SCIENTIFIC REPORTS

OPEN

SUBJECT AREAS:
TRANSCRIPTION
EMBRYONIC STEM CELLS
CELLULAR NOISE
SINGLE-CELL IMAGING

Received
7 August 2014

Accepted
30 October 2014

Published
20 November 2014

Correspondence and requests for materials should be addressed to H.O. (ochiai@hiroshima-u.ac.jp)

Stochastic promoter activation affects Nanog expression variability in mouse embryonic stem cells

Hiroshi Ochiai[1], Takeshi Sugawara[1], Tetsushi Sakuma[2] & Takashi Yamamoto[1,2]

[1]Research Center for the Mathematics on Chromatin Live Dynamics (RcMcD), Hiroshima University, Higashi-Hiroshima 739-8530, Japan, [2]Department of Mathematical and Life Sciences, Graduate School of Science, Hiroshima University, Higashi-Hiroshima 739-8526, Japan.

Mouse embryonic stem cells (mESCs) are self-renewing and capable of differentiating into any of the three germ layers. An interesting feature of mESCs is the presence of cell-to-cell heterogeneity in gene expression that may be responsible for cell fate decisions. Nanog, a key transcription factor for pluripotency, displays heterogeneous expression in mESCs, via mechanisms that are not fully understood. To understand this variability, we quantitatively analyzed *Nanog* transcription and found that *Nanog* was both infrequently transcribed, and transcribed in a pulsatile and stochastic manner. It is possible that such stochastic transcriptional activation could contribute to the heterogeneity observed in *Nanog* expression as "intrinsic noise." To discriminate the effects of both intrinsic noise from other (extrinsic) noise on the expression variability of *Nanog* mRNA, we performed allele-specific single-molecule RNA fluorescent *in situ* hybridization in a reporter cell line and found that intrinsic noise contributed to approximately 45% of the total variability in *Nanog* expression. Furthermore, we found that *Nanog* mRNA and protein levels were well correlated in individual cells. These results suggest that stochastic promoter activation significantly affects the Nanog expression variability in mESCs.

mESCs are clonal cell lines derived from the pre-implantation epiblast. They are capable of self-renewing in media containing leukemia inhibitory factor (LIF), can differentiate into three germ layers, and display heterogeneity in gene expression. Uncovering the nature of this heterogeneity at the molecular level is important to the understanding of how stem cells modulate their stemness/differentiation balance. Nanog, a key transcription factor for pluripotency, exhibits a broad range of expression levels, and its heterogeneity seems to be related to the stemness/differentiation balance[1–4].

The autocrine fibroblast growth factor (FGF)/extracellular signal-regulated kinase (Erk) signaling that induces lineage commitment and inhibits Nanog expression is thought to be a source of Nanog heterogeneity[5–7], along with the positive[8] and negative feedback loops[9] in the gene regulatory networks responsible for pluripotency. Recently, it has been reported that *Nanog* is prone to be monoallelically transcribed (i.e., transcribed from only one copy of *Nanog*) in mESCs cultured in standard mESC medium containing serum and LIF (serum condition), even though mice have two copies of *Nanog*[9–11]. If this effect is a result of an intrinsically stochastic effect on promoter activation, it may also contribute to Nanog heterogeneity[12,13].

In this study, to gain insight into the contribution of *Nanog* transcriptional activity to Nanog heterogeneity, we quantitatively analyzed the transcription dynamics of endogenous *Nanog* and distribution of *Nanog* mRNA in a mESC population at single-cell resolution and observed infrequent and stochastic switching on and off of *Nanog* promoter states. Furthermore, we found that expression noise stemming from such promoter dynamics significantly affected heterogeneity in Nanog expression.

Results

Establishment of a Nanog-MS2 mESC line. To quantitatively analyze *Nanog* transcription in mESCs, we applied the MS2 system[14,15]. The transcribed MS2 sequence derived from MS2 bacteriophage forms a stem-loop structure, which is known to be bound by the MS2 coat protein (MCP) as a dimer (Figure 1a). Therefore, the integration of the 24 tandem MS2 sites into a specific gene of interest and expression of MCP fused with fluorescent protein enables the visualization of mRNA transcription as bright spots in the nucleus[16–18]. We applied transcription activator-like effector nuclease (TALEN)-mediated targeted integration of the MS2 repeat sequence into the

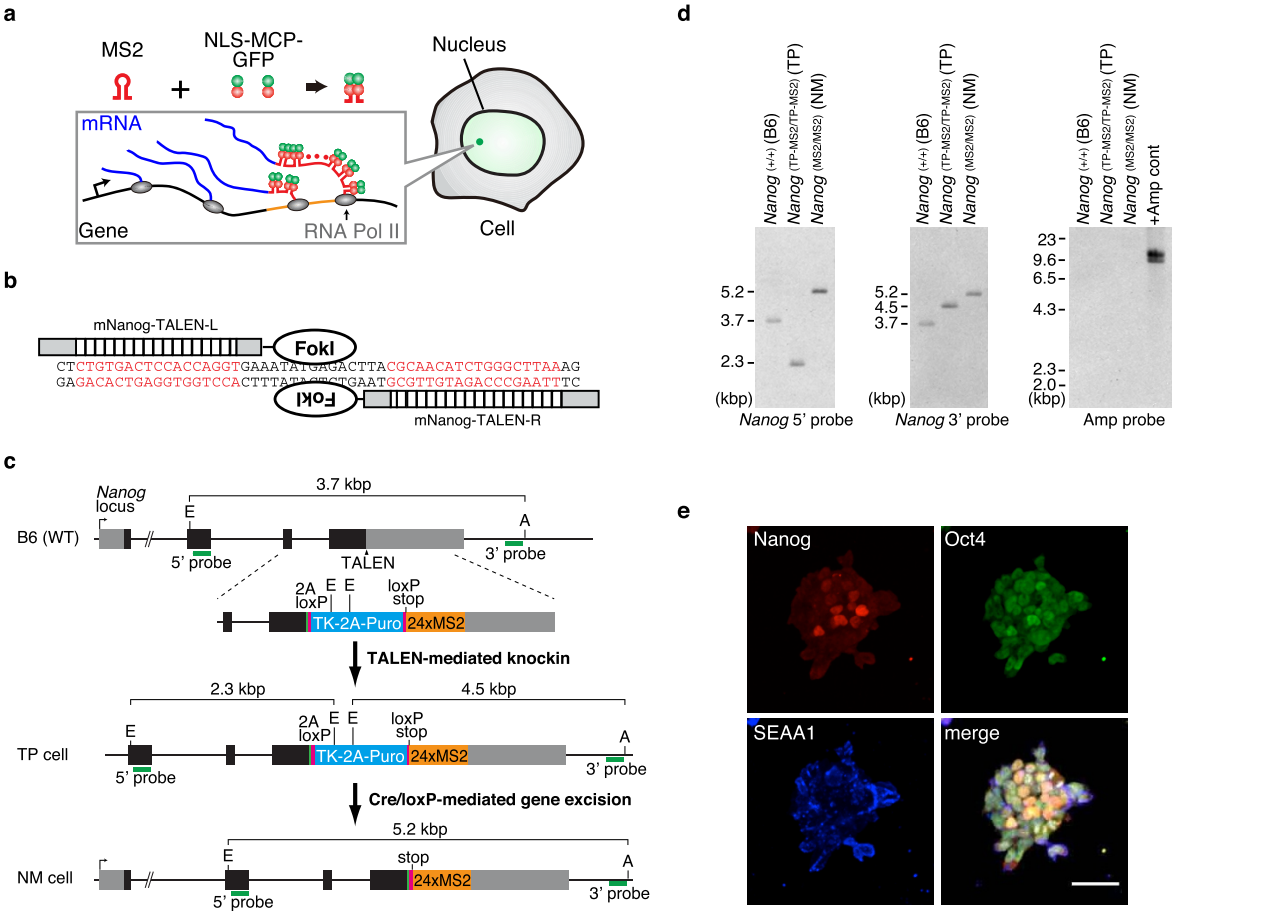

Figure 1 | Establishment of MS2-targeted ES cell line. (a) Schematic representation of the MS2 system. In the inset, black, orange, blue, and red lines represent a gene, its integrated MS2 repeat, the transcribed nascent mRNAs, and the transcribed MS2 repeats, respectively. Because of the accumulation of nuclear localization signal (NLS)-MS2 coat protein (MCP)-GFP fusion protein at the transcription site, the bright fluorescent spot could be observed at the site in the nucleus. (b) Schematic representations of transcription activator-like effector nucleases (TALENs) used in this study and their respective target nucleotide (highlighted in red letters). (c) The strategy for biallelic targeted gene integration into the *Nanog* locus. Grey, black, green, and magenta boxes indicate untranslated regions, Nanog coding sequences, and coding sequences for 2A peptide and loxP site, respectively. Green bars represent the positions of Southern blot probes. E, EcoNI; A, AflII. (d) Southern blot analyses showing TALEN-mediated biallelic insertion and Cre-mediated biallelic excision of the selection cassettes in modified mESC clones. (e) Immunofluorescence displays expression of pluripotent markers in the targeted cell line (NM cell). Each image is a maximum intensity projection of image stacks. Scale bar, 50 μm.

