the function of the genes located in this locus and leads to a spectrum of disorders (**table 4.1**).

	matUPD14	Paternal deletion of *Dlk1-Gtl2*	Paternal deletion of *Wars25-snoRNAs*	patUPD14	Maternal deletion of *Dlk1-Gtl2*	Maternal deletion of *Wars25-snoRNAs*
Predicted gene dosage[b]	*DLK1*: 0x *ncRNAs*: 2x *RTL1*: 0x	*DLK1*: 0x *ncRNAs*: 1x *RTL1*: 1x	*DLK1*: 0x *ncRNAs*: 1x *RTL1*: 0x	*DLK1*: 2x *ncRNAs*: 0x *RTL1*: 4x	*DLK1*: 1x *ncRNAs*: 0x *RTL1*: 4x	*DLK1*: 1x *ncRNAs*: 0x *RTL1*: 2x
Facial appearance	High forehead, short philtrum, downturned mouth	Facial dysmorphism similar to matUPD14	Facial dysmorphism similar to matUPD14	Face—hairy forehead, blepharophimosis, micrognathia and prominent philtrum	Facial dysmorphism as in patUPD14	Facial dysmorphism similar to (but milder than) patUPD14
Growth	Pre- and postnatal growth retardation	Pre- and postnatal growth retardation	Pre- and postnatal growth retardation	Normal prenatal growth;[c] severe postnatal growth retardation	Normal prenatal growth;[c] postnatal growth retardation	Normal prenatal growth[c]
Motor system	Joint laxity; muscle hypotonia; Motor delay	Motor delay	Motor delay	Developmental delay; joint contractures; feeding difficulties	Developmental delay; joint contractures; feeding difficulties	Developmental delay; feeding difficulties
Skeletum	Reduced height (by accelerated skeletal maturation); dysmorphic face, hands and feet	Reduced height; dysmorphic features of face, hands and feet	Reduced height; dysmorphic features of face, hands and feet	Small and typical bell-shaped thorax with 'coat-hanger' appearance of the ribs	Small and typical bell-shaped thorax with 'coat-hanger' appearance of the ribs	Small and typical bell-shaped thorax with 'coat-hanger' appearance of the ribs
Placenta	nd	nd	nd	Placentomegaly	Placentomegaly	Placentomegaly
Neurological processes	Normal/mild mental retardation	No mental retardation	No mental retardation	Moderate to severe mental retardation	ND	ND
Specific characteristics	Early puberty, recurrent otitis media; occasional obesity	Early puberty, occasional obesity	No early puberty	Complicated pregnancies with polyhydramnios and premature delivery; abdominal wall defects	Complicated pregnancies with polyhydramnios and premature delivery; abdominal wall defects	Complicated pregnancies with polyhydramnios and premature delivery; mild abdominal wall defects

[a]Dlk1-Dio3, delta-like 1 homolog-type III iodothyronine deiodinase; ncRNA, noncoding RNA; ND, not determined.
[b]Gene dosage of the *DLK1-DIO3* domain in humans is difficult to monitor because these genes are weakly or not expressed in leukocytes; the predicted gene dosage is based mainly on expression studies in human placentas and from predictions based on the known dosage of these genes in humans and mice and considering the extent of the deletions.
[c]patUPD14 patients are frequently delivered prematurely; their growth is normal compared with gestational age-matched reference.

Table 4.1 Pathological anomalies in human harbouring genetic defects at the *Dlk1-Dio3* locus. matUPD, maternal uniparental disomy; patUPD, paternal uniparental disomy; snoRNA, small nucleolar RNA (Table adapted from Rocha et al., 2008)

In our previous study (Xu et al., 2011), we observed that the GC marker genes are activated earlier than endogenous pluripotency marker genes during somatic cell reprogramming. These results promoted us to investigate the role of GC marker genes in establishment of authentic pluripotency. In the present study (Xu et al., 2012, manuscript under review), we show that all iPS cell clones generated in presence of *Dppa3* show normal imprinting at the *Dlk1-Dio3* imprinting cluster. At the molecular level, we show that Dppa3 counteracts the binding of Dnmt3a, a *de novo* DNA methyltransferase, to this region thereby probably preventing the aberrant DNA methylation at this region during reprogramming. In line with a recent report, which showed that vitamin C can abolish the imprinting defect during reprogramming (Stadtfeld et al., 2011), we also observed that vitamin C can prevent the loss of imprinting at the *Dlk1- Dio3* locus. Moreover, we observed a positive correlation between vitamin C and Dppa3, where the addition of vitamin C dramatically activated the expression of Dppa3 during very early stages of reprogramming. This early activation of Dppa3 might be sufficient for the maintenance/establishment of imprinting at the *Dlk1-Dio3* locus. Thus, our study establishes Dppa3 as a genetic factor responsible for proper maintenance of imprinting during reprogramming and for the establishment of authentic pluripotent stem cells.

4.3.1. Functional analysis of Dppa3 during somatic cell reprogramming by inducible transgenic mouse model

To gain more insights into the function of *Dppa3* in establishment of authentic pluripotency and to study in which stage of reprogramming Dppa3 is important. We took advantage of doxycyclin-inducible system where the expression of Dppa3 can be temporally controlled. Towards this end, we constructed pTre_Dppa3 expression vector by cloning mouse *Dppa3* open reading frame into pTre 3G vector. An ectopic hemagglutinin (HA) protein tag was added to the C-terminus of *Dppa3* in order to facilitate immunoprecipitation assays afterwards (**Fig. 4.5 A**). The pTre_Dppa3 expression construct was then used to generate a transgenic mouse line. At the same time, an additional transgenic mouse line carrying the transactivator was generated. Further, the double transgenic mouse line was generated by inter-breeding of the two transgenic lines and currently, we are on the way to produce embryonic fibroblast cells and to perform iPS cells generation as outlined (**Fig. 4.6**). Briefly, after OSKM transduction, the *Dppa3* transgene expression will be induced at indicated time point of reprogramming using doxycycline. The reprogramming time course samples in which the Dppa3 was activated at different time points will be analysed for *Gtl2* expression. Moreover, the genome wide transcriptome might be necessarily performed from different samples to address the question of boosting reprogramming efficiency.

