microRNAs and metabolites in naïve to primed human embryonic stem cell transition

Hans Henrik Bertil Sperber

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2014

Reading Committee
Hannele Ruohola Baker, Chair
Carl Anthony Blau
Gabriele Varani
Dustin Maly

Program Authorized to Offer Degree:
Chemistry

University of Washington

Abstract

microRNAs and metabolites in naïve to primed human embryonic stem cell transition

Hans Henrik Bertil Sperber

Chair of the Supervisory Committee:

Professor Hannele Ruohola-Baker

Department of Biochemistry

This dissertation research is focused on the small molecules of the cell: metabolites and miRNAs. The purpose is to gain a deeper understanding of how they control the cell by investigating the role of secondary structure during miRNA biogenesis and of the changes that occur in the metabolome during naïve to primed human embryonic stem cell (hESC) transition.

The first chapter of the thesis focuses on miRNA regulation: Through analysis of miRNA secondary structure and miRNA expression levels in cells with different levels of the enzyme Drosha I have found that miRNAs with mismatches in the precursor region of the miRNA, 9-12 nucleotides from the Drosha cutting site show a higher sensitivity to changes in Drosha levels than miRNAs that lack mismatches in said regions. Through mutagenesis experiments I have shown that by altering the miRNA secondary structure its sensitivity to changes in Drosha levels

can be changed. This shows how the cell can selectively regulate a group of miRNAs relative to another by changing Drosha expression and may explain the impact of point mutations outside of the seed region of certain miRNAs. miRNA therapies are also being developed, where the small size of miRNAs makes them excellent candidates for gene therapy. The results of this research can be applied to miRNA therapies by increasing the potency of select miRNAs in their target tissues through structural modifications. In cases where the target tissue has different levels of Drosha compared to surrounding tissues (such as in many types of cancer) the miRNA can also be designed to minimize side effects in surrounding healthy tissues.

The second chapter of the thesis focuses on the metabolome in human embryonic stem cells: With the use of mass spectrometry combined with gene expression data I have analyzed the metabolome of naïve and primed hESCs. I show that the two developmental stages have distinct profiles in metabolite set up and I have identified pathways that show a change in activity during naïve to primed hESC transition. I show that tryptophan degradation is increased in primed cells leading to an accumulation of the tryptophan degradation product kynurenine, which is known to inhibit the antitumor response of the immune system by activting the transcription factor AhR and could be a mechanism which prevents rejection of the recently implanted embryo. I also show that primed hESCs display a decrease in activity of the enzyme NNMT, which uses the metabolite S-adenosyl methionine (SAM) to convert nicotinamide into 1-methyl nicotinamide. The decrease of NNMT activity makes SAM available as a substrate for other methylation mechanisms and increases H3K27me3 methylation. When knocking down NNMT the H3K27me3 marks are increased, supporting the hypothesis that the metabolome regulates the epigenetic landscape during naïve to primed hESC transition.

Supplemental file includes Supplementary Table 1 for Chapter II.

Table of contents

Materials and Methods

Introduction

Human embryonic stem cells (hESCs) have the capacity to regenerate and differentiate into any tissue type in the human body and show great potential for therapeutic applications such as growing tissues or even complete organs *in vitro* which can be used for transplantation and/or repair of damaged organs. With the recent development of induced pluripotent stem cells these therapies could also be performed by reprogramming the patients own cells (Takahashi and Yamanaka, 2006), which would help minimizing the risk of an autoimmune response and rejection of the transplant.

Until recently all stem cell lines derived from human represented a state equivalent to that of the postimplantation embryo (primed), but recently we and other groups have created stable hESC lines in a pre-implantation (naïve) state (Chan et al., 2013;Gafni et al., 2013;Theunissen et al., 2014;Valamehr et al., 2014;Ware et al., 2014). Naïve hESCs are more pluripotent than primed and could therefore improve current and future stem cell therapies (Gafni et al., 2013;Ware et al., 2014). However, many things are unknown about the function of naïve and primed hESCs, particularly the details of what happens during the transition from one state to the other.

Characterization of cell lines can be done using a wide range of technologies. Typically the genome is studied to get information about variability in the DNA sequence. The transcriptome is analyzed for information about the RNA that is being transcribed. The proteome can also be mapped to show how much of each protein is ultimately present in the cell. Recent developments of mass spectrometry has now made it possible to gain detailed information about the metabolome of a cell, i.e. the levels of small molecules such as nucleotides, amino acids, lipids, as well as metabolic intermediates and degradation products, which provides information about the activity of various metabolic pathways. The study of the metabolome during naïve to primed hESC transition is of particular interest since metabolite levels are one of the first things that changes during cell development (Bracha et al., 2010;Yanes et al., 2010;Folmes et al., 2011;Greer et al., 2012;Panopoulos et al., 2012;Rafalski et al., 2012;Zhou et al., 2012;Mathieu et al., 2014), but also since identification of metabolites capable of stabilizing cell states and determining cell fate could provide for simpler and cheaper methods to reprogram cells, which currently requires the use of viruses to overexpress a set of reprogramming factors.

miRNA is another group of small molecules that has been shown to be important for proper stem cell function (Hatfield et al., 2005) and acts by fine tuning translation of mRNA into protein by binding to the 3'UTR (Kuppusamy et al., 2013). The precursor molecules of the mature miRNA have a distinct secondary structure which is frequently mismatched. Because of their high impact on cell function combined with their small size, miRNAs are good candidates for gene therapies. Gaining a deeper understanding of the mechanisms of miRNA biogenesis and how to maximize the potency of select miRNAs will help in the development of miRNA tools and clinical applications.

In this dissertation I will investigate the changes that occur in the metabolome during naïve to primed hESC transition as well as how the miRNA secondary structure affects its sensitivity to changes in Drosha levels. Together, the results provide further insight into the regulatory ability of metabolites and miRNAs.

Chapter I: miRNA sensitivity to Drosha levels correlates with pre-miRNA secondary structure

Abstract

microRNAs (miRNAs) are crucial for cellular development and homeostasis. In order to better understand regulation of miRNA biosynthesis, we studied cleavage of primary-miRNAs by Drosha. While Drosha knockdown triggers an expected decrease of many mature miRNAs in human embryonic stem cells (hESC), a subset of miRNAs are not reduced. Statistical analysis of miRNA secondary structure and fold change of expression in response to Drosha knockdown showed that absence of mismatches in the central region of the hairpin, 5 and 9-12nt from the Drosha cutting site conferred decreased sensitivity to Drosha knockdown. This suggests that, when limiting, Drosha processes miRNAs without mismatches more efficiently than mismatched miRNAs. This is important because Drosha expression changes over cellular development and the fold change of expression for miRNAs with mismatches in the central region correlates with Drosha levels. To examine the biochemical relationship directly, we overexpressed structural variants of miRNA-145, -137, -9 and -200b in HeLa cells with and without Drosha knockdown; for these miRNAs, elimination of mismatches in the central region increased, and addition of mismatches decreased their expression in an *in vitro* assay or in cells with low Drosha expression. Change in Drosha expression can be a biologically relevant mechanism by which eukaryotic cells control miRNA profiles. This phenomenon may explain the impact of point mutations outside the seed region of certain miRNAs.

Introduction

MicroRNAs (miRNA) are a group of short, non-coding RNAs that bind target mRNAs by incomplete complementarity to either inhibit the translation, or reduce the stability of target mRNAs (Kim and Nam, 2006). miRNAs are involved with the onset of various diseases, immunoregulation, angiogenesis, neural growth and stem cell renewal and maintenance (Bartel, 2004;Hatfield et al., 2005;Stark et al., 2005;Cullen, 2006;Mattick and Makunin, 2006;Shcherbata et al., 2006;Kuehbacher et al., 2007;Stadler and Ruohola-Baker, 2008;Anokye-Danso et al., 2011). Currently 1,872 human microRNAs have been identified and many of these are evolutionarily conserved (Griffiths-Jones et al., 2006).

miRNA biogenesis is regulated on multiple levels (Kuppusamy et al., 2013). Individual miRNA expression levels vary between tissues (Lagos-Quintana et al., 2001), and while transcription of primary-miRNAs (pri-miRNAs) is regulated by transcription factors, e.g. c-Myc (O'Donnell et al., 2005), this variability is found even within polycistronic miRNAs transcribed as a single unit (Davis and Hata, 2009). pri-miRNAs are cleaved by the Microprocessor, which principally consists of Drosha and DGCR8 (Lee et al., 2003;Gregory et al., 2004). Canonically, Drosha and DGCR8 bind pri-miRNAs, and Drosha cleaves pri-miRNAs 11bp from the base of the stem and liberates a structure known as the precursor microRNA (pre-miRNA), which is subsequently exported to the cytoplasm and processed into mature miRNA (Zeng et al., 2005). The cleavage by the Microprocessor complex is regulated by several components which are part of the Microprocessor, such as the DEAD-box helicases p68 and p72, which have been proposed to stabilize the Microprocessor complex (Fukuda et al., 2007). The Microprocessor can include additional proteins in a signal-dependent fashion, such as SMAD proteins, which have been found to selectively upregulate certain miRNAs, e.g. miR-21 (Davis et al., 2008). Furthermore, autoregulation and Ago involvement has been observed at the let7 Drosha processing step (Zisoulis et al., 2012). A group of miRNAs named mirtrons are processed by the spliceosome (Ruby et al., 2007), making them independent of the canonical miRNA biogenesis pathway where the pri-miRNA is cleaved by the Microprocessor.

In this paper we describe Drosha expression levels as a mechanism to selectively regulate a group of miRNAs with a particular secondary structure. The fold change of expression of miRNAs with mismatches located 5 and 9-12 nucleotides from the Drosha cutting site is greater than with more rigid miRNAs, when Drosha levels are low. This can be observed in cells with shRNA knockdown of Drosha (DroshaKD), in tissues with different Drosha levels and in *in vitro* assays and suggests a mechanism where the cell can selectively alter the levels of mismatched miRNAs through changes in Drosha expression.

Results

Differences in change of miRNA expression due to Drosha knockdown in hESC (H1) correlate with differences in predicted pri-miRNA secondary structures

Previous research in the laboratory used H1 human embryonic stem cells with DroshaKD to study the response to a general decrease of miRNAs in stem cells (Bar et al., 2008;Qi et al., 2009;Stadler et al., 2010). As expected, qPCR of 220 microRNAs in Drosha KD hESCs showed a general decrease of miRNA expression. However, a group of miRNAs was still processed efficiently, despite the 3-fold reduction of Drosha expression (Fig.1). It seems logical that mirtrons (which bypass Drosha processing) would be unaffected by Drosha knockdown, however, few mirtrons are known (Berezikov et al., 2007) and none of the known mirtrons are in the 220 miRNAs assayed by qPCR. mirtrons are generally located within small introns, but no statistically significant correlation could be verified between genomic location and fold change of miRNA expression (p=0.27); the mean fold change was 2.8 for intergenic miRNAs, and 2.4 for the intronic miRNAs (Fig.2).

Given that bulges in the stem of RNA hairpins seem to prevent processing by the miRNA pathway in general, and that others have shown Drosha to differentially process miRNAs based on structural features (Han et al., 2006;Ritchie et al., 2007;Feng et al., 2011), it seems logical that differences in the number of mismatches in miRNA hairpins would alter their interaction with the Microprocessor, and could explain the variation of fold change in response to Drosha knockdown. In order to compare the three-dimensional shape of the pri-miRNAs in our dataset, we developed a novel method of assigning position within the miRNA hairpin in order to give a better representation of the mismatch locations relative to each other and the Drosha cutting site in three dimensional space (Fig.3A). Position assignment and mismatch counting was performed by defining the location of each mismatch in terms of distance from the Drosha cutting site using predicted secondary structure, counting asymmetric mismatches as the length of the shorter strand; this is different from previously described methods to calculate mismatch positions in pri-miRNAs, where asymmetry was not taken into account (Han et al., 2006).

The secondary structures for the 220 miRNAs from the H1 dataset were calculated using the ViennaRNA package (Hofacker et al., 1994), and exact locations of cutting sites were found through miRBase. The dataset contained 220 miRNAs, and we selected 202 by sorting out any miRNAs with ambiguous secondary structure predictions. The dataset was split in half based on fold change (Fig.1, arrow), and the sums of mismatches at each nucleotide position were compared between the two groups. The summation revealed that miRNAs with a high fold change when Drosha levels were reduced had more mismatches in position(s) 5 and 10-12 than miRNAs with low fold change, and that these differences were statistically significant (Fig.3B,C). A knockdown of Dicer did not result in the same trend (Suppl.Fig.1) In addition p-values were calculated for the quartiles at the extreme ends of the DroshaKD dataset, with a sample size extrapolated to that of the total data set (Fig.3C). For subsequent experiments and analysis position 9 was also included since it follows the same trend as positions 10-12. This takes into account any lack of specificity in our method of assigning positions.

Correlation of Drosha expression and miRNA secondary structure profile in mature tissues

Having shown that pri-miRNA processing by Drosha is affected by the pri-miRNA secondary structure in cell culture we also investigated if this mechanism is used *in vivo* as a mechanism to selectively regulate groups of miRNAs based on secondary structure. Since Drosha expression alters the expression of miRNAs with bulges in the central region more than miRNAs without (Fig.3), and because of the crucial role of miRNAs for cellular biology, we were interested in determining whether or not Drosha expression levels vary across cell type or according to developmental stage in mammals. To find naturally occurring examples of Drosha variation, we used the Gene Expression Omnibus (GEO) to find microarray data of relevant gene expression in mature and differentiating mouse tissues.

From the GEO expression datasets GDS3052 and -3142, we extracted and analyzed expression data for Drosha and common housekeeping genes in multiple mouse tissues. In both datasets the expression of Drosha is dramatically lower in liver and submaxillary/salivary gland (4-10 fold) than in brain tissues when normalized to EEF2 (Fig.4A-B). To test whether the difference in Drosha expression was dependent on the housekeeping gene used for normalization, we analyzed the fold change of expression between liver and brain tissues for five housekeeping genes; EEF2, RPL13A, SDHA, TBP, and UBC. The data revealed that the expression change for Drosha is greater than variation among known housekeeping genes (Fig.4C-D), showing that Drosha is differentially expressed regardless of the housekeeping gene used for normalization.

We validated the differential expression of Drosha in mouse brain and liver by qPCR, normalized to UBC and verified that Drosha expression was 5.6 fold higher in the brain than in the liver (Fig.4E; n=3, $p<0.01$). Furthermore, Western blotting revealed 9.6 or 3.2 fold higher Drosha protein expression in brain than in liver when normalized to total protein or to β-actin expression, respectively (Coomassie Blue staining or to β-actin Western; Fig.4F,G; Suppl.Fig.2)

Having validated that Drosha levels are higher in brain than in liver, we were interested in analyzing if miRNA expression varied according to the number of mismatches in the 9-12nt region of the miRNAs in the two tissues. We analyzed miRNA expression data from mouse brain and liver (Takada et al., 2006) and found that miRNAs with mismatches in positions 5 and 9-12 are enriched in brain but not in liver and that the difference was larger the more mismatches present (Fig.4H), supporting the hypothesis that miRNAs with central mismatches are not processed efficiently when Drosha is limiting. This trend for miRNA expression did not correlate with mismatches in the other positions (Suppl.Fig.3)

Coordination of Drosha and miRNA expression by microRNA secondary structure during T-cell development and neuronal differentiation

Since Drosha levels vary between mature tissue types and expression levels of miRNAs with mismatches in the central region are high relative to rigid miRNAs when Drosha levels are high, we wanted to test if this mechanism is used to dynamically alter the miRNA expression profiles during cell development. In order to establish this relationship we first analyzed Drosha expression levels during T-cell differentiation and found a 3-fold decrease in Drosha expression from double-negative T-cells (DN3) to CD4+ cells (Chong et al., 2010) (Fig.5A). In neuronal lineage, we compared Drosha and Eef2 expression during the proliferation of mouse neural progenitors (GDS3442) and showed a significant decrease ($p<0.01$) in Drosha expression as the cells proliferate and mature in embryonic mouse brains (Fig.5B).

We further tested whether the miRNA groups correlate with the Drosha levels and found that the lower Drosha expression in CD4+ compared to DN3 cells correlates with lower expression of miRNAs with 5 mismatches in the region 5, 9-12nt from the Drosha cutting site and the ratio of expression between DN3 and CD4+ for miRNAs grouped by the number of mismatches in positions 5 and 9-12 shows that the difference in Drosha expression between DN3 and CD4+

cells correlates with a decrease in miRNA expression based on number of mismatches (Fig.5C). In DN3 cells Drosha levels are relatively high, and mismatched miRNAs are highly expressed; in CD4+ cells Drosha levels are low, and predominantly rigid miRNAs are produced. This agrees with previous results that when Drosha levels are low the levels of mature miRNAs generated from mismatched hairpins go down disproportionately compared to rigid miRNAs (Fig.3B-C, 4C-D). Mismatch counting of the complete hairpin, to control for the difference being local to the 5, 9-12 region, did not result in a similar trend (Suppl.Fig.4)

We also analyzed Drosha levels in hESC derived neuronal differentiation process. RNAseq data from early neuronal differentiation of H1 hESC shows an increase in Drosha, with a greater change occuring over day 0 to day 6 than day 6 to day 12 (Fig.6A). Canonical and non-canonical components of the miRNA biogenesis pathway have a significantly greater change in expression between days 0 and 6 than between days 6 and 12 of the neural differentiation protocol (Suppl.Fig.5). miRNA microarray data from the same period, day 0 to day 6, when Drosha levels increase, shows that the miRNAs with positive fold-change have significantly more mismatches in the region(s) 5, 9-12 than miRNAs with negative fold change (Fig.6B). Between days 6 and 12, when changes in Drosha expression are significantly lower, no significant difference in number of mismatches can be found between miRNAs grouped by fold-change (Fig.6C). This shows that when Drosha levels increase in early neuronal differentiation the expression of miRNAs generated from hairpins with mismatches in the 5, 9-12 region go up relative to more rigid hairpins, and when Drosha levels stabilize mismatched miRNAs no longer behave differently from rigid ones (Fig.6B-C). Whether increasing or decreasing during cell development, changes in Drosha levels affect groups of miRNAs based on their secondary structure. To control for the difference being local to the 5, 9-12 region, mismatches were counted in the remaining positions for a subset of 166 miRNAs, where no similar trend was observed (Suppl.Fig.6)

Biochemical validation of the relationship between pri-miRNA secondary structure, miRNA expression, and Drosha expression *in vivo*

Given the spatial and developmental heterogeneity of Drosha expression and its' correlation with miRNA expression based on secondary structure (Fig.3-6), we hypothesized that mismatches in a miRNA hairpin approximately 9-12nt from the Drosha cutting site would confer increased sensitivity to low levels of Drosha. Conversely, the absence of mismatches would confer decreased sensitivity to Drosha expression. In order to test our hypothesis, we compared the relative over expression of four miRNAs in two HeLa cell lines; one with a stable Drosha knockdown, the other transduced with a control virus (Control). Drosha protein levels were significantly reduced in the KD line (5-fold reduction; Suppl.Fig.7) resulting in downregulation of endogenous miRNAs (Suppl.Fig.8). miR-145 and miR-200b were selected for over-expression in these cells due to the clear presence of mismatches 9-12 nt from the Drosha cutting site (Fig.7A-B), making them ideal candidates to represent miRNAs at one end of the spectrum with regards to secondary structure. For the same reason, miR-137 and miR-9-1 were selected due to the clear absence of mismatches 9-12 nt from the Drosha cutting site (Fig.7C-D). Modifications of the 9-12nt region were designed to either insert or remove mismatches without disrupting the rest of the hairpin, while preserving the primary sequence of the mature miRNA, changing each miRNA from one extreme structure to the other while allowing for detection by traditional methods (Fig.7A-D). Since alteration of the secondary structure close to the Drosha cutting site has previously been shown to completely disrupt processing and may affect strand selection during RISC loading, we did not modify this region (Zeng and Cullen, 2003;Noland et al., 2011) We measured fold change of miRNA over expression (KD/ctrl; referred to as "relative over expression") by qPCR to determine whether or not mismatches in the central region determine sensitivity to Drosha *in vivo*.

For miR-145, mutation from a mismatched to a rigid stem increased relative over expression by 1.52-fold, $p<0.01$ (Fig 8A). For miR-137, mutation from a rigid stem to a mismatched stem decreased relative over expression by 3.84-fold, $p<0.05$ (Fig.8B). For miR-200b, mutation from a mismatched to a rigid stem increased relative over-expression by 3.76-fold, $p=0.05$ (Fig.8C). For miR-9-1, mutation to create mismatches in the 9-12nt region decreased relative over-expression by 4.66-fold (Fig.8D). These data show that the addition or removal of mismatches did affect the level of miRNA overexpression in DroshaKD compared to control, with more mismatches generating a lower ratio than the corresponding miRNA without mismatches 9-12nt from the Drosha cutting site.

***in vitro* processing of modified miRNAs**

Having shown the differential processing of rigid and mismatched miRNAs *in* vivo, we proceeded by testing this relationship *in vitro.* Linearized plasmids containing wild type (wt) and modified versions of mir-137 (203nt) and mir-200b (430nt) were *in vitro* transcribed and radiolabeled, followed by Drosha processing for 10, 30 and 90 minutes using nuclear extract from 293T cells. The resulting RNA was purified and visualized on a 10% urea-polyacrylamide gel (Fig.9A). Intensities of the bands corresponding to the pre-miRNA were quantified and normalized to the intensities of the regions containing pri-miRNA. Negative controls using only protein buffer and no lysate showed no detection in the bands used to quantify pre-miRNAs. Calculation of expression ratios of normalized 10min and 90 min samples showed a consistently higher ratio for rigid miRNAs than for their mismatched counterparts (Fig.9B). While the fold changes are smaller in *in vitro* assay compared to in vivo experiments (compare Fig.8B-C to Fig.9), the trend is the same, consistently showing that mismatched miRNAs have a higher sensitivity to reduced Drosha activity.

Discussion

Here we show that when Drosha levels change, the change in expression levels of miRNAs with mismatches in the hairpin, 9-12nt from the Drosha cutting site is greater than the change in expression levels of miRNAs that lack mismatches in the same region. We have found this correlation in mature mouse tissues and during development. We also successfully altered the sensitivity to changes in Drosha expression for 4 different miRNAs by changing their secondary structure. The results from the *in vitro* assay showed the same trend as the *in vivo* experiments.. Overall these data suggest that secondary structure of miRNA affects miRNA's processing efficiency when Drosha is limiting and that differentiating cells may use this as a means of miRNA regulation

Previous work on standardizing the miRNA secondary structure and cataloguing mismatch locations has used the approach of assigning mismatch positions for each strand of the hairpin separately (Han et al., 2006). In order to improve the representation of mismatch locations in three dimensional space, we further developed the mismatch mapping technique by taking both strands into consideration simultaneously in the analysis. This counting method resulted in a very consistent measurement of 21nt between the Drosha and Dicer cutting sites, which is consistent with previous reports indicating that Dicer processing of pre-miRNAs with asymmetric mismatches seems to explain the heterogeneity of length observed amongst mature miRNAs, and is also consistent with the idea that Dicer processing utilizes 'molecular rulers' which are size invariant (Starega-Roslan et al., 2010), which indicates that our counting method provides a good representation of the shape of the miRNA.

A relationship between secondary structure and pri-miRNA processing is well established, but no previous studies have focused specifically on the 9-12nt region within the pre-miRNA portion of the pri-miRNA and processing. pri-miRNAs are generally shorter, have fewer bulges and internal loops, and overall less mismatched bases within the stem region than non-miRNA-generating RNA transcripts, implying that secondary structure prediction is a reliable predictor of Drosha and Dicer processing (Ritchie et al., 2007). Furthermore, the secondary structure of pri-miRNAs controls the specificity and efficiency of Drosha processing, *in vitro* and *in vivo*, suggesting that Drosha substrate specificity may act as a mechanism for global regulation of miRNA expression profiles (Feng et al., 2011). On the other hand, it has also been shown that polymorphisms in pri-miRNA secondary structure do not affect processing (Diederichs and Haber, 2006), but only one of the examined single nucleotide polymorphisms (SNPs) was within the pre-miRNA portion of the pri-miRNA structure. In a wider search of human miRNA genes, it was determined that SNPs did affect the processing and function of miRNAs and most of these SNPs were found between the Dicer and Drosha cutting sites (Sun et al., 2009). The results presented herein could explain the significance of SNPs outside the seed region, but within the pre-miRNA structure, of some cancer and disease related miRNAs, such as the mutation in the stem of hsa-mir-125a (Duan et al., 2007), or in miR-181b-2, miR-208, miR-520e (Wu et al., 2008).

The miRNA biogenesis is regulated on multiple levels. The transcription of primary miRNAs (pri-miRNA) is regulated by common transcription factors, e.g. c-Myc (Abdelmohsen et al.), and some miRNAs are also found to be subject to methylation (Brueckner et al., 2007). miRNAs are transcribed in the nucleus by RNA polymerase II or III (Borchert et al., 2006;Kim and Nam, 2006) as part of the pri-miRNA, whose length is highly variable, ranging from ~200 up to several thousand nucleotides (Du and Zamore, 2005;Cullen, 2006). Drosha and DGCR8 bind the pri-miRNA, and Drosha cleaves pri-miRNAs at the base of the stem-loop and liberates a structure known as the precursor microRNA (pre-miRNA), which is ~60-70 nucleotides in length and forms a frequently mismatched hairpin structure with a 2 nucleotide 3’ overhang (Basyuk et al., 2003;Lee et al., 2003;Ritchie et al., 2007). The DGCR8 mRNA has stem loop structures that can be cleaved by the microprocessor, accordingly giving the microprocessor a self regulating mechanism (Han et al., 2009). The pre-miRNA is transported from the nucleus to the cytoplasm by Exportin-5 and is subsequently cleaved by the enzyme Dicer with its cofactor trans-activator RNA (tar)-binding protein (TRBP) (Chendrimada et al., 2005), with a notable exception in mir-451, which bypasses Dicer cleavage and is processed by Ago instead (Cheloufi et al., 2010). Dicer cleavage is further regulated by Lin28, which binds to pre-miRNAs from the let-7 family (Nam et al., 2011). Dicer acts by binding the 3'-overhang and cleaves the pri-miRNA ~22nt from the Drosha-cutting site to remove the terminal loop, resulting in an imperfect ~22nt called miRNA:miRNA* (Hutvagner et al., 2001;Bernstein and Allis, 2005;Feng et al., 2012). The miRNA enters the RNA-induced silencing complex (RISC), and the miRNA* is degraded, in a process where strand selection is variable (Bartel, 2004;Du and Zamore, 2005;Cullen, 2006;Kim and Nam, 2006). Although several mechanisms have been identified by which expression levels of select miRNAs are affected, and though Drosha is known to regulate overall Drosha expression, it has never previously been shown that the cell selectively can change miRNA expression profiles by altering levels of Drosha expression.

We show that Drosha levels vary between tissues and throughout cellular development, and that miRNAs without mismatches in the 9-12nt region are overrepresented in cells with low levels of Drosha, while highly mismatched miRNAs are highly expressed in cells with high levels of Drosha, and that this behavior can be altered by changing the miRNA secondary structure. This supports our hypothesis that the cell selectively regulates rigid and mismatched miRNAs by altering Drosha expression. We propose that this mechanism may contribute to the phenotypic relevance of specific levels of Drosha expression in development and terminal differentiation. In addition to other examples of post-transcriptional regulation, this can explain why some polycistronic miRNAs that are driven by the same promoter have very different expression levels, and the impact of SNPs, as stated above. This mechanism may be utilized while designing artificial miRNAs to increase stability, control expression levels throughout differentiation, or fine-tune miRNA expression in different tissue types.

Materials and methods

Counting method

Position 1 was assigned to the first nucleotide pair after the Drosha cutting site. The nucleotides that would form the single stranded overhang after Drosha processing were assigned negative positions (-1, -2 etc.) and if the cutting site was known for only one strand the overhang was assumed to be 2nt. Positions were then assigned for each nucleotide pair in ascending order moving towards the Dicer cleavage site. The average hairpin contains several mismatches, which occasionally contain more nucleotides in one strand than in the other. Whenever such a length discrepancy occurred, the shorter strand was used to assign the mismatch position, using the later position whenever an asymmetric mismatch was located between two positions. Each position where the nucleotides were bound to each other was assigned the score 0, while positions with at least one nucleotide mismatched were given the score 1.

Genomic location

Data from miRBase was used to determine whether each miRNA was intronic or intergenic. The fold change values in the two groups are visualized in Fig 2. Significance was calculated using a t-test. Information regarding whether each miRNA was intronic or intergenic was retrieved from miRBase and fold change distribution was analyzed for each of the two groups and differences in distribution was calculated using a t-test.

miRNA machinery and structure in neuronal differentiation

H1 cells were differentiated into neurons using 90% confluent cells on Matrigel. The differentiation was performed using a modified version of the dual SMAD inhibition protocol (Chambers et al., 2009), using the TGF-β inhibitor SB432542, ALK inhibitor LDN193189, and cyclopamine in CDM Basal media, which was changed daily for 12 days. At time points 0, 6, and 12 days starting from the first addition of inhibitors, cells were harvested. mRNA and miRNA were purified using the RNeasy Plus Mini Kit, and the miRNeasy Mini Kit, respectively, according to manufacturer instructions (Qiagen). Expression levels were measured using RNAseq (Expression Analysis Inc) and single-color miRNA microarrays (LC Sciences) respectively. RNAseq expression data were then count- and length-normalized to transcripts per million (TPM) after exclusion of mitochondrial genes, and miRNA expression data were Lowess normalized for cross-chip comparisons.

H1 cell culture

H1 human embryonic stem cell lines were cultured on a feeder layer (MEFs) in Dulbecco's modified Eagle's medium (DMEM)/Ham's F-12 medium containing GlutaMax supplemented with 20% serum replacer (SR), 1 mM sodium pyruvate, 0.1 mM non-essential amino acids, 50 U/ml penicillin, 50 ug/ml streptomycin, 0.1 mM β-mercapto- ethanol (Sigma-Aldrich, St. Louis), and 2 ng/ml basic fibroblast growth factor (Qi et al 2009). Knockdowns of Drosha and Dicer were generated by infection with lentiviral vectors containing shRNA constructs against Drosha and Dicer respectively. Control cells were infected with control viruses. After passaging the next day, infected cells were treated with blasticidin (0.5 ug/ul) for selection until discrete clusters of cells appeared. If cells showed less stability with respect to Drosha knockdown, cells were repeatedly selected with antibiotics (Qi et al., 2009). For the neuronal differentiation experiment H1 cells were cultured feeder-free on Matrigel (BD Biosciences) with mTeSR1 (Stemcell Technologies).

Generation of HeLa line with Drosha knockdown

HeLa cells were transduced with a pL6-Tet lentiviral vector containing shRNA against Drosha (Qi et al., 2009) (Titer 1.79E+08, MOI of 1:10; 10µg/ml blasticidin selection for 4 days). Generation of the control cells was done with a pLKO.1 control lentivirus, (MOI 1:10; 10µg/ml puromycin selection for 4 days).

Drosha qPCR analysis

Drosha and DGCR8 levels were measured in Drosha KD and Control HeLa cells using the SYBR-GREEN Q-PCR kit. Drosha Ct values were normalized to GAPDH (ΔCt: Drosha–GAPDH). The ΔΔCt values were calculated by subtracting the Drosha ΔCt values in the Control cells from those values in the KD cells. Fold change was calculated using the (2^(-ΔΔCt)) method.

Total RNA extraction and quantitative RT-PCR analysis

Brain and liver tissues were dissected from WT mice (n=3). Tissues were homogenized using mortar and pestle in a liquid nitrogen bath. RNA was extracted from cells adherent on plates or from tissue homogenate using Trizol reagent (Invitrogen) and cDNA was synthesized with the

Omniscript RT kit (Qiagen). Beta-actin or UBC was used as endogenous control for normalization. Quantitative PCR reactions were performed in triplicate using the SyberGreen (Applied Biosystems) with the 7300 real time PCR system (Applied Biosystems). Primers used in our study are listed in Supplemental Table 1. Normalized Drosha levels in the brain were compared to those in the liver and the fold change was calculated using the 2^(-ΔΔCt) methods. qPCR of miRNAs was conducted using TaqMan® miRNA assays (Applied Biosystems). miRNA expression was measured in both DroshaKD and Control HeLa cells. Raw Ct values for miRNAs were first normalized to RNU66 (endogenous snoRNA, internal control). Bleomycin values were normalized to GAPDH. To normalize to transfection efficiency, the normalized miRNA values were re-normalized to the normalized Bleomycin values (normalized miRNA ΔCt - normalized Bleomycin ΔCt). This new normalized value from control cells was subtracted from the corresponding value in the knockdown cells to generate the ΔΔCt values. Fold change with respect to the control cells was calculated using the 2^(-ΔΔCt) method.

Protein extraction and Western blot

With the exception of mouse tissues, protein extraction and Western blot analysis followed procedures that were described previously (Zhou et al., 2011). Briefly, cells were washed with Dulbecco's PBS (Sigma Aldrich, St. Louis, MO) and directly lysed on culture dish using M-PER Mammalian Protein Extraction Reagent (Thermo Scientific, Rockford, IL; 0.5ml per 35mm plate) and protease inhibitor cocktail (Complete Mini, Roche Applied Science, Germany). 20 µg of protein extracts were loaded, separated by 4-20% SDS-PAGE using Mini-PROTEAN TGX Precast Gels (Bio-Rad, Hercules, CA), and transferred to polyvinylidene difluoride transfer membranes (Thermo Scientific). Membranes were blocked with 5% nonfat dry milk for at least 60 minutes at room temperature, and incubated overnight at 4°C with primary antibody. Blots were incubated for one hour with horseradish peroxidase-conjugated Blotting Grade Affinity Purified secondary antibodies (Bio-Rad) and were visualized by enhanced chemiluminescence (Millipore Corp, Billerica, MA). Protein expression levels in the gel were quantified by densitometry implemented in Image-J (National Institutes of Health, Bethesda http://rsb.info.nih.gov/ij/, 1997–2013). Antibodies used in this study are: rabbit anti-Drosha antibody (Cell Signaling Technology, Beverly, MA) diluted at 1:1000, mouse anti-β-actin antibody (Santa Cruz Biotechnology, Santa Cruz, CA) diluted at 1:5000, goat anti-rabbit antibody (Bio-Rad Laboratories, Hercules, CA) diluted at 1:10000 and goat anti-mouse antibody (Bio-Rad Laboratories, Hercules, CA). For mouse tissues, the protocol above was followed for protein extraction, gel loading, and the quantification of Drosha expression. As a control, either Western using actin Ab or Coomassie Brilliant Blue R (Sigma Aldrich) staining was used. Entire-cell protein expression was quantified using Image-J.

