THE EMBO JOURNAL

Manuscript EMBO-2011-78829

Human Long Non-coding RNAs Promote Pluripotency and Neuronal Differentiation by Association with Chromatin Modifiers and Transcription Factors

Shi-Yan Ng, Rory Johnson and Lawrence W. Stanton

Corresponding author: Lawrence W. Stanton, Genome Institute of Singapore

Review timeline:

Submission date:	16 July 2011
Editorial Decision:	15 August 2011
Revision received:	26 October 2011
Editorial Decision:	08 November 2011
Revision received:	11 November 2011
Accepted:	17 November 2011

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 15 August 2011

Thank you very much for submitting your research manuscript on potential involvement of long non-coding RNA's in pluripotency and cellular differentiation for consideration to The EMBO Journal editorial office.
I did receive comments from three scientists that confirm high interest and timeliness in your findings. At the same time, they also reveal significant shortcomings that relate to issues as proper annotation and characterization of lncRNAs-properties. There most important concerns center on missing molecular-mechanistic insight (target genes?), directness of lncRNA's interplay with pluripotency-transcription factors and indeed functional evidence for their role in pluripotency in general and neurogenesis preferably outside the established differentiation model. As these are relatively strong concerns and given the competitiveness of the topic, it needs focusing your efforts to go as far as possible during the relatively short period granted for major revisions.
Following the refs' requests hESC-knockdowns of lncRNAs upregulated in NPCs to assess their differentiation potential would be informative. Some ChIP-seq data (that might also address the issue of TF's interplay) and microarrays to elucidate (direct) targets or effects on more general signaling pathways would be desired to substantiate the findings. Finally, evidence for the expression/functionality(?) of the indicated lncRNA's in neuronal tissues would address physiological significance.

I am aware that these are relatively strong demands and I would thus be happy to discuss/clarify specific points (also in light of relevant new literature) in more detail. Please do not hesitate to contact me in case of further questions (preferably via E-mail).

Lastly, I do have to formerly remind you that it is EMBO_J policy to allow a single round of revisions only and that the ultimate decision depends on the content and strength of your adequately modified version.

I am very much looking forward to your response and remain with best regards

Sincerely yours,

Editor
The EMBO Journal

REFEREE COMMENTS

Referee #1:

This is a solid and interesting paper that examines the expression and functional role of long noncoding RNAs (lncRNAs) exhibiting stage-specific expression in human embryonal stem cells (hESCs) and induced pluripotent stem cells (iPSCs), compared (1) to derived neural progenitor cells (NPCs) expressing neural progenitor markers and radial glia markers, and (2) to dopamine neurons (DANs) that were secondarily derived from these NPCs.

The paper reports two main findings: (1) that siRNA-mediated knockdown of 4 lncRNAs that are exclusively expressed in hESCs / iPSCs and that are downregulated upon differentiation of these stem cells to NPCs causes a degree of loss of pluripotency in hESCs, which was more pronounced when all were coincidentally targeted; and (2) that siRNA-mediated knockdown in NPCs of 4 lncRNAs that were highly expressed in mature dopamine neurons and whose sequences are relatively conserved (3 of which are nuclear-localized) altered subsequent differentiation fate from a primarily neurogenic to a gliogenic program. As the authors note, the findings not only indicate a role for these RNAs in pluripotency and neuronal differentiation, but also point to the wider value of hESCs and iPSCs as model systems to study the roles of lncRNAs in human development, which adds to the significance of the paper.

Specific comments:

While the authors acknowledge the possibility, if not likelihood, that at least some of these lncRNAs may be acting via the regulation of the epigenetic processes that determine cell phenotype and fate, they do not cite recent articles that review this emerging area, nor acknowledge the previous study on the differential expression and epigenetic regulation of lncRNAs during neuronal-glial fate specification and oligodendrocyte lineage maturation (Mercer et al., BMC Neuroscience 2010).

