_____

Comparative Analysis of Embryonic Stem Cells and Multipotent Adult Germline Stem Cells at the Level of Transcriptome and Proteome

_____

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades

"Doctor rerum naturalium"

an der Georg-August-Universität Göttingen

vorgelegt von

Sandra Meyer

aus Minden

Göttingen, 2010

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

Korreferent: Prof. Dr. Sigrid Hoyer-Fender

Tag der mündlichen Prüfung:

Table of Contents

1. Zusammenfassung

Bei der Pluripotenz verschiedener Stammzelltypen handelt es sich um einen komplexen biologischen Zustand, der der Zelle erlaubt, sich unbegrenzt zu teilen und zu proliferieren (Selbst-Erneuerung), der die Zelle aber auch auf eine Differenzierung in Gewebe aller drei Keimblätter und zu Keimzellen vorbereitet. Pluripotenz wird durch das Vorhandensein von speziellen transkriptionellen regulatorischen Netzwerken aufrechterhalten, wobei verschiedene Transkriptionsfaktoren an DNA binden oder auf Protein-Protein-Ebene miteinander interagieren. Diese Transkriptionsfaktoren sind daher wichtig für die Aufrechterhaltung der Pluripotenz. Allerdings wurden die Komponenten dieser Transkriptionsfaktornetzwerke bis jetzt noch nicht vollständig identifiziert. Die Analyse von Mechanismen, die an der Regulation von Pluripotenz beteiligt sind, wurde hauptsächlich in embryonalen Stammzellen (ESCs) durchgeführt. Vor kurzem konnte ein neuer Typus von pluripotenten Stammzellen identifiziert werden, nämlich multipotente adulte Keimbahnstammzellen (maGSCs). In der vorliegenden Arbeit zielten wir darauf ab, Ähnlichkeiten und Unterschiede im Genexpressionsmuster von ESCs und maGSCs herauszuarbeiten. Dabei sollte ein besonderer Fokus auf die Gene gelegt werden, die bekanntermaßen an der Regulation von Pluripotenz beteiligt sind. Ein weiteres Ziel war die Identifizierung putativer neuer Faktoren, die Pluripotenz regulieren.

Im ersten Teil dieser Arbeit führten wir deswegen eine Microarray-Analyse durch, um ESCs, maGSCs und deren differenzierte Abkömmlinge auf RNA-Ebene zu vergleichen. Es konnte gezeigt werden, dass die beiden undifferenzierten Zelltypen weder anhand der Expression bekannter Pluripotenzgene noch anhand ihres globalen Genexpressionsmusters unterschieden werden können. Beim Vergleich differenzierter und undifferenzierter Zellen konnte die erwartete Veränderung im Genexpressionsmuster dargestellt werden. Dahingegen wiesen die zwei Zelltypen nach ihrer Differenzierung wieder eine sehr ähnliche Genexpression auf.

Im zweiten und dritten Teil dieser Doktorarbeit wird der Vergleich von ESCs und maGSCs auf Proteinebene beschrieben. Hierdurch war es möglich, die Ähnlichkeiten zwischen den Zelltypen im undifferenzierten Zustand zu bestätigen. Allerdings konnten sowohl qualitative als auch quantitative Unterschiede in der Proteinausstattung der Zellen nach der Differenzierung beider Zelltypen gefunden werden. Zusätzlich konnten wir zeigen, dass die post-translationale Modifikation ‚Hypusinierung von Eif5a' einen Einfluss auf die Proliferation von ESCs und maGSCs hat, wohingegen sie sich nicht auf die Pluripotenz der Zellen auswirkt.

Der vierte Teil dieser Arbeit beschreibt die funktionelle Charakterisierung des putativen Pluripotenzgens Stra8. Es konnte gezeigt werden, dass Stra8 alle Charakteristika eines Proteins erfüllt, das an der transkriptionellen Regulation der Pluripotenz beteiligt ist. Diese

Charakteristika sind das Vorhandensein einer Transkriptionsfaktordomäne, die Expression in undifferenzierten pluripotenten Stammzellen und in Embryonen im Präimplantationsstadium, sowie die Abnahme der Expression während der Differenzierung der pluripotenten Stammzellen. Wir konnten zeigen, dass eine Veränderung im Proteingehalt von Stra8, und zwar siRNA-vermittelte Herunterregulation und stabile Überexpression, eine Veränderung des Expressionslevels sowohl verschiedener bekannter Pluripotenzgene als auch von Markergenen für die Differenzierung in die drei Keimblätter bewirkt.

Zusätzlich wird die Identifizierung und Charakterisierung weiterer putativer Pluripotenzgene dargestellt. Hierfür wurden die Ergebnisse der Transkriptomanalyse, die im ersten Teil der Arbeit beschrieben werden, neu ausgewertet. Damit sollten Transkriptionsfaktoren identifiziert werden, deren Expression während der Differenzierung von ESCs und maGSCs herunterreguliert wird. Diese Kandidatengene wurden anhand ihres Expressionsmusters in pluripotenten Zellen und adulten Organen weiter analysiert. Dabei konnte gezeigt werden, dass ein Gen, Kbtbd8, ein Expressionsprofil zeigt, das charakteristisch ist für Gene, die an der Regulation von Pluripotenz beteiligt sind.

Zusammenfassend kann gesagt werden, dass die hier präsentierten Ergebnisse Einblicke in die Charakteristika von maGSCs geben, wodurch deren Ähnlichkeiten zu ESCs gezeigt werden können. Außerdem wurden zwei Faktoren identifiziert, die möglicherweise an der Regulation von Pluripotenz beteiligt sind. Die Rolle dieser Faktoren in der Aufrechterhaltung der Pluripotenz *in vitro* und *in vivo* muss allerdings durch funktionelle Analysen noch genauer untersucht werden.

1. Summary

Pluripotency of different stem cell types is a complex biological state which allows the cells to continuously proliferate (self-renewal) but also primes them for differentiation into all germ layers and germ cells. Regulation of pluripotency involves the presence of transcriptional regulatory networks, in which specific transcription factors interact via DNA-binding and protein-protein interaction, thereby being functionally important for maintenance of pluripotency. These transcription factor networks are, however, until now incompletely defined. In the past, the analysis of mechanisms that control pluripotency was mostly performed in embryonic stem cells (ESCs). Recently, a new type of pluripotent stem cells, namely multipotent adult germline stem cells (maGSCs), was identified. In the present study, we aimed at the identification of similarities and differences in gene expression patterns of ESCs and maGSCs with a special focus on genes known to be involved in the regulation of pluripotency. Another goal was the identification of putative new pluripotency-regulating factors.

In the first part of the study, we therefore performed whole genome microarray analysis to compare undifferentiated and differentiated ESCs and maGSCs with each other at RNA-level. It could be shown that the undifferentiated cell lines are not only indistinguishable from each other based on their expression of known pluripotency-regulating factors but also based on their global gene expression pattern. We could find that, as expected, the cell types change their gene expression profile during differentiation. However, after differentiation both cell types again show a very similar gene expression pattern.

In the second and third part of the thesis, the comparison of ESCs and maGSCs at protein level is described. Herewith, it was possible to confirm the similarities between both cell types in their undifferentiated state. However, differences in protein abundance could be found after differentiation of the cell lines. Additionally, we could show that the post-translational modification 'hypusination of Eif5a' has an effect on the proliferation potential of ESCs and maGSCs, whereas it did not influence the pluripotency of the cells.

The fourth part of the study describes the functional characterization of the putative pluripotency-regulating factor Stra8. It was found that Stra8 fulfills all the criteria for a protein involved in transcriptional regulation of pluripotency, namely the presence of a transcription factor domain, the expression in undifferentiated pluripotent stem cells and preimplantation stage embryos and decrease of expression during differentiation of pluripotent cells. We could show that a change in protein level of Stra8, that are siRNA-mediated knockdown and stable overexpression, results in a change of expression level of known pluripotency regulators as well as marker genes for differentiation into the three germ layers.

In addition, the identification and characterization of further putative pluripotency-regulating factors are shown. Therefore, a reanalysis of the results of the transcriptional profiling experiments described in the first part of the thesis was performed to identify transcription factors whose expression is downregulated during differentiation of ESCs and maGSCs. These candidate genes were further analyzed according to their expression pattern in pluripotent cell lines and adult organs. Herewith, we could show that one gene, Kbtbd8, possesses an expression profile characteristic for genes involved in the regulation of pluripotency.

In conclusion, the data presented here provides insights into the characteristics of maGSCs, thereby showing their ESC-like nature. Additionally, two factors that might be involved in transcriptional regulation of pluripotency were identified. However, further experiments are necessary to prove their function in regulation of pluripotency *in vitro* and *in vivo*.

2. Introduction

2.1 Stem cells

Stem cells possess the capacity to proliferate indefinitely (self-renew) and to differentiate into different tissues or cell types. These cells can be found in most tissues, and they are responsible for tissue maintenance and repair.

Stem cell types can either be distinguished by their origin or by their potency. Concerning their differentiation potential, they can be categorized as totipotent, pluripotent, multipotent or unipotent stem cells. Totipotent cells have the ability to differentiate into cells of all three germ layers (mesoderm, endoderm, ectoderm) and to trophoblast cells, thereby being able to generate a viable organism after transfer to an uterus. Only the fertilized oocyte (zygote) and blastomeres of cleavage stage embryos till eight-cell-stage are totipotent. The individual blastomeres lose their totipotency when embryonic development progresses. This is due to the first irreversible differentiation into trophoblast and inner cell mass (ICM) at blastocyst stage. The cells of the inner cell mass can give rise to pluripotent cells. These cells are not able to form a complete viable organism, but they have the ability to differentiate into cells of all three germ layers *in vivo* and *in vitro* (reviewed in: Sell, 2004). Multipotent stem cells can differentiate into a small number of cell types. This differentiation is limited to tissues belonging to only one germ layer. Unipotent stem cells are even more restricted in their differentiation potential, which means they can only differentiate into one specialized cell type.

According to their origin, pluripotent stem cells can be further classified as embryonic carcinoma cells (ECCs) (Kleinsmith and Pierce, 1964; Finch and Ephrussi, 1967; Kahan and Ephrussi, 1970), embryonic stem cells (ESCs) (Evans and Kaufmann, 1981; Martin, 1981), embryonic germ cells (EGCs) (Matsui et al., 1992; Resnick et al., 1992), multipotent germline stem cells, either derived from neonatal (mGSCs) (Kanatsu-Shinohara et al., 2004) or adult mouse testis (maGSCs) (Guan et al., 2006), epiblast stem cells (EpiSCs) (Brons et al., 2007; Tesar et al., 2007) and induced pluripotent stem cells (iPSCs) (Takahashi and Yamanaka, 2006). An overview of mouse pluripotent stem cells and their respective origin is given in figure 1.
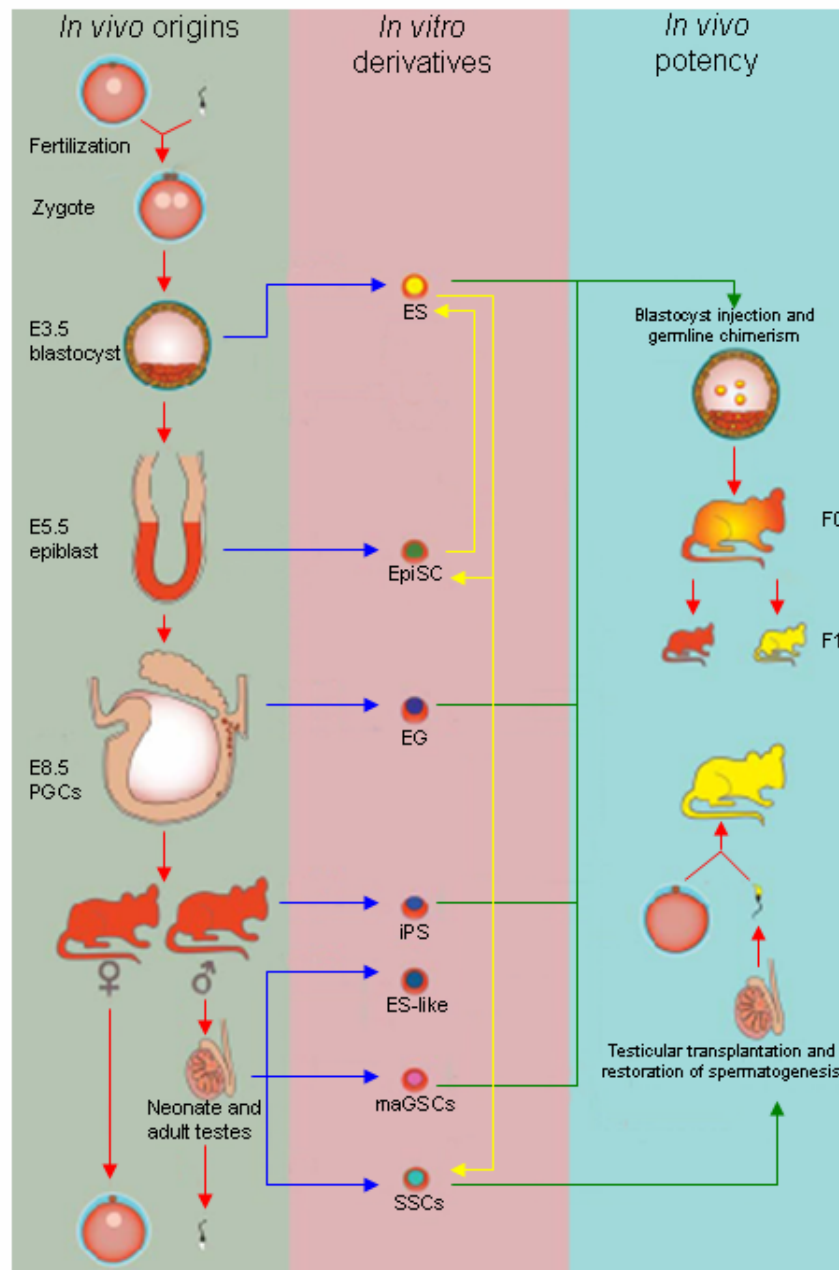
Figure 1: Origins of mouse pluripotent stem cells. In the left column the origins of pluripotent stem cells and in the middle column the stem cell types, which can be derived from them, are shown. The right column displays the functional properties of the respective cell lines. Blue arrows indicate the derivation of cell lines from their respective tissues, yellow arrows show possible *in vitro*-differentiation of one cell type into another and green arrows display *in vivo*-properties of cultured cells. (Figure taken from: Kujik et al., 2010, modified.)

2.1.1 Embryonic stem cells (ESCs)

Embryonic stem cells are derived from the inner cell mass of blastocysts and are able to proliferate indefinitely under defined culture conditions *in vitro*. They are able to differentiate into derivatives of all three germ layers *in vivo* as well as *in vitro*. The first ESC-lines were derived from mouse blastocysts (Evans and Kaufmann, 1981; Martin,

1981), whereas later it even became possible to obtain these cells from individual blastomeres from cleavage stage mouse embryos (Chung et al., 2006; Wakayama et al., 2007). ESC-lines could also be generated from blastocysts of two non-human primates, the rhesus monkey and the common marmoset (Thomson et al., 1995; Thomson et al., 1996), as well as from human (Thomson et al., 1998). Comparable to mouse ESC-lines, it was also possible to derive human ESCs from morula stage embryos (Strelchenko et al., 2004), single blastomeres (Klimanskaya et al., 2006) and parthenogenetic embryos (Lin et al., 2007; Mai et al., 2007; Revazova et al., 2007).

To prevent spontaneous differentiation of mouse ESCs *in vitro*, they are cultured on a "feeder layer" of mitotically inactive murine embryonic fibroblasts (MEFs) (Evans and Kaufmann, 1981; Martin, 1981). These MEFs secrete a cocktail of growth factors, which includes the cytokine LIF (leukemia inhibitory factor). LIF causes the ESCs to proliferate and at the same time inhibits their spontaneous differentiation (Martin and Evans, 1975). However, it is sufficient to only supplement the culture medium with LIF to keep mouse ESCs in an undifferentiated state (Smith et al., 1988; Williams et al., 1988). If mouse ESCs are cultured in the absence of a "feeder layer" and LIF, they spontaneously differentiate into derivatives of all three germ layers and germ cells (Doetschman et al., 1985; Suda et al., 1987; Geijsen et al., 2004; Nayernia et al., 2006). Addition of retinoic acid (RA) to the culture medium induces and accelerates differentiation into several lineages (e.g. Doetschmann et al., 1985; Slager et al., 1993; Dinsmore et al., 1996; Dani et al., 1997; Drab et al., 1997).

2.1.2 Pluripotent stem cells in the testis

The mammalian testis harbors a population of unipotent stem cells, the spermatogonial stem cells (SSCs). These cells are responsible for maintaining spermatogenesis throughout the life of a male individual, thereby continuously self-renewing and differentiating into spermatozoa (Brinster and Avarbock, 1994).

SSCs from mouse neonatal testis can be cultured *in vitro* in the presence of glial cell line derived neurotrophic factor (GDNF), leukemia inhibitory factor (LIF), epidermal growth factor (EGF) and basic fibroblast growth factor (bFGF) for an extended time period without losing their ability to colonize seminiferous tubules of the testis and to differentiate into functional sperm after transplantation into testes of infertile mice (Kanatsu-Shinohara et al., 2003). However, under these culture conditions cells can occasionally be found that resemble the morphology of ESCs. These colonies can be maintained when subsequently cultured under standard ESC culture conditions. These multipotent germline stem cells (mGSCs) express marker genes characteristic for ESCs and are able to differentiate into derivatives of the three germ layers *in vitro* and *in vivo*, as determined by teratoma

formation after subcutaneous injection into nude mice and into testes of germ cell-depleted mice as well as chimera formation after blastocyst injection (Kanatsu-Shinohara et al., 2004).

When applying standard ESC culture conditions, it is also possible to generate pluripotent cells from adult mouse testis (Guan et al., 2006). These multipotent adult germline stem cells (maGSCs) show ESC-like morphology, express marker genes characteristic for ESCs and can differentiate *in vitro* spontaneously into different cell types of all three germ layers. After subcutaneous injection into immunodeficient mice maGSCs form teratomas and after injection into blastocysts they contribute to chimera formation (Guan et al., 2006). The derivation of pluripotent stem cells from adult mouse testis was confirmed by several other groups (Seandel et al., 2007; Kanatsu-Shinohara et al., 2008; Izadyar et al., 2008; Ko et al, 2009).

Recently, the derivation of pluripotent stem cells from spermatogonial stem cells from adult human testis was also reported (Conrad et al., 2008; Golestaneh et al., 2009; Kossack et al., 2009; Mizrak et al., 2009). These cells show similarities to human embryonic stem cells, are able to form teratomas after transplantation into immunodeficient mice and differentiate into derivatives of the three germ layers *in vitro*. However, until now it was not possible to generate clonal, yet proliferating cell lines from adult human testis.

2.2 Analyses of pluripotent stem cells

The properties of ESCs have been extensively characterized, mainly at the level of transcriptome and proteome. Mostly, these studies aimed at the identification of genes specifically expressed in pluripotent cells, thereby attempting to identify regulatory pathways that define pluripotency.

When comparing ESCs with multipotent hematopoietic and neural stem cells at transcriptome level, it could be found that all these stem cell types share a similar gene expression pattern, maybe representing a common set of regulatory pathways (Ivanova et al., 2002). However, only a small number of genes belonging to this gene set are not at all expressed in terminally differentiated cell lines. This reveals that the genes which are enriched in stem cell populations are not exclusively expressed by them (Ramalho-Santos et al., 2002). Besides these similarities between multipotent and pluripotent stem cells, each type of stem cells can be distinguished from the others by a specifically expressed set of genes (Ramalho-Santos et al., 2002).

Furthermore, ESCs have been studied by proteome analysis and reference maps of expressed proteins could be generated (Elliott et al., 2004; Nagano et al., 2005). In 2006, it was even possible to identify 743 proteins exclusively expressed in undifferentiated

mouse ESCs in a comparison between undifferentiated ESCs and ESCs differentiated for 12 days in the absence of a "feeder layer" and LIF. This ESC-specific gene set included known ESC-specific transcription factors but also a high number of proteins not associated with pluripotency (van Hoof et al., 2006).

Despite these studies, the mechanisms that control pluripotency remain largely unknown. To gain more insight into the general characteristics of pluripotent cells, it is necessary to analyze the similarities and differences of several pluripotent cell lines. This approach was taken by different groups in transcriptomics and proteomics studies (e.g. Takahashi and Yamanaka, 2006; Sharova et al., 2007; Buhr et al., 2008). Although different pluripotent stem cells are not morphologically distinguishable from each other, further analyses revealed that the cell lines can be discerned based on their gene expression profiles. Using global gene expression profiling at RNA-level, Sharova et al. (2007) could show that EGCs and ESCs show a very similar global gene expression pattern. This pattern distinguishes them from mulitpotent or unipotent stem cells. The global differences between different pluripotent stem cell types were even smaller than the differences between stem cells derived from different mouse backgrounds. However, when analyzing different ESC- and EGC-lines, it was possible to identify subsets of genes specific for each cell type. In case of ESCs and EGCs 20 and 84 "signature genes", respectively, which are consistently higher expressed in one cell type, could be detected (Sharova et al., 2007). Examination of the nuclear proteome of ESCs and EGCs revealed a similar result. Although commonly expressed proteins could be found in both cell types, around 30% of proteins identified in EGCs are specifically expressed only in this cell type (Buhr et al., 2008).

In addition to extensive comparisons between ESCs and EGCs, the analyses of pluripotent stem cells were recently extended to iPSCs. These cells were routinely compared to ESCs at transcriptome level by the groups who reported the successful reprogramming of differentiated cells into a pluripotent state (e.g. Takahashi and Yamanaka, 2006; Takahashi et al., 2007; Wernig et al., 2007; Lowry et al., 2008). iPSCs have been referred to as "nearly similar" to ESCs of mouse and human, but a small percentage of genes is differentially expressed (e.g. Takahashi et al., 2007; Yu et al., 2007; Lowry et al., 2008; Zhou et al., 2009). These differences are found in all analyzed iPSC-lines, suggesting that iPSCs possess a gene expression signature that can distinguish them from ESCs but is common for iPSC-lines generated from different species, from different somatic origins and following different protocols for reprogramming (Chin et al., 2009). The differentially expressed genes can be classified according to their biological functions, showing that genes which are related to processes associated with differentiation are higher abundant in iPSCs (Chin et al., 2009). This suggests that iPSCs

have not completely silenced the expression pattern of their somatic cell origin. The differences are strongest in early passage iPSCs, whereas iPSCs of later passages cluster more closely with ESCs (Chin et al., 2009).

These alterations in gene expression are not due to differences in global histone modification. The global analysis of trimethylation of histones H3K27 and H3K4 reveals nearly no differences between ESCs and iPSCs, i.e. only 1% of genes are differentially methylated (Maherali et al., 2007; Chin et al., 2009). In contrast to this, global microRNA (miRNA) expression profiling showed consistently a few miRNAs that are differentially expressed between iPSCs and ESCs (Chin et al., 2009; Wilson et al., 2009). These miRNAs include a group of miRNAs previously classified as ESC-specific (Card et al., 2008). The impact of the reported differences has neither been elucidated in case of EGCs nor in case of iPSCs.

While pluripotent stem cell lines like EGCs and iPSCs have been studied in great detail, the characterization of maGSCs is limited to the confirmation of their pluripotent state by teratoma and chimera formation assays as well as basic analyses of gene expression of pluripotency-related genes (Guan et al., 2006; Nolte, 2008). Recently, it was shown that miRNAs which were thought to be specific for ESCs (Houbaviy et al., 2003) are also expressed in maGSCs (Zovoilis et al., 2008). However, the expression level varies between the two cell types, showing a lower amount of miRNAs in maGSCs (Zovoilis et al., 2008). A slight difference could also be observed by comparative analyses of global and gene-specific DNA-methylation patterns. Whereas the global DNA-methylation as well as the DNA-methylation of promoters of pluripotency-regulating genes are very similar in ESCs and maGSCs, the two cell types can be discriminated by their DNA-methylation of imprinted gene loci (Zechner et al., 2009). Although these results give first hints about the ESC-like nature of maGSCs, it still has to be elucidated if they are similar or even identical to ESCs at the level of gene expression.

2.3 Regulation of pluripotency

The mechanisms by which pluripotent cell lines retain their pluripotency during culture are until now rather unknown. However, it is known that a complex network of transcription factors and extrinsic signals that activate signal transduction cascades are involved in the regulation of pluripotency and self-renewal of stem cells.

Inside this intrinsic pluripotency network are the three transcription factors Oct4, Nanog and Sox2 that are indispensable for maintaining pluripotency in ESCs and early embryos (Nichols et al., 1998; Niwa et al., 2000; Avilion et al., 2003; Chambers et al., 2003; Mitsui et al., 2003). Inactivation of one of these genes leads to a loss of pluripotency and to differentiation of the cells. The three transcription factors act by binding to the same

promoter regions of genes, thereby regulating these genes' expression (Loh et al., 2006; Chen et al., 2008; Sharov et al., 2008). Besides these three factors, the pluripotency network consists of several other proteins which were recently identified. For example, it was shown that amongst others the transcription factors Sall4 (Zhang et al., 2006), Zscan10 (Wang et al., 2007a; Wang et al., 2007b) and Zic3 (Lim et al., 2007) play an important role in the core pluripotency network.

One of the extrinsic regulators is the cytokine LIF that binds to heterodimers of LIF-receptor and Gp130 (David et al., 1993). This interaction results in the activation of Stat3- (Niwa et al., 1998), PI(3)K-Akt- (Kobayashi et al., 1999) and MAPK-signaling (Paling et al., 2004). These pathways connect the LIF-signal to the core network of intrinsic pluripotency-regulating factors: Stat-signaling activates Klf4 and Sox2, PI(3)K-Akt-signaling activates Tbx3 and Nanog (Niwa et al., 2009). In the absence of fetal calf serum (FCS) supplementation of the growth medium with LIF is not sufficient to prevent the cells from differentiation, but the induction of cells with bone morphogenetic protein 4 (BMP4) is required (Ying et al., 2003). BMP4 inhibits differentiation by activation of Smad transcription factors which subsequently induce gene expression of Id gene family members (Ying et al., 2003). Although several genes which are involved in the maintenance of pluripotency have been identified in the past years, the complete pluripotency-regulating network remains largely unknown and needs further studies to identify more of its components.

2.4 Objectives of this study

The aim of this study was a further detailed molecular characterization of maGSC-lines isolated by our group (Guan et al., 2006). These cells are known to express the same set of pluripotency-regulating genes, contribute to chimeras at the same level as ESCs (Nolte, 2008) and express the same set of microRNAs (Zovoilis et all., 2008). However, it is known that differences between pluripotent stem cell lines, e.g. between ESCs and EGCs or between ESCs and iPSCs, at the level of transcriptome and proteome exist (e.g. Sharova et al., 2007; Buhr et al., 2008; Chin et al., 2009). Therefore, the maGSCs should be compared to ESC-lines derived from the same genetic backgrounds.

For that purpose, one wildtype and one transgenic maGSC- and ESC-line should be examined at transcriptome level using whole-genome microarray analyses. The genes found to be differentially expressed should be subsequently classified according to the molecular functions they perform and biological processes they are involved in. The results should be compared with previously published data comparing undifferentiated EGC- and ESC-lines (Sharova et al., 2007). With this approach it should be proven whether a transcriptome common for pluripotent cell lines exists.

The same cell lines should also be compared at proteome level. Therefore, 2D-gelectrophoresis should be performed to generate 2D-reference maps for maGSCs and ESCs and the identified proteins should be classified according to their known functions. The ESC-like properties of maGSCs should be further validated using 2D-difference gel electrophoresis (2D-DIGE).

After spontaneous differentiation of maGSCs and ESCs from wildtype mouse background, transcriptome and proteome analyses should be repeated to monitor the gene expression changes during RA-induced differentiation. Additionally, it should be investigated whether the cell types differentiate spontaneously in a similar way or if the origin of maGSCs influences their gene expression after differentiation. The genes which are downregulated after differentiation of the cells should be further characterized using bioinformatics. The classification should be based on the cellular localization and the postulated function of the encoded protein products.

In the last part of this study, the role of the gene Stra8 in pluripotent cells should be further specified. Stra8 is known to play a role in initiation of meiosis in mammals (Anderson et al., 2008) and was recently found to be expressed in ESC-lines where it shuttles between cytoplasm and nucleus (Tedesco et al., 2009). Expression analyses of Stra8 in different pluripotent and differentiated cell lines as well as in embryonic stages should be performed. The role of Stra8 in transcriptional regulation concerning the pluripotency network should be evaluated by transient downregulation using small interfering RNA (siRNA) and stable overexpression.

3. Results

The present study deals with the detailed molecular characterization of multipotent adult germline stem cells (maGSCs) at the level of transcriptome and proteome and the identification of genes involved in the regulation of pluripotency not only in maGSCs but also in embryonic stem cells (ESCs). The results of this thesis have been summarized in the following four manuscripts, of which two have been published and another two are in submission stage:

3.1 Pluripotent embryonic stem cells and multipotent adult germline stem cells reveal similar transcriptomes including pluripotency-related genes

3.2 Multipotent adult germline stem cells and embryonic stem cells: comparative proteomic approach

3.3 Multipotent adult germline stem cells and embryonic stem cells functional proteomics revealed an important role of translation initiation factor 5A (Eif5a) in stem cell differentiation

3.4 Germ cell specific gene Stra8 has an impact on the pluripotency network

Each chapter within the results starts with a brief description of the aim and a summary of the conclusions of the particular manuscript in context of the complete thesis, the status of the manuscript, the authors and their contributions to the work.

3.1 Pluripotent embryonic stem cells and multipotent adult germline stem cells reveal similar transcriptomes including pluripotency-related genes

In the first part of the thesis, the comparison of ESCs and maGSCs at transcriptome level is described. The analyses revealed that maGSCs are very similar to ESCs based on their global gene expression pattern, and the undifferentiated cell lines show a 97-99% identity dependent on the analyzed mouse background. No difference in gene expression could be found in genes involved in the regulation of pluripotency. Also after spontaneous differentiation of both cell lines, their transcriptomes are nearly identical, i.e. show a similarity of 95%, suggesting that the cells differentiate spontaneously in the same direction. Most interestingly, when comparing gene expression of maGSCs with previously published gene expression data for another pluripotent stem cell line, namely embryonic germ cells (EGCs), we found a high similarity between both cell types. These results give a hint that a common transcriptome for pluripotent cell lines exists.

Sandra Meyer, Jessica Nolte, Lennart Opitz, Gabriela Salinas-Riester and Wolfgang Engel
Status: Published in *Molecular Human Reproduction*, Volume 16 (2010), pp. 846-855 (Impact Factor: 3.005)

Author contributions to the work:

1. **Sandra Meyer**: conception and design of experiments, performance of cell culture, preparation of cells for transcriptional profiling, characterization of cells using qRT-PCR and FACS, performance of transcriptional profiling (partial), data analysis, confirmation of data from transcriptional profiling using qRT-PCR, preparation of manuscript.

2. Jessica Nolte: conception and design of experiments, generation of cell lines, characterization of cells using Western Blot, involved in preparation of manuscript.

3. Lennart Opitz: bioinformatical analysis of raw data from microarray.

4. Gabriela Salinas-Riester: conception and design of microarray experiments.

5. Wolfgang Engel: conception and design of experiments, preparation of manuscript.

MHR Basic science of reproductive medicine   **ORIGINAL RESEARCH**

Pluripotent embryonic stem cells and multipotent adult germline stem cells reveal similar transcriptomes including pluripotency-related genes

S. Meyer[1], J. Nolte[1], L. Opitz[2], G. Salinas-Riester[2], and W. Engel[1,*]

[1]Institute of Human Genetics, Georg-August-University Göttingen, Heinrich-Düker-Weg 12, D-37073 Göttingen, Germany [2]DNA Microarray Facility, Georg-August-University Göttingen, 37073 Göttingen, Germany

*Correspondence address. Tel: +49-551-397589; Fax +49-551-399303; E-mail: wengel@gwdg.de

ABSTRACT: DNA microarray analysis was performed with mouse multipotent adult germline stem cells (maGSCs) and embryonic stem cells (ESCs) from different genetic backgrounds cultured under standard ESC-culture conditions and under differentiation-promoting conditions by the withdrawal of the leukemia inhibitory factor (LIF) and treatment with retinoic acid (RA). The analyzed undifferentiated cell lines are very similar based on their global gene expression pattern and show 97–99% identity dependent on the analyzed background. Only 621 genes are differentially expressed in cells derived from mouse 129SV-background and 72 genes show differences in expression in cells generated from transgenic Stra8-EGFP/Rosa26-LacZ-background. Both maGSCs and ESCs express the same genes involved in the regulation of pluripotency and even show no differences in the expression level of these genes. When comparing maGSCs with previously published signature genes of other pluripotent cell lines, we found that maGSCs shared a very similar gene expression pattern with embryonic germ cells (EGCs). Also after differentiation of maGSCs and ESCs the transcriptomes of the cell lines are nearly identical which suggests that both cell types differentiate spontaneously in a very similar way. This is the first study, at transcriptome level, to compare ESCs and a pluripotent cell line derived from an adult organism (maGSCs).

Key words: embryonic stem cells / genome-wide transcriptional profiling / multipotent adult germline stem cells / pluripotency / pluripotency marker genes

Introduction

Embryonic stem cells (ESCs) derived from the inner cell mass (ICM) of blastocysts are pluripotent cells (Evans and Kaufman, 1981; Martin, 1981). They possess the capacity to self-renew, that means they can produce daughter cells with an identical phenotype. They also have the ability to generate all types of differentiated cells (Suda et al., 1987), including germ cells (Geijsen et al., 2004; Nayernia et al., 2006).

Spermatogonial stem cells (SSCs) are located in the seminiferous tubules of the testis and they are responsible for maintaining spermatogenesis throughout life in the male. Pluripotent stem cell lines were generated from SSCs derived from neonatal mouse testis (Kanatsu-Shinohara et al., 2004) as well as from adult mouse testis (Guan et al., 2006) suggesting that the germline lineage retains the ability to generate pluripotent cells. SSCs isolated from adult mouse testis acquire ESC properties, give rise to pluripotent stem cell lines in culture (multipotent adult germline stem cells, maGSCs) and are able to differentiate into derivatives of the three germ layers *in vitro*

(Guan et al., 2006). The cultured cells can be expanded and remain karyotypically stable for more than 100 passages (own unpublished results). Injected into blastocysts, maGSCs give rise to germline-transmitting chimaeras (Guan et al., 2006). After subcutaneous injection of maGSCs into immunodeficient mice or into the testes of adult mice, they generate teratomas (Guan et al., 2006; own unpublished results). In contrast, the freshly isolated SSCs are able to regenerate spermatogenesis after transplantation into germ cell-depleted adult mouse testis (Guan et al., 2006).

Both maGSCs and ESCs express the same set of pluripotency-related genes, such as Oct4, Nanog and Sox2, and the expression of these genes ceases during differentiation of the cells. Recently, it was found that maGSCs and ESCs also express the same set of microRNAs (miRNAs) (Zovoilis et al., 2008), which were previously thought to be ESC-specific (Houbaviy et al., 2003), and share the same profiles concerning their proteomes (Dihazi et al., 2009). These findings provide a strong indication that SSCs can be reprogrammed to pluripotent ESC-like cells under

in vitro-culture conditions, which was recently confirmed by several other groups (Seandel *et al.*, 2007; Izadyar *et al.*, 2008; Kanatsu-Shinohara *et al.*, 2008; Ko *et al.*, 2009).

Mouse ESCs have been studied by global gene expression profiling: Several groups compared hematopoietic stem cells, neural stem cells and ESCs and identified genes commonly expressed in all types of the analyzed cell lines. This reveals the presence of conserved regulatory pathways (Ivanova et al., 2002), but every stem cell type can be distinguished by the highly enriched expression of several genes which are not expressed in other stem cell types (Ramalho-Santos *et al.*, 2002). With the comparison of ESCs and terminally differentiated mouse embryonic fibroblasts (MEF) Tanaka et al. (2002) showed the presence of a subset of genes specifically expressed in pluripotent ESCs (Tanaka et al., 2002). By analyzing ESCs with SAGE analysis Anisimov et al. (2002) identified several transcripts with an expression profile unique to ES RI cells (Anisimov et al., 2002). In 2006, it was shown that ESC-cultures consist of different subpopulations which can be distinguished by their gene expression profiles (Furusawa et al., 2006). Mansergh et al. (2009) compared three undifferentiated ESC-lines derived from different mouse strains. They found that the ESC-lines showed a very similar expression pattern when the cells were cultured under identical culture conditions (Mansergh *et al.*, 2009). Although these approaches led to a clear insight in the global gene expression pattern of mouse ESCs, the comparison of different pluripotent cell lines is on the one hand limited to ESCs and embryonic germ cells (EGCs; Sharova et al., 2007) and on the other hand to ESCs and induced pluripotent stem cells (iPSCs; Takahashi and Yamanaka, 2006) until now. Even though maGSCs are pluripotent stem cell lines based on the criteria reported above, it still has to be elucidated whether they also share the global gene expression pattern common to other pluripotent cell lines like ESCs, EGCs and iPSCs.

To address this question, we first compared the global gene expression profiles of maGSCs and ESCs derived from different mouse strains. In order to check for differences and similarities between the cell lines and mouse strains, we concentrated on genes involved in the regulation of pluripotency and genes known to be specifically present in germ cells. Second, we analyzed maGSCs and ESCs cultured under differentiation-promoting conditions for 20 days, namely in the absence of Feeder Layer and leukemia inhibitory factor (LIF), but in the presence of retinoic acid (RA). In this approach, we concentrated on the down-regulation of expression of pluripotency-related genes. With the first approach we wanted to identify a transcriptome which is common for different types of pluripotent cells. With the second approach, the comparison of differentiated and undifferentiated maGSCs and ESCs, it is possible to check if the gene expression of both cell types is similar also after spontaneous differentiation also.

Materials and Methods

Cell culture

In this study, we used two mouse maGSC-lines and the respective ESC-lines derived from wildtype 129SV- and transgenic Stra8-EGFP/Rosa26-LacZ-background. The used Stra8-EGFP/Rosa26-LacZ-background is a mixed background consisting of FVB-, C57Bl- and 129SV-mouse inbred strains. The establishment of maGSC-lines was performed as follows:

testicular cell suspensions of 129SV- and of transgenic mice were cultured under ESC-conditions. The maGSCs-lines (maGSC 129SV, maGSC Stra8) were generated without genetic selection and identified morphologically on the basis of formation of ESC-like colonies (Guan et al., 2006, 2009). The ESC-line from Stra8-EGFP/Rosa26-LacZ-background was derived as previously described (Cheng et al., 2004) and designated as 'ES Stra8'. From 129SV-background we used the established 'ES RI' cell line (Nagy et al., 1993).

The undifferentiated cell lines were maintained on Mitomycin C—inactivated MEF and cultured in standard ESC medium consisting of DMEM (PAN, Aidenbach, Germany) supplemented with 20% defined fetal bovine serum (PAN), 1% penicillin/streptomycin (Gibco BRL, Eggenstein, Germany), 0.1 mM non-essential amino acids (Gibco), 2 mM L-glutamine (Gibco), 1 mM sodiumpyruvate (Gibco), 0.1 mM β-mercaptoethanol (Gibco) and 1000 U/ml LIF (Chemicon, Temecula, CA, USA). For the differentiation of maGSC- and ESC-lines, we cultured the cells for 20 days on gelatine-coated culture dishes with standard ESC medium without LIF, but in the presence of 10^{-6} M RA (Sigma, Deisenhofen, Germany). Comparable passages of cells were used for subsequent analyses, namely ES RI P21, ES RI diff P22 and P23, maGSC 129SV P20, maGSC 129SV diff P18, ES Stra8 P18 and maGSC Stra8 P15.

Microarray analysis using the GeneChip® Mouse Gene 1.0 ST Array

RNA isolation was performed from three biological replicates from cell lines ES RI, ES Stra8, maGSC 129SV, maGSC Stra8, ES RI diff and maGSC 129SV diff using the Trizol (Invitrogen) method according to the manufacturer's recommendations. Afterwards, the samples were DNAse I (Sigma) treated in order to remove DNA contamination. RNA quality was determined using the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) microfluidic electrophoresis. Only samples with comparable RNA integrity numbers were selected for microarray analysis.

Total RNA of 0.3 μg was used as starting material to prepare cDNA. The synthesis of double-stranded cDNA was done with the WT Target Labeling and Control Reagents (Affymetrix, Santa Clara, CA, USA). The cleanup of double-stranded cDNA was done using the GeneChip® Sample Cleanup module (Affymetrix).

The *in vitro* transcription was conducted with the WT Target Labeling Kit (Affymetrix). The total amount of the reaction product was purified with the GeneChip® cRNA Sample Cleanup Module (Affymetrix) and quantified using the NanoDrop ND-1000. A cDNA synthesis (ss) was performed using the WT Target Labeling Kit (Affymetrix). Total ssDNA of 5.5 μg was cleaved into fragments of 35–200 bases by enzymatic processes. The degree of fragmentation and the length distribution of the ssDNA were checked by capillary electrophoresis using the Agilent 2100 Bioanalyzer (Agilent Technologies). A terminal labeling reaction (Biotin) was performed after fragmentation using the WT Labeling Kit (Affymetrix).

Biotinylated fragmented ssDNA was hybridized onto the GeneChip® Mouse Gene 1.0 ST Array (Affymetrix) according to the manufacturer's recommendation. The hybridization was performed for 16 h at 1g and 45°C in the GeneChip® Hybridization Oven 640 (Affymetrix). Washing and staining of the arrays were done on the Gene Chip® Fluidics Station 450 (Affymetrix) according to the manufacturer's recommendation. The antibody signal amplification and washing and staining protocol were used to stain the arrays with streptavidin R-phycoerythrin (SAPE; Invitrogen). To amplify staining, SAPE solution was added twice with a biotinylated anti-streptavidin antibody (Vector Laboratories, Burlingame, CA, USA) staining step in between. Arrays were scanned using the GeneChip® Scanner 3000 7G (Affymetrix).

Microarray data analysis

Intensity data were extracted using the Affymetrix AGCC Software (version 2.0) (Affymetrix) and analyzed using the affy (Gautier et al., 2004) and Limma package (Smyth, 2004) of Bioconductor (Gentleman et al., 2004). The microarray data discussed in this paper were generated conforming to the MIAME guidelines and are deposited in the Gene Expression Omnibus (GEO) database at http://www.ncbi.nlm.nih.gov/geo/ (accession-number GSE22605).

The microarray data analysis consists of the following steps: 1. between-array normalization, 2. probe summary, 3. global clustering and PCA-analysis, 4. fitting the data to a linear model and 5. detection of differential gene expression. Quantile-normalization was applied to the log2-transformed intensity values as a method for between-array normalization, to ensure that the intensities had similar distributions across arrays (Irizarry et al., 2003). For the summary of probes we used a median polish procedure.

For cluster analysis we used a hierarchical approach with the average linkage-method. Distances were measured as 1 − Pearson's Correlation Coefficient. The PCA was performed using the princomp-function in the R software. To estimate the average group values for each gene and assess differential gene expression, a simple linear model was fit to the data, and group-value averages and standard deviations for each gene were obtained.

To find genes with significant expression changes between groups, empirical Bayes statistics were applied to the data by moderating the standard errors of the estimated values (Smyth, 2004).

P-values were obtained from the moderated t-statistic and corrected for multiple testing with the Benjamini−Hochberg method (Benjamini and Hochberg, 1995). The P-value adjustment guarantees a smaller number of false positive findings by controlling the false discovery rate (fdr). For each gene, the null hypothesis, that there is no differential expression between degradation levels, was rejected when its fdr was lower than 0.05.

Bioinformatics

The classification of the identified transcripts according to their biological functions and localization in cellular compartments was performed using DAVID bioinformatics (http://david.abcc.ncifcrf.gov/) (Supplementary Tables S2 and S5). Data from NCBI Databases Gene and PubMed are shown in Supplementary Tables S3 and S4. Comparison of gene lists was carried out using Whitehead BaRC public tools (http://jura.wi.mit.edu/bioc/tools/).

Results

Characterization of maGSCs and ESCs

We first examined the expression of known pluripotency-related marker genes to show that maGSCs and ESCs used in our experiments are both pluripotent cell lines. Western blot analysis revealed that Oct4 (Okamoto et al., 1990; Schöler et al., 1990), Nanog (Mitsui et al., 2003), Sall4 (Zhang et al., 2006), Sox2 (Avilion et al., 2003) and Zfp206 (Wang et al., 2007a, b) are expressed in equal amounts in maGSCs and ESCs but not in MEF (Supplementary Fig. S1A). Additionally, FACS-analysis was performed to determine the percentage of cells that express the pluripotency-related gene Sox2 (Supplementary Fig. S1B). It was found that in maGSC 129SV and ES R1 92 and 98% of cells, respectively, were positive for Sox2, showing that the cell lines were a homogenous population of pluripotent cells.

For differentiation of maGSC 129SV and ES R1 the cells were cultured on gelatine-coated dishes in the absence of LIF but in the presence of RA. After 20 days of culture we isolated whole protein and performed western blot analysis for expression of pluripotency marker gene Oct4. No Oct4-expression could be detected at this time point (Supplementary Fig. S1C). To make sure that a spontaneous differentiation has taken place the expression of marker genes for the three germ layers (Vimentin, Nestin, Hnf4) was studied by qRT−PCR. It was found that the expression of these marker genes strongly increased during differentiation (Supplementary Fig. S1D). Additionally, the morphology of the cells changed during differentiation (Supplementary Fig. S1E). maGSC 129SV and ES R1 after 20 days of culture under differentiation-promoting conditions lost their colony-like appearance, but still showed a homogenous differentiated phenotype.

These results clearly demonstrate that maGSCs and ESCs used in these experiments expressed pluripotency marker genes. During a 20 day differentiation period the expression of these marker genes diminished whereas the expression of genes characteristic for the three germ layers increased. Therefore these undifferentiated and differentiated maGSCs and ESCs were used for transcriptome analysis.

Global comparison of the maGSC- and ESC-transcriptomes

Using a Mouse Whole Genome Microarray (Mouse Gene 1.0 ST Array, Affymetrix) we obtained the global gene expression profiles of two different mouse maGSC- and ESC-lines. In addition, the expression profiles of one maGSC- and one ESC-line cultured under differentiation-promoting conditions were analyzed.

To first get an overall impression of the global similarities between the different cell lines we performed principal component analysis (PCA) of the three analyzed replicates of maGSC 129SV, maGSC Stra8, ES R1, ES Stra8 and their differentiated counterparts (Fig. 1). We found very similar gene expression profiles in undifferentiated maGSCs and ESCs derived from the same mouse background. Small strain-specific differences could be observed when comparing the cell types derived from different mouse backgrounds. The expression pattern of both maGSCs and ESCs changed considerably during their RA-induced spontaneous differentiation. However, the differentiated cell types maGSC 129SV diff and ES R1 diff show a very similar transcriptome, suggesting that they differentiate spontaneously towards the same lineages.

