An-Najah National University

Faculty of Graduate Studies

Identification of Grhl3 Target Genes During the Mesenchymal to Epithelial Transition

By

Diana Ahmad Saleh Zeid

Supervisors

Dr. Ashraf Sawafta

Dr. Hani Alotaibi

This Thesis is Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Biology, Faculty Graduate of Studies, An-Najah National University, Nablus, Palestine.

2021

Identification of Grhl3 Target Genes During the Mesenchymal to Epithelial Transition

By

Diana Ahmad Saleh Zeid

This Thesis was defended successfully on 4/11/2021 and approved by:

Defense Committee Members	**Signature**
1. Dr. Ashraf Sawafta/ Supervisor	
2. Dr. Hani Alotaibi/ Second supervisor	
3. Dr. Raed Alkawni (Internal examiner)	
4. Dr. Osama Abdallah (External examiner)	

Dedication

This thesis is dedicated to Allah, my Creator, and Lord, Mohammed (peace be upon him), my great teacher and messenger who taught us the meaning of life, and Palestine, the warmest womb.

This work is also dedicated to my parents, Ahmad and Amnah, who have always loved me unconditionally and inspired me to work hard for the things I want to achieve by their wonderful example.

My cherished brothers and sisters, especially my favorite sisters, Kawther and Çağri, who is always at my side when things appear to be grim.

I dedicate this research to my beloved aunt Sana and my friends who inspire and support me, as well as all the people in my life who have touched my heart.

Acknowledgment

My deepest gratitude goes to Allah, who has provided all that was needed to complete this research.

Also, my big appreciation goes to my first supervisor, Dr. Ashraf Sawafta, for his great feedback, excellent encouragement, and guidance. Thank you so much.

I greatly appreciate my second supervisor, Dr. Hani Alotaibi, for his marvelous supervision, guidance, and encouragement. Sincere gratitude is extended to his generous participation in guiding, constructive feedback, kind support, and advice during my Thesis work. Thank you very much.

Thanks to my doctors, teachers, and all members of the biology department at An-Najah national university.

Many thanks to all of the staff members in Dokuz Eylül University, Izmir International Biomedicine, and Genome Institute, and Izmir Biomedicine and Genome Center for their kind support during my thesis project. As well as my colleagues and kind friends at IBG Canan Büşra Ersayar, Elif Hayden.

And I'd like to express my gratitude to my roommate at Izmir Çağrı Buldu, who stood by me and encouraged me when I was away from home; thank you so much. This work was supported by The Scientific and Technological Research Council of Turkey, Grant Number 117Z223.

إقرار

V

إقرار

أنا الموقع أدناه مقدم الرسالة التي تحمل عنوان:

Identification of Grhl3 Target Genes During the Mesenchymal to Epithelial Transition

أقر ما اشتملت عليه هذه الرسالة إنما هي نتاج جهدي الخاص، باستثناء ما تمت الإشارة اليه حيثما ورد، وان هذه الرسالة ككل او اي جزء منها لم يقدم من قبل لنيل أي درجة أو لقب علمي أو بحث لدى أي مؤسسة تعليمية أو بحثية أخرى.

Declaration

The work provided in this thesis, unless otherwise referenced, is the researcher's work and has not been submitted elsewhere for any other degree or qualification.

Student's Name: Diana Zeid ..... **اسم الطالب:** ديانا زيد

Signature: .... **التوقيع:** .....

Date: 4/11/2021 ..... **التاريخ:** ٤/١١/٢٠٢١

Table of Contents

List of Tables

List of figures

Table of Abbreviations

EMT	Epithelial to Mesenchymal Transition
MET	Mesenchymal to Epithelial Transition
Grhl3	Grainy head like 3
ECM	Extra Cellular Matrix
TGFβ	Transforming Growth Factor Beta
HGF	hepatocyte growth factor
EGF	epidermal growth factor
FGF	fibroblast growth factor
E-cad	Epithelial-cadherin
N-cad	Neural-cadherin
FGFR	Fibroblast Growth Factor Receptor
iPS	pluripotent stem
NMuMG	Normal Murine Mammary Gland
Hnf4a	Hepatocyte Nuclear Factor 4 alpha
GO	Gene Ontology
V(D)J	Variable (Diversity) Joining gene segments recombination
Nusap1	Nucleolar and Spindle Associated Protein 1
Cenpf	Centromere Protein F
IBG	Izmir Biomedicine and Genome Center
DMEM	Dulbecco's Modified Eagle's Medium
FBS	Fetal Bovine Serum
NEAA	Non-essential amino acids
siRNA	Small Interfering RNA
BME	Beta-mercaptoethanol
Oip5	Opa Interacting Protein 5 gene
CSV	Comma Separated Values
Hmgb2	High Mobility Group Box 2 gene
Deseq2	Deferential expression Sequencing
SCR	Scrambled control
Knl1	Kinetochore Scaffold 1
CNS	Conserved Sequences
RNA-Seq	RNA sequencing
NIS	sodium iodide symporter

Identification of Grhl3 target genes during the Mesenchymal to Epithelial Transition

By
Diana Ahmad Saleh Zeid
Supervisors
Dr. Ashraf Sawafta
Dr. Hani Alotaibi

Abstract

The epithelial to mesenchymal transition (EMT) and the mesenchymal to epithelial transition (MET) are essential processes that occur throughout embryonic development and are critical for organogenesis and morphogenesis, resulting in dynamic changes in cell identity. The transcription factors Grhl3 and other regulatory genes play a key role in regulating E-cad expression during the MET process, which is a hallmark of the procedure. This work aims to identify Grhl3's transcriptional targets during MET and assess their potential relevance.

To test the hypothesis that a transcriptional network that controls the stability of the epithelial state controls the dynamic regulation of MET and that this network may also play a role in the initiation of MET, Analysis of RNA sequencing data containing 14055 genes differentially expressed genes was carried out using pipelines for RNA sequencing analysis such as DESeq2 to identify Grhl3 target genes during MET bioinformatically. Then, using a loss of function approach, where Grhl3 was downregulated by siRNAs, putative Grhl3 target genes were validated by qPCR.

The results showed that cells could not initiate MET when Grhl3 was silenced. There is a correlation between the expression of the knocked-down Grhl3 and the chosen target genes in both RNA-Seq results and the experimental results, where they all were downregulated. Cenpf was the most downregulated with the Grhl3 RNA interference from the RNA-Seq. On the other hand, Knl1 was the most downregulated experimentally.

These findings suggest that Grhl3 is one of the essential transcription factors that direct the MET process, as well as that Grhl3, has many target genes that are differentially correlated with its gene expression and that Hmgb2, Nusap1, Cenpf, and Knl1 could be good candidates as Grhl3 target genes, which should be further investigated and more experiments carried out to emphasize the results.

Chapter One

Introduction

Chapter One
Introduction

1.1 Overview of Epithelial-Mesenchymal transition.

The epithelial to mesenchymal transition (EMT), which was first perceived as a characteristic of embryogenesis by Betty Hay during the 1980s (Lamouille, Xu, et al. 2014), and its reverse process mesenchymal to epithelial transition (MET) are critical processes used at several stages of embryonical organogenesis and morphogenesis in which cell identity change dynamically. They become morphologically discernible when singular cells or tissues are created by cell delamination during EMT or by cell assembling and re-epithelialization during MET. Since it was classed in the nineteenth century, two essential phenotypic cell morphologies are epithelial and mesenchymal cell morphology (Vanhoutteghem, Duval, et al 1989). Epithelial cells are set up as single-cell layers or multi-facet tissues with different tasks. Epithelial cells show apical-basal polarity, cohere, and interact through specific intercellular intersections such as tight junctions, adherens junctions, desmosomes, and gap junctions and are situated on a base membrane that assists with characterizing their physiology (Lamouille, Xu, Derynck 2014). on the other hand, mesenchymal cells don't make up a structured cell layer, nor do they have the similar apical–basolateral association and polarization of the cell-surface molecules and the actin cytoskeleton as epithelial cells. They contact adjoining mesenchymal cells just centrally and are not generally connected with the basal lamina. In

culture, mesenchymal cells have a spindle-shaped, fibroblast-like morphology, while epithelial cells develop as groups of cells that keep up with complete cell-cell coherence with their neighbors. (Thiery, Sleeman. 2006)

EMT process is an arranged series of interactions where cell-cell and cell-extracellular matrix (ECM) synergies are changed for the emission of the epithelial cells from the encompassing tissue, the cytoskeleton is redesigned to give the capacity to travel through a three-dimensional ECM, and another transcriptional program is prompted to preserve the mesenchymal form (Radisky 2005). The execution of an EMT is motioned by the degradation of the bottom-line of the basement membrane and the development of a mesenchymal cell that can move away from the epithelial layer where it emerged from (Kalluri, Weinberg, et al 2009). Cells going through EMT become transitory and intrusive, acquire resilience from apoptosis, increment secretion of degradative catalysts, and condensation basic extracellular network (ECM) along with an expanded discharge of new ECM constitutes (Banyarda, and Bielenberga 2015).

EMT is ordered into three types (figure 1.1); developmental (type I) EMT is needed for mesoderm development, neural crest delamination, the foundation of the heart valve, palatogenesis, and myogenesis (Thiery, Acloque, et al. 2009). After tissue injury, EMT gets actuated during fibrosis and wound recuperating (type II), but at the same time is unusually enacted during tumorigenesis when malignant cells begin to spread, attack, and form metastases (type III) (Kalluri and Weinberg 2009, Brabletz 2012).

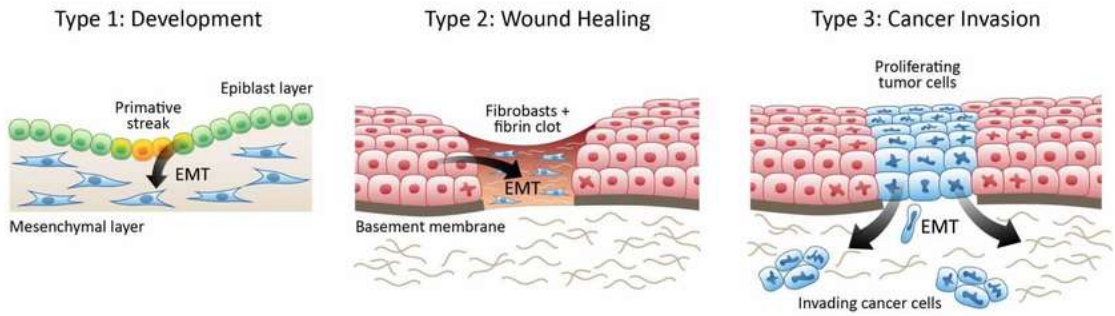


Figure 1.1 There are three different categories of EMT. Type 1 EMT is responsible for the formation of mesoderm and endoderm, as well as mobile neural crest cells. Epiblast, in particular, gives birth to primary mesenchyme by EMT. By using MET, the main mesenchyme can be re-induced to produce secondary epithelia. Type 2 During fibrosis and wound healing, EMT is activated after tissue injury. Type 3 EMT has the potential to develop into cancer, which can then spread by invasion and metastasis (Kalluri and Weinberg, 2009).

The Transforming Growth Factor Beta (TGFβ) is the most known EMT inducer and acts to actuate other EMT-related transcription factors (Illman et al., 2006; Xu, Lamouille, and Derynck, 2009). EMT can likewise be actuated by hepatocyte growth factor (HGF), epidermal growth factor (EGF), and fibroblast growth factor (FGF) (Grünert, 2003; Hay, 1995; Huber, 2005). When EMT is prompted with immoderate signalizing of TGFβ, it can lead to adverse conditions, for example, tumorigenesis and fibrosis (Tan, Olsson, and Moustakas, 2015; Gilbert, Vickaryous and Viloria-Petit, 2016)

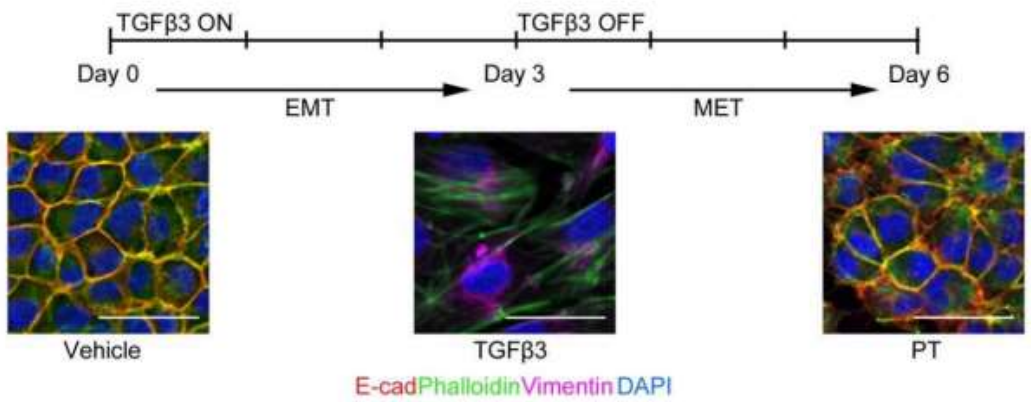


Figure 1.2 Immunofluorescent labeling was used to visualize EMT-MET pathways. Day 0 Vehicle indicates epithelial NMuMG cells, with epithelial marker E-cad (red) and actin staining phalloidin visible on the membrane (green). Cells take on a mesenchymal phenotype after being treated with TGF for 72 hours on Day 3 (TGF3). Mesenchymal marker vimentin can be observed surrounding the nucleus in purple, whereas epithelial marker E-cad is not visible. The cells grew in a normal medium for 72 hours after TGF was removed from the medium, in PT condition, which represents After treatment, cells returned to an epithelial state, and the epithelial marker E-cad (red) was visible in the membrane, just as it was in Vehicle. (Alotaibi lab project #114Z245, 2017)

Cytoskeletal rearrangements and modulation of the expression of many different genes, including cell adhesion molecules, are the hallmarks of EMT and MET. As for the adhesion molecules, the classical representatives are the Epithelial-cadherin (E-cad) and Neural-cadherin (N-cad). As per for E-cad (the Ca2+-dependent adhesion molecule), it belongs to the cell-cell adhesion molecules family, and it has a very significant role in the upkeep of cell proliferation and the ordinary structure of epithelial tissues (Graziano, Humar et, al. 2003), E-cad work requires a fine-tuned interaction with the catenin–cytoskeleton complex in the cytoplasmic space and the E-cadherin dimers of adjoining cells in the intercellular space, the group in explicit layer belts-like to shape the adherents intersections, the most pervasive sort of intercellular coherence (Handschuh G, Candidus S, Luber B et al. 1999). Revocation of the E-cadherin work Instigates loss of adherent's intersections and weakness

of cell adhesiveness and cell multiplication signalizing pathways (Mayer, Johnson, et al. 1993). Moreover, suppressed E-cadherin expression causes unusual morphogenesis and structure of epithelial tissues, loss of cell polarity and contact hindrance, unregulated development, and intrusion of nearby tissues have been exhibited in tumor cell composition (Vleminckx, Vakaet, Mareel, et al. 1991), which is why its downregulated during EMT. On the other hand, N-cad is an adhesion molecule that is common in non-epithelial tissues and is expressed in various sorts of cells like neural cells, endothelial cells, stromal cells, and osteoblasts (Van Roy 2014, Mrozik, Blaschuk et, al. 2018). Also, N-cad functions as an indicator of progressing EMT, and its expression has been related to the occurrence of different kinds of carcinoma (Wang, Ren et, al. 2016, Jennbacken, Tešan et, al. 2010, Hulit, Suyama et, al. 2007, Hui, Zhang et, al. 2013). N-cad intersection empowers the adjustment of fibroblast growth factor receptors (FGFR) to upgrade cell endurance and relocation (Loh, Chai et, al. 2019) which explains its upregulation during EMT. And so, as a result, the key evidence of a proper EMT is the downregulation of E-cadherin (E-cad) and the activation of N-cadherin (N-cad) expression that causes the loss of cell polarity, adherent structure, and of the epithelial gene signs (Stemmler 2008, Thiery, Acloque, et al. 2009), where they obtain an unpolarized mesenchymal structure in combination with increased cell motility and mesenchymal gene signs including fibronectin and vimentin expression (Kalluri and Weinberg 2009).

