מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

Thesis for the degree

Doctor of Philosophy

Submitted to the Scientific Council of the

Weizmann Institute of Science

Rehovot, Israel

עבודת גמר (תזה) לתואר

דוקטור לפילוסופיה

מוגשת למועצה המדעית של

מכון ויצמן למדע

רחובות, ישראל

By

Leehee Weinberger

מאת

ליהי ויינברגר

אפיון המצב הנאיבי בתאי גזע אנושיים והתמיינותם לתאי מין קדמוניים

Characterization of human naïve pluripotent stem cells and their differentiation to primordial germ cells

Advisor: Dr. Jacob H. Hanna

מנחה: דר. יעקוב חנא

January 2016

שבט, התשע"ו

Contents

List of Abbreviation 4

Abstract 6

תקציר 7

Introduction 8

Embryonic stem cells and pluripotency 8

The origin of pluripotent stem cells 9

Naïve and primed states of pluripotency 11

Pluripotency network and naïve/primed states in mouse PSCs 12

Epigenetic characteristics of naïve and primed states in the mouse 14

Pluripotency of human cells 15

Primordial germ cells (PGC) development in mouse 17

Development of human PGCs (hPGCs) *in vivo* 18

In vitro differentiation of mouse PSCs to PGCs 18

In vitro differentiation of hPSCs to PGCs 20

Results 22

Part I: Establishment, maintenance and characterization of human ESCs in a naïve state of pluripotency 22

Screening for conditions that maintain hPSCs in the presence of 2i/Lif 22

NHSM maintains PSCs pluripotency and enable ESCs derivation 24

Gene expression profile of hPSCs grown in NHSM 26

Epigenetic landscape of NHSM-grown hPSCs 27

Cross species chimerism of hPSCs grown in NHSM 32

Improvement of NHSM medium to a more ICM-like state 32

Part II: Human ESCs differentiation to Primordial germ cells-like-cell (hPGCLC) 36

Establishment of hPGCLCs from hESCs 36

Gene expression profile of hPGCLCs 38

PGCLC differentiation efficiency in naïve and primed 40

Differences between human and mouse PGCLCs 42

Integrin β3 is a specific PGCLC marker 43

Knock out of BLIMP1 precludes hPGC specification 44

The role of SOX17 in hPGCs 45

Efforts to differentiate hPGCLCs to spermatozoa 47
Discussion 50
Materials and methods 54
Cultivation of human naïve pluripotent stem cells 54
Culture of conventional/primed human ESCs and iPSCs 55
Mouse naive and primed stem cell lines and cultivation 56
Flow Cytometry and immune-fluorescence staining of cells 57
ES cell derivation from human blastocysts 57
Teratoma assay 58
Reverse transcription and quantitative PCR analysis 58
Protein western blotting analysis 59
Microarrays analysis of gene expression 59
Chromatin immune-precipitation and sequencing library preparation 60
Chromatin IP sequencing data analysis 61
PGCLC induction 63
Genomic Modifications in hESCs 63
Collection of samples and analysis for RNA–seq 65
Microinjection to mouse testis 65
References 67
List of publications during the course of the PhD 70
Student Declaration 71
Acknowledgements 72

List of Abbreviation

ICM	Inner Cell Mass
EBs	Embryonic Bodies
MEFs	Mouse Embryonic Fibroblasts
SCNT	Somatic Cell Nuclear Transfer
iPSC	Induced Pluripotency Stem Cell
ESC	Embryonic Stem Cell
PSCs	Pluripotent Stem Cells, including ESCs and iPSCs
hESCs	Human Embryonic Stem Cells
mESCs	Mouse Embryonic Stem Cells
EpiSCs	Epiblast Stem Cells
EpiLCs	Epiblast Stem Cell Like Cells
OSKM	Oct4, Sox2, Klf4 and c-Myc
PGCs	Primordial Germ Cells
PCLCs	Primordial Germ Cell Like Cells
EGCs	Embryonic Germ Cells
EpiLCs	Epiblast Stem Cell Like Cells
Dox	Doxycycline
KSR	Knock-out Serum Replacement
FBS	Fetal Bovine Serum
2i	PD0325901 and CHIR99021 (inhibitors of the ERK1/2 and GSK3β)
2i/Lif	serum free medium with addition of Lif, Erki and Gsk3βi
PE	Proximal Enhancer of OCT4 gene
DE	Distal Enhancer of OCT4 gene
TNAP	Tissue Non-specific Alkaline Phosphatase
IVF	*In-Vitro* Fertilization
ICSI	Intra-Cytoplasmic Sperm Injection
NHSM	Naïve Human Stem Cells Medium
ENHSM	Enhanced Naïve Human Stem Cells Medium
MHC	Major Histocompatibility Complex

HLA	Human Leukocyte Antigen
FDR	False Discovery Rate
RPKM	Reads Per Kilobase per Million reads
RQ	Relative Quantity
CRISPR	Clustered Regularly-Interspaced Short Palindromic Repeats

Abstract

The pluripotent state of human embryonic stem cells (hESCs) and iPSCs is different in many characteristics from their mouse counterparts, although the cells originate from the same developmental stage. Since human stem cells resemble the mouse primed state of pluripotency and not the naïve state, we sought to define conditions that can stabilize human cells under ERK and GSK3β inhibitors (2i) containing conditions that enable growth of only naïve mouse cells. Indeed, a combination of small inhibitors of signaling pathways and cytokines (named NHSM) allowed hESCs proliferation in the presence of the two inhibitors. This combination supported human cells in undifferentiated state and allowed for derivation of new cell lines directly from human pre-implantation embryos. Next, hESCs were characterized in NHSM conditions in terms of gene expression and epigenetic modifications and compared with mouse naïve and primed ESCs. The comparison revealed that human ESCs resemble the naïve state of mouse pluripotency and not the primed state. NHSM-naïve hESCs also show expression pattern that is similar to human blastocyst, therefore NHSM medium render the human cells a pluripotency state that is more naïve than conventional hESCs. We also found that modification of NHSM can increase the expression of naïve-related transcripts.

Using Fluorescence reporter lines for the NANOS3 gene, we were able to differentiate naïve hESCs to Primordial Germ Cells (PGCs) *in vitro*, while conventional/primed PGCs were inefficient in this specification. The gene expression profile of the differentiated cells was similar to that of authentic gonadal human embryonic PGCs, showing expression of early, but no late, germ line markers. A specific surface marker, Integrin β3, can be used to identify differentiated cells without the use of fluorescent reporter. This system enabled us to study human-specific regulators of the germ lineage and to assign SOX17 a role as a key transcription factor in human PGCs, in contrast to mouse PGCs.

תקציר

המצב הפלוריפוטנטי של תאי גזע אנושיים שונה במאפיינים רבים מהמצב של תאי גזע עובריים עכבריים, למרות המקור ההתפתחותי הזהה של התאים. כיוון שתאי גזע אנושיים דומים יותר לתאי גזע עכבריים במצב פריימד ולא במצב נאיבי, חיפשנו תנאים המאפשרים לתאי גזע אנושיים לגדול בנוכחות מעכבים של ERK ושל GSK3β (2i), שמאפשרים גידול רק של תאי עכבר נאיביים ולא פריימד. אכן, קומבינציה של מעכבים וציטוקינים (הנקראת NHSM) איפשרה גדילה של תאי גזע אנושיים בנוכחות 2i. צירוף זה איפשר פרוליפרציה, שמר את תאי הגזע האנושיים במצב בלתי ממויין, ואיפשר יצירה של שורות תאי גזע חדשות ישירות מבלסטוציסט אנושי. בהמשך, ביטוי גנים וסמנים אפי גנטיים אופיינו בתאי גזע אנושיים שגדלו במדיום NHSM והושוו עם אלו של תאי גזע עכבריים, נאיביים ופריימד. ההשוואה הראתה שתאי אנושיים שגודלו ב-NHSM דומים לתאים עכבריים נאיביים. כמו כן, תאי גזע אנושיים נאיביים שגודלו בNHSM הראו ביטוי גנים דומה לזה של בלסטוציסט אנושי ולכן ניתן לומר שמדיום NHSM גורם לתאי גזע אנושיים לעבור למצב פלוריפוטנטי נאיבי. שיפורים קלים בהרכב של NHSM גרמו לעליה נוספת ברמת הביטוי של גנים הקשורים במצב הנאיבי.

תאי מין קדמוניים הם התאים העובריים שיהפכו בעתיד לתאי מין: ביצית וזרע. בעזרת גן מדווח פלואורוסנטי יכולנו לזהות תאי גזע אנושיים נאיביים שהתמיינו לתאי מין קדמוניים במעבדה. לעומת התאים הנאיביים, תאי גזע במצב פריימד לא הצליחו להתמיין ביעילות לתאי מין קדמוניים. תבנית ביטוי הגנים של תאי המין הקדמוניים שהתמיינו במעבדה דומה לתבנית הביטוי של תאי מין קדמוניים אותנטיים שהופקו מעובר אדם, ומראים ביטוי של גנים מוקדמים של תאי מין, אך לא מאוחרים. סמן ספציפי לשטח פני התא (אינטגרין β3) מאפשר זיהויי תאים שהתמיינו לתאי מין קדמוניים, ללא שימוש בגן מדווח. מערכת התמיינות זו איפשרה מחקר של רגולטורים הייחודיים להתמיינות תאי מין ראשוניים באדם, ובעזרתה נמצא תפקיד לגנים BLIMP1 ו-SOX17 כחלבוני שיעתוק מרכזיים בהתפתחות תאי מין אנושיים.

Introduction

Embryonic stem cells and pluripotency

Stem cell is a general term for cells that bear the potential to differentiate into different somatic cell types and self-renew. Under this wide definition, we can set apart two types of stem cells: adult and embryonic stem cells. Adult or tissue-specific stem cells are cells that reside in the adult body as well as in the fetus, and are able to differentiate to cells of a specific lineage such as the hematopoietic lineage. Adult stem cells are considered multipotent – baring the ability to turn into multiple types of cells. Embryonic stem cells, on the other hand, are not present in the organism per se. Instead, they can be derived from the epiblast cells of the early embryo, maintain the unlimited differentiation capacity of the embryonic cells and acquire the ability to proliferate indefinitely. Therefore, embryonic stem cells can be maintained in the lab and have the ability to self renew indefinitely and differentiate into all types of embryonic cells and tissues except extra embryonic tissues, thus are considered pluripotent.

Embryonic stem cells were first discovered in the 1980's, when the inner cell mass (ICM) of mouse embryos were cultured with conditioned medium[1]. Cells that grew from these embryos could form teratoma, a tumor containing cells from the three germ lineages - endoderm, mesoderm and ectoderm and were able to self renew indefinitely in the lab. Upon injection into a host mouse blastocyst, These embryonic stem cells (ESCs) were able to contribute to different lineages including the germ line, leading to the formation of a chimeric animal. *In-vitro* differentiation of pluripotent stem cells was possible either spontaneously when these adherent cells were detached from the dish and formed embryonic bodies (EBs) that contained various cell types, or through specific protocols directed to differentiate the cells to particular cell types such as neurons and cardiomyocytes. ESCs can also create a whole embryo (which developed into an adult animal) in tetraploid complementation assay. In this assay, ESCs are injected into a blastocyst that can only contribute to the extra-embryonic tissues and the resulting embryo is composed solely of cells that originated from the injected population.

Tetraploid complementation provides a solid confirmation for the pluripotency of ESCs and their ability to create a complete healthy and fertile animal.

In the following years, mouse embryonic stem cells were grown in conditioned media or on mitotically inactive mouse embryonic fibroblasts (MEFs) in the presence of fetal bovine serum (FBS). In 1988, Leukemia Inhibitory Factor (Lif) was found to be the cytokine that stabilizes mouse ESCs (mESCs)[2], and in 1998, human ESCs (hESCs) where first derived from human embryos[3]. The hESCs were able to form EBs *in vitro* and teratomas when injected sub-coetaneous into immune-deficient mouse. For ethical reasons, the ability of hESCs to contribute to human embryos was never tested. Therefore hESCs originate from the ICM of human embryo and can differentiate to cells of the three germ layers, similar to mESCs. However, their ability to contribute to adult tissues remained to be studied. In contrast to mESCs, conventional hESCs are not dependent on LIF signaling but require FGF2 and TGFβ for their maintenance.

Pluripotent stem cell research can be used for several purposes. First, *in-vitro* construction of differentiated tissues for tissue replacement therapy is an ultimate goal that could aid in many diseases, such as diabetes and Parkinson. Second, pluripotent stem cells can be used to model diseases in the lab, drug screening and studying genetics of different diseases, as long as the differentiation protocol for the tissue in interest is robust. The fact that ESCs are easy to be genetically manipulated in the lab, together with their capacity to contribute to chimeric animal and to the germ line, make them an ideal tool to use in genetic engineering and formation of model animal for various diseases and genetic conditions. Differentiated ESCs can be also used to induce tissue response during transplantation, instead of direct replacement of the damaged tissue. Finally, pluripotent stem cells can aid in studying the early embryonic development because of their similarities to the ICM of the blastocyst.

The origin of pluripotent stem cells

Embryonic stem cells were the first type of pluripotent stem cells to be identified, but the use of embryonic stem cells in cell therapy is limited for ethical reasons such as requirement for fresh early embryos, and technical reasons such as genetic background mismatch with patient. Already in 1962, Gurdon et al[4] have shown that oocytes have

the ability to "reset" the nucleus of terminally differentiated somatic cells. When nucleus-free oocyte was injected with somatic cell nucleus and stimulated, this zygote-like cell was able to continue to develop and to form and embryo, which later developed into a whole animal. This technique, called Somatic Cell Nucleus Transfer (SCNT), has provided the evidence that each cell in the organism maintain the potential to become pluripotent and even totipotent. ESCs could be derived from SCNT blastocyst and these cells resemble ESCs derived from normally fertilized eggs. SCNT is technically inefficient and requires a large number of non-fertilized oocytes, therefore it is not routinely used.

Since SCNT revealed the developmental potential hidden in somatic cells, a search for different factors that can elicit totipotency or pluripotency in somatic cells began. In 2006, Shinya Yamanaka and colleges[5] found that transfection of somatic cells with 4 factors: Oct4, Sox2, Klf4 and Myc (OSKM) reprogrammed the cells to pluripotency. These induced pluripotent stem cells (iPSCs) have differentiation potential identical to embryonic stem cells and could proliferate *in vitro* endlessly. iPSCs could be derived from different somatic cells, including skin fibroblasts and B cells, and the reprogramming process, that takes 1-4 weeks in conventional conditions, completely erases the transcriptional and epigenetic program that was present in these somatic cells. Different combinations of factors can be used to reprogram cells to pluripotency and in all cases the iPSCs are stable and display gene expression, epigenetic state and developmental capacity extremely similar to ESCs. The discovery of iPSCs has revolutionized the stem cells research field and enabled the use of patient specific iPSCs for disease study.

While most somatic cells require genetic manipulation in order to regain pluripotency, specific cell types maintain the ability to become pluripotent after the blastocyst stage. Such cells are the Primordial Germ Cells (PGCs), which are the embryonic cells that will develop into the germ cells (oocytes and sperm) later in development. PGCs are not pluripotent themselves and can give rise only to germ line cells, but they convey some pluripotent characteristics that will later allow the gametes to induce totipotency in the fertilized egg. Therefore, PGCs express some of the pluripotency network genes such as

Oct4 and Nanog and are presumably closer in their epigenetic state to pluripotent cells than other somatic cells. Indeed, mouse PGCs (mPGCs) have the ability to be reprogramed to pluripotency without the addition of any factor: isolated mPGCs can be seeded in mouse ESCs conditions which turn them into pluripotent cells[6]. The resulting cells, which are called EGCs (Embryonic Germ Cells), show high similarity to ICM derived ESCs and are able to contribute to chimeric animals. A prominent difference between ESCs and EGCs is the state of DNA methylation on imprinted genes. These genes are constantly methylated in somatic cells and ESCs according to their parental/maternal origin and this methylation is lost during PGCs differentiation and re-written later. EGCs display aberrant DNA methylation on imprinted genes, which makes them unsuitable for whole embryo formation through tetraploid complementation assay. Nevertheless, EGCs are able to donate to the germ line in chimeras[7].

Most of the cells in adult animal completely lose the ability to become pluripotent without genetic interference. Nevertheless, spermatogonia cells that were isolated from an adult mouse testis could become pluripotent *in vitro* without external addition of reprogramming factors[8]. It is possible that the propensity of the germ lineage cells to become totipotent after fertilization, also provide them with the epigenetic state that support transition to pluripotency in the lab.

A recent report suggest that somatic cells can be reprogrammed to pluripotency solely through the use of small molecule inhibitors that affect the epigenetic status of the cells, resulting in a less constrained state that allows the propagation of the pluripotency network[9].

Naïve and primed states of pluripotency

In 2007, close to the iPSCs discovery, two groups showed that pluripotent stem cells could also be derived from post implantation embryos[10, 11]. Isolated embryonic cells from E5.5-E7.5 mouse embryos readily gain pluripotency in the lab and became a new type of embryonic stem cells named Epiblast Stem Cells (EpiSCs). EpiSCs are quite different from previously characterized mESCs and iPSCs and cannot contribute efficiently to chimeric animal through blastocyst injection, although their ability to form teratomas is not compromised. It is not clear whether the inefficiency in forming

chimeras is due to lower differentiation potential of EpiSCs, or because of expression of adhesion molecules that are relevant to the advanced embryo, which prevent cell-cell contact and integration of the EpiSCs into the blastocyst. It is also important to note that EpiSCs are dependent on FGF2 and Activin/Nodal signaling for their survival, similar to hESCs and in contrast to mESCs and iPSCs that depend on Lif.

Many studies of pluripotent stem cells (PSCs, include both ESCs and iPSCs) were done on the mouse, and specifically in 129 mouse strain. ESCs derivation from some mouse strains was unsuccessful under conventional conditions that include MEF coated plates and medium with serum and Lif. Other species such as rat, monkey and pig also couldn't yield ESCs under these conditions. As hPSCs require different conditions from mPSCs, it is possible that different species require different conditions to support their pluripotency state, although the origin of the cells is identical.

Pluripotency network and naïve/primed states in mouse PSCs

Conventional mPSCs require only serum and Lif for their proliferation and maintenance in an undifferentiated state. Attempts to find a defined medium that support mPSCs led to the discovery of two inhibition signaling pathways: Erk and β-catenin. Addition of small molecule inhibitors for Erk and Gsk3β enables cell proliferation and prevents differentiation in a defined medium without the addition of serum, MEFs or other animal products[12]. The 2i/Lif medium (serum free medium with addition of Lif, Erki and Gsk3βi) prevents differentiation by blocking intrinsic and extrinsic differentiation cues leading to a more homogenous cell population, with higher levels of pluripotency gene expression. While Lif is an activator of the STAT3/JAK pathway that is activated in the early embryo and in pluripotency, the roles of the Erk and β-catenin pathways in pluripotency remain to be studied.

Alternative 2i conditions, involving small molecule inhibitors for Gsk3β and Src pathways, have been shown to yield murine germ-line competent ESCs[13]. Atypical PKC inhibitor (aPKCi) Gö6983 is another stimulator for isolating murine ESCs together with LIF and/or MEK inhibitors[14]. These findings support the notion that PSCs tend

to spontaneously differentiate in neutral conditions, and that inhibition of signaling pathways is required to sustain PSCsundifferentiated.

In contrast to mESCs and iPSCs, mEpiSCs that are isolated from a later stage embryos, depend on FGF2 and Activin signaling. EpiSCs show higher levels of expression of lineage or differentiation related genes such as OTX2, ZIC2/3, T (also called Brachyury)[15] and very low propensity for blastocyst integration. These characteristics of EpiSCs lead researchers to articulate EpiSCs as "primed" pluripotent stem cells, in contrast to PSCs in serum+Lif or 2i/Lif that are considered to reside in a naïve state of pluripotency[16]. Many differences were characterized between naïve and primed pluripotency. In general, primed EpiSCs are more similar to post implantation epiblast while naïve PSCs are more similar to pre-implantation epiblast. Surprisingly, the naïve and primed states are completely dependent on medium condition such that change of media can lead to a transition of cells from one state to another. mESCs derived from blastocyst can be set to the primed state after couple of passages in primed medium and even more amazingly, mEpiSCs can transit to the naïve state only through change to a naïve medium, without genetic manipulation[17]. iPSCs can be reprogrammed in naïve medium to reach the naïve state or in primed medium to reach the primed state. Even though the two states are distinct in many features, external signals can completely set the pluripotency state of a cell.

