Epigenomic Comparison Reveals Activation of “Seed” Enhancers during Transition from Naive to Primed Pluripotency

Daniel C. Factor,1,6 Olivia Corradin,1,6 Gabriel E. Zentner,1,6 Alina Saiakhova,1 Lingyun Song,6 Josh G. Chenoweth,4 Ronald D. McKay,4 Gregory E. Crawford,6 Peter C. Scacheri,1,2,7,* and Paul J. Tesar1,2,3,7,*

1Department of Genetics and Genome Sciences, Case Western Reserve University School of Medicine, Cleveland, OH 44106, USA
2Case Comprehensive Cancer Center, Case Western Reserve University School of Medicine, Cleveland, OH 44106, USA
3National Center for Regenerative Medicine, Case Western Reserve University School of Medicine, Cleveland, OH 44106, USA
4Lieber Institute for Brain Development, Baltimore, MD 21205, USA
5Institute for Genome Sciences & Policy, Duke University, Durham, NC 27709, USA
6Co-first author
7Co-senior author
8Present address: Basic Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA
*Correspondence: peter.scacheri@case.edu (P.C.S.), paul.tesar@case.edu (P.J.T.)
http://dx.doi.org/10.1016/j.stem.2014.05.005

SUMMARY

Naive mouse embryonic stem cells (mESCs) and primed epiblast stem cells (mEpiSCs) represent successive snapshots of pluripotency during embryogenesis. Using transcriptomic and epigenomic mapping we show that a small fraction of transcripts are differentially expressed between mESCs and mEpiSCs and that these genes show expected changes in chromatin at their promoters and enhancers. Unexpectedly, the cis-regulatory circuitry of genes that are expressed at identical levels between these cell states also differs dramatically. In mESCs, these genes are associated with dominant proximal enhancers and dormant distal enhancers, which we term seed enhancers. In mEpiSCs, the naive-dominant enhancers are lost, and the seed enhancers take up primary transcriptional control. Seed enhancers have increased sequence conservation and show preferential usage in downstream somatic tissues, often expanding into super enhancers. We propose that seed enhancers ensure proper enhancer utilization and transcriptional fidelity as mammalian cells transition from naive pluripotency to a somatic regulatory program.

INTRODUCTION

At the molecular level, pluripotency is under the control of a complex array of regulatory mechanisms that maintain chromatin in a state permissive to differentiation into each of the early somatic and germ cell lineages. Recent evidence shows that pluripotency is not a single entity and can be maintained in either a “naive” or “primed” state (Brons et al., 2007; Chenoweth et al., 2010; Nichols and Smith, 2009; Tesar et al., 2007). Naive pluripotent cells, typified by mouse embryonic stem cells (mESCs), represent the preimplantation inner cell mass and are widely utilized for developmental genetics because they are capable of extensive contribution to chimeric animals upon reintroduction back into the blastocyst (Bradley et al., 1984; Evans and Kaufman, 1981; Martin, 1981). On the other hand, primed pluripotent cells, typified by mouse epiblast stem cells (mEpiSCs) and human embryonic stem cells (hESCs), represent the postimplantation epiblast, the next successive stage of pluripotency, which occurs immediately prior to differentiation into the three germ cell lineages at gastrulation (Brons et al., 2007; Najm et al., 2011; Tesar et al., 2007; Thomson et al., 1998). There is tremendous interest in understanding the differences between the naive and primed pluripotent states because they provide a direct window into the epigenetic dynamics in placental mammals that function to maintain pluripotency while simultaneously preparing to transition to a somatic regulatory program.

