Title
The Regulation of X Chromosome Inactivation in Human Development and Pluripotent Stem Cells

Author
Patel, Sanjeet

Publication Date
2014

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

The Regulation of X Chromosome Inactivation in Human Development and Pluripotent Stem Cells

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Molecular Biology

by

Sanjeet Patel

2014

ABSTRACT OF DISSERTATION

The regulation of X chromosome inactivation in human development and pluripotent stem cells

by

Sanjeet Patel

Doctor of Philosophy in Molecular Biology

University of California, Los Angeles 2014

Professor Kathrin Plath, Chair

The following thesis addresses how the epigenetic state of the female X chromosome reflects the developmental stage captured by human embryonic stem cells (hESCs). This topic has been controversial, as many epigenetic states of the X chromosome have been proposed to exist in female human pluripotent stem cells; however, none of the described states captures the state of the X chromosome in the human embryo. To investigate the basis for this disconnect between human pluripotent stem cells and human development, we used single cell resolution techniques in conjunction with mapping of DNA methylation and global transcription to examine the epigenome and transcriptome of pre-implantation embryos, cells during derivation, and in established hESCs. We demonstrate that X chromosome inactivation (XCI) initiates early during derivation of female hESCs, and unlike mouse development, progresses from a naïve state where *XIST* is expressed on two active chromosomes to a state where *XIST* is expressed on one inactive X chromosome. Therefore, the majority of

established female hESCs display a pattern resembling the mouse post-implantation embryo, with one X chromosome silenced, which is associated with the spreading and coating of the non-coding RNA (ncRNA) *XIST*. However, this pattern erodes during progressive culture of female hESCs, with Xist RNA loss and subsequent partial activation of genes on the Xi, to an extent that is unpredictable. We also find that de novo XCI can be stochastically disrupted during hESC derivation, leading to a state where *XIST* is permanently silenced thus that XCI cannot occur *in vitro*. Accordingly some established female hESC lines carry two active X chromosomes in the self-renewing state, but, unlike the prevailing model suggests, do not undergo X chromosome inactivation (XCI) nor undergo dosage compensation upon induction of differentiation. Finally, we find that no existing method is capable of reproducing the epigenetic environment of the pre-implantation embryo in established human pluripotent stem cell lines, and conclude that a female human pluripotent stem cell that is competent to undergo XCI has not been attained. My work places the seemingly confusing findings regarding human pluripotency and the X chromosome into context and also highlights the need for a revision of culture conditions to stabilize the naïve, pre-implantation state of pluripotency in human cells.

The dissertation of Sanjeet Patel is approved.

Donald Kohn

Amander Clark

William Lowry

S. Lawrence Zipursky

Kathrin Plath, Committee Chair

University of California, Los Angeles

2014

DEDICATIONS

To Kathrin, for being a fantastic mentor and supporting all these scientific endeavors these past four years, no matter how high risk. She has not only provided with the means and guidance in order to complete the work within this thesis, but has created an outstanding training environment for which I was able to learn so much.

To my friends and family for always being supportive and understanding.

To my lab mates, these past four years have been memorable.

To my committee for advice and inclusion in collaborative opportunities

To the Department of Surgery for providing me the time and funding to pursue this goal.

TABLE OF CONTENTS

Figures and Tables……...vii-viii

Acknowledgements……..…...ix

Vita……...x-xi

Chapter 1:

Introduction…….……..1-8

References……..…..…..…9-11

Chapter 2:

Pluripotency and the transcriptional inactivation of the female mammalian X chromosome…….……..12-32

References……...…33-38

Chapter 3:

X Chromosome Inactivation in female human embryos and embryonic stem cells……..…….…...39-97

References……....98-101

Chapter 4:

Current states of naïve human pluripotency diverges significantly from ground state mouse pluripotency……102-144

References……..145-149

Chapter 5:

Conclusions……150-153

FIGURES AND TABLES

Chapter 2

Figure 2-1. Mouse and human XCI in development and reprogramming
Figure 2-2. Xist activators and repressors regulate initiation of XCI in mESCs

Chapter 3

Figure 3-1A. Human pluripotent stem cells display varied X inactivation patterns
Figure 3-1B. RNA FISH allows for single cell resolution of X inactivation patterns
Figure 3-2A. X inactivation patterns established in pluripotency are maintained in differentiation.
Figure 3-2B. The pattern of XCI is maintained in differentiation through retinoic acid or directed into germ lineage specific cell types.
Figure 3-2C. Differentiation propensities do not correlate with the degree of erosion in female human pluripotent stem cells.
Figure 3-2D. Genomewide CNV analysis does not show abnormal karyotypes in cardiomyocyte differentiated cells
Figure 3-3A. Genomic approaches to assaying for X inactivation patterns reveal hotspots of erosion in the XaXeXist- lines as well as a loss of dosage compensation for both XaXe and XaXa lines.
Figure 3-3B. Expression from the X chromosome using cumulative distribution frequencies demonstrate that no published pre-XCI line has the pattern of an XaXa human ES line.
Figure 3-4. Tracing the progression of Xist loss during early passage of human ES cells demonstrated that Xist is loss stochastically over time and that XaXa derivation cultures are a mix of XaXa and XaXiXist+ cells
Figure 3-5A. The human pre-implantation blastocyst is XaXaXist++ in female embryos and XaXist+ in male embryos however early in the course of derivation of an ES line, X-inactivation occurs and is subject to aborted products producing XaXaXist- cells.
Figure 3-5B. Human ES derivation results in stochastic kinetics of XCI during cell fate changes.
Movie 1 Cardiomyocyte like cell lines are responsive to beta-adrenergic stimulation
Movie 2 Female pre-implantation Blastocysts have uniform patterns of XCI

Chapter 4

Figure 4-1. The state of XCI in female human embryonic stem cells do not respond to alterations in oxygen tension nor antioxidant treatment.
Figure 4-2. Attempts at converting pre-existing primed cells to XaXaXist- cells based on the published literature fail to reactivate the Xi.
Figure 4-3. Both somatic cell reprogramming and reprogramming from primed pluripotency is not possible using typical combinations of media, small molecules, matrix, and growth factors described to date.

Figure 4-4. Optimal conditions of high insulin, 2i, human LIF and Vitamin C do not result in NANOG positive cells but rather enrich for neurospheres.
Figure 4-5. Naïve pluripotent stem cells from the Hanna lab fail to show evidence of an X inactivation pattern similar to human blastocysts
Figure 4-6. The transcriptome of human naïve cells lines demonstrate that they do not acquire the gene expression program of pre-implantation blastocysts.

ACKNOWLEDGEMENTS

Chapter 2 is a version of a manuscript published and authored by Sanjeet Patel, Alissa Minkovsky and Kathrin Plath with permission for reprint through license number 3398310351342.

Chapter 3 is a version of a manuscript in preparation for publication authored by Sanjeet Patel, Rachel Kim, Giancarlo Bonora, Rhys Skelton, Matteo Pelligrini, William Lowry, Reza Ardehali, Amander Clark, and Kathrin Plath.

Chapter 4 are experiments proposed and completed in work funded by the California Institute of Regenerative Medicine Clinical Fellow Training Grant (TG-01169).

This work was supported by the CIRM Clinical Fellow Training Grant (TG-01169) awarded to Sanjeet Patel and the CIRM Basic Biology 3 Grants (RB3-05080 and RB4-06133) awarded to Kathrin Plath.

VITA

Education

2000-2004	Bachelor of Arts in Kinesiology: Health Sciences Rice University Houston, TX
2004-2008	Doctor of Medicine Baylor College of Medicine Houston, TX
2008-current	Resident in General Surgery University of California Los Angeles Los Angeles, CA

Honors and Awards

2000	High School Salutatorian
2000	National Merit Scholar. Rice University.
2000	Rice/Baylor Medical Scholar Program (combined B.A./M.D.).
2002,2003	President's Honor Roll. Rice University.
2005	AMA Medical Student Research Grant in Aid
2006	American Federation of Aging Research student grant
2007	BCM Dan L Duncan Cancer Center Grant in Aid
2008	Outstanding Medical Student in Surgery.
2010-2014	Outstanding Student Teaching Award in General Surgery.
2010-2013	CIRM Clinical Fellow Training Grant
2014	Molecular Biology IDP / Amgen Dissertation Year Award Winner

Selected Publications

Agrawal N*, Iyer D*, Patel SG, Sekhar RV, Phillips TM, Schubert U, Oplt T, Buras ED, Samson SL, Couturier J, Lewis DE, Rodriguez-Barradas MC, Jahoor F, Kino T, Kopp J, Balasubramanyam A. HIV-1 Vpr Induces Adipose Dysfunction *In Vivo* via Reciprocal Effects on PPAR/GR Coregulation: Mechanism of HIV-associated Fat Metabolic Defects. Science Translational Medicine: 2013 November. PMID : 24285483

Sridharan R,Gonzales-Cope M, Chronis C, Bonora G, McKee R, Huang C, Patel S ,Lopez D, Mishra N, Pellegrini M ,Carey M ,Garcia BA, Plath K. Proteomic and genomic approaches reveal critical functions of H3K9 methylation and heterochromatin protein-1 in reprogramming to pluripotency. Nature Cell Biology. 2013 June. PMID: 23748610

Patel SG*, Hsu J*, Jahoor F, Coraza I, Bain J, Stevens R, Iyer D, Nalini R, Ozer K, Hampe C, Newgard C, Balasubramanyam A. Pathogenesis of A-β+ Ketosis-Prone Diabetes. Diabetes. 2012 May. PMID: 23160531

Donahue TR, Tran LM, Hill R, Li Y, Kovochich A, Calvopina JH, Patel SG, Wu N, Hindoyan A, Farrell JJ, Li X, Dawson DW, Wu H. Integrative survival-based molecular profiling of human pancreatic cancer. Clin Cancer Res. 2012 Mar PMID: 22261810.

Diaz Perez SV, Kim R, Li Z, Marquez VE, Patel S, Plath K, Clark AT. Derivation of new human embryonic stem cell lines reveals rapid epigenetic progression in vitro that can be prevented by chemical modification of chromatin. Hum Mol Genet. 2012 Feb 15. PMID: 22058289

Minkovsky A*, Patel SG*, Plath K. Pluripotency and the transcriptional inactivation of the female mammalian X chromosome. Stem Cells. Invited Review. 2012 Jan. PMID: 21997775

Couturier J, Patel SG, Balasubramanyam A, Lewis D. Human peripheral blood monocytes accelerate preadipocyte proliferation and inhibit differentiation in association with aberrant CCAAT/enhancer binding protein (C/EBP) alpha expression. Obesity. 2012 Feb. PMID : 21869759

Cooper A, Patel S, Senadheera S, Plath K, Kohn DB, Hollis RP. Highly efficient large scale lentiviral vector concentration by tandem tangential flow filtration. Journal of virological methods. 2011 Oct. PMID: 21784103

Sekhar RV, Patel SG, Guthikonda A, Reid M, Balasubramanyam A, Taffet G, Jahoor F. Deficient synthesis of glutathione plays a key role in aging-related oxidative stress, and can be corrected by dietary cysteine and glycine supplementation. American Journal of Clinical Nutrition. 2011 Sep. PMID: 21795440

Sekhar RV, McKay SV, Patel SG, Guthikonda AP, Reddy VT, Balasubramanyam A. Jahoor F. The effect of uncontrolled type 2 diabetes on glutathione synthesis and oxidative stress. 2011 Jan. Diabetes Care. PMID: 20929994.

Sekhar RJ, Patel SG, D'Amico AD, Jianjian S, Balasubramanyam A, Rehman K, Jahoor F, Visnergarwala F. Effects of rosiglitazone on abnormal lipid kinetics in HIV associated dyslipidemic lipodystrophy: a stable isotope study. Metabolism. 2010. July PMID: 20832829

Haaland W, Scaduto D, Maldonado M, Mansouri D, Nalini R, Iyer D, Patel SG, Guthikonda A, Hampe C, Balasubramanyam A, Metzker M. The A-β- subtype of Ketosis-Prone Diabetes (KPD) is not predominantly a monogenic diabetic syndrome. Diabetes Care. 2009 Feb 19. PMID: 19228875

Chang E, Sekhar R, Patel S, Balasubramanyam A. Dysregulated energy expenditure in HIV-infected patients: a mechanistic review. Clin Infect Dis. 2007, Jun; 44(11):1509-17. PMID: 17479951.

CHAPTER 1

INTRODUCTION

The X chromosome is epigenetically unstable in female human stem cells

Human embryonic stem cells are pluripotent and thus by definition possess the capacity to differentiate into all three germ lineages: serving as a remarkable tool for both basic and translational science [1]. Pluripotent stem cells have allowed for deep conceptual advances in our understanding of early embryonic development. Given their unique ability to change cell fate on differentiation cues, they are the model system for understanding the complex interplay of genetic and epigenetic factors required to change cell identity. However, human pluripotent stem cells have a number of undesirable traits including a propensity for genomic and epigenomic aberrancies: many arising stochastically in usual culture conditions. The mechanisms surrounding this instability are largely unknown but given that human pluripotent stem cells are being considered as a source for regenerative medicine in human patients, efforts in clarifying these phenomena are incredibly important. This thesis will focus on the epigenetic state of the X chromosome in human pluripotent stem cells, and reveal an amazing epigenetic heterogeneity between and amongst these cell lines. Our approach using RNA FISH in conjunction with genomics based methods allows us to investigate the activity of one chromosome at the single cell level and use those findings to guide the population based assays yielding an incredible viewpoint of the epigenetic regulation of the X chromosome in human pluripotent stem cells. Through these investigations we find that X inactivation occurs *during* the process of human embryonic stem cell derivation, that it is liable to errors of completion and that the ascribed essential role of dosage compensation resulting from this process is actually *not* required for in vitro differentiation. These results argue for a revision of the prevailing model of X

chromosome inactivation in human pluripotent stem cells, and at least for female human pluripotent stem cells an appeal for the inclusion of epigenetic fitness when evaluating these cells for both scientific and translational goals.

X Chromosome Inactivation in mice and humans

Traditionally the fundamental process of inactivating one of the two X chromosomes in female embryos has been studied in the context of development in mice and the mechanisms have been teased apart in mouse embryonic stem cells [2-4]. The inner cell mass of the mouse pre-implantation blastocyst is the source for mouse embryonic stem cells [5]. During mouse embryogenesis changes on the X chromosome parallel changing stages of development. The paternally contributed X chromosome is initially inactive from the 4-8 cell stage and remains that way for cells giving rise to the placental structures [2,7-9] through the expression, coating and silencing of the X chromosome by the non-coding RNA *Xist*. However, cells of the inner cell mass (epiblast progenitors – cells giving rise to the embryo proper) reactivate the paternally inactive X chromosome through silencing of *Xist* (reviewed in chapter 2), resulting in cells with two active X chromosomes. During implantation, these epiblast progenitor cells inactivate one of the two X chromosomes to produce an adult female with one active X chromosome. This process of random X inactivation is recapitulated in vitro during mouse embryonic stem cell differentiation [10-11]. Hence the capacity for mouse embryonic stem cells to model the initiation of X inactivation relies on 1) the correct reflection of epigenetic and transcriptomic features characteristic of mouse pre-implantation blastocyst epiblast progenitor cells, and 2) the capacity for changing cell

identify upon stimulation towards differentiation. The process of X inactivation has been described as essential given genetic manipulations that defunctionalize *Xist* and thus prevent the initiation of X chromosome inactivation leads to embryonic lethality in vivo and cell death in vitro [2, 12- 13].

How and when the X chromosome inactivates in human development is less well known given the genetic manipulation afforded in mice are unethical in humans. Additionally, the observation that mouse embryos undergo X inactivation near the time of implantation is lost in human studies given that it is impossible to retrieve the embryo at that time of development. Although human embryonic stem cells are derived from an equivalent state of embryonic development as mouse embryonic stem cells, the state of the X chromosome in these stem cells is not as straightforward. Multiple states of X inactivation exist in these lines, and the prevailing model suggests that there is 1) a hierarchy with transitions between the reported states, and 2) the highest state in this hierarchy has two active X chromosomes with no expression of *XIST* and is capable of undergoing X inactivation [14]. Chapter 3 of this thesis will demonstrate that no reported state of human embryonic stem cells can initiate X inactivation, and that human embryonic stem cells are reflective of a developmental stage past the window for initiation of X inactivation (post-XCI). This conclusion is in agreement with the similarities in culture conditions as well as molecular phenotypes of mouse epiblast stem cells that are pluripotent yet post-XCI, and enforces the notion that multiple classes of pluripotency exist and the relationship of each to a developmentally identifiable stage in vivo is complex.

States of pluripotency

Pluripotency is defined as the capability of cells to differentiate into any cell type of the organism, and although initially described to occur during pre-implantation of teratocarcinoma cells, it has been best defined by embryonic stem cells. One of the most intriguing questions regarding pluripotency is the divergent phenotype of mouse and human embryonic stem cells. Mouse embryonic stem cells were first derived from the 129 mouse strain and shown to self-renew indefinitely in the presence of the cytokine LIF [15-17]. They are genetically tractable with high rates of homologous recombination in vitro as well as capable to contribute to a developing embryo and its germline. Therefore mouse embryonic stem cells are a remarkable tool for understanding biology in mammals.

However, not all states of pluripotency are equivalent. The state of pluripotency in mouse embryonic stem cells has been described to be in the naïve state reflecting the *late* pre-implantation blastocyst and for female embryos possess two active X chromosomes with no *Xist* RNA coating. These cells are grown in serum and in the presence of the class I cytokine LIF. A more primitive variant of this state can be derived through small molecule modulation of two signaling pathways Wnt and MAPK, produce cell lines that resemble the *early* pre-implantation embryo and have been termed the ground state of naïve pluripotency or ground state [19]. These two states are both pre-XCI and reflect pre-implantation development, however stem cells derived from the post-implantation mouse embryo (EpiSCs) can also give rise to pluripotent stem cells *in vitro*. These stem cells are liable to karyotypic abnormalities, are incapable of contributing to the developing embryo and are dependent on the cytokine bFGF for

self-renewal [20-22]. These cells in female embryos have undergone X inactivation and have *Xist* RNA coating and inactivation of one X chromosome. Molecular analysis demonstrates that these cells are exceptionally distinct from naïve mouse embryonic stem cells. It is the culmination of these attributes that has been described as the primed state of pluripotency. Unfortunately the aforementioned naïve state in serum conditions is now confusingly being referred to as the primed state of naïve pluripotency to distinguish from the ground state of naïve pluripotency. Interestingly, although Wnt stimulation and MAPK inhibition are not required to derive embryonic stem cells from the more common 129 or C57/B6 mice, one or both are required to derived embryonic stem cells from most other strains of mice. Supporting the idea that naïve pluripotency is a restricted capability, no other mammal has reproducibly shown the relatively high degree of karyotypic stability in the naïve state *in vitro* as that from the 129 strain of mouse [23].

Primed pluripotency is however easy to establish in multiple species. It was this state of pluripotency that James Thomson initially had stabilized in non-human primates and from that work was able to derive the first human embryonic stem cell lines [1]. Like mouse EpiSCs, human embryonic stem cells have poor single cell cloning efficiency, are liable to karyotypic abnormalities in culture, are dependent on bFGF for self-renewal and high levels of ascorbic acid to limit spontaneous differentiation. They tend to express or have genes epigenetically poised for activation that are typically required for the initiation of lineage commitment and thus this state of pluripotency is developmentally more progressed than naïve pluripotency: hence primed. These states

in mice are interconvertible, in vitro, through manipulations of culture conditions or overexpression of pluripotency markers.

These identified culture conditions allowed for the derivation of human induced pluripotent stem cells (hiPSCs) where pluripotent stem cells were created through somatic cell reprogramming from human fibroblasts by overexpression of pluripotency specific genes including Oct4, Sox2, Klf4 and cMyc [24-25]. Additionally, in the standard hiPSC culture conditions all female human induced pluripotent stem cells begin as possessing one *XIST* coated X chromosome and through time in culture lose *XIST* [26]. Attempts at growing human pluripotent stem cells in mouse ground state conditions often led to the development of transgene dependent, non-autonomously self-renewing cells that were difficult to replicate across labs. However, in the past year a flurry of reports claiming the establishment of naïve pluripotency in human cells have arisen, yet all fall short upon close inspection of the X chromosome as well as the transcriptome. In essence no state of pluripotency described for human cells recapitulates the pattern of X inactivation seen in human pre-implantation blastocysts supporting the belief that these newly described cells are not truly at the naïve state. This argument is reinforced by the simple observation that key genes expressed in the human blastocyst, with homologs having core function in mouse naïve pluripotent stem cells, are simply not expressed in these putative naïve human cells. These results will be touched upon in chapter 3 and described in detail in chapter 4 [27-29]. Chapter 2 will review the foundations and principles X chromosome in activation in development and the mechanistic insights afforded through mouse and human embryonic stem cells and embryos. Chapter 3 will focus on the states of X inactivation in human embryonic stem

cells including aberrations that arise during culture of these cells. Chapter 4 will demonstrate attempts at stabilizing the naïve state in human pluripotency and potential avenues to continue trying to identify the correct state of culture. Chapter 5 will provide a summary of these results and introduce potential future directions.

REFERENCES:

1. Marahrens, Y., Panning, B., Dausman, J., Strauss, W. and Jaenisch, R. (1997) Xist-deficient mice are defective in dosage compensation but not spermatogenesis. Genes Dev. 1997 Jan 15;11(2):156-66

2. Lyon, M. Gene action in the X chromosome of the mouse (Mus musculus). Nature. 1961. 190, 372–373.

3. Takagi, N. and Sasaki, M. (1975) Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. Nature. 1975 Aug 21;256(5519):640-2.

4. Kay GF, Penny GD, Patel D, Ashworth A, Brockdorff N, Rastan S. Expression of Xist during mouse development suggests a role in the initiation of X chromosome inactivation. Cell. 1993 Jan 29;72(2):171–182.

5. Huynh KD, Lee JT. Inheritance of a pre-inactivated paternal X chromosome in early mouse embryos. Nature. 2003 Dec 18;426(6968):857-62.

6. Patrat C, Okamoto I, Diabangouaya P, Vialon V, Le Baccon P, Chow J, Heard E. Dynamic changes in paternal X-chromosome activity during imprinted X-chromosome inactivation in mice. Proc Natl Acad Sci U S A. 2009 Mar 31;106(13):5198-203.

7. Okamoto I, Otte AP, Allis CD, Reinberg D, Heard E. Epigenetic dynamics of imprinted X inactivation during early mouse development. Science. 2004 Jan 30;303(5658):644-9

8. Kay GF, Barton SC, Surani MA, Rastan S. Imprinting and X chromosome counting mechanisms determine Xist expression in early mouse development. Cell. 1994 Jun 3;77(5):639-50.

9. Mak W, Nesterova TB, de Napoles M, Appanah R, Yamanaka S, Otte AP, Brockdorff N. Reactivation of the paternal X chromosome in early mouse embryos. Science. 2004 Jan 30;303(5658):666-9.

10. Okamoto I, Otte AP, Allis CD, Reinberg D, Heard E. Epigenetic dynamics of imprinted X inactivation during early mouse development. Science. 2004 Jan 30;303(5658):644-9.

11. Penny GD, Kay GF, Sheardown SA, Rastan S, Brockdorff N. Requirement for Xist in X chromosome inactivation. Nature. 1996 Jan 11;379(6561):131-7.

12. Mugford JW, Yee D, Magnuson T. Failure of extra-embryonic progenitor maintenance in the absence of dosage compensation. Development. 2012 Jun;139(12):2130-8.

13. Silva SS, Rowntree RK, Mekhoubad S, Lee JT. X-chromosome inactivation and epigenetic fluidity in human embryonic stem cells. Proc Natl Acad Sci U S A. 2008 Mar 25;105(12):4820-5.

14. Evans MJ, Kaufman MH. Establishment in culture of pluripotential cells from mouse embryos. Nature. 1981 Jul 9;292(5819):154-6.

15. Martin GR. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci U S A. 1981 Dec;78(12):7634-8.

16. Brook FA, Gardner RL. The origin and efficient derivation of embryonic stem cells in the mouse. Proc Natl Acad Sci U S A. 1997 May 27;94(11):5709-12.

17. Nichols J, Chambers I, Taga T, Smith A. Physiological rationale for responsiveness of mouse embryonic stem cells to gp130 cytokines. Development. 2001 Jun;128(12):2333-9.

18. Ying QL, Wray J, Nichols J, Batlle-Morera L, Doble B, Woodgett J, Cohen P, Smith A. The ground state of embryonic stem cell self-renewal. Nature. 2008 May 22;453(7194):519-23.

19. Rossant J, Gardner RL, Alexandre HL. Investigation of the potency of cells from the postimplantation mouse embryo by blastocyst injection: a preliminary report. J Embryol Exp Morphol. 1978 Dec;48:239-47.

20. Brons IG,Smithers LE, Trotter MW, Rugg-Gunn P, Sun B, Chuva de Sousa Lopes SM, Howlett SK, Clarkson A, Ahrlund-Richter L, Pedersen RA, Vallier L. Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature. 2007 Jul 12;448(7150):191-5.

21. Tesar, PJ, Chenoweth JG, Brook FA, Davies TJ, Evans EP, Mack DL, Gardner RL, McKay RD. New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature. 2007 Jul 12;448(7150):196-9.

22. Kawase, E, Suemori H, Takahashi N, Okazaki K, Hashimoto K, Nakatsuji N. Strain difference in establishment of mouse embryonic stem (ES) cell lines. Int J Dev Biol. 1994 Jun;38(2):385-90.

23. Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM. Embryonic stem cell lines derived from human blastocysts. Science. 1998 Nov 6;282(5391):1145-7.

24. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, Yamanaka S. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell. 2007 Nov 30;131(5):861-72.

25. Lowry WE, Richter L, Yachechko R, Pyle AD, Tchieu J, Sridharan R, Clark AT, Plath K. Generation of human induced pluripotent stem cells from dermal fibroblasts. Proc Natl Acad Sci U S A. 2008 Feb 26;105(8):2883-8.

26. Tchieu J, Kuoy E, Chin MH, Trinh H, Patterson M, Sherman SP, Aimiuwu O, Lindgren A, Hakimian S, Zack JA, Clark AT, Pyle AD, Lowry WE, Plath K. Female human iPSCs retain an inactive X chromosome. Cell Stem Cell. 2010 Sep 3;7(3):329-42.

27. 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.

28. Chan YS, Göke J, Ng JH, Lu X, Gonzales KA, Tan CP, Tng WQ, Hong ZZ, Lim YS, Ng HH. Induction of a human pluripotent state with distinct regulatory circuitry that resembles preimplantation epiblast. CelStem Cell. 2013 Dec 5;13(6):663-75.

29. Ware, C. Caliani AJ,Deng X, Cavanaugh C, Cook S, Tesar PJ, Okada J, Margaretha L, Sperber H, Choi M, Blau CA, Treuting PM, Hawkins RD, Cirulli V, Ruohola-Baker H. Derivation of naïve human embryonic stem cells. Proc Natl Acad Sci U S A. 2014 Mar 25;111(12):4484-9.

CHAPTER 2

Pluripotency and the Transcriptional Inactivation of the female X chromosome

Abstract

X chromosome inactivation (XCI) is a striking example of developmentally regulated, wide-range heterochromatin formation that is initiated during early embryonic development. XCI is a mechanism of dosage compensation unique to placental mammals whereby one X chromosome in every diploid cell of the female organism is transcriptionally silenced to equalize X-linked gene levels to XY males. In the embryo, XCI is random with respect to whether the maternal or paternal X chromosome is inactivated and is established in epiblast cells upon implantation of the blastocyst. Conveniently, *ex vivo* differentiation of mouse embryonic stem cells (mESCs) recapitulates random XCI and permits mechanistic dissection of this stepwise process that leads to stable epigenetic silencing. Here, we focus on recent studies in mouse models characterizing the molecular players of this female-specific process with an emphasis on those relevant to the pluripotent state. Further, we will summarize advances characterizing XCI states in human pluripotent cells, where surprising differences from the mouse process may have far-reaching implications for human pluripotent cell biology.

The noncoding RNA *Xist* controls the initiation of random XCI

The importance of XCI is demonstrated by the fact that ablation of the master regulator of this process, *Xist* (X-inactive specific transcript), leads to female-specific lethality early in embryonic development in mice [1,2]. The X-linked *Xist* gene encodes an approximately 17 kb spliced and polyadenylated transcript that is essential for heterochromatin formation on the X chromosome from which it is transcribed [1–4]. In

the embryo, XCI is random based on the parent-of-origin for the inactive X (Xi), such that female organisms are mosaic for which X chromosome is expressed. *In vivo*, random XCI is initiated in epiblast cells of the inner cell mass (ICM) of the blastocyst soon after implantation and, *in vitro*, upon induction of differentiation in mESCs, which are derived from epiblast cells of the pre-implantation blastocyst. Upon initiation of XCI, *Xist* is transcriptionally upregulated on the future Xi [5,6]. It has been suggested that the transcription factor Yin-Yang 1 (YY1) tethers *Xist* RNA to its site of transcription by binding directly to both *Xist* RNA and DNA7. The RNA then spreads and creates an '*Xist* RNA cloud' demarcating the nuclear domain of the inactivating X yet the regulation of the release of Xist RNA from the Yy1 tether at the site of transcription is still unknown.

As *Xist* RNA molecules coat the X, they trigger transcriptional silencing with immediate exclusion of RNA polymerase II [8]. This is followed by loss of active chromatin marks and establishment of silencing chromatin marks, which occur in an ordered sequence of events and include, for example, trimethylation of histone H3 lysine 27 (H3K27me3) by the Polycomb complex PRC2, DNA methylation of promoter regions, and recruitment of the repressive histone variant macroH2A [9]. The result is the Xi is maintained late replicating in S phase through the lifetime of the organism. *Xist* transcription and coating of the Xi continues in somatic cells, with *Xist* RNA dissociating from the Xi in mitosis and re-coating the X in early G1 of the cell cycle [10]. Though *Xist* depletion during initiation of XCI leads to reversal of X chromosome silencing and heterochromatin formation, its deletion in somatic cells has only minor effects on Xi

reactivation as the RNA acts synergistically with other repressive chromatin modifications that accumulate on the Xi during differentiation [11,12].

Transcription and spreading of *Xist* RNA along the X is a prerequisite for silencing, which is not X-restricted as silencing can spread across X:autosome translocations and transgenic *Xist* can induce silencing of neighboring autosomal DNA [12]. The spread of *Xist* RNA-mediated silencing into autosomal regions is variable and has been proposed to correlate with the density of retrotransposons belonging to the family of long interspersed elements (L1) [13]. A recent report suggested that the silencing of X-linked L1s occurs prior to X-linked gene silencing and may promote the nucleation of heterochromatin. Conversely, specifically a subset of young L1 elements becomes transcribed upon *Xist* RNA coating and may help the local propagation of XCI [14]. In support of a functional role for L1 elements in XCI, the human X chromosome has a two-fold enrichment in L1 elements relative to autosomes [15]. Still it remains to be seen whether the behavior of these repetitive elements is a functionally important means of *Xist*-dependent facultative heterochromatin formation. In the following sections of the review we will discuss how *Xist* is regulated in pluripotent cells of the mouse.

Acquisition of pluripotency in mouse is coupled to Xi reactivation

In the mouse, XCI occurs in two forms that differ in parent-of-origin effect and in the developmental timing of initiation. Imprinted XCI, where the paternal X chromosome (Xp) is inactivated, is established in the mouse pre-implantation embryo at the four-cell stage and occurs in all cells of the pre-implantation embryo (Fig1) [16–21]. As the mid-blastocyst stage is reached (prior to implantation), imprinted XCI is reversed only in the

subset of cells in the ICM that give rise to the epiblast, so that the cells that form the future embryo carry two active X chromosomes (XaXa) without *Xist* RNA coating [16,18,21,22] (Fig1). Reactivation of the Xp is a prerequisite for subsequent random XCI in the epiblast upon implantation of the blastocyst [16,18]. In contrast, the imprinted form of XCI is maintained in the extraembryonic tissues. Random and imprinted XCI differ in the molecular requirements for initiation and reactivation. *In vivo* evidence shows that, though *Xist* RNA coats the Xp, it is not required when imprinted XCI first occurs at the four-cell stage (as it is for random XCI). Rather, *Xist* RNA coating is needed to complete and stabilize the silencing of the imprinted Xi [17,19,20]. With respect to Xi reactivation, a recent study demonstrates that the reactivation of the imprinted Xp occurs in two steps, with induction of biallelic expression of X-linked genes preceding the disappearance of *Xist* RNA coating, in agreement with the notion that *Xist* RNA coating and silencing of the Xp are uncoupled at this point in development [21]. The mechanisms that lead to gene activation on the Xp and *Xist* silencing are still unclear but linked to the specification of the epiblast lineage, as pre-implantation embryos lacking the pluripotency transcription factor Nanog are unable to specify the epiblast lineage and do not induce the loss of *Xist* RNA coating and Polycomb protein enrichment on the Xi [22]. Nanog appears to be directly involved in the regulation of *Xist* because pre-implantation embryos with a genetically engineered overexpression of Nanog lose *Xist* RNA more rapidly, though without affecting the timing of Xp reactivation [21]. However, Nanog may not be sufficient for this effect on *Xist* as Nanog is already present in the Xi-bearing cells of the late morula and becomes restricted as the

pluripotent XaXa epiblast lineage forms, indicating that other epiblast-linked mechanisms must synergize with Nanog to control *Xist* repression [21,22].

It is now appreciated that X chromosome reactivation (XCR) also occurs during the experimentally induced acquisition of pluripotency through either transcription factor-induced reprogramming to induced pluripotent stem cells (iPSCs), somatic cell nuclear transfer, or ESC/somatic cell fusion [23–25]. XCR during reprogramming of mouse somatic cells to iPSCs leads to loss of heterochromatic marks of the Xi and *Xist* repression, such that random XCI is observed upon differentiation of mouse (m) iPSCs, as in mESCs [23] (Fig1). It has been demonstrated that XCR is a late event in miPSC reprogramming, occurring at around the time of pluripotency gene activation [26], but insight into the mechanism and the events leading to Xi reactivation is still lacking. Nevertheless, the establishment of pluripotency both *in vitro* via reprogramming and *in vivo* during the establishment of the epiblast lineage in pre-implantation embryos, is coupled to XCR and *Xist* repression. Therefore, the XaXa state is a key attribute of the pluripotent state of mESCs and miPSCs.

Importantly, studies with a doxycycline-inducible *Xist* transgene have shown that *Xist*-dependent gene silencing is possible in undifferentiated male and female mESCs, but no longer after induction of differentiation or in somatic cells [12]. This observation illustrates that *Xist* function is context-dependent but not with respect to sex, as factors required for the silencing process are present in male and female undifferentiated mESCs. Since the active state of the X chromosomes must therefore be ensured by strong transcriptional repression of *Xist* in mESCs, one can view initiation of XCI upon differentiation of mESCs from the perspective of loss of *Xist* repression.

Xist is regulated by its antisense transcript *Tsix*

A major antagonizing factor to *Xist* in mESCs is another long noncoding RNA, *Tsix*, transcribed antisense to *Xist* specifically in mESCs and downregulated first on the Xi and then on the Xa during differentiation [27] (Fig2). Loss of *Tsix* function on one of the two female X's leads to slight upregulation of *Xist* transcript levels in undifferentiated mESCs and skewing of XCI towards the *Tsix*-deleted X upon differentiation [28,29]. These observations suggest that *Tsix* mainly regulates the monoallelic induction of *Xist* in the choice aspect of XCI. In support of this idea, live-cell imaging of differentiating female ESCs carrying X chromosomes tagged with a tetO array bound by a tetR-mCherry fusion confirmed a previously shown transient pairing of homologous *Xist/Tsix* regions of the two X chromosomes and demonstrated that this interaction is associated with exclusive deafening of the *Tsix* allele on the future Xi, which is proposed to allow upregulation of *Xist* [30–32]. *Tsix* antagonism of *Xist* requires transcription through the *Xist* locus and the mechanism is suggested to involve change in the chromatin structure around the *Xist* 5' regulatory region [33,34]. Together these findings indicate that *Tsix* is not the only repressor of *Xist* in pluripotency and other factors must be involved in keeping *Xist* downregulated (Fig 2-2).

Pluripotency transcription factors directly repress XCI in ESCs

Oct4, Sox2, and Nanog form a transcription factor triad that is key to maintaining ESC identity by activating genes of the self-renewal program and repressing lineage commitment genes. An attractive hypothesis for how pluripotency is directly linked to

Xist repression has come from a study that demonstrates binding of Oct4, Sox2 and Nanog to the first intron (intron I) of *Xist* in male and female mESCs and loss of this interaction upon differentiation [35]. Intriguingly, depletion of Nanog or Oct4 leads to inappropriate *Xist* upregulation in male mESCs or biallelic *Xist* upregulation in differentiating female mESCs [35,36]. It is still an open question whether specific binding at intron1 is at the heart of this XCI phenotype as these pluripotency transcription factors bind and regulate thousands of loci in the genome to maintain pluripotency. Mechanistically, the repressive function binding to intron1 has on *Xist* expression remains unclear, though one possibility is modification of the three-dimensional chromatin configuration within the *Xist* locus [37].