1.2 Mesenchymal-Epithelial transition

MET, to some point, is considered as the reverse process of EMT, which alludes to the reformist phenotypic change from unattached motile cells to firmly bound cells with a particular apical-basal polarity, and it's a fundamental embryonic program as well (Vestweber, Kemler, et al. 1985, Li, Zheng, et al. 2011). Here, mesenchymal cells acquire epithelial characteristics, including loss of N-cad and activation of E-cad expression (Kalluri and Weinberg 2009, Thiery, Acloque, et al. 2009, Brabletz 2012). In addition to orchestrating morphogenetic events during embryogenesis, the process of MET is also utilized by disseminating tumor cells required for colonization and formation of metastasis at distant sites (Thiery, Acloque, et al. 2009, Brabletz 2012). Moreover, when somatic reprogramming is induced in mesenchymal cells to generate induced pluripotent stem (iPS) cells, MET is necessary. It precedes the activation of the endogenous loci of the core pluripotency network (Li, Liang, et al. 2010). This is in part established by the activation of E-cad expression via direct binding of exogenous Klf4 to specific sites at the promoter (Chen, Yuan, et al. 2010). Interestingly, E-cad supports the initiation of MET and increases reprogramming efficiency (Redmer, Diecke, et al. 2011). Recently it was shown that the expression of E-cad is indispensable for keeping the pluripotent state of embryonic stem cells and that its depletion decreases the potential for somatic cell reprogramming during the generation of induced pluripotent stem cells due to failure to initiate MET.

Given discoveries from a scope of model techniques that cover parts of the intervening process, numerous procedures were proposed of MET that covers the whole-cell activity from (1) promoting the particularization of epithelial cells through a generative process or miniature natural signs, (2) foundation of cell polarity, (3) proliferation of MET through tissue, and (4) adjustment of new tissue structure. The exterior stresses and environs firmness alongside intercellular pressure influence the pace of MET in a tissue-wide means. With intersection and coherence adjustment, the returning of the cells to the epithelial state is achieved. (Kim, Jackson, and Davidson 2016). Figure1.3

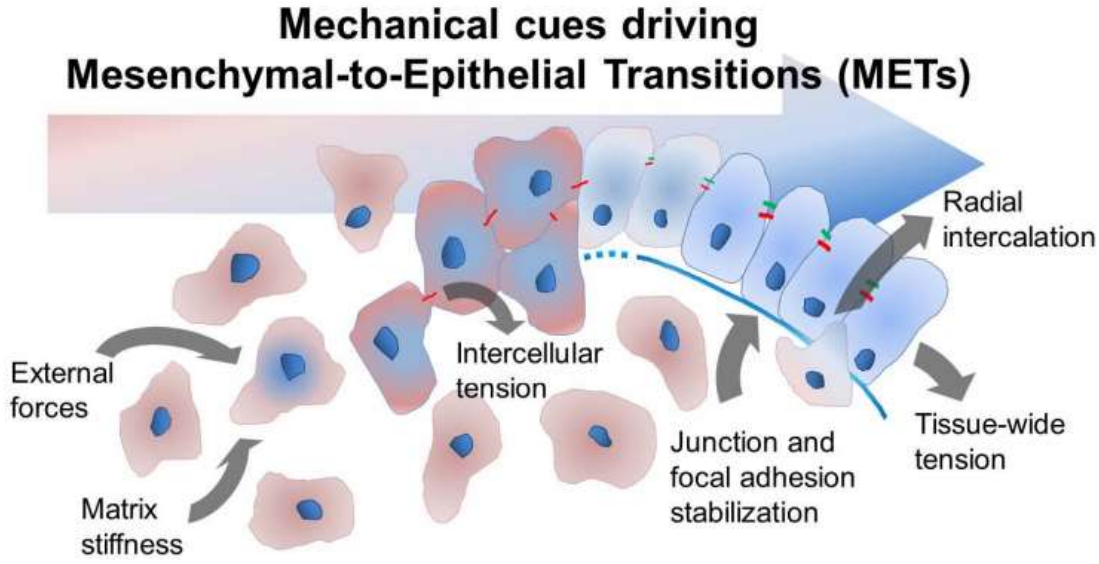


Figure 1.3 Mechanism for MET was suggested that cover the whole-cell activity from (1) boosting epithelial cell specialization through a generative process or miniature natural signs, (2) cell polarity foundation, (3) MET proliferation through tissue, and (4) tissue structure adjustment. External stressors and the hardness of the environment, and intercellular pressure have a tissue-wide impact on the rate of MET. The return of the cells to the epithelial state is achieved through intersection and coherence correction. (Kim, Jackson, and Davidson 2016).

In all EMT and MET processes, the transitions between epithelial and mesenchymal states are usually associated with switching in the expression state of E-cad among changes in other molecular signatures. An essential step of EMT is the downregulation of E-cad, whereas its expression is rapidly restored during MET. Induction of EMT is integrated by many different signals but shares the activation of intracellular EMT-inducers like Snail, Slug, Twist, Zeb1, Zeb2, and others, which are at the core of a transcriptional regulatory network controlling the EMT program (Thiery, Acloque, et al. 2009, Yilmaz and Christofori 2009, Brabletz 2012). In agreement with a required rapid downregulation of E-cad, all of these transcription factors are, in fact, repressors of E-cad expression. They all bind to several evolutionary conserved E-boxes present in the proximal promoter (Batlle, Sancho, et al. 2000, Cano, Perez-Moreno, et al. 2000, Carver, Jiang, et al. 2001, Comijn, Berx, et al. 2001, Bolos, Peinado, et al. 2003, Yang, Mani, et al. 2004, Eger, Aigner, et al. 2005). However, the immediate activation of the Cdh1 locus during MET can only partly be explained by the release and downregulation of these repressors. The Cdh1 promoter alone, including all known E-box elements, was shown to be insufficient for conferring strict cell-type specificity (Stemmler, Hecht, et al. 2003). On the other hand, intron 2 of E-cad has emerged as an essential regulator of Cdh1 expression (Stemmler, Hecht, et al. 2005).

The EMT generates all mesenchymal cells in early embryogenesis. Contrary to what was believed previously, MET isn't the exact inverse of EMT

regarding gene regulation but instead governed by a different set of regulators. Although the phenotypical properties of the cells revert into their epithelial express, the mechanism behind the interaction is unique but to be specified entirely. E-cad relocalization to the cell membrane, the reconstruction of cell-medium bonds, revamping the cytoskeleton coupled with apicobasal polar morphology, and the upregulation of E-cad and downregulation mesenchymal factors are adequately not to clarify how MET functions (Kim, Jackson, and Davidson, 2017). Beginning to end, both EMT and MET are constrained by numerous transcriptional, post-transcriptional, and post-translational controllers on each progression (Lamouille et al., 2013).

1.3 The NMuMG cell line (Normal Murine Mammary Gland cell line) is a well-known model for investigating EMT-MET transitions.

(Miettinen, Ebner, et al. 1994). Following TGF treatment, these cells will change from epithelial to mesenchymal, losing E-cad expression, displacing cortical actin to stress fibers, and upregulating N-cad and Vimentin. When TGF is removed from the culture medium, the cells use MET to return to an epithelial condition. The substantial changes in gene signature are represented in the easily observed phenotypic traits, making NMuMG cells an ideal model for studying several aspects of EMT-MET transitions. It was feasible to identify important elements of E-cad regulation and discover novel enhancers inside the second intron of E-cad using NMuMG cells.

1.4 Grainy-head like 3

The transcription factors Grainy head-Like 3 (GRHL3; likewise alluded to as GET1) and other regulating genes contribute significantly to the regulation of E-cad expression during MET, which is a hallmark of the process. Grhl3 is a gene that encodes a protein that acts as a transcriptional regulatory factor during development, which is a member of the grainy head-like family where it was first discovered in mutated Drosophila that fails to undergo occlusion of the neural tube and unusual granular scleritis where it gained the phenotype and name of grainy- head like family (Frisch, Farris, and Pifer 2017). Grhl3 attaches straightforwardly to the consensus DNA sequence 5'- AACCGGTT-3' functioning as an activator and repressor on some particular target genes (Luck, Kim 2020). Mice with a total Grhl3 knockout showed substantial epidermal hyper- multiplication and sensitivity to chemical carcinogenesis, showing a powerful tumor-suppressive impact of this Grainy head-like group member (Darido, Georgy, et al 2012). Grhl3 is fundamental for early-stage epidermal dissemination and barriers arrangement, mature epidermal repairing, actuates gene expression programs needed for cell coherence, lipid formation, cornified envelope development, and protein crosslinking (Lin KK, Spencer, et al. 2006 and Gordon, Zeller, et al 2014). Intriguingly, GRHL3 is likewise fundamental for typical keratinocyte relocation during eyelid closing and wound recuperating (Caddy, Wilanowski, et al 2010 and Bhandari, Mannik, et al 2008 and Hislop, Caddy, et al 2008) where it regulates gene expression programs that control the motility of keratinocytes

and inhibit the precursor and differentiation states. However, GRHL3 regulation of both delamination and movement of a single cell type remains ineffectively comprehended (Klein, Lin, et al 2017).

Researchers pinpoint important features of E-cad regulation and discover novel enhancers within the second intron of E-cad using NMuMG cells. The transcription factors Grhl3 and Hnf4a were found to have a substantial role in controlling E-cad expression during MET. Grhl3 was discovered to be an absolute need for the start of MET. This is partly due to their inability to upregulate E-cad expression (Alotaibi, Basilicata, et al. 2015), as Grhl3 was also discovered to have a role in regulating Hnf4a by binding to a well-conserved motif within the Hnf4a promoter P2 (Alotaibi, Basilicata, et al. 2015). Hnf4a is an essential component of various epithelia's transcription factor family; it was recently discovered to have a crucial function in liver biology and can also repress EMT regulators such as Snail and Slug.

Furthermore, both in cell culture and in vivo, its downregulation activates the mesenchymal program in differentiated hepatocytes (Santangelo, Marchetti, et al. 2011). Hnf4a deficiency has also been linked to the progression of liver tumors (Lazarevich, Shavochkina, et al. 2010). However, little is known about the role of Hnf4a and its transcriptional partners in tumorigenesis-driven MET development. The features of Grhl3 regulation during MET have recently been established. Elf3 was discovered to be an absolute necessity for a successful MET. Its impact on the advancement of MET was twofold: on the one hand, it was necessary for the appropriate migration of E-cadherin to

the plasma membrane, and on the other hand, it was required for the activation of the Grhl3 locus (Sengez, Aygun, et al. 2019).

Grhl3 was proven to be an absolute need for the start of MET. RNAi-mediated knockdown of Grhl3 prevented cells from initiating the MET program. Their failure to upregulate E-cad expression is one reason for this (Alotaibi, Basilicata, et al. 2015) figure 1.4.

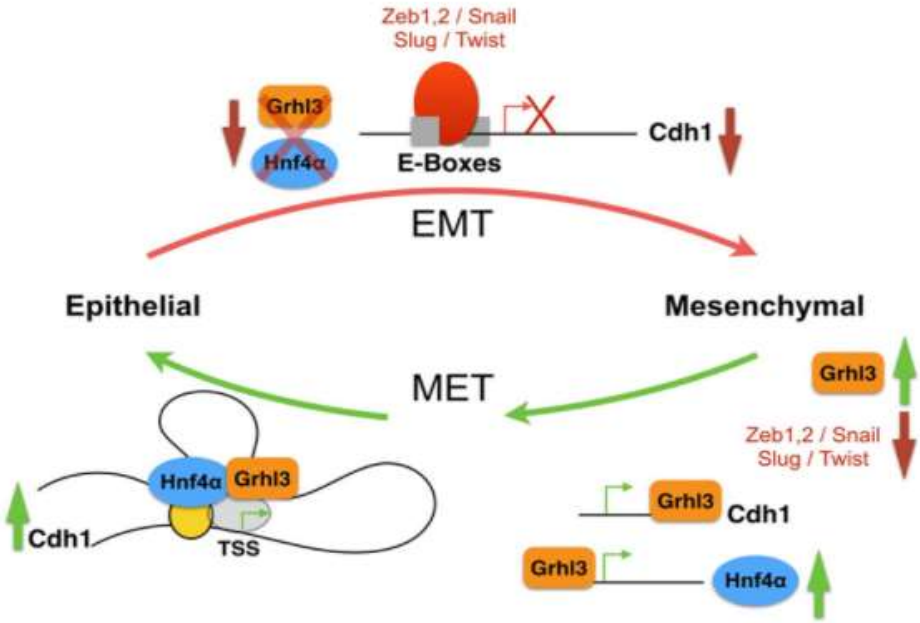


Figure1.4 The activation of intracellular EMT-inducers like Snail, Slug, Zeb1, and others, which are transcription factors acting as repressors of the Ca2+-dependent adhesion molecule (E-cadherin), which is at the center of EMT-MET, must be successfully downregulated for EMT to proceed, and its expression must be quickly restored for proper MET to occur. A feed-forward loop composed of the transcription factors Grhl3 and Hnf4a has recently been discovered to be critical for the advancement of MET. Grhl3 has emerged as a prerequisite for the start of MET. RNAi-mediated knockdown of Grhl3 prevented cells from initiating the MET program. Their failure to upregulate E-cad expression is one reason for this. Furthermore, we discovered that Grhl3 contributes to Hnf4 regulation by binding to a well-conserved region within the Hnf4 promoter P2 (Alotaibi, Basilicata, et al. 2015).

1.5 The control of abnormally regulated EMT is a critical goal in disease progression monitoring.

There is currently a significant number of data in the literature studying the EMT program during development and carcinogenesis. The MET program is still lacking in a similar understanding. MET is required for the colonization of metastatic tumor cells at distal locales, much as EMT is required for tumor development. As a result, a greater understanding of how MET is regulated is critical for improved metastatic disease management.

The unpublished data demonstrate that MET is dynamically regulated and that MET is not just the inverse of EMT, as current literature suggests. We hypothesize that a transcriptional network that controls the stability of the epithelial state controls the dynamic regulation of MET and that this network may also play a role in the initiation of MET.

Chapter Two
Materials and Methods

Chapter Two

Materials and Methods

2.1 Type of the Study

This study seeks a logical and computational methodology using in vitro generated experiments.

