Title

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

Authors

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

Affiliations

[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

* Corresponding author

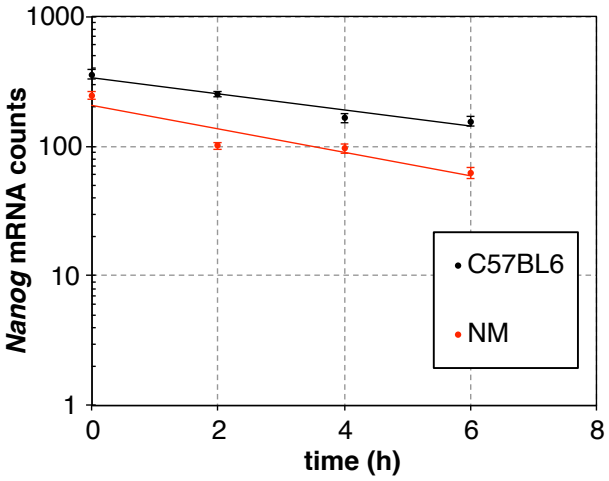
Tel.: +81-82-424-5529

Fax: +81-82-424-5529

Hiroshi Ochiai:

E-mail: ochiai@hiroshima-u.ac.jp

Supplementary Figures



***Nanog* mRNA half-life:**

C57BL6 mESCs : 4.72 (4.15 - 5.29) h

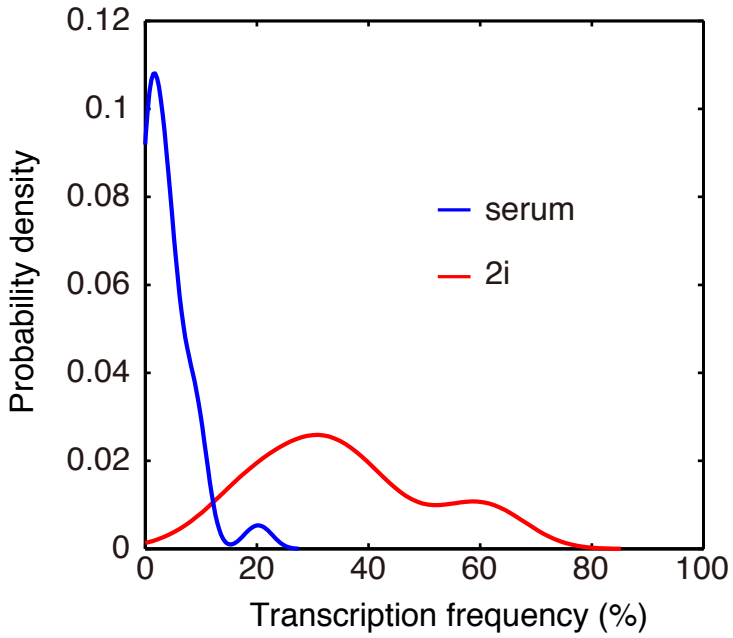
NM cells : 3.30 (3.01 - 3.59) h

***Nanog* mRNA degradation rate:**

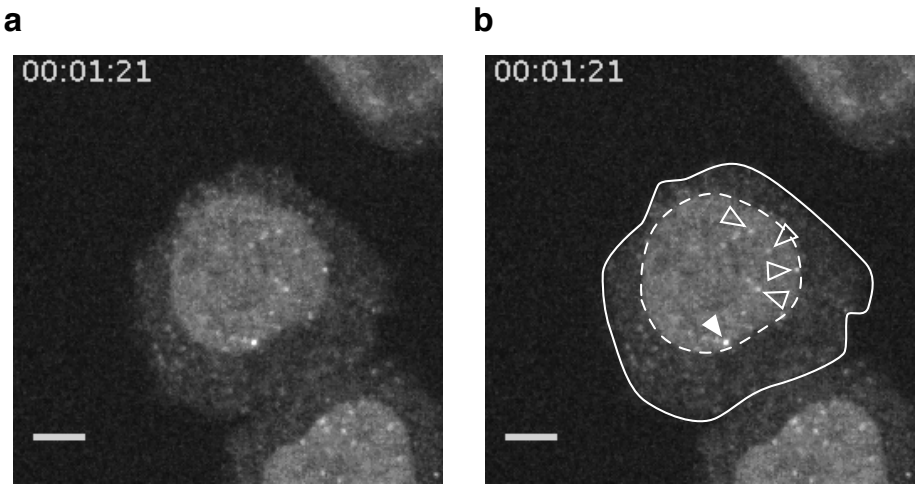
C57BL6 mESCs : 0.147 (0.13 - 0.164) / h

NM cells : 0.209 (0.191 - 0.227) / h

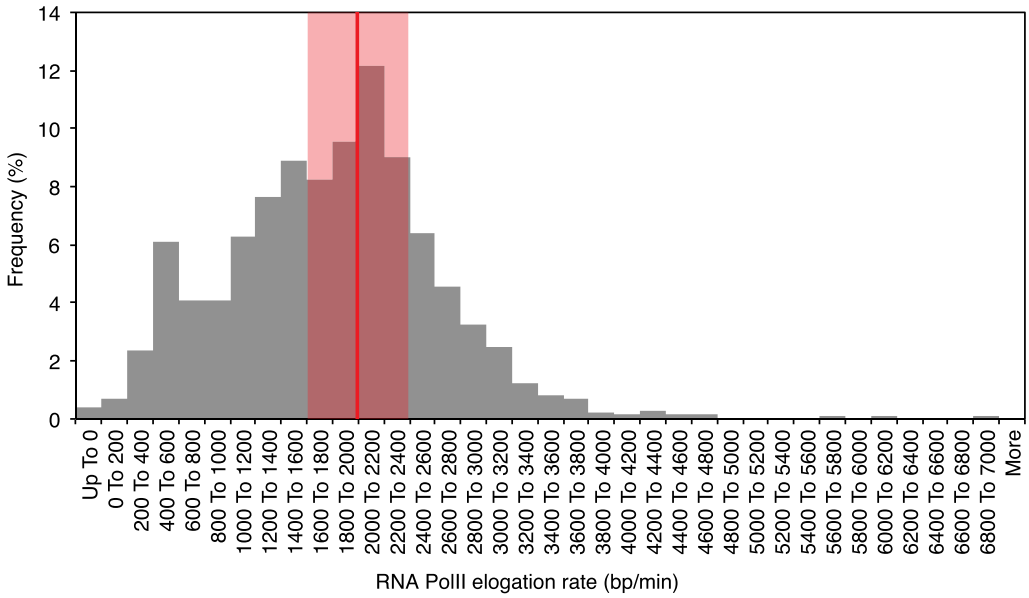
Supplementary Figure S1, Half-lives of WT and MS2-integrated *Nanog* mRNA. *Nanog* mRNAs in NM and wild-type (WT) C57BL6 mESCs were counted by smFISH using *Nanog* exonic probes over 6 hours after transcription inhibition with actinomycin D (averages and 95% confidence intervals are shown). Half-lives and degradation rates are shown at the right side of the Figure. n > 100 sample points. 95% confidence limits are in parentheses.



