SUPPLEMENTARY INFORMATION

Single-cell RNA-Seq reproducibility and sensitivity

As single-cell RNA-Seq is a developing technology that relies on amplification of a small amount of starting material, establishing limits for its sensitivity and accuracy is crucial to interpreting these data[1]. In order to determine thresholds above which we judged our single-cell RNA-Seq data to accurately reflect gene expression, we analyzed data from two replicate experiments in which mESCs grown in serum+LIF from the same culture were captured on separate Fluidigm C1 plates and from which material was processed, amplified, and sequenced independently. We parsed our single-cell RNA-Seq data using the fraction of cells within the population in which a gene was detected (α), and the mean expression level among cells in which the gene was detected (μ), as previously described[1]. We observed that correlations in μ between the two replicate samples reached a maximum when only genes detected in $\geq 21\%$ of cells were included in the analysis, and therefore judged genes detected in at least 21% of cells to be reliably detected **(Extended Data Fig. 1B)**. A subset of genes, including several lineage regulators and signaling factors that have been previously identified as targets of the Polycomb family of epigenetic regulators in embryonic stem cells[2], were undetected in a majority of cells in the population by single-cell RNA-Seq but showed relatively high expression in a small fraction of cells **(Figure 1)**. Amplification artifacts could in principle lead to occasional overestimation of expression levels for genes expressed in low abundance, but sufficiently high expression levels have been shown to be reproducible and less likely to be due to such artifacts[3]. We therefore sought to determine threshold expression levels

above which expression estimates were reproducible between plates, in order to determine which polycomb target genes truly show sporadically high expression. When comparing the two replicates, many genes were detected in one plate (typically at low levels), but not the other (points on the x- and y-axes **Extended Data Fig. 1C**). By setting a maximum expression threshold of ln(TPM) ≥ 2.48, 95% of these irreproducible expression estimates were eliminated. On the basis of these two criteria for reproducibility, we judged genes detected in at least 21% of cells, or showing a maximum expression level of ln(TPM) ≥ 2.48 and detected in at least 5% of cells, to be reliably detected. Genes passing these criteria showed strong agreement in both α **(Extended Data Fig. 1D)** and μ **(Extended Data Fig. 1E)** between the two replicates. Single-cell QPCR showed that α remained relatively consistent between independent biological cultures grown at different times under the same conditions **(Extended Data Fig. 3F)**.

Quantitative comparison between single-molecule FISH and single-cell RNA-Seq indicated a detection efficiency for individual transcripts by single-cell RNA-Seq of ~20%, in line with previous estimates[1, 4-6] (**Extended Data Fig. 3C**). Taking into account this limit of detection, the fraction of cells in which a gene was detected showed good agreement between biological replicate samples and across different technologies for measuring gene expression (**Extended Data Fig. 3**). In total, 12,768 protein-coding genes were reliably detected by single-cell RNA-Seq as being expressed in mESCs cultured in serum+LIF, with a core set of 2,482 genes detected in all cells profiled, a further 3,829 genes detected in at least 80% of cells, an additional 5476 genes detected in

20–80% of cells, and 981 genes detected in 5–20% of cells **(SI Table 2)**. An additional 558 long non-coding RNAs and 259 pri-miRNAs were detected **(SI Table 7)**.

Analysis of mESCs cultured in serum+LIF for the presence of distinct subpopulations

Multiple lines of evidence confirm that the observed sporadic expression of certain genes (which we define as those for which expression is detected in 1–20% of the cell population), including Polycomb target genes, was not simply due to the presence of differentiated cells in the mESC populations being profiled: 1) single-molecule FISH verified sporadic expression of a subset of genes, including a number of Polycomb target genes **(Extended Data Figs. 3–4)**, and cells showing expression were not obviously differentiated by appearance but rather were members of morphologically normal ESC colonies; 2) although expression of any one of these genes was relatively rare, virtually every individual cell profiled expressed several members of this class of genes at detectable levels (see **Extended Data Fig. 8A**); 3) sporadic gene expression was not associated with reduced expression of the master pluripotency regulator *Oct4*; and 4) individual mESCs sorted on the basis of expression of the pluripotency cell surface marker SSEA1 also showed sporadic gene expression by single-cell qPCR **(Extended Data Fig. 6B)**.

