Epigetics 7:6, 664-666; June 2012; © 2012 Landes Bioscience

Review of *Non-coding RNAs and the epigenetic regulation of gene expression*

A book edited by Kevin Morris

Marguerite Blignaut

Department of Botany and Zoology; Stellenbosch University; Stellenbosch, South Africa

Keywords: epigenetics, non-coding RNA, epigenomics, gene regulation, gene expression

Advances in sequencing and detection technology over the past two decades, highlighted by the data explosion brought about by the human genome project, have transformed what was previously assumed to be a relatively simple genetic landscape into a new picture where the so-called "dark matter" of the genome has stolen the spotlight from the not so hip protein-coding genes. The simplified central dogma of molecular biology, in which a gene encodes for a protein via a messenger RNA (mRNA), is still at the core of genetics but is now caught in a much more complex web of regulation by the genomic region previously known as "junk" DNA. Books such as *Non-coding RNAs and epigenetic regulation of gene expression*, published by Caister Academic Press, become essential guidelines to help us understand the current status of the very fast paced field of RNA research, which has only just started to uncover the roles of non-coding RNAs (ncRNAs) in the regulation of gene expression.

Edited by Kevin V. Morris (The Scripps Research Institute, La Jolla, CA USA), who discovered how ncRNAs could influence splicing by inhibition of transcription initiation, the book presents a comprehensive review of the role that ncRNAs play on the epigenetic regulation of gene expression, focusing on ncRNA function during disease development and their potential use in the design of therapeutic approaches. Comprised of 13 chapters written by an array of distinguished scientists, the book is organized in three sections that group together different aspect of ncRNA research: (1) form, functions and diversity; (2) gene regulation and epigenetics; and (3) disease and therapeutics.

The book starts with a comprehensive introduction to sense-antisense transcripts (SATs), which were initially discovered in the early 1980s. These transcripts, also referred to as natural antisense transcripts (NATs), are produced by the simultaneous transcription of both sense and antisense DNA strands. It is only very recently, through the incredible advances of sequencing technology and the announcement of the complete human and mouse genomes and full-length cDNA studies, that SATs have been recognized as a universal phenomenon in the plant, animal and fungi kingdoms. These transcripts contain sequences that are complementary to other endogenous RNAs and are capable of forming double-stranded RNA (dsRNA) with their sense-counterparts. Conservation of these transcripts between humans and mice indicate that they likely play a regulatory role. They have been shown to trigger alternative splicing and RNA editing within the nucleus while contributing to RNA stability in the cytoplasm of a cell. These RNA transcripts serve as precursors for many proposed gene regulation mechanisms, such as the regulation of protein-coding gene expression through the epigenetic processes of chromatin remodeling and DNA methylation.

The introductory chapter sets the stage for a much more detailed discussion, found later in the book, about the functions and regulatory roles of ncRNAs derived from NATs within different study systems. The best-known ncRNAs are called endogenous small interfering RNAs (endo-siRNAs). Endo-siRNAs, together with microRNAs (miRNAs) and piwi RNAs (piRNAs), play an instrumental regulatory and defensive role in organisms. All three classes of small RNAs show overlap with regards to their structure, synthesis and biological role. Endo-siRNAs are involved in gene regulation and transposon silencing, although the latter mechanism is still not understood. It should also be noted that a small but significant number of endo-siRNAs originate from pseudogenes via conserved NATs, which suggests not only that these transcripts are under selection but also a possible regulatory function for siRNAs. The majority of endo-siRNAs derive from non-coding DNA NATs and, in an attempt to integrate endo-siRNAs with NAT expression, an attractive but speculative model is proposed by which siRNAs could induce transcriptional silencing. This idea fits well with the overarching hypothesis of the book, which looks at the finely balanced and intricate relationship of ncRNAs and epigenetic processes. In plants, siRNAs, which act via the RNA interference (RNAi) pathway, are very effective against viral infection and have a similar defensive role as that of piRNAs and miRNAs in animals.

The elegant relationship between ncRNAs and epigenetics is further emphasized in the third chapter, which deals with the epigenetic regulation of ncRNA gene expression. Genomic imprinting is an epigenetic mechanism that triggers parent-of-origin specific expression not only of/for protein-coding genes,

Correspondence to: Marguerite Blignaut; Email: mblig@sun.ac.za
Submitted: 03/27/12; Accepted: 03/27/12
http://dx.doi.org/10.4161/epi.20170
Book review to: Morris KV, ed. Non-coding RNAs and the epigenetic regulation of gene expression. Norfolk: Caister Academic Press, 2012.

but also for a great number of ncRNA genes. Some of the ncRNA genes under regulation by this mechanism include box C/D small nucleolar RNAs (snoRNAs) and miRNA genes, of which mammalian genomes express hundreds, if not thousands, as well as long imprinted ncRNA. These ncRNAs can also control genomic imprinting by inducing local silencing of flanking protein-coding genes. Genomic imprinting, a mechanism limited to mammals, is suggested to derive from an epigenetic mechanism originally dedicated to silencing repeat DNA. This mechanism is also at work during X-chromosome inactivation, a process that offers an archetype for investigating the role of ncRNA in the regulation of the epigenome. During X-chromosome inactivation, chromatin is remodeled by utilizing a RNA co-factor as a sequence-specific guide while ncRNAs not only regulate gene expression but also control other ncRNAs, such as long non-coding RNAs (lncRNAs). Although RNA-directed DNA methylation via the RNAi pathway is also utilized, this mechanism probably represents just one of a plethora of ways by which ncRNAs play a role in X-chromosome inactivation, which in itself offers an exciting platform on which to study the various functions of ncRNAs.