When comparing the global gene expression patterns we found a 97−99% identity in undifferentiated cells of 129SV- and transgenic Stra8-EGFP/Rosa26-LacZ-background, respectively. These findings show that the transcriptomes of maGSCs and ESCs are similar but not totally identical. To further examine these small differences we analyzed the number of genes which are differentially expressed between the cell lines. We found that only 621 genes out of 20 986 genes analyzed show a more than 2-fold change in gene expression when comparing ES R1 and maGSC 129SV and only 72 genes when comparing ES Stra8 and maGSC Stra8, respectively (Fig. 2A). Higher expressed in ES R1 are 518 genes, whereas in maGSC 129SV 103 genes are predominantly transcribed. In the transgenic Stra8-EGFP/Rosa26-LacZ background 17 and 55 transcripts are stronger expressed in ES Stra8 and maGSC Stra8, respectively. We classified

Figure 1 PCA showing the similarities and differences between global gene expression in undifferentiated and differentiated maGSC- and ESC-lines. Each biological replicate analyzed is depicted by a single rhomb. Undifferentiated cell lines derived from the same mouse background are nearly identical, whereas small differences can be observed between cells derived from different mouse backgrounds. Both differentiated cell lines have a very similar gene expression pattern but are very different from their undifferentiated counterparts.
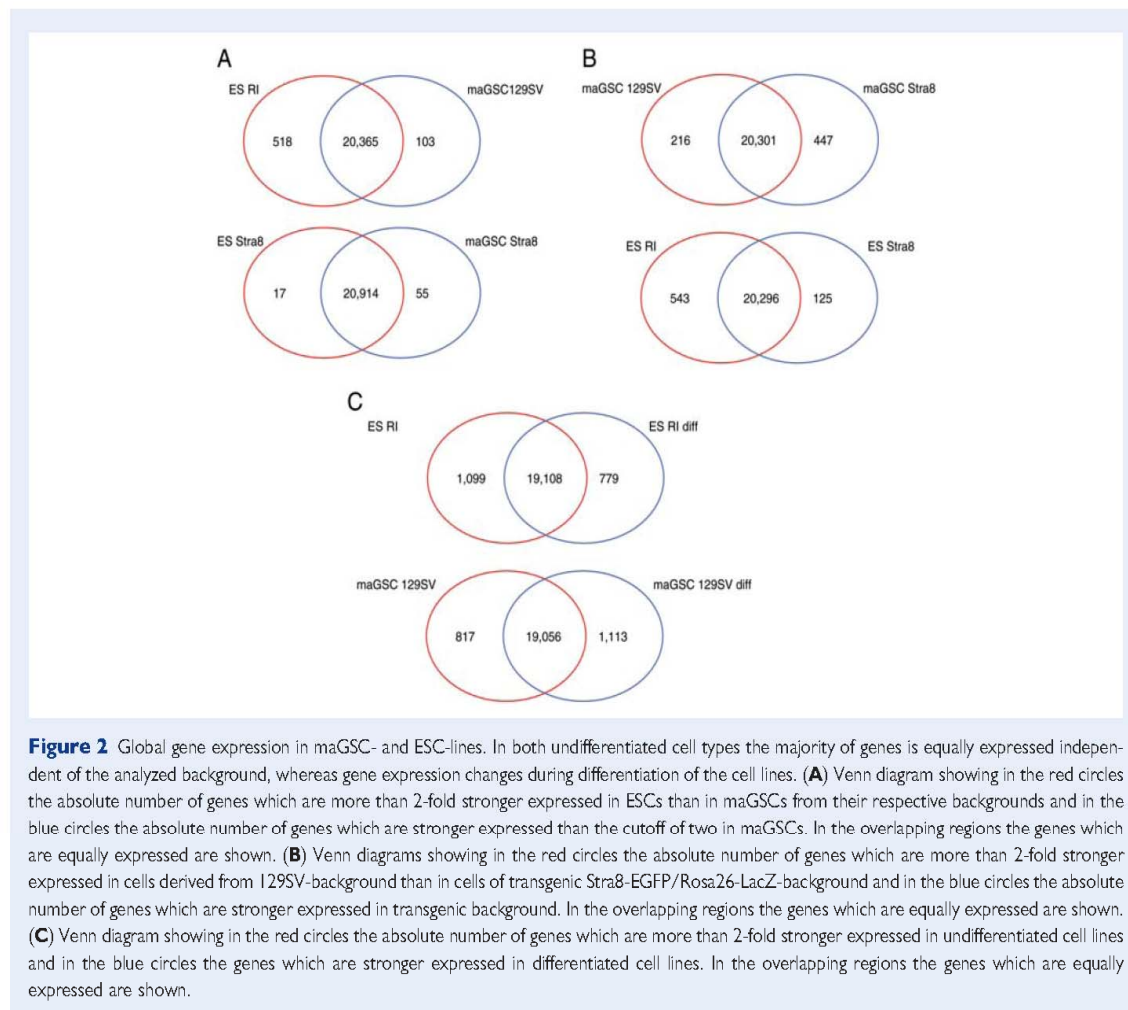
the differentially expressed genes according to their biological functions based on Gene Ontology (GO)-terms and found mostly genes with unknown function. The remaining genes belong to various categories (Supplementary Table SII). We further analyzed the genes which show the largest differences in gene expression (at least 4-fold) (Supplementary Table SIII). Differentially expressed genes playing a role in signal transduction and regulation of transcription could be found. This suggests that the cell type-specific transcripts define biological pathways that distinguish maGSCs and ESCs. Higher expressed in maGSCs are mostly transcripts with an unknown function, whereas higher expressed in ESCs are genes predominantly involved in immune response like Cxcl10, Ifit3 and various major histocompatibility complex-genes (e.g. H2-Q6, H2-T22 and H2-K1). Next, we wanted to figure out if transcripts exist which show differential expression between multiple maGSC- and ESC-lines and are consistently higher expressed in one of the cell types. These genes then could be used to distinguish between maGSCs and ESCs. Surprisingly, no such transcripts could be found in either maGSCs or ESCs (Table I). These results suggest that maGSCs as well as ESCs from different mouse backgrounds show differences in their gene expression pattern.

Thereafter, we evaluated the differences in gene expression between both maGSC- and ESC-lines. With this approach we could find that 663 genes are differentially expressed when comparing

both maGSC-lines (216 genes are stronger expressed in maGSC 129SV, 447 are stronger expressed in maGSC Stra8) (Fig. 2B). In both ESC-lines a similar result was found: 136 genes are up-regulated in ES Stra8, whereas 554 genes are higher expressed in ES RI (Fig. 2B). Next, we checked for genes which could be used to differentiate between 129SV- and Stra8-EGFP/Rosa26-LacZ-background. These genes should be consistently higher expressed in both cell types from one of the mouse backgrounds and additionally should show differential expression between backgrounds. With this procedure it was possible to identify genes which are overexpressed in both cell types from the same background, namely 11 genes in 129SV- and 50 genes in Stra8-EGFP/Rosa26-LacZ-background, respectively (Table I). The commonly up-regulated genes belong to various biological functions and cellular compartments with a large fraction being localized in the plasma membrane and being involved in signal transduction (Supplementary Table SIV).

Finally, we compared gene expression in undifferentiated and differentiated maGSCs and ESCs. We identified 817 genes higher expressed in undifferentiated maGSCs and 1113 genes higher expressed in differentiated maGSCs. For ES RI we found 1099 genes being up-regulated in undifferentiated cells and 779 genes being up-regulated in differentiated cells (Fig. 2C). Next, we analyzed whether gene expression of both cell types changed similarly during differentiation. Therefore, we selected genes which were up-regulated in both types of undifferentiated or

Meyer *et al.*

Figure 2 Global gene expression in maGSC- and ESC-lines. In both undifferentiated cell types the majority of genes is equally expressed independent of the analyzed background, whereas gene expression changes during differentiation of the cell lines. (**A**) Venn diagram showing in the red circles the absolute number of genes which are more than 2-fold stronger expressed in ESCs than in maGSCs from their respective backgrounds and in the blue circles the absolute number of genes which are stronger expressed than the cutoff of two in maGSCs. In the overlapping regions the genes which are equally expressed are shown. (**B**) Venn diagrams showing in the red circles the absolute number of genes which are more than 2-fold stronger expressed in cells derived from 129SV-background than in cells of transgenic Stra8-EGFP/Rosa26-LacZ-background and in the blue circles the absolute number of genes which are stronger expressed in transgenic background. In the overlapping regions the genes which are equally expressed are shown. (**C**) Venn diagram showing in the red circles the absolute number of genes which are more than 2-fold stronger expressed in undifferentiated cell lines and in the blue circles the genes which are stronger expressed in differentiated cell lines. In the overlapping regions the genes which are equally expressed are shown.

differentiated cells, respectively (Table I). A total of 455 genes were up-regulated in both undifferentiated cell lines, and 501 up-regulated transcripts could be identified in both differentiated cell lines. These results reveal that both cell types differentiate spontaneously towards the same lineages. The biological functions of these genes up-regulated in both types of undifferentiated or differentiated cells, respectively, are shown in Supplementary Table SV. During the differentiation of both cell types the gene expression of several biological categories changed. The biggest differences in expression could be found in genes with unknown function. Changes were also detected in genes involved in the regulation of transcription; most known pluripotency-regulating genes belong to this group whose expression was expected to cease during differentiation of the cell lines.

Expression analysis of genes involved in regulation of pluripotency

Since ESCs and maGSCs are both pluripotent cell lines, we decided to include genes in our study that are involved in the maintenance of pluripotency or are known to be ESC-specific. The fold changes in expression for 24 known pluripotency-related genes in

undifferentiated cell lines are shown in Table II. Most of these genes are equally expressed in maGSCs and ESCs which means that they show no significant difference in gene expression as based on the applied cutoff. Only two genes, Nr0b1 and Tex19.1, show a slightly higher than 2-fold difference in expression in 129SV-background, whereas these genes are equally expressed in transgenic Stra8-EGFP/Rosa26-LacZ-background. This finding additionally shows that there are slight differences between the different mouse backgrounds. These differences could also be found in lower amount in expression of Dppa4, Fgf4 and others. The expression of these genes also varies between the wildtype and transgenic backgrounds. However, taken together these results suggest that maGSCs and ESCs possess the same pathways for the maintenance of pluripotency. For several of the pluripotency-related genes expression profiling by qRT–PCR was also performed. Herewith it was possible to validate the results of the Microarray, which means that none of the analyzed pluripotency-regulating genes showed a statistically significant difference in gene expression pattern in the four analyzed pluripotent cell lines (Supplementary Fig. S2).

After differentiation of maGSC 129SV and ES RI with RA for 20 days, we found a down-regulation of expression of most

Table I Global comparison of genes which are up-regulated in more than one cell type.

(A) Comparison of the genes which are up-regulated (↑) in ESCs and maGSCs from wildtype 129SV- (first column) and transgenic mouse background (second column). No genes are commonly up-regulated in neither maGSCs from different mouse backgrounds nor in ESCs from different mouse backgrounds (third column)

	129SV-background	Stra8-EGFP/ Rosa26-LacZ-background	Transcripts commonly up-regulated in both backgrounds
ESCs	ES RI ↑ 518	ES Stra8 ↑ 17	0
maGSCs	maGSC 129SV ↑ 103	maGSC Stra8 ↑ 55	0

(B) Comparison of the genes which are up-regulated (↑) in maGSCs (first column) and ESCs (second column) derived from the same mouse background identifies genes which are overexpressed in both cell types (third column)

	maGSCs	ESCs	Transcripts commonly up-regulated in both cell types
129SV-background	↑ 216	↑ 543	11
Stra8-EGFP/ Rosa26-LacZ-background	↑ 447	↑ 125	50

(C) Comparison of the genes which are up-regulated in both types of undifferentiated (first line) and differentiated cells (second line), respectively, reveals the presence of commonly up-regulated transcripts

	maGSCs	ESCs	Transcripts commonly up-regulated in undifferentiated (line 1) or differentiated (line 2) cell lines
Undifferentiated	↑ 817	↑ 1099	455
Differentiated	↑ 1113	↑ 779	501

pluripotency-related genes analyzed (Table III). In most cases this down-regulation is stronger in ESCs than in maGSCs. This suggests that ESCs differentiate faster and stronger than maGSCs. Surprisingly, in both cell types several pluripotency-related genes could be found that did not change their expression pattern during differentiation, whereas even three genes, SSEA1, Trap1a and Zfx showed a slight increase in expression during differentiation. These results were again validated by qRT–PCR (Supplementary Fig. S2).

Discussion

In this study, we compared the transcriptional profiles of pluripotent stem cell lines derived from different developmental stages: cells from the ICM of blastocysts (ESCs) and maGSCs from SSCs of adult mouse testis. In addition, the transcriptomes of differentiated cells derived from maGSCs and ESCs were analyzed. As far as we know, this is the first study comparing ESCs with a pluripotent cell line derived from an adult organism on transcriptional level.

One of the main questions we wanted to address was if maGSCs and ESCs are distinguishable by global gene expression patterns. Although some differences between both cell types could be found, it is evident that 97% of genes are identically expressed in maGSCs and ESCs of 129SV-background and even 99% of genes are identically expressed in Stra8-EGFP/Rosa26-LacZ-background, which means that cells of transgenic background are more similar than those of wildtype background. This fact was also noticed by Zechner et al. (2009), who identified a more similar DNA methylation pattern in transgenic Stra8-EGFP/Rosa26-LacZ-background when comparing them with wildtype cell lines (Zechner et al., 2009).

Owing to their different developmental origins we expected that maGSCs and ESCs show differences in expressed genes after differentiation.

However, after differentiation both cell types still showed a very similar gene expression suggesting that maGSCs and ESCs differentiated towards the same lineages under the differentiation conditions used in this study (Doetschmann et al., 1985). These similarities support the hypothesis that both cell types are developmentally derived from the same precursor cells. This is consistent with the theory of Zwaka and Thomson (2005) who hypothesized that pluripotent cell lines with common properties are derived from a common origin, namely from germ cells (Zwaka and Thomson, 2005). This hypothesis is also supported by the fact that we could find different germ cell specific genes to be expressed at the same level in ESCs as in maGSCs (Supplementary Table SVI).

Since maGSCs and ESCs are pluripotent cell lines we analyzed the expression patterns of 24 known pluripotency marker genes. Apart from Nr0b1 and Tex19.1, which show a slight difference in gene expression in 129SV-background, all other pluripotency marker genes were found to be equally expressed in both cell types. According to the microarray data Nr0b1 is stronger expressed in ES RI than in maGSC 129SV, whereas Tex19.1 is stronger expressed in maGSC 129SV (Table II). Both proteins are suggested to be involved in the Oct4-dependent pluripotency regulating pathway: Nr0b1 is known to regulate pluripotency by binding to the Oct4-protein and inhibiting the promoter-binding activity of Oct4, thereby leading to a loss of pluripotency (Sun et al., 2009). Tex19.1 is colocalized with Oct4 in the ICM of blastocysts, and expressed in different pluripotent cell lines. A Nanog-binding site could be identified in the 3′-proximal region of Tex19.1 (Kuntz et al., 2008).

Despite these minor differences in 129SV-background, our results demonstrate that in maGSCs as well as in ESCs the genes involved in the regulation of pluripotency are equally expressed. This provides strong evidence that the mechanisms and pathways that control pluripotency are identical in both cell types. In addition, as expected, for

Table II Comparison of expression levels between undifferentiated maGSCs and ESCs of 24 randomly chosen known pluripotency marker genes.

Gene name	NCBI ID	Annotation	Fold difference ES RI – maGSC 129SV	Fold difference ES Stra8 – maGSC Stra8
Dppa4	NM_028610	Developmental pluripotency-associated 4	−1.26	1.83
Dppa5	NM_025274	Develpopmental pluripotency-associated 5	1.16	1.31
Epha2	NM_010139	Eph receptor A2	1.30	1.21
Esrrb	NM_011934	Estrogen-related receptor, beta	1.04	1.38
Fgf4	NM_010202	Fibroblast growth factor 4	−1.08	1.49
Gdf3	NM_008108	Growth differentiation factor 3	1.21	1.24
Jarid1b	NM_152895	Jumonji, AT rich interactive domain 1B	1.21	−1.13
Klf4	NM_010637	Kruppel-like factor 4	1.06	1.29
Lin28	NM_145833	Lin-28 homolog	−1.30	−1.28
Nanog	NM_028016	Nanog homeobox	1.18	1.83
Nr0b1	NM_007430	Nuclear receptor subfamily 0, group B, member 1	2.17	1.42
Sall4	NM_201395	Sal-like 4	−1.12	1.02
Slc2a3	NM_011401	Solute carrier family 2 (facilitated glucose transporter), member 3	1.64	1.53
Sox2	NM_011443	SRY-box containing gene 2	1.09	1.27
Tbx3	NM_011535	T-box 3	1.56	1.09
Tcf3	NM_001079822	Transcription factor 3	−1.95	1.02
Tcl1	NM_009337	T-cell lymphoma breakpoint 1	−1.23	1.65
Tex19.1	NM_028602	Testis expressed gene 19, transcript 1	−2.87	1.24
Trap1a	NM_011635	Tumor rejection antigen P1A	−1.19	1.32
Zfp42	NM_009556	Zinc finger protein 42	−1.14	1.61
Zfp143	NM_20841	Zinc finger protein 143	1.23	−1.12
Zfp281	NM_226442	Zinc finger protein 281	1.00	−1.48
Zic3	NM_009575	Zinc finger protein of the cerebellum 3	1.16	1.16
Zscan10	NM_001033425	Zinc finger and SCAN domains 10	−1.71	−1.03

Most of the analyzed genes show a <2-fold change in gene expression and are therefore considered as equally expressed. Numbers shown with a minus (−) sign are higher expressed in maGSCs while numbers without a sign are higher expressed in ESCs.

nearly all pluripotency marker genes studied a down-regulation of their expression was observed after differentiation of both cell lines. It is noticeable that down-regulation generally occurs stronger in ESCs than in maGSCs, indicating that maGSCs differentiate slower than ESCs. This observation is supported by the results of Zovoilis et al. (2008), who found that expression of the pluripotency-specific micro-RNA family 290 is more stable in maGSCs during differentiation than in ESCs cultured under the same conditions (Zovoilis et al., 2008). However, for SSEA1 and Trap1a no down-regulation could be observed during differentiation of maGSCs and ESCs, but their expression even slightly increased during differentiation.

Surprisingly, differences in gene expression were larger in cells derived from different mouse strains than between maGSCs and ESCs. When comparing both maGSC- and ESC-lines 686 and 691 genes, respectively, were found to be differentially expressed (data not shown), whereas maGSC 129SV and ES RI or maGSC Stra8 and ES Stra8 are distinguishable only by 621 and 72 genes, respectively (Fig. 2). These results indicate that cell lines derived from different mouse strains show relevant differences in gene expression, which is consistent with the findings of Sharova et al. (2007) who reported

the same for the comparison of ESCs and EGCs. Despite these small differences the transcriptomes of both maGSC- and both ESC-lines show a high similarity of 97%, which is higher than the reported value of 70–87% similarity between transcriptomes of different ESC-lines (Mise et al., 2008). This suggests that a common expression pattern of pluripotent cell lines cultured under identical culture conditions exists.

Sharova et al. (2007) compared EGCs and ESCs at the global gene expression level and they found that both pluripotent cell types are indistinguishable. However, it was possible to identify a subset of 'signature genes' which showed a consistent differential expression between the two cell types analyzed (Sharova et al., 2007). We wanted to find out whether maGSCs are even more closely related to EGCs than to ESCs due to their common origin. Therefore, we compared Sharova's 'signature genes' with our transcriptional profiling results. When we checked the published 20 'ESC-signature genes' against our transcriptional profiling we found that only one gene, namely Grb10, showed a significantly different expression between maGSC 129SV and ES RI. Grb10 encodes a cytoplasmic protein that is involved in signal transduction and regulation of glucose import in

Table III Comparison of expression levels of 28 randomly chosen pluripotency-related genes in undifferentiated and differentiated cell lines.

Gene name	NCBI ID	Annotation	Fold difference ES RI – ES RI diff	Fold difference maGSC 129SV – maGSC 129SV diff
Dppa4	NM_028610	Developmental pluripotency-associated 4	4.00	3.43
Dppa5	NM_025274	Develpopmental pluripotency-associated 5	2.33	2.10
Epha2	NM_010139	Eph receptor A2	2.10	1.57
Esrrb	NM_011934	Estrogen-related receptor, beta	4.32	3.23
Fgf4	NM_010202	Fibroblast growth factor 4	4.96	4.72
FoxD3	NM_010425	Forkhead box 3D	1.49	1.77
Gdf3	NM_008108	Growth differentiation factor 3	4.32	2.85
Jarid1b	NM_152895	Jumonji, AT rich interactive domain 1B	1.68	1.33
Klf2	NM_008452	Kruppel-like factor 2	3.58	2.08
Lin28	NM_145833	Lin-28 homolog	1.95	2.50
Nanog	NM_028016	Nanog homeobox	2.48	1.52
Nr0b1	NM_007430	Nuclear receptor subfamily 0, group B, member 1	13.9	4.53
Oct4	NM_013633	POU domain, class 5, transcription factor 1	8.51	5.78
Sall4	NM_201395	Sal-like 4	1.87	1.64
Slc2a3	NM_011401	Solute carrier family 2 (facilitated glucose transporter), member 3	1.35	−1.36
Sox2	NM_011443	SRY-box containing gene 2	4.26	2.97
SSEA1	NM_010242	Stage-specific embryonic antigen 1	−1.27	−1.67
Tbx3	NM_011535	T-box 3	1.46	1.12
Tcf3	NM_001079822	Transcription factor 3	−1.19	1.33
Tcl1	NM_009337	T-cell lymphoma breakpoint 1	2.36	2.71
Tex19.1	NM_028602	Testis expressed gene 19, transcript 1	1.33	2.20
Trap1a	NM_011635	Tumor rejection antigen P1 A	−1.57	−1.34
Utf1	NM_009482	Undifferentiated embryonic cell transcription factor 1	11.24	5.39
Zfp42	NM_009556	Zinc finger protein 42	3.86	2.68
Zfp281	NM_226442	Zinc finger protein 281	1.33	1.39
Zfx	NM_001044386	Zinc finger protein X-linked	−1.41	−1.45
Zic3	NM_009575	Zinc finger protein of the cerebellum 3	5.78	4.14
Zscan10	NM_001033425	Zinc finger and SCAN domains 10	4.89	6.50

The marker genes are down-regulated during differentiation of maGSCs and ESCs. Numbers shown with a minus (−) sign are higher expressed in differentiated cells while numbers without a sign are higher expressed in undifferentiated cells.

the cell. Next to this gene only 10 others were slightly stronger expressed in ESCs (Supplementary Table SVIIA). Out of the 84 'EGC-signature genes' 12 were differentially expressed between maGSCs and ESCs, all of them unexpectedly stronger in ESCs (Supplementary Table SVIIB). Next, we compared the 'EGC-signature genes' with the 103 genes which we found to be up-regulated in maGSC 129SV. Genes which were up-regulated in both cell types could not be identified. When comparing our 519 ESC-overexpressed genes with the published 'ESC-signature genes' only one transcript, namely Grb10, could be found that was included in both gene lists. These results clearly show that all analyzed pluripotent cell lines are very similar and indistinguishable from a viewpoint of global gene expression. However, every cell type can be distinguished by a subset of specifically regulated genes which are equally expressed in other pluripotent

cell lines. Since the ESC-specific genes we identified and the ones published by Sharova et al. (2007) did not match, it can be concluded that these 'signature genes' are specific for comparison with only one other pluripotent cell line, in this case maGSCs or EGCs, respectively. This shows that it is necessary to further compare different pluripotent cell lines on the level of gene expression to further identify subsets of genes which can be used to distinguish them from each other.

Conclusion

In summary, in this study we have shown that maGSCs and ESCs are very similar cell lines based on their global gene expression profiles. They express the same set of pluripotency-related genes, and show almost no differences in expression levels of these genes. Even after

differentiation the cells show only minor differences in their global gene expression profiles which further supports their similarity.

Authors' roles

Supplementary data

Supplementary data are available at http://molehr.oxfordjournals.org/.

Acknowledgements

Funding

References

Anisimov SV, Tarasov KV, Tweedie D, Stern MD, Wobus AM, Boheler KR. SAGE identification of gene transcripts with profiles unique to pluripotent mouse R1 embryonic stem cells. *Genomics* 2002;**79**:169–176.

Avilion AA, Nicolis SK, Pevny LH, Perez L, Vivian N, Lovell-Badge R. Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev* 2003;**17**:126 140.

Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Statist Soc B* 1995; **57**:289–300.

Cheng J, Dutra A, Takesono A, Garrett-Beal L, Schwartzberg PL. Improved generation of C57BL/6J mouse embryonic stem cells in a defined serum-free media. *Genesis* 2004;**39**:100–104.

Dihazi H, Dihazi GH, Nolte J, Meyer S, Jahn O, Müller GA, Engel W. Multipotent adult germline stem cells and embryonic stem cells: comparative proteomic approach. *J Proteome Res* 2009;**8**:5497–5510.

Doetschmann TC, Eistetter H, Katz M, Schmidt W, Kemler R. The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *J Embryol Exp Morph* 1985;**87**:27–45.

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

Furusawa T, Ikeda M, Inoue F, Ohkoshi K, Hamano T, Tokunaga T. Gene expression profiling of mouse embryonic stem cell subpopulations. *Biol Reprod* 2006;**75**:555–561.

Gautier L, Cope L, Bolstad BM, Irizarry RA. affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 2004;**20**:307–315.

Geijsen N, Horoschak M, Kim K, Gribnau J, Eggan K, Daley GQ. Derivation of embryonic germ cells and male gametes from embryonic stem cells. *Nature* 2004;**427**:148–154.

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 2004;**5**:R80.

Guan K, Nayernia K, Maier LS, Wagner S, Dressel R, Lee JH, Nolte J, Wolf F, Li M, Engel W et al. Pluripotency of spermatogonial stem cells from adult mouse testis. *Nature* 2006;**440**:1199–1203.

Guan K, Wolf F, Becker A, Engel W, Nayernia K, Hasenfuss G. Isolation and cultivation of stem cells from adult mouse testes. *Nat Protoc* 2009;**4**:143–154.

Houvabiy HB, Murray MF, Sharp PA. Embryonic stem cell-specific MicroRNAs. *Dev Cell* 2003;**5**:351–358.

Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 2003;**4**:249–264.

Ivanova NB, Dimos JT, Schaniel C, Hackney JA, Moore KA, Lemsichka IR. A stem cell molecular signature. *Science* 2002;**298**:601–604.

Izadyar F, Pau F, Marh J, Slepko N, Wang T, Gonzales R, Ramos T, Howerton K, Sayre C, Silva F. Generation of multipotent cell lines from a distinct population of male germ line stem cells. *Reproduction* 2008;**135**:771–784.

Kanatsu-Shinohara M, Inoue K, Lee J, Yoshimoto M, Ogonuki N, Miki H, Baba S, Kato T, Kazuki Y, Toyokuni S et al. Generation of pluripotent stem cells from neonatal mouse testis. *Cell* 2004;**119**:1001–1012.

Kanatsu-Shinohara M, Lee J, Inoue K, Ogonuki N, Miki H, Toyokuni S, Ikawa M, Nakamura T, Ogura A, Shinohara T. Pluripotency of a single spermatogonial stem cell in mice. *Biol Reprod* 2008;**78**:681–687.

Ko K, Tapia N, Wu G, Kim JB, Bravo MJ, Sasse P, Glaser T, Ruau D, Han DW, Greber B et al. Induction of pluripotency in adult unipotent germline stem cells. *Cell Stem Cell* 2009;**5**:3–4.

Kuntz S, Kieffer E, Bianchetti L, Lamoureux N, Fuhrmann G, Viville S. Tex19, a mammalian-specific protein with a restricted expression in pluripotent stem cells and germ line. *Stem Cells* 2008;**26**:734–744.

Mansergh FC, Daly CS, Hurley AL, Wride MA, Hunter SM, Evans MJ. Gene expression profiles during early differentiation of mouse embryonic stem cells. *BMC Dev Biol* 2009;**9**:5.

Martin GR. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci USA* 1981;**78**:7634–7638.

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

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

Nagy A, Rossant J, Nagy R, Abramov-Newerly W, Roder JC. Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. *Proc Natl Acad Sci USA* 1993;**90**:8424–8428.

Nayernia K, Nolte J, Michelmann HW, Lee JH, Rathsack K, Drusenheimer N, Dev A, Wulf G, Ehrmann IE, Elliott DJ et al. In

vitro-differentiated embryonic stem cells give rise to male gametes that can generate offspring mice. *Dev Cell* 2006;**11**:125–132.

Okamoto K, Okazawa H, Okuda A, Sakai M, Muramatsu M, Hamada H. A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells. *Cell* 1990;**60**:461–472.

Ramalho-Santos M, Yoon S, Matsuzaki Y, Mulligan RC, Melton DA. 'Stemness': transcriptional profiling of embryonic and adult stem cells. *Science* 2002;**298**:597–600.

Schöler HR, Dressler GR, Balling R, Rohdewohld H, Gruss P. Oct-4: a germline-specific transcription factor mapping to the mouse t-complex. *EMBO J* 1990;**9**:2185–2195.

Seandel M, James D, Shmelkov SV, Falciatori I, Kim J, Chavala S, Scherr DS, Zhang F, Torres R, Gale NW *et al.* Generation of functional multipotent adult stem cells from GPR125+ germline progenitors. *Nature* 2007; **449**:346–350.

Sharova LV, Sharov AA, Piao Y, Shaik N, Sullivan T, Stewart CL, Hogan BLM, Ko MSH. Global gene expression profiling reveals similarities and differences among mouse pluripotent stem cells of different origins and strains. *Dev Biol* 2007;**307**:446–459.

Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004;**3**:Article 3.

Suda Y, Suzuki M, Ikawa Y, Aizawa S. Mouse embryonic stem cells exhibit indefinite proliferative potential. *J Cell Physiol* 1987;**133**:197–201.

Sun C, Nakatake Y, Akagi T, Ura H, Matsuda T, Nishiyama A, Koide H, Ko MS, Niwa H, Yokota T. Dax1 binds to Oct3/4 and inhibits its transcriptional activity in embryonic stem cells. *Mol Cell Biol* 2009; **29**:4574–4583.

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

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

Wang ZX, Teh CH, Kueh JL, Lufkin T, Robson P, Stanton LW. Oct4 and Sox2 directly regulate expression of another pluripotency transcription factor, Zfp206, in embryonic stem cells. *J Biol Chem* 2007a;**282**:12822–12830.

Wang ZX, Kueh JL, Teh CH, Rossbach M, Lim L, Li P, Wong KY, Lufkin T, Robson P, Stanton LW. Zfp206 is a transcription factor that controls pluripotency of embryonic stem cells. *Stem Cells* 2007b; **25**:2173–2182.

Zechner U, Nolte J, Wolf M, Shirneshan K, Haji NE, Weise D, Kaltwasser B, Zovoilis A, Haaf T, Engel W. Comparative methylation profiles and telomerase biology of mouse multipotent adult germline stem cells and embryonic stem cells. *Mol Hum Reprod* 2009;**15**:345–353.

Zhang J, Tam WL, Tong GQ, Wu Q, Chan HY, Soh BS, Lou Y, Yang J, Ma Y, Chai L *et al.* Sall4 modulated embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. *Nat Cell Biol* 2006;**8**:1114–1123.

Zovoilis A, Nolte J, Drusenheimer N, Zechner U, Hada H, Guan K, Hasenfuss G, Nayernia K, Engel W. Multipotent adult germline stem cells and embryonic stem cells have similar microRNA profiles. *Mol Hum Reprod* 2008;**14**:521–529.

Zwaka TP, Thomson JA. A germ cell origin of embryonic stem cells? *Development* 2005;**132**:227–233.

3.2 Multipotent adult germline stem cells and embryonic stem cells: comparative proteomic approach

In this second part of the thesis, we report about the comparative analysis of undifferentiated ESCs and maGSCs at the proteome level. We generated 2D-reference maps and were able to identify a total number of 409 peptides, corresponding to 166 non-redundant proteins. These proteins were classified and found to belong to various biological categories and cellular compartments. Using 2D-DIGE, it was possible to highlight the ESC-like nature of maGSCs also at protein level. In this experiment, only 18 proteins were found to be differentially expressed between the two cell types, showing that ESCs and maGSCs from the same mouse strain are more similar than the same cell lines (e.g. ESCs or maGSCs) derived from different mouse strains.

Author contributions to the work:

1. Hassan Dihazi: conception and design of experiments, data analysis and interpretation, preparation of manuscript.

2. Gry H. Dihazi: isolation of proteins, performance of 2D-gelelectrophoresis and 2D-DIGE (partial), performance of Western Blot (partial).

3. Jessica Nolte: generation of cell lines, involved in data interpretation, involved in preparation of manuscript.

4. **Sandra Meyer**: performance of cell culture, preparation of cells for proteome analysis, involved in data interpretation, involved in preparation of manuscript.

5. Olaf Jahn: collection and assembly of mass spectrometry data.

6. Gerhard A. Müller: financial and administrative support.

7. Wolfgang Engel: conception and design of experiments, involved in preparation of manuscript.

Multipotent Adult Germline Stem Cells and Embryonic Stem Cells: Comparative Proteomic Approach

Hassan Dihazi,*,[†] Gry H. Dihazi,[†] Jessica Nolte,[†] Sandra Meyer,[†] Olaf Jahn,[§,‖]
Gerhard A Müller,[†] and Wolfgang Engel[‡]

*Department of Nephrology and Rheumatology, Georg-August University Goettingen, Robert-Koch-Strasse 40,
D-37075 Goettingen, Germany, Institute of Human Genetics, Georg-August University Goettingen,
Heinrich-Dueker-Weg 12, D-37073 Goettingen, Germany, Proteomics Group, Max-Planck-Institute of
Experimental Medicine, Hermann-Rein-Strasse 3, D-37075 Goettingen, Germany, and Deutsche
Forschungsgemeinschaft Research Center for Molecular Physiology of the Brain, Humboldtallee 23, D-37073
Goettingen, Germany*

Spermatogonial stem cells isolated from the adult mouse testis acquire under certain culture conditions pluripotency and become so-called multipotent adult germline stem cells (maGSCs). They can be differentiated into somatic cells of the three germ layers. We investigated a subset of the maGSCs and ESCs proteomes using cell lines derived from two different mouse strains, narrow range immobilized pH gradients to favor the detection of less abundant proteins, and DIGE to ensure confident comparison between the two cell types. 2-D reference maps of maGSCs and ESCs in the p*I* ranges 3−6 and 5−8 were created, and protein entities were further processed for protein identification. By peptide mass fingerprinting and tandem mass spectrometry combined with searches of protein sequence databases, a set of 409 proteins was identified, corresponding to a library of 166 nonredundant stem cell-associated proteins. The identified proteins were classified according to their main known/postulated functions using bioinformatics. Furthermore, we used DIGE to highlight the ESC-like nature of maGSCs on the proteome scale. We concluded that the proteome of maGSCs is highly similar to that of ESCs as we could identify only a small subset of 18 proteins to be differentially expressed between the two cell types. Moreover, comparative analysis of the cell line proteomes from two different mouse strains showed that the interindividual differences in maGSCs proteomes are minimal. With our study, we created for the first time a proteomic map for maGSCs and compared it to the ESCs proteome from the same mouse. We confirmed on the proteome level the ESC-like nature of maGSCs.

Keywords: multipotent adult germline stem cells • embryonic stem cell • 2DE • mass spectrometry

Introduction

Stem cells (SCs) and particularly embryonic stem cells (ESCs) are undifferentiated cells generally characterized by their functional capacity to both self-renew and to differentiate into multiple lineages.[1−3] The potential of self-renewing enables SCs to form all cells in the adult body once they receive the right signals. The ability to control and direct ESC differentiation may open new opportunities in regenerative medicine. Despite the tremendous possibilities and chances brought by ESCs for the medicine especially in replacing damaged tissue, there are still many challenges to overcome before application for disease treatment becomes possible. In addition to the immune

reaction after transplantation, ethical issues regarding the usage of embryos represent another type of obstacles which ESCs face on their long way to clinical usage.

Recent studies demonstrated that the germline lineage has the ability to generate pluripotent cells. In 2004, ESC-like cells were found in germ SC cultures established from neonatal mouse testis, designated as multipotent germline stem cells.[4] In 2006, we have isolated and cultured for the first time spermatogonial stem cells (SSCs) from the adult mouse testis which respond to culture conditions and acquire ESC properties.[5] We proved that the pluripotency and plasticity of these cells, which were named multipotent adult germline stem cells (maGSCs), were similar to ESCs. They are able to spontaneously differentiate into derivatives of the three embryonic germ layers in vitro, to generate teratomas in immunodeficient mice and to contribute to the development of various organs when injected into an early blastocyst. In a recent study, we confirmed the ESC-like nature of maGSCs on the microRNA level[6] and demonstrated their potential as an alternative source of pluripotent cells from nonembryonic tissues. The pluripotency

* To whom correspondence should be addressed. Telephone: +049-551-3991221. Fax: +049-551-3991039. E-mail: dihazi@med.uni-goettingen.de.
  † Department of Nephrology and Rheumatology, Georg-August University Goettingen.
  ‡ Institute of Human Genetics, Georg-August University Goettingen.
  § Max-Planck-Institute of Experimental Medicine.
  ‖ Deutsche Forschungsgemeinschaft Research Center for Molecular Physiology of the Brain.

of mouse SSCs was confirmed by Seandel et al.[7] and others. More recently, Conrad et al. generated pluripotent SCs from adult human testis.[8] The established human adult GSCs were able to differentiate into various types of somatic cells of all three germ layers when cultured under conditions that stimulate differentiation of human ESCs.[8]

Considering the many challenges in SC biology, there is a continuous need for the implementation of cutting-edge techniques. The field of proteomics has matured immensely in recent years,[9,10] now allowing investigation at reasonable throughput in all areas of cell biology. After having analyzed maGSCs at the microRNA level,[6] the aim of the present study was to provide first insights into the proteome of maGSCs by using narrow range 2-DE (zooming gels) that improves the detection of low-abundance proteins.[11–13] We report here the 2-D reference maps of maGSCs and ESCs together with a library of 166 nonredundant stem-cell associated proteins. By combining narrow range 2-DE with the DIGE technology, we were able to show that the proteome of maGSCs is highly similar to that of ESCs, a finding that further substantiates the ESC-like nature of maGSCs.

Experimental Section

Culture of Mouse maGSC and ESC Lines. The culture of maGSC lines from both mouse lines 129/Sv (maGSC 129SV) and the transgenic one Stra8-EGFP/ROSA26 (maGSC Stra8) was described previously.[5] In this case, both maGSC lines were generated without genetic selection but by morphological criteria only. The ESC R1 line was derived from the 129/Sv mouse line.[14] The cell line ESC Stra8 was generated from the transgenic Stra8-EGFP/ROSA26 mouse as described previously.[15] Additionally maGSC FVB, and maGSC C57BL/6 from the testes of adult FVB mice, and adult C57BL/6 mice and their corresponding ESCs[16] were used in the Western blot analyses for the confirmation of the proteomics data. To maintain maGSCs and ESCs in an undifferentiated state, the cells were cultured under standard ESC culture conditions: DMEM (PAN, Aidenbach, Germany) supplemented with 20% fetal calf serum (PAN), 2 mM L-glutamine (PAN), 50 mM β-mercaptoethanol (Gibco/Invitrogen, Eggenstein, Germany), $1\times$ nonessential amino acids (Gibco/Invitrogen), sodium pyruvate (Gibco/Invitrogen), penicillin/streptomycin (PAN). maGSCs and ESCs were cultured on a feeder layer of mitomycin C-inactivated mouse embryonic fibroblasts (MEFs) in the presence of 1000 U/mL recombinant mouse leukemia inhibitory factor (LIF) (Chemicon, Temecula, CA).

Protein Extraction. The protein extraction for 2D gel electrophoresis was performed as described previously.[17] Briefly, 75% confluent cultures were scraped and washed three times with PBS. The cells were harvested by centrifugation at $200g$ for 10 min, the pellet was treated with 0.3–0.5 mL lysis buffer (9.5 M urea, 2% CHAPS (w/v), 2% ampholytes (w/v), 1% DTT, 10 mM PMSF). Ampholytes, DTT, pepstatin (to a final concentration of 1.4 μM), and Complete from Roche Diagnostic (according to the manufacturer's protocol) were added before use. After adding the lysis buffer, the samples were incubated for 30 min at 4 °C. To remove the cell debris, sample centrifugation was carried out at 13 000g and 4 °C for 45 min. Supernatant was recentrifuged at 13 000g and 4 °C for an additional 45 min to get maximal purity. The resulting samples were used immediately or stored at −80 °C until use.

Protein Precipitation. To reduce the salt contamination and to enrich the proteins, methanol–chloroform precipitation was

performed according to Wessel and Flugge.[18] Briefly 0.4 mL of methanol (100%) was added to 0.1 mL aliquots of protein samples and mixed together. A total of 0.1 mL of chloroform was added to the mixture and vortexed. Subsequently, 0.3 mL of water was added and the solution was vortexed and centrifuged at 13 000g for 1 min. The aqueous layer was removed, and another 0.4 mL of methanol (100%) was added to the rest of the chloroform and the interphase with the precipitated proteins. The sample was mixed and centrifuged for 2 min at 13 000g and the supernatant was removed. The pellet was vacuum-dried and dissolved in rehydration buffer.

Total protein concentration was estimated using the Bio-Rad protein assay (Bio-Rad, Hercules, CA) according to Bradford.[19] Bovine serum albumin (Sigma, Steinheim, Germany) was used as a standard.

2D Gel Electrophoresis (2-DE). For large-scale 2-DE, IPG strips (11 cm, pI 3–10, pI 5–8 or pI 3–6) were passively rehydrated in 185 μL of solution containing 150 μg of protein in a rehydration buffer (8 M urea, 1% CHAPS, 1% DTT, 0.2% ampholytes, and a trace of bromophenol blue) for 12 h. The IEF step was performed on the PROTEAN IEF Cell (Bio-Rad). With the temperature controlled at 20 °C, the voltage was set to 500 V for 1 h, increased to 1000 V for 1 h, 2000 V for 1 h and left at 8000 V until a total of 50 000 Vh was reached. Prior to SDS-PAGE, the IPG strips were each reduced for 20 min at room temperature in SDS equilibration buffer containing 6 M urea, 30% glycerol, 2% SDS, 0.05 M Tris-HCl, and 2% DTT on a rocking table. The strips were subsequently alkylated in the same solution with 2.5% iodacetamide substituted for DTT, and a trace of bromophenol blue. For the SDS-PAGE, 12% BisTris Criterion precast gels (Bio-Rad) were used according to manufacturer's instructions. The gels were run at 150 V for 10 min following by 200 V until the bromophenol blue dye front had reached the bottom of the gel.

Gel Staining. For image analysis, 2-DE gels were fixed in a solution containing 50% methanol and 12% acetic acid overnight and stained with Flamingo fluorescent gel stain (Bio-Rad) for minimum 5 h. After staining, gels were scanned at 50 μm resolution on a Fuji FLA5100 scanner. The digitalized images were analyzed using Delta 2D 3.4 (Decodon, Braunschweig, Germany). 2-DE gels were poststained with collodial Coomassie (Roti-Blue, Roth, Karlsruhe, Germany) overnight to enable manual spot picking for protein identification.

2D-DIGE. Protein extraction and methanol–chloroform precipitation were performed as described above. The resulting pellet was solubilized in labeling buffer (30 mM Tris-HCl, pH 8.5, 9.5 M urea, 2% CHAPS, 10 mM PMSF), centrifuged (5 min, 13 000g), and the protein concentration of the supernatant was determined as described above.

CyDyes (GE Healthcare) were dissolved in anhydrous N,N-dimethylformamide (DMF) (Sigma-Aldrich, St. Louis, MO) to yield a stock solution containing 1000 pmol/μL. One volume of CyDye stock solution was added to 1.5 vol of DMF to make a 400 μM CyDye working solution. For minimal labeling, 400 pmol of the amine-reactive cyanine dyes Cy3 and Cy5 was added to 50 μg of proteins from maGSCs and ESCs, respectively. The labeling reaction was carried out at 4 °C in the dark for 30 min and the reaction was terminated by addition of 10 nmol lysine at 4 °C in the dark for 10 min. Equal volumes of $2\times$ sample buffer (30 mM Tris-HCl, pH 8.5, 9.5 M urea, 2% CHAPS, 10 mM PMSF, 130 mM DTT and 2% ampholytes 3–10) were added to each of the labeled protein samples. To avoid the dye-specific protein labeling, every pair of protein samples from

Similar Proteoms of maGSCs and ESCs   research articles

two independent cell extract preparation was processed in duplicate while swapping the dyes. Thereby four replicate gels were obtained, allowing to monitor regulation factors down to 2-fold changes.[20] A total of 50 μg of an internal standard consisting of a mixture of all four stem cell samples under investigation was labeled with 400 pmol Cy2 and included on all gels to facilitate gel matching, thereby eliminating artifacts from experimental variation. The three differentially labeled fractions were pooled. Rehydration buffer (8 M urea, 1% CHAPS, 13 mM DTT and 1% ampholytes 3–10) was added to make up the volume to 185 μL prior to IEF. The 2-DE was performed as described above. The CyDye-labeled protein gels were scanned at 50 μm resolution on a Fuji FLA5100 scanner (Fuji Photo, Kanagawa, Japan) with excitation light 473 nm (Cy2), 575 nm (Cy3) and 635 nm (Cy5). Fluorescent images were acquired in 16-bit TIFF files format. Spot matching across gels and normalization based on the internal standard was performed with Delta2D software (Decodon, Greifswald, Germany). To analyze the significance of protein regulation, a Student's t-test was performed, and statistical significance was assumed for p-values less than 0.01. 2D gels were poststained as described above and differentially regulated proteins were excised and processed for identification by MS.

Protein Identification. Manually excised gel plugs were subjected to an automated platform for the identification of gel-separated proteins[21] as described in the framework of recent DIGE-based[22] and large-scale proteome studies.[23] An Ultraflex MALDI-TOF mass spectrometer (Bruker Daltonik) was used to acquire both peptide mass fingerprint (PMF) and fragment ion spectra, resulting in confident protein identifications based on peptide mass and sequence information. Database searches in the Swiss-Prot primary sequence database restricted to the taxonomy *Mus musculus* were performed using the MASCOT Software 2.2 (Matrix Science). Carboxamidomethylation of Cys residues was specified as fixed and oxidation of Met residues as variable modifications. One trypsin missed cleavage was allowed. Mass tolerances were set to 100 ppm for PMF searches and to 100 ppm (precursor ions) and 0.7 Da (fragment ions) for MS/MS ion searches. The minimal requirement for accepting a protein as identified was at least one peptide sequence match above identity threshold in coincidence with at least 20% sequence coverage in the PMF.

Alternatively, tryptic peptides were subjected to mass spectrometric sequencing using a Q-TOF Ultima Global mass spectrometer (Micromass, Manchester, U.K.). For that purpose, gel plugs were excised from 2-D gels and digested as described previously.[24] After digestion, the supernatant was removed and saved, and the additional peptides were extracted with different acetonitrile/ trifluoroacetic acid ratio under sonication. All supernatants were pooled together, dried in a vacuum centrifuge, and redissolved in 0.1% formic acid. The mass spectrometric sequencing was performed as described previously.[25]

Processed data were searched against MSDB and Swiss-Prot databases through Mascot search engine using a peptide mass tolerance of 50 ppm and fragment mass tolerance of 0.1 Da. Protein identifications with at least two peptides sequenced were considered significant.