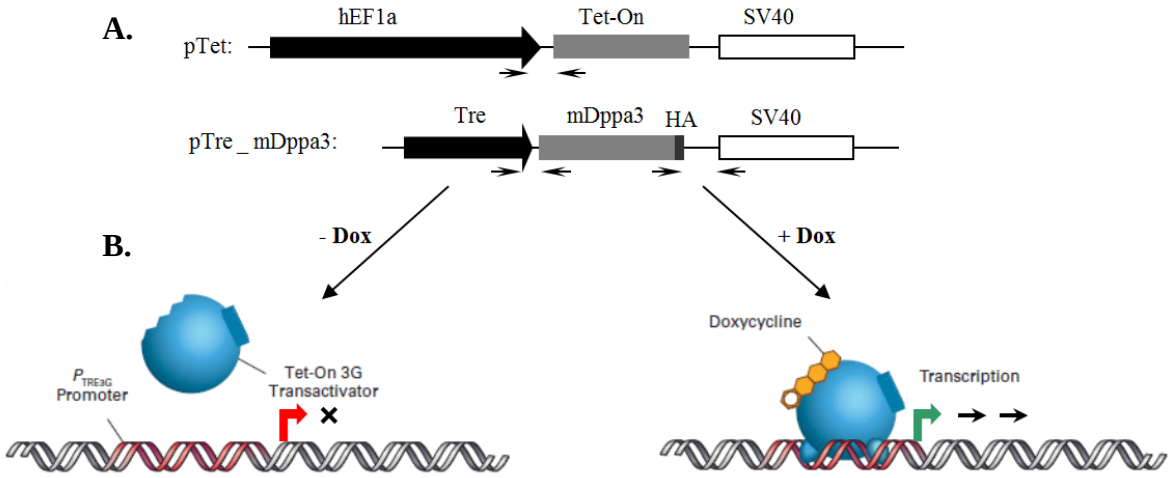


Figure 4.5. Schematic representation of Tet-on 3G system and Dppa3 transgenic mouse model generation. (**A**) Representative images showing the construct structure of Tet-on vector and pTre_Dppa3. The genotyping primers are indicated by arrows. (**B**) In the absence of doxycycline (-dox), the generated double transgenic embryonic fibroblasts will have no *Dppa3* transgene expression (left panel), whereas the addition of doxycycline (+dox) to the culture medium induced the activation of Dppa3. (Images were adapted and modified from www.clonetech.com)

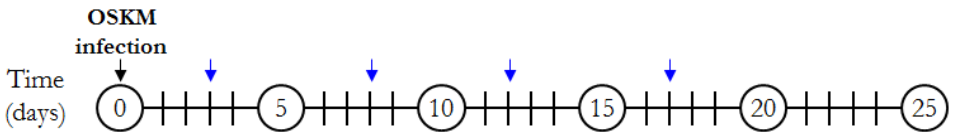


Figure 4.6. Schematic representation of induced Dppa3 expression during somatic cell reprogramming. The classic reprogramming factors (OSKM) will be delivered by retroviral infection. The expression of *Dppa3* will be activated at indicated time points during the reprogramming by adding doxycycline (blue arrows) to the culture medium.

4.4. Mouse Dazl and its novel splice variant function in translational repression of target mRNAs in mouse embryonic stem cells

ESCs are regarded as a suitable model to study germ cell development *in vitro* and hold a great potential for reproductive medicine (Daley, 2007). *In vitro* generation of haploid gametes has been partly successful. The sperm-like cells derived from ESCs were able to fertilize oocytes by intracytoplasmic sperm injection (ICSI) (Hübner et al., 2003; Toyooka et al., 2003; Clark et al., 2004; Geijsen et al., 2004; Nayernia et al., 2006). A recent publication showed that the overexpression of Dazl, a PrM marker gene, in mouse ESCs leads to the generation of both motile tailed-sperm and oocytes (Yu et al., 2009). Subsequent functional studies showed that these artificial gametes could be used to produce early embryos until 8-cell stage (Yu et al., 2009). This study further supports our assumption (Xu et al., 2011) that the expression of PrM marker genes is important for the germ cell lineage commitment of pluripotent stem cells.

Since 2004, artificial sperm has been derived *in vitro* in our group using several pluripotent cell types (Nayernia et al., 2004; Nayernia et al., 2006; Nolte et al., 2010). In general, our protocol is based on retinoic acid induced differentiation, analysis for the activation of the EGFP reporter from Stra8 promoter and subsequent separation of cells using fluorescence activated cell sorting (FACS). To overcome the transgene overexpression and reporter based approaches, which might cause the insertional mutagenesis, we attempted to use protein transduction strategy to generate germ cell-like cells. Previous studies have demonstrated that the conjugation of a short peptide such as HIV-Tat to a recombinant protein can facilitate the delivery of target protein into cells (Becker-Hapak et al., 2001; Albarran et al., 2005; Rayapureddi et al., 2010). To generate recombinant Dazl protein that can penetrate through the cell plasma membrane, we decided to fuse HIV-Tat polypeptide to the C terminus of Dazl (**Fig. 4.7 A**). The recombinant protein was expressed in *E. coli* and subsequently purified by using HisTrap purification

system (**Fig. 4.7 B**). The Western blot analysis of the purified recombinant protein confirmed the successful purification of Dazl-Tat protein (**Fig. 4.7 C**). In order to test the function of Dazl-Tat recombinant protein in terms of cell permeability and stability, we treated immortal mouse fibroblast (NIH-3T3) cells. We observed that the Dazl-Tat recombinant protein easily penetrates cells (**Fig. 4.7 D**) and is stable inside of cells for at least 48 hours (data not shown).

