Investigation of regulation of promoter proximal pausing in YAP driven tumorigenesis

A Thesis

submitted to

Indian Institute of Science Education and Research Pune in partial fulfilment of the requirements for the BS-MS Dual Degree Programme

By

Bhavesh Sudhakar Vasave

Indian Institute of Science Education and Research Pune

Dr. Homi Bhabha Road, Pashan, Pune 411008, INDIA.

March, 2025

Supervisor: Dr Madhura Kulkarni

Translational Research Lead and DBT-Ramalingaswami Fellow at

Centre for Translational Cancer Research (CTCR) a joint initiative by

Indian Institute of Science Education and Research (IISER) Pune

and Prashanti Cancer Care Mission (PCCM), Pune

From May 2024 to Mar 2025

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH PUNE

Certificate

This is to certify that this dissertation entitled “Investigation of regulation of promoter proximal pausing in YAP driven tumorigenesis” towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Bhavesh Vasave at Indian Institute of Science Education and Research under the supervision of Dr Madhura Kulkarni, CTCR, PCCM, Pune during the academic year 2024-2025.

Dr Madhura Kulkarni

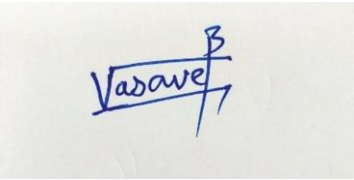
Committee:

Dr Madhura Kulkarni

Professor Kundan Sengupta

Declaration

I hereby declare that the matter embodied in the report entitled “Investigation of regulation of promoter proximal pausing in YAP driven tumorigenesis” are the results of the work carried out by me at the Department of Centre for translational cancer research, Indian Institute of Science Education & Research (IISER) Pune, under the supervision of Dr Madhura Kulkarni, and the same has not been submitted elsewhere for any other degree. Wherever others contribute, every effort is made to indicate this clearly, with due reference to the literature and acknowledgement of collaborative research and discussions.



Bhavesh Sudhakar Vasave

20201048

Table of Contents

List of Tables

List of Figures

Abstract

Cancer is a multifaceted and diverse disease marked by unregulated cell division and expansion. These cancerous cells cause genetic alterations and enables hitching of signalling pathways and regulatory networks. The Hippo-YAP signaling pathway an evolutionarily conserved pathway that is involved in controlling organ size and development. Dysregulation of the Hippo-YAP pathway is associated with a large number of human cancers and there are multiple efforts to target key components of the pathway for disease intervention. The Hippo-YAP pathway has been implicated in several types of cancer, involving liver, lung, breast, and colon cancer. Downstream regulator of Hippo pathway – YAP (Yes-Associated Protein) acts as co-activator during transcription and that plays a vital role in regulating cellular growth, differentiation and proliferation, and which is a key effector in the pathway.

In eukaryotes, RNA polymerase II (RNAPII) is responsible for the transcription of protein-coding genes as well as many non-coding genes. Among the three steps transcript elongation is discontinuous and can be perturbed by intrinsic regulatory barriers, such as promoter-proximal pausing (PPP), nucleosomes, secondary structures of RNA and the underlying DNA sequence. The temporary halting of RNA polymerase Ⅱ during the elongation phase is known as promoter-proximal pausing (PPP), regulated by a complex known as negative elongation factor (NELF). Previous work in the lab, discovered that PPP regulates tumorigenesis by regulating Yki transcription, we are investigating if PPP regulated YAP targets are crucial towards neoplastic transformation in mammalian cells as well. Here we explored effect of PPP mediated regulation of YAP target gene expression on breast cancer outcomes using TCGA and an Indian cohort of breast cancer patient samples. We further investigated mechanism of PPP in YAP transcription regulation using breast cancer cell lines.

I dedicate this thesis to my beloved parents, whose unconditional love, unwavering support, and constant encouragement have been my greatest strength throughout this journey.

Acknowledgements

I cannot thank enough to my master's thesis supervisor, **Dr. Madhura Kulkarni**, from the Centre for Translational Cancer Research (CTCR) Pune and Prashant Cancer Care Mission (PCCM) Pune because of her invaluable guidance and support. She has played a pivotal role in shaping my scientific curiosity and confidence, particularly in the field of translational cancer research. Her guidance has been instrumental in shaping my scientific thinking, and her emphasis on research freedom has allowed me to explore and grow as a researcher. Immensely thankful for making my scientific journey a smooth ride alike transcription without perturbations.

I extend my sincere thanks to my thesis expert **Professor Kundan Sengupta** and **Professor L. S. Shashidhara** for their invaluable inputs in the initial days of the project. I am also grateful to the faculty members at IISER Pune for fostering a dynamic research environment. I am also thankful to **Sanket N** for his constant support in the project. **Rishabh K** for helping me out with data analysis. My heartfelt thanks go to my old and current lab mates. I'm grateful to **Aadya A** for mentoring me from the scratch. It was also wonderful to get acquainted with lab rules and discipline with **Pooja V** and **Gomathi V**. **Khushi S** has always been a pillar to all my ups and downs. Thanks to **Devaki K** and **Jisha J** for always spreading cheerfulness in and out of the lab. I am especially grateful to **Vedant and sampada**, for being with me in stressful days and nights, especially during deadlines :). A nerve-racking or trying day couldn't be called without a cup of tea and an incy-wincy gossip session, where I can only think of **Sakshi D**. We had quite a lot of extended members in the lab but I have and will always appreciate **Manav S** for being.

A healthy lab environment is a good mental exercise; hence I would like to give credit to **Irene, Disha, Anuvind, Smera, Bhavesh O, Reva** and **Anuradha** for making it possible. Thanks to **Malavika S** for making my eyes shine and heart smile. I would also like to express my gratitude to non-academic staff at IISER with especially to IISER biology staff who have made administrative work a breeze for us.

There are enormous number of people (in literal sense) I would like to praise and acknowledge but alas it's not possible so last but not the least, I would like to thank my family – my darling mom (**Kapila**) and dad (**Sudhakar**) for standing by me during the all these years. Can't end without appreciating my lovely and supportive sister **Anjali**.

Contributions

Contributor name	Contributor role
Bhavesh vasave, Dr Madhura Kulkarni, Sanket Nagarkar, Rishabh Kulkarni	Conceptualization Ideas
Bhavesh Vasave, Dr Madhura Kulkarni, Rishabh Kulkarni	Methodology
Bhavesh Vasave, Rishabh Kulkarni	Software
Dr Madhura Kulkarni	Validation
Bhavesh Vasave	Formal analysis
Bhavesh Vasave	Investigation
Dr Madhura Kulkarni	Resources
Bhavesh Vasave, Dr Madhura Kulkarni	Data Curation
Bhavesh Vasave	Writing - original draft preparation
Dr Madhura Kulkarni	Writing - review and editing
Bhavesh Vasave	Visualization
Dr Madhura Kulkarni	Supervision
Dr Madhura Kulkarni	Project administration
Dr Madhura Kulkarni	Funding acquisition

Chapter 1

Introduction

1.1 Cancer and Breast Cancer

Cancer is a disease which is characterized by uncontrolled cell growth, with the potential to spread from its original site to other parts of the body(Schwartz, 2024). It remains a major global health, societal, and economic challenge in the 21st century, accounting for nearly one in six deaths (16.8%) worldwide and 22.8% of deaths caused by noncommunicable diseases (NCDs)(Bray et al., 2024). Among the most common cancers affecting women is breast cancer (BC), a significant public health concern that impacts millions worldwide(Obeagu & Obeagu, 2024). It is the most frequently diagnosed cancer in women, with about 2.3 million new cases annually, surpassing lung cancer. In 2020, BC accounted for 2,308,897 cases (23.8% of all female cancers) and resulted in 665,684 deaths globally(Bray et al., 2024).

Breast cancer can be categorized as either non-invasive i.e., benign or invasive i.e., malignant. Between 50% and 70% of breast cancer cases fall under these two classifications. A common non-invasive form is ductal carcinoma in situ (DCIS), where abnormal epithelial cells grow within the breast ducts but remain confined by the myoepithelial-basement membrane. Although DCIS is a precursor lesion to invasive cancer, research suggests that 50–70% of DCIS cases do not progress further(Gibson et al., 2023). Another non-invasive form is lobular carcinoma in situ (LCIS). It originates in the lobules of the breast but does not spread. Despite being labelled as "carcinoma," LCIS is not taken as true breast cancer but rather a benign breast condition. Other benign tumors include rest hyperplasia, fibroma, and phyllodes tumors.

Invasive Ductal Carcinoma (IDC) is the most prevalent form of invasive cancer. In IDC, abnormal epithelial cells grow uncontrollably, breaking through the ductal basement membrane and spreading into surrounding breast tissues(Gibson et al., 2023). These cancerous cells can enter the lymphatic system or bloodstream, leading to metastatic breast cancer, which can affect other organs. IDC is further classified into molecular subtypes based on the presence or absence of specific hormone receptors, determined through immunohistochemical analysis. These subtypes include:

- (ER+) Estrogen receptor-positive
- (PR+) Progesterone receptor-positive (Inic et al., 2014)
- (HER2+) Human epidermal growth factor receptor 2-positive
- (TNBC) Triple-negative breast cancer, which lacks expression of all three receptors (Orrantia-Borunda et al., 2022)

Among these subtypes, TNBC is considered the most aggressive, as it has a high risk of metastasis, recurrence, and poor survival rates (Foulkes et al., 2010) (Kassam et al., 2009) In India, breast cancer remains a significant health issue, with 192,020 new cases and 98,337 deaths (~51% mortality rate) in a single year (Bray et al., 2024).This makes breast cancer the leading cause of cancer-related deaths among Indian women. Globally, TNBC accounts for 10–12% of all breast cancer cases. However, in India, its prevalence is notably higher, ranging from 20–30%. This presents a major clinical challenge, as TNBC lacks specific targetable treatments and is often diagnosed at a more advanced stage(Lee, 2023). Additionally, a larger proportion of Indian TNBC cases occur in younger women compared to global trends, further complicating disease management. The absence of effective targeted therapies highlights the urgent need for improved treatment strategies, particularly in regions where TNBC is more prevalent(Vaid et al., 2022). Among all the key challenges in TNBC, the dysregulation of transcriptional mechanisms that drive cancer progression by sustaining through oncogenic signalling. Understanding how mechanisms of transcription and cell signalling contribute to TNBC pathophysiology could reveal novel therapeutic targets, offering new avenues for intervention in this aggressive cancer subtype along with all the studies.

1.2 Hippo- YAP Pathway and Deregulation in Cancer

The Hippo-pathway is a highly conserved signalling pathway across higher-order vertebrates that modulates target genes to regulate a multitude of biological processes including cell survival, cell proliferation, differentiation, cellular fate determination, organ size control, and tissue homeostasis(Boopathy & Hong, 2019). The YES-Associated Protein (YAP1 or YAP) is an oncoprotein that has a pivotal role in the Hippo signalling pathway, acting as its principal effector (Calses et al., 2019). Along with Transcriptional Co-activator with PDZ-binding Motif (TAZ), YAP functions as a transcriptional co-regulator and downstream mediator of this pathway. The Hippo

pathway primarily governs YAP activity, which is crucial for tissue homeostasis, regeneration, tumour formation, progression, metastasis, invasion, apoptosis suppression, and therapy resistance(Szulzewsky et al., 2021).

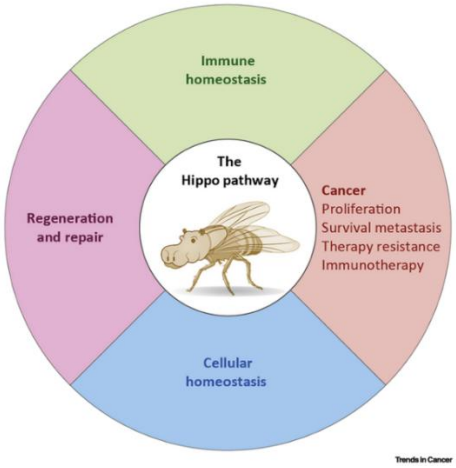


Figure 1: *Hippo pathway discovered in drosophila melanogaster, its mechanisms along with effects of its aberrations (Calses et al., 2019)*

This highly conserved signalling pathway has been extensively studied, particularly in Drosophila, where the YAP homolog, Yorkie (Yki), regulates tissue growth by promoting cell proliferation while inhibiting apoptosis. In mammals, YAP activity is controlled mainly through phosphorylation by Large Tumour Suppressor kinases (LATS1 and LATS2). These kinases are activated by Mammalian STE-like kinases (MST1 and MST2) or Mitogen-Activated Protein Kinase Kinase 4 (MAP4K4)(Ibar & Irvine, 2020).

Upon phosphorylation, YAP-TAZ proteins are sequestered in the cytoplasm by binding to 14-3-3 proteins, which leads to their degradation through SCFβ-TRCP E3 ubiquitin ligase. However, this phosphorylation process is highly dynamic and can be quickly reversed by phosphatases. When the Hippo pathway is inactive, YAP-TAZ accumulate in the nucleus, where they interact with DNA-binding proteins, such as TEAD, to activate gene transcription.(Lamar et al., 2012)

External factors, such as cell density, shape, and adhesion, regulate the Hippo pathway. When activated, Hippo signalling phosphorylates MST1/2, which in turn phosphorylates LATS1/2 and their cofactor MOB, with assistance from the adaptor protein Sav. This LATS/MOB complex phosphorylates YAP, leading to its proteasomal degradation or cytoplasmic sequestration via 14-3-3 proteins. Additionally, several

modulators directly influence YAP/TAZ activity. For instance, AMPK, a key regulator of cellular metabolism and energy balance, inhibits YAP by phosphorylating multiple sites, including Ser94, which disrupts the YAP-TEAD interaction. In contrast, phosphorylation by CDK1 or NLK enhances the nuclear localization and activation of YAP-TAZ. Notably, NLK-mediated phosphorylation at Ser128 disrupts 14-3-3 binding, promoting nuclear retention and activation of YAP(Koo & Guan, 2018).

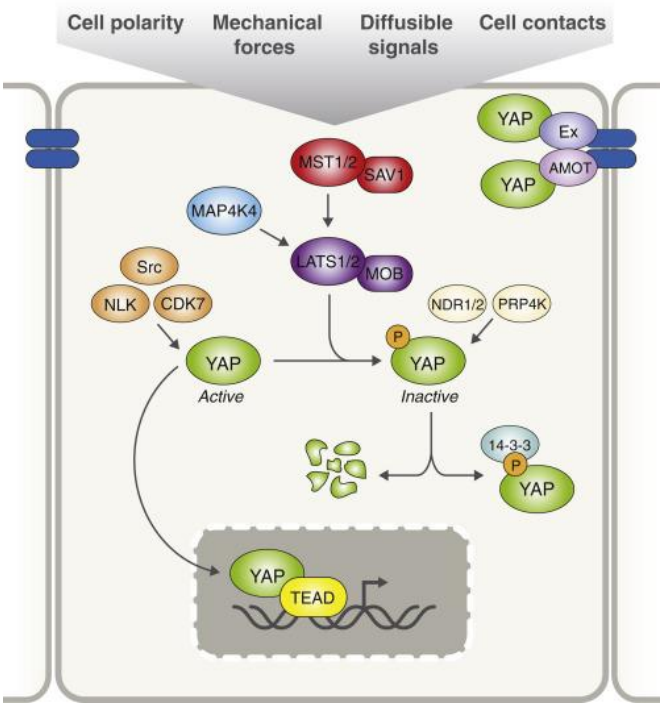


Figure 2: *Cascade of proteins involved in canonical Hippo signalling pathway (Ibar & Irvine, 2020)*

Research suggests that a suppressed Hippo pathway increases YAP/TEAD-dependent gene expression, which contributes to tumour growth and metastasis by enhancing oncogenic processes at both the primary tumour site and metastatic locations. YAP has various downstream transcriptional targets that regulate the oncogenic transformation of cells(Zanconato et al., 2016).