The demonstration of phenotypic impact is central, and this study adds to the still limited but growing number of reports demonstrating the importance of lncRNAs in differentiation and development processes. However, given the expectations of the readers of EMBO J, it is a pity that the study did not go further and include chIP-seq and RNA-seq analysis of the cells before and after / with and without lncRNA knockdown to obtain molecular insight into the processes involved, both locally (in the vicinity of the lncRNAs, as a number of studies suggest that the epigenetic effects of their expression are felt locally) and globally, especially at loci that are indicative of and/or involved in the specification of pluripotency and lineage-specificity.

It is also a pity that the study did not include knockdown in hESCs of lncRNAs that are upregulated in NPCs, to determine whether this altered the differentiation into the latter, as was done for NPCs > DAN. That is, the paper presents the effects of lncRNA knockdown on hESC pluripotency (hESC > hESC) and DAN differentiation from NPCs (NPC > DAN) but not for the intermediate step of hESC > NPC, which would have been equally interesting, and perhaps more interesting than the NPC > DNA transition.

In this context, the paper states that the 4 lncRNAs studied in the NPC > neuron differentiation were "highly expressed in the mature dopamine neurons", implying that they are among the few (see Figure 3, total 35? - see below) that are not highly expressed in the more glia-like NPCs, although

this was not explicitly stated. This should be clarified.

The paper states (p. 16), in presenting the case for the selection of 4 lncRNAS for knockdown studies (from an initial 35 candidate lncRNAs that were highly expressed in DANs) that "12 out of these 25 lncRNAs were highly conserved among vertebrates, indicating that they could be functional transcripts. In addition, RMST (3 alternatively spliced transcripts AK056164, AF429305 and AF429306), lincRNA-ZFHX4 (AK124684), lincRNA-CACNA2D1 (AK055040), and BLID-OT or BLID overlapping transcript (AK091713) were very highly conserved". This implies that the latter were more highly conserved than the former": Is this the case? If so, the quantitative basis for this assessment should be made clear.

Referee #2:

In the manuscript, "Long non-coding RNAs are required for differentiation of human neurons", the authors develop an effective differentiation method from human stem cells into dopamine neurons. Samples were taken at three stages during the differentiation process and were subjected to custom designed arrays in order to identify putative long non-coding RNAs (lncRNAs). The authors identified four lncRNAs expressed specifically in undifferentiated pluripotent stem cells that may be associated with a pluripotent role. Knockdown of the pluripotent lncRNAs resulted in decreased mRNA levels of known pluripotent markers. Similarly, the authors identify four lncRNAs highly expressed in dopamine neurons. Knockdown experiments of these RNAs resulted in reduced neuronal differentiation. While it is exciting that one can potentially manipulate lncRNAs to guide pluripotency or neuronal differentiation, the mechanism of action still remains uncharacterized. Key controls for the experiments are also lacking, limiting the enthusiasm for this work.

Concerns:

1. The authors developed a modified SDIA differentiation method for hESCs into dopamine neurons. While the authors demonstrated that these neurons express markers characteristic of their nature by immunostaining in addition to showing that these cells are capable of releasing dopamine in vitro, it would be more convincing if they could compare the global expression of these cells to mature dopamine neurons and other types of neurons.

2. The authors designed a human lncRNA microarray using annotated lncRNAs from published sources. RNA extracted at different stages of the differentiation process was used in the custom-made microarray in order to identify populations of lncRNAs uniquely expressed at particular stages. Although the lncRNAs identified by this study are validated in the same cell culture system, it would be interesting to known whether these characterized lncRNAs are expressed in other hES cell lines as well as in authentic brain tissue. In addition, it would be interesting to known whether iPS cells expressed the identified pluripotent lncRNAs.

3. The annotation of transcripts into lncRNAs is not well described-- the curation and annotation needs to be carefully described and made public for a custom platform. Although the authors summarize well the microarray findings, additional details about the nature of the lncRNAs that they decide to follow-up on could benefit this study. What are the criteria to lncRNAs noncoding, and what is the evidence that the criteria are accurate? Are the lncRNAs intronic? What is their length? What is the distance between these lncRNAs and the proximal gene? How is their genomic localization associated with known chromatin marks in hESCs?

4. In the siRNA experiments, it seems that only one siRNA was used per lncRNA, which raises the concern of off target effects. Multiple independent siRNAs should be tested. Also, the knockdowns are incomplete, and some of the observed effects are modest, raising concerns for nonspecific effects.