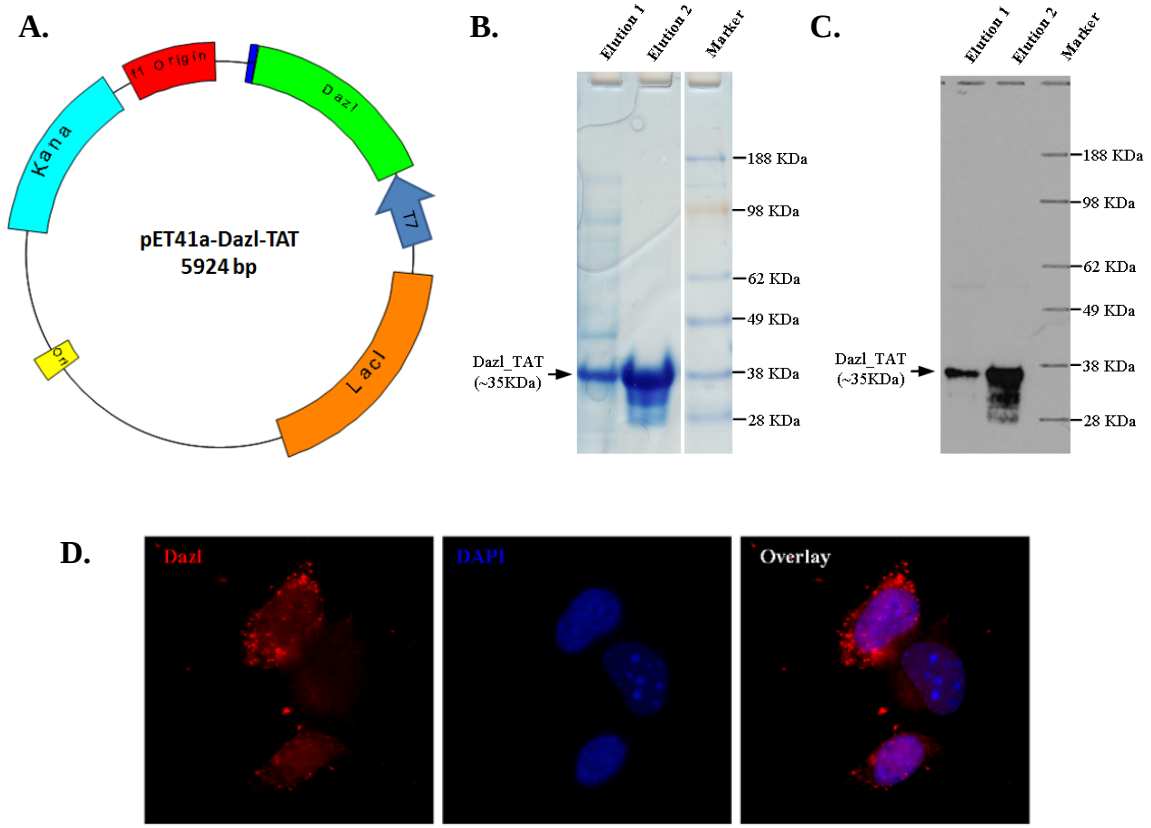


Figure 4.7. Dazl-Tat recombinant protein purification and functional test. (**A**) The representation of pET41a-Dazl-Tat construct, HIV-Tat polypeptide (blue) was tagged on to C terminus of mouse *Dazl* open reading frame. (**B**) Recombinant protein elution fraction Nr.1 and Nr.2 were analyzed on SDS-PAGE gel and staining with commassie blue. (**C**) The identity of recombinant protein elution fraction Nr.1 and Nr.2 were assessed by Western blot with Dazl antibody. (**D**) Dazl_Tat recombinant protein can penetrate the cell membrane and showed cytoplasmic locolization.

We then applied Dazl-Tat recombinant protein to ESCs in germ cell differentiation medium (Yu et al., 2009). However, we were unable to identify any sperm-like or oocyte-like cells in our system. This could be due to the Dazl-Tat protein concentration levels which might not have been sufficient to trigger the germ cell differentiation. In contrast, the study by Yu et al (2009) employed lentiviral expression system, which gives the expression several orders of magnitudes than our protein delivery system.

During the course of this study, we identified a novel splice variant of mouse Dazl (*Dazl_Δ8*), which lacks exon 8 of Dazl. We then characterized the expression and localization pattern of this novel Dazl splice variant in pluripotent stem cells as well as in germ cells (Xu et al., 2012, manuscript under review). By comparison of protein domain structures, we found that a part of proline-rich (PR) region was lacking in the Dazl_Δ8 isoform. It is known that PR regions mediate protein-protein interactions (Williamson, 1994). Therefore, it might be possible that Dazl_Δ8 isoform is not able to interact with proteins such as Poly (A) binding family proteins and might lead to interaction with other unknown proteins thereby to regulate different targets. In contrast to the known function of translation stimulation in germ cells, we observed rather a translation repression function of Dazl in ESCs. In sub-cellular localization studies, we detected cytoplasmic localization of both isoforms, as it was reported for Dazl_FL protein (Lee et al., 2006; Kim et al., 2012). Interestingly, many cells showed stress granule (SG)-like as well as diffused cytoplasmic pattern for both Dazl isoforms. These results are in agreement with a recent report, which showed the Dazl-mediated SG formation and translation repression during induced stress conditions (Kim et al., 2012). In support of Dazl role in translation repression, through Co-immunoprecipitation (Co-IP) followed by mass spectrometry analysis, we could identify a heat shock protein 70 (*HSP70*), which is known to regulate the formation of TIA1-mediated translational repression (Gilks et al., 2004), as a putative interaction partner of Dazl. These observations led us to suggest that under stress circumstances, formation of SGs is induced probably by HSP and Dazl complex to repress the translational event and to protect and store mRNAs. However, further studies are necessary to test this hypothesis and to show the mechanisms of translational repression function of Dazl in ESCs.

4.5. Future endeavors and perspectives

In the present study, we elucidated the function of selected GC marker gene (*Dppa3*) and PrM marker gene (*Dazl*) in pluripotent stem cells. To gain deeper insights into the function of GC/PrM marker genes during reprogramming, it might also be necessary to establish iPS cell lines from GC/PrM overexpressed or knockedout fibroblasts to examine the gain or loss of pluripotency in comparison to wild type cells. The observation of GC marker genes expression in human fibroblast cells led us to hypothesize that mesenchymal-to-epithelial transition (MET) will be activated early during reprogramming, as MET and GC marker genes are the responsive targets of the BMP signaling pathway. Therefore, there might be a crosstalk between the networks of MET genes and GC marker genes. Future studies aimed at investigation of MET in human fibroblast reprogramming might help to address the expression of GC marker genes in human fibroblast cells.

We observed an early activation of *Dppa3* during somatic cell reprogramming and our further results implied its role in maintenance of the *Dlk1-Dio3* imprinting. Although, the DNA binding ability of *Dppa3* has been documented, its global genomic targets are still largely unknown. Therefore, the genome-wide screening for direct target based on ChIP followed by next generation sequencing techniques (ChIP-Seq) might help us to better understand Dppa3 role in epigenetic regulation of imprinting and gene expression. Dppa3 is a well-known maternally inherited epigenetic regulator and preferentially protects the maternal allele during zygotic DNA demethylation event. The study of Dppa3 binding to the *Dlk1-Dio3* locus during reprogramming of fibroblasts isolated from F1 hybrid offspring of Mus musculus domesticus (C57BL6/J) and Mus musculus castaneus (CAST/EiJ) crosses might unravel the allele specific binding of Dppa3.

Although Dazl has been regarded as a post-transcriptional regulator, the repressive role of Dazl is still controversial and needs to be further studied. It might be necessary to perform electrophoretic mobility shift assays (EMSA) for studying the binding specificity of Dazl to its target RNA sequences. It was reported that the regulatory role of RNA-binding protein mainly accounts for interaction proteins. Therefore, identification of other interaction proteins of Dazl in ESCs might be useful to dissect the translational repression role of Dazl.

