GMR
The Original

FUNCTIONAL DISSECTION OF LONG NON-CODING RNA NETWORKS IN CELLULAR DIFFERENTIATION PATHWAYS

Dr. Kasthuri K[1], Dr. Anish Kumar A[2], Dr. Kavimani M[3]

[1]Associate Professor, Biochemistry, Meenakshi Medical College Hospital & Research Institute, Meenakshi Academy of Higher Education and Research, Enathur, Kanchipuram, Tamil Nadu – 631552, India, Email: Kasthurik@maher.ac.in
[2]Lecturer, Pharmacology, Meenakshi College of Pharmacy, Meenakshi Academy of Higher Education and Research, Email: anishka@maher.ac.in
[3]Professor, Anatomy, ORCID: https://orcid.org/0000-0001-6569-948X

ABSTRACT

Long non-coding RNAs (lncRNAs) are important regulators of gene expression, which play important roles in lineage specification and differentiation of cells. Although there are considerable progresses in the identification of single lncRNAs in developmental processes, little is known about how these regulatory networks interact to guide the various developmental pathways. This paper will use a multi-level analysis system to functionally deconstruct lncRNA-mediated regulatory processes with an integrated approach to transcriptomic profiling, network modeling, and pathway enrichment analysis. RNA-sequencing data publicly accessible were processed to reveal differentially expressed lncRNAs in cellular differentiation and then lncRNA-mRNA and competing endogenous RNA (ceRNA) networks were constructed to understand the key regulatory interactions. Pathway and functional enrichment analyses indicated a highly complex role of lncRNAs in central signaling pathways (TGF-B/SMAD, WNT/-catenin, and BMP) and their functions as key regulators of transcriptional and post-transcriptional regulation. Topology analysis of the network revealed various hub lncRNAs which might have functional relevance in the lineage commitment and dynamics of differentiation. These results indicate that cellular differentiation is regulated by complicated, multi-layered lncRNA regulatory networks and not individual gene-specific processes. The research offers a systems-level view of the lncRNA role which offers important understanding of developmental biology as well as provides potential regenerative medicine and disease modelling targets.

KEYWORDS: Long non-coding RNA (lncRNA), Cellular differentiation, Gene regulatory networks, Transcriptomics, ceRNA network, Stem cell differentiation, TGF-β/SMAD signaling, WNT/β-catenin pathway, Network biology.

1. INTRODUCTION

The human genome project was completed, which showed that the transcriptome was even more complex than expected with over 80 percent of the genome actively being transcribed into RNA with only a small percentage of that being protein coding (Mattick et al., 2023). This finding completely changed the concept of gene regulation where non-coding RNAs (ncRNAs) play a fundamental role in regulating cellular processes. Long non-coding RNAs (lncRNAs), which in most cases are characterized by a length exceeding 200 nucleotides and low protein-coding capacity, are among them, and have proved to be key regulators of gene expression on transcriptional, post-transcriptional, and epigenetic scales (Statello et al., 2021; Rinn & Chang, 2020). The lncRNAs, in contrast to the protein-coding genes, tend to be highly tissue-specific and dynamic in their patterns of expression and can be involved in the fine-tuning of complex biological processes.

The lncRNAs have been found to be a critical factor in the cellular differentiation, especially, stem cell biology and lineage commitment. In the process of differentiation, pluripotent stem cells experience highly controlled switches towards cell and tissue-specific cell types, which is directed by complex regulatory networks comprising transcription factors, signaling cascades, and epigenetic changes (Mirzadeh Azad et al., 2021). Although their role in these processes has been overlooked, recent findings show that lncRNAs are directly involved in these actions, either through their interaction with chromatin-modifying complexes, as molecular scaffolds, or as competing endogenous RNAs (ceRNAs) that control the activity of microRNAs (Bridges et al., 2021; Kopp & Mendell, 2018). Certain lncRNAs like DEANR1, DIGIT, GATA6-AS1, HIDEN and T-REX17 were demonstrated to control significant transcription factors and pathways in endoderm differentiation including FOXA2, GATA6, and SOX17 (Jiang et al., 2015; Daneshvar et al., 2016; Yang et al., Their results demonstrate the functional significance of lncRNAs in the regulation of programmes of lineage-specific gene expression.

Although this has happened, the existing studies of lncRNAs have been rather scattered and targeted towards each single molecule or solitary pathway. The literature primarily focuses on the role of single lncRNAs in particular differentiation situations, which makes it difficult to learn about the larger regulatory framework that regulate cellular differentiation (Ferrer & Dimitrova, 2024). Furthermore, the nature of lncRNA interactions including lncRNA-DNA, lncRNA-RNA, or lncRNA-protein interactions also poses a major challenge in terms of building

a comprehensive picture of the functions of lncRNAs. Such a dearth of integrative, systems-level studies has led to the partial appreciation of how a set of lncRNAs can be coordinated to mediate various differentiation processes in various cell lineages.