To further test the hypothesis that a subset of differentiated cells in our cultures accounts for the observed sporadic expression of lineage regulators, we performed principal component analysis on the single-cell RNA-Seq data from the 183 mESCs profiled under

serum+LIF culture conditions. The first three components accounted for 10% of the total variation between cells, with the first component accounting for 5.3% of the total variation. These axes defined a set of 14 cells that clustered away from the majority of cells **(Extended Data Fig. 2B)**. These cells did not form a distinct group as measured by principal component analysis, but rather were defined by their distance from the majority of cells. To test if these outlier cells represented differentiated cells in the culture, we determined which of the top 5% of genes contributing to the difference in these cells along the principal component axes were either pluripotency or lineage regulators, and whether these were differentially expressed between the cells. Of the pluripotency regulators examined, only *Nr0b1* was among the top 5% of genes contributing to the first three components and showed statistical significance for differential expression between the cell groups, with *Nr0b1* expression being higher in the majority of cells as compared to the outlier cells **(Extended Data Fig. 2C)**. Among 206 lineage regulators examined, only *Pax3* contributed to the top 5% of variation explained by the first three principal components and was differentially expressed, with *Pax3* expression actually being higher in the majority of cells as compared to the outlier cells. Notably, the outlier cells displayed similar *Oct4* expression levels to the majority of cells, indicating that they do not represent a differentiated subpopulation of cells.

Previous work has suggested that *Krt8*, *Krt18*, and *Hes1* mark a subpopulation of epiblast stem cells (EpiSCs) within cultures of mESCs grown in serum+LIF[7]. To test if the outlier cells we identified are epiblast stem cells, we first examined expression of *Krt8*, *Krt18*, and *Hes1* within these cells, and found that *Krt8* and *Krt18* were indeed among the top

5% of genes contributing to the first principal component. However, examination of canonical markers of EpiSCs[8, 9] within these cells such as *Lefty1*, *Eomes*, *Nodal*, *Otx2*, *Gata6*, and *Fgf5* did not reveal significantly higher expression of these genes within the outlier cells as compared to the majority of cells **(Extended Data Fig. 2C)**. Expression levels of *Dppa3* (*Stella*), which have been shown to be markedly lower in EpiSCs as compared to mESCs, were in fact higher in the outlier cells than in the majority of cells. Therefore, these outlier cells are not EpiSCs.

Previous studies have postulated the existence of distinct stem cell subpopulations characterized by reduced levels of particular pluripotency regulators that fluctuate within the population, including *Nanog*[10], *Rex1(Zfp42)*[11], *Dppa3(Stella)*[12], and *Hes1*[13]. All of these transcription factors, along with the nuclear hormone receptor and pluripotency regulator *Nr0b1*, showed bimodal expression patterns in our single-cell RNA-Seq data **(Extended Data Fig. 2D)**. Expression states of these regulators did not define distinct subpopulations of cells, as the outlier cells identified by the principal component analysis did not cluster in a distinct group and contained cells in both high and low states of these genes. As a class, however, these outlier cells were more likely to be in a low *Nr0b1* (Hypergeometric P-value = 6e-4), low *Nanog* (P = 9e-6), or low *Zfp42(Rex1)* (P = 0.02) state as compared to the majority of cells. Furthermore, while the outlier cells did not show statistically significant higher expression of any single lineage regulator as compared to the majority of cells, as a class they showed higher average Polycomb target gene expression (P-value < 0.001, Kolmogorov-Smirnov test) **(Extended Data Fig. 2E)**. Further analysis (described in **Figure 2F–G**, **Extended Data Fig. 5**, and under

'**Correlation between gene expression states'**, below) shows that low expression of *Nr0b1* increases the probability of a cell being in a low *Nanog* or low *Zfp42* state, and that low expression states of all three pluripotency regulators are associated with an increased probability of high Polycomb target gene expression. Our data therefore indicate that fluctuating expression of pluripotency regulators does not define coherent stem cell subpopulations, but rather functions to influence gene expression in a probabilistic manner that may enable exploratory decision making within the population.