Nanog loci[19,20] (Figure 1b, c), and established a biallelically targeted mESC line (Figure 1d). Then, selection cassettes were removed by transient expression of Cre recombinase (Figure 1c, d). Southern blot analysis showed the expected band patterns and no random integrations (Figure 1d). We refer to the obtained cell line as Nanog-MS2 (NM) cells. The cell line expressed the undifferentiated embryonic stem-cell markers SSEA1 and Oct4 (also known as Pou5f1) (Figure 1e).

Next, to check the cell-to-cell variability of *Nanog* mRNA copy numbers in the NM cell line, we performed single-molecule fluorescent *in situ* hybridization (smFISH) using *Nanog* exonic probes (Figure 2a–c)[21]. In the serum conditions used (standard mESC culture medium containing LIF and serum), the mean count of *Nanog* mRNA copies in NM cells was slightly, but significantly, lower than that of the parental mESC line (133 for the parental mESC line and 77 for the NM cell line) (Figure 2c). However, the degrees of variability between the parental and derived lines were comparable (coefficient of variation [CV] for the parental and NM cell lines were 0.85 and 0.89, respectively). One of the possible causes of the difference in expression is the change in mRNA stability due to insertion of the MS2 repeats. To explore the possibility, we examined the half-lives of *Nanog* mRNA in NM and the parental mESCs (Supplementary Figure S1), and found that insertion of MS2 repeats slightly destabilizes *Nanog* mRNA. Therefore, the mean number of *Nanog* mRNA in NM cells appears lower than that in the parental mESCs.

It has been reported that the cytokine FGF4 is secreted by undifferentiated cells, acts in an autocrine manner[22,23], and represses *Nanog* expression through Erk signaling, suggesting that FGF/Erk signaling is a contributor to the transcription factor heterogeneity[6,24]. When cell lines were cultured in medium containing FGF/Erk and GSK3β inhibitors (2i conditions), the means of *Nanog* mRNA copies were increased, as expected (Figure 2c). Furthermore, the variability of these cells was lower than cells grown in the serum condition (CV for parental and NM cell lines in 2i conditions were 0.43 and 0.42, respectively). The CVs in parent and NM cells were comparable (Figure 2c). These findings suggest that the NM cell line displays similar heterogeneity to that of the parent cell line.

Next, to visualize transcriptional activity, we established a cell line that constitutively expresses nuclear localization signal (NLS)-MCP-green fluorescent protein (GFP) (mNeonGreen) (NM-G cell line). In this cell line, either none, one or two bright fluorescent spots were observed (Figure 2d). To confirm whether these spots corresponded with the *Nanog* transcription sites, we performed smFISH using a *Nanog* intronic probe set (Figure 2a, e). Although a majority of the smFISH and NLS-MCP-GFP spot signals coincided in cells cultured in 2i conditions (88%, n = 101), only one type of signal was detected

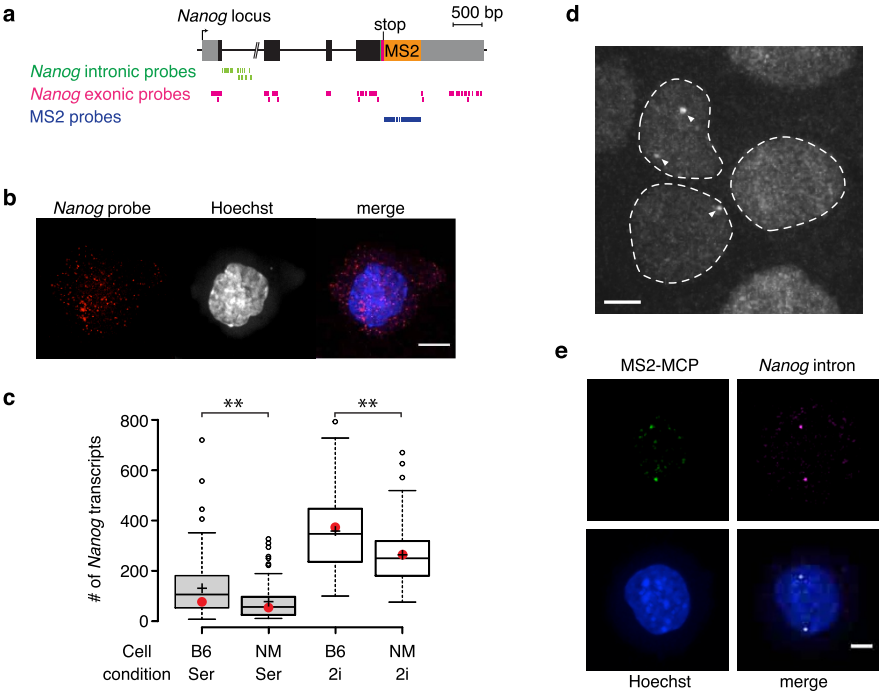

Figure 2 | Nanog-MS2 mESC line is useful for quantifying the *Nanog* transcription dynamics. (a) An illustration of the *Nanog* genomic locus and positions of smFISH probes used in this study. (b) Single-molecule fluorescent *in situ* hybridization (smFISH) analysis. *Nanog* mRNAs in a cell were visualized using *Nanog* exonic smFISH probes. Nuclei were counter-stained with Hoechst33342. Scale bar, 10 μm. (c) Distributions of *Nanog* mRNAs in mESC lines cultured in either serum (Ser) or 2i conditions (2i). The number of *Nanog* mRNAs in parental C57BL/6 (B6) or NM cell lines (NM) was counted by smFISH using *Nanog* exonic probes. In the box plot, center lines show the medians; box limits indicate the 25th and 75th percentiles; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles; outliers are represented by dots; crosses represent sample means. n = 129, 109, 102, and 119 sample points. Statistical significance of differences was assessed by a two-sample Kolmogorov-Smirnov test (**$P \ll 0.01$). Red dots represent the mean counts of *Nanog* transcripts predicted by the random telegraph model (see the main text). (d) Bright fluorescent spots as observed in NM-G cells. Representative NM-G cells in 2i conditions were subjected to live imaging. The image shown is a maximum intensity projection of image stacks. The dashed lines delineate individual nuclei of cells. The white arrowheads point out bright fluorescent spots assumed to be *Nanog* transcription sites. Scale bar, 5 μm. (e) Results of smFISH analysis in NM-G cells. A representative image shows the overlap of fluorescent spots of nuclear localization signal (NLS)-MS2 coat protein (MCP)-mNeonGreen (mNG) and *Nanog* intronic smFISH probes, indicating that the MS2-MCP spot signals coincide with *Nanog* transcription sites. Scale bar, 5 μm.

in some of the spots. This limitation could be attributed to the distance between the sites of *Nanog* intronic probes and the MS2 repeat (~4.5 kb, Fig. 2a). The RNA Pol II elongation rate in mESCs was reported to be 0.5–4 kb/min[25]; therefore, all signals of smFISH and MS2-MCP may not be simultaneously detected at the same position. These findings suggest that the MS2-MCP spot signals indicate the transcription of MS2-integrated *Nanog* alleles and that the NM-G cell line is useful for quantifying the *Nanog* transcription dynamics.

Quantification of *Nanog* transcription dynamics. To quantify *Nanog* transcription dynamics, we performed live imaging of NM-G cells cultured in serum or 2i conditions at 2-min intervals for 4 h (Figure 3a, b, Supplementary Video S1 and S2). In both conditions, *Nanog* was transcribed in a pulsatile manner, as reported in other model systems[26,27]. The transcription frequency over the 4-h time period was higher in 2i conditions than in serum conditions, consistent with previous reports (Figure 3c)[10,11].