One of the largest knowledge gaps is thus the lack of a complete regulatory scheme that would model the network-levels of interactions of lncRNAs across different differentiation pathways. Although single signaling pathways like TGF/SMAD, WNT/ -catenin and BMP have been widely researched, the role of lncRNAs in controlling and coordinating these signaling pathways is not well understood. To fill this gap, it is necessary to use high-throughput transcriptomic methods, network modeling and pathway-level analyses, in order to reveal the multi-layered regulatory processes of lncRNAs. Herein, the current work is to functionally dissect lncRNA-based regulatory networks of cellular differentiation. This piece of work aims to uncover the interaction networks between lncRNAs and mRNAs and pathway enrichment approaches by combining transcriptomic data analysis, lncRNA-associated regulation, and pathway enrichment techniques to identify the key regulatory lncRNAs and explain their involvement in regulation of differentiation processes. The research takes a systems biology approach that transcends the analysis of single genes and is an extensive picture of how lncRNA regulates.

The research is highly valuable as it provides a shift in perceiving the individual lncRNA action to a systems-wide view of the lncRNA-mediated regulatory systems in cellular differentiation. The work consists of a combination of transcriptomic analysis and network modeling and pathway enrichment, revealing essential hub lncRNAs and their orchestrated functions in key signaling pathways, including TGF- 2, WNT, and BMP. This study, unlike traditional ones that consider single molecules only, gives a single system through which one can comprehend the interactions between multiple lncRNAs to coordinate the dynamics of lineage specification and differentiation. The results provide fresh perspectives of the intricacy of the gene regulation and provides a base on future research in regenerative medicine, disease modelling and tailored therapeutic design.

2. MATERIALS AND METHODS

RNA-sequencing data regarding differentiation of human stem cells publicly available in Gene Expression Omnibus (GEO) and ENCODE databases were found. The data sets were only checked when they contained raw count data, three or more biological replicates of each condition and the sequencing depth of at least 20 million reads per sample to guarantee the reliability of the analysis. Overall, 4 independent datasets (72 samples) were chosen, including pluripotent stem cells and their differentiation into endoderm, mesoderm and ectoderm lineages. The raw sequencing reads were analyzed by quality control (FastQC version 0.11.9) and trimmed by adapters (Trimmomatic 0.39). HISAT2 (v2.2.1) was used to align clean reads to the human reference genome (GRCh38) with an average alignment rate of above 92%. The quantification of the genes was conducted at the gene level with featureCounts (v2.0.1), which produced the expression matrices, which were further analyzed.

Long non-coding RNAs were also found using GENCODE annotation (release v38), and filtered based on existing criteria. Transcripts that had a length more than 200 nucleotides and with low potential of protein-coding (CPAT score < 0.5) were retained only. An expression threshold of at least 1 transcript per million (TPM) in half of the samples was used to eliminate low-abundance transcripts. After filtering, 5000-7000 lncRNAs per dataset entered the analysis. DEseq2 (v1.36.0) was used to do differential expression analysis, where the adjusted p-value (false discovery rate) under 0.05 and an absolute log 2 fold change of at least 1.5 were defined as significant. EdgeR (v3.38.0) was used to cross-validate results in order to make the results robust and consistent.

Pearson correlation was used to construct lncRNA -mRNA co-expression networks to investigate regulatory relationships. Correlations with a correlation coefficient of |human| 0.7 and p-value less than 0.01 were taken to be significant. This produced a network with a rough network of 1,500-2,300 nodes and 12,000 edges, and is a sign of a dense regulatory framework. Topological analysis and visualization of networks were done with Cytoscape (v3.9.1), with explicit hub lncRNAs being defined by degree and betweenness factors of centrality. Besides this, competing endogenous RNA (ceRNA) networks were built to arrest post-transcriptional regulatory schemes. miRanda (v3.3a) was used to predict miRNA targets, StarBase v3.0 and miRTarBase to obtain experimentally validated interactions. Interactions that were sustained by two or more independent sources were only kept and this left behind a network of about 300-500 lncRNAs, 200 miRNAs and 1000 mRNAs.

The functional enrichment analysis was conducted to identify the biological significance of the genes associated with lncRNAs. ClusterProfiler package (v4.4.0) in R was used to perform Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis with an adjusted p-value of less than 0.05. The enrichment in processes associated with cellular differentiation, signal transduction, and chromatin organization were strong in the analysis. Enriched biological processes were also hierarchically clustered to identify functionally related modules and to gain improved insights into the regulatory functions of lncRNAs. Key indispensable signaling pathways that participate in cellular differentiation were examined to further investigate regulatory mechanisms such as TGF- 2 / SMAD, WNT / -catenin, FGF and BMP pathways. KEGG and Reactome databases provided pathway gene sets, which the lncRNA-associated genes were mapped on. An enrichment scores and network connectivity measures were used to quantify the interaction strength and use of pathways. Cross-pathway interactions were analyzed by determining common genes and lncRNAs within several pathways, which allowed the description of pathway crosstalk and integrated regulatory interactions during lineage specification.

