DOWNSTATE
HEALTH SCIENCES UNIVERSITY

University Hospital of Brooklyn
College of Medicine
School of Graduate Studies
College of Nursing
School of Health Professions
School of Public Health

December 22, 2023

Dr. David Christini
Interim Dean, School of Graduate Studies
SUNY Downstate Health Sciences University
450 Clarkson Avenue, Box 60
Brooklyn, NY 11203-2098

Dear Dr. Christini:

I am pleased to inform you that Carlos Velez submitted master's thesis document, "**Unraveling the Mystery of Non-Coding Genomic Content: Evolution, Regulation, and Functional Significance**" has been determined as sufficient to pass by these faculty members of the School of Graduate Studies. Upon completion of his remaining degree requirements, he should therefore be granted the Master's Degree in Physiology.

Sincerely yours,

Juan Marcos Alarcon
Committee Member

Henri Tiedge
Committee Member

John Kubie
Committee member and
Program Director

State University of New York Downstate Medical Center
450 Clarkson Avenue, MSC 41, Brooklyn, NY 11203-2098 • Phone 718 270-2739 Fax 718 270-3378
E-mail: ethrockmorton@downstate.edu

Unraveling the Mystery of Non-Coding Genomic Content: Evolution, Regulation, and Functional Significance

by

Carlos Velez, BA

Master's Thesis
Submitted to the
Department of Physiology
SUNY Downstate Health Sciences University
in partial fulfillment of the
Requirements for the Degree of
Masters in Physiology (Concentration in Neuroscience)
December 2023

Abstract

This comprehensive review explores the diverse realm of non-coding genomic content, shedding light on its crucial functions in intricate organisms. Once considered mere "junk DNA", non-coding genomic elements have now proven to be pivotal regulators in genetic coordination. The primary focus of this review is centered around understanding the indispensability of non-coding genomic content for complex organisms. To further unpack this, an in depth look of key non-coding elements, including long non-coding RNAs (lncRNAs), microRNAs (miRNAs), small cytoplasmic RNAs (scRNAs), nucleolar RNAs (noRNAs) and transposable elements (TEs) was done. The possible evolutionary and regulatory role of non-coding genomic content was also explored. Specifically, gene regulatory network formation, and cell specific regulation. Understanding these non-coding elements is pivotal not only for the understanding of evolutionary biology, but for the development of our own precision medicine and innovative strategies in fields like conservation and agriculture. The multifaceted functions of non-coding DNA in complex organisms emphasizes its central significance in the intricate genetic framework. Ultimately this genomic content serves as a fundamental and dynamic component of the genomic landscape. This article intends to encourage additional research and allow for a deeper appreciation for the role of non-coding genomic content in the realm of complex life forms.

Acknowledgments

I wish to express my sincere appreciation to the esteemed members of my thesis committee for their time, effort, and support. Their willingness to help and support is immensely appreciated and valued.

Committee Members:

Dr. Juan Marcos Alarcon

Dr. Henri Tiedge

Dr. John Kubie

I extend special gratitude to two distinguished mentors, Dr. Sabina Hrabetova and Dr. Ivan Hernandez. Their influence, scholarly wisdom, and generous allocation of time and resources throughout my time at SUNY Downstate have greatly enhanced my understanding of becoming a scientist in the persuit of knowledge.

Introduction

The Human Genome Project, initiated in 1990 and completed in 2003, marked a significant milestone in mapping and sequencing the entire human genome. This world changing project provided researchers with a comprehensive blueprint of the human genetic code, including extensive regions of non-coding DNA. Thus, following the completion of the Human Genome Project, there was an intensified focus on deciphering the functional roles of non-coding genomic elements. Recent technological advancements have greatly supported high-throughput sequencing and bioinformatics research tools, enabling more comprehensive explorations into non-coding regions, previously referred to as "junk DNA."[1] Geneticist Susumu Ohno popularized the term "junk DNA" in 1972, proposing that non-functional evolutionary remnants could be responsible for the non-coding regions of the genome that do not code for proteins. This was held as the dominant viewpoint during the time and thus it was believed that protein-coding genes should be the focus of genetic research, pushing these genomic regions to the side [2,3].

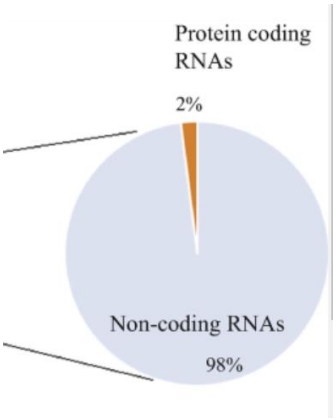

Figure 1Protein coding RNA vs Non-coding RNAs

As genomics has progressed however, it has become clear that non-coding genomic content is essential, with this being more apparent in complex organisms. In fact, in humans only about 2% of the genome codes for proteins, leaving about 98% of the genome under non coding content (Fig.1)[1].

Thus, this realization has prompted an exploration into the reasons behind their significance. This article is organized into several main sections, each focused on examining a different aspect of the significance of non-coding genomic content. The aim being to provide a

clearer picture as to the importance of non-coding genomic content as well as their evolutionary role and importance in complex organisms.

In this review, we will start by exploring important topics in non-coding RNA biology, honing on specific non-coding RNAs such as long non-coding RNAs (lncRNAs), microRNAs (miRNAs), small cytoplasmic RNAs (scRNAs) and small nucleolar RNAs (snoRNAs). This is further explored via specific examples. For instance, many lncRNAs, such as HOX Transcript Antisense Intergenic RNA (HOTAIR), act as master regulators, fine-tuning gene expression networks and controlling complex cellular processes, such as cell differentiation and organ development[5]. miRNAs like miR-155, plays a crucial role in balancing pro-inflammatory and anti-inflammatory signals, illustrating the significance of precise immune regulation[6]. Processes like cellular stress response and signal transduction, and neuronal gene expression via translation initiation regulation are heavily impacted by scRNAs. The dysregulation of these processes can serve as a potential driving force in diseases such as lupus, adding another layer of complexity to certain organisms [7]. snoRNAs such as U3 are vital in ribosomal RNA (rRNA) processing, thus demonstrating the value they hold in the regulation of other genetic content[8]. Transposable elements, specifically Alu elements, introduce genetic diversity, aiding species' adaptation to changing environments[9]. The focus with these examples being to highlight the complex regulatory landscape that shapes the biology of many living organisms.

We then discuss the evolutionary benefits of non-coding genomic content, explaining how it allows for complex gene regulatory networks, precise cell type-specific regulation, regulation of other non-coding genetic content, and the shaping of genetic variability within populations. Additional analysis to investigate the significant influence of non-coding DNA on the architecture of the genome, with particular attention to chromatin remodeling complexes and

chromatin loop formation is also explored.

Finally, we discuss the wider ramifications of understanding non-coding genomic content, which include biotechnology, agriculture/ecological management and precision medicine[10-13]. Through this scientific review paper, we hope to stimulate more research in the field of genomics while also adding to the expanding body of knowledge in this area.

1. The Functional and Regulatory Significance of Non-Coding Genomic Content

The evolution of different non-coding genomic content involves various classes of elements, each with its unique characteristics and evolutionary trajectories. John Mattick, a molecular biologist, has suggested that much of the non-coding DNA, particularly the non-coding RNAs, may be involved in a hidden genetic regulatory program. Mattick's "Hidden Genetic Program" hypothesis suggests that regardless of origin, it is irrefutable that these non-coding RNAs act as master regulators, controlling and fine-tuning complex processes such as cell differentiation, immune responses, organelle development, and responses to environmental changes[14-17].

1.1 Long Non-Coding RNAs (lncRNAs)

1.1.1 Structure, Characteristics and Function

Long non-coding RNAs (lncRNAs) form a diverse class of RNA molecules, typically characterized by lengths exceeding 200 nucleotides and displaying considerable size variability, ranging from a few hundred to tens of thousands of nucleotides[18]. Unlike shorter non-coding RNAs such as microRNAs and small interfering RNAs, lncRNAs exhibit multifaceted functions due to their size diversity. Structurally, they feature secondary motifs like hairpin loops, bulges, and stem-loop structures, with some possessing modular domains with distinct functions[19].

lncRNAs are often transcribed by RNA polymerase II just like mRNAs, and showcase similar features such as 5′-end m7G caps and 3′-end poly(A) tails. Recently however, differences in the transcription, processing, and fate of lncRNAs compared to mRNAs has been noted more[20].