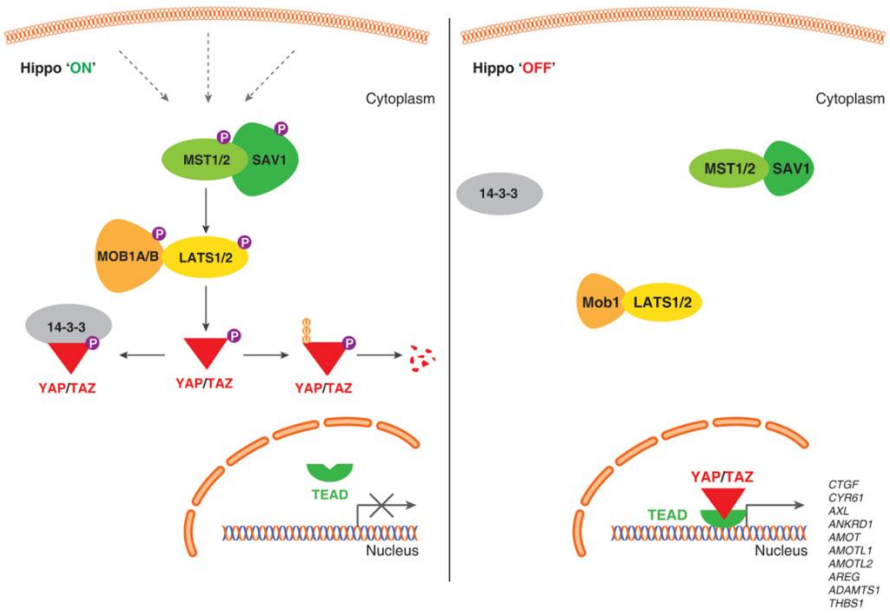


Figure 3: *YAP activity during active and inactive state of Hippo-YAP pathway with its downstream transcriptional targets (Boopathy & Hong, 2019)*

1.3 Causality of Promoter proximal Pausing (PPP) during transcription

Gene expression is a tightly regulated and dynamic process that ensures the steady production of some proteins while allowing for rapid changes in the levels of others(Rodríguez-Molina et al., 2023). In most cellular functions, including development, proliferation, and responses to environmental signals, transcription serves as the primary rate-limiting step.

In eukaryotes, RNA polymerase II (Pol II) drives transcription through three fundamental stages: initiation, elongation, and termination. However, the transcription process is not always smooth. Recent advancements in research have revealed that transcript elongation is highly dynamic, with Pol II moving in a discontinuous and irregular manner along genes. This process is marked by transcriptional pausing, which can either slow down or support elongation, aid in RNA processing, or, in some cases, lead to premature termination of transcription (Noe Gonzalez et al., 2021)

To comprehend the factors leading to transcription stalling and arrest—collectively referred to as transcription stress—it is essential to investigate the structural alterations and enzymatic mechanisms governing transcript elongation. Various factors influence RNA polymerase II (Pol II) stalling, including nucleosomes, promoter-

proximal pausing, and DNA sequence barriers that regulate transcription progression (Noe Gonzalez et al., 2021)

According to a conventional understanding in elongation phase during transcription RNA Pol II proceeded only after passing promoter. Studies in higher organisms have shown that polymerase II can remain associated with induced gene promoters even after stimulation and still transcriptionally active (Gilmour & Lis, 1986). Further research revealed that a rate-limiting step exists between transcription initiation and elongation, and in many genes o metazoan, polymerase II pauses and gets stacked along the DNA template shortly after the promoter (Jonkers & Lis, 2015).This pausing typically occurs 30-60 base pairs downstream from the transcription start site and is referred to as promoter-proximal pausing (PPP) (Kwak et al., 2013).

PPP is regulated by a complex network of molecular interactions that control gene expression. The three most critical regulatory complexes involved in this process are:

1. The 7SK small nuclear ribonucleoprotein (snRNP) complex(Peterlin et al., 2012)
2. The Negative Elongation Factor (NELF) complex(Core & Adelman, 2019)
3. The Positive Transcription Elongation Factor Complex (P-TEFb) (Nguyen et al., 2001)

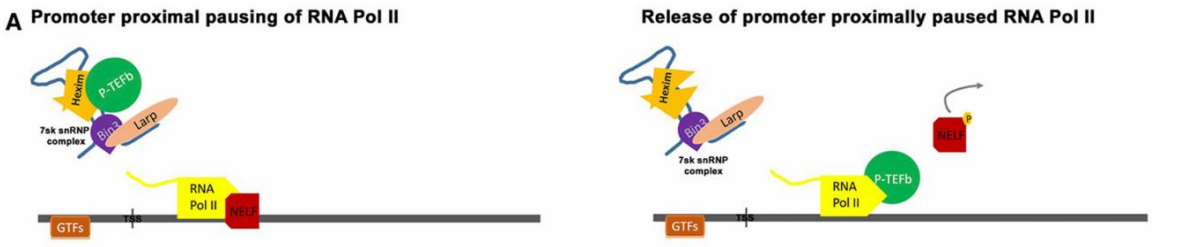


Figure 4: *RNA polymerase II Paused by the NELF complex until 7SKsnRNP complex which is sequestered by P-TEFb gets released to resume the RNA Pol II activity*

For transcription to proceed efficiently, RNA Pol II must transition from its paused state. This pause is regulated by the NELF complex and DRB-sensitivity-inducing factor (DSIF), which associate with RNA Pol II and stall its movement just downstream of the promoter(Muse et al., 2007). The release from pausing is facilitated by P-TEFb, a

complex composed of cyclin T and cyclin-dependent kinase CDK9. Upon recruitment, CDK9 phosphorylates NELF and DSIF, causing NELF to dissociate from the complex, while DSIF converts into a positive elongation factor that supports RNA Pol II in productive elongation(Yamaguchi et al., 1999).

The regulation of P-TEFb itself has a crucial part in transcription control. This complex is sequestered by the 7SK snRNP complex, which prevents it from mediating pause release. The 7SK snRNP complex, which consists of 7SK RNA, HEXIM1/2, LARP7, and MePCE, controls the availability of P-TEFb. Various signalling pathways, including ERK and TCR, can trigger the release of P-TEFb, highlighting how its sequestration and activation are context-dependent processes essential for gene regulation(Yik et al., 2003).

1.4 Promoter-Proximal Pausing (PPP), YAP/Yorkie, and Cancer Progression

The Drosophila ortholog of YAP1, known as Yorkie (Yki), functions as a transcriptional cofactor in the Hippo tumour suppressor pathway, a signalling cascade that is evolutionarily conserved from flies to humans. In Drosophila, Yki regulates genes that influence cell growth and survival, including Diap1, dMyc, and bantam. In humans, YAP target genes such AREG, CTGF and Cyr61 are known as the EGFR-ligand. Where these genes contribute to Yki/YAP-driven growth, they are not solely responsible for its full effects, suggesting the involvement of additional regulators(Nagarkar et al., 2020).

A genetic screen conducted in *Drosophila* identified numerous genes that, when suppressed, enhanced Yki and EGFR-driven growth(Groth et al., 2020). Notably, some of these gene's function as tumour suppressors, as their downregulation—when combined with Yki or EGFR overexpression—led to neoplastic growth. Among the identified genes, a few were specifically linked to the regulation of promoter-proximal transcriptional pausing, highlighting a potential connection between transcriptional control and tumour progression. As both Hippo signalling – Yki pathway and Promoter Proximal Pausing are highly conserved across metazoans these observations led us to look into function of YAP and its oncogenic mechanisms(Groth et al., 2020; Nagarkar et al., 2020).

1.4 Previous data (clinical association of NELFA and YAP) and other TFs

Research on the Negative Elongation Factor (NELF) complex has explored its role in gene expression regulation in breast cancer cells extended this investigation to axonal levels, using exon arrays to analyse transcriptional regulation and RNA processing beyond conventional gene-level profiling. Their findings revealed that NELF is crucial for sustaining the expression of key cell cycle regulators, thereby supporting proper cell proliferation and progression. The study demonstrated that NELF facilitates transcription by maintaining chromatin accessibility and aiding in the assembly of preinitiation complexes (PICs) at gene promoters. Rather than merely suppressing transcription elongation, NELF promotes sustained RNA polymerase II (RNAPII) engagement with active genes, ensuring their proper expression. Chromatin analysis further revealed that NELF depletion reduces histone modifications associated with active transcription, such as H3K9Ac. This suggests that NELF plays a role in recruiting histone-modifying enzymes to establish an epigenetic environment conducive to transcription(Sun & Li, 2010).

Interestingly, NELF was found to exert gene-specific regulatory effects. It activates certain genes, those involved in the cell cycle, while repressing others, like *JUNB*, indicating that its function depends on chromatin context, promoter strength, and gene-specific factors. A key discovery was that while knocking down NELF-B had minimal impact, depletion of NELF subunits A, C, and E significantly affected multiple cell cycle genes, suggesting distinct functional roles among the subunits. Computational analysis further identified NELF-regulated exon targets, particularly in genes with alternative promoters, highlighting its complex influence on transcription(Sun & Li, 2010).

A recent study also highlighted the critical role of the NELF complex in breast cancer progression and metastasis, regardless of the cancer subtype. Researchers found that NELF facilitates epithelial-mesenchymal transition (EMT) and mammary stem cell programs by interacting with the transcription factor SLUG and the histone acetyltransferase KAT2B. This interaction drives key gene expression changes necessary for cancer metastasis. The study further revealed that NELF is essential for

tumour development, as its loss significantly reduces the tumorigenic properties of breast cancer cells. Specifically, NELFE was found to interact with major EMT regulators, forming a complex with SLUG and KAT2B to activate genes like *ZEB1*, which promote EMT and metastasis. While SLUG is typically recognized as a transcriptional repressor, these findings suggest that in conjunction with NELF and KAT2B, it can also function as a transcriptional activator(Zhang et al., 2023).

Since NELF and KAT2B are not required for normal tissue function, disrupting their interaction in breast cancer could present a promising therapeutic approach. The link between Yorkie overexpression causing tumorigenesis and the deregulation of promoter-proximal pausing underscores the importance of transcriptional regulation in cancer progression. Notably, the discovery that the NELF complex and 7SK snRNP complex function in humans similarly to their Drosophila counterparts reinforces their significance as potential targets in cancer research(Zhang et al., 2023).

These parallels motivated us to start with investigation of cancer patient data in context to the NELF complex subunit- NELFA in conjunction with YAP in an Indian cohort. For that total 85 FFPE samples of breast tumours along with their demographic data were procured with appropriate ethics approvals. Out 85, 75 patient samples were stained for YAP and NELFA expression. FFPE tumour sections were then stained for YAP and NELFA by immunohistochemistry (IHC). The staining was scored for intensity and percent expression. Survival analysis was done using the Kaplan-Meier method to assess association of NELFA and YAP expression with patient outcomes. Among subtypes high NELFA and high YAP patients especially in TNBC show low survival outcomes whereas patients with high YAP and low NELFA expression show high recurrence rate, especially in TNBC. These findings led us to investigate underlying mechanism in mammalian breast cancer cell lines.

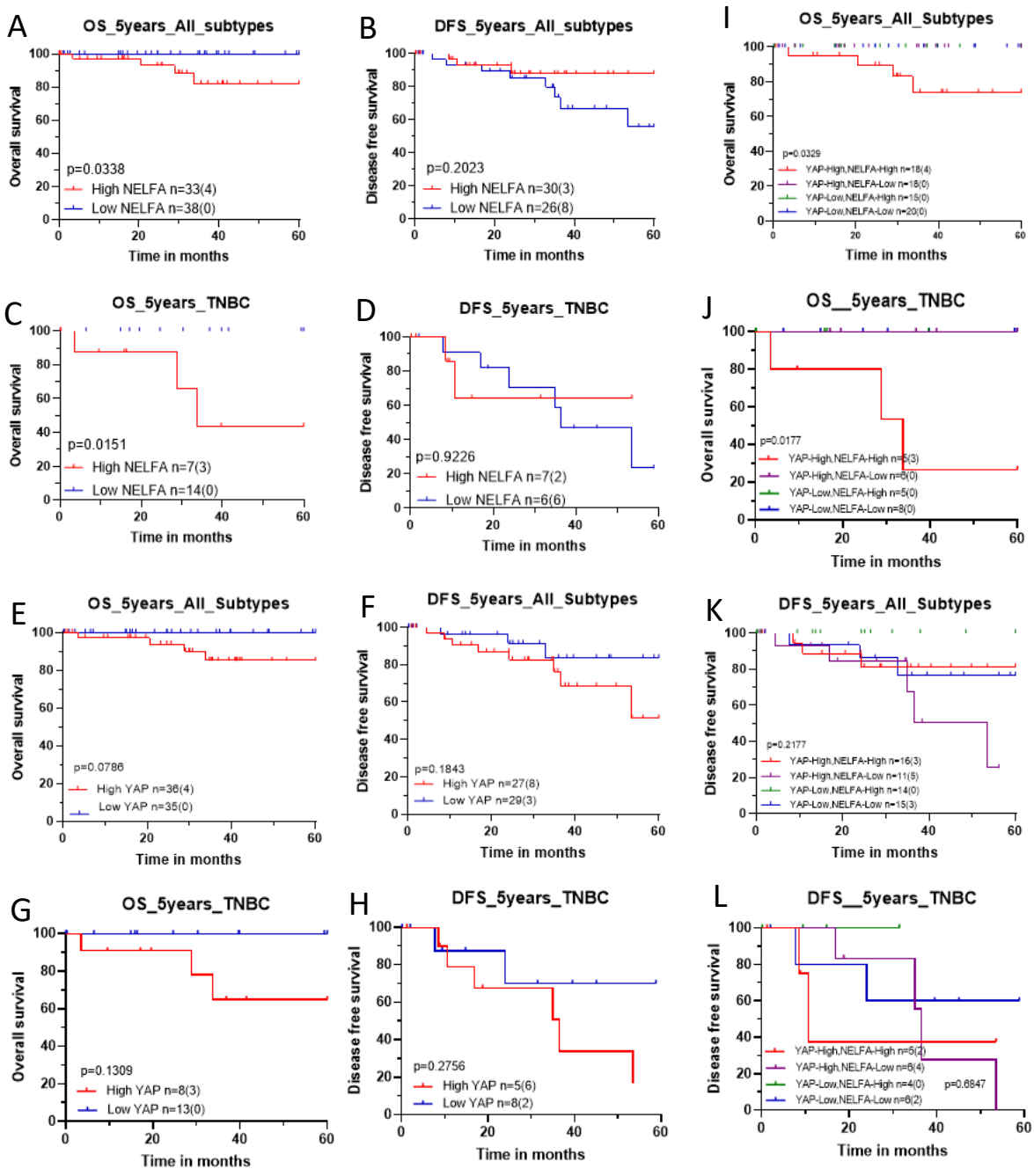


Figure 5: Gene expression and its association with survival outcomes*. (A)-(L) KM plots of Indian breast cancer patients divided based on expression of YAP and NELFA. Based on a set cut off high and low expression of genes was determined (High- red and Low- blue). High and Low NELFA overall and disease-free survival outcomes are plotted in all subtypes and in TNBC (A)-(D). High and Low YAP expression in all subtypes and in TNBC (E)-(H). KM plots of Breast cancer patients categorized based on a combination of high YAP expression with NELFA (I)-(L).*

1.5 Aims and objectives

Aims:

- To investigate role of PPP in regulating YAP mediated transcription in TNBC cell lines.
- To send RNA samples for whole transcriptome sequencing in order to identify a global transcriptional co-regulation by YAP and PPP.

Objectives:

- To knock down NELFA and YAP using siRNA in a TNBC breast cancer cell line and validate the knock down and gene expression of YAP targets using RT-qPCR and western blot.
- Identify and select RNA samples from the knockdown studies and send them for total RNA sequencing and to analyse the data in order to look at global level changes of NELF and also understand why patient outcomes are in a certain way.

Chapter 2

Materials and Methods

2.1 Cell Culture

2.1.1 Cell lines and reagents

Two mammalian cell lines – HEK293Ts and MDA-MB-231 were used in the study. **HEK293T** is a human embryonic kidney cell line in which large SV40 T antigen is expressed. The cell line has been widely used in research due to its reliable fast growth and propensity for transfection. It is also used in gene therapy and bioprocessing. **MDA-MB-231** is a triple-negative breast cancer (TNBC) cell line which is highly aggressive with metastatic ability and resistant to drugs. These cell lines were cultured in (DMEM) Dulbecco's Modified Eagle Medium High glucose *(HIMEDIA- #AL066A)* supplemented with 1x sodium pyruvate *(GIBCO- Cat. No. 1136007)*, 10% FBS *(Fetal Bovine Serum, qualified, Brazil, GIBCO – Cat. No. 10270106)* and 1% Penicillin-Streptomycin *(Penicillin-Streptomycin, Sigma-Aldrich, Cat. No. P4333)* in standard conditions incubated at 37° C and 5% CO2. Three types of plates (6cm – 60mm, 10cm – 100mm and 6 well) (TPP tissue culture plates) were used to maintain the cells as per the experimental requirement. 6cm plate was used for RNA extraction after 70-80% confluency of cells and 10cm plate was used for protein extraction after the similar confluency. 6 well plate was used for knockdown – siRNA experiments. The method used for maintaining the cells and passaging them was either reviving them or splitting them into further passages. The steps are as follows:

2.1.2 Reviving cells

Cryo vials from -80 degrees or liquid nitrogen were carried in an ice box to cell culture. The plates, falcons and serological pipettes to be used were UV ed in prior. The vial was speed thawed with the help of water bath. When the vial was 40% thawed, 1ml of media was used to pipette mix and resuspend the cells which were frozen using media and DMSO. Once everything was resuspended properly the 2ml resuspended cells were pelleted down and resuspended in fresh 1ml complete media (Serum free media+ supplements) and were seeded in respective plates drop wise. 8ml, 4ml and 3/4ml complete media was used in a 10cm, 6cm and 6 well plate respectively. Post

seeding plate was observed under microscope and confluency and cell death was noted every day for the record. Seeding density of HEK293Ts is 2M and MDA-MB-231 is 1.2M.