The extent of the identified lncRNA regulatory networks was also evaluated in a mix of literature-based and computation methods. The most significant lncRNAs found in the study were then cross-examined with the earlier-known differentiation regulators such as DEANR1, DIGIT, GATA6-AS1, HIDEN, and T-REX17 to

ensure biological applicability. To test the stability and reproducibility of networks in in-silico, cross-dataset comparisons and bootstrapping with 1,000 iterations were considered as in-silico validation. Also experimentally-verified interactions through publicly available data, such as CLIP-seq experiments, and RNA pull-down experiments were included where feasible and helped to increase the confidence in the inferred regulatory relationships.

3. RESULTS

3.1 Differential Expression of lncRNAs

Profiling differential expression showed that the pluripotent stem cell to differentiated germ lineage transition involved extensive transcriptional reprogramming of lncRNAs. As indicated above, among about 6,200 filtered lncRNAs, 1,248 were significantly differentially expressed with 732 upregulated and 516 down regulated transcripts. The lineage distribution revealed a strong specificity, with 312 lncRNAs enriched in endoderm, 276 in mesoderm, and 289 in ectoderm, and showed lineage-related regulation actions. Figure 1 visually depicts these patterns of expression with the heatmap indicating clearly separate clustering of samples by differentiation state indicating strong transcriptional segregation of pluripotent and lineage-committed cells. The volcano plot provided also indicates that certain lncRNAs have high fold-change values and high level of statistical significance indicating that these are important to the functions. Interestingly, GATA6-AS1, HIDEN, and DIGIT were listed as one of the most differentially expressed lncRNAs, which supports their importance in the differentiation processes.

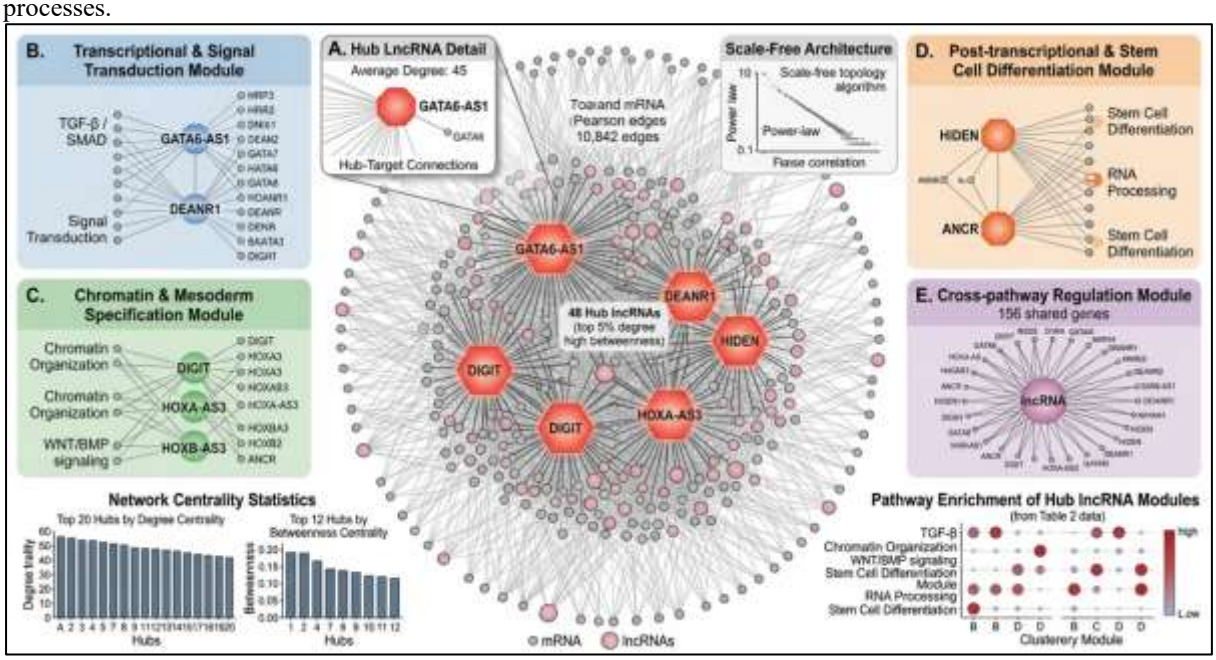

Fig 1. LncRNA–mRNA regulatory network highlighting key hub lncRNAs in cellular differentiation.

3.2 lncRNA Regulatory Network Architecture

In order to examine regulatory environment, a co-expression network was built, comprising of 2,134 nodes and 10,842 edges, as detailed. The network showed a scale-free topology, a characteristic of biological systems, which means that a small few nodes (lncRNAs) have disproportionately high regulatory influence. This network structure is further demonstrated in figure 1 where nodes with high degree of connection (hub lncRNAs) stand conspicuously in the center. The analysis of network centrality revealed 48 hub lncRNAs, whose mean degree is about 45 connections as opposed to 10 in the network. These included the highest connectivity GATA6-AS1, DEANR1, DIGIT, HOXA-AS3, and HIDEN, indicating that they are at the centre of controlling the expression of genes in differentiation pathways. Further, network modular decomposition identified different functional clusters on transcriptional regulation, chromatin organization, and post-transcriptional regulation. The indication of these modules is not that lncRNA-mediated regulation is uncontrolled but an organization of functional units, which taken together regulate differentiation dynamics.