5. No mechanism is offered for the actions of the lncRNAs. The authors show decreased expression of pluripotent lncRNAs when pluripotent transcription factors (TFs) are knockdown. Similarly, knockdown of lncRNAs results in reduced expression of TFs. In addition, the authors address published ChIP-seq studies in hESCs revealing pluripotent TFs binding sites near transcription start sites (TSS) of the lncRNAs. Do the lncRNAs directly interact with the pluripotent TFs?

6. If one knock down one of the lncRNAs needed to maintain pluripotency, can its function be rescued when a different pluripotency ncRNA is overexpressed? Does the same phenomenon occur when a ncRNA needed for the differentiation into domapine neurons is knockdown? Can it be rescued by another ncRNA with potentially redundant function?

7. The title and abstract contain statements that are too strong. The results at best show that lncRNAs influence neuronal differentiation in vitro. The authors do not provide evidence that lncRNAs are required or even involved in human neurogenesis in vivo.

Referee #3:

In this study the authors have used a custom array based approach to identify lncRNAs that are misregulated upon the differentiation of hES cells to neural progenitor cells or dopamine neurons. Four hESC-specific lncRNAs were identified which have potential roles in maintenance of pluripotency and four additional lncRNAs were identified that appear to be indispensable for neuronal differentiation. The study of lncRNAs is a very timely and exciting topic of investigation and the authors have identified several candidate RNAs that may have important roles in the maintenance of pluripotency or differentiation. However, the authors have not proven that these candidates are bonafide lncRNAs that do not produce small peptides. Furthermore, these potential candidates have not been studied in enough depth to be sure that the authors are observing a direct rather than indirect effect. The authors should focus their efforts on one or two candidates and provide significant mechanistic insights into their precise function.

1. Throughout the MS, the authors refer to transcripts as lncRNAs. What is the evidence that these are truly "non-coding"? For example, what is their codon-substitution frequency? Also, it would be good to see experimental evidence (e.g. in vitro translation) that significant ORFs of their key candidates are not translated.

2. The authors have not addressed the actual transcript size of each of these lncRNAs and their abundance in cells.

3. All RNAs are purely based on GeneBank entries. There is no cloning and RACE data on any of the candidate RNAs to validate transcript structure, transcription start and transcription end sites. Existing RNA-Seq data (ENCODE project on the UCSC browser) supports that some of the transcripts (e.g. lincRNA-CD83; lincRNA-ZNF281) may in fact form part of longer transcripts.

4. The data supporting a role for the transcription factor POU5F1 in regulating several of the lncRNAs is reasonable, however a role for NANOG in regulating lncRNAs 281 and HESRG is weak as knock-down of NANOG only resulted in a 20-30% down regulation of these lncRNAs.

5. In Figures 4 and 5 only some markers of a given lineage are up-/downregulated upon lncRNA candidate knock-down, while the other markers of the target lineage show hardly any change. This seems unexpected. How do lineage markers change in an Oct4 and / or Nanog control knock-down?

6. The authors claim to achieve a larger effect when they K-D 4 lncRNA vs individual ones and suggest that there may be redundancy among the lncRNAs. However, there is no biological evidence to suggest that any of these lncRNAs function in a similar way. In fact their biological functions are unknown.

7. In Figure 6 immunofluorescence staining for MAP2 in lncRNA knock-down cells over a large field of view is virtually abolished (panel A), but there is little or no change in the corresponding MAP2 transcript levels. How can this be explained?

8. The authors point out (p.18) that BLID-OT overlaps known miRNAs that were shown to function

in neural differentiation. Are these miRNAs lost upon BLID-OT knock-down, and can this explain claimed phenotypes?

9. "POU5F1 and NANOG binding sites are referred to near the transcription start sites of the lncRNAs" (p.13). 5' RACE needs to be done to determine/confirm the respective transcription start sites to make this claim.

10. The authors state that some hESC colonies are morphologically different upon siRNA transfection targeting candidate lncRNAs (p15). This should be supported by providing quantitative data on the fraction of abnormal colonies in siRNA vs. control conditions. Similar quantitation should be provided for the number of colonies with down-regulated GFP(Oct4) reporter fluorescence (Figure 4A).

