UTS UNIVERSITY OF TECHNOLOGY SYDNEY

The impact of the Human papillomavirus type 16 on non-coding RNAs in Head and Neck cancer

by Dayna Grace Sais

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

under the supervision of

Dr Nham Tran
Dr Valerie Gay

University of Technology Sydney
Faculty of Engineering and Information Technology

Mar 2022

Certificate of original authorship

I, Dayna Grace Sais, declare that this thesis is submitted in fulfilment of the requirements for the award of Doctor of Philosophy in the School of Biomedical Engineering, Faculty of Engineering, and Information Technology at the University of Technology Sydney.

This thesis is wholly my work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Signature: Production Note:
Signature removed prior to publication.

Date:11/03/2022

Acknowledgements:

The past few years have been such a unique experience, and I have learnt so much about myself and what it means to be a researcher. The work presented in this thesis is something I am genuinely proud of and was made possible by many individuals' expertise, creativeness, and unwavering support.

Firstly, I would like to give thanks to my supervisors, A/Prof Nham Tran and A/Prof Valerie Gay, who provided me with their academic expertise and encouraged me to explore my own ideas. To my primary Supervisor, A/Prof Nham Tran, who not only taught me all the molecular techniques to complete my work but pushed me in my professional development to become the researcher I am today. He encouraged me to expand my research portfolio, to present my ideas and research at conferences and to collaborate with others. I have learnt many invaluable skills, and I appreciate the immense time and effort he has put into supervising me over the years.
To my co-supervisor, A/Prof Valerie Gay, thank you for guiding me through the PhD experience and always being someone, I could come and talk to.

Within the School of Biomedical Engineering at UTS, I would like to thank the members and past affiliates of the Tran Lab, including Dr Samantha Khoury, Meredith Hill and Fiona Deutsch. All of whom have given me great technical advice, created a collaborative research environment and provided many opportunities to laugh. We have shared many great memories in the lab, becoming true friends.
Similarly, from the School of Life Sciences, members and past affiliates of the Donnelly Lab, including A/Prof Sheila Donnelly, Alison Ricafrente, Susel Myra Loli Quinteros, and Inah Camaya, were equally supportive when sharing their technical knowledge.

From the Graduate School of Biomedical Sciences at Tufts University Boston, I would like to thank the members of the Munger Lab, including Surendra Sharma and Miranda Grace. They provided their expertise in lncRNAs and guided me through my international research experience. Profound thanks to A/Prof Karl Munger for welcoming me to join the team in Boston and providing your continual support throughout my PhD. Working in the Munger Lab will be an unforgettable experience, and I gained so much from my time here.

Outside of the research space, I would like to give a big thanks to all of my family and friends, especially to my parents, Darryl and Nicole. I could not have done this without your support and encouragement. Finally, I would love to thank my partner, Dylan, who has been there from day one of this incredible journey. He has been the strong emotional pillar of support for me; he motivated me and encouraged me to work hard. He even dropped everything to support me and join my international endeavours. The last few years have been unpredictable and challenging, but Dylan was there for me every step of the way; thank you for being there for me. I could not have achieved this without you.

Statement & list of papers and conferences

This thesis originates from collected works featured in (Previous name Dayna Mason):

Research articles:

1. Mason, D., Munger, K. and Tran, N., 2021. The dynamic interactome of microRNAs and the human papillomavirus in Head and Neck cancers. Current Opinion in Virology, 51, pp.87-95

2. Mason, D., Hill, M., Marques, T.M., Carvalho, M.G. and Tran, N., 2020. Developing a virus-microRNA interactome using Cytoscape. *MethodsX*, *7*, p.100700.

3. Mason, D., Zhang, X., Marques, T.M., Rose, B., Khoury, S., Hill, M., Deutsch, F., Lyons, J.G., Gama-Carvalho, M. and Tran, N., 2018. Human papillomavirus 16 E6 modulates the expression of miR-496 in oropharyngeal cancer. *Virology*, *521*, pp.149-157

Conferences:

1. Mason D, Zhang X, Marques TM, Gama-Carvalho M, Tran N, “MiR-496 expression is altered in Oropharyngeal cancers by HPV16 E6” *Combio*, Sydney, 2018

2. Mason D, Zhang X, Marques TM, Rose B, Khoury S, Hill M, Deutsch F, Lyons JG. Gama-Carvalho M, Tran N, "Human papillomavirus 16 E6 modulates the expression of miR-496 in oropharyngeal cancer” *International Papilloma virus conference*, Sydney, 2018

3. Mason D, Zhang X, Marques TM, Gama-Carvalho M, Tran N, "The HPV16 E6 oncogene alters the expression of miR-496 in Oropharyngeal cancers” *Translation cancer research network symposium*, Sydney, 2018

4. Mason D, Tran N, “The HPV16 E6 oncogene alters the expression of miR-496 in Oropharyngeal cancer” *Chris O'Brien Head and Neck Symposium*, Sydney 2019

5. Mason D, Zhang X, Marques TM, Gama-Carvalho M, Tran N, “Uncovering miR-33a and SREBF2 relationship in HPV16 positive Oropharyngeal cancers” *Keystone Symposia: Noncoding RNAs: Mechanisms, function, and therapies*, 2020

6. Mason D, Tran N, “Using a viral interactome to uncover mechanistic links between HPV16 and specific microRNAs in Head and neck cancers” *Lorne Cancer*, Online Conference, 2021

7. Mason D, Tran N, “Uncovering mechanistic links between HPV16 and specific microRNAs in Head and Neck cancers” RNA 2021, *RNA Society*, Online Conference, 2021

COVID-19 impact statement:

Due to the COVID-19 pandemic and enforced restrictions, I experienced two extended periods of suspended experimental work during my candidature. These COVID-19 restrictions were due to NSW-wide lockdowns, the first period being from March 2020 to July 2020 and the second from July 2021 to November 2021. I was also under stricter COVID-19-related restrictions due to high transmission rates in my area of residence and residing with an immunocompromised individual.

During the two COVID-19 lockdowns, it was highly recommended that all research activities be conducted from home. Additionally, laboratory access was restricted to those with highly time-sensitive research and animal works that were unable to be reorganised or postponed. Due to my research not fitting within this category and travel restrictions, I was unable to conduct experiments during this time. This meant that a proportion of the experiments within this thesis were delayed, and particular experiments were not achieved.

These issues particularly affected Chapter 5 and Chapter 7 of this thesis. Due to delays, I was unable to complete the western blot validation for the SREBF2 and HPV16 E6/E7 knockdowns. I did optimise the SREBF2 antibody in control cells, but HPV16 E6/E7 antibodies need to be optimised, and experimental samples need to be tested. Due to the COVID-19 delays, I also was unable to repeat the RNA sequencing of the HPV16 E6/E7 knockdown to account for variation between biological replicates. Additionally, the decrease in air freight and the associated increase in boat freight due to COVID-19 prolonged the time for essential western blot reagents to arrive. In this case the shipment delay was greater than three months. The delivery of RNA samples for total RNA sequencing was also postponed due to travel restrictions. The sequencing service provider was also changed. As a result, the RNA sequencing results, and western blot reagents were received within the last four months of candidature. The differentially expressed lncRNAs were not followed up with confirmatory RT-qPCR or functional analysis, and antibody optimisation was delayed. To account for changes in the experimental timeline and available resources, the aims of this thesis were adjusted. The time delays were mitigated by shifting the main focus towards a more computational and bioinformatics approach for the synthesis of this thesis.

Table of Contents:

Supplementary/Appendix material

The supplementary tables listed below can be found via:

https://doi.org/10.26195/mdzm-qc54

Supplementary Table 4.1

Supplementary Table 7.1

Supplementary Table 7.2

Supplementary Table 7.3

Supplementary Table 7.4

Chapter 7:

List of figures:

Chapter 1:

Chapter 2:

Chapter 3:

Chapter 4:

Chapter 5:

Chapter 6:

Chapter 7:

Chapter 8:

List of tables:

Chapter 1:

Chapter 2:

Chapter 4:

Chapter 5:

Chapter 6:

Chapter 7:

Abbreviations:

Abbreviation	Term
ABCC1	ATP Binding Cassette transporter sub-family C
ACTB	Actin Beta
Ago	Argonaut protein
Air	Antisense of IGF2R non protein coding RNA
ANRIL	Antisense Non-Coding RNA in the INK4 locus
AS	Antisense
ATP	Adenosine triphosphate
B2M	Beta-2-Microglobulin
BACH1	BTB Domain and CNC homolog 1
BAK1	BCL2 Antagonist/Killer 1
BAN	4-Bromoanisole
BEND4	BEN domain containing 4
BRCA1	Breast Cancer type 1 susceptibility protein
C19MC	Chromosome 19 microRNA cluster
CAV1	Caveolin 1
CCEPR	Cervical Carcinoma Expressed PCNA Regulator
CDC	Centre for Disease Control
CDK	Cyclin-dependent kinase
CDKN1A	Cyclin-dependent kinase inhibitor 1
CDKN2B	Cyclin dependent kinase inhibitor 2B
cDNA	complementary DNA
CEBP	CCAAT Enhancer Binding Protein
ceRNA	competitive endogenous RNA
CO2	Carbon dioxide
COLCA	Colorectal Cancer Associated RNA
CREBBP	CREB Binding Protein
CTCF	CCCTC Finding factor
CTDSP	CTD small phosphatase like 2
Ctl	Control
DAVID	Database for Annotation Visualisation and Integrated Discovery
DGCR8	Di George critical region 8
DE	Differential Expression
dH2O	deionised H2O
DINO	Damage induced noncoding
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxy ribonucleic acid
dNTP	deoxyribonucleotide triphosphates
DOX	Doxycycline
E2F	E2F transcription factor
EBV	Epstein Barr virus
ECT2	Epithelial Cell Transforming 2 oncogene

EGFL7	EGF like domain multiple 7
EGFR-AS1	Epidermal Growth Factor Receptor antisense RNA 1
EGOT	Eosinophil Granule Ontogeny Transcript
EMT	Epithelial Mesenchymal Transition
EP300	Histone acetyltransferase p300
FACS	Fluorescence Activated Cell Sorting
FAM83H-AS1	Family with sequence similarity 83 member H antisense RNA transcript 1
FC	Fold Change
FCS	Foetal calf serum
FFPE	Fresh Frozen paraffin embedded
FOXC1	Foxhead Box C1
FOXCUT	FOXC1 promoter upstream transcript
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GAS5	Growth Arrest Specific 5
GATA2	GATA binding protein 2
GFP	Green Fluorescent Protein
GO	Gene ontology
GSk3B	Glycogen synthase kinase 3 beta
HERC5	HECT and RLD Domain containing E3 Ubiquitin Protein Ligase 5
HFKs	Human Foreskin Keratinocytes
HMGCR	3-Hydroxy-2-Methylglutaryl-CoA Reductase
HMGCS	3-Hydroxy-2-Methylglutaryl-CoA Synthase
HNC	Head and Neck cancer
HNSCC	Head and Neck Squamous cell carcinomas
HOTAIR	HOX Transcript Antisense intergenic RNA
HOTS2	H19 opposite tumour suppressor
HPV	Human papilloma virus
HPV+	HPV positive
HPV-	HPV negative
HRP	horseradish peroxidase
hrs	hours
HSC70	Heat shock cognate 70
HSCP90	Heat shock protein 90
IDT	Integrated DNA Technologies
IGF2	Insulin like growth factor 2
iHFK	Immortalised human foreskin keratinocytes
iNOK	Immortalised normal oral keratinocyte
IRF3	Interferon regulatory factor 3
ISH	Immunohistology chemistry
kb	Kilobase
Kcnq1ot1	KCNQ1 Opposite strand/Antisense transcript 1

KEGG	Kyoto Encyclopedia of Genes and Genomes
KSFM	Keratinocyte serum free medium
LDLR	Low density lipoprotein receptor
lncRNA	Long non-coding RNAs
MAFK	MAF BZIP transcription factor K
MALAT1	Metastasis Associated Lung Adenocarcinoma transcript 1
MCM7	Minichromosomal maintenance complex component 7
MEG3	maternally expressed gene 3
miRNA or miR	MicroRNA
MRE	miRNA response element
mRNA	messenger RNA
mTOR	Mechanistic target or rapamycin kinase
MYC	MYC proto-oncogene, BHLH transcription factor
NCAN	Neurocan
ncRNA	Non-coding RNA
NEAT1	Nuclear paraspeckle assembly transcript 1
oncomiR	Oncogenic miRNA
OPC	Oropharyngeal cancer
OSCC	Oral squamous cell carcinoma
PANDA	Promoter of CDKN1A antisense DNA damage activated RNA
PANDAR	Promoter of CDKN1A antisense DNA damage activated RNA
PBS	Phosphate buffered saine
PCAT7	Prostate Cancer associated transcript 7
PCNA	Proliferative cell nuclear antigen
PCR	polymerase chain reaction
PI3K	Phosphatidylinositol-3-Kinase
piRNAs	PIWI interacting RNAs
pRB	Retinoblastoma protein
pre-miRNA	precursor miRNA
pri-miRNA	primary miRNA
PRINS	Psoriasis Associated non protein coding RNA induced by stress
PRNCR1	Prostate Cancer associated non codding RNA 1
PTEN	Phosphatase and tensin homolog
PTENP1	Phosphatase and tensin homolog pseudogene 1
PTOV1-AS1	PTOV1 antisense RNA 1
PTPN1	Tyrosine-protein phosphatase non receptor type 1
PURPL	P53 upregulated regulator or p53 levels
PVT1	Pvt1 oncogene
RAN	Ras-related nuclear protein
RBL1	Retinoblastoma like protein 1
RISC	RNA induced silencing complex
RNA	Ribonucleic acid

RNA pol	RNA polymerase
RNA-GTP	RNA guanosine triphosphate
RNF157	Ring Finger Protein 157
RNU6B	RNA, U6 small nuclear 6
RPM	Reads per million
RT	Reverse transcription
RT-qPCR	Reverse transcription quantitative polymerase chain reaction
SAM	Significance Analysis of Microarrays
SCC	Squamous cell carcinoma
siRNA	Small interfering RNA
SIRT7	Sirtuin 7
SLC7A5	Solute Carrier Family 7-member 5
SNHG12	Small Nucleolar RNA host gene 12
snoRNA	Small nucleolar RNA
SNPs	Single nucleotide polymorphisms
SREBF2	Sterol regulatory binding factor 2
STAT5A	Signal Transducer and activator of transcription 5A
STRAP	Serine/Threonine kinase receptor associated protein
T-UCR	Transcribed ultra-conserved regions
TCGA	The Cancer Genome Atlas
TFBS	Transcription factor binding sites
TFs	Transcription factors
TIRARP	TIR domain containing adaptor protein
TMPOP2	Thymopoietin pseudogene 2
TP53	Tumour protein p53
TRINGS	TP53 regulated inhibitor of necrosis under glucose starvation
TSCC	Tongue squamous cell carcinoma
TTTY15	Testis-Specific Transcript, Y-linked 15
UBR4	Ubiquitin Protein Ligase E3 Component N-Recognin-4
UCA1	Urothelial cancer associated 1
UCSC	University of California Santa Cruz
UNSW	university of new south wales
USF1	Upstream Transcription factor 1
UTR	untranslated region
WT	Wild type
XIST	X-inactive specific transcripts
XPO	Exportin
YY1	Yin yang 1

Abstract:

The Human Papillomavirus is a major risk factor for Head and Neck cancers, having a strong association with the Oropharyngeal cancer (OPC) subtype. The high-risk variant, HPV type 16 (HPV16), is the cause for 90% of all HPV related OPCs. Non-coding RNAs, such as microRNAs (miRNAs) and long non-coding RNAs (lncRNAs), have been shown to have dysregulated expression in OPC. The dysregulated expression on non-coding RNAs has severe consequences on a cell and can lead to tumour development. The aim of this study was to identify specific non-coding RNAs that HPV16 and its viral oncogenes E6 and E7 target in HNC, specifically OPC, and the mechanisms behind their altered expression.

To further our understanding of HPV16s impact on miRNAs, we investigated the expression levels of miRNAs in OPC tissue. We compared miRNA expression between HPV16+ and HPV16- samples using an LNA miRNA array. MiR-496 and miR-33a were found to be the most significantly deregulated. Using a bioinformatics approach, we constructed a novel HPV16-non-coding RNA interactome to identify mechanistic links between miR-33a, SREBF2 and miR-496. Interestingly, SREBF2 harbours miR-33a, and we believe there is a correlation whereby E6 and E7 regulate the expression of SREBF2/miR-33a, which then has a downstream impact on miR-496. TCGA analysis showed that SREBF2 was significantly higher in the HPV+ OPC and not in cervical samples. This suggests this pathway is unique to OPCs. We knocked out HPV16 E6/E7 and SREFB2 using siRNAs and showed that the viral oncogenes could regulate SREBF2 and miR-33a. Although with the knockdown of SREBF2, we saw no change in miR-33a expression, suggesting that miR-33a is not co-expressed with its host gene and is under its own form of regulation by HPV16. This study has identified a regulatory pathway involving HPV16 E6/E7 and specific miRNAs. We also sought to expand our research to include other regulatory RNAs such as long non-coding RNAs.

The lncRNA, TRINGS (TP53 regulated induced by glucose stress), was investigated in HPV16 E6/E7 expressing cells. It was determined that TRINGS expression is decreased with the presence of the viral genes. Using a siRNA knockdown system, we demonstrated that the modulation of TRINGS expression will reduce cell viability

of cervical cancer cell lines and that TRINGS expression could not be stimulated by glucose starvation alone. TRINGS was also determined to be specific for HPV related cervical cancers. We believe that HPV16 targets a different cohort of lncRNAs in HPV related OPC. Given this, we assessed raw RNA-sequencing data from TCGA to identify differentially expressed lncRNAs in HPV16+ OPC. In future research we aim to investigate the function of these lncRNAs in HPV related OPC.

This study has identified the complexity of molecular pathways between HPV16 E6/E7 and non-coding RNAs and the consequences in Head and Neck cancers. The construction of viral interactomes combined with in vitro approaches provided novel insights for discovering mechanistic links in HPV16+ Head and Neck cancers.

Chapter 1: Introduction and literature review

Part of this chapter has been published in the journal “Current Opinions in Virology”. *Mason, D., Munger, K. and Tran, N., 2021. The dynamic interactome of microRNAs and the human papillomavirus in head and neck cancers. Current Opinion in Virology, 51, pp.87-95*

Head and Neck cancers (HNC) accounted for 890,000 new cases and 450,000 deaths in 2018, making it the 7th most common cancer worldwide [1]. The most common form of Head and Neck cancer is the squamous cell carcinoma (HNSCC), which are typically found in the mucosal lining of the larynx, nasal cavity, paranasal sinus, pharynx, oral cavity and salivary glands [2]. The main risk factors for HNSCC are alcohol consumption and smoking [3], but recently it has emerged that the Human Papillomavirus (HPV) is also a significant risk factor for HNSCC. This virus has been identified in approximately 25% of HNSCC cases [4, 5]. The number of cases of HNSCC associated with HPV has been increasing over the past few years and is strongly related to Oropharyngeal carcinomas (OPC) [6], a subtype of HNC [3].

The increasing incidence demonstrates that HPV is an important etiological agent of HNSCC, particularly OPC. Our knowledge of the specific molecular pathways that drive HPV infected OPCs, is lacking. The elucidation of molecular mechanisms involved and the processes behind this disease may assist in developing novel diagnostic and treatment strategies. Specifically, we will be investigating regulatory non-coding RNAs, including the small non-coding RNAs microRNAs (miRNAs) and long non-coding RNAs (lncRNAs).

The overarching aim of this study is to determine the impact of HPV16 on the non-coding RNAs involved in the development of Head and Neck cancers. Once these alterations are identified, we can then gain a better understanding of which molecular pathways are influence by the virus. With this new information, new targets can be identified for future novel treatments.

1.1. Head and Neck cancers

1.1.1. Prevalence of Head and Neck cancers

Head and Neck squamous cell carcinomas (HNSCC) are a heterogeneous disease characterised by the accumulation of genetic alterations leading to the modification of cellular processes [7]. HPV has become a significant risk factor for HNSCC, and the viruses presence has been observed frequently in the oropharyngeal region [3, 6] (Figure 1.1). Squamous cell carcinomas (SCC) represent greater than 95% of tumours of the oropharynx [8] and 60-70% of Oropharyngeal cancers (OPC) is attributed to HPV [2, 4, 9]. HPV variants are classified as high risk or low risk based on clinical outcome. The dominant variant causing OPC is HPV type 16, accounting for up to 85% of HPV associated OPCs [10]. In the 1980s, in the United States, 16.3% of HNSCC were diagnosed as HPV+ OPC; this rose to 72.7% in the 2000s [11]. We are not expected to see a decrease in the incidence of HPV+ OPC until 2060 due to lower vaccine awareness and uptake in the most impacted group, older males [12]. Due to the fast-growing rates of HPV+ OPC, it is essential to understand the similarities and differences between positive and negative HNSCC in terms of diagnosis, treatment, clinical and biological features.

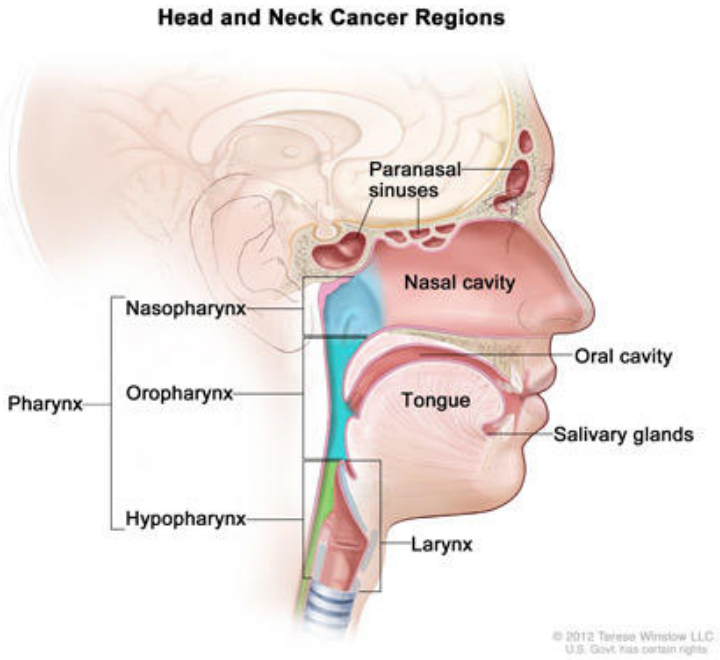


Figure 1. 1: Head and Neck squamous cell carcinoma subtypes*:* Head and Neck cancer is divided into several subtypes, including, Nasopharyngeal, Oropharyngeal, Hypopharyngeal, Laryngeal and Oral cavity. Diagram from: “Head and Neck Cancers” published by the National Cancer Institute Credit to Terse Winslow

1.1.2. Clinical treatment and outcome

Current methods for the first line of treatment of HNSCC include a combination of surgery, chemotherapy and radiotherapy. Approximately one-third of HNSCC patients present with early-stage disease and are treated with surgery or radiotherapy, with cure rates between 70-90% [13]. Despite this success rate with early-stage disease, the majority of patients present advanced stages of cancer, which have a high rate of treatment failure [14]. Currently, radiotherapy is the standard treatment for HNSCC, when surgery is not a viable option. Due to the anatomical location of these tumours, treatment options present many challenges to the patients. The treatment of HNSCC can be toxic to the patient with severe side effects such as mucositis, dermatitis, dysphagia and loss of swallowing function or speech [15]. Within two years of tumour removal, there is a 50% chance of relapse [13]. These methods are not preventative and are associated with significant morbidity. Interestingly patients with HPV associated OPC have a significantly better outcome than patients with an HPV-disease [16, 17] and have an enhanced response to treatment methods [18].

Methods for the early detection, diagnosis and treatment of OPC, both positive and negative for HPV, need to be developed. Most patients are diagnosed in the later stages with a less favourable outcome. It is also essential to understand the risk factors that may cause OPC, as HPV positive OPC present different clinical outcomes to negative OPC.

1.1.3. Risk factors from Head and Neck cancer

The number of HNSCC related to smoking and alcohol consumption have declined [19] and a new risk factor, HPV, has emerged. Of the HPV negative (HPV-) HNSCC, smoking accounts for approximately 70% of cases and alcohol consumption 30% in developed countries [20]. In developing countries such as Southern Asia, HNSCC is primarily associated with smokeless or chewing tobacco [21, 22]. The Human Papillomavirus (HPV) has been identified in approximately 25-50% of HNSCC cases [2, 4] with an increasing incidence. Developed countries usually see a higher percentage of HNSCC associated with HPV, making the virus the primary causative agent.

HPV positive (HPV+) HNSCC cancer patients have shown a different demographic pattern to the HPV- patients, typically being younger (40-60 years of age) and with no history of smoking [23, 24]. There are numerous differences between HPV+ and HPV- HNSCC, both clinically and biologically. When HPV was reported in OPCs, it remained controversial until several landmark publications revealed that HPV+ and negative OPCs are two distinct disease entities [25].

1.1.4. HPV positive vs. HPV negative Head and Neck cancers

With HPV arising as a key risk factor for OPC, numerous studies have demonstrated that HPV+ and HPV- OPC are two distinct entities. These two cancer types represent uniquely dysregulated gene expression networks and proteomic consequences, contributing to tumorigenesis.

Demographically HPV+ OPC patients are generally younger with no smoking history [16]. In the United States, higher rates of HPV positivity were found in OPC groups with greater socioeconomic status and higher educational levels [16]. The observed socioeconomic and regional differences may be due to the differing sexual behaviour across regions [26] and accessibility to health care and education. Studies have shown that HPV+ OPC have an enhanced response to treatment methods [18] and the difference in survival rates appear to be significantly driven by the HPV tumour status [27, 28].

This trend may be attributed to the biological and molecular difference between the HPV+ and HPV- HNSCC. The specific mechanisms by which the HPV virus achieves a favourable outcome are yet to be discovered. Table 1.1 summarises the differences observed between HPV+ and HPV- HNSCC, highlighting the importance of molecular studies of this disease.

Table 1. 1: Comparison of HPV positive and HPV negative HNSCC

Characteristic	HPV Positive	HPV negative	Reference
Main risk factor	Sexual behaviour and exposure to the virus	Smoking and alcohol consumption	[20, 29]
Incidence	Increasing	Decreasing	[30, 31]
Predominant subsite	Oropharyngeal	All HNC	[2, 4] [6]
Histology	Non-keratinised	Keratinised	[23, 32]
Age	40-60	>60	[16]
Gender	Predominantly Male	Varied	[16, 33]
P53 status	Wild type	Mutated	[34, 35]
P16 expression	Overexpressed	Reduced expression	[28]
Chemo-radiotherapy and recurrence	Better response to therapy	Worse response to therapy	[36] [18, 37]
Overall Survival	Improved overall survival	Worse overall survival	[27, 28, 36]

Whole exome sequencing of tumour tissues showed that HPV- tumours contained more mutations than HPV+ [34]. HPV- HNSCC are characterised by deletions of whole or large parts of chromosomal arms, whereas HPV+ HNSCC only showed a minimal chromosomal loss [35]. HPV+ HNSCC typically maintain wild type TP53 (p53), whereas 75% of the HPV- HNSCC present TP53 mutations. The overexpression of p16 are also typically molecular characteristics of HPV+ HNSCC [38-40].

An investigation into differential gene expression in OPC demonstrated over 59 differentially expressed genes between HPV+ and HPV- OPC [41]. During proteomic studies comparing HPV+ and HPV- OPC, Selbo *et al.* [42] showed 309 proteins having a higher level in HPV+ OPC than the normal epithelium and 319 proteins having a 2-fold difference between HPV+ and HPV- OPC. Those proteins enriched in HPV+ OPC were associated with DNA initiation and replication and cell cycle control, demonstrating the viral promotion of proliferation of cells. These vast genetic differences between HPV+ and HPV- HNSCC is incredibly evident, highlighting the need to study them individually.

HPV is a primary etiological agent for HNSCC, but there are gaps in our knowledge concerning the molecular mechanisms involved in HPV+ HNSCC. To develop novel diagnostic and treatment strategies for this disease, we need to investigate the role HPV has to play and the impact this virus has on the molecular pathways within a cell.

1.2. Non-coding RNAs in cancers

Two-thirds of the human genome is estimated to be transcribed, but only 2% of these transcribed RNAs are translated into proteins [43, 44]. It has become increasingly apparent that the non-protein-coding portion of the genome (or non-coding RNAs) plays a significant role in normal cell functioning and has great implications in disease development.

Many types of non-coding RNAs have been discovered. Non-coding RNAs are divided up into two groups according to their size. Small non-coding RNA group includes microRNAs, transcribed ultra-conserved regions (T-UCR), PIWI-interacting RNAs (piRNAs) and small nucleolar RNAs (snoRNAs) [45]. Non-coding RNAs that are 200nt or greater in length are referred to as Long non-coding RNAs (lncRNAs) [45].

All of these non-coding RNAs have distinct functions with the overall aim to regulate biological pathways within a cell to promote normal cell functioning. In this study, we aimed to focus on the small non-coding RNA, microRNAs, and long non-coding RNAs.

1.3. Small non-coding RNAs: microRNAs

1.3.1. MicroRNAs and their Discovery

MicroRNAs are short single-stranded non-coding RNAs involved in the regulation of genes [46]. The target genes regulated by miRNAs are associated with numerous biological processes, such as developmental timing [47], stem cell division [48, 49] and apoptosis [50]. MiRNAs were first identified in 1974 [51], there is still much about these gene regulators that remains unknown to the scientific community.

In 1974 miRNAs were first identified during the genetic studies of *C.elegans* and were shown to be involved in regulating developmental processes [51]. *Lee and Ambros* [52] went on to define the mechanisms behind miRNAs actions in 1993. They found

that the miRNA, lin-4, was a non-protein-coding RNA that contained sequences complementary to the 3`UTR of the lin-14 messenger RNA (mRNA). *Lee and Ambros* (1993) [52] suggested that lin-4 bound to mRNA transcript of lin-14 prior to translation, resulting in the regulation of the expression of the lin-14 protein.

The miRNA Let-7 was also one of the first miRNAs to be found, in C.elegans, with complementary sequences within the 3'UTR of 39 genes, including lin-14, lin-28, lin-41 and lin-42 [47]. Lin-4 and Let-7 were later identified in several species, including humans, with a temporal expression similar to that seen in C.elegans [53].

More recent studies have established that the binding of the miRNA to the 3`UTR of target mRNA sequences will induce mRNA degradation and translational repression [54]. Now hundreds of miRNAs have been discovered in humans, some being highly conserved. It is now widely accepted that miRNAs are crucial components in gene regulation, and their deregulation can lead to the development of human diseases such as cancer. Hence, it is vital to understand how miRNAs are produced, regulated and carry out their function.

1.3.2. The biogenesis of microRNAs

The biogenesis of miRNAs involves several stages of processing. Production begins within the nucleus, where the miRNA is transcribed; they are then exported into the cytoplasm and undergo maturation to form the effector miRNA. Figure 1.2 depicts the miRNA biogenesis process.

Once cleaved, the pre-miRNA is then transported to the cytoplasm via a nuclear transport factor, exportin-5 and RNA-GTP [55, 56]. The pre-miRNA is then further processed by Dicer, which forms single strands from the double-stranded molecule [57, 58]. Dicer cleaves both the pre-miRNA strands near the loop to generate a miRNA duplex with a 2nt 3' overhang on each end [59]. One strand of the miRNA duplex the 'passenger strand', which is degraded and the other, the 'guide strand', forms the mature miRNA and carries out its function [60]. The miRNA duplex is then loaded onto an Argonaut protein (Ago), rendering it single-stranded, with assistance from chaperon proteins, HSC70 and HSCP90 [61].

This forms the RNA induced silencing complex (RISC), which is the effector for mediated miRNA activity. RISC facilitates gene silencing post-transcriptionally via the degradation of mRNA (by cleavage or deadenylating) or the repression of translation [62]. The fate of the mRNA is determined by the degree of base pairing with its miRNAs [63].

Perfect base pairing results in mRNA degradation, whereas imperfect pairing can lead to translational repression [62]. MiRNAs are stable when loaded onto the Argonaut protein and have a half-life of days [64]. The region for which miRNAs bind to their target mRNA is 2-8nt from the 5`end of the miRNA, referred to as the seed region.

The biogenesis pathway is tightly regulated at numerous stages such as miRNA transcription, Drosha and Dicer processing, transportation into the cytoplasm and RISC binding. During transcription, miRNAs transcribed from introns are controlled by the same transcriptional regulators as the mRNA of their host gene [65], and other miRNAs are under post-transcriptional control.

Any alteration to the biogenesis machinery or the functioning of miRNAs has been shown to have severe consequences and can lead to the development of cancer. It has been shown that miRNAs have altered expression levels in HNSCC and that this altered expression may be due to the risk factor HPV.

1.4. Manuscript: The dynamic interactome of microRNAs and the Human papillomavirus in Head and Neck cancers

1.4.1. Copyright information:

This chapter has been published as a **Review** article in ***Current Opinions in Virology***:

Current Opinion in Virology 2021, 51:87–95
https://doi.org/10.1016/j.coviro.2021.09.013

The text presented here is the accepted version of the manuscript. The numbering of sections, style of referencing, the numbering of tables and figures are altered to align with the formatting of the thesis.

Copyright Declaration:

1.4.2. Author contribution

Manuscript: The dynamic interactome of microRNAs and the human papillomavirus in head and neck cancers *Current Opinion in Virology*

Authors: Dayna Mason, Karl Munger and Nham Tran

Author Contribution:

Dayna Mason (graduate student) is the first author of this Review paper. Her contribution to this manuscript is demonstrated by the following roles/tasks:

Dayna Mason researched the field and consolidated information from the literature to form the review of HPV16 induced miRNAs in OPC presented in Table 1.1 and Figure 1. She designed and performed the bioinformatic and network analysis and predictive target analysis featured in Figure 2 and Figure 3. In addition, she was actively involved in the writing of the initial draft, the iterative editing, and the review process.

Nham Tran and Karl Munger provided critical analysis of the conceptual data advice, provided expert opinions, and revised the manuscript.

Signatures:

Dayna Mason (graduate student) Production Note: Signature removed prior to publication.

Nham Tran (DM PhD Principal Supervisor) Production Note: Signature removed prior to publication.

Karl Munger (Co-Investigator) Production Note: Signature removed prior to publication.

1.4.3. Manuscript: Abstract

The Human Papillomavirus type 16 is a major etiologic factor for a subset of Head and Neck cancers. These cancers of the oropharyngeal region are growing, and it is expected to exceed cervical cancers in the near future. The major oncogenes E6 and E7 mediate many early transformation stages targeting p53 and other tumour suppressor genes. The majority of this regulation is centred on protein-coding genes, but more recently, small non-coding RNAs, such as miRNAs, are also regulated by HPV16. However, the system-wide impact of HPV16 on miRNAs is yet to be fully understood. Several studies have devised dynamic interactomes encompassing viral oncogenes, miRNAs, and gene targets to gauge the overall relationship between HPV16 and miRNAs fully. These interactomes map potential pathways, which permit the identification of possible mechanistic links. Our review will discuss the latest developments in using viral interactomes to understand viral mechanisms and how these approaches may aid in the elucidation of potentially druggable pathways.

1.4.4. Manuscript: HPV rates are increased in Head and Neck cancers

Head and Neck cancers (HNC) were the seventh most common cancer worldwide in 2018 (890,000 new cases and 450,000 deaths) [1]. The incidence of this cancer type has increased over the past decade. They are typically observed as squamous cell carcinomas (SCC) within the mucosal lining of the larynx, nasal cavity, paranasal sinus, pharynx, oral cavity and salivary glands.

Smoking and the consumption of alcohol are considered the main risk factors for HNC [26], but cases attributed to these have been declining. More recently, HNC associated with HPV have been increasing, with a strong association with oropharyngeal carcinomas (OPC) [66], a subtype of HNC. The oropharyngeal region consists of the soft palate, base of the tongue, pharyngeal wall, tonsils and tonsillar region. Approximately 60-90% of OPC are attributed to HPV. The number of HPV- OPC has been declining by ~50% in the United States over the past decade [67]; this trend is similar across other developed countries. The CDC [68] reported there were approximately 19,800 OPC cases per year in the US; 14,000 (70%) were attributed to HPV (between 2013-2017). In comparison, only 11,000 HPV related cervical cancers

are reported per year. OPC made up 39% and cervical cancers 30.6% of all HPV related cancer cases reported per year in the US between 2013-2017 [68]. Of the HPV+ OPC cases, up to 85% are infected with the high-risk HPV16 [10]. There are numerous differences between HPV+ and HPV- OPC, both clinically and biologically, classifying them as two separate entities [25].

These cancers types represent uniquely dysregulated gene expression networks [41] and proteomic profiles [42]. Studies show that HPV+ OPC has an enhanced response to treatment methods [18] and differences in survival rates appear to be significantly driven by the HPV tumour status [27, 28]. A lower proportion of distant metastasis seen in HPV+ patients (54.2% in HPV- patients to 34.2% in HPV+ patients)[16]. This trend may be attributed to the biological and molecular difference between the HPV+ and HPV- OPC. Specific mechanisms by which the HPV virus achieves a favourable outcome are yet to be discovered, hence highlighting the importance of molecular studies of this disease.

1.4.5. Manuscript: Human Papillomavirus

Human Papillomaviruses (HPVs) are non-enveloped viruses with circular double-stranded DNA genomes that infect undifferentiated, dividing basal cells of the squamous epithelium [69]. Members of the alpha genus HPVs have been categorized as “low” or “high” risk based on the clinical outcome. Low-risk types cause benign genital warts, whereas high-risk cause lesions that have a propensity for malignant progression. Currently, 13 HPV types (HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 66) have been classified as class I carcinogens. It is estimated that up to 80% of the population in developed countries come into contact with HPV (either high or low risk) at some point in their lifetime, highlighting the global impact of these viruses [70, 71]. Moreover, HPVs are associated with approximately 4.8% of the total cancer burden globally; this is the highest rate in comparison to other viruses [72].

The 8kb genome is divided up into several regions [73]. The early (E) region contains up to seven genes, E1, E2, E4, E5, E6, E7 and E8, that code for non-structural proteins, with various levels of involvement in viral infection. The late (L) region encodes the major and minor capsid proteins L1 and L2. The long control region contains the viral origin of replication and enhancer and promoter elements that control

the transcription of the viral genome [73]. High-risk HPVs encodes the oncogenic proteins, E6 and E7, which are consistently expressed in cancer cells. E6 binds to and causes degradation of the tumour suppressor p53 through the ubiquitin-proteasome pathway (Figure 1.2). E7 targets the retinoblastoma tumour suppressor, pRB, for degradation, which causes uncontrolled cell cycle progression.

These oncogenes have many additional cellular targets and affect nearly all known cancer hallmarks. Importantly, cancer cells remain “addicted” to the expression of the E6 and E7 proteins and, thus, the pathways that they affect are viable targets for therapy of high-risk HPV-associated lesions and cancers. While much of the previous research on HPV cellular targets have focused on protein-coding genes, there is an increasing interest in determining the impact of HPVs on non-coding RNAs, such as microRNAs (miRNAs, miRs).

1.4.6. Manuscript: MicroRNAs: tiny gene regulators

MicroRNAs are short, 21-22 nucleotides (nt), non-coding RNAs that act as gene regulators [51, 74-76]. They typically control gene expression by binding to specific sequences in the 3` untranslated region (UTR) [77]. This binding prevents the translation of RNA messages into proteins or can lead to the degradation of mRNAs, impacting processes such as cell differentiation, development, growth and cell death. Approximately >60% of all protein-coding genes are influenced by a miRNA, with a single miRNA having the ability to control hundreds of targets [78, 79]. The biogenesis of miRNAs involves several stages of processing, beginning with transcription in the nucleus, export to the cytoplasm and maturation to form the effector miRNA [80]. For a schematic of the miRNA, biogenesis processing, refer to Figure 1.2.

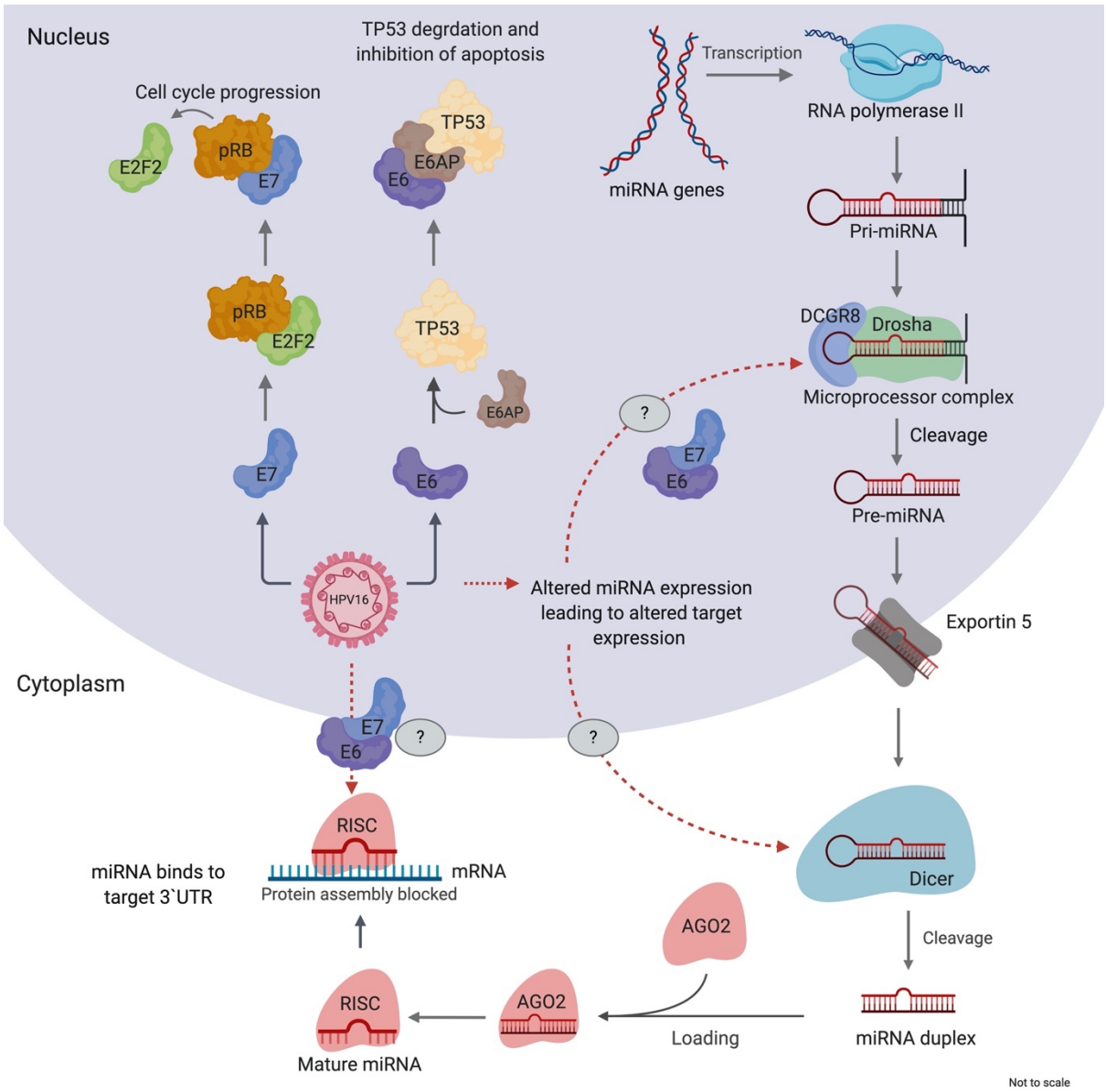


Figure 1. 2: The relationship between HPV16 oncogenes and miRNA biogenesis. miRNAs are transcribed in the nucleus by RNA polymerase II to form the pri-miRNA. The pri-miRNA is then cleaved by the microprocessor complex to form the pre-miRNA, which is then exported into the cytoplasm. The pre-miRNAs are cleaved by Dicer to form a miRNA duplex which is then loaded onto an Argonaut protein. The miRNA duplex unwinds to a single-stranded target miRNA to form the RNA induced silencing complex. This complex binds to target messenger RNA at the 30UTR to block protein assembly. HPV16 has two major viral oncogenes**,** E6 and E7. E7 binds to pRb, which releases the transcription factor E2F2, leading to cell cycle progression. E6 recruits E6AP to bind TP53, inducing its degradation and inhibiting apoptosis. HPV16 can mediate miRNA expression at various steps

We know that cancer progression is a multistep process by which normal cells develop genetic damage over time; this leads to abnormal cell growth and development. Due to their regulatory nature, miRNAs play an integral role in multiple cellular processes and alterations to their levels can trigger pathways associated with the tumorigenic process [81].

This was first shown in B-cell lymphocytic leukaemia, whereby miR-15 and miR-16 were frequently deleted or downregulated in leukaemia patients [74]. The aberrant expression of miRNAs can be classified into two groups, depending on their expression status. Overexpressed miRNAs that suppress tumour suppressor genes are termed oncogenic miRNAs or oncomiRs.

The upregulation of these miRNAs can lead to uncontrolled cell proliferation and metastasis. The archetype oncomiR is miR-21, which is overexpressed in most cancer types, including HNC [82-86]. MicroRNAs that are frequently downregulated in cancer are classified as tumour suppressor miRNAs. An example is the Let-7 family, which can negatively regulate the expression of RAS [87].

In one of the first studies to associate miRNAs in HNC cell lines [82], up to 261 mature miRNAs were investigated. A large proportion of these miRNAs were dysregulated, including miR-21. Another study confirmed that miR-21 with miR-55, let-7i, miR-142-3p, miR-125b and miR-375 were also dysregulated in HNC [85]. It is now well accepted that miRNAs can induce transformative changes and activate tumorigenic pathways associated with HNC. They can be used for as biomarkers for the prediction of patient disease recurrence and overall survival [88, 89] and prediction of radiotherapy response to optimise radiation strategies [90] or even as a treatment method with the use of tumour suppressive miRNAs [91].

1.4.7. Manuscript: The dynamic regulation of the miRNAome by HPV16

Given the impact of HPV16 on OPC and its global emergence, we still lack a fundamental understanding of the regulatory impact of the virus on the miRNA milieu (miRNAome). Although there are a vast number of findings investigating miRNAs in

HNC, there are very few studies investigating HPV related OPC specifically (Table 1.2).

One of the first studies comparing HPV+ cell lines to HPV- identified several miRNAs with differential expression: miR-363, miR-33 and miR-497 were significantly upregulated while miR-155, miR-181a, miR-181b, miR-29a, miR-218, miR-222, miR-221 and miR-142-5p were downregulated. They also suggested that E6 could influence the expression of miR-363, miR-181a, miR-218, and miR-29a in Human foreskin keratinocytes overexpression systems (HFKs) [92]. When we reviewed the data from the literature, there are very few miRNAs that share a similar expression pattern. For example, miR-107 was seen to be upregulated in one study [93] but shown to be downregulated by others [94, 95]. These discrepancies could be due to different experimental designs, methods of miRNA detection, type of samples and differences in the patient cohort (age, gender, size, smoking and drinking habits). The miRNAs that have similar expression patterns across various studies (highlighted in Table 1.2) are the upregulated miRNAs: miR-363, miR-150, miR-20b and miR-9 and the downregulated miRNAs: miR-181b and miR-31. The majority of recent studies shown in Table 1.2 catalogue the differences in the miRNAome in HPV+ and HPV- tissues.

The differences in miRNA expression between HPV+ and HPV- OPC has the potential to elicit different phenotypic responses due to HPV inducing a widespread dysregulation of miRNAs and gene expression. These HPV dysregulated miRNAs and genes are largely attributed to cell cycle regulation [96]. This may be responsible for the distinct clinical behaviour of HPV associated OPC. HPV+ OPC tends to be non-keratinised [23, 32], respond better to treatment [18], have lower levels of metastasis [16] and show improved overall patient survival [17]. Identifying the pathways with which these miRNAs are associated with may provide molecular clues for these differences.

Table 1. 2: Summary of microRNA expression profiles in HPV+ oropharyngeal and oral cancers.

Paper	Sample Type	Number of samples	Methodology for miRNA detection	Upregulated miRNAs	Downregulated miRNAs
Wald *et al.* (2011) [92]	Cell lines	Cell lines=6 HPV+=4 HPV- =2	mirVana miRNA Bioarray V2, RT-qPCR	**miR-363**, miR-33	miR-155, miR-181a, **miR-181b**, miR-29a, miR-218, miR-222, miR-221, miR-142-5p
Lajer *et al.* (2011) [97]	Fresh frozen	Normal=39 HPV- =49 HPV+=9	Affymetrix miRNA microarray (miRbase v.11) RT-qPCR	**miR-363**, miR-195-star	miR-127-3p, miR-376, miR-125a-5p, miR-432, miR-143, miR-145, miR-409-5p, miR-33, miR-381, miR-199a-3p, miR-199b-3p, miR-26b, miR-199b-5p, miR-143-star, miR-1201, miR-126, miR-409-3p, miR-101, miR-517a
Lajer *et al.* (2012) [98]	Fresh Frozen FFPE	Normal=13 HPV- =46 HPV+=10	Affymetrix GeneChip miRNA array RT-qPCR	**miR-363**, miR-21, **miR-150,** miR-146b-5p, miR-15a, **miR-20b**, miR-146a, let-7g, miR-625, miR-155, miR-15b, miR-29a, let-7f, miR125b, miR-26b, miR-342-3p	miR-16, miR-768-3p, miR-34a-star, miR-596, miR-20b-star, miR-598, miR-200c, let-7e, miR-1275, **miR-181b**, miR-324-5p, miR-423-5p, miR-125a-3p, miR-99b, miR-744, miR-877, miR-99b-star, miR-1180, **miR-31**, miR-193b-star

Gao *et al.* (2013) [99]	FFPE	HPV- =150 HPV+=82	RT-qPCR	**miR-9**, miR-155	**miR-31**, miR-223, miR-18a
Hui *et al.* (2013) [94]	FFPE	Normal=7 HPV- =88 HPV+=56	Taqman Low Density Array Human Micro-RNA panel v1.0	**miR-9**, **miR-20b**, miR-422a, miR-492, miR-545, miR-591	miR-107, miR-193b
Miller *et al.* (2015) [93]	FFPE TMA, cell lines, TCGA data	HPV- =109 HPV+=63	Qiagen miRNome qPCR, microRNA ISH	**miR-9**, miR-15b, miR-16-2, **miR-20b**, miR-25, miR-29c, miR-93, miR-106a, miR-106b, miR-107, miR-148a, **miR-150**, miR-222, miR-320a, miR-335, **miR-363**, miR-378, mR-598, miR-625	miR-126, miR-143, miR-145, miR-193b, miR-199a, miR-199b, miR-214, miR-337, miR-574
Mirghani *et al.* (2016) [95]	Fresh Frozen	HPV- =26 HPV+ =11	3D-Gene human miRNA microarray (miRbase v17) RT-qPCR	miR-7-2-3p	miR-324-5p, miR-4764-3p, miR-107, miR-1234, miR-3144-3p, miR-3176, miR-3177-3p, miR-4267, miR-4418, miR-615-3p, miR-668, miR-99b-3p, miR-675-3p. miR-584-5p, miR-212-3p, miR-18b-5p, miR-18a-5p, miR-138-1-3p, miR-135b-5p, miR-1246, miR-675-3p, miR-476-3p, miR-857

Quabius *et al.* (2017) [100]	FFPE	HPV+=53 HPV- =73	Affymetrix miRNA 3.0 array RT-qPCR	**miR-363**, miR-21, miR-378c, miR-378i, miR-210, miR-378f, miR-422a, **miR-150**, miR-15a, miR-378, miR-182, miR-200a, **miR-20b**, miR34a, miR-34c-5p, miR-141, miR-130a, miR378d, miR-3188, miR-155, miR-378g, miR-30a, miR-27b, miR-378-star	**miR-31**, miR-193b-star, miR-3126.3p, miR-181a, **miR-181b** miR-1275, miR-3126-3p miR-4749-3p, miR-486-3p, miR-1976, miR-342-5p
Mason *et al.* 2018) [101]	Fresh Frozen	HPV+=15 HPV- =15	LNA microarray (miRbase v7.1) RT-qPCR	miR-33, miR-210, miR-146b, miR-342, miR-142-3p	miR-518c-star, miR-505, miR-520a-star, miR-362, miR-519e, miR-504, miR-520d, miR-496

1.4.8. Manuscript: Interactomes: mapping the system-wide impact of HPV

The interactions between the HPV16 viral oncogenes and human proteins is well understood, but a large piece is missing from this puzzle, non-coding RNAs.
It is evident that HPV16 has a major impact on miRNA expression in OPC, this may translate to thousands of gene targets regulated by either HPV and miRNAs. Recently, system-wide integration or mapping of these interactions into an HPV-miRNA interactome has provided further mechanistic insights [101]. The full methods on how to develop such as interactome are described in Hill *et al*. [102]. In summary, miRNAs from Table 1.2, which are seen across 3 or more studies, were linked to their mRNA targets. These mRNA targets were filtered to include only those that are also directly targeted by HPV16 viral genes, using online databases.

Lapa *et al.* [103] described the use of a miRNA-mRNA interactome to identify core sets of miRNAs with regulatory roles in laryngeal cancers as well as identifying drugs that can act on these miRNAs and their targets. Another interactome showed the relationships between different RNA families, mRNAs, miRNAs and long non-coding RNAs. Numerous studies have now constructed similar non-coding RNA interactomes to uncover genes associated with overall patient survival [104-107]. The creation of an interactome for these types of analysis is relatively easy to do, given the enormous amount of data that is currently online. One of the major benefits is that it allows the visualisation of pathways. We believe this *in silico* approach will be the first step in both discovering new pathways but also linking up potential drug targets for clinical development.

From the list of miRNAs described (Table 1.2), using an HPV-miRNA interactome can now reveal the relationships between specific miRNAs, HPV oncogenes and their targets (Figure 1.3). In this figure, miR-20b-5p [93, 94, 98, 100], miR-150-5p [93, 98, 100] (upregulated in HPV+ cells) and miR-181b-5p [98, 100, 108] (downregulated in HPV+ cells) seem to have a greater regulatory impact in HPV related OPC than the other miRNAs.

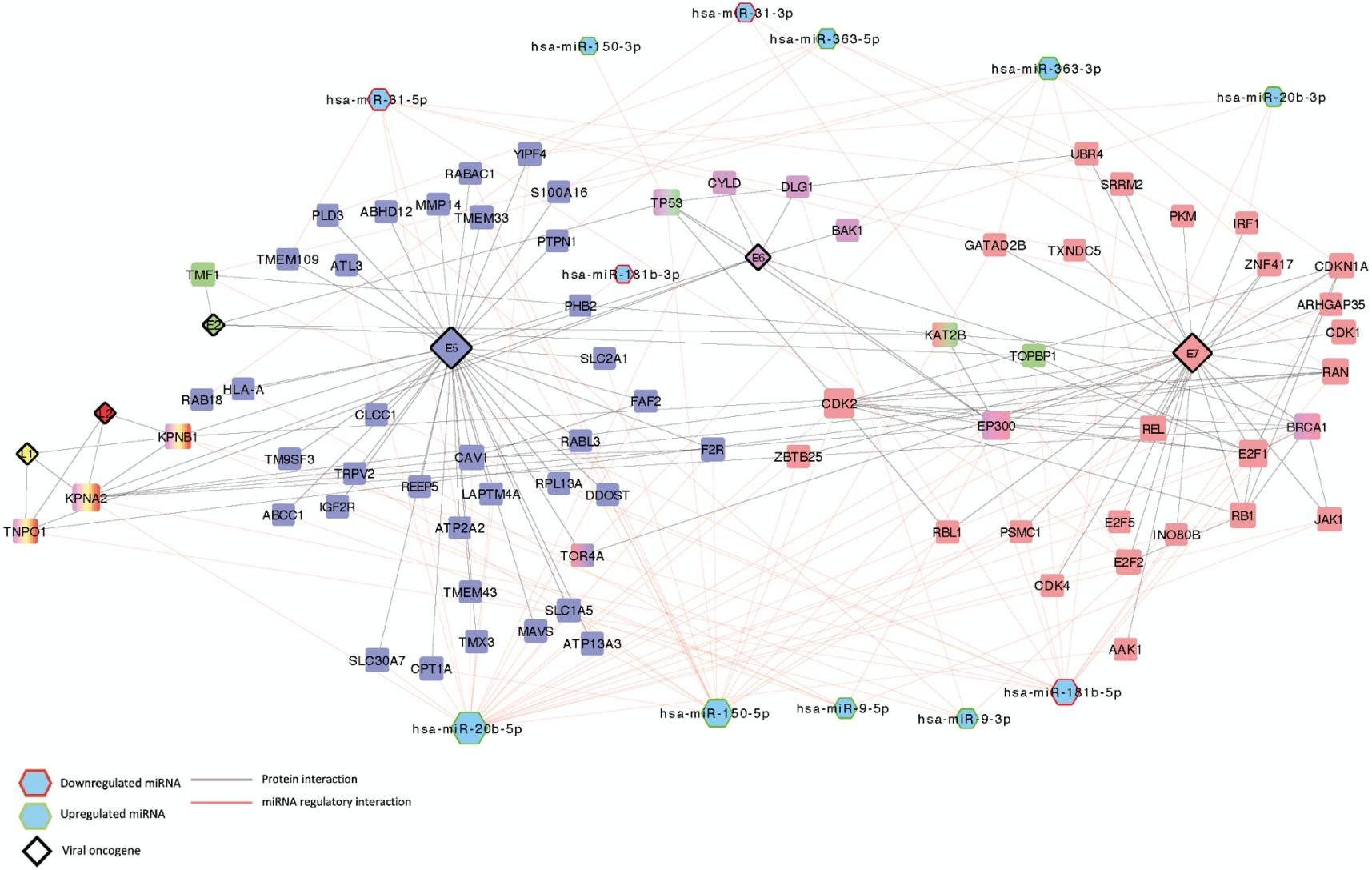


Figure 1. 3: HPV-miRNA interactome for oropharyngeal cancers*:* Modelling of the HPVP16 interactome showing the HPV16 viral genes (L1, L2, E2, E5, E6 and E7), the proteins which HPV directly target and the miRNAs which can target these genes.

For the above-mentioned miRNAs, miR-20b, miR-150 and miR-181b have been shown to regulate the cell cycle transcription factor E2F family (E2F1 and E2F2). The E2F family play a critical role in cell cycle regulation, and its activity is regulated by its association with pRb. Of these miRNAs, miR-20b has been shown to be in an auto-regulatory feedback loop with E2F1 to regulate proliferation and differentiation [109] and modulate the epithelial-mesenchymal transition [110]. This viral interactome can be used to postulate additional genes in previously identified models. We know that miR-150 is upregulated in OPC and targets both CDK2 [111] and p53. From the interactome, we can extend and connect other genes. Of note, miR-181b and miR-20b may regulate RBL1 and UBR4, respectively, which may then potentially affect p53 levels (Figure 1.3). When we apply interactomes in this manner, they can be powerful tools that can find mechanistic associations on an individual basis. Filtering steps can further pinpoint single gene-miRNA tumorigenic pathways, which can provide new information on the impact of HPV driven OPCs.

This analysis can also be extended to understand if these miRNAs have any phenotypic or prognostic significance. For those miRNAs which were observed to be downregulated miR-31 and miR-181b, they have been associated with HNC survival and metastasis respectively [112, 113]. For the upregulated miRNAs, miR-9 is enriched in HPV+ HNC exosomes and associated with radiosensitivity and survival [114, 115]. High levels of miR-363 in HNC were associated with better overall survival [116-118].

1.4.9. Manuscript: An HPV-interactome for understanding drug mechanisms of action

Interactomes can be utilised to map potential gene targets to compounds with the aim of repurposing known drugs for new OPC treatments. For example, we can trace the pathway from miRNA to gene target, then match it to a known drug for possible applications (Figure 1.4). Figure 1.4 was developed using methods from Hill *et al.* [102] as well as online drug databases such as Drugbank [119]. Through this association, the drugs Alvocidib, 5-Fluorouracil, Camptothecin, Alanine, Curcumin, Estrogen, Etoposide, Resveratrol, Polypropylene Glycol and Butyrate all possess binding capacities for the HPV regulated genes involved in cell cycle regulation (E2F1, CDK1, CDK2, CDK4, EP300, RAN, CAV1, PTPN1). For example, Alvocidib acts by inhibiting

cyclin-dependant kinases, arresting cell division and leading to apoptosis [120]. CDK1, CDK2 and CDK4 are all targets of the miRNAs from the HPV interactome and potential targets for Alvocidib in OPC. The drug 5-Fluorouracil is a common agent in chemotherapy which interferes with DNA synthesis. This would have an impact on numerous cancer genes such as RBL1, CDKN1A, CDK1, CDK2, CDK4, E2F1, BRCA1, BAK1, TP53, ABBC1, which are regulated by miRNAs in HPV+ OPC (Figure 1.4). Furthermore, 5-Fluoroucrail is being interrogated in de-escalation clinical trials of HPV related HNC (NCT00122460, NCT03107182, NCT03944915, NCT01221753), with the aim to reduce the dose of radiation combined with standard chemotherapy [121].

Interestingly, the fact that many of the targets of HPV regulated miRNAs may also be targets of currently used drugs could explain why HPV16 OPC patients are better responders to current therapies. The use of an HPV-miRNA interactome could further our understanding of chemoresistant genes. Of note, the ATP binding cassette transporter ABCC1 involved in multi-drug resistance is a potential target of miR-9-5p. ABCC1 has previously been described to be upregulated by HPV18 [122]. From the mapping, ABCC1 could be targeted by several drugs such as Alvocidib, 5-Fluorouracil, Camptothecin and Etoposide, or the use of miRNA mimics to reduce the expression of ABCC1. The overall clinical aim would be to neutralise this protein transporter and render the cell sensitive to drug compounds. Thus, the use of interactomes in drug development has the potential to a) link is known drugs to possible new gene targets and b) to provide a clearer understanding of the mechanism of action of known or new drugs at the molecular level. With this information, we can develop better clinical strategies regarding dosage and identify the side effects of new and known compounds.

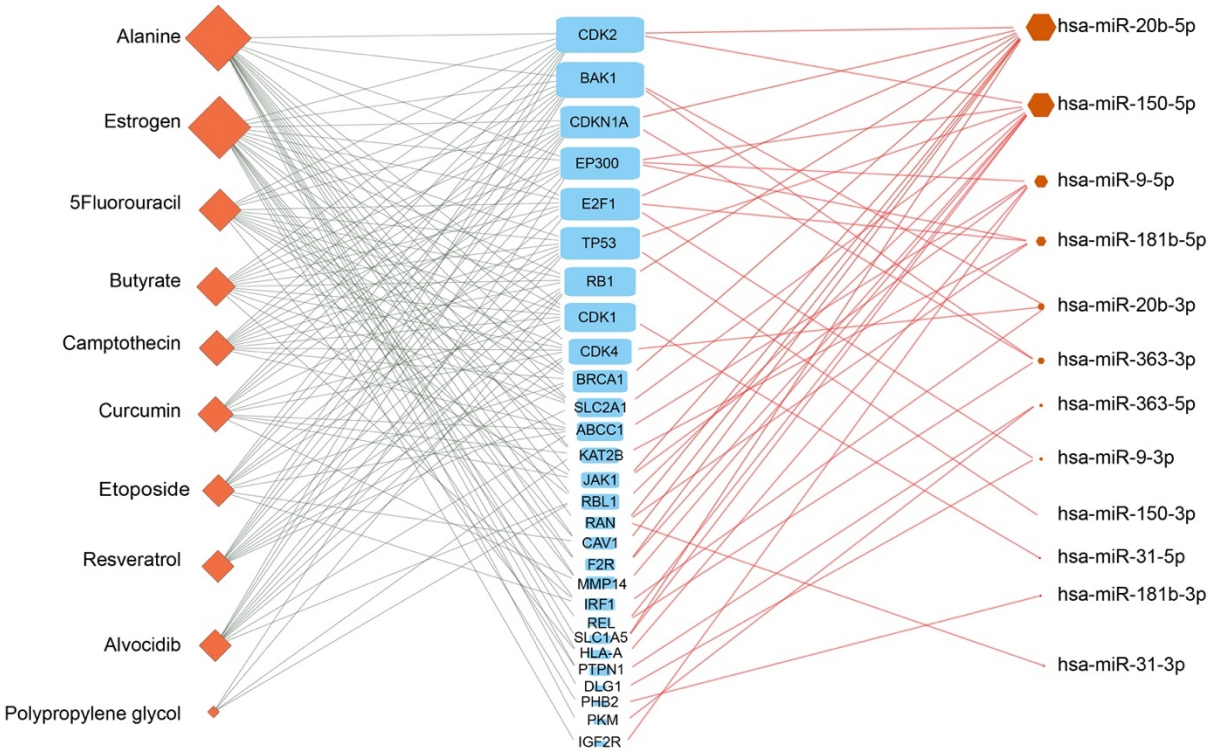


Figure 1. 4: Using interactomes to map chemotherapeutic drug associations with mRNAs targets. Drugs database can be compiled into these interactomes to visualise the possible mode of drug action, understanding side effects and evaluating or repurposing known drugs

1.4.10. Manuscript: Conclusion

HPV is a major risk factor for OPC, with high-risk HPV16 accounting for an overwhelming number of cases, posing a major global health concern. It has been demonstrated that HPV16 OPC is a separate entity to virus-free OPC, with HPV16 patients having better overall survival.

The pathways for HPV mediated transformation are numerous, and its interactions with specific proteins are well understood, but the interaction between HPV and ncRNAs is an exciting area. By integrating expression data and protein associations into a visual interactome, we can identify key target gene hubs and new associations in HPV16 OPC. The marrying of these interactions with drug databases may also pave the way forward to design possible new drugs to target key hub genes, re-purpose existing drugs to target specific genes or evaluate side effects by understanding which genes are most affected. Given the very high cost of drug development and clinician trials, this may be a feasible avenue to screen possible targets.

