© 2015. Published by The Company of Biologists Ltd | Development (2015) 142, 3468-3477 doi:10.1242/dev.123950

RESEARCH ARTICLE

STEM CELLS AND REGENERATION

Dynamic transcriptional symmetry-breaking in pre-implantation mammalian embryo development revealed by single-cell RNA-seq

Junchao Shi[1,2,*], Qi Chen[1,3,*,‡], Xin Li[1,*], Xiudeng Zheng[1,*], Ying Zhang[1], Jie Qiao[4], Fuchou Tang[4], Yi Tao[1,‡], Qi Zhou[1,‡] and Enkui Duan[1,‡]

ABSTRACT

During mammalian pre-implantation embryo development, when the first asymmetry emerges and how it develops to direct distinct cell fates remain longstanding questions. Here, by analyzing single-blastomere transcriptome data from mouse and human pre-implantation embryos, we revealed that the initial blastomere-to-blastomere biases emerge as early as the first embryonic cleavage division, following a binomial distribution pattern. The subsequent zygotic transcriptional activation further elevated overall blastomere-to-blastomere biases during the two- to 16-cell embryo stages. The trends of transcriptional asymmetry fell into two distinct patterns: for some genes, the extent of asymmetry was minimized between blastomeres (monostable pattern), whereas other genes, including those known to be lineage specifiers, showed ever-increasing asymmetry between blastomeres (bistable pattern), supposedly controlled by negative or positive feedbacks. Moreover, our analysis supports a scenario in which opposing lineage specifiers within an early blastomere constantly compete with each other based on their relative ratio, forming an inclined 'lineage strength' that pushes the blastomere onto a predisposed, yet flexible, lineage track before morphological distinction.

KEY WORDS: Transcriptional symmetry-breaking, Pre-implantation embryo development, Lineage divergence, Monostable model, Bistable model

INTRODUCTION

Mammalian pre-implantation embryos provide a unique opportunity to study how a single cell diverges from a totipotent state into different fates within limited rounds of cell division (Zernicka-Goetz et al., 2009). Morphologically, distinct cell lineage patterning first emerges during the fourth cleavage division to generate a 16-cell embryo. At this stage, some blastomeres divide symmetrically, contributing two daughter cells to the outside region of the embryo, whereas others divide asymmetrically and contribute one daughter cell to the outside and another to the inside region (Rossant and Tam, 2009). This difference in cell allocation has long been considered to be the first emergence of asymmetry in early embryo development; the 'inside cells' will contribute to the inner cell mass (ICM), whereas the 'outside cells' will contribute to the trophectoderm (TE) (Rossant and Tam, 2009; Takaoka and Hamada, 2012; Zernicka-Goetz et al., 2009).

However, contrary to this 'late-deciding' model, recent emerging evidence from studies on morphologically indistinguishable blastomeres before the 16-cell stage has revealed distinctive molecular markers in blastomeres during the four- to eight-cell stages that are indicative of future lineage segregations, leading to the idea that the late-onset event of 'inside' and 'outside' cell divergence is not a random choice but rather an anticipated result rooted in the embryonic division history (Torres-Padilla et al., 2007). As early as the four-cell embryo stage, different levels of CARM1-mediated histone 3 arginine methylation can be found in blastomeres, with a higher level of methylation directed to the ICM lineages (Torres-Padilla et al., 2007). At the four- to eight-cell stage, two distinct patterns of OCT4 (POU5F1) kinetics are observed between blastomeres, reflecting the different levels of OCT4 binding affinity to DNA, with the slower OCT4 kinetics (higher binding ability to DNA) tending to give rise to the ICM (Plachta et al., 2011). Further evidence obtained using the state-of-the-art Rainbow lineage-tracing system has shown that four-cell blastomeres do indeed display developmental biases in contributing to either the ICM or TE (Tabansky et al., 2013). These advancements have promoted the concept that the molecular differences in blastomeres emerge much earlier (as early as the four-cell stage) than the morphological clue.

Despite these observations, the open question remains regarding when the very first molecular asymmetry emerges, and how such initial asymmetry develops with the potential to direct lineage bifurcation. The emerging technology of single-cell RNA-seq has provided the opportunity to study the transcriptome-wide dynamics of the transcriptional symmetry-breaking process during pre-implantation mammalian development. In the present study, combining single-blastomere transcriptome analysis from human and mouse pre-implantation embryos (Deng et al., 2014; Yan et al., 2013), we show that random segregation at the first cleavage division, known as 'partitioning errors', provides an important source of initial blastomere-to-blastomere heterogeneity. During the two- to 16-cell stages, zygotic transcriptional activation generates different molecular feedbacks that minimize or enhance the initial blastomere-to-blastomere biases, resulting in different gene clusters with either a 'monostable' (ubiquitous expression between blastomeres) or 'bistable' (strongly asymmetric expression between blastomeres) pattern. For those genes with a bistable pattern, the relative ratio of opposing lineage specifiers in each blastomere creates an inclined 'lineage strength', providing a transcriptional clue to guide future cell

[1]State Key Laboratory of Stem Cell and Reproductive Biology, Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, 100101 Beijing, China. [2]University of Chinese Academy of Sciences, 100049 Beijing, China. [3]Department of Physiology and Cell Biology, University of Nevada School of Medicine, Reno, NV 89557, USA. [4]Biodynamic Optical Imaging Center and Center for Reproductive Medicine, College of Life Sciences, Third Hospital, Peking University, 100871 Beijing, China.
*These authors contributed equally to this work

‡Authors for correspondence (cqi@medicine.nevada.edu; yitao@ioz.ac.cn; qzhou@ioz.ac.cn; duane@ioz.ac.cn)

Received 4 March 2015; Accepted 17 August 2015

Development (2015) 142, 3468-3477 doi:10.1242/dev.123950

fates ahead of morphological distinction. We believe that these analyses represent a major step in understanding the very first mammalian embryonic symmetry-breaking process, as a result of both chance distribution and dynamic transcriptional regulation before morphological distinction.

RESULTS

Small biases initially rise in two-cell blastomeres following an approximately binomial distribution pattern

When considering the source of cell-to-cell heterogeneity, there are two major theoretical models. One is partitioning error, which occurs when, at cell division, cellular substances are unevenly distributed into the two daughter cells (Huh and Paulsson, 2011a). The other is 'stochastic transcription noise', which is derived from random transcriptional starts within each cell cycle (Cai et al., 2006; Elowitz et al., 2002; Raj et al., 2006). These two models are usually mixed-up in a real cell system, both contributing to cell-to-cell heterogeneity, and are therefore very difficult to analyze separately (Huh and Paulsson, 2011b). However, unlike normal mitotic cell division during which gene transcription is always active, mammalian zygotic transcription is mostly silent before the first embryo cleavage (Li et al., 2013; Tadros and Lipshitz, 2009). Therefore, the first mammalian embryo cleavage division provides a unique system in which partitioning error comprises the dominant source of biases between two-cell blastomeres, and its contribution to initiation of blastomere-to-blastomere biases can be calculated mathematically and tested by single-blastomere sequencing data.

By considering mammalian zygotes as a homogenous sphere and the first embryonic cleavage division as a binomial distribution system, we calculated the theoretical chance (Eqn 1 in Materials and Methods) of a biased distribution (with 10%, 15% and 20% differences between two daughter blastomeres) for a putative substance with regard to their initial copy numbers, and showed that the chance of an uneven distribution depends on their initial quantity (Fig. 1A): the substances present in small quantities are more likely to be differentially distributed into the two daughter cells, whereas those present in larger quantities have a greater chance of being equally divided (Fig. 1A). To test this binomial model in a real mammalian embryo cleavage division, we analyzed single-blastomere RNA-seq data from mouse and human two-cell blastomeres (Tang et al., 2011; Yan et al., 2013). As shown in a representative two-cell mouse embryo, the overall comparative transcriptome profile between two-cell blastomeres showed that highly expressed genes are usually similarly expressed, whereas differential expression is more often observed for genes with lower expression (Fig. 1B), which is similar to the mathematical prediction. To test more stringently the accuracy of this binomial model in a two-cell embryo, especially the extent of bias distribution with regard to the copy number of a gene transcript, we next estimated the transcript numbers in each two-cell mouse blastomere by combining qPCR and single-cell transcriptome reads (Materials and Methods). This estimate showed that every five transcripts approximately equaled one RPKM (reads per kilobase transcriptome per million reads) value in a two-cell blastomere, which was similar to what we previously reported (Tang et al., 2009). Then, by using the estimated transcript number for each gene, the blastomere-to-blastomere heterogeneity at the transcript level could be tested against the theoretical curve under binominal partition (Materials and Methods). The statistics of transcript partitioning at the two-cell embryo stage showed that the RNA copy numbers detected in two-cell blastomeres were approximately aligned with the binomial prediction (Fig. 1C,D; Fig. S1A,B),

