T.C.
DOKUZ EYLÜL UNIVERSITY

IZMIR INTERNATIONAL BIOMEDICINE AND GENOME INSTITUTE

FUNCTIONAL ANALYSIS OF LONG NON-CODING RNAs CRITICAL FOR THE MESENCHYMAL TO EPITHELIAL TRANSITION

AYÇA ÖZKAN

DEPARTMENT OF BIOMEDICINE AND HEALTH TECHNOLOGIES

Biomedicine and Health Technologies

MASTER OF SCIENCE THESIS

IZMIR-2024

THESIS CODE: DEU.IBG.MSc-2020850030

T.C.
DOKUZ EYLÜL UNIVERSITY

IZMIR INTERNATIONAL BIOMEDICINE AND GENOME INSTITUTE

FUNCTIONAL ANALYSIS OF LONG NON-CODING RNAS CRITICAL FOR THE MESENCHYMAL TO EPITHELIAL TRANSITION

AYÇA ÖZKAN

DEPARTMENT OF BIOMEDICINE AND HEALTH TECHNOLOGIES

Biomedicine and Health Technologies

MASTER OF SCIENCE THESIS

Supervisor: Assoc. Prof. Hani Alotaibi

This work was supported by
TÜBİTAK grant 219Z034

IZMIR-2023

THESIS CODE: DEU.IBG.MSc-2020850030

Dokuz Eylül University Izmir International Biomedicine and Genome Institute

Department of Biomedicine and Health Technologies Biomedicine and Health Technologies graduate program Master of Science student Ayça ÖZKAN

has successfully completed her Master of Science thesis titled

"FUNCTIONAL ANALYSIS OF LONG NON-CODING RNAS CRITICAL FOR THE MESENCHYMAL TO EPITHELIAL TRANSITION" on 19/01/2024.

Assoc. Prof. Hani Alotaibi

Dokuz Eylül University

CHAIR

Prof. Dr. Şerife Esra Erdal Bağrıyanık

Dokuz Eylül University

MEMBER

Prof. Dr. Güneş Özhan

Izmir Institute of Technology

MEMBER

Prof. Dr. Gökhan Karakülah

Dokuz Eylül University

SUBSTITUTE MEMBER

Asst. Prof. Ayşe Banu Demir

Izmir University of Economics

SUBSTITUTE MEMBER

Dokuz Eylül Üniversitesi İzmir Uluslararası Biyotıp ve Genom Enstitüsü

Genom Bilimleri ve Moleküler Biyoteknoloji Anabilim Dalı, Biyotıp ve Sağlık teknolojileri Yüksek Lisans programı öğrencisi, Ayça ÖZKAN

"**FUNCTIONAL ANALYSIS OF LONG NON-CODING RNAS CRITICAL FOR THE MESENCHYMAL TO EPITHELIAL TRANSITION**" konulu Yüksek Lisans tezini 19/01/2024 tarihinde başarılı olarak tamamlamıştır.

Doç. Dr. Hani Alotaibi

Dokuz Eylül Universitesi

BAŞKAN

Prof. Dr. Şerife Esra Erdal Bağrıyanık

Dokuz Eylül Üniversitesi

ÜYE

Prof. Dr. Güneş Özhan

İzmir Yüksek Teknoloji Enstitüsü

ÜYE

Prof. Dr. Gökhan Karakülah

Dokuz Eylül Üniversitesi

YEDEK ÜYE

Dr. Öğr. Üyesi Ayşe Banu Demir

İzmir Ekonomi Üniversitesi

YEDEK ÜYE

TABLE OF CONTENTS

TABLE OF CONTENTS V
LIST OF TABLES VII
LIST OF FIGURES VIII
LIST OF ABBREVIATIONS IX
ACKNOWLEDGMENTS XI
ABSTRACT XII
ÖZET XIII
1. INTRODUCTION 1
1.1. Statement and Importance of the Problem 1
1.2. Aim of Study 1
1.3. Hypothesis of Study 1
2. GENERAL INFORMATION 2
2.1. Cellular Plasticity 2
2.2. Epithelial-Mesenchymal Transition and Mesenchymal-Epithelial Transition 3
2.2.1 Epithelial-Mesenchymal Transition (EMT) 3
2.2.2. Mesenchymal Epithelial Transition (MET) 4
2.3 Long non-coding RNA 4
2.4. Preliminary data 6
3. MATERIALS AND METHODS 8
3.1. Type of Study 8
3.2. Time and Place of Study 8
3.3. Materials of Study 8
3.3.1 Cell Culture Materials, Reagents & Cell Lines 8
3.3.2 PCR, cDNA, and qPCR Reagents 9
3.3.3 SDS gel materials for IVTT Asay 10
3.3.4 Molecular Cloning Reagents 12
3.3.5 Bacterial Reagents 12
3.3.6 Plasmid Construction 13
3.3.7 siRNA Sequence Information 14
3.3.8 Immunofluorescence Labelling Reagents 14
3.3.9 Other Consumables and Commercial kits 14

3.3.10 Hardware & Machines 15
3.4. Variables of Study 16
3.5. Tools For Data Collection 16
3.6 Study Plan and Calendar 16
3.7 Data Evaluation 16
3.7.1 Cell Culture and EMT/MET Experiment Model 16
3.7.2 Total RNA isolation and cDNA synthesis 17
3.7.3 RT-PCR 17
3.7.4 Restriction digestion of PCR Products 18
3.7.5. In Vitro Transcription and Translation Assay 19
3.7.6 Enhancer Analysis 19
3.7.7 Construction of Reporter Vector Sequences 19
3.7.8 Luciferase Reporter Activity Measurement 20
3.7.9 siRNA Knockdown and qPCR 20
3.7.10 Immunofluorescence Labeling and Confocal Microscopy 21
3.8. Limitation of the Study 21
3.9. Ethics Committee Approval 21
4. RESULTS 22
4.1. Expression of lncRNAs 22
4.2. Genomic DNA Contamination Control 24
4.3. Restriction Digestion of PCR Products 24
4.4. Protein Coding Capacity Confirmation 25
4.5 Enhancer lncRNAs 26
4.6. Cloning of Selected Enhancer lncRNAs in pGL4.23 33
4.7. Luciferase Reporter Assay 34
4.8 siRNA Silencing of NH.3067 35
5. DISCUSSION 37
6. CONCLUSION AND SUGGESTIONS 40
7. REFERENCES 41
8. APPENDIX 46

LIST OF TABLES

Table 1. Hub transcription factors and the associated lncRNAs..... 7
Table 2. Cell culture reagents 9
Table 3. PCR, cDNA, and qPCR Reagents 9
Table 4. Recipe of TBE Buffer 10
Table 5. SDS gel materials 10
Table 6. Recipe for SDS-PAGE Gel..... 11
Table 7. Buffer and Solution Recipes List 11
Table 8. Molecular Cloning Reagents 12
Table 9. Bacterial Reagents..... 12
Table 10. Recipes of LB..... 13
Table 11. siRNA Sequences 14
Table 12. Immunofluorescence Labelling Reagents 14
Table 13. Other Chemicals, Consumables, and Commercial kits 14
Table 14. Hardware and machines used in this study..... 15
Table 15. Conditions for Taq and Q5 polymerases..... 17
Table 16. Components for PCR 18
Table 17. List of restriction enzymes..... 18
Table 18. PCR conditions for colony PCR 20
Table 19. Components for colony PCR..... 20
Table 20. List of RT-PCR Primers..... 46
Table 21. Primers used in colony PCR 47
Table 22. List of qPCR primers 47

LIST OF FIGURES

Figure 1. Identification and annotation of gene co-expression modules associated with EMT-MET stages. 6

Figure 2. Separation of lncRNAs expressed in the MET process into stages according to the RNA-Seq data we obtained from the 114Z245 project 7

Figure 3. EMT and MET on NMuMG cells by TGFβ3. 8

Figure 4. pGL4.23 plasmid map..... 13

Figure 5. Expression of selected lncRNAs in the EMT/MET process..... 22

Figure 6. Expression of selected lncRNAs in the EMT/MET process..... 23

Figure 7. Expression of unknown lncRNAs in the EMT/MET process..... 23

Figure 8. Control of genomic DNA contamination by RT-PCR..... 24

Figure 9. Restriction digestion results of PCR products..... 25

Figure 10. PCR results for IVTT..... 26

Figure 11. IVTT result on SDS-Page gel..... 26

Figure 12. Integrative approach which combinates H3K27Ac ChIP-seq data with RNA-seq data for elncRNA NH.3067 27

Figure 13. Integrative approach which combinates H3K27Ac ChIP-seq data with RNA-seq data for elncRNA NH.3934 28

Figure 14. Integrative approach, which combinates H3K27Ac ChIP-seq data with RNA-seq data for elncRNA NH.15444 29

Figure 15. An Integrative approach that combinates H3K27Ac ChIP-seq data with RNA-seq data for elncRNA NH.11835 30

Figure 16. Integrative approach which combinates H3K27Ac ChIP-seq data with RNA-seq data for elncRNA NH.19540 and NH.19543 31

Figure 17. Integrative approach which combinates H3K27Ac ChIP-seq data with RNA-seq data..... 32

Figure 18. Colony PCR of NH.15444, NH11835, NH.19540 and NH.14958 34

Figure 19. Luciferase results of novel lncRNAs showing strong regulator activity. 34

Figure 20. Silencing of NH.3067 in MET stage. 35

Figure 21. The siRNA-mediated silencing of NH.3067 the immunofluorescence staining of NMuMG cells in MET 36

Figure 22. Effective silencing of NH.3067 resulted in altered expression of epithelial and mesenchymal markers. 36

LIST OF ABBREVIATIONS

HF	High Fidelity
IVTT	in vitro transcription and translation assay
NMuMG	Normal Murine Mammary Gland
lncRNA	long non-coding RNA
EMT	epithelial to mesenchymal transition
MET	mesenchymal to epithelial transition
E-cad	Epithelial cadherin
N-cad	Neural cadherin
TGFβ	Transforming growth factor beta
PCR	Polymerase Chain Reaction
qRT-PCR	Quantitative Real-Time Polymerase Chain Reaction
WGCNA	Weighted gene co-expression network analysis
ATAC-seq	Assay for Transposase-Accessible Chromatin using sequencing

For my uncle Hasan and my grandmother Nadire Yeşiltuna,

I hope I made you proud.

ACKNOWLEDGMENTS

I would like to extend my gratitude to my advisor, Assistant Prof. Dr. Hani Alotaibi, who has been supportive throughout my master's education and always encouraged me to work for excellence. I couldn't have accomplished anything without his observations and suggestion. Thank you for your guidance, patience and support. I am profoundly thankful to him, particularly for introducing me to the incredible individuals in his lab. They have evolved beyond mere lab colleagues, transforming into my extended family.

I want to express my gratitude to all members of the Alotaibi lab, both form past and present. Specifically, I want to thank Büşra, Seray, and Şeyma for consistently providing unwavering support and comfort during times of need. They will be the sisters I seek when I face difficulties. I would like to thank Seray for her contribution more than anyone else during the thesis writing process. Also, Büşra for her contributions in calculations for experiments. I'd still be calculating the dilutions right now. I'd also like to express my gratitude to Zeynep, even though we know each other for a short time, you made me feel like I have known you all for years. Also, I would like to thank Şimal and Gkamze for their support. Furthermore, you have consistently supported me through the most challenging moments in my life, and I am endlessly grateful.

I want to express my gratitude to my best friends Özge and Gözde. Your support is truly invaluable. Merve, for always being there for me my free psychologist aka my friend to sister.

Lastly, I would like to express my gratitude to my parents, Nilgün and Aytekin, my brother Aytaç for their unconditional support. Your boundless love and unwavering support made me who I am now My cousin Cansu, for providing support and entertainment that never fails to bring a smile to my face. My aunt Dilek, for her academic knowledge and aunt Gülgün (Dilek) for her love.

This thesis was supported by TÜBİTAK grant number 219Z034.

Functional Analysis of Long Non-coding RNAs Critical for the Mesenchymal to Epithelial Transition

Ayça Özkan, İzmir International Biomedicine and Genome Institute, Dokuz Eylul University Health Campus, Balçova 35340 - Izmir / TURKEY

ABSTRACT

Epithelial-mesenchymal transition (EMT) and mesenchymal-epithelial transition (MET) are biological events crucial for physiological and pathological processes like normal development, wound healing, and tumorigenesis. Previously, MET was viewed merely as a reversible process of EMT, with decreased expression of EMT-inducing factors considered responsible for MET regulation. Recently, there has been an increase in research describing the role of non-coding RNAs in EMT. However, the role of lncRNA in MET is still unknown. There are novel lncRNAs that are MET-associated, which indicates their potential role in MET regulation similar to the EMT process. In this thesis, the aim was to identify lncRNAs that are critical for MET. We employed an integrated approach involving high-throughput data generation, gene modulation, and bioinformatics. Our results indicate that MET-associated lncRNAs play crucial roles in the cell cycle, chromosome segregation, and DNA damage repair and contribute to a regulatory network controlling MET. From our experiments, NH.3067 emerged as a critical initiator of MET, warranting further exploration through advanced experiments. Beyond its regulatory roles, this lncRNA holds promise as a potential cancer biomarker.

Keywords: MET, lncRNA, EMT, Transcription

Mezenkimal Epitelyal Dönüşüm Sürecinde Kritik Olan Uzun Kodlamayan RNA'ların İşlevsel Analizleri

Ayça Özkan, İzmir Uluslararası Biyotıp ve Genom Enstitüsü, Dokuz Eylül Üniversitesi Sağlık Yerleşkesi, Balçova 35340 - İzmir / TÜRKİYE

ÖZET

Epitelyal-mezenkimal dönüşüm (EMT) ve mezenkimal-epitelyal geçiş (MET), normal gelişim, yara iyileşmesi ve tümör oluşumu gibi çeşitli fizyolojik ve patolojik süreçler için çok önemli olan farklı biyolojik olaylardır. Daha önce MET, yalnızca EMT'nin tersinir bir süreci olarak görülüyordu; MET düzenlemesinden sorumlu olduğu düşünülen EMT'yi indükleyen faktörlerin ekspresyonu azalmıştı. Ayrıca EMT'de kodlamayan RNA'larla ilgili bazı yayınlar olmasına rağmen MET ile ilgili lncRNA'larla ilgili yayınlarda hala eksiklikler mevcut. EMT sürecine benzer şekilde MET düzenlemesindeki potansiyel rollerini gösteren MET ile ilişkili yeni lncRNA'lar vardır. Bu tezde amaç MET için kritik olan lncRNA'ları tanımlamaktı. Yüksek verimli veri üretimi, gen modülasyonu ve biyoinformatik içeren entegre bir yaklaşım kullandık; sonuçlarımız MET ile ilişkili lncRNA'ların hücre döngüsünde, kromozom ayrımında, DNA hasarı onarımında önemli roller oynadığını ve MET'i kontrol eden düzenleyici bir ağa katkıda bulunduğunu gösteriyor. Deneylerimizden NH.3067, MET'in kritik başlatıcısı olarak ortaya çıktı ve gelişmiş deneyler yoluyla daha fazla araştırması gerekiyor. Düzenleyici rollerinin ötesinde, bu lncRNA'lar potansiyel kanser biyobelirteçleri olarak umut vaat ediyor.

Anahtar kelimeler: MET, EMT, lncRNA, Transkripsiyon

1. INTRODUCTION

1.1. Statement and Importance of the Problem

The MET and EMT processes represent biological events crucial for physiological and pathological phenomena like normal development, wound healing, and tumorigenesis. Historically, MET was perceived solely as a reversible counterpart of EMT, with the belief that the regulation of MET primarily involved a decrease in the expression levels of EMT-inducing factors. However, recent investigations reveal that, unlike EMT, MET undergoes regulation through a transcriptional network with dynamic interactions within the chromatin infrastructure. While research on regulatory pathways and networks has traditionally focused on the activities of protein-coding genes, recent studies underscore the significance of the non-coding genome. Despite numerous publications on EMT-related long non-coding RNAs (lncRNAs), a similar understanding of MET-related lncRNAs is missing. Recognizing the considerable influence of long non-coding RNAs (lncRNAs) is essential to comprehend their impact on the EMT and MET processes.

1.2. Aim of Study

This project mainly aims to study the novel lncRNAs in mesenchymal to epithelial transition using an integrated approach, including bioinformatics and gene modulation, to understand the role of lncRNAs in MET with NMuMG cell line in vitro.

Aim:

To study the novel lncRNAs in mesenchymal to epithelial transition.

Objectives:

1. To define the lncRNAs by their expression and non-coding capacity.
2. To determine the effect on MET by using the RNAi-mediated gene silencing in MET.
3. To study the enhancer capacity of lncRNAs.

1.3. Hypothesis of Study

The study hypothesizes the presence of novel lncRNAs involved in regulating MET either by their expression or enhancer capacity.

2. GENERAL INFORMATION

2.1. Cellular Plasticity

Developing vertebrate embryos requires precise cellular differentiation processes that firmly regulate evolving cell adhesion and dynamic interactions among cells. Gene regulatory networks govern cellular phenotypes, and the dynamic alterations among molecules like genes, proteins, and microRNAs within these networks dictate the flexibility between a stable and variable cell phenotype.

A widely recognized mechanism influencing both cellular shape and the dynamic interactions between cells is cell adhesion, which is governed by E-cadherin. The regulated expression of the E-cadherin gene Cdh1 is crucial for the dynamic alterations and adaptability observed in this process. In biological events like epithelial-mesenchymal transition (EMT) and mesenchymal-epithelial transition (MET), the Cdh1 locus undergoes significant inactivation during EMT and subsequent reactivation during MET. It is established that transcription factors like Snai1/2, Twist1, and Zeb1/2, recognized as inducers of epithelial-mesenchymal transition (EMT), have the ability to inhibit the expression of E-cadherin during the EMT process. Beyond their suppressive actions on E-cadherin, these factors influence various EMT stages (Gupta et al., 2019). The complex nature of the epithelial-mesenchymal transition (EMT) process mirrors the diversity in the transcriptional networks overseeing this phenomenon and the signaling pathways essential for initiating and sustaining EMT.

A crucial signaling pathway in the epithelial-mesenchymal transition (EMT) is the TGFβ signaling pathway. In both adult and embryonic cells, TGFβ plays a critical role in fundamental cellular processes, including proliferation, differentiation, and apoptosis. In addition to Smad proteins, it is recognized that TGFβ can engage in signaling through alternative pathways like PI3K and MAPK. In embryonic morphogenesis events such as mesoderm formation and neural crest development, epithelial cells exhibit significant adaptability and can transform into mesenchymal cells through the epithelial-mesenchymal transition (EMT) process (Tsai and Yang, 2013).

