



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Transcription factor heterogeneity in pluripotent stem cells

Citation for published version:

Torres-Padilla, ME & Chambers, I 2014, 'Transcription factor heterogeneity in pluripotent stem cells: A stochastic advantage', *Development*, vol. 141, no. 11, pp. 2173-2181. <https://doi.org/10.1242/dev.102624>

Digital Object Identifier (DOI):

[10.1242/dev.102624](https://doi.org/10.1242/dev.102624)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

Development

Publisher Rights Statement:

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



REVIEW

Transcription factor heterogeneity in pluripotent stem cells: a stochastic advantage

Maria-Elena Torres-Padilla^{1,*} and Ian Chambers^{2,*}

ABSTRACT

When pluripotent cells are exposed to a uniform culture environment they routinely display heterogeneous gene expression. Aspects of this heterogeneity, such as *Nanog* expression, are linked to differences in the propensity of individual cells to either self-renew or commit towards differentiation. Recent findings have provided new insight into the underlying causes of this heterogeneity, which we summarise here using *Nanog*, a key regulator of pluripotency, as a model gene. We discuss the role of transcription factor heterogeneity in facilitating the intrinsically dynamic and stochastic nature of the pluripotency network, which in turn provides a potential benefit to a population of cells that needs to balance cell fate decisions.

KEY WORDS: Pluripotency, Epiblast, Embryonic stem cell (ESC), Heterogeneity, Transcription, Monoallelic

Introduction

Pluripotency is the capacity of individual cells to give rise to daughter cells that can differentiate into representative tissues of each of the primary embryonic germ layers. This property is held by cells of the early embryo, with pluripotent cells being present in mammals from the pre-implantation epiblast stage and beyond implantation until the onset of somitogenesis (Beddington, 1983; Osorno et al., 2012). Pluripotency is also a defining feature of cell lines that can be established from peri-implantation mammalian embryos. In the mouse, these are embryonic stem cells (ESCs) from the pre-implantation epiblast (Evans and Kaufman, 1981; Martin, 1981; Brook and Gardner, 1997) and so-called epiblast stem cells (EpiSCs) from the post-implantation epiblast (Brons et al., 2007; Tesar et al., 2007), although EpiSCs can also be obtained by explanting pre-implantation embryos into EpiSC culture media (Najm et al., 2011).

Pluripotency depends upon the coordinated action of a gene regulatory network (GRN; see glossary in Box 1) assembled from transcription factors and subject to modulation by signalling pathways responding to environmental cues (Ng and Surani, 2011; Levine and Davidson, 2005). This GRN is not present at fertilisation but rather is synthesised in an apparently stochastic manner upon formation of the pluripotent cells within the blastocyst (Dietrich and Hiiragi, 2007; Guo et al., 2010). The activity of the pluripotency GRN persists in epiblast cells (although not uniformly in all cells) until they either gastrulate to form mesoderm or until the start of somitogenesis. At this time, the declining level of

pluripotency transcription factors, notably OCT4 (also known as POU5F1) and NANOG, drops below a hypothesised threshold concentration, at which point pluripotency can no longer be sustained (Osorno et al., 2012).

One of the most intriguing properties of ESCs is the heterogeneity in gene expression that they display. Revealing the molecular basis of this heterogeneity is not only important for understanding the flexible nature of the pluripotent state but might also serve as a model to understand heterogeneity in other systems. Heterogeneity of pluripotent cells may arise via different modes of transcriptional regulation, as well as post-transcriptional events such as protein synthesis and cell cycle dynamics, but the contribution of these different steps to heterogeneity is largely unknown. Recent advances in understanding transcription factor heterogeneity have come from work that focuses on the regulation of *Nanog*, one of the core components of the pluripotency transcriptional network that shows heterogeneity in ESC cultures. Allele switching at the *Nanog* locus was reported in ESCs and may contribute to the observed heterogeneity of NANOG in these cell types (Miyanari and Torres-Padilla, 2012). However, other data showing that NANOG protein is

Box 1. Gene expression terms

Buffering. To lessen the effect of cell-to-cell variability. Buffering implies that, over a period of time, differences in gene expression among individual cells become averaged because all cells will have gone through similar patterns of transcriptional firing. Buffering also refers to space, whereby uncoordinated activity at the single-cell level in a cell population or a tissue becomes 'coordinated' when one considers the population as a single entity.

Bursting. A mode of transcriptional activity that designates occasional firing separated by long periods of inactivity.

Firing. Production of a nascent transcript on the chromatin template that reflects promoter activity in the nucleus.

Gene regulatory network (GRN). An interrelated composition of transcription factor genes and signals that impinge upon such genes to define a functional cell state, for example ESCs.

Half-life. Refers to the average time required to degrade 50% of the mRNA or protein and therefore reflects stability. Most data on half-life measurements analyse a population of cells and therefore ignore cell-to-cell variability.

mRNA. Refers to the mature and processed messenger transcript and reflects most often steady-state levels in the cytoplasm. Steady-state levels depend on transcriptional activation rates and mRNA maturation, stability and degradation.

Pulsatile transcription. Refers to those genes that fire transcription with a period of inactivity (long or short) between firing events.

Transcriptional activation. Refers to promoter activity and leads to the generation of a transcript from either or both alleles. It is the first clearly productive and obligate step in the gene expression cascade.

2i. Refers to dual inhibition of MEK-ERK and GSK3, resulting in inhibition of FGF and stimulation of Wnt signalling pathways in ESC cultures. This produces cells that are referred to as being in a pluripotent 'ground state'.

¹Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM U964, Université de Strasbourg, Cité Universitaire de Strasbourg, Illkirch F-67404, France. ²MRC Centre for Regenerative Medicine, Institute for Stem Cell Research, School of Biological Sciences, University of Edinburgh, 5 Little France Drive, Edinburgh EH16 4UU, UK.

*Authors for correspondence (metp@igbmc.fr; ichambers@ed.ac.uk)

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

heterogeneous across the population propose that NANOG protein levels derive equally from both alleles (Filipczyk et al., 2013). Although at first sight contradictory, these data might be explained by considering the different levels at which gene expression can be regulated and suggest that there are multiple molecular steps that underpin heterogeneity.

In this Review, we will summarise the evidence supporting transcription factor heterogeneity and its functional significance in stem cell populations. Using *Nanog* as an example, we discuss the various mechanisms that can influence heterogeneity in pluripotent stem cell populations, including allelic transcriptional regulation, variations in mRNA concentration and half-life (see glossary in Box 1) at the single-cell level as well as feedback mechanisms and post-transcriptional events.

Transcription factor heterogeneity

Several transcription factors are expressed heterogeneously in undifferentiated ESCs (see Table 1). Initial findings using immunofluorescence detection showed that NANOG was not expressed in all OCT4-expressing undifferentiated ESCs (Fig. 1A) (Hatano et al., 2005; Chambers et al., 2007). Using gene targeting to introduce fluorescent protein reporters at the *Nanog* locus, either by direct placement of GFP at the *Nanog* translation initiation codon (Chambers et al., 2007) or by engineering fluorescent protein fusions to the NANOG C-terminus (Filipczyk et al., 2013), this heterogeneity could be recapitulated, with populations of ESCs expressing the surrogate fluorescent protein heterogeneously and in a bimodal manner (Chambers et al., 2007; Filipczyk et al., 2013). Importantly, such ESCs could be purified by fluorescence activated cell sorting (FACS) of cells that did or did not express Nanog-GFP. Moreover, replating of single GFP-negative cells produced colonies containing GFP-positive cells at high efficiency, suggesting that the non-expressing cells in a heterogeneous population can revert to a *Nanog*-expressing state. Subsequent replating of populations demonstrated that the original distribution of fluorescence levels could be restored by placing cells in culture for 1-2 weeks (Chambers et al., 2007; Kalmar et al., 2009). Interestingly, the GFP-negative cells take slightly longer to reach the original distribution than GFP-positive cells due to the fact that the

GFP-negative cells are more prone to undergo a commitment event that will initiate their differentiation and prevent their contribution to the undifferentiated population. The time taken to reach the equilibrium of the original population depends on the purity, or conversely, the degree of heterogeneity remaining, within the sorted fraction.