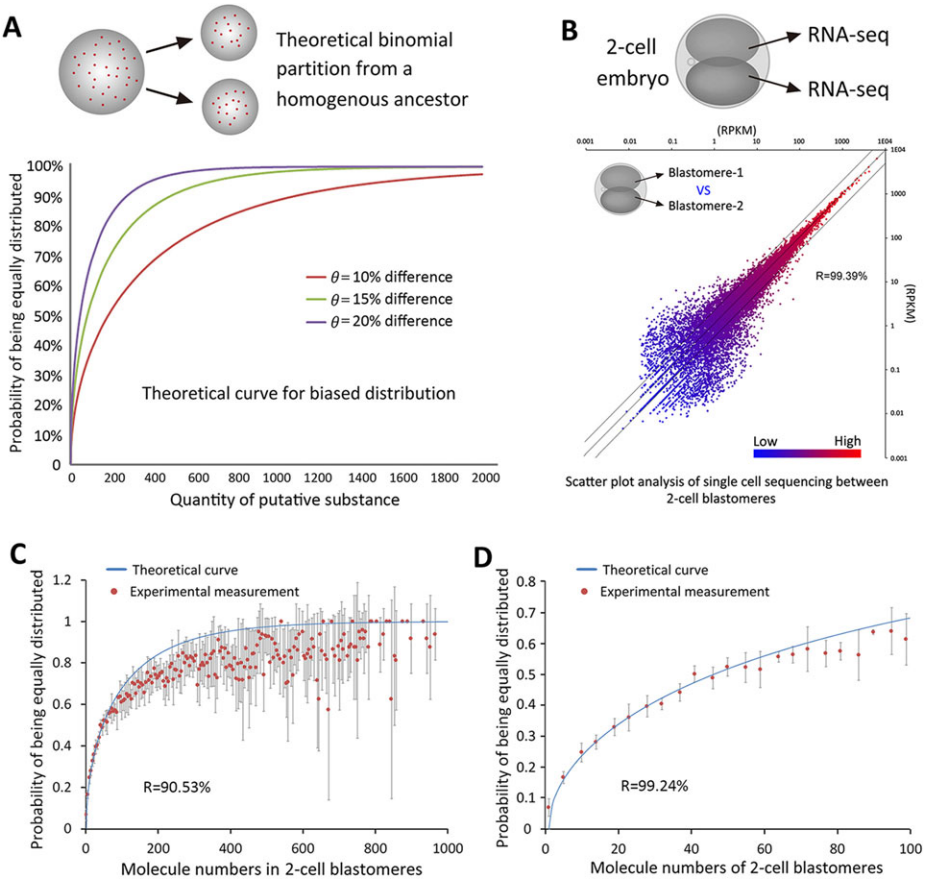


Fig. 1. Expression biases in two-cell blastomeres approximately follow a binomial distribution pattern. (A) Theoretical curve of binomial distribution from a homogenous ancestor, with thresholds of differential distribution of 10%, 15% and 20% of putative molecules. (B) Scatter plot comparison analysis of single-blastomere transcriptomes of two-cell mouse blastomeres. R, correlation coefficient. (C,D) Transcript counts in two-cell blastomeres were tested against the theoretical curve (for 20% unequal distribution) under binominal partition and a scale of 1000 counts (C) and 100 counts (D). Error bars indicate s.e.m., with average values from four independent embryos. Correlation coefficient R represents the fitness of data to the theoretical curve.

Development (2015) 142, 3468-3477 doi:10.1242/dev.123950

especially when the copy numbers were low (<100) (Fig. 1D; Fig. S1B). This analysis demonstrated that during the first embryo cleavage division, small biases initially arise following a binomial distribution pattern, with the low-quantity transcripts being more likely to be asymmetrically distributed.

Zygotic transcriptional activation and multiplied partitioning errors elevate overall blastomere-to-blastomere biases during embryo development

Once the initial bias is generated in two-cell blastomeres under binomial partitioning, where will it end? It has been previously shown that transcription noise is an important source of cell-to-cell heterogeneity (Raj et al., 2006). Therefore, it is our hypothesis that the increasing level of zygotic transcriptional activation after the middle-to-late two-cell-embryo stage should increase the level of blastomere-to-blastomere bias. To test this hypothesis, we utilized two methods to analyze quantitatively the stage-specific blastomere-to-blastomere heterogeneity using single-blastomere RNA-seq datasets of two-cell mouse embryos at the early, middle and late stages (Deng et al., 2014) (Fig. 2A).

First, we calculated the expression noise between RNA-seq technical repeats of one oocyte/blastomere and those from two-cell blastomeres based on a previously published formula (Swain et al., 2002) (Eqn 2 in Materials and Methods). This calculation clearly showed that the extent of blastomere-to-blastomere expression bias from the early two-cell embryo was higher than that of the single-oocyte/blastomere technical repeats (Fig. 2B). It can be assumed that this difference represents the contribution of partitioning error because early two-cell blastomeres have only just been formed by cleavage division and zygotic transcription is almost silent at this stage (Lee et al., 2014) so the contribution of transcriptional noise is minimal. The early-to-late embryo progression showed an ever increasing expression bias, which presumably represents the contribution of transcriptional noise (Fig. 2B) because during this period there are no cleavage divisions, therefore transcriptional noise becomes the only source of increased asymmetry. These analyses dissect the relative contribution of partitioning errors and transcriptional noise that contribute to transcriptional asymmetry between two-cell blastomeres in a stage-dependent manner, supporting the hypothesis illustrated in Fig. 2A.

Our second method for representing blastomere-to-blastomere expression asymmetry is as follows: for a gene detected in a two-cell embryo, one blastomere will show E_{higher} and the other will show E_{lower} (where E represents gene expression level), and the value of E_{higher}/E_{lower} will represent the extent of asymmetry for this gene between the two blastomeres. When the two examined blastomeres are more 'identical', the calculated E_{higher}/E_{lower} values for more genes should be close to 1, whereas if the two blastomeres are more 'distinct', more genes will show increased E_{higher}/E_{lower} values. In this context, the numbers of genes with $1<E_{higher}/E_{lower}<1.5$ and $E_{higher}/E_{lower}>2$ were calculated for each two-cell embryo (early, middle and late stages) and for the technical repeats (Fig. 2C). The results clearly showed that the number of genes with $1<E_{higher}/E_{lower}<1.5$ was highest in the technical repeats and continually decreased during the early-to-late embryo progression, whereas this trend was reversed for the number of genes with $E_{higher}/E_{lower}>2$ (Fig. 2C). These analyses confirmed the conclusion drawn from the data in Fig. 2B, both showing that the first embryo cleavage division generates blastomere-to-blastomere bias as a result of partitioning error, and that during the early-to-late two-cell embryo progression, the blastomere-to-blastomere bias continually increases owing to zygotic transcriptional activation.

To analyze further the expression asymmetry in four-cell and eight-cell embryos, we performed a pairwise comparison for all blastomeres within an embryo to reveal the expression asymmetry between each pair of blastomeres. As shown in Fig. 3A,B and Fig. S2, the extent of blastomere-to-blastomere asymmetry showed an overall increase from the two-cell stage to the eight-cell stage in both mouse and human embryos, which could be due to continuous transcriptional activities and multiplied partitioning error (Eqns 3-5 in Materials and Methods; Fig. 3A) during embryo development.

Different trends of transcriptional asymmetry during the two- to 16-cell embryo stages represents either a monostable or bistable pattern

In addition to the observed overall increase in gene expression asymmetry along embryonic development (two-cell to eight-cell stage) (Fig. 3B; Fig. S2), it is notable that during continuous embryonic cleavage, the available amount of embryonic sample for each RNA-seq decreases and this will cause increased technical noise, especially for genes expressed at a low level (Streets et al., 2014), which might partially account for the observed asymmetry increase. Because highly expressed genes are less affected by both technical noise and partition errors, we further selected a subgroup

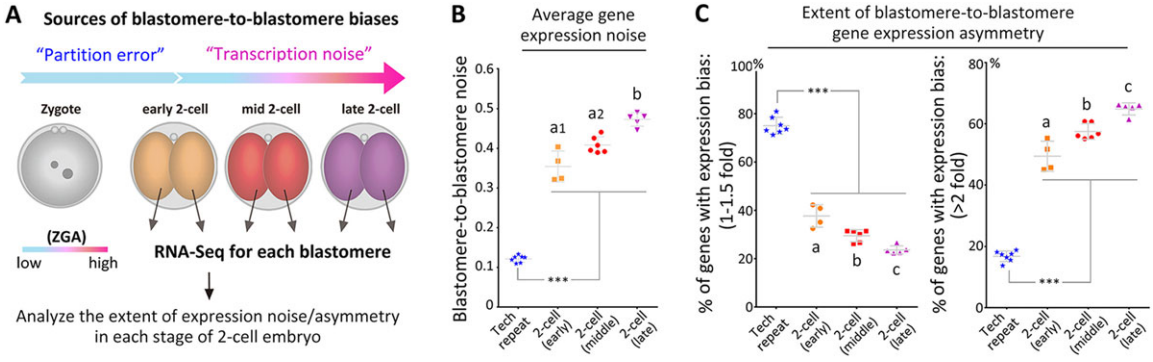


Fig. 2. Zygotic transcriptional activation elevates initial blastomere-to-blastomere biases. (A) Illustration of the strategy to analyze separately the two sources of blastomere-to-blastomere heterogeneity, namely partitioning errors and transcriptional noise, in an early-to-late two-cell embryo in a stage-specific manner. (B) Blastomere-to-blastomere expression noise calculated with a previously published formula using a single-blastomere transcriptome from an early-to-late two-cell embryo as well as using technical repeats of RNA-seq data from early blastomeres (Eqn 2 in Materials and Methods). (C) The extent of blastomere-to-blastomere gene expression asymmetry is represented by E_{higher}/E_{lower}. If the gene expression in the two blastomeres is more similar, the value of E_{higher}/E_{lower} will be closer to 1. ***P<0.001; different letters indicate statistical difference, P<0.01; same letters indicate P>0.05; one-way ANOVA.