Already one study reports no effect of heterozygous deletion of intron1 and a very subtle skewing of XCI to the intron1-deleted X chromosome late in differentiation [38]. Conceivably, synergism of pluripotency factor binding to intron1 of *Xist* as well as other regulatory regions could suppress XCI in mESCs. In line with this model, *Tsix* transcription, particularly transcriptional elongation, is dependent upon binding of the pluripotency transcription factors Rex1, Klf4, and cMyc, within a mini-satellite region of the regulatory region of the gene, and to a lesser extent by binding of Oct4 and Sox2, with the latter being somewhat debated [36,39]. Thus, the pluripotency network may directly repress *Xist* and activate *Tsix*, which in turn contributes to the suppression of *Xist* and XCI (Fig2), an idea that could be tested with double knockout studies of intron1 and *Tsix*. Nevertheless, it may be challenging to pinpoint a role of pluripotency regulators in XCI especially as additional *Xist* activators and repressors are discovered

(see below) and transactivation or repression of these other factors by pluripotency regulators may indirectly exert XCI effects.

XCI in differentiating female mouse ESCs is governed by a balance of *Xist* activators and repressors

The mechanisms governing *Xist* upregulation during XCI must also ensure that only one X is silenced in female cells during differentiation. In addition to the X:X pairing model described above, another model proposes that in random XCI every individual X has an independent probability to initiate silencing, and this probability is proportional to the X:autosome ratio, keeping one X active per diploid chromosome set [40]. Accordingly, repressors of XCI would be autosomally encoded and activators would be X-linked. In XX cells, the double dose of the activator would stimulate *Xist* upregulation and XCI on one X, and the reciprocal *cis* silencing of the X-linked activator gene would in turn protect the other X from inactivation [40].

Rnf12, the first such characterized X-linked activator of XCI, resides ~500 kilobases from *Xist*, and encodes an E3 ubiquitin ligase bearing a RING domain. In line with a role in the initiation of XCI, Rnf12 protein levels increase in differentiation and overexpression of Rnf12 stimulates ectopic XCI [41]. The heterozygous mutation of *Rnf12* in female mESCs reduces the number of female cells undergoing XCI, however, it remains unclear if there is an essential requirement for Rnf12 in random XCI as the two published homozygous knockout strategies show contrasting results of delayed differentiation and dramatic loss of XCI [38,41,42]. These differences may be attributed to differentiation protocols as the late appearance of *Xist* RNA cloud-positive cells

suggests a selective outgrowth of cells undergoing XCI independently of *Rnf12*. Gene expression profiling suggests Rnf12 acts on *Xist*, as *Xist* was the only transcript significantly downregulated in Rnf12 knockout cells [38]. Proteomic studies will likely be necessary to see if Rnf12 plays an indirect role in XCI through ubiquitylation targets.

Two additional noncoding RNAs have recently also been identified as X-linked *Xist* activators. *Jpx*, located upstream of *Xist*, escapes XCI and increases ~10-fold during mESC differentiation. Its heterozygous deletion leads to loss of XCI and subsequent cell death upon embryoid body differentiation of female XaXa mESCs [43]. These phenotypes can be rescued by an autosomal *Jpx* transgene, indicating that this novel gene can function in *trans*, which contrasts *Xist* and *Tsix* [43]. Strikingly, the double knockout of *Jpx* and *Tsix* completely restores XCI kinetics and viability and thus it will be exciting to see how this observation and the mechanistic action of *Jpx* is explained [43]. Like *Jpx*, the noncoding transcript encoded by the neighboring *Ftx* gene is also transcriptionally upregulated with female mESC differentiation. Targeted deletion of *Ftx* suggests that its role is in controlling the chromatin structure of the *Xist* promoter [44]. It is tempting to speculate that continuous expression of these noncoding transcripts may be necessary for *Xist* itself to escape XCI. *Rnf12* and *Jpx* are both bound by Oct4, Sox2, and Nanog in mESCs, suggesting that pluripotency factors could also act on XCI through these X-linked activators [45].

In summary, the activation of *Xist*, repression of *Tsix*, and XCI during mESC differentiation depends on the downregulation of pluripotency factors and the expression of X-linked activators such as Rnf12, *Jpx* and *Ftx*, linking XCI status to the

global pluripotency gene-expression network and ensuring sex-specificity of the developmental process.

XCI in human development

Studies on XCI in human pluripotent cells have been more limited in scope because of technical challenges in manipulating human pre-implantation embryos and the ethical challenges of acquiring them. However, studies of XCI in the human system remain essential because the XCI process appears to be different from that in mouse. For instance, human pre-implantation embryos demonstrate *XIST* expression from both X chromosomes and human full-term placentas have random, rather than imprinted XCI found in mouse [46,47] (Fig 2-1).

RNA fluorescence in situ hybridization (FISH) shows *XIST* activation as a transition from a pinpoint signal to a '*XIST* RNA cloud' that can be appreciated in human female pre-implantation embryos as early as the eight-cell stage [48]. In one study, the majority of these *XIST* RNA-coated chromosomes show features of transcriptional silencing and enrichment of *XIST*-dependent repressive histone marks in the morula [48]. Contradictory results come from a more recent study, which finds that the trophectoderm and the inner cell mass of both female and male human pre-implantation blastocysts carry active X chromosomes coated by *XIST* RNA [49]. The discrepancy between the two studies may be due to different culture conditions as well as hybridization efficiencies in the FISH procedure. Regardless, it appears there is no imprinted XCI in human embryogenesis, that human XCI has different developmental

timing, and that *XIST* RNA coating of the X and XCI are uncoupled in early human embryos (Fig1).

Studies of additional factors involved in human XCI are limited to *TSIX*, which may not play a functional role in human cells. *TSIX* is transcribed in fetal cells, term placenta, and human ESCs but is truncated and lacks the CpG island essential for expression in mouse cells [50,51]. Since in human pre-Implantation development *XIST* expression appears to be uncoupled from XCI, *TSIX*-mediated regulation may be unnecessary. However, *TSIX* has not been studied in human pre-implantation blastocysts nor during initiation of XCI, therefore a potential role may have been missed [52]. Other modulators of XCI in mouse, namely *JPX, FTX*, and RNF12, have been mapped in the human genome but their functions have not yet been tested, mostly due to the lack of an *in vitro* system that allows their mechanistic dissection (see below).

Different XCI states are found in human ESCs

XCI state in human (h) ESCs is complicated by a gradual drift so that one hESC line can exhibit different states of XCI [53–56]. hESCs are grouped into three classes to describe the XCI states that are typically observed (Fig 2-1) [53]. Class I hESCs are XaXa and upregulate *XIST* and undergo XCI upon differentiation, similar to mESCs. This class seems to be the most difficult to stabilize *in vitro* because they readily transition to class II, which have initiated XCI already in the undifferentiated state and carry a *XIST*-coated Xi. Class II hESCs often further transition to class III where the silent state of the Xi is largely maintained but *XIST* is lost from the Xi along with the *XIST*-dependent histone mark H3K27me3 which leads to partial reactivation of some Xi-

linked genes [54,56]. *XIST* likely becomes silenced by methylation of its promoter region, and class III hESCs do not re-express *XIST* upon differentiation [57,53]. Given that both class I and III hESCs do not express *XIST* and lack an Xi enrichment of H3K27me3, extrapolating the XCI state solely on the basis of lack of *XIST* RNA FISH or H3K27me3 signal or even global gene expression data, has obfuscated the collective understanding of XCI in hESCs. Rather, characterization of XCI in hESC requires validation against the gold-standard assays of RNA FISH for mono- or biallelic expression of X-linked genes in addition to *XIST*.

hESCs derived and maintained in hypoxia, which is thought to better represent physiologic oxygen tension in development, preferentially remain in class I as demonstrated by RNA FISH for *XIST* and X-linked genes [56]. A switch to atmospheric oxygen tensions leads to irreversible transition to class II and subsequently to class III, strengthening the observation that female hESCs are unstable with respect to their XCI state (Fig 2-1) [56]. It will be important to determine whether this fluctuating XCI status is indicative of global epigenetic instability in hESCs.

X chromosome state in human iPSCs

Like in the mouse, human (h) iPSCs are similar to their hESC equivalent based on functional assays of pluripotency, genome-wide expression and chromatin analysis, and XCI state. At early passage, hiPSCs are class II (XaXi with *XIST* RNA coating) which readily transition to class III as *XIST* RNA is lost from the Xi (Fig 2-1) [58]. The same X chromosome is inactivated in all cells of a given hiPSC line reflecting the origin

from a single somatic cell [58,59]. These results suggest the absence of Xi reactivation during human cell reprogramming and enable the generation of hiPSC lines expressing either only the Xm or Xp [58]. Such approaches have allowed for generation of genetically-matched hiPSC lines expressing either the mutant or wild-type X-linked gene MECP2 from fibroblasts of female patients with Rett syndrome [59,60]. However, complete skewing of XCI to one X chromosome occurs upon extended passaging of fibroblasts, preventing the generation of hiPSC lines with different X chromosomes inactivated [59]. Two contradictory studies that report Xi reactivation in a subset of hiPSC lines have not performed the single cell FISH analysis of X-linked gene expression, and the skewed XCI in neurons generated from hiPSCs in one of the studies would be consistent with the lack of Xi reactivation [61,62]. Nevertheless, these results do not exclude that different culture and reprogramming conditions could lead to XCR during hiPSC induction.

Naïve versus primed pluripotency

The different XCI states in mouse and human ESCs and iPSCs suggest that either there have been significant changes to XCI in mammalian evolution or, alternatively, that these XCI states are reflective of two different developmental states 'suspended' *ex vivo* through current ESC culturing techniques. Although pluripotent cells by definition can give rise to cells of all three germ layers, distinct states of pluripotency have recently been described *in vitro*, represented by mESCs and mouse epiblast stem cells (mEpiSCs). mESCs, derived from epiblast cells of pre-implantation

blastocysts, are cultured in the presence of the cytokine LIF whereas mEpiSCs are obtained from post-implantation epiblast and cultured in the growth factor bFGF, in the absence of LIF. Since mEpiSCs express genes associated with early events in differentiation they are considered to be in the “primed” pluripotent state, whereas the typical mESC is in the “naïve” pluripotent state [63]. mEpiSCs resemble class II hESC/iPSCs in many aspects including their flat colony morphology, bFGF culture requirement, and the presence of an Xi coated by *Xist* RNA and enriched for H3K27me3 and the Polycomb protein Ezh2 [64–66]. XiXa mEpiSCs can also be generated from pre-implantation blastocysts cultured with bFGF (just like hESCs), differentiated from mESCs with bFGF and Activin A, and obtained via reprogramming of fibroblasts with Oct4, Sox2, Klf4 and cMyc in bFGF-containing media as opposed to LIF [66–68] (Fig1B). Together, the parallels between hESCs and mEpiSCs suggest that the culture of human pluripotent cells has been optimized for the primed state and not for the naïve state.

More research is necessary to molecularly define whether mEpiSCs exhibit different types of XCI states as do hESCs/iPSCs. Interestingly, it appears that compared with mouse fibroblasts, the form of XCI in mEpiSCs is a developmental intermediate and more labile with regard to reactivation based on studies transplanting nuclei into xenopusgerminal vesicles [65]. In this reprogramming system, the Xi of female mEpiSCs is receptive to nuclear reprogramming whereas the mouse fibroblast macroH2A-enriched Xi is resistant [65].

Molecular manipulation can transition mEpiSCs to the naïve pluripotent state and these approaches have been extended to the human system to generate XaXa hESCs

and hiPSCs. The reprogramming of mEpiSCs to an mESC-like state is achieved through a combination of ectopic expression of any one of the transcription factors Klf4, cMyc, Stat3 or Nanog and addition of LIF and 2i (a combination of two small molecules inhibiting GSK3β in the Wnt signaling pathway and mitogen-activated protein kinase signaling, which is thought to promote naïve pluripotency) (Fig 2-1) [22,66,69–71]. A subsequent study applied this approach to hESCs and found similar requirements for acquisition of naïve pluripotency in primed hESCs when Klf4 and Klf2 or Klf4 and Oct4 are overexpressed [72]. Prolonged maintenance of the naïve human pluripotent state appears to depend on constitutive overexpression of the reprogramming factors, indicating that the naïve human state is metastable [72,59]. As expected from the mouse system, naïve human pluripotent stem cells are XaXa without *XIST* expression, and diverge from primed pluripotent cells in both culture requirements and molecular profile as determined by gene expression microarrays [72]. As in mouse, *XIST* is re-expressed and random XCI initiated upon differentiation of naïve human cells [59,72]. The derivation of XaXa human pluripotent cells, either in the primed state under hypoxic conditions or in the naïve state, should in the future allow the modeling of initiation of XCI *ex vivo*. Yet, the relevance of modeling human XCI *ex vivo* for the XCI process occurring during human embryonic development is still unclear. During derivation and culture of human pluripotent cells, the XCI state diverges from that described for pre-implantation embryos, as the XaXa pattern with biallellic *XIST* coating of pre-implantation embryos has not been detected in cell cultures *ex vivo*. Therefore more studies are warranted but, with the approaches of these recent studies, we can already

begin to define the molecular interplay of pluripotency and XCI, akin to the mouse system, and extend these findings to optimize reprogramming to pluripotency.

Figure 2-1: **Mouse and human XCI in development and reprogramming**.

Left Panel: Species differences during normal embryogenesis with respect to XCI converge to random X inactivation in the embryo.

Right Panel: Mouse and human reprogramming can produce both naïve and primed pluripotent stem cells. * The naïve human state requires continuous ectopic expression of reprogramming factors for stability. **This figure has been modified to reflect diploid chromosomes.

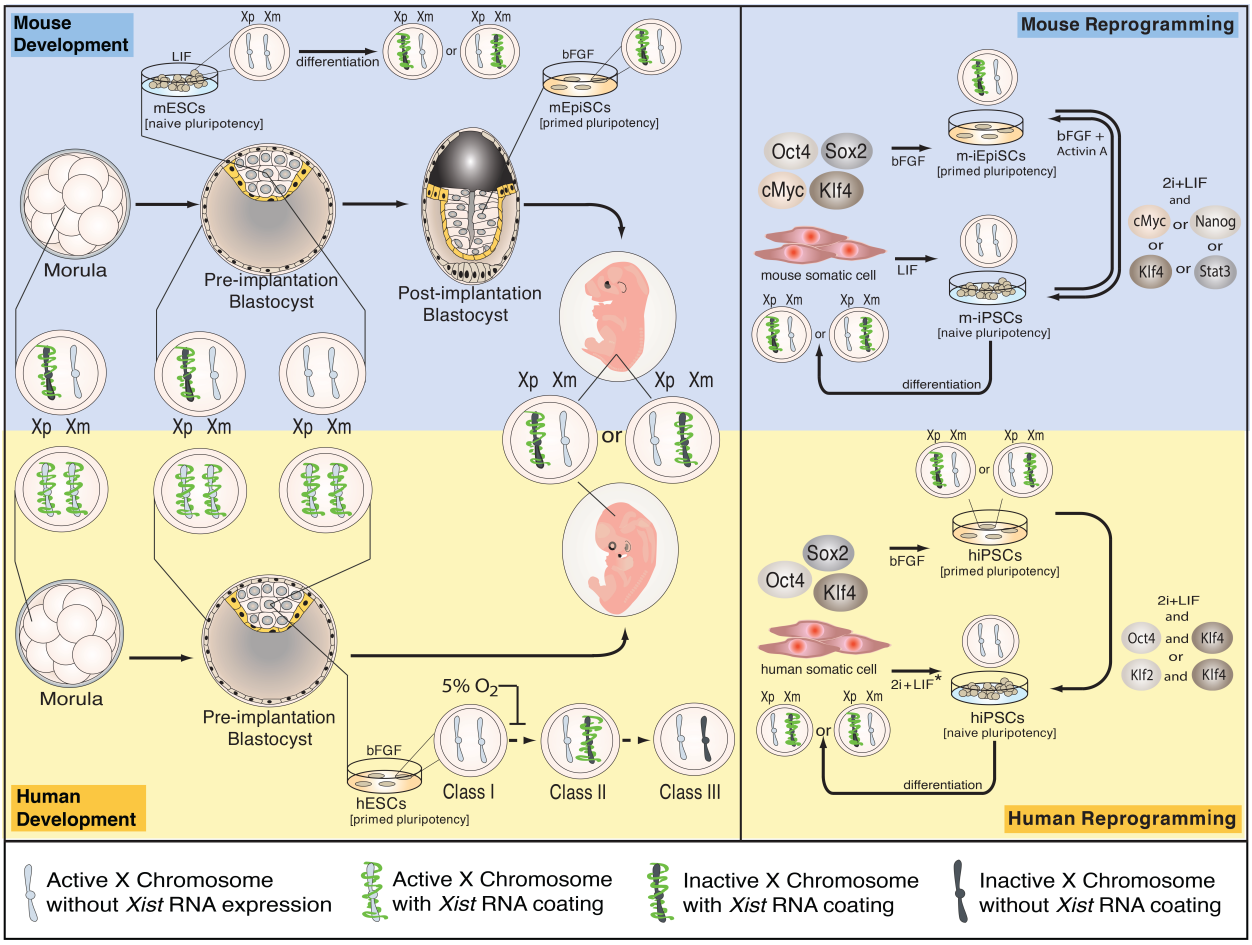
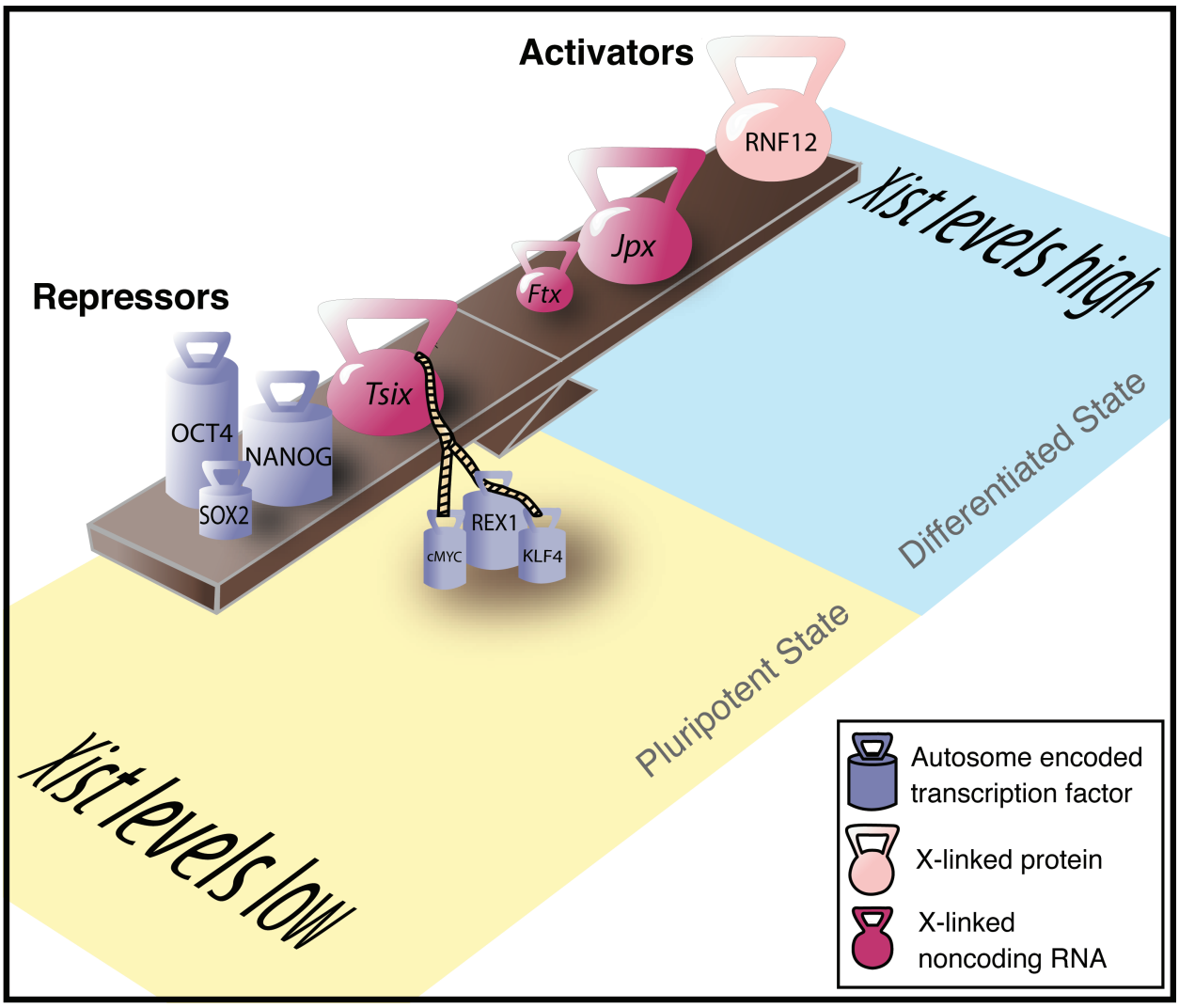

Mouse Development
Mouse Reprogramming
LIF
Xp Xm
differentiation
or
bFGF
mESCs
[naive pluripotency]
mEpiSCs
[primed pluripotency]
Morula
Pre-implantation Blastocyst
Post-implantation Blastocyst
Oct4
Sox2
cMyc
Klf4
bFGF
m-iEpiSCs
[primed pluripotency]
bFGF + Activin A
2i+LIF
and
cMyc or Nanog
or
Klf4 or Stat3
mouse somatic cell
LIF
m-iPSCs
[naive pluripotency]
differentiation
Human Development
Human Reprogramming
5% O2
bFGF
hESCs
[primed pluripotency]
Class I
Class II
Class III
hiPSCs
[primed pluripotency]
human somatic cell
2i+LIF*
hiPSCs
[naive pluripotency]
Oct4 and Klf4
or
Klf2 and Klf4
Active X Chromosome without Xist RNA expression
Active X Chromosome with Xist RNA coating
Inactive X Chromosome with Xist RNA coating
Inactive X Chromosome without Xist RNA coating

Figure 2-2: Xist activators and repressors regulate initiation of XCI in mESCs.

Xist levels are low in the undifferentiated state before onset of XCI, because of pluripotency transcription factors repressing Xist directly or indirectly via Tsix. X-linked Xist activators increase Xist levels during differentiation, as they themselves are upregulated.


Activators
RNF12
Jpx
Ftx
Repressors
Tsix
OCT4
NANOG
SOX2
REX1
cMYC
KLF4
Xist levels high
Differentiated State
Pluripotent State
Xist levels low
Autosome encoded transcription factor
X-linked protein
X-linked noncoding RNA

REFERENCES:

1. Penny GD, Kay GF, Sheardown SA, et al. Requirement for Xist in X chromosome inactivation. Nature. 1996, 379(6561):131–137.

2. Marahrens Y, Panning B, Dausman J, et al. Xist-deficient mice are defective in dosage compensation but not spermatogenesis. Genes and Development. 1997, 11(2):156–166.

3. Brockdorff N, Ashworth A, Kay GF, et al. Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. Nature. 1991, 351(6324):329–331.

4. Brown CJ, Ballabio A, Rupert JL, et al. A gene from the region of the human X inactivation center is expressed exclusively from the inactive X chromosome. Nature. 1991, 349(6304):38–44.

5. Sun BK, Deaton AM, Lee JT. A transient heterochromatic state in Xist preempts X inactivation choice without RNA stabilization. Molecular Cell. 2006, 21(5):617–628.

6. Panning B, Jaenisch R. DNA hypomethylation can activate Xist expression and silence X-linked genes. Genes and Development. 1996;10(16):1991–2002.

7. Jeon Y, Lee JT. YY1 tethers Xist RNA to the inactive X nucleation center. Cell. 2011;146(1):119–133.

8. Chaumeil J, Le Baccon P, Wutz A, et al. A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. Genes and Development. 2006;20(16):2223–2237.

9. Chow J, Heard E. X inactivation and the complexities of silencing a sex chromosome. Current Option in Cell Biology. 2009;21(3):359–366.

10. Clemson CM, McNeil JA, Willard HF, et al. XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. The Journal of Cell Biology. 1996;132(3):259–275.

11. Csankovszki G, Nagy A, Jaenisch R. Synergism of Xist RNA, DNA methylation, and histone hypoacetylation in maintaining X chromosome inactivation. The Journal of Cell Biology. 2001;153(4):773–784.

12. Wutz A. A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. Molecular Cell. 2000 Apr;5(4):695–705.

13. Lyon MF. X-chromosome inactiavtion: a repeat hypothesis. Cytogenetics and Cell Genetics. 1998;80(1–4):133–137.

14. Chow JC, Ciaudo C, Fazzari MJ, et al. LINE-1 Activity in Facultative Heterochromatin Formation during X Chromosome Inactivation. Cell. 2010;141(6):956–969.

15. Bailey JA, Carrel L, Chakravarti A, Eichler EE. Molecular evidence for a relationship between LINE-1 elements and X chromosome inactivation: the Lyon repeat hypothesis. Proceedings of The National Academy of Sciences. 200;97(12):6634–6639.

16. Okamoto I, Otte A, Allis C, et al. Epigenetic dynamics of imprinted X inactivation during early mouse development. Science. 2004;303(5658):644–649.

17. Kalantry S, Purushothaman S, Bowen RB, et al. Evidence of Xist RNA-independent initiation of mouse imprinted X-chromosome inactivation. Nature. 2009;460(7255):647–651.

18. Mak W, Nesterova TB, de Napoles M, et al. Reactivation of the paternal X chromosome in early mouse embryos. Science. 2004;303(5658):666–669.

19. Namekawa SH, Payer B, Huynh KD, et al. Two-Step Imprinted X Inactivation: Repeat versus Genic Silencing in the Mouse. Molecular and Cellular Biology. 2010;30(13):3187–3205.

20. Patrat C, Okamoto I, Diabangouaya P, et al. Dynamic changes in paternal X-chromosome activity during imprinted X-chromosome inactivation in mice. Proceedings of the National Academy of Sciences. 2009;106(13):5198–5203.

21. Williams LH, Kalantry S, Starmer J, et al. Transcription precedes loss of Xist coating and depletion of H3K27me3 during X-chromosome reprogramming in the mouse inner cell mass. Development. 2011;138(10):2049–2057.

22. Silva J, Nichols J, Theunissen TW, et al. Nanog Is the Gateway to the Pluripotent Ground State. Cell. 2009;138(4):722–737.

23. Maherali N, Sridharan R, Xie W, et al. Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. Cell Stem Cell. 2007;1(1):55–70.

24. Eggan K, Akutsu H, Hochedlinger K, et al. X-Chromosome inactivation in cloned mouse embryos. Science. 2000 Nov 24;290(5496):1578–1581.

25. Tada M, Takahama Y, Abe K, et al. Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. Current Biology. 2001;11(19):1553–1558.

26. Stadtfeld M, Maherali N, Breault DT, Hochedlinger K. Direct molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. Cell Stem Cell. 2008;2(3):230–240.

27. Lee JT, Davidow LS, Warshawsky D. Tsix, a gene antisense to Xist at the X-inactivation centre. Nature Genetics. 1999;21(4):400–404.

28. Lee JT, Lu N. Targeted mutagenesis of Tsix leads to nonrandom X inactivation. Cell. 1999;99(1):47–57.

29. Sado T, Wang Z, Sasaki H, et al. Regulation of imprinted X-chromosome inactivation in mice by Tsix. Development. 2001;128(8):1275–1286.

30. Masui O, Bonnet I, Le Baccon P, et al. Live-Cell Chromosome Dynamics and Outcome of X Chromosome Pairing Events during ES Cell Differentiation. Cell. 2011;145(3):447–458.

31. Xu N, Tsai C, Lee J. Transient homologous chromosome pairing marks the onset of X inactivation. Science. 2006;311(5764):1149–1152.

32. Bacher C, Guggiari M, Brors B, et al. Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation. Nature Cell Biology. 2006;8(3):293–299.

33. Sado T, Hoki Y, Sasaki H. Tsix silences Xist through modification of chromatin structure. Developmental Cell. 2005;9(1):159–165.

34. Navarro P, Pichard S, Ciaudo C, et al. Tsix transcription across the Xist gene alters chromatin conformation without affecting Xist transcription: implications for X-chromosome inactivation. Genes and Development. 2005;19(12):1474–1484.

35. Navarro P, Chambers I, Karwacki-Neisius V, et al. Molecular coupling of Xist regulation and pluripotency. Science. 2008;321(5896):1693–1695.

36. Donohoe ME, Silva SS, Pinter SF, et al. The pluripotency factor Oct4 interacts with Ctcf and also controls X-chromosome pairing and counting. Nature. 2009;460(7251):128–132.

37. Tsai C, Rowntree R, Cohen D, et al. Higher order chromatin structure at the X-inactivation center via looping DNA. Developmental Biology. 2008;319(2):416–425.
38. Barakat TS, Gunhanlar N, Pardo CG, et al. RNF12 activates Xist and is essential for X chromosome inactivation. Plos Genetics. 2011;7(1):e1002001.

39. Navarro P, Oldfield A, Legoupi J, et al. Molecular coupling of Tsix regulation and pluripotency. Nature. 2010;468(7322):457–460.

40. Monkhorst K, Jonkers I, Rentmeester E, et al. X inactivation counting and choice is a stochastic process: evidence for involvement of an X-linked activator. Cell. 2008;132(3):410–421.

41. Jonkers I, Barakat T, Achame E, et al. RNF12 is an X-Encoded dose-dependent activator of X chromosome inactivation. Cell. 2009;139(5):999–1011.

42. Shin J, Bossenz M, Chung Y, et al. Maternal Rnf12/RLIM is required for imprinted X-chromosome inactivation in mice. Nature. 2010;467(7318):977–981.

43. Tian D, Sun S. The Long Noncoding RNA, Jpx, Is a Molecular Switch for X Chromosome Inactivation. Cell. 2010;143(3):390–403.

44. Chureau C, Chantalat S, Romito A, et al. Ftx is a non-coding RNA which affects Xist expression and chromatin structure within the X-inactivation center region. Human Molecular Genetics. 2011;20(4):705–718.

45. Chen X, Xu H, Yuan P, et al. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell. 2008;133(6):1106–1117.

46. Daniels R, Zuccotti M, Kinis T, et al. XIST Expression in Human Oocytes and Preimplantation Embryos. The American Journal of Human Genetics. 1997;61(1):33–39

47. de Mello J, de Araújo É, Stabellini R, et al. Random X Inactivation and Extensive Mosaicism in Human Placenta Revealed by Analysis of Allele-Specific Gene Expression along the X Chromosome. Plos One. 2010;5(6):e10947.

48. van den Berg IM, Laven JSE, Stevens M, et al. X Chromosome Inactivation Is Initiated in Human Preimplantation Embryos. American Journal of Human Genetics. 2009;84(6):771–779.

49. Okamoto I, Patrat C, Thépot D, et al. Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development. Nature. 2011;472(7343):370–374.

50. Migeon BR, Chowdhury AK, Dunston JA, et al. Identification of TSIX, encoding an RNA antisense to human XIST, reveals differences from its murine counterpart: implications for X inactivation. American Journal of Human Genetics. 2001;69(5):951–960.

51. Migeon BR, Lee CH, Chowdhury AK, et al. Species differences in TSIX/Tsix reveal the roles of these genes in X-chromosome inactivation. American Journal of Human Genetics. 2002;71(2):286–293.

52. Migeon BR. Is Tsix repression of Xist specific to mouse? Nature Genetics. 2003;33(3):337. author reply 337–338.

53. Silva S, Rowntree R, Mekhoubad S, et al. X-chromosome inactivation and epigenetic fluidity in human embryonic stem cells. Proceedings of the National Academy of Sciences. 2008;105(12):4820–4825.

54. Shen Y, Matsuno Y, Fouse SD, et al. X-inactivation in female human embryonic stem cells is in a nonrandom pattern and prone to epigenetic alterations. Proceedings of the National Academy of Sciences. 2008;105(12):4709–4714.

55. Dvash T, Lavon N, Fan G. Variations of X Chromosome Inactivation Occur in Early Passages of Female Human Embryonic Stem Cells. Plos One. 2010;5(6):e11330.

56. Lengner CJ, Gimelbrant AA, Erwin JA, et al. Derivation of Pre-X Inactivation Human Embryonic Stem Cells under Physiological Oxygen Concentrations. Cell. 2010;141(5):872–883.

57. Dvash T, Fan G. Epigenetic regulation of X-inactivation in human embryonic stem cells. Epigenetics. 2009;4(1):19–22.

58. Tchieu J, Kuoy E, Chin MH, et al. Female Human iPSCs Retain an Inactive X Chromosome. Cell Stem Cell. 2010;7(3):329–342.

59. Pomp O, Dreesen O, Leong D, et al. Unexpected X Chromosome Skewing during Culture and Reprogramming of Human Somatic Cells Can Be Alleviated by Exogenous Telomerase. Cell Stem Cell. 2011;9:156–165.

60. Cheung AYL, Horvath LM, Grafodatskaya D, et al. Isolation of MECP2-null Rett Syndrome patient hiPS cells and isogenic controls through X-chromosome inactivation. Human Molecular Genetics. 2011;20(11):2103–2115.

61. Marchetto MCN, Carromeu C, Acab A, et al. A Model for Neural Development and Treatment of Rett Syndrome Using Human Induced Pluripotent Stem Cells. Cell. 2010;143(4):527–539.

62. Kim K-Y, Hysolli E, Park IH. Neuronal maturation defect in induced pluripotent stem cells from patients with Rett syndrome. Proceedings of the National Academy of Sciences. 2011;108(33):1–6.

63. Nichols J, Smith A. Naive and Primed Pluripotent States. Cell Stem Cell. 2009;4(6):487–492.

64. Tesar P, Chenoweth J, Brook F, et al. New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature. 2007;448(7150):196–199.
65. Pasque V, Gillich A, Garrett N, et al. Histone variant macroH2A confers resistance to nuclear reprogramming. The Embo Journal. 2011;30(12):2373–2387.

66. Guo G, Yang J, Nichols J, et al. Klf4 reverts developmentally programmed restriction of ground state pluripotency. Development. 2009;136(7):1063–1069.

67. Najm FJ, Chenoweth JG, Anderson PD, et al. Isolation of epiblast stem cells from preimplantation mouse embryos. Cell Stem Cell. 2011;8(3):318–325.

68. Han DW, Greber B, Wu G, et al. Direct reprogramming of fibroblasts into epiblast stem cells. Nature Cell Biology. 2010;13(1):66–71.

69. Bao S, Tang F, Li X, et al. Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. Nature. 2009;461(7268):1292–1295.

70. Hanna J, Markoulaki S, Mitalipova M, et al. Metastable Pluripotent States in NOD-Mouse-Derived ESCs. Cell Stem Cell. 2009;4(6):513–524.

71. Yang J, Oosten ALV, Theunissen TW, et al. Stat3 Activation Is Limiting for Reprogramming to Ground State Pluripotency. Cell Stem Cell. 2010;7(3):319–328.

72. Hanna J, Cheng AW, Saha K, et al. Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. Proceedings of The National Academy of Sciences. 2010;107(20):9222–9227.

CHAPTER 3

X CHROMOSOME INACTIVATION IN FEMALE HUMAN EMBRYOS AND EMBRYONIC STEM CELLS

ABSTRACT:

The mechanisms underlying X chromosome inactivation (XCI) in humans is unknown as most of our understanding of the process of XCI derives from studies of mouse cells (reviewed in chapter 2). XCI in mouse embryonic development is well understood beginning with two active X chromosomes in the undifferentiated, pluripotent state (XaXaXist-), and changes through the expression and chromosome-wide coating of *XIST* RNA on one X chromosome upon induction of differentiation, resulting in the inactivated state (XaXiXist+). This process occurs in early mouse development in vivo and can be recapitulated in vitro using female mouse pluripotent stem cells. Conversely, human pluripotent stem cells have been argued to display multiple states of XCI. In human embryonic stem cells (hESCs), the XaXaXist-, XaXiXist+, and XaXiXist- states have been described and placed in a hierarchy with XaXaXist- cells at the top (see Figure 2-1). To study XCI in female human ESCs, we established hESC lines with an XaXaXist- state, however, we found they are incapable of undergoing X inactivation. We identified that early in the derivation of female hESCs the pre-XCI state of human blastocyst showing a pattern of XaXaXist++ where two active X chromosomes both exhibit XIST RNA clouds, transitions into an XaXiXist+ state but also, stochastically, gives rise to XaXaXist- cells. The XaXiXist+ state would then subsequently silence *XIST* and re-express previously inactivated genes consistent with an erosion of the inactive X chromosome. However the XaXaXist- cells are stable in culture and thus we speculate this state is a product of disrupted XCI that arises during derivation of hESCs. Our findings show that to date no culture condition stabilizes the human pre-XCI state. Rather upon derivation and expansion of hESC lines various XCI

states arise, with none of the states allowing de novo XCI and where the state of the X chromosome present in the undifferentiated state is maintained upon differentiation. This result does indicate that XCI is not required for *in vitro* differentiation. In the future it will be interesting to uncover which aspect of X chromosome dosage correction is required for development in vivo.

