SCIENCE CHINA
Life Sciences

• RESEARCH PAPER •

December 2014 Vol.57 No.12: 1183–1190
doi: 10.1007/s11427-014-4717-z

Expression and function of natural antisense transcripts in mouse embryonic stem cells

ZHAO Tong, WU ZhaoTing, WANG SiYu & CHEN LingYi*

State Key Laboratory of Medicinal Chemical Biology, Collaborative Innovation Center for Biotherapy, 2011 Collaborative Innovation Center of Tianjin for Medical Epigenetics and College of Life Sciences, Nankai University, Tianjin 300071, China

Received April 16, 2014; accepted May 29, 2014; published online September 5, 2014

Non-coding RNAs (ncRNAs), such as microRNAs and large intergenic non-coding RNAs, have been shown to play essential roles in regulating pluripotency. Yet, it is not clear the role of natural antisense transcripts (NATs), also belonging to ncRNAs, in embryonic stem cells. However, the role of NATs in embryonic stem cells remains unknown. We further confirmed the expression of the NATs of three key pluripotency genes, *Oct4*, *Nanog* and *Sox2*. Moreover, overexpression of *Sox2-NAT* reduces the expression of Sox2 protein, and slightly enhances the *Sox2* mRNA level. Altogether, our data indicated that like other ncRNAs, NATs might be involved in pluripotency maintenance.

natural antisense transcripts, embryonic stem cells, Sox2, Oct4, pluripotency

Citation: Zhao T, Wu ZT, Wang SY, Chen LY. Expression and function of natural antisense transcripts in mouse embryonic stem cells. Sci China Life Sci, 2014, 57: 1183−1190, doi: 10.1007/s11427-014-4717-z

The primary function of RNAs is well understood as intermediate of genetic information for protein synthesis. Recent years, non-coding RNAs (ncRNAs) have drawn a lot of attention, and the diverse functions of ncRNAs in various biological processes have been revealed [1–6]. ncRNAs can be further divided into two classes: structural and regulatory ncRNAs. Structural ncRNAs, including transfer RNA (tRNA), ribosomal RNA (rRNA), and small nucleolar RNA (snoRNA), are involved in mRNA translation and rRNA modification, whereas regulatory ncRNAs, such as small interfering RNA (siRNA), microRNA (miRNA), PIWI-interacting RNA (piRNA), and long non-coding RNA (lncRNA), play important roles in regulating gene expression.

Natural antisense transcripts (NATs), belonging to the lncRNA family, are transcribed from the opposite DNA strand to other transcripts, and at least partially overlap with sense RNAs [7,8]. NATs were first discovered in bacteria [9,10], and then in eukaryotes [11,12]. With the development of genomic approaches, it became clear that NATs are widespread throughout the genomes of many species, including mouse and human [13–15]. Yet, NATs are generally expressed in low abundance, normally more than 10–fold lower than sense transcripts [15,16]. NATs carry out the regulatory functions through distinct mechanisms, such as DNA replication interference, chromatin remodeling, transcriptional regulation, RNA masking, double-stranded RNA (dsRNA) mediated siRNA mechanism and translation interference [13].

Embryonic stem cells (ESCs) are derived from the inner cell mass of the blastocyst [17–19]. They are able to self-renew indefinitely and have the potential to differentiate into all types of cells in an organism. Thus, ESCs have great application value in basic developmental biology research, drug discovery, and cell replacement therapies. Three transcription factors, Nanog, Oct4 and Sox2, form a core regulatory circuit for pluripotency maintenance [20–24].

*Corresponding author (email: lingyichen@nankai.edu.cn)

ncRNAs also play critical roles in maintaining the pluripotency of ESCs, as well as in establishing the pluripotency during induced pluripotent stem (iPS) cell generation. Dicer knockout ESCs, in which siRNA and miRNA generation is blocked, are deficient in proliferation and differentiation [25]. Similarly, ESCs lacking DGCR8, a RNA-binding protein required for miRNA biogenesis, are unable to fully down-regulate pluripotency markers, and retain ESC colony morphology upon differentiation [26]. It has been shown that dozens of large intergenic non-coding RNAs (lincRNAs) are required for maintaining ESC unique gene expression profiles [27]. Both miRNAs and lincRNAs can facilitate the derivation of iPS cells [28,29]. Moreover, miRNAs alone, without exogenous reprogramming transcription factors, could reprogram somatic cells to the pluripotent state [30,31].

It becomes clear that ncRNAs play essential roles in pluripotent stem cells. NATs have been detected in human and mouse ESCs [32,33]. However, due to the experimental technique and focus, no genome-wide analysis of NATs has been performed in ESCs. In this study, through digital gene expression (DGE) profiling, we found that NATs are also genome-wide expressed in mouse ESCs. Next, we confirmed the expression of the NATs of three key pluripotency genes, *Nanog*, *Oct4* and *Sox2* by strand-specific reverse transcription-PCR (RT-PCR). The ends of these pluripotency gene NATs were determined by 3′- and 5′-Rapid amplification of cDNA end (RACE). We further demonstrated that *Sox2-NAT* suppresses the expression of Sox2 protein at the post-transcriptional level, whereas *Oct4-NAT* does not affect the expression of *Oct4*. All together, our data demonstrated the widespread expression of NATs in mouse ESCs, and that NATs may be involved in pluripotency maintenance through regulating the expression of pluripotency genes.