Consistent with the smFISH data (Figure 2c), transcription frequencies were apparently variable among cells even in 2i conditions (Figure 3a–c). 2i conditions inhibit autocrine FGF/Erk signaling, one potential source of Nanog heterogeneity. This finding suggests that factors other than FGF/Erk signaling may significantly affect the expression variability of Nanog. Furthermore, transcription frequency during the 4-h imaging period seemed to show multimodal distributions in both conditions (Supplementary Figure S2). This suggests the existence of multiple states of *Nanog* expression that switch at intervals longer than 4 hours, as recently reported[28]. On the other hand, the distributions of transcriptional signal intensities in the two conditions were comparable (Figure 3d), suggesting that *Nanog* transcription was regulated not by modulation of the number of transcripts per transcriptional pulse, but rather by transcription frequency[27]. This kind of transcription dynamics might be explained by the random telegraph model[29–31] (Figure 3e). In this model, promoter states stochastically switch "ON", permitting transcription, and "OFF", deactivating transcription (Figure 3f). The ON and OFF times are predicted to be exponentially distributed. To determine those parameters from imaging data, transcriptional dynamics of each allele should be separately quantified. Although non-transcribing alleles were not actually traceable, their position is roughly estimated by the somewhat uneven nuclear background signals of NLS-MCP-mNeonGreen and changes in nuclear shape (Supplementary Video S1 and S2). By this qualitative estimation, distributions of ON and OFF time durations were obtained (Figure 3g). ON and OFF time distributions were well fitted by exponential distributions, consistent with the Poisson stochastic processes expected in the random telegraph model[30]. Furthermore, the mean duration of ON time in 2i conditions was longer than that in serum conditions. Conversely, the mean duration of OFF time in 2i conditions was shorter than that in serum conditions. These results are consistent with the increase in transcription frequency in 2i conditions (Figure 3c). Collectively, these findings suggest that *Nanog* transcription might be regulated by modulation of promoter states.

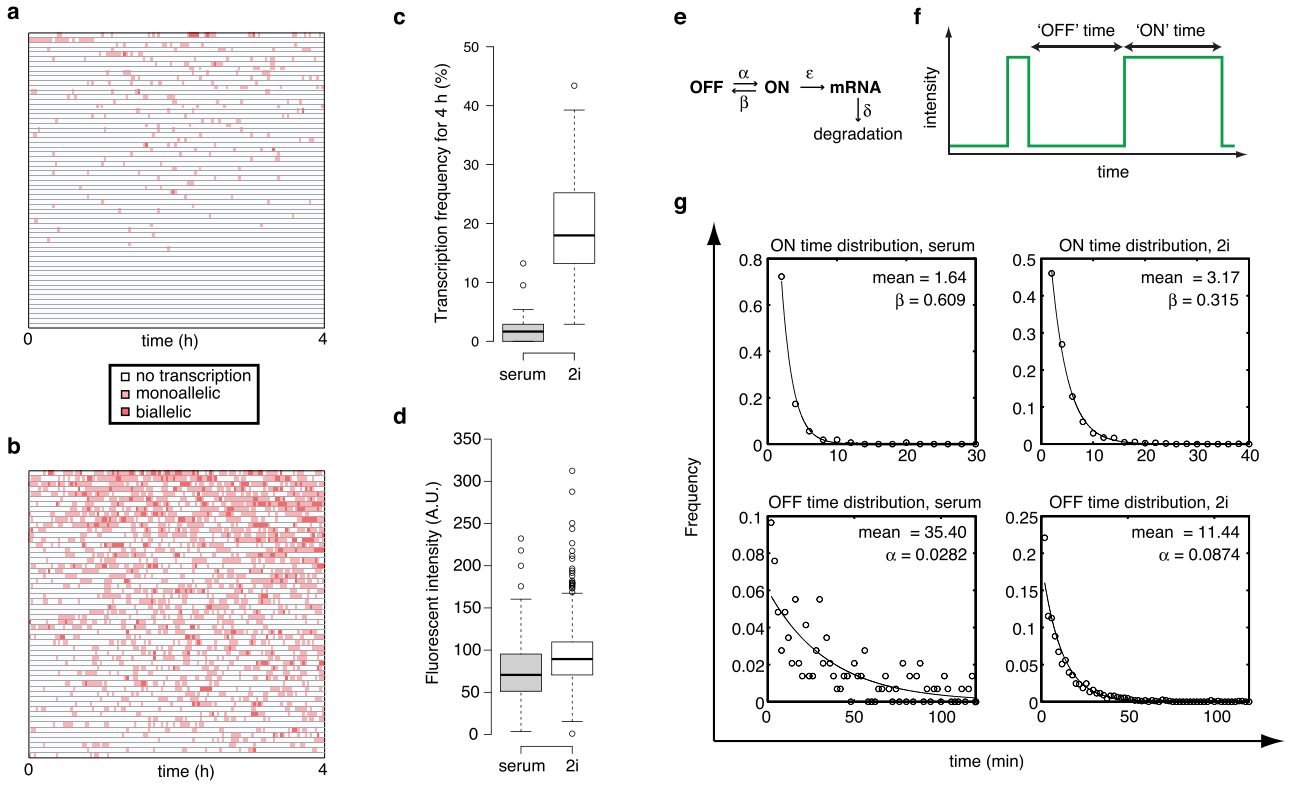

Figure 3 | Transcription dynamics of *Nanog* in mESCs. (a, b) NM-G cells cultured in serum (a) and 2i conditions (b) were imaged at 2-min intervals for 4 h. Each line of the color-coded graph represents the transcription dynamics of each cell. (c) Distribution of transcription frequencies over a 4-h period for NM-G cells cultured in serum or 2i conditions [(total counts of transcribed alleles in each cell at 121 time points)/242] (n = 62 and 56 for serum and 2i conditions, respectively). (d) Distribution of fluorescence intensity at the transcription site in cells cultured in serum or 2i conditions. Each value was subtracted by the mean fluorescent intensity of each nucleus (n = 171 and 565 sample points). (e) A schematic presentation of the random telegraph model. "OFF" and "ON" depicts the OFF and ON states of promoters, respectively. OFF to ON and ON to OFF transition rates were defined as α and β, respectively. mRNA is assumed to be transcribed from only ON state promoter at rate ε. mRNA is degraded at rate δ. (f) A schematic representation of fluorescence intensity dynamics of an allele. (g) Distributions of ON and OFF time of *Nanog* transcription in cells cultured in serum or 2i conditions. Each probability distribution was fitted with an exponential function. Mean values are displayed in each graph.

If the *Nanog* transcription kinetics could be explained by the random telegraph model, the mean quantity of mRNA at steady state can be predicted as $<mRNA> = \gamma(\alpha/(\alpha+\beta))(\varepsilon/\delta)$ (Figure 3e). Here, γ is the copy number of *Nanog* gene; α and β are the reciprocal of the means of OFF and ON time durations, respectively (Figure 3g). The degradation rate of MS2-integrated and wild-type (WT) *Nanog* mRNA, δ, was determined to be 0.00348 and 0.00245 min^{-1}, respectively (Supplementary Figure S1). A least square fit of the random telegraph model to the experimentally obtained means of *Nanog* transcripts in NM and parent cells cultured in serum and 2i conditions reveals that the transcription rate at ON state ε = 2.11 min^{-1} (Figure 2c). In both cell types cultured in serum conditions, the fitted means of *Nanog* mRNAs were considerably lower than those of the experimentally obtained values. Although we performed live imaging at 2-min intervals, the estimated mean duration of ON time in serum conditions was shorter than 2 min (1.64 min, Figure 3g). Therefore, it is possible that not all the transcriptional pulses had been detected and that we underestimated the transcriptional frequency in serum conditions. To confirm whether the value of ε is realistic or not, we performed further quantitative analysis. The value of ε could be predicted from the number of nascent mRNAs remaining at each transcription site. When we imaged the NM-G cells at higher magnification, not only transcriptional bright spots but also other relatively weak and fast-moving signals, which are assumed to be individual mRNAs, were observed (Supplementary Video S3 and Figure S3).

From the comparison of individual mRNA signals and transcriptional spot signals, we estimate that there are 2.79 ± 0.55 [mean ± standard deviation, n = 15] nascent NLS-MCP-mNeonGreen-tagged mRNAs per transcription site. The sum length of an MS2 repeat and the 3′ untranslated region of *Nanog* mRNA is 2.55 kb; therefore, each nascent mRNA is roughly estimated to be transcribed by RNA Pol II at 946 ± 184-bp intervals. Accordingly, the RNA Pol II elongation rate was estimated to be 1.99 ± 0.39 kb/min.Recently, the RNA Pol II elongation rate was reported to be 0.5–4 kb/min (mean and median are 1.793 and 1.824 kb/min, respectively) in mESCs[25] (Supplementary Figure S4), suggesting that the estimated ε seems to be a realistic value.

