

The Transcriptional and Functional Properties of Mouse Epiblast Stem Cells Resemble the Anterior Primitive Streak

Yoji Kojima,^{1,6} Keren Kaufman-Francis,¹ Joshua B. Studdert,¹ Kirsten A. Steiner,¹ Melinda D. Power,¹ David A.F. Loebel,^{1,5} Vanessa Jones,¹ Angelyn Hor,¹ Gustavo de Alencastro,² Grant J. Logan,² Erdahl T. Teber,³ Oliver H. Tam,¹ Michael D. Stutz,⁴ Ian E. Alexander,^{2,5} Hilda A. Pickett,^{4,5} and Patrick P.L. Tam^{1,5,*}

¹Embryology Unit

²Gene Therapy Research Unit

³Bioinformatics Group

⁴Cancer Research Unit

Children's Medical Research Institute, Westmead NSW 2145, Australia

⁵Sydney Medical School, University of Sydney, NSW 2006, Australia

⁶Present address: Institute for Integrated Cell-Material Sciences, Kyoto University, Kyoto 606-8501, Japan

*Correspondence: ptam@cmri.org.au

<http://dx.doi.org/10.1016/j.stem.2013.09.014>

SUMMARY

Mouse epiblast stem cells (EpiSCs) can be derived from a wide range of developmental stages. To characterize and compare EpiSCs with different origins, we derived a series of EpiSC lines from pregastrula stage to late-bud-stage mouse embryos. We found that the transcriptomes of these cells are hierarchically distinct from those of the embryonic stem cells, induced pluripotent stem cells (iPSCs), and epiblast/ectoderm. The EpiSCs display globally similar gene expression profiles irrespective of the original developmental stage of the source tissue. They are developmentally similar to the ectoderm of the late-gastrula-stage embryo and behave like anterior primitive streak cells when differentiated in vitro and in vivo. The EpiSC lines that we derived can also be categorized based on a correlation between gene expression signature and predisposition to differentiate into particular germ-layer derivatives. Our findings therefore highlight distinct identifying characteristics of EpiSCs and provide a foundation for further examination of EpiSC properties and potential.

INTRODUCTION

Epiblast stem cells (EpiSCs) are self-renewing multipotent stem cells that can be derived from the epiblast of postimplantation mouse embryos (Brons et al., 2007; Tesar et al., 2007). In contrast to mouse embryonic stem cells (mESCs), EpiSCs are inefficient in contributing to the tissue of chimeras following blastocyst injection, show less active expression of “naive” pluripotency-related genes, and exhibit increased Oct4 occupancy of the proximal enhancer of *Pou5f1* (Tesar et al., 2007). Like human ESCs, EpiSCs are dependent on FGF/ERK and Activin(Nodal)/

Smad signaling activities for derivation and maintenance (Vallier et al., 2009a, 2009b, 2009c).

EpiSCs derived from E5.5 to E6.5 epiblast display the gene expression characteristics of gastrula-stage epiblast/ectoderm (epi/ect) (Brons et al., 2007; Hayashi et al., 2011; Iwafuchi-Doi et al., 2012; Tesar et al., 2007). EpiSCs show monoallelic expression of imprinted genes (Sun et al., 2012) and display X-inactivation in XX lines (Bao et al., 2009; Guo et al., 2009; Hayashi et al., 2008). The presence of histone methylation marks in the *Stella* (*Dppa3*), *Otx2*, *Rex1*, and *Fbxo15* loci (Bao et al., 2009; Hayashi and Surani, 2009; Tesar et al., 2007) further suggests that EpiSCs are distinct from mESCs. EpiSCs are multipotent and can generate a multitude of germ-layer-derived tissues in teratomas and during in vitro differentiation. Although there are variations in the neural potential among EpiSC lines (Bernemann et al., 2011), they readily differentiate into neuronal lineages after removal of supporting growth factors (Iwafuchi-Doi et al., 2012; Najm et al., 2011b). When grafted into gastrula-stage mouse embryo, EpiSCs were incorporated into the derivatives of all three germ layers (Huang et al., 2012), revealing that the EpiSCs may be the in vitro equivalent of gastrula-stage epiblast cells. However, not all cells display similar lineage potential. A subset of cells (0.1%–0.5% of the colony) exhibit occupancy of the distal enhancer of *Pou5f1* gene that is characteristic of the naive pluripotency state and are able to contribute to chimeras like the ESCs (Han et al., 2010). Also, some cells in the EpiSC colonies express *Stella* and are competent to produce primordial germ cells in vitro (Hayashi et al., 2008; Hayashi and Surani, 2009). These cells therefore may retain the potency of germ cell formation and are functionally similar to the epiblast at (or before) E6.25–E6.5 in vivo when germ cells are specified (Ohinata et al., 2009).

Recently, it has been shown that EpiSCs can be derived from mouse embryos of a wide range of developmental stages from E3.5 blastocysts (Najm et al., 2011a) and from epi/ect of E6.5–E8.0 presomite-stage embryos (Osorno et al., 2012). It is not clear if EpiSCs derived from these cellular sources display similar or different states of pluripotency and lineage characteristics. In this study, we examined the molecular properties and the differentiation potential of EpiSCs generated from embryos

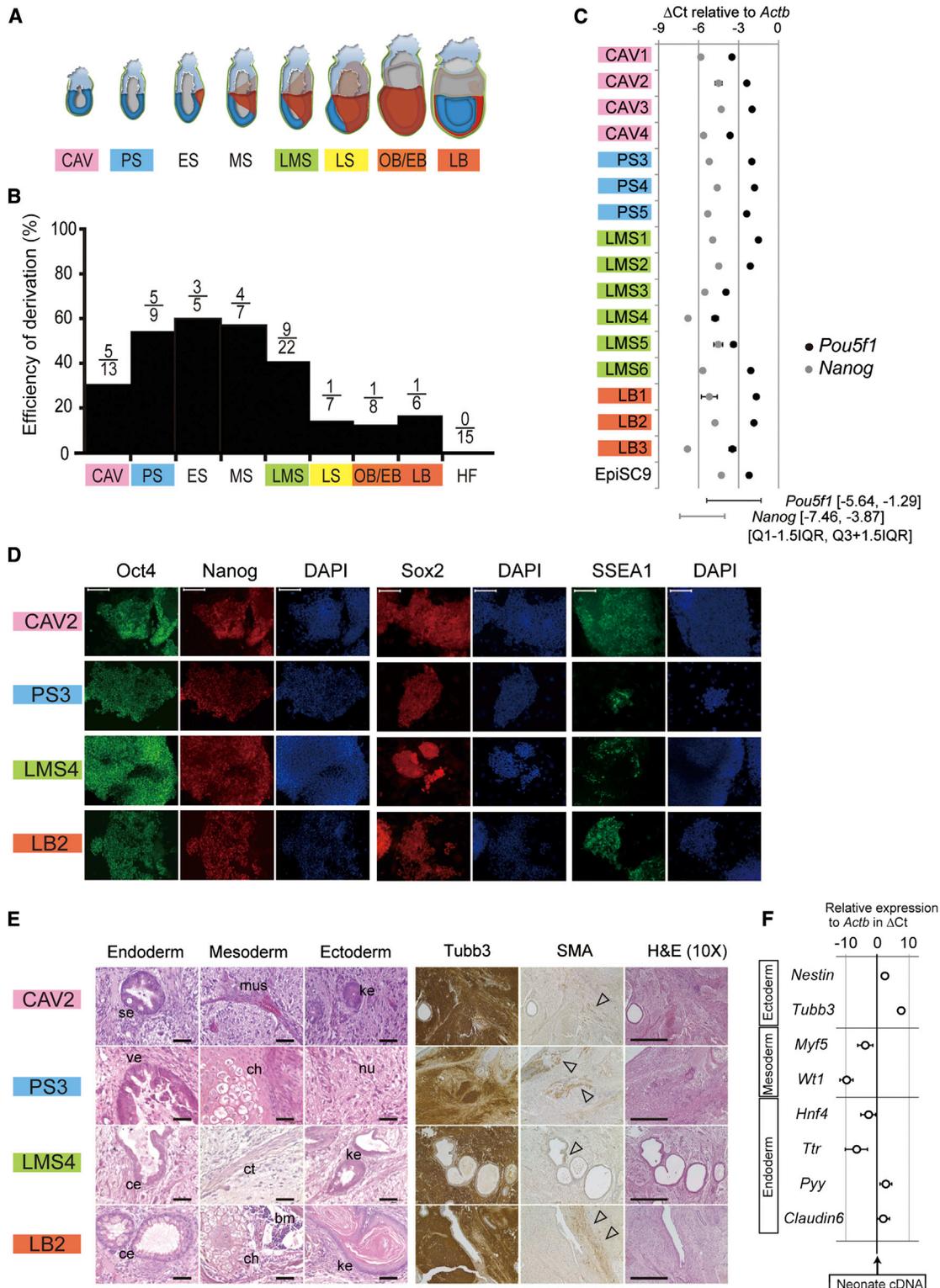


Figure 1. Derivation of EpiSCs from Mouse Embryos at Immediate Postimplantation Stages of Development

(A) Staging of embryos by the morphology of the epiblast and proamniotic cavity before gastrulation, the shape of the epiblast/ectoderm (epi/ect, blue), formation of the mesoderm, the primitive streak, and the allantois (red), and the development of the extraembryonic mesoderm and the chorion (gray). See [Supplemental Information](#) for staging criteria. CAV, cavity; PS, prestreak; ES, early-streak; MS, midstreak; LMS, late midstreak; LS, late streak; OB/EB, no bud/early bud; LB, late bud; HF, head fold.

(B) The efficiency of EpiSC derivation from epi/ect of CAV- to HF-stage embryos. Fraction above the bar = number of established lines/number of explants.

(legend continued on next page)

of a range of developmental stages before and after the onset of gastrulation to test the hypothesis that EpiSCs from different developmental stages might retain the characteristics of their original cellular source. Interestingly, we found that EpiSCs derived from epi/ect of different developmental stages display similar transcriptome characteristics that are most closely related to the late-gastrulation-stage ectoderm. EpiSCs also display the properties of the anterior primitive streak. Furthermore, specific EpiSC lines have an inherently different propensity for germ-layer differentiation, which could be related to the expression of a signature gene set. Our study has therefore pinpointed the ectoderm cells of late-gastrulation-stage embryo as the *in vivo* developmental counterparts of EpiSCs and identified salient features of the cellular and molecular phenotypes of EpiSCs that may correlate with their germ-layer differentiation potential and their resemblance to the anterior primitive streak.