1 Materials and methods

1.1 Cell Culture

V6.5 mouse ESCs were cultured in growth medium consisting of 85% DMEM (high-glucose, Invitrogen, USA), 15% FBS (Hyclone, USA), 2 mmol L^{-1} L-glutamine, 5000 U mL^{-1} penicillin and streptomycin, 0.1 mmol L^{-1} non- essential amino acids (Invitrogen, USA), 0.1 mmol L^{-1} β-mercaptoethanol (Sigma, USA), and 1000 U mL^{-1} LIF (ESGRO, Chemicon, USA).

1.2 Digital Gene Expression (DGE) Profiling

DGE profiling and bioinformatic analysis were carried out by BGI Tech, China.

1.3 Strand-specific RT-PCR

Total RNA was extracted from cells using the RNeasy mini kit (Qiagen, Germany). cDNA synthesis was performed using the Tanscriptor First Strand cDNA Synthesis Kit (Roche, Switzerland) with strand-specific primer Table S1 in Supporting Information according to the manufacturer's instructions. PCR reactions were performed with EasyTaq (Transgen, China) in a Bio-Rad cycler. PCR cycling conditions were 95°C for 5 min, 35 cycles of 95°C for 15 s, 55°C for 15 s, and 72°C for 30 s.

1.4 RACE

Rapid amplification of cDNA end (RACE) was performed by the SMART RACE cDNA Amplification Kit (Clontech, USA). The RACE products were cloned into pEASY-T3 vector (Transgen, China) and then sequenced.

1.5 Quantitative RT-PCR

Total RNA was extracted from cells using the RNeasy mini kit (Qiagen, Germany). cDNA synthesis was performed using the Tanscriptor First Strand cDNA Synthesis Kit (Roche, Switzerland) with random primers according to the manufacturer's instruction. PCR reactions were performed with FastStart Universal SYBR Green Master (Roche, Switzerland) in a Bio-Rad iQ5 system. PCR cycling conditions were 95°C for 2 min, 40 cycles of 95°C for 15 s, 58°C for 15 s, and 72°C for 30 s, and then a melting curve of the amplified DNA was acquired. Quantification of target genes was normalized with *β-Actin*. Primer information was listed in Table S1 in Supporting Information.

1.6 Western blot Analysis

Cells were lysed in lysis buffer (Beyotime, China), and protein concentration was measured using a BCA protein assay kit (Beyotime, China) to ensure equal loading. The samples were resolved by SDS-PAGE, followed by transferring onto a PVDF membrane (Millipore, USA). Membranes were probed with anti-Sox2 (Cell Signaling Technology, USA), anti-β-Tubulin (Huada, China), and anti-Oct3/4 (Santa Cruz Biotechnology, USA). Bound primary antibodies were recognized by HRP-linked secondary antibodies (GE Healthcare, Germany). Immunoreactivity was detected by ECL Plus (Beyotime, China) and Kodak light film. Digital images of films were taken with Bio-Rad Molecular Imager Gel Doc XR. The intensity of bands was quantified with the Quantity One analysis software (Bio-Rad, USA).

1.7 Statistical Analysis

Data were analyzed by Student's *t* test. Statistically significant *P* values were indicated in figures as follows: **, $P<0.01$; *, $P<0.05$.

2 Results

2.1 Widespread expression of NATs in mouse ESCs

To determine the expression of NATs in mouse ESCs, duplicated mouse ESC RNA samples were subjected to DGE analysis. Solexa sequencing of the two DGE libraries yielded 6003694 and 5832791 clean sequence tags, respectively. Around 55% of the sequence tags were unambiguously mapped to ~13000 sense genes, while about 5% of the tags were mapped to ~8000 antisense genes (Table S3 in Supporting Information). The genes associated with antisense tags were listed in Table S4 in Supporting Information. The data demonstrated the widespread expression of NATs in mouse ESCs. Consistent with previous reports, NATs are generally expressed 10–fold lower than their sense counterparts.