Western Blot Analysis. A selection of the proteins found to be at different levels in the maGSCs and ESCs were tested with Western blot in four different maGSC cell line types and their corresponding ESCs isolated from four different mouse lines (Stra8-EGFP/ROSA26, 129/Sv, FVB and C57BL/6) according to Towbin et al.[26] Rabbit monoclonal anti-Eif5a, rabbit polyclonal

anti-Tardbp, rabbit polyclonal anti-Stmn1 (Abcam, U.K.), rabbit polyclonal anti-Sod1 antibody (Abnova, Taiwan), and mouse monoclonal anti-Tuba antibody (Abcam, U.K.) were used as primary antibodies. Molecular Probes Alexa Fluor 647 goat antimouse IgG antibody or Alexa Fluor 647 goat anti-rabbit IgG were used as secondary antibodies. Before imaging, the blots were dried in the dark. The blot membranes were scanned at 50 μm resolution on a Fuji FLA5100 scanner (Fuji Photo, Kanagawa, Japan) with single laser-emitting excitation light at 635 nm. The isoelectric focusing for 2DE-Western blot was performed using 7 cm IPG strips. The second dimension, the blotting and protein detection were carried out as described above. Comparative statistically analyses of Western blot were assessed using student t-test for paired samples using Sigma-Stat-software 2.03 (Jandel Scientific, San Rafael, CA). A P-value of <0.05 was considered to be statistically significant.

Bioinformatics. The classification of the identified proteins according to their main known/postulated functions was carried out using DAVID bioinformatics (www.david.abcc. ncifcrf.gov). This was used, together with the official gene symbol (given in Table 1), to investigate and categorize the gene ontology (GO)-annotations (cellular components, molecular functions, biological processes).

Results

Analysis of the Mouse ESC and maGSC Proteome by 2-DE. First, we evaluated the experimental reproducibility of the 2-DE by labeling four independent samples of maGSCs lysates with Cy3 dye. DIGE was performed as described in the Experimental Section. Analysis of the gels was performed with Delta 2D software. Approximately 500 spots from a well-resolved area of the gel were selected (those spots could be matched on all gels) and the spot-volume-percentage-data were quantified by Delta 2D software. The resulting data were plotted on a logarithmic scale on both axes and correlation coefficients were calculated for the data sets using Sigma Stat version 8. We observed good experimental reproducibility by comparing the different gel images, with correlation coefficients between 0.89 and 0.95, confirming that consistent protein profiles could be obtained from maGSCs using DIGE.

To establish gel-based reference maps of maGSCs and ESCs from mice, we subjected protein extracts from these two cell lines to conventional 2-DE. The performed analyses comprise IEF in immobilized pH gradients (IPG) with broad range (pI 3–10) and overlapping narrow range (i.e., 3–6, 5–8) IPGs in the first dimension and SDS-PAGE in the second dimension. The protein profiles of both cell lines were highly reproducible as revealed by the overlay of four 2-DE replicates, each visualized by Flamingo staining. The spots detected by image analysis in each Flamingo-stained narrow range gel and in the corresponding regions of the broad range gel are presented in Supplemental Figure 1. Around 880 protein spots were detected in the pH 4–9 region of the 3–10 2-D gel for maGSCs whereas about 1800 distinct spots could be detected on the combined narrow range IPG gels when the overlap was excluded so that every protein was only referenced once (Supplemental Figure 1). Analysis of the pI 3–6 2-DE gel revealed the presence of around 600 spots, whereas only 400 could be detected in the same region of the pI 3–10 2-DE gel. For the pI 5–8 2-DE gel, 1100 different spots were resolved and detected compared with 500 in the corresponding region of the pI 3–10 2-DE gel. In summary, narrow range IPG gels allowed us to visualize at least 2-fold more spots than classical broad range IPG gels.

research articles   Dihazi et al.

Table 1. List of Nonredundant Proteins Identified from maGSCs Using Different Mass Spectrometry Techniques and Database Comparisons[a]

#	protein name	gene name	accession	UniProtKB/ Swiss-Prot	Swiss-Prot	nominal mass	CPI	PMF- score	PMF sequence coverage	MS/MS- score	MS/MS- sequence coverage	ESI score	ESI peptides
1	Actin, aortic smooth muscle	Acta2	gi|49968	P62737	ACTA_MOUSE	41982	5.24	123	53	212	12	72	5
6	Actin, cytoplasmic 1	Actb	gi|195277270	P60710	ACTB_MOUSE	42052	5.2	59	49	101	29	378	23
39	1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase	Adi1	gi|110946936	Q99JT9	MTND_MOUSE	21567	5.2	101	58	208	31		
40	Adenylate kinase isoenzyme 1	Ak1	gi|77020262	Q9R0Y5	KAD1_MOUSE	21640	5.6	99	58	151	18		
41	Adenylate kinase 2	Ak2	gi|13046247	Q8C719	KAD2_MOUSE	25817	7.8	105	42	188	13		
42	Aldose reductase	Akr1b1	gi|6753036	P45376	ALDR_MOUSE	36052	6.9	73	25	171	10		
43	Aldehyde dehydrogenase, mitochondrial	Aldh2	gi|6671539	P47738	ALDH2_MOUSE	56502	8.6	212	77	260	19		
46	Fructose-bisphosphate aldolase A	Aldoa	gi|13435747	P05064	ALDOA_MOUSE	39787	9.2	112	50	206	25		
47	Rho GDP-dissociation inhibitor 1	Arhgdia	gi|23272966	Q99PT1	GDIR1_MOUSE	23393	5	227	56	278	10		
51	ATP synthase subunit beta, mitochondrial	Atp5b	gi|116741459	P56480	ATPB_MOUSE	56265	5.1	107	69	196	36		
54	ATP synthase subunit d, mitochondrial	Atp5h	gi|56605979	Q9DCX2	ATP5H_MOUSE	18795	5.4					56	5
56	Transcription factor BTF3	Btf3	gi|3334247	Q64152	BTF3_MOUSE	22017	10	83	51	445	46		
57	Complement component 1 Q subcomponent-binding protein, mitochondrial	C1qbp		O35658	C1QBP_MOUSE	31336	4.7	63	30	143	9		
59	Calmodulin	Calm1	gi|12851157	P62204	CALM_MOUSE	16827	3.9	38	30	211	22		
60	Calreticulin	Calr	gi|6680836	P14211	CALR_MOUSE	48136	4.2	144	36	114	10		
64	Calumenin	Calu	gi|6680840	O35887	CALU_MOUSE	37041	4.3	99	30	207	26		
68	F-actin-capping protein subunit alpha-1	Capza1	gi|595917	P47753	CAZA1_MOUSE	33090	5.2	111	55	174	27	49	3
69	F-actin-capping protein subunit beta	Capzb	gi|74184034	Q3TRH8	CAPZB_MOUSE	31611	5.4	102	54	192	11		
71	Chromobox protein homologue 1	Cbx1	gi|33115171	Q7TPM0	CBX1_MOUSE	21519	4.7	58	31	125	23		
74	Chromobox protein homologue 5	Cbx5	gi|6671694	Q61686	CBX5_MOUSE	22343	5.7			56	6		
75	T-complex protein 1 subunit alpha A	Cct1	gi|74142471	P11983	TCPA2_MOUSE	60411	5.8	261	71	57	4		
76	T-complex protein 1 subunit beta	Cct2	gi|468546	P80314	TCPB_MOUSE	57783	6	159	49	201	10		
77	T-complex protein 1 subunit epsilon	Cct5	gi|6671702	P80316	TCPE_MOUSE	59586	5.7	178	39	135	7		
79	Cofilin-1	Cfl1	gi|6680924	P18760	COF1_MOUSE	18776	9.1	102	57	331	36		
80	Clathrin light chain A	Clta	gi|21312564	O08585	CLCA_MOUSE	25541	4.5					33	2
81	Calponin-3	Cnn3	gi|6681071	Q9DAW9	CNN3_MOUSE	36577	5.4	69	39	43	3		
82	Cystatin-B	Cstb	gi|6753618	Q62426	CYTB_MOUSE	11039	7.8	83	51	127	29		
83	D-dopachrome decarboxylase	Ddt	gi|114916554	O35215	DOPD_MOUSE	13183	6.1	131	88	56	10		
84	Protein dpy-30 homologue	Dpy30	gi|8347092	Q99LT0	DPY30_MOUSE	11206	4.7	30	47	73	36		
85	DUTPase	Dut		Q9JJ44		17404	5.7	36	35	49	19		
86	Elongation factor 1-alpha 1	Eef1a1	gi|556301	P10126	EF1A1_MOUSE	50424	9.7	53	36	150	5		
89	Elongation factor 1-beta	Eef1b	gi|5902663	O70251	EF1B_MOUSE	24849	4.4	91	61	201	23		
90	Eukaryotic translation initiation factor 1A	Eif1a	gi|158631225	Q60872	IF1A_MOUSE	16606	4.9	79	38	27	16	100	3
91	Eukaryotic translation initiation factor 3 subunit G	Eif3g	gi|31980808	Q9Z1D1	EIF3G_MOUSE	35901	5.6	65	34	9	2		
92	Eukaryotic translation initiation factor 4H	Eif4h	gi|4972951	Q9WUK2	IF4H_MOUSE	27381	7.5	46	35	80	10		
96	Eukaryotic translation initiation factor 5A-1	Eif5a	gi|148680528	P63242	IF5A1_MOUSE	17049	4.9	104	58	252	31		
101	Alpha-enolase	Eno1	gi|55491	P17182	ENOA_MOUSE	47111	6.4	192	51	362	15	51	4
110	Beta-enolase	Eno3	gi|6679651	P21550	ENOB_MOUSE	47453	6.4	81	40	85	10	662	32
111	Enhancer of rudimentary homologue	Erh	gi|182993904	P84089	ERH_MOUSE	12422	5.6	57	75	58	18		
112	Electron transfer flavoprotein subunit alpha, mitochondrial	Etfa	gi|473012	P68040	ETFA_MOUSE	35330	9.5	76	53	63	51	41	7
113	Fatty acid-binding protein, heart	Fabp3	gi|6753810	P11404	FABPH_MOUSE	14810	6.1	63	39	66	24		
115	FK506-binding protein 4	Fkbp4	gi|6753882	P30416	FKBP4_MOUSE	51939	5.4	110	46	22	5		
117	Ferritin light chain 1	Ftl1	gi|1120524	P29391	FRIL1_MOUSE	20790	5.6	86	49	107	15		
119	Glyceraldehyde-3-phosphate dehydrogenase	Gapdh	gi|56188	P04797	G3P_MOUSE	36072	9.2	137	55	248	19	119	5
127	Lactoylglutathione lyase	Glo1	gi|19354350	Q9CPU0	LGUL_MOUSE	20967	5.1	196	70	170	20		

Similar Proteoms of maGSCs and ESCs research articles

Table 1. Continued

	protein name	gene name	accession	UniProtKB/ Swiss-Prot	Swiss-Prot	nominal mass	CPI	PMF score	PMF sequence coverage	MS/MS score	MS/MS sequence coverage	ESI score	ESI peptides
128	Glutamate dehydrogenase 1, mitochondrial	Glud1	gi|6680027	P26443	DHE3_MOUSE	61298	8.8	115	30	60	6		
129	Guanine nucleotide-binding protein subunit beta-2-like 1	Gnb2l1	gi|475012	P68040	GBLP_MOUSE	35511	8.9	222	89	258	18		
132	Growth factor receptor-bound protein 2	Grb2	gi|6680083	Q60631	GRB2_MOUSE	25336	5.9	129	60	105	17		
133	Glutathione transferase omega-1	Gsto1	gi|6754090	O09131	GSTO1_MOUSE	27480	7.7	126	54	208	16	72	3
134	Nuclear protein Hcc-1	Hcc1	gi|12858661	Q9D113	HCC1_MOUSE	23518	6.3	104	37	80	10	115	6
135	Histone H2B type 1-P	Hist1h2bp	gi|223096	Q8CGP2	H2B1B_MOUSE	13944	10.8	67	32	93	19	47	7
136	High mobility group protein B2	Hmgb2	gi|6680229	P30681	HMGB2_MOUSE	24318	7.6	72	49	93	12		
138	Heterogeneous nuclear ribonucleoproteins A2/B1	Hnrnpa2b1	gi|3329498	O88569	ROA2_MOUSE	37437	9.3	101	47	238	16		
139	Heterogeneous nuclear ribonucleoprotein A3	Hnrnpa3	gi|31559916	Q8BG05	ROA3_MOUSE	39856	9.6	70	38	98	8		
140	Heterogeneous nuclear ribonucleoprotein A/B	Hnrnpab	gi|6754222	Q99020	ROAA_MOUSE	30812	8.7	82	30	107	12	105	6
143	Heterogeneous nuclear ribonucleoproteins C1/C2	Hnrnpc	gi|12860377	Q9Z204	HNRPC_MOUSE	34421	4.8	57	24	110	10		
147	Heterogeneous nuclear ribonucleoprotein F	Hnrnpf	gi|19527048	Q9Z2X1	HNRPF_MOUSE	45701	5.2	159	52	337	16	164	10
151	Heterogeneous nuclear ribonucleoprotein H	Hnrnph1	gi|10946928	O35737	HNRH1_MOUSE	49454	5.9	184	55	216	14		
153	Heterogeneous nuclear ribonucleoprotein H2/H1	Hnrnph2/1	gi|9845253	P70333	HNRH1_MOUSE	49168	5.9	48	21	111	11		
157	Heterogeneous nuclear ribonucleoprotein K	Hnrnpk	gi|13384620	P61979	HNRPK_MOUSE	51230	5.3	146	36	138	5		
159	Heterogeneous nuclear ribonucleoprotein D-like	Hnrpdl	gi|81917948	Q9Z130	HNRDL_MOUSE	33709	7.5	73	22	139	19		
162	Heat shock 70 kDa protein 4	Hspa4	gi|2495342	Q61316	HSP74_MOUSE	94872	5	21	13	33	4		
163	78 kDa glucose-regulated protein	Hspa5	gi|31981722	P20029	GRP78_MOUSE	72492	4.9	208	40	293	10	155	10
170	Heat shock cognate 71 kDa protein	Hspa8	gi|74220592	P63017	HSP7C_MOUSE	71055	5.2	150	40	393	10		
207	Stress-70 protein, mitochondrial	Hspa9	gi|435838	P38647	GRP75_MOUSE	73483	5.8	243	51	252	10		
208	Heat shock protein beta-1	Hspb1	gi|424143	P14602	HSPB1_MOUSE	23057	6.1	81	39	189	20		
210	60 kDa heat shock protein, mitochondrial	Hspd1	gi|51455	P63038	CH60_MOUSE	61088	5.8	279	67	316	14		
216	60 kDa heat shock protein, mitochondrial	Hspd1	gi|51455	P63038	CH60_MOUSE	61088	5.8	302	65	311	14		
218	10 kDa heat shock protein, mitochondrial	Hspe1	gi|400542	P26772	CH10_MOUSE	10956	9.1	71	54	207	19		
219	Inosine-5'-monophosphate dehydrogenase 2	Impdh2	gi|309413	P24547	IMDH2_MOUSE	55780	7	132	35	211	10		
220	Far upstream element-binding protein 2	Khsrp	gi|163954948	Q3U0V1	FUBP2_MOUSE	76762	7.1	71	20	94	4		
223	LIM and SH3 domain protein 1	Lasp1	gi|6754508	Q61792	LASP1_MOUSE	30374	6.7	64	33	82	14		
226	Galectin-1	Lgals1	gi|52890	P16045	LEG1_MOUSE	14856	5.2	42	25	192	20		
229	Lamin-B1	Lmnb1	gi|26348034	P14733	LMNB1_MOUSE	47461	5	86	26	77	9		
231	Malate dehydrogenase, mitochondrial	Mdh2	gi|89574115	AAA62644	MDHM_MOUSE	36045	8.9	60	41	37	3		
232	macrophage migration inhibitory factor	Mif	gi|694108		MIF_MOUSE	12667	7.7	51	41	138	17		
233	Myosin light polypeptide 6	Myl6	gi|33620739	Q60605	MYL6_MOUSE	17090	4.4	50	35	62	20		
238	Nascent polypeptide-associated complex subunit alpha	Naca	gi|41350312	Q60817	NACA_MOUSE	23370	4.4	56	32	59	18	60	5
239	Nucleolin	Ncl	gi|13529464	P09405	NUCL_MOUSE	76734	4.5	113	27	318	5		
245	Nucleoside diphosphate kinase A	Nme1	gi|387496	P15532	NDKA_MOUSE	17197	7.7	102	58	123	27		
247	Nucleoside diphosphate kinase B	Nme2	gi|6679078	Q01768	NDKB_MOUSE	17466	7.8	121	79	246	41		
250	Purine nucleoside phosphorylase	Np	gi|388921	P23492	PNPH_MOUSE	32541	5.8	171	73	78	13		
252	Nucleophosmin	Npm1	gi|6679108	Q61937	NPM_MOUSE	32711	4.5	105	42	311	22		
255	Nuclear migration protein nudC	Nudc	gi|6754910	O35685	NUDC_MOUSE	38334	5	71	33	99	13		
256	Nuclear pore glycoprotein p62	Nup62	gi|236261	Q63850	NUP62_MOUSE	53336	5.1	36	23	103	10		
259	Protein disulfide-isomerase	P4hb	gi|129729	P09103	PDIA1_MOUSE	57507	4.6	231	50	225	12		
260	Proliferation-associated protein 2G4	Pa2g4	gi|790470	P50580	PA2G4_MOUSE	44013	6.4	69	38	94	11		
261	Polyadenylate-binding protein 1	Pabpc1	gi|129535	P29341	PABP1_MOUSE	70826	10	94	36	95	8		
262	Protein DJ-1	Park7	gi|55741460	Q99LX0	PARK7_MOUSE	20236	6.4	116	61	137	30	37	5
264	Poly(rC)-binding protein 1	Pcbp1	gi|6754994	P60335	PCBP1_MOUSE	37474	6.8	214	73	314	19		
265	Proliferating cell nuclear antigen	Pcna	gi|53600	P17918	PCNA_MOUSE	29108	4.5	92	39	126	9		

research articles   Dihazi et al.

Table 1. Continued

	protein name	gene name	accession	UniProtKB/Swiss-Prot	Swiss-Prot	nominal mass	CPI	PMF-score	PMF-sequence coverage	MS/MS-score	MS/MS-sequence coverage	ESI score	peptides
267	Pyruvate dehydrogenase E1 component subunit beta, mitochondrial	Pdhb	gi|12805431	Q9D051	ODPB_MOUSE	39254	6.5	123	56	231	16		
270	Protein disulfide-isomerase A3	Pdia3	gi|26353794	P27773	PDIA3_MOUSE	57099	5.8	230	52	168	11		
272	Phosphatidylethanolamine-binding protein 1	Pebp1	gi|74222953	P70296	PEBP1_MOUSE	20817	5.1	159	80	252	32	220	9
273	Prefoldin subunit 2	Pfdn2	gi|3212116	O70591	PFD2_MOUSE	16581	6.2	52	37	83	21		
275	Phosphoglycerate mutase 1	Pgam1	gi|8248819	P25113	PGAM1_MOUSE	28814	6.8	155	74	260	21		
276	Phosphoglycerate kinase 1	Pgk1	gi|202423	P09411	PGK1_MOUSE	44921	9	240	73	241	16		
278	Prohibitin	Phb	gi|6679299	P67778	PHB_MOUSE	29859	5.5	85	36	40	11		
279	Pyruvate kinase isozymes M1/M2	Pkm2	gi|551295	P52480	KPYM_MOUSE	58378	7.9	233	68	154	9		
283	Inorganic pyrophosphatase	Ppa1	gi|27754065	Q9D819	IPYR_MOUSE	33102	5.3	108	43	109	20		
286	Peptidyl-prolyl cis-trans isomerase A	Ppia	gi|6679439	P17742	PPIA_MOUSE	18131	9.1	78	56	370	30		
289	Peptidyl-prolyl cis-trans isomerase C	Ppic			PPIC_MOUSE	22780	7.8	49	27	43	5		
292	Peroxiredoxin-1	Prdx1	gi|6754976	P35700	PRDX1_MOUSE	22390	9.2	171	68	225	19		
293	Peroxiredoxin-2	Prdx2	gi|3603241	Q61171	PRDX2_MOUSE	21936	5.1	140	53	319	22		
296	Thioredoxin-dependent peroxide reductase, mitochondrial	Prdx3	gi|6680690	P20108	PRDX3_MOUSE	28337	8	64	32	103	14		
297	Peroxiredoxin-4	Prdx4	gi|7948999	O08807	PRDX4_MOUSE	31033	6.8	101	49	224	20		
298	Prospero homeobox protein 1	Prox1	gi|13386026	P48437	CN166_MOUSE	28135	6.5	93	46	84	18		
299	Proteasome subunit alpha type-1	Psma1	gi|26353732	Q8BTU5	PSA1_MOUSE	29528	6	79	30	158	16		
300	Proteasome subunit alpha type-2	Psma2	gi|1709759	P49722	PSA2_MOUSE	26023	9.1	56	35	93	17		
302	Proteasome subunit alpha type-5	Psma5	gi|7106387	Q9Z2U1	PSA5_MOUSE	26565	4.6	83	50	197	29		
303	Proteasome subunit alpha type-6	Psma6	gi|6755198	Q9QUM9	PSA6_MOUSE	27811	6.4	88	34	191	21		
305	Proteasome subunit beta type-2	Psmb2	gi|9910832	Q9R1P3	PSB2_MOUSE	22892	6.6	68	29	80	17		
306	Proteasome subunit beta type-3	Psmb3	gi|6755202	Q9R1P1	PSB3_MOUSE	23235	6.2	110	51	247	30		
307	Proteasome subunit beta type-4	Psmb4	gi|3914439	P99026	PSB4_MOUSE	29211	5.3	60	48	118	18		
308	Proteasome subunit beta type-6	Psmb6	gi|74204932	Q60692	PSB6_MOUSE	25591	4.8	54	34	47	14		
310	Proteasome subunit beta type-7	Psmb7	gi|6755206	P70195	PSB7_MOUSE	29872	9.1	88	38	55	5		
311	26S protease regulatory subunit 7	Psmc2	gi|33859604	Q8BVQ9	PRS7_MOUSE	49016	5.6	192	55	63	8		
312	26S protease regulatory subunit 6B	Psmc4	gi|1709797	P54775	PRS6B_MOUSE	47252	5	98	24	90	11		
313	26S proteasome non-ATPase regulatory subunit 14	Psmd14	gi|2505940	O35593	PSDE_MOUSE	34726	6.1	97	55	30	3		
314	Proteasome activator complex subunit 3	Psme3	gi|6755214	P61290	PSME3_MOUSE	29602	5.6	110	51	70	17		
315	Prothymosin alpha	Ptma		P26350	PTMA_MOUSE	12247	3.7					44	2
316	Parathymosin	Ptms		Q9D0J8	PTMS_MOUSE	11423	4.1					86	2
317	tRNA pseudouridine synthase A	Pus1		Q9WU56		47472	9					33	3
319	GTP-binding nuclear protein Ran	Ran	gi|74208259	Q3ULW0	RAN_MOUSE	24408	7.8	93	40	286	27		
320	Ran-specific GTPase-activating protein	Ranbp1	gi|431422	P34022	RANG_MOUSE	23753	5	49	26	69	10		
321	Putative RNA-binding protein 3	Rbm3	gi|6093936	O89086	RBM3_MOUSE	16595	7.7	77	37	138	24		
322	60S acidic ribosomal protein P0	Rplp0	gi|13277927	P14869	RLA0_MOUSE	34366	5.9	111	50	147	11		
323	60S acidic ribosomal protein P2	Rplp2	gi|83745120	P99027	RLA2_MOUSE	11644	4.2	87	70	386	66		
326	40S ribosomal protein S12	Rps12	gi|149250091	P63323	RS12_MOUSE	14858	7.7	55	34	74	27	56	5
327	40S ribosomal protein SA	Rpsa	gi|31560560	P14206	RSSA_MOUSE	32931	4.6	90	50	25	10		
329	Septin-2	Sept2	gi|6754816	P42208	SEPT2_MOUSE	41727	6.1	81	39	93	18		
330	Plasminogen activator inhibitor 1 RNA-binding protein	Serbp1	gi|12836024	Q9CY58	PAIRB_MOUSE	44801	9.3	53	27	91	7		
331	Protein SET	Set	gi|3953617	Q9EQU5	SET_MOUSE	33358	4.1	77	22	225	18		
334	Splicing factor, arginine/serine-rich 1	Sfrs1	gi|26383576	Q6PDM2	SFRS1_MOUSE	27842	10.8	116	39	65	17		
337	Splicing factor, arginine/serine-rich 3	Sfrs3	gi|2125864	P84104	SFRS3_MOUSE	19546	12.3	45	27	80	14		
338	Superoxide dismutase [Cu-Zn]	Sod1	gi|201006	P08228	SODC_MOUSE	15933	6	85	38	226	24		

Table 1. Continued

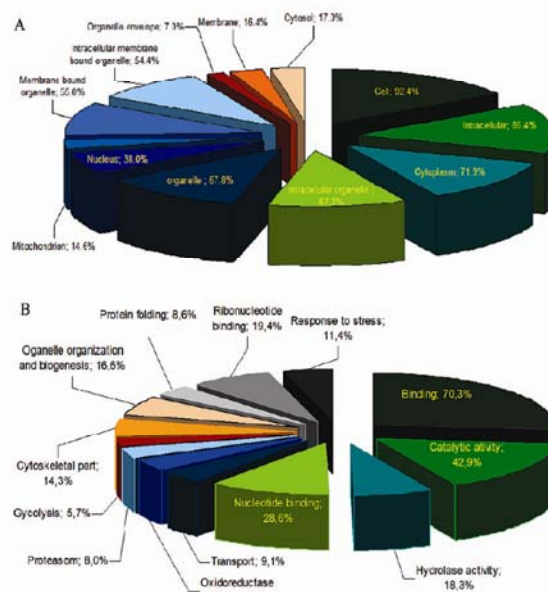|	protein name	gene name	accession	UniProtKB/Swiss-Prot	Swiss-Prot	nominal mass	CPI	PMF-score	PMF sequence coverage	MS/MS-score	MS/MS-sequence coverage	ESI score	peptides
339	Superoxide dismutase [Mn], mitochondrial	Sod2	gi|7390379	P09671	SODM_MOUSE	24816	9.5	114	68	281	30		
340	Spermidine synthase	Srm	gi|6678131	Q64674	SPEE_MOUSE	34543	5.2	191	64	101	10		
341	Hsc70-interacting protein	St13		Q99L47	F10A1_MOUSE	41630	5.2					115	4
342	Stress-induced-phosphoprotein 1	Stip1	gi|13277819	Q60864	STIP1_MOUSE	63170	6.4	70	30	42	5		
344	Stathmin	Stmn1	gi|9789995	P54227	STMN1_MOUSE	17264	5.7	47	30	110	16	124	9
345	Small ubiquitin-related modifier 2	Sumo2	gi|19111164	P61957	SUMO2_MOUSE	10864	5.3	18	20	43	12	77	1
346	Transaldolase	Taldo1	gi|33859640	Q93092	TALDO_MOUSE	37534	6.7	103	30	88	6		
348	TAR DNA-binding protein 43	Tardbp	gi|21704096	Q921F2	TADBP_MOUSE	44918	6.3	102	38	175	10		
350	Transcription elongation factor B polypeptide 2	Tceb2	gi|62869	P62869	ELOB_MOUSE	13162	4.9					48	3
352	Triosephosphate isomerase	Tpi1	gi|1864018	P17751	TPIS_MOUSE	27038	7.7	114	56	239	24		
353	Tropomyosin alpha-1 chain	Tpm1	gi|148634199	P58771	TPM1_MOUSE	32718	4.5	57	18	57	7	93	8
354	Tropomyosin beta chain	Tpm2	gi|50190	P58774	TPM2_MOUSE	32817	4.5	66	26	60	7		
355	Tropomyosin alpha-3 chain	Tpm3	gi|26341416	Q6PDM2	TPM3_MOUSE	28760	4.5	109	41	57	14	87	5
357	Tropomyosin alpha-4 chain	Tpm4	gi|47894398	Q6IRU2	TPM4_MOUSE	28450	4.5	95	36	144	18	237	15
359	Translationally controlled tumor protein	Tpt1	gi|6678437	P63028	TCTP_MOUSE	19450	4.6	36	22	100	22		
367	Transcription intermediary factor 1-beta	Trim28	gi|1621617	Q62318	TIF1B_MOUSE	88791	5.4	42	15	117	5		
368	Tubulin alpha-1A chain	Tuba1a	gi|202210	P05213	TBA1A_MOUSE	50788	4.8	64	26	41	3		
370	Tubulin alpha-1B chain	Tuba1b	gi|202210	P05213	TBA1B_MOUSE	50120	4.8	104	28	276	13		
373	Tubulin alpha-1C chain	Tuba1c	gi|6678469	P68368	TBA1C_MOUSE	49877	4.8	118	32	274	12		
374	Tubulin beta-2C chain	Tubb2c	gi|5174735	Q9CWF2	TBB2C_MOUSE	50255	4.6	146	56	128	11		
382	Tubulin beta-5 chain	Tubb5	gi|7106439	P99024	TBB5_MOUSE	49639	4.6	140	29	199	10		
384	Thioredoxin	Txn	gi|6755911	P10639	THIO_MOUSE	11668	4.6	54	42	103	20	106	5
386	Thioredoxin-like protein 1	Txnl1	gi|74181480	Q3UBF4		32217	4.8					137	9
387	Ubiquitin carboxyl-terminal hydrolase isozyme L1	Uchl1			UCHL1_MOUSE	24822	5	33	23	31	17		
388	Cytochrome b-c1 complex subunit 1, mitochondrial	Uqcrc1	gi|14548301	Q9CZ13	QCR1_MOUSE	53420	5.7	58	24	30	3		
389	Prefoldin subunit 3	Vbp1	gi|61759	P61759	PFD3_MOUSE	22592	6	65	37	29	6		
390	voltage-dependent anion channel 1	Vdac1	gi|6755963	Q60932	VDAC1_MOUSE	32502	9.2	179	77	136	27		
391	Vimentin	Vim	gi|55291	P20152	VIME_MOUSE	53712	4.9	126	47	259	14	245	18
408	14-3-3 protein gamma	Ywhag	gi|3065929	P61982	1433G_MOUSE	28456	4.7	62	36	122	21		
409	14-3-3 protein theta	Ywhaq			1433T_MOUSE	28046	4.7	135	52		4		

[a] The accession number in Swiss-Prot and NCBI are given. The PMF and MS/MS information are also given. The numbering was performed according to the Supplemental Table 1.

research articles   Dihazi et al.
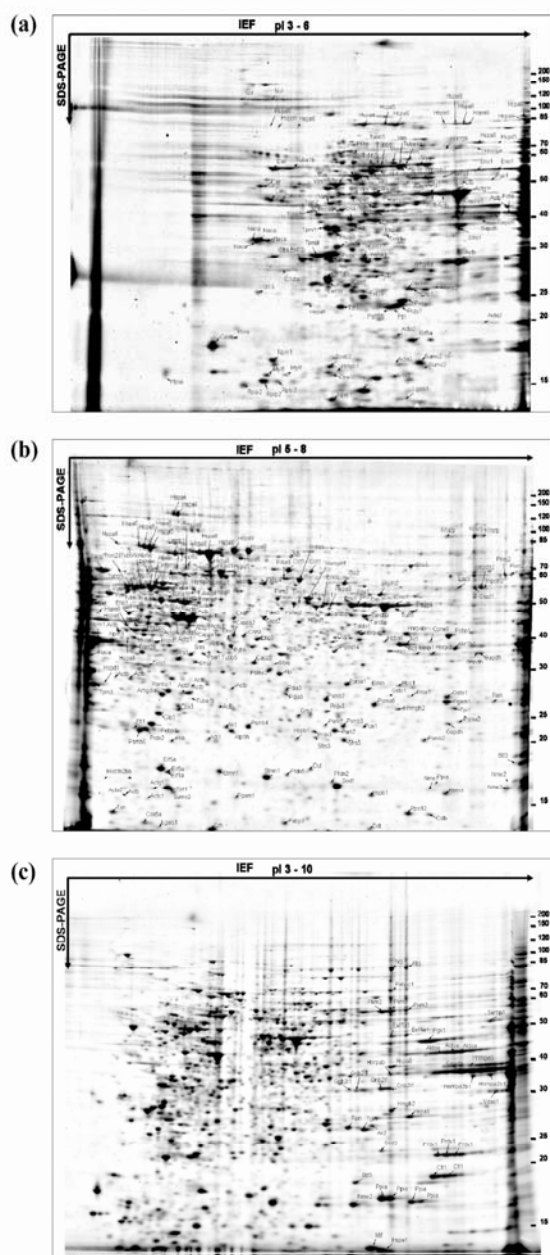
Figure 1. 2-DE reference maps of maGSCs protein extracts (mouse line 129/SV). A total of 150 μg of protein was loaded on an 11-cm IPG strip with a linear pH gradient pI 3–6 (a), pI 5–8 (b) and 3–10 (c) for isoelectric focusing, and a 12% SDS-polyacrylamide gel was used for the SDS-PAGE. Proteins were stained with Flamingo fluorescent gel stain. Identified spots were assigned a gene name.

Figure 2. Bioinformatic analysis of the identified proteins. The list of genes to be analyzed was uploaded into Gene list Manager Window of DAVID bioinformatics and the species info was added. The functional annotations associated with groups of genes or with each individual gene were then displayed in chart. The Chart Report is an annotation-term-focused view with the genes under study. To avoid over counting duplicated genes, the Fisher Exact statistics is calculated based on corresponding DAVID gene IDs by which all redundancy in original IDs is removed. All results of Chart Report has to pass the thresholds (by default, Max-.Prob.≤0.1 and Min.Count ≥ 2) in Chart Option section to ensure only statistically significant once displayed. (A) GO analysis of cellular components, top 12 categories with the highest hits for cellular locations are shown in form of a pie chart. (B) GO analysis of molecular function, the top 13 with the highest hits for molecular functions are displayed in pie chart.

Gel plugs (protein spots) from different pI ranges were excised and subjected to tryptic in-gel digestion. For high confidence protein identification, we used two complementary mass spectrometric approaches, MALDI-TOF/TOF-MS/MS[21] and online LC-MS/MS.[25] This allowed the identification of 409 proteins from maGSCs and ESCs (Supplemental Table 1), of which 128 were identified in pI range 3–6 (Figure 1a), 233 in pI range 5–8 (Figure 1b), and 43 in pI range 8–10 (Figure 1C). Probably because of alternative splicing and post-translational

modifications, 57 proteins were identified in more than one spot, a finding typical for 2-DE. Thus, the 409 proteins identified corresponded to a library of 166 nonredundant maGSCs- and ESCs-associated proteins (Table 1).

Bioinformatic Analysis of the Identified Proteins. To gain more information on the biological mechanisms associated with the identified proteins, we combined DAVID bioinformatics with information on the putative function of the protein found in the UniProt and GenBank databases. Thereby, we were able to annotate all 166 gene products. The tissue distribution analysis of the identified proteins classified the proteins in a total of 24 different tissue types. The majority of the proteins were classified as bone marrow proteins. An analysis of subcellular distribution of the maGSCs proteins based on the Gene Ontology (GO) terms allowed the differentiation of 12 different categories (Figure 2A); the majority of proteins were intracellular (85.4%) with the nuclear genes representing the most abundant part (38%). The rest of the terms were cytoplasmic, intracellular organelle, mitochondria, membrane bound organelles, intracellular membrane bound organelles, organelle envelopes, membranes and cytosol. The percentage of the GO annotation often exceed 100; this is because of the multifunction or subcellular localization of some proteins resulting in counting these proteins in different GO categories. The classification of the gene products based on

Table 2. List of Proteins Found to Be Present in Different Amount in maGSCs and ESCs from One Mouse Strain When Compared to Their Counterparts from the Second Mouse Strain[a]

gene name	maGSC 129SV, ESC R1	maGSC Stra8, ESC Stra8	function
Khsrp		2↑	Binds to the dendritic targeting element and may play a role in mRNA trafficking
Stip1		1.8↑	Mediates the association of the molecular chaperones HSC70 and HSP90
Hspb1		2↑	Inhibits HSPA1A chaperone activity, interferes with ubiquitination and inhibits chaperone-assisted degradation
Krt8		2.5↑	Cytoskeleton
Ncl	2.4↑		synthesis and maturation of ribosomes
Trim28	1.8↑		Forms a complex with a KRAB-domain transcription factor and increases the efficiency of KRAB-mediated repression. Silences transcription through an interaction with HP1 proteins. Acts as a corepressor of transcription for the KRAB zinc finger proteins and as a moderator of the repression activity.
Hnrnpk	1.9↑		One of the major pre-mRNA-binding proteins. Binds tenaciously to poly(C) sequences. Likely to play a role in the nuclear metabolism of hnRNAs, particularly for pre-mRNAs that contain cytidine-rich sequences.
Prdx6	2.2↑		Involved in redox regulation of the cell
Nme2		2.3↑	Acts as a transcriptional activator of the c-Myc gene; binds DNA nonspecifically
Sod1		2↑	Destroys radicals which are normally produced within the cells and which are toxic to biological systems
Rps12		1.6↑	Ribosomal protein

[a] The arrows (↑) indicate higher level of expression; maGSC 129SV/ESC R1 are from mouse line 129/SV and maGSC Stra8/ESC Stra8 are from mouse line Stra8-EGFP/ROSA26.

their molecular function resulted in 13 different categories (Figure 2B), where the binding functions with 70.3% represents the largest group. The proteins were classified in the following categories: binding, catalytic activity, hydrolase activity, nucleotide binding, transport, oxidoreductase activity, proteasom, glycolysis, cytoskeletal part, organelle organization and biogenesis, protein folding, ribonucleotide binding and response to stress.

Proteomic Analysis of the Interindividual Differences of maGSCs and ESCs. To investigate possible interindividual differences of the stem cell types under study, we first performed a DIGE analysis of maGSCs cells originating from two different mouse lines, 129SV and Stra8-EGFP/ROSA26. This experiment led to the identification of 402 proteins and, more importantly, revealed that the proteomes of the maGSCs from the two cell lines were almost identical as only 11 proteins were identified as differentially expressed (Table 2). This finding indicates a conserved maGSC proteome in the species mouse. Similar results with respect to interindividual differences were obtained when the ESCs from the respective mouse lines were compared in the same way (Table 2).

Differential Proteomic Analysis of maGSCs and ESCs. Recent analysis of microRNA profiles in maGSCs and ESCs confirmed that maGSCs also contain miRNAs from the 290 and 302 miRNA groups, which have been previously classified as ESC specific. This finding already indicated an ESC-like nature of maGSCs.[5] To fortify the ESC-like nature of maGSCs and to identify potentially proteins of altered abundance, differential proteomic analysis of the two stem cell types of cells was performed. For this purpose, we used the 2D-DIGE technology to compare the proteomes of maGSCs and ESCs from the same mouse strain, thereby excluding interindividual differences. The protein profiles of maGSCs and ESCs were found to be highly similar (Figure 3a,b) with a quantitative and qualitative match of over 96%. We identified only 18 proteins which were differentially expressed in the two cell types: Eif5a (two out of three isoforms), Gapdh, Glo1, Lgals1, Psma6, Tardbp were present in greater levels in maGSCs compared to ESCs, whereas Atp5h, Cstb, Eif5a (1), Hnrnpab, Khsrp, Park7, Prdx2, Psmb4,

Ranbp, Rps12, Stip1, and Stmn1 were present in greater levels in ESCs compared to maGSCs (Table 3). Eif5a is a protein involved in protein biosynthesis. We identified three forms of Eif5a which we labeled Eif5a(1) Eif5a(2) and Eif5a(3), the latter being equally abundant in both cell types. In contrast, Eif5a(1) and Eif5a(2), which have similar isoelectric points but differ in molecular mass, were found to be differentially expressed. Eif5a(1) was found to be present in lower level in maGSCs compared to ESCs, whereas the Eif5a(2) was present in greater levels in maGSCs compared to ESCs (Figure 3c). A majority of the proteins that differ between the two cell line types are proteins involved in protein biosynthesis or degradation, possibly indicating a difference in protein turnover in the two cell types. Moreover, the DAVID Bioinformatics clustering of these proteins showed that Eif5a, Prdx2, Glo1, Lgals1, Hnrnpab, and Stmn1 are involved in cell differentiation processes.

Western Blot Analysis of Proteins Present in Different Levels in maGSCs and ESCs. A selection of the proteins which were found to differ in abundance in the maGSCs and ESCs were validated with fluorescent Western blot detection. In this experiment, maGSCs isolated from four different mouse lines (Stra8-EGFP/ROSA26, 129/Sv, FVB, and C57BL/6) were tested together with their corresponding ESCs. Fluorescent Western blotting using one-step fluorescence imaging of Alexa Fluor Dyes-conjugated antibodies offers a high signal stability and a wide linear dynamic range. Therefore, it allows for accurate quantification which can effectively complement DIGE data. Our results showed that the differences in protein expression between maGSCs and ESCs depend on the cell origin. For example, the higher levels of Tardbp in maGSCs seen in the DIGE analysis were found to hold true only for cells which originated from Stra8-EGFP/ROSA26 and 129/Sv mouse lines, whereas cells from FVB mice C57BL/6 mice did not exhibit any significant differences in Tardbp expression (Figure 4A,B). Similar results were also observed for Stmn1 and Sod1 (Figure 4A,B). Likewise, both 1D and 2D Western blot detection of Eif5a revealed that the expression differences between maGSCs and ESCs are mouse line-dependent. The different forms of Eif5a found in 2-DE gels could be confirmed only for the cells derived
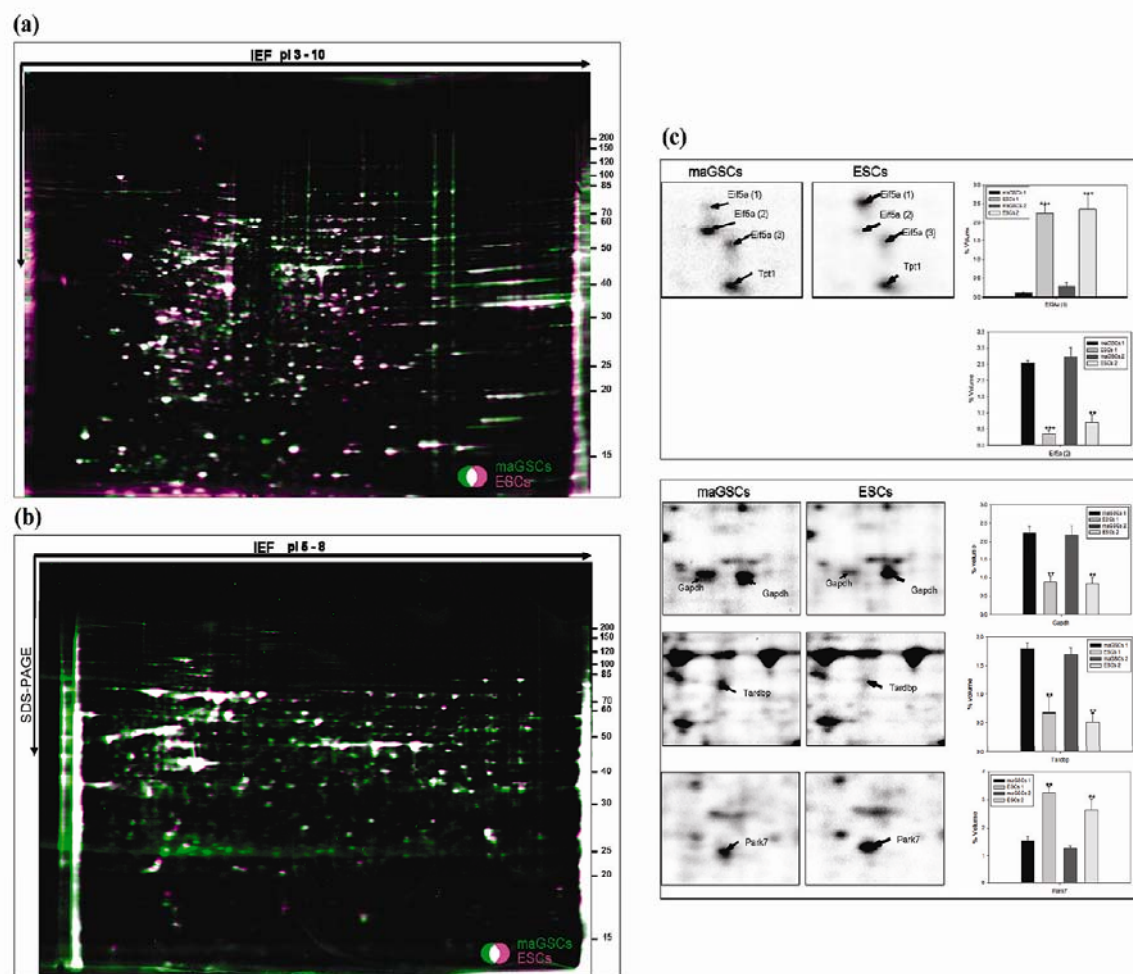
Figure 3. Differential proteome analysis of the maGSCs and ESCs. (a) Overlap of 2-DE maps of maGSCs and ESCs proteomes from the mouse line Stra8-EGFP/ROSA26 pI 3–10. (b) After the labeling step (50 μg of protein were labeled with 400 pmol of Cy3 or Cy5, the internal standard was labeled with Cy2), the samples were subjected to DIGE; 2D-DIGE map of maGSCs and ESCs proteomes (mouse line 129/SV) are shown: Cy3-labeled proteins are false colored in green (representing maGSCs proteome) and Cy5-labeled proteins are false colored in purple (representing ESCs proteome). (c) Graphics represent separation of the detection channels and enlargement of the gel regions of interest showing protein spots found to be differentially expressed. The protein expression quantification for selected proteins is given in form of bar diagrams. On the y-axis, the spot-volume-percentage is given and the x-axis shows the corresponding gene name. Labeling of the graphics corresponds to the gene names listed in Table 1. The quantification was performed for the same proteins in cell lines deriving from two different mice. maGSCs 1 and ESCs 1 are from the mouse line 129/SV and maGSCs 2 and ESCs 2 are from the mouse line Stra8-EGFP/ROSA26. Statistical analysis was performed by Detla 2D software (Decodon AG) with the following parameters: Student's t-test was performed, and statistical significance was assumed for p-values less than 0.01: (**)$p < 0.01$, (***)$p < 0.001$.

from Stra8-EGFP/ROSA26 and 129/Sv mouse lines. Here, the 2D Western blot carried out with a mixture of cell extracts of maGSCs and ESCs from the same mice showed two distinct spots (Figure 4Cii). In contrast, only one major spot corresponding to Eif5a was observed when mixed extracts from FVB and C57BL/6 mouse lines were analyzed (Figure 4Cii).