INTRODUCTION:

X inactivation is a developmentally regulated process where in early female development one of the two X chromosomes is randomly chosen and transcriptionally silenced [1]. This process of random X inactivation occurs early in human development [2], yet many of the mechanistic factors leading to this dramatic change in the epigenome of the human X chromosome are unknown. By virtue of their ability to undergo differentiation to all three germ lineages, human embryonic stem cells are an attractive model system to understand mechanisms of early development and lineage commitment. As opposed to the single pre-XCI state identified in mouse ESCs with two active X chromosomes and no accumulation of *Xist* (XaXaXist-), human ESCs were reported to have multiple states of XCI both between and within ESC lines [3]. Three classes of XCI were identified using RNA FISH for the non-coding RNA (ncRNA) *XIST* and for measuring transcription of repetitive RNA with fluorescently labeled Cot-1 DNA probe. XaXaXist- hESCs showed two active X chromosomes in self-renewing conditions with a lack of *XIST* and no focal loss of Cot-1 expression, a characteristic of the inactive X chromosome. They were described to undergo XCI upon induction of differentiation. XaXiXist+ ESCs have already undergone XCI and have *XIST* coating the inactive X chromosome in the self-renewing, pluripotent state. XaXiXist- hESCs had lost *XIST* expression but maintained an inactive X chromosome. These latter two states were maintained after differentiation. The Plath lab had demonstrated a similar phenomena in female human induced pluripotent stem cells (hiPSCs) where all human iPSCs begin as XaXiXist+ and then subsequently became XaXiXist- [4]. A later study demonstrated through a clever drug selection model that in hiPSCs after *XIST* loss, the

inactive X chromosome can “erode” and reactivate the *HPRT* gene. They went on to show that through gene expression analysis from microarray data that many female human pluripotent stem cell lines had increased expression along the X chromosome in XaXeXist- lines [5]. However despite these similarities between certain states of XCI in hESCs and hiPSCs, we never observed the XaXaXist- state in hiPSCs and thus we asked if this was a state particular to hESCs. We were interested in establishing these lines given they were argued to be competent to undergo XCI upon differentiation and thus could serve as a model system to study the mechanisms behind human XCI.

To study the state of the X chromosomes in females cells, we initially focused on an approach (RNA FISH) which gives single cell resolution of the expression state of XCI, because mixed populations of XCI states can confound results from PCR based assays. Additionally, to discriminate lines with eroded genes from lines with uniformly active X chromosomes, we assayed multiple genes along both arms. We find that hESCs possess multiple states of XCI, but unlike the prevailing model all states are maintained upon differentiation and thus no class of hESC is competent to undergo de novo X inactivation. We confirm findings from the Laurent and Eggan labs that female hESCs erode the inactive X chromosome in the absence of *XIST* but show at a higher resolution that the regions harbor multiple genes that are expressed at elevated levels. However, we were unable to identify a dramatic effect of the state of XCI on the *function* of the differentiated cells and thus argue any potential detrimental effects of erosion per se can be masked by standard metrics of lineage function. Furthermore, we show for the first time that a completely XaXaXist- cell line can differentiate to three germ lineages and maintain an elevated state of X-linked gene expression. Therefore this

argues against a necessity for dosage compensation in the differentiated state in vitro and implies that the biallelic state maintains a 2:1 ratio of X-linked to Autosome gene expression in differentiated cells [6-9]. We further show that a homogenously pre-XCI state in human blastocysts rapidly changes upon derivation and heterogenous populations of cells. Thus because the naïve state of the human blastocyst is giving rise to primed state of human embryonic stem cells, the state of the X chromosome is changing alongside the identity of the cell. Together, these findings argue for the XaXaXist- state in hESCs is not pre-XCI but rather an artifact of aborted XCI in culture.

RESULTS:

Different XCI states in hESCs are detectable by single cell RNA FISH

Our lab had previously used RNA FISH against one X-linked gene to define the state of the X chromosome in human iPSCs [4]. This method allowed for single cell resolution of allelic gene expression but was limited by assaying only a small portion of the X chromosome. Previous studies on the states of XCI were limited to three defined states [3], however in the interim period a fourth state was discovered where an erosion of the Xi after loss of *XIST* RNA had been demonstrated in human iPSCs [5] and was implied from DNA methylation and array gene expression analysis in hESCs and hiPSCs [10]. Therefore, we decided to investigate at the single cell level 10 female pluripotent hESC lines derived at UCLA (UCLA 1, 3, 4, 5, 8, 9, 14, 16, 17, 18) and 3 female and 1 male pluripotent hESC lines derived at the WiCell (H9, H7, ESI03 and H1 (male)), by RNA FISH to test if erosion or biallelic expression of the X can be visualized (Fig 3-1A). These lines were all considered pluripotent by expressing markers of

pluripotency, showing normal karyotypes and passing the highest threshold for differentiation capacity: teratoma formation with all three germ lineages present. The emphasis on single cell resolution for studying XCI comes from the realization that SNP based assays for allelic gene expression cannot adequately distinguish between homogenous populations of biallelic cells versus heterogeneous populations of monoallelic cells expressing from either allele.

We identified genes that are subject to X-inactivation and escape X-inactivation (escapees) by using RNA FISH against normal human female fibroblasts. Nine genes (*Mid1, Cask, Huwe, Atrx, Ammecr1, Wdr44, Thoc2, Gpc3, Mtmr1:* red signals in representative FISH images) subject to X-inactivation and spanning the X chromosome were used for assaying the presence of either monoallelic signal (implying that the other allele is silent and inactivated) or biallelic signal (implying that both are active) (Fig 3-1A and 3-1B). We also ensured the genes were long and both rich in exon and intronic lengths to maintain a high signal to noise ratio and were of general housekeeping function to avoid stochastic fluctuations in expression between experimental conditions. We assessed specificity of the probe by restriction digest of the source BAC DNA to confirm its identity and tested the resultant labeled probe for a solitary signal in male H1 hESCs (Fig 3-1B Panel B). Finally, to control for potential changes in ploidy within the culture, escapee genes (*Utx,* yellow signals in representative FISH images) were used to identify both X chromosomes.

Importantly, we found that three states of XCI are apparent in these hESC lines. An XaXiXist- state where X-linked genes subject to inactivation (red signal) showed a predominantly monoallelic signal despite a lack of *XIST* (Fig 3-1A panel A). An

XaXeXist- state where some genes on both long and short arms of the X chromosome were biallelic, yet others were monoallelic, implying erosion of the biallelic regions (Fig 3-1A panel B and D). This eroded state showed some X-linked genes were highly susceptible to biallelism upon *XIST* loss (*Mid1, Cask, Gpc3*), whereas some X-linked genes were resistant (*Thoc2, Huwe*). Additionally, within the eroded lines themselves, not all genes were uniformly biallelic suggesting that the process of erosion was dynamic and not consistently manifested by all the cells within a given culture, and thus enforcing the notion that this phenomena is best approached through a single cell assay. Three lines, UCLA 9, 17 and 18 displayed a homogenous XaXaXist- state as all assayed genes showed a remarkably high and *consistent* biallelic signal throughout the population of cells (Fig 3-1A panel C and 3-1B panel A). Lastly we found a variant of erosion that could easily be misinterpreted as biallelic if only assaying a single X-linked gene. UCLA 4 showed some genes to be completely biallelic (*Wdr44*, *Gpc3*, *Cask*) whereas others remained monoallelic (*Huwe*). We refer to this extensively eroded state as XaXEXist-. In the interim period of these experiments the Yamanka group had published that reprogramming somatic cells to induced pluripotent stem cells (hiPSCs) or culturing pre-existing hIPSCs on mouse LIF secreting immortalized STO fibroblasts (SNL fibroblasts) could induce an XaXaXist- state in hiPSCs which, surprisingly, resulted in an XaXiXist- state upon differentiation [11]. Their claim derived from increased X-linked gene expression from gene expression microarray and biallelic expression of a single X-linked gene (*Pgk*) by RNA FISH. We received three of these lines as live cultures and 72 hours after receipt used RNA FISH with our panel of X-linked genes and found these lines to be severely eroded and thus classified as

XaXEXist- and not XaXaXist- (Fig 3-1 D). We suspect that the use of a moving average method of quantifying X-linked gene expression and the lack of an authentic XaXaXist- control line led them to erroneously conclude they had XaXaXist- hIPSCs.

Importantly, we believed that the 3 XaXaXist- ESC lines, UCLA 9, 17 and 18, were examples of the prevailing model's pristine state of human ESC capable of undergoing X inactivation upon differentiation. Therefore we decided to investigate the mechanisms of human X inactivation using these lines. However we immediately noticed that these cells had slightly different phenotypes then what was claimed. First, using immunofluorescence for NANOG, a marker restricted to the pluripotent state, no NANOG+ XaXiXist+ cells could be found within the cultures. This contradicted the claim that the XaXaXist- cells are labile and spontaneously undergo X chromosome inactivation in self-renewing conditions. Additionally, spontaneous differentiation can be found in most human pluripotent stem cell culture but we found no NANOG-, XaXiXist+ cells also indicating that these cells upon loss of the pluripotency marker were not upregulating *XIST*. (3-1B Panel C). Given these findings were reproducibly seen across all three lines, we hypothesized that perhaps the XaXaXist- cells do not undergo XCI upon differentiation. Therefore we used retinoic acid to induce differentiation and identify whether these lines were capable of upregulating *XIST* and inactivating the X chromosome upon induction of differentiation.

The pattern of XCI in the pluripotent state is maintained upon differentiation

XCI has been studied extensively during mouse ESC differentiation and therefore we wanted to identify the kinetics of XCI in the human system using retinoic acid (RA)

differentiation. To date little mechanistic insight into the process of human XCI exists because somatic cell lines cannot recapitulate the required developmental transitions in vitro [7-10]. The XaXaXist- had previously been described to undergo X inactivation upon differentiation and thus although spontaneous differentiation of these lines did not upregulate *XIST*, we asked if a more complete method of differentiation, using retinoic acid, could induce *XIST*. By using immunofluorescence for the pluripotency marker OCT4 and RNA FISH for the same panel of X-linked genes, we assayed the state of the X chromosomes in female ESCs in differentiated conditions (Fig 3-2A). We found that the proportion of X-linked gene expression demonstrating monoallelic (light grey bars) or biallelic RNA FISH signals (black bars) in OCT4-negative cells were largely similar to the pluripotent cell of origin (Fig 3-2A, Fig 3-2B).

The XaXiXist- state has previously been shown to maintain this state in differentiation, and the XaXeXist- state implied to persist through studies of DNA methylation of X-linked genes [10]. We confirmed these findings using RNA FISH. However unlike the previous description of the XaXaXist- state [3], these lines also did not undergo X inactivation: no OCT4- negative cells demonstrated *XIST* upregulation nor silencing of one of the two X chromosomes (Fig 3-2B panel B). Enforcing the hypothesis that XCI cannot occur in established human ESC lines undergoing in vitro differentiation, the enrichment of H3K27Me3 on the Xi, observed when *XIST* is expressed, was absent in all retinoic acid differentiations but present as expected in normal female dermal fibroblasts (Fig 3-2B panel B). The finding of a complete lack of de novo XCI upon induction of differentiation of hESCs has never been described

before predominantly due to most published findings investigating the X chromosome in hESCs did not investigate differentiation of pure XaXaXist- hESC lines.

This was surprising given the process of XCI was described as essential for loss of self-renewal and early development in mice. However, both spontaneous differentiation and retinoic acid induced differentiation may produce cells with unknown *in vivo* counterparts, and therefore we differentiated these cell lines to defined cell types with assays that could describe cell specific function. We differentiated XaXaXist-, XaXeXist-, and XaXiXist- cells towards endoderm lineage (hepatocyte-like cells) and mesoderm lineage (cardiomyocyte-like cells). For hepatocyte-like cells we used immunofluorescence for human albumin and found again the state of XCI was maintained (Fig 3-2A Panel C, Fig 3-2B panel C) with no evidence to support and upregulation of *XIST* either directly using RNA FISH or indirectly through immunofluorescence for an H3K27Me3 foci. These hepatocyte-like cells produced human albumin at levels similar to each other implying that the original state of XCI did not affect the synthetic capacity of the differentiated cells (Fig 3-2A Panel B). For cardiomyocyte-like cells, we found that the cardiac troponin I positive cells also maintained the state of XCI (Fig 3-2A Panel B). These cardiomyocyte like cells would spontaneously contract, show appropriate calcium transients implying a functioning sarcoplasmic reticulum and respond appropriately to beta-adrenergic stimulation (Fig 3-2A Panel B and C, Fig 3-2B Panel D and E, Movie Fig 3-1). The cardiomyocyte-like lines were further evaluated by genome-wide SNP/CNV array and found to be similar to the respective starting pluripotent line with no overt increase in chromosomal copy number indicating the biallelic expression of genes was from erosion and not changes in

ploidy (Fig 3-2D). We were also able to derive neurons with robust neuronal projections and outgrowths from the XaXaXist- line, UCLA9, to confirm that neuroectodermal differentiation is not grossly affected by the complete lack of XCI (Fig 3-2A). These findings are supported by the following two published reports. First, the Eggan lab had demonstrated using female and male hIPSCs that model the X-linked disorder Lesch-Nyhan Syndrome, that female hIPSCs that are isogenic but differ in regard to which HPRT is expressed have different neural differentiation propensities. They show that female hIPSC that have an $Xa^{HPRT\text{-}WT}Xi^{HPRT\text{-}mut}$ state can produce normal neurons but female hIPSCs that have an $Xa^{HPRT\text{-}mut}Xi^{HPRT\text{-}WT}$ state have decreased neurite projections and decreased propensity for neuronal differentiation. The latter state mimicked the differentiation of male $Xa^{HPRT\text{-}mut}$ hIPSCs. However, upon extended culturing the $Xa^{HPRT\text{-}mut}Xi^{HPRT\text{-}WT}$ hIPSCs would silence *XIST* and erode the Xi leading to reactivation of the wildtype HPRT and loss of the neuronal differentiation defect. This argues in their system that erosion, per se, was not detrimental to neuronal differentiation but rather the expression of the mutant form of HPRT was the source of the defect [5]. Another example implying erosion of the Xi may not drive defective differentiation comes from a large study from the Meissner lab (Fig 3-2C Panel A and B) [12]. Using their data, we demonstrate that for a panel of human pluripotent stem cells with different degree of erosion, there is no strong correlation between erosion and differentiation propensity.

Although our derivative hepatocyte-like and cardiomyocyte-like cell lines showed evidence of attaining cell type specific function (Movie 3-1), we cannot conclude from these experiments if more in depth analysis would reveal subtle functional defects

between XaXa, XaXe, XaXE, and XaXi cells. However, the capability of hESCs to produce these cells in vitro with normal karyotypes, and additionally a competence to differentiate in teratoma assays, imply that the current measures of human pluripotency are incapable of capturing many aspects of epigenetic heterogeneity. How this impacts more defined investigations using derivatives of these lines is an important yet understudied area.

DNA methylation patterns across the X-chromosome capture XCI states

RNA FISH for X-linked genes at the single cell level provides a resolution that cannot be recapitulated in population based assays, however the sensitivity for the activity of the entire X chromosome with its nearly 1000 genes is impossible to quantitate by FISH, and so we decided to use a genomics approach to complement the RNA FISH results and identify if across substantial differences existed between eroded and XaXaXist- lines across the chromosome. We used reduced representation bisulfite sequencing (RRBS) to map the methylation profiles for CpG islands of genes on the X chromosome in our lines both in self-renewing and differentiated conditions as had previously been shown to loosely correlate with increased gene expression [10]. We mapped DNA methylation in XaXiXist-, XaXeXist- and XaXaXist- lines, clustered the results according to the level of methylation within CpG islands, and found that our classification system from RNA FISH was reproducibly confirmed by RRBS (Fig3-3A Panel C). The XaXiXist- lines showed a highly uniform pattern of hemi-methylation (yellow) across the X chromosome (Fig 3-3A Panel D). The XaXeXist- lines showed

patches of demethylation (green) consistent with focal areas of erosion, and many seemed to occur in “hotspots” within the middle of the long and short arms (Fig3-3A panel A and D). The XaXEXist- hESC line (UCLA4) showed maintenance of methylation near the centromeric and telomeric regions of the X chromosome (Fig 3-3A panel A and D) and the hIPSC lines K3F1 and K3S5F4 from the Yamanaka lab grown on SNL feeders also demonstrated these remnants of methylation (Fig 3-3B Panel E) confirming their status as severely eroded and not like the XaXaXist- lines derived at UCLA [11].

The XaXaXist- lines possessed a methylation pattern that mirrored the male ESC suggesting that **both** X chromosomes are demethylated to same extent as the solitary male X chromosome (Fig3-3A Panel and D). This confirmed our findings from RNA FISH where the homogeneity in biallelic FISH signal was arising from a pure population of XaXaXist- cells. As seen earlier by RNA FISH, after retinoic acid induced differentiation, all XCI states were maintained and the differentiated cell’s RRBS patterns clustered with the self-renewing pattern on the X-chromosome (Fig3-3A Panel A), but not for the autosomes (Fig 3-3A panel B). This indicates that the methylation pattern on the X chromosome is conferred to the differentiated cell, but the cell fate change is dramatic and thus the clustering on the autosomes are by cell state. Furthermore cardiac differentiation and neuronal differentiation mirrored the parental line (Fig 3-3A panel A), in agreement with the results from the retinoic acid induced differentiation. These assays verified the observations from RNA FISH and strengthen the conclusion that the XaXaXist- state of human ESCs does not undergo XCI upon

differentiation but rather maintains that state just as all other states of XCI do in human ESCs.

Dosage compensation is not required for pluripotent or differentiated cells *in vitro*

XCI has traditionally been described as an essential process in normal female development, which leads to the inactivation of the diploid X chromosome such that dosage compensation to the solitary male chromosome can be achieved. However, this leaves the sex chromosomes genetically haploinsufficient and thus multiple studies have investigated the mechanism of compensation to the autosomes as haploinsufficiency is usually not tolerated. We find that both our XaXeXist- and XaXaXist- cell lines in differentiated conditions have a loss of dosage compensation and increased gene expression from the X chromosome (Fig3-3A panels E and F).

To identify whether erosion resulted in an overall increased gene expression from the X-chromosome, and thus also involved a loss of dosage compensation, we used RNA-seq and enumerated first-strand synthesized transcripts using a digital barcoding scheme resistant to bias created from PCR to identify if demethylated regions produce an increase in total expression and not compensated by autosomal expression. The X:A ratios from median X-linked gene expression and median autosomal gene expression in the undifferentiated cell lines is 2:1 in the XaXa and XaXe lines, whereas the XaXe lines are 1.5 and the XaXi lines 1.0. (Fig3-3A panels E and F). This verifies that dosage compensation is not occurring through other mechanisms involving down-regulation of autosomal gene expression in the pluripotent state and agrees with the

typical finding that adult females with appropriate XCI are in effect genetically hemizygous for the X chromosome but are not haploinsufficient by gene dosage and maintain a 1:1 ratio of X to autosomal gene expression [6-9]. However, in this case the lack of XCI in the XaXaXist-, and to a lesser extent XaXeXist+, is in effect producing a functional tetraploidy in terms of gene expression where the two X chromosomes are delivering a gene dose that is twice that of the autosomes. How this is regulated in the inner cell mass of developing humans is unknown but it has been suggested that the global demethylation found in XaXa female mESCs not identified in male mouse ESCs is to potentially compensate for the increased gene expression from the two active X chromosomes to maintain an X:A ratio less than 2:1 [13-15]. However, we found no evidence of profound autosomal demethylation in XaXaXist- cell lines by RRBS nor between male control and female XaXaXist- lines, indicating even if this phenomena does occur in response to two active X chromosomes in female mouse embryonic stem cells, it is not prevalent in human embryonic stem cells with two active X chromosomes.

We used cumulative distribution frequency plots to determine if the increased expression was limited to specific genes or across the X chromosome and saw that in the XaXe and XaXa lines the expression was increased at all levels of gene expression (Fig 3-3A panel E). These distributions were significantly different and increased compared to the other lines assayed (two-tailed KS test $p < 0.01$). We used the identified eroded regions (based on DNA methylation) for each line (Fig 3-3A Panel C) and performed an expression analysis to determine if the eroded regions in the XaXeXist- lines showed an expression pattern similar to the same region in the XaXaXist- lines and to determine if the inactivated regions showed an expression

pattern similar to the same region in the XaXiXist- lines (Fig 3-3A Panel H). We find that in the hESC lines UCLA1, 8, 4 and ESI02, the regions that are demethylated and thus eroded, express genes at a level similar to the XaXaXist- lines, while the regions that maintain hemi-methylation at a level similiar to the XaXiXist- lines (KS test $p>0.05$).

Pluripotent stem cells may be able to handle this increased gene expression from the X chromosome but the consequence may be more significant in the differentiated state so we used the same analysis on the cardiac differentiated cell lines. This showed that the expression from the X chromosome was increased (Fig 3-3A panel I) and the expression from regions eroded and inactivated followed the same pattern of differential expression as in the pluripotent state (Fig 3-3A panel J). The X:A ratios were elevated for the XaXa and XaXE cardiac lines (Fig 3-3A Panel K, $p < 0.05$) and the differentiation was determined to be complete from the perspective of gene expression differences between the pluripotent and cardiac differentiated lines (Fig 3-3 Panel L, $p < 0.05$) arguing against a possible maintenance of a pluripotent state leading to persistence of the increased gene expression from the eroded regions.

Naive human pluripotent stem cell lines described to date do not display the XaXa gene expression signature

Culture conditions for derivation and maintenance of human embryonic stem cells have undergone many iterations, from the original description of feeder based culture in serum to new formulations attempting to promote increased stability and ease of passaging [1, 16]. However, the pursuit of a state of human pluripotency that is more immature than primed pluripotency has arisen largely from the success of ground state

culture conditions in stabilizing mouse embryonic stem cells from recalcitrant strains and extension into other species like rats. These initial successes in rodentia were quickly followed by a paper from the Jaenisch lab where human iPSCs cultured in ground state conditions supplemented with the cAMP agonist forskolin were argued to be in the naïve state of pluripotency [17]. To match the proposed state, the study authors argued that these cells were competent to undergo X inactivation using RNA FISH for *XIST* as well as complementary evidence using bisulfite PCR for the *XIST* promoter. We were very interested in using this proposed XaXa system capable of undergoing de novo XCI given all XaXa hESCs lines we had established and analyzed were incapable of undergoing X inactivation. Therefore, we assayed the gene expression profile of the Jaenisch lab lines from the published datasets in conjunction with the gene expression profiles for arrays on the same platform as the lines derived at UCLA. We were interested in identifying if the increased gene expression we observed for the XaXaXist- lines were also found in these naïve lines. We further wanted to identify if these lines were truly biallelic by determining whether they possessed an increased X-linked gene expression in erosion resistant regions: defined by the consistent difference in DNA methylation found in the UCLA XaXaXist- lines versus the XaXEXist- lines, enriched at the peri-centromeric and peri-telomeric regions. We found that the lines generated in the 2010 paper did not meet the increase in gene expression we see in our XaXaXist- line (Figure 3-3B panel B; KS test $p < 0.05$) but rather the C1 hIPSC line was similar to the highly eroded UCLA4 line. By limiting the analysis to genes that are resistant to erosion in our hands, no line generated in the 2010 paper had the increased gene expression seen in these regions (Figure 3-3B panel C).

Using more refined conditions, in 2013, this same authors were able to demonstrate that the original formulation was inadequate and that conversion to the naïve state required additional constant bFGF and Tgf-beta stimulation as well as more broad MAPK and JNK inhibition [18]. This new version of naïve human pluripotent stem cells also was argued to undergo X inactivation upon differentiation by appearance of an H3K27Me3 foci in differentiated cells. Intriguingly, these cells were argued to be able to undergo chimerism in mouse embryos where the isogenic primed pluripotent lines could not, potentially reflecting the true naïve state of human pluripotency in vitro. We were fortunate to receive two hESC lines completely derived in these refined naïve conditions as described in 2013 from the author and found that by RNA FISH the male line has a solitary X chromosome with no expression of *XIST* and the female line has a pattern that is XaXiXist- (Fig 3-4B Panel C). We can see through the expression profiles that these new naïve lines also do not express genes along the X chromosome as highly as our XaXaXist- hESC lines (KS test $p < 0.01$) after normalization of the two platforms using quantile normalization. Additionally, the C1 hIPSC line derived in 2010 and cultured in the refined conditions seems to express the same degree of erosion induced expression implying the increase in expression is carried through from the original derivation and not a result of the new culture conditions. In retrospect, this study claimed a new state of pluripotency in human cells by using assays that at first glance seemed appropriate: utilization of the Oct4 distal enhancer, chimerism to early embryos (mouse), X inactivation upon differentiation, a global demethylation pattern seen in ground state mouse embryonic stem cells and high rates of single cell cloning efficiencies. However, the preferential use of the distal enhancer in human blastocysts

have never been demonstrated, the chimerism was weak at best and that assay can support teratocarcinoma cell engraftment [19] questioning its ability to discriminate pluripotent embryonic stem cells from transformed cells, the H3K27Me3 pattern was multi-focal and *XIST* RNA was never shown to increase upon differentiation, the claimed demethylation was nowhere near the degree seen in mouse ground state conditions and the single cell cloning efficiency had been described before in a simpler formulation [20].

A separate study in 2010 argued that the derivation of hESCs in hypoxia could stabilize the XaXaXist- state (at that time believed to be labile in hESC cultures) and, could undergo de novo XCI upon transfer to normoxic conditions [21]. This study argued that normoxia induces a stress on the cells that induces premature XCI from a pre-XCI state (XaXaXist-) in hESCs. We received live cultures of the lines WIBR2 and WIBR3 at low passage, cultured them briefly in hypoxia, and assayed them using RNA FISH. We found that they were XaXiXist- and XaXiXist+, respectively, and thus hypothesized that the SNP based assay used to determine biallelism was describing erosion and not true biallelism as the report was on cells 40-50 passages beyond those we had assayed at the single cell level. Interestingly, at low passage, the expression arrays showed little to no difference in X-linked gene expression between normoxic and hypoxic states and thus further argued against these cells from being XaXaXist- (Fig 3-3B Panel D). Finally we used these erosion resistant regions to identify if a culture condition using SNL feeders published by the Yamanaka lab was inducing a pattern of erosion or met the methylation and RNA FISH signature of XaXaXist- cells. Three lines K3F1, K3F2 and K3FS54 from the Yamanaka group harvested from live cultures within

24 hours of receipt were processed for RRBS or plated and fixed within 72 hours for RNA FISH showed evidence of extensive erosion but no evidence of a state consistent with our XaXa lines (Fig 3-3A panel E, Fig 3-1 panel D) [11]. From our results, we conclude that using gene expression microarrays and observing biallelism of a single X-linked gene using RNA FISH is insufficient to distinguish XaXEXist- cells from XaXaXist- cells. Therefore, we argue that culture on SNL fibroblasts leads to erosion of the Xi and not reactivation.

Given that these profoundly different culture conditions are both creating confusion regarding the state of the X chromosome but also incapable of reflecting a state that has two active X chromosomes, we decided to investigate how the X chromosome state changes during the process of hESC derivation itself. We hypothesized that given the XaXaXist- cells do not undergo X inactivation then the prevailing model describing the sequence of events from the blastocyst to ESC line derivation is incorrect. We thus wanted to identify which characteristics of the state transitions that had been reported in the literature could be reproduced in our hands as well as identify what the sequence of events surrounding the establishment of XCI in hESCs. Therefore we turned our attention to the events that occur during early derivation and passage of hESCs to search for the time point when the initiation of XCI was occurring.

XaXiXist+ cells undergo erosion during early passages

The prevailing model describes a hierarchy XCI states beginning with XaXaXist- cells transitioning into XaXiXist+ and then XaXiXist-. However, we have observed the

former transition in our lines. Namely, the XaXaXist- state was stable and not subject to drift towards the XaXiXist- state. We thus focused on the state of XCI across time in early derivation cultures to identify how the population could change. To address this question we used RNA FISH to follow the state of the X chromosome in early passage hESC lines. In two lines beginning as a predominantly *XIST* positive state early after derivation we observed a progression to an *XIST* negative state during standard culturing. These lines (UCLA8 and UCLA14) lost *XIST* quickly and heterogeneously over time but at different rates (Fig 3-4A panel A). The loss of *XIST* did not coincide with the erosion of *ATRX* into a biallelic signal at these early passages but remained monoallelic. However, from the panel of lines studied earlier this gene was not subject to early erosion and thus likely was not representing the expression of the entire X chromosome after the loss of *XIST*. Thus we split and subcloned UCLA8 into *XIST* + and *XIST* - clones and along with the XaXaXist- hESC line, UCLA9, used RNA-seq to quantify the expression of X-linked genes. We identified that the *XIST* + subclones of UCLA 8: 3,10,23 exhibited different proportions of *XIST* + cells by RNA FISH. Additionally, the median CDF of X-linked gene expression inversely correlated with the degree of *XIST* loss implying that erosion of gene expression occurs quite rapidly after *XIST* loss (Fig 3-4A Panel B).

When following the three XaXaXist- derived at UCLA over time, we found that this state remained stable over time in culture and was present already in the earliest passages (Fig 3-4A Panel C). This was in contrast to what was previously thought regarding the XaXa state as extremely labile and subject to inactivation upon stress [21]. Rather, all three lines across extended culturing did not show evidence of X

inactivation and were remarkably stable as XaXa. Furthermore, hESC lines UCLA17 and UCLA18 were followed from the derivation culture (Passage 0, 3 weeks after plating) by RNA FISH through subsequent passages (Fig 3-4A Panel D). In the initial derivation culture, UCLA17 showed a vast majority of cells with no evidence of XCI and only 20 nuclei could be found in over 800 with an *XIST* cloud and evidence of silencing from an X-linked gene (*HUWE1*). UCLA18 on the other hand had 11% Xist positive nuclei with evidence of XCI demonstrated by H3K27Me3 foci as well as absence of x-linked gene expression (Fig 3-4A Panel E). However, upon examining the entire culture, the Xist+ cells clustered together, implying these cells arose from a common progenitor(s). Given the majority of the culture was XaXaXist-, and at UCLA lines are derived in a subcloning manner where pieces rather than whole derivation cultures are used, the purity of our XaXaXist- cell lines were easily explained.

Upon careful review of the original paper describing XCI states in human ESCs [3], the previously ascribed XaXaXist- cell lines began as a mixed culture of XaXaXist- and XaXiXist+ cells and thus our findings could be placed in context. XaXaXist- lines are derived from XaXaXist- cells that can be found mixed in a derivation culture with XaXiXist+ cells and thus the method for propagation of a line from derivation cultures will result in a mixed population in the cell line as in the case of previous publications or a pure cell line in the case of our derivation procedures.

The XaXaXist- hESC state appears in very early derivation cultures at a time when de novo X inactivation occurs

The pristine state of human ESCs as determined from XCI largely arose from two observations 1) that the cells began with two active X chromosomes as seen in mouse ESCs, and 2) upon differentiation one X chromosome was inactivated and coated with *XIST*. This model was largely guided with the known progression of XCI in mouse ES and early development. We decided to identify the state of the human blastocyst in IVF donated embryos at UCLA and track the progression of XCI during early derivation to clarify how the naïve embryo transitions from a pre-XCI to post-XCI state during human ESC derivation. Both female and male pre-implantation embryos showed *XIST* coating their respective X chromosomes and using RNA FISH for *XIST* and an X-linked gene resistant to erosion in established lines, *HUWE1*, showed active X chromosomes (Fig 3-5 Panel A; Movie 3-2). This was consistent with the findings of Okamoto et al and implied that the pre-XCI state of humans is very different from rodents and is mirrored by indirect evidence using PCR for *XIST* from non-human primates [22-23]. During the first 48 hours, approximately 24 hours after the blastocyst settles onto the feeders the majority of cells maintain the XCI state of the pre-implantation embryo (Fig 3-5A panel B and C). At 96 hours after plating, various states of XCI were identified and although many could be placed in a logical sequence from the XaXaXist++ to XaXiXist+ state, approximately 30% have silenced Xist on both alleles and by RNA FISH resemble the XaXa human ESC state. The kinetics of XCI was not consistent as a other experiments had shown XCI occurring at 48 hours after plating and largely complete by 96 hours (Fig 3-5B Panel A and B). A logical sequence would then entail loss of Xist on both chromosomes and subsequent upregulation on one with later silencing of that allele to become the Xi (Fig 3-5A Panel D); however, we cannot exclude the possibility that the

XaXaXist- state is not within the normal sequence of human XCI. The large fraction of cells with the XaXaXist- state observed during the early course of derivation could lead to a greater chance of an XaXaXist- cell line to arise. We speculate the establishment of multiple classes of XCI within the same line results from the progression of the naïve pluripotent pre-implantation blastocyst (XaXaXist++) to the primed human ESC state (XaXiXist+ and XaXaXist-). Importantly, the naïve pattern of the female human blastocyst showing an XaXaXist++ state has never been described in currently published naïve cells and this metric may aid in identifying a pluripotent stem cell that closely mirrors the pre-implantation human blastocyst and thus serve as a useful tool for studying human X inactivation in vitro.

DISCUSSION:

Our work describes a revision of the currently understood model of XCI states in human embryonic stem cells and shows that the primed state human embryonic stem cells is marked by several patterns of XCI never seen in the pre-implantation blastocyst. we show that the female X chromosome in human embryonic stem cells are subject to lability in patterns and this arises very early in derivation culture. Additionally, we demonstrate that loss of self-renewal as well as differentiation to specific cell types do not induce X inactivation but rather maintain the state that existed in the pluripotent cell of origin. The most interesting finding specific to human embryonic stem cells is that the XaXaXist- cell types do not undergo dosage compensation upon differentiation. This finding indicates that dosage compensation is not required for cell survival in the differentiated state, in vitro. These findings also re-emphasize that for studying X inactivation in the human system, human embryonic stem cells are an inadequate model and the stabilization of naïve pluripotency which would capture the pre-implantation blastocyst cell state and thus be pre-XCI is vital.

Female human embryonic stem cells derived at UCLA as well as through the WiCell can thus be loosely categorized into four classes: XaXaXist- where both gene expression and DNA methylation of X-linked genes resemble the male solitary X yet the dose of genes expressed is 2 fold higher, XaXiXist+ where both gene expression and DNA methylation show an inactive and an active X chromosome, and then XaX(i/e)Xist- state where loss of Xist leads to reactivation of portions of the X chromosome and result in both a loss of DNA methylation in those regions as well as increased gene

expression. This latter state has been described in human induced pluripotent stem cells to occur after Xist loss [5] as well as the DNA methylation patterns and described in both human embryonic stem cells as well human induced pluripotent stem cells [10]. We show that the regions themselves that are eroded show a higher expression than the ones that remain inactivated implying a tight link between gene expression levels and maintenance of X inactivation.

This work hopes to clarify some aspects of confusion in the literature regarding the state of the X chromosome in human embryonic stem cells. First, the initial description of the classes of XCI in hESC was based on the model developed from our understandings of mouse XCI. This work and others have shown that human pre-implantation blastocysts are not XaXaXist- but actually XaXaXist++. Therefore the idea that the XaXaXist- state found in human ESCs to be equivalent to the pre-XCI state found in mouse ESCs is flawed. Additionally, the original description of XaXaXist- cells included that upon differentiation *XIST* is upregulated on one X chromosome and random X inactivation occurs. However, this study began with a mixed population of cells resembling the derivation cultures we have seen and therefore upon differentiation the data suggested that some of those XaXiXist+ cells simply came through. We were fortunate to derive human ESCs as near clonal populations and thus were able to identify that class we cells do not upregulate Xist neither through serial culture nor through differentiation.

This the notion that the XaXaXist- cells are representative of the human blastocyst led to a number of papers describing the acquisition of that state of XCI and further claimed multiple mechanisms to stabilize that state including hypoxic growth

conditions [21], antioxidant treatment [21], MAPK and GSK3B antagonism [17], HDAC treatment [24], growth on mouse LIF producing feeder fibroblasts [11]. However no study was able to convincingly demonstrate that a uniform population of XaXaXist- cells became a near uniform population of XaXiXist+ cells upon differentiation. Furthermore, only assaying single genes on the X chromosome using RNA FISH, incorrect conclusions regarding cells with erosion versus true biallelism can be easily appreciated. The use of both a population based assay such as RRBS with single cell assays such as RNA FISH allows for a true representation of the heterogeneity in culture from eroded lines.