2.2 Validating the NATs of pluripotency genes

Given the importance of pluripotency genes in ESC self-renewal, we are particularly interested in the NATs of pluripotency genes, such as *Nanog*, *Oct4* and *Sox2*. Indeed, sequence tags corresponding to these pluripotency gene NATs were identified (Table 1). The DGE data showed that *Oct4-NAT* and *Sox2-NAT* are expressed at higher levels than *Nanog-NAT*. To confirm the expression of the pluripotency gene *NATs*, strand-specific RT-PCR were performed. For each *NAT*, multiple primer pairs were used to avoid false-negative result due to RNA splicing. Meanwhile, multiple primer pairs also facilitated us to narrow down the 3′ and 5′-ends of the *NATs*. With this strategy, we validated the expression of *Nanog-*, *Oct4-* and *Sox2-NATs*, with N2, N3, O2-O5, and S2 primer pairs. The antisense transcripts were not detected with N1, N4, O1, O6, S1, and S3 primer pairs (Figure 1A and B), indicating that these regions are not transcribed in the antisense direction. With this information, we carried out 3′- and 5′-RACE experiments to precisely determine the 3′- and 5′-ends of these *NATs* (Figure 1C). Except for the 3′-end of *Nanog-NAT*, all the ends were identified (indicated with arrows in Figure 1A, and listed in Table S2 in Supporting Information). The failure to map the 3′-end of *Nanog-NAT* is probably due to the low abundance of the *Nanog-NAT* transcript. Interestingly, three different 5′-ends of *Oct4-NAT* were identified, likely caused by different transcription starting sites.

After defining the 3′- and 5′-ends of *Oct4-* and *Sox2-NATs*, we amplified the full length NATs by RT-PCR and sequenced them. The result showed that *Oct4-NAT* mRNA is spliced in the same pattern as its sense counterpart, despite the opposite transcription orientation. Both *Sox2* and *Sox2-NAT* are single exon genes. Based on the overlapping pattern between sense and antisense transcripts, NATs can be divided into three different types: head to head, tail to tail, and full overlap [34]. Both *Oct4-NAT* and *Sox2-NAT* belong to the full overlap type. Thus far, we have demonstrated the expression of *Nanog-*, *Oct4-* and *Sox2-NATs* in mouse ESCs by DGE, strand-specific RT-PCR, and cloning the full length NATs.

2.3 Regulatory function of *Oct4-* and *Sox2-NATs* in ESCs

To investigate whether *Oct4-* and *Sox2-NATs* regulate the expression of their sense transcripts, the full length *Oct4-* and *Sox2-NATs* were overexpressed in mouse ESCs by lipofectamine transfection, and the expression of *Oct4* and *Sox2* mRNA were measured by quantitative RT-PCR. The overexpression of *Sox2-NAT* slightly enhances the expression of *Sox2* mRNA, while *Oct4-NAT* does not affect the *Oct4* mRNA level (Figure 2A). Next, we asked whether *Oct4-* and *Sox2-NATs* regulate the expression of Oct4 and Sox2 at protein level. Western blot results showed that *Sox2-NAT* overexpression significantly down-regulates the level of Sox2 protein, whereas Oct4 protein remains unchanged upon *Oct4-NAT* overexpression (Figure 2B and 2C). Thus, *Sox2-NAT* regulates the expression of Sox2 protein post-transcriptionally.

Table 1 Antisense sequence tags of key pluripotency genes

Gene	Tag	Abundance			
		Tag copy number		TPM (Tag)	
		ES-1	ES-2	ES-1	ES-2
Nanog	CATGTCAGTGTGATGGCGAGG	2	6	0.3	1.0
	CATGGTGGCTCACAACCATAC	5	0	0.8	0.0
Oct4	CATGGGAGAGCCCAGAGCAGT	125	121	20.8	20.74
	CATGGTCTCCAGACTCCACCT	11	12	1.8	2.1
	CATGTTCTTAAGGCTGAGCTG	5	7	0.8	1.2
Sox2	CATGGATTCTCGGCAGCCTGA	249	210	41.5	36.0
	CATGGACATTTTTTTTTTGCC	8	6	1.3	1.0
	CATGTTTTCCTTTTGTACAAT	22	15	3.7	2.6
	CATGTTTATCTCGATAAATAC	25	17	4.2	2.9
	CATGGTCCGATTCCCCCGCCC	0	3	0.0	0.5