Furthermore, the signals that support each of the pluripotency states often antagonize each other: while FGF2 and TGFβ signaling support the primed state, they lead to differentiation when added to naïve cells. Similarly, addition of Erk or Gsk3β inhibitors to primed cells leads to rapid differentiation. An exception to this rule is the addition of Gsk3 inhibitor CHIR to primed EpiSCs together with IWR1, Tankirase inhibitor that lead to β-catenin retention in the cytoplasm. This combination leads to high levels of cytoplasmic β-catenin and enables primed mPSCs proliferation, while maintaining primed characteristics[18].

Pluripotency of a cell is set by the expression of a network of transcription factors that bind to promoters of thousands of genes and to each other, this way stabilizing the

cellular state. This network includes Oct4, Nanog, Klf4, Klf5, Sox2, Dppa3, Dppa5, Esrrb, Tfcp2l1, and more[19]. While naïve mouse pluripotent cells express all these genes in a high level, primed PSCs express Oct4 and Sox2 but only express low levels of Nanog, Klfs, Dppas and Esrrb, showing rewiring of the network that occurs during the transition between the states. Comparison to *in-vivo* states, naïve and primed pluripotency gene expression show high similarities to post- and pre- implantation embryos, respectively. Accordingly, the binding sites of Oct4 in naïve and primed ESCs differ substantially.

Other differences between naïve and primed cells include flat colony morphology for primed cells in contrast to domed colonies of the naïve cells, lower expression of MHC class I molecules on the surface of naïve cells [20], lower clonogenity of single cells of primed PSCs and higher heterogeneity of primed PSCs within a cell line population[21].

Epigenetic characteristics of naïve and primed states in the mouse

PSCs exhibit four unique epigenetic features: DNA methylation level, X-chromosome activation state, poised enhancers and bivalent promoters.

In terms of global DNA methylation, naïve mPSCs that are grown in 2i/Lif show extremely low level of DNA methylation, similar to that present in the blastocyst, while primed cells, but also naïve cells in serum+Lif medium, show high levels of global DNA methylation, similar to somatic cells[22].

All cells in a female body exhibit inactivation of one of the X chromosomes. This state is denoted XaXi, when one of the X chromosomes is inactive in terms of gene expression, covered by Xist LincRNA, marked by high level of H3K27me3 modifications and dense to a bar body in the cell. Naïve PSCs are an exception to this role, when both X chromosomes are active (XaXa), similar to Blastocyst cells. Primed ESCs keep one chromosome in an inactive state[17].

Naïve PSCs are also characterized by general low level of H3K27me3. This histone modification is present mostly in developmentally-regulated genes in somatic cells and renders them inactive or silent[23]. Primed PSCs, on the other hand, display both H3K27me3 inhibitory signal and H3K4me3 activating signal at the same time on many

developmental genes. This unique combination of activating and silencing modification renders these genes inactive but primed for activation such that they could be quickly turned on during upcoming differentiation.

Similarly, poised enhancers exhibit both enhancer-related features such as H3K4me1 and p300 binding together with H3K27me3 – an inactive mark[24]. Poised enhancers do not activate transcription of their target genes. Instead, it is thought that the enhancers are set to standby mode and able to activate target genes in a fast manner, once the inactive mark is removed.

The expression level of Oct4, the master regulator of pluripotency, is similar between naïve and primed states, but its expression is regulated differently. The enhancer of this gene has two regulatory elements, a proximal enhancer (PE) and a distal one (DE). In the ICM, Oct4 transcription is regulated through both enhancers, with the DE being more dominant. After implantation, Oct4 is regulated mostly through the proximal enhancer[25]. Similarly, in naïve ESCs the distal enhancer is utilized, while primed cells predominantly activate the distal enhancer[16].

In summary, differences between naïve and primed states of pluripotency in the mouse are present at different levels of regulation, including signaling pathways, epigenetic and transcriptional. Generally, primed PSCs resemble the post implantation epiblast while the naïve PSCs resemble the blastocyst ICM, and display features of a more prone-to-differentiation state. Within the two extreme pluripotency states (2i\Lif naïve and FGF2+Activin primed), different variations such as Serum+Lif and "alternative 2i" exist, with varying levels of pluripotency markers. Therefore, it seems that pluripotency *in vitro* consists a spectrum, rather than two discrete states.

Pluripotency of human cells

While ESCs are readily derived from some inbred mouse strains, other strains such as NOD and other species like rat did not yield ESCs in Serum+Lif medium. The derivation of ESCs from these strains and species required inhibition of differentiation signals through the use of Gsk3β and Erk inhibitors in the 2i/Lif medium[26, 27]. It is

therefore possible that some species require harsh regulation in order to maintain PSCs in a naïve state.

While mouse and human embryonic stem cells are derived from the same embryonic state, human ESCs differ from their mouse counterparts in many aspects. While some aspects can be attributed to differences between the ICM of the two species, other differences require further investigation.

hESCs depend on FGF2 and TGFβ for their growth, similar to primed mouse cells. Furthermore, addition of 2i/Lif components to hESCs, either together with FGF2+TGFβ or alone, leads to rapid differentiation and cell death. In terms of gene expression, conventional hESCs express the basal pluripotency network genes OCT4, NANOG and SOX2, but do not express or express only low levels of KLF4, KLF5, TFCP2L1 and DPPAs. hPSCs also express some lineage related genes, show flat colony morphology and have low level of single cell clonality[16].

Therefore, in most characteristics human stem cells resemble primed mouse cells rather than naïve cells. It is now widely accepted that hPSCs indeed reside in a pluripotency state that is more primed than the mouse 2i/Lif naïve cells[16, 28, 29]. Since some species like rat require strict regulation in order to control their ESCs, it is possible that hESCs go through a stage of transition from the naïve to the primed state shortly after their isolation from the ICM. A proof for the existence of a more naïve state for hPSCs was shown in 2010, when constitutive over expression of pluripotency factors such as OCT4+KLF4 or KLF2+KLF4 in hESCs enabled their growth in naive 2i/Lif medium without differentiation[20]. These reverted cells were more similar to mouse naïve cells and to human ICM than conventional hPSCs in both transcriptional and epigenetic parameters. When the expression of the exogenous factors was removed, the cells reverted back to a primed-like state and rapidly differentiated in 2i/Lif medium. This proof of concept for the ability of hPSCs to acquire a naïve-like state led us to hypothesis that hPSCs could be maintained in a naïve state of pluripotency without genetic modifications through restriction of differentiation using different combinations of inhibitors and cytokines.

During my PhD I screened combination of inhibitors and cytokines that were previously described to have a role in pluripotency, in order to enable hPSCs to grow in the presence of ERK and GSK3β inhibition. I devised a combination that allowed hPSCs to grow and remain pluripotent. Next, I characterized hPSCs grown in this condition and found that they exhibit a variety of naïve traits, consistent with a naïve state of pluripotency that is endowed upon human cells in this condition.

Primordial germ cells (PGC) development in mouse

Germ cells are the carriers of genetic information to the next generations. Therefore, the specification and differentiation of these cells are of high importance and are highly regulated. PGCs consists a group of precursor cells that are specified during early development (from E6.5 in the mouse) by Bmp signal from surrounding extra-embryonic ectoderm cells, and will later differentiate into oocytes and spermatozoa. Only a group of ~40 cells respond to Bmp signal and activate a transcriptional program that specifies their fate to become PGCs[30]. This program is based on Blimp1 (encoded by Prdm1) and Prdm14 together with Ap2γ (encoded by Tfap2c), a direct target of Blimp1 that acts combinatorially to repress somatic genes, induce expression of PGC genes (Nanos3, Dppa3), re-induce pluripotency genes and initiate the epigenetic program of PGCs[31].

During their differentiation, PGCs go through massive epigenetic modifications including increased level of H3K27me3, reduced H3k9me2 level and complete erasure of DNA methylation together with X-chromosome reactivation in females[30]. Actually, late PGCs is the only stage in which DNA methylation is removed from imprinted (parental and maternal) genes[32]. Another epigenetic mark of PGCs is Histone 4 methylation on Arginine residues, formed by Prmt5, a methyl-transferase that is specifically active in PGCs[33]. After PGCs are specified they migrate through the allantois to the developing gonads (genital ridges), where surrounding cells supply further support for the later stages of differentiation to germ cells (GCs) and meiosis.

Interestingly, mPGCs display many characteristics of embryonic stem cells, including Oct4 and Nanog expression, and stain positively for alkaline phosphatase (TNAP) and

SSEA1 markers. PGCs are also the only cells of the late post-implantation embryo (after E7) that could be reprogrammed to naïve pluripotency when plated in appropriate conditions *in vitro*. The resulting EGCs resemble naïve ESCs in all qualities except for differences in DNA methylation in imprinted loci.

Development of human PGCs (hPGCs) *in vivo*

Mammalian development and specification of PGCs were mostly studied in the mouse model. Mouse and other rodents show unique early post implantation morphology of an egg like cylinder while other mammals, including humans, have flat disk morphology at this stage. Since PGCs specification is heavily dependent on signals from surrounding tissues, the different shape of the embryo could have great effect on the signaling requirements during PGCs specification[34]. Since PGCs resemble ESCs, the differences between human and mouse pre-implantations and the primed state of human PSCs could lead to differences in PGCs specification and development between human and mouse. hPGCs are specified only on the third week after fertilization, although the location of the cells is similar to that in the mouse. Due to technical and ethical limitations in processing early post-implantations human embryos, not much is known about the first steps in hPGC specification *in vivo*.

The gene expression and methylation profiles of late hPGCs were recently described, confirming that hPGCs express many of the transcription factors that are present in mPGCs, including PGC-related genes such as BLIMP1 and AP2γ, and pluripotency genes such as OCT4 and NANOG, although SOX2 is not express in humans in contrast to mPGCs[35, 36]. This genome-wide analysis shed the light on late stages of PGC development, including steps of global DNA demethylation and X-chromosome inactivation, but the earlier stages of hPGC specification remained obscure.

In vitro differentiation of mouse PSCs to PGCs

The nature of the epigenetic changes that occur during PGC differentiation *in vivo*, although extensively studied, is not completely defined due to technical limitations, mainly because the number of PGCs is very small, making some experimental techniques impossible to apply. In 2011, Saitou group succeeded in differentiating

murine ESCs into primordial germ cells *in vitro*, providing large quantity of PGC-like cells for the study of PGCs specification[37]. The differentiation protocol imitates *in-vivo* development: naïve ESCs were first exposed to prime-promoting conditions (FGF2 and Activin for 2 days), then to a high concentration of BMP4, the cytokine that is responsible to PGC differentiation in the embryo, together with SCF, EGF and Lif. The *in-vitro* derived PGCs, which were entitle PGC-Like Cells (PGCLCs), showed similar properties to E9.5 mPGCs in terms of gene expression and removal of some of the DNA methylation on imprinted loci. Notably, primed mPSCs were unable to produce PGCLC, but a stage of priming for two days was important for the differentiation protocol.

Functional evidence of the nature of these *in vitro* differentiated cells came when PGCLCs that were injected into mouse testis could further differentiate and give rise to viable, fertile sperm cells[37]. Female mPGCLCs could also differentiate to oocytes when they were mixed with embryonic supporting gonadal cells and injected to recipient mouse ovary[38]. mPGCLC derived oocytes could be fertilized through *in vitro* fertilization (IVF) after *in-vitro* maturation (IVM) process, while mPGCLCs-derived sperm required intracytoplasmic injection (ICSI) in order to fertilize an oocyte. Nevertheless, most of the resulting embryos did not show abnormalities and were fertile themselves. This ability to transfer the genetic information through PGCLCs could provide technical assistance in genetic engineering techniques in the future.

Furthermore, PGCLCs are a source for large quantity of primordial germ cells, such that experimental methods that require large cell number such as ChIP-Seq can be applied to study early PGCs. Indeed, quantitative dynamics of epigenetic modifications such as H3K27me3 and H3K9me2 were studied during the differentiation process of mPGCLC from ESCs and reveled some of the mechanisms through which Blimp1 deploy the PGC fate[39].

Therefore, the establishment of mPGCLCs is useful for studying early stages of development in the lab, it provides large quantities of cells for genome-wide studies, and could also provide insights for the developmental potential of different pluripotent

states. The generation of fertile gametes from PGCLCs could supply tools for genetic engineering and for studies of fertility conditions.

In vitro differentiation of hPSCs to PGCs

While mPGCLCs could provide insights into mammalian PGC development, the differences in pluripotency and in embryonic development between human and mouse require further investigation in human systems. Furthermore, fertility treatments will require human germ cells and not mouse. Several groups have shown that hESCs have a tendency to spontaneously differentiate to the germ lineage with or without addition of BMP4[40-43], but the robustness of this differentiation and the full characteristics of the differentiated cells was not studied deeply enough. First, the differentiation of human cells to PGCLCs occurred at very low percentage during a long period of time. Second, systemic comparison of these *in-vitro* derived cells to human and mouse *in vivo* PGCs was not shown and the identification of differentiated cells was based on markers that are not PGC-specific. Direct differentiation of hPSCs to PGCLCs could provide a system for studying early hPGC specification when human cells are unreachable for study. Such study could identify human or primate specific regulators of PGCs and epigenetic differences and similarities with rodents. Since primed mESCs are inefficient in PGCLC differentiation, specification of PGCLCs could also help us to evaluate the pluripotency state of human stem cells. Finally, further differentiation of PGCLCs to gametes could prove beneficial in fertility treatment in patients.

In my work, I present an efficient *in vitro* system for human naïve PSCs differentiation into PGCs. Using fluorescence knock-in reporters I was able to identify cells that differentiate to human PGCLCs, characterize their gene expression in comparison to authentic hPGCs and identify the effect of pluripotency state of the cells on differentiation efficiency. I also studied the effect of BLIMP1 and SOX17 in regulation of hPGCLCs.

Most of the work described in my thesis was done by me. Tissue culture work, including optimization of human naïve mediums (NHSM and ENHSM) was done by me, with the help from Dr Jacob Hanna. Analysis of gene expression data was done by Yair Manor (both micro-array and RNA-seq data). Analysis of ChIP-Seq data was done by Dr. Noa

Novershtern. Cloning and construction of reporter lines and mutations were done with Sergey Viukov. Cell injections to mouse testis were done me together with Mirie Zerbib. Derivation of new hESCs lines was done in LIS maternity hospital in Dr. Dalit Ben Yosef lab. The second part of the work (hPGCs project) was done in collaboration with Prof. Azim Surani lab in University of Cambridge, UK and specifically with Dr. Naoko Irie.

Results

Part I: Establishment, maintenance and characterization of human ESCs in a naïve state of pluripotency

Most of the work described in this chapter was published in

"*Derivation of novel human ground state naive pluripotent stem cells*". Nature, 2013[44].

Screening for conditions that maintain hPSCs in the presence of 2i/Lif

Since 2i/Lif conditions cannot support hPSCs, we set to define additional conditions that would support human cells and transform them to a naïve state. We used the human ES cell line H1, which harbors OCT4-GFP knock-in reporter, to test several cytokines and small molecule inhibitors previously reported to be involved in pluripotency maintenance. First, we tested the effect of p38 inhibition, as it was shown that BMP4 signaling inhibits this pathway and promote pluripotency in the mouse[45]. When p38 inhibitor (p38i) was added to conventional (primed) hESC medium the colony morphology and the level of OCT4-GFP reporter did not change, therefore p38 inhibition is not detrimental to human cells (figure 1A). Next, we screened different combination of inhibitors and cytokines, looking for their effect on differentiation status, through measurement of OCT4-GFP percentage of the cells and colony morphology, looking for conditions that enable hESCs growth in the presence of ERKi and GSK3βi. We found that many conditions reduced the level of OCT4-GFP reporter or led to differentiation, as seen by lose of colony morphology. Between the conditions that supported pluripotency, the combination of Lif, 2i, p38i and FGF2 increased the percentage of GFP+ cells and resulted in more domed-shaped colony morphology (figure 1B-C). ACTIVIN A, which activates the TGFβ receptor, also increased survival and prevented differentiation. On the other hand, FGF2 inhibition led to rapid differentiation or cell death in all the tested conditions. Since FGF2 and TGFβ diverse in their control of pluripotent genes between human and mice[46], we speculated that

addition of FGF2 and TGFβ to the medium, together with 2i/Lif and p38i, could maintain the cells in a pluripotent state that would be similar to the mouse naïve state.

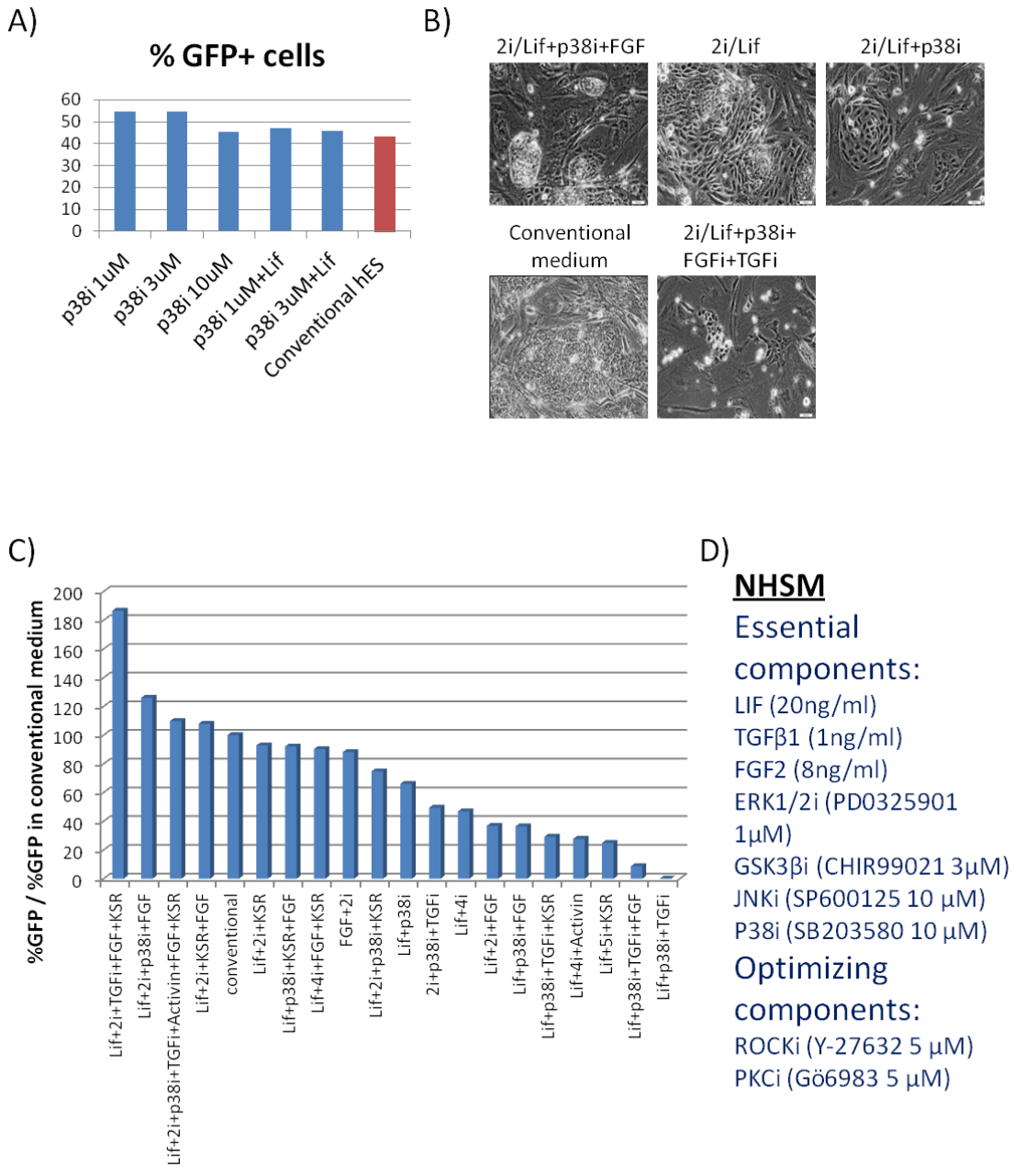


Figure 1: Screening for conditions that support human ESCs pluripotency. A) Addition of p38 inhibitor (SB203580) to conventional ES medium does not interfere with pluripotency level. H1 cells harboring OCT4-GFP reporter were grown in the presence

of different concentrations of p38i for 4 passages and the percentage of GFP positive cells was measured by flow cytometry. B) Representative pictures of colony morphology in 5 conditions after 3 passages in the conditions. C) Screening for the effect of combinations of 2i, p38i, FGF2, ACTIVIN A, FGFi, TGFβi and KSR on human ES pluripotency. H1 cell line was grown for 8 passages in the presence of the different combinations, and the percentage of GFP positive cells, corrected for the percentage in conventional medium, is plotted. D) Composition of the NHSM medium.