Single-cell analysis of neural precursor cells derived from embryonic stem cells

To examine the impact of lineage specification on population heterogeneity, we directed the differentiation of mESCs into neural precursor cells (NPCs) using established protocols[14, 15] and profiled the resulting cells by single-cell RNA-Seq **(SI Table 8)**. As observed for pluripotency regulators in mESCs, neural regulators showed both unimodal and bimodal expression patterns in NPCs. Some regulators of neural stem cell maintenance, such as *Hes1*[16], were detected in virtually all NPCs and show unimodal expression **(Extended Data Fig. 4)**. A subset of these unimodally expressed factors, including *Pax6*[17] and *Prrx1*[18], are known targets of Polycomb-group proteins and show sporadic expression in ESCs. Other regulators of neuronal subtype specification that are targeted by Polycomb-group proteins in ESCs show bimodal expression in NPCs. Included among this set of genes are the homeobox protein *Msx1*, which is involved in the differentiation of dopaminergic neurons[19], and *Runx1*, which is involved in specification of dorsal root ganglion neurons[20, 21], suggesting that heterogeneous expression of these factors may influence cell-fate decisions upon induction of NPC

differentiation. A subset of lineage regulators targeted by Polycomb-group proteins in ESCs showed sharply reduced expression in NPCs, including the musculoskeletal regulator *Pax3*, which reflects the loss of developmental potential upon differentiation of ESCs to NPCs. In contrast to mESCs cultured in serum+LIF, which were predominantly classified into one group by clustering and principal component analysis, and which shift into a ground state upon inhibition of external signaling pathways[22], the NPC population showed four separate states and displayed greater heterogeneity as compared to ESCs **(Extended Data Fig. 4)**. Functional categories that were enriched among genes discriminating these four states including cell cycle regulation, eye development, and cell adhesion. These substates may represent previously undescribed NPC subclasses that are capable of differentiation into distinct neuronal subtypes, or partially differentiated intermediate states. Thus, characterization of NPCs derived from ESCs reveals greater heterogeneity and more clearly defined substates within this population as compared to ESCs. This expanded number of states, increased heterogeneity, and variable expression of key regulatory factors may contribute to the observed tendency of these progenitor cells to give rise to multiple neuronal types upon induction of differentiation[14, 15].

Coupling between variable regulators

Our data show a moderate degree of coupling between factors, which we interpret as reflecting the probabilistic nature of gene expression governed by multiple regulatory factors. Importantly, Polycomb target genes in serum+LIF mESCs do not show a linear dependence on the mRNA level of individual pluripotency regulators, but rather are influenced by the expression state (high or low) of variable pluripotency regulators such

as *Esrrb*, *Nr0b1*, *Zfp42 (Rex1)*, and *Nanog* **(Figure 2G** and **Extended Data Fig. 5)**. The heritability of expression states observed for particular highly variable pluripotency regulators **(Figure 2A–E)** indicates that entire colonies or portions of colonies contain higher and lower protein levels of these regulators depending on gene expression state, and we find that it is this expression state that governs the probability of a cell expressing a Polycomb target gene. We also find a relationship between expression states of fluctuating pluripotency regulators **(Figure 2F** and **Extended Data Fig. 5)**, such that low *Nr0b1* cells are more likely to be in a low *Zfp42* and/or low *Nanog* state as compared to high *Nr0b1* cells, placing *Nr0b1* within a hierarchy of transcription factors governing *Zfp42* and *Nanog* expression.

Polycomb target gene expression in ground and transition states

Polycomb target genes showed an overall increase in the fraction of cells in which they were detected in 2i+LIF and *Dgcr8-/-* mESCs **(Extended Data Fig. 8A)**, and lineage regulator expression showed a greater similarity between *Dgcr8-/-* cells grown in serum+LIF and wild-type cells grown in 2i+LIF than those cultured in serum+LIF **(Extended Data Fig. 7C)**. While Polycomb target gene expression under serum+LIF culture conditions had shown a dependence on the expression of particular pluripotency regulators, the expression of these genes in 2i+LIF and *Dgcr8-/-* mESCs was less strongly correlated with pluripotency factor expression **(Extended Data Fig. 8B)**. This may reflect the reduced levels of H3K27me3 at these genes in 2i+LIF and *Dgcr8-/-* mESCs **(Figure 4E)**.