Discussion

Embryonic germ cells as well as germline stem cells from neonatal mouse testis have been described to behave like ESCs and are thus capable of differentiating into various other cell types.[4,27] Recently three groups isolated and characterized spermatogonial stem cells from adult testis.[7,28,29] We used spermatogonia specific promoter Stra8 to isolate the cells from adult testis. The isolated and cultured cells named multipotent

adult germline stem cells (maGSCs)[5] respond to culture conditions, acquire pluripotent embryonic stem cell properties and differentiate into derivatives of the three embryonic germ layers. Seandel et al. confirmed the pluripotency of spermatogonial stem cells and isolated the cells using a specific surface marker, GPR125. GPR125 is exclusively expressed in testicular spermatogonia and is lacking in differentiated germ cells.[7] The GPR125+ cells were termed spermatogonial progenitor cells; they could differentiate into contractile cardiac tissue and form functional blood vessels in vitro.[7]

The therapeutic interest in pluripotent adult spermatogonial stem cells raises the understanding of the mechanisms underlying pluripotency to a fundamental issue. In this study, we characterized the maGSCs proteome and provide for the first time 2-DE reference map of maGSCs and ESCs from the same

Table 3. List of Proteins Found to Be Present in Different Amount in maGSCs When Compared to ESCs[a]

gene name	maGSCs	ESCs	function
Atp5h		2↑	Mitochondrial membrane ATP synthase
Cstb		2↑	Intracellular thiol proteinase inhibitor
Eif5a (1)		8↑	protein biosynthesis
Eif5a (2)	7↑		protein biosynthesis
Gapdh	2.5↑		carbohydrate metabolism also involved in membrane trafficking in the early secretory pathway
Glo1	2.8↑		prevent oxidative stress and cell aging
Hnrnpab		1.8↑	Binds single-stranded RNA
Khsrp		2↑	Binds to the dendritic targeting element and may play a role in mRNA trafficking
Lgals1	1.6↑		Regulate cell apoptosis and cell differentiation
Park7		2↑	Redox-sensitive chaperone and a sensor for oxidative stress. cell-growth and transformation
Prdx2		1.8↑	Involved in redox regulation of the cell
Psma6	2.3↑		Part of proteasom, protein degradation part
Psmb4		2.5↑	Part of proteasom, protein degradation part
Ranbp1		1.8↑	Protein degradation, an E3 ubiquitin-protein ligase
Rps12		1.8↑	protein synthesis
Stip1		2.2↑	Mediates the association of the molecular chaperones HSC70 and HSP90
Stmn1		2.5↑	cell proliferation and differentiation
Tardbp	2.7↑		DNA and RNA-binding protein which regulates transcription and splicing

[a] The arrows (↑) indicate higher level of expression; maGSCs and ESCs are from mouse line 129/Sv.
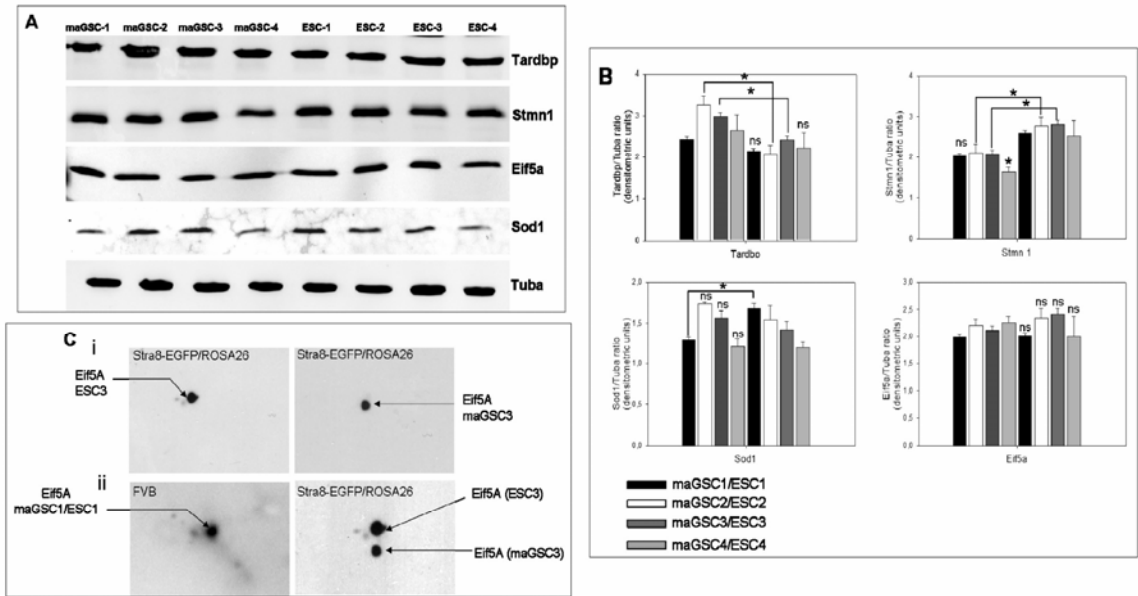
Figure 4. Fluorescent Western blot analyses of proteins found to be differentially expressed between maGSCs and ESCs. (A) The protein extraction and estimation were carried out as described in the Experimental Section. The Western blots were performed with antibodies against the Eif5a, Stmn1, Tardbp, Sod1, and Tuba in cell extracts from cells originating from four different mice lines: maGSC1/ESC1 from FVB mouse line, maGSC2/ESC2 from C57BL/6 mouse line, maGSC3/ESC3 and maGSC4/ESC4 from Stra8-EGFP/ROSA26 and 129/Sv mice lines, respectively. (B) Western blot quantification was performed by densitometry and is represented in form of histograms. The bar charts are represented as the ratio (in densitometric units) of the corresponding protein and tubulin (Tuba) as a loading control. Comparative statistic analyses of Western blots were assessed using *t* test for paired samples, *: *P* < 0.05. ns: non significant. (C) 2D Western blots were performed with anti-Eif5a antibody. The cell extract preparation from maGSCs and ESCs and the 2-D Western blot using 7 cm IPG strips were performed as described in Experimental Section. (i) The maGSC3 and ESC3 protein extracts were separated in different IPG strips and the Eif5a was detected using anit-Eif5a monoclonal antibody. (ii) Equal protein amounts from maGSCs and ESCs cell extracts were mixed and the focusing was carried out on the same IPG strip. The second dimension was performed on the same gel. The Eif5a was then detected using rabit monoclonal Eif5a antibody. (ii) Left panel, maGSC1/ESC1 from FVB mouse line; (ii) right panel, maGSC3/ESC3 from Stra8-EGFP/ROSA26 mouse line.

mouse strain. We have separated the stem cell-derived proteins by narrow range 2-DE to facilitate the detection of low-abundance proteins and combined this approach with mass spectrometric protein identification. By combining narrow range 2-DE with the DIGE technology, we were able to show

that only a very small subset of proteins is differentially expressed in ESCs and maGSCs, thereby substantiating the ESC-like nature of maGSCs.

By our gel-based approach, we identified 409 proteins from maGSCs and ESCs, representing a library of 166 nonredundant

research articles   Dihazi et al.
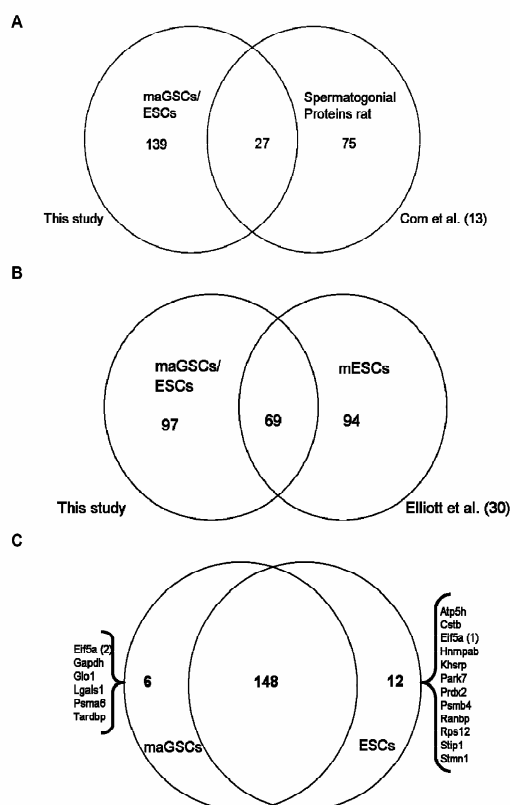
Figure 5. Comparison of maGSCs proteome with mESCs and spermatogonia proteome libraries. (A) Comparison of the maG-SCs associated proteins identified in this study with the spermatogonia associated proteins identified previously;[13] (B) Venn diagram comparing the number of maGSCs associated proteins identified by MS after separation by 2-DE in the present study and the mESCSs associated proteins reported previously based on the same technique;[30] (C) comparison of maGSCs and ESCs proteomes from this study. Both cell lines were from the same mouse (Line 129/Sv).

stem-cell associated proteins. When we compared our compendium to other protein libraries that were established for ESCs (30) or spermatogonia (13) by using the same methods (2-DE and MS), we found that 27 of the 166 proteins identified from maGSCs were reported to be also spermatogonia-associated proteins in rat,[13] whereas 139 were only identified in maGSCs (Figure 5A). A relatively low overlap was also found with a mouse ESC proteome published earlier as 69 of the 166 proteins identified from maGSCs were reported to be associated with ESCs[30] (Figure 5B). In contrast, our data revealed an overlap of more than 96% between the proteomes of ESCs and maGSCs (Figure 5C). This high overlap may result from the fact that the cells used in our study were established from the same mouse strain and have been cultured for almost identical time periods before proteome analysis.

Protein annotation using DAVID bioinformatics substantiates the high similarity between the maGSCS and ESCs. The majority of the identified proteins in both cell types are intracellular with the nucleus (38%) being the most represented organelle. This observation is consistent with the morphology of the stem cells, which grow as compact colonies with a high nucleus-to-cytoplasm ratio and prominent nucleoli. The resemblance between the two cell types is further reflected by

the protein functional distribution as the majority of the 166 identified hits are classified as binding proteins (70.3%). Particularly nucleotide binding proteins are well represented (28.6%), and within this group, the majority is formed by the ribonucleotide binding proteins (19.4%). The latter are mainly represented by heterogeneous nuclear ribonucleoproteins (Hnrnp) with their different identified forms Hnrnpab, Hnrnpa3, Hnrpdl, Hnrnpf, Hnrnph1, Hnrnph2, Hnrnpk, Hnrnpa2b1, Hnrnpc. Hnrnp are a family of ubiquitously expressed ribonucleoproteins that were originally found as proteins bound to nascent RNA transcripts in the form of ribonucleoprotein particles. These proteins are associated with pre-mRNAs in the nucleus and appear to influence pre-mRNA processing and other aspects of mRNA metabolism, stability and transport. They are also thought to have a role during cell cycle progression.[31,32] Their high abundance in maGSCs may reflect the high protein turnover present in these cells. Heat shock cognate 71 kDa protein (Hspa8) was also found to be highly expressed and present in different forms in maGSCs and ESCs as reflected by spots at different pI ranges. Hspa8 belongs to the heat shock protein 70 family, which contains heat shock inducible and the so-called heat shock cognate proteins. Hspa8 is a member of the latter group; it has an ATPase activity and binds to nascent polypeptides to facilitate their proper folding. Hspa8 was found to be expressed on the surface of human ESCs and its expression was altered upon differentiation.[33,34] It was also suggested to be a cell-surface marker for undifferentiated human ESCs. The role of the Hspa8 in maGSCs remains unclear and should be a matter of further investigations, but their presence in different forms in both cell types support the ESC-like nature of maGSCs. An important number of the identified proteins are involved in defending the cells against oxidative stress and cell injury. Oxidative stress may lead to increased damage of the genome[35] explaining the necessity of protecting mechanisms and the expression of proteins involved in oxidative stress defense in maGSCs.

Our differential proteomic analysis revealed only 18 proteins that were differentially expressed in maGSCs and ESCs and we could not identify proteins that were exclusively present in only one cell type. Eif5a (2), Gapdh, Glo1, Lgals1, Psma6, Tardbp were found to be present with greater levels in maGSCs compared to ESCs. Three of these proteins (Eif5a (2), Glo1, Lgals1) are involved in cell differentiation. In case of Eif5a, we identified three isoforms, of which (1) and (2) were found to be differentially expressed in the investigated cells. The precise role of Eif5a in protein biosynthesis is not known, but it functions by promoting the formation of the first peptide bond. Eif5a is also known to be essential for cell proliferation and it seems to be the only eukaryotic protein which is subjected to a type of post-translational modification called hypusination.[36] During hypusination, one of the Eif5a lysine residues is modified by the addition of a butylamino group from spermidine to form hypusine, which is an unusual amino acid found in all eukaryotes and in some archaea, but not in bacteria.[37] Hypusine is essential to the function of Eif5a. Thus, hypusine and Eif5a appear to be important for the viability and proliferation of eukaryotic cells.[38] With respect to the electrophoretic behavior of the different Eif5a species in the 2D-gels of both maGSCs and ESCs, the Western blot analyses performed in maGSCs and ESCs from four different mouse lines could in part confirm the 2D gel finding. Moreover, the data showed that the Eif5a spot patterns found for maGSCs and ESCs are mouse line dependent. It needs to be established whether the

Similar Proteoms of maGSCs and ESCs

<div style="columns">

characteristic spot pattern in some way is related to hypusi-nation. Tardbp is a DNA- and RNA-binding protein which regulates transcription and splicing.[39] Tardbp may also be involved in microRNA biogenesis, apoptosis and cell division.[40] Of the proteins found to be present in lower levels in maGSCs compared to ESCs, Eif5a (1), Hnrnpab, Stmn1 and Prdx2 are involved in cell differentiation as assigned by bioinformatic analysis. As to Park7, the product of this gene (protein DJ-1) may function as a redox-sensitive chaperone and as a sensor for oxidative stress, and has been related to male fertility and Parkinson disease.[41]

So far, several studies have shown that maGSCs and ESCs share a high similarity in pluripotency.[5-7,28,42,43] In recent studies, we have confirmed the ESC-like nature of maGSCs[6] on the microRNA level and further demonstrated that these cell types have comparative methylation profiles and telom-erase biology.[16] Our proteome data presented in this study support the potential of maGSCs as an alternative source of pluripotent cells from nonembryonic tissues. We found almost identical proteomes for maGSCs and ESCs (more than 96% overlap), and among the small subset of 18 differentially expressed proteins, none was exclusively present in only one of the cell types. Western blot validation showed that the differences in protein expression cannot be generalized as they were found to be dependent on the mouse line from which the stem cells were derived. In conclusion, our data strongly support the high similarity between maGSCs and ESCs and suggest the use of maGSCs as alternative to overcome ethical issues regarding the usage of embryos.

Supporting Information Available: Proteins identified from maGSCs and ESCs. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Weissman, I. L. Translating stem and progenitor cell biology to the clinic: barriers and opportunities. *Science* **2000**, *287* (5457), 1442–6.

(2) Prelle, K.; Zink, N.; Wolf, E. Pluripotent stem cells--model of embryonic development, tool for gene targeting, and basis of cell therapy. *Anat. Histol. Embryol.* **2002**, *31* (3), 169–86.

(3) Gepstein, L. Derivation and potential applications of human embryonic stem cells. *Circ. Res.* **2002**, *91* (10), 866–76.

(4) 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. Generation of pluripotent stem cells from neonatal mouse testis. *Cell* **2004**, *119* (7), 1001–12.

(5) Guan, K.; Nayernia, K.; Maier, L. S.; Wagner, S.; Dressel, R.; Lee, J. H.; Nolte, J.; Wolf, F.; Li, M.; Engel, W.; Hasenfuss, G. Pluripo-tency of spermatogonial stem cells from adult mouse testis. *Nature* **2006**, *440* (7088), 1199–203.

(6) Zovoilis, A.; Nolte, J.; Drusenheimer, N.; Zechner, U.; Hada, H.; Guan, K.; Hasenfuss, G.; Nayernia, K.; Engel, W. Multipotent adult germline stem cells and embryonic stem cells have similar microRNA profiles. *Mol. Hum. Reprod.* **2008**, *14* (9), 521–9.

(7) Seandel, M.; James, D.; Shmelkov, S. V.; Falciatori, I.; Kim, J.; Chavala, S.; Scherr, D. S.; Zhang, F.; Torres, R.; Gale, N. W.; Yancopoulos, G. D.; Murphy, A.; Valenzuela, D. M.; Hobbs, R. M.; Pandolfi, P. P.; Rafii, S. Generation of functional multipotent adult stem cells from GPR125+ germline progenitors. *Nature* **2007**, *449* (7160), 346–50.

(8) Conrad, S.; Renninger, M.; Hennenlotter, J.; Wiesner, T.; Just, L.; Bonin, M.; Aicher, W.; Buhring, H. J.; Mattheus, U.; Mack, A.; Wagner, H. J.; Minger, S.; Matzkies, M.; Reppel, M.; Hescheler, J.;

<div style="column 2">

Sievert, K. D.; Stenzl, A.; Skutella, T. Generation of pluripotent stem cells from adult human testis. *Nature* **2008**, *456* (7220), 344–9.

(9) Maltman, D. J.; Przyborski, S. P. Can large-scale analysis of the proteome identify effective new markers for embryonic stem cells. *Regen. Med.* **2007**, *2* (4), 465–9.

(10) Cox, J.; Mann, M. Is proteomics the new genomics. *Cell* **2007**, *130* (3), 395–8.

(11) Wildgruber, R.; Harder, A.; Obermaier, C.; Boguth, G.; Weiss, W.; Fey, S. J.; Larsen, P. M.; Gorg, A. Towards higher resolution: two-dimensional electrophoresis of Saccharomyces cerevisiae proteins using overlapping narrow immobilized pH gradients. *Electrophore-sis* **2000**, *21* (13), 2610–6.

(12) Hoving, S.; Voshol, H.; van Oostrum, J. Towards high performance two-dimensional gel electrophoresis using ultrazoom gels. *Elec-trophoresis* **2000**, *21* (13), 2617–21.

(13) Com, E.; Evrard, B.; Roepstorff, P.; Aubry, F.; Pineau, C. New insights into the rat spermatogonial proteome: identification of 156 additional proteins. *Mol. Cell. Proteomics* **2003**, *2* (4), 248–61.

(14) Nagy, A.; Rossant, J.; Nagy, R.; Abramow-Newerly, W.; Roder, J. C. Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90* (18), 8424–8.

(15) Cheng, J.; Dutra, A.; Takesono, A.; Garrett-Beal, L.; Schwartzberg, P. L. Improved generation of C57BL/6J mouse embryonic stem cells in a defined serum-free media. *Genesis* **2004**, *39* (2), 100–4.

(16) Zechner, U.; Nolte, J.; Wolf, M.; Shirneshan, K.; Hajj, N. E.; Weise, D.; Kaltwasser, B.; Zovoilis, A.; Haaf, T.; Engel, W. Comparative methylation profiles and telomerase biology of mouse multipotent adult germline stem cells and embryonic stem cells. *Mol. Hum. Reprod.* **2009**, *15* (6), 345–53.

(17) Dihazi, H.; Kessler, R.; Muller, G. A.; Eschrich, K. Lysine 3 acetylation regulates the phosphorylation of yeast 6-phosphof-ructo-2-kinase under hypo-osmotic stress. *Biol. Chem.* **2005**, *386* (9), 895–900.

(18) Wessel, D.; Flugge, U. I. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal. Biochem.* **1984**, *138* (1), 141–3.

(19) Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976**, *72*, 248–54.

(20) Karp, N. A.; Griffin, J. L.; Lilley, K. S. Application of partial least squares discriminant analysis to two-dimensional difference gel studies in expression proteomics. *Proteomics* **2005**, *5* (1), 81–90.

(21) Jahn, O.; Hesse, D.; Reinelt, M.; Kratzin, H. D. Technical innova-tions for the automated identification of gel-separated proteins by MALDI-TOF mass spectrometry. *Anal. Bioanal. Chem.* **2006**, *386* (1), 92–103.

(22) Werner, H. B.; Kuhlmann, K.; Shen, S.; Uecker, M.; Schardt, A.; Dimova, K.; Orfaniotou, F.; Dhaunchak, A.; Brinkmann, B. G.; Mobius, W.; Guarente, L.; Casaccia-Bonnefil, P.; Jahn, O.; Nave, K. A. Proteolipid protein is required for transport of sirtuin 2 into CNS myelin. *J. Neurosci.* **2007**, *27* (29), 7717–30.

(23) Reumann, S.; Babujee, L.; Ma, C.; Wienkoop, S.; Siemsen, T.; Antonicelli, G. E.; Rasche, N.; Luder, F.; Weckwerth, W.; Jahn, O. Proteome analysis of Arabidopsis leaf peroxisomes reveals novel targeting peptides, metabolic pathways, and defense mechanisms. *Plant Cell* **2007**, *19* (10), 3170–93.

(24) Dihazi, H.; Kessler, R.; Eschrich, K. High osmolarity glycerol (HOG) pathway-induced phosphorylation and activation of 6-phospho-fructo-2-kinase are essential for glycerol accumulation and yeast cell proliferation under hyperosmotic stress. *J. Biol. Chem.* **2004**, *279* (23), 23961–8.

(25) Dihazi, H.; Asif, A. R.; Agarwal, N. K.; Doncheva, Y.; Muller, G. A. Proteomic analysis of cellular response to osmotic stress in thick ascending limb of Henle's loop (TALH) cells. *Mol. Cell. Proteomics* **2005**, *4* (10), 1445–58.

(26) Towbin, H.; Staehelin, T.; Gordon, J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: pro-cedure and some applications. 1979. *Biotechnology* **1992**, *24*, 145–149.

(27) Matsui, Y.; Zsebo, K.; Hogan, B. L. Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. *Cell* **1992**, *70* (5), 841–7.

(28) Kanatsu-Shinohara, M.; Muneto, T.; Lee, J.; Takenaka, M.; Chuma, S.; Nakatsuji, N.; Horiuchi, T.; Shinohara, T. Long-term culture of male germline stem cells from hamster testes. *Biol. Reprod.* **2008**, *78* (4), 611–7.

(29) Izadyar, F.; Pau, F.; Marh, J.; Slepko, N.; Wang, T.; Gonzalez, R.; Ramos, T.; Howerton, K.; Sayre, C.; Silva, F. Generation of multipotent cell lines from a distinct population of male germ line stem cells. *Reproduction* **2008**, *135* (6), 771–84.

research articles   Dihazi et al.

(30) Elliott, S. T.; Crider, D. G.; Garnham, C. P.; Boheler, K. R.; Van Eyk, J. E. Two-dimensional gel electrophoresis database of murine R1 embryonic stem cells. *Proteomics* **2004**, *4* (12), 3813–32.

(31) Chou, M. Y.; Rooke, N.; Turck, C. W.; Black, D. L. hnRNP H is a component of a splicing enhancer complex that activates a c-src alternative exon in neuronal cells. *Mol. Cell. Biol.* **1999**, *19* (1), 69–77.

(32) Choi, Y. D.; Grabowski, P. J.; Sharp, P. A.; Dreyfuss, G. Heterogeneous nuclear ribonucleoproteins: role in RNA splicing. *Science* **1986**, *231* (4745), 1534–9.

(33) Son, Y. S.; Park, J. H.; Kang, Y. K.; Park, J. S.; Choi, H. S.; Lim, J. Y.; Lee, J. E.; Lee, J. B.; Ko, M. S.; Kim, Y. S.; Ko, J. H.; Yoon, H. S.; Lee, K. W.; Seong, R. H.; Moon, S. Y.; Ryu, C. J.; Hong, H. J. Heat shock 70-kDa protein 8 isoform 1 is expressed on the surface of human embryonic stem cells and downregulated upon differentiation. *Stem Cells* **2005**, *23* (10), 1502–13.

(34) Baharvand, H.; Fathi, A.; Gourabi, H.; Mollamohammadi, S.; Salekdeh, G. H. Identification of mouse embryonic stem cell-associated proteins. *J. Proteome Res.* **2008**, *7* (1), 412–23.

(35) Suganuma, R.; Yanagimachi, R.; Meistrich, M. L. Decline in fertility of mouse sperm with abnormal chromatin during epididymal passage as revealed by ICSI. *Hum. Reprod.* **2005**, *20* (11), 3101–8.

(36) Park, M. H.; Wolff, E. C.; Smit-McBride, Z.; Hershey, J. W.; Folk, J. E. Comparison of the activities of variant forms of eIF-4D. The requirement for hypusine or deoxyhypusine. *J. Biol. Chem.* **1991**, *266* (13), 7988–94.

(37) Park, M. H.; Cooper, H. L.; Folk, J. E. Identification of hypusine, an unusual amino acid, in a protein from human lymphocytes and of spermidine as its biosynthetic precursor. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78* (5), 2869–73.

(38) Schnier, J.; Schwelberger, H. G.; Smit-McBride, Z.; Kang, H. A.; Hershey, J. W. Translation initiation factor 5A and its hypusine modification are essential for cell viability in the yeast Saccharomyces cerevisiae. *Mol. Cell. Biol.* **1991**, *11* (6), 3105–14.

(39) Buratti, E.; Baralle, F. E. Characterization and functional implications of the RNA binding properties of nuclear factor TDP-43, a novel splicing regulator of CFTR exon 9. *J. Biol. Chem.* **2001**, *276* (39), 36337–43.

(40) Gerhard, D. S.; Wagner, L.; Feingold, E. A.; Shenmen, C. M.; Grouse, L. H.; Schuler, G.; Klein, S. L.; Old, S.; Rasooly, R.; Good, P.; Guyer, M.; Peck, A. M.; Derge, J. G.; Lipman, D.; Collins, F. S.; Jang, W.; Sherry, S.; Feolo, M.; Misquitta, L.; Lee, E.; Rotmistrovsky, K.; Greenhut, S. F.; Schaefer, C. F.; Buetow, K.; Bonner, T. I.; Haussler, D.; Kent, J.; Kiekhaus, M.; Furey, T.; Brent, M.; Prange, C.; Schreiber, K.; Shapiro, N.; Bhat, N. K.; Hopkins, R. F.; Hsie, F.; Driscoll, T.; Soares, M. B.; Casavant, T. L.; Scheetz, T. E.; Brownstein, M. J.; Usdin, T. B.; Toshiyuki, S.; Carninci, P.; Piao, Y.; Dudekula, D. B.; Ko, M. S.; Kawakami, K.; Suzuki, Y.; Sugano, S.; Gruber, C. E.; Smith, M. R.; Simmons, B.; Moore, T.; Waterman, R.; Johnson, S. L.; Ruan, Y.; Wei, C. L.; Mathavan, S.; Gunaratne, P. H.; Wu, J.; Garcia, A. M.; Hulyk, S. W.; Fuh, E.; Yuan, Y.; Sneed, A.; Kowis, C.; Hodgson, A.; Muzny, D. M.; McPherson, J.; Gibbs, R. A.; Fahey, J.; Helton, E.; Ketteman, M.; Madan, A.; Rodrigues, S.; Sanchez, A.; Whiting, M.; Madari, A.; Young, A. C.; Wetherby, K. D.; Granite, S. J.; Kwong, P. N.; Brinkley, C. P.; Pearson, R. L.; Bouffard, G. G.; Blakesly, R. W.; Green, E. D.; Dickson, M. C.; Rodriguez, A. C.; Grimwood, J.; Schmutz, J.; Myers, R. M.; Butterfield, Y. S.; Griffith, M.; Griffith, O. L.; Krzywinski, M. I.; Liao, N.; Morin, R.; Palmquist, D.; Petrescu, A. S.; Skalska, U.; Smailus, D. E.; Stott, J. M.; Schnerch, A.; Schein, J. E.; Jones, S. J.; Holt, R. A.; Baross, A.; Marra, M. A.; Clifton, S.; Makowski, K. A.; Bosak, S.; Malek, J. The status, quality, and expansion of the NIH full-length cDNA project: the Mammalian Gene Collection (MGC). *Genome Res.* **2004**, *14* (10B), 2121–7.

(41) Honbou, K.; Suzuki, N. N.; Horiuchi, M.; Niki, T.; Taira, T.; Ariga, H.; Inagaki, F. The crystal structure of DJ-1, a protein related to male fertility and Parkinson's disease. *J. Biol. Chem.* **2003**, *278* (33), 31380–4.

(42) Kanatsu-Shinohara, M.; Morimoto, T.; Toyokuni, S.; Shinohara, T. Regulation of mouse spermatogonial stem cell self-renewing division by the pituitary gland. *Biol. Reprod.* **2004**, *70* (6), 1731–7.

(43) Kanatsu-Shinohara, M.; Lee, J.; Inoue, K.; Ogonuki, N.; Miki, H.; Toyokuni, S.; Ikawa, M.; Nakamura, T.; Ogura, A.; Shinohara, T. Pluripotency of a single spermatogonial stem cell in mice. *Biol. Reprod.* **2008**, *78* (4), 681–7.

PR900565B

3.3 Multipotent adult germline stem cells and embryonic stem cells functional proteomics revealed an important role of translation initiation factor 5A (Eif5a) in stem cell differentiation

In this chapter, the comparison of undifferentiated and differentiated ESCs and maGSCs at proteome level is described. 2D-gelectrophoresis- and 2D-DIGE-reference maps were created to screen for changes in protein profiles which are associated with stem cell differentiation. Herewith, it was possible to identify 36 proteins with a difference in protein level before and after differentiation in both ESCs and maGSCs. Among these proteins, 18 were downregulated during differentiation and another 18 show a higher expression in differentiated cell lines. Despite the similarities at proteome level in undifferentiated cell lines, the differentiated ESCs and maGSCs could be distinguished by 55 differentially expressed proteins. One of the proteins which were downregulated during retinoic acid (RA)-induced differentiation in both cell types was Eif5a. This protein was previously described to play an important role in cell proliferation and differentiation, and its activation is governed by the post-translational modification hypusination. Alteration of hypusination of Eif5a resulted in a reduced proliferation of ESCs and maGSCs, whereas it did not have an effect on the pluripotency of the cell lines.

Hassan Dihazi[*], Gry H. Dihazi[*], Olaf Jahn, **Sandra Meyer**, Jessica Nolte, Abdul R. Asif, Gerhard A. Müller and Wolfgang Engel

[*]These authors contributed equally to this work.

Author contributions to the work:

1a. Hassan Dihazi: conception and design of experiments, data analysis and interpretation, preparation of manuscript.

1b. Gry H. Dihazi: isolation of proteins, performance of 2D-gelelectrophoresis and 2D-DIGE (partial), performance of Western Blot (partial).

3. Olaf Jahn: collection and assembly of mass spectrometry data.

4. **Sandra Meyer**: performance of cell culture, differentiation of cell lines, preparation of cells for proteome analysis, preparation of cells for MTT assay and for examination of effect of hypusination on stem cell differentiation, involved in data interpretation, involved in preparation of manuscript.

5. Jessica Nolte: generation of cell lines, involved in preparation of manuscript.

6. Abdul R. Asif: collection and assembly of mass spectrometry data.

7. Gerhard A. Müller: financial and administrative support.

8. Wolfgang Engel: conception and design of experiments, involved in preparation of manuscript.

Multipotent adult germline stem cells and embryonic stem cells functional proteomics revealed an important role of translation initiation factor 5A (Eif5a) in stem cell differentiation

Hassan Dihazi[*#1], Gry H. Dihazi[*1], Olaf Jahn[2,3], Sandra Meyer[4], Jessica Nolte[4], Abdul R. Asif[5], Gerhard A. Müller[1] and Wolfgang Engel[4]

[1]Department of Nephrology and Rheumatology, University of Göttingen, 37075 Göttingen, Germany.
[2]Proteomics Group, Max-Planck-Institute of Experimental Medicine, 37075 Göttingen, Germany.
[3]Deutsche Forschungsgemeinschaft Research Center for Molecular Physiology of the Brain, 37073 Göttingen, Germany.
[4]Institute of Human Genetics, University of Göttingen, 37073 Göttingen, Germany.
[5]Department of Clinical Chemistry, University of Göttingen, 37075 Göttingen, Germany.

*These authors contributed equally to this work.
#Corresponding author.

Abstract

Multipotent adult germline stem cells (maGSCs) are pluripotent cells that can be differentiated into somatic cells of the three primary germ layers. In order to highlight the protein profile changes associated with stem cell differentiation, retinoic acid (RA) treated mouse stem cells (maGSCs and ESCs) were compared to untreated cells. 2-DE and DIGE reference maps were created, and differentially expressed proteins were further processed for identification. In both stem cell types the RA-induced differentiation resulted in alteration of 36 proteins of which 18 were downregulated and might be potential pluripotency associated proteins, whereas the other 18 proteins were upregulated. These might be correlated to stem cell differentiation. Surprisingly, eukaryotic initiation factor 5A (Eif5a), a protein which is essential for cell proliferation and differentiation, was significantly downregulated under RA-treatment. A time dependent investigation of Eif5a showed that the RA-treatment of stem cells resulted in a significant upregulation of Eif5a in the first 48h followed by a progressive downregulation thereafter. This effect could be blocked by the hypusination inhibitor ciclopirox olamine (CPX). The alteration of Eif5a-hypusination, as confirmed by mass spectrometry, exerts an antiproliferative effect on ESCs and maGSCs in vitro, but does not affect the cell pluripotency. Our data highlights the important role of Eif5a and its hypusination for stem cell differentiation and proliferation.

Introduction

Stem cells (SCs) and especially embryonic stem cells (ESCs) are pluripotent cells characterized by their functional capacity of self-renewal and their ability to generate all types of differentiated cells [1-3]. The therapeutic use of ESCs could be impeded by problems regarding immune rejection due to genetic differences between the patient and the donor ESCs, as well as ethical issues associated with the use of embryos [4].

Spermatogonial stem cells (SSCs) are self-renewing single cells located in the periphery of the seminiferous tubules whose continuous division maintains spermatogenesis throughout the life of a male individual [5]. Several studies have revealed that the germline lineage retains the potential to generate pluripotent cells. In 2004, ESC-like cells were found in cell cultures of germ cells established from murine neonatal testis, designated as multipotent germline stem cells (mGSCs) [6]. In 2006, we isolated and cultured spermatogonial stem cells (SSCs) from the adult mouse testis which responded to culture conditions and acquired ESC properties [7]. We proved that the pluripotency and plasticity of these cells, which were named multipotent adult germline stem cells (maGSCs), were similar to ESCs. They were able to spontaneously differentiate into derivatives of the three embryonic germ layers *in vitro*, to generate teratomas in immunodeficient mice and to contribute to the development of various organs when injected into an early blastocyst. Our results were confirmed by other groups [8, 9]. In 2008, we could confirm with respect to microRNA (miRNA) expression, the ESC-like nature of maGSCs [10], and their potential as an alternative source of pluripotent cells from non-embryonic tissues. The miRNAs were constantly expressed in maGSCs and downregulated after long exposure to differentiation conditions [10]. It has been shown that maGSCs share important pluripotency features with male ESCs such as telomerase activity, telomere length and hypomethylation of pluripotency marker genes. This strengthens the view of maGSCs as pluripotent cells [11]. More recently, Conrad and colleagues [12] generated pluripotent SCs from adult human testis. The established human adult GSCs were able to differentiate into various types of somatic cells of all three germ layers when cultured under conditions that stimulated human ESCs to differentiation [12].

In a recent study, we created a proteomic map for maGSCs and compared it to the ESCs proteome from the same mouse. We confirmed on the proteome level the ESC-like nature of maGSCs [13]. To identify new potential pluripotency correlated proteins or differentiation associated proteins, we investigated the effects of retinoic acid (RA) treatment on the protein expression profiles of maGSCs and ESCs and identified proteins potentially involved in the SCs differentiation. We demonstrated the important role of Eif5a in stem cell differentiation and highlighted the role of the hypusination as a stem cell differentiation and cell cycle control step.

Material and Methods

Derivation and culture of maGSC- and ESC-lines

The derivation and culture of maGSCs 129/SV (maGSC 129/SV) and the transgenic cell line Stra8-EGFP/ROSA26 (maGSC Stra8) was described previously [7]. In this case, both maGSC-lines were generated without genetic selection, only by morphological criteria. The ESC R1 line was derived from the 129/SV mouse [14]. The cell line ESC Stra8 was generated from the transgenic Stra8-EGFP/ROSA26 mouse as described previously [15]. To maintain maGSCs and ESCs in an undifferentiated state, the cells were cultured under standard ESC culture conditions: DMEM (PAN, Aidenbach, Germany) supplemented with 20% fetal calf serum (PAN), 2 mM L-glutamine (PAN), 50 mM β-mercaptoethanol (Gibco/Invitrogen, Eggenstein, Germany), 1x non-essential amino acids (Gibco/Invitrogen), sodium pyruvate (Gibco/Invitrogen), penicillin/streptomycin (PAN). ESCs and maGSCs were cultured on a feeder layer of mitomycin C-inactivated mouse embryonic fibroblasts (MEFs) in the presence of 1000 U/ml recombinant mouse leukemia inhibitory factor (LIF) (Chemicon, Temecula, USA). ESCs were isolated as described previously, and male ESC lines were identified and selected by PCR amplification of *Sry* gene-specific sequences [15]. Male ESCs and maGSCs were cultured under the same conditions as described above. In order to differentiate maGSCs and male ESCs, the cells were plated on gelatine-coated dishes and culture medium was supplemented with 10^{-6} M RA (Sigma-Aldrich, Steinheim, Germany) instead of LIF. Cells were cultured for 20 days before they were lysed and the proteins were extracted.

Protein extraction

The protein extraction for 2D gel electrophoresis was performed as described previously [16]. Briefly, 75% confluent cultures were trypsinized and washed three times with 1x PBS. The cells were harvested by centrifugation at 200 x g for 10 min, the pellet was treated with 0.3-0.5 ml lysis buffer (9.5 M urea, 2% CHAPS (w/v), 2% ampholytes (w/v), 1% DTT). Ampholytes and DTT were added before use. After adding the lysis buffer the samples were incubated for 30 min at 4°C. For removing the cell debris sample centrifugation was carried out at 13,000 x g and 4°C for 45 min. The supernatant was recentrifuged at 13,000 x g and 4°C for an additional 45 min to get maximal purity. The resulting samples were used immediately or stored at -80°C until use.

Protein precipitation

To reduce the salt contamination and to enrich the proteins, methanol-chloroform-precipitation according to Wessels and Flugge [17] was performed. Briefly, 0.4 ml of methanol (100%) was added to 0.1 ml aliquots of protein samples and mixed together. 0.1

ml chloroform was added to the samples and the mixture was vortexed. Subsequently, 0.3 ml water was added and the solution was vortexed and centrifuged at 13,000 x g for 1 min. The aqueous layer was removed, and another 0.4 ml methanol (100%) was added to the rest of the chloroform and the interphase with the precipitated proteins. The sample was mixed and centrifuged for 2 min at 13,000 x g and the supernatant was removed. The pellet was vacuum dried and dissolved in lysis buffer.

Total protein concentration was estimated using the Bio-Rad protein assay (Bio-Rad, Hercules, USA) according to Bradford [18]. Bovine serum albumin (Sigma-Aldrich) was used as a standard.

2D Gel Electrophoresis (2-DE)

IPG strips (11 cm, pI 5-8) were passively rehydrated in 185 µl solution containing 150 µg protein in a rehydration buffer (8 M urea, 1% CHAPS, 1% DTT, 0.2% ampholytes, and a trace of bromophenol blue) for 12 h. The IEF step was performed on the PROTEAN® IEF Cell (Bio-Rad). Temperature-controlled at 20°C, the voltage was set to 500 V for 1 h, increased to 1,000 V for 1 h, 2,000 V for 1 h and left at 8,000 V until a total of 50,000 Vhours was reached. Prior to SDS-PAGE, the IPG strips were reduced for 20 min at room temperature in SDS equilibration buffer containing 6 M urea, 30% glycerol, 2% SDS, 0.05 M Tris-HCl, and 2% DTT on a rocking table. The strips were subsequently alkylated in the same solution with 2.5% iodacetamide substituted for DTT, and a trace of bromophenol blue. For the SDS-PAGE 12% BisTris Criterion precast gels (Bio-Rad) were used according to manufacturer's instructions. The gels were run at 150 V for 10 min, followed by 200 V until the bromophenol blue dye front had reached the bottom of the gel.

Gel staining

For image analysis, 2-DE gels were fixed in a solution containing 50% methanol and 12% acetic acid over night and fluorescent stained with Flamingo fluorescent gel stain (Bio-Rad) for minimum 5 h. After staining, gels were scanned at 50 µm resolution on a Fuji FLA-5100 scanner (Fuji Photo, Kanagawa, Japan). The digitalized images were analyzed using Delta 2D 3.4 (Decodon, Braunschweig, Germany). For protein identification, 2-DE gels were additionally stained with colloidal Coomassie blue, Roti-blue (Roth, Karlsruhe, Germany) overnight.

2D-DIGE

Protein extraction and methanol-chloroform-precipitation were performed as described above. The resulting pellet was dissolved in labeling buffer (30 mM Tris-HCl pH 8.5, 9.5 M

urea, 2% CHAPS), centrifuged (5 min, 13,000 x g), and the protein concentration of the supernatant was determined as described above.

Each dye was freshly dissolved in anhydrous N,N-dimethylformamide (DMF) (Sigma-Aldrich, St. Louis, USA) to a stock solution containing 1000 pmol/µl. One volume of CyDye stock solution was added to 1.5 volumes of high grade DMF to make a 400 µM CyDye solution. For minimal labeling 400 pmol of the amine-reactive cyanine dyes Cy3 and Cy5 was added respectively to 50 µg proteins from each maGSCs and ESCs, following the manufacturer's protocol (GE Healthcare, Munich, Germany). The labeling reaction was carried out at 4°C in the dark for 30 min and the reaction was terminated by addition of 10 nmol lysine at 4°C in the dark for 10 min. Equal volumes of 2x sample buffer (30 mM Tris-HCl pH 8.5, 9.5 M urea, 2% CHAPS, 10 mM PMSF, 130 mM DTT and 2% ampholytes 3-10) were added to each of the labeled protein samples. To avoid the dye-specific protein labeling, every pair of protein samples from two independent cell extract preparations were processed in duplicate while swapping the dyes. Thereby four replicate gels were obtained, allowing to monitor regulation factors down to twofold changes [19]. 50 µg of an internal standard consisting of a mixture of all samples under investigation were labeled with 400 pmol Cy2 and included on all gels to facilitate gel matching, thereby eliminating artifacts from experimental variation. The three differentially labeled fractions were pooled. Rehydration buffer (8M urea, 1% CHAPS, 13 mM DTT and 1% ampholytes 3-10) was added to make up the volume to 185 µl prior to IEF. The 2-DE was performed as described above. The CyDye-labeled gels were scanned at 50 µm resolution on a Fuji FLA5100 scanner (Fuji Photo) with laser excitation light at 473 nm and long pass emission filter 510LP (Cy2), 532 nm and long pass emission filter 575LP (Cy3), and 635 n, and long pass emission filter 665LP (Cy5). Fluorescent images were acquired in 16-bit TIFF files format. Spot matching across gels and normalization based on the internal standard was performed with Delta 2D software (Decodon). To analyze the significance of protein regulation, a Student's t-test was performed, and statistical significance was assumed for p-values less than 0.01. For protein identification, 2-DE were post stained with colloidal Coomassie blue (Roti-Blue) overnight. Differentially regulated proteins were excised and processed for identification by mass spectrometry.

Protein identification

Manually excised gel plugs were subjected to an automated platform for the identification of gel-separated proteins [20] as described in the framework of recent DIGE-based [21] and large-scale proteome studies [22]. An Ultraflex MALDI-TOF-TOF mass spectrometer (Bruker Daltonik) was used to acquire both PMF and fragment ion spectra, resulting in confident protein identifications based on peptide mass and sequence information.

Database searches in the Swiss-Prot primary sequence database restricted to the taxonomy *mus musculus* were performed using the MASCOT Software 2.2 (Matrix Science). Carboxamidomethylation of Cys residues was specified as fixed and oxidation of Met as variable modifications. One trypsin missed cleavage was allowed. Mass tolerances were set to 100 ppm for PMF searches and to 100 ppm (precursor ions) and 0.7 Da (fragment ions) for MS/MS ion searches. The minimal requirement for accepting a protein as identified was at least one peptide sequence match above identity threshold in addition to at least 20% sequence coverage in the PMF.

Western Blot analysis

The confirmation of the protein expression differences between differentiated and undifferentiated maGSCs and ESCs detected during the proteomic analysis data was done performing Western Blot analyses according to Towbin et al. [23]. Rabbit monoclonal anti-Eif5a (Abcam, UK), rabbit polyclonal anti-Tardbp (Abcam), rabbit polyclonal anti-stathmin 1 (Abcam), rabbit anti-laminin (Sigma-Aldrich), rabbit polyclonal anti Pdlim (Abcam), mouse monoclonal anti-Trim28 (Abnova), rabbit polyclonal anti-peroxoredoxin 6 (Abcam), rabbit anti-nucleolin (Sigma-Aldrich), rabbit polyclonal anti-enolase (Santa Cruz), rabbit poyclonal anti-Oct4 (Abcam), rabbit polyclonal anti-Sox2 (Abcam) and mouse monoclonal anti-Tuba antibody (Sigma-Aldrich) were used as primary antibodies. Molecular Probes® Alexa Fluor 647 goat anti-mouse IgG antibody, Alexa Fluor 647 goat anti-rabbit IgG, Alexa Fluor 680 goat anti-mouse IgG antibody and/or Alexa Fluor 680 goat anti-rabbit were used as secondary antibodies. Before imaging, the blots were dried in the dark. The blot membranes were scanned at 50 μm resolution on a Fuji FLA-5100 scanner (Fuji Photo) with single laser-emitting excitation light at 635 nm and 670 nm, respectively.

Bioinformatics

The classification of the identified proteins according to their main known/postulated function was carried out using DAVID bioinformatics (http://www.david.abcc.ncifcrf.gov). This classification together with the official gene symbol (given in Supplemental Table 1) was used to investigate and categorize the gene ontology (GO)-annotations (biological processes and molecular functions). To predict potential protein-protein interaction between the altered protein expressions and to understand the connection to cell differentiation network generation was performed using STRING 8.2 (Search Tool for the Retrieval of Interacting Genes/Proteins, http://string.embl.de) [24].

MTT assay for cell proliferation

For the cell proliferation assay the cell proliferation kit I (MTT) from Roche was used. ESCs and maGSCs were plated into 96-well flat-bottomed microtiter plates (Becton Dickinson, Heidelberg, Germany) with 10^4 cells/well in 150 µl of their respective media. To investigate the effect of ciclopirox olamine (CPX) on cell proliferation, the cells were preincubated for 24h before treatment with increasing concentrations of CPX (0-15 µM) for different incubation times. Furthermore, RA (1 µM) treated cells were subjected to 2 µM CPX and the cell proliferation was monitored in a time dependent manner. All analyses were performed in triplicate. The cell viability in each well was assayed according to the manufacturer's recommendation. The cells were incubated with the yellow MTT solution (10 µl solution to 100 µl medium) for 4h at 37°C. After the incubation period, purple formazan salt crystals were formed. These salt crystals are insoluble in aqueous solution, but may be solubilized by adding the solubilization solution (100 µ/slit) and incubating these overnight in a humidified atmosphere (37°C, 5% CO_2). The quantification of the formazan product is performed spectrophotometrically by an ELISA-reader (PerkinElmer). An increase/decrease of the number of viable cells results in an increase/decrease of the total metabolic activity in the sample. This increase/decrease correlates directly to the amount of purple formazan crystals which are formed.

Monitoring the effect of hypusination inhibition on stem cell differentiation and proliferation

ESCs and maGSCs were cultured in 24-well plates for 24 h. Subsequently, the cells were treated with CPX (2 µM) for 72 h. The CPX was then removed by replacing the culture medium with fresh medium containing RA (1 µM), LIF (10^3 units/ml) or none of the two. The impact of the treatment on cell differentiation and proliferation was followed for 4 days using a Zeiss Axiophot microscope (Carl Zeiss, Jena, Germany) and the AnalySIS software (Soft Imaging Systems, Leinfelden, Germany).