11. "Fewer and shorter neurite extensions" (p16) were observed in cells transfected with siRNAs against lncRNA candidates. However, there is no data presented to support this claim, and quantitation as in the point above, should be performed.

1st Revision - authors' response 26 October 2011

Response to Reviewer's Comments

We provide a substantially revised version of our manuscript taking into the consideration the insightful and constructive issues raised by the three referees. In particular, we have added additional data that speak to the key issues raised by each reviewer about the mechanism of action for the lncRNAs that we show are important in regulating neuronal lineage specification. We have also addressed each of the minor issues and points of clarification raised by the reviewers. A detailed response (in *italic)* to each reviewer comment follows.

We trust you will find that we have seriously and adequately addressed all issues raised. In so doing we have undoubtedly improved the quality of this paper, making it a high impact study suitable for publication in the EMBO Journal.

Reviewer #1

This is a solid and interesting paper that examines the expression and functional role of long noncoding RNAs (lncRNAs) exhibiting stage-specific expression in human embryonal stem cells (hESCs) and induced pluripotent stem cells (iPSCs), compared (1) to derived neural progenitor cells (NPCs) expressing neural progenitor markers and radial glia markers, and (2) to dopamine neurons (DANs) that were secondarily derived from these NPCs.

The paper reports two main findings: (1) that siRNA-mediated knockdown of 4 lncRNAs that are exclusively expressed in hESCs / iPSCs and that are downregulated upon differentiation of these stem cells to NPCs causes a degree of loss of pluripotency in hESCs, which was more pronounced when all were coincidentally targeted; and (2) that siRNA-mediated knockdown in NPCs of 4 lncRNAs that were highly expressed in mature dopamine neurons and whose sequences are relatively conserved (3 of which are nuclear-localized) altered subsequent differentiation fate from a primarily neurogenic to a gliogenic program. As the authors note, the findings not only indicate a role for these RNAs in pluripotency and neuronal differentiation, but also point to the wider value of hESCs and iPSCs as model systems to study the roles of lncRNAs in human development, which adds to the significance of the paper.

1. While the authors acknowledge the possibility, if not likelihood, that at least some of these lncRNAs may be acting via the regulation of the epigenetic processes that determine cell phenotype and fate, they do not cite recent articles that review this emerging area, nor acknowledge the previous study on the differential expression and epigenetic regulation of

lncRNAs during neuronal-glial fate specification and oligodendrocyte lineage maturation (Mercer et al., BMC Neuroscience 2010).

We have mentioned the relevant recent references (Mercer et al., BMC Neuroscience 2010; Guttman et al., Nature 2011), in the introduction (p. 2).

2. The demonstration of phenotypic impact is central, and this study adds to the still limited but growing number of reports demonstrating the importance of lncRNAs in differentiation and development processes. However, given the expectations of the readers of EMBO J, it is a pity that the study did not go further and include chIP-seq and RNA-seq analysis of the cells before and after / with and without lncRNA knockdown to obtain molecular insight into the processes involved, both locally (in the vicinity of the lncRNAs, as a number of studies suggest that the epigenetic effects of their expression are felt locally) and globally, especially at loci that are indicative of and/or involved in the specification of pluripotency and lineage-specificity.

   All three reviewers have raised similar concerns that our paper lacked mechanisitic insight. We have thus included results from several new experiments to provide mechanistic insight regarding the lncRNAs in this study. For one, we performed the lncRNA knockdowns, as well as a nonspecific RNAi control, and analyzed changes in the transcriptome by means of a microarray, instead of RNA-seq suggested by the reviewer. These additional results are presented in Figure 4G and 4H, as well as Supplementary Table S6.
   In addition, RNA-immunoprecipitation studies are now included and we elucidated protein partners of the lncRNAs, achieving the mechanistic goals set by the reviewer. These extensive new experimental results are now presented in Figures 5 and 7.

   We also discussed the local (cis) versus global (trans) mode of action (p. 14-15) and supplemented with observations from our study (Figure 7B and Supplementary Figure S8.)