Analysis of cell cycle-dependent variability

To determine the fraction of the observed gene expression variability that was due to cell cycle-dependent effects, we isolated individual cells in the G_0/G_1 phase of the cell cycle (described under 'Single-cell qRT-PCR' methods below) and profiled their expression by single-cell qRT-PCR as shown in Figure 3 ('Hoechst' sample). The subset of cells in this phase of the cell cycle showed only four genes significantly changed in expression as compared to single cells cultured in the same conditions that were not sorted on the basis of their cell cycle phase **(Extended Data Fig. 11B)**. Coupled with the observed persistence of expression states through multiple cell divisions (Figure 2), this finding indicates that a substantial portion of the gene expression variability we observed was cell cycle-independent.

Perturbation analysis and state classification of PSCs

For mESCs exposed to different perturbations, those missing the chromatin regulator *Mbd3*[23], grown in 2i+LIF culture, or lacking mature miRNAs showed the greatest number of significant gene expression changes as compared to cells grown in serum+LIF culture **(Extended Data Fig. 6C)**. A substantial fraction of *Dgcr8-/-* and *Dicer* knockout mESCs were assigned to the ground state using our classification algorithm. Although differentiation defects have been observed in both *Dgcr8-/-* and *Dicer* knockout mESCs[24, 25], their similarity to cells cultured under ground state conditions is unlikely to solely reflect this blockade, as mESCs lacking DNA methyltransferases[26], the *Eed* Polycomb-group protein[27], and the chromatin modifier *Mbd3*[23] also show impaired differentiation but were not classified as being in the ground state. The assignment of a substantial

fraction of *Mbd3* knockout mESCs to the primed state may reflect previously observed upregulation of trophectoderm markers upon impairment of *Mbd3* function[28, 29].

Single-cell correlation between expression of predicted let-7 or miR-152 target genes

To determine whether targets of let-7 and miR-152 were more highly correlated in 2i than non-targets, we computed the pairwise Pearson correlation coefficients among miRmap predicted target genes, and among non-target genes, and performed a Student's T-test to compare the mean correlations of predicted targets and non-targets. Predicted let-7 and miR-152 targets were observed to be more highly correlated in expression at a single-cell level in 2i+LIF as compared to non-target genes (Mean let-7 targets (ngenes = 4,403) PCC: 6.229830e-04, Mean non-let-7 targets (n = 18040) PCC: -1.673788e-06, P-value = 3.805817e-58; Mean miR-152 targets (ngenes = 1,533) PCC: 5.855215e-04, Mean non-miR-152 targets (n = 20,910) PCC: 3.275421e-05, P-value = 9.669e-08), suggesting that these genes and miRs comprise a distinct regulatory module.

Perturbation of the *c-myc* / *Lin28* / let-7 axis, miRNA balance, and cell state

We find that both Lin28a knockdown **(Figure 5D)** and induced let-7 expression **(Figure 5E)** in serum+LIF culture lead to a higher proportion of compact colonies showing uniformly positive AP staining **(Figure 5F)**. However, more overall colonies are formed upon Lin28a knockdown, while fewer colonies are observed upon let-7 induction. Our model suggests that this is a consequence of let-7 levels induced by the two perturbations. Lin28a knockdown relieves the block to processing of multiple let-7 family members into their mature form, resulting in more uniformly physiological let-7 levels such as those