Another interesting function that has only recently come to light in a system initially described in 1977, is the role of ncRNAs in alternative splicing. Alternative splicing, originally described as a mechanism that generates protein diversity, was inititally thought to only affect 20% of the human genome; today, we know that up to 90% of the human genome is subjected to this process. The mechanism of gene transcription is thus much more complicated than previously thought and requires ncRNAs not only to identify, initiate or inhibit the genes to be transcribed but also to regulate the inclusion or exclusion of certain exons, as well as to control transcriptional gene silencing through histone modification and DNA methylation.

The following chapter within this section picks up on lncRNAs, previously mentioned with regards to X-chromosome inactivation. Because this specific type of RNAs can act as important regulators in several processes where epigenetics plays a role, such as histone modification, methylation and splicing, they have been proposed as possible master regulators of gene expression and regulation in eukaryotes. Furthermore, the high level of conservation seen in vertebrates implicates lncRNAs as key developmental regulators that are essential in both epigenomic and epigenetic mechanisms. The field of ncRNAs, especially lncRNAs and their functions, is rapidly changing our view on evolutionary theories. Chapter 7 suggests a theory of inheritance based on Lamarckism (with a few small modifications) that argues that acquired traits (brought about by environmentally driven epigenetic changes and paramutation) can be mediated and inherited through generations via intermediate molecules such as RNA. This picks up on the potential of ncRNAs as drivers of natural selection and highlights the complexity of regulatory and developmental functions associated with lncRNA. The web of complexity associated with ncRNAs is aptly summarized in a figure, also used as the front cover of the book, which shows a proportional increase in the amount of non-coding DNA (aka "junk" DNA, which today we know codes for ncRNA) in the genome as we climb the evolutionary ladder toward more complex organisms.

Even though non-coding DNA is only present in very small amounts in prokaryotes when compared with complicated system such as humans and vertebrates, ncRNAs play a just as important role in gene expression in these organisms. In the case of *Plasmodium falciparum* (the protozoan parasite responsible for the most severe form of human malaria), most of the genome is dedicated to immune evasion that contributes to virulence. Although the finer details of this regulatory pathway are still unknown, ncRNAs have been implicated in the control of the expression of the genes encoding the primary virulence factors expressed by these parasites and the epigenetic marks associated with them, enabling the parasites to avoid the immune response of their human hosts.

Moving the focus back to humans, the third section of the book focuses on the role of ncRNAs in diseases such as cancer and neurological disorders, as well as the possibility to employ this knowledge in the entirely new field of pharmacopoeia. In the case of cancer and many other diseases such as psoriasis, mental disorders and autism, lncRNAs again step to the forefront. Although many intronically expressed lncRNAs have been identified in association with different cancers such as prostate cancer, melanoma and thyroid cancer, the function of many of these are still unknown. This is in contrast to the much better understood long intergenic ncRNAs (lincRNAs), which can be altered in cancer and play a central role in cellular regulation. These cancer-associated lincRNAs have furthermore been associated with imprinting, tissue-specificity and stress. They can also either encode or be regulated by smaller ncRNAs, such as miRNAs. Although it is known that lncRNAs are associated with many different kinds of cancer, the big question still remains about what role these lncRNAs play in cancer and whether they might make suitable targets for treatment.

Zooming further out, it seems relevant to ask what roles of NATs play during transcriptional and posttranscriptional regulation of tumor-suppressor genes (TSG). These genes encode for a group of proteins that guard the fidelity of the genome; loss of function on any of these genes can result in carcinogenesis. These genes are epigenetically silenced in tumor cells through NATs and, once the multilayered mechanisms through which these NATs regulate TSG are better understood, they could possibly also serve as biomarkers for cancer risk determination or treatment targets.

The very well studied tumors suppressor p53 is a good example of a TSG that is controlled by ncRNAs. The gene is subject to transcriptional, posttranscriptional and posttranslational control, highlighting the importance of ncRNAs at every level of control. The regulatory manner through which NATs, macroRNAs and small ncRNAs act, and their role in epigenetic processes such as chromatin modification, genomic imprinting, alternative splicing and methylation, are concisely summarized in the second to last chapter, which binds the book together before the concluding thoughts from Kevin Morris.

The last chapter discusses the implications of understanding that ncRNAs, whether long or short, drive and guide epigenetic changes and, therefore, have major potential as therapeutic targets. Furthermore, the knowledge that these targeted epigenetic

changes are heritable has significant impact with regards to selective pressures driven by the environment. This forces us to revisit Darwinian evolution by considering that ncRNAs can indeed improve fitness through selection by acting as intermediates in evolution. Highlighting the concluding thoughts found in every other chapter, Morris concludes: "Many questions remain regarding the role of ncRNAs and their driving forces but once these are understood [this knowledge] will reshape our models and current understanding of gene regulation."

Following the thread of ncRNAs throughout the book, it is clear that they are not only intricately wound up with epigenetic mechanisms, but that they also direct these mechanisms. This does not only change how we view the genome, epigenome and epigenetics, but also further our understanding of cellular fitness and environmental response. This book brings together more than a decade's worth of research by leaders in the field of ncRNAs and epigenetics. Each chapter is presented in a compressed and well-balanced format that can stand alone as a review article, including the history behind NATs and ncRNAs, the technical advances made, as well as current examples and discussions relevant to the chapter topic. The well-referenced and up to date text is further supported by explanatory, clearly illustrated figures, and is a must-have for any post-graduate student or researcher in the field of epigenetics and RNA.