

REVIEW ARTICLE

A link between long non-coding RNA (lncRNA) and thalassaemia: A review

Afshan SUMERA^{1*}, Ammu K RADHAKRISHNAN², Abdul AZIZ BABA¹, Elizabeth GEORGE³

¹School of Medicine, International Medical University, Bukit Jalil, 57000 Kuala Lumpur, Malaysia.

²Jeffrey Cheah School of Medicine and Health Sciences, Monash University Malaysia, 47500 Bandar Sunway, Malaysia. ³Assunta Hospital, Petaling Jaya, Selangor, Malaysia.

Abstract

The long non-coding RNAs (lncRNAs) are the most prevalent and functionally diverse member of the non-coding RNA (ncRNA). The lncRNA has previously been considered to be a form of transcriptional “noise” but recent studies have found that the lncRNA to be associated with various disease conditions. It has also been found to play important roles in various physiological processes such as haemopoiesis, where lncRNA is reported to act as a fine-tuner of this very important process. To date, the effects of dysregulated lncRNA in thalassaemia has not been fully explored. This review article focuses on the possible roles of dysregulated lncRNAs in the pathogenesis of thalassaemia.

Keywords: Thalassaemia, long non-coding RNA (lncRNA), erythropoiesis

INTRODUCTION

Thalassaemia is a progressive global disorder due to single gene defects. It is expected that there would be around 900,000 births of clinically significant thalassaemia disorders in the next 20 years.¹ Thalassaemia is a quantitative disease, where the production of globin chains that are required to produce haemoglobin is decreased due to mutations in the human *globin* gene. Mutated genes fail to produce adequate amounts of haemoglobin, which causes the patient to present with several clinical symptoms related to lack of oxygen in the body. There are many types of thalassaemia depending on the site of mutation in the human *globin* gene. More recently, thalassaemic individuals are classified into two categories, which are based on the need for blood transfusion, i.e. (i) transfusion-dependent and (ii) transfusion independent.¹ At present, a diagnosis of thalassaemia is made by clinical history, physical examination, complete blood count (CBC), haemoglobin (Hb) analysis by electrophoresis or high-performance liquid chromatography (HPLC). However, there is a need to improve the detection and diagnosis of thalassaemia. One approach that could be used to achieve this would be to study the irregular transcription patterns in erythrocytes from

thalassaemic individuals, which could help the clinicians and scientist to understand better the pathogenesis of this disease, which is associated with defects in DNA sequences.² This review article focuses on possible roles of dysregulated long-chain non-coding RNAs (lncRNAs) in the pathogenesis of thalassaemia.

RNA types and their function

In the early 1970s, it was discovered that the genome contains protein-coding micro-molecules known as ribonucleic acid (RNA), which are single-stranded biopolymers that form an essential element in all living cells.³ The RNA carries genetic information from the DNA stored in the nucleus of a cell, which is then translated into proteins. Initially, there were only three types of RNA were known; i.e. (i) messenger RNA (mRNA); (ii) transfer RNA (tRNA); and (iii) ribosomal RNA (rRNA).³ The tRNA and rRNA are present in an organelle found in the cytoplasm of cells, known as the ribosome. The tRNA serves as an adapter molecule as such has two functional ends. One end reads and interacts with the triplet code present in the mRNA while the other end interacts with the right amino acid. The main function of the rRNA is to attach new amino acids to the peptide chain that is

*Address for correspondence: Afshan Sumera, School of Medicine, International Medical University, Bukit Jalil, 57000 Kuala Lumpur Malaysia. Tel: +60327317264 Fax: +60386567229. Email: AfshanSumera@imu.edu.my

being newly synthesised.⁴ The mRNA carries a sequence, which serves as a template to encode amino acid sequences that form a peptide chain, required to form a protein.

The pre-mRNA or unprocessed mRNA, which is known as the “heterogeneous nuclear RNAs” (hnRNAs) changes into mature mRNA after extensive processing.⁵ Previously it was widely reported that half of hnRNAs are limited to the nucleus and do not contain coding sequences.⁶⁻⁸ However, in 1977, it was discovered that introns, represents just a small portion of the non-coding sequences.^{9,10} Soon after, in the 1980s, small nuclear RNA (snRNA) and small nucleolar RNAs (snoRNAs) were identified as essential players in RNA post-transcriptional processing.¹¹ The snRNA can only be found in eukaryotes.¹²

Non-coding RNA

Approximately two-thirds of the human genome is reported to be composed of “junk” or non-coding DNA^{13,14}, which includes transposons, pseudogenes etc. and only the remaining one-third make up the protein-coding regions that range from 20,000-30,000 genes.¹⁵ In addition, there is a heterogeneous class of non-coding RNA (ncRNA), which forms the vast majority of the human genome was considered as “junk” or “waste” matter because these were not directly involved in the translation process.¹¹ What these findings mean is that only a small portion of the human genome, which is transcribed into mRNA results in protein-encoding.

There are two major types of ncRNA, which are (i) small RNA (sRNA) and (ii) long non-coding RNAs (lncRNAs).¹¹ The lncRNAs are reported to be the most prevalent and functionally diverse member of the ncRNA.^{11,16,17} At present, there is no single definition for lncRNAs based on their biological functions. The most widely

used definition of lncRNA is based on their location and size; where an ncRNA that is more-than or equal-to (\geq) 200 nucleotides of length is known as lncRNA while those that are less than 200 nucleotides are defined as sRNA.¹⁸ As shown in Table 1, the lncRNA can be classified into five mutually non-exclusive groups based on their genomic position while the sRNA can be classified into three groups. One way of differentiating lncRNA from sRNA is based on their function as a primary or spliced transcript.¹⁷ The lncRNA is reported to be present in almost all species including yeast¹⁹, prokaryotes²⁰, viruses²¹, plants²² and animals.^{23,24}

What is long non-coding RNA?