Transfections

The mutations were performed on the passenger strand, leaving the mature strand unchanged in order to be able to detect both the mutated and non-mutated miRNAs using the same TaqMan® assay. The pri-miRNA inserts were amplified from genomic DNA by PCR, including the hairpin and ~200nt flanks on each side. The inserts were ligated into a modified pcDNA3.1+ plasmid with the neomycin selection marker replaced with a bleomycin resistance gene. Mutated versions of the constructs, with altered mismatches in the 9-12 region, were ordered from Genscript. The miRNA constructs were transiently transfected into HeLa cells using Lipofectamine2000. RNA was extracted after 48 hours using Trizol.

Analysis of Drosha and miRNA expression in mouse tissues

For GDS3052 and GDS3142: These are two microarray analyses done in mouse using an Affymetrix Mouse Genome 430 2.0 chip. Drosha expression values were extracted from the datasets and were normalized to a gene that in both studies showed minimal variation between tissues; Eef2 (eukaryotic translation elongation factor 2). Drosha expression was represented as percent of Eef2 in the tissues.

For Takada et al. 2006: The miRNAs expressed in each tissue were first sorted according to number of clones (highest number of clones being highest expression and lowest number of clones being lowest expression). The miRNAs were then divided into groups based on the number of mismatches in the 5, 9-12 region (X-axis) and within each group the average expression level (Y-axis) was calculated by geometric mean.

***in vitro* processing of miRNAs**

Plasmids for wild type (wt) and modified versions of mir-137 and mir-200b were linearized, gel purified and *in vitro* transcribed using the Riboprobe (Promega) protocol and labeled with α^{32}P-CTP. Transcripts were purified using the RNAeasy mini kit (Qiagen) and then allowed to fold by addition of KAc, heating to 95°C for 5 min and then kept on ice till the processing reaction. Processing was performed by mixing folded pri-microRNA transcripts in 20µl reactions with 8µg 293T cell nuclear extract (ProteinOne, Catalog Ref. P0003-01; nuclear extract has been confirmed to support pri-miRNA processing, Chen et al., in preparation), 6.4mM $Mg(Ac)_2$, 1mM

ATP, 20mM Creatine phosphate, 20mM Tris, 100mM KAc, 0.5mM DTT, 0.2mM PMSF, 20U RiboLock RNase inhibitor (ThermoScientific) and 1µg tRNA for 10, 30 and 90 minutes at 37°C followed by Trizol extraction and run on a 10% polyacrylamide-urea gel for 3 hours at 390V. Gel exposures were quantified using ImageJ64 software where intensities of bands corresponding to pre-miRNAs were normalized to regions containing pri-miRNA transcripts.

Statistical analysis

Throughout the paper, P values for expression levels were calculated using Student's t-test. *, $P< 0.05$; and **, $P< 0.01$. Bars show Standard error of the mean (SEM) for at least 3 separate experiments.

P-values for differences in mismatch trends were calculated in R using a chi-square test with Yate's continuity correction.

Figure legends

Figure 1.

Fold change of miRNA expression between Drosha knockdown and control H1 hESC for 220 miRNAs assayed by qPCR. Arrow marks median (0.495).

Figure 2.

Fold change of expression for miRNAs between Drosha knockdown and Control H1 hESC for 220 miRNAs assayed by qPCR and grouped by genomic origin (intergenic/intronic).

Figure 3.

Counting of mismatches in miRNAs in H1 cells with DroshaKD and control cells. Positions regarded as mismatched are marked by numerals in light grey (A). Mismatches summed for positions counted from the Drosha cutting site for a set of 202 miRNAs split in two equal groups based on fold change in the Drosha knockdown (B). P-values were calculated using Pearson's Chi-squared test with Yates' continuity correction, giving log p-values < -2 for positions 5, 10, 11 and 12 for the full data set split in half (dark grey) and for the quartiles at the extreme ends with sample size extrapolated to that of the full data set (C).

Figure 4: Differential expression of Drosha in mouse tissues.

A and B. Expression levels in Drosha in mouse tissues, normalized to EEF2; GEO(GDS3052) (A), GEO(GDS3142) (B). C and D. Fold change of Drosha expression in liver, compared to the average expression of EEF2, RPL13A, SDHA, TBO and UBC; GDS3052 (C), GDS3142 (D). qPCR analysis of Drosha mRNA levels in mouse brain and liver (E), Quantification of Drosha protein levels show 9.6 fold increase in brain compared to liver tissue ($p<0.05$) (F), Western blot of Drosha protein in mouse brain and liver normalized to total protein (Coomassie Blue) (G), Average expression of miRNAs with different numbers of mismatches in the 5, 9-12 region in mouse liver and brain (Takada et al., 2006) (H)

Figure 5: Drosha levels and miRNA profiles during T-cell differentiation

Drosha levels in T-cells measured by qPCR, normalized to β-actin (Chong et al., 2010) (A). Drosha levels in developing mouse brain (GDS3442), normalized to EEF2 (B). Ratio (CD4SP/DN3) of average expression for each miRNA group based on number of mismatches in the 5, 9-12 region (C)

Figure 6: Changes in gene expression levels during neuronal differentiation

Drosha expression in neuronal differentiation of H1 hESC into Pax6+ neural cells (RNAseq) (A). Number of mismatches in miRNAs with high and low fold change through days 0 - 6 (multiArray) (B). Number of mismatches in miRNAs with high and low fold change through days 6 - 12 (multiArray) (C).

Figure 7.

Predicted structures of wild-type and modified hairpins of hsa-mir-145, -137, -9-1 and 200b, demonstrating canonical "mismatched" and "rigid" miRNA hairpins. Entropy curves of the nucleotide sequence, starting at the 5' end, indicate stability changes being local to the changed regions.

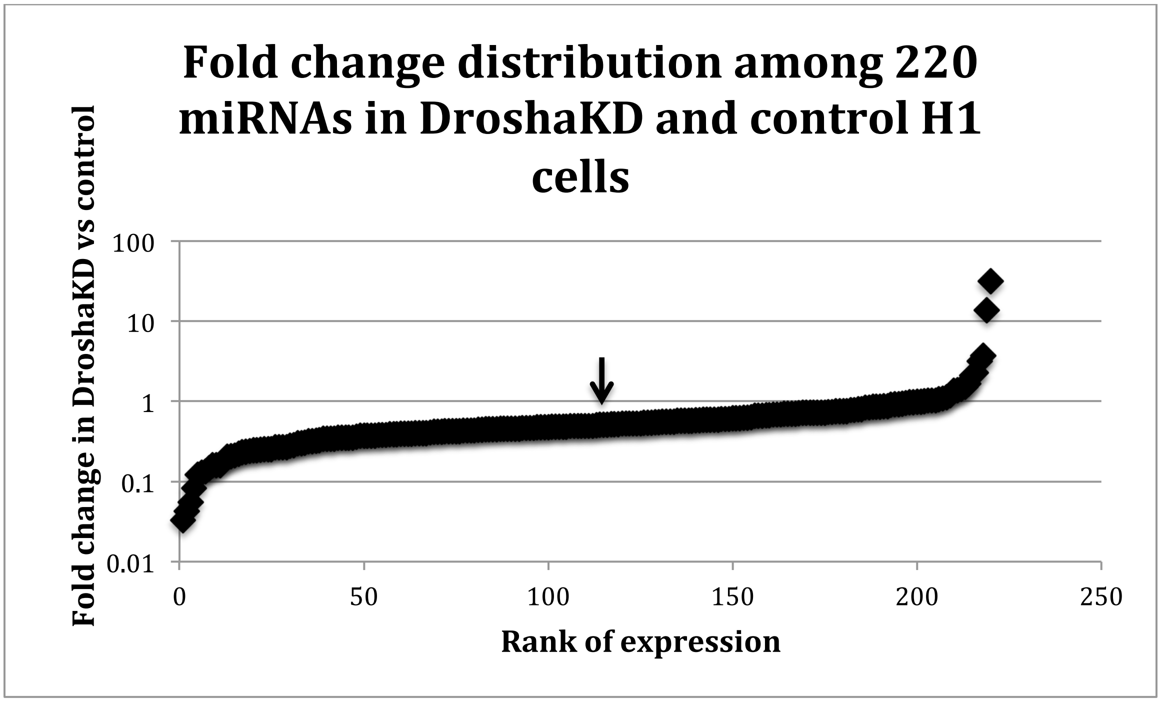
Figure 8.

Fold change of over expression levels of wild type and mutated versions of miR-145 (A), -137 (B), -200b (C) and -9 (D) between Drosha knockdown and control HeLa cells, assayed by TaqMan® probes and normalized to RNU66. (n=3). Transfection efficiency was accounted for by SYBRgreen® qPCR of the bleomycin resistance gene in the vector, normalized to GAPDH and subtracted from the ΔCt values from the TaqMan® assays. Final normalization was made to the fold change of over expression for each wt construct. Statistical significance was calculated using a student's t-test, resulting in p-values of <0.01 (mir-145), <0.05 (mir-137), =0.05 (mir-200b).

Figure 9: *in vitro* processing of miRNAs using nuclear extract

A. *in vitro* transcribed miRNAs, processed by 293T cell nuclear extract and visualized on a 10% urea-polyacrylamide gel. Bands corresponding to processed pre-miRNAs are marked in grey box, and areas containing pri-miRNAs used for normalization are labeled with black box. B. 10/90 minute ratios of quantified pre-miRNA levels normalized to each respective pri-miRNA levels as indicated in A.

Figure 1



Fold change distribution among 220 miRNAs in DroshaKD and control H1 cells
Fold change in DroshaKD vs control
100
10
1
0.1
0.01
0
50
100
150
200
250
Rank of expression

Figure 2

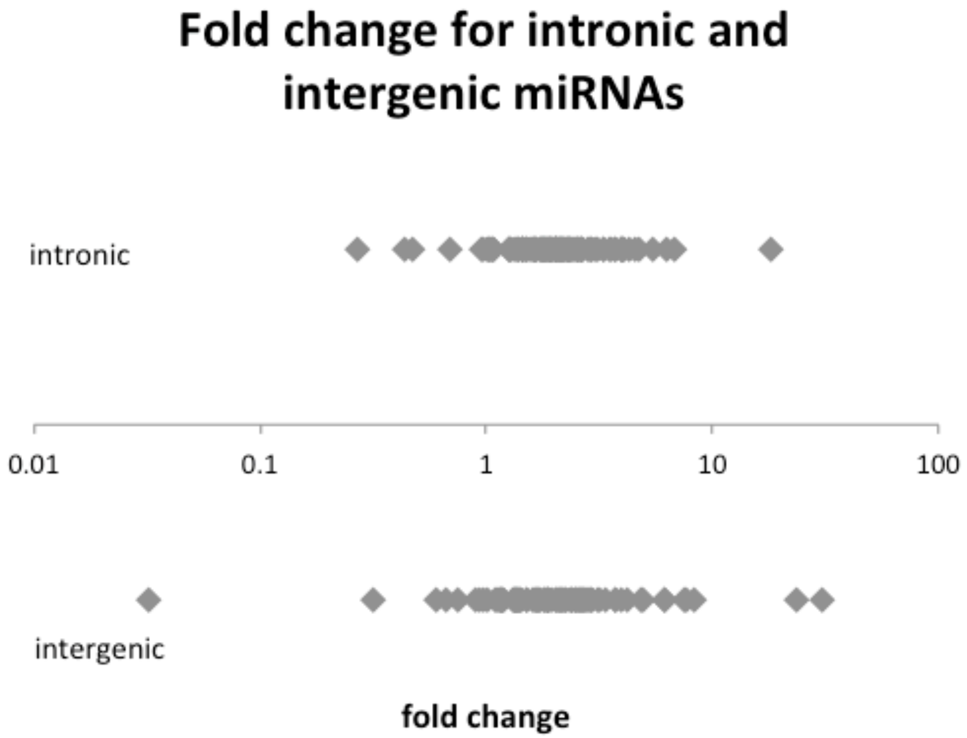
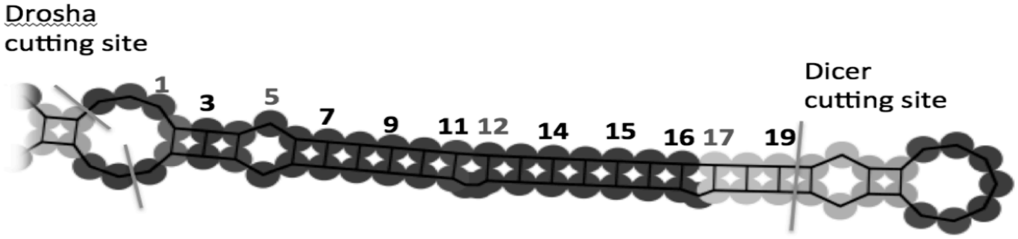
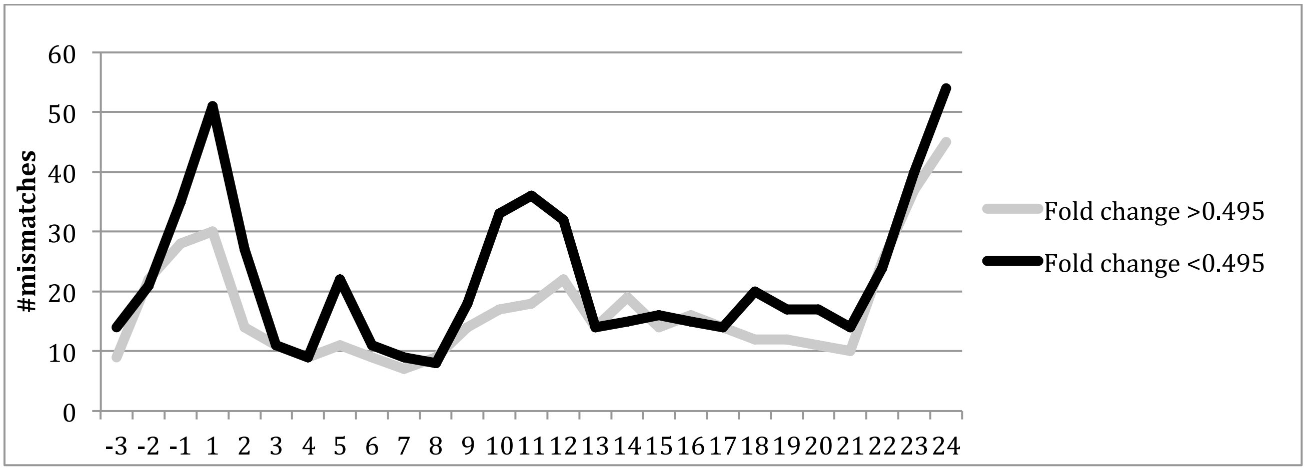

Fold change for intronic and intergenic miRNAs
intronic
0.01
0.1
1
10
100
intergenic
fold change

Figure 3

A


Drosha cutting site
1
3
5
7
9
11
12
14
15
16
17
19
Dicer cutting site

B


60
50
40
30
20
10
0
#mismatches
-3 -2 -1 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Fold change >0.495
Fold change <0.495

C

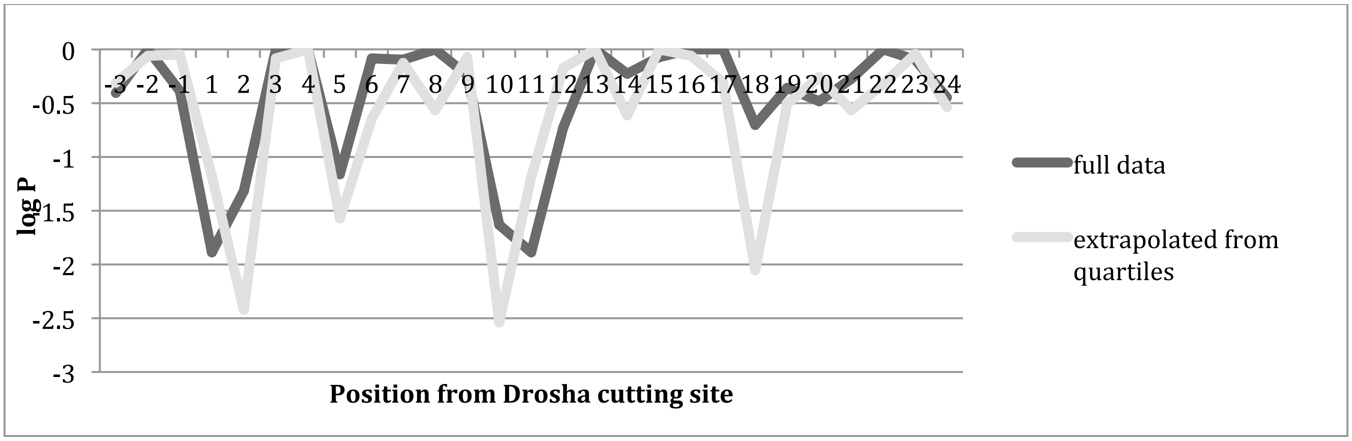

0
-0.5
-1
-1.5
-2
-2.5
-3
log P
-3 -2 -1 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Position from Drosha cutting site
full data
extrapolated from quartiles

Figure 4

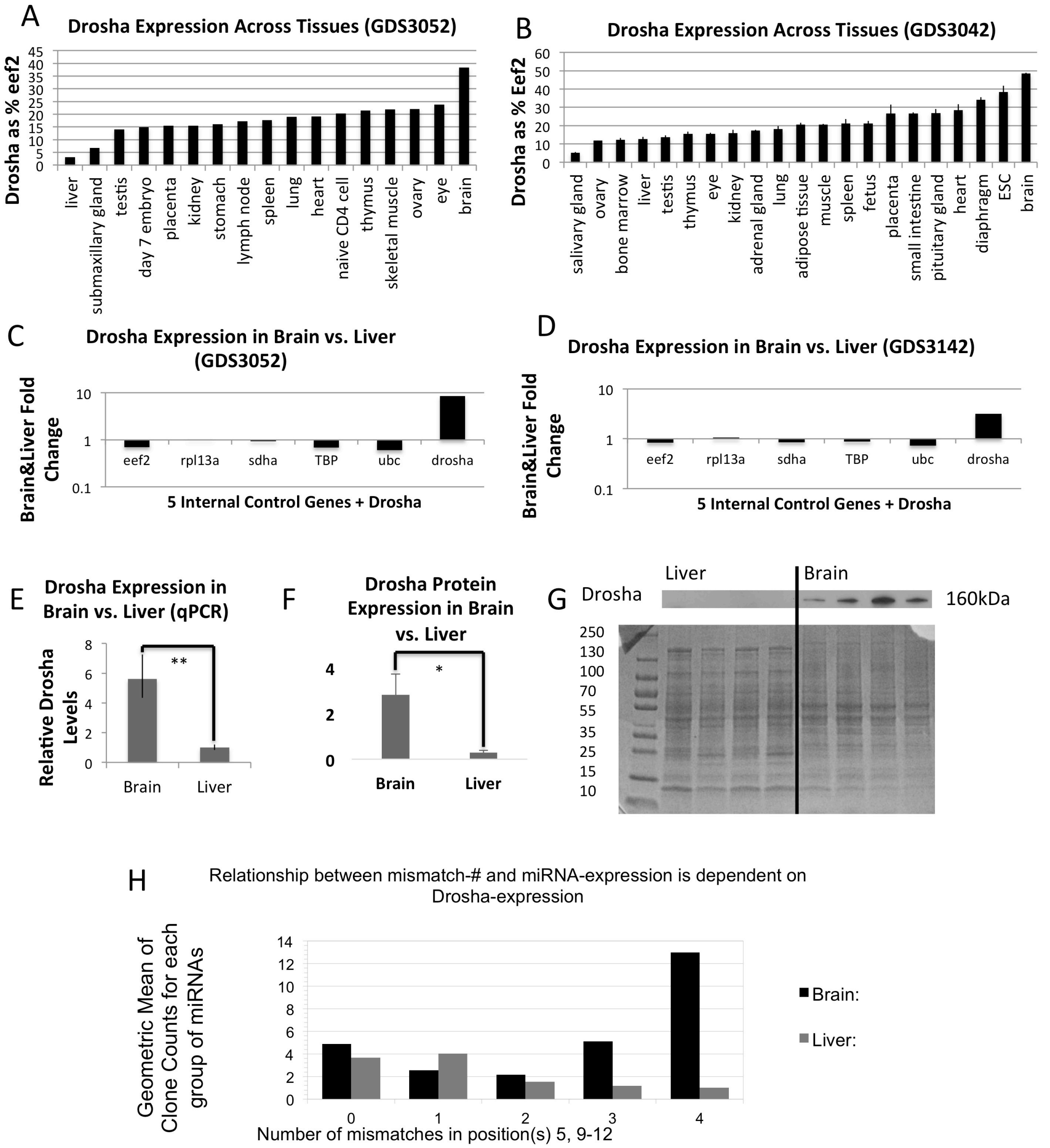

A
Drosha Expression Across Tissues (GDS3052)
Drosha as % eef2
45
40
35
30
25
20
15
10
5
0
liver
submaxillary gland
testis
day 7 embryo
placenta
kidney
stomach
lymph node
spleen
lung
heart
naive CD4 cell
thymus
skeletal muscle
ovary
eye
brain
B
Drosha Expression Across Tissues (GDS3042)
Drosha as % Eef2
60
50
40
30
20
10
0
salivary gland
ovary
bone marrow
liver
testis
thymus
eye
kidney
adrenal gland
lung
adipose tissue
muscle
spleen
fetus
placenta
small intestine
pituitary gland
heart
diaphragm
ESC
brain
C
Drosha Expression in Brain vs. Liver (GDS3052)
Brain&Liver Fold Change
10
1
0.1
eef2
rpl13a
sdha
TBP
ubc
drosha
5 Internal Control Genes + Drosha
D
Drosha Expression in Brain vs. Liver (GDS3142)
Brain&Liver Fold Change
10
1
0.1
eef2
rpl13a
sdha
TBP
ubc
drosha
5 Internal Control Genes + Drosha
E
Drosha Expression in Brain vs. Liver (qPCR)
Relative Drosha Levels
8
6
4
2
0
**
Brain
Liver
F
Drosha Protein Expression in Brain vs. Liver
4
2
0
*
Brain
Liver
G
Liver
Brain
Drosha
160kDa
250
130
100
70
55
35
25
15
10
H
Relationship between mismatch-# and miRNA-expression is dependent on Drosha-expression
Geometric Mean of Clone Counts for each group of miRNAs
14
12
10
8
6
4
2
0
0
1
2
3
4
Number of mismatches in position(s) 5, 9-12
Brain:
Liver:

Figure 5

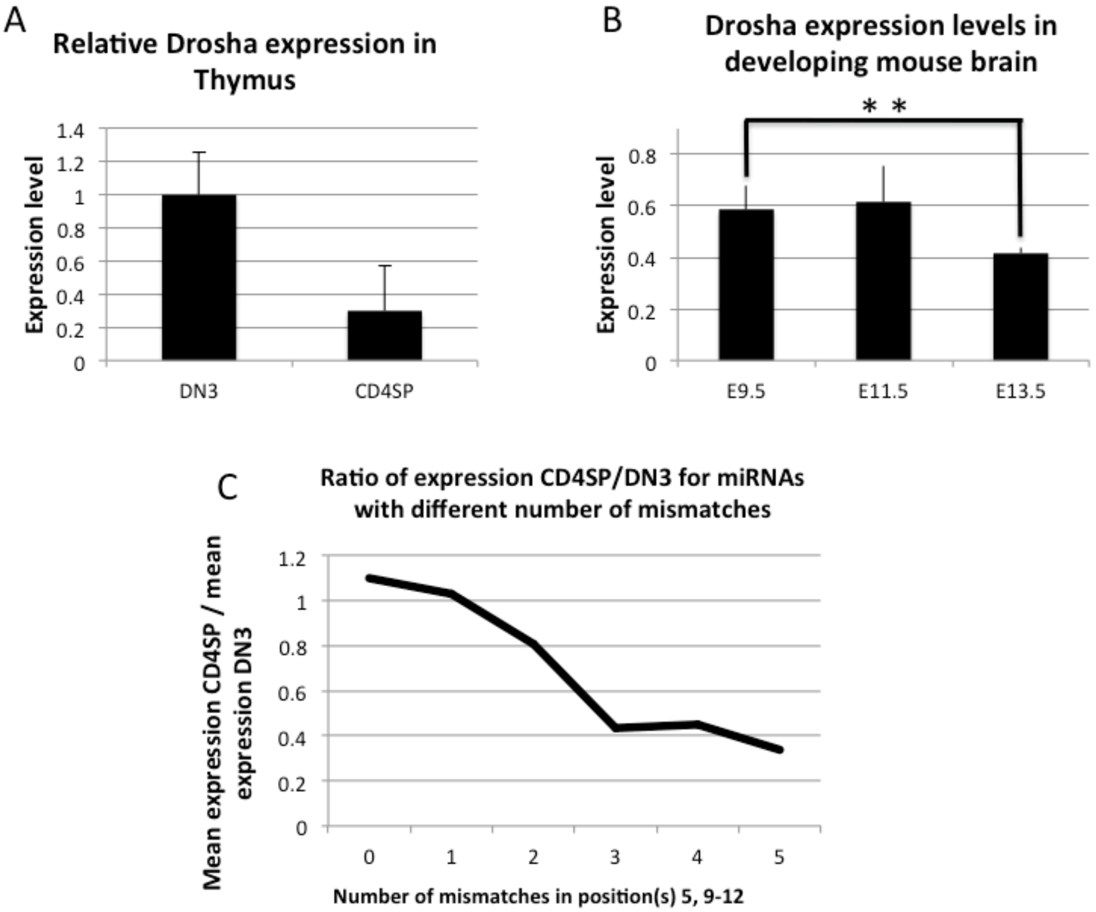

A
Relative Drosha expression in Thymus
Expression level
1.4
1.2
1
0.8
0.6
0.4
0.2
0
DN3
CD4SP
B
Drosha expression levels in developing mouse brain
* *
Expression level
0.8
0.6
0.4
0.2
0
E9.5
E11.5
E13.5
C
Ratio of expression CD4SP/DN3 for miRNAs with different number of mismatches
Mean expression CD4SP / mean expression DN3
1.2
1
0.8
0.6
0.4
0.2
0
0
1
2
3
4
5
Number of mismatches in position(s) 5, 9-12

Figure 6

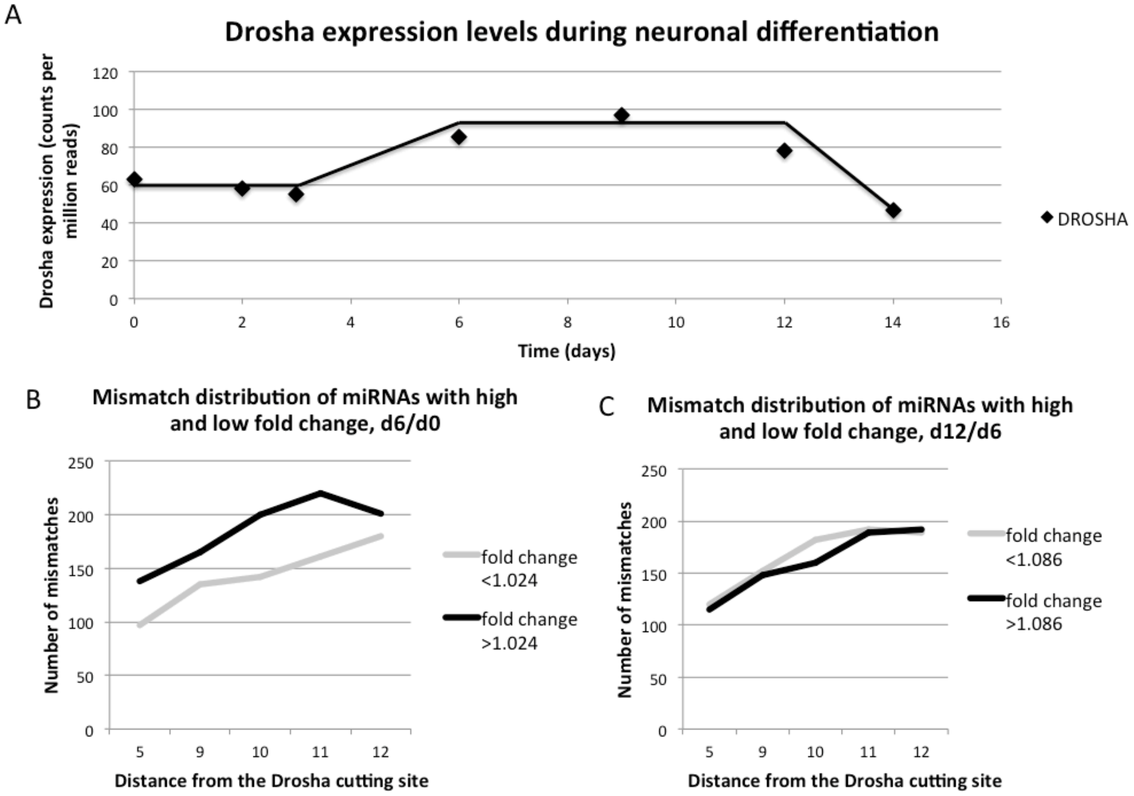

A
Drosha expression levels during neuronal differentiation
Drosha expression (counts per million reads)
120
100
80
60
40
20
0
0
2
4
6
8
10
12
14
16
Time (days)
◆ DROSHA
B
Mismatch distribution of miRNAs with high and low fold change, d6/d0
Number of mismatches
250
200
150
100
50
0
5
9
10
11
12
Distance from the Drosha cutting site
fold change <1.024
fold change >1.024
C
Mismatch distribution of miRNAs with high and low fold change, d12/d6
Number of mismatches
250
200
150
100
50
0
5
9
10
11
12
Distance from the Drosha cutting site
fold change <1.086
fold change >1.086

Figure 7

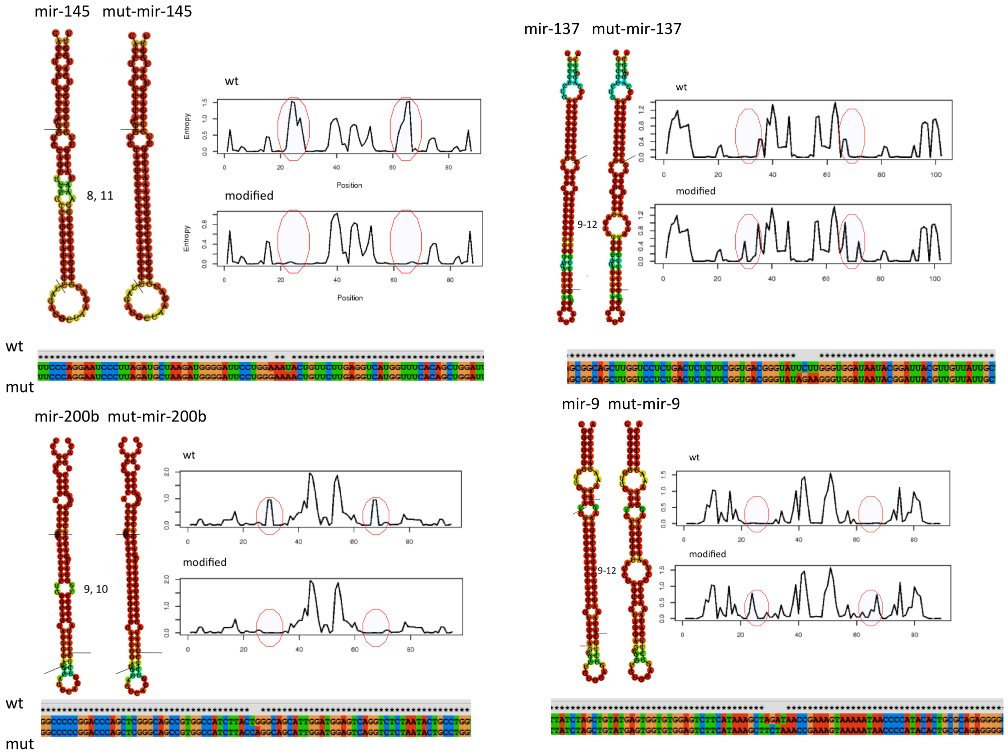

mir-145
mut-mir-145
8, 11
wt
modified
Entropy
Position
mir-137
mut-mir-137
9-12
mir-200b
mut-mir-200b
9, 10
mir-9
mut-mir-9
wt
mut

Figure 8

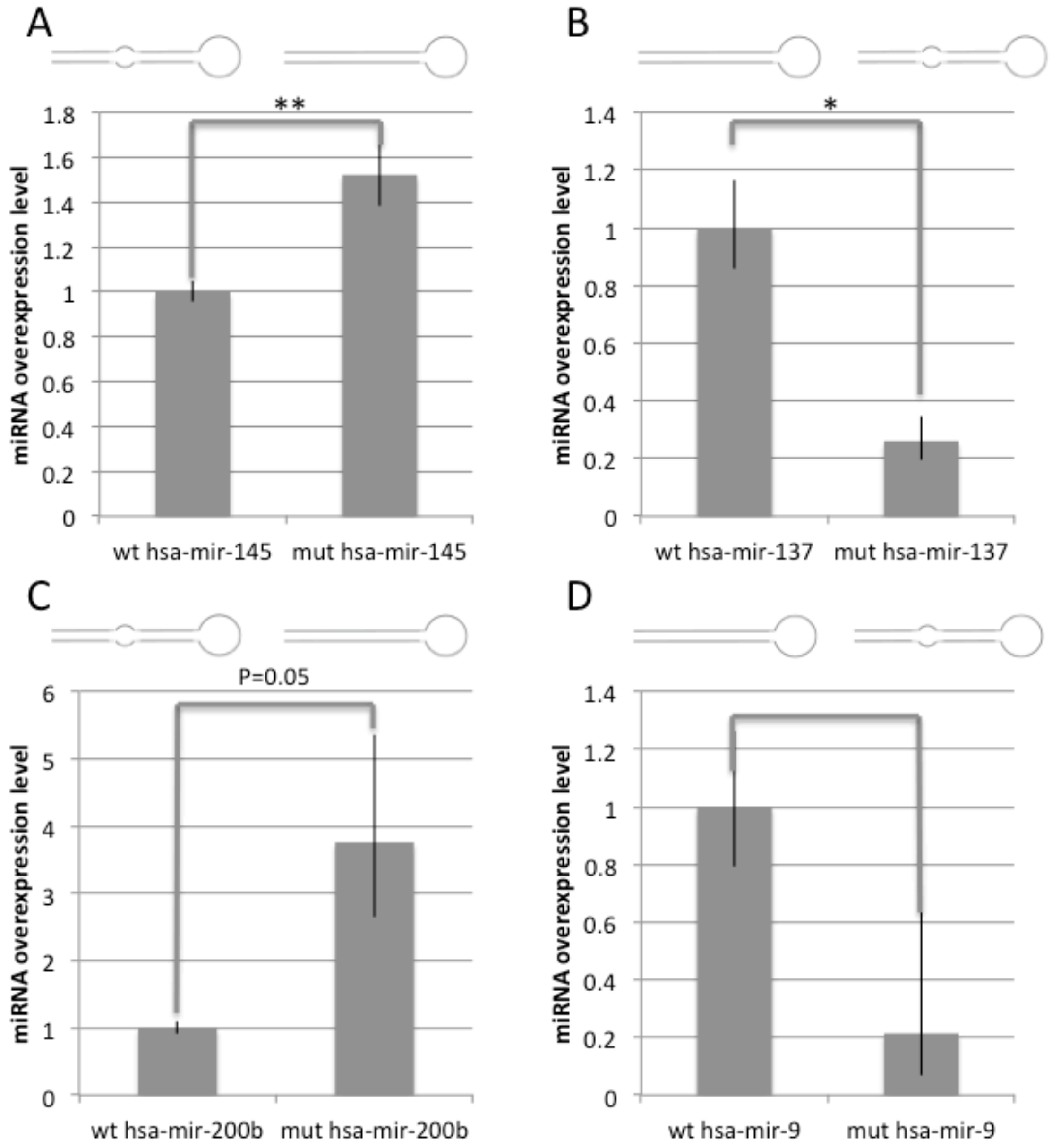
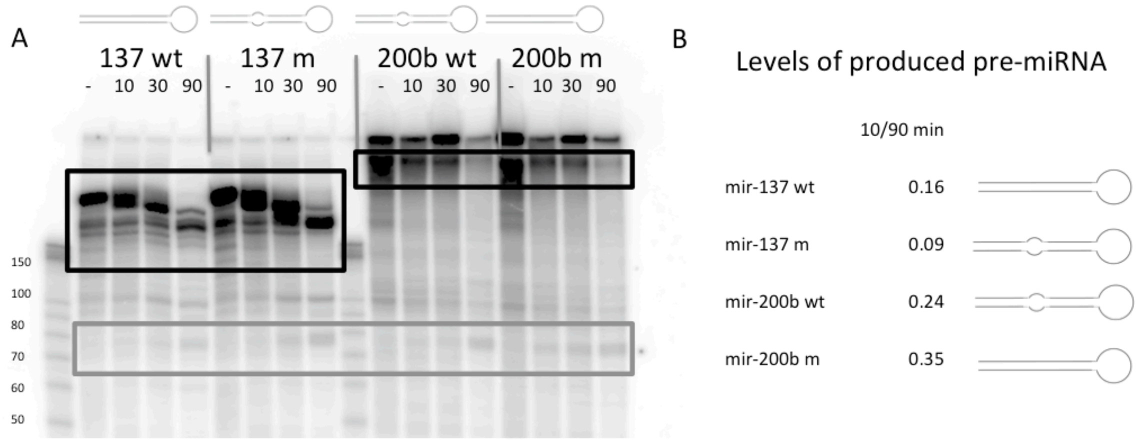

A
**
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
miRNA overexpression level
wt hsa-mir-145
mut hsa-mir-145
B
*
1.4
1.2
1
0.8
0.6
0.4
0.2
0
miRNA overexpression level
wt hsa-mir-137
mut hsa-mir-137
C
P=0.05
6
5
4
3
2
1
0
miRNA overexpression level
wt hsa-mir-200b
mut hsa-mir-200b
D
1.4
1.2
1
0.8
0.6
0.4
0.2
0
miRNA overexpression level
wt hsa-mir-9
mut hsa-mir-9

Figure 9



	10/90 min
mir-137 wt	0.16
mir-137 m	0.09
mir-200b wt	0.24
mir-200b m	0.35

Supplemental Figure legends

Supplemental Figure 1: Counting of mismatches in miRNAs in H1 cells with DicerKD and control cells. Mismatches summed for positions counted from the Drosha cutting site for a set of 202 miRNAs split in two equal groups based on fold change (median=0.395) in the Dicer knockdown

Supplemental Figure 2: Drosha protein expression levels in mouse brain and liver. Protein levels of Drosha and β-actin measured by western blots (A). Quantification of Protein levels using triplicates extrapolated from different gels. Fold change = 3.2, $p<0.05$ (B).