2.1.3 Splitting cells

Once the revived cells in the plate were 70-80% confluent they were split into 6cm/10cm/6 well plate as per the requirement. 2ml trypsin was added to 10cm plate and incubated for 2-3 min (incubation time varies with cell lines) post two PBS washes. After trypsinization the cells were collected and added to the 15ml flacon containing 6ml media. Hence the media: trypsin ratio stays 3:1. The falcon was then spun down in a centrifuge at 1000rmp for 5 mins. Before putting it for centrifuge 10ul of cell suspension was taken for the cell count. Haemocytometer was used for counting the cells (representative calculation in the fig). After the centrifuge pelleted cells were resuspended in fresh complete media and was seeded according to the calculations. An extra 10cm plate was always maintained for further passages and extra cells were frozen in cryo vials using complete media+ DMSO in 9:1 ratio.

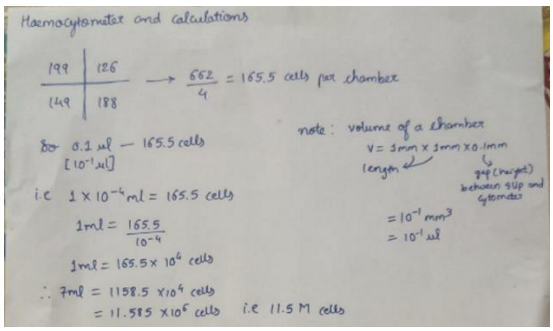


Figure 6: *A mock calculation for cell count used during passaging using haemocytometer*

RNA extracted from cells was used for cDNA followed by RT-qPCR to look at mRNA expression level of desired genes. Similarly, protein extracted was used for western blotting.

2.2 RNA extraction and RT-qPCR

2.2.1 RNA isolation

RNA was isolated using TRIzol *(Invitrogen, Cat. No. 15596026)* method. Following steps were followed for RNA extraction-

Cells were given two 1X PBS washes.1ml TRIzol was used for 6cm plate. By pipetting, cells were thoroughly resuspended in TRIzol and collected in a 1.5ml Eppendorf tube using a cut tip. 200ul of chloroform was added to the 1.5ml tube and incubated at room temperature for 20 min. Post 20 min incubation the tube was centrifuged at 14000rpm at 4 degrees for 10 min. The centrifugation formed three layers: Top- aqueous, Middle- interphase, Bottom- organic phase. The aqueous layer was collected in a fresh 1.5ml tube to which 500ul of isopropanol was added for RNA precipitation and incubated for 10 min at room temperature.

After the incubation the tube was centrifuged at 14000rpm at 4 degrees for 10 min. Post 10 min of centrifuge the pellet was marked and supernatant was discarded. 80% of ethanol in DEPC water was added to the tube and was again centrifuged at 14000rpm at 4 degrees for 10 min to give a wash. The step was repeated twice and supernatant was discarded. The tube was kept inverted on tissue for air dry. Once the tube was complete dried off the pellet was resuspended in 20ul of RNase free water. NanoDrop was used to check the concentration and quality of RNA extracted.

2.2.2 cDNA synthesis

cDNA was synthesised from the extracted RNA using iScript cDNA synthesis kit *(BIO-RAD, Cat. No. #1708891)*. 10ul of reaction mix is prepared from the kit protocol and 1ug of RNA goes in 10ul of reaction mix. Following is the representative image for the contents added and used for reaction mixture preparation plus the protocol followed to set up in a thermal cycler.

Component	Volume per Reaction, µl
5x iScript Reaction Mix	4
iScript Reverse Transcriptase	1
Nuclease-free water	Variable
RNA template (100 fg–1 µg total RNA)*	Variable
Total volume	**20**

* When using larger amounts of input RNA (>1 µg), the reaction should be scaled up (for example, 40 µl reaction for 2 µg, or 100 µl reaction for 5 µg) to ensure optimum synthesis efficiency.

Reaction Protocol

Incubate the complete reaction mix in a thermal cycler using the following protocol:

Priming	5 min at 25°C
Reverse transcription	20 min at 46°C
RT inactivation	1 min at 95°C
Optional step	Hold at 4°C

Figure 7: *Content chart and reaction protocol used for cDNA synthesis*

2.2.3 qPCR (real-time PCR)

The cDNA, primers (forward and reverse), and 2X SYBR *(iTaq-universal SYBR Green supermix 2X- BIO-RAD, Cat. No. #38220090)* were kept on ice. PCR plate and sealer were kept aside. The PCR plate was labelled according to the prepared plate map. Then 600µl Eppendorf tubes were labelled according to the genes. Master mix was prepared for each gene.

For 1 PCR reaction, the master mix is prepared by adding-

-Nuclease-free water (3.6µl)

-100nM forward and reverse primer mix (0.2µl R + 0.2µl F)

-2X SYBR (5µl)

After making the master mix, Eppendorf was short-spun to get rid of bubbles, if any. 1ul cDNA was added to wells according to the plate map, followed by 9ul of master mix added to each appropriate well. It was ensured that the master mix was resuspended and mixed thoroughly with a pipette while avoiding the formation of bubbles before adding the wells. (This was done to ensure homogeneity of the master mix as SYBR Green is highly viscous, settles down on spinning.) The final volume in each well - 1µl cDNA + 9µl master mix = 10µl. Plate was sealed with the sealer, and the gel cutter was used to seal the edges properly. It was followed by giving the PCR plate a spin at 2000 rpm for 1 minute to ensure no sample is on the walls. Later the PCR plate was placed in Bio-Rad Real-Time PCR System. The reactions were run in

duplicates and were analysed in Excel and GraphPad Prism. Normalization of RT-PCR was computed using 2^-dCt values and fold change with respect to housekeeping gene GAPDH. Primer sequences and cycling conditions are shown in Table 1.

Gene		Sequence	Bp
YAP	forward	ATCCCAGCACAGCAAATTCT	20
	reverse	TGGATTTTGAGTCCCACCAT	20
Cyr61	forward	GTGTGAAGAAATACCGGCCC	20
	reverse	CTGTAGAAGGGAAACGCTGC	20
CTGF	forward	GGCCCAGACCCAACTATGAT	20
	reverse	TGGGAGTACGGATGCACTTT	20
ANKRD1	forward	TGAATCCACAGCCATCCACT	20
	reverse	TCCTTCTCTGTCTTTGGCGT	20
HEXIM1	forward	CATGACTCCGAGGCCAGTAA	20
	reverse	AGGCTCTGTTTCTCGTCGAA	20
HEXIM2	forward	CAGGGAACCACCAGAGTCAT	20
	reverse	ACCGCCTGTAATGCAGAGTC	20
MePCE	forward	AGGCAGAGCACCACATCATA	20
	reverse	GGAGCGGACACATCAGTCTT	20
NELFA	forward	TGGATGATCTCCATTAGGGC	20

Table no. 1: *Primer sequences used for RT-qPCR*

The same procedure was followed for all target genes. The rest of the target genes were done in 3Ns and the results were combined. For multiple biological replicates, i.e., N1, N2, and N3, a combined plot can be prepared, for which an unpaired t-test can be done to check the significance of the results. P value < 0.05 is considered to be significant.

2.3 Protein extraction and Immunoblotting

Whole cell lysates of cell lines were extracted from cell pellets by taking equal volumes of modified RIPA buffer. RIPA buffer was prepared in house by taking 20 mM Tris-HCl pH 8.0, 420 mM NaCl, 10% Glycerol, 0.5% NP-40, 0.1 mM EDTA, with 1 mM DTT, 10 mM PMSF and 20 mM protease inhibitor. The cells were incubated with RIPA buffer and inhibitors cocktail for 15 min on mini-Twist at 4 degrees. Post incubation they were again incubated on ice for 40 mins and were resuspended at regular interval of 10 min. Extracts were then centrifuged at 12000 rpm for 15 minutes. The supernatants i.e., protein was estimated using Bradford assay

After protein estimation, samples for loading were prepared using 4X Laemmli loading dye. SDS-PAGE was made using 30% Acrylamide/Bis-acrylamide solution (29:1), lower Tris buffer (pH 8.6) and upper Tris buffer (pH 8.3) for resolving and stacking gel respectively, 10% APS and TEMED and Milli-Q**.** Volume table is given below (Table no. 2). The solidified gel was then placed in the cassette the tank filled with reservoir buffer. Samples were heated in Thermomixer at 95 degrees before loading in the wells. Reservoir buffer was prepared using Tris base, glycine and 10% SDS having pH of 8.3. The gel run was kept at 70V for stacking of the samples and it entered the resolving part the voltage was changed to 150V. After 1.5hrs of gel run the gel was taken out, stacking was cut out and resolving part of the gel was placed in between blotting papers and foam pads held with gel holder cassette with the PVDF membrane. The cassette was then placed in buffer tank along with electrode assembly. The tank was filled with cold transfer buffer and a small cooling unit was placed I the tank. With constant current of 330mA the transfer was kept for 2 hours.

For one -1mm gel		
	Resolving 10% (5ml)	Stacking 5% (3ml)
MilliQ	2.1 ml	1.59 ml
Acrylamide/Bis-acrymalide	1.67 ml	510 ul
4X lower Tris buffer	1.25 ml	NA

4X Upper Tris Buffer	NA	750 ul
10% APS	50 ul	15 ul
TEMED	4 ul	6 ul

Table no. 2: *Reagents used for making 1mm SDS page*

Proteins separated by SDS-PAGE and transferred onto PVDF membrane were probed with primary antibodies after blocking it with milk for 30 min. Blots were subsequently incubated with secondary antibodies. Stripping buffer prepared with glycine, SDS and Tween20 was used to strip the blot if different proteins were probed on a single blot. Blots were imaged using the ImageQuant™ LAS 4000 biomolecular imager where bands were detected using chemiluminescence. The intensity of bands was quantified using ImageJ software. The list of antibodies used is provided in Table no. 3

1° Antibody	Source: Cat. No.	Dilution	2° Antibody	Source: Cat. No.	Dilution
GAPDH	Santa Cruz: sc-32233	1:1000	Anti-Mouse HRP (m-IgG Fc BP-HRP)	Santa Cruz: sc-525409	1:10000
YAP1	AbCam: ab52771	1:1000	Mouse anti-Rabbit HRP (IgG- HRP)	Santa Cruz: sc-2357	1:10000
NELFA	Bethy: A301-910A	1:500	Mouse anti-Rabbit HRP (IgG- HRP)	Santa Cruz: sc-2357	1:5000

Table no. 3: List of antibodies used for western blotting

2.4 Small interfering RNA transfection

2.4.1 Reagents and protocol

For knockdown studies RNA interference was performed in MDA-MB-231 cell line. siRNA against NELF-A *(ON-TARGETplus Human NELFA siRNA smartpool #L-012156-00-0005)* YAP *(ON-TARGETplus Human YAP1 siRNAsmartpool #L-012200-00-0005)* and control siRNA *(Non-targeting Pool #D-001810-10-05)* were ordered from Dharmacon and transfected in MDA-MB-231 cell line using DharmaFECT Transfection Reagent *(#T-2001-02)* according to manufacturer's protocol. Standardization of siRNA concentration and RNA extraction time point was done using different parameters. Cells were seeded in 6 well plate and were transfected with 25nM, 50nM and 100nM siRNA concentration as a final concentration.

After calculating the volume of siRNA to be added, mastermix was prepared in a falcon for transfection reagent DharmaFECT plus serum free media. Mastermix for siControl, siNELFA and siYAP with serum free media were prepared in three different tubes. Both the mastermix were gently mixed and incubated for 5 min at room temperature. After the incubation transfecting reagent was added to each tube, resuspended gently and later tubes were incubated at RT for 20 min. Post incubation the mixture was added to respective wells and cells were incubated at 37 degrees in 5% CO2. Cells were harvested at 80% confluency and RNA was extracted on two different time points - 48hrs and 72hrs after transfection. Extracted RNA was used to prepare cDNA and was followed by RT-qPCR and analysis using 2^-dCt values.

2.4.2 Standardization

Cells were initially transfected with only 25nM of siRNA according to manufacturer's protocol. No knockdown of genes was observed with 25nM of siRNA transfected (Figure 8).

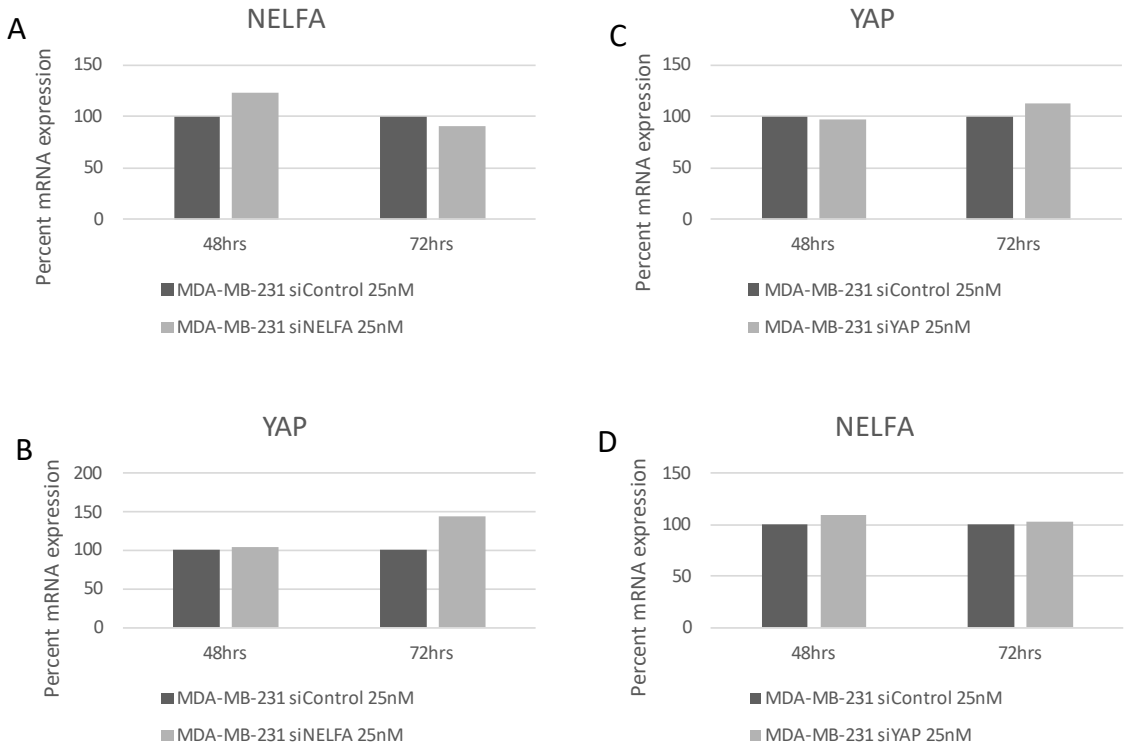


Figure 8: Percent mRNA expression of NELFA and YAP (N0). *NELFA (A) and YAP (B) in NELFA Knockdown MDA-MB-231, YAP (C) and NELFA (D) expression in YAP knockdown MDA-MB-231 with 25nM of siRNA concentration*

Since 25nM siRNA concentration did not work, 50nM and 100nM siRNA concentration were chosen with two time points- 48hrs and 72hrs for RNA extraction. NELFA and YAP were knocked down at varying percentages at different time points and concentrations. We also checked mRNA expression levels of YAP dependent gene expression Cyr61, CTGF and ANKRD1 (Figure 9) which were shortlisted form various YAP targets based on a previous study in lab.