Given human embryonic stem cells and human induced pluripotent stem cells in the primed state are the most frequently used cell model for understanding mechanisms of early differentiation as well a potential source for regenerative medicine, a realization of this potential for epigenetic drift in culture and mechanisms surrounding this drift are important future questions. Although we did not observe progressive erosion, brief studies on three UCLA lines attempting to induce those phenomena failed to significantly erode over 10 passages (data not shown). Thus specific events in culture may dominate further erosion, or perhaps genetic biases inherent to some lines may drive erosion over a short period of time. However, given that the resulting increased gene expression from eroded regions carries through differentiation, the assessment and consequences of these phenomena should be considered and impact on the investigations using these lines identified.

Experimental Procedures:

Human ES derivation

Human embryos donated from consenting couples through IVF treatment were thawed at or derived to the blastocyst stage under hypoxia as previously reported [25]. hESC lines were derived on inactivated CF1 mouse embryonic fibroblasts feeders and Manual passaging from the derivation culture yielded subclones of which one was chosen based on growth characteristics and banked by the UCLA hESC core. All lines in this study showed normal karyotypes, capacity to produce teratomas with three germ lineages, and expressed markers of pluripotency. hESC lines H1, H7, H9, ESI02 and ESI03 were obtained from the WiCell Research Institutes' WISC Bank.

Human ES culture and differentiation

Human ES lines derived on mitomycin C inactivated CF1 feeder fibroblasts were mef depleted and plated onto hesc-qualified matrigel (BD) plates in mTesr1 (Stemcells) and passaged 2 – 3 times prior to analysis. hESCs were typically passaged using versene (Lonza). Retinoic acid differentiation was done by plating at 50% confluency and switching to media containing 15% FBS in DMEMF12 with 1 uM Retinoic Acid (Tocris) for 10 days. Directed differentiation into hepatocyte like cells [26], cardiomyocyte like cells [27] and neurons [28] were as described previously. Albumin concentrations were measured by ELISA (Bethyl) from one well of a six well after 24 hours of conditioning.

Calcium transients were measured using a custom built contact fluorescence imaging system and voltage and Ca++ signals measured and analyzed as described [29].

RNA FISH and Immunofluorescence

RNA FISH, immunofluorescence and acquisition of images were generated as described previously [4]. The following BACs were obtained from CHORI and used to make RNA FISH probe using the random priming method: Xist (RP11-13M9), Mid1 (RP11-798D20), Cask (Rp11-154-P12, RP11-820G10, and RP11-977L20), Huwe1 (RP11-975N19), Atrx (RP11-1145J4), Gpc3 (RP11-585F14), Mtmr1 (WI2-2809A11 and WI2-922F21), Wdr44 (WI2-2159C5, WI-2 2160O10, WI2-585C17). RNA FISH on blastocysts was performed as described (Nagakawa et al). Antibodies for immunofluorescence used include Nanog (R&D AF1997), cTnI (Abcam ab47003), Pou5F1 (Santa Cruz C10), H3K27Me3 (Millipore 07449), Afp (Sigma SAB3300007), Albumin (Sigma A6684), Tuj1 (Abcam). Secondary antibodies used were Alexa-flours pre-cleared and against the appropriate primary (Life).

Reduced Representation Bisulfite Sequencing

RRBS was performed as described (12). DNA Methylation calling was performed using BS-Seeker2 (30.) using Bowtie (version 0.12.9) to perform read alignments. CpG islands (CGIs) were obtained from UCSC [31] and parsed against the Gencode database to identify associated genes. Preliminary data processing was performed using custom PERL scripts (http://perl.org). Clustering, and heat map generation (Fig 3-3A-C,E) was performed using custom R scripts (http://www.r-project.org). Only sites

covered by at least 5 reads across all samples under consideration were used in an effort to obtain reliable methylation levels. The methylation levels of samples were hierarchically clustered using complete linkage and the Euclidean distance metric. CpG sites that shows constitutive hypermethylation (>= 75%) or hypomethylation (<=15%) across all samples were discarded for presentation purposes. Cell lines were classified into “erosion groups” (Fig 3-3C) based on clustering results with the maximum methylation level at each CpG across all samples within a group taken as the representative methylation level. This is the most conservative assessment of potential demethylation because in order to see low methylation (erosion) within a group, all members of the group have to exhibit it. Erosion hotspots (Fig 3-3D) were identified by first smoothing the data by taking the running mean of methylation levels within windows of 100Kb around each CpG with at least 5X coverage across all samples. CpGs falling within CpG islands were deemed to be eroded if they exhibited a running mean methylation level of < 15%, while showing a hemi-methylated status (20% <= methylation <= 60%) in the reference XaXi group and being hypomethylated in the XY group (methylation <= 20%). Subsequent expression analysis (Fig 3-3G-J) was done at the level of CGIs based on Gencode genes (TSS +/- 2 Kb) that overlapped these regions: Only CGIs harboring at least 5 CpGs with at least 5X coverage across samples were considered and the average methylation level calculated. CGIs were deemed to be eroded if they exhibited an average methylation of < 15%, when the reference XaXi line (ESI03) showed a hemi-methylated status (20% <= average methylation <= 60%) and were hypomethylated (average methylation <= 20%) in the XY reference line (UCLA10). Erosion resistant islands were those whose average methylation was > 15%

across all lines. RRBS data from Bock et. al. [12] filtered for constitutive and hypomethylated regions and defined as hypomethylated, hemi-methylated and fully methylated as before.

RNA-SEQ

For gene expression analysis of the pluripotent stem cells and cardiac differentiated derivative cells (Fig 3-3), a digital barcoding was used to enumerate transcripts using a pool of template switching oligos with random 8-mer barcodes. Paired end reads of 50 bp in length were aligned to the human transcriptome from defined by GenCode version 16 using STAR. Unique barcodes were identified from aligned pairs by counting those sequences within 2 Hamming Distances as non-unique using custom awk scripts. Normalization for read depth was performed using the DE-Seq package after a lower bound of 5 counts were filtered out. Cumulative distribution functions were plotted using R. Statistical significance was tested using various packages including two tailed Kolmogorov-Smirnov tests using R (http://www.r-project.org). For expression analysis of the early passage, subcloned UCLA8 and UCLA9 lines (Fig 3-4), libraries were generated using the standard strand specific Illumina Tru-seq v2 RNA-Seq workflow and sequenced as single end reads of 50 bp.

Gene Expression Microarray analysis

RNA from UCLA lines were extracted and process under feeder free conditions and hybridized to Affymetrix Human Genome U133 plus 2.0 arrays under standard conditions. Array data was processed using R library affy and further normalized using

mas5. For external data, normalization was done in tandem with the UCLA data and after assurance of adequate normalization, expression values compared using cumulative distribution functions (R). Data generated on other arrays such as the HTGene arrays, raw data was normalized to UCLA data using quantile normalization and compared. Agilent arrays were only analyzed within experiments using package limma after QC filters for green and red channels.

SNP/CNV Copy Number Array Analysis

Genomic DNA from pluripotent stem cells and cardiac differentiated derivates were extracted according to standard protocol. Genotyping was done on an Affymetrix GenomeWIde SNP 6.0 arrays analysis as per manufacturers instructions. SNP genotyping was performed using the default parameters on the Affymetrix Genotyping Console. All lines assayed showed correct gender calls and passed QC filters.

FIGURE 3-1A: Human pluripotent stem cells display varied X inactivation patterns

Each bar graph shows the percent of nuclei that are scored at the single cell level with a monoallelic (light grey), biallelic (black) and no-signal (dark grey). Each gene is subject to X inactivation except *UTX*, which escapes and serves a surrogate for number of X chromosomes and Xist (absent in this cohort of cell lines). A) UCLA1, H9 and ESI03 RNA FISH results for allelic expression of X-linked genes are shown as percentages of scored nuclei. Representative RNA FISH Nuclei from two of these lines shows monoallelic expression of *ATRX* but indicates erosion of *GPC3* in UCLA1 B) The XaXe and D) XaXE lines show increased biallelism in some genes. C) The XaXaXist- lines are predominantly biallelic with homogenously high percentages for each gene scored. Representative RNA FISH images show biallelism of X-linked genes subject to inactivation in proximity to X-linked genes subject to escape inactivation (*UTX)*. At least 200 nuclei were scored for each gene, for each line and many lines were scored in replicate experiments (not shown).

Figure 3-1A

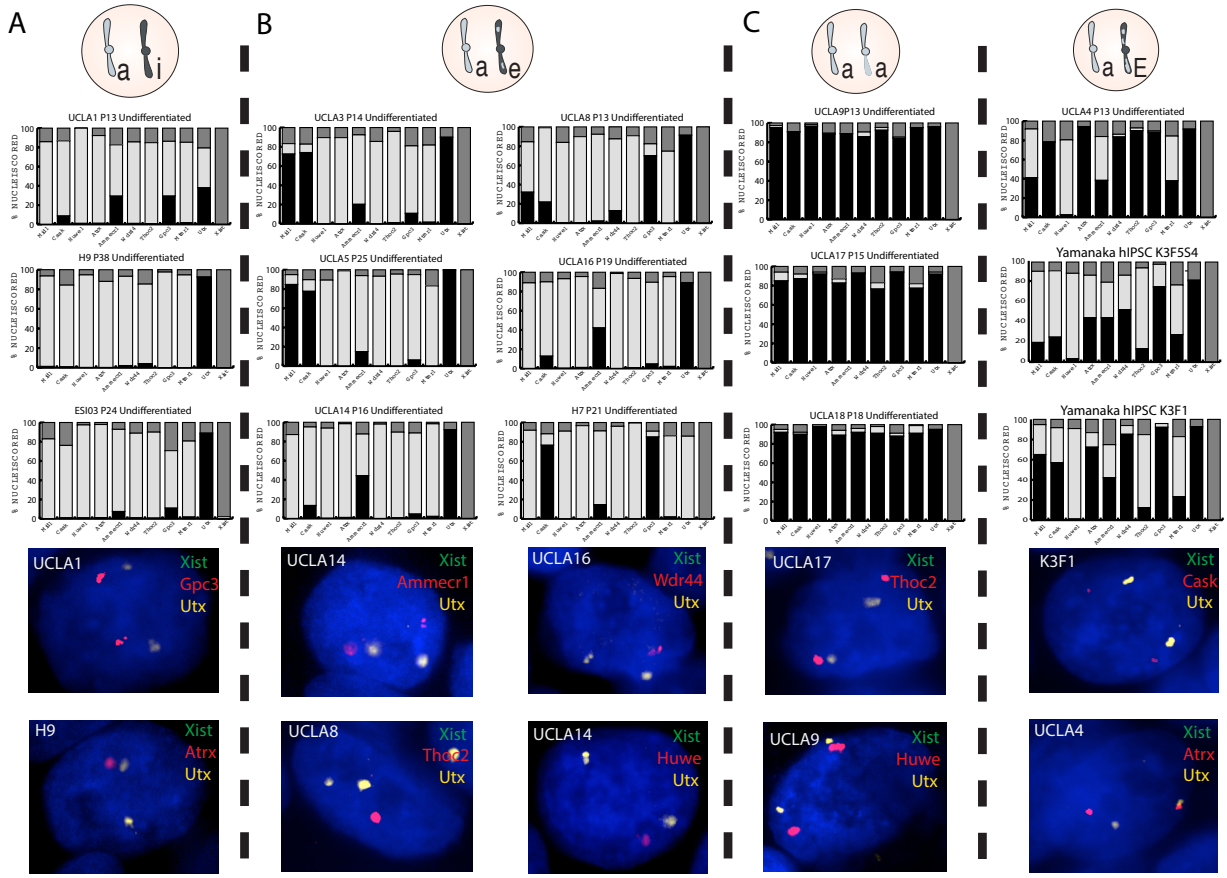
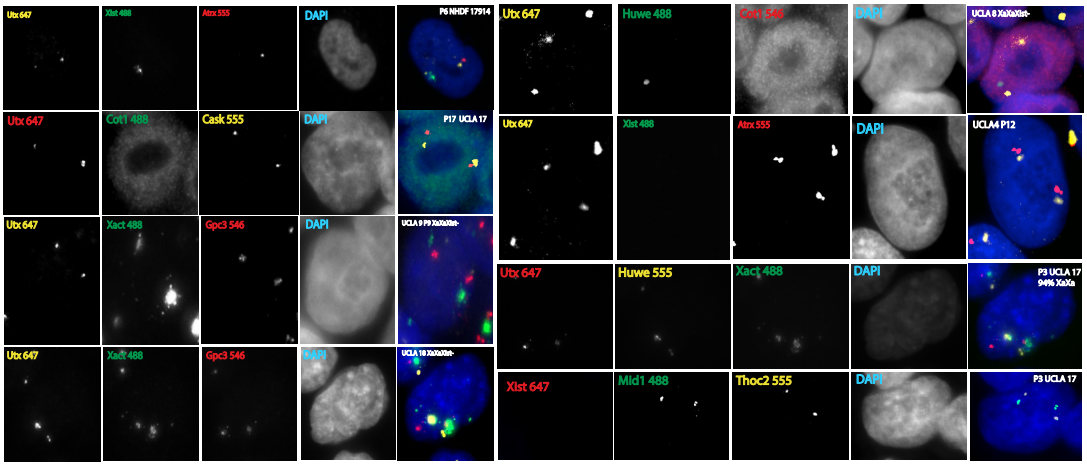

A
B
C
a i
a e
a a
a E
UCLA1 P13 Undifferentiated
UCLA3 P14 Undifferentiated
UCLA8 P13 Undifferentiated
UCLA9P13 Undifferentiated
UCLA4 P13 Undifferentiated
H9 P38 Undifferentiated
UCLA5 P25 Undifferentiated
UCLA16 P19 Undifferentiated
UCLA17 P15 Undifferentiated
Yamanaka hIPSC K3F5S4
ESI03 P24 Undifferentiated
UCLA14 P16 Undifferentiated
H7 P21 Undifferentiated
UCLA18 P18 Undifferentiated
Yamanaka hIPSC K3F1
% NUCLEI SCORED
UCLA1
Xist
Gpc3
Utx
UCLA14
Xist
Ammecr1
Utx
UCLA16
Xist
Wdr44
Utx
UCLA17
Xist
Thoc2
Utx
K3F1
Xist
Cask
Utx
H9
Xist
Atrx
Utx
UCLA8
Xist
Thoc2
Utx
UCLA14
Xist
Huwe
Utx
UCLA9
Xist
Huwe
Utx
UCLA4
Xist
Atrx
Utx

Figure 3-1B: RNA FISH allows for single cell resolution of X inactivation patterns

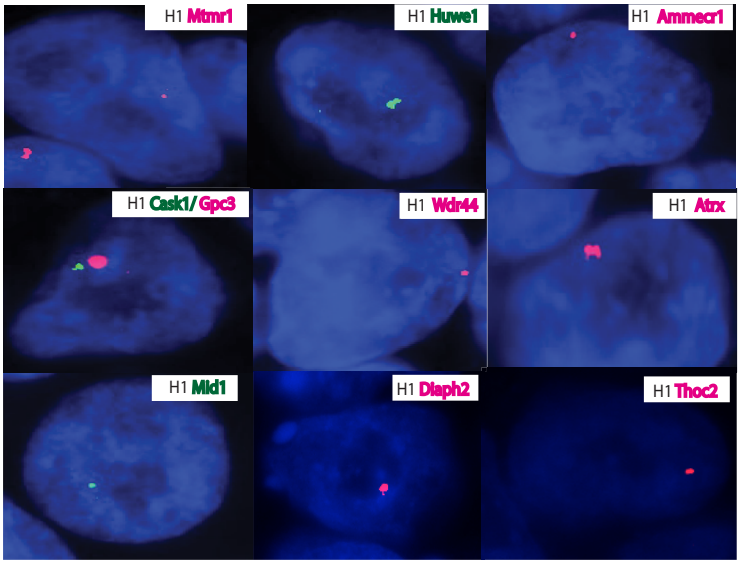
A) Representative RNA FISH images for XaXaXist- lines for genes not depicted in Fig 3-1A showing biallelic gene expression. NHDF17914 female fibroblast showing control RNA FISH image of monoallelic X-linked gene expression in red and a solitary *XIST* cloud in green. B) RNA FISH probes showing monoallelic expression in male hESC line C) Spontaneous differentiation of a biallelic line showing DAPI channel at 5X magnification and then nuclei with RNA FISH signal for *XIST* (absent), *GPC3* and *ATRX* (subject to inactivation, here shown biallelic) and NANOG (Marker of pluripotency) at 63X magnification. Dashed line indicates example pluripotent nuclei, solid line indicates loss of pluripotency.

Fig 3-1B

A


Utx 647
DAPI
P6 NHDF 17914
Utx 647
Huwe 488
DAPI
UCLA 8 XaXaXist-
Utx 647
Cot1 488
Cask 555
DAPI
P17 UCLA 17
Utx 647
Xist 488
Atrx 555
DAPI
UCLA4 P12
Utx 647
Xact 488
Gpc3 546
DAPI
Utx 647
Huwe 555
Xact 488
DAPI
P3 UCLA 17
94% XaXa
Utx 647
Xact 488
Gpc3 546
DAPI
Xist 647
Mid1 488
Thoc2 555
DAPI
P3 UCLA 17

B


H1 Mtmr1
H1 Huwe1
H1 Ammecr1
H1 Cask1/ Gpc3
H1 Wdr44
H1 Atrx
H1 Mid1
H1 Diaph2
H1 Thoc2

C

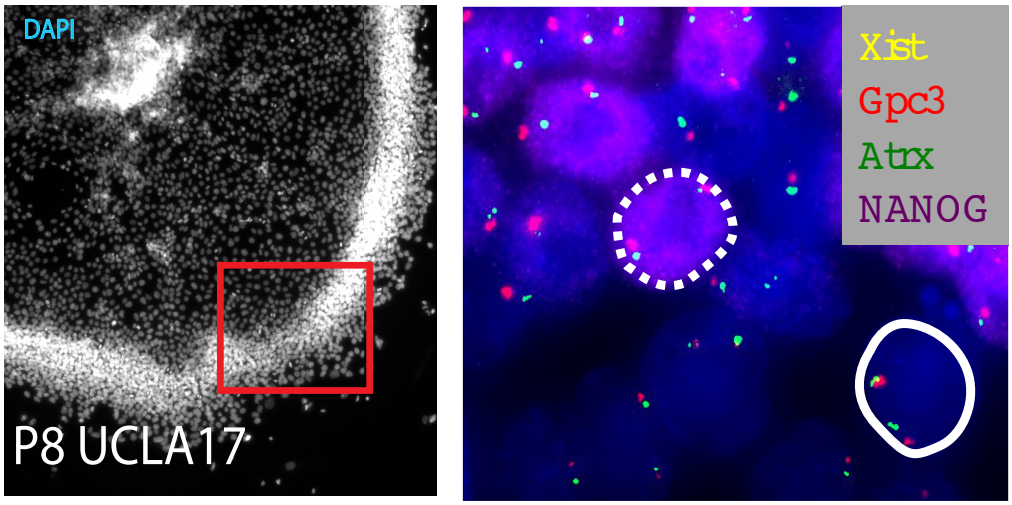

DAPI
P8 UCLA17
Xist
Gpc3
Atrx
NANOG

FIGURE 3-2A: X inactivation patterns established in pluripotency are maintained in differentiation.

A) Allelic expression as a percentage of all nuclei scored on left in pluripotent state and on right after 10 days of retinoic acid induced differentiation. Representative RNA FISH nuclei adjacent to graphs showing presence or absence for the pluripotency marker OCT4 (purple) B) Top: Allelic expression of genes in hepatocyte like cells after completion of protocol and albumin concentrations in ng/mL of each line collected 24 hours after feeding fresh media. Bottom: Allelic expression in cardiomyocyte like cells by RNA FISH. C) Left: Hepatocyte like cell as individual nuclei showing RNA FISH pattern corresponding to state of XCI with ALB in yellow, X-linked gene in red. Lower magnification of same cultures stained for two markers of hepatocyte like cells, ALB in green, AFP in red, *XIST* in purple (absent). Middle: Cardiomyocyte like cells showing striated patterns of cTnI after greater than 2 months of culture from initiation of spontaneous beating. RNA FISH of nuclei show XCI pattern specific to line with X-linked gene *HUWE1* in red. UCLA9 shown as maximum image projection (inset) in right corner to demonstrate biallelism evident in the z-axis. Bottom panels show absence of *XIST. Right*: Neuronal Tuj1 positive (green) cells differentiated from UCLA9 showing biallelic expression of *HUWE1* (red), absence of *XIST* (yellow) and bottom panel showing robust differentiation with no H3K27Me3 (green) foci in TUJ1 positive cells (yellow).

Figure 3-2A

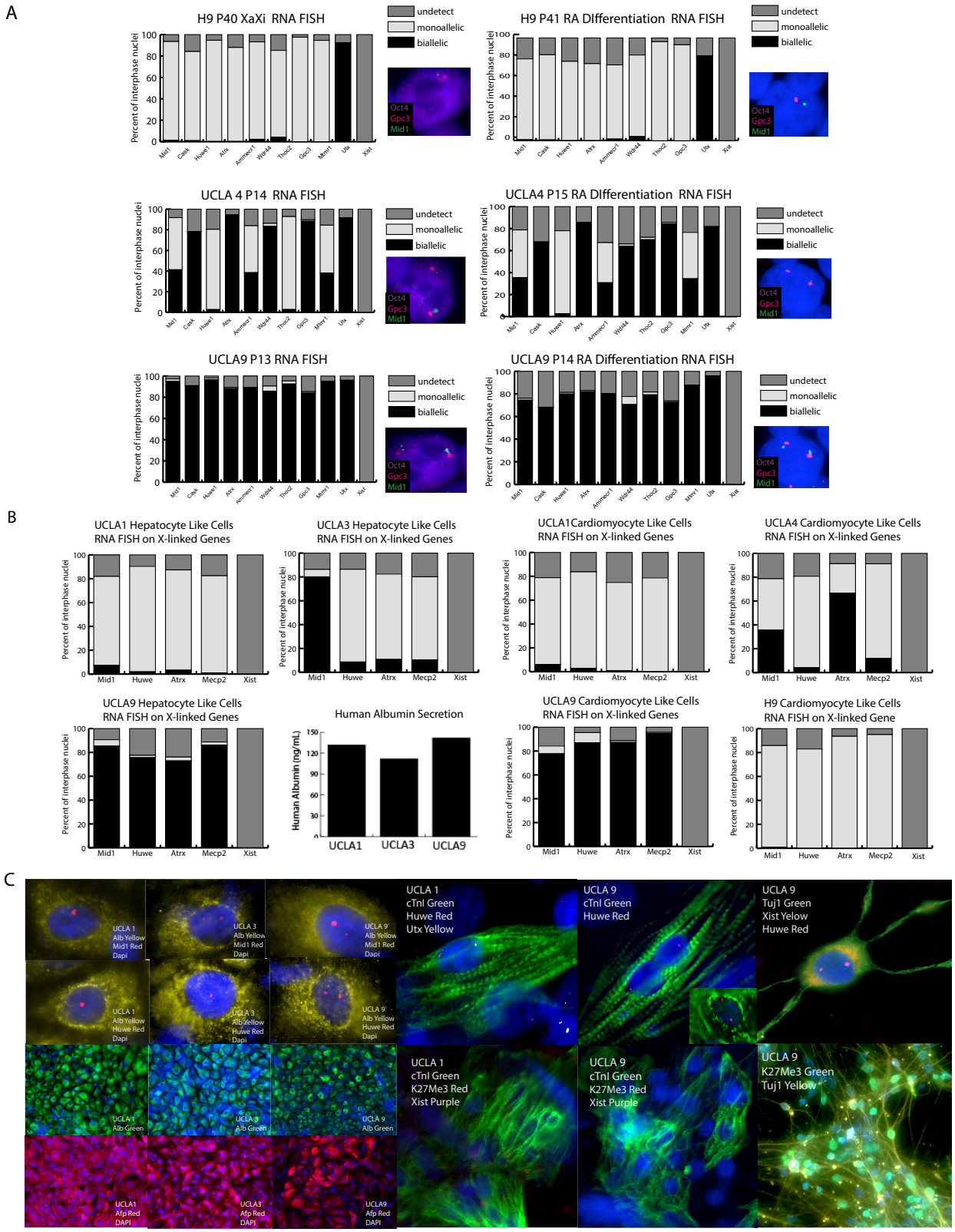

A
H9 P40 XaXi RNA FISH
H9 P41 RA DIfferentiation RNA FISH
UCLA 4 P14 RNA FISH
UCLA4 P15 RA DIfferentiation RNA FISH
UCLA9 P13 RNA FISH
UCLA9 P14 RA Differentiation RNA FISH
undetect
monoallelic
biallelic
Percent of interphase nuclei
Oct4
Gpc3
Mid1
B
UCLA1 Hepatocyte Like Cells
RNA FISH on X-linked Genes
UCLA3 Hepatocyte Like Cells
RNA FISH on X-linked Genes
UCLA1Cardiomyocyte Like Cells
RNA FISH on X-linked Genes
UCLA4 Cardiomyocyte Like Cells
RNA FISH on X-linked Genes
UCLA9 Hepatocyte Like Cells
RNA FISH on X-linked Genes
Human Albumin Secretion
Human Albumin (ng/mL)
UCLA1
UCLA3
UCLA9
UCLA9 Cardiomyocyte Like Cells
RNA FISH on X-linked Genes
H9 Cardiomyocyte Like Cells
RNA FISH on X-linked Gene
Mid1
Huwe
Atrx
Mecp2
Xist
C
UCLA 1
Alb Yellow
Mid1 Red
Dapi
UCLA 3
Alb Yellow
Mid1 Red
Dapi
UCLA 9'
Alb Yellow
Mid1 Red
Dapi
UCLA 1
Alb Yellow
Huwe Red
Dapi
UCLA 3
Alb Yellow
Huwe Red
Dapi
UCLA 9'
Alb Yellow
Huwe Red
Dapi
UCLA 1
Alb Green
UCLA 3
Alb Green
UCLA 9
Alb Green
UCLA1
Afp Red
DAPI
UCLA3
Afp Red
DAPI
UCLA9
Afp Red
DAPI
UCLA 1
cTnI Green
Huwe Red
Utx Yellow
UCLA 9
cTnI Green
Huwe Red
UCLA 9
Tuj1 Green
Xist Yelow
Huwe Red
UCLA 1
cTnI Green
K27Me3 Red
Xist Purple
UCLA 9
cTnI Green
K27Me3 Red
Xist Purple
UCLA 9
K27Me3 Green
Tuj1 Yellow

Figure 3-2B: The pattern of XCI is maintained in differentiation through retinoic acid or directed into germ lineage specific cell types.

A) RNA FISH for genes on the X chromosome for the rest of the lines in the study showing allelic expression as percentage of nuclei scored. B) H3K27Me3 immunofluorescence (green) stains for representative RA induced differentiation nuclei. C) Immunofluorescence and RNA FISH for H3K27Me3 (green) in albumin (yellow) positive cells of hepatocyte like cells differentiation showing no foci and absence of *XIST* (purple). D) Beat frequency at baseline and after addition of 1 uM isoproterenol in XaXa UCLA9 and control H9 derived cardiomyocyte like cells E) Calcium transient waveforms of the same lines to match supplementary movie 1.

Figure 3-2B

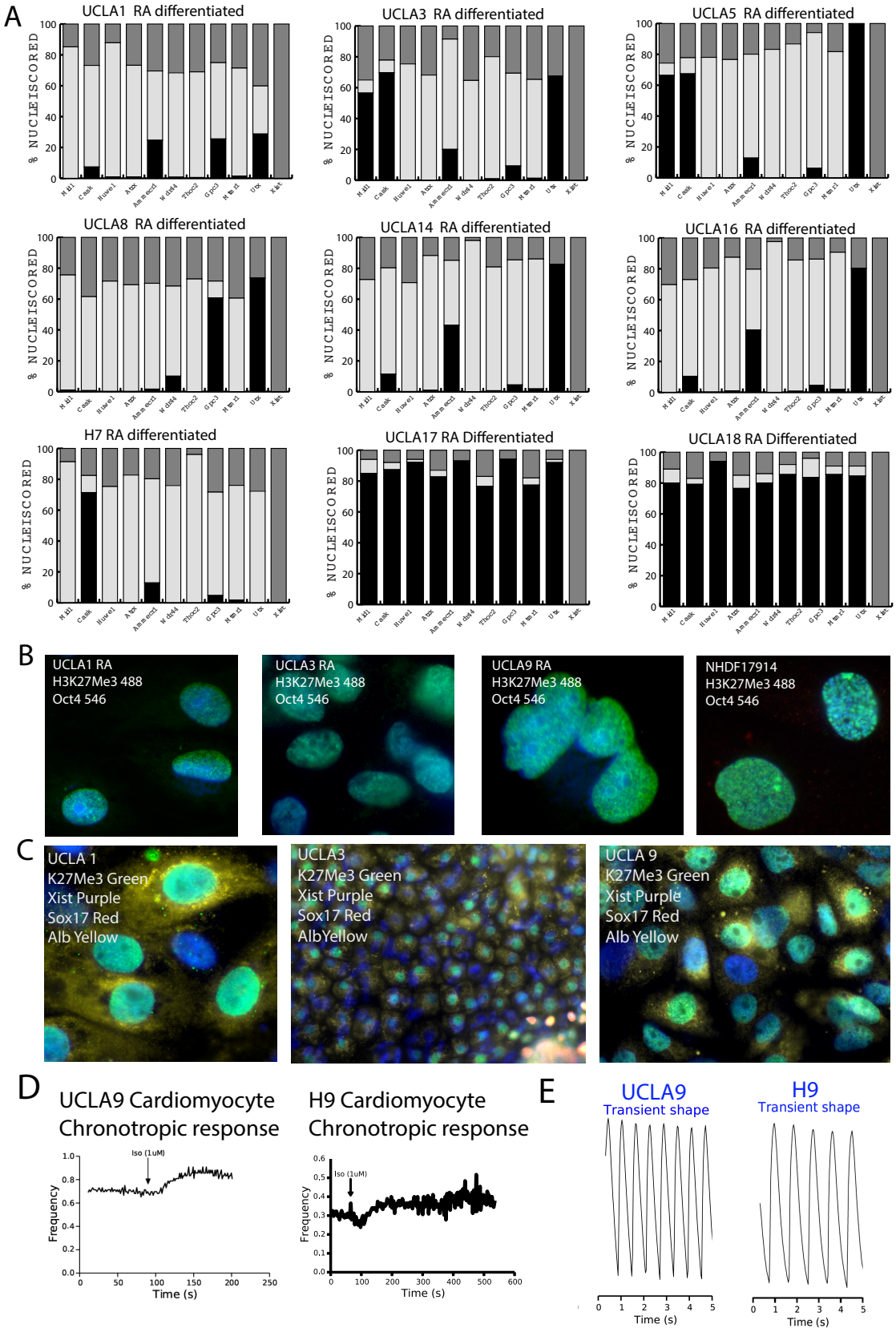
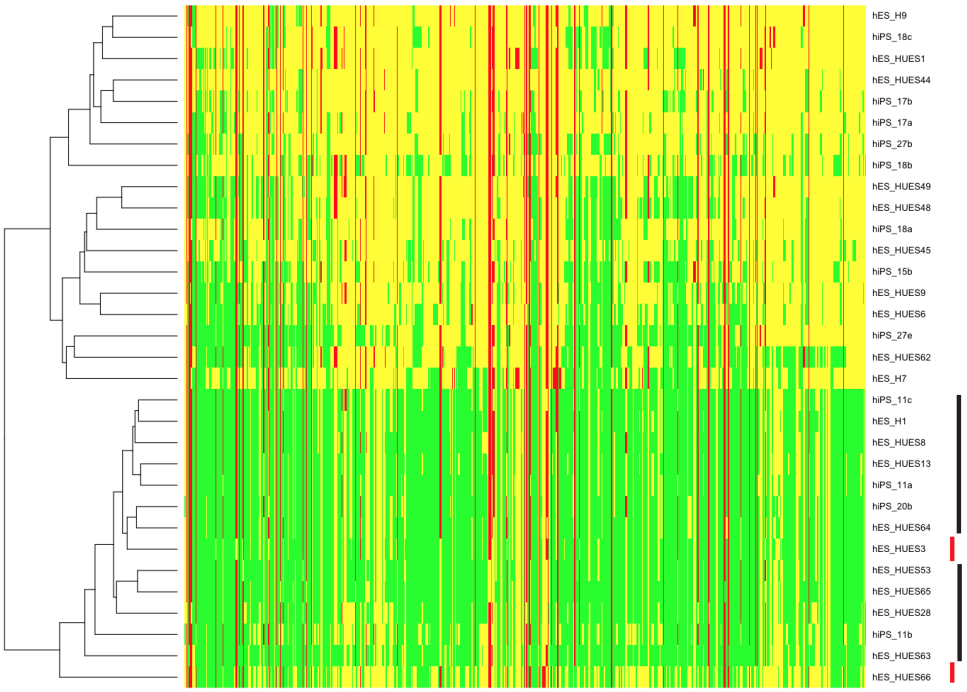

A
UCLA1 RA differentiated
UCLA3 RA differentiated
UCLA5 RA differentiated
UCLA8 RA differentiated
UCLA14 RA differentiated
UCLA16 RA differentiated
H7 RA differentiated
UCLA17 RA Differentiated
UCLA18 RA Differentiated
% NUCLEISCORED
B
UCLA1 RA
H3K27Me3 488
Oct4 546
UCLA3 RA
H3K27Me3 488
Oct4 546
UCLA9 RA
H3K27Me3 488
Oct4 546
NHDF17914
H3K27Me3 488
Oct4 546
C
UCLA 1
K27Me3 Green
Xist Purple
Sox17 Red
Alb Yellow
UCLA3
K27Me3 Green
Xist Purple
Sox17 Red
AlbYellow
UCLA 9
K27Me3 Green
Xist Purple
Sox17 Red
Alb Yellow
D
UCLA9 Cardiomyocyte
Chronotropic response
H9 Cardiomyocyte
Chronotropic response
Iso (1uM)
Frequency
Time (s)
E
UCLA9
Transient shape
H9
Transient shape
Time (s)

Figure 3-2C: Differentiation propensities do not correlate with the degree of erosion in female human pluripotent stem cells.

A) Heatmap demonstrating the degree and distribution of erosion in female lines across the X chromosome from a large study identifying lineage propensities of human pluripotent stem cells lines. (Green is demethylated, Yellow hemi-methylated and Red constitutively methylated, 888 cytosines passed filter on 32 samples). The black cluster represent male pluripotent stem cells and lack of significant DNA methylation on the solitary active X chromosome. The orange clusters represent the most highly eroded lines in the study for which no significant pattern of differentiation bias could be identified. B) Scatterplot of erosion as median methylation on the X chromosome and lineage score reported. Correlation determined using Spearman's rho. Three lines consistently showed different behavior and corresponding dots labeled.