Supplemental Figure 3: Control mismatch counts for mouse brain and liver. Average expression of miRNAs with different numbers of mismatches in the 1-21 region in mouse liver and brain (Takada et al., 2006)

Supplemental Figure 4: Control mismatch ratios in T-cells. Ratio (CD4SP/DN3) of average expression for each miRNA group based on number of mismatches in the 5, 9-12 region.

Supplemental Figure 5: Fold change of miRNA machinery genes in neuronal development. Fold changes for day6 vs. day0, and day12 vs. day6 for a set of genes involved with miRNA biogenesis, based on RNAseq data (Suppl.Table1).

Supplemental Figure 6: Control mismatch counts during neuronal differentiation. Number of mismatches outside the 5, 9-12 region in miRNAs with high and low fold change (multiArray) in neuronal differentiation through days 0-6 and 6-12.

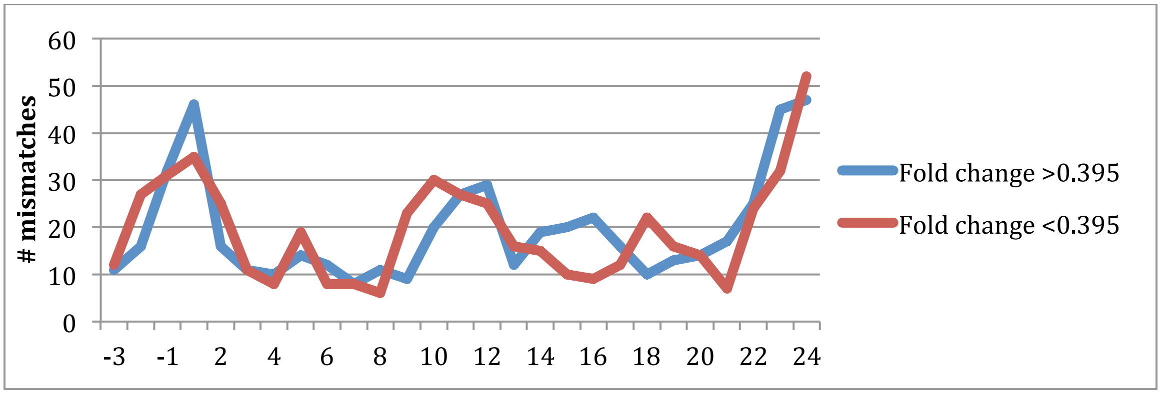
Supplemental Figure 7: DroshaKD in HeLa cells, verification of expression levels. Drosha levels measured by qPCR and normalized to GAPDH. 7 biological replicates showed an average 9-fold knock down (log $p < -5$) (A). Drosha protein expression levels measured by western blotting, normalized to β-actin. 3 biological replicates showed an average 5-fold knockdown ($p<0.05$) (B)

Supplemental Figure 8: Endogenous miRNA levels in DroshaKD and control cells. Endogenous levels for a random selection of miRNAs were measured by TaqMan qPCR, normalized to RNU66. The miRNAs are sorted right to left by increasing number of mismatches in region 9-12 nucleotides from the Drosha cutting site.

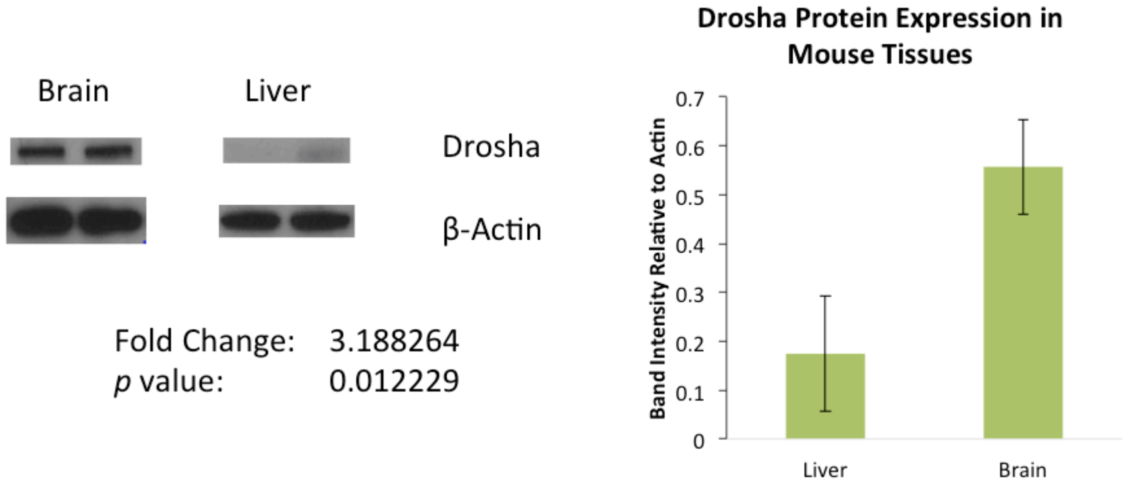
Supplemental Table 1: qPCR primer sequences

Supplemental Table 2: Expression levels of genes involved with miRNA biogenesis during neuronal differentiaiton. The genes of interest are listed by rows with normalized mean TPM (transcript per million) values reported in the table. Across columns are the time points of in vitro cortical neural induction, from day 0 to day 12 incremented by three days.

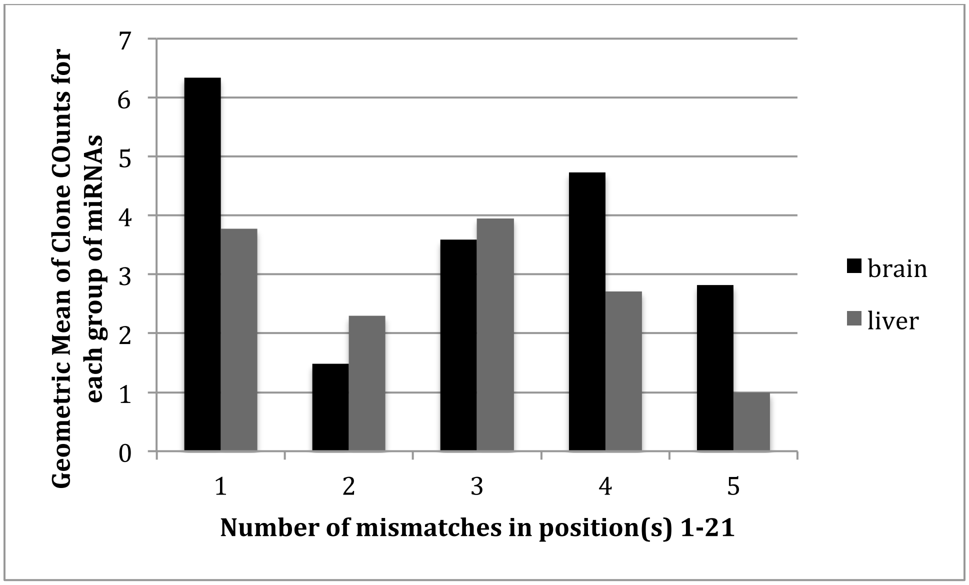
Supplemental Figure 1

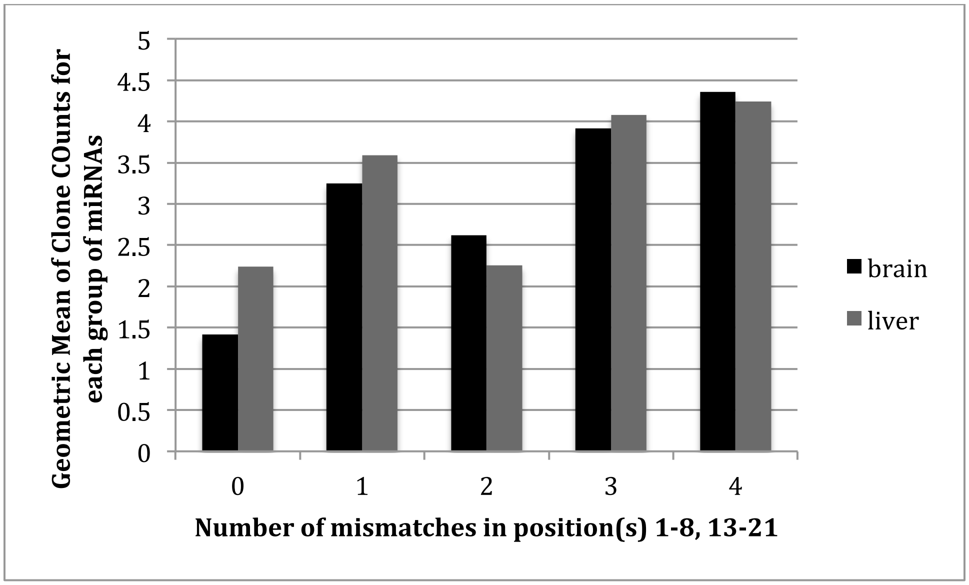


Supplemental Figure 2

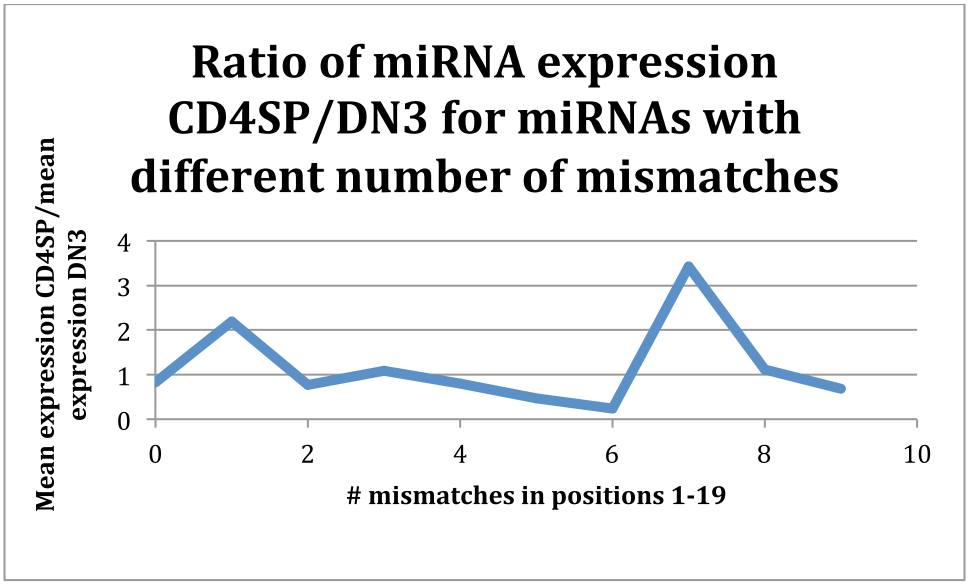


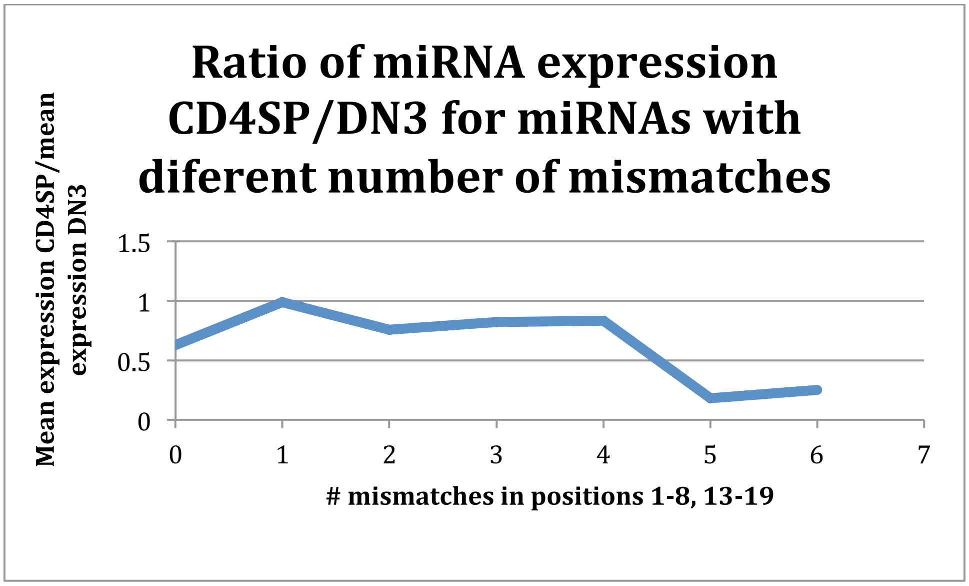
Supplemental Figure 3



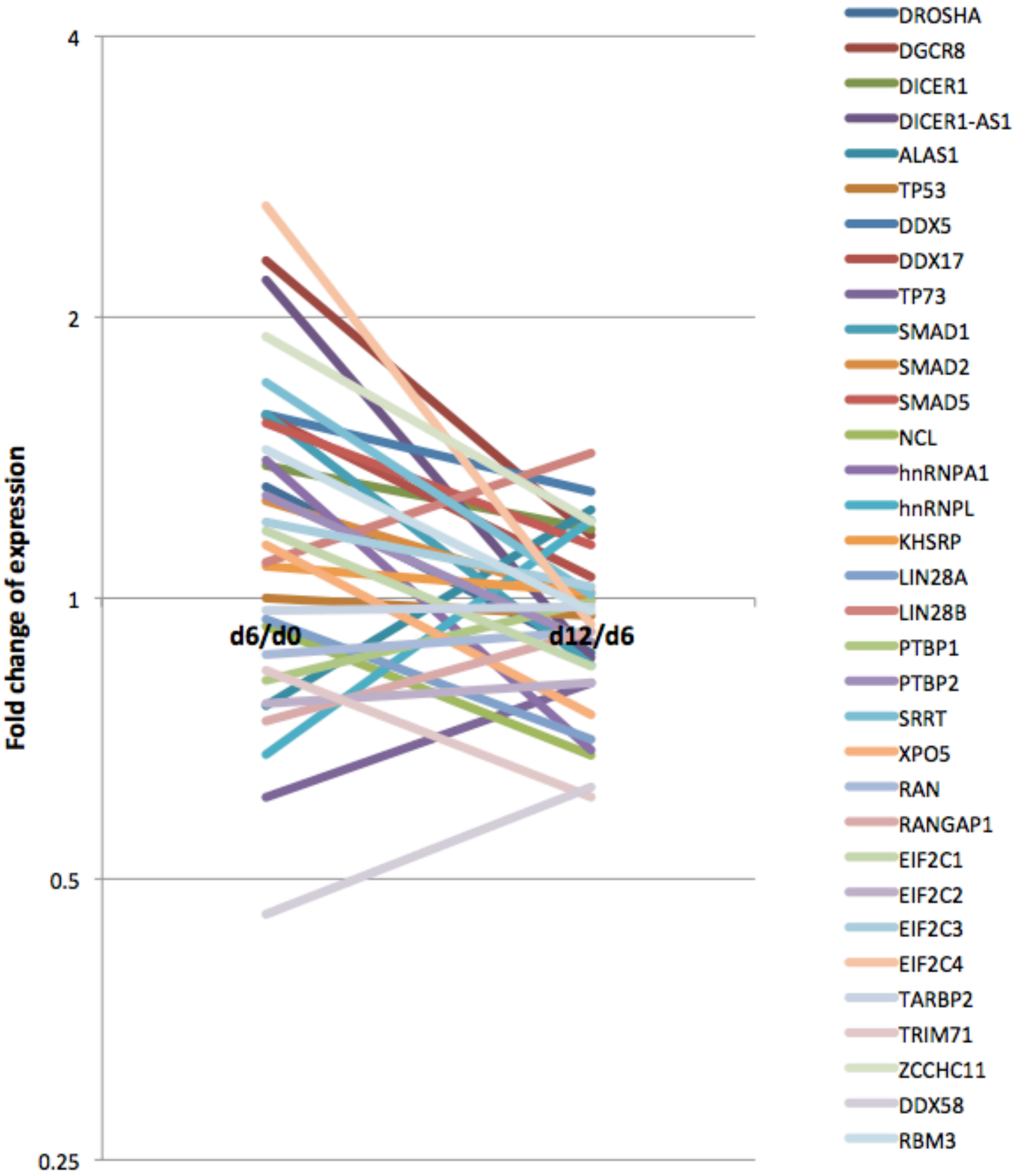


Supplemental Figure 4

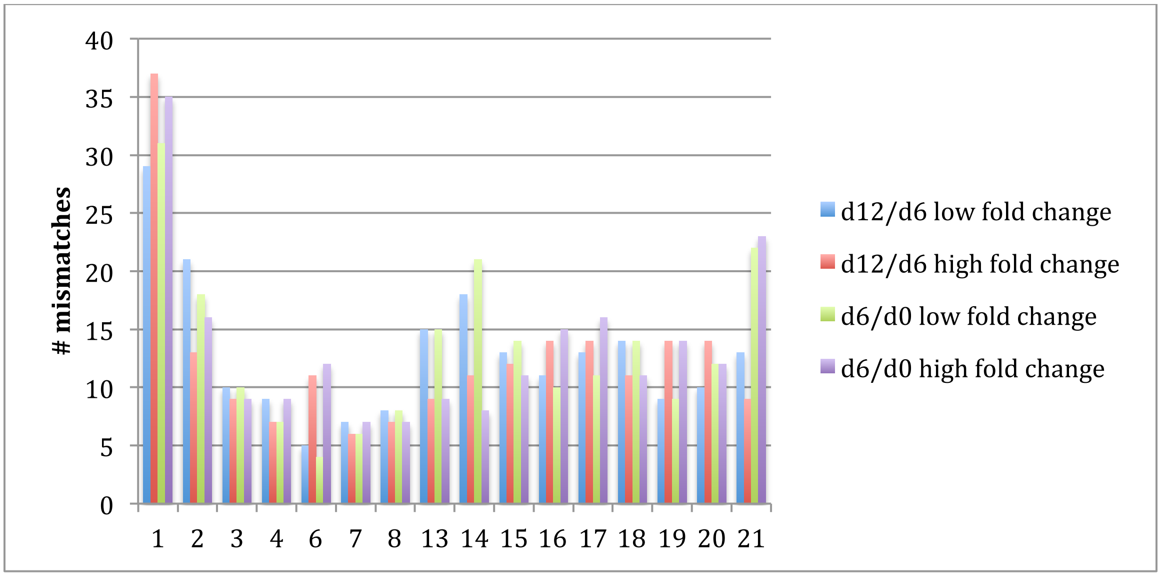




Supplemental Figure 5

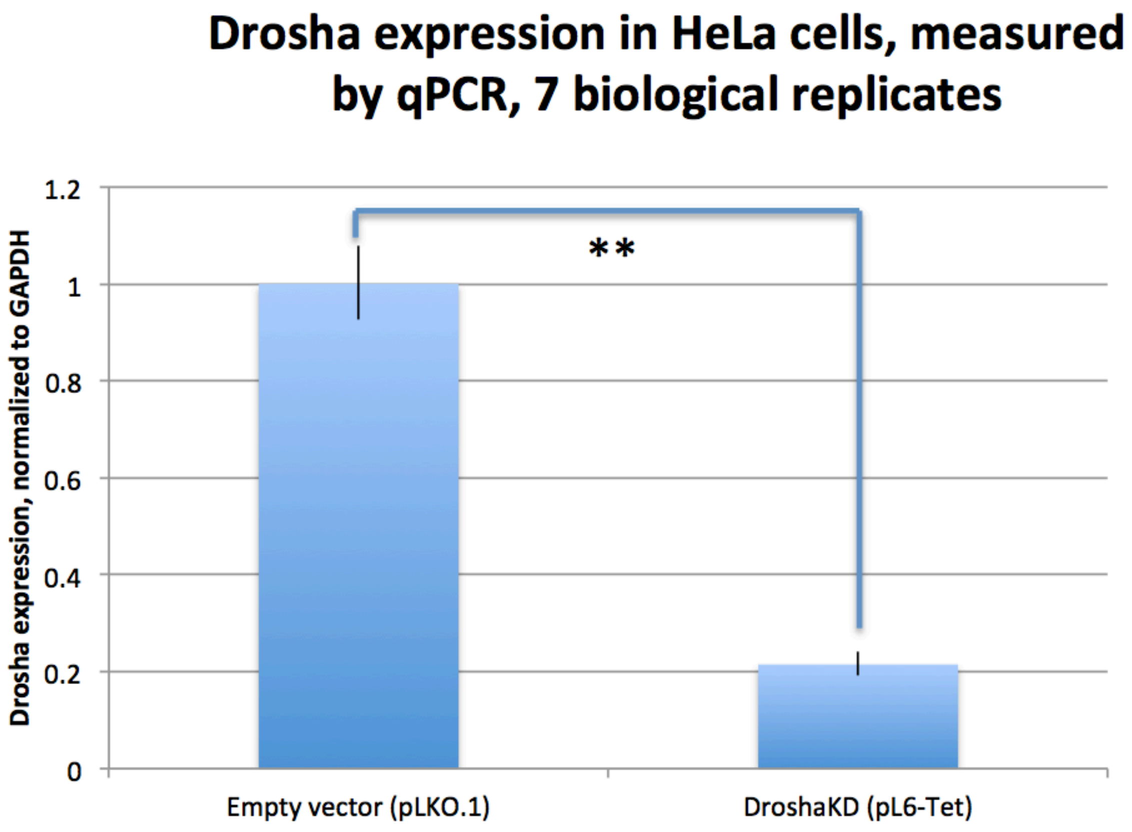


Supplemental Figure 6

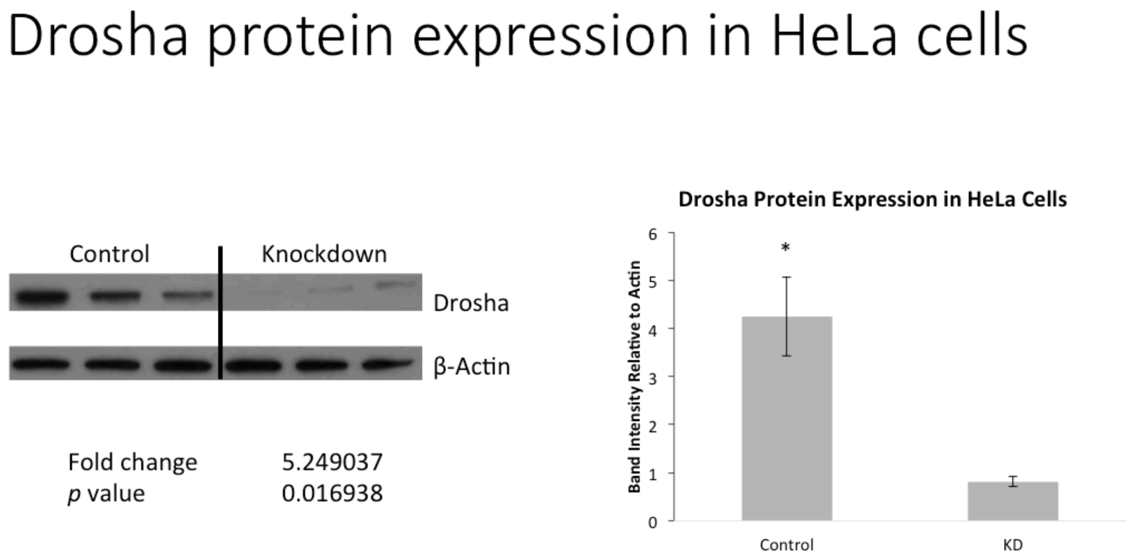


Supplemental Figure 7

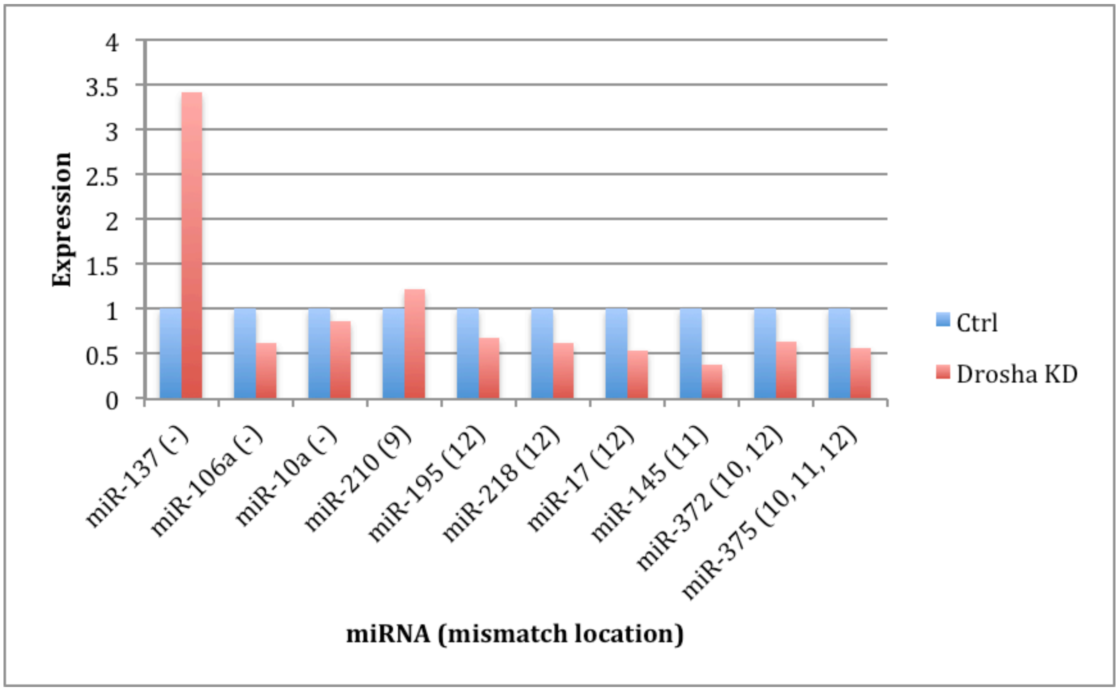
A



B



Supplemental Figure 8



Supplemental Table 1

Primer	Sequence
Drosha	SS-AGGAGTACGCCATAACCAACG AS-CAATCGTGGAAAGAAGCAGACA
β-actin	F-TCCCTGGAGAAGAGCTACG R-GTAGTTTCGTGGATGCCACA
Gapdh	SS-TGATGACATCAAGAAGCTGGTGAAG AS-TCCTTGGAGGCCATGTGGGCCAT
Bleomycin	F-AGCTGTACGCCGAGTGGTC R-CGTGTCAGTCCTGCTCCTC
Ubc	F-ATTTGGGTCGCGGTTCTTG R-TGCCTTGACATTCTCGATGGT

Supplemental Table 2

	ylw.d0	ylw.d3	ylw.d6	ylw.d9	ylw.d12
DROSHA	6.61E-05	6.48E-05	8.70E-05	9.82E-05	7.60E-05
DGCR8	2.98E-05	3.53E-05	6.86E-05	3.48E-05	8.01E-05
DICER1	3.83E-05	3.97E-05	5.31E-05	5.35E-05	6.29E-05
DICER1-AS1	1.04E-06	1.27E-06	2.28E-06	7.71E-06	1.97E-06
ALAS1	5.67E-05	4.96E-05	4.35E-05	4.89E-05	5.40E-05
TP53	0.000180321	0.000175473	0.000180299	0.000139297	0.000172342
DDX5	0.000351979	0.000339056	0.000554114	0.000523609	0.000720827
DDX17	0.000299841	0.000280771	0.000470649	0.000511757	0.000495815
TP73	2.50E-06	3.75E-06	1.53E-06	6.32E-07	1.24E-06
SMAD1	2.56E-05	2.73E-05	4.03E-05	2.88E-05	3.41E-05
SMAD2	3.15E-05	3.08E-05	4.00E-05	7.85E-05	3.98E-05
SMAD5	3.51E-05	4.80E-05	5.41E-05	5.79E-05	6.17E-05

NCL	0.001463131	0.001273564	0.001363324	0.000693873	0.000925033
HNRNPA1	0.003868665	0.004616134	0.005426	0.005911903	0.003730878
HNRNPL	0.000845167	0.000742587	0.000574737	0.000845478	0.000695673
KHSRP	0.00027734	0.000293381	0.000300315	0.000282278	0.000304843
LIN28A	0.001320614	0.001683845	0.001253371	0.001345877	0.000883196
LIN28B	0.000100119	0.000101629	0.000109379	0.000130416	0.000156404
PTBP1	0.000443271	0.000406256	0.000361956	0.000354963	0.000359039
PTBP2	5.33E-05	5.50E-05	6.86E-05	7.87E-05	6.16E-05
SRRT	0.000141233	0.000136973	0.000240314	0.000122967	0.000242403
XPO5	9.42E-05	9.11E-05	0.000107547	5.61E-05	8.06E-05
RAN	0.001606789	0.001492132	0.001397514	0.000888721	0.001287702
RANGAP1	0.000169424	0.000162503	0.00012513	0.000116093	0.000115304
EIF2C1	2.25E-05	1.99E-05	2.66E-05	2.62E-05	2.25E-05
EIF2C2	3.41E-05	2.66E-05	2.63E-05	3.30E-05	2.13E-05
EIF2C3	7.74E-06	7.98E-06	9.33E-06	1.54E-05	9.60E-06

EIF2C4	7.27E-06	8.88E-06	1.91E-05	3.10E-05	1.79E-05
TARBP2	6.72E-05	7.65E-05	6.52E-05	5.03E-05	6.38E-05
TRIM71	0.000146825	0.000137212	0.000122665	0.000126995	7.51E-05
ZCCHC11	2.87E-05	3.47E-05	5.47E-05	0.000133505	6.61E-05
DDX58	8.38E-06	6.66E-06	3.84E-06	3.17E-06	2.41E-06
RBM3	9.62E-05	9.09E-05	0.000138597	7.46E-05	0.000134169

Chapter II: The metabolome regulates the epigenetic landscape during naïve to primed human embryonic stem cell transition

Abstract

For nearly a century developmental biologists have recognized that cells from embryos can differ in their potential to differentiate into distinct cell types. Recently, it has been recognized that embryonic stem cells derived from both mice and humans display two stable yet epigenetically distinct states of pluripotency, naïve and primed. We now show that nicotinamide-N-methyl transferase (NNMT) and metabolic state regulate pluripotency in hESCs. Specifically, in naïve hESC NNMT and its enzymatic product 1-methylnicotinamide (1-MNA) are highly upregulated, and NNMT is required for low H3K27me3 repressive state. NNMT consumes SAM in naïve cells, making it unavailable for histone methylation that represses Wnt and activates HIF pathway in primed hESC. These data support the hypothesis that the metabolome regulates the epigenetic landscape of the earliest steps in human development.