Development (2015) 142, 3468-3477 doi:10.1242/dev.123950

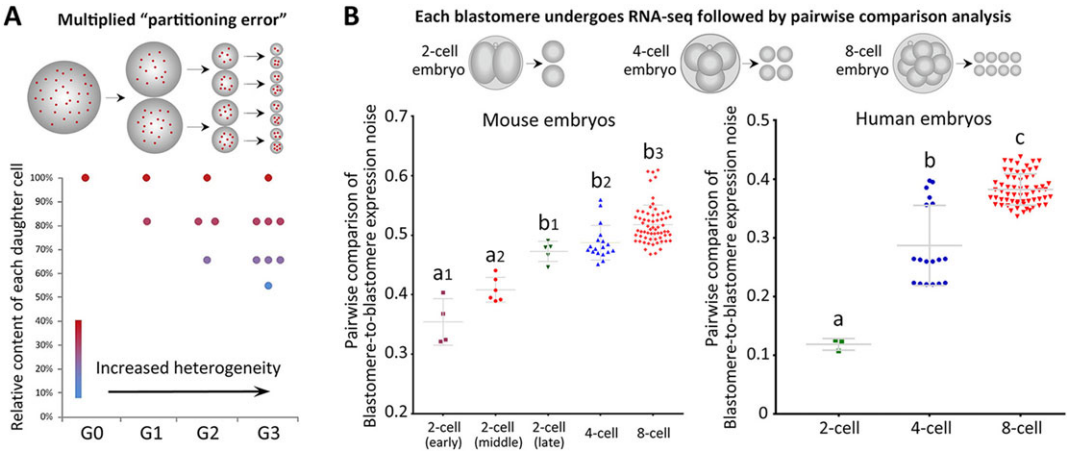


Fig. 3. Multiplied partitioning errors and elevated overall blastomere-to-blastomere biases during the two- to 16-cell stages. (A) Illustration showing that multiplied partitioning errors at each cleavage division drive the continuously increasing cell-to-cell bias, each time creating one daughter cell bearing a higher quantity of a specific substance and another with a lower quantity of that substance (Eqns 3-5 in Materials and Methods). The graph shows the amplification effect of a small bias (10% co-efficiency) during the first rounds of cleavage division; G0-3 indicates generations 0-3. (B) Single-blastomere RNA-seq and blastomere-to-blastomere expression noise in two- to eight-cell embryos in both mouse and human. Each dot represents a comparison of two blastomeres within one embryo. Datasets from mouse (Deng et al., 2014) and human (Yan et al., 2013) were used for the calculations. Different letters indicate statistical difference, $P<0.01$; same letters indicate $P>0.05$; one-way ANOVA.

of genes from each mouse embryonic stage with both a strong expression bias (top 20%) and high expression level (top 20%), which we believe will more faithfully represent the level of transcription and minimize the potential influence of technical bias (which affects to a greater degree genes with lower expression). For these selected genes (highly expression level with strong bias) (Table S1), we analyzed the trends of blastomere-to-blastomere asymmetry during the two- to 16-cell stages, and found that the extent of asymmetry at a certain stage proceeds with two distinct patterns: for some genes, the extent of asymmetry tends to be minimized between blastomeres, whereas other genes become increasingly asymmetrically distributed (Fig. 4A-C). Notably, at the eight-cell stage, only very few genes with high asymmetry showed decreased asymmetry onward (Fig. 4C), suggesting that, compared with two- to four-cell embryos, the transcriptional bias between eight-cell blastomeres has become increasingly irreversible and deterministic. Interestingly, for those genes with strong bias distribution at the 16-cell stage, gene ontology analysis revealed that most of these genes are cataloged in, for example, 'positive regulation of cell differentiation', 'positive regulation of developmental process' (Fig. 4D), implicating their potential in directing different cell fates. In addition to the overall analysis, the actual asymmetry trends of three typical genes were examined during the two-cell to 16-cell stages: a housekeeping gene (*Tubb2c*, which encodes the tubulin beta-2C chain; also known as *Tubb4b*), a transcription factor for self-renewal (*Pou5f1*, also known as *Oct4*), and a lineage specifier with a defined role starting at the four-cell stage (*Carm1*, the histone arginine methylase). Each of these genes exhibits a different pattern of gene expression asymmetry, which may be described as overall-flat (*Tubb2c*), wave-like (*Pou5f1*) or ever-increasing (*Carm1*) (Fig. 4E), suggesting dynamic transcriptional regulation for each gene.

The observed distinct asymmetry dynamics for different genes are, in general, supposed to be regulated by series of molecular chain reactions (MacArthur et al., 2012) with either negative or positive transcriptional circuits. This could be simplified as two distinct models involving a negative feedback-controlled 'monostable pattern' (Fig. 4F) or a positive feedback-controlled 'bistable pattern' (Fig. 4G) (Materials and Methods; Fig. S3), with the latter (such as the pattern of *Carm1*) being a potential driving force for transcriptional symmetry-breaking in embryo development.

Clues to cell fate bifurcation may be found in the ratio of opposing lineage specifiers, which is influenced by cleavage history and *de novo* transcription

In a real biological system, the divergence of lineage is usually decided by more than one factor, and recent evidence from stem cell research has revealed that lineage specifiers with counteracting effects work in balance to maintain pluripotency, whereas a tilted ratio leads to lineage differentiation (Montserrat et al., 2013; Shu et al., 2013). Such a scenario similarly exists in the context of early mouse embryos, in which biased lineage-driving forces are crucial in guiding future cell fates (Bruce and Zernicka-Goetz, 2010). Here, we proposed that the relative ratio of a pair of opposing lineage specifiers is influenced by both partitioning error at cleavage division, and transcriptional regulation within the cell cycle. By supposing that two lineage specifiers have an initial ratio that maintains the cell in an undifferentiating state (Fig. 5A), we first calculated the probability of tilting the initial ratio (to an extent of 10%) of opposing lineage specifiers based on partitioning error (Eqn 6 in Materials and Methods), and showed that the odds of breaking the initial ratio are inversely correlated with the level of either one or both substances before cleavage (Fig. 5B), suggesting that the copy number of lineage specifiers deposited in the previous cycle can influence their distribution pattern (relative ratio with other lineage specifiers) in the daughter cells, thus contributing to the bifurcation of future cell fate.

Transcriptional regulation is another important contributor that can tilt the ratio of lineage specifiers. For example, Cdx2 and Carm1 are a known pair of lineage specifiers, which guide different cell fates in mouse early embryo. The expression of Cdx2 at the two- to four-cell stages is very low (mostly inherited from maternal storage), but its expression increases drastically in the eight-cell embryo (Fig. 5C); thus, *de novo* transcription at this stage becomes the dominant factor affecting the ratio of Cdx2 to Carm1. Whereas for Carm1, transcription of which increases only mildly at the eight-cell stage (Fig. 5C), the Carm1 to Cdx2 ratio is affected by both