Other transcription factors have also been demonstrated to be expressed heterogeneously (Table 1). These include *Esrrb* (van den Berg et al., 2008), *Rex1* (Toyooka et al., 2008), *Stella* (Hayashi et al., 2008), *Hex* (Canham et al., 2010), *Klf4* and *Tbx3* (Niwa et al., 2009). In the case of *Stella*, the ESCs required co-culture with fibroblasts to display heterogeneity, as otherwise the expression of this protein is vanishingly low. Functional distinctions between cells displaying heterogeneous expression of some of these factors have also been made. For example, *Hex*-positive cells have an increased propensity to contribute to extra-embryonic endoderm when reintroduced into blastocysts, with the opposite propensity observed for *Hex*-negative cells (Canham et al., 2010). In addition, ESC cultures routinely harbour a minor proportion (<5%) of cells that express *Zscan4*, a gene expressed in embryos at the 2-cell stage (Zalzman et al., 2010). Using a reporter system directed by an endogenous retroviral long terminal repeat specific for the 2-cell stage, it was shown that essentially all ESCs in a culture can cycle in and out of a state resembling the 2-cell stage (Macfarlan et al., 2012). While in this state, ESCs are capable of contributing to trophectoderm (Macfarlan et al., 2012).

An important question that remains unanswered is the extent to which these heterogeneities are interdependent. The fact that both *Esrrb* and *Klf4* are downstream targets of NANOG (Festuccia et al., 2012) could suggest that heterogeneity in *Esrrb* and *Klf4* might be eliminated if *Nanog* were expressed homogeneously. Additionally, ESRRB has been reported to cooperate with OCT4 in the transcriptional activation (see glossary in Box 1) of *Nanog*, so the converse relationship might apply (van den Berg et al., 2008). Interestingly, TBX3 has been reported to influence the expression of *Zscan4* via indirect effects on DNA methylation and histone methylation (Macfarlan et al., 2011; Dan et al., 2013). Further experiments will provide a more detailed understanding of the relationships between individual heterogeneously expressed transcription factors.

Transcriptional regulation of heterogeneity

Global analyses of chromatin localisation indicate that individual pluripotency transcription factors are located at thousands of sites in the chromatin of populations of predominantly undifferentiated ESCs. These sites are in many cases close, or immediately adjacent, to sites where additional transcription factors bind (Loh et al., 2006; Chen et al., 2008; Kim et al., 2008). In some cases, such as for OCT4 and SOX2 proteins, the biochemical basis for the juxtaposed interaction is clear as it occurs through direct interaction (Ambrosetti et al., 1997; Williams et al., 2004). In many other cases this remains to be established. Nevertheless, the idea that genes are commonly regulated through the co-binding of distinct transcription factors is in line with the notion of combinatorial control of gene expression (Ptashne and Gann, 2001). As many of the co-bound loci are adjacent to genes that encode components of the pluripotency GRN, these observations have led to the notion that binding of a transcription factor to a target gene, X, in populations of cells in which gene X is expressed, means that the transcription factor positively regulates gene X (Jaenisch and Young, 2008). This implies that the core network is composed solely of self-reinforcing interactions that maintain the activity of the pluripotency GRN. However, negative autoregulatory interactions are also possible and these could contribute to the heterogeneous expression of some transcription factors. A prime example of this is NANOG. Using

Table 1. Transcription factors heterogeneously expressed in ESCs

Protein	Comments	References
NANOG	NANOG levels fluctuate in culture and determine ESC self-renewal efficiency	Chambers et al., 2007; Kalmar et al., 2009; Abranches et al., 2013; Filipczyk et al., 2013
REX1 (ZFP42)	Rex1:GFP-positive and -negative cells show distinct colony morphologies	Toyooka et al., 2008
STELLA (DPPA3)	Fluorescent protein reporter expression requires fibroblast co-culture	Hayashi et al., 2008
ESRRB	Direct transcriptional target of NANOG	van den Berg et al., 2008; Festuccia et al., 2012
KLF4	Downstream of LIF; direct transcriptional target of NANOG	Niwa et al., 2009; Festuccia et al., 2012
TBX3	Downstream of LIF; modulates <i>Zscan4</i> indirectly	Niwa et al., 2009; Dan et al., 2013
HEX (HHEX)	<i>Hex</i> ⁺ cells contribute preferentially to primitive endoderm	Canham et al., 2010; Morgani et al., 2013
ZSCAN4	Transiently expressed in <5% of ESCs	Zalzman et al., 2010

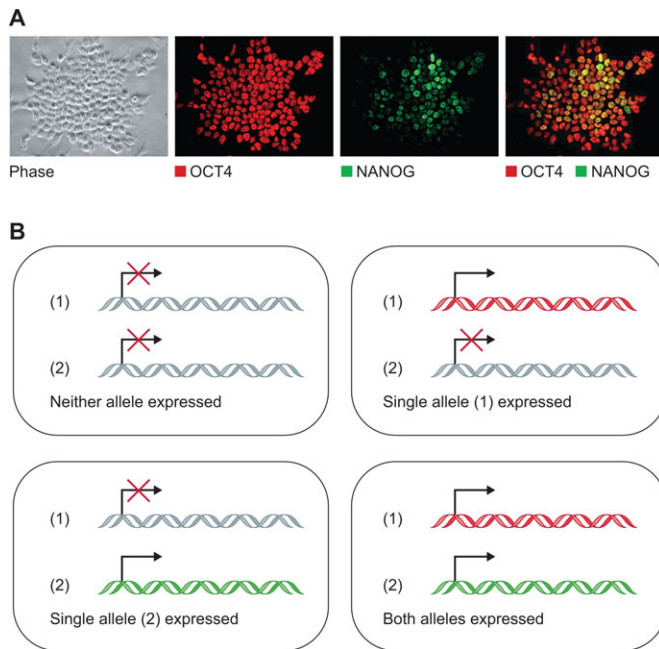


Fig. 1. Embryonic stem cells (ESCs) show heterogeneous levels of NANOG protein and transcriptional firing. (A) ESC colony showing heterogeneous expression of NANOG protein in cells that express relatively homogeneous levels of OCT4. Phase contrast and immunofluorescence images are shown for OCT4 (red) and NANOG (green). (B) Detection of *Nanog* nascent transcription has revealed four different types of ESCs: those that do not express *Nanog*, those that express from either allele, and those that express from both alleles. The existence of these four different modes of expression suggests that transcriptional firing provides a means to generate heterogeneity in ESC populations. Note that the proportion of the four modes of expression changes according to whether ESCs are cultured in serum/LIF or are supplemented with 2i/LIF; in the latter, most cells express both *Nanog* alleles and the proportion of cells that do not express *Nanog* is negligible.

inducible gain- and loss-of-function genetic approaches, NANOG was shown to decrease the level of transcription of the *Nanog* gene (Navarro et al., 2012). *Zfp281* has been implicated in the repressive effect of NANOG on the *Nanog* locus and has been proposed to act via recruitment of the NuRD complex (Fidalgo et al., 2012). Without NANOG, the degree of heterogeneity in *Nanog* gene expression in ESCs is reduced, with a higher proportion of cells expressing the *Nanog* gene (Navarro et al., 2012). Interestingly, a separate study found that transient knockdown of the endogenous *Nanog* gene combined with constitutive transgenic expression of *Nanog* resulted in a reduced variability in expression of pluripotency GRN components, suggesting that feedback loops among the GRN components regulate ESC heterogeneity (MacArthur et al., 2012).

The case of *Nanog*: allele switching may contribute to heterogeneity

Nanog was originally identified in a genetic screen for molecules that facilitate ESC self-renewal, specifically in the absence of LIF (Chambers et al., 2003; Mitsui et al., 2003). However, overexpression of *Nanog* also enhances ESC self-renewal in the presence of LIF. Indeed, subsequent analyses of a *Nanog* allelic series showed a direct relationship between the NANOG level and the self-renewal efficiency, indicating that the NANOG level acts as a differentiation rheostat (Chambers et al., 2007).