Introduction

Pluripotent stem cells are able to self-renew and have the capacity to regenerate all tissues in the body. Unveiling the molecular mechanisms through which pluripotency is maintained holds tremendous promise for understanding early animal development and for developing therapies in regenerative medicine. Recent findings have revealed that pluripotency does not represent a single defined state; subtle states of pluripotency, with differences in measurable characteristics relating to gene expression, epigenetics and cellular phenotype, provide an experimental system for studying potential key regulators that constrain or expand the developmental capacity of pluripotent cells(Tesar et al., 2007;Buecker et al., 2014;Factor et al., 2014). Two stable pluripotent states have been derived in the mouse, and now in humans; preimplantation naïve and postimplantation primed ESC states.(Chan et al., 2013;Gafni et al., 2013;Theunissen et al., 2014;Valamehr et al., 2014;Ware et al., 2014) Since naïve, preimplantation human embryonic stem cells (hESCs) show higher developmental potential than postimplantation, primed hESCs(Gafni et al., 2013;Ware et al., 2014), it is critical to understand the key molecular differences between these pluripotent cell types.

Metabolic signatures are highly characteristic for a cell and are proposed to act as a leading cause for cell fate changes(Bracha et al., 2010;Yanes et al., 2010;Folmes et al., 2011;Greer et al., 2012;Panopoulos et al., 2012;Rafalski et al., 2012;Zhou et al., 2012;Mathieu et al., 2014). Recent data have shown that pluripotent stem cells have a unique metabolic pattern. The naïve to primed mouse ESC transition accompanies a dramatic metabolic switch from bivalent to highly glycolytic state(Zhou et al., 2012). However, primed state of inert mitochondria rapidly changes to highly respiring mitochondria during further differentiation. It is not yet understood how and why the pluripotent cells enter the highly glycolytic, metabolically cancer-like (Warburg effect) state and how a differentiating cell leaves this state.

In mouse embryonic stem cells (mESCs) threonine and S-adenosyl methionine (SAM) metabolism are coupled resulting in regulation of histone methylation marks(Shyh-Chang et al., 2013). Methionine and SAM are also required for the self renewal of hESCs, since depletion of SAM leads to reduced H3K4me3 marks and defects in maintenance of the hESC state(Shiraki et al., 2014). SAM therefore is shown to be a key regulator for maintaining ESC undifferentiated state and regulating their differentiation. However, little is known about SAM levels or its regulation during the transition between naïve and primed human embryonic states. Recent derivation of naïve human ESCs allows a deeper analysis of the human naïve to primed transition(Chan et al., 2013;Gafni et al., 2013;Theunissen et al., 2014;Valamehr et al., 2014;Ware et al., 2014). These studies have already revealed that the epigenetic landscape changes from the naïve to primed state through increased H3K27me3 repressive methylation marks. However, the regulation of this process or the metabolomics of this transition have not been dissected.

We now show that the increased H3K27me3 repressive methylation marks during naïve to primed hESC transition are regulated by metabolites and NNMT. The epigenetic repressive marks regulate HIF and Wnt pathways and thereby control the key metabolic switch observed during naïve to primed pluripotent transition. We show increased lipid biosynthesis, reduced beta-oxidation and reduced mitochondrial oxygen consumption in primed hESCs. Metabolic and gene expression analysis revealed that nicotinamide-N-methyl-transferase (NNMT) and its enzymatic product 1-methylnicotinamide (1-MNA) were highly upregulated, while the substrates nicotinamide and SAM were downregulated in naïve state, correlating with reduced H3K27me3 marks. Knockdown of NNMT in naïve hESCs increased H3K27me3 repressive marks in developmental as well as key metabolic genes that regulate the metabolic switch in naïve to primed transition. These data show that NNMT consumes SAM in naïve cells, making it unavailable for histone methylation. Histone methylation further regulates the key signaling pathways important for the metabolic changes that are necessary for early human development.

Results

A dramatic metabolic switch occurs in mouse ESCs between pre-implantation (naïve) and post-implantation (primed) state(Zhou et al., 2012). To assess the metabolic profiles of the human ESC (hESC) counterparts we analyzed the naïve and primed hESCs using a SeaHorse extracellular flux analyzer. As seen previously in mouse ESCs(Zhou et al., 2012), we detected an increase in oxygen consumption rate (OCR) after FCCP injection in the newly derived naïve hESCs Elf1(Ware et al., 2014), while little increase was observed in primed hESCs (H1 and H7) (Fig.1A-B, Suppl.Fig.1A). Likewise, cells "toggled" back to a more naïve state using GSK3 and MEK inhibitors plus bFGF (H1 2iF(Ware et al., 2014)) or GSK3, MEK, JNK and p38 inhibitors in addition to LIF, IGF and FGF (H1 4iLIF) showed increased OCR in response to FCCP to a level similar to mESCs (Fig1A-B, Suppl.Fig1B-C). These results indicate that the primed hESCs have a lower mitochondrial respiration capacity than naïve hESCs.

The higher mitochondrial capacity of naïve hESCs (Elf1) reflect neither more mature mitochondria(Ware et al., 2014), nor an increase in mitochondrial DNA copy number compared to primed hESCs (H7, H1) (Fig.1C, Suppl.Fig1D). Further, no obvious increase in mitochondrial mutation rate was detected in primed compared to naïve hESCs (Fig.1D, Suppl.Fig1E), suggesting that reduction of oxidative respiration in primed hESCs is not caused by a deteriorating mitochondrial genome. However, consistent with the mouse data(Zhou et al., 2012), HIF1α is stabilized in primed but not in naïve hESC (Fig.1E), correlating with a significant change in expression of prolyl hydroxylase domain-containing protein 2, PHD2 (EGLN1), the primary regulator of HIF1α steady state levels(Berra et al., 2003;Simonson et al., 2010) (Suppl.Table.1A). Further support for HIF1α stabilization and activity at the primed state comes from our proteomic analysis revealing a significant increase in the protein expression of HIF targets, Ldha and Jarid2 at primed hESC state (Elf1 AF compared to Elf1; Fig.1F-G; Suppl.Table.1B; Suppl.Fig.2). Also consistent with the mouse data, RNA-seq data from our study and microarray data from Gafni et al.(Gafni et al., 2013) and Theunissen et al.(Theunissen et al., 2014) showed that expression of most mitochondrial electron transport chain complex IV-Cytochrome oxidase c (COX) genes is significantly downregulated in the primed state compared to the naïve state (Fig.1H; Suppl.Fig.1F). A paired Wilcoxon rank-sum test resulted in a p-value of 1.73*10-4 in our data and 8.27*10-3 in Gafni et al.(Gafni et al., 2013) data and 1.49*10-8 in Theunissen et al.(Theunissen et al., 2014) data. These results suggest that as seen before with mouse pre-implantation and post-implantation state ESCs in vitro and ICM and Epi-state in vivo, post-implantation hESCs reduce their mitochondrial activity by reducing the abundance of complex IV of the electron transport chain.

Differential metabolites between naïve and primed embryonic stem cells

To search for critical metabolites that control the metabolic transitions between naïve and primed mouse and human ESCs, we performed metabolic profiling using non-targeted GC-TOF, LC-QTOF and targeted LC-QQQ mass spectrometry (MS) analysis. For the non-targeted MS analysis naïve and primed human and mouse ESCs were harvested using acetonitrile, sonicated and cleared by centrifugation to prepare for LC-QTOF and GC-TOF analysis. Spectral peaks were annotated using an in house annotation library and final area under the curve (AUC) values were normalized to average intensity, showing a good consistency between biological replicates as well as a clear difference in metabolite profiles between each condition (Fig.2A; Suppl.Table.1C-D).

Principal component analysis (PCA) reveals a difference in metabolite profiles between naïve and primed cells, regardless of species (Fig.2B-C, Suppl.Fig.3). Multiple naïve and primed cell lines from human and mouse separated clearly by naïve vs. primed state based on the PCA plot of GC metabolomics data (Fig.2B; Suppl.Fig.3). The first PC, which represents the separation of naïve vs. primed cell lines, explains 50.2% of the variance, whereas the second PC explains 14.5%. Stearic acid and cholesterol are the metabolites that contribute the most to the separation within the first PC, which indicates that when ESCs transition from naïve to primed state, a major switch occurs in the lipid metabolism. A similar trend of naïve and primed ESC separation is observed in PCA plots of the LC metabolomics data (Fig.2C). Interestingly though, H1 2iF, which is a primed cell line "toggled" towards the naïve state, clustered midway between truly naïve cell line (R1) and primed cell lines (R1AF, EpiSC and H1) in LC analysis, suggesting that H1 2iF has not reached the true naïve state with respect to lipid signature.

Based on the GC-MS analysis we identified a higher level of succinate in the mouse primed state compared to naïve (Fig.2D-E; Suppl.Table.1C). Succinate can act as an inhibitor for prolyl hydroxylase, thereby allowing HIF stabilization in the primed state. Hence succinate upregulation may be critical for primed state acquisition since HIF is stabilized in the primed state and this stabilization is shown to be sufficient to drive naïve to primed transition in mouse ESC(Zhou et al., 2012) (Fig1E).

In addition, targeted analysis of metabolites was performed using LC-QQQ-MS using Elf1 and H1 hESCs (Suppl.Table.1E-F). Metabolites consistently upregulated in primed state include fructose(1,6/2,6)-bisphosphate (F16BP), lactate, methionine, nicotinamide and kynurenine (Fig.2F-G). Fold change analysis of glycolysis metabolites detected by targeted LC-QQQ-MS shows an increase in primed H1 hESCs for metabolites in the early steps of glycolysis, and a decrease in metabolites in the late steps of glycolysis, relative to naïve Elf1 hESCs (Fig.2H-I). Upregulation of F16BP is in concord with highly active glycolysis, however, phosphoenolpyruvate (PEP), the downstream metabolite of F16BP, does not increase in primed hESCs (Fig.2I). Intermediates prior to PEP can be conserved for biosynthetic purposes: 3-phosphoglycerate (3PG) can be diverted to serine and glycine synthesis, which can supply one-carbon units to multiple methylation reactions (e.g., re-generation of methionine from homocysteine in the SAM cycle); dihydroxyacetone phosphate (DHAP) can be converted to glycerol, which serves as the backbone of glycerolipids. Therefore we tested potential changes in lipid/fatty acid metabolism and amino acid pathways (Fig.2H).

Differential Fatty Acid metabolism in Naïve and Primed ESCs

Further lipid analysis was performed using a LC-QTOF instrument on naïve Elf1 and primed H1 cells, grown on matrigel for one passage and harvested using a 9:1 methanol:chloroform solution (Suppl.Table.1G). All lipidomic features were classified into two groups: more abundant in H1 or more abundant in Elf1. For 119 features with identified molecular formulae and structures, lipids more abundant in H1 have higher numbers of carbons (Wilcoxon rank sum test p-value 4.30e-4, Fig.3A). For 320 features with just identified mass, lipids more abundant in H1 are heavier (p-value=1.66e-10, Fig.3B). In mESCs, when sorting based on level of unsaturation lipids more abundant in primed R1AF have a higher number of double bonds than lipids more abundant in naïve R1 (p-value=0.044, Fig.3C).

In concordance with the significant increase in long carbon chain lipids observed in primed mouse and human ESCs, we also detected a significant increase in accumulation of lipid droplets in the primed state, as observed by Oil Red O and BODIPY staining (Fig.3D-E,). These data indicate an increased synthesis and/or decreased beta-oxidation in primed cells.

To identify the cause for accumulation of lipids in primed ESCs, we searched for differences among the enzymes involved in fatty acid metabolism. Interestingly, several of the enzymes involved in fatty acid transport into the mitochondria and fatty acid beta-oxidation are significantly downregulated in primed human ESCs, as well as in mouse in vivo post-implantation state (Fig.3F-G; Suppl.Fig4A-B; Suppl.Table.1G). Carnitine acyltransferase 1 (CPT1) transfers long chain acyl groups to carnitine. This enzyme is responsible for a very important step of mitochondrial fatty acid beta-oxidation by facilitating the initial step in acyl transfer to the mitochondrial matrix. Three isoforms have been described: CPT1A, CPT1B and CPT1C; however, CPT1C is not involved in fatty acid beta-oxidation(Wolfgang et al., 2006). Interestingly, the rate limiting fatty acid transporter CPT1A is downregulated in both mouse *in vivo* post-implantation and human primed ESC state (Fig.3F-G). The decrease in CPT1A expression in the primed state could be due to increased H3K27me3 and decreased H3K4me3 and H3K27ac marks observed in CPT1A promoter in primed hESCs(Chan et al., 2013;Gafni et al., 2013;Ware et al., 2014) (Fig3H; Suppl.Fig5A). miRNA analysis of Elf1 compared to primed H1 cells reveals that several of the miRNAs predicted to target CPT1A and other enzymes involved in beta-oxidation are up-regulated in primed hESCs (e.g., miR-9, miR-33a-3p, miR-124; Fig.3I; Suppl.Table.1H). Moreover, microRNAs predicted by Targetscan and miRTarBase(Lewis et al., 2005;Hsu et al., 2014) to target enzymes involved in fatty acid synthesis (Fig.3I) were downregulated in primed cells (e.g., miR-10a and miR-193). Concomitantly, key fatty acid synthesis genes were up in primed H1 hESCs compared to naïve Elf1 state(Watkins et al., 2002) (SLC25A1, ACLY, ACACA, FASN, and SREBP-1c; Suppl.Fig6). We further validated some of the miRNAs by qPCR analysis and showed that miR-9, predicted to target CPT1A, was upregulated, while miR-10a, predicted to target SREBP-1c (a regulator of fatty acid and cholesterol synthesis), was downregulated in the human primed state (Fig.3J).

To test the level of fatty acid beta-oxidation in naïve and primed human and mouse ESCs, we performed a palmitate-oxidation assay in the Seahorse metabolic flux analyzer(Nguyen-Tran et al., 2014). Importantly, both mouse and human naïve ESCs were capable of utilizing palmitate as an energy source, while primed mouse or human ESCs were not (Fig.3 K-M; Suppl.Fig5B). This result suggests that primed human and mouse ESCs are not capable of significant beta-oxidation and, in combination with increased fatty acid synthesis, may explain the accumulation of lipids observed in this state (Fig.3D-E).

Differential amino acid metabolism in naïve and primed ESCs

In addition to glycolysis and fatty acid metabolism, primed cells show changes in amino acid metabolism pathways (Fig.4A-B). In primed vs. naïve hESCs we observed a large enrichment of the tryptophan degradation product kynurenine, which can act as a ligand for the nuclear receptor AHR(Opitz et al., 2011) (Fig.4B). Interestingly, tryptophan is shown to be critical for primed hESCs growth(Shiraki et al., 2014). RNAseq and qPCR data show a large increase of the tryptophan metabolizing enzyme IDO1 in primed hESCs (Fig.4C,F), providing further evidence of a major change in the tryptophan degradation pathway. IDO1 is also upregulated in primed hESCs compared to the in vivo 8 cell human embryo(Yan et al., 2013) (Suppl.Fig.7A). After peaking in primed hESCs IDO1 levels quickly drop when the hESCs begin to differentiate, indicating that the function of IDO1 is specific for the primed state(Ulanovskaya et al., 2013) (Suppl.Fig.7B-C). In addition to the consistent increase from the naïve to primed state in kynurenine vs. tryptophan ratios observed in intracellular metabolite levels (Fig.4D), secreted kynurenine can be observed in the media of primed hESCs (Fig.4E). Future identification of the function of kynurenine in this cell type might reveal why tryptophan is essential for primed hESC growth.

Methionine and nicotinamide downregulation along with 1-methyl-nicotinamide (1MNA) upregulation in the naïve state correlates with upregulation of Nicotinamide N-methyltransferase (NNMT), shown previously to create a metabolic methyl sink, thereby promoting epigenetic remodeling in cancer (Ulanovskaya et al., 2013) (Fig.4A; Fig.4G-H; Suppl.Fig8A-B). Primed hESCs show an increase in SAM and a decrease in SAH levels compared to the naïve state (Fig.4I; Suppl.Table.1I). The increase in SAM correlates with the sharp decrease in NNMT enzyme levels observed in primed hESCs(Chan et al., 2013;Gafni et al., 2013;Valamehr et al., 2014) (RNA-seq, microarray and qPCR data; Fig.4F-G; Suppl.Fig8B), suggesting that SAM levels may be reduced in the naïve state by high NNMT activity. Further, significant expression changes of NNMT among various tissues reveal that NNMT is dynamically regulated during development and suggest that NNMT might act as regulator of SAM levels also in a developmental context, not only in cancer(Kraus et al., 2014) (Suppl.Fig.9).

NNMT regulates repressive histone modifications

Reduction of NNMT levels during the naïve (Elf1) to primed transition correlates with a significant increase in SAM levels and in H3K27me3 histone methylation marks in 648 developmentally regulated genes(Ware et al., 2014) (Fig.5A). This correlation is observed in other naïve published lines(Chan et al., 2013;Gafni et al., 2013;Theunissen et al., 2014) (Fig.5A-B). Moreover, Western blot analysis revealed an overall increase of H3K27me3 and H3K9me3 marks in primed hESCs (H7) compared to naïve hESCs (Elf1), while the H3K9/K14 acetylation marks remained unchanged (Fig.5C). ChIP-seq analysis of other marks (H3K4me1, H3K4me3 and H3K27ac) did not show any significant change between primed and naïve hESCs(Chan et al., 2013;Gafni et al., 2013) (Suppl.Fig.10). RNA-seq analysis of histone methyltransferases and histone demethylases involved in H3K27 and H3K9 methyl marks did not show changes in expression levels that could explain the significant increase in repressive methylation marks observed at the primed state (Suppl.Fig.11). Furthermore, we showed by Western analysis that the protein levels of PRC2 regulator, EED is not significantly changed between the two pluripotent states (Fig.5C).

To test whether NNMT is causal for low H3K27me3 and H3K9me3 levels in naïve hESC, we knocked down NNMT in Elf1 using siRNA (80% reduction of NNMT; Fig.5D) and analyzed the histone methylation marks. This direct reduction of NNMT levels in naïve hESCs significantly increased H3K27me3 and H3K9me3 marks, as analyzed by Western blots, while H3K9/K14 acetylation marks did not change (Fig.5E), suggesting that NNMT is a key regulator of the histone repressive methylation marks in the naïve state of pluripotency.

NNMT levels were also altered by inhibiting the LIF/STAT pathway. The LIF/STAT pathway was activated in naïve hESCs since they were grown in media supplemented with LIF. LIF is known to activate STAT3(Graf et al., 2011), which has been shown to bind to the NNMT promoter and activate its transcription(Tomida et al., 2008). We hypothesized that treating naïve hESCs with a STAT3 inhibitor might affect NNMT expression and the repressive histone methylation pattern of those cells. qPCR analysis showed a reduction of NNMT expression on Elf1 cells as early as 6h after STAT3 inhibitor addition (Suppl.Fig12A). Importantly, reduction of NNMT in naïve hESCs by STAT3 inhibitor also increased H3K27me3 and H3K9me3 marks, as shown by Western blot analysis (Fig.5F). We characterized H3K27me3 in naïve hESCs by ChIP-seq analysis and observed a significant increase in H3K27me3 marks at promoters after 6h STAT3 inhibitor treatment (Fig.5G; Suppl.Fig.12B; Suppl.Table.1J-K). Interestingly, over 25% (313 genes) of genes with primed-enriched H3K27me3 marks began to increase H3K27me3 marks after 6 hour STAT3 inhibitor treatment (Fig.5N; Suppl.Table.1L). Windowed chromatin heatmaps revealed a dramatic increase in H3K27me3 marks close to the transcription start site of these 313 genes in the naïve to primed transition as well as after 6h STAT3 inhibitor treatment(Gafni et al., 2013;Ware et al., 2014) (Fig.5H).

To determine the functional consequence of H3K27me3 marks after STAT3 inhibitor treatment, we analyzed the overlapping 313 genes through RNA-seq and ChIP-seq data. The majority of these genes are developmentally regulated genes that are repressed in primed compared to naïve state (Suppl.Table.1L). Importantly, among the 313 genes with repressive H3K27me3 marks are Wnt ligand Wnt5, and Wnt targets, ZEB1, ZEB2 and SLUG. Further, the majority of Wnt ligands and target genes are downregulated in primed compared to naïve hESC, suggesting that

Wnt pathway might be inactivated during the naïve to primed transition (Fig.5I-J; Suppl.Fig.13; Fig.1G ROES and VIME). The Wnt pathway may additionally be downregulated by the primed state enriched miRNAs, miR-33a and miR-200 (Fig.3I; Suppl.Fig.15A), which are predicted regulators of Wnt and Zeb respectively (Targetscan). In addition, miR-155-5p, miR-148-3p and miR-130a-3p which have been shown to target JARID2(Hafner et al., 2010;Xu et al., 2010) were all upregulated in naïve compared to primed state hESCs, consistent with the observed upregulation of Jarid2 protein in primed hESCs (Fig.1G). Previous studies have revealed that in human and mouse primed ESC the Wnt pathway is not active and forced activation of the pathway leads to differentiation(ten Berge et al., 2011;Davidson et al., 2012). We now analyzed the activity of the Wnt pathway in naïve hESC using Wnt-pathway activity reporter(Davidson et al., 2012) and revealed that while the reporter is not activated in primed hESC, strong activation is observed in naïve hESCs (Fig.5K; Suppl.Fig.14). We furthermore show that the Wnt-activity in naïve hESC is dependent on β-catenin since siRNA(β-cat) or XAV treatment dramatically downregulated the reporter activity (Fig.5L). In addition, the Wnt ligand is produced by the naïve hESC since IWP2, an inhibitor that represses Wnt palmitylation also represses the reporter activity in naïve cells (Fig.5L). These data reveal that the robust Wnt activity in naïve hESC is among the earliest responders to the repressive H3K27me3 marks during naïve to primed hESC transition.

Another gene category of interest in the group of 313 early H3K27me3 responders is metabolic genes. A dramatic increase of H3K27me3 marks was observed in prolyl hydroxylase 2 (EGLN1), ECHS1, HIGD1 and miR-193 promoters in the primed state as well as in STAT3 inhibitor treated Elf1 cells, compared to naïve state (Fig.5M-N; Suppl.Table.1L). The repressive marks in these promoters correlated with the observed reduced gene expression in the primed state (Suppl.Table.1L). Since EGLN1 induces HIF1 degradation, its repression in primed state (Fig.5M) suggests that HIF1 is stabilized. Our analysis of HIF1 protein in the naïve and primed states supports this hypothesis (Fig.1E). miR-193 is predicted to target enzymes involved in oleic acid biosynthesis (FADS and PTPRT, TargetScan; Mirtarbase). Our metabolomic analysis showed that oleic acid levels are up in primed hESCs compared to naïve, suggestive of increased fatty acid synthesis (Suppl.Table.1C; Suppl.Fig.15B). Furthermore, repressive epigenetic marks and reduced expression of ECHS1 and HIGD1a correlate with reduced beta-oxidation and electron transport chain activity in primed hESC (Fig.5N).

In summary, downregulated NNMT expression results in increased levels of SAM and induction of H3K27me3 repressive marks. Jarid2 protein upregulation in primed state may give further regulation/target specificity for the H3K27me3 marks generated by PRC2 (Fig.1F) since Jarid2, the catalytically inactive demethylase is an essential component of PRC2 in ESC(Landeira and Fisher, 2011;Escobar et al., 2014). NNMT downregulation induced repression of Wnt pathway and other developmentally regulated genes as well as genes involved in the metabolic switch support the dramatic reduction of mitochondrial activity observed in the primed state and suggest that NNMT reduction moves the cells towards the primed state.

Discussion

We show that human naïve and primed ESCs display distinct metabolic profiles. We further show that switching between these metabolic states is regulated by NNMT, which controls the amount of SAM available for PRC2 dependent H3K27me3 histone methylation. Repressive histone methylation then controls the primed hESC specific metabolism through the Wnt and HIF pathways. The naïve to primed hESC transition shows a reduction in Wnt signaling and fatty acid beta-oxidation and increase in mechanisms involved in lipid biosynthesis and HIF1a stabilization. In naïve hESC NNMT and its enzymatic product 1-MNA are highly upregulated, while the substrates, nicotinamide and SAM are downregulated, correlating with reduced H3K27me3 marks. Inhibition of STAT3 in naïve hESCs increases H3K27me3 repressive marks in developmental and metabolic genes, including Wnt signaling and HIF1 repressor, prolyl hydroxylase EGLN1. These data show that NNMT consumes SAM in naïve cells, making it unavailable for histone methylation that represses Wnt pathway and fatty acid beta-oxidation and activates HIF pathway and lipid synthesis, facilitating the metabolic switch in the naïve to primed hESC transition (Fig.5N). The working hypothesis is that differential metabolites between pluripotent states may control epigenetic dynamics and signaling.

Primed ESCs are dependent on glycolysis, as has been previously described(Prigione and Adjaye, 2010;Folmes et al., 2011;Varum et al., 2011;Zhang et al., 2011;Panopoulos et al., 2012;Zhou et al., 2012). We now show that while early glycolysis metabolites are upregulated, the downstream metabolites are downregulated in primed state hESCs. We hypothesize that in primed hESCs, midway through the glycolysis, metabolites are being channeled off to increase the amount of glycerol backbone available for biosynthesis of lipids in primed cells, which - along with reduced beta-oxidation - would explain the observed buildup of large lipids in primed cells. The same build up of glycolysis intermediates (e.g. 3-phosphoglycerate) can be channeled to serine and glycine synthesis, which provide one-carbon unit for many methylation reactions, including methyl group donation to homocysteine to regenerate methionine. This would explain the observed build up of SAM in primed hESCs, with methionine as a substrate. In addition to the amino acid based regulation of SAM in hESCs, SAM can also be regulated by NNMT, which is dramatically downregulated in primed compared to naïve hESCs, making SAM available as a substrate for DNA and histone methylation. We show that a difference in SAM levels between naïve and primed hESCs correlates with dramatic changes in H3K27me3 marks. These findings could explain the dramatic difference in H3K27me3 marks observed between naïve and primed hESC(Chan et al., 2013;Gafni et al., 2013;Theunissen et al., 2014;Valamehr et al., 2014;Ware et al., 2014) and reveal NNMT as a key regulator of these changes. While previous studies have shown SAM dependent regulation of histone methylation in stable primed hESCs, the effect was mainly in H3K4me3, not in H3K27me3 marks(Shiraki et al., 2014). It is plausible that H3K27me3 marks, once established are less dynamic in primed hESCs than H3K4me3 marks. Indeed, previous studies have suggested that H3K27me3 might be more stable than other methylation marks(Trojer and Reinberg, 2006), suggesting that transition from the naïve to primed state requires increased substrate levels for H3K27me3 methyltransferases, but once these marks exist, reduction of SAM levels in the primed state has no immediate effect on these marks.

While H3K27me3 marks are reduced in naïve compared to primed hESCs, interestingly the enzymes required for this methylation mark (EZH2/EED) are not downregulated. This prompts the following question: how are the methylases regulated if not by expression level in these stem cell states? One possibility is the presence of compounds that control their activity and/or access to sufficient amounts of substrate. We have shown that NNMT is required for naïve hESC epigenetic regulation. NNMT can reduce SAM, a substrate for EZH2, and hence potentially reduce the number of H3K27me3 marks. Furthermore the PRC2 recruiting protein Jarid2 is upregulated in primed hESCs compared to naïve, which may give further specificity to PRC2 action in naïve to primed hESC transition(Landeira and Fisher, 2011;Escobar et al., 2014)(Fig.1G, 5M). miR-155-5p, miR-148-3p and miR-130a-3p, which are confirmed to target JARID2 show an inverse expression pattern compared to Jarid2, suggesting that upregulation of Jarid2 protein in primed compared to naïve state may be under microRNA control.

In this study we show a direct impact of STAT3 pathway and NNMT function on histone marks in naïve hESC through knockdown experiments, revealing that changes in the metabolic profile of hESCs shape the epigenetic landscape during hESC development. We propose that these epigenetic changes in hESC govern the signaling pathway activities required for the metabolic switch in the naïve to primed transition. Previous studies have revealed the importance of Wnt and HIF pathways in mouse naïve to primed pluripotent stem cell transition(Mazumdar et al., 2010;ten Berge et al., 2011;Zhou et al., 2012). Further studies are required to dissect the exact requirements and potential interactions of these two pathways in human naïve to primed transition. These studies pave the way for further understanding and utilization of metabolite specific molecular mechanisms for cell fate changes in general.

Materials and methods

Culture of primed and naïve embryonic stem cells

Primed human ESCs [H1 (WA-01) and H7 (WA-07)] and naïve hESCs [Elf-1(NIHhESC-12-

0156)] were cultured as previously described(Ware et al., 2006;Ware et al., 2014). Briefly, the cells were cultured on a feeder layer of irradiated primary mouse embryonic fibroblasts (MEF) in hESC media: high glucose (3.151g/L) DMEM/F-12 media supplemented with 20% serum replacer, 1 mM sodium pyruvate, 0.1 mM nonessential amino acids, 50 U/ml penicillin, 50 μg/ml streptomycin (all from Invitrogen, Carlsbad, CA) and 0.1 mM β-mercaptoethanol (Sigma-Aldrich, St. Louis, MO). hESC media was supplemented with 4 ng/ml basic fibroblast growth factor (bFGF) for primed hESCs and 1μM GSK3 inhibitor (CHIR99021, Selleckchem), 1μM of MEK inhibitor (PD0325901, Selleckchem), 10ng/mL human LIF (Chemicon), 5ng/mL IFG1 (Peprotech), 10ng/mL bFGF for naïve hESCs, Elf1 (2iLIF). One passage prior to the experiments, the cells were transferred to growth factor reduced Matrigel (Becton Dickinson, Moutainview, CA) in MEF conditioned media (CM). Dispase and Trypsin/EDTA (Invitrogen) were used to passage primed and naïve hESCs, respectively. To reverse toggle, H1 hESCs were first cultured for 3 passages in presence of HDAC inhibitors [50nM SAHA (Cayman) and 0.1mM Butyrate (Sigma-Aldrich)], followed by 1μM CHIR99021, 1μM PD0325901 and 10ng/mL bFGF (H1 2iF) for 3 passages. Alternatively, H1 cells were pushed toward a more naïve state by culture in 4iLIF media: 1μM GSK3 inhibitor (CHIR99021), 1μM of MEK inhibitor (PD0325901), JNK inhibitor 10μM (SP600125, Selleck), 10μM p38 inhibitor (SB203580, Selleck), 10ng/mL human LIF, 5ng/mL IFG and 10ng/mL bFGF for 3 passages. In addition, human naïve Elf1 cells were pushed toward a more primed state by culturing them in Activin (10ng/mL) and bFGF (10ng/mL) for 3 passages (Elf1 AF).

Naïve mouse ESC (R1) were cultured in DMEM media supplemented with 20% FBS (ES qualified, Invitrogen), 1.5μM CHIR99021 and 1μM PD0325901 and mouse LIF at

1,000units/mL (Chemicon). Primed mouse ESCs (EpiSCs) were cultured in hESC media supplemented with activin A (10ng/mL) and bFGF (10ng/mL). mESC R1 were toggled to a primed state with addition of activin A (10ng/mL) and bFGF (10ng/mL) (R1 AF) for 3 passages. All cells were grown at 37 degres, 5%CO_2 and 5% O_2.

OCR measurement using SeaHorse Cellular Flux assays

Naïve and primed ESCs were seeded in their specific growth media onto 96-well SeaHorse plates pre-coated with Matrigel at $25x10^4$ or $40x10^4$ cells/well. Culture media were exchanged for base media (unbuffered DMEM (Sigma D5030) supplemented with sodium pyruvate (Gibco, 1mM) and with 25mM glucose (for MitoStress assay), 25mM glucose and 50µM carnitine (for palmitate assay), or 2mM glutamine (for glucose stress assay) 1 h prior to the assay and for the duration of the measurement. Substrates and selective inhibitors were injected during the measurements to achieve final concentrations of glucose (2.5mM), 4-(trifluoromethoxy)phenylhydrazone (FCCP, 300nM-500nM), oligomycin (2.5 µM), antimycin (2.5µM), rotenone (2.5µM), palmitate (50µM in BSA), BSA and ETO (50µM). Mitochondrial stress protocol starts with the measurement of baseline oxygen consumption rate (OCR) followed by measurement of OCR changes in response to injection of oligomycin, FCCP and finally antimycin and rotenone. Palmitate assay starts with the measurement of baseline OCR followed by measurement of OCR changes in response to injection of palmitate or BSA in negative controls) and ETO. The OCR values were further normalized to the number of cells present in each well, quantified by the Hoechst staining (HO33342; Sigma-Aldrich) as measured using fluorescence at 355nm excitation and 460nm emission (Envision 2104 Multilabel reader, Perkin Elmer). Changes in OCR in response to substrates and inhibitors addition were defined as the maximal change after the chemical injection compared to the last OCR value before the injection.

Mitochondrial DNA copy number measurement

The DNA was isolated using DNAzol (Invitrogen) following manufacturer's protocol. TaqMan primers were used to quantify mitochondrial and genomic DNA (Suppl.Table.2). The ratio of mtDNA to genomic DNA was determined using a standard curve for each primer. Each reaction contained 2 ng of DNA extract, 1x TaqMan Universal PCR Master Mix No AmpErase UNG, 500nM of each primer, and 200nM of the TaqMan probe. Using 7300 real time PCR system (Applied Biosystems), the reactions were amplified by incubation at 50°C for 2 minutes, 95°C for 10 minutes, and then 40 cycles of 15 seconds at 95°C followed by 1 minute at 60°C where the intensity of fluorescence was measured.