In terms of transcription processes some lncRNAs are transcribed by phosphorylation-dysregulated Pol II. The phosphorylation of the Pol II carboxy-terminal domain seems to play a large role in the weak co-transcriptional splicing of these lncRNAs as well as their transcription termination independent of polyadenylation signals. This results in the subcellular and temporal accumulation

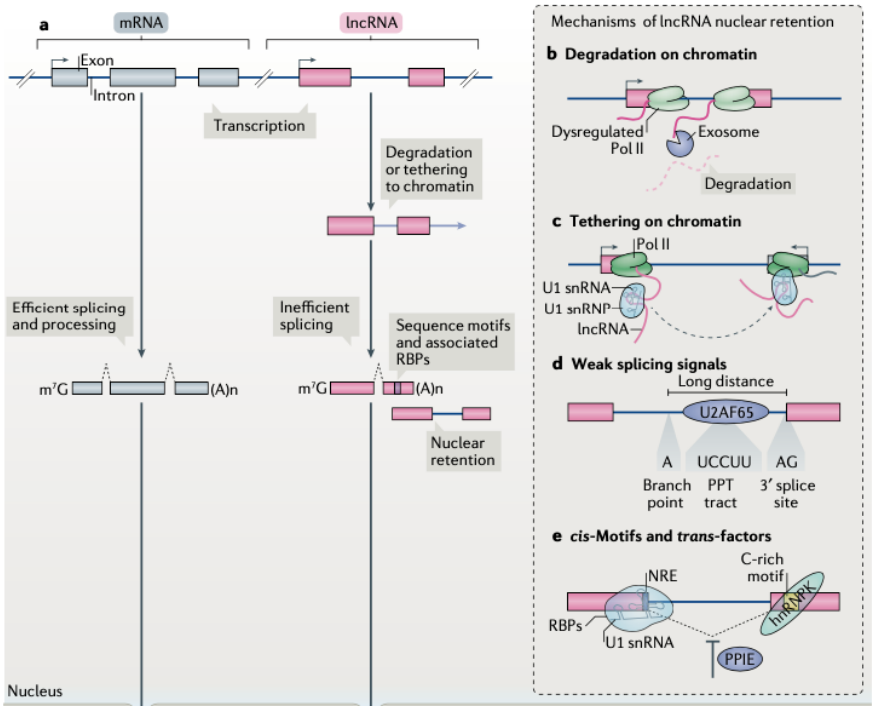

Figure 2. Biogenesis and Cellular Fates of Long Non-Coding RNAs. Adapted from "Gene Regulation by Long Non-Coding RNAs and Its Biological Functions" by Statello et al., Nature Publishing Group, 2020. DOI: 10.1038/s41580-020-00315-9

degradation by the RNA exosome (Fig 2b). Weaker internal splicing signals, as well as longer distances between the 3’ splice site and the branch point ultimately lead to lncRNAs being spliced less efficiently than mRNA (Fig 2d)[20].

The splicing process relies on the accurate recognition of specific splicing signals, including the branch point and the 3' splice site. Weaker signals in lncRNAs may make it more challenging for the splicing machinery, including the spliceosome, to precisely identify and assemble at these crucial sites, leading to less efficient splicing [21,22].

In addition, weaker splicing signals can result in suboptimal interactions with splicing factors and regulatory proteins. This diminished interaction may slow down the splicing process, reducing the efficiency of intron removal during transcription. Incomplete or inefficient splicing may leave intronic sequences uncleaved, triggering certain mechanisms that maintain lncRNAs in the nucleus. These nuclear retention mechanisms prevent the export of incompletely processed lncRNAs to the cytoplasm and ensures that only fully mature transcripts are released into from the nucleus into the cellular space [22].

Nevertheless, there are instances like with chromatin-tethered lncRNAs in which this degradation system may not always target and degrade the nucleus lncRNAs. Some chromatin-localized lncRNAs contain binding sites for U1 small nuclear RNA, which recruits the U1 small nuclear ribonucleoprotein (U1 snRNP) to Pol II (Fig 2c). This results in the tethering of numerous non-coding RNAs to chromatin which correlates with augmented nuclear retention. While not the focus of this review, it is important to note that not all lncRNAs stay within the nucleus, as they make their way into the cytosol and possibly undergo specific sorting processes that modulate different lncRNAs to specific organelles [21,23].

Conservation of lncRNAs varies, some being evolutionarily conserved across species, underscoring fundamental functions, while others are species-specific, indicating context-dependent roles[24,25]. When alternative splicing occurs in the transcripts of lncRNA genes, it can lead to the production of different lncRNA isoforms with distinct structural and functional properties. The specific combination of exons included or excluded during alternative splicing can determine the characteristics and functions of the resulting lncRNA isoforms. This process contributes to the diversity of lncRNAs observed across different tissues and cell types. This

nuanced comprehension of lncRNA characteristics and functions sets the stage for a deeper exploration of their regulatory power.

1.1.1. HOX Transcript Antisense Intergenic RNA

An example of the regulatory power in some lncRNA's is the HOX Transcript Antisense Intergenic RNA (HOTAIR). HOTAIR plays a role in regulating the expression of HOXD genes during limb development. HOXD genes are essential for determining the positional identity of body parts throughout embryonic development [5]. HOTAIR acts as an enhancer lncRNA which enhances the expression of HOXD genes by recruiting chromatin modifying complexes to the regulatory regions of these genes. The HOTAIR lncRNA interacts with the Polycomb Repressive Complex 2 (PRC2), which is involved in adding repressive histone marks to genes, thereby silencing their expression. However, instead of silencing, HOTAIR brings PRC2 to specific loci within the HOXD gene cluster, where it plays a role in establishing an activating chromatin state. This leads to increased expression of the HOXD genes, which are essential for the proper development of limbs. HOTAIR ultimately contributes to the precise spatial and temporal expression patterns of HOXD genes during limb development, ensuring that the fingers and toes form in the correct places and in the appropriate numbers. It is in this example that we find one of the most evident reasons for needing regulation. This is because if HOTAIR is not upregulated or if the expression is disrupted during development, it can have significant consequences on the regulation of HOXD genes and limb development[26,27]. The failure of this enhancer to regulate the expression of HOXD genes during limb development can lead to limb malformations or abnormalities such as the improper formation of fingers, toes, and other limb structures as well as a manifestation of disrupted spatial and temporal expression patterns in the manner of potential duplications/deletions of digits[28].

1.2 MicroRNAs (miRNAs):

1.2.1 Structure, Characteristics and Function

Another noncoding RNA that plays a regulatory role in the human system is MicroRNA(miRNAs). MicroRNAs are small, non-coding RNA molecules are typically 21-23 nucleotides in length and serve as post transcriptional regulators[6]. miRNAs are transcribed by RNA polymerase II, similar to lncRNAs and mRNAs, but their processing involves distinct mechanisms with enzymes such as Drosha and Dicer. The primary transcript being transcribed by RNA Pol II is called a primary miRNA (pri-miRNA). The pri-miRNA undergoes processing in the nucleus by an enzyme complex called the "microprocessor", which consists of two main components: Drosha and DGCR8 (DiGeorge syndrome critical region. Drosha is a ribonuclease III enzyme that cleaves the pri-miRNA to generate a hairpin-shaped precursor miRNA (pre-miRNA), which is approximately 70 nucleotides long[29].

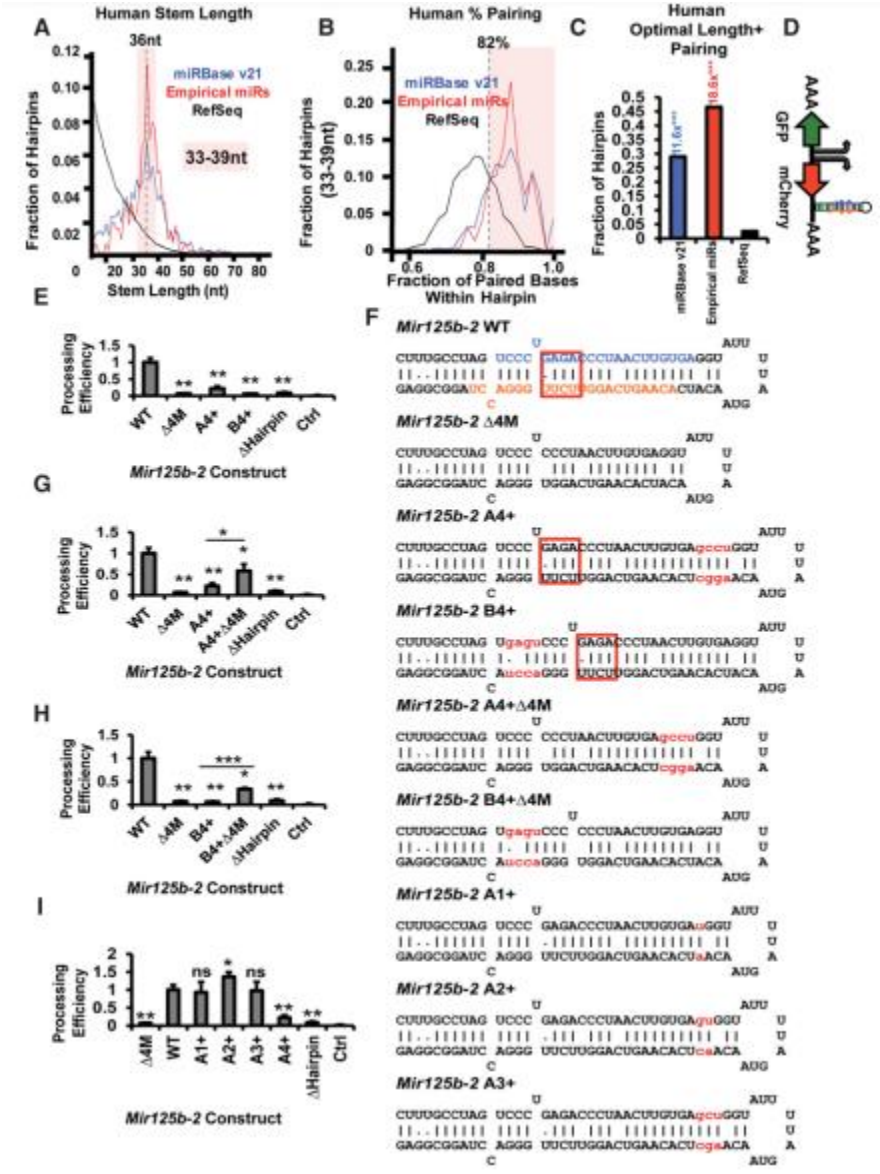

*Figure 3. MicroRNA Processing Determinants. Adapted from "Novel determinants of mammalian primary microRNA processing revealed by systematic evaluation of hairpin-containing transcripts and human genetic variation" by Roden et al., 2017, *Genome Res.* D*