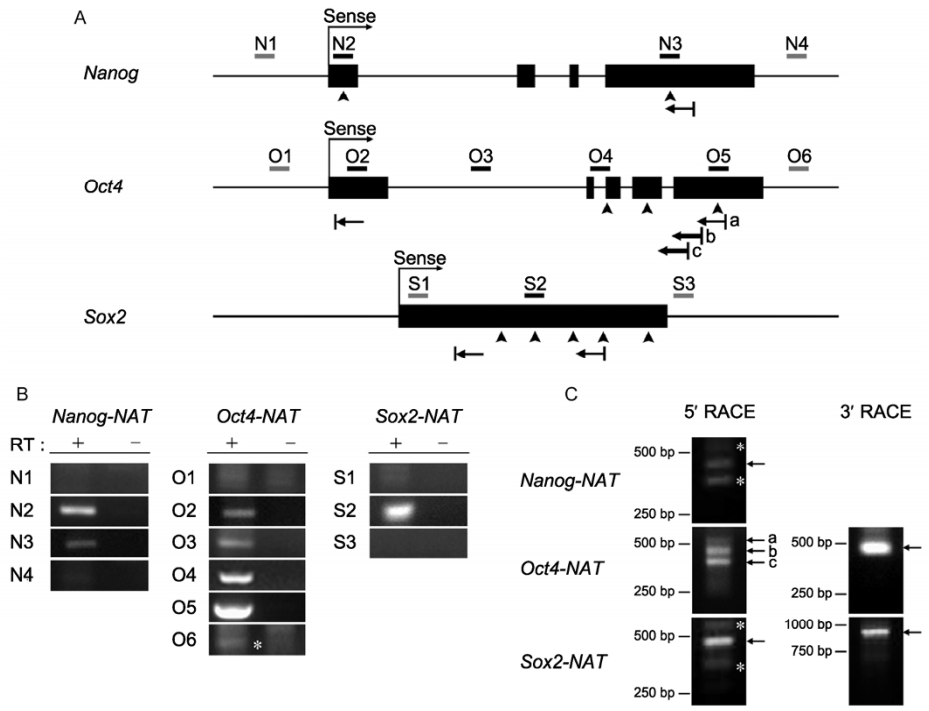

Figure 1 Validating the expression of *Nanog*-, *Oct4*- and *Sox2-NATs* in mouse ESCs. A, Schematic illustration of the *Nanog*, *Oct4* and *Sox2* loci. Exons are shown in black boxes, and arrows with "sense" indicate the sense transcription direction. Filled triangles represent sequence tags identified in DGE analysis. Short fragments N1-N4, O1-O6, and S1-S3, represent amplified regions in strand-specific RT-PCR. Black short fragments indicate detectable NATs, whereas grey short fragments mean undetectable NATs. The 5′- and 3′- ends of NATs are marked by arrows with short vertical lines. B, Confirmation the expression of *Nanog*, *Oct4* and *Sox2-NATs* by strand-specific RT-PCR. White asterisks mark non-specific bands proved by sequencing. C, Determination of the 3′- and 5′-ends *Nanog*, *Oct4* and *Sox2-NATs* by RACE. White asterisks mark non-specific bands proved by sequencing.

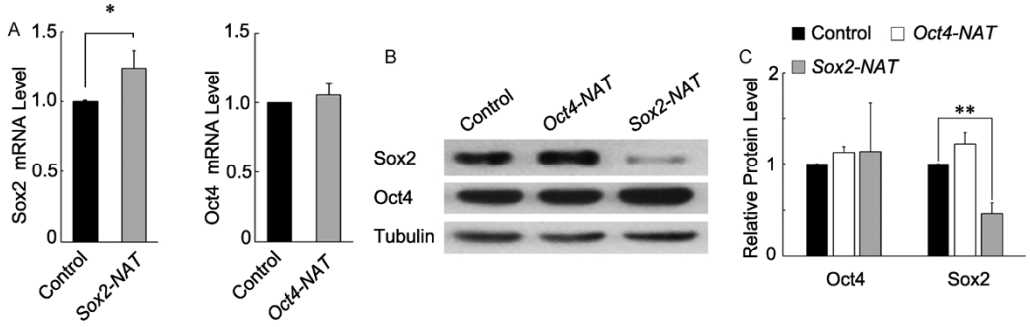

Figure 2 *Sox2-NAT* down-regulates Sox2 protein. A, Mouse ESCs were transfected relative with empty vector (control), vectors expressing *Oct4-NAT* or *Sox2-NAT*. 48 h after transfection, cells were harvested for RNA purification. And the expression of *Oct4* and *Sox2* mRNA were measured by quantitative RT-PCR. B, Mouse ESCs were transfected with empty vector (control), vectors expressing *Oct4-NAT* or *Sox2-NAT*. 48 h after transfection, cells were harvested and subjected to Western Blot. C, Quantification of Western Blot results in (B). Averages and standard deviations of three independent experiments are shown. Statistically significant *P* values were indicated in figures as follows: **, $P<0.01$; *, $P<0.05$.

2.4 *Sox2-NAT* functions as a ncRNA

Sequence analysis of *Sox2-NAT* revealed a putative open reading frame (ORF) at the 3′-end, even though lacking a stop codon (Figure 3). To rule out the possibility that the putative encoded protein accounts for the regulatory function of *Sox2-NAT*, a series of truncation mutants of *Sox2-NAT* were constructed. Only the F800 and R800 mutants remain the ability to reduce Sox2 protein expression, when they are overexpressed (Figure 3). Since 57 amino acid residues of the putative ORF, more than 1/3 of the putative protein, were deleted in the F800 mutant, it suggested that the regulatory function of *Sox2-NAT* is independent of the putative encoded protein. Moreover, the R600 mutant, harboring the intact ORF, failed to decrease Sox2 protein level, further ruling out that the putative encoded protein is required for the function of *Sox2-NAT*. Thus, *Sox2-NAT* functions as a ncRNA to down-regulate Sox2 protein.