2.2 Time and Place of the Study

This study was directed at Dokuz Eylül University, Izmir International Biomedicine and Genome Institute (DEU-IBG). Experiments were carried out in Izmir Biomedicine and Genome Center (IBG) between March 2020 and February 2021.

2.3 Population and Samples of the Study

None of the experiments of the research were carried on human beings. The examination was done in vitro utilizing commercial cell lines and in silico techniques using openly accessible data sets.

2.4 Tools and Databases

2.4.1 Gene Pattern

2006- 2020 Regents of the University of California, Broad Institute, MIT.

Website: https://www.genepattern.org/.

Gene Pattern is a well-established logical work process framework that gives access to many genomic analysis instruments. Utilize these examination

apparatuses as building squarest configure complex investigation pipelines that catch the strategies, boundaries, and information used to create examination results.

For 14055 genes used to assess Grhl3 candidate target genes, differential gene expression was carried using pipelines for RNA-Seq, DESeq2. One prevalent undertaking in this bioinformatics tool is analyzing high-throughput RNA-seq information to discover differentially expressed genes across clusters of tests or phenotypes. DESeq2 is a famous calculation for analyzing RNA-Seq information, which assesses the change mean depending on high-throughput check information, and decides differential expression dependent on a negative binomial circulation. DESeq2 enhances the recently distributed DESeq calculation by improving the steadiness and interpretability of expression gauges.

2.4.2 The gene ontology resource (GO Enrichment Analysis)

Current release 2020-06-01: 44.411 GO terms | 7.975.639 annotations 1.558.956 gene products | 4.611 species. http://geneontology.org/.

An ontology is a proper portrayal of an assortment of information inside a given area. For the most part, ontologies comprise many classes (or terms or ideas) with relations that work between them. The Gene Ontology (GO) knowledgebase is the world's most significant data source on gene function. This information is understandable and machine-readable and is an

establishment for computational analysis of extensive molecular science and genome tests in biomedical studies.

For 14055 genes included in this study, GO enrichment analysis was carried to interpret the functions of the genes and gene products in terms of molecular function, cellular components, and biological processes in correlation to the Grhl3 gene.

2.4.3 mVista

a bioinformatic analytical tool from the Joint Genome Institute of the University of California, mVista, was used to visualize, determine and analyze the conserved regions of Grhl3 and some of its target genes among different species.

The location of the gene region of Grhl3 in Mus musculus on chromosome 4 with the latest mouse genome mm10 (December 2011) release is as follows; chr4:135,541,888-135,573,630.

2.4.4 UCSC Genome browser

UCSC Genome browser from the University of California Santa Cruz was used. It is an intuitive site offering admittance to genome grouping information from an assortment of vertebrate and invertebrate species and significant model creatures, incorporated with an enormous variety of adjusted annotations. Up to 200 out of 14055 candidate Grhl3 target genes were analyzed using the genome browser tool by graphically viewing the sequences of the genes, their location, and annotations in correlation to Grhl3.

75% downregulation percentage was considered for measuring the correlation of gene expression concerning Grhl3 from the DESeq results. The criteria were followed to select the target genes from the list of the genes was:

- Grhl3 binding to the target gene.
- Presence of TATAA box to determine the promoter region and if the Grhl3 gene binds to it.
- Presence of conserved regions for the target genes.
- Binding of Grhl3 at the conserved regions of the target genes.

2.5 Tools for Data Collection

Normal Murine Mammary Gland (NMuMG) Cell Line

The NMuMG cell line utilized in this work was obtained commercially (ATCC: The Worldwide Bioresource Center/CRL-1636).

2.6 Cell Culture

Cell culture tests that were done to grow the cells utilizing High Glucose Dulbecco" s Adjusted Eagle" s Medium (DMEM) (Gibco 41965-039) with adding in of 10% Fetal Bovine Serum (FBS) (Gibco 1347D), 1% Non-essential amino acids (NEAA) (Gibco 11140-035) and 1% Penicillin/Streptomycin antibiotic agents (Pen/Strep) (Gibco 15140-122) on 100 mm cell culture plates. The composition of the complete medium is presented in Table 2.1.

2.7 Preparation of Complete DMEM.

The medium contained 10% Fetal bovine serum.

1% Non-essential amino acids and 1% antibiotic

Table 2.1 Preparation of Complete DMEM. The medium consists of 10% Fetal bovine serum, 1% non-essential amino acids, and 1% antibiotics

Medium components	Concentrations/ mL
DMEM	88%
FBS	10%
NEAA	1%
Pen/Strep	1% (0.1mg/mL)

Incubation of the cells was done in CO2 cell culture (Eppendorf New Brunswick Cosmic system 170S) at 37oC with 10% CO2. Handling of the cells was carried out in a BSL-2 biosafety cabinet. As indicated by the confluency of the cell culture plate that the cells were filled in, when the confluency is around 80%, the subculture of the cells was done. The technique is as per the following:

- remove the medium from the plate
- Wash the cells with 10 mL 1X PBS twice
- Aspirate the PBS from the plate
- Add 1 mL of 0.25% Trypsin-EDTA onto the cells
- Incubate at 37oC inside the incubator for 1-2 minutes until the cells are detached from the plate surface
- Add 9 mL of Complete DMEM to the cells to inactivate Trypsin-EDTA

- Collect the cells in a falcon tube and centrifuge at 1200 rpm for 2 minutes
- Aspirate the medium keeping the cell pellet intact
- Flick the falcon gently to soften the cells pellet
- Add 10 mL of Complete DMEM and mix to resuspend the cells by pipetting
- Prepare another plate with 9 mL of Complete DMEM
- Add 1 ml of the cell combination into the plate

In each new subculture of NMuMG cells, 1μL/mL (10μg/mL) of insulin (Sigma-Aldrich/ I9278-5ML) was freshly added into the medium. The subsequent visualization of the cells was done utilizing inverted phase-contrast microscopy.

2.8 EMT-MET

NMuMG cells were treated with 5ng/mL TGFβ3 for 72 hours to change their form into mesenchymal cells. After EMT, the cells would return to their epithelial state after the elimination of TGFβ3 by washing the cells with 1X PBS multiple times and changing the media with the addition of insulin.

2.9 Gene Silencing with siRNAs

Epithelial NMuMG cells were held under standard culture conditions, and the cells that were younger than the 6th passage were picked to use for the gene knockdown experiment. The cells were transfected with siRNAs utilizing

Lipofectamine RNAiMAX Transfection Reagent (#1836079 Invitrogen) to evaluate the distinction of expression levels of Grhl3 on MET.

The protocol is as follows:

- NMuMG cells were gathered from the plate similarly as they would have been while subculturing and were calculated.
- The cells were cultivated into 6-well plates as 100,000 cells/well
- 1ul/ml of insulin and 0,5 ul/ml TGFβ3 were added to each well
- Plates were placed into a 37oC 10% CO2 incubator for 72 hours for EMT
- After 72 hours when cells were in the mesenchymal state, MET was initiated by removing TGFβ, immediately followed by siRNA transfection using Lipofectamine RNAiMAX
- After transfection, cells were incubated at 37oC for 48 hours, and the media changed with complete DMEM
- After media change, cells were further incubated at 37oC for 24 hours and collected at 72 hours as the PT72 condition was met.
- The cells were then collected for RNA isolation.
- RNAs were then converted into cDNAs, and gene expression levels were done by qPCR.

2.10 RNA isolation

The RNA isolation was done utilizing the NucleoSpin RNA set (Macherey-Nagel #740955.50), as follows:

- The cells in the 6-well plate were washed with 1X PBS multiple times.
- 3,5 µL BME (Betamerchaptoethanol) was added onto 350 µL of RA1 buffer per well, and this would be the lysis buffer to separate the cells from the base surface of the wells.
- The cells were gathered from the wells and put into the pink-sectioned filter microcentrifuge tubes.
- They were centrifuged at 11000 G for 1 minute.
- The supernatant was kept, and pink columns were disposed of.
- 350 µL 70% Ethanol was added onto the supernatants and pipetted 5-6 times.
- The mixture was stacked onto new blue-sectioned column tubes.
- They were centrifuged at 11000 G for 30 seconds.
- The blue column was taken out and placed into a new lidless tube, and the leftover supernatant was disposed of.
- 350 µL of MDB buffer was added onto the blue-columned filter tubes and centrifuged at 11000 G for 1 minute.

- The excess fluid at the lower part of the tube was disposed of, and the tube with the column was centrifuged again at 11000 G for 30 seconds to dispose of the surplus Ethanol.
- DNAse mix was set up as follows; 10 µL rDNAse 90 µL Reaction buffer.
- From this 100 uL mix/well, 95 µL of it was added onto the column in the tube and incubated at room temperature for 15 minutes.
- 200 µL of RAW2 buffer was added onto the columns and centrifuged at 11000 G for 30 seconds.
- The columns from the tubes at that point were put onto new lidless microcentrifuge tubes.
- 600 µL of RA3 buffer was added onto the columns and centrifuged at 11000 G for 30 seconds.
- The excess fluid at the lower part of the tube was disposed of, and the columns were centrifuged again at 11000 G for 30 seconds.
- 250 µL of RA3 buffer was added onto the columns and centrifuged at 11000 G for 2 minutes.
- The columns were taken out and placed into new microcentrifuge tubes with covers.
- A proper measure of nuclease-free water was added onto the columns and centrifuged at 11000 G for 1 moment to elute the RNAs.

- RNA concentrations were then estimated with a spectrophotometer (NanoDrop).

2.11 cDNA Synthesis

cDNA production was achieved utilizing Maxima First Strand cDNA Synthesis Kit (Thermo Fisher#K1642). The protocol is described in Table 2.2.

Table 2.2 cDNA synthesis mix. The mixture was set up as per the protocol of the producer. template RNA was utilized in 1 ug concentration.

cDNA mix	Concentrations
5X Reaction Mix	4 µL
Maxima Enzyme Mix	2 µL
Template RNA	1 µg
Nuclease-free water	up to 20 µL
Total volume	20 µL

RNAs that were isolated and measured were set up as 1ug per tube for cDNA synthesis. The mixture above was ready for each trial exclusively and put into RNAase-free tubes and kept on ice. The thermo-cycler was set to proper conditions;

- 25oC for 10 minutes
- 50oC for 30 minutes
- 85oC for 5 minutes
- 4oC at infinity.

After the cDNA synthesis reaction was done, cDNAs could be utilized straightforwardly for qPCR or kept at - 20oC for as long as a month or at -86oC for longer timeframes.

2.12 qPCR (Real-Time PCR)

The cDNAs that would be utilized for qPCR concentrations were reduced 1/20 before use and kept at 4oC. The qPCR mixture was set up as depicted in table 2.3 and added at each well of the 96-well qPCR plate (MicroAmp Fast Optical 96-well reaction plate from Applied Biosystems #4346906).

Table 2.3 qPCR master mix

Master mix	concentrations
TaqMan Mastermix	6.25 µL
Forward Primer	0.5 µL
Reverse Primer	0.5 µL
Probe	0.25 µL
Nuclease-free water	2.5 µL
Total volume	10 µL

After placing the combination in each well, 2.5 µL of cDNA were added onto their wells that were related to every one of them.

Table 2.4 Forward and reverse sequences of the cDNA probes.

Probe	Forward	Reverse	ProbeID
Mhprt	tcctcctcagaccgct	cctggttcatcatcgctaatc	95
Cenpf	ggagctctcccagcttga	cagcgctttgttctcagtca	108
Grhl3	aaggaagatgtcgaatgaacttg	tcgtcctcattactgtagggaaa	100
Hmgp2	agtctctctgcggaggtctg	gtcacccttgcccatgac	38
Nusap1	gttttctaaacttgggaacaataaaag	tggattccattttcttaaaacga	41
Oip5	aaggttttctcaagtccagcac	ggcagaaatggcctctca	46
Knl1	tcatcatcacctttggagagc	cagagactgaaaactgagtgcatta	4

Chapter Three
Results

Chapter Three

Results

3.1 Gene Pattern

To use gene patterns, one must convert the data file format to match the module of interest since each requires a specific format. When running the analysis module, visualization module, or pipeline, Gene Pattern shows the parameters for the chosen modules. Frequently, at least one of these parameters is input documents, which must have a specific format.

To get the Differential Expression workflow for RNA-Seq data based on the DESeq2 package from Bioconductor, several formats needed to be done and converted.

One main format is the GCT file format which is a tab-delimited file format that describes an expression dataset. For 14055 genes, files were converted from comma-delimited file CSV to tab-delimited values (GCT format). The GCT form is organized in an interchangeable structure where the columns represent genes rows represent samples.

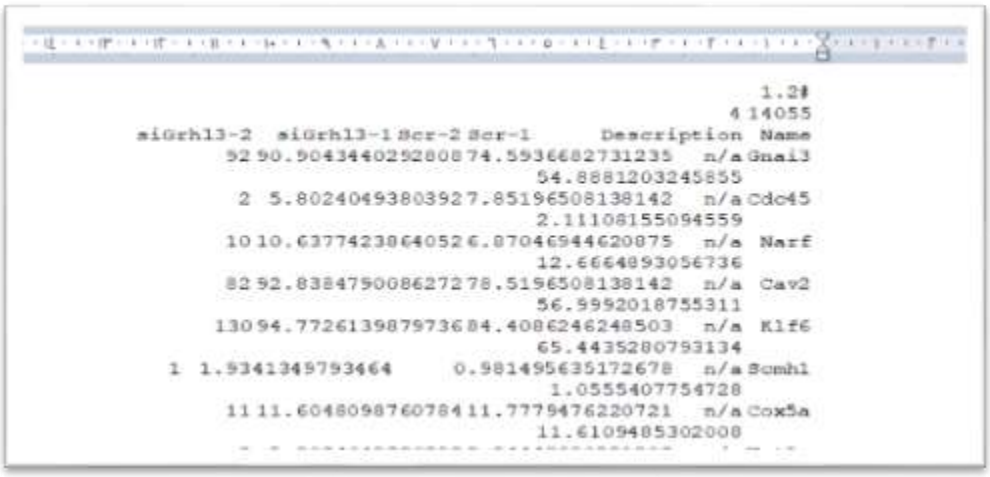


Figure 3.1 GCT file format.

Another important format is the CLS file format using the cls file creator module. The CLS format design characterizes phenotype (class or format) marks and couples each example in the expression information with a name. It utilizes spaces or tabs to isolate the fields.

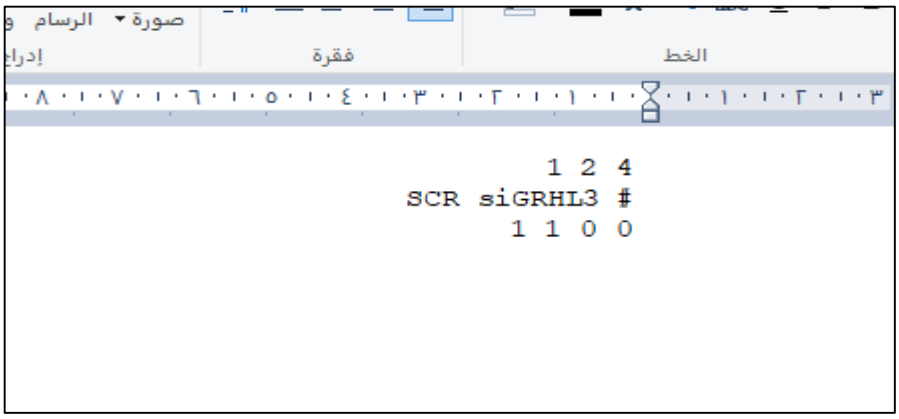


Figure 3.2 CLS file format

Then the GCT and Cls files were uploaded into their specific fields, and then ran the analysis.