Recent research by Roden et.al has identified specific features in miRNA hairpins, specifically the optimal stem length ranging which in panel A of their figure ranges from 33 to 39 nucleotides. Panel B looks at how well the two halves of the stem match, with a good match being 82% or more. As displayed in panels (E–I), when utilizing an in vivo reporter vector for pri-miRNA processing, they observed a decrease in processing efficiency for mutant transcripts of mouse miR125b-2, with stem lengths of either 31 or 39 nucleotides, compared to the wild-type miR125b-2 with a 35-nucleotide stem. This emphasizes the crucial role of stem length in the precise processing of pri-miRNA[30]. The pre-miRNA is then exported from the nucleus to the cytoplasm by the exportin-5/Ran-GTP complex. In the cytoplasm, the pre-miRNA is further processed by an enzyme called Dicer, another ribonuclease III enzyme. Dicer cleaves the pre-miRNA to generate a double-stranded RNA molecule, typically about 22 nucleotides in length[30]. The double-stranded miRNA (miRNA duplex) is unwound, and one strand, known as the guide strand, is selected for incorporation into the RNA-induced silencing complex (RISC) while the other is usually degraded. The guide strand of the mature miRNA is loaded into the RISC. The miRNA within the RISC recognizes target mRNAs through base pairings, usually in the 3' untranslated region (UTR) of the mRNA. Binding of the miRNA to its target mRNA can then lead to translational repression or mRNA degradation, depending on factors like complementarity [31]. This level of control extends to critical cellular processes, including development, cell differentiation, and responses to environmental stimuli.

When speaking about miRNAs, it's essential to note that while they often target the 3' UTRs of mRNAs as mentioned above, this is mainly seen in humans. Some miRNAs in other species may target regions beyond the 3' UTR.

With some exceptions, a notable feature of miRNAs is their conserved nature across species. Their consistent ability to regulate gene expression underscores their fundamental roles across a multitude of organisms.

1.2.2 MicroRNA-155

Among these regulatory miRNA, miR-155 emerges as a noteworthy mention, particularly when speaking on immune system homeostasis. miR-155 is crucial in balancing the scales between proinflammatory and anti-inflammatory signaling pathways. miR-155 can target and downregulate the expression of genes involved in promoting pro-inflammatory responses[6]. For example, it may target transcripts of cytokines, such as tumor necrosis factor-alpha (TNF-α)[33] and interleukin-1 beta (IL-1β)[34], which are key mediators of inflammation.

Tumor Necrosis Factor-alpha (TNF-α) is initially produced as an inactive form called pro-TNF-α, anchored to the cell surface. To become active and biologically functional, pro-TNF-α undergoes a process called cleavage. This cleavage releases the soluble and active form of TNF-α. The activation of TNF-α occurs in response to various triggers such as infection, tissue injury, or immune system activation[33].

MiR-155 has been shown to target the mRNA of TNF-α by binding to the 3' UTR of TNF-α mRNA and as a result inhibiting its translation. In a recent experiment, scientists began looking deeper, as they wanted to understand how miR-155 influences the creation of special cells (foam cells) and the body's inflammatory response, particularly in a type of immune cell called macrophages[34]. Using digital programs, they analyzed data from various

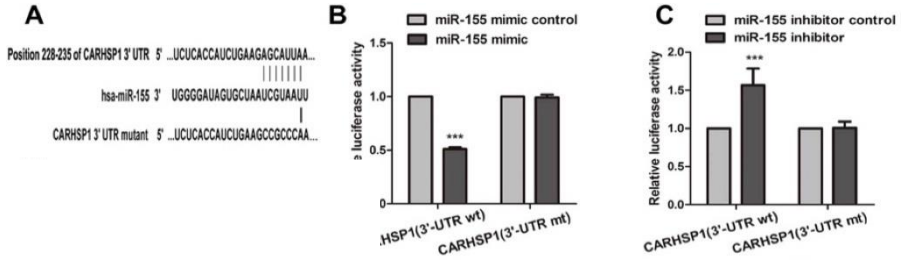

Figure 4. Validation of CARHSP1 as a functional target of miR-155. Adapted from "Mir-155 acts as an anti-inflammatory factor in atherosclerosis-associated foam cell formation by repressing calcium-regulated heat stable protein 1" by Li, X., Kong, D., Chen, H., Liu, S., Hu, H., Wu, T., Wang, J., Chen, W., Ning, Y., Li, Y., & Lu, Z.

databases to find out which genes miR-155 might directly affect. CARHSP1 was one of the genes that came up as it usually helps to stabilize TNF-α, which is involved in inflammation. The scientists then found a specific spot where miR-155 binds to CARHSP1. What was interesting was that the binding is the same in humans, mice, rats, rabbits, and horses.

Following this, the investigators co-transfected a pRL-TK plasmid carrying either a wild-type CARHSP1 3′-UTR sequence or a mutant 3′-UTR sequence lacking the miR-155 binding site into HEK-293T cells. This part of the experiment aimed to prove that the predicted miRNA binding site in the 3′-UTR of CARHSP1 was accurate. After transfection with the miR-155 mimic or miR-155 inhibitor along with the wild-type or mutant 3′-UTR sequences, the luciferase activity was measured. The results showed a significant reduction in luciferase activity when the miR-155 mimic was added (Fig. 4B), indicating that miR-155 suppressed the activity of CARHSP1. On the other hand, the addition of the miR-155 inhibitor caused a significant increase in luciferase activity (Fig. 4C), which perhaps suggested that inhibiting miR-155 led to increased CARHSP1 activity [35].

Like TNF-α, Interleukin-1 beta (IL-1β) is produced as a dormant precursor in response to infection or injury. It is activated when there is the formation of an inflammasome, which is essentially a complex of proteins responding to cellular stress. Within the inflammasome, a protein called caspase-1 becomes active and carries out a cleavage process. Caspase-1 cleaves pro-IL-1β into its active form, enabling IL-1β to exert its proinflammatory effects. Similarly to the process for TNF-α, miR-155 can target the mRNA of IL-1β via the 3' UTR of IL-1β mRNA. This ultimately leads to the downregulation of IL-1β expression [36].

By repressing the expression of proinflammatory genes, miR-155 helps prevent an excessive inflammatory response. This ability for adjustments, is essential for controlling

inflammation and avoiding tissue damage associated with prolonged or uncontrolled inflammation. Thus, miR-155 acts as a critical regulator that helps balance pro-inflammatory and anti-inflammatory signals in the immune system. Its ability to target specific genes involved in inflammation and immune cell differentiation allows for precise control of the immune response, contributing to overall immune homeostasis and avoiding excessive or prolonged inflammation.

1.3 Small cytoplasmic RNAs (scRNAs):

1.3.1 Structure, Characteristics and Function

Small cytoplasmic RNAs (scRNAs) constitute a diverse category of RNA molecules, typically characterized by their short lengths which often range from 20 to 30 nucleotides[7]. They belong to a broad class of non-coding RNA molecules that are present in cell cytoplasm, and unlike lncRNAs, scRNAs exhibit a more uniform and consistent size range. Despite this, there is a lot of variability in their functionality, with common examples of scRNAs including microRNAs (miRNAs) and small interfering RNAs (siRNAs) [7]. Structurally, scRNAs are characterized by their short sequences and the ability to form double-stranded structures when engaged in RNA interference pathways.

In terms of transcription and processing, scRNAs are often transcribed by RNA polymerase II, like mRNAs. The tight regulation of scRNA processing is crucial for their functional roles in gene regulation, mRNA degradation, and translational control [37]. Despite their small size, scRNAs play pivotal roles in gene regulation and cellular homeostasis. siRNAs for example, are involved in the RNA interference pathway, guiding the degradation of complementary RNA molecules.

scRNAs are also efficiently processed during transcription as their small size and defined structure contribute to the precision of splicing machinery, ensuring accurate and efficient processing [7,37].

As we move forward and look into post-transcriptional regulation, the involvement of cytoplasmic RNA terminal uridylyl transferases becomes evident. These enzymes play a pivotal role in modifying microRNAs and their precursors by adding either single UMP residues or oligo(U) tails. In essence, the involvement of cytoplasmic RNA terminal uridylyl transferases adds a layer of precision post-transcriptionally, where even subtle modifications play an important role in determining the fate of RNA molecules [38]. The addition of oligo(U) tails serves as a regulatory switch, dictating whether the RNA undergoes further processing, contributing to its functional maturation, or enters a pathway leading to degradation. Essentially, this process fine-tunes the fate of RNA molecules within the cellular context.