In addition to the various biological occurrences in normal developmental processes, it is understood that the alterations in cell phenotype and cellular adaptability observed in the tumorigenesis process are also connected to the epithelial-mesenchymal transition (EMT) and mesenchymal-epithelial transition (MET) processes (Jolly et al., 2015). These processes govern the transformation between epithelial and mesenchymal states (Liao and Yang, 2017). In the context of tumorigenesis, TGFβ-dependent cellular reactions exhibit distinctions. Specifically, in the epithelial-mesenchymal transition (EMT), TGFβ enhances cell mobility, promotes

invasion, and facilitates metastasis. In this process, the reduction of E-cadherin expression leads to the accumulation of β-catenin in the cytoplasm and nucleus, triggering the activation of Wnt target genes (Jayachandran et al., 2016).

2.2. Epithelial-Mesenchymal Transition and Mesenchymal-Epithelial Transition

Epithelial-mesenchymal transition (EMT) and mesenchymal-epithelial transition (MET) are vital biological processes that induce adaptable alterations in epithelial cells. Epithelial cells establish cohesive cell structures with apical-basal polarity, tightly adhering to one another (Tan et al., 2015). In contrast, mesenchymal cells lack polarity and exhibit individual movement capabilities as they do not form intercellular connections (Brabletz, 2012). Epithelial cells maintain close associations with neighboring cells and the membrane. In contrast, mesenchymal cells, lacking apicobasal polarity, exhibit limited physical interactions with neighboring cells and are more predisposed to cell migration (Shu and Pei, 2014).

2.2.1 Epithelial-Mesenchymal Transition (EMT)

Epithelial-mesenchymal transformation (EMT) is a reversible biological event that characterizes the transition of cells from a polarized, less invasive state to a more invasive, non-polarized mesenchymal state (Diepenbruck and Christofori, 2016). This process holds significance in both embryogenesis and the development of cancer (Craene and Berx, 2013).EMT is categorized into three classes. The first type, Type 1 EMT, is essential for the formation of mesoderm, neural crest, heart valves, palate, and muscles in the developmental process (Thiery et al., 2009). Type 2 EMT is triggered in response to tissue damage, particularly during fibrosis and wound healing. Type 3 EMT is activated in the context of tumorigenesis, specifically during the dissemination and metastasis of cancer cells (Brabletz, 2012; Kalluri and Weinberg, 2009).

EMT is a mechanism in which the suppression of E-cadherin expression leads to the loss of cellular polarity, adherent morphology, and epithelial gene expressions. This is facilitated by the activation of transcription factors, including Snai1/2, Twist1, and Zeb1/2, commonly known as EMT inducers (Tan et al., 2015). These transcription factors play a role in inhibiting the expression of epithelial genes like E-cadherin while promoting the activation of mesenchymal genes such as N-cadherin (Diepenbruck and Christofori, 2016; Stemmler, 2008; Thiery et al., 2009). Various signaling pathways, including TGFβ, Wnt, and Notch signaling pathways, contribute to the initiation of EMT. The TGFβ signaling pathway is particularly identified as the principal signaling pathway in EMT, as it directly activates key EMT regulators like Snail, Slug, Twist, and Zeb1/2. The regulation of EMT involves numerous steps influenced by epigenetic modifications, transcriptional and post-translational regulations,

alternative splicing, and the involvement of non-coding RNAs (Craene and Berx, 2013; Lamouille et al., 2014; Nieto et al., 2016).

2.2.2. Mesenchymal Epithelial Transition (MET)

MET controls morphogenetic processes in embryonic development and is crucial for the metastasis and colonization of tumor cells in distant locations (Craene and Berx, 2013). Moreover, MET plays a critical role in the early stages of the induction of pluripotent stem cells (iPSC) from fibroblasts in somatic reprogramming (Li et al., 2010; Polo and Hochedlinger, 2010; Shu and Pei, 2014). While traditional definitions characterize the MET process as the opposite of the EMT process, recent studies underscore the intricate and dynamic nature of the MET process.

The expression of E-cadherin, suppressed in the EMT process, promptly reverts to its original state during the MET process. At the same time, there is an upswing in the expression of epithelial genes, coupled with a decline in the expression of mesenchymal genes. This sequence facilitates the restoration of cells' apicobasal polarity. Reactivation of E-cadherin expression is critical for MET. At the beginning of the MET process, forced suppression of Cdh1 expression may cause cells to remain in a mesenchymal state (Alotaibi et al., 2015; Stemmler, 2008; Stemmler et al., 2003)(Lamouille et al., 2014).

Like the EMT process, the MET process is subject to regulation by various signaling pathways, transcription factors, and epigenetic modifications. The specific parameters controlling these processes may vary under physiological and pathological conditions.

2.3 Long non-coding RNA

The cellular transcriptome contains various RNA molecules, including mRNA, rRNA, and tRNA (Karapetyan et al., 2013). The non-coding RNAs, constituting almost 95% of the total cellular RNA and once labeled as 'junk' RNA, are now recognized as essential epigenetic and transcriptional regulators, contrary to long-standing beliefs (Diamantopoulos et al., 2018). Non-protein-coding RNAs are categorized into two primary groups: short non-coding RNAs, such as microRNAs (miRNAs), and long non-coding RNAs (lncRNAs) (McCabe and Rasmussen, 2021). MicroRNAs are extensively understood for their ability to trigger mRNA degradation and inhibit mRNA translation through the RNA interference pathway (Hubé and Francastel, 2018). In contrast, there is limited information regarding the functions of long non-coding RNAs (lncRNAs) compared to microRNAs (Uszczynska-Ratajczak et al., 2018).

Long non-coding RNAs (lncRNAs) are commonly accepted as RNA molecules longer than 200 nucleotides in length, lacking protein-coding potential, or having the capacity to encode peptides shorter than 100 amino acids (Cao, 2014; M. Chen et al., 2015; He et al., 2019; Youness and Gad,

2019). Long non-coding RNAs (lncRNAs) are present in animals, plants, yeast, prokaryotes, and even viruses (Ma et al., 2013). Long non-coding RNAs (lncRNAs) share numerous characteristics with mRNAs. Key similarities include their exon-intron structure, transcription by RNA polymerase II, post-transcriptional processing processes (such as splicing), and sequence content (Gugnoni and Ciarrocchi, 2019). Despite these resemblances, they exhibit lower conservation across species, and their expression levels are generally lower than those of mRNAs (Ma et al., 2013). Long non-coding RNAs (lncRNAs) can be categorized based on various criteria (St.Laurent et al., 2015). The primary classification criteria include transcript length, association with identified protein-coding genes, correlation with DNAs possessing known functions and/or biochemical pathways, sequence and structural conservation, and connections with intracellular structures. Additionally, classifications can be made based on their functions. When categorizing lncRNAs according to their functions, four prominent classes emerge:

lncRNAs with enhancer function (non-coding (ncRNA) activating lncRNAs)
miRNA (microRNA) precursor transcripts
piRNA (Piwi interacting RNA) precursor transcripts
ceRNAs (competing endogenous RNAs).

The expression of lncRNAs with enhancer (eRNA) function is predominantly associated with the mRNA of the enhancer's target gene (Font-Cunill et al., 2018). It is established that eRNA transcripts can also regulate the activity of target genes in numerous instances.

The significance and variety of biological functions carried out by lncRNAs, which comprise the majority of non-coding RNAs, are too substantial to ignore. The primary reason behind this importance is the potential of lncRNAs to regulate gene expression. It is established that, through their interactions with DNA, RNA, and/or proteins, they can serve as signals for transcription, act as regulators influencing the expression of transcription factors, guide chromatin-modifying enzymes to their target genes, and function as structural frameworks that bring together various proteins to create functional ribonucleoprotein complexes (Ma et al., 2013). It has been identified that specific lncRNAs are involved in modulating the expression levels of protein-coding RNAs, while others can alter gene expression patterns by influencing the chromatin infrastructure through the formation of complexes with chromatin regulatory proteins (Böhmdorfer and Wierzbicki, 2015; Dykes and Emanueli, 2017; Karapetyan et al., 2013; Marchese et al., 2017). Given that non-coding RNAs can act as trans regulators through intermolecular interactions, it is prevalent for lncRNAs to associate with enhancer, promoter, and intron regions or exert their function by acting on the antisense strand of a recognized gene (Rutenberg-Schoenberg et al., 2016).

2.4. Preliminary data

We first utilized data from RNA-seq bioinformatic analyses previously conducted by Doğa Eskier within the TÜBİTAK grant number 114Z245 frame. This data included the Identification and Annotation of Putative Long Non-coding RNAs expressed during the Mesenchymal-Epithelial Transition. From this analysis, there are 1397 lncRNA that are related to EMT-MET processes. Because of this high number, some selections were made based on different criteria. With the help of Weighted Gene Co-expression Network Analysis (WGCNA), lncRNAs were classified into different modules (Figure 1). Subsequently, the obtained modules were examined regarding the relationship between lncRNA and transcription factors. Due to the large number of genes (transcription factors and lncRNAs), genes with a minimum kME score of 0.95 (a calculation indicating an eigen-based module relationship) were explored, and a 'hub gene' list was generated. Hub genes exhibit a high connectivity (relationship) level within their respective modules. In Table 1, unidentified lncRNAs and the transcription factors believed to be associated with them are shown. Also, these lncRNAs are divided into stages of MET, which are early, middle, and late stages (Figure 2).

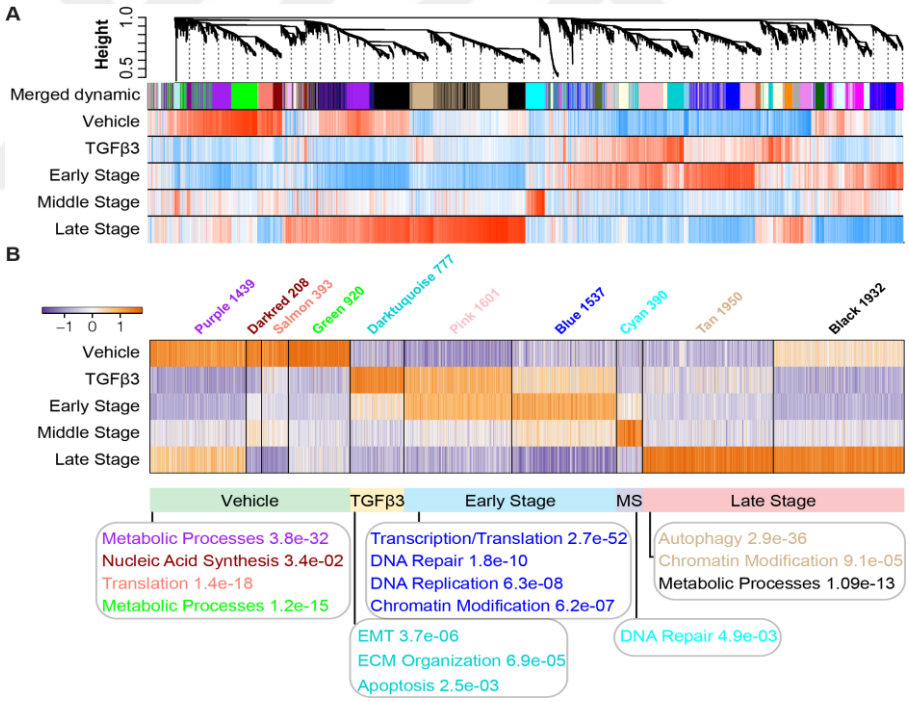

Figure 1. Identification and annotation of gene co-expression modules associated with EMT-MET stages. A. Dendrogram, module membership, and trait (stage) correlation heatmap of genes transcribed during the EMT-MET process. Each branch of the dendrogram indicates a single gene. Blue cells in the heatmap indicate a negative correlation, red cells indicate a positive correlation and white cells indicate zero correlation. B. Z-score heatmap of the expression values of genes in the most strongly associated module for each EMT-MET stage and the biological process GO enrichment summaries of each module.

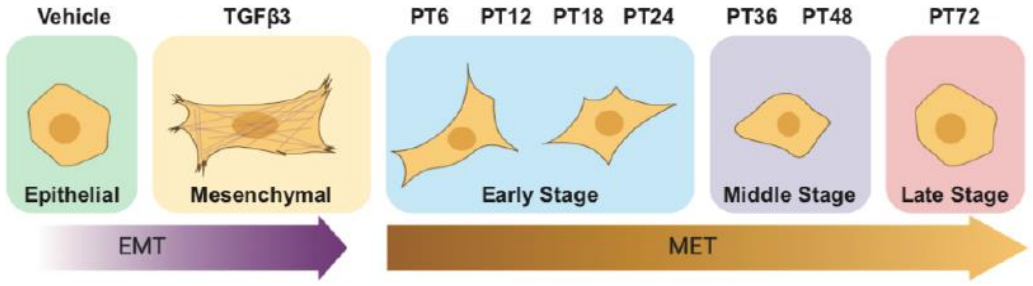

Figure 2. Separation of lncRNAs expressed in the MET process into stages according to the RNA-Seq data we obtained from the 114Z245 project

Table 1. Hub transcription factors and the associated lncRNAs

stages	Hub TF	kME score	Novel lncRNAs	kME score	Novel lncRNAs	kME score
Vehicle	Irf6	0.9832866	NH.14622	0.9859	NH.3067	0.9641
	Cebp-α	0.9812477	NH.18336	0.9853	NH.6937	0.9629
	Hnf4a	0.9706679	NH.8369	0.9784	NH.20522	0.9613
			NH.16434	0.9747	NH.4846	0.9613
			NH.23801	0.9745	NH.9741	0.9602
			NH.7062	0.9710	NH.20519	0.9575
			NH.14621	0.9686	NH.17858	0.9560
			NH.8918	0.9684	NH.13885	0.9557
			NH.9740	0.9674	NH.11898	0.9548
			NH.9401	0.9651	NH.15998	0.9523
Early stage	Prdm4	0.9831504	NH.3259	0.9881	NH.20155	0.9574
	Etv4	0.9608307	NH.13502	0.9712	NH.7756	0.9561
	Smarca5	0.9569540	NH.1020	0.9640	NH.20938	0.9538
			NH.10286	0.9616	NH.14182	0.9528
			NH.23540	0.9608	NH.5736	0.9528
			NH.8180	0.9606	NH.8144	0.9516
			NH.20253	0.9592	NH.15614	0.9505
Late stage	Tfcp2l1	0.9561734	NH.5041	0.9812	NH.10008	0.9571
	Sox9	0.9517030	NH.14677	0.9782	NH.202	0.9568
			NH.6969	0.9742	NH.2	0.9563
			NH.184	0.9735	NH.26246	0.9563
			NH.10006	0.9720	NH.26239	0.9562
			NH.16699	0.9713	NH.1266	0.9561
			NH.24922	0.9705	NH.16130	0.9558
			NH.13505	0.9686	NH.22652	0.9556
			NH.8346	0.9649	NH.26237	0.9554
			NH.834	0.9649	NH.24831	0.9549
			NH.3949	0.9643	NH.16700	0.9543
			NH.14687	0.9642	NH.6921	0.9542
			NH.25506	0.9629	NH.15088	0.9542
			NH.14607	0.9627	NH.25474	0.9538
			NH.965	0.9616	NH.10007	0.9537
			NH.11114	0.9602	NH.18716	0.9536
			NH.13439	0.9593	NH.16575	0.9534
			NH.14608	0.9580	NH.20173	0.9528
			NH.5757	0.9572	NH.26269	0.9526

3. MATERIALS AND METHODS

3.1. Type of Study

This research study is experimental.

3.2. Time and Place of Study

This study was performed in Alotaibi Transcriptional Regulation Laboratory in Izmir Biomedicine and Genome Center between March 2021 and January 2024.

3.3. Materials of Study

3.3.1 Cell Culture Materials, Reagents & Cell Lines

All the experiments used the NMuMG cell line procured from ATCC (ATCC® CRL-1636) in 2017. This cell line is the most commonly used model for EMT-MET studies because morphological changes can be observed easily (Miettinen et al., 1994). EMT and MET studies can be done by treating cells with TGFβ. With TGFβ treatment, cells undergo EMT by downregulating Cdh1 and upregulating N-cadherin and vimentin, which are mesenchymal genes. Also, actin stress fiber can be seen in the mesenchymal state of cells. When TGFβ is removed from the environment, cells undergo MET, an epithelial state (Figure 3).

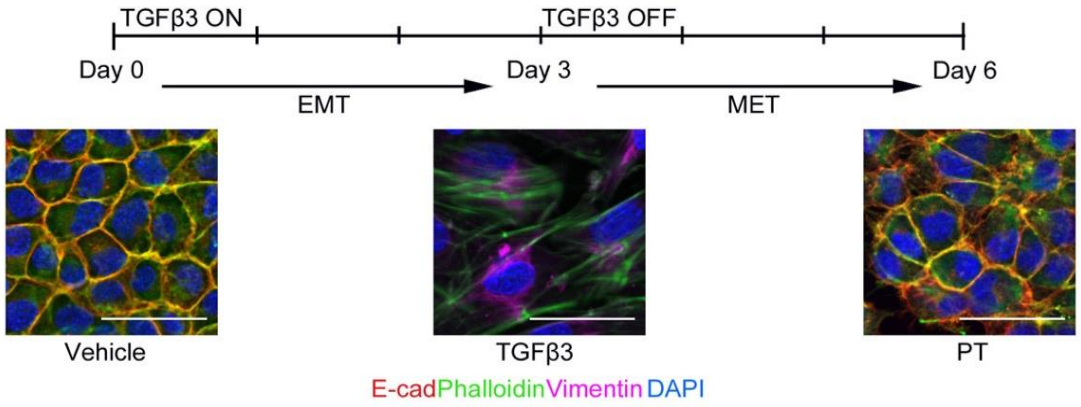

Figure 3. EMT and MET on NMuMG cells by TGFβ3. Red represents E-cadherin on the cell membrane. Green represents actin fibers by phalloidin staining. Blue represents the nucleus by DAPI staining. In vehicle control, cells are in the epithelial state, and E-cad can be seen on the cell membrane. TGFβ3 cells for 3 days are in a mesenchymal state, and actin fibers can be seen as long for the mesenchymal phenotype. Also, vimentin, a mesenchymal marker, can be seen in purple in TGFβ3 cells for 3 days. There is no visible red color of E-cad in TGFβ3 cells for 3 days. When removal of TGFβ3 for 3 days, cells can be seen in the epithelial state, and E-cad can be seen in the cell membrane. Actin fibers can be seen as in the epithelial phenotype.