Enhancer elements establish and maintain expression patterns that drive normal development and cell identity. In comparison to promoters, the chromatin state of enhancers is divergent across different cell types. Even genes expressed broadly across different cell types can show dramatic differences in enhancer usage (Kieffer-Kwon et al., 2013). Recent evidence suggests that large genomic domains containing clusters of active enhancers, variously referred to as “super enhancers,” “stretch enhancers,” or “multiple enhancer variants” are particularly cell type specific, and they are proposed to mediate transcription of genes that are important for controlling cell identity (Corradin et al., 2014; Hnisz et al., 2013; Lovén et al., 2013; Parker et al., 2013; Whyte et al., 2013). These discoveries have largely been garnered from comparisons of regulatory landscapes of cell types derived from very different tissues and distinct stages of development. Here, we employ genomic approaches to directly characterize the regulatory landscapes of two closely related cell types, mESCs and mEpiSCs. These cell types represent successive snapshots of early development, share the core property of pluripotency, and largely share a common transcriptional program; however, their maintenance relies on distinct signaling pathways. Our analyses show that enhancer usage differs not only for genes that are differentially expressed...
between mESCs and mEpiSCs, but also for those that are similarly expressed between the two cell types. While enhancers unique to mESCs are decommissioned following the transition to primed pluripotency, those unique to mEpiSCs, which we term seed enhancers, are present in naive pluripotency, become active in primed pluripotency, and retain activity in somatic derivatives, often contributing to super enhancers.

RESULTS

Enhancer Profiles Distinguish Mouse Pluripotent States
To understand the differences in transcriptional regulation between mESCs and mEpiSCs, we performed epigenomic and transcriptome profiling of these two pluripotent cell types using high-quality ChIP-seq and RNA-seq data sets (see Table S1 and Figure S1A available online). We focused our epigenomic analysis on cis-regulatory regions known to be marked by specific chromatin features: H3K4me1, associated with putative enhancer elements (Heintzman et al., 2007, 2009; Wang et al., 2008); H3K4me3, associated with transcription start sites (TSSs) (Heintzman et al., 2007, 2009; Wang et al., 2008); H3K27ac, enriched at active promoters and enhancers (Creyghton et al., 2010; Rada-Iglesias et al., 2011; Wang et al., 2008; Zentner et al., 2011); H3K27me3, generally associated with transcriptionally repressed regions of chromatin (Rada-Iglesias et al., 2011; Schwartz and Pirrotta, 2007; Zentner et al., 2011); and DNase-seq, indicative of open regions of chromatin (Song et al., 2011).

Through analysis of the RNA-seq and ChIP-seq data, we determined the number of transcripts differentially expressed between mESCs and mEpiSCs, as well as the number of promoters and enhancers that showed chromatin state differences between the two cell types (Figure 1A). The transcriptomes of mESCs and mEpiSCs were remarkably similar to one another ($R^2 = 0.83; n = 4$ biological replicates per cell type), with less than 6% (852 out of 15,198) of expressed transcripts (fragments per kilobase per million mapped reads [FPKM] > 0.25) showing a significant difference in abundance between mESCs and mEpiSCs. Among this list were transcripts of genes known to be mESC specific, including Esrb, Zfp42, Dppa3, and Klf4, as well as those known to be mEpiSC specific, including Fgf5, Cer1, and Lefty1 (Figure 1B; see also Figure 1D and Table S2). Global promoter states were also largely similar, with over 73% of 10,560 active (H3K27ac+) promoters overlapping between the two cell types (Figures 1A and 1B). Even promoters of differentially expressed genes (referred to as mESC-enriched and mEpiSC-enriched) showed a similar state of DNase hypersensitive open chromatin (Figures 1D–1F). In stark contrast, chromatin states at the 22,156 H3K4me1+ H3K27ac+ enhancer loci were much more distinctive, showing only 27% overlap (Figures 1A and 1C). These data suggest that the enhancer landscape may play a dominant role in defining differences between pluripotent cell types on a molecular level.