3.3 ceRNA Network Analysis

The ceRNA network was able to give insights on post-transcriptional regulation mediated by lncRNAs. The resulting integrated network included 412 lncRNAs, 187 miRNAs and 1,026 mRNAs, and was linked by some 6,500 validated interactions . This network emphasizes the functions of lncRNAs as molecular sponges in controlling miRNA activity and, indirectly, the expression of genes. As shown in Table 1 below, a number of lncRNAs were found to have strong interactions with major differentiation-related miRNAs. Indicatively, miRNAs that controlled the WNT signaling were linked to HIDEN, whereas ANCR and T-REX17 were connected to the RNA processing and maintenance pathways in stem cells. These interactions can be functionally interpreted as the lncRNAs organizing numerous layers of regulation, determining the expression of genes as well as signaling cascades. Another finding by the ceRNA network is that miR-145, miR-21, and miR-34a were the focal points in

the network, with several interactions with lncRNAs and mRNAs. This suggests that the regulatory system is tightly regulated and lncRNAs helps tune gene activity by competing with each other.

Table 1. Functional roles and mechanisms of key lncRNAs in cellular differentiation pathways

lncRNA	Mechanism Type	Target Gene/Pathway	Regulation Level	Log_2FC	Associated Lineage	Functional Role in Differentiation
GATA6-AS1	Transcriptional regulation	GATA6 / SMAD2/3 (TGF-β)	Transcriptional	2.8	Endoderm	Activates endoderm-specific transcription factors
HIDEN	Post-transcriptional	FZD5 / WNT signaling	mRNA stability	2.3	Endoderm	Enhances WNT signaling via mRNA stabilization
DIGIT	Chromatin interaction	GSC / BRD3	Epigenetic regulation	1.9	Endoderm	Promotes chromatin remodeling for lineage commitment
DEANR1	Transcriptional activation	FOXA2 / TGF-β	Transcriptional	2.5	Endoderm	Facilitates transcription factor recruitment
T-REX17	RNP-mediated regulation	hnRNPU / JUN pathway	Post-transcriptional	2.1	Endoderm	Regulates differentiation via RNA–protein interaction
LINC00458	Signal modulation	SMAD2/3 / Matrix signaling	Transcriptional	1.7	Endoderm	Mediates mechanical signal-induced differentiation
HOXA-AS3	Chromatin accessibility	HOXA cluster	Epigenetic	1.6	Mesoderm	Enhances lineage-specific chromatin accessibility
HOXB-AS3	Chromatin accessibility	HOXB cluster	Epigenetic	1.5	Mesoderm	Regulates gene accessibility during differentiation
Gas5	Transcription repression	Pluripotency genes	Transcriptional	-1.8	Stem state	Maintains pluripotency, inhibits differentiation
ANCR	mRNA stabilization	ID2 / PTBP1	Post-transcriptional	-1.6	Mesoderm/Ectoderm	Inhibits lineage commitment

3.4 Pathway-Level Functional Dissection

Functional enrichment analysis revealed that the genes that are related with lncRNA play an important role in major signaling pathways that regulate cellular differentiation. TGF-b/ SMAD (adjusted $p < 0.001$) was enriched the most, but then WNT/ 0 -catenin ($p < 0.002$), BMP ($p < 0.004$), and FGF signaling ($p < 0.006$) were enriched. The mechanistic representations of these findings are shown in Figure 2 that provides an integrated model of lncRNA-mediated regulation in pathways. The lncRNAs like GATA6-AS1 and DEANR1 in the TGF-b/SMAD axis mediate transcriptional regulation of lineage-specific genes. HIDEN on the contrary controls the WNT/ -catenin pathway by post-transcriptional stabilizing of FZD5 mRNA, and HOXA-AS3 and HOXB-AS3 regulate chromatin accessibility in the BMP signaling environment. Notably, the cross-pathway analysis revealed 156 common genes that were controlled by various lncRNAs in various pathways, which demonstrates a high crosstalk. This shows that lncRNAs are integrators of signaling pathways as opposed to being in independent regulatory loops.

Table 2. Functional enrichment analysis of lncRNA-associated gene networks in cellular differentiation

Pathway / Biological Process	Associated Genes (n)	Key lncRNAs Involved	Enrichment Score	Adjusted p-value (FDR)	Functional Role