Table 2. Cell culture reagents

Materials	Cat. No	Vendor
DMEM, high glucose	41965039	Gibco
Trypsin-EDTA (0.25%), phenol red	25-200-056	Gibco
DPBS, no calcium, no magnesium	14190094	Gibco
Fetal Bovine Serum, qualified, Brazil	10270106	Gibco
MEM Non-Essential Amino Acids Solution (100X)	11140035	Gibco
Sodium pyruvate 100 mM	p04-43100	PAN-Biotech
Penicillin-Streptomycin (10,000 U/mL)	15140122	Gibco
FuGENE® HD Transfection Reagent	E2311	Promega
Lipofectamine™ RNAiMAX Transfection Reagent	13778150	Invitrogen
DMSO	A3672,0100	Applichem

3.3.2 PCR, cDNA, and qPCR Reagents

Table 3. PCR, cDNA, and qPCR Reagents

Materials	Cat. No	Vendor
Taq DNA Polymerase with Standard Taq (Mg-free) Buffer	M0320S	NEB
Q5® High-Fidelity DNA Polymerase	M0491L	NEB
Deoxynucleotide (dNTP) Solution Mix	N0447S	NEB
AMPIGENE® qPCR Probe Mix No-ROX	ENZ-NUC107-1000	ENZO
MicroAmp™ Fast Optical 96-Well Reaction Plate with Barcode, 0.1 mL	4346906	Applied Biosystems
MicroAmp™ Optical Adhesive Film	4360954	Applied Biosystems

Table 4. Recipe of TBE Buffer

Reagents for TBE	Recipe (for 10x)
Tris	108 gr
Boric acid	55 gr
0.5 M pH:8 EDTA	40 ml
ddH2O	in 1 L

3.3.3 SDS gel materials for IVTT Asay

Table 5. SDS gel materials

Materials	Cat. No	Vendor
30% Acrylamide bis	A3574-100ML	Sigma
Sodium dodecyl sulfate (SDS)	74255-250G	Sigma
APS	161-0702	Bio-Rad
TEMED	161-0800	Bio-Rad
β-Mercaptoethanol	60-24-2	Sigma
Glacial acetic acid	64-19-7	ISOLAB
PageRuler Plus Prestained Protein Ladder	26619	Thermo Scientific
Trizma base	T1503-1KG	Sigma
coomaasie brilliant blue	A1092	Applichem
methanol	32213-2,5L	Sigma
bromophenol blue	115-39-9	Applichem
glycine	453807	Carlo ERBA
Color Prestained Protein standard, broad range	P7719S	NEB
IVTT	E6800	NEB

Table 6. Recipe for SDS-PAGE Gel

Reagents	stacking gel (5%)	resolving gel (12%)
ddH2O	4.1 ml	3.3 ml
30% Acyrlamide bis	1 ml	4 ml
1 M pH:6.8 Tris	750 µl	---
1.5 M pH:8.8 Tris	----	2.5 ml
10% SDS	60 µl	100 µl
10% APS	60 µl	100 µl
TEMED	6 µl	4 µl
Total volume	6 ml	10 ml

Table 7. Buffer and Solution Recipes List

Reagent	Recipe
Laemmli buffer (4x)	1.514 g Tris base
	4 g SDS
	20 ml glycerol
	10 ml β-mercaptoethanol
	200 mg bromophenolblue
	50 ml total
Staining buffer	450 ml methanol
	400 ml ddH2O
	100 ml acetic acid
	2 g Coomassie brilliant blue
	1L Total
Destaining buffer	500 ml methanol
	400 ml ddH2O
	100 ml acetic acid
	1 liter total
Running buffer	15.1 g Tris base
	94 g Glycine
	50 ml 10 %SDS
	ddH2O up to 1 L

3.3.4 Molecular Cloning Reagents

Table 8. Molecular Cloning Reagents

Materials	Cat. No	Vendor
HindIII-HF	R3104S	NEB
XhoI	R0146S	NEB
BglII	R0144S	NEB
NcoI-HF	R3193S	NEB
KpnI-HF	R3142S	NEB
SacI	R3156S	NEB
BsaI-HF	R3733S	NEB
NheI-HF	R3131S	NEB
SmaI	R0141S	NEB
NotI-HF	R3189S	NEB
T4 DNA Ligase	M0202S	NEB

3.3.5 Bacterial Reagents

Table 9. Bacterial Reagents

Materials	Cat. No	Vendor
Ampicillin Sodium Salt BioChemica	A0839,0010	AppliChem
Yeast Extract	Y1625-250G	Sigma
Agar	05039-50G	Sigma
Tryptone BioChemica	1553.0500	AppliChem
NaCl	1.06404.1000	Merck
NucleoBond Xtra Maxi kit for transfection-grade plasmid DNA	740414.10	Macherey-Nagel (MN)

Table 10. Recipes of LB

reagents	recipe for LB Agar	recipe for LB Broth
Tryptone	10 g	10 g
Yeast extract	5 g	5 g
Agar	15 g	---
Sodium Chloride	5 g	5 g
ddH2O	in 1 L	in 1 L

3.3.6 Plasmid Construction

NH.3067, NH.14958, NH.3934, NH.15444, NH.19540, NH.19543, NH.11835, NH.13885 genes were cloned into the luciferase reporter vector pGL4.23 (Promega).

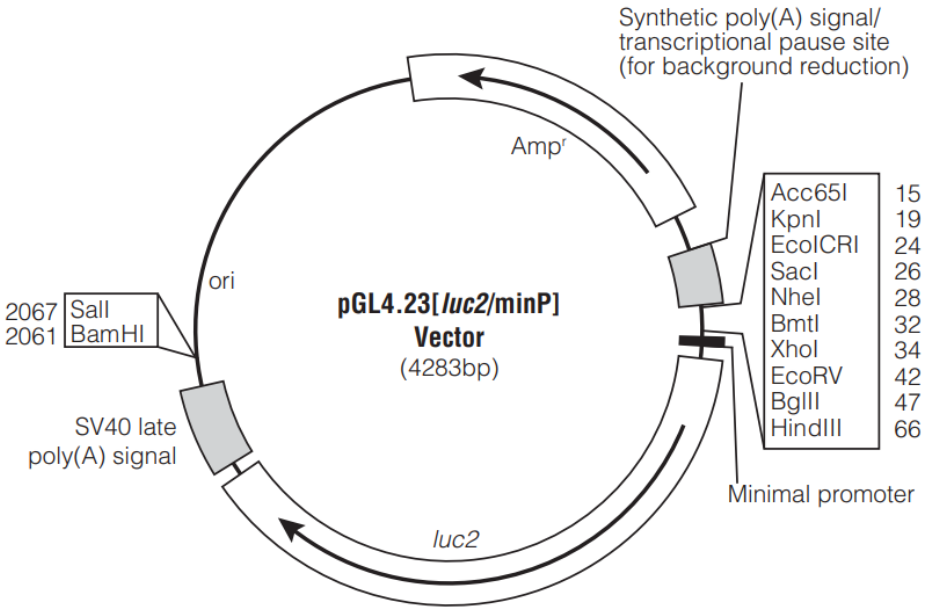

Figure 4. pGL4.23 plasmid map.

3.3.7 siRNA Sequence Information

Table 11. siRNA Sequences

siRNA name	sense strand	antisense strand
siRNA_3067-1	5'CCUUUAAAUAAUGGAUCAGAAUUCT3'	5'AGAAUUCUGAUCCAUUAUUUAAAGGAU3'
siRNA_3067-2	5'CAACCUUAUGUUCAUCAGCAUCUTT3'	5'AAAGAUGCUGAUGAACAUAAGGUUGAC3'
siRNA_3934-1	5'ACUAAGACGAGACAUUCACUAGAAA3'	5'UUUCUAGUGAAUGUCUCGUCUUAGUCG3'
siRNA_3934-2	5'AGCUAGCUCUCCAGUGAAGACUGGG3'	5'CCCAGUCUUCACUGGAGAGCUAGCUCG3'

3.3.8 Immunofluorescence Labelling Reagents

Table 12. Immunofluorescence Labelling Reagents

Materials	Cat. No	Vendor
Formaldehyde	F8775	Sigma
Triton X-100	T8787	Sigma
E-cadherin antibody	610181	BD
Alexa594-conjugated goat anti-mouse IgG	A-11005	Invitrogen
DAPI	D1306	Invitrogen
Alexa Fluor 488 Phalloidin	A12379	Invitrogen
Prolong Diamond Antifade Mountant solution	P36962	Thermo Fisher Scientific

3.3.9 Other Consumables and Commercial kits

Table 13. Other Chemicals, Consumables, and Commercial kits

Materials	Cat. No	Vendor
NucleoSpin RNA, Mini kit for RNA purification	740955.50	Macherey-Nagel (MN)
NucleoSpin Gel and PCR Clean-up, Mini kit for gel extraction and PCR clean-up	740609.50	Macherey-Nagel (MN)
Dual-Glo® Luciferase Assay System	E2940	Promega
GeneRuler 1 kb DNA Ladder, ready-to-use	SM0313	Thermo Scientific
GeneRuler 50 bp DNA Ladder	SM0372	Thermo Scientific
Agarose	A9536-500G	Sigma

Absolute Ethanol	9.200.262.500	Isolab
Ultrapure DNAse/RNAse free distilled water	10977035	Thermo Scientific
2- Propanol	961.023.2500	Isolab
Safe View	G108	ABM
Boric acid	10043-35-3	Merck
EDTA	E5134-500G	Sigma
luciferase assay plate	655075	Greiner
Recombinant Human TGF-β3	100-36E-10UG	PEPROTECH

3.3.10 Hardware & Machines

Table 14. Hardware and machines used in this study

Machine	**vendor**
pH Meter	HANNA Instruments
Toploader Balance 820 g x 0.01g	Sartorius
SimpliAmp™ Thermal Cycler	Applied Biosystems
Owl™ EasyCast™ B1A Mini Gel Electrophoresis Systems	ThermoFisher Scientific
microwave oven	Beko
mySPIN™ 6 Mini Benchtop Centrifuge	ThermoFisher Scientific
Vortex Mixer	ThermoFisher Scientific
Hotplate Stirrer Thermolyne	ThermoFisher Scientific
Waterbath	Nuve
Centrifuge 5810R	Eppendorf
Centrifuge MicroCL 17R	ThermoFisher Scientific
Incubator MaxQ 4000	ThermoFisher Scientific
ABI 7500 Fast Thermal Cycler	ThermoFisher Scientific
NanoDrop 2000	ThermoFisher Scientific
Centro Microplate Luminometer XS3 LB960	Berthold technologies
Gel Doc XR+ System with Image Lab Software	Bio-Rad

3.4. Variables of Study

The time points of the NMuMG cell model of EMT and MET are independent variables. The expression values of identified novel lncRNAs and their enhancer capacity are dependent variables. Dependent variables include the siRNAs used in silencing research gene enrichment profiles in NGS methods, which include RNA-seq, ATAC-seq, and ChIP-seq FANTOM-5 Cage-seq.

3.5. Tools For Data Collection

All data was collected by Luminometer, confocal microscopy, and phase contrast microscopy. Calculations were done in Prism graphpad.

3.6 Study Plan and Calendar

Mar 2021- Sep 2021	Literature Search for MET and lncRNAs
Oct 2021- Mar 2022	Expression validation of lncRNAs by PCR and restriction digestion
Apr 2022- Jul 2022	IVTT assays for enhancers
Aug 2022- Sep 2022	lncRNA enhancer analysis
Oct 2022- Mar 2023	Molecular cloning experiments for enhancer lncRNAs
Apr 2023- Jun 2023	luciferase assay for enhancer lncRNAs
Jul 2023- Sep 2023	siRNA knockdown experiments
Oct2023- Nov2023	site-directed mutagenesis
Nov 2023- Dec 2023	luciferase assay for site directed mutagenesis
Dec 2023- Jan2024	writing

3.7 Data Evaluation

3.7.1 Cell Culture and EMT/MET Experiment Model

NMuMG cells were used in the experiments. Cells were cultured with Gibco high glucose DMEM, 10% FBS, 1% penicillin-streptomycin, 1% non-essential amino acids, and 1% sodium pyruvate. The growth conditions of these cells are 10% CO2 and 37°C. EMT and MET studies were done in NMuMG cells by treating with 2.5 μM of TGF-β3. TGFβ3 was dissolved in a buffer containing pure water with a final concentration of 7.5% BSA and 1.2 M

HCl and stored at -80°C. Treating cells with TGFβ3 for 72 hours caused EMT, and an additional 72 hours without TGFβ3 caused MET on NMuMG cells. Samples were collected from non-treated cells with TGFβ which are named Vehicle, 72 hours TGF-β3 treated cells, which are TGFβ372, 72 hours TGFβ3 treated, and additional 24, 4,8 and 72 hours without TGFβ3 treated cells, which are PT24, PT48, and PT72, respectively.

3.7.2 Total RNA isolation and cDNA synthesis

Cells grown in six-well cell culture dishes were isolated using the "Nucleospin RNA II kit (Macherey-Nagel GmbH & Co.)" as specified in the user manual and then measured using the "NanoDrop 2000" (Thermo Scientific) device. Using 1 ug total RNA, cDNA synthesis was performed with the "Maxima First Strand cDNA Synthesis kit for RT-qPCR" (Thermo Scientific) as specified in the user manual.

3.7.3 RT-PCR

All PCR reactions were performed using the "Applied Biosystems/SimpliAmp thermal cycler" device with Taq polymerase or Q5 polymerase polymerases. The most used conditions are in Table 15, and the reaction components in Table 16. After amplification was completed, each gene was checked in agarose gel.

Table 15. Conditions for Taq and Q5 polymerases

Taq polymerase		**Q5 polymerase**	
95°C 5 min	1 cycle	98°C 30 sec	1 cycle
95°C 30 sec 58°C 15 sec 72°C 40sec	40 cycles	98°C 30 sec 58°C 15 sec 72°C 40sec	35 cycles
72°C 5min	1 cycle	72°C 5min	1 cycle

Table 16. Components for PCR

Taq polymerase			
Reaction mixture (1.5 mM MgCl2)	**Amount**	**Reaction mixture (3 mM MgCl2)**	**Amount**
H2O	16.25 µl	H2O	14.75 µl
10X Reaction Buffer	2.5 µl	10X Reaction Buffer	2.5 µl
25 mM MgCl2	1.5 µl	25 mM MgCl2	3 µl
dNTP	0.5 µl	dNTP	0.5 µl
Forward Primer	1 µl	Forward Primer	1 µl
Reverse Primer	1 µl	Reverse Primer	1 µl
Taq polymerase	0.25 µl	Taq polymerase	0.25 µl
cDNA (1:10 diluted)	2 µl	cDNA (1:10 diluted)	2 µl
Total	25 µl	Total	25 µl
Q5 Polymerase			
Reaction mixture (2 mM MgCl2)	**Amount**	**Reaction mixture (3 mM MgCl2)**	**Amount**
H2O	14.75 µl	H2O	13.75 µl
5X Reaction Buffer	5 µl	5X Reaction Buffer	5 µl
		25 mM MgCl2	1 µl
dNTP	0.5 µl	dNTP	0.5 µl
Forward Primer	1.25 µl	Forward Primer	1.25 µl
Reverse Primer	1.25 µl	Reverse Primer	1.25 µl
Q5 polymerase	0.25 µl	Q5 polymerase	0.25 µl
cDNA (1:10 diluted)	2 µl	cDNA (1:10 diluted)	2 µl
Total	25 µl	Total	25 µl

3.7.4 Restriction digestion of PCR Products

10 µl of PCR product was digested with 0.5 µl of the appropriate enzyme (Table 17) for 3 hours at the indicated temperature. Then, products were loaded on agarose gel for visualization.

Table 17. List of restriction enzymes

Gene ID	**Restriction enzyme**	**Incubation temperature**
NH.3067	SmaI	25°C
NH.15444	NcoI	37°C
NH.13885	XbaI	37°C
NH.3934	NheI	37°C
NH.14958	BsaI	37°C
NH.2	HindIII	37°C
NH.184	SacI	37°C
NH.26267	NcoI	37°C
NH.963	NcoI	37°C
NH.10007	NcoI	37°C
NH.10008	NcoI	37°C

3.7.5. In Vitro Transcription and Translation Assay

PURExpress In Vitro Protein Synthesis kit was used. Real-time PCR results concentration was measured using ThermoFisher Scientific's NanoDrop 2000. It was transferred to Eppendorf at 250 ng. 10 µl of Solution A and 7.5 µl of Solution B were added. The reaction was completed with 25 µl of water. For the DHFR positive control, 5.5 µl of water was added from 2 µl of control plasmid. It was incubated for 2 hours at 37°C. Then, it was run on an acrylamide gel and incubated with Coomassie Brilliant Blue Solution for 30 minutes at room temperature. It was incubated with a destaining solution for 30 minutes at room temperature. This process continued until the blue color of the gel disappeared.

3.7.6 Enhancer Analysis

To avoid potential bias arising from the density of enhancer markers in promoter regions when utilizing the ROSE enhancer identification algorithm, gene region annotations for the mm10 mouse genome were acquired in UCSC RefSeq format. The narrow peak BED files obtained through MACS2 were converted into a 9-column GFF format to make them suitable for analysis by the ROSE tool. The ROSE tool was executed using the command “ROSE_main.py -g MM10 -i <peaks.gff> -r <peaks.bam> -c <input.bam> -o <output> -s 12500 -t 2500”. Peaks situated within 12,500 base pairs of each other and more than 2,500 nucleotides away from the annotated transcription start points underwent merging. Following the merging process, reads in the peaks.bam file were normalized based on input.bam, and the signals in the peak regions were sorted according to RPM (reads per million) values. The intersections of the identified enhancer regions with the expressed lncRNA regions in the examples were computed using the “bedtools intersect -wo -a <ROSE_allStitched_output.bed> -b <annotated_transcriptome.bed>” command. The results were organized and presented in a format compatible with the UCSC Genome Browser.