We next used a computational approach called PreSTIGEouse to assign enhancer elements to each pluripotency-enriched gene using a computational approach called PreSTIGEouse (see Experimental Procedures). Of the 602 pluripotency-enriched genes, 97% showed evidence of differential enhancer usage between the two pluripotent cell types. An example is shown in Figure 2C. Here, the five enhancer elements predicted to regulate Kdm5b, an exemplar pluripotency-enriched gene, are highlighted in blue (Figures 2C and 2D). All five enhancers are located in open chromatin and contain high levels of H3K4me1 and H3K27ac (Figure 2C). In mEpiSCs these five enhancers are virtually devoid of H3K4me1 and H3K27ac, and a different enhancer (highlighted in red) is predicted to regulate Kdm5b.

We assigned enhancer elements to each pluripotency-enriched gene using a computational approach called PreSTIGEouse (see Experimental Procedures). Of the 602 pluripotency-enriched genes, 97% showed evidence of differential enhancer usage between the two pluripotent cell types. An example is shown in Figure 2C. Here, the five enhancer elements predicted to regulate Kdm5b, an exemplar pluripotency-enriched gene, are highlighted in blue (Figures 2C and 2D). All five enhancers are located in open chromatin and contain high levels of H3K4me1 and H3K27ac (Figure 2C). In mEpiSCs these five enhancers are virtually devoid of H3K4me1 and H3K27ac, and a different enhancer (highlighted in red) is predicted to regulate Kdm5b.

Pluripotency-Enriched Genes Show Dramatic Enhancer Differences between Pluripotent States
To further explore the global enhancer differences between mESCs and mEpiSCs, we asked whether genes expressed at similar levels between mESCs and mEpiSCs, including many pluripotency-related factors, are regulated by distinct enhancers. It is known that the Pou5f1, also known as Oct3/4, locus is controlled by distinct enhancers in the preimplantation inner cell mass versus the postimplantation epiblast in vivo as well as in mESCs versus mEpiSCs in vitro (Tesar et al., 2007; Yeom et al., 1996). However, it was not known whether this represents a global regulatory phenomenon. To test this hypothesis, we first defined a set of “pluripotency-enriched” genes using RNA-seq data sets from mESC, mEpiSCs, and a panel of 18 developmental and adult mouse cell and tissue types. We set stringent metrics to ensure that genes within this class are expressed at similar levels in mESCs and mEpiSCs (p > 0.05 and 2-fold change between mESCs and mEpiSCs) and enriched in the two pluripotent cell types in comparison to a panel of 18 different mouse tissues (Figure 2A and Table S3). As expected, this pluripotency-enriched gene class contained gene ontology terms consistent with pluripotency phenotypes (Figure S1B) and the genes share a nearly identical active chromatin state at their promoters (Figure 2B). Consistent with their role in regulating key cell identity genes, 25% of these pluripotency-enriched genes were associated with super enhancers in either of the two cell types (Figure S1C). However, the locations of these super enhancers in the two cell types were largely nonoverlapping, indicating a distinct control mechanism even for these similarly expressed pluripotency-enriched genes.

We assigned enhancer elements to each pluripotency-enriched gene using a computational approach called PreSTIGEouse (see Experimental Procedures). Of the 602 pluripotency-enriched genes, 97% showed evidence of differential enhancer usage between the two pluripotent cell types. An example is shown in Figure 2C. Here, the five enhancer elements predicted to regulate Kdm5b, an exemplar pluripotency-enriched gene, are highlighted in blue (Figures 2C and 2D). All five enhancers are located in open chromatin and contain high levels of H3K4me1 and H3K27ac (Figure 2C). In mEpiSCs these five enhancers are virtually devoid of H3K4me1 and H3K27ac, and a different enhancer (highlighted in red) is predicted to regulate Kdm5b.

We next selected all enhancers that, like those associated with Kdm5b in mESCs, were predicted to regulate pluripotency-enriched genes exclusively in mESCs. These enhancers, which we call “naive-dominant,” generally contained high levels of H3K4me1 and H3K27ac in mESCs relative to mEpiSCs, where enhancer-histone signals were low or near background levels (Figures 2E, 2F, and S2A). Thus, most naive-dominant
Figure 1. Enhancer Profiles Distinguish Mouse Pluripotent States

(A) Venn diagrams showing the number and overlap of expressed transcripts (left), H3K27ac+ promoters of expressed genes (center), and H3K27ac+ enhancers (right) detected in mESCs and mEpiSCs.