While small cytoplasmic RNAs contribute significantly to cellular processes, a subset known as Brain Cytoplasmic RNAs (BC RNAs) introduces a distinctive layer of complexity, especially within neuronal cells.

1.3.2 Brain Cytoplasmic RNA and Lupus

BC RNAs are a particular subset of small cytoplasmic RNAs that includes rodent BC1 and primate BC200. Whereas BC200 RNA comes from SRP RNA, BC1 RNA descends from tRNAAla. Despite their differences, they both share structural similarities such as a three-stem-loop domain, and a central A-rich region. They function as negative regulators of translation initiation, which has a significant effect on the expression of neuronal genes. By suppressing the formation of 48S initiation complexes, BC RNAs' translational control mechanism prevents

translation from beginning. Poly(A) binding protein (PABP) and eukaryotic initiation factors (eIFs) 4A and 4B are the main targets of this repression. After neuronal stimulation and receptor activation, translation is reversibly decreased, which essentially switches BC RNA translational control from a repressive to a permissive state[39].

BC RNAs are interesting, as autoantibodies to BC RNAs cause disruptions in people with Systemic Lupus Erythematosus (SLE). Known as anti-BC antibodies, these antibodies target dendritic targeting elements (DTEs) in the 5' stem-loop domains of BC RNAs. While it is the C-loop architectural motifs in the BC RNA 3' domain that are responsible for the translational control, the delivery to the specific areas needed is at this 5'stem loop. The RNA transport factor hnRNPA2 and these anti-BC autoantibodies engage in competitive activity to gain access to this essential structural motif. When SLE anti-BC antibodies successfully displace hnRNPA2, they ultimately impede BC RNA delivery to synaptodendritic sites of function. The profound influence of BC RNAs in preserving neural homeostasis is revealed by this disruption. We notice this with the epileptic susceptibility and cognitive dysfunction demonstrated in animal models with BC1 RNA knockout. The results point to possible links between molecular dysregulation and the emergence of neurological disorders by indicating that the complexity of BC RNAs can be disturbed by autoimmune reactivity and have observable effects on neurological function[39].

1.4 Small nucleolar RNAs (snoRNAs):

1.4.1 Structure, Characteristics and Function

Small nucleolar RNAs, or snoRNAs, are a class of RNA molecules that are almost exclusively located in the nucleolus, a specialized area within the nucleus. snoRNAs play a crucial part in the modification of other RNAs, particularly small nuclear RNA (snRNA) and

ribosomal RNA (rRNA)[41]. Usually ranging in length from 60 to 300 nucleotides, small nucleolar ribonucleoprotein (snoRNP) complexes are formed when snoRNAs interact with other proteins.

When speaking in regards to rRNA, the primary function of snoRNAs is to direct the site-specific alteration of rRNA. At specific locations along the rRNA sequence, uridine undergoes a metamorphosis into pseudouridine or methyl groups are carefully added[42]. Due to their precise nature, snoRNAs enhance both the structural complexity and the efficiency of ribosome activity during the synthesis of proteins. Furthermore, snoRNAs actively take part in the maturation and processing of snRNAs, which are important components of RNA splicing.

RNA splicing involves cutting out non-coding sections of RNA transcripts called introns and joining the remaining coding portions together known as exons. Given that small nuclear RNAs (snRNAs) are essential for RNA splicing, the role that small nucleolar RNAs (snoRNAs) play in the maturation and processing of snRNAs is especially important. Together with other proteins, they form complexes to form the spliceosome, which is necessary for precisely and effectively removing introns and guaranteeing the correct assembly of mature RNA molecules. In eukaryotic cells, this exact splicing method is essential for producing a variety of functional proteins.

1.4.2 U3 snoRNA

An example of snoRNA in action is the U3 snoRNA. U3 snoRNA is well-known for its role in the early processing events of the precursor rRNA, where it guides the cleavage of the rRNA transcript to generate the mature 18S rRNA, a key component of the small ribosomal subunit[43]. By base-pairing with the precursor rRNA and directing the recruitment of protein complexes that aid in the cleavage and modification activities, U3 snoRNA achieves this. The

correct construction and operation of ribosomes, the vital cellular machinery for protein synthesis, depend on the precise processing of rRNA by snoRNAs. Consequently, snoRNAs play a major role in the complex coordination of cellular functions associated with protein synthesis and gene expression subgroup.

U3 snoRNA intricately interacts with precursor rRNA, guiding the recruitment of protein ensembles to execute precise cleavage and modification sequences. The accuracy of snoRNA-mediated rRNA processing is crucial for the efficient assembly and optimal functioning of ribosomes. This is crucial as ribosomes are the essential players in protein synthesis.

2. Evolutionary Strategies/ Regulatory Advantages Behind Non-Coding Genomic Content

The examples and roles of non-coding genomic content mentioned above gives insight as to their functional and regulatory importance. As such, it is evident that in higher order species such as humans this non-coding genomic content is essential. Being that this phenomenon appears to almost exclusively be seen in more complex organisms, it is within reason to believe that they developed as an evolutionary adaptation. As organisms evolve, they encounter different environmental challenges and selective pressures. Non-coding DNA might provide a reservoir for evolutionary innovations, allowing for the emergence of new regulatory elements that could fine tune gene expression and adapt to changing environmental conditions.

The "Hidden Genetic Program" hypothesis suggests that these non-coding RNAs act as regulators of gene expression networks and control complex cellular processes. New interactions between non-coding regulatory elements and coding genes can lead to the emergence of complex traits and phenotypes. Thus, an increase in non-coding elements for coding regions to interact

with can help in the formation of complex gene regulatory networks and cell specificity needed for more sophisticated species.

2.1 Gene Regulatory Networks

As life on Earth continually faces diverse challenges, the precise regulation of gene expression becomes crucial for the survival and adaptation of an organism. One such key player in this intricate regulatory machinery is HOTAIR, a long non-coding RNA (lncRNA) found in mammals. A sequence analysis of HOTAIR reveals its existence in mammals with poorly conserved sequences yet considerably conserved structures, evolving faster than nearby HoxC genes[44]. This accelerated evolution may be linked to its role in guiding the Polycomb Repressive Complex 2 (PRC2) to specific loci within the HOXD gene cluster. The dynamic interplay between HOTAIR (non-coding), PRC2, and the precise regulation of (coding) gene expression, particularly during embryonic development, could have evolved over time through a combination of evolutionary processes like random genetic changes and natural selection.
The evolution of HOX genes is considered a critical step in the development of complex body plans, providing a means to specify changes along the anterior-posterior axis. As animals diversified and evolved more intricate structures, the need for fine-tuned regulation of HOX genes increased. Non-coding DNA elements, including lncRNAs like HOTAIR, may have emerged as regulatory elements to finely tune the expression of HOX genes. Despite poorly conserved sequences, the conserved structures in HOTAIR suggest functional significance. Over time, these non-coding RNAs may have gained the ability to interact with chromatin-modifying complexes like PRC2. This interaction could have facilitated the recruitment of PRC2 to specific genomic loci, adding an additional layer of regulation to the HOX gene clusters. The integration

of non-coding elements, exemplified by HOTAIR, allows for a more sophisticated control of gene expression, offering precise regulation over timing, magnitude, and duration. This intricate regulatory network, involving both coding and non-coding elements, showcases the complexity and adaptability of the genetic machinery that has evolved to meet the challenges of diverse and changing environments.

The timing of gene expression, regulated by non-coding elements like HOTAIR, needs to be coordinated with other molecular events and signals to ensure the proper integration of various aspects of development. This coordination is essential for the overall harmonious development of the more complex organism.

The magnitude of this increase is finely tuned by the regulatory actions of HOTAIR, ensuring that the appropriate level of gene expression is achieved for limb development. The influence of HOTAIR on the chromatin state and the activation of HOXD genes is not a momentary event but is sustained over specific developmental periods. This sustained impact contributes to the duration of gene expression necessary for the proper progression of limb development.

The temporal control exerted by HOTAIR ensures that the expression of HOXD genes persists for the duration required to coordinate the intricate processes involved in the formation of fingers and toes. Organisms with more precise regulation of HOX gene expression would have had advantages in terms of developmental precision. This precision could lead to better adapted structures, increasing the likelihood of evolutionary survival and reproduction.

2.2 Cell Type-Specific Regulation

The evolution of non-coding DNA has also equipped organisms with the ability to achieve cell type specific gene regulation. In multicellular organisms, the diversity of cell types requires precise control over gene expression to maintain distinct cellular identities. Different cell types within an organism have distinct functions and responsibilities, and each requires a unique repertoire of gene expression patterns to fulfill their roles effectively.

The regulation of gene expression in a cell type-specific manner is critical for the proper functioning of an organism and thus, non-coding regulatory elements play a central role. During embryonic development for example, noncoding elements contribute to the establishment of cell lineages. Enhancers and other regulatory elements guide the differentiation of stem cells into specific cell types by activating lineage specific genes and silencing others. Enhancers are compact DNA segments, typically spanning tens to hundreds of base pairs, and are defined by the presence of densely concentrated binding sites for transcription factors (TF)[45] .