3.7.7 Construction of Reporter Vector Sequences

NH.3067, NH.3934, NH.14958, NH.11835, NH.13885, NH.15444, NH.19540 and NH.19543 genes were cloned into pGL4.23 vector (Promega) to investigate the enhancer capability of NH.3067, NH.3934, NH.14958, NH.11835, NH.13885, NH.15444, NH.19540 and NH.19543 genes. Genes were amplified with Taq and Q5 polymerases. Conditions and reaction mixture were the same as in RT-PCR. First, the amplified fragment had enzymatic cleavage, and then they were inserted into the pGL4.23 vector; the resulting clones were confirmed by colony PCR. Mini culture and miniprep were done. Positive colonies are sent to Sanger sequencing. The enhancer capacity of genes was analyzed by the Dual-Glo Luciferase Test System.

Table 18. PCR conditions for colony PCR

95°C 5 min	1 cycle
95°C 30 sec 58°C 15 sec 72°C 40sec	40 cycles
72°C 5min	1 cycle

Table 19. Components for colony PCR

Reaction mixture (1.5 mM MgCl2)	**Amount**
H2O	17.25 µl
10X Reaction Buffer	2.5 µl
25 mM MgCl2	1.5 µl
dNTP	0.5 µl
Forward Primer	1 µl
Reverse Primer	1 µl
Taq polymerase	0.25 µl
Ligation product	1 µl
Total	25 µl

3.7.8 Luciferase Reporter Activity Measurement

NMuMG cells were seeded on a 24-well cell culture plate and were transfected with the plasmid DNA using the “Fugene HD” transfection reagent. Transfection experiments were performed with 200 ng DNA containing 2.5 ng pRL-TK. Firefly luciferase reporter values of cells were measured using a luminometer. Both expression plasmid and empty vector were measured and normalized to the values of the control Renilla luciferase. The fold change was calculated by proportioning the values of the empty vector.

3.7.9 siRNA Knockdown and qPCR

siRNA for the NH.3067 was purchased from IDT by custom design. NMuMG cells were treated with TGFβ-3 for 72 hours, then transfected with 50 nM siRNA with Lipofectamine RNAi Max (Invitrogen) and collated 72 hours later. RNA isolation was made with Nucleospin RNA II kit for controls and transfected cells. Using NanoDrop 2000, the concentration of these RNAs was measured, and 1 µg RNA was converted to cDNA using the Maxima First Strand cDNA Synthesis Kit. cDNAs were diluted 1:10 and used in qPCR with AMPIGENE qPCR probe mix (Enzo) and Universal Probe Library probes by user manual. Analysis and quantification of qPCR were performed using the ΔΔCt method.

3.7.10 Immunofluorescence Labeling and Confocal Microscopy

Cells were washed with PBS before being fixed for 10 minutes at room temperature with 4% formaldehyde. After that, it was permeabilized with 0.25% Triton X-100 for 10 minutes and washed three times with PBS. After, cells were incubated with E-cadherin antibody with 5% BSA (1:200 diluted) for 1 hour at room temperature in the dark. Then, samples were washed three times with PBS and incubated with Alexa Fluor 488 Phalloidin for one hour at room temperature in the dark. Nuclei were visualized with DAPI. Prolong Diamond Antifade Mountant solution was used to mount slides. Visualizing was done with Confocal microscopy LSM 880 microscope and ZEN software.

3.8. Limitation of the Study

No limitations are to be reported.

3.9. Ethics Committee Approval

This study was approved by the IBG local ethics committee with ethics protocol number 2022-010.

4. RESULTS

4.1. Expression of lncRNAs

PCR of 22 selected lncRNAs was made with each time point. Gel electrophoresis was done with 1.5% agarose gel, and results were obtained in Figures 5B, 6B, and 7B. These PCR results were compared with RNA-seq data in Figures 5A, 6A, and 7A. Expression change of each gene was consistent with RNA-seq data.

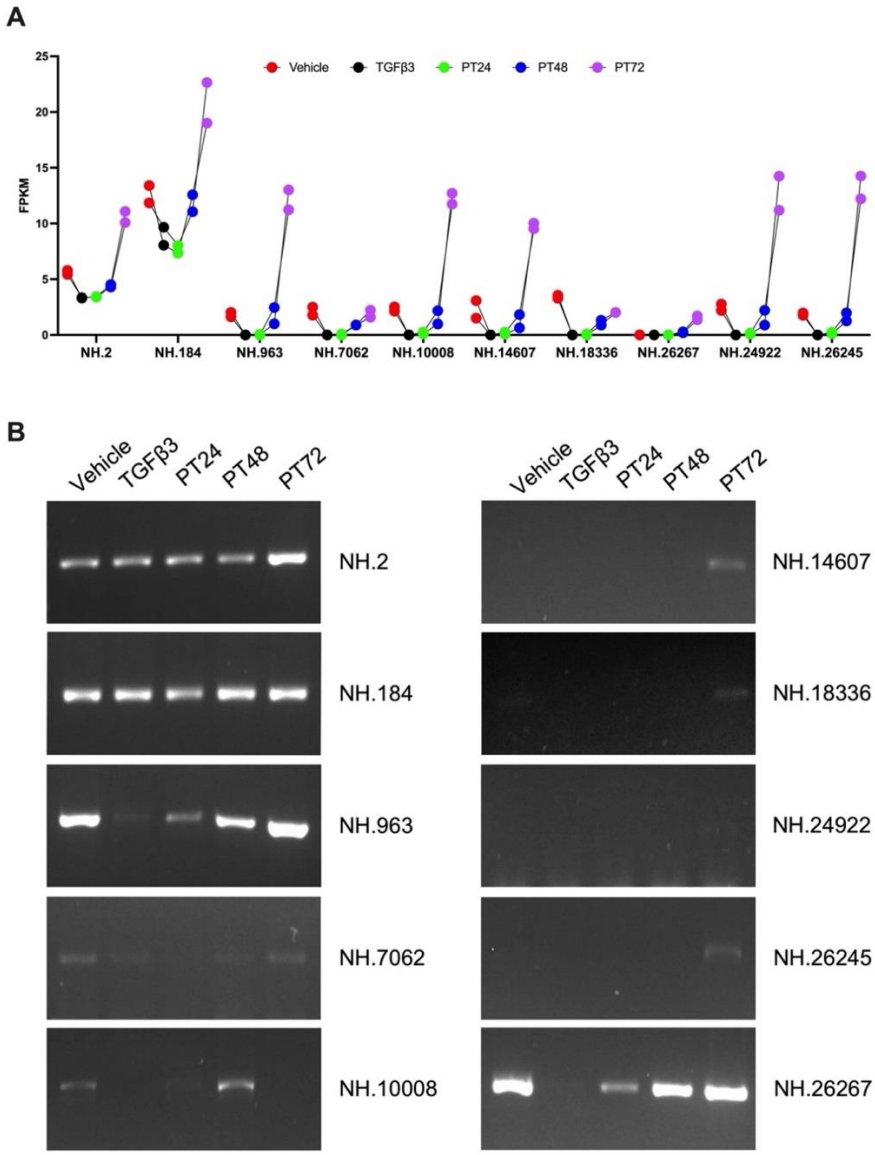

Figure 5. Expression of selected lncRNAs in the EMT/MET process. A. Expression analysis in RNA-Seq data. B. Gene expression results obtained by semi-quantitative RT-PCR method during the EMT/MET process.

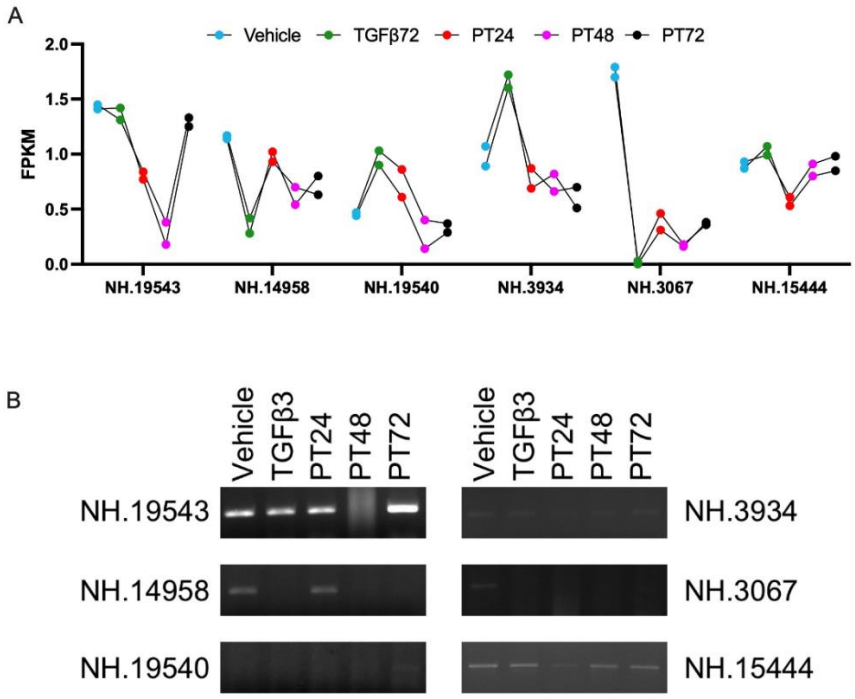

Figure 6. Expression of selected lncRNAs in the EMT/MET process. A. Expression analysis in RNA-Seq data. B. Gene expression results obtained by semi-quantitative RT-PCR method during the EMT/MET process.

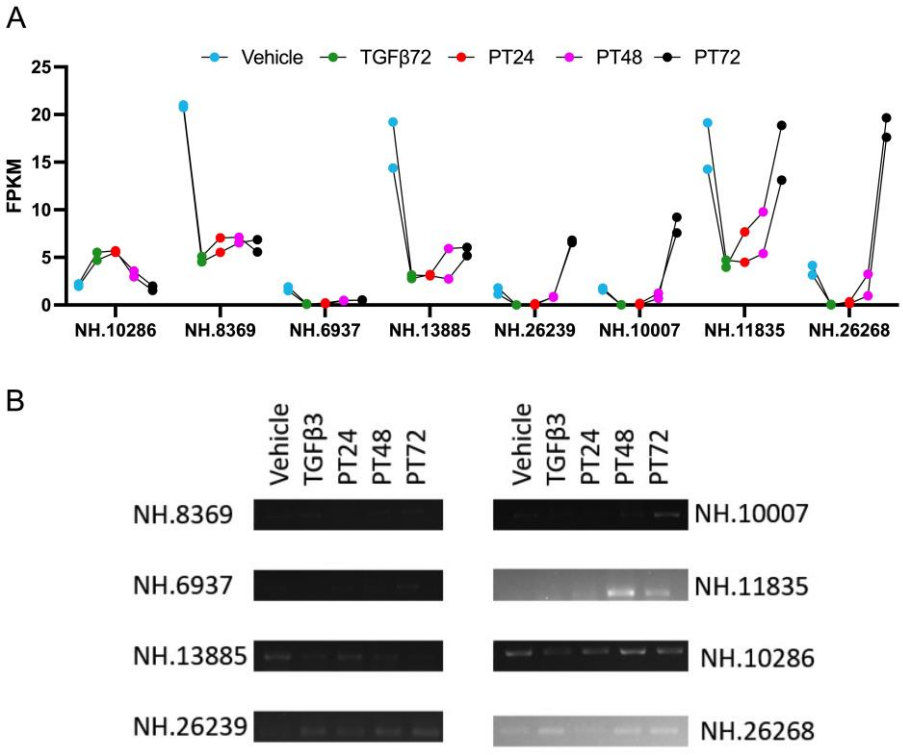

Figure 7. Expression of unknown lncRNAs in the EMT/MET process. A. Expression analysis in RNA-Seq data**. B.** Gene expression results obtained by semi-quantitative RT-PCR method during the EMT/MET process.

4.2. Genomic DNA Contamination Control

Because most genes were mono exonic, genomic DNA contamination was a concern. For that, a cDNA reaction was performed with and without the reverse transcriptase enzyme. There are no bands in PCR done with cDNA without the enzyme, and there are bands in PCR done with cDNA with the enzyme. NH.19540, NH.3067, NH.3934, NH.13885, NH.19543, NH.14958, NH.15444 genes were selected for genomic DNA contamination. Results can be seen as regular PCR and control PCR (Figure 8).

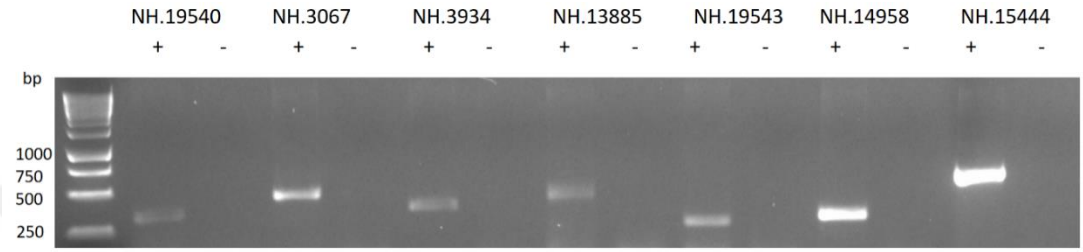

Figure 8. Control of genomic DNA contamination by RT-PCR

4.3. Restriction Digestion of PCR Products

PCR products were digested with related restriction enzymes. All of the genes were digested except NH.13885 (Figure 9). Some of the genes were not digested with restriction digestion because there was no proper restriction digestion enzyme available to digest them. NH.3067 was digested with SmaI, and bands were observed in 101 bp and 327 bp. NH.3934 was digested with NheI, and bands were observed in 146 bp and 184 bp. NH.14958 was digested with BsaI, and bands were observed in 51 bp and 229 bp. NH.15444 was digested with NcoI, and bands were observed in 193 bp and 405 bp. NH.184 was digested with SacI, and bands were observed in 287 bp and 78 bp. NH.2 was digested with HindIII, and bands were observed at 248 bp and 150 bp. NH.26267 was digested with NcoI, and bands were observed at 366 bp and 117 bp. NH.963 was digested with NcoI, and bands were observed in 367 bp and 133 bp. NH.10007 was digested with NcoI, and bands were observed in 848 bp and 272 bp.

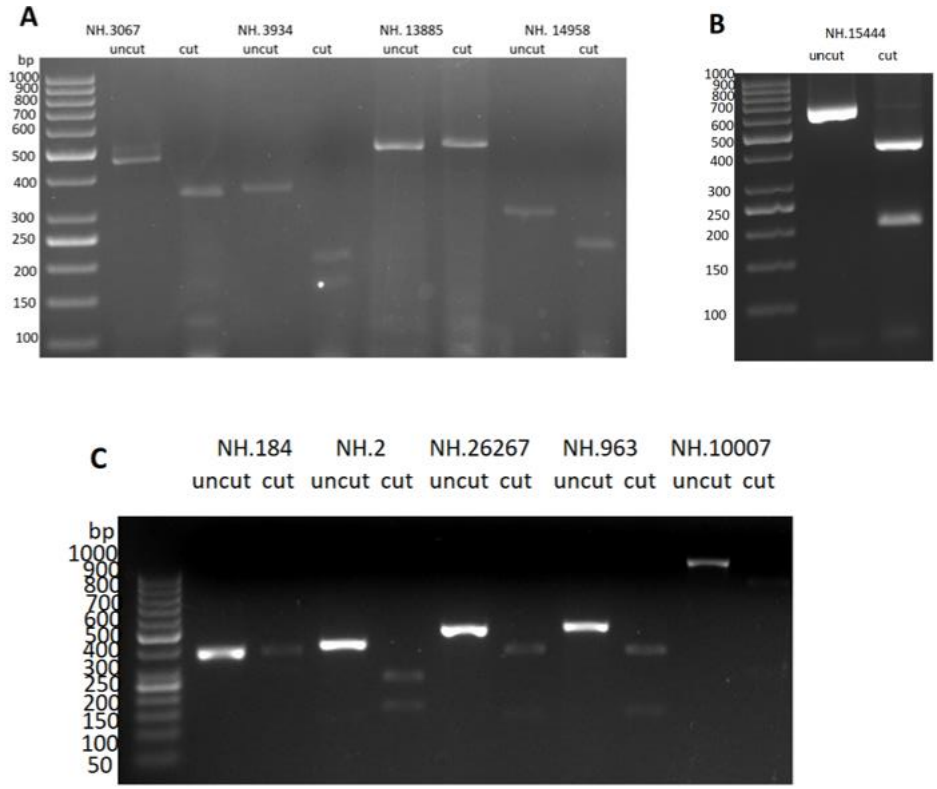

Figure 9. Restriction digestion results of PCR products A. Results of NH.3067, NH.3934, NH.13885, NH.14958 **B.** Result of NH.15444 **C.** Results of NH.184, NH.2, NH.26267, NH.963 and NH.10007

4.4. Protein Coding Capacity Confirmation

PCR was done on 4 selected genes (NH3067, NH.3934, NH.10286 and NH.13885). Successful amplifications were observed in each gene (Figure 10). IVTT assay was done with these PCR products. The PCR clean-up procedure was done on NH.10286 (850 bp), NH.3067 (512 bp), and NH.13885 (519 bp). Gel extraction and then PCR clean-up procedure were done to NH.3934 (414 bp) to remove the band at 1000 bp. Protein production was achieved as expected in the positive control sample (DHFR), and the bands in the negative control (-) sample belong to proteins involved in the transcription and translation stages in the reaction. The same bands were observed in negative control and genes that were selected. The extra band was only observed in positive control (Figure 11). Protein was only produced in DHFR positive control.