RESULTS

EpiSCs Display Stable Stem Cell Characteristics

EpiSCs were generated using 77 explants of epi/ect of postimplantation mouse embryos at developmental stages between pregastrulation (E5.5) to late-bud (E8.25) stage (Figure 1A; see Supplemental Experimental Procedures, Staging of Embryos available online; Downs and Davies, 1993; Rivera-Pérez et al., 2010). Twenty-nine lines were established (Figure 1B), giving an overall efficiency rate of 37.7% (similar to that reported for pre- to early-streak-stage epiblast: Brons et al., 2007; Tesar et al., 2007). The efficiency was higher (40%–60%) for the epi/ect of pre- to midgastrula stages (cavity to late mid streak) and lower (10%–20%) for late-gastrula stages (late streak to late bud). No stable lines were established from ectoderm of embryos at the head fold stage or later.

The expression of pluripotency factors (*Pou5f1* and *Nanog*; Figure 1C; *Oct4*, *Nanog*, *Sox2*, and *SSEA-1*; Figure 1D) did not vary significantly among EpiSC lines. All EpiSC lines generated teratomas containing a multitude of germ-layer-derived tissues (Figure 1E). In contrast to the teratomas generated from ESCs, EpiSC teratomas contained more abundant β III tubulin-positive neural tissue but less smooth muscle actin positive tissue (Figure 1E). The relative proportion of tissue types found in the EpiSC-derived teratomas was compared to that of whole neonate (day 1) mice (Figure 1F). qPCR analysis of the expression level of tissue-specific gene expression showed that the teratomas contained more

abundant ectoderm tissues (*Nestin* and *Tubb3*), less mesoderm tissues (*Myf5* and *Wt1*), and similar amount of endoderm tissues (*Hnf4*, *Ttr*, *Pyy*, and *Claudin6*) compared with the neonate mice.

Phenotypic stability of EpiSC lines was assessed by telomere length maintenance. The telomere lengths of EpiSC lines (at 12th to 18th passages) were between 28.0 and 41.0 kb pairs (Figure S1A), which were similar to those of MEFs, shorter than mESCs (\approx 50 kb) (Varela et al., 2011) and longer than differentiated tissues (\approx 20 kb) such as lung (Figure S1A), heart, liver, kidney, and spleen (data not shown). We examined if telomere length may be maintained by a combination of trimming and telomerase-mediated lengthening activity. None of the EpiSC lines or the mESCs examined displayed any telomere trimming that would be detected by the presence of T-circles (Pickett et al., 2009). In contrast, T-circles were present in terminally differentiated tissues such as the liver (Figure S1B). Telomerase enzyme activity was detected by TRAP assay in both mESCs and EpiSCs (Figure S1C). However, *mTert* expression, which is tightly linked to telomerase activity, was lower in EpiSCs than ESCs (Figure S1D). Karyotype analysis of 23 EpiSC lines showed that 17 lines (73.9%) had the normal number of 40 chromosomes in over 70% of metaphases (Figure S1E). A comparison of the transcriptome of EpiSCs revealed that the gene expression profile was remarkably consistent among triplicate cultures of individual lines, and the transcriptome of two EpiSC lines derived from different halves of one embryo remained similar to each other after ten passages (Figure S1F). Several EpiSC lines were cryopreserved and re-established in culture as triplicate cultures. The transcriptomes of these sublines were hierarchically separated from the parental lines but were still consistent among the triplicates (Figure S1F). Taken together, these findings highlight that the EpiSCs display stable stem cell properties and are multipotent for germ-layer differentiation.

EpiSCs Derived from Epiblast/Ectoderm of Different Developmental Stages Display Similar Transcriptomes

The hypothesis underpinning our study is that EpiSCs are different from other pluripotent stem cells and that EpiSCs derived from epi/ect of embryos of different developmental stages are inherently different in their cellular and molecular characteristics, which reflect those of the source tissue. This was tested by analyzing the transcriptomes of the EpiSCs, mESCs, and mouse iPSCs (miPSCs) (Figure 2A). Unsupervised hierarchical clustering of the transcriptomes showed that EpiSCs were clustered in a group different from the mESCs

(C) Expression of *Pou5f1* and *Nanog* relative to *Actb* in Δ Ct determined by qPCR ($n = 3$) for EpiSC lines derived from epi/ect of CAV-, PS-, LMS-, and LB-stage embryos and the EpiSC9 line from Dr. Tesar. Data were expressed as mean \pm SEM (not shown when SEM is less than 0.25). Bars below the graph indicate the range of insignificant variation by outlier analysis [Tukey's method, 25th percentile $-1.5 \times$ interquartile range (IQR) to 75th percentile $+1.5 \times$ (IQR)].

(D) Immunofluorescence analysis of the expression of *Oct4*, *Nanog*, *Sox2*, and *SSEA-1* (counterstained with DAPI) in examples of EpiSCs derived from epi/ect of CAV, PS, LMS, and LB stages. Scale bar, 100 μ m.

(E) Examples of germ-layer derivatives in teratomas generated from EpiSCs of CAV, PS, LMS, and LB stages (histology sections at 40 \times magnification; scale bar, 50 μ m) and immunostaining results for *Tubb3* and smooth muscle actin (SMA), with adjacent H&E-stained sections (at 10 \times magnification; scale bar, 200 μ m); Δ marks tissues with positive SMA staining. bm, bone marrow; ce, ciliated epithelium; ch, chondrocytes; ct, fibrous connective tissues; ecy, epithelial cyst; ke, keratinized epithelium; mus, muscle; nu, neural tissues; se, secretory epithelium; ste, stratified epithelium; ve, villous epithelium.

(F) qPCR analysis for assessing the tissue composition of 36 independent samples of teratomas generated from 12 EpiSC lines (three each of CAV, PS, LMS, and LB stage, and each sample analyzed in duplicates), with fold differences in the amount of tissue-specific transcripts relative to that in cDNA sample of whole day 1 neonate mouse. Data for each gene are shown as mean \pm SEM, normalized against the level of *Actb*.

See also Figure S1.

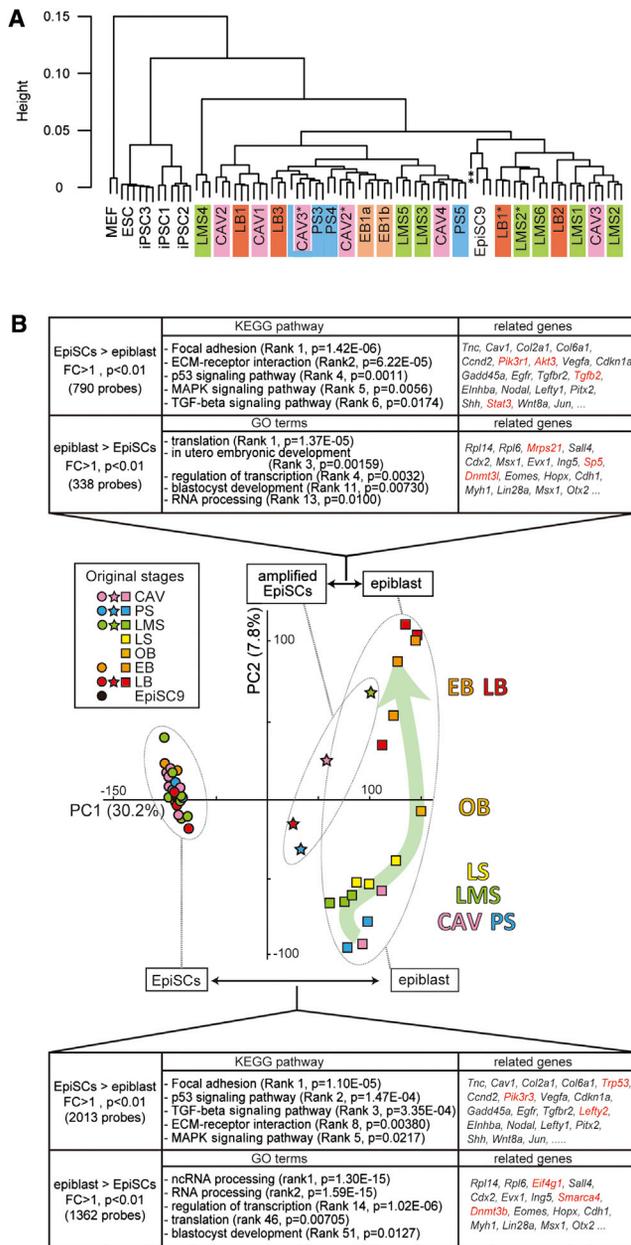


Figure 2. EpiSC Lines Display Similar Transcriptome Characteristics

(A) Unsupervised hierarchical clustering of the transcriptome of EpiSC lines derived from epi/ect of CAV, PS, LMS, EB, and LB stages using data of all annotated probes (illuminaMousev2.db package, n = 31,646) by Pearson distance and complete linkage. Triplicate cultures of each EpiSC line were analyzed. All stem cell lines were different from the MEF. CAV2*, CAV3*, LMS2*, and LB1* are the sublines of CAV2, CAV3, LMS2, and LB1, respectively. **RNA sample prepared from EpiSC9 maintained in the Tesar laboratory and analyzed in parallel with samples of the same line maintained in our laboratory.

(B) PC1-PC2 planes showing clusters of EpiSC lines (four EpiSC lines with amplified RNA sample) and 17 epi/ect samples. The green arrow through the epi/ect cluster indicates the advancing “developmental axis.” The plots are colored according to the stages of source tissue (see color legend). EpiSCs, amplified EpiSCs, and epi/ect are represented in circles, stars, and squares, respectively. Differentially expressed genes between amplified RNA samples of EpiSC lines and epi/ect (top panel) and between nonamplified

and miPSCs. EpiSC lines were not clustered in relation to the original developmental stages but scattered across the hierarchical group. An established EpiSC line (EpiSC9, Najm et al., 2011b) also clustered with our set of EpiSC lines, and it displayed a similar transcriptome when maintained independently in two laboratories (Figure 2A), further attesting the stability of EpiSCs. Our transcriptome data strongly suggest that EpiSCs derived from epi/ect of embryos at different developmental stage acquire similar transcriptome characteristics after they were established in vitro and that there is no evident correlation of the molecular properties of the EpiSCs with the developmental stages of the source tissue.

EpiSCs Display Gene Expression Profiles that Are Similar to the Ectoderm of Late-Gastrula-Stage Embryos

In order to pinpoint the in vivo developmental stage that EpiSCs are most closely related to, the transcriptomes of EpiSCs were compared to those of epi/ect of cavity (CAV), prestreak (PS), late mid streak (LMS), late streak (LS), no bud (OB), early bud (EB), and late-bud (LB)-stage embryos. We first performed a principal component analysis (PCA) and observed that the epi/ect were aligned in a developmental order along the PC2 axis (Figure 2B). In this axis, the CAV- and PS-stage epiblast samples that clustered closely were indistinguishable from each other. CAV and PS epiblast were clustered separately from both the LMS and LS epi/ect. The LMS-LS epi/ect were positioned separately from EB and LB epi/ect on the PC2 axis, with the OB epi/ect in between them (Figure 2B). The alignment of these epi/ect transcriptomes therefore delineates a developmental axis of gastrulation.