When analysing nascent *Nanog* transcription by RNA fluorescent *in situ* hybridisation (RNA-FISH), it was found that ~59% of ESCs grown in LIF/serum engage in *Nanog* transcriptional firing (see

glossary in Box 1) (Miyanari and Torres-Padilla, 2012). However, only a very low percentage of cells (14%) fire from both alleles, with the remaining 45% of cells instead displaying only one active allele, implying that, at a given time, ESCs are more likely to transcribe only one *Nanog* allele (Miyanari and Torres-Padilla, 2012). Similarly, an independent study by Navarro and colleagues reported ~10% of ESCs with biallelic firing (Navarro et al., 2012). Likewise, studies of single-molecule RNA-FISH (smRNA-FISH) using intronic probes to measure firing of transcription identified four patterns of *Nanog* transcription in ESCs at any given moment in time: those containing either of the two intronic signals (one of two alleles active), those containing two (two alleles active) and those containing none (no active allele) (Hansen and van Oudenaarden, 2013) (Fig. 1B). Counting the number of transcription sites using intronic probes revealed a strong monoallelic signal in ESCs grown in LIF/serum. Importantly, however, examination of individual cells using exonic probes showed the simultaneous presence of mature cytoplasmic mRNA transcripts from both alleles, even when only one or none of the *Nanog* alleles was firing (Hansen and van Oudenaarden, 2013). Thus, there is a discrepancy between the allelic nature of transcriptional firing and the accumulation of mature mRNA transcripts from both alleles. In agreement with the biallelic accumulation of exonic transcripts in single cells, ESC lines encoding the fluorescent protein reporter fusions NANOG-Katushka and NANOG-Venus knocked in to each of the two *Nanog* alleles revealed the overall presence of NANOG fusion proteins from both alleles in individual cells (Filipczyk et al., 2013) (Fig. 2). The allele switching of *Nanog* expression reported previously anticipated just such a scenario: if allele switching occurs on a rapid time-scale relative to the life time of the mRNA and protein, the presence of NANOG protein from both alleles in individual cells would be expected (Fig. 2) (Miyanari and Torres-Padilla, 2012). Indeed, destabilised fluorescent protein reporters inserted downstream of the *Nanog* coding region, but preceded by a self-cleavable 2A peptide, revealed frequent allele switching with irregular time intervals between switches (Miyanari and Torres-Padilla, 2012). If the destabilised self-cleavable fluorescent protein-based reporters have a shorter protein half-life than NANOG, they will reflect transcriptional firing rather than protein content. Furthermore, the range of *Nanog* mRNA concentrations in individual cells is highly variable, a potential outcome of the discontinuous accumulation of transcripts from both alleles in single cells with time. A discrepancy between allelic firing and mature mRNA transcripts is therefore expected if the frequency of firing and allele switching is irregular in time and of variable strength in individual cells within the population.

Previous analysis using reporters with stable GFP have reported a degree of mismatching between profiles of fluorescent protein and endogenous NANOG protein levels (Chambers et al., 2007; Kalmar et al., 2009; Abranches et al., 2013). However, it is important to note that the degree of matching depends upon how the cells are treated. Undifferentiated fluorescent protein-negative cells can become contaminated in at least two ways. First, without sorting for a second undifferentiated cell marker, clearly differentiated cells that no longer express the reporter will contaminate the undifferentiated population. Second, gene conversion can replace the reporter allele with a second wild-type allele, resulting in the covert existence of essentially wild-type cells in the fluorescent protein-negative population. FACS-sorted OCT4-positive cells show a good correlation between NANOG and Nanog:GFP in TNG cells, which are mouse ESCs in which a GFP open reading frame has been placed at the *Nanog* translation initiation codon via homologous recombination (Festuccia and Chambers, 2011; Chambers et al., 2007). Moreover, the utility of a given reporter allele depends upon what it is being deployed for: an allele that has a

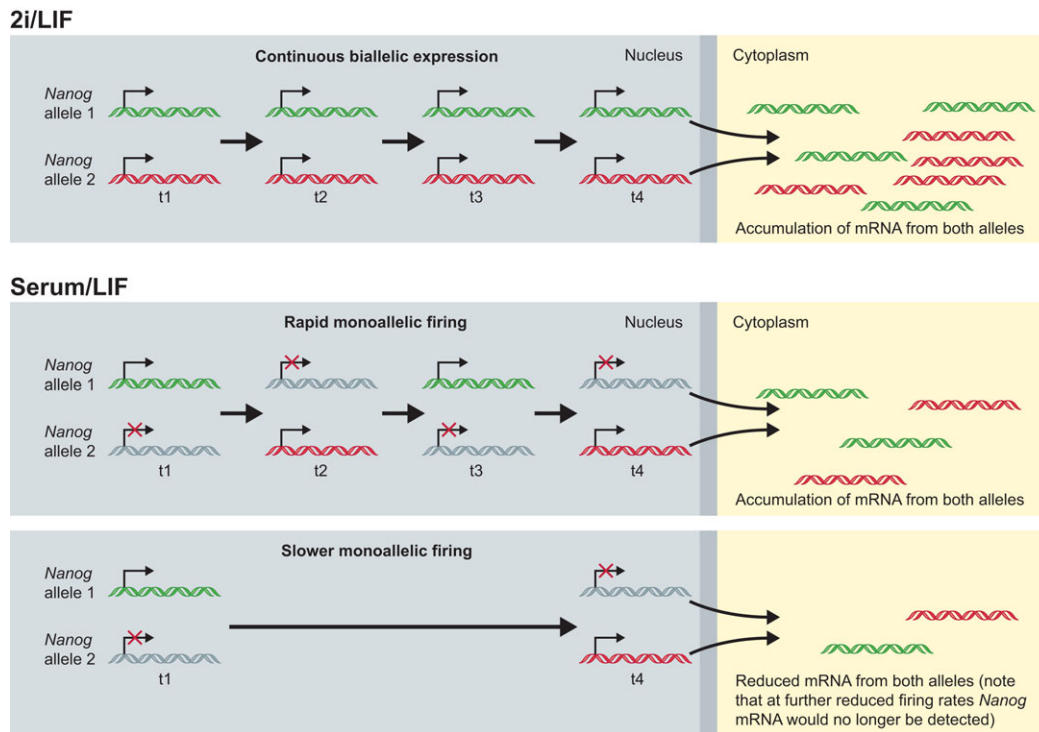


Fig. 2. *Nanog* monoallelic firing and allele switching is predominant in mouse ESCs. Allelic transcriptional firing and allele switching in ESCs grown in 2i/LIF (top) or serum/LIF (middle and bottom), over time (indicated by the arrows; t1, t2, etc.). Given the nature of *Nanog* allele switching over time, cells are expected to accumulate mature mRNA transcripts from both alleles (green and red) with the subsequent production of biallelic NANOG protein. When ESCs are cultured in 2i/LIF conditions, both *Nanog* alleles tend to be expressed at the same time (green and red), leading to accumulation of mRNA from both alleles and increased expression of *Nanog* overall in the ESC population. When ESCs are cultured in serum/LIF, the transcription rate of *Nanog* decreases. If the rate of monoallelic firing is rapid, this leads to an accumulation of higher levels of *Nanog* mRNA from both alleles (middle). When the rate of monoallelic firing is reduced, this leads to a reduction in the concentration of *Nanog* mRNA from both alleles. At further reduced firing rates *Nanog* mRNA would no longer be detectable.

degree of mismatch between fluorescent protein and endogenous protein half-lives is still useful in the purification of cells from the population extremes for comparative kinetic analysis following replating. Nevertheless, it is not surprising that other recent fluorescent protein reporters used in ESCs might not fully recapitulate the endogenous protein (Faddah et al., 2013). This drawback has been tackled by generating fluorescent protein fusions expressed from the two endogenous *Nanog* loci, therefore serving as an allelic reporter at the protein level (Filipczyk et al., 2013). This analysis confirmed heterogeneity at the protein level in individual cells of the population. The half-life of the resulting fusion proteins matches that measured for NANOG protein by the same authors (~5.5 h). However, this does not agree with the half-life of NANOG in other studies of cells cultured in LIF/serum, which was determined to be ~2 h (Chae et al., 2012; Abranches et al., 2013). The origin of this discrepancy is unclear as Filipczyk et al. use the same genetic background and culture conditions as some of these other studies. It is also notable that human NANOG protein examined in a human cell also has a half-life of ~2 h (Ramakrishna et al., 2011). Putting aside these concerns, the Filipczyk study using NANOG fusions provides the best approximation available to date for addressing NANOG protein dynamics in single cells. The range in the relative allelic expression of the two fusion proteins in single cells matches that of the exonic mRNA reported by Hansen and van Oudenaarden (2013), with the exception that the dispersion from the mean is wider for the mRNA than for the protein. This detail implies that a significant parameter accounting for the lack of correlation between monoallelic firing and biallelic protein content may arise from the stability, concentration and/or half-life of the mRNA itself. Indeed,

the number of *Nanog* mRNA molecules present in individual cells ranges from 0 to 500 (median of ~220), with a peak of cells with no or low *Nanog* mRNA (Hansen and van Oudenaarden, 2013).