3. It is also a pity that the study did not include knockdown in hESCs of lncRNAs that are upregulated in NPCs, to determine whether this altered the differentiation into the latter, as was done for NPCs > DAN. That is, the paper presents the effects of lncRNA knockdown on hESC pluripotency (hESC > hESC) and DAN differentiation from NPCs (NPC > DAN) but not for the intermediate step of hESC > NPC, which would have been equally interesting, and perhaps more interesting than the NPC > DAN transition.

   We agree that it would be interesting to also delve more deeply into the ES > NPC transition. However, we chose to focus on the two developmental processes of ESC pluripotency and NPC>DAN differentiation, because they are of particular therapeutic / biotechnological importance. While we could provide some preliminary evidence that the NPC-lncRNAs affect differentiation of hESCs, it would not have been a definitive analysis without the subsequent mechanistic elucidation, which would be an entire new line of research. To reconcile Referee #1's suggestion to look at more lncRNAs in another developmental window and Referee #3's suggestion to focus on one or two candidates, we have focused on elucidating the mechanism for the lncRNAs we have originally set out to test.

4. In this context, the paper states that the 4 lncRNAs studied in the NPC > neuron differentiation were "highly expressed in the mature dopamine neurons", implying that they are among the few (see Figure 3, total 35? - see below) that are not highly expressed in the more glia-like NPCs, although this was not explicitly stated. This should be clarified.

   We have clarified the expression patterns of the 4 neuronal lncRNAs (p. 10-11). "From the microarray, we identified a group of 35 neuronal lncRNAs, which were highly expressed in mature neurons (more than 3-fold) compared to hESCs and NPCs (Supplementary Figure S2 and Supplementary Table 5). …… As a proof of concept, we focused on 4 neuronal lncRNAs for functional studies, namely, RMST (AK056164, AF429305 and AF429306), lncRNA_N1 (AK124684), lncRNA_N2 (AK091713), and lncRNA_N3 (AK055040)."

5. The paper states (p. 16), in presenting the case for the selection of 4 lncRNAS for knockdown studies (from an initial 35 candidate lncRNAs that were highly expressed in DANs) that "12 out of these 25 lncRNAs were highly conserved among vertebrates, indicating that they could

be functional transcripts. In addition, RMST (3 alternatively spliced transcripts AK056164, AF429305 and AF429306), lincRNA-ZFHX4 (AK124684), lincRNA-CACNA2D1 (AK055040), and BLID-OT or BLID overlapping transcript (AK091713) were very highly conserved". This implies that the latter were more highly conserved than the former": Is this the case? If so, the quantitative basis for this assessment should be made clear.

We note that lncRNAs, even well characterized functional transcripts, are poorly conserved across species (p. 3). Sequence conservation therefore is not a good indicator of functional prediction of lncRNAs. In view of this, we have removed the discussion of sequence conservation of lncRNAs.

Reviewer #2

In the manuscript, "Long non-coding RNAs are required for differentiation of human neurons", the authors develop an effective differentiation method from human stem cells into dopamine neurons. Samples were taken at three stages during the differentiation process and were subjected to custom designed arrays in order to identify putative long non-coding RNAs (lncRNAs). The authors identified four lncRNAs expressed specifically in undifferentiated pluripotent stem cells that may be associated with a pluripotent role. Knockdown of the pluripotent lncRNAs resulted in decreased mRNA levels of known pluripotent markers. Similarly, the authors identify four lncRNAs highly expressed in dopamine neurons. Knockdown experiments of these RNAs resulted in reduced neuronal differentiation. While it is exciting that one can potentially manipulate lncRNAs to guide pluripotency or neuronal differentiation, the mechanism of action still remains uncharacterized. Key controls for the experiments are also lacking, limiting the enthusiasm for this work.

1. The authors developed a modified SDIA differentiation method for hESCs into dopamine neurons. While the authors demonstrated that these neurons express markers characteristic of their nature by immunostaining in addition to showing that these cells are capable of releasing dopamine in vitro, it would be more convincing if they could compare the global expression of these cells to mature dopamine neurons and other types of neurons.