In the random telegraph model, mRNAs are transcribed in busts during promoter switches from an inactive to active state, and the average size of the transcriptional bursts should be described as ε/β[30]. Because of the increase in the mean ON time duration in 2i conditions, the burst size in 2i conditions is larger than that in serum conditions (3.46 and 6.69 for serum and 2i conditions, respectively). Collectively, these suggest that *Nanog* transcription can be explained by the random telegraph model.

Stochastic promoter activation significantly contributes to expression variability of Nanog. Heterogeneity in gene expression is induced by several factors[12]. One factor is the presence of stochasticity that is inherent to the biochemical process of gene expression, called the 'intrinsic noise'. Therefore, the stochastic promoter activation observed in *Nanog* transcription could be a result of intrinsic noise. On the other hand, other effects, including variability of cellular components (such as RNA Pol II or other regulatory molecules), asynchronous cell cycle, and heterogeneous

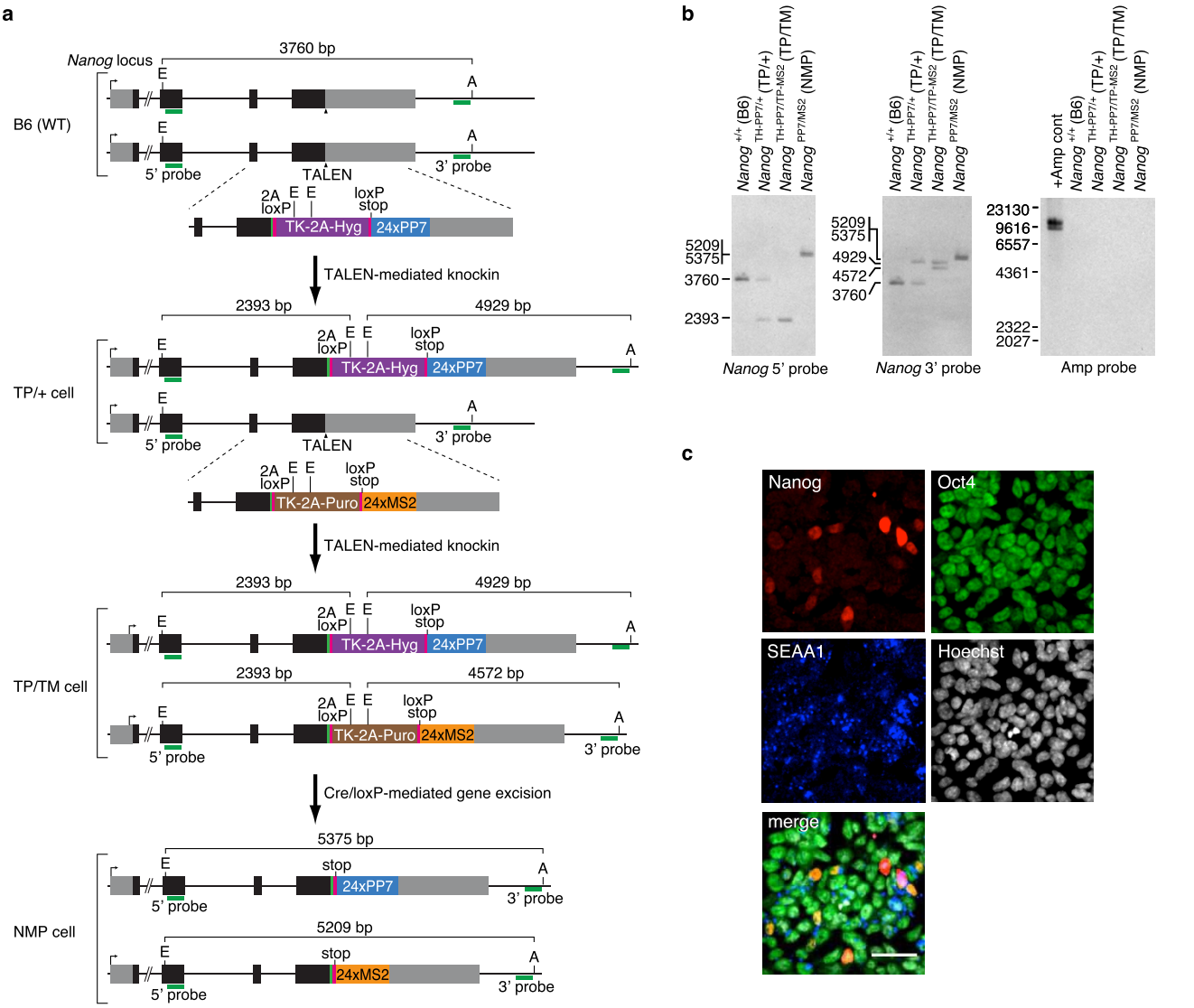

Figure 4 | Establishment of a reporter mESC line (NMP). (a) The strategy for targeted gene integration into the *Nanog* locus for establishment of MS2-PP7 targeted mouse embryonic stem cell (mESC) line (NMP). Grey, black, green, and magenta boxes indicate untranslated region, Nanog coding sequences, and coding sequences for 2A peptide and loxP site, respectively. Green bars represent the positions of Southern blot probes. E, EcoNI; A, AflII. (b) Southern blot analyses showing transcription activator-like effector nuclease (TALEN)-mediated insertion and Cre-mediated excision of the selection cassettes in modified mESC clones. (c) Immunofluorescence showed expression of pluripotent markers in the targeted cell line (NMP cells). Each image is a maximum intensity projection of image stacks. Scale bar, 50 μm.

inter-cellular signaling also affect gene expression variability and are called 'extrinsic noise'. These two types of noise can be discriminated by two-reporter assays[12,13]. If different reporters are integrated into each allele, distribution of each reporter enables the determination of how intrinsic noise contributes to the total expression variability in the population[12,13].

To determine the type of noise that is dominant in the expression of *Nanog*, we established another cell line (NMP) in which repeats of MS2 and PP7 were integrated immediately upstream of each allele of the endogenous *Nanog* stop codon (Figure 4a). PP7 is used to visualize mRNAs[32]. Southern blot analysis showed the expected band patterns and no random integrations (Figure 4b). The derived cells express the undifferentiated embryonic stem-cell markers, SSEA1 and Oct4 (Figure 4c). To quantify *Nanog* transcription dynamics on each allele, we introduced NLS-MCP-mNeonGreen and NLS-PP7 coat protein (PCP)-red fluorescent protein (RFP), which binds to the PP7 RNA stem loop. However, a tendency toward nucleoli localization of NLS-PCP-RFP prevents quantification of the transcription dynamics of *Nanog-PP7*. Using *Nanog* exonic and MS2 probes, we performed smFISH analysis of NMP cells cultured in serum and 2i conditions (Figure 2a and 5a). Among the *Nanog* probe-positive spots, those with an MS2 probe signal intensity above the threshold value were assumed to be mRNAs expressed from the *Nanog-MS2* allele; otherwise, mRNAs were considered to be expressed from the *Nanog-PP7* allele (Figure 5b, Supplementary Figure S5). To confirm the accuracy of this methodology, we also performed the smFISH using MS2 and PP7 probes in NMP cells cultured in 2i conditions (Supplementary Figure S5). The distribution of mRNA counts was similar to that obtained using *Nanog* exonic and MS2 probes (Figure 5b), suggesting that the method is reasonably accurate.

Intrinsic noise η^2_{int}, extrinsic noise η^2_{ext}, and total noise η^2_{tot} (sum of intrinsic and extrinsic noises) can be calculated using obtained allele-specific mRNA distribution data[12,13] (Figure 5b). Consistent with other reports, the total noise (heterogeneity in *Nanog* expression) in serum conditions was larger than that in 2i conditions (Figure 5b)[6]. Surprisingly, intrinsic noise contributed to approximately 45% of the total noise in both conditions, sug-