The F400, F600, R400 and R600 mutants all lost the regulatory function on Sox2 protein (Figure 3), implying that

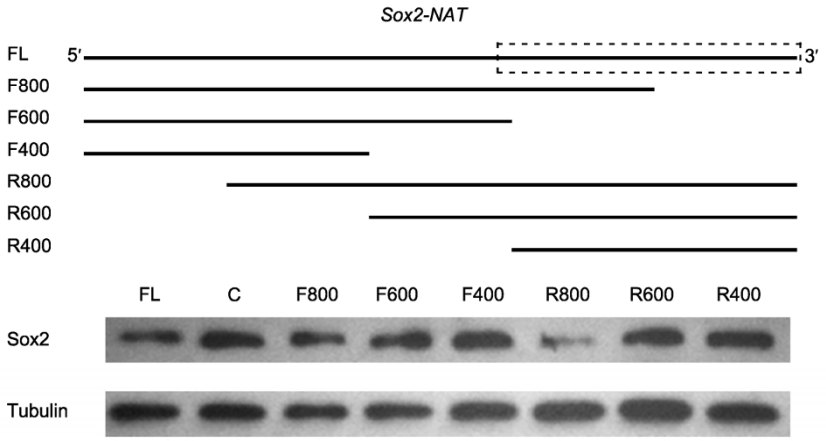

Figure 3 *Sox2-NAT* functions as an ncRNA. Black lines represent full length (FL) *Sox-NAT* and a series of truncation *Sox-NAT* mutants. The dotted box on the FL *Sox-NAT* marks the putative ORF. FL *Sox-NAT* and truncation *Sox-NAT* mutants were overexpressed in mouse ESCs for 48 h. Then Cells were harvested and subjected to Western Blot. The F800 and R800 mutants, as well as FL *Sox-NAT*, suppress the expression of Sox2 protein, while the rest mutants do not affect Sox2 protein expression. C: Control vector.

Sox2-NAT does not carry out its function through a dsRNA mediated siRNA mechanism. Instead, the proper folding of RNA and certain domain(s) are essential for *Sox2-NAT*'s function. We then used RNAfold (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi) to predict the secondary structures of the full length and truncated *Sox2-NATs*. Notably, the structures of mutants with large truncations, such as F400, F600, R400 and R600, are quite different from that of the full length *Sox2-NAT* (Figure 4A). It might explain why these truncation mutants are defect in regulating Sox2 protein. In contrast, the overall structures of the F800 and R800 mutants are similar to that of the full length *Sox2-NAT*, especially the central region highlighted with dotted rectangles (Figure 4A). This region might be involved in interacting with other protein and/or RNA molecules, thus indispensible for the function of *Sox2-NAT*.

3 Discussion

Over the past few years, it became clear that NATs are widely expressed in many species, and that they exert their regulatory functions through distinct mechanisms in various biological processes [7,8]. Thirty NATs, including *OCT4*-, *NANOG*- and *LIN28-NATs*, have been identified in human ESCs by reverse serial analysis of gene expression (SAGE) [32]. A recent study on lncRNAs identified 202 and 143 NATs in human and mouse ESCs, respectively [33]. In this study, we first demonstrated widespread expression of NATs in mouse ESCs by DGE analysis. Around 8000 genes are associated with antisense transcription. The discrepancy might be due to the conventional sequencing method used in the first study and the stringent call for large noncoding RNAs in the second study. The high sensitivity of DGE analysis might also facilitate detecting low abundant NATs. We then confirmed the expression of the NATs of three key pluripotency genes, *Nanog*, *Oct4* and *Sox2*, by strand-specific RT-PCR, 3′- and 5′-RACE, and cloning the full length NATs. Furthermore, we found that *Sox2-NAT* functions as a ncRNA to down-regulate Sox2 protein expression. Therefore, like other ncRNAs, NATs might be involved in regulating self-renewal and differentiation of ESCs. NATs may prevent overexpression of pluripotency genes, such as *Sox2*, thus facilitating ESCs exiting from the pluripotency state. On the other hand, NATs may contribute to ESC self-renewal by suppressing of differentiation genes.

Even though DGE analysis allows identification of NATs, it cannot provide the full-length information of NATs. To obtain the full-length information of NATs, RNA sequencing with the second generation sequencing technology (RNA-seq) is required. However, due to the low abundance of NATs, deeper sequencing is necessary to comprehensively detect genome-wide NATs.

Sox2-NAT regulates Sox2 protein expression post-transcriptionally (Figure 4B). Two evidences support that *Sox2-NAT* carries out this regulatory function as a ncRNA, but not through its putative encoded protein. First, truncation of the putative ORF does not impair the function of the F800 mutant. Second, the R600 mutant, which deletes the 5′-end of *Sox2-NAT* and keeps the ORF intact, loses the regulatory function. The proper folding of *Sox2-NAT* seems to be essential to down-regulate Sox2 protein. How *Sox2-NAT* suppresses Sox2 protein expression, remains to be explored. It has been shown that PU.1 antisense RNA inhibits the synthesis of PU.1 protein by blocking translation elongation [35]. It is possible that *Sox2-NAT* also exerts its regulatory function at the protein synthesis step. Alternatively, *Sox2-NAT* might interact with *Sox2* mRNA, as well as other proteins, to retain *Sox2* mRNA in the nucleus. Consequently, protein synthesis is blocked. Overexpression