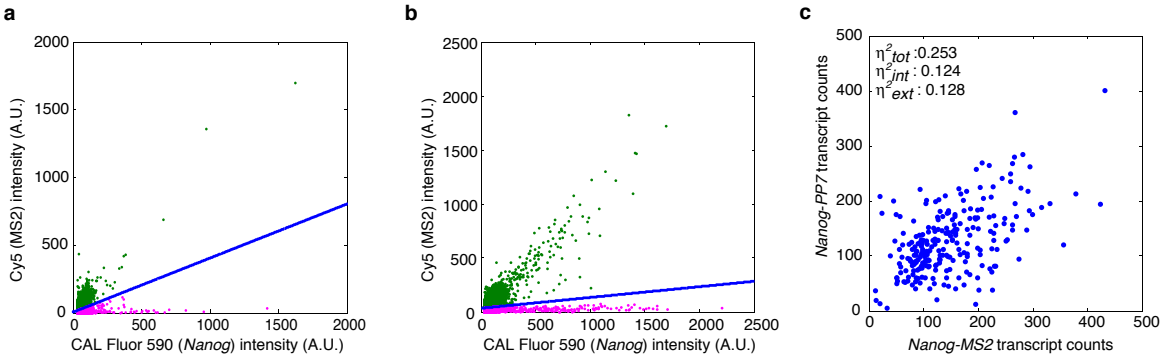
Supplementary Figure S2, Multidmodal distributions of transcription frequency. Probability density functions of transcription frequency over 4 hours in NM-G cells cultured in serum and 2i conditions (Figure 3c) are shown. Continuous probability density functions were generated using a kernel smoothed density estimate (ksdensity, Matlab).



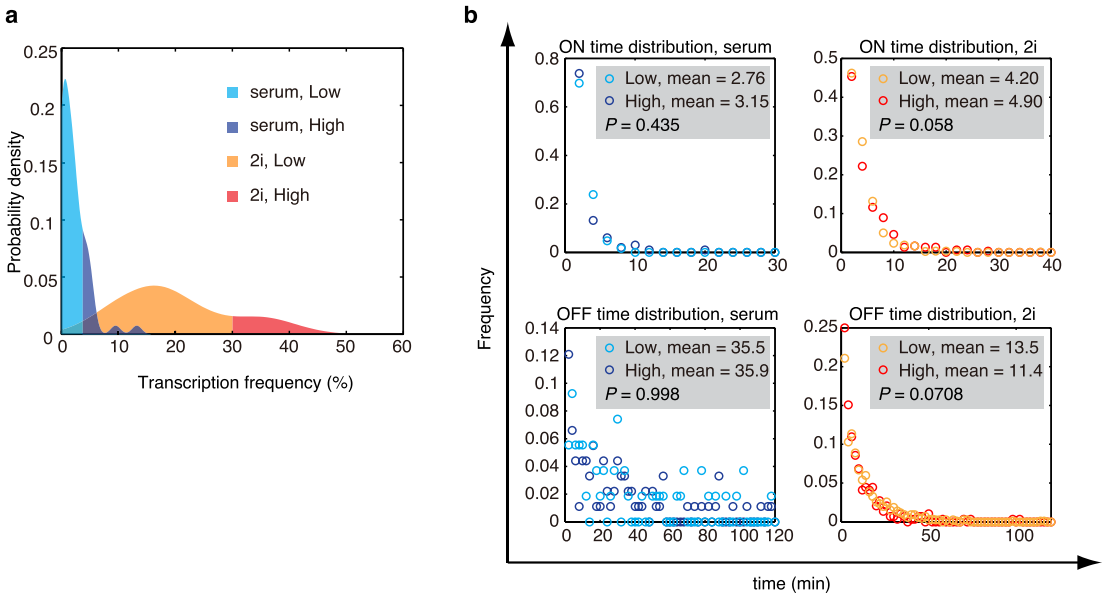
Supplementary Figure S3, Simultaneous visualization of individual mRNA molecules and transcription site. (a, b) The images are the same snapshot of live imaging of an NM-G cell with (b) or without (a) annotations (Supplementary Video S3). Solid and dashed lines indicate the edges of cell and nuclear membranes, respectively. Filled and unfilled arrowheads point to a transcription site and individual mRNA spots, respectively. Scale bar, 5 μm.