The lncRNAs are reported to be inadequately moderated among various species when compared to other well-studied RNAs^{25,26}, which raises the question of whether the lncRNA has functional roles in all species; or does it have species-particular attributes. Due to its’ weak expression, some researchers have proposed lncRNA to be transcriptional noise.²⁷⁻²⁹ However, to date, there are several reports on their biological role in processes that range from transcription to translation²⁷⁻³⁵, protein localisation^{36,37}, cellular structure integrity^{38,39}, imprinting⁴⁰⁻⁴², cell cycle^{43,44}, apoptosis^{21,45}, stem cell pluripotency⁴⁶, stem cell reprogramming⁴⁷ and heat shock response.^{48,49}

In the recent year, studies on lncRNAs appear to be gaining momentum; especially their role in diseases that occur following their up or down-regulation. Many of these studies have provided some insights to understand the biological roles of lncRNA. In the early 1970s, the lncRNA was considered to be “junk”.¹⁶ This concept was found to be untrue as more studies were carried out to investigate the role of lncRNA. For instance,

TABLE 1: Classification of lncRNA and sRNA

lncRNA	sRNA
1. Stand-alone lncRNAs	1. micro-RNA (miRNA)
2. Natural antisense transcripts	2. small nucleolar RNA (snoRNAs)
3. Pseudogenes	3. Piwi-interacting RNAs (piRNAs) for silencing of genetic material
4. Long intronic ncRNAs	
5. Divergent transcripts, promoter-associated transcripts, and enhancer RNAs	

lncRNA: long non-coding RNA; sRNA: small RNA; RNA: ribonucleic acid; [Adapted from]¹⁶

in the early 1990s, an association of lncRNA and gene expression was reported in H19⁵⁰ and Xist.^{23,51} More recent studies using advanced technologies such as microarray hybridisation and deep sequencing have found that these non-coding sequences to be an important part of the genome that are transcribed at some point during development.⁵²⁻⁵⁴ So, what are the known biological roles of lncRNA?

Role of lncRNA in cellular processes

To date, there is not much known about the biological functions of lncRNAs. However, their role in various cellular processes cannot be denied. For instance, lncRNA can provide potential and target recognition by folding itself into higher-order structures.⁵⁵⁻⁵⁷ Other important roles of lncRNA can be seen in gene expression, genomic imprinting, nuclear organisation and gene dosage compensation.⁵⁸ Some studies have shown that the lncRNA can act as recruiters, scaffolders and tethers in gene expression.^{59,60} Genome-wide studies show that lncRNA can recruit chromatin-modifying factors such as polycomb repressive complex 2 (PRC2), which are needed for regulation of chromatin states by interaction.⁶¹ Another important role is to silence ribosomal DNA (rDNA) by directing the methylation process.⁶²

In the genome, lncRNA are transcribed from a single locus, and they can control allele and locus in cis.⁶³ During transcription, lncRNA attaches to chromatin via DNA:RNA hybridisation and can act as cis-acting molecular tethers.⁶³ In addition, during transcription, lncRNA can function as decoyers, inhibitors and co-regulators.⁶³ The lncRNA can also inhibit polymerase II (Pol II) directly by inhibiting the activity of the dihydrofolate reductase (DHFR) during transcription.⁶⁴ Alternatively, the lncRNA could decoy transcription factors, co-regulate or co-activated several nuclear steroid receptors.⁶⁵ Another role for lncRNA is in the regulation of the nuclear compartment by interacting with nuclear bodies.⁶⁶ The lncRNAs can also affect mRNA destabilisation through interfering the micro-RNA (miRNA)-mediated process by hindering spliceosome formation or masking the binding sites miRNA.⁶⁷

Role of lncRNA in Haemopoiesis

Haemopoiesis is an important process that must happen throughout the life of a human. This is a life-long process that continuously produces new blood cells. One of the important regulators

of this process is a transcription factor known as PU.1, which is considered to be a master haemopoietic regulator.⁶⁸ This protein directs myeloid and lymphoid stem cell differentiation and suppresses leukaemic proliferation. One of the lncRNA, known as anti-sense to PU.1 is reported to affect the expression of this protein.²⁷ The presence of this lncRNA can result in haematological malignancies as it has been found to negatively regulate the translation of the PU.1 mRNA.⁵⁶

The lncRNA can also act as fine-tuners of cellular processes at various stages of cell differentiation and lineage commitment; thereby regulating the fate of a cell by the positive or negative feedback loop.^{27,69-71} For instance, the *lncRNAES1* and *lncRNAES2* are known as stem cell pluripotency-associated lncRNAs as these lncRNAs can silence the expression of the *Sox2* gene in human embryonic stem cells (ESC).⁷² The *Sox2* gene is essential for ESC to maintain the pluripotent state.⁷² Silencing of the *Sox2* gene is shown to compromise the pluripotent state of both mouse and human ESC.⁷² This is one of the ways that lncRNA may be able to control the fate of the cell i.e. by deciding between cell differentiation and proliferation.

One of the important lncRNA that has been studied in haematopoietic stem cells is the lncRNA-EPS (erythroid pro-survival).⁷³ This lncRNA can activate transcription by indirect repression of apoptosis-associated speck-like protein (ASC) (also known as Pycard), which is a key modulator of caspase-mediated apoptosis.⁷⁴ Indirect repression is achieved by producing transcriptional repressors.⁷⁵ Over-expression of Pycard will inhibit proliferation of erythroid cells; thereby promoting apoptosis in these cells as well as interfere with their terminal differentiation and enucleation.⁵⁶ The exact mechanism of repression has not yet been elucidated.