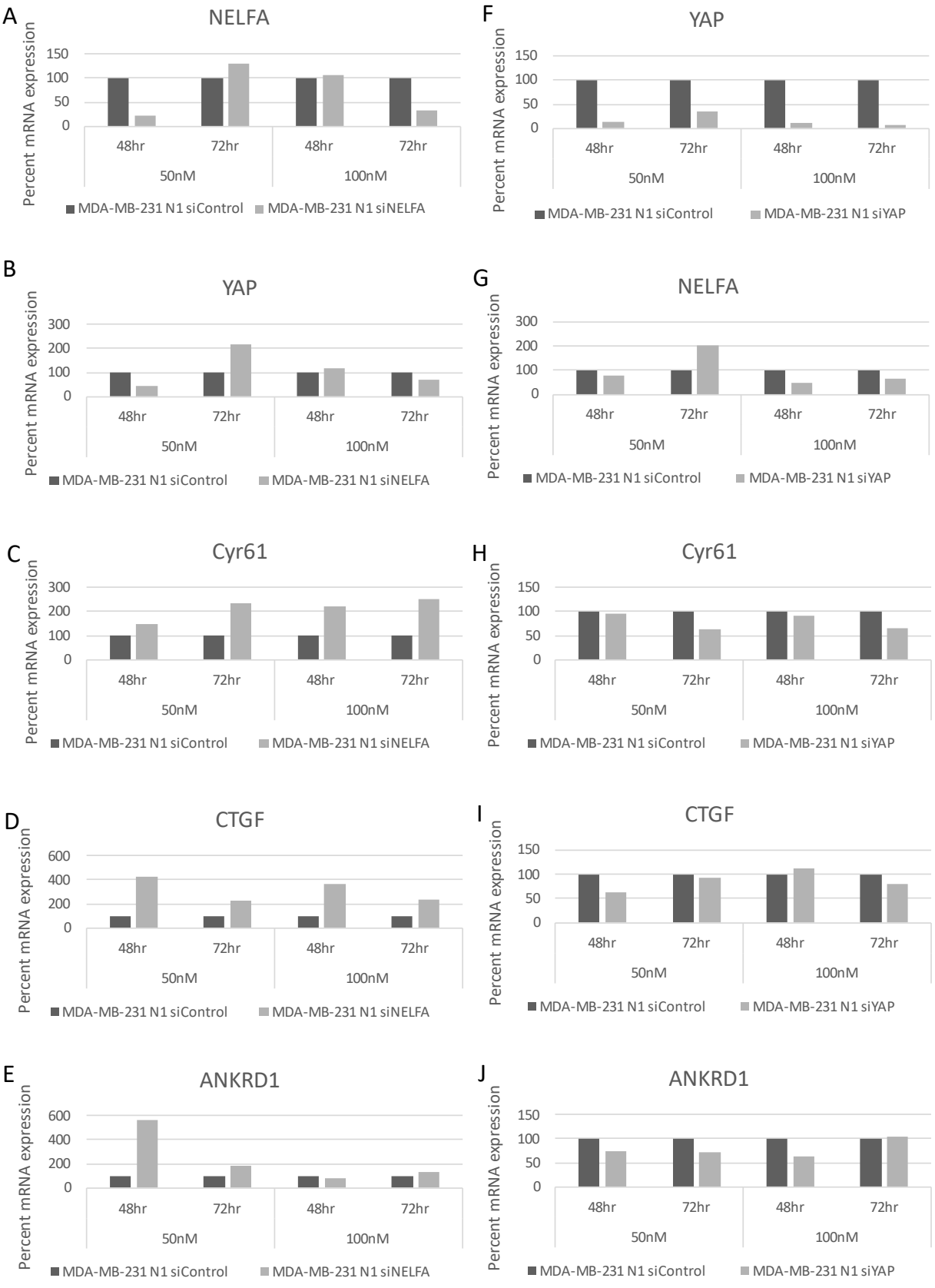

A
NELFA
Percent mRNA expression
150
100
50
0
48hr
72hr
48hr
72hr
50nM
100nM
MDA-MB-231 N1 siControl
MDA-MB-231 N1 siNELFA
F
YAP
Percent mRNA expression
150
100
50
0
48hr
72hr
48hr
72hr
50nM
100nM
MDA-MB-231 N1 siControl
MDA-MB-231 N1 siYAP
B
YAP
Percent mRNA expression
300
200
100
0
48hr
72hr
48hr
72hr
50nM
100nM
MDA-MB-231 N1 siControl
MDA-MB-231 N1 siNELFA
G
NELFA
Percent mRNA expression
300
200
100
0
48hr
72hr
48hr
72hr
50nM
100nM
MDA-MB-231 N1 siControl
MDA-MB-231 N1 siYAP
C
Cyr61
Percent mRNA expression
300
200
100
0
48hr
72hr
48hr
72hr
50nM
100nM
MDA-MB-231 N1 siControl
MDA-MB-231 N1 siNELFA
H
Cyr61
Percent mRNA expression
150
100
50
0
48hr
72hr
48hr
72hr
50nM
100nM
MDA-MB-231 N1 siControl
MDA-MB-231 N1 siYAP
D
CTGF
Percent mRNA expression
600
400
200
0
48hr
72hr
48hr
72hr
50nM
100nM
MDA-MB-231 N1 siControl
MDA-MB-231 N1 siNELFA
I
CTGF
Percent mRNA expression
150
100
50
0
48hr
72hr
48hr
72hr
50nM
100nM
MDA-MB-231 N1 siControl
MDA-MB-231 N1 siYAP
E
ANKRD1
Percent mRNA expression
600
400
200
0
48hr
72hr
48hr
72hr
50nM
100nM
MDA-MB-231 N1 siControl
MDA-MB-231 N1 siNELFA
J
ANKRD1
Percent mRNA expression
150
100
50
0
48hr
72hr
48hr
72hr
50nM
100nM
MDA-MB-231 N1 siControl
MDA-MB-231 N1 siYAP

Figure 9: Percent mRNA expression of NELFA and YAP (N1). *NELFA (A), YAP (B) and YAP targets Cyr61 (C), CTGF (D), ANKRD1 (E) in NELFA Knockdown MDA-MB-231 and YAP (F), NELFA (G) and YAP targets Cyr61 (H), CTGF (I), ANKRD1 (J) expression in YAP knockdown MDA-MB-231 with 50nM and 100nM of siRNA concentration at two 48hrs and 72hrs time point*

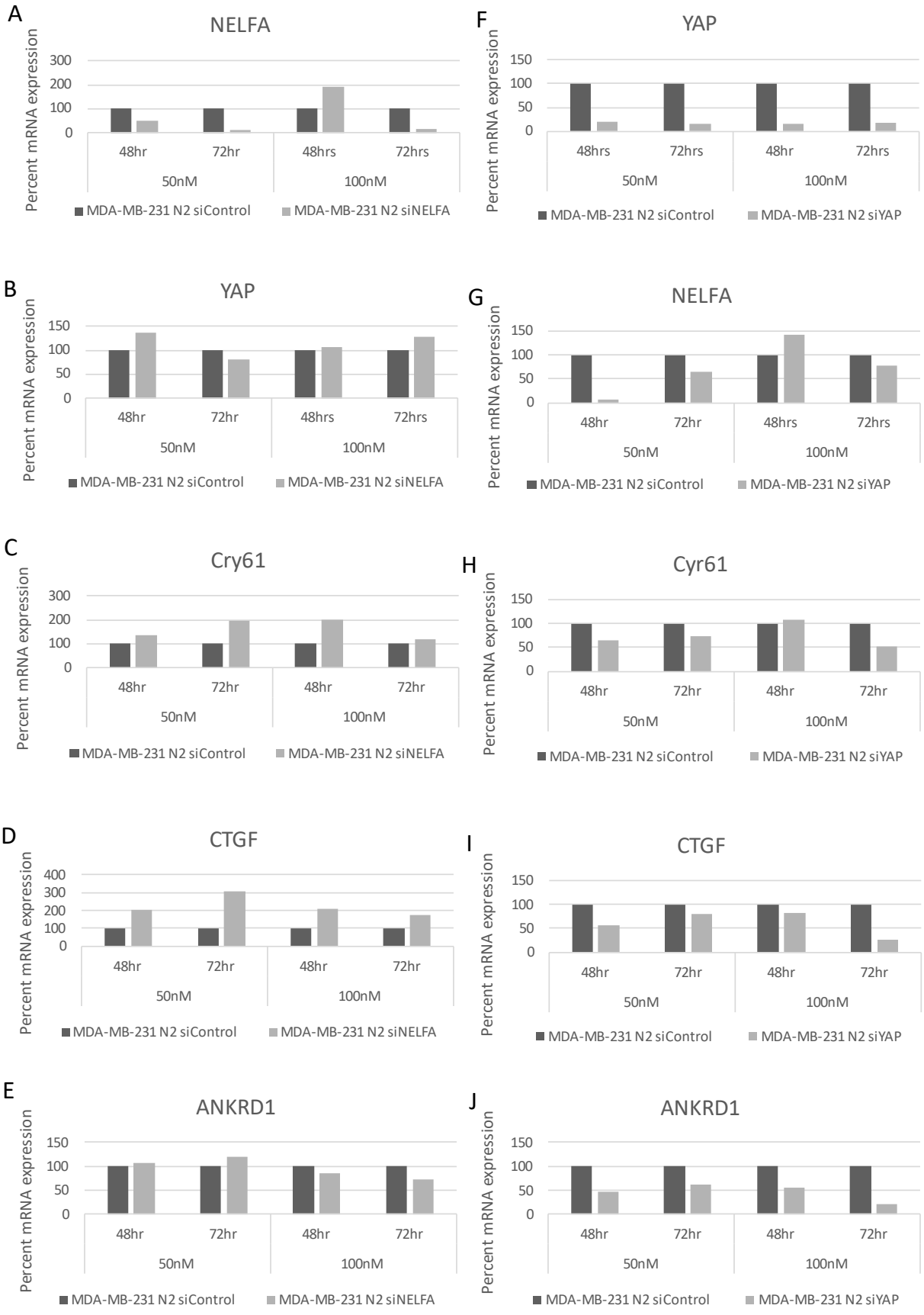
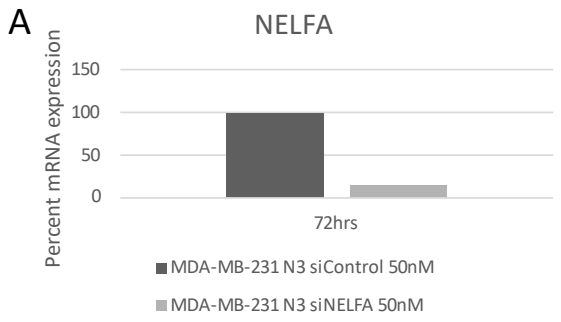
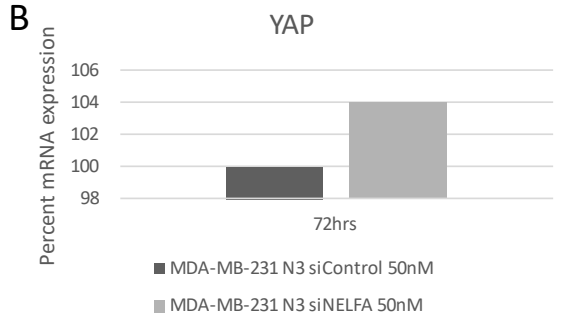
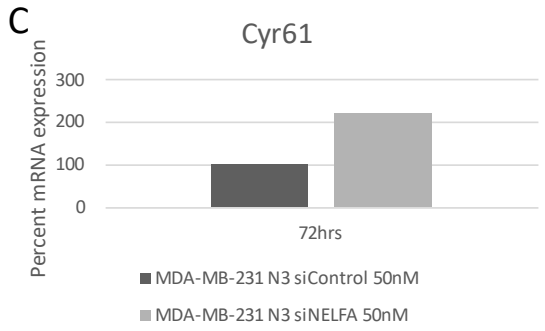

A
NELFA
Percent mRNA expression
300
200
100
0
48hr
72hr
48hrs
72hrs
50nM
100nM
MDA-MB-231 N2 siControl
MDA-MB-231 N2 siNELFA
F
YAP
Percent mRNA expression
150
100
50
0
48hrs
72hrs
48hr
72hrs
50nM
100nM
MDA-MB-231 N2 siControl
MDA-MB-231 N2 siYAP
B
YAP
Percent mRNA expression
150
100
50
0
48hr
72hr
48hrs
72hrs
50nM
100nM
MDA-MB-231 N2 siControl
MDA-MB-231 N2 siNELFA
G
NELFA
Percent mRNA expression
150
100
50
0
48hr
72hr
48hrs
72hrs
50nM
100nM
MDA-MB-231 N2 siControl
MDA-MB-231 N2 siYAP
C
Cry61
Percent mRNA expression
300
200
100
0
48hr
72hr
48hr
72hr
50nM
100nM
MDA-MB-231 N2 siControl
MDA-MB-231 N2 siNELFA
H
Cyr61
Percent mRNA expression
150
100
50
0
48hr
72hr
48hr
72hr
50nM
100nM
MDA-MB-231 N2 siControl
MDA-MB-231 N2 siYAP
D
CTGF
Percent mRNA expression
400
300
200
100
0
48hr
72hr
48hr
72hr
50nM
100nM
MDA-MB-231 N2 siControl
MDA-MB-231 N2 siNELFA
I
CTGF
Percent mRNA expression
150
100
50
0
48hr
72hr
48hr
72hr
50nM
100nM
MDA-MB-231 N2 siControl
MDA-MB-231 N2 siYAP
E
ANKRD1
Percent mRNA expression
150
100
50
0
48hr
72hr
48hr
72hr
50nM
100nM
MDA-MB-231 N2 siControl
MDA-MB-231 N2 siNELFA
J
ANKRD1
Percent mRNA expression
150
100
50
0
48hr
72hr
48hr
72hr
50nM
100nM
MDA-MB-231 N2 siControl
MDA-MB-231 N2 siYAP

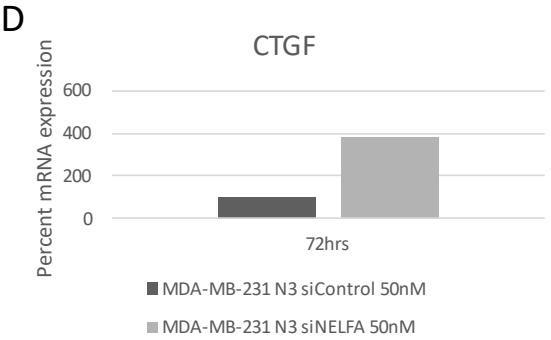
Figure 10: Percent mRNA expression of NELFA and YAP (N2). *NELFA (A), YAP (B) and YAP targets Cyr61 (C), CTGF (D), ANKRD1 (E) in NELFA Knockdown MDA-MB-231 and YAP (F), NELFA (G) and YAP targets Cyr61 (H), CTGF (I), ANKRD1 (J) expression in YAP knockdown MDA-MB-231 with 50nM and 100nM of siRNA concentration at two 48hrs and 72hrs time point*

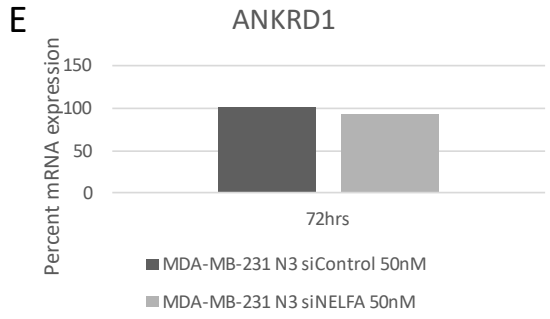
After the first round of knock down confirmation we repeated the same set of parameters for the second round of transfection for an additional validation. The second round of transfection also showed similar results with NELFA and YAP knock down depicted in percent mRNA expression (Figure 10). YAP targets Cyr61, CTGF and ANKRD1 also downregulated with siYAP and were upregulated siNELFA at different percentages.