Supplementary Figure S4, Distribution of RNA Pol II elongation rate. Histogram shows distribution of RNA Pol II elongation rate in mESCs determined by Jonker et al. (Jonker et al., *eLife*, 2014). Red line and light red rectangle represent the mean (1.99 kb/min) and standard deviation (±0.39 kb/min) of RNA Pol II elongation rate estimated in this study.



Supplementary Figure S5, Allele-specific single-molecule fluorescence *in situ* hybridization (smFISH) in a reporter mESC line. (a, b) smFISH was performed in NMP cells cultured in serum (a) or 2i conditions (b) using *Nanog* exonic probes (CAL Fluor 590 conjugated) and MS2 probes (Cy5 conjugated). The fluorescence intensities of CAL Fluor 590 above the threshold and Cy5 of individual mRNA spots were plotted. Among the *Nanog* probe-positive spots, those with an MS2 probe signal intensity above a given threshold (blue line) were assumed to be mRNAs expressed from an MS2-integrated allele (green dots); otherwise, mRNAs were assumed to be expressed from PP7 alleles (magenta dots). The threshold lines were manually determined to equally separate all dots. (c) Scatter plot of *Nanog-MS2* and *-PP7* transcripts determined by smFISH using MS2 and PP7 probes.



Supplementary Figure S6, ON and OFF time for *Nanog* transcription is not significantly affected by differences in state. (a) Probability density functions of transcription frequency over 4 hours in NM-G cells cultured in serum and 2i conditions (Figure 3c) are shown (see also Supplementary Figure S2). The functions are divided into two groups with higher or lower transcription frequencies relative to threshold values (3.75% and 30% for serum and 2i conditions, respectively). (b) Distributions of ON and OFF time for *Nanog* transcription in cells with higher or lower transcription frequencies. Mean values are displayed in each graph. Differences between the two groups within the same culture conditions were not significant ($P > 0.05$, two-sample Kolmogorov–Smirnov test).

mNanog-TALEN-L

MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAAVDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAV
KYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASHDGGK
QALETVQRLLPVLCQDHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPAQVVAIA
SNGGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPD
QVVAIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLLPVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQD
HGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLL
PVLCQDHGLTPDQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNNGGKQALE
TVQRLLPVLCQAHGLTPAQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGRPALESIVAQLSRPDPALAALTNDHLVALACLG
GRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAGSQLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKH
LGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMERYVEENQTRDKHLNPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNHI
TNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINFRS*

mNanog-TALEN-R

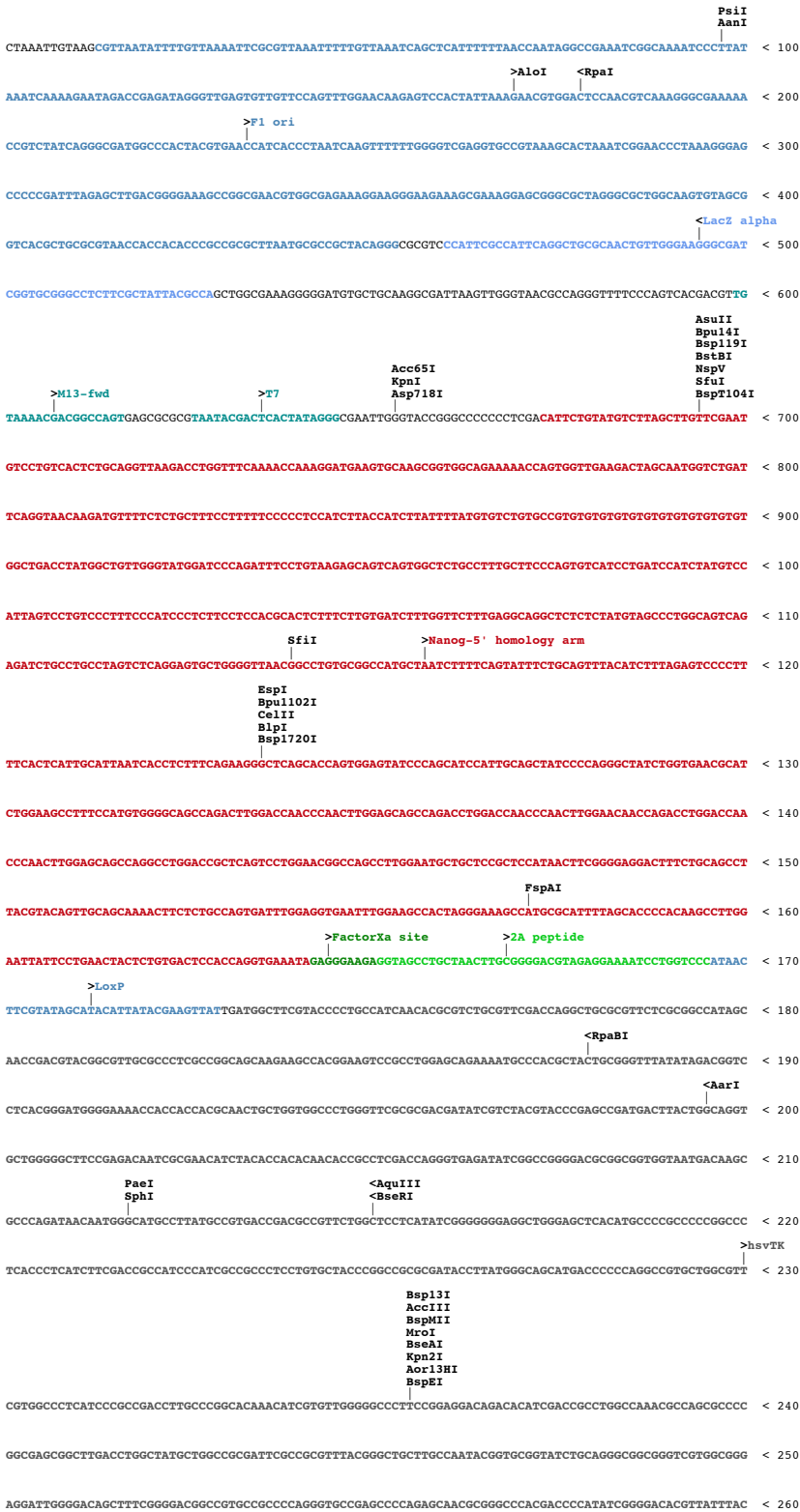
MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAAVDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAV
KYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLNLTPDQVVAIASNGGGK
QALETVQRLLPVLCQDHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVVAIA
SNIGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPD
QVVAIASHDGGKQALETVQRLLPVLCQAHGLTPAQVVAIASHDGGKQALETVQRLLPVLCQDHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQD
HGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPAQVVAIASNGGGKQALETVQRLL
PVLCQDHGLTPDQVVAIASNNGGKQALETVQRLLPVLCQDHGLTPEQVVAIASNGGKQALETVQRLLPVLCQAHGLTPDQVVAIASNGGKQALE
TVQRLLPVLCQAHGLTPAQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLG
GRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVAGSQLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKH
LGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVKENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFKGNYKAQLTRLNRK
TNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINFRS*