The lncRNAs have also been implicated to play a role in lymphopoiesis, where these molecules facilitate development and function of adaptive immunity by modulating differentiation of memory T-cell and lymphocyte activation.⁷⁶

Role of lncRNA in Erythropoiesis

Haemopoiesis is a well-regulated complicated process. This process starts when the haematopoietic stem cells (HSC) in the bone marrow are stimulated with various growth factors to differentiate into the different types of mature blood cells. Several studies have found

that differentiation of erythroid progenitors and megakaryocytes start early in the hierarchy from HSC compared to other cells.^{77,78}

Erythropoiesis starts with early committed precursors, which are the burst-forming unit-erythroid (BFU-E) and colony-forming unit-erythroid (CFU-E) cells. Following this, these cells will be differentiated into erythroid precursors that possess more distinct morphology. The different stages of erythroid differentiation include pro-erythroblast, basophilic, polychromatophilic and orthochromatic erythroblast. The nucleus ejects at orthochromatic erythroblast and changes into reticulocyte to accommodate more Hb within the cell.⁷⁹

To date, approximately 1100 potential lncRNA genes, which may be involved in the regulation of erythropoiesis have been identified.⁸⁰ In adults, most of these lncRNAs are particularly seen in the late stages of this process, which range from accommodation of Hb within a cell to ejection of the nucleus. Among these lncRNAs, a total of 96 are found to be exclusively expressed during erythropoiesis.⁸¹ Most of these lncRNAs are associated with regulation of final stages of erythropoiesis.

All the different types of lncRNA such as intergenic, antisense, intronic, and enhancer loci, pseudogenes have been reported to play a role in erythropoiesis.⁸¹ However, the level of expression of the lncRNA in these cells varies according to different studies.^{80,81} Majority of foetal liver erythroid lncRNAs were expressed in adult human erythroblasts and also in foetal and adult mouse erythroblast.^{78,81} The lncRNA binds to the promoter region by GATA and

other transcription factors in erythroblast and this causes developmental stage-specific expression.⁸¹ The importance of inter-species orthologous region expression of lncRNA has also been studied. Researchers found that almost 85% of coding genes expressed in mouse were also expressed in human erythroblast.⁸⁰ However, the expression of lncRNA was found to be strictly species-specific because only 15% of mouse lncRNA was observed in humans.⁸⁰

The lncRNA-EPS is a 2531 nucleotides lncRNA, which is expressed during differentiation of foetal erythroblast in mouse.⁷⁴ The lncRNA-EPS is located intergenically in the spleen, bone marrow and foetal liver.⁷⁴ Its expression is observed during the transition of CFU-Es to haemoglobin synthesising cells.⁸² In normal erythropoiesis, the levels of lncRNA-EPS are usually in an inverse relationship with Pycard, which allow cells to differentiate more.⁸³ However, lncRNA-EPS can control apoptosis in an erythroid generation. Downregulation of lncRNA-EPS results in the increased apoptosis of erythroid progenitors by targeting Pycard, which significantly reduce the number of cells (Figure 1).⁸² Furthermore, knockdown of lncRNA-EPS resulted in higher levels of Pycard, thereby favouring cell death through unknown mechanisms.⁸² These findings strongly suggest that lncRNA-EPS nurtures erythropoiesis by inhibiting apoptosis (Fig. 1a and b).

Another important lncRNA expressed in erythropoiesis is the lncRNA Fas-antisense 1 (Fas-AS1 or Saf), which is needed for the maintenance of erythrocyte production⁸⁴. The GATA-1 binds to the promoter region of lncRNA

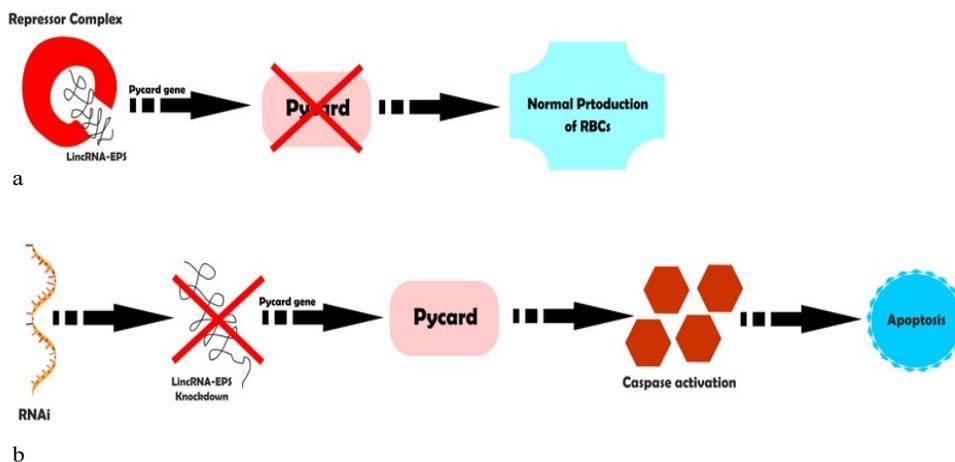


FIG.1: LincRNA-EPS and erythropoiesis. (a) LincRNA-EPS represses Pycard gene to decrease apoptosis of precursors of red blood cells. (b) RNAi knockdown of LincRNA to activate intercellular caspases and cause apoptosis. [Adapt from (75)]