The apparent lack of correlation between *Nanog* transcriptional firing and the mRNA and protein contents is not unprecedented. Although mRNA content generally correlates with protein levels, there are some important exceptions. Studies both in bacteria and mammalian cells have shown that, whereas the mRNA can show a large degree of variation between cells in a population, the variations at the protein level in the same population can be much smaller (Raj et al., 2006; Sigal et al., 2006). This implies that although variations in the analysis of mRNA levels can be easily seen in a cell population, they will not necessarily translate into variations at the protein level, particularly when the protein is comparatively stable. Interestingly, whether genes for which transcription is activated in bursts (see glossary in Box 1) show a general correlation between promoter activity and mRNA and/or protein levels in individual cells remains an open question. Genome-wide analyses in fission yeast have revealed that whereas changes in mRNA and protein concentrations are concordant for induced genes, genes that are rapidly repressed upon a stimulus, for example stress, can have uncoupled mRNA and protein levels (Lackner et al., 2012). Indeed, discordant half-lives can give apparently uncoupled mRNA and protein levels, particularly for a repressed gene for which the mRNA half-life is shorter than that of the corresponding protein. On a case-by-case basis, cell-to-cell variability in protein level will depend on whether the mRNA and protein have dissimilar half-lives: in cases in which the protein has a

longer half-life than the mRNA, mRNA variability may not translate into protein variability. However, in cases where the protein half-life is shorter than, or similar to, that of its mRNA, variability at the mRNA level will generate cell-to-cell variability at the protein level. The above considerations caution against drawing oversimplified parallels between mRNA production and protein concentrations. This is particularly important for genes subject to pulsatile transcriptional activation (see glossary in Box 1), allele switching or genes that become rapidly repressed in response to signalling or other stimuli.

Non-transcriptional regulation of transcription factor heterogeneity

In addition to the factors discussed above, protein synthesis and other post-transcriptional events may also have a major effect on cellular function and heterogeneity. Studies that measured relative synthesis and degradation rates in C2C12 myotubes have suggested that protein synthesis, rather than mRNA production, can be a dominant factor for cell differentiation (Kristensen et al., 2013). Protein synthesis rates are relatively understudied in models for mammalian cell differentiation or *in vivo*. In particular, the rate of NANOG synthesis has not been addressed in detail. However, ablation of *Gsk3* has been reported to enhance translation of NANOG, while the half-life of NANOG is reduced in cells cultured in 2i/LIF (see glossary in Box 1) compared with LIF/serum, suggesting that protein dynamics might make complex contributions to ESC self-renewal (Sanchez-Ripoll et al., 2013).

Fluctuations in gene expression can be driven by intrinsic and extrinsic factors. Extrinsic fluctuations derived from environmental changes, such as signalling inputs, will affect both alleles. By contrast, intrinsic factors result from randomness inherent to transcription and/or translation. Therefore, intrinsic fluctuations can potentially affect the two alleles of a gene independently (Elowitz et al., 2002; Raj and van Oudenaarden, 2008). The period of inactivity between transcriptional firing has, in some cases, been associated with remodelling of the chromatin (Harper et al., 2011). Indeed, when assessing the transcriptional activation of identical promoters in single cells using two different reporters, the dynamics were shown to be distinct and independent of cell cycle and were proposed to be due to different chromatin configurations. In general, two types of genes could be distinguished: fully activated genes that displayed constant transcriptional activity and those that are activated only in a particular context. Fluctuations in gene expression, known as pulses, are globally inherent to genes that are not fully activated.

Non-transcriptional regulation of heterogeneity might also occur, for example through modulation of signalling pathways or via the cell cycle phase of single cells in a population. Here, it is likely that these extracellular stimuli act through generating changes in transcription. The cell cycle has been shown to affect stochastic transcriptional cycles in cell populations (Zopf et al., 2013). However, opposite results have been obtained using different systems (Harper et al., 2011), implying that the influence of the cell cycle on cell-to-cell variability might differ and that conclusions on cell cycle effects will need to be drawn on a case-by-case basis.

Cell-to-cell variability in mRNA and protein levels in *Saccharomyces cerevisiae* colonies has been explained through differences in the cell cycle, where the phase of the cell cycle dominates extrinsic noise and is the main determinant of heterogeneity within a population of cells (Zopf et al., 2013). This work also postulated that transcriptional burst frequency increases with high expression levels. The impact of the cell cycle on transcription pulses has also been addressed in individual pituitary cells (Harper et al., 2011). The transcriptional bursts of the prolactin gene displayed by a

population of pituitary cells was not coordinated between different cells, and tracking the activity of two different reporters in live imaging allowed the authors to ask whether such bursts were temporally coordinated or out of phase within a single cell cycle. Remarkably, the dynamics of identical promoters in single cells were shown to differ and to be independent of cell cycle. Instead, the period of inactivity between transcriptional firing was associated with remodelling of the chromatin (Harper et al., 2011).

The cell cycle has recently been shown to affect the differentiation capacity of human ESCs (Pauklin and Vallier, 2013). Single human ESCs initiate differentiation towards endoderm or neuroectoderm lineages according to whether they depart from early G1 or late G1 phase, respectively. These results imply that the asynchronous differentiation observed in human ESC cultures might be related to differences in the cell cycle phase between individual cells. However, no correlation between transcription factor expression and cell cycle phasing was noted. Indeed, human ESCs express pluripotency transcription factor (OCT4) or differentiation factors (SOX1, SOX17) in the same proportions as the whole population, independent of their cell cycle stage (Pauklin and Vallier, 2013). Thus, it remains to be seen whether the cell cycle has a direct effect on regulating transcription factor heterogeneity.

Functional significance of transcription factor heterogeneity

Heterogeneity of gene expression might have a functional role in cell fate decisions. This notion was put forward for *in vivo* embryonic development some years ago, and stipulated that stochastic changes in gene expression, including that of *Nanog*, might allow a window of opportunity to direct lineage allocation (Chazaud et al., 2006; Kurimoto et al., 2006; Dietrich and Hiiragi, 2007; Frankenberg et al., 2011). *In vitro*, some models accounting for *Nanog* heterogeneity in ESCs that integrate allele switching have been proposed recently (Wu and Tzanakakis, 2013). Using a multiscale stochastic population balance equation model, it was proposed that allelic control of *Nanog* allows ESCs to restore a characteristic equilibrium population with a constant fraction of the four types of cells discussed above – that is, expressing one or other *Nanog* allele, both, or neither. In this scenario, allelic control of *Nanog* transcription is proposed to be a prime determinant of stem cell population heterogeneity under LIF/serum culture conditions (Wu and Tzanakakis, 2013). When allelic switching was not included in the model parameters, the population reached an equilibrium that did not reflect the experimental data generated by Miyanari and Torres-Padilla (2012) and by Hansen and van Oudenaarden (2013). Therefore, this model not only recapitulates experimental data but, more importantly, surmises that ESCs in any ‘state’ of *Nanog* allelic firing can give rise to ESC populations with the same heterogeneity in terms of *Nanog* expression. It will be important to test this prediction experimentally: if allele switching is prevented at the *Nanog* loci, do ESC populations become more homogeneous? In ESCs grown in 2i/LIF, allele switching is not expected to occur (Miyanari and Torres-Padilla, 2012). Therefore, the fact that ESCs cultured in 2i/LIF express NANOG more homogeneously (Ying et al., 2008; Wray et al., 2011) might provide some support to the prediction of the model. Moreover, NANOG protein itself plays a central role in the emergence of a heterogeneous *Nanog* expression profile. Indeed, *Nanog* null ESC populations carrying a Nanog:GFP reporter allele have a reduced ability to generate *Nanog* heterogeneity, as reported by GFP expression. Interestingly, this is the case regardless of whether NANOG activity is maintained in either an ON or OFF state (Navarro et al., 2012).

Pulsatile transcriptional activity can confer robustness

At the molecular level, transcription of a gene is a discrete event defined by the assembly of the pre-initiation complex and the binding of one RNA polymerase II (RNAPII) molecule to the promoter of a gene, followed by promoter melting, initial elongation to a site of promoter proximal pausing, and subsequent processive elongation (Core and Lis, 2008; Brookes and Pombo, 2009). Thus, transcription occurs in molecular steps, or pulses, rather than in a continuous fashion. When the rate of pre-initiation complex and RNAPII binding and processivity is high, transcription of a gene becomes a sustained event, with only short periods of inactivity between the production of transcripts from the given gene. Often, however, transcription of genes is a rather discontinuous event, with only a few pulses occurring over a discrete time period (Muramoto et al., 2012).