5. References

Aflatoonian, B., Ruban, L., Jones, M., Aflatoonian, R., Fazeli, A., Moore, H. D., 2009. In vitro post-meiotic germ cell development from human embryonic stem cells. Hum Reprod. 24**,** 3150-9.

Albarran, B., To, R., Stayton, P. S., 2005. A TAT-streptavidin fusion protein directs uptake of biotinylated cargo into mammalian cells. Protein Eng Des Sel. 18**,** 147-52.

Aravind, L., Koonin, E. V., 2000. SAP - a putative DNA-binding motif involved in chromosomal organization. Trends in biochemical sciences. 25**,** 112-4.

Bao, S., Leitch, H. G., Gillich, A., Nichols, J., Tang, F., Kim, S., Lee, C., Zwaka, T., Li, X., Surani, M. A., 2012. The germ cell determinant blimp1 is not required for derivation of pluripotent stem cells. Cell Stem Cell. 11**,** 110-7.

Becker-Hapak, M., McAllister, S. S., Dowdy, S. F., 2001. TAT-mediated protein transduction into mammalian cells. Methods. 24**,** 247-56.

Bilic, J., Izpisua Belmonte, J. C., 2012. Concise review: Induced pluripotent stem cells versus embryonic stem cells: close enough or yet too far apart? Stem Cells. 30**,** 33-41.

Bortvin, A., Goodheart, M., Liao, M., Page, D. C., 2004. Dppa3 / Pgc7 / stella is a maternal factor and is not required for germ cell specification in mice. BMC Dev Biol. 4**,** 2.

Brons, I. G., Smithers, L. E., Trotter, M. W., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S. M., Howlett, S. K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R. A., Vallier, L., 2007. Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature. 448**,** 191-5.

Bucay, N., Yebra, M., Cirulli, V., Afrikanova, I., Kaido, T., Hayek, A., Montgomery, A. M., 2009. A novel approach for the derivation of putative primordial germ cells and sertoli cells from human embryonic stem cells. Stem Cells. 27**,** 68-77.

Byrne, A., (2011-06-29), J., 2011. "The definition and etymology of the word pluripotency". eJournal of Cellular Biotechnology, 1;eP2.

Cauffman, G., Van de Velde, H., Liebaers, I., Van Steirteghem, A., 2005. DAZL expression in human oocytes, preimplantation embryos and embryonic stem cells. Mol Hum Reprod. 11**,** 405-11.

Chan, Y. S., Yang, L., Ng, H. H., 2011. Transcriptional regulatory networks in embryonic stem cells. Prog Drug Res. 67**,** 239-52.

Chu, L. F., Surani, M. A., Jaenisch, R., Zwaka, T. P., 2011. Blimp1 expression predicts embryonic stem cell development in vitro. Curr Biol. 21**,** 1759-65.

Da Rocha, S. T., Edwards, C. A., Ito, M., Ogata, T., Ferguson-Smith, A. C., 2008. Genomic imprinting at the mammalian Dlk1-Dio3 domain. Trends Genet. 24**,** 306-16.

Daley, G. Q., 2007. Gametes from embryonic stem cells: a cup half empty or half full? Science. 316**,** 409-10.

De Felici, M., Dolci, S., Pesce, M., 1992. Cellular and molecular aspects of mouse primordial germ cell migration and proliferation in culture. Int J Dev Biol. 36**,** 205-13.

Eguizabal, C., Montserrat, N., Vassena, R., Barragan, M., Garreta, E., Garcia-Quevedo, L., Vidal, F., Giorgetti, A., Veiga, A., Izpisua Belmonte, J. C., 2011. Complete meiosis from human induced pluripotent stem cells. Stem Cells. 29**,** 1186-95.

Eguizabal, C., Shovlin, T. C., Durcova-Hills, G., Surani, A., McLaren, A., 2009. Generation of primordial germ cells from pluripotent stem cells. Differentiation. 78**,** 116-23.

Evans, M. J., Kaufman, M. H., 1981. Establishment in culture of pluripotential cells from mouse embryos. Nature. 292**,** 154-6.

Finch, B. W., Ephrussi, B., 1967. Retention of multiple developmental potentialities by cells of a mouse testicular teratocarcinoma during prolonged culture in vitro and their extinction upon hybridization with cells of permanent lines

Proc Natl Acad Sci U S A. 57**,** 615-21.

Fujiwara, Y., Komiya, T., Kawabata, H., Sato, M., Fujimoto, H., Furusawa, M., Noce, T., 1994. Isolation of a DEAD-family protein gene that encodes a murine homolog of Drosophila vasa and its specific expression in germ cell lineage. Proc Natl Acad Sci U S A. 91**,** 12258-62.

Gardner, R. L., Brook, F. A., 1997. Reflections on the biology of embryonic stem (ES) cells. Int J Dev Biol. 41**,** 235-43.

Geijsen, N., Horoschak, M., Kim, K., Gribnau, J., Eggan, K., Daley, G. Q., 2004. Derivation of embryonic germ cells and male gametes from embryonic stem cells. Nature. 427**,** 148-54.

Gilks, N., Kedersha, N., Ayodele, M., Shen, L., Stoecklin, G., Dember, L. M., Anderson, P., 2004. Stress granule assembly is mediated by prion-like aggregation of TIA-1. Mol Biol Cell. 15**,** 5383-98.

Ginsburg, M., Snow, M. H., McLaren, A., 1990. Primordial germ cells in the mouse embryo during gastrulation. Development. 110**,** 521-8.

Gore, A., Li, Z., Fung, H. L., Young, J. E., Agarwal, S., Antosiewicz-Bourget, J., Canto, I., Giorgetti, A., Israel, M. A., Kiskinis, E., Lee, J. H., Loh, Y. H., Manos, P. D., Montserrat, N., Panopoulos, A. D., Ruiz, S., Wilbert, M. L., Yu, J., Kirkness, E. F., Izpisua Belmonte, J. C., Rossi, D. J., Thomson, J. A., Eggan, K., Daley, G. Q., Goldstein, L. S., Zhang, K., Somatic coding mutations in human induced pluripotent stem cells. Nature. 471**,** 63-7.

Guan, K., Nayernia, K., Maier, L. S., Wagner, S., Dressel, R., Lee, J. H., Nolte, J., Wolf, F., Li, M., Engel, W., Hasenfuss, G., 2006. Pluripotency of spermatogonial stem cells from adult mouse testis. Nature. 440**,** 1199-203.

