Translating embryology to embryonic stem cells (ESCs): mesendoderm (ME) specification

A number of diseases that represent a major cause of morbidity and mortality in the developed world are, at their root, diseases of cellular deficiency. Most of the cells that build our tissues and organs share a common origin during the early events of embryogenesis: they all derive from a transient structure known as the primitive streak (PS). The specification of the distinct PS populations is tightly regulated in a spatial and temporal manner, giving rise to both endoderm and mesoderm cells. ESCs have the potential to produce any of the three germ layer-derived cells and represent an ideal and relevant in vitro system to recapitulate the early embryological events. During their in vitro differentiation into mesoderm and endoderm, ESCs transition an intermediate stage called ME, which is equivalent to the PS (1-3). The gene regulatory networks (GRNs) that dictate ME specification are under the control of distinct combinations of transcription factors (TFs) that are necessary to elicit cell fate and lineage determination. Emerging evidence suggests that GRN activity is, in addition to TFs, under the control of complex regulatory non-coding RNA networks. A better understanding of the non-coding RNA networks implicated in ME specification is needed in order to provide important insights into the regulatory mechanisms operating in the early embryo and provide novel approaches to improve stem cell commitment for regenerative medicine.

Long noncoding RNAs (lncRNAs)

Among all the classes of non-coding RNAs, lncRNAs are operationally defined as RNA molecules that are longer than 200 nucleotides in length and exhibit minimal protein coding potential. Currently, the best characterized lncRNAs are typically polymerase II (Pol II) transcribed, multi-exonic, 5’ capped, alternatively spliced and polyadenylated. LncRNAs are usually expressed at a lower level than mRNAs but importantly are characterized by a greater tissue and cell specificity (4-7). Recent studies have shown that lncRNAs regulate various biological processes such as chromosome inactivation (8), imprinting (9), splicing (10) and transcriptional control (11). LncRNAs represent a very dynamic class of molecules that are emerging as key regulators of every aspect of GRN activity, including...
transcriptional control, post-transcriptional processing and chromatin remodeling (12). Recently, lncRNAs have been shown to be involved in the fine-tuning of GRNs controlling cell-fate and lineage specification. Here, we review the lncRNAs that have been implicated in regulating ME specification in vitro and/or in vivo.

**lncRNAs in ME specification**

The first lncRNA shown to be required for controlling ME specification, specifically into the cardiac fate, was *Braveheart (Bvht)* (13). Depletion of *Bvht* in mouse ESCs decreases their ability to activate a core GRN directing cardiovascular cell fate. This lncRNA acts upstream of the TF mesoderm posterior protein 1 (MESP1), master regulator of cardiac mesoderm specification. *Bvht* interacts with the epigenetic silencer polycomb repressive complex 2 (PRC2), possibly displacing PRC2 from the *MesP1* locus to activate cardiac mesoderm-specific genes. Very recently, the mechanism of action of *Bvht* has been characterized in detail. The secondary structure of this lncRNA was determined and the 5’ asymmetric G-rich internal loop (AGIL) was shown to be essential for *Bvht* function (14). Indeed, mouse ESCs carrying a specific deletion of AGIL showed impaired cardiogenic differentiation. Importantly, Bvht interacts through AGIL with CNBP, a zinc finger TF which acts as negative regulator of the cardiac gene program. Our group has characterized a lncRNA associated with the same locus controlling *Bvht* in mouse. This novel lncRNA called *CARMEN, (C)ardiac (M)esoderm (E)nhancer-associated (N)oncoding RNA, was identified as one of the most up-regulated lncRNAs during the cardiac differentiation of human cardiac progenitor cells (CPCs) isolated from the fetal human heart (15). As opposed to *Bvht*, *Carmen* is conserved in human and is associated to an active super-enhancer in both the fetal and human heart. *Carmen* loss of function (LoF) by shRNAi in mouse P19CL6s resulted in a decreased ability of the cells to differentiate into the cardiac lineage. Interestingly, *Carmen* depletion resulted in the down-regulation of *Eomesodermin (Eomes)*, a crucial TF known to dictate mesoderm and definitive endoderm (DE) specification. Thus, *Carmen* acts upstream of *Eomes* in the GRNs driving early ME specification.