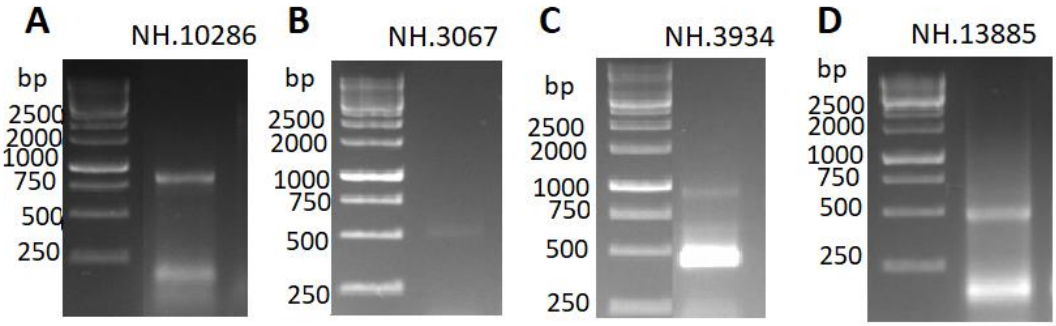

Figure 10. PCR results for IVTT A. NH.10286 PCR product **B.** NH.3067 PCR product **C.** NH.3934 PCR product **D.** NH.13885 PCR product

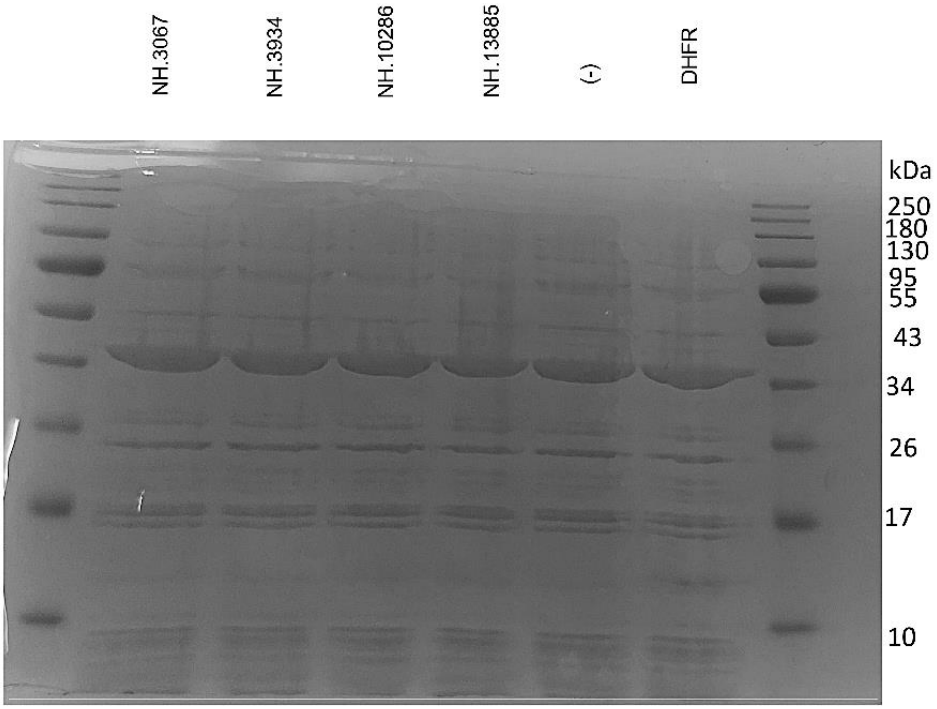

Figure 11. IVTT result on SDS-Page gel

4.5 Enhancer lncRNAs

In addition to lncRNAs associated with hub transcription factors, the study was deepened by investigating the enhancer potential of selected lncRNAs. To identify enhancer lncRNAs, an integrative approach was used where a comparative analysis of H3K27Ac ChIP-Seq data with RNA-Seq data from NMuMG cells during MET was performed. ATAC-Seq data with the conservation track from the Genome Browser and K27Ac and RNA-Seq data were used to increase the confidence of the data. In total, 3366 potential enhancer regions were identified, of which those overlapping lncRNA transcripts were selected. By examining the expression profiles of these

candidate enhancer lncRNAs during the EMT and MET processes, priority was given to those with a certain level of expression (FPKM > 1) and those adjacent to genes that define epithelial or mesenchymal character, considering their location in the genome. Overlapping results from these three data were selected as enhancer lncRNAs (Figure 12 A, Figure 13 A, Figure 14 A, Figure 15 A, Figure 16 A, and Figure 17 A).

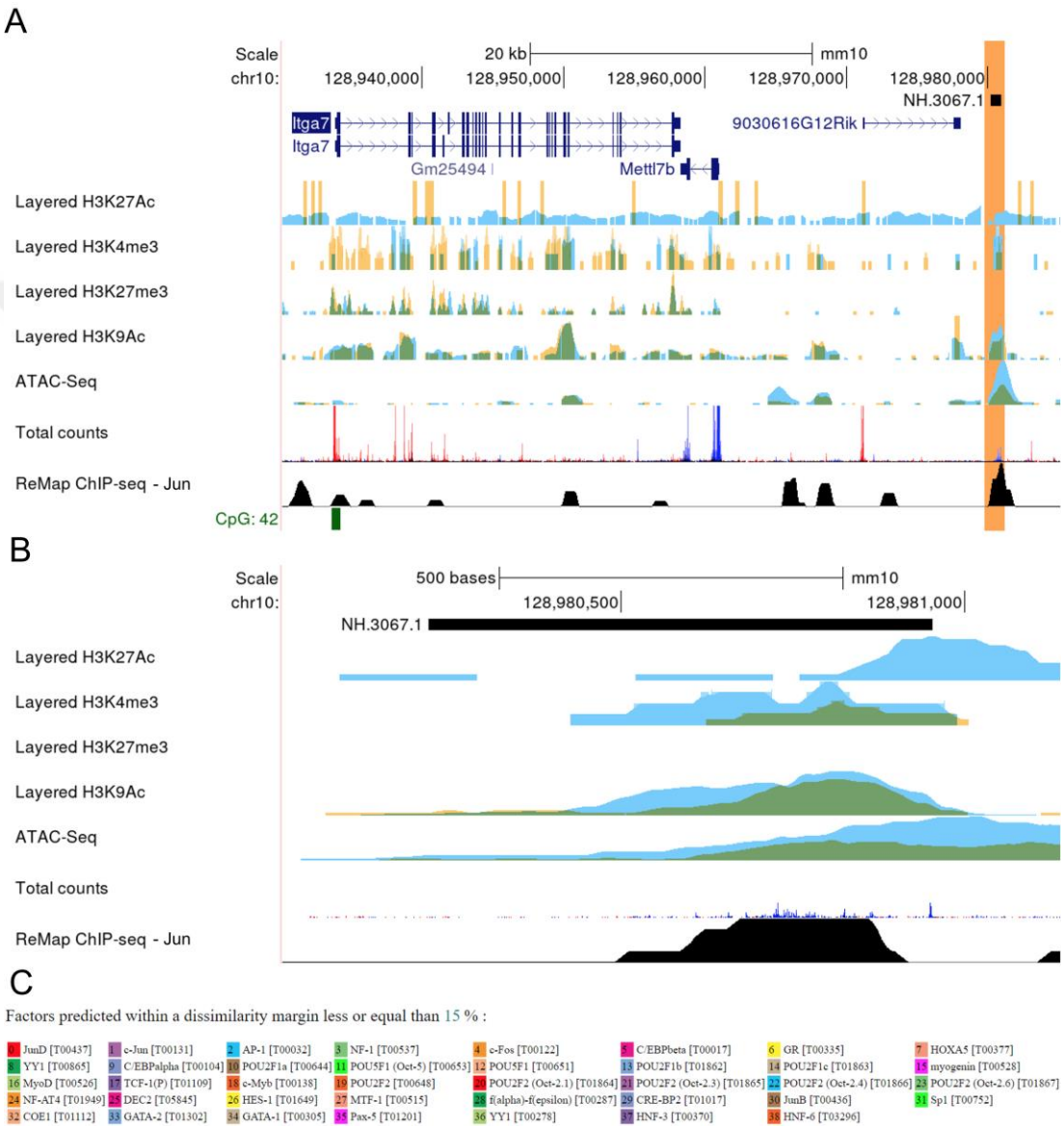

Figure 12. Integrative approach which combinates H3K27Ac ChIP-seq data with RNA-seq data for elncRNA NH.3067 A. elncRNA NH.3067 might have a relation with Itga7, which regulates cell behavior by cell-cell adhesion or cell-extracellular matrix adhesion B. Enhancer lncRNA NH.3067 compatibility with transcription start point FANTOM-5. C. Transcription factor binding site prediction of NH.3067

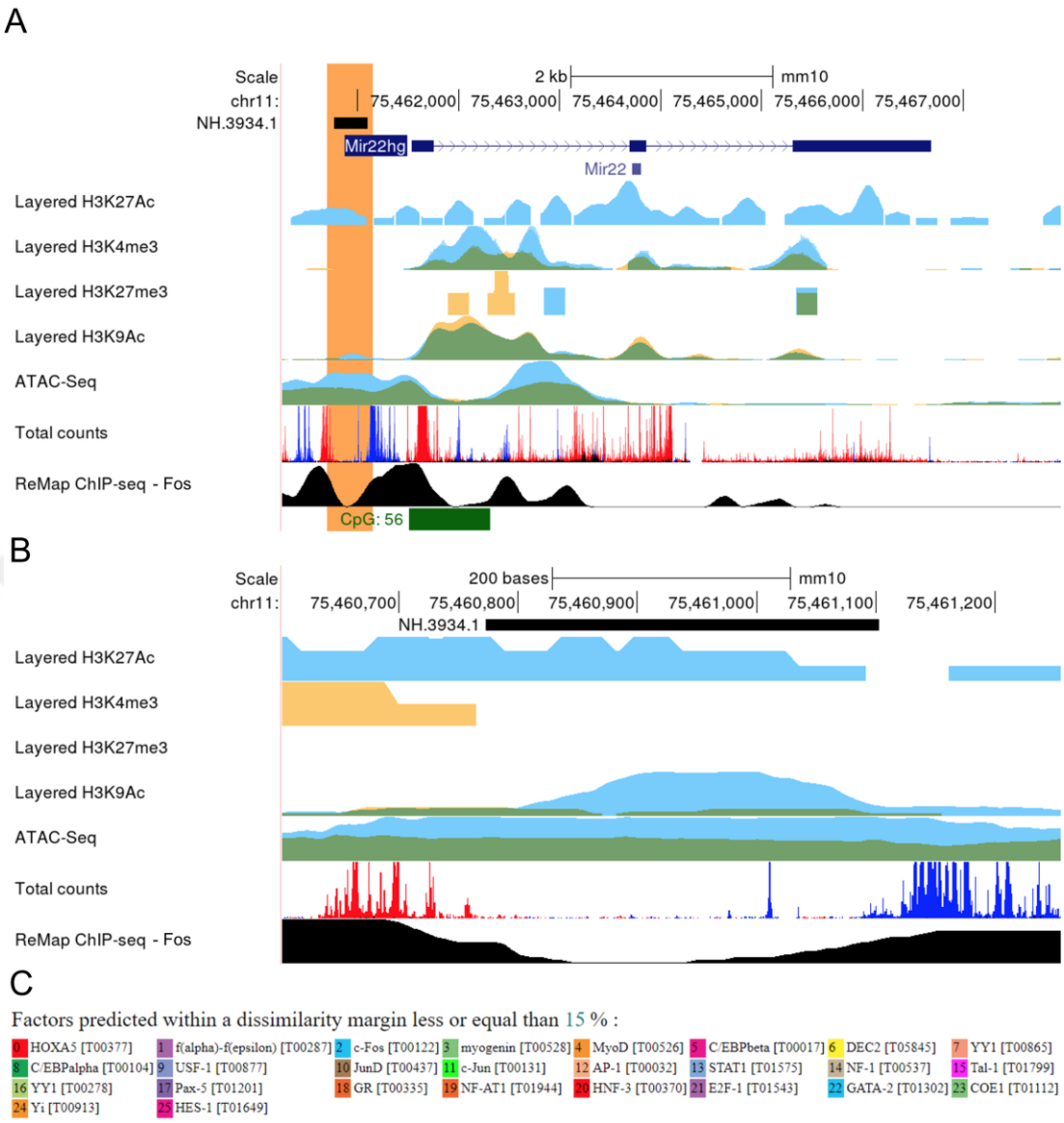

Figure 13. Integrative approach which combinates H3K27Ac ChIP-seq data with RNA-seq data for elncRNA NH.3934 A. elncRNA NH.3934 might have a relation with Mir22hg, which is an important miRNA for EMT/MET process. B Enhancer lncRNA NH.3934 compatibility with transcription start point FANTOM-5. C. Transcription factor binding site prediction of NH.3934

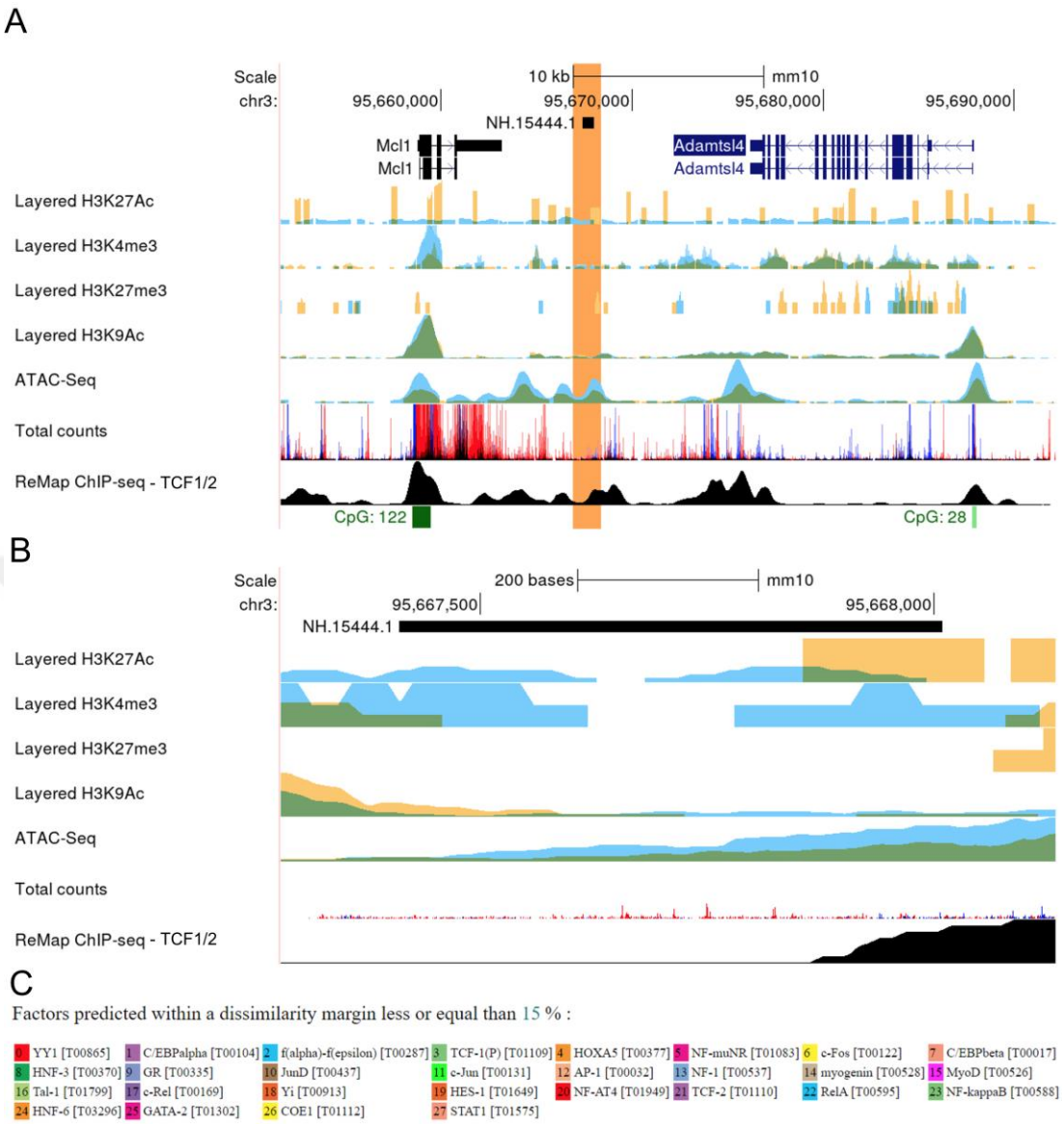

Figure 14. Integrative approach, which combinates H3K27Ac ChIP-seq data with RNA-seq data for elncRNA NH.15444 A. elncRNA NH.15444 might have a relationship with apoptosis. B. Enhancer lncRNA NH.15444 compatibility with transcription start point FANTOM-5. C. Transcription factor binding site prediction of NH.15444

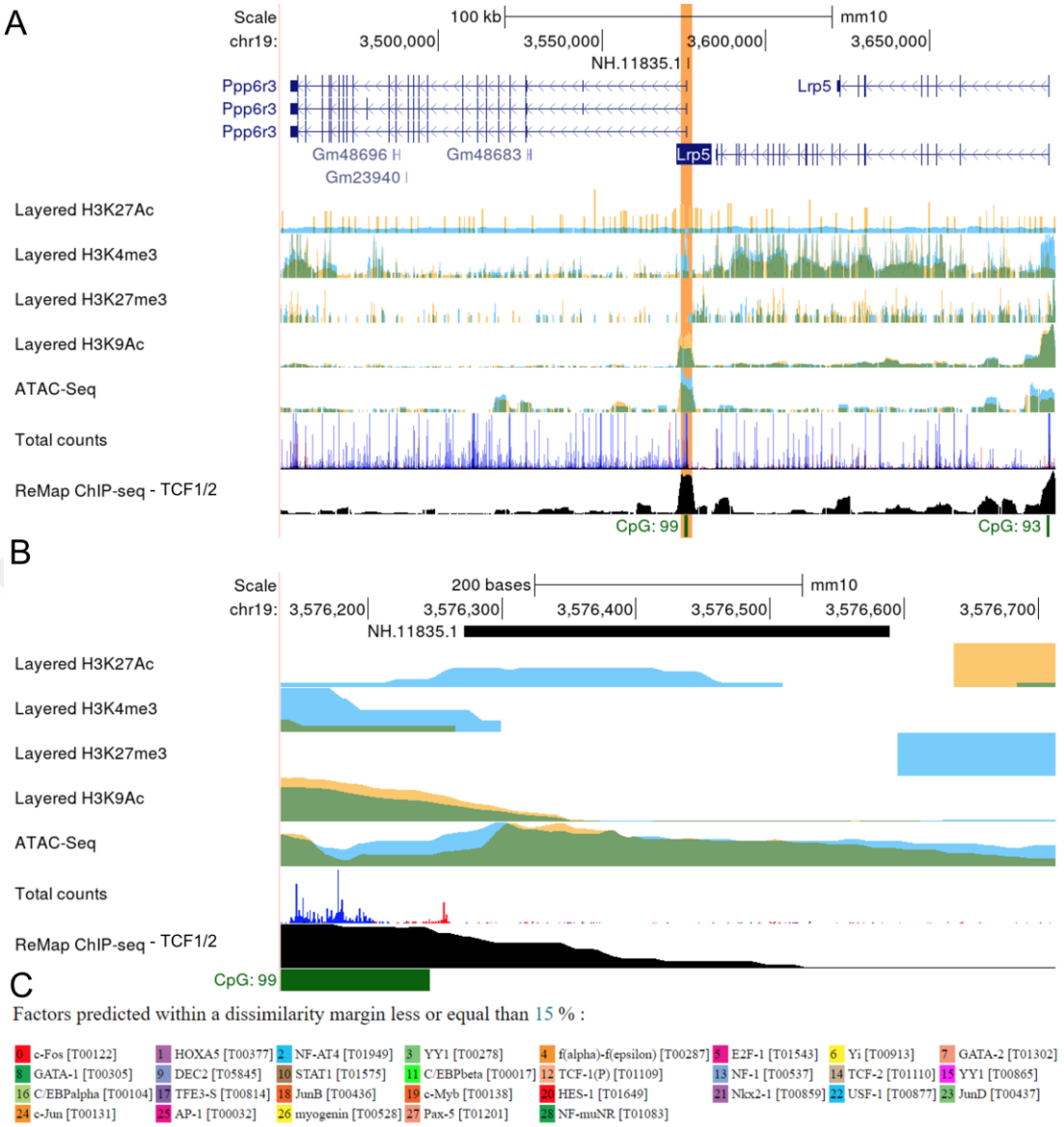

Figure 15. An Integrative approach that combinates H3K27Ac ChIP-seq data with RNA-seq data for elncRNA NH.11835 A. elncRNA NH.11835 might relate to Lrp5, which is essential for Wnt signaling. B Enhancer lncRNA NH.11835 compatibility with transcription start point FANTOM-5. C. Transcription factor binding site prediction of NH.11835