Results

Comparative analysis of the proteome of the undifferentiated and the differentiated maGSCs and ESCs by 2D-DIGE

In a recent study we confirmed on the proteome level the ESC-like nature of maGSCs. 166 different non-redundant proteins were identified [13] of which 18 showed a different expression in maGSCs than in ESCs. To explore the proteome changes in stem cells caused by differentiation, and to highlight the pluripotency associated proteins, the effect of the differentiation stimulating agent retinoic acid (RA) on the protein expression in maGSCs and ESCs was investigated. DIGE analysis of maGSCs and ESCs originating from two different mouse lines (129/SV and Stra8-EGFP/ROSA26) was performed, and

compared to the RA stimulated maGSCs and ESCs. The DIGE images were analyzed with the Delta2D software (Decodon); interesting protein spots were excised and analyzed with MALDI-TOF-TOF. Using the MASCOT Database (Matrix Science), we could identify 77 different protein spots (Supplemental Table 1). Comparing both cell types and taking the proteins, which were regulated in the same manner in both cell types (maGSCs and ESCs) in account, the RA treatment resulted in the alteration of 36 non-redundant proteins in the pI-range 5-8 in ESCs as well as in maGSCs. Among these proteins, 18 were found to be downregulated (Atp5h, Eif5a, Eno1, Eno3, Hnrpab, Hnrpf, Hnrnph1, Hnrpdl, Hspd1, Khsrp, Ncl, Nup50, Nup62, Pgam1, Ranbp1, Stmn1, Tardbp, and Trim28), and 18 were upregulated (Actb, Capg, Cstb, Eno1, Erp29, Etfa, Fabp3, Lamc1, Lgals1, Lrpap1, Pdia3, Pdlim, Prdx6, Txndc5, Pgk1, Tubb5, Atp5a1, and Vim) under RA-stimulation. We noticed that one isoform of Eno1 was downregulated and another isoform of Eno1 was upregulated under RA treatment (Figure 1A, 1B, Table 1).
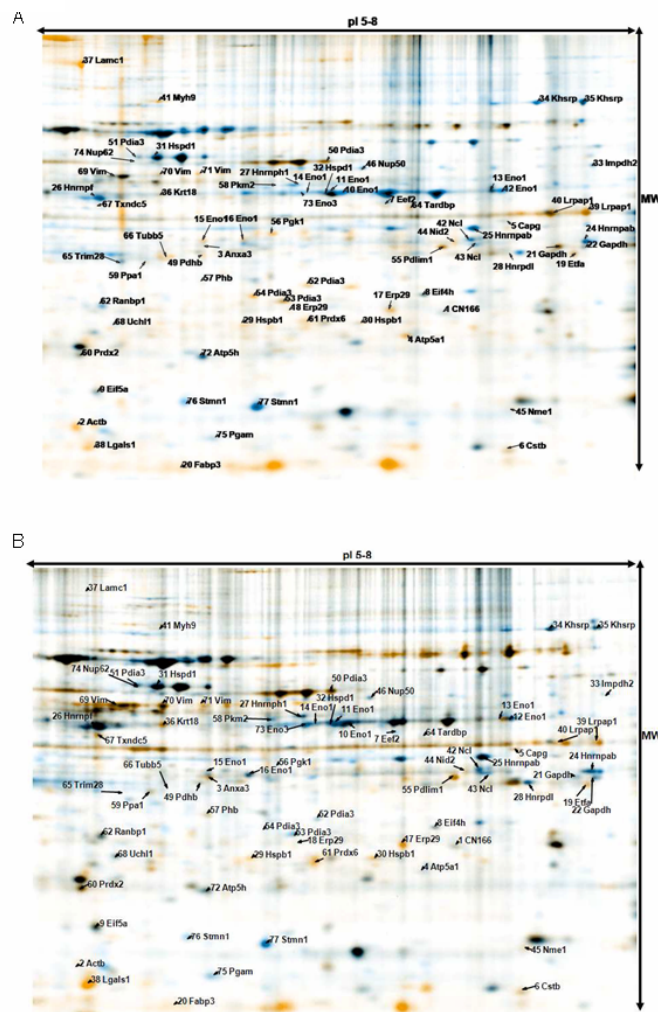
Figure 1: Differential proteome analysis of maGSCs/maGSCs-RA and ESCs/ESCs-RA. After the labeling step (50 µg protein were labeled with 400 pmol of Cy3 respectively Cy5, the internal standard was labeled with Cy2) the samples were subjected to DIGE. A) 2D-DIGE map of maGSCs and maGSCs-RA proteomes (mouse line 129/SV) are shown: Cy3-labeled proteins are false colored in blue (representing

maGSCs proteome) and Cy5-labeled proteins are false colored in orange (representing maGSCs-RA proteome). B) 2D-DIGE map of ESCs and ESCs-RA proteomes (mouse line 129/SV) are shown: Cy3-labeled proteins are false colored in blue (representing ESCs proteome) and Cy5-labeled proteins are false colored in orange (representing ESCs-RA proteome). Gene names of the identified proteins are indicated.

Some of the proteins were behaving differently in maGSCs and in ESCs under RA treatment. Gapdh, Uchl1 and some isoforms of Pdia3 were upregulated in ESCs under RA treatment, but slightly downregulated in maGSCs (Supplemental Table 2). Atp5a1, Pgk1, and Tubb5 showed an upregulation in ESCs but no regulation in maGSCs. Txndc5 showed an upregulation in maGSCs but no regulation in ESCs (Table 1).

Table 1: Listed proteins found to be differently expressed in maGSCs and ESCs after RA-treatment. Presented are proteins, which showed similar behavior in both cell types. The regulation fold is given.

Down-regulated protein spots				Up-regulated protein spots		
Gene name	ESC	maGSC		Gene name	ESC	maGSC
9 Eif5a	0.4	0.6		2 Actb	21	2.1
10 Eno1(1)	0.4	0.5		4 Atp5a1	2.7	
11 Eno1(1)	0.7	0.4		5 Capg	2.3	1.9
12 Eno1(1)	0.2	0.6		6 Cstb	4.7	2.7
24 Hnrnpab	0.4	0.4		15 Eno1(2)	3.7	1.0
25 Hnrnpab	0.3	0.4		17 Erp29	3.1	2.4
26 Hnrnpf	0.5	0.7		19 Etfa	2.5	2.2
27 Hnrnph1	0.4	0.6		20 Fabp3	>100	>100
28 Hnrpdl	0.5	0.6		37 Lamc1	>100	1.0
31 Hspd1	0.5	0.6		38 Lgals1	81	22
34 Khsrp	0.3	0.4		39 Lrpap1	4.0	2.1
35 Khsrp	0.3	0.4		40 Lrpap1	3.7	2.9
42/43 Ncl	0.7	0.6		54 Pdia3	10.7	
46 Nup50	0.4	0.6		55 Pdlim1	3.5	3.0
62 Ranbp1	0.5	0.9		56 Pgk1	2.2	
64 Tardbp	0.8	0.5		61 Prdx6	9.7	6.3
65 Trim28	0.6	0.3		66 Tubb5	2.0	
72 Atp5h	0.6	0.8		67 Txndc5		2.0
73 Eno3	0.2	0.4		69 Vim	1.4	2.9
74 Nup62	0.6	0.4		70 Vim	1.9	2.5
75 Pgam1	0.5	0.2		71 Vim	5.2	5.6
76 Stmn1	0.1	0.2				
77 Stmn1	0.3	0.2				

Comparative analysis of the proteome of the differentiated maGSCs and ESCs

Because of the high proteome similarity between the maGSCs and ESCs [13], DIGE analysis of the cell extracts from the two cell types stimulated with RA was performed. The achieved DIGE-gels of the stimulated cells indicated a significant difference in protein pattern, showing that 55 protein spots were differently expressed, when maGSCs/RA was compared to ESCs/RA (Supplemental Figure 1). RA treatment resulted in different behavior in the investigated cell lines.

<u>Western blot analysis of the proteins presented in different levels in undifferentiated and</u>
<u>differentiated SCs</u>

A selection of the proteins found to be downregulated after stimulation with RA was validated with fluorescent Western Blot analysis. maGSCs isolated from two different mouse lines (Stra8-EGFP/ROSA26 and 129/SV) stimulated by RA and their corresponding ESCs were tested together with their corresponding non-stimulated samples. Fluorescent Western blot using one-step fluorescence imaging with Alexa Fluor Dye-conjugated antibodies offer high signal stability and a wide linear dynamic range. Therefore, it allows for accurate quantification which can effectively complement the DIGE data.

Our results verified the data found in the DIGE experiments. Eif5a, Tardbp, Ncl, Stmn1, and Trim28 were downregulated after stimulation with RA (Figure 2A, 2C), whereas Lamc1, and Pdlim were upregulated (Figure 2B, 2D). Tuba was used as an internal standard. The differentiation of the stem cells was monitored by stem cell markers Oct4 and Sox2, which were significantly downregulated upon RA treatment (Figure 2A). Regarding Eno1, we observed that one isoform (higher mass) of Eno1 was downregulated after stimulation (Figure 2A), whereas another isoform with lower mass was upregulated after RA-treatment (Figure 2B). These differences in regulation of the proteins were observed in both mouse lines and in maGSCs as well as in ESCs. The confirmation of the DIGE data by the Western blot experiments strengthens the described protein alterations during stem cell differentiation.
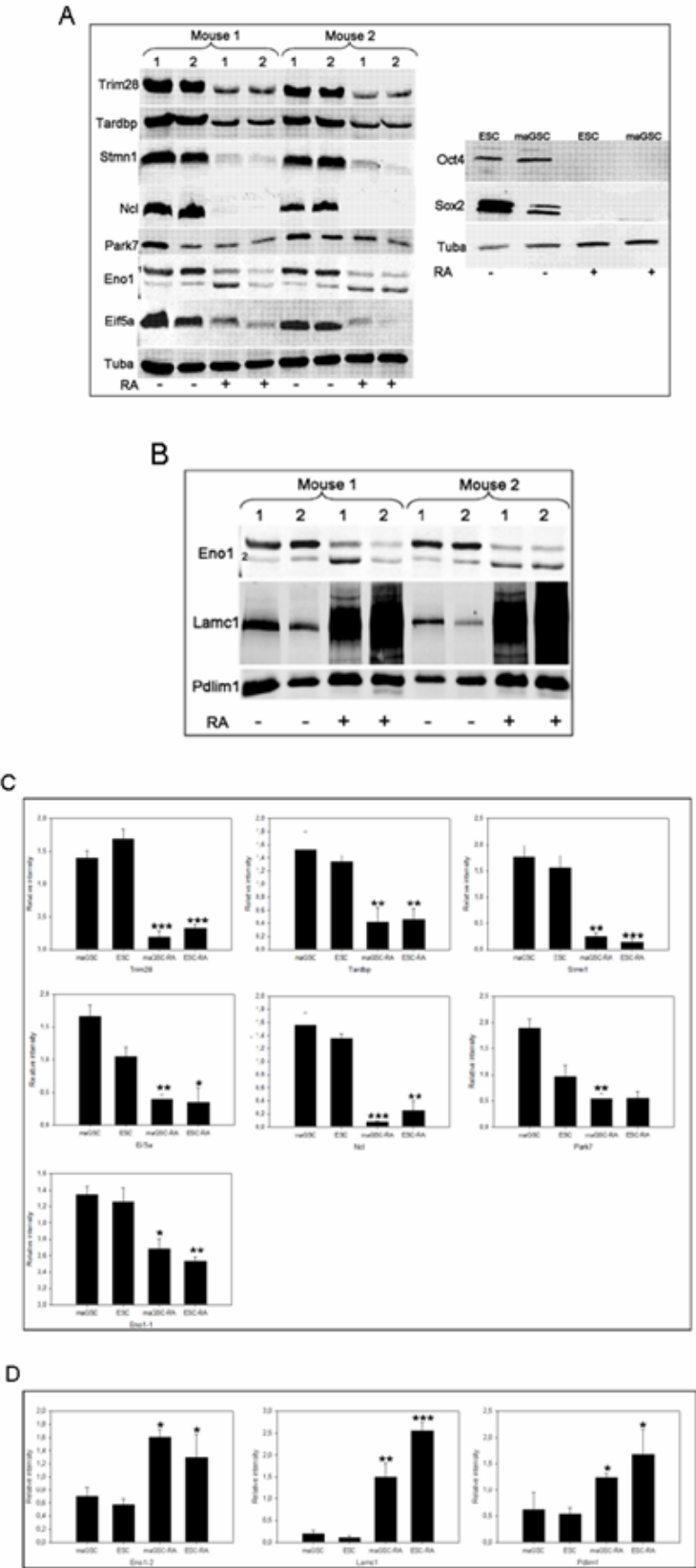
Figure 2: Fluorescent Western blot analyses of proteins which expression was found to be altered as a result of cell differentiation. A) The protein extraction and estimation were carried out as described in the materials and methods. The Western blots were performed with antibodies against selected proteins from the two different mice lines Stra8-EGFP/ROSA26 and 129/SV mice lines respectively were analyzed. Tuba was used as an internal standard, and the stem cell markers Oct4 and Sox2 were used as controls. B) Three of

the upregulated proteins were analyzed with Western blots: Eno1, Lamc1, Pdlim1 in cell extracts from cells originating from two mice lines mentioned above were analyzed. C, D) Western blot quantification was performed by densitometry and is represented in form of histograms. The bar charts represent the ratio (in densitometric units) of the corresponding protein, tubulin (Tuba) was used as a loading control. Comparative statistic analyses of the Western blots were assessed using t-test for paired samples, (*) $p<0.05$, (**) $p<0.01$, (***) $p<0.001$

To gain more information on the biological mechanisms, molecular functions and cellular locations associated with the downregulated proteins, DAVID bioinformatics, Gene Ontology (GO) and PANTHER (Protein ANalysis THrough Evolutionary Relationships) as classification systems were used to categorize the identified proteins. The 18 downregulated proteins were analyzed (Atp5h, Eif5a, Eno1, Eno3, Hnrpab, Hnrpf, Hnrph1, Hnrpdl, Hspd1, Khsrp, Ncl, Nup50, Nup62, Pgam1, Ranbp1, Stmn1, Tardb, and Trim28), and of these proteins 58% were located in the nucleus (Supplemental Figure 2). 14 of the downregulated proteins were nucleic acid binding, and were found to be involved in the mRNA transcription regulation, 8 were involved in the mRNA splicing and 9 in the nucleoside, nucleotide and nucleic acid metabolism (Supplemental Figure 3, Supplemental Table 3). Furthermore, 8 proteins were involved in the KRAB box transcription factor (transcriptional repressor), and 3 in other zinc finger transcription factors (Supplemental Figure 4, Supplemental Table 4). The functional analysis of the downregulated proteins revealed that the larger part of these proteins are involved in transcription regulation and have potential to be pluripotency associated proteins.

To bring more light in the data generated and to analyze the proteins in the context of biological processes, pathways and networks, the data were compared using STRING: functional protein association networks 8.2. The database and web-tool STRING is a metaresource that aggregates most of the available information on protein-protein association, scores these, and augments this within predicted interactions as well as with the results of automatic literature-mining searches [24]. The basic interaction of STRING is the functional association, which in this database, is defined as the specific and meaningful interaction between two proteins which contribute to the same function process. To generate an interaction network, the 166 identified non-redundant proteins from the former study [13] including the proteins found to be downregulated upon RA treatment were given into the database which created an interaction network of the predicted associations of the proteins. The proteins which were downregulated under RA treatment were highlighted on the network image. The interactions among these downregulated proteins were higher than the interactions among the other proteins (Supplemental Figure 5A). This finding might explain their similar behavior under RA treatment. Moreover, we used STRING 8.2 to investigate the interaction partners of Eif5a. We generated another network map illustrating Eif5a and its interaction partners among

the downregulated proteins. Predicted functional partners of Eif5a were added to the interaction map (Supplemental Figure 5B). The generated network showed a strong interaction between Eif5a and the proteins which were involved in nuclear protein transport. This finding confirms the possible nuclear translocation of Eif5a.

Inhibition of Eif5a hypusination impacts the stem cell proliferation and differentiation

The DIGE and Western blot data showed a downregulation of Eif5a upon RA-treatment. Eif5a was described to be involved in cell proliferation and differentiation [25]. To highlight the possible role of this protein in stem cell differentiation, a time dependent expression regulation upon RA-treatment was investigated performing Western blot analysis. Short time treatment of RA, 24-48 h, resulted in a significant increase in the Eif5a expression, whereas a prolonged stimulation led to a significant decrease in Eif5a expression (Figure 3).
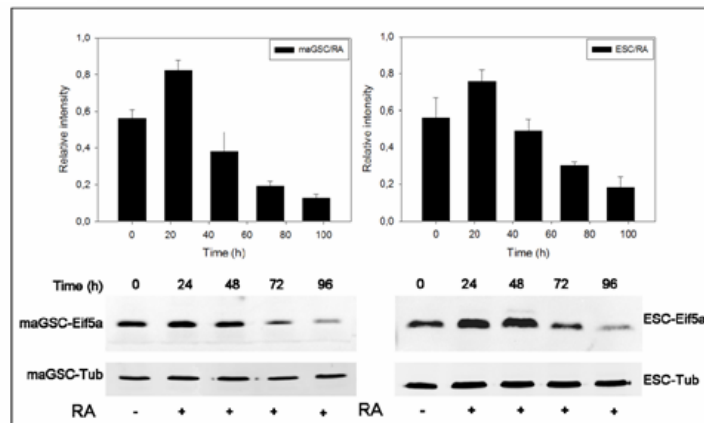
Figure 3: Fluorescent Western blot analyses of Eif5a expression alteration during stem cell differentiation. The protein extraction and estimation were carried out as described in the material and methods. The Western blots were performed with anti-Eif5a antibody in cell extracts from cells treated with RA for different incubation times. Western blot quantification was performed by densitometry and is represented in form of histograms. The bar charts represent the ratio (in densitometric units) of the corresponding protein and tubulin (Tuba) was used as a loading control. Comparative statistical analyses of the Western blots were assessed using t-test for paired samples.

Eif5a activation is controlled by a posttranslational modification called hypusination. This modification requires two steps which are controlled by two different enzymes. CPX is a hypusination inhibitor which controls the second step of the modification, which is catalyzed by deoxyhypusine hydroxylase [26]. We investigated the effect of this inhibitor on the cellular viability and proliferation of ESCs and maGSCs using time- and concentration dependent treatment of the cells and MTT-assay. ESCs and maGSCs were treated with increasing CPX concentrations during different incubation times. The hypusination

inhibition was controlled by mass spectrometry and the cell viability was monitored by the MTT assay (Figure 4).
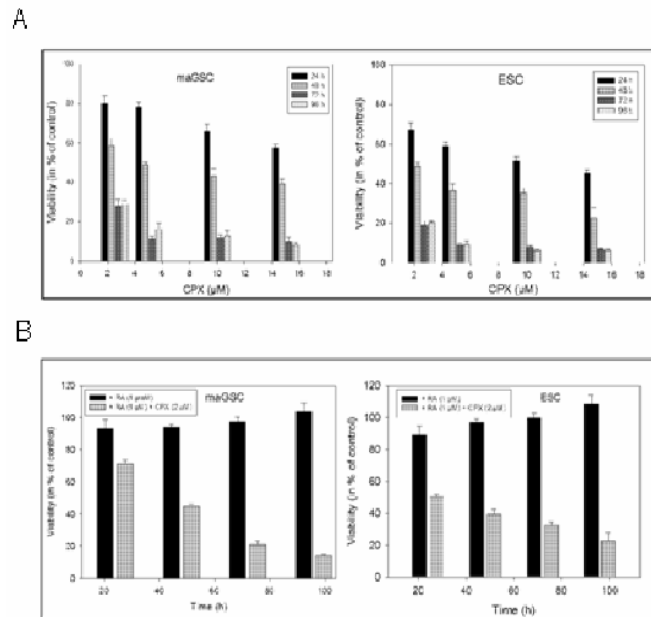

Figure 4: Effect of CPX on stem cell proliferation and viability. A) ESCs and maGSCs treated with different CPX concentrations (0-15 µM) for different incubation times. The cell viability was assessed using MTT-assay. B) Effect of hypusination inhibition on the cell proliferation was assessed using MTT-assay. Stem cells were treated either with RA (1 µM) or with RA-CPX (1 µM/2 µM). The cell viability was tested.

The overall cell viability of the pluripotent cells was found to be strongly dependent on the CPX concentration as well as the incubation time in both cell lines, ESCs and maGSCs. In the presence of 2.5 µM CPX the cell viability of the maGSCs in culture was 79.6 ± 4.1%, 58.34 ± 5.5%, 27.3 ± 3.4% and 28.6 ± 2.1% after 24, 48, 72 and 96 h, respectively (Figure 5B). A higher CPX concentration (15 µM) resulted in a significant decrease in the cell viability. After 24 h of incubation, the cell viability was 56 ± 2.9%, and only 8 ± 2.4% survived after 96 h treatment (p<0.05, Student's t-test) (Figure 4A). The CPX had a similar effect on the ESCs; the viability of the ESCs was significantly affected by the CPX in a time- and concentration-dependent manner (p<0.05, Student's t-test). Moreover, the ESCs were found to be more sensitive to CPX than the maGSCs; the cell viability of the ESCs was less than 75% after 24 h of incubation with 2.5 µM CPX (Figure 4A). The combination of the CPX treatment with RA-stimulation did not impair the effect of CPX on the cells. In contrast to RA treated cells, CPX-RA treatment resulted in CPX-similar effects without RA treatment, namely cell cycle and cell proliferation arrest (Figure 4B).

To investigate the impact of Eif5a hypusination inhibition on maGSCs and ESCs proliferation and differentiation, we treated the cells with CPX (2 µM) for 72 h. Subsequently, the CPX was removed by replacing the culture medium with fresh medium, containing either RA (1 µM), LIF (10³ units/ml) or none of the two. The impact of the

treatment on cell morphology was followed for 4 days using a Zeiss Axiophot microscope (Carl Zeiss) and the AnalySIS software (Soft Imaging Systems). When compared to the control, CPX treated ESCs and maGSCs showed no change in cell differentiation. In contrast, LIF and RA treated cells built colonies and showed normal proliferation, altogether the RA treated cells showed differentiation (Figure 5A, C). The effect of CPX could not be impaired by addition of RA. Both maGSCs and ESCs treated with RA and CPX showed similar behavior to the CPX treated cells, which means slow proliferation and no differentiation (Figure 5A, C). The CPX treatment altered significantly the cell differentiation and proliferation but did not affect the cell pluripotency. The cells were able to differentiate when CPX was removed from the culture medium or replaced by RA or LIF (Figure 5B, D, E, F).

Parallel to the proliferation inhibition, Eif5a showed no significant alteration in the expression level in cells treated with CPX, not even when the cells were treated with RA for 96 h (Figure 6).
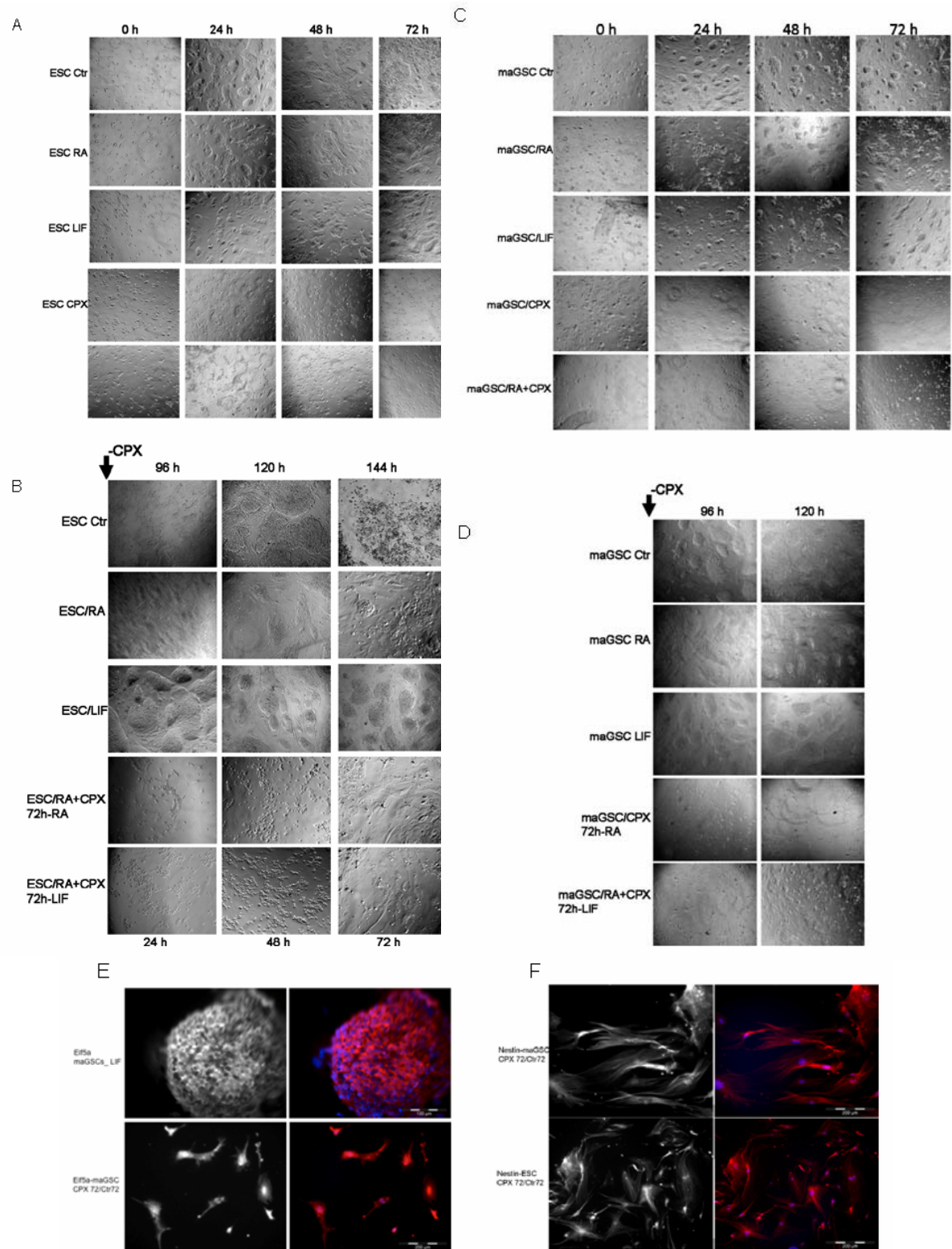
Figure 5: Microscopic analyses of maGSCs and ESCs. A, C) Microscope analyses of ESCs and maGSCs differentiating cells treated with different compounds (RA, LIF, and CPX) for 72 h. B, D) CPX was removed from the respective wells and cell differentiation was monitored with Zeiss Axiophot microscope equipped with the AnalySIS software. E, F) In parallel, the cells were treated with CPX for 72 h followed by 72 h without treatment and immunofluorescence was performed using Eif5a (E) or nestin (F) antibodies. Samples were analyzed on an immunofluorescence Zeiss Axiophot microscope equipped with the AnalySIS software.
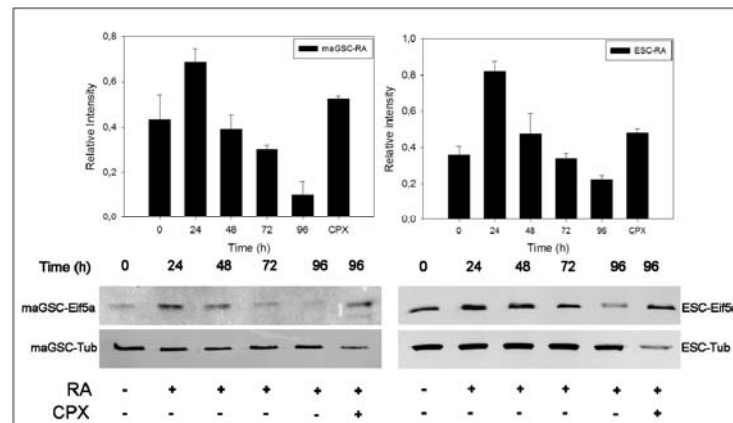
Figure 6: Fluorescent Western blot analyses of Eif5a expression alteration during stem cell differentiation. The protein extraction and concentration determination were carried out as described in the materials and methods. The Western blots were performed with anti-Eif5a antibody in cell extracts from cells treated with either CPX for 96 h or RA for different incubation times. Western blot quantification was performed by densitometry and is represented in form of histograms. The bar charts are represented as the ratio (in densitometric units) of the corresponding protein and tubulin (Tuba) as a loading control. Comparative statistic analyses of Western blots were assessed using t-test for paired samples.

Discussion

ESCs and maGSCs share high similarity in pluripotency [7, 8, 10, 27-29]. In our previous work we studied the proteomes from both cell types. We used DIGE to highlight the ESC-like nature of maGSCs on the proteome level. The maGSCs proteome was found to be highly similar to that of ESCs [13]. We also demonstrated that the inter-individual differences of maGSCs proteomes are minimal. As previously described, these two types of cell lines can differentiate into somatic cells of the three primary germ layers [7]. In this study, we investigated the proteomes of differentiated cells and compared these to the proteome on undifferentiated cells to identify pluripotency associated proteins and proteins, which might correlate to differentiation. RA, as the oxidized form of vitamin A, acts through Hox genes, which are involved in early developmental stages. Depending on the target cells, RA stimulates the transcription of different sets of genes involved in cell differentiation [30, 31]. The stimulation of stem cells with RA resulted in alteration of the expression of 36 proteins, which were classified in two groups according to their expression status (up- or downregulated). Functional analysis of the proteins revealed that a major part of the downregulated proteins were nucleic acid binding proteins, which were found to be involved in the mRNA transcription regulation, in mRNA splicing and in nucleoside, nucleotide and nucleic acid metabolism (Supplemental Table 2, Supplemental Figure 2). Moreover, some of these proteins were involved in transcription repression, which might highlight their potential role in pluripotency. Trim28 is a member of the tripartite motif (TRIM) family, which is involved in transcription regulation. Trim28 is a well-characterized transcriptional repressor. It forms a complex with a KRAB-domain transcription factor and

increases the efficiency of KRAB-mediated repression [32, 33]. It can also silence transcription through an interaction with heterochromatin protein 1 (HP1). Trim28 also functions as a transcriptional corepressor for the KRAB zinc finger proteins and as a moderator of the repression activity [34-36]. Its role in transcription repression and in interaction with KRAB-domain, may explain its role in pluripotency and the downregulation during stem cell differentiation. Tardbp is a DNA- and RNA-binding protein, which is involved in transcription and splicing regulation. Tardbp may also be involved in microRNA biogenesis, apoptosis, cell division and in HIV-1 transcription repression [37, 38]. The downregulation of Tardbp in stem cells upon RA treatment is still not clear, but it may be explained by its potential role as a transcription regulator in pluripotency. The upregulated proteins were predicted to be involved in different processes which are related to cell differentiation (proliferation, oxidative stress, cell adhesion, cell extracellular matrix interaction and cytoskeleton).

Translation elongation factors facilitate protein synthesis by the ribosome. Eif5a is a universal translation elongation factor which is highly conserved in all cells and it is the only eukaryotic protein known to have the unusual amino acid hypusine. The hypusine is the result of a post-translational modification catalyzed by two enzymes: deoxyhypusine synthase and deoxyhypusine hydroxylase [39]. During hypusination, one of the Eif5a lysine residues is modified by the addition of a butylamino group from spermidine to form hypusine, which is found in all eukaryotes and in some archaea, but not in bacteria [40]. Hypusine is essential to the function of Eif5a and is involved in protein biosynthesis by promoting the formation of the first peptide bond and translation elongation [41]. Eif5a was downregulated upon prolonged RA-treatment, whereas short time stimulation resulted in an upregulation of the protein. Eif5a has been shown to be associated with translation, viability and proliferation processes [25, 42-44]. Similar to our results, Luchessi and colleagues[45] observed an increase in Eif5a gene expression in the first 48 h after stem cells from rat skeletal muscles were submitted to differentiation. Moreover, they demonstrated that the inhibition of Eif5a hypusination altered the satellite cell differentiation. The inhibition of maGSC and ESC differentiation during the treatment of the cells with the hypusination inhibitor CPX revealed the importance of Eif5a according to the differentiation. The inhibitory effect of CPX on cell differentiation was found to be reversible and was not associated to apoptosis. The Eif5a expression was upregulated in the first period of cell differentiation and went back to a lower expression level compared to the Eif5a expression level of undifferentiated cells. The hypusination inhibition did not have any impact on the expression of Eif5a, but blocked the cell differentiation completely. The Eif5a hypusination is important for the protein function and for the promotion of translation elongation [41].

Summary

The data presented in this work highlights a set of proteins, which are differentially expressed during stem cell differentiation. Some of the identified proteins were found to be downregulated in accordance with downregulation of the pluripotency markers. This might underline the role of these proteins causing pluripotency. Proteins, which were upregulated accompanying the cell differentiation, were also presented. These proteins might have an important role during stem cell differentiation. Furthermore, our functional proteomics data also highlight the role of hypusination by affecting the Eif5a activity, the stem cell proliferation and differentiation.

Acknowledgement

The authors declare no potential conflict of interest.

References

[1] Weissman IL. Translating stem and progenitor cell biology to the clinic: barriers and opportunities. Science. 2000; 287:1442-1446.

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

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

[4] Vats A, Bielby RC, Tolley NS, et al. Stem cells. Lancet. 2005;366:592-602.

[5] Spradling A, Drummond-Barbosa D, Kai T. Stem cells find their niche. Nature. 2001;414:98-104.

[6] Kanatsu-Shinohara M, Inoue K, Lee J, et al. Generation of pluripotent stem cells from neonatal mouse testis. Cell. 2004;119:1001-1012.

[7] Guan K, Nayernia K, Maier LS, et al. Pluripotency of spermatogonial stem cells from adult mouse testis. Nature. 2006;440:1199-1203.

[8] Seandel M, James D, Shmelkov SV, et al. Generation of functional multipotent adult stem cells from GPR125+ germline progenitors. Nature. 2007;449:346-350.

[9] Izadyar F, Pau F, Marh J, et al. Generation of multipotent cell lines from a distinct population of male germ line stem cells. Reproduction. 2008;135:771-784.

[10] Zovoilis A, Nolte J, Drusenheimer N, et al. Multipotent adult germline stem cells and embryonic stem cells have similar microRNA profiles. Mol Hum Reprod. 2008;14:521-529.

[11] Zechner U, Nolte J, Wolf M, et al. Comparative methylation profiles and telomerase biology of mouse multipotent adult germline stem cells and embryonic stem cells. Mol Hum Reprod. 2009.

[12] Conrad S, Renninger M, Hennenlotter J, et al. Generation of pluripotent stem cells from adult human testis. Nature. 2008.

[13] Dihazi H, Dihazi GH, Nolte J, et al. Multipotent adult germline stem cells and embryonic stem cells: comparative proteomic approach. J Proteome Res. 2009;8:5497-5510.

[14] Nagy A, Rossant J, Nagy R, et al. Derivation of completely cell culture-derived mice from early passage embryonic stem cells. Proc Natl Acad Sci U S A. 1993;90:8424-8428.

[15] Cheng J, Dutra A, Takesono A, et al. Improved generation of C57BL/6J mouse embryonic stem cells in a defined serum-free media. Genesis. 2004;39:100-104.

[16] Dihazi H, Kessler R, Muller GA, et al. Lysine 3 acetylation regulates the phosphorylation of yeast 6-phosphofructo-2-kinase under hypo-osmotic stress. Biol Chem. 2005;386:895-900.

[17] Wessel D, Flugge UI. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. Anal Biochem. 1984;138:141-143.

[18] Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem. 1976;72:248-254.

[19] Karp NA, Griffin JL, Lilley KS. Application of partial least squares discriminant analysis to twodimensional difference gel studies in expression proteomics. Proteomics. 2005;5:81-90.

[20] Jahn O, Hesse D, Reinelt M, et al. Technical innovations for the automated identification of gelseparated proteins by MALDI-TOF mass spectrometry. Anal Bioanal Chem. 2006;386:92-103.

[21] Werner HB, Kuhlmann K, Shen S, et al. Proteolipid protein is required for transport of sirtuin 2 into CNS myelin. J Neurosci. 2007;27:7717-7730.

[22] Reumann S, Babujee L, Ma C, et al. Proteome analysis of Arabidopsis leaf peroxisomes reveals novel targeting peptides, metabolic pathways, and defense mechanisms. Plant Cell. 2007;19:3170-3193.

[23] Towbin H, Staehelin T, Gordon J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. 1979. Biotechnology. 1992;24:145-149.

[24] Jensen LJ, Kuhn M, Stark M, et al. STRING 8--a global view on proteins and their functional interactions in 630 organisms. Nucleic Acids Res. 2009;37:D412-416.

[25] Schnier J, Schwelberger HG, Smit-McBride Z, et al. Translation initiation factor 5A and its hypusine modification are essential for cell viability in the yeast Saccharomyces cerevisiae. Mol Cell Biol. 1991;11:3105-3114.

[26] Balabanov S, Gontarewicz A, Ziegler P, et al. Hypusination of eukaryotic initiation factor 5A (eIF5A): a novel therapeutic target in BCR-ABL-positive leukemias identified by a proteomics approach. Blood. 2007;109:1701-1711.

[27] Kanatsu-Shinohara M, Muneto T, Lee J, et al. Long-term culture of male germline stem cells from hamster testes. Biol Reprod. 2008;78:611-617.

[28] Kanatsu-Shinohara M, Morimoto T, Toyokuni S, et al. Regulation of mouse spermatogonial stem cell self-renewing division by the pituitary gland. Biol Reprod. 2004;70:1731-1737.

[29] Kanatsu-Shinohara M, Lee J, Inoue K, et al. Pluripotency of a single spermatogonial stem cell in mice. Biol Reprod. 2008;78:681-687.

[30] Marshall H, Morrison A, Studer M, et al. Retinoids and Hox genes. Faseb J. 1996;10:969-978.

[31] Holland LZ. Developmental biology: a chordate with a difference. Nature. 2007;447:153-155.

[32] Agata Y, Matsuda E, Shimizu A. Two novel Kruppel-associated box-containing zinc-finger proteins, KRAZ1 and KRAZ2, repress transcription through functional interaction with the corepressor KAP-1 (TIF1beta/KRIP-1). J Biol Chem. 1999;274:16412-16422.

[33] Friedman JR, Fredericks WJ, Jensen DE, et al. KAP-1, a novel corepressor for the highly conserved KRAB repression domain. Genes Dev. 1996;10:2067-2078.

[34] Le Douarin B, Nielsen AL, Garnier JM, et al. A possible involvement of TIF1 alpha and TIF1 beta in the epigenetic control of transcription by nuclear receptors. Embo J. 1996;15:6701-6715.

[35] Schultz DC, Ayyanathan K, Negorev D, et al. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. Genes Dev. 2002;16:919-932.

[36] Schultz DC, Friedman JR, Rauscher FJ, 3rd. Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2alpha subunit of NuRD. Genes Dev. 2001;15:428-443.

[37] Buratti E, Baralle FE. Characterization and functional implications of the RNA binding properties of nuclear factor TDP-43, a novel splicing regulator of CFTR exon 9. J Biol Chem. 2001;276:36337-36343.

[38] Buratti E, Dork T, Zuccato E, et al. Nuclear factor TDP-43 and SR proteins promote in vitro and in vivo CFTR exon 9 skipping. Embo J. 2001;20:1774-1784.

[39] Park MH, Wolff EC, Smit-McBride Z, et al. Comparison of the activities of variant forms of eIF-4D. The requirement for hypusine or deoxyhypusine. J Biol Chem. 1991;266:7988-7994.

[40] Park MH, Cooper HL, Folk JE. Identification of hypusine, an unusual amino acid, in a protein from human lymphocytes and of spermidine as its biosynthetic precursor. Proc Natl Acad Sci U S A. 1981;78:2869-2873.

[41] Saini P, Eyler DE, Green R, et al. Hypusine-containing protein eIF5A promotes translation elongation. Nature. 2009;459:118-121.

[42] Park MH, Lee YB, Joe YA. Hypusine is essential for eukaryotic cell proliferation. Biol Signals. 1997;6:115-123.

[43] Jasiulionis MG, Luchessi AD, Moreira AG, et al. Inhibition of eukaryotic translation initiation factor 5A (eIF5A) hypusination impairs melanoma growth. Cell Biochem Funct. 2007;25:109-114.

[44] Jakus J, Wolff EC, Park MH, et al. Features of the spermidine-binding site of deoxyhypusine synthase as derived from inhibition studies. Effective inhibition by bis- and mono-guanylated diamines and polyamines. J Biol Chem. 1993;268:13151-13159.

[45] Luchessi AD, Cambiaghi TD, Hirabara SM, et al. Involvement of eukaryotic translation initiation factor 5A (eIF5A) in skeletal muscle stem cell differentiation. J Cell Physiol. 2009;218:480-489.

Manuscript submission

Journal of Proteome Research

Multipotent Adult Germline Stem Cells and Embryonic Stem Cells Functional Proteomics Revealed an Important Role of Eukaryotic Initiation Factor 5A (Eif5a) in Stem Cell Differentiation

Journal:	*Journal of Proteome Research*
Manuscript ID:	pr-2010-012015
Manuscript Type:	Article
Date Submitted by the Author:	01-Dec-2010
Complete List of Authors:	Dihazi, Hassan; University Medical Centre Goettingen, Nephorology-Rheumatology Dihazi, Gry; University Medical Centre Goettingen, Nephorology-Rheumatology Jahn, Olaf; Max-Planck-Institute of Experimental Medicine, Proteomics Group, Meyer, Sandra; University Medical Centre Goettingen, Institute of Human Genetics Nolte, Jessica; University Medical Centre Goettingen,, Institute of Human Genetics Asif, Abdul; University Medical Centre Goettingen, Clincal Chemitry Mueller, Gerhard; University Medical Centre Goettingen, Nephorology-Rheumatology Engel, Wolfgang; University Medical Centre Goettingen,, Institute of Human Genetics

3.4 Germ cell specific gene Stra8 has an impact on the pluripotency network

Transcriptome and proteome analyses of differentiated and undifferentiated pluripotent stem cells can be used as a tool for the identification of unknown putative pluripotency-regulating genes. These genes should show a high expression in all analyzed undifferentiated cell lines, and the expression should be strongly downregulated during differentiation of the cells. Since most known pluripotency-regulating genes act as transcription factors, the presence of protein domains known to be involved in transcriptional regulation can provide a further indication whether the downregulated genes encode for proteins involved in regulation of pluripotency.

One of the putative pluripotency-regulating genes identified in a global transcriptome analysis is Stra8 (stimulated by retinoic acid gene 8). In this last part of the thesis, Stra8 was further characterized. It could be shown that Stra8 is expressed in preimplantation stage embryos and in nucleus and cytoplasm of pluripotent cells, whereas its expression is downregulated during spontaneous RA-induced differentiation of these cells. The protein domain structure of Stra8 includes a helix-loop-helix DNA binding domain which suggests a role in transcriptional regulation. We propose that this regulation is performed on genes involved in the maintenance of pluripotency. siRNA-mediated knockdown of Stra8 in ESCs resulted in an increased expression of pluripotency-regulating genes, whereas marker genes related to differentiation of the cells are downregulated. In contrast to this, stable overexpression of Stra8 in ESCs caused a downregulation of expression of genes involved in the pluripotency network. It can be concluded from these results that Stra8 plays a role in the control of the balance between pluripotency and differentiation of pluripotent stem cells.

Jessica Nolte[*], **Sandra Meyer**[*], Tatjana Khromov, Romy Graner, Linda Kotzenberg, Kaomei Guan, Arvind Dev, Krishna D.V. Pantakani and Wolfgang Engel
[*] These authors contributed equally to this work.

Author contributions to the work:
1a. Jessica Nolte: conception and design of experiments, performance of immunocyto-chemistry (partial), conventional RT-PCR, generation of constructs for overexpression, preparation of manuscript.
1b. **Sandra Meyer**: conception and design of experiments, cell culture and transfection of cells with Stra8-siRNA, generation of stable Stra8-overexpressing cell lines, qRT-PCRs

for Stra8-siRNA-treated cells and for stable Stra8-overexpressing cells, involved in preparation of manuscript.

3. Tatjana Khromov: performance of ChIP-analyses and subsequent qRT-PCRs.

4. Romy Graner: LacZ-staining of embryonic stages.

5. Linda Kotzenberg: performance of immunocytochemistry (partial).

6. Kaomei Guan: conception and design of experiments.

7. Arvind Dev: generation of one transgenic mouse line.

8. Krishna D.V. Pantakani: Western Blot for Stra8 in Stra8-siRNA-treated cells.

9. Wolfgang Engel: conception and design of experiments, involved in preparation of manuscript.

Germ cell specific gene *Stra8* has an impact on the pluripotency network

Jessica Nolte[*#1], Sandra Meyer[*1], Tatjana Khromov[1], Romy Graner[1], Linda Kotzenberg[1], Kaomei Guan[2], Arvind Dev[1], Krishna D.V. Pantakani[1] and Wolfgang Engel[1].

[1]Institute of Human Genetics, University of Göttingen, 37073 Göttingen, Germany.
[2]Department of Cardiology and Pneumology, University of Göttingen, 37075 Göttingen, Germany.

[*]These authors contributed equally to this work.
[#]Corresponding author.

Abstract

Stra8 (stimulated by retinoid acid gene 8) is known as a germ cell marker gene that is exclusively expressed in premeiotic spermatogonia of adult male and in premeiotic follicles of adult female mice. Very recently it was shown that Stra8 is also expressed in embryonic stem cells (ESCs) and that the protein is not exclusively localized in cytoplasm but can shuttle between the cytoplasm and the nucleus. Here, we describe that *Stra8* has an impact on the pluripotency network. The expression of *Stra8* is limited to pluripotent cell lines compared to non pluripotent ones, and is downregulated during differentiation of these cells. Immunocytochemistry demonstrated the presence of STRA8 protein in all preimplantation stages including blastocyst stage. This result could be confirmed by two independent transgenic mouse models. During preimplantation stages STRA8 shows cytoplasmic localisation whereas in ESCs it could also be found in the nucleus. STRA8 protein sequence includes a helix-loop-helix DNA binding domain containing a nuclear localisation signal, suggesting a putative role in transcriptional regulation. siRNA-mediated knockdown of *Stra8* in ESCs caused upregulation pluripotency regulating genes like Oct4, Nanog, Klf4, Lin28 and Zfp206 and downregulation of differentiation markers like Hnf4, Nestin and Vimentin. Stable overexpression of *Stra8* in ESCs resulted in the expected downregulation of the pluripotency network and in an increase of expression of mesodermal, but not of ecto- and endodermal differentiation markers. Taken together our results suggest a regulatory function of *Stra8* in the control of the balance between pluripotency and differentiation of embryonic stem cells.

Introduction

The germ cell specific gene *Stra8* (Stimulated by retinoic acid gene 8) was identified in P19 embryonal carcinoma cells as retinoic acid responsive [1]. *Stra8* expression could be induced by retinoic acid in F9 teratocarcinoma cells as well as in embryonic stem cells (ESCs) [2, 3] and its expression was found to be crucial for the initiation of meiosis of

germ cells in both, male and female [4-6]. In female germ cells, *Stra8* is expressed in an anterior to posterior wave corresponding to the wave of expression of early meiotic genes [3]. In male mice, STRA8 protein is localized in the cytoplasm of type A and B spermatogonia, preleptotene spermatocytes as well as in early leptotene spermatocytes [2, 7]. Both *Stra8*-deficient male and female mice are infertile [4, 8, 9]. Taken together, all these data suggest that the decision to enter meiosis is determined by retinoic acid (RA) induction of *Stra8* preceding premeiotic DNA replication.