Fig 3-2C

A


hES_H9
hiPS_18c
hES_HUES1
hES_HUES44
hiPS_17b
hiPS_17a
hiPS_27b
hiPS_18b
hES_HUES49
hES_HUES48
hiPS_18a
hES_HUES45
hiPS_15b
hES_HUES9
hES_HUES6
hiPS_27e
hES_HUES62
hES_H7
hiPS_11c
hES_H1
hES_HUES8
hES_HUES13
hiPS_11a
hiPS_20b
hES_HUES64
hES_HUES3
hES_HUES53
hES_HUES65
hES_HUES28
hiPS_11b
hES_HUES63
hES_HUES66

B

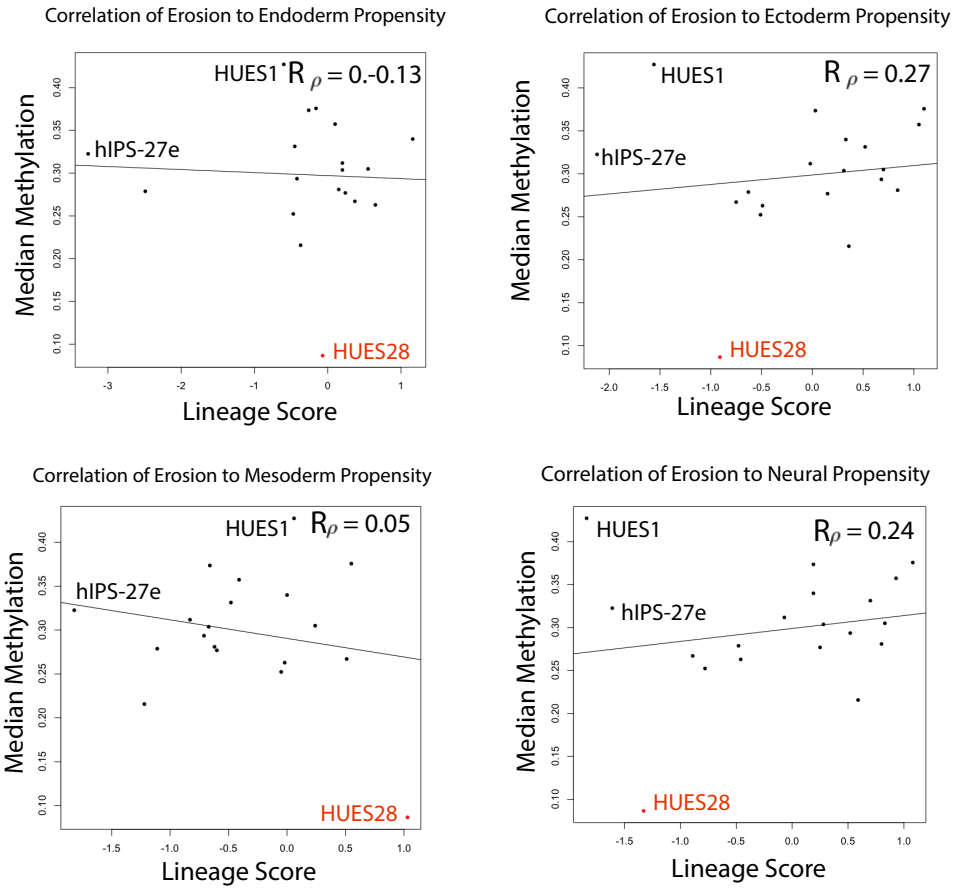
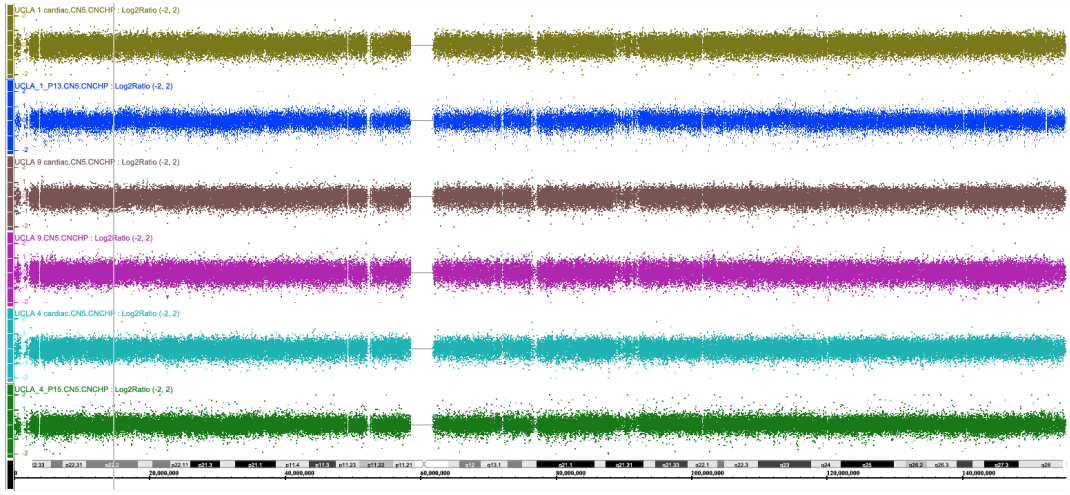

Correlation of Erosion to Endoderm Propensity
HUES1
R ρ = 0.-0.13
hIPS-27e
HUES28
Median Methylation
Lineage Score
Correlation of Erosion to Ectoderm Propensity
HUES1
R ρ = 0.27
hIPS-27e
HUES28
Median Methylation
Lineage Score
Correlation of Erosion to Mesoderm Propensity
HUES1
Rρ = 0.05
hIPS-27e
HUES28
Median Methylation
Lineage Score
Correlation of Erosion to Neural Propensity
HUES1
Rρ = 0.24
hIPS-27e
HUES28
Median Methylation
Lineage Score

Figure 3-2D: Genomewide CNV analysis does not show abnormal karyotypes in cardiomyocyte differentiated cells

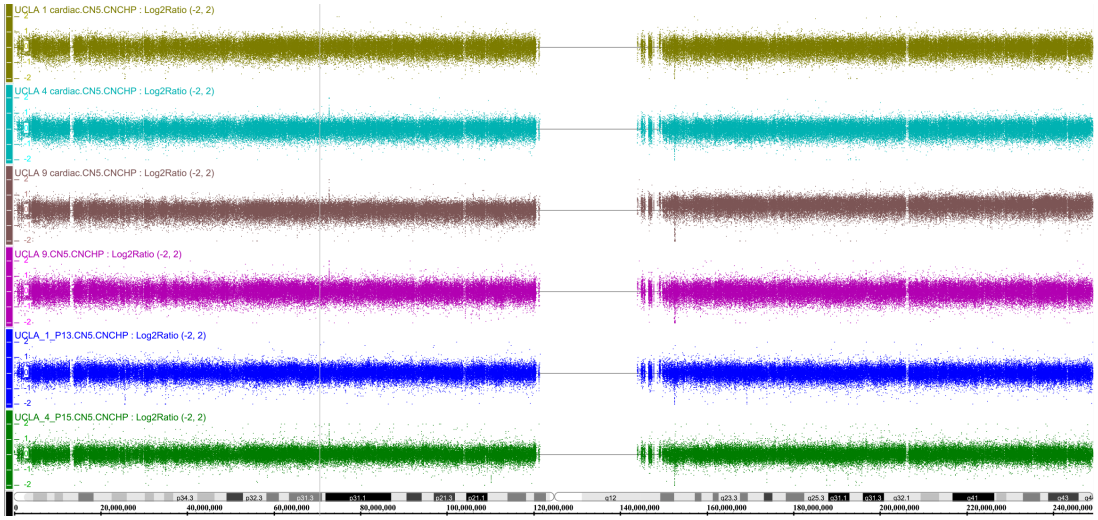
Genomewide arrays for CNV were used to identify if karyotypic abnormalities on the X chromosome could explain the biallelic pattern seen by RNA FISH. Each chromosome is shown as a log2 ratio of array signal indicating diploid state. UCLA 1 cardiac (amber line) is XaXiXist-, UCLA4 cardiac (turquoise line) is XaXEXist-, and UCLA 9 (brown line) is XaXa Xist-. The parental lines are UCLA 9 (purple), UCLA 1 (blue) and UCLA 4 (green).

Fig 3-2D

Chromosome X



Chromosome 1



Chromosome 12

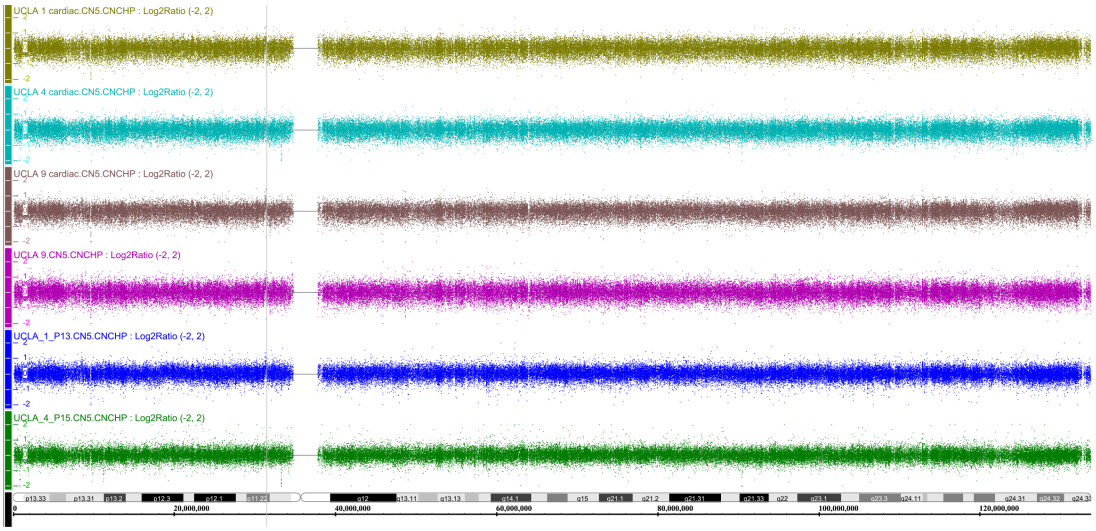


Figure 3-3A: Genomic approaches to assaying for X inactivation patterns reveal hotspots of erosion in the XaXeXist- lines as well as a loss of dosage compensation for both XaXe and XaXa lines.

A) Reduced representation bisulfite sequencing reveals hotspots of DNA demethylation (green) in regions on the short and long arm of the chromosome amongst the XaXeXist- lines (brown cluster). The XaXiXist- lines show hemi-methylation (yellow) along the X chromosome (blue cluster). The XaXaXist- lines show no DNA methylation in these same regions and mirror the solitary X chromosome found in male human ES cells (green cluster). The XaXE line UCLA4 shows that maintenance of DNA methylation is enriched in areas near the centromere and telomere (orange cluster). B) The autosomal genes show clustering by cell state rather than by line suggesting that the X chromosome confers its DNA methylation pattern upon differentiation but this is not due to inadequate differentiation. C) The hESC lines could be grouped into five categories resembling the RNA FISH interpretations and the pattern of XCI on the chromosome is demonstrated from the raw signal of DNA methylation for representative lines (D). E) RNA-seq for X-linked genes show that although the distribution of gene expression is roughly the same after normalization the median is higher in XaXE lines (UCLA4) and XaXa lines (UCLA9). The X:A ratio of the various lines show that XaXa and XaXE have lost dosage compensation relative to the expression of the autosomes. F) The distribution of X-linked genes is higher in XaXe and XaXa lines at all levels of expression implying the increased gene expression is a result of global derepression and not focal hyperactivity of a few genes. G) XaXe lines show gene expression levels non-signficantly different from XaXa (grey) lines in the eroded regions (red) and non-

signficantly different from XaXi (blue) lines in the inactivated regions (black). H) The gene expression of cardiac differentiation also shows a global increase in XaXa and XaXe lines with eroded regions following the same pattern as in the pluripotent state (I). The X:A ratios in the cardiac lines imply the biallelic expression from the XaXa and XaXe is not compensated by a reduced expression on the autosomes and thus dosage compensation is not required for differentiation in vitro. K) Log2 ratios of pluripotent versus cardiac derived cell lines for pluripotent specific (blue) and cardiac specific (red) genes.

Fig 3-3A

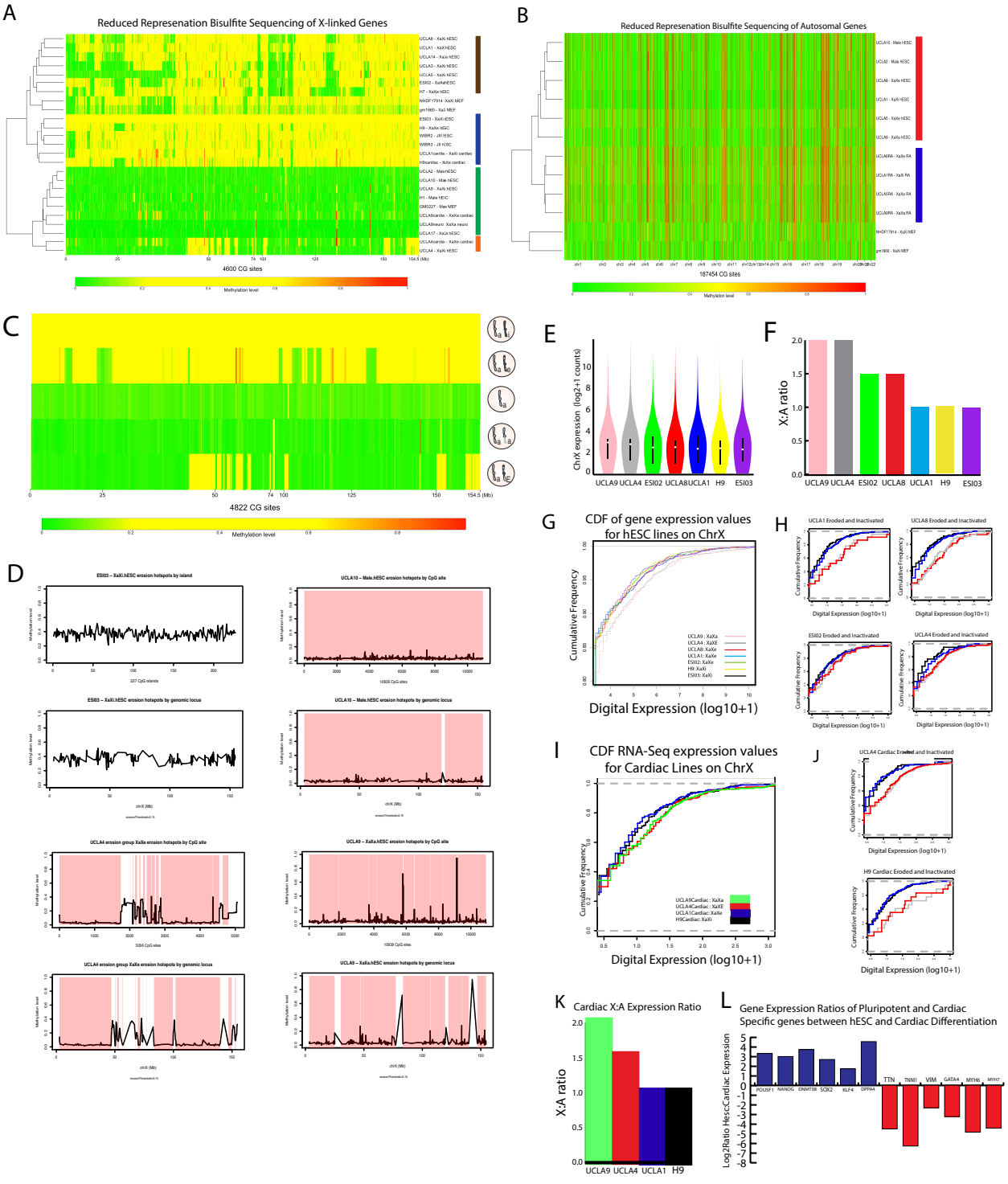

A
Reduced Represenation Bisulfite Sequencing of X-linked Genes
4600 CG sites
Methylation level
B
Reduced Represenation Bisulfite Sequencing of Autosomal Genes
187454 CG sites
Methylation level
C
4822 CG sites
Methylation level
D
E
ChrX expression (log2+1 counts)
UCLA9 UCLA4 ESI02 UCLA8 UCLA1 H9 ESI03
F
X:A ratio
UCLA9 UCLA4 ESI02 UCLA8 UCLA1 H9 ESI03
G
CDF of gene expression values for hESC lines on ChrX
Cumulative Frequency
Digital Expression (log10+1)
H
UCLA1 Eroded and Inactivated
UCLA8 Eroded and Inactivated
ESI02 Eroded and Inactivated
UCLA4 Eroded and Inactivated
Cumulative Frequency
Digital Expression (log10+1)
I
CDF RNA-Seq expression values for Cardiac Lines on ChrX
Cumulative Frequency
Digital Expression (log10+1)
J
UCLA4 Cardiac Eroded and Inactivated
H9 Cardiac Eroded and Inactivated
Cumulative Frequency
Digital Expression (log10+1)
K
Cardiac X:A Expression Ratio
X:A ratio
UCLA9 UCLA4 UCLA1 H9
L
Gene Expression Ratios of Pluripotent and Cardiac Specific genes between hESC and Cardiac Differentiation
Log2Ratio HescCardiac Expression
POU5F1 NANOG DNMT3B SOX2 KLF4 DPPA4 TTN TNNI1 VIM GATA4 MYH6 MYH7

Figure 3-3B: Expression from the X chromosome using cumulative distribution frequencies demonstrate that no published pre-XCI line has the pattern of an XaXa human ES line.

A) Cumulative distribution functions for the autosomal gene expression from our human ES and human ES derived cardiomyocyte like cells show a non-significant difference between the lines implying adequate normalization after RNA-Seq. B) Naïve human iPS generated from the Jaenisch lab in 2010 show no evidence of biallelism but rather severe erosion. All lines published are depicted in green with one line (C1 naïve) expressing at a higher level than others. Our XaXi (blue) and XaXa (red) lines serve as a metric to gauge the level of reactivation. Furthering the conclusion the C1 naïve line is merely eroded the expression in regions that are resistant to erosion but biallelically active only in the XaXa lines are shown in right panel. Here the C1 naïve line is no different from the other lines also indicating this is not a true XaXa line but merely severely eroded. C) New naïve lines generated in the Hanna lab in 2013 show mostly XaXi pattern (green lines) however the C1 naïve line approaches (blue) is indistinguishable from our most severely eroded line (grey). The red and pink lines indicate the expression level of our XaXa lines and show that these are not biallelic by expression. Two lines naïve embryonic stem cells lines derived from blastocysts were examined by RNA FISH and both are indicative of an XCI state common to primed pluripotency i) the male line shows monoallelic X linked gene expression with no Xist and ii) the female line shows predominantly monoallelic X linked gene expression with a solitary Xist cloud competently silencing that chromosome. D) The lines WIBR2 and WIBR3 grown in hypoxia show no evidence of an XaXa state by RNA FISH or by the

expression data published. E) Patterns of erosion can be seen in our lines from mild (UCLA 1) to severe (UCLA4), the maintenance of pericentromeric and peritelomeric methylation (dashed lines) distinguishes human pluripotent stem cell lines as eroded when compared to the XaXa lines. The human iPS lines derived on mouse LIF expressing fibroblasts and argued to be XaXa from the Yamanaka group show a pattern consistent with severe erosion.

Figure 3-3B

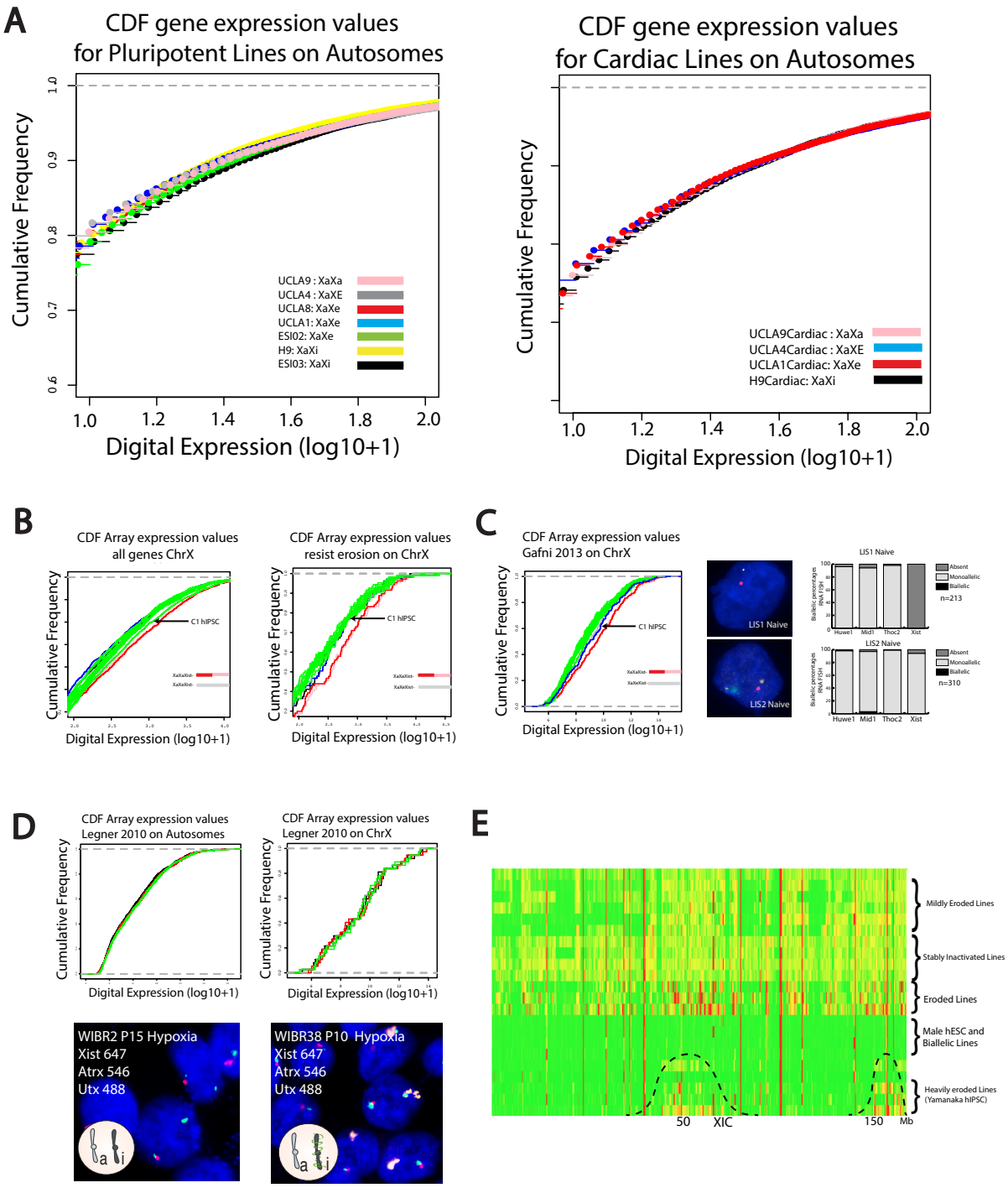

A
CDF gene expression values
for Pluripotent Lines on Autosomes
Cumulative Frequency
Digital Expression (log10+1)
UCLA9 : XaXa
UCLA4 : XaXE
UCLA8: XaXe
UCLA1: XaXe
ESI02: XaXe
H9: XaXi
ESI03: XaXi
CDF gene expression values
for Cardiac Lines on Autosomes
Cumulative Frequency
Digital Expression (log10+1)
UCLA9Cardiac : XaXa
UCLA4Cardiac : XaXE
UCLA1Cardiac: XaXe
H9Cardiac: XaXi
B
CDF Array expression values
all genes ChrX
CDF Array expression values
resist erosion on ChrX
Cumulative Frequency
Digital Expression (log10+1)
C1 hIPSC
C
CDF Array expression values
Gafni 2013 on ChrX
LIS1 Naive
LIS2 Naive
Absent
Monoallelic
Biallelic
n=213
n=310
Huwe1
Mid1
Thoc2
Xist
D
CDF Array expression values
Legner 2010 on Autosomes
CDF Array expression values
Legner 2010 on ChrX
WIBR2 P15 Hypoxia
Xist 647
Atrx 546
Utx 488
WIBR38 P10 Hypoxia
Xist 647
Atrx 546
Utx 488
E
Mildly Eroded Lines
Stably Inactivated Lines
Eroded Lines
Male hESC and
Biallelic Lines
Heavily eroded Lines
(Yamanaka hIPSC)
50
XIC
150
Mb

Figure 3-4: Tracing the progression of Xist loss during early passage of human ES cells demonstrated that Xist is loss stochastically over time and that XaXa derivation cultures are a mix of XaXa and XaXiXist+ cells

A) The loss of Xist over time is stochastic with respect to kinetics between lines starting as XaXaXist+. B) RNA-seq of expression from the X chromosome shows that subcloned lines from an Xist positive cell line has an expression along the X chromosome that is associated with the degree of loss of Xist in the population suggesting erosion begins to occur quickly after Xist loss. C) The stability of the XaXa state is seen in there lines derived at UCLA however at early passage Xist positive cells were seen in the culture. D) Our procedure at UCLA is to subclone through manual passage portions of the derivation culture and thus we are able to generate pure populations of XaXaXist- lines from mixed cultures arising from the derivation of a single human blastocyst. The white dotted lines represent the area that was use to derive the line and the adjacent regions were pure XaXaXist-. E) The encircled red regions shows an example patch of H3K27Me3 positive cells that was also Xist positive and away from the region used to derive UCLA18.

Figure 3-4

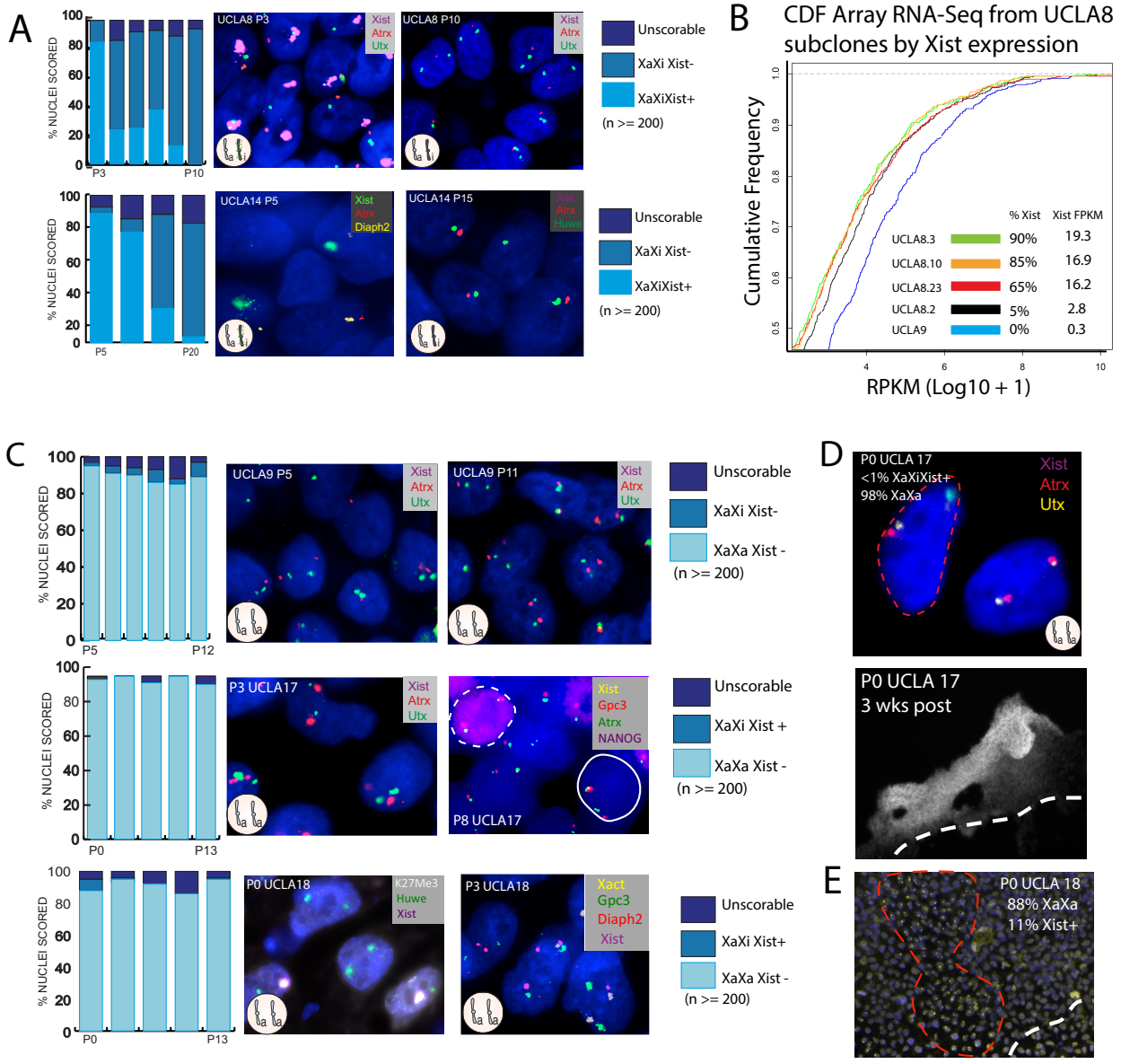
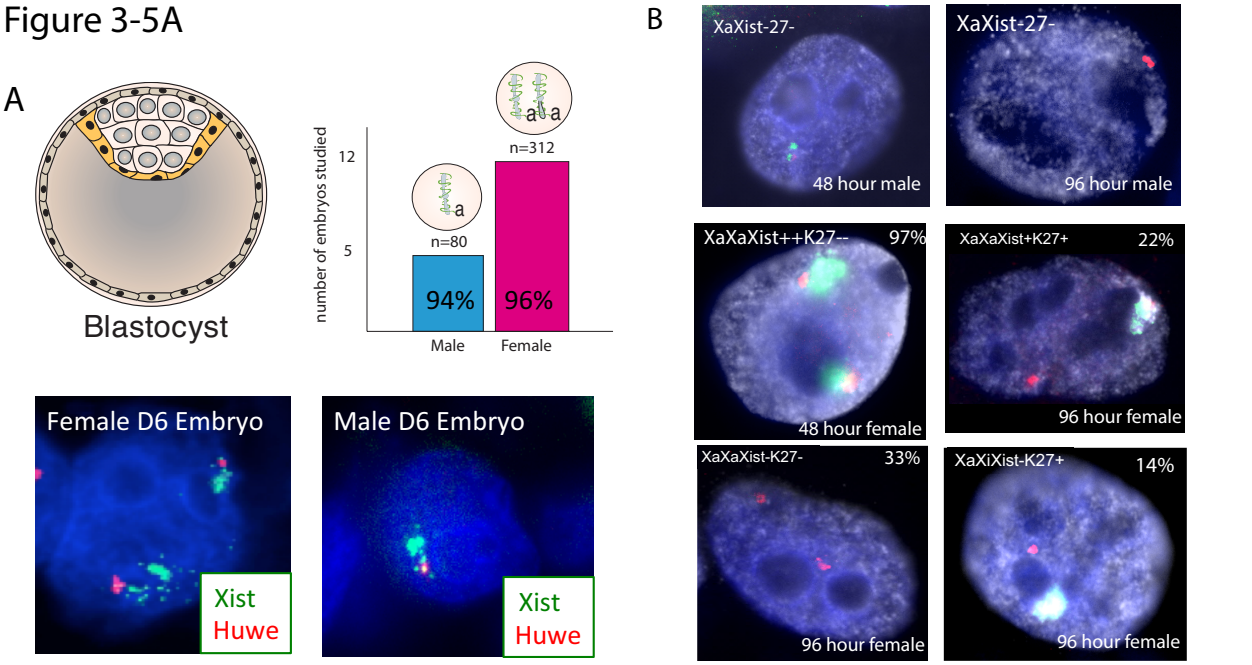
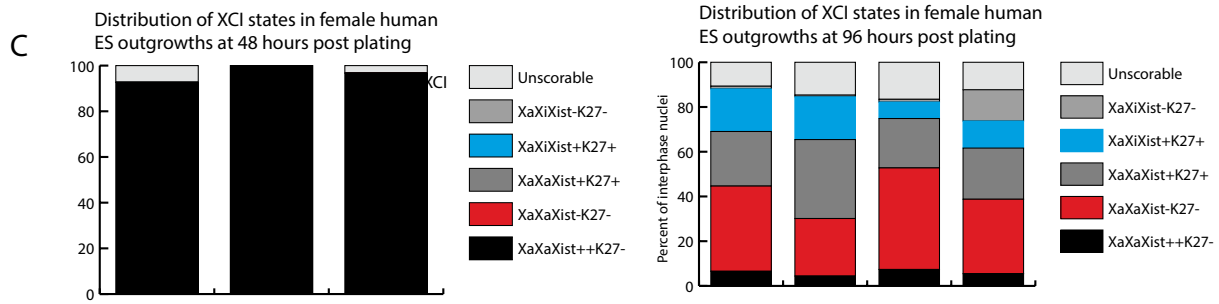

A
% NUCLEI SCORED
P3
P10
UCLA8 P3
Xist
Atrx
Utx
UCLA8 P10
Unscorable
XaXi Xist-
XaXiXist+
(n >= 200)
P5
P20
UCLA14 P5
Diaph2
UCLA14 P15
B
CDF Array RNA-Seq from UCLA8 subclones by Xist expression
Cumulative Frequency
RPKM (Log10 + 1)
% Xist
Xist FPKM
UCLA8.3
90%
19.3
UCLA8.10
85%
16.9
UCLA8.23
65%
16.2
UCLA8.2
5%
2.8
UCLA9
0%
0.3
C
P5
P12
UCLA9 P5
UCLA9 P11
XaXa Xist -
P0
P13
P3 UCLA17
Gpc3
Atrx
NANOG
P8 UCLA17
XaXi Xist +
P0 UCLA18
K27Me3
Huwe
P3 UCLA18
Xact
Gpc3
Diaph2
XaXa Xist+
D
P0 UCLA 17
<1% XaXiXist+
98% XaXa
3 wks post
E
P0 UCLA 18
88% XaXa
11% Xist+

Figure 3-5A: The human pre-implantation blastocyst is XaXaXist++ in female embryos and XaXist+ in male embryos however early in the course of derivation of an ES line, X-inactivation occurs and is subject to aborted products producing XaXaXist- cells.

A) The majority of nuclei in human pre-implantation embryos show an Xist+ signal with an active X chromosome representing an XCI pattern never seen in human ES nor in mouse embryos. B) The pattern of X inactivation arises in derivation cultures as they progress from the naïve state of the human blastocyst to the primed state of human ES. C) In this experiment XCI had not occurred by 48 hours of plating however by 96 hours multiple forms of XCI were found with a substantial portion showing the XaXaXist- state. D) The model of human ES derivation we put forth shows XCI occurring in the majority of human ES derivations and in a few cases XaXa cells are produced which can give rise to class we human ES cells. Only the pre-implantation blastocyst XCI pattern is competent to undergo x inactivation and the XaXaXist- state is simply an intermediate of this process. It is unclear if the XaXaXist- cells are a side product or part of usual sequence of human XCI.