Supplementary Figure S6, Amino acid sequences of TALENs.

The amino acid sequences of TALENs (mNanog-TALEN-L and mNanog-TALEN-R) used in this study are shown. Repeat variable diresidues (RVDs) are shown in red.

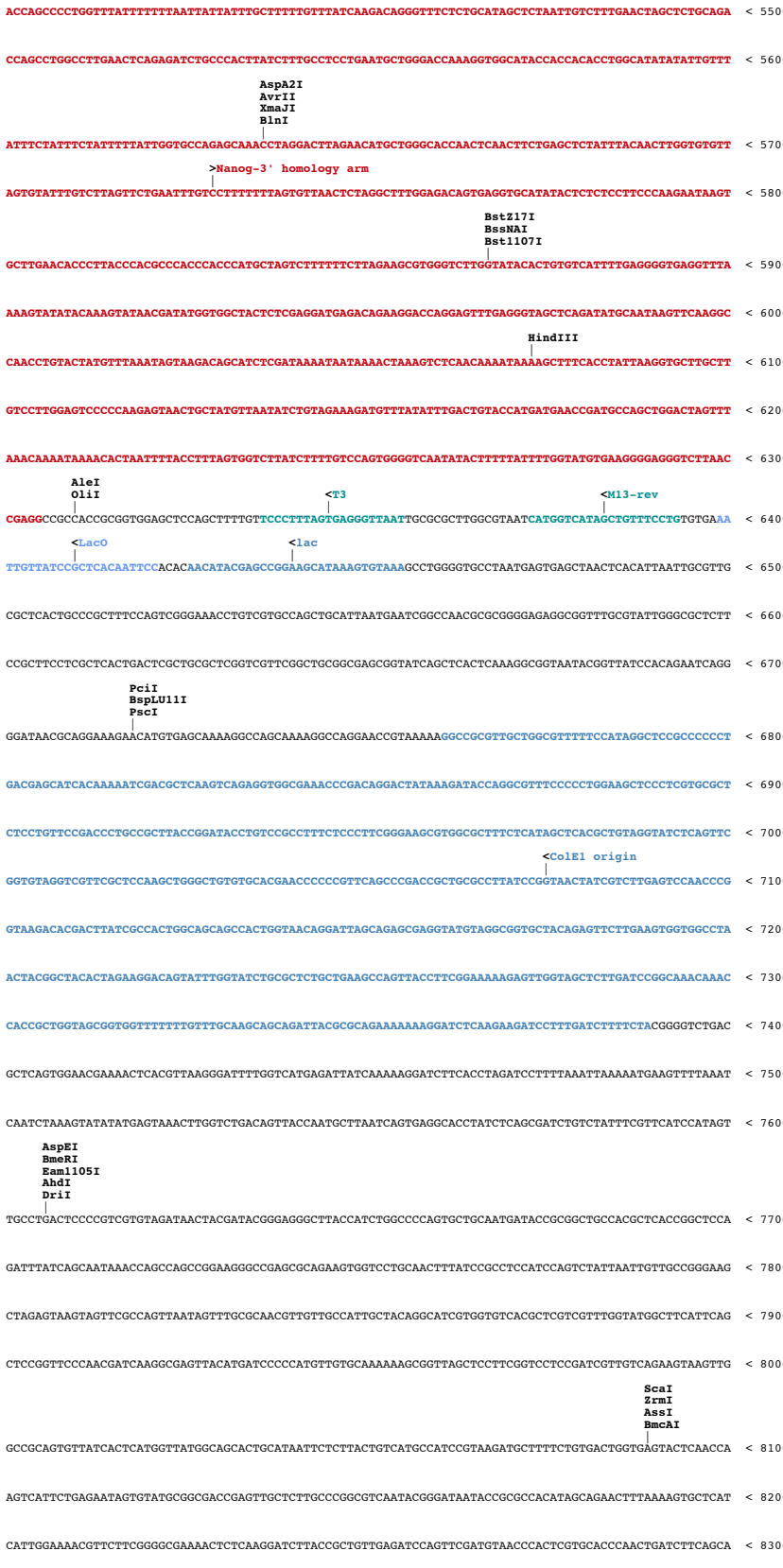
Supplementary Fig. S7, Nucleotide sequence of pTV-mNanog-PMS used in this study.