These enhancers play a crucial role in recruiting transcription factor and chromatin regulatory complexes, contributing to tissue-specificity and meticulous control of transcriptional levels for target genes. This process allows for the precise modulation of gene expression. While enhancers frequently contain evolutionarily conserved sequences, such as transcription factor binding sites, the overall sequences of enhancers exhibit

miRNA	Functions	Target	References
miR-146a	Blocking differentiation	Suppressing IRAK1 and TRAF6	Bissels et al. [21]
miR-10a	Blocking differentiation	ND	
miR-29a	Regulating extracellular matrix	ND	
	Repressing apoptosis	Blocking BAK1	
miR-29b	Regulating extracellular matrix	ND	
miR-23	Inhibiting B-cell differentiation	Blocking B-cell receptor	
miR-24	Inhibiting B-cell differentiation	Blocking B-cell receptor	
	Inhibiting apoptosis	Suppressing caspase 9 and ALK4	
miR-125a/b	Inhibiting B-cell differentiation	Blocking B-cell receptor	
	Inhibiting apoptosis	Blocking BAK1 and p53	
	Remodeling cytoskeleton	ERBB2 and ERBB3	
miR-142-5p	Blocking differentiation	Suppressing UPS	Bissels et al. [21]
miR-142-3p			
miR-191			
miR-484			
miR-425			
miR-17	Blocking differentiation at early	ND	Georgantas et al. [23]
miR-24	progenitor stage		

Table 1miRNAs expressed in Hematopoietic Stem Cells (HSCs) with detailed insights. Adapted from Nassiri, S. M., Ahmadi Afshar, N., & Almasi, P. (2023). Stem Cell Res Ther, 14(1), https://doi.org/10.1186/s13287-023-03504-3"

lower conservation, leading to considerable variation in their locations between different species enhancers and their dynamics. During the differentiation of hematopoietic stem cells into erythrocytes (red blood cells) for example, specific enhancers located in non-coding regions become active. These enhancers interact with transcription factors that are characteristic of the erythroid lineage. As a result, they contribute to the activation of genes involved in hemoglobin production and other erythrocyte specific functions.

Simultaneously, non-coding elements like miRNAs associated with other lineages are silenced, ensuring that we have a commitment of the cell to that red blood cell lineage (Table 1)[45]. This cell type-specific regulation modulated by non-coding elements ensures that each blood cell type expresses the appropriate set of genes, allowing for the diversity and functionality of distinct cell lineages within the blood.

If circumstances change and nutrient availability is different, more complex organisms need a way to be able to adapt and still maintain their complexity. Thus, in conditions of nutrient scarcity, microRNAs may also be upregulated or downregulated to modulate the expression of genes involved in hematopoiesis. This allows the organism to adjust the production of different blood cell types based on the available nutrients. In times of where oxidative stress occurs via exposure to environmental pollutants, toxins, or even radiation, there is an imbalance between the production of reactive oxygen species (ROS) and the body's ability to detoxify them. Non-coding DNA could modulate the hematopoietic system to enhance the production of white blood cells, particularly those involved in combating oxidative stress, such as certain types of immune cells. This adaptation helps the organism better cope with increased oxidative challenges in the environment. Ultimately, the flexibility provided by non-coding elements in the regulation of hematopoiesis allows organisms to adjust their blood cell composition in response to various

environmental challenges. This is achieved through the precise balance of gene expression by non-coding elements.

2.3 Processing non-coding RNAs via other non-coding RNAs:

SnoRNAs, in addition to their well-known role in guiding modifications in ribosomal RNA (rRNA), also contribute significantly to the processing of other non-coding RNAs (ncRNAs). This broader involvement highlights the multifaceted impact of snoRNAs on the cellular transcriptome (complete set of RNA molecules, including messenger RNA (mRNA), non-coding RNA (ncRNA), and other RNA species, present in a cell at a specific time). One notable example is their participation in the maturation of small nuclear RNAs (snRNAs), a class of ncRNAs crucial for RNA splicing46.

In the intricate process of pre-mRNA splicing, snRNAs are integral components of small nuclear ribonucleoprotein (snRNP) complexes, which catalyze the removal of introns and the joining of exons to generate mature mRNA. SnoRNAs guide modifications in snRNAs, ensuring their proper folding and functionality within snRNPs. Modifications like pseudo uridylation or 2'-O-methylation, play a key role in stabilizing snRNAs and enhancing their binding affinity to target pre-mRNA sequences during splicing events47.

Moreover, snoRNAs have been implicated in the processing of other small non-coding RNAs, such as small nucleolar RNAs themselves and certain microRNAs (miRNAs). The intricate base-pairing interactions made possible by snoRNAs play a crucial role in guiding the cleavage and maturation of these non-coding transcripts. This involvement in the processing of diverse ncRNAs highlights the versatility of snoRNAs as molecular guides in shaping the functional characteristics of the non-coding transcriptome.

2.4 Role of Alu Elements

Elements that have evolutionary origins linked to ancient viruses, and thus have been there from the beginning can also be a contributing factor to the abundance of our non-coding genomic content. Transposable elements, a type of repetitive DNA sequence that can move within the genome can have advantageous and deleterious impacts for example. Some transposable elements, particularly Alu elements, can be transcribed into non-coding RNAs known as Alu RNAs.

Alu RNAs are a type of retrotransposons (link with retroviruses) which after transcription into RNA, reverse transcribe back into DNA, and the resulting DNA is integrated into a new genomic location. When they move, they might influence nearby genes and impact how genes are regulated[48]. Alu elements have undergone multiple rounds of amplification and insertion throughout evolution, contributing to the expansion of non-coding regions in the primate and human genome. Mutations, insertions, and deletions within Alu elements result in variations that have shaped the genomic landscape. While some Alu elements can be functionally neutral or even detrimental, some can have functional roles in the genome. For example, Alu elements can serve as regulatory elements, contribute to alternative splicing, and influence gene expression. These functional roles may contribute to their evolutionary persistence. The presence of transposable elements, such as Alu elements, contributes to genetic variation within populations which can be then be subject to natural selection. This means it will ultimately play a role in adaptation to different environments and evolutionary diversity.

An example of this is seen with the amylase enzyme that breaks down starch into sugars. The amylase gene is crucial for digestion, particularly in the context of dietary starch consumption, but different populations have varying levels of amylase gene copies, and this

variability is partly influenced by the presence of Alu elements[49]. Populations with higher starch consumption tend to have more copies of the amylase gene. The insertion of Alu elements in the vicinity of the amylase gene seems to contribute to genetic variability, potentially influencing our ability to adapt to different dietary habits. Essentially, these transposable elements appear to act as enhancers or promoters for this gene, contributing to the formation of novel regulatory networks and enabling the evolution of starch consumption via regulation of the amylase gene.

3. Genome Architecture: The Stabilizer

3.1. Chromatin Remodeling Complexes: Architects of Three-Dimensional Genomic Organization

John Mattick and other researchers have also proposed that non-coding DNA plays a significant role in the dynamic regulation of genome architecture. This refers to the three dimensional organization of the genome within the cell nucleus, which can influence gene expression patterns and cellular processes. Chromatin remodeling complexes, such as SWI/SNF and ISWI, respond to cues from non-coding regions, modifying chromatin structure to mediate gene accessibility[50]. As these complexes change the chromatin landscape, they allow for the integration of enduring epigenetic marks such as DNA methylation and histone modications. The changes induced by chromatin remodeling, especially in non-coding DNA regions, serve as a way for epigenetic modifications unfold.

These modification capabilities can be clearly seen in the CDKN2A gene for example. CDKN2A is a critical regulator of cell cycle progression and tumor suppression and distant non-coding enhancers can relay signals to the SWI/SNF chromatin remodeling complex to maintain an accessible chromatin state at the CDKN2A promoters. This includes the promoters governing the transcription of p16INK4a and p14ARF. p16INK4A and the p14ARF proteins both function

as tumor suppressors, thus, these promoters are crucial for the initiation of transcription of tumor suppression[51].

However, in conditions such as cancer, issues in non-coding regions induce heightened DNA methylation and reduced histone acetylation at the CDKN2A promoters, particularly affecting the regulatory elements associated with p16INK4a and p14ARF. Potential dysregulation of the SWI/SNF complex can occur at the same time and may also exacerbate this epigenetic issue, resulting in a repressive chromatin landscape and transcriptional silencing of the CDKN2A gene. This silencing, in turn, contributes to uncontrolled cellular proliferation, essentially leading to cancer progression.

It is important to note that epigenetic modifications, such as DNA methylation or histone modifications, can be inherited and influence chromatin structure. As mentioned previously, chromatin remodeling complexes, such as SWI/SNF, can actively modify the structure of chromatin. By repositioning nucleosomes and altering the accessibility of DNA, these complexes influence the placement of epigenetic marks. On the other hand, epigenetic marks can recruit or influence the activity of chromatin remodeling complexes, and as a result facilitating or hindering the action of chromatin remodeling complexes.