Figure 3-5A


A
Blastocyst
number of embryos studied
12
5
n=312
n=80
94%
96%
Male
Female
Female D6 Embryo
Xist
Huwe
Male D6 Embryo
Xist
Huwe
B
XaXist-27-
48 hour male
XaXist-27-
96 hour male
XaXaXist++K27--
97%
48 hour female
XaXaXist+K27+
22%
96 hour female
XaXaXist-K27-
33%
96 hour female
XaXiXist-K27+
14%
96 hour female


C
Distribution of XCI states in female human
ES outgrowths at 48 hours post plating
100
80
60
40
20
0
Unscorable
XaXiXist-K27-
XaXiXist+K27+
XaXaXist+K27+
XaXaXist-K27-
XaXaXist++K27-
Distribution of XCI states in female human
ES outgrowths at 96 hours post plating
Percent of interphase nuclei
100
80
60
40
20
0
Unscorable
XaXiXist-K27-
XaXiXist+K27+
XaXaXist+K27+
XaXaXist-K27-
XaXaXist++K27-

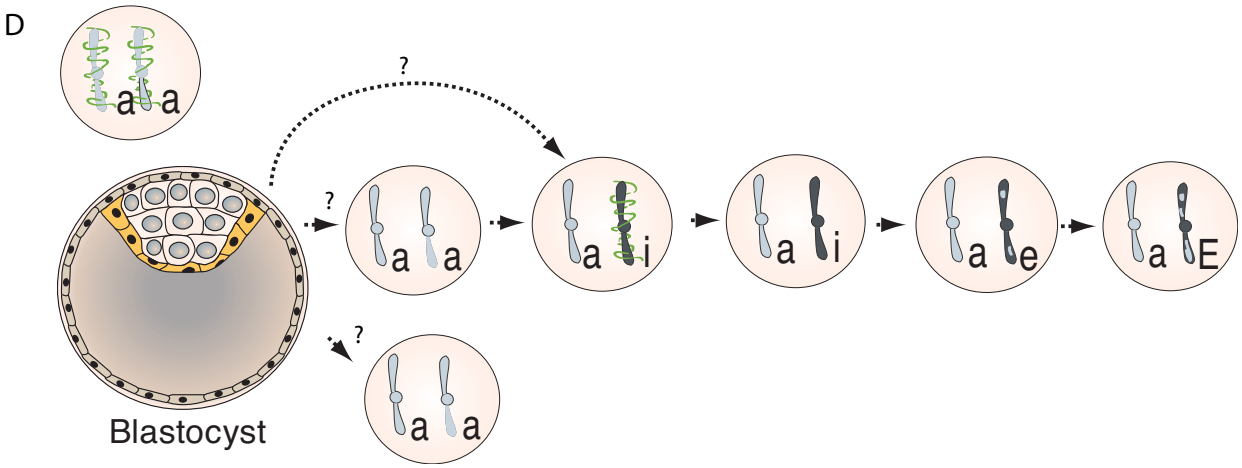
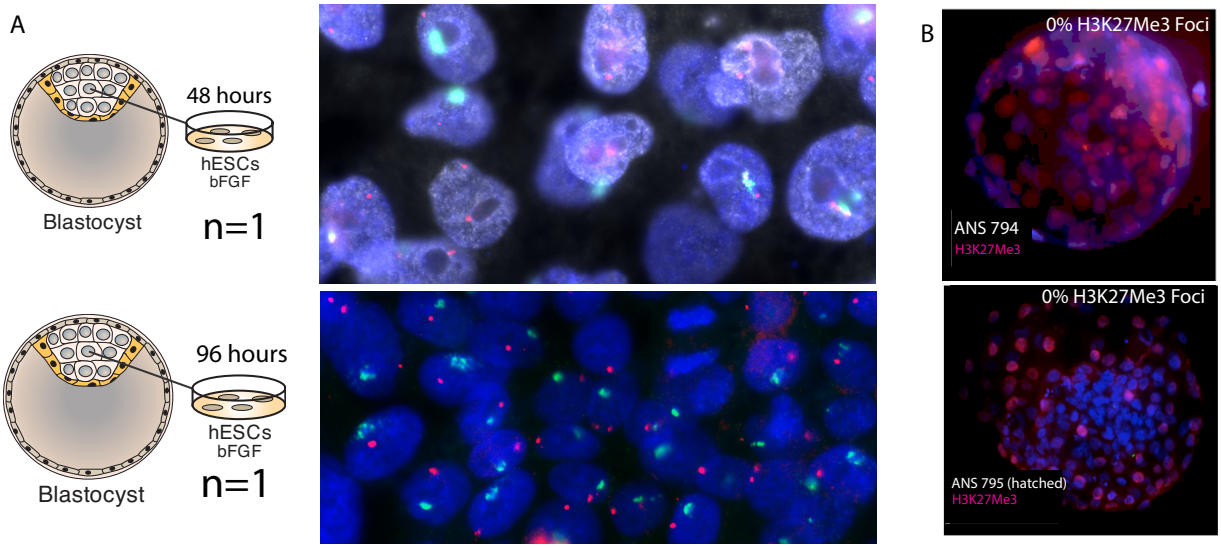
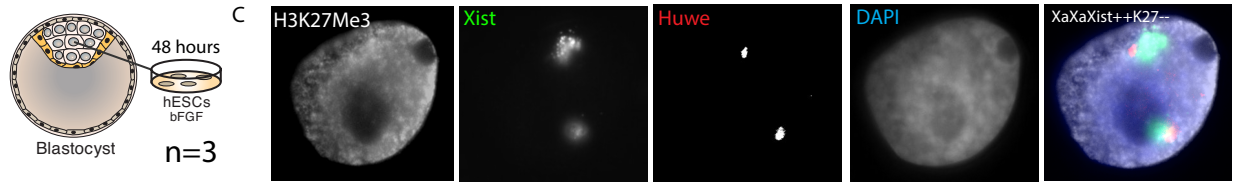

D
a a
?
?
a a
a i
a i
a e
a E
?
a a
Blastocyst

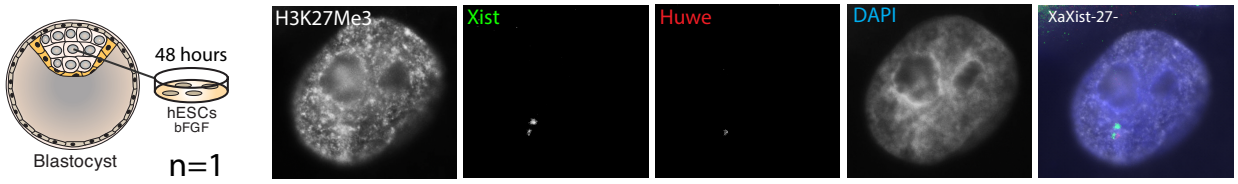
Figure 3-5B: Human ES derivation results in stochastic kinetics of XCI during cell fate changes.

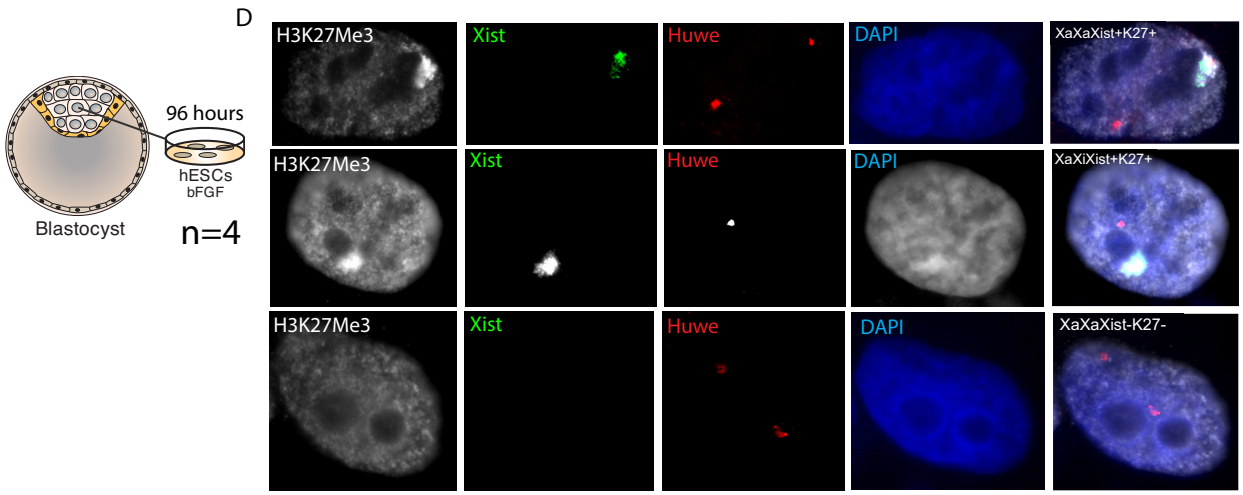
A) This experiment showed a dramatic increase in XCI patterns at the 48 hours mark and had largely completed by the 96 hour mark implying XCI occurs upon derivation however is subject to different kinetics. B) The pre-implantation blastocyst does not have H3K27Me3 foci (purple) within intact (top) or hatched embryos (bottom). C) Nuclei by channel for examples of pre-XCI patterns of Xist, X-linked gene expression (Huwe) and H3K27Me3 and D) post-XCI patterns.

Fig 3-5B


A
48 hours
hESCs
bFGF
Blastocyst
n=1
96 hours
hESCs
bFGF
Blastocyst
n=1
B
0% H3K27Me3 Foci
ANS 794
H3K27Me3
0% H3K27Me3 Foci
ANS 795 (hatched)
H3K27Me3


C
48 hours
hESCs
bFGF
Blastocyst
n=3
H3K27Me3
Xist
Huwe
DAPI
XaXaXist++K27--


48 hours
hESCs
bFGF
Blastocyst
n=1
H3K27Me3
Xist
Huwe
DAPI
XaXist-27-


D
96 hours
hESCs
bFGF
Blastocyst
n=4
H3K27Me3
Xist
Huwe
DAPI
XaXaXist+K27+
H3K27Me3
Xist
Huwe
DAPI
XaXiXist+K27+
H3K27Me3
Xist
Huwe
DAPI
XaXaXist-K27-

Movie 1: Cardiomyocyte like cell lines are responsive to beta-adrenergic stimulation

Phase movies of UCLA4, H9 and UCLA9 showing baseline spontaneous beating and an expected increase in frequency after addition of 1 uM isoproterenol.

Movie 3-2: Female pre-implantation Blastocysts have uniform patterns of XCI

Z-stack movie of Female Blastocyst after RNA FISH for Xist (green) and X-linked gene (HUWE1). Autoflourescent debris seen with off-color orange spots typical of human pre-implantation blastocysts. Three dimensional projection of one nuclei demonstrating Xist expression is not limited to site of transcription consistent with spreading but X-linked gene expression maintained consistent with an absence of silencing.

REFERENCES:

1. Lyon M. F. Gene action in the X-chromosome of the mouse (Mus musculus L.). Nature. 1961 Apr 22;190:372-3.

2. Lyon MF Sex chromatin and gene action in the mammalian X-chromosome. Am J Hum Genet. 1962 Jun;14:135-48.

3. Silva SS, Rowntree RK, Mekhoubad S, Lee JT. X-chromosome inactivation and epigenetic fluidity in human embryonic stem cells. Proc Natl Acad Sci U S A. 2008 Mar 25;105(12):4820-5

4. Tchieu J, Kuoy E, Chin MH, Trinh H, Patterson M, Sherman SP, Aimiuwu O, Lindgren A, Hakimian S, Zack JA, Clark AT, Pyle AD, Lowry WE, Plath K. Female human iPSCs retain an inactive X chromosome. Cell Stem Cell. 2010 Sep 3;7(3):329-42.

5. Mekhoubad S, Bock C, de Boer AS, Kiskinis E, Meissner A, Eggan K. Erosion of dosage compensation impacts human iPSC disease modeling. Cell Stem Cell. 2012 May 4;10(5):595-609.

6. Deng, Berletch JB, Ma W, Nguyen DK, Hiatt JB, Noble WS, Shendure J, Disteche CM. Mammalian X upregulation is associated with enhanced transcription initiation, RNA half-life, and MOF-mediated H4K16 acetylation. Dev Cell. 2013 Apr 15;25(1):55-68.

7. Kharchenko PV, Xi R, Park PJ. Evidence for dosage compensation between the X chromosome and autosomes in mammals. Nat Genet. 2011 Nov 28;43(12):1167-9;

8. Yildirim, E, Sadreyev RI, Pinter SF, Lee JT. X-chromosome hyperactivation in mammals via nonlinear relationships between chromatin states and transcription. Nat Struct Mol Biol. 2011 Dec 4;19(1):56-61.

9. Lin H, Halsall JA, Antczak P, O'Neill LP, Falciani F, Turner BM. Relative overexpression of X-linked genes in mouse embryonic stem cells is consistent with Ohno's hypothesis. Nat Genet. 2011 Nov 28;43(12):1169-70;

10. Nazor KL, Altun G, Lynch C, Tran H, Harness JV, Slavin I, Garitaonandia I, Müller FJ, Wang YC, Boscolo FS, Fakunle E, Dumevska B, Lee S, Park HS, Olee T, D'Lima DD, Semechkin R, Parast MM, Galat V, Laslett AL, Schmidt U, Keirstead HS, Loring JF, Laurent LC. Recurrent variations in DNA methylation in human pluripotent stem cells and their differentiated derivatives. Cell Stem Cell. 2012 May 4;10(5):620-34.

11. Tomoda K, Takahashi K, Leung K, Okada A, Narita M, Yamada NA, Eilertson KE,Tsang P, Baba S, White MP, Sami S, Srivastava D, Conklin BR, Panning B, Yamanaka S. Derivation conditions impact X-inactivation status in female human induced pluripotent stem cells. Cell Stem Cell. 2012 Jul 6;11(1):91-9.

12. Bock C, Kiskinis E, Verstappen G, Gu H, Boulting G, Smith ZD, Ziller M, Croft GF, Amoroso MW, Oakley DH, Gnirke A, Eggan K, Meissner A. Reference Maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. Cell. 2011 Feb 4;144(3):439-52

13. Zvetkova I, Apedaile A, Ramsahoye B, Mermoud JE, Crompton LA, John R, Feil R, Brockdorff N. Global hypomethylation of the genome in XX embryonic stem cells. Nat Genet. 2005 Nov;37(11):1274-9.

14. Lin H, Gupta V, Vermilyea MD, Falciani F, Lee JT, O'Neill LP, Turner BM.Dosage compensation in the mouse balances up-regulation and silencing of X-linked genes. PLoS Biol. 2007 Dec;5(12):e326.

15. Schulz EG, Meisig J, Nakamura T, Okamoto I, Sieber A, Picard C, Borensztein M, Saitou M, Blüthgen N, Heard E. The two active X chromosomes in female ESCs block exit from the pluripotent state by modulating the ESC signaling network. Cell Stem Cell. 2014 Feb 6;14(2):203-16.

16. Ludwig TE, Levenstein ME, Jones JM, Berggren WT, Mitchen ER, Frane JL,Crandall LJ, Daigh CA, Conard KR, Piekarczyk MS, Llanas RA, Thomson JA. Derivation of human embryonic stem cells in defined conditions. Nat Biotechnol. 2006 Feb;24(2):185-7.

17. Hanna J, Cheng AW, Saha K, Kim J, Lengner CJ, Soldner F, Cassady JP, Muffat J, Carey BW, Jaenisch R. Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. Proc Natl Acad Sci U S A. 2010 May 18;107(20):9222-7.

18. 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.

19. Rossant J, Papaioannou VE. Outgrowth of embryonal carcinoma cells from injected blastocysts in vitro correlates with abnormal chimera development in vivo. Exp Cell Res. 1985 Jan;156(1):213-20

20. Xu Y, Zhu X, Hahm HS, Wei W, Hao E, Hayek A, Ding S. Revealing a core signaling regulatory mechanism for pluripotent stem cell survival and self-renewal by

small molecules. Proc Natl Acad Sci U S A. 2010 May 4;107(18):8129-34.

21. Lengner CJ, Gimelbrant AA, Erwin JA, Cheng AW, Guenther MG, Welstead GG, Alagappan R, Frampton GM, Xu P, Muffat J, Santagata S, Powers D, Barrett CB, Young RA, Lee JT, Jaenisch R, Mitalipova M. Derivation of pre-X inactivation human embryonic stem cells under physiological oxygen concentrations. Cell 2010 May 28;141(5):872-83.

22. Tachibana M, Ma H, Sparman ML, Lee HS, Ramsey CM, Woodward JS, Sritanaudomchai, Masterson KR, Wolff EE, Jia Y, Mitalipov SM. X-chromosome inactivation in monkey embryos and pluripotent stem cells. Dev Biol. 2012 Nov 15;371(2):146-55.

23. Okamoto I, Patrat C, Thépot D, Peynot N, Fauque P, Daniel N, Diabangouaya P,Wolf JP, Renard JP, Duranthon V, Heard E. Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development. Nature. 2011 Apr 21;472(7343):370-4.

24. Ware CB, Wang L, Mecham BH, Shen L, Nelson AM, Bar M, Lamba DA, Dauphin DS, Buckingham B, Askari B, Lim R, Tewari M, Gartler SM, Issa JP, Pavlidis P, Duan Z, Blau CA. Histone deacetylase inhibition elicits an evolutionarily conserved self-renewal program in embryonic stem cells. Cell Stem Cell. 2009 Apr 3;4(4):359-69.

25. Diaz Perez SV, Kim R, Li Z, Marquez VE, Patel S, Plath K, Clark AT. Derivation of new human embryonic stem cell lines reveals rapid epigenetic progression in vitro that can be prevented by chemical modification of chromatin. Hum Mol Genet. 2012 Feb 15;21(4):751-64.

26. Medine CN, Lucendo-Villarin B, Zhou W, West CC, Hay DC. Robust generation of hepatocyte-like cells from human embryonic stem cell populations. 2012 Jun 10.StemBook [Internet]. Cambridge (MA): Harvard Stem Cell Institute; 2008.

27. Lian X, Hsiao C, Wilson G, Zhu K, Hazeltine LB, Azarin SM, Raval KK, Zhang J, Kamp TJ, Palecek SP. Robust cardiomyocyte differentiation from human pluripotent stem cells via temporal modulation of canonical Wnt signaling. Proc Natl Acad Sci U S A. 2012 Jul 3;109(27):E1848-57.

28. Chambers SM, Fasano CA, Papapetrou EP, Tomishima M, Sadelain M, Studer L. Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. Nat Biotechnol. 2009 Mar;27(3):275-80.

29. Chang MG, Chang CY, de Lange E, Xu L, O'Rourke B, Karagueuzian HS, Tung L, Marbán E, Garfinkel A, Weiss JN, Qu Z, Abraham MR. Dynamics of early afterdepolarization-mediated triggered activity in cardiac monolayers. Biophys J. 2012 Jun 20;102(12):2706-14

30. Chen PY, Cokus SJ, Pellegrini M. BS Seeker: precise mapping for bisulfite sequencing. BMC Bioinformatics. 2010 Apr 23;11:203

31. Gardiner-Garden M, Frommer M. CpG islands in vertebrate genomes. J Mol Biol. 1987 Jul 20;196(2):261-82.

CHAPTER 4

CURRENT STATES OF NAÏVE HUMAN PLURIPOTENCY DIVERGES SIGNFICANTLY FROM GROUND STATE MOUSE PLURIPOTENCY

ABSTRACT

States of pluripotency represent a relatively new concept in early developmental biology where as opposed to the traditional solitary state of pluripotency defined by mouse embryonic stem cells, multiple stages have been isolated through modifications of culture conditions [1-6]. In essence the originally derived mouse embryonic stem cell represents the in vitro equivalent of the late mouse pre-implantation blastocyst inner cell mass and the features within is termed the naïve state of pluripotency. As described earlier, this state of pluripotency has been an incredible resource to biologists because aside from practical advantages such as ease of culture and stability across passages, the naïve state of mouse pluripotency has been integral in the creation of transgenic mice: arguably the most important model system to study mammalian biology.

Naïve pluripotency has broad implications for human stem cell biology namely it would provide the ability to interrogate the mechanisms of early human development in vitro. These cells would potentially supply equivalent gains in understanding biology that mouse naïve pluripotent cells have which include high cloning efficiency as single cells, robust homologous recombination rates allowing for better genetic manipulation, faster proliferation, increased genomic stability over time, minimal differentiation bias, and perhaps a stronger utility for regenerative medicine and studies of developmental biology than the current state of human embryonic stem cells (hESCs). In this context we were interested in human naïve pluripotency because it would allow us to mechanistically dissect human X inactivation as had worked so well in the mouse system (see Chapter 3 for kinetics of human XCI *in vitro* and Chapter 2 for a review of mouse XCI). The experiments in this chapter were an attempt to stabilize this state of

pluripotency with approach and endpoints changing as the field evolved. Therefore the initial experiments focused on replication of the described work in the attempt to stabilize the XaXaXist- state in human induced pluripotent stem cells (hIPSCs) but the latter experiments were an attempt to stabilize an XaXaXist++ state as had been clarified in the middle of my graduate work to be the pre-XCI state of female humans. The conclusions from these experiments is that the we were unable to reproduce much of the published literature's findings in regards to the creation of an XaXaXist+ state from XaXiXist- states and that current methods for producing naïve pluripotency do not resemble the pre-implantation blastocyst neither through our assays focusing on the X chromosome nor through the data published from these reports.

INTRODUCTION:

Mouse naïve pluripotency encompasses two distinct states: the traditional serum and LIF dependent naïve state and a subclassification grown in chemically defined media termed the ground state of pluripotency [3]. The ground state is marked by a profound global DNA demethylation [7-10], loss of the characteristic bivalent (H3K4Me3 and H3K27Me3) at lineage specific genes [11], maintained self-renewal capacity at clonal densities and a transcriptome that mirrors the early pre-implantation blastocyst. This state of pluripotency can be achieved in mouse embryonic stem cells by growing in the presence of a Wnt agonist (usually the GSK3B inhibitor CHIR99021) and a MAPK antagonist (usually PD0325901) [3]. These cells potentially also harbor a subpopulation of totipotent cells that have a higher propensity for contribution to extra-embryonic tissues then embryonic stem cells grown in serum [12]. This is in contrast to usual state

of mouse embryonic stem cells (mESCs) grown in serum and LIF where these cells are rarely capable of contributing to the extra-embryonic tissue in chimerism assays[13-17]. Early work from the Smith lab showed that this combination of small molecules could stabilize a transgene free rat embryonic stem cell: which until then had not been able to be derived in usual naïve culture conditions. Since that report, rat embryonic stem cells grown in ground state conditions have been used to expand the utility of this model species in a similar fashion as had been previously enjoyed only by the mouse [18-24]. This extension of capturing naive pluripotency in other species within rodentia and the findings that not all strains of mice are competent to produce embryonic stem cells in serum and many require 2i conditions [25-29] led to a hypothesis that perhaps human naïve pluripotency could be captured using similar conditions. The Jaenisch lab published in 2010 that naïve human pluripotent stem cells (hPSCs) could be stabilized through exogenous transgene expression of the Yamanaka reprogramming factors, Oct4, Sox2, Klf4 and cMyc in 2i conditions or for a limited set of passages in the absence of exogenous transgene expression with the addition of a cAMP agonist, forskolin [30]. In 2013 the same study author essentially overturned this proposed technique of adapting 2i conditions to human stem cells where now multiple inhibitors affecting a multitude of pathways are required to achieve naïve human pluripotency without exogenous transgenes [31]. This new technique emphasizes the addition of basic FGF and TGF-beta, signaling molecules that are essential for primed pluripotency and used in the absence of LIF to convert naïve mESCs to primed pluripotent cells. Without these essential signaling molecules the conditions are susceptible to yield neural stem cell differentiation from primed human stem cell [32]. This chapter will

focus on experiments that initially attempted to recapitulate the proposed mechanisms for achieving naïve pluripotency and then conclude with the finding that no method to data has achieved a state of pluripotency that resembles the human pre-implantation blastocyst. I will also discuss potential endpoints that can be used as metrics for identifying the conditions required to potentially achieve this state.

RESULTS:

Neither hypoxia, antioxidants, nor passaging methods affect the stability of the XaXiXist- state in human embryonic stem cells

Culture of cells in vitro is typically done in atmospheric concentrations of oxygen (21%), which are vastly above what is believed to be the oxygen tension in tissues (~5%). This difference was argued to account for increased reactive oxygen species generation in cultured cells and served the basis for a series of findings arguing for hypoxic culture conditions stabilizing the pristine state of hESCs [33]. Additionally, stress from freeze-thaw cycles or enzymatic dissociation for passaging was also argued to push cells from an XaXiXist+ state to an XaXiXist- state. At the time the both the XaXa and XaXiXist+ state were considered to be labile and we were fortunate to have derived two hESC lines from the same donors and thus had genetically matched lines with two distinctly different states of XCI at passage 1 after derivation. UCLA8 was a dominantly XaXiXist- state and UCLA9 was a pure XaXaXist- state (Fig 4-1 A). Thus we could interrogate the contribution of hypoxia and antioxidant treatment for maintenance of the XaXaXist- state and the contribution of stress from enzymatic dissociation for the maintenance of the XaXiXist- state. We found that the XaXaXist- state is remarkably stable and potential reasons are elaborated in chapter 3 of this thesis (Fig 4-1 B/C). We also found that neither oxygen concentrations of the culture environment nor passaging method influenced the loss of Xist+ cells in UCLA8 and that although stress of culture may contribute to this loss, hypoxia nor manual passaging are

preventative (Fig 4-1 D/E/F). The emphasis of the previous work regarding XaXaXist- cells as labile and stabilized by hypoxia was reinforced by findings regarding the use of antioxidants in the media. Also argued to prevent a transition of these cells into XaXiXist+ cells in self-renewing conditions, we attempted to use antioxidants described by Lengner et. al. at the same concentrations described. We found that no antioxidant reported to prevent this transition (Fig 4-1 G/H), and given the stability of XaXaXist- cells in the usual culture conditions, no apparent differences with the inclusion of antioxidants on these cells. This demonstrates that the use of hypoxia to derive and culture hESCs has no remarkable affect on the usual kinetics of *XIST* loss, perhaps supporting faster silencing. Additionally, the prevailing model of XaXaXist- cells as labile has been argued to be not the case in chapter 3, and therefore the use of physiological oxygen or antioxidants do not appreciably change these cells in short term culture.

Histone Deacetylase inhibition do not capture human naïve pluripotency

X chromosome inactivation involves multiple sequential events leading to stable gene repression on the future Xi. Epigenetic mechanisms for enforcement of silencing involve histone modifications that are conducive for repression including H3K27Me3 placed by the PRC complexes and DNA methylation of gene promoters [reviewed in 34]. In 2009, a report of HDAC inhibition on female human ES cells suggested that reactivation of the inactive X chromosome could occur and thus we were interested in this relatively simple mechanism for creating a pre-XCI state. UCLA1 and UCLA3 from

the hESC core were passaged into hESC media containing 0.2 mM Sodium Butyrate for 3 passages. RNA FISH for an X-linked gene (*ATRX*) subject to inactivation showed that reactivation did not occur in human ES cells grown in this media (Fig 4-2 D) in any scored nuclei. Additionally, as noticed by morphology the colonies in sodium butyrate had a subjectively tighter appearance with better-defined edges than those grown in usual culture conditions with vehicle (water). Incidentally (Fig 4-2A bottom right panel) shows that using sodium butyrate during reprogramming led to polyploidy with multiple H2K27Me3 foci and corresponding supernumerary Atrx FISH signals during reprogramming from fibroblasts. These experiments demonstrate that the use of the HDAC inhibitor Sodium Butyrate does not lead to a transition of post-XCI XaXiXist- hESC to an XaXaXist+ hESC.

The addition of HDAC inhibitors nor culture on mouse LIF secreting fibroblasts are insufficient to establish an XaXaXist- state in human somatic cell reprogramming.

Somatic cell reprogramming to induced pluripotency yields cells with characteristics similar to hESCs. We were unable to use HDAC inhibition to produce an XaXaXist- hESC from an XaXiXist- line, but speculated that the state may be extremely stable as these hESC lines have been cultured for many passages after derivation. Therefore we attempted to identify if any reported condition that yields an XaXa state could induce an XaXaXist – state through somatic cell reprogramming. In order to test these conditions we used concentrated lentivirus expressing the 4 reprogramming

factors on normal human dermal fibroblasts (Fig 4-1 B). We evaluated the percent of pluripotent stem cells that were created during this process using the pluripotency marker Tra 1-60, and the percent of cells that may have lost *XIST* by using immunofluorescence for focal enrichment of H3K27Me3. We tested three variables, the basal media used in standard hESC culture versus that sold by Reprocell for which we believed was the media that Dr. Yamanaka, 0.2 um Sodium Butyrate, and the feeder fibroblasts used in standard (CF1 mouse derived) culture or in Yamanaka reprogramming conditions (SNL a STO cell line that overexpresses mouse LIF). We found that sodium butyrate did increase the number of Tra 1-60 positive colonies in reprogramming cultures but also led to some increase in the ploidy of these colonies as determined by multiple H3K27Me3 foci (Fig 4-1B/A). We did we see a single colony with more than 5 nuclei that possessed an XaXa Atrx RNA FISH signal in Yamanaka conditions (SNL feeders with Reprocell primate media and 4 ng/mL bFGF). We subsequently passaged those cells in the Yamanaka conditions for 10 passages and never saw a rise in the number of XaXa positive cells, and attributed these scant cells to be byproducts of reprogramming conditions and certainly not entire clonal colonies. We concluded that perhaps the reprogramming conditions described were not adequately encompassing the details necessary in order to reproduce the published findings. Therefore we attempted to focus on the naïve reprogramming, given that had the most promise in terms of the complementary evidence of achieving a developmental stage that may be similar to the pre-XCI human blastocyst.

Naïve pluripotency conditions using 2i + human LIF were not capable of inducing a naïve human pluripotent state in the context of somatic cell reprogramming.

The previous experiments focused on interconverting states of XCI within the primed pluripotent state, however at the time newer reports regarding potential naïve pluripotency in human cells were intriguing. At this time, the state of XCI in human female blastocysts was unknown but the prevailing model in primed human ES cells was that the XaXaXist- cells were at the top of a hierarchy of X inactivation states and was reflective the pre-implantation human blastocyst. This conclusion was drawn from known XCI states during early mouse embryogenesis and the finding in that the original publication where during EB differentiation, an increase in the percent of *XIST* positive cells within the culture [35]. The parallels between culture conditions and molecular phenotypes of hESCs and mEpi-SCs are rather strong (reviewed in Chapter 2), and indicate that the hESCs in the primed state resembles the pluripotent stem cell derived from post-implantation mouse embryos. Chapter 3 serves to provide an explanation for this resemblance by looking at the progression of XCI during derivation and identifying a cell fate change from the naïve human blastocyst to the primed human ES cell. Therefore, given that in mice, X inactivation has not occurred in the pre-implantation blastocyst, and naïve cell culture of mESCs stabilize a state of pre-XCI, an impetus for creating naïve pluripotent stem cells in humans was apparent. In 2010, the Jaenisch lab had shown somatic cell and primed hESC reprogramming to the naïve state was possible by using mouse ground state culture conditions (1 uM PD0325901 – MAPK inhibitor, 3 uM CHIR990221 – the GSK3B inhibitor, and human LIF, albeit with constant

transgene expression of the four reprogramming factors: Oct4, Sox2, Kl4 and cMyc. The metrics for this state included the utilization of the distal Oct4 enhancer, rapid cell growth, single cell dissociation capacity, transcriptome that correlated with naïve mouse ES cells and active STAT3 signaling through a positive phosphorylation found upon exposure to human LIF. In addition, using RNA FISH they were able to demonstrate a process of de novo XCI in these cells where in self-renewing conditions, the cells were absent for *XIST* but then upon reconversion to primed pluripotency or through embryoid body mediated differentiation: upregulation of *XIST* clouds indicating de novo XCI. We were very interested in reproducing and establishing this system given this could allow us to interrogate the mechanisms of human XCI. To that end we initially attempted conversion of human fibroblasts using concentrated lentivirus expressing the four reprogramming factors under 2i+LIF conditions (Fig 4-1 A). Although marked loss of cells in the reprogramming culture was generally noted during the course of exposure to 2i+ LIF, few colonies would arise that resembled mESCs by morphology by Day 10 post infection. However these colonies would crash by day 16 and no stable colony could be maintained in these conditions. We used RNA FISH for *XIST* and *ATRX* and immunofluorescence for H3K27Me3 (enriched on the inactive X when Xist is coating the chromosome) on the scant colonies left by day 19 in these conditions. We found that the few colonies alive in naïve culture conditions during somatic cell reprogramming did not have biallelic gene expression for the x-linked gene, however many would interestingly lose H3K27Me3 suggesting a loss of Xist.

We hypothesized that perhaps the conversion of somatic cells to naïve pluripotency may be extremely inefficient given somatic cell reprogramming to primed

conditions is already very inefficient when using lentivirus on unmodified cells. This report from the Jaenisch lab had used a secondary reprogramming system where the cells had undergone a conversion to primed pluripotency through lentiviral based reprogramming, subsequently differentiated to fibroblast-like cells, and then re-induced with reprogramming factors from the integrated constructs. Additionally, the clone used in this paper had an unusually high kinetic of reprogramming during the first phase, where unlike most colonies that arose at 8 weeks post transduction of three factors of hOCT4, hSOX2 and hKLF4, the C1 clone was isolated at 2 weeks post transduction [36] implying a specific feature of this specific clone to rapidly attain a pluripotent state.

We initially asked if simple addition of the Jaenisch media formulation onto human primed iPS cells could induce any conversion, if even at very low levels. We had a unique system in the lab where human iPS cells were generated from a female that were a carriers for an HPRT mutation, an X-linked gene that when active in males leads to the Lesch-Nyhan disease. This system allows for the generation of human iPS cells that are either expressing the WT OR the mutant HPRT on the active X chromosome given somatic cell reprogramming to the primed state, in our hands, always leads to an XaXiXist+ state. Upon exposure of human iPSCs that are $Xa^{(HPRT-MUT)}$ $Xi^{(HPRT-WT)}$ to the nucleotide salvage pathway inhibitor cocktail HAT (hypoxanthine, aminopterin and thymidine) the iPS cells rapidly differentiate and die. However unlike cells expressing WT HPRT they are resistant to the chemotherapeutic agent 6-thioguanine (6TG). Conversely, $Xa^{(HPRT-WT)}Xi^{(HPRT-MUT)}$ hIPSCs are resistant to HAT but susceptible to 6TG. Therefore by using clones that were $Xa^{(HPRT-MUT)}$ $Xi^{(HPRT-WT)}$, we could identify if certain culture conditions successfully reactivated the in active X

chromosome by culturing in the presence of 6TG, for which they will be resistant in baseline conditions and die in reactivated conditions (Fig1C). We used the aforementioned naïve conditions published form the Jaenisch Lab (2i+LIF) as well as just human LIF and identified that the 2i+LIF conditions yielded profound differentiation and in both 6TG and HAT implying a loss of pluripotency and that human LIF yielded no difference from baseline conditions implying that the addition of LIF to the media does not result in reactivation of the X

Forced overexpression of the reprogramming factors in human ES and human fibroblasts leads to ES like colonies that are not stable nor demonstrate an XaXa phenotype by RNA FISH

Simple conversion of media from primed pluripotent specific conditions to naïve conditions often led to profound cell death, however that phenomena had been described in conversion of mouse primed cells to naïve cells and was implied from the Hanna et. al [30]. Exogenous transgene expression of Klf4 in mouse primed cells was reported to increase the frequency of conversion to mouse naïve cells [37]. Moreover, the use of a cAMP agonist forskolin allowed Dr. Hanna to stabilize his naïve cells grown in 2i+LIF in the absence of exogenous transgenes. Therefore we decided to attempt the conversion of human cells in the presence of consistent transgene expression as well as in the presence of forskolin. We used a doxycycline inducible system to express the reprogramming factors in somatic cell reprogramming as well as using the same transgene cassette published by Dr. Hanna in primed human ES reprogramming. We

also cloned a reprogramming cassette that expressed L-myc as suggested by Dr. Yamanaka to reduce transformation potential in human somatic cell reprogramming over c-Myc and attached a GFP marker to identify when the transgene was expressed during the course of reprogramming. Human primed pluripotent stem cells were transfected with a plasmid that contained human Oct4 and human Klf4 under a CAGGS promoter and then 48 hours later naïve conditions (2i + LIF, 2i + LIF + FK) were applied. Again massive differentiation and death was observed but interestingly mESC like colonies by morphology could also be seen and many continued to grow during the subsequent two weeks (Fig 4-1E). However RNA FISH for Xist and Atrx demonstrated that these colonies were monoallelic for X linked genes and thus no different from the starting population of primed pluripotent stem cells. Somatic cell reprogramming with individual and the polycistronic cassette also yielded mouse ES like colonies however none of the demonstrated a reactivation of the X chromosome by RNA FISH but almost all showed no H3K27Me3 foci enrichment implying a loss of Xist (Fig 4-1F). This showed that the inclusion of the pluripotency factors in the form of a transgene was required to observe a morphological change, however the X chromosome was not reactivated as had been described by the original publication. We hypothesized that the levels of the reprogramming factors may be insufficient in these experiments, as that notion had been proposed in the field of somatic cell reprogramming in general around the time these experiments were underway [38].

Lentiviral reprogramming from somatic and primed cells into putative naïve conditions involve cell fate changes that may silence viral vectors

We realized that these experiments were underpowered because we had no marker of reprogramming progression during the course of the experiment. Furthermore, the use of serial immunofluorescence during the reprogramming process was an option to identify if any XaXa colony was arising, even transiently but felt this was too labor intensive and subject to false negatives. Regardless, the new lentiviral cassette we were using had a GFP flourophore attached through viral 2A sequence, and what became quite evident was despite a high percentage of cells fluorescing green in the early days after initiation of reprogramming, as soon as media exchange occurred to primed or naïve conditions, the GFP transgene would silence. This was reproducibly seen and implied that the entire cassette was being silenced.

We expected this to occur late in reprogramming as viral vectors are known to be silenced in pluripotent states, however in these experiments this event was occurring extremely early. Additionally, primed reprogramming conditions would also show a loss of GFP after conversion from fibroblast to ES media and a cessation of cell growth and morphological changes indicative of ongoing reprogramming compared to the previously validated control vector. We decided to test the hypothesis of viral transgene silencing by creating a tet inducible GFP and tdTomato lentiviral construct and infect cells during reprogramming. Near the time of media conversion a complete loss of fluorescence was observed in the cultures consistent with silencing of either the reverse transactivator or the tet inducible flourophore. A collaboration with the Don Kohn lab showed that the splice acceptor in the rtta construct had been destroyed during cloning and resulted in aberrantly spliced products (Fig 4-3D). Thus we decided to re-clone the

M2rtta in the original FuW vector with the correct splice acceptor as well as switch the promoter for one that has been demonstrated to resist silencing in multiple cell states (A2UCOE) [39]. With this new vector we were able to improve the expression levels through the course of reprogramming by using a co-transduced flourophore. Additionally to reduce variability in between experiments we decided to utilize the Kohn lab to create tangential flow concentrated lentivirus of the following tet inducible factors (Fig 3) such that the same high-titer virus could be used to address concerns that perhaps the levels or stoichiometry of the factors was at fault. We cloned into a doxycycline-inducible FUW based lentiviral backbone the usual reprogramming factors: hOCT4, hSOX2, hKLF4, hCMYC, hNANOG as well as factors implied to important for mouse pluripotency: mStat3-CA (constitutively active), hPRDMT14, hKLf2 and factors involved in a report of naïve pluripotency from the Sanger institute [40]: hLRH1 and hRARy, as well as combinations thereof mStat3CA – 2A – NANOG, hKLF2 -2A hKLF4, hKLF2 – 2A – mKLF4, hKLF4 - 2A - hOCT4. These viruses were validated for expression of nuclear protein and titers from immunofluorescence ranged from 1-5 x 10^7 TU/mL (Fig 4-4A). They were used in both somatic cell reprogramming and reprogramming from the primed state with the rtta cassette driven off the A2UCOE promoter.