   Using our modified SDIA approach, we showed highly efficient dopaminergic differentiation. To date, no microarray or RNA-seq data are available from enriched populations of human dopaminergic neurons, as these cells are most difficult to attain in reasonable yields and purity. Comparing global expression of our neurons with mixed cultures would not be insightful. In view of this limitation still to address this concern, we performed a gene ontology analysis of the genes upregulated in our dopaminergic neurons. The enriched biological processes from gene ontology studies are provided in Table I. The top 10 terms are associated with neuronal differentiation. In addition, we also performed a qPCR to confirm that markers for other neuronal subtypes (GABAergic, motor, sertonergic and noradrenergic) were absent/ weakly expressed (Figure 2J). Together these data support our conclusions that these cells are being preferentially directed towards the dopaminergic lineage.

2. The authors designed a human lncRNA microarray using annotated lncRNAs from published sources. RNA extracted at different stages of the differentiation process was used in the custom-made microarray in order to identify populations of lncRNAs uniquely expressed at particular stages. Although the lncRNAs identified by this study are validated in the same cell culture system, it would be interesting to known whether these characterized lncRNAs are expressed in other hES cell lines as well as in authentic brain tissue. In addition, it would be interesting to known whether iPS cells expressed the identified pluripotent lncRNAs.

   Expression profiling of lncRNAs in different somatic tissues and pluripotent stem cells was performed to address this concern and is now presented in Figure 3A (for pluripotency lncRNAs) and Figure 6A (for neuronal lncRNAs). The pluripotency lncRNAs identified in H1

hESCs were also expressed in iPSCs (Figure 3A). These data thus provide independent validation of the cell-type specificity of the lncRNAs.

3. The annotation of transcripts into lncRNAs is not well described-- the curation and annotation needs to be carefully described and made public for a custom platform. Although the authors summarize well the microarray findings, additional details about the nature of the lncRNAs that they decide to follow-up on could benefit this study. What are the criteria to lncRNAs noncoding, and what is the evidence that the criteria are accurate? Are the lncRNAs intronic? What is their length? What is the distance between these lncRNAs and the proximal gene? How is their genomic localization associated with known chromatin marks in hESCs?

   The catalogue of lncRNAs used in this paper comes from the publication Jia et al (RNA. 2010 Aug;16(8):1478-87). In that publication, the properties of the lncRNAs were described, including their protein-coding status. We have corrected the present manuscript to now provide more detailed annotation of the lncRNAs and describe more fully the curation process. Protein-coding potential was determined using the Coding Potential Calculator (CPC) for reasons stated in the text (p. 8). The genomic location, transcript length, type of lncRNA (intronic/ intergenic etc), and CPC scores are presented in Tables II and III. Histone marks and RNA Pol II occupancy was also examined and shown in Supplementary Figure S3.

4. In the siRNA experiments, it seems that only one siRNA was used per lncRNA, which raises the concern of off target effects. Multiple independent siRNAs should be tested. Also, the knockdowns are incomplete, and some of the observed effects are modest, raising concerns for nonspecific effects.

   We used 2 siRNAs for knockdown of each lncRNA (Methods, Supplementary Table S2, Supplementary Figures S5 and S7). We determined the most effective siRNA by qPCR and subsequent RNAi experiments were performed using the most effective siRNA.

5. No mechanism is offered for the actions of the lncRNAs. The authors show decreased expression of pluripotent lncRNAs when pluripotent transcription factors (TFs) are knockdown. Similarly, knockdown of lncRNAs results in reduced expression of TFs. In addition, the authors address published ChIP-seq studies in hESCs revealing pluripotent TFs binding sites near transcription start sites (TSS) of the lncRNAs. Do the lncRNAs directly interact with the pluripotent TFs?

   As we discussed above in response to Reviewer 1, to obtain mechanistic insight into lncRNA function, we performed RNA-immunoprecipitation to determine if the lncRNAs physically associate with chromatin modifiers and pluripotent transcription factors (Figures 5 and 7). Pluripotent lncRNAs were shown to interact with SUZ12 and SOX2, but not OCT4, and a possible regulatory mechanism is proposed in Fig. 8.

6. If one knock down one of the lncRNAs needed to maintain pluripotency, can its function be rescued when a different pluripotency ncRNA is overexpressed? Does the same phenomenon occur when a ncRNA needed for the differentiation into domapine neurons is knockdown? Can it be rescued by another ncRNA with potentially redundant function?