We noted in the review that very few miRNAs were consistent across the different studies. This is mainly due to variations in sampling handling and the different technologies used to measure the expression of these miRNAs. Perhaps, interactomes can be used to map all these mRNAs to identify hub genes and miRNAs. This would, to a certain point, address the variable published findings and create a platform for further analysis. These interactomes, if used in the correct framework, have the potential to integrate expression data from various sources into a single coherent gene mining tool.

There are other species of ncRNAs that can be included in these viral interactomes. The long non-coding RNA family, which accounts for 80% of the human genome, can act as miRNA sponges or protein tethers to modulate function. The inclusion of these long ncRNAs would provide a greater genome-wide perspective for HPV16 regulation of the RNA world. Adding in new components such as other ncRNAs other viruses such as EBV, these atlases can become a visual repository for inspecting viral gene interactions.

In this review, we summarised all the pertinent miRNAs describe for HPV16 OPCs and reviewed our understanding of these regulatory RNAs by using a viral interactome. Key miRNAs and hub genes regulated by HPV16 can be identified using this visual tool. Merging a drug database may yield important information for developing alternative treatments. One of the obvious caveats is that interactomes rely on expression data generated by published studies. This fact must also be considered when interrogating these interactomes. The role or impact of HPV16 for virally driven oncogenesis in Head and Neck cancers must be viewed from a wider gene regulatory perspective and include other forms of non-coding RNAs such as long non-coding RNAs. We have surpassed the basic model of an oncogene regulating a single target. The construction of these dynamic interactomes will drive new basic and clinical developments in Head and Neck research.

1.5. Long non-coding RNAs

It has been demonstrated that HPV can alter the expression of miRNAs in numerous cancers, including HNSCC. This alteration to miRNAs has a severe effect on its downstream targets and the overall functioning of a cell. Given that there are many other forms of gene regulation by non-coding RNAs, such as long non-coding RNAs, we need to also include these in our analysis. There is currently little research into the effects of HPV on long non-coding RNAs in HNSCC, and the majority of these studies have focused on cervical cancers.

1.5.1. Long non-coding RNAs: discovery, biogenesis, and function

Long non-coding RNAs (lncRNA) are 200 nucleotides or greater in length [45] with a vast array of functions. LncRNAs can form secondary structures, enabling them to interact with DNA, RNA and protein molecules. The act in transcriptional and post transcription regulation and as regulators of chromatin organisation. LncRNAs are usually poorly conserved and are expressed at relatively low levels, with cell type-specific expression in response to a range of stimuli. The lncRNA expression pattern depends on cell state, differentiation, and disease development. LncRNAs may act as enhancers, decoys or scaffolds by physically interacting with proteins and other RNA types, resulting in changes to cell signalling cascades.

Any changes such as deficiency, overexpression or mutations to lncRNAs have been associated with multiple human diseases. Currently, there are 17944 lncRNA genes (GENCODE v38). For many lncRNAs, their function remains unknown, highlighting the importance of research into their mechanistic function and their role in cancer.

In the 1990s, the first cellular lncRNAs were discovered, the H19 and X-inactive specific transcripts (XIST) [123, 124]. LncRNA expression studies in mice also revealed that they have precise expression patterns dependent on tissue type, cell type and subcellular compartments [125]. Initially, lncRNAs were described as mRNA-like transcripts that do not encode proteins, but recent studies have assisted in distinguishing between lncRNAs and mRNA. In terms of their biogenesis, lncRNAs are very similar to mRNAs. They are transcribed by RNA polymerase II (some are

transcribed by RNA polymerase III) [126]. LncRNAs often are 5'-capped with a promoter structure and are 3`polyadenylated. These non-coding RNAs can also have a multi-exonic composition and undergo splicing into different isoforms with distinct functional outcomes [127, 128].

Distinguishing features of lncRNAs is generally they are shorter than mRNA with fewer exons, and they are under weaker selective constraints during evolution [129]. The majority lack a conserved open reading frame [130], although a few examples of lncRNAs contain cryptic open reading frames [131].

LncRNAs are often in very low abundance, highlighting the difficulty in characterising them. The expression levels of lncRNAs are highly cell type and tissue-specific. Many lncRNAs also have very well-defined subcellular localisation in either the cytoplasm or nucleus. For example, the lncRNA named XIST is localised to the inactive X chromosome [132]. Regarding nuclear versus cytoplasmic localisation, lncRNAs are enriched in the nucleus with functions in epigenetic regulation of chromatin. More lncRNAs by transcript numbers are present in the cytoplasm [133].

The synthesis of lncRNAs can be classified into several ways, based on their genomic proximity to their neighbouring transcripts. These classifications are; sense strand synthesis, antisense strand synthesis, intergenic synthesis and intronic synthesis [134]. Intergenic lncRNAs are transcribed from intergenetic regions, and those transcribed from introns of protein-coding genes are termed intronic lncRNAs.
Sense lncRNAs are transcribed from the sense strand of protein-coding genes, containing exons from those genes and overlapping parts of the protein-coding. Antisense lncRNAs are transcribed from the antisense strand of protein-coding genes.

In order to understand the impacts of lncRNA on cancer initiation and development, we must understand their development, structure and activity in a cell. LncRNAs are highly diverse and can perform a multitude of functions, highlighting their importance as essential gene regulators

1.5.2. The classification of long non-coding RNAs

There are various mechanisms by which lncRNAs carry out their function that are not fully understood. Many lncRNAs have been characterised, and several modes of actions have been proposed, such as guide, signalling, decoy, enhancer or scaffold lncRNAs [135].

Signalling lncRNAs act as a molecular signal to regulate transcription, usually in response to stimuli [135]. Examples of signalling lncRNAs include Kcnq1ot1 and Air, which mediate transcriptional silencing of multiple genes through the interaction with chromatin to recruit chromatin-modifying machinery [136, 137]. Guide lncRNAs interact with ribonucleoprotein complexes, directing them to their specific target genes, playing an essential role in localisation [135]. LncRNAs can guide changes in gene expression in either the *cis* (neighbouring genes) or *trans* (distantly located genes) [136]. Decoy lncRNAs reduce the availability of regulatory factors such as transcription factors, chromatin-modifying complexes, miRNAs and catalytic proteins [135]. Decoy lncRNAs do this by binding to these molecules, preventing them from exerting their function.

An example is the tumours suppressor pseudogene PTENP1 which acts as a decoy for PTEN targeting miRNAs. These miRNAs will bind to PTENP1, reducing their regulatory control of PTEN [138]. Enhancers lncRNAs influence chromatin interactions or spatial organisation of DNA [135] and scaffolding lncRNAs have a structural role by providing a means for assembly of multiple proteins to form ribonucleoprotein complexes [135]. For example the lncRNA HOTAIR can associate E3 ubiquitin ligases and its ubiquitination substrates to facilitate ubiquitination of Ataxin-1 and Snurportin-1 to accelerate their degradation [139]. Figure 1.5 is a schematic of the different lncRNA archetypes and their functions. LncRNAs can also be classified as non-functional lncRNAs that may result from transcriptional noise [133].

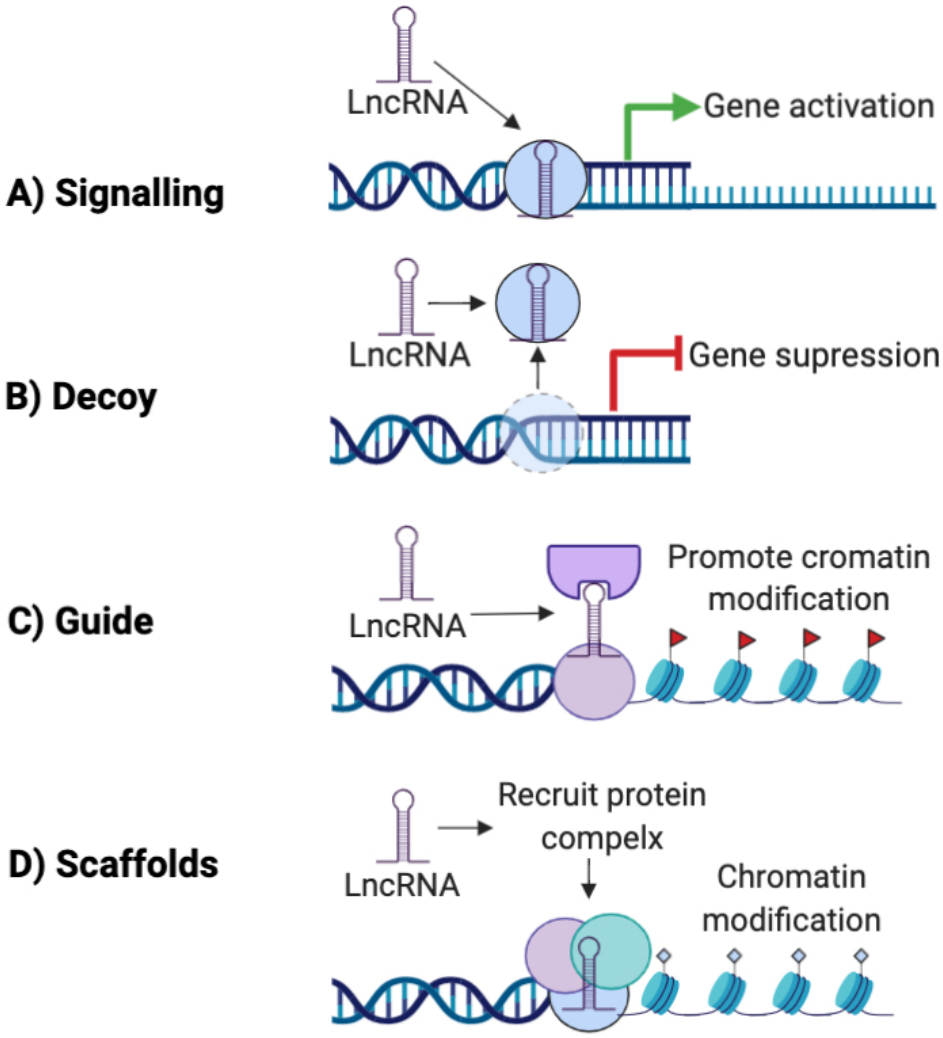


Figure 1. 5: The main archetypes of long non-coding RNA classification: A) Signalling lncRNAs represent those that can lead to the activation of gene transcription. B) Decoy lncRNAs suppress or activate gene expression, generally acting as a sponge. C) Guide lncRNAs promote chromatin modification by assisting protein binding, D) Scaffold lncRNAs act as components to assist multiprotein complexes [135].

It is evident that there is a wide range of cellular processes which lncRNAs can regulate. This regulation includes those associated with cancer development, growth, migration, invasion and differentiation [130]. This is very little research into lncRNAs role in HNSCC and even less into HPV related HNSCC. If we want to gain a holistic understanding as to how HPV influences all molecular pathways within a cell. We would also need to investigate the different types of non-coding RNAs, such as microRNAs and lncRNAs. Especially as non-coding RNAs have been considered potential targets for gene therapies due to their specific functions and specific expression relative to disease. Uncovering the gene expression networks, such as lncRNAs, associated with cancer initiation and development is critical but remains largely uncompleted.

1.5.3. The role of long non-coding RNAs in cancer

The dysregulation of Long non-coding RNA expression has been associated with many human diseases, but our focus is their role in cancers. LncRNAs are involved in cancers such as breast cancer [140], Glioblastomas [141], colorectal cancer [142], liver cancer [143] and Head and Neck cancer [144]. LncRNA involvement in cancer progression has been associated with cancer processes such as tumour initiation, growth and metastasis. A summary of well-studied cancer-related lncRNAs is described in Table 1.3

In a comprehensive analysis on lncRNA alterations in 5037 human tumours across 13 different cancer types, Yan *et al*. [129] determined that the dysregulation in expression pattern of lncRNAs is highly specific for the cancer type. Only a handful of lncRNAs showed dysregulation across all cancer types analysed. LncRNAs identified to have dysregulated expression across numerous cancer types were PCAT7, PVT1 and HOTAIR, which were upregulated in various cancer types. This highlights the uniqueness of the lncRNA expression network for each cancer type, showing we need to investigate them as specific entities.

The dysregulation of lncRNAs in cancer can occur in several ways, such as genetic and epigenetic changes. Dysregulation due to genetic and epigenetic modifications includes chromosomal translocations, copy number alterations and single nucleotide polymorphisms (SNPs). Emerging studies describe an association between mutations

in lncRNA genes and cancer. In osteosarcomas, copy number alterations are frequently seen in the lncRNA LOC285194 and BC040587, resulting in their loss of expression. Genetic modifications to LOC285194 was associated with poor survival in patients [145]. SNPs also contribute to dysregulated lncRNA expression in cancer. In prostate cancer, SNPs between rs1456315 and rs7463708 are associated with cancer susceptibility and expression of the oncogenic lncRNA PRNCR1 [146]. Many more lncRNAs have been shown to have dysregulated expression in cancers due to epigenetic changes.

LncRNAs may also be dysregulated in cancers due to regulation by specific oncogenic and tumour suppressive related signals and regulatory factors. For example, lincRNA-p21 acts as a repressor in p53-dependent transcriptional response [147]. Hung *et al.* [148] identified 216 putative lncRNAs derived from the promoter region of cell cycle genes, which showed periodic expressing during the cell cycle and had altered expression in human cancers. LncRNAs also act as decoys for miRNAs to prevent them from carrying out their functions. This role also has implications in cancer development. GUARDIN is a lncRNA transcribed from the promoter region of the p53 target miRNA miR-34a and is activated upon DNA damage response, promoting cell survival and controlling genome stability.

The uniqueness of lncRNA profiles in cancer could be applied for HNSCC research. LncRNA expression profiles could be used to assist in distinguishing between HPV+ and HPV- HNSCC. Traditional methods for the detection of HPV in HSNCC is identification of HPV genetic material and p16 staining. P16 staining and detection of HPV DNA do no indicate if there is a transcriptionally active viral infection. These methods have also been shown to have discrepancies in results with low sensitivity rates [149-151]. The gold standard for HPV testing is quantitative PCR detection of viral E6/E7 mRNA, although this is labour intensive and expensive [151]. LncRNAs may be utilised in combination with traditional methods to increase specificity and sensitivity for the detection of HPV in HNSCC. In other cancer types lncRNAs have also been shown to distinguish between subgroups within cancer patient cohorts identifying response to treatment, progression of disease, overall survival rate and identify chemo-resistance [152-154].

Table 1. 3: Examples of specific lncRNAs and their associated impact on the hallmarks of cancer

LncRNA	Associated Hallmark of cancer	Associated cancers	Reference
HOTAIR	Increased expression leads to increased invasion metastasis and promotes drug resistance.	Breast, lung, liver, colorectal, cervical, ovarian, prostate, gastric, endometrial, Head and Neck	[155-158]
MALAT1	Overexpression increased metastasis, proliferation and invasion of tumour cells.	Gastric, gallbladder, oesophageal, pancreatic, adenocarcinoma, liver, colorectal, glioblastoma, cervical, ovarian, breast, urothelial, bladder	[159]
NEAT1	Elevated levels correlate with increased proliferation, invasion and metastasis. Knockdown in laryngeal cancers suppressed cell proliferation and increased apoptosis	Lung, oesophageal, laryngeal, colorectal, liver	[160]
GAS5	In breast, cancer downregulation is associated with more aggressive cancers. GAS5 sponges miR-196a-5p suppressing cell invasion. Metabolism reprogramming by suppressing the transcription of glucocorticoid receptors. repression of anti-apoptosis genes	Bladder	[161, 162]
MEG3	Low expression in cervical cancers promotes hypermethylation, and in colorectal cancers, increased cell invasion and metastasis	Colorectal, cervical, renal, liver, lung, thyroid, gastric, ovarian, urothelial, tongue, bladder	[163]

	Inhibits cell proliferation through TP53 regulation and inhibition of Cyclin D1 to induce cell cycle arrest Induces apoptosis by inhibiting BCL-2 expression		
PANDA	Associated with invasion and metastasis. Higher expression in Colorectal and renal cancers showed to have poorer overall survival. Knockdown in colorectal and renal cancers decreased cell proliferation, cell cycle progression, migration and invasion. Decrease of PANDA as a result of decrease TP53 led to increased proliferation but decreased apoptosis in lung cancers	Colorectal, lung, renal, breast, liver, osteosarcomas, thyroid, gastric, bladder,	[164]
XIST	Higher expression associated with increased cell proliferation and metastasis	Gastric, liver, nasopharyngeal, lung, glioma, pancreatic, cervical, osteosarcoma, colorectal, breast, ovarian	[165]
ANRIL	Higher expression associated with increased metastasis	Colorectal, nasopharyngeal, gallbladder, liver, lung, cervical, gastric, ovarian, thyroid	[166]

1.5.4. The impact of HPV on long non-coding RNAs expression

Cancers related to HPV have been shown to have altered molecular profiles due to the activity of the viral infection. LncRNAs altered by viruses have been described as both pro-viral, enhancing invasion of the virus and viral life cycle [167, 168], as well as anti-viral lncRNAs which regulate innate and adaptive immune response [169]. Due to the site-specific response of lncRNAs to various stimuli, HPV infection plays a part in lncRNAs expression. HPV has been shown to alter the expression of lncRNAs in cervical cancers, including CCEPR, MEG3, EGFR-AS1 [170], TMPOP2 [171], HOTS2 [172] and FAM83H-AS1 [173]. The effect of lncRNA expression dysregulation, induced by HPV, impacts many cellular pathways to promote tumour growth in this cancer type [170].

The Cervical carcinoma expressed PCNA regulatory (CCEPR) lncRNA has been shown to be highly upregulated in cervical cancers [174] as well as in HFKs expressing HPV16 E6 [175]. The overexpression of CCEPR has been associated with increased cellular proliferation [174, 175] and poor patient prognosis [176]. MEG3, the maternally expressed gene 3 lncRNA, is lowly expressed in cervical cancer tissue and correlates with HPV expression [177]. MEG3 overexpression in cervical cancer cell lines led to inhibition in cell proliferation, increased apoptosis, and reduced tumorigenicity in xenograft models [178].

Cervical cancer patients who respond well to erlotinib therapy tend to have a higher level of the lncRNA EGFR-AS1. In HPV16 E6/E7 transformed cells, the levels of this lncRNA are reduced compared to normal cells and result in a worse response to erlotinib [170]. HPV16 E6/E7 can induce an increase in the expression of TMPOP2 [171] in cervical cancers. TMPOP2 can sequester the tumour repressor miRNAs, miR-375 and miR-139, which target HPV16 E6/E7 mRNA, resulting in their upregulation. Knockdown of TMPOP2 reduced cell cycle genes, induced cell cycle arrest, and inhibited cell proliferation of cervical cancer cells [171]. HOST2 can inhibit apoptosis and promote cellular proliferation of cervical cancers via the blockage of the miRNA, let-7b. HOST2 is high in HPV+ cervical cancers, and let-7b is low [172].

Finally, the lncRNA FAM83H-AS1 plays a role in the apoptotic process in cervical cancer cells, with its overexpression in HPV16 positive cervical cancers. This lncRNA is regulated by the E6-Ep300 pathway independently of p53 action [173]. It is evident the HPV virus and its major viral oncogenes E6/E7 can induce changes in lncRNA expression resulting in alterations to the hallmarks of cancers.

1.5.5. Long non-coding RNAs and their role in Head and Neck cancer

LncRNA regulation and their role in tumour development are understudied compared to miRNAs. The lncRNA expression profile for each cancer type are also unique, and the HPV virus has been shown to alter lncRNAs expression in cervical cancers. Hence it is important to investigate the role lncRNAs in HNSCC, both HPV16 positive and negative.

Gibb *et al*.[179] provided the first lncRNA expression map of the oral mucosa. Of the 325 lncRNAs found to be expressed in the oral mucosa, 60% showed differential expression in premalignant lesions suggesting they may play a role in tumorigenesis. Since then, there have been numerous studies describing lncRNA differential expression oral squamous cell carcinomas [180, 181] as well as other HNSCC subtypes such as laryngeal [182], tongue [183] oesophageal [184].

Several studies have determined that the expression levels of lncRNAs can influence HNSCC cell proliferation, migration and apoptosis, all pathways heavily altered in cancer development. For example, the lncRNA FOXCUT (FOXC1 promoter upstream transcript) is overexpressed in OSCC along with its adjacent gene FOXC1. Knockdown of FOXCUT leads to the downregulation of FOXC1 and a reduction in cell proliferation and cell migration [185]. Fang *et al*. [183] identified the lncRNA UCA1 to be upregulated in Tongue SCC (TSCC), and its overexpression was correlated with lymph node metastasis.

MALAT1 has been heavily investigated and is involved in metastasis in HNSCC. Metastasis associated lung adenocarcinoma transcript 1 (MALAT1) was first detected in non-small cell lungs, with a high expression level, contributing to cell migration and invasion [186]. In Oral SCC, MALAT1 has been shown to be overexpressed, serving

as a new prognostic factor of OSCC. Knockdown of MALAT1 demonstrated that the lncRNA is required for epithelial-mesenchymal transition (EMT) mediated cell migration and invasion [187]. MALAT1s involvement in EMT, invasion and migration has also been demonstrated in tongue SCC [188].

Similarly, to miRNAs, very few studies have investigated the impact HPV has on lncRNA expression in HNSCC specifically. Considering that HPV+ HNSCC is recognised as a separate entity to HPV- HNSCC, it is important to understand the impact HPV has on lncRNAs in this cancer type. Nohata *et al*. [144] analysed the TCGA data and were the first to compare lncRNA expression in HPV+ to HPV- HNSCC (includes the subtypes: Pharyngeal, Laryngeal and Oral), identifying 140 lncRNA transcripts to be differentially expressed when the virus was present. Of which LINC01305, LINC0189 and PTOV1-AS1 were significantly upregulated in the HPV+ tumours from the TCGA. TP53 is frequently mutated in HNSCC due to tobacco use or by degradation by HPV, hence why it is crucial to determine lncRNAs associated with TP53. Nohata *et al*. [144] found 30 lncRNAs with differential expression between TP53 mutated and wild type tumours. Among these lncRNAs, 19 transcripts were associated with HPV+ tumours. Although it is essential to identify lncRNAs with altered expression due to HPV, it is also important to understand the impact of this change in expression.

Salayakina and Tsinoremas [189] analysed differentially expressed lncRNAs from HPV+ HNSCC in the TCGA cohort but also investigated lncRNA predicted mRNA targets. The majority of the lncRNAs in their study showed a positive correlation in expression with their predicted targets in the same loci, and many tumour suppressors and oncogenes served as potential targets for differentially expressed lncRNAs. Recently Song *et al*. [190] also analysed the TCGA dataset comparing lncRNA expression profiles between HPV+ and negative patients. They identified 177 lncRNAs with HPV associated differential expression (75 upregulated and 102 downregulated) of these 8 lncRNAs (RP11-635N19.3, CCDC144NL-AS1, AC007879.2, RP11-30P6.6, AC104534.2, CD81-AS1, AC006946.16 and LINC00504) that are associated with HPV infection and improved prognosis. The lnc-IL17RA-11 was associated with a better response to radiotherapy. This lncRNA has targets enriched for genes related to the functions such as cell cycle, p53 signalling and DNA replication. HPV increases the expression of the estrogen receptor alpha, which acts as a transcription factor to

lnc-IL17RA-11, increasing its expression. This promotes the co-expression of genes associated with these cancer pathways, increasing the cells' sensitivity to radiation. These findings could suggest why HPV-positive HNSCC are more sensitive to radiotherapy; perhaps lncRNAs play a significant role in this clinical outcome.

More recent studies on lncRNAs in HPV HNSCC Cu *et al.* [191] defined a signature of 15 lncRNAs with prognostic significance for recurrence-free survival for both HPV+ and negative cohorts. EGOT [192] eosinophil granule ontogeny transcript lncRNAs upregulated in HPV+ HNSCC, its mechanism was not reported.

Finally, Kopcyznska *et al.* [169] determine the impact of HPV infection on the immune systems, particularly on the lncRNA PRINS, psoriasis susceptibility related RNA gene induced by stress. PRINS was upregulated in HPV+ cancers compared to HPV- and was associated with better overall survival. Higher levels of PRINS resulted in changes in the expression of genes related to the immune and antiviral response, as well as a higher number of immune cells within tumours. Kopczynska et al. [169] also observed other differentially expressed lncRNAs in the HPV+ cells, including CDKN2B-AS1 (up), TTTY14 (up), TTTY15 (up), MEG3 (down) and H19 (down).

We have only scratched the surface of the role lncRNAs in HPV+ HNSCC. LncRNAs have many different modes of regulation. One function we are interested in is lncRNAs acting as decoys, more specifically as sponges for miRNAs. Numerous miRNAs are differentially expressed in HPV+ HNSCC, but the cause for the dysregulation for many of these miRNAs is not fully understood, and perhaps lncRNAs play a role. The function of lncRNA acting as miRNA sponges has not yet been described in HPV+ HNSCC. We aim to determine the impact of HPV16 on lncRNAs in Oropharyngeal cancers specifically and determine the downstream effect this is on lncRNA targets such as microRNAs.

Table 1. 4: LncRNAs dysregulation in Head and Neck Squamous Cell Carcinomas and their function

LncRNA	HNSCC subtype	Function	Expression in HNSCC	Reference
NEAT1	Oral (HPV-) Laryngeal (HPV-)	higher expression in metastatic tumours and later clinical-stage, knockdown reduced cell proliferation, induce apoptosis and cell cycle arrest at G1 phase, regulates CDK6	Up	[193, 194]
FOXCUT	Oral (HPV-) Nasopharyngeal (HPV-)	Overexpressed with its adjacent gene FOX1, regulates cell proliferation and invasion	Up	[128, 195, 196]
UCA1	Tongue (HPV-) Hypopharyngeal (HPV-)	Overexpression correlates with lymph node metastasis Functions as an oncogene promote cell proliferation invasion and prevent apoptosis	Up	[183, 197]
MALAT1	Oral (HPV-) Tongue (HPV-)	Metastasis of HNC, associated with the epithelial-mesenchymal transition to mediated migration and invasion	Up	[187, 188]
LINC00152	Tongue (HPV-) Laryngeal (HPV-) Oral (HPV-)	High expression associated tumour progression, invasion and metastasis promote cell cycle progression, apoptosis, EMT.	Up	[198-200]
LINC01305	HPV+ HNSCC	Not characterised	Up	[144]

lnc-IL17RA-11	HPV+ HNSCC	Associated with improved response to radiotherapy, targets related to p53 signalling and DNA replication. Induces genomic stability	Up	[190]
PTENP1	HNSCC	Low expression is associated with a history of alcohol use and worse overall survival. PTENP1 overexpression inhibited cell proliferation, colony formation and migration of cells and tumours	Down	[201]
EGOT	HPV+ HNSCC	Activates invasion and metastasis. Promotes cell proliferation	Up	[192]
MEG3	HNSCC (HPV-) Nasopharyngeal (HPV-)	Overexpression inhibited cell proliferation, migration and invasion. Regulates EMT by sponging miR-421, can inhibit autophagy by regulating miR-21 and PTEN	Down	[169, 202, 203]
PRINS	HPV+ HNSCC	Associated with better overall survival, associated with genes involved in immune and antiviral response. Upregulated increase immune cell numbers within tumours	Up	[169]

1.6. Hypothesis

We put forward the notion that the HPV16 oncogenes E6 and E7 will dysregulate the expression of non-coding RNAs, specifically miRNAs and long non-coding RNAs, in Oropharyngeal cancers, and this will result in a downstream effect on their specific targets.

1.7. Aims

1. To develop a method for visualising and interrogating large genomic datasets using an HPV-non-coding RNA interactome. This approach will discern potential regulatory pathways in HPV+ OPC and lay the foundation for future mechanistic studies.
2. To establish a direct regulatory role between HPV16 E6 and E7 and specific miRNAs and their target genes in OPC.
3. To explore the role of lncRNAs in cancer and determine their possible impact on Head and Neck cancer.

.

Chapter 2: Materials and methods

2.1. Materials

Below are lists of reagents, commercially available kits, TaqMan probes and antibodies used in this study including the catalogue number and manufacturer.

Table 2. 1: Commercially available kits and related reagents used in this study

Kit/Reagents	Catalogue number	Manufacturer
Dual-Luciferase® Reporter 1000 Assay System	E1980	Promega, Australia
ECL Western Blotting Substrate	W1001	Promega, Australia
Fast SYBR™ Green Mastermix	4385612	ThermoFisher Scientific, Australia
High-Capacity cDNA Reverse Transcription	4374966	ThermoFisher Scientific, Australia
Multiscribe™ Reverse Transcriptase	4311235	ThermoFisher Scientific, Australia
Plasmid Midi Kit	12143	Qiagen, USA
QuantiTect® Reverse Transcription Kit	205311	Qiagen, USA
Qubit™ Protein Assay Kit	Q33212	ThermoFisher Scientific, Australia
Quick-RNA™ Miniprep Kit	R1055	Zymo Research
Random Primers	48190011	ThermoFisher Scientific, USA
TaqMan Fast Advanced Master Mix	444557	ThermoFisher Scientific, Australia
TaqMan Universal PCR Master Mix	4324018	ThermoFisher Scientific, Australia

Table 2. 2: Reagents used in this study

Item	Catalogue number	Manufacturer
20x Bolt™ Bis-Tris Transfer Buffer	BT0006	Invitrogen, ThermoFisher Scientific, Australia
20x Bolt™ MES-SDS Running Buffer	B0002	Invitrogen, ThermoFisher Scientific, Australia
Ampicillin	A0166-5G	Sigma-Aldrich, Australia
BAN Phase Separation Reagent (1-Bromo-4-methoxybenzene)	BN191	Molecular Research Centre, USA
Bolt™ (4-12%) Bis-Tris Plus Mini-Gel	NW04120BOX	Invitrogen, ThermoFisher Scientific, Australia
Bovine Serum Albumin	A7906	Sigma-Aldrich, Australia
Cell Lytic M Cell Lysis reagent	C2978	Sigma-Aldrich, Australia
Dimethyl Sulfoxide (DMSO)	67-68-5	Sigma, Aldrich, Australia
Dulbecco's Modified Eagle Medium (DMEM), high glucose, GlutaMAX™ Supplement	10566016	Gibco, ThermoFisher Scientific, Australia
Dulbecco's Modified Eagle Medium (DMEM), no glucose	11966025	Gibco, ThermoFisher Scientific, Australia
Ethanol	64-17-5	Sigma-Aldrich, Australia
Foetal Bovine Serum (FBS)	10082147	Gibco, Thermofisher Scientific, Australia or Invitrogen, USA
Glycogen	AM9510	ThermoFisher Scientific, Australia
Hyper ladder 1KB	BIO-33053	Bioline, UK
Isopropanol	190764	Sigma-Aldrich, Australia
Keratinocyte Serum free medium (KSFM)	17005042	Gibco, ThermoFisher, USA
KSFM Supplement: Human recombinant epidermal growth factor (EGF1-53) and Bovine Pituitary Extract (BPE)	37000015	Gibco, ThermoFisher, USA
Lipofectamine® 2000	11668019	ThermoFisher Scientific, Australia
Lipofectamine® 3000	L300015	ThermoFisher Scientific, Australia Invitrogen??
Lipofectamine® RNAiMax	13778150	ThermoFisher Scientific, Australia
Luria-Bertani (LB)	J106-1KG	Ameresco, USA
Methanol	5005-10L	ThermoFisher Scientific, Australia

Negative control DsiRNA, 5nmol	51-01-14-04	Integrated DNA Technologies
ON-TARGET plus Non-targeting Pool	D-001810-10-05	Thermo Scientific Dharmacon
Opti-MEM	31985070	Gibco, ThermoFisher Scientific, Australia
PageRuler™ Plus Pre-stained Protein Ladder	26619	Thermofisher scientific, Australia
Penicillin Streptomycin	15070063	Gibco, ThermoFisher, USA
Phosphate Buffered Saline (PBS)	P4417-50TAB	Sigma-Aldrich, Australia
Resazurin	R7017	Sigma-Aldrich, Australia
RNAase-free water	10977015	ThermoFisher Scientific, Australia
RNAzol RT	RN190-200ml	Molecular Research Centre, USA
Serine protease inhibitor	A6664-10MG	Sigma-Aldrich, Australia
TO-PRO3 Iodide	R37170	Invitrogen, USA
TrypLE	12563029	Gibco, ThermoFisher Scientific, Australia
Tween-20	P7949	Sigma-Aldrich, Australia
TYE563™ Transfection Control DsiRNA 5nmol	51-01-20-20	Integrated DNA Technologies

Table 2. 3: TaqMan probes used in this study (Applied Biosystems, ThermoFisher Scientific, USA)

miRNA/Gene	Catalogue
ACTB	Hs01060665_g1
B2M VIC™ Labelled	Hs00187842_m1
E2F2	Hs00918090_m1
GAPDH	Hs02786624_g1
hsa-miR-33a	002135
hsa-miR-496	001953
RNU6B VIC™ Labelled	001093
SREBF2	Hs018081784_m1

Table 2. 4: Antibodies used in this study for western blotting

Antibody	Catalogue number	Manufacturer
ACTB-HRP	Ab20272	Abcam, USA
Anti-rabbit-HRP	7074	Cell Signal Technologies, Australia
GAPDH	Ab9382	Abcam, USA
HPV16 E6	Ab70	Abcam, USA
SREBF2	Ab30682	Abcam, USA

2.2. Methods

2.2.1. Cell lines

This study used the following cell lines for transfections, SCC090 (HPV16+, Tongue SCC) and SCC154 (HPV16+, Tongue SCC), SCC4 (HPV- tongue SCC), SiHa (HPV16+ cervical SCC) and Caski (HPV16+ cervical SCC).
All these cell lines were maintained using DMEM media (Invitrogen, USA), supplemented with 10% Foetal Calf Serum (FCS) (Invitrogen, USA) and GlutaMAX™, (Invitrogen, USA). These cell lines were grown at 37°C in a humidified incubator (Binder, Germany) with a 5% CO_2 environment.

Immortalised Human Foreskin keratinocytes(iHFKs) and immortalised Normal oral keratinocytes (iNOKs) were a gift from Dr. Karl Munger. The iHFKs and iNOKs were grown and maintained in keratinocyte serum-free medium (KSFM) supplemented with human recombinant epidermal growth factor and bovine pituitary extract (Invitrogen, USA). HPV16 E6/E7 expressing iHFKs and iNOKs were created by retroviral infection with recombinant retroviruses, provided by Karl Munger [204, 205]. Both control and HPV16 E6/E7 expressing iHFKs and iNOKs were used in this study.

HCT116 colorectal cell lines cells lines were included wild type p53 and p53 null cells, a gift from Dr. Karl Munger, cells were grown in McCoy's 5A media (Invitrogen, USA) supplemented with 10% FCS (Invitrogen, USA). These cell lines were grown at 37°C in a humidified incubator (FORMA Scientific) with a 5% CO_2 environment.

2.2.2. Designing HPV16 E6/E7 and TRINGS siRNAs and transient reverse transfection of siRNA

For the HPV16 viral oncogenes E6 and E7 alone and in combination, siRNAs were designed and purchased from Integrated DNA Technologies (IDT). Predesigned siRNA targeting SREBF2 were purchased from IDT along with a non-template binding control. Sequences for these all siRNAs are described in their corresponding chapters (Chapter 5 and Chapter 7).
Several siRNAs targeting TRINGS were purchased from Dharmacon™, (ON-TARGETplus SMARTpools) sequences were extracted from Khan *et al.* [206] and are

displayed in relevant chapter (Chapter 6). A non-template binding control was also purchased from Dharmcon™.

For SREBF2, HPV16 and TRINGS siRNA transfections, cells were seeded in a 12-well plate (Invitrogen, USA), 24hrs prior to transfection at $5.5x10^5$ cells/mL. To each well, 1mL of the desired cells at a predetermined concentration was added. 1mL of appropriate media was also added to each well to give a final volume of 2mL.

After 24hrs, transfection was performed using the Lipofectamine® 3000 Reagent Protocol (Invitrogen, USA). Each well was transfected with 20pmol of desired siRNA and 2.0μL of RNAimax (Invitrogen, USA). The transfected cells were then incubated at 37°C with 5% CO_2 in a humidified incubator. 24 and 48hrs after transfection the RNA and proteins were isolated from the SCC090 and SCC154 cells. RNA was harvested and cell viability was measured 24, 48, 72, 96hrs post transfection for the SiHa, Caski, iHFKs, iNOKs and HCT116 cell lines.

2.2.3. HPV16 E6/E7 plasmid overexpression

2.2.3.1. Plasmids preparation from bacterial stab cultures

The plasmids p1321.HPV-16 E6/E7, p1322.HPV-16 E6, p1324.HPV-16 E7, 46949 pCIneoEGFP, and 13680.pB-actin, were purchased from Addgene, USA. The stab culture was spread onto LB agar plates, supplemented with 100mg/mL of Ampicillin. Plates were incubated overnight at 37°C and inspected for bacterial clones. The 13680 pB-actin was purchased (Addgene, USA) to be used as the control backbone vector. Linear representations of these plasmids are shown in Figure 2.1.

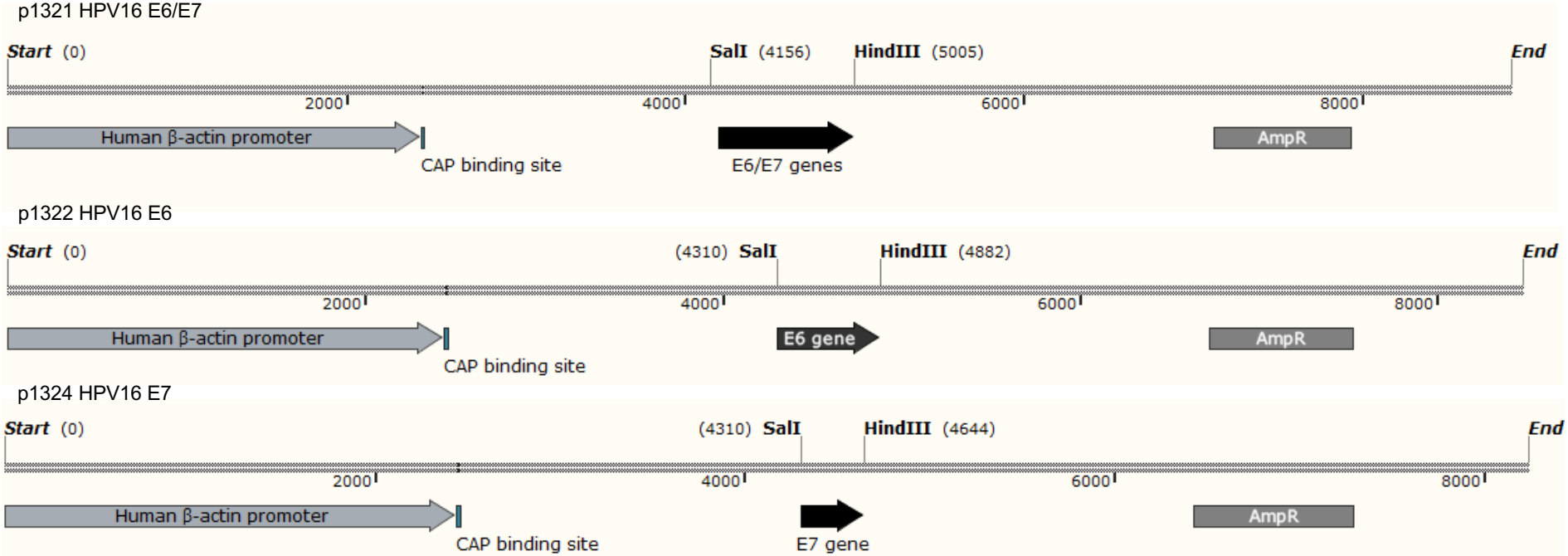


Figure 2. 1 Linear representation of all plasmid vectors: Each plasmid is Ampicillin resistant (represented by grey box, AmpR). P1321, p1322 and p1324 all contain a Human β-actin promoter. All HPV16 gene inserts are represented as the black arrow with restriction sites on either side (represented as a black line). P1321HPV16 E6/E7, p1322 HPV16 E6 and p1324 HPV16 E7 are all SalI, HindIII insert

2.2.3.2. DNA midiprep preparation

Isolated colonies for p1321.HPV-16 E6/E7, p1322.HPV-16 E6, p1324.HPV-16 E7, and 13680.pB-actin, were selected from the plates and cultured in 25mL of LB broth supplemented with 100mg/mL of Ampicillin, incubated at 37^{o}C in a InforsHT Multitron pro (Infors HT, Switzerland) 200rpm, overnight.

Plasmid DNA from overnight cultures were isolated using the Qiagen Plasmid Midi kit (QIAGEN, USA), as per suggested protocol. The DNA pellet was resuspended in 50µL of dH2O and quantitated using a Nanodrop^{TM} UV-VIS spectrophotometer (ThermoFisher Scientific, USA).

2.2.3.3. Transfection efficiency and fluorescence-activation sorting (FACS) flow cytometry

Fluorescence-activated cell sorting (FACS) flow cytometry was used to determine the most efficient concentration of plasmid DNA for transfection. For plasmid DNA optimisation, 100ng, 200ng, 500ng and 1000ng of the plasmid 46949 pCIneoEGFP and the pCIneoE backbone control were transfected into the SCC4 cell lines, in duplicate wells. Cells were harvested 24hrs and 48hrs after transfection. Cells transfected with the vector harbouring GFP were imaged on an Evos FL cell imaging system (ThermoFisher, USA) prior to harvesting.

In summary cells were harvested by adding 200µL of Trypsin to each well and incubating at 37^{o}C for 5 minutes. Trypsin was neutralised with 1mL of DMEM +10% FCS. Resuspended cells were transferred to a 2mL Eppendorf Safe-Lock tube^{TM} (Eppendorf, Germany), triplicate wells for each sample were pooled together, and spun down for 5 minutes at 4^{o}C at 1500rpm. The non-transfected samples were treated with 500µL of 70% EtOH and were left to stand for 2 minutes at room temperature. These samples act as a positive control for the viability marker TO-PRO3 Iodide (Invitrogen, USA).

The supernatant was removed, and cells are washed twice with 700µL of PBS by spinning for 5 minutes at 4^{o}C at 15000rpm. The supernatant was removed, and pellet washed twice with 700µL of FACS stain wash solution (PBS/BSA 15%/Az 0.05%/ 2%

HI-FBS) by spinning for 5 minutes at 4°C at 15000rpm. The pellet was resuspended in 500µL of FACS stain wash solution and filtered through filter gauze into falcon tubes. 0.5µL of the viability marker TO-PRO3 Iodide (Invitrogen, USA), diluted to 1:100, was added to each sample. Transfection efficiency was measured using the BD™LSRII flow cytometer (BD Biosciences, USA) and the lasers used were ALexaFluor®-488 (Invitrogen, USA) (excitation maxima 488nm, emission maxima at 519nm) and APC (excitation maxima 650nm, emission maxima 660nm). These lasers were selected based on ability to detect the fluorescence of GFP (excitation maxima 480nm, emission maxima 510nm) and TO-PRO3 Iodine (excitation maxima 640nm, emission maxima 650nm) and to reduce spectral overlap.

Each cell sample was sorted for 10,000 events. The data was recorded on the BDFACSDiva™ software and analysed using the FlowJo® (FlowJo LLC, USA) software package.

2.2.3.4. Plasmid transient reverse transfection

For HPV16 transfections, cells were seeded in a 12-well plate (Invitrogen, USA), 24hrs prior transfection at 5x104 cells/mL; for the cell line SCC4. To each well, 1mL of the desired cells at a predetermined concentration was added. 1mL of DMEM media (Invitrogen, USA) supplemented with 10%FCS (Invitrogen, USA) and GLutaMax™ (Invitrogen, USA) was also added to each well to give a final volume of 2mL.

After 24hrs, transfection was performed using the Lipofectamine® 3000 Reagent Protocol (Invitrogen, USA). Each well was transfected with 500ng of plasmid DNA, 1µL of P3000™ Reagent (Invitrogen, USA) and 1.5µL of Lipofectamine™ 3000 reagent (Invitrogen, USA). The transfected cells were then incubated at 37°C with 5% CO_2 in a humidified incubator (Binder, Germany) 48hrs after transfection the RNA was isolated from the cells.

2.2.4. Harvesting cells

Cells that were transfected were harvested 24hrs and 48hrs post transfection. The media was removed, cells were washed with 1mL of PBS, and PBS was removed. 200µL of Trypsin was added to each well and cells were dislodged using a cell scraper.

The trypsin was then neutralised with 1mL DMEM + 10%FCS. The cells were transferred to an Eppendorf tube and spun down at 1200g for 10 minutes. The cell pellet was washed with PBS and split into two 2mL Eppendorf Safe-Lock tube™ (Eppendorf, Germany). One tube was used for RNA isolation and the second for protein isolation.

2.2.5. Total RNA isolation

Total RNA isolation was performed on the SCC4, SCC090 and SCC154 cell lines that were transiently transfected. The protocol was adapted from the RNAzol RT protocol (Molecular Research Centre, USA). In summary 0.5mL (for large cell pellets) or 250uL (for 12 well plate) of RNAzol (Sigma-Aldrich, USA) was used to resuspend each cell pellet. 0.4mL of deionised water (dH_2O) was added to the solution and incubated for 5 minutes. The mixture was centrifuged at 12,000g and 4°C for 15 minutes.

Following centrifugation 80% of the total supernatant was transferred to a new Eppendorf tube. For further purification, dH_2O was added to supernatant to make a total volume of 1mL and 5µL of 4-bromoanisole (Molecular Research Centre, USA) was added. This mixture was left to stand for 5 minutes then centrifuged at 12,000g and 4°C for 15 minutes.

The RNA was precipitated by removing 80% of the supernatant (800µL), transferring this to another tube and mixing with 800µL of isopropanol and 5µL of glycogen. This mixture was stored on dry ice for 1hr. The precipitation was centrifuged at 12,000g and 4°C for 10 minutes. The supernatant was removed and discarded, and the pellet was washed twice with 500µL of 75% ethanol (Sigma-Aldrich, USA). Between each wash the mixture was centrifuge at 12,000g and 4°C for 5 minutes. The ethanol was removed and discarded, and the pellet was left to dry. The RNA pellet was then resuspended in 20µL of dH_2O and quantitated using a Nanodrop™ UV-VIS spectrophotometer (ThermoFisher Scientific, USA).

2.2.6. Analysis of gene expression using quantitative Polymerase Chain Reaction

2.2.6.1. Random primer and specific miRNA cDNA synthesis

Random primer cDNA (Table 2.5) and specific miRNA cDNA (Table 2.6) were synthesised from isolated RNA for use in TaqMan© Quantitative-Polymerase Chain reactions (qPCR). This was performed using the Applied Biosystems High-Capacity cDNA Reverse Transcription Kit for Random primers (Life Technologies, USA) and the miRNA cDNA synthesis (Life technologies, USA). The cDNA reaction was set-up as per the manufacturer's protocol. For random primer synthesis the total volume was increase to 20µL with nuclease-free water, instead of the suggested 10µL. Thermal cycling was run on a VapoProtect Mastercycler (Eppendorf, USA) and reaction cycle conditions were as per the manufacturer's protocol. Following the thermal cycling, the cDNA mixture was diluted 1 in 4 by adding nuclease-free water. Samples were stored at -20^{o}C until required.

Table 2. 5: Random Primer cDNA synthesis protocol

Reagent	Volume (1x) (µL)
10x Reverse Transcriptase Buffer, 1.0mL	2.0
25x dNTP Mix 100mM, 200µL	0.8
RNase Inhibitor 100 µL, 20Units/µL	1.0
RNA Input (500ng/ µL)	1.0
10x RT Random Primer, 1.0mL	2.0
MultiScribe™ Reverse Transcriptase 100µL, 50units/µL	1.0
Nuclease-free water	12.2
Total	**20.0µL Final Volume**

Table 2. 6: Specific microRNA primer cDNA Synthesis protocol

Reagent	Volume (1x) (µL)
10x Reverse Transcriptase Buffer, 1.0mL	1.5
25x dNTP Mix 100mM, 200µL	0.15
RNase Inhibitor 100 µL, 20Units/µL	0.2
RNA Input (500ng/µL)	1.0
5 x miRNA RT Primer	6µL in total
MultiScribe™ Reverse Transcriptase 100µL, 50units/µL	1.0
Nuclease-free water	5.15
Total	**15.0µL Final Volume**

2.2.6.2. TaqMan qPCR approach and design

The TaqMan© qPCR assay was used for the analysis of the miRNAs miR-496, miR-33a and RNU6B and the genes SREBF2, E2F2, HPV16 E6, HPV16 E7, GAPDH, ACTB, B2M. Reaction set up was as per TaqMan® qPCR protocol, modifications are displayed in Table 2.7, this was used for all gene expression analysis and miRNA analysis. The smaller volume of the qPCR reaction increases the sensitivity for detecting the specific miRNAs and genes as described in Khoury and Tran [207]. All reactions were performed in triplicate using the StepOnePlus™ or the QuantStudio 6K Flex System (Life Technologies, USA). The collected data was analysed using the ΔCT or ΔΔCT method [208].

Table 2. 7: TaqMan© quantitative Polymerase Chain reaction protocol

Reagent	Manufacturers Protocol (1x) (µL)	Modified protocol (1x) (µL
TaqMan® Assay Gene Probe (20x)	1.0	0.5
TaqMan® Master Mix (2x)	10.0	2.5
Diluted cDNA	9.0	1.0
RNAse-free Deionized water	0.0	1.0
Total	**20.0**	**5.0**

2.2.6.3. SYBR green approach and design

The Fast SYBR green™ qPCR assay was used for the analysis of the lncRNA TRINGS and reference gene GAPDH. Reaction set up (Table 2.8) was as per Fast SYBR™ (Applied Biosystems, USA) qPCR protocol. All reactions were performed in triplicate using the StepOnePlus™ (Life Technologies, USA). The collected data was analysed using the ΔCT or ΔΔCT method [208].

Table 2. 8: Fast SYBR™ Green quantitative Polymerase Chain reaction protocol

Reagent	Modified protocol (1x) (µL)
Fast SYBR™ Green master mix (2x)	10.0
Forward Primer (10µM)	1.0
Reverse Primer (10µM)	1.0
RNAse-free Deionized water	4.0
cDNA template	4.0
Total	**20.0**

2.2.7. Protein isolation and western blot analysis

Transfected cells were harvested and lysed with CellLytic M cell lysis reagent (Sigma-Aldrich) supplemented with 0.1% serine protease inhibitor (Sigma-Aldrich).
The protein concentration was determined using Qubit™ Protein assay kit on the Qubit 2.0 fluorometer (Life Technologies) as per manufactures instructions.
20µg of protein was loaded onto a precast Bolt™ (4 to 12%) Bis-tris Plus mini gel (Invitrogen) in Bolt™ (MES)-SDS running buffer (50mM MES, 50mM Tris Base, 0.1% SDS, 1mM EDTA, pH 7.3) as per manufactures protocol. After electrophoresis, the proteins were transferred onto a polyvinylidene difluoride (PVDF) membrane (Pierce) using Bolt™ Bis-Tris Transfer buffer (25mM Bicine, 25mM Bis-Tris, 1mM EDTA, pH 7.2). The membranes were incubated with SREBF2 (1:200) and HPV16 E6 (1:500). followed by conjugation with a secondary antibody containing horseradish peroxidase (HRP). The membranes were also incubated with ACTB (1:2000) was used as a control protein, pre labelled with HRP. Chemiluminescence with the ECL Western Blotting Substrate (Promega) was used to visualise the specific protein bands. Antibodies were purchased from Abcam, SREBF2 (ab30682), HPV16 E6 (ab70) and Actin (Ab20272). Chemiluminescence with the ECL Western Blotting Substrate (Promega) was used to visualise the specific protein bands.

2.2.8. The Cancer Genome Atlas

We downloaded RNA-seq and miRNA-seq BAM files from 48 Head and Neck cancer patient samples (HPV16+ n=33 , HPV-n= 15), specifically Oropharyngeal cancers (tonsil, base of tongue, pharynx and oropharynx), with permission from The Cancer Genome Atlas (TCGA) from the GDC Data portal (https://portal.gdc.cancer.gov/). HPV positivity was determined using HPV by ISH clinical data and the RNA sequencing results from TCGA [209].

Chapter 3: Developing a virus-microRNA interactome using Cytoscape

This chapter has been published in the journal “MethodsX”. Mason, D., Hill, M., Marques, T.M., Carvalho, M.G. and Tran, N., 2020. Developing a virus-microRNA interactome using cytoscape. *MethodsX*, *7*, p.100700.

3.1. Copyright information

This chapter has been published as a **Methodology article** in ***MethodsX***:

MethodsX 2020, 7:100700 https://doi.org/10.1016/j.mex.2019.10.011

The text presented here is the accepted version of the manuscript. The numbering of sections, style of referencing, the numbering of tables and figures are altered to align with the formatting of the thesis.

3.2. Authors contribution

Paper: Developing a virus-microRNA interactome using Cytoscape *MethodsX*
Authors: Dayna Mason*, Meredith Hill*, Tania Monteiro Marques*, Margarida Gama Carvalho, Nham Tran

*These authors contributed equally to the work

Authors Contributions:
These authors contributed equally to the work; Dayna Mason (graduate student) is an equal first author of this Methodology paper. Her contribution to this manuscript is demonstrated by the following roles and tasks:

Under the guidance of and collaboration with Tania Monteiro Marques, Dayna assisted in developing the methodology to produce an HPV16-non-coding RNA interactome using Cytoscape. She broke down the steps for the bioinformatics pipeline to produce the interactome and documented each process contributing to Figure 1 (Graphical abstract), Figure 2, Figure 3, Figure 4, Figure 5, and Figure 6. In addition, she was actively involved in the writing of the initial draft and the iterative editing and review process.

Nham Tran, Tania Monteiro Marques and Margarida Gama Carvalho provided guidance on the bioinformatics pipelines to create Figure 5 as well as critical analysis of the pipeline, writing and editing of the manuscript. Meredith Hill made contributions to the writing of the initial draft and editing of the manuscript.

Signatures:

Dayna Mason (graduate student) Production Note: Signature removed prior to publication.

Meredith Hill (Co-investigator) Production Note: Signature removed prior to publication.

Tania Monteiro Marques (Co-investigator) Production Note: Signature removed prior to publication.

Margarida Gama Carvalho (Co-investigator) Production Note: Signature removed prior to publication.

Nham Tran (DM PhD Principal Supervisor) Production Note: Signature removed prior to publication.

3.3. Abstract

It is currently difficult to determine the effect of oncogenic viruses on the global function and regulation of pathways within mammalian cells. A thorough understanding of the molecular pathways and individual genes altered by oncogenic viruses is needed for the identification of targets that can be utilised for early diagnosis, prevention, and treatment methods. We detail a logical step-by-step guide to uncover viral-protein-miRNA interactions using publicly available datasets and the network building program, Cytoscape. This method may be applied to identify specific pathways that are altered in viral infection and contribute to the oncogenic transformation of cells. To demonstrate this, we constructed a gene regulatory interactome encompassing Human Papillomavirus Type 16 (HPV16) and its control of specific miRNAs. This approach can be broadly applied to understand and map the regulatory functions of other oncogenic viruses and determine their role in altering the cellular environment in cancer.

Availability and Implementation Cytoscape [210, 211] is freely available at: https://cytoscape.org/

3.4. Graphical Abstract

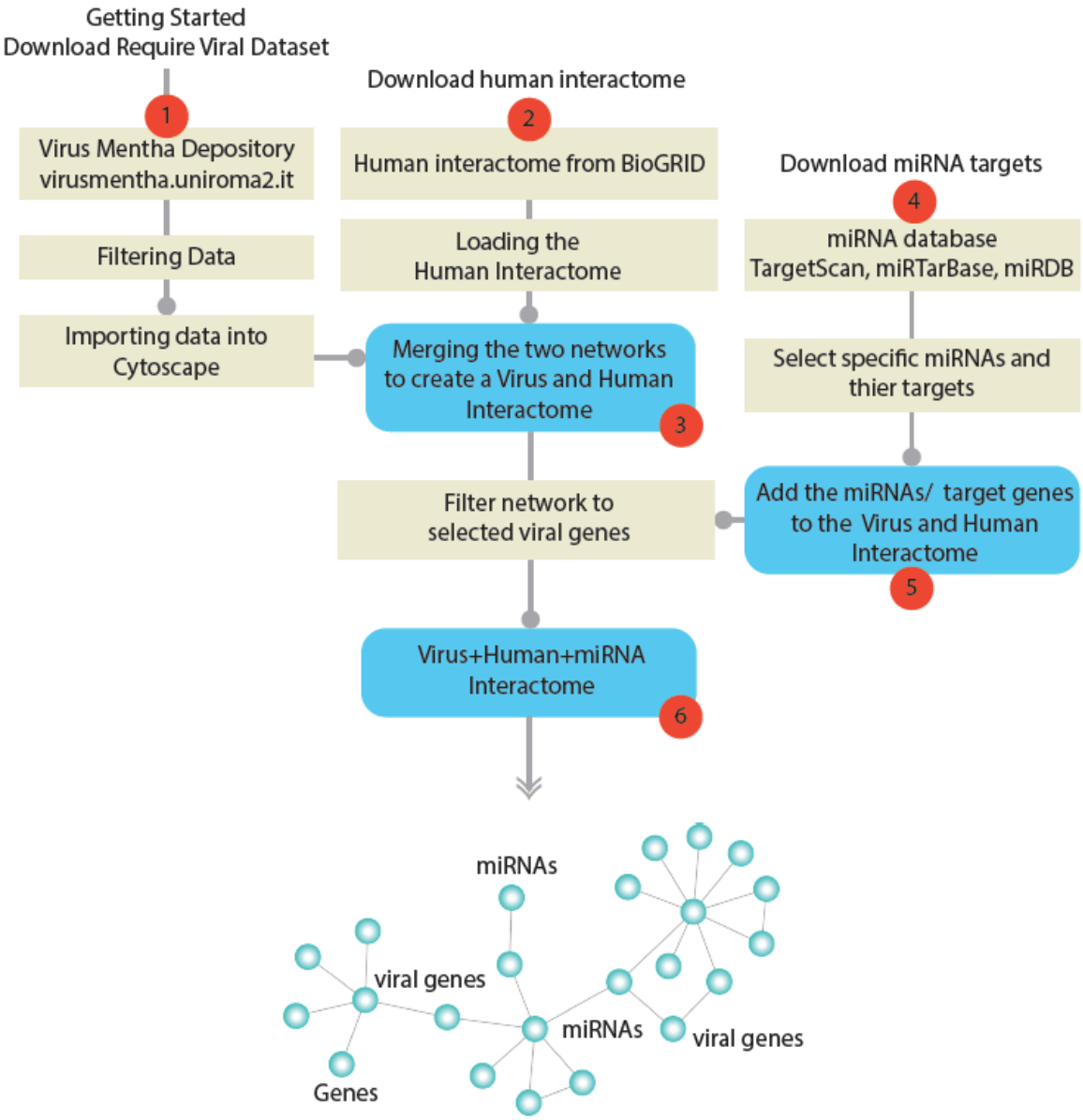


Figure 3. 1: Graphical abstract depicting the methodology for the development of a viral-miRNA interactome

3.5. Building a viral miRNA human interactome

An example of a viral-human miRNA interactome was created using the guidelines described in this paper. The virus of choice was HPV16, with a particular emphasis on the effect of the viral oncoproteins E6 and E7 on miRNA and gene regulation. For this interactome, we specifically focused upon hsa-miR-33a, hsa-miR-496, and the transcription factor SREBF2. This method can be adapted to a specific virus and miRNA(s) of interest.

3.6. Downloading required viral dataset

To build this network, several datasets were downloaded from multiple publicly available websites. The VirusMentha depository website (virusmentha.uniroma2.it) contains data listing the interactions between a virus for a specific set of viruses with human proteins [212, 213]. The HPV16 viral protein interactions were downloaded from the VirusMentha website and filtered in Excel by taxon 333760 (HPV16) and 9606 (H. sapiens) (Figure 3.2A).

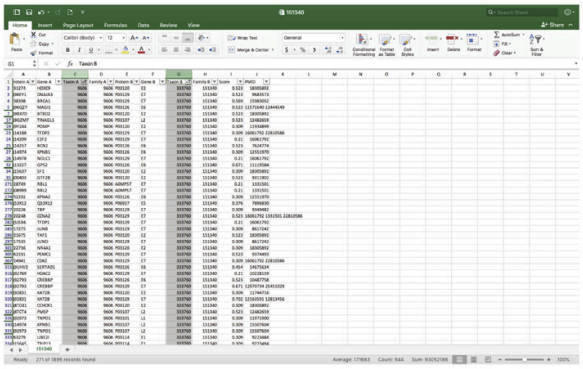
3.7. Importing data into Cytoscape

Files can be directly uploaded into Cytoscape using the import network tool (**File > Import> Network> Choose File**). When importing our data, the gene names of humans were selected as the source nodes and the HPV16 oncogene names as the target node. As there is no information about the directionality of the interactions, this choice is arbitrary, and the source and target nodes can be attributed to either human-virus or virus-human. The created network only contains the direct interactions between the viral and host proteins (Figure 3.2B). Next, the viral interactome is merged with the human interactome to create a more complete network.

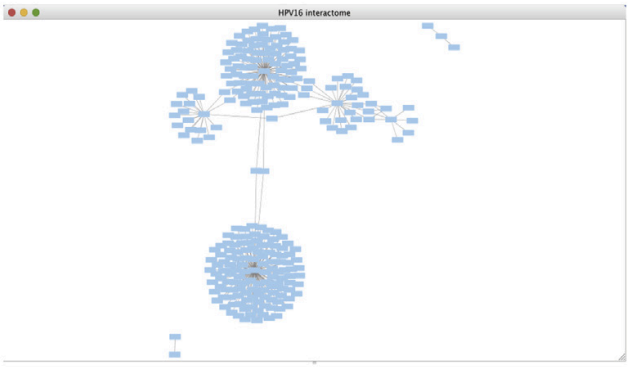
3.8. Downloading the human interactome

Cytoscape version 3.3.0 already contains the Human interactome from BioGRID. If using a newer version, the human interactome will need to be downloaded from BioGRID [214] (https://downloads.thebiogrid.org/BioGRID/Release-Archive/BIOGRID-3.4.163/) and imported into Cytoscape. To load the Human interactome in Cytoscape v3.3.0, go to the toolbar and select **Help > Show Welcome Screen > Select H. sapiens**. Import the Human interactome as a new collection network. This will load the entire human interactome (Figure 3.2C). This was utilised to determine the direct and indirect effects of the HPV16 viral oncoproteins on gene expression.

A



B



C

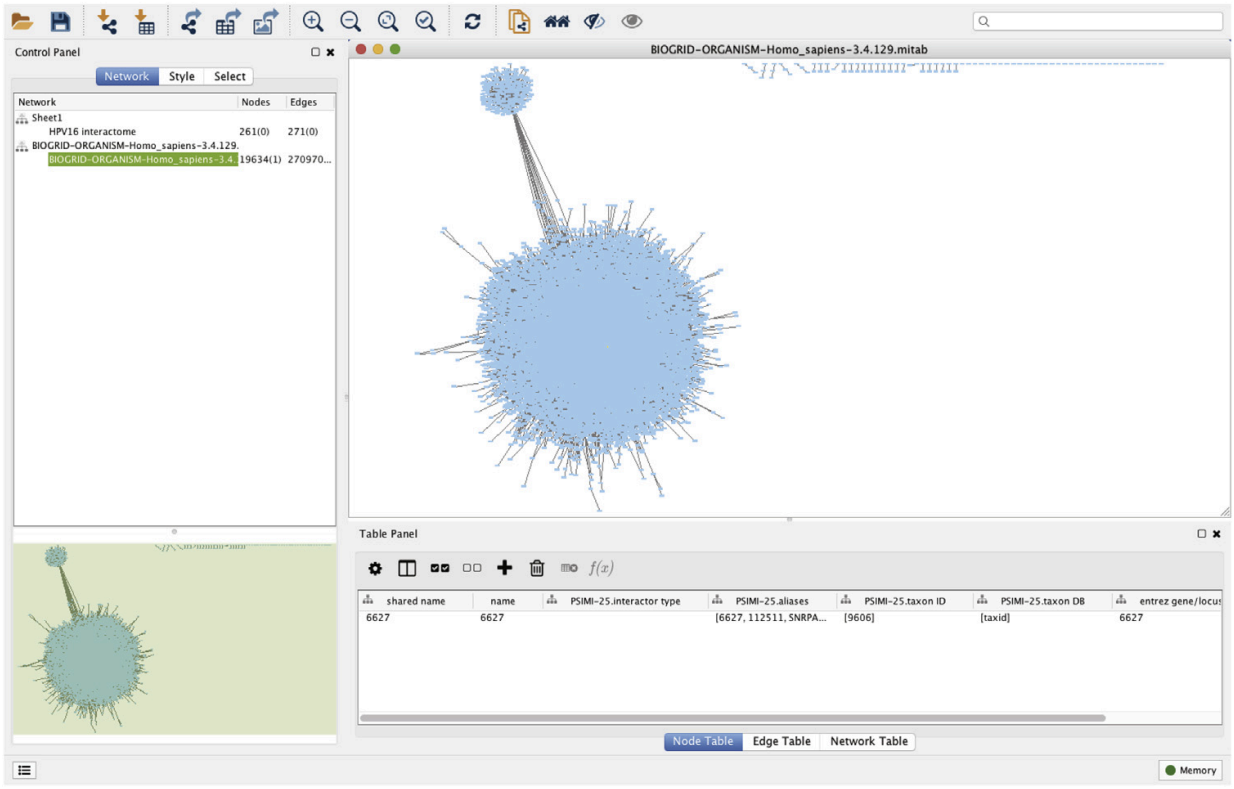


Figure 3. 2 Importing Raw Human and HPV Data. A) Screen capture of Excel workbook highlighting the filtered columns for Taxon A and Taxon B to select for HPV16 protein interactions specific for H. sapiens. B) Raw HPV16 interactome as visualised in Cytoscape. C) BIOGRID complete Human interactome as viewed in Cytoscape.