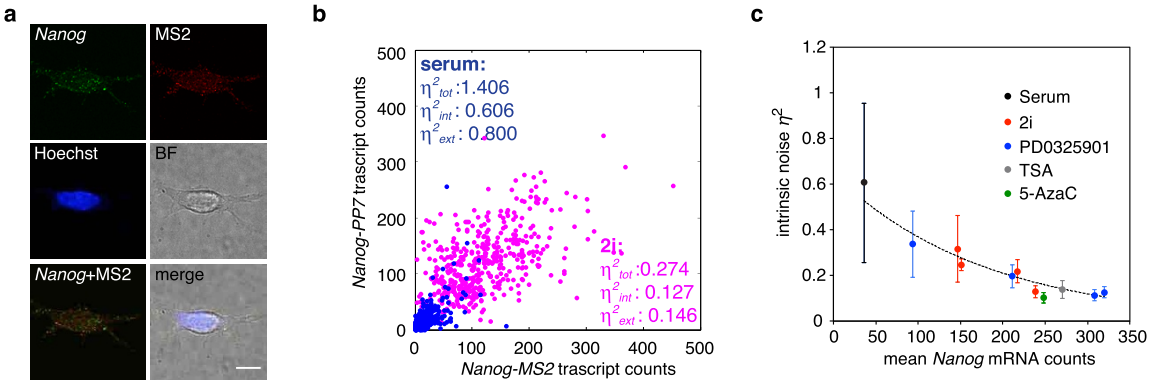

Figure 5 | Intrinsic noise significantly affects expression variability of *Nanog*. (a) Allele-specific single-molecule fluorescent *in situ* hybridization (smFISH) analysis. NMP cells cultured in serum condition were subjected to smFISH analysis using *Nanog* exonic smFISH probes and MS2 probes. Cellular nuclei were counter-stained with Hoechst33342. Each image is a maximum intensity projection of image stacks. The probe position is described in Figure 2a. Scale bar, 10 μm. (b) Scatter plot of *Nanog-MS2* and *-PP7* transcripts. *Nanog-MS2* and *-PP7* transcripts were counted in each cell by allele-specific smFISH. η^2_{int}, η^2_{ext}, and η^2_{tot} represent intrinsic, extrinsic and total noise, respectively. The Fano factors of *Nanog* mRNA are 39.9 and 50.3 in serum and 2i conditions, respectively. (c) Scatter plot of mean *Nanog* mRNA counts and intrinsic noise. Distributions of *Nanog-MS2* and *-PP7* mRNAs in cells cultured in several conditions were obtained by allele-specific smFISH analyses as in (b). Afterward, the mean *Nanog* mRNA counts and noise values were calculated and plotted. Dashed line represents a trend line obtained from data of cells cultured in media containing 2i and PD0325901 as calculated using Microsoft Excel. Error bars are 95% confidence intervals obtained by bootstrapping. Concentrations of inhibitors in each culture condition are listed in Supplementary Table S1.

gesting that intrinsic noise significantly affects *Nanog* expression variability.

Models of stochastic gene expression predict that intrinsic noise should increase as the amount of transcript decreases[33]. To change the mean expression level of *Nanog* mRNA, we cultured NMP cells in culture media containing several concentrations of 2i inhibitors and PD0325901, an FGF/Erk signal inhibitor (one of the 2i inhibitors) that increases *Nanog* expression[10] (Figure. 5c and Supplementary Table S1). As expected, intrinsic noise monotonously increased as the mean *Nanog* mRNA counts decreased (Figure 5c), suggesting that intrinsic noise significantly affects *Nanog* expression variability. We investigated the effects of histone acetylation and DNA methylation on intrinsic noise by treatments with Trichostatin A (TSA, a histone deacetylase inhibitor)[34] and 5-azacytidine (5-AzaC, a DNA-demethylating agent)[35] in 2i conditions (Fig. 5c), as these modifications influence gene activity. However, intrinsic noises in cells cultured with TSA and 5-AzaC were not strongly deviated from a trend line obtained from data of cells cultured in media containing 2i and PD0325901 (Fig. 5c), suggesting that histone acetylation and DNA methylation might not affect *Nanog* intrinsic noise at least in 2i conditions.

Given the above results, we evaluated whether the mRNA expression level in individual cells correlates with Nanog protein levels. To address this question, we performed smFISH using *Nanog* exonic probes followed by immunofluorescence using Nanog antibody (Figure 6), and found that *Nanog* mRNA and Nanog protein levels were well correlated in both serum and 2i conditions ($r = 0.85$ and $r = 0.72$ for serum and 2i conditions, respectively) (Figure 6). This finding suggests that Nanog protein heterogeneity originates from *Nanog* mRNA heterogeneity.

Discussion

We have demonstrated that the stochastic promoter activation significantly affects expression variability of Nanog in mESCs. Nanog expression variability is observed not only in mESCs cultured *in vitro* but also in pre-implantation embryonic inner cell mass (ICM) cells[36]. The biological meaning of heterogeneity is not fully understood, but some researches suggest that it may play a functional role in cell fate decisions[3,4,37]. It has been reported that the genome-wide epigenetic status of mESCs cultured in serum and 2i conditions resemble that of later ICM cells before differentiation and that of early ICM cells, respectively[38]. In addition, a recent report suggests that expression fluctuations observed in several genes including *Nanog* in both early- and late-stage ICM cells underlie lineage choices[39].

Our data indicated substantial expression variability of *Nanog* mRNA not only in serum but also in 2i conditions (Figure 2c and 5b). The Fano factor is the ratio between the variance and the mean of the mRNA copy number distribution, and is a key parameter used to quantify the deviation from Poisson statistics. For a complete Poisson distribution, the Fano factor equals 1. Distribution of *Nanog* mRNA transcripts in NMP cells were non-Poissonian (Fano factor $\gg 1$), suggesting that heterogeneity in *Nanog* expression is relatively high in both conditions. This observation is consistent with a previous report by Abranches et al. on the expression dynamics of the Nanog protein; using a bacterial artificial chromosome transgenic reporter and smFISH analysis, significant variability was observed in Nanog expression in mESCs cultured in 2i conditions[37]. Furthermore, we found that intrinsic noise derived from stochastic promoter activation significantly affects expression variability. Recently, Singer et al. reported that expression levels of genes, including *Nanog*, fluctuate between cells due not only to stochastic gene expression, but also to transitions between states with different gene activation potential[28]. This is consistent with our observation of multimodal distribution of transcription frequency over 4 h (Supplementary Figure S2). To compare the *Nanog* transcription dynamics among cells belonging to different states, we divided the cells into two groups according to their transcription frequency (lower or higher; Supplementary Figure S6). However, the distributions of ON and OFF time of *Nanog* transcription between the groups did not show statistically significant differences in any of the culture conditions (Supplementary Figure S6). Singer et al. suggested that metastable states are correlated with DNA methylation in mESCs[28]. However, inhibition of DNA methylation by DNA methyltransferase inhibitor 5-AzaC (our study) or inactivation of all three DNA methyltransferases[10,28] scarcely affects *Nanog* expression. Further analysis is necessary to understand what mechanisms determine the various *Nanog* expression states.

In our system, insertion of MS2 repeats into the *Nanog* locus slightly destabilizes *Nanog-MS2* mRNA (Supplementary Figure S1). In NMP cells, in which MS2 and PP7 repeats were integrated into the *Nanog* locus, *Nanog-MS2* and *Nanog-PP7* mRNA showed similar expression levels (Supplementary Figure S5), suggesting that