Hanna, J. H., Saha, K., Jaenisch, R., 2010. Pluripotency and cellular reprogramming: facts, hypotheses, unresolved issues. Cell. 143**,** 508-25.

Haston, K. M., Tung, J. Y., Reijo Pera, R. A., 2009. Dazl functions in maintenance of pluripotency and genetic and epigenetic programs of differentiation in mouse primordial germ cells in vivo and in vitro. PLoS One. 4**,** e5654.

Hay, B., Jan, L. Y., Jan, Y. N., 1990. Localization of vasa, a component of Drosophila polar granules, in maternal-effect mutants that alter embryonic anteroposterior polarity. Development. 109**,** 425-33.

Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S., Saitou, M., 2011. Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. Cell. 146**,** 519-32.

Hübner, K., Fuhrmann, G., Christenson, L. K., Kehler, J., Reinbold, R., De La Fuente, R., Wood, J., Strauss, J. F., 3rd, Boiani, M., Scholer, H. R., 2003. Derivation of oocytes from mouse embryonic stem cells. Science. 300**,** 1251-6.

Hussein, S. M., Batada, N. N., Vuoristo, S., Ching, R. W., Autio, R., Narva, E., Ng, S., Sourour, M., Hamalainen, R., Olsson, C., Lundin, K., Mikkola, M., Trokovic, R., Peitz, M., Brustle, O., Bazett-Jones, D. P., Alitalo, K., Lahesmaa, R., Nagy, A., Otonkoski, T., 2011. Copy number variation and selection during reprogramming to pluripotency. Nature. 471**,** 58-62.

Hustin, J., Collette, J., Franchimont, P., 1987. Immunohistochemical demonstration of placental alkaline phosphatase in various states of testicular development and in germ cell tumours. Int J Androl. 10**,** 29-35.

Imamura, M., Aoi, T., Tokumasu, A., Mise, N., Abe, K., Yamanaka, S., Noce, T., 2010. Induction of primordial germ cells from mouse induced pluripotent stem cells derived from adult hepatocytes. Mol Reprod Dev. 77**,** 802-11.

Kahan, B. W., Ephrussi, B., 1970. Developmental potentialities of clonal in vitro cultures of mouse testicular teratoma. J Natl Cancer Inst. 44**,** 1015-36.

Kanatsu-Shinohara, M., Inoue, K., Lee, J., Yoshimoto, M., Ogonuki, N., Miki, H., Baba, S., Kato, T., Kazuki, Y., Toyokuni, S., Toyoshima, M., Niwa, O., Oshimura, M., Heike, T., Nakahata, T., Ishino, F., Ogura, A., Shinohara, T., 2004. Generation of pluripotent stem cells from neonatal mouse testis. Cell. 119**,** 1001-12.

Kim, B., Cooke, H. J., Rhee, K., 2012. DAZL is essential for stress granule formation implicated in germ cell survival upon heat stress. Development. 139**,** 568-78.

Kleinsmith, L. J., Pierce, G. B., Jr., 1964. Multipotentiality of Single Embryonal Carcinoma Cells. Cancer Res. 24**,** 1544-51.

Lacham-Kaplan, O., Chy, H., Trounson, A., 2006. Testicular cell conditioned medium supports differentiation of embryonic stem cells into ovarian structures containing oocytes. Stem Cells. 24**,** 266-73.

Lange, U. C., Adams, D. J., Lee, C., Barton, S., Schneider, R., Bradley, A., Surani, M. A., 2008. Normal germ line establishment in mice carrying a deletion of the Ifitm/Fragilis gene family cluster. Mol Cell Biol. 28**,** 4688-96.

Lange, U. C., Saitou, M., Western, P. S., Barton, S. C., Surani, M. A., 2003. The fragilis interferon-inducible gene family of transmembrane proteins is associated with germ cell specification in mice. BMC developmental biology. 3**,** 1.

Lawson, K. A., Dunn, N. R., Roelen, B. A., Zeinstra, L. M., Davis, A. M., Wright, C. V., Korving, J. P., Hogan, B. L., 1999. Bmp4 is required for the generation of primordial germ cells in the mouse embryo. Genes Dev. 13**,** 424-36.

Lee, K. H., Lee, S., Kim, B., Chang, S., Kim, S. W., Paick, J. S., Rhee, K., 2006. Dazl can bind to dynein motor complex and may play a role in transport of specific mRNAs. EMBO J. 25**,** 4263-70.

Lerou, P. H., Daley, G. Q., 2005. Therapeutic potential of embryonic stem cells. Blood Rev. 19**,** 321-31.

Li, R., Liang, J., Ni, S., Zhou, T., Qing, X., Li, H., He, W., Chen, J., Li, F., Zhuang, Q., Qin, B., Xu, J., Li, W., Yang, J., Gan, Y., Qin, D., Feng, S., Song, H., Yang, D., Zhang, B., Zeng, L., Lai, L., Esteban, M. A., Pei, D., 2010. A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. Cell Stem Cell. 7**,** 51-63.

Liu, L., Luo, G. Z., Yang, W., Zhao, X., Zheng, Q., Lv, Z., Li, W., Wu, H. J., Wang, L., Wang, X. J., Zhou, Q., 2010. Activation of the imprinted Dlk1-Dio3 region correlates with pluripotency levels of mouse stem cells. The Journal of biological chemistry. 285**,** 19483-90.

Martin, G. R., 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci U S A. 78**,** 7634-8.

Matsui, Y., 2009. The molecular mechanisms regulating germ cell development and potential. Journal of andrology. 31**,** 61-5.

Matsui, Y., Zsebo, K., Hogan, B. L., 1992. Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. Cell. 70**,** 841-7.

McLaren, A., 1994. Introduction. In: Germline development.Ciba Foundation Symposium 182.Chichester: John Wiley& Sons,1994:1.

McLaren, A., 2001. Mammalian germ cells: birth, sex, and immortality. Cell Struct Funct. 26**,** 119-22.

Mise, N., Fuchikami, T., Sugimoto, M., Kobayakawa, S., Ike, F., Ogawa, T., Tada, T., Kanaya, S., Noce, T., Abe, K., 2008. Differences and similarities in the developmental status of embryo-derived stem cells and primordial germ cells revealed by global expression profiling. Genes Cells. 13**,** 863-77.

Nayernia, K., Li, M., Jaroszynski, L., Khusainov, R., Wulf, G., Schwandt, I., Korabiowska, M., Michelmann, H. W., Meinhardt, A., Engel, W., 2004. Stem cell based therapeutical approach of male infertility by teratocarcinoma derived germ cells. Hum Mol Genet. 13**,** 1451-60.