Because the epi/ect samples for the microarray analysis were subject to linear amplification, triplicate RNA samples from four EpiSC lines (CAV2, PS5, LMS3, and LB2) sourced from different developmental stages and that clustered separately (Figure 2A) were similarly amplified for comparison. Principal component analysis (PCA) of the whole transcriptome data set revealed that the epi/ect and amplified EpiSCs clustered separately along the PC1 axis (Figure 2B). Relative to epi/ect, the amplified EpiSC samples were enriched for gene transcripts associated with cell adhesion, matrix production, and signaling activity of the transforming growth factor (TGF)-β, MAPK, and WNT pathways but displayed lower expression of genes associated with chromatin modification and transcriptional regulation (Figure 2B). A comparison of the differentially expressed genes between the epi/ect and all unamplified EpiSC samples revealed an expanded list of transcripts associated with the same categories of cellular function, signaling, and transcription regulatory activity as those of the amplified EpiSC samples (Figure 2B). These results showed that the amplification step has not significantly altered the outcome of the transcriptome analysis and that the

EpiSCs and epi/ect (bottom panel) are selected by average fold change (FC) and significant p value with Limma package on R. Overrepresented KEGG pathways and related genes are shown for EpiSCs > epiblasts probes, and GO terms and related genes for epiblasts > EpiSCs probes analyzed with DAVID Bioinformatics Resources 6.7 (<http://david.abcc.ncifcrf.gov>). In the panels, genes with microarray results influenced by RNA amplification are marked in red.

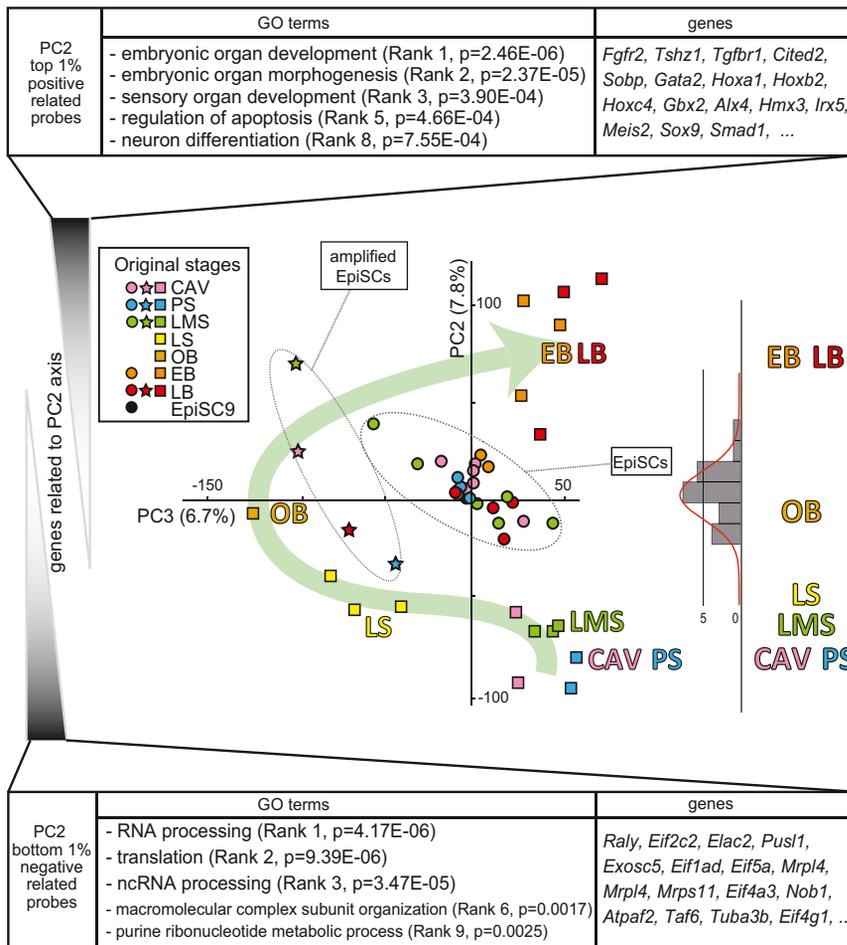


Figure 3. PCA for Delineating the Relevant Developmental Stages of EpiSCs

PC2-PC3 plane displays the range of 25 EpiSC lines along the “developmental axis” (green arrow). Top panel: Probes with top 1% (n = 316) PC2-positive-related loadings. Gene ontology terms and related genes are shown. Bottom panel: Probes with bottom 1% (n = 316) negative-related PC2 loadings are listed. Gene ontology terms and related genes are shown. Histogram presents the number of EpiSC lines distributed along the PC2 axis. Intervals of PC2 score values are in steps of 10, the vertical line equals a score of five EpiSC samples, and the correlation shown with the normal distribution curve (red) (p = 0.815 by Shapiro-Wilk test).

this analysis, the top 1% genes that were differentially expressed in the epi/ect at all developmental stages were selected based on Hotelling T² score (Table S1). The developmental profile of the expression of these genes followed four distinctive stage-related patterns: (1) lineage commitment: progressively increasing expression, (2) neuroectoderm differentiation: gradual followed by elevated expression, (3) pluripotency related: rapid followed by slow decline in expression, and (4) maintenance of epiblast during development: progressively decreasing expression (Figures 4A and S2). Gene ontology (GO) analysis of these genes revealed similar terms and identity to

established EpiSC lines display a profile of gene expression that is globally different from the epi/ect.

In the PC2 axis, the position of the amplified EpiSC lines overlapped with the unamplified EpiSC lines, and they were positioned between LS and EB/LB stages in the developmental axis (Figures 2B). To further evaluate the significance of the PCA presentation of developmental progression, we examined the differentially expressed genes identified in the epi/ect by probe loadings along PC2 axis (Figure 3). In this axis, the bottom 1% negative-related probes (n = 316) that were associated with the epiblast of younger developmental stages were mainly regulators of RNA processing (Figure 3). In contrast, the top 1% positive-related probes that were associated with the epi/ect of the advanced developmental stages were genes related to lineage commitment (Figure 3). The distribution of the epi/ect samples along the PC2 axis therefore correlates with the progression of development. When the transcriptomes of the EpiSC lines (amplified and unamplified samples, including the EpiSC9 line) were mapped to this developmental axis, they were found in a normal distribution (by Shapiro-Wilk test) around the no bud stage (Figure 3). This finding suggests that the EpiSCs may be developmentally equivalent to the ectoderm cells of the late-gastrula-stage embryo.

The developmental changes in the epi/ect transcriptome were further examined using Bioconductor Timecourse package. For

those that signify the developmental axis identified by PCA (Figures 3 and 4B). Heatmaps of the expression profiles (Figure 4A) and the expression levels (Figure S2A) of these developmentally regulated genes in amplified EpiSC samples indicated that they matched most closely with those of the epi/ect between LS and EB/LB stages (Figure 4C). Unsupervised hierarchical clustering of the whole transcriptome also showed that these EpiSC lines clustered closely with OB- to LB-stage ectoderm and were hierarchically separated from CAV and PS epiblast (Figure S2B).

These results therefore point to a strong likelihood that the EpiSCs derived from embryos of different postimplantation stages have converged to a cellular state comparable to the epi/ect of embryos at the late-gastrula stages.

EpiSCs Acquire Anterior Primitive Streak Properties

The finding of the enhanced expression of genes related to FGF/MAPK, WNT, and Activin/GDF/Nodal signaling (Figures 2B and S2A) raised the possibility that EpiSCs might have acquired cellular properties that are characteristic of the primitive streak. To follow up this observation, we mined the EpiSCs and epi/ect transcriptomes for an annotated set of genes that are expressed in the primitive streak of the mouse and the chick embryos (Alev et al., 2010; Pfister et al., 2007). This analysis revealed that many genes of the anterior mesendoderm and definitive endoderm (AME/ADE) (*Lefty1, Cited2, Cer1, Lbh*, and

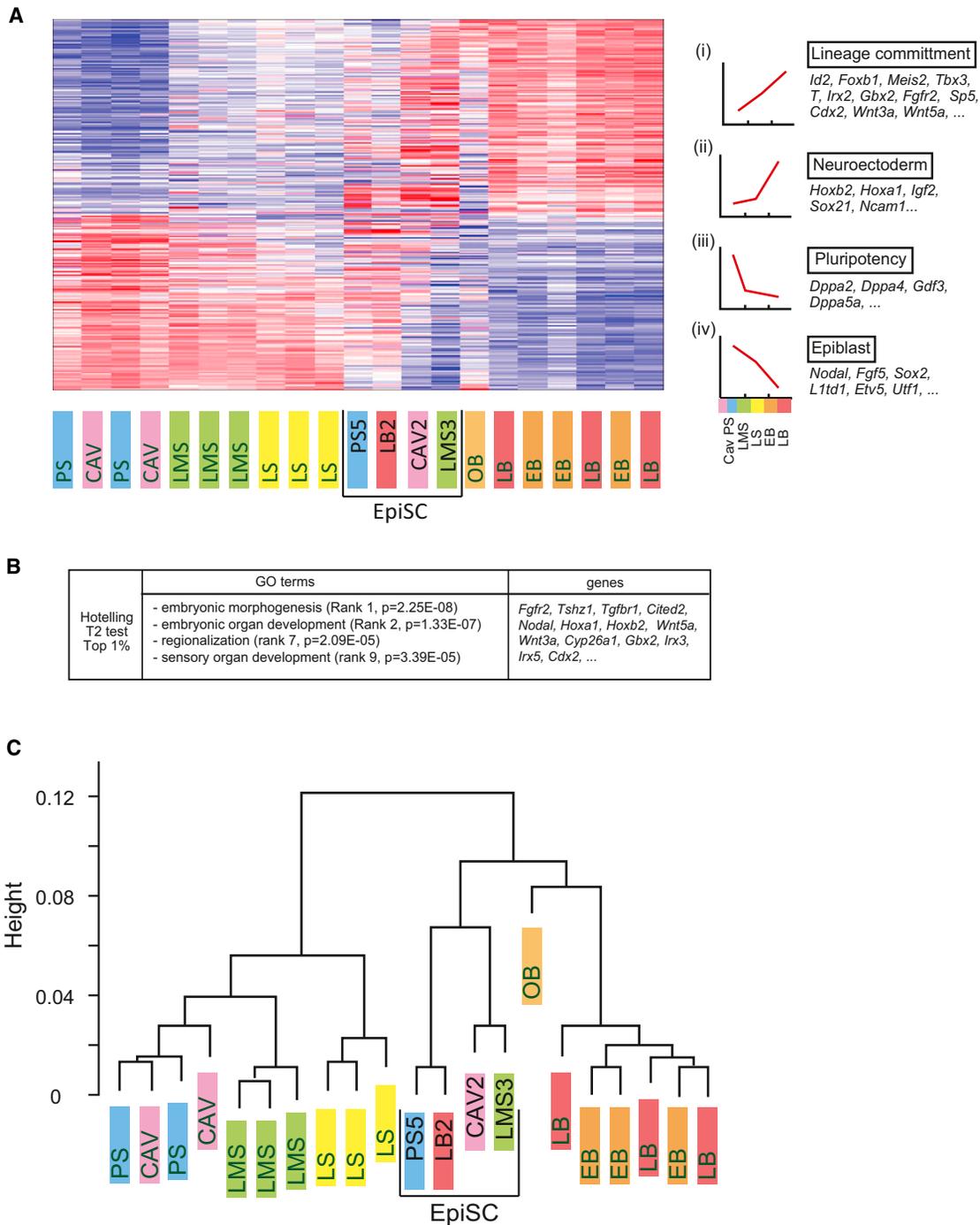


Figure 4. Developmental Gene Expression Profile Relates EpiSCs to Late-Gastrula-Stage Epiblast/Ectoderm

(A) Heatmap of amplified RNA samples generated by GeneAnswers package on R. Top 1% ($n = 316$) differentially expressed probes from Hotelling T^2 analysis were used. Gene expression pattern is classified in four groups: (1) increasing progressively from CAV/PS to LB stage, (2) increasing between LMS to LB, and decreasing between (3) CAV, PS to LMS or (4) LMS to LB. See also Figure S2.

