A regulatory function of long non-coding RNAs in red blood cell development

Klaudia Kulczyńska and Mirosława Siatecka

Department of Genetics, Institute of Experimental Biology, Faculty of Biology, Adam Mickiewicz University in Poznan, Poznań, Poland

In recent years it has been discovered that long non-coding RNAs are important regulators in many biological processes. In this review, we summarize the role of lncRNA in erythropoiesis. LncRNA are crucial for regulation of gene expression during both, proliferation and differentiation stages of red blood cell development. Many are regulated by erythroid-specific transcription factors and some are expressed in a developmental stage-specific manner. The majority of individually studied IncRNAs are involved in regulating the terminal maturation stages of red cell differentiation. Their regulatory function is accomplished by various mechanisms, including direct regulation in cis or trans by the IncRNA product or by the cis-localized presence of the IncRNA transcription itself. These add additional levels of regulation of gene expression during erythropoiesis.

Key words: long non-coding RNA, erythropoiesis, red blood cells

Received: 03 June, 2016; revised: 24 July, 2016; accepted: 15 August, 2016; available on-line: 17 November, 2016

INTRODUCTION

Erythropoiesis is a process characterized by commitment of pluripotent hematopoietic stem cells to the erythroid lineage. During mammalian development, erythropoiesis occurs successively in the yolk sac, the fetal liver and the bone marrow (Barker, 1968). Erythrocytes are short-lived, continuously replenished cells. Erythropoiesis comprises several developmental stages, during which hematopoietic stem cells are committed to progenitor and precursor cells with gradually restricted potential. In erythroid lineage development, a bipotential megakaryocytic-erythroid progenitor gives rise to a burst-forming unit-erythroid (BFU-E), which differentiates into a colony-forming unit-erythroid (CFU-E) and finally to erythroblasts. In hematopoietic tissue, these cells differentiate in the context of a specialized niche, the erythroblastic island, where erythroblasts are attached in concentric rings to one or more central macrophages. At this stage, cells decrease their size, nucleus is condensed and ultimately is expelled; cells become reticulocytes. The final stage of erythroid differentiation involves maturation of the reticulocytes into circulating, mature red blood cells (RBC). The reticulocytes dismantle their ribosomal machinery, expel organelles, and assume a biconcave discoid shape. (Manwani & Bieker, 2008; An & Mohandas, 2011; Pals, 2014).

To ensure a continuous and controlled production of red cells, this process has to be tightly regulated. Fundamental to the homeostasis of the hematopoietic system is the correct balance of progenitor cell proliferation versus lineage committed differentiation (Orkin & Zon, 2008). Growth factors play a critical role in preventing apoptosis and in inducing proliferation, while successive combination of a complex network of transcription factors drives differentiation (Novershtern et al., 2011).

In recent years a new class of small, endogenous non-coding RNAs (ncRNAs) emerged as important regulators of gene expression at the post-transcriptional level. MicroRNAs are the best-known family of ncRNAs. These molecules are 19–23 nucleotides long and bind to specific sites within the 3’-UTR of regulated transcripts. MicroRNAs can decrease gene expression of various mRNAs by either inhibiting translation or directly causing degradation of the transcript. MicroRNAs demonstrate evolutionary conservation as well as tissue and developmental stage specific expression patterns (Wienholds & Plasterk, 2005). In 2005, Lu and coworkers found for the first time that miRNAs expression changes during erythropoiesis (Lu et al., 2005; for reviewed see: Lawrie, 2009; Listowski et al., 2013; Zhao et al., 2010).

This review focuses on long non-coding RNAs (lncRNAs) and their involvement in regulation of red blood cell development, particularly the late stages of adult erythropoiesis that comprise maturation of erythrocytes characterized by hemoglobinization, cell size reduction and extrusion of nucleus.