Genes that pulse can display different probabilities of firing, differences in the duration of the pulse, the frequency and the intensity. A sustained, non-binary transcriptional activity might be better suited for housekeeping functions (Muramoto et al., 2012). By contrast, stochastic switch behaviour might allow room for developmental decisions, an idea supported by studies from both *Dictyostelium* and *Drosophila* (Muramoto et al., 2012; Little et al., 2013). In *Drosophila* embryos, developmental RNAs analysed by smRNA-FISH displayed differences of up to 44% in expression between neighbouring nuclei (Little et al., 2013). Importantly, the variation between cells was about 6-fold higher for nascent transcripts than for cytoplasmic transcripts. Thus, similar to observations from smRNA-FISH of *Nanog* in ESCs (Hansen and van Oudenaarden, 2013), cytoplasmic mRNA measures have considerably less noise than nascent transcription site counts. In the *Drosophila* study, the authors concluded that a robust and precise temporal expression pattern in the early syncytium is achieved through stochastic decisions in transcription at the single-cell level (Li and Xie, 2011; Little et al., 2013). More recent evidence using quantitative live nascent RNA imaging in *Drosophila* further suggests that developmental patterns arise as a result of buffering (see glossary in Box 1) single-cell expression patterns over time and space, and not because cells display similar transcriptional activity as a whole (Garcia et al., 2013). Therefore, allelic firing might represent a way in which stochasticity is generated at the single-cell level in a specific developmental context. Transcriptional burst frequency is expected to increase with high expression levels (Zopf et al., 2013). Of note for *Nanog*, higher expression levels in ESCs cultured in 2i/LIF correlate with a higher burst frequency and with the probability of both alleles firing (Miyanari and Torres-Padilla, 2012; Hansen and van Oudenaarden, 2013). If transcription occurs with bursts or as discontinuous firing events, a high variability in the number of mRNA molecules produced over time in individual cells would be expected (Raj et al., 2006; Raj and van Oudenaarden, 2008). Thus, the experimental data documenting high variability in the number of *Nanog* mRNA molecules per cell fit well with a mode of regulation for *Nanog* involving discontinuous firing. The data of several studies (Miyanari and Torres-Padilla, 2012; Navarro et al., 2012; Hansen and van Oudenaarden, 2013) provide evidence for such bursts and an experimental framework to explain such heterogeneity, since at a given time allele firing occurs only in a proportion of the cell population. In a fluctuation regulatory model, a positive-feedback mechanism can amplify such fluctuations (Little et al., 2013). The positive-feedback loops within the pluripotency GRN discussed above are in accord with this type of regulation. Furthermore, since NANOG autoregulates its own expression in a negative-feedback mechanism (Navarro et al., 2012), this might provide a means for

the cell to limit the extent of fluctuations and for buffering cell-to-cell variations in protein concentrations. Importantly, the autorepression of NANOG is independent of OCT4 and SOX2 activity, which may provide a means to mechanistically separate the autoregulatory loop involved in maintaining *Nanog* heterogeneity from the regulation of *Oct4* and *Sox2* (Navarro et al., 2012). Thus, it is possible that a regulated balance of positive and negative feedback might be necessary to maintain the steady-state equilibrium of *Nanog* expression over time in a given ESC population as a whole.

In microbial populations, it has been known for some time that competence for stress response entails phenotypic variability (Nester and Stocker, 1963). A relevant question is, therefore, whether heterogeneity would be of any advantage for a stem cell population. Here, it would be useful to learn from what has been inferred in other systems that are also dependent on signalling. Indeed, environmental signals control the pluripotent cell type captured *in vitro* at the time of cell isolation from the embryo and dramatically influence the type of pluripotent cells established (Festuccia et al., 2013). A model for pulsatile stimulation has been described for the signalling response to NF- κ B, whereby a population of cells shows a more efficient response in downstream transcriptional activation to repeated pulses of TNF α stimulation as opposed to a continuous stimulus (Ashall et al., 2009). A broader conclusion from these observations is that the cellular heterogeneity that arises from pulsatile transcriptional responses confers robustness to a cell population (Paszek et al., 2010). In other words, generating enhanced cell heterogeneity at the single-cell level can control the stability of a population of cells by decreasing fluctuations over time at the population level. This might confer an advantage by both preserving the identity of the population and simultaneously enabling the population to respond to signalling stimuli (Paszek et al., 2010). In this scenario, fluctuating gene expression may enable a faster response to signalling.

Heterogeneity as a means to impart developmental robustness

An alternative, non-mutually exclusive view of stem cell heterogeneity based on ideas from statistical mechanics reaches a similar conclusion: that fluctuations in single-cell properties 'prime' pluripotent cells to respond to certain cues important for differentiation (MacArthur and Lemischka, 2013). According to this model, a pluripotent population can be robustly maintained without there being a fixed definition of the pluripotent state at the molecular level in single cells.

Put into a broader context, development can be considered an intrinsically noisy system due to fluctuations in transcriptional regulation. We must not forget that ESCs are, after all, a developmental model system and it is therefore expected that ESCs that are not locked in a self-renewing state by pharmacological inhibition retain such a characteristic fluctuating behaviour. RNA-FISH (Miyanari and Torres-Padilla, 2012) and genome-wide analysis of allele-specific transcripts (Deng et al., 2014) revealed that at the early stages of development, between the 4- and the 8-cell stage, cells fire primarily one *Nanog* allele, presumably generating a large degree of variability among cells (Dietrich and Hiiragi, 2007), which might be potentiated by the generation of secondary heterogeneities. Following the above considerations, cell-to-cell fluctuations can be buffered across time and space through translation and protein degradation rates. In an ESC population, even if individual cells change their expression of *Nanog*, for example through allele firing, the population reaches an equilibrium over time, which is that of a typical 'snapshot' observed when we perform static experiments to address gene expression in single cells by RNA-FISH, smRNA-FISH, immunostaining or FACS.

Together, these inputs provide a pluripotent population with the robustness necessary to self-renew under the appropriate culture conditions, but, importantly, with the simultaneous responsiveness required to allow a rapid transition into differentiation both *in vitro* and *in vivo*.

Fluctuations in *Nanog* expression are extremely sensitive to alterations in signalling (Luo et al., 2012; Karwacki-Neisius et al., 2013), and this most likely reflects the native signalling environment in the embryo around the time when ESCs are derived. Lineage choice between the epiblast and the primitive endoderm is governed by fibroblast growth factor (FGF) and downstream signalling activity (Cheng et al., 1998; Chazaud et al., 2006; Nichols et al., 2009; Yamanaka et al., 2010; Frankenberg et al., 2011). A pulsatile or burst reaction to activate or repress *Nanog* expression might allow an efficient allocation towards these two lineages by allowing individual cells within the pluripotent population to portray a different response, in spite of being submerged in the same signalling environment. Responsiveness to FGF is required for ESCs to generate *Nanog* heterogeneity and initiate differentiation. If FGF signalling is blocked, ESCs cannot demonstrate the pluripotency crucial for normal development. Whether ESCs cultured in pharmacological inhibitors or in LIF/serum reflect a precise stage in the peri-implantation development programme is presently unclear. Indeed, the expression of genes characteristic of the totipotent 2-cell stage blastomeres in ESCs cultured in LIF/serum or LIF/2i suggests that neither culture regime faithfully and uniformly recapitulates a specific point in developmental time (Macfarlan et al., 2012). Therefore, caution should be exercised when extrapolating the behaviour of ESCs grown in either condition with the behaviour of pluripotent cells of the embryo. The above scenarios require a NANOG threshold dictating permissiveness for differentiation versus self-renewal, which has not only been suggested in the literature (Chambers et al., 2003, 2007; Silva et al., 2009; Frankenberg et al., 2011; Silva and Smith, 2008) but without which normal development does not occur (Nichols et al., 2009). Note that although *Nanog*^{+/-} mice develop to term, the formation of inner cell mass derivatives is delayed (Miyazari and Torres-Padilla, 2012), suggesting that in the presence of only one allele the epiblast acquires full functionality at a later timepoint.