(B) Gene ontology terms and representative genes from Hotelling T^2 analysis (see Figure S2).

(C) Hierarchical clustering of amplified RNA samples from four EpiSC lines and epi/ect of CAV, PS, LMS, LS, OB, EB, and LB stages using data of all annotated probes ($n = 31,646$) by Pearson distance and complete linkage.

See also Figure S2 and Table S1.

Sox17) and anterior primitive streak (*Sfrp1, Foxa2, Chrd, Acvr1b*, and *Fzd8*) of the late-gastrula embryo were expressed higher in EpiSCs than the epi/ect. Genes expressed broadly in the primi-

tive streak such as *T, Wnt8a, Hoxb1*, and *Lefty2* were also expressed in EpiSCs, except *Wnt3a*. The expression of genes associated with the posterior primitive streak, such as *Cdx2*,

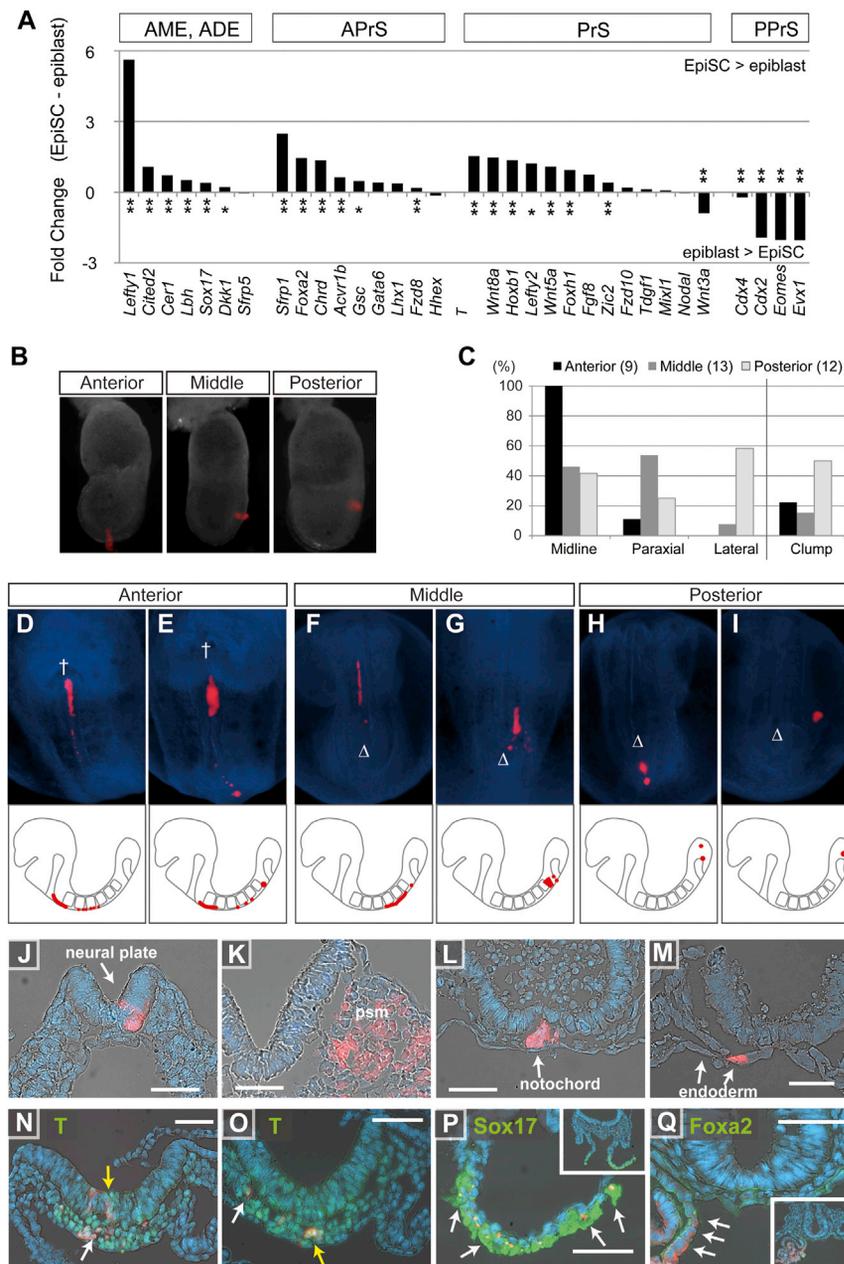


Figure 5. EpiSCs Display Characteristics of Anterior Primitive Streak Cells

(A) Fold change between EpiSCs and epiblasts of the genes expressed in anterior mesendoderm (AME) and anterior definitive endoderm (ADE), anterior primitive streak (APrS), whole primitive streak (PrS), and posterior primitive streak (PPrS). Significant changes at **p < 0.01; *p < 0.05 by t test.

(B) RFP-expressing EpiSCs grafted to anterior-, mid-, and posterior-third segment of the primitive streak of OB/EB-stage embryo.

(C) Percentage of host embryos showing the distribution of graft-derived cells in the midline, paraxial and lateral tissues, and the frequency of clumping of cells. Numbers on the top indicates numbers of embryo grafted to each site in the primitive streak.

(D–I) Representative embryos showing distribution of graft-derived cells 24 hr after transplantation of RFP-expressing EpiSCs to anterior (D and E), middle (F and G), and posterior streak (H and I). Embryos were viewed from the ventral side with the anterior to the top. The symbols † and Δ indicate the anterior and posterior intestinal portal, respectively. Lower panels: The distribution of grafted cells in the host embryos illustrated in (D)–(I).

(J–Q) EpiSC-derived cells (marked by RFP expression) in the neuroepithelium (J), paraxial mesoderm (K), notochord (L), and endoderm (M). Expression of T in EpiSC-derived cells in the primitive streak (yellow arrow) and the nascent mesoderm (white arrow) (N) and the axial mesoderm that forms the notochord (yellow arrow) and paraxial mesoderm (white arrow) (O). EpiSC-derived cells in the endoderm expressed Sox17 (P) and Foxa2 (white arrow) (Q). Insets show low magnification image of the sections. Scale bar, 50 μm.

See also Figure S3 and Table S2.

Cdx4, *Eomes*, and *Evx1*, was significantly lower in EpiSCs (Figure 5A). These results suggest that some cells in the EpiSC lines have acquired the molecular and signaling properties of primitive streak cells, with a bias toward those of the anterior primitive streak.

To test if the EpiSCs behave like the native primitive streak cells of the gastrulating embryo, EpiSCs were assessed for their pattern of tissue colonization following grafting into the midsegment of the primitive streak of OB/EB-stage (E7.5) embryos. Eight EpiSC lines (CAV1, CAV2, CAV3, CAV4, PS3, LMS1, LMS4, and LMS5) were tested. After 24 hr of in vitro development, descendants of the EpiSCs (marked by the lipophilic CM-Dil fluorescent dye) were found to have dispersed from the graft site and incorporated into the germ-layer tissues of the

host embryo (Table S2). To evaluate the ability of the EpiSCs to integrate properly into the host tissues derived from the primitive streak, another series of experiments was performed by grafting EpiSCs expressing the red fluorescent protein (CMV-RFP) to three sites (anterior-, middle- and posterior-third segment) of the primitive streak (Figure 5B). Cells derived from the grafts were distributed along the body axis of the host embryo after 24 hr of culture (Figures 5C and 5D–5I). EpiSCs grafted to anterior and middle sites showed a higher rate of incorporation than the posterior site. EpiSCs from the anterior site dispersed anteriorly and colonized the foregut invagination and the midline mesendoderm (Figures 5D and 5E). Cells grafted into the middle segment of the primitive streak populated the paraxial and lateral mesoderm, extending from midtrunk to the level of the posterior intestinal portal (Figures 5F and 5G). In contrast, EpiSCs grafted to the posterior site tended to clump (Figures S3A and S3B). This clumping phenotype may reflect the lower compatibility of the EpiSCs with the posterior primitive streak. However, for EpiSCs