Despite its undisputable role in initiation of male and female meiosis, recent findings put light on its possible function in pluripotent stem cells. Todesco et al. (2009) [10] described the shuttle of STRA8 between the cytoplasm and the nucleus in ESCs. As this shuttling could play a basic role for the regulation of protein activity important for cell proliferation and signal transduction [11, 12] this raises the question about the function of *Stra8* in pluripotent cells. STRA8 protein has a helix-loop-helix domain that is localized at the N-terminus and is highly conserved during evolution [10]. Such domains are known to be involved in protein-protein interactions [13] and may include functional sequences for the active shuttle of proteins between nucleus and cytoplasm [14]. Basic helix-loop-helix proteins (bHLH) are also known as a group of eukaryotic transcription factors that are involved in a variety of different developmental pathways [15]. It was shown that STRA8 protein is endowed with a functional nuclear localisation signal (NLS) within its helix-loop-helix domain and cross-linking experiments also suggest a putative binding of STRA8 to DNA [10]. The active import of STRA8 into the nucleus let the authors suggest that this protein may act as a transcription factor or at least as a transcriptional co-regulator.

As we used Stra8 promoter activity before to enrich spermatogonial stem cells from adult mouse testis which could be converted to pluripotent maGSCs (multipotent adult germline stem cells) [16] we asked the question, whether *Stra8* is only expressed in pluripotent stem cells or if there is a direct connection to the pluripotency network. We mainly used ESCs and not maGSCs in this study as maGSCs are derived from germ cells, the main acting place of Stra8, and therefore the credibility of the statement for all pluripotent cells could be challenged by someone.

Materials and methods

Cell culture

Pluripotent ESC- and maGSC- cell lines were cultured as previously described [17]. For differentiation, cells were plated on 0.1% gelatine coated dishes without LIF (leukaemia inhibitory factor) and in the presence of 10^{-6}M retinoic acid (RA) for 20 days.

Immunostaining of cells and embryos

Cells were grown on culture slides till 80% confluence and than fixed with 4% paraformaldehyde (PFA) in PBS for 10 min at room temperature. Preimplantation embryos were collected from superovulated wild type or *Stra8-EGFP* transgenic mice by flushing the uterus in due course post-fertilization. Embryos were washed in PBS and than also fixed in 4% PFA in PBS for 10 min at room temperature. Fixed cells as well as fixed embryos were washed twice with PBS and then incubated for 1h in PBS containing 0.1% BSA and 0.1% Tween20 to block for unspecific binding. Cells were incubated with either Stra8 (ab49405, Abcam) or Oct4 (MAB4419, Millipore) antibodies diluted 1:500 in blocking buffer in a humidified chamber at 4°C overnight. Either anti-rabbit IgG Cy3, anti-mouse IgG-Cy3 or anti-mouse IgG-FITC (all SIGMA-Aldrich) were used as secondary antibodies, depending on the specific requirements. For nuclear staining, cells and embryos were covered with DAPI (Vectashield). Pictures were taken using confocal laser scanning microscope (Olympus). To perform negative controls, cells or embryos were incubated with IgG from rabbit serum (I8140, SIGMA-Aldrich) or IgG from mouse serum (I8465; SIGMA-Aldrich) instead of primary antibody.

Transfection of eukaryotic cells with siRNA

Transfection of pluripotent cell lines with siRNA oligonucleotides at a final concentration of 60 nM was performed using Lipofectamine™2000 (Invitrogen) according to manufacturer's instructions. In brief, 500.000 cells per well were plated on 6 well tissue culture plates one day prior to transfection. On the next day 10 µl of siRNA (Invitrogen) (sequence can be found in supplementary table S1) and 10 µl of Lipofectamine™2000 were mixed with 50 µl OptiMEM (Gibco/Invitrogen) each and incubated for 5 min at room temperature. Then both mixtures were combined and incubated for further 20 min at room temperature. Meanwhile, cells were washed with PBS and medium was changed. OptiMEM-siRNA-Lipofectamine™2000-mixture was added drop-by-drop to the cells. After 24h at 37°C, 5% CO_2, medium was changed to normal culture medium and cells were used for isolation of RNA or protein at indicated time points after transfection. For control the same procedure was performed using BLOCK-iT™ Fluorescent Oligo (Invitrogen).

Gene expression studies

Total RNA was extracted from cells or preimplantation embryos by using Trifast Reagent (Peqlab) according to manufacturer's protocol. Prior to reverse transcription using Superscript II cDNA synthesis kit and oligo(dT) primer (Invitrogen), RNA was treated with DNase I (Sigma) to avoid any genomic DNA contamination. Quantitative RealTime-PCR (qRT-PCR) was performed on ABI Prism 7900 HT Fast Detection System (Applied

Biosystems Inc.) as previously described (Kurrasch et al., 2004). Each 10 µl reaction was performed in 384-well format using SYBRgreen PCR Master Mix (Qiagen) and 3 µM of each PCR primer. All reactions were performed from two biological replicates, analysed in one run in triplicate and runs were repeated three times (n=18). Levels of mRNA expression were normalized to those of the mouse housekeeping genes *Sdha* (succinate dehydrogenase) and *HPRT* (Hypoxanthine-guanine phosphoribosyltransferase). Oligonucleotide primers for qRT-PCR were obtained from Eurofins MWG Operon. Primer sequences used for gene amplification can be found in supplementary material (Supplementary Table S2). Conventional RT-PCR amplification was performed using specific primers for *HPRT*, *Stra8* and *Zfp206*. RT-PCR was achieved after 33 cycles at 94ºC, 30 s; 50ºC–62ºC, 30 s; 72ºC, 45 s, depending on the primer sets. Primer sequences are given in Supplementary Table S2.

Isolation of proteins and Western blot

Cells were resuspended in lysis buffer (10 mM Tris/HCl, pH 8.1 mM EDTA, 2.5% SDS), supplemented with 1 mM phenylmethanesulphonylfluoride (PMSF) and protease inhibitors and sonicated. Protein extracts (20 µg) were denatured at 70°C in NuPage SDS sample buffer (Invitrogen), separated on a NuPage 10% Bis-Tris Gel (Invitrogen) and transferred to a Hybond-C Extra membrane (GE Healthcare Europe). Blots were blocked for unspecific binding with 5% nonfat dry milk in blocking buffer (25 mM Tris, 0.15 M NaCl, 0.1% Tween20) and incubated overnight at 4°C with primary and - after washing in blocking buffer for 1 h at 4°C - with secondary HRP-conjugated antibody. Protein bands were visualised using enhanced chemiluminescence as described by the manufacturer (Santa Cruz Biotechnology).

Construction of plasmids for Stra8 overexpression in ESCs

To obtain a vector for the overexpression of *Stra8*, the hEF1alpha- promoter was cut out from the pBOS-vector (BD Bioscience) using HindIII/EcoRI and cloned into pEGFP-1 (Clontech). The resulting phEF1alpha-EGFP vector served as control in all experiments. From this vector, ORF of EGFP was replaced by PCR-amplified ORF of *Stra8* using EcoRI/NotI sites resulting in phEF1alpha-Stra8 (figure 4A).

Generation of transgenic mice

The Stra8-EGFP promoter vector and resulting transgenic mice were previously described [18]. To create pStra8-Cre, the 1.4 kb promoter of Stra8 was PCR-amplified and sequenced with primers containing *XbaI* or *EcoRI* sequences. pBS-Cre construct was produced by Cre-excision from pMC-Cre (kindly provided by Klaus Rajewsky [19]) by

EcoRI/XhoI and cloning into pBluescript II KS+ (Stratagene). Sequence verified 1.4 kb Stra8 promoter was cloned into pBS-Cre. Resulting Stra8-Cre construct was linearized by *XbaI* and injected into pronuclei of fertilized oocytes from FVB mice. Resulting transgenic animals were bred to create a colony of Stra8-Cre mice. When homozygous, the mice were bred with R26R reporter strain [20]. Whenever Stra8 promoter is or was active, cells will become lacZ positive (figure S3). All of the experimental procedures complied with National regulations for the Care and Use of Laboratory Animals (similar to the U.S. National Research Council guidelines).

Chromatin immunoprecipitation (ChIP) real time PCR assays

To cross-link the cells, they were cultured for 10 min at room temperature in 1% formaldehyde in culture medium. Cross-linking was stopped by glycine at a final concentration of 125 mM. Cells were lysed by incubation in buffer containing 10mM Tris-HCl pH7.5, 10mM NaCl, 3mM $MgCl_2$, 0.5% IGEPAL and 1mM PMSF followed by centrifugation and incubation in lysis buffer supplemented with 1mM $CaCl_2$ and 4% IGEPAL. The chromatin was then sonicated using Branson Sonifier 250 to obtain an average DNA fragment length of ~200-500 bp. Soluble chromatin was incubated with and without (negative-control) antisera (3-5µg) directed against H3K4me3 (39159; Active Motif), H3K9me3 (07-442; Millipore), H3K9ac (07-352; Millipore) and H3K27me3 (07-449; Millipore) for 3h at 4°C and complexed with protein-A sepharose beads overnight at 4°C. Next, the beads were washed with wash buffer I (20mM Tris-HCl pH8.0, 2mM EDTA, 1.0% Triton X-100, 150mM NaCl, and 1mM PMSF) followed by washing with wash buffer II (wash buffer I supplemented with 0.1% SDS, 500mM NaCl) and wash buffer III (10mM Tris-HCl pH8.0, 1mM EDTA, 250mM LiCl, 0.5% IGEPAL, 0.5% Deoxycholate). Finally, the chromatin was eluted from the beads by adding elution buffer (25mM Tris-HCl, 10 mM EDTA, 0.5% SDS) and incubating for 1h at 65°C. Reverse cross-linking was performed by adding elution buffer containing Proteinase K followed by incubation for 1h at 65°C. The DNA was than extracted, purified and eluted in 40 µl of TE buffer using the Wizard SV Gel and PCR clean-up System (Promega). Quantification of precipitated DNA was carried out using SYBR green (Invitrogen)-based real-time qPCR amplification. Primer sequences are given in supplementary table S2. The qRT-PCR data of two biological and two independent technical replicates were calculated (n=12) and expressed as percentage of input DNA.

Results

<u>Expression analysis of Stra8 in pluripotent cell lines</u>

Immunocytochemistry (ICC) with Stra8 antibody on different pluripotent cell types (maGSCs and ESCs) from different genetic backgrounds (C57Bl/6 and FVB) revealed expression of STRA8 in all analyzed cells. Also the different subcellular localization which was recently described by Tedesco et al. (2009) [10] could be depicted (Fig. 1).
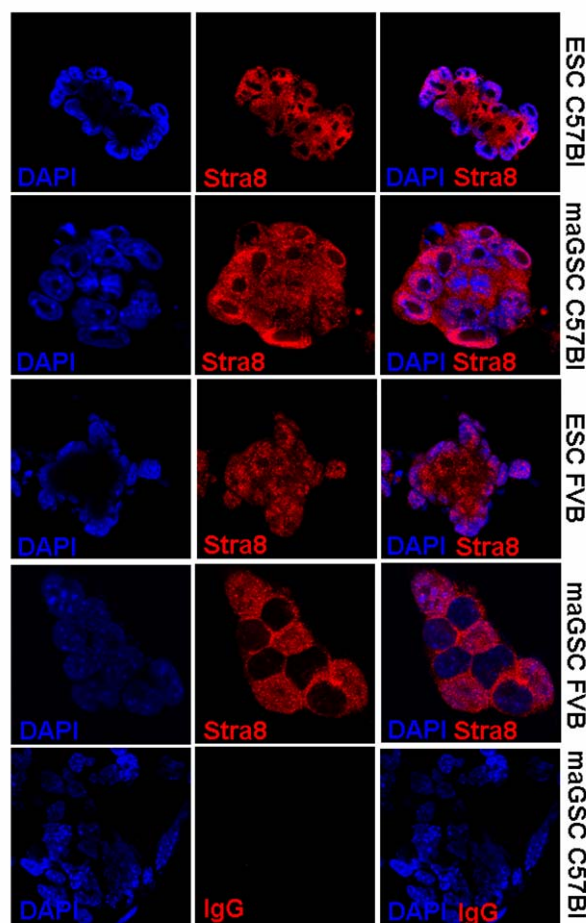

Figure 1: Expression and subcellular localization of STRA8 protein in different pluripotent cell types. Cells were analyzed by immunostaining with a STRA8 specific antibody. For comparative analysis two pluripotent cell types (ESCs and maGSCs) were analyzed from two different genetic backgrounds (C57Bl/6 and FVB). maGSCs from C57Bl/6 were used for negative control (IgG) staining. Nuclear as well as cytoplasmic localization of STRA8 protein could be detected. Magnification: 600X

Compared to OCT4, nuclear localization of STRA8 looked similar in ES-RI cells (Fig. S1). *Stra8* expression analysis on RNA level revealed that *Stra8* is downregulated upon RA induced differentiation after 20 days (Fig. 2). Because Stra8 is expressed in undifferentiated ESCs and downregulated upon differentiation, it can be assumed that Stra8 is connected to the pluripotency network.
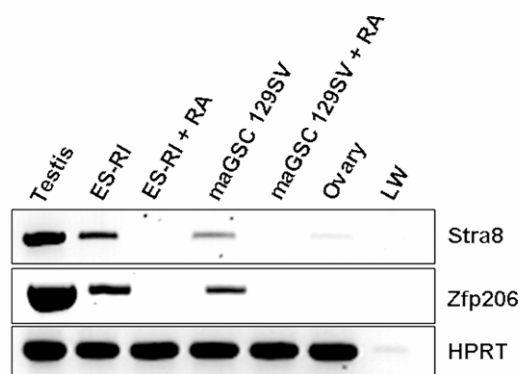
Figure 2: Expression of Stra8 in undifferentiated and differentiated ESCs and maGSCs. Conventional RT-PCR was done on cDNA of undifferentiated and differentiated ESCs and maGSCs from the same genetic background (129Sv). To induce differentiation, cells were cultivated for 20d on gelatine coated dishes in the absence of LIF and in the presence of RA (+RA). cDNA from WT-testis served as a positive control. Comparative expression of pluripotency gene *Zfp206* served as control. *Stra8* shows the same pattern as *Zfp206* – its expression is downregulated after differentiation.
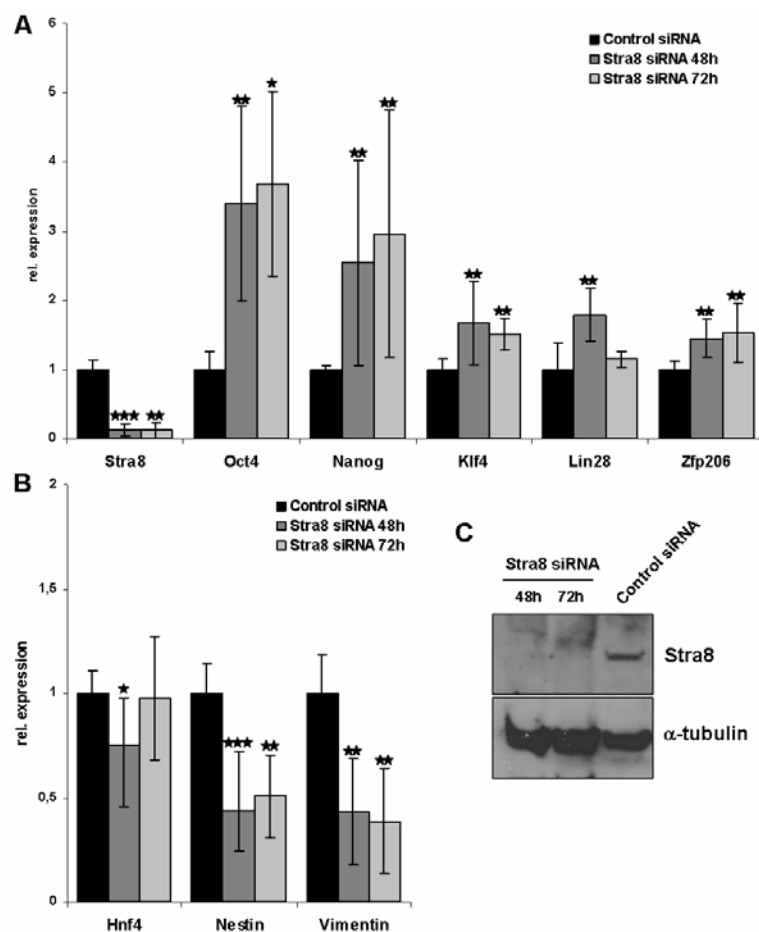
Figure 3:   Characterization of ESCs with knock down of *Stra8* expression using siRNA. (A) Expression of *Stra8* and pluripotency markers was analyzed by quantitative real-time PCR in *Stra8* siRNA treated ESCs. Data are presented as relative expression level to control siRNA transfected ESCs. (B) Relative expression of differentiation markers in the *Stra8* knock down cells. Each embryonic germ layer is represented by one marker gene (endoderm – *Hnf4*, mesoderm – *Vimentin*, ectoderm – *Nestin*) (C) Western blot using Stra8 specific antibody confirmed knockdown of STRA8 at protein level. qRT-PCR data

are depicted as the average +/- SD of two biological replicates. Statistical significances were analyzed using Mann-Whitney U-Test (*, p<0.05; **, p<0.005; ***, p<0.0005).

Influence of *Stra8* expression on the pluripotency network

To further support our hypothesis we checked the influence of Stra8 on the pluripotency network. We performed siRNA mediated knockdown of *Stra8* in ES-RI cell line. *Stra8* expression was checked by quantitative Real Time-PCR analysis (qRT-PCR) (Fig. 3A) and the efficient downregulation at the protein level was confirmed by Western blot analysis (Fig. 3C). Interestingly, the Stra8 downregulation in ESCs resulted in significant upregulation of the analyzed pluripotency genes (namely Oct4, Nanog, Klf4, Lin28 and Zfp206) (Fig. 3A) both 48h and 72h after siRNA transfection. As expected from this result, the expression of the differentiation markers Hnf4, Nestin and Vimentin was significantly decreased (Fig 3B).

These results let us ask the question, what influence *Stra8* overexpression may have in pluripotent cells. To answer this question, we constructed two vectors. In the control vector EGFP is expressed under the control of the human elongation factor 1 alpha promoter (hEF1alpha-EGFP) whereas in the second vector Stra8 is expressed under the same promoter (hEF1alpha-Stra8) (Fig. 4A). ES-RI cell were stably transfected either with the control or the Stra8 vector. Level of overexpression of different clonal cell lines was determined by qRT-PCR (data not shown). Finally, one clonal cell line, in which Stra8 showed a two fold overexpression compared to EGFP control cells, was chosen for further analysis (Fig. 4B). Coincident with the results from the downregulation, *Stra8* overexpression resulted in a significant downregulation of the pluripotency markers (Fig. 4B). Additionally, only the mesoderm differentiation markers (vimentin and eomesodermin) are significantly upregulated (Fig. 4C) suggesting a directed mesodermal differentiation of *Stra8* overexpressing ESCs. After culturing the cells for further four passages, also the morphology of the cells changed to now resembling differentiating cells (Fig. S2).
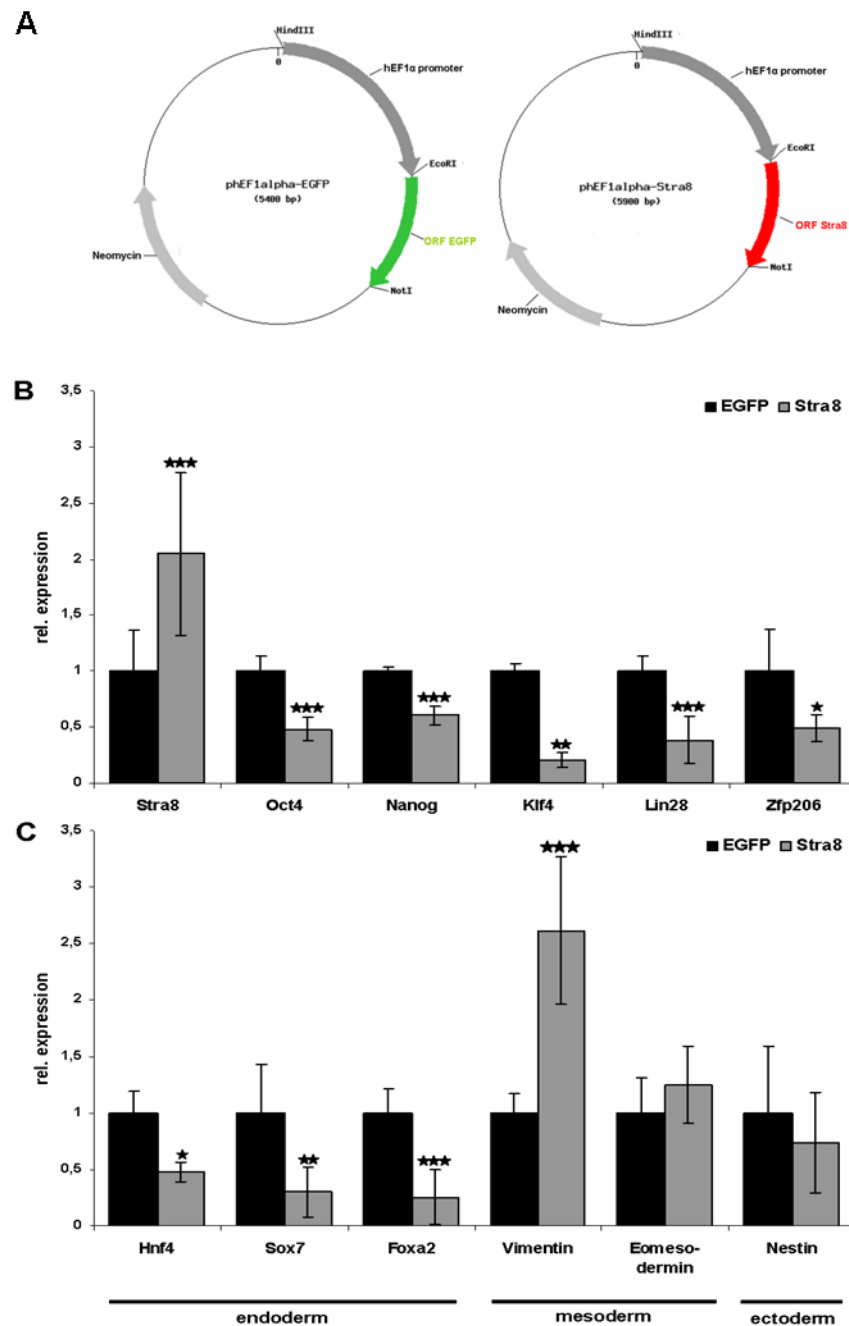
Figure4: Characterization of ESCs with stable *Stra8* overexpression under the hEF1α promoter. (A) The 1.2 kb promoter of *hEF1α* was inserted into the EGFP reporter vector pEGFP-1. This vector served as a control. For overexpression of *Stra8*, EGFP was cut out and full-length cDNA of *Stra8* was inserted. (B) Expression of *Stra8* and pluripotency markers was analyzed by quantitative real-time PCR in *Stra8* overexpressing ESCs. Data are presented as relative expression level to cells expressing EGFP with the same vector system. (C) Relative expression of some differentiation markers in the *Stra8* overexpressing cells. Markers are grouped by their affiliation to different embryonic germ layers.

qRT-PCR data are depicted as the average +/- SD of two biological replicates. Statistical significances were analyzed using Mann-Whitney U-Test (*, p<0.05; **, p<0.005; ***, p<0.0005).

Expression of _Stra8_ in preimplantation embryos

Pluripotency regulating genes are expressed in preimplantation embryos and especially in the inner cell mass (ICM) of blastocysts. Therefore we performed ICC for STRA8 on different stages of preimplantation development. STRA8 protein could be detected in all investigated stages starting from fertilized oocytes till blastocysts (Fig. 5).
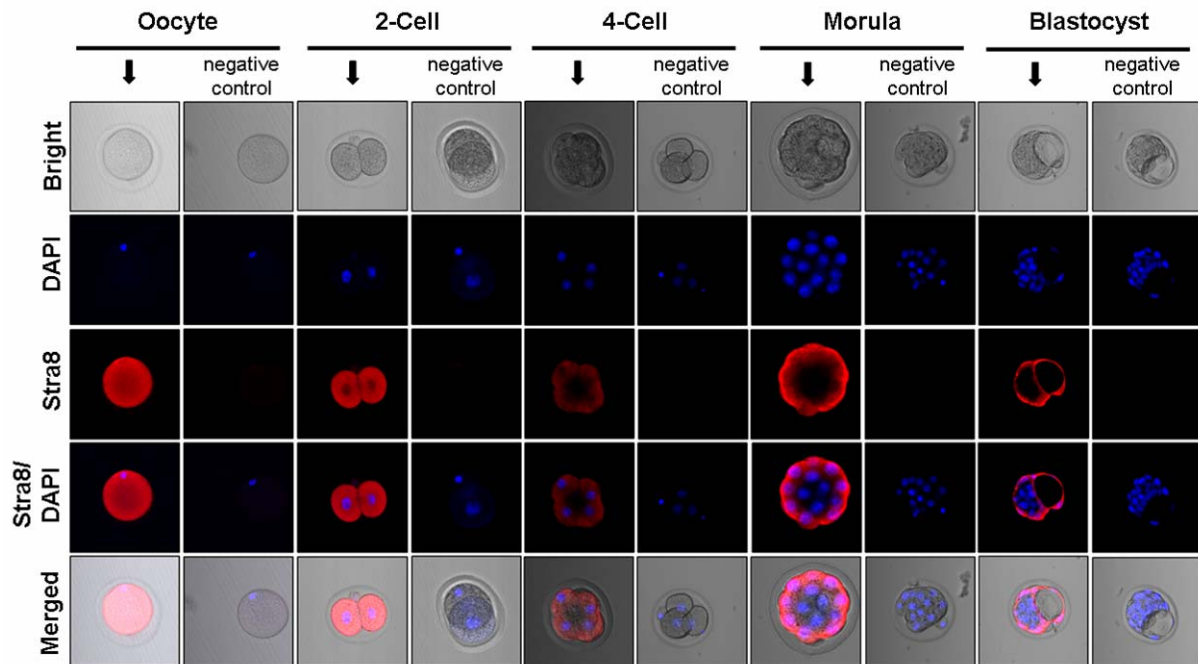
Figure 5: Expression of STRA8 in preimplantation embryos. Preimplantation embryos as indicated were stained with STRA8 specific antibody. Nuclei were stained with DAPI. Results indicate that STRA8 is already present in unfertilized oocytes and is expressed throughout pre-implantation development till blastocyst stage. Magnification: 200x

STRA8 protein seems to be localized in all stages in the cytoplasm. To substantiate the expression of STRA8 in the ICM, we repeated staining and investigated the blastocysts in a higher magnification (Fig. 6A). Here it became clear that STRA8 is also but not exclusively expressed in the ICM. Double staining with an OCT4-antibody on morula stages (Fig. 6B) revealed that STRA8 and OCT4 are expressed in the same cell but in different compartments at least in preimplantation embryos. To confirm these results we used two transgenic mouse models which were established in our lab before. The first one expresses EGFP under the control of the 1.4 kb Stra8 promoter (Fig. S3A) [18] and was used in other studies as well (e.g. [16]). Analysis of blastocysts of homozygous Stra8-EGFP breedings revealed EGFP expression in all analyzed embryos (Fig. S3B). Interestingly, EGFP expression in blastocysts was restricted to the ICM. The next transgenic mouse line used to confirm _Stra8_ expression in preimplantation development, expresses _Cre recombinase_ under the control of the same 1.4 kb promoter used in the

other mouse line (Fig. S4). Breeding of these mice with the R26R reporter strain [20] resulted in LacZ expression. LacZ staining of blastocysts from double homozygous breeding resulted in positive signal (Fig. S4C) whereas the control blastocysts (single homozygous) remained unstained (Fig. S4A+B). Also whole mount LacZ staining on 13.5 dpc embryos supported this result as embryos from the Stra8-Cre/R26R mice are lacZ positive (Fig. S4D). Taken together all these results confirmed that Stra8 is indeed expressed in preimplantation embryos and in the ICM of blastocysts even though not exclusively and not in the nucleus. These results let us suggest, that the shuttle of STRA8 to the nucleus is restricted to pluripotent cell lines [10].
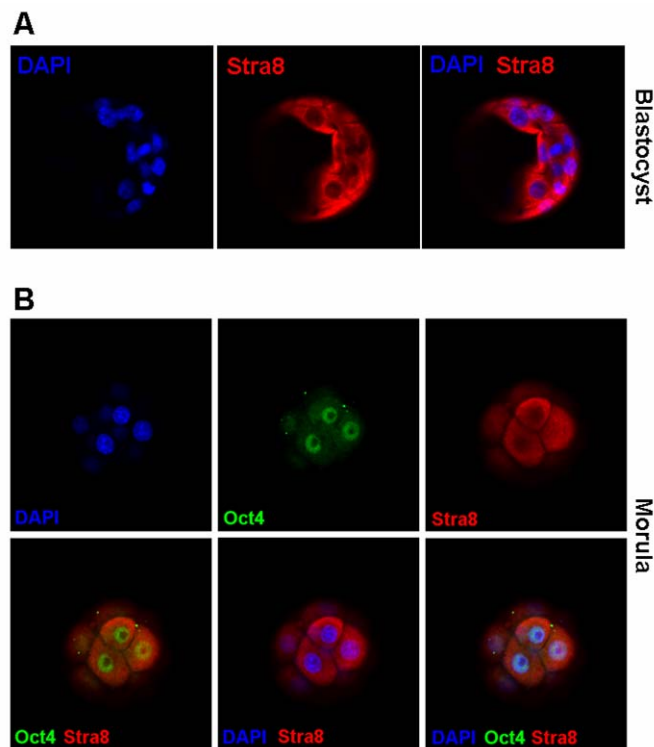

Figure 6: Expression of STRA8 and co-localization with Oct4 in the Inner Cell Mass (ICM) of the blastocyst. (A) To further substantiate expression of STRA8 in cells of the ICM, blastocyst was stained with STRA8 specific antibody and investigated with higher magnification (400x). (B) Double staining with OCT4 and STRA8 specific antibodies on morula stage supports that cells expressing OCT4 in the nucleus are expressing STRA8 in the cytoplasm. Magnification: 400x

Analysis of histone modifications on the Stra8 promoter

Next, we investigated gene specific histone modifications by chromatin immunoprecipitation (ChIP) real-time PCR analysis in ES-RI cells (Fig. 7). For pluripotency regulators in ESCs like Oct4 and Nanog it is known that the promoters of these genes are enriched for activating histone modifications including H3K4me3 and H3K9ac and depleted of repressive ones like H3K27me3 and H3K9me3 [21]. In addition, Hoxa11, an early gene for embryonic development has ESC-typical bivalent chromatin

conformation with activating (H3K4me3) as well as repressive (H3K27me3) modifications. The histone modifications on the Stra8 promoter showed a bivalent modification pattern but with a much less prominent repressing modification compared to Hoxa11.
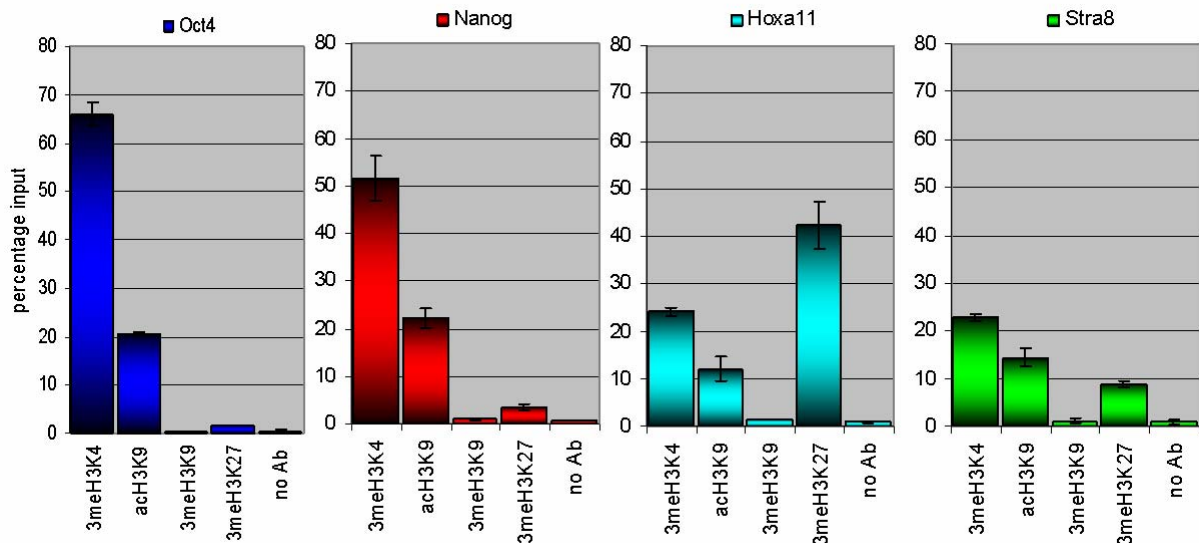
Figure 7: Analysis of histone modification levels at *Oct4*, *Nanog*, *Hoxa11* and *Stra8* promoters in ESCs. The chromatin of ES-RI lines was subjected to ChIP analysis using antibodies against four different histone modifications. The precipitated DNA was analyzed by real-time qPCR for abundance of the various modifications at the indicated promoter regions. The qPCR data are presented as percentage of input DNA. qRT-PCR data are depicted as the average +/- SD of two biological replicates.

Discussion

Germ cell specific genes became more and more important in the regulation of pluripotency. Even Oct4, the main key player of pluripotency, was first described to have a germline-specific expression in adult mice [22]. In nascent germ cells, pluripotency regulating genes like Oct4, Sox2 and Nanog are either maintained or reactivated during germ cell specification and it is suggested that the expression of these genes is essential for the establishment and the maintenance of functional totipotency in the germ cell lineage [23, 24]. More recently Sall4, a member of the SALL gene family that encodes a group of putative developmental transcription factors, was found to be exclusively expressed in the adult mouse in testis and ovary [25]. Later it was shown that murine Sall4 is essential for inner cell mass (ICM) formation and knock down of Sall4 in ESCs leads to loss of pluripotency [26]. As Sall4 acts as an activator of Oct4 and interacts with Nanog by forming a protein-protein complex, it seems that SALL4 can be a core factor for the pluripotency network [27, 28]. Recent data suggest that Sall4 is not essential for the pluripotency itself but for stabilization of ESCs by repressing aberrant trophectoderm gene expression [29].

Another example for a germ cell specific gene regulating pluripotency is the transcription factor Zfp206. This gene encodes for a zinc finger and SCAN- domain-containing protein and was found to play an important role in the control and regulation of pluripotency. The knockdown of *Zfp206* expression by RNA interference was found to induce differentiation in ESCs while the stable overexpression prevents ESCs from differentiation upon retinoic acid (RA) treatment [30-32]. It was further shown that *Zfp206* is only expressed in the testis of adult mice but not in other organs [32], suggesting again a connection between germ cell specific and pluripotency related genes. Very recently it was shown that Zfp206 physically interacts with Oct4 and Sox2, two of the main key players in pluripotency [33]. Similar to Sall4 and e, Stra8 is a germ cell specific gene which is also expressed in pluripotent cells. Therefore, we asked the question, whether germ line specific Stra8 could also have an impact on the pluripotency network. Expression analysis of different pluripotent cell types showed that *Stra8* is widely and strongly expressed and that this expression is decreased upon differentiation of stem cells (Figs. 1+2). These results are supported by several comparative transcriptome analyses in stem cells. As an example, Stra8 appears in several lists of genes that are predominantly expressed in stem cells and embryos and was classified as a pluripotency related gene [34, 35]. Stra8 was found to be downregulated in embryonic stem and embryonic germ cells after withdrawal of LIF for three days [35] supporting our result of complete shut down of expression after 20 days of differentiation (Fig. 2). STRA8 expression in early embryonic development and the co-expression with OCT4 in blastomers supports our hypothesis that *Stra8* is not restricted to male germ cells in mouse testis. Because the *Stra8* deficient mouse is not lethal but infertile [4, 8, 9], we suggest that *Stra8* has no crucial role in preimplantation development and for the establishment and maintenance of pluripotency. However, siRNA mediated knock down of *Stra8* in ESCs demonstrated a clear effect on pluripotency related genes (Fig. 3). But, as these genes are significantly upregulated and differentiation markers are downregulated, we conclude, that *Stra8* might play a role in the maintenance of the balance between pluripotency and differentiation in stem cells. To support this hypothesis, we show also the opposite effect: stable overexpression of *Stra8* in ESCs leads to downregulation of the pluripotency network (Fig.4). Interestingly, endodermal marker expression is significantly downregulated while *vimentin*, a mesodermal marker gene, is significantly upregulated and ectodermal markers are not affected. This leads us to the conclusion that a well defined level of *Stra8* expression is essential for the stabilization of stem cells and that the twofold overexpression of *Stra8* supports mesodermal differentiation. Such level-of-expression-dependency is also known for *Oct4* where a 1.5 fold overexpression of the OCT4 protein in ESCs resulted in differentiation into a population of cells expressing both endodermal and mesodermal markers [36]. Taken

together, the data from knock-down and overexpression experiments indicate, that *Stra8* is connected to the pluripotency regulating network. This is also supported by the data of Sharov et al. (2008) [37] who could identify *Stra8* as one of the genes that respond to suppression of *Oct4* and *Sox2* in ESCs as well as to the overexpression of *Nanog.* Another support for the hypothesis that Stra8 level is important for stability of pluripotency is the decoration of the Stra8 promoter with partially bivalent chromatin marks (Fig. 7). The bivalent chromatin pattern appears to be unique for pluripotent stem cells [38] and was identified at about 15 % of the examined sites [39]. Many of the genes endowed with these marks were found to encode for transcription factors that play an important role in lineage specification during embryonic development like Hoxa11. As this bivalent decoration on the Stra8 promoter is only partial, means that the repressive mark (3meH3K27) is not as dominant as for Hoxa11, we suggest that even by the chromatin modifications the level of Stra8 expression is controlled.

We have shown that Stra8 is expressed in preimplantation embryos as well as in different pluripotent cell types and becomes downregulated upon differentiation. The expression pattern could be confirmed in transgenic embryos of two different strains, as well. Functional analyses in ESCs showed that overexpression as well as knock down of Stra8 has an impact on the pluripotency network, and we suggest a role of Stra8 in the stabilization and maintenance of pluripotency.

References

[1] Bouillet, P., Oulad-Abdelghani, M., Vicaire, S. et al. (1995). Efficient cloning of cDNAs of retinoic acid-responsive genes in P19 embryonal carcinoma cells and characterization of a novel mouse gene, Stra1 (mouse LERK-2/Eplg2). Dev Biol., 170(2), 420-433.

[2] Oulad-Abdelghani, M., Bouillet, P., Décimo, D. et al. (1996). Characterization of a premeiotic germ cell-specific cytoplasmic protein encoded by Stra8, a novel retinoic acid-responsive gene. J Cell Biol., 135(2): 469-477.

[3] Menke. D.B., Koubova, J., Page, D.C. (2003). Sexual differentiation of germ cells in XX mouse gonads occurs in an anterior-to-posterior wave. Dev Biol., 262(2), 303-312.

[4] Koubova, J., Menke, D.B., Zhou, Q. et al. (2006). Retinoic acid regulates sex-specific timing of meiotic initiation in mice. Proc Natl Acad Sci U S A., 103(8), 2474-2479.

[5] Bowles, J., Knight, D., Smith, C. et al. (2006). Retinoid signaling determines germ cell fate in mice. Science., 312(5773), 596-600.

[6] Anderson, E.L., Baltus, A.E., Roepers-Gajadien, H.L. et al. (2008). Stra8 and its inducer, retinoic acid, regulate meiotic initiation in both spermatogenesis and oogenesis in mice. Proc Natl Acad Sci U S A., 105(39), 14976-14980.

[7] Zhou, Q., Li, Y., Nie, R. et al. (2008). Expression of stimulated by retinoic acid gene 8 (Stra8) and maturation of murine gonocytes and spermatogonia induced by retinoic acid in vitro. Biol Reprod., 78(3), 537-545.

[8] Baltus, A.E., Menke, D.B., Hu, Y.C. et al. (2006). In germ cells of mouse embryonic ovaries, the decision to enter meiosis precedes premeiotic DNA replication. Nat Genet., 38(12), 1430-1434.

[9] Mark, M., Jacobs, H., Oulad-Abdelghani, M. et al. (2008). STRA8-deficient spermatocytes initiate, but fail to complete, meiosis and undergo premature chromosome condensation. J Cell Sci., 121(Pt 19), 3233-3242.

[10] Tedesco, M., La Sala, G., Barbagallo, F. et al. (2009). STRA8 shuttles between nucleus and cytoplasm and displays transcriptional activity. J Biol Chem., 284(51), 35781-93.

[11] Fried, H., Kutay, U. (2003). Nucleocytoplasmic transport: taking an inventory. Cell Mol Life Sci., 60(8), 1659-1688.

[12] Gama-Carvalho, M., Carmo-Fonseca, M. (2001). The rules and roles of nucleocytoplasmic shuttling proteins. FEBS Lett., 498(2-3), 157-163.

[13] Norton, J.D. (2000). ID helix-loop-helix proteins in cell growth, differentiation and tumorigenesis. J Cell Sci., 113 (Pt 22), 3897-3905.

[14] Black, B.E., Holaska, J.M., Rastinejad, F. et al. (2001). DNA binding domains in diverse nuclear receptors function as nuclear export signals. Curr Biol., 11(22), 1749-1758.

[15] Littlewood, T.D., Evan, G.I. (1995). Transcription factors 2: helix-loop-helix. Protein Profile., 2(6), 621-702.

[16] Guan, K., Nayernia, K., Maier, L.S. et al. (2006). Pluripotency of spermatogonial stem cells from adult mouse testis. Nature., 440(7088), 1199-1203.

[17] Zovoilis, A., Nolte, J., Drusenheimer, N. et al. (2008). Multipotent adult germline stem cells and embryonic stem cells have similar microRNA profiles. Mol Hum Reprod., 14(9), 521-529.

[18] Nayernia, K., Li, M., Jaroszynski, L. et al. (2004). Stem cell based therapeutical approach of male infertility by teratocarcinoma derived germ cells. Hum Mol Genet., 13(14), 1451-1460.

[19] Gu, H., Zou, Y.R., Rajewsky, K. (1993). Independent control of immunoglobulin switch recombination at individual switch regions evidenced through Cre-loxP-mediated gene targeting. Cell., 73(6), 1155-1164.

[20] Soriano, P. (1999). Generalized lacZ expression with the ROSA26 Cre reporter strain. Nat Genet., 21(1), 70-71.

[21] Azuara, V., Perry, P., Sauer, S. et al. (2006). Chromatin signatures of pluripotent cell lines. Nat Cell Biol., 8(5), 532-538.

[22] Schöler, H.R., Hatzopoulos, A.K., Balling, R. et al. (1989). A family of octamer-specific proteins present during mouse embryogenesis: evidence for germline-specific expression of an Oct factor. EMBO J., 8(9), 2543-2550.

[23] Pesce, M., Schöler, H.R. (2001). Oct-4: gatekeeper in the beginnings of mammalian development. Stem Cells., 19(4), 271-278.

[24] Surani, M.A., Hayashi, K., Hajkova, P. (2007). Genetic and epigenetic regulators of pluripotency. Cell., 128(4), 747-762.

[25] Kohlhase, J., Heinrich, M., Liebers, M. et al. (2002). Cloning and expression analysis of SALL4, the murine homologue of the gene mutated in Okihiro syndrome. Cytogenet Genome Res., 98(4), 274-277.

[26] Zhang, J., Tam, W.L., Tong, G.Q. et al. (2006). Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. Nat Cell Biol., 8(10), 1114-1123.

[27] Yang, J., Chai, L., Fowles, T.C. et al. (2008). Genome-wide analysis reveals Sall4 to be a major regulator of pluripotency in murine-embryonic stem cells. Proc Natl Acad Sci U S A., 105(50), 19756-19761.

[28] Liang, J., Wan, M., Zhang, Y. et al. (2008). Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. Nat Cell Biol., 10(6), 731-739.

[29] Yuri, S., Fujimura, S., Nimura, K. et al. (2009). Sall4 is essential for stabilization, but not for pluripotency, of embryonic stem cells by repressing aberrant trophectoderm gene expression. Stem Cells., 27(4), 796-805.

[30] Zhang, W., Walker, E., Tamplin, O.J. et al. (2006). Zfp206 regulates ES cell gene expression and differentiation. Nucleic Acids Res., 34(17), 4780-4790.

[31] Wang, Z.X., Teh, C.H., Kueh, J.L. et al. (2007). Oct4 and Sox2 directly regulate expression of another pluripotency transcription factor, Zfp206, in embryonic stem cells. J Biol Chem., 282(17), 12822-12830.

[32] Wang, Z.X., Kueh, J.L., Teh, C.H. et al. (2007). Zfp206 is a transcription factor that controls pluripotency of embryonic stem cells. Stem Cells., 25(9),2173-2182.

[33] Yu, H.B., Kunarso, G., Hong, F.H. et al. (2009). Zfp206, Oct4, and Sox2 are integrated components of a transcriptional regulatory network in embryonic stem cells. J Biol Chem., 284(45), 31327-31335.

[34] Sharov, A.A., Piao, Y., Matoba, R. et al. (2003). Transcriptome analysis of mouse stem cells and early embryos. PLoS Biol., 1(3), E74.

[35] Sharova, L.V., Sharov, A.A., Piao, Y. et al. (2007). Global gene expression profiling reveals similarities and differences among mouse pluripotent stem cells of different origins and strains. Dev Biol., 307(2), 446-459.

[36] Niwa, H., Miyazaki, J., Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat Genet., 24(4), 372-376.

[37] Sharov, A.A., Masui, S., Sharova, L.V. et al. (2008). Identification of Pou5f1, Sox2, and Nanog downstream target genes with statistical confidence by applying a novel algorithm to time course microarray and genome-wide chromatin immunoprecipitation data. BMC Genomics., 9, 269.

[38] Bernstein, B.E., Mikkelsen, T.S., Xie, X. et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell., 125(2), 315-326.

[39] Mikkelsen, T.S., Ku, M., Jaffe, D.B. et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature., 448(7153), 553-560.

Manuscript submission

Manuscript Number: STCR-302

Title: Germ cell specific gene Stra8 has an impact on the pluripotency network

Article Type: Full-length Reports

Corresponding Author: Jessica Nolte, PhD

Corresponding Author's Institution:

First Author: Jessica Nolte, PhD

Order of Authors: Jessica Nolte, PhD;Sandra Meyer;Tatjana Khromov;Romy Graner;Linda Kotzenberg;Kaomei Guan;Arvind Dev;Krishna Pantakani D.V.;Wolfgang Engel

Suggested Reviewers: Stefan Schlatt
Stefan.Schlatt@ukmuenster.de
specialist for male germ cells

David Conrad Page
dcpage@wi.mit.edu
Specialist for Stra8

Frank Buchholz
buchholz@mpi-cbg.de
Stemcells

Opposed Reviewers: Hans Robert Schöler
schoeler@mpi-muenster.mpg.de
personal conflict to one of the coauthors

4. Discussion

As described above (2.1.1), ESCs are pluripotent cells that possess the capacity to self-renew unlimitedly and to differentiate into derivatives of all three germ layers *in vitro* and *in vivo*. Because of this plasticity, ESC-therapies promise great advantages in regenerative medicine. However, the use of ESCs in therapy is limited due to ethical and immunological reasons. The generation of maGSCs, i.e. pluripotent stem cells derived from an adult organism, is an important advance in overcoming these obstacles, since it allows obtaining patient-specific pluripotent cell lines without the use of embryos. However, until now the features and characteristics of maGSCs are incompletely defined. To figure out whether maGSCs are similar or even identical to ESCs, we performed analyses of their gene expression pattern. For this purpose we studied their properties at transcriptome and proteome level. We could show that nearly no significant differences exist between both cell types, especially the genes known to be involved in the regulation of pluripotency were found to be identically expressed and present at the same level.