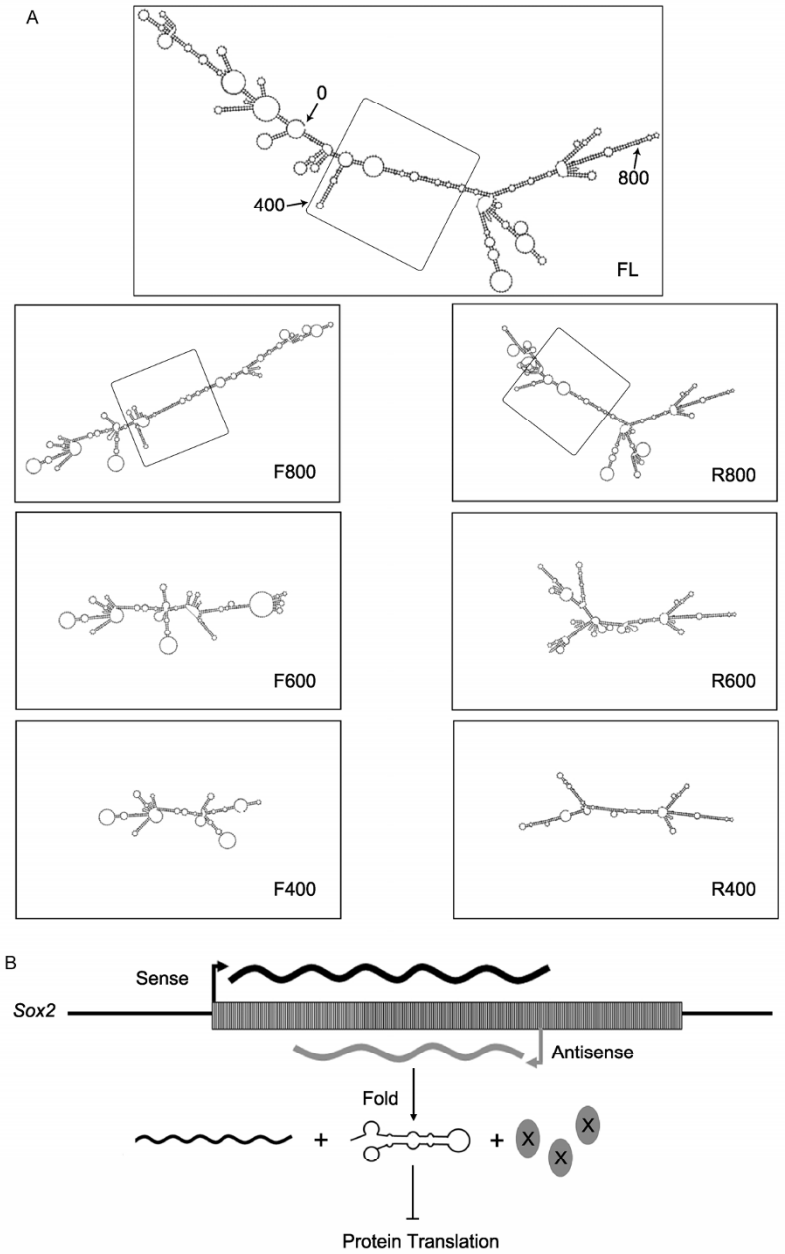

Figure 4 Proper folding of *Sox2-NAT* RNA may be essential for its regulatory function. A, Predicted secondary structures of *Sox2-NAT* and its truncation mutants. RNAfold (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi) was used to predict RNA secondary structure. Rectangles highlight the relatively conserved structure shared by FL, F800 and R800. B, A proposed model for the regulatory function of *Sox2-NAT*. *Sox2-NAT* should be properly folded, and interact with other RNA and protein molecules to suppress the translation of Sox2 protein.

of *Sox2-NAT* slightly enhances the level of *Sox2* mRNA. The formation of RNA duplex might stabilize *Sox2* mRNA, thus elevating the expression of *Sox2* mRNA. Another possibility is that reduced Sox2 protein level induces a compensatory response at transcriptional level, and more *Sox2* mRNA is transcribed.

Even though we did not find any function of *Oct4-NAT*, it does not mean that *Oct4-NAT* is just a product of transcriptional noise without any function. Overexpression of *Oct4-NAT* only allows us to identify the regulatory function in *trans*, but not in *cis*. It remains possible that *Oct4-NAT* functions in *cis* to regulate the transcription of the *Oct4* gene. To study the in *cis* function of *Oct4-NAT*, modulating the endogenous *Oct4-NAT* transcription is necessary. However, *Oct4-NAT* completely overlaps with *Oct4* sense transcript. Thus, with conventional siRNA or shRNA, it is impossible to specifically down-regulate *Oct4-NAT*, without affecting *Oct4*. It has been reported that chemically modified single-stranded siRNA (ss-siRNA) can achieve allele-specific silencing [36,37]. Yet, the ss-siRNA technology is not widely accessible to the research community. In the future, application of ss-siRNA might facilitate the

functional study of NATs.