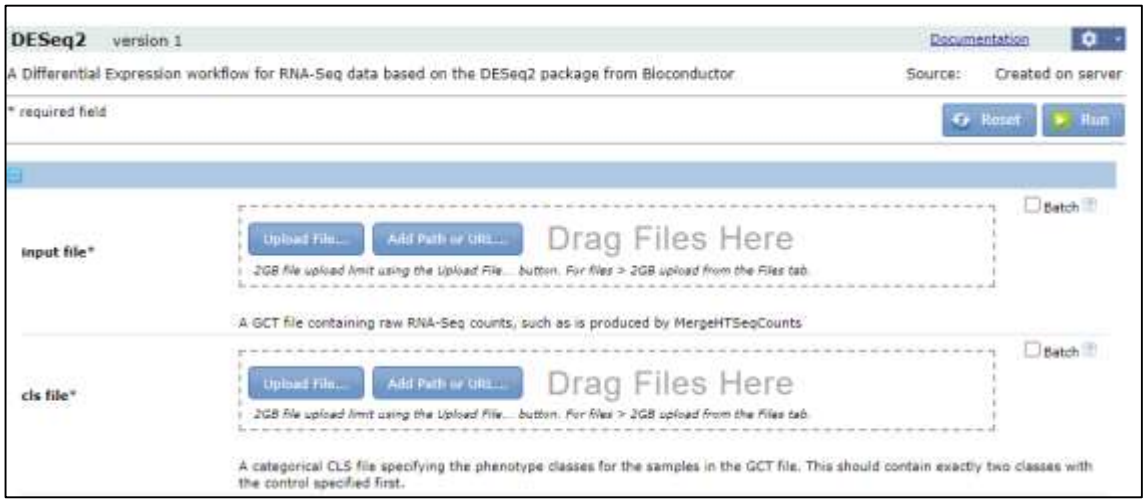


Figure 3.3 A differential Expression workflow for RNA-Seq data based on DESeq package from Bioconductor.

The analysis output consists of the following files.

- de-list-deseq2.tsv: Table containing the significantly differentially expressed genes. The columns include
    - baseMean = the average of the normalized counts taken overall samples.

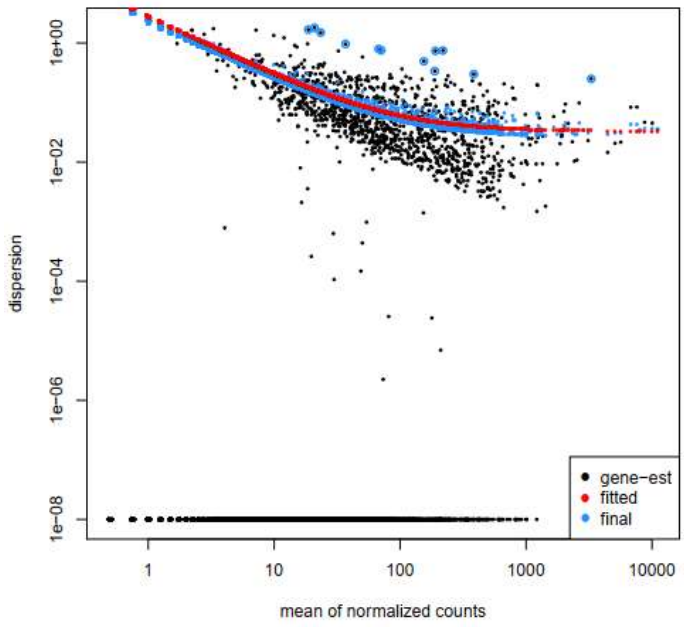


Figure 3.4 the average of the normalized counts taken over all 14055 samples.

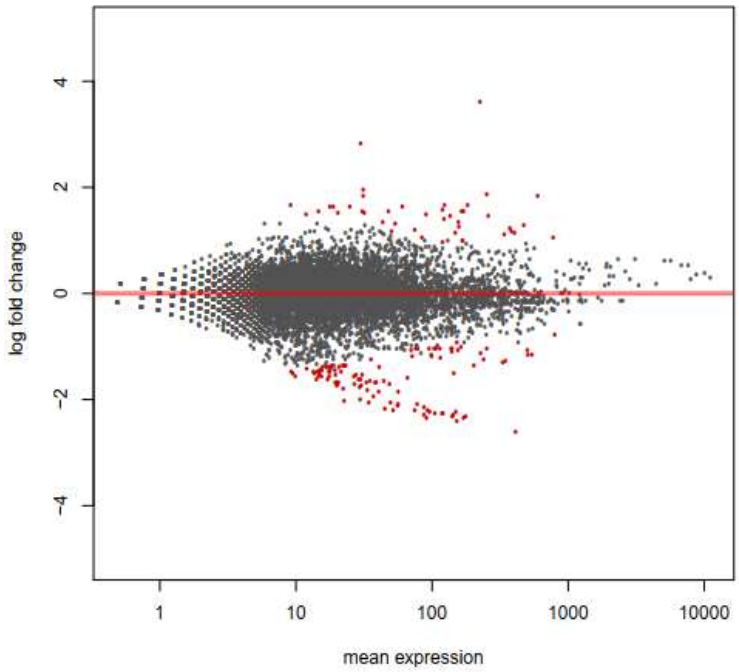
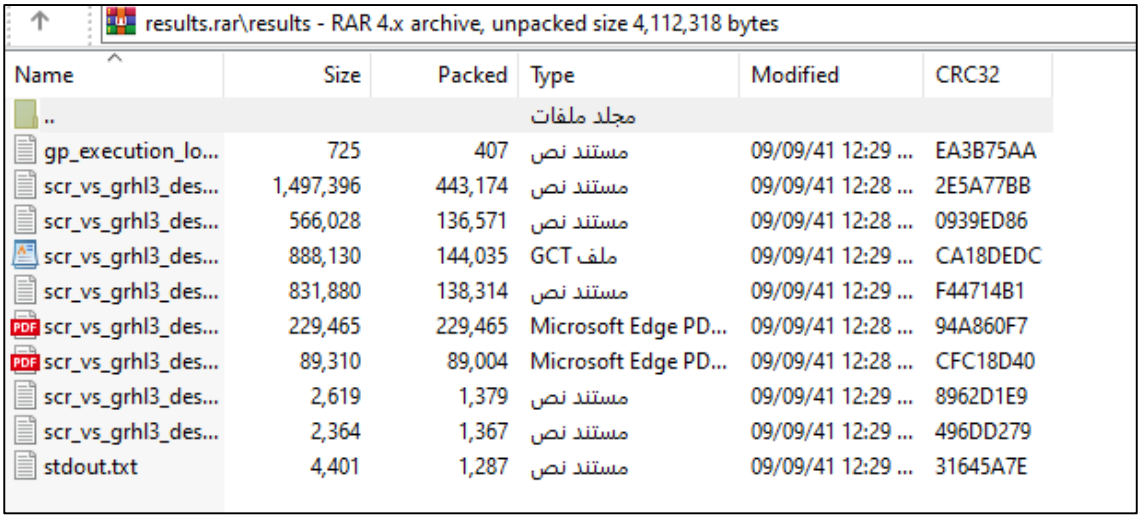


Figure 3.5 Log2 fold change between the groups (SCR vs. siGRhl3).

- Change = log2 fold change between the groups
- lfcSE = standard error of the log2FoldChange estimate
- stat = Wald statistic
- p-value = Wald test p-value
- padj = Benjamini-Hochberg adjusted p-value

- de-list-deseq2.bed: The BED version of the results table contains genomic coordinates and log2 fold change values.
- summary.txt: Textual summary of the differential expression results, including information on filtering and outliers.
- A deseq2_report. pdf: A PDF file containing:

- MA scatter plot where the significantly differentially expressed genes are highlighted.
- A plot of dispersion estimates at different count levels
- A plot of the raw and adjusted p-value distributions of the statistical test.



Name	Size	Packed	Type	Modified	CRC32
..			مجلد ملفات		
gp_execution_lo...	725	407	مستند نص	09/09/41 12:29 ...	EA3B75AA
scr_vs_grhl3_des...	1,497,396	443,174	مستند نص	09/09/41 12:28 ...	2E5A77BB
scr_vs_grhl3_des...	566,028	136,571	مستند نص	09/09/41 12:28 ...	0939ED86
scr_vs_grhl3_des...	888,130	144,035	GCT ملف	09/09/41 12:29 ...	CA18DEDC
scr_vs_grhl3_des...	831,880	138,314	مستند نص	09/09/41 12:29 ...	F44714B1
scr_vs_grhl3_des...	229,465	229,465	Microsoft Edge PD...	09/09/41 12:28 ...	94A860F7
scr_vs_grhl3_des...	89,310	89,004	Microsoft Edge PD...	09/09/41 12:28 ...	CFC18D40
scr_vs_grhl3_des...	2,619	1,379	مستند نص	09/09/41 12:29 ...	8962D1E9
scr_vs_grhl3_des...	2,364	1,367	مستند نص	09/09/41 12:29 ...	496DD279
stdout.txt	4,401	1,287	مستند نص	09/09/41 12:29 ...	31645A7E

Figure 3.6 Results from DESeq2

3.2 Comparative Marker selection Module.

Another essential module was used for the analysis, which is the comparative marker selection. The Comparative Marker Selection module takes as info a dataset of expression profiles from tests having a place with two classes (SCR and siGrhl3 in this research) and, executing the statistical tests a test statistic (e.g., t-test) which surveys the relationship of the gene's expression profile with a class format. Subsequently, it recognizes marker genes that distinguish between the classes. Again, the GCT and Cls files are required.

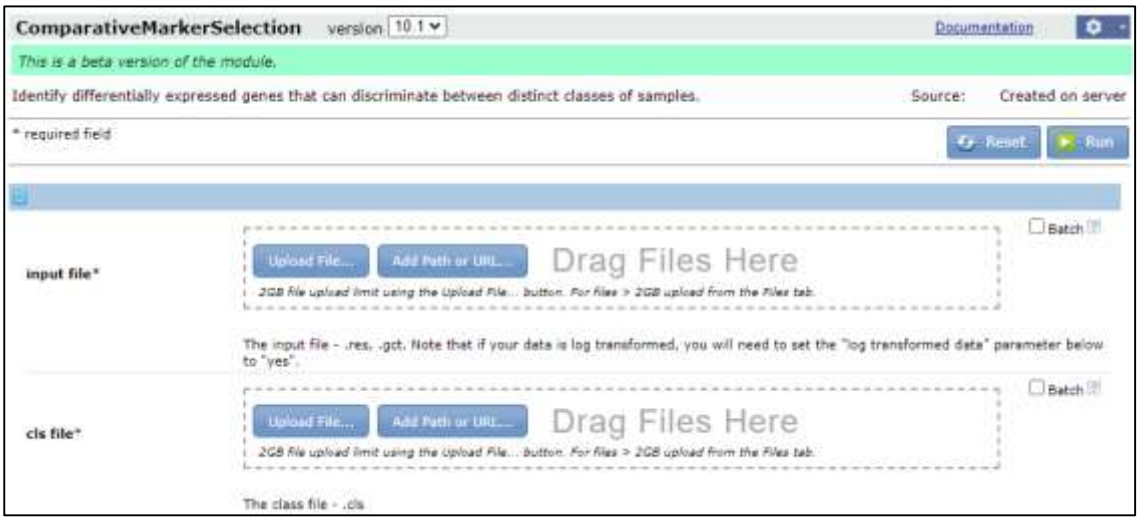


Figure 3.7 Identifying differentially expressed genes that can discriminate between distinct classes of samples.

The results from the Comparative Marker Selection algorithm can be viewed with the Comparative Marker Selection Viewer in addition to a Heat-Map that shows a heat map design where the highest qualities are shown as the reddest (hot), the lowest qualities are shown as the bluest (cool), and the midmost qualities are a lighter shade of one or the other blue or red. This is an exceptionally advantageous approach to show three-dimensional information (tests, highlights, and qualities).

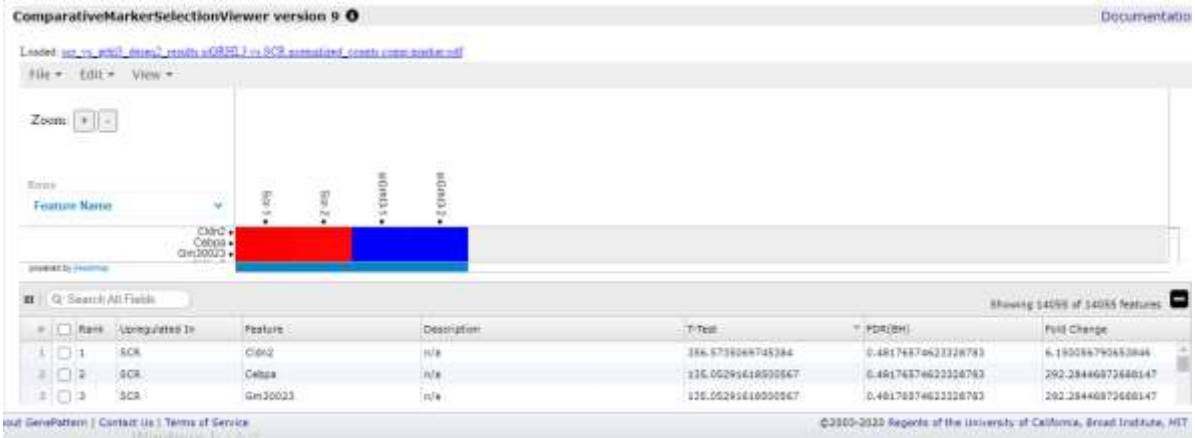


Figure 3.8 Comparative marker selection viewer showing the Algorithm, in addition to a heatmap, T-Test, FDR, and Fold change.

Extract Comparative Marker Results creates a derived dataset and feature list file from the results of Comparative Marker Selection.

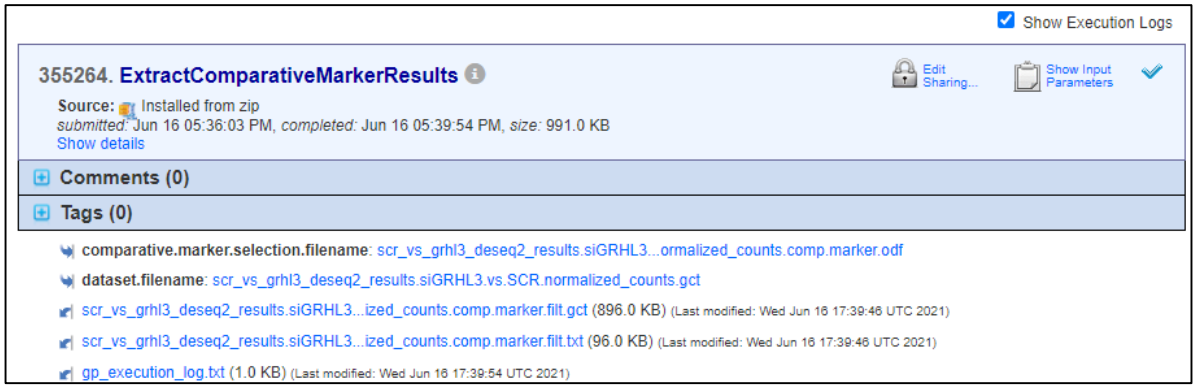


Figure 3.9 Extract Comparative Marker Selection results.