observed in 2i culture. Let-7 induction is a less controlled process that may lead to excessive let-7 levels in some cells. This in turn may lead to growth arrest or the formation of colonies that are too small to be counted in our assay, or result in differentiation or cell death. Cells that express an appropriate amount of let-7 are driven into the ground state, as evidenced by the morphology, AP staining pattern, and myc expression levels in the colonies that are formed **(Figure 5E)**. Population average let-7 levels upon induction were comparable to those found under 2i conditions (fold-change of let7-g in induced versus uninduced cells as compared to cumulative fold-change of all let-7 family members measured in 2i versus serum, with both showing ~500-fold induction, **Extended Data Fig. 9**), indicating that let-7 expression at these levels is able to drive phenotypic changes. Overall, let-7 levels measured in 2i conditions were ~16-fold lower than those measured in MEFs (based on expression determined by QPCR normalized to reference small RNAs), with expression of some let-7 family members approaching levels found in MEFs (**Figure 5A** and **Extended Data Fig. 9**). Our model postulates that let-7 and ESCC miRNAs balance each other through their function in the context of regulatory networks, and builds upon previous work showing that let-7 miRNAs directly repress a set of genes indirectly activated by ESCC miRNAs. Considering these network and indirect effects, there is no reason to believe that let-7 and ESCC miRNAs must be expressed at equivalent molar ratios in order to functionally balance one another.

It is surprising to find that factors previously thought to act solely to promote differentiated states, such as the let-7 family of miRNAs and miR-152, function to counterbalance the effect of ESCC miRNAs in the ground state. miRNAs have been

considered to buffer against gene expression noise through direct attenuation of aberrant transcription, and may also act to generate thresholds in target gene expression[30, 31]. Here, we observed that removal of this layer of gene regulation resulted in a shift towards ground state self-renewal, suggesting that by virtue of their function in the context of genetic regulatory networks, miRNAs mediate gene expression fluctuations that may enable cell fate decision-making in PSCs. A dramatic decrease in expression of several let-7 family members has been observed during the derivation of ESCs from the inner cell mass (ICM), suggesting that let-7 may function transiently during early development to help promote ground state pluripotency[32].

Ground-state culture involves treatment with inhibitors of both Erk and GSK3 signaling pathways. To test which pathway governs regulation of the *c-myc* / *Lin28* / let-7 axis, we cultured mESCs with inhibitors of the two signaling pathways individually or in combination, and compared expression of relevant genes under these conditions to culture in serum+LIF. Inhibition of Erk signaling alone or in combination with GSK3 inhibition resulted in similar levels of repression of *c-myc*, *Lin28a*, *Lin28b*, *Dnmt3b*, *Dnmt3l*, and *Bmp4* compared to serum culture, indicating that Erk signaling is the principal external regulator of the *c-myc* / *Lin28* / let-7 axis in mESCs **(Extended Data Figure 9)**.

Our results suggest that Erk signaling mediates the effect of differentiation-inducing signals in serum on PSCs in part through *c-myc* / *Lin28*, and interfering with this axis helps shield PSCs from inductive signals in the environment and drives ground-state self-

renewal. Erk has been shown to stabilize c-myc[33] and suppress let-7[34], and c-myc in turn is a known transcriptional activator of Lin28, which blocks let-7 processing[34-38]. This positive feedback loop acts to stabilize *c-myc* and *Lin28* expression, which is promoted by ESCC miRNAs, which themselves are bound by c-myc **(Figure 5H)**. This loop can be broken by inhibition of Erk signaling, which destabilizes c-myc leading to reduction of Lin28 and upregulation of let-7, which in turn feeds back to negatively regulate *Lin28*, *c-myc*, and other genes indirectly activated by ESCC miRNAs including *Dnmt3b*, *Dnmt3l*, and *Bmp4*, with the net result being ground-state self-renewal. Inhibition of Lin28a **(Figure 5D)** or sustained forced expression of let-7 **(Figure 5E)** is also able to activate this positive feedback circuit and reprogram a portion of mESCs into the ground state even in the presence of serum. Ground-state self-renewal can also be achieved by ablation of ESCC miRNAs, as in *Dgcr8-/-* mESCs, thereby removing an activating force driving *c-myc* and *Lin28* expression. The interplay of the *c-myc* / *Lin28* / let-7 axis with other mechanisms contributing to entry and exit from ground state pluripotency[39] will be an important area for further study.