Restriction map of pTV-mNanog-PMS



Supplementary Fig. S7, continued

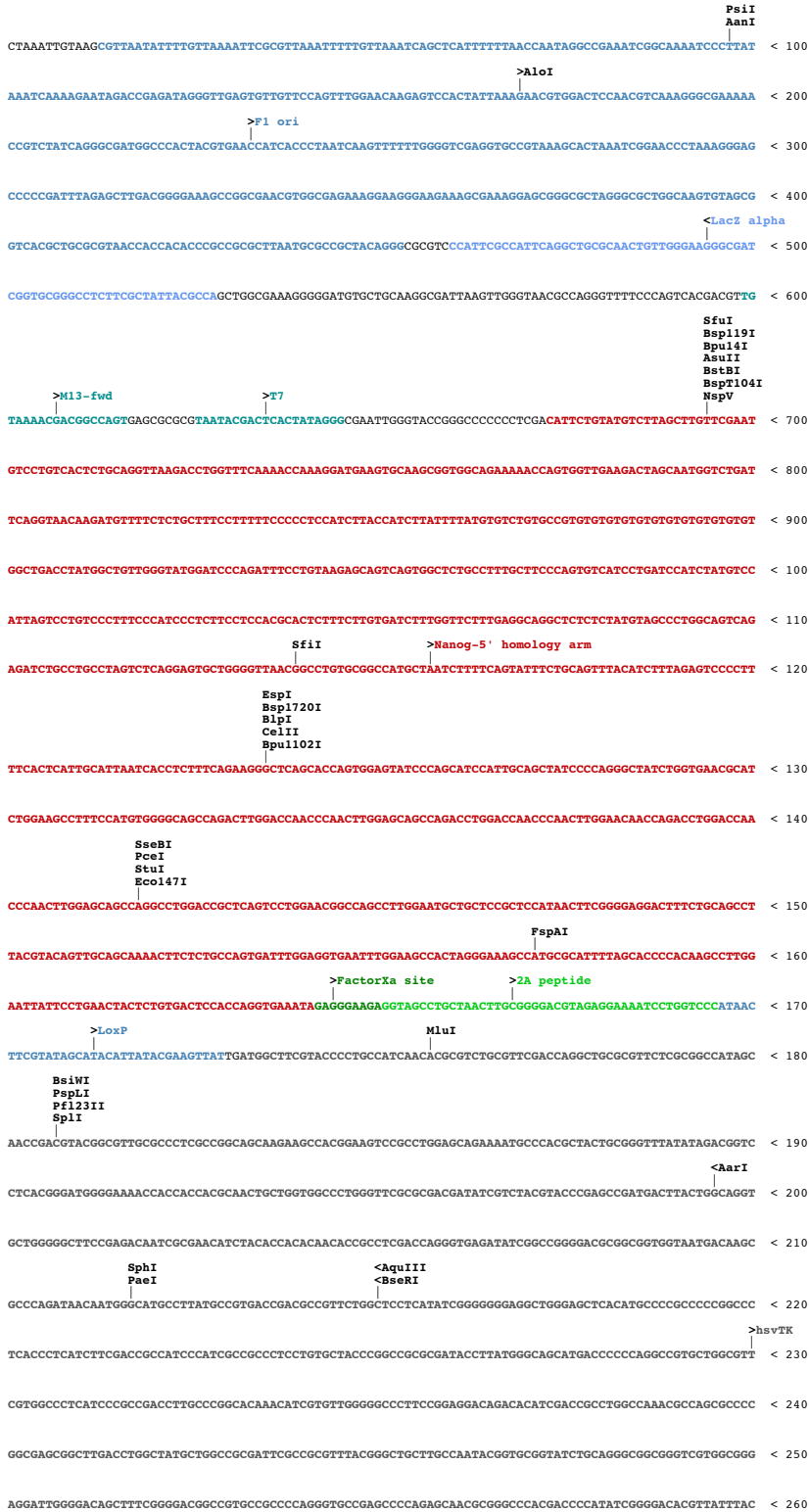
Supplementary Fig. S7, continued