Mitochondrial mutation frequency analysis

Droplet digital PCR (ddPCR) was performed on naïve hESCs (Elf1 2iL at passage 19: 0.4 uM PD0325901, 3.0 uM CHIR99021 and 10 ng/m) and primed hESCs H1 (passage 65) as previously described (Taylor et al., 2014). DNA was isolated with the DNeasy Blood and Tissue Kit (QIAGEN) according to kit instructions. Rare mutation-bearing molecules were selectively enriched through endonucleolytic destruction of wild-type target sites. DNA was digested with 100 U of TaqI (New England Biolabs, Ipswich, MA, USA) at 65°C for 10 h. An additional 100 U of TaqI was added to each reaction every hour. The final concentration of digested DNA was adjusted to yield less than ~3500 positive molecules per µL, which is within the range of linearity for the Poisson calculation (Pinheiro et al., 2012). Reaction droplets were made by applying ddPCR Master Mix (Bio-Rad, Hercules, CA, USA), 250 nM TaqMan probe, 900 nM of each flanking primer, and TaqI-digested DNA to a droplet generator DG8 cartridge (Bio-Rad) for use in the QX100 Droplet Generator (Bio-Rad). Following droplet generation, the droplet emulsion was transferred to a Twin.tec semi-skirted 96-well PCR plate (Eppendorf, Hamburg, Germany) and fragments for point mutation detection and mtDNA copy number measurement were amplified as follows: 95°C for 10 min, followed by 50 cycles of 94°C for 30 s, and 63.5°C for 2 min. The thermally cycled droplets were analyzed by flow cytometry in a QX100TM Droplet DigitalTM Reader (Bio-Rad) for fluorescence analysis and quantification of mutation frequencies. Point mutation or deletion frequency was quantified using a primer set that flanks the test region and measures the concentration of mutation-bearing molecules and a second primer set that flanks a region in the mitochondrial genome that bears no restriction recognition sites (described in Suppl.Table.2). Point mutation frequency per base pair is calculated by taking the ratio of the normalized concentrations of mutation-bearing mtDNA molecules to the total mtDNA molecules screened, divided by the number of bases per target site. Reactions that yielded < 10 positive droplets per well were scored conservatively as having no positives above background.(Pinheiro et al., 2012)

Proteomics

Naïve hESCs (Elf1 2iLIF) and primed hESCs (Elf1 AF) were washed in 1xPBS and flash frozen. Cell pellets were lysed in 1M urea, 50mM ammonium bicarbonate, pH 7.8, and heated to 50°C for 20min. Cell debris was removed by centrifugation (1000g, 2min). Following a BCA assay, normalized quantities of protein were reduced with 2mM DTT, alkylated with 15mM iodoacetamide, and digested overnight with a 1:50 ratio of trypsin to total protein. The resulting peptides were desalted on Waters Sep-Pak C18 cartridges.

Peptides were measured by nano-LC-MS/MS on a Thermo Scientific Fusion. Peptides were separated online by reverse phase chromatography using a heated 50°C 30cm C18 columns (75mm ID packed with Magic C18 AQ 3µM /100A beads) in a 180min gradient (1% to 45% acetonitrile with 0.1% formic acid) separated at 250nL/min. The Fusion was operated in the data-dependent mode with the following settings: 60000 resolution, 400-1600 m/z full scan, Top Speed 3 seconds, and an 1.8 m/z isolation window.

Identification and label free quantification of peptides was done with MaxQuant 1.5 using a 1% false discovery rate (FDR) against the human Swiss-Prot/TrEMB database downloaded from Uniprot on October 11th, 2013. The databases contained forward and reverse mouse sequences as well as common contaminants. We analyzed two biological and three technical replicates per condition. Peptides were searched using a 5ppm mass error and a match between run window of 2min. Proteins that were significantly regulated between conditions were identified using a permutation-based t-test (S1, FDR 5%) in Perseus 1.4.1.3.

Non-targeted GC-TOF and LC-QTOF analysis for metabolites in mouse and human ESCs

Cells (5 x 10 cm plates for each H1 replicate - equivalent of 5-10 million cells, and 2 x 10 cm plate for each R1 replicate - equivalent of 10-20 million cells) were scraped in PBS, pelleted in 6 replicates per condition, quick frozen and stored at -80 ºC, then thawed on ice and mixed with 2 mL of ice-cold degassed acetonitrile followed by vortexing for 20 sec, sonication for 5 min, and vortexing for 20 sec. 1 mL material was taken and centrifuged for 5 min at 14,000 rcf and the remaining material was saved. Supernatant was divided into a 500 µL (GC-TOF) and a 250 µL (LC-QTOF) aliquot, which were lyophilized.

GC-TOF MS Extraction and Analysis

Lyophilized material was resuspended in 500 µL of 1:1 acetonitrile:H_2O and centrifuged for 2 min at 14,000 rcf to remove membrane lipids and triglycerides. Supernatant was collected and lyophilized. GC-TOF analysis and data processing was performed as previously described (Meissen et al., 2012) using a Leco Pegasus IV time of flight mass spectrometer (Leco Corporation, St. Joseph, MI) coupled to an Agilent 6890 gas chromatograph (Agilent Technologies, Santa CA) equipped with a 30m long 0.25mm id Rtx5Sil-MS column and a Gerstel MPS2 automatic liner exchange system (Gerstel GMBH & Co. KG, Mülheim an der Ruhr, Germany).

LC-QTOF MS Extraction and Analysis

Lyophilized material was redissolved in 100 µL initial LC gradient solvent and analyzed within 24 hrs. HILIC and Reversed Phase LC-QTOF analysis and data processing was performed as previously described (Meissen et al., 2012) using an Agilent 1200 series HPLC equipped with either Agilent Zorbax Eclipse Plus C18 2.1x150mm column for Reversed phase or a Waters 1.7 µm Acquity BEH HILIC 2.1x150mm column. LC eluents were analyzed with an Agilent 6530 accurate mass Q-TOF mass spectrometer. The LipidBlast database was used for identifications, also as previously described(Meissen et al., 2012;Kind et al., 2013).

LC-QTOF for Elf1 and H1 hESCs

For lipid extraction Elf1 and H1 cells were grown on matrigel coated 35 mm plates for one passage. Cells were washed with PBS, followed by a 2 sec wash with 37 °C deionized water and the addition of 0.5ml of a -75 °C solution and internal standards as previously described.(Nygren et al., 2011) The plates were incubated on dry ice for 15 min before scraping the plates and transferring everything into eppendorf tubes. 1ml of chloroform was added to the tubes, followed by another 15 min incubation on dry ice. The mixture was spun for 5 min at 4 °C at 18000 rcf, after which the lower phase was collected and stored at -80 °C.

LC-QTOF-MS experiments were performed using an Agilent 1200 SL LC system coupled online with an Agilent 6520 Q-TOF mass spectrometer (Agilent Technologies, Santa Clara, CA). Each prepared sample (4 μL for positive ESI ionization, 8 μL for negative ESI ionization) was injected onto an Agilent Zorbax 300 SB-C8 column (2.1 x 50 mm, 1.8-micron), which was heated to 50 ^{o}C. The flow rate was 0.4 mL/min. Mobile phase A was 5 mM ammonium acetate and 0.1% formic acid in water, and mobile phase B was 5% water in ACN containing 5 mM ammonium acetate and 0.1% formic acid. The mobile phase composition was kept isocratic at 35% B for 1 min, and was increased to 95% B in 19 min; after another 10 min at 95% B, the mobile phase composition was returned to 35% B. The ESI voltage was 3.8 kV.

Targeted LC-QQQ-MS analysis for water soluble metabolites

Elf1 and H1 cells were grown on matrigel coated 35mm plates (3 plates per replicate) for one passage. Cells were washed with PBS, followed by a 2 second wash with ice cold deionized water and the addition of a -75 °C 0.75 mL 9:1 methanol:chloroform solution. The plates were incubated on dry ice for 15 min before scraping the plates and transferring the cellular debris into eppendorf tubes, which were spun at 18000 rcf for 5 min at 4 °C. All soluble extract was transferred into a new eppendorf tube and vacuum dried. Samples were stored in -80 °C before preparing for MS analysis.

Chromatography conditions: dried samples were reconstituted in 200 µL 5 mM ammonium acetate in 40% water/60% acetonitrile + 0.2% acetic acid, and filtered through 0.45 µm PVDF filters (Phenomenex, Torrance, CA) prior to LC-MS analysis. LC-MS/MS was performed using an Agilent 1260 LC AB-Sciex 5500 QQQ MS (Dai et al., 2014). Both chromatographic separations were performed in HILIC mode on two SeQuant ZIC-cHILIC columns (150 x 2.1 mm, 3.0 µm particle size, Merck KGaA, Darmstadt, Germany). While one column was performing the separation, the other column was reconditioned for the next injection. The flow rate was 0.300 mL/min, auto-sampler temperature was kept at 4 °C, the column compartment was set at 40 °C, and total separation time for both ionization modes was 20 min. The mobile phase was composed of Solvents A (5 mM ammonium acetate in 90%H_2O/ 10% acetonitrile + 0.2% acetic acid) and B (5 mM ammonium acetate in 90%acetonitrile/ 10% H_2O + 0.2% acetic acid). The gradient conditions for both separations were identical and were as follows: 0-2 min, 25% A – 2-5 min, 25% to 70% A, linear gradient – 5-7 min, 70% A – 9-11 min 70% to 25% A, linear gradient – 11-20 min, The chromatographic separation, MS ionization and data acquisition was performed using an AB Sciex QTrap 5500 mass spectrometer equipped with electrospray ionization (ESI) source. The instrument was controlled by Analyst 1.5 software (AB Sciex, Toronto, ON, Canada). Targeted data acquisition was performed in multiple-reaction-monitoring (MRM) mode. We monitored 98 and 59 MRM transitions in negative and positive mode, respectively (157 transitions total).

The extracted MRM peaks were integrated using MultiQuant 2.1 software (AB Sciex, Toronto, ON, Canada).

Targeted HILIC-QTOF mass spectrometry metabolite quantifications of methionine metabolites

Elf1 and H1 cells were grown on matrigel coated 10cm plates for one passage (4 replicates each with 4 plates per H1 replicate and 3 plates per Elf replicate). Cells were washed with PBS, and scraped into fresh PBS at room temperature, pelleted and flash frozen in liquid nitrogen. Samples were extracted by adding 1mL cold 3:1 cold methanol:water (Fisher Scientific, Waltham, MA) to the cell pellet. Samples were vortexed, placed at -20^{oC} for 30 minutes, centrifuged for 10 minutes at 14000 rcf. The supernatant was transferred then centrifuged again, then supernatant was evaporated to dryness. Samples were re-suspended in 80:20 acetonitrile:water (Fisher Scientific) containing the following labeled-internal standard mixture for external quality controls: L-arginine HCl (GUANIDO-15N2) and L-glutamine-2,3,3,4,4-d5 (CDN Isotopes, Quebec, Canada) at 14.8 μg and 0.85 μg respectively. Standard curve dilutions for quantifications were prepared using mixture of 1-methylnicotinamide HCl (1-MNA), S-methyl-5'-thioadenosine (MTA), S-adenosyl methionine (SAM), S-adenosyl homocysteine (SAH), Methionine, Kynurenine and Tryptophan (Sigma Aldrich, St Louis, MO).

Hydrophilic interaction chromatography (HILIC) analysis of standard curve and samples was performed using an Agilent 1290 Infinity Ultrahigh Pressure Liquid Chromatography stack

(Agilent Technologies, Santa Clara, Ca) equipped with an auto-sampler (4°C) using 5 µL injections into an Acquity UPLC BEH Amide column (Waters Corporation, Milford, MA) maintained at 45 °C with 1.7 µm particles and 2.1 mm i.d. x 150mm length. Mobile phases where prepared with 10 mM ammonium formate and 0.125 % formic acid (Sigma Aldrich) in either 100% LCMS grade Water (Fisher Scientific) for mobile phase A or 95:5 acetonitrile:water (Fisher Scientific) for mobile phase B. Gradient elution was performed from 100%B at 0-2 min to 70%B at 7.7 min, 40%B at 9.5 min, 30%B at 10.25 min, 100%B at 12.75min, isocratic until 16.75 min with a column flow of 0.4 mL/min. Metabolites were detected and quantified by an Agilent 6530 accurate mass quadrupole time-of-flight (QTOF) mass spectrometer with a jet stream ESI source in positive ion mode (Agilent). Mass calibration was maintained by constant reference ion infusion, with MS data collected at 4 spectra/s. Data files were analyzed using Agilent Mass Hunter TOF Quantitative Analysis software. Metabolites were quantified at the precursor ion level set at a mass window of 0.02 Da. Peak filtering was performed manually to eliminate peaks with a signal to noise ratio of less than 3. Retention times and major adducts for each compound are as follows: 1-MNA (m/z 137.0715) 6.345 min, MTA (m/z 297.0896) 2.583 min M+H, tryptophan (m/z 204.0899) 6.904 min M+H, kynurenine (m/z 208.0848) 6.971 min M+H & M+Na, methionine (m/z 149.0511) 7.493 min M+H & M+2Na+H, SAH (m/z 384.1216) 8.810 min M+H, SAM (m/z 399.1451) 9.768 min.

Metabolite levels were sum-normalized for each sample using the methionine metabolite values (nicotinamide, MTA, 1-MNA, SAM and SAH). P-values were calculated using a 1-tailed t-test.

Transcriptomic data analysis

The approximately 92 million paired-end reads from RNA-seq (H1 and Elf1 hESCs, SeqWright) were aligned using Gsnap(Wu and Nacu, 2010) to version hg19/GRCh37 of the human genome. Reads aligned to genes defined by the Ensembl 74 annotations(Flicek et al., 2014) were counted and the DESeq R/Bioconductor package(Anders and Huber, 2010) was used to evaluate differential expression.

Microarray data from a Gafni et al. (PMID:24172903, accession number GSE46872) and Valamehr et al. (PMID: 24672758, accession number GSE50868) were downloaded from Gene Expression Omnibus and processed with Robust Multi-array Average(Irizarry et al., 2003). Processed expression data table was downloaded from supplemental Table S1 of Theunissen et al (PMID: 25090446). RNA-seq data was downloaded from Chan et al (PMID: 24315441, accession number E-MTAB-2031). PCA analysis revealed that one 3iL hESC sample was very far away from the other two 3iL hESC samples and was removed from downstream analysis.

Global metabolomic data analysis

Alll global metabolomic data was mean-centered within each sample(Wang et al., 2003). *prcomp* function in R (R Core Team, 2013.) is used for Principle Component Analysis of metabolomics data. Differentially abundant metabolites were defined as metabolites with 2 fold change in abundance and Benjamini-Hochberg adjusted false discovery rate <0.2.

For the lipidomics data, features missing in more than half of all samples (4 or more out of 6) were removed from further analysis. Missing values were replaced with minimum detected values within each sample before mean-centering.

ChIP-seq data analysis

ChIP-seq data of H3K27me3 H3K4me3, H3K9me3 and H3K27ac modifications from Chan et al (PMID: 24315441), Gafni et al (PMID:24172903), Theunissen et al (PMID: 25090446) and Bernstein et al (PMID: 20944595) was downloaded from Array Express (accession number E-MTAB-2041) and GEO (accession numbers GSE52617 and GSE59435). Reads were aligned to hg19 using Bowtie version 1.0.0.(Langmead et al., 2009) allowing 1 mismatch (-N 1). ngsplot was used to generate plots of reads around 5 KB of transcription start sites of *a priori* defined developmental genes. Reads of replicate samples for the same cell type wre merged for ngsplot. Reads with mapping quality above 20 were used by ngsplot. Differentially marked genomic regions were identified with diffReps version 1.55.4(Shen et al., 2013) and annotated to the closest genes. Genes associated with at least one significant genomic region (FDR less than 0.05 and fold change >1) were classifiedas differentially marked. When a gene is annotated with multiple significant genomic regions, the most significant one is assigned to that gene.

Lipid droplet visualization using Oil Red O and bodipy staining

Naïve and primed ESCs were fixed with 4% PFA at room temperature for 10min, washed twice with PBS and stained with Oil Red O dye (Sigma) for 10min at room temperature. Alternatively, lipid droplets were also stained using BODIPY 493/503 (Molecular Probes) for 15min on a rocking platform at room temperature. Pictures were taken using a fluorescent microscope (Leica).

Protein extraction and Western blot analysis

Cellular extracts were prepared using a lysis buffer containing 20 mM Tris HCl (pH 7.5), 150 mM NaCl, 15% glycerol, 1% Triton, 25mM β-glycerolphosphate, 50mM NaF, 10mM NaPyrophosphate, orthovanadate, PMSF (all chemicals are from Sigma-Aldrich, St. Louis, MO), Protease inhibitor cocktail (Roche) and 2% SDS. 25 U of Benzonase® Nuclease (EMD Chemicals, Gibbstown, NJ) and 20mM of DTT (Sigma) were added to the lysis buffer right before use. Protein concentrations were then determined by the method of Bradford. 15 μg of protein extracts were loaded, separated by 4-20% SDS-PAGE, and transferred to polyvinylidene difluoride membranes (Hybond-N+, Amersham Pharmacia Biotech, Buckinghamshire, England). Membranes were blocked with 5% nonfat dry milk for 60 minutes at room temperature, and incubated overnight at 4°C with primary antibody. Finally, after blots had been incubated for one hour with horseradish peroxidase-conjugated secondary antibodies, they were visualized by enhanced chemiluminescence (Millipore Corp, Billerica, MA). Antibodies used in this study are: H3K27me3 (1/1000, Active Motif), H3K9me3 (1/1000, Abcam), H3K9/14Ac (1/1000, Cell Signaling), EED (1/1000, gift from Dr. Bomsztyk), HIF1α (1/2000, BD Biosciences) and γ-tubulin (1/10000, Sigma).

RNA extraction and qPCR analysis

Total RNA was extracted using trizol and subsequently analyzed by SYBRgreen qPCR with the 7300 real time PCR system (Applied Biosystems) as described previously(Sperber et al., 2014) as well as TaqMan qPCR (Applied Biosystems). Primers used are listed in Suppl.Table.2.

qPCR of miRNAs was conducted using TaqMan miRNA assays (Applied Biosystems). Raw Ct values for miRNAs were normalized to RNU66 (endogenous snoRNA, internal control). Linear expression values for all qPCR experiments were calculated using the $2(-\Delta Ct)$ method. P-values were calculated using a student's t-test (*p<0.05, **p<0.01, ***p<0.001)

ChIP-seq analysis

Naïve hESCs Elf1 2iLIF grown on matrigel were treated with 100μM of STAT3 inhibitor (Selleckchem) for 6h or 24h and analyzed for methylation marks by Western blot and ChIP Seq. For ChIP-seq analysis, cells were crosslinked and chromatin processed as previously described(Hawkins et al., 2010) with minor modifications. Briefly, cells were harvested with accutase and crosslinked in suspension with 1% formaldehyde solution for 10min at room temperature. Reaction was quenched with glycine and crosslinked cells were rinsed with ice-cold PBS. Nuclei were isolated and chromatin sonicated using a Covaris E210 to approximately 200-500bp size range. ChIP-seq was conducted as previously described(Hawkins et al., 2010) with minor modifications. Briefly, magnetic Dynabeads were incubated overnight rotating at 4C with antibody against H3K27me3 (Active Motif, cat # 39155). Sonicated chromatin from approximately 200 thousand cells was added to the bead-bound-antibodies and allowed to incubate at 4C rotating overnight. Beads were washed to remove unbound chromatin. Bound chromatin was eluted from beads and reverse crosslinked overnight. Purified DNA was prepared for next-generation sequencing via end repair, A-tailing, ligation of custom Y-adapters and PCR amplification to generate final DNA library following gel size selection.

Generation of BAR-Elf reporter cell line

Elf1 cells grown in naïve media (2iL or 2iLIF) or primed media (AF) were infected with BAR reporter lentivirus(Biechele et al., 2009;Davidson et al., 2012) and seeded onto matrigel-coated plates in MEF-CM with 10μM Y-27632, and 1μM Thiazovivin (ROCK inhibitors, Tocris).

Transduced Elf cells were cultured for a week on matrigel, then passaged onto MEF plates for further selection and expansion. Elf1 naïve reporter cells were harvested as single cells via TrypLE Express and FACS sorted for the population with both Venus and DsRed positive signals. DsRed positive colonies of Elf1 primed reporter cells were manually dissected, transferred onto MEF plates, and the same positive selection was repeated one more round or two depending on the selection efficiency. Negative colonies were manually removed as a negative selection. Once the Ds-Red positive lines were isolated and stable lines established, the primed Elf1 reporter cells were passaged with dispase.

Wnt secretion was inhibited in Elf1 naïve reporter cells by treatment with IWP2 (72h, 2µM, Torcis) or XAV939 (5µM, Sigma). Wnt pathway was activated in primed Elf1 AF reporter cells using a GSK3 inhibitor, CHIR99021 (72h, 10µM, AxonMedChem). Both IWP2 and CHIR99021 were reconstituted in DMSO.

Production of Conditioned Medium (LCM and Wnt3A-CM)

L and L-Wnt3A cells (ATCC) were cultured in 15cm plate in 10% FBS/DMEM media until ~90% confluent. Medium was collected every 48hrs for three batches. Biological activity of secreted Wnt3A in the individual batches of the conditional medium was confirmed in 293T-BAR reporter cells(Davidson et al., 2012), then batches were pooled and filtered. Primed (Elf1 AF) reporter cells were grown on matrigel with 50% LCM or 50% Wnt3A-CM for 3 days prior taking bright field and fluorescent pictures (Leica microscope).

RNA interference experiments

Naïve Elf1 2iLIF cells were transfected in matrigel coated plates in MEF-CM supplemented with ROCK inhibitors (Torcis) using Lipofectamine RNAiMAX (Life Technologies). siRNA targeting NNMT (Hs-NNMT-8) was purchased from Qiagen as Flexitube siRNA premix, and siRNA targeting luciferase was used as control. siRNAs against NNMT and luciferase were used at 50nM final concentration. Proteins and RNAs were extracted 72h after transfection. siRNA targeting beta-catenin (Invitrogen, *CTNNB1*, Silencer Select ID s437) and Silencer Select Negative Control 1 (Invitrogen) were transfected in naïve Elf1 2iLIF cells at 10nM final concentration following a reverse transfection protocol. Bright field and fluorescence images were taken after 3 days. The efficacy of NNMT and beta-catenin siRNAs was confirmed by qPCR analysis.

Detection of Kynurenine in media

The biological activity of IDO was evaluated by measuring the level of tryptophan degradation product, L-kynurenine, present in the medium of primed and naïve hESCs, as previously described (Rezakhanlou et al., 2010). Proteins in the medium were precipitated with 30% trichloroacetic acid. After centrifugation at 8000 g for 5 min, the supernatant was incubated with an equal volume of Ehrlich's reagent (1% 4-dimethylaminobenzaldehyde in glacial acetic acid) at room temperature for 10 min. Optic density was measured at 492 nm, using a Molecular Devices Spectramax Plus 384 microplate reader. The concentration of kynurenine in the medium was calculated according to a standard curve of defined kynurenine concentration (0–100 µM). P-values were calculated using a t-test (*$p<0.05$)

Figure legends

Figure 1: Naïve and primed ESCs are metabolically different

A: Metabolic profile of naïve and primed human pluripotent stem cells (naïve: Elf1 and H1 4iLIF; primed: H1 plated at 25k or 40k/well). A representative trace of OCR changes is shown under a MitoStress protocol (sequential addition of oligomycin, FCCP, and antimycin/rotenone). B: Primed hESCs (H7 and H1) have reduced OCR changes in response to FCCP following oligomycin treatment compared to naïve hESCs (Elf1 and H1 4iLIF). C: Primed hESCs (H7) do not have lower mitochondrial DNA copy number than naïve hESCs (Elf1). D: Primed hESCs (H1) and naïve hESCs (Elf1) have similar mitochondrial mutation frequencies. E: HIF1α is stabilized in hESCs (H7). F: Proteomic workflow used to identify differentially regulated protein expression in prime vs. naive hESCs. G: Volcano plot of differentially expressed proteins in primed hESCs (right, green; Elf1 AF) vs naïve cells (left, blue, Elf1 2iLIF). Significant hits are shown (FDR<0.05). Proteins were quantified by nano-LC-MS/MS on a Fusion Orbitrap. H: log2 fold expression change of all cytochrome C oxidases (COX) genes between primed and naïve states are shown in both our data set (left) and Gafni et al (right). Consistently, COX genes are downregulated in the primed state.

Figure 2: Metabolomic analysis of naïve and primed ESCs

A: Scheme of mass spectrometry experiments performed for metabolites on mouse and human naïve (pre-implantation) and primed (post-implantation) ESCs. B-C: naïve and primed stem cells can be clearly separated based on their metabolic profiles. (B) PCA plot of water-soluble untargeted GC-MS metabolomics data. The first principal component (PC), which separates the primed cell types (left) from the naïve cell types (right) explained 50.5% of total variance. (C) PCA plot of untargeted LC metabolomics data. 3 clusters are along the first PC: primed cells (left), primed cells toggled back to naïve cells (middle) and naïve cells (right). The first PC explained 68.2% of total variance. D-G: volcano plots of differentially abundant metabolites between primed and naïve cells in mESCs (D, E) and hESCs (F-G). x-axis is log2 fold change of abundance, y-axis is negative log10 of p-value. Metabolites of biological interest for further analysis are labeled. H-I: Fold change of glycolysis metabolites in H1 vs. Elf1 detected by targeted LC-QQQ-MS.

Figure 3: Primed ESCs accumulate lipids while naïve ESCs use fatty acids as a source of energy A-B: More abundant lipids in primed cells (H1) have more carbon atoms (A) and larger mass (B) than more abundant lipids in naïve cells (Elf1). C: More abundant lipids in primed cells (R1AF) are more unsaturated than more abundant lipids in naïve cells (R1). D: Oil Red O staining shows an increase of lipid droplet accumulation in primed hESCs (H7) compared to naïve hESCs (Elf1). E: BODIPY 493/503 staining shows an increase of lipid droplet accumulation in primed hESCs (H7, H1, Elf1 AF) and mESCs (EpiSC) compared to naïve hESCs (Elf1, H1 2iF, H1 4iLIF) and mESCs (R1). F-G: Relative fold change of expression of genes involved in transport of fatty acids in mitochondria and genes involved in 4 steps of fatty acid beta-oxidation (RNASeq data, E: H1 vs. Elf1, F: mouse Epiblasts vs. ICM, Zhou et al.). Only genes with significant change (FDR<0.05, fold change >1) or genes expressed (10% of most highly expressed genes within the pathway) are plotted. A full plot of all genes is shown in Supplementary Figure 5. H: ChIP-seq analysis of CPT1A gene shows more repressive H3K27me3 marks and less active H3K4me3 and H3K27ac marks in primed hESCs (C1,

WIBR3(Gafni et al., 2013), H1, H9(Bernstein et al., 2010)) than naïve hESCs (Elf1(Ware et al., 2014); naïve C1, naïve BGO1, naïve WIBR3(Gafni et al., 2013)). I: volcano plot representation of microRNAs expression in naïve hESCs (Elf1) and primed hESCs (H1, ENCODE). J: qPCR expression of microRNA miR-9 (predicted to target CPT1A) and miR-10a (predicted to target FASN) showing a 34-fold higher expression of hsa-miR-10a in Elf1 compared to H1, and 4-fold higher expression of hsa-miR-9 in H1 compared to Elf1. K-M: Seahorse palmitate assay shows that naïve human and mouse ESCs use fatty acids as a source of energy. A representative trace of OCR changes is shown after addition of 2 doses of palmitate or BSA vehicle followed by 2 doses of ETO in human ESCs (naïve Elf1 and primed H7, K) and mouse ESCs (naïve R1 and primed EpiSCs, L). Changes after ETO injections were quantified in M.

Figure 4: Amino acids methionine and tryptophan are differentially regulated in naïve and primed hESCs

A: Model of SAM pathway and NNMT. Metabolites in red are up-regulated in primed hESCs while metabolites and enzymes in blue are up-regulated in naïve hESCs. B: Model of Tryptophan-Kynurenine pathway. C: IDO1 is highly expressed in primed hESCs (H1, H7, Elf1 AF) as compared to naïve hESCs (Elf1 2iLIF, H1 2iF, H1 4iLIF), as shown by qPCR analysis. D: The kynurenine vs. tryptophan ratio is higher in primed hESCs (H1) than naïve hESCs (Elf1), as detected in three independent targeted mass spectrometry experiments and one non targeted mass spectrometry experiment. $p<0.05$ E: Kynurenine is secreted in the media by primed hESCs (H7). F: Volcano plot of RNA-seq data from naïve hESCs (Elf1) and primed hESCs (H1). Genes with greater than 2 fold change and false discovery rate <0.05 are colored. NNMT and IDO1 are among the most differentially expressed genes. G: NNMT is highly up-regulated in naïve hESCs

(Elf1 2iLIF, H1 2iF, H1 4iLIF) compared to primed hESCs (H1, H7, Elf1 AF), as shown with qPCR analysis. H: Naïve hESCs (Elf1) have higher amounts of the NNMT product, 1MNA, than primed hESCs (H1). I: SAM levels are higher in primed hESCs (H1) than in naïve hESCs (Elf1), while SAH levels are lower in naïve hESCs (Elf1) than primed hESCs (H1).

Figure 5: High NNMT expression in naïve hESCs regulates histone methylation status

A-B: H3K27me3 reads mapped 5kb around transcription start sites (TSS) of 648 developmental genes were plotted for Ware et al., Gafni et al., Theunissen et al., Bernstein et al., 2010 (A), as well as Chan et al. (B) ChIP-seq data sets. All showed that primed cells have more H3K27me3 repressive marks around TSS. C: Western blot analysis show higher H3K27me3 and H3K9me3 in primed hESCs (H7) than naïve hESCs (Elf1) while H3K9/K14Ac does not change. D: qPCR analysis shows a knock-down regulation of NNMT using siRNA (50 nM, 72h) in naïve hESCs (Elf1). E: Western blot analysis of histone marks in Elf1 cells treated siRNA against NNMT or siRNA against luciferase as a control. F: Western blot analysis of histone modifications after treatment of Elf1 cells with 100 μM of STAT3 inhibitor. G: 6h treatment with STAT3 inhibitor (100 μM) in Elf1 cells increases H3K27me3 marks, as shown by ChipSeq analysis on all genes. H: Windowed chromatin heatmaps of H3K27me3 profile +/- 5kb of promoters of 313 overlapping genes with increased H3K27me3 mark in primed vs. naïve hESCs (Gafni et al., Theunissen et al., Bernstein et al.), and Elf1 cells treated for 6h with 100uM STAT3 inhibitor vs. Elf1 cells. I-J: Wnt ligands and Wnt targets are up-regulated in naïve (Elf1) compared to primed hESCs (H1), as detected by RNA Seq in our data set (I) and Chan et al. (J). Genes names are indicated in Suppl.Fig.13A K: Wnt is activated in naïve hESCs. Endogenous Wnt signaling in naïve (Elf1 2iLIF) and primed (Elf1 AF) BAR-reporter cells. Scale bars represent 200μm. L:

siRNA against β-catenin, Wnt inhibitor IWP2 (2μM) and Wnt antagonist XAV939 (5μM) inhibit the reporter activity in naïve Elf1 2iLIF cells after 72h. Scale bars represent 200μm. M: screen shot of RNA expression and H3K27me3 marks of EGLN1 (PHD2) in naïve hESCs [Elf1 (Ware et al.), WIRB3 naïve and BGO1 naïve (Gafni et al.)], primed hESCs [WIRB3 primed (Gafni et al.), H1 and H9 (Bernstein et al.) and Elf1 treated with STAT3 inhbitor (100 μM) for 6h. N: Model of the intricate relationship between metabolism and epigenetic in hESCs. Primed hESCs and naïve hESCs with STAT3 inhibited have downregulated NNMT expression compared to naïve hESCs, resulting in increase levels of SAM and induction of H3K27me3 repressive marks on genes involved in metabolism switch.