Transcriptional heterogeneity and regulatory network architecture

PSCs cultured in serum+LIF are exposed to conflicting signals promoting both self-renewal and differentiation. We find that a subset of genes, particularly those involved in signaling and developmental regulation, fluctuate in the presence of these signals, with heritable expression over multiple cell divisions and abrupt transitions between states. This pronounced variability of genes that mediate cellular responses to environmental signals may represent a mechanism for cells in the same milieu to respond differentially

to external stimuli. Early studies on gene expression variability, particularly in microbial systems, highlighted a dominant role for expression level focused on inevitably noisy expression arising from stochastic production of low numbers of mRNAs from genes expressed at low levels[40-43]. As observed here and in other recent studies, it is becoming increasingly evident that even highly expressed mammalian genes can display considerable variability[1, 3, 44, 45].While mammalian genes have been found to show widely varying transcriptional bursting kinetics[46], it remains unclear how this property affects the biological function and regulation of genes within transcriptional networks, or how natural selection has shaped this feature of gene expression. Further work will be needed to fully address these questions.

SUPPLEMENTAL INFORMATION REFERENCES

1. Shalek, A.K. et al. Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature* **510**, 363-369 (2014).

2. Ku, M. et al. Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet* **4**, e1000242 (2008).

3. Shalek, A.K. et al. Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. *Nature* **498**, 236-240 (2013).

4. Grun, D., Kester, L. & van Oudenaarden, A. Validation of noise models for single-cell transcriptomics. *Nat Methods* **11**, 637-640 (2014).

5. Islam, S. et al. Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Res* **21**, 1160-1167 (2011).

6. Islam, S. et al. Highly multiplexed and strand-specific single-cell RNA 5' end sequencing. *Nat Protoc* **7**, 813-828 (2012).

7. Islam, S. et al. Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat Methods* **11**, 163-166 (2014).

8. Brons, I.G.M. et al. Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448**, 191-195 (2007).

9. Tesar, P.J. et al. New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* **448**, 196-199 (2007).

10. Chambers, I. et al. Nanog safeguards pluripotency and mediates germline development. *Nature* **450**, 1230-1234 (2007).

11. Toyooka, Y., Shimosato, D., Murakami, K., Takahashi, K. & Niwa, H. Identification and characterization of subpopulations in undifferentiated ES cell culture. *Development* **135**, 909-918 (2008).

12. Hayashi, K., Lopes, S.M.C.d.S., Tang, F. & Surani, M.A. Dynamic Equilibrium and Heterogeneity of Mouse Pluripotent Stem Cells with Distinct Functional and Epigenetic States. *Cell stem cell* **3**, 391-401 (2008).

13. Kobayashi, T. et al. The cyclic gene Hes1 contributes to diverse differentiation responses of embryonic stem cells. *Genes & development* **23**, 1870-1875 (2009).

14. Lee, S.H., Lumelsky, N., Studer, L., Auerbach, J.M. & McKay, R.D. Efficient generation of midbrain and hindbrain neurons from mouse embryonic stem cells. *Nat Biotechnol* **18**, 675-679 (2000).

15. Okabe, S., Forsberg-Nilsson, K., Spiro, A.C., Segal, M. & McKay, R.D. Development of neuronal precursor cells and functional postmitotic neurons from embryonic stem cells in vitro. *Mechanisms of development* **59**, 89-102 (1996).

16. Nakamura, Y. et al. The bHLH gene hes1 as a repressor of the neuronal commitment of CNS stem cells. *J Neurosci* **20**, 283-293 (2000).

17. Ericson, J. et al. Pax6 controls progenitor cell identity and neuronal fate in response to graded Shh signaling. *Cell* **90**, 169-180 (1997).

18. Shimozaki, K., Clemenson, G.D., Jr. & Gage, F.H. Paired related homeobox protein 1 is a regulator of stemness in adult neural stem/progenitor cells. *J Neurosci* **33**, 4066-4075 (2013).

19. Andersson, E. et al. Identification of intrinsic determinants of midbrain dopamine neurons. *Cell* **124**, 393-405 (2006).

20. Kobayashi, A., Senzaki, K., Ozaki, S., Yoshikawa, M. & Shiga, T. Runx1 promotes neuronal differentiation in dorsal root ganglion. *Mol Cell Neurosci* **49**, 23-31 (2012).

21. Theriault, F.M. et al. Role for Runx1 in the proliferation and neuronal differentiation of selected progenitor cells in the mammalian nervous system. *J Neurosci* **25**, 2050-2061 (2005).