(B) Heatmap of expression differences between mESCs and mEpiSCs (log2 transformed, [average of mESC replicates]/[average of mEpiSC replicates]) ranked from high to low (left). Known mESC- and mEpiSC-enriched genes are listed to the left. Windowed chromatin heatmaps comparing DNase HS, H3K4me3, H3K27ac, and H3K27me3 profiles ± 5 kb of promoters in mESCs and mEpiSCs are ranked in the same order as expression data (right).

(C) Windowed heatmaps contrasting DNase HS, H3K4me1, H3K27ac, and H3K27me3 signal ± 5 kb from the midpoint of DNase-centered putative enhancers identified in mESCs or mEpiSC. Enhancers are ranked from most to least mESC-specific (compared to mEpiSCs) based on H3K4me1 peak intensities.

(legend continued on next page)
enhancers are largely inactivated, or “lost” upon transition of mESCs to the mEpiSC state. We next selected “primed-dominant” enhancers, or those predicted to regulate the pluripotency-enriched genes exclusively in mEpiSCs. As expected, these enhancers contained high levels of the active enhancer marks H3K4me1 and H3K27ac in mEpiSCs (Figures 2E, 2F, and S2B). However, unexpectedly, these same enhancers were enriched for H3K4me1 and H3K27ac in mESCs, albeit at lower levels than in mEpiSCs (Figures 2E, 2F, and S2B). These enhancers were not assigned to gene targets in mESCs due to the relatively low levels of the associated enhancer histone marks. As a result of their apparent switch from dormancy in the naive state to active transcriptional regulation in the primed state, we refer to these as “seed enhancers.” We compared seed enhancers to a class of “shared enhancers,” i.e., enhancers predicted to regulate pluripotency-enriched genes in both cell types. Despite their relatively low signal intensity in mESCs, seed enhancers contain the highest levels of H3K4me1 and H3K27ac among all mEpiSC enhancers of pluripotency-enriched genes (Figure S2C), suggesting the importance of the seed enhancer in the active regulation of the pluripotency-enriched genes.

To test whether the chromatin signal at seed enhancers in mESCs could be an artifact caused by metastable heterogeneity within mESC cultures, we tested if the seed enhancers were present in mESCs grown under defined “2i” culture conditions, which are thought to promote a more homogeneous population of naive cells (Ying et al., 2008). We find that, similar to mESCs grown in standard conditions, nearly all seed enhancer loci are marked by H3K4me1 in mESCs grown in 2i conditions and that half of these sites are also marked with H3K27ac, while the other half are marked with H3K27me3 (Figures 2E and S2B). Additionally, the levels of Pecam1 expression can distinguish mESCs in culture that are more naive-like from cells that are more epiblast-like (Furusawa et al., 2004; Hayashi et al., 2008). We used FACS to separate mESCs into Pecam1-high and Pecam1-low populations (Figure S2D), followed by immediate fixation and ChIP-seq. The seed enhancers were found to be present at nearly identical levels in both populations (Figure S2E), further supporting the idea that the presence of seed enhancers in mESCs is not the result of a contaminating epiblast-like population. Using publicly available chromatin interaction maps generated through Hi-C experiments, we validated the fact that compared to naive-dominant enhancers, seed enhancers rarely physically interact with the promoters of pluripotency-enriched genes in mESCs (Figure 2G). Additionally, compared to naive-dominant enhancers, seed enhancers are infrequently occupied by components of the mediator-cohesin complex (Med12, Nipbl, and Smc1) in mESCs, which are known to physically link enhancers with their target promoters (Figure 2H) (Kagey et al., 2010).