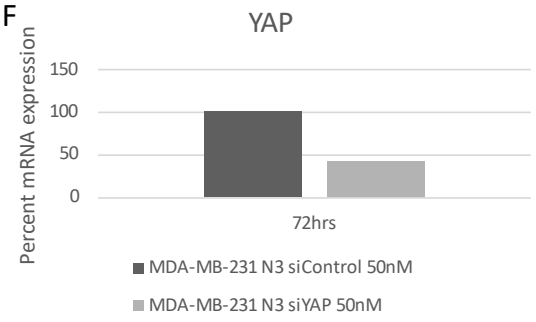

A
NELFA
Percent mRNA expression
150
100
50
0
72hrs
MDA-MB-231 N3 siControl 50nM
MDA-MB-231 N3 siNELFA 50nM

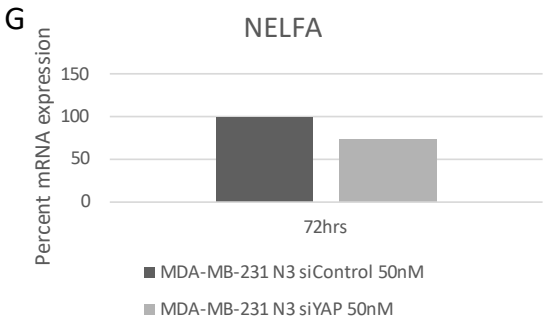

B
YAP
Percent mRNA expression
106
104
102
100
98
72hrs
MDA-MB-231 N3 siControl 50nM
MDA-MB-231 N3 siNELFA 50nM

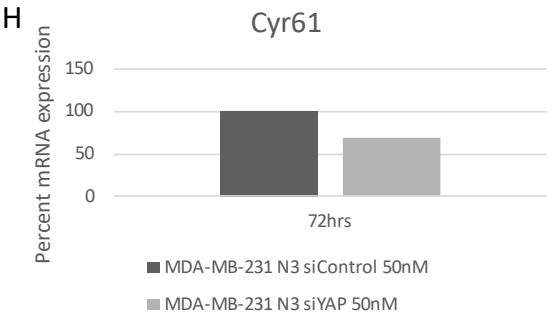

C
Cyr61
Percent mRNA expression
300
200
100
0
72hrs
MDA-MB-231 N3 siControl 50nM
MDA-MB-231 N3 siNELFA 50nM

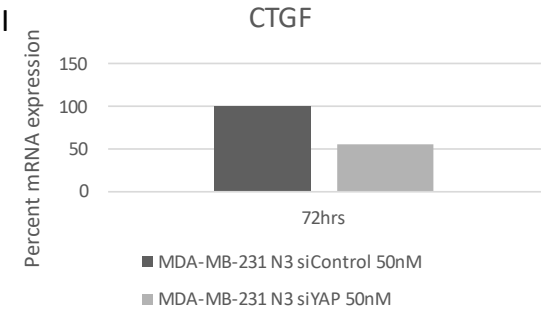

D
CTGF
Percent mRNA expression
600
400
200
0
72hrs
MDA-MB-231 N3 siControl 50nM
MDA-MB-231 N3 siNELFA 50nM


E
ANKRD1
Percent mRNA expression
150
100
50
0
72hrs
MDA-MB-231 N3 siControl 50nM
MDA-MB-231 N3 siNELFA 50nM


F
YAP
Percent mRNA expression
150
100
50
0
72hrs
MDA-MB-231 N3 siControl 50nM
MDA-MB-231 N3 siYAP 50nM


G
NELFA
Percent mRNA expression
150
100
50
0
72hrs
MDA-MB-231 N3 siControl 50nM
MDA-MB-231 N3 siYAP 50nM


H
Cyr61
Percent mRNA expression
150
100
50
0
72hrs
MDA-MB-231 N3 siControl 50nM
MDA-MB-231 N3 siYAP 50nM


I
CTGF
Percent mRNA expression
150
100
50
0
72hrs
MDA-MB-231 N3 siControl 50nM
MDA-MB-231 N3 siYAP 50nM

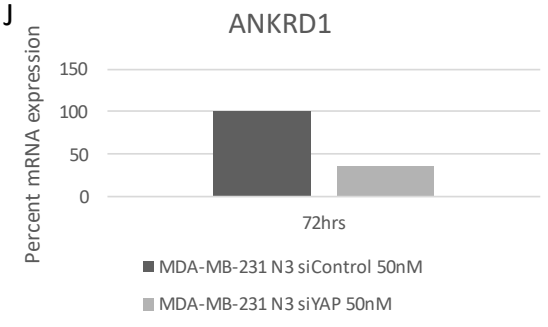

J
ANKRD1
Percent mRNA expression
150
100
50
0
72hrs
MDA-MB-231 N3 siControl 50nM
MDA-MB-231 N3 siYAP 50nM

Figure 11: Percent mRNA expression of NELFA and YAP (N3). *NELFA (A), YAP (B) and YAP targets Cyr61 (C), CTGF (D), ANKRD1 (E) in NELFA Knockdown MDA-MB-231 and YAP (F), NELFA (G) and YAP targets Cyr61 (H), CTGF (I), ANKRD1 (J) expression in YAP knockdown MDA-MB-231 with 50nM of siRNA concentration at 72hrs time point*

With these two rounds of transfection we observed that 50nM and 72hrs has worked across both the sets for NELFA and YAP and also for YAP targets that we were looking at and hence decided to go with 50nM and 72hrs for our third round of transfection. YAP, NELFA and YAP targets – Cyr61, CTGF and ANKRD1 were successfully knocked down more than 50% (Figure 11). We also co-transfected the cells with siYAP+ siNELFA to observe the trend. Co- transfection was also repeated thrice (Figure 12,13,14)

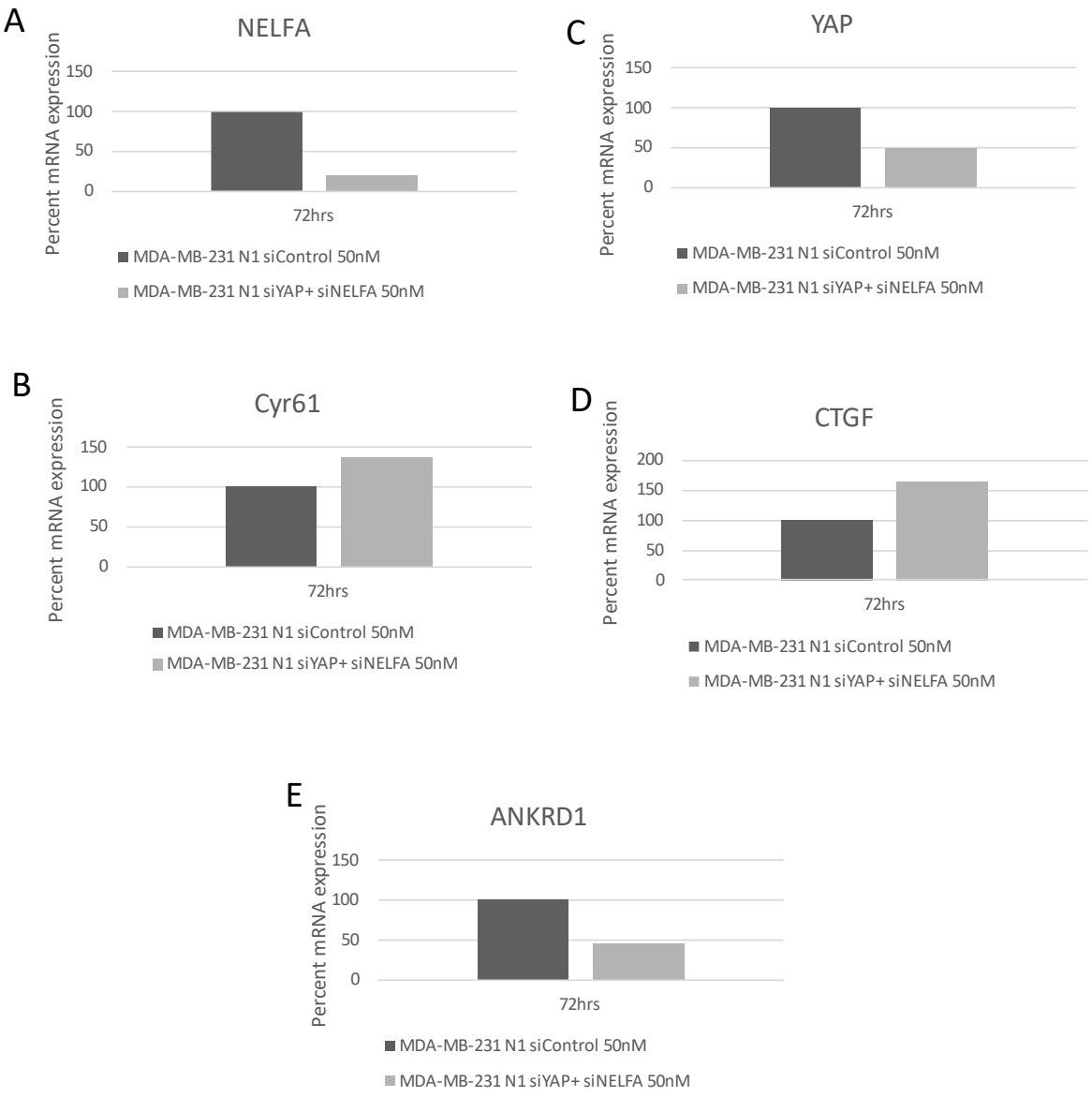


Figure 12: Percent mRNA expression of genes after co-transfection (N1) *NELFA (A), Cyr61 (B), YAP (C), CTGF (D) and ANKRD1 (E) in co-transfected MDA-MB-231 with 50nM of siRNA concentration at 72hrs time point*

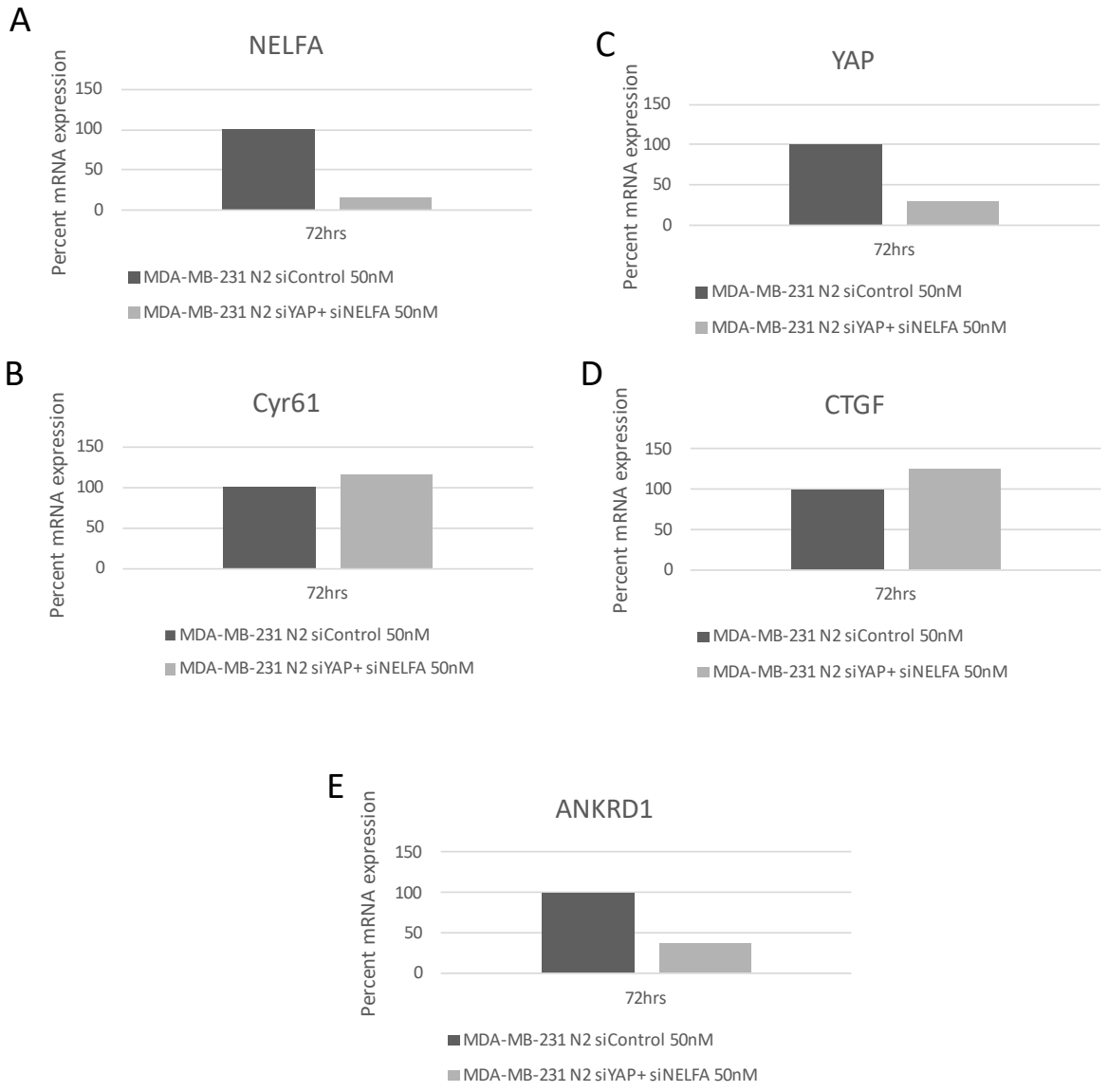


Figure 13: Percent mRNA expression of genes after co-transfection (N2). *NELFA (A), Cyr61 (B), YAP (C), CTGF (D) and ANKRD1 (E) in co-transfected MDA-MB-231 with 50nM of siRNA concentration at 72hrs time point*

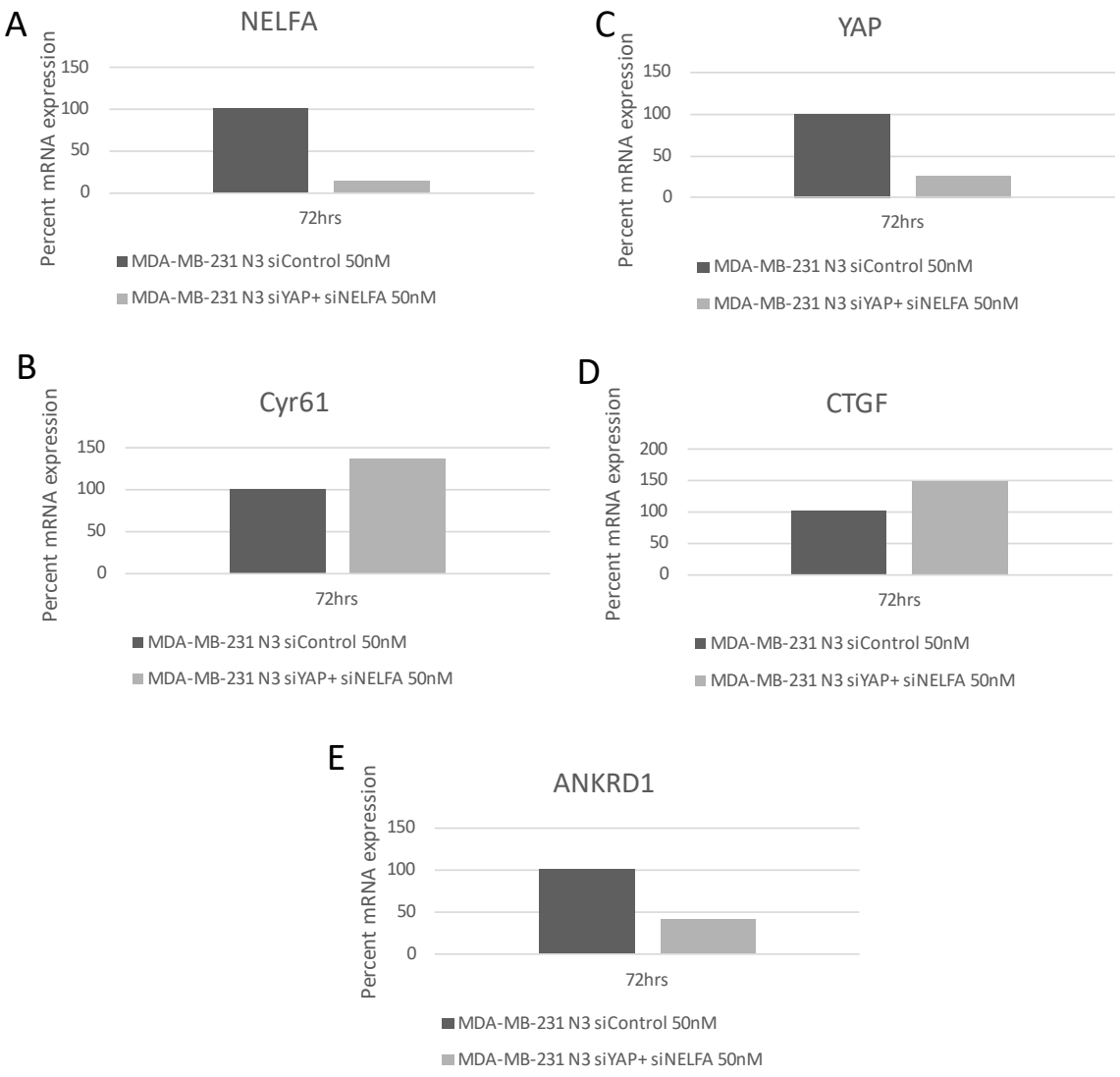


Figure 14: Percent mRNA expression of genes after co-transfection (N3). *NELFA (A), Cyr61 (B), YAP (C), CTGF (D) and ANKRD1 (E) in co-transfected MDA-MB-231 with 50nM of siRNA concentration at 72hrs time point*

2.5 RNA seq analysis

Step 1: RNA-Seq and Read Processing (Done by Medgenome)

RNA sequencing (RNA-seq) was performed on 12 samples across four experimental conditions: siControl, siNELFA, siYAP, and siNELFA+siYAP, with three biological replicates per condition. Sequencing reads underwent quality control using FastQC, adapter trimming with fastq-mcf (v1.05) and Cutadapt (v2.5), and contamination removal (non-polyA RNAs, mitochondrial, and ribosomal RNA) via Bowtie2. Paired-end reads were aligned to the human reference genome (GRCh38/hg38) using STAR (v2.7.3a). Gene expression levels were quantified using feature Counts (v2.0.0) for raw read counts and Cufflinks (v2.2.1) for FPKM normalization. Differential gene expression (DGE) analysis was performed using the DESeq2 package in R.