Development (2015) 142, 3468-3477 doi:10.1242/dev.123950

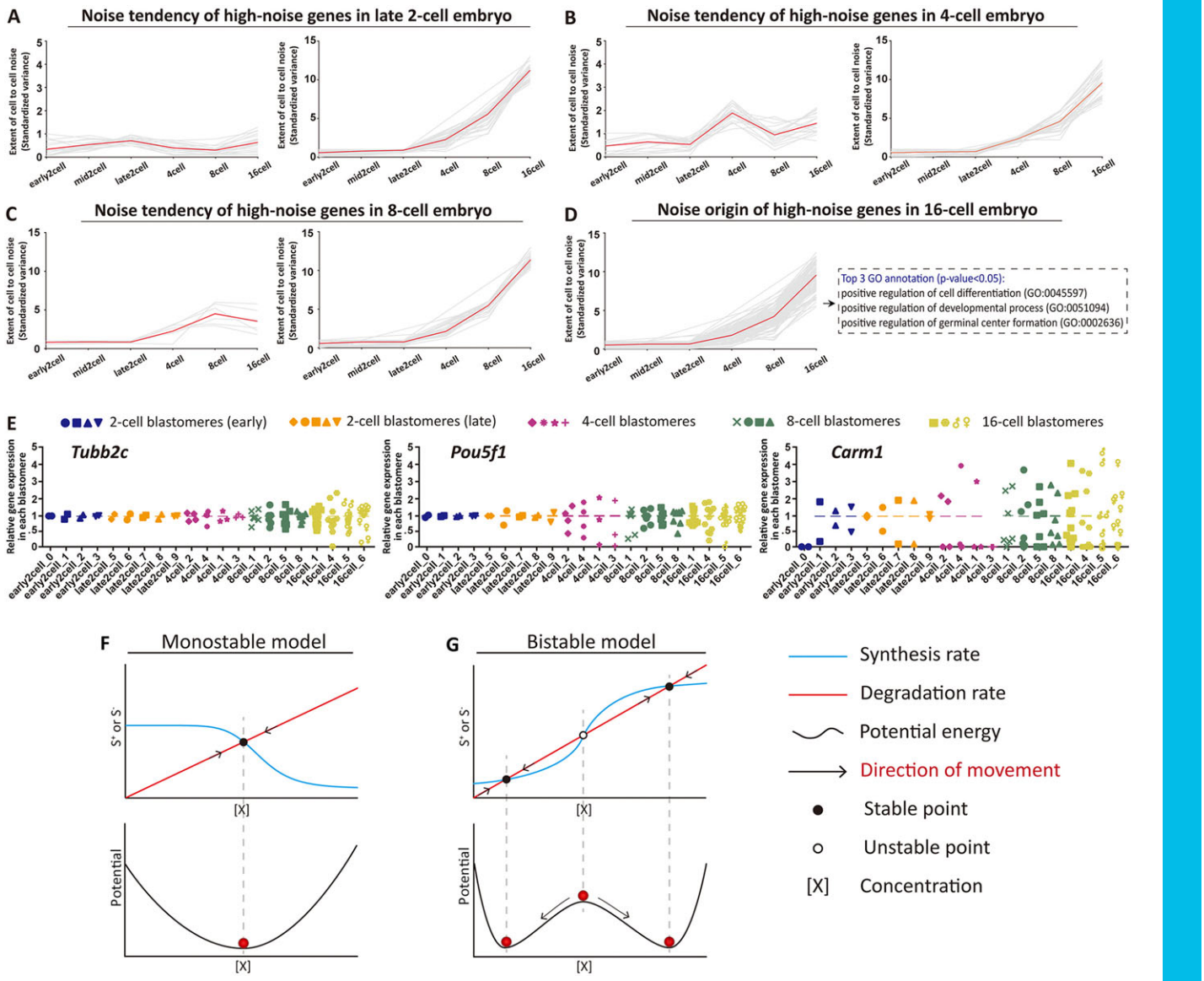
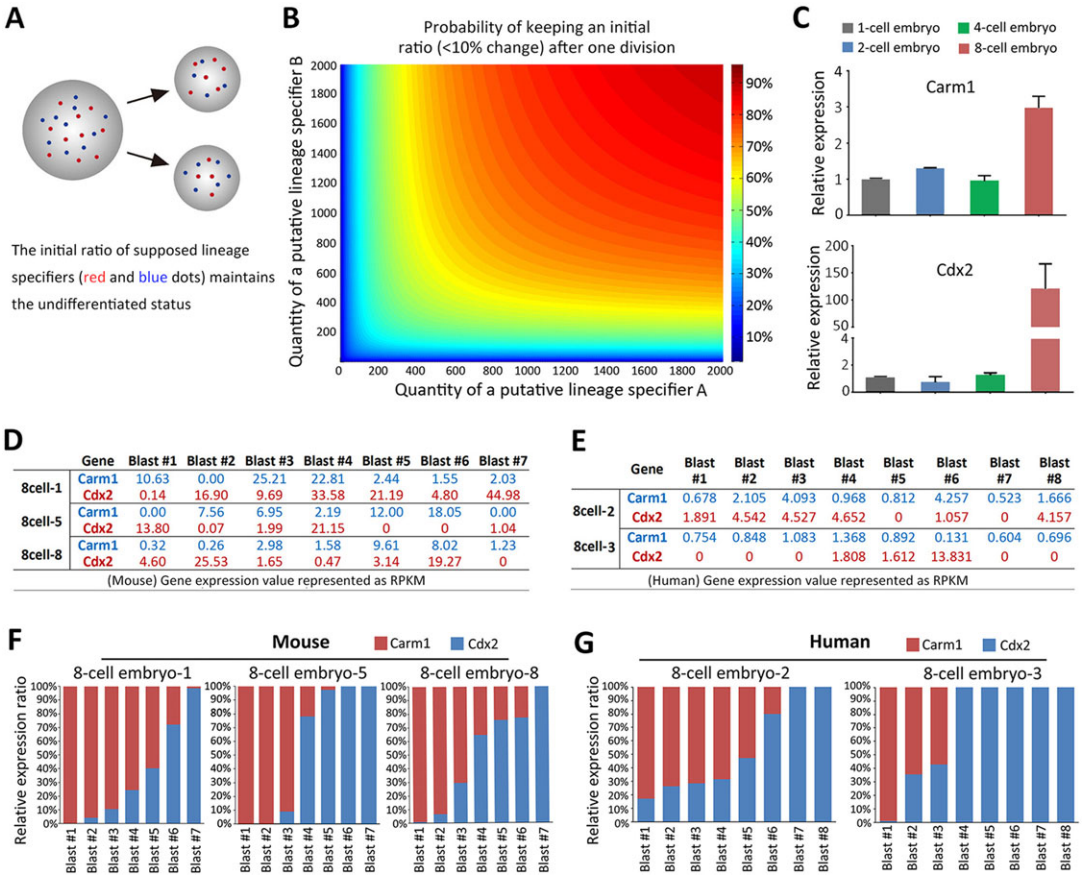


Fig. 4. Trends of transcriptional asymmetry during the two- to 16-cell embryo stages reveals monostable and bistable patterns. (A-D) The trends of blastomere-to-blastomere asymmetry for genes with high expression level (top 20%) and strong bias (top 20%) at the two-cell (A), four-cell (B), eight-cell (C) and 16-cell (D) stages. (E) The expression asymmetries of *Tubb2c*, *Carm1* and *Pou5f1* were analyzed using single-blastomere expression profiles from two-cell to 16-cell embryo stages in mouse. The average expression level of each gene in each embryo is normalized to 1, and the relative expression level of each gene in each blastomere is shown using different shapes/colors, as illustrated. The dispersion degree along the *y*-axis can be directly visualized in each embryo and represents the overall extent of expression asymmetry between different blastomeres. Each dot represents the relative value of each blastomere to the average RPKM of all blastomeres within one embryo. (F,G) Dynamics of single gene expression systems with negative/positive feedback regulation. The change of protein X over time (*dX*/*dt*) can be described by a differential equation which is synthesis rate S^+ minus degradation rate S^-. The top two panels are the plot of the functions S^+ and S^- within the same graph. The lower two panels illustrate the corresponding potential function of the system. From these graphs, there are one (F) or three (G) intersection points, indicating the steady states in which *dX*/*dt* is equal to 0. The directions of the arrows indicate the movement towards equilibrium. If two arrows move away from each other, this represents an unstable steady state. Therefore, a single gene expression system with negative feedback regulation (F) has only one stable steady states and a single gene expression system with positive feedback regulation (G) has two stable steady states: one in which the level of X is low and one in which the level of X is high; the system could shift into either of these two states from the unstable point.

transcripts inherited from previous divisions and those newly transcribed in the current cell cycle.