3.9. Merging the two networks to create a virus and human interactome

The human-viral and H. sapiens interactomes need to be merged to visualise their connections. To merge the two networks, go to the option **Tools> Merge > Networks**. The HPV16 interactome and BIOGRID Human genome interactome were merged by 'shared name' and 'PSMI-25. alias' (Figure 3.3A). The merge type was assigned 'Union', as the aim of this project was to integrate the HPV16 genes with those of the human genome.

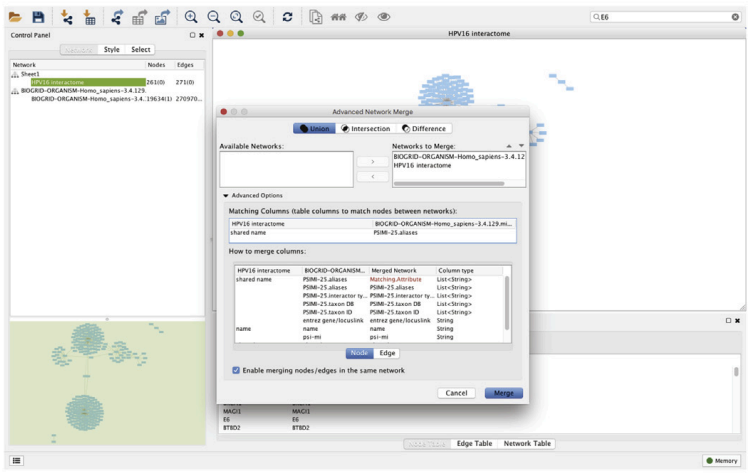
3.10. Filtering the viral/human interactome

The HPV16 viral oncoproteins, E6 and E7, were searched for using the search network tool, and their respective nodes were moved away from the main interactome. This ensures easier identification of the two genes for the subsequent creation of a smaller network.

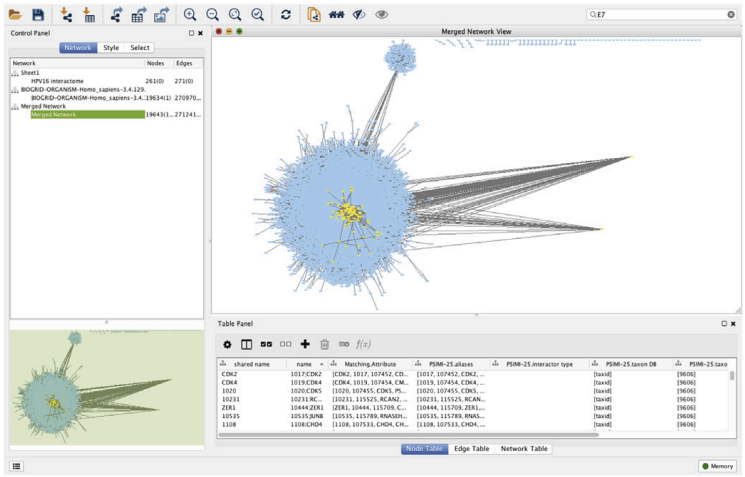
For our example on HPV16, the two major oncoproteins, E6 and E7, were selected, along with their direct interactors (first neighbours). This was done by searching for these two HPV16 oncogenes within the network and selecting their first neighbours. Once all nodes of interest are selected, a new network is created by selecting **File > New > Network > from Selected nodes, Selected edges**. To tidy up this new network, select **Edit>Remove Duplicated edges**, and **Edit>Remove Self-Loops.**

This produces a smaller secondary network, as shown in Figures 3.3B and C. This interactome contains the genes that directly interact with the viral oncoproteins and their relationship with other human genes. It was this smaller interactome that was used for further network filtering and annotation after the removal of duplicated edges and self-loops.

A



B



C

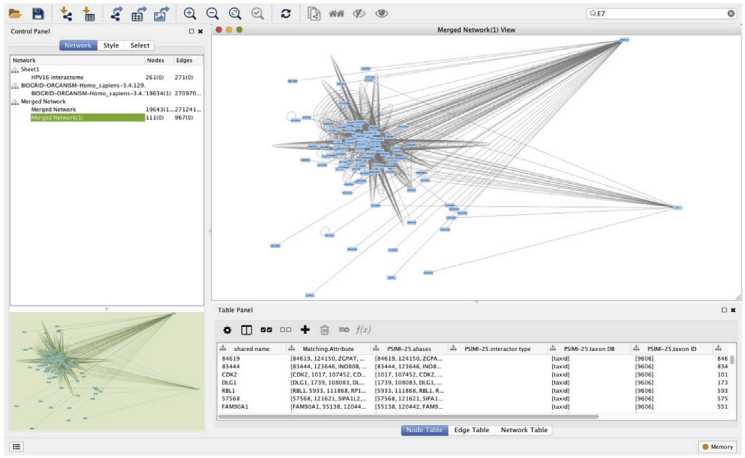


Figure 3. 3: Merged and Filtered Network. A) The pop-up screen containing the settings for the merging of the HPV16 interactome and the BIOGRID human interactome. B) The merged HPV16 and BIOGRID human interactome network. To the right of the network, the visualisation panel are the two oncogenes, E6 and E7. The yellow-coloured nodes represent those genes that are direct interactors of E6 and E7. C) The smaller network containing E6 and E7 and their association with human genes.

3.11. Assigning gene names and transcription factors

The gene targets of our miRNAs of interest, hsa-miR-33a and hsa-miR-496, were extracted from miRanda and saved as a .txt file (Figure 3.4A). The transcription factors for hsa-miR-33a, hsa-miR-496 and SREBF2 were inferred from the UCSC Genome Browser and saved as a .txt file (Figure 3.4B).

To select nodes from the miRNA target lists, which are downloaded as gene names, all nodes need to be changed to their respective gene name. This was done using the online tool DAVID ID (https://david.ncifcrf.gov/conversion.jsp) [215, 216]. Export the merged interactome node list from Cytoscape and remove all columns except the Entrez ID. Import this Entrez ID list into DAVID to identify the gene names and abbreviations for each node. Download the gene names from DAVID and import them into Cytoscape using file > import. Choose “name” as the key identifier. Once completed, all nodes in the interactome should now show their gene name (Figure 3.4C). Alternatively, if using the human interactome from the BIOGRID website, the gene names are used automatically, rather than Entrez ID.

3.12. Downloading the miRNAs and their targets

MiRNAs are important gene regulators therefore it is essential include them in the interactome. The targets for the miRNAs of interest were retrieved from www.microRNA.org using the miRanda algorithm. Other data repositories can be used to obtain miRNA target information, such as TargetScan [217], miRTarBase [218] and miRDB [219]. Download the most recent file for human miRNA targets that are considered “highly conserved and good”. Convert the miRNA-target files to a .txt file. For our example, we investigated the miRNAs miR-496 and miR-33a.

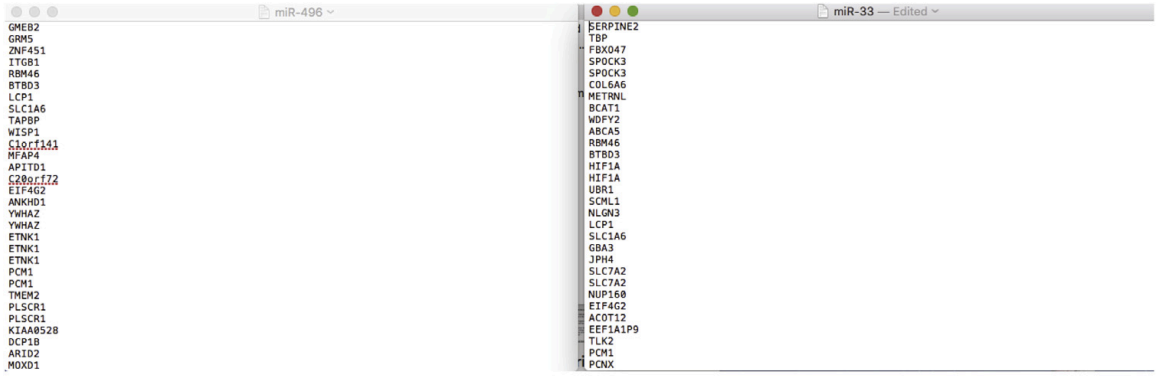
The transcription factors for each miRNA were inferred using the UCSC Genome Browser (https://genome.ucsc.edu) [220]. The UCSC predicted transcription factors are kept in separate lists to the miRNA targets. Each miRNA is searched individually using the FEB2009(GRCH37/hg19) dataset. The track containing the transcription factors is titled “transcription factor chip-seq (162 factors) from ENCODE with Factor book motifs”. To view the full track right click and select “full”. The potential transcription factors were extracted 6kb upstream of the coding regions of the miRNAs

of interest (Figure 3.4B). The produced list of transcription factors for each miRNA were saved as a .txt. In our example, we noted that miR-33a was encoded within the gene SREBF2, therefore, transcription factors regulating SREBF2 were also identified using the UCSC genome browser.

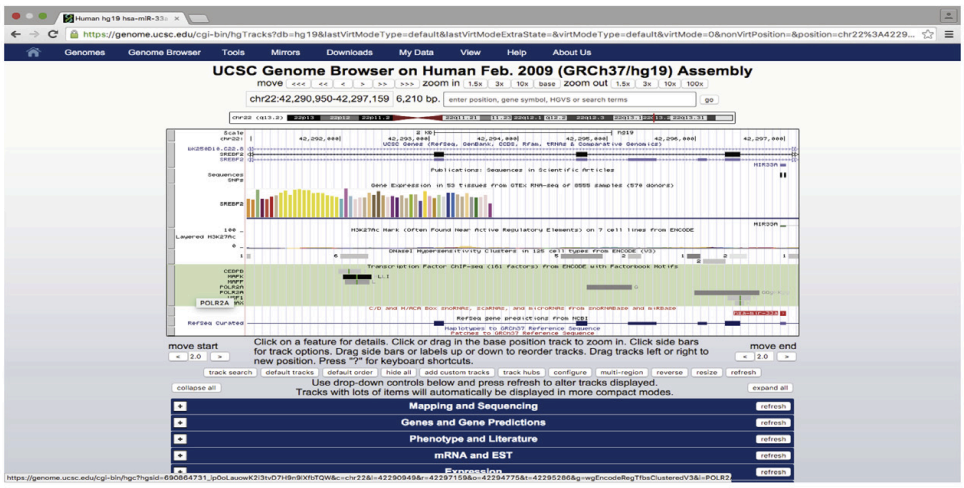
3.13. Addition of miRNAs and target genes to the virus and human interactome

The viral-human interactome needs to be filtered further to contain only targets and transcription factors of our miRNAs of interest. For our HPV16 (E6/E7) example, the gene lists containing the targets and transcription factors of miR-496, miR-33a and SREBF2 were used to select nodes. This was performed by going to the Cytoscape menu **Select > Nodes > From ID lists file** and choosing the gene list of interest. After this step, the adjacent edges of these nodes are selected by clicking **Select > Edges > Adjacent edges**. The nodes connected to those edges are then selected by going to **Select > Nodes > Nodes connected by selected edges** (Figure 3.5A). After this step, a new network containing only the nodes and edges of interest can be created by selecting **File > New > Network > from selected nodes, selected edges**.
Secondary neighbours (or more) of these genes can also be included, if desired, by repeating the previous steps. Once the network is filtered to contain the relevant genes, the nodes for the miRNAs of interest need to be added manually by **right clicking > add > node**. The edges connecting these miRNAs to their targets and transcription factors also need to be added manually by **right clicking > add > edge connecting nodes**. This process needs to be repeated for each miRNA target (Figure 3.5B). The target and transcription factor ID lists can be used to identify the nodes that require the addition of edges connecting to the miRNA of interest.

A



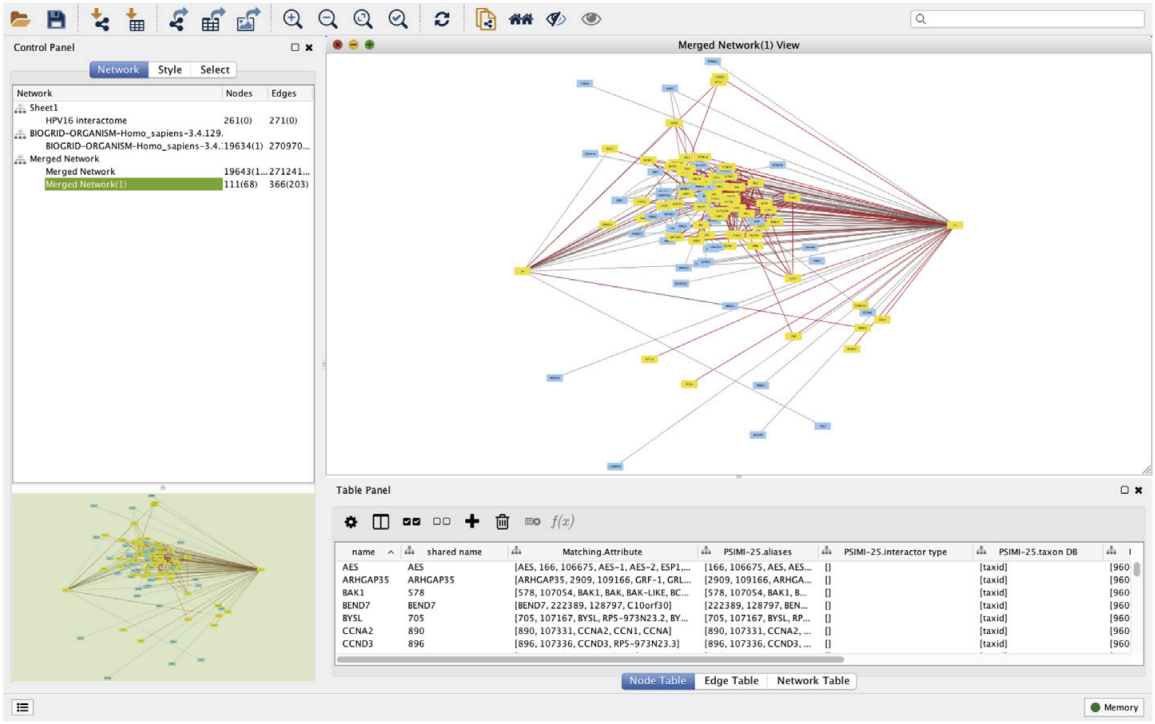
B



C

Figure 3. 4 Screenshots of the. txt files created for the miRNAs of interest. A) Gene target list for hsa-miR-33a (Left Panel). Transcription factor list of hsa-miR-496 as extracted from the UCSC Genome Browser (Right Panel). B) Screenshot of the UCSC Genome Browser highlighting the location of the transcription factors of the selected gene or miRNA. C) Screenshot of the DAVID output of the conversion from Entrez ID to Gene Name.

A



B

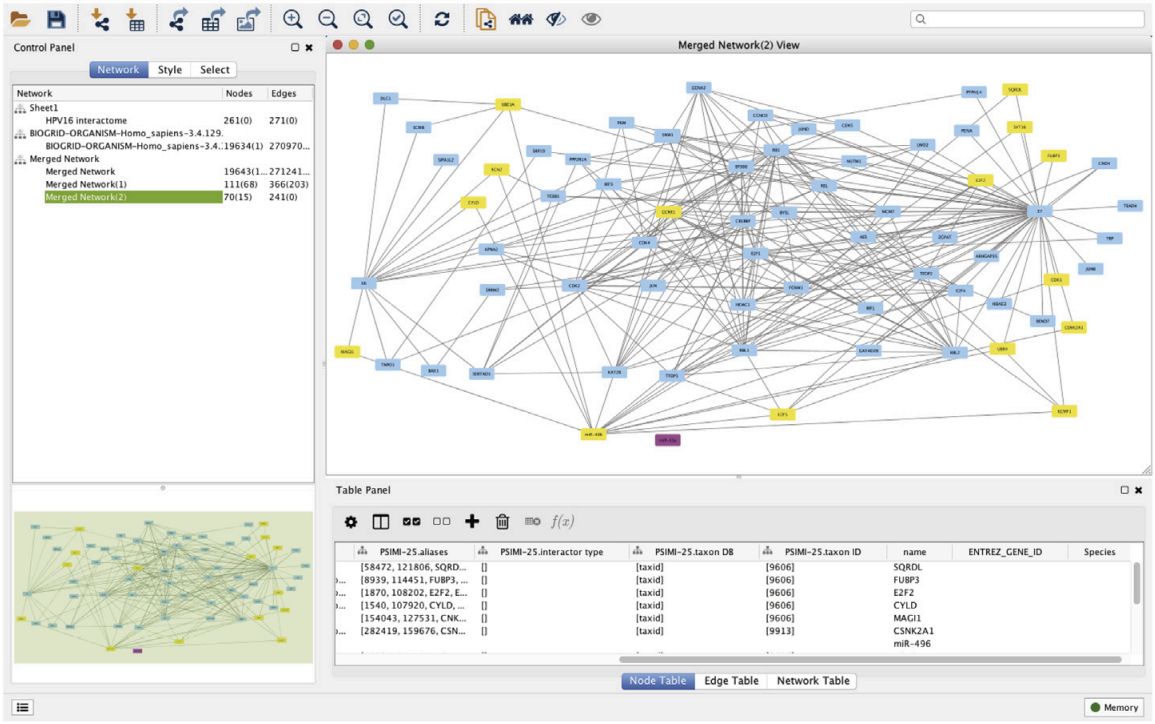


Figure 3. 5: Selecting nodes A) Screenshot of the created viral interactome highlighting the nodes selected from the gene target and transcription factor lists and their direct interactors. B) Screenshot of the network, highlighting miR-496 and its connections to the highlighted nodes (yellow). The other miRNA, miR-33a, is shown in purple.

3.14. Annotation of genes and final virus/human/miRNA interactome

Once the miRNAs were added and connected to their respective targets, node annotation was performed to visually delineate between transcription factors, gene targets, and miRNAs. The ‘gene name’ column from the exported node table was used in a separate document to classify the nodes according to their connection to the HPV16 oncoproteins, and their characteristics (Figure 3.6A).

To annotate the network, the created excel table was imported into Cytoscape, and the samples were matched according to ‘gene name’ (Figure 3.6B). The inclusion of these annotations allows for the alteration of the visual properties of the network, such as varying the colour of nodes according to their targets (Figure 3.6C), the shape of the nodes to indicate their biological role, and the colour of edges according to their regulatory interaction. The final interactome, which includes these features, in addition to edge annotations indicating the direction of the transcription factor interactions, is shown in Figure 3.7.

A

B

C

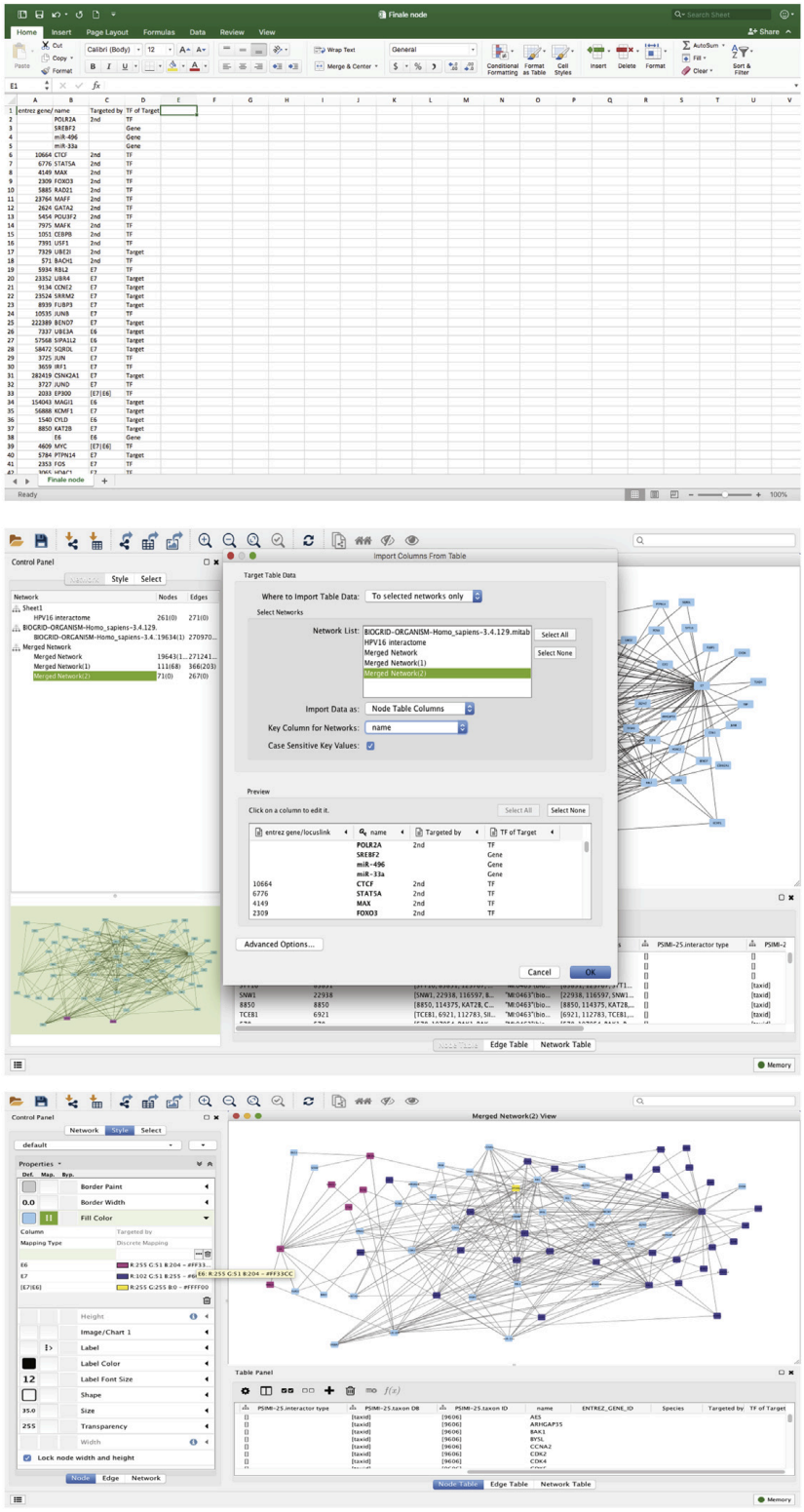


Figure 3. 6: Annotating the network A) Screenshot of the Microsoft Excel file used to annotate the network. B) Screenshot of the import table screen in Cytoscape, indicating the way in which the annotations will be correlated with the data already saved within the program. C) The colour annotation of the nodes within the created HPV-miRNA interactome. Pink indicates targets of E6, blue indicates targets of E7, and turquoise represents those that are targets of both E6 and E7.

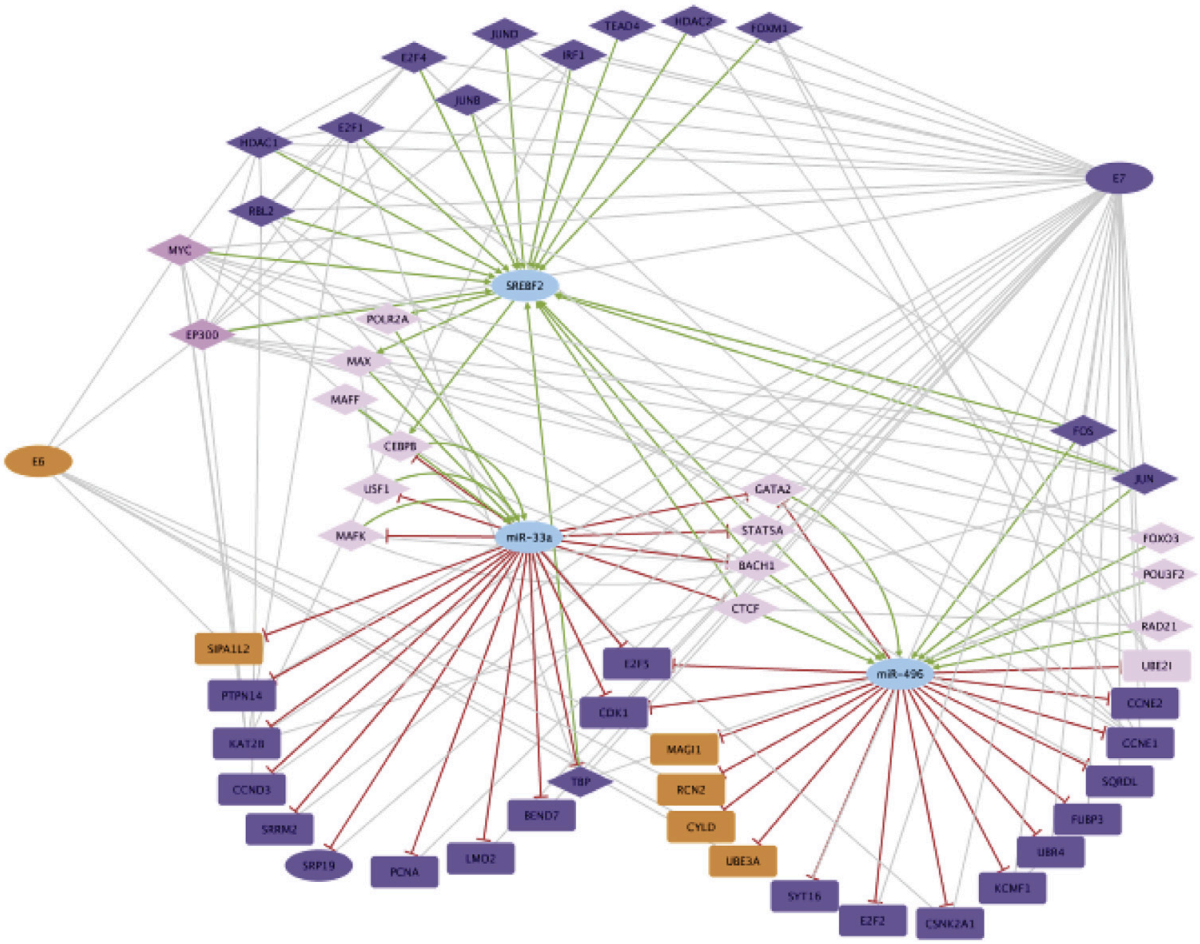


Figure 3. 7:The final HPV16-miRNA interactome created using the described Cytoscape methodology, indicating both the impact of transcription factors and genes on mRNA and miRNA expression. Orange represents targets of E6, blue indicates those targeted by E7, and the pink nodes are targeted by both E6 and E7. The lighter pink nodes represent secondary interactions to both E6 and E7, those proteins that are downstream targets of proteins directly targeted by these oncogenes. Additionally, the diamonds, rectangles and ovals represent the transcription factors, genes, and targets of interest respectively. Edge annotations were also added to indicate the direction of transcription factor control as either activating (green) or inhibitory (red).

3.15. Additional information

Changes to the human genome and the expression of its regulators, such as miRNAs, in response to viruses is highly complex. This is of particular importance in the case of virally driven oncogenesis, where further modifications to the regulatory network may compound pre-existing tumorigenic characteristics. Using the mapping software, Cytoscape, we developed a method to integrate the viral and human genome, along with miRNA regulators, which can be used to identify novel pathways and interactions. Our described method will enable researchers to readily identify targets and pathways of interest in the context of human-viral infection and the development of disease.

Chapter 4: Human papillomavirus 16 E6 modulates the expression of miR-496 in oropharyngeal cancer

This chapter has been published in the journal “Virology”. Mason, D., Zhang, X., Marques, T.M., Rose, B., Khoury, S., Hill, M., Deutsch, F., Lyons, J.G., Gama-Carvalho, M. and Tran, N., 2018. Human papillomavirus 16 E6 modulates the expression of miR-496 in oropharyngeal cancer. *Virology*, *521*, pp.149-157.

4.1. Copyright information

This chapter has been published as a **Research Article** in ***Virology***:

Virology 2018, 521:149–157 https://doi.org/10.1016/j.virol.2018.05.022

The text presented here is the accepted version of the manuscript. Numbering of sections, style of referencing, numbering of tables and figures are altered to align with the formatting of the thesis.

4.2. Author contribution

Paper: Human papilloma virus 16 E6 modulates the expression of miR-496 in oropharyngeal cancer *Virology*

Authors: Dayna Mason*, Xiaoying Zhang*, Tania Monteiro Marquess, Barbara Rose, Samantha Khoury, Meredith Hill, Fiona Deutsch, J. Guy Lyons, Margarida Gama-Carvalho, Nham Tran

* These authors contributed equally to the work

Authors Contributions:
These authors contributed equally to the work. Dayna Mason (graduate student) is an equal first author of this research article. Her contribution to this manuscript is demonstrated by the following roles/tasks:

Specifically, she collated the data describing patient clinical information in Table 1. She designed, performed, and analysed experiments shown in Figure 3, and Supplementary Figure 3. She also contributed to the bioinformatic analysis to produce the HPV-miRNA interactome (Figure 4, Table 3, and Supplementary table 1) collaboratively with Tania Monteiro Marquess. She confirmed the results seen in Figure 1 and Table 2 to determine if the miRNA expression pattern is robust in The Cancer Genome Atlas as well as repeated in the patient cohort. This generated supporting data as well as investigation into other miRNAs not featured in the manuscript. In addition, she was actively involved in the writing of the initial draft and the iterative editing and review process.

Xiaoying conducted the LNA array on patient samples and performed the analysis to produce, Figure 1 and Table 2. Xiaoying also performed the E6 overexpression analysis for Figure 2. Nham Tran, Tania M. Marques, Margarida Gama-Carvalho contributed to the bioinformatic pipeline for the HPV-interactome analysis in Figure 4, Table 3, and Supplementary table 1. J Guy, Lyons conducted the GFP expression plasmid transfection in cells to form Supplementary Figure 1. Nham Tran conducted

the bioinformatics analysis and 3D terrain mapping described in Supplementary Figure 2.

Tania Monteiro Marques, Samantha Khoury, Meredith Hill, Fiona Deutsch, Margarida Gama-Carvalho and Nham Tran made contributions to the conception of the work, critical analysis of data, writing and editing of the manuscript.

Signatures:

Dayna Mason (graduate student) Production Note: Signature removed prior to publication.

Xiaoying Zhang (Co-investigator)
Ivy Zhang

Tania Monteiro Marquess (Co-investigator) Production Note: Signature removed prior to publication.

Barbara Rose (Co-investigator)
B.Rose

Samantha Khoury (Co-investigator) Production Note: Signature removed prior to publication.

Meredith Hill (Co-investigator) Production Note: Signature removed prior to publication.

Fiona Deutsch (Co-investigator) Production Note: Signature removed prior to publication.

J. Guy Lyons (Co-investigator) Production Note: Signature removed prior to publication.

Margarida Gama-Carvalho (Co-investigator) Production Note: Signature removed prior to publication.

Nham Tran (DM PhD Principal Supervisor) Production Note: Signature removed prior to publication.

4.3. Abstract

Human papillomavirus (HPV), notably type 16, is a risk factor for up to 75% of oropharyngeal squamous cell carcinomas (SCC). It has been demonstrated that small non-coding RNAs known as microRNAs play a vital role in the cellular transformation process. In this study, we used an LNA array to further investigate the impact of HPV16 on the expression of microRNAs in oropharyngeal (tonsillar) cancer. Several miRNAs were found to be deregulated, with miR-496 showing a four-fold decrease. Over-expression of the high risk E6 oncoprotein down-regulated miR-496, impacting upon the post-transcriptional control of the transcription factor E2F2. These HPV specific miRNAs were integrated with the HPV16 interactome to identify possible mechanistic pathways. These analyses may provide new insights into the novel molecular interactions between HPV16 and miRNAs in oropharyngeal cancers.

4.4. Introduction

Head and neck cancer is one of the top six malignancies affecting men worldwide [221]. The majority of these cancers arise from squamous cells that line the mucosal surfaces of the oral cavity, larynx, and oropharynx. Despite improvements in clinical care, survival rates of approximately 50% have remained unchanged for the past several decades [221]. Traditionally, the main risk factors for head and neck cancer globally have been tobacco and alcohol [222]. This view has now changed with epidemiological data showing that human papillomavirus (HPV) has a causal association with a subset of head and neck cancers [66, 223]. The main anatomical site for HPV associated cancers is the oropharynx, most notably the tonsils and base of tongue, with HPV positivity rates of up to 75% reported [224, 225]

HPV16 has been identified overwhelmingly as the major contributing type, with positivity rates dependent in part on cultural background and ethnicity [226]. Over the last two decades there has been an increase in the prevalence for HPV16 in oropharyngeal cancers across both Europe and North America [227, 228]. Of notable concern is that HPV16 positivity increases the odds ratio by more than 200 times. In comparison, smoking alone results in a 6.82 times increased risk of disease [229]. Despite these increases in prevalence and risk, it is well known that the median survival for HPV-positive patients is significantly longer than for HPV-negative patients.

It is known that HPV16 induces cellular transformation by the production of its two major oncoproteins, E7 and E6 [230]. Both early genes can alter the expression of the tumour suppressors pRb and p53 [231-233]. Although this is the classical modification induced by HPV, we are only beginning to unravel the interactions between the virus and other regulatory elements.

The non-coding milieu, which includes small RNAs, has been shown to be an important modulator of disease progression [234-237]. These small non-coding RNAs, often known as microRNAs or miRNAs, are regulatory small RNAs (~ 22 nts) which are highly conserved and have been predicted to regulate 40% to 60% of all human genes [238, 239]. To date over 1800 miRNAs have been identified within the human genome [240] and control gene expression by binding to the target 3´UTR and eliciting post-transcriptional gene silencing [241]. This modality of gene silencing is finely tuned

and underpins the biological networks which drive vital processes such as development, growth, differentiation, and cell death. Disruption of this miRNA regulatory network, via the accumulation of genomic instability, can lead to human diseases (As reviewed in Khoury and Tran [242]).

Our previous studies have shown that the dysregulation of specific miRNAs may play a vital role in the development of squamous cell carcinoma (SCC) of the head and neck [82, 243, 244]. Wald *et al* published one of the first studies describing the effect of E6 overexpression and its impact on miRNA levels in human foreskin keratinocytes [245]. It is likely that HPV16 alters specific miRNAs in oropharyngeal cancers, and this affects their regulatory targets. This regulatory cascade is often not well characterised in oropharyngeal cancers and given that a single miRNA can potentially regulate hundreds of targets, the disruption of these regulatory processes by HPV may have profound affects.

To further our understanding of the regulatory impact of HPV16, we performed a genome wide miRNA screen in HPV16+ oropharyngeal cancers. Using a stringent threshold, thirteen miRNAs were identified to be deregulated in HPV+ cancers. We showed that the overexpression of E6 modulated the expression of miR-33 and miR-496. Moreover, the transcription factor E2F2, which is vital for HPV16 replication, is directly influenced by miR-496 levels. Network analysis using the HPV interactome revealed a possible mechanism for the regulation of these miRNAs by HPV16.

4.5. Methods

4.5.1. Study population

Investigations were carried out on 30 patients treated for SCC of the oropharynx at Royal Prince Alfred Hospital, Sydney between 2002 and 2006 (Table 4.1). The study was approved by the Research Ethics Committee at Royal Prince Alfred Hospital, Sydney, Australia (Approval X05-0270). Informed consent was obtained for the collection of fresh tissues. Immediately after surgical resection, tissues were snap frozen on dry ice and stored at -70^{o}C until use. The histology of tissues was assessed by the hospital pathologists.

4.5.2. HPV status

The presence of HPV16 was determined through the detection of E6 using TaqMan assay (Applied BioSystems), with 16S gene as an internal sample control and SiHa total RNA containing HPV E6 mRNA as an E6 positive control. Approximately 90% of HPV+ oropharyngeal cancers from our centre are caused by type 16 [246]. However, to exclude the possibility that a small proportion of HPV+ cancers were induced by HPV types other than type 16, all samples were tested for p16 by semiquantitative immunohistochemistry as previously described [247].

Table 4. 1: Cohort for the study included 30 tumour samples. Of these 14 samples were arrayed in duplicate and 25 samples were used for qPCR validation. All samples were pathologically examined, checked for HPV status by PCR and all HPV positive (HPV+) samples were immunohistochemically stained for p16 expression. Average age was 59 with a range of 38–90 years. Abbreviations: MD, WD, PD and UND: moderately, well, poorly differentiated and un-differentiated. SCC: squamous cell carcinoma. BOT: base of tongue, R: right and L: left, HPV+: HPV positive, HPV-: HPV negative.

	Sample	Gender	Age	Location	Pathology	HPV byE6	P16 staining	Used for qPCR	Used for array
1	HN02-04T	M	90	Tonsil-R	SCC, WD	HPV-	P16-	Yes	
2	HN03-05T	M	80	Tonsil-L	SCC, WD	HPV-	P16-	Yes	Yes
3	HN05-04T	M	70	BOT	SCC, WD	HPV-	P16-	Yes	
4	HN06-03T	M	80	Tonsil-L	SCC	HPV-	P16-		Yes
5	HN11-05T	M	59	Tonsil-L	SCC, MD	HPV-	P16-		Yes
6	HN12-02T	M	58	BOT	SCC, MD	HPV-	P16-	Yes	
7	HN12-03T	F	59	BOT	SCC, WD	HPV-	P16-	Yes	
8	HN17-02T	M	53	BOT	SCC, MD	HPV-	P16-	Yes	
9	HN18-02T	M	55	Tonsil-R	SCC, MD	HPV-	P16-	Yes	
10	HN21-03T	M	75	BOT	SCC, PD	HPV-	P16-	Yes	
11	HN23-02T	M	52	Tonsil-L	SCC, MD	HPV-	P16-	Yes	Yes
12	HN23-03T	M	51	Tonsil	SCC, MD	HPV-	P16-	Yes	Yes
13	HN30-06T	M	70	BOT	SCC, MD	HPV-	P16-	Yes	
14	HN35-02T	M	49	Tonsil-R	SCC, MD	HPV-	P16-	Yes	
15	HN39-05T	M	63	Tonsil-R	SCC, PD	HPV-	P16-	Yes	
16	HN03-03T	M	61	BOT	SCC, MD	HPV+	p16+	Yes	
17	HN15-04T	M	56	Tonsil-L	SCC, MD	HPV+	p16+	Yes	
18	HN15-06T	F	55	BOT	SCC, PD	HPV+	p16+	Yes	
19	HN22-04T	M	41	Tonsil-R	SCC	HPV+	p16+	Yes	Yes
20	HN24-02T	M	61	BOT	SCC	HPV+	p16+	Yes	
21	HN25-02T	M	38	Tonsil-L	SCC, MD	HPV+	p16+	Yes	Yes
22	HN27-02T	M	47	Tonsil-L	SCC, WD	HPV+	p16+		Yes
23	HN29-02T	M	62	Tonsil-L	SCC, MD	HPV+	p16+	Yes	Yes
24	HN31-02T	M	38	BOT	SCC, PD	HPV+	p16+	Yes	
25	HN31-05T	M	58	Tonsil-L	SCC, MD	HPV+	p16+		Yes
26	HN32-05T	M	65	Tonsil-L	SCC, MD	HPV+	p16+	Yes	Yes
27	HN33-03T	F	53	Tonsil-L	SCC,	HPV+	p16+	Yes	Yes
28	HN36-02T	M	54	OPH	SCC, MD	HPV+	p16+	Yes	
29	HN37-02T	M	53	Tonsil-R	SCC, PD	HPV+	p16+		Yes
30	HN43-05T	M	71	Tonsil-L	SCC, M-PD	HPV+	p16+	Yes	Yes

4.5.3. E6*I and E6*II constructs and stable transfectants

E6*I and E6*II ORFs were amplified by PCR (forward primer: atagatctgccaccatgcaccaaaagagaactgcaatgtt, reverse primer: attctagattaca gctgggtttctctacgtgttctt) from SiHa cells and cloned into pCEFL2 eukaryotic expression vectors [248]. The plasmids were linearised and transfected into the Oropharyngeal Squamous Cell Carcinoma, HN6 cell line [249] using Fugene 6 (Roche) according to the manufacturer's instructions. Cells that had stably incorporated the plasmid were selected by resistance to 1mg/ml G418 (Invitrogen), and those expressing the GFP-tagged E6*I and E6*II were purified by fluorescence-activated cell sorting (FACS) using a FACSAria (Becton Dickinson) [250]. Given these E6 variants have GFP tags, expression levels were then assessed with fluorescence imaging (Appendix Figure 4.1)

4.5.4. Cell culture and transfection

The following HPV- cell lines were use in this study included SCC003, SCC089, UMSCC22B, SCC4 and HN6 [251]. All cells were maintained in Dulbecco's Modified Eagle Medium (DMEM) with 1% glutamine, 10% foetal calf serum and 100mg/L penicillin/streptomycin (Invitrogen) in a 37°C incubator with humidified 5% CO2. All transfections were performed using LipofectamineTM RNAiMAX (Invitrogen) in a 6-well plate format and in triplicate, according to manufacturer's instructions. Pre-miR-496, and pre-miR-negative control#1 (Applied BioSystems) were reverse transfected at 10nM into 2.5 X 105 cells. Transfections from replicate wells were combined and divided into two equal parts for RNA extraction. Transfection efficiency of the pre-miR's was determined using qPCR. HN6 stable transfectants were monitored by GFP expression with fluorescent microscopy.

4.5.5. RNA isolation and microarray

100mg of fresh frozen tissue was diced with a surgical blade, homogenised with a pestle and mortar, and then rinsed with 1 ml of Trizol reagent (Invitrogen), or 5 X 10^6 cultured tissue cells were homogenised by adding 1 ml of Trizol. Total RNA was quantified using a NanoDrop ND 1000 (Thermo Fisher Scientific). A commercial LNA-modified oligonucleotide library (Exiqon) based on miRBase release 7.1, covering 371 human and mouse miRNAs, was utilised for expression profiling. Features were

deposited onto GAPS II slides (Amersham) at a concentration of 10 μM (Australian Genome Research Facility). Individual miRNA LNA probes were printed four times on each array. In addition, all samples were arrayed in technical duplicate. Total RNAs and a mixture of 371 synthetic DNA reference nucleotides (Sigma-Genosys) containing complementary sequences to all LNA probes were labelled separately and arrayed simultaneously, as described previously [243].

Raw data manipulation and statistical analysis (terrain mapping, hierarchical clustering) were performed using the TM4 suite (http://www.tm4.org) [252]. Normalisation of the raw data involved Lowes correction followed by in-slide replicate analysis. The data was then filtered by a percentage cut-off of 95% and subjected to statistical data mining. Two-class un-paired hierarchical clustering of samples was constructed with average linkage and Pearson correlation. Differential miRNA expression was then analyzed by Significance Analysis of Microarrays (SAM).

4.5.6. Quantitative PCR (qPCR) and statistical analysis

Total RNAs were briefly treated with RNase-free DNase (Promega). Expression of individual miRNAs was determined via qPCR. In brief, miRNA specific primers plus U6 primer (Applied BioSystems) were combined in equimolar concentrations into 500ng of total RNA. cDNA was generated with the Hi-Capacity cDNA Reverse Transcription Kit (Applied BioSystems). The cDNA was then diluted in a 1:4 ratio for 5.0 μl qPCR reactions. All qPCR reactions (including non-RT and no template controls) were conducted in triplicate using either the StepOne or 7900 Real-Time PCR System (Applied BioSystems) at the recommended cycle conditions using the TaqMan Universal PCR Master Mix, No AmpErase UNG (2x) (ABI, USA). Similarly, TaqMan assays were employed to determine the expression of the target genes, E2F2 (Hs00918090_m1), ECT2 (Hs00978168_m1), E2F7 (Hs00987777_m1), p16 (Hs00923894_m1) and E6 (custom design). The endogenous reference gene for these assays was B2M (Hs00187842_m1) labelled with a VIC™ probe for multiplexing or VIC™ labelled 16S (Hs02596860). Mean Ct values were normalised using nuclear RNA U6 (001093) for miRNA genes and B2M for target genes, and displayed as the Δ mean Ct. The relative expression level of a given gene was calculated using $2^{-\Delta\Delta Ct}$ and presented as fold change [253]. The significance of difference between gene expression levels was evaluated by a two tailed Student T-test.

4.5.7. Luciferase assays

The vectors, 29469.psiCHECK2-E2F2-3'UTR, 29468.psiCHECK2-E2F1-3'UTR, and 29470 psiCHECK2-E2F3-3'UTR were used for the Luciferase assay. Cells were seeded in triplicate at 4x104 into a 24-well plate. 0.5mL of the desired cell line at a predetermined concentration was added. After 24hrs, transfection was performed using Lipofectamine®.3000 Reagent. Each well was transfected with 500ng of plasmid DNA, 0.5µL of P3000TM reagent, 22.5µL of Lipofectamine®.3000 reagent and 20pmol of miR-496 mimic. A Pre-miRTM Negative control#1 (Ambion, USA) was used as a control. Cells were harvested 24hrs after transfection as per the Dual-Luciferase® Reporter 1000 Assay System (Promega, Australia), suggested protocol. Firefly activity was measured on the Tecan Infinite®200 Pro spectrophotometer. After this measurement Renilla activity was measured. Each Luminescence reading had a 2s pre-measurement delay followed by a 10s-measurement period. The Luciferase activity was expressed as a ratio of Renilla/Firefly.

4.5.8. HPV-Human interactome network

A network representing putative interactions between hsa-miR-33a-5p, hsa-miR-496-5p, HPV E6 and E7 proteins was built from publicly available datasets, using predicted miR targets and candidate regulatory transcription factors (TFs). The complete list of miRanda-predicted 'good mirSVR score and conserved miRNA' target sites for H. sapiens (August 2010 release) was downloaded from microRNA.org and filtered to include only the entries for hsa-miR-33a-5p and hsa-miR-496-5p. The putative TFs regulating expression of both miRs were inferred from transcription factor binding sites (TFBS) up to 6 kb upstream of miR-33a and miR-496 genes using UCSC Genome Browser version GRCh37/hg19 annotations. Since miR-33a is an intragenic microRNA encoded in intron 16 of the Sterol Regulatory Element Binding Transcription Factor 2 (SREBF2), the TFBS for its host gene were also considered.

The interactome data for the Papillomaviridae family was downloaded from the virus mentha website (virusmentha.uniroma2.it) and filtered to retain only the interactions between HPV16 and human proteins. The resulting interaction table (HPV16-host interactome) was loaded into Cytoscape v.3.3.0. The complete human interactome was loaded from BioGrid (version 3.4.129) through Cytoscape. The HPV16-host interactome was then merged with the human interactome to identify second order

neighbours of viral proteins. This expanded HPV-16-host interaction network was filtered to generate a E6/E7-focused regulatory network of putative miR-33a and miR-496 targets and TFs and connecting nodes.

4.6. Results

4.6.1. Patient cohort and HPV status

Fifteen of the 30 (50%) cancers were HPV16+ by qPCR detection of the HPV16 E6 gene. These same 15 samples were also positive for p16 as detected by immunohistochemistry staining (Table 4.1). The remaining 15 samples tested negative for both HPV16 DNA and p16.

4.6.2. miRNA signatures can differentiate HPV positive and HPV negative oropharyngeal SCCs into two separate clusters

We measured the expression of miRNAs in 14 cancers (nine HPV+ and five HPV-) using an LNA miRNA array. To determine if the miRNA signatures could differentiate HPV+ samples from HPV- cancers, we mined the data using 3D terrain maps, which depict related samples as peaks [254]. From this we noted a cluster of peaks in the centre of the terrain map (Appendix Figure 4.2A). These peaks corresponded to HPV+ cancers, whereas the surrounding peaks represented the HPV- samples. To identify specific miRNAs and further support these expression profiles, a two class SAM analysis with a two-fold cut-off and an FDR of 0% (Figure 4.1A) revealed five up-regulated and eight down-regulated miRNAs in HPV+ cancers (Table 4.2). Moreover, the miRNAs altered in the HPV16+ samples were distinct from those in the negative samples. The next stage was to validate these observed miRNAs using an alternative detection method. To this end, we measured the expression of selected differentially expressed miRNAs in 26 of the 30 samples using a probe-based qPCR method.

The miRNAs, miR-33, miR-210, miR-142-3p and miR-33 were shown to be significantly different in their expression levels between HPV16+ and HPV16- oropharyngeal cancers (Figure 4.1B-F). These qPCR results demonstrated a strong concordance with the array expression data. However, qPCR measurements for miR-520d* showed no fold change in contrast to a three-fold decrease in miR-520d* as measured by the array platform.

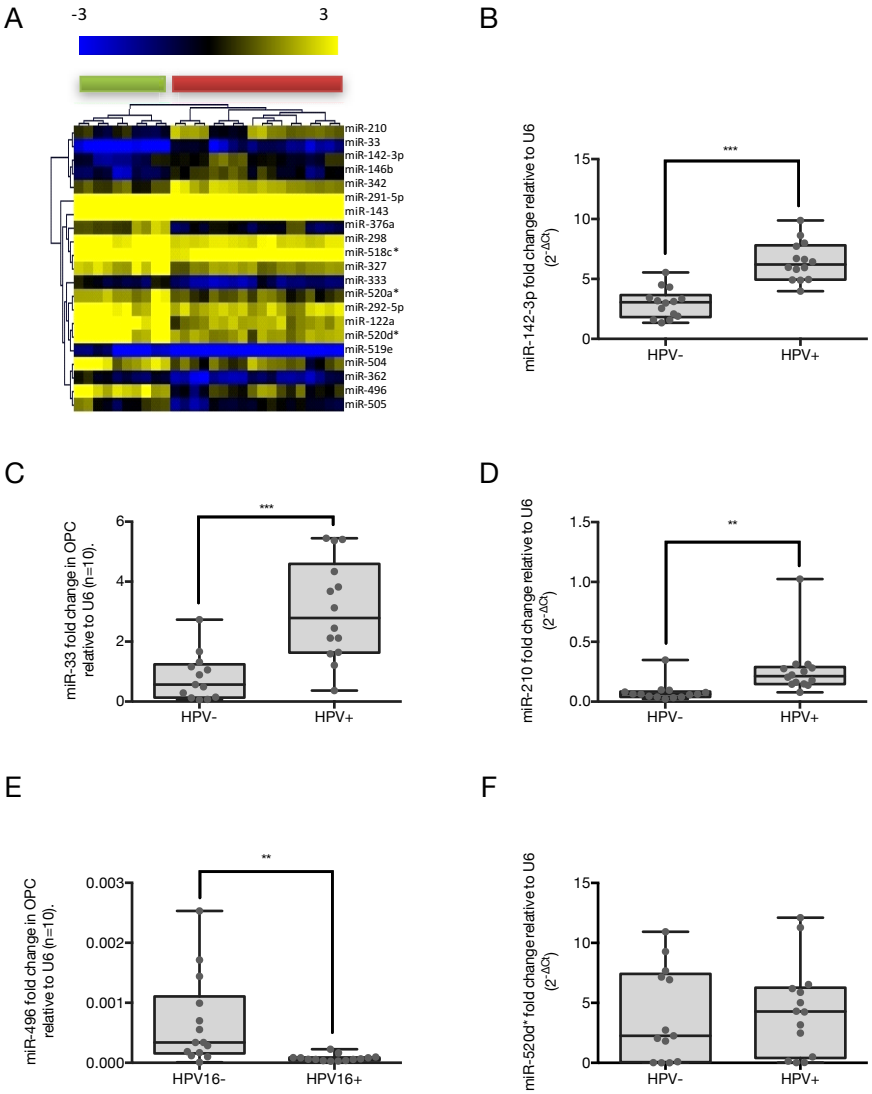


Figure 4. 1: Deregulation of specific miRNAs in HPV positive and HPV negative oropharyngeal (tonsillar) cancers. A) A constructed heat map showing differential miRNA expression between HPV+ (red bar) and HPV- oropharyngeal tumours (green bar). miRNAs shown in yellow are up-regulated, whilst miRNA shown in blue are down-regulated. B) Box Whisker plots showing qPCR confirmation of miR-142–3p in HPV+ and HPV- samples. This validation was then repeated for C) miR-33, D) miR-210, E) miR-496 and F) miR-520d*. Normalisation of miRNA expression was achieved by using the calibrator U6 and fold change expressed relative to U6 using the 2-ΔCt calculation. Statistical analysis was performed using two tailed Student Ttest, *p < 0.05, **p < 0.01, ***p < 0.005.

Table 4. 2: Identification of miRNAs in HPV+ tonsil SCC were compared to HPV- tonsil SCC tissue. A two-fold cut-off was imposed resulting in five up- regulated and eight downregulated miRNAs.

Name	Array fold change
miR-33	4.74
miR-210	2.73
miR-146b	2.50
miR-342	2.29
miR-142-3p	2.06
miR-518c*	-2.08
miR-505	-2.12
miR-520a*	-2.27
miR-362	-2.31
miR-519e	-2.50
miR-504	-3.28
miR-520d*	-3.56
miR-496	-4.30

4.6.3. HPV16 E6 regulates expression of miR-33 and miR-496

The HPV16 oncogene E6 is a major contributor to HPV-induced carcinogenesis [255, 256]. Over-expression of E6 in the epithelial basal layer of transgenic mice causes benign skin tumours [257]. Furthermore, E6 can also modify the malignant phenotype of these tumour cells [257]. The E6 transcript, through RNA splicing, gives rise to two alternative species: E6*I and E6*II. Full length E6 and E6*I can interact with E7 and Ras to transform cells *in vitro*, whereas E6*II cannot [258]. To determine the role of E6 in altering the expression of specific miRNAs in HPV16 oropharyngeal cancers, we transfected E6*I or E6*II into HPV- HN6 cells (Figure 4.2). These cell lines were previously characterised and shown to be expressing the E6 variants [250].

We reasoned that the most deregulated miRNAs, miR-33 and miR-496, may be more partial to E6 regulation. On this basis, stable over-expression of E6*1 and E6*II was marked by a concomitant reduction in miR-496 when compared to control cells. A similar pattern showing an increase in miR-33 in the presence of E6*1 and E6*II was also noted (Figure 4.2A). This suggests that both E6*I and E6*II could modulate miR-33 and miR-496 levels in head and neck cancer cells.

A single miRNA can regulate hundreds of targets and identifying the correct target genes can be problematic. To understand the regulatory role of these HPV16 specific miRNAs, we utilised the web-based computational tool TargetScan [238] to identify potential targets. For this study, we focused on miR-496 (the highest down-regulated miRNA) as there are no studies describing its regulatory role in oropharyngeal cancers. Our analysis indicated that E2F transcription factor 2 (E2F2) and the epithelial cell transforming 2 oncogenes (ECT2) (Appendix Figure 4.2B) could be targets for miR-496 regulation.

We then examined the expression of E2F2 and ECT2 in the same E6 over-expressing cells (Figure 4.2B). One of the consequences of reduced miRNA expression would be loss of regulatory control, observed by an increase in the target gene. From the results, a reduction of miR-496 was accompanied by a marked increase in RNA levels of E2F2. In contrast, there was no significant change in ECT2. The other targets E2F7 and p16 do not contain any miR-496 binding sites and accordingly did not alter their expression.

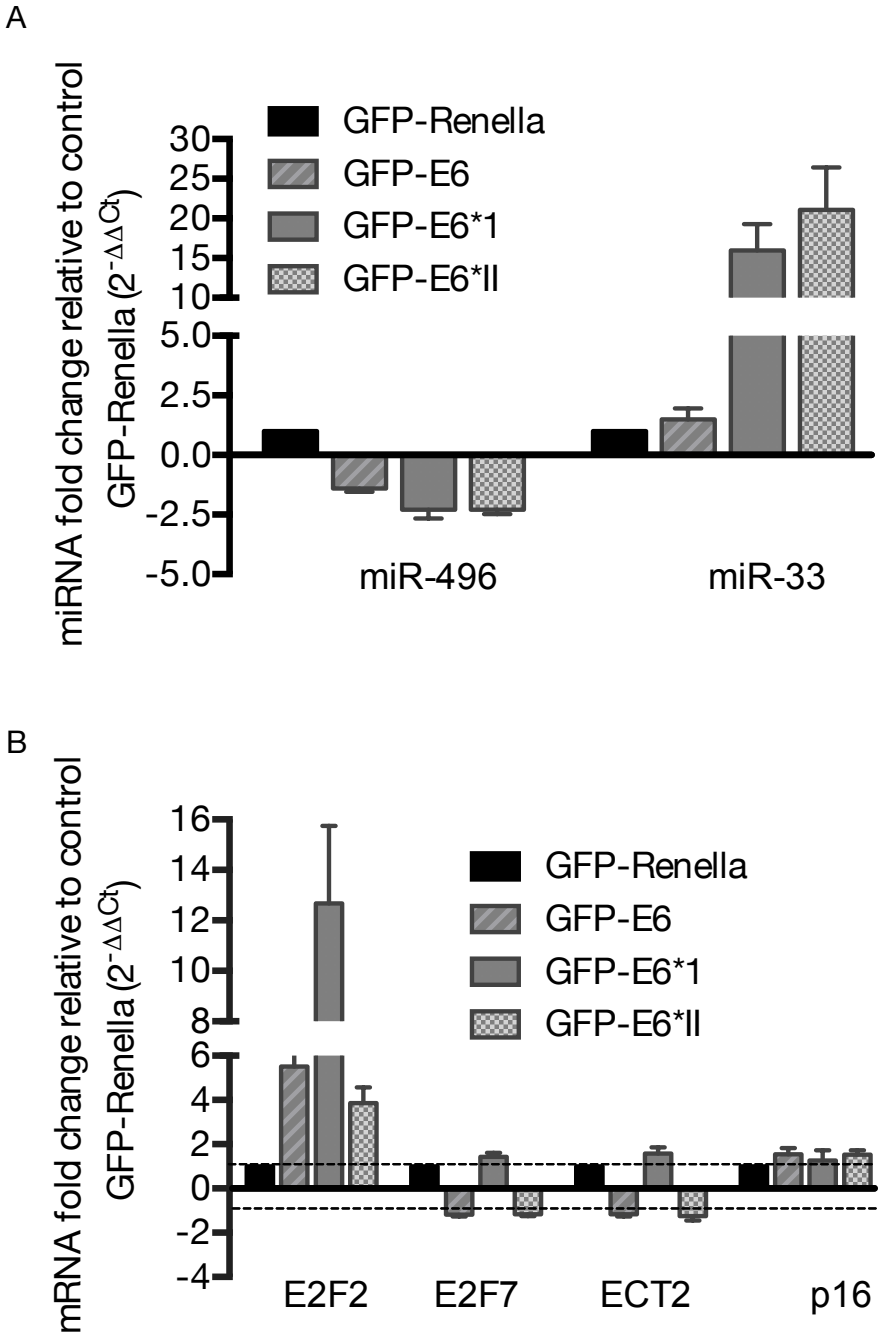


Figure 4. 2: Overexpression of E6 can regulate miR-33 and miR-496. A) The stable overexpression of the high risk HPV16 E6 and its variants was able to increase expression of miR-33 and decrease expression of miR-496. The levels of these two miRNAs were normalised to U6 with fold change expressed relative to the control cells GFP-Renilla using the 2-ΔΔCt method. B) RNA expression of miR-496 targets in the same cells overexpressing E6 and its variants. There is a marked increase for E2F2 while E2F7, ECT2 and p16 remain unchanged. Measurements of the RNA was normalised using B2M with fold change expressed relative to the control cells GFP-Renilla using the 2-ΔΔCt method. The dotted line represents a fold value of 1.0 which is indicative of no change in expression

4.6.4. Direct post-transcriptional regulation of E2F2 by miR-496 in Head and Neck cancer cell lines

To prove a direct relationship between miR-496 and E2F2, we transfected the miR-496 mimic into two HPV16- oropharyngeal SCC cell lines, SCC003 and SCC089. RNA levels of E2F2 were then measured using qPCR (Figure 4.3A and 3B). Over-expression of miR-496 caused a significant and robust decrease in the RNA expression of E2F2 in SCC003 and SCC089 cells. As expected, there was no change in E2F7, ECT2 or p16. To further consolidate these findings, we delivered the miR-496 into other head and neck cancer cell lines (UMSCC22B and SCC4) and observed a similar reduction in E2F2 mRNA (Appendix Figure 4.3A-D).

The 3´UTR of the E2F2 harbours a single miRNA binding site for miR-496. As shown above, the overexpression of miR-496 results in the downregulation of E2F2 RNA. To assess whether miR-496 can directly bind to and regulate E2F2 we performed a Dual Luciferase Reporter Assay. Cells were transfected with the psiCHECK-2 luciferase vector harbouring the 3´UTR of E2F2 in combination with either miR-496 mimic or Pre-miR Negative control. Additionally, cells were also transfected with psiCHECK-2 plasmids harbouring the 3´UTR of E2F1 and E2F3 to demonstrate that miR-496 only binds to a single member of the E2F family, E2F2. In the UMSCC22B cells, we noted a 30% reduction in luciferase activity in the cells containing psiCHECK-E2F2 and miR-496, compared to the scramble control (Figure 4.3C). In contrast, cells transfected with psiCHECK-E2F1 or psiCHECK-E2F3 with miR-496 showed no reduction in luciferase activity. This indicates that miR-496 only targets the 3´UTR of E2F2. The same results were also observed in a second independent cell line (Figure 4.3D). These results support the model that miR-496 can directly regulate the transcription factor E2F2 at the post-transcriptional level.

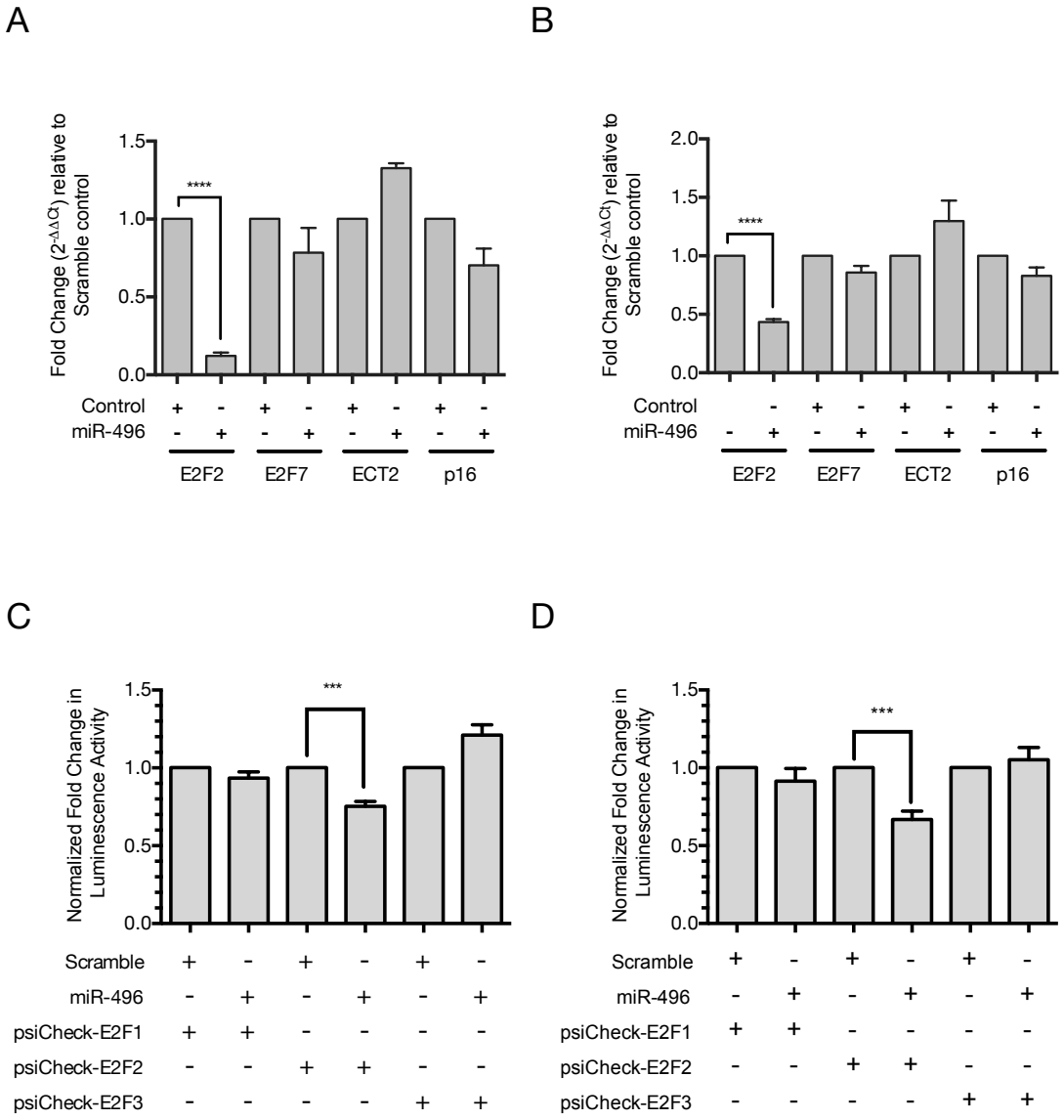


Figure 4. 3: Regulation of E2F2 by miR-496 in various head and neck cancer cells. A) SCC003 cells were transfected with the miR-496 mimic and the scramble control mimic. E2F2 was significantly reduced at the RNA levels in cells overexpressing miR-496. B) SCC089 transfected with the miR-496 mimic and the scramble control mimic. E2F2 RNA was significantly reduce in the presence of miR-496. The other gene targets did not show reduction in RNA levels. All transfections were performed in biological triplicate with fold change measured relative to the control scramble mimic. C) Direct regulation of the E2F2 3 UTR by miR-496 in UMSCC22B cells. D) Direct regulation of the E2F2 3 UTR by miR-496 in SCC4 cells. Both cell lines were transfected with either psiCheck-E2F1, psiCheck-E2F2, and psiCheck-E2F3 in combination with miR-496 or the scramble control mimic. These plasmids harboured the 3 UTR of E2F1, E2F2, E2F3. There was a 30% reduction luminescence activity in cells containing psiCHECK-E2F2 and miR-496. There was no reduction on the other E2F family members, E2F1 and E2F3. All statistical analysis was performed using two tailed Student T-test, *p < 0.05, **p < 0.01, ***p < 0.005, ****p < 0.0

4.6.5. HPV16 interactome analysis identified potential mechanisms regulating miR expression

We then sought to gains some insight into the interplay between HPV infection and miR regulation, focusing on HPV16 E7 and E6 proteins, and miR-33a and miR-496 ass the top-most deregulated miRNAs found in our screen.