In this study, we only investigated the narrowly defined NATs, also known as *cis*-NATs, which refer to RNAs transcribed from the opposite DNA strand of other transcripts, and at least overlapping partially with their sense counterparts. A broader definition of NATs refers to RNAs containing sequences complementary to other transcripts [38]. According to their origin, the broadly defined NATs can be divided into two classes, *cis*-NATs and *trans*-NATs. *cis*-NATs are transcribed from the same genomic loci as their *trans*-counterparts, while *trans*-NATs and their complementary sense transcripts are derived from separate genomic loci [38]. Further studies are required to investigate the abundance of *trans*-NATs in mouse ESCs, and the roles of *trans*-NATs in pluripotency maintenance.

In summary, our study confirmed the expression of NATs in mouse ESCs, and demonstrated a regulatory function of *Sox2-NAT*. This is just a beginning to understand the role of NATs in pluripotent stem cells. Further studies and new experimental technologies are required to elucidate the functions of NATs in pluripotency maintenance.

The authors declare that they have no conflict of interest.

This work was supported by the National Natural Science Foundation of China (31271547), the Natural Science Foundation of Tianjin, China (14JYBJC23600), the National Key Basic Research and Development Program of China (2010CB833603), the Program for New Century Excellent Talents (NCET-13-0293), the 111 Project Grant (B08011), and the Funds for National Basic Science Personnel Training (J1103503).

1 Fatica A, Bozzoni I. Long non-coding RNAs: new players in cell differentiation and development. Nat Rev Genet, 2014, 15: 7–21
2 Pasquinelli AE. MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship. Nat Rev Genet, 2012, 13: 271–282
3 Luteijn MJ, Ketting RF. PIWI-interacting RNAs: from generation to transgenerational epigenetics. Nat Rev Genet, 2013, 14: 523–534
4 Esteller M. Non-coding RNAs in human disease. Nat Rev Genet, 2011, 12: 861–874
5 Castel SE, Martienssen RA. RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. Nat Rev Genet, 2013, 14: 100–112
6 Pauli A, Rinn JL, and Schier AF. Non-coding RNAs as regulators of embryogenesis. Nat Rev Genet, 2011, 12: 136–149
7 Faghihi MA, Wahlestedt C. Regulatory roles of natural antisense transcripts. Nat Rev Mol Cell Biol, 2009, 10: 637–643
8 Pelechano V, Steinmetz LM. Gene regulation by antisense transcription. Nat Rev Genet, 2013, 14: 880–893
9 Itoh T, Tomizawa J. Formation of an RNA primer for initiation of replication of ColE1 DNA by ribonuclease H. Proc Natl Acad Sci USA, 1980, 77: 2450–2454
10 Lacatena RM, Cesareni G. Base pairing of RNA I with its complementary sequence in the primer precursor inhibits ColE1 replication. Nature, 1981, 294: 623–626
11 Spencer CA, Gietz RD, Hodgetts RB. Overlapping transcription units in the dopa decarboxylase region of Drosophila. Nature, 1986, 322: 279–281
12 Nepveu A, Marcu KB. Intragenic pausing and anti-sense transcription within the murine c-myc locus. Embo J, 1986, 5: 2859–2865
13 Li K, Ramchandran R. Natural antisense transcript: a concomitant engagement with protein-coding transcript. Oncotarget, 2010, 1: 447–452
14 Katayama S, Tomaru Y, Kasukawa T, Waki K, Nakanishi M, Nakamura M, Nishida H, Yap CC, Suzuki M, Kawai J, Suzuki H, Carninci P, Hayashizaki Y, Wells C, Frith M, Ravasi T, Pang KC, Hallinan J, Mattick J, Hume DA, Lipovich L, Batalov S, Engström PG, Mizuno Y, Faghihi MA, Sandelin A, Chalk AM, Mottagui-Tabar S, Liang Z, Lenhard B, Wahlestedt C; RIKEN Genome Exploration Research Group; Genome Science Group (Genome Network Project Core Group); FANTOM Consortium. Antisense transcription in the mammalian transcriptome. Science, 2005, 309: 1564–1566
15 He Y, Vogelstein B, Velculescu VE, Papadopoulos N, Kinzler KW. The antisense transcriptomes of human cells. Science, 2008, 322: 1855–1857
16 Ge X, Wu Q, Jung YC, Chen J, Wang SM. A large quantity of novel human antisense transcripts detected by LongSAGE. Bioinformatics, 2006, 22: 2475–2479
17 Evans MJ, Kaufman MH. Establishment in culture of pluripotential cells from mouse embryos. Nature, 1981, 292: 154–156
18 Martin GR. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci USA, 1981, 78: 7634–7638
19 Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, Jones JM. Embryonic stem cell lines derived from human blastocysts. Science, 1998, 282: 1145–1147
20 Nichols J, Zevnik B, Anastassiadis K, Niwa H, Klewe-Nebenius D, Chambers I, Schöler H, Smith A. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell, 1998, 95: 379–391
21 Chambers I, Colby D, Robertson M, Nichols J, Lee S, Tweedie S, Smith A. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell, 2003, 113: 643–655
22 Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M, Takahashi K, Maruyama M, Maeda M, Yamanaka S. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell, 2003, 113: 631–642
23 Avilion AA, Nicolis SK, Pevny LH, Perez L, Vivian N, Lovell-Badge R. Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev, 2003, 17: 126–140
24 Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, Gifford DK, Melton DA, Jaenisch R, Young RA. Core transcriptional regulatory circuitry in human embryonic stem cells. Cell, 2005, 122: 947–956
25 Kanellopoulou C, Muljo SA, Kung AL, Ganesan S, Drapkin R, Jenuwein T, Livingston DM, Rajewsky K. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev, 2005, 19: 489–501
26 Wang Y, Medvid R, Melton C, Jaenisch R, Blelloch R. DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. Nat Genet, 2007, 39: 380–385
27 Guttman M, Donaghey J, Carey BW, Garber M, Grenier JK, Munson G, Young G, Lucas AB, Ach R, Bruhn L, Yang X, Amit I, Meissner A, Regev A, Rinn JL, Root DE, Lander ES. lincRNAs act in the circuitry controlling pluripotency and differentiation. Nature, 2011, 477: 295–300
28 Judson RL, Babiarz JE, Venere M, Blelloch R. Embryonic stem cell-specific microRNAs promote induced pluripotency. Nat Biotechnol, 2009, 27: 459–461
29 Loewer S1, Cabili MN, Guttman M, Loh YH, Thomas K, Park IH, Garber M, Curran M, Onder T, Agarwal S, Manos PD, Datta S, Lander ES, Schlaeger TM, Daley GQ, Rinn JL. Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. Nat Genet, 2010, 42: 1113–1117
30 Miyoshi N, Ishii H, Nagano H, Haraguchi N, Dewi DL, Kano Y, Nishikawa S, Tanemura M, Mimori K, Tanaka F, Saito T, Nishimura