Figure 1

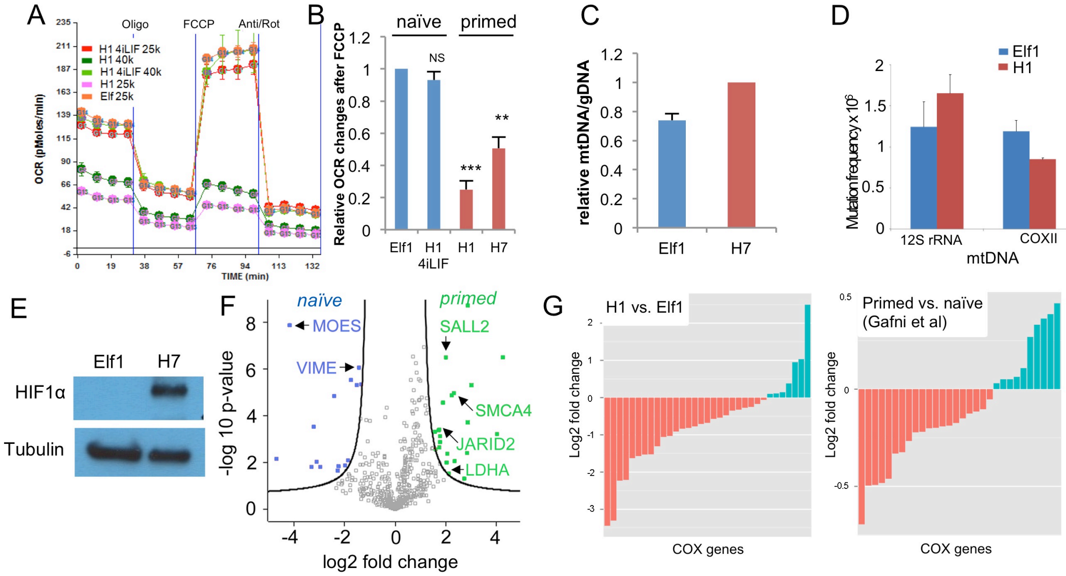

A
Oligo
FCCP
Anti/Rot
H1 4iLIF 25k
H1 40k
H1 4iLIF 40k
H1 25k
Elf 25k
OCR (pMoles/min)
235
211
187
163
139
115
90
66
42
18
-6
0
19
38
57
76
94
113
132
TIME (min)
B
naïve
primed
Relative OCR changes after FCCP
1.2
1
0.8
0.6
0.4
0.2
0
NS
***
**
Elf1
H1
H1
H7
4iLIF
C
relative mtDNA/gDNA
1.2
1
0.8
0.6
0.4
0.2
0
Elf1
H7
D
Elf1
H1
Mutation frequency x 10^6
2
1.5
1
0.5
0
12S rRNA
COXII
mtDNA
E
Elf1
H7
HIF1α
Tubulin
F
naïve
primed
MOES
VIME
SALL2
SMCA4
JARID2
LDHA
-log 10 p-value
8
6
4
2
0
-4
-2
0
2
4
log2 fold change
G
H1 vs. Elf1
Log2 fold change
2
1
0
-1
-2
-3
COX genes
Primed vs. naïve
(Gafni et al)
Log2 fold change
0.5
0
-0.5
COX genes

Figure 2

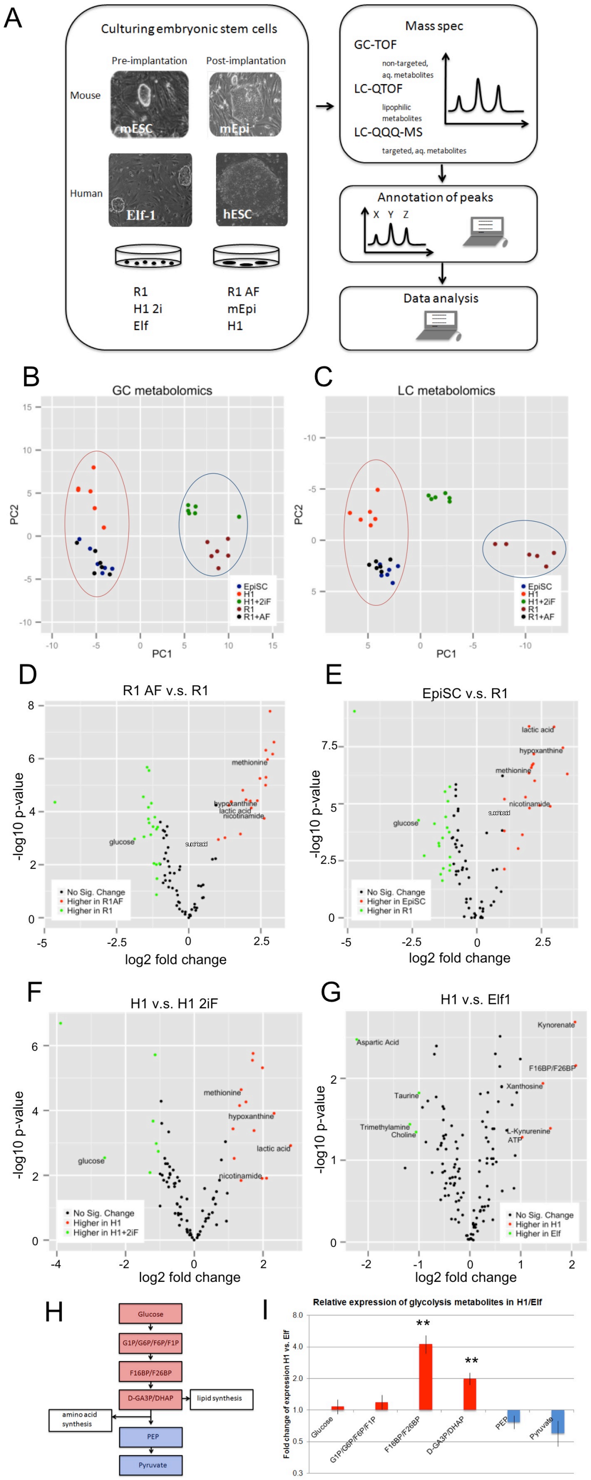

A
Culturing embryonic stem cells
Pre-implantation
Post-implantation
Mouse
mESC
mEpi
Human
Elf-1
hESC
R1
H1 2i
Elf
R1 AF
mEpi
H1
Mass spec
GC-TOF
non-targeted, aq. metabolites
LC-QTOF
lipophilic metabolites
LC-QQQ-MS
targeted, aq. metabolites
Annotation of peaks
X Y Z
Data analysis
B
GC metabolomics
PC1
PC2
EpiSC
H1
H1+2iF
R1
R1+AF
C
LC metabolomics
PC1
PC2
EpiSC
H1
H1+2iF
R1
R1+AF
D
R1 AF v.s. R1
-log10 p-value
log2 fold change
methionine
hypoxanthine
lactic acid
nicotinamide
glucose
No Sig. Change
Higher in R1AF
Higher in R1
E
EpiSC v.s. R1
-log10 p-value
log2 fold change
lactic acid
hypoxanthine
methionine
nicotinamide
glucose
No Sig. Change
Higher in EpiSC
Higher in R1
F
H1 v.s. H1 2iF
-log10 p-value
log2 fold change
methionine
hypoxanthine
lactic acid
nicotinamide
glucose
No Sig. Change
Higher in H1
Higher in H1+2iF
G
H1 v.s. Elf1
-log10 p-value
log2 fold change
Kynorenate
Aspartic Acid
F16BP/F26BP
Xanthosine
Taurine
Trimethylamine
Choline
L-Kynurenine
ATP
No Sig. Change
Higher in H1
Higher in Elf
H
Glucose
G1P/G6P/F6P/F1P
F16BP/F26BP
D-GA3P/DHAP
lipid synthesis
amino acid synthesis
PEP
Pyruvate
I
Relative expression of glycolysis metabolites in H1/Elf
Fold change of expression H1 vs. Elf
8.0
4.0
2.0
1.0
0.5
0.3
**
**
Glucose
G1P/G6P/F6P/F1P
F16BP/F26BP
D-GA3P/DHAP
PEP
Pyruvate

Figure 3

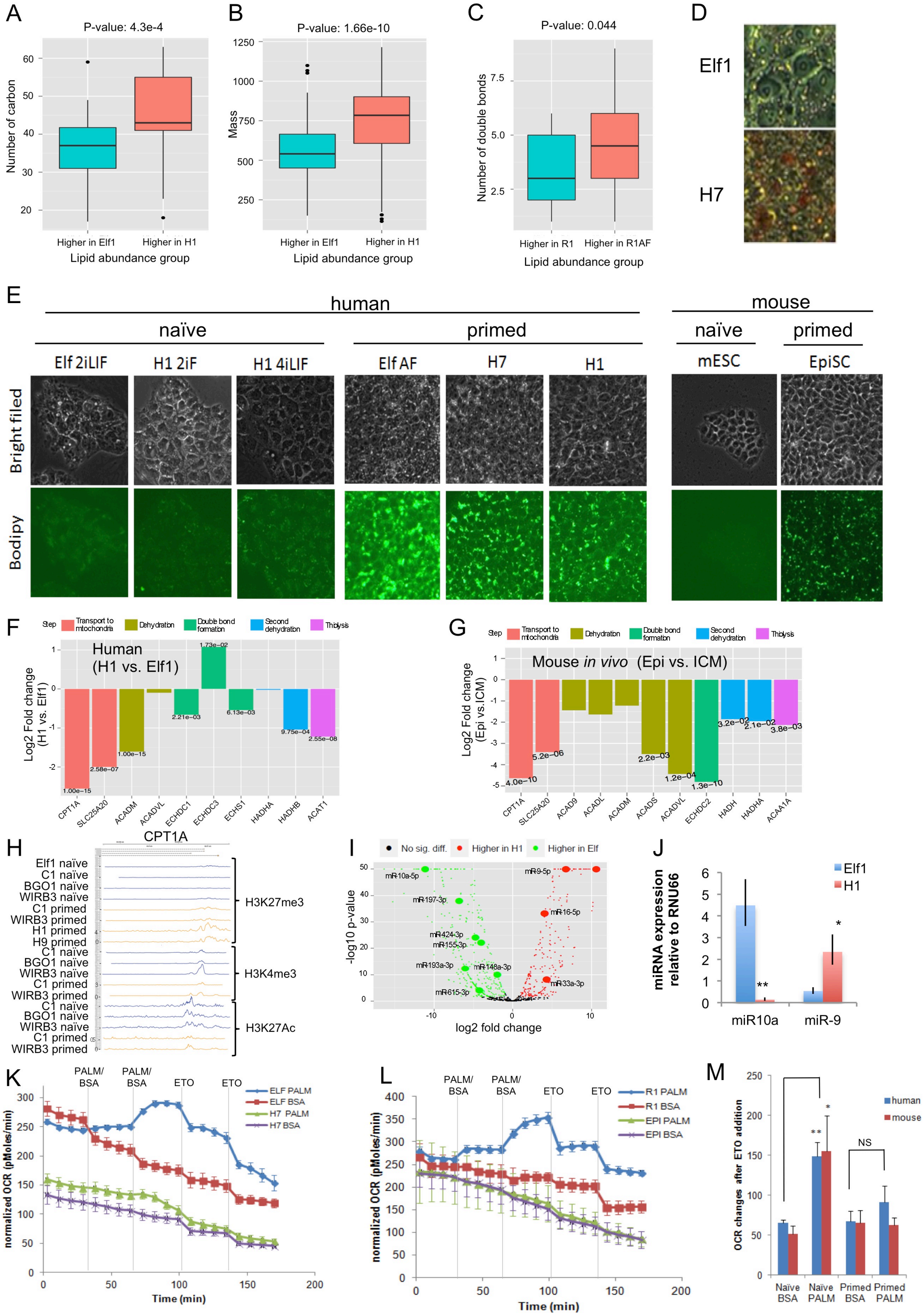

A
P-value: 4.3e-4
Number of carbon
Higher in Elf1
Higher in H1
Lipid abundance group
B
P-value: 1.66e-10
Mass
Higher in Elf1
Higher in H1
Lipid abundance group
C
P-value: 0.044
Number of double bonds
Higher in R1
Higher in R1AF
Lipid abundance group
D
Elf1
H7
E
human
mouse
naïve
primed
naïve
primed
Elf 2iLIF
H1 2iF
H1 4iLIF
Elf AF
H7
H1
mESC
EpiSC
Bright filed
Bodipy
F
Step
Transport to mitochondria
Dehydration
Double bond formation
Second dehydration
Thiolysis
Human
(H1 vs. Elf1)
Log2 Fold change (H1 vs. Elf1)
CPT1A
SLC25A20
ACADM
ACADVL
ECHDC1
ECHDC3
ECHS1
HADHA
HADHB
ACAT1
G
Mouse in vivo (Epi vs. ICM)
Log2 Fold change (Epi vs.ICM)
CPT1A
SLC25A20
ACAD9
ACADL
ACADM
ACADS
ACADVL
ECHDC2
HADH
HADHA
ACAA1A
H
CPT1A
Elf1 naïve
C1 naïve
BGO1 naïve
WIRB3 naïve
C1 primed
WIRB3 primed
H1 primed
H9 primed
H3K27me3
H3K4me3
H3K27Ac
I
No sig. diff.
Higher in H1
Higher in Elf
-log10 p-value
log2 fold change
miR-10a-5p
miR-197-3p
miR-424-3p
miR-155-3p
miR-193a-3p
miR-148a-3p
miR-615-3p
miR-9-5p
miR-16-5p
miR-33a-3p
J
Elf1
H1
miRNA expression relative to RNU66
miR10a
miR-9
K
PALM/ BSA
ETO
ELF PALM
ELF BSA
H7 PALM
H7 BSA
normalized OCR (pMoles/min)
Time (min)
L
R1 PALM
R1 BSA
EPI PALM
EPI BSA
M
human
mouse
OCR changes after ETO addition
NS
Naive BSA
Naive PALM
Primed BSA
Primed PALM

Figure 4

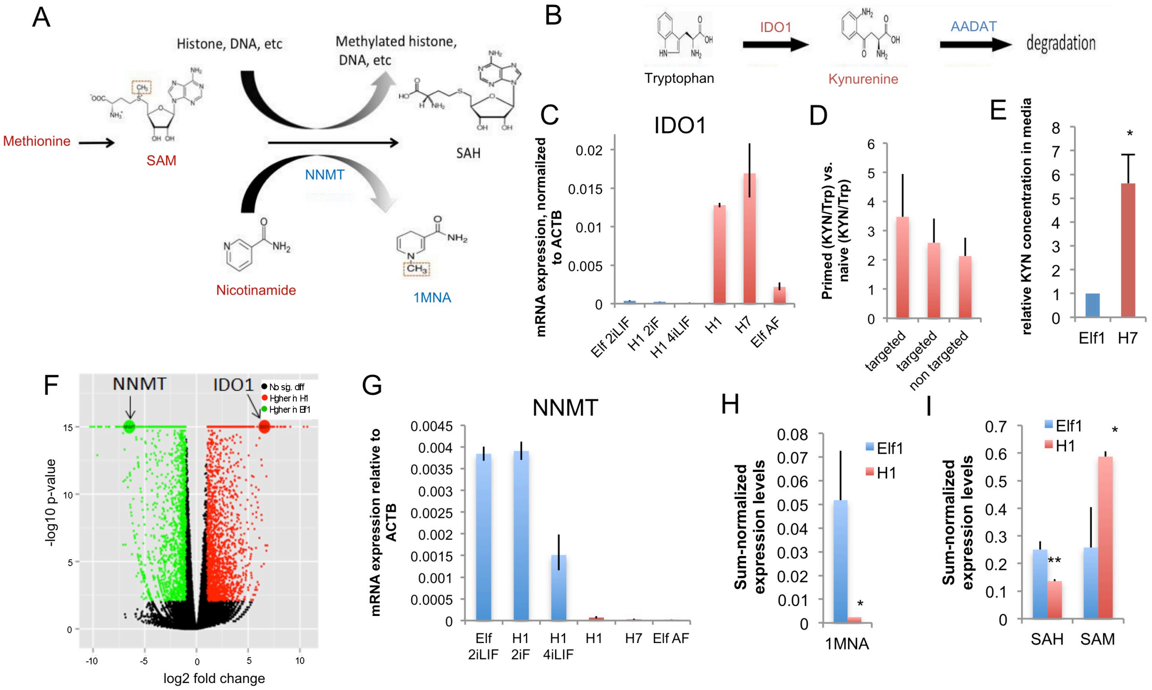
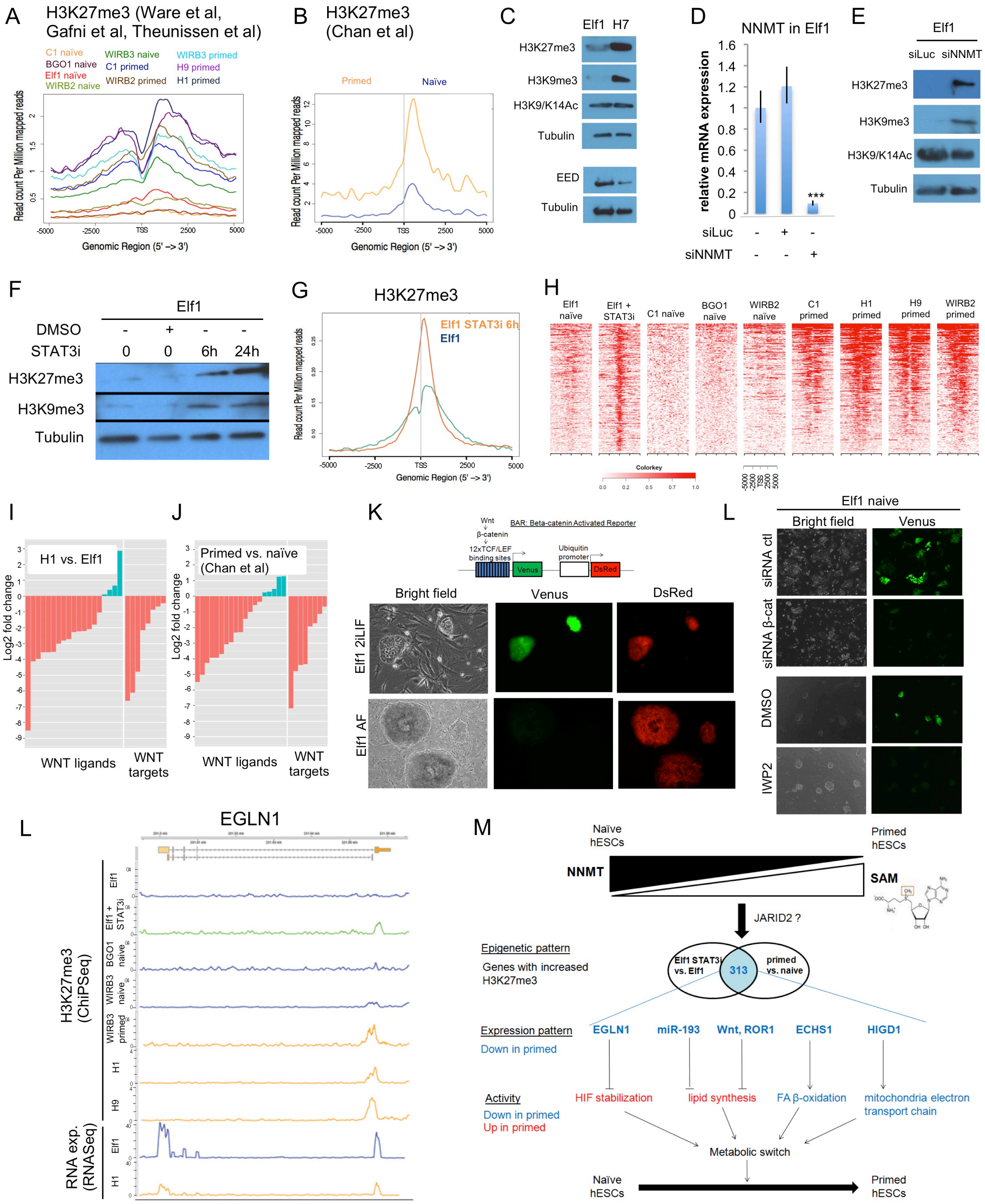

A
Methionine
SAM
Histone, DNA, etc
Methylated histone, DNA, etc
SAH
NNMT
Nicotinamide
1MNA
B
Tryptophan
IDO1
Kynurenine
AADAT
degradation
C
IDO1
mRNA expression, normalized to ACTB
0.02
0.015
0.01
0.005
0
Elf 2iLIF
H1 2iF
H1 4iLIF
H1
H7
Elf AF
D
Primed (KYN/Trp) vs. naive (KYN/Trp)
6
5
4
3
2
1
0
targeted
targeted
non targeted
E
relative KYN concentration in media
8
7
6
5
4
3
2
1
0
*
Elf1
H7
F
NNMT
IDO1
No sig. diff
Higher in H1
Higher in Elf1
-log10 p-value
15
10
5
0
-10
-5
0
5
10
log2 fold change
G
NNMT
mRNA expression relative to ACTB
0.0045
0.004
0.0035
0.003
0.0025
0.002
0.0015
0.001
0.0005
0
Elf 2iLIF
H1 2iF
H1 4iLIF
H1
H7
Elf AF
H
Sum-normalized expression levels
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
Elf1
H1
*
1MNA
I
Sum-normalized expression levels
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Elf1
H1
**
*
SAH
SAM

Figure 5


A
H3K27me3 (Ware et al, Gafni et al, Theunissen et al)
C1 naïve
BGO1 naive
Elf1 naïve
WIRB2 naive
WIRB3 naive
C1 primed
WIRB2 primed
WIRB3 primed
H9 primed
H1 primed
Read count Per Million mapped reads
Genomic Region (5' -> 3')
B
H3K27me3 (Chan et al)
Primed
Naïve
C
Elf1 H7
H3K27me3
H3K9me3
H3K9/K14Ac
Tubulin
EED
Tubulin
D
NNMT in Elf1
relative mRNA expression
siLuc
siNNMT
***
E
Elf1
siLuc siNNMT
H3K27me3
H3K9me3
H3K9/K14Ac
Tubulin
F
Elf1
DMSO
STAT3i
0 0 6h 24h
H3K27me3
H3K9me3
Tubulin
G
H3K27me3
Elf1 STAT3i 6h
Elf1
H
Elf1 naïve
Elf1 + STAT3i
C1 naïve
BGO1 naïve
WIRB2 naïve
C1 primed
H1 primed
H9 primed
WIRB2 primed
Colorkey
I
H1 vs. Elf1
Log2 fold change
WNT ligands
WNT targets
J
Primed vs. naïve (Chan et al)
Log2 fold change
WNT ligands
WNT targets
K
Wnt
β-catenin
12xTCF/LEF binding sites
Venus
Ubiquitin promoter
DsRed
BAR: Beta-catenin Activated Reporter
Bright field
Venus
DsRed
Elf1 2iLIF
Elf1 AF
L
Elf1 naive
Bright field
Venus
siRNA ctl
siRNA β-cat
DMSO
IWP2
L
EGLN1
H3K27me3 (ChIPSeq)
Elf1
Elf1 + STAT3i
BGO1 naive
WIRB3 naive
WIRB3 primed
H1
H9
RNA exp. (RNASeq)
Elf1
H1
M
Naïve hESCs
Primed hESCs
NNMT
SAM
JARID2 ?
Epigenetic pattern
Genes with increased H3K27me3
Elf1 STAT3i vs. Elf1
313
primed vs. naive
Expression pattern
Down in primed
EGLN1
miR-193
Wnt, ROR1
ECHS1
HIGD1
Activity
Down in primed
Up in primed
HIF stabilization
lipid synthesis
FA β-oxidation
mitochondria electron transport chain
Metabolic switch
Naïve hESCs
Primed hESCs

Supplemental figure legends

Supplemental figure 1

A: Representative trace of OCR changes in primed hESCs (H7) and naïve hESCs (Elf1). B-C: Metabolic profile of primed hESCs (H1), hESCs toggled toward a more naïve state (H1 2iF) and naïve mouse embryonic stem cells (R1). A representative trace of OCR changes is shown (A). Primed hESCs (H1) have reduced OCR changes in response to FCCP following glucose treatment compared to naïve human and mouse ESCs (H1 2iF and R1, respectively; B). D:Mitochondrial DNA copy numbers quantified in naïve hESCs (Elf1) and primed hESCs (H1). E: Mitochondrial DNA deletion frequency in naïve hESCs (Elf1) and primed hESCs (H1). F: COX genes are downregulated in the primed state in Theunissen et al.

Supplemental figure 2

Label-free quantification of protein expression is reproducible. Tryptic digestions of naïve hESCs (Elf1 2iLIF) or primed hESCs (Elf1 AF) were analyzed in triplicate by nano-LC MS/MS, with an average Pearson correlation of 0.86.

Supplemental figure 3

Heatmap generated through hierarchical Pearson clustering of metabolite expression from GC-TOF shows that naïve ESCs cluster away from primed ESCs, regardless of origin (mouse or human).

Supplemental figure 4

Relative fold change of expression of genes involved in transport of fatty acids into mitochondria and genes involved in 4 steps of fatty acid beta-oxidation from RNASeq analysis in human ESCs (H1 vs. Elf1, A) and mouse ESCs (Epiblasts vs.ICM, B).

Supplemental figure 5

A: H3K27me3 reads mapped 5kb around transcription start site (TSS) of CPT1A were plotted for Chan et al ChIP-seq data sets. Primed cells have more H3K27me3 repressive marks around TSS of CPT1A. B: Representative trace of OCR changes under Seahorse palmitate assay with addition of 2 doses of palmitate or BSA vehicule followed by 2 doses of ETO in hESCs H1 pushed toward a more naïve state using 4iLIF (H1 4iLIF) and primed hESCs H1.

Supplemental figure 6

Expression of key genes involved in fatty acid synthesis from RNA-seq analysis in naïve (Elf1) and primed (H1) hESCs.

Supplemental figure 7

A: RNA expression of IDO1 in human 8-cell embryo and primed hESCs at passage 0 (hESCp0) and 10 (hESCp10) analyzed by single cell RNA Seq data (Yan et al., 2013). B: RNA expression of IDO1, IDO2, TDO2 and AADAT in naïve hESCs Elf1, primed hESCs H1 and H1 differentiated during 4 days (Stadler et al., 2010) detected by microarray. C: IDO expression in H1 cells and H1 differentiated toward various lineages (Chadwick, 2012).

Supplemental figure 8

A: Fold change of methionine and nicotinamide levels observed between primed hESCs (H1) and naïve or pushed toward naïve hESCs (Elf1 and H1 2i), as detected by GC-TOF (H1 vs. H1 2i) and LC-QQQ-MS (H1 vs. Elf1). B: Expression level of NNMT is consistently observed to be higher in naïve cells than in primed cells by 3 independent groups (this study, Chan et al., Gafni et al.) with different gene expression measurement technologies

Supplemental figure 9

A: NNMT expression in various tissues analyzed by RNASeq (Yu et al., 2014). B: NNMT expression in various organs of rats over time (from 2 to 104 weeks) analyzed by RNASeq (Yu et al., 2014). C: NNMT expression in H1 cells and H1 differentiated toward various lineages (Chadwick, 2012).

Supplemental figure 10

A-D: Reads for H3K4me1 (A), H3K4me3 (B), H3K9me3 (C) and H3K27ac (D) mapped 5kb around transcription start sites (TSS) were plotted for Gafni et al. ChIP-seq data set. E-F: Reads for H3K4me3 (E) and H3K27ac (F) mapped 5kb around TSS were plotted for Chan et al ChIP-seq data set. G: Reads for H3K4me3 mapped 5kb around TSS were plotted for Theunissen et al. ChIP-seq data set.

Supplemental figure 11

A-B: RNA-seq data of histone methyltransferases (A) and histone demethylases (B) in naïve (Elf1) and primed (H1) hESCs.

Supplemental figure 12

A: qPCR analysis of NNMT expression after treatment of Elf1 cells with 100uM of STAT3 inhibitor. B: H3K27me3 reads from ChIP-seq data mapped 5kb around transcription start sites (TSS) were plotted for naïve hESCs (C1, WIBR3, BGO1 from Gafni et al., and Elf1 from Ware et al.), primed hESCs (C1, WIBR3 from Gafni et al., H1 from Ware et al.), and naïve hESCs Elf1 treated for 6h with 100µM of STAT3 inhibitor. STAT3 inhibitor treatment increases H3K27me3 on 313 overlapping genes between primed vs. naïve hESCs (Gafni et al.) and STAT3i vs. Elf1.

Supplemental figure 13

A: Gene names of Wnt ligands and Wnt targets presented in Fig.5I (H1 vs. Elf1) and Fig.5J (primed vs. naïve, Chan et al.). B-C: Wnt ligands are up-regulated in naïve hESCs compared to primed hESCs, as detected by microarray in Gafni et al. (B) and Theunissen et al. (C). D-E: Wnt ligands are up-regulated in naïve hESCs compared to primed hESCs, as detected by microarray in Gafni et al (D) and Theunissen et al. (E).

Supplemental figure 14

A: BAR reporter is activated in naïve hESCs (Elf1 2iL). B: treatment of primed hESCs (Elf1 AF) with Wnt3a CM or GSK3 inhibitor (CHIR99021, 10µM) for 3 days induces differentiation and re-activation of the BAR reporter.

Supplemental figure 15

A: qPCR analysis of miR-200b in naïve hESCs (Elf1) and primed hESCs (H1). B: details of the model presented in Fig.5N showing some of the genes with H3K27me3 mark found in the overlap between primed hESCs and naïve hESCs treated with STAT3 inhibitor, and their possible involvement in the metabolic switch occurring between naïve and primed hESCs. Shown are miRNAs predicted to regulate those genes and that have a change in expression inversely correlating with the expression of those genes.

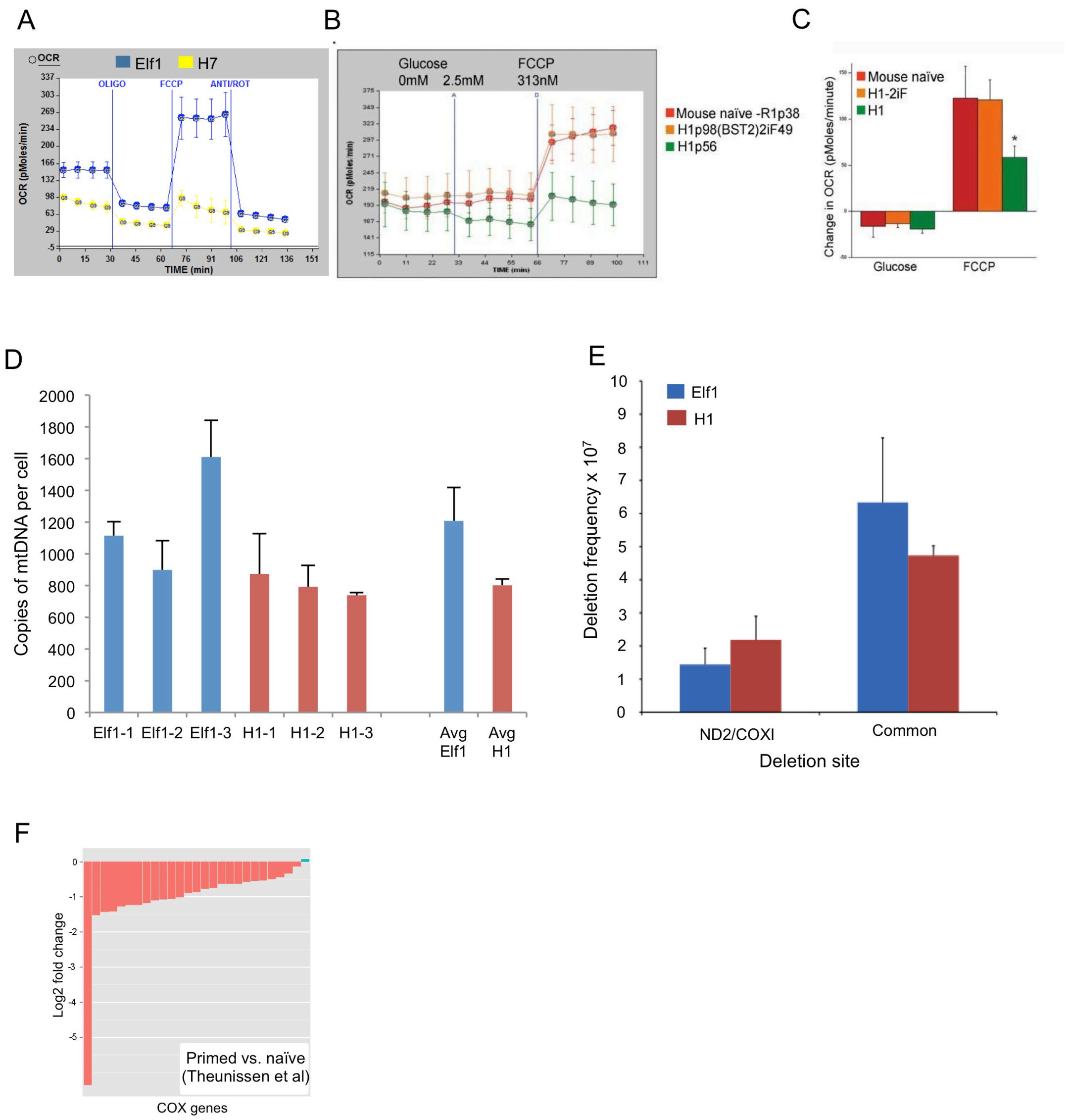
Supplemental table 1

A: mRNA-seq data of H1 and Elf1 human embryonic stem cells. B: Proteomics mass spec data of H1 and Elf1 human embryonic stem cells. C: Mass spec data from non-targeted GC-TOF analysis of mouse R1, epi and R1 AF ESCs, and human H1 and H1 2i ESCs. D: Mass spec data from non-targeted LC-QTOF analysis of mouse R1, epiSC and R1 AF ESCs, and human H1 and H1 2i ESCs. EF: Mass spec data from targeted LC-QQQ-MS of H1 and Elf1 hESCs. G: Mass spec data from lipidomic analysis (LC-QTOF) of H1 and Elf1 human embryonic stem cells. H: miRNA-seq data of H1 and Elf1 human embryonic stem cells. I: Targeted LC-QTOF mass spec analysis of metabolites in methionine and tryptophan metabolism in H1 and Elf1 human embryonic stem cells. J: Top GO-terms for genes that are upregulated in Elf1 human embryonic stem cells after 6h STAT3 inhibitor treatment. K: Top GO-terms for genes that are downregulated in Elf1 human embryonic stem cells after 6h STAT3 inhibitor treatment. L: List of 313 genes that make up the intersection between genes that go down in Elf1 human embryonic stem cells after 6h STAT3 inhibitor treatment, and genes that have increased H3K27me3 marks in primed H1 compared to naïve Elf1 human embryonic stem cells.

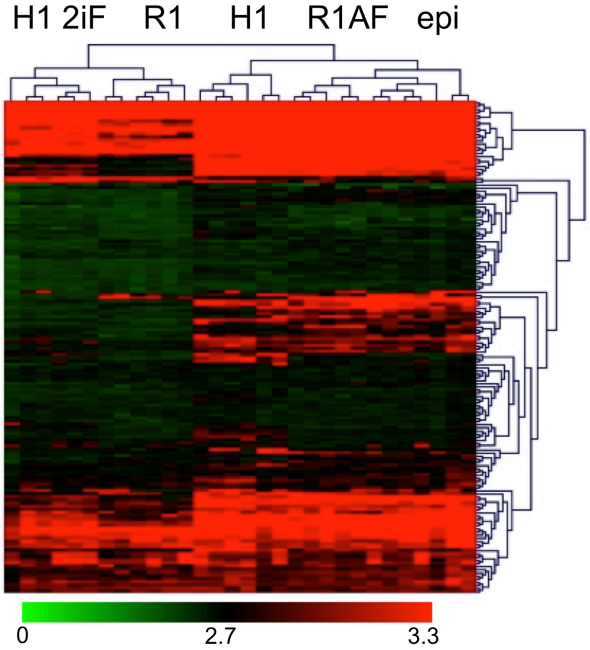
Supplemental table 2

List of primers and Taqman probes used for amplification and quantification of genomic DNA, mitochondrial DNA and cDNA.