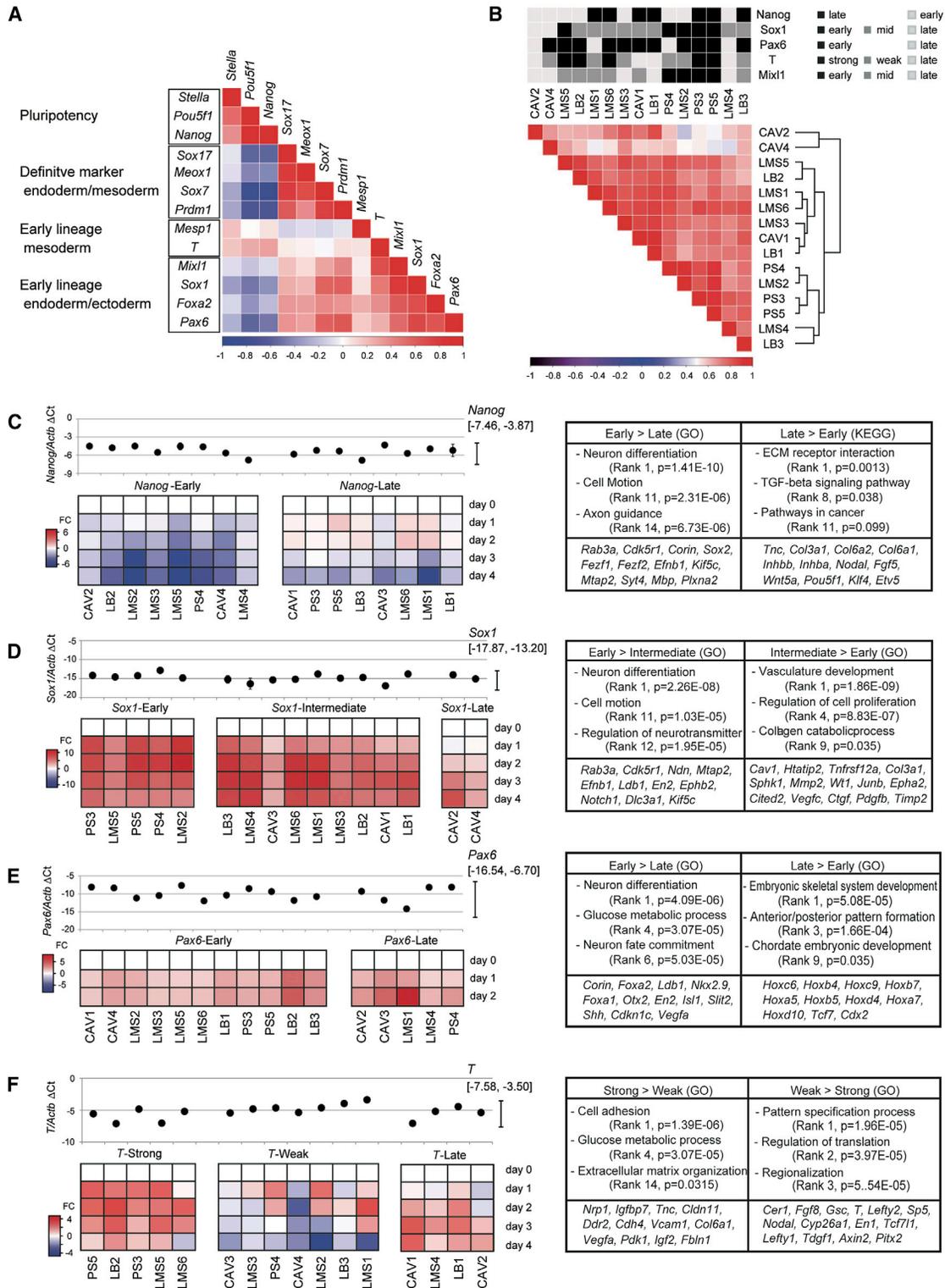


Figure 6. Gene Expression Profiles in EpiSCs during Short-Term In Vitro Differentiation

(A) Coefficient matrix of Pearson correlation analysis of qPCR results of 13 genes (relative to *Actb*) in 16 EpiSC lines over 4 days of in vitro differentiation. Clustering of gene expression data reveals syn-expression groups for pluripotency-related, definitive endoderm/mesoderm, mesoderm progenitors, and endoderm/ectoderm progenitors.

(B) Correlation matrix of Pearson correlation analysis and hierarchical clustering of qPCR results of the gene expression pattern of the same 13 genes in the EpiSC lines during in vitro differentiation. Top panel shows the association between the different grouping of EpiSCs by the expression pattern of *Nanog*, *Sox1*, *Pax6*, *T*, and *Mixl1*.

(legend continued on next page)

that were incorporated successfully, they colonized the paraxial and lateral mesoderm adjacent to the hindgut invagination (Figures 5H and 5I). Histological examination revealed proper integration of the EpiSC-derived cells in the tissues of all three germ layers including the neural plate (Figure 5J), the paraxial mesoderm (Figure 5K), the notochord (Figure 5L), and the definitive endoderm (Figure 5M). EpiSCs that were integrated in the primitive streak initiated *T* expression (Figure 5N) and those in the mesoderm and the notochord expressed *T* strongly (Figures 5N, 5O, and S3C). EpiSC-derived cells in the primitive streak did not express *Foxa2* but did so when they were incorporated into the precursor of the notochord (Figure S3D). EpiSC-derived cells in the endoderm expressed *Sox17* and *Foxa2* (Figures 5P, 5Q, and S3E–S3H). EpiSCs therefore expressed appropriate lineage markers when they were incorporated into the host tissues. The morphogenetic behavior and the pattern of tissue colonization displayed by the EpiSCs are reminiscent of cells of the anterior primitive streak of the late-streak-stage embryo (Kinder et al., 1999). This further indicated that, consistent with the transcriptome characteristics, EpiSCs might have acquired the properties of the anterior primitive streak cells.

Temporal Gene Expression Pattern during Short-Term Differentiation Reflects the Germ-Layer Potential of the EpiSC Lines

To test the germ-layer potential of the EpiSCs, their immediate response to induction of differentiation was examined by assaying the expression of lineage marker genes during a short-term (4 day) culture. Sixteen EpiSC lines (four CAV, three PS, six LMS, and three LB lines, in triplicate samples) were studied by qPCR analysis of the expression of markers of pluripotency (*Pou5f1* and *Nanog*), progenitors of endoderm (*Mixl1*, *Sox17*, *Foxa2*, and *Sox7*), mesoderm (*T*, *Meox1*, and *Mesp1*), and neuroectoderm (*Sox1* and *Pax6*) and primordial germ cells (*Prdm1* and *Stella*) and two reference genes (*Actb* and *Tbp*).

At day 0, the level of expression of the 14 marker genes, normalized against *Actb*, did not vary significantly among the 16 EpiSC lines (Figure S4). However, upon induction of differentiation, EpiSCs displayed different patterns of up- and downregulation of the markers (summarized as heatmaps by fold changes of each gene over the 4 day period, Figure S4), whereas the expression of the housekeeping gene, *Tbp*, remained constant. The general responses were the upregulation of ectoderm (*Sox1* and *Pax6*) and endoderm (*Mixl1*, *Sox17*, *Sox7*, and *Foxa2*) genes, and the downregulation of pluripotency-related genes. A correlation matrix analysis revealed that the expression patterns of the markers could be categorized into four syn-expression groups (Figure 6A). During EpiSC differentiation, the expression of the PGC marker, *Stella*, followed that of the pluripotency genes *Pou5f1* and *Nanog*. The other PGC-related gene, *Prdm1*, behaved more like definitive endoderm genes and

therefore could be marking endoderm differentiation (Chang et al., 2002). Mesoderm precursor genes (*Mesp1* and *T*) and the early endoderm and ectoderm lineage markers (*Mixl1*, *Sox1*, *Foxa2*, and *Pax6*) made up the other two syn-expression clusters. However, hierarchical clustering and correlation matrix analysis of the expression of the marker genes collectively for the syn-expression groups (Figure 6B) did not reveal any apparent subgrouping of these EpiSC lines.

To assess if subsets of EpiSC lines could be distinguished by the expression of genes in a syn-expression group, we further analyzed the expression pattern of *Nanog*, *Sox1*, *Pax6*, *T*, and *Mixl1* by K-Medians clustering of the qPCR data set. EpiSC lines, which displayed similar levels of *Nanog* expression prior to differentiation, can be distinguished by the changes in *Nanog* expression during differentiation: seven lines showed rapid and robust downregulation (the “*Nanog*-early” group); eight lines showed moderate increase in first 2 days and then decreased expression (the “*Nanog*-late” group). Based on this classification, we reanalyzed the transcriptome of these EpiSC lines. The “*Nanog*-early” group expressed a higher level of neuronal development genes, whereas the “*Nanog*-late” group expressed a higher level of extracellular matrix (ECM) genes and pluripotency-related genes including *Pou5f1* and *Klf4* and signaling factors such as *Nodal* and *Inhba*, which codes for Activin A (Figure 6C; Table S3). The downregulation of *Nanog* expression therefore reflects the ability of EpiSCs to exit from pluripotency to undergo neural differentiation.

Next, the predisposition of neural lineage was examined by analyzing *Sox1* expression (Figure 6D). The peak expression of *Sox1* was observed at day 1 in the “*Sox1*-early” group, days 2–3 in the “*Sox1*-intermediate” group, and at day 4 in the “*Sox1*-late” group. The transcriptome of *Sox1*-early lines displayed a neural predisposition, whereas that of *Sox1*-intermediate and *Sox1*-late lines revealed a mesendoderm predisposition (Figure 6D; Table S3). Because the *Pax6* response to serum-induced differentiation did not discriminate between EpiSC lines (Figure S3), we further tested the response of EpiSCs to the withdrawal of the growth factors in a serum-free condition that drives the differentiation of EpiSCs to the anterior neural plate-like state (Iwafuchi-Doi et al., 2012). Based on the pattern of *Pax6* expression, EpiSC lines could be segregated into the “*Pax6*-early” group showing peak *Pax6* expression at day 1 and the “*Pax6*-late” group showing peak expression at day 2 or later (Figure 6E; Table S3). The *Pax6*-early EpiSC lines expressed higher level of neural lineage genes, whereas *Pax6*-late group expressed *Hox* family genes, suggesting *Pax6*-early response may be correlated with the anterior-biased ectoderm characteristics, whereas *Pax6*-late responses are associated with the propensity to generate posterior neural progenitors.

To assess the potential of mesoderm differentiation, EpiSC lines were grouped by *T* expression (Figure 6F): the “*T*-strong”

(C–F) EpiSC lines grouped on temporal gene expression pattern of (C) *Nanog*, (D) *Sox1*, (E) *Pax6*, and (F) *T* by K-Medians clustering and analyzed for differentially expressed genes and GO terms. Top panel: Plots of mean expression level at day 0 of triplicate samples normalized to *Actb* (SEM of less than 0.50 is not shown) generated with Microsoft Excel. Heatmaps display the temporal pattern of fold changes of the expression of each gene relative to day 0 of the experiment. EpiSC lines grouped as (C) *Nanog*-early and *Nanog*-late, (D) *Sox1*-early, *Sox1*-intermediate, and *Sox1*-late, (E) *Pax6*-early and *Pax6*-late, and (F) *T*-strong, *T*-weak, and *T*-late. The functional ontology and related genes ($p < 0.01$) revealed in the comparison of the transcriptomes of EpiSCs grouped by *Nanog*, *Sox1*, *Pax6*, and *T* were shown in the boxes.

See also Figure S4 and for detailed gene lists, see Table S3.