J, Takemasa I, Mizushima T, Ikeda M, Yamamoto H, Sekimoto M, Doki Y, Mori M. Reprogramming of Mouse and Human Cells to Pluripotency Using Mature MicroRNAs. Cell Stem Cell, 2011, 8: 633–638
31 Anokye-Danso F, Trivedi CM, Juhr D, Gupta M, Cui Z, Tian Y, Zhang Y, Yang W, Gruber PJ, Epstein JA, Morrisey EE. Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. Cell Stem Cell, 2011, 8: 376–388
32 Richards M1, Tan SP, Chan WK, Bongso A. Reverse serial analysis of gene expression (SAGE) characterization of orphan SAGE tags from human embryonic stem cells identifies the presence of novel transcripts and antisense transcription of key pluripotency genes. Stem Cells, 2006, 24: 1162–1173
33 Sigova AA, Mullen AC, Molinie B, Gupta S, Orlando DA, Guenther MG, Almada AE, Lin C, Sharp PA, Giallourakis CC, Young RA. Divergent transcription of long noncoding RNA/mRNA gene pairs in embryonic stem cells. Proc Natl Acad Sci USA, 2013, 110: 2876–2881
34 Osato N, Suzuki Y, Ikeo K, Gojobori T. Transcriptional interferences in cis natural antisense transcripts of humans and mice. Genetics, 2007, 176: 1299–1306
35 Ebralidze AK, Guibal FC, Steidl U, Zhang P, Lee S, Bartholdy B, Jorda MA, Petkova V, Rosenbauer F, Huang G, Dayaram T, Klupp J, O'Brien KB, Will B, Hoogenkamp M, Borden KL, Bonifer C, Tenen DG. PU.1 expression is modulated by the balance of functional sense and antisense RNAs regulated by a shared cis-regulatory element. Genes Dev, 2008, 22: 2085–2092
36 Lima WF, Prakash TP, Murray HM, Kinberger GA, Li W, Chappell AE, Li CS, Murray SF, Gaus H, Seth PP, Swayze EE, Crooke ST. Single-stranded siRNAs activate RNAi in animals. Cell, 2012, 150: 883–894
37 Yu D, Pendergraff H, Liu J, Kordasiewicz HB, Cleveland DW, Swayze EE, Lima WF, Crooke ST, Prakash TP, Corey DR. Single-stranded RNAs use RNAi to potently and allele-selectively inhibit mutant huntingtin expression. Cell, 2012, 150: 895–908
38 Lapidot M, Pilpel Y. Genome-wide natural antisense transcription: coupling its regulation to its different regulatory mechanisms. EMBO Rep, 2006, 7: 1216–1222

Supporting Information

Table S1 Primers used in the study

Table S2 5′- and 3′-end sequences of *Nanog*-, *Oct4*- and *Sox2-NAT*s

Table S3 Summary of the Solexa sequencing results

Table S4 Antisense sequence tags identified by the DGE profiling of mouse ESCs

The supporting information is available online at life.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.