Supplemental figure 1

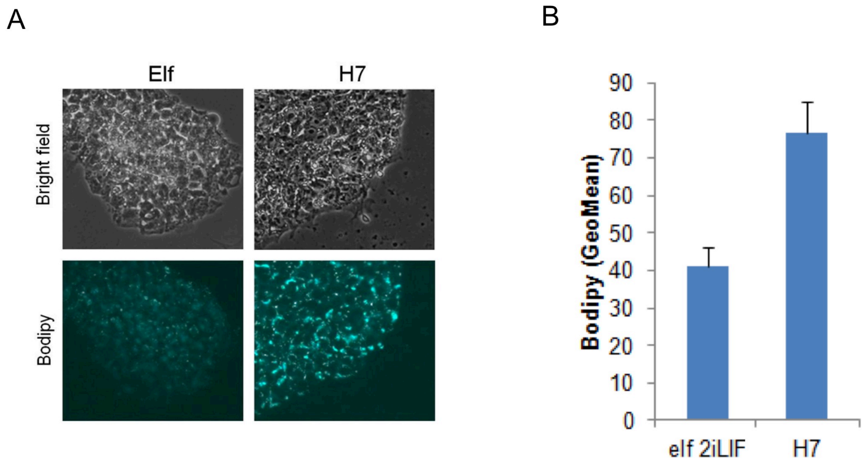


Supplemental figure 2

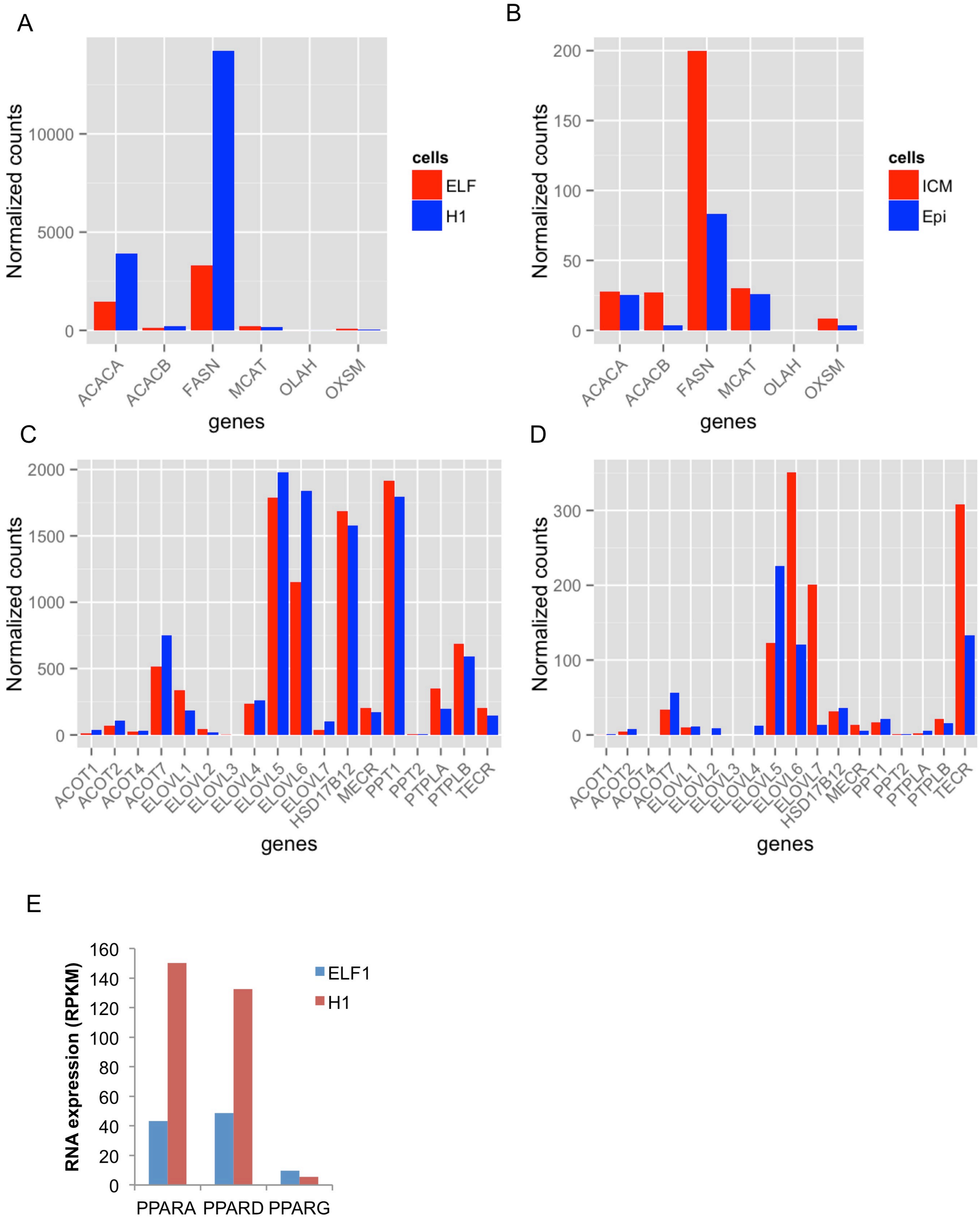


Heatmap of metabolite expression from GC-TOF

Supplemental figure 3

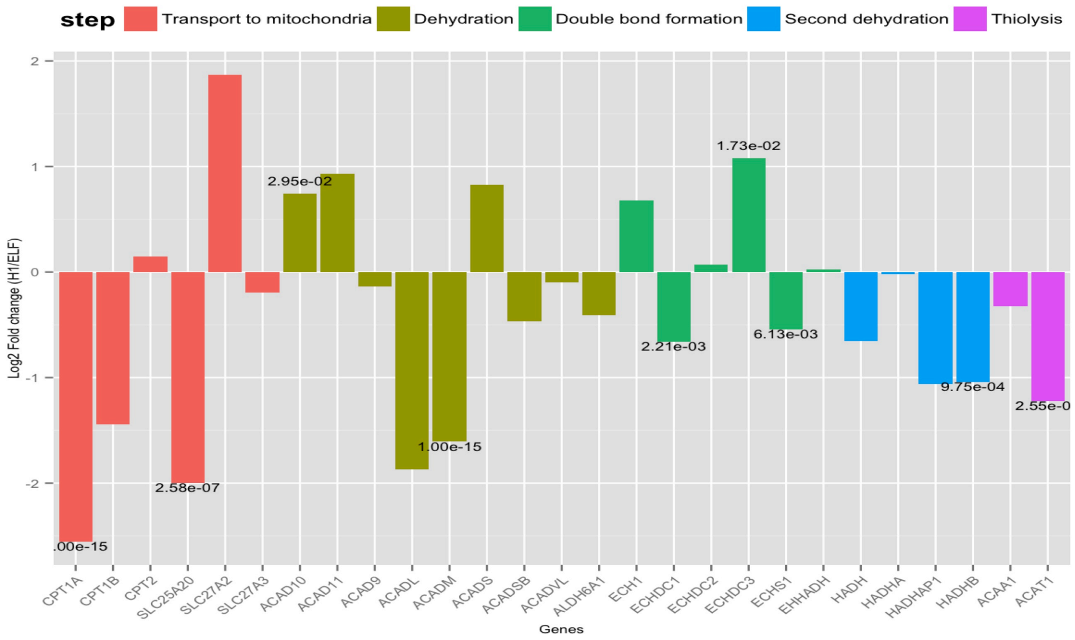


Supplemental figure 4

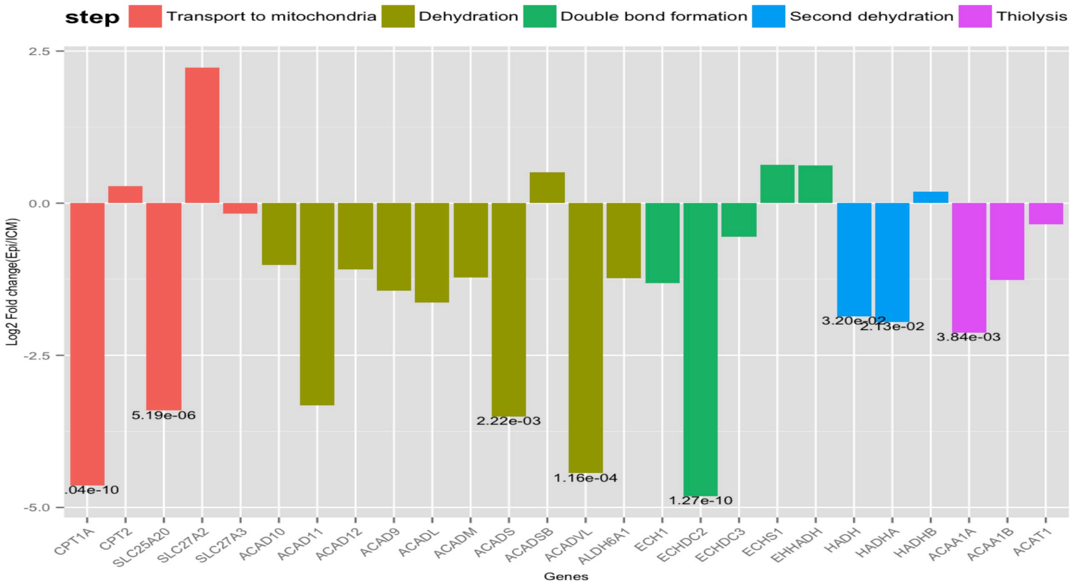


Supplemental figure 5

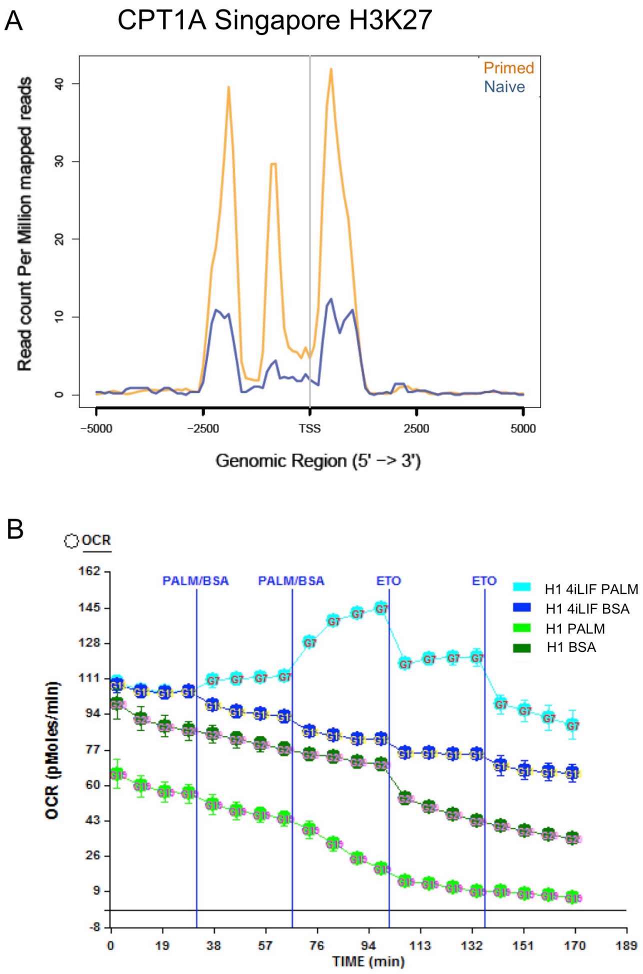
A Genes involved in fatty acid transport and beta oxidation (human: H1 vs. Elf1)



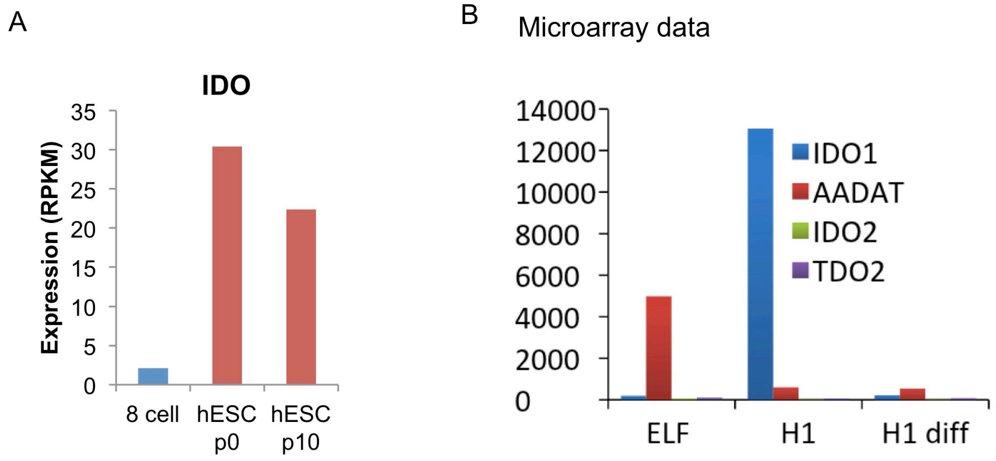
B Genes involved in fatty acid transport and beta oxidation (mouse: Epi vs. ICM)



Supplemental figure 6

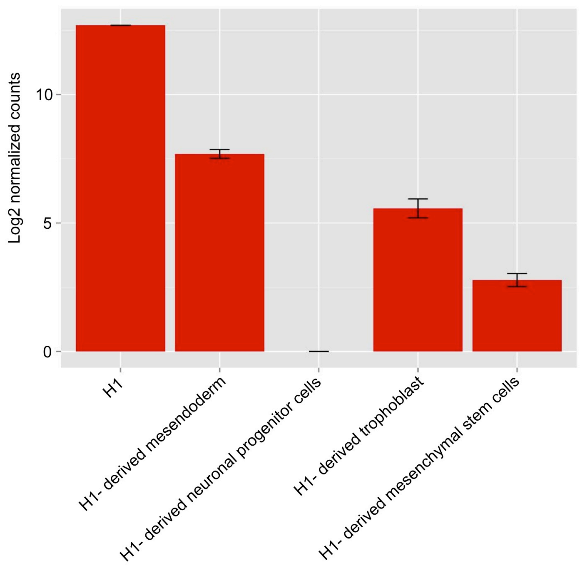


Supplemental figure 7



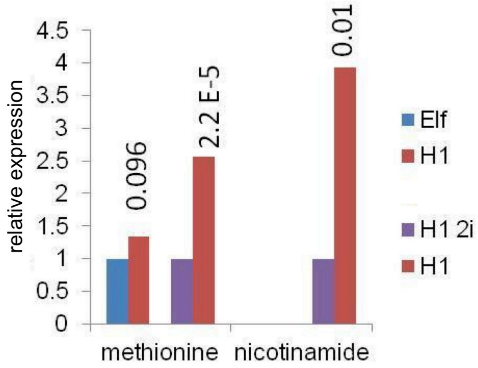
Microarray data

C IDO1 expression in various lineages



Supplemental figure 8

A



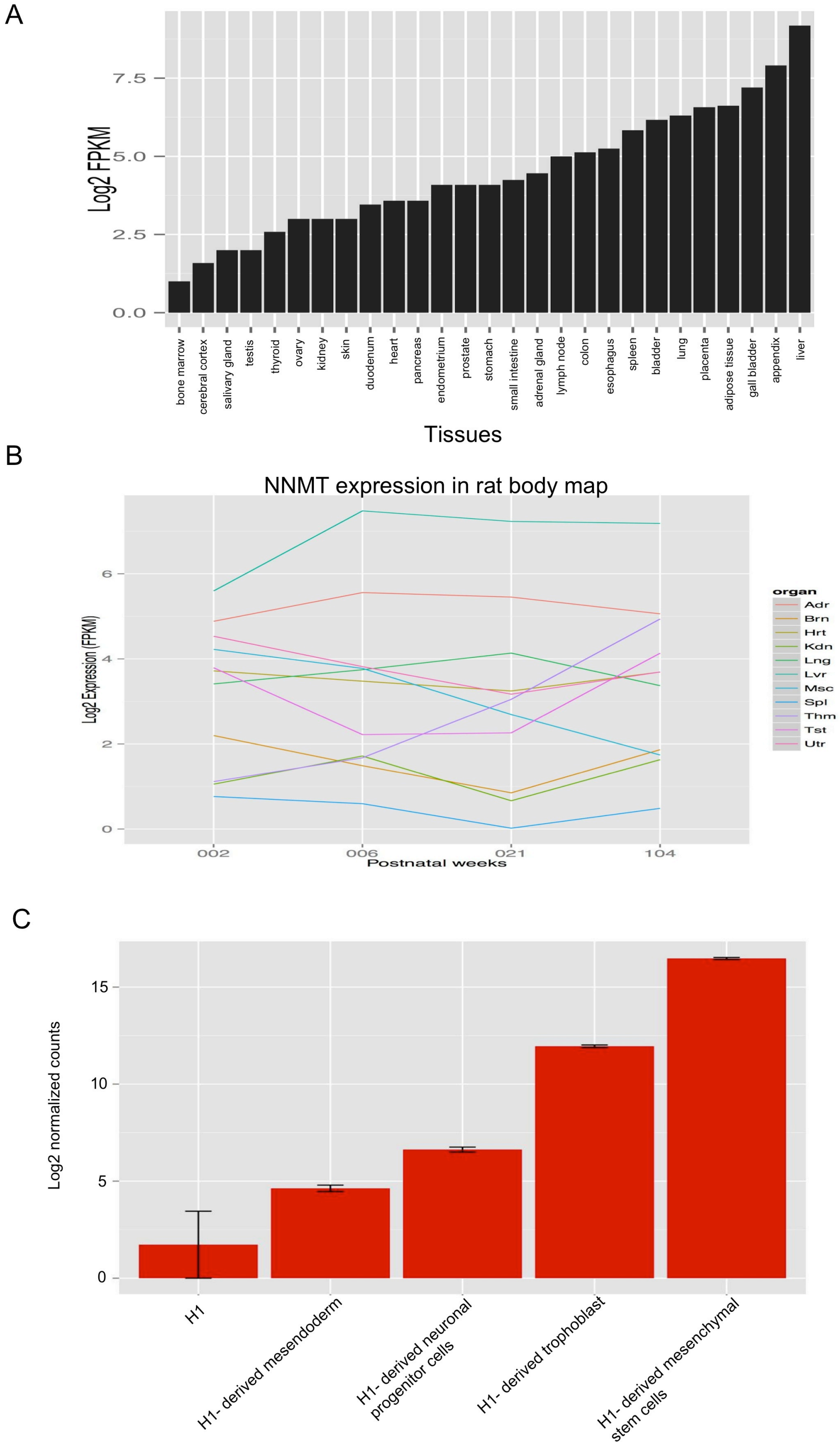
B

NNMT expression

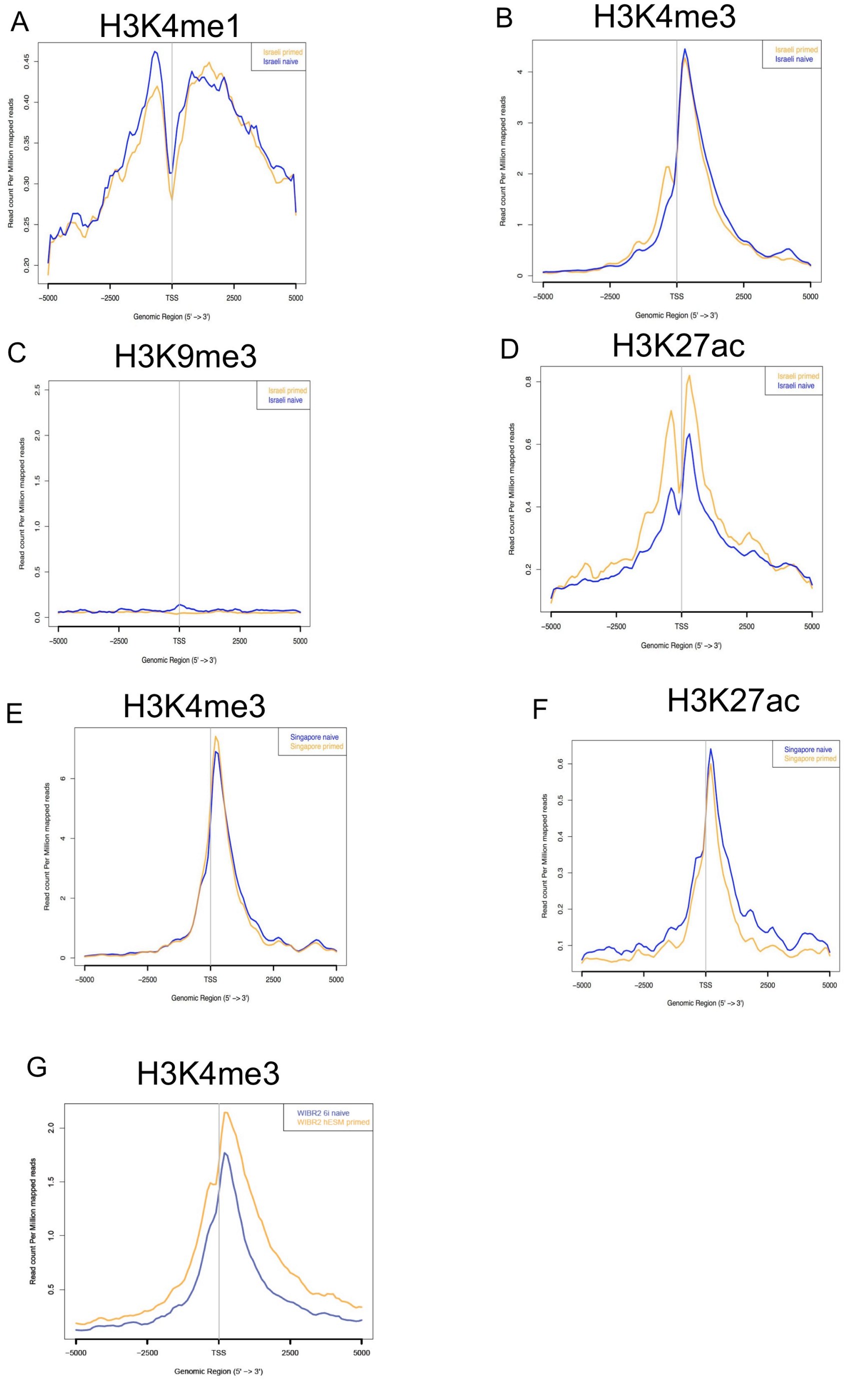
Group	Technology	Primed	Naïve
This study	RNA-seq	2.71	243.21
Chan et al	RNA-seq	0.07	7.54
Gafni et al	Microarray	7.24	8.07

Supplemental figure 9

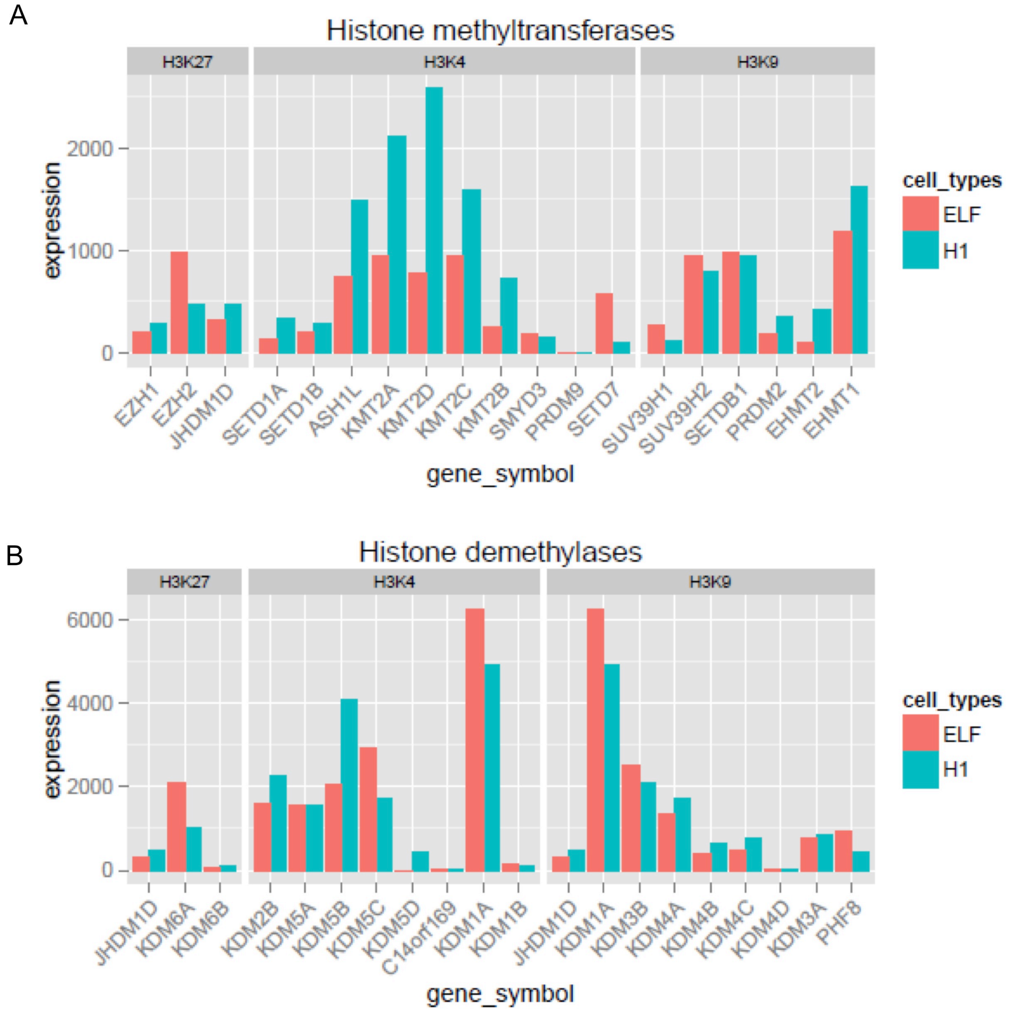
Expression of NNMT in various tissues



Supplemental figure 10



Supplemental figure 11



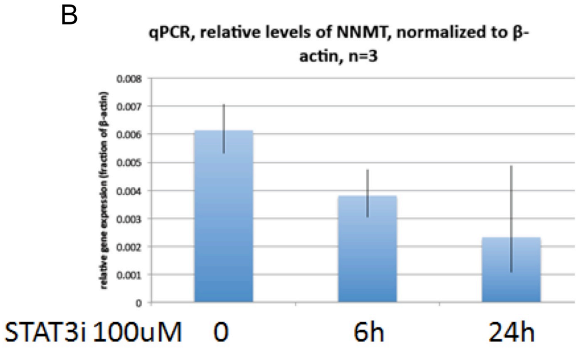
Supplemental figure 12

A

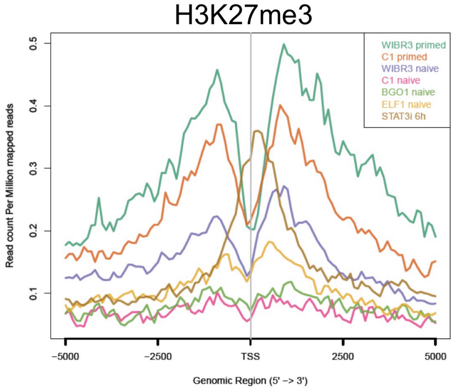
STAT3 is up-regulated
in naïve hESCs

TF	Log2 FC
EP300	-1.02
STAT1	-0.45
STAT3	-2.81
JUN	1.92

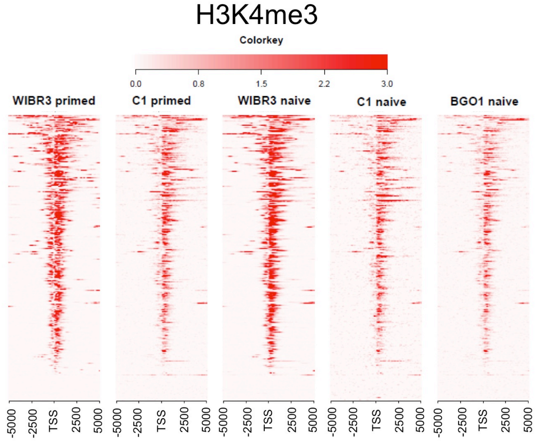
B



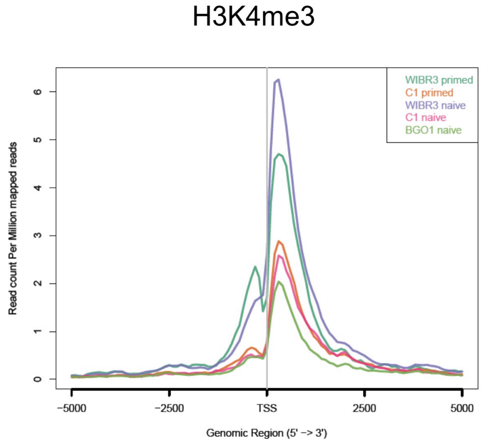
C



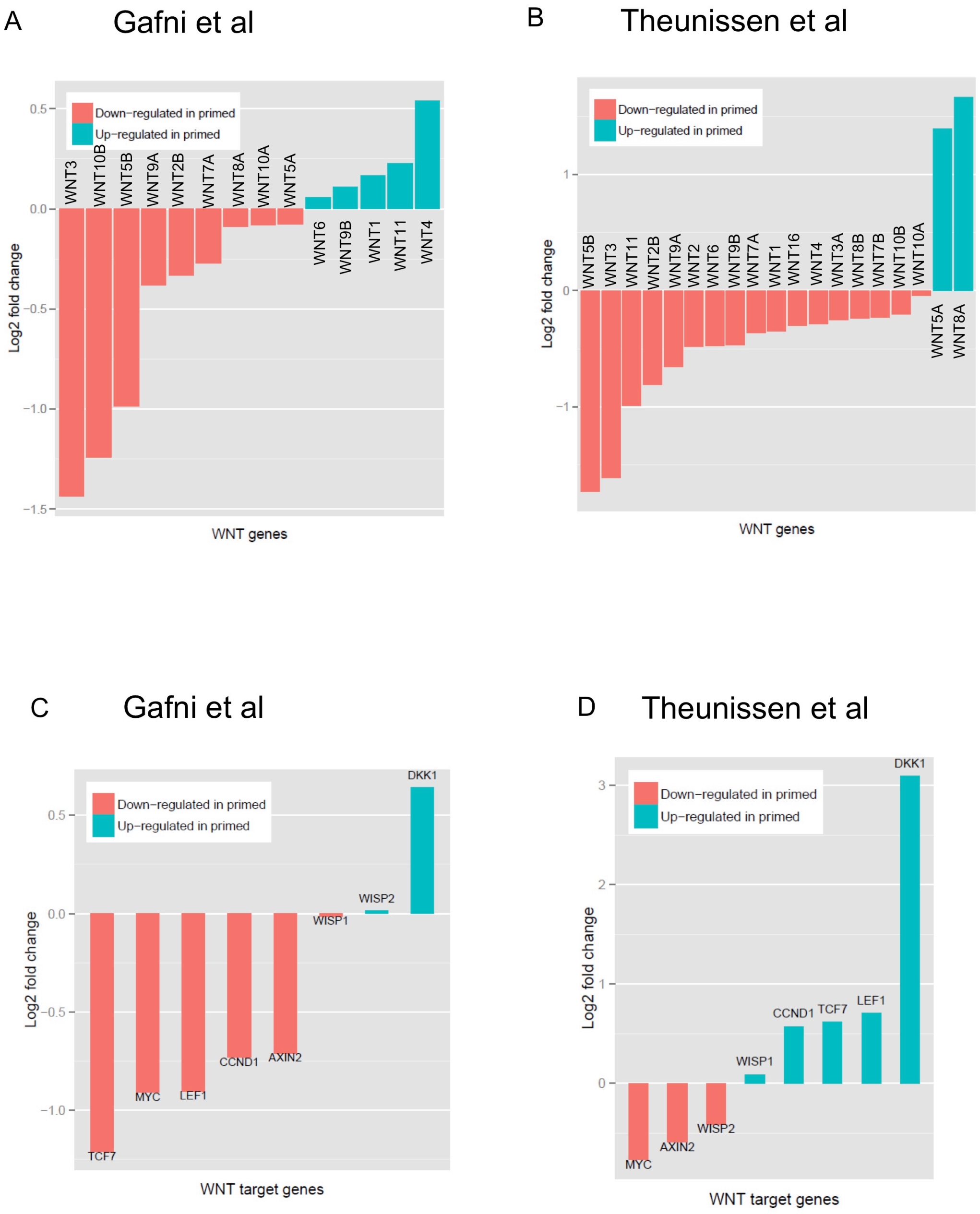
D



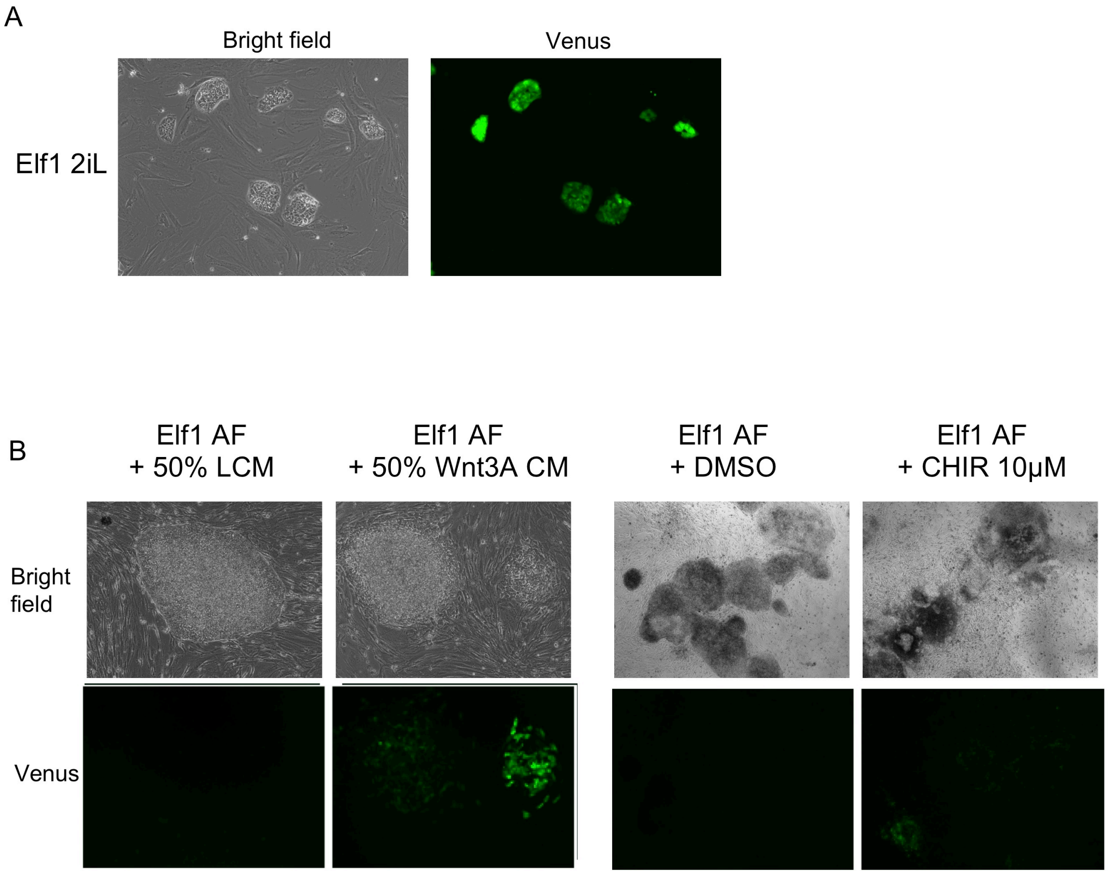
E



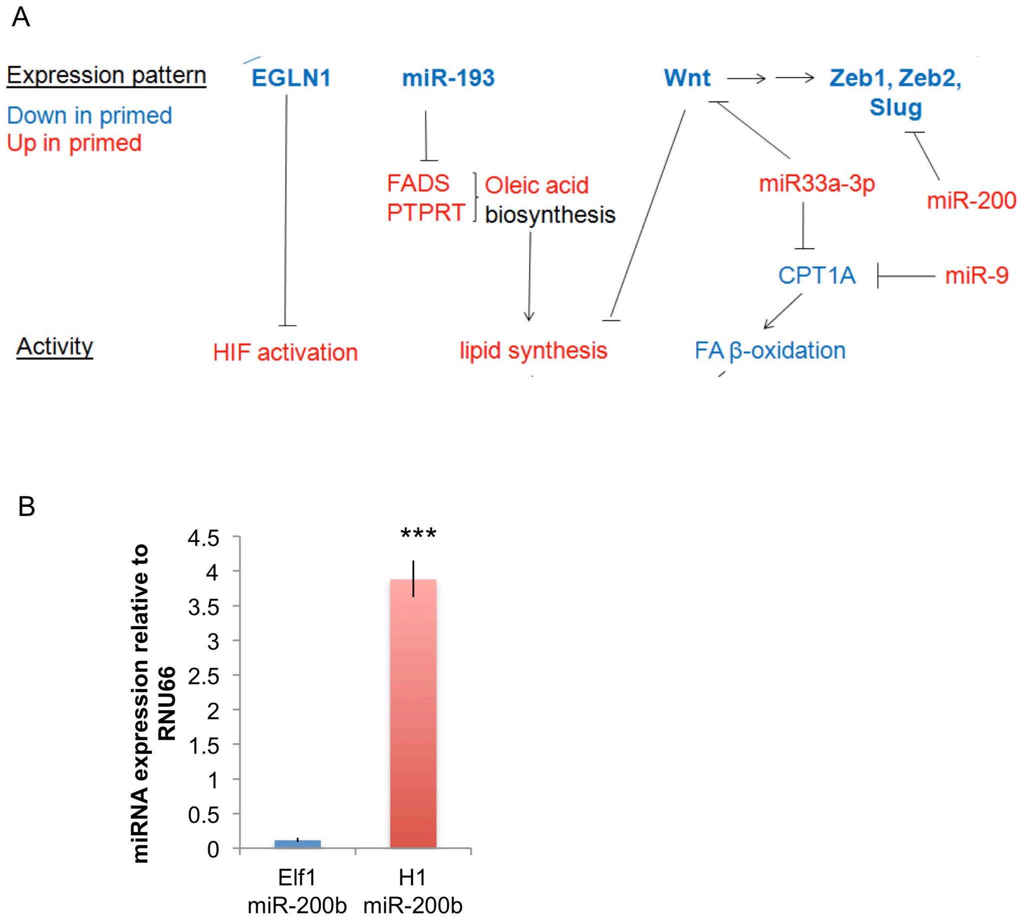
Supplemental figure 13



Supplemental figure 14



Supplemental figure 15



Supplemental Table 2

Primers used to quantify mitochondria copy number in human cells:

18s1546F (5’TAGAGGGACAAGTGGCGTTC3’)

18s1650R, (5’CGCTGAGCCAGTCAGTGT3’)

for genomic DNA with TaqMan probes

VIC-5’AGCAATAACAGGTCTGTGATG3’-TAMRA

ND13485F (5’CCCTAAAACCCGCCACATCT3’)

ND13553R, (5’GAGCGATGGTGAGAGCTAAGGT3’)

for mtDNA with TaqMan probes

6FAM-5’CCATCACCCTCTACATCACCGCCC3’-TAMRA

The following primer/probe sets were used with human total DNA for mtDNA mutation detection and copy number measurement:

mtDNA control site:

5′-CTA AAA ATA TTA AAC ACA AAC TAC CAC CTA CCTC-3′ (forward primer)

5′-GTT CAT TTT GGT TCT CAG GGT TTG TTA TAA-3′ (reverse primer)

5′-6FAM- CCT CAC CAA AGC CCA TA-MGB-3′ (probe)

ND1/ND2 deletion site:

5′-CGC CAC ATC TAC CAT CACC-3′ (forward primer)

5′-GAT TAT GGA TGC GGT TGC TT-3′ (reverse primer)

5′-6FAM-TTG ATG GCA GCT TCT GT-MGB-3′ (probe)

Common deletion site:

5′-TAC CCC CTC TAG AGC CCA CT-3′ (forward primer)

5′-GAG GAA AGG TAT TCC TGC TAA TGCT-3′ (reverse primer)

5′-6FAM-TGG CCC ACC ATA ATTMGB-3′ (probe)

12S rRNA point mutation site:

5’-CCA GAA CAC TAC GAG CCA CA-3’ (forward primer)

5’-TAC GTG GGT ACT TGC GCT TAC-3’ (reverse primer)

5’-6FAM-TTG CGC TTA CTT TGT A-MGB-3’ (probe)

COXII point mutation site:

5’-TGG CCA CCA ATG GTA CTG AAC CTA-3’ (forward primer)

5’-ACG TCC GGG AAT TGC ATC TGT T-3’ (reverse primer)

5’-6FAM-TCG CAG GTC GCC TGG TTC TAG GAA-BHQ1-3’ (probe)

RNaseP site for nuclear DNA quantification:

5’-AGA TTT GGA CCT GCG AGC-3’ (forward primer)

5’-GAG CGG CTG TCT CCA CAA GT-3’ (reverse primer)

5’-VIC-CCG CGC AGA GCC TTC-MGB-3’ (probe) qPCR primers for gene expression analysis:

NNMT (Forward TCCTTTAAGGAGATCGTCGTCAC; Reverse GCTTGACCGCCTGTCTCAAC)

IDO1 (Forward TTCAGTGCTTTGACGTCCTG; Reverse TGGAGGAACTGAGCAGCAT)

β-actin (Forward TCCCTGGAGAAGAGCTACG; Reverse GTAGTTTCGTGGATGCCACA)

References

Abdelmohsen, K., Srikantan, S., Kang, M.J., and Gorospe, M. (2012). Regulation of senescence by microRNA biogenesis factors. *Ageing Res Rev* 11**,** 491-500.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol* 11**,** R106.