Further optimization resulted in addition of JNK inhibitor to the medium, which we termed NHSM, for Human Naïve Stem cell Medium (Figure 1D). To account for different genetic background of hESCs and iPSCs, we grew 6 different cell lines, including both ESCs and iPSCs, in NHSM medium that containing 2i, Lif, p38i, JNKi, FGF2 and TGFβ. ROCKi (Y27632) and PKCi (Gö6983) can be optionally added to enhance growth and characteristics of human naïve PSCs. All cell lines were expanded for over 30 passages and maintained domed-shape morphology (Figure 2B). Two cell lines, H1 and BGO1, that harbor an OCT4-GFP reporter, demonstrated high level of GFP even after long term passaging. The NHSM medium could be switched back to conventional hPSCs medium and the cells remained pluripotent as noticed by the OCT4-GFP reporters. Based on the fact that NHSM combination of factors enabled the growth of hESCs under ERK and GSK3β inhibition, we decided to further characterize cells grown in this condition to see if their epigenetic state resembles the mouse naïve state more than hPSCs grown in the conventional hPSC medium.

NHSM maintains PSCs pluripotency and enable ESCs derivation

First, we wanted to confirm that cells grown in NHSM are pluripotent. For this purpose, we injected six different cell lines subcutaneously to Skid mice. All injected lines formed teratomas that were composed of cells of the three germ layers: endoderm, mesoderm and ectoderm (Figure 2C). Plating the cells in suspension, in medium without cytokines and inhibitors led to formation of embryonic bodies (EBs), where pluripotency

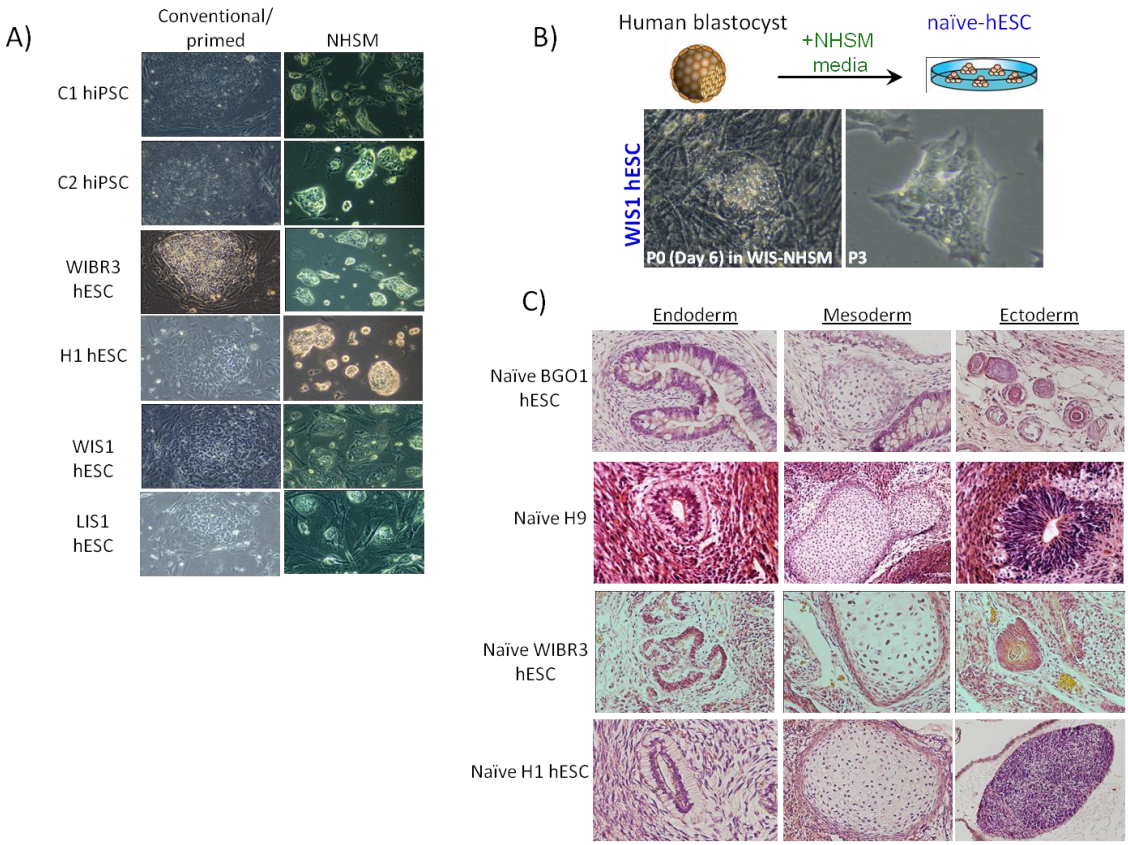


Figure 2: Pluripotency of hPSCs grown in NHSM medium. A) Colony morphology of different human ESCs and iPSCs lines in primed medium (KSR+FGF2) and NHSM medium, showing flat colony morphology in primed conditions compared to a more domed shape in NHSM. B) Derivation of new hESCs lines with NHSM medium. Human blastocyst was plated with NHSM medium, and 6 days later first outgrowths appeared, that later formed a stable ES cell line. C) hPSCs grown in NHSM are able to differentiate to cells of the three germ lineages. H&E staining of Teratomas from four human cell lines that were grown in NHSM for 10-20 passages, showing tissues of the three germ layers.

genes expression is down regulated and the expression of markers of the three germ layers is upregulated. These results, together with staining for pluripotency markers, demonstrate that hESCs grown in NHSM remain pluripotent, even in the presence of 2i. The karyotype of 6 different cell lines that were grown in NHSM for over 10 passages was normal, therefore NHSM can maintain chromosomal stability for long periods. In collaboration with LIS maternity hospital IVF unit, donated human blastocysts were plated in NHSM medium and after few days outgrowths began to appear. These cells

were trypsinized and further expanded in NHSM medium, resulting in derivation of 4 different hESCs cell lines (Figure 2B). The derived cell lines have normal karyotype, show domed morphology similar to other hESC lines in NHSM medium and could form teratomas upon subcutaneous injection. Since NHSM medium can support the derivation and establishment, as well as the maintenance, of hPSCs in conditions that include ERK and GSK3β inhibition, we referred to these cells as naïve pluripotent human cells.

Gene expression profile of hPSCs grown in NHSM

In order to discover differences in gene expression between NHSM-grown cells and conventional hPSCs, we used Affymetrix microarrays to profile gene expression of 9 cell lines that were grown in NHSM and 7 cell lines grown in conventional medium, including both hESCs and hiPSCs. Both naïve and conventional lines expressed similar high levels of pluripotency genes, such as NANOG, OCT4 and SOX2, across all cell lines. On the other hand, some lineage specific genes such as ZIC1 were upregulated specifically in conventional medium (Figure 3B). Unbiased clustering of the genome-wide expression patterns between the different samples clustered according to the growth conditions (NHSM vs. conventional medium), regardless of the genetic background of the cells (Figure 3C). Also, cells grown in NHSM for over 20 passages and were transferred back to conventional medium for 4 passages clustered with conventional PSCs and separately from NHSM cells, proving that NHSM-dependent pluripotency can be reverted back to conventional/primed pluripotency. GO enrichment showed that genes that were down regulated in NHSM were related to several biological processes, including "regulation of organismal development" and "system development". On the other hand, some lineage markers such as BRACHYURY (T) were upregulated in NHSM. Next we tested whether NHSM-grown cells resemble the human ICM cells. For this purpose we compared our data to previously published dataset that measured the gene expression of the human pre-implantation embryo[47]. Amazingly, cells that were grown in NHSM clustered together with human ICM, while conventional primed PSCs clustered with the previously published hESCs and away from the ICM (Figure 3D). This result signifies that NHSM indeed brings hPSCs to a pluripotent state that is closer to the *in vivo* naïve state, compared with conventional hPSCs medium.

Since hPSCs that are grown in NHSM show some characteristics of naïve pluripotency, a comparison to the mouse naïve and primed PSCs could further strengthen the notion that NHSM force naïve characteristics on human cells. To this end, cross-species unbiased clustering of the human samples with previously published naïve and primed mouse samples[20] for 9,803 ortholog genes found high correlation between NHSM samples and mouse naïve (2i/Lif) samples, compared to the correlation between primed hESCs and naïve mouse cells. On the other hand, conventional hPSCs displayed higher correlation with primed mESCs (Figure 3E).

We then looked at the expression of MHC class I (HLA) on the cell surface of hESCs that were grown in NHSM. Similar to mouse naïve cells[20], NHSM-grown human cells express very low level of HLA proteins, while conventional hESCs express intermediate level of HLA proteins and differentiated line (HEK293) expressed high levels of these proteins (Figure 3A). Naïve hESCs also showed low apoptosis levels when dissociated to single cells without the presence of Rock inhibitor, which is necessary for survival of conventional hESCs during trypsinization to single cells.

Epigenetic landscape of NHSM-grown hPSCs

In order to assess the chromatin state of human naïve cells, we conducted ChIP-seq analysis for a couple of chromatin marks: H3K4me3 that marks active transcription, H3K4me1 marking enhancers, H3K27me3 marking repressed genes and poised enhancers, H3K27Ac marking active enhancers and H3K9me3, marking heterochromatin. The positions of these marks was tested in five NHSM-grown human cell lines and four conventional medium grown human cell lines, one naïve mouse sample and one primed mouse sample. The genome-wide distribution of H3K9me3, a chromatin mark that is localized predominantly to silence DNA, is similar between the human naïve samples and the human primed samples. When focusing on the X chromosome, in female cell lines the level of H3K9me3 is significantly lower in NHSM samples, compared to primed human samples. The level is significantly lower in naïve cells in 3 different female cell lines, but male line (WIS1) display comparable levels of H3K9me3 on X chromosome (Figure 4A). We conclude that H3K9me3 marks inactive

X chromosome in conventional human cells, and that NHSM can relieve this mark from human cells, a sign for X chromosome reactivation.

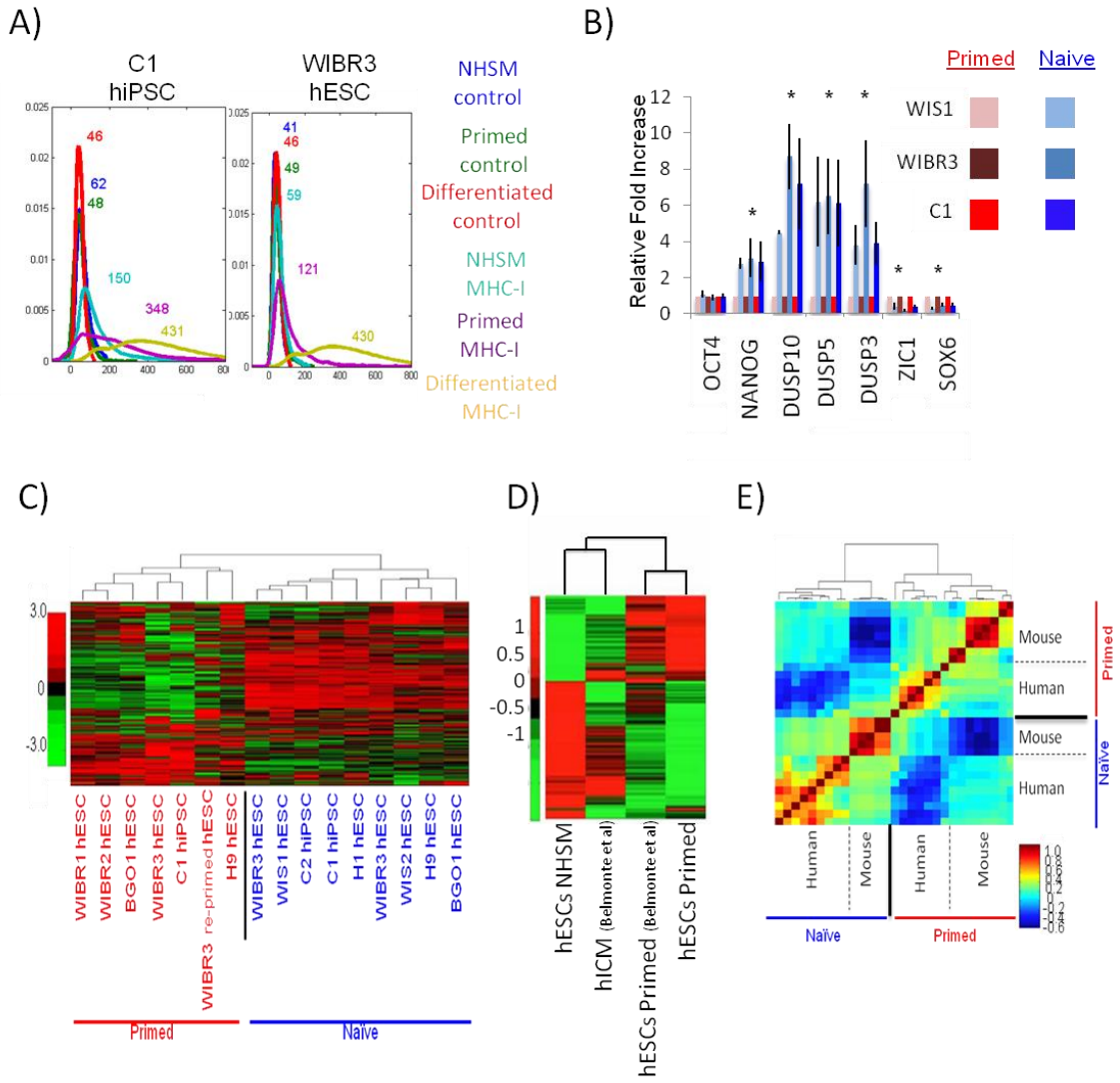


Figure 3: Gene expression of naïve hPSCs. A) Histogram of surface expression of MHC class I (HLA) using FACS analysis on the indicated naïve (NHSM), primed and differentiated (HEK293 cells) states, in C1 and WIBR3 cell lines . B) Gene expression levels of pluripotency (OCT4, NANOG), DUSP family (DUSP10, DUSP5, DUSP3) and lineage specific (ZIC1, SOX6) genes, measured by Affymetrix micro-arrays in naïve (NHSM) and primed/conventional states in 3 different cell lines. C) Hierarchical clustering of genome-wide (n=12,071) transcriptional profiles of the indicated primed and naive hPSC lines measured by Affymetrix microarrays using Pearson correlation. Heat map shows row-normalized expression levels (log2 scale) with red and green colors representing up and down-regulated genes, respectively. Re-primed: cells were

transferred from NHSM medium to primed medium for 4 passages before RNA extraction. RNA from WIBR3 cell line was extracted in two biological replications. D) Transcriptional comparison of *in vitro* and *in vivo* isolated human pluripotent cells. Hierarchical clustering of the mean expression profile of differentially expressed genes between naive and primed samples (FDR, 0.05), in the different hPSCs groups: naive (NHSM), primed, previously published primed ES cells, and human ICM (Belmonte and colleagues[47]), using Spearman correlation. E) Unbiased global transcriptional (n=9,803 ortholog genes) cross-species hierarchical clustering of naive and primed pluripotent cells from mice and humans. Correlation matrix of gene expression was clustered using Spearman correlation and average linkage. Color bar indicates correlation strength. Each row/column represents an independent cell line or clone. The human cell lines are the ones shown in (C), and the mouse gene profiles were taken from Hanna et al [20].

Next we compared the active and repressive transcription marks H3K4me3 and H3K27me3, respectively. Since systematic comparison of these chromatin marks between naïve and primed state in the mouse was not published yet, we first looked at the position of these marks in mouse ESCs. As expected, mouse naïve cells show global low level of H3K27me3 in gene bodies and even decreased level on developmental genes. Many developmental genes retain a bivalent promoter architecture that contains both H3K27me3 and H3K4me3 signals in primed state. The amount of H3K4me3 on genes and promoters is also lower in the mouse naïve state, and again, more predominantly in developmental genes (Figure 4B,D). These results are similar to previous comparison between mouse naïve and serum+Lif conditions, thus emphasizing the naïve state of cells grown in the presence of 2i.

Looking at five different hPSC lines, we found that the level of both H3K4me3 and H3K27me3 marks around genes is generally lower in the NHSM samples compared with samples of conventional human medium. This phenomenon is comparable to the differences between mouse naïve and primed cells, although to a lower extent (Figure 4B). Focusing on gene groups, developmental genes showed marked reduction of H3K27me3 in NHSM condition, which was more extensive than the average reduction (Figure 4D). For example, HOX genes such as HOXA1 and HOXA9, and the mesoderm and endoderm lineages genes GATA6 and NKX2-5 show a prominent reduction in H3K27me3, almost to the background level in naïve lines, together with reduction in

H3K4me3 level (Figure 4E). The low level of H3k27me3 leads to massive decrease in bivalent genes in human cells grown in NHSM, that only have a minimal number of 226 bivalent genes, while conventional hPSCs have on average 3,013 bivalent genes. The total amount of K27me3 methylated histone 3 is similar between human cells in NHSM and conventional human cells, as judged by western blot (Figure 4C). It is possible that in cells grown in NHSM these methylated histones are redistributed to intergenic areas instead of genes and promoters, or that they are not bound to the DNA.

Enhancers are control sequences residing more than 1000bp from adjacent genes, and are recognized by H3K4me1 chromatin mark and p300 binding. It was shown that conventional hESCs have two types of enhancers: class I, composed of H3K4me1 and H3K27Ac that are associated with active transcription of nearby genes, and Class II, composed of H3K4me1 and H3K27me3 marks, that is associated with low expression of nearby genes that are considered poised for activation[24]. Amazingly, ChIP-seq for H3K4me1 enhancer mark together with H3K27me3 and H3K27Ac in mouse naïve and primed ESCs revealed that naïve mESCs completely lack class II enhancers, while class I enhancers are present in both states and increase in number in the naïve state (Figure 5A). Similarly, hPSCs samples in conventional medium have 5.8-fold higher amount of class II enhancers over NHSM samples, while the number of class I enhancers is similar between primed and NHSM samples, with slight increase in the NHSM sample (Figure 5A). The genes that are associated with class II enhancers are significantly enriched for developmental functions and morphogenesis, again emphasizing the primed state of the these genes that are ready to be activated only in primed pluripotency (Figure 5B). Therefore, the presence of class II poised enhancers, similar to bivalent genes, is a mark of the primed state of pluripotency in mouse. In hPSCs, the level of class II enhancers is high in conventional/primed PSCs and low when NHSM medium is applied.