Seed enhancers show increased sequence conservation and are generally located farther from the TSSS of genes they control than naive-dominant enhancers are (Figures S2F–S2H). Motif analysis of the two classes revealed that many of the transcription factors enriched at seed enhancers overlap with those found at naive-dominant enhancers (Table S3). It is interesting that naive-dominant enhancers exclusively are enriched for both known naive-specific factors, such as Esrrb and Tcfcp21, and pluripotency factors, such as Oct4, Nanog, and Sox2. Additionally, naive-dominant enhancers are enriched for motifs of Smad2/3 and Smad4, key downstream mediators of the Activin/Nodal pathway required for mEpiSC maintenance (Brons et al., 2007; Nomura and Li, 1998; Tesar et al., 2007; Weinstein et al., 1998). Collectively, these results suggest a mechanism by which Activin/Nodal signaling may be required for repression of naive-dominant enhancers during the transition to primed pluripotency.

We asked if the control of pluripotency-enriched genes by seed enhancers was a mouse specific molecular feature or if it extended more broadly to other mammalian pluripotent cells such as hESCs. Although embryo-derived hESCs exist in a primed pluripotent state similar to mEpiSCs, recent work has shown that hESCs can be converted into a naive-like state using extrinsic factors (Chan et al., 2013; Gafni et al., 2013; Ware et al., 2014). We took advantage of available enhancer-histone modification ChIP-seq data from this model system to test if the dynamic enhancer changes observed in mESCs and mEpiSCs are recapitulated in human pluripotent stem cells. To do this, a set of human pluripotency-enriched genes were selected and assigned to enhancers using methods similar to those used in mouse. Similar to our observations in mouse cells, these human cells contained robust naive-dominant enhancers that lack chromatin markers of enhancer activity in the primed state, as well as seed-like enhancers with more robust activity in the primed state than the naive cell state (Figures S3A–S3I). These data suggest that an early developmental transition from a naive pluripotency-reinforcing epigenomic state to a primed somatic differentiation-capable state may be a general phenomenon of mammalian development.

Seed Enhancers Are Utilized in Downstream Tissues Where They Expand into Enhancer Clusters

Collectively, our findings suggest that the global transcriptional control of pluripotency genes is quite distinct between the naive and primed phases of pluripotency. While genes have been shown to be controlled by distinct enhancers in completely different cell types (Kieffer-Kwon et al., 2013), we found it puzzling as to why genes expressed at virtually identical levels in successive cell stages would undergo enhancer switching. This question prompted us to investigate additional features of seed enhancers. Given that seed enhancers show greater sequence conservation than typical enhancers, we asked if they might play a role at later stages of development. Using available H3K27ac ChIP-seq data from 15 different mouse embryonic and adult tissues, we found that 21% of seed enhancers were significantly enriched for H3K27ac in at least one tissue (Figures 3A and 3B). The rate of seed enhancer usage in downstream tissues was nearly double that of naive-dominant enhancers, (D) Row normalized expression of transcripts differentially expressed (also denoted as enriched) between mESCs and mEpiSCs, ranked as in (B).

(E) Aggregate plots depicting average ChIP-seq and DNase-seq signals at promoters and enhancers of mESC-enriched genes in mESCs and mEpiSCs.

(F) Same as (E), but for mEpiSC-enriched genes.

See also Figure S1, Table S1, and Table S2.
shared enhancers, and a control set of mouse embryonic fibroblast enhancers (indicative of the background rate of enhancer usage in multiple tissues) (Figure 3C). This trend held true across all 15 tissues (Figure 3D).