   The reviewer raises a valid concern that, it may be too speculative to suggest fucntional redundancy of lncRNAs. Therefore, we have removed any discussion on redundancy of action of the lncRNAs.

7. The title and abstract contain statements that are too strong. The results at best show that lncRNAs influence neuronal differentiation in vitro. The authors do not provide evidence that lncRNAs are required or even involved in human neurogenesis in vivo.

   We have amended the title, and the abstract accordingly to better represent our work.

Referee #3

In this study the authors have used a custom array based approach to identify lncRNAs that are misregulated upon the differentiation of hES cells to neural progenitor cells or dopamine neurons. Four hESC-specific lncRNAs were identified which have potential roles in maintenance of pluripotency and four additional lncRNAs were identified that appear to be indispensable for neuronal differentiation. The study of lncRNAs is a very timely and exciting topic of investigation and the authors have identified several candidate RNAs that may have important roles in the maintenance of pluripotency or differentiation. However, the authors have not proven that these candidates are bonafide lncRNAs that do not produce small peptides. Furthermore, these potential candidates have not been studied in enough depth to be sure that the authors are observing a direct rather than indirect effect. The authors should focus their efforts on one or two candidates and provide significant mechanistic insights into their precise function.

1. Throughout the MS, the authors refer to transcripts as lncRNAs. What is the evidence that these are truly "non-coding"? For example, what is their codon-substitution frequency? Also, it would be good to see experimental evidence (e.g. in vitro translation) that significant ORFs of their key candidates are not translated.

   As mentioned above, we have added clarification on the use of the Coding Potential Calculator as an indication of the "non-coding" potential of the lncRNAs, which predicts with an accuracy of 95% (Kong et al., 2007). The lncRNAs represented on the microarray in this study were from a genome-wide identification of human long non-coding RNA genes (Jia et al., RNA 2010).). In the original publication, the properties of the lncRNA catalogue, including protein-coding status, were thoroughly examined.

2. The authors have not addressed the actual transcript size of each of these lncRNAs and their abundance in cells.

   Transcript sizes of lncRNAs are now presented in Tables II and III, while the abundance of lncRNAs in cells is addressed in Figures 3B and 6B.

3. All RNAs are purely based on GeneBank entries. There is no cloning and RACE data on any of the candidate RNAs to validate transcript structure, transcription start and transcription end sites. Existing RNA-Seq data (ENCODE project on the UCSC browser) supports that some of the transcripts (e.g. lncRNA_ES1; lncRNA_ES2) may in fact form part of longer transcripts.

   Many of the lncRNAs in this study have been successfully cloned in a previous study (Imanishi et al., PLoS Biol. 2004), and we have now clarified this in the text (p. 11). To validate transcript start and end sites, we performed and include the results of RNA-seq analyses on the lncRNAs in the relevant cell types (Supplementary Figures S4 and S6). We also acknowledge that some of these transcripts may form part of longer transcripts.

4. The data supporting a role for the transcription factor POU5F1 in regulating several of the lncRNAs is reasonable, however a role for NANOG in regulating lncRNA_ES2 is weak as knock-down of NANOG only resulted in a 20-30% down regulation of these lncRNAs.

   We agree the original results were less than convincing, therefore the NANOG RNAi experiment was repeated. The new results show that downregulation of specific lncRNAs upon NANOG RNAi was statistically ($p<0.05$) significant (Figure 3E).

5. In Figures 4 and 5 only some markers of a given lineage are up-/downregulated upon lncRNA candidate knock-down, while the other markers of the target lineage show hardly any change. This seems unexpected. How do lineage markers change in an Oct4 and / or Nanog control knock-down?

   We repeated the lncRNA RNAi experiment in ES cells, and performed a genome-wide expression profiling. We now show in Figure 4H that pluripotency markers were downregulated, and early lineage markers were upregulated, clearly confirming that there was

cellular differentiation. OCT4 and NANOG RNAi knockdowns were also performed in the same experiment, and the changes in expression of the lineage markers are shown in the same heatmap. Changes in expression of the markers for OCT4 RNAi and NANOG RNAi were similar to the knockdown of the lncRNAs.