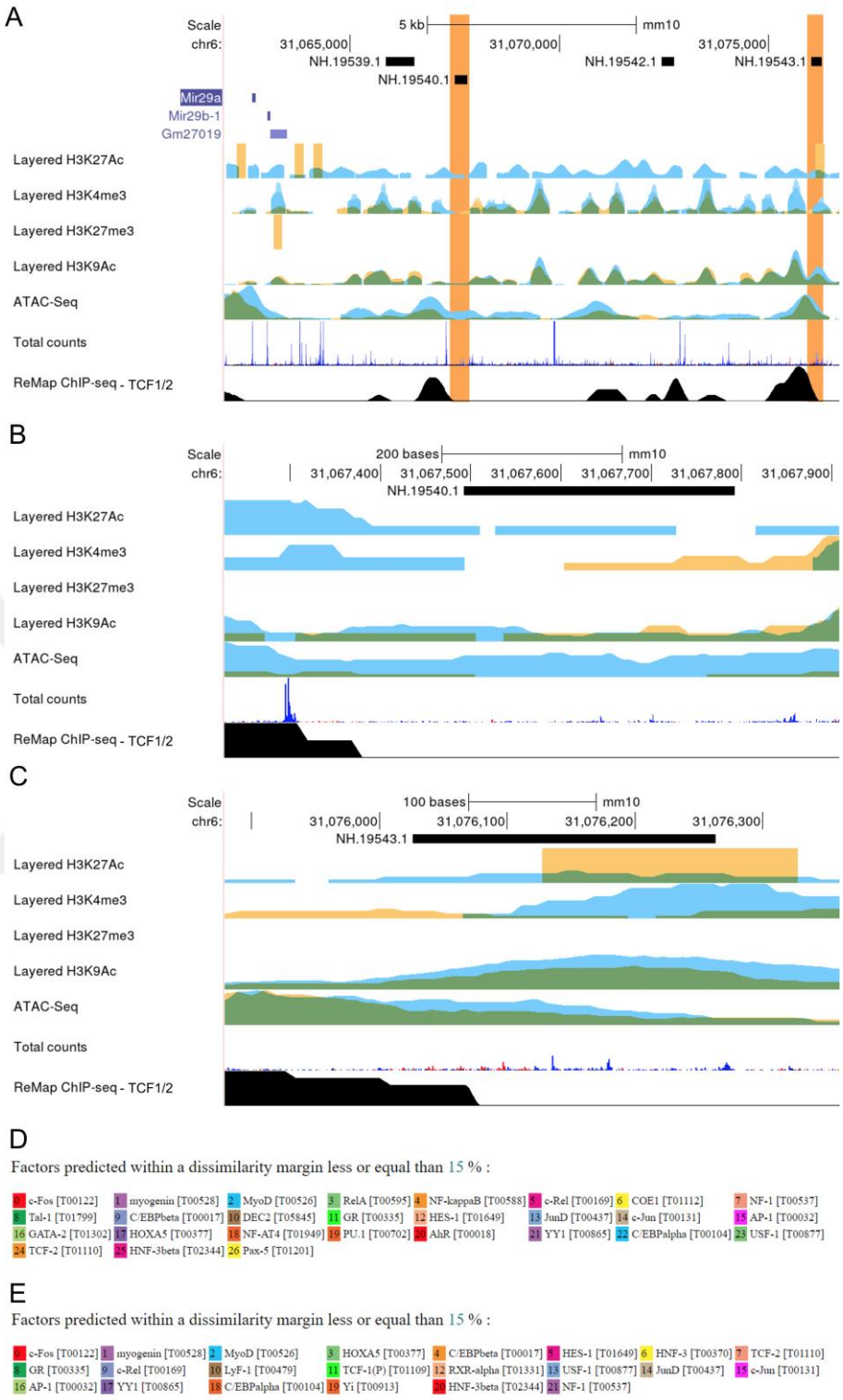

Figure 16. Integrative approach which combinates H3K27Ac ChIP-seq data with RNA-seq data for elncRNA NH.19540 and NH.19543 A. elncRNA NH.19543 might have a relation with Mir29, which is an important miRNA for apoptosis. B. Enhancer lncRNA NH.19540 compatibility with transcription start point FANTOM-5. C Enhancer lncRNA NH.19543 compatibility with transcription start point FANTOM-5. D. Transcription factor binding site prediction of NH.19540 E. Transcription factor binding site prediction of NH.19543

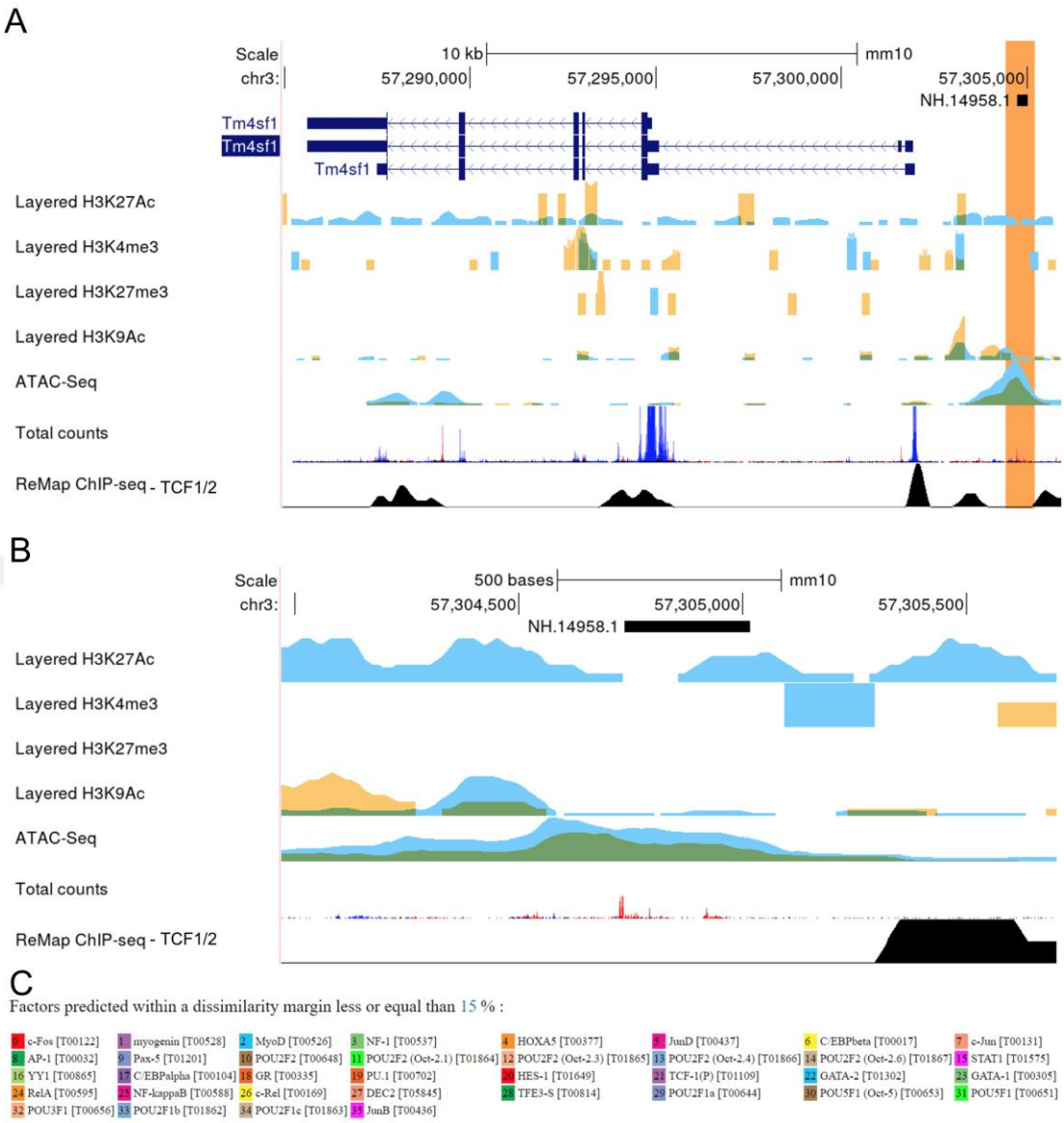

Figure 17. Integrative approach which combinates H3K27Ac ChIP-seq data with RNA-seq data A. elncRNA NH.14958 might have a relation with Tm4sf1, which contributes to the control of cellular development, activation, growth, and motility. B. Enhancer lncRNA NH.14958 compatibility with transcription start point FANTOM-5. C. transcription factor binding site prediction of NH.14958

To confirm the de novo transcriptome assembly performed for the purpose of identifying unknown lncRNA using all transcriptome RNA-Seq data obtained from samples collected at different time intervals of EMT and MET, FANTOM-5 was used. LncRNA regions were compared with FANTOM-5 transcription start data. From 1185 unknown lncRNA, only 320 of them were correlated with FANTOM-5 data. Comparison of the region identified with mm97147.1 with the Cage-Seq "total counts" data of FANTOM-5 and our H3K4me3 data from the NMuMG MET process indicates the promoter/TSS region of NH.3067 in Figure 12 B. For NH.3934, promoter/TSS can be seen in

Figure 13 B with the help of Cage-Seq "total counts" data of FANTOM-5 and our H3K4me3 data from the NMuMG MET process. The promoter/TSS of NH.15444 can be seen in Figure 14 B with the help of Cage-Seq "total counts" data of FANTOM-5 and our H3K4me3 data from the NMuMG MET process. For NH.11835, promoter/TSS can be seen in Figure 15 B with the help of Cage-Seq "total counts" data of FANTOM-5 and our H3K4me3 data from the NMuMG MET process. The promoter/TSS of NH.19540 can be seen in Figure 16 B with the help of Cage-Seq "total counts" data of FANTOM-5 and our H3K4me3 data from the NMuMG MET process. For NH.19543, promoter/TSS can be seen in Figure 16 C with the help of Cage-Seq "total counts" data of FANTOM-5 and our H3K4me3 data from the NMuMG MET process. For NH.14958, promoter/TSS can be seen in Figure 17 B with the help of Cage-Seq "total counts" data of FANTOM-5 and our H3K4me3 data from the NMuMG MET process.

Transcription factor binding site prediction was made in https://alggen.lsi.upc.es/recerca/frame-recerca.html website. JUN transcription factor has a potential binding site in NH.3067 gene in Figure 12 C. Fos transcription factor has a potential binding site in NH.3934 gene in Figure 13 C. TCF transcription factor has a potential binding site in NH.15444 gene in Figure 14 C, NH.11835 gene in Figure 15 C, NH.14958 gene in Figure 17 C, NH.19540 gene in Figure 16 D, and NH.19543 gene in Figure 17 E.

4.6. Cloning of Selected Enhancer lncRNAs in pGL4.23

NH.3067, NH.3934, NH.19543, NH.11835, NH.19540, NH.15444, NH.14958 genes cloned into pGL4.23 vector with minimal promoter and luciferase reporter gene. All genes were amplified with PCR. Then NH.3067, NH.14958, NH.19543, NH.11835, NH.19540, and NH.15444 were digested with XhoI/HindIII, NH.3934 were digested with XhoI/BglII.

After ligation and transformation, colony PCR was done to check the positive colonies. For NH.15444, 778 bp bands were observed in colonies 1 and 2, which were positive. 212 bp bands were observed in colonies 3 and 4, which were negative. For NH.11835, 497 bp bands were observed in colonies 2 and 4, which were positive. 212 bp bands were observed in colonies 1 and 3, which were negative. For NH.19540, a 486 bp band was observed in colony 4, which was positive. 212 bp bands were observed in colonies 1, 2, and 3, which were negative. For NH.14958, 460 bp bands were observed in colonies 1, 2, and 3, which were positive. 212 bp band was observed in colony 4, which was negative. (Figure 18). Other genes were checked in the same way.

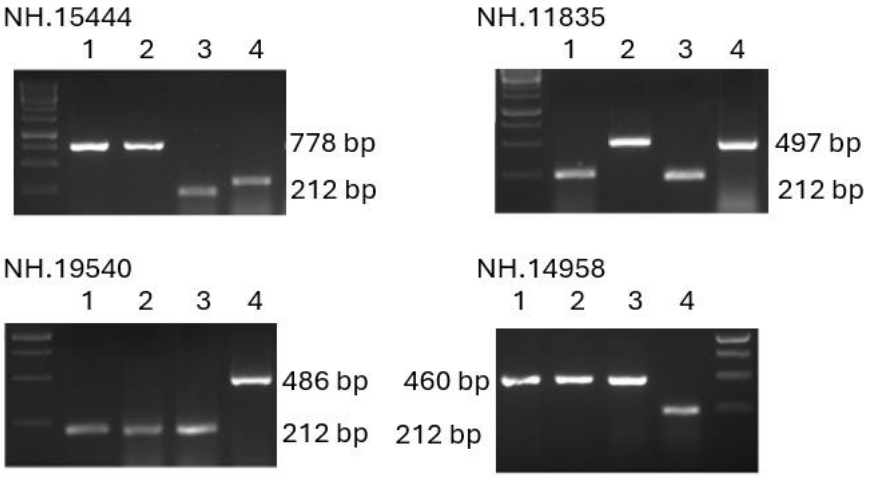

Figure 18. Colony PCR of NH.15444, NH11835, NH.19540 and NH.14958

4.7. Luciferase Reporter Assay

To check whether NH.3067, NH.3934, NH.19543, NH.11835, NH.19540, NH.15444, and NH.14958 genes have enhancer capacity or not, luciferase reporter assay was done in NMuMG cells. NH.3067 and NH.19543 genes increased the activity of the minimal promoter and, therefore, the luciferase gene in pGL4.23 by approximately 11 and 6 times, respectively. According to the results of other transcripts, instead of increasing reporter activity, they caused a very strong decrease (Figure 19).

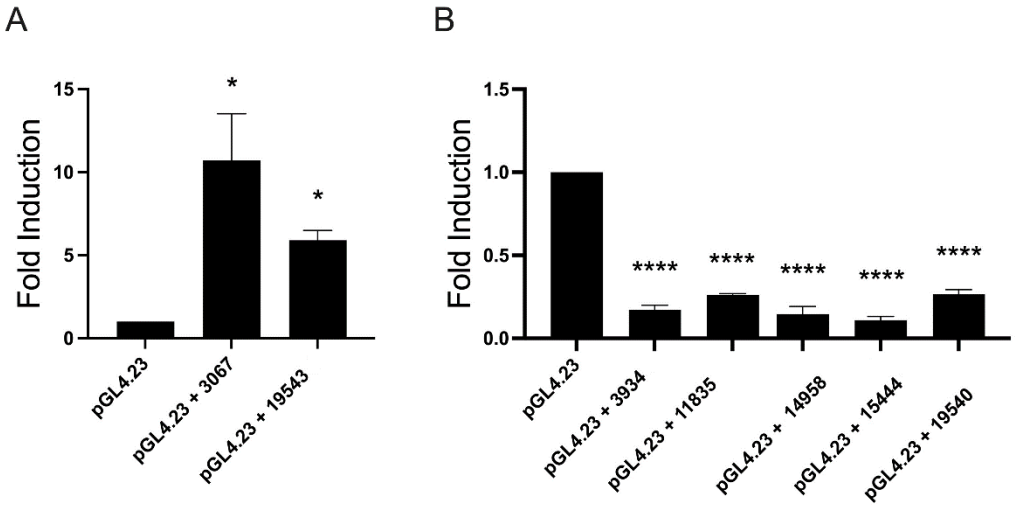

Figure 19. Luciferase results of novel lncRNAs showing strong regulator activity. A. Genes which increased the activity of the minimal promoter. B. Genes that decreased the activity of the minimal promoter.

4.8 siRNA Silencing of NH.3067

Silencing experiments were carried out with the help of siRNA during the MET process of the NH.3067 gene. After NMuMG cells entering the MET process were transfected with siRNA designed for NH.3067, whether there were morphological changes was checked. Compared to control siRNAs, NH.3067 silencing led to cell phenotypic changes (Figure 20). These morphological changes can also be seen in confocal images (Figure 21). Therefore, after silencing, epithelial and mesenchymal markers were checked by the qPCR method. Effective silencing was achieved with both siRNAs we used. Due to silencing, a severe decrease in the expression of Cdh1, Elf3, and Grhl3 was observed (Figure 22).