Upon visual inspection of the seed enhancers in the downstream tissues, we noticed that several were contained within large domains of chromatin broadly marked with H3K4me1. For example, the Cct3 gene is a pluripotency-enriched gene regulated by two seed enhancers in mEpiSCs (Figure 3E) that is also expressed in embryonic brain (Figure 3F). Both seed enhancers become components of a large enhancer domain in embryonic brain. Domains like these are likely composed of multiple individual enhancer elements and are reminiscent of super and stretch enhancers. To test the significance of these observations, we determined the number of seed enhancers that lie within an enhancer cluster (defined as four or more enhancers within 100 kb) in a downstream tissue (Figure 3G). Seed enhancers were more likely than naive-dominant enhancers to lie within enhancer clusters in four different tissues: embryonic brain, cortex, olfactory bulb, and embryonic heart (Figure 3G). By comparison, naive-dominant enhancers were not significantly enriched in enhancer clusters of any downstream tissue tested.

Given the propensity for seed enhancers to lie within neural-related enhancer clusters, we identified how often a seed enhancer gives rise to an enhancer cluster in at least one of the four neural tissues in our panel (embryonic brain, cortex, olfactory bulb, and cerebellum) and found that 29% of all seed enhancers fall within an enhancer cluster in one of these neural tissues (Figure 3H). We found the same enrichment for seed enhancers in regions that become super enhancers in these four neural tissues (Figure 3I). Interestingly, given the tendency for seed enhancers to contribute to super enhancers or enhancer clusters in neural tissues, seed enhancers that resolve in any downstream tissue are enriched for motifs associated with neural lineage transcription factors, including Ascl2, Sox6, Olig2, and Neurod1 (Figure 3J and Table S3).

**DISCUSSION**

Here we compared the transcriptomes and epigenomic landscapes of naive mESCs and primed mEpiSCs. We report two unexpected observations. First, pluripotency-enriched genes shared between mESCs and mEpiSCs are regulated by distinct enhancer elements in each cell type. Thus, the known differential enhancer regulation of the Oct3/4 locus between naive and primed states may represent a well-studied example of a more general mechanism of enhancer switching at other pluripotency genes. Second, we show that enhancers actively regulating pluripotency-enriched genes in mEpiSCs often exist as dormant seed enhancers in mESCs. The seed enhancers are not in physical contact with the promoters of the pluripotency-enriched genes in mESCs, but they appear to take up primary transcripational control in the sequential mEpiSC state. Seed enhancers appear to be epigenetically and functionally distinct from previously described “poised” and “latent” enhancers. Poised enhancers are comarkered with H3K4me1 and H3K27me3 in mESCs, whereas many seed enhancers lack H3K27me3 in mESCs (Rada-Iglesias et al., 2011; Zentner et al., 2011). Additionally, poised enhancers are not active in mESCs, nor do they regulate genes that are transcriptionally active in pluripotent cells. In contrast to seed enhancers that take over control of already expressed genes, poised enhancers regulate genes that become active upon differentiation. Latent enhancers altogether lack the signature enhancer-histone marks and acquire them only upon response to cell stimulation (Ostuni et al., 2013). By comparison, seed enhancers are present but not engaged in mESCs, and they persist through the mEpiSC state into more terminally differentiated cell types.

This raises the questions of why seed enhancers exist and what their precise function is. While further studies are clearly required, one possibility is that seed enhancers function in mESCs to ensure transcriptional robustness of the associated target gene. This hypothesis has been proposed for shadow enhancers in *Drosophila melanogaster* (Hong et al., 2008; Perry et al., 2010). However, if seed enhancers ensure proper gene expression, we might expect them to be in physical contact with their target gene promoters in mESCs, which was clearly not the case. A more attractive hypothesis is that seed enhancers act as “placeholders” in the naive mESC state to ensure that the proper enhancer is utilized in the primed pluripotent state and in subsequent stages of development. It will be particularly

---

**Figure 2. Pluripotency-Enriched Genes Show Dramatic Enhancer Differences between Pluripotent States**

(A) Heatmap depicting row normalized expression of pluripotency-enriched genes. These genes are not differentially expressed (p > 0.05), have <2-fold change between mESCs and mEpiSCs, and are relatively specific to mESCs and mEpiSCs compared to downstream tissues (see Experimental Procedures).