As Carm1 and Cdx2 have been reported to direct distinct cell fates (Jedrusik et al., 2008; Torres-Padilla et al., 2007), it is our hypothesis that their relative ratio within an eight-cell blastomere provides clues for the future lineage specification. We therefore analyzed the relative expression levels of both *Carm1* and *Cdx2* in each blastomere from eight-cell mouse and human embryos, and showed that different eight-cell blastomeres indeed have a 'wax and wane' phenomenon regarding the relative ratio of *Carm1* and *Cdx2*. As shown in Fig. 5D-G, most blastomeres are either dominated by *Carm1* or *Cdx2*, suggesting that a predisposed lineage strength has been formed in these blastomeres. It is notable that some blastomeres showed a similar expression level of both *Carm1* and *Cdx2*, thus representing an undecided, bi-potent state as suggested previously (Bruce and Zernicka-Goetz, 2010). According to our analysis, when these bi-potent blastomeres proceed into next cell cycle, the relative ratio of *Carm1* and *Cdx2* could be further tilted as a result of both

Development (2015) 142, 3468-3477 doi:10.1242/dev.123950



	Gene	Blast #1	Blast #2	Blast #3	Blast #4	Blast #5	Blast #6	Blast #7
8cell-1	Carm1	10.63	0.00	25.21	22.81	2.44	1.55	2.03
	Cdx2	0.14	16.90	9.69	33.58	21.19	4.80	44.98
8cell-5	Carm1	0.00	7.56	6.95	2.19	12.00	18.05	0.00
	Cdx2	13.80	0.07	1.99	21.15	0	0	1.04
8cell-8	Carm1	0.32	0.26	2.98	1.58	9.61	8.02	1.23
	Cdx2	4.60	25.53	1.65	0.47	3.14	19.27	0

	Gene	Blast #1	Blast #2	Blast #3	Blast #4	Blast #5	Blast #6	Blast #7	Blast #8
8cell-2	Carm1	0.678	2.105	4.093	0.968	0.812	4.257	0.523	1.666
	Cdx2	1.891	4.542	4.527	4.652	0	1.057	0	4.157
8cell-3	Carm1	0.754	0.848	1.083	1.368	0.892	0.131	0.604	0.696
	Cdx2	0	0	0	1.808	1.612	13.831	0	0

Fig. 5. Cleavage history and *de novo* transcription contribute to the relative ratio of opposing lineage specifiers. (A) Illustration of the tilted ratio of two substances as a result of partitioning error. (B) Colors indicate the probability of tilting an initial ratio (10% threshold) of a pair of lineage specifiers with counteracting functions (Eqn 6 in Materials and Methods). (C) The stage-specific expression level (RPKM value) of *Carm1* and *Cdx2* during the one- to eight-cell embryo stages. *n*=3 for each column. Data used to generate these columns are derived from a previous publication (Xie et al., 2010). (D,E) The RPKM values of *Carm1* and *Cdx2* in each eight-cell blastomere examined from mouse (D) and human (E) embryos. (F,G) The relative ratios of *Carm1* and *Cdx2* in different eight-cell blastomeres form distinct lineage strengths in both mouse (F) and human (G) embryos. The graphs in F and G were generated from the raw numbers in D and E, respectively. Previously published mouse (Deng et al., 2014) and human (Yan et al., 2013) datasets were used to generate the table and figure.

partitioning error and transcriptional regulation, eventually creating a situation in which one side wins out on the other.

Moreover, we have further provided additional lists of genes showing the 'wax and wane' phenomenon with regard to Carm1 and Cdx2 in eight-cell blastomeres (Table S2). These genes could be potential lineage specifiers and their function warrants further investigations.

Finally, based on our single-blastomere transcriptome data analysis, a scenario of early transcriptional symmetry-breaking and lineage specification is summarized in Fig. 6. In this model, a static view before the final lineage destination will give an impression of an undecided lineage combat (Fig. 6), showing a heterogeneous level of opposing transcription factors (i.e. a 'stochastic appearance') that could be either inherited from a previous division or produced by *de novo* transcription, i.e. also having a 'traceable origin'.

DISCUSSION

The mechanisms of early mammalian cell fate determination remain debatable as it has not yet been established whether the first bifurcation of cell fate in mammalian embryo emerges randomly at the morula stage (Dietrich and Hiiragi, 2007; Wennekamp and Hiiragi, 2012) or if the molecular clues of differentiation emerge before morphological distinction (Plachta et al., 2011; Tabansky et al., 2013; Torres-Padilla et al., 2007; Zernicka-Goetz, 2013). In the present study, our single-blastomere transcriptome analysis of early embryo development have showed that the earliest blastomere-to-blastomere biases emerge with the first cleavage division, owing to random segregation, but subsequent zygotic transcriptional activation triggers transcriptional regulation that fine-tunes these small biases in a more defined manner, minimizing or amplifying the initial biases with negative or positive feedback mechanisms. We believe that only those transcriptional biases that finally develop into a bistable pattern (strongly asymmetric between blastomeres), bear the potential to guide lineage fates. Moreover, our analysis supports a scenario in which opposing lineage specifiers within an early blastomere constantly compete with each other based on their relative ratio, pushing the blastomere onto a predisposed, yet flexible, lineage track before morphological distinction. These analyses revealed mammalian early embryo symmetry-breaking as a continuous process rather than a sudden emergence, with the driving force involving both chance separation and transcriptional circuits.

As shown in our analysis, at the first cleavage division, small biases will inevitably arise in a binomial distribution pattern, with lower-quantity substances bearing a greater chance of being asymmetrically distributed (Fig. 1A-D). Such quantity-dependent asymmetric distribution in two-cell blastomeres, in our opinion, is an important step in resolving the dilemma of how two-cell-stage

Development (2015) 142, 3468-3477 doi:10.1242/dev.123950

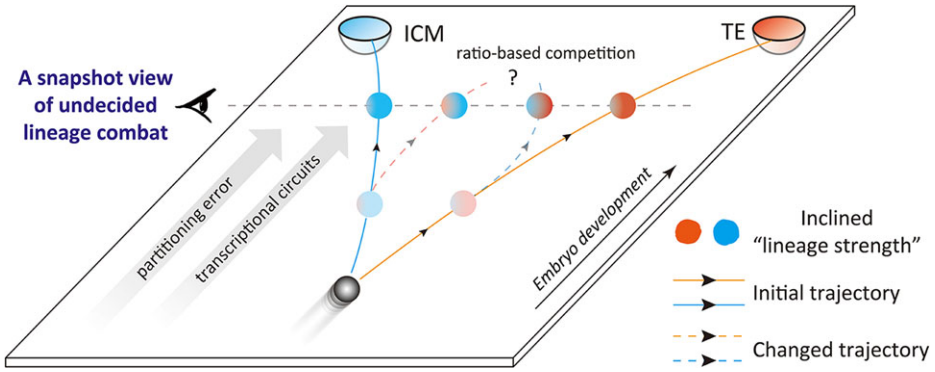


Fig. 6. A symmetry-breaking and lineage-competition model involving both random segregation and transcriptional regulation. The partitioning error at each cleavage division and the transcriptional circuits along embryo development are two major sources for initial generation and subsequent fine-tuning of blastomere-to-blastomere heterogeneity. These two factors also change the initial ratio of opposing lineage specifiers in mother or daughter blastomeres, thus changing the initial trajectory of lineage fate. The changes in color density represent the gradually reinforced lineage strength of either side. Note that an undecided impression could be observed at a static time point, representing a snapshot of the continuous lineage competition. Clues to cell fate bifurcation have been developed by the ratios of opposing lineage specifiers, with the wax and wane of either side (illustrated by different colors in each sphere) finally leading to the dichotomy of cell fate.

blastomeres can be both 'identical' (in totipotency) and 'different' (in guiding future lineages) (Zernicka-Goetz, 2006). In a real biological system such as a zygote, the substance found in greater quantity is usually required for maintaining basic cell properties, whereas those present in small amounts are usually elements with fine-tuning regulatory functions (transcriptional factor, non-coding RNAs, etc.). Therefore, these regulatory factors tend to be unequally distributed in two-cell blastomeres as a result of partitioning errors, which may trigger downstream events that influence future cell fate. Moreover, apart from the biased transcript distribution, it could be repeatedly observed that, among the two-cell blastomeres (which are generated at the same time), one cell always divides before the other (Fig. S4), supporting the idea that at the late two-cell stage, the two-cell blastomeres already bear slight biological differences regarding their intrinsic properties.

Recently, it is notable that two groups also used single-blastomere transcriptome analysis to study the early symmetry-breaking process. One report showed that as early as the two-cell embryo stage, some genes have already become strongly asymmetrically expressed in different blastomeres and it has been suggested that these genes may be involved in directing future lineages (Biase et al., 2014). The other report further showed that the blastomeres have become strikingly variable at eight-cell stage owing to the amplification effect of transcriptional activities (Piras et al., 2014). These conclusions are in general in accordance to our analysis, as we showed an overall increase in blastomere-to-blastomere biases driven by both zygotic transcriptional activation and multiplied partitioning errors during ongoing embryonic cleavage (Fig. 2B,C and Fig. 3A,B). However, when we performed more stringent analysis by filtering out genes with low expression, which in large ruled out the influence of technical biases, we found a more dynamic framework of transcriptional regulation during the two- to 16-cell stages, showing two distinct patterns of progression: for some genes, the extent of blastomere-to-blastomere asymmetry tends to be minimized, displaying a monostable pattern [in general, the single stable equilibrium in gene expression dynamics is a result of negative feedback regulation (Alon, 2007)], whereas for other genes, the extent of asymmetry becomes increasingly larger, displaying a bistable pattern, as a result of positive feedback regulation as previously pointed out (Smits et al., 2006).