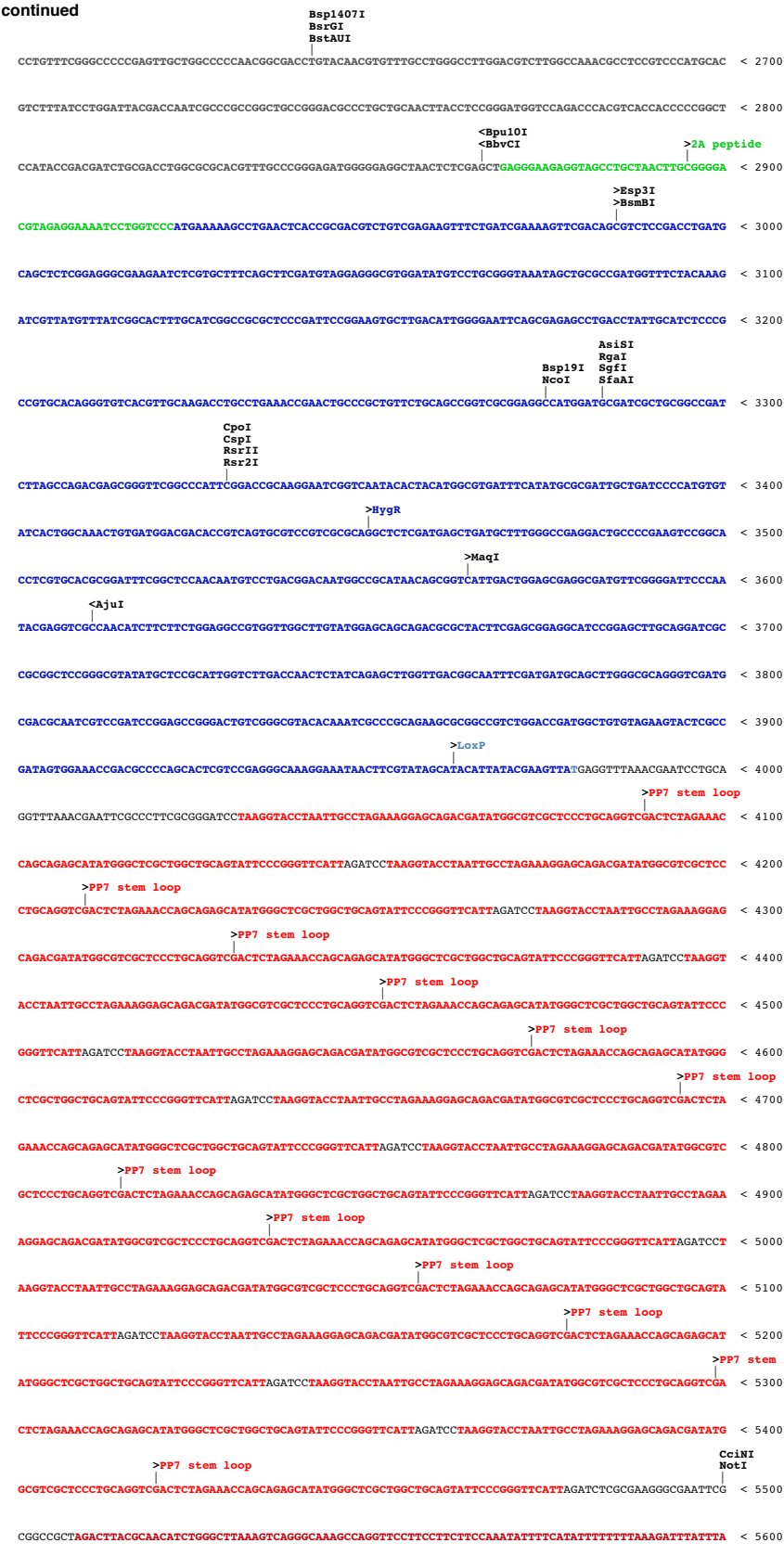
Supplementary Fig. S7, continued

Supplementary Fig. S8 Nucleotide sequence of pTV-mNanog-HPP used in this study.