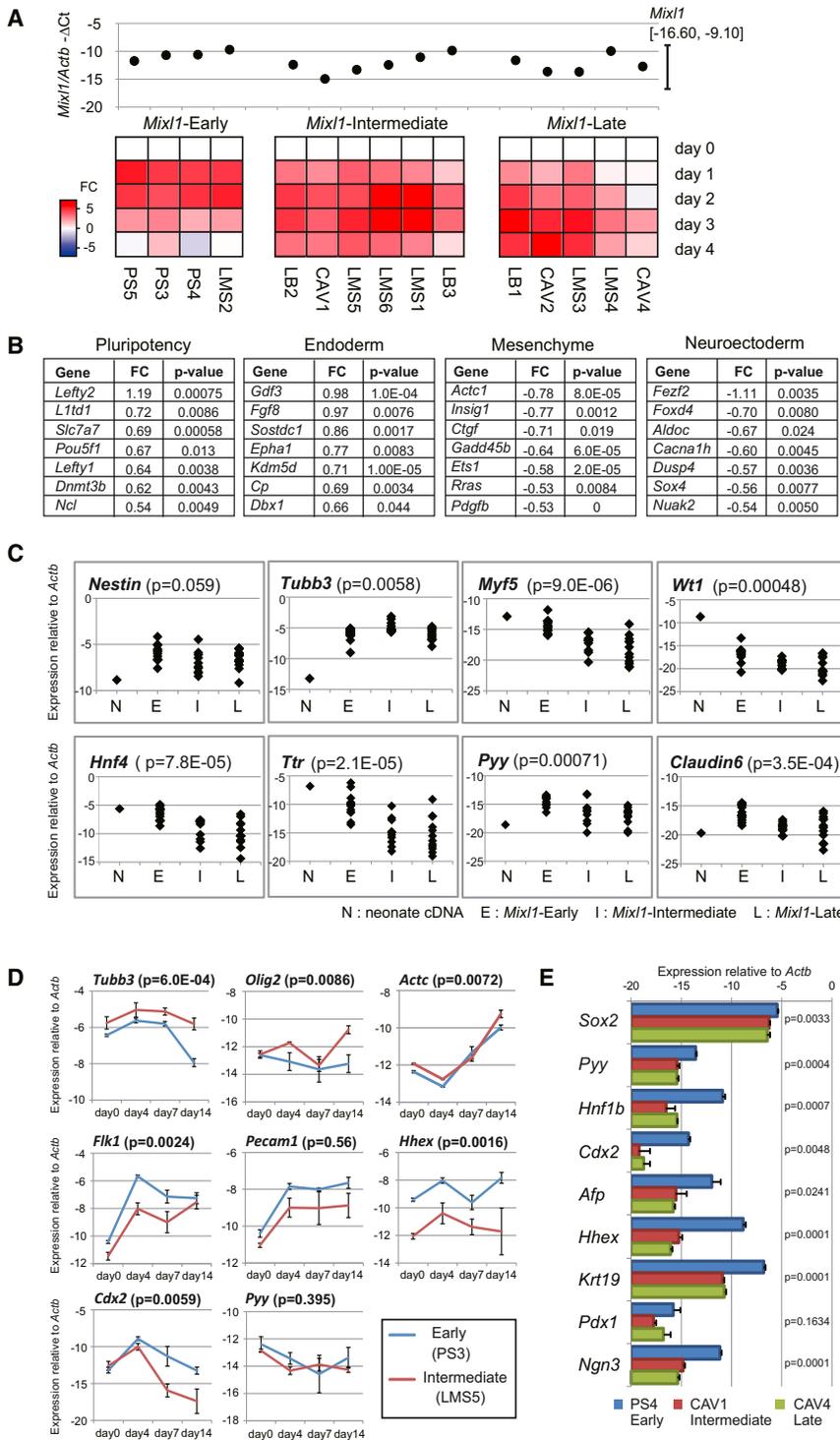


Figure 7. EpiSC Lines Displaying Different Propensity of Endoderm Differentiation

(A) EpiSC lines grouped as *Mixl1*-early (right), *Mixl1*-intermediate (middle), and *Mixl1*-late (left) by K-Medians clustering of *Mixl1* expression profiles during in vitro differentiation. Top panel: Plots of the expression level (mean of triplicate) relative to *Actb* at day 0 (SEM < 0.50 is not shown). The heatmaps display the temporal pattern of fold change of *Mixl1* expression relative to day 0 during in vitro differentiation.

(B) The gene list extracted from top 50 differentially expressed genes ($p < 0.01$): overrepresented (pluripotency- and endoderm-related genes) and underrepresented (mesenchyme- and neuroectoderm-related genes) in *Mixl1*-early compared to *Mixl1*-late (for detailed *Mixl1*-related gene lists, see Table S3; for validation of microarray data, see Table S4).

(C) qPCR analysis of the tissue-specific markers for assessing tissue composition of teratomas generated from EpiSC lines of *Mixl1*-early (E), *Mixl1*-intermediate (I), and *Mixl1*-late (L) groups; p value for variances in marker expression among the three *Mixl1* groups computed by ANOVA.

(D) qPCR result of the gene expression changes during 14 day differentiation of EpiSC lines PS3 (*Mixl1*-early) and LMS5 (*Mixl1*-intermediate); p values computed by nested ANOVA.

(E) qPCR analysis of directed endoderm differentiation of PS4 (*Mixl1*-early), CAV1 (*Mixl1*-intermediate), and CAV4 (*Mixl1*-late) EpiSC lines; error bar = SEM, p values computed by ANOVA.

See also Figures S4, S5, and Tables S3 and S4.

group showing robust upregulation, “*T*-late” group showing late upregulation, and the “*T*-weak” group showing moderate upregulation or downregulation. The correlation analysis (Figure 6A) indicated that the strong *T* expression was accompanied by upregulation of the mesoderm marker, *Mesp1*, suggesting a tendency for mesoderm differentiation. Consistent with this finding, the transcriptome of the *T*-strong group EpiSC lines show higher

expression of genes related to cell adhesion and extracellular matrix, whereas *T*-weak group expressed genes related to primitive streak and anterior mesoderm/ectoderm differentiation (Figure 6F; Table S3). EpiSCs of the *T*-weak group therefore show anterior primitive streak-like properties, whereas the *T*-strong group is predisposed for mesoderm differentiation. The *T*-late group showed higher expression of pluripotency-related genes (*Inhba*, *Inhbb*, *Klf4*) and the genes coding extracellular matrix and focal adhesion (Table S3).

In the embryo, *Mixl1* expression marks the emergence of the progenitors of mesoderm and endoderm in the primitive streak and loss of *Mixl1* leads to absence of definitive endoderm (Hart et al., 2002).

To evaluate the propensity of endoderm differentiation, EpiSCs were categorized by the timing of upregulation of *Mixl1* into “*Mixl1*-early,” “*Mixl1*-intermediate,” and “*Mixl1*-late” groups (Figure 7A). The transcriptomes of both the *Mixl1*-early and *Mixl1*-intermediate EpiSCs showed higher expression of pluripotency-related genes, but the *Mixl1*-early EpiSCs showed higher expression of endoderm-related genes. The *Mixl1*-late EpiSCs,

when compared to the other two groups, showed stronger expression of mesenchyme- and neural-related genes (Figure 7B; Table S3). All four *Mixl1*-early EpiSC lines (PS4, LMS2, PS3, and PS5) displayed the “*Sox1*-early” expression pattern (Figure 6B). The expression of lineage-related genes determined by the microarray analysis was validated by qPCR in 19 EpiSC lines (Table S4). Of the 43 genes tested, 40 showed significant correlation of array and qPCR results ($p < 0.0001$). In the four *Mixl1*-early and five *Mixl1*-late EpiSCs, these lineage-related genes, including six genes that were strongly expressed in the *Mixl1*-early group (*Lefty2*, *Dnmt3b*, *L1td1*, *Gdf3*, *Pou5f1*, and *Epha1*; Figure 7B), also showed significant correlation between the microarray and qPCR results (Table S4). Prior to in vitro differentiation, EpiSC colonies of the three *Mixl1* groups showed the presence of T-positive cells, but *Sox17*⁺ and *Foxa2*⁺ cells were not detected, suggesting that these cells had not initiated endoderm differentiation (Figures S5A and S5B). Fluorescence-activated cell sorting analysis further showed the presence of more PlexinA2-expressing (neural-related) cells in the *Mixl1*-late EpiSC colonies (Figure S5C), which is consistent with the higher level of expression of neuroectoderm-related markers (Figure 7B). A comparison of the differentially expressed genes associated with *Mixl1* grouping (Figure 7B; Table S3) of four EpiSC lines by hierarchical clustering analysis revealed a consistency of the gene expression pattern among the parental and its sublines (Figure S5D), which may account for the stability of the differentiation propensity of the EpiSC line at different passages.

Lineage Differentiation Potential of EpiSCs Grouped by *Mixl1* Expression Pattern

To test the lineage differentiation potential of EpiSCs of the three *Mixl1* groups, we analyzed the tissue composition of teratomas generated from these EpiSCs by qPCR analysis (Figure 7C). The lines from *Mixl1*-early group exhibited higher expression of genes associated with the endoderm tissues (*Hnf4*, *Ttr*, *Pyy*, and *Claudin6*) and mesoderm tissues (*Myf5* and *Wt1*) but a similar level of gene expression for ectoderm tissues (*Nestin* and *Tubb3*) when compared with the *Mixl1*-late teratomas. The gene expression profiles of *Mixl1*-intermediate teratomas were more closely related to the *Mixl1*-late group.

To test more directly if the *Mixl1* expression pattern may be correlated with differences in their endoderm differentiation potential, one EpiSC line each was selected from the *Mixl1*-early and the *Mixl1*-intermediate group to track their differentiation in serum-containing medium. Neuroectoderm markers *Tubb3* and *Olig2* were expressed higher in the *Mixl1*-intermediate EpiSCs. In contrast, the *Mixl1*-early EpiSCs displayed higher expression of genes of mesoderm and endoderm derivatives, such as *Flk1*, *Pecam1*, *Hhex*, and *Cdx2* (Figure 7D). We further subjected five different EpiSC lines of the *Mixl1*-early, *Mixl1*-intermediate, and *Mixl1*-late groups to culture conditions for directed differentiation to the endoderm lineage. After 4 days of culture, the *Mixl1*-early EpiSC lines tended to express *Sox17* at a higher level than EpiSCs of the other two categories (Figure S5E). Upon further directed differentiation, the *Mixl1*-early line (PS4) expressed significantly higher expression of genes for foregut (*Sox2*, *Pyy*, and *Hnf1b*), hindgut (*Cdx2*), liver (*Afp*, *Hhex*, and *Krt19*) and pancreas (*Ngn3*) lineage when compared to *Mixl1*-intermediate (CAV1) and *Mixl1*-late (CAV4) line (Fig-

ure 7E). These results show that early upregulation of *Mixl1* during serum-mediated differentiation of the EpiSCs is correlated with a predisposition to respond more robustly to the induction of differentiation to endoderm lineage by Activin A.

DISCUSSION

Derivation of epiblast stem cells (EpiSCs) is limited to a window of postimplantation development (CAV: E5.5 to LB: E8.0) that coincides with the progressive decline of the expression of the pluripotency-related genes in the epiblast/ectoderm (epi/ect) during gastrulation (this study; Osorno et al., 2012). We showed that the transcriptomes of these EpiSC lines resemble each other regardless of the original developmental stage of the source tissue. Principal component analysis and hierarchical clustering of the transcriptomes showed that EpiSCs as a group display a gene expression profile that is most similar to that of the ectoderm of embryos at late gastrulation. At the late-gastrula stage, the ectoderm is poised for neuroectoderm differentiation. EpiSCs may therefore be at a state of “primed” pluripotency that is distinct from “naive” pluripotency of mESCs (Nichols and Smith, 2009). The primed pluripotency state of the EpiSCs may also account for their inefficiency to participate in chimera formation, and the poor response to induction of primordial germ cells (Hayashi et al., 2011).