Nayernia, K., Nolte, J., Michelmann, H. W., Lee, J. H., Rathsack, K., Drusenheimer, N., Dev, A., Wulf, G., Ehrmann, I. E., Elliott, D. J., Okpanyi, V., Zechner, U., Haaf, T., Meinhardt, A., Engel, W., 2006. In vitro-differentiated embryonic stem cells give rise to male gametes that can generate offspring mice. Dev Cell. 11**,** 125-32.

Nichols, J., Silva, J., Roode, M., Smith, A., 2009. Suppression of Erk signalling promotes ground state pluripotency in the mouse embryo. Development. 136**,** 3215-22.

Nichols, J., Smith, A., 2009. Naive and primed pluripotent states. Cell Stem Cell. 4**,** 487-92.

Nichols, J., Smith, A., 2011. The origin and identity of embryonic stem cells. Development. 138**,** 3-8.

Nishikawa, S., Goldstein, R. A., Nierras, C. R., 2008. The promise of human induced pluripotent stem cells for research and therapy. Nature reviews. Molecular cell biology. 9**,** 725-9.

Nolte, J., Michelmann, H. W., Wolf, M., Wulf, G., Nayernia, K., Meinhardt, A., Zechner, U., Engel, W., 2010. PSCDGs of mouse multipotent adult germline stem cells can enter and progress through meiosis to form haploid male germ cells in vitro. Differentiation. 80**,** 184-94.

Ohinata, Y., Payer, B., O'Carroll, D., Ancelin, K., Ono, Y., Sano, M., Barton, S. C., Obukhanych, T., Nussenzweig, M., Tarakhovsky, A., Saitou, M., Surani, M. A., 2005. Blimp1 is a critical determinant of the germ cell lineage in mice. Nature. 436**,** 207-13.

Panula, S., Medrano, J. V., Kee, K., Bergstrom, R., Nguyen, H. N., Byers, B., Wilson, K. D., Wu, J. C., Simon, C., Hovatta, O., Reijo Pera, R. A., 2011. Human germ cell differentiation from fetal- and adult-derived induced pluripotent stem cells. Hum Mol Genet. 20**,** 752-62.

Park, T. S., Galic, Z., Conway, A. E., Lindgren, A., van Handel, B. J., Magnusson, M., Richter, L., Teitell, M. A., Mikkola, H. K., Lowry, W. E., Plath, K., Clark, A. T., 2009. Derivation of primordial germ cells from human embryonic and induced pluripotent stem cells is significantly improved by coculture with human fetal gonadal cells. Stem Cells. 27**,** 783-95.

Payer, B., Saitou, M., Barton, S. C., Thresher, R., Dixon, J. P., Zahn, D., Colledge, W. H., Carlton, M. B., Nakano, T., Surani, M. A., 2003. Stella is a maternal effect gene required for normal early development in mice. Curr Biol. 13**,** 2110-7.

Peng, J. X., Xie, J. L., Zhou, L., Hong, Y. H., Gui, J. F., 2009. Evolutionary conservation of Dazl genomic organization and its continuous and dynamic distribution throughout germline development in gynogenetic gibel carp. J Exp Zool B Mol Dev Evol. 312**,** 855-71.

Qing, T., Shi, Y., Qin, H., Ye, X., Wei, W., Liu, H., Ding, M., Deng, H., 2007. Induction of oocyte-like cells from mouse embryonic stem cells by co-culture with ovarian granulosa cells. Differentiation. 75**,** 902-11.

Rayapureddi, J. P., Tomamichel, W. J., Walton, S. T., Payne, R. M., 2010. TAT fusion protein transduction into isolated mitochondria is accelerated by sodium channel inhibitors. Biochemistry. 49**,** 9470-9.

Reijo, R. A., Dorfman, D. M., Slee, R., Renshaw, A. A., Loughlin, K. R., Cooke, H., Page, D. C., 2000. DAZ family proteins exist throughout male germ cell development and transit from nucleus to cytoplasm at meiosis in humans and mice. Biol Reprod. 63**,** 1490-6.

Ren, B., Chee, K. J., Kim, T. H., Maniatis, T., 1999. PRDI-BF1/Blimp-1 repression is mediated by corepressors of the Groucho family of proteins. Genes Dev. 13**,** 125-37.

Resnick, J. L., Bixler, L. S., Cheng, L., Donovan, P. J., 1992. Long-term proliferation of mouse primordial germ cells in culture. Nature. 359**,** 550-1.

Ruggiu, M., Speed, R., Taggart, M., McKay, S. J., Kilanowski, F., Saunders, P., Dorin, J., Cooke, H. J., 1997. The mouse Dazla gene encodes a cytoplasmic protein essential for gametogenesis. Nature. 389**,** 73-7.

Saiti, D., Lacham-Kaplan, O., 2007. Mouse Germ Cell Development in-vivo and in-vitro. Biomark Insights. 2**,** 241-52.

Saitou, M., Barton, S. C., Surani, M. A., 2002. A molecular programme for the specification of germ cell fate in mice. Nature. 418**,** 293-300.

Saitou, M., Yamaji, M., 2009. Germ cell specification in mice: signaling, transcription regulation, and epigenetic consequences. Reproduction. 139**,** 931-42.

Samavarchi-Tehrani, P., Golipour, A., David, L., Sung, H. K., Beyer, T. A., Datti, A., Woltjen, K., Nagy, A., Wrana, J. L., 2010. Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. Cell Stem Cell. 7**,** 64-77.

Sato, M., Kimura, T., Kurokawa, K., Fujita, Y., Abe, K., Masuhara, M., Yasunaga, T., Ryo, A., Yamamoto, M., Nakano, T., 2002. Identification of PGC7, a new gene expressed specifically in preimplantation embryos and germ cells. Mech Dev. 113**,** 91-4.

Schlegel, P. N., 2009. Evaluation of male infertility. Minerva Ginecol. 61**,** 261-83.

Smith, K. P., Luong, M. X., Stein, G. S., 2009. Pluripotency: toward a gold standard for human ES and iPS cells. J Cell Physiol. 220**,** 21-9.

Spahn, L., Barlow, D. P., 2003. An ICE pattern crystallizes. Nature genetics. 35**,** 11-2.

Stadtfeld, M., Apostolou, E., Akutsu, H., Fukuda, A., Follett, P., Natesan, S., Kono, T., Shioda, T., Hochedlinger, K., 2010. Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. Nature. 465**,** 175-81.

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**,** 861-72.