LncRNAs belong to a novel heterogeneous class of ncRNAs that includes thousands of different species identified by a high-throughput sequencing technologies that allow for the sequencing of the genome and transcriptome at an unprecedented depth (Paralkar & Weiss, 2011; Alvarez-Dominguez et al., 2014; Paralkar et al., 2014). LncRNAs are longer than 200 nt, often polyadenylated and devoid of evident ORFs (Wilusz et al., 2009; Kung et al., 2013). LncRNAs can be classified based on genomic location relative to the well-established markers, such as protein-coding genes. Six classes can be distinguished (Table 1). The first class is located intergenically and does not overlap with any protein coding genes. The second class consists of lincRNAs situated within intronic regions (lincRNA). In the third class, lincRNAs are transcribed from the antisense strand (alscRNA). The fourth class consists of enhancer lncRNAs (elncRNA), which are expressed from active enhancers. There is also a class named shlncRNA, where lncRNA hosts small RNA (sRNA) - microRNAs. A final class comprises pseudogene lncRNAs (pIcnRNA) (Alvarez-Dominguez et al., 2014; Rinn & Chang, 2012). Unlike small ncRNAs,
lncRNAs can fold into higher ordered structures to provide greater potential and versatility for target recognition (Batista & Chang, 2013; Rinn & Chang, 2012; Guttman & Rinn, 2012). lncRNAs affect many biological processes, such as genomic imprinting, gene dosage compensation, gene expression, and nuclear organization. Several models have been proposed for the function of the lncRNAs: (1) as regulatory signals, (2) as a decoy system, (3) as guides of regulatory components to the genomic targets, and (4) as a ribonucleoprotein scaffold. The evidence indicates that lncRNAs exert their function mostly through modulation of chromatin-remodeling complexes (Wang & Chang, 2011; Guttman & Rinn, 2012; Rinn & Chang, 2012; Arriaga-Canon et al., 2014).

Two major approaches have been applied to study involvement of lncRNA in erythropoiesis: 1) global analysis of lncRNA expression throughout RBC development using recent technologies, i.e., microarrays and next generation sequencing, 2) direct functional tests of various individual lncRNAs.

### GLO BAL ANALYSIS AND CHANGES IN lncRNA EXPRESSION DURING ERYTHROPOIESIS

In the last couple of years several research groups performed a vast analysis of long non-coding RNAs involved in red blood cell development. Owing to newly developed technologies, such as high-throughput next generation sequencing and computational methods, they obtained comprehensive data sets from RNA-seq experiments. Diverse sources of erythroid cells were used for analyses covering different stages of red blood cell development, different source organs, and also various species. Researchers compared transcriptomes obtained from bipotential megakaryocytic-erythroid progenitors to lineage-committed megakaryocytes and erythroblasts using murine embryonic day 14.5 (E14.5) fetal liver and bone marrow cells. They identified 1109 potential lncRNA genes (including 683 transcribed in erythroblasts); around half of them were not annotated (Paralkar et al., 2014).

There were also studies of lncRNAs related to particular stages of erythroid development. Transcriptomes from BFU-E and CFU-E progenitors purified from mouse fetal liver along with differentiated TER119 positive erythroblasts were compared. As a result, 96 lncRNAs were identified that are differentially expressed during erythropoiesis (Alvarez-Dominguez et al., 2014), thus showing a dynamic regulation of lncRNAs expression during erythroid maturation. In addition, separate RNA-seq analyses were done for the TER119 positive cells comparing poly(A)+ and poly(A)− transcripts. In total, 9512 coding messenger RNAs (mRNA) genes and 655 lncRNA genes were identified (Alvarez-Dominguez et al., 2014). lncRNA included genes of all types of genomic location: intergenic, antisense, intronic, and enhancer loci, but also pseudogenes (Alvarez-Dominguez et al., 2014).

In general, these studies revealed that mRNAs were expressed at higher levels than lncRNAs, but lncRNAs were more developmental stage-restricted than coding mRNAs (Alvarez-Dominguez et al., 2014). ElncRNA-EC1, lncRNA-EC9, and alncRNA-EC3 are expressed in erythroblasts but not in the closely related megakaryocyte or megakaryocyte-erythroid progenitors (Alvarez-Dominguez et al., 2014).

---

**Table 1. Classification of lncRNAs based on their genomic location (taken from (Alvarez-Dominguez et al., 2014)) and examples of lncRNA involved in erythropoiesis.**