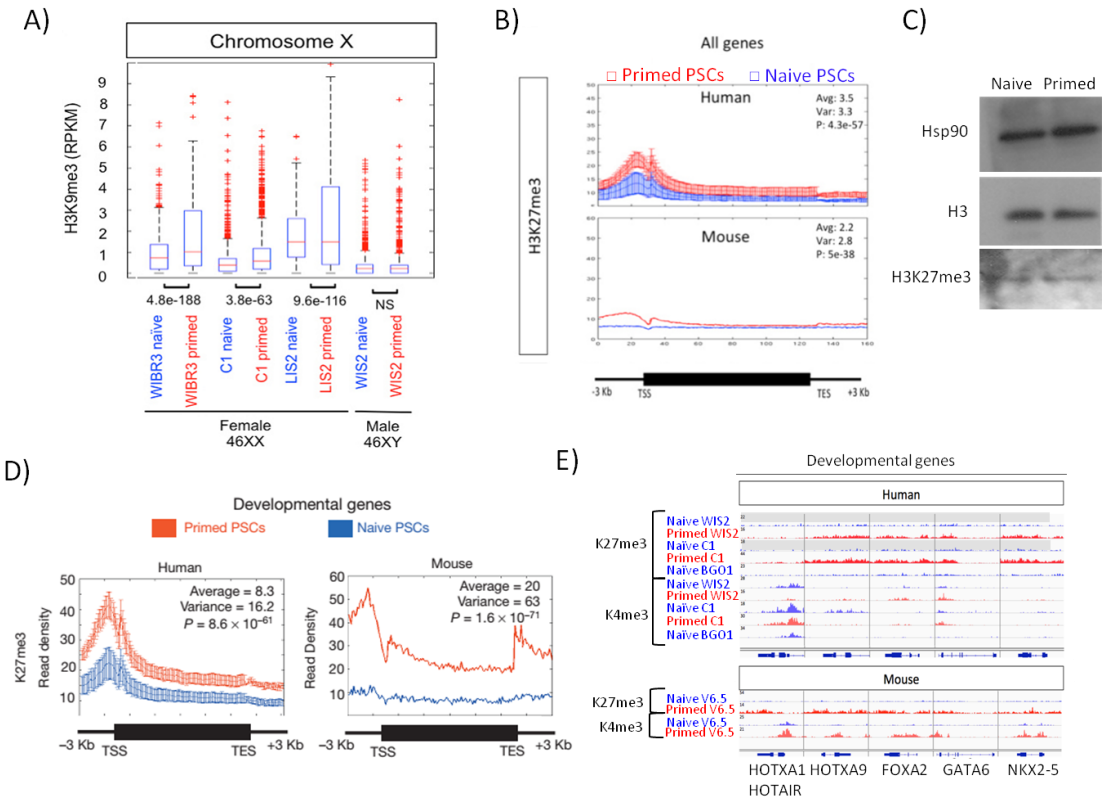


Figure 4: Epigenetic landscape of naïve pluripotency in human. A) RPKM (Reads Per Kilobase per Million reads) of H3K9me3 levels in chromosome X genes, measured in 4 different human cell lines. Boxes, 25th and 75th percentiles; horizontal lines, median; crosses, outliers. RPKM were measured for each gene on the region starting 1kb upstream of the TSS and ending at the TES. Distributions of H3K9me3 in female primed cells are significantly higher compared to their naive counterparts, whereas in male cells they are the same. P values were calculated with one-tail paired-sample t-test. B) Global H3K27me3 and H3K4me3 deposition in naive and primed pluripotent cells. Profiles of H3K27me3 and H3K4me3 chromatin modifications of all RefSeq genes in human (n=43,463) and mouse (n=30,480), naive (blue) and primed (red), represented as normalized read-density. Human profiles indicate average and s.d. (error-bars) calculated over 5 different cell lines for naïve condition and 4 cell lines for primed conditions. Average difference between plots is indicated alongside variance and P values (calculated with paired-sample t-test). C) Western blot for H3K27me3 in Lis39 cell line naïve and primed cells, showing similar levels of the modification. D) H3K27me3 distribution only in developmental genes (human n=5,922, mouse n=420), showing marked reduction in both mouse and human naïve states. E) Representative examples of H3K27me3 and H3K4me3 modifications on developmental genes in naïve and primed hPSCs lines.

Cross species chimerism of hPSCs grown in NHSM

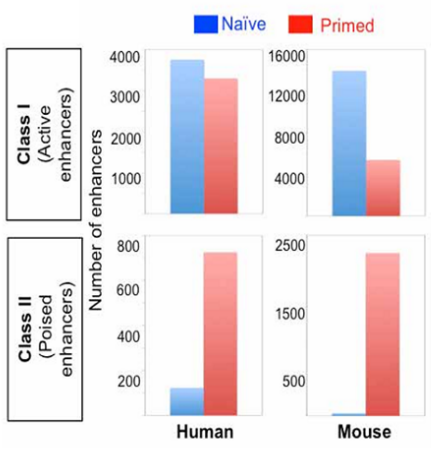
In mouse cells, naïve pluripotency can easily be tested through injection to early embryos that result in chimeric embryos. Injections to human embryos are, of course, practically and ethically forbidden and therefore the ability of hPSCs to contribute to all germ layers *in vivo* can be tested only through cross-species chimerism. Mouse-rat chimeras through ESCs injections were reported, although the level of contribution of the injected cells was low, probably due to inter-species differences in signaling and developmental patterns [48]. Ohad Gafni and Mirie Zerbib from my lab injected naïve human cells into mouse morula, and found that the cells integrated into the mouse embryo. The human GFP-positive cells managed to survive, proliferate and integrate in the embryo. *In-toto* analysis of the injected embryos revealed integration of human cells in several places in the mouse E10.5 embryo. Therefore, NHSM provides hPSCs with the ability to incorporate into the mouse embryo.

To conclude, the novel NHSM medium maintain hPSCs in a pluripotent state that demonstrate many naïve characteristics when compared to conventional medium, including gene expression, epigenetic landscape, surface marker expression and the ability to integrate in mouse blastocyst. Therefore, hPSCs in NHSM acquire a naïve state of pluripotency that might not be identical to mPSCs 2i/Lif medium but clearly display some naïve characteristics.

Improvement of NHSM medium to a more ICM-like state

It was previously reported that hESCs, as well as mouse primed cells, can survive in the presence of the GSK3β inhibitor CHIR when the Tankyrase inhibitor IWR1 was added to the medium, resulting in high level of cytoplasmic β-catenin[18]. We therefore speculated that addition of IWR1 to NHSM medium could replace the primed-related cytokines FGF2 and TGFβ.

A)



B)

Class II enhancers of human primed

GO annotation	P Value
Transcription factor activity	5.96E-77
Multicellular organismal development	5.15E-61
RNA metabolic process	1.75E-57
Developmental process	5.56E-49
System development	6.88E-45
Organ development	2.01E-35
Anatomical structure morphogenesis	1.41E-32
Nervous system development	2.02E-32
Embryonic development	1.40E-25
Organ morphogenesis	4.22E-24
Embryonic morphogenesis	6.25E-24
Central nervous system development	6.92E-20
Pattern specification process	8.42E-20
Neurogenesis	1.72E-19
Brain development	7.40E-18
Cell differentiation	3.01E-16
Cell fate commitment	4.35E-15
Regionalization	1.08E-14

Figure 5: Enhancer classes in human and mouse naïve and primed. A) Number of enhancers of class I (active transcription, H3K4me and H3K27Ac) and class II (poised, H3K4me1 and H3K37me3) in naive and primed, human and mouse cells. For human, numbers correspond to common enhancers in lines WIBR3 and C1. A marked decrease in class II enhancers can be seen in the naive cells of both human and mouse. B) GO annotation of genes that are related to class II enhancers of primed human samples from (A). 18 selected annotations are shown (from the top 25 annotations), showing enrichment for transcription factors and developmental processes, as expected from primed enhancers.

In order to quantify the effect of IWR1, as well as other factors, we utilized three hESCs reporter lines: H9-NANOG-GFP, WIBR3-OCT4-GFP knock-in, and WIBR3-ΔPE-OCT4-GFP knock-in. The latter is a mutated version of the WIBR3-OCT4-GFP line, where the OCT4 proximal enhancer was deleted and only the distal enhancer, which is active in the naïve state, remained[49].

At first we noticed that addition of IWR1 together with PKCi to NHSM resulted in smoother and more homogenous colony morphology (Figure 6A). This combination enabled hESCs survival without FGF2 and TGFβ but morphological heterogeneity between colonies was observed. The heterogeneity was also noticeable by immunostaining for pluripotency markers and from the varying levels of fluorescent

reporters. We then observed that addition of high dose ACTIVIN A (a ligand from the TGFβ family) to the medium (20ng/ml) resulted in increased percentage of OCT4+ cells and NANOG+ cells (Figure 6D). Immunostaining also supported a more homogenous expression of pluripotency genes within the cell population (Figure 6B). Interestingly, the expression of OCT4 was dependent on the distal enhancer, as witnessed in the high percentage of GFP+ cells (~60%) in the WIBR3-ΔPE-OCT4-GFP line. Other TGFβ ligands, such as TGFβ (2ng/ml) , GDF3 and GDF9 did not fully recapture the effect of high dose ACTIVIN A (20ng/ml). We set the term ENHSM - Enhanced NHSM – for the new naïve medium, which included 6 inhibitors: MEKi, GSK3i, JNKi, p38i, PKCi and IWR1, together with LIF and ACTIVIN A. RT-PCR analysis of cells that were grown in ENHSM showed high levels of the naïve related transcripts STELLA, KLF2, KLF4 and TFCP2L1 (figure 6C).

One prominent mark of mouse naïve state is the ability to overcome knock-out of epigenetic regulators, such as DNMT1, EED, G9A, METTL3 and more. In many of the knock-outs, the mouse naïve stem cells are normal but show defects in differentiation and lethality during development. In contrast, conventional hPSCs cannot survive DNMT1 knock-out and rapidly die when the expression of this gene is brought to a halt[50].

In order to test the ability of ENHSM to withstand removal of epigenetic modifications, we utilized a cell line that encompasses a knock-out of DNMT1 together with TET-OFF copy of the gene. Addition of Doxycycline (DOX) to the cells results is complete impediment of DNMT1 expression and removal of DNA methylation, as DNMT1 is the sole maintenance methyl-transferase of the DNA[50]. Remarkably, when DOX was added to hESCs grown in ENHSM, the growth rate of the cells was decreased, but they continued to grow for over 5 weeks (6 passages) and remained pluripotent, as seen by SSEA4 staining (Figure 6E). This result shows that ENHSM renders human cells resistant to epigenetic modifications, supplying another evidence for a pluripotency state that is similar to the naïve state of the mouse.

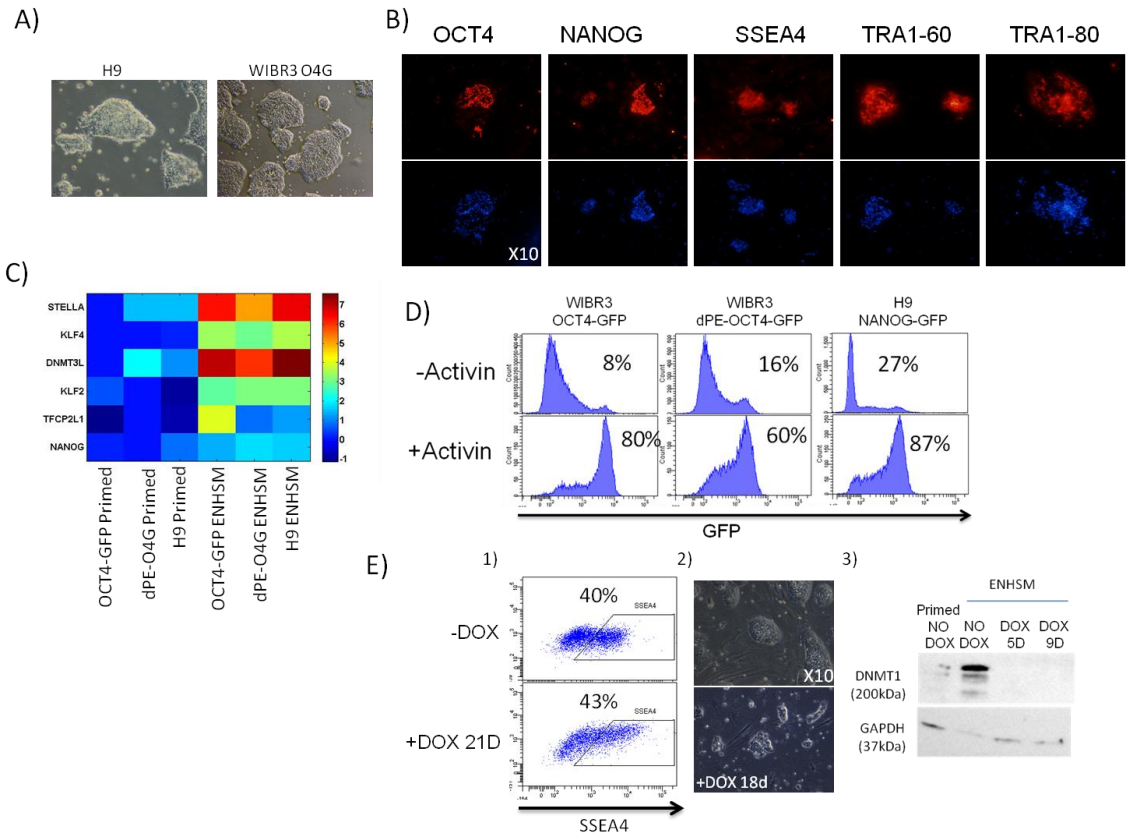


Figure 6: ENHSM – naïve human medium without FGF2 and TGFβ. A) Colony morphology two hESC lines (H9 and WIBR3) in ENHSM medium. B) Immunofluorescence staining of WIBR3 hESCs for different pluripotency markers in WIBR3 hESCs grown in ENHSM. Blue-DAPI. C) Expression level by RT-PCR of naïve related transcripts in primed and ENHSM mediums in three cell lines. Colors represent log(RQ). D) Histogram of fluorescence level of three reporters: NANOG-GFP, OCT4-GFP and ΔPE-OCT4-GFP in ENHSM with and without ACTIVIN. E) HuES64 cell line and DNMT1 knock-out. 1) FACS analysis of cells with and without induction of DNMT1 knockout for 21 days. Cells were stained with SSEA4 antibody, a marker for human pluripotent cells. 2) Colony morphology of cells 18 days after induction of DNMT1 knock out. 3) Western blot showing DNMT1 protein levels after DOX addition.

In general, my work supports a reversible naïve state of pluripotency in human cells, which was not shown before without exogenous gene expression. Differences in the development of mouse and human, such as expression of ACTIVIN in human blastocyst only and not in mouse[51], can explain some of the differences in the composition of the human and mouse naïve mediums. Further characterization and fine tuning of ENHSM components will continue to contribute to the study of human development and stem cells.

Part II: Human ESCs differentiation to Primordial germ cells-like-cell (hPGCLC)

Most of the data presented here was published in

"*SOX17 is a critical specifier of human primordial germ cell fate"*. Cell, 2015[52].

Establishment of hPGCLCs from hESCs

After characterization of conditions that support hESCs naïve state of pluripotency, we sought to find out whether naïve hESCs are more prone for *in vitro* differentiation to primordial germ cells (PGC). For this purpose, an approach similar to the one that was published for mESCs to PGCs was chosen[37]. This protocol includes priming with FGF2+ ACTIVIN A of naïve stem cells for 2 days resulting in EpiLC (EbiSC-Like-Cells), followed by seeding in suspension plates to form small cell aggregates (2000-4000 cells/well) in PGC medium that includes high concentration of BMP4, together with SCF, EGF and LIF (Figure 7A). Indeed, when human naïve ESCs (NHSM) were differentiated using this protocol, the expression of several PGC related genes, such as BLIMP1, NANOS and AP2γ, was increased (Figure 7B). In order to identify cells that commit to germ lineage program, we created a knock-in reporter for NANOS3-mCherry (Figure 7C). NANOS3 is a conserved RNA-binding protein that is expressed in vertebrates PGCs specifically, both in males and females[53]. Knock-down of NANOS3 results in reduced number of PGCs and reduced expression of PGC related genes, providing evidence for the functional role of NANOS3 in germ cell development. Thus NANOS3 reporter can assist in identification of cell populations that began the commitment to the germ lineage. We inserted the NANOS3-mCherry reporter into 3 different hESC lines: WIBR3 (female), WIS1 and LIS38 (male). When cells were grown in conventional or naïve medium, the reporter did not turn on, as expected. Amazingly, 3 days after the PGC medium was added, a small percentage of the cells turned on the NANOS3 reporter, hence presumably differentiated into Primordial Germ Cell Like Cells (PGCLCs). The percentage and intensity of the reporter was further increased in the following days (days 2+4 to 2+6) to 24%±4% and then decreased (Figure 8A). The reporter activity after PGC induction was seen in all the 3 different cell lines, female and

male, at similar kinetics and percentages (Figure 8B). Removal of BMP4 from the PGC differentiation medium resulted in lack of fluorescence signal, while replacement of BMP4 with BMP2 stimulated reporter activity to a lower level than the level of BMP4 (Figure 8A,C). Therefore hESCs begin to express NANOS3 when exposed to conditions that induce germ-lineage differentiation.

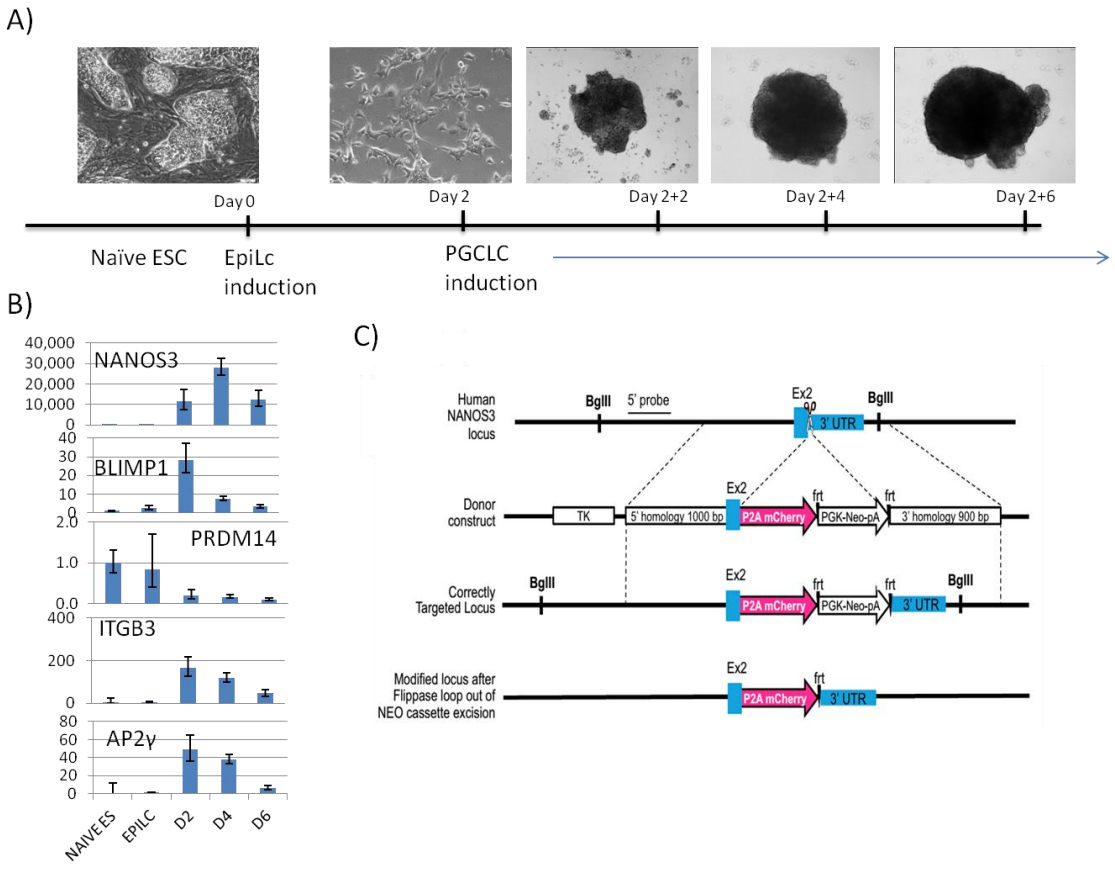


Figure 7: Differentiation of hESCs to PGCs. A) Timeline of differentiation of human naïve cells to PGCs, through a 2 days priming step (EpiLC) and 4-6 days of growth in suspension in hPGC medium, resulting in floating embryonic bodies (EBs). B) RT-PCR analysis for PGC-related genes during hPGC differentiation. C) Targeting strategy of generation of NANOS3-mCherry knock-in reporter hESC lines. P2A-mCherry sequence was inserted in frame with the last exon of the human NANOS3 gene.