This bidirectional interplay allows organisms to dynamically respond to environmental changes. Chromatin remodeling complexes and epigenetic marks together facilitate the fine-tuning of gene expression patterns in response to external stimuli. Over time, heritable changes in epigenetic marks can provide a mechanism for adaptation.

3.2. Formation of Chromatin Loops: Facilitating Precise Gene Regulation

Certain non-coding DNA elements, such as enhancers, interact with gene promoters through the formation of chromatin loops. Chromatin looping involves physical interactions between distant genomic regions, bringing them into close proximity. These loops bring distant genomic regions into close physical proximity, allowing for enhancers to activate or repress target gene expression more effectively. Such interactions are crucial for precise gene regulation during development and cellular differentiation. While not limited to, it is possible that the non-coding genetic content in human DNA has been impacted by chromatin loops via the work of pseudogenes.

Pseudogenes are DNA sequences that resemble functional genes but have lost their ability to encode functional proteins or RNAs. Pseudogenes emerge as a result of the deterioration of genes that initially came from duplication events during the course of evolution. Their deterioration encompasses point mutations, insertions, deletions, misplacement of stop codons, or frameshifts within a gene[52].

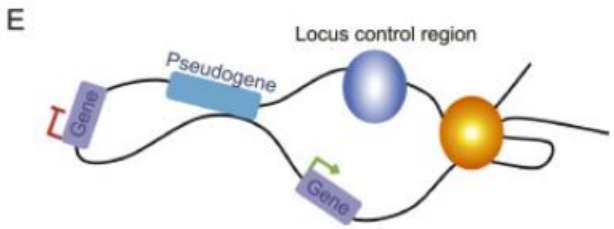

Figure 5: Models of pseudogene functioning in hematopoiesis. Adapted from Ma, Y., Chen, Z., & Yu, J. (2021). Exp Hematol, 103, 24–29. https://doi.org/10.1016/j.exphem.2021.09.001"

Pseudogenes, especially those with retained regulatory elements, could participate in the formation of chromatin loops (Fig 5)[52]. This interaction might influence the regulation of nearby genes or genomic regions. For example, a pseudogene that originated from a gene with regulatory elements, such as enhancers or promoters can over time lose its coding function. Despite this, it may still contain functional regulatory elements. These elements could participate in the formation of chromatin loops,

bringing the pseudogene into physical proximity with other genomic regions which can aid in enhancing or repressing their expression.

One well-known example of a pseudogene with regulatory functions involving chromatin loop formation is the PTENP1 pseudogene and its interaction with the PTEN tumor suppressor gene. Studies have indicated that PTENP1 can form chromatin loops with the PTEN gene[53]. This physical interaction brings PTENP1 into close proximity with the PTEN gene's regulatory elements. PTENP1 basically functions as a competitive endogenous RNA (ceRNA) and competes with PTEN mRNA for binding to microRNAs that share common binding sites between the two. By seizing miRNAs that would otherwise target PTEN mRNA, PTENP1 essentially acts as a sponge, preventing miRNA-mediated degradation of PTEN mRNA. This results in increased stability and expression of the PTEN tumor suppressor gene.

Conclusion

The exploration of non-coding genomic content reveals a complex and dynamic landscape integral to the regulatory machinery of higher-order organisms. The emergence and persistence of non-coding RNAs, exemplified above by lncRNA, miRNAs, snoRNAs, etc. underscore their regulatory significance. These non-coding elements contribute to the formation of gene regulatory networks, exerting regulatory control during development and ensuring precise cell type-specific regulation as well.

Transposable elements, specifically Alu elements, represent a fascinating link to evolutionary history. While some elements may be functionally neutral, others actively contribute to genetic variation, impacting adaptability to different environments. The case of the

amylase gene highlights how Alu elements can influence dietary adaptations, showcasing the intricate interplay between non-coding elements and functional genes.

The regulatory influence of these non-coding genetic components extends beyond the coding regions, reaching into the three-dimensional organization of the genome. Chromatin remodeling complexes respond to cues from non-coding regions, mediating gene accessibility and influencing epigenetic modifications. The bidirectional interplay between chromatin remodeling complexes and epigenetic marks offers a mechanism for dynamic responses to environmental changes, facilitating fine-tuning of gene expression over time.

Chromatin loops, facilitated by non-coding elements like enhancers and even pseudogenes, provide a mechanism for precise gene regulation. Pseudogenes, despite losing coding function, retain regulatory elements and contribute to chromatin loop formation. The example of PTENP1 and its interaction with the PTEN tumor suppressor gene illustrates how pseudogenes can modulate gene expression through physical interactions, adding another layer of complexity to the non-coding landscape.

Thus, the non-coding genomic content represents a sophisticated and integrated regulatory network that has evolved to meet the demands of complex organisms. From evolutionary adaptations to intricate gene regulatory networks and three-dimensional genome architecture, non-coding elements play pivotal roles in shaping the biology and diversity of life. Understanding the nuances of non-coding elements is vital to our comprehension of organisms and life as a whole.

Discussion

In this review article, we've discussed findings regarding non-coding genomic content, emphasizing the pivotal roles of various non-coding elements. We have illuminated the multifaceted roles and evolutionary implications of non-coding genomic content, spanning from regulatory elements to architectural influences on genome organization. Through an exploration of evolutionary strategies, gene regulatory networks, cell type-specific regulations, and the stabilizing role of genome architecture, we unravel the genetic intricacies that shape complex organisms. Thus, it is only logical to follow through and highlight the importance in terms of the field of genetics as it pertains to what has been discussed above.

Implications for Genomics

The implications for the field of genomics are profound. Understanding non-coding genomic content allows for a more comprehensive grasp of the intricate regulatory mechanisms that govern gene expression. It provides insights into the layers of complexity within the genome, leading to the potential development of more precise and targeted therapeutic interventions for various diseases. By unraveling the role of non-coding DNA in evolutionary adaptation, we are better able to understand how organisms adapt to diverse environments and ecological niches which in turn, can inspire innovative strategies for agriculture, and ecological management.

1. Precision Medicine

The utilization of microRNA (miRNA) in therapeutics, specifically by miR-34a, stands as a prominent illustration of how non-coding genetic content is valuable in precision medicine. MiR-

34a is recognized as a tumor suppressor miRNA that intricately regulates cellular processes such as apoptosis and cell cycle arrest. Its pivotal role in controlling the expression of genes associated with cancer development and progression underscores its significance in targeted therapeutic interventions. mi-34a has been a focal point for therapeutic targeting in cancer due to its downregulation in various cancers.

Researchers have developed synthetic miR-34a mimics, which when delivered to cancer cells, restore normal miRNA levels. Clinical trials have been pivotal in investigating the therapeutic potential of these miR-34a mimics, demonstrating promise in treating diverse types of cancer, including advanced solid tumors. By understanding the role of miRNAs like miR-34a in the intricate biology of cancer, therapeutic interventions can be tailored to the unique genetic and molecular characteristics of a patient's tumor. This personalized approach is crucial for maximizing efficacy while minimizing adverse effects.

In a recent study by Di Martino et al. the translational success of miR-34a is evident in both in vitro and in vivo studies, particularly when targeting multiple myeloma (MM) cells. Transient expression of synthetic miR-34a mimics and lentivirus-based miR-34a-stable expression demonstrated growth inhibition and apoptosis in MM cells. In vivo studies, including SCID mice bearing MM xenografts and SCID-synth-hu mice with complex humanized environments, further validated the anti-MM activity of miR-34a, emphasizing its potential for precision medicine applications. This example of non-coding genetic content and cancer therapeutics is one of many, and highlights how knowledge about non-coding genetic content, specifically miRNAs, can be harnessed for precision medicine[54].

2. *Insights into Evolutionary Adaptation and Agricultural Product Consumption*

The revelation of non-coding DNA's role in evolutionary adaptation provides a deeper understanding of how organisms adapt to diverse environments. Transposable elements, such as Alu elements, introduce genetic diversity that can be advantageous for species facing environmental challenges. For example, the presence of Alu insertions in genomic regions associated with dietary adaptations can contribute to the differential processing of dietary components. This has implications in helping us understand how humans have adapted to different diets throughout history. This is vital as it helps shed light on the genetic basis of dietary preferences and nutritional efficiency.

Building upon this, a recent study conducted on sixty randomly recruited individuals from the general population (RASIG) and thirty two offspring of healthy nonagenarians (GO) as part of the MARK-AGE project delved into the intricate relationship between Alu methylation, dietary factors, and aging-related genomic instability. The investigation revealed that Alu hypomethylation, a specific behavior of Alu elements, was associated with lower Alu CpG1 methylation in individuals over 65 years within the RASIG group compared to their genetically predisposed counterparts in the GO group. Interestingly enough, Alu CpG1 methylation emerged as a potential marker of aging[55].