However, considering the fact that the mechanisms and pathways that control pluripotency and stem cell self-renewal are still largely unknown, it is very important to identify new pluripotency-regulating genes. In the present study, we were able to recognize Stra8 as a gene putatively involved in the maintenance of pluripotency. We characterized the expression profile of Stra8 and analyzed the subcellular localization of the protein encoded by this gene. We further described its role in pluripotency by performing Stra8-overexpression and -downregulation studies *in vitro*. Herewith, we could find that a change in the expression level of Stra8 affects not only the expression of known pluripotency-regulating genes but also the expression of several lineage-specific genes.

4.1 Identification of putative pluripotency-regulating genes using transcriptome analyses

Transcriptional profiling is not only useful for comparing two or more cell lines to evaluate their overall similarities and differences but it can also be applied for the identification of transcripts that are characteristic of one cell line. To recognize a gene expression profile specific for pluripotent stem cells, several studies were performed that compared ESCs to differentiated cells (e.g. Bhattacharya et al., 2004; Brandenberger et al., 2004), to other pluripotent and multipotent stem cells (e.g. Ivanova et al., 2002; Ramalho-Santos et al., 2002; Tanaka et al., 2002) or to ESCs derived from other species (e.g. Ginis et al., 2004). Several groups reported about the comparison of human ESCs and terminally differentiated cells. Bhattacharya et al. (2004) analyzed six human ESC-lines and RNA pooled from different human organs with whole genome oligonucleotide microarrays.

Herewith, it was possible to identify 92 genes specifically expressed in all the analyzed ESC-lines. These genes included known pluripotency-regulating factors, like Oct4 (Nichols et al., 1998; Niwa et al., 2000) and Nanog (Chambers et al., 2003; Mitsui et al., 2003), as well as transcripts for ribosomal proteins and DNA-repair enzymes, modulators of WNT- and retinoic acid-signaling and cell cycle regulators (Bhattacharya et al., 2004). Differences between human ESCs and differentiated subpopulations derived from them could also be found in FGF- and Nodal-signaling which components were strongly expressed in undifferentiated cells. The presence of this variety of signal transduction factors suggests that ESCs are able to respond to diverse cellular signals while maintaining an undifferentiated state (Brandenberger et al., 2004).

Comparison of ESCs to multipotent stem cell types resulted in identification of gene clusters that are equally expressed in different types of stem cells, independent of their potency, but also of several genes that are specifically expressed in one cell type (Ivanova et al., 2002; Ramalho-Santos et al., 2002; Tanaka et al., 2002). For example, 216 genes were found to be strongly expressed in mouse hematopoietic, neural and embryonic stem cells. These genes could be classified in various biological categories with most genes playing a role in signal transduction, stress response with upregulated DNA repair and protein folding, as well as chromatin remodeling (e.g. DNA methylases, DNA helicases, histone deacetylases) (Ramalho-Santos et al., 2002). However, only four out of these genes were not detectable at all in terminally differentiated cells, providing an indication that the properties of stem cells are caused rather by the simultaneous enrichment of several genes than by the presence of individual factors (Ramalho-Santos et al., 2002).

The approach described above was also taken by Ginis et al. (2004), who compared human and mouse ESCs to identify genes that can distinguish between both cell types. For this, they used cDNA microarrays containing 96 genes specific for cell cycle and apoptosis. Expression of genes involved in cell cycle showed that both cell lines proliferate, and that they are distinguishable only by their expression of different MCM genes (Ginis et al., 2004). The factors encoded by these genes bind to the DNA after mitosis and induce DNA-replication (Lei and Tye, 2001; Nishitani and Lygerou, 2002). However, larger differences could be found in genes involved in apoptosis. Although only a small number of apoptosis-related genes were expressed at all, only 8% of these transcripts were present in ESCs from both species. Many of the genes which were only detectable in mouse cells belong to the p53-pathway. This pathway responds to DNA damage and induces cell growth inhibition and cell death (Gottifredi et al., 2000; Schultz et al., 2000). In contrast to this, human cells expressed factors which inhibit the function of p53 (Ginis et al., 2004). Although mouse and human ESCs both show similarities in

expression of marker genes for pluripotency, they can be distinguished by differences in expression of FoxD3, which is crucial for pluripotency in mouse ESCs but not detectable and therefore dispensable in human ESCs (Ginis et al., 2004). Moreover, mouse and human ESCs differ in the expression of embryonic antigens (Henderson et al., 2002) and the LIF-receptor (Carpenter et al., 2004; Ginis et al., 2004; Rosler et al., 2004). Differences in gene expression between mouse and human ESCs could also be found by Bhattacharya et al. (2004), who compared the human ESC-specific genes identified in their study with previously published genes specific for mouse ESCs (Ivanova et al., 2002; Ramalho-Santos et al., 2002; Tanaka et al., 2002). Only 13-35% of mouse "stemness" genes are also overexpressed in human ESCs, indicating quite large differences in gene expression in ESCs derived from different species (Bhattacharya et al., 2004). These differences in gene expression might be explained by the fact that human and mouse ESCs are derived from different developmental stages, what is demonstrated by the fact that human ESCs are able to differentiate into trophectoderm lineage, whereas mouse ESCs are not (reviewed in Odorico et al., 2001). Another explanation might be that mouse ESCs are treated with LIF to inhibit differentiation, whereas human ESCs keep their pluripotent state without administration of LIF (reviewed in Kujik et al., 2010).

Taken together, these results show that signature profiles for individual cell lines or cell types can be developed applying transcriptional profiling. In the present study, we therefore aimed at the identification of genes specifically expressed in pluripotent stem cells, which might play a role in the regulation of pluripotency or self-renewal. Therefore, we combined the approaches of the previously published studies mentioned above, analyzing two different types of pluripotent stem cells with their respective differentiated counterparts instead of either analyzing two different types of stem cells or ESCs and differentiated cell lines derived from them. Since we presume that the mechanisms and pathways that control pluripotency are identical in different pluripotent cell lines from one species, with this approach it should be possible to more specifically identify genes whose products are involved in regulation of pluripotency.

To follow this approach, we reanalyzed the results of the whole genome transcriptional profiling experiments we performed (3.1) and concentrated on the genes whose expression was downregulated at least twofold during differentiation of both ESCs and maGSCs. We could identify 795 genes that fulfilled this criterion. To narrow down the number of potential candidate genes, we reviewed the literature for properties of known pluripotency-regulating genes. We found that most of these factors play a role in transcriptional regulation (e.g. reviewed in: Chen and Daley, 2008; Jaenisch and Young, 2008; Kim et al., 2008). For example, the core pluripotency-regulators, Oct4, Nanog and Sox2, all act as transcription factors, and genome-wide binding site analysis revealed that

they interact with a high number of genes (Boyer et al., 2005; Loh et al., 2006). In addition, a large number of other transcription factors, like Esrrb, Tbx3, Foxd3, Klf4, Sall4 (Ivanova et al., 2006; Zhang et al., 2006; Jiang et al., 2008; Liu and Labosky, 2008; van den Berg et al., 2008; Zhang et al., 2008; Han et al., 2010; Zhang et al., 2010) and others, were recently identified to play a role in pluripotency. A summary of several known pluripotency-regulating transcription factors and their protein domains are shown in table 1.

Table 1: 14 exemplary chosen pluripotency-regulating genes that function via transcriptional regulation. The protein domains present in the gene products as predicted using Pfam database (http://pfam.sanger.ac.uk/) are presented as are the publications in which the respective gene's role in pluripotency regulation were first described.
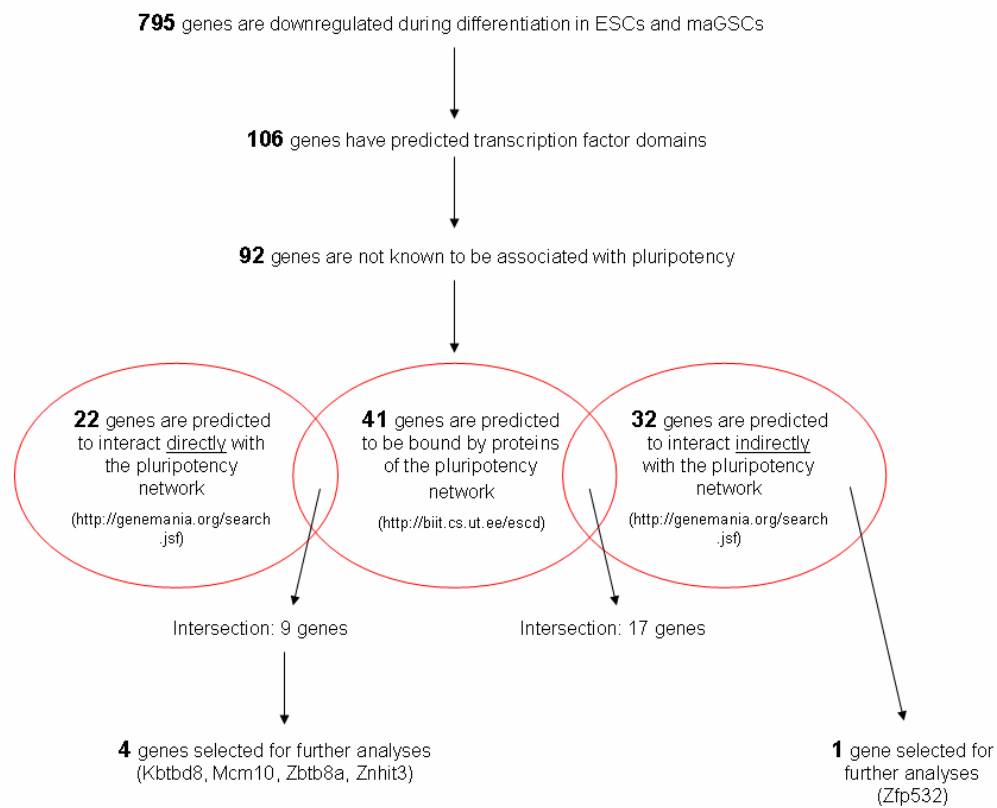
Symbol	NCBI ID	Name	Protein domains involved in transcriptional regulation (predicted with Pfam database; http://pfam.sanger.ac.uk)	Literature
Esrrb	26380	estrogen related receptor, beta	zinc finger, C4 type	Ivanova et al., 2006; van den Berg et al., 2008; Zhang et al., 2008
Foxd3	15221	forkhead box D3	fork head domain	Liu and Labosky, 2008
Klf4	16600	kruppel-like factor 4 (gut)	zinc finger, C2H2 type	Jiang et al., 2008; Zhang et al., 2010
Lin28	380669	lin-28 homolog (C. elegans)	'Cold-shock' DNA-binding domain	Xu et al., 2009
Nanog	71950	nanog homeobox	homeobox domain	Chambers et al., 2003; Mitsui et al., 2003
Oct4	18999	POU domain, class 5, transcription factor 1	homeobox domain, POU domain	Nichols et al., 1998; Niwa et al., 2000
Sall4	99377	sal-like 4 (Drosophila)	zinc finger, C2H2 type	Zhang et al., 2006
Sox2	20674	SRY-box containing gene 2	HMG box	Avilion et al., 2003
Tbx3	21386	t-box 3	T-box	Ivanova et al., 2006; Han et al., 2010
Zfp281	226442	zinc finger protein 281	zinc finger, C2H2 type	Wang et al., 2008
Zfp42	22702	zinc finger protein 42	zinc finger, C2H2 type	Scotland et al., 2009
Zfx	22764	zinc finger protein x-linkes	zinc finger, C2H2 type	Galan-Caridad et al., 2007

Table 1: Continued

| Zic3 | 22773 | zinc finger protein of the cerebellum 3 | zinc finger, C2H2 type | Lim et al., 2007 |
| Zscan10 | 332221 | zinc finger and SCAN domain containing 10 | zinc finger, C2H2 type | Wang et al., 2007a; Yu et al., 2009 |

Because of these properties of pluripotency-regulating genes, we examined the presence of protein domains known to be specific for transcription factors using Pfam database (Finn et al., 2010; http://pfam.sanger.ac.uk/) in the 795 genes identified in our transcriptional profiling. After this restriction, 689 genes that were not transcription factors could be neglected, so that 106 putative pluripotency-regulating genes remained. Within this cluster were 14 genes that are already known to play a role in the regulation of pluripotency, so that 92 putative candidate genes remained. These 92 candidate genes were subsequently classified using GeneMania (http://www.genemania.org/search.jsf) according to their predicted physical interactions with proteins of the pluripotency-network or co-expression with these respective proteins. GeneMania is a tool that helps to identify genes that are related to a set of input genes by using published functional association data, which includes interactions at protein and gene level, co-expression and co-localization of the proteins in question (Warde-Farley et al., 2010). Using Embryonic Stem Cells database (http://biit.cs.ut.ee/escd) it was additionally checked whether a binding of Oct4, Nanog, Sox2 or Klf4 occurred at gene level. The transcription factor binding data in the database was assembled from different published ChIP-PET-, ChIP-Chip- and ChIP-seq-experiments (e.g. Boyer et al., 2005; Chen et al., 2008). These analyses revealed that 22 of the candidate proteins are predicted to directly interact with proteins of the pluripotency-network, namely Oct4, Nanog, Sox2, Zscan10, Zfp296, Lin28 or Klf4, whereas 32 genes indirectly interact with these proteins. Furthermore, 41 of the 92 genes are bound by the transcription factors Oct4, Nanog, Sox2 or Klf4 at DNA level. From these genes, 5 were selected for further characterization (Fig. 2). These genes are: Zbtb8a, which directly interacts with Oct4 and indirectly with Sox2 and Nanog at protein level and is bound by Oct4 at DNA level; Kbtbd8, which colocalizes with Zscan10 and Nanog and its promoter is additionally bound by Nanog; Mcm10, which is coexpressed with Oct4, Nanog, Sox2, Utf1 and Lin28 and its promoter is also bound by Nanog; Znhit3 is also bound by Nanog at DNA level and the gene is coexpressed with Oct4, Nanog and Sox2; and finally Zfp532 which interacts indirectly with Oct4 and Sox2 (Table 2).

795 genes are downregulated during differentiation in ESCs and maGSCs

106 genes have predicted transcription factor domains

92 genes are not known to be associated with pluripotency

22 genes are predicted to interact <u>directly</u> with the pluripotency network

(http://genemania.org/search.jsf)

41 genes are predicted to be bound by proteins of the pluripotency network

(http://biit.cs.ut.ee/escd)

32 genes are predicted to interact <u>indirectly</u> with the pluripotency network

(http://genemania.org/search.jsf)

Intersection: 9 genes   Intersection: 17 genes

4 genes selected for further analyses (Kbtbd8, Mcm10, Zbtb8a, Znhit3)

1 gene selected for further analyses (Zfp532)

Figure 2: Schematic overview showing the procedure for the identification of putative pluripotency-regulating genes. The genes which were found to be downregulated during differentiation of ESCs and maGSCs were analyzed for the presence of transcription factor domains. Genes, whose encoded products are known to be involved in the regulation of pluripotency, were excluded. With the remaining factors, database searches evaluating their predicted interactions with known pluripotency-regulating genes at protein and DNA level were performed. Several genes were selected for further analyses.

Table 2: Characteristics of the genes selected for further analyses after following the strategy for identification of putative pluripotency-regulating genes described above. The fold changes of expression during differentiation of ESCs and maGSCs are shown. The interactions of the encoded proteins with components of the pluripotency network and the binding of pluripotency-associated transcription factors at DNA level are presented.

| Symbol	NCBI ID	Name	Fold change ESC – ESC diff	Fold change maGSC – maGSC diff	Interaction with proteins of pluripotency network (predicted with GeneMania)	Binding of transcription factors at gene level (predicted with Embryonic Stem Cells Database)
Kbtbd8	243574	kelch repeat and BTB (POZ) domain containing 8	2.5	2.3	Nanog, Zscan10	Nanog

Table 2: Continued

Mcm10	70024	minichromosome maintenance deficient 10 (S. cerevisiae)	2.8	2.4	Lin28, Nanog, Oct4, Sox2, Utf1	Nanog
Zbtb8a	73680	zinc finger and BTB domain containing 8a	2.9	2.9	Oct4, via Mnat1 with Nanog and Sox2	Oct4
Zfp532	328977	zinc finger protein 532	2.4	2.3	via Pou2f1 with Sox2, via Ewsr1 or Ilf2 or Sumo1 with Oct4	none
Znhit3	448850	zinc finger, HIT type 3	1.7	2.6	Nanog, Oct4, Sox2	Nanog

These genes should be further characterized according to their gene expression pattern in differentiated and undifferentiated pluripotent stem cells and in different adult tissues and organs. We presumed that pluripotency-regulating genes are strongly expressed in pluripotent cell lines, whereas they are downregulated during differentiation and therefore not detectable in differentiated cell lines as well as in adult organs. However, for several pluripotency-regulating genes an expression in spermatogonia of adult testis has been described, e.g. Sall4 (Kohlhase et al., 2002), Tex19 (Kuntz et al., 2008), Zscan10 (Wang et al., 2007a). Therefore, expression in testis was not a criterion for exclusion of genes from the study.

First, we performed conventional RT-PCR to validate the whole genome microarray results and to check whether the gene expression of Kbtbd8, Mcm10, Zbtb8a, Zfp532 and Znhit3 decreased during differentiation of the cells with RA for 20 days (Fig. 3A). We could find a strong decrease in expression in ESCs and in maGSCs after differentiation in all analyzed candidate genes except Znhit3 and Zbtb8a. In case of Zbtb8a in ESCs the expression level remained constant before and after differentiation, whereas a significant decrease could be observed in maGSCs. For Znhit3 no expression at all could be detected in undifferentiated ESCs, whereas strong expression was present in differentiated ESCs. Although these differences between microarray- and RT-PCR-data could be found, we next performed expression analyses using conventional RT-PCR for all the candidate genes in adult mouse organs (Fig. 3B). The results of the gene expression profiling revealed that Kbtbd8 and Mcm10 are specifically expressed in adult mouse testis, whereas the other candidate genes were present in all analyzed organs, thus providing an indication of their ubiquitous expression. Since their expression was not

restricted to pluripotent cell lines, it was unlikely that Zbtb8a, Zfp532 and Znhit3 play a critical role in the regulation of pluripotency. Therefore, they were excluded from our set of candidate genes. To next evaluate whether the expression of Kbtbd8 and Mcm10 can be found in the spermatogonial stem cell population in testis, we performed conventional RT-PCR-analyses with cDNA derived from testicular developmental stages (Fig. 3C). Kbtbd8 could be detected from day 5 onwards, which shows that its expression starts in spermatogonia. In contrast to this, expression of Mcm10 could only be found from day 15 onwards, what suggests an expression in spermatocyte stage.
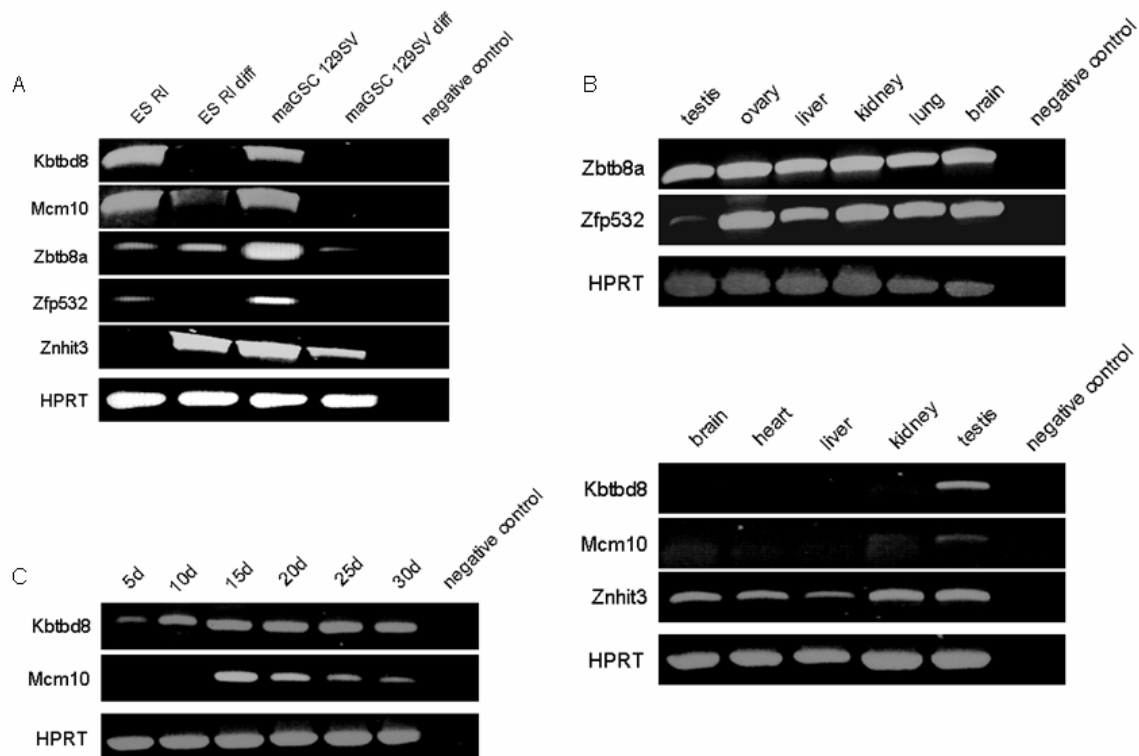
Figure 3: Analyses of gene expression of the 5 putative pluripotency-regulating genes identified in our whole genome transcriptional profiling using conventional RT-PCR. (A) Expression analyses in undifferentiated (ES RI and maGSC 129SV) and 20d differentiated ESCs and maGSCs derived from 129SV mouse background (ES RI diff and maGSC 129SV diff). (B) Expression analyses in different adult mouse organs. (C) Expression analyses in testis developmental stages (from 5 days after birth to 30 days after birth). Amplification of HPRT is used as a positive control for presence of cDNA in analyzed samples.

Next, we reviewed the literature for published data about functional aspects of Kbtbd8 and Mcm10. Mcm10 is described as belonging to the highly conserved family of mini-chromosome maintenance (MCM) proteins that play a role in the initiation of DNA-replication in eukaryotes (Homesley et al., 2000). Mcm10 is localized at replication origins of DNA and might be involved in the formation of replication forks and the recruitment of other DNA replication related proteins, like all six subunits of the Mcm2-Mcm7-complex (Lei and Tye, 2001). Mcm10 is not only extensively classified in literature but it was also

found to be highly abundant in mouse ESCs before (Ginis et al., 2004). This and the fact, that we did not find Mcm10 to be present in spermatogonial stem cells of adult testis, led us to conclude that it does not play a role in pluripotency-regulation. In contrast to this, murine Kbtbd8 is not described in the literature until now. Only one publication exists that associates human KBTBD8 with idiopathic short stature (ISS) in a Korean population (Kim et al., 2010). Because the gene has not yet been functionally classified and its expression is specific for pluripotent cells and stem cell populations in the adult body, we concluded that this gene might be a marker gene for the pluripotent state. However, gene expression analyses of Kbtbd8 that show its predominant expression in pluripotent stem cells do not provide sufficient evidence to deduce that this gene plays a role in the regulation of pluripotency. For that purpose, functional studies, like downregulation of the gene's expression using RNA interference (RNAi), have to be performed. With this combination of expression profiling and RNAi-experiments Ivanova et al. (2006) were able to identify several genes involved in the regulation of pluripotency (Ivanova et al., 2006).

Using a two-color whole genome microarray system (Agilent, Santa Clara, USA) (data deposited at GEO database at http://www.ncbi.nlm.nih.gov/geo (accession number GSE15861)), it was also possible in the present work to identify the downregulation of transcripts for the transcription factors Morc1, Stra8 and Zfp819 as well as for Lrrc34, which contains a protein-protein-interaction domain, during differentiation. The properties of Stra8 are described in chapter 3.4 and will be further discussed in chapter 4.4. The functional characterization of Morc1 was performed in a diploma thesis (Stolp, 2009) and the latter two proteins are presently analyzed in the context of two individual PhD-theses and will therefore not be discussed here.

4.2 Identification of putative pluripotency-regulating genes using proteome analyses

The goal of proteomics is identical to that of transcriptional profiling, namely the quantitative comparative description of gene expression in two or more cell lines. However, the advantage of describing the overall similarities and differences of cell lines or cells under different culture conditions at protein level is that the protein is the functional unit of the cell on which cell behavior depends, whereas the mRNA quantified in transcriptional profiling is just an intermediate stage.

To identify the characteristics of pluripotent cell lines, several proteomic studies using 2D-gelelectrophoresis (2-DE) and subsequent mass spectrometry were performed (Baharvand et al., 2007; Buhr et al., 2007; Baharvand et al., 2008). Although herewith protein reference maps specific for ESCs were generated, with none of the studies it was possible to identify known pluripotency-regulating genes. In our proteomics study

comparing undifferentiated and differentiated ESCs and maGSCs (Dihazi et al., submitted; 3.3) it was only possible to identify 18 proteins that were specific for both pluripotent cell lines. One of these was the known pluripotency-regulating transcription factor Trim28 (Fazzio et al., 2008; Seki et al., 2010). Trim28 is known to be involved in chromatin remodeling by inducing formation of heterochromatin during differentiation of stem cells (Kurisaki et al., 2005). When Trim28 is present in mouse ESCs, it was found to be phosphorylated at the C-terminal serine residue 824. The phosphorylated form associates with Oct4 and subunits of the ESC-specific chromatin remodeling complex esBAF. This complex localizes at active chromatin marks and induces the expression of ESC-specific genes, thereby maintaining the undifferentiated state of ESCs (Seki et al., 2010). To identify further putative pluripotency-regulating genes we concentrated on the 18 proteins found to be downregulated during differentiation in both pluripotent stem cell lines. However, these proteins are mostly well characterized and were already found to be downregulated during differentiation of ESCs. For example, the decrease in expression in Eno1, Eno3, Hspd1 and Stmn1 was also found in the study of Baharvand et al. (2008), who compared two different mouse ESC-lines with their differentiated counterparts after 16 days of culture in the absence of a feeder layer and LIF (Baharvand et al., 2008). Hspd1 belongs to the family of heat shock proteins and functions as a component of the protein-folding system located in the mitochondrial matrix. However, it could also be found at the cell surface (Soltys and Gupta, 1997; Soltys and Gupta, 1999; Nunomura et al., 2005). The downregulation of Hspd1 was not only reported by Baharvand et al. (2008) but was also found by several other groups who performed proteomic studies (Kurisaki et al., 2005; Wang and Gao, 2005; van Hoof et al., 2006). This high abundance in different proteomic studies suggests that Hspd1 can be used as a marker for undifferentiated ESCs, although it does not play a role in the regulation of pluripotency. In contrast to this, Stmn1 encodes a cytosolic phosphoprotein which is responsible for cell proliferation in different lineages (Rowlands et al., 1995). The role of Stmn1 during differentiation of pluripotent cells is not yet clear, however, it was proposed that it regulates cell division by instabilizing the interphase microtubules leading to formation of the mitotic spindle (Holmfeldt et al., 2001; Iancu et al., 2001). Eno1 and Eno3 are proteins involved in energy metabolism. The downregulation of these proteins reveals a highly active metabolism in pluripotent cells (Baharvand et al., 2008). Next to these proteins, we could identify different components of ribonucleoproteins (RNPs) (Hnrpab, Hnrpf, Hnrnph1, Hnrpdl) and nucleolins (Nup50, Nup62). RNPs are associations of proteins and RNAs. The proteins identified here possess a RNA-binding domain and are implicated in performing a function in RNA-splicing (reviewed in Han et al., 2010). Nucleoporins act as the main structural components of the nuclear pore complex, which is a structure that regulates the flow of

macromolecules between nucleus and cytoplasm (reviewed in Wälde and Kehlenbach, 2010).

As described above, it was neither possible to identify proteins containing transcription factor domains nor known pluripotency-regulating factors, except for Trim28. Because of these results, we decided to repeat our proteomic analyses after prefractionation of the cells. For this, we analyzed the nuclear proteomes of differentiated and undifferentiated ESCs and maGSCs. We expected to get an enrichment of transcription factors with this methodology. This approach was also taken by Buhr et al. (2008) who compared undifferentiated ESCs and EGCs with MEFs. They could find 30 proteins that were present only in pluripotent cell lines, of which 60% had a nuclear localization. 17 of these proteins were previously identified in other proteomic studies. However, among these 30 proteins only one was known to be involved in the regulation of pluripotency, namely Dppa5 (Tanaka et al., 2006). In our study, we could find 59 proteins which were downregulated during differentiation. Also in this experiment around 60% of proteins were in fact localized in the nucleus, indicating a nuclear enrichment according to percentages found in literature (Buhr et al., 2008). Surprisingly, it was not possible to identify known pluripotency regulators among these 59 proteins, not even Trim28, which could be found in our complete cell proteome analysis. However, again several RNPs (Hnrnpab, Hnrnpa3, Hnrpdl, Hnrnph1, Hnrnpk, Hnrnpl, Hnrnpa2b1), nucleoporins (Nup62, Nup43, Nup54) and histone clusters (Hist1h2af, Hist2h2ac, Hist1h2bb) could be identified.

Using other proteomic methods without 2D-gelelectrophoresis, it was generally possible to generate reference maps from pluripotent stem cells with a much larger number of identified proteins (e.g. Nagano et al., 2005; van Hoof et al., 2006). Nagano et al. (2005) described the proteomic analysis of murine ESCs. For this, they isolated whole protein from cultured ESCs and subjected it directly to tryptic digestion. These digested proteins were then analyzed using a micro-scale 2-D LC MS/MS. Herewith it was possible to identify a subset of 1790 proteins, including a variety of transcription factors (Nagano et al., 2005). This subset of transcription factors also contained known pluripotency-regulating proteins, namely Sox2 (Avilion et al., 2003), Utf1 (Nishimoto et al., 2005) and Oct4 (Nichols et al., 1998; Niwa et al., 2000). In conformance with our data, they could identify a high number of nuclear proteins like nucleolins, histones and components of heterogenous nuclear RNPs. The expression levels of these proteins were found to be at the same level as those of known cytosolic house-keeping proteins (Nagano et al., 2005), which are the most abundant proteins in other cell types (Jacobs et al., 2004). These results indicate that pluripotent stem cells express exceptionally high levels of these nuclear proteins, which might be explained by the fact that ESCs form colonies with a high nuclear-cytoplasmic ratio (Nagano et al., 2005). A similar method without performance of

2D-geleclectrophoresis was taken by van Hoof et al. (2006). They isolated whole protein from mouse and human ESCs as well as from their differentiated derivatives, performed 1D SDS-PAGE and cut each lane in 26 parts which were digested by incubation with trypsin. Finally, the digested peptides were analyzed by nanoflow liquid chromatography and FT-ICR-MS/MS (van Hoof et al., 2006). Herewith, 1871 proteins could be identified in undifferentiated mouse ESCs and 1552 proteins in their differentiated counterparts. From these set of proteins, 743 could be uniquely identified in mouse pluripotent stem cells. These subset included not only known pluripotency-regulating factors like Oct4 (Nichols et al., 1998; Niwa et al., 2000) and Utf1 (Nishimoto et al., 2005) but also other proteins characteristic for ESCs like alkaline phosphatase (ALPL). Comparable to the result of our proteomic analysis (Dihazi et al., submitted; 3.3), van Hoof et al. (2006) could find an enrichment of the heat shock protein Hspd1 in undifferentiated mouse ESCs.

Although a limited number of identified proteins overlapped between our study (Dihazi et al., submitted; 3.3) and the published large-scale proteomics approaches (Nagano et al., 2005; van Hoof et al., 2006), our results clearly show that the methodology applied by us contains limitations in matters of the quantity of proteins and the types of proteins which can be identified. We hypothesized that the lack of known pluripotency-regulating genes found not only in our study but also in other published studies (Baharvand et al., 2007; Buhr et al., 2007; Baharvand et al., 2008) is due to the limitations of 2D-gelelectrophoresis. 2-DE couples isoelectric focusing (IEF) using immobilized pH-gradients (IPGs) in the first dimension with SDS-PAGE in the second dimension. Thereby, it enables the separation of a protein mixture according to the molecular mass, isoelectric point (pI) and relative abundance of a single protein (reviewed in: Görg et al., 2004). However, limitations in 2-DE include the detection of proteins which strongly differ in their abundance in one sample, detection of very acidic or very basic proteins and detection of low-abundance or membrane proteins (Görg et al., 2009). Proteins which are differentially expressed are hardly detectable, because the minor spots, representing low-abundance proteins, can be obscured by stronger ones, which are generated by high abundance proteins with similar molecular mass and isoelectric point (Anderson and Anderson, 1998). Very acidic or very basic proteins possess a low or a high pI, respectively, which are separated in the IEF. For this step, typically IPGs spanning a gradient between pI 3-10 are used, so that proteins with a lower or higher pI are not detectable. In addition, when analyzing total protein derived from eukaryotic cells using a single wide-range pH gradient, only a small percentage of the complete proteome can be identified. The same is true for proteins with different molecular masses: proteins with a very high or low molecular mass are not detectable due to the resolution of the SDS-PAGE (Görg et al., 2009). Membrane proteins, which are usually very hydrophobic proteins, are also

underrepresented on 2D-gels. This can be explained by the fact that hydrophobic proteins possess a low solubility in aqueous solution which leads to aggregation and precipitation of the proteins. However, even if the proteins are solubilized and properly separated according to their pI during IEF, it is possible that the proteins will not elute from the IPG-strip during the transfer from first to second dimension (Görg et al., 2009). Many membrane proteins are expressed in low copy numbers and possess basic pIs, what additionally complicates their detection (Wilkins et al., 1998; Corthals et al., 2000). To find out why it was not possible to detect known pluripotency-regulating factors using 2-DE, we analyzed their molecular masses and calculated their pIs (Table 3).

Table 3: Depiction of 30 exemplary chosen pluripotency-regulating factors. The molecular masses in kDa as indicated at NCBI (http://www.ncbi.nlm.nih.gov/protein) and the theoretical isoelectric points (pI) as calculated using ExPASY Proteomics Server (http://expasy.org/tools/pi_tool.html) are given.

Symbol	Name	UniProt KB Accession Number	Molecular mass [kDa]	Theoretical pI
Dppa4	developmental pluripotency-associated 4	Q8CCG4	32.7	8.97
Dppa5a	developmental pluripotency-associated 5a	Q9CQS7	13.8	6.15
Epha2	Eph receptor A2	Q03145	106.5	5.82
Esrrb	estrogen related receptor, beta	Q61539	48.4	8.27
Fgf4	fibroblast growth factor 4	P11403	19.0	10.24
FoxD3	forkhead box D3	Q61060	46.3	5.64
Gdf3	growth differentiation factor 3	Q3TUX1	41.6	9.44
Jarid1b	lysine(K)-specific demethylase 5b	Q80Y84	175.6	6.03
Klf2	kruppel-like factor 2 (lung)	Q3V293	37.6	8.84
Klf4	kruppel-like factor 4 (gut)	Q60793	51.9	8.66
Lin28	Lin28 homolog A (C. elegans)	Q8K3Y3	22.7	8.85
Nanog	Nanog homeobox	Q80Z64	34.2	6.32
Nr0b1	nuclear receptor subfamily 0, group B, member 1	Q61066	52.6	5.77
Oct4	POU domain, class 5, transcription factor 1	P20263	38.2	6.05
Sall4	sal-like 4 (Drosophila)	Q8BX22	113.2	8.02
Slc2a3	solute carrier family 2 (facilitated glucose transporter), member 3	P32037	53.5	4.95
Sox2	SRY-box containing gene 2	P48432	34.5	9.81
SSEA1	stage-specific embryonic antigen 1	Q11127	49.5	9.95
Tbx3	T-box 3	P70324	79.3	8.05

Table 3: Continued

Tcf3	transcription factor 3	P15806	67.7	5.94
Tcl1	T-cell lymphoma breakpoint 1	P56280	14.1	5.36
Tex19.1	testis expressed gene 19.1	Q99MV2	40.4	4.61
Trap1a	tumor rejection antigen P1A	P19473	25.8	3.83
Utf1	undifferentiated embryonic cell transcription factor 1	Q6J1H4	36.4	10.05
Zfp42	zinc finger protein 42	P22227	32.4	8.42
Zfp143	zinc finger protein 143	O70230	69.0	5.74
Zfp281	zinc finger protein 281	Q99LI5	96.7	8.77
Zfx	zinc finger protein x-linked	P17012	90.0	5.75
Zic3	zinc finger protein of the cerebellum 3	Q3UYV1	50.4	8.89
Zscan10	zinc finger and SCAN domain containing 10	Q3URR7	88.4	8.37

Herewith it became clear that 18 of these 30 known pluripotency regulating proteins possess pIs that are not in the range of the used IPG-strip, namely between 5 and 8, whereas the molecular masses of the proteins lie within the detection range of the SDS-PAGE gel. While the lack of these 18 proteins can be explained by the limitations of the used IPG-strip, the absence of the other 12 proteins can only be explained by their low abundance or by other methodological limitations. However, first preliminary results generated by prefractionation of pluripotent stem cells and subsequent analyses of their nuclear proteomes using 1D SDS-PAGE followed by mass spectrometry as described by van Hoof et al. (2006) revealed the possibility to identify a significantly higher quantity of proteins, including transcription factors and several known pluripotency-regulating genes (data not shown). This analysis will be subsequently extended to prefractionated differentiated cells to identify a reliable pluripotent stem cell-specific proteome.

4.3 Similarities and differences in genes identified in transcriptome and proteome analyses

Despite the obvious limitations of proteome analysis applying 2-DE described above (4.2), we wanted to evaluate similarities and differences in genes identified in transcriptomics and proteomics. Therefore, we compared the proteins found to be downregulated during differentiation in maGSCs and ESCs in proteomics with those found in transcriptomics: in the comparison between the 795 genes whose expression was downregulated in transcriptomics and the 18 proteins which were higher abundant in whole cell lysates from undifferentiated cells we could identify only two factors that were found to be similarly regulated with both methods. These factors were Eif5a, which we propose to play an

important role in stem cell proliferation and differentiation (Dihazi et al., submitted; 3.3), and Eno1, which is a subunit of the phosphopyruvate dehydratase, an enzyme required for glycolysis. The downregulation of Eno1 was previously recognized by Baharvand et al. (2008). Next we compared the 795 'transcriptomic-genes' with the 59 downregulated proteins from our nuclear proteome analysis and found 14 commonly regulated factors (Table 4).

Table 4: List of genes whose expression was found to be downregulated during differentiation in the transcriptional profiling experiments and in the analyses of nuclear proteomes in both maGSCs and ESCs.

Symbol	NCBI ID	Name
Actl6a	56456	actin-like 6A
Hist2h2ac	319176	histone cluster 2, H2ac
Grwd1	101612	glutamate-rich WD repeat containing 1
C1qbp	12261	complement component 1, q subcomponent binding protein
Polr2e	66420	polymerase (RNA) II (DNA directed) polypeptide E
Nup43	69912	nucleoporin 43
Cbx1	12412	chromobox homolog 1 (Drosophila HP1 beta)
Hspe1	15528	heat shock protein 1 (chaperonin 10)
Phb	18673	prohibitin
Pa2g4	18813	proliferation-associated 2G4
Sod2	20656	superoxide dismutase 2, mitochondrial
Eif5a	276770	eukaryotic translation initiation factor 5A
Lsm2	27756	LSM2 homolog, U6 small nuclear RNA associated (S. cerevisiae)
Hist1h2bb	319178	histone cluster 1, H2bb

This list again includes the proliferation-inducing factor Eif5a, several histone clusters and the nucleoporin Nup43 (4.2). However, in summary it can be concluded that the overlap in genes whose expression is downregulated during differentiation found in transcriptomics and in different proteomics approaches is very low.

It is obvious that a much higher number of differentially expressed genes could be detected using transcriptional profiling, whose products could not be detected at all using our proteomic approach. The same phenomenon was also observed by Nagano et al. (2005) who performed a large-scale proteomic approach and compared their results with previously published transcriptional profiling data (Ramalho-Santos et al., 2002). For them, it was only possible to identify the products of 60 genes among 485 transcripts

identified in transcriptional profiling (Nagano et al., 2005). They argued that their proteome-analysis covered only a subset of the cells' complete proteomes whereas the transcriptional profiling discovered all transcripts present in a cell line. Another explanation could be that the translation of mRNA into protein is a highly regulated process, so that changes in mRNA expression do not necessarily reflect changes in protein levels (e.g. Anderson and Seilhamer, 1997; Nagano et al., 2005; Unwin et al., 2006). A low correlation between data derived at RNA- and protein level was also found when performing Northern blots for detection of differentially expressed RNA and high-performance liquid chromatography (HPLC) for detection of proteins (Tew et al., 1996). This indicates that the obtained results are not due to methodological problems with transcriptomics and proteomics (Anderson and Anderson, 1998). Additionally, it was found by performing transcriptional profiling of ESCs that these cells express transcripts associated with differentiation (e.g. Ginis et al., 2004). These transcripts were present in a very low amount, but they were detectable by transcriptional profiling using microarray analyses or serial analysis of gene expression (SAGE) (Anisimov et al., 2002; Sharov et al., 2003; Brandenberger et al., 2004; Richards et al., 2004). However, these transcripts are not necessarily translated into proteins, and when they are, they are too low abundant to be detected using 2-DE. In contrast to this, presence of gene products that can be detected by proteomics but not by transcriptional profiling might be due to the fact that proteins from FCS cannot be completely eliminated when performing proteomics from cultured cells (Miller et al., 2006; Buhr et al., 2007). FCS is used in cell culture to keep the cells in a proliferative state. It contains a wide variety of proteins, however, its exact composition is not known. Buhr et al. (2007) could show that around 7% of their identified proteins were derived from serum in the cell culture media. These proteins could not be disregarded even after four washing steps with PBS (Buhr et al., 2007). Moreover, it was shown by Miller et al. (2006) that even extensive washing cannot eliminate all FCS-proteins but changes the cells' protein expression profile because of induction of stress (Miller et al., 2006). These summarized results clearly show that due to the differences in the resolution limits of the applied methods and the uncommon properties of pluripotent stem cells the divergences in identified gene products observed in our studies are not surprising.

4.4 Functional characterization of the putative pluripotency-regulating gene Stra8

The murine gene for Stra8 (stimulated by retinoic acid gene 8) is located on chromosome 6 and consists of 10 exons. Stra8 is highly conserved in amniotes at gene as well as at protein level. Prediction of protein domains using PSIPRED-database (McGuffin et al.

(2000; http://www.psipred.net/psiform.html) for Stra8 revealed the presence of a N-terminal Helix-loop-helix- (HLH) domain and three nuclear export signals (NES) (Tedesco et al., 2009). The HLH-domain is found in transcription factors, and HLH-proteins function in the regulation of gene expression, cell cycle and cell differentiation (Norton, 2000). The HLH-domains often include functional sequence motifs which are responsible for shuttling of proteins between nucleus and cytoplasm (Black et al., 2001). This is also the case in the HLH-domain of Stra8 which includes a nuclear localization signal (NLS) (Tedesco et al., 2009).

Stra8 was first identified in P19 embryonic carcinoma cells as a gene upregulated in response to induction with RA (Bouillet et al., 1995). In the mouse, expression of Stra8 could be detected in the male developing gonad from 12.5 dpc onwards and in cytoplasm of spermatogonia and preleptotene spermatocytes in the adult testis (Oulad-Abdelghani et al., 1996). Later, Menke et al. (2003) could show that Stra8 is also expressed in mouse embryonic ovaries between 12.5 dpc and 16.5 dpc (Menke et al., 2003). Male as well as female Stra8$^{-/-}$ mice are infertile because Stra8-deficient germ cells do not progress through meiosis (Baltus et al., 2006; Koubova et al., 2006). Stra8 expression could also be found in different pluripotent cell lines, like ESCs (Oulad-Abdelghani et al., 1996), F9 teratocarcinoma cells (Oulad-Abdelghani et al., 1996) and maGSCs (Guan et al., 2006) and its expression strongly increased upon short time-stimulation with RA. Stra8-promoter activity was previously used as a tool to enrich spermatogonial stem cell populations either from adult mouse testis suspensions (Guan et al., 2006) or from cultured pluripotent stem cells (Nayernia et al., 2004; Nayernia et al., 2006; Nolte et al., 2010). In the first case, a transgenic mouse carrying a Stra8/EGFP-promoter construct was used. Testis cell suspensions derived from these mice were subjected to fluorescence activated cell sorting (FACS) to obtain an enriched spermatogonial stem cell (SSC) population which could be recognized by its green fluorescence. This cell population was subsequently reprogrammed to pluripotency by culture of the cells in ESC-specific conditions and called maGSCs (Guan et al., 2006). In the second case, mouse pluripotent cell lines, namely teratocarcinoma cells, ESCs and maGSCs, were stably transfected with a Stra8/EGFP-promoter construct. After RA-induction the EGFP-positive cells were sorted out by FACS. Thereby so-called SSC lines could be generated from pluripotent cells. These SSCs were subsequently able to undergo meiosis and generate haploid male gametes that were functional. SSCs derived from teratocarcinoma cells were injected into germ cell-depleted mouse testis and could give rise to mature spermatozoa *in vivo*. These spermatozoa were functional and were able to fertilize oocytes after intracytoplasmatic sperm injection (ICSI) (Nayernia et al., 2004). In case of SSCs derived from ESCs (Nayernia et al., 2006) and maGSCs (Nolte et al., 2010), the differentiation towards male gametes was pursued *in*

vitro and led to the generation of haploid cells. These haploid cells fertilized oocytes after ICSI, embryonic development progressed and offspring was generated.

These studies clearly show that Stra8 has an important role in germ cell development and the initiation of meiosis *in vivo* and *in vitro*. However, recent findings provide evidence that Stra8 has an additional function in pluripotent stem cells as well as in germ cells. Stra8 protein was first described to localize in cytoplasm (Oulad-Abdelghani et al., 1996). However, Tedesco et al. (2009) could show that in the developing ovary at 13.5 dpc in around 50% of cells Stra8 is present in cytoplasm as well as in nucleus, whereas it is exclusively present in nucleus in the other 50% of cells. In pluripotent stem cells Stra8 is even predominantly present only in the nuclear compartment (Tedesco et al., 2009). This occurrence in different cellular compartments could be explained by the presence of the NLS inside Stra8's HLH-domain, which is responsible for the shuttle between nucleus and cytoplasm (Black et al., 2001). This shuttle between cellular compartments is common for proteins that act as transcriptional regulators and gives a hint that Stra8 might also act as a transcription factor or a transcriptional co-regulator that interacts with other transcription factors (Tedesco et al., 2009). These studies reveal that Stra8 fulfills all the requirements for a protein involved in the regulation of pluripotency mentioned above (4.1). These are, in brief, the presence of a transcription factor domain and therefore nuclear localization of the protein and the restricted expression in pluripotent cell lines and premeiotic stages of germ cells in adult testis. Additionally, using whole genome microarray analyses we could show that the expression of Stra8 is strongly downregulated during spontaneous RA-induced differentiation of the cells, confirming the exclusive expression of Stra8 in stem cell populations. Further indications for the role of Stra8 in pluripotency regulation are the results of our knockdown and overexpression experiments. They showed that changes in Stra8 protein levels influenced not only expression of genes required for maintaining the pluripotent state but also expression of lineage-specific genes (Nolte et al., submitted; 3.4). These results reveal that Stra8 likely plays a role in maintaining the balance between self-renewal and differentiation in pluripotent stem cells.