(B) Aggregate plots of DNase hypersensitivity and histone marks at the promoters of pluripotency-enriched genes (depicted in A) in mESCs (left) and in mEpiSCs (right).

(C) UCSC Browser image depicting the Kdm5b locus and the naive-dominant enhancers (highlighted in blue) predicted to regulate its expression in mESCs and the seed enhancers (highlighted in red) predicted to regulate its expression in mEpiSCs using the PreSTIGEous methodology (see Experimental Procedures). Gray boxes identify the Kdm5b promoter. The Rabif gene was not predicted to be regulated by these enhancers.

(D) Expression levels (mean ± SD) of pluripotency-enriched transcript Kdm5b (top) and nonpluripotency-enriched transcript Rabif (bottom) in indicated cell types.

(E) Aggregate plots of enhancers associated with the pluripotency-enriched genes. Naive-dominant enhancers are predicted to regulate expression of pluripotency-enriched genes in the mESC state, but not in the mEpiSC state. Seed enhancers are predicted to regulate pluripotency-enriched genes in mEpiSCs, but not in mESCs. mESC genes grow in standard conditions and 2i conditions are shown.

(F) Boxplots depicting differences in the levels of enhancer histone marks between the two cell types as measured by RPKM (reads per kilobase per million mapped reads) at naive-dominant enhancers (top) and seed enhancers (bottom) (paired sample Wilcoxon signed rank test, *p < 0.0001).

(G) Percentage of pluripotency-enriched genes associated with naive-dominant enhancers and seed enhancers in mESC Hi-C data sets (Fisher’s Exact Test, *p < 0.002). See also Figures S2 and S3.
Pluripotent Stem Cell States Reveal Seed Enhancers

A. Seed enhancer resolution

B. Activity by tissue

C. Percent of enhancers active in any downstream tissue

D. Percent of enhancers active in downstream tissue

E. Scale chr3:

F. Non-predicted target: Rhgb

G. Predicted target: Cct3

H. Percent resolved in enhancer cluster

I. Percent resolved in super enhancers

J. Gene expression levels

(legend on next page)
interesting to test whether seed enhancers are a phenomenon specific to the unique preimplantation to postimplantation transition of placental mammals or are something more general within developmental and stem cell hierarchies.

**EXPERIMENTAL PROCEDURES**

**Cell Culture**

Pluripotent stem cells were cultured as previously described (Chenoweth and Tesar, 2010; Najm et al., 2011; Tesar et al., 2007). For details see Supplemental Experimental Procedures.

**Sequencing Experiments**

ChIP and DNase sequencing experiments were performed as previously described (Schmidt et al., 2009; Song et al., 2011). RNA sequencing libraries were prepared using the Illumina TruSeq RNA Sample Preparation Kit according to the manufacturer’s protocol. For details see Supplemental Experimental Procedures.

**Data Sets**

Data sets used in this study are summarized in Table S1 (Buecker et al., 2014; ENCODE Project Consortium, 2011; Gafni et al., 2013; Kagey et al., 2010; Marks et al., 2012; Selvaraj et al., 2013; Shen et al., 2012; Wamstad et al., 2012).

**Predicting Gene Targets of Enhancers with PreSTIGEose**

The PreSTIGE (Predicting Specific Tissue Interactions of Genes and Enhancers) bioinformatics algorithm was utilized to associate enhancer/gene pairs (Corradin et al., 2014). PreSTIGE predicts enhancer-gene pairs based on genomic proximity (<100 kb) and shared specificity of enhancer H3K4me1 signal and gene expression as compared to a panel of 12 tissues. PreSTIGE was initially developed to infer human enhancer-gene pairs and was adapted for application to mouse (PreSTIGEose). Predictions are available at http://genetics.case.edu/prestige/. For details see Supplemental Experimental Procedures.

**ACCESSION NUMBERS**

mEpiSC ChIP-seq, mEpiSC DNase-seq, and mESC and mEpiSC RNA-seq data sets generated for this publication are available on the NCBI Gene Expression Omnibus website in GEO Series GSE57409.