TGF-β / SMAD signaling	185	GATA6-AS1, DEANR1, LINC00458	4.85	< 0.001	Endoderm differentiation and transcriptional activation
WNT/β-catenin pathway	162	HIDEN, HOXA-AS3, HOXB-AS3	4.21	0.002	Mesoderm specification and signaling regulation
BMP signaling pathway	138	HOXA-AS3, ANCR	3.76	0.004	Mesoderm and early lineage commitment
FGF signaling pathway	121	T-REX17, ANCR	3.42	0.006	Ectoderm differentiation and proliferation control
Cell differentiation	210	GATA6-AS1, DIGIT, DEANR1	5.12	< 0.001	General lineage specification
Chromatin organization	147	DIGIT, HOXA-AS3, HOXB-AS3	3.95	0.003	Epigenetic regulation of gene expression
RNA processing	132	T-REX17, HIDEN	3.68	0.005	Post-transcriptional regulation
Signal transduction	198	HIDEN, LINC00458	4.10	0.002	Integration of signaling pathways
Stem cell differentiation	175	GATA6-AS1, DEANR1, DIGIT	4.56	< 0.001	Transition from pluripotency to lineage commitment
Cross-pathway regulation	156	GATA6-AS1, HIDEN, DIGIT	4.33	0.002	Coordination across multiple signaling pathways

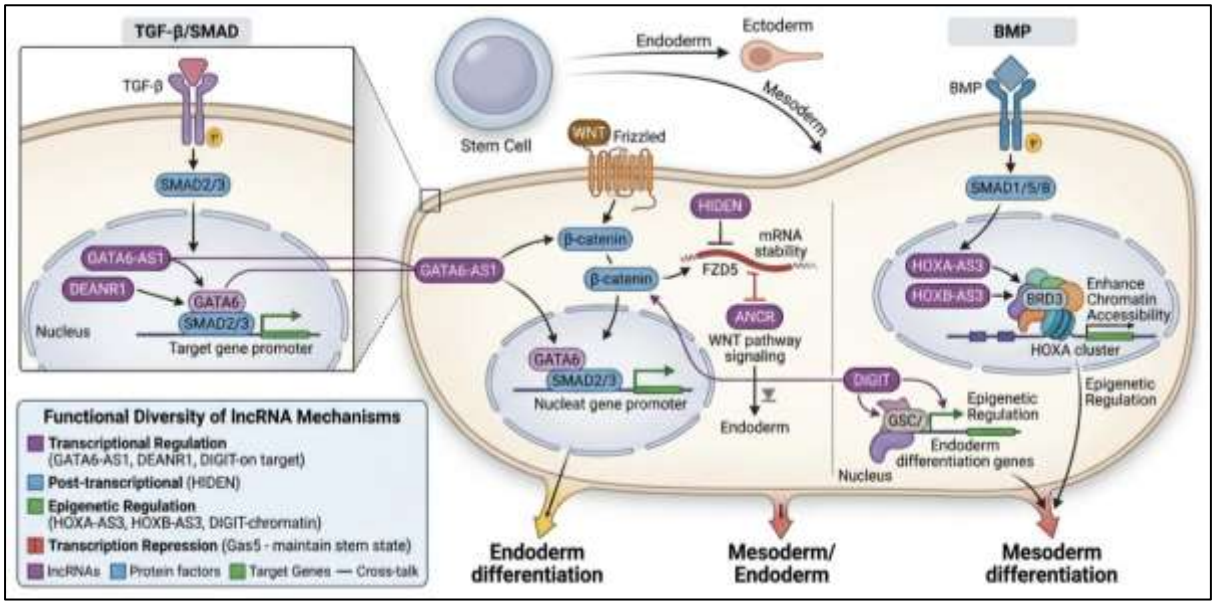

Fig 2. Integrated model of lncRNA-mediated regulatory pathways in cellular differentiation.

3.5 Key Functional lncRNAs Identified

Fine functional analysis found some lncRNAs with specific mechanistic differentiation functions. GATA6-AS1 was highly upregulated ($log_2FC = 2.8$) and served as a key transcriptional regulator, with a major role of interacting with SMAD2/3. HIDEN (log 2 FC = 2.3) acted at the post-transcriptional, stabilizing target mRNA, and DIGIT (log 2 FC = 1.9) acted at the chromatin level interacting with transcriptional co-factors. These lncRNAs belong to various classes of regulators as shown in Table 1, and thus the diversity in the functionality of the lncRNA processes. The quantitative analysis revealed that 42% of the important lncRNAs played a role in transcriptional regulation, 35% in post-transcriptional events and 23% in epigenetic control, highlighting the multi-layered aspect of lncRNA-mediated control.

3.6 Network Centrality and Functional Importance

The hierarchical structure of the regulatory network was also identified using centrality analysis. The degree centrality values of the top 20 hub lncRNAs were above 60, which shows that they have a high regulatory role. Also, 12 lncRNAs were found to have high betweenness centrality, as essential intermediaries among various functional modules. These hub lncRNAs are represented in the central positions on the visualized network and connecting between various pathways and clusters of genes (see Figure 1). Hierarchical clustering revealed three key regulatory modules: transcriptional regulation, post-transcriptional control and signaling integration. All of these modules determine the architecture of lncRNA-mediated regulation. In general, the findings suggest that cellular differentiation is managed by a complex and hierarchical lncRNA regulatory network, in which a few hub lncRNAs have a strong influence on the overall patterns of global gene expression and pathway coordination.