lncRNA that are transcribed on the opposite strand to their neighboring protein coding gene are named divergent lncRNAs and comprise ~20% of the total lncRNAs in the mammalian genome. The first example of a divergent lncRNA implicated in early development was *Fendrr*, transcribed divergently from the TF forkhead box F1 (FOXF) (16). *Fendrr* is specifically expressed in the lateral plate mesoderm of the developing embryo and interacts with both activating (TrxG/MLL) and repressive (PCR2) chromatin modifying complexes to modulate expression of specific TFs implicated in the development of the heart. *Fendrr* KO resulted in embryonic lethality at embryonic day 13.75 with abdominal wall defects and pooling of blood in the right atrium. An important feature of *Fendrr* is its long-term effect. Epigenetic signatures, like the ones established via *Fendrr* activity, persist through several stages of differentiation. *Fendrr* is therefore a modulator of the epigenetic landscape during early development. The crossstalk between *Fendrr* and *Foxf1* represents an example of an intriguing transcriptional and functional coupling of a divergent lncRNA with its adjacent PCG, both of which are essential for development of the same embryonic tissue. Another example of a divergent lncRNA implicated in early ME specification is *Evsx1as*, transcribed on the opposite strand of its nearby gene Even-Skipped Homeobox 1 (*Evsx1*) (17). *Evsx1* is an important homeodomain TF that promotes ME specification during mouse ESC differentiation and its expression is highly correlated with its divergent lncRNA *Evsx1as*. Both *Evsx1as* genomic deletion and post-transcriptional knockdown significantly repressed *Evsx1* expression. To evaluate whether *Evsx1as* induced *Evsx1* transcription via a cis-regulatory mechanism, the authors artificially activated *Evsx1as* expression using a CRISPR-ON system. Consistent with a cis-regulatory mechanism, activation of *Evsx1as* induced *Evsx1* expression. Finally, using an elegant dCas9-mediated approach to tether the *Evsx1as* RNA to the promoter of *Evsx1as/Evsx1*, the authors demonstrated a direct role for *Evsx1as* RNA in controlling *Evsx1* transcription through *cis* regulation. Very interestingly, the depletion of *Evsx1* had a lesser impact on ME specification than the depletion of its divergent lncRNA, suggesting the possibility that *Evsx1as* acts also in trans independently of *Evsx1*. *DIGIT* represents an alternative interesting antisense lncRNA, divergent to *Goosecoid (GSC)*, an important regulator of ME specification. Knock-down of *DIGIT* in human ESCs impacted upon DE differentiation and expression of important developmental protein coding genes including *SRY-Box 17 (SOX17)* and *Forkhead Box A2 (FOXA2)*. *DIGIT* has an ortholog in mouse expressed in early development at E6.5, a stage corresponding to the formation of the anterior PS. Importantly, *Digit* expression peaks at E.7.5, when the DE is formed. Depletion of *DIGIT* via post-transcriptional approaches (shRNAs and LNAs)
and the insertion of a poly(A) termination sequences in human ESCs demonstrated that DIGIT RNA rather than the act of transcription at the DIGIT locus is essential for GSC expression. Importantly, ectopic re-expression of the DIGIT transcript in DIGIT-deficient cells induced GSC expression, as well as expression of other ME markers such as FOXA2 and SOX17. This suggests that DIGIT regulates GSC expression via a trans-regulatory mechanism as it does not need to be transcribed in close proximity to GSC to elicit its regulatory role.

Another lncRNA implicated in early development is Alien. Alien starts to be expressed at the PS stage, showing maximal expression in the lateral plate mesoderm of E8.5 mouse embryos (18). This lncRNA is proximal to Foxa2 and is conserved in human. Alien expression correlates with crucial ME genes such as Brachyury (T) and Mix Paired-Like Homeobox (Mix1). LoF of Alien is associated with severe defects in multiple anatomic structures, revealing its importance during ME specification. In another study, the same lncRNA (on this occasion named DEANR1) was shown to be an important regulator of Hox expression. Depletion of DEANR1 in human ESCs affects endoderm differentiation but is not required for maintaining the stemness features of ESCs. Interestingly, DEANR1 regulates FOXA2 expression by facilitating SMAD2/3 binding at the FOXA2 promoter.

Deng and colleagues identified and characterized HoxBlinc, a lncRNA transcribed in the Homeobox B (boxb) gene locus upstream of the boxb4 gene (19). HoxBlinc depletion during mouse ESCs differentiation affects boxb genes and other crucial regulators of ME specification such as Eomes and Mesp1. Using RNA immune precipitation (RIP) and chromatin isolation by RNA purification (ChIRP), the authors demonstrated that the HoxBlinc transcript binds to the promoters of the boxb1-boxb6 genes and recruits Set1/MLL1 complexes, driving H3K4me3 to the promoters of the boxb genes. Importantly, HoxBlinc RNA regulates the chromatin looping of the boxb gene cluster, locking the boxb locus in a permissive conformation for being actively transcribed and drive ME specification into the hematopoietic lineages.

Conclusions

The assumption that the GRNs controlling cell fate specification were primarily protein-based regulatory systems has been somewhat premature. We are now beginning to appreciate and decipher the complex roles of lncRNAs in contributing to the regulation of GRNs dictating ME specification. In the near future, we envisage that the manipulation of this new class of regulatory molecules may facilitate the development of more efficient cell differentiation protocols for regenerative medicine. Finally, understanding the roles of lncRNAs in development and cell differentiation is crucial to develop therapeutic approaches to treat human diseases for which lncRNA dysregulation is reported.

Acknowledgements

Financial support was granted by the Research Commission, Faculty of Biology and Medicine, University of Lausanne, Switzerland (T Pedrazzini). Additional funding was provided by the Pierre Mercier Foundation, Lausanne, Switzerland (S Ounzain).

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

References


Cite this article as: Alexanian M, Pedrazzini T, Ounzain S. New Lncs to mesendoderm specification. J Thorac Dis 2017;9(Suppl 1):S5-S8. doi: 10.21037/jtd.2017.03.70

© Journal of Thoracic Disease. All rights reserved.