**SUPPLEMENTAL INFORMATION**

Supplemental Information for this article includes three figures, three tables, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.stem.2014.05.005.

**AUTHOR CONTRIBUTIONS**

D.C.F., O.C., P.C.S., and P.J.T. designed the study, analyzed and interpreted all data, and wrote the manuscript. D.C.F., G.E.Z., J.G.C., R.D.M., and P.J.T. cultured cells and performed all ChIP-seq and RNA-seq experiments. L.S. and G.E.C. performed DNase-seq experiments. O.C., A.S., and P.C.S. developed and implemented the PreSTIGEose software package. All authors edited and approved the final manuscript.

**ACKNOWLEDGMENTS**

This research was supported by funding from the NIH (R01HD056369 and R01CA160356), the New York Stem Cell Foundation, the Mt. Sinai Health Care Foundation, the Genomics and Cytometry core facilities of the Case Comprehensive Cancer Center (P30CA043703), and CWRU Cellular and Molecular Biology training grant T32GM008056 (O.C.). P.J.T. is a New York Stem Cell Foundation Robertson Investigator. We are grateful to Thomas LaFranchise, Cindy Bartels, Simone Edelheit, and Neil Molyneaux for technical assistance, James Morrow and Tyler Miller for discussion, and Joanna Wysocka and Jacob Hanna for sharing data.

Received: March 3, 2014
Revised: May 1, 2014
Accepted: May 13, 2014
Published: June 5, 2014

**REFERENCES**


**Figure 3. Seed Enhancers Are Utilized in Downstream Tissues Where They Expand into Enhancer Clusters**

(A) Pie chart depicting the fraction of seed enhancers that show a chromatin state indicative of enhancer activity (H3K27ac+) in downstream tissues (red). (B) Heatmap displaying the downstream tissues in which each seed enhancer is active. Each row represents a seed enhancer. For details see Supplemental Experimental Procedures. (C) Percentage of seed enhancers (red) active in at least one downstream tissue compared to naive-dominant enhancers (blue), MEF enhancers (gray), and enhancers shared between mESCs and mEpiSCs (white). (D) Percent of seed enhancers active in each downstream tissue (red) compared to naive-dominant enhancers (blue), MEFS enhancers (gray), and enhancers common to mESCs and mEpiSCs (white). (E) Genome browser image depicting the Cct3 gene and the enhancers predicted to regulate its transcription in mESCs and mEpiSCs. Gray boxes demarcate active promoters, while blue boxes identify naive-dominant enhancers. The red boxes highlight two seed enhancers, which become components of a super enhancer in embryonic brain (black box), but not bone marrow. (F) Expression (mean ± SD) of Cct3 (right) is high in pluripotent cells and embryonic brain, but low in bone marrow. None of the enhancers in this region are predicted to target neighboring gene Rhbg (left). (G) Percent of enhancers located in a cluster of enhancers (defined as four or more active enhancer elements within a 100 kb window) in a downstream tissue. Seed enhancers (red) are significantly more likely to occur in enhancer clusters than naive-dominant enhancers (blue) in embryonic brain, cortex, olfactory bulb, and embryonic heart (Fisher’s Exact Test, *p < 0.003). The background rate for all mESC and mEpiSC enhancers (gray) is included for comparison. (H) Percentage of seed enhancers (red), naive-dominant enhancers (blue), and all enhancers (gray) within a region that becomes an enhancer cluster (defined as in G) in one of the four neural tissues in the panel (embryonic brain, cortex, olfactory bulb, and cerebellum; Fisher’s Exact Test, *p < 0.003). (I) As in (H), but for regions that are super enhancers in neural tissues. (J) Motifs enriched among seed enhancers that are active in downstream tissues. See also Table S3.

Supplemental Information for this article includes three figures, three tables, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.stem.2014.05.005.