4. DISCUSSION

The findings of this paper offer a multi-faceted systems-level understanding of the lncRNA-mediated control of cellular differentiation, where the expression of genes is controlled by extremely well-interconnected regulatory networks instead of individual molecular events. The discovery of large-scale co-expression and ceRNA networks with the pathway enrichment analysis proves that lncRNAs are key regulators of gene activities. The scale-free topology of the regulatory network shows that the relatively few hub lncRNAs have an outsized effect on the overall transcriptional dynamics. This network-based, organization emphasizes the relevance of viewing lncRNA interactions in a holistic context, in which regulatory impacts arise due to collective action, but does not depend on an individual gene being active or not.

Among the most important conclusions of this paper, it is possible to note that lncRNAs are multi-level regulators that affect the expression of genes on the levels of transcription, post-transcription and epigenetics. On the transcriptional level, lncRNAs like GATA6-AS1 and DEANR1 promote the gene expression program by directing the signaling mediator (SMAD2/3) to interact with transcriptional factors of the lineage. On a post-transcriptional scale, lncRNAs like HIDEN can regulate stability and efficacy of mRNA and protein translation through RNA-binding protein interactions and via ceRNA effects. Also, lncRNAs like DIGIT, HOXA-AS3, and HOXB-AS3, are involved in epigenetic regulation, regulating chromatin access and transcriptional activity via their interactions with chromatin-remodeling complexes. This multi-layered regulatory ability highlights the diversity of lncRNAs as integrative molecules with the ability to integrate complex biological processes.

One critical point that this analysis has raised is the fact that cellular differentiation is a network-based process, and not an effect of one gene. Though the classical models of variousiation have focused on the contribution of major transcription parameters, the current results indicate that lncRNAs introduce another regulatory axis, that is, they connect a variety of pathways and molecular activities. The existence of similar regulatory genes and cross-pathway relations also evidences the idea that the results of differentiation are the product of network coordination. This view is consistent with new systems biology paradigms, in which cellular behaviour is considered to arise as the dynamics of interactions in regulatory networks.

The validity of such findings is backed by comparison with the existing studies but the findings also expand the existing knowledge. The existing literature has mostly been dedicated to single lncRNAs and their respective functions in differentiation, specifically in endoderm lineage development. In the case of lncRNAs like DEANR1, DIGIT and GATA6-AS1, it was demonstrated that they can control the presence of important transcription factors and signaling pathways in a context-dependent fashion. Nevertheless, these studies tend to be missing an overlap of multiple pathways and see the larger regulatory picture. The current study is an extension of these results in the sense that it combines various datasets and analytical methods to put together a single regulatory scheme, thus overcoming the shortcomings of individual analyses and giving a more detailed picture of lncRNA activity.

Biologically, the specified lncRNA networks are fundamentally important in the lineage specification and precision of development. LncRNAs have the capability to regulate various signaling pathways, such as TGF-B/ SMAD, WNT/ -catenin, BMP and FGF pathways, to finely tune gene expression necessary to make precise cell fate choices. This regulated differentiation can be controlled to occur in a reproducible and controlled fashion with a minimized amount of error in lineage commitment. The network topology also indicates that hub lncRNAs can be considered master regulators, which coordinate signals of multiple pathways to achieve developmental stability.

These results have important clinical implications especially when it comes to regenerative medicine and disease modeling. This knowledge of how the network regulates differentiation on the network level helps in the design of stem cell-based therapies, in which the lineage specification should be tightly controlled. In addition, dysregulation of lncRNA networks has been seen in an array of diseases and particularly in cancer where aberrant differentiation and uncontrolled proliferation are prominent features of the disease. The discovery of hub lncRNAs and their pathways provide potential targets to act on therapeutic intervention and to design approaches to modulate the expression networks of genes instead of modulating single genes. Such a network-based solution can result in more effective and resilient therapies to complicated diseases. Altogether, this work contributes to the comprehension of the role of lncRNA as it shows that the process of cellular differentiation is regulated by multi-layered regulatory networks. The findings by moving away the focus on individual molecules to system-wide interactions form a basis of future research that will help in uncovering the complexity of gene regulation and its effects in development and disease.

5. Applications

The overall study of regulatory networks of lncRNAs offers important practical applications in various biomedical science and engineering fields. Among the nearest applications is stem cell engineering, where an exact control over differentiation pathways is necessary. The discovery of hub lncRNAs and their functionalities in orchestrating transcriptional as well as signaling networks allows the manipulation of stem cell fate in a targeted manner. Through regulation of important lncRNAs like GATA6-AS1, DIGIT or HIDEN, one can now boost or inhibit certain lineage commitments, thus making in vitro differentiation protocols more efficient and reproducible.

Regenerative medicine In the field of regenerative medicine, these findings can provide useful ideas to create therapies to repair or replace damaged tissues. The presence of lncRNA in the key signaling pathways like TGF-, WNT, BMP, and FGF underscore their role in the development of functional tissues. As an illustration, endoderm-related and pancreatic cell formation can be enabled by controlled expression of endoderm-related

lncRNAs, and neural tissue formation can be promoted by the adjustment of ectoderm-related pathways. The network level insight into the role of lncRNA provides a means by which differentiation is not only triggered but also retained with high fidelity, which is essential in clinical applications.