It is important to consider whether heterogeneity in *Nanog* expression affects either the transcriptional or epigenetic status of ESC populations (Ficz et al., 2013). As mentioned above, cells cultured in 2i that fire transcription from two alleles for most of the time and have a more homogeneous distribution of NANOG and other pluripotency transcription factors (Ying et al., 2008; Kalmar et al., 2009; Wray et al., 2011; Miyazari and Torres-Padilla, 2012; Hansen and van Oudenaarden, 2013) do have a different gene expression profile as a population (Marks et al., 2012; Leitch et al., 2013). Moreover, ESCs cultured in 2i display global differences in chromatin and have globally hypomethylated DNA (Ficz et al., 2013; Habibi et al., 2013; Leitch et al., 2013). Whether these epigenetic changes occur homogeneously throughout the population remains to be seen.

Perspectives and conclusions

Heterogeneity between individual cells is an inherent feature of dynamic cellular processes. The known heterogeneity of transcription factor expression that we have discussed in ESCs suggests that the transcriptional control of ESC fate is a dynamic process. Therefore, pluripotency should be seen as a dynamic stochastic process that can potentially impart a benefit to a population of cells in the process of undergoing cell fate decisions in response to signalling.

We have discussed transcription factor heterogeneity in general and some of the molecular principles that may underlie fluctuations in *Nanog* expression in ESCs in particular. Allelic switching provides a framework for such variability, and is most likely potentiated by changes in mRNA stability and distribution in single cells. Although other genes have also been reported to be expressed heterogeneously in ESCs (Hayashi et al., 2008; Toyooka et al., 2008; Festuccia and Chambers, 2011; Morgani et al., 2013), most of the data described above are based on findings related to *Nanog*, perhaps because *Nanog* is the most extensively studied. Whether *Nanog* is a unique case in stem cell regulation will be important to establish. Furthermore, it will be interesting to explore whether the recently described naïve human ESCs (Gafni et al., 2013; Chan et al., 2013) display fluctuations in *Nanog* expression.

Mathematical modelling approaches have started to contribute to our understanding of stochastic gene expression and its impact on the stability and heterogeneity of cell populations. Although such analyses allow the projection of population states along time and the testing of some of the known molecular regulatory parameters (provided that a sufficient amount of quantitative data is available), they do not predict unknown parameters. For example, we cannot infer whether there is any influence from individual components of the pluripotency GRN in maintaining and generating heterogeneity. How is this type of analysis affected by the negative autoregulation that NANOG exerts on *Nanog*? Is the largely heterogeneous distribution of *Nanog* mRNA observed at the single-cell level relevant? What is the effect of the cell cycle, if any, on ESC heterogeneity? Although re-establishment of a steady-state heterogeneous population has been documented, the time taken for re-equilibration varies (Chambers et al., 2007; Kalmar et al., 2009; Abranches et al., 2013). Some of these differences might be due to non-matching half-lives, but, as mentioned earlier, the purity of the sorted population is crucial here. Similar to experiments performed with haematopoietic cells (Pina et al., 2012), future experiments should address the re-equilibration ability of single cells obtained from the extremes of the distribution. Imaging data with destabilised fluorescent protein reporters acquired over several generations under relevant culture conditions should also provide a more comprehensive understanding of how *Nanog* transcriptional activity affects heterogeneity. Moreover, compiling data on NANOG protein concentrations, fluctuations and synthesis rate should also be considered in the future. This will help to determine whether there is a key molecular event underlying heterogeneity in ESCs or, rather, whether heterogeneity results from a combined effect of each of the steps discussed above.

Much remains to be learned regarding the molecular control of heterogeneous transcription factor expression in pluripotent cell types (Cahan and Daley, 2013). It will be instructive to examine the multiple levels of regulation of protein and RNA heterogeneity and the degree to which heterogeneity can be brought under control by endogenous signalling proteins. Approaches to investigate gene regulation in single cells will be key to these analyses. Much of the information currently available about pluripotent cell gene regulation has come from the advances brought about by the deployment of powerful high-throughput sequencing approaches to measure gene expression and the chromatin localisation of transcription factors, chromatin modifiers and histone modifications in a given population of cells. However, from the arguments raised above, it is clear that in order to understand pluripotency it is necessary to study individual cells rather than population averages. In particular, it is crucial to remember that the most significant landmark discoveries in stem cell biology have assessed the functional capabilities of single cells (Siminovitch et al., 1963; Kleinsmith and Pierce, 1964). Therefore, techniques that

examine the molecular details within individual cells are essential to a proper understanding of stem cell biology.

We anticipate that the regulatory principles that apply to pluripotent cells will be relevant to the control of heterogeneous physiological responses in other stem cell systems and possibly also to understanding how regulatory control is lost in some pathological conditions. As such, a detailed knowledge of the regulation of *Nanog*, and an understanding of how NANOG activity delivers enhanced self-renewal, will provide important parallels for other, less biochemically tractable stem cell systems.

Acknowledgements

We thank members of our labs for critical comments on the manuscript.

Competing interests

The authors declare no competing financial interests

Funding

Work in the M.-E.T.-P. laboratory is funded through the ANR-09-Blanc-0114, EpiGeneSys NoE, ERC-Stg 'NuclearPotency' and the FP7 Marie-Curie Actions ITN Nucleosome4D; work in the I.C. laboratory is funded by the Medical Research Council of the U.K.