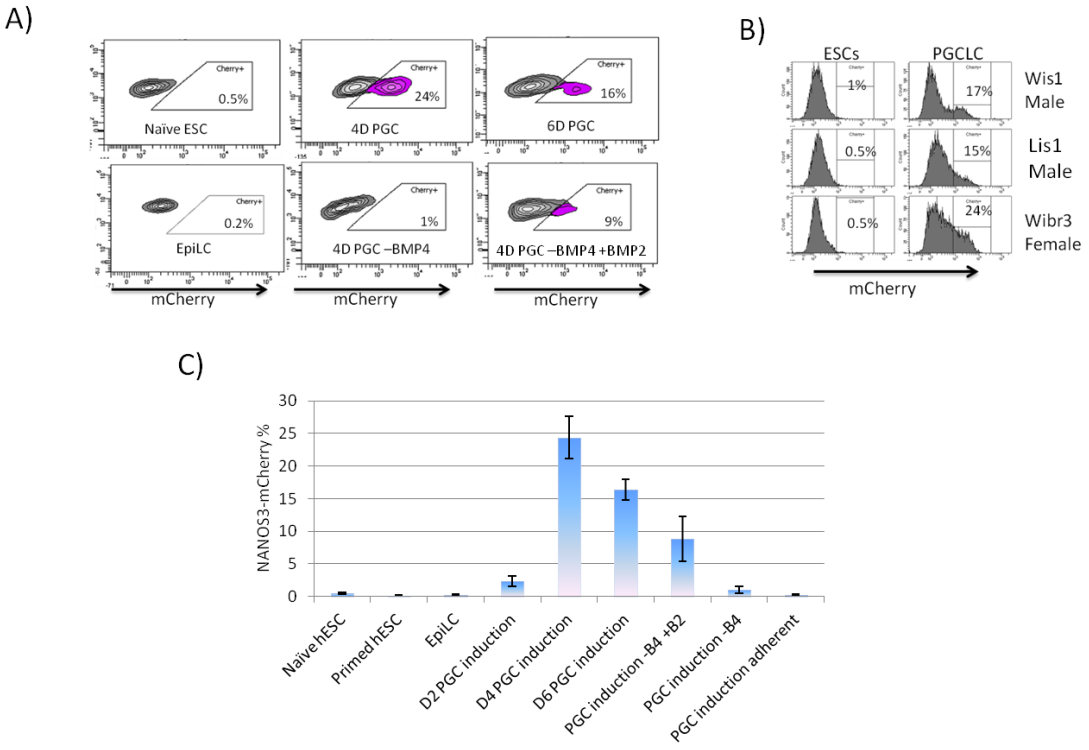


Figure 8: Activity of NANOS3-mCherry reporter during hESCs differentiation to hPGCs. A) WIS1 hESCs carrying NANOS3-mCherry reporter was differentiated to PGCLCs and NANOS3-mCherry level was measured through flow cytometry at different steps of differentiation and with interruptions: without BMP4 or with BMP2 instead of BMP4. B) NANOS3-mCherry reporter was introduced to 3 hESC lines (WIS1, LIS1, WIBR3) and the fluorescence level was measured after 4 days of PGCLC induction. C) NANOS3-mCherry level in WIS1 cell line in different conditions. Shown is average of at least 3 biological repeats. Error bars represent standard error.

Gene expression profile of hPGCLCs

We went on to test whether the *in vitro* differentiated PGCLCs indeed resemble human gonadal PGCs and are indeed hPGCLCs. For this purpose, RNA was extracted from WIS1 hESCs at days 2, 4 and 6 of hPGC induction (days 4 and 6 were sorted for NANOS3-mCherry+), as well as from human naïve (NHSM), human primed, and human EpiLCs (after 2 days of priming with FGF2 and ACTIVIN A). The RNA was

subjected to RNA-Seq analysis, using high-throughput sequencing in Ilumina Hi-seq machine and the recovered reads were aligned to the human genome. PGC specific gene expression can be divided to 3 categories: early PGC genes, late PGC genes and pluripotency genes. Human NANOS3-mcherry+ PGCLCs express high levels of early PGC genes such as BLIMP1, AP2γ, ITGB3, KIT, NANOS3, as well as pluripotency genes: OCT4 and NANOG, most of them already from day 2 of differentiation, before the NANOS3-mCherry reporter turned on (Figure 9C). Expression of late PGC genes that are related to meiosis such as BOLL, VASA, DAZL were absent from the differentiated cells. Next, we used hierarchical clustering to identify similarities between samples. Indeed, differentiated hPGCLCs clustered together, specifically samples from day 4 and day 6 of differentiation, while naïve ESCs clustered with cells that were primed for 2 days (EpiLC sample) and with primed ESCs (Figure 9B). Principle component analysis revealed the pattern of differentiation through the 1^{st} and 2^{nd} components. In this PCA plot of the six samples, naïve ESCs and PGCLCs day 6 were placed at the extreme positions of the graph, while the path of differentiation, through 2 days priming, PGCLCs day 2 and PGCLCs day 4 were ordered between the beginning and end states, while primed hESCs settled further aside, thus the trajectory of PGCLC differentiation can be inferred from the PCA analysis (Figure 9A).

Authentic human early PGCs are unavailable for research for technical and ethical reasons, and therefore it is impossible to compare PGCLCs to their *in vivo* counterpart. Instead, human NANOS3-mCherry PGCLCs were compared to late hPGCs that were isolated from weeks 5.5-9 of pregnancy. In this stage, the PGCs already finished their migration, occupy the embryonic gonadal ridge and began sexual differentiation, similar to mouse E11-E13 PGCs. A comparison of a dataset of RNA-seq of hPGCs from 10 embryos with hPGCLCs revealed that hPGCLCs are indeed more similar to authentic late PGCs compared to somatic gonadal cells, while hESCs are similar to PGCLCs but correlate with PGCs to a low extent (Figure 9D). Thus, NANOS3-mCherry human cells can be considered hPGCLCs, at least in terms of gene expression, and probably represents the *in-vitro* counterpart of early hPGCs (E16-E28), similar to mPGCLC that represent early mPGCs (E9).

jkgh

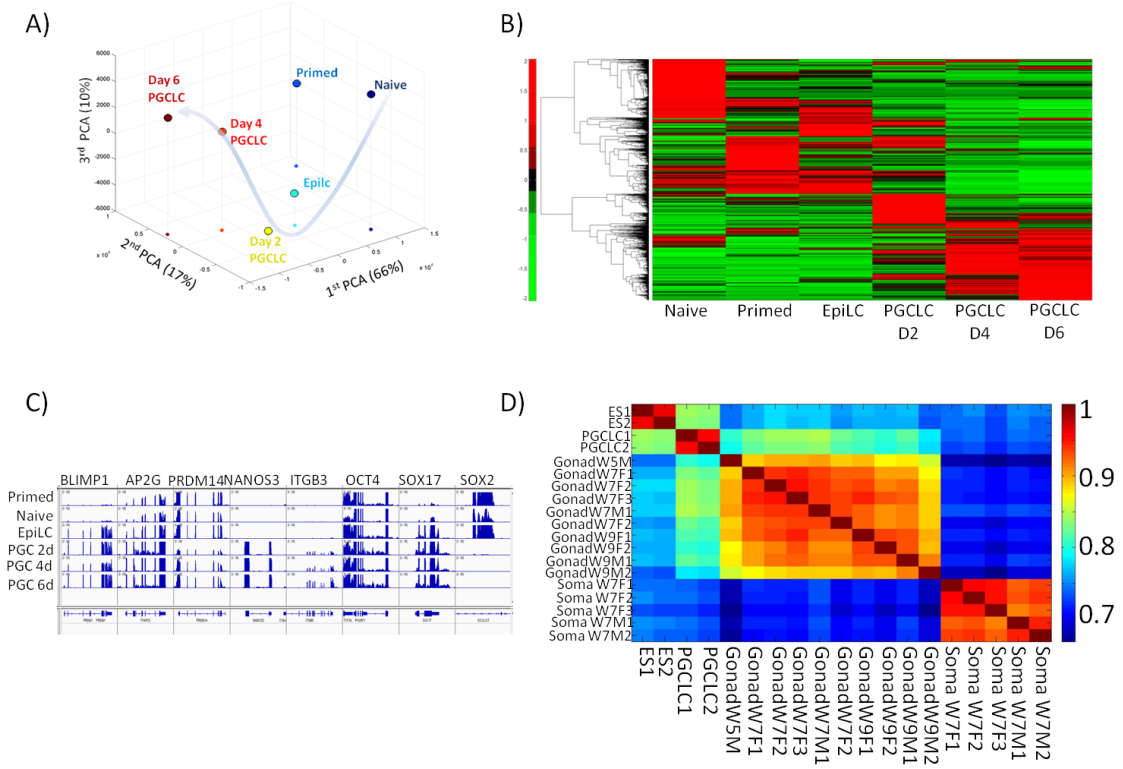


Figure 9: Gene expression profiles of hPGCs. WIS1 hESCs were differentiated to PGCLCs and RNA was collected from the differentiated cells. At day 4 and 6 RNA was extracted only from the NANOS3-mCherry sorted population. A) PCA of gene expression (RNA-Seq) of naïve and primed hESCs (WIS1) and during PGCLC differentiation. B) Hierarchical clustering of gene expression (RNA-Seq) of different cell states during hESCs differentiation to PGCLCs. C) Examples of the expression of PGC related genes during hPGCLC differentiation, shown in IGV. D) Correlation matrix of gene expression (RNA-Seq) between hESCs (H9), hPGCLCs (WIS1) and 10 different human gonadal PGCs and gonadal soma cells from human embryos. W-week of pregnancy. F/M - female/male (data taken from Tang et al.[35])

PGCLC differentiation efficiency in naïve and primed

The ability to differentiate to PGCLCs in high efficiency is unique to naïve cells in the mouse, and requires two days of priming. In order to test whether this is also the case in human, hESCs carrying the NANOS3-mcherry were transferred from the naïve NHSM medium to conventional (primed) human medium with FGF2 for 4 passages, resulting in a homogenous population of primed cells. Next, the primed cells were differentiated to

PGCLCs, either with priming step or without. In both cases, only a minor fraction of the cells (up to 5%) turned on the NANOS3 reporter (Figure 10A). RT-PCR for germ cell genes on all the cell population, compared to all the population of naïve cells that were differentiated showed that the primed cells indeed do not turn on the PGC genes (Figure 10B). In order to determine possible reasons for the inability of primed cells to differentiate to PGCLCs, analysis of ChIP-Seq data for chromatin marks on naïve and primed human samples was performed. Amazingly, primed hESCs display a high level of the repressive chromatin mark H3K27me3 on the gene body of NANOS3 and SOX17, two crucial factors in hPGC lineage (Figure 10C). Therefore, inhibitory chromatin state could be responsible for the inefficiency of human primed cells to differentiate to the germ lineage.

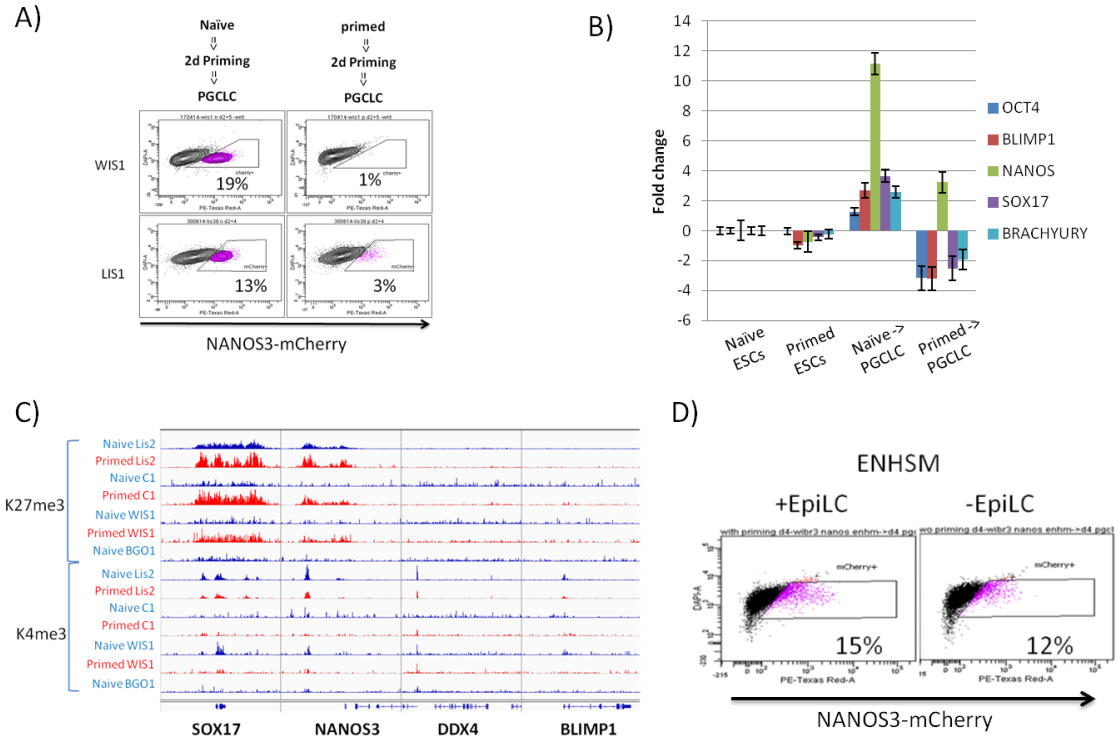


Figure 10: Primed hESCs are inefficient in PGCLC differentiation. A) Naïve (NHSM) and primed hESCs (WIS1 and LIS1 cell lines) were induced to differentiate to PGCs, and NANOS3-mCherry level was measured through FACS. B) RT-PCR analysis of PGC related genes in naïve (NHSM) and primed hESCs and during their differentiation to PGCLCs. RNA was extracted from all the population at day 4 of differentiation.

Results are in logarithmic scale. Error bars represent RQ min and max. C) H3K27me3 and H3K4me3 levels around the TSS of early (SOX17, NANOS3, BLIMP1) and late (DDX4) PGC genes in naïve (NHSM) and primed human cell lines (4 lines: LIS2, C1, WIS1, BGO1). D) WIBR3 ESCs were grown in ENHSM medium and differentiated to PGCs either with 2 days priming (+EpiLC) or without (-EpiLC). NANOS3-mCherry level was measured by FACS.

Next, we looked whether a priming step is absolutely required for naïve human ESCs differentiation to PGCLCs. In mouse cells, 2 days priming step enhances the differentiation efficiency. On the other hand, when we directly transferred hESCs to PGC conditions, without priming, they were able to differentiate in similar efficiency to cells that were primed for 2 days. It is possible that human pluripotency states are divergent from the mouse ones, and represent an intermediate state that already acquired the "priming" necessary for PGC differentiation. It is also possible that the differences in embryonic development and PGC specification between human and mouse are responsible for the difference in priming requirement of the two species. Other naïve mediums, including ENHSM that does not include the priming factors FGF2 and TGFβ (but include ACTIVIN A) could also support differentiate to PGCLCs either with or without priming (Figure 10D).

Differences between human and mouse PGCLCs

PRDM14 was attributed a main role in the transcription network for mouse PGCs specification, together with AP2γ and BLIMP1[31]. In the human differentiated cells PRDM14 was expressed to a low level, much lower than in ESCs. Indeed, PRDM14 is expressed in human both in naïve and primed states, while in the mouse it is expressed only in naïve state, and authentic hPGCs from weeks 5-9 of pregnancy also show low level of PRDM14, compared with ESCs (Figure 9C). A new paper showed that knock-down of PRDM14 had no effect on human PGC differentiation[54]. Therefore, PRDM14 seems to have a different role in human germ cells than in mouse cells.

Another gene that unexpectedly showed low-absent levels of expression in the NANOS3-mcherry hPGCLCs is SOX2. SOX2 is a pluripotency gene that is highly

expressed in human and mouse PGCs. While mouse PGCs and PGCLCs present high level of SOX2, similar to other pluripotency-related genes (OCT4, NANOG, KLF4), human PGCLCs do not express SOX2, in accordance to human authentic PGCs that show absence of SOX2 both at the mRNA level and at the protein level[35, 36]. On the other hand, SOX17 is highly expressed in hPGCLC (Figure 9C). SOX17 is a master regulator of the Endoderm lineage and has no role in mouse PGCs development[55]. Since SOX2 is known to have a role in mPGC development, it is possible that another member of the SOX gene family, such as SOX17, is replacing SOX2 in human germ line.

Integrin β3 is a specific PGCLC marker

Identification of specific surface markers for somatic lineages can assist in isolation of certain cell types, without genetic manipulation such as insertion of fluorescent reporters. Such surface markers can be utilized for cell sorting of cells from various sources, including pre-established mutant lines and patient-specific iPSCs, so we sought to discover markers that are specific to hPGCLCs. For this purpose hPGCLCs that carry the NANOS3-mCherry reporter were stained with a variety of pluripotency and germ line associated antibodies. The hPGCLCs population (mCherry+) specifically stained for integrin β3 (ITGB3, CD61) similar to mPGCLCs, while Alkaline-phosphatase (TNAP) was highly expressed in the PGCLCs population but also in hESCs (Figure 11A). Next, the hiPSC line FX17 and the hESC line BGO1 were differentiated to PGCLCs, stained with ITGB3 antibody and FACS-sorted. ITGB3+ population in both lines expressed high levels of PGC related transcripts such as BLIMP1 and AP2γ compared to the ITGB3- population (Figure 11B-C). The hESCs line BGO1 carry an OCT4-GFP reporter, which was expressed specifically in the ITGB3+ population, as expected in hPGCLCs. Therefore Integrin β3 can be utilized to isolate hPGCLCs from different human ESCs and iPSCs lines. The ability to isolate PGCLC population without the use of a reporter can support future study of different cell lines, including iPSCs from patients that suffer infertility conditions.

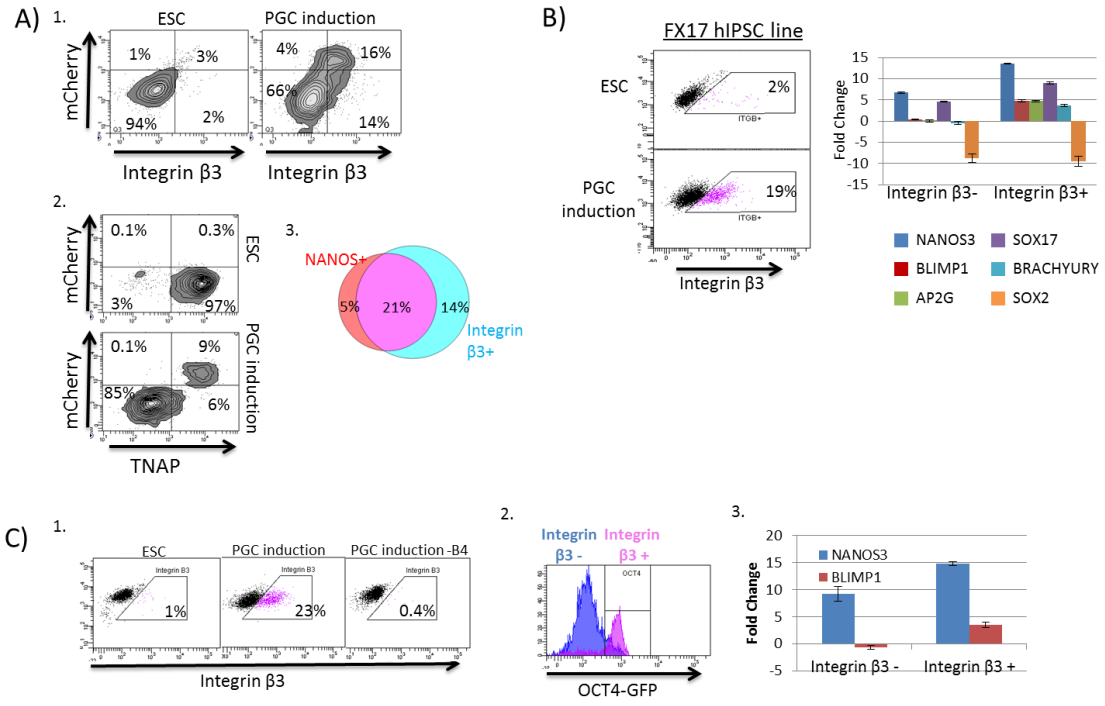


Figure 11: PGCLC identification through surface markers. A) Integrin β3 and Alkaline phosphatase (TNAP) are specific surface markers for NANOS3-mCherry expressing PGCLCs. WIS1 hESCs were induced to differentiate to PGCLCs and stained with: 1. Integrin β3 antibody or 2. TNAP antibody, followed by FACS analysis. NANOS3 expressing population overlaps with Integrin β3 and with Alkaline phosphatase expressing cells. 3. Venn diagram illustrating the percentage of NANOS-mCherry and Integrin β3 populations in PGCLCs. B) hiPSC line FX17 was differentiated to PGCLCs and stained for Integrin β3. The ITGB3 positive and negative population were FACS-sorted, RNA was extracted and RT-PCR analysis for PGC-related gene expression was carried. Results are shown on a logarithmic scale. Error bars represent RQ min and max. D) hESC line BGO1, which carries OCT4-GFP reporter was induced to differentiate to PGCLCs. 1. Naïve ESC and differentiated cells, with and without BMP4 were stained for Integrin β3 expression. 2. Histogram of OCT4-GFP reporter levels in Integrin β3 positive and negative populations in PGCLCs. 3. RT-PCR analysis for NANOS3 and BLIMP1 was performed on BGO1 derived hPGCLCs, sorted by Integrin β3 expression. Results are shown on a logarithmic scale. Error bars represent RQ min and max.