Step 2: FPKM Analysis and Target Gene Identification

Average FPKM values were calculated for each condition, and genes with distinct expression patterns in siNELFA, siYAP, and siNELFA+siYAP compared to siControl were identified. CTGF, CYR61, YAP, and NELFA emerged as key genes based on their expression trends, which aligned with prior qPCR validation. These genes displayed expected regulatory interactions, where NELFA suppression led to specific transcriptional shifts, corroborating its role in modulating YAP target genes. Data visualization was performed using ggplot2 in R.

Step 3: Differential Expression and Overlap Analysis

Differentially expressed genes (DEGs) were identified using DESeq2, with a significance threshold of p-value & lt; 0.05 and an absolute log2 fold change (|log2FC|) ≥ 1. DEGs were identified for siNELFA_vs_siControl, siYAP_vs_siControl, and siNELFA+siYAP_vs_siControl comparisons. Gene overlap analysis was conducted to examine co-regulation patterns, classifying DEGs into four groups based on upregulation or downregulation across conditions. InteractiVenn (https://www.interactivenn.net/) was used to generate Venn diagrams, while heatmap and dplyr facilitated clustering, data sub setting and visualization. This analysis reinforced the regulatory role of NELFA in YAP-mediated transcription and highlighted gene subsets involved in cellular proliferation and transcriptional control.

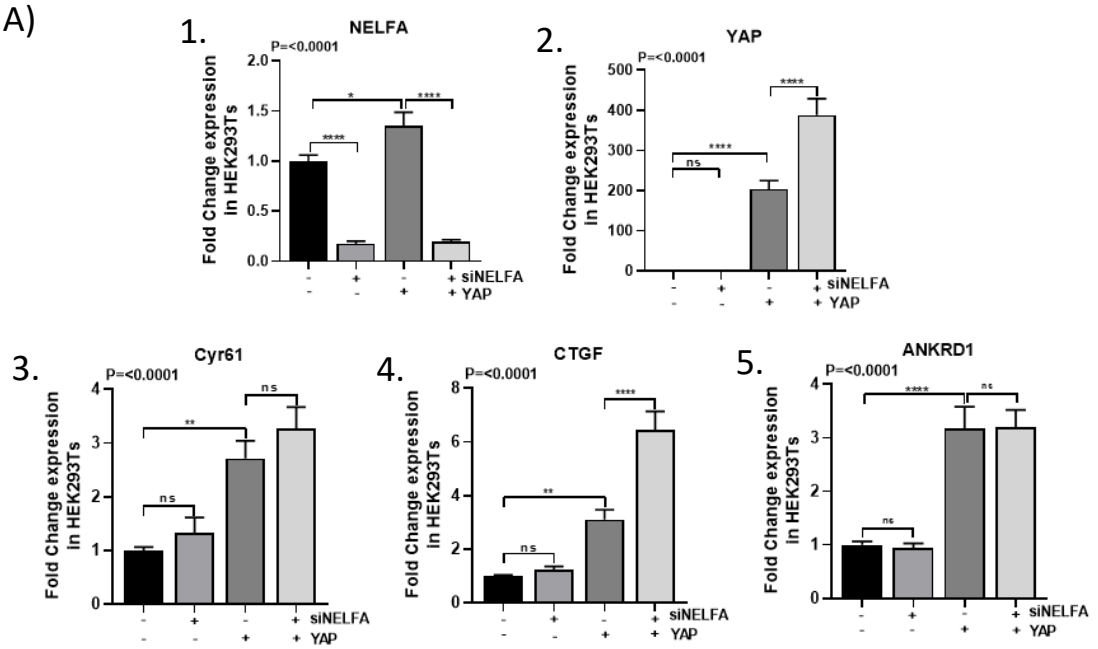
Chapter 3

Results

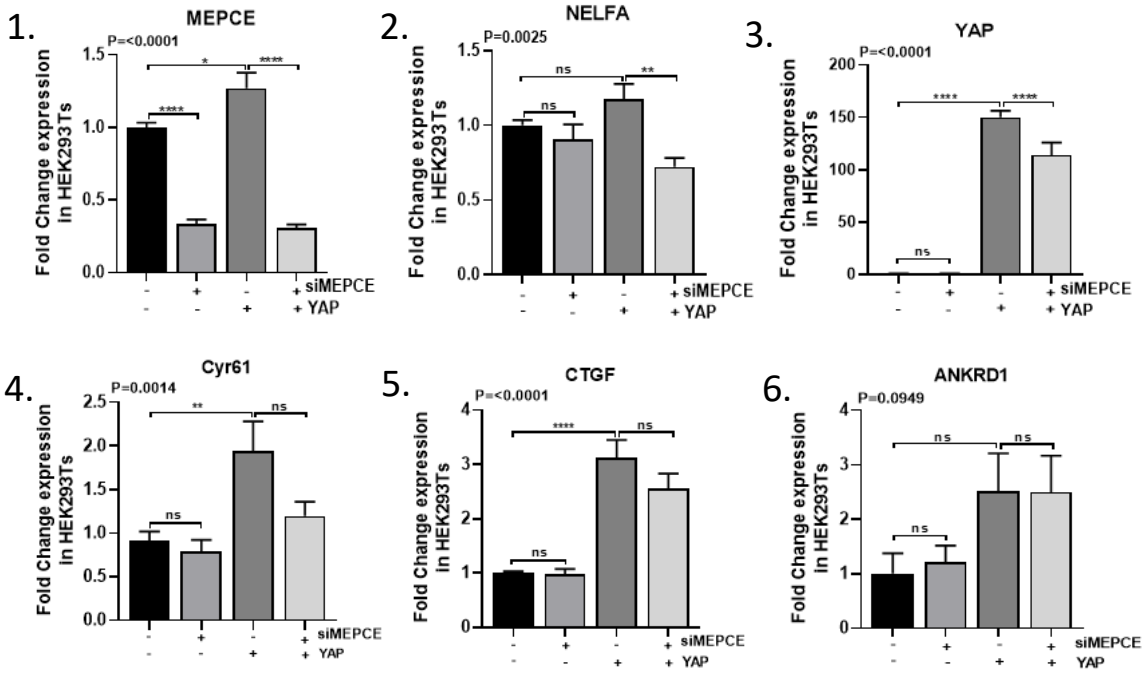
3.1 Co- regulation of YAP targets by YAP and NELFA in HEK293Ts

To investigate the role and impact of the PPP subcomplex subunits NELFA and 7SK snRNP on YAP- mediated transcription in mammalian cells, YAP was expressed alongside siRNA-mediated depletion of NELFA and key 7SK snRNP components, MEPCE and HEXIMs. NELFA was knocked down and YAP was induced in HEK293T by a previous lab member. The target gene expression data generated by RT-PCR was analysed by me, where we observed that NELFA was successfully knocked down in HEK293T cell line (Figure 15, A-1). NELFA knockdown had no significant effect on YAP and its targets - Cyr61, CTGF and ANKRD1. Depleting NELFA alongside over-expression of YAP significantly increased the expression of YAP target genes, Cyr61, CTGF, and ANKRD1 beyond the levels observed with YAP expression alone (Figure 15, A-3,4,5). Knockdown of 7SKsnRNP subunit MePCE was 70-80% (Figure 15, B) and had no significant change on YAP target gene expression whereas it also had no effect of NELFA and YAP mRNA expression. Similarly, HEXIM1 had a 50-60% of knockdown whereas HEXIM2 had almost 90% (Figure 15, C-1,2). Again, no significant difference was observed either in the YAP target gene expression or in siNELFA and YAP O/E individually. Depleting components of the 7SK snRNP complex either had no effect on the expression of YAP target genes or led to a slight decrease in their expression.

A)


1.
NELFA
P=<0.0001
Fold Change expression in HEK293Ts
2.0
1.5
1.0
0.5
0.0
*
****
***
+ siNELFA
+ YAP
2.
YAP
P=<0.0001
500
400
300
200
100
0
****
****
ns
3.
Cyr61
P=<0.0001
4
3
2
1
0
**
ns
ns
4.
CTGF
P=<0.0001
8
6
4
2
0
****
**
ns
5.
ANKRD1
P=<0.0001
4
3
2
1
0
****
ns
ns

B)


1.
MEPCE
P=<0.0001
Fold Change expression in HEK293Ts
1.5
1.0
0.5
0.0
*
****
****
+ siMEPCE
+ YAP
2.
NELFA
P=0.0025
1.5
1.0
0.5
0.0
ns
**
ns
3.
YAP
P=<0.0001
200
150
100
50
0
****
****
ns
4.
Cyr61
P=0.0014
2.5
2.0
1.5
1.0
0.5
0.0
**
ns
ns
5.
CTGF
P=<0.0001
4
3
2
1
0
****
ns
ns
6.
ANKRD1
P=0.0949
4
3
2
1
0
ns
ns
ns

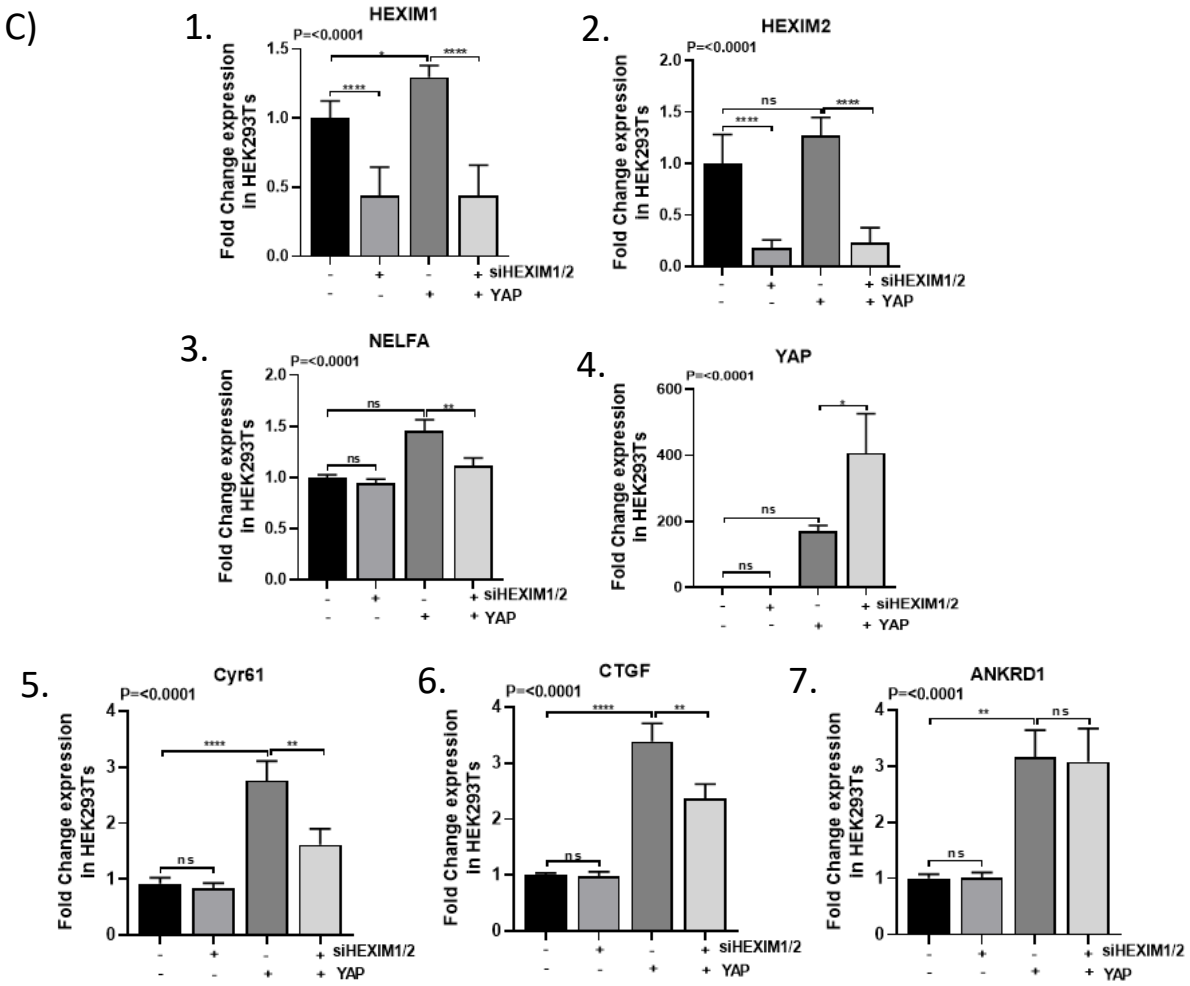


Figure 15: Co-regulation of YAP targets by YAP and NELFA. *A) The impact of NELFA knockdown, both alone and in combination with constitutively active YAP, on YAP target genes was analysed using RT-qPCR on total RNA extracted from 293T cells. It represents fold change in NELFA (1), YAP (2) and its targets gene expression-Cyr61 (3), CTGF (4) and ANKRD1 (5). B) represents fold change for YAP targets expression (4)-(6) in HEK293Ts along with MePCE (1), NELFA (2) and YAP (3) under YAP overexpression and MePCE knockdown. C) Effect of HEXIM knockdown (1)– (2) represented in fold change values for YAP targets expression (5)-(6) in HEK293Ts. Error bars depict SEM where Statistical significance is calculated using one way ANOVA test, with p value < 0.0001. p values are denoted above the bar graphs for respective comparisons.*

3.2 Comparative gene expression of NELFA, YAP and YAP target

Based on our clinical patient data where we observed a significant association between high YAP and low NELFA expression with poor prognosis specifically in TNBC subtype. Hence, we chose MDA-MB-231 cell lines to work on as a TNBC breast cancer cell line model. It's a commonly used TNBC cell line to model late-stage breast cancer in which YAP is also endogenously active. We confirmed the presence of our gene of interest – NELFA, YAP and Cyr61 by looking at their mRNA expression level using RT-qPCR (Figure 2 A) and protein level expression using western blotting (Figure 2 B) in MDA-MB-231 as compared to that of HEK293T.