References

- Abranches, E., Bekman, E. and Henrique, D. (2013). Generation and characterization of a novel mouse embryonic stem cell line with a dynamic reporter of Nanog expression. *PLoS ONE* **8**, e59928.
- Ambrosetti, D.-C., Basilico, C. and Dailey, L. (1997). Synergistic activation of the fibroblast growth factor 4 enhancer by *sox2* and *oct-3* depends on protein-protein interactions facilitated by a specific spatial arrangement of factor binding sites. *Mol. Cell. Biol.* **17**, 6321-6329.
- Ashall, L., Horton, C. A., Nelson, D. E., Paszek, P., Harper, C. V., Sillitoe, K., Ryan, S., Spiller, D. G., Unitt, J. F., Broomhead, D. S. et al. (2009). Pulsatile stimulation determines timing and specificity of NF-kappaB-dependent transcription. *Science* **324**, 242-246.
- Beddington, R. S. (1983). Histogenetic and neoplastic potential of different regions of the mouse embryonic egg cylinder. *J. Embryol. Exp. Morphol.* **75**, 189-204.
- Brons, I. G. M., Smithers, L. E., Trotter, M. W. B., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S. M., Howlett, S. K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R. A. et al. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448**, 191-195.
- Brook, F. A. and Gardner, R. L. (1997). The origin and efficient derivation of embryonic stem cells in the mouse. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 5709-5712.
- Brookes, E. and Pombo, A. (2009). Modifications of RNA polymerase II are pivotal in regulating gene expression states. *EMBO Rep.* **10**, 1213-1219.
- Cahan, P. and Daley, G. Q. (2013). Origins and implications of pluripotent stem cell variability and heterogeneity. *Nat. Rev. Mol. Cell Biol.* **14**, 357-368.
- Canham, M. A., Sharov, A. A., Ko, M. S. H. and Brickman, J. M. (2010). Functional heterogeneity of embryonic stem cells revealed through translational amplification of an early endodermal transcript. *PLoS Biol.* **8**, e1000379.
- Chae, H.-D., Lee, M.-R. and Broxmeyer, H. E. (2012). 5-Aminoimidazole-4-carboxamide ribonucleoside induces G(1)/S arrest and Nanog downregulation via p53 and enhances erythroid differentiation. *Stem Cells* **30**, 140-149.
- Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S. and Smith, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**, 643-655.
- Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L. and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. *Nature* **450**, 1230-1234.
- Chan, Y.-S., Göke, J., Ng, J.-H., Lu, X., Gonzales, K. A., Tan, C.-P., Tng, W.-Q., Hong, Z.-Z., Lim, Y.-S. and Ng, H.-H. (2013). Induction of a human pluripotent state with distinct regulatory circuitry that resembles preimplantation epiblast. *Cell Stem Cell* **13**, 663-675.
- Chazaud, C., Yamanaka, Y., Pawson, T. and Rossant, J. (2006). Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. *Dev. Cell* **10**, 615-624.
- Chen, X., Fang, F., Liou, Y.-C. and Ng, H.-H. (2008). Zfp143 regulates Nanog through modulation of Oct4 binding. *Stem Cells* **26**, 2759-2767.
- Cheng, A. M., Saxton, T. M., Sakai, R., Kulkarni, S., Mbamalu, G., Vogel, W., Tortorice, C. G., Cardiff, R. D., Cross, J. C., Muller, W. J. et al. (1998). Mammalian Grb2 regulates multiple steps in embryonic development and malignant transformation. *Cell* **95**, 793-803.
- Core, L. J. and Lis, J. T. (2008). Transcription regulation through promoter-proximal pausing of RNA polymerase II. *Science* **319**, 1791-1792.
- Dan, J., Li, M., Yang, J., Li, J., Okuka, M., Ye, X. and Liu, L. (2013). Roles for Tbx3 in regulation of two-cell state and telomere elongation in mouse ES cells. *Sci. Rep.* **3**, 3492.
- Deng, Q., Ramskold, D., Reinius, B. and Sandberg, R. (2014). Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science* **343**, 193-196.
- Dietrich, J.-E. and Hiragi, T. (2007). Stochastic patterning in the mouse pre-implantation embryo. *Development* **134**, 4219-4231.
- Elowitz, M. B., Levine, A. J., Siggia, E. D. and Swain, P. S. (2002). Stochastic gene expression in a single cell. *Science* **297**, 1183-1186.
- Evans, M. J. and Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292**, 154-156.
- Faddah, D. A., Wang, H., Cheng, A. W., Katz, Y., Buganim, Y. and Jaenisch, R. (2013). Single-cell analysis reveals that expression of *nanog* is biallelic and equally variable as that of other pluripotency factors in mouse ESCs. *Cell Stem Cell* **13**, 23-29.
- Festuccia, N. and Chambers, I. (2011). Quantification of pluripotency transcription factor levels in embryonic stem cells by flow cytometry. *Curr. Protoc. Stem Cell Biol.* Chapter 1, Unit 1B 9.
- Festuccia, N., Osorno, R., Halbritter, F., Karwacki-Neisius, V., Navarro, P., Colby, D., Wong, F., Yates, A., Tomlinson, S. R. and Chambers, I. (2012). *Esrrb* is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells. *Cell Stem Cell* **11**, 477-490.
- Festuccia, N., Osorno, R., Wilson, V. and Chambers, I. (2013). The role of pluripotency gene regulatory network components in mediating transitions between pluripotent cell states. *Curr. Opin. Genet. Dev.* **23**, 504-511.
- Ficz, G., Hore, T. A., Santos, F., Lee, H. J., Dean, W., Arand, J., Krueger, F., Oxley, D., Paul, Y.-L., Walter, J. et al. (2013). FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. *Cell Stem Cell* **13**, 351-359.
- Fidalgo, M., Faiola, F., Pereira, C.-F., Ding, J., Saunders, A., Gingold, J., Schaniel, C., Lemischka, I. R., Silva, J. C. R. and Wang, J. (2012). Zfp281 mediates Nanog autorepression through recruitment of the NuRD complex and inhibits somatic cell reprogramming. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 16202-16207.
- Filipczyk, A., Gkatzis, K., Fu, J., Hoppe, P. S., Lickert, H., Anastassiadis, K. and Schroeder, T. (2013). Biallelic expression of nanog protein in mouse embryonic stem cells. *Cell Stem Cell* **13**, 12-13.
- Frankenberg, S., Gerbe, F., Bessonard, S., Belville, C., Pouchin, P., Bardot, O. and Chazaud, C. (2011). Primitive endoderm differentiates via a three-step mechanism involving Nanog and RTK signaling. *Dev. Cell* **21**, 1005-1013.
- Gafni, O., Weinberger, L., Mansour, A. A., Manor, Y. S., Chomsky, E., Ben-Yosef, D., Kalma, Y., Viukov, S., Maza, I., Zviran, A. et al. (2013). Derivation of novel human ground state naive pluripotent stem cells. *Nature* **504**, 282-286.
- Garcia, H. G., Tikhonov, M., Lin, A. and Gregor, T. (2013). Quantitative imaging of transcription in living Drosophila embryos links polymerase activity to patterning. *Curr. Biol.* **23**, 2140-2145.
- Guo, G., Huss, M., Tong, G. Q., Wang, C., Li Sun, L., Clarke, N. D. and Robson, P. (2010). Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. *Dev. Cell* **18**, 675-685.
- Habibi, E., Brinkman, A. B., Arand, J., Kroeze, L. I., Kerstens, H. H., Matarese, F., Lepikhov, K., Gut, M., Brun-Heath, I., Hubner, N. C. et al. (2013). Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. *Cell Stem Cell* **13**, 360-369.
- Hansen, C. H. and van Oudenaarden, A. (2013). Allele-specific detection of single mRNA molecules in situ. *Nat. Methods* **10**, 869-871.
- Harper, C. V., Finkstädt, B., Woodcock, D. J., Friedrichsen, S., Semprini, S., Ashall, L., Spiller, D. G., Mullins, J. J., Rand, D. A., Davis, J. R. E. et al. (2011). Dynamic analysis of stochastic transcription cycles. *PLoS Biol.* **9**, e1000607.
- Hatano, S.-y., Tada, M., Kimura, H., Yamaguchi, S., Kono, T., Nakano, T., Suemori, H., Nakatsuji, N. and Tada, T. (2005). Pluripotential competence of cells associated with Nanog activity. *Mech. Dev.* **122**, 67-79.
- Hayashi, K., Chuva de Sousa 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.
- Jaenisch, R. and Young, R. (2008). Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell* **132**, 567-582.
- Kalmar, T., Lim, C., Hayward, P., Muñoz-Descalzo, S., Nichols, J., Garcia-Ojalvo, J. and Martinez Arias, A. (2009). Regulated fluctuations in nanog expression mediate cell fate decisions in embryonic stem cells. *PLoS Biol.* **7**, e1000149.
- Karwacki-Neisius, V., Göke, J., Osorno, R., Halbritter, F., Ng, J. H., Weisse, A. Y., Wong, F. C., Gagliardi, A., Mullin, N. P., Festuccia, N. et al. (2013). Reduced Oct4 expression directs a robust pluripotent state with distinct signaling activity and increased enhancer occupancy by Oct4 and Nanog. *Cell Stem Cell* **12**, 531-545.
- Kim, J., Chu, J., Shen, X., Wang, J. and Orkin, S. H. (2008). An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* **132**, 1049-1061.