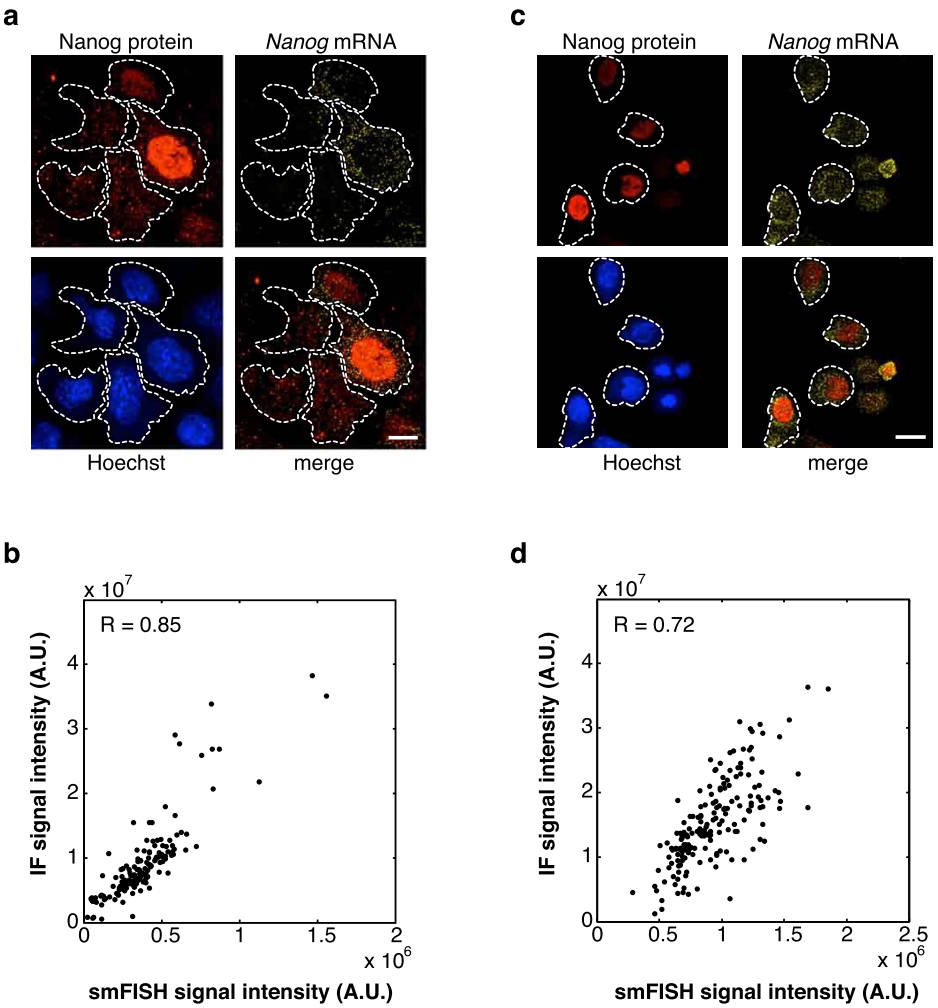

Figure 6 | Nanog protein heterogeneity originates from *Nanog* mRNA heterogeneity. C57BL/6 mESCs cultured in serum (a, b) and 2i conditions (c, d) were subjected to smFISH using *Nanog* exonic probes and followed by immunofluorescence (IF) using Nanog antibody. (a, c) Nuclei were counterstained with Hoechst33342 (Hoechst). Dashed lines represent edges of cell membranes of nondividing cells. Scale bar, 20 μm. (b, d) Scatter plot of fluorescence intensities of IF (Nanog protein) and smFISH (*Nanog* exonic probes).

integration of PP7 also affects *Nanog* mRNA degradation rate. Therefore, WT *Nanog* mRNA stays intact longer than *Nanog-MS2* (or *-PP7*) mRNA in cells. This might mask the cell-to-cell expression variability in *Nanog* mRNA in WT mESCs; in other words, it is possible that intrinsic noise of *Nanog* mRNA in WT mESCs is lower than that in NMP cells. However, it has been recently reported that considerable intrinsic noise in *Nanog* mRNA expression seemed to exist in hybrid mESCs without integration of reporter genes[11]. Therefore, it is possible that intrinsic noise has a considerable effect on Nanog expression variability in WT mESCs as well as in NMP cells.

Although pulsatile transcriptional events or "transcriptional bursting" has been reported in several model systems, the underlying molecular mechanisms are still elusive[15]. One of the models of transcriptional bursting is the chromatin-based model[13,15,40]. In this model, the efficiency of transcription depends on the absence of nucleosomes, which compete with the binding of transcription factors immediately upstream of the transcription start site (TSS). Therefore, the promoter activation timescale depends on relatively slower nucleosome turnover[41]. Recently, it has been reported that nucleosome occupancy immediately upstream of the *Nanog* TSS is inversely correlated with Nanog expression level[42], implying that the stochastic promoter activation of *Nanog* may originate from the relatively slow nucleosome dynamics.

Another potential source of transcriptional bursting includes DNA conformation changes involving efficient transcription. Some genes are regulated via long-range interaction between promoters and enhancers; because such long-range interactions seem to be variable among cells[43], the regulatory mechanism could be a source of expression variability[44]. It has been reported that the *Nanog* promoter region is associated with several regions genome-wide[45], suggesting that genome-wide stochastic association between *Nanog* promoters and enhancers may underlie transcriptional bursting. Further investigation is needed to understand the molecular mechanism of the *Nanog* promoter activation.

In summary, *Nanog* transcription dynamics were quantitated using the MS2 system in mESCs. We found that the promoter activation occurs in a pulsatile and stochastic manner. Furthermore, allele-specific smFISH analysis revealed that intrinsic noise considerably contributes to the Nanog heterogeneity. Therefore, we conclude that stochastic processes of promoter activation might be a key source of the intrinsic noise, and hence significantly affect Nanog expression variability. The combination of the MS2 system and smFISH analysis seems to be useful for evaluating stochastic promoter activation and expression variability at single-cell resolution. The techniques used in the present report will help further the understanding of the molecular basis of allelic expression in mESCs[46] and their heterogeneity.

Methods

Cell culture. Bruce 4 C57BL/6 mESCs (Merk Millipore, Billerica, MA) (and later derivatives) were cultured in 2i conditions (StemSure D-MEM [Wako Pure Chemicals, Osaka, Japan], 15% fetal bovine serum, 0.1 mM β-mercaptoethanol, 1 × MEM nonessential amino acids [Wako Pure Chemicals], 2 mM L-alanyl-L-glutamine solution [Wako Pure Chemicals], 1000 U/mL LIF [Wako Pure Chemicals], 20 μg/mL gentamicin [Wako Pure Chemicals], 3 μM CHIR99021, and 1 μM PD0325901) on a 0.1% gelatin-coated dish. Before each experiment, cells were passaged two times and cultured in 2i conditions as well as serum conditions (StemSure D-MEM, 15% fetal bovine serum, 0.1 mM β-mercaptoethanol, 1 × MEM nonessential amino acids, 2 mM L-alanyl-L-glutamine solution, 1000 U/mL LIF, 20 μg/mL gentamicin), or in serum conditions containing 2i inhibitors or PD0325901 at several concentrations, as described in Supplementary Table S1. TSA and 5-AzaC were added to the cells at a final concentration of 50 nM and 50 μM, respectively. They were applied for 72 hours.

Plasmid construction. TALEN expression vectors were constructed as described in Ochiai et al[20]. A part of the vector was derived from pTALEN_v2 (NG) vector (plasmid 32190, Addgene, Cambridge, MA)[47]. The TALEN target sequences and their amino acid sequences are described in Figure 1b and Supplementary Figure S7, respectively. Targeting vectors containing either 2A-loxP-hsvTK-2A-Puro-loxP-24×MS2 (pTV-mNanog-PMS) or 2A-loxP-hsvTK-2A-Hyg-loxP-24×PP7 (pTV-mNanog-HPP) were constructed by polymerase chain reaction (PCR) and standard cloning techniques. The 24× MS2 and 24× PP7 sites in the targeting vectors were derived from pCR4-24XMS2SL-stable (Addgene plasmid 31865) and the pCR4-24XPP7SL (Addgene plasmid 31864), respectively[14,48]. The hsvTK in the targeting vector was derived from the pLOX-TERT-iresTK vector (Addgene plasmid 12245)[49]. The nucleotide sequences of the targeting vectors are provided in Supplementary Figure S8 and S9. To construct pPB-LR5-CAG-MCP-mNeonGreen-IRES-Neo, the CAG-MCP-mNeonGreen-IRES-Neo cassette was inserted into the NheI/XhoI site of pPB-LR5[50]. The cDNA of MCP in pPB-LR5-CAG-MCP-mNeonGreen-IRES-Neo was derived from the phage-ubc-nls-ha-tdMCP-gfp vector (Addgene plasmid 40649)[51].

Gene targeting. To generate *Nanog*$^{MS2/MS2}$ cells, C57BL/6 mESCs (0.5×10^5) were plated in 24-well plates the day before transfection. The next day, the cells were transfected with 1 μg of pTV-mNanog-PMS and 250 ng each of TALEN expression vectors using Lipofectamine 2000 (Life Technologies, Gaithersburg, MD) according to the manufacturer's instructions. After 24 h, the cells were transferred to 10-cm plates, incubated for 72 h, and then subjected to puromycin selection (1 μg/mL). Homologous recombination was verified by PCR and Southern blotting. Then, to excise the selection cassette flanked by loxP sites, 500 ng of Cre expression vector (pCAG-Cre) (Addgene plasmid 13775)[52] was transfected into the obtained clone (*Nanog*$^{TP\text{-}MS2/TP\text{-}MS2}$). The genotype of the resultant ganciclovir-resistant ESC (*Nanog*$^{MS2/MS2}$, NM clone) was confirmed by Southern blotting. For constitutive expression of NLS-MCP-mNeonGreen, pPB-LR5-CAG-MCP-mNeonGreen-IRES-Neo plasmid was transfected with pCMV-hyPBase[53] into NM cells, and G418-resistant clones were obtained. To generate *Nanog*$^{PP7/MS2}$ (NMP cells), 1 μg of pTV-mNanog-HPP was transfected with 250 ng of TALEN expression vectors into C57BL/6 ESCs, as described above. The resultant hygromycin-resistant clone (*Nanog*$^{TH\text{-}PP7/+}$) was subsequently subjected to a second targeting with pTV-mNanog-PMS and TALEN expression vectors. The resultant puromycin-resistant clone (*Nanog*$^{TH\text{-}PP7/+}$) was subsequently subjected to a second targeting with pTV-mNanog-PMS and TALEN expression vectors. The genotype of the resultant puromycin-resistant mESC (*Nanog*$^{TH\text{-}PP7/TP\text{-}MS2}$) was confirmed by Southern blot. Then, to excise the selection cassette flanked by loxP sites, pCAG-Cre was transfected into the obtained clone (*Nanog*$^{TH\text{-}PP7/TP\text{-}MS2}$). The genotype of the resultant ganciclovir-resistant NMP cell was confirmed by Southern blot. Southern blots were performed as described previously[20].