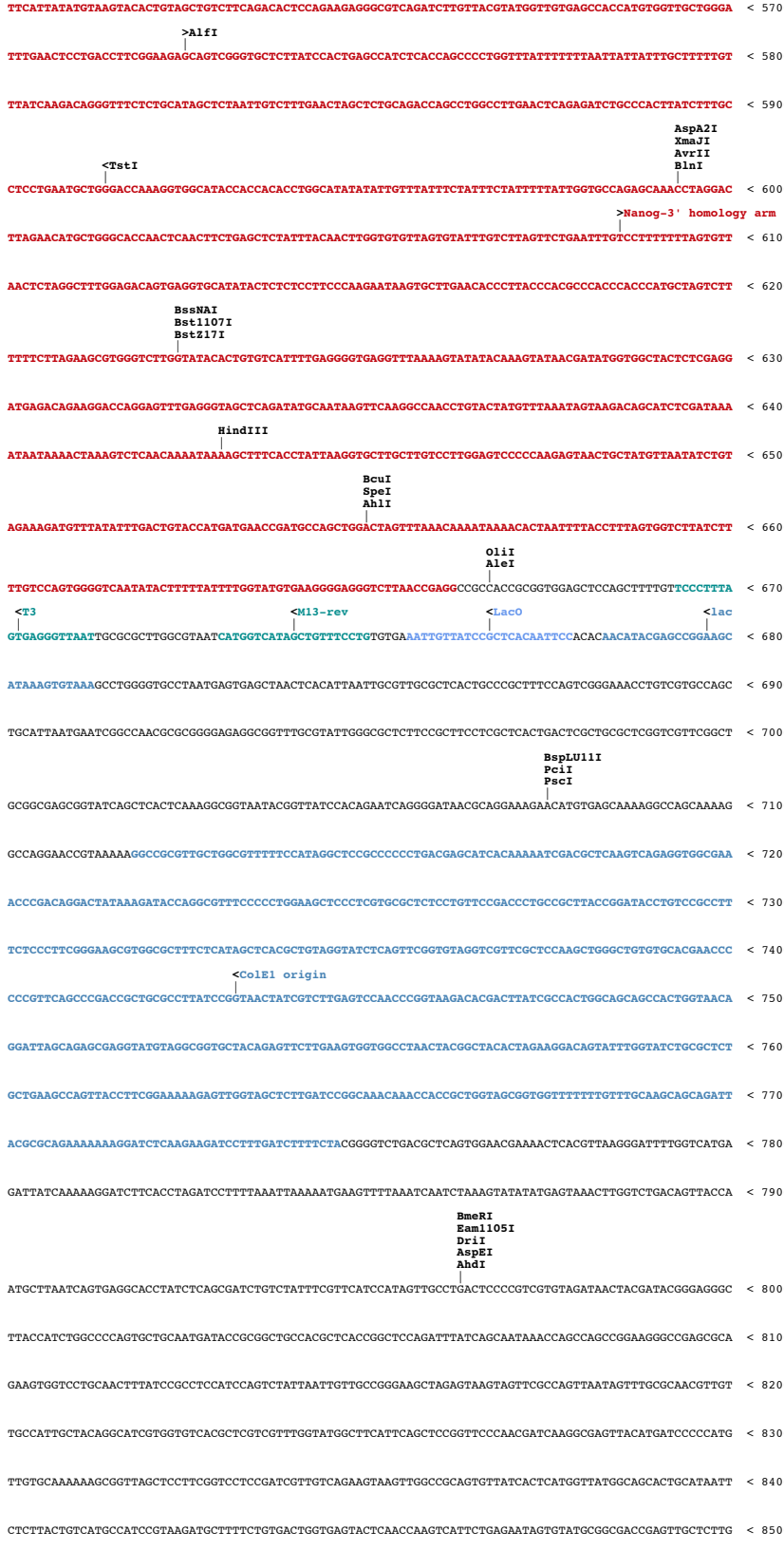
Restriction map of pTV-mNanog-HPP



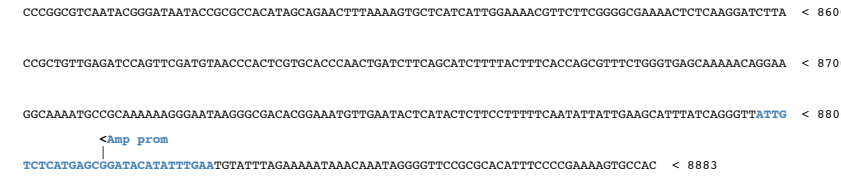
Supplementary Fig. S8, continued



Supplementary Fig. S8, continued



Supplementary Fig. S8, continued



Features :

T3	: [6712 : 6693 - CCW]
M13-fwd	: [599 : 616 - CW]
M13-rev	: [6750 : 6730 - CCW]
T7	: [626 : 645 - CW]
ColE1 origin	: [7746 : 7118 - CCW]
F1 ori	: [13 : 453 - CW]
LacZ alpha	: [528 : 460 - CCW]
LacO	: [6778 : 6756 - CCW]
LoxP	: [1696 : 1729 - CW]
LoxP	: [3946 : 3979 - CW]
Amp prom	: [8825 : 8797 - CCW]
lac	: [6812 : 6783 - CCW]
PP7 stem loop	: [4032 : 4146 - CW]
PP7 stem loop	: [4153 : 4267 - CW]
PP7 stem loop	: [4274 : 4388 - CW]
PP7 stem loop	: [4395 : 4509 - CW]
PP7 stem loop	: [4516 : 4630 - CW]
PP7 stem loop	: [4637 : 4751 - CW]
PP7 stem loop	: [4758 : 4872 - CW]
PP7 stem loop	: [4879 : 4993 - CW]
PP7 stem loop	: [5000 : 5114 - CW]
PP7 stem loop	: [5121 : 5235 - CW]
PP7 stem loop	: [5242 : 5356 - CW]
PP7 stem loop	: [5363 : 5477 - CW]
2A peptide	: [1642 : 1695 - CW]
2A peptide	: [2869 : 2922 - CW]
hsvTK	: [1732 : 2868 - CW]
HygR	: [2923 : 3978 - CW]
FactorXa site	: [1639 : 1650 - CW]
Nanog-5' homology arm	: [673 : 1641 - CW]
Nanog-3' homology arm	: [5509 : 6662 - CW]

Supplementary Table S1, Concentrations of chemical inhibitors.

	Concentrations of inhibitors (μM)		mean # of mRNA	Intrinsic noise	95% CI
	PD0325901	**CHIR99021**			
Serum condition	0	0	35.915	0.606	0.350
2i	0.2	0.6	146.620	0.316	0.079
	0.5	1.5	217.626	0.218	0.060
	0.75	2.25	151.372	0.246	0.059
	1	3	238.511	0.128	0.020
PD0325901	0.2	0	93.616	0.337	0.145
	0.5	0	211.453	0.196	0.050
	0.75	0	308.127	0.112	0.025
	1	0	319.824	0.126	0.026

Supplementary Table S2, Nucleotide sequences of smFISH probes used in this study.