- Kleinsmith, L. J. and Pierce, G. B. (1964). Multipotentiality of single embryonal carcinoma cells. *Cancer Res.* **24**, 1544-1552.
- Kristensen, A. R., Gsponer, J. and Foster, L. J. (2013). Protein synthesis rate is the predominant regulator of protein expression during differentiation. *Mol. Syst. Biol.* **9**, 689.
- Kurimoto, K., Yabuta, Y., Ohinata, Y., Ono, Y., Uno, K. D., Yamada, R. G., Ueda, H. R. and Saitou, M. (2006). An improved single-cell cDNA amplification method for efficient high-density oligonucleotide microarray analysis. *Nucleic Acids Res.* **34**, e42.
- Lackner, D. H., Schmidt, M. W., Wu, S., Wolf, D. A. and Bahler, J. (2012). Regulation of transcriptome, translation, and proteome in response to environmental stress in fission yeast. *Genome Biol.* **13**, R25.
- Leitch, H. G., McEwen, K. R., Turp, A., Encheva, V., Carroll, T., Grabole, N., Mansfield, W., Nashun, B., Knezovich, J. G., Smith, A. et al. (2013). Naive pluripotency is associated with global DNA hypomethylation. *Nat. Struct. Mol. Biol.* **20**, 311-316.
- Levine, M. and Davidson, E. H. (2005). Gene regulatory networks for development. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 4936-4942.
- Li, G.-W. and Xie, X. S. (2011). Central dogma at the single-molecule level in living cells. *Nature* **475**, 308-315.
- Little, S. C., Tikhonov, M. and Gregor, T. (2013). Precise developmental gene expression arises from globally stochastic transcriptional activity. *Cell* **154**, 789-800.
- Loh, Y.-H., Wu, Q., Chew, J.-L., Vega, V. B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J. et al. (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat. Genet.* **38**, 431-440.
- Luo, Y., Lim, C. L., Nichols, J., Martinez-Arias, A. and Wernisch, L. (2012). Cell signalling regulates dynamics of Nanog distribution in embryonic stem cell populations. *J. R. Soc. Interface* **10**, 20120525.
- MacArthur, B. D. and Lemischka, I. R. (2013). Statistical mechanics of pluripotency. *Cell* **154**, 484-489.
- MacArthur, B. D., Sevilla, A., Lenz, M., Müller, F.-J., Schuld, B. M., Schuppert, A. A., Ridden, S. J., Stumpf, P. S., Fidalgo, M., Ma'ayan, A. et al. (2012). Nanog-dependent feedback loops regulate murine embryonic stem cell heterogeneity. *Nature Cell Biol.* **14**, 1139-1147.
- Macfarlan, T. S., Gifford, W. D., Agarwal, S., Driscoll, S., Lettieri, K., Wang, J., Andrews, S. E., Franco, L., Rosenfeld, M. G., Ren, B. et al. (2011). Endogenous retroviruses and neighboring genes are coordinately repressed by LSD1/KDM1A. *Genes Dev.* **25**, 594-607.
- Macfarlan, T. S., Gifford, W. D., Driscoll, S., Lettieri, K., Rowe, H. M., Bonanomi, D., Firth, A., Singer, O., Trono, D. and Pfaff, S. L. (2012). Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487**, 57-63.
- Marks, H., Kalkan, T., Menafra, R., Denisov, S., Jones, K., Hofmeister, H., Nichols, J., Kranz, A., Stewart, A. F., Smith, A. et al. (2012). The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* **149**, 590-604.
- Martin, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **78**, 7634-7638.
- Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M. and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**, 631-642.
- Miyazari, Y. and Torres-Padilla, M.-E. (2012). Control of ground-state pluripotency by allelic regulation of Nanog. *Nature* **483**, 470-473.
- Morgani, S. M., Canham, M. A., Nichols, J., Sharov, A. A., Migueles, R. P., Ko, M. S. H. and Brickman, J. M. (2013). Totipotent embryonic stem cells arise in ground-state culture conditions. *Cell Rep.* **3**, 1945-1957.
- Muramoto, T., Cannon, D., Gierlinski, M., Corrigan, A., Barton, G. J. and Chubb, J. R. (2012). Live imaging of nascent RNA dynamics reveals distinct types of transcriptional pulse regulation. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 7350-7355.
- Najm, F. J., Chenoweth, J. G., Anderson, P. D., Nadeau, J. H., Redline, R. W., McKay, R. D. G. and Tesar, P. J. (2011). Isolation of epiblast stem cells from preimplantation mouse embryos. *Cell Stem Cell* **8**, 318-325.
- Navarro, P., Festuccia, N., Colby, D., Gagliardi, A., Mullin, N. P., Zhang, W., Karwacki-Neisius, V., Osorno, R., Kelly, D., Robertson, M. et al. (2012). OCT4/SOX2-independent Nanog autorepression modulates heterogeneous Nanog gene expression in mouse ES cells. *EMBO J.* **31**, 4547-4562.
- Nester, E. W. and Stocker, B. A. (1963). Biosynthetic latency in early stages of deoxyribonucleic acid transformation in bacillus subtilis. *J. Bacteriol.* **86**, 785-796.
- Ng, H.-H. and Surani, M. A. (2011). The transcriptional and signalling networks of pluripotency. *Nat. Cell Biol.* **13**, 490-496.
- Nichols, J., Silva, J., Roode, M. and Smith, A. (2009). Suppression of Erk signalling promotes ground state pluripotency in the mouse embryo. *Development* **136**, 3215-3222.
- Niwa, H., Ogawa, K., Shimosato, D. and Adachi, K. (2009). A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. *Nature* **460**, 118-122.
- Osorno, R., Tsakiridis, A., Wong, F., Cambrey, 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.
- Paszek, P., Ryan, S., Ashall, L., Sillitoe, K., Harper, C. V., Spiller, D. G., Rand, D. A. and White, M. R. (2010). Population robustness arising from cellular heterogeneity. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11644-11649.
- Pauklin, S. and Vallier, L. (2013). The cell-cycle state of stem cells determines cell fate propensity. *Cell* **155**, 135-147.
- Pina, C., Fugazza, C., Tipping, A. J., Brown, J., Soneji, S., Teles, J., Peterson, C. and Enver, T. (2012). Inferring rules of lineage commitment in haematopoiesis. *Nat. Cell Biol.* **14**, 287-294.
- Ptashne, M. and Gann, A. (2001). Transcription initiation: imposing specificity by localization. *Essays Biochem.* **37**, 1-15.
- Raj, A. and van Oudenaarden, A. (2008). Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* **135**, 216-226.
- Raj, A., Peskin, C. S., Tranchina, D., Vargas, D. Y. and Tyagi, S. (2006). Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.* **4**, e309.
- Ramakrishna, S., Suresh, B., Lim, K.-H., Cha, B.-H., Lee, S.-H., Kim, K.-S. and Baek, K.-H. (2011). PEST motif sequence regulating human NANOG for proteasomal degradation. *Stem Cells Dev.* **20**, 1511-1519.
- Sanchez-Ripoll, Y., Bone, H. K., Owen, T., Guedes, A. M. V., Abranches, E., Kumpfmüller, B., Spriggs, R. V., Henrique, D. and Welham, M. J. (2013). Glycogen synthase kinase-3 inhibition enhances translation of pluripotency-associated transcription factors to contribute to maintenance of mouse embryonic stem cell self-renewal. *PLoS ONE* **8**, e60148.
- Sigal, A., Milo, R., Cohen, A., Geva-Zatorsky, N., Klein, Y., Liron, Y., Rosenfeld, N., Danon, T., Perzov, N. and Alon, U. (2006). Variability and memory of protein levels in human cells. *Nature* **444**, 643-646.
- Silva, J. and Smith, A. (2008). Capturing pluripotency. *Cell* **132**, 532-536.
- Silva, J., Nichols, J., Theunissen, T. W., Guo, G., van Oosten, A. L., Barrandon, O., Wray, J., Yamanaka, S., Chambers, I. and Smith, A. (2009). Nanog is the gateway to the pluripotent ground state. *Cell* **138**, 722-737.
- Siminovitch, L., McCulloch, E. A. and Till, J. E. (1963). The distribution of colony-forming cells among spleen colonies. *J. Cell. Physiol.* **62**, 327-336.
- 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. G. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* **448**, 196-199.
- Toyooka, Y., Shimosato, D., Murakami, K., Takahashi, K. and Niwa, H. (2008). Identification and characterization of subpopulations in undifferentiated ES cell culture. *Development* **135**, 909-918.
- van den Berg, D. L. C., Zhang, W., Yates, A., Engelen, E., Takacs, K., Bezstarosti, K., Demmers, J., Chambers, I. and Poët, R. A. (2008). The Estrogen Related Receptor Beta interacts with Oct4 to positively regulate Nanog gene expression. *Mol. Cell. Biol.* **28**, 5986-5995.
- Williams, D. C., Jr, Cai, M. and Clore, G. M. (2004). Molecular basis for synergistic transcriptional activation by Oct1 and Sox2 revealed from the solution structure of the 42-kDa Oct1 · Sox2 · Hoxb1-DNA ternary transcription factor complex. *J. Biol. Chem.* **279**, 1449-1457.
- Wray, J., Kalkan, T., Gomez-Lopez, S., Eckardt, D., Cook, A., Kemler, R. and Smith, A. (2011). Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. *Nat. Cell Biol.* **13**, 838-845.
- Wu, J. and Tzanakakis, E. S. (2013). Distinct allelic patterns of nanog expression impart embryonic stem cell population heterogeneity. *PLoS Comput. Biol.* **9**, e1003140.
- Yamanaka, Y., Lanner, F. and Rossant, J. (2010). FGF signal-dependent segregation of primitive endoderm and epiblast in the mouse blastocyst. *Development* **137**, 715-724.
- Ying, Q.-L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P. and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519-523.
- Zalzman, M., Falco, G., Sharova, L. V., Nishiyama, A., Thomas, M., Lee, S.-L., Stagg, C. A., Hoang, H. G., Yang, H.-T., Indig, F. E. et al. (2010). Zscan4 regulates telomere elongation and genomic stability in ES cells. *Nature* **464**, 858-863.
- Zopf, C. J., Quinn, K., Zeidman, J. and Maheshri, N. (2013). Cell-cycle dependence of transcription dominates noise in gene expression. *PLoS Comput. Biol.* **9**, e1003161.