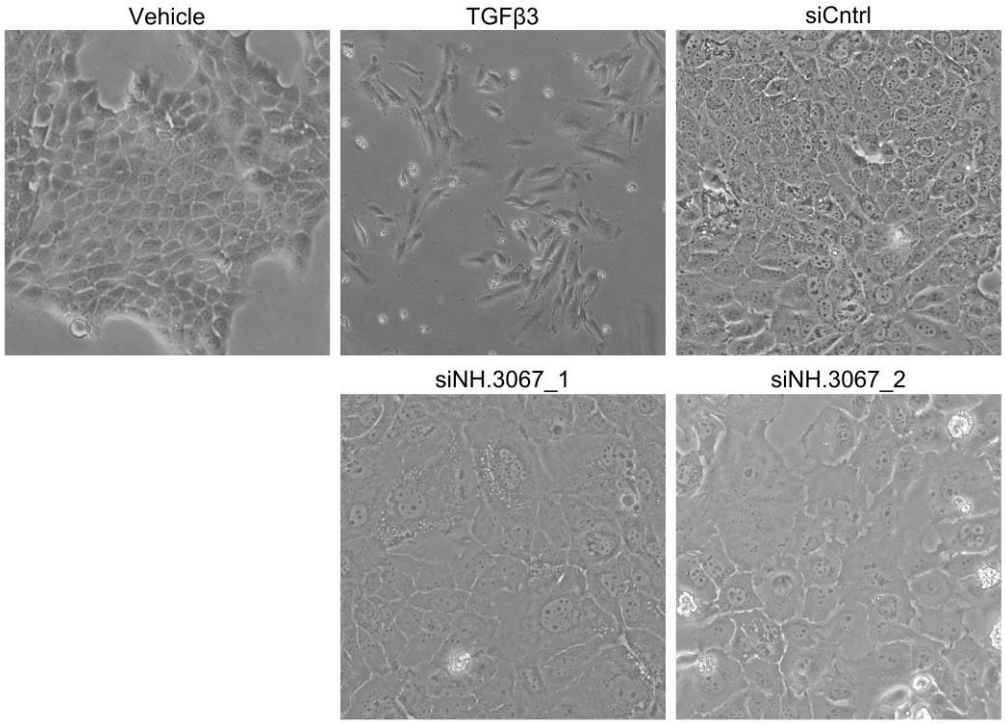

Figure 20. Silencing of NH.3067 in MET stage. Phase contrast images show the morphological changes in siRNA-mediated NH.3067 silenced cells.

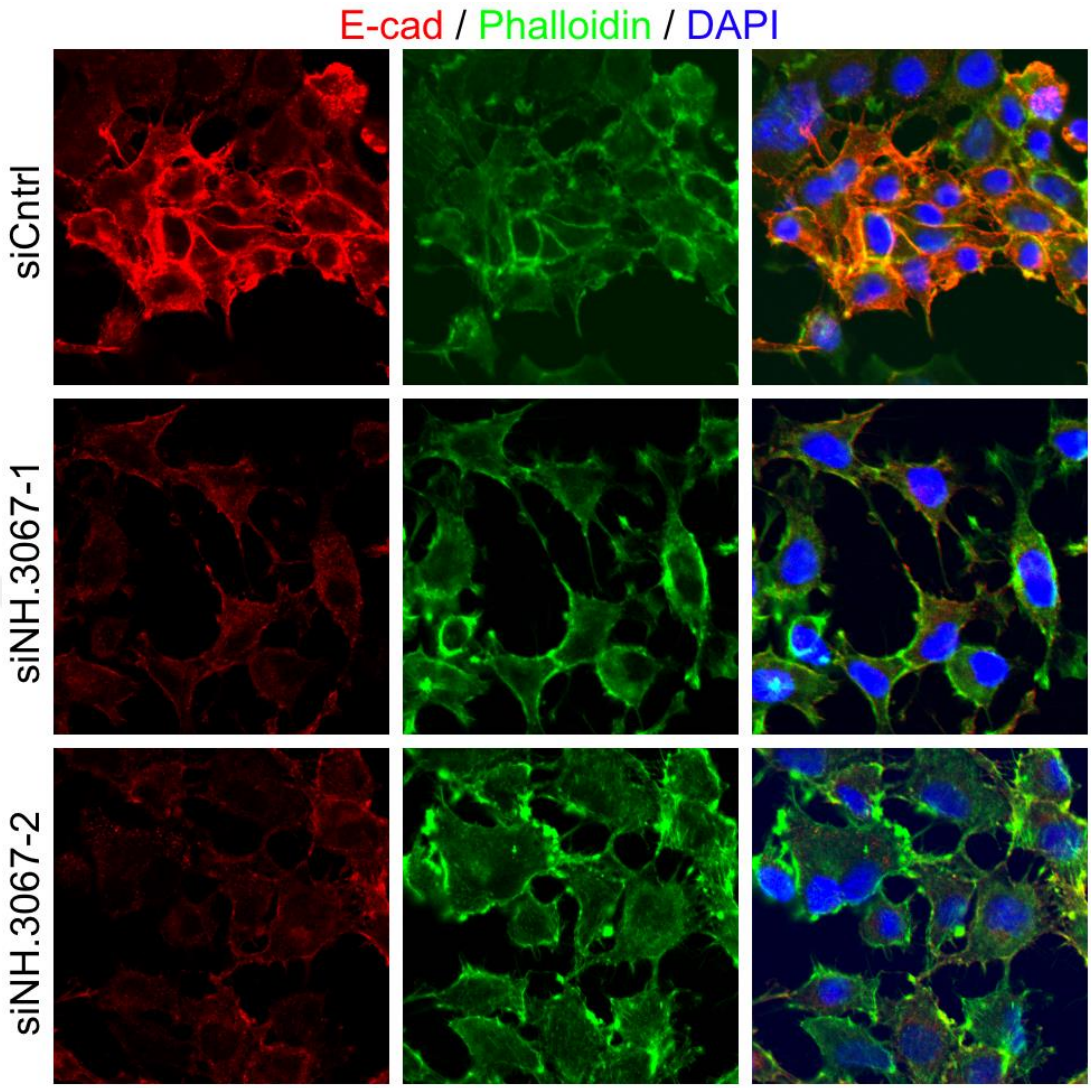

Figure 21. The siRNA-mediated silencing of NH.3067 the immunofluorescence staining of NMuMG cells in MET. Confocal images show the morphological changes in siRNA-mediated NH.3067 silenced cells.

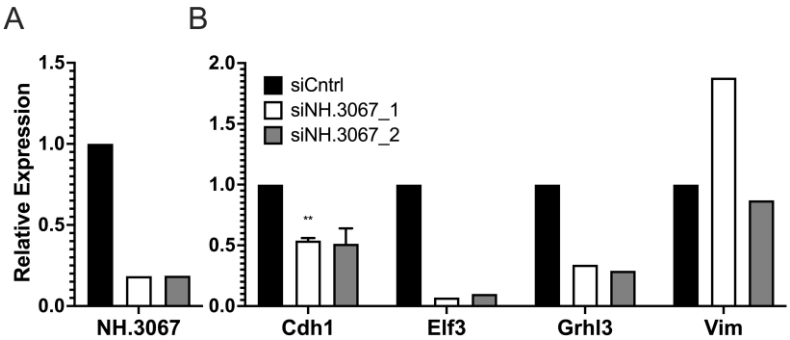

Figure 22. Effective silencing of NH.3067 resulted in altered expression of epithelial and mesenchymal markers.

5. DISCUSSION

In normal circumstances, the EMT and MET processes play a crucial role in physiological conditions, ensuring the smooth progression of embryonic development and wound healing (Kalluri and Weinberg, 2009). These processes are known to be fundamental for various biological functions, including the maintenance of cellular phenotype and the facilitation of plasticity (Lamouille et al., 2014). In pathological conditions like tumorigenesis and metastasis, tumor cells leverage these processes to acquire an aggressive phenotype (Heerboth et al., 2015). The tumor's ability to metastasize, either locally or distantly, stands as a primary factor contributing to cancer-related deaths (Spano et al., 2012). While the complete mechanism of the metastasis process remains incompletely understood, it is widely believed that EMT/MET processes play a central role in this phenomenon (Seyfried and Huysentruyt, 2013).

The regulation of EMT involves factors such as Snail/Slug and Zeb1/2, which suppress the expression of E-cadherin. These factors collaboratively contribute to a significant reduction in EMT followed by a rapid re-increase in E-cadherin gene expression during MET (Voutsadakis, 2015). While MET is traditionally considered the reversible counterpart of EMT, its regulation is governed by a more dynamic and intricate mechanism. EMT-inducing factors may account for only a portion of the activation of E-cadherin gene expression, a central aspect of the MET process (Kim et al., 2018).

EMT and MET are intricate cellular reprogramming processes regulated at various levels. Despite progress in understanding EMT, there is still a gap in comprehending the molecular mechanisms governing MET. Our investigation utilized NMuMG cells, known for well-characterized EMT and MET responses, primarily focusing on MET progression. Sample collection occurred at different intervals spanning three days of MET, accompanied by confirmatory experiments to validate the seamless transition between epithelial and mesenchymal states. The three-day MET timeline was categorized into distinct phases, each corresponding to a day of MET, based on unique morphological features. Subsequently, we conducted RNA sequencing and meticulously analyzed the expression profiles of NMuMG cells throughout MET. The overall shifts in gene expression aligned with the MET stages designated according to morphological changes. Furthermore, the expression patterns of various epithelial and mesenchymal markers affirmed the successful transitions between EMT and MET. Turning our attention to non-coding genes, we identified numerous lncRNAs with limited expression during any process stage.

The functional annotation of lncRNAs necessitates combining various in silico and experimental methods. A frequently employed strategy involves association, where the function of a lncRNA is deduced from its expression

correlation with protein-coding genes, coupled with gene set enrichment analysis (Lefever et al., 2017). Decode (Lefever et al., 2017) and WGCNA (Langfelder and Horvath, 2008) are commonly used for these analyses.

Interpreting the roles of the newly identified lncRNA candidates presents a challenge due to the limited conservation of lncRNAs across species. The sequence conservation of lncRNAs is notably low, but their functional conservation remains significant. In our study, the MET model employed was based on a mouse cell line, while many cancer datasets are sourced from human samples or xenografts, where human tumors are transplanted into mouse models. Determining human orthologs of specific lncRNAs found in other species often requires approaches beyond sequence comparison, such as syntenic comparisons, which consider shared genomic contexts between lncRNA species and the conservation of adjacent genes (Amaral et al., 2018) or secondary structure domains (Bussotti et al., 2013). This method relies on partial primary sequence conservation in the genome. An alternative approach for identifying lncRNA relationships involves recognizing shared interactions through network analysis. However, such network analysis demands a comprehensive transcriptomic examination of human cancer samples, encompassing the non-coding transcriptome.

In this study, we initially evaluated the expression of lncRNAs in vehicle, EMT, and various stages of MET. Alterations in gene expression were observed, and this pattern was corroborated by RNA-seq data, indicating compatible changes in each gene's expression. To determine whether these lncRNAs originated from genomic DNA or isolated RNA, we conducted a confirmation by excluding the reverse transcriptase enzyme from the cDNA reaction. The results confirmed the absence of genomic DNA contamination, validating the accuracy of our RT-PCR results.

For further validation, we performed restriction digestion on each amplification, with expected results for all lncRNAs except NH.13885, where the band was not visible, possibly indicating an incomplete reaction. Non-coding capacity was confirmed through an IVTT assay, demonstrating that only the positive control DHFR produced protein, substantiating the non-coding nature of these genes.

Employing the ROSE enhancer identification algorithm, we identified potential enhancers, subsequently confirmed by ATAC-seq and H3K27Ac ChIP-Seq data, indicating open chromatins (Grandi et al., 2022). We selected specific lncRNAs based on their proximity to enhancer targets. The NH.3067 gene potentially acts as an enhancer for ITGA7, with the JUN transcription factor, a known regulator of ITGA7, possibly exerting regulatory influence on NH.3067. This is indicated by a potential binding site identified through transcription factor binding site prediction. Moreover, Itga7, a neighboring gene that regulates cell behavior, could be influenced by NH.3067(Bai et al., 2019; Z. Chen et al., 2023). Furthermore, the JUN transcription factor is implicated in the repression of ITGA7 expression (Lerch et al., 2014). Additionally, there

is a potential interaction between the JUN transcription factor and β-catenin, as suggested by Bayes' theorem (Hwang et al., 2017). Considering the significant role of β-catenin in EMT through the β-catenin signaling pathway (Kato et al., 2007), this interaction could contribute to the observed regulatory dynamics. Cloning NH.3067 into pGL4.23 revealed enhancer potential, confirmed by an 11-fold increase in luciferase activity.

NH.3934, with Mir22 as a neighbor, displayed high luciferase inhibition. It is likely linked to Mir22/Mir22hg adjacent to its sequence because Mir22 inhibits EMT (Xia et al., 2022). NH.3934 may play a role in regulating Mir22, as indicated by the presence of a binding site for the Fos transcription factor in the NH.3934 gene, a known regulator of Mir22 expression. Since Mir22 inhibits EMT (Xia et al., 2022), the regulatory dynamics of NH.3934 and Fos may contribute to EMT regulation. Fos transcription factor has been shown to be involved in the regulation of Mir22, which is associated with EMT progression. Moreover, the overexpression of FOSL1, a member of the Fos family, induces EMT (Ahmad et al., 2017).

Similarly, NH.15444, NH.11835, NH.14958, NH.19540, and NH.19543 genes might regulate Mcl1, Lrp5, Tm4sf1, and Mir29b respectively, as five genes contain potential binding sites for the TCF transcription factor, known to regulate the expressions of Mcl1, Lrp5, Tm4sf1, and Mir29b. TCF plays a role in Wnt signaling (Kolligs et al., 2002), and Lrp5, being a coreceptor of Wnt ligands, is a major component of the Wnt pathway (Ren et al., 2021). Mcl1 is regulated by TCF and β-catenin signaling pathway (Ruiz-Vela et al., 2008). Tm4sf1 promotes EMT by Wnt/β-catenin signaling pathway by TCF transcription factor (Tang et al., 2020). Mir29b holds a crucial position in the regulation of Wnt/β-catenin signaling (Shin et al., 2014). Only NH.19543 revealed enhancer potential, confirmed by a 5-fold increase in luciferase activity. Four of them displayed high luciferase inhibition.

Upon NH.3067 knockdown, noticeable changes in cell morphology suggested disrupted MET progression. Subsequent examination revealed a significant decrease in Cdh1, Elf3, and Grhl3 expression and an increase in Vimentin expression, underscoring the crucial role of NH.3067 in MET.

6. CONCLUSION AND SUGGESTIONS

The experimental validation of lncRNAs capable of regulating the MET process has been successfully illustrated as part of the thesis project. Conducting thorough and advanced experiments to explore the function of NH.3067 is crucial as it plays a vital role in regulating MET. The lncRNAs identified as significant in the MET process exhibit associations with the cell cycle, chromosome segregation, and DNA damage repair. MET transcriptional factors and lncRNAs are integral components of the MET regulatory network.

In the context of the thesis, numerous MET-specific lncRNAs were identified and subjected to detailed analysis. These investigations are pivotal for the expansion of the MET transcriptional network. Furthermore, our data is believed to bridge the publication gap in the existing literature regarding the transcriptional and epigenetic regulation of MET. The identified lncRNAs have potential applications as cancer biomarkers.

The correlation of these identified lncRNAs and transcription factors can be analyzed in the future. For example, Human orthologs can be identified. Cut&Run assay can be done for further investigations in vehicle state, EMT state, MET state, and silenced NH.3067 state (in MET). Further experiments can be done to understand the functions of identified lncRNAs.

7. REFERENCES

Ahmad, H. M., Muiwo, P., Muthuswami, R., Bhattacharya, A. 2017. "FosB regulates expression of miR-22 during PMA induced differentiation of K562 cells to megakaryocytes". Biochimie.

Alotaibi, H., Basilicata, M. F., Shehwana, H., Kosowan, T., Schreck, I., Braeutigam, C., … Stemmler, M. P. 2015. "Enhancer cooperativity as a novel mechanism underlying the transcriptional regulation of E-cadherin during mesenchymal to epithelial transition". Biochimica et Biophysica Acta - Gene Regulatory Mechanisms.

Amaral, P. P., Leonardi, T., Han, N., Viré, E., Gascoigne, D. K., Arias-Carrasco, R., … Kouzarides, T. 2018. "Genomic positional conservation identifies topological anchor point RNAs linked to developmental loci". Genome Biology.

Bai, X., Gao, C., Zhang, L., Yang, S. 2019. "Integrin α7 high expression correlates with deteriorative tumor features and worse overall survival, and its knockdown inhibits cell proliferation and invasion but increases apoptosis in breast cancer". Journal of Clinical Laboratory Analysis, *33*(8).

Böhmdorfer, G., Wierzbicki, A. T. 2015. "Control of Chromatin Structure by Long Noncoding RNA". Trends in Cell Biology, *25*(10), 623–632.

Brabletz, T. 2012. "To differentiate or not — routes towards metastasis". Nature Reviews Cancer, *12*(6), 425–436.

Bussotti, G., Notredame, C., Enright, A. J. 2013. "Detecting and comparing non-coding RNAs in the high-throughput era". International Journal of Molecular Sciences.

Cao, J. 2014. "The functional role of long non-coding RNAs and epigenetics". Biological Procedures Online.

Chen, M., Chen, J., Zhang, D. 2015. "Exploring the secrets of long noncoding RNAs". International Journal of Molecular Sciences.

Chen, Z., Kelly, J. R., Morales, J. E., Sun, R. C., De, A., Burkin, D. J., McCarty, J. H. 2023. "The alpha7 integrin subunit in astrocytes promotes endothelial blood–brain barrier integrity". Development (Cambridge).

Craene, B. De, Berx, G. 2013. "Regulatory networks defining EMT during cancer initiation and progression". Nature Reviews Cancer, *13*(2), 97–110.

Diamantopoulos, M. A., Tsiakanikas, P., Scorilas, A. 2018. "Non-coding RNAs:

the riddle of the transcriptome and their perspectives in cancer". Annals of Translational Medicine, *6*(12), 241–241.

Diepenbruck, M., Christofori, G. 2016. "Epithelial–mesenchymal transition (EMT) and metastasis: yes, no, maybe?". Current Opinion in Cell Biology, *43*, 7–13.

Dykes, I. M., Emanueli, C. 2017. "Transcriptional and Post-transcriptional Gene Regulation by Long Non-coding RNA". Genomics, Proteomics and Bioinformatics.

Font-Cunill, B., Arnes, L., Ferrer, J., Sussel, L., Beucher, A. 2018. "Long Non-coding RNAs as Local Regulators of Pancreatic Islet Transcription Factor Genes". Frontiers in Genetics.

Grandi, F. C., Modi, H., Kampman, L., Corces, M. R. 2022. "Chromatin accessibility profiling by ATAC-seq". Nature Protocols.

Gugnoni, M., Ciarrocchi, A. 2019. "Long noncoding RNA and epithelial mesenchymal transition in cancer". International Journal of Molecular Sciences.

Gupta, P. B., Pastushenko, I., Skibinski, A., Blanpain, C., Kuperwasser, C. 2019. "Phenotypic Plasticity: Driver of Cancer Initiation, Progression, and Therapy Resistance.". Cell Stem Cell, *24*(1), 65–78.