These analyses revealed that although strong biases of transcripts could be observed in early blastomeres as early as the two- to four-cell stages, it is not a guarantee that these early observed biases will be kept at the later embryo stages; alternatively, only those crucial molecules (mRNA, protein or non-coding RNA) with the ability to trigger self-reinforcement or gene regulatory feedbacks (Davidson, 2010) bear the ability to drive initial small biases to produce tangible biological differences.

Despite the observed dynamic transcriptional symmetry-breaking during early embryo development, we would like to emphasize that even if a lineage specifier is asymmetrically distributed between early blastomeres, they could only 'guide', but not 'decide' the lineage track until other more definitive clues such as inner-outer position, cell polarity, cell-cell contacts (Lorthongpanich et al., 2012) emerge. However, once a blastomere has acquired dominant lineage specifiers (for example, Carm1) at the four-cell stage, the chance of keeping the dominant position in its daughter generation will be higher than those blastomeres with low level expression. This is, in principle, consistent with the biological observations that these Carm1-dominant blastomeres showed a biased cell fate as previously published (Tabansky et al., 2013; Torres-Padilla et al., 2007). Also, previous experiments have shown that if the transcript of a lineage specifier is exogenously injected in a subset of blastomeres, it will guide the lineage direction in a more definite manner (Cockburn et al., 2013; Jedrusik et al., 2008; Torres-Padilla et al., 2007), most probably because the injection of exogenous transcripts secured its dominant expression level, which overwhelms the fluctuations caused by other factors (such as partitioning error and transcription regulation) in its daughter cells.

Regarding the lineage competition by opposing lineage specifiers in the early embryo, it remains unclear at what time the competition first begins and by what mechanisms the winner steadily establishes its cell lineage. In the present study, we revealed an interesting phenomenon that the relative ratio of Carm1 and Cdx2 (which guide ICM or TE specifications, respectively) in each eight-cell blastomere is different, suggesting that inclined lineage strengths might have been formed in these eight-cell blastomeres. However, the lineage track in these blastomeres have not been fully decided, because in the next round of division, the ratio of Carm1 to Cdx2 could be reversed as a result of both random segregation and transcriptional regulation. Nonetheless, such a ratio-dependent

Development (2015) 142, 3468-3477 doi:10.1242/dev.123950

lineage battle will continue, until one side finally becomes strong enough to consolidate its dominant position and define the lineage fate (Fig. 6) by introducing other more definitive clues, such as inner-outer position, cell polarity, etc. Similar lineage competition and wax and wane of opposing lineage specifiers (Fig. 6) may also exist in the second wave of cell specification of the ICM into the epiblast and primitive endoderm (Ohnishi et al., 2014), which would involve both a traceable cell origin from their lineage ancestors (Morris et al., 2010) and the plasticity to change cell fates before the final differentiation (Yamanaka et al., 2010).

In summary, our single-blastomere transcriptome analysis revealed a dynamic transcriptional symmetry-breaking process occurring far earlier than morphological distinction, contributed to by both random segregations at each embryo cleavage division and transcriptional regulatory feedbacks. In the future, more detailed verification of our analysis with regard to lineage specification will depend on our improved ability to examine precisely the quantity of molecules (proteins, RNAs) in an amplification-free manner, such as by single-molecule fluorescent *in situ* hybridization detection of selected genes at single-cell resolution (Itzkovitz et al., 2012) with time-lapse observation, as well as functional examination of candidate genes regarding their roles in lineage specification.

MATERIALS AND METHODS

Single-blastomere RNA-seq data of human and mouse pre-implantation embryos

We analyzed single-blastomere RNA-seq datasets from pre-implantation human embryos (Yan et al., 2013; data available in Gene Expression Omnibus under accession number GSE36552). Single-blastomere RNA-seq data from mouse pre-implantation embryos have been previously published (Tang et al., 2011; Deng et al., 2014; data available in Gene Expression Omnibus under accession numbers GSE22182 and GSE45719, respectively). Gene expression was calculated as RPKM (reads per kilobase transcriptome per million reads; Audic and Claverie, 1997; Mortazavi et al., 2008).

Mouse embryo collection and *in vitro* culture under time-lapse recording

Eight- to ten-week-old C57BL/6 female mice were stimulated to superovulate by standard methods (Wan et al., 2013) and then mated with DBA2 male mice. The animal-use protocols in the present study were approved by the Animal Research Committee of the Institute of Zoology, Chinese Academy of Sciences. Two-cell embryos were collected and cultured *in vitro* in M2 medium under mineral oil, and their progression to four-cell stage was video recorded by time-lapse microscopy (UltraVIEW VoX 3D Live-Cell Imaging System). The time lag between the three-cell and four-cell stages of each embryo was recorded and analyzed.

Clustering of gene noise pattern

Genes with both high noise (top 20%) and relative high expression level (top 20%) at each embryonic stage were selected, and the pattern of these genes' noise during the progression from two-cell to 16-cell stages was clustered by Genesis (Sturn et al., 2002), using a hierarchical clustering method, with 'Complete linkage clustering' agglomeration parameter in the software.

Gene Ontology (GO) analysis

GO enrichment was analyzed using DAVID (http://david.abcc.ncifcrf.gov/). A hypergeometric test was performed using the default parameters to adjust the P value.

Model construction and data analysis

Calculating the real counts of mRNAs in one two-cell blastomere

The real count of *Hprt* in one MII mouse oocyte is ~4829.2, according to a previous report (Steuerwald et al., 2000), and the relative gene expression of *Hprt* between an MII oocyte and a two-cell-stage blastomere was obtained by real-time PCR. The timing of oocyte and two-cell embryo retrieval was the same as we previously described (Tang et al., 2011). The calculation showed that 1 RPKM that corresponded to 4.465 counts for one two-cell blastomere.

Testing the theoretical pattern using experimental data

To test the probability of asymmetrical distribution upon an equal cleavage division from a homogenous ancestor, we built an integral equation to approximate the probability (P) of a specific substance with a quantity N to show the same distribution in the daughter cells as in the mother cell. The coefficient θ represents the extent of asymmetry between the two daughter cells. Eqn (1) is shown below, based on previous publications (Berg, 1978; Rigney, 1979).

$$P\left(\left|\frac{N_A - N_B}{N}\right| \le \theta\right) = \int_{((1-\theta)/2)N}^{((1+\theta)/2)N} \frac{1}{\sqrt{(1/2)\pi N}} e^{-(N_A-(1/2)N)^2/(1/2)N} dN_A. \quad (1)$$

In the present equation, we set θ and all N values as non-negative real parameters. N_A and N_B represent the quantities in daughter cell A and daughter cell B, respectively. Different values of θ (10%, 15% and 20%) were used to simulate the probability curves corresponding to different quantities of N. (Fig. 1A). Then, we counted the frequency of gene numbers between two-cell blastomeres under the given threshold θ, with one count resolution of the real counts of the mRNAs. For Fig. 1C,D and Fig. S1, each dot represents the average value of four independent embryos.

Measuring noise between technical repeats and different blastomeres

To measure the noise between technical repeats or two blastomeres, we used Eqn (2) adapted from a previous publication (Swain et al., 2002):

$$\eta^2_{avg} = \frac{1}{n}\sum_{i=1}^{n} \frac{1/2(N_{iA}^2 + N_{iB}^2) - [1/2(N_{iA} + N_{iB})]^2}{[1/2(N_{iA} + N_{iB})]^2}. \quad (2)$$

In this equation, η_{avg} represents the average noise of each expressed gene in two blastomeres, and n is the number of expressed genes. N_A and N_B represent the quantities in cell A and cell B, respectively. i is a positive integer number. In Fig. 2, The RNA-seq datasets used for analyzing technical replicates were five mouse oocytes and two four-cell blastomeres (previously published; Tang et al., 2009); datasets for analyzing single two- to eight-cell embryo blastomeres were previously published (Deng et al., 2014). In Fig. 3B and Fig. S2, each dot represents one pair of blastomeres in a two- to eight-cell embryo. For example, if four blastomeres were sequenced in a four-cell embryo, this could generate six pairs of comparisons [C(4,2)=6], thus six dots on the chart.

Multiplied partition error during ongoing cleavage division

Because each cleavage division will generate small biases in a quantity-dependent manner, as demonstrated in Eqn (1) and Fig. 1A, under constant cleavage after n times, it will spontaneously generate one daughter cell with the highest content of a specific substance and another daughter cell with the lowest content of that substance N. The $N_{highest}$ and N_{lowest} in a daughter cell after n cleavages can be calculated using Eqns (3) and (4), as shown below.

$$N_{highest} = N \cdot \left(\frac{1}{2}\right)^n \prod_{i=1}^{n} (1 + \theta_i) \quad (3)$$

$$N_{lowest} = N \cdot \left(\frac{1}{2}\right)^n \prod_{i=1}^{n} (1 - \theta_i). \quad (4)$$

When supposing that each cleavage has the same, unequal coefficient, the relative levels of a specific substance in each blastomere can be calculated with Eqn (5):

$$content\% = \frac{(1+\theta)^{\alpha}(1-\theta)^{n-\alpha}}{(1+\theta)^n}. \quad (5)$$

The α and n are non-negative integers, and $\alpha \in [0, n]$. When calculating the substance content for each blastomere after n cleavage divisions, the number

Development (2015) 142, 3468-3477 doi:10.1242/dev.123950

of blastomeres in an embryo after n cleavage divisions is 2^n. Under these conditions, α should be used $C_n^i(i \in [0,n])$ times to obtain the content value for each blastomere. When θ=10%, the relative levels of each blastomere from two-, four- and eight-cell embryos are determined as demonstrated in Fig. 3A.