22. Ying, Q.-L. et al. The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519-523 (2008).

23. Kaji, K. et al. The NuRD component Mbd3 is required for pluripotency of embryonic stem cells. *Nature cell biology* **8**, 285-292 (2006).

24. Kanellopoulou, C. et al. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes & development* **19**, 489-501 (2005).

25. Wang, Y., Medvid, R., Melton, C., Jaenisch, R. & Blelloch, R. DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. *Nature genetics* **39**, 380-385 (2007).

26. Tsumura, A. et al. Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. *Genes Cells* **11**, 805-814 (2006).

27. Chamberlain, S.J., Yee, D. & Magnuson, T. Polycomb repressive complex 2 is dispensable for maintenance of embryonic stem cell pluripotency. *Stem cells* **26**, 1496-1505 (2008).

28. Latos, P.A. et al. NuRD-dependent DNA methylation prevents ES cells from accessing a trophectoderm fate. *Biology open* **1**, 341-352 (2012).

29. Zhu, D., Fang, J., Li, Y. & Zhang, J. Mbd3, a component of NuRD/Mi-2 complex, helps maintain pluripotency of mouse embryonic stem cells by repressing trophectoderm differentiation. *PloS one* **4**, e7684 (2009).

30. Ebert, M.S. & Sharp, P.A. Roles for microRNAs in conferring robustness to biological processes. *Cell* **149**, 515-524 (2012).

31. Mukherji, S. et al. MicroRNAs can generate thresholds in target gene expression. *Nature genetics* **43**, 854-859 (2011).

32. Tang, F. et al. Tracing the Derivation of Embryonic Stem Cells from the Inner Cell Mass by Single-Cell RNA-Seq Analysis. *Cell stem cell* **6**, 468-478 (2010).

33. Sears, R. et al. Multiple Ras-dependent phosphorylation pathways regulate Myc protein stability. *Genes & development* **14**, 2501-2514 (2000).

34. Paroo, Z., Ye, X., Chen, S. & Liu, Q. Phosphorylation of the human microRNA-generating complex mediates MAPK/Erk signaling. *Cell* **139**, 112-122 (2009).

35. Chang, T.C. et al. Lin-28B transactivation is necessary for Myc-mediated let-7 repression and proliferation. *Proc Natl Acad Sci U S A* **106**, 3384-3389 (2009).

36. Dangi-Garimella, S. et al. Raf kinase inhibitory protein suppresses a metastasis signalling cascade involving LIN28 and let-7. *The EMBO journal* **28**, 347-358 (2009).

37. Viswanathan, S.R. & Daley, G.Q. Lin28: A microRNA regulator with a macro role. *Cell* **140**, 445-449 (2010).

38. Viswanathan, S.R., Daley, G.Q. & Gregory, R.I. Selective blockade of microRNA processing by Lin28. *Science* **320**, 97-100 (2008).

39. Betschinger, J. et al. Exit from pluripotency is gated by intracellular redistribution of the bHLH transcription factor Tfe3. *Cell* **153**, 335-347 (2013).

40. Balazsi, G., van Oudenaarden, A. & Collins, J.J. Cellular decision making and biological noise: from microbes to mammals. *Cell* **144**, 910-925 (2011).

41. Blake, W.J. et al. Phenotypic consequences of promoter-mediated transcriptional noise. *Mol Cell* **24**, 853-865 (2006).

42. Elowitz, M.B., Levine, A.J., Siggia, E.D. & Swain, P.S. Stochastic gene expression in a single cell. *Science* **297**, 1183-1186 (2002).

43. Newman, J.R. et al. Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. *Nature* **441**, 840-846 (2006).

44. Deng, Q., Ramskold, D., Reinius, B. & Sandberg, R. Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science* **343**, 193-196 (2014).

45. Wu, H. et al. Cellular resolution maps of x chromosome inactivation: implications for neural development, function, and disease. *Neuron* **81**, 103-119 (2014).

46. Suter, D.M. et al. Mammalian Genes Are Transcribed with Widely Different Bursting Kinetics. *Science* **332**, 472-474 (2011).