3.3 Grhl3 target genes identification.

After that, up to 200 out of 14055 candidate Grhl3 target genes from the DESeq2 results were analyzed using the Genome Browser by graphically viewing the sequences of the genes, their location, and annotations in correlation with Grhl3. 75% downregulation percentage was considered for measuring the correlation of gene expression with Grhl3. The criteria were followed to select the target genes from the list of the genes was:

- Grhl3 binding to the target gene
- Presence of TATAA box to determine the promoter region and so if the Grhl3 gene binds to it
- Presence of conserved regions for the target genes
- Binding of Grhl3 at the conserved regions of the target genes

As per the Genome Browser, selected genes were analyzed by entering the gene's name and then got the genome sequence from the mouse, NCBI RefSeq

genes, candidate cis-regulatory elements, mouse ESTs that have been spliced, multi alignments of 60 vertebrates, placental mammals' conservation and annotations.

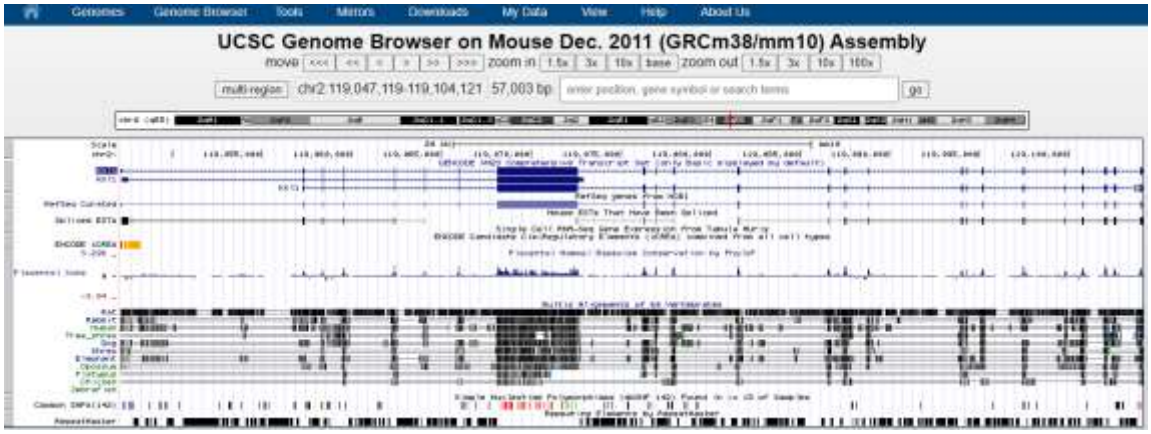


Figure 3.10 UCSC Genome browser on mouse Dec. 2011 (GRCm38/mm10) Assembly.

After that, the strand of interest was chosen specifically to check the presence of Grhl3 binding sites in the relevant promoter. Then from the next page, there is sequence retrieval region options and sequence formatting options, as the following figure shows

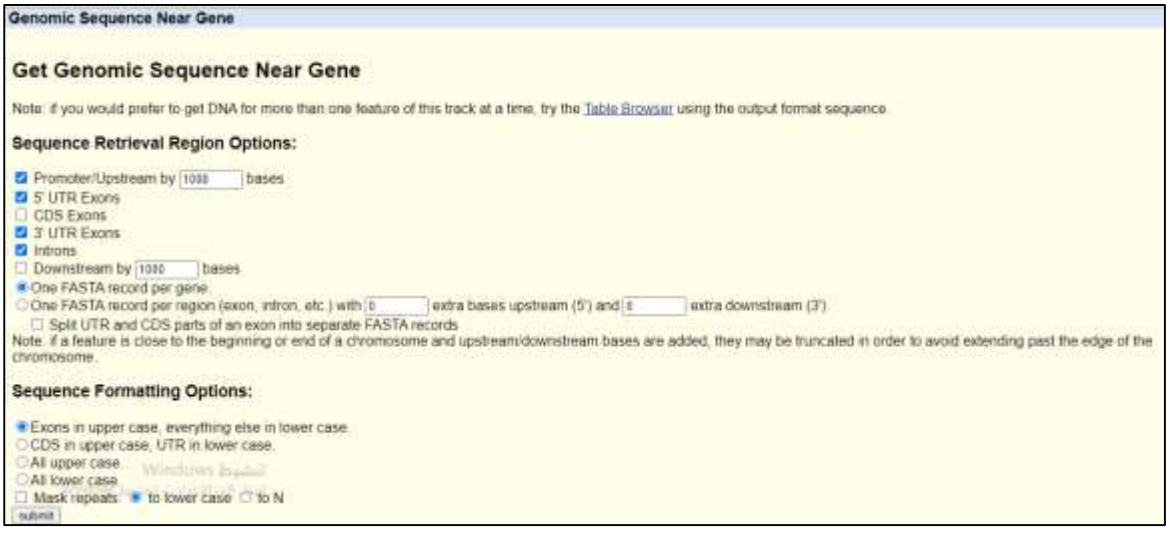


Figure 3.11 Genomic Sequence Near gene for sequence retrieving.

Following the selection of the options of the gene of interest, the sequence was retrieved and was used for further analysis.

```
>mm10_knownGene_ENSMUST00000028799.11 range=chr2:119046150-119072919 5'pad=0 3'pad=0 strand=+ repeatMasking=none
acaaatcccatagcttgcttaagggctggaagaggtggcatttgtaatgc
aatatggtacttcacgatcaggacaggtattaaacccccaaaatagcgat
tatgggggtatagctttgtaatcctagcatttgggaaggcgagacaggat
gatttttaacctggaggctagactgctacacagtaaaactgtgtctcaaa
ataacaaaaacaaaacaaagccctcggagtaagagctcattaaagactga
aagcattcatctcagagctaggacatttcatgttccagaccaccactgaa
catgtcctgttatttttactttgtaagaaattaaattcctagaaaaaagtc
tctttatctgctttttcttcatctgtgagtcataacattccaaaagcagc
ctttcaaatatctcaattccttcagctgatttcaaatatcaaattccttc
aaattacaatatcaccaactaaaacacctcgatgctgagactaattgtgt
tagatagaacctgatcggaattaaaacccgaagcaatatttaatcacttc
catcctgtggattataaattcacattacactaattagcaaaagacacttt
cgacgaggatgggattcgaacccatgcgtgcagagcacaatggattagca
gtccatcgcctttaaccactcggccacctcgtcccagaagctagtttcttc
ctggtagctatagggagtagacacagagcgctgtgacgcaataccatgga
agtcgccattttgtacggaagccaacagtgcgattgtttactctgcaaac
actcgaaagaagatggccgttctcagaataacagagacgaagaaaattct
gtagaaaagaagaaaaacccgagggaacttaagtcagagcagaggaggac
agagaagtttatgtcttaggggcggggccggtctggagccgaaggggtgg
ggtctgcgcggcgcgcagcactgagcatgcgcatcactagggcggttaga
GTATCTAAGCAATAGAAACTATTTCTACTTACTATAGGTTAGTGAAATAT
```

Figure 3.12 Retrieved sequence of a gene showing the chromosome range, the gene ID, and strand position.

3.4 CLC analysis and conserved sequences obtaining

At the CLC program, the sequences that were retrieved from the genome browser were used to look for the conserved sequences of the Grhl3 target genes in addition to other important sequences such as the TATAA box and Grhl3 binding site, the Grhl3 sequence was obtained from JASPER figure 3.13 and the conserved sequences from mVISTA tool, by simply entering the needed gene position as in figure 3.14 and then from the tool section, the CNS was chosen and acquired, where the conserved sequences were selected among the six vertebrate species in mVISTA figure 3.15.

Figure 3.13 Grhl3 binding site sequence profile from JASPER

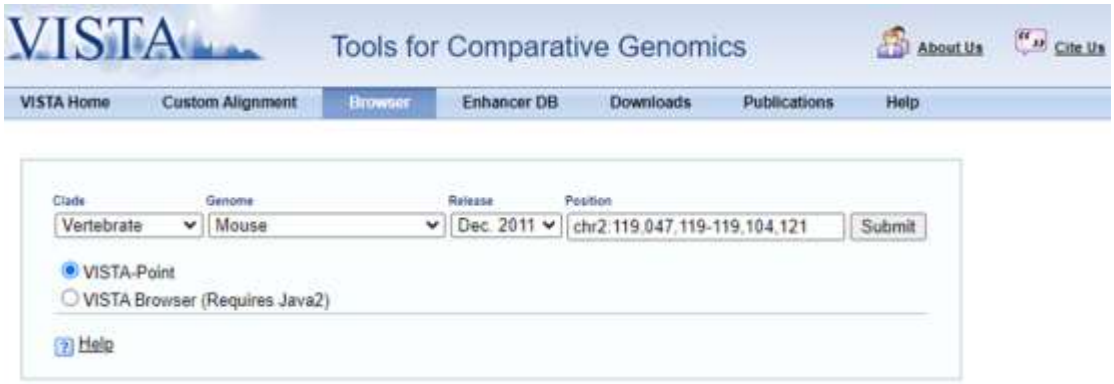


Figure 3.14 mVISTA tools for comparative genomics main page

Figure 3.15 the tool section from mVISTA where the conserved sequences are selected among the six vertebrate species.

Later at the CLC sequence viewer, target genes sequences data were entered, and from the search bar, the sequence of CNS was able to be determined whether they are found or not.

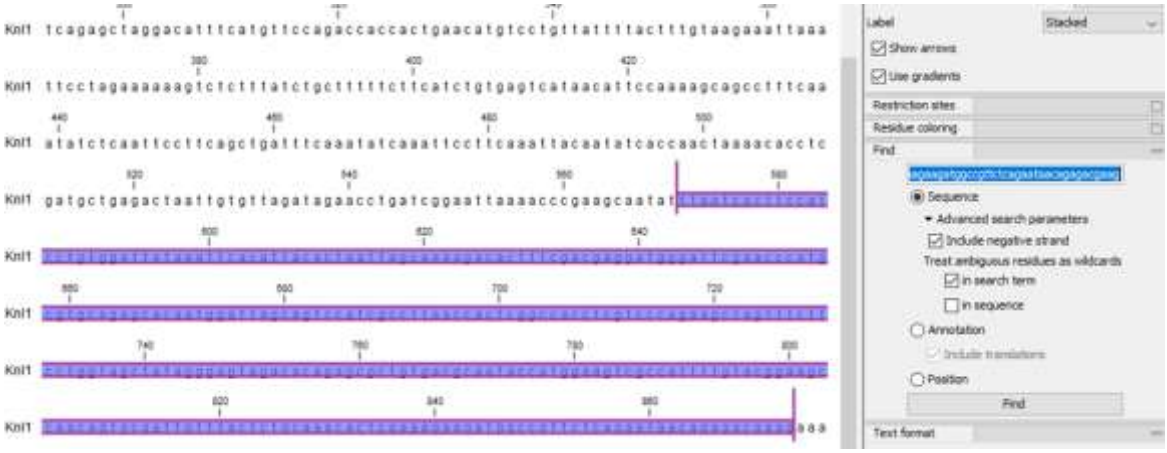


Figure 3.16 CLC sequence viewer.

For more emphasizing on the specificity of the Grhl3 target genes binding sites conservation, Grhl3 was looked if it is bound to the conserved sequences of the target genes or not; as a result, several genes were found to follow the criteria as follows: Nusap1, Hmgb2 in addition to other three genes Oip5, Cenpf, and Knl1. Later these sequences were used to design primers for wet-lab experiments for the Grhl3 RNA interference.

Nusap1

mouse agggaccgcg acaggtccca ggcgaggtgc

human AGGGACCGCG ACAGGTCCCA GGCGAGGTGC

chimp AGGGACCGCG ACAGGTCCCA GGCGAGGTGC

cow AGGGACCGCG ACAGGTCCCA GGCGAGGTGC

dog AGGGACCGCG ACAGGTCCCA GGCGAGGTGC

Rhesus AGGGACCGCG ACAGGTCCCA GGCGAGGTGC

Hmgb2

mouse ttcctctcta gaccggttcc aacgggtctg

human TTCCTCTCTA GACCGGTTCC AACGGGTCTG

chimp TTCCTCTCCT GACCGGTTCT AACGGGTCTG

dog TTCCTCTCCT GACCGGTTCT AACGGGTCTG

cow TTCCTCTCCT GACCGGTTCT AACGGGTCTG

Figure 3.17 Grhl3 conserved binding sites among species at the transcription factors target genes of Nusap1 and Hmgb2.

3.5 Grhl3 target genes correlation of gene expression from RNAseq results.

The chart below shows the correlation of Grhl3 target genes when it was silenced from the results obtained from RNA-seq to be compared with the experimental results.

Where Cdh1was downregulated to~ 0.17, Hmgb2 to~ 0.22, Nusap1 to~ 0.18, Oip5 to~ 0.29, Knl1 to~ 0.18, and Cenpf to zero

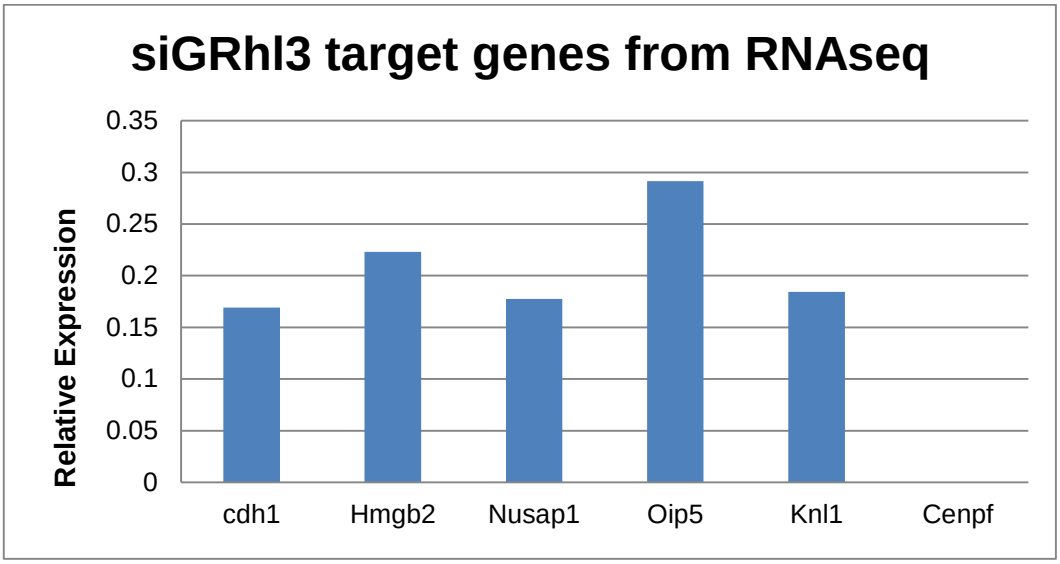


Figure 3.18 Relative expressions of Grhl3 target genes when Grhl3 was silenced from RNAseq results showing the fold change of gene expression of the target genes where Cenpf was dropped significantly to 0.