No published condition allowed for stable reprogrammed naïve cells lines from somatic cells nor from primed human ES.

Since the first description of naïve pluripotency in human cells, many conditions have been clarified both by that study author as well as others in the field. However

capturing human naïve pluripotency should involve the reflection of epigenetic and transcriptomic states that are similar to the pre-implantation blastocyst. For mouse naïve cells, a striking feature is ability to undergo X inactivation upon loss of pluripotency. These cells begin as XaXa and result in a random inactivation with respect to allele of XaXiXist+ cells. This same sequence was thought to occur in human development, and thus served the basis for a number of papers claiming naïve pluripotency. It was not until clarifying experiments from the Heard laboratory showing that pre-implantation blastocysts from humans are not XaXa but rather XaXaXist++. In chapter 3 of this thesis I further show that upon derivation of human ES cells, the XaXaXist++ state transitions into an XaXiXist+ state but also frequently can stall as an XaXaXist- state. Therefore, our earlier attempts for inducing naïve pluripotency by searching for the XaXa state were misguided, and under those as well as the conditions described in Fig 3D we can not find a state of human pluripotency that captures the pre-implantation blastocyst state. Manipulations regarding the basal media, the extracellular matrix, serum replacements, small molecules and reprogramming factors have all yielded either a loss of pluripotency or metastable pluripotent states that do not have the features of human pre-implantation blastocysts with regard to expression of pluripotency factors or the pattern of X inactivation (Fig 4-4 B-D). Experiments yielding mESC like colonies in somatic cell reprogramming were also tested to see if similar morphological changes occurred from primed hESC using the lines UCLA1, 5 and H9. These reprogramming experiments yielded neurosphere like cells with significantly faster kinetics then reprogramming from somatic cells. The exclusion of TGF-beta in all experiments could explain why we had such rapid induction of primitive neural cells in

these cultures as this has now been argued to be essential to prevent a neuroectodermal conversion of human pluripotent stem cells into the neural fate in the presence of 2i [30, 41].

Recent methods of establishing a naïve state fail at recapitulating early human development in regards to the pre-XCI state

Human naïve pluripotency possess an X inactivation state that is not like the mouse embryo nor mESCs, however most publications arguing for achieving this state include within their definition an XaXaXist- cell type that upon differentiation, upregulates *XIST*. The experiments performed in my hands to attain the naïve state have never resulted in this state of XCI. However it is conceivable that the XaXaXist- state is in a continuum coming from the XaXaXist++ naïve state to the XaXiXist- state, which is potentially stabilized in these reported culture conditions. Therefore through an initial collaborative effort, we were fortunate to receive fixed samples of these lines from the Hanna lab grown in his reported media formulation [30] and used RNA FISH for X-linked genes and *XIST* to identify if these cells are representative of the state of XCI in human blastocysts or intermediates in primed hESC culture implying they are reflective of an intermediate between primed and naïve pluripotency. No cell line sent to us deviated from the previous stages of primed pluripotent stem cells, and thus we concluded that this state, although having characteristics of mouse ES cells in terms of clonogenicity, growth rates and morphology, the *epigenetic* charecteristics associated with human naïve pluripotency were absent in these cells. To confirm the findings were

not limited to naïve conversion from previously established primed cells, we received his ES lines derived in his media, cultured in our laboratory briefly and assayed for a pre-XCI state. Both lines #38 (female) and #39 (male) showed a pattern of XCI seen in typical primed pluripotent cells and in particular the female line had many XaXiXist+ cells. This demonstrated to us that the state of pluripotency that Gafni et al had achieved may have brought some technical benefits in ease of culturing, the claim of resembling the pre-implantation blastocyst could not be validated.

Highly expressed gene in human blastocysts are not expressed in the naïve cells generated using extended MAPK inhibition

Given the limited material and difficulty obtaining human embryos for research, the ability to understand the regulatory networks important for naïve human pluripotency has been understudied. Recent work from the Tang lab succeeded in using a first generation RNA-seq method to profile human blastocysts at the single cell level. This data has allowed for a comparison of existing naïve and primed pluripotent stem cells to a standard of human naïve pluripotency seen in the pre-implantation embryo. Unsupervised clustering demonstrates that the epiblast and initial human ES derivation cultures have transcriptomes that are very similar, but interestingly at passage 10 after derivation human ES cells adopt a global transcriptome that is very distinct. This in conjunction with results presented in the Chapter 3 suggests that a naïve state of pluripotency can be identified in early derivation cultures but changes towards a different cells state during derivation. To avoid platform specific issues we used the

array data from Gafni et al [31] and re-analysed it to the array data using the same array platform from human blastocysts from Vassena et. al. [42]. We focused our attention on genes that were involved in pluripotency specific to the human blastocyst and found in ground state mESCs (Fig 4-6 A and B). The de novo methyltransferases DNMT3B and DNMT3A were downregulated in human blastocysts, but to a lesser extent in the naïve cells. This point has been argued by the authors to be due to insuffiecient p38 MAPK inhibition and involves changing the inhibitor in later versions of the media (see Hanna Lab website: protocols tab). The genes involved in naïve pluripotency in both human and mouse embryos were dynamically higher in human blastocysts but unchanged in the Gafni naïve cells. This argued that whichever state the Gafni naïve cells were in, they could were still very distinct from the human blastocyst and resembled human primed cells. This finding correlated well with our RNA FISH results for the state of XCI in these cells.

Unsupervised clustering of the average expression from these single cells by cell type show that relatively high correlations can be found, regardless of the cell type of origin (Fig 4-6 I). However, using K-means clustering (Fig 4-6 J) on the data a minimal set of genes that serves to distinguish the cell types can be identified. The naïve set includes well-known regulators of naïve pluripotency including DNMT3L and DPPA3. Figure 4-6 panel K also demonstrates a complete lack of clustering of expression values between Hanna's naïve lines and the blastocyst expression values from either the Belmonte human blastocyst array dataset [42] or Yan single cell RNA-seq datasets [43]. The Pearson correlation between differentially expressed genes in the human inner cell mass versus primed hESC and the naïve stem cells produced with version 1 of Hanna's

media cocktail and his primed hESCs is 0.003. Additionally, the top 500 genes differentially expressed in human blastocysts versus primed hESCs share only 18 genes in common with the top 500 genes differentially expressed in the Gafni et. al. naïve hESCs versus primed hESCs: of which none of them have mouse homologs that are highly expressed in naïve mESCs. This clearly demonstrates that the gene expression differences in naïve human pre-implantation blastocysts compared with primed hESCs are not addressed in the Hanna naïve lines. The original description of these lines indicated a high correlation between the naïve pluripotent stem cells and the human blastocyst, however correlations between arrays can be driven by non-biologically significant gene expression and thus likely resulted in the over-estimation of the correlation between these two cell states. Transcription factors with known function in the naïve pluripotency network like ESRRB, TFCP2L1, and DNMT3L, are not expressed in the Gafni et. al. naïve lines.

A separate group had identified a combination of small molecules for which they argue more conservatively is similar to the human blastocyst, but not reaching naïve pluripotency per se [42.1]. This group has demonstrated using a WNT agonist (BIO), a BMP antagonist (Dorsomorphin), the commonly used MEK inhibitor (PD0325901) and human LIF to be able to drive a cell state that is dependent on LIF signaling, and re-expresses genes associated with the human blastocyst, namely DPPA3, DNMT3L, and interestingly one associated with endodermal fates, GATA6. They were unable to evaluate the X chromosome given they recognized the starting population were eroded, yet the findings were very convincing that a number of pluripotency markers known to be important in mouse naïve pluripotency were upregulated. Therefore, we downloaded

the data available for the H1 male line in the two states and initially just confirmed the expression values to ensure that the absolute values were reasonably high as noted in the human blastocyst [43]. Figure 4-6 panel C shows that although the relative expression for these genes are high in the 3iL state, the absolute expression values (Panel D) are not. This is due predominantly from the lack of sequencing reads supporting the expression of these genes (given as coverage of the gene underneath the absolute values), and therefore the inferred values are likely driven from spurious read alignments in the primed state as opposed to truly high expression from the 3iL state. To date there is only one RNA-seq dataset available for human blastocysts and the relative and absolute values for these same genes are presented in Fig 4-6 panel F where aside from ESRRB, the relative values look similar to the Ng lab's naïve cells, however the absolute expression values are quite different. Interestingly, GATA6 was expressed in large percentage of 3iL cells and co-expressed with NANOG but unfortunetly with Tra 1-60 (which is not highly expressed in human blastocysts). The study authors described this as an indicator of a more primitive state where the pluripotent stem cells are also expressing primitive endoderm genes. This feature is reminiscent of previous work suggesting ground state mouse pluripotency also has is enriched for cells that express primitive endoderm genes [12,14]. However, despite these intriguing findings, no *XIST* could be found in the sequencing data, suggesting that these cells are not developmentally immature enough to recapitulate the pre-XCI state.

Several other aspects of naïve pluripotency specific to the human blastocyst seem to be absent in the culture of these lines including the reliance of IL6 signaling in

the human blastocyst as the dominant class I cytokine given most naïve cocktails described to date use LIF. Fig 4-6 L is where we place the recently described naïve pluripotent stem cells in the developmental pathway, close to primed but optimistically, slightly in the correct direction towards naïve pluripotency. These results highlight the necessity for careful re-evaluation of genes sharing high expression in human pre-implantation blastocysts amongst all published datasets in order to adequately evaluate the validity of naïve pluripotency in the human system.

DISCUSSION:

Naive pluripotency has allowed for incredible gains in our understanding of both developmental biology as well as the basis for the incredibly powerful biological tool : the transgenic mouse. However, up until work that identified key signaling pathways driving mouse pluripotency the application of naïve pluripotency was limited to particular strains of mice. Since the discovery of the 2i cocktail of small molecules naïve pluripotency and thus the ability to use homologous recombination in the setting of transgenic animals has extended through rodentia. This has also allowed for a finer dissection of pluripotency and has revealed that mouse embryonic stem cells grown in serum and LIF represent an in vitro equivalent of late stage pre-implantation blastocysts whereas mouse embryonic stem cells grown in 2i+ LIF are more similar to early stage pre-implantation blastocysts. This difference has profound consequences as the existence of bivalent promoters (combined H3K27Me3: repressive and H3K4Me3 : activating) on lineage specific genes is lost in the ground state and that global DNA methylation is much higher in the naïve state then in ground state conditions. These distinct states are convertible and thus extremely tractable for further experimentation and has yielded incredible insights into the mechanisms surrounding DNA methylation and early development

All other species reported outside of rodentia capable of producing pluripotent stem cells in culture exist in a very distinct state: the primed state of pluripotency. This state is marked by growth factor requirements that are very different from the naïve (or ground) state of pluripotency. More importantly, for scientific study the epigenetic and transcriptomic change when going from naïve to primed states of pluripotency are

profound and marked by a steep rise in global DNA methylation as well as an increase in the spontaneous expression of lineage specific genes. Mouse embryos derived form the post-implantation blastocyst or cultured in primed conditions from pre-implantation blastocysts can also be stabilized in the primed state of pluripotency, and thus to date the only convincing comparison of naïve and primed states of pluripotency are available for mice. Since 2010 reports of naïve pluripotency have arisen for human cells, and many arguments forged from a misunderstanding of human pre-implantation biology secondary to a lack of available data. Included in this confusion was the idea that human pluripotent stem cells with two active X chromosomes and no expression of Xist was the in vitro equivalent of pre-implantation blastocysts. Since 2012 and further elaborated in chapter 3 of this thesis we know that this is not the case, and that the female human pre-implantation blastocyst has two active X chromosomes with *TWO* Xist *clouds* that are not silencing. Therefore all claims to naïve pluripotency in human cells that argue for an XaXaXist- state in the undifferentiated state are largely invalid and a naïve state marked by an XaXaXist++ state has yet to be described. We attempted to derive human iPS cells from both somatic and primed pluripotent starting populations, but were unsuccessful in ever achieving a state that showed the X inactivation pattern of the human pre-implantation blastocyst. Moreover, our analytical methods applied to cells and fixed coverslips from the Hanna lab further demonstrate that the naïve state of human pluripotency has yet to be achieved.

Many of identified additions to the mouse ground state media found to help induce larger percentages of mouse es like colonies including the use of Vitamin C, high doses of Insulin, addition of small amounts of bFGF, and are now additions or critical

components of the evolving naïve media proposed by the Hanna lab (as of May 2015, 3 versions exist with more inhibitors of lineage specific pathways with no concrete evidence of achieving a more immature state in later versions). Additionally, given the experiments are constrained by hypothesis derived form mouse pluripotency (e.g. the use of 2i or the importance of ascorbic acid, the transcription factors used) it is not surprising that we are not able to create this state. The use of these same chemicals in mouse blastocysts leads to an increase in the epiblast progenitors and decrease in hypoblast cells within the embryo for which naïve mESCS are derived, but does NOT result in an increase in epiblast progenitors or limit hypoblast formation in human blastocysts [44-46]. This implies that the naïve pluripotent stem cells within the human blastocyst are maintained in that state using signaling pathways that are divergent from mouse cells and, potentially attempts at augmenting that environment within the embryo should precede attempts to stabilize the state in vitro. Fortunately, attempts to understand the biology of the human pre-implantation blastocyst are making progress.

Whether any animal is truly capable outside of rodentia for a stable naïve state in vitro is controversial. The cytokine produced by fibroblast feeders (LIF) found to be essential for the stability of mouse embryonic stem cells may have divergent mechanisms in species outside of rodentia. Unlike mouse embryos, which express the LIFR, human blastocysts express only the IL6R and thus the use of a cytokine that is capable of activating the IL6 receptor in the absence of the costimulatory receptor gp130 may give rise to a more stable from of STAT3 activation [46]. Additionally, as more in depth studies regarding embryonic stem cells arise, the genetic differences between mice and humans are becoming more apparent. For example repetitive

elements that can re-wire the regulatory regions surrounding pluripotency factors are extremely different between mice and human and thus naïve pluripotency may be "hardcoded" within the mouse genome and absent from other species.

Lastly, for the purposes of regenerative medicine, a major concern with human embryonic stem cells is that they are liable to gather genetic and epigenetic aberrations in culture that seem to occur at rates much more frequently than in mouse embryonic stem cells. However, recently investigators using ground state mouse embryonic stem cells are also identifying recurrent aberrations surrounding this state of pluripotency and therefore which state (if any) would be best for clinical applications of human stem cells. In conclusion, the identification of states of pluripotency has dramatically increased our knowledge of early development, and this has allowed for a rapid expansion of tools and technology for both basic and applied biological science. However, issues regarding latent tumorigenicity, inadequate culture conditions and limited differentiation propensities for human pluripotent stem cells persist. Further work on identifying key aspects of human pre-implantation biology will be required to help guide the efforts in stabilizing human naïve pluripotency in vitro.

EXPERIMENTAL PROCEDURES:

Human ES derivation

Human embryos donated from consenting couples through IVF treatment were thawed at or derived to the blastocyst stage under hypoxia as previously reported [25]. hESC lines were derived on mitomycin C (Calbiochem) inactivated CF1 mouse embryonic fibroblasts feeders and Manual passaging from the derivation culture yielded subclones of which one was chosen based on growth characteristics and banked by the UCLA hESC core as the final line. All lines in this study showed normal karyotypes, capacity to produce teratomas with three germ lineages, and expressed markers of pluripotency. All human ESC work was approved by the University of California Los Angeles Embryonic Stem Cell Research Oversight Committee (ESCRO). All animal research was performed with approval from the Institutional Animal Care and Use Committee (IACUC).

RNA FISH and Immunofluorescence

RNA FISH, immunofluorescence and acquisition of images were generated as described previously [48]. The following BACs were obtained from CHORI and used to make RNA FISH probe using the random priming method: Xist (RP11-13M9), Mid1 (RP11-798D20), Cask (Rp11-154-P12, RP11-820G10, and RP11-977L20), Huwe1 (RP11-975N19), Atrx (RP11-1145J4), Gpc3 (RP11-585F14). RNA FISH on blastocysts was performed as described (Nagakawa et al). Antibodies for immunofluorescence used include Nanog (R&D AF1997), Pou5F1 (Santa Cruz C10), H3K27Me3 (Millipore

07449), and Nestin (Abcam ab6142). Secondary antibodies used were Alexa-flours pre-cleared and against the appropriate primary (Life).

Gene Expression Microarray analysis

RNA from UCLA lines were extracted and process under feeder free conditions and hybridized to Affymetrix Human Genome U133 plus 2.0 arrays under standard conditions. Array data was processed using R library affy and further normalized using mas5. For external data, normalization was done in tandem with the UCLA data and after assurance of adequate normalization, expression values compared using cumulative distribution functions (R). Data generated on other arrays such as the HTGene arrays, raw data was normalized to UCLA data using quantile normalization and compared. Agilent arrays were only analyzed within experiments using package limma after QC filters for green and red channels.

Somatic cell reprogramming

Human NHDF17914 fibroblasts were obtained from the Coriell Cell Repository and grown in DMEM with 15% fetal bovine serum (Hyclone). SNL fibroblasts were obtained from the ATCC, expanded and inactivated with mitomycin C. To produce primed hIPSCs, the reprogramming factors were introduced using tangential flow concentrated lentivirus as previously described [47]. Media was exchanged for primed pluripotency media which consisted of DMEM:F12 (Gibco) with 20% Knockout Serum Replacement as previously described [48]. Experiments were typically done on #1 thickness coverslips for further assay. Naïve reprogramming was attempted with various

combinations of the following reagents. Base media was DMEM F12 (Gibco) or KO DMEM (Gibco). Matrix involved 0.3% porcine skin gelatin (Sigma), hESC qualified matrigel at a 1:100 dilution in basal media coated overnight at 4C and then warmed to 37C for two hours (BD). Laminin 521 was coated as recommended (Biolamina), Vitronectin XF (BD), or FBS coated overnight at 37C and then washed four times in large volumes of PBS (Hyclone). Serum supplementation involved using N2B27 with and without vitamin A at 1X concentrations (Life). Knockout serum replacement was used at 20% (v/v) (Life) or at 15% with 5% FBS (Hyclone). Albumax I was used at a concentration of 1% (v/v) (life). Small molecules and hormones included PD0325901 (1uM Tocris), CHIR99021 (3 uM Tocris), Y27632 (10 uM Tocris), Hydrocortisone (x uM, Tocris), A-83-01 (0.5 uM Tocris), SB203580 (5 uM Tocris), SP600125 (10 uM Tocris), L-Ascorbic Acid 2-phosphate (100 uM, Sigma), Holo-transferrin (Sigma), human LIF (recombinant made in house from e. coli over-expressed in BL21 DE), mouse LIF (recombinant made in house), recombinant human insulin (20 ug/ml Sigma), Sodium Butyrate (0.2 uM Sigma), Trolox (5 uM Sigma), Morin Hydrate (5 uM Sigma), alpha-Lipoic acid (5 uM Sigma), N-acetyl Cysteine (100 uM Sigma).

Primed hESC reprogramming

hESC were typically maintained in feeder based culture or on feeder free matrigel coated plates in mTesr1 (Stemcells). Lentiviral infections of human ES cells were most optimally performed by overnight ROCKi incubation, dispersion to single cells, incubation with concentrated lentivirus with occasional resuspension in the presence of ROCKi, spinning off virus after 4-6 hours of infection and plating onto freshly coated

matrigel or feeder based plates. Transduction efficiencies were verified using immunofluorescence for the transgene used. Media was exchanged to naïve condition 48 hours after plating when colonies were approximately 4-6 cell clusters.

RNA-Seq analysis

Average Single cell gene expression from Yan et. al. were taken as processed counts in the form of RPKMs. RNA-Seq data from Chan et. al. were processed using Sailfish [49] against the known genes track of the UCSC hg19 Genome. Values were normalized to each other using quantile normalization to reduce batch specific effects. Relative and absolute values for expression of genes were parsed from individual alignments and quantification files using the bias corrected option and plots produced using the R statistical package (http://www.r-project.org). Pearson correlations of average lineage specific expression was calculated using R. K-means clustering of the average single cell expression data for each lineage in the blastocyst was initially optimized using an custom R script to calculate the minimum number of clusters that maximizes the adjusted R^2 for the data. Then with elbow determined visually, K-means clusters 1 and 4 were visually shown to be the highest expressed in the epiblast progenitors and genes within listed. Unsupervised hierarchical clustering after quantile normalization was done using R.

Figure 4-1: The state of XCI in female human embryonic stem cells do not respond to alterations in oxygen tension nor antioxidant treatment.

A) Schematic of intent for experiment demonstrating the use of hypoxia, methods of dissociation and antioxidants to prevent the XaXaXist- (black) state transitioning to an XaXiXist+ (dark grey) state as well as and XaXiXist+ state transitioning to an XaXiXist- (light grey) state. RNA FISH counts are presented as percentages of total counts with bars indicating XCI pattern. Unscorable nuclei percentages are represented at the top of each stacked column B-E) Left panel: Manual passaging over time for indicated sublines with percentage of XCI patterns presented as stacked columns. Middle Panel: Collagenase based passaging for matched lines. Right Panel to graphs: Representative RNA FISH images of each experiment with an X-linked gene *ATRX* in red, *XIST* in orange and an escaping gene for total X chromosome count *UTX* in green. F) Passaging with the EZ tool showed near equivalent loss of XaXiXist+ cells with representative RNA FISH images below (*ATRX* in red, *XIST* in green). G-H) RNA FISH based percentages of XCI states over three passages in hypoxia as a control for normoxia with antioxidants : Morin Hydrate, N-acetyl Cysteine, Trolox and alpha-Lipoic acid.

Fig 4-1

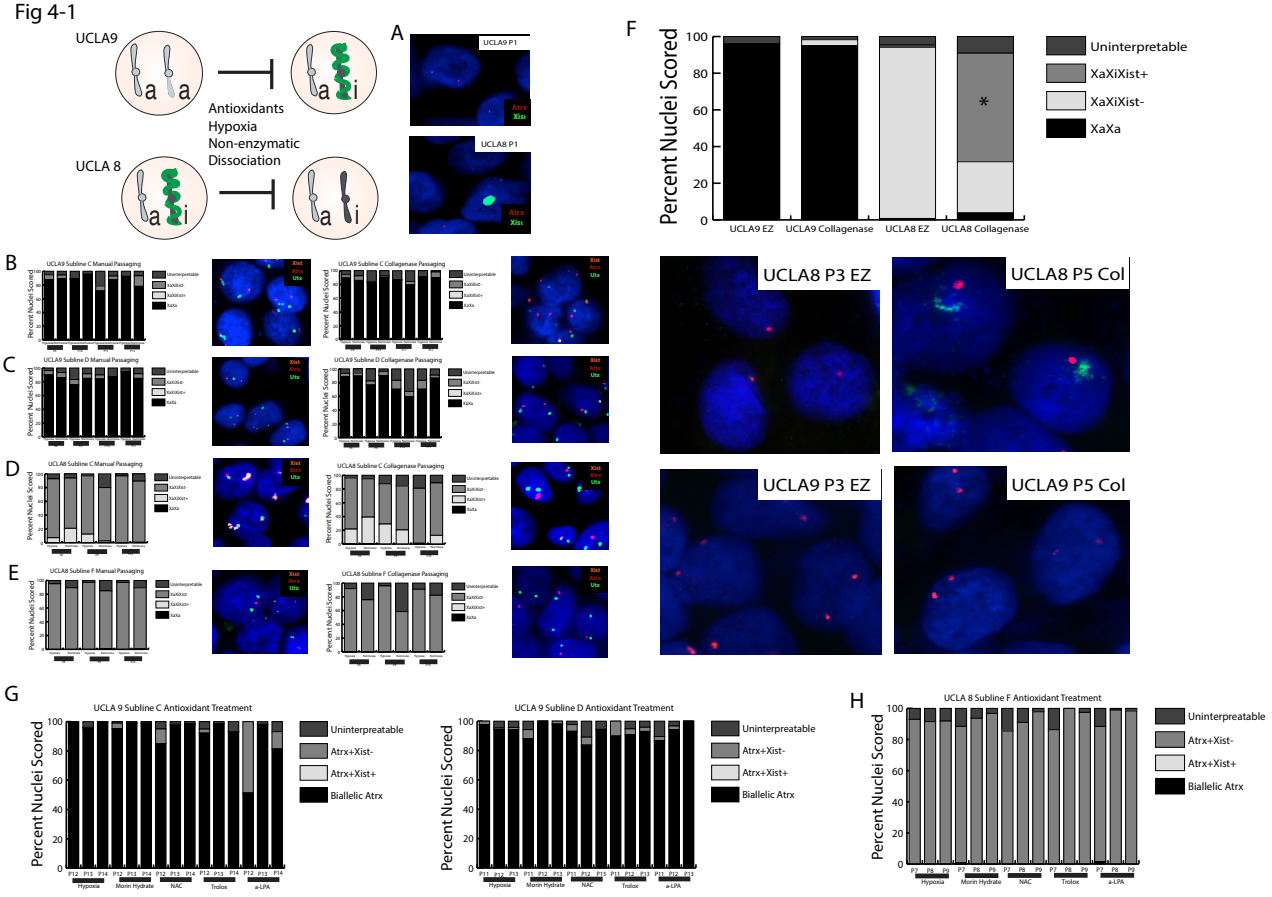

UCLA9
UCLA 8
Antioxidants
Hypoxia
Non-enzymatic
Dissociation
A
UCLA9 P1
UCLA8 P1
F
Percent Nuclei Scored
UCLA9 EZ
UCLA9 Collagenase
UCLA8 EZ
UCLA8 Collagenase
Uninterpretable
XaXiXist+
XaXiXist-
XaXa
UCLA8 P3 EZ
UCLA8 P5 Col
UCLA9 P3 EZ
UCLA9 P5 Col
G
UCLA 9 Subline C Antioxidant Treatment
UCLA 9 Subline D Antioxidant Treatment
Uninterpreatable
Atrx+Xist-
Atrx+Xist+
Biallelic Atrx
Hypoxia
Morin Hydrate
NAC
Trolox
a-LPA
H
UCLA 8 Subline F Antioxidant Treatment
Uninterpreatable
Atrx+Xist-
Atrx+Xist+
Biallelic Atrx

Figure 4-2: Attempts at converting pre-existing primed cells to XaXaXist- cells based on the published literature fail to reactivate the Xi.

A) Phase contrast images with RNA FISH and immunofluorescence for colonies formed during at the end of reprogramming experiments. Line and condition specified in subpanel captions B) Schematic of reprogramming conditions tested. Percentage of Tra 1-60 positive cells produced in each condition. Immunofluorescence based counts for colonies with H3K27Me3 foci (black bars indicate one foci, light grey bar indicate loss of foci, dark grey bar indicates more than one foci). C) Phase contrast images of drug selection and naïve conditions used on LNS hIPSCs. D) Phase contrast and RNA FISH for X-linked genes on UCLA1 and UCLA3 in sodium butyrate E) Phase contrast, immunofluorescence for pluripotency marker Tra 1-60 and absence of H3K27Me3 in reprogramming primed pluripotent stem cells using transgenes and 2i + LIF naïve conditions. RNA FISH demonstrating representative nuclei with monoallelic expression of X-linked gene *ATRX* F) Somatic cell reprogramming of normal female dermal fibroblasts in naïve conditions showing mouse es like morphology, presence of pluripotency marker Tra 1-60 and lack of H3K27Me3 foci.

Fig 4-2

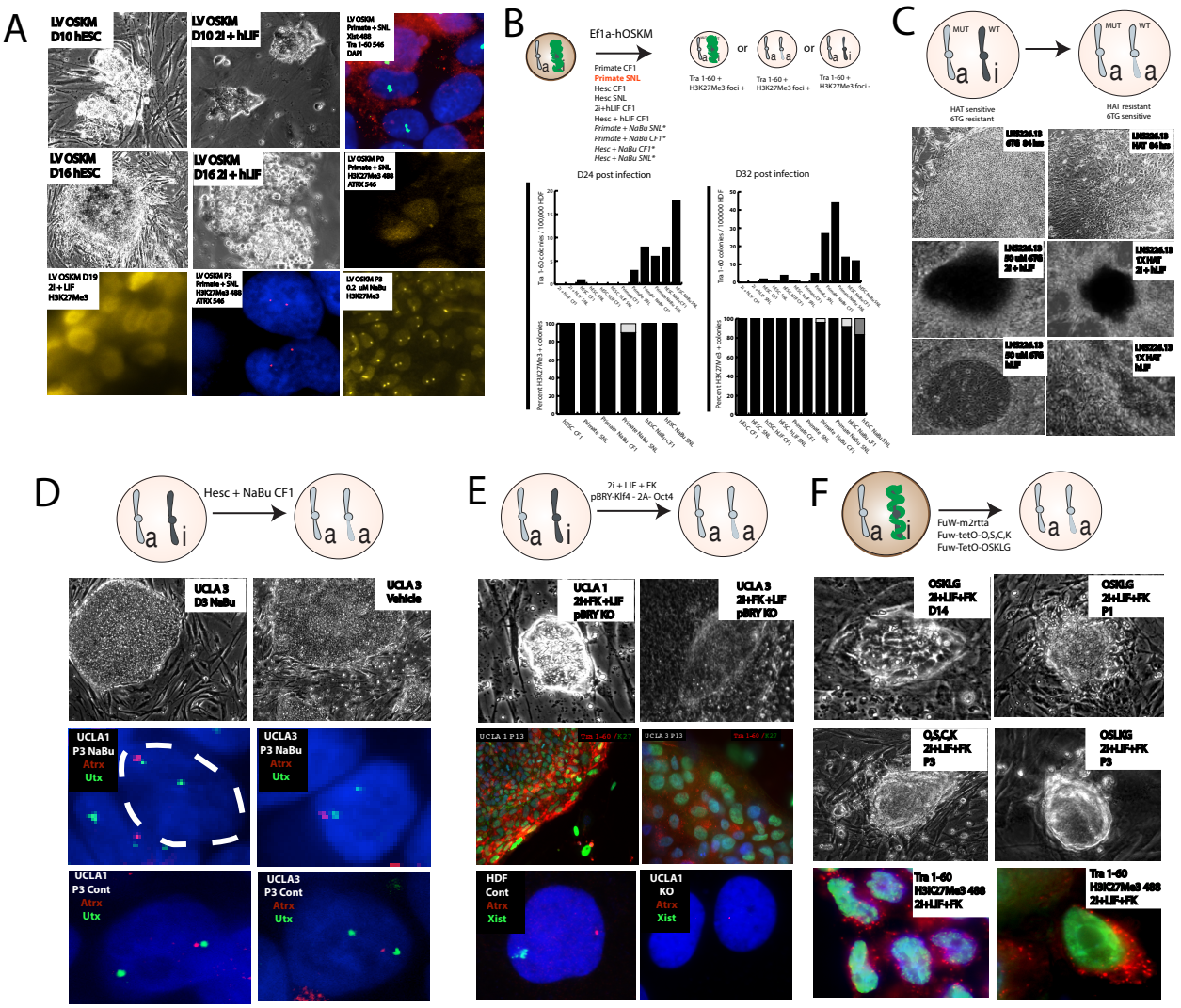
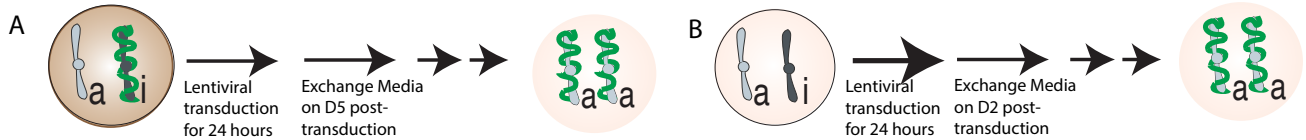

A
LV OSKM D10 hESC
LV OSKM D10 2i + hLIF
LV OSKM D16 hESC
LV OSKM D16 2i + hLIF
LV OSKM D19 2i + LIF H3K27Me3
B
Ef1a-hOSKM
Primate CF1
Primate SNL
Hesc CF1
Hesc SNL
2i+hLIF CF1
Hesc + hLIF CF1
Primate + NaBu SNL*
Primate + NaBu CF1*
Hesc + NaBu CF1*
Hesc + NaBu SNL*
D24 post infection
D32 post infection
C
MUT
WT
HAT sensitive 6TG resistant
HAT resistant 6TG sensitive
D
Hesc + NaBu CF1
UCLA 3 D8 NaBu
UCLA 3 Vehicle
UCLA1 P3 NaBu Atrx Utx
UCLA3 P3 NaBu Atrx Utx
UCLA1 P3 Cont Atrx Utx
UCLA3 P3 Cont Atrx Utx
E
2i + LIF + FK
pBRY-Klf4 - 2A- Oct4
UCLA 1 2i+FK+LIF pBRY KO
UCLA 3 2i+FK+LIF pBRY KO
HDF Cont Atrx Xist
UCLA1 KO Atrx Xist
F
FuW-m2rtta
Fuw-tetO-O,S,C,K
Fuw-TetO-OSKLG
OSKLG 2i+LIF+FK D14
OSKLG 2i+LIF+FK P1
O,S,C,K 2i+LIF+FK P3
OSLKG 2i+LIF+FK P3
Tra 1-60 H3K27Me3 488 2i+LIF+FK

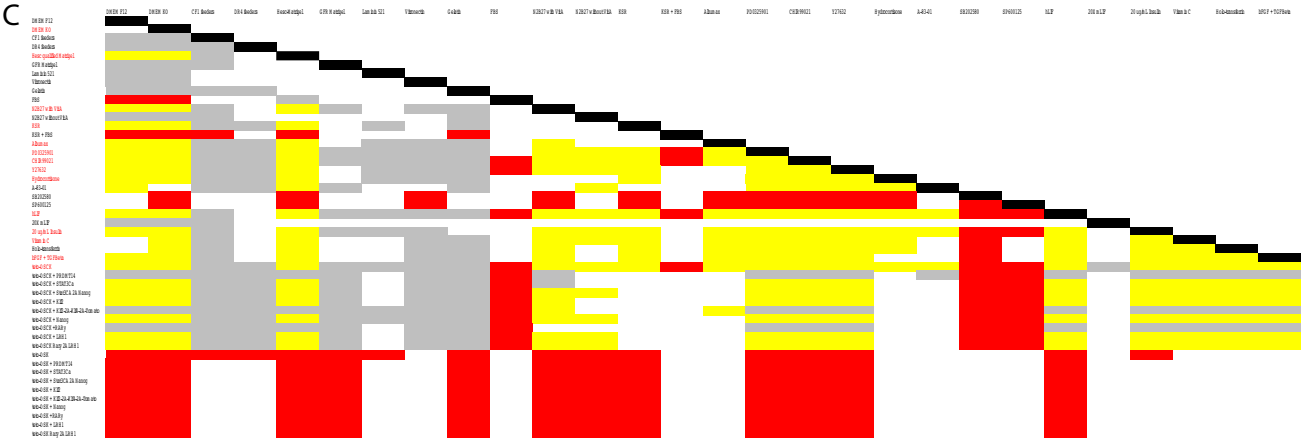
Figure 4-3: Both somatic cell reprogramming and reprogramming from primed pluripotency is not possible using typical combinations of media, small molecules, matrix, and growth factors described to date.

A) Schematic of intended outcomes for somatic cell reprogramming to pre-XCI state B) Schematic for intended outcomes for primed hESC reprogramming to pre-XCI state C) Combinations of viral transcription factors, base media, growth factors, small molecules and matrices were tested in multiple reprogramming experiments. The matrix depicted here is relaying three outcomes i) Red: no conversion of somatic cells to a reprogrammed state as assessed by lack of appreciable change in morphology ii) Grey: colonies would arise but would not be stable either succumbing in the same reprogramming culture or die upon initial passaging iii) Colonies were produced and large enough to determine if the Xi state had changed from the source cell. D) Optimal conditions for mouse es like colonies formed during virus based reprogramming are listed, however as indicated in figure 4-4 none of these conditions yielded an X-inactivation pattern resembling the pre-implantation human embryo.