For this purpose, a network representing potential interactions between miR-33 and miR-496, their predicted targets and putative TFs and viral proteins was built in Cytoscape using publicly available data (see Appendix Table 4.1). Key interactions with regulatory potential identified using this approach and are resented in Figure 4.4. Our analysis did not identify any direct interactors of E6 or E6 (1st neighbours') that could act as a putative miR-33a TFs, while two putative miR-496 TFs (Fos and Jun) have reported direct interactions with the E7 protein [259]. We further identified as first neighbours of this protein multiple TFs that present binding sites upstream of the miR-33a host gene SREBF2. Among these, the transcriptional factors EP300 and MYC are also first neighbours of the E6 protein (Figure 4.4). This evidence suggests that miR-33a expression could be regulated indirectly through the transcriptional control of SREBF2. Interestingly, a connection between HPV16-E6 and SREB signalling activation has been previously hypothesised [260].

In order to get a deeper insight into the transcriptional regulation of miR-33a and miR-496 in connection to HPV16 infection, we expanded our network to include second order neighbours of these proteins. This led to the identification of seven TFs with binding sites upstream of miR-496 gene and six potential direct regulators of miR-33a transcription, all of which have indirect, second order interactions with both E6 and E7 proteins. Some of these TFs (MAFK, USF1 and CEBP for miR-33a and GATA2 for miR-496) may be involved in negative regulatory feedback loops as miR-targets. Most interestingly, miR-33a is predicted to target four of the candidates miR-496 transcription factors (CTCF, BACH1, STAT5A and GATA2), creating the potential for negative regulatory interaction between both microRNAs.

To further explore the impact of these two miRNAs, we performed a functional enrichment analysis of miR-33a and miR-496 specific targets in our network against

the human genome, using the DAVID online tools [261]. Interestingly, miR-496 targets that were captured as part of the E6 and E7 interactome display a significant association to pathways and biological processes related to viral carcinogenesis (Table 4.3), whereas miR-33a targets were found to be associated to the control of transcription by RNA polymerase II, namely in the context of retroviral infection.

These observations suggest that the downregulation of miR-496 is likely to have positive impact on the progression of HPV infection and viral carcinogenesis.

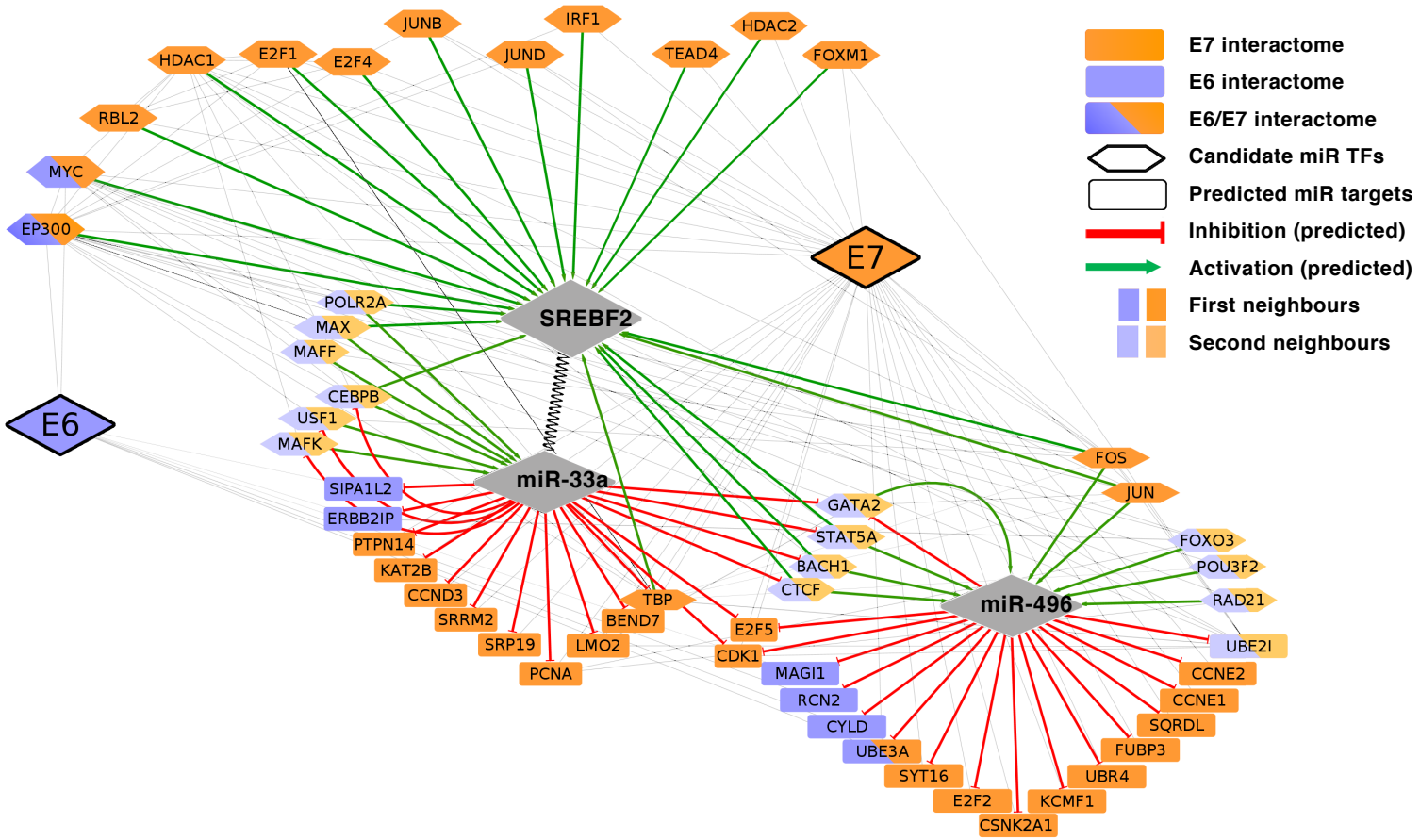


Figure 4. 4 Building an HPV interactome. Modelling of a HPV16 interactome describing 1st and 2nd TFs neighbours for E7 and E7. The network also incorporates miR-33 and miR-496 to discover possible mechanistic linkages and associations.

Table 4. 3: Enrichment analysis of miR targets

Category	Term	p value	Genes
Enriched terms for miR-496 targets			
KEGG pathway	hsa05203: Viral carcinogenesis	0,00189376	CCNE2, CCNE1, UBE3A, UBR4
	hsa05206: microRNAs in cancer	0,00484572	CCNE2, CCNE1, E2F2, UBE2I
	hsa05222: Small cell lung cancer	0,00509009	CCNE2, CCNE1, E2F2
	hsa05215: Prostate cancer	0,00544687	CCNE2, CCNE1, E2F2
GO term Biological Process	0051726: regulation of cell cycle	0,00284822	CCNE2, CCNE1, E2F2
	0006511: ubiquitin-dependent protein catabolic process	0,00602552	CYLD, UBE3A, UBE2I
	0016055: Wnt signaling pathway	0,00635072	CCNE1, CYLD, CSNK2A1
Enriched terms for miR-33a targets			
KEGG pathway	hsa05166: HTLV-I infection	0,00157462	KAT2B, CCND3, PCNA, TBP
GO term Biological Process	0045944: positive regulation of transcription from RNA polymerase II promoter	0,00274308	KAT2B, CEBPB, LMO2, MAFK, USF1
	0006366: transcription from RNA polymerase II promoter	0,00389425	CEBPB, TBP, MAFK, USF1

4.7. Discussion

Human papillomavirus 16 is now considered a key etiologic factor for SCC of the oropharynx [66, 256, 262]. It is a complex disease and we do not fully understand how regulatory pathways contribute to the development of oropharyngeal cancers. Given that large regions of the human genome are non-coding, our study has focused on the relationship between HPV16 with specific miRNAs and their targets.

The array data in our study was able to separate HPV16+ oropharyngeal tumours from samples without the HPV virus. This was demonstrated by the tight cluster of peaks in the terrain map and the HCL clustering. This screen identified only a small pool of deregulated miRNAs and several of these, including miR-33, have been previously described [85, 245, 263]. However, the discovery that miR-496 is lowly expressed in oropharyngeal cancers is unique to this study. The two most deregulated microRNAs, miR-33 and miR-496, were then independently validated using qPCR.

We then explored if these miRNAs could be regulated by the viral protein E6. Using head and neck cancer cell lines already shown to be overexpressing different E6 variants [250], it was demonstrated that E6 could increase miR-33 levels. The regulation of miR-33 by HPV has been previously shown and our data reinforces these initial studies [245]. As miR-496 has not been characterised in oropharyngeal cancers, we show that the viral oncoprotein E6 can decrease the mature levels for miR-496. The changes in both miR-33 and miR-496 using this *in-vitro* cell model are in agreement with the biological changes observed in our clinical samples.

To further understand the functional role of miR-496 in oropharyngeal cancers we examined several possible targets, E2F2 and ECT2. E2F2 is part of the E2F transcription factor family which interacts with the retinoblastoma protein (pRb) to promote cell proliferation. Degradation of pRb through interaction with the HPV oncoprotein releases E2F, allowing recruitment of proteins needed for DNA replication, leading to overexpression of p16 through a feedback loop. Notably, increased levels of E2F2 were needed for HPV genome replication in immortalised human foreskin keratinocytes [264]. ECT2 is a proto-oncogene which may activate the

Rho signalling pathway to promote malignant transformation [265]. Additionally, ECT2 was shown to be elevated in HPV+ head neck cancers [266, 267].

Overexpression of the miR-496 mimic significantly reduced the RNA levels of E2F2 but had no effect on ECT2 or the other control gene targets. This observation was then repeated using two different head and neck cancer cell lines at different miR-496 concentrations with similar results. To ascertain a direct binding role between E2F2 and miR-496, cancer cell lines were transfected with miR-496 and the psiCHECK-2 plasmid harbouring the 3`UTR of E2F2. Furthermore, the 3`UTR of other E2F members, E2F1 and E2F3, were used to determine miR-496 specificity towards only E2F2. In both the UMSCC22Bs and the SCC4s, there was a significant 30% reduction in luciferase activity when miR-496 was present. There was no reduction observed in cells co-transfected with miR-496 and the plasmids harbouring the 3`UTR of E2F1 and E2F3. This reduction in luminescence suggests that miR-496 is able to bind to and regulate only E2F2. Interestingly, the reduction was only observed at 30%, but the reduction in E2F2 RNA was more pronounced. Taken together, we are confident that miR-496 exerts post-transcriptional control on the expression of E2F2 in oropharyngeal cancer cells.

There are no current studies investigating miRNA regulation of E2F2 in Head and Neck cancer and no studies showing E2F2 regulation by miR-496 specifically. Lal *et al* investigated E2F2 regulation by miR-24 in non-cancerous cell lines and showed a 50% reduction in luminescence [268], a rate comparable with our findings and is consistent with the theme that miRNAs can regulate E2F2. E2F2 plays an important role in viral pathogenesis, as reductions in E2F2 can decrease the copy number of HPV by up to 60% without affecting cell proliferation [264]. Our study has identified a potential pathway whereby HPV16 downregulates miR-496, with a consequent effect on E2F2 expression.

To provide a broader understanding of other possible pathways and the association between miRNAs and HPV16, we constructed a merged HPV -human interactome. This network analysis allowed us to visualise all possible connections. We believe there is a plausible connection between HPV16, miR-33 and miR-496 (the highest deregulated miRNAs in this study). From the network mapping, viral E7 and E6 can

directly interact with EP300 and MYC. These transcription factors are known to be associated with SREBP2, which contains the sequence for miR-33. From our results, we show that miR-33 is elevated in oropharyngeal cancer, and E6 can increase miR-33 levels. Based on the modelling, we proposed that the elevation of SREBP2 gives rise to higher miR-33 levels. The predicted targets for miR-33 included transcription factors (CTCF, BACH1, STAT5A, and GATA2) which are upstream of miR-496. An increase in miR-33 targets these transcription factors, resulting in a decrease in miR-496 levels. This speculative pathway suggests that miR-33 can regulate the expression of miR-496 via these specific transcription factors. Further studies will be required to prove this direct relationship.

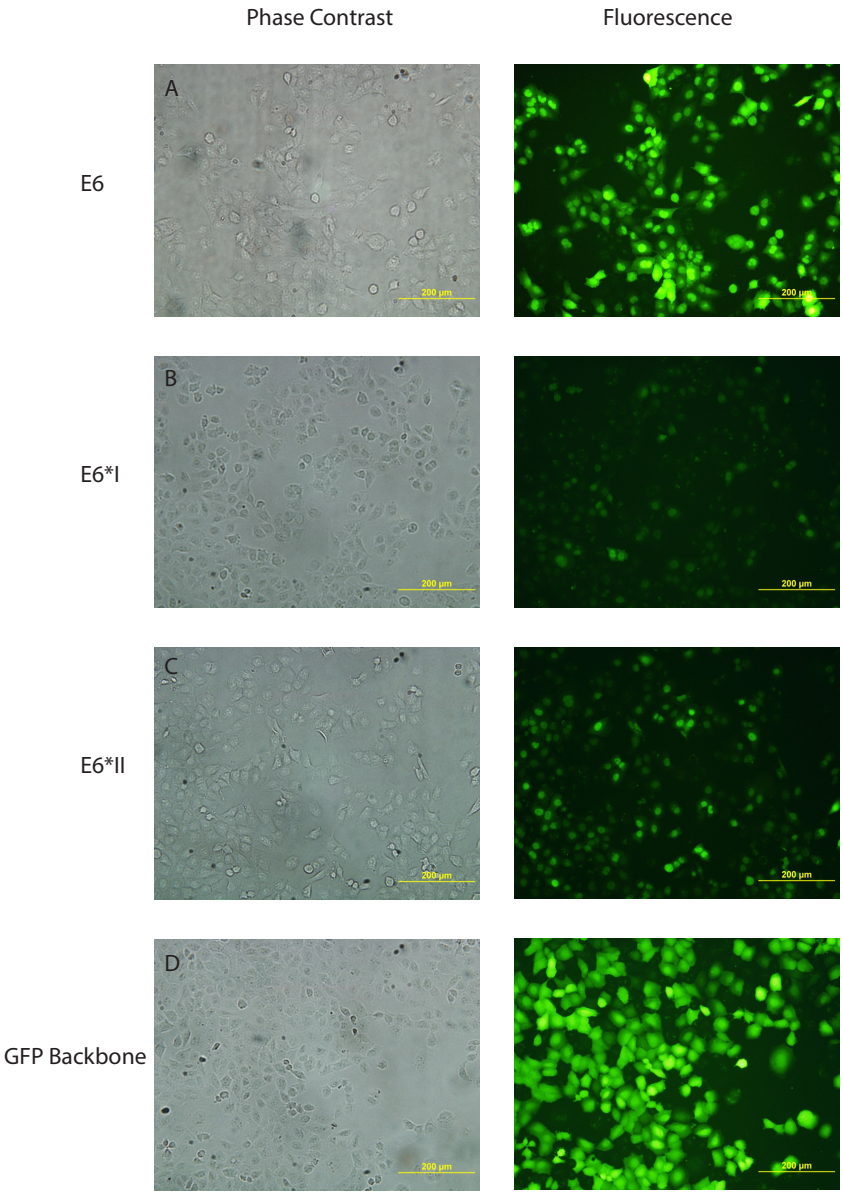
Most interestingly, we find that a large number of predicted miR-496 targets in addition to E2F2 are direct interactors of the E7 protein and have been linked to the promotion of viral carcinogenesis. These observations suggest that the regulatory interplay between the E6 protein and miR-33a moves the regulatory environment of the cell towards a pro-oncogenic environment through the down regulation of miR-496

Although this type of network analysis is speculative, it does have the power to provide a targeted framework to discern potential regulatory pathways. By identifying specific interactors, this type of approach can lay the foundation for future mechanistic studies. The mapping itself is amenable to either a focused approach to identify specific pathways as shown in our study or can be modified to provide a more complex understanding of global interactions.

In summary, our study has provided a greater understanding of the regulatory relationship between HPV16, miRNAs and target genes in oropharyngeal cancer. We identified a specific regulatory pathway involving E6, miR-496 and E2F2. These events may be important for viral replication and ultimately transformation of the cell. Furthermore, an interactome was created to offer a greater appreciation of these relationships. This approach identified one possible mechanistic link which warrants further investigations. Taken together, the unique interactions uncovered between HPV and miRNAs in this study may provide a significant insight into the cellular mechanisms in HPV+ oropharyngeal SCC, and the aetiology of the disease.

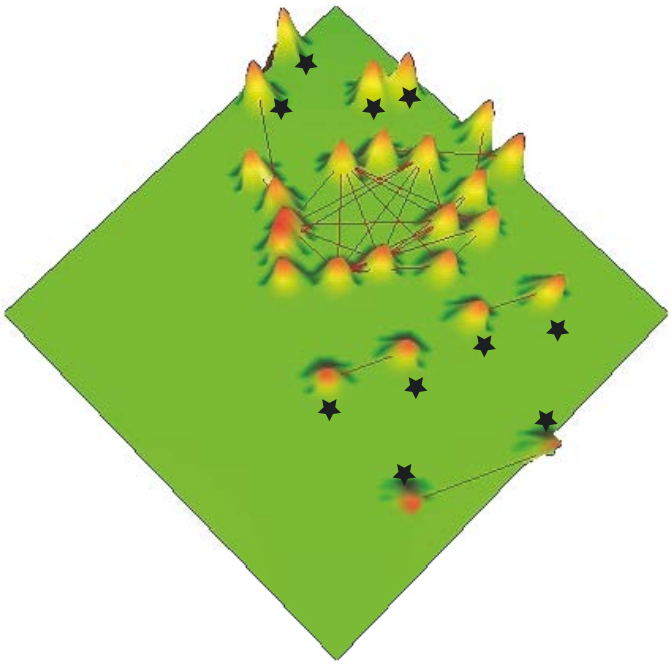
4.8. Appendix

As this chapter is published as a Research article in the Journal *Virology* all supplementary Figures are represented here in the appendix and all supplementary tables can be found via: https://doi.org/10.26195/mdzm-qc54

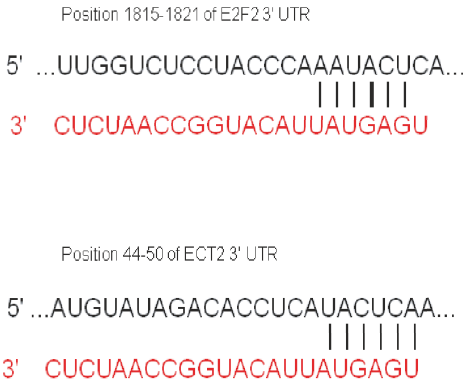


Appendix Figure 4. 1: Fluorescently sorted GFP/E6s and GFP HN6 cells. The fluorescent intensities for the positive GFP control cells (Panel D) are greater compared to the EGFP/E6s-transfected cells (Panels A, B and C). However, approximately 90% of sorted cells were considered to be positively fluorescence with the expression of EGF.

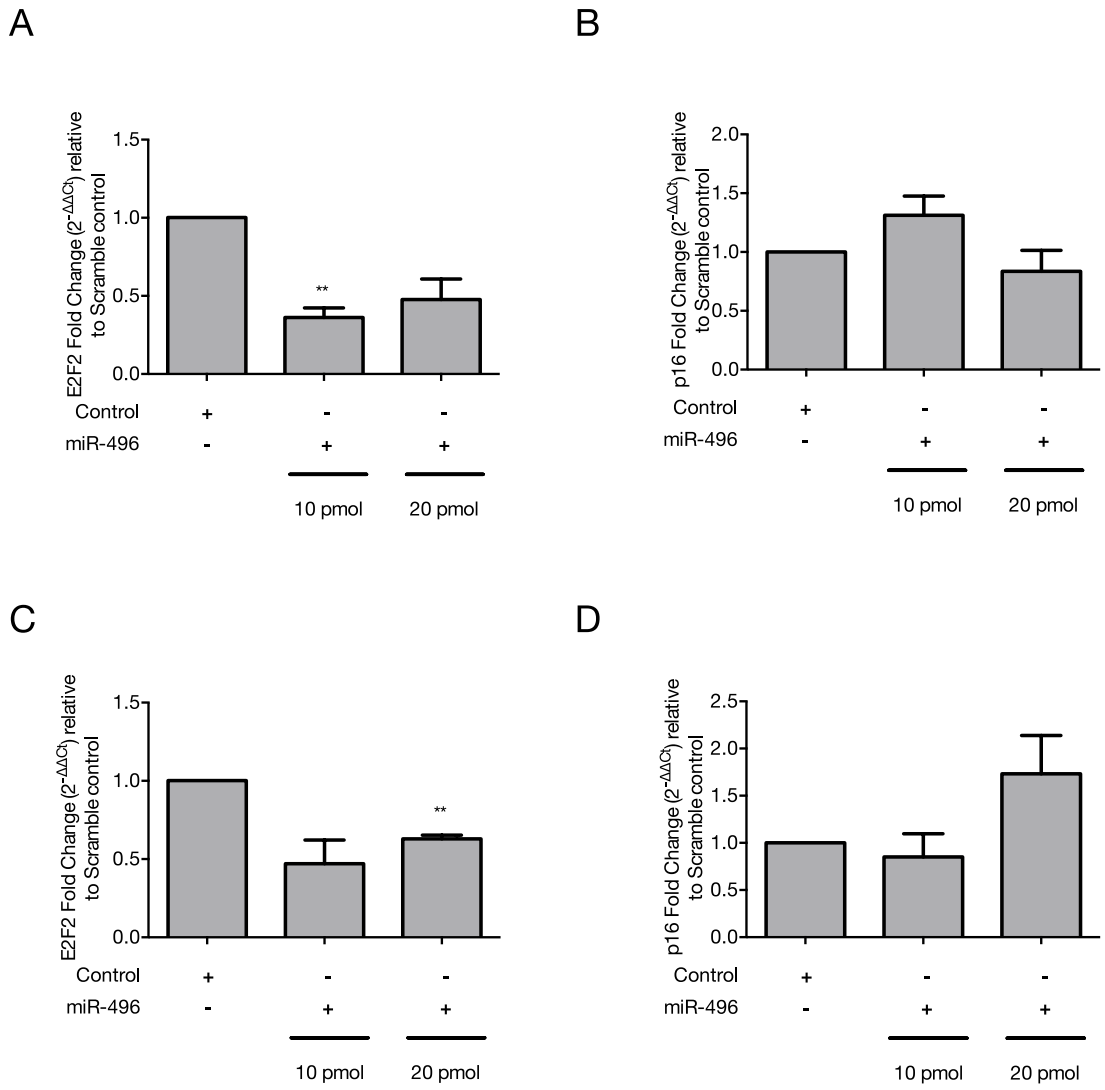
A



B



Appendix Figure 4. 2 A) 3D Terrain map showing a cluster of central peaks representing HPV+ oropharyngeal tumours. Surrounding peaks are HPV- samples and denoted by black stars. B) Schematic representation of potential miR-496 binding sites on the 3'UTR of the target genes E2F2 and ECT2.



Appendix Figure 4. 3: E2F2 and p16 fold change with miR-496 overexpression. A: E2F2 fold change in UMSCC22B, B: p16 fold change in UMSCC22B, C: E2F2 fold change in SCC4, D: p16 fold change in SCC4. E2F2 RNA is downregulated in both the UMSCC22B and the SCC4s. In the UMSCC22B downregulation at 10pmol is statistically significant and in the SCC4 downregulation at 20pmol is statistically significant ($P<0.005$: **). There was no statistically significant change in p16 RNA expression levels in both cell lines. The levels of these two genes were normalised to B2M with fold change expressed relative to the cells transfected with a scramble control, using the $2^{-\Delta\Delta Ct}$ method.

Chapter 5: HPV16 E6 and E7 oncogenes can alter the expression of SREBF2 and miR-33a in Head and Neck cancers

5.1. Introduction

The Human papillomavirus (HPV) accounts for up to 60-90% of all OPC cases, and the incidence of HPV associated OPC is increasing every year [4, 9]. The high-risk variant HPV16 encodes for two main oncogenes, E6 and E7, which are consistently expressed in cancer cells and are necessary for induction and maintenance of the transformed state. Non-coding genes such as microRNAs (miRNAs/miRs) have been reported to be targeted by HPV16 [108, 269]. Our current understanding of miRNAs and HPV16 in OPC is limited, and we are only beginning to uncover these interactions.

Previously our lab conducted an LNA-array on oropharyngeal tissues, comparing HPV16+ vs HPV- (HPV16 and HPV18 negative) tumours, to identify critical miRNAs that HPV16 may regulate [101]. From this, we identified miR-33a to be the most upregulated and miR-496 to be the most downregulated miRNAs in HPV16+ OPC. The expression of these miRNAs were altered with E6 overexpression. We also developed an HPV-miRNA interactome to visualise novel mechanistic links between the HPV16 viral oncogenes and the miRNAs. This current study will expand on this observation and validate potential mechanistic pathways between HPV16 and these specific miRNAs. Of note, miR-33a is encoded in the intron of the Sterol regulatory binding factor 2 (SREBF2). The SREBF2 transcription factor is a crucial protein in cholesterol metabolism through the transactivation of cholesterol biosynthesis genes and the LDL receptor [270, 271].

In particular, we will determine if SREBF2 and miR-33a expression can be influenced by HPV in OPC and if the viral oncogenes E6 and E7 can regulate SREBF2 expression. Furthermore, we will assess if SREBF2 and miR-33a are co-expressed in HPV16+ OPC. These results will provide evidence of a mechanistic pathway between HPV16 E6 and E7 and their effect on specific miRNAs via SREBF2.

5.2. Methods

5.2.1. Mining: The Cancer Genome Atlas

The databases TCGA-Head and Neck cancer and the TCGA-Cervical cancer datasets https://www.cancer.goc/tcga were mined to explore specific expression data. These datasets were filtered using the online visualisation tool XenaBrowser [272]. The Head and Neck cancer dataset was filtered to include oral and oropharyngeal cancers that have been tested for HPV. The cervical cancer dataset was also filtered for HPV positivity. HPV status and type were also further validated by RNA seq analysis done by the TCGA for HPV genomes [209, 273].

5.2.2. Cell lines

Isolated RNA from the following HPV16+ cell lines were used in this study, SCC090 (tongue) and SCC154 (tongue). The RNA from the following HPV- cell lines were used in this study, UMSCC22B (hypopharyngeal), UMSCC38 (tonsil), HN12 (tongue), HN13 (tongue), HN31 (pharynx lymph node), HN17 (laryngeal), HN30 (pharynx), HN26 (oral cavity) and SCC6 (oropharynx).

5.2.3. Cell culture and transfections

The HPV16+ cells SCC154 and SCC090 were transfected with siRNA that were designed for the targeting of SREBF2 and HPV16 E6/E7; sequences for these siRNAs are described in Table 5.1 (IDT). siRNAs were forward transfected at 20pmol into $5x10^5$ cells. Cells were harvested 24hr and 48hr post-transfection for RNA and western blot analysis.

The HPV- HNC cell lines SCC4 were transfected with 500ng of the plasmids P1321 HPV16 E6/E7, P1324 HPV16 E7, P1322 HPV16 E6 and the control plasmid 3680.pB-actin, (Addgene) into $5x10^5$ cells. Cells were harvested 48hr post-transfection for RNA analysis.

Table 5. 1: Sequences for siRNA used in this study

Target gene	siRNA name	Sequence
SREBF2	siSREBF2 1	Sequence 1: 5`GGCCAUUGAUUACAUCAAAUACUUG`3 Sequence 2: 5`CAAGUAUUUGAUGUAAUCAAUGGCCUU`3
	siSREBF2 2	Sequence 1: 5`AUCGAGGACUUUAAUCAGGAUGUCC`3 Sequence 2: 5`GGACAUUCUGAUUAAAGUCCUCGAUCU`3
	siSREBF2 3	Sequence 1: 5`UUGGUCUGUUGGGGUAUUAAAUGA`3 Sequence 2: 5`UCAUUUAAUACCCCAACAGACCAA`3
HPV16 E6/E7	siHPV16 E6/E7 1 (E61)	Sequence 1: 5` GAGCUGCAAACAACUAUACAUGAUA`3 Sequence 2: 5`UAUCAUGUAUAGUUGUUUGCAGCUC3`
	siHPV16 E7/E7 2 (E7 1)	Sequence 1: 5` GAGAUACACCUACAUUGCAUGAAUA`3 Sequence 2: 5`UAUUCAUGCAAUGUAGGUGUAUCUC`3
	siHPV16 E6/E7 3 (E6 2)	Sequence 1: 5` GAUAAAUGUUUAAAGUUUUAUUCU A'3 Sequence 2: 5`UAGAAUAAAACUUUAAACAUUUAUC`3
	siHPV16 E6/E7 4 (E7 2)	Sequence 1: 5` GAUCUCUACUGUUAUGAGCAAUUAA`3 Sequence 2: 5`UUAAUUGCUCAUAACAGUAGAGAUC3`

5.2.4. Designing TaqMan assays for miRNA detection

A specific TaqMan assay for the targeting of hsa-miR-33a, consisting of stem-loop primers, PCR primers and TaqMan probes, was designed and based on the method described by Varkonyi-Gasic *et al.* [274]. This stem-loop primer consists of a Universal TaqMan sequence, a miRNA specific forward primer and a universal reverse primer for the miRNA. These primers were ordered from IDT, and sequences used are shown in Table 5.2

Table 5. 2: miRNA step loop and primer sequences

Component	Sequences 5`-3`
miR-33a forward primer	GTC TCC AGT GCA GGG T
miR-33a RT Stem-loop	5`-GTC GTA TCC AGT GCA GGG TCC GAG GTA TTC GCA CTG GAT ACG ACT GCA AT-3`
Universal Reverse Primer	AAC ACG CGT GCA TTG TAG TT
Universal TaqMan Probe	5`-/56-FAM/GTA TCC AGT/ ZEN/ GCG AA/31ABkFQ/-3`

MiRNA stem-loop cDNA (Table 5.3) was synthesised from isolated RNA and used for TaqMan© Quantitative-Polymerase Chain reactions (qPCR). cDNA synthesis was performed using the Applied Biosystems High-Capacity cDNA Reverse Transcription Kit (Life Technologies, USA) with the addition of the designed miRNA RT stem-loop primer. Prior to the cDNA stem-loop, RT stem-loop was denatured at 75^{o}C. Thermal cycling was run on a VapoProtect Mastercycler (Eppendorf, USA), and reaction cycle conditions were as per the manufacturer's protocol. Following the thermal cycling, the cDNA mixture was diluted 1 in 4 by adding nuclease-free water.

Table 5. 3: Stem-loop microRNA cDNA Synthesis protocol

Reagent	Volume (1x) (µL)
10x Reverse Transcriptase Buffer, 1.0mL	2.0
25x dNTP Mix 100mM, 200µL	0.8
RNase Inhibitor 100 µL, 20Units/µL	1.0
RNA Input (500ng/µL)	1.0
miRNA Stem loop primer (25µM)	2.0
MultiScribe™ Reverse Transcriptase 100µL, 50units/µL	1.0
Nuclease-free water	12.2
Total	**20.0µL Final Volume**

The TaqMan© qPCR assay was used for the analysis of the miR-33a. Table 5.4 represents the TaqMan® PCR protocol, modified for the miRNA stem-loop PCR reactions. All reactions were performed in triplicate using the StepOnePlus™ (Life Technologies, USA). The annealing temperature for miR-33a synthesis was reduced to 55^{o}C due to primer melting temperatures. The collected data were analysed using the ΔCT or $\Delta\Delta$CT method [208].

Table 5. 4: TaqMan© quantitative Polymerase Chain reaction protocol

Reagent	Modified protocol (1x) (µL)
Universal miRNA Probe (10µM)	0.2
miR-33a Primer pool (15µM) (Reverse and forward primers)	2.0
TaqMan® Master Mix (2x)	10.0
Diluted cDNA	2.0
RNAse-free Deionized water	5.8
Total	**20.0**

5.3. Results

5.3.1. The regulation of SREBF2 and specific miRNAs in HPV16 positive oropharyngeal cancers

Our previous *in vitro* data suggested a regulatory relationship between SREBF2, E2F2, miR-33a and miR-496 [101]. To explore if this connection had any clinical relevance, we mined the TCGA dataset and filtered the samples to only include HNC of the oral and oropharyngeal region as these are commonly infected with HPV. Data was also filtered to only include samples that have been determined to be HPV+ or negative by HPV-ISH and to be infected with the high-risk HPV16 [209].

From the TCGA, we retrieved 61 samples, twenty of which (33%) were HPV16+ (Appendix Table 5.1). We then plotted the Log2(Norm count+1) for SREBF2, E2F2 and Log2 (RPM+1) for miR-33a and miR-496 to determine if their expression was influenced by the presence of the virus (Figure 5.1). The analysis showed that SREBF2, miR-33a and E2F2 were upregulated, and miR-496 was downregulated in the HPV+ OPC tumours. This suggests that HPV may influence the expression of these specific genes and miRNAs. This pattern of expression is similar to our own clinical samples, as reported in a previous study [101].

We then asked if this expression pattern was evident in other HPV cancer types, such as cervical cancers. As HPV16 and HPV18 are major causes of cervical cancers, both high risk variants were assessed. We filtered the TCGA data for cervical cancers (CC), comparing HPV- to HPV16+ and HPV18+ (Figure 5.2). Of the 150 CC samples, 101 were HPV16+ (6.7%) and 26 (17%) were HPV18+. HPV status and type were determined by PCR and RNA sequencing [273].

Interestingly, we did not observe the same expression pattern of SREBF2, miR-496 and E2F2 in OPC for CC patients. There was no change in SREBF2, miR-496 or E2F2 in the CC tumours across different HPV statuses. We did, however, see an increase in miR-33a in the HPV16+ CC compared to the HPV- CC, but this was not seen for HPV18 infected tumours.

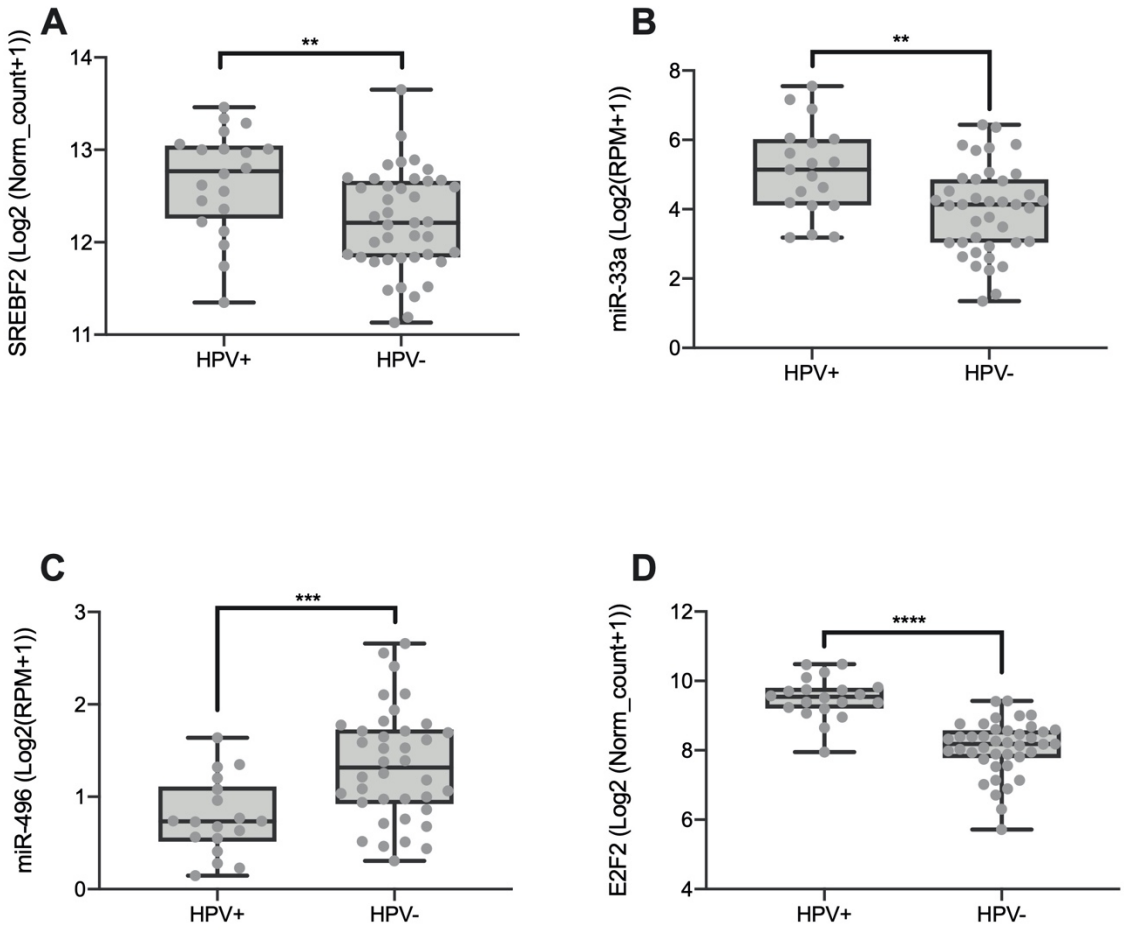


Figure 5. 1: Expression levels of specific genes and miRNAs in HPV16+ and HPV- Oropharyngeal cancer from TCGA database. Box whisker plots showing the Log2 of normalised counts for genes and normalised reads per million (RPM) for miRNAs comparing HPV16+ oropharyngeal tumour samples (n=20) and HPV- tumour samples (n=41). A) SREBF2 B) miR-33a C) miR-496 and D)E2F2. Statistical analysis was performed using two-tailed student T-test, **p<0.005, ***p<0.0005, ****p<0.0001.

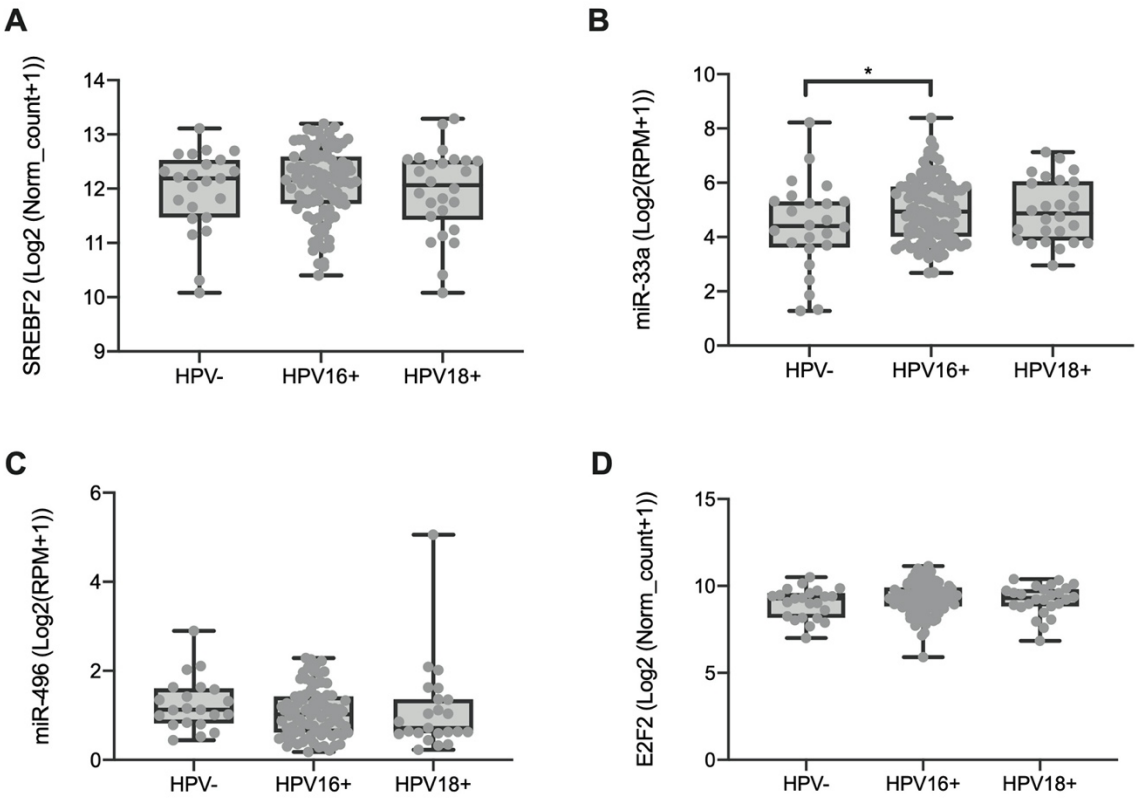


Figure 5. 2: Expression levels of specific genes and miRNAs in HPV+ (HPV16 and HPV 18) and HPV- Cervical cancer from TCGA database. Box whisker plots showing the Log2 of normalised counts for genes and normalised reads per million (RPM) for miRNAs between HPV- (n=23), HPV16+ (n=101) and HPV18+ (26) cervical cancer tumour samples. A) SREBF2 B) miR-33a C) miR-496 and D) E2F2 Statistical analysis was performed using two-tailed student T-test p*< 0.05.

5.3.2. Optimisation of PCR TaqMan assays for HPV16 E6 and E7 and miR-33a

For this chapter, we were required to design and optimise our own TaqMan assays for the detection of the HPV16 viral genes E6 and E7 as well as hsa-miR-33a. For HPV16, E6 and E7, custom TaqMan assays were designed and synthesised. These primers were then tested in a cell line panel consisting of the HPV16+ cells SCC090, SCC154 and SiHa, the HPV18+ cell line HeLa (to test for specificity to HPV type 16) and the HPV- cell lines MCF-7 and UMSCC38 as negative controls (Figure 5.3). HPV16 E6 and E7 were detected in the HPV16+ cell lines only. The average CT values for E6 and E7 were 26.3 and 26.6 for SiHa, 32.8 and 33.2 for SCC154, 28.2 and 29.8 for SCC090, respectively. HPV16 E6 and E7 remained undetected in control cell lines.

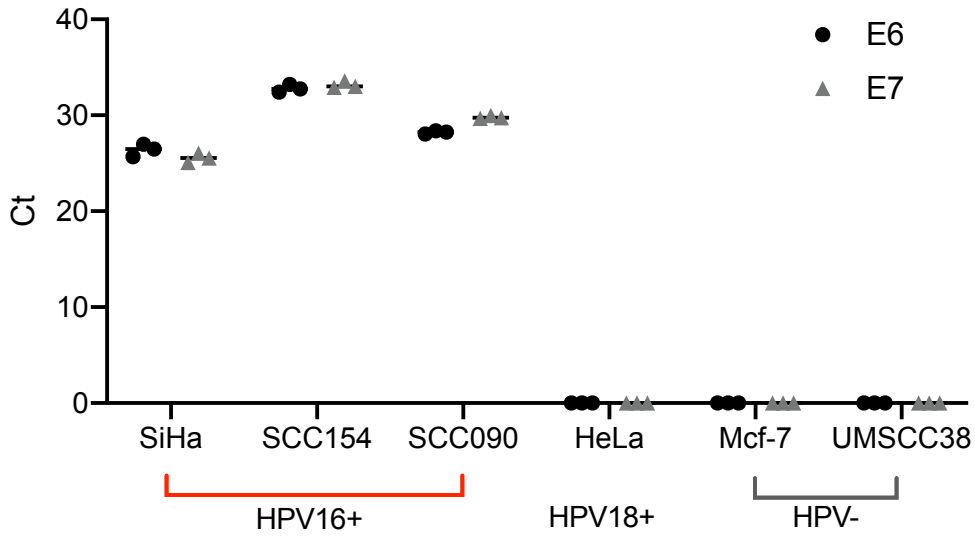


Figure 5. 3: Detection of HPV16 E6 and E7 in cell lines panel. Box whisker plots showing the CT values of E6 and E7 in the HPV16+ cell lines SiHa, SCC15, SCC090, the HPV18 positive cell lines HeLa and the HPV- cell lines MCF-7 and SCC38.

To measure the expression levels of miR-33a in our cell lines, we designed a custom RT stem-loop, PCR primes and specific TaqMan probes. These PCR components were optimised in various serial dilution experiments to determine the optimal conditions for miR-33a detection.

For the reverse transcriptase reaction (RT), the RT-stem loop (20μM, 25μM and 30μM) was denatured at 75°C prior to reaction as recommended Varkonyi-Gasic *et al.* [274]. During the PCR reaction, the annealing temperature was reduced to 50°C; due to the melting temperatures of the primers being 56.9°C and 55.6°C, using the standard 60°C would result in no amplification of any product. PCR primers were tested from 5μM, 10μM, 15μM and 20μM. The concentration of the Universal miRNA TaqMan probe (10μM) was based on the recommendations by Varkonyi-Gasic *et al.* [274]. Figure 5.4 represents the optimisation of the miR-33a PCR assay.

For the 20μM RT stem-loop, the average CTs which each primer concentrations (5μM, 10μM, 15μM and 20μM) were 33.8, 31.5, 27.5 and 33.0, respectively. For the 25μM RT stem-loop, the average CTs were 34.6, 29.4, 25.8 and 32.2, and for the 30μM RT stem-loop, the average CTs were 33.8, 31.4, 25.6 and 32.8. A lower concentration of PCR primers (5μM-10μM) resulted in too much variation between triplicate samples as well as high CT values (less product). From the optimisation, it was decided that 25μM of RT-stem loop and 15μM of the PCR primers gave optimal detection of miR-33a.

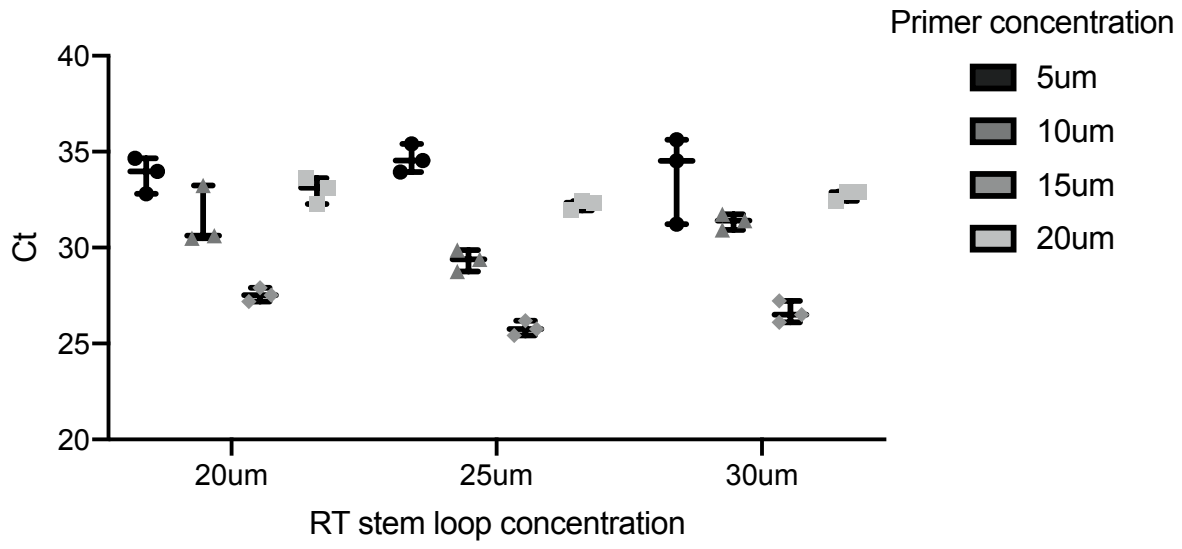


Figure 5. 4: Primer optimisation for miR-33a detection. Box whisker plots showing the CT values for miR-33a detection in a serial dilution of the RT-stem loop (20μM, 25μM and 30μM) and of the PCR primers (5μM, 10μM, 15μM and 20μM)

5.3.3. SREBF2 and miR-33a are overexpressed in HPV16 positive Head and Neck cancer cell lines

We know that miR-33a is encoded in the intron of SREBF2, to further explore their possible mechanistic relationship with HPV16, we profiled a panel of HPV+ and HPV- Head and Neck cancer cell lines. In Figure 5.5, the relative expression of SREBF2 and miR-33a was measured in HPV16+ HNC cell lines, SCC090 and SCC154 compared to the HPV- HNC cells UMSCC22B, UMSCC38, HN12, HN13, HN31, HN17, HN30, HN26 and SCC6.

HPV16 status for all cells was determined by RT-qPCR by measuring HPV16 oncogenes E6 and E7 (Appendix Figure 5.1). From these measurements, we can see that SREBF2 and miR-33a are both elevated in the HPV16+ cell lines (SCC090 and SCC154) compared to HPV- cells. The expression of SREBF2 and miR-33a are not always to a similar level in HNC cell lines, there is a greater level of miR-33a in UMSCC22B cells and greater SREBF2 in SCC6 and HN29 cell lines. This suggested SREBF2 and miR-33a may not be co-expressed. The elevated expression of both SREBF2 and miR-33a were statistically significant when cell lines were grouped together according to HPV16 status (Figure 5.5 B and C). This further supports the notion that HPV16 may influence the expression of SREBF2 and miR-33a in Head and Neck cancer cells.

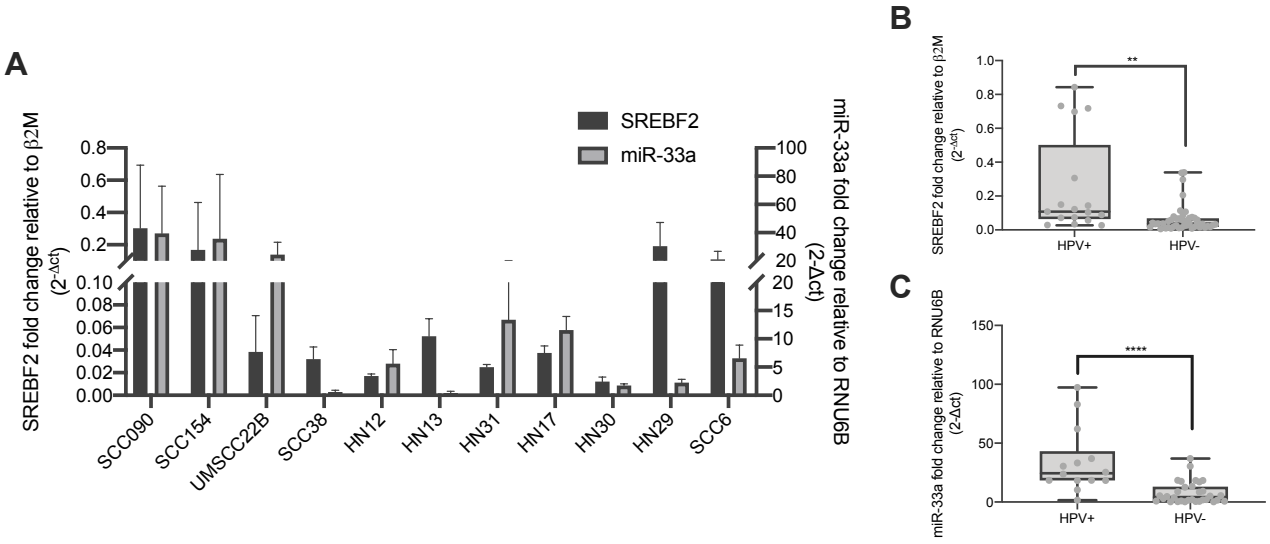


Figure 5. 5: SREBF2 and miR-33a levels in HPV16+ Head and Neck cancer cell lines. A) SREBF2 (black bars) and miR-33a (grey bar) are overexpressed in the HPV16+ Head and Neck cancer cell lines (SCC090, SCC154) compared to the HPV16- Head and Neck cancer cell lines (UMSCC22B, SCC38, HN12, HN13, HN31, HN17, HN30, HN29 and SCC6). B) Box whisker plot representing SREBF2 relative expression from the same cell lines C) B) Box whisker plot representing miR-33a relative expression from the same cell lines Normalisation of gene expression was achieved using the calibrator β2M and miRNA expression using RNU6B and fold relative to B2M and RNU6B using the $2^{-\Delta CT}$ calculation. Statistical analysis was performed using a two-tailed student T-test **p<0.005, ****p<0.0001.

5.3.4. Decrease in SREBF2 expression does not always lead to an alteration in miR-33a expression in Head and Neck cancer cell lines

There appears to be a concomitant relationship between SREBF2, and its intron encoded miR-33a in our HPV+ cell lines. Previously it has been demonstrated that both are overexpressed in response to sterol depletion [275]. Due to these observations, we reasoned that miR-33a expression must be linked to SREBF2 levels. Using the same cells, we investigate if depletion of SREBF2 would affect the expression of miR-33a. In SCC154 and SCC090, we transfected a synthetic si-SREBF2 and measured RNA expression at 24hrs and 48hrs post-transfection.

At 24hrs post-transfection (Appendix Figure 5.2), we observed 20-30% knockdown of SREBF2 in the SCC154 and a 40-50% knockdown in the SCC090 cells, with no significant change in miR-33a. In contrast, the silencing of SREBF2 was greater at 48 hrs in both cell lines (Figure 5.6). In the SCC154, transfection of si-SREBF2 #1 led to a 60% decrease, and si-SREBF2 #2 showed a 40% decrease. When used in combination, we noted a 50% decrease in SREBF2 RNA expression relative to the control. In the SCC090 (Figure 5.6C), all siRNA knockdowns led to a 90% decrease in SREBF2 RNA relative to the control cells. The third siRNA (si-SREBF2 #3) did not result in a reduction in SREBF2 expression levels and both 24 and 48hrs in either cell line; this data is not shown. Overall, we did not see a corresponding decrease in miR-33a levels except for one specific transfection using si-SREBF2 #1 in the SCC154 cells. Collectively these results indicate that depletion of SREBF2 by siRNAs does not affect the levels of miR-33a expression.

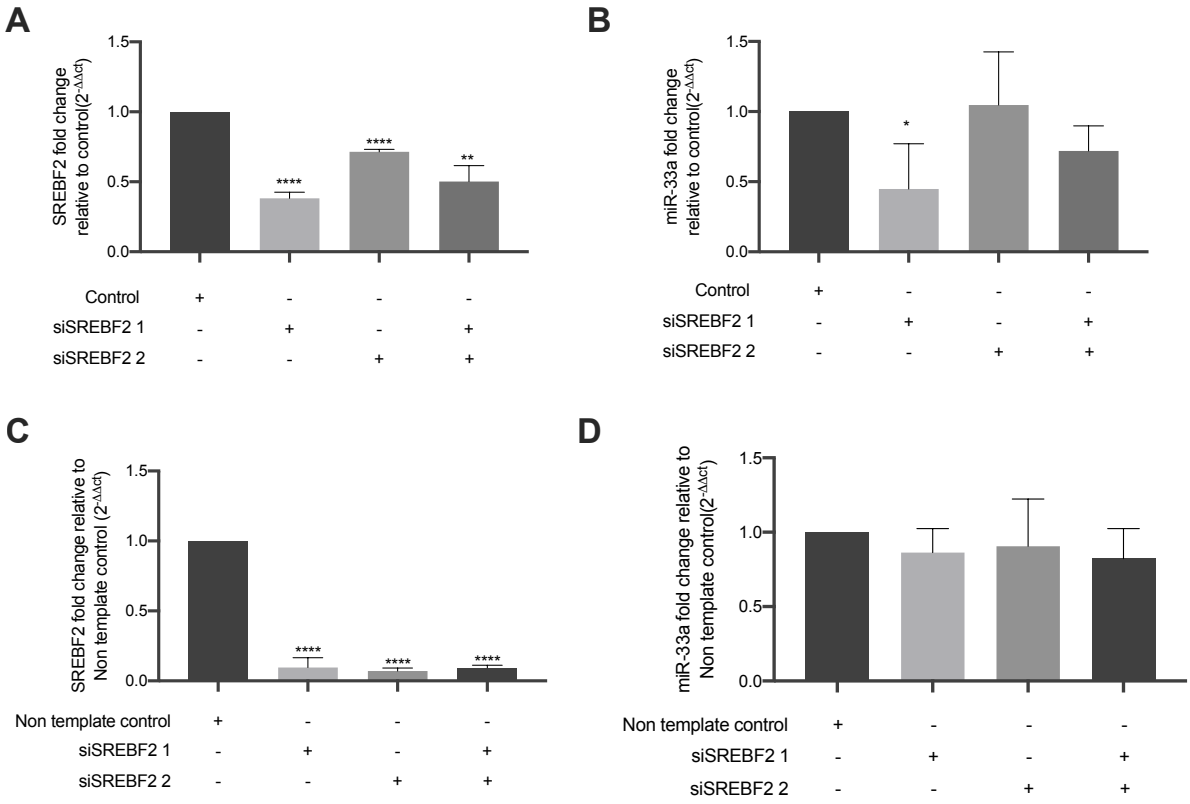


Figure 5. 6: Knockdown of SREBF2 using siRNA in the HPV16+ Head and Neck cancer cell lines SCC154 and SCC090 at 48 hours post-transfection. Two siRNAs (1, 2) were used in combination. A) SREBF2 expression in the SCC154 B) miR-33a expression in the siSREBF2 SCC154 cells C) SREBF2 in the SCC090 D) miR-33a expression in the siSREBF2 SCC090 cells. The normalisation of gene expression was achieved using the calibrator β2M and miRNA expression using RNU6B and fold relative to the non-template control (NC) using the $2^{-\Delta\Delta CT}$ calculation. Statistical analysis was performed using a two-tailed student T-test ***p<0.0005, ****p<0.0001.

5.3.5. HPV16 major viral oncogenes E6 and E7 can influence the expression of SREBF2 and miR-33

We then decided to investigate if the HPV16 viral oncogenes E6 and E7 could influence the overall levels of SREBF2 RNA, thereby affecting the expression of miR-33a. The two cell lines, SCC154 and SCC090, were transfected with siRNAs against E6 and E7 and harvested at 24hr and 48hr post-transfection (Figure 5.7).

In the SCC154 cells, at 48hr post-transfection, both E6 and E7 were knockdown by 60%. This reduction in HPV16 E6/E7 led to the decrease in both SREBF2 and miR-33a RNA expression. In the other cell line, SCC090, both E6 and E7 were knockdown by 90% with either siRNA, resulting in a decrease for both SREBF2 and miR-33a expression. These results are quite different from the previous SREBF2 transfections, and we now believe that the regulation of the nascent SREBF2 RNA is required to alter any miR-33a levels. The silencing of E6 and E7 had directly decreased the levels of nascent SREBF2, thus affecting miR-33a levels overall. This indicated that the HPV16 viral oncogenes E6 and E7 can influence the expression of both SREBF2 and miR-33a in HNC cells.

We also harvested RNA at 24hrs (Appendix figure 5.3) and observed only a minor decrease in SREBF2 and miR-33a expression with a decrease in HPV16 E6/E7 expression, but this was not a significant change.

To further demonstrate this observation, another two siRNAs were designed to target HPV16 E6/E7 (Appendix Figure 5.4) and then transfected into the SCC154 cell line. In these cells, we saw a similar trend, whereby a decrease in HPV16 E6/E7 led to a significant decrease in SREBF2 and miR-33a, particularly 48hrs post-transfection. This further supports the notion that HPV16 E6 and E7 may regulate SREBF2 and miR-33a in HNC cells.

These observations were made in a depleted HPV16 E6 and E7 background. To further confirm this regulatory pathway, we decided to overexpress E6 and E7 and then measure its impact on SREBF2. In the HPV16- cell line, SCC4, plasmids harbouring HPV16 E6 and E7 alone as well as E6/E7 in combination were transiently

transfected and RNA harvest 48hr post-transfection (Figure 5.8). Optimisation of plasmid transfections was conducted using a plasmid harbouring GFP and conducting FACS. This determined 500ng of plasmids would give a 48% transfection rate (Figure 5.8A). The viral oncogenes HPV16 E6/E7 were successfully overexpressed individually and together compared to control cells. In the cells overexpressing HPV16 E6/E7, SREBF2 was also significantly upregulated compared to the control cells. This further supports the idea that HPV16 E6/E7 can influence the expression of SREBF2 in HNC cells.

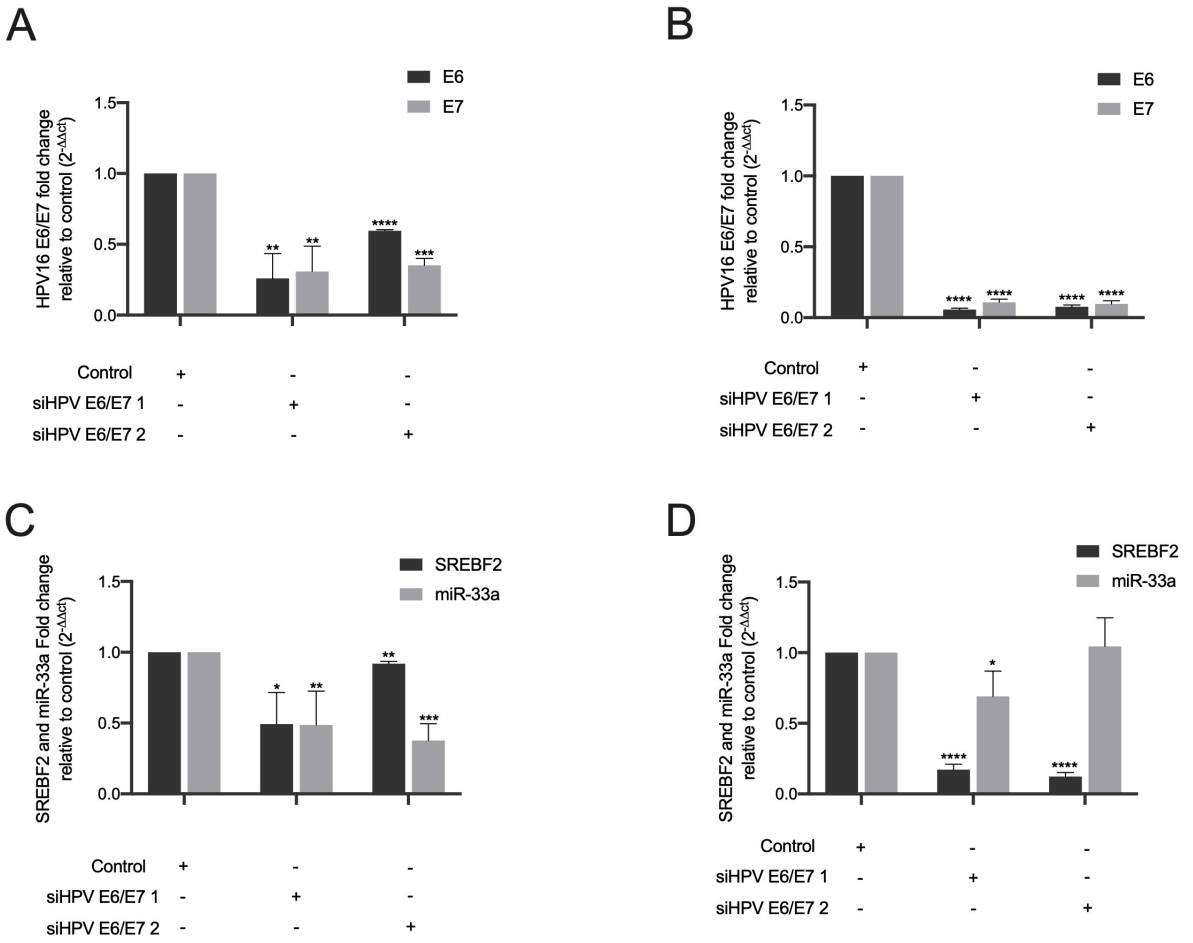


Figure 5. 7: Knockdown of HPV16 E6 and E7 using several siRNA in the HPV16+ Head and Neck cancer cell lines SCC154 and SCC090 at 48 hours post-transfection. A) E6 and E7 were knockdown in the SCC154 cell line B) E6 and E7 were knockdown in the SCC090 cell line C) SREBF2 and miR-33a relative expression in the siHPV16 E6/E7 SCC154 cells D) SREBF2 and miR-33a relative expression in the siHPV16 E6/E7 SCC090 cells. The normalisation of gene expression was achieved using the calibrator β2M and miRNA expression using RNU6B and fold relative to the non-template control (NC) using the $2^{-\Delta\Delta CT}$ calculation. Statistical analysis was performed using two-tailed student T-test *p <0.05, **p<0.005, ***p<0.0005, ****p<0.0001.

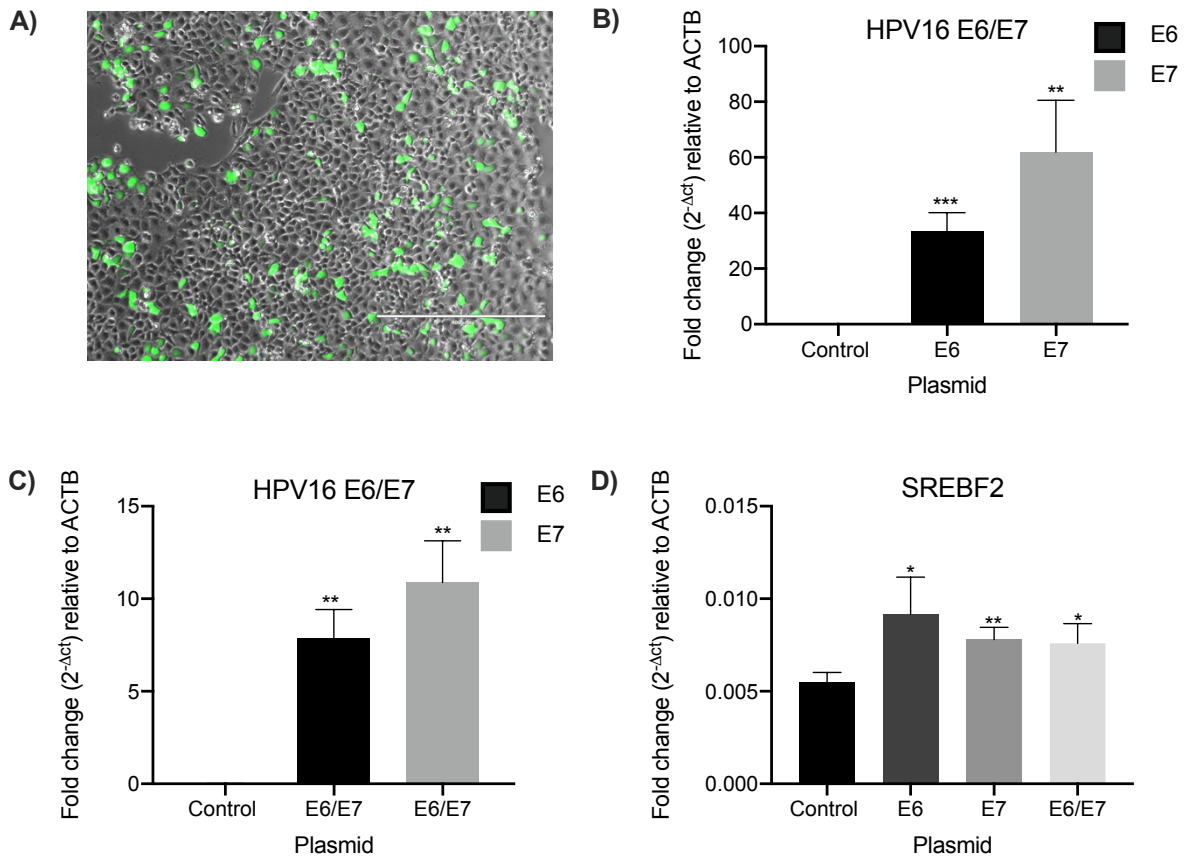


Figure 5. 8: The overexpression of HPV16 E6 and E7 led to an increase in the expression levels of SREBF2 at 48 hours post-transfection in the Head and Neck cancer cell line SCC4. A) SCC4 cells transfected with 500ng of a plasmid harbouring GFP and assessed by FACS determined a 48% transfection efficiency. Validation of HPV16 B) E6 and E7 overexpression individually and C) E6/E7 in combination, D) SREBF2 expression was significantly increased in E6, E7 and E6/E7 expressing cells compared to control cells. The normalisation of gene expression was achieved using the calibrator β2M to calculate relative fold chance using the $2^{-\Delta CT}$ calculation. Statistical analysis was performed using a two-tailed student T-test *p<0.05, **p<0.005.