3.6 (RNA interference): Targeting Grhl3 with siRNAs.

NMuMG cells were treated with TGFβ (5 ng/ml) for 72 hours to induce EMT, and then TGFβ withdrawal was followed to allow MET progression. Cells undergoing MET treated with both control siRNA and siRNA targeting Grhl3. Seventy-two hours later, cells were collected and RNA was isolated. First, the EMT and MET transitions were validated by analyzing the expression of Cdh1 by qPCR. qPCR was also used to validate the knockdown efficiency using primers for Grhl3. For Grhl3 target genes, validation primers of Cenpf, Knl1, Oip5, Hmgb2, Nusap1, was used to measure the gene expression levels. As shown in the figure below, when Grhl3 was knocked down, Cdh1 was also reduced significantly.

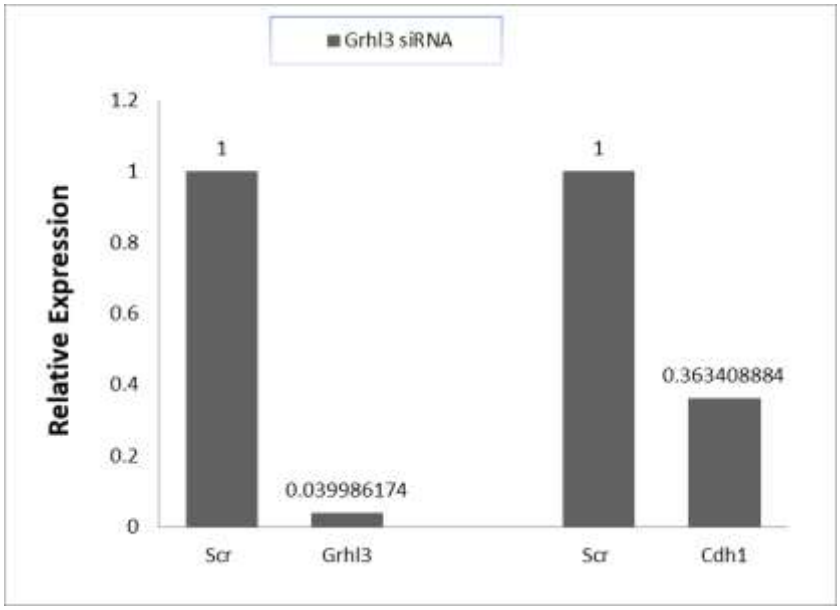


Figure 3.19 Relative expressions of Grhl3 and Cdh1 when Grhl3 was knocked down. The expression of Grhl3 was shown to drop when Grhl3 was silenced. The expression of E-cad (Cdh1) was also shown to be dropped significantly when Grhl3 was silenced.

To evaluate the phenotypical changes in the cells in contrast to epithelial NMuMG, microscopy images were taken after siRNA transfection at the 72nd-hour figure 3.20 and figure 3.21.

`

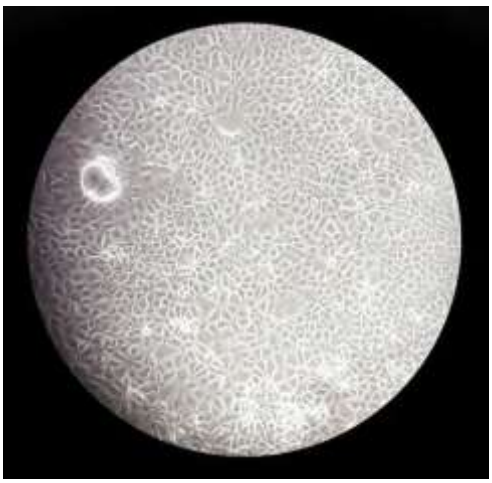


Figure 3.20 epithelial NMuMG cells.

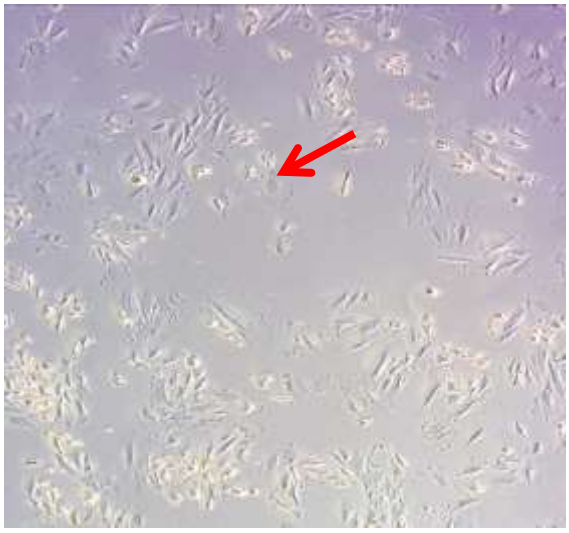


Figure 3.21 NMuMG cells in MET/Grhl3 was silenced

Cells were trapped in the mesenchymal state even after the removal of TGFβ, which shows that cells were unable to revert to the epithelial state with the Grhl3 knockdown in addition to the disruption of the cells where they appeared in a squishy look as the figure shows, the exact reason for this still unknown.

3.7 Grhl3 target genes correlation of gene expression

After getting the results from the RNAseq of the bioinformatics analysis, and choosing the best Grhl3 target genes that followed the criteria, Hmgb2, Nusap1, Oip5, Cenpf, and Knl1 primers were used for the qPCR analysis when Grhl3 has silenced to measure the correlation of gene expression with Grhl3 and compare to the results from RNAseq.

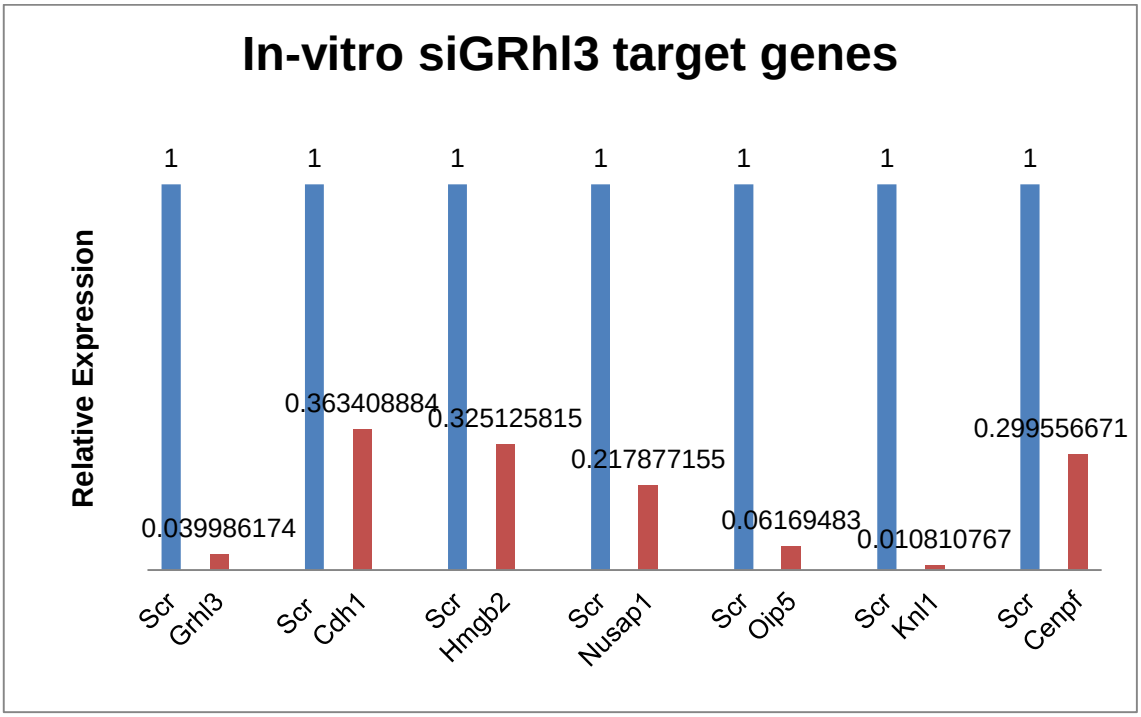


Figure 3.22 Grhl3 knockdown results utilizing primers for Grhl3 target genes. It demonstrates that relative gene expression control (Scr) equals 1 when compared to target gene expression, and since target gene values were less than 1, it means that it is all downregulated. Cdh1 was reduced to 0.36, Hmgb2 to 0.33, Nusap1 to 0.22, Oip5 to 0.06, Knl1 was significantly reduced to 0.01, and Cenpf was reduced to 0.30.

The chart above shows the results of Grhl3 knockdown using primers for Grhl3 target genes. It shows the relative gene expression control (Scr) equals 1 compared to the gene expression of target genes, and since values of target genes were less than 1, which means it's all downregulated. siGrhl3 expression was lowered to approximately 0.04, Cdh1 to~ 0.36, Hmgb2 to ~ 0.33, Nusap1 to~ 0.22, Oip5 to~ 0.06, Knl1segnificantly to~ 0.01, and Cenpf to~ 0.30.

Chapter Four
Discussion

Chapter Four

Discussion

Epithelial-mesenchymal transition (EMT) and its opposite interaction mesenchymal-epithelial transition (MET) are basic developmentally monitored components utilized at various phases of morphogenesis and organogenesis to produce the body plan of metazoans (Pei, Liu, et al 2019). EMT includes the cumulative loss of epithelial cells polarity, downregulation of junctional components, and redesign of the actin cytoskeleton to present the transitory aggregate important to execute the overall cell developments that will bring about the germ layers and later to tissues (Nieto, Zhao, et al 2016). MET is utilized to create epithelia at various formative stages. During MET, mesenchymal cells sequentially build up apicobasal polarity through developmentally monitored clusters of proteins and definite yet distinct mechanisms (Boulan, Macara, et al 2014). MET is a crucial process in histogenesis to gather mesenchymal-like cells into coherent units. Without MET, embryos can't involve in gastrulation or the subsequent body plan development. Besides the foundation of epithelial polarity, MET in reprogramming is involved with metabolic exchanging, epigenetic adjustments, and cellular fate changes, giving a remarkable chance to explore the connections between these cell interactions (Wu, Ocampo, & Belmonte 2016).

The EMT procedure is controlled by the elements known as EMT inducers that suppress the E-cadherin activity like Snai1/Snail, Snai2/Slug, Twist1/2,

Zeb, and KLF8 (Thuault et al., 2008; Lin and Baritaki, 2010; Lo. H.W., 2007; Jung et al., 2008; Lamouille, Xu and Derynck, 2014; Peinado et al., 2007; Wang et al., 2007), regardless of whether by cellular alteration and variation of gene expression straightforwardly or by implication by the transcription inactivation (Yang and Weinberg, 2008; Sobrado et al., 2009). These factors are liable for downregulating the E-cad expression quickly in EMT and upregulating it back again promptly during MET (Voutsadakis, 2016). Although MET can resemble the converse of EMT, it is controlled by more integrated and dynamic mechanisms. Even though it isn't completely understood how these mechanisms work during MET, the EMT inducers may take part in initiating E-cad expression just to an obscure degree at this point (Kim et al.; 2017).

Since Grhl3, in addition to other transcription factors, are critical for MET progression, cells could not initiate MET when Grhl3 was silenced; that's because they failed to upregulate the E-cad, where Grhl3 appeared as a principal requirement for MET process initiation and progression. In our lab, our primer data recommends the presence of a different integrated transcriptional node having Grhl3 and its target genes in the focal point of a larger network and E-cadherin as the significant result. Our primary objective is to affirm the transcriptional relations between the determined transcription factors and draw a more inclusive image of the MET network. Accordingly, our main aim is to recognize a major MET network, which, eventually, may lead to identifying therapeutic targets for the control of MET during

metastasis. MET is additionally a basic process during the reprogramming of iPSCs, hence recognizing a core MET network could be used to work on the effectiveness and quality of the reprogramming procedure.

After knowing the significant roles that Grhl3 plays in the cell, identifying its target genes during MET will provide insight into understanding how MET is regulated and may create possibilities for developing novel therapeutics.

In this study, different bioinformatics analyses were used to identify Grhl3 target genes, GO enrichment analysis for the annotations, Genepattern for differential gene expression, and extracting the RNAseq, in addition to producing a heat map of the results. Finally, Grhl3 knockdown experiments to identify Grhl3 targets during MET followed by qPCR analysis using primers for Grhl3, Cdh1, Cenpf, Hmgb2, Nusap1, Oip5, and Knl1. The last five target genes were selected from the RNAseq results from Genepattern analysis since they met the criteria set for the best consideration of Grhl3 target genes.

Hmgb2 (High Mobility Group Box 2) encodes an individual from the non-histone chromosomal high-portability proteins family (Murugesapillai, McCauley 2016). The proteins of this family are chromatin-related and universally circulated in the nucleus of higher eukaryotic cells. In vitro examines have exhibited that this protein can effectively twist DNA and leads to the formation of DNA circles. This protein was likewise being found to be engaged with the last ligation step in DNA end-joining mechanisms of DNA double-strand break repairing and V(D)J recombination (Paull, Haykinson

1993). Hmgb2 interfaces with the minor section of DNA and twists the DNA portion to which it is bound. It has no arrangement particularity except for can be enlisted through protein-protein associations by an assortment of sequence-precise DNA restricting proteins to give a 'structural' action, it works with or balances out the congregation of multiprotein combinations on DNA, both by forcing the right geometrical way onto the double-stranded helix and by securing other DNA-restricting proteins (Ronfani, Ferraguti 2001).

Nusap1 (Nucleolar and Spindle Associated Protein 1) is a gene encoding a 55 kD protein, a cell-cycle-subordinate protein expressed during mitosis at high levels. It has microtubule-restricting and DNA-restricting action, which underlies its best-portrayed capacity as a controller of the mitotic mechanical assembly in partitioning cells (Raemaekers, Ribbeck, et al 2003). The downregulation of NUSAP1 is related to misassembly or misfunction of the mitotic spindle fibers during mitosis, causing a genetic imbalance in progeny cells. Serves in cell relocation, intrusion, and metastases are likewise found and utilitarian interactivity with pro-oncogenic pathways. NUSAP1 is overexpressed in a few disease types and is viewed as a novel prognostic biomarker. When Nusap1 is experimentally downregulated, it regularly restrained cell multiplication in a few disease settings. NUSAP1 is likewise proposed as a remedial objective, with the possibility to work on the result of therapies in mixed treatment for certain malignancy types (Damizia, Lavia, 2020).

Moreover, Overexpression of NUSAP1 caused microtubule folding and cell cycle capture at the G2/M checkpoint (Vanden, Raemaekers, et al 2010). As a result of DNA disruption, NUSAP1 is phosphorylated by ATM/ATR kinases initiating a mitotic arrest (Matsuoka, Ballif, et al 2007 and Xie, Li, et al 2011). Also, NUSAP1 has been demonstrated to be reduced because of UV damage (Emanuele, Elia, et al 2011).