Immunostaining. Immunostaining was performed on fixed cells (4% paraformaldehyde in BBS [50 mM BES, 280 mM NaCl, 1.5 mM $Na_2HPO_4{\cdot}2H_2O$] with 1 mM $CaCl_2$, for 15 min), washed, and blocked for 30 min in BBT-BSA buffer (BBS with 0.5% BSA, 0.1% Triton, and 1 mM $CaCl_2$). Cells with primary antibodies were incubated overnight at 4°C at the following dilutions: anti-Nanog (1 : 500; MLC-51, eBioscience, San Diego, CA), anti-Oct4 (1 : 500; ab19857, Abcam, Cambridge, MA), anti-SSEA-1 (1 : 250, ab16285, Abcam). Cells were washed and blocked in BBT-BSA and then incubated with Hoechst33342 (1 : 1000, Life Technologies) and Alexa-conjugated secondary antibodies (1 : 500, Life Technologies). Images were acquired using an Olympus IX83 microscope (Olympus, Tokyo, Japan) with a CSU-W1 confocal unit (Yokogawa, Tokyo, Japan).

smFISH. Trypsinized cells were transferred onto Laminin-511 (BioLamina, Stockholm, Sweden)-coated round coverslips and cultured for 1 h at 37°C and 5% CO_2. Cells were washed with phosphate-buffered saline (PBS), fixed with 4% paraformaldehyde in PBS for 10 min, and washed with PBS two times. Then, cells were permeabilized in 70% ethanol at 4°C overnight. Following a wash with 10% formamide dissolved in 2× SSC, the cells were hybridized to probe sets in 60 μL of hybridization buffer containing 2× SSC, 10% dextran sulfate, 10% formamide, and each probe set. Hybridization was performed for 4 h at 37°C in a moist chamber. The coverslips were washed with 10% formamide in 2× SSC solution, and then with 10% formamide in 2× SSC solution with Hoechst33342 (1 : 1000). Hybridized cells were mounted in catalase/glucose oxidase containing mounting media, as described previously[54]. Images were acquired using an Olympus IX83 microscope with a CSU-W1 confocal unit, a 100× Olympus oil immersion objective of 1.40 NA, and an iXon3 EMCCD camera (Andor, Belfast, UK), with laser illumination at 405 nm, 561 nm, and 637 nm, and were analyzed using Metamorph software (Universal Imaging Corporation, West Chester, PA); 101 z planes per site spanning 15 μm were acquired. Images were filtered with a one-pixel diameter 3D median filter and subjected to background subtraction via a rolling ball radius of 5 pixels, using ImageJ software. Detection and counting of smFISH signals were performed using Imaris software (Bitmap, Zurich, Switzerland) as described by Yunger et al[55]. Mixtures of *Nanog* exonic and intronic probes conjugated with CAL Fluor Red 590 were obtained from BioSearch Inc (Novato, CA) and used at 0.25 μM. MS2 probes conjugated with Cy5, and PP7 probes conjugated with Cy3 were obtained from Operon Technologies Inc. (Alameda, CA) and used at 0.52 μM each. Probe sequences are shown in Supplementary Table S2. To estimate mRNA half-life, NM and C57BL6 mESCs were cultured in 2i medium containing actinomycin D (5 μg/ml) for 0, 2, 4, and 6 hours, and subjected to smFISH analysis using *Nanog* probes.

smFISH-immunostaining. Trypsinized cells were transferred onto μ-Slide 8-well (Ibidi, Martinsried, Germany) coated with Laminin-511 and cultured for 1 h at 37°C and 5% CO_2. Then, cells were fixed and subjected to smFISH as described above using *Nanog* exonic probes. After the image acquisition, cells were subjected to immunostaining, as described in the immunostaining section above, using the anti-Nanog antibody and Alexa Fluor 647 goat anti-rat IgG (H+L) (1 : 500, Life Technologies). The images were acquired using an Olympus IX83 microscope with a CSU-W1 confocal unit, 60× Olympus oil-immersion objective of 1.42 NA; 101 z planes per site, spanning 15 μm, were analyzed. smFISH images were filtered using ImageJ, with a one-pixel diameter 3D median filter, and subjected to background subtraction via a rolling ball radius of 2 pixels. Further fluorescence subtraction (100, 16-bit pixel unit) for almost complete removal of background intensity and maximum-intensity projection was subsequently performed. Immunostained images were filtered with a two-pixel diameter 3D median filter, and subjected to fluorescence subtraction (300, 16-bit pixel unit) for almost complete removal of background intensity and maximum-intensity projection, using ImageJ. The freehand selection tool of ImageJ was used for measurements of integrated signal intensity in each cell.

Live imaging. Trypsinized cells were transferred onto μ-Slide 8-well coated with Laminin-511 and cultured overnight at 37°C and 5% CO_2. The next day, after the medium was changed, cells were subjected to live imaging by an Olympus IX83 microscope with a CSU-W1 confocal unit and a 60× Olympus oil-immersion objective of 1.42 NA for quantification of transcriptional dynamics. Fluorescence images were captured using an iXon3 EMCCD camera, equipped with a 488-nm laser, a stage-top microscope incubator (5% CO_2 at 37°C; Tokai Hit, Shizuoka, Japan), and an ASI MS-2000 piezo stage (ASI, Lyon, France), and were analyzed using Metamorph software; 41 z planes per site, spanning 12 μm, were acquired with a 2-min interval time for 4 h. Acquired images were filtered with a one-pixel diameter 3D median filter using ImageJ. Detection of fluorescent spots was performed using the "Spot" function in Imaris with a spot diameter set at 0.75 μm (semi-automatic detection). For observation of individual mRNA spots, a 100× Olympus oil-immersion objective of 1.40 NA was used. See Supplementary Video S3 and Figure S3 for details.

1. Mitsui, K. *et al.* The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**, 631–642 (2003).
2. Chambers, I. *et al.* Nanog safeguards pluripotency and mediates germline development. *Nature* **450**, 1230–1234 (2007).
3. Kalmar, T. *et al.* Regulated Fluctuations in Nanog Expression Mediate Cell Fate Decisions in Embryonic Stem Cells. *PLoS Biol* **7**, e1000149 (2009).
4. Torres-Padilla, M. E. & Chambers, I. Transcription factor heterogeneity in pluripotent stem cells: a stochastic advantage. *Development* **141**, 2173–2181 (2014).
5. Nichols, J., Silva, J., Roode, M. & Smith, A. Suppression of Erk signalling promotes ground state pluripotency in the mouse embryo. *Development* **136**, 3215–3222 (2009).
6. Wray, J., Kalkan, T. & Smith, A. G. The ground state of pluripotency. *Biochem. Soc. Trans.* **38**, 1027–1032 (2010).
7. Herberg, M., Kalkan, T., Glauche, I., Smith, A. & Roeder, I. A Model-Based Analysis of Culture-Dependent Phenotypes of mESCs. *PLoS ONE* **9**, e92496 (2014).
8. Festuccia, N. *et al.* Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells. *Cell Stem Cell* **11**, 477–490 (2012).
9. Navarro, P. *et al.* OCT4/SOX2-independent Nanog autorepression modulates heterogeneous Nanog gene expression in mouse ES cells. *EMBO J* **31**, 4547–4562 (2012).
10. Miyanari, Y. & Torres-Padilla, M.-E. Control of ground-state pluripotency by allelic regulation of Nanog. *Nature* **483**, 470–473 (2012).
11. Hansen, C. H. & van Oudenaarden, A. Allele-specific detection of single mRNA molecules in situ. *Nat Methods* **10**, 869–871 (2013).
12. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science* **297**, 1183–1186 (2002).