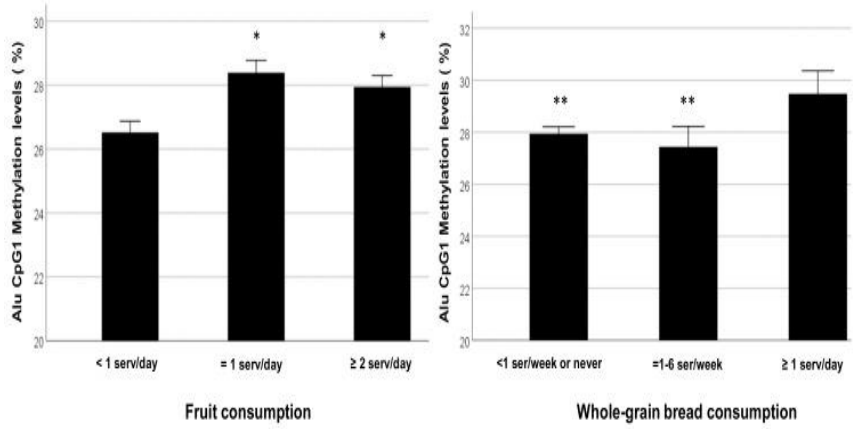

Figure 6 Adapted from Giacconi, R., Malavolta, M., Bürkle, A., Moreno-Villanueva, M., Franceschi, C., Capri, M., ... Cardelli, M. (2019). Nutrients, 11(12), 2986. https://doi.org/10.3390/nu11

The study went on to explore associations between Alu CpG1 methylation and specific dietary components. Fruit and whole-grain bread consumption were positively correlated with Alu CpG1 methylation (Fig 6), emphasizing

the potential influence of dietary choices on the epigenetic changes in Alu elements. Additionally, circulating LDL2-Cholesterol and plasma copper levels showed associations with Alu CpG1 methylation, suggesting a complex interplay between dietary intake and epigenetic modifications.

This research not only aligns with the broader understanding of non-coding DNA's role in evolutionary adaptation but also provides a concrete example of how Alu elements, through their epigenetic changes, may contribute to the genetic basis of aging and dietary preferences in human populations. This is vital as it offers a nuanced perspective on the intricate dynamics between genetic diversity, environmental adaptation, and dietary evolution.

Future Research and Open Questions:

Advancements in understanding the roles of non-coding genomic content have been significant, yet numerous complex and fascinating inquiries remain unanswered. These unresolved questions set the groundwork for future genomic research and offer the potential to gain deeper insights into non-coding elements.

Precise Mechanisms

What currently undiscovered interplay exists among different classes of non-coding RNAs, such as long non-coding RNAs, microRNAs, and nucleolar RNAs, in the intricate regulation of gene expression? How can understanding these cross-interactions advance our understanding of cellular processes and pave the way for innovative therapeutic interventions in complex diseases? Getting to the bottom of these key questions is necessary for maximizing the potential of non-coding RNA.

Functional Roles of Non-Coding Genomic Content

A persistent debate focuses on determining the functional significance of non-coding genomic content. Distinguishing between relevant and non-relevant non-coding DNA is a complex task. Researchers are employing immense genomic studies as well as other methods in order to identify patterns in the activity of non-coding elements but more work is left to be done

Roles in Diseases and Human Adaptation

Investigating the involvement of non-coding genomic content in diseases and human adaptation is a promising research area. This includes exploring the roles of non-coding RNAs in disease like lupus, cancer, and even memory impacted diseases.

How have non-coding RNAs evolved across species, and to what extent do evolutionary patterns of non-coding RNAs contribute to the susceptibility or resilience to diseases in diverse organisms? In exploring the evolutionary dynamics of non-coding RNAs, how can we uncover novel insights into the molecular basis of diseases and potential adaptive responses that could inform therapeutic strategies?"

Final Remark:

It is vital for questions such as these to continue to be broken down and explored. All these endeavors are centered around bringing the next wave of genomic discoveries, reshaping our understanding of genetics and its applications in biology, medicine, and in the world.

References

1. Ohno S. (1972) So much "junk" DNA in our genome. Brookhaven Symp Biol. ;23:366-70. PMID: 5065367.
2. Comings DE (1972). The structure and function of chromatin. *Advances in human genetics*. Springer. pp. 237–431.
3. Gilbert, Walter (1978). Why genes in pieces?. *Nature*. 271 (5645): 501 doi:10.1038/271501a0
4. Fu, Q., Liu, C.-J., Zhai, Z.-S., Zhang, X., Qin, T., & Zhang, H.-W. (2018). Single-cell non-coding RNA in embryonic development. *Single Cell Biomedicine*, 19–32. https://doi.org/10.1007/978-981-13-0502-3_3
5. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, et al. (June 2007). "Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs". *Cell*. **129** (7): 13123. doi:10.1016/j.cell.2007.05.022.
6. Schulte LN, Westermann AJ, Vogel J (Jan 2013). "Differential activation and functional specialization of miR-146 and miR-155 in innate immune sensing". *Nucleic Acids Research*. **41** (1): 542–553. doi:10.1093/nar/gks1030
7. Choudhuri, S. (2003). The path from nuclein to human genome: A brief history of DNA with a note on human genome sequencing and its impact on future research in biology. *Bulletin of Science, Technology & Society, 23*, 360–367.
8. Cléry, A.; Senty-Ségault, V.; Leclerc, F.; Raué, A.; Branlant, C. (Feb 2007). "Analysis of sequence and structural features that identify the B/C motif of U3 small nucleolar RNA as the recognition site for the Snu13p-Rrp9p protein pair". *Molecular and Cellular Biology*. **27** (4): 1191–1206. doi:10.1128/MCB.01287-06
9. Kramerov, D; Vassetzky, N (2005). "Short Retroposons in Eukaryotic Genomes". *International Review of Cytology*. **247**: 165–221. doi:10.1016/S0074-7696(05)47004-7
10. Eisenstein, M. (2021). Drug hunters uncloak the non-coding 'hidden' genome. *Nature Biotechnology*, *39*(10), 1169–1171. https://doi.org/10.1038/s41587-021-01088-y
11. Ekine-Afolabi, B. A., Njan, A. A., Rotimi, S. O., R. I., A., Elbehi, A. M., Cash, E., & Adeyeye, A. (2020). The impact of Diet on the involvement of non-coding RNAS, extracellular vesicles, and gut microbiome-virome in colorectal cancer initiation and progression. *Frontiers in Oncology*, *10*. https://doi.org/10.3389/fonc.2020.583372
12. Law, P.-P., & Holland, M. (2018). Deciphering the role of the non-coding genome in regulating gene-diet interactions. *Nutrients*, *10*(12), 1831. https://doi.org/10.3390/nu10121831
13. Yadav, A., Vidal, M., & Luck, K. (2020). Precision Medicine — networks to the rescue. *Current Opinion in Biotechnology*, *63*, 177–189. https://doi.org/10.1016/j.copbio.2020.02.005
14. Ahmad, I., Valverde, A., Ahmad, F., & Naqvi, A. R. (2020). Long noncoding RNA in myeloid and lymphoid cell differentiation, polarization and function. *Cells*, *9*(2), 269. https://doi.org/10.3390/cells9020269

15. Chini, A., Guha, P., Malladi, V.S. *et al.* (2023) Novel long non-coding RNAs associated with inflammation and macrophage activation in human. *Sci Rep* **13**, 4036. https://doi.org/10.1038/s41598-023-30568-1
16. Miguel, V., Lamas, S., & Espinosa-Diez, C. (2020). Role of non-coding-rnas in response to environmental stressors and consequences on human health. *Redox Biology*, *37*, 101580. https://doi.org/10.1016/j.redox.2020.101580
17. Holland, M. (2018). Deciphering the role of the non-coding genome in regulating gene-diet interactions. *Nutrients*, *10*(12), 1831. https://doi.org/10.3390/nu10121831
18. Wang, K. C., & Chang, H. Y. (2011). Molecular mechanisms of long noncoding RNAS. *Molecular Cell*, *43*(6), 904–914. https://doi.org/10.1016/j.molcel.2011.08.018
19. Li, R., Zhu, H., & Luo, Y. (2016). Understanding the functions of long non-coding RNAs through their higher-order structures. *International Journal of Molecular Sciences*, *17*(5), 702. https://doi.org/10.3390/ijms17050702
20. Statello, L., Guo, C. J., Chen, L. L., & Huarte, M. (2020, December 22). Gene regulation by long non-coding RNAs and its biological functions. *Nature Reviews Molecular Cell Biology*. Advance online publication. https://doi.org/10.1038/s41580-020-003159
21. Ponting CP, Oliver PL, Reik W. (2009). Evolution and functions of long noncoding RNAs. *Cell.*;136:629–641.
22. Wilusz JE, Sunwoo H, Spector DL. (2009). Long noncoding RNAs: functional surprises from the RNA world. *Genes Dev.*;23:1494–1504.
23. Basu K, Dey A, Kiran M.(2023). Inefficient splicing of long non-coding RNAs is associated with higher transcript complexity in human and mouse. *RNA Biol*. Jan;20(1):563-572. doi: 10.1080/15476286.2023.2242649.
24. Johnsson P, Lipovich L, Grandér D, Morris KV. (2014). Evolutionary conservation of long non-coding RNAs; sequence, structure, function. *Biochim Biophys Acta.* Mar;1840(3):1063-71. doi: 10.1016/j.bbagen.2013.10.035.
25. Struhl K. (2007).Transcriptional noise and the fidelity of initiation by RNA polymerase II. *Nature Structural & Molecular Biology*;14:103–105.
26. Snajdr P, Grim M, Liska F. (2010). HOX geny a vývoj koncetin v klinické medicíne i v experimentu [HOX genes and the limb development in the clinical praxis and in the experiment]. *Cas Lek Cesk*.;149(1):4-9. Czech. PMID: 20662453.
27. Pineault KM, Wellik DM. (2014). Hox genes and limb musculoskeletal development. *Curr Osteoporos Rep*. Dec;12(4):420-7. doi: 10.1007/s11914-014-0241-0.
28. Liska F. (2010). HOX genes and the limb development in the clinical praxis and in the experiment. *Cas Lek Cesk*.;149(1):4-9. Czech. PMID: 20662453.
29. Roden, C. (2017). Novel determinants of mammalian primary microRNA processing revealed by systematic evaluation of hairpin-containing transcripts and human genetic variation. *Genome Research*. Advance online publication. https://doi.org/10.1101/gr.208900.116
30. Adams, L. (2017). Pri-miRNA processing: structure is key. *Nat Rev Genet* 18, 145. https://doi.org/10.1038/nrg.2017.6
31. Huntzinger, E., Izaurralde, E. (2011). Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat Rev Genet* 12, 99–110. https://doi.org/10.1038/nrg2936