The transcriptome of the EpiSCs differs from that of the epi/ect regarding the activity of genes associated with FGF/MAPK, TGF- β , and WNT signaling. This phenomenon may reflect the response and the adaptation of the epi/ect cells to the Activin and FGF supplemented culture conditions. FGF signaling, upstream of the MAPK pathway, is crucial to restrain the neuronal differentiation of EpiSCs (Greber et al., 2010), and to stimulate feeder cells to produce Activin A, the TGF- β signaling ligand supporting the self-renewal of EpiSCs (Greber et al., 2007). The activation of these signaling pathway genes and the enrichment of primitive streak genes raise the possibility that the EpiSCs may have acquired the properties of primitive streak cells (Tam and Loebel, 2007). The enrichment of anterior streak genes versus the downregulation of posterior streak genes, and the specific morphogenetic behavior and pattern of tissue contribution displayed by EpiSCs following grafting to the primitive streak (Huang et al., 2012 and this study) further highlight that EpiSCs are showing the lineage characteristics of anterior primitive streak cells (Kinder et al., 2001; Tam and Beddington, 1987).

Our differentiation profiling of the EpiSC lines has revealed that certain lineage signatures embedded in the transcriptome can be correlated with the predisposition for germ-layer differentiation. For example, EpiSCs that robustly and rapidly upregulate *Mixl1* when induced to differentiate are enriched for transcripts characteristic of endoderm progenitors and pluripotency and differentiate more efficiently into endoderm derivatives in vitro and in vivo as teratomas. We have focused on the EpiSCs of the *Mixl1* group and demonstrated that there is a discernible endoderm lineage bias. This inherent propensity for endoderm differentiation is likely to be entrained as it persists even when cells are subject to conditions of directed differentiation to endoderm. It is presently not known whether the lineage bias of the EpiSC lines is the results of a uniform tendency of all cells to lean toward a specific path of differentiation or whether there

are differences in the composition of the stem cell colonies. Our results so far have shown that the apparent endoderm bias of the *Mixl1*-early EpiSC lines could be due to the preponderance of cells that are primed for endoderm differentiation, which can respond more effectively to induction by the TGF- β signals. Consistent with the concept of heterogeneity of cell types, EpiSCs of the *Mixl1*-early group also display *Sox1*-early phenotype, suggesting that the endoderm-biased population may coexist with neural-biased cells. In contrast to the *Mixl1*-early lines, the *Mixl1*-late lines show an overall enhanced expression of the neural lineage genes and contain a larger fraction of neural progenitor-like cells. Although we have not tested the lineage bias of cell lines grouped by other gene expression patterns, it is possible that they may also display some bias in lineage potential. A caveat of the population-based transcriptome and differentiation analysis in the present study is that it only provides an average of the cellular properties that would mask the inherent variations among cells. In this regard, future work should be directed to single-cell analysis of the transcriptome and using sorted cell population of common characteristics for assessing the innate lineage potency of the stem cells. Induced pluripotent stem cells (iPSCs) generated from different types of somatic cells using different reprogramming approaches may have different lineage potential. This phenomenon may be associated with the memory and de novo creation of epigenetic marks in the genome during cellular reprogramming (Kim et al., 2011; Polo et al., 2010; Bar-Nur et al., 2011; Liang and Zhang, 2013). It is therefore likely the predisposition of lineage potential in the EpiSCs may also involve modifications of the epigenome.

The categorization of the lineage potency of multipotent stem cells could be applied to the choice of cell lines for directed differentiation of the desirable lineage. Our findings open an experimental avenue to characterize the potential of lineage differentiation of stem cells derived from other embryonic tissues or by reprogramming of somatic cells. It would be feasible to evaluate the cells by the specific gene expression signature in conjunction with the response to the induction of differentiation. In view of the observation that EpiSCs are dependent on signaling activity similar to that for the maintenance of human ESCs (Vallier et al., 2009b, 2009c; Kim et al., 2013), the knowledge gleaned from EpiSCs may be applicable for predicting the lineage potential of human pluripotent stem cells.

EXPERIMENTAL PROCEDURES

Generation of EpiSCs

E5.5–E8.5 129 strain embryos were staged (see Supplemental Experimental Procedures: Staging of Embryos) and used for generating EpiSC lines and sampling tissues for transcriptome analysis. Whole epiblast of pregastrulation embryos and the epiblast/ectoderm (epi/ect, excluding the primitive streak) of gastrula-stage embryos were dissected (see Supplemental Experimental Procedures: Dissection and Culture of Epiblast and Ectoderm). The explants were cultured on mouse embryonic fibroblasts in EpiSC culture medium (see Supplemental Experimental Procedures: Culture Media and Preparation of MEFs). Additional EpiSC lines were also generated for characterization of the molecular properties and differentiation potential: six from 129 embryos (stages: two CAV, one PS, one LMS, and two LB), two from 129SvJ embryo harboring the CMV-RFP transgene (one each of CAV and PS stage; Vintersten et al., 2004), and three CAV stage from ICR;C57BL/6 embryos expressing the Kikume Green-Red (KikGR) transgene (Ando et al., 2002). Eleven EpiSC lines were cultured for seven to ten passages, 14 lines between ten and 20

passages, and 15 lines for more than 20 passages. The use of animals for this project has been approved by the Animal Ethics Committee of the Children's Medical Research Institute and the Children Hospital at Westmead.

Characterization of EpiSCs

For analysis of the expression of pluripotency factors, immunocytochemistry and qPCR were performed. Karyotypes were determined on metaphases prepared from the EpiSCs. Telomere length and telomere trimming were assessed by the terminal restriction fragments assays, and telomerase activity was examined by PCR amplified telomere extension assay. The differentiation potential of EpiSCs was examined by assessing the tissue composition of teratomas derived from the EpiSCs by histology and histochemistry (see Supplemental Experimental Procedures: Immunocytochemistry, Karyotyping, Telomere Analysis, and Teratoma Analysis and qPCR Analysis). RNA was extracted using RNeasy Micro and Mini Kits for cultured cell and Lipid Kits for teratoma samples (QIAGEN) and used for qPCR analysis of tissue-specific markers (see Supplemental Experimental Procedures: qPCR Analysis of Gene Expression).

Gene Expression Profiling

For microarray analysis of the gene expression profile for the epi/ect samples, RNA from single epiblast was extracted with RNeasy Micro kit with *E. coli* ribosomal RNA (Roche) as a carrier RNA and amplified with MessageAmpII aRNA kit (Ambion) following manufacturer's protocol. RNA samples of four selected EpiSC lines were amplified using MessageAmp II aRNA kit (Ambion). Ten nanogram amplified RNA obtained was labeled using the Illumina Total Prep RNA Amplification Kit (Ambion) but using second round primers recommended by the manufacturer in the MessageAmp II aRNA kit. For other EpiSC lines, 230 ng RNA samples were labeled with the Illumina Total Prep RNA Amplification Kit. All the samples were analyzed on Illumina Sentrix Mouse WG-6 v2 Expression Beadchip at Australian Genome Research Facility, and analyses were performed using R application. The intensity data were log transformed and quantile normalized. Unannotated probes were removed, and the remaining 31,646 probes were used for clustering and heatmap analyses.

Differentiation of EpiSCs in Host Embryos

EpiSCs were examined for their germ-layer differentiation potential by transplantation to the primitive streak of host ARC/s OB/EB-stage (E7.5) embryos (Franklin et al., 2007). EpiSC colonies in cultures were washed once with PBS followed by PBS containing 2 μ M CM-Dil (in alcoholic solution, Molecular Probes, C7000). Cells were incubated at 37°C for 5 min followed by 10 min on ice. After one wash with PBS, EpiSC colonies were treated with collagenase and cell clumps were collected for transplantation to host embryos. The grafted embryos were cultured in heat-inactivated rat serum for 24 hr in 5% carbon dioxide, 5% oxygen, 90% nitrogen at 37°C (Rivera-Pérez et al., 2010) in a rolling culture chamber. After 24 hr of in vitro development, the patterns of tissue contribution by the descendants of EpiSCs (labeled with CM-Dil dye or expressing red fluorescent protein [RFP]) were visualized in whole embryos by fluorescence microscopy. The embryos were fixed in 4% PFA for 30 min, the yolk sac and amnion were dissected away, and then the embryos were embedded in OCT compound and cryosectioned. For immunofluorescence imaging of cryosections counterstained with DAPI, the same antibodies as for immunocytochemistry were used.

Testing Lineage Potential of EpiSCs by In Vitro Differentiation

EpiSCs were tested for in vitro differentiation to endoderm derivatives (see Supplemental Experimental Procedures: Generating Embryoid Body from EpiSCs and In Vitro Differentiation of EpiSCs: Serum-Induced Differentiation). For short-term differentiation, day 0 RNA samples were processed also for microarray, and samples of embryoid bodies were collected daily for 4 days. Statistical analysis of the qPCR data and the generation of correlation matrix and the heatmaps were done in R (see Figure 6 legend for packages used). The cellular composition of the EpiSC colonies was analyzed by immunocytochemistry and flow cytometry (see Supplemental Experimental Procedures: Immunocytochemistry and Flow Cytometry Analysis). For long-term differentiation, embryoid bodies after 4 day culture were plated down on a gelatin-coated plate. Samples were collected on days 0, 4, 7, and 14. For directed differentiation, EpiSC colonies were dissociated into single cells,

and approximately 500 cells were placed in each microwell on the AggreWell400 plate (STEMCELL Technologies) to generate embryoid bodies, which were then cultured with a modified protocol for directed differentiation to endoderm cells (see [Supplemental Experimental Procedures: In Vitro Differentiation of EpiSCs: Directed Differentiation](#)). Samples were collected on fifth and tenth days of in vitro culture. qPCR of RNAs extracted from the cultures was performed using customized RT2 Profiler PCR Array (QIAGEN).

ACCESSION NUMBERS

The gene expression data sets reported in this paper have been deposited in the NCBI Gene Expression Omnibus (GEO) database under accession number GSE46227.

SUPPLEMENTAL INFORMATION

Supplemental Information for this article includes Supplemental Experimental Procedures, five figures, and four tables and can be found with this article online at <http://dx.doi.org/10.1016/j.stem.2013.09.014>.

AUTHOR CONTRIBUTIONS

Y.K., K.K.-F., J.B.S., K.A.S., I.E.A., H.A.P., and P.P.L.T. designed the project; Y.K., K.K.-F., J.B.S., K.A.S., M.D.P., V.J., A.H., G.d.A., G.J.L., M.D.S., and H.A.P. conducted the experiments; Y.K., D.A.F.L., E.T.T., and O.H.T. performed the bioinformatic analysis; and Y.K., K.K.-F., and P.P.L.T. wrote the paper.