Saf during transcription. It is repressed during early stages of erythropoiesis, but expression increases during late stages.⁸⁴ Other important lncRNAs have been identified in erythropoiesis such as AlncRNA-EC7, which regulates expression of erythrocyte membrane gene (*SLC4A1*).⁸⁵

Role of lncRNA in beta-thalassaemia

To date, the diversity in the phenotypic presentation of beta-thalassaemia (beta-thal) with same mutations in a *beta-globin* gene is not well understood. However, the role of epigenetic regulators and modifiers including lncRNA on haemoglobin synthesis have been studied.⁸⁶⁻⁸⁸ There have not been many studies reported on the role of dysregulated lncRNA in beta-thal in humans. Many tissue-specific lncRNAs are expressed during haemopoiesis and erythropoiesis^{80-82,89,90} The lncRNA that may play a role in changing the expression of the human *globin* gene has not been studied.⁹¹ The probable mechanisms proposed so far on the role of lncRNA on beta-thal include (i) recruitment of chromatin enzymes during epigenetic process^{69,92}; (ii) scaffolding of transcription factors^{69,93}; and (iii) enhancing the interaction between enhancers and promoters.^{94,95}

It has been proposed that lncRNAs may act by regulating protein-coding genes in such pathways in the emergence of histanoxia, which is caused by elevated foetal haemoglobin (HbF) levels and dysontogenesis in thalassaemia major patients.⁹⁶ Literature suggests that lncRNAs may be involved in the pathogenesis of beta-thal by affecting the functions of microRNAs (miRNA).⁹⁶ In normal cells, lncRNA prevents binding of miRNA, thereby suppressing the function of miRNA.⁷⁰ Hence, by studying the functions of well-characterised miRNAs, it might be possible to identify the role of lncRNA in thalassaemia. Further research is warranted to study the role of lncRNA in Hb disorders.

High levels of HbF were reported to be induced by dysregulated lncRNAs and miRNAs in hereditary persistence of foetal Hb (HPFH) and beta-thal carriers.⁹⁶ One of the possible mechanisms proposed for induction of high levels of HbF levels by lncRNA is through the activation of Hb-subunit epsilon-1 (HBE1) and haemopoietic cell lineage-inducible molecules.⁹⁶ The lncRNA can also inhibit expression of apoptosis-inducible molecules.⁹⁶ Another mechanism could be through regulation of *gamma-globin* gene (HBG) expression by the

HBS1L-MYB intergenic long noncoding RNA (HMI-lncRNA).⁹¹ The HMI-lncRNA is studied as an enhancer RNA (eRNA) and functions by facilitating the interactions between the distal enhancer and the *myeloblastosis* (*MYB*) gene promoter.⁹¹ Downregulation of this lncRNA resulted in a significant increase of HbF levels through increased expression of HBG.⁹¹ However, the exact mechanism needs to be elucidated by future studies. Furthermore, elucidating the function of lncRNA in inducing HbF levels may help researchers to design therapeutic agents that can target this pathway in sickle cell disease and beta-thal.

Other important lncRNAs that have been recently studied in beta-thal include (i) metastasis-associated lung adenocarcinoma transcript 1 (MALAT1); (ii) myocardial infarction associated transcript (MIAT) and (iii) antisense non-coding RNA in the inhibitors of CDK4 (INK) locus (ANRIL). Previously, MALAT1, which is expressed in normal cells and tissues, was studied mainly in lung cancers metastasis.⁹⁷ MIAT is involved in RNA splicing and regulation of gene expression⁹⁸, is implicated in disorders of the microvasculature^{99,100} while ANRIL's role is reported to be epigenetic silencing of other genes in its' cluster.¹⁰¹ A recent study showed that differential expression of these three important lncRNAs (MALATI, MIAT and ANRIL) in beta-thal and healthy individuals, strongly suggests a possible role for these lncRNAs in disease pathophysiology and phenotypic features.¹⁰¹ Some of the important lncRNAs associated with the pathogenesis of thalassaemia are summarised in Table 2.

Potential roles of lncRNA in other diseases

Dysregulated lncRNAs have been linked with numerous human diseases, including malignancies, such as neoplasms arising from liver, ovaries and kidney breast cancer.⁸⁸ It appears that dysregulation of lncRNA favours tissue invasion and metastasis by neoplastic progression and can cause malignant transformation. One example is lncRNA p15AS, which silences *p15* gene by initiating chromatin condensation, which results in acute haematological malignancies such as acute myeloid leukaemia (AML) and acute lymphoblastic leukaemia (ALL).¹⁰² Recently it was reported that dysregulated lncRNAs are involved in the development of malignant T-lymphocytes in *Mycosis fungoides* and Sézary syndrome.¹⁰³ Differential expression of lncRNA