5.3.6. The regulation of E2F2 by HPV16 E6/E7

Next, we wanted to determine if HPV16 could also alter the expression of other target genes of our proposed pathway, such as the transcription factor E2F2. We previously demonstrated that E2F2 is a direct target of miR-496 and that miR-496 expression can be altered by HPV16 E6 [101] (Chapter 4). In the same transfected SCC154 and SCC090 cells, we noted a decrease in E2F2 expression (Figure 5.9).

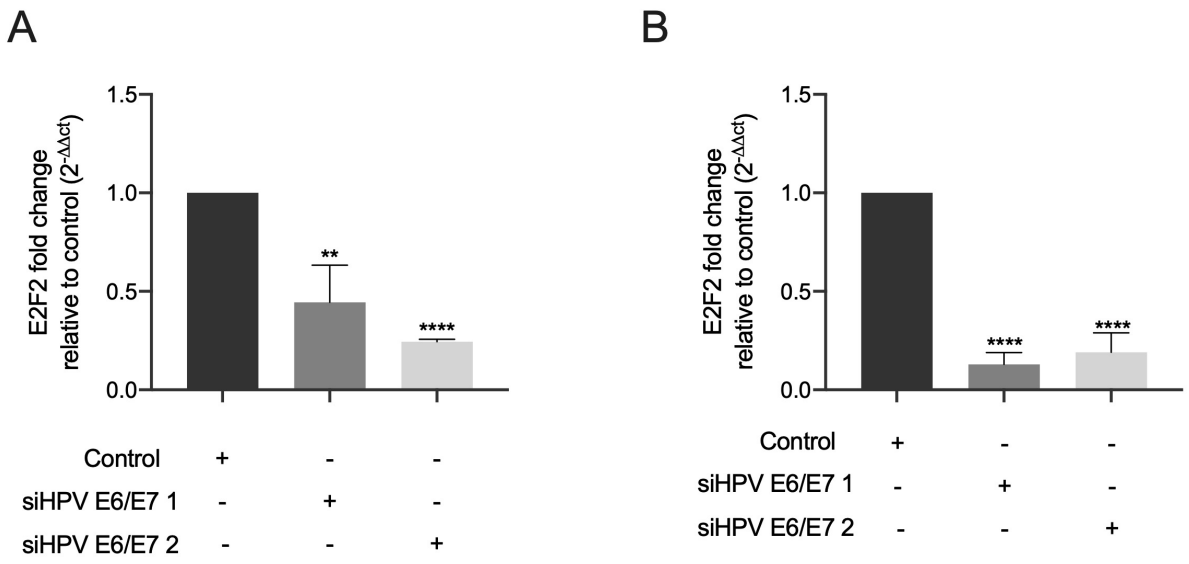


Figure 5. 9: The knockdown of HPV16 E6 and E7 led to a decrease in the expression levels of E2F2 at 48 hours post-transfection, in the Head and Neck cancer cell lines A) SCC154 and B) SCC090. The normalisation of gene expression was achieved using the calibrator β2M and fold relative to the non-template control (NC) using the $2^{-\Delta\Delta CT}$ calculation. Statistical analysis was performed using two-tailed student T-test **p<0.005, ***p<0.0005, ****p<0.0001.

5.3.7. There is a positive correlation between SREBF2 and miR-33a across different cancer types

It has previously been described that there is a correlation between SREBF2 and miR-33a expression in various tissue types [276]. To expand on this observation, we sought to understand if this correlation was common across other different cancer types. Using the TCGA, we examine the expression of both SREBF2 and miR-33a. Table 5.5 shows all cancers with a moderate correlation, and Appendix Table 5.2 describes all other cancers with a weak correlation. In all the cancer types described in Table 5.5, there is a positive correlation between SREBF2 and miR-33a. That is, when SREBF2 is elevated, there is a strong association with high miR-33a levels. This observation may suggest that miR-33a levels are linked with the expression of SREBF2.

Table 5. 5: Correlation between SREBF2 and miR-33a expression in different cancer types from the TCGA.

Cancer type	r	p-value	Interpretation
Cholangiocarcinoma	0.41	1.20E-02	moderate
Thymoma	0.43	1.00E-06	moderate
Mesothelioma	0.45	1.20E-05	moderate
Oesophageal carcinoma	0.46	5.70E-11	moderate
Adrenocortical carcinoma	0.47	1.60E-05	moderate
Lung squamous cell carcinoma	0.55	<2.2e-16	moderate
Pheochromocytoma and paraganglioma	0.57	<2.2E-16	moderate
Uterine carcinosarcoma	0.58	3.20E-06	moderate

5.4. Discussion

5.4.1. Co-expression of SREBF2 and miR-33a in cancer cells

The Human papillomavirus is considered a major risk factor for Head and Neck cancers, in particular those of the oropharyngeal region. HPV16 and its major viral oncogenes have been shown to influence the expression of its host genes as well as non-coding RNAs such as miRNAs in other cancers. Previously our lab aimed to uncover the impact of HPV16 on miRNAs in HNC and discovered a novel pathway via an interactome study [101]. This interactome revealed a potential pathway in which HPV16 E6 and E7 could indirectly influence the expression of the gene SREBF2, leading to its upregulation. This may cause the upregulation of the miRNA, miR-33a, which is encoded within this gene. The focus of this study was to uncover the relationships between the HPV16, the gene SREBF2 and its encoded miR-33a in Head and Neck cancers.

The RNA sequencing data from The Cancer Genome Atlas allowed us to compare the expression profile of specific genes and miRNAs between HPV16+ and HPV- oral and oropharyngeal cancers. These results demonstrated that SREBF2 was upregulated in the HPV16+ cancers, along with miR-33a. This miRNA is located within intron 16 of the human SREBF2 gene, and it is highly conserved across numerous animal species [277]. We also noted that miR-496 was downregulated in the HPV16+ samples, and its target, E2F2, no longer under miRNA control, was upregulated in the positive HPV samples.

miR-33a overexpression due to HPV infection has previously been described [108] but SREBF2 overexpression in HPV+ HNC is unique to this study. The results from TCGA further support miR-33a, miR-496 and E2F2 expression profiles we observed in our own patient cohort [101]. Correlation analysis of SREBF2 and miR-33a across different cancer types from the TCGA also revealed that SREBF2 and miR-33a typically show a positive correlation, further supporting the assumption that they are co-expressed in cancer.

We also investigated the expression profiles of these genes and miRNAs comparing HPV16+, HPV18+ and HPV- cervical cancer tumours. There was no change in

expression for miR-496, SREBF2 and E2F2, indicating their altered expression in specific for HPV infected HNC. When comparing the different virus types, miR-33a had an increased expression in the HPV16+ cervical cancer but not the HPV18+ cervical cancers. This demonstrates that only HPV16 can influence its expression and can do so across various cancer types. HPV16 is the dominant HPV variant in Head and Neck cancers [4, 9]. Studies have also previously shown that the miRNA expression profile between HPV+ cervical and Head and Neck cancers is vastly different, and the two cancer types show a unique expression pattern, although there are some core miRNAs the virus targets [269].

In the cell panel, when HPV16 was present, we noted an increased expression for both SREBF2 and miR-33a. SREBF2 overexpression in HPV+ HNC has not previously been characterised. Although SREBF2 has been shown to be upregulated in oesophageal cancers, this is a very different type of HNC anatomically [278, 279] and this was not due to HPV infections. The changes we see *in vitro* in cell lines is in agreement with the observations in the clinical samples from the TCGA and our previous patient cohort [101]. From these independent observations, it appears that HPV16 can promote the expression of SREBF2 in HNC.

5.4.2. Depletion of SREBF2 may not affect miR-33a levels

To further understand the relationship between SREBF2 and miR-33a in HPV16+ HNC, we conducted a siRNA knockdown of SREBF2 in two HPV16+ HNC cell lines (SCC090 and SCC154). A decrease in SREBF2 expression led to a decrease in miR-33a expression of only one specific siRNA (siSREBF2 1) in SCC154.

We noted that there was no significant change in miR-33a expression for other siRNA, there is a downward trend when both siRNA are used in combination, but this is not statistically significant. The decrease in SREBF2 expression did not lead to any significant change in miR-33a expression in the SCC090. Together, these results may indicate that miR-33a expression is not dependent on SREBF2 expression in HPV16+ HNC, and these two genes may not co-transcribed as previously assumed.

There is strong evidence that miR-33a and SREBF2 are concomitantly expressed in normal tissue [276] as well as in certain cancer types in response to cholesterol levels. In colorectal cancers, when cells were treated with high cholesterol, both miR-33a and SREBF2 expression were reduced [280] and glioblastoma miR-33a and SREBF2 are both elevated [281]. However, in prostate cancers, SREBF2 and miR-33a had differing expression levels [282, 283]. SREBF2 had an increased level of expression, although miR-33a had no such change or, even in some cases, a decrease in expression. This expression pattern was also seen in liver cells treated with various polyphenols, miR-33a expression was significantly decreased, but there was no change in SREBF2 [284]. This implies that there is some form of post-transcriptional regulation of miR-33a expression in these cancers.

The majority of the studies describing SREBF2 and miR-33a co-expression [276, 277, 280] demonstrate that this is induced by cholesterol levels. Our results suggest that there is some form of mechanism that protects miR-33a in HPV16+ HNC cells. Kim and Kim [285] described a phenomenon whereby intronic miRNAs can be spliced by Drosha from unspliced intronic regions before splicing catalysis. They concluded that splicing is not required for intronic miRNA processing, and unspliced introns may serve as a substrate for the microprocessor complex. SREBF2 and miR-33a were used as an example in this study [285] as well as several intronic miRNAs such as miR-126 being processed by Drosha from the EGFL7 gene prior to the miRNA-encoding intron being spliced out [285]. Perhaps this is the case in our study; miR-33a is being processed by Drosha before our siRNA have a chance to act in HNC cells.

SREBF2 and miR-33a co-expression pattern has been assumed, given there is a strong positive correlation in their expression across different tissue types [276, 277, 280]. There is increasing evidence that not all intronic miRNAs are co-transcribed with their host genes [286-288]. It has been found that over 1/3 of intronic miRNAs may possess their own independent promoter sequences [289, 290]. For example, miR-26b was found to have an independent transcription promoter sequence and transcriptional factor from its host gene CTDSP [286]. The expression of this miRNA and host gene were negatively correlated. Another example is, in the intron of insulin-like growth factor 2 (Igf2), miR-483-2p is embedded [291]. This miRNA is co-expressed in some hepatocellular carcinomas cells (Hepa1-6) [292], but not others

(HUVECs) [293]; these differences are potentially due to differing promoter regions [291]. This mechanism could be explored further to determine how miR-33a and SREBF2 are expressed; miR-33a may possess its own promoter sequences, separate from SREBF2.

5.4.3. E6 and E7 oncogenes can potentially regulate SREBF2/miR-33a expression

Next, we wanted to determine what the impact of the two major viral oncogenes E6 and E7 are on SREBF2 and miR-33a. We focused on the oncogenes E6 and E7 because they are consistently expressed in cancer cells [294]. E6 and E7 share the same promoter and expressed of the same mRNA; hence a single siRNA can reduce the expression of both viral oncogenes.

Two siRNA targeting HPV16 E6/E7 for its knockdown were used in this study. In the SCC154, both E6 and E7 were knockdown by 50-60% and in the SCC090s, 90%. When we decreased E6 and E7 expression, this led to the decrease in both SREBF2 and miR-33a expression. This indicates that the HPV16 viral oncogenes E6 and E7 can influence the expression of SREBF2 and miR-33a in HNC.

We further demonstrated this interaction using an HPV16 E6/E7 plasmid overexpression system. Not only could the overexpression of E6/E7 together lead to an upregulation in SREBF2, but the overexpression of E6 and E7 individually also increased SREBF2 expression in HNC cells. This data also further supports our notion that miR-33a expression can be altered by HPV16 E6 overexpression [101], but the nature of this targeting is yet to be determined. SREBF2 and miR-33a are in a negative feedback loop dependant on sterol levels. Perhaps this is a mechanism in which HPV16 targets in HNCs to “feed” cancer cells and promote cellular growth.

The viral miRNA interactome can provide further mechanistic insights. Looking at other genes from our interactome, we observed that the knockdown of E6 and E7 also led to a decrease in E2F2. This relationship has previously been described in HPV+ cervical cancers, where E2F2 is overexpressed with HPV infection, possibly due to its relationship with the cell cycle [295]. We believe E2F2 is also under the control of miR-496, a miRNA that is downregulated by HPV16. When we knockout E6/E7, miR-496

is now expressed and can regulate its target E2F2 resulting in a decrease in E2F2 expression. Other transcription factors from our proposed model also had altered expression due to HPV16, evident in the TCGA data.

The transcription factors from the model that miR-33a regulate and that might promote miR-496 expression, GATA2 and BACH1, are both downregulated in the HPV+ tumours (Appendix Figure 5.5). If they are downregulated by HPV16, then they cannot promote the expression of miR-496, and hence its targets (E2F2) are overexpressed. This highlights the complexity of molecular networks in cancers.

From our interactome, there were many transcription factors upstream of SREBF2, which are direct targets of E6 and E7; these include MYC and EP300, which are targeted by both oncogenes. Mining the TCGA data to compare HPV+ and HPV- HNC revealed that neither of these transcription factors would promote the transcription of SREBF2 as there was a decrease in MYC and no change for EP300 in the HPV+ tumours (Appendix Figure 5.5). There may be some other regulatory factor between HPV16 and SREBF2.

5.4.4. Significance and Future directions

SREBF2 and miR-33a altered expression has been described to be associated with numerous cancer pathways such as cell proliferation, migration and invasion of various cancer types [296-298]. This association has not been reported for HNC. It has previously been described that HPV and its viral oncogenes manipulate metabolic pathways to promote a neoplastic phenotype, and the metabolic activity of a cell has been described to be fundamental to the viruses' oncogenicity [299].

Our knockdown of SREBF2 and HPV16 E6/E7 were assessed at an RNA level. To determine if our knockdowns are functionally active, the levels of SREBF2 and HPV16 E6/E7 need to be assessed at a protein level. Due to delays caused by COVID-19 and time constraints; this could not be conducted but is a future consideration. Currently, the protein analysis via western blots for this project has been started, optimisation for SREBF2 and Actin as an internal control determined that 20ug of protein is necessary

for protein detection (Appendix figure 5.6). In future research, SREBF2, HPV16 E6 and E7 will be measured at a protein level in our knockdown systems.

It would also be beneficial to measure the protein levels of downstream targets of these genes to determine if this knockdown has an impact on the cellular pathways, they regulate The targets we suggest are HMG-CoA reductase (HMGCR), HMG-CoA synthase (HMGCS) and LDL receptor (LDLR) for SREBF2 and p53 and pRb for HPV16 E6 and E7, respectively. This study would also benefit from RNA pull-down assay of SREBF2 to determine if E6 or E7 directly bind to SREBF2 or if this is an indirect interaction via transcription factors as predicted in our interactome.

The increase in activity of PI3K/Akt/mTOR pathways has been linked to an increase in SREBF2s expression to promote lipid biosynthesis in other cancers [300]. This pathway is altered by several of the HPV early-stage genes (E6, E7 and E5) [299].
In this study, we have provided evidence that HPV16 E6 and E7 can lead to the upregulation of SREBF2 as well as its encoded miRNA, miR-33a. Whether this upregulation is directly or indirectly due to upstream pathways like the PI3K/Akt/mTOR is yet to be determined.

With the upregulation of SREBF2 and miR-33a, it is expected that cholesterol and lipid levels will be much higher in these cells. This has been shown in cervical cancers that HPV+ dysplastic cells had a higher lipid content compared to HPV- and normal cells [301]. Increased synthesis of lipids in cancer cells has been shown to stimulate cell proliferation and block apoptosis [302, 303]. It may be plausible that the upregulation of SREBF2 and miR-33a by HPV16 E6/E7 can lead to an increase in lipid levels within HNC cancer cells to promote cell proliferation, migration and invasion.

Under sterol depletion, the co-expression of SREBF2 and miR-33a is said to promote the synthesis of cholesterol and decreases the efflux of cholesterol out of a cell. This mRNA and its embedded miRNA work cooperatively to boost cellular cholesterol levels. It has always been assumed that SREBF2 and miR-33a are co-expressed due to a positive correlation in their expression patterns in various tissue types. However, there are a few reports that this is not the case for SREBF2 and miR-33a [282-284].

Currently, the exploration of the impact of HPV16 E6/E7 on intronic miRNAs is very limited. Only one study has previously described the interaction between HPV16 E6 with an intronic miR-23b and its host gene C9orf3. It was demonstrated that HPV16 E6 downregulated miR-23b via the hypermethylation of the promoter of its host gene C9orf3 in cervical cancers. This, in turn, promoted the expression of c-MET and prevented apoptosis of cervical cancer cells [304].

This study has provided another example of HPV16 E6/E7 regulation of an intronic miRNA. This type of co-expression adds another level of complexity when studying the impact and mode of action behind HPV16 effects on miRNAs. Not only do we need to consider the downstream effects of altered miRNA expression on their targets, but we also need to take into consideration the genomic location of these miRNAs and the severe consequences this might have on a cell.

Introns host around 60% of all known miRNAs, and these miRNAs provide functional synergistic or antagonistic roles with their host gene [305]. It is estimated that 35-50% of intronic miRNAs can be transcribed by independent promoters [306, 307] and older or more conserved intronic miRNAs are more likely to be transcribed by an independent promoter [307]. The mechanisms driving HPV16s influence on intronic miRNAs have not yet been described. Assessing other DE-miRNAs (UCSC genome browser) from our LNA array (Chapter 4), 54% of these are also located within the introns of either protein-coding genes (miR-33a/SREBF2, miR-342/EVL, miR-505/ATP11C, miR-362/CLCN5 and miR-504/FGF13), or long non-coding RNAs (miR-210/MIR210HG, miR-142-3p/TSPOAP1-AS1).

Even several of our DE-miRNAs from the interactome analysis in our literature review (Chapter 1) were identified to be in intronic regions of lncRNAs (miR181b/MIR181A1, miR-31/LOC554202 and miR-9/MIR9HG). This clearly demonstrates that HPV16 may have an impact on intronically located miRNAs. It would be interesting to understand how HPV16 may affect these miRNAs and the implications of this in cancer.

One example of HPV16 regulating an intronic miRNAs is its targeting of miR-139-3p via methylation of its host genes promoter. In both HPV16+ cervical cancer and HNC patient samples, miR-139-3p is downregulated.

This downregulation is reported to be due to increased DNA methylation of miR-1399-3p host gene PDE2A promoter [308] and methylation is induced by HPV16. The regulation of miRNAs via epigenetic regulation such as methylation has been shown to be a mode of action by HPV to induce altered miRNA expression [309]. In the case of intronic miRNAs, this methylation can occur in their own miRNA promoter sequences or within their host gene promoter.

HPV16 can also interact with transcription factors to promote their binding and regulation of miRNAs. The E7 oncogene degrades the tumour suppressor protein, pRB, freeing the transcription factors E2F. Many miRNAs contain E2F binding sites within their promoter regions. E2F can transactivate the expression of miRNAs including miR15b/16-2, miR-17-92, miR-106b-25, let-7a-d and let-7i [310-312] . E2F can also induce the expression of an intronic miRNAs host gene and thus the miRNA such as MCM7 and its intronic miR-106b/93/25 [313]. These two modes of action suggest a possible mechanism for our miR-33a regulation by HPV16.

It is suggested from our data that HPV16 E6/E7 can lead to the upregulation or increased transcription of the entire mRNA of SREBF2 containing miR-33a. However, we identified that miR-33a does not necessarily require SREBF2 to be transcribed during HPV16 infection, and the viral oncogenes E6/E7 induces an increase in miR-33a expression alone. MiR-33a is perhaps cleaved out or processed and undergoes maturation under its own form of transcriptional regulation. This highlights the crucial role this miRNA may have in HPV viral tumorigenesis.

In summary, our study has identified that HPV16 can influence the expression of SREBF2 and E2F2 as well as miR-33a. The data we have provided supports the model from our previous interactome analysis (Figure 5.10). This model describes HPV16 E6/E7 indirectly influencing the expression of SREBF2 and miR-33a, leading to the downregulation of transcription factors that no longer promote miR-496 expression, which will lead to an increase in the miRNA target E2F2.

This mechanistic model provides a mode of action behind HPV16 altering miR-33a and miR-496 expression, and it assists in uncovering regulatory targets of HPV and

their involvement in crucial cellular pathways such as cholesterol metabolism and cell cycle pathways.

We demonstrated that SREBF2 and miR-33a are overexpressed in HPV16+ HNC, although miR-33a expression may not be co-expressed with SREBF2 as previously assumed. HPV16 will still drive miR-33a overexpression despite SREBF2 knockdown, perhaps due to cleavage of the unspliced miR-33a by Drosha or miR-33a possessing its own promoter sequence.

We also demonstrated that HPV16 viral oncogenes, specifically E6 and E7, could influence the expression of SREBF2 and miR-33a in HNC. Due to the role SREBF2 and miR-33a play in cholesterol biosynthesis, this suggests that HPV16 alters the metabolic environment of HNCs and may promote tumorigenesis.

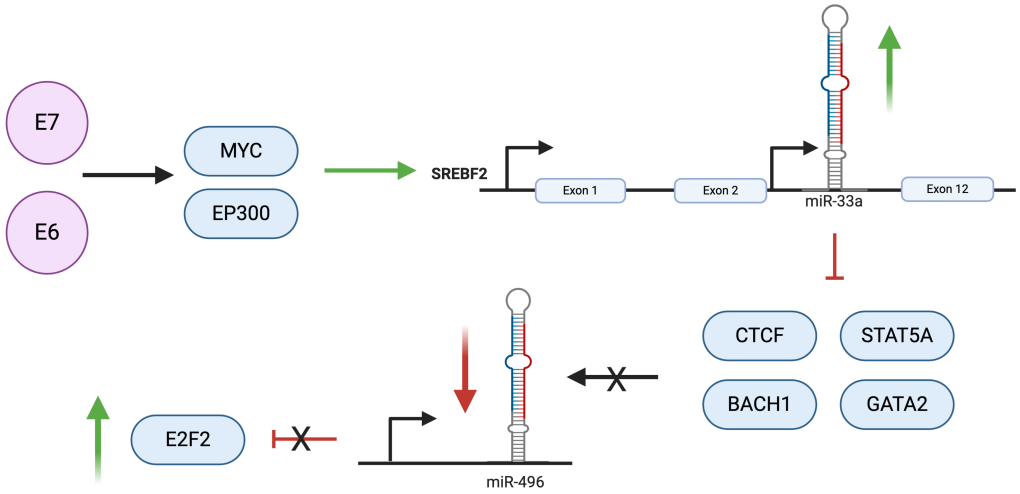


Figure 5. 10: Proposed model for miR-33a and miR-496 dysregulation by HPV16. The HPV16 viral oncogenes directly bind to the transcription factors MYC and EP300. These transcription factors promote the expression of SREBF2; this will lead to the increase of miR-33a expression as it is located within intron 16 of SREBF2. Several transcription factors (CTCF, STAT5A, BACH1 and GATA2) are under the regulation of miR-33a. These transcription factors will no longer activate the transcription of miR-496, and its levels will decrease. This will lead to the overexpression of E2F2 as it is no longer under the direct regulation of miR-496.

5.5. Appendix

This chapter has been written with the intention to be submitted to the journal of *Virology.* Below is the appendix material that will form the supplementary information for the submitted paper.

Appendix Table 5. 1: The cancer gene atlas, Head and Neck cancer cohort: patient barcode IDs and relevant clinical information:

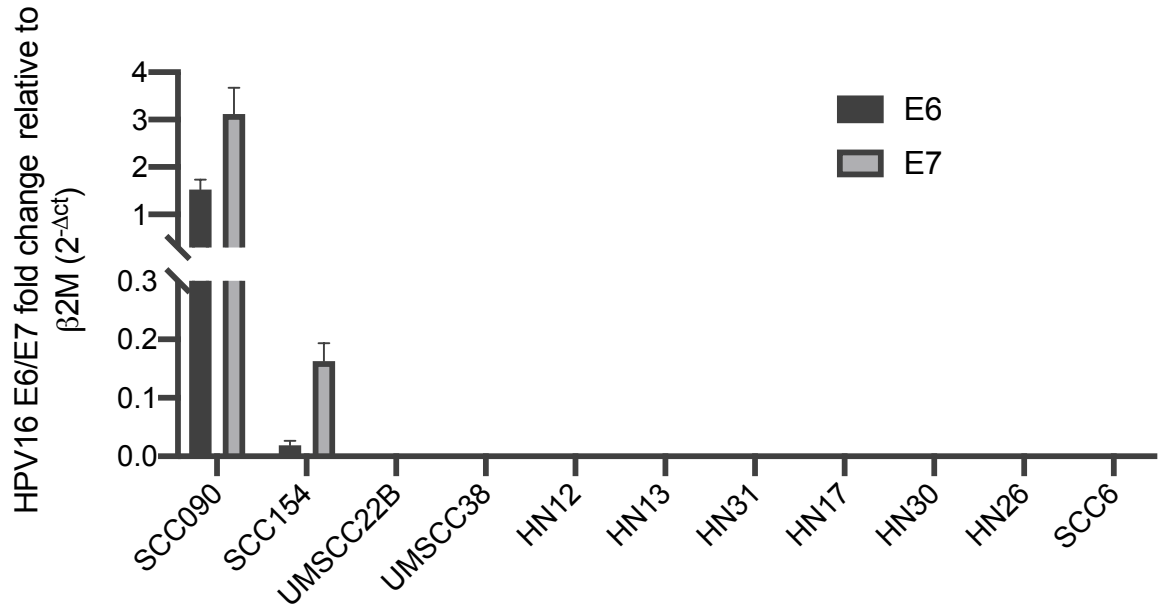
Sample ID	HPV status by ish testing	Clinical stage	Gender	Anatomical neoplasm subdivision
TCGA-CN-A6V6-01	Positive	Stage IVA	MALE	Base of tongue
TCGA-CN-A6UY-01	Positive	Stage IVA	MALE	Base of tongue
TCGA-BB-7861-01	Positive	Stage III	MALE	Base of tongue
TCGA-CR-5250-01	Positive	Stage II	MALE	Base of tongue
TCGA-HL-7533-01	Positive	Stage IVA	MALE	Oral Cavity
TCGA-BB-A6UM-01	Positive	Stage III	MALE	Tonsil
TCGA-CN-A49C-01	Positive	Stage IVA	MALE	Tonsil
TCGA-CR-7404-01	Positive	Stage IVA	MALE	Tonsil
TCGA-CN-A6V7-01	Positive	Stage IVA	MALE	Tonsil
TCGA-BA-A4IH-01	Positive	Stage IVA	MALE	Tonsil
TCGA-QK-A6IF-01	Positive	Stage IVA	MALE	Tonsil
TCGA-CR-6480-01	Positive	Stage IVA	MALE	Tonsil
TCGA-CR-6481-01	Positive	Stage IVA	MALE	Tonsil
TCGA-T2-A6X0-01	Positive	Stage IVA	MALE	Tonsil
TCGA-CN-5374-01	Positive	Stage IVA	FEMALE	Tonsil
TCGA-BB-7866-01	Positive	Stage III	MALE	Tonsil
TCGA-CR-6487-01	Positive	Stage II	MALE	Tonsil
TCGA-QK-A6V9-01	Positive	Stage II	MALE	Tonsil
TCGA-TN-A7HI-01	Positive	Stage I	MALE	Tonsil
TCGA-CN-A63Y-01	Positive	Stage I	MALE	Tonsil
TCGA-CN-A499-01	Positive	Stage I	FEMALE	Tonsil
TCGA-CR-6477-01	Negative	Stage IVA	FEMALE	Base of tongue
TCGA-CR-6491-01	Negative	Stage IVA	MALE	Floor of mouth
TCGA-BA-A6DD-01	Negative	Stage IVA	MALE	Floor of mouth
TCGA-BA-A6D8-01	Negative	Stage IVA	MALE	Floor of mouth

TCGA-BA-A6DF-01	Negative	Stage IVA	FEMALE	Floor of mouth
TCGA-CN-6016-01	Negative	Stage IVA	MALE	Floor of mouth
TCGA-CN-5359-01	Negative	Stage IVA	MALE	Floor of mouth
TCGA-CN-5358-01	Negative	Stage III	MALE	Floor of mouth
TCGA-QK-A6IJ-01	Negative	Stage II	MALE	Floor of mouth
TCGA-CN-5373-01	Negative	Stage II	FEMALE	Floor of mouth
TCGA-CR-6492-01	Negative	Stage IVA	MALE	Hard Palate
TCGA-CN-6020-01	Negative	Stage IVA	MALE	Oral Cavity
TCGA-CR-7373-01	Negative	Stage IVA	MALE	Oral Cavity
TCGA-CR-7379-01	Negative	Stage IVA	FEMALE	Oral Cavity
TCGA-BB-A5HU-01	Negative	Stage IVA	MALE	Oral Cavity
TCGA-CR-6484-01	Negative	Stage IVA	FEMALE	Oral Cavity
TCGA-CR-7376-01	Negative	Stage II	MALE	Oral Cavity
TCGA-CR-7395-01	Negative	Stage II	FEMALE	Oral Cavity
TCGA-CR-7394-01	Negative	Stage IVA	MALE	Oral Tongue
TCGA-BB-7872-01	Negative	Stage IVA	MALE	Oral Tongue
TCGA-BA-A6DG-01	Negative	Stage IVA	MALE	Oral Tongue
TCGA-T2-A6WZ-01	Negative	Stage IVA	MALE	Oral Tongue
TCGA-CR-7392-01	Negative	Stage IVA	FEMALE	Oral Tongue
TCGA-CN-6024-01	Negative	Stage IVA	MALE	Oral Tongue
TCGA-CN-6998-01	Negative	Stage IVA	MALE	Oral Tongue
TCGA-CN-4742-01	Negative	Stage IVA	FEMALE	Oral Tongue
TCGA-CN-6017-01	Negative	Stage III	MALE	Oral Tongue
TCGA-CN-6019-01	Negative	Stage IVA	MALE	Oral Tongue
TCGA-CR-7390-01	Negative	Stage III	MALE	Oral Tongue
TCGA-BA-A6DE-01	Negative	Stage III	FEMALE	Oral Tongue
TCGA-CN-4737-01	Negative	Stage II	MALE	Oral Tongue
TCGA-CR-7372-01	Negative	Stage II	MALE	Oral Tongue
TCGA-CN-4725-01	Negative	Stage II	MALE	Oral Tongue
TCGA-CN-5370-01	Negative	Stage II	MALE	Oral Tongue
TCGA-BA-A6DB-01	Negative	Stage II	FEMALE	Oral Tongue
TCGA-IQ-A61L-01	Negative	Stage II	FEMALE	Oral Tongue
TCGA-CN-4733-01	Negative	Stage I	MALE	Oral Tongue
TCGA-CR-7393-01	Negative	Stage I	MALE	Oral Tongue
TCGA-CR-7391-01	Negative	Stage I	FEMALE	Oral Tongue
TCGA-BA-A8YP-01	Negative	Stage IVB	MALE	Oropharynx

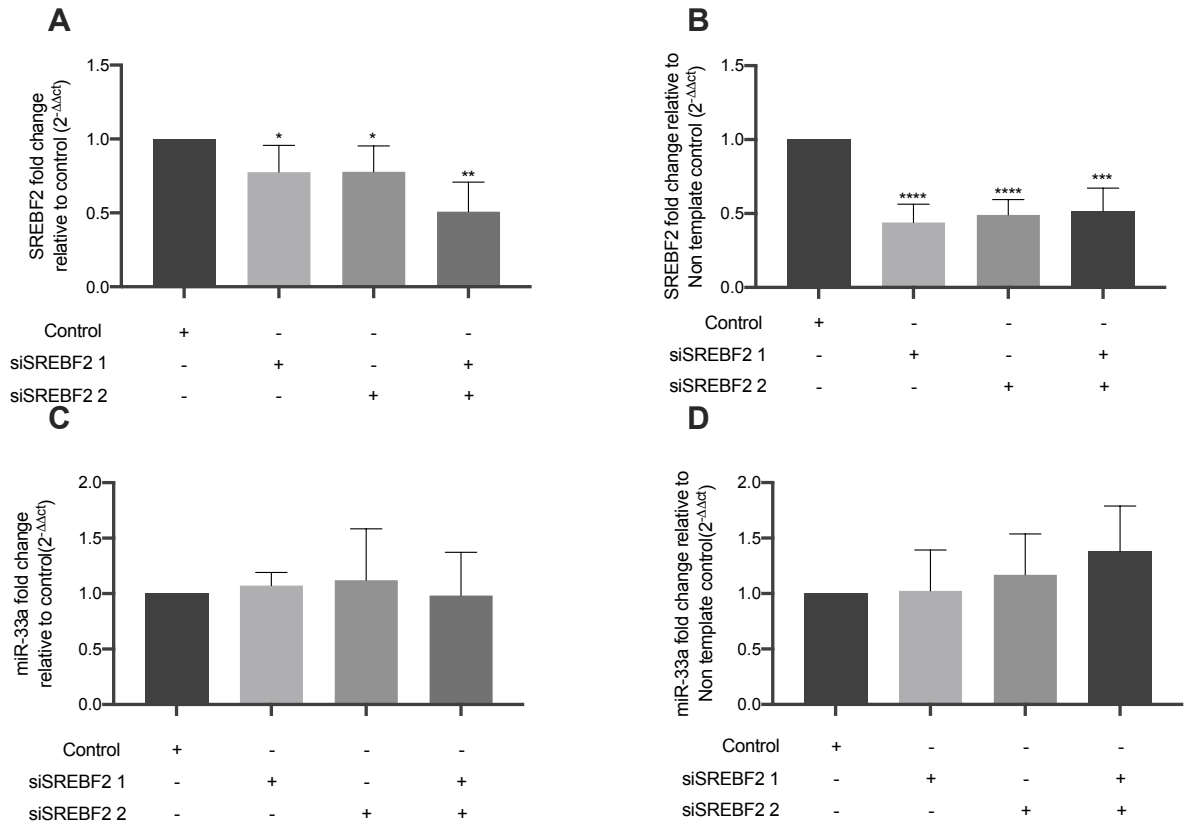
TCGA-BA-A4IF-01	Negative	Stage IVA	MALE	Oropharynx
TCGA-BA-A6DL-01	Negative	Stage III	MALE	Oropharynx
TCGA-MT-A51W-01	Negative	Stage I	FEMALE	Tonsil

Appendix Table 5. 2: Correlation between SREBF2 and miR-33a expression in different cancer types from the TCGA

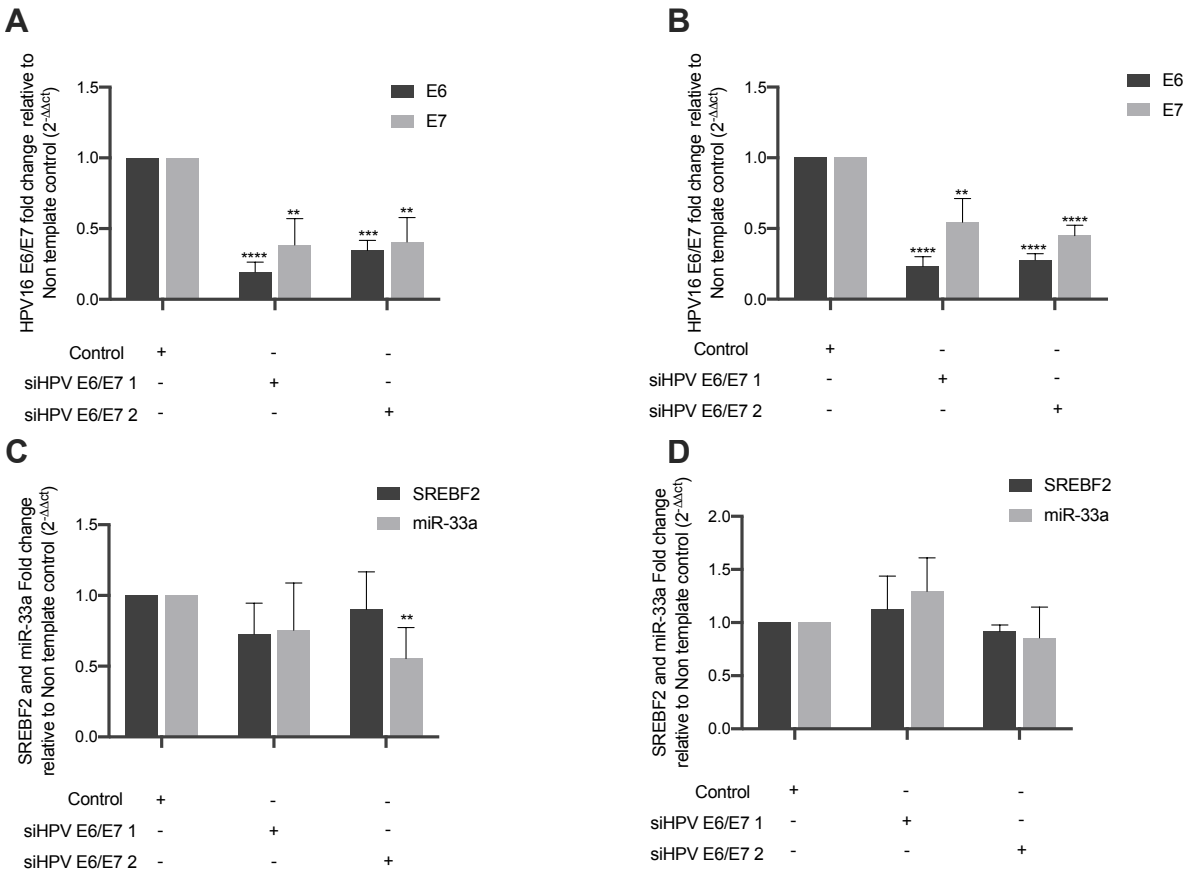
Cancer type	r (tumour)	p-value (tumour)	Interpretation
Prostate Adenocarcinoma	0.13	3.60E-03	very weak
Kidney Renal Clear Cell Carcinoma	0.15	2.30E-02	very weak
Lung Adenocarcinoma	0.19	5.3-E05	very weak
Ovarian Serous Cystadenocarcinoma	0.19	1.60E-01	very weak
Pancreatic Adenocarcinoma	0.19	1.40E-02	very weak
Rectum Adenocarcinoma	0.21	5.20E-02	weak
Lymphoid Neoplasm Diffuse Large B Cell Lymphoma	0.23	1.20E-01	weak
Testicular Germ Cell Tumours	0.24	5.50E-03	weak
Uveal Melanoma	0.28	1.20E-02	weak
Brain Lower-grade Glioma	0.29	1.20E-11	weak
Sarcoma	0.3	7.60E-07	weak
Kidney Renal Papillary Cell Carcinoma	0.31	1.40E-07	weak
Skin Cutaneous melanoma	0.31	2.50E-03	weak
Breast Invasive Carcinoma	0.32	< 2.2e-16	weak
Head and Neck Squamous Cell Carcinoma	0.34	2.30E-14	weak
Stomach Adenocarcinoma	0.34	1.20E-11	weak
Thyroid Carcinoma	0.34	2.10E-15	weak
Colon Adenocarcinoma	0.35	1.40E-08	weak
Bladder Urothelial Carcinoma	0.36	1.30E-13	weak
Kidney Chromophobe	0.37	2.20E-03	weak
Liver Hepatocellular Carcinoma	0.38	8.90E-14	weak
Uterine Corpus Endometrial Carcinoma	0.38	2.30E-15	weak



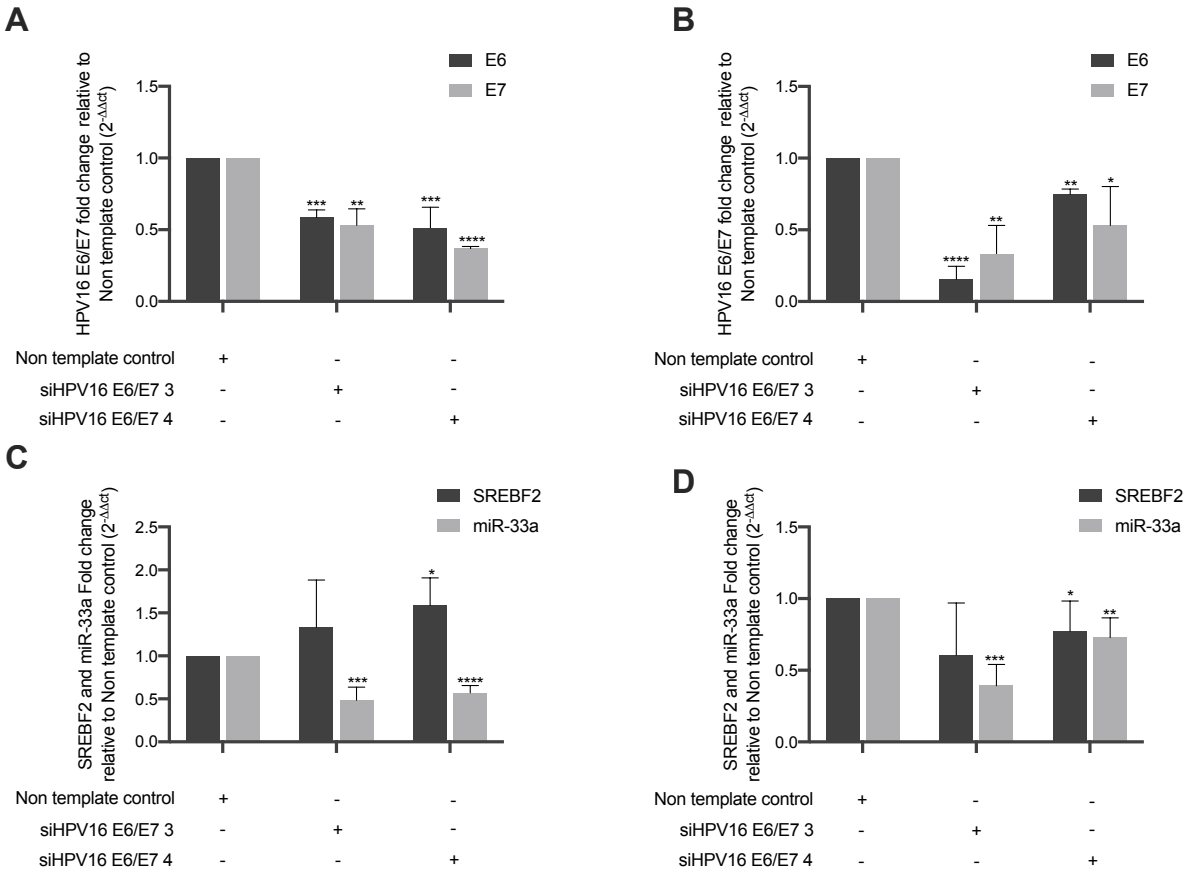
Appendix Figure 5. 1: The HPV16 status of the Head and Neck cancer cell lines panel. The oncogenes E6/E7 were detected in the HPV16+ HNC cell lines SCC090 and SCC154. E6/E7 were not detected in the HPV- cell lines, UMSCC22B, UMSCC38, HN12, Hn13, HN31, HN17, HN30, HN26 and SCC6. The normalisation of gene expression was achieved using fold change relative to the calibrator β2M using the $2^{-\Delta CT}$ calculation.



Appendix Figure 5. 2: Knockdown of SREBF2 using siRNA in HPV16+ HNC cell lines SCC154 and SCC090 at 24hrs post-transfection. A) SREBF2 expression was knockdown to 20-50% in the SCC154 B) SREBF2 expression was knockdown to 50% in the SCC090 C and D) There was no significant decrease in miR-33a expression with SREBF2 knockdown. The normalisation of gene expression was achieved using the calibrator β2M and miRNA expression using RNU6B and fold relative to the non-template control (NC) using the $2^{-\Delta\Delta CT}$ calculation. Statistical analysis was performed using two-tailed student T-test *p<0.05, **p<0.005, ***p<0.0005, ****p<0.0001.

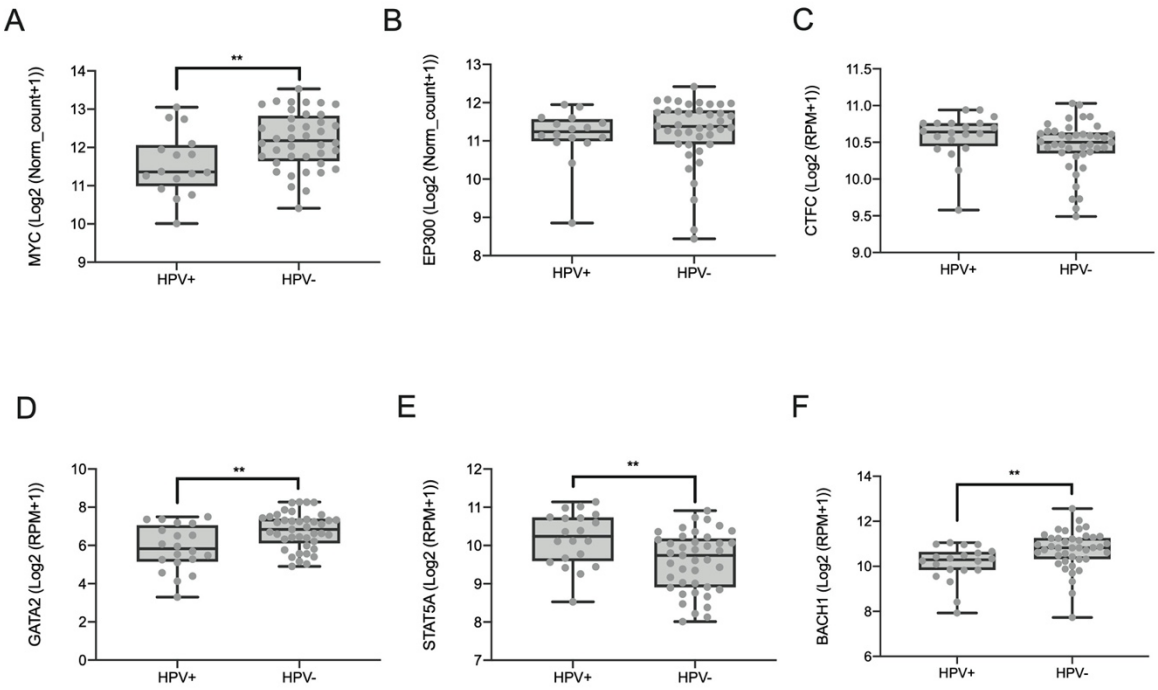


Appendix Figure 5. 3: Knockdown of HPV16 E6/E7 using siRNA in HPV16+ HNC cell lines SCC154 and SCC090 at 24hrs post-transfection. A) HPV16 E6/7 expression was knockdown to 50-80% in the SCC154 B) HPV16 E6/E7 expression was knockdown to 50-80% in the SCC090 C) There was a decrease in miR-33a expression with siHPV16 E6/E7 2 in the SCC154, but no significant with other siRNA and no change in SREBF2 expression D) There was no significant decrease in miR-33a or SREBF2 expression. The normalisation of gene expression was achieved using the calibrator β2M and miRNA expression using RNU6B and fold relative to the non-template control (NC) using the $2^{-\Delta\Delta CT}$ calculation. Statistical analysis was performed using two-tailed student T-test **p<0.005, ***p<0.0005, ****p<0.0001.

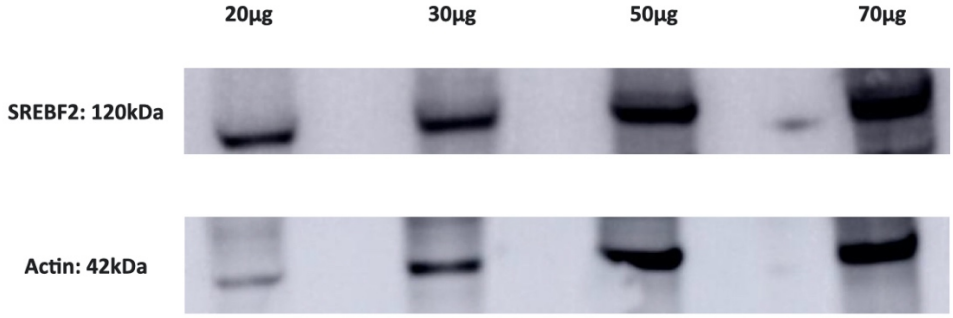


Appendix Figure 5. 4: Knockdown of HPV16 E6/E7 using siRNA in HPV16+ HNC cell lines SCC154 at 24hrs and 48hrs post-transfection. A) HPV16 E6/7 expression was knockdown to 50-60% at 24hrs and B) 20-90% at 48hrs C) There was a decrease in miR-33a expression but no significant decrease in SREBF2 expression, SREBF2 was upregulated by siHPV16E6/E7 4 D) There was a significant decrease in miR-33a expression with both siRNA and a decrease of SREBF2 using siHPV16 E6/E7 4. Normalisation of gene expression was achieved using the calibrator β2M and miRNA expression using RNU6B and fold relative to the non-template control (NC) using the $2^{-\Delta\Delta CT}$ calculation. Statistical analysis was performed using two-tailed student T-test **p<0.005, ***p<0.0005, ****p<0.0001

.



Appendix Figure 5. 5: Expression levels of specific genes in HPV+ and HPV- Oropharyngeal cancer from TCGA database. Box whisker plots showing the Log2 of normalised counts of genes comparing HPV16+ oropharyngeal tumour samples (n=19) and HPV- tumour samples (n=41). A) MYC was significantly down in the HPV+ samples, B) EP300, C) CTFC showed no significant change. D)GATA2 was significant, E) STAT5A was significantly up, and BACH1 was significantly down in HPV+ tumours. Statistical analysis was performed using a two-tailed student T-test, **p<0.005.



Appendix Figure 5. 6: Western blot optimisation for SREBF2 protein analysis in HUH7 cells: Serial dilution of protein isolated from HUH7 cells was loaded at the concentrations 20µg, 30µg, 50µg and 70µg. The antibodies for the detection of SREBF2 and the internal control Actin were tested, and 20µg was determined to be sufficient for protein detection.

Chapter 6: The lncRNA TRINGS influences cell viability of HPV16 positive cervical cancers despite glucose starvation

6.1. Introduction

The other aim of this thesis was to investigate the impact of HPV16 on the long non-coding RNA family (lncRNAs). These RNAs are transcripts that are >200 nucleotides in length with no or limited protein coding potential (<100 amino acids) [45]. LncRNAs have a vast array of functions due to their ability to form secondary structures, enabling them to interact with DNA, RNA and protein molecules. LncRNAs can act transcriptionally or post-transcriptionally and as regulators of chromatin organisation. Their functions range from acting as enhancers, decoys or scaffolds by physically interacting with their targets.

The expression of lncRNAs is cell type-specific and usually in response to a range of stimuli. Any changes such as deficiency, overexpression or mutations to lncRNAs have been associated with multiple human diseases, including cancers [314]. Numerous studies describe an alteration in lncRNA expression in HPV associated cancers [170], particularly in cervical cancers.

The dysregulation of lncRNAs expression may contribute to HPV related carcinogenesis, similar to that of miRNAs. The HPV16 E6 and E7 viral oncogenes have been shown to be regulators of some lncRNAs, including MALAT1 [315], PVT1 [316], SNHG12 [317], DINO [318] and CCEPR [175]. Targeting these lncRNAs by HPV E6 and E7 impacts critical cellular pathways, including cell proliferation, migration, invasion, DNA damage repair and altering the chemosensitivity of cancer cells. This demonstrates that HPV E6/E7 can have a direct impact on lncRNA expression, with consequences leading to the promotion of tumorigenesis.

To further my interest in the lncRNA area of research, I was awarded an endeavour leadership fellowship to spend time in Prof Karl Munger Laboratory, Tufts University, USA. The expertise in this research group allowed me to understand the important function of lncRNAs and design mechanistic models behind HPV16 influencing their expression. Although the expertise of this lab is mainly focused on cervical cancers, I gained valuable skills and knowledge that I could apply to HNCs.
RNA sequencing data was obtained from HPV16 E6/E7 expressing primary human foreskin keratinocytes (HFKs), and 1453 lncRNAs with at least 2-fold alteration in

expression were identified in the HPV16 E6/E7 expressing cells compared to the control [170]. From this list, the lncRNA TRINGS was identified as one of the most significantly downregulated lncRNAs in HPV16 E6/E7 expressing cells, with a decrease greater than 2-fold. TRINGS, or P53- regulated inhibitor of necrosis under glucose starvation, has previously been identified by Khan et al.[206] who conducted a microarray on p53 expression cells. TRINGS expression was described as directly upregulated by p53 under glucose starved conditions; this increase of TRINGs expression protected cancer cells from necrosis. In other cancer types Khan et al. [206] found TRINGS expression tends to be lower when comparing patients with wild type or mutated p53.There are currently no other studies investigating TRINGS in cancer or in relation to HPV, this lncRNA research is highly novel.

Given that p53 is a major target of HPV16 E6 and TRINGS expression was lower in HPV16 E6/E7 expressing cells, we aimed to identify the role of TRINGS in HPV related cancers. We also aimed to determine if glucose starvation could induce a change in TRINGS expression and have a protective effect over HPV infected cancer cells. This mode of action is described in Figure 6.1.

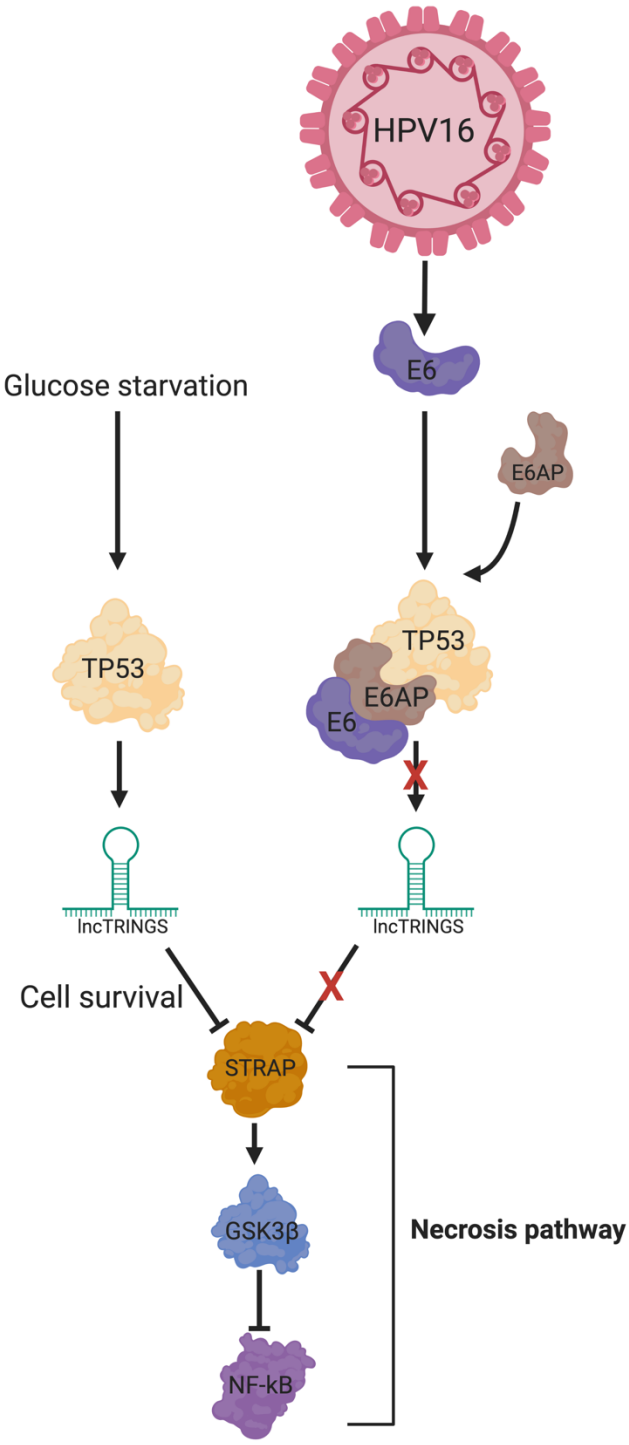


Figure 6. 1: TRINGS regulation of E6 degradation of p53 and its impact on the necrotic cell pathway.

6.2. Methods

6.2.1. RNA interference

Several custom, target specific siRNAs (ON-TARGETplus SMARTpools; Thermo Scientific Dharmacon) were used. Sequences for the siRNA are represented in Table 6.1. These sequences were obtained from Khan *et al.* [206]. A non-template binding control (D-001810-10) was also used (nontargeting pool; Thermo Scientific Dharmacon).

Table 6. 1: siRNAs used in the knockdown of TRINGS

Name	Sequence
siTRINGS 1	Sequence 1: 5`-GCACTGCCCTACTTACCTACT-3` Sequence 2: 5`-AGTAGGTAAGTAGGGCAGTGC-3`
siTRINGS 2	Sequence 1: 5`-AACCAGGTAGTCGTACAGTGT-3` Sequence 2: 5`-ACACTGTACGACTACCTGGTT-3`

All cell lines (Control and HPV16 E6/E7 iHFK and iNOKs, Caski, SiHa, HCT116 WT p53 and HCT116 P53 -/-) were seeded into a 12-well plate (Invitrogen, USA), 24hrs prior to transfection at $1x10^5$ cells/mL in triplicate wells. To each well, 1mL of the desired cells at a predetermined concentration in the media of choice was added. After 24hrs, transfection was performed using the Lipofectamine® 3000 Reagent Protocol (Invitrogen, USA). Each well was transfected with 20pmol of desired siRNA, 2.0μL of RNAimax (Invitrogen, USA). The transfected cells were then incubated at 37°C with 5% CO_2 in a humidified incubator (FORMA Scientific). Cell viability assays were measured, and RNA was harvested at 24-hour intervals post-transfection up to 96 hours.

6.2.2. RNA isolation and quantitative PCR

Total RNA was isolated using Quick-RNA™ Miniprep Kit (Zymo Research). Cells were lysed using 300uL of RNA lysis buffer in the 12 well plate using a cell scraper. RNA isolation was performed as per the suggested protocol, except RNA was eluted in 20µl of DNase/RNase-Free dH_2O and quantitated using a Nanodrop™ UV-VIS spectrophotometer (ThermoFisher Scientific, USA).

Random primer cDNA was synthesised from 500ng of total RNA. This was performed using the QuantiTect® Reverse Transcription Kit as per the manufacturer's protocol. Thermal cycling was run on a 2720 Thermal Cycler (Applied Biosystems, USA). Following the thermal cycling, the cDNA mixture was diluted 1 in 4 by adding nuclease-free water. Quantitative PCR (qPCR) was performed in triplicate using SYBR Green PCR Master Mix (Applied Biosystems) reagents in a StepOne Plus (Applied Biosystems) thermocycler system. For all qPCRs in this study, thermocycler settings of 20s at 95°C followed by 40 cycles of 3s at 95°C and 30s at 60°C were used. Table 6.2 represents the qPCR primer sequences were used in this study

Table 6. 2 :qPCR primers used in this study:

Target gene	Primer sequence
TRINGS	Forward: 5`- GGACAAGTATGCACTGCCCT -3` Reverse: 5' AAGAACGCATGAGCTGGTGA -3`
GAPDH	Forward: 5`-GATTCCACCCATGGCAAATCC-3` Reverse: 5`-TGGGATTTCCATTGATGACAA-3`

6.2.3. Cell viability assay

Cell viability was measured at 24-hour intervals post-transfection up to 96 hours. Media was removed from cells and 10µg/ml of Resazurin (Sigma) was added to determine redox fitness. Cells were incubated with the dye at 37°C with 5% CO_2 for two hours. Sample fluorescence was read in triplicate wells with 100µL of each transfection well on a 96 well plate. Sample fluorescence was read using 560nm excitation and 590nm emission filters on a Synergy H1 microplate reader (BioTek).

6.2.4. Interactome and gene ontology analysis

A network representing putative interactions between TRINGS, P53 and HPV16 E6 proteins was built from publicly available datasets. TRINGS predictive targets were obtained from Khan *et al.* [206] RNA pull-down assay and those genes 10kb upstream and downstream from TRINGS using UCSC Genome Browser version GRCh37/hg19 annotations. The interactome data for the Papillomavirade family was downloaded from the virus mentha website (virusmentha.uroma2.it) and filtered to retain only the interactions between HPV16 E6 and human proteins. The resulting table (HPV16 E6-host interactome and TRINGS-targets interactome) were loaded into Cystoscope v3.3.0. The complete human interactome was loaded from BioGrid (version 3.4.129) through Cystoscope. The HPV16 E6 interactome and the TRINGs interactome were merged with the human interactome to identify second-order neighbours of the viral proteins and TRINGS targets as well as extract the P53-target interactome. This expanded interactome was filtered to generate an HPV16 E6/TRINGS/P53 focused regulatory network and filtered to include only shared targets between these main genes of interest. Gene Ontology analysis was conducted using the web-based gProfiler [319] with a cut off p-value of 0.05. GO.

6.3. Results

6.3.1. LncRNA TRINGS expression level is independent of glucose starvation but is dependent on p53 levels

In previous studies, TRINGS has been described to be dependent on p53 expression under glucose starved conditions in various cancer cells [206]. We investigated if this relationship existed in HPV16+ cervical cancer cells as p53 is the main target of HPV16 E6. First, we measured the relative expression levels of TRINGS across various cells lines (Figure 6.2). In the cell lines with no or low levels of p53, there is a lower expression of TRINGS. When comparing HCT116 with wild type p53 and HCT116 with null p53, TRINGS relative expression is lower in the p53 null cells. The expression level of TRINGS is lower in iNOKs expressing HPV16 E6/E7 compared to the control cells. This trend is not as distinct in the iHFKs, but there is still a lower level of TRINGS in the E6/E7 expressing cells compared to control cells.

In the cervical cancer cell lines, SiHa and Caski, both expressing HPV16, there is a relatively low expression level of TRINGS with extremely low levels in Caski. In HPV16 E6/E7 expressing iNOKs, E6/E7 iHFKs, SiHa and Caski cells, there are lower expression levels of TRINGS due to the lower expression of p53 levels, as seen in the HCT116 p53 null cells.

We then asked if the pattern of HPV16 E6/E7 expression and low TRINGS expression holds true over time in replicating cells. Over a span of 4 days, we observed that the expression levels of TRINGS remained much lower in E6/E7 expressing iNOKs and HFKs compared to control cells (Figure 6.3). This was under normal glucose conditions. In the iNOKs, the lower expressing levels of TRINGS in E6/E7 expressing cells was statistically significant across all time points. In the iHFKs, there was only a statistically significant change at 48hrs, but the trend of lower expression TRINGS in E6/E7 expressing cells appears across time points. This perhaps is due to there already being such a low relative expression of TRINGS.

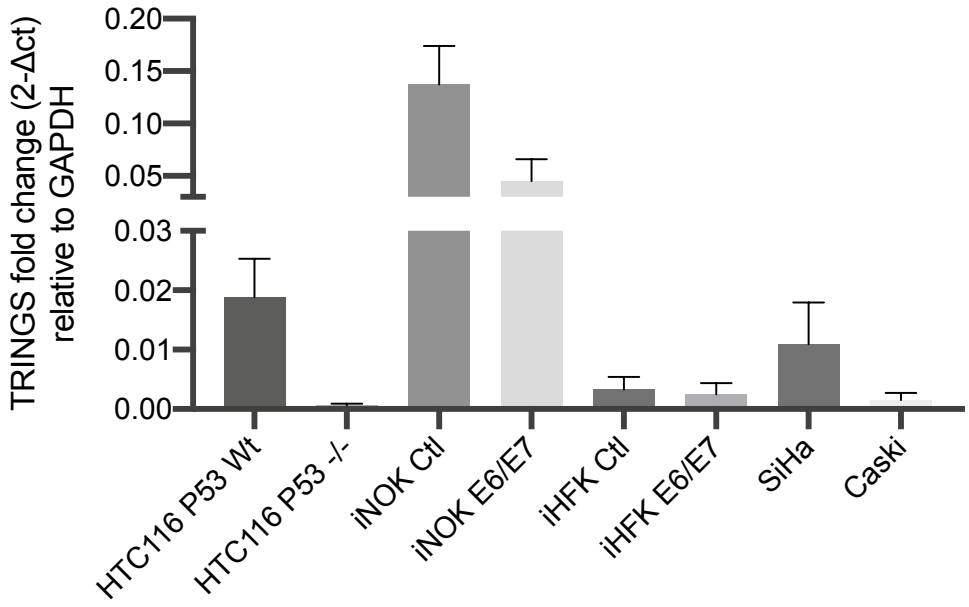


Figure 6. 2: Relative expression levels of TRINGS in a cell lines panel Cell lines with a higher level of p53 (HCT116 P53 WT, iNOK Ctl and iHFK Ctl) all appear to have a greater level of TRINGS expression compared to cell lines with lower p53 levels (HCT116 p53 -/-, iNOK E6/E7, iHFK E6/E7, SiHa and Caski). The normalisation of gene expression was achieved using the calibrator GAPDH expression and fold relative to GAPDH using $2^{-\Delta Ct}$ calculation.