Knock out of BLIMP1 precludes hPGC specification

BLIMP1 is the first and key regulator of mPGCs, and its loss of function abrogates PGC fate[31, 56]. Knock-down of BLIMP1 in human cells was also shown to suppress PGC

differentiation *in vitro*[41]. In order to test the role of BLIMP1 in depth, BLIMP1 loss of function mutant was generated in the NANOS3-mCherry hESC background (WIS1 line) using gRNA and CRISPR (Figure 12C). Notably, NANOS3-mCherry reporter did not turn on in the BLIMP1 mutant cells (Figure 12A). The expression of OCT4, NANOS3 and SOX17 was not increased during PGCLC induction, in contrast to WT cells, and SOX2 expression remained high and was not downregulated. On the other hand, KIT and AP2γ expression was maintained in the BLIMP1 knock-out cells (Figure 12C). This observation is very similar to the effects of Blimp1 mutation on mPGC specification and suggests a conserved role for BLIMP1 in the regulation of the early steps of germ cell specification.

The role of SOX17 in hPGCs

SOX17 was one of the genes that were highly expressed early during hPGCLC induction. In the mouse germ lineage, Sox17 was shown to be expressed in PGCs but its expression in the PGCs themselves is dispensable for the intrinsic function of PGCs, while expression by surrounding cells is important for PGCs migration to the gonads[55]. Since we speculated that SOX17 can replace SOX2 in hPGCs and in order to study the role of SOX17 in the system, a knock out mutation of SOX17 was inserted in the NANOS3-mCherry WIS1 cell line (Figure 12E). After verification of the knock-out, PGCLC differentiation was carried out. Similar to BLIMP1 knock-out, the NANOS3-mCherry reporter did not turn on during the induction (Figure 12A). RT-PCR analysis revealed that PGC related transcripts were not upregulated and SOX2 was not down regulated (Figure 12D). BLIMP1 did not turn on in the SOX17 knock-out cells, suggesting that SOX17 is required for BLIMP1 induction. Since BLIMP1 mutant also prevented the expression of SOX17, a feedback loop might be required for stabilization of these two transcription factors and maintenance of germ lineage fate. In order to test whether the effect of SOX17 in cell autonomous, SOX17 knock-out cells were mixed with WT cells (LIS1 hESC line) that were labeled with GFP in a 1:1 ratio, PGCLC differentiation was carried out and the cells were stained with ITGB3 antibody after 4 days. Remarkably, the GFP+ population (WT) expressed ITGB3 (Integrin β3) in 10% of the population, while the GFP- population ($SOX17^{-/-}$) did not express ITGB3 at all, proving that SOX17 is required intrinsically in order to employ germ cell differentiation

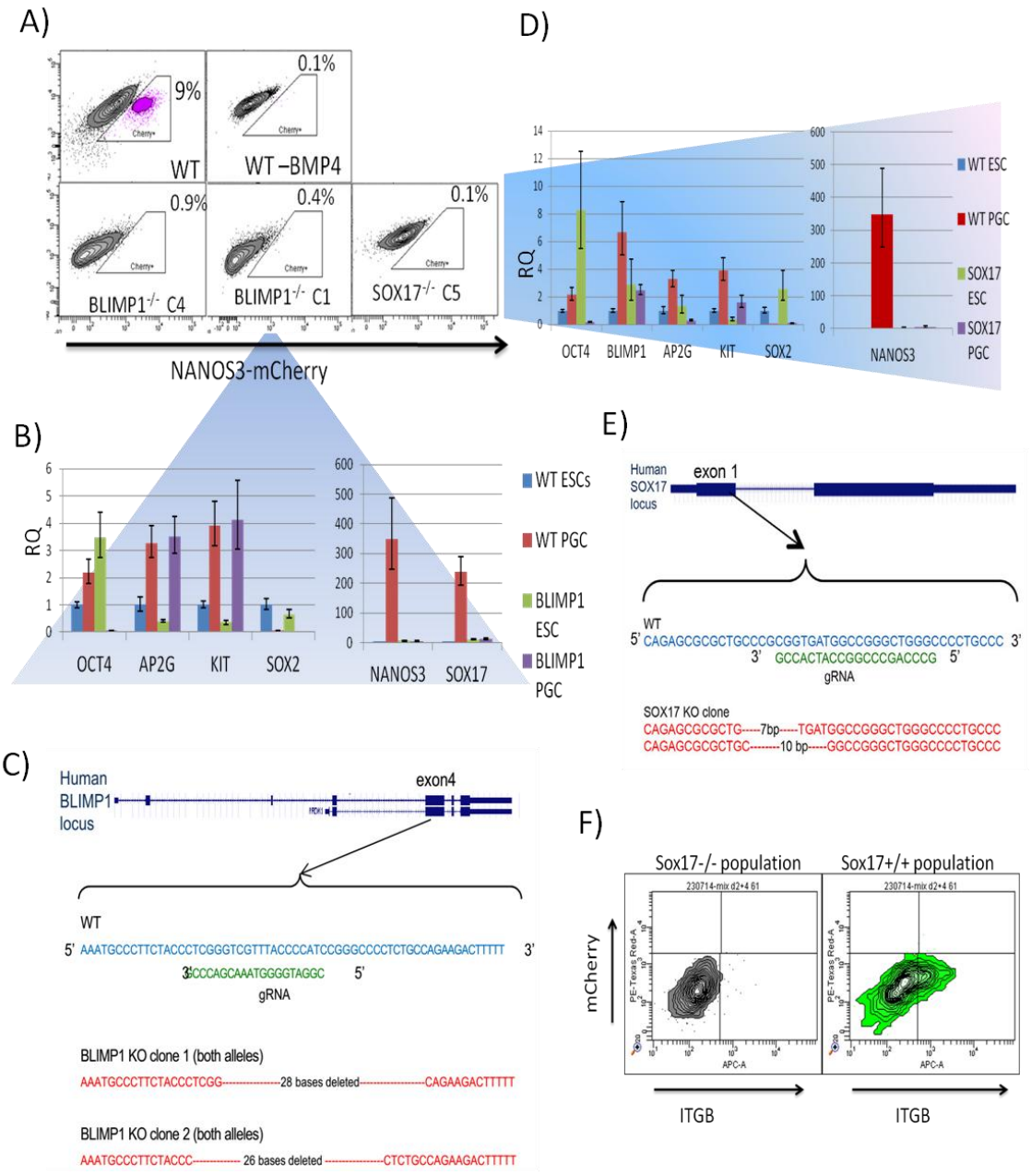


Figure 12: BLIMP1 and SOX17 play important roles in PGC specification. A) NANOS3-mCherry levels in WT, BLIMP1 knock-out and SOX17 knock-out (WIS1 cell line) after PGCLC induction. B) RT-PCR analysis of the expression of PGC related transcripts in WT and BLIMP1 knock-out, from all the population after PGCLC induction. C) Schematic representation and sequence of BLIMP1 knock out. D) RT-PCR analysis of the expression of PGC related transcripts in WT and SOX17 knock-out, from all the population after PGCLC induction. E) Schematic representation and sequence of SOX17 knock-out. F) SOX17 effect is cell autonomous: WT cells (LIS1 expressing constitutive GFP) and SOX17 knock-out (WIS1, carrying NANOS3-

mCHERRY reporter) cells were mixed in a ratio of 1:1 and differentiated to PGCLCs. In the WT (GFP+) population ~10% of the cells were stained for integrin β3. In contrast, in the KO (GFP-) population, NANOS3-mCherry reporter was not turned on, and integrin β3 staining was negative.

program (Figure 12F). We conclude that SOX17 is an early regulator of PGC differentiation that is specific to human and not to the mouse.

Efforts to differentiate hPGCLCs to spermatozoa

mPGCLCs are capable of fully developing to fertile sperm after injection to the testes of young (7-10 days old) sterile mice[37]. Seemingly, the spermatid niche inside the testes tubules and the surrounding cells can drive PGCLCs to further differentiation, which is currently unreachable *in vitro*. In order to test whether hPGCLCs can acquire a further differentiated state, we attempted to recapitulate the testes injection method[57] in our hands, using human PGCLCs instead of mouse. The testis of young pups carrying the W\W-V mutation in the KIT gene that render them sterile, where injected with hPGCLCs (WIS1 cell line) that were isolated through MACS separation, separating ITGB3 expressing cells only. In order to test whether the injection procedure is steady, mPGCLCs (BVSC cell line) were also injected to W\W-V sterile mice, after FACS sorting for BLIMP1-GFP+ population, similar to previously published procedure. Since FACS sorting can reduce cell viability, in some cases mouse cells were injected directly after PGCLC induction without separation of the differentiated population. A total number of 35 testes were injected with mouse cells and 12 with human cells (Table 1). Technical difficulties occurred during the injection procedure, mostly due to the young age of the pups such that the post-injection viability was low.

The injected testes were fixed and stained in Hematoxylin & Eosin 8-16 weeks after injection. The W\W-V testes are small and show no signs of spermatogenesis, and the injected testes could not be distinguished from the non-injected. Sperm could not be isolated from any of the injected testes, and the morphology was very similar between injected and non-injected testes. Staining for the late germ cell marker Mvh marked few cells in both injected and non-injected testes, while WT testes stain was intense and homogeneous across the tubules. Since the injection of mouse cells were unsuccessful and did not show signs of spermatogenesis, we conclude that technical difficulties

limited our ability to state whether hPGCLCs can further differentiate to spermatogonia in the mouse testis environment.

To conclude, I showed that naïve human PSCs are able to efficiently differentiate to PGC-like-cells that probably represent an early stage of hPGCs specification. This is in contrast to primed hPSCs that are inefficient in the PGCs differentiation. BLIMP1 and SOX17 play important role in human germ cells specification and Integrin β3 is a specific surface marker for hPGCLCs. Technical difficulties prevented us from studying further differentiation of the PGCLCs to late germ cells.

Table 1: List of injected mouse testis, sorted by origin of injected cells and population of cells

Species of injected cells	Cell population	number of injected testis	Results from testis analysis
Mouse	PGCLC only	23	No spermatogenesis by H&E
Mouse	All population of differentiated cells	8	No spermatogenesis by H&E
Human	PGCLC only	15	No spermatogenesis by H&E
Human	All population of differentiated cells	8	No spermatogenesis by H&E

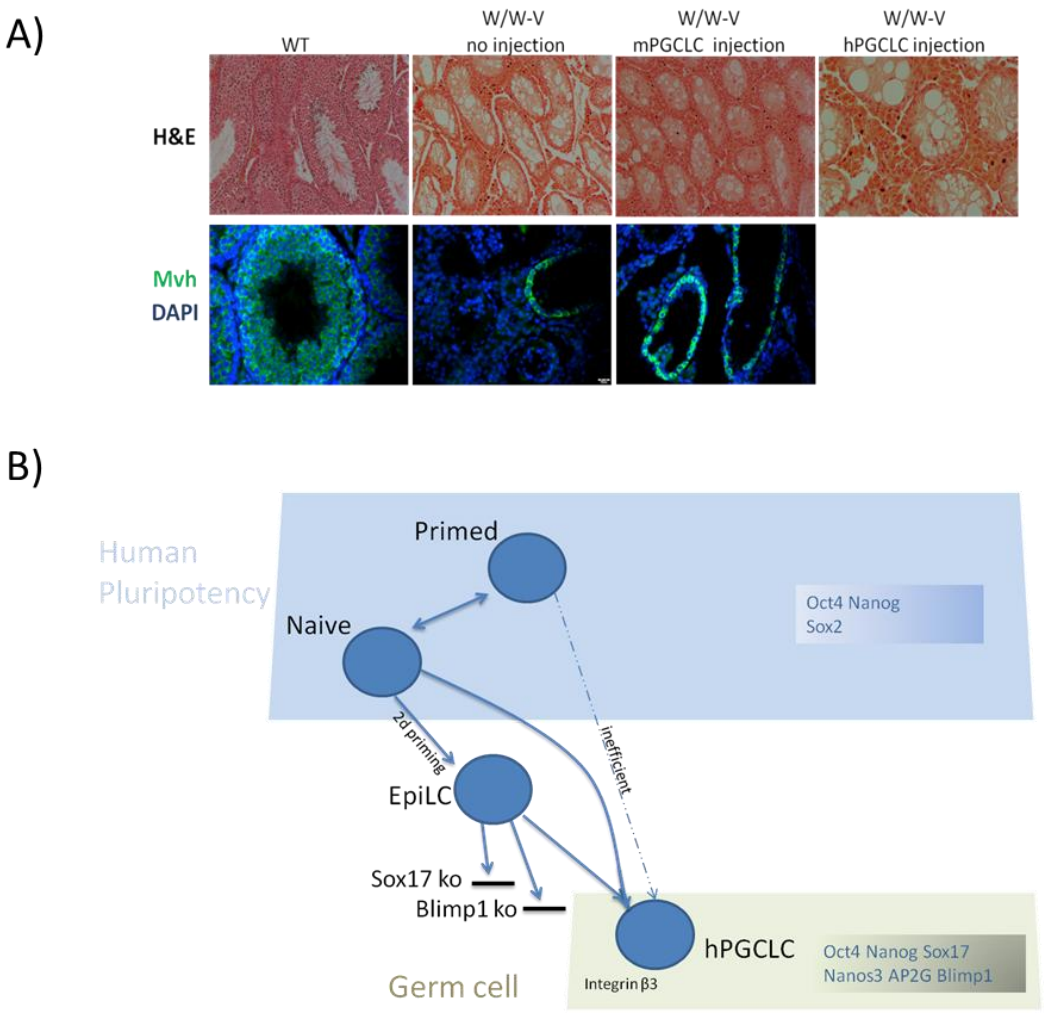


Figure 13: Injections of human and mouse PGCLC to mouse testis and overview of hPGCLCs differentiation. A) Human and mouse PGCLCs were injected into the testis of 10 days old sterile W\W-V mouse pups. Testis were removed and analysis by H&E staining and Mvh (VASA) immunostaining in order to asses spermatogenesis. B) Overview of hPSC differentiation to PGCs *in-vitro*.

Discussion

In the first part of my work, I described the NHSM, a novel defined growth condition for hPSCs that supports a MEK/ERK independent naïve state of pluripotency. First, a screen was carried out for inhibitors and cytokines that could enable hPSC growth in the presence of GSK3β and MEK/ERK inhibitors. Once a combination of factors was attained, several lines of human ESCs and iPSCs were characterized in this medium. The NHSM medium maintains hPSCs in a pluripotent state, as noted by teratoma assay and the expression of pluripotency factors, and allows derivation of hESC line directly from embryos. hPSCs that are maintained in NHSM show characteristics of naïve pluripotency in their gene expression profile, as micro-array analysis revealed that hPSCs cluster closer to human ICM cells and to mouse naïve cells, in comparison to hESCs that were grown in conventional medium. The chromatin landscape of NHSM-grown hPSCs also show marks of naïve pluripotency such as low level of the inhibitory mark H3K27me3 around the gene body of all genes and specifically in developmental genes. H3K4me3 was also reduced, and the repressive H3K9me3 modification was lower in X chromosome of female lines, suggesting re-activation of the X chromosome. Finally, naïve hESCs that were grown in NHSM were able to integrate into a mouse embryo when injected to the morula, forming an inter-species chimera. Further improvement of the NHSM medium through stabilization of cytoplasmic β-catenin and addition of ACTIVIN A (ENHSM) permitted hPSCs growth without FGF2 and TGFβ that was associated with high expression of naïve related transcripts, even in feeder free conditions.

It is important to notice that the pluripotency state, at least *in vitro*, does not consist of two distinct states but is rather a spectrum of states and drawing the line between naïve and primed can be hard even in mouse cells, where cells in serum+LIF exhibits some primed-related characters but are considered naïve, while cells in IWR1+CHIR are considered primed although they survive with GSK3β inhibition.

It is now widely accepted that conventional hPSCs represent a primed state of pluripotency, but they still express NANOG and REX1, which are not present in mouse

primed EpiSCs[58, 59]. Also, only little is known about the early post implantation human embryo and therefore comparisons to this state is limited. There are several differences between mouse and human ICM, such as expression of Esrrb and Klf2 only in the mouse and expression of ACTIVIN receptors in human[51]. Due to these differences, the naïve states between mouse and human can be different and direct comparison should take these differences into consideration.

Therefore, it is reasonable that during hESCs derivation the cells continue to differentiate. Specific signals can prevent this differentiation and revert the cells back to the ICM-like state. NHSM and ENHSM represent a combination of factors that support hPSC proliferation and maintain a stable state that displays several characteristics o a naïve state.

After the publication of the NHSM medium, several groups published other medium combinations that can support naïve pluripotency in human cells. The work of Chan et al claims that inhibition of BMP results in naïve hESCs that express high levels of naïve-related transcripts, but the medium was only tested on one hESC line[60]. Because of the large genetic variation between hPSCs that originate from different sources, it is important to test any medium on several lines. The Smith group showed that 3i/Lif medium, that consists of addition of PKC inhibitor to a titrated 2i/Lif medium, supported hESCs in a naïve state with near complete removal of DNA methylation[61]. In this work, the cells required exogenous expression of KLF2+NANOG in order to transit to the naïve state and the ability of this condition to support pluripotency and naivety without expression of transgenes remain to be tested. The Jaenisch group published a combination of 6 inhibitors, together with LIF and ACTIVIN A, which supports hESCs with naive characteristics such as gene expression[49]. Since cells in this 6i medium seem to be chromosomally unstable, it is possible that further improvements will be required before this medium could be widely used.

The study of naïve pluripotency is important in order to understand the advantages and limitations of laboratory growth of embryonic stem cells. As stem cells are considered a great promise for regenerative medicine, understanding the molecular state of the cells could provide insights to the *in vitro* differentiation potential of hPSCs. Since naïve

PSCs show greater differentiation potential, either through inter-species chimerism or through *in vitro* differentiation protocols such as PGCLC induction, improvement of the culture medium holds a great promise for future improvement of stem cell associated medical treatments.

In the second part of my work, I describe a differentiation protocol that allows human PSCs differentiation to early primordial germ cell like cells (PGCLCs). Using a fluorescent reporter for NANOS3 expression, I showed a short induction of PGC fate through addition of BMP4 together with EGF and SCF to non-adherent stem cells resulted in ~20% differentiation to PGCLCs. The differentiated cells showed gene expression pattern that is similar to authentic embryo-derived PGCs, and probably represent the early stages of PGCs specification, that are not available for research in humans. Conventional primed hESCs were inefficient in driving PGCLC differentiation, possibly because of non-permissive chromatin state. Integrin β3 (ITGB3) is a surface marker that can be used to isolate the hPGCLC population, even without the use of fluorescent reporter. Knock-out of BLIMP1 and SOX17 indicated an important role for these transcription factors in hPGC specification.

Since PGCs develop at early stage of human embryonic development, the study of early germ cell development in human is very limited. The *in vitro* model presented here can recapitulate the first steps of PGC specification and enables study of hPGCs in a simple and robust manner. Indeed, SOX17 was found to play a major role in human germ cells, while the PGCs of $Sox17^{-/-}$ mice seem to be unaffected, emphasizing differences between the species that could be studied only in human model.

Similar to mouse, the differentiation of human cells to PGCLCs is dependent on the pluripotency state and is more robust when the ESCs are in a naïve state of pluripotency. This result converges the two parts of my work and applies a functional role for naïve pluripotency in human. The fact that naïve cells differentiate better to PGCLCs can also explain the failure of previous works to efficiently differentiate hPGCs *in vitro*.

In addition to the function of hPGCLCs in studying early human germ cell specification, hPGCLCs could be the first step toward the production of human gametes in the lab and

study of human fertility disease. Although technical problems limited our ability to test sperm differentiation in the mouse testis, it is possible that in the future, *in vitro* differentiation of hPGCLCs to late germ cells and fully developed gametes will be possible. Such advancement could provide patients that suffer from fertility diseases with fertile sperm or oocytes for reproduction.

Materials and methods

Cultivation of human naïve pluripotent stem cells

The following conditions, termed NHSM (naive human stem cell medium) conditions were used to isolate, generate, derive and stabilize naive human pluripotent stem cells (iPSCs and ESCs). NHSM includes:

- 475ml knock-out DMEM (Invitrogen), 5g AlbuMAX I (Invitrogen), 50ml KSR (Thermo),1 mM glutamine (Invitrogen), 1% non-essential amino acids (Invitrogen), 0.1mM b-mercaptoethanol (Sigma), Penicillin-Streptomycin (Invitrogen)
- Cytokines: 10µg of recombinant human LIF (Peprotech), 8ng/ml recombinant FGF2 (Peprotech) and 1ng/ml recombinant TGFβ1 (Peprotech).
- Small molecule inhibitors: PD0325901 (1µM, ERK1/2i, Axon Medchem); CHIR99021 (3µM, GSK3βi, Axon Medchem); SP600125 (10µM, JNKi, TOCRIS) and SB203580 (10µM, p38i, TOCRIS). After further optimization we alternatively used SB202190 (5µM, p38i, Axon Medchem) or BIRB796 (2µM,p38i Axon Medchem) for enhanced p38 inhibition.