He, R. Z., Luo, D. X., Mo, Y. Y. 2019. "Emerging roles of lncRNAs in the post-transcriptional regulation in cancer". Genes and Diseases.

Heerboth, S., Housman, G., Leary, M., Longacre, M., Byler, S., Lapinska, K., … Sarkar, S. 2015. "EMT and tumor metastasis". Clinical and Translational Medicine.

Hubé, F., Francastel, C. 2018. "Coding and non-coding RNAs, the frontier has never been so blurred". Frontiers in Genetics.

Hwang, J. R., Chou, C.-L., Medvar, B., Knepper, M. A., Jung, H. J. 2017. "Identification of β-catenin-interacting proteins in nuclear fractions of native rat collecting duct cells". American Journal of Physiology-Renal Physiology, *313*(1), F30–F46.

Jayachandran, A., Dhungel, B., Steel, J. C. 2016. "Epithelial-to-mesenchymal plasticity of cancer stem cells: Therapeutic targets in hepatocellular carcinoma". Journal of Hematology and Oncology.

Jolly, M. K., Boareto, M., Huang, B., Jia, D., Lu, M., Onuchic, J. N., … Ben-Jacob, E. 2015. "Implications of the hybrid epithelial/mesenchymal phenotype in metastasis". Frontiers in Oncology, *5*.

Kalluri, R., Weinberg, R. A. 2009. "The basics of epithelial-mesenchymal transition". Journal of Clinical Investigation, *119*(6), 1420–1428.

Karapetyan, A. R., Buiting, C., Kuiper, R. A., Coolen, M. W. 2013. "Regulatory roles for long ncRNA and mRNA". Cancers.

Kato, N., Shimmura, S., Kawakita, T., Miyashita, H., Ogawa, Y., Yoshida, S., … Tsubota, K. 2007. "β-catenin activation and epithelial-mesenchymal transition in the pathogenesis of pterygium". Investigative Ophthalmology and Visual Science.

Kim, D. H., Xing, T., Yang, Z., Dudek, R., Lu, Q., Chen, Y. H. 2018. "Epithelial mesenchymal transition in embryonic development, tissue repair and cancer: A comprehensive overview". Journal of Clinical Medicine.

Kolligs, F. T., Bommer, G., Göke, B. 2002. "Wnt/beta-catenin/Tcf signaling: A critical pathway in gastrointestinal tumorigenesis". Digestion.

Lamouille, S., Xu, J., Derynck, R. 2014. "Molecular mechanisms of epithelial–mesenchymal transition". Nature Reviews Molecular Cell Biology, *15*(3), 178–196.

Langfelder, P., Horvath, S. 2008. "WGCNA: an R package for weighted correlation network analysis". BMC Bioinformatics, *9*(1), 559.

Lefever, S., Anckaert, J., Volders, P.-J., Luypaert, M., Vandesompele, J., Mestdagh, P. 2017. "decodeRNA— predicting non-coding RNA functions using guilt-by-association". Database, *2017*.

Lerch, J. K., Martínez-Ondaro, Y. R., Bixby, J. L., Lemmon, V. P. 2014. "cJun promotes CNS axon growth". Molecular and Cellular Neuroscience, *59*, 97–105.

Li, R., Liang, J., Ni, S., Zhou, T., Qing, X., Li, H., … Pei, D. 2010. "A Mesenchymal-to-Epithelial Transition Initiates and Is Required for the Nuclear Reprogramming of Mouse Fibroblasts". Cell Stem Cell, *7*(1), 51–63.

Liao, T. T., Yang, M. H. 2017. "Revisiting epithelial-mesenchymal transition in cancer metastasis: the connection between epithelial plasticity and stemness". Molecular Oncology.

Ma, L., Bajic, V. B., Zhang, Z. 2013. "On the classification of long non-coding RNAs". RNA Biology, *10*(6), 924–933.

Marchese, F. P., Raimondi, I., Huarte, M. 2017. "The multidimensional mechanisms of long noncoding RNA function". Genome Biology.

McCabe, E. M., Rasmussen, T. P. 2021. "lncRNA involvement in cancer stem

cell function and epithelial-mesenchymal transitions". Seminars in Cancer Biology, *75*, 38–48.

Nieto, M. A., Huang, R. Y.-J., Jackson, R. A., Thiery, J. P. 2016. "EMT: 2016". Cell, *166*(1), 21–45.

Polo, J. M., Hochedlinger, K. 2010. "When Fibroblasts MET iPSCs". Cell Stem Cell, *7*(1), 5–6.

Ren, Q., Chen, J., Liu, Y. 2021. "LRP5 and LRP6 in Wnt Signaling: Similarity and Divergence". Frontiers in Cell and Developmental Biology, *9*.

Ruiz-Vela, A., Aggarwal, M., De La Cueva, P., Treda, C., Herreros, B., Martín-Pérez, D., … Piris, M. A. 2008. "Lentiviral (HIV)-based RNA interference screen in human B-cell receptor regulatory networks reveals MCL1-induced oncogenic pathways". Blood.

Rutenberg-Schoenberg, M., Sexton, A. N., Simon, M. D. 2016. "The Properties of Long Noncoding RNAs That Regulate Chromatin". Annual Review of Genomics and Human Genetics.

Seyfried, T. N., Huysentruyt, L. C. 2013. "On the origin of cancer metastasis". Critical Reviews in Oncogenesis.

Shin, J., Shin, Y., Oh, S.-M., Yang, H., Yu, W.-J., Lee, J.-P., … Kim, H.-S. 2014. "MiR-29b controls fetal mouse neurogenesis by regulating ICAT-mediated Wnt/β-catenin signaling". Cell Death & Disease, *5*(10), e1473–e1473.

Shu, X., Pei, D. 2014. "The function and regulation of mesenchymal-to-epithelial transition in somatic cell reprogramming". Current Opinion in Genetics & Development, *28*, 32–37.

Spano, D., Heck, C., De Antonellis, P., Christofori, G., Zollo, M. 2012. "Molecular networks that regulate cancer metastasis". Seminars in Cancer Biology.

St.Laurent, G., Wahlestedt, C., Kapranov, P. 2015. "The Landscape of long noncoding RNA classification". Trends in Genetics.

Stemmler, M. P. 2008. "Cadherins in development and cancer". Molecular BioSystems, *4*(8), 835.

Stemmler, M. P., Hecht, A., Kinzel, B., Kemler, R. 2003. "Analysis of regulatory elements of E-cadherin with reporter gene constructs in transgenic mouse embryos". Developmental Dynamics.

Tan, E. J., Olsson, A. K., Moustakas, A. 2015. "Reprogramming during epithelial to mesenchymal transition under the control of TGFβ". Cell

Adhesion and Migration.

Tang, Q., Chen, J., Di, Z., Yuan, W., Zhou, Z., Liu, Z., … Di, M. 2020. "TM4SF1 promotes EMT and cancer stemness via the Wnt/β-catenin/SOX2 pathway in colorectal cancer.". Journal of Experimental & Clinical Cancer Research : CR, *39*(1), 232.

Thiery, J. P., Acloque, H., Huang, R. Y. J., Nieto, M. A. 2009. "Epithelial-Mesenchymal Transitions in Development and Disease". Cell, *139*(5), 871–890.

Tsai, J. H., Yang, J. 2013. "Epithelial-mesenchymal plasticity in carcinoma metastasis". Genes and Development.

Uszczynska-Ratajczak, B., Lagarde, J., Frankish, A., Guigó, R., Johnson, R. 2018. "Towards a complete map of the human long non-coding RNA transcriptome". Nature Reviews Genetics.

Voutsadakis, I. A. 2015. "The network of pluripotency, epithelial–mesenchymal transition, and prognosis of breast cancer". Breast Cancer: Targets and Therapy.

Xia, S., Wang, X., Wu, Y., Zhou, T., Tian, H., Liu, Z., … Zhang, G. 2022. "MiR-22 Suppresses EMT by Mediating Metabolic Reprogramming in Colorectal Cancer through Targeting MYC-Associated Factor X". Disease Markers.

Youness, R. A., Gad, M. Z. 2019. "Long non-coding RNAs: Functional regulatory players in breast cancer". Non-Coding RNA Research, *4*(1), 36–44.

8. APPENDIX

Table 20. List of RT-PCR Primers.

Gene ID	forward	reverse	Restriction site
NH.3067	attcctcgagTTGGGGCAGTAATTAGGGA	cactaagcttGCGTGGCCGAGCGGTCTA	XhoI/HindIII
NH.13885	attcctcgagCGTCTACGTTACTGCTTTG	cactagatctGTTCCCGGTAGCTCGCCGT	XhoI/BglII
NH.3934	aattcagatctTCTAAAAGTGAGGGCGAAT	cactaagcttAGTGAGTAGAACTAAAGAT	BglII/HindIII
NH.14958	aattcagatctTCAGCTTGAACCGTTAGC	cactaagcttCACATCTCCAATTTCCTGA	XhoI/HindIII
NH.15444	attcctcgagGCTTTGTTTATTTTTAAGTACCGG	cactaagcttCTGAATTTTGTTTTGCAGAGTG	XhoI/HindIII
NH.19543	aattcagatctATCAGCTCCAATGCCAGAG	cactaagcttCCCAGCCTCTACCTGCTTG	XhoI/HindIII
NH.11835	attcctcgagCTCAAATGGTAGAGCGCTCG	cactaagcttGTTAAGTGCAAGACTTGAG	XhoI/HindIII
NH.19540	attcctcgagACCTGCCTCAGCTTCCCCA	cactaagcttTTACCAGTCACCATCTTCA	XhoI/HindIII
NH.6937	TGATGAAAAACTCCAAAGC	CCCACAGATCCAACCCATGG	None
NH.184	CAGTATATGTTACAGAATTGG	GAGACTAGCCTCGTGAGA	None
NH.2	GTTTCAAGGAAGTGTTTGC	AAGTCTTGGAAGCTTATGG	None
NH.963	AGGGAGATTGTGAAGGGCTC	AGCCACAAATTGTGGTGT	None
NH.24922	CTCCATTTCCCACTTCCGTG	AGAGTAGCAGGCAAGTGGAG	None
NH.26267	GCTCACTTTCTCACAGGATGG	GCCACAAATTGTGGTGTTAC	None
NH.7062	GTGGTTTTCCGGGGATGGCTG	GGAACTGATTCTCAGCCTGTG	None
NH.10008	TGAAAAGGACCCACTCATTG	CAGCAAGGGTGGAAAACTGG	None
NH.14607	ATATGTCTGTCATGTTTG	CTCAAAGACTATCTCTATG	None
NH.26245	ACCTCTATAACACAAAAGC	ATATGTCTGTCATGTTTG	None
NH.10007	GATCGTAATTCCTAATCAG	CTCAAAGACTATCTCTATG	None
NH.5736	CCAGGCTTTGCACCCCACTCC	CGAAGGTCAGGAGTTCAA	None
NH.10286	CACAGGACCTGACTGCAGGC	TGAGTCTGGTCCCCAGTGTTGG	None
NH.26239	CTGGTCTATGAAAATAAAA	TCGCGACCCCAAGAATTTA	None

NH.26268	GTTGAGACCACTACTTCT	TTGTCAATTAACTAATCC	None
NH.18336	TTTCGTCTTTATATTTCTAAG	GAACCAATTCTCCCAAGTTG	None
NH.22652	GCCAGAATCTTGGCTCAA	ATATGTCTGTCATGTTTG	None
NH.8369	GATGAGATCAAATGCTCTTAG	AATGGACAAAGGTTTAAT	None
NH.8369	GATGAGATCAAATGCTCTTAG	AATGGACAAAGGTTTAAT	None
NH.3067_IVTT	GCGAATTAATACGACTCACTATAGGGCTTAAGTATAAGGAGGAAAAAATACTAAGGCTTTGACTTATGGCTAGCTC	AAACCCCTCCGTTTAGAGAGGGGTTATGCTAGTTACTCAGTCCCTGATGCAGTAGGGCG	None
NH.3934_IVTT	GCGAATTAATACGACTCACTATAGGGCTTAAGTATAAGGAGGAAAAAATTCTAAAAGTGAGGGCGAATCCTTTGTT	AAACCCCTCCGTTTAGAGAGGGGTTATGCTAGTTAAGTGAGTAGAACTAAAGATGGAAG	None
NH.10286_IVTT	GCGAATTAATACGACTCACTATAGGGCTTAAGTATAAGGAGGAAAAAATCACAGGACCTGACTGCAGGCACGCACA	AAACCCCTCCGTTTAGAGAGGGGTTATGCTAGTTATGAGTCTGGTCCCCAGTGTTGGAC	None
NH.13885_IVTT	GCGAATTAATACGACTCACTATAGGGCTTAAGTATAAGGAGGAAAAAATTTTTTTAAATCACTTCGGCGTCTACGT	AAACCCCTCCGTTTAGAGAGGGGTTATGCTAGTTAGTTCCCGGTAGCTCGCCGTGATCG	None

Table 21. Primers used in colony PCR

GL2	CTTTATGTTTTTGGCGTCTTCCA
RV3	CTAGCAAAATAGGCTGTCCC

Table 22. List of qPCR primers

Gene ID	forward	reverse	Probe no
3934	CCAGTGAAGATTGGGCCTGT	TCGGGGTATAGGGACCACTG	57
3067	GCACCAATTCTTTCCAGGCG	TGATGTCAGACACAGCTCACC	5
CDH1	ATCCTCGCCCTGCTGATT	ACCACCGTTCTCCTCCGTA	18
psmd4	GGTCTTGGTGCCAGTCACTT	TAGAGACTCGAAGGGCCAGG	118

İZMİR BİYOTIP VE GENOM MERKEZİ
GİRİŞİMSEL OLMAYAN ARAŞTIRMALAR ETİK KURULU (İBG-GOEK)
KARARI

Toplantı Tarihi	: 04/03/2022	**Toplantı Günü**	: Cuma
Toplantı Sayısı	: 4	**Toplantı Saati**	: 10:30

Sayın Dr. Hani ALOTAIBI,

2022-010 Protokol No'lu; sorumlusu olduğunuz "Mezenkimal Epitelyal Dönüşüm Sürecinde Kritik Olan Uzun Kodlamayan RNA'ların İşlevsel Analizleri" başlıklı araştırmanın uygulanmasında etik açıdan sakınca olmadığına oy birliği ile karar verilmiştir.

Bilgilerinizi ve gereğini rica ederiz.

Prof. Dr. H. Alper BAĞRIYANIK **Başkan**	**Dr. Serap ERKEK ÖZHAN** **Başkan Yardımcısı**
Prof. Dr. Sedef AKGÜNGÖR **Üye**	**Prof. Dr. İnci ALACACIOĞLU** **Üye** (Katılmadı)
Prof. Dr. Gülgün OKTAY **Üye**	**Prof. Dr. Hilmi ORHAN** **Üye** (Katılmadı)
Prof. Dr. Şermin GENÇ **Üye** (Katılmadı)	**Doç. Dr. Canan Aslı YILDIRIM** **Üye**
Doç. Dr. Can KÜÇÜK **Üye**	**Dr. Öğr. Üyesi Yavuz OKTAY** **Üye**



AYÇA ÖZKAN

Personal Information

Contact Information

Contact Address	52 / [illegible] Apartmanı Dai[illegible] [illegible]hallesi
Telephone	(5[illegible]
Email	[illegible]u.tr
	ayc[illegible]
Web Page	

Education Information

05 March 2021 - Now (3 Year)
Yüksek Lisans, Tezli Program, DOKUZ EYLÜL ÜNİVERSİTESİ, TÜRKİYE
İZMİR ULUSLARARASI BİYOTIP VE GENOM ENSTİTÜSÜ, BİYOTIP VE SAĞLIK TEKNOLOJİLERİ (YL) (TEZLİ) (İNGİLİZCE)
Diploma Number: -

05 August 2015 - 27 October 2020 (5 Year 3 Month)
Lisans, Anadal/Normal Öğretim, İZMİR YÜKSEK TEKNOLOJİ ENSTİTÜSÜ, TÜRKİYE
FEN FAKÜLTESİ, MOLEKÜLER BİYOLOJİ VE GENETİK PR. (İNGİLİZCE)
Diploma Number: 4923
Cumulative Grade Point Average: 2.67 / 4.0

Experience Information

01 March 2021 - Now (3 Year) (Tam Zamanlı)
DİĞER, YÜKSEK LİSANS ÖĞRENCİSİ, İZMİR BİYOTIP VE GENOM MERKEZİ
Administrative Job: DİĞER

R&D Competency

Conference Papers

A. ÖZKAN, D. ESKİER, G. KARAKÜLAH & H. ALOTAİBİ, Functional Analysis of Noncoding RNAs Critical to the Mesenchymal Epithelial Transformation Process, Poster Sunumu, 8th International Congress of the Molecular Biology, 09 June 2022, 12 June 2022.

Generated Time: 16 February 2024 10:36:43

TÜBİTAK Scholarships and Fundings

Project Info	122Z886, Mezenkimal Epitelyal Dönüşüm Sürecinde Rol Alan Mikrorna'Ların Tanımlanması Ve Fonksiyonel Analizleri, Burslu, Yürürlükte, KBAG - Kimya Biyoloji Araştırma Destek Grubu, 1001 - Araştırma, ARDEB, Project Participation/Leave Dates: 9/8/23 - 1/19/24, Project Start/End Dates: 3/15/23 - 3/15/26. 219Z034, Mezenkimal Epitelyal Dönüşüm Sürecinde Kritik Olan Uzun Kodlamayan RNA'ların İşlevsel Analizleri, Burslu, Sonuçlandı, KBAG - Kimya Biyoloji Araştırma Destek Grubu, 1001 - Araştırma, ARDEB, Project Participation/Leave Dates: 3/19/21 - 7/15/23, Project Start/End Dates: 7/15/20 - 7/15/23.
BIDEB Supports	AYÇA ÖZKAN, Etkinlik Destekleri ve Eğitim Bursları Müdürlüğü, 2210-A Genel Yurt İçi Yüksek Lisans Burs Programı, Başvurusu Reddedildi, 2021 - 1, 3/1/21 - 2/28/23.

Number of Panelist/Observer/Reporter Jobs

Refereeing/Panelist/Outer Consultant Count	ARDEB/BİDEB 0	TEYDEB 0	Total 0
Number of Observer/Consultant Jobs	ARDEB/BİDEB 0	TEYDEB 0	Total 0
Number of Reporter Jobs	ARDEB/BİDEB 0	TEYDEB 0	Total 0