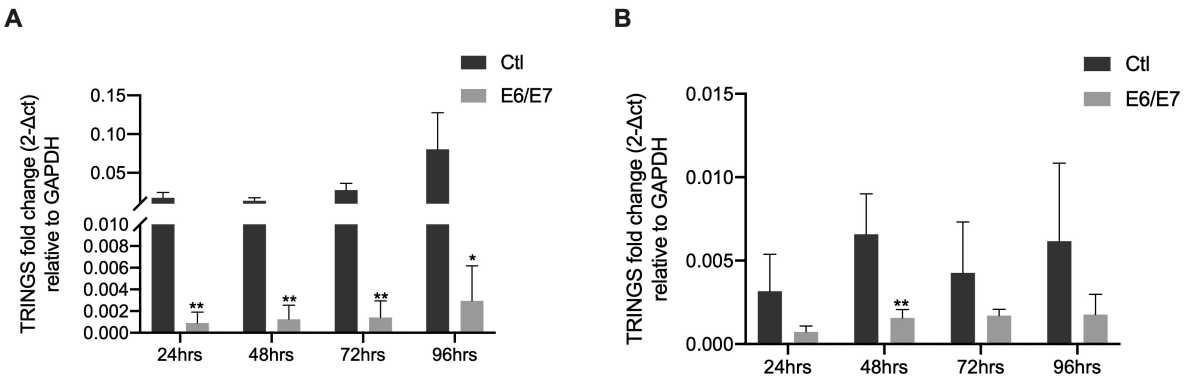


Figure 6. 3: Relative expression levels of TRINGS over time in immortalised keratinocyte cell lines comparing control (Ctl) cells to cells expressing HPV16 E6/E7. The cells lines A) iNOK and B) iHFK grew with normal glucose conditions for 24hrs, 48hrs, 72hrs and 96hrs. The normalisation of gene expression was achieved using the calibrator GAPDH and fold relative to GAPDH using the $2^{-\Delta Ct}$ calculation Statistical analysis was performed using a two-tailed student T-test, *$p \leq 0.05$, **$p \leq 0.01$.

.

These results may suggest a relationship between HPV16 E6/E7 expression and TRINGS expression, most likely due to low p53 levels. Khan *et al.* [206] described TRINGS expression as being dependent on glucose starvation in cancer cells. With this understanding, our next aim was to determine if glucose starvation impacts the expression levels of TRINGS in HPV16+ cervical cancer cells (Figure 6.4). In two different HPV16+ cervical cancer cell lines, SiHa and Caski cells were treated with glucose (DMEM with GlutaMAX, 4500mg/L) or no glucose media over the course of 72 hours.

There was no statistically significant difference in TRINGS expression between cells under glucose normal and glucose starved conditions in both cervical cancer cell lines. As cells proliferated over time, the expression levels of TRINGS increased over time (compared to day 1), but there was no change when comparing glucose normal and glucose starved cells.

To determine if this observation is due to the low levels of p53 not activating TRINGS expression (due to the degradation by HPV16 E6), we conducted the same experiment in HCT116 cells with wild type (WT) p53 (Figure 6.5). There was no statistically significant change in TRINGS expression when cells were treated with normal glucose and starvation conditions. At 48 and 72hrs of glucose starvation, there does appear to be an increase in TRINGS expression, although this is not statistically significant.

These results indicate that glucose starvation alone does not stimulate the expression of TRINGS in HPV16+ cervical cancer cell lines and that p53 is required to promote TRINGS expression. P53 is usually activated in response to external stresses on a cell, such as DNA damage or nutrient starvation [320-322]. Perhaps, in this case, glucose starvation may activate the expression of p53, leading to an increase in TRINGS expression, as with WT p53 in the HCT116 cell lines, we see an increase of TRINGS post 48hrs of glucose starvation (Figure 6.5). We cannot be definitive about this as there was no statistically significant change. We can conclude that glucose starvation alone does not lead to the activation of TRINGS expression.

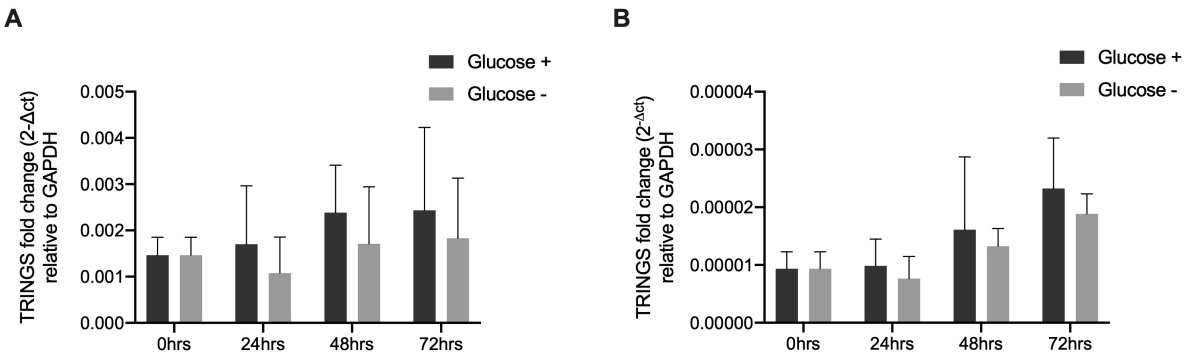


Figure 6. 4: Relative expression levels of TRINGS over time in cervical cancer cell lines under glucose starved conditions. The cells lines A) SiHa and B) Caski were treated with normal or no glucose media for 0hrs, 24hrs, 48hrs and 72hrs and TRINGS expression was measured at each time point. The normalisation of gene expression was achieved using the calibrator GAPDH and fold relative to GAPDH using the $2^{-\Delta Ct}$ calculation. Statistical analysis was performed using a two-tailed student T-test.

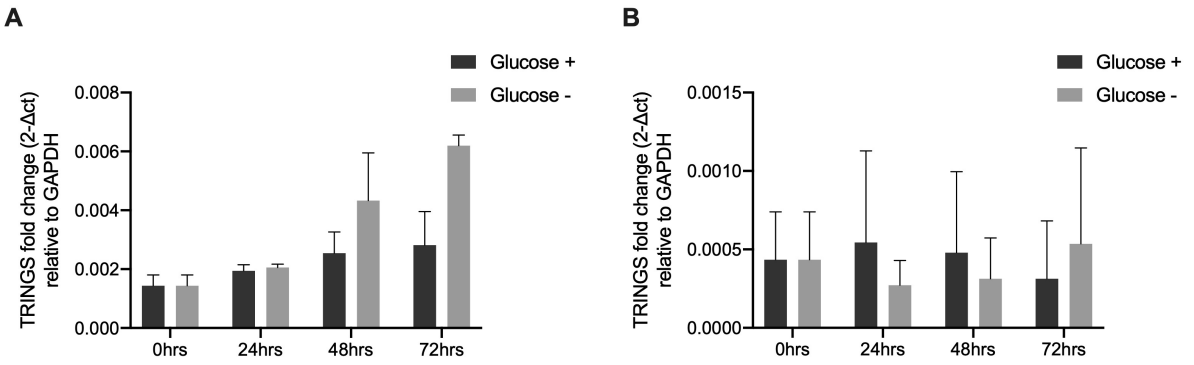


Figure 6. 5: Relative expression levels of TRINGS over time in colorectal cancer cell line HCT116 under glucose starved conditions comparing cells A) Wild type P53 and B) Null P53. Cells were treated with normal or no glucose media for 0hrs, 24hrs, 48hrs and 72hrs, and TRINGS expression was measured at each time point. The normalisation of gene expression was achieved using the calibrator GAPDH and fold relative to GAPDH using the $2^{-\Delta Ct}$ calculation. Statistical analysis was performed using a two-tailed student T-test.

6.3.2. Modulation of TRINGS expression can influence cell viability of cancer cells

TRINGs has been linked to promoting cell survival under glucose starved conditions in cancer cells [206]. We aimed to determine if TRINGS played a role in cancer cell survival when cells are infected with HPV16 and have inactive p53.
We utilised a siRNA knockdown system, using two siRNA in combination. These siRNAs were validated to knockdown TRINGS by 50-60% individual (Appendix figure 6.1). Cells were grown in either media with 4500mg/L glucose or glucose-free media.

In the HPV16+ cervical cancer cell line SiHa (Figure 6.6), we successfully knockdown TRINGS expression level by 70-80% over four days, under glucose and no glucose conditions. With this decrease in TRINGS expression, we saw that there was a significant decrease in cell viability across each time point, relative to control cells at day 0. From a 10% decrease at 24hrs to a 70% decrease in viability at 96hrs post-transfection. This decrease in cell viability occurred under normal and glucose starved conditions at a similar rate. We observed a similar trend in another HPV16+ cervical cancer cell line, Caski, with a 40-50% reduction in TRINGS expression under both glucose conditions. Under normal glucose conditions, a decrease in TRINGS led to a steady decline in cell viability from 5% at 24hrs to 50% 96hrs post-transfection. Under glucose starved conditions, we see the same trend; however, only 24hrs post-transfection (0hrs glucose starvation) and 48hrs post-transfection (12hrs glucose starvation) was this change statistically significant. We see a decrease in cell viability at other time points, but these changes were not significant.

It is evident from the results that a reduction in the expression of TRINGS will decrease cell viability over time in HPV16+ cervical cancers. This decrease in cell viability is due to TRINGS expression and not glucose starvation, as we noted the same trend in cells growing under ‘normal’ glucose conditions.

The cell protective effect of TRINGS has been shown to only occur in cancer cells and not in normal cell lines [206]. We expected this to be the case in our non-cancer cell lines. We aimed to determine if HPV16 expression alone could alter TRINGS effect on cell viability. To this end, we conducted our siRNA knockdown system in two non-

cancerous cell lines, iNOKs and iHFKs, comparing control and HPV16 E6/E7 expressing cells.

In the iNOK cells, TRINGS was successfully reduced to 80-90% across all time points. With the control cells, there was no decrease in cell viability with the reduction of TRINGS. There was a 10% decrease in viability for iNOKs expressing E6/E7 at 72hrs post-transfection, but no significant change at all other time points. These observations suggest that in non-cancerous cells with wild type p53, the knockdown of TRINGS will have little to no impact on cell viability. We obtained variable results when this was repeated in another non-cancerous cell line. In the iHFKs, we successfully knockdown TRINGS by 60-70% in the control cells and 50-80% in the E6/E7 expressing cells. In the control cells, we see a decrease in cell viability of approximately 40-60%, at 48, 72 and 96hrs post-transfection. These cells have a wild type p53, but the reduction of TRINGS still decreased cell viability. In the E6/E7 expressing cells there was a 10% decrease in cell viability at 48 and 72 hrs post-transfection. Again, these cells have low p53 levels, and a decrease in TRINGS had a minor impact on cell viability, although not as significant when compared to the cancer cell lines. With these findings, we cannot conclusively state that TRINGS expression does play a role in the cell viability of non-cancerous cells, as between two different cell lines, we see differing results; TRINGS regulation of cell viability may be independent of HPV16 and p53.

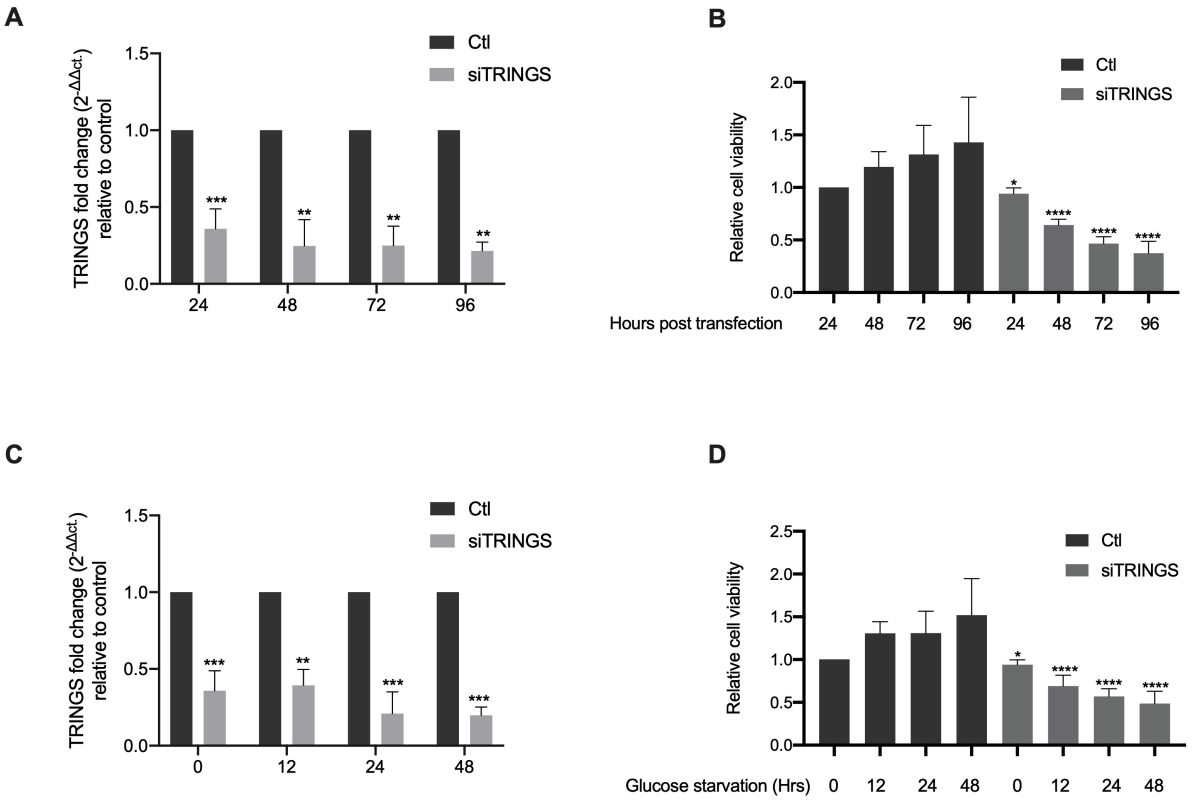


Figure 6. 6: Knockdown of TRINGS using siRNA in HPV16 positive cell line SiHa at 24hrs, 48hrs, 72hrs, and 96hrs post-transfection under normal glucose and glucose starved conditions. A and C) normalised TRINGS expression at different time points under normal glucose (A) and glucose starved (C) conditions. B and D) Relative cell viability of these TRINGS knockdown cells under normal glucose (B) and glucose starved (D) conditions. The normalisation of gene expression was achieved using the calibrator GAPDH and fold relative to the control (Ctl) cells using the $2^{-\Delta\Delta Ct}$ calculation. Cell viability was calculated relative to the Day 1 control cells. Statistical analysis was performed using two-tailed student T-test ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$ and *$p \leq 0.05$.

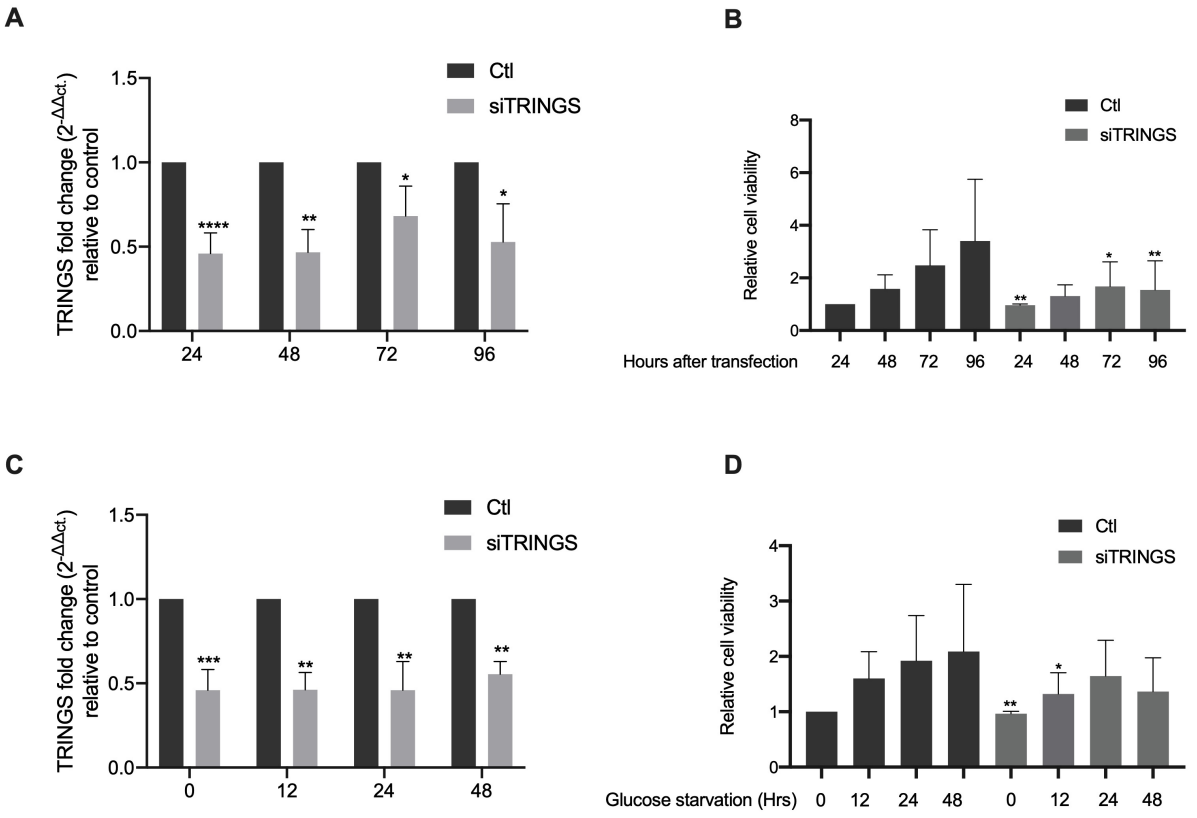


Figure 6. 7: Knockdown of TRINGS using siRNA in HPV16 positive cell line Caski at 24hrs, 48hrs, 72hrs, and 96hrs post-transfection under normal glucose and glucose starved conditions. A and C) normalised TRINGS expression at different time points under normal glucose (A) and glucose starved (C) conditions. B and D) Relative cell viability of these TRINGS knockdown cells under normal glucose (B) and glucose starved (D) conditions. The normalisation of gene expression was achieved using the calibrator GAPDH and fold relative to the control (Ctl) cells using the $2^{-\Delta\Delta Ct}$ calculation. Cell viability was calculated relative to the Day 1 control cells. Statistical analysis was performed using two-tailed student T-test ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$ and *$p \leq 0.05$.

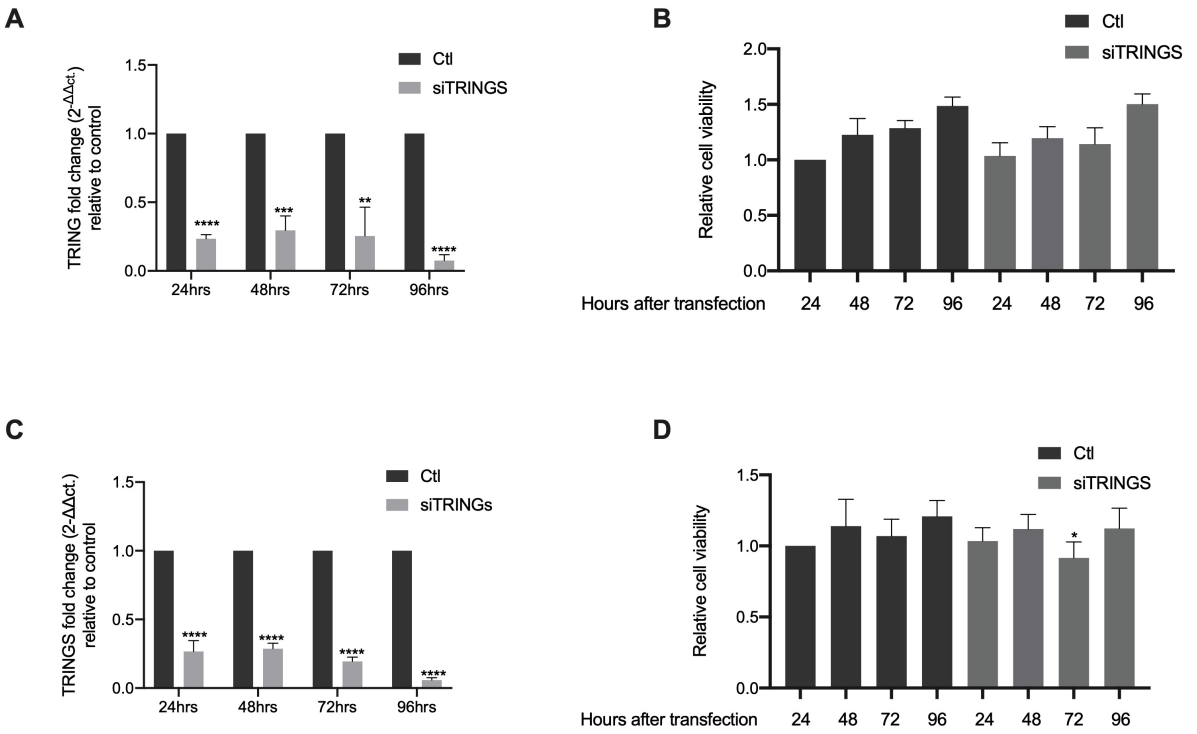


Figure 6. 8: Knockdown of TRINGS using siRNA in iNOK cells at 24hrs, 48hrs, 72hrs, and 96hrs post-transfection comparing control cells and cells expressing HPV16 E6/E7 A and C) normalised TRINGS expression at different time points under normal glucose (A) and glucose starved (C) conditions. B and D) Relative cell viability of these TRINGS knockdown cells under normal glucose (B) and glucose starved (D) conditions. The normalisation of gene expression was achieved using the calibrator GAPDH and fold relative to the control (Ctl) cells using the $2^{-\Delta\Delta Ct}$ calculation. Cell viability was calculated relative to the Day 1 control cells. Statistical analysis was performed using two-tailed student T-test ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$ and *$p \leq 0.05$.

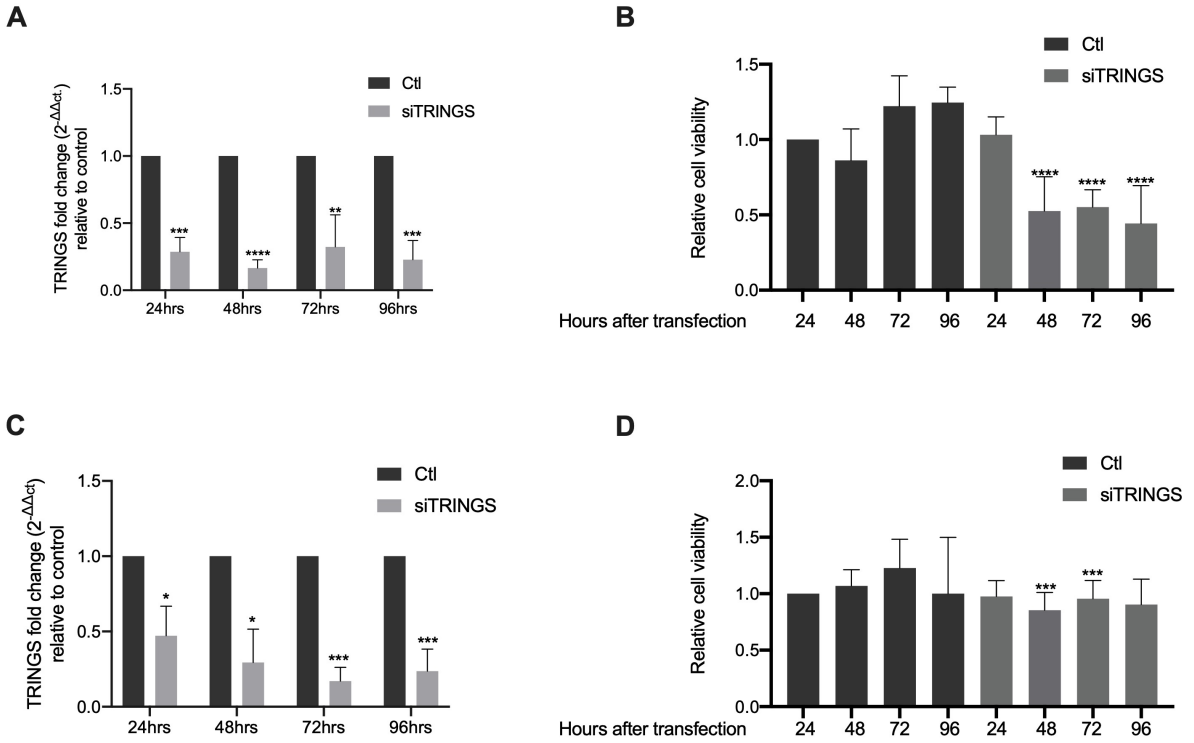


Figure 6. 9: Knockdown of TRINGS using siRNA in iHFK cells at 24hrs, 48hrs, 72hrs and 96hrs post-transfection comparing control cells and cells expressing HPV16 E6/E7. A and C) normalised TRINGS expression at different time points under normal glucose (A) and glucose starved (C) conditions. B and D) Relative cell viability of these TRINGS knockdown cells under normal glucose (B) and glucose starved (D) conditions. The normalisation of gene expression was achieved using the calibrator GAPDH and fold relative to the control (Ctl) cells using the $2^{-\Delta\Delta Ct}$ calculation. Cell viability was calculated relative to the Day 1 control cells. Statistical analysis was performed using two-tailed student T-test ****$p \leq 0.0001$, ***$p \leq 0.001$, **$p \leq 0.01$ and *$p \leq 0.05$.

6.3.3. The modulation of TRINGS in cells with wild type p53

Since TRINGS expression is dependent on p53 expression, we sought to determine if the reduction in TRINGS would still decrease cell viability if cancer cells had wild type (WT) p53 levels. The rationale is that in the cervical cancer cells we observed, TRINGS still plays a role in cancer cell survival without p53. In cells with WT p53, we wanted to observe if the knockdown TRINGS will change cell viability. If these cancer cells survive, then p53 recruits other components to increase cell survival and is not reliant on TRINGS expression.

In the colorectal cancer cell line with wild type p53, HCT116, (Figure 6.10) we knocked down TRINGS under normal glucose conditions by 60-70% at 24 and 48hrs post-transfection and by 40-70% under glucose starved conditions across all time points. There was no statistically significant change in cell viability, with the reduction of TRINGS, in either the normal or glucose starved cells. This indicates that cancer cells with wild type p53 will remain protected even with TRINGS reduction and despite glucose starvation. P53 is potentially recruiting some other cellular component to protect these cancer cells.

We did conduct our siRNA knockdown in HCT115 p53 null cells to determine if TRINGS protected these cells for comparison. Still, the relative expression of TRINGS was so low we could not draw any conclusions from these results (Appendix Figure 6.2); we cannot further knockdown a gene that is already so lowly expressed in a cell.

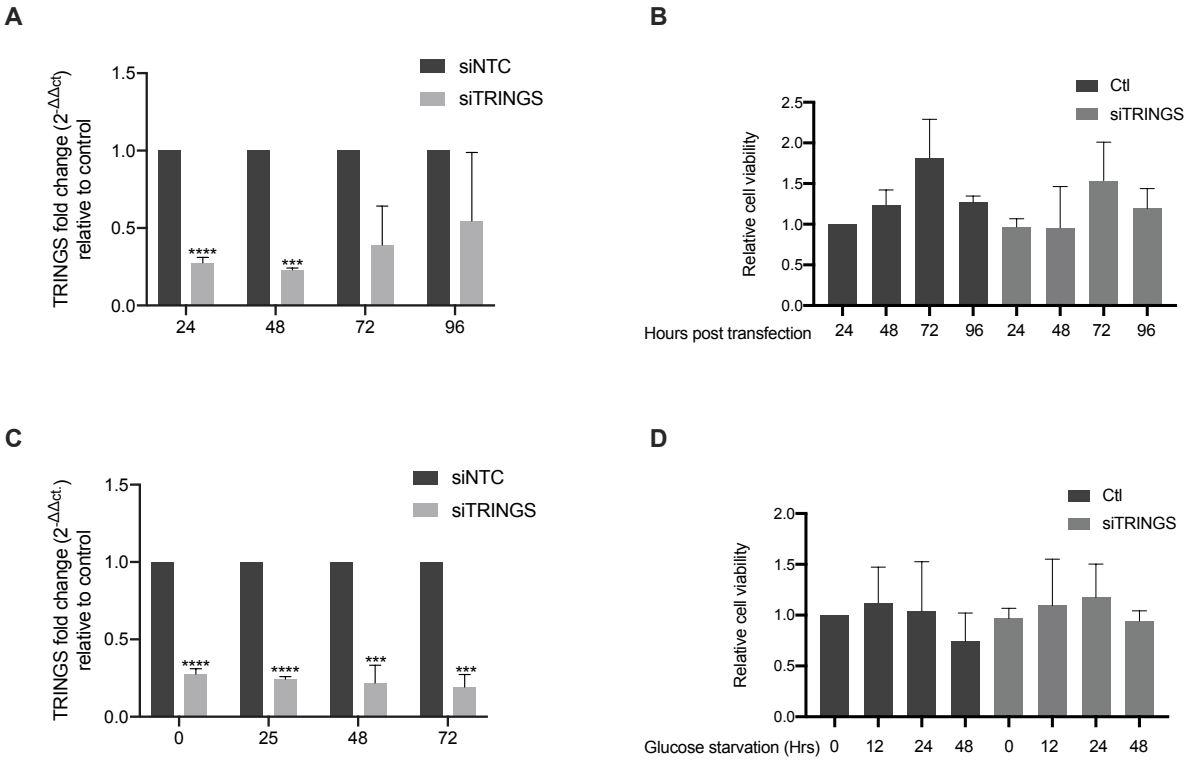


Figure 6. 10: Knockdown of TRINGS using siRNA in the cell line HCT116 with wild type p53 at 24hrs, 48hrs, 72hrs, and 96hrs post-transfection under normal glucose and glucose starved conditions. A) Relative expression levels of TRINGS at different time points under normal glucose conditions and B) Relative cell viability of these cells under normal glucose conditions. C) Relative expression levels of TRINGS under glucose starved conditions and D) Relative cell viability of these glucose starved cells. The normalisation of gene expression was achieved using the calibrator GAPDH and fold relative to the control (Ctl) cells using the $2^{-\Delta\Delta Ct}$ calculation. Cell viability was calculated relative to the Day 1 control cells. Statistical analysis was performed using a two-tailed student T-test ****$p \leq 0.0001$, ***$p \leq 0.001$.

6.3.4. TRINGS expression in Head and Neck cancers

The Human papillomavirus type 16 is also an important etiological agent in Head and Neck Cancers and warrants investigation into TRINGS role in this cancer type We aimed to determine if TRINGS expression was also altered in HPV16 infected Head and Neck cancers.

Using an HNC cell lines panel (Figure 6.11), we found that the SCC154 and SCC090 (HPV16+ cells) appear to have a lower expression of TRINGS, although several other HPV- HNC cell lines (UMSCC22B, SCC38, HN13, HN12 and SCC6) also show low levels of TRINGS. For the majority of HNC, p53 mutations are common. Given TRINGS expression levels have a positive correlation to p53 expression and Khan et al [206] demonstrated wild type p53 is required to upregulation of TRINGS, we would expect a low level of TRINGs in cells with mutated p53. Of these cells, UMSCC22B [323], SCC38 [324], HN12 [325], HN13 [325] and SCC6 [326] have been shown to have mutations to p53, and perhaps explains why we observe low level of TRINGS expression. In our other cell lines HN31 [325], HN30 [325], HN26 [325] all have WT p53, although HN17 [325] has a mutated p53. In the HN17 cells, we would expect TRINGS to have a low expression level due to p53 mutation; perhaps there is another factor in this cell line leading to a greater level of TRINGS expression.

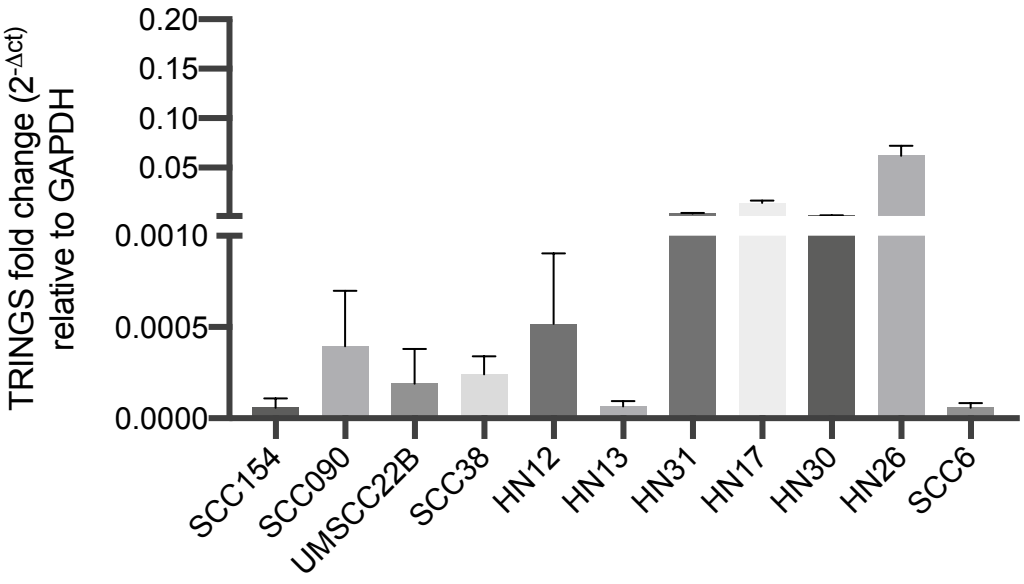


Figure 6. 11: Relative expression levels of TRINGS in a Head and Neck cancer cell line panel SCC154 and SCC090 are HPV16+ cells, and SCC22B, SCC38, HN12, HN13, HN31, HN17, HN30, HN26 and SCC6 are all HPV- cell lines Normalisation of gene expression was achieved using the calibrator GAPDH expression and fold relative to GAPDH using $2^{-\Delta Ct}$ calculation.

To further investigate if TRINGS has a clinical association with HPV16, we measured the relative expression level of TRINGS in an HNC patient cohort [101] (Figure 6.12). There was no significant change in TRINGS expression when comparing normal tissue to HPV- and HPV16+ tumour tissue. Although there is a downward trend in HPV16+ tumours compared to normal, this was not significant. We then interrogated another patient cohort, using The Cancer Genome Atlas data (Figure 6.13), comparing both HNC and cervical cancer cohorts. In the cervical cancer cohort, we noted a significant decrease in TRINGS expression between HPV- tumours and tumours infected with both HPV16 and HPV18; this supports the results from our cervical cell line data. We saw no change in TRINGS expression between normal and tumour samples in the HNC cohort, both HPV+ and negative. However, the was a slight increase in the expression of TRINGS when comparing HPV16+ to HPV16- tumours. This is the opposite trend compared to the cervical cancer cohort. This may be due to p53 being commonly mutated in HPV- HNC whereas, in the HPV16+ HNC, there are low levels of p53, but it is still the wild type p53, and residual p53 may still be having an effect on TRINGS expression.

These results taken together suggest that the impact of TRINGS may not be as significant across other HPV+ cancer types, particularly comparing HNC and cervical cancer. HPV itself may not directly have an effect on TRINGS expression, and p53 is the main determinant of TRINGS expression.

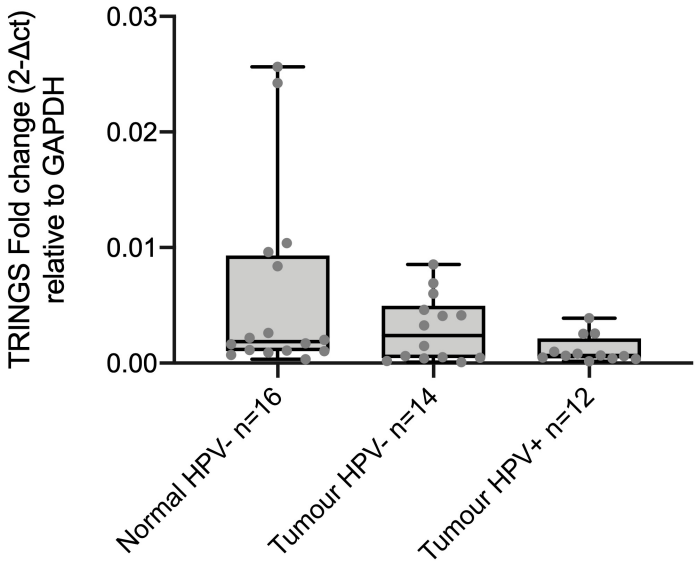


Figure 6. 12: Relative expression levels of TRINGS in a Head and Neck cancer patient tumours comparing normal tissue (n=16) to tumour tissue both HPV- (n-14) and HPV16+ (n=12). There was no statistically significant difference in the expression levels of TRINGS across all groups. The normalisation of gene expression was achieved using the calibrator GAPDH expression and fold relative to GAPDH using $2^{-\Delta Ct}$ calculation.

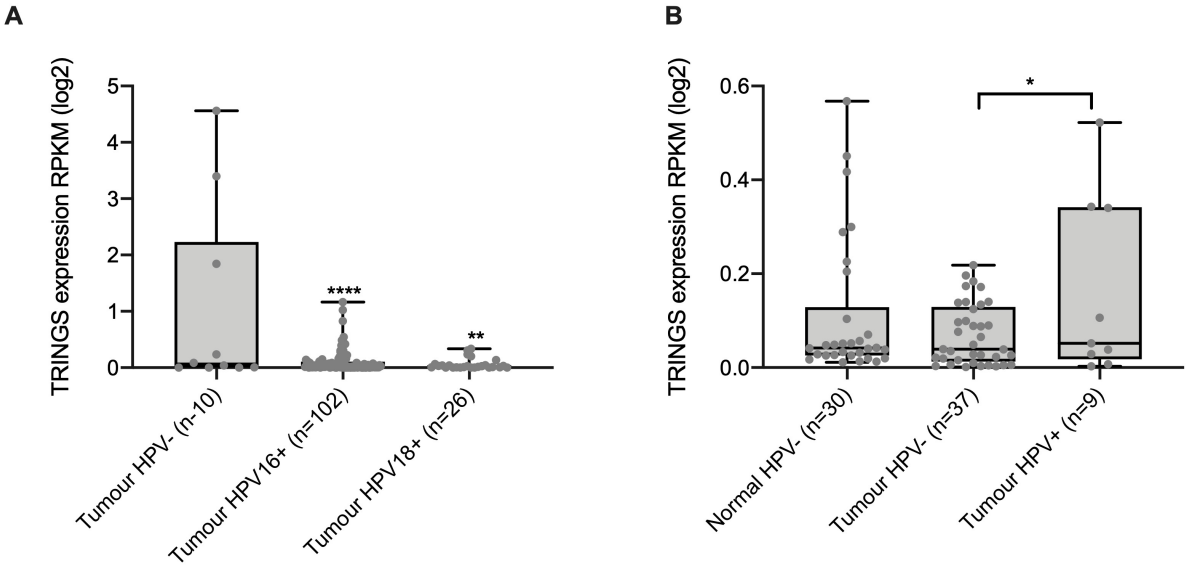


Figure 6. 13: Expression levels of TRINGS in comparing tumours that are HPV+ and HPV- in cervical and oropharyngeal cancers from TCGA database: Box whisker plots showing the normalised counts for genes in comparing HPV status in A) Cervical cancers and B) Oropharyngeal cancers. Statistical significance was determined using a student T-test ****$p \leq 0.0001$,**$p \leq 0.01$ and *$p \leq 0.05$.

6.3.5. Determining the relationship between TRINGS and HPV16 using an interactome

Our results revealed TRINGS expression is associated with p53, a major target of HPV16 E6, and is involved in the cell viability pathway in HPV+ cervical cancers. In this next section, we aimed to uncover the molecular interactions occurring in HPV16+ cervical cancers. First, we obtained targets of TRINGS identified by Khan *et al.* [206] who used an RNA pulldown assay. Of these targets Khan *et al.* [206] confirmed TRINGS could bind to STRAP to regulate the necrotic signalling pathway. In addition to this we aimed to identify the pathways other TRINGS targets were involved in. We also identified potential novel TRINGS targets that are 10kb upstream and downstream from TRINGS using the UCSC Genome Browser [220]. The TRINGS targets were then matched to p53 and E6 targets using Biogrid [214] to identify p53 targets and VirusMentha [327] for E6 targets (Figure 6.14).

None of TRINGS predicted targets are directly regulated by E6, but HERC5 and YY1 are targeted directly by both TRINGS and p53. There are some secondary interactions between TRINGS targets and E6, whereby E6s targets SIRT7, CREBBP, EP300, BRCA1 and IRF3, which all interact with one or more of TRINGS, predicted targets. There are also many secondary interactions between TRINGS targets and p53 targets.

This evidence suggests p53 has more influence on TRINGS than E6, as it shares more targets with TRINGS than the viral oncogene. There is also a much bigger network between TRINGS and p53 compared to HPV16 E6.

To explore the functional impact of TRINGS and its network, we performed a Gene Ontology analysis using g: Profiler [319] (Figure 6.15). The top three biological processes of these genes in the TRINGS/p53/E6 interactome are all associated with the regulation of metabolic processes. Regulation of metabolic processes is crucial for cell viability and survival, and perhaps TRINGS association with these genes from the interactome are pathways to investigate in future research.

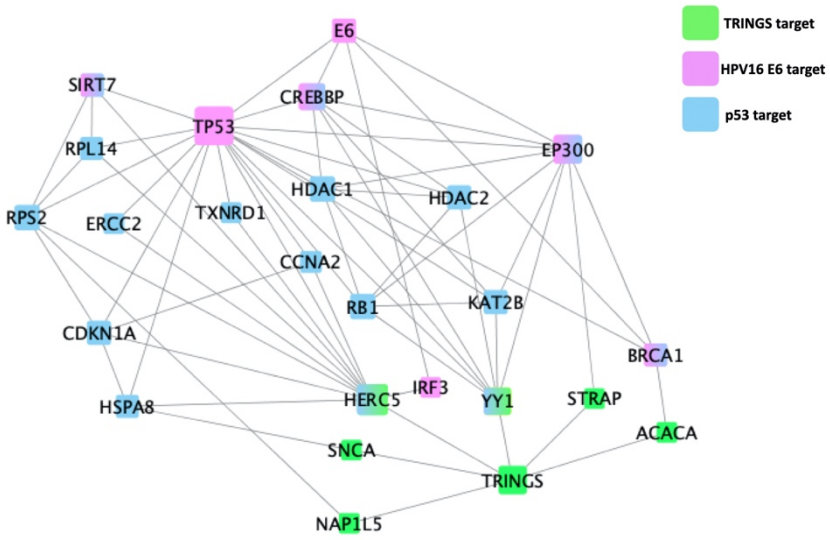


Figure 6. 14: Interactome depicting the lncRNA TRINGS, HPV16 E6, p53, each of their targets and how they interact with each other. Each square represents a gene, colour coded according to its regulatory factor (TRINGS, E6 and P53). Green represents those genes targeted by TRINGS; Pink represent those targeted by HPV16 E6, and blue indicate those targeted by P53. The size of the squares indicates the number of interactions that gene has; the larger the size, the more interactions.

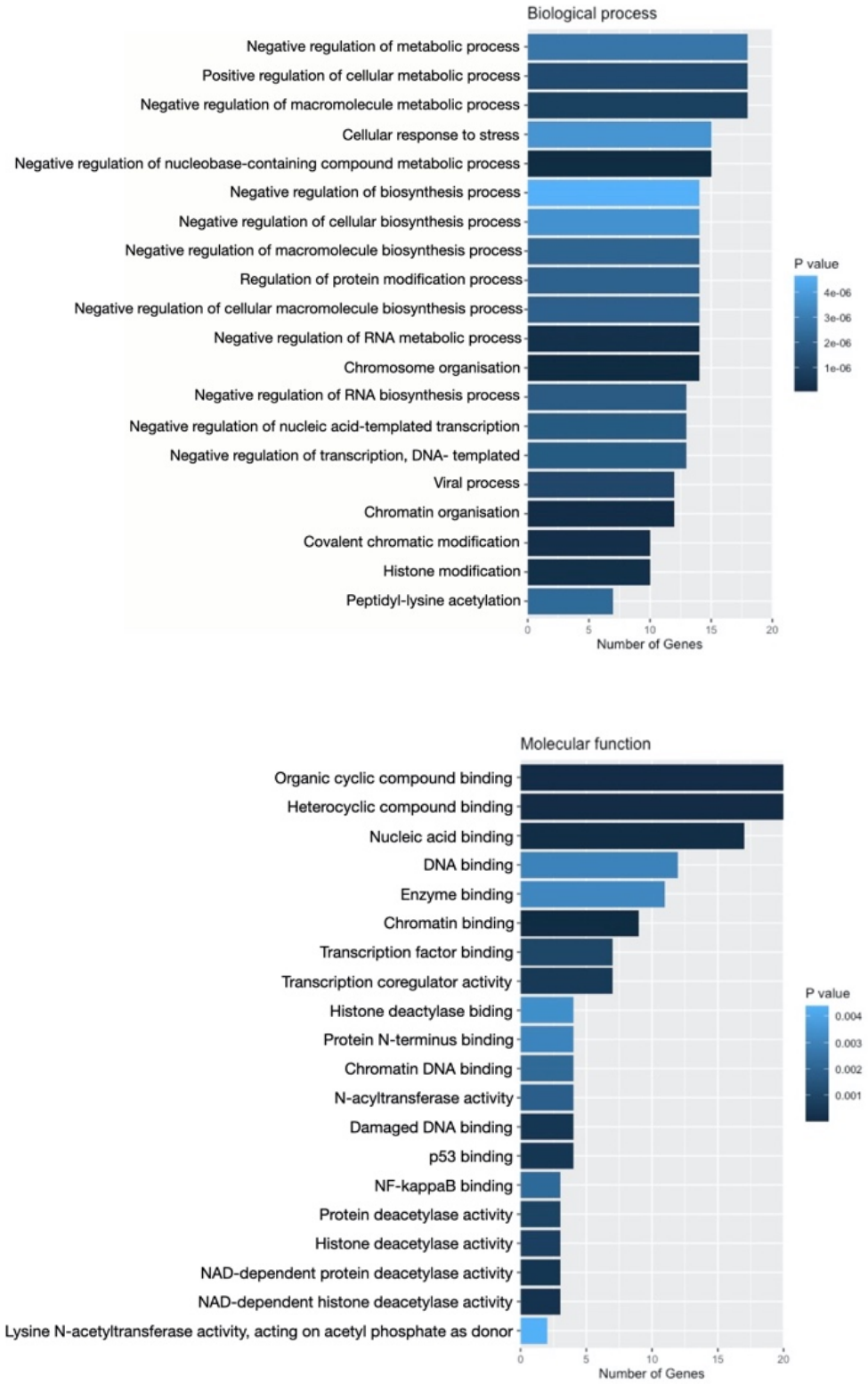


Figure 6. 15: Gene Ontology of TRINGS/p53/HPV16 E6 interactome*:*

6.4. Discussion

Long non-coding RNAs have a vast array of functions within cells. The Human papillomavirus is a major etiological agent for many cancer types, including cervical and Head and Neck cancers. It has been demonstrated that this virus can influence the expression of various genes and non-coding RNAs in different cancers types. For this chapter, we sought to determine the influences of HPV16 and its viral oncogenes E6/E7 on a specific lncRNA known as TRINGS [170]. This lncRNA TRINGS has previously been described to promote cell survival under glucose starved conditions, and its expression is directly regulated by the tumour suppressor p53 [206]. Given p53 is a major target of HPV16, specifically E6, we aimed to evaluate the cell survival role of TRINGS in HPV16+ cervical cancers as well as if TRINGS expression could be modulated by glucose starvation.

In the cell line panel, we measured the relative expression level of TRINGS. For cells that have no or low levels of p53, there is a lower expression of TRINGS. In iNOKs and iHFKs, there is a lower level of TRINGS in the E6/E7 expressing cells compared to control cells. This observation was consistent when cells were grown over an extended period of time. The expression levels of TRINGS remained much lower in E6/E7 expressing iNOKs and iHFKs compared to control cells.

These results indicate the TRINGS expression levels can be influenced by HPV16 E6/E7 expression. Although this is not a direct interaction, this may be due to p53 levels, as it is the main target of HPV16 E6. In our colorectal cancer cells, HCT116, p53 null cells had a significantly lower TRINGS expression compared to wild type p53 cells, indicating that TRINGS expression does follow a similar trend to p53 expression levels and suggests why we see a decrease in TRINGS expression in the presence of HPV16.

Khan *et al.* [206] demonstrated that TRINGS expression followed a similar trend to p53 expression in a series of cells line and the TCGA data. There is evidence of p53 inducing the expression of other lncRNAs such as PANDA, HN19, PURPL and PVT1 in cancer cells to promote cell survival [328-330].

6.4.1. The impact of glucose starvation on TRINGS expression

The presence of glucose has been demonstrated to affect TRINGS expression. It was shown that under glucose starvation, TRINGS expression is increased to protect cancer cells [206]. Thus, we aimed to determine if TRINGS expression could be stimulated in glucose starved HPV16 cervical cancer cells, despite low levels of p53.

The cervical cancer cell lines SiHa and Caski were treated with glucose and no glucose conditions over 72hrs. There was no difference in the expression levels of TRINGS between glucose normal and glucose starved cells over time. This indicates that the activation of TRINGS under glucose starved conditions does not occur in HPV16+ cervical cancer cells.

This may be due to the low levels of p53, not being able to promote TRINGS expression in these cells lines. To determine if p53 is activated by glucose starvation and not TRINGS, we repeated the same conditions in HCT116 cells with wild type p53. There was no significant change in TRINGS expression between glucose normal and glucose starved cells across all timepoints. However, there does appear to be an increase in TRINGS expression after 48hrs of glucose starvation. These results demonstrate that glucose starvation alone does not stimulate the expression of TRINGS; perhaps it requires extended glucose starvation over several days before p53 is activated to promote TRINGS expression. Khan *et al.* [206] tested several other treatments, including glucose starvation, FBS, serine or glutamine depravations and found that only glucose starvation could elicit the expression of TRINGS in p53 positive cells. This suggests that p53 itself is activated by glucose starvation and not TRINGS alone.

The activation of TRINGS under glucose starved conditions has been shown to promote cell survival of cancer cells [206]. We conducted siRNA knockdowns of TRINGS in our cervical cancer cell lines to determine if this had an effect on cell viability for HPV16 infected cells. In both HPV16+ cervical cancer cell lines, there was a significant decrease in cell viability with the knockdown of TRINGS. This trend was observed despite glucose conditions. Visually under the microscope, the knockdown of TRINGS led to a significant reduction in cell numbers, indicating cell death. This

was then confirmed with a cell viability assay using Resazurin and Trypan blue. (Appendix Figure 6.3).

In the HCT116 cells with wild type p53, there was no significant decrease in cell viability in a TRINGS depleted background under both glucose conditions. These results suggest that cancer cells with wild type p53 will remain protected even with TRINGS reduction and despite glucose starvation.

In our cervical cancer cells lines, we demonstrated that the modulation or decrease in TRINGS expression resulted in a decrease in cell viability. Khan *et al.* [206] have also demonstrated that increased TRINGS expression protected cells from cell death, and a decrease in TRINGS expression would reduce cell survival. In this paper, they also reported a protective effect over cancer cells was due to TRINGS involvement in the necrotic cell pathway via destabilisation of STRAP. Under glucose starvation, p53 would activate TRINGS expression, TRINGS would destabilise STRAP, and this, in turn, would inhibit the STRAP-GSK3 β-NF-κB necrotic signalling pathway, protecting cancer cells from cell death [206]. It would be beneficial in our research to measure the expression levels of STRAP in our HPV16+ cervical cancer cell lines to determine if the knockdown of TRINGS is leading to the stabilisation of STRAP, promoting this necrotic signalling pathway, resulting in the cell death we observed

6.4.2. The effect of TRINGS on normal cells

This protective effect by TRINGS has been observed to only occur in cancer cell lines. We expected this would be the case in non-cancerous cells. To this end, we conducted our siRNA knockdown system in two non-cancerous cell lines, iNOKs and iHFKs, comparing control and HPV16 E6/E7 expressing cells. These two cell lines gave differing results. In the iNOKs, there was no decrease in cell viability with a knockdown of TRINGS for both control cells and E6/E7 expressing cells, as we expected. Opposingly, in the iHFKs, the knockdown of TRINGS in the control cells led to a significant decrease in cell viability. In the E6/E7 expressing cells, there was a slight decrease in viability at 48 and 72hrs post-transfection.

In the original sequencing data, the overexpression of HPV16 E6/E7 was performed in the HFKs, and we observed a decrease in TRINGS expression; we would expect to see a decrease in TRINGS expression in cells as well as a decrease in cell viability with TRINGS knockdown. The differing results between iHFKs and iNOKs could be due to these two cell lines originating from very different cell types, foreskin keratinocytes and oral keratinocytes. These two cell types present different genomic and transcriptional patterns and thus will show different results with the modulation of TRINGS expression. With our differing results between two cells lines, we cannot definitively state that TRINGS impact on cell viability and survival is exclusive to cancer cells. In future research, it would be interesting to investigate TRINGS expression on another 'normal' cell type such as HEKa (epidermal keratinocytes) to gain more evidence to support this theory.

6.4.3. The expression of TRINGS in Head and Neck cancers

HPV16 impacts numerous cancer types, the second cancer type with the biggest burden of HPV infection is Oropharyngeal cancers. As there was a decrease in TRINGS expression in cervical cancers, we wanted to understand if the same trend existed in HPV+ oropharyngeal cancers. In cell lines, we observed TRINGS to have a low level of expression in HPV16+ cells, but there was variation in the HPV- cell lines. There appeared to be a trend of TRINGS expression being lower in cell lines with p53 mutations. In the TCGA patient cohort, there was a slight increase in TRINGS expression in the HPV+ compared to the HPV- tumours, although this comparison was nine patients (HPV16+) vs. 37 patients (HPV-), our sample size should be increased for a more reliable comparison. We would expect TRINGS levels to be lower in the HPV16+ HNC patients, similar to what we see in the cervical cancer patients. The TCGA HPV- cohort could be a combination of patients with and without mutated p53, and this could be influencing the trend we see in TRINGS expression. Khan *et al.* [206] found mutated p53 failed to upregulate TRINGS in cell lines. The mutation of p53 has also been shown to impede p53 transcriptional activation of MDM2 [331], p21 and MSH2 [332]. Given p53 is commonly mutated in HNC [333, 334] and the trend we see in our cell lines data, p53 mutations may be also be leading to a decreased expression level of TRINGS in HPV- HNC as p53 can no longer bind to the p53 response element

and promote TRINGs expression. Hence TRINGS is not a reliable biomarker for HPV detection in HNCs but is a potential biomarker candidate for cervical cancers.

6.4.4. Mechanistic models for the TRINGS – p53 network

An interactome with TRINGS/p53/HPV16 E6 was built to provide visual clues to possible mechanistic pathways. This interactome includes TRINGS predicted RNA targets as well as their association with p53 and HPV16 E6. We found that none of the TRINGS predicted targets are directly regulated by HPV16 E6. There are a few secondary interactions between TRINGS targets with another protein which is then bound by E6. Given this current mapping of known associations. This indicates that HPV16 E6 itself is not a direct regulatory of TRINGS; the main influence of TRINGS expression remains p53, as a result of HPV infection. Two of TRINGS predicted targets, HERC5 and YY1 are also direct targets of P53.

Of these, Yin Yang 1 or YY1 is a transcription factor that regulates many genes associated with cell differential, DNA repair, autophagy, cell survival and apoptosis [335]. YY1 is involved in both glucose pathways and necrotic/apoptotic pathways, similar to TRINGS. YY1 can enhance the transcriptional activity of ATF6, which is involved in inducing glucose-regulated protein genes [336]. YY1 is also a target of NF-kB; YY1 activation by NF-kB led to the repression of Bim, a pro-apoptotic gene [337], promoting the survival of multiple myeloma cells. Perhaps TRINGS targets many components of the STRAP/GSk3B/NF-kB pathway like YY1 in HPV driven cancers. YY1 has been shown previously to be overexpressed in HPV16, and HPV18 positive cervical cancers [338] and depletion of YY1 in these cancers leads to an increase in p53 expression and apoptosis [339]. It would be plausible to suggest there is a feedback loop between YY1/p53/TRINGS in HPV related cancers that needs to be investigated further.

LncRNAs like TRINGS add a critical layer to the regulation of p53 pathways in cancer, including HPV related cancers. Collectively the data described in this chapter indicated that TRINGS expression is not stimulated by glucose starvation alone but is dependent on p53 expression. In HPV16+ cancer cells, we see that with a low level of p53, there will be a low expression of TRINGS, but TRINGS still has some role in cell viability

and survival despite p53 silencing. Glucose starvation does not impact the function of TRINGS in HPV16+ cells, and it may be that there are other metabolic factors or stimuli which activate TRINGS expression in HPV+ cervical cancers.

6.4.5. Future directions:

In this study, we identified that the modulation of lncRNA TRINGS will have an impact on the cell death of cervical cancer cells. In future research, it would be interesting to investigate if this trend is due to the destabilisation of STRAP by TRINGS, as previously described [206]. We would expect that with the knockdown of TRINGS expression, there would be an increase in STRAP expression as it is no longer under the lncRNAs control. Investigation into the other downstream components of the necrosis pathway (GSK3 β and NF-κB) would indicate if the whole necrosis pathway is activated when it is no longer under TRINGS regulatory control.

From our interactome analysis, we also identified several genes (HERC5 and YY1) that are potential targets of TRINGS and that are involved in cell differentiation, DNA repair, autophagy, cell survival and apoptosis [335]. We also proposed a feedback loop between YY1/P53/TRINGS, given this gene is a target of both p53 and TRINGS. RNA pull-downs or ChIP assays will determine if TRINGS is bound to HERC5 or YY1 at a protein or RNA level and confirm they are under the lncRNAs control.

An overexpression system of TRINGS using a plasmid vector would complement our findings. We would expect the inverse relationship, with the overexpression of TRINGS in cancer cells, leading to cell survival upon various stressors. We could overexpress TRINGS and treat cells to glucose starvation treatment to attempt to induce cell death. Using overexpression of TRINGS, the potential TRINGS targets (STRAP, HERC5, YY1) could also be validated using RT-qPCR; the targets would be predicted to decrease with an increase in TRINGS.

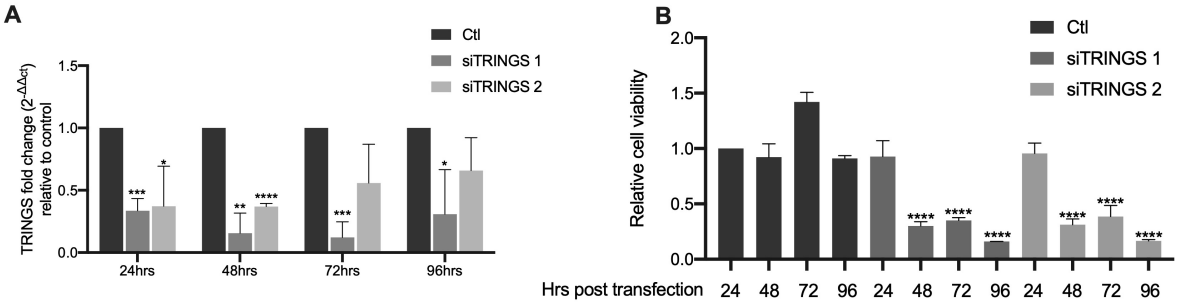
Given the differing result we observed in TRINGS impact on cell viability of iNOKs and iHFKs, performing the TRINGS knockdown in another cell line such as HEKa could give us more of an indication as to the impact of TRINGS modulation in normal cells. Another comparison that could be made in measuring the relative expression of

TRINGS in the HPV- cervical cancer cell line CA33 to confirm further TRINGS expression is lower in the HPV+ cervical cancer cell lines.

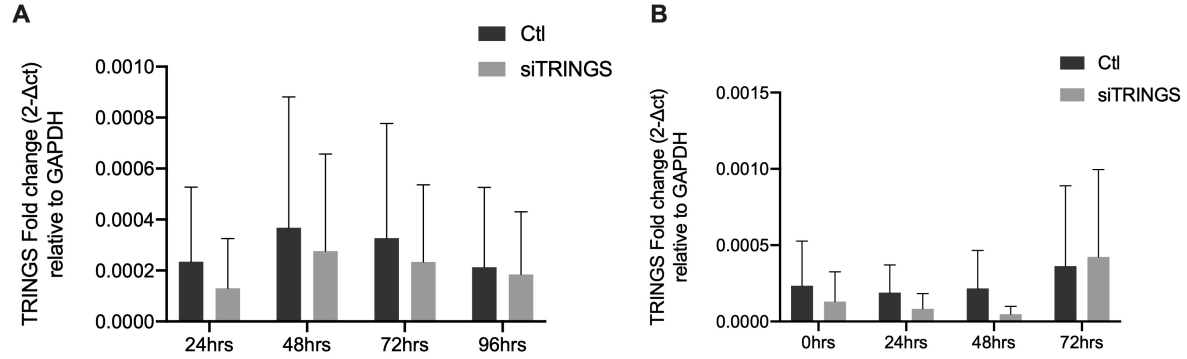
In our model (Figure 6.1), we describe TRINGS decrease in expression to be a result of E6 targeting and degrading p53. To further validate this, we should assess cells expressing the two viral oncogenes E6 and E7 individually. In cancer cells, E7 expression without E6 leads to a stabilisation of p53 levels [340, 341]. Hence in E7 alone expressing cells, we would also presume to see an increase in TRINGS expression. This would further confirm the notion that E6 degradation of p53 inhibits TRINGS expression in cervical cancers.

In this study, we have identified that the lncRNA TRINGS has altered expression induced by HPV16 E6/E7 in cervical cancer cells. We also demonstrated that the modulation of TRINGS will have an impact on cancer cell viability. Although in this chapter we focused on cervical cancers, it did reveal that HPV16 and its viral oncogenes can influence the expression of lncRNAs in cancer and that lncRNAs potentially play an important role in cancer cell pathways. We aimed to apply the same investigative methods in this chapter to the investigation into lncRNAs in HNC. Given that there was no significant change in TRINGS expression in HNC patient samples, we postulated that HPV16 might target a different cohort of lncRNAs in HNC compared to cervical cancers. Next, we wanted to explore the impact of HPV16 on non-coding RNA expression in Head and Neck cancers and determine possible cancer-related pathways these lncRNAs are associated with.

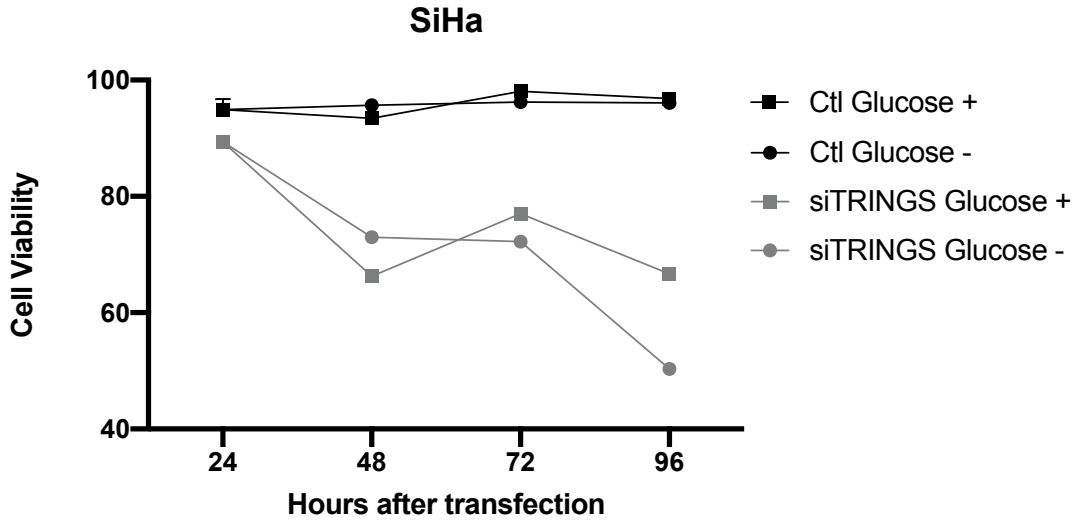
6.5. Appendix



Appendix Figure 6. 1: Knockdown of TRINGS using two siRNA in HPV16 positive cell line SiHa at 24hrs, 48hrs, 72hrs and 96hrs post transfection. A) Relative expression levels of TRINGS at different time points and B) Relative cell viability of knockdown cells over time. Normalisation of gene expression was achieved using the calibrator GAPDH and fold relative to the control (Ctl) cells using the $2^{-\Delta\Delta Ct}$ calculation. Cell viability was calculated relative to the Day 1 control cells. Statistical analysis was performed using two tailed student T-test ****p≤0.0001, ***p≤0.001, **p≤0.01 and *p≤0.05.



Appendix Figure 6. 2: Knockdown of TRINGS siRNA in HCT116 TP53 null cells at 24hrs, 48hrs, 72hrs and 96hrs post transfection under A) normal glucose conditions and B) glucose starved conditions. Normalisation of gene expression was achieved using the calibrator GAPDH and fold relative to the GAPDH using the $2^{-\Delta Ct}$ calculation. Statistical analysis was performed using two tailed student T-test.