Fig 4-3



C



D Optimal Conditions for creation of passageable colonies:

Virus	Small molecules	Base	Matrix	Serum substitute	Growth Factors:
tetO O,S,C,K UCOE-rtta	PD0325901 CHIR99021 Hydrocortisone	DMEM F12 or KO	Matrigel	Albumax 5 g/L or KSR 20% N2B27 with Vit A	bFGF 4 ng/ml Tgf-beta 1 ng/ml hLIF Insulin 20 ug/ml Vitamin C 100 uM

Figure 4-4: Optimal conditions of high insulin, 2i , human LIF and Vitamin C do not result in NANOG positive cells but rather enrich for neurospheres

A) PCR product of expressed rtta showing mis-spliced ORF (Photo courtesy of Aaron Cooper, Don Kohn Lab) B) Phase contrast of reprogrammed cells from normal human dermal fibroblasts in optimal conditions in the absence of bFGF or TGF-beta C) Immunofluorescence with RNA FISH showing Nestin positive colonies, heterogeneously maintaining H3K27Me3 foci implying an accelerated loss of *XIST* D) Immunofluorescence for OCT4 and H3K27Me3 in mouse like colonies from primed cells and somatic cells showing maintenance of XCI patterns.

FIG 4-4

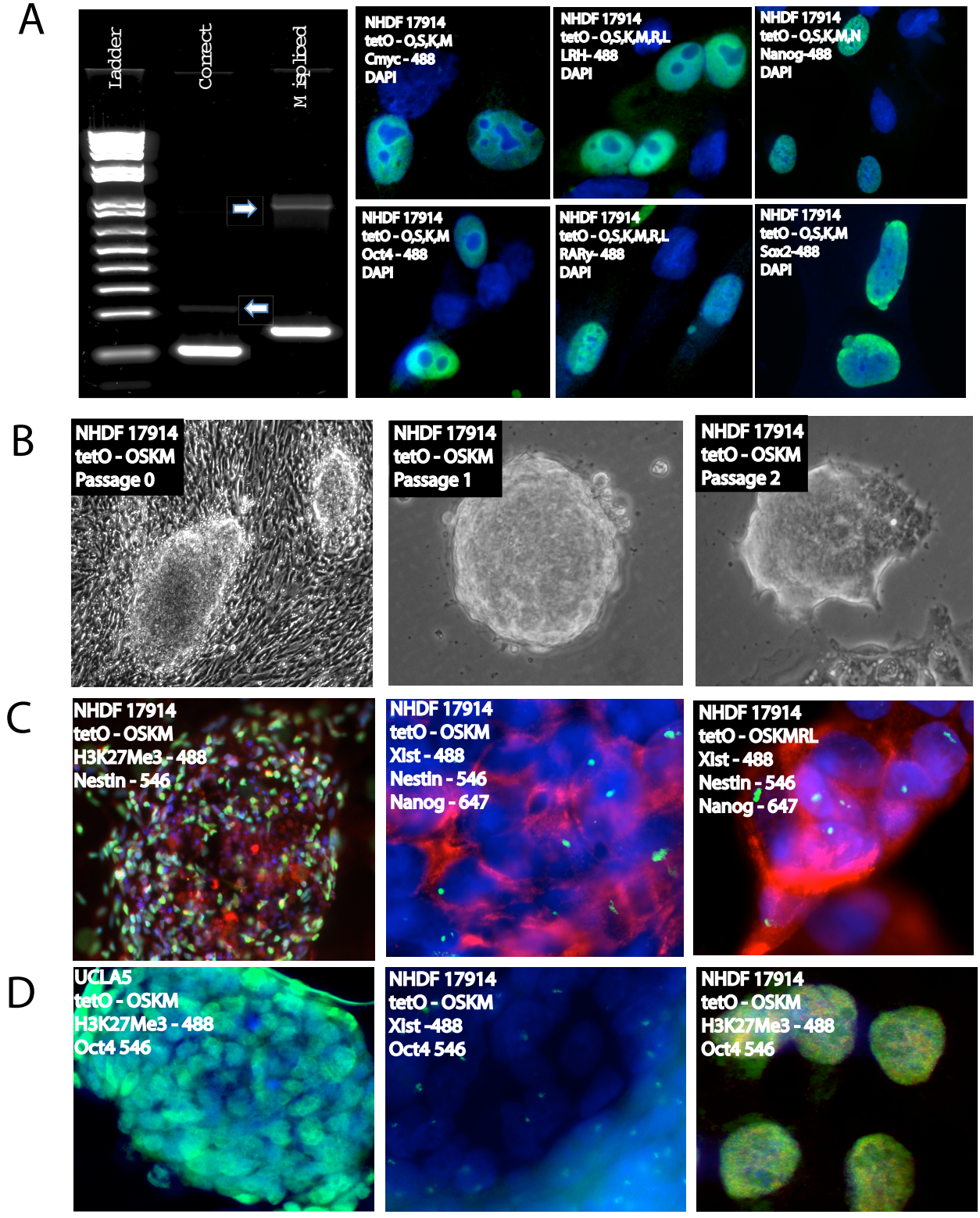
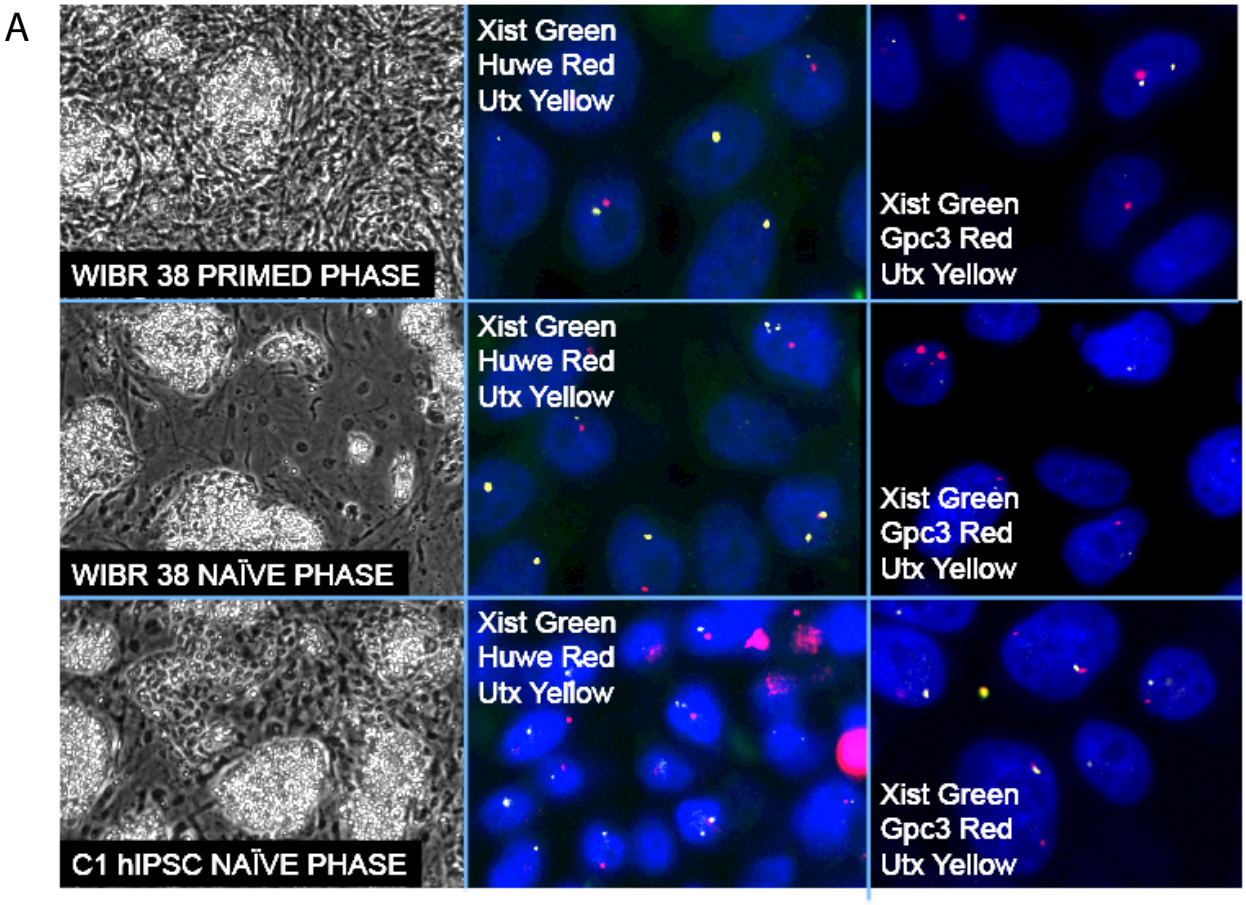
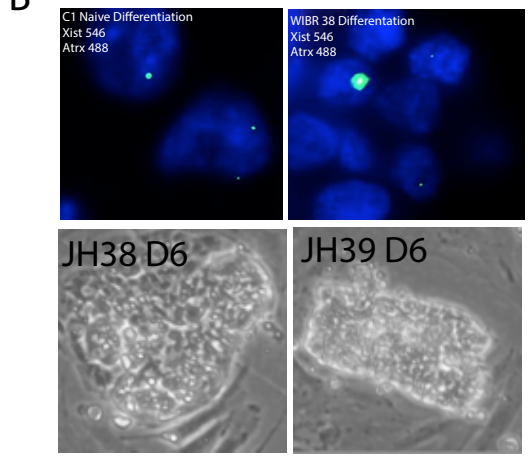

A
Ladder
Correct
M ispliced
NHDF 17914
tetO - O,S,K,M
Cmyc - 488
DAPI
NHDF 17914
tetO - O,S,K,M,R,L
LRH- 488
DAPI
NHDF 17914
tetO - O,S,K,M,N
Nanog-488
DAPI
NHDF 17914
tetO - O,S,K,M
Oct4 - 488
DAPI
NHDF 17914
tetO - O,S,K,M,R,L
RARγ- 488
DAPI
NHDF 17914
tetO - O,S,K,M
Sox2-488
DAPI
B
NHDF 17914
tetO - OSKM
Passage 0
NHDF 17914
tetO - OSKM
Passage 1
NHDF 17914
tetO - OSKM
Passage 2
C
NHDF 17914
tetO - OSKM
H3K27Me3 - 488
Nestin - 546
NHDF 17914
tetO - OSKM
Xist - 488
Nestin - 546
Nanog - 647
NHDF 17914
tetO - OSKMRL
Xist - 488
Nestin - 546
Nanog - 647
D
UCLA5
tetO - OSKM
H3K27Me3 - 488
Oct4 546
NHDF 17914
tetO - OSKM
Xist -488
Oct4 546
NHDF 17914
tetO - OSKM
H3K27Me3 - 488
Oct4 546

Figure 4-5: Naïve pluripotent stem cells from the Hanna lab fail to show evidence of an X inactivation pattern similar to human blastocysts

A) Phase and RNA FISH images of XCI patterns consistent with erosion in both primed and naïve cells from Hanna lab. B) Differentiation from the naïve state as performed by the Hanna lab showed no evidence of upregulation and coating of Xist consistent with the lack of evidence of the pre-XCI state in panel A. Bottom panels show morphology naïve lines JH38 and JH39 in version 1 naïve media 6 days after thaw C) De novo derived lines in the naïve media JH38 (male) and JH39 (female) cultured in our lab also did not show pre-implantation blastocyst patterns of XCI but rather were consistent with early passage primed human ES. RNA FISH for x-linked genes described in captioned subpanels.

Fig 4-5


A
WIBR 38 PRIMED PHASE
Xist Green
Huwe Red
Utx Yellow
Xist Green
Gpc3 Red
Utx Yellow
WIBR 38 NAÏVE PHASE
Xist Green
Huwe Red
Utx Yellow
Xist Green
Gpc3 Red
Utx Yellow
C1 hIPSC NAÏVE PHASE
Xist Green
Huwe Red
Utx Yellow
Xist Green
Gpc3 Red
Utx Yellow


B
C1 Naive Differentiation
Xist 546
Atrx 488
WIBR 38 Differentation
Xist 546
Atrx 488
JH38 D6
JH39 D6

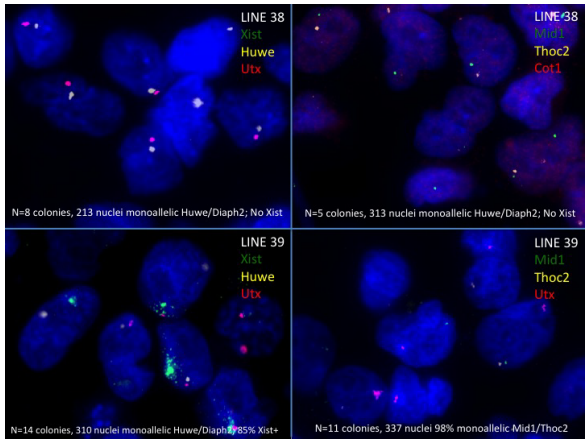
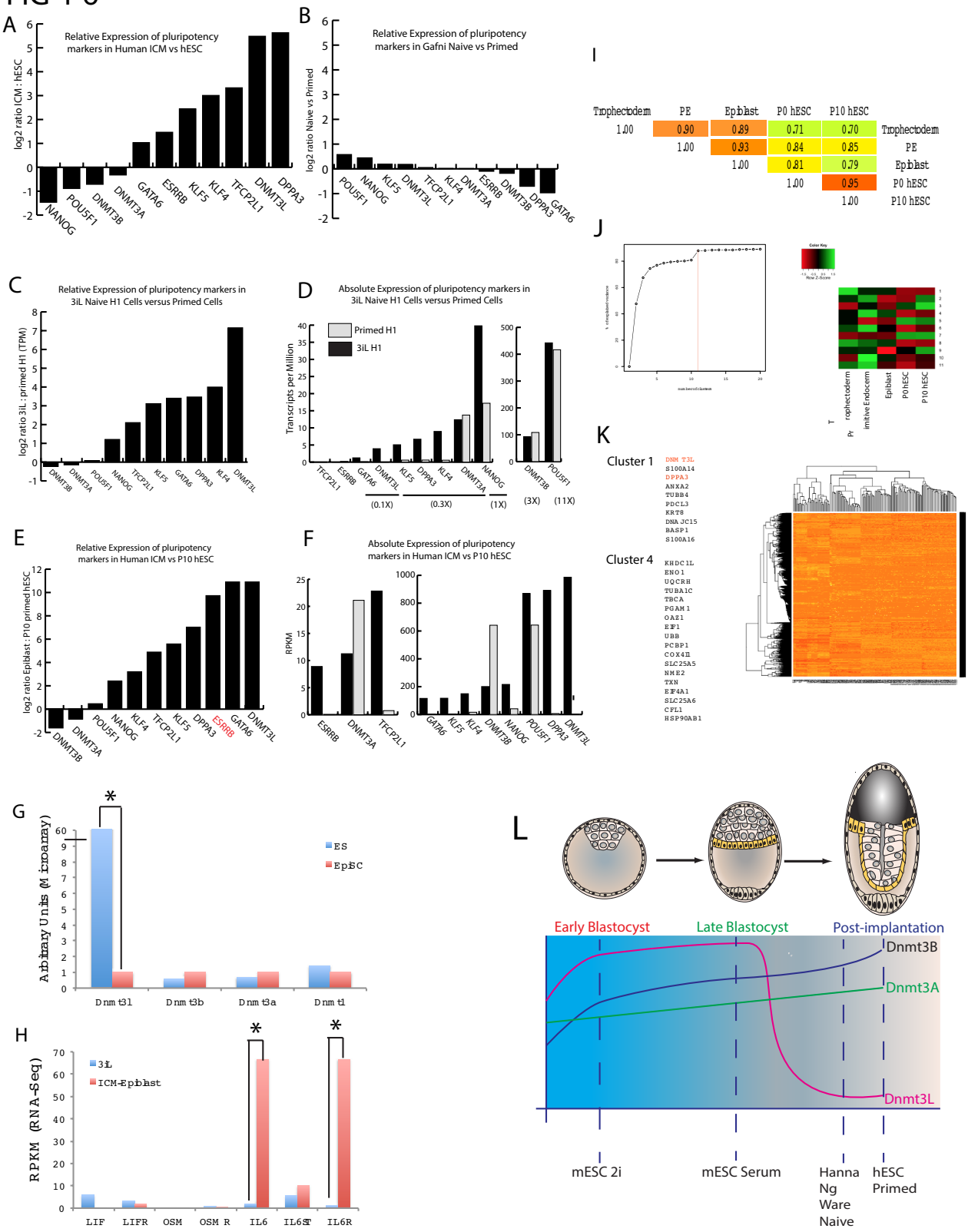

C
LINE 38
Xist
Huwe
Utx
N=8 colonies, 213 nuclei monoallelic Huwe/Diaph2; No Xist
LINE 38
Mid1
Thoc2
Cot1
N=5 colonies, 313 nuclei monoallelic Huwe/Diaph2; No Xist
LINE 39
Xist
Huwe
Utx
N=14 colonies, 310 nuclei monoallelic Huwe/Diaph2 85% Xist+
LINE 39
Mid1
Thoc2
Utx
N=11 colonies, 337 nuclei 98% monoallelic Mid1/Thoc2

Figure 4-6: The transcriptome of human naïve cells lines demonstrate that they do not acquire the gene expression program of pre-implantation blastocysts.
A) Log2 ratio of normalized microarray data for pluripotency markers and de novo methyltransferases in human blastocysts versus hESCs and B) Gafni naïve versus primed stem cells C) RNA-seq from 3iL naïve H1 cells compared to primed showing relative expression of same markers and D) absolute expression. Panel is split to show low and high expressed cells E) Relative expression of same markers from average single cell RNA-seq data in epiblast progenitor cells of the human inner cell mass F) Relative expression of the same markers, split for presentation of low and highly expressed genes G) De novo methyltransferase levels from microarray data of mouse ESCs versus mouse EpiSCs H) Class I cytokine expression levels from 3iL naïve cells versus single cell RNAseq of epiblast progenitor cells from the human inner cell mass I) Correlations between expression of all genes aligned from RNA-seq of biologically different lineages in the human blastocyst and human ES cells J) K-means clustering from an optimal number of clusters identified by iterative variance estimates K) Clusters with most significant different correlations showing naïve genes are enriched L) Model of known developmental trajectories of the de novo methyltransferases and speculated placement of reported naïve human stem cells.

FIG 4-6


A
Relative Expression of pluripotency markers in Human ICM vs hESC
log2 ratio ICM : hESC
NANOG
POU5F1
DNMT3B
DNMT3A
GATA6
ESRRB
KLF5
KLF4
TFCP2L1
DNMT3L
DPPA3
B
Relative Expression of pluripotency markers in Gafni Naive vs Primed
log2 ratio Naive vs Primed
POU5F1
NANOG
KLF5
DNMT3L
TFCP2L1
KLF4
DNMT3A
ESRRB
DNMT3B
DPPA3
GATA6
I
Trophectoderm PE Epiblast P0 hESC P10 hESC
1.00 0.90 0.89 0.71 0.70 Trophectoderm
1.00 0.93 0.84 0.85 PE
1.00 0.81 0.79 Epiblast
1.00 0.95 P0 hESC
1.00 P10 hESC
C
Relative Expression of pluripotency markers in 3iL Naive H1 Cells versus Primed Cells
log2 ratio 3iL : primed H1 (TPM)
DNMT3B
DNMT3A
POU5F1
NANOG
TFCP2L1
KLF5
GATA6
DPPA3
KLF4
DNMT3L
D
Absolute Expression of pluripotency markers in 3iL Naive H1 Cells versus Primed Cells
Primed H1
3iL H1
Transcripts per Million
TFCP2L1
ESRRB
GATA6
DNMT3L
KLF5
DPPA3
KLF4
DNMT3A
NANOG
DNMT3B
POU5F1
(0.1X)
(0.3X)
(1X)
(3X)
(11X)
J
Color Key
Row Z-Score
Trophectoderm
Primitive Endoderm
Epiblast
P0hESC
P10 hESC
K
Cluster 1
DNMT3L
S100A14
DPPA3
ANXA2
TUBB4
PDCL3
KRT8
DNAJC15
BASP1
S100A16
Cluster 4
KHDC1L
ENO1
UQCRH
TUBA1C
TBCA
PGAM1
OAZ1
EIF1
UBB
PCBP1
COX4I1
SLC25A5
NME2
TXN
EEF4A1
SLC25A6
CFL1
HSP90AB1
E
Relative Expression of pluripotency markers in Human ICM vs P10 hESC
log2 ratio Epiblast : P10 primed hESC
DNMT3B
DNMT3A
POU5F1
NANOG
KLF4
TFCP2L1
KLF5
DPPA3
ESRRB
GATA6
DNMT3L
F
Absolute Expression of pluripotency markers in Human ICM vs P10 hESC
RPKM
ESRRB
DNMT3A
TFCP2L1
GATA6
KLF5
KLF4
DNMT3B
NANOG
POU5F1
DPPA3
DNMT3L
G
Arbitrary Units (Microarray)
ES
EpiSC
Dnmt3l
Dnmt3b
Dnmt3a
Dnmt1
H
RPKM (RNA-Seq)
3iL
ICM-Epiblast
LIF
LIFR
OSM
OSM R
IL6
IL6ST
IL6R
L
Early Blastocyst
Late Blastocyst
Post-implantation
Dnmt3B
Dnmt3A
Dnmt3L
mESC 2i
mESC Serum
Hanna Ng Ware Naive
hESC Primed

REFERENCES:

1. Evans MJ, Kaufman MH. Establishment in culture of pluripotential cells from mouse embryos. Nature. 1981 Jul 9;292(5819):154-6.

2. Martin GR. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci U S A. 1981 Dec;78(12):7634-8.

3. Ying QL, Wray J, Nichols J, Batlle-Morera L, Doble B, Woodgett J, Cohen P, Smith A. The ground state of embryonic stem cell self-renewal. Nature. 2008 May 22;453(7194):519-23.

4. Rossant J, Gardner RL, Alexandre HL. Investigation of the potency of cells from the postimplantation mouse embryo by blastocyst injection: a preliminary report. J Embryol Exp Morphol. 1978 Dec;48:239-47.

5. Brons, IG, Smithers LE, Trotter MW, Rugg-Gunn P, Sun B, Chuva de Sousa Lopes SM, Howlett SK, Clarkson A, Ahrlund-Richter L, Pedersen RA, Vallier L. Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature. 2007 Jul 12;448(7150):191-5.

6. Tesar, PJ, Chenoweth JG, Brook FA, Davies TJ, Evans EP, Mack DL, Gardner RL, McKay RD. New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature. 2007 Jul 12;448(7150):196-9.

7. Ficz G, Hore TA, Santos F, Lee HJ, Dean W, Arand J, Krueger F, Oxley D, Paul YL, Walter J, Cook SJ, Andrews S, Branco MR, Reik W. FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. Cell Stem Cell. 2013 Sep 5;13(3):351-9.

8. Habibi E, Brinkman AB, Arand J, Kroeze LI, Kerstens HH, Matarese F, Lepikhov K, Gut M, Brun-Heath I, Hubner NC, Benedetti R, Altucci L, Jansen JH, Walter J, Gut IG, Marks H, Stunnenberg HG. Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. Cell Stem Cell. 2013 Sep 5;13(3):360-9

9. Hackett JA, Dietmann S, Murakami K, Down TA, Leitch HG, Surani MA. Synergistic Mechanisms of DNA Demethylation during Transition to Ground-State Pluripotency. Stem Cell Reports. 2013 Dec 17;1(6):518-31.

10. Blaschke K, Ebata KT, Karimi MM, Zepeda-Martínez JA, Goyal P, Mahapatra S, Tam A, Laird DJ, Hirst M, Rao A, Lorincz MC, Ramalho-Santos M. Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. Nature. 2013 Aug 8;500(7461):222-6.

11. Marks H, Kalkan T, Menafra R, Denissov S, Jones K, Hofemeister H, Nichols J, Kranz A, Stewart AF, Smith A, Stunnenberg HG. The transcriptional and epigenomic foundations of ground state pluripotency. Cell. 2012 Apr 27;149(3):590-604.

12. Morgani SM, Canham MA, Nichols J, Sharov AA, Migueles RP, Ko MS, Brickman JM. Totipotent embryonic stem cells arise in ground-state culture conditions. Cell Rep. 2013 Jun 27;3(6):1945-57.

13. Beddington, RS, Robertson EJ. An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo. Development. 1989 Apr;105(4):733-7.

14. Canham MA, Sharov AA, Ko MS, Brickman JM. Functional heterogeneity of embryonic stem cells revealed through translational amplification of an early endodermal transcript. PLoS Biol. 2010 May 25;8(5):e1000379.

15. Lallemand, Y, Brûlet P. An in situ assessment of the routes and extents of colonisation of the mouse embryo by embryonic stem cells and their descendants. Development. 1990 Dec;110(4):1241-8.

16. Macfarlan, TS, Gifford WD, Driscoll S, Lettieri K, Rowe HM, Bonanomi D, Firth A, Singer O, Trono D, Pfaff SL. Embryonic stem cell potency fluctuates with endogenous retrovirus activity. Nature. 2012 Jul 5;487(7405):57-63.

17. H. Suemori, Y. Kadodawa, K. Goto, I. Araki, H. Kondoh, N. Nakatsuji. A mouse embryonic stem cell line showing pluripotency of differentiation in early embryos and ubiquitous beta-galactosidase expression. Cell Differ. Dev., 29 (1990), pp. 181–186.

18. Buehr, M., Meek, S., Blair, K., Yang, J., Ure, J., Silva, J., McLay, R., Hall, J., Ying, Q.L., and Smith, A. Capture of authentic embryonic stem cells from rat blastocysts. *Cell*. 2008; 135: 1287–1298’

19. Men H, Bauer BA, Bryda EC. Germline transmission of a novel rat embryonic stem cell line derived from transgenic rats. Stem Cells Dev. 2012 Sep 20;21(14):2606-12.

20. Hong J, He H, Weiss ML. Derivation and characterization of embryonic stem cells lines derived from transgenic Fischer 344 and Dark Agouti rats. Stem Cells Dev. 2012 Jun 10;21(9):1571-86.

21. Meek S, Buehr M, Sutherland L, Thomson A, Mullins JJ, Smith AJ, Burdon T. Efficient gene targeting by homologous recombination in rat embryonic stem cells. PLoS One. 2010 Dec 3;5(12):e14225

22. Kawamata M, Ochiya T. Generation of genetically modified rats from embryonic stem cells. Proc Natl Acad Sci U S A. 2010 Aug 10;107(32):14223-8.

23. Li P, Tong C, Mehrian-Shai R, Jia L, Wu N, Yan Y, Maxson RE, Schulze EN, Song H, Hsieh CL, Pera MF, Ying QL. Germline competent embryonic stem cells derived from rat blastocysts. Cell. 2008 Dec 26;135(7):1299-310.

24. Li W, Li X, Li T, Jiang MG, Wan H, Luo GZ, Feng C, Cui X, Teng F, Yuan Y, Zhou Q, Gu Q, Shuai L, Sha J, Xiao Y, Wang L, Liu Z, Wang XJ, Zhao XY, Zhou Q. Genetic modification and screening in rat using haploid embryonic stem cells. Cell Stem Cell. 2014 Mar 6;14(3):404-14.

25. Kawase, E, Suemori H, Takahashi N, Okazaki K, Hashimoto K, Nakatsuji N. Strain difference in establishment of mouse embryonic stem (ES) cell lines. Int J Dev Biol. 1994 Jun;38(2):385-90

26. Silva, J, Barrandon O, Nichols J, Kawaguchi J, Theunissen TW, Smith A. Promotion of reprogramming to ground state pluripotency by signal inhibition. PLoS Biol. 2008 Oct 21;6(10):e253. doi: 10.1371/journal.pbio.0060253.

27. Hanna, J, Markoulaki S, Mitalipova M, Cheng AW, Cassady JP, Staerk J, Carey BW,Lengner CJ, Foreman R, Love J, Gao Q, Kim J, Jaenisch R. Metastable pluripotent states in NOD-mouse-derived ESCs. Cell Stem Cell. 2009 Jun 5;4(6):513-24.

28. Reinholdt, LG, Howell GR, Czechanski AM, Macalinao DG, Macnicoll KH, Lin CS, Donahue LR, John SW. Generating embryonic stem cells from the inbred mouse strain DBA/2J, a model of glaucoma and other complex diseases. PLoS One. 2012;7(11):e50081.

29. Nichols, J, Jones K, Phillips JM, Newland SA, Roode M, Mansfield W, Smith A, Cooke A. Validated germline-competent embryonic stem cell lines from nonobese diabetic mice. Nat Med. 2009 Jul;15(7):814-8.

30. Hanna J, Cheng AW, Saha K, Kim J, Lengner CJ, Soldner F, Cassady JP, Muffat J, Carey BW, Jaenisch R. Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. Proc Natl Acad Sci U S A. 2010 May 18;107(20):9222-7.

31. 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

32. Lengner CJ, Gimelbrant AA, Erwin JA, Cheng AW, Guenther MG, Welstead GG, Alagappan R, Frampton GM, Xu P, Muffat J, Santagata S, Powers D, Barrett CB, Young RA, Lee JT, Jaenisch R, Mitalipova M. Derivation of pre-X inactivation human embryonic stem cells under physiological oxygen concentrations. Cell. 2010

May 28;141(5):872-83.

33. Chan YS, Göke J, Ng JH, Lu X, Gonzales KA, Tan CP, Tng WQ, Hong ZZ, Lim YS, Ng HH. Induction of a human pluripotent state with distinct regulatory circuitry that resembles preimplantation epiblast. Cell Stem Cell. 2013 Dec 5;13(6):663-75.

34. Wutz A. Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. Nat Rev Genet. 2011 Jul 18;12(8):542-53.

35. Silva SS, Rowntree RK, Mekhoubad S, Lee JT. X-chromosome inactivation and epigenetic fluidity in human embryonic stem cells. Proc Natl Acad Sci U S A. 2008 Mar 25;105(12):4820-5.

36. Hockemeyer D, et al. A drug-inducible system for direct reprogramming of human somatic cells to pluripotency. Cell Stem Cell. 2008;3:346–353.

37. Guo G, et al. Klf4 reverts developmentally programmed restriction of ground state pluripotency. Development. 2009;136:1063–1069.

38. Carey BW, Markoulaki S, Hanna JH, Faddah DA, Buganim Y, Kim J, Ganz K, Steine EJ, Cassady JP, Creyghton MP, Welstead GG, Gao Q, Jaenisch R. Reprogramming factor stoichiometry influences the epigenetic state and biological properties of induced pluripotent stem cells. Cell Stem Cell. 2011 Dec 2;9(6):588-98.

39. Zhang F, Frost AR, Blundell MP, Bales O, Antoniou MN, Thrasher AJ. A ubiquitous chromatin opening element (UCOE) confers resistance to DNA methylation-mediated silencing of lentiviral vectors. Mol Ther. 2010 Sep;18(9):1640-9.

40. Wang W, Yang J, Liu H, Lu D, Chen X, Zenonos Z, Campos LS, Rad R, Guo G, Zhang S, Bradley A, Liu P. Rapid and efficient reprogramming of somatic cells to induced pluripotent stem cells by retinoic acid receptor gamma and liver receptor homolog 1. Proc Natl Acad Sci U S A. 2011 Nov 8;108(45):18283-8.

41. Li W, Sun W, Zhang Y, Wei W, Ambasudhan R, Xia P, Talantova M, Lin T, Kim J, Wang X, Kim WR, Lipton SA, Zhang K, Ding S. Rapid induction and long-term self-renewal of primitive neural precursors from human embryonic stem cells by small molecule inhibitors. Proc Natl Acad Sci U S A. 2011 May 17;108(20):8299-304.

42. Vassena R, Boué S, González-Roca E, Aran B, Auer H, Veiga A, Izpisua Belmonte JC. Waves of early transcriptional activation and pluripotency program initiation during human preimplantation development. Development. 2011 Sep;138(17):3699-709.

42.1 Chan YS, Göke J, Ng JH, Lu X, Gonzales KA, Tan CP, Tng WQ, Hong ZZ, Lim YS, Ng HH. Induction of a human pluripotent state with distinct regulatory circuitry that resembles preimplantation epiblast. Cell Stem Cell. 2013 Dec 5;13(6):663-75.

43. Yan L, Yang M, Guo H, Yang L, Wu J, Li R, Liu P, Lian Y, Zheng X, Yan J, Huang J, Li M, Wu X, Wen L, Lao K, Li R, Qiao J, Tang F. Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. Nat Struct Mol Biol. 2013 Sep;20(9):1131-9.

44. Roode M, Blair K, Snell P, Elder K, Marchant S, Smith A, Nichols J. Human hypoblast formation is not dependent on FGF signalling. Dev Biol. 2012 Jan 15;361(2):358-63.

45. Kuijk EW, van Tol LT, Van de Velde H, Wubbolts R, Welling M, Geijsen N, Roelen BA. The roles of FGF and MAP kinase signaling in the segregation of the epiblast and hypoblast cell lineages in bovine and human embryos. Development. 2012 Mar;139(5):871-82.

46. Rakemann T, Niehof M, Kubicka S, Fischer M, Manns MP, Rose-John S, Trautwein C. The designer cytokine hyper-interleukin-6 is a potent activator of STAT3-dependent gene transcription in vivo and in vitro. J Biol Chem. 1999 Jan 15;274(3):1257-66.

47. Cooper AR, Patel S, Senadheera S, Plath K, Kohn DB, Hollis RP. Highly efficient large-scale lentiviral vector concentration by tandem tangential flow filtration. J Virol Methods. 2011 Oct;177(1):1-9.

48. Tchieu J, Kuoy E, Chin MH, Trinh H, Patterson M, Sherman SP, Aimiuwu O, Lindgren A, Hakimian S, Zack JA, Clark AT, Pyle AD, Lowry WE, Plath K. Female human iPSCs retain an inactive X chromosome. Cell Stem Cell. 2010 Sep 3;7(3):329-42

49. Patro R, Mount SM, Kingsford C. Sailfish enables alignment-free isoform quantification from RNA-seq reads using lightweight algorithms. Nat Biotechnol. 2014 May;32(5):462-4.

CHAPTER 5

CONCLUSIONS

This dissertation investigates the state of the X chromosome in female human pluripotent stem cells and focuses on the ability of the X chromosome to place these cells in their developmental context. Published work has largely held that the state of the X chromosome in human pluripotent stem cells mirrors that of mouse development however we find that this is not the case and the human system of X inactivation must follow dramatically different mechanisms of initiation. Furthermore, we identify that although human embryonic stem cells begin as an *XIST* expressing XaXi state, this can rapidly change in usual culture conditions. Finally the most surprising findings surround the idea that X inactivation is not required for in vitro differentiation and that the state of the X chromosome established in pluripotency persists in differentiation.

Chapter 2 focused on the published literature regarding the mechanisms surrounding X inactivation in mouse models and the dearth of information in the human system. The mechanisms surrounding initiation of X inactivation have been thoroughly studied in mice and thus serve as a system to interrogate the function of several mediators of this process. In humans, however, the lack of an appropriate system available to study x inactivation hampers the ability to understand this mechanism. Furthermore, the confusion surrounding classes of X inactivation, as well as the epigenetic drift that accumulates on human pluripotent stem cells in culture further creates a difficult system to study. Approaches to identify a naïve state, for which the cells would be considered pre-XCI, have not been successfully reproduced and therefore many questions regarding this process are left unanswered.

Chapter 3 focused on the state of XCI in primed pluripotent stem cells derived at UCLA and at the WiCell. We find that many states of XCI exist and most intriguingly that these states of XCI are maintained upon differentiation. This substantiates the concept that human embryonic stem cells are capturing a post-implantation state of pluripotency that has already passed the stage of X inactivation. We identify that there exists multiple mechanisms to appreciate the state of XCI, and use these orthogonal assays to fortify our understanding of these states. Lastly, we show that human ES derivation itself undergoes a transition from pre-XCI in the embryo to post-XCI in the resulting line, and that this process is subject to producing multiple endpoints. We postulate that mixed lines in terms of XCI derived in other derivation centers is due to the fact they passage entire derivation cultures, where here at UCLA we subclone and thus enrich for specific XCI states. This procedure helped to identify that the current model of XCI in human ES is mistaken and we propose an alternative model that encompasses our data and places the existing one in congruent context.

Chapter 4 showed that the naïve state of human pluripotency is yet to be acquired. Previous reports claiming this state were largely irreproducible in my experiments or simply did not reflect the developmental state argued. These attempts to replicate the published work were not simply due to technical issues but instead demonstrated that all attempts thus far to stabilize human naïve pluripotency have not achieved an epigenetic nor transcriptomic similarity to the pre-implantation blastocyst. Unfortunately, given the lack of available data on a fine dissection of stages of human development the current claims for human naïve pluripotency cannot be placed within a known sequence of development, and thus must optimistically placed after true naïve

pluripotency but perhaps slightly closer to it than current primed pluripotent human stem cells.

In conclusion this dissertation has investigated states of X chromosome inactivation in human pluripotent stem cells and human development. These findings not only help clarify the model of X inactivation in human pluripotent stem cells but also demonstrates that studying X inactivation can provide an incredible insight into the epigenetic makeup of pluripotent stem cells. This work will hopefully identify strengths and weaknesses to the approach for capturing human naïve pluripotency and serves to rationalize further studies on X inactivation in human pluripotent stem cells.