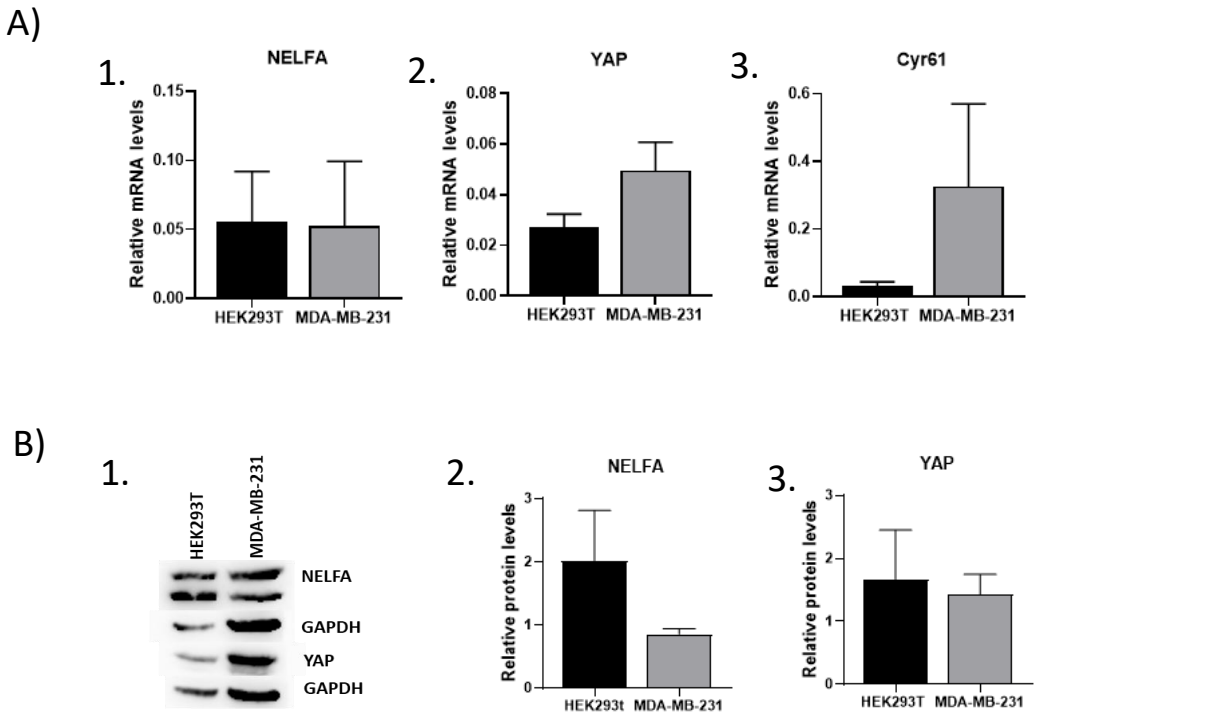


Figure 16: Comparative gene expression of NELFA, YAP and YAP target*. A) RNA was extracted from HEK293Ts and MDA-MB-231. Using the extracted RNA RT-qPCR was done for examination of the above genes- NELFA (1), YAP (2) and Cyr61 (3). B) Protein lysates were used to run a SDS page and figure B-1 is the blot for desired proteins.*

3.3 Silencing NELFA affects YAP targets at varying degrees in MDA-MB-231

NELFA and YAP siRNA mediated knocked down was standardized for concentration and time. The concentration and the time points were chosen on the basis of the percent knockdown of NELFA and YAP. Samples with 70-80% knockdown of NELFA and YAP were shortlisted and compiled in GraphPad prism. One way ANOVA test with multiple comparison filter was used for the analysis. NELFA and YAP both were knocked down with a significant value. (Figure 18. A, B) When NELFA was silenced in MDA-MB-231 YAP target Cyr61 and CTGF were significantly upregulated with no significant upregulation was observed in ANKRD1 gene expression. Upon YAP knockdown the targets were also downregulated confirming their dependency on YAP expression. Knock down of NELFA and YAP together (Co-transfection) did not completely downregulate the targets neither were they upregulated compared to their individual knockdown effects. (Figure 18, C, D, E). CTGF had more effect with NELFA and YAP knockdown. Cyr61 had a moderate effect whereas ANKRD1 had very less or no effect with YAP and NELFA knockdown. We also validated our findings at protein level by performing western blot. The parameters for knocking down NELFA and YAP were 50nM and 72hrs as per the previous conclusion. We observed similar findings with respect to NELFA and YAP knockdown in MDA-MB-231 cell lines. NELFA and YAP bands were not seen with their knockdown. YAP expression in co-transfection was slightly visible since it wasn't 100% knocked down which matched our RT-qPCR results (Figure 18 A). 2Ns were analysed in ImageJ using intensity of the bands and data was compiled in GraphPad prism. In the representative image (Figure 18. B)

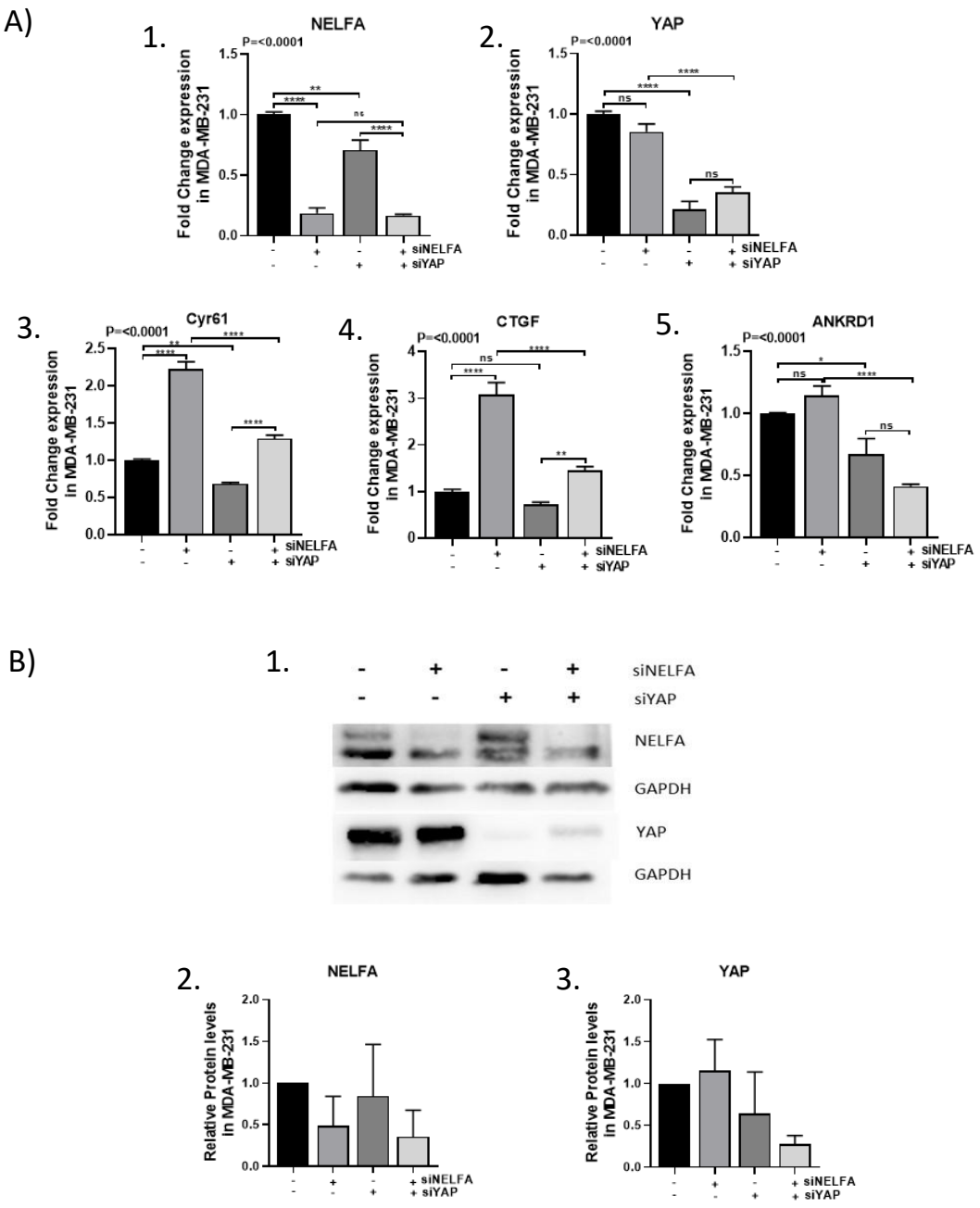


Figure 17: Silencing NELFA affects YAP targets at varying degrees in MDA-MB-231. *NELFA (A-1), YAP (A-2), Cyr61 (A-3), CTGF (A-4) and ANKRD1 (A-5) fold change expression upon knockdown of NELFA and YAP alone and upon co-transfection of NELFA and YAP. Post mRNA level validation (B-1) represents a western blot of proteins. The bands were analysed in ImageJ and NELFA (B-2), YAP (B-3) were plotted.*

3.4 siRNA-Induced Transcriptional Changes: Concordance Between RT-qPCR and RNA-Seq Data

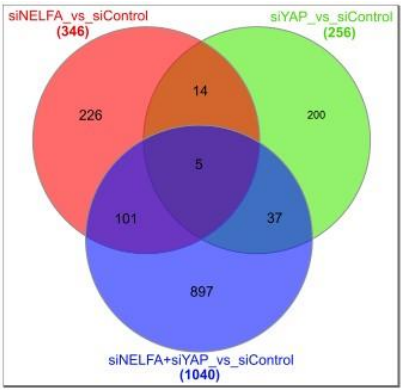
To validate the transcriptional effects observed in our RT-qPCR analysis done for siRNA mediated knockdown of NELFA, YAP and NELFA+YAP we performed RNA sequencing on the same siRNA-treated MDA-MB-231 samples. The RNA-seq data showed a high degree of concordance with our RT-qPCR results, reinforcing the reliability of our initial observations.

RNA-seq confirmed effective knockdown of NELFA and YAP, consistent with RT-qPCR findings. The Expression Changes in YAP Target Genes, CYR61 and CTGF were upregulated upon NELFA knockdown and downregulated upon YAP knockdown, mirroring the RT-qPCR results. ANKRD1 expression remained unchanged with NELFA knockdown, which was also reflected in the RNA-seq data because it was not listed in the data sent by Sequencing company. To further analyze the impact of NELFA and YAP knockdown on the transcriptome, we generated Expression plots for key genes, demonstrating consistent trends across both methods (Figure 18 C- 1,2,3,4). Heatmaps to visualize global expression changes upon NELFA and YAP perturbation (Figure 18 B-1,2,3). Venn diagrams to identify overlapping and unique differentially expressed genes across conditions, highlighting shared regulatory pathways (Figure 18 A-1,2)

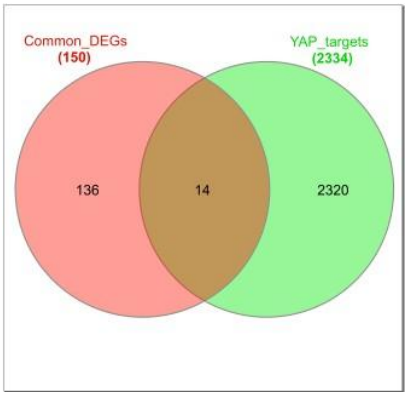
The strong agreement between our previous data and RNA-seq results underscores the robustness of our findings. The transcriptomic analysis further refines our understanding of NELFA's role in modulating YAP target genes and suggests potential mechanisms through which promoter-proximal pausing influences YAP-driven transcriptional programs.

A)

1.

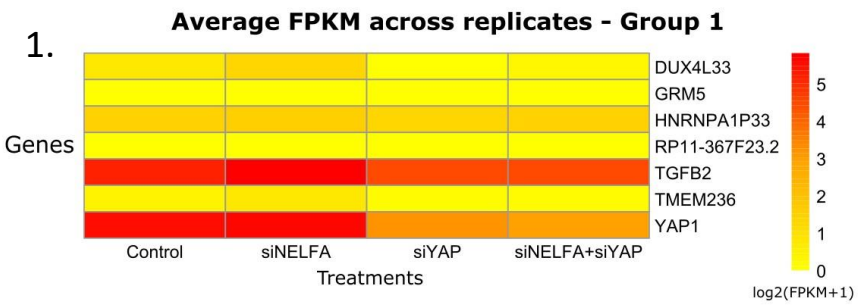

siNELFA_vs_siControl
(346)
siYAP_vs_siControl
(256)
siNELFA+siYAP_vs_siControl
(1040)
226
14
200
5
101
37
897

2.

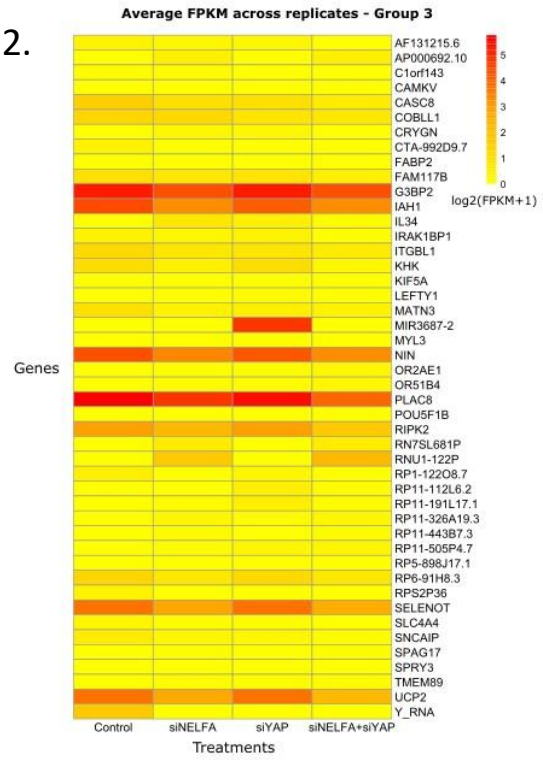

Common_DEGs
(150)
YAP_targets
(2334)
136
14
2320

B)

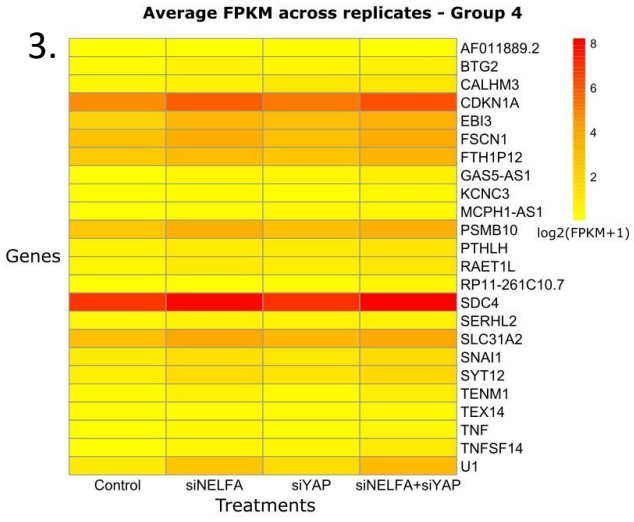
1.


Average FPKM across replicates - Group 1
Genes
DUX4L33
GRM5
HNRNPA1P33
RP11-367F23.2
TGFB2
TMEM236
YAP1
Control
siNELFA
siYAP
siNELFA+siYAP
Treatments
log2(FPKM+1)

2.


Average FPKM across replicates - Group 3
Genes
Control
siNELFA
siYAP
siNELFA+siYAP
Treatments
log2(FPKM+1)

3.