A double integral model to simulate ratio change between two molecules

To test the probability of breaking an initial ratio of a pair of counteracting lineage specifiers, we built a double integral model in which N and M represent the two putative lineage specifiers. The probability (P) of N and M maintaining the initial ratio [with an asymmetry permission coefficient of θ in the daughter cells (A and B)] after an embryo cleavage division can be calculated using Eqn (6):

$$P\left(\left|\frac{N_A}{M_A}-\frac{N}{M}\right|\le\theta\right)=\int_1^N\frac{1}{\sqrt{(1/2)\pi N}}e^{-(N_A-(1/2)N)^2/(1/2)N}dN_A \int_{(1-(\theta/2))(M/N)N_A}^{(1+(\theta/2))(M/N)N_A}\frac{1}{\sqrt{(1/2)\pi M}}e^{-(M_A-(1/2)M)^2/(1/2)M}dM_A. \quad (6)$$

The method of calculation is the same for daughter cell B. When θ=10%, the probability density is as shown in Fig. 5B.

Monostable and bistable models

The effect of negative feedback regulation is keeping the expression level around one stable state (monostable), whereas an auto-activating positive-feedback loop is able to exhibit bistable states, which may have three fixed points, including two local stable states (one high and one low) and one unstable state.

We simulated the monostable and bistable process based on the following assumptions:

1. The whole process is divided into two parts: determinate dynamics behavior during the cell cycle and mock process during the cell division period (Fig. S3). Fluctuations from gene expression are neglected, i.e. the process of gene expression is considered to be determinate as is its ordinary differential equation (ODE) dynamics.
2. Duration (or period) of cell cycle T is long enough that the system state can reach the corresponding local stable state in each cell cycle. In mathematical terms, that is $\gamma, \Gamma \ll T$.
3. Production of gene expression, i.e. protein and mRNA, are both considered.
4. The total volumes of the embryo are unchanged in the first six cell divisions (from one cell to 32 cells). Only this cell period is considered.

Determinate gene expression process during the cell cycle

An absolutely open loop of gene expression (without any feedback regulation) is impossible in a real biological system (Smits et al., 2006; Sprinzak and Elowitz, 2005). According to the regulation effect, genetic regulation can be divided into two types: positive and negative regulations. The regulation can be auto-feedback regulation (directed regulation), but also can be indirect regulation by production of other genes. For simplicity, we only consider single auto-feedback regulation of the gene expression network, i.e.:

$$\begin{cases}\dfrac{dm}{dt}=F(p)-\gamma m\\ \dfrac{dp}{dt}=Km-\Gamma p\end{cases}, \quad (7)$$

where m and p are mRNA concentration and protein concentration, respectively, and γ and Γ are their respective degradation rates; $F(p)$ is the mRNA transcription rate, which is defined as a function of p; and K is the translation rate per mRNA. For the mRNA transcription rate $F(p)$, the sign of $dF(p)/dp$ determines positive or negative feedback regulation; here, we take it as a Hill-type function $F(p)=(k_{max}/(k_H^n+p^n))+k_0$ or $F(p)=(k_{max}p^n/(k_H^n+p^n))+k_0$ for the case of negative or positive regulation, respectively. Here, k_{max} is the maximum transcription rate, n the Hill coefficient, k_H the Hill constant, and k_0 the basal transcription rate with $k_0 \ll k_{max}$.

The parameter values we selected in simulation were: K=30 h^{-1}, k_0=0.1 h^{-1}, Υ=30 h^{-1}, Γ=1 h^{-1}, k_H=10 (Vilar et al., 2002; Zheng et al., 2011).

Stochastic partitioning at cell division

We only consider the case in which segregation is independent, which means that each molecule has an independent probability (50%) of being in either daughter cell. For independent segregation, the partitioning error is already known to be binomial and could be postulated directly. In stochastic simulation, we specify a simple dynamic process (Markov process) in which fluctuations in the stationary state generate binomial partitioning error. The detailed Markov processes used here are merely mock processes: partitioning errors for independent partitioning could be calculated from the irreversible process

$$\begin{aligned}(y,L)&\xrightarrow{\lambda y}(y-1,\ L+1)\\ (y,L)&\xrightarrow{\lambda y}(y-1,\ L),\end{aligned} \quad (8)$$

where the process starts with y=x and the partitioning error is calculated at the end of the process, y=0. The rates have y multiplied to a constant, which sets a natural boundary such that when y=0, the process spontaneously terminates, rather than imposing boundary conditions externally. The time of the process does not necessarily correspond to any physical interpretation. By solving the stationary fluctuation-dissipation relations (FDRs), then we get the statistical partitioning error $Q_x^2=1/\langle x\rangle$ (Huh and Paulsson, 2011b). The simulation result is shown in Fig. S3A-C.

Candidate gene selection with a similar ratio pattern between Carm1 and Cdx2

We select genes that show an inverse expression relationship with Cdx2 (similar to Carm1), and those showing inverse expression relationship with Carm1 (similar to Cdx2) (Table S2) based on the correlation coefficient R of each gene in three eight-cell embryos. Genes were selected following the rules below:

$$\begin{aligned}A_i&=\left\{\frac{exp(candidate\ gene)_j}{exp(candidate\ gene)_j+exp(reference\ gene\ a)_j}\right\}\\ B_i&=\left\{\frac{exp(reference\ gene\ b)_j}{exp(reference\ gene\ b)_j+exp(reference\ gene\ a)_j}\right\}\end{aligned}. \quad (9)$$

Array A_i lists the percentages between candidate gene and reference gene a of each blastomere in order number i embryo. j represents the order number of blastomeres.

Array B_i contains the percentages between reference gene b and reference gene a of each blastomere in order number i embryo. In the present manuscript, genes a and b can be Carm1(Cdx2) and Cdx2(Carm1). The genes with correlation coefficient R between array A_i and B_i >0.7 in every embryo were selected, deemed as highly correlated with the expression pattern of reference gene b.

Statistical analysis

Statistical analysis was conducted using GraphPad Prism 6.0 software. One-way ANOVA was used for statistical analysis. For all statistical analyses, a value of P<0.05 was considered statistically significant.

Acknowledgements

We thank Drs Lei Li and Haibin Wang at the Institute of Zoology, Chinese Academy of Sciences for helpful discussions; and Mr Ming Ge at the Institute of Zoology, Chinese Academy of Sciences and Ms Yue Wang at the National Institute of Biological Sciences, Beijing, for their help in mouse breeding and embryo manipulation.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Q.C. and J.S. conceived the project and analyzed the single-cell datasets. J.S., X.Z. and Y.T. constructed the model. Q.C., J.Q., F.T., Y.Z., Y.T., Q.Z. and E.D. contributed to the development and analysis of the model. X.L. performed mouse embryo related experiments. Q.C., J.S. and X.Z. contributed to the writing of the manuscript. Y.T., Q.Z. and E.D. supervised the whole project and proofread the manuscript.

Funding

This research is supported by the National Basic Research Program of China [2015CB943000, 2011CB944401];the Strategic Priority Research Program of the Chinese Academy of Sciences [XDA01000000, XDA04020202-20]; the National Natural Science Foundation of China [81490741, 31200879, 31300957, 11401562]; and a Fellowship of Youth Innovation Promotion Association, Chinese Academy of Sciences [201306 to Q.C.].

Supplementary information

Supplementary information available online at
http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.123950/-/DC1

References

Alon, U. (2007). Network motifs: theory and experimental approaches. *Nat. Rev. Genet.* **8**, 450-461.

Audic, S. and Claverie, J. M. (1997). The significance of digital gene expression profiles. *Genome Res.* **7**, 986-995.

Berg, O. G. (1978). A model for the statistical fluctuations of protein numbers in a microbial population. *J. Theor. Biol.* **71**, 587-603.

Biase, F. H., Cao, X. and Zhong, S. (2014). Cell fate inclination within 2-cell and 4-cell mouse embryos revealed by single-cell RNA sequencing. *Genome Res.* **24**, 1787-1796.

Bruce, A. W. and Zernicka-Goetz, M. (2010). Developmental control of the early mammalian embryo: competition among heterogeneous cells that biases cell fate. *Curr. Opin. Genet. Dev.* **20**, 485-491.