<table>
<thead>
<tr>
<th>Classes of lncRNA</th>
<th>mRNA</th>
<th>lncRNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intergenic lncRNA (lincRNA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intronic overlapping lncRNA (ilncRNA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antisense lncRNA (alncRNA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enhancer lncRNA (elncRNA)</td>
<td>Enhancer</td>
<td></td>
</tr>
<tr>
<td>sRNA-host lncRNA (shlncRNA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudogene lncRNA (plncRNA)</td>
<td></td>
<td>Pseudogene</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IncRNAs - role in erythropoiesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linc RNA-EPS</td>
</tr>
<tr>
<td>IncRNA-aGT</td>
</tr>
<tr>
<td>Upregulation of adult α-globin</td>
</tr>
<tr>
<td>IncRNA Fas-antisense 1 (IncRNA-Saf)</td>
</tr>
<tr>
<td>AlncRNA-EC7</td>
</tr>
<tr>
<td>ElncRNA-EC3</td>
</tr>
<tr>
<td>ShlncRNA-EC6/DLEU2 (host to microRNAs 15a and 16-1)</td>
</tr>
<tr>
<td>Downregulation of Rac1-PIP5K</td>
</tr>
</tbody>
</table>
Paralkar and coworkers compared IncRNA expression between fetal and adult erythropoiesis. More than 85% of fetal liver erythroid IncRNAs were detected in adult erythroblasts implying that most of mouse erythroid IncRNAs are expressed in both, fetal liver and adult bone marrow erythroblasts (Paralkar et al., 2014). However, Alvarez-Dominguez et al. showed that IncRNAs were expressed at different levels in mouse fetal and adult erythroblasts (Alvarez-Dominguez et al., 2014).

IncRNAs expressed in multiple species are less conserved in primary nucleotide sequence than coding genes (Paralkar et al., 2014). Paralkar et al. identified human orthologous regions for 95% of all transcribed mouse genes. Approximately 85% of coding genes expressed in mouse erythroblasts were also expressed in humans, but only 20% of erythroid IncRNA genes expressed in mouse erythroblasts were expressed in matched human samples. Similar results were obtained from an opposite analysis revealing that most mouse erythroid IncRNA genes are not transcribed in human erythroblasts. Only 15% of mouse IncRNAs are expressed in humans and vice versa, reflecting a dramatic species specificity (Paralkar et al., 2014). Most mouse erythro-megakaryocytic IncRNAs are transcribed from conventional gene promoters regulated by key hematopoietic transcription factors. Around 75% of erythro-megakaryocytic IncRNAs are transcribed from regions with promoter-like signatures and 25% from enhancer-like signatures (Paralkar et al., 2014).

In erythroblasts, 60 of 96 differentially expressed IncRNAs are bound at their promoters by erythroid specific transcription factors such as GATA, TAL1, or KLF1 (Alvarez-Dominguez et al., 2014). Transcriptional binding for these factors coincides with DNase I hypersensitive sites, with RNA pol II binding and active epigenetic chromatin marks. As a consequence of regulation by cell-type specific transcription factors, IncRNA expression can be highly developmental stage-specific (Alvarez-Dominguez et al., 2014).

Sun and coworkers conducted RNA-seq studies comparing transcriptomes obtained from embryonic fetal liver tissues from wild-type and Klf1 knockout mice (Sun et al., 2012). Klf1 is a gene encoding an essential erythroid transcription factor (Miller & Bieker, 1993; Siatecka et al., 2011). Klf1 knockout mice die from anemia by day E15, with severe defects in differentiation, hemoglobinization, enucleation, and membrane cytoskeleton organization of red blood cells (Nuez et al., 1995; Perkins et al., 1995). Transcriptome analyses identified 13 novel IncRNAs that showed significantly different expression between the wild-type and Klf1 knockout conditions (Sun et al., 2012). Ten IncRNAs were repressed versus three activated after Klf1 was knocked out. As described below, these may play functional roles in the development of erythroid cells.

**LncRNA-EPS** — lncRNA-EPS (erythroid gosurvival) was identified during erythroid differentiation of mouse fetal liver progenitors. LncRNA-EPS is located intergenically. It is 2531 nt long, consists of 4 exons and 3 introns and has a 5'end cap structure and a 3' poly(A) tail. LncRNA-EPS is enriched in hematopoietic organs, such as spleen, bone marrow and fetal liver cells. It is strongly induced during the transition from erythroid precursors CFU-Es to hemoglobin synthesizing TER119+ cells (Hu et al., 2011; Paralkar & Weiss, 2011). The knockdown of LncRNA-EPS by shRNAs significantly reduced the accumulation of erythroid cells by inducing massive apoptosis, as revealed by cell cycle analysis. A significant fraction of analyzed cells was located in the sub-G1 population representing cells that underwent apoptosis and/or necrosis. Conversely, ectopic expression of LncRNA-EPS resulted in an increased number of cells localized in the S and G2/M phases, indicating that overexpression of LncRNA-EPS protects erythroblasts from apoptosis. The time frame of LncRNA-EPS induction correlates well with the time window at which erythropoietin (Epo) exerts its biological function. This suggests that LncRNA-EPS’s anti-apoptotic ability could contribute to cell survival mediated by Epo (Hu et al., 2011).