Takahashi, K., Yamanaka, S., 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 126**,** 663-76.

Tam, P. P., Snow, M. H., 1981. Proliferation and migration of primordial germ cells during compensatory growth in mouse embryos. J Embryol Exp Morphol. 64**,** 133-47.

Tanaka, S. S., Toyooka, Y., Akasu, R., Katoh-Fukui, Y., Nakahara, Y., Suzuki, R., Yokoyama, M., Noce, T., 2000. The mouse homolog of Drosophila Vasa is required for the development of male germ cells. Genes Dev. 14**,** 841-53.

Tanaka, S. S., Yamaguchi, Y. L., Tsoi, B., Lickert, H., Tam, P. P., 2005. IFITM/Mil/fragilis family proteins IFITM1 and IFITM3 play distinct roles in mouse primordial germ cell homing and repulsion. Dev Cell. 9**,** 745-56.

Tedesco, M., La Sala, G., Barbagallo, F., De Felici, M., Farini, D., 2009. STRA8 shuttles between nucleus and cytoplasm and displays transcriptional activity. J Biol Chem. 284**,** 35781-93.

Tesar, P. J., Chenoweth, J. G., Brook, F. A., Davies, T. J., Evans, E. P., Mack, D. L., Gardner, R. L., McKay, R. D., 2007. New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature. 448**,** 196-9.

Tilgner, K., Atkinson, S. P., Golebiewska, A., Stojkovic, M., Lako, M., Armstrong, L., 2008. Isolation of primordial germ cells from differentiating human embryonic stem cells. Stem Cells. 26**,** 3075-85.

Toyooka, Y., Tsunekawa, N., Akasu, R., Noce, T., 2003. Embryonic stem cells can form germ cells in vitro. Proceedings of the National Academy of Sciences of the United States of America. 100**,** 11457-62.

Toyooka, Y., Tsunekawa, N., Takahashi, Y., Matsui, Y., Satoh, M., Noce, T., 2000. Expression and intracellular localization of mouse Vasa-homologue protein during germ cell development. Mechanisms of development. 93**,** 139-49.

Tsui, S., Dai, T., Roettger, S., Schempp, W., Salido, E. C., Yen, P. H., 2000. Identification of two novel proteins that interact with germ-cell-specific RNA-binding proteins DAZ and DAZL1. Genomics. 65**,** 266-73.

Vincent, S. D., Dunn, N. R., Sciammas, R., Shapiro-Shalef, M., Davis, M. M., Calame, K., Bikoff, E. K., Robertson, E. J., 2005. The zinc finger transcriptional repressor Blimp1/Prdm1 is dispensable for early axis formation but is required for specification of primordial germ cells in the mouse. Development. 132**,** 1315-25.

West, F. D., Machacek, D. W., Boyd, N. L., Pandiyan, K., Robbins, K. R., Stice, S. L., 2008. Enrichment and differentiation of human germ-like cells mediated by feeder cells and basic fibroblast growth factor signaling. Stem Cells. 26**,** 2768-76.

Williamson, M. P., 1994. The structure and function of proline-rich regions in proteins. Biochem J. 297 (Pt 2)**,** 249-60.

Xu, X., Pantakani, D. V., Luhrig, S., Tan, X., Khromov, T., Nolte, J., Dressel, R., Zechner, U., Engel, W., 2011. Stage-specific germ-cell marker genes are expressed in all mouse pluripotent cell types and emerge early during induced pluripotency. PLoS One. 6**,** e22413.

Yang, S., Bo, J., Hu, H., Guo, X., Tian, R., Sun, C., Zhu, Y., Li, P., Liu, P., Zou, S., Huang, Y., Li, Z., 2012. Derivation of male germ cells from induced pluripotent stem cells in vitro and in reconstituted seminiferous tubules. Cell Prolif. 45**,** 91-100.

Yu, J., Angelin-Duclos, C., Greenwood, J., Liao, J., Calame, K., 2000. Transcriptional repression by blimp-1 (PRDI-BF1) involves recruitment of histone deacetylase. Mol Cell Biol. 20**,** 2592-603.

Yu, Z., Ji, P., Cao, J., Zhu, S., Li, Y., Zheng, L., Chen, X., Feng, L., 2009. Dazl promotes germ cell differentiation from embryonic stem cells. J Mol Cell Biol. 1**,** 93-103.

Zhao, T., Zhang, Z. N., Rong, Z., Xu, Y., 2011. Immunogenicity of induced pluripotent stem cells. Nature. 474**,** 212-5.

Zhu, Y., Hu, H. L., Li, P., Yang, S., Zhang, W., Ding, H., Tian, R. H., Ning, Y., Zhang, L. L., Guo, X. Z., Shi, Z. P., Li, Z., He, Z., 2012. Generation of male germ cells from induced pluripotent stem cells (iPS cells): an in vitro and in vivo study. Asian J Androl. 14**,** 574-9.

Zwaka, T. P., Thomson, J. A., 2005. A germ cell origin of embryonic stem cells? Development. 132, 227-33.

6. Abbreviations

°C	degree Celsius
x g	gravity
µ-	micro
Δ8	Exon 8 deletion
Ac	Acetylation
Al	Allantois
AP	Alkaline Phosphatase
AVE	Anterior visceral endoderm
Blimp1	B- lymphocyte- induced maturation protein-1
BMP	Bone morphogenetic protein
bp	base pair
c-	centi
cDNA	complementary DNA
ChIP	Chromatin Immunoprecipitation
ChIP-Seq	ChIP- Sequencing
Co-IP	Co-Imuunoprecipitation
DAPI	4',6 -diamidino-2-phenylindole
Dazl	Deleted in azoospermia-like
Dazl-FL	Dazl Full length
DE	Distal endoderm
Dio3	type III iodothyronine deiodinase
Dlk1	delta-like 1 homolog
DN	Down regulation
DNA	Deoxyribonucleic acid
Dox	Doxycycline
dpc	days post coitum
dpp	days post partum
Dppa3	Developmental pluripotency-associated protein 3
DVE	Distal visceral endoderm
ECCs	Embryonic carcinoma cells
EGCs	Embryonic germ cells
EGFP	Enhanced green fluorescent protein