TABLE 2: lncRNA involved in the pathogenesis of thalassaemia

LncRNA	Normal function	Effect /role in thalassaemia	Mechanism	References
NR_001589, NR_120526 and T315543	Facilitating the interactions between the distal enhancer and the myeloblastosis (<i>MYB</i>) gene promoter	High levels of HbF	Activation of haemoglobin subunit epsilon-1 (HBE1) and haemopoietic cell lineage-inducible molecules Inhibit expression of apoptosis-inducible molecules	<i>Lai K et al.</i> ⁹⁶
HMI-lncRNA (<i>MYB</i> enhancer RNA)	Regulates erythroid cell proliferation, maturation, and foetal haemoglobin (HbF) expression.	High levels of HbF	Downregulation increased <i>gamma-globin</i> gene (HbG) expression in erythroid cells	<i>Morrison TA et al.</i> ⁹¹
Metastasis-associated lung adenocarcinoma transcript 1 (MALAT1)	Splicing regulation	MALAT1 over-expression in thalassaemia as an adaptive response to increased DNA damage, one of the major players in thalassaemia phenotype and complications	MALAT1 expression is part of an adaptive response to hypoxia	<i>Fakhr-Eldeen A et al.</i> ¹⁰¹
myocardial infarction associated transcript (MIAT)	RNA splicing and regulation of gene expression	Upregulated expression may be associated to the increased severity of β -thalassaemia,	Increased MIAT expression in β -thal is associated with endothelial dysfunction	<i>Tsuiji H et al.</i> ⁹⁸
antisense non-coding RNA in the INhibitors of CDK4 (INK) locus (ANRIL)	Epigenetic silencing of other genes in its' cluster	Haematopoietic differentiation, proliferation, and stress response	Not known	<i>Fakhr-Eldeen A</i> ¹⁰¹
DQ583499, X-inactive specific transcript (Xist), lin-cRNA-TPM1, MRFS16P, and lincRNA-RUNX2-2	Not known	Epigenetic regulators and modifiers on Hb synthesis	Abnormal expression levels of lncRNAs and mRNA in β -thalassaemia cases may be correlated with its various clinical phenotypes	<i>Ma J et al.</i> ⁸⁸

lncRNA: long-non-coding RNA

HbF: Foetal haemoglobin

was also observed in several in rare inherited diseases like facioscapulohumeral muscular dystrophy¹⁰⁴, brachydactyly type E¹⁰⁵, Prader-Willi¹⁰⁶ and haemolysis, elevated liver enzymes, low platelet count (HELLP) syndromes.¹⁰⁵

Conclusion and Future work

The roles played by lncRNAs as regulators of multiple cellular processes⁶⁹, biomarkers for disease prognosis⁹³ and potential therapeutic targets¹⁰⁷ cannot be denied. In humans, many types of anaemia are associated with an increased rate of apoptosis of erythrocyte precursors in the bone marrow. If research can find suitable ways to silence the expression of these pro-apoptotic pathways, it would be possible to design and develop better drugs to treat anaemia secondary to thalassaemia, which is resistant to EPO therapy. Exploring into greater depth about the LncRNA-EPS-Pycard axis could also aid the development of pharmacological agents that can manipulate this pathway to impact erythropoiesis.

Authors' contribution: Sumera authored the manuscript, Radhakrishnan, Aziz Baba and George reviewed the manuscript.

Conflict of interest: The authors declare no conflict of interest.

REFERENCES

- Cappellini, M, Cohen, A, Eleftheriou, A, Piga, A, Porter, J, Taher, A. Guidelines for the Clinical Management of Thalassaemia, 2nd Revised edition. 2nd ed. Nicosia (CY); 2008.
- Taghavifar F, Hamid M, Shariati G. Gene expression in blood from an individual with beta-thalassemia: An RNA sequence analysis. *Mol Genet Genomic Med.* 2019; 7(7): e00740.
- Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. The RNA world and the origins of life In: *Molecular Biology of the Cell.* 4th edition. Garland Science; 2002
- Clancy S. Translation: DNA to mRNA to protein. *Nat Edu.* 2008; 1(1): 101.
- Martinez NM, Gilbert WV. Pre-mRNA modifications and their role in nuclear processing. *Quantitative Biology.* 2018; 6(3): 210-27.
- Hames BD. Gene expression 2, eukaryotic chromosomes by B Lewin. pp 1160. John Wiley & Sons, Chichester. 1980. ISBN 0-471-01977-1. *Biochemical Education.* 1982; 10(4): 158.
- Pierpont ME, Yunis JJ. Localisation of chromosomal RNA in human G-banded metaphase chromosomes. *Exp Cell Res.* 1977; 106(2): 303-8.
- Holmes DS, Mayfield JE, Sander G, Bonner J. Chromosomal RNA: Its properties. *Science.* 1972; 177(4043): 72-4.
- Chow LT, Gelinas RE, Broker TR, Roberts RJ. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell.* 1977; 12(1): 1-8.
- Berget SM, Moore C, Sharp PA. Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc Natl Acad Sci USA.* 1977; 74(8): 3171-5.
- Kulczyńska K, Siatecka M. A regulatory function of long non-coding RNAs in red blood cell development. *Acta Biochim Pol.* 2016; 63(4): 675-80.
- Valadkhan S, Gunawardane LS. Role of small nuclear RNAs in eukaryotic gene expression. *Essays Biochem.* 2013; 54: 79-90.
- Thomas Jr CA. The genetic organisation of chromosomes. *Annu Rev Genet.* 1971; 5(1): 237-56.
- Ohno S. So much 'junk' DNA in our genome. In: *Evolution of Genetic Systems, Brookhaven Symp. Biol.* 1972; 366-70.
- de Koning AJ, Gu W, Castoe TA, Batzer MA, Pollock DD. Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genetics.* 2011; 7(12): e1002384.
- Kung JT, Colognori D, Lee JT. Long noncoding RNAs: Past, present, and future. *Genetics.* 2013; 193(3): 651-69.
- Amaral PP, Clark MB, Gascoigne DK, Dinger ME, Mattick JS. lncRNAdb: A reference database for long noncoding RNAs. *Nucleic Acids Res.* 2010; 39: D146-51.
- Cao J. The functional role of long non-coding RNAs and epigenetics. *Biol Proc Online.* 2014; 16(1): 42.
- Houseley J, Rubbi L, Grunstein M, Tollervey D, Vogelauer M. A ncRNA modulates histone modification and mRNA induction in the yeast GAL gene cluster. *Mol Cell.* 2008; 32(5): 685-95.
- Bernstein HD, Zopf D, Freymann DM, Walter P. Functional substitution of the signal recognition particle 54-kDa subunit by its escherichia coli homolog. *Proc Natl Acad Sci USA.* 1993; 90(11): 5229-33.
- Reeves MB, Davies AA, McSharry BP, Wilkinson GW, Sinclair JH. Complex I binding by a virally encoded RNA regulates mitochondria-induced cell death. *Science.* 2007; 316(5829): 1345-8.
- Swiezewski S, Liu F, Magusin A, Dean C. Cold-induced silencing by long antisense transcripts of an arabidopsis polycomb target. *Nature.* 2009; 462(7274): 799.
- Brown CJ, Hendrich BD, Rupert JL, *et al.* The human XIST gene: Analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localised within the nucleus. *Cell.* 1992; 71(3): 527-42.
- Clemson CM, McNeil JA, Willard HF, Lawrence JB. XIST RNA paints the inactive X chromosome at interphase: Evidence for a novel RNA involved in nuclear/chromosome structure. *J Cell Biol.* 1996; 132(3): 259-75.
- Pang KC, Frith MC, Mattick JS. Rapid evolution of noncoding RNAs: Lack of conservation does not mean lack of function. *Trends in Genetics.* 2006; 22(1): 1-5.
- Wang J, Zhang J, Zheng H, *et al.* Mouse transcrip-