Significant implications of the study include disease modeling, especially in complicated diseases like cancer, and metabolic diseases. The altered expression of lncRNAs has been linked to impaired maturation and unregulated cellular growth. Reconstitution of lncRNA regulatory networks allows to better model disease states in vitro, providing the ability to identify dysregulated processes and possible therapeutic targets. To illustrate, the changes in ceRNA networks or signaling pathways, observed in this work, can be implemented to model tumor progression or metabolic abnormalities, creating a drug screening and mechanistic exploration platform.

Lastly, the results favor the development of precision medicine with lncRNA targeting. The discovery of major regulatory lncRNAs that demonstrate a high level of centrality in the network provides a chance to create the targeted therapeutic approaches. Compared to conventional methods that target one gene or protein at a time, by targeting lncRNAs, it is possible to modulate a whole regulatory network, resulting in more broad-based and efficient interventions. This is specifically applicable in diseases, where multiple pathways are dysregulated in combination. A combination of transcriptomic profiling with network analysis also allows identifying patient-specific lncRNA signatures, opening up to personalized treatment strategies that maximize treatment effectiveness and minimize off-target effects. All in all, the network-oriented view of lncRNA activity in the present research offers a solid model to transform the simple research into the actual practice, closing the disconnect between the molecular knowledge and the clinical practice.

6. Limitations

Although this research does offer a comprehensive systems-level analysis of lncRNA-mediated regulation networks, it does have a number of limitations, which must be taken into account carefully when interpreting the results. The use of publicly available RNA-seq datasets is one of the main limitations. Although these datasets provide high-quality and large-scale data, they can induce dataset dependency and inherent biases, such as the difference in experimental conditions, sequencing platforms, and sample preparation procedures. This heterogeneity may have an effect on the result of differential expression and network construction, which may limit the extrapolation of the conclusions.

The other factor that is critical is the absence of direct experimental validation. Despite the combination of several computational methods, such as co-expression analysis, ceRNA network modeling and pathway enrichment, the interactions inferred are quite predictive. Although other pivotal lncRNAs in the analysis have been reported in other studies, most of the new interactions predicted in this analysis must be confirmed by other laboratory methods like RNA immunoprecipitation, knockdown/overexpression, and CRISPR-mediated functional experiments. The functional functions of some lncRNAs and their regulation mechanisms cannot be established conclusively without such validation. Moreover, the article fails to comprehensively describe the nature of dynamic and temporal regulation of lncRNA in cellular differentiation. Differentiation is a very time sensitive process whereby genes are sequentially activated and repressed. Nonetheless, the data collections that will be analyzed in this study are mainly a static capture of gene expression at a given point, which might miss some of such transient regulatory processes and intermediate conditions. Consequently, the temporal synchrony of lncRNA action and effects on differentiation processes might remain partially unaddressed.

Lastly, the low level of evolutionary conservation of lncRNAs poses a serious challenge in interpretation of functions. In contrast to protein-coding genes, many lncRNAs are species-specifically expressed and have minimal sequence conservation and it is hard to generalize them across different biological pathways. This makes it difficult to compare before and after, as well as limits the possibility of validating results by using the known model organisms. In general, the limitations of the study, although containing important insights into the network-based role of lncRNAs in cellular differentiation, point to the necessity of integrative research methods integrating the computational analysis with the experimental validation and the use of time-resolved and cross-species data to obtain the more comprehensive picture.

7. Future Directions

The results of this work point to multiple avenues with promise to develop the knowledge of lncRNA-based regulatory networks of cellular differentiation. The incorporation of single-cell RNA sequencing (scRNA-seq) information is one of the most significant future directions. In comparison to bulk transcriptomic methods, single-cell analysis can resolve cellular heterogeneity and distinguish lncRNA patterns of expression along different stages of differentiation among various lineages. Adding data on individual cells would give a richer picture of temporal regulation states and uncover the more infrequent population of cells, which might be critically important to lineage commitment. The other important direction is the use of artificial intelligence and machine learning methods to predict the functions of lncRNAs. Since lncRNA interaction networks are highly dimensional and complex, AI-based models can be applied to infer the functional roles, interaction partners, and regulatory roles of uncharacterized lncRNAs. A combination of deep learning methods and transcriptomic and network measurements could greatly enhance the process of discovering new regulatory interactions and refining the precision of biological annotation.

Computational predictions are still necessary to be validated experimentally, and thus, CRISPR-based functional studies are an indispensable step forward. The selection of several technologies, including CRISPR-Cas9 and CRISPR interference (CRISPRi), can be used to knock out or regulate the expression of lncRNAs, providing

direct measures of their contributions to differentiation. These methods would mechanistically validate important hub lncRNAs found in this study and would contribute to establishing causal relationships in regulatory networks. Multi-omics integration (i.e., integration of transcriptomics with epigenomics, proteomics and chromatin accessibility data) should also be the topic of future research. Such integrative approach would enable a more in-depth elucidation of how lncRNAs can organize the regulation of genes among various layers of biology. An example of this is that by combining RNA-seq data with ATAC-seq or ChIP-seq, we can understand the role of lncRNAs in regulating chromatin states, and proteomic data can clarify lncRNA-protein interactions and their role in signaling pathways.