Average FPKM across replicates - Group 4
Genes
Control
siNELFA
siYAP
siNELFA+siYAP
Treatments
log2(FPKM+1)

C)

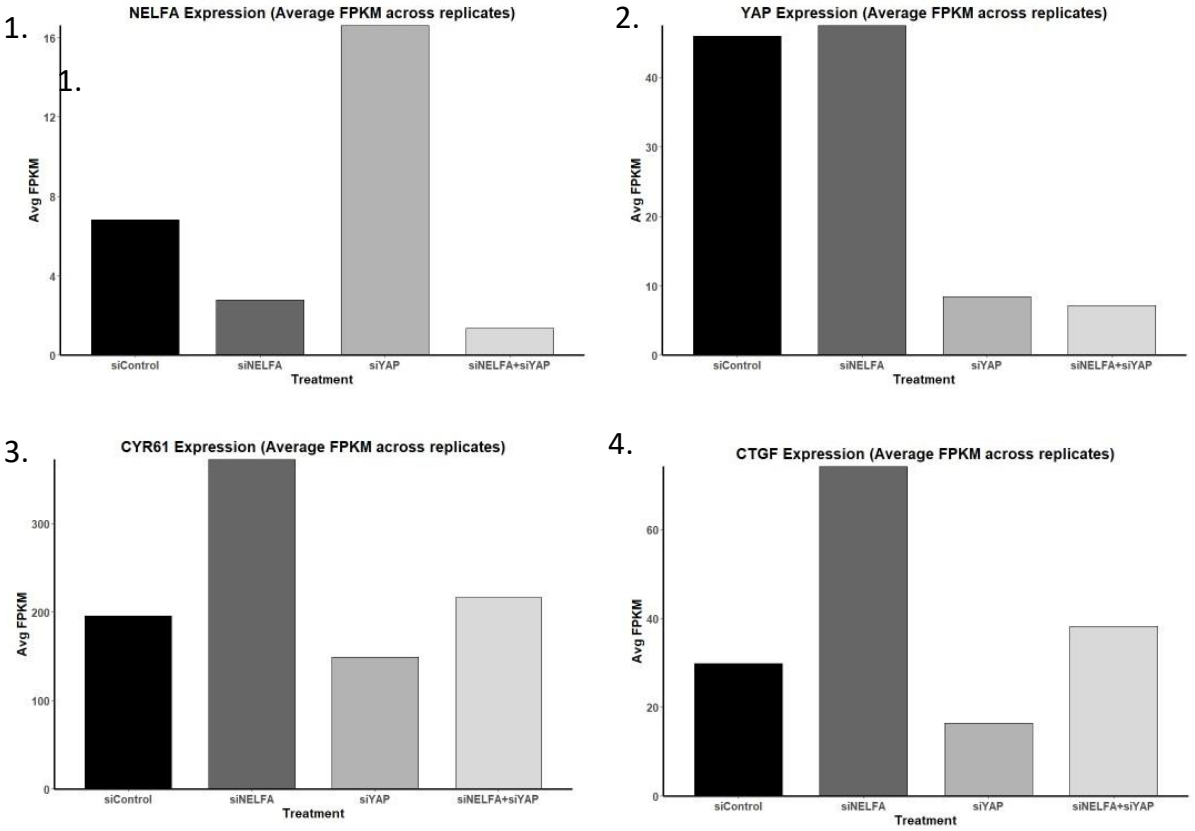


Figure 18: siRNA-Induced Transcriptional Changes: Concordance Between RT-qPCR and RNA-Seq Data. *Venn diagram (A-1,2) to check the overlap among various genes those get affected. Three different Heatmaps to see how many genes follow the trend (B-1,2,3). FPKM gene expression of NELFA (C-1), YAP (C-2) and YAP targets Cyr61 (C-3) and CTGF (C-4) upon NELFA, YAP and NELFA+YAP knockdown.*

Chapter 4

Discussion

Our findings reveal that the transcriptional regulatory complex subunit NELFA plays a dynamic and context-dependent role in transcriptional regulation. NELFA depletion enhances the expression of YAP target genes such as Cyr61 and CTGF upon YAP induction but does not affect ANKRD1. In contrast, silencing 7SK snRNP components MePCE and HEXIM1/2 did not alter YAP target gene expression in HEK293T cells. This suggests that promoter-proximal pausing (PPP) modulates YAP-driven transcription. While PPP has been previously implicated in cancer(Rahl et al., 2010). our study provides direct evidence linking PPP to YAP-mediated transcriptional activity in the MDA-MB-231 breast cancer cell line, which exhibits endogenous active YAP. NELFA depletion in this system led to similar trends in YAP target gene expression, reinforcing its role in YAP-driven transcriptional regulation. Given that these YAP target genes are involved in cell proliferation and tumour growth (Kulkarni et al., 2018), our findings suggest a broader oncogenic impact of PPP perturbation.

Previous studies have established the role of the Negative Elongation Factor (NELF) complex in regulating gene expression in breast cancer cells(Sun & Li, 2010).NELF is essential for maintaining the expression of key cell cycle regulators, thereby supporting cell proliferation and tumour progression. Our RNA-seq analysis in MDA-MB-231 cells revealed that while NELFA depletion influenced YAP target gene expression, the effect was context-dependent, as not all YAP-regulated genes were affected.

Our findings highlight how the interplay between YAP and PPP disruption uncovers a novel physiological function of PPP, where YAP target genes are regulated independently and at varying levels. In contrast, 7SK snRNP, previously implicated in gastric cancer through P-TEFb-mediated transcriptional deregulation (Cheng et al., 2012) exhibited no significant effect on YAP target gene expression upon depletion, suggesting a limited role in YAP-mediated transcriptional regulation.

In Drosophila, studies on the Hippo pathway and Yorkie (Yki) demonstrated that PPP restricts Yki's oncogenic potential. While Yki promotes hyperplastic growth, neoplastic

transformation occurs only when 7SK snRNP or NELF is depleted, leading to enhanced Yki activity(Nagarkar et al., 2020). Among these findings NELF study align with our observations on NELF depletion in mammalian systems.

Furthermore, the NELF complex has been shown to drive breast cancer progression and metastasis by facilitating epithelial-mesenchymal transition (EMT) and mammary stem cell programs. This occurs through NELF's interaction with SLUG and KAT2B, which orchestrates gene expression changes essential for metastasis across multiple breast cancer subtypes (Zhang et al., 2023). Analysis of an Indian breast cancer patient cohort further supports the oncogenic role of NELFA, with poorer survival outcomes observed in overall survival (OS) plots. Interestingly, in disease-free survival (DFS) analyses, NELFA appears to exhibit tumour-suppressive properties, suggesting a context-dependent dual role. Integrating previous studies on Indian breast cancer dataset, we propose that while NELF generally displays oncogenic properties, it may act as a tumour suppressor in specific contexts.

Our findings underscore the context-specific tumour-suppressive roles of NELF, which may restrain YAP-driven tumorigenesis. The specificity of YAP function suggests cooperative mechanisms in cancer, warranting further exploration in mouse models to elucidate the conserved role of NELF in YAP-driven tumorigenesis. Examination of TCGA data would also shed light upon our findings.

References

Boopathy, G. T. K., & Hong, W. (2019). Role of Hippo Pathway-YAP/TAZ Signaling in Angiogenesis. *Frontiers in Cell and Developmental Biology*, *7*, 49. https://doi.org/10.3389/fcell.2019.00049

Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R. L., Soerjomataram, I., & Jemal, A. (2024). Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, *74*(3), 229–263. https://doi.org/10.3322/caac.21834

Calses, P. C., Crawford, J. J., Lill, J. R., & Dey, A. (2019). Hippo Pathway in Cancer: Aberrant Regulation and Therapeutic Opportunities. *Trends in Cancer*, *5*(5), 297–307. https://doi.org/10.1016/j.trecan.2019.04.001

Cheng, Y., Jin, Z., Agarwal, R., Ma, K., Yang, J., Ibrahim, S., Olaru, A. V, David, S., Ashktorab, H., Smoot, D. T., Duncan, M. D., Hutcheon, D. F., Abraham, J. M., Meltzer, S. J., & Mori, Y. (2012). LARP7 is a potential tumor suppressor gene in gastric cancer. *Laboratory Investigation*, *92*(7), 1013–1019. https://doi.org/10.1038/labinvest.2012.59

Core, L., & Adelman, K. (2019). Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation. *Genes & Development*, *33*(15–16), 960–982. https://doi.org/10.1101/gad.325142.119

Foulkes, W. D., Smith, I. E., & Reis-Filho, J. S. (2010). Triple-negative breast cancer. *The New England Journal of Medicine*, *363*(20), 1938–1948. https://doi.org/10.1056/NEJMra1001389

Galli, G. G., Carrara, M., Yuan, W.-C., Valdes-Quezada, C., Gurung, B., Pepe-Mooney, B., Zhang, T., Geeven, G., Gray, N. S., de Laat, W., Calogero, R. A., & Camargo, F. D. (2015). YAP Drives Growth by Controlling Transcriptional Pause Release from Dynamic Enhancers. *Molecular Cell*, *60*(2), 328–337. https://doi.org/10.1016/j.molcel.2015.09.001

Gibson, S. V, Roozitalab, R. M., Allen, M. D., Jones, J. L., Carter, E. P., & Grose, R. P. (2023). Everybody needs good neighbours: the progressive DCIS microenvironment. *Trends in Cancer*, *9*(4), 326–338. https://doi.org/10.1016/j.trecan.2023.01.002

Gilmour, D. S., & Lis, J. T. (1986). RNA polymerase II interacts with the promoter region of the noninduced hsp70 gene in Drosophila melanogaster cells. *Molecular and Cellular Biology*, *6*(11), 3984–3989. https://doi.org/10.1128/mcb.6.11.3984-3989.1986

Groth, C., Vaid, P., Khatpe, A., Prashali, N., Ahiya, A., Andrejeva, D., Chakladar, M., Nagarkar, S., Paul, R., Kelkar, D., Eichenlaub, T., Herranz, H., Sridhar, T., Cohen, S. M., & Shashidhara, L. (2020). Genome-Wide Screen for Context-Dependent Tumor

Suppressors Identified Using in Vivo Models for Neoplasia in *Drosophila*. *G3 Genes|Genomes|Genetics*, *10*(9), 2999–3008. https://doi.org/10.1534/g3.120.401545

Ibar, C., & Irvine, K. D. (2020). Integration of Hippo-YAP Signaling with Metabolism. *Developmental Cell*, *54*(2), 256–267. https://doi.org/10.1016/j.devcel.2020.06.025

Inic, Z., Zegarac, M., Inic, M., Markovic, I., Kozomara, Z., Djurisic, I., Inic, I., Pupic, G., & Jancic, S. (2014). Difference between Luminal A and Luminal B Subtypes According to Ki-67, Tumor Size, and Progesterone Receptor Negativity Providing Prognostic Information. *Clinical Medicine Insights. Oncology*, *8*, 107–111. https://doi.org/10.4137/CMO.S18006

Jonkers, I., & Lis, J. T. (2015). Getting up to speed with transcription elongation by RNA polymerase II. *Nature Reviews. Molecular Cell Biology*, *16*(3), 167–177. https://doi.org/10.1038/nrm3953

Kassam, F., Enright, K., Dent, R., Dranitsaris, G., Myers, J., Flynn, C., Fralick, M., Kumar, R., & Clemons, M. (2009). Survival outcomes for patients with metastatic triple-negative breast cancer: implications for clinical practice and trial design. *Clinical Breast Cancer*, *9*(1), 29–33. https://doi.org/10.3816/CBC.2009.n.005

Koo, J. H., & Guan, K.-L. (2018). Interplay between YAP/TAZ and Metabolism. *Cell Metabolism*, *28*(2), 196–206. https://doi.org/10.1016/j.cmet.2018.07.010

Kulkarni, M., Tan, T. Z., Syed Sulaiman, N. B., Lamar, J. M., Bansal, P., Cui, J., Qiao, Y., & Ito, Y. (2018). RUNX1 and RUNX3 protect against YAP-mediated EMT, stem-ness and shorter survival outcomes in breast cancer. *Oncotarget*, *9*(18), 14175–14192. https://doi.org/10.18632/oncotarget.24419

Kwak, H., Fuda, N. J., Core, L. J., & Lis, J. T. (2013). Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science (New York, N.Y.)*, *339*(6122), 950–953. https://doi.org/10.1126/science.1229386

Lamar, J. M., Stern, P., Liu, H., Schindler, J. W., Jiang, Z.-G., & Hynes, R. O. (2012). The Hippo pathway target, YAP, promotes metastasis through its TEAD-interaction domain. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(37), E2441-50. https://doi.org/10.1073/pnas.1212021109

Lee, J. (2023). Current Treatment Landscape for Early Triple-Negative Breast Cancer (TNBC). *Journal of Clinical Medicine*, *12*(4). https://doi.org/10.3390/jcm12041524

Muse, G. W., Gilchrist, D. A., Nechaev, S., Shah, R., Parker, J. S., Grissom, S. F., Zeitlinger, J., & Adelman, K. (2007). RNA polymerase is poised for activation across the genome. *Nature Genetics*, *39*(12), 1507–1511. https://doi.org/10.1038/ng.2007.21

Nagarkar, S., Wasnik, R., Govada, P., Cohen, S., & Shashidhara, L. S. (2020). Promoter Proximal Pausing Limits Tumorous Growth Induced by the Yki Transcription Factor in *Drosophila*. *Genetics*, *216*(1), 67–77. https://doi.org/10.1534/genetics.120.303419

Nguyen, V. T., Kiss, T., Michels, A. A., & Bensaude, O. (2001). 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. *Nature*, *414*(6861), 322–325. https://doi.org/10.1038/35104581

Noe Gonzalez, M., Blears, D., & Svejstrup, J. Q. (2021). Causes and consequences of RNA polymerase II stalling during transcript elongation. *Nature Reviews. Molecular Cell Biology*, *22*(1), 3–21. https://doi.org/10.1038/s41580-020-00308-8

Obeagu, E. I., & Obeagu, G. U. (2024). Breast cancer: A review of risk factors and diagnosis. *Medicine*, *103*(3), e36905. https://doi.org/10.1097/MD.0000000000036905

Orrantia-Borunda, E., Anchondo-Nuñez, P., Acuña-Aguilar, L. E., Gómez-Valles, F. O., & Ramírez-Valdespino, C. A. (2022). Subtypes of Breast Cancer. In *Breast Cancer* (pp. 31–42). Exon Publications. https://doi.org/10.36255/exon-publications-breast-cancer-subtypes

Peterlin, B. M., Brogie, J. E., & Price, D. H. (2012). 7SK snRNA: a noncoding RNA that plays a major role in regulating eukaryotic transcription. *Wiley Interdisciplinary Reviews. RNA*, *3*(1), 92–103. https://doi.org/10.1002/wrna.106

Rahl, P. B., Lin, C. Y., Seila, A. C., Flynn, R. A., McCuine, S., Burge, C. B., Sharp, P. A., & Young, R. A. (2010). c-Myc Regulates Transcriptional Pause Release. *Cell*, *141*(3), 432–445. https://doi.org/10.1016/j.cell.2010.03.030

Rodríguez-Molina, J. B., West, S., & Passmore, L. A. (2023). Knowing when to stop: Transcription termination on protein-coding genes by eukaryotic RNAPII. *Molecular Cell*, *83*(3), 404–415. https://doi.org/10.1016/j.molcel.2022.12.021

Schwartz, S. M. (2024). Epidemiology of Cancer. *Clinical Chemistry*, *70*(1), 140–149. https://doi.org/10.1093/clinchem/hvad202

Sun, J., & Li, R. (2010). Human Negative Elongation Factor Activates Transcription and Regulates Alternative Transcription Initiation. *Journal of Biological Chemistry*, *285*(9), 6443–6452. https://doi.org/10.1074/jbc.M109.084285

Szulzewsky, F., Holland, E. C., & Vasioukhin, V. (2021). YAP1 and its fusion proteins in cancer initiation, progression and therapeutic resistance. *Developmental Biology*, *475*, 205–221. https://doi.org/10.1016/j.ydbio.2020.12.018

Vaid, P. M., Puntambekar, A. K., Jumle, N. S., Banale, R. A., Ansari, D., Reddy, R. R., Unde, R. R., Namewar, N. P., Kelkar, D. A., Shashidhara, L. S., Koppiker, C. B., & Kulkarni, M. D. (2022). Evaluation of tumor-infiltrating lymphocytes (TILs) in molecular subtypes of an Indian cohort of breast cancer patients. *Diagnostic Pathology*, *17*(1), 91. https://doi.org/10.1186/s13000-022-01271-y

Yamaguchi, Y., Takagi, T., Wada, T., Yano, K., Furuya, A., Sugimoto, S., Hasegawa, J., & Handa, H. (1999). NELF, a Multisubunit Complex Containing RD, Cooperates with DSIF to

Repress RNA Polymerase II Elongation. *Cell*, *97*(1), 41–51. https://doi.org/10.1016/S0092-8674(00)80713-8

Yik, J. H. N., Chen, R., Nishimura, R., Jennings, J. L., Link, A. J., & Zhou, Q. (2003). Inhibition of P-TEFb (CDK9/Cyclin T) Kinase and RNA Polymerase II Transcription by the Coordinated Actions of HEXIM1 and 7SK snRNA. *Molecular Cell*, *12*(4), 971–982. https://doi.org/10.1016/S1097-2765(03)00388-5

Zanconato, F., Cordenonsi, M., & Piccolo, S. (2016). YAP/TAZ at the Roots of Cancer. *Cancer Cell*, *29*(6), 783–803. https://doi.org/10.1016/j.ccell.2016.05.005

Zhang, J., Hu, Z., Chung, H. H., Tian, Y., Lau, K. W., Ser, Z., Lim, Y. T., Sobota, R. M., Leong, H. F., Chen, B. J., Yeo, C. J., Tan, S. Y. X., Kang, J., Tan, D. E. K., Sou, I. F., McClurg, U. L., Lakshmanan, M., Vaiyapuri, T. S., Raju, A., … Tee, W.-W. (2023). Dependency of NELF-E-SLUG-KAT2B epigenetic axis in breast cancer carcinogenesis. *Nature Communications*, *14*(1), 2439. https://doi.org/10.1038/s41467-023-38132-1