In pluripotent stem cell lines Stra8 was found to be associated with chromatin (Tedesco et al., 2009). However, it is not clear whether this association is due to a direct or an indirect interaction. This renders possible that Stra8 might be involved in modulating chromatin structure during differentiation of stem cells and might be associated with subunits of chromatin remodeling complexes. This mechanism was also found for the pluripotency-regulating transcription factor Trim 28, which promotes a stem cell specific gene expression profile by binding to active chromatin marks in undifferentiated cells and inducing heterochromatin formation during differentiation of stem cells (Kurisaki et al., 2005; Seki et al., 2010). Taken together, the results described above reveal that Stra8 not

only plays a role in the initiation of germ cell development and the regulation of pluripotency but may also play a role in epigenetic regulation of gene expression. However, the possible function of Stra8 in chromatin remodeling and epigenetic regulation has to be further analyzed.

4.5 Future endeavors and perspectives

The present study provides first insights in the gene expression profile of maGSCs at RNA and protein level. It was possible to show that maGSCs are nearly identical to ESCs at transcriptome and proteome level. Additionally, it could be shown that the cell types differentiate spontaneously in the same directions. While analyzing the changes in gene expression during differentiation we were able to identify several transcription factors that might act as pluripotency regulators. One of these candidate genes, Stra8, was extensively characterized by performing functional studies, like knockdown and overexpression of Stra8-protein. Another putative pluripotency-regulating gene, the transcription factor Kbtbd8, has been studied based on its gene expression profile in different cell lines, organs and testicular developmental stages. Based on its predominant expression in pluripotent stem cells, we propose that Kbtbd8 possesses a role in the regulation of pluripotency. However, since expression analysis alone does not provide enough evidence for conclusions about the functional properties of a protein, it is necessary to study its role in pluripotency and self-renewal of stem cells. This could be done by using siRNA-mediated knockdown of Kbtbd8 to analyze the changes in gene expression and properties of pluripotent stem cells after downregulation of the putative pluripotency-regulating gene. This approach was undertaken by Ivanova et al. (2006) who were able to identify several regulators of self-renewal which were unknown until then, namely Esrrb, Tbx3 and Tcl1 (Ivanova et al., 2006). In contrast to this, overexpression studies of Kbtbd8 should be performed in pluripotent stem cells as well as in differentiated cell lines to check whether the cells keep their pluripotent state under differentiation-promoting conditions or are reprogrammed to a pluripotent state, respectively. Interaction partners of the protein should be determined to find potential association with proteins of the pluripotency network. In our approach to identify genes playing a role in pluripotency-maintenance we concentrated on proteins which are downregulated during differentiation and additionally include a transcription factor domain. We took this approach because it is known that transcription factors play an important role in maintaining pluripotency. However, further studies showed that regulatory proteins and functional RNAs, like histone modulators, DNA-methylases and siRNAs, are differentially expressed between ESCs and their differentiated derivatives (Lei et al., 1996; Bernstein et al., 2003; Lee et al., 2004). Additionally, ESCs possess a high abundance of DNA repair factors (Sato et

al., 2003). This provides an indication that these factors might act as molecular switches that activate gene expression programs required for pluripotency and differentiation (Miura et al., 2004). Therefore, the results of the transcriptional profiling should be reanalyzed to identify chromatin modifying factors and DNA repair proteins that are downregulated during differentiation. These proteins might play a role in regulating pluripotency and defining the stem cell state. Finally, the proteomic analyses should be extended to large-scale approaches, thereby providing more insights into pluripotent stem cell specific proteomes. These proteomic approaches could be used to identify abundance of proteins or posttranslational modifications involved in signal transduction pathways. Herewith, the presence and activity of pathways specific for pluripotent stem cells could be analyzed.

5. References

Anderson EL, Baltus AE, Roepers-Gajadien HL, Hassold TJ, de Rooij DG, van Pelt AM, Page DC (2008) Stra8 and its inducer, retinoic acid, regulate meiotic initiation in both spermatogenesis and oogenesis in mice. Proc Natl Acad Sci USA 105(39): 14976-14980

Anderson L, Seilhamer J (1997) A comparison of selected mRNA and protein abundances in human liver. Electrophoresis 18(3-4): 533-537

Anderson NL, Anderson NG (1998) Proteome and proteomics: New technologies, new concepts, and new words. Electrophoresis 19(11):1853-1861

Anisimov SV, Tarasov KV, Tweedie D, Stern MD, Wobus AM, Boheler KR (2002) SAGE identification of gene transcripts with profiles unique to pluripotent mouse R1 embryonic stem cells. Genomics 79(2): 169-176

Avilion AA, Nicolis SK, Pevny LH, Perez L, Vivian N, Lovell-Badge R (2003) Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev 17(1): 126-140

Baharvand H, Fathi A, van Hoof D, Salekdeh GH (2007) Trends in stem cell proteomics. Stem Cells 25(8): 1888-1903

Baharvand H, Fathi A, Gourabi H, Mollamohammadi S, Salekdeh GH (2008) Identification of mouse embryonic stem cell-associated proteins. J Proteome Res 7(1): 412-423

Baltus AE, Menke DB, Hu YC, Goodheart ML, Carpenter AE, de Rooij DG, Page DC (2006) In germ cells of mouse embryonic ovaries, the decision to enter meiosis precedes premeiotic DNA replication. Nat Genet 38(12): 1430-1434

Bernstein E, Kim SY, Carmell MA, Murchison EP, Alcorn H, Li MZ, Mills AA, Elledge SJ, Anderson KV, Hannon GJ (2003) Dicer is essential for mouse development. Nat Genet 35(3): 215-217

Bhattacharya B, Miura T, Brandenberger R, Mejido J, Luo Y, Yang AX, Joshi BH, Ginis I, Thies RS, Amit M, Lyons I, Condie BG, Itskovitz-Eldor J, Rao MS, Puri RK (2004) Gene expression in human embryonic stem cell lines: unique molecular signature. Blood 103(8): 2956-2964

Black BE, Holaska JM, Rastinejad F, Paschal BM (2001) DNA binding domains in divers nuclear receptors function as nuclear export signals. Curr Biol 11(22): 1749-1758

Bouillet P, Oulad-Abdelghani M, Vicaire S, Garnier J-M, Schuhbaur B, Dolle P, Chambon P (1995) Efficient cloning of cDNAs of retinoic acid-responsive genes in P19 embryonal carcinoma cell and characterization of a novel mouse gene, Stra1 (mouse LERK-2/Epig2). Dev Biol 170(2): 420-433

Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, Gifford DK, Melton DA, Jaenisch R, Young RA (2005) Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 122(6): 947-956

Brandenberger R, Wei H, Zhang S, Lei S, Murage J, Fisk GJ, Li Y, Xu C, Fang R, Guegler K, Rao MS, Mandalam R, Lebkowski J, Stanton LW (2004) Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation. Nat Biotechnol 22(6): 707-716

Brinster RL, Avarbock MR (1994) Germline transmission of donor haplotype following spermatogonial transplantation. Proc Natl Acad Sci USA 91(24): 11303-11307

Brons IG, Smithers LE, Trotter MW, Rugg-Gunn P, Sun B, Chuva de Sousa Lopes SM, Howlett SK, Clarkson A, Ahrlund-Richter L, Pedersen RA, Vallier L (2007) Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature 448(7150): 191-195

Buhr N, Carapito C, Schaeffer C, van Dorsselaer A, Viville S (2007) Proteome analysis of the culture environment supporting undifferentiated mouse embryonic stem and germ cell growth. Electrophoresis 28(10): 1615-1623

Buhr N, Carapito C, Schaeffer C, Kieffer E, van Dorsselaer A, Viville S (2008) Nuclear proteome analysis of undifferentiated mouse embryonic stem and germ cells. Electrophoresis 29(11): 2381-2390

Card DA, Hebbar PB, Li L, Trotter KW, Komatsu Y, Mishina Y, Archer TK (2008) Oct4/Sox2-regulated miR-302 targets cyclin D1 in human embryonic stem cells. Mol Cell Biol 28(20): 6426-6438

Carpenter MK, Rosler ES, Fisk GJ, Brandenberger R, Ares X, Miura T, Lucero M, Rao MS (2004) Properties of four human embryonic stem cell lines maintained in a feeder-free culture system. Dev Dyn 229(2): 243-258

Chambers I, Colby D, Robertson M, Nichols J, Lee S, Tweedie S, Smith A (2003) Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell 113(5): 643-655

Chen L, Daley GQ (2008) Molecular basis of pluripotency. Hum Mol Genet 17(R1): R23-27

Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, Wong E, Orlov YL, Zhang W, Jiang J, Loh YH, Yeo HC, Yeo ZX, Narang V, Govindarajan KR, Leong B, Shahab A, Ruan Y, Bourque G, Sung

WK, Clarke ND, Wei CL, Ng HH (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell 133(6): 1106-1117

Chin MH, Mason MJ, Xie W, Volinia S, Singer M, Peterson C, Ambartsumyan G, Aimiuwu O, Richter L, Zhang J, Khvorostov I, Ott V, Grunstein M, Lavon N, Benvenisty N, Croce CM, Clark AT, Baxter T, Pyle AD, Teitell MA, Pelegrini M, Plath K, Lowry WE (2009) Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. Cell Stem Cell 5(1): 111-123

Chung Y, Klimanskaya I, Becker S, Marh J, Lu SJ, Johnson J, Meisner L, Lanza R (2006) Embryonic and extraembryonic stem cell lines derived from single mouse blastomeres. Nature 439(7073): 216-219

Corthals GL, Wasinger VC, Hochstrasser DF, Sanchez JC (2000) The dynamic range of protein expression: a challenge for proteomic research. Electrophoresis 21(6): 1104-1115

Dani C, Smith AG, Dessolin S, Leroy P, Staccini L, Villageois P, Darimont C, Ailhaud G (1997) Differentiation of embryonic stem cells into adipocytes *in vitro*. J Cell Sci 110(11): 1279-1285

David S, Aldrich TH, Stahl N, Pan L, Taga T, Kishimoto T, Ip NY, Yancopoulos GD (1993) LIFRβ and gp130 as heterodimerizing signal transducers of the tripartite CNTF receptor. Science 260(5115): 1805-1808

Dinsmore J, Ratliff J, Deacon T, Pakzaban P, Jacoby D, Galpern W, Isacson O (1996) Embryonic stem cells differentiated in vitro as a novel source of cells for transplantation. Cell Transplant 5(2): 131-143

Doetschman TC, Eistetter H, Katz M, Schmidt W, Kemler R (1985) The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. J Embryol Exp Morphol 87: 27-45

Drab M, Haller H, Bychkov R, Erdmann B, Lindschau C, Haase H, Morano I, Luft FC, Wobus AM (1997) From totipotent embryonic stem cells to spontaneously contracting smooth muscle cells: a retinoic acid and db-cAMP in vitro differentiation model. FASEB J 11(11): 905-915

Evans MJ, Kaufmann MH (1981) Establishment in culture of pluripotential cells from mouse embryos. Nature 292(5819): 154-156

Fazzio TG, Huff JT, Panning B (2008) An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. Cell 134(1): 162-174

Finch BW, Ephrussi B (1967) Retention of multiple developmental potentialities by cells of a mouse testicular teratocarcinoma during prolonged culture *in vitro* and their extinction upon hybridisation with cells of permanent lines. Proc Natl Acad Sci 57(3): 615-621

Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, Holm L, Sonnhammer ELL, Eddy SR, Bateman A (2010) The Pfam protein families database. Nucl Acids Res 38(suppl 1): D211-D222

Galan-Caridad JM, Harel S, Arenzana TL, Hou ZE, Doetsch FK, Mirny LA, Reizis B (2007) Zfx controls the self-renewal of embryonic and hematopoietic stem cells. Cell 129(2): 345-357

Geijsen N, Horoschak M, Kim K, Gribnau J, Eggan K, Daley GQ (2004) Derivation of embryonic germ cells and male gametes from embryonic stem cells. Nature 427(6970): 148-154

Ginis I, Luo Y, Miura T, Thies S, Brandenberger R, Gerecht-Nir S, Amit M, Hoke A, Carpenter MK, Itskovitz-Eldor J, Rao MS (2004) Differences between human and mouse embryonic stem cells. Dev Biol 269(2): 360-380

Görg A, Drews O, Lück C, Weiland F, Weiss W (2009) 2-DE with IPGs. Electrophoresis 30(Suppl 1): S122-S132

Görg A, Weiss W, Dunn MJ (2004) Current two-dimensional electrophoresis technology for proteomics. Proteomics 4(12): 3665-3685

Gottifredi V, Shieh SY, Prives C (2000) Regulation of p53 after different forms of stress and at different cell cycle stages. Cold Spring Harbor Symp Quant Biol 65: 483-488

Guan K, Nayernia K, Maier LS, Wagner S, Dressel R, Lee JH, Nolte J, Wolf F, Li M, Engel W, Hasenfuss G (2006) Pluripotency of spermatogonial stem cells from adult mouse testis. Nature 440(7088): 1199-1203

Han J, Yuan P, Yang H, Zhang J, Soh BS, Li P, Lim SL, Cao S, Tay J, Orlov YL, Lufkin T, Ng HH, Tam WL, Lim B (2010) Tbx3 improves the germ-line competency of induced pluripotent stem cells. Nature 463(7284): 1096-1100

Han SP, Tang YH, Smith R (2010) Functional diversity of the hnRNPs: past, present and perspectives. Biochem J 430(3): 379-392

Henderson JK, Draper JS, Baillie HS, Fishel S, Thomson JA, Moore H, Andrews PW (2002) Preimplantation human embryos and embryonic stem cells show comparable expression of stage-specific embryonic antigens. Stem Cells 20(4): 329-337

Holmfeldt P, Larsson N, Segerman B, Howell B, Morabito J, Cassimeris L, Gullberg M (2001) The catastrophe-promoting activity of extopic Op18/stathmin is required for disruption of mitotic spindles but not interphase microtubules. Mol Biol Cell 12(1): 73-83

Houbaviy H, Murray M, Sharp P (2003) Embryonic stem cell-specific microRNAs. Dev Cell 5(2): 351-358

Homesley L, Lei M, Kawasaki Y, Sawyer S, Christensen T, Tye BK (2000) Mcm10 and the MCM2-7 complex interact to initiate DNA synthesis and to release replication factors from origins. Genes Dev 14(8): 913-926

Iancu C, Mistry SJ, Arkin S, Wallenstein S, Atweh GF (2001) Effects of stathmin inhibition on the mitotic spindle. J Cell Sci 114(Pt 5): 909-916

Ivanova NB, Dimos JT, Schaniel C, Hackney JA, Moore KA, Lemischka IR (2002) A stem cell molecular signature. Science 298(5593): 601-604

Ivanova N, Dobrin R, Lu R, Kotenko I, Levorse J, DeCoste C, Schafer X, Lun Y, Lemischka IR (2006) Dissecting self-renewal in stem cells with RNA interference. Nature 442(7102): 533-538

Izadyar F, Pau F, Marh J, Slepko N, Wang T, Gonzales R, Ramos T, Howerton K, Sayre C, Silva F (2008) Generation of multipotent cell lines from a distinct population of male germ line stem cells. Reproduction 135(6): 771-784

Jacobs JM, Mottaz HM, Yu LR, Anderson DJ, Moore RJ, Chen WN, Auberry KJ, Strittmatter EF, Monroe ME, Thrall BD, Camp DG 2nd, Smith RD (2004) Multidimensional proteome analysis of human mammary epithelial cells. J Proteome Res 3(1): 68-75

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

Jiang J, Chan YS, Loh YH, Cai J, Tong GQ, Lim CA, Robson P, Zhong S, Ng HH (2008) A core Klf circuitry regulates self-renewal of embryonic stem cells. Nat Cell Biol 10(3): 353-360

Kahan BW, Ephrussi B (1970) Developmental potentialities of clonal *in vitro* cultures of mouse testicular teratoma. J Natl Cancer Inst 44(5): 1015-1036

Kanatsu-Shinohara M, Ogonuki N, Inoue K, Miki H, Ogura A, Toyokuni S, Shinohara T (2003) Long-term proliferation in culture and germline transmission of mouse male germline stem cells. Biol Reprod 69(2): 612-616

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(7): 1001-1012

Kanatsu-Shinohara M, Lee J, Inoue K, Ogonuki N, Miki H, Toyokuni S, Ikawa M, Nakamura T, Ogura A, Shinohara T (2008) Pluripotency of a single spermatogonial stem cell in mice. Biol Reprod 78(4): 681-687

Kim J, Chu J, Shen X, Wang J, Orkin SH (2008) An extended transcriptional network for pluripotency of embryonic stem cells. Cell 132(6): 1049-1061

Kim JJ, Lee HI, Park T, Kim K, Lee JE, Cho NH, Shin C, Cho YS, Lee JY, Han BG, Yoo HW, Lee JK (2010) Identification of 15 loci influencing height in a Korean population. J Hum Genet 55(1): 27-31

Kleinsmith LJ, Pierce Jr. GB (1964) Multipotentiality of single embryonal carcinoma cells. Cancer Res 24: 1544-1551

Klimanskaya I, Chung Y, Becker S, Lu SJ, Lanza R (2006) Human embryonic stem cell lines derived from single blastomeres. Nature 444(7118): 481-485

Kohlhase J, Heinrich M, Liebers M, Fröhlich-Archangelo L, Reardon W, Kispert A (2002) Cloning and expression analysis of SALL4, the murine homologue of the gene mutated in Okihiro syndrome. Cytogenet Genoome Res 98(4): 274-277

Ko K, Tapia N, Wu G, Kim JB, Bravo MJ, Sasse P, Glaser T, Ruau D, Han DW, Greber B, Hausdörfer K, Sebastiano V, Stehling M, Fleischmann BK, Brüstle O, Zenke M, Schöler HR (2009) Induction of pluripotency in adult unipotent germline stem cells. Cell Stem Cell 5(1): 87-96

Koubova J, Menke DB, Zhou Q, Capel B, Griswold MD, Page DC (2006) Retinoic acid regulates sex-specific timing of meiotic initiation in mice. Proc Natl Acad Sci USA 103(8): 2474-2479

Kujik EW, Chuva de Sousa Lopes SM, Geijsen N, Macklon N, Roelen BAJ (2010) The different shades of mammalian pluripotent stem cells. Hum Reprod Update, doi: 10.1093/humupd/dmq035

Kuntz S, Kieffer E, Bianchetti L, Lamoureux N, Fuhrmann G, Viville S (2008) Tex19, a mammalian-specific protein with a restricted expression in pluripotent stem cells and germ line. Stem Cells 26(3): 734-744

Kurisaki A, Hamazaki TS, Okabayashi K, Iida T, Nishine T, Chonan R, Kido H, Tsunasawa S, Nishimura O, Asashima M, Sugino H (2005) Chromatin-related proteins in pluripotent mouse embryonic stem cells are downregulated after removal of leukemia inhibitory factor. Biochem Biophys Res Commun 335(3): 667-675

Lee JH, Hart SR, Skalnik DG (2004) Histone deacetylase activity is required for embryonic stem cell differentiation. Genesis 38(1): 32-38

Lei H, Oh SP, Okano M, Juttermann R, Goss KA, Jaenisch R, Li E (1996) De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. Development 122(10): 3195-3205

Lei M, Tye BK (2001) Initiating DNA synthesis: from recruiting to activating the MCM complex. J Cell Sci 114(Pt 8): 1447-1454

Lim LS, Loh YH, Zhang W, Li Y, Chen X, Wang Y, Bakre M, Ng HH, Stanton LW (2007) Zic3 is required for maintenance and pluripotency of embryonic stem cells. Mol Biol Cell 18(4): 1348-1358

Lin G, Ou Yang Q, Zhou X, Gu Y, Yuan D, Li W, Liu G, Liu T, Lu G (2007) A highly homozygous and parthenogenetic human embryonic stem cell line derived from a one-pronuclear oocyte following *in vitro* fertilization procedure. Cell Res 17(12): 999-1007

Liu Y, Labosky PA (2008) Regulation of embryonic stem cell self-renewal and pluripotency by Foxd3. Stem Cells 26(10): 2475-2484

Loh YH, Wu Q, Chew JL, Vega VB, Zhang W, Chen X, Bourque G, George J, Leong B, Liu J, Wong KY, Sung KW, Lee CWH, Zhao XD, Chiu KP, Lipovich L, Kuznetsov VA, Robson P, Stanton LW, Wei CL, Ruan Y, Lim B, Ng HH (2006) The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat Genet 38(4): 431-440

Lowry WE, Richter L, Yachechko R, Pyle AD, Tchieu J, Sridharan R, Clark AT, Plath K (2008) Generation of human induced pluripotent stem cells from dermal fibroblasts. Proc Natl Acad Sci USA 105(8): 2883-2888

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

Mai Q, Yu Y, Li T, Wang L, Chen MJ, Huang SZ, Zhou C, Zhou Q (2007) Derivation of human embryonic stem cell lines from parthenogenetic blastocysts. Cell Res 17(12): 1008-1019

Martin GR, Evans MJ (1975) Differentiation of clonal lines of teratocarcinoma cells: formation of embryoid bodies *in vitro*. Proc Natl Acad Sci USA 72(4): 1441-1445

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

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

McGuffin LJ, Bryson K, Jones DT (2000) The PSIPRED protein structure prediction server. Bioinformatics 16(4): 404-405

Menke DB, Koubova J, Page DC (2003) Sexual differentiation of germ cells in XX mouse gonads occurs in an anterior-to-posterior wave. Dev Biol 262(2): 303-312

Miller I, Radwan M, Strobl B, Müller M, Gemeiner M (2006) Contribution of cell culture additives to the two-dimensional protein patterns of mouse macrophages. Electrophoresis 27(8): 1626-1629

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

Nagano K, Taoka M, Yamauchi Y, Itagaki C, Shinkawa T, Nunomura K, Okamura N, Takahashi N, Izumi T, Isobe T (2005) Large-scale identification of proteins expressed in mouse embryonic stem cells. Proteomics 5(5): 1346-1361

Nayernia K, Li M, Jaroszynski L, Khusainov R, Wulf G, Schwandt I, Korabiowska M, Michelmann HW, Meinhardt A, Engel W (2004) Stem cell based therapeutical approach of male infertility by teratocarcinoma derived germ cells. Hum Mol Genet 13(14): 1451-1460

Nayernia K, Nolte J, Michelmann HW, Lee JH, Rathsack K, Drusenheimer N, Dev A, Wulf G, Ehrmann IE, Elliott DJ, 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(1): 125-132

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

Nishimoto M, Miyagi S, Yamagishi T, Sakaguchi T, Niwa H, Muramatsu M, Okuda A (2005) Oct-3/4 maintains the proliferative embryonic stem cell state via specific binding to a variant octamer sequence in the regulatory region of the UTF1 locus. Mol Cell Biol 25(12): 5084-5094

Nishitani H, Lygerou Z (2002) Control of DNA replication licensing in a cell cycle. Genes Cells 7(6): 523-534

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

Niwa H, Ogawa K, Shimosato D, Adachi K (2004) A parallel circuit of LIF signaling pathways maintains pluripotency of mouse ES cells. Nature 460(7251): 118-122

Nolte (2008) Zur Pluripotenz spermatogonialer Stammzelllinien. Doctoral thesis.

Nolte J, Meyer S, Khromov T, Graner R, Kotzenberg L, Guan K, Dev A, Pantakani KDV, Engel W (2010) Germ cell specific gene Stra8 has an impact on the pluripotency network. Manuscript in submission stage.

Nolte J, Michelmann HW, 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(4-5): 184-194

Norton JD (2000) ID helix-loop-helix proteins in cell growth, differentiation and tumorigenesis. J Cell Sci 113(Pt 22): 3897-3905

Nunomura K, Nagano K, Itagaki C, Taoka M, Okamura N, Yamauchi Y, Sugano S, Takahashi N, Izumi T, Isobe T (2005) Cell surface labeling and mass spectrometry reveal diversity of cell surface markers and signaling molecules expressed in undifferentiated mouse embryonic stem cells. Mol Cell Proteomics 4(12): 1968-1976

Pantakani KDV, Zheng Y, Khromov T, Wolf M, Meyer S, Nolte J, Zechner U, Engel W (2010) Stra8 is an epigenetic regulator: Insights from interaction partners Setd8 and Arid4b. Manuscript in submission stage.

Odorico JS, Kaufman DS, Thomson JA (2001) Multilineage differentiation from human embryonic stem cell lines. Stem Cells 19(3): 193-204

Oulad-Abdelghani M, Bouillet P, Decimo D, Gansmuller A, Heyberger S, Dolle P, Bronner S, Lutz Y, Chambon P (1996) Characterization of a premeiotic germ cell-specific cytoplasmic protein encoded by Stra8, a novel retinoic acid-responsive gene. J Cell Biol 135(2): 469-477

Ramalho-Santos M, Yoon S, Matsuzaki Y, Mulligan RC, Melton DA (2002) "Stemness": transcriptional profiling of embryonic and adult stem cells. Science 298(5593): 597-600

Resnick JL, Bixler LS, Cheng L, Donovan PJ (1992) Long-term proliferation of mouse primordial germ cells in culture. Nature 359(6395): 550-551

Revazova ES, Turovets NA, Kochetkova OD, Kindarova LB, Kuzmichev LN, Janus JD, Pryzhkova MV (2007) Patient-specific stem cell lines derived from human parthenogenetic blastocysts. Cloning Stem Cells 9(3): 432-449

Richards M, Tan SP, Tan JH, Chan WK, Bongso A (2004) The transcriptome profile of human embryonic stem cells as defined by SAGE. Stem Cells 22(1): 51-64

Rosler ES, Fisk GJ, Ares X, Irving J, Miura T, Rao MS, Carpenter MK (2004) Long-term culture of human embryonic stem cells in feeder-free conditions. Dev Dyn 229(2): 259-274

Rowlands DC, Harrison RF, Jones NA, Williams A, Hubacher SG, Brown G (1995) Stathmin is expressed by the proliferating hepatocytes during liver regeneration. Clin Mol Pathol 48(2): M88-M92

Sato N, Sanjuan IM, Heke M, Uchida M, Naef F, Brivanlou AH (2003) Molecular signature of human embryonic stem cells and its comparison with the mouse. Dev Biol 260(2): 404-413

Schultz LB, Chehab NH, Malikzay A, DiTullio Jr, RA, Stavridi ES, Halazonetis TD (2000) The DNA damage checkpoint and human cancer. Cold Spring Harbor Symp Quant Biol 65: 489-498

Scotland KB, Chen S, Sylvester R, Gudas LJ (2009) Analysis of Rex1 (zfp42) function in mouse embryonic stem cell differentiation. Dev Dyn 238(8): 1863-1877

Seandel M, James D, Shmelkov SV, Falciatori I, Kim J, Chavala S, Scherr DS, Zhang F, Torres R, Gale NW, Yancopoulos GD, Murphy A. Valenzuela DM, Hobbs RM, Pandolfi PP, Rafii S (2007) Generation of functional multipotent adult stem cells from GPR125[+] germline progenitors. Nature 449(7160): 346-350

Seki Y, Kurisaki A, Watanabe-Susaki K, Nakajima Y, Nakanishi M, Arai Y, Shiota K, Sugino H, Asashima M (2010) TIF1beta regulates the pluripotency of embryonic stem cells in a phosphorylation-dependent manner. Proc Natl Acad Sci USA 107(24): 10926-10931

Sell S (2004) Stem Cells. Stem Cell Handbook, Humana Press 1-18

Sharov AA, Masui S, Sharova LV, Piao Y, Aiba K, Matoba R, Xin L, Niwa H, Ko MSH (2008) Identification of Pou5f1, Sox2, Nanog downstream target genes with statistical confidence by applying a novel algorithm to time course microarray and genome-wide chromatin immunoprecipitation data. BMC Genomics 9: 269

Sharov AA, Piao Y, Matoba R, Dudekula DB, Qian Y, VanBuren V, Falco G, Martin PR, Stagg CA, Bassey UC, Sharova Y, Tanaka TS, Kimber WL, Yoshikawa T, Jaradat SA, Pantano S, Nagaraja R, Boheler KR, Taub D, Hodes RJ, Longo DL, Schlessinger D, Keller J, Klotz E, Kelsoe G, Umezawa A, Vescovi AL, Rossant J, Kunath T, Hogan BL, Curci A, D'Urso A, Kelso J, Hide W, Ko MS (2003) Transcriptome analysis of mouse stem cells and early embryos. PLoS Biol 1(3): E74

Sharova LV, Sharov AA, Piao Y, Shaik N, Sullivan T, Stewart CL, Hogan BLM, Ko MSH (2007) Global gene expression profiling reveals similarities and differences among mouse pluripotent stem cells of different origins and strains. Dev Biol 307(2): 446-459

Slager HG, Freund E, Buiting AM, Feijen A, Mummery CL (1993) Secretion of transforming growth factor-beta isoforms by embryonic stem cells: isoform and latency are dependent on direction of differentiation. J Cell Physiol 156(2): 247-256

Smith AG, Heath JK, Donaldson DD, Wong GG, Moreau J, Stahl M, Rogers D (1988) Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. Nature 336(6200): 688-690

Soltys BJ, Gupta RS (1997) Cell surface localization of the 60 kDa heat shock chaperonin protein (hsp60) in mammalian cells. Cell Biol Int 21(5): 315-320

Soltys BJ, Gupta RS (1999) Mitochondrial-matrix proteins at unexpected locations: are they exported. Trends Biochem Sci 24(5): 174-177

Stolp K (2009) Identifizierung und Charakterisierung putativer Pluripotenzgene. Diploma thesis. Strelchenko N, Verlinsky O, Kukharenko V, Verlinsky Y (2004) Morula-derived human embryonic stem cells. Reprod Biomed Online 9(6): 623-629

Suda Y, Suzuki M, Ikawa Y, Aizawa S (1987) Mouse embryonic stem cells exhibit indefinite proliferative potential. J Cell Physiol 133(1): 197-201

Sun C, Nakatake Y, Akagi T, Ura H, Matsuda T, Nishiyama A, Koide H, Ko MS, Niwa H, Yokota T (2009) Dax1 binds to Oct3/4 and inhibits its transcriptional activity in embryonic stem cells. Mol Cell Biol 29(16): 4574-4583

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

Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131(5): 1-12

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

Tanaka TS, Lopez de Silanes I, Sharova LV, Akutsu H, Yoshikawa T, Amano H, Yamanaka S, Gorospe M, Ko MS (2006) Esg1, expressed exclusively in preimplantation embryos, germline, and embryonic stem cells, is a putative RNA-binding protein with broad RNA targets. Dev Growth Differ 48(6): 381-390

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(51): 35781-35793

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

Tew KD, Monks A, Barone L, Rosser D, Akerman G, Montali JA, Wheatley JB, Schmidt DE Jr (1996) Glutathione-associated enzymes in the human cell lines of the national cancer institute drug screening program. Mol Pharmacol 50(1): 149-159

Thomson JA, Kalishman J, Golos TG, Durning M, Harris CP, Becker RA, Hearn JP (1995) Isolation of a primate embryonic stem cell line. Proc Natl Acad Sci USA 92(17): 7844-7848

Thomson JA, Kalishman J, Golos TG, Durning M, Harris CP, Hearn JP (1996) Pluripotent cell lines derived from common marmoset (*Callithrix jacchus*) blastocysts. Biol Reprod 55(2): 254-259

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

Van den Berg DL, Zhang W, Yates A, Engelen E, Takacs K, Bezstarosti K, Demmers J, Chambers I, Poot RA (2008) Estrogen-related receptor beta interacts with Oct4 to positively regulate Nanog gene expression. Mol Cell Biol 28(19): 5986-5995

Van Hoof D, Passier R, Ward-Van Oostwaard D, Pinkse MWH, Heck AJR, Mummery CL, Krijgsveld J (2006) A quest for human and mouse embryonic stem cell-specific proteins. Mol Cell Proteomics 5(7): 1261-1273

Wakayama S, Hikichi T, Suetsugu R, Sakaide Y, Bui HT, Mizutani E, Wakayama T (2007) Efficient establishment of mouse embryonic stem cell lines from single blastomeres and polar bodies. Stem Cells 25(4): 986-993

Wälde S, Kehlenbach RH (2010) The part and the whole: functions of nucleoporins in nucleocytoplasmic transport. Trends Cell Biol 20(8): 461-469

Wang D, Gao L (2005) Proteomic analysis of neural differentiation of mouse embryonic stem cells. Proteomics 5(17): 4414-4426

Wang ZX, Kueh JLL, Teh CHL, Rossbach M, Lim L, Li P, Wong KY, Lufkin T, Robson P, Stanton LW (2007a) Zfp206 is a transcription factor that controls pluripotency of embryonic stem cells. Stem Cells 25(9): 2173-2182

Wang ZX, Teh CHL, Kueh JLL, Lufkin T, Robson P, Stanton LW (2007b) Oct4 and Sox2 directly regulate expression of another pluripotency transcription factor, Zfp206, in embryonic stem cells. J Biol Chem 282(17): 12822-12830

Wang ZX, The CH, Chan CM, Chu C, Rossbach M, Kunarso G, Allapitchay TB, Wong KY, Stanton LW (2008) The transcription factor Zfp281 controls embryonic stem cell pluripotency by direct activation and repression of target genes. Stem Cells 26(11): 2791-2799

Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, Franz M, Grouios C, Kazi F, Lopes CT, Maitland A, Mostafavi S, Montojo J, Shao Q, Wright G, Bader GD, Morris Q (2010) The GeneMANIA prediction server: biological network integration for gene priorization and predicting gene function. Nucleic Acids Res 38(suppl 2): W214-W220

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

Wilkins MR, Gasteiger E, Sanchez JC, Bairoch A, Hochstrasser DF (1998) Two-dimensional gel electrophoresis for proteome projects: the effects of protein hydrophobicty and copy number. Electrophoresis 19(8-9): 1501-1505

Williams RL, Hilton DJ, Pease S, Willson TA, Stewart CL, Gearing DP, Wagner EF, Metcalf D, Nicola NA, Gough NM (1988) Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. Nature 336(6200): 684-687

Wilson KD, Venkatasubrahmanyam S, Jia F, Sun N, Butte AJ, Wu JC (2009) MicroRNA-profiling of human-induced pluripotent stem cells. Stem Cells Dev 18(5): 749-758

Xu B, Zhang K, Huang Y (2009) Lin28 modulates cell growth and associates with a subset of cell cycle regulator mRNAs in mouse embryonic stem cells. RNA 15(3): 357-361

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

Yu HB, Kunarso G, Hong FH, Stanton LW (2009) Zfp206, Oct4, and Sox2 are integrated components of a transcriptional regulatory network in embryonic stem cells. J Biol Chem 284(45): 31327-31335

Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, Tian S, Nie J, Jonsdottir GA, Ruotti V, Stewart R, Slukvin II, Thomson JA (2007) Induced pluripotent stem cells derived from human somatic cells. Science 318(5858): 1917-1920

Zechner U, Nolte J, Wolf M, Shirneshan K, El Hajj N, Weise D, Kaltwasser B, Zovoilis A, Haaf T, Engel W (2009) Comparative methylation profiles and telomerase biology of mouse multipotent adult germline stem cells and embryonic stem cells. Mol Hum Reprod 15(6): 345-353

Zhang J, Tam WL, Tong GQ, Wu Q, Chan HY, Soh BS, Lou Y, Yang J, Ma Y, Chai L, Ng HH, Lufkin T, Robson P, Ling B (2006) Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. Nat Cell Biol 8(10): 1114-1123

Zhang P, Andrianakos R, Yang Y, Liu C, Lu W (2010) Kruppel-like factor 4 (Klf4) prevents embryonic stem (ES) cell differentiation by regulating Nanog gene expression. J Biol Chem 285(12): 9180-9189

Zhang X, Zhang J, Wang T, Esteban MA, Pei D (2008) Esrrb activates Oct4 transcription and sustains self-renewal and pluripotency in embryonic stem cells. J Biol Chem 283(51): 35825-35833

Zhou H, Wu S, Joo JY, Zhu S, Han DW, Lin T, Trauger S, Bien G, Yao S, Zhu Y, Siuzdak G, Schöler HR, Duan L, Ding S (2009) Generation of induced pluripotent stem cells using recombinant proteins. Cell Stem Cell 4(5): 1-4

Zovoilis A, Nolte J, Drusenheimer N, Zechner U, Hada H, Guan K, Hasenfuss G, Nayernia K, Engel W (2008) Multipotent adult germline stem cells and embryonic stem cells have similar microRNA profiles. Mol Hum Reprod 14(9): 521-529

6. Abbreviations

°C	degree Celsius
x g	gravity
µ-	micro
1D	one dimensional
2D	two dimensional
2D-DIGE	two dimensional difference gel electrophoresis
2DE	two dimensional gel electrophoresis
ALPL	alkaline phosphatase
bFGF	basic fibroblast growth factor
bHLH	basic helix-loop-helix proteins
BMP4	bone morphogenetic protein 4
bp	base pair
BSA	bovine serum albumin
c-	centi
cDNA	complementary DNA
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate
ChIP	Chromatin ImmunoPrecipitation
CPX	ciclopirox
Da	Dalton
DAPI	4',6-diamidino-2-phenylindole
DMEM	Dulbecco's modified Eagle's medium
DMF	N, N-dimethylformamide
DNA	deoxyribonucleic acid
dpc	days post coitum
DTT	dithiothreitol
ECCs	embryonic carcinoma cells
EDTA	ethylenediaminetetraacetic acid
EGCs	embryonic germ cells
EGF	epidermal growth factor
EGFP	enhanced green fluorescent protein
ELISA	enzyme-linked immunosorbent assay
EpiSCs	epiblast stem cells
ESCs	embryonic stem cells
FACS	fluorescence activated cell sorting
FCS	fetal calf serum

FDR	false discovery rate
FT-ICR-MS/MS	Fourier transform ion cyclotron resonance mass spectrometry
g	gramm
GDNF	glial cell line derived neurotrophic factor
GEO	gene expression omnibus
GO	gene ontology
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
HMT	histone methyltransferase
HPLC	high-performance liquid chromatography
HRP	horseradish peroxidase
ICC	immunocytochemistry
ICM	inner cell mass
ICSI	intracytoplasmic sperm injection
IEF	isoelectric focusing
IPG	immobilized pH-gradients
iPSCs	induced pluripotent stem cells
ISS	idiopathic short stature
k-	kilo
l	liter
LC-MS/MS	liquid chromatography tandem mass spectrometry
LIF	leukemia inhibitory factor
m-	milli
m	meter
M	molar
maGSCs	multipotent adult germline stem cells
MALDI	matrix-assisted laser desorption/ionization
MCM	mini-chromosome maintenance
MEFs	mouse embryonic fibroblasts
mGSCs	multipotent germline stem cells

min	minute
miRNA	micro RNA
mRNA	messenger RNA
MS	mass spectrometry
n-	nano
NES	nuclear export signals
NLS	nuclear localization signal
ORF	open reading frame
p-	pico
PANTHER	Protein ANalysis THrough Evolutionary Relationships
PBS	phosphate buffered saline
PCA	principal component analysis
PCR	polymerase chain reaction
PFA	paraformaldehyde
PGCs	primordial germ cells
pH	preponderance of hydrogen ions
pI	isoelectric point
PMF	peptide mass fingerprint
PMSF	phenylmethanesulphonylfluoride
ppm	parts per million
qRT-PCR	quantitative RealTime-PCR
RA	retinoic acid
RNA	ribonucleic acid
RNAi	RNA interference
RNPs	ribonucleoproteins
RT-PCR	reverse transcriptase PCR
SAGE	serial analysis of gene expression
SAPE	streptavidin R-phycoerythrin
SCs	stem cells
SDS	sodium dodecylsulfate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
siRNA	small interfering RNA
SSCs	spermatogonial stem cells
ssDNA	single strand DNA
TOF	time of flight
V	Volt

7. Appendix

7.1 Publications

Dihazi H, Dihazi GH, Nolte J, **Meyer S**, Jahn O, Müller GA, Engel W (2009) Multipotent adult germline stem cells and embryonic stem cells: comparative proteomic approach. J Proteome Res 8(12): 5497-5510

Meyer S, Nolte J, Opitz L, Salinas-Riester G, Engel W (2010) Pluripotent embryonic stem cells and multipotent adult germline stem cells reveal similar transcriptomes including pluripotency-related genes. Mol Hum Reprod 16(11): 846-855

Dihazi H, Dihazi GH, Jahn O, **Meyer S**, Nolte J, Asif AR, Müller GA, Engel W (2010) Multipotent adult germline stem cells and embryonic stem cells functional proteomics revealed an important role of translation initiation factor 5A (Eif5a) in stem cell differentiation. Journal of Proteome Research: Manuscript submitted

Nolte J, **Meyer S**, Khromov T, Graner R, Kotzenberg L, Dev A, Guan K, Pantakani K, Engel W (2010) Germ cell specific gene Stra8 has an impact on the pluripotency network. Stem Cell Reviews and Reports: Manuscript submitted

7.2 Posters

Nolte J, Drusenheimer N, **Meyer S**, Zovoilis A, Hertwig B, Ludwig H, Dev A, Michelmann HW, Wulf G, Meinhardt A, Zechner U, Haaf T, Engel W (2007) Differentiation of pluripotent spermatogonial stem cells to haploid male germ cells. 18th Annual Meeting of the German Society of Human Genetics, March, Bonn

Meyer S, Nolte J, Kaltwasser B, Engel W (2008) Transcriptional profiling of embryonic and spermatogonial stem cells. 19th Annual Meeting of the German Society of Human Genetics, April, Hannover

Meyer S, Nolte J, Kaltwasser B, Pantazi A, Engel W (2008) Transcriptional profiling of mouse embryonic and spermatogonial stem cells. 2nd International Congress on Stem Cells and Tissue Formation, July, Dresden

Nolte J, **Meyer S**, Zechner U, Michelmann HW, Meinhardt A, Engel W (2008) *In vitro* spermatogenesis of mouse pluripotent stem cells. 2nd International Congress on Stem Cells and Tissue Formation, July, Dresden

Meyer S, Dihazi H, Nolte J, Dihazi GH, Khromov T, Opitz L, Salinas-Riester G, Engel W (2009) Transcriptomics and proteomics of mouse embryonic and multipotent adult germline stem cells. 5th International Meeting of the Stem Cell Network North Rhine Westphalia, March, Aachen

Meyer S, Nolte J, Zheng Y, Kotzenberg L, Becker M, Engel W (2009) Germ cell-specific protein Stra8 possesses a role in regulation of pluripotency. 7th Annual Meeting of the International Society for Stem Cell Research, July, Barcelona, Spain

Nolte J, **Meyer S**, Michelmann HW, Kaltwasser B, Ulrich J, Wulf G, Zechner U, Haaf T, Nayernia K, Meinhardt A, Engel W (2009) *In vitro* spermatogenesis of mouse multipotent adult germline stem cells (maGSCs). 7th Annual Meeting of the International Society for Stem Cell Research, July, Barcelona, Spain

Nolte J, **Meyer S**, Kotzenberg L, Becker M, Zheng Y, Pantakani K, Engel W (2009) Germ cell-specific gene Stra8 has a regulatory function in pluripotency and differentiation. World Conference on Regenerative Medicine, October, Leipzig

Meyer S, Dihazi H, Nolte J, Dihazi GH, Engel W (2009) Proteomics of mouse embryonic stem cells and multipotent adult germline stem cells. World Conference on Regenerative Medicine, October, Leipzig

Meyer S, Lührig S, Nolte J, Engel W (2010) *In vitro*-spermatogenesis of mouse pluripotent stem cells. 1st German DFG Reproductive Science Network Meeting, April, Gießen

Khromov T, **Meyer S**, Wolf M, Dihazi H, Nolte J, Zechner U, Engel W (2010) Comparative analysis of mouse embryonic stem cells and multipotent adult germline stem cells. 3rd International Congress on Stem Cells and Tissue Formation, July, Dresden

Lührig S, Nolte J, **Meyer S**, Drusenheimer N, Nayernia K, Engel W (2010) *In vitro* spermatogenesis of mouse pluripotent stem cells. 3rd International Congress on Stem Cells and Tissue Formation, July, Dresden

Nolte J, **Meyer S**, Pantakani K, Khromov T, Zheng Y, Zechner U, Engel W (2010) Mouse germ cell specific gene Stra8 has an impact on the pluripotency network. 3rd International Congress on Stem Cells and Tissue Formation, July, Dresden

8. Curriculum Vitae

Personal details:

Name: Sandra Meyer

Permanent address: Marschnerstr.2

                         30167 Hannover

E-mail: meyer-sandra@gmx.de

Date of birth: 12.12.1981

Place of birth: Minden

Nationality: German

Educational background:

Aug 1988 – Jul 2001:   Primary and secondary school

Oct 2001 – Jan 2006:   Course of Studies: Biology

                         Georg-August-University Göttingen, Germany

Jan 2006 – Jan 2007:   Diploma thesis: "Versuche zur Differenzierung von Teratocarcinomazellen und Fibroblasten zu männlichen Keimzellen."

                         (engl: "On the differentiation of teratocarcinoma cells and fibroblasts to male germ cells.")

                         Institute of Human Genetics, Georg-August-University Göttingen, Germany

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

since Feb 2007:   Doctoral thesis: "Comparative analysis of embryonic stem cells and multipotent adult germline stem cells at the level of transcriptome and proteome."

                         Institute of Human Genetics, Georg-August-University Göttingen, Germany

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

9. Acknowledgements

I wish to express my deepest gratitude to Prof. Dr. med. Dr. h. c. Wolfgang Engel for his excellent scientific supervision, encouragement, for valuable scientific discussions and financial support for my PhD-study. It has been a pleasure working in such an inspiring atmosphere as he has created at the Institute of Human Genetics.

I sincerely thank Prof. Dr. Sigrid Hoyer-Fender for being my co-referee. I also extend my sincere thanks to the other members of my thesis committee Prof. Dr. Wolfgang Brück, Prof. Dr. Gregor Bucher, Prof. Dr. Jürgen Wienands and Prof. Dr. Wolfgang Wuttke for taking time to serve as my examiners and for their critical reading of my thesis despite their busy schedules.

I would like to thank Dr. Gabriela Salinas-Riester, Lennart Opitz und Susanne Luthin for their help in performance and analysis of various transcriptional profiling experiments.

I would also like to thank Dr. Hassan Dihazi and Gry H. Dihazi for a good cooperation and continuous support concerning proteomic experiments. I extend my thanks to Prof. Dr. Georg J. Arnold for unhesitatingly providing intellectual and technical support for proteome analysis.

I am very grateful to Dr. Pawel Grzmil and Dr. Arne Zibat for giving an introduction in quantitative RealTime-PCR and use of Statistica software.

Additionally, I would like to express my sincere thanks to Dr. Jessica Nolte, Sandra Lührig and Nadine Kramann for critically proof-reading the present manuscript.

I would like to thank all my current and former institute colleagues and friends for their help, numerous advices, constant support and fantastic work atmosphere. Especially, in this context I would like to point out the current and former members of our group Birgit, Britta, Heike, Janine, Jessica, Katy, Krishna, Kristin, Linda, Lukasz, Nadja, Sandra, Tan, Tanja and Xingbo. Thank you for spending so much time with me inside and outside the lab.

A very special acknowledgement to my parents, whose love, moral support and total confidence in me were always present and highly appreciated. Thank you for helping me overcome the most stressful und difficult moments of my PhD.