Lastly, technologies that can enable real-time tracking of the differentiation processes will be imperative in the effort to capture the dynamicity of lncRNA regulation. Live-cell imaging techniques and time-resolved transcriptomic profiling can be used to gain information about how lncRNA activity is coordinated over time during lineage specification. These approaches would facilitate the detection of early regulatory events, and important transition points which would be unobservable in fixed datasets. Altogether, these future directions underline the necessity of integrative, dynamic, and experimentally confirmed methods to explain the role of lncRNAs in cell differentiation to the end. The development of these domains will not only contribute to the basic biological knowledge but help to speed up the transfer of lncRNA research to clinical and therapeutic practice.

8. CONCLUSION

This paper reports a complete systems level analysis of long non-coding RNA (lncRNA) regulatory networks that control cellular differentiation pathways and that play a key role in orchestrating gene expression within a central box. Through the combination of large-scale transcriptomic data and network modeling and pathway enrichment techniques, the results reveal lncRNAs as multi-layer regulators, shaping differentiation pathways by coordinating transcriptional, post-transcriptional, and epigenetic responses. The broader engagement of key hub lncRNAs in key signaling pathways, such as TGF-b/SMAD, WNT/bcatenin, and BMP, highlight the complexity and interconnectedness of regulatory systems that drive lineage specification. Notably, the work turns the conventional emphasis on regulating single-gene to a network-based perspective, where cellular differentiation is regulated by dynamic and integrated molecular interactions. These lessons offer a solid basis on the future of developmental biology research, and show promise in the future of regenerative medicine, disease modeling, and the creation of targeted and network based therapeutic approaches.

REFERENCES

Azad, F. M., Polignano, I. L., Proserpio, V., & Oliviero, S. (2021). Long noncoding RNAs in human stemness and differentiation. Trends in Cell Biology, 31(7), 542-555.

1. Bridges, M. C., Daulagala, A. C., & Kourtidis, A. (2021). LNCcation: lncRNA localization and function. Journal of Cell Biology, 220(2), e202009045.
2. Daneshvar, K., Pondick, J. V., Kim, B. M., Zhou, C., York, S. R., Macklin, J. A., & Mullen, A. C. (2016). DIGIT is a conserved long noncoding RNA that regulates GSC expression to control definitive endoderm differentiation of embryonic stem cells. Cell reports, 17(2), 353-365.
3. Ferrer, J., & Dimitrova, N. (2024). Transcription regulation by long non-coding RNAs: mechanisms and disease relevance. Nature Reviews Molecular Cell Biology, 25(5), 396-415.
4. Jiang, W., Liu, Y., Liu, R., Zhang, K., & Zhang, Y. (2015). The lncRNA DEANR1 facilitates human endoderm differentiation by activating FOXA2 expression. Cell reports, 11(1), 137-148.
5. Kopp, F., & Mendell, J. T. (2018). Functional classification and experimental dissection of long noncoding RNAs. Cell, 172(3), 393-407.
6. Landshammer, A., Bolondi, A., Kretzmer, H., Much, C., Buschow, R., Rose, A., & Meissner, A. (2023). T-REX17 is a transiently expressed non-coding RNA essential for human endoderm formation. Elife, 12, e83077.
7. Lu, P., Yang, J., Li, M., Wen, S., Zhang, T., Yan, C., & Jiang, W. (2023). A desert lncRNA HIDEN regulates human endoderm differentiation via interacting with IMP1 and stabilizing FZD5 mRNA. Genome Biology, 24(1), 92.
8. Luo, H., Lv, W., Tong, Q., Jin, J., Xu, Z., & Zuo, B. (2021). Functional non-coding RNA during embryonic myogenesis and postnatal muscle development and disease. Frontiers in cell and developmental biology, 9, 628339.
9. Mattick, J. S., Amaral, P. P., Carninci, P., Carpenter, S., Chang, H. Y., Chen, L. L., & Wu, M. (2023). Long non-coding RNAs: definitions, functions, challenges and recommendations. Nature reviews Molecular cell biology, 24(6), 430-447.
10. Rinn, J. L., & Chang, H. Y. (2020). Long noncoding RNAs: molecular modalities to organismal functions. Annual review of biochemistry, 89(1), 283-308.
11. Statello, L., Guo, C. J., Chen, L. L., & Huarte, M. (2021). Gene regulation by long non-coding RNAs and its biological functions. Nature reviews Molecular cell biology, 22(2), 96-118.
12. Wang, K. C., & Chang, H. Y. (2011). Molecular mechanisms of long noncoding RNAs. Molecular cell, 43(6), 904-914.
13. Yang, J., Lu, P., Li, M., Yan, C., Zhang, T., & Jiang, W. (2020). GATA6-AS1 regulates GATA6 expression to modulate human endoderm differentiation. Stem cell reports, 15(3), 694-705.