Cenpf (The centromere protein F) is located on chromosome 1q41 and encodes for CENPF protein, and it is engaged in chromosome dissociation during cell division. It also plays a role in the direction of microtubules to form cellular cilia (Waters A, et al 2015). Cenpf functions as a part of the centromere kinetochore components and a part of the nuclear set during G2 of interphase (Bomont, Maddox, et al 2005). CENPF expression is in a cell-cycle-subordinate way, steadily escalated during the interaction of the cell cycle, then, at that point arrived at top levels in the G2/M stage, at last Cenpf will be decreased at the end of mitosis (Shahid, Lee, et al 2018). The protein level is low in the G1 stage but rises intensely in the S stage as a nuclear protein. After the continuation of the M stage, Cenpf is hyperphosphorylated and shows a localization at the kinetochores, spindle poles, and midbody (YANG, GUO, et al 2003). Expression changes of CENPF have been displayed to correspond with the progression of different diseases. Practically talking, cell multiplication and apoptosis are known to be influenced by CENPF (Du, Lu, et al 2018).

The results from the Differential Expression workflow for RNA-Seq data for

14055 target genes were analyzed and obtained using Gene pattern program; it revealed which genes were up or down-regulated, it also showed with the agreement with previous studies the downregulation of central genes such as Hnf4α with the Grhl3 knocking down (where it was downregulated significantly from 22 in the control to 4.7 when Grhl3 was silenced) (Sengez, Aygün, et al 2019).

From the comparative marker selection module, the resultant RNA-seq Grhl3 target genes were shown graphically as a heatmap in addition to their P-value and other filters for better viewing and confirmation of the differentially expressed target genes.

Changes in RNA expression levels define transcriptional relationships due to gene silencing, yet this was needed to be validated by looking for transcription factor binding sites in the promoters of target genes or by identifying enhancers involved in the transcriptional regulation. First of all, transcription factor binding sites were searched within the promoters of target genes from RNA-seq results, 200 target genes were looked at the genome browser for the identification of the regulatory elements and binding site of Grhl3, with the help of the CLC sequence viewer, Grhl3 binding was detected if it's at the promoter region taking 1000 bp upstream promoter region, in addition to TATAA box.

For the Grhl3 targets with no relevant binding site in the promoter region, prediction analyses were performed to search for enhancers. It is rationalized that a conserved expression pattern reflects a conserved regulation through

conserved enhancers; thus, analysis of evolutionarily conserved sequences in non-coding sequences might reveal the existence of enhancers. This methodology has previously used and identified enhancers within the introns of the sodium iodide symporter (NIS) and in the second intron of E-cadherin (Alotaibi et al 2010, and Alotaibi et al, 2015). The conserved sequences were obtained from the mVista, chosen among the six vertebrate species in mVista. This bioinformatics tool is used because of the updated provisions, multifaceted utilization, and detailed visualization of the aligned conserved sequences of different species. Several Grhl3 target genes were obtained following that criteria efficiently, including Hmgb2 and Nusap1. Later they were used to design primers for Grhl3 interference for measuring the correlation of gene expression of target genes.

The Grhl3 knockdown in-vitro experiment results during MET showed that with Grhl3 silencing, Cdh1 expression was also downregulated to (0.4). Cells in the microscopy images appeared disrupted and trapped in the mesenchymal state, which means cells could not initiate MET and revert to the epithelial state in agreement with the previous studies and literature review. Also, it supports the hypothesis for the role of Grhl3 during MET and the importance of identifying its target genes.

As per for comparing the siGrhl3 bioinformatics results from RNA-seq with experimental results, Grhl3 target genes showed a correlation of gene expression, where the 5 target genes were downregulated compared to the control (Scr) with fold change equals 1. The difference of the correlation of

gene expression between the RNA-seq and experimental results was significant for the Cdh1 with a difference of 0.19, 0.11 for the Hmgb2, and 0.04 for Nusap1, whereas the correlation of Cenpf wasn't compatible; thus, in the RNA-seq is zero but 0.30 experimentally. Knl1 showed the most significant downregulation to 0.011, with a difference of 0.20 from bioinformatics results. Knl1 has two essential functions during mitosis: it is crucial for spindle-assembly checkpoint signaling and correct chromosome alignment. Directly links spindle checkpoint proteins to kinetochores. Also, it's a Part of the MIS12 complex, which may be fundamental for kinetochore formation and proper chromosome segregation during mitosis. Also, Knl1 acts in coordination with CENPF to recruit a proteins complex to the outer kinetochore (Visel, Thaller, Eichele 2004 and Shi, Qalieh, et al 2019). Overall, these results indicate that the two target genes not only correlated with Grhl3 expression but also with each other.

Chapter Five
Conclusion and Suggestions

Chapter Five

Conclusion and Suggestions

Grhl3 is one of the fundamental transcription factors that direct the MET process. Its relationship or association with the other EMT-MET-related factors and identifying its target genes will reveal insight into the MET mechanism.

Grhl3 has many target genes that are differentially correlated with its gene expression; most of them have a common role of being involved in cell cycle, chromosome segregation, and mitosis when their annotations are searched at Gene Ontology.

Hmgb2, Nusap1, Cenpf, and Knl1 can be good candidates as Grhl3 target genes, which should be studied further.

References

- 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 (BBA)-Gene Regulatory Mechanisms, 1849(6), 731-742.
- Alotaibi, H., Yaman, E., Salvatore, D., Di Dato, V., Telkoparan, P., Di Lauro, R., & Tazebay, U. H. (2010). Intronic elements in the Na+/I-symporter gene (NIS) interact with retinoic acid receptors and mediate the initiation of transcription. Nucleic acids research, 38(10), 3172-3185.
- Banyard, J., & Bielenberg, D. R. (2015). The role of EMT and MET in cancer dissemination. Connective tissue research, 56(5), 403-413.
- Batlle, E., Sancho, E., Francí, C., Domínguez, D., Monfar, M., Baulida, J., & De Herreros, A. G. (2000). The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. Nature cell biology, 2(2), 84-89.
- Bolós, V., Peinado, H., Pérez-Moreno, M. A., Fraga, M. F., Esteller, M., & Cano, A. (2003). The transcription factor Slug represses E-cadherin expression and induces epithelial to mesenchymal transitions: a comparison with Snail and E47 repressors. Journal of cell science, 116(3), 499-511.

- Bomont, P., Maddox, P., Shah, J. V., Desai, A. B., & Cleveland, D. W. (2005). Unstable microtubule capture at kinetochores depleted of the centromere-associated protein CENP-F. The EMBO journal, 24(22), 3927-3939.
- Brabletz, T. (2012). EMT and MET in metastasis: where are the cancer stem cells?. Cancer cell, 22(6), 699-701.
- Brabletz, T. (2012). MiR-34 and SNAIL: another double-negative feedback loop controlling cellular plasticity/EMT governed by p53. Cell Cycle, 11(2), 215-215.
- Caddy, J., Wilanowski, T., Darido, C., Dworkin, S., Ting, S. B., Zhao, Q., ... & Jane, S. M. (2010). Epidermal wound repair is regulated by the planar cell polarity signaling pathway. Developmental cell, 19(1), 138-147.
- Cai, Y. R., Zhang, H. Q., Qu, Y., Mu, J., Zhao, D., Zhou, L. J., Yan, H., Ye, J. W., & Liu, Y. (2011). Expression of MET and SOX2 genes in non-small cell lung carcinoma with EGFR mutation. Oncology reports, 26(4), 877–885. https://doi.org/10.3892/or.2011.1349.
- Cano, A., Pérez-Moreno, M. A., Rodrigo, I., Locascio, A., Blanco, M. J., del Barrio, M. G., ... & Nieto, M. A. (2000). The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. Nature cell biology, 2(2), 76-83.
- Carver, E. A., Jiang, R., Lan, Y., Oram, K. F., & Gridley, T. (2001). The mouse snail gene encodes a key regulator of the epithelial-mesenchymal

transition. Molecular and cellular biology, 21(23), 8184-8188.

- Chen, H. F., Ma, R. R., He, J. Y., Zhang, H., Liu, X. L., Guo, X. Y., & Gao, P. (2017). Protocadherin 7 inhibits cell migration and invasion through E-cadherin in gastric cancer. Tumor Biology, 39(4), 1010428317697551.
- Comijn, J., Berx, G., Vermassen, P., Verschueren, K., van Grunsven, L., Bruyneel, E., & Van Roy, F. (2001). The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. Molecular cell, 7(6), 1267-1278.
- Damizia, M., Lavia, P., (2020) "NUSAP1 (nucleolar and spindle associated protein 1)". 15:48:01 CEST 2021.
- Dworkin, S., Darido, C., Georgy, S. R., Wilanowski, T., Srivastava, S., Ellett, F., ... & Jane, S. M. (2012). Midbrain-hindbrain boundary patterning and morphogenesis are regulated by diverse grainy head-like 2-dependent pathways. Development, 139(3), 525-536.
- Eger, A., Aigner, K., Sonderegger, S., Dampier, B., Oehler, S., Schreiber, M., ... & Foisner, R. (2005). DeltaEF1 is a transcriptional repressor of E-cadherin and regulates epithelial plasticity in breast cancer cells. Oncogene, 24(14), 2375-2385.
- Emanuele, M. J., Elia, A. E., Xu, Q., Thoma, C. R., Izhar, L., Leng, Y., ... & Elledge, S. J. (2011). Global identification of modular cullin-RING ligase substrates. Cell, 147(2), 459-474.

- Frisch, S. M., Farris, J. C., & Pifer, P. M. (2017). Roles of Grainyhead-like transcription factors in cancer. Oncogene, 36(44), 6067-6073.
- Gilbert, R. W., Vickaryous, M. K., & Viloria-Petit, A. M. (2016). Signaling by transforming growth factor-beta isoforms in wound healing and tissue regeneration. Journal of developmental biology, 4(2), 21.
- Gordon, W. M., Zeller, M. D., Klein, R. H., Swindell, W. R., Ho, H., Espetia, F., ... & Andersen, B. (2014). A GRHL3-regulated repair pathway suppresses immune-mediated epidermal hyperplasia. The Journal of clinical investigation, 124(12), 5205-5218.
- Graziano, F., Humar, B., & Guilford, P. (2003). The role of the E-cadherin gene (CDH1) in diffuse gastric cancer susceptibility: from the laboratory to clinical practice. Annals of oncology, 14(12), 1705-1713.
- Grünert, S., Jechlinger, M., & Beug, H. (2003). Diverse cellular and molecular mechanisms contribute to epithelial plasticity and metastasis. Nature reviews Molecular cell biology, 4(8), 657-665.
- Handschuh, G., Candidus, S., Luber, B., Reich, U., Schott, C., Oswald, S., & Becker, K. F. (1999). Tumor-associated E-cadherin mutations alter cellular morphology, decrease cellular adhesion and increase cellular motility. Oncogene, 18(30), 4301-4312.
- Hay, E. D. (1995). An overview of epithelial-mesenchymal transformation. Cells Tissues Organs, 154(1), 8-20.
- Hislop, N. R., Caddy, J., Ting, S. B., Auden, A., Vasudevan, S., King, S.

L., ... & Jane, S. M. (2008). Grhl3 and Lmo4 play coordinate roles in epidermal migration. Developmental biology, 321(1), 263-272.

- Huber, M. A., Kraut, N., & Beug, H. (2005). Molecular requirements for epithelial-mesenchymal transition during tumor progression. Current opinion in cell biology, 17(5), 548-558.
- Hulit, J., Suyama, K., Chung, S., Keren, R., Agiostratidou, G., Shan, W., ... & Hazan, R. B. (2007). N-cadherin signaling potentiates mammary tumor metastasis via enhanced extracellular signal-regulated kinase activation. Cancer Research, 67(7), 3106-3116.
- Illman, S. A., Lehti, K., Keski-Oja, J., & Lohi, J. (2006). Epilysin (MMP-28) induces TGF-β mediated epithelial to mesenchymal transition in lung carcinoma cells. Journal of cell science, 119(18), 3856-3865.
- Jennbacken, K., Tešan, T., Wang, W., Gustavsson, H., Damber, J. E., & Welén, K. (2010). N-cadherin increases after androgen deprivation and is associated with metastasis in prostate cancer. Endocrine-related cancer, 17(2), 469-479.
- Jung H. Lee K. Park S. Park J. et. al. TMPRSS4 promotes invasion, migration, and metastasis
- of human tumor cells by facilitating an epithelial-mesenchymal transition, Oncogene 2008;
- vol: 27 (18) pp: 2635-2647
- Kalluri, R., & Weinberg, R. A. (2009). The basics of epithelial-

mesenchymal transition. The Journal of clinical investigation, 119(6), 1420-1428.

- Kim, H. Y., Jackson, T. R., & Davidson, L. A. (2017, July). On the role of mechanics in driving mesenchymal-to-epithelial transitions. In Seminars in cell & developmental biology (Vol. 67, pp. 113-122). Academic Press.
- Kim, I.-H., and Nam, T.-J. Enzyme-treated Ecklonia cava extract inhibits adipogenesis through the downregulation of C/EBPα in 3T3-L1 adipocytes. International Journal of Molecular Medicine, 2017;39(3), 636–644.
- Klein, R. H., Lin, Z., Hopkin, A. S., Gordon, W., Tsoi, L. C., Liang, Y., ... & Andersen, B. (2017). GRHL3 binding and enhancers rearrange as epidermal keratinocytes transition between functional states. PLoS genetics, 13(4), e1006745.
- Lamouille, S., Subramanyam, D., Blelloch, R., & Derynck, R. (2013). Regulation of epithelial-mesenchymal and mesenchymal-epithelial transitions by microRNAs. Current opinion in cell biology, 25(2), 200-207.
- Lamouille, S., Xu, J., & Derynck, R. (2014). Molecular mechanisms of epithelial-mesenchymal transition. Nature reviews Molecular cell biology, 15(3), 178-196.
- Lazarevich, N. L., Shavochkina, D. A., Fleishman, D. I., Kustova, I. F.,

Morozova, O. V., Chuchuev, E. S., & Patyutko, Y. I. (2010). Deregulation of hepatocyte nuclear factor 4 (HNF4) as a marker of epithelial tumors progression. Exp Oncol, 32(3), 167-71.

- 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.
- Lin, K., Baritaki, S., Militello, L., Malaponte, G., Bevelacqua, Y., & Bonavida, B. (2010). The Role of B-RAF Mutations in Melanoma and the Induction of EMT via Dysregulation of the NF-κB/Snail/RKIP/PTEN Circuit. Genes & cancer, 1(5), 409-420.
- Lo, H. W., Hsu, S. C., Xia, W., Cao, X., Shih, J. Y., Wei, Y., ... & Hung, M. C. (2007). Epidermal growth factor receptor cooperates with signal transducer and activator of transcription 3 to induce epithelial-mesenchymal transition in cancer cells via up-regulation of TWIST gene expression. Cancer Research, 67(19), 9066-9076.
- Loh, C. Y., Chai, J. Y., Tang, T. F., Wong, W. F., Sethi, G., Shanmugam, M. K., ... & Looi, C. Y. (2019). The E-cadherin and N-cadherin switch in epithelial-to-mesenchymal transition: signaling, therapeutic implications, and challenges. Cells, 8(10), 1118.
- Lu, Y., Jin, X., Chen, Y., Li, S., Yuan, Y., Mai, G., & Cheng, J. (2010). Mesenchymal stem cells protect islets from hypoxia/reoxygenation-induced injury. Cell biochemistry and function, 28(8), 637-643.