Hu and coworkers found evidence indicating that Pycard, a signaling molecule that promotes cell death by activating caspases (Ohtsuka et al., 2004), is a target of LncRNA-EPS. Ectopic expression of LncRNA-EPS resulted in a dramatic repression of Pycard (Hu et al., 2011). During normal erythropoiesis, the expression of Pycard is inversely correlated to LncRNA-EPS (Hu et al., 2011). Moreover, overexpression of Pycard inhibits proliferation of erythroid cells, promotes their apoptosis and interferes with their terminal differentiation and enucleation. The exact repression mechanism has not yet been identified (Hu et al., 2011; Paralkar & Weiss, 2011).

Summarizing, the results obtained in the described research indicate that LncRNA-EPS modulates apoptosis at least in part through repressing Pycard expression. This pathway of regulation is required for the proper generation of mature red blood cells in response to various physiological and pathological stimuli.

**ShlncRNA-EC6**, also called DLEU2, is localized at chromosome 14. It belongs to IncRNA transcripts hosting microRNAs. ShlncRNA-EC6 hosts microRNAs 15a and 16-1. However, it has been suggested that the function of DLEU2 is independent of microRNA generation (Lerner et al., 2009; Klein et al., 2010; Alvarez-Dominguez et al., 2014). ShlncRNA-EC6 knockout or ectopic expression shows a stronger phenotype compared to miR-15a/16-1 knockout or misexpression. ShlncRNA-EC6 is broadly expressed. In erythroblasts, it is transcribed from a different specific promoter than in other cell types (Alvarez-Dominguez et al., 2014). Devoid of poly(A) tail shlncRNA-EC6/DLEU2 is predominantly induced during erythropoiesis. It promotes red blood cell maturation to least in part by cis-acting, IncRNA-directed control of expression of neighboring genes (Alvarez-Dominguez et al., 2014). Inhibition of DLEU2 caused up-regulation of SPRYD7/CLLD6, residing 45 kb away. No function is known for the SPRYD7 protein, although an RNA binding role has been proposed (Ponting et al., 1997).

Another study regarding shlnc-EC6 revealed its connection with the Rac1-PIPK5 pathway. Knockdown of shlnc-EC6 in purified mouse fetal liver erythroid progenitors significantly blocked erythroid enucleation, which led to a significantly upregulated expression of Rac1 (Wang et al., 2015). Rac1 is a GTPase, and it has been

**THE ROLE OF INDIVIDUAL IncRNAs IN ERYTHROID MATURATION**

Thus far, the majority of individually studied IncRNAs are involved in regulation of the terminal maturation stage in red cell development. The hallmarks of this stage of erythropoiesis are: expression of the TER119 marker, cell size reduction, progressive nuclear condensation and subsequent enucleation. Functions of IncRNAs were investigated through the knockdown procedure using shRNA.
K. Kulczyńska and M. Siatecka reported that deregulation of Rac GTPase during the late stage of erythropoiesis blocks enucleation of cultured mouse fetal erythroblasts without affecting their proliferation or differentiation (Ji & Lodish, 2010). Shlnc-EC6 negatively regulates Rac1 at the posttranscriptional level via specific binding to sites within the 3' UTR of Rac1 mRNA (Wang et al., 2015).

Consistently, overexpression of Rac1 and subsequent upregulation of its downstream component PIP5K strongly inhibited erythroid enucleation, which resembled the inhibitory effects of shlnc-EC6 knockdown (Villamizar et al., 2016). These results suggest that high expression of shlnc-EC6 at late-stage of red cell development helps erythroblasts to remove their nuclei through the Rac1-PIP5K pathway (Wang et al., 2015).