32. Jang DI, Lee AH, Shin HY, Song HR, Park JH, Kang TB, Lee SR, Yang SH. (2021). The Role of Tumor Necrosis Factor Alpha (TNF-α) in Autoimmune Disease and Current TNF-α Inhibitors in Therapeutics. *Int J Mol Sci*. Mar 8;22(5):2719. doi: 10.3390/ijms22052719.
33. Yee D, Shah KM, Coles MC, Sharp TV, Lagos D. MicroRNA-155 induction via TNF-α and IFN-γ suppresses expression of programmed death ligand-1 (PD-L1) in human primary cells. J Biol Chem. 2017 Dec 15;292(50):20683-20693. doi: 10.1074/jbc.M117.809053.
34. Ceppi M, Pereira PM, Dunand-Sauthier I, Barras E, Reith W, Santos MA, Pierre P. (2009). MicroRNA-155 modulates the interleukin-1 signaling pathway in activated human monocyte-derived dendritic cells. *Proc Natl Acad Sci U S A*. Feb 24;106(8):2735-40. doi: 10.1073/pnas.0811073106
35. Li, X., Kong, D., Chen, H., Liu, S., Hu, H., Wu, T., Wang, J., Chen, W., Ning, Y., Li, Y., & Lu, Z. (2016, February 22). Mir-155 acts as an anti-inflammatory factor in atherosclerosis-associated foam cell formation by repressing calcium-regulated heat stable protein 1. *Nature News*. https://www.nature.com/articles/srep21789
36. Choi YJ, Kim C, Choi EW, Lee SH, Chae MK, Jun HO, et al. (2022) MicroRNA-155 acts as an anti-inflammatory factor in orbital fibroblasts from Graves' orbitopathy by repressing interleukin-2-inducible T-cell kinase. PLoS ONE 17(8): e0270416. https://doi.org/10.1371/journal.pone.0270416
37. Lu C, Tej SS, Luo S, Haudenschild CD, Meyers BC, Green PJ (September 2005). "Elucidation of the small RNA component of the transcriptome". *Science*. **309** (5740): 1569. doi:10.1126/science.1114112
38. Tang TTL, Passmore LA. (2020). Recognition of Poly(A) RNA through Its Intrinsic Helical Structure. *Cold Spring Harb Symp Quant Biol*. 2019;84:21-30. doi: 10.1101/sqb.2019.84.039818. Epub Apr 15. PMID: 32295929; PMCID: PMC7116106.
39. Eom, T., Muslimov, I. A., Iacoangeli, A., & Tiedge, H. (2018). Dendritic targeting and regulatory RNA control of local neuronal translation. *The Oxford Handbook of Neuronal Protein Synthesis*. https://doi.org/10.1093/oxfordhb/9780190686307.013.8
40. Muslimov, I. A., Iacoangeli, A., Eom, T., Ruiz, A., Lee, M., Stephenson, S., Ginzler, E. M., & Tiedge, H. (2019). Neuronal BC RNA transport impairments caused by systemic lupus erythematosus autoantibodies. *The Journal of Neuroscience*, *39*(39), 7759–7777. https://doi.org/10.1523/jneurosci.1657-18.2019
41. Malik, S. S., Masood, N., Sherrard, A., Bishop, P. N., (2019). Small non-coding RNAs as a tool for personalized therapy in familial cancers. In B. Mallick (Ed.), AGO-Driven Non-Coding RNAs (pp. 179-208). *Academic Press*. ISBN 978-0128156698.
42. Jády, B. E., & Kiss, T. (2001). A small nucleolar guide RNA functions both in 2'-O-ribose methylation and pseudouridylation of the U5 spliceosomal RNA. *EMBO Journal*, *20*(3), 541-551. https://doi.org/10.1093/emboj/20.3.541
43. *Cléry, A.; Senty-Ségault, V.; Leclerc, F.; Raué, A.; Branlant, C. (Feb 2007).* Analysis of sequence and structural features that identify the B/C motif of U3 small nucleolar RNA as the recognition site for the Snu13p-Rrp9p protein pair". *Molecular and Cellular Biology*. ***27*** *(4): 1191–1206.*

44. Fu, Q., Liu, C.-J., Zhai, Z.-S., Zhang, X., Qin, T., & Zhang, H.-W. (2018). Single-cell non-coding RNA in embryonic development. Single Cell Biomedicine, 19–32. https://doi.org/10.1007/978-981-13-0502-3_3
45. Nassiri, S. M., Ahmadi Afshar, N., & Almasi, P. (2023). Insight into micrornas' involvement in hematopoiesis: Current standing point of findings. Stem Cell Research & Therapy, 14(1). https://doi.org/10.1186/s13287-023-03504-3
46. Deryusheva, S., Talross, G. J. S., & Gall, J. G. (2021). SnoRNA guide activities: real and ambiguous. *RNA*, *27*(11), 1363-1373. https://doi.org/10.1261/rna.078916.121
47. Kiss, T. (2001). A small nucleolar guide RNA functions both in 2'O-ribose methylation and pseudouridylation of the U5 spliceosomal RNA. *EMBO Journal*, *20*(3), 541-551. https://doi.org/10.1093/emboj/20.3.541
48. Szmulewicz, Martin N; Novick, Gabriel E; Herrera, Rene J (1998). "Effects of *Alu* insertions on gene function". *Electrophoresis*. **19** (8–9): 1260–4. doi:10.1002/elps.1150190806
49. Samuelson LC, Phillips RS, Swanberg LJ (1996) Amylase gene structures in primates: retroposon insertions and promoter evolution. Mol Biol Evol 13:767–779.
50. Tsuda, M., Fukuda, A., Kawai, M., Araki, O., & Seno, H. (2021). The role of the SWI/SNF chromatin remodeling complex in pancreatic ductal adenocarcinoma. *Cancer Science*, *112*(2), 490-497. https://doi.org/10.1111/cas.14768
51. Brown, V. L., Harwood, C. A., Crook, T., Cronin, J. G., Kelsell, D. P., & Proby, C. M. (2004). p16INK4a and p14ARF tumor suppressor genes are commonly inactivated in cutaneous squamous cell carcinoma. *Journal of Investigative Dermatology*, *122*(5), 1284-1292. https://doi.org/10.1111/j.0022-202X.2004.22501.x
52. Ma, Y., Chen, Z., & Yu, J. (2021). Pseudogenes and their potential functions in hematopoiesis. *Experimental Hematology*, *103*, 24–29. https://doi.org/10.1016/j.exphem.2021.09.001
53. Haddadi, N., Lin, Y., Travis, G., Simpson, A. M., Nassif, N. T., & McGowan, E. M. (2018). PTEN/PTENP1: 'Regulating the regulator of RTK-dependent PI3K/Akt signalling', new targets for cancer therapy. *Molecular Cancer*, *17*(1), 37. https://doi.org/10.1186/s12943-018-0803-3
54. Di Martino, M. T., Leone, E., Amodio, N., Foresta, U., Lionetti, M., Pitari, M. R., Cantafio, M. E., Gullà, A., Conforti, F., Morelli, E., Tomaino, V., Rossi, M., Negrini, M., Ferrarini, M., Caraglia, M., Shammas, M. A., Munshi, N. C., Anderson, K. C., Neri, A., … Tassone, P. (2023). *Data from Synthetic Mir-34a Mimics as a Novel Therapeutic Agent for Multiple Myeloma:* In Vitro *and* in Vivo *Evidence*. https://doi.org/10.1158/1078-0432.c.6520806.v1
55. Giacconi, R., Malavolta, M., Bürkle, A., Moreno-Villanueva, M., Franceschi, C., Capri, M., ... Cardelli, M. (2019). Nutritional Factors Modulating Alu Methylation in an Italian Sample from The Mark-Age Study Including Offspring of Healthy Nonagenarians. *Nutrients*, *11*(12), 2986. https://doi.org/10.3390/nu11122986