ACKNOWLEDGMENTS

We thank BioResources staff at CMRI for the care of animals, Dr. Paul Tesar for providing the EpiSC9 line and the RNA sample, and Life Technologies Corporation (Melbourne branch) for assistance with the PCR arrays. Our work is supported by the NHMRC of Australia (grants 632776, 1022498, and 1051897 to G.L., I.E.A., and P.P.L.T.) and Mr. James Fairfax (to CMRI). Y.K. is supported by the Manpei Suzuki Diabetes Foundation Fellowship and the Research Fellowship of Uehara Memorial Foundation. H.A.P. is a Career Development Fellow of Cancer Institute NSW, and P.P.L.T. is a NHMRC Senior Principal Research Fellow (grant 1003100).

Received: April 9, 2013

Revised: July 22, 2013

Accepted: September 23, 2013

Published: October 17, 2013

REFERENCES

Alev, C., Wu, Y., Kasukawa, T., Jakt, L.M., Ueda, H.R., and Sheng, G. (2010). Transcriptomic landscape of the primitive streak. *Development* **137**, 2863–2874.

Ando, R., Hama, H., Yamamoto-Hino, M., Mizuno, H., and Miyawaki, A. (2002). An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein. *Proc. Natl. Acad. Sci. USA* **99**, 12651–12656.

Bao, S., Tang, F., Li, X., Hayashi, K., Gillich, A., Lao, K., and Surani, M.A. (2009). Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. *Nature* **461**, 1292–1295.

Bar-Nur, O., Russ, H.A., Efrat, S., and Benvenisty, N. (2011). Epigenetic memory and preferential lineage-specific differentiation in induced pluripotent stem cells derived from human pancreatic islet beta cells. *Cell Stem Cell* **9**, 17–23.

Bernemann, C., Greber, B., Ko, K., Sternecker, J., Han, D.W., Araúzo-Bravo, M.J., and Schöler, H.R. (2011). Distinct developmental ground states of epiblast stem cell lines determine different pluripotency features. *Stem Cells* **29**, 1496–1503.

Brons, I.G., Smithers, L.E., Trotter, M.W., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S.M., Howlett, S.K., Clarkson, A., Ahrlund-Richter, L.,

Pedersen, R.A., and Vallier, L. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448**, 191–195.

Chang, D.H., Cattoretti, G., and Calame, K.L. (2002). The dynamic expression pattern of B lymphocyte induced maturation protein-1 (Blimp-1) during mouse embryonic development. *Mech. Dev.* **117**, 305–309.

Downs, K.M., and Davies, T. (1993). Staging of gastrulating mouse embryos by morphological landmarks in the dissecting microscope. *Development* **118**, 1255–1266.

Franklin, V.J., Bildsoe, H., and Tam, P.P. (2007). Fate-mapping technique: grafting fluorescent cells into gastrula-stage mouse embryos at 7–7.5 days post-coitum. *CSH Protoc.* **2007**, pdb.prot4892.

Greber, B., Lehrach, H., and Adjaye, J. (2007). Fibroblast growth factor 2 modulates transforming growth factor beta signaling in mouse embryonic fibroblasts and human ESCs (hESCs) to support hESC self-renewal. *Stem Cells* **25**, 455–464.

Greber, B., Wu, G., Bernemann, C., Joo, J.Y., Han, D.W., Ko, K., Tapia, N., Sabour, D., Sternecker, J., Tesar, P., and Schöler, H.R. (2010). Conserved and divergent roles of FGF signaling in mouse epiblast stem cells and human embryonic stem cells. *Cell Stem Cell* **6**, 215–226.

Guo, G., Yang, J., Nichols, J., Hall, J.S., Eyres, I., Mansfield, W., and Smith, A. (2009). Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development* **136**, 1063–1069.

Han, D.W., Tapia, N., Joo, J.Y., Greber, B., Araúzo-Bravo, M.J., Bernemann, C., Ko, K., Wu, G., Stehling, M., Do, J.T., and Schöler, H.R. (2010). Epiblast stem cell subpopulations represent mouse embryos of distinct pregastrulation stages. *Cell* **143**, 617–627.

Hart, A.H., Hartley, L., Sourris, K., Stadler, E.S., Li, R., Stanley, E.G., Tam, P.P., Elefanty, A.G., and Robb, L. (2002). Mixl1 is required for axial mesoderm morphogenesis and patterning in the murine embryo. *Development* **129**, 3597–3608.

Hayashi, K., and Surani, M.A. (2009). Self-renewing epiblast stem cells exhibit continual delineation of germ cells with epigenetic reprogramming in vitro. *Development* **136**, 3549–3556.

Hayashi, K., Lopes, S.M., Tang, F., and Surani, M.A. (2008). Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. *Cell Stem Cell* **3**, 391–401.

Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S., and Saitou, M. (2011). Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* **146**, 519–532.

Huang, Y., Osorno, R., Tsakiridis, A., and Wilson, V. (2012). In vivo differentiation potential of epiblast stem cells revealed by chimeric embryo formation. *Cell Reports* **2**, 1571–1578.

Iwafuchi-Doi, M., Matsuda, K., Murakami, K., Niwa, H., Tesar, P.J., Aruga, J., Matsuo, I., and Kondoh, H. (2012). Transcriptional regulatory networks in epiblast cells and during anterior neural plate development as modeled in epiblast stem cells. *Development* **139**, 3926–3937.

Kim, K., Zhao, R., Doi, A., Ng, K., Unternaehrer, J., Cahan, P., Huo, H., Loh, Y.H., Aryee, M.J., Lensch, M.W., et al. (2011). Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. *Nat. Biotechnol.* **29**, 1117–1119.

Kim, H., Wu, J., Ye, S., Tai, C.-I., Zhou, X., Yan, H., Li, P., Pera, M., and Ying, Q.-L. (2013). Modulation of b-catenin function maintains mouse epiblast stem cell and human embryonic stem cell renewal. *Nature Commun.* **4**, 2403, <http://dx.doi.org/10.1038/ncomms3403>.

Kinder, S.J., Tsang, T.E., Quinlan, G.A., Hadjantonakis, A.K., Nagy, A., and Tam, P.P. (1999). The orderly allocation of mesodermal cells to the extraembryonic structures and the anteroposterior axis during gastrulation of the mouse embryo. *Development* **126**, 4691–4701.

Kinder, S.J., Tsang, T.E., Wakamiya, M., Sasaki, H., Behringer, R.R., Nagy, A., and Tam, P.P. (2001). The organizer of the mouse gastrula is composed of a dynamic population of progenitor cells for the axial mesoderm. *Development* **128**, 3623–3634.

Liang, G., and Zhang, Y. (2013). Genetic and epigenetic variations in iPSCs: potential causes and implications for application. *Cell Stem Cell* **13**, 149–159.

- Najm, F.J., Chenoweth, J.G., Anderson, P.D., Nadeau, J.H., Redline, R.W., McKay, R.D., and Tesar, P.J. (2011a). Isolation of epiblast stem cells from preimplantation mouse embryos. *Cell Stem Cell* 8, 318–325.
- Najm, F.J., Zaremba, A., Caprariello, A.V., Nayak, S., Freundt, E.C., Scacheri, P.C., Miller, R.H., and Tesar, P.J. (2011b). Rapid and robust generation of functional oligodendrocyte progenitor cells from epiblast stem cells. *Nat. Methods* 8, 957–962.
- Nichols, J., and Smith, A. (2009). Naive and primed pluripotent states. *Cell Stem Cell* 4, 487–492.
- Ohinata, Y., Ohta, H., Shigeta, M., Yamanaka, K., Wakayama, T., and Saitou, M. (2009). A signaling principle for the specification of the germ cell lineage in mice. *Cell* 137, 571–584.
- Osorno, R., Tsakiridis, A., Wong, F., Cambray, N., Economou, C., Wilkie, R., Blin, G., Scotting, P.J., Chambers, I., and Wilson, V. (2012). The developmental dismantling of pluripotency is reversed by ectopic Oct4 expression. *Development* 139, 2288–2298.
- Pfister, S., Steiner, K.A., and Tam, P.P. (2007). Gene expression pattern and progression of embryogenesis in the immediate post-implantation period of mouse development. *Gene Expr. Patterns* 7, 558–573.
- Pickett, H.A., Cesare, A.J., Johnston, R.L., Neumann, A.A., and Reddel, R.R. (2009). Control of telomere length by a trimming mechanism that involves generation of t-circles. *EMBO J.* 28, 799–809.
- Polo, J.M., Liu, S., Figueroa, M.E., Kulalert, W., Eminli, S., Tan, K.Y., Apostolou, E., Stadtfeld, M., Li, Y., Shioda, T., et al. (2010). Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nat. Biotechnol.* 28, 848–855.
- Rivera-Pérez, J.A., Jones, V., and Tam, P.P. (2010). Culture of whole mouse embryos at early postimplantation to organogenesis stages: developmental staging and methods. *Methods Enzymol.* 476, 185–203.
- Sun, B., Ito, M., Mendjan, S., Ito, Y., Brons, I.G., Murrell, A., Vallier, L., Ferguson-Smith, A.C., and Pedersen, R.A. (2012). Status of genomic imprinting in epigenetically distinct pluripotent stem cells. *Stem Cells* 30, 161–168.
- Tam, P.P., and Beddington, R.S. (1987). The formation of mesodermal tissues in the mouse embryo during gastrulation and early organogenesis. *Development* 99, 109–126.
- Tam, P.P., and Loebel, D.A. (2007). Gene function in mouse embryogenesis: get set for gastrulation. *Nat. Rev. Genet.* 8, 368–381.
- 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.
- Vallier, L., Mendjan, S., Brown, S., Chng, Z., Teo, A., Smithers, L.E., Trotter, M.W., Cho, C.H., Martinez, A., Rugg-Gunn, P., et al. (2009a). Activin/Nodal signalling maintains pluripotency by controlling Nanog expression. *Development* 136, 1339–1349.
- Vallier, L., Touboul, T., Brown, S., Cho, C., Bilican, B., Alexander, M., Cedervall, J., Chandran, S., Ahrlund-Richter, L., Weber, A., and Pedersen, R.A. (2009b). Signaling pathways controlling pluripotency and early cell fate decisions of human induced pluripotent stem cells. *Stem Cells* 27, 2655–2666.
- Vallier, L., Touboul, T., Chng, Z., Brimpari, M., Hannan, N., Millan, E., Smithers, L.E., Trotter, M., Rugg-Gunn, P., Weber, A., and Pedersen, R.A. (2009c). Early cell fate decisions of human embryonic stem cells and mouse epiblast stem cells are controlled by the same signalling pathways. *PLoS ONE* 4, e6082.
- Varela, E., Schneider, R.P., Ortega, S., and Blasco, M.A. (2011). Different telomere-length dynamics at the inner cell mass versus established embryonic stem (ES) cells. *Proc. Natl. Acad. Sci. USA* 108, 15207–15212.
- Vintersten, K., Monetti, C., Gertsenstein, M., Zhang, P., Laszlo, L., Biechele, S., and Nagy, A. (2004). Mouse in red: red fluorescent protein expression in mouse ES cells, embryos, and adult animals. *Genesis* 40, 241–246.