Rock inhibitor Y-27632 (5µM, Axon Medchem) was added to human cells 4-24 hours before and after splitting. PKC (protein kinase C) inhibitor Go6983 (5mM, PKCi, TOCRIS) can be permanently used in NHSM conditions and result in positive metabolomics effect and reduction in background apoptosis levels in the cell cultures. Naive hESCs/hiPSCs were grown on gelatin (0.2%) + Vitronectin (1µg/ml) -coated plates or on mouse ICR-DR4 embryonic fibroblast (MEF). Media was replaced every 24h. Cells were passaged by single-cell trypsinization (0.05% trypsin EDTA) every 3–4 days.

Enhanced naïve medium (ENHSM) was used in some of the studies:

- 1:1 mix of Neurobasal Medium (Invitrogen 21103-049) and DMEM/F12 (Invitrogen 21331), 1mM glutamine (Invitrogen), 1% non-essential amino acids

(Invitrogen), 0.1mM β-mercaptoethanol (Sigma), Penicillin-Streptomycin (Invitrogen), 5ml KSR, 10ml B27 supplement (Invitrogen 17504-044), 100X Fraction V 7.5% Solution Gibco 15260-037, L-ascorbic acid 2-phosphate (Sigma A8960, 50 µg/ml final concentration)

- N2 components: Insulin (Sigma I-1882, 12.5µg/ml), Apo-transferrin (Sigma T-1147, 100µg/ml), Progesterone (Sigma P8783, 0.02µg/ml), Putrescine (SigmaP5780, 16µg/ml), Sodium selenite (Sigma S5261, add 5µL of 3mM stock solution per 500ml)
- Cytokines: 10µg of recombinant human LIF (Peprotech), Human ACTIVIN A - (Peprotech 120-14E, 20ng/ml)
- Small molecule inhibitors: PD0325901 (1µM, ERK1/2i, Axon Medchem); CHIR99021 (0.1-3µM, GSK3βi, Axon Medchem); SP600125 (5µM, JNKi, TOCRIS), BIRB796 (2µM, p38i Axon Medchem), IWR1 (Tankirasei, TOCRIS – 3532, 5µM), Go6983 (PKCi, TOCRIS 2285, 2µM) , CGP77675 (SRCi , Axon 2097, 1.5µM)

The following already established human ES cells and iPS cell lines were used: human induced pluripotent stem cells C1, C2, FX17 human iPS cell lines and the human embryonic stem cell (human ES cell) lines BGO1 (National Institutes of Health ID code BG01;BresaGen), H1, H9, WIBR3, WIBR3-MBD3mut, WIBR3 OCT4-GFP, WIBR3 ΔPE-OCT4-GFP (gift from Prof. Rudolph Jaenisch lab), H9 NANOG-GFP (WICELL), HUES64 DNMT1-Tet off (Meissner lab). The following hESCs lines were derived from human embryos using NHSM medium: WIS1, WIS2, LIS1, LIS2.

Culture of conventional/primed human ESCs and iPSCs

Human ES cells were maintained on irradiated mouse embryonic fibroblast (MEF) feeder layers or gelatin+vitronectin-coated plates, in human ES cell medium:

- 425ml knock-out DMEM or DMEM-F12 (Invitrogen), 20% KSR (Invitrogen), 1mM glutamine (Invitrogen), 1% nonessential amino acids (Invitrogen), 0.1 mM β-mercaptoethanol (Sigma),

- Cytokines: FGF2 (Peprotech, 8ng/ml)

Recombinant human TGF-β1 (Peprotech, 1ng/ml) was added in feeder-free conditions. Cultures were passaged every 5–7 days manually using Collagenase and scraping or by trypsinization (ROCK inhibitor at 10mM concentration was added 4-24h pre and 24h trypsinization).
For transfection of hiPSC and hESC lines, cells were collected with 0.25% trypsin-EDTA solution (Invitrogen), and resuspended in PBS with 100mg DNA constructs (Gene Pulser Xcell System; Bio-Rad; 250 V, 500mF, 0.4-cm cuvettes). Antibiotics selection was applied 48h after transfection.

Mouse naive and primed stem cell lines and cultivation

Murine naive V6.5 (C57B6/129sJae) pluripotent ESCs were maintained mostly in mESCs culture medium:

- 500ml DMEM (Invitrogen), 1mM glutamine (Invitrogen), 1% nonessential amino acids (Invitrogen), 0.1mM bmercaptoethanol (Sigma), penicillin/streptomycin (Invitrogen)
- Cytokines: human LIF (Peprotech, 0.02ng/ml)
  For serum free conditions, mESCs were transferred to N2B27 chemically defined medium for 2-3 passages before experiment:
- 500ml KO-DMEM, 5ml B27 supplement (Invitrogen), 1mM glutamine (Invitrogen), 1% nonessential amino acids (Invitrogen), 0.1mM bmercaptoethanol (Sigma), penicillin/streptomycin (Invitrogen), 5mg/ml BSA (Sigma)
- N2 components (as in ENHSM medium)
  Naive 2i/LIF conditions for murine PSCs included 0.02ng/ml human LIF (Peprotech), CHIR99021 (CH, 3mM Axon Medchem) and PD0325901 (PD, 1mM TOCRIS).

Primed 129Jae EpiSC line (derived from E6.5 embryos) or C57BL6/129sJae were expanded in N2B27 with 8ng/ml human FGF2 (Peprotech) and 20 ng/ml recombinant human ACTIVIN A (Peprotech) and 1% KSR. Cell lines were routinely checked for

mycoplasma contaminations every month (Lonza, mycoalert kit), and all samples were not contaminated.

Flow Cytometry and immune-fluorescence staining of cells

Cells were dissociated with trypsin-EDTA and fluorescence level was tested using Aria III flow cytometer and sorter. For immune-staining followed by flow cytometry, dissociated cells were washed once with PBS+0.5% BSA, then incubated for 30' with the following antibodies: Pacific blue anti-human HLA A,B,C (W6/32) (Biolegend 311418), h/m SSEA4-Allophycocyanin (R&D FAB1435A), ITGB3-Alexa Fluor 647 (anti-human CD61 Biolegend BLG-336408), TNAP-APC (Biolegend, 327308). After incubation, cells were washed twice with PBS+0.5% BSA.

For immune-fluorescence staining, Cells were grown for 2 days on glass cover slips and then fixed with 4% paraformaldehyde/phosphate buffer for 10 min at room temperature, washed three times with PBS, and permeabilized in PBS+0.1% Triton for 10 min. Cells were blocked with blocking solution (2% normal donkey serum, 0.1% BSA in PBS+0.05% Tween) and incubated with primary antibody diluted in blocking solution for 1-2h in room temperature. Cells were then washed three times with PBS+0.05% Tween, incubated with secondary antibodies for 1 h at room temperature, washed in PBS+0.05% Tween, counterstained with DAPI (1:1000), mounted with Shandon Immu-Mount (Thermo Scientific), and imaged. The antibodies used for staining: mouse anti-TRA-1-60 (Abcam ab16288, 1:500), mouse anti-TRA-1-81 (Abcam ab16289, 1:500), mouse anti-SSEA4 (MC813) (Abcam ab16287, 1:50), rabbit anti-Nanog (Bethyl A300-397A, 1:200), mouse anti-Oct4 (C-10 clone) (Santa Cruz SC5279, 1:200). Secondary antibodies were: RRX-Donkey Anti-Rabbit IgG, RRX-Donkey Anti-Mouse IgG, RRX-Donkey Anti-mouse IgM (Jackson, 1:200).

ES cell derivation from human blastocysts

The use of human pre-implantation embryos for ES cell derivation was performed in compliance with protocols approved by a Weizmann Institute and LIS hospital ESCRO committees, Lis hospital Institutional review committee and Israeli National Ethics Committee (7/04-043) and following the acceptance of a written informed consent. The

couples' participation in the study was voluntary after signing informed consent forms and there was no monetary compensation for their embryo donation. Inner cell masses were isolated mechanically by laser-assisted micromanipulation from spare *in vitro* fertilized embryos, at day 6–7 following fertilization. The intact inner cell mass clumps were placed on a feeder cell layer of irradiation treated DR4 mouse embryonic fibroblasts and cultured in NHSM media. Initial outgrowths of proliferating ES cells were evident by day 6, and were trypsinized into single cells, 6–10 days following inner cell mass plating. The newly established cell lines were further propagated by trypsin and then either frozen or used for further analysis.

Teratoma assay

For teratoma formation and analysis, naive human ESCs and iPSCs were collected by trypsinization before injection. Cells were injected sub-cutaneously into NSG mice (Jackson laboratories). Tumors generally developed within 4–6 weeks and animals were euthanized before tumor size exceeded 1.5 cm in diameter. All animal studies were conducted according to the guideline and following approval by the Weizmann Institute IACUC (approval 00960212-3). Tumor mass extracted and fixed in 4% paraformaldehyde overnight. Slides were prepared from the paraffin embedded fixed tissue, which were next Hematoxylin & Eosin stained and inspected for representation of all three germ layers.

Reverse transcription and quantitative PCR analysis

Total RNA was isolated using TRIZOL (Invitrogen). 1ug RNA was reverse transcribed using a first strand synthesis kit (Invitrogen) and ultimately re-suspended in 50 ul of water. Quantitative PCR analysis was performed in triplicate using 1/10 of the reverse transcription reaction in a Viia7 platform (Applied Biosystems). Bars represent RQ (relative quantity) or Log2(fold change), as indicated. Error bars indicate standard deviation of triplicate measurements for each sample and represent one repeat of 3 biological repeats for each analysis. For expression level in ENHSM medium (Figure 6C) MATLAB software was used to represent the Log2(Fold change) value in color code. RT–PCR primers used herein:

Gene	Forward	Reverse
hSOX2	CACTGCCCCTCTCACACATG	TCCCATTTCCCTCGTTTTTCT
hKIT	CCAACACCGGCAAATACACGTG	CCCATACAAGGAGCGGTCAACAA
hBLIMP1	ACATACATTGTGAACGACCACCCCT	CTGTTGGTGGCATACTTGAAAAGCA
hITGB3	TGACTCCGACTGGACCGGCT	CCCATAGGAGCCCGGCTGGA
hPRDM14	CTACCGAGCCCGAGTGGCCTAC	TAGAGCCATCCCGGGACCGCA
hAP2γ	CGCTCATGTGACTCTCCTGACATCC	TGGGCCGCCAATAGCATGTTCT
hNANOS3	CCCGAAACTCGGCAGGCAAGA	AAGGCTCAGACTTCCCGGCAC
hBRACHYURY	ACCCAGTTCATAGCGGTGAC	CCATTGGGAGTACCCAGGTT
hSOX17	AGCAGAATCCAGACCTGCAC	TTGTAGTTGGGGTGGTCCTG
hACTINB	CCACGAAACTACCTTCAACTCC	GTGATCTCCTTCTGCATCCTGT
hGAPDH	GATGACATCAAGAAGGTGGTGA	GCTGTTGAAGTCAGAGGAGACC

Protein western blotting analysis

Whole-cell protein extracts were isolated from human ES cells. Blots were incubated with the following antibodies in 3% BSA/TBST or PBST: HSP90 (CA1016; 1:5,000; Calbiochem), Histone H3 tri methyl K27 (Abcam ab6002; 1:200), Histone H3 (Abcam ab1791; 1:1000), GAPDH (Epitomics, 2251-1; 1:1000), DNMT1 (Cell signaling, #5032; 1:1000). Secondary antibodies were horseradish peroxidase-linked goat anti-mouse, goat anti rabbit and rabbit anti-goat (1:10,000; Jackson), Blots were developed using ECL (Thermo).

Microarrays analysis of gene expression

Total RNA was isolated from indicated primed and naive (in NHSM with 5mM SB203580). The concentration of RNA was quantified and subjected to quality control on Agilent Bioanalyzer. cDNA was fragmented, labeled and hybridized to Affymetrix Human Gene 1.0 ST GeneChip (Affymetrix). Labeling and hybridization was done in Biological services unit, Weizmann institute. Transcripts levels were processed from Affymetrix CEL files and CDF file (version V1.r3, which maps probes into 33,252 probe sets) using Robust Multi-Array Average method. All analyses were done using Matlab (Version R2012b) and its Bioinformatics toolbox. Data was further filtered to include probes that have at least one call higher than 64 (526), and mapped to unique Entrez IDs using Affymetrix annotations (NetAffx Annotation Files version 33.1) and NCBI sites. Probes targeting the same gene are represented by their median, resulting in

13,894 genes. Microarray data are available at the series accession number GSE46872. The samples were hierarchically clustered using average linkage and either Spearman or Pearson correlation as a distance matrix, with similar results. To compare our samples to hICM cells, human pre-implantation data from [47] were included. This data, from Affymetrix HuGene 1.0 st microarrays described before, was passed through the same processing, resulting in 16,953 genes before and 12,062 genes after intersection with the previous gene list. Batch effect was corrected by normalization of the new data to the mean values of its ES cell samples, making it consistent with the normalization of our data.

Cross-species gene expression analysis was conducted on human arrays described above and previously described mouse ES cell and EpiSC gene expression data sets on an Agilent 4344 k array platform (GSE15603) containing 45,018 probes. Mouse data were processed as described above, resulting in 17,885 unique genes[20]. Human–mouse orthology was downloaded from MGI (http://www. informatics.jax.org, on April 2013) containing 17,772 pairs of orthologous genes. Of these, 9,803 were mapped to our expression data. The expression values from mouse and human were transformed separately into relative abundance values: For each gene, the relative abundance value is the expression value divided by the mean of expression values within the same gene across samples in the same species. The resulting expression matrix was subjected to hierarchical clustering (Spearman correlation, average linkage).

Chromatin immune-precipitation and sequencing library preparation

Chromatinimmunoprecipitation followed by deep sequencing (ChIP-seq) was measured for H3K4me3, H3K27me3, H3K4me1, H3K27ac and H3K9me3 in mouse and human pluripotent cells (ES cells, EpiSCs and/or iPS cells) expanded in fetal bovine free and feeder free naive or primed/conventional growth conditions (FBS and feeder free expansion, on gelatin/vitronectin-coated plates). Approximately $40X10^6$ cells were cross-linked in formaldehyde (1% final concentration, 10min at room temperature), and then quenched with glycine (5min at room temperature). Fixed cells were lysed in 50mM HEPES KOH pH 7.5, 140mM NaCl, 1mM EDTA, 10% glycerol, 0.5% NP-40 alternative, 0.25% Triton supplemented with protease inhibitor at 4°C (Roche), centrifuged at 950g for 10min and resuspended in 0.2% SDS, 10mM EDTA, 140mM

NaCl and 10mM Tris-HCL. Cells were then fragmented with a Branson Sonifier (model S-450D) to size ranges between 200 and 800bp, and precipitated by centrifugation. 10mg of each antibody was pre-bound by incubating with Protein-G Dynabeads (Invitrogen100-07D) in blocking buffer (PBS supplemented with 0.5% TWEEN and 0.5% BSA) for 2h at room temperature. Beads were added to the chromatin lysate, and then incubated overnight. Samples were washed 5 times with RIPA buffer (Tris HCl pH 8.0 10mM, EDTA 1mM, NaCl 140mM, Triton 1%, SDS 0.1%, DOC0.1%) twice with RIPA buffer supplemented with 500mM NaCl, twice with LiCl buffer (10mM TE, 250mM LiCl, 0.5% NP-40, 0.5% DOC), once with TE (10mM Tris-HCl pH 8.0, 1mM EDTA), and then eluted in 0.5% SDS, 300mM NaCl, 5mM EDTA, 10mM Tris Hcl pH 8.0 at 65°C. Eluate was treated sequentially with RNase A (Roche) for 30 min and proteinase K (NEB) for 2h, and then incubated at 65°C for 4h. DNA was purified with Agencourt AMPure XP system (Beckman Coulter Genomics, A63881). Libraries of cross reversed ChIP DNA samples were prepared according to a modified version of the Illumina Genomic DNA protocol. Briefly, ChIP DNA was ligated to Illumina adaptors and subjected to 14 cycles of PCR amplification. Amplified products between 200 and 800bp were purified by using magnetic bead size selection. 5pmole of DNA library was then applied to each lane of the flow cell and sequenced on Illumina Hiseq2000 sequencer according to standard Illumina protocols, at the INCPM unit, Weizmann institute. The following antibodies were used for chromatin-IP experiments: control IgG (ChIP grade, ab46540, Abcam), anti-H3K4me3 (ab8580, Abcam), anti-H3K27me3 (07-449, Millipore), anti-H3K4me1 (ab8895, Abcam), anti-H3K27acetyl (ab4729, Abcam), anti-H3K9me3 (ab8898, Abcam).

Chromatin IP sequencing data analysis

The chromatin markers H3K27me3, H3K4me3, H3K4me1, H3K27Ac and H3K9me3 were measured in the different human pluripotent cell lines: C1, WIBR3, LIS2 (naive and primed), BGO1 and WIBR3-MBD3mut (naive). WIS1 naive and primed hES cells were profiled for the H3K9me3 mark only. In addition, H3K27me3, H3K4me3, H3K4me1 and H3K27Ac were measured in mouse V6.5 naive and 129 primed EpiSCs. H3K27me3 and H3K4me3 measurements in naive mouse ES were previously published[62]. Each sample was accompanied by control sequencing experiment of

whole-cell extract input. We used bowtie software version 1.0.0 to align human reads to human reference genome hg19 (UCSC, February 2009) and mouse reads to mouse mm9 reference genome (UCSC, July 2007). To rule out sequencing depth bias, the aligned sequences were down-sampled such that all samples had the same number of aligned reads. Human samples of the marks H3K4me3, H3K27me3, H3K4me1 and H3K27Ac, as well as whole-cell extract, were down-sampled to include 3,750,000 aligned reads. H3K9me3 samples were down-sampled to 5,900,000 aligned reads. Mouse samples of the marks H3K4me3, H3K27me3, H3K4me1 and H3K27Ac, were down-sampled to 3,420,000 aligned reads. Chromatin profiles were calculated over all RefSeq genes (n=543,463), and over developmental genes in the following way: (1)read densities were calculated between 3 kb upstream to TSS, and 3 kb downstream to TES. (2) Each gene was divided to 100 bins of identical size, and the sum of reads in each bin was calculated. (3) Average profile was calculated over all genes, where the gene body, which is of changing size, was represented by 100 quantiles. Genes of size less than 1 kb were filtered out. (4) Profiles of human samples represent mean and s.d. (error bars) of primed and naive samples. Lastly, developmental genes were selected if they have a GO annotation that is related to development or differentiation. Using this criterion, we had 5,922 RefSeq human developmental genes, and 420 RefSeq mouse developmental genes. Concrete examples of genes were processed and visualized using IGV software version 2.0. To measure the distribution of H3K9me3 accumulation in chromosome X genes, RPKM (reads per kilobase per 5.9 million reads) was calculated for each gene (between 1kb upstream to TSS and TES).P-values between distributions were calculated with one-tail paired-samplet-test.

Enhancers were detected following the guidelines set in [24]. Shortly, enhancers of type one are genomic intervals that contain H3K4me1 and H3K27Ac marks, do not contain H3K4me3 or H3K27me3 marks and are at least 500bp away from any TSS. Enhancers of type two are genomic intervals that contain H3K4me1 and H3K27me3 marks, do not contain H3K27Ac mark, and are at least 500bp away from any TSS. To find those enhancers we first identified enriched intervals of the marks above using MACS version 1.4.1. We used sequencing of whole cell extract as control in order to define a background model. Duplicate reads aligned to the exact same location are excluded by

MACS default configuration. Enriched intervals (peaks) that overlap by at least 1bp were considered as overlapping, and their union was defined as the enhancer interval, unless at least 10% of it overlapped with any of the excluded marks (for example, H3K27me3 in the case of type one enhancers) or with TSS. To calculate the number of enhancers in human samples we only considered enhancers that are common to 2 primed cell lines C1 and WIBR3, or common to 4 naive cells lines: C1, WIBR3, WIBR3-MBD3mut and BGO1. The expression level of genes associated with enhancers was calculated by mapping each enhancer to the nearest gene (as long as it is at most 100 kb away), converting it to Entrez Gene and taking its expression value from the raw gene expression data set described above.