EM	Embryonic mesoderm
EMSA	Electrophoretic mobility shift assay
Epi	Epiblast
EpiSCs	Epiblast stem cells
ESCs	Embryonic stem cells
ExE	Extra-embryonic ectoderm
ExM	Extra-embryonic mesoderm
FACS	Fluorescence activated cell sorting
FB	Fibroblast
FCS	Fetal calf serum
FITC	Fluorescein Isothiocyanate
GC	germ cell
GSCs	Germline stem cells
h	hour
H3K27	histone 3 lysine 27
H3K27me3	trimethylation of histone 3 lysine 27
H3K4	histone 3 lysine 4
H3K4me3	trimethylation of histone 3 lysine 4
H3K9ac	acetylation of histone 3 lysine 9
H3K9me3	trimethylation of histone 3 lysine 9
HA	hemagglutinin
hiPSCs	human iPSCs
HRP	Horseradish peroxidase
HSP70	Heat shock protein 70
ICC	Immunocytochemistry
ICM	Inner cell mass
ICRs	Imprinting control regions
ICSI	Intracytoplasmic sperm injection
Ifitm3	Interferon-induced transmembrane protein 3
IG-DMR	Intergenic-differentially methylated region
iPSCs	induced pluripotent stem cells
k-	kilo
kDa	kilo Dalton
KO	Knock-out

l	liter
LIF	Leukemia inhibitory factor
m	meter
M	Molar
m-	milli
maGSCs	multipotent adult germline stem cells
MEFs	Mouse embryonic fibroblasts
MET	Mesenchymal-to-epithelial transition
mGSCs	multipotent germline stem cells
min	minute
miRNA	micro-RNA
mRNA	messenger RNA
Mvh	Mouse vasa homologue
OE	Overexpression
ORF	Open reading frame
OSKM	Oct4, Sox2, Klf4 and c-Myc
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PGC	primordial germ cell
pH	preponderance of hydrogen ions
PR	proline-rich
PrM	pre- meiotic
qRT-PCR	quantitative RealTime-PCR
RA	Retinoic acid
RNA	Ribonucleic acid
RRM	RNA recognition motif
RT-PCR	Reverse transcriptase-PCR
SCs	Stem cells
SDS	Sodium dodecylsulfate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
SG	Stress granule
siRNA	small interfering RNA
SSCs	Spermatogonial stem cells
VE	Visceral endoderm

7. Acknowledgments

First of all, I wish to offer my sincerest gratitude to **Prof. Dr. med. Dr. h. c. Wolfgang Engel**, who has offered me such a invaluable opportunity to study in a prestigious stem cell research group, for his constant excellent guidance, inspiration, encouragement and for providing me with a harmonious atmosphere during my PhD study. It has been my great honor to study under his supervision and his personality and rigorous scholarship will motivate me for whole research life.

I sincerely thank all my thesis committee members, **Prof. Dr. Sigrid Hoyer-Fender, Prof. Dr. Detlef Doenecke, Prof. Dr. Michael Zeisberg, Prof. Dr. Michael Kessel, and Prof. Dr. Ahmed Mansouri** for evaluation of my thesis and serving as my thesis committee.

Next, I would like express my deepest thank to my supervisor **Dr. Krishna Pantakani** for his continuous guidance, understanding, generosity, patience and helping me to develop my background in molecular biology and all techniques. I have been so lucky to have him as my group leader, friend and brother to company me throughout my PhD time. Without him this thesis would not have been completed or written.

My special thanks go to **Dr. Jessica Nolte** for her constitutive advice and discussions throughout my study.

I would like to thank **Lukasz Smorag** and **Joanna Jakubiczka Smorag** for their kindly help in my hardest time during PhD study. I am thankful to my colleagues **Pawel, Tanja, Nadine**, **Sandra L**, **Janine**, **Sandra M**, **Gungee**, **Chiru**, **Belal**, **Mannar** and all my institute members for their advice, discussions, help and providing an amicable working environment.

I would not have been here without them, so I would like to particularly thank **Prof. Dr. Karim Nayernia** and **Christian Mueller**, who recommended and introduced me to study at the Institute of Human Genetics and built a bridge from my Master's to PhD studies.

Last but not least, I want to express my deepest gratitude to my parents for supporting me in starting my scientific study, always put faith in me and helped to pass the hardest time during my study. I am grateful to my beloved wife, **Xiaoying Tan** who always backed me up, shared my ups and downs and gave me a family in Göttingen.

8. Curriculum Vitae

Personal details:

Name:	Xingbo Xu
Date of birth:	22.06.1983
Place of birth:	Tongliao, China
Nationality:	Chinese
Address:	02-03, 22H, Albrecht-Thaer-Weg, Göttingen
E-Mail:	xuhover@hotmail.com

Education:

Sep. 1995 – Jul. 2001:	Secondary School Experimental Middle School, Tongliao, China
Sep.2001 – Jul.2005:	Bachelors in Biology Diploma thesis: "The study of cell apoptosis induced by potassium chromate" College of Life Science, Dalian Nationalities University, China
Sep.2007 – Aug.2008:	Masters in Medical and Molecular Biosciences (MRes) Master thesis: "The derivation of male meiotic germ line cells from mouse embryonic stem cells" Medical School, Newcastle University, Newcastle upon Tyne, U.K
Since Jun. 2009:	PhD thesis: "Stage-specific germ cell marker genes function in establishment and germ cell lineage commitment of pluripotent stem cells" Institute of Human Genetics, Georg-August-University, Göttingen, Germany

9. List of Publications

1. **Xu X**, Pantakani DV, Luhrig S, Tan X, Khromov T, Nolte J, Dressel R, Zechner U, Engel W. (2011) Stage-specific germ-cell marker genes are expressed in all mouse pluripotent cell types and emerge early during induced pluripotency. PLoS One 6(7): e22413.

10. List of Manuscripts in Submission Stage

1. **Xu X**, Nakamura T, Smorag L, Dressel R, Fitzner A, Linke M, Nolte J, Zechner U, Engel W, Pantakani DV. Dppa3 binds to the IG-DMR of the *Dlk1-Dio3* imprinting cluster and prevents its imprinting loss during iPS cell generation. (Manuscript under review).

2. **Xu X**, Tan X, Lin Q, Schmidt B, Engel W, Pantakani DV. (2012) Dazl and its novel splice variant functions in translational repression of target mRNAs in mouse embryonic stem cells. (Manuscript under review).

3. Nyamsuren G[*], **Xu X**[*], Taka O, Doerfel L, Pantakanai DV, Engel W, Adham I. PELO functions in establishment of pluripotency and development of primitive endoderm through the activation of BMP-signaling pathway. (*Equal contribution, Manuscript under review).

4. Tan X, **Xu X**, Zechner U, Nolte J, Engel W, Pantakanai DV. Zfp819, a novel KRAB-zinc finger protein, interacts with KAP1 and functions in genomic integrity maintenance of mouse embryonic stem cells. (Manuscript under review).