LncRNA Fas-antisense 1 (Fas-AS1 or Saf) is encoded on the antisense strand of the first intron of the human Fas receptor gene (also called APO-1 or CD95) on chromosome 10 (Yan, 2005). LncRNA Saf takes part in the maintenance of RBC production. Its promoter contains binding sites for the essential erythroid transcription factors GATA-1 and KLF1, as well as for NF-κB. Experimental data support the significance of these sites for LncRNA Saf transcription. During the early stages of erythroblast expansion, NF-κB signaling is involved in the repression of LncRNA Saf expression. Subsequently, at late stages of RBC maturation, expression of LncRNA Saf increases and it coincides with elevated expression of GATA-1 and KLF1, suggesting that LncRNA Saf could be regulated by these factors (Villamizar et al., 2016).

A cell culture model of human erythropoiesis revealed that induction of LncRNA Saf confers resistance to Fas-mediated cell death. LncRNA Saf interacts with Fas premRNA and human splicing factor 45 (SPF45). It facilitates splicing and production of a soluble Fas protein that protects cells against Fas-mediated apoptosis (Villamizar et al., 2016). Overexpression of LncRNA Saf in erythroblasts derived from CD34+ hematopoietic stem/progenitor cells reduced surface levels of Fas receptor and protected cells from Fas-mediated apoptosis signals. LncRNA Saf expression may therefore provide a means to regulate cell death during erythroid maturation (Villamizar et al., 2016).

AlncRNA-EC7 is an enhancer transcript that regulates expression of SLC4A1 gene encoding Band 3 protein. Band 3 is a structural component of the erythrocyte cell membrane and the primary anion exchanger responsible for mediating the exchange of chloride (Cl−) with bicarbonate (HCO3−) across plasma membrane (Alper, 2009). AlncRNA-EC7 is located 10 kb upstream from SLC4A1 gene locus and is involved in regulation of its expression by a cis mode of action. Experimental data suggested a model for looping of the alncRNA-EC7 enhancer to SLC4A1 gene locus with subsequent activation of Band 3 expression in erythroid cells (Fig. 1). Knockdown of alncRNA-EC7 was associated with an 80% decrease in Band 3 gene mRNA expression and severely impaired erythrocyte maturation, inhibiting cell size reduction and subsequent enucleation (Alvarez-Dominguez et al., 2014). As SLC4A1 gene is mutated in hereditary hemolytic anemias, this data predicts that alncRNA-EC7 is a novel disease-relevant locus (Sánchez-López et al., 2010).

ElncRNA-EC3 belongs to the enhancer class of LncRNA and is transcribed from an erythroid-restricted enhancer co-bound at multiple sites by GATA1 and TAL1. Activation of elncRNA-EC3 in erythroblasts coincides with a more than two-fold up-regulation of its neighbor KIF2A gene that is located 40 kb away. Thus, elncRNA involved in erythropoiesis.

Figure 1. Modes of action of LncRNA involved in erythropoiesis. Regulation by LncRNA product: a) upregulation of a cis target gene expression by induction of chromosomal looping or b) downregulation by posttranscriptional inhibition via specific binding of LncRNA in trans to sites within the 3' UTR of the targeted mRNA. Regulation by LncRNA transcription process. The LncRNA gene overlaps with the targeted genes (α type globin genes; Gavrilov & Razin, 2008) located on the same strand of DNA. Transcription of LncRNA recruits the epigenetic machinery to the chromatin which permits opening of its structure and allows for the transcription of the target genes (as delineated) along the same locus. P, promoter; LCR, locus control region.
CONCLUSIONS

Many studies have recently identified and discussed the role of various lncRNAs that are involved in red blood cell development. LncRNAs add an additional level of regulation of gene expression. As opposed to microRNAs, their mechanism of action is more complex. Based on examples described above, modes of action of lncRNAs involved in erythropoiesis can be divided into two categories (Fig. 1): 1) regulation (up or down) by lncRNA product and 2) regulation by lncRNA transcription. These additional components that are important for erythroid regulation may also contribute to disease, but they may also provide new opportunities to use them as diagnostic markers and therapeutic targets for treatment of erythroid disorders.

Acknowledgements

The authors thank Dr. James Bieker from the Icahn School of Medicine at Mt.Sinai in New York, Dr. Barbara Imioleckzy from the Institute of Bioorganic Chemistry Polish Academy of Sciences and Dr. Karol Szkaradkiewicz from the Adam Mickiewicz University for helpful discussions and critical comments on this manuscript.

Acknowledgements of financial support

This work was supported by the OPUS grant 2013/09/B/NZ1/01879 of the Polish National Science Center, and the KNOW RNA Research Centre in Poznań (No. 01/KNOW2/2014).

REFERENCES