PGCLC induction

For pre-induction, naïve (NHSM) hESCs were dissociated with TrypLE and $3x10^6$ cells/well were plated on fibronectin coated 6-well plates in N2B27 medium with TGF-β1 (Peprotech, 1ng/ml) or 20ng/ml ACTIVIN A (Peprotech) and 10μM ROCK inhibitor. Medium was changed on day 1. After 2 days of pre-induction, the cells are dissociated with TrypLE and plated to ultra-low cell attachment U-bottom 96-well plates (Corning, 7007), at a density of 2000-4000 cells/well, in 150μl PGCLC medium. PGCLC medium is composed of Glasgow's MEM (GMEM, GIBCO), 15% KSR, 0.1mM nonessential amino acids, 0.1mM 2-mercaptoethanol, 100U/ml Penicillin, 0.1mg/ml Streptomycin, 2mM L-Glutamine, 1mM Sodium pyruvate and the following cytokines: 500ng/ml BMP4 (R&D Systems or Peprotech), 0.02ng/ml human LIF (Peprotech), 100ng/ml SCF (Peprotech), 50ng/ml EGF (Peprotech) and 10μM ROCK inhibitor.

When BMP2 was used, the BMP4 in the hPGC medium was replaced with 500 ng/ml recombinant hBMP2 (Peprotech).

Genomic Modifications in hESCs

In order to introduce 2A-mCherry sequence immediately downstream and in frame with the coding sequence of NANOS3, a donor construct was produced as depicted in Figure 7C. Homology arms were amplified using the primers: 50 homology forward with Not I restriction enzyme site: at gcggccgc gtgcctgtagtcccagctacttgggag, 50 homology reverse

with SalI restriction enzyme site: acc gtcgactctagagtcgcgaccgctgta ggtggacatggagggagagcagg, 30 homology forward with HpaI restriction enzyme site: at gttaac gaggctgcctacacctgggca and 30 homology reverse with HpaI restriction enzyme site: cg gttaac aagatctggaggtggaggaggcag. TALEN expressing constructs were generated using GoldenGate TALEN kit 2.0 (Addgene cat#1000000024). TALEN's repeats sequences targeting human NANOS3 stop codon were: forward HD HD NG NH HD NG HD NG HD HD HD NG HD HD NI NG, and reverse NG NH HD HD HD NI HD HD NG NH NG NI NH NH HD NI. WIS1 and LIS1 male hESC and WIBR3 female hESC lines were electroporated with a pair of TALEN coding plasmids and donor construct. After selection with G418 (150mg/ml) and ganciclovir (2mM), genomic DNA was extracted from 96 clones for each cell line. Targeting efficiency was about 50% 60% in all the experiments as revealed by PCR and Southern Blot analysis. Southern blot with internal anti-mCherry probe did not show non-specific insertions in 80% of correctly targeted clones. In order to delete PGK-Neo cassette, correctly targeted clones were transfected with flippase expressing plasmid and subcloned. Neo cassette excision was confirmed by PCR. Karyotyping analysis of correctly targeted clones was performed by G-Banding on ASI platform, and confirmed normal karyotype in all clones used. In order to knock out BLIMP1 and SOX17 genes, oligos encoding gRNAs targeting these genes were inserted into px330 vector. Unique gRNA sequences were chosen with the help of Zhang Lab websitehttp://www.genome-engineering.org/ crispr/(Figures S5A andS6A). 100mg of resulting construct and 10mg of GFP expressing vector were electroporated into WIS1-NANOS3-mCherry hESCs. 3 or 4 days later, GFP expressing cells were sorted by FACS and seeded at low density. 9 days after seeding 88 colonies were picked for each experiment and genomic DNA was extracted. DNA was analyzed by High Resolution Melt assay (HRM) using MeltDoctor reagent (Life Technologies) and the clones that showed reduced Tm for both alleles compared to wild-type controls were expanded. In these selected clones targeted locus was amplified and sequenced. Primers for HRM are; BLIMP1: CGATGACTTTAGAAGACGTGGAGCC; CGTAGGCCAGGGAAGCTTTCAA, SOX17: GCCAGTGACGACCAGAGCCAG; TCACCTT CATGTCCCCGATGG. Gene targeting plasmids are available through Addgene.

LIS1 CCAGS-GFP line was prepared by electroporation of plasmid with GFP under the constitutive CAAGS promoter, followed by sort for strong GFP expression.

Collection of samples and analysis for RNA–seq

Day 4 and day 6 (after priming) hPGCLC from WIS1 cell line were sorted for the expression of NANOS3-mCherry reporter. Other samples were collected directly from plates into TRIZOL. Total RNA was extracted using DirectZol RNA mini-prep (Zymo research, R2052). RNA integrity was evaluated on Tapestation (Agilent). Libraries were prepared by the INCPM unit in the Weizmann Institute of Science according to Illumina's instructions accompanying the TruSeq RNA Sample Preparation Kit v2 (RS-122-2001). Sequencing was carried out on Illumina HiSeq2500 according to the manufacturer's instructions, using 10pM template per sample for cluster generation, and sequencing kit V2 (Illumina), resulting in ~40 million paired-end reads per sample. Data was subjected to analyses with Tophat software and mapped to the human reference genome (GRCh37/hg19). Reads were counted with HTSeq-count and Matlab was used for all further analysis. Specific genes were shown using IGV viewer.

Microinjection to mouse testis

Mouse (BVSC cell line) or human ESCs were induced to differentiate to PGCLC for 4-6 days. The injected cells were dissociated using TripLE and then either FACS-Sorted for PGC markers (Blimp1-Venus for mouse), separated by MACS column, or taken as a whole. MACS separation was done by addition of biotinilated ITGB3 untibody to the cells (Biolegend), followed by incubation with Miltenyi anti-biotin magnetic microbeads (130-090-485), and separation with Miltenyi MS column on magnet.

Cells were centrifuged and resuspended in PBS and Trypan blue was added to the cells at a 1:10 dilution in order to identify injected testis.

Microinjection was done Similarly to a previously published protocol[57]. W/W-V mice (Jackson Laboratory) pups, at age of 9-12 days were used, and the mutant mice were identified by the white coat color. The pups were anesthetized, and the testis was exteriorized through a midline abdominal cut and aligned to injection glass pipette needle. The pipette was filled with germ cell solution by using a thin needle, and the

pipette was fastened in a pipette holder attached to a pressure injector. Injection was directed to the testis rete, between the large veins. After injection the testis was returned to the abdominal cavity and the cut was glued. 8-16 weeks after injection, the testis were removed for analysis, fixed over night in Bouin solution, washed with increasing concentrations of ethanol and paraffin embedded. Next, the testes were Hematoxylin & Eosin stained and inspected for morphology and spermatogenesis. For Mvh staining, testis were fixed with 4% PFA, paraffin embedded and immune-stained with Mvh antibody (Abcam, ab13840) and counter-stained with DAPI.

References

1. Martin, G.R., *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**(12): p. 7634-8.
2. Smith, A.G., et al., *Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides.* Nature, 1988. **336**(6200): p. 688-90.
3. Thomson, J.A., et al., *Embryonic stem cell lines derived from human blastocysts.* Science, 1998. **282**(5391): p. 1145-7.
4. Gurdon, J.B., *The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles.* J Embryol Exp Morphol, 1962. **10**: p. 622-40.
5. Takahashi, K. and S. Yamanaka, *Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors.* Cell, 2006. **126**(4): p. 663-76.
6. Matsui, Y., K. Zsebo, and B.L. Hogan, *Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture.* Cell, 1992. **70**(5): p. 841-7.
7. Labosky, P.A., D.P. Barlow, and B.L. Hogan, *Mouse embryonic germ (EG) cell lines: transmission through the germline and differences in the methylation imprint of insulin-like growth factor 2 receptor (Igf2r) gene compared with embryonic stem (ES) cell lines.* Development, 1994. **120**(11): p. 3197-204.
8. Kanatsu-Shinohara, M., et al., *Generation of pluripotent stem cells from neonatal mouse testis.* Cell, 2004. **119**(7): p. 1001-12.
9. Hou, P., et al., *Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds.* Science. **341**(6146): p. 651-4.
10. Brons, I.G., et al., *Derivation of pluripotent epiblast stem cells from mammalian embryos.* Nature, 2007. **448**(7150): p. 191-5.
11. Tesar, P.J., et al., *New cell lines from mouse epiblast share defining features with human embryonic stem cells.* Nature, 2007. **448**(7150): p. 196-9.
12. Ying, Q.L., et al., *The ground state of embryonic stem cell self-renewal.* Nature, 2008. **453**(7194): p. 519-23.
13. Shimizu, T., et al., *Dual inhibition of Src and GSK3 maintains mouse embryonic stem cells, whose differentiation is mechanically regulated by Src signaling.* Stem Cells. **30**(7): p. 1394-404.
14. Dutta, D., et al., *Self-renewal versus lineage commitment of embryonic stem cells: protein kinase C signaling shifts the balance.* Stem Cells. **29**(4): p. 618-28.
15. Acampora, D., L.G. Di Giovannantonio, and A. Simeone, *Otx2 is an intrinsic determinant of the embryonic stem cell state and is required for transition to a stable epiblast stem cell condition.* Development. **140**(1): p. 43-55.
16. Hanna, J.H., K. Saha, and R. Jaenisch, *Pluripotency and cellular reprogramming: facts, hypotheses, unresolved issues.* Cell. **143**(4): p. 508-25.
17. Bao, S., et al., *Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells.* Nature, 2009. **461**(7268): p. 1292-5.
18. Kim, H., et al., *Modulation of beta-catenin function maintains mouse epiblast stem cell and human embryonic stem cell self-renewal.* Nat Commun. **4**: p. 2403.

19. Festuccia, N., et al., *The role of pluripotency gene regulatory network components in mediating transitions between pluripotent cell states.* Curr Opin Genet Dev. **23**(5): p. 504-11.
20. Hanna, J., et al., *Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs.* Proc Natl Acad Sci U S A. **107**(20): p. 9222-7.
21. Bernemann, C., et al., *Distinct developmental ground states of epiblast stem cell lines determine different pluripotency features.* Stem Cells. **29**(10): p. 1496-503.
22. Marks, H., et al., *The transcriptional and epigenomic foundations of ground state pluripotency.* Cell. **149**(3): p. 590-604.
23. Creyghton, M.P., et al., *Histone H3K27ac separates active from poised enhancers and predicts developmental state.* Proc Natl Acad Sci U S A. **107**(50): p. 21931-6.
24. Rada-Iglesias, A., et al., *A unique chromatin signature uncovers early developmental enhancers in humans.* Nature. **470**(7333): p. 279-83.
25. Yeom, Y.I., et al., *Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells.* Development, 1996. **122**(3): p. 881-94.
26. Hanna, J., et al., *Metastable pluripotent states in NOD-mouse-derived ESCs.* Cell Stem Cell, 2009. **4**(6): p. 513-24.
27. Buehr, M., et al., *Capture of authentic embryonic stem cells from rat blastocysts.* Cell, 2008. **135**(7): p. 1287-98.
28. Nichols, J. and A. Smith, *Naive and primed pluripotent states.* Cell Stem Cell, 2009. **4**(6): p. 487-92.
29. Li, W. and S. Ding, *Human pluripotent stem cells: decoding the naive state.* Sci Transl Med. **3**(76): p. 76ps10.
30. Saitou, M. and M. Yamaji, *Germ cell specification in mice: signaling, transcription regulation, and epigenetic consequences.* Reproduction. **139**(6): p. 931-42.
31. Magnusdottir, E., et al., *A tripartite transcription factor network regulates primordial germ cell specification in mice.* Nat Cell Biol. **15**(8): p. 905-15.
32. Saitou, M., S. Kagiwada, and K. Kurimoto, *Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells.* Development. **139**(1): p. 15-31.
33. Nagamatsu, G., et al., *A germ cell-specific gene, Prmt5, works in somatic cell reprogramming.* J Biol Chem. **286**(12): p. 10641-8.
34. Irie, N., W.W. Tang, and M. Azim Surani, *Germ cell specification and pluripotency in mammals: a perspective from early embryogenesis.* Reprod Med Biol. **13**(4): p. 203-215.
35. Tang, W.W., et al., *A Unique Gene Regulatory Network Resets the Human Germline Epigenome for Development.* Cell. **161**(6): p. 1453-67.
36. Perrett, R.M., et al., *The early human germ cell lineage does not express SOX2 during in vivo development or upon in vitro culture.* Biol Reprod, 2008. **78**(5): p. 852-8.
37. Hayashi, K., et al., *Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells.* Cell. **146**(4): p. 519-32.
38. Hayashi, K., et al., *Offspring from oocytes derived from in vitro primordial germ cell-like cells in mice.* Science. **338**(6109): p. 971-5.
39. Kurimoto, K., et al., *Quantitative Dynamics of Chromatin Remodeling during Germ Cell Specification from Mouse Embryonic Stem Cells.* Cell Stem Cell. **16**(5): p. 517-32.
40. Gkountela, S., et al., *The ontogeny of cKIT+ human primordial germ cells proves to be a resource for human germ line reprogramming, imprint erasure and in vitro differentiation.* Nat Cell Biol. **15**(1): p. 113-22.

41. Lin, I.Y., et al., *Suppression of the SOX2 neural effector gene by PRDM1 promotes human germ cell fate in embryonic stem cells.* Stem Cell Reports. **2**(2): p. 189-204.
42. Chuang, C.Y., et al., *Meiotic competent human germ cell-like cells derived from human embryonic stem cells induced by BMP4/WNT3A signaling and OCT4/EpCAM (epithelial cell adhesion molecule) selection.* J Biol Chem. **287**(18): p. 14389-401.
43. Kee, K., et al., *Human DAZL, DAZ and BOULE genes modulate primordial germ-cell and haploid gamete formation.* Nature, 2009. **462**(7270): p. 222-5.
44. Gafni, O., et al., *Derivation of novel human ground state naive pluripotent stem cells.* Nature. **504**(7479): p. 282-6.
45. Qi, X., et al., *BMP4 supports self-renewal of embryonic stem cells by inhibiting mitogen-activated protein kinase pathways.* Proc Natl Acad Sci U S A, 2004. **101**(16): p. 6027-32.
46. Greber, B., et al., *Conserved and divergent roles of FGF signaling in mouse epiblast stem cells and human embryonic stem cells.* Cell Stem Cell. **6**(3): p. 215-26.
47. Vassena, R., et al., *Waves of early transcriptional activation and pluripotency program initiation during human preimplantation development.* Development. **138**(17): p. 3699-709.
48. Kobayashi, T., et al., *Generation of rat pancreas in mouse by interspecific blastocyst injection of pluripotent stem cells.* Cell. **142**(5): p. 787-99.
49. Theunissen, T.W., et al., *Systematic identification of culture conditions for induction and maintenance of naive human pluripotency.* Cell Stem Cell. **15**(4): p. 471-87.
50. Liao, J., et al., *Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells.* Nat Genet. **47**(5): p. 469-78.
51. Blakeley, P., et al., *Defining the three cell lineages of the human blastocyst by single-cell RNA-seq.* Development. **142**(20): p. 3613.
52. Irie, N., et al., *SOX17 is a critical specifier of human primordial germ cell fate.* Cell, 2015. **160**(1-2): p. 253-68.
53. Julaton, V.T. and R.A. Reijo Pera, *NANOS3 function in human germ cell development.* Hum Mol Genet. **20**(11): p. 2238-50.
54. Sugawa, F., et al., *Human primordial germ cell commitment in vitro associates with a unique PRDM14 expression profile.* EMBO J. **34**(8): p. 1009-24.
55. Hara, K., et al., *Evidence for crucial role of hindgut expansion in directing proper migration of primordial germ cells in mouse early embryogenesis.* Dev Biol, 2009. **330**(2): p. 427-39.
56. Kurimoto, K., et al., *Complex genome-wide transcription dynamics orchestrated by Blimp1 for the specification of the germ cell lineage in mice.* Genes Dev, 2008. **22**(12): p. 1617-35.
57. Ogawa, T., et al., *Transplantation of testis germinal cells into mouse seminiferous tubules.* Int J Dev Biol, 1997. **41**(1): p. 111-22.
58. Manor, Y.S., R. Massarwa, and J.H. Hanna, *Establishing the human naive pluripotent state.* Curr Opin Genet Dev. **34**: p. 35-45.
59. De Los Angeles, A., et al., *Accessing naive human pluripotency.* Curr Opin Genet Dev. **22**(3): p. 272-82.
60. Chan, Y.S., et al., *Induction of a human pluripotent state with distinct regulatory circuitry that resembles preimplantation epiblast.* Cell Stem Cell. **13**(6): p. 663-75.
61. Takashima, Y., et al., *Resetting transcription factor control circuitry toward ground-state pluripotency in human.* Cell. **158**(6): p. 1254-69.
62. Mansour, A.A., et al., *The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming.* Nature. **488**(7411): p. 409-13.

List of publications during the course of the PhD

1. Maza I, Caspi I, Zviran A, Chomsky E, Rais Y, Viukov S, Geula S, Buenrostro JD, **Weinberger L**, Krupalnik V, Hanna S, Zerbib M, Dutton JR, Greenleaf WJ, Massarwa R, Novershtern N, Hanna JH. *Transient acquisition of pluripotency during somatic cell transdifferentiation with iPSC reprogramming factors*. Nat Biotechnol. 2015 Jun 22. doi: 10.1038/nbt.3270.
2. Irie N*, **Weinberger L***, Tang WW*, Kobayashi T, Viukov S, Manor YS, Dietmann S, Hanna JH, Surani MA. *SOX17 is a critical specifier of human primordial germ cell fate*. Cell. 2015 Jan 15;160(1-2):253-68. doi: 10.1016/j.cell.2014.12.013. Epub 2014 Dec 24
3. Gafni O*, **Weinberger L***, Mansour AA*, Manor YS*, Chomsky E*, Ben-Yosef D, Kalma Y, Viukov S, Maza I, Zviran A, Rais Y, Shipony Z, Mukamel Z, Krupalnik V, Zerbib M, Geula S, Caspi I, Schneir D, Shwartz T, Gilad S, Amann-Zalcenstein D, Benjamin S, Amit I, Tanay A, Massarwa R, Novershtern N, Hanna JH. *Derivation of novel human ground state naive pluripotent stem cells*. Nature. 2013 Dec 12;504(7479):282-6.
4. Rais Y*, Zviran A*, Geula S*, Gafni O, Chomsky E, Viukov S, Mansour AA, Caspi I, Krupalnik V, Zerbib M, Maza I, Mor N, Baran D, **Weinberger L**, Jaitin DA, Lara-Astiaso D, Blecher-Gonen R, Shipony Z, Mukamel Z, Hagai T, Gilad S, Amann-Zalcenstein D, Tanay A, Amit I, Novershtern N, Hanna JH. *Deterministic direct reprogramming of somatic cells to pluripotency*. Nature. 2013 Oct 3;502(7469):65-70. doi: 10.1038/nature12587. Epub 2013 Sep 18.
5. Mansour AA*, Gafni O*, **Weinberger L**, Zviran A, Ayyash M, Rais Y, Krupalnik V, Zerbib M, Amann-Zalcenstein D, Maza I, Geula S, Viukov S, Holtzman L, Pribluda A, Canaani E, Horn-Saban S, Amit I, Novershtern N, Hanna JH. The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming. Nature. 2012 Aug 16;488(7411):409-13. doi: 10.1038/nature11272

*Equal first authors

Student Declaration

I hereby declare that the thesis presented summarizes my independent research work under the supervision of Dr. Jacob H. Hanna at the Department of Molecular Genetics, the Weismann Institute of Science. Tissue culture work, including optimization of human naïve mediums (NHSM and ENHSM) was done by me together with Dr Jacob Hanna and Yoav Rechavi. Analysis of gene expression data was done by Yair Manor (both micro-array and RNA-seq data). Analysis of ChIP-Seq data was done by Dr. Noa Noverstern. Injection of human naïve cells to mouse morula was done by Ohad Gafni and Mirie Zerbib. Processing and imaging of the resulting chimeric embryos were done by Dr. Rada Massarwa. Cloning and construction of reporter lines and mutations were done with Sergey Viukov. Cell injections to mouse testis were done me together with Mirie Zerbib. Derivation of new hESCs lines was done in LIS maternity hospital in Dr. Dalit Ben Yosef lab. The second part of the work (hPGCs project) was done in collaboration with Prof. Azim Surani lab in Cambridge, UK and specifically with Dr. Naoko Irie.

Acknowledgements

To my mentor, Jacob Hanna who supported me through the hard times, thanks for your guidance. To Yair Manor, Sergey Viyokov, Mirie Zerbib, Ohad Gafni, Rada Massarwa and Noa Noverstern, this work would not have been done without you. To Naoko Irie and Azim Surani lab for pushing the germ cell project forward. To all Hanna lab members, for helping, assisting, teaching and inspiring me during the last 4.5 years. I also want to thank my family, Eyal and my children Tom and Shachar for understanding and waiting for me all the long weekends and evenings I spent in the lab.