- tome: Neutral evolution of 'non-coding' complementary DNAs. *Nature*. 2004; 431(7010): 1-7.
27. Derrien T, Johnson R, Bussotti G, *et al*. The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Res*. 2012; 22(9): 1775-89.
 28. van Bakel H, Nislow C, Blencowe BJ, Hughes TR. Most "dark matter" transcripts are associated with known genes. *PLoS Biology*. 2010; 8(5): e1000371.
 29. Cabili MN, Trapnell C, Goff L, *et al*. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev*. 2011; 25(18): 1915-27.
 30. Martianov I, Ramadass A, Barros AS, Chow N, Akoulitchev A. Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. *Nature*. 2007; 445(7128): 666.
 31. Rinn JL, Kertesz M, Wang JK, *et al*. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell*. 2007; 129(7): 1311-23.
 32. Rintala-Maki ND, Sutherland LC. Identification and characterisation of a novel antisense non-coding RNA from the RBM5 gene locus. *Gene*. 2009; 445(1-2): 7-16.
 33. Tripathi V, Ellis JD, Shen Z, *et al*. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol Cell*. 2010; 39(6): 925-38.
 34. Muddashetty RS, Khanam T, Kondrashov A, *et al*. Poly (A)-binding protein is associated with neuronal BC1 and BC200 ribonucleoprotein particles. *J Mol Biol*. 2002; 321(3): 433-45.
 35. Beltran M, Puig I, Pena C, *et al*. A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snai1-induced epithelial-mesenchymal transition. *Genes Dev*. 2008; 22(6): 756-69.
 36. Watanabe Y, Yamamoto M. S. pombe mei2 encodes an RNA-binding protein essential for premeiotic DNA synthesis and meiosis I, which cooperates with a novel RNA species meiRNA. *Cell*. 1994; 78(3): 487-98.
 37. Campalans A, Kondorosi A, Crespi M. Enod40, a short open reading frame-containing mRNA, induces cytoplasmic localisation of a nuclear RNA binding protein in medicago truncatula. *Plant Cell*. 2004; 16(4): 1047-59.
 38. Kloc M, Wilk K, Vargas D, Shirato Y, Bilinski S, Etkin LD. Potential structural role of non-coding and coding RNAs in the organisation of the cytoskeleton at the vegetal cortex of xenopus oocytes. *Development*. 2005; 132(15): 3445-57.
 39. Sunwoo H, Dinger ME, Wilusz JE, Amaral PP, Mattick JS, Spector DL. MEN epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Res*. 2009; 19(3): 347-59.
 40. Brockdorff N, Ashworth A, Kay GF, *et al*. Conservation of position and exclusive expression of mouse xist from the inactive X chromosome. *Nature*. 1991; 351(6324): 329.
 41. Brown CJ, Ballabio A, Rupert JL, *et al*. A gene from the region of the human X inactivation centre is expressed. *Nature*. 1991; 349: 3.
 42. Bartolomei MS, Zemel S, Tilghman SM. Parental imprinting of the mouse H19 gene. *Nature*. 1991; 351(6322): 153.
 43. Ginger MR, Shore AN, Contreras A, *et al*. A noncoding RNA is a potential marker of cell fate during mammary gland development. *Proc Natl Acad Sci USA*. 2006; 103(15): 5781-6.
 44. Mourtada-Maarabouni M, Hedge VL, Kirkham L, Farzaneh F, Williams GT. Growth arrest in human T-cells is controlled by the non-coding RNA growth-arrest-specific transcript 5 (GAS5). *J Cell Sci*. 2008; 121: 939-46.
 45. Huarte M, Guttman M, Feldser D, *et al*. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell*. 2010; 142(3): 409-19.
 46. Sheik Mohamed J, Gaughwin PM, Lim B, Robson P, Lipovich L. Conserved long noncoding RNAs transcriptionally regulated by Oct4 and nanog modulate pluripotency in mouse embryonic stem cells. *RNA*. 2010; 16(2): 324-37.
 47. Loewer S, Cabili MN, Guttman M, *et al*. Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. *Nat Genet*. 2010; 42(12): 1113.
 48. Espinoza CA, Allen TA, Hieb AR, Kugel JF, Goodrich JA. B2 RNA binds directly to RNA polymerase II to repress transcript synthesis. *Nature structural & molecular biology*. 2004; 11(9): 822.
 49. Shamovsky I, Ivannikov M, Kandel ES, Gershon D, Nudler E. RNA-mediated response to heat shock in mammalian cells. *Nature*. 2006; 440(7083): 556.
 50. Brannan CI, Dees EC, Ingram RS, Tilghman SM. The product of the H19 gene may function as an RNA. *Mol Cell Biol*. 1990; 10(1): 28-36.
 51. Brockdorff N, Ashworth A, Kay GF, *et al*. The product of the mouse xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell*. 1992; 71(3): 515-26.
 52. Fantom Consortium, RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature*. 2002; 420(6915): 563.
 53. Rinn JL, Euskirchen G, Bertone P, *et al*. The transcriptional activity of human chromosome 22. *Genes Dev*. 2003; 17(4): 529-40.
 54. Bertone P, Stolc V, Royce TE, *et al*. Global identification of human transcribed sequences with genome tiling arrays. *Science*. 2004; 306(5705): 2242-6.
 55. Guttman M, Rinn JL. Modular regulatory principles of large non-coding RNAs. *Nature*. 2012; 482(7385): 339.
 56. Morlando M, Ballarino M, Fatica A. Long non-coding RNAs: New players in hematopoiesis and leukemia. *Frontiers in medicine*. 2015; 2: 23.
 57. Liu S, Hu C, Zhang J. Regulatory effects of cotranscriptional RNA structure formation and transitions. *Wiley Interdisciplinary Reviews: RNA*. 2016; 7(5): 562-74.
 58. Fatica A, Bozzoni I. Long non-coding RNAs: New players in cell differentiation and development. *Nature Reviews Genetics*. 2014; 15(1): 7.