Nanog **intronic probes**	**Dye: CAL Fluor Red 590**
Probe #	**Probe (5'-> 3')**
1	TTACTGGGTTCTTCGGGGAC
2	TTTTTTCTACTCTTACCCTA
3	AGAAGCAATAACCCTTCAGC
4	CCCGCTTATGTTAATGACTA
5	GGGTTTCCAGAAGAGTGATA
6	CAGACTAGAAGGCCAACGTA
7	TTATATTGCTCCGTCCTGTG
8	TAGGATGTTAGGTCTCCCTG
9	AAATGGGGTGCTCATTCCAA
10	CTAACTGTATAACCTCACCA
11	AAACGGCCATTTGGGCAAAT
12	AATGCTAACTGCTTCTGCTG
13	TAAGTGACATCCATATTCCC
14	TGAGCTCACAAACCCAGAAC
15	CTCCAGATGCTAGCTATAAG
16	AGACAATGAGCTTCAGACCT

Nanog **exonic probes**	**Dye: CAL Fluor Red 590**
Probe #	**Probe (5'-> 3')**
1	AGGTTTTAGGCAACAACCAA
2	ATGGCGAGGGAAGGGATTTC
3	AACTAGGCAAACTGTGGGGA
4	TTCCCAGAATTCGATGCTTC
5	AACTGCAGGCATTGATGAGG
6	AGCAAGAATAGTTCTCGGGA
7	AGAGCATCTCAGTAGCAGAC
8	TTCAGAGGAAGGGCGAGGAG
9	GAATCAGGGCTGCCTTGAAG
10	CTTTTGTTTGGGACTGGTAG
11	TGTCAGCCTCAGGACTTGAG
12	TGAGAGAACACAGTCCGCAT
13	CTGTCCTTGAGTGCACACAG
14	GAGGTACTTCTGCTTCTGAA
15	GAGAGTTCTTGCATCTGCTG
16	ATAGCTCAGGTTCAGAATGG
17	GAAACCAGGTCTTAACCTGC
18	TTGCACTTCATCCTTTGGTT
19	TCTTCAACCACTGGTTTTTC
20	TTCTGAATCAGACCATTGCT

21	GATACTCCACTGGTGCTGAG
22	GGGGATAGCTGCAATGGATG
23	ACATGGAAAGGCTTCCAGAT
24	AAGTTGGGTTGGTCCAAGTC
25	AGGTCTGGTTGTTCCAAGTT
26	GAAGTTATGGAGCGGAGCAG
27	CTGTACGTAAGGCTGCAGAA
28	CTGGCAGAGAAGTTTTGCTG
29	CTTCCAAATTCACCTCCAAA
30	TAAAATGCGCATGGCTTTCC
31	ATAATTCCAAGGCTTGTGGG
32	GTAAGTCTCATATTTCACCT
33	TGACTTTAAGCCCAGATGTT
34	CATGTTCTAAGTCCTAGGTT
35	AGCTCAGAAGTTGAGTTGGT
36	ACACTAACACACCAAGTTGT
37	TCTCCAAAGCCTAGAGTTAA
38	GAGAGAGTATATGCACCTCA
39	GGTGTTCAAGCACTTATTCT
40	AGACCCACGCTTCTAAGAAA
41	AAACCTCACCCCTCAAAATG
42	CCACCATATCGTTATACTTT
43	TTCTGTCTCATCCTCGAGAG
44	TATCTGAGCTACCCTCAAAC
45	CGAGATGCTGTCTTACTATT
46	GCAAGCACCTTAATAGGTGA
47	ACATAGCAGTTACTCTTGGG
48	ATCGGTTCATCATGGTACAG

MS2 probes*	**Dye: Cy5**
Probe #	**Probe (5'-> 3')**
1	TTTCTTGGCAATAAGTACCGTA
2	TTTGAAGATTCGACCTGGAG
3	GATGAACCCTGGAATACTGGAG

*each probe is labeled at both ends

PP7 probes*	**Dye: Cy3**
Probe #	**Probe (5'-> 3')**
1	TTCTAGGCAATTAGGTACCTTA
2	TTTCTAGAGTCGACCTGCAG
3	AATGAACCCGGGAATACTGCAG

*each probe is labeled at both ends

Supplementary Video S1, Live imaging of NM-G cells cultured in serum condition. NM-G cells cultured in serum condition were subjected to live imaging with a spinning-disc confocal microscope at 2-min intervals for 4 h. Z series of 41 focal planes with a step size of 0.25 μm were acquired. Maximum-intensity image projections are shown. Scale bar, 10 μm.

Supplementary Video S2, Live imaging of NM-G cells cultured in 2i condition. NM-G cells cultured in 2i condition were subjected to live imaging with a spinning-disc confocal microscope at 2-min intervals for 4 h. Z series of 41 focal planes with a step size of 0.25 μm were acquired. Maximum-intensity image projections are shown. Scale bar, 10 μm.

Supplementary Video S3, Simultaneous visualization of individual mRNA molecules and transcription site in the NM-G cell line. NM-G cells cultured in 2i conditions were subjected to live imaging with a spinning-disc confocal microscope using a 100× objective. Z series of 3 focal planes with a step size of 0.15 μm were continuously acquired at minimum intervals with 30-ms exposure time for about 2.5 s. The approximated collection time is indicated. Maximum-intensity image projections are shown. Scale bar, 5 μm.