Cai, L., Friedman, N. and Xie, X. S. (2006). Stochastic protein expression in individual cells at the single molecule level. *Nature* **440**, 358-362.

Cockburn, K., Biechele, S., Garner, J. and Rossant, J. (2013). The hippo pathway member Nf2 is required for inner cell mass specification. *Curr. Biol.* **23**, 1195-1201.

Davidson, E. H. (2010). Emerging properties of animal gene regulatory networks. *Nature* **468**, 911-920.

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 Hiiragi, 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.

Huh, D. and Paulsson, J. (2011a). Random partitioning of molecules at cell division. *Proc. Natl. Acad. Sci. USA* **108**, 15004-15009.

Huh, D. and Paulsson, J. (2011b). Non-genetic heterogeneity from stochastic partitioning at cell division. *Nat. Genet.* **43**, 95-100.

Itzkovitz, S., Lyubimova, A., Blat, I. C., Maynard, M., van Es, J., Lees, J., Jacks, T., Clevers, H. and van Oudenaarden, A. (2012). Single-molecule transcript counting of stem-cell markers in the mouse intestine. *Nat. Cell Biol.* **14**, 106-114.

Jedrusik, A., Parfitt, D.-E., Guo, G., Skamagki, M., Grabarek, J. B., Johnson, M. H., Robson, P. and Zernicka-Goetz, M. (2008). Role of Cdx2 and cell polarity in cell allocation and specification of trophectoderm and inner cell mass in the mouse embryo. *Genes Dev.* **22**, 2692-2706.

Lee, M. T., Bonneau, A. R. and Giraldez, A. J. (2014). Zygotic genome activation during the maternal-to-zygotic transition. *Annu. Rev. Cell Dev. Biol.* **30**, 581-613.

Li, L., Lu, X. and Dean, J. (2013). The maternal to zygotic transition in mammals. *Mol. Aspects Med.* **34**, 919-938.

Lorthongpanich, C., Doris, T. P. Y., Limviphuvadh, V., Knowles, B. B. and Solter, D. (2012). Developmental fate and lineage commitment of singled mouse blastomeres. *Development* **139**, 3722-3731.

MacArthur, B. D., Sevilla, A., Lenz, M., Müller, F.-J., Schuldt, 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. *Nat. Cell Biol.* **14**, 1139-1147.

Montserrat, N., Nivet, E., Sancho-Martinez, I., Hishida, T., Kumar, S., Miquel, L., Cortina, C., Hishida, Y., Xia, Y., Esteban, C. R. et al. (2013). Reprogramming of human fibroblasts to pluripotency with lineage specifiers. *Cell Stem Cell* **13**, 341-350.

Morris, S. A., Teo, R. T. Y., Li, H., Robson, P., Glover, D. M. and Zernicka-Goetz, M. (2010). Origin and formation of the first two distinct cell types of the inner cell mass in the mouse embryo. *Proc. Natl. Acad. Sci. USA* **107**, 6364-6369.

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5**, 621-628.

Ohnishi, Y., Huber, W., Tsumura, A., Kang, M., Xenopoulos, P., Kurimoto, K., Oles, A. K., Araúzo-Bravo, M. J., Saitou, M., Hadjantonakis, A.-K. et al. (2014). Cell-to-cell expression variability followed by signal reinforcement progressively segregates early mouse lineages. *Nat. Cell Biol.* **16**, 27-37.

Piras, V., Tomita, M. and Selvarajoo, K. (2014). Transcriptome-wide variability in single embryonic development cells. *Sci. Rep.* **4**, 7137.

Plachta, N., Bollenbach, T., Pease, S., Fraser, S. E. and Pantazis, P. (2011). Oct4 kinetics predict cell lineage patterning in the early mammalian embryo. *Nat. Cell Biol.* **13**, 117-123.

Raj, A., Peskin, C. S., Tranchina, D., Vargas, D. Y. and Tyagi, S. (2006). Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.* **4**, e309.

Rigney, D. R. (1979). Stochastic model of constitutive protein levels in growing and dividing bacterial cells. *J. Theor. Biol.* **76**, 453-480.

Rossant, J. and Tam, P. P. L. (2009). Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. *Development* **136**, 701-713.

Shu, J., Wu, C., Wu, Y., Li, Z., Shao, S., Zhao, W., Tang, X., Yang, H., Shen, L., Zuo, X. et al. (2013). Induction of pluripotency in mouse somatic cells with lineage specifiers. *Cell* **153**, 963-975.

Smits, W. K., Kuipers, O. P. and Veening, J.-W. (2006). Phenotypic variation in bacteria: the role of feedback regulation. *Nat. Rev. Microbiol.* **4**, 259-271.

Sprinzak, D. and Elowitz, M. B. (2005). Reconstruction of genetic circuits. *Nature* **438**, 443-448.

Steuerwald, N., Cohen, J., Herrera, R. J. and Brenner, C. A. (2000). Quantification of mRNA in single oocytes and embryos by real-time rapid cycle fluorescence monitored RT-PCR. *Mol. Hum. Reprod.* **6**, 448-453.

Streets, A. M., Zhang, X., Cao, C., Pang, Y., Wu, X., Xiong, L., Yang, L., Fu, Y., Zhao, L., Tang, F. et al. (2014). Microfluidic single-cell whole-transcriptome sequencing. *Proc. Natl. Acad. Sci. USA* **111**, 7048-7053.

Sturn, A., Quackenbush, J. and Trajanoski, Z. (2002). Genesis: cluster analysis of microarray data. *Bioinformatics* **18**, 207-208.

Swain, P. S., Elowitz, M. B. and Siggia, E. D. (2002). Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc. Natl. Acad. Sci. USA* **99**, 12795-12800.

Tabansky, I., Lenarcic, A., Draft, R. W., Loulier, K., Keskin, D. B., Rosains, J., Rivera-Feliciano, J., Lichtman, J. W., Livet, J., Stern, J. N. H. et al. (2013). Developmental bias in cleavage-stage mouse blastomeres. *Curr. Biol.* **23**, 21-31.

Tadros, W. and Lipshitz, H. D. (2009). The maternal-to-zygotic transition: a play in two acts. *Development* **136**, 3033-3042.

Takaoka, K. and Hamada, H. (2012). Cell fate decisions and axis determination in the early mouse embryo. *Development* **139**, 3-14.

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B. B., Siddiqui, A. et al. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nat. Methods* **6**, 377-382.

Tang, F., Barbacioru, C., Nordman, E., Bao, S., Lee, C., Wang, X., Tuch, B. B., Heard, E., Lao, K. and Surani, M. A. (2011). Deterministic and stochastic allele specific gene expression in single mouse blastomeres. *PLoS ONE* **6**, e21208.

Torres-Padilla, M.-E., Parfitt, D.-E., Kouzarides, T. and Zernicka-Goetz, M. (2007). Histone arginine methylation regulates pluripotency in the early mouse embryo. *Nature* **445**, 214-218.

Vilar, J. M. G., Kueh, H. Y., Barkai, N. and Leibler, S. (2002). Mechanisms of noise-resistance in genetic oscillators. *Proc. Natl. Acad. Sci. USA* **99**, 5988-5992.

Wan, H., He, Z., Dong, M., Gu, T., Luo, G. Z., Teng, F., Xia, B., Li, W., Feng, C., Li, X. et al. (2013). Parthenogenetic haploid embryonic stem cells produce fertile mice. *Cell Res.* **23**, 1330-1333.

Wennekamp, S. and Hiiragi, T. (2012). Stochastic processes in the development of pluripotency in vivo. *Biotechnol. J.* **7**, 737-744.

Xie, D., Chen, C. C., Ptaszek, L. M., Xiao, S., Cao, X., Fang, F., Ng, H. H., Lewin, H. A., Cowan, C. and Zhong, S. (2010). Rewirable gene regulatory networks in the preimplantation embryonic development of three mammalian species. *Genome Res.* **20**, 804-815.

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.

Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, R., Liu, P., Lian, Y., Zheng, X., Yan, J. et al. (2013). Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. *Nat. Struct. Mol. Biol.* **20**, 1131-1139.

Zernicka-Goetz, M. (2006). The first cell-fate decisions in the mouse embryo: destiny is a matter of both chance and choice. *Curr. Opin. Genet. Dev.* **16**, 406-412.

Zernicka-Goetz, M. (2013). Development: do mouse embryos play dice? *Curr. Biol.* **23**, R15-R17.

Zernicka-Goetz, M., Morris, S. A. and Bruce, A. W. (2009). Making a firm decision: multifaceted regulation of cell fate in the early mouse embryo. *Nat. Rev. Genet.* **10**, 467-477.

Zheng, X.-D., Yang, X.-Q. and Tao, Y. (2011). Bistability, probability transition rate and first-passage time in an autoactivating positive-feedback loop. *PLoS ONE* **6**, e17104.