6. The authors claim to achieve a larger effect when they K-D 4 lncRNA vs individual ones and suggest that there may be redundancy among the lncRNAs. However, there is no biological evidence to suggest that any of these lncRNAs function in a similar way. In fact their biological functions are unknown.

   This concern was also voiced by Reviewer 2 and we agree that it was premature to suggest redundancy among the lncRNAs. Therefore, we removed any discussion speculating redundancy of lncRNAs, and focused on their mechanistic functions instead.

7. In Figure 6 immunofluorescence staining for MAP2 in lncRNA knock-down cells over a large field of view is virtually abolished (panel A), but there is little or no change in the corresponding MAP2 transcript levels. How can this be explained?

   We repeated the experiment involving knockdown of the neuronal lncRNAs followed by differentiation and immunofluorescence (Figure 6C). We stained for both TUJ1 (marker or early neurons) and MAP2 (mature neurons) to ensure that lncRNAs blocked neurogenesis, rather than progression of early neurons to maturation. We used a different set of MAP2 primers for qPCR, and both TUJ1 and MAP2 transcript levels were decreased upon lncRNA knockdown (Figure 6I).

8. The authors point out (p.18) that lncRNA_N2 overlaps known miRNAs that were shown to function in neural differentiation. Are these miRNAs lost upon lncRNA_N2 knock-down, and can this explain claimed phenotypes?

   We performed the experiment to profile microRNA expression following lncRNA_N2 knockdown. This is shown in Figure 7B.

9. "OCT4 and NANOG binding sites are referred to near the transcription start sites of the lncRNAs" (p.13). 5' RACE needs to be done to determine/confirm the respective transcription start sites to make this claim.

   The OCT4 and NANOG binding sites were derived from a previous ChIP-seq study in human ESCs (Chia et al., Nature 2011). The transcription start sites (TSS) were defined by our new RNA-seq data, and the transcription factor binding sites are still near the newly defined TSS.

10. The authors state that some hESC colonies are morphologically different upon siRNA transfection targeting candidate lncRNAs (p15). This should be supported by providing quantitative data on the fraction of abnormal colonies in siRNA vs. control conditions. Similar quantitation should be provided for the number of colonies with down-regulated GFP (Oct4) reporter fluorescence (Figure 4A).

    We repeated the lncRNA knockdown experiment in hESCs, and quantified percentage of OCT4-expressing cells by FACS. The results are now shown in Figure 4B. These experiments also indicate that the ESC pluripotency is lost following siRNA knockdown of the indicated lncRNAs.

11. "Fewer and shorter neurite extensions" (p16) were observed in cells transfected with siRNAs against lncRNA candidates. However, there is no data presented to support this claim, and quantitation as in the point above, should be performed.

    Similarly, we repeated the lncRNA knockdown experiment in neural stem cells, and quantified percentage of $TUJ1^+$ cells, now shown in Figure 6D. Again, these experiments support and strengthen the original conclusions.

2nd Editorial Decision 08 November 2011

Your revised manuscript has now been re-assessed by one of the original referees with comments enclosed below. This scientist only asks for slight moderations and openly discussing some remaining gaps that demand further experimentation. I do leave such minor modifications up to your discretion and would be grateful for a modified word file as soon as possible to enable efficient formal acceptance and production.
Looking forward to your timely response.

Sincerely yours,

Editor
The EMBO Journal

REFEREE COMMENTS

Referee #2:

The authors have improved the manuscript in several ways to bolster the claim that lncRNAs are involved in maintenance of pluripotency of ES cells and neurogenesis. As this is one of the first evidence in human ESCs, this work is important and supports recent data from mouse ESCs suggesting a similar idea. The addition of RIP data indicating some of the candidate lncRNAs interact with Polycomb, Sox2, or REST is welcomed.

However, at points the claims are too strong. The authors do not define how said lncRNAs regulate Polycomb or Sox2 function, nor their epistatic relationships. The model showing co-binding of two proteins to one lncRNA is also not proven by data. As such, the authors should moderate their text and acknowledge gaps in the data.