59. Lee JT. Epigenetic regulation by long noncoding RNAs. *Science*. 2012; 338(6113): 1435-9.
60. Engreitz JM, Ollikainen N, Guttman M. Long non-coding RNAs: Spatial amplifiers that control nuclear structure and gene expression. *Nature Reviews Molecular Cell Biology*. 2016; 17(12): 756.
61. Campos EI, Reinberg D. Histones: Annotating chromatin. *Annu Rev Genet*. 2009; 43: 559-99.
62. McStay B, Grummt I. The epigenetics of rRNA genes: From molecular to chromosome biology. *Annu Rev Cell Dev Biol*. 2008; 24: 131-57.
63. Lee JT. Lessons from X-chromosome inactivation: Long ncRNA as guides and tethers to the epigenome. *Genes Dev*. 2009; 23(16): 1831-42.
64. Schnell JR, Dyson HJ, Wright PE. Structure, dynamics, and catalytic function of dihydrofolate reductase. *Annu Rev Biophys Biomol Struct*. 2004; 33: 119-40.
65. Hung T, Wang Y, Lin MF, *et al*. Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. *Nat Genet*. 2011; 43(7): 621.
66. Mao YS, Zhang B, Spector DL. Biogenesis and function of nuclear bodies. *Trends in Genetics*. 2011; 27(8): 295-306.
67. Sone M, Hayashi T, Tarui H, Agata K, Takeichi M, Nakagawa S. The mRNA-like noncoding RNA gomafu constitutes a novel nuclear domain in a subset of neurons. *J Cell Sci*. 2007; 120(Pt 15): 2498-506.
68. DeKoter RP, Singh H. Regulation of B lymphocyte and macrophage development by graded expression of PU.1. *Science*. 2000; 288(5470): 1439-41.
69. Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. *Annu Rev Biochem*. 2012; 81: 145-66.
70. Cesana M, Cacchiarelli D, Legnini I, *et al*. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell*. 2011; 147(2): 358-69.
71. Ebert MS, Sharp PA. Roles for microRNAs in conferring robustness to biological processes. *Cell*. 2012; 149(3): 515-24.
72. Zhang S, Cui W. Sox2, a key factor in the regulation of pluripotency and neural differentiation. *World J Stem Cells*. 2014; 6(3): 305-11.
73. Morceau F, Chateauvieux S, Gaigneaux A, Dicato M, Diederich M. Long and short non-coding RNAs as regulators of hematopoietic differentiation. *International journal of molecular sciences*. 2013; 14(7): 14744-70.
74. Atianand MK, Hu W, Satpathy AT, *et al*. A long noncoding RNA lincRNA-EPS acts as a transcriptional brake to restrain inflammation. *Cell*. 2016; 165(7): 1672-85.
75. Paralkar VR, Weiss MJ. A new 'linc' between noncoding RNAs and blood development. *Genes Dev*. 2011; 25(24): 2555-8.
76. Ng S, Johnson R, Stanton LW. Human long non-coding RNAs promote pluripotency and neuronal differentiation by association with chromatin modifiers and transcription factors. *EMBO J*. 2012; 31(3): 522-33.
77. Sanjuan-Pla A, Macaulay IC, Jensen CT, *et al*. Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature*. 2013; 502(7470): 232.
78. Notta F, Zandi S, Takayama N, *et al*. Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science*. 2016; 351(6269): aab2116.
79. Nandakumar SK, Ulirsch JC, Sankaran VG. Advances in understanding erythropoiesis: Evolving perspectives. *Br J Haematol*. 2016; 173(2): 206-18.
80. Paralkar VR, Mishra T, Luan J, *et al*. Lineage and species-specific long noncoding RNAs during erythro-megakaryocytic development. *Blood*. 2014; 123(12): 1927-37.
81. Alvarez-Dominguez JR, Hu W, Yuan B, *et al*. Global discovery of erythroid long noncoding RNAs reveals novel regulators of red cell maturation. *Blood*. 2014; 123(4): 570-81.
82. Hu W, Yuan B, Flygare J, Lodish HF. Long noncoding RNA-mediated anti-apoptotic activity in murine erythroid terminal differentiation. *Genes Dev*. 2011; 25(24): 2573-8.
83. Gheitis R, Manoochchri H, Gharib D, Dalil N, Gharib A. LncRNAs: A new trend in molecular biology of diseases; A review. *Chronic Diseases Journal*. 2019; 7(3): 195-206.
84. Villamizar O, Chambers CB, Mo Y, *et al*. Fas-antisense long noncoding RNA is differentially expressed during maturation of human erythrocytes and confers resistance to fas-mediated cell death. *Blood Cells, Molecules, and Diseases*. 2016; 58: 57-66.
85. Alper SL. Molecular physiology and genetics of na⁺-independent SLC4 anion exchangers. *J Exp Biol*. 2009; 212(Pt 11): 1672-83.
86. Hanly DJ, Esteller M, Berdasco M. Interplay between long non-coding RNAs and epigenetic machinery: Emerging targets in cancer? *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2018; 373(1748): 20170074.
87. Kim A, Kiefer CM, Dean A. Distinctive signatures of histone methylation in transcribed coding and noncoding human beta-globin sequences. *Mol Cell Biol*. 2007; 27(4): 1271-9.
88. Ma J, Liu F, Du X, Ma D, Xiong L. Changes in lncRNAs and related genes in β -thalassemia minor and β -thalassemia major. *Frontiers of Medicine*. 2017; 11(1): 74-86.
89. Shi L, Lin Y, Sierant M, *et al*. Developmental transcriptome analysis of human erythropoiesis. *Hum Mol Genet*. 2014; 23(17): 4528-42.
90. Alvarez-Dominguez JR, Knoll M, Gromatzky AA, Lodish HF. The super-enhancer-derived alncRNA-EC7/bloodline potentiates red blood cell development in trans. *Cell reports*. 2017; 19(12): 2503-14.
91. Morrison TA, Wilcox I, Luo H, *et al*. A long non-coding RNA from the HBS1L-MYB intergenic region on chr6q23 regulates human fetal hemoglobin expression. *Blood Cells, Molecules, and Diseases*. 2018; 69: 1-9.
92. Khalil AM, Guttman M, Huarte M, *et al*. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect

- gene expression. *Proc Natl Acad Sci USA*. 2009; 106(28): 11667-72.
93. Ørom UA, Shiekhattar R. Noncoding RNAs and enhancers: Complications of a long-distance relationship. *Trends in Genetics*. 2011; 27(10): 433-9.
 94. Batista PJ, Chang HY. Long noncoding RNAs: Cellular address codes in development and disease. *Cell*. 2013; 152(6): 1298-307.
 95. Bergmann JH, Spector DL. Long non-coding RNAs: Modulators of nuclear structure and function. *Curr Opin Cell Biol*. 2014; 26: 10-18.
 96. Lai K, Jia S, Yu S, Luo J, He Y. Genome-wide analysis of aberrantly expressed lncRNAs and miRNAs with associated co-expression and ceRNA networks in beta-thalassemia and hereditary persistence of fetal hemoglobin. *Oncotarget*. 2017; 8(30): 49931-43.
 97. Sallé-Lefort S, Miard S, Nolin M, *et al*. Hypoxia upregulates Malat1 expression through a CaMKK/AMPK/HIF-1 α axis. *Int J Oncol*. 2016; 49(4): 1731-6.
 98. Tsuiji H, Yoshimoto R, Hasegawa Y, Furuno M, Yoshida M, Nakagawa S. Competition between a noncoding exon and introns: Gomafu contains tandem UACUAAAC repeats and associates with splicing factor-1. *Genes to Cells*. 2011; 16(5): 479-90.
 99. Sun C, Huang L, Li Z, *et al*. Long non-coding RNA MIAT in development and disease: A new player in an old game. *J Biomed Sci*. 2018; 25(1): 23.
 100. Yan B, Yao J, Liu J, *et al*. lncRNA-MIAT regulates microvascular dysfunction by functioning as a competing endogenous RNA. *Circ Res*. 2015; 116(7): 1143-56.
 101. Fakhr-Eldeen A, Toraih EA, Fawzy MS. Long non-coding RNAs MALAT1, MIAT and ANRIL gene expression profiles in beta-thalassemia patients: A cross-sectional analysis. *Hematology*. 2019; 24(1): 308-17.
 102. Yu W, Gius D, Onyango P, *et al*. Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA. *Nature*. 2008; 451(7175): 202.
 103. Lee CS, Ungewickell A, Bhaduri A, *et al*. Transcriptome sequencing in Sezary syndrome identifies Sezary cell and mycosis fungoides-associated lncRNAs and novel transcripts. *Blood*. 2012; 120(16): 3288-97.
 104. Noguembor MV, Jothi M, Gabellini D. Long noncoding RNAs, emerging players in muscle differentiation and disease. *Skeletal muscle*. 2014; 4(1): 8.
 105. Maass PG, Rump A, Schulz H, *et al*. A misplaced lncRNA causes brachydactyly in humans. *J Clin Invest*. 2012; 122(11): 3990-4002.
 106. McCann KL, Baserga SJ. Long noncoding RNAs as sinks in Prader-Willi syndrome. *Mol Cell*. 2012; 48(2): 155-7.
 107. Matsui M, Corey DR. Non-coding RNAs as drug targets. *Nature reviews Drug discovery*. 2017; 16(3): 167.