13. Raser, J. M. & O'Shea, E. K. Control of stochasticity in eukaryotic gene expression. *Science* **304**, 1811–1814 (2004).
14. Bertrand, E. *et al.* Localization of ASH1 mRNA particles in living yeast. *Mol Cell* **2**, 437–445 (1998).
15. Lionnet, T. & Singer, R. H. Transcription goes digital. *EMBO Rep* **13**, 313–321 (2012).
16. Janicki, S. M. *et al.* From silencing to gene expression: real-time analysis in single cells. *Cell* **116**, 683–698 (2004).
17. Chubb, J. R., Trcek, T., Shenoy, S. M. & Singer, R. H. Transcriptional pulsing of a developmental gene. *Curr Biol* **16**, 1018–1025 (2006).
18. Lionnet, T. *et al.* A transgenic mouse for in vivo detection of endogenous labeled mRNA. *Nat Methods* **8**, 165–170 (2011).
19. Sakuma, T. *et al.* Repeating pattern of non-RVD variations in DNA-binding modules enhances TALEN activity. *Sci Rep* **3**, 3379 (2013).
20. Ochiai, H. *et al.* TALEN-mediated single-base-pair editing identification of an intergenic mutation upstream of BUB1B as causative of PCS (MVA) syndrome. *Proc Natl Acad Sci USA* **111**, 1461–1466 (2014).
21. Raj, A., van den Bogaard, P., Rifkin, S. A., van Oudenaarden, A. & Tyagi, S. Imaging individual mRNA molecules using multiple singly labeled probes. *Nat Methods* **5**, 877–879 (2008).
22. Ambrosetti, D. C., Basilico, C. & Dailey, L. 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 (1997).
23. Kunath, T. *et al.* FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development* **134**, 2895–2902 (2007).
24. Herberg, M., Kalkan, T., Glauche, I., Smith, A. & Roeder, I. A model-based analysis of culture-dependent phenotypes of mESCs. *PLoS ONE* **9**, e92496 (2014).
25. Jonkers, I., Kwak, H. & Lis, J. T. Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *eLife* **3**, e02407 (2014).
26. Muramoto, T. *et al.* Live imaging of nascent RNA dynamics reveals distinct types of transcriptional pulse regulation. *Proc Natl Acad Sci USA* **109**, 7350–7355 (2012).
27. Larson, D. R. *et al.* Direct observation of frequency modulated transcription in single cells using light activation. *eLife* **2**, e00750–e00750 (2013).
28. Singer, Z. S. *et al.* Dynamic Heterogeneity and DNA Methylation in Embryonic Stem Cells. *Mol Cell* **55**, 319–331 (2014).
29. Peccoud, J. & Ycart, B. Markovian modeling of gene-product synthesis. *Theoretical population biology* **48**, 222–234 (1995).
30. Raj, A., Peskin, C. S., Tranchina, D., Vargas, D. Y. & Tyagi, S. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol* **4**, e309 (2006).
31. Larson, D. R., Singer, R. H. & Zenklusen, D. A single molecule view of gene expression. *Trends Cell Biol.* **19**, 630–637 (2009).
32. Hocine, S., Raymond, P., Zenklusen, D., Chao, J. A. & Singer, R. H. Single-molecule analysis of gene expression using two-color RNA labeling in live yeast. *Nat Methods* **10**, 119–121 (2012).
33. Swain, P. S., Elowitz, M. B. & Siggia, E. D. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc Natl Acad Sci USA* **99**, 12795–12800 (2002).
34. Yoshida, M., Horinouchi, S. & Beppu, T. Trichostatin A and trapoxin: novel chemical probes for the role of histone acetylation in chromatin structure and function. *Bioessays* **17**, 423–430 (1995).
35. Veselý, J. Mode of action and effects of 5-azacytidine and of its derivatives in eukaryotic cells. *Pharmacol. Ther.* **28**, 227–235 (1985).
36. Dietrich, J.-E. & Hiiragi, T. Stochastic patterning in the mouse pre-implantation embryo. *Development* **134**, 4219–4231 (2007).
37. Abranches, E. *et al.* Stochastic NANOG fluctuations allow mouse embryonic stem cells to explore pluripotency. *Development* **141**, 2770–2779 (2014).
38. Ficz, G. *et al.* FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. *Cell Stem Cell* **13**, 351–359 (2013).
39. Ohnishi, Y. *et al.* Cell-to-cell expression variability followed by signal reinforcement progressively segregates early mouse lineages. *Nat Cell Biol* **15**, 1–13 (2013).
40. Brown, C. R., Mao, C., Falkovskaia, E., Jurica, M. S. & Boeger, H. Linking Stochastic Fluctuations in Chromatin Structure and Gene Expression. *PLoS Biol* **11**, e1001621 (2013).
41. van Royen, M. E., Zotter, A., Ibrahim, S. M., Geverts, B. & Houtsmuller, A. B. Nuclear proteins: finding and binding target sites in chromatin. *Chromosome Res.* **19**, 83–98 (2011).
42. Wu, C. Y., Feng, X. & Wei, L. N. Coordinated repressive chromatin-remodeling of Oct4 and Nanog genes in RA-induced differentiation of embryonic stem cells involves RIP140. *Nucleic Acids Res* **42**, 4306–4317 (2014).
43. Nagano, T. *et al.* Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* **502**, 59–64 (2014).
44. Noordermeer, D. *et al.* Variegated gene expression caused by cell-specific long-range DNA interactions. *Nat Cell Biol* **13**, 944–951 (2011).
45. Apostolou, E. *et al.* Genome-wide chromatin interactions of the Nanog locus in pluripotency, differentiation, and reprogramming. *Cell Stem Cell* **12**, 699–712 (2013).
46. Eckersley-Maslin, M. A. *et al.* Random Monoallelic Gene Expression Increases upon Embryonic Stem Cell Differentiation. *Dev Cell* **28**, 351–365 (2014).
47. Sanjana, N. E. *et al.* A transcription activator-like effector toolbox for genome engineering. *Nat Protoc* **7**, 171–192 (2012).
48. Larson, D. R., Zenklusen, D., Wu, B., Chao, J. A. & Singer, R. H. Real-time observation of transcription initiation and elongation on an endogenous yeast gene. *Science* **332**, 475–478 (2011).
49. Salmon, P. *et al.* Reversible immortalization of human primary cells by lentivector-mediated transfer of specific genes. *Mol Ther* **2**, 404–414 (2000).
50. Yusa, K., Rad, R., Takeda, J. & Bradley, A. Generation of transgene-free induced pluripotent mouse stem cells by the piggyBac transposon. *Nat Methods* **6**, 363–369 (2009).
51. Wu, B., Chao, J. A. & Singer, R. H. Fluorescence fluctuation spectroscopy enables quantitative imaging of single mRNAs in living cells. *Biophysical Journal* **102**, 2936–2944 (2012).
52. Matsuda, T. & Cepko, C. L. Controlled expression of transgenes introduced by in vivo electroporation. *Proc Natl Acad Sci USA* **104**, 1027–1032 (2007).
53. Yusa, K., Zhou, L., Li, M. A., Bradley, A. & Craig, N. L. A hyperactive piggyBac transposase for mammalian applications. *Proc Natl Acad Sci USA* **108**, 1531–1536 (2011).
54. Lyubimova, A. *et al.* Single-molecule mRNA detection and counting in mammalian tissue. *Nat Protoc* **8**, 1743–1758 (2013).
55. Yunger, S., Rosenfeld, L., Garini, Y. & Shav-Tal, Y. Quantifying the transcriptional output of single alleles in single living mammalian cells. *Nat Protoc* **8**, 393–408 (2013).

Acknowledgments

We thank Dr. Feng Zhang for providing the pTALEN_v2 (NG) vector (Addgene plasmid 32190); Dr. Robert Singer for providing the pCR4-24XMS2SL-stable vector (Addgene plasmid 31865), the pCR4-24XPP7SL (Addgene plasmid 31864), and the phage-ubc-nls-ha-tdMCP-gfp vector (Addgene plasmid 40649); Dr. Didier Trono for supplying the pLOX-TERT-iresTK vector (Addgene plasmid 12245); Dr. Connie Cepko for supplying the pCAG-Cre vector (Addgene plasmid 13775); Dr. Allan Bradley for supplying the pPB-LR5 vector; Dr. Nancy L. Craig for supplying the pCMV-hyPBase vector; Drs. Yusuke Miyanari, Ken-ichi Suzuki, Shinichi Tate, and Akinori Awazu for valuable discussions; and Ms. Noriko Sumiyoshi for technical assistance. This work was supported by Platform for Dynamic Approaches to Living System from the Ministry of Education, Culture, Sports, Science and Technology, Japan (to H.O., T. Sugawara and T.Y.), and Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology (to H.O.).

Author contributions

H.O. designed and performed the experiments; T.Sugawara, T.Sakuma and T.Y. contributed new reagents/analytic tools; and H.O. wrote the paper.

Additional information

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Ochiai, H., Sugawara, T., Sakuma, T. & Yamamoto, T. Stochastic promoter activation affects Nanog expression variability in mouse embryonic stem cells. *Sci. Rep.* **4**, 7125; DOI:10.1038/srep07125 (2014).