Anokye-Danso, F., Trivedi, C.M., Juhr, D., Gupta, M., Cui, Z., Tian, Y., Zhang, Y., Yang, W., Gruber, P.J., Epstein, J.A., and Morrisey, E.E. (2011). Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. *Cell Stem Cell* 8**,** 376-388.

Bar, M., Wyman, S.K., Fritz, B.R., Qi, J., Garg, K.S., Parkin, R.K., Kroh, E.M., Bendoraite, A., Mitchell, P.S., Nelson, A.M., Ruzzo, W.L., Ware, C., Radich, J.P., Gentleman, R., Ruohola-Baker, H., and Tewari, M. (2008). MicroRNA discovery and profiling in human embryonic stem cells by deep sequencing of small RNA libraries. *Stem Cells* 26**,** 2496-2505.

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116**,** 281-297.

Basyuk, E., Suavet, F., Doglio, A., Bordonne, R., and Bertrand, E. (2003). Human let-7 stem-loop precursors harbor features of RNase III cleavage products. *Nucleic Acids Res* 31**,** 6593-6597.

Berezikov, E., Chung, W.J., Willis, J., Cuppen, E., and Lai, E.C. (2007). Mammalian mirtron genes. *Mol Cell* 28**,** 328-336.

Bernstein, B.E., Stamatoyannopoulos, J.A., Costello, J.F., Ren, B., Milosavljevic, A., Meissner, A., Kellis, M., Marra, M.A., Beaudet, A.L., Ecker, J.R., Farnham, P.J., Hirst, M., Lander, E.S., Mikkelsen, T.S., and Thomson, J.A. (2010). The NIH Roadmap Epigenomics Mapping Consortium. *Nat Biotechnol* 28, 1045-1048.

Bernstein, E., and Allis, C.D. (2005). RNA meets chromatin. *Genes Dev* 19, 1635-1655.

Berra, E., Benizri, E., Ginouves, A., Volmat, V., Roux, D., and Pouyssegur, J. (2003). HIF prolyl-hydroxylase 2 is the key oxygen sensor setting low steady-state levels of HIF-1alpha in normoxia. *EMBO J* 22, 4082-4090.

Biechele, T.L., Adams, A.M., and Moon, R.T. (2009). Transcription-based reporters of Wnt/beta-catenin signaling. *Cold Spring Harb Protoc* 2009, pdb prot5223.

Borchert, G.M., Lanier, W., and Davidson, B.L. (2006). RNA polymerase III transcribes human microRNAs. *Nat Struct Mol Biol* 13, 1097-1101.

Bracha, A.L., Ramanathan, A., Huang, S., Ingber, D.E., and Schreiber, S.L. (2010). Carbon metabolism-mediated myogenic differentiation. *Nat Chem Biol* 6, 202-204.

Brueckner, B., Stresemann, C., Kuner, R., Mund, C., Musch, T., Meister, M., Sultmann, H., and Lyko, F. (2007). The human let-7a-3 locus contains an epigenetically regulated microRNA gene with oncogenic function. *Cancer Res* 67, 1419-1423.

Buecker, C., Srinivasan, R., Wu, Z., Calo, E., Acampora, D., Faial, T., Simeone, A., Tan, M., Swigut, T., and Wysocka, J. (2014). Reorganization of enhancer patterns in transition from naive to primed pluripotency. *Cell Stem Cell* 14, 838-853.

Chadwick, L.H. (2012). The NIH Roadmap Epigenomics Program data resource. *Epigenomics* 4, 317-324.

Chambers, S.M., Fasano, C.A., Papapetrou, E.P., Tomishima, M., Sadelain, M., and Studer, L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat Biotechnol* 27, 275-280.

Chan, Y.S., Goke, J., Ng, J.H., Lu, X., Gonzales, K.A., Tan, C.P., Tng, W.Q., Hong, Z.Z., Lim, Y.S., and Ng, H.H. (2013). Induction of a human pluripotent state with distinct regulatory circuitry that resembles preimplantation epiblast. *Cell Stem Cell* 13**,** 663-675.

Cheloufi, S., Dos Santos, C.O., Chong, M.M., and Hannon, G.J. (2010). A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* 465**,** 584-589.

Chendrimada, T.P., Gregory, R.I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., and Shiekhattar, R. (2005). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* 436**,** 740-744.

Chong, M.M., Zhang, G., Cheloufi, S., Neubert, T.A., Hannon, G.J., and Littman, D.R. (2010). Canonical and alternate functions of the microRNA biogenesis machinery. *Genes Dev* 24**,** 1951-1960.

Cullen, B.R. (2006). Viruses and microRNAs. *Nat Genet* 38 Suppl**,** S25-30.

Dai, D.F., Karunadharma, P.P., Chiao, Y.A., Basisty, N., Crispin, D., Hsieh, E.J., Chen, T., Gu, H., Djukovic, D., Raftery, D., Beyer, R.P., Maccoss, M.J., and Rabinovitch, P.S. (2014). Altered proteome turnover and remodeling by short-term caloric restriction or rapamycin rejuvenate the aging heart. *Aging Cell* 13**,** 529-539.

Davidson, K.C., Adams, A.M., Goodson, J.M., Mcdonald, C.E., Potter, J.C., Berndt, J.D., Biechele, T.L., Taylor, R.J., and Moon, R.T. (2012). Wnt/beta-catenin signaling promotes differentiation, not self-renewal, of human embryonic stem cells and is repressed by Oct4. *Proc Natl Acad Sci U S A* 109**,** 4485-4490.

Davis, B.N., and Hata, A. (2009). Regulation of MicroRNA Biogenesis: A miRiad of mechanisms. *Cell Commun Signal* 7**,** 18.

Davis, B.N., Hilyard, A.C., Lagna, G., and Hata, A. (2008). SMAD proteins control DROSHA-mediated microRNA maturation. *Nature* 454, 56-61.

Diederichs, S., and Haber, D.A. (2006). Sequence variations of microRNAs in human cancer: alterations in predicted secondary structure do not affect processing. *Cancer Res* 66, 6097-6104.

Du, T., and Zamore, P.D. (2005). microPrimer: the biogenesis and function of microRNA. *Development* 132, 4645-4652.

Duan, R., Pak, C., and Jin, P. (2007). Single nucleotide polymorphism associated with mature miR-125a alters the processing of pri-miRNA. *Hum Mol Genet* 16, 1124-1131.

Escobar, T.M., Kanellopoulou, C., Kugler, D.G., Kilaru, G., Nguyen, C.K., Nagarajan, V., Bhairavabhotla, R.K., Northrup, D., Zahr, R., Burr, P., Liu, X., Zhao, K., Sher, A., Jankovic, D., Zhu, J., and Muljo, S.A. (2014). miR-155 activates cytokine gene expression in Th17 cells by regulating the DNA-binding protein Jarid2 to relieve polycomb-mediated repression. *Immunity* 40, 865-879.

Factor, D.C., Corradin, O., Zentner, G.E., Saiakhova, A., Song, L., Chenoweth, J.G., Mckay, R.D., Crawford, G.E., Scacheri, P.C., and Tesar, P.J. (2014). Epigenomic Comparison Reveals Activation of "Seed" Enhancers during Transition from Naive to Primed Pluripotency. *Cell Stem Cell* 14, 854-863.

Feng, Y., Zhang, X., Graves, P., and Zeng, Y. (2012). A comprehensive analysis of precursor microRNA cleavage by human Dicer. *RNA* 18, 2083-2092.

Feng, Y., Zhang, X., Song, Q., Li, T., and Zeng, Y. (2011). Drosha processing controls the specificity and efficiency of global microRNA expression. *Biochim Biophys Acta* 1809, 700-707.

Flicek, P., Amode, M.R., Barrell, D., Beal, K., Billis, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fitzgerald, S., Gil, L., Giron, C.G., Gordon, L., Hourlier, T., Hunt, S., Johnson, N., Juettemann, T., Kahari, A.K., Keenan, S., Kulesha, E., Martin, F.J., Maurel, T., Mclaren, W.M., Murphy, D.N., Nag, R., Overduin, B., Pignatelli, M., Pritchard, B., Pritchard, E., Riat, H.S., Ruffier, M., Sheppard, D., Taylor, K., Thormann, A., Trevanion, S.J., Vullo, A., Wilder, S.P., Wilson, M., Zadissa, A., Aken, B.L., Birney, E., Cunningham, F., Harrow, J., Herrero, J., Hubbard, T.J., Kinsella, R., Muffato, M., Parker, A., Spudich, G., Yates, A., Zerbino, D.R., and Searle, S.M. (2014). Ensembl 2014. *Nucleic Acids Res* 42**,** D749-755.

Folmes, C.D., Nelson, T.J., Martinez-Fernandez, A., Arrell, D.K., Lindor, J.Z., Dzeja, P.P., Ikeda, Y., Perez-Terzic, C., and Terzic, A. (2011). Somatic oxidative bioenergetics transitions into pluripotency-dependent glycolysis to facilitate nuclear reprogramming. *Cell Metab* 14**,** 264-271.

Fukuda, T., Yamagata, K., Fujiyama, S., Matsumoto, T., Koshida, I., Yoshimura, K., Mihara, M., Naitou, M., Endoh, H., Nakamura, T., Akimoto, C., Yamamoto, Y., Katagiri, T., Foulds, C., Takezawa, S., Kitagawa, H., Takeyama, K., O'malley, B.W., and Kato, S. (2007). DEAD-box RNA helicase subunits of the Drosha complex are required for processing of rRNA and a subset of microRNAs. *Nat Cell Biol* 9**,** 604-611.

Gafni, O., Weinberger, L., Mansour, A.A., Manor, Y.S., 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., and Hanna, J.H.

(2013). Derivation of novel human ground state naive pluripotent stem cells. *Nature* 504**,** 282-286.

Graf, U., Casanova, E.A., and Cinelli, P. (2011). The Role of the Leukemia Inhibitory Factor (LIF) - Pathway in Derivation and Maintenance of Murine Pluripotent Stem Cells. *Genes (Basel)* 2**,** 280-297.

Greer, S.N., Metcalf, J.L., Wang, Y., and Ohh, M. (2012). The updated biology of hypoxia-inducible factor. *EMBO J* 31**,** 2448-2460.

Gregory, R.I., Yan, K.P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. *Nature* 432**,** 235-240.

Griffiths-Jones, S., Grocock, R.J., Van Dongen, S., Bateman, A., and Enright, A.J. (2006). miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res* 34**,** D140-144.

Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jr., Jungkamp, A.C., Munschauer, M., Ulrich, A., Wardle, G.S., Dewell, S., Zavolan, M., and Tuschl, T. (2010). Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* 141**,** 129-141.

Han, J., Lee, Y., Yeom, K.H., Nam, J.W., Heo, I., Rhee, J.K., Sohn, S.Y., Cho, Y., Zhang, B.T., and Kim, V.N. (2006). Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* 125**,** 887-901.

Han, J., Pedersen, J.S., Kwon, S.C., Belair, C.D., Kim, Y.K., Yeom, K.H., Yang, W.Y., Haussler, D., Blelloch, R., and Kim, V.N. (2009). Posttranscriptional crossregulation between Drosha and DGCR8. *Cell* 136**,** 75-84.

Hatfield, S.D., Shcherbata, H.R., Fischer, K.A., Nakahara, K., Carthew, R.W., and Ruohola-Baker, H. (2005). Stem cell division is regulated by the microRNA pathway. *Nature* 435, 974-978.

Hawkins, R.D., Hon, G.C., Lee, L.K., Ngo, Q., Lister, R., Pelizzola, M., Edsall, L.E., Kuan, S., Luu, Y., Klugman, S., Antosiewicz-Bourget, J., Ye, Z., Espinoza, C., Agarwahl, S., Shen, L., Ruotti, V., Wang, W., Stewart, R., Thomson, J.A., Ecker, J.R., and Ren, B. (2010). Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell* 6, 479-491.

Hofacker, I.L., Fontana, W., Stadler, P.F., Bonhoeffer, L.S., Tacker, M., and Schuster, P. (1994). Fast folding and comparison of RNA secondary structures. *Monatsh Chem* 125, 167-188.

Hsu, S.D., Tseng, Y.T., Shrestha, S., Lin, Y.L., Khaleel, A., Chou, C.H., Chu, C.F., Huang, H.Y., Lin, C.M., Ho, S.Y., Jian, T.Y., Lin, F.M., Chang, T.H., Weng, S.L., Liao, K.W., Liao, I.E., Liu, C.C., and Huang, H.D. (2014). miRTarBase update 2014: an information resource for experimentally validated miRNA-target interactions. *Nucleic Acids Res* 42, D78-85.

Hutvagner, G., Mclachlan, J., Pasquinelli, A.E., Balint, E., Tuschl, T., and Zamore, P.D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* 293, 834-838.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., and Speed, T.P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4, 249-264.

Kim, V.N., and Nam, J.W. (2006). Genomics of microRNA. *Trends Genet* 22, 165-173.

Kind, T., Liu, K.H., Lee Do, Y., Defelice, B., Meissen, J.K., and Fiehn, O. (2013). LipidBlast in silico tandem mass spectrometry database for lipid identification. *Nat Methods* 10**,** 755-758.

Kraus, D., Yang, Q., Kong, D., Banks, A.S., Zhang, L., Rodgers, J.T., Pirinen, E., Pulinilkunnil, T.C., Gong, F., Wang, Y.C., Cen, Y., Sauve, A.A., Asara, J.M., Peroni, O.D., Monia, B.P., Bhanot, S., Alhonen, L., Puigserver, P., and Kahn, B.B. (2014). Nicotinamide N-methyltransferase knockdown protects against diet-induced obesity. *Nature* 508**,** 258-262.

Kuehbacher, A., Urbich, C., Zeiher, A.M., and Dimmeler, S. (2007). Role of Dicer and Drosha for endothelial microRNA expression and angiogenesis. *Circ Res* 101**,** 59-68.

Kuppusamy, K.T., Sperber, H., and Ruohola-Baker, H. (2013). MicroRNA Regulation and role in stem cell maintenance, cardiac differentiation and hypertrophy. *Curr Mol Med* Submitted.

Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. *Science* 294**,** 853-858.

Landeira, D., and Fisher, A.G. (2011). Inactive yet indispensable: the tale of Jarid2. *Trends Cell Biol* 21**,** 74-80.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10**,** R25.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S., and Kim, V.N. (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425**,** 415-419.

Lewis, B.P., Burge, C.B., and Bartel, D.P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 120**,** 15-20.

Mathieu, J., Zhou, W., Xing, Y., Sperber, H., Ferreccio, A., Agoston, Z., Kuppusamy, K.T., Moon, R.T., and Ruohola-Baker, H. (2014). Hypoxia-inducible factors have distinct and stage-specific roles during reprogramming of human cells to pluripotency. *Cell Stem Cell* 14**,** 592-605.

Mattick, J.S., and Makunin, I.V. (2006). Non-coding RNA. *Hum Mol Genet* 15 Spec No 1**,** R17-29.

Mazumdar, J., O'brien, W.T., Johnson, R.S., Lamanna, J.C., Chavez, J.C., Klein, P.S., and Simon, M.C. (2010). O2 regulates stem cells through Wnt/beta-catenin signalling. *Nat Cell Biol* 12**,** 1007-1013.

Meissen, J.K., Yuen, B.T., Kind, T., Riggs, J.W., Barupal, D.K., Knoepfler, P.S., and Fiehn, O. (2012). Induced pluripotent stem cells show metabolomic differences to embryonic stem cells in polyunsaturated phosphatidylcholines and primary metabolism. *PLoS One* 7**,** e46770.

Nam, Y., Chen, C., Gregory, R.I., Chou, J.J., and Sliz, P. (2011). Molecular basis for interaction of let-7 microRNAs with Lin28. *Cell* 147**,** 1080-1091.

Nguyen-Tran, D.H., Hait, N.C., Sperber, H., Qi, J., Fischer, K., Ieronimakis, N., Pantoja, M., Hays, A., Allegood, J., Reyes, M., Spiegel, S., and Ruohola-Baker, H. (2014). Molecular mechanism of sphingosine-1-phosphate action in Duchenne muscular dystrophy. *Dis Model Mech* 7**,** 41-54.

Noland, C.L., Ma, E., and Doudna, J.A. (2011). siRNA repositioning for guide strand selection by human Dicer complexes. *Mol Cell* 43**,** 110-121.

Nygren, H., Seppanen-Laakso, T., Castillo, S., Hyotylainen, T., and Oresic, M. (2011). Liquid chromatography-mass spectrometry (LC-MS)-based lipidomics for studies of body fluids and tissues. *Methods Mol Biol* 708**,** 247-257.

O'donnell, K.A., Wentzel, E.A., Zeller, K.I., Dang, C.V., and Mendell, J.T. (2005). c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* 435**,** 839-843.

Opitz, C.A., Litzenburger, U.M., Sahm, F., Ott, M., Tritschler, I., Trump, S., Schumacher, T., Jestaedt, L., Schrenk, D., Weller, M., Jugold, M., Guillemin, G.J., Miller, C.L., Lutz, C., Radlwimmer, B., Lehmann, I., Von Deimling, A., Wick, W., and Platten, M. (2011). An endogenous tumour-promoting ligand of the human aryl hydrocarbon receptor. *Nature* 478**,** 197-203.

Panopoulos, A.D., Yanes, O., Ruiz, S., Kida, Y.S., Diep, D., Tautenhahn, R., Herrerias, A., Batchelder, E.M., Plongthongkum, N., Lutz, M., Berggren, W.T., Zhang, K., Evans, R.M., Siuzdak, G., and Izpisua Belmonte, J.C. (2012). The metabolome of induced pluripotent stem cells reveals metabolic changes occurring in somatic cell reprogramming. *Cell Res* 22**,** 168-177.

Pinheiro, L.B., Coleman, V.A., Hindson, C.M., Herrmann, J., Hindson, B.J., Bhat, S., and Emslie, K.R. (2012). Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification. *Anal Chem* 84**,** 1003-1011.

Prigione, A., and Adjaye, J. (2010). Modulation of mitochondrial biogenesis and bioenergetic metabolism upon in vitro and in vivo differentiation of human ES and iPS cells. *Int J Dev Biol* 54**,** 1729-1741.

Qi, J., Yu, J.Y., Shcherbata, H.R., Mathieu, J., Wang, A.J., Seal, S., Zhou, W., Stadler, B.M., Bourgin, D., Wang, L., Nelson, A., Ware, C., Raymond, C., Lim, L.P., Magnus, J., Ivanovska, I., Diaz, R., Ball, A., Cleary, M.A., and Ruohola-Baker, H. (2009). microRNAs regulate human embryonic stem cell division. *Cell Cycle* 8**,** 3729-3741.

Rafalski, V.A., Mancini, E., and Brunet, A. (2012). Energy metabolism and energy-sensing pathways in mammalian embryonic and adult stem cell fate. *J Cell Sci* 125**,** 5597-5608.

Rezakhanlou, A.M., Habibi, D., Lai, A., Jalili, R.B., Ong, C.J., and Ghahary, A. (2010). Highly efficient stable expression of indoleamine 2,3 dioxygenase gene in primary fibroblasts. *Biol Proced Online* 12**,** 9028.

Ritchie, W., Legendre, M., and Gautheret, D. (2007). RNA stem-loops: to be or not to be cleaved by RNAse III. *RNA* 13**,** 457-462.

Ruby, J.G., Jan, C.H., and Bartel, D.P. (2007). Intronic microRNA precursors that bypass Drosha processing. *Nature* 448**,** 83-86.

Shcherbata, H.R., Hatfield, S., Ward, E.J., Reynolds, S., Fischer, K.A., and Ruohola-Baker, H. (2006). The MicroRNA pathway plays a regulatory role in stem cell division. *Cell Cycle* 5**,** 172-175.

Shen, L., Shao, N.Y., Liu, X., Maze, I., Feng, J., and Nestler, E.J. (2013). diffReps: detecting differential chromatin modification sites from ChIP-seq data with biological replicates. *PLoS One* 8**,** e65598.

Shiraki, N., Shiraki, Y., Tsuyama, T., Obata, F., Miura, M., Nagae, G., Aburatani, H., Kume, K., Endo, F., and Kume, S. (2014). Methionine metabolism regulates maintenance and differentiation of human pluripotent stem cells. *Cell Metab* 19**,** 780-794.

Shyh-Chang, N., Locasale, J.W., Lyssiotis, C.A., Zheng, Y., Teo, R.Y., Ratanasirintrawoot, S., Zhang, J., Onder, T., Unternaehrer, J.J., Zhu, H., Asara, J.M., Daley, G.Q., and Cantley, L.C. (2013). Influence of threonine metabolism on S-adenosylmethionine and histone methylation. *Science* 339**,** 222-226.

Simonson, T.S., Yang, Y., Huff, C.D., Yun, H., Qin, G., Witherspoon, D.J., Bai, Z., Lorenzo, F.R., Xing, J., Jorde, L.B., Prchal, J.T., and Ge, R. (2010). Genetic evidence for high-altitude adaptation in Tibet. *Science* 329**,** 72-75.

Sperber, H., Beem, A., Shannon, S., Jones, R., Banik, P., Chen, Y., Ku, S., Varani, G., Yao, S., and Ruohola-Baker, H. (2014). miRNA sensitivity to Drosha levels correlates with pre-miRNA secondary structure. *RNA* 20**,** 621-631.

Stadler, B., Ivanovska, I., Mehta, K., Song, S., Nelson, A., Tan, Y., Mathieu, J., Darby, C., Blau, C.A., Ware, C., Peters, G., Miller, D.G., Shen, L., Cleary, M.A., and Ruohola-Baker, H. (2010). Characterization of microRNAs involved in embryonic stem cell states. *Stem Cells Dev* 19**,** 935-950.

Stadler, B.M., and Ruohola-Baker, H. (2008). Small RNAs: keeping stem cells in line. *Cell* 132**,** 563-566.

Starega-Roslan, J., Krol, J., Koscianska, E., Kozlowski, P., Szlachcic, W.J., Sobczak, K., and Krzyzosiak, W.J. (2010). Structural basis of microRNA length variety. *Nucleic Acids Res* 39**,** 257-268.

Stark, A., Brennecke, J., Bushati, N., Russell, R.B., and Cohen, S.M. (2005). Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution. *Cell* 123**,** 1133-1146.

Sun, G., Yan, J., Noltner, K., Feng, J., Li, H., Sarkis, D.A., Sommer, S.S., and Rossi, J.J. (2009). SNPs in human miRNA genes affect biogenesis and function. *RNA* 15**,** 1640-1651.

Takada, S., Berezikov, E., Yamashita, Y., Lagos-Quintana, M., Kloosterman, W.P., Enomoto, M., Hatanaka, H., Fujiwara, S., Watanabe, H., Soda, M., Choi, Y.L., Plasterk, R.H., Cuppen, E., and Mano, H. (2006). Mouse microRNA profiles determined with a new and sensitive cloning method. *Nucleic Acids Res* 34**,** e115.

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

Taylor, S.D., Ericson, N.G., Burton, J.N., Prolla, T.A., Silber, J.R., Shendure, J., and Bielas, J.H. (2014). Targeted enrichment and high-resolution digital profiling of mitochondrial DNA deletions in human brain. *Aging Cell* 13**,** 29-38.

Ten Berge, D., Kurek, D., Blauwkamp, T., Koole, W., Maas, A., Eroglu, E., Siu, R.K., and Nusse, R. (2011). Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells. *Nat Cell Biol* 13**,** 1070-1075.

Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and Mckay, R.D. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* 448**,** 196-199.

Theunissen, T.W., Powell, B.E., Wang, H., Mitalipova, M., Faddah, D.A., Reddy, J., Fan, Z.P., Maetzel, D., Ganz, K., Shi, L., Lungjangwa, T., Imsoonthornruksa, S., Stelzer, Y., Rangarajan, S., D'alessio, A., Zhang, J., Gao, Q., Dawlaty, M.M., Young, R.A., Gray, N.S., and Jaenisch, R. (2014). Systematic Identification of Culture Conditions for Induction and Maintenance of Naive Human Pluripotency. *Cell Stem Cell*.

Tomida, M., Ohtake, H., Yokota, T., Kobayashi, Y., and Kurosumi, M. (2008). Stat3 up-regulates expression of nicotinamide N-methyltransferase in human cancer cells. *J Cancer Res Clin Oncol* 134**,** 551-559.

Trojer, P., and Reinberg, D. (2006). Histone lysine demethylases and their impact on epigenetics. *Cell* 125**,** 213-217.

Ulanovskaya, O.A., Zuhl, A.M., and Cravatt, B.F. (2013). NNMT promotes epigenetic remodeling in cancer by creating a metabolic methylation sink. *Nat Chem Biol* 9**,** 300-306.

Valamehr, B., Robinson, M., Abujarour, R., Rezner, B., Vranceanu, F., Le, T., Medcalf, A., Lee, T.T., Fitch, M., Robbins, D., and Flynn, P. (2014). Platform for Induction and Maintenance of Transgene-free hiPSCs Resembling Ground State Pluripotent Stem Cells. *Stem Cell Reports* 2**,** 366-381.

Varum, S., Rodrigues, A.S., Moura, M.B., Momcilovic, O., Easley, C.a.T., Ramalho-Santos, J., Van Houten, B., and Schatten, G. (2011). Energy metabolism in human pluripotent stem cells and their differentiated counterparts. *PLoS One* 6**,** e20914.

Wang, W., Zhou, H., Lin, H., Roy, S., Shaler, T.A., Hill, L.R., Norton, S., Kumar, P., Anderle, M., and Becker, C.H. (2003). Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards. *Anal Chem* 75**,** 4818-4826.

Ware, C.B., Nelson, A.M., and Blau, C.A. (2006). A comparison of NIH-approved human ESC lines. *Stem Cells* 24**,** 2677-2684.

Ware, C.B., Nelson, A.M., Mecham, B., Hesson, J., Zhou, W., Jonlin, E.C., Jimenez-Caliani, A.J., Deng, X., Cavanaugh, C., Cook, S., Tesar, P.J., Okada, J., Margaretha, L., Sperber, H., Choi, M., Blau, C.A., Treuting, P.M., Hawkins, R.D., Cirulli, V., and Ruohola-Baker, H. (2014). Derivation of naive human embryonic stem cells. *Proc Natl Acad Sci U S A*.

Watkins, S.M., Reifsnyder, P.R., Pan, H.J., German, J.B., and Leiter, E.H. (2002). Lipid metabolome-wide effects of the PPARgamma agonist rosiglitazone. *J Lipid Res* 43**,** 1809-1817.

Wolfgang, M.J., Kurama, T., Dai, Y., Suwa, A., Asaumi, M., Matsumoto, S., Cha, S.H., Shimokawa, T., and Lane, M.D. (2006). The brain-specific carnitine palmitoyltransferase-1c regulates energy homeostasis. *Proc Natl Acad Sci U S A* 103**,** 7282-7287.

Wu, M., Jolicoeur, N., Li, Z., Zhang, L., Fortin, Y., L'abbe, D., Yu, Z., and Shen, S.H. (2008). Genetic variations of microRNAs in human cancer and their effects on the expression of miRNAs. *Carcinogenesis* 29**,** 1710-1716.

Wu, T.D., and Nacu, S. (2010). Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics* 26**,** 873-881.

Xu, G., Fewell, C., Taylor, C., Deng, N., Hedges, D., Wang, X., Zhang, K., Lacey, M., Zhang, H., Yin, Q., Cameron, J., Lin, Z., Zhu, D., and Flemington, E.K. (2010). Transcriptome and targetome analysis in MIR155 expressing cells using RNA-seq. *RNA* 16**,** 1610-1622.

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., Qiao, J., and Tang, F. (2013). Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat Struct Mol Biol* 20**,** 1131-1139.

Yanes, O., Clark, J., Wong, D.M., Patti, G.J., Sanchez-Ruiz, A., Benton, H.P., Trauger, S.A., Desponts, C., Ding, S., and Siuzdak, G. (2010). Metabolic oxidation regulates embryonic stem cell differentiation. *Nat Chem Biol* 6**,** 411-417.

Yu, Y., Fuscoe, J.C., Zhao, C., Guo, C., Jia, M., Qing, T., Bannon, D.I., Lancashire, L., Bao, W., Du, T., Luo, H., Su, Z., Jones, W.D., Moland, C.L., Branham, W.S., Qian, F., Ning, B., Li, Y., Hong, H., Guo, L., Mei, N., Shi, T., Wang, K.Y., Wolfinger, R.D., Nikolsky, Y., Walker, S.J., Duerksen-Hughes, P., Mason, C.E., Tong, W., Thierry-Mieg, J., Thierry-Mieg, D., Shi, L., and Wang, C. (2014). A rat RNA-Seq transcriptomic BodyMap across 11 organs and 4 developmental stages. *Nat Commun* 5**,** 3230.

Zeng, Y., and Cullen, B.R. (2003). Sequence requirements for micro RNA processing and function in human cells. *RNA* 9**,** 112-123.

Zeng, Y., Yi, R., and Cullen, B.R. (2005). Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. *EMBO J* 24**,** 138-148.

Zhang, J., Khvorostov, I., Hong, J.S., Oktay, Y., Vergnes, L., Nuebel, E., Wahjudi, P.N., Setoguchi, K., Wang, G., Do, A., Jung, H.J., Mccaffery, J.M., Kurland, I.J., Reue, K., Lee, W.N., Koehler, C.M., and Teitell, M.A. (2011). UCP2 regulates energy metabolism and differentiation potential of human pluripotent stem cells. *EMBO J* 30**,** 4860-4873.

Zhou, W., Choi, M., Margineantu, D., Margaretha, L., Hesson, J., Cavanaugh, C., Blau, C.A., Horwitz, M.S., Hockenbery, D., Ware, C., and Ruohola-Baker, H. (2012). HIF1alpha induced switch from bivalent to exclusively glycolytic metabolism during ESC-to-EpiSC/hESC transition. *EMBO J* 31**,** 2103-2116.

Zisoulis, D.G., Kai, Z.S., Chang, R.K., and Pasquinelli, A.E. (2012). Autoregulation of microRNA biogenesis by let-7 and Argonaute. *Nature* 486**,** 541-544.