- Luck, K., Kim, D. K., Lambourne, L., Spirohn, K., Begg, B. E., Bian, W., Brignall, R., Cafarelli, T., Campos-Laborie, F. J., Charloteaux, B., Choi, D., Coté, A. G., Daley, M., Deimling, S., Desbuleux, A., Dricot, A., Gebbia, M., Hardy, M. F., Kishore, N., Knapp, J. J., … Calderwood, M. A. (2020). A reference map of the human binary protein interactome. Nature, 580(7803), 402–408. https://doi.org/10.1038/s41586-020-2188-x.
- Matsuoka, S., Ballif, B. A., Smogorzewska, A., McDonald, E. R., Hurov, K. E., Luo, J., ... & Elledge, S. J. (2007). ATM and ATR substrate analysis reveal extensive protein networks responsive to DNA damage. science, 316(5828), 1160-1166.
- Mayer, B., Jauch, K. W., Schildberg, F. W., Funke, I., Günthert, U., Figdor, C. G., & Johnson, J. P. (1993). De-novo expression of CD44 and survival in gastric cancer. The Lancet, 342(8878), 1019-1022.
- Miettinen, P. J., Ebner, R., Lopez, A. R., & Derynck, R. (1994). TGF-beta induced transdifferentiation of mammary epithelial cells to mesenchymal cells: involvement of type I receptors. The Journal of cell biology, 127(6), 2021-2036.
- Mrozik, K. M., Blaschuk, O. W., Cheong, C. M., Zannettino, A. C. W., & Vandyke, K. (2018). N-cadherin in cancer metastasis, its emerging role in hematological malignancies, and potential as a therapeutic target in cancer. BMC Cancer, 18(1), 1-16.

- Murugesapillai, D., McCauley, M. J., Maher, L. J., & Williams, M. C. (2017). Single-molecule studies of high-mobility group B architectural DNA binding proteins. Biophysical Reviews, 9(1), 17-40.
- Nieto, A., Zhao, J. M., Han, Y. H., Hwang, K. H., & Schoenung, J. M. (2016). Microscale tribological behavior and in vitro biocompatibility of graphene nanoplatelet reinforced alumina. Journal of the mechanical behavior of biomedical materials, 61, 122-134.
- Paull, T. T., Haykinson, M. J., & Johnson, R. C. (1993). The nonspecific DNA-binding and-bending proteins HMG1 and HMG2 promote the assembly of complex nucleoprotein structures. Genes & development, 7(8), 1521-1534.
- Pei, Y. F., Liu, J., Cheng, J., Wu, W. D., & Liu, X. Q. (2019). Silencing of LAMC2 reverses epithelial-mesenchymal transition and inhibits angiogenesis in cholangiocarcinoma via inactivation of the epidermal growth factor receptor signaling pathway. The American journal of pathology, 189(8), 1637-1653.
- Pieters, T., & Van Roy, F. (2014). Role of cell-cell adhesion complexes in embryonic stem cell biology. Journal of cell science, 127(12), 2603-2613.
- Radisky, D. C. (2005). Epithelial-mesenchymal transition. *Journal of cell science*, *118*(19), 4325-4326.
- Peinado H. Olmeda D. Cano A, Snail, Zeb and bHLH factors in tumor

progression: an alliance against the epithelial phenotype. Nature Reviews Cancer 2007;vol: 7 (6) pp: 415-428

- Redmer, T., Diecke, S., Grigoryan, T., Quiroga-Negreira, A., Birchmeier, W., & Besser, D. (2011). E-cadherin is crucial for embryonic stem cell pluripotency and can replace OCT4 during somatic cell reprogramming. EMBO reports 12(7), 720-726.
- Raemaekers, T., Ribbeck, K., Beaudouin, J., Annaert, W., Van Camp, M., Stockmans, I., ... & Carmeliet, G. (2003). NuSAP, a novel microtubule-associated protein involved in mitotic spindle organization. The Journal of cell biology, 162(6), 1017-1029.
- Rodriguez-Boulan, E., & Macara, I. G. (2014). Organization and execution of the epithelial polarity program. Nature reviews Molecular cell biology, 15(4), 225-242.
- Ronfani, L., Ferraguti, M., Croci, L., Ovitt, C. E., Scholer, H. R., Consalez, G. G., & Bianchi, M. E. (2001). Reduced fertility and spermatogenesis defects in mice lacking chromosomal protein Hmgb2. Development, 128(8), 1265-1273.
- Santangelo, L., Marchetti, A., Cicchini, C., Conigliaro, A., Conti, B., Mancone, C., ... & Tripodi, M. (2011). The stable repression of the mesenchymal program is required for hepatocyte identity: a novel role for hepatocyte nuclear factor 4α. Hepatology, 53(6), 2063-2074.
- Sengez, B., Aygün, I., Shehwana, H., Toyran, N., Tercan Avci, S., Konu,

O., ... & Alotaibi, H. (2019). The transcription factor Elf3 is essential for a successful mesenchymal to epithelial transition. Cells, 8(8), 858.

- Shahid, M., Lee, M. Y., Piplani, H., Andres, A. M., Zhou, B., Yeon, A., ... & Kim, J. (2018). Centromere protein F (CENPF), a microtubule-binding protein, modulates cancer metabolism by regulating pyruvate kinase M2 phosphorylation signaling. Cell Cycle, 17(24), 2802-2818.
- Shi, L., Qalieh, A., Lam, M. M., Keil, J. M., & Kwan, K. Y. (2019). Robust elimination of genome-damaged cells safeguards against brain somatic aneuploidy following Knl1 deletion. Nature communications, 10(1), 1-14.
- Sobrado V. Moreno-Bueno G. Cubillo E. Holt L. et. al., The class I bHLH factors E2-2A and E2-2B regulate EMT, Journal of Cell Science 2009; vol: 122 (7) pp: 1014-1024
- Stemmler, M. P. (2008). Cadherins in development and cancer. Molecular bioSystems, 4(8), 835-850.
- 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: an official publication of the American Association of Anatomists, 227(2), 238-245.
- Stemmler, M. P., Hecht, A., & Kemler, R. (2005). E-cadherin intron 2 contains cis-regulatory elements essential for gene expression.
- Tan, E. J., Olsson, A. K., & Moustakas, A. (2015). Reprogramming

during epithelial to mesenchymal transition under the control of TGFβ. Cell adhesion & migration, 9(3), 233-246.

- Thiery, J. P., Acloque, H., Huang, R. Y., & Nieto, M. A. (2009). Epithelial-mesenchymal transitions in development and disease. cell, 139(5), 871-890.
- Thiery, J. P., & Sleeman, J. P. (2006). Complex networks orchestrate epithelial-mesenchymal transitions. *Nature reviews Molecular cell biology*, *7*(2), 131-142.
- Thuault, S., Tan, E. J., Peinado, H., Cano, A., Heldin, C. H., & Moustakas, A. (2008). HMGA2 and Smads co-regulate SNAIL1 expression during induction of epithelial-to-mesenchymal transition. Journal of Biological Chemistry, 283(48), 33437-33446.
- Vanden Bosch, A., Raemaekers, T., Denayer, S., Torrekens, S., Smets, N., Moermans, K., ... & Carmeliet, G. (2010). NuSAP is essential for chromatin-induced spindle formation during early embryogenesis. Journal of cell science, 123(19), 3244-3255.
- Vanhoutteghem, A., Maciejewski-Duval, A., Bouche, C., Delhomme, B., Hervé, F., Daubigney, F., & Djian, P. (2009). Basonuclin 2 has a function in the multiplication of embryonic craniofacial mesenchymal cells and is orthologous to disco proteins. Proceedings of the National Academy of Sciences, 106(34), 14432-14437.
- Vestweber, D., & Kemler, R. (1985). Identification of a putative cell

adhesion domain of uvomorulin. The EMBO journal, 4(13A), 3393-3398.

- Visel, A., Thaller, C., & Eichele, G. (2004). GenePaint. org: an atlas of gene expression patterns in the mouse embryo. Nucleic acids research, 32(suppl_1), D552-D556.
- Vleminckx, K., Vakaet Jr, L., Mareel, M., Fiers, W., & Van Roy, F. (1991). Genetic manipulation of E-cadherin expression by epithelial tumor cells reveals an invasion suppressor role. Cell, 66(1), 107-119.
- Voutsadakis, I. A. (2016). Epithelial-mesenchymal transition (EMT) and regulation of EMT factors by steroid nuclear receptors in breast cancer: a review and in silico investigation. Journal of clinical medicine, 5(1), 11.
- Wang, M., Ren, D., Guo, W., Huang, S., Wang, Z., Li, Q., ... & Peng, X. (2016). N-cadherin promotes epithelial-mesenchymal transition and cancer stem cell-like traits via ErbB signaling in prostate cancer cells. International journal of oncology, 48(2), 595-606.
- Wang X. Zheng M. Liu G. Xia W. et. al. Kruppel-Like Factor 8 Induces Epithelial to Mesenchymal Transition and Epithelial Cell Invasion, Cancer Research 2007 vol: 67 (15) pp: 7184-7193
- Waters, A. M., Asfahani, R., Carroll, P., Bicknell, L., Lescai, F., Bright, A., Chanudet, E., Brooks, A., Christou-Savina, S., Osman, G., Walsh, P., Bacchelli, C., Chapgier, A., Vernay, B., Bader, D. M., Deshpande, C., O' Sullivan, M., Ocaka, L., Stanescu, H., Stewart, H. S., … Beales, P. L. (2015). The kinetochore protein, CENPF, is mutated in human ciliopathy

and microcephaly phenotypes. Journal of medical genetics, 52(3), 147–156. https://doi.org/10.1136/jmedgenet-2014-102691.

- Wu, J., Ocampo, A., & Belmonte, J. C. I. (2016). Cellular metabolism and induced pluripotency. Cell, 166(6), 1371-1385.
- Xie, P., Li, L., Xing, G., Tian, C., Yin, Y., He, F., & Zhang, L. (2011). ATM-mediated NuSAP phosphorylation induces mitotic arrest. Biochemical and biophysical research communications, 404(1), 413-418.
- Yang, J., & Weinberg, R. A. (2008). Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. Developmental cell, 14(6), 818-829.
- Yang, Z. Y., Jing, G. U. O., Ning, L. I., Min, Q. I. A. N., Wang, S. N., & Zhu, X. L. (2003). Mitosin/CENP-F is a conserved kinetochore protein subjected to cytoplasmic dynein-mediated poleward transport. Cell research, 13(4), 275-283.
- Yang, J., Mani, S. A., Donaher, J. L., Ramaswamy, S., Itzykson, R. A., Come, C., ... & Weinberg, R. A. (2004). Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. cell, 117(7), 927-939.
- Yilmaz, M., & Christofori, G. (2009). EMT, the cytoskeleton, and cancer cell invasion. Cancer and Metastasis Reviews, 28(1), 15-33.
- Yu, Z., Lin, K. K., Bhandari, A., Spencer, J. A., Xu, X., Wang, N., ... & Andersen, B. (2006). The Grainyhead-like epithelial transactivator Get-

1/Grhl3 regulates epidermal terminal differentiation and interacts functionally with LMO4. Developmental biology, 299(1), 122-136.

- Yu, Z., Bhandari, A., Mannik, J., Pham, T., Xu, X., & Andersen, B. (2008). Grainy head-like factor Get1/Grhl3 regulates the formation of the epidermal leading edge during eyelid closure. Developmental biology, 319(1), 56-67.

جامعة النجاح الوطنية

كلية الدراسات العليـــا

تحديد جينات مستهدفة خاصة ب Grhl3 أثناء عملية تحول الخلايا من وسيطية إلى طلائية

إعداد

ديانا أحمد صالح زيد

إشراف

د. أشرف صوافطة

د. هاني العتيبي

قدمت هذه الأطروحة استكمالاً لمتطلبات الحصول على درجة الماجستير في العلوم الحياتية من كلية الدراسات العليا في جامعة النجاح الوطنية في نابلس، فلسطين.

2021

تحديد جينات مستهدفة خاصة ب Grhl3 أثناء عملية تحول الخلايا من وسيطية إلى طلائية

إعداد

ديانا أحمد صالح زيد

إشراف

د. أشرف صوافطة

د. هاني العتيبي

الملخص

تعتبر عملية تحول الخلايا من طلائية إلى وسيطية وتحول الخلايا من وسيطية إلى طلائية (MET) عمليات مهمة تحدث خلال عملية تكوين الأعضاء الجنينية وتشكلها وتتضمن تغييرات ديناميكية في هوية الخلية. يلعب الجين الرأس المحبب 3 (Grhl3) وجينات تنظيمية أخرى دورًا رئيسيًا في تنظيم التعبير عن الكادهيرين الطلائي (E-cad) أثناء عملية MET ، وهي سمة مميزة لهذه العملية. الهدف من هذا البحث هو تحديد أهداف النسخ Grhl3 خلال MET وتقييم أهميتها المحتملة.

لاختبار الفرضية التي تنص بأن شبكة النسخ التي تتحكم في استقرار الحالة الطلائية للخلايا تتحكم في التنظيم الديناميكي لـ MET وأن هذه الشبكة قد تلعب أيضًا دورًا في بدء MET ، تم إجراء تحليل لبيانات تسلسل الحمض النووي (RNA) التي تحتوي على 14055 جينًا معبرًا تمايزيا باستخدام برامج لتحليل تسلسل الحمض النووي (RNA) مثل DESeq2 لتحديد الجينات المستهدفة من قبل Grhl3 أثناء عملية MET باستخدام المعلوماتية البيولوجية، بعد ذلك ، باستخدام نهج فقدان الوظيفة ، حيث تم تقليل التعبير عن Grhl3 بواسطة siRNAs، تم تحليل وتقييم الجينات المستهدفة المفترضة الخاصة ب Grhl3 بواسطة qPCR.

أظهرت النتائج أن الخلايا لا تستطيع بدء عملية التحول من وسيطية الى طلائية MET عندما تم تعطيل Grhl3 هناك علاقة بين التعبير عن Grhl3 المعطلة والجينات المستهدفة المختارة في كل من نتائج RNA-Seq والنتائج التجريبية ، حيث تم تقليل تنظيمها جميعًا. كان Cenpf هو أكثر جين تم تقليل تنظيمَه مع تعطيل Grhl3 من RNA-Seq من ناحية أخرى، كان Knl1 هو أكثر جين تم تقليل تنظيمه تجريبًا .

تشير هذه النتائج إلى أن Grhl3 هو أحد عوامل النسخ الأساسية التي توجه عملية MET ، بالإضافة إلى أن Grhl3لديه العديد من الجينات المستهدفة التي ترتبط بشكل تفاضلي بتعبيرها الجيني وأن Hmgb2 و Nusap1 و Cenpf و Knl1 يمكن أن تكون جينات مرشحة جيدة كجينات مستهدفة لـGrhl3 ، والتي يجب إجراء مزيد من التحقيق فيها وإجراء المزيد من التجارب للتأكيد على النتائج .