Appendix Figure 6. 3: Trypan blue staining of SiHa knockdown of TRINGS siRNA at 24hrs, 48hrs, 72hrs and 96hrs post transfection under normal glucose conditions and glucose starved conditions. Cell viability is presented as a percentage of the number of dead cells. There is a significant decrease in cell viability in the siTRINGS cells compared to the control cells in both the normal glucose and glucose starved cells.

Chapter 7: The interaction between microRNAs and long non-coding RNAs in HPV16 related Oropharyngeal cancers

7.1. Introduction

The incidence of human papillomavirus (HPV) associated oropharyngeal cancers (OPC) has been rising, making it one of the most common HPV related malignancies in western countries [11, 23, 342]. HPV+ OPC are classified as their own distinct entity when compared to HPV- OPC [25]. The molecular pathways underpinning this disease still require further elucidation and given the expanding universe of non-coding RNA (ncRNA) families, it is paramount we begin to understand their contribution.

In HPV+ cancers, the research has primarily focused on the viral impact on microRNA (miRNA) expression. HPV can lead to the altered expression of numerous miRNAs in OPC [343]. Another form of ncRNA that can induce oncogenic phenotypes are the long non-coding RNAs (lncRNA). These are RNA transcripts greater than 200 nucleotides in length and have no coding potential. They can interact with RNA and DNA via base pairing or form secondary and tertiary structures for binding to proteins, providing numerous possibilities for their functionality [344]. The specific molecular mechanisms behind lncRNAs' action and regulation in HNC is highly complex and not fully understood.

LncRNAs and miRNAs are described as independent master regulators; however, increasing evidence suggests the two ncRNAs can interact with each other. One of the recently discovered functions of lncRNAs is that they can act as miRNA “sponges”. That is, a lncRNA can bind to a miRNA to prevent it from regulating its targets, making lncRNAs positive regulators of mRNA transcription. This type of lncRNA/miRNA interaction is also known as competitive endogenous RNA (ceRNA) network. The majority of lncRNAs capture miRNAs using regions close to the 3`end denoted the miRNA response element (MRE), this region is complementary with Ago binding sites [345]. This ceRNA network theory has since been proven in several different cancer types [346-348]. The interaction between lncRNAs and miRNAs forms a large-scale regulatory network across the transcriptome which has been described for HNC [349-352].

Few studies have investigated the impact of HPV on ceRNA networks in HNC, and these studies have grouped all HNC subtypes together [144, 189]. HPV is associated

with up to 50-70% of OPC [11, 33, 353] and up to 85% of these are infected with the high-risk variant, HPV type 16 [354]. Given this, we conducted a comprehensive analysis of lncRNA/miRNA interactions in this specific subtype, OPC, infected with the HPV16 variant.

This chapter analysed the RNA-sequencing data from The Cancer Genome Atlas (TCGA) to identify differentially expressed lncRNA, miRNAs, and mRNAs in HPV16+ OPC. We then explored the specific functions of these lncRNAs in relation to miRNA sponging and the impact of altered expression on the RNA axis, which includes: lncRNA/miRNA/mRNA.

7.2. Methods

7.2.1. TCGA RNA-seq data analysis

Raw RNA-sequencing data was downloaded from TCGA (https://www.cancer.gov/tcga) via the GDC data portal (https://portal.gdc.cancer.gov/). We retrieved RNA-seq BAM files of 49 patients, all of which were primary oropharyngeal tumours (tonsil, base of tongue, oropharynx) from the head and neck cancer cohort. The OPC cohorts was also filtered for those tested for HPV16 by ISH, or identified to contain HPV16 by RNA sequencing [209]. In this study, 34 samples were identified as HPV16+ and 15 as HPV-. Samples were annotated according to their Barcode ID and available clinical information, including tumour site, sex, HPV status, age and clinical state, Appendix Table 7.1. MiRNA-sequencing data was downloaded for the 49 patients as raw count txt files from the TCGA. Total RNA-seq data were aligned to the human genome sequences (GRCh38p13) from GENCODE (release 38).

To perform a comprehensive survey of human lncRNAs, we obtained the genomic coordinates of 17958 human lncRNAs from GENCODE Resource (release 38) and extracted the raw RNA-seq counts for these lncRNAs. Differential expression analysis of individual mRNAs, lncRNAs and miRNAs was performed using the DESeq2 Bioconductor package (https://bioconductor.org/packages/release/bioc/html/DESeq2.html).

Raw counts were extracted from the OPC samples, and DESeq2 was employed to normalise the counts and identify differentially expressed mRNA, lncRNAs and miRNAs between HPV16+ and HPV- OPC. Only transcripts (mRNA, miRNA, lncRNA) with $\geq$10 counts across all samples and those with an adjusted p-value $\leq$0.05 were included.
A 5-fold change cut off was applied to the lncRNAs and mRNAs, and a 2-fold change cut off was applied to the miRNAs. The methodology for TCGA RNA-sequencing data analysis is described in Figure 7.1.

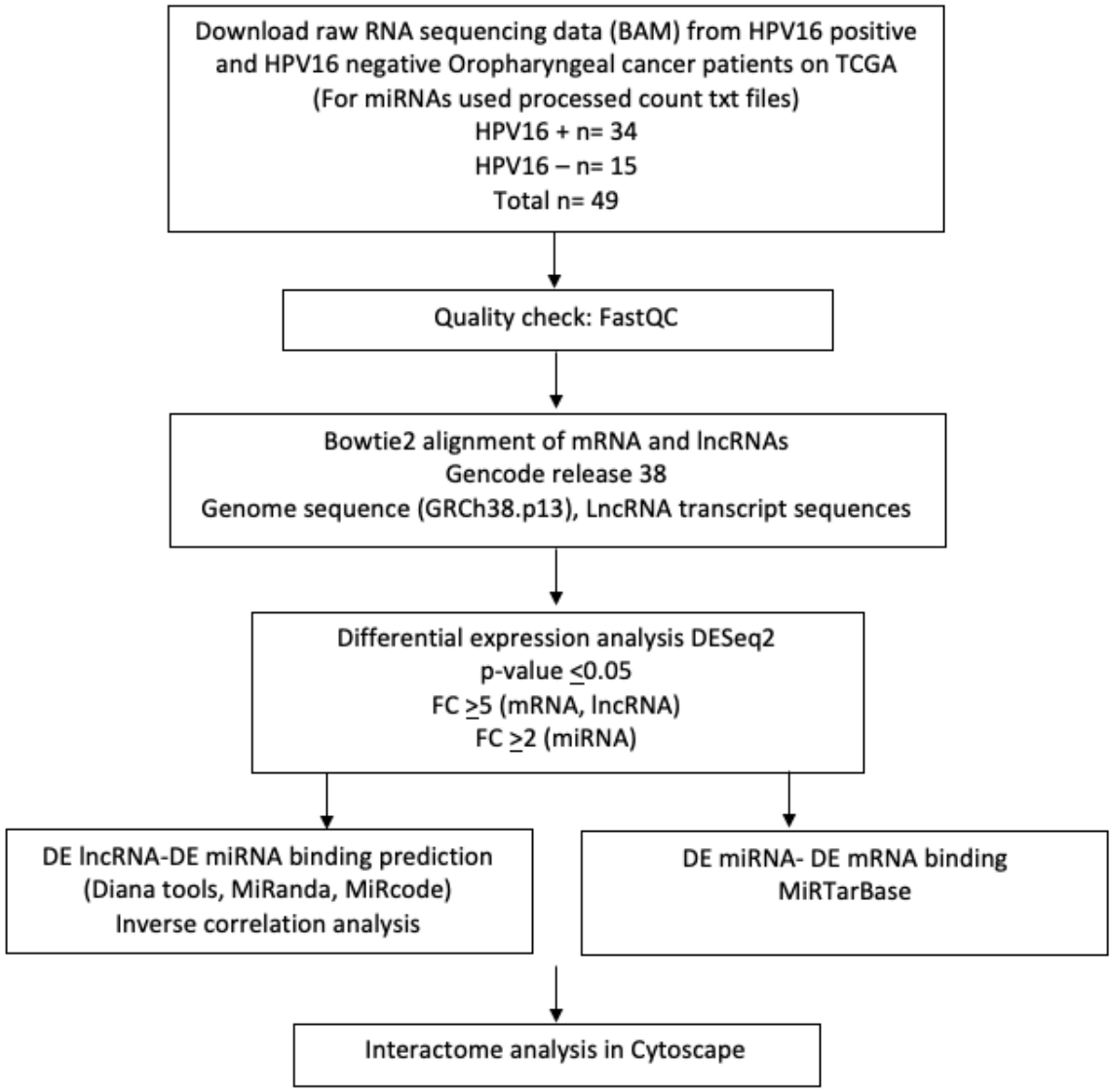


Figure 7. 1: Flow diagram representing the methodology for processing RAW RNA-seq data from TCGA and development of lncRNA-miRNA-mRNA interactome

7.2.2. LncRNA-miRNA sponging prediction and HPV16 non-coding RNA interactome analysis

In order to identify predictive lncRNA-miRNA sponging interactions three online tools were utilised, Diana tools LncRNA prediction tool v2 (https://carolina.imis.athena-innovation.gr/diana_tools/web/index.php?r=lncbasev2%2Findex-predicted), using the MiRanda parameters on miRNAconsTarget (https://arn.ugr.es/srnatoolbox/mirconstarget) and the mirCode database (http://www.mircode.org/). The miRNA-lncRNA interactions found using miRanda were filtered according to a score of ≥150 and a minimum free energy less than -20Cal/mol to identify reliable targets [355].

Each of our differentially expressed lncRNAs (FC≥5) were analysed using each of these databases or tools, and their miRNAs targets were identified. The lncRNA-miRNA targets were filtered to only include those that were differentially expressed (miRNAs with an FC≥2) in our TCGA RNA-sequencing dataset. The lncRNA-miRNA targets were filtered again to include miRNAs with an inverse differential expression to the lncRNAs. The mRNA targets of these miRNAs were then identified using MiRTarBase [356], this database describes experimentally validated miRNA-target interactions. The DE miRNAs-mRNA targets were filtered for inverse DE patterns in the TCGA RNA-seq dataset (mRNAs with an FC ≥2). The DE lncRNAs and their DE miRNA targets were imported into the network mapping program Cytoscape v3.9. The DE miRNA and their DE mRNA targets were added to this interactome for visualisation and network analysis.

7.2.3. Gene Ontology and KEGG analysis

To gain a biological understanding of those gene sets statistically significantly associated with HPV16+ OPC, we carried out pathway enrichment analysis and disease association analysis using various programs. The BiNGO [357] Cytoscape software was utilised for gene ontology analysis and g:Profiler (https://biit.cs.ut.ee/gprofiler/gost) for KEGG analysis.

7.2.4. Cell culture and transfections

The HPV16+ cells SCC090 were maintained in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% foetal calf serum in a 37°C incubator with humidified 5% CO2. All transfections were performed using Lipofectamine™ RNAiMAX (Invitrogen) in a 12-well plate format in duplicate, according to the manufacturer's instructions. siRNA targeting HPV16 E6/E7 (IDT) were forward transfected at 20pmol into $5x10^5$ cells. The sequences for the siRNA are described below in Table 7.1:

Table 7. 1: Sequences for siRNA used in this study

Target/name	Sequence
siHPV16 E6/E7 1	Sequence 1: 5`GAGCUGCAAACAACUAUACAUGATA`3 Sequence 2: 5`UAUCAUGUAUAGUUGUUUGCAGCUCUG`3
siHPV16 E6/E7 2	Sequence 1: 5`GAGAUACACCUACAUUGCAUGAATA`3, Sequence 2: 5`UAUUCAUGCAAUGUAGGUGUAUCUCCA`3

Transfection from replicate wells were used for analysis, and RNA was harvested 48hr post-transfection. Total RNA isolation was performed according to the RNAzol RT manufactures protocol (Molecular Research Centre). One sample from each condition (siHPV16 E6/E7 1, siHPV16 E6/E7 2 and control) from two biological replicates was used in RNA-sequencing.

7.2.5. Quantitative-PCR (qPCR) and statistical analysis

The expression of individual genes was determined via qPCR. In brief, cDNA was prepared using random primers to 500ng of total RNA. Complementary DNA was generated with the Hi-Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific). The cDNA was diluted in a 1:4 ratio for a 5.0µL qPCR reaction. All qPCR reactions (including non-RT and no-template controls) were conducted in triplicate on a 96-well plate using the StepOne PCR systems (Applied BioSystems) at the recommended cycle conditions using the TaqMan Universal PCR Master Mix, No AmpErase UNG (2x) (Thermo Fisher Scientific). TaqMan assays were employed to determine the expression of the target genes HPV16 E6 and HPV16 E7 (custom

design), and the endogenous reference gene for these assays was β2M (Hs00187842_m1) labelled with a VIC^{TM} probe. Mean Ct values were normalised using β2M. The relative expression level was calculated using $2^{-\Delta\Delta CT}$ [208] and presented as fold change. The significant difference between gene expression levels was evaluated by two tailed Student T-test.

7.2.6. RNA sequencing of HPV16 E6/E7 knockdown

Total RNA (1ug) the siHPV6 E7/E7 knockdown were sent to Ramaciotti Centre for Genomics UNSW Sydney for total RNA sequencing via Illumina HiSeq at a read depth of 120M total reads. Figure 7.2 depicts the processing of the raw RNA sequencing data. In summary, total RNA-seq data was aligned to the human genome sequences (GRCh38p13) from GENCODE (release 38), and lncRNAs were extracted using the lncRNA transcript coordinates from GENCODE (release 38). Differential expression analysis of individual lncRNAs was carried out using the DESeq2 Bioconductor package. Raw counts were extracted from the samples, and DESeq2 was employed to normalise the counts and identify differentially expressed lncRNAs between HPV16 E6/E7 knockdown cells to control cells. Only transcripts with $\geq$10 counts across all samples and those with a p-value $\leq$0.05 were included. Due to variation between biological replicates (Appendix Figure 7.1), the p-value was used instead of the adjusted p-value for DESeq2 analysis.

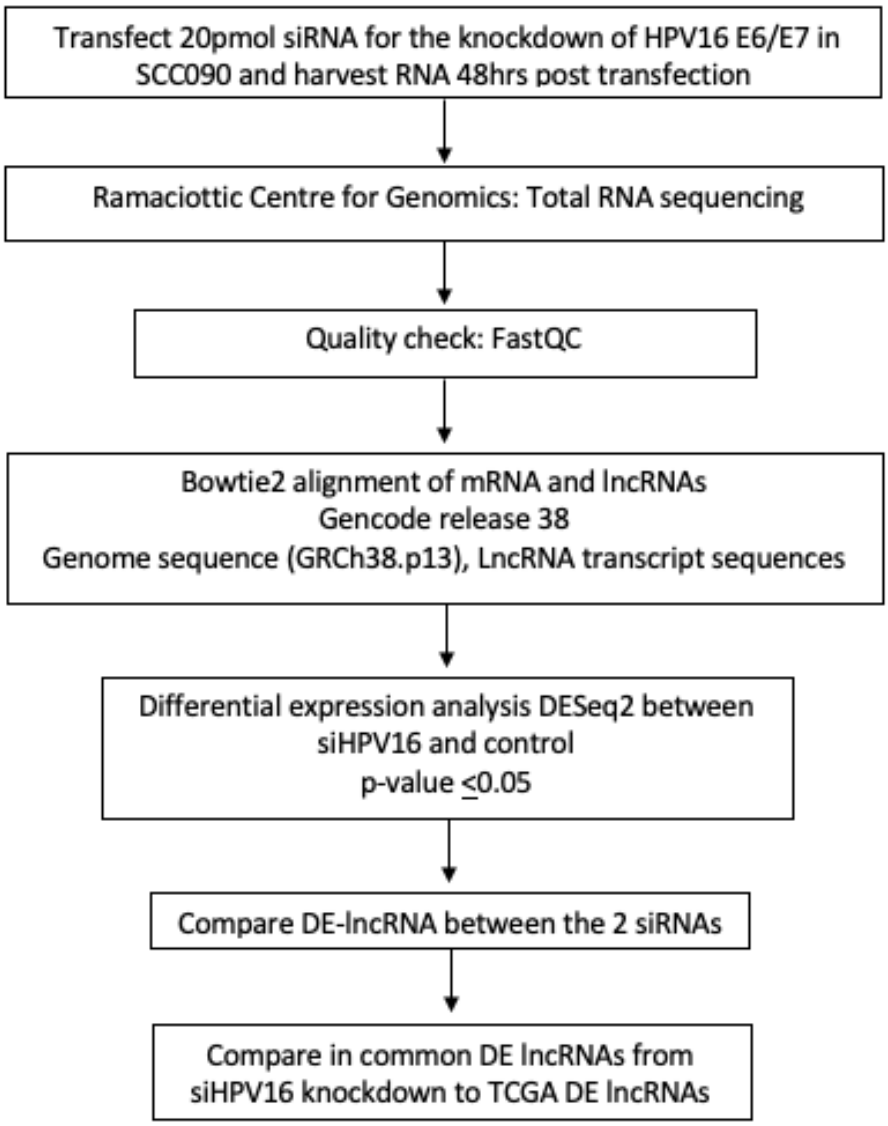


Figure 7. 2: Flow chart representing the processing and analysis of RNA sequencing data from HPV16 E6/E7 knockdown.

7.3. Results

7.3.1. Identifying differentially expressed lncRNAs, miRNAs and mRNAs between HPV16+ and HPV16- Oropharyngeal cancers

First, we utilised 49 OPC samples in the TCGA RNA sequencing database generated by Illumina HiSeq 2000. These samples were filtered for HPV16+ (n=34) or HPV- (n=15) status. Differential expression analysis was performed with Deseq2, lncRNA profiles revealed 1929 transcripts differentially expressed between the HPV16+ and the negative samples. Of these, 1176 were upregulated, and 753 were downregulated in the HPV16+ tumours (p-value of 0.05 or less Appendix table 7.2, Figure 7.3A and B).

To identify the most significant lncRNA transcripts in HPV associated OPC, we applied a 5-fold change cut off. This showed 37 lncRNAs transcripts that were differentially expressed in the HPV16+ compared to the HPV- OPC. Of these, 23 were upregulated, and 14 were downregulated, Figure 7.3C and described in Appendix table 7.2. Clustering revealed a distinct separation between the HPV16+ and negative cohorts, indicating the DE lncRNAs are distinct between the groups. The most significantly upregulated lncRNA was the novel-ENSG00000282816.1 (6.5 log2 FC). In the comparison, lncRNA novel-ENSG00000231683.6 was the most significantly downregulated (-7.4 log2 FC). Out of the 37 most differentially expressed lncRNAs, there were ten antisense lncRNAs, thirteen long intergenic non-coding RNAs and 14 novel transcripts remain uncharacterised.

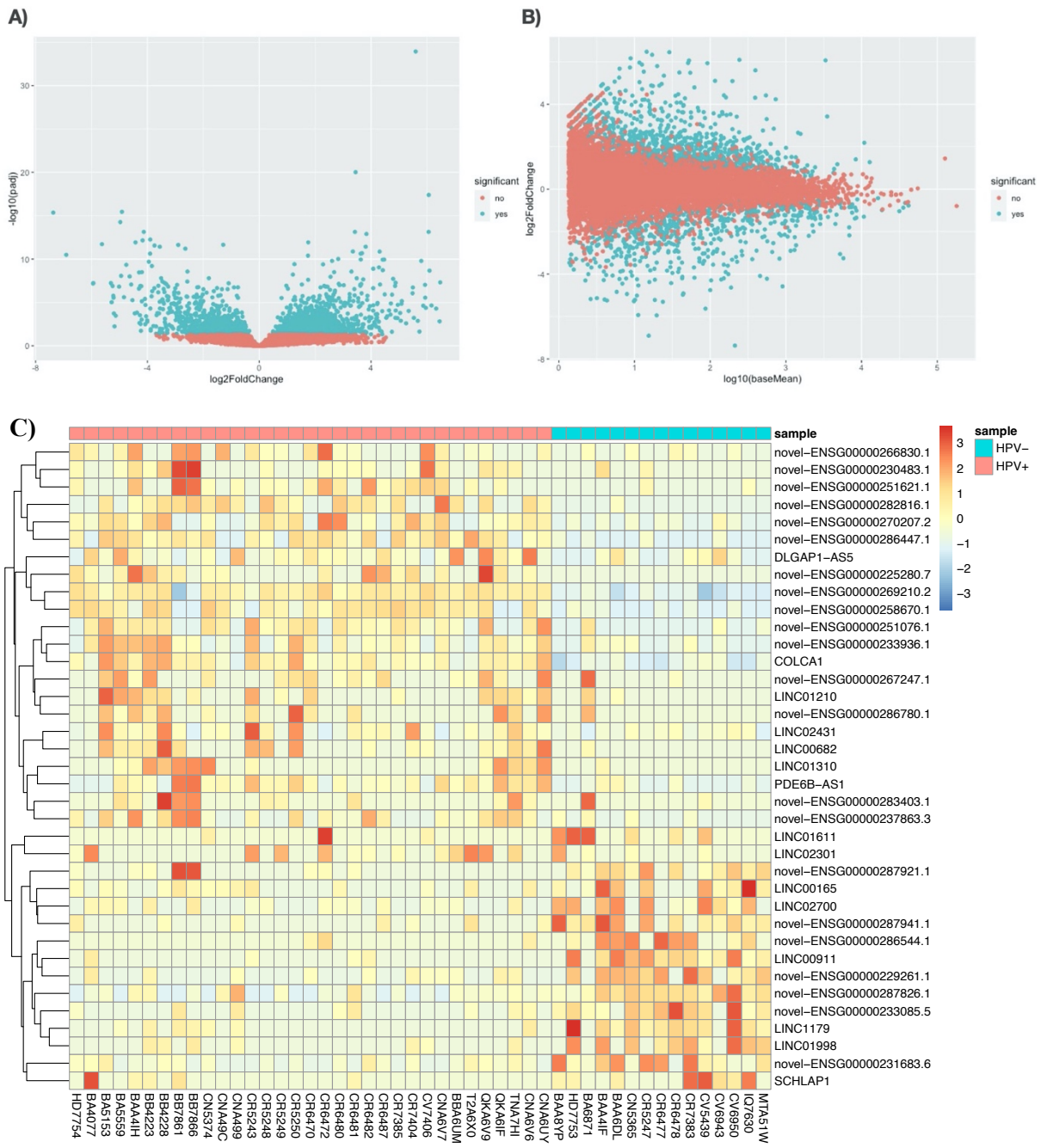


Figure 7. 3: Differentially expressed lncRNAs between HPV16+ and HPV- OPC patients. A) Volcano plot showing log2Fold change for specific lncRNAs B) MA plot representing the change in lncRNAs comparing Log2 Fold change on the Y axis and Log10 of the mean average normalised counts. C) Heat map showing the differentially expressed lncRNAs between HPV+ (pink bar) and HPV- (blue bar). LncRNAs shown in red are upregulated, whilst lncRNAs in blue are downregulated. The lncRNA name/Ensembl ID are represented as rows and the patient samples as columns with the TCGA patient identifier on the bottom. This heatmap represents all lncRNAs with a fold change greater than or equal to 5-fold.

For each of the samples the expression profiles of miRNAs and mRNAs were also assessed comparing HPV+ to HPV- OPC. A total of 161 miRNAs (Figure 7.4A and B) and 9836 mRNAs (Figure 7.5A and B) were identified to be differentially expressed in the HPV+ tumours compared to HPV-. When applying a 2-Fold change cut off for miRNAs, this resulted in 32 differentially expressed miRNAs (Figure 7.4C). Of these, 14 miRNAs were upregulated and 18 were downregulated. The most significantly upregulated miRNAs were miR-125b (3.9 log2FC), miR-6510 (3.8 log2FC) miR-20b (3.5 log2FC) and miR-499a (3.4 log2FC). The miRNAs miR-516a-1 (-4.6 log2FC), miR-116a-2 (-4.4 log2FC), miR-519a-1 (-4.0 log2FC) and miR-522 (-3.0 log2FC) were the most significantly downregulated.

Of the downregulated miRNAs, a large number are a part of the miRNA cluster on chromosome 19 (C19MC), one of the largest miRNAs clusters containing 46 miRNAs in total. These miRNAs include miR-520b, miR-518e, miR-517b, miR-517a, miR-519a-2, miR-518a-2, miR-518b, miR-520f, miR-522, miR-519a-1, miR-516a-1, miR-516a-2.

To identify the most significant target mRNAs that HPV16 may regulate, a 5-fold change cut off was applied and this resulted in 60 differentially expressed mRNAs (Figure 7.5C). A total of 39 mRNAs were upregulated and 21 mRNAs were downregulated. The most significantly upregulate mRNAs were FEZF2(22.4 log2FC), GBX1 (8.5 log2FC), NKX2-4 (8.1 log2FC) and ZPBP2 (7.9 log2FC). The most significantly downregulated mRNAs were PRR9 (-7.1 log2FC), FOLR3 (-6.6 log2FC), CNHB1 (-6.2 log2 FC) and LCE1A (-6.1 long2FC).

Of the mRNAs one had an antisense lncRNA that were also differentially expressed in the HPV+ OPC, ENSG00000225280.7 is antisense to NKX-2-4, both of which were upregulated.

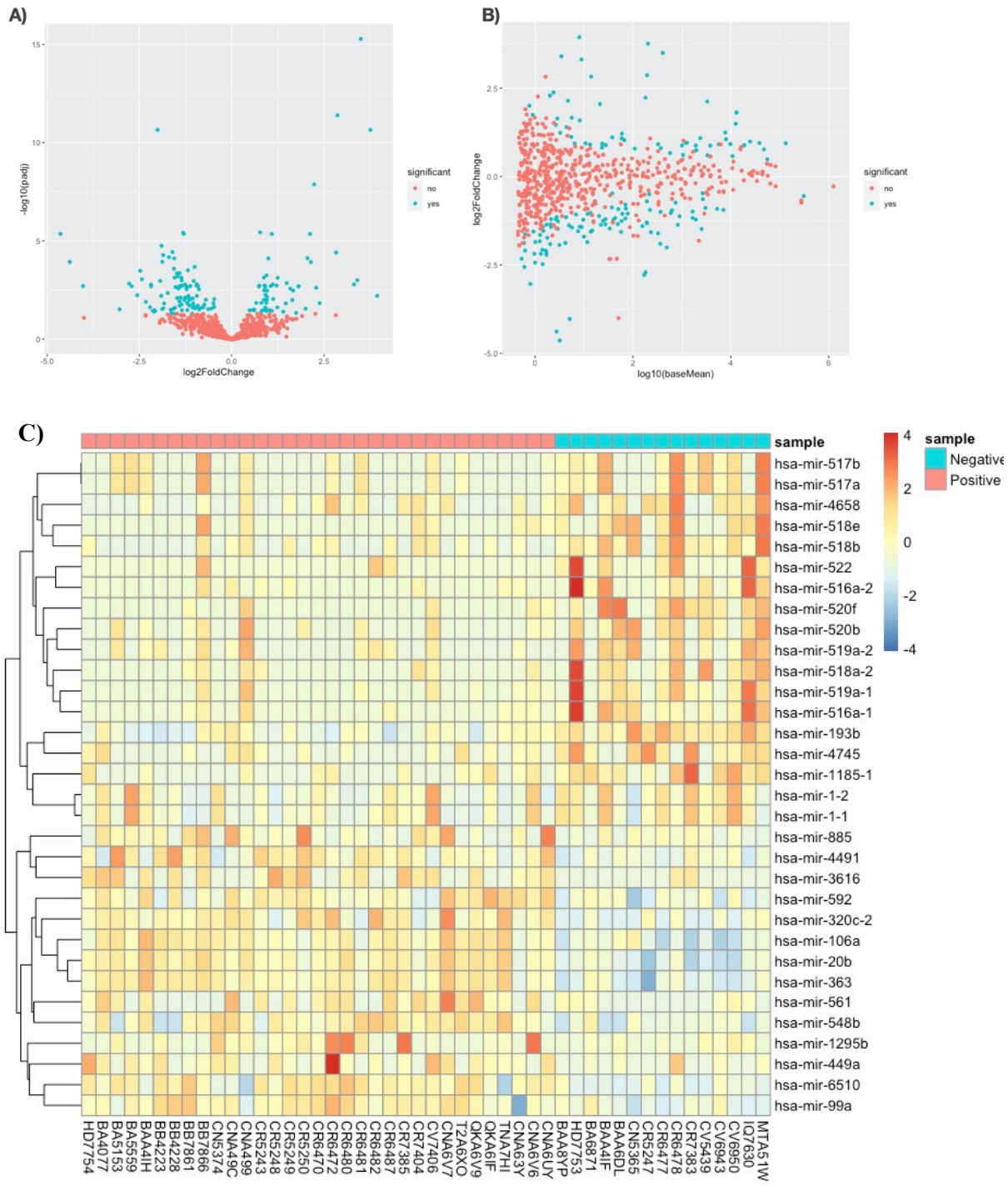


Figure 7. 4: Differentially expressed miRNAs between HPV16+ and HPV- OPC patients. A) Volcano plot shows the difference in log2Fold change on miRNAs. B) MA plot representing the differences between miRNAs expression comparing Log2 Fold change on the Y axis and Log10 of the mean average normalised counts. C) Heat map displaying the differentially expressed miRNAs between HPV16+ (pink bar) and HPV- (blue bar) oropharyngeal tumours. This heatmap represents all miRNAs with a fold change greater than or equal to 2-fold. miRNAs shown in red are upregulated, whilst miRNAs in blue are downregulated. TCGA Patient IDs are represented along the bottom of the heatmap.

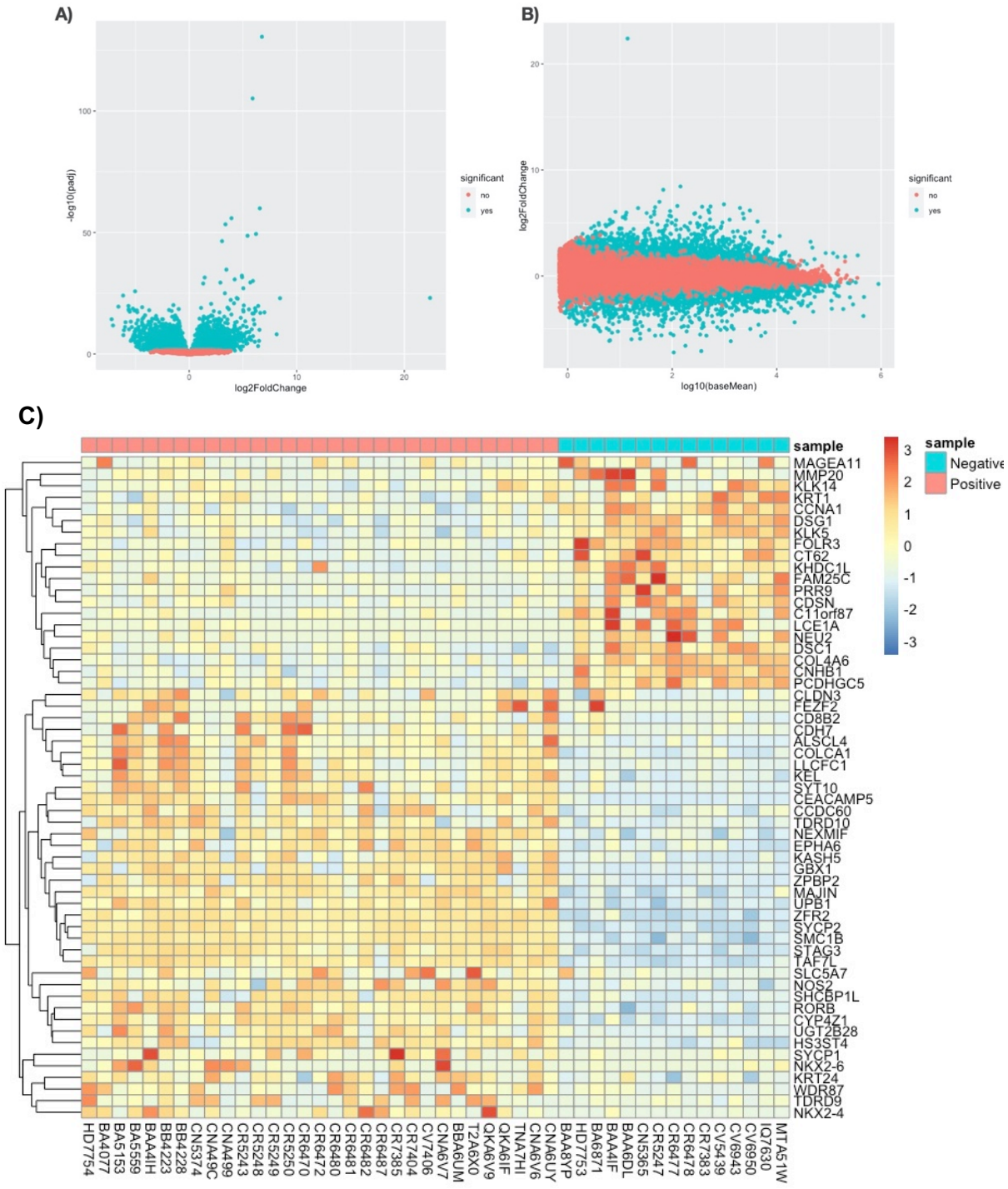


Figure 7. 5: Differentially expressed mRNAs between HPV16+ and HPV negative OPC patients. A) Volcano plot shows the difference in log2Fold change on mRNAs. B) MA plot representing the differences between mRNAs expression in HPV16+ and HPV negative OPC patients, comparing Log2 Fold change on the Y axis and Log10 of the mean average normalised counts. C) Heat map displaying the differentially expressed mRNAs between HPV16+ (pink bar) and HPV- (blue bar) oropharyngeal tumours. This heatmap represents all mRNAs with a fold change greater than or equal to 5-fold. miRNAs shown in red are upregulated, whilst mRNAs in blue are downregulated. TCGA Patient IDs are represented along the bottom of the heatmap.

7.3.2. Interactome analysis of differentially expressed lncRNAs-miRNAs-mRNAs in HPV16 OPC

Next, we wanted to determine if these differentially expressed lncRNAs had the potential to sponge miRNAs in HPV16+ OPC. In order to do this, we utilised the online tools (Diana LncBase Prediction tool v2, miRanda, MirCode) to predict miRNA binding sites on each of our lncRNAs. LncRNA-miRNA targets were filtered for those miRNAs that also had differential expression (FC≥2) in HPV16+ OPC and filtered again for those with the inverse expression pattern to the binding DE-lncRNAs. These DE lncRNA-mRNA interactions are described in Appendix Table 7.4. The DE lncRNAs with their DE miRNA targets were imported into the network mapping program Cytoscape software v3.3. Upregulated lncRNA (Figure 7.6) and Downregulated lncRNAs (Figure 7.7) were depicted in separate networks.

From the 37 DE-lncRNAs, novel-ENSG00000283403 had the most DE-miRNA targets (12 upregulated and 16 downregulated with a two-fold cut-off). This novel transcript is approximately 3113 bp in length, one of the largest transcripts from the DE lncRNAs, perhaps why it contains the most miRNA bindings sites. When this was filtered for lncRNA-miRNA interactions with inverse expression, the novel-ENSG00000287941 (downregulated) lncRNA could potentially sponged the most miRNAs (13 upregulated miRNAs) and ENSG00000283403 (upregulated) sponged the most downregulated miRNAs (16 miRNAs). Of the 32 most differentially expressed miRNAs (FC≥2) hsa-miR-6510-5p was most “sponged”, targeted by 25 different lncRNAs. Of the inverse lncRNA-miRNA hsa-miR-6510-5p (upregulated) was sponged by the most downregulated lncRNAs (n=13) and hsa-miR-4745-5p (downregulated) was sponged by the most upregulated lncRNAs (n=13).

The mRNA targets of these miRNAs were then identified using MiRTarBase [356], this database describes experimentally validated miRNA-target interactions. The DE miRNAs, mRNA targets were filtered for those with an inverse DE pattern in the TCGA RNA-seq dataset (mRNAs with a FC ≥2) (Appendix table 7.2). The DE miRNA, DE mRNA targets were also added to this interactome for visualisation and network analysis. Figure 7.6 describes the lncRNAs that were upregulated in HPV16+ OPC, their sponging interaction if downregulated miRNAs and the binding of miRNAs to their

mRNA targets (expected to be downregulated due to binding of the miRNA). The second interactome (Figure 7.7) depicts upregulated lncRNAs, their binding to downregulated miRNAs and those miRNA mRNA targets that are now upregulated in HPV16 OPC potentially due to lncRNA-miRNA sponging interactions.

From the interactome, hsa-miR-106a-5p (upregulated) followed by hsa-miR-20b-5p (upregulated) had the most mRNA targets that were also downregulated in the HPV+ OPC. Of the downregulated miRNAs hsa-miR-4658 and hsa-miR-516a-3p had the most upregulated mRNAs targets in the HPV+ OPC. CDK6 (downregulated) was targeted by the most upregulated miRNAs (miR-20b-5p, miR-4491, miR-499a and miR-561-5p) and the two upregulated mRNAs NCAN (miR-518e-5p, miR-519a-5p, miR-522-5p) and BEND4 (miR-518e-5p, miR-519a-5p, miR-522-5p) are targeted by the most downregulated miRNA.

Of the DE miRNA sponge, miR-517b-3p, miR-517a-3p, miR-516a-5p, miR-518b, miR-520b-5p, miR-519a-2-5p, miR-1-5p (all downregulated), miR-885-5p, miR-3616-5p, miR-320c and miR-592 (all upregulated miRNAs) did not have any mRNA targets that were also inversely differentially expressed in the TCGA cohort.

All mRNAs from Figures 7.6 and 7.7 were searched in the database VirusMentha (https://virusmentha.uniroma2.it/index.php), to determine if any HPV16 viral genes directly target these mRNAs. Only two mRNAs, SLC7A5 (miR-4757-5p target) and CAV1 (miR-106b-5p and miR-20b-5p target) are current known targets of the viral gene E5. The main viral genes HPV16 E6/E7 do not directly regulate any on the sponged miRNA-mRNA targets from out interactomes. The altered expression of these mRNAs is perhaps due to the lncRNA/miRNA/mRNA interactions.

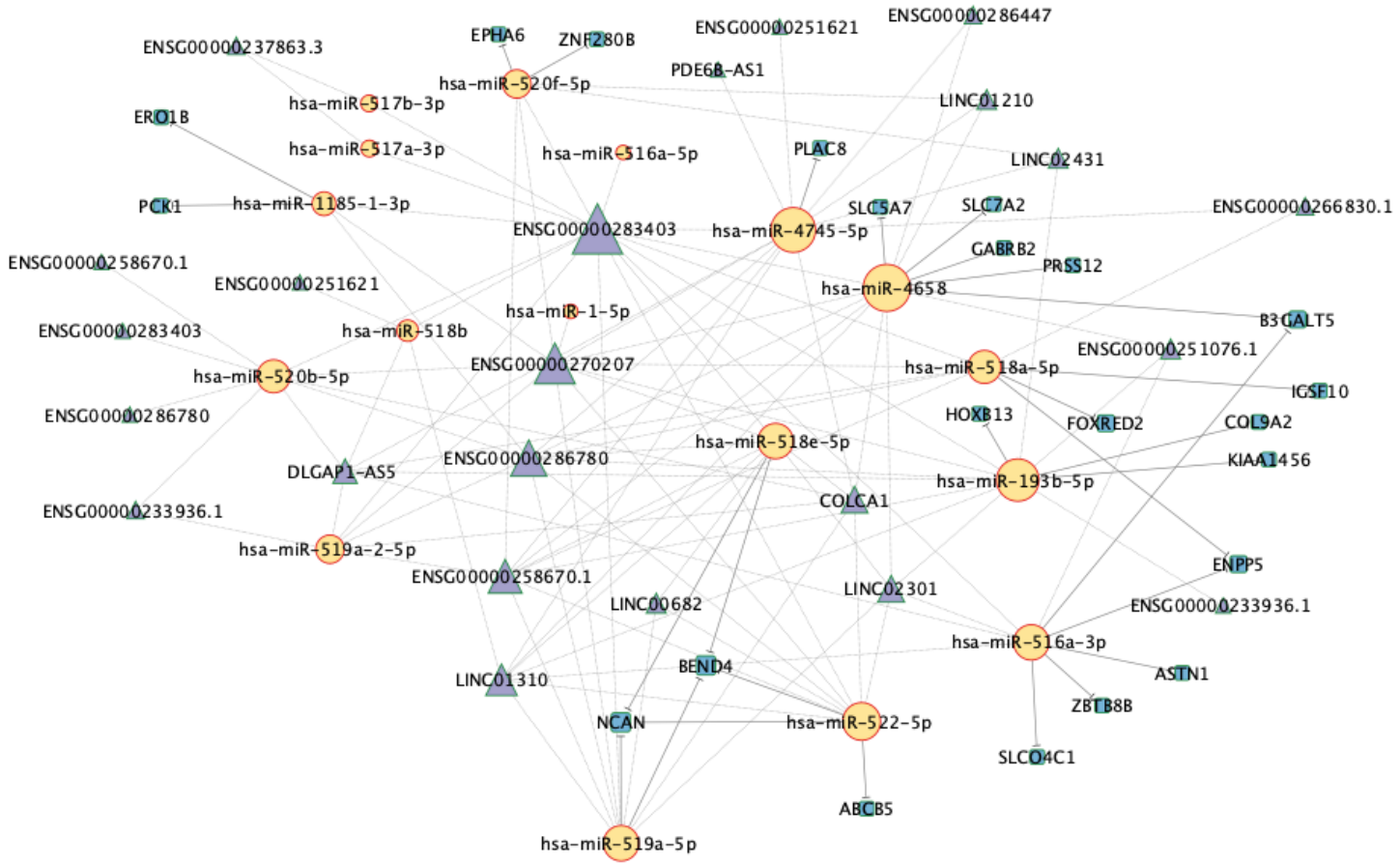


Figure 7. 6: HPV16 upregulated lncRNA and the predicted miRNAs they sponge: This interactome depicts the differentially expressed lncRNA that were upregulated in HPV16+ oropharyngeal cancers compared to HPV- tumours. The lncRNAs are linked to their predicted target miRNAs, these miRNAs are downregulated in HPV16+ oropharyngeal cancers. The miRNA known mRNA targets, that are upregulated in HPV16+ oropharyngeal cancers, are also depicted. Triangles represent lncRNAs, circles represent miRNAs and squares represent mRNA. Green node borders represent upregulated expression and red represent downregulated expression pattern. The size of the nodes indicates number of interactions each node has, the larger the size the more interactions.

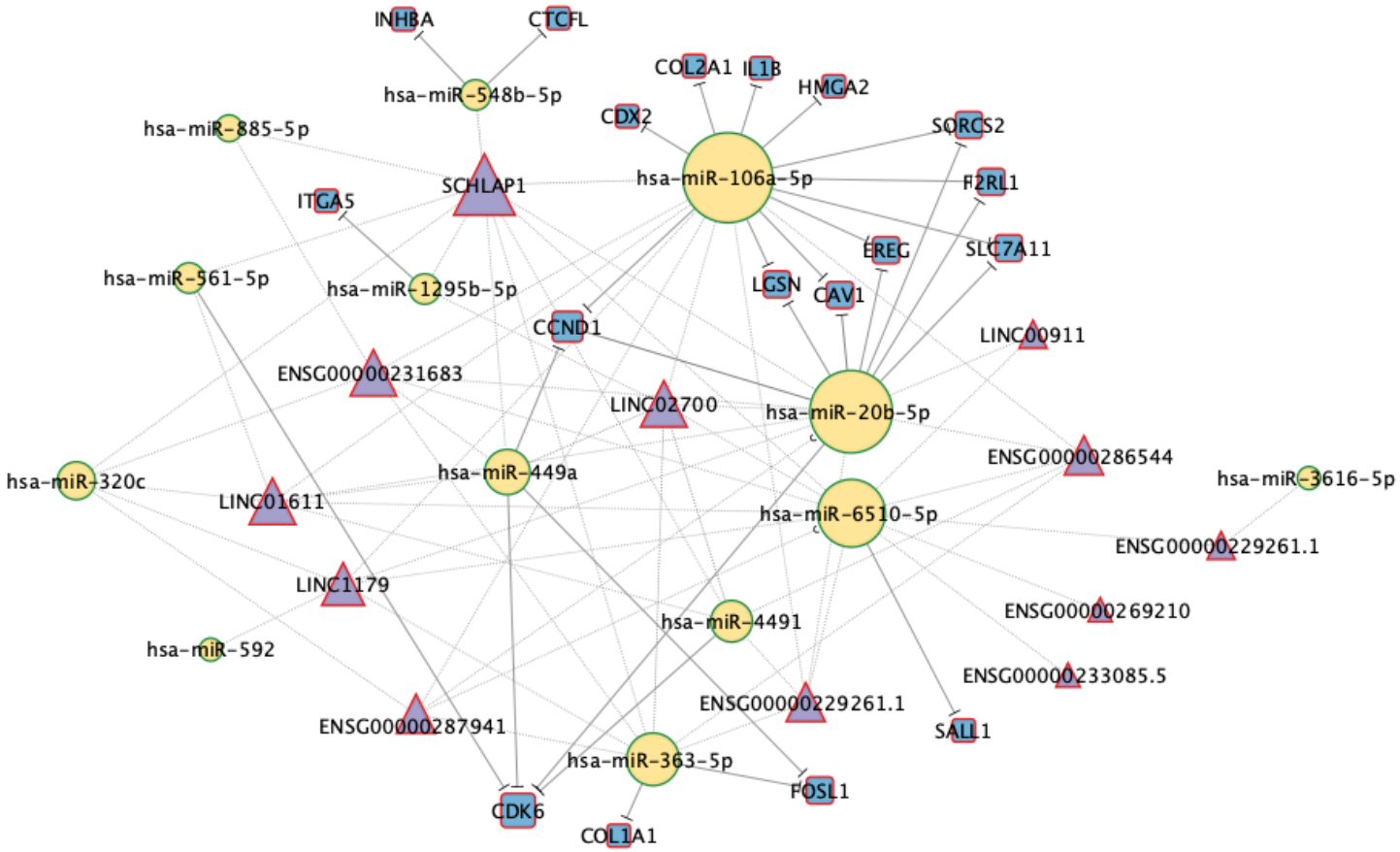
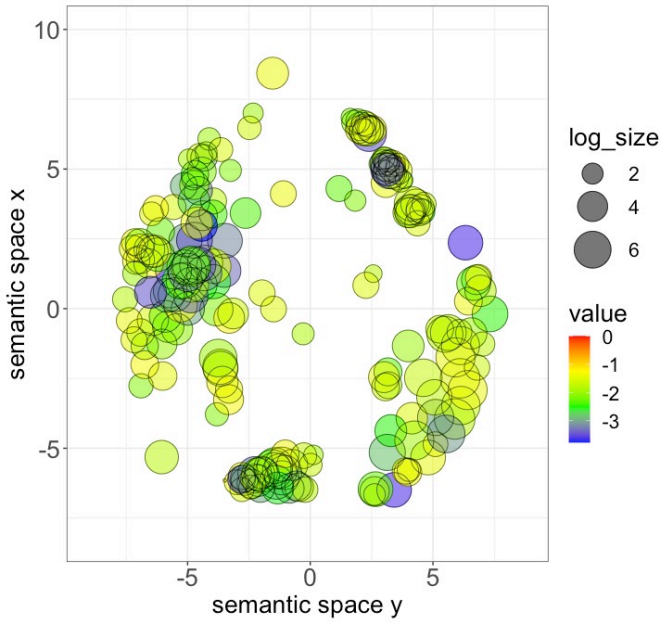


Figure 7. 7: HPV16 down regulated lncRNA and the predicted miRNAs they sponge: This interactome depicts the differentially expressed lncRNA that were down regulated in HPV16+ oropharyngeal cancers compared to HPV- tumours. The lncRNAs are linked to their predicted target miRNAs, these miRNAs are upregulated in HPV16+ oropharyngeal cancers. The miRNA known mRNA targets, that are down regulated in HPV16+ oropharyngeal cancers, are also depicted. Triangles represent lncRNAs, circles represent miRNAs and squares represent mRNA. Green node borders represent upregulated expression and red represent downregulated expression pattern. The size of the nodes indicates number of interactions each node has, the larger the size the more interactions.

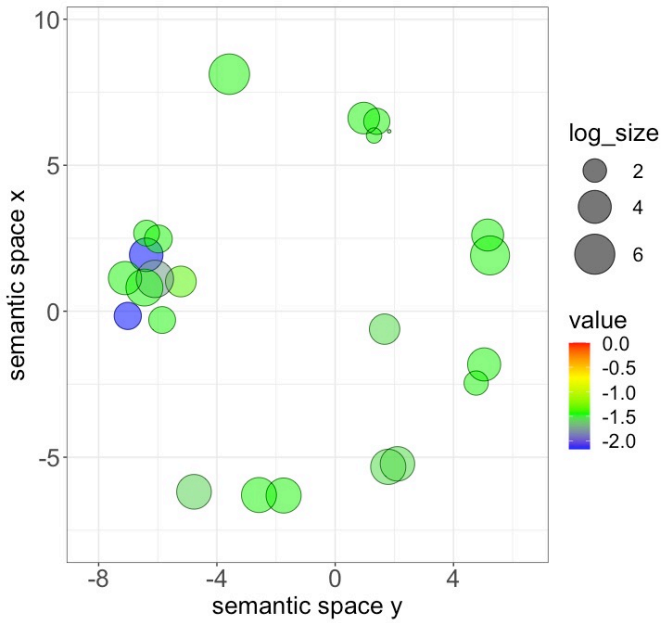
To gauge the impact of these interactions, each network was assessed for Gene Ontology enrichment analysis for the protein coding genes that were regulated by the differentially expressed miRNAs. This was performed using BINGO and plotted in Revigo. For the downregulated lncRNAs interactome (Figure 7.7), the DE mRNAs were significantly associated to “positive regulation of cell cycle”, "negative regulation of cellular process", "reproductive process", “reproduction” and "negative regulation of cell population proliferation" under the biological process category (Figure 7.8a). For molecular processes the DE-mRNAs were significantly associated with "platelet-derived growth factor binding", "growth factor receptor binding" and "signalling receptor binding" (Figure 7.8b). The DE mRNAs were also significantly associated to “fibrillar collagen trimer”, “cyclin-dependent protein kinase holoenzyme complex" and “collagen trimer" under cellular component. For the upregulated lncRNAs interactome (Figure 7.6) there was only one GO biological process that was significantly associated with the mRNA targets, "neuron cell-cell adhesion".

KEGG pathway enrichment analysis of the downregulated lncRNA interactome (Figure 7.7) suggests that those DE-mRNA are significantly linked to “focal adhesion”, “PI3K-Akt signalling pathway”, “Human papillomavirus infection”, “Proteoglycans in cancer” and “ECM-receptor interaction” (Figure 7.8D). Again, the upregulated lncRNA interactome (Figure 7.6) did not reveal any significant association to KEGG pathways. The lncRNA/miRNA/mRNA interactions from this network and their involvement in cellular pathways need to be analysed individually, the GO and KEGG analysis revealed these lncRNA/miRNAs/mRNAs don’t collectively work together in one cellular pathway.

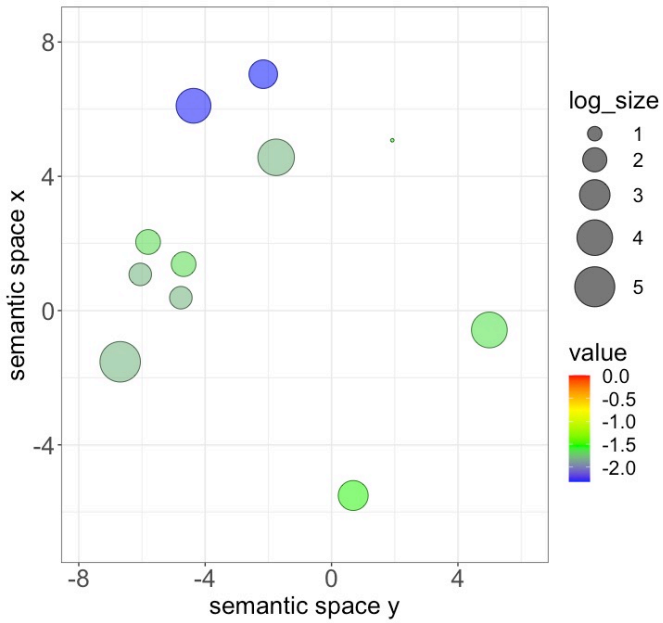
A)


semantic space x
semantic space y
log_size
2
4
6
value
0
-1
-2
-3

B)


semantic space x
semantic space y
log_size
2
4
6
value
0.0
-0.5
-1.0
-1.5
-2.0

C)


semantic space x
semantic space y
log_size
1
2
3
4
5
value
0.0
-0.5
-1.0
-1.5
-2.0

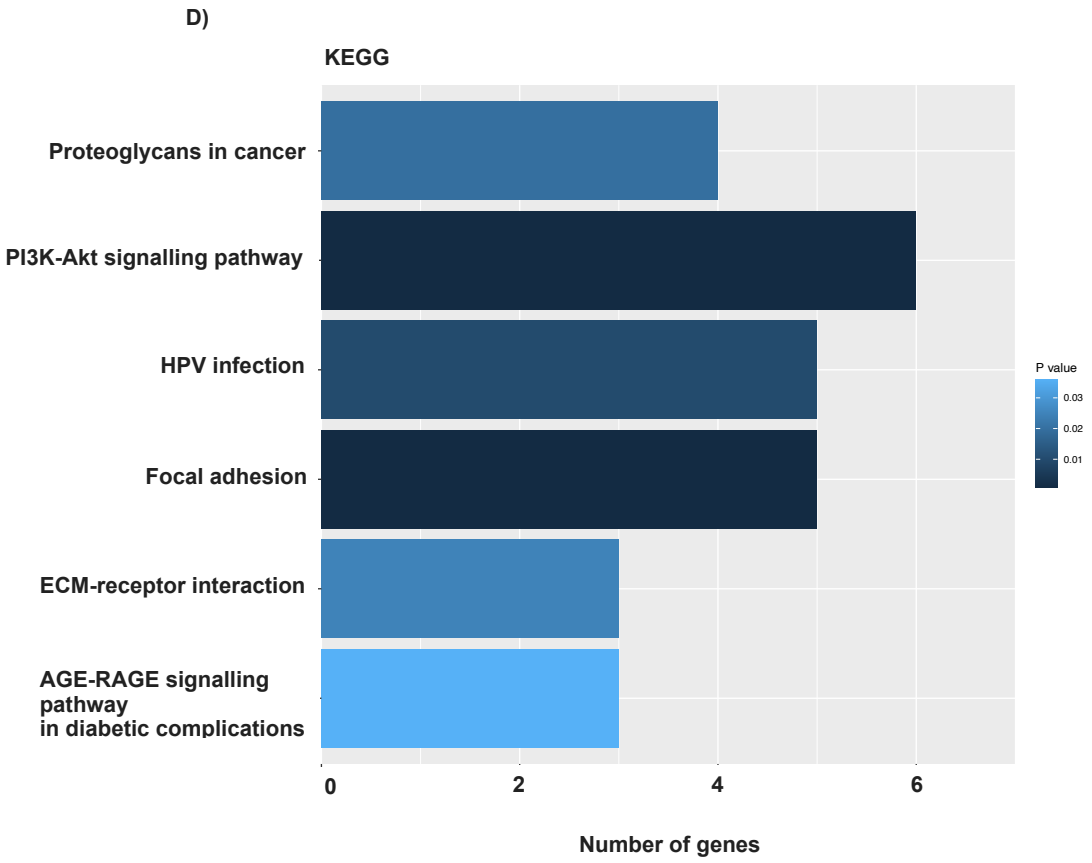


Figure 7. 8 Gene ontology and KEGG analysis of mRNAs targeted by upregulated miRNA sponged by downregulated lncRNAs in HPV16+ Oropharyngeal cancers: A) GO terms under Biological process, B) GO terms under Molecular Function, C) GO terms under cellular component, D) KEGG terms.

7.3.3. Knockdown of HPV16 viral oncogenes E6/E7 reveal lncRNAs to be targeted in OPC

From the TCGA, we have suggested that lncRNAs do have a differential expression pattern between HPV16+ and HPV- OPC tumours. We then sought to understand if the lncRNA expression alteration was due to the two major viral oncogenes HPV16 E6/E7.

Towards this end, we designed two siRNAs for the knockdown of HPV16 E6/E7 and transfected these into the oropharyngeal cancer cell line, SCC090 (Figure 7.9). Both HPV16 E6/E7 oncogenes were reduced by 90% using siHPV16 E6/E7 1 and 60-80% using siHPV16 E6/E7 2, 48hrs post transfection. The RNA from these cells were sent for RNA sequencing and the raw data sets (reads) were aligned to the lncRNA transcript genome using the same pipeline as the TCGA data. Differential expression analysis was then conducted using Deseq2.

A PCA plot was created to determine the variability between biological replicates (Appendix figure 7.1). Separate biological replicates were used to account for inconsistencies that might been seen across cell culture cohorts. In our RNA sequencing data, we did observe that our two biological replicates were highly variable and represented two different cohorts. This variability may be due to the expression of lncRNA being highly specific and sensitive to the conditions of the cells at the time of endpoint collection. Due to this variability between biological replicates, the p-adjusted values were high and showed no change in any lncRNAs expression. For the DeSeq2 analysis we decided to use the unadjusted p-value for statistical analysis, this is not as accurate as adjusted p-value but may still provide some information.

To perform this analysis, each knockdown (siRNA HPV16 E6/E7 1 and siRNA HPV16 E6/E7 2) were assessed for differentially expressed lncRNAs individually (Appendix Figure 7.2) and lncRNA in common between the two siRNAs were extracted (Figure 7.10). Individually siHPV16 E6/E7 1 resulted in 64 DE lncRNAs (23 downregulated and 41 upregulated) and siHPV16 E6/E7 2 resulted in 38 DE lncRNAs (18 downregulated and 20 upregulated).

When comparing the DE-lncRNAs between the two HPV16 E6/E7 knockdown resulted in 12 DE lncRNAs of which 9 were upregulated and 3 were downregulated (Figure 7.12). Of these lncRNAs TIRARP-AS1 was the most significantly upregulated (6.0 Log2 FC average) and RNF157-AS1 was the most significantly downregulated (-2.2 log2 FC average).

The DE-lncRNAs from the HPV16 E6/E7 knockdown were compared to the DE lncRNAs from TCGA (Table 7.2, Appendix Figure 7.3). There is only one lncRNA in common between the TCGA HPV+ OPC samples and in the HPV16 E6/E7 knockdown that had a statistically significant change in the TCGA cohort, RNF157-AS1. The RNF157-AS1 lncRNA only had a -1.9-Fold change in the TCGA data so it did not meet our FC cut off criteria in previous analysis.

An interactome analysis of RNF157-AS1 was created depicting its predicted miRNA targets that were differentially expressed in the TCGA data as well as miRNA-mRNA interactions (Figure 7.11).

The downregulated lncRNA RNF157-AS1 was predicted to sponge to 8 miRNAs that are up in the HPV16 OPC patient cohort. Of these miRNAs miR-20b-5p and miR-106a-5p had the most mRNA targets and a lot of these were shared between these miRNAs.

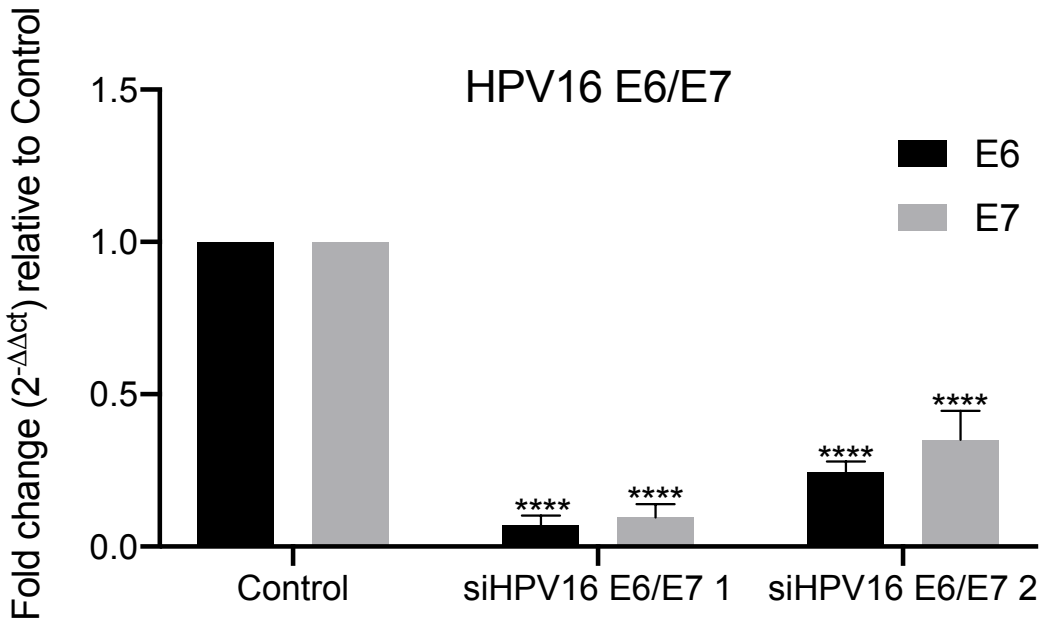


Figure 7. 9 Quantitative PCR validation of HPV16 E6/E7 knockdown using several siRNA in the HPV16+ HNC cell line, SCC090: E6 and E7 were knockdown to 90% by siHPV16 E6/E7 1 and 70-80% by siHPV16 E6/E7 2, this decrease in E6/E7 expression was statistically significant compared to control cells. Normalisation of gene expression was achieved using the calibrator β2M and fold relative to the control using the 2-ΔΔCt calculation. Statistical analysis was performed using two tailed student T-test ****p<0.0001.

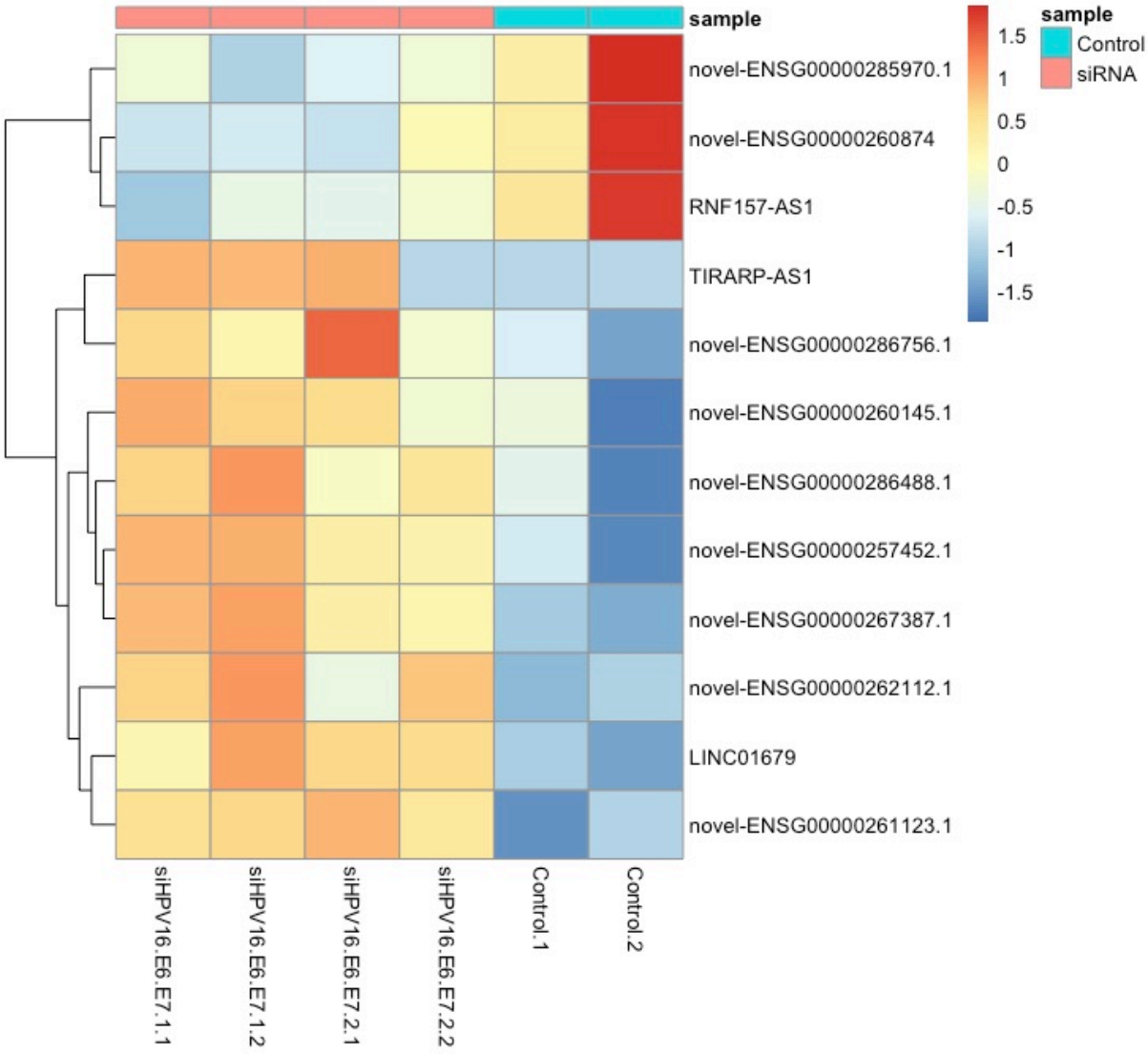


Figure 7. 10: A constructed heat map representing differentially expressed lncRNAs in common between siRNAHPV16 E6/E7 1 and 2 (pink bar) compared to the control (blue bar) in SCC090 cell line. This heatmap represents all lncRNAs with a p-value greater than 0.05. LncRNAs shown in red are upregulated, whilst lncRNAs in blue are downregulated.

Table 7. 2: Summary of DE lncRNAs from siRNA knockdown compared to TCGA

LncRNA Transcript	TCGA log2FC	TCGA p-value	siHPV16 E6/E7 average log2 FC	siHV16 E6/17 average p-value
LINC01679	0.229	7.61E-01	2.7385	1.73E-02
TIRARP-AS1	0.144	8.29E-01	5.9985	9.30E-03
ENSG00000257452.1	-0.689	1.34E-01	3.5065	1.31E-04
ENSG00000260145.1	0.585	6.00E-01	2.5415	1.41E-02
ENSG00000260874.5	0.419	1.90E-01	-1.814	1.65E-02
ENSG00000261123.1	2.311	1.06E-04	4.1275	1.61E-02
ENSG00000262112.1	-0.32	2.95E-01	1.71	2.53E-02
RNF157-AS1	**1.573**	**1.44E-04**	**-2.223**	**1.35E-02**
ENSG00000267387.1	-0.082	8.48E-01	2.946	6.00E-04
ENSG00000285970.1	0.504	1.85E-01	-1.7785	3.80E-02
ENSG00000286488.1	0.136	8.53E-01	2.7815	2.28E-02
ENSG00000286756.1	-1.369	2.84E-01	3.8405	1.31E-02

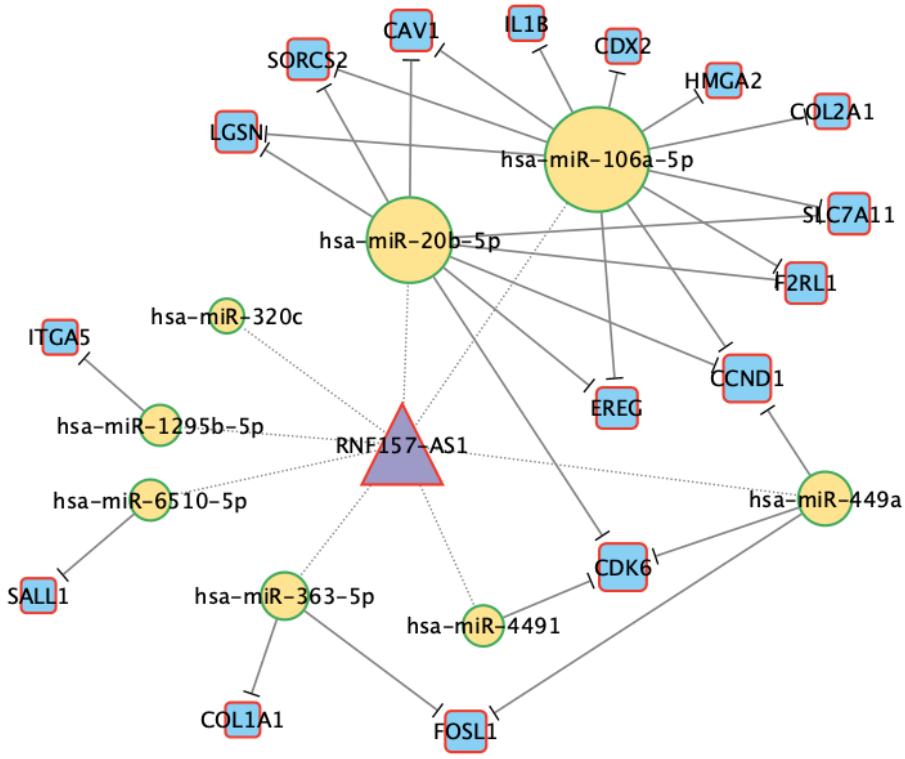


Figure 7. 11: RNF157-AS1 and sponged miRNA interactome: This interactome depicts the differentially expressed lncRNA RNF157-AS1 that was down regulated in HPV16+ oropharyngeal cancers and upregulated with HPV16 E6/E7 were knocked down. This lncRNAs is linked to its predicted target miRNAs, these miRNAs are upregulated in HPV16+ oropharyngeal cancers. The miRNA known mRNA targets, that are down regulated in HPV16+ oropharyngeal cancers, are also depicted. Triangles represent lncRNAs, circles represent miRNAs and squares represent mRNA. Green node borders represent upregulated expression and red represent downregulated expression pattern. The size of the nodes indicates number of interactions each node has, the larger the size the more interactions.

7.4. Discussion

7.4.1. The characterisation of lncRNAs in HPV16 positive OPC

In this study, we focused on the impact of HPV16 on the long non-coding RNA family in OPC. There is increasing evidence that lncRNAs are associated with a wide range of essential functions in cell biology, including the hallmarks of cancers [358, 359]. To date, very few studies have investigated lncRNAs in HPV related HNSCC [144, 189] and none have investigated HPV related OPC specifically. Previous studies investigating the impact of HPV in TCGA HNC cohort examined all HNC subtypes, including oral cavity, OPC and laryngeal and not OPC by itself. Clinically HNCs are diagnosed and treated according to their specific subtype, and this fact can also be applied to the genomic analysis [360].

We identified 1929 differentially expressed lncRNAs when comparing HPV16+ to HPV- OPC from the TCGA dataset. We applied a 5-fold cut-off to identify the most significant HPV OPC involved lncRNAs, reducing the list to 37 DE-lncRNAs. Our most upregulated lncRNA was ENSG00000282816.1, and downregulated was ENSG00000231683.6. The function of these lncRNAs have not previously been reported in any cancer.

Similarly to previous studies we saw the lncRNAs COLCA1 [189], PDE6B-AS1 [189], ENSG00000267247.1 [189], ENSG00000251076.1 [189], ENSG00000258670 [144] and ENSG00000233936 [144] to be upregulated and the lncRNA ENSG00000233085.5 [189] to be downregulated. All other lncRNA with a 5-fold cut off were unique to this study. Our unique lncRNAs could be due to using an updated version of the lncRNAs database, or these DE-lncRNAs are specific to the HPV16+ OPC subtype. Some of the lncRNAs below 2-fold change may also be seen in other studies; we could not make this comparison as other studies did not provide a complete list of all DE lncRNAs.

To minimise the nomenclature for these lncRNAs, we decided to use the HUGO Gene Nomenclature Committee (HGNC) and then Ensembl IDs for novel transcripts that have not been characterised. All other possible gene names of these lncRNAs across different databases can be found in Appendix Table 7.3.

Of these common DE-lncRNAs, COLCA1 was shown to be co-regulated with its predicted target COLCA2 in HPV16+ HNC [189], similar to our investigation. COLCA1 or colorectal cancer-associated 1 has also been described to be upregulated and co-regulated with COLCA2 in colon cancer [361-363] and primary biliary cholangitis [364], and renal carcinomas [365]. COLCA1 expression has been observed in cells that play a role in inflammation and cancer, particularly eosinophils, which were the highest expressors of COLCA1 [361], but COLCA1 targets have yet to be identified. Perhaps COLCA1 expression is upregulated in response to HPV infection to promote the recruitment of inflammatory cells in OPC.

7.4.2. Building an RNA axis interactome to gauge functionality

Given that most of these DE-lncRNAs are novel lncRNA transcripts and their functions are unknown, we created a lncRNA/miRNA/mRNA interactome analysis to elucidate the molecular interactions and cellular pathways. One relatively new function of LncRNAs is their ability to sponge miRNAs, preventing them from regulating their target genes creating a competitive endogenous RNA interaction. We utilised various online predictive databases to determine lncRNA-miRNA binding interactions for our list of DE-lncRNAs.

The lncRNA-miRNA pairings were filtered for those with inverse expression patterns in the TCGA cohort. Nearly all our differentially expressed lncRNAs had the potential to bind to one or more DE-miRNA. Only LINC00165 and ENSG00000225280 did not sponge any of our DE-miRNAs, although they still possess the ability to sponge other miRNAs not seen in this study.

The lncRNA ENSG00000287941 sponged the most miRNAs (13 upregulated miRNAs), and ENSG00000283403 sponged the most downregulated miRNAs (16 miRNAs). These lncRNAs have not been described as possible sponges; however, it is a bioinformatics prediction that will require further investigation and *in vitro* validation.

Of our DE-miRNAs from HPV16+ OPC, miR-20b, miR-363, miR-193b, miR-517a have all previously been described to have similar differential expression in HPV16+ OPC

[343]. All other DE-miRNAs (28 miRNAs with 2-fold change) were unique to our analysis.

Of the downregulated miRNAs, many are a part of the miRNA cluster on chromosome 19 (C19MC), one of the largest miRNAs clusters containing 46 miRNAs. Altered expression of the chromosome 19 microRNA cluster has been associated with numerous cancer types, including breast cancers [366], parathyroid tumours [367], thyroid adenomas, hepatocellular carcinoma [368], embryonal tumours [369] and testicular carcinomas [370].

Members of this cluster have been individually assessed, and several are tumour suppressor miRNAs and have some involvement in cell proliferation and invasion [371-373]. For example, miR-517a is a tumour suppressor in bladder cancers; its overexpression inhibited cell proliferation and induced apoptosis [371] and in colorectal cancers miR-517a regulated cell migration and invasion [372]. In HNSCC miR-520b is also seen to be downregulated in HNC cell lines, leading to increased tumour growth [373] and miR-519a is reported to be a tumour suppressor by targeting of DNA replication associated protein MCM7 [374]. In ovarian cancers its overexpression inhibited cell proliferation and promoted apoptosis [375].

Several of these C19MC miRNAs are also associated with chemosensitivity of tumours. miR-522 is downregulated in DOX resistance colon cancer cells, the overexpression of this miRNA did restored DOX sensitivity [376]. miR-519a expression could promote the cisplatin sensitivity of lung cancer cells, this cisplatin sensitivity was mediated by miR-519a sponging by the lncRNA LINC00221, creating cisplatin resistant cells [377]. This highlights another role lncRNAs might play in HPV related OPC, the sponging of miRNAs to influence the drug sensitivity of cancer cells to treatment. Interestingly, all our upregulated lncRNAs were predicted to regulate multiple members of this miRNA cluster, demonstrating that that C19MC may play and important role in HPV16 infection in OPC.

Of the sponged miRNAs hsa-miR-6510-5p (upregulated) was targeted by 13 downregulated lncRNAs and hsa-miR-4745-5p (downregulated) was sponged by the most upregulated lncRNAs (n=13). Neither of these miRNAs have previously been

described to be differentially expression in HPV16+ OPC. Although in HPV- tumours the miR-6510-3p has been identified to be downregulated when compared to normal tissue [378, 379]. It would be worth comparing lncRNA expression between HPV16+ OPC tissue to normal tissue to identify if miR-6510-5p is upregulated in HPV16+ tumours compared to normal or if HPV16 retains normal expression levels or miR-6510-5p and this miRNA is specific to HPV16- OPC. As for miR-4745 this has not been previously identified in HNSCC, but in colon cancers this miRNA had elevated expression levels [380], indicating its role in various cancer types, its function has not yet been reported.

7.4.3. Predicting the impact on gene regulation by these DE-lncRNAs/miRNAs interactions

To gain a greater understanding of the impact of these DE-lncRNA/miRNAs interactions we investigated the known mRNA targets of our DE-miRNAs and filtered these for those with an inverse expression pattern in HPV16+ OPC. The miRNAs with the most DE-mRNA targets in HPV16+ OPC, hsa-miR-106a-5p (upregulated) and hsa-miR-20b-5p (upregulated) have previously been described with a similar expression pattern in HPV16+ OPC [343]. The downregulated miRNAs has-miR-4658 and hsa-miR-516a-3p differential expression are novel to this study. The upregulated miRNAs miR-106a, miR-20b are part of the same cluster and have previously been shown to be upregulated in HPV16+ OPC [381] playing an important role in cancer cell proliferation.

In HPV+ cervical cancers, HPV E7 can promote the expression of miR-106a via increasing DCGR8, to promote cellular proliferation, reduce autophagy and increase radiation sensitivity of cells [382, 383]. miR-20b has been shown to be overexpressed in HPV+ HNC [384, 385] and cervical cancers, and is specifically regulated by the E6 oncogene. Its inhibition was shown to decreased cell migration and invasion [386, 387]. In colorectal cancers, miR-4658 was sponged by the lncRNA RP11-362K14.5 upon germline mutations [388]. This miRNA has also been seen to be sponged by the lncRNA SNHG12 in hepatocellular carcinomas, this interaction led to an increase in cellular proliferation, migration, EMT and reduced apoptosis [389]. These examples illustrate the continuing importance and impact of miRNAs-lncRNA sponging on cancer pathways.

7.4.4. In-vitro knockdown of HPV16 E6/E7 to confirm the regulation of these DE-lncRNAs

So far, the data is a mapping exercise using bioinformatics and associated tools. To determine if our TCGA prediction were valid, we downregulated E6/E7 using siRNAs in a specific cell line. The RNA was sent for sequencing, and we identified 12 DE lncRNAs of which 9 were upregulated and 3 were downregulated. Of these lncRNAs, TIRARP-AS1 was the most significantly upregulated and RNF157-AS1 was the most significantly downregulated.

TIRARP-AS1 has not previously been described but RNF157-AS1 was discovered by Ota *et al.* [390]. RNA157-AS1 is downregulated in thyroid neoplasms [391] but its function has not yet been reported. The 12 DE-lncRNA from our E6/E7 knockdown were compared to the 1929 DE-lncRNAs in the TCGA cohort. Only RNF157-AS1 was found to be in common between the TCGA HPV+ OPC samples and in the HPV16 E6/E7 knockdown that had a statistically significant change.

This differences in lncRNA expression between patient tissue and cell lines may be due to the sample type and the fact that lncRNA expression is highly specific to the environment of a cell. We know that cell lines consist of one homogenous cell type which have uniform genotypic and phenotypic characteristics. In comparison, tissue samples exhibit heterogeneity in that a solid tumour mass may consist of cancer cells, normal cells, immune cells, fibroblasts, stromal cells, extracellular matrix, and blood vessels. This heterogeneity in cellular composition may contribute to the variability in lncRNA and miRNA expression. The TCGA selection criteria for their analysis is 60% of tumour nuclei in the sample, thus it is plausible that the TCGA tissues do harbour various cell types.

Another factor which could influence these signatures is the synchronicity of the cells, that is we have a mixture of cells at different cell cycle stages. We know for a fact that, cells at different cycles have a different pattern of miRNA and mRNA expression [392]. Thus, a cell line even though it is homogenous in composition, unsynchronised cells can also affect the expression of RNAs and genes.

7.5. Conclusion

We have described a unique lncRNA/miRNA/mRNA axis in HPV16 related OPC. Building this interactome, we can begin to visualise the complex relationship between coding and the non-coding RNA families. This type of mapping provides in silico evidence that HPV16 can altered expression of lncRNAs in oropharyngeal cancers and that these DE-lncRNAs can potentially sponge miRNAs involved in numerous cellular pathways related to cancer development. This approach will allow the identification of other mechanistic pathways between the virus and our RNA axis.

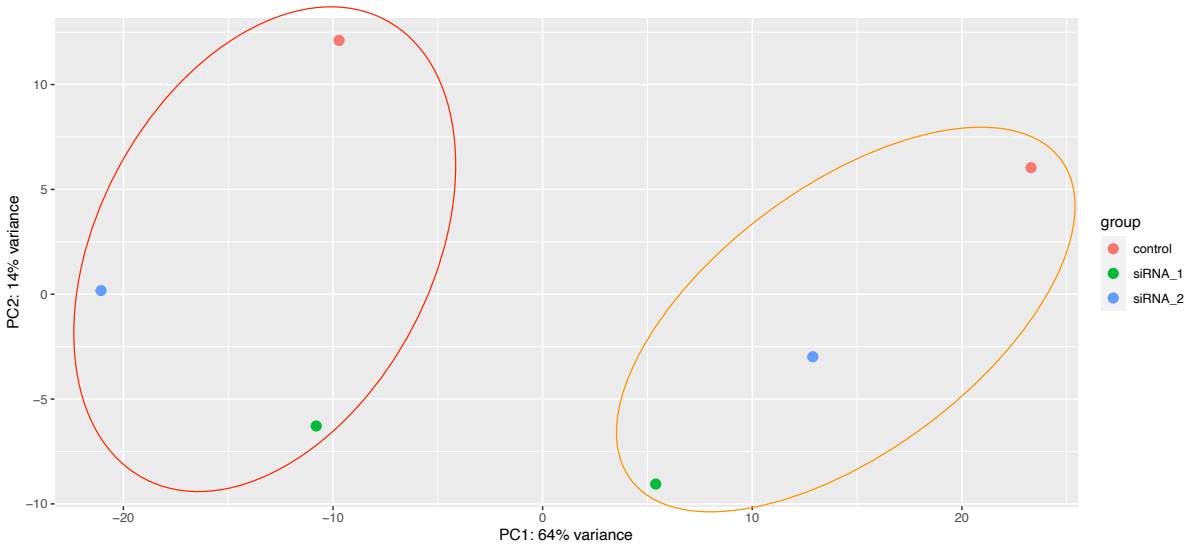
It would be beneficial to investigate these pathways observed in our network with *in vitro* evidence. For our study we observed the interactions between lncRNA and miRNAs from our network analysis. Measuring the expression levels of miRNAs in an lncRNA overexpression system would indicate to us that these specific lncRNAs can regulate our predicted miRNAs. Performing a luciferase assay would confirm the lncRNA binding and sponging of the miRNAs from the interactome. RNA pull-downs of DE-lncRNA from the TCGA cohort would identify mRNA targets of our specific lncRNAs, further uncovering the functionality of the lncRNAs and the cellular pathways they are involved in, in HPV16+ OPC. Finally, we would need to further validate if HPV16 and its viral oncogenes E6/E7 can induce lncRNA expression in OPC. This could be done by performing a HPV16 E6/E7 overexpression in an HPV- HNC cell line. Our interactome analysis could also be further expanded by including lncRNA-mRNA interactions or other forms of non-coding RNAs such as circular-RNAs, which also have the potential to sponge miRNAs. The possibilities are endless with network analysis, and this is a great starting tool for researchers for the investigation of HPV driven cancers.

Overall, these results have demonstrated to us that HPV16 does lead to altered expression of lncRNAs in OPC. Through an interactome analysis we also identify novel lncRNA-miRNA sponging interactions in HPV16+ OPC. These molecular interactions have the potential to be key targets of HPV16 and play an important role in HPV virally driven cancers.

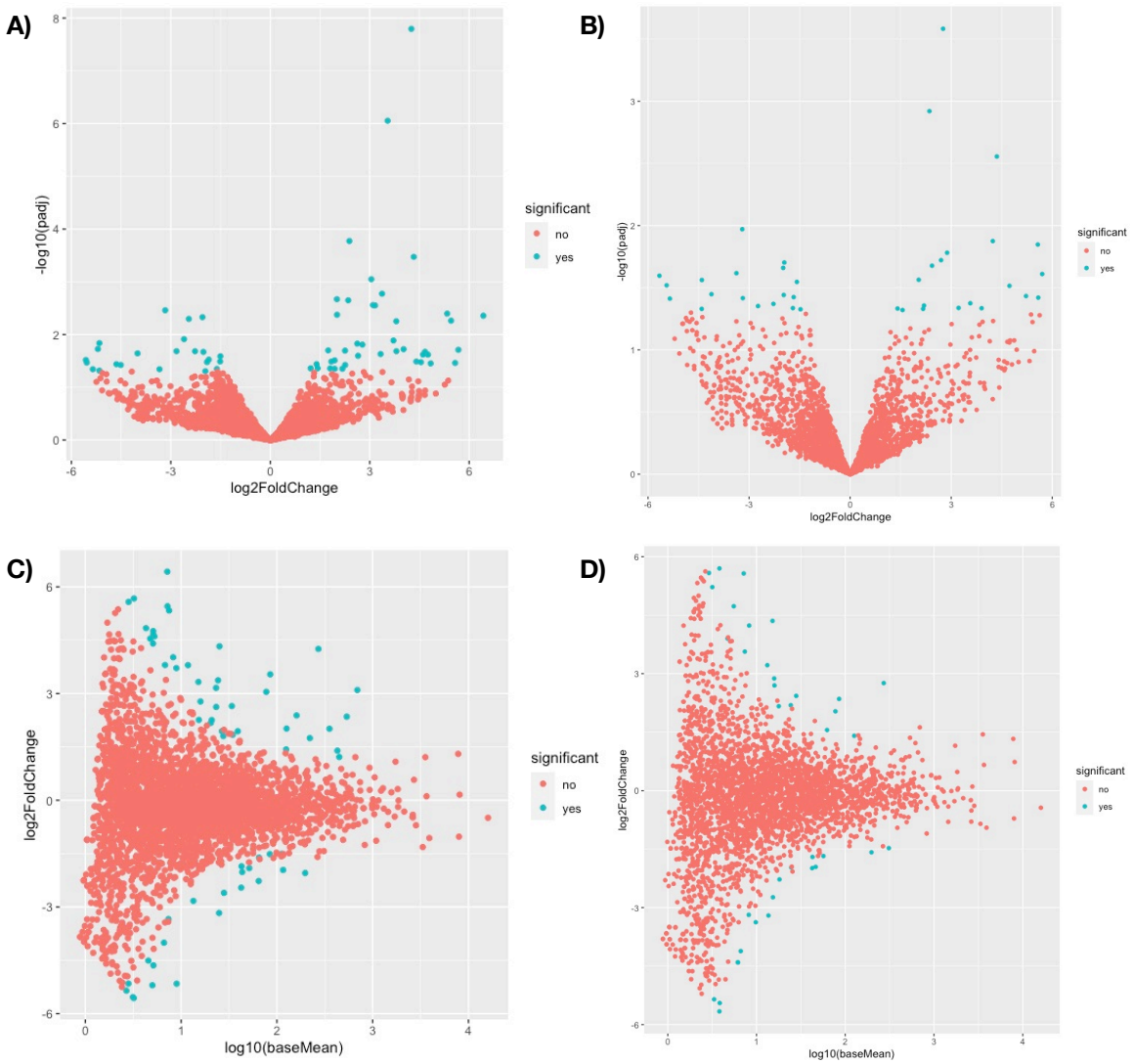
7.6. Appendix

This chapter was written with the intention to be submitted as a research article to the journal of *Molecular cancer.* Below is the appendix for this chapter which will form the supplementary information for the submitted research article.

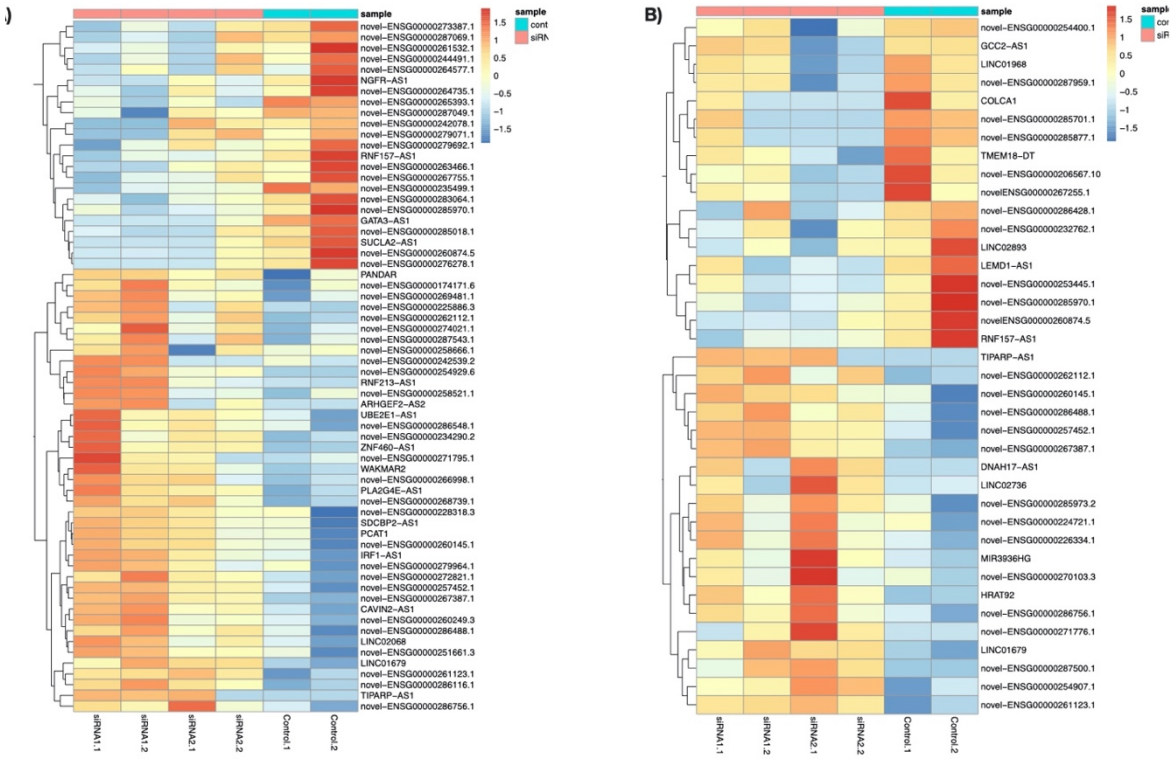
All Appendix tables can be found via: https://doi.org/10.26195/mdzm-qc54



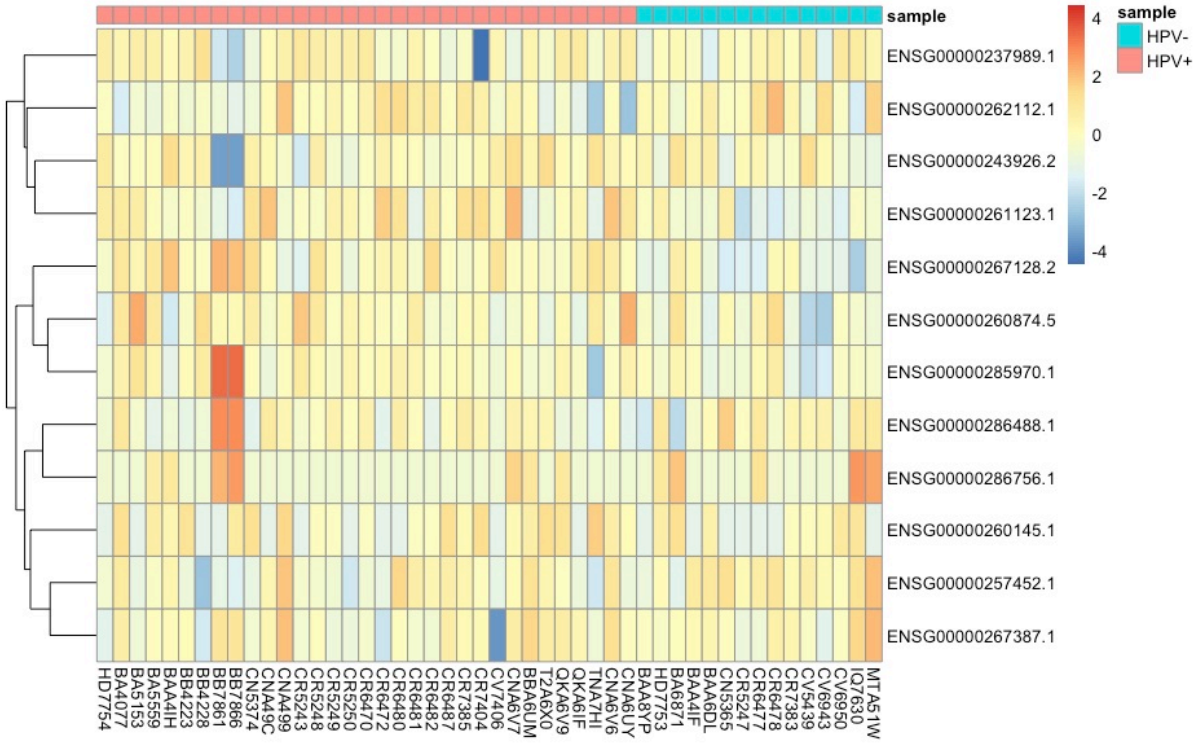
Appendix Figure 7. 1: Plot PCA of RNA sequencing data on HPV16 E6/E7 siRNA knockdown in SCC090: The Two circles represent each biological replicate. There is strong variability between the two biological replicate groups.



Appendix Figure 7. 2: Differentially expressed lncRNAs between HPV16 E6/E7 siRNA knockdown and control SCC090 cells. A) Volcano plot shows the difference in log2Fold change on lncRNAs in siHPV16 E6/E7 1 cells on the X axis and significance (log2 of adjusted p-value) on the Y axis, B) Volcano plot shows the difference in log2Fold change on lncRNAs in siHPV16 E6/E7 2 cells on the X axis and significance (log2 of adjusted p-value) on the Y axis, C) MA plot representing the differences between mRNAs expression in siHPV16 E6/E7 1 cells, comparing Log2 Fold change on the Y axis and Log10 of the mean average normalised counts, D) MA plot representing the differences between mRNAs expression in siHPV16 E6/E7 2 cells, comparing Log2 Fold change on the Y axis and Log10 of the mean average normalised counts.



Appendix Figure 7. 3: Heatmaps representing differentially expressed lncRNAs in HPV16 E6/E7 siRNA knockdown with A) siRNA 1 and B) siRNA 2 vs control, p-value ≤0.05



Appendix Figure 7. 4: Differentially expressed lncRNAs from HPV16 E6/E7 siRNA knockdown in TCGA cohort: Only lncRNA showed a statistically significant change

Chapter 8: Overall discussion

8.1. An HPV-RNA interactome is a valuable tool for uncovering novel mechanistic insights

A major challenge in cancer research is understanding how a virus interacts with a host cell on a genome-wide level. These interactions will have an effect on the system-wide regulatory pathways in a cell that may be driving viral oncogenesis. Virus interactions with host genes are highly complex, with many levels of regulation and control. At present, there are currently limited tools for the interrogation of these viral regulatory networks.

In this thesis, we have proposed a network analysis that includes viral oncogenes, host genes and proteins, and their regulators such as transcription factors, miRNAs, and lncRNAs. In this thesis we provide the methodology for developing such a network analysis, this is described in Chapter 3. This type of mapping analysis will provide a complete picture of the molecular interactions driving virally associated cancers. A thorough understanding of the molecular pathways altered by oncogenic viruses is needed to identify targets that can be utilised for early diagnosis, prevention, and treatment methods.

The human papillomavirus is a major etiological agent for HNC [4, 9, 11]. However, we still lack a fundamental understanding of the molecular interactions that drive these cancers, particularly the impact of HPV on the non-coding RNA milieu. Up to 85% of OPC, a subtype of HNC, are infected with the high-risk HPV variant HPV16; hence this was the primary focus of the research [10]. The interactions between HPV16 and its viral oncogenes E6 and E7 with human proteins is well understood and characterised [73]. The impact of HPV16 E6 and E7 expression on non-coding RNAs and their associated pathways in OPC are largely unreported. As we demonstrated in our literature review (Chapter 1), multiple studies have identified HPV16 as a modulator for the differential expression patterns of miRNAs in OPC [93-95, 97-101, 108]. A singular miRNA can potentially regulate thousands of gene targets that involve numerous cellular pathways that could lead to tumour development. By creating an HPV16-miRNA interactome, we demonstrated the influence HPV and its viral

oncogenes have on miRNA expression via direct or indirect interactions with multiple miRNA targets.

We provided a logical step-by-step process to produce a viral-ncRNA interactome, using publicly available datasets and the network building program, Cytoscape (Chapter 3) [102]. This interactome depicts interactions between HPV16 viral oncoproteins and the human host proteins, with the addition of the gene regulators, miRNAs, and their predictive and known targets. This methodology was utilised to produce the HPV/miRNA interactome in Chapters 1 and 4 and was adapted and employed throughout this research. We also expanded the interactome to include other ncRNAs, such as lncRNAs (Chapter 6 and 7). This network analysis would provide insight into the interplay between HPV16 E6/E7 and the specific miRNAs, miR-33a and miR-496 (Chapter 4) [101]. Using our HPV16 viral/miRNA interactome, a novel mechanistic pathway was extracted and validated *in vitro* (Chapter 5). We predicted that HPV16 E6 and E7 could indirectly alter the expression of SREBF2, leading to the upregulation of the encoded miRNA, miR-33a; this interaction was novel and was then experimentally validated.

The strength of a molecular network map is determined by the quality of the dataset used in its production. For this thesis the VirusMentha [213] dataset was selected for identification of viral and host protein-protein interactions. VirusMentha is a freely available dataset curated from five different databases from the IMEx consortium (MINT [393], IntAct [394], DIP [395], MatrixDB [396] and Biogrid [214]). VirusMentha is updated automatically on a weekly basis. As VirusMentha is the most relevant and comprehensive viral/host dataset available it was selected for the basis of our interactome analysis. At the time of our methodology publication [102], the dataset used from VirusMentha consisted of 82 publications with the earliest dating from 1990 and the latest from 2016. It is important to note that not all publications describing HPV16 interactions with host proteins are included in Virusmentha [397, 398]. From the literature addition human protein binding interactions with HPV16 E6 (ADMR1, EIF3D, GOPC, HERC2, MTA2, NOTCH1, OASL, PDZD11, PLEKHA5, PSME4, TNKS1BP1, TRIM24, UCHL5) and HPV16 E7 (ABHD10, CUL2, SARS2) need to also be considered in network analysis. A combination of interactions from Virusmentha and the relevant literature should be utilised when accessing viral-host interactions. The molecular landscape of HPV driven cancers is constantly changing, with new viral-

host interactions being discovered as detection methods are improving. When developing a viral-host interactome it is important to bare this in mind and regularly updated our interactome analysis based on new publications and new findings that have been discovered.

Although the methodology was initially developed with miRNAs as the primary focus, we have expanded this analysis to lncRNAs (Chapter 6 and Chapter 7). The HPV-ncRNA interactome was utilised to identify potential feedback loops associated with cancer cell survival between the lncRNA TRINGS, p53 and HPV16 E6. We also explored the role lncRNA-miRNA sponging interactions may play in HPV16 associated OPC and identified the potential for lncRNAs to target miRNAs associated with pathways, including cell cycle, cell proliferation and receptor binding.

We acknowledge that this type of network analysis is speculative, and the predicted interactions require validation *in vitro*. The caveat of this type of analysis is that filtering the larger interactome may create bias, and various information may be lost. This analysis is also highly dependent on the available datasets at the time. Despite this, the HPV-ncRNA interactome is a powerful tool in providing a targeted framework to uncover novel regulatory pathways.

This methodology can be modified for the investigation of other viruses. Seven viruses are recognised as the causative agent for cancers, including the Epstein-Barr virus, the Human papillomavirus, Hepatises B virus, Hepatitis C virus, Polyomavirus, and Kaposi sarcoma herpesvirus and Human T-lymphotropic virus type 1 [399, 400]. In 2014, 15% of the new cancer cases were attributed to carcinogenic infections [401]. These cancer-causing viruses all belong to multiple viral families with different routes of transmission and infection that contribute to the development of cancers. Each one of these viruses could be assessed in a viral-ncRNA interactome to evaluate their molecular impact, elucidating how these viruses lead to an oncogenic phenotype.

In this research, we have touched on applying drug target analysis combined with interactome analysis (Chapter 1). By using databases such as DrugBank [119], the impact of current or new drugs on cellular pathways can be assessed. The viral-ncRNA interactome can be developed as a predictive tool to understand the global

implications of gene-targeted therapies. The idea of combining molecular network interactomes with modern therapeutics is a relatively new field of research termed network medicine [402]. This is an emerging tool that can explore the molecular complexity of the disease and uncover the interplay between disease and drugs. Network medicine integrates the genome, proteome, cellular environment and pathological phenotypes via network analysis. It can assist in predicting treatment side effects and can be utilised in the future for predictive, personalised medicine approaches.

Network medicine may also be utilised to identify the impact of novel ncRNA therapeutic strategies. Non-coding RNAs act as oncogenes and tumour suppressors, targeting multiple cellular pathways simultaneously. Due to these characteristics, the manipulation of ncRNAs may be a viable option for therapeutics in HPV driven cancers. For example, the two major targets of HPV16 E6 and E7 are p53 and prB respectively, both have been identified as targets for ncRNA therapeutics. HPV16 E6 prevents apoptosis of cancer cells via the ubiquitin ligation of E6AP, causing the degradation of p53. In a recent publication by Martinez-Noël [403], 32 different miRNAs, including miR-375-3p, were transfected into HPV18 positive cervical cancer cell lines. Several of these miRNAs stabilised p53 levels, decreased E6AP protein and induced apoptosis on the HPV positive cancer cells. MiR-375 has also been show to stabilise prB levels in HPV16 and HPV18 cervical cancer cell lines [404]. Similarly lncRNAs have been identified to stabilise p53 levels with the Damaged Induced Long Noncoding RNA (DINO), being reported to override HPV16 E6/UBE2A mediated p53 degradation in cervical cancer cell lines [318]. DINOS expression sensitised the HPV positive cells to chemotherapy agents and rendered them vulnerable to metabolic stress [318]. This demonstrates the use of ncRNAs to target proteins like p53 and prB to stabilise or reverse the impacts of HPVs oncogenes.

Our interactome analysis may assist in identifying additional candidates such as SREBF2 and E2F2 to be targeted by ncRNAs as a treatment strategy. Anti-ncRNA or ncRNA replacement therapies is a long-term goal for molecular studies of HPV induced HNC. To achieve this, we first must improve our understanding as to the molecular functions and the cellular pathways in which these ncRNAs are embedded this can assist with interactome analysis.

There are very few studies that have utilised an HPV-interactome analysis.
The majority have either mapped only miRNA-mRNA interactions in HNC [103, 405] or HPV-protein [406] interactions. Similar to Farooq *et al.* [406], our interactome analysis found that the viral oncogenes E7, E6 and E5 have the most interactions with host proteins. These interactions are primarily associated with cell cycle pathways.
Our interactome analysis is unique as we are the first to integrate the interactions between HPV-miRNAs, HPV-mRNAs, miRNAs-mRNAs, with putative transcription factors. The integration of each of these components into one interactome provides a more holistic view of the molecular environment of a cell during HPV infection.

Our methodology for producing a viral-ncRNA interactome will allow researchers to identify specific new targets for gene therapy or potential ncRNA biomarkers for the detection of viral diseases. This methodology will also uncover the specific impact of viruses on genes and ncRNAs in numerous cancer types. Our approach can assist in identifying interactors and can lay the foundation for future mechanistic studies. Moreover, the mapping and data in the interactome are amenable to either a broad or more focused approach in identifying pathways.

8.2. The human papillomavirus and its viral oncogenes E6 and E7 can influence the expression of miRNAs in Oropharyngeal cancers: Determining the role SREBF2 and miR-33a play in HPV16 positive OPC

The interactions between HPV16 and its viral oncogenes E6 and E7 with human proteins is well understood and characterised; however, we still only have a basic understanding of how viruses interact or regulate ncRNAs. During our literature review, we demonstrated that HPV16 infection could lead to the altered expression of miRNAs in OPC (Chapter 1). However, the molecular mechanism behind this observation remains unknown. There is a suggestion that the two major viral oncogenes E6/E7, can alter the expression of miRNAs in cervical cancers and HFKs [60, 407, 408]. We aimed to explore the idea of the HPV16 viral oncogenes E6/E7 altering miRNA expression in OPC and determine the targets and pathways these miRNAs were involved in (Chapter 4).

Previous research from our laboratory identified 13 miRNAs (5 upregulated and 8 downregulated) to have differential expression patterns in HPV16+ compared to HPV- OPC via an LNA array (Chapter 4) [101]. Of these, miR-496 was the most significantly downregulated, and miR-33a was the most upregulated. The miRNAs expression pattern was further validated in OPC patient cohort from TCGA (Chapter 5). Previously it was determined that miR-496 and miR-33a could be regulated by the viral oncogene E6 (and its two isoforms E6*I and E6**II) in an overexpression system in HNC cells. Target prediction identified several novel mRNA targets for miR-496 (ECT2, p16 and E2F2). The overexpression of miR-496 revealed only E2F2 to be downregulated, and a luciferase assay confirmed the binding and regulation of E2F2 by miR-496.

To further uncover the relationship behind miR-496, miR-33a and HPV16, we developed and produced an HPV16-miRNA interactome for network analysis. It was revealed through this analysis that miR-33a is encoded within intron 16 of the transcription factor SREBF2. This discovery led us to investigate the relationship between miR-33a and SREBF2 in HPV16 positive OPC (Chapter 5). SREBF2 had not previously been examined in this cancer type. Both SREBF2 and miR-33a were shown to be overexpressed in HPV16+ patient tissue (TCGA data) and HPV16+ HNC cell lines. The knockdown and overexpression system of HPV16 E6 and E7 led to altered SREBF2 and miR-33a expression, further supporting the notion that these viral genes could regulate specific genes and miRNAs in OPC.

There was evidence of a positive correlation between SREBF2 and miR-33a expression in our own data set (Chapter 5) and other cancer types [280] [281]. Due to this positive correlation, the literature has described that SREBF2 and miR-33a are co-transcribed [276, 277, 280]. We aimed to determine if this was the case in HPV16 positive OPC. During our knockdown of SREBF2 at an RNA level, there was no consistent decrease in miR-33a expression across cell lines. Collectively the results may suggest that miR-33a is under its own form of regulation, potentially being spliced out of SREBF2 prior to siRNA silencing. Drosha will likely cleave out the primary miRNA as this has been known to be a very efficient process.

This study demonstrated that HPV16 will still lead to miR-33a overexpression despite gene targeting of SREBF2. Targeting intronic miRNAs by HPV16 adds a further layer

of complexity to the HPV regulatory network. We demonstrated that HPV16 could induce the expression of intronic miRNAs independently of their host gene. Investigations into the exact mechanisms behind this regulation would add value to this research.

In future research, it would be interesting to perform a functional analysis of SREBF2 and miR-33a and the role each play in HPV+ OPC. SREBF2 and miR-33a are involved in the homeostatic regulation of cholesterol levels within a cell [270, 271]. Under sterol depletion, SREBF2 is expressed to promote the biosynthesis of cholesterol and miR-33a is expressed to prevent the efflux of cholesterol out of a cell. A cholesterol efflux assay or metabolic assay could be performed on HPV16 E6/E7 knockout cells to determine the role cholesterol metabolism plays in HPV+ OPC.
These assays would assist in determining the impact of altered miR-33a and SREBF2 expression on metabolic pathways in HPV+ OPC.
It has been reported that circulating levels of miR-33a can be altered in adipose and by the changing lipid profile of a cell [409]. HPV16 E6 and E7 are reported to alter the expression of miRNAs packaged into extracellular vesicles in HFK cells [410]. In HPV+ HNC, perhaps the virus may impact circulating miR-33a levels, either directly or indirectly via lipid pathways. It would be interesting to characterise circulating miR-33a levels in response to HPV infection in HNC. The idea of using circulating miRNAs as biomarkers for HPV detection in cervical cancers has been explored [411]. MiR-33a may serve as a circulating biomarker candidate for HPV infected HNC.

Downstream targets of SREBF2 and miR-33a could also be measured in our knockdown systems to determine the molecular consequences of E6/E7 targeting these genes and miRNAs. SREBF2 cholesterol biosynthesis targets that could be investigated further include HMG-CoA reductase, HMG-CoA synthase, the LDL receptor, and miR-33a, ABCA1 and APOa1 could also be measured due to their involvement in cholesterol metabolic pathways.

In addition, a cell viability (resazurin) assay and cell proliferation assay would provide evidence as to the impact of SREBF2 and miR-33a regulation of metabolism on cancer cell pathways. Overall, a functional analysis would provide insight into cholesterol metabolism's role in HPV16 associated OPC. Uncovering SREBF2/miR-33a

regulation of metabolism may elucidate a crucial pathway for HPV16 associated OPC cell survival. This pathway could be used as a target for drug targeted treatment.

This project primarily focused on two host components (SREBF2 and miR-33) and two viral components (E6 and E7) from our HPV16/miRNA interactome (Chapter 5). We provided a mode of action behind HPV16 E6/E7 targeting miR-33a and miR-496. The regulatory relationship between miR-496 and its novel target gene E2F2 was experimentally validated. Our model also contained numerous intermediate transcription factors that warrant further investigation and validation.

Upstream of SREBF2 are several transcription factors that are direct targets of E7(RBL2, HDAC1, E2F1, E2F4, JUNB, JUND, IRF1, TEAD4, HDAC2, FOXM1) or both E6/E7 (EP300 and MYC). HPV16 targeting of these transcription factors potentially induces the expression of SREBF2. These transcriptional factors also have the potential to drive miR-33a independently and may serve as the mechanistic link between HPV16 E6/E7 and this miRNA.

We briefly investigated these components in the HPV16+ OPC patient cohort from TCGA; in order to promote SREBF2/miR-33a expression, the transcription factors need to be upregulated. HDAC1, E2F1, IRF1, HDAC2, FOXM1, TBP were all upregulated in HPV16+ OPC and are potential candidates for further investigation. Bioinformatic analysis can be conducted to identify transcription factor binding elements upstream of SREBF2 and chromatin immunoprecipitation (ChIP) to confirm transcriptional binding to SREBF2 or miR-33a promoters.

Our HPV16-miRNA interactome analysis also revealed several transcription factors that miR-33a regulates that lie upstream of miR-496 (GATA2, BACH1, STAT5A and CTFC), potentially promoting the expression of this miRNA. Of these, GATA2 and BACH1 were downregulated in HPV16+ OPC TCGA samples, suggesting their targeting by miR-33a and no longer being able to promote miR-496 expression as seen in our model. Experimental validation of miR-33a regulation of these transcription factors could be performed using a luciferase assay and ChIP assays to confirm transcriptional factor binding to miR-496 promoters. This investigation would provide

a mechanistic link between our two miRNAs of interest, miR-33a and miR-496, highlighting the downstream consequences of miRNA targeting by HPV16 in OPC.

HPV16 contains several other early viral genes (E1, E2, E4 and E5), which play essential roles in viral infection and replication. There are few reports describing the involvement of these viral genes (E1, E2, E4 and E5) in cancer [412-415] and our understanding as to their impact on miRNA expression is lacking. It has been demonstrated that HPV16 E5 can downregulate the expression of miR-196a in cervical cancers, impacting cellular proliferation and apoptosis [416]. This example shows that other HPV16 viral early genes have the potential to alter miRNA expression, having a tumorigenic effect and warrants further investigation in HPV16 associated HNC.

The HPV16 E1, E2, E4 and E5 viral genes could be added to our HPV16-miRNA interactome analysis as a starting point to identify novel mechanistic pathways driving HPV16 early infection. This type of analysis will reveal the role miRNAs potentially play at all stages of HPV16 infection and tumour development in the Head and Neck region.

Taken together, these unique interactions uncovered between HPV16 E6/E7, miRNAs (miR-33a and miR-496) and genes (SREBF2 and E2F2) in this study provide significant insight into the cellular mechanisms behind HPV16 infection in OPC. The role of miR-496, miR-33a and SREBF2 has not previously been reported in HPV16+ OPC, representing a novel interaction. This research also provides insight into cellular pathways that HPV16 targets in OPC, such as cholesterol metabolism and cell cycle regulation. These events may be necessary for HPV16 viral replication and the transformation of a cell and pose as potential to be targeted for novel gene therapies to treat this devastating disease.

Given the vast effect that HPV16 has on miRNA expression and their downstream targets, we wanted to explore the notion that HPV16 infection will also impact other forms of non-coding RNA expression. We also chose to investigate the newly discovered long non-coding RNAs (lncRNAs) and their potential role in HPV driven cancers.

8.3. LncRNAs can play an important role in cancer cell survival in HPV related cancers: an in-depth investigation into the lncRNA TRINGS

To begin my journey into the investigation of lncRNAs in HPV driven cancers, I conducted a 6-month research project under the guidance of Dr Karl Munger at Tufts University, Boston. The primary focus of this research was investigating the impact of HPV16 E6/E7 expression on lncRNAs and identifying the roles the differentially expressed lncRNAs play in cervical cancer (Chapter 6). Another goal of this project is to identify the function of lncRNAs and how they interact with HPV16 E6/E7 in virally infected Head and Neck cancers.

Previously the Munger laboratory conducted RNA sequencing on HFKs expressing HPV16 E6/E7 to identify differentially expressed lncRNAs induced by these viral oncogenes [170]. The data analysis revealed that the lncRNA, TRINGS, was one of the most significantly downregulated lncRNAs in E6/E7 expressing cells. TRINGS, or TP53 regulated lncRNA induced under glucose starvation, was first characterised by *Khan et al.* [206] and was reported to be regulated by TP53, having a protective function of cancer cells under glucose starved conditions. Given that TP53 is a major target of HPV16 E6 and TRINGS is lowly expressed in E6/E7 expressing cells, this study aimed to determine if the modulation of TRINGS expression could increase the susceptibility of E6/E7 expressing cells to cell death upon metabolic stress.

In a series of cell lines, we observed that TRINGS expression is lower in E6/E7 expressing cells (iNOKs, iHFKs) and TP53 nulls cells (HCT116). This supported the RNA sequencing data and reports in the literature that TRINGS expression tends to follow TP53 expression [206]. Various cells lines were treated to normal or glucose starved conditions over 72hrs, and it was revealed that TRINGS expression was not stimulated by glucose starvation alone. This observation could be due to TP53 levels being low in cervical cancer cells (due to HPV16 E6 expression) [417] as TP53 is the main factor stimulated by glucose starvation. TP53 will then go on and promote TRINGs expression in cancer cells. It would be beneficial to extend the glucose starvation beyond 72hrs in our cell line panel to determine if a change in TRINGS expression is observed after a particular extended period of metabolic stress.

Several siRNA knockdown systems were performed to determine if TRINGS played a role in cell survival. It was observed across numerous cell lines that the decrease in TRINGS significantly reduced cancer cell survival and induced cell death. This supports the notion that TRINGS is involved in the necrotic cellular pathway [206] and demonstrates that the modulation of TRINGS impacts cervical cancer cell survival.

We found differing results across cell lines in keratinocytes stably expressing HPV16 E6/E7. In the iNOKs, reduction of TRINGS did not alter cell viability, but in iHFKs, a decrease in cell viability was observed. It appeared that HPV16 E6/E7 expression does not impact TRINGS ability to regulate cell survival. That modulation of TRINGS expression affects the viability of cancer cells and 'normal' cell types.

We also examine the clinical relevance of TRINGS and its expression levels. It was demonstrated that TRINGS expression was significantly decreased in HPV16 and HPV18 infected cervical cancer cells compared to HPV- (TCGA data). There was no significant change in the HPV16+ OPC cohort. TRINGS was downregulated by two high-risk HPV variants, demonstrating the lncRNA is not limited to HPV16. TRINGS dysregulated expression may be necessary for all high-risk HPV viral infections in cervical cancers. We could perform survival analysis using Kaplan-Meyer plots to determine the relationship between low TRINGS expression levels and patient survival for clinical relevance.

Given that TP53 is the main target of HPV16 E6, it would be highly interesting to investigate the expression levels and function of TRINGS in E6 and E7 alone expressing cells. Cells expressing HPV16 E7 have been reported to have restored levels of TP53 due to the stabilisation of the protein [340, 341]. These cells are also sensitised to cell death upon metabolic stress [418, 419]. It has been reported the sensitivity of HPV16 E7 expressing cells to metabolic stress is rescued by the lncRNA DINO depletion [419]. Perhaps in E7 alone expressing cells, p53 levels are restored, under metabolic stress, p53 can promote TRINGS expression, which will protect cancer cells from necrotic cell death. An updated model of this mode of action is described in Figure 8.1. This demonstrates the role of lncRNAs in p53 pathway regulation and the impact of the modulation of lncRNA levels on cancer cell survival.

As p53 is a tumour suppressor, activating TRINGS expression [206] suggests a feedback loop mechanism. Under metabolic stress, p53 has some protective effect over cancer cells via its promotion of TRINGS expression. A knockdown or overexpression system of TRINGS in HPV16 E6 and E7 individually expressing cells whilst treating cells to normal and glucose starved conditions will potentially reveal the nature of these feedback mechanisms. This is significant as it will reveal a mechanism in which viral oncogenes can individually protect and promote tumour development via lncRNAs. The modulation of lncRNA expression may elicit a clinical response to drugs that cause metabolic stress in HPV associated cancers.

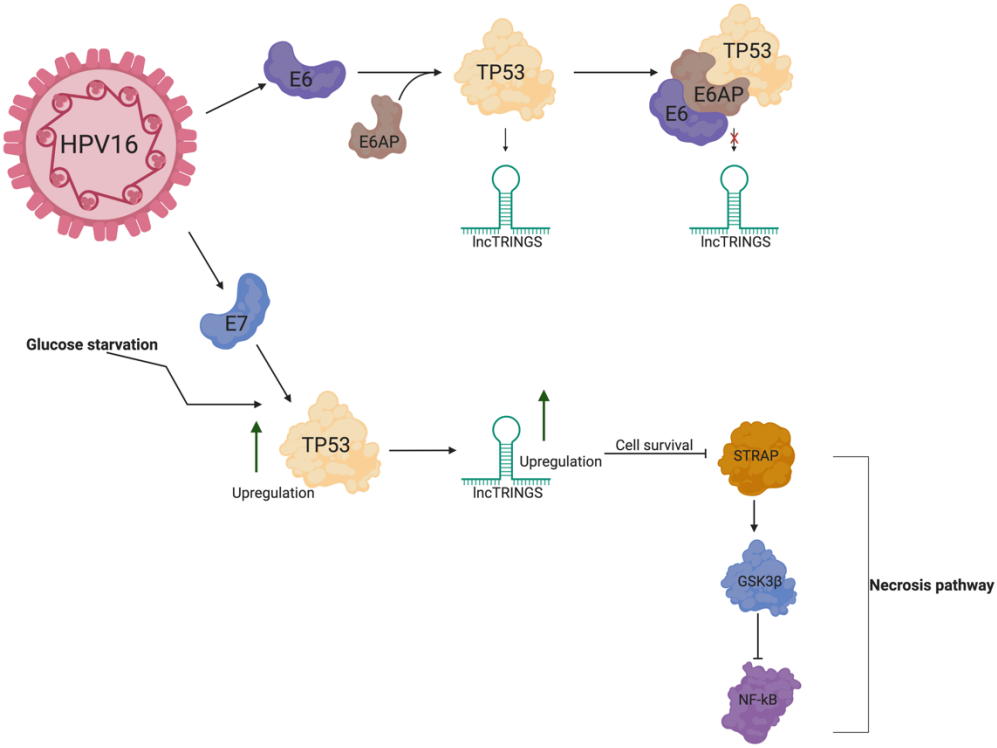


Figure 8. 1A model depicting HPV16 E6 and E7 interactions with TP53 and the lncRNA TRINGS

Identification and validation of TRINGS targets will also provide insight into the mechanism behind the lncRNAs action and the cellular pathways it is involved in. We utilised an interactome analysis to potentially identify mechanistic pathways between HPV16 E6, TP53 and TRINGS and their target genes. Only one TRINGS target, STRAP, has been investigated experimentally, demonstrating TRINGS regulation of

the necrotic cell pathway [206]. Of the predictive TRINGS targets from our interactome analysis, HERC5 is of particular interest as it is also targeted by p53 and is a known regulator of the antiviral immune response [420, 421]. Interestingly, expression of HERC5 promotes an antiviral immune response, so are p53 and TRINGS important promoters of HERC5 gene expression. By HPV16 downregulating p53 and TRINGS, they can no longer promote HERC5 expression, preventing an antiviral immune response during HPV infection.

To further investigate other TRINGS-RNA interactions, siRNA techniques or CRISPR can determine if the modulation of TRINGS expression will alter predicted target genes. TRINGS-protein interactions could also be investigated by RNA pull-downs, CLIP or RNA immunoprecipitation would validate RNA protein binding interactions. There are very few lncRNAs that have been characterised in HPV related cancers. It will be essential to understand and identify the nature of these interactions as the ncRNA family are highly abundant.

Overall the results from this study demonstrate that the HPV16 viral oncogenes E6/E7 can influence the expression of lncRNAs, and this interaction may represent an additional mechanism driving HPV associated cancers. We observed that any modulation to these lncRNAs expression would have an impact on the cancer cellular pathways they are involved in. In this chapter, we described HPV16 E6/E7 expressing cells showing a decrease in the lncRNA, TRINGS expression, and the downregulation of TRINGS led to decreased cancer and normal cell survival.

The alteration of lncRNAs may not be due to a direct binding or interaction between the HPV16 viral genes and a lncRNA but instead is a result of HPV16 targeting their promoter genes, like p53. During this investigation, we identified that TRINGS had altered expression in cervical cancer patient tissue, but there was no significant change in HPV related HNC. This finding led us to investigate if HPV16 and its viral oncogenes E6 and E7 can alter the expression of lncRNAs in HNC and, if so, what are the cellular pathways those lncRNAs are involved in during tumour development.

8.4. Exploring the impact of HPV16 on long non-coding RNA expression in Oropharyngeal cancers.

During our study into the lncRNA TRINGS, we wanted to determine if this lncRNA presented the same differential expression pattern in OPC (Chapter 6). It was identified that TRINGS did not show a statistically significant change in expression between HPV16+ and HPV- OPC patients. These results indicated that HPV16 might lead to differential expression of a separate cohort of lncRNAs in OPC compared to cervical cancers. This is has been observed for miRNA expression between HPV16+ cervical cancer and OPC [98]. In an explorative approach, we identified differentially expressed (DE) lncRNAs in HPV16+ OPC and determined if the oncogenes HPV16 E6/E7 could induce this altered lncRNA expression.

Analysis of RNA sequencing data from The Cancer Genome Atlas revealed that 1929 lncRNAs to be DE in the HPV16+ OPC compared to the negative (Chapter 7). Applying a 5-fold cut off to reduce the number of lncRNAs, to identify those most involved in HPV16 related OPC revealed 37 lncRNA transcripts. As well as lncRNAs, we also investigated miRNA and mRNA and identified 32 DE miRNAs (2-fold cut off) and 60 mRNA (5-fold cut off) in the HPV16+ OPC patient samples. Sponging interactions between our DE-lncRNAs and DE-miRNAs were predicted, as well as DE-miRNA and DE-mRNA regulatory interactions. These interactions were merged for an interactome analysis to uncover lncRNA regulatory pathways that might drive HPV16 associated OPC.

LncRNA sponging of miRNAs has been implicated in the development of numerous cancer types [422-424]; these regulatory relationships could be vital in uncovering the progression of tumour development. Current lncRNA-miRNA sponging interactions have been described in HNC and impact patient survival and progression [351, 405, 425]. HPV16 related HNC is its own distinct entity and presents a very different molecular pattern than HPV- HNC, hence the need to investigate them separately. Our HPV16/lncRNA/miRNA interactome analysis is the first to depict lncRNA-miRNA sponging interactions in HPV16 induced OPC, and these interactions are key to fully understanding HPVs involvement in tumorigenesis.

The lncRNA/miRNA/mRNA interactions from our network analysis were associated with cell cycle regulation, cellular proliferation and signalling receptor binding. KEGG pathway enrichment analysis revealed the lncRNA/miRNA/mRNA interactions were associated with focal adhesion, PI3K-Akt signalling and HPV infection. The regulation of the cell cycle and cellular proliferation is the primary function of HPV16 E6/E7 to promote tumorigenesis [426, 427]. The PI3K-Akt signalling is also consistently dysregulated in HPV related cancers [428]. It is not surprising the DE-lncRNAs, and their miRNA/mRNA targets are also involved in these pathways.

The involvement of lncRNA/mRNA in cell cycle pathways in HPV+ HNC has previously been reported, as well as splicesome, endocytosis MAPK and Wnt signalling pathways [144, 189]. The association between these lncRNAs/miRNA/mRNAs interactions and these cellular pathways demonstrate the potential role these DE lncRNAs may play in HPV16 tumorigenesis.

One example of a mechanistic pathway derived from our analysis is the downregulation of SChLAP1 by the virus. This lncRNA has previously been reported to be involved in cell proliferation and metastasis of prostate cancers [429-431]. It has also been described to possess the capacity to sponge miRNAs, including miR-524-5p [431] and miR-198 [429]. In our study, SChLAP1 had the most interactions in our downregulated lncRNA cohort (predicted to sponge 10 different miRNAs). One of the miRNAs, SChALP1, was predicted to sponge miR-20b-5p, which is shown to be upregulated in HPV16+ OPC. The upregulation of miR-20b-5p potentially is due to it no longer being under the control and regulation of SChLAP1, allowing miR-20b-5p to act and regulate its targets CDK6 and CCND1. These miR-20b-5p mRNA target interactions have been associated with cell cycle, migration and invasion pathways in different cancer types [432, 433]. This example demonstrates the impact and downstream consequences of differential lncRNA expression in HPV16 associated OPC. Numerous molecular pathways similar to SChLAP1 can be extracted from our interactome analysis. Of course, these mechanistic interactions need to be validated experimentally to understand their involvement in HPV16 related cancers fully.

As shown with miRNAs (Chapter 4, 5), we aimed to determine if the HPV16 E6/E7 viral oncogenes could alter lncRNA expression specifically. Analysis of an E6/E7 knockdown system revealed the lncRNA RNF157-AS1 to have consistent altered

expression in the E6/E7 knockdown and in the HPV16 patient cohort. This lncRNA function has not previously been reported, but it has been identified to be upregulated in ovarian cancers and is a positive predictor for the survival of ovarian cancer patients [434, 435]. Some suggest that the nucleotide complementarity of antisense lncRNAs allows them to have regulative functions over their sense genes [436].

Antisense-lncRNAs are more likely to regulate the expression of their corresponding protein-coding genes and their functions being both oncogenic or tumour suppressive [436]. The RNAF157-AS1 is an antisense lncRNAs to the RNF157 gene, which has been linked to the promotion of cell cycle progression in cancer cells [437]. Perhaps RNF157-AS1 is also associated with the regulation of these pathways. It is crucial for us to summarise lncRNA roles and molecular mechanisms in cancer development and progression to understand how OPC may develop fully.

It would be beneficial in future research to repeat the E6/E7 knockdown system accompanied by overexpression of E6/E7 as well as other viral oncogenes (E4 and E5) in OPC cell lines to validate RNF157-AS1s expression pattern further as well as identify other lncRNAs that these viral oncogenes specifically regulate.

Currently, the molecular interactions we have observed are predictive models. Still, they do provide an insight into the mechanistic pathways of lncRNAs and their involvement in cancer pathways of HPV16+ OPC. The majority of the lncRNA differentially expressed in the HPV16+ OPC cohort have not previously been reported or characterised. In future research, identifying lncRNA functions can begin with determining their location, are the lncRNAs present in the nucleus or cytoplasm. The subcellular location of lncRNAs is critical to understanding their functions in terms of their interactome partners, post or co-transcriptional regulatory modifications, external stimuli and their role in homeostasis [438].

In the nucleus, lncRNAs function to modulate transcriptional programs through chromatic interactions and remodelling and spacial organisation via scaffolding [438]. In the cytoplasm, lncRNAs function to regulate translational programs, signal transduction pathways and post-transcriptional control of gene expression [438]. Identification of this HPV16 DE-lncRNAs localisation will indicate the level of regulation

of our lncRNAs, their mode of action, and type of targets (DNA, RNA, proteins and miRNAs). In our current analysis, we used predictive-based methods to identify lncRNA-miRNA interactions.

The lncRNA-miRNA sponging interactions that we predicted in our interactome analysis provide further insightful molecular mechanisms that may drive HPV16 alteration of miRNA expression, potentially via alteration of lncRNA sponging interactions. To further enrich our understanding of the function of these lncRNAs, the lncRNA-miRNA interactions will be validated experimentally. This can be done via luciferase assays to determine the direct binding of miRNAs to our DE-lncRNAs and qPCR analysis of downstream miRNA targets with lncRNA overexpression systems.

Our analysis focused on lncRNA-miRNA interactions, but lncRNAs have the potential to regulate DNA, RNA and proteins. In future research, it would be interesting to investigate predictive mRNA targets of our DE-lncRNAs, add this to our interactome analysis, and perform an RNA pulldown assay to identify proteins that are directly bound to the lncRNAs in HPV16+ OPC.

Overall we have identified DE-lncRNAs in HPV16+ OPCs and recognise predictive miRNA targets that potentially are sponged by the DE-lncRNAs. A visual interactome was constructed to visualise these associations. We have mapped the lncRNA/miRNA/mRNA interactions to extrapolate novel mechanistic links between HPV16, lncRNAs and miRNAs that may drive virally associated OPC. LncRNAs have been identified as crucial regulators related to tumorigenesis and this explorative study has laid the groundwork for more exciting discoveries

8.5. Overall conclusions

The Human Papillomavirus is one of the main drivers for Head and Neck cancer, particularly Oropharyngeal cancers. The vast impact of this virus on the molecular functioning of a cell is evident. Non-coding RNAs, such as miRNAs and lncRNAs, are important gene regulators. The effects of HPV16 infection on these RNA family members provide an additional layer of regulatory complexity when uncovering the driving factors of HPV16 HNC development. In this study, we aimed to discover the

impact of HPV16 on non-coding RNAs in HNC and provide a better understanding as to how HPV leads to this disease.

Towards this end, we developed a methodology for interrogating large genomic datasets to elucidate novel mechanistic pathways that may drive HPV16 associated HNC. We characterised the multitude of interactions between types non-coding RNAs and identified novel mechanistic links between the HPV16 viral oncogenes, miRNAs, lncRNAs, gene targets, as well as drug interactions. Our findings also supported the view that HPV16 and its oncogenes E6/E7 can alter various forms of non-coding RNAs in HNC developmental pathways. The findings from our study provide a fuller picture of the complexity of the molecular networks driving HPV16 infection in HNC.

References

1. Bray, F., et al., *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.* CA: a cancer journal for clinicians, 2018. **68**(6): p. 394-424.
2. AIHW, *Head and neck cancers in Australia* in *Cancer series 83*. 2014: Canberra.
3. Chatuvedi A.K., et al., *Incidence Trends for Human Papillomavirus-related and-urelated Oral Squamous Cell Carcinomas in the United States.* Journal of Clinical Oncology, 2008. **28**(4): p. 612-.
4. Spence, T., et al., *HPV associated head and neck cancer.* Cancers, 2016. **8**(8): p. 75.
5. de Martel, C., et al., *Worldwide burden of cancer attributable to HPV by site, country and HPV type.* International journal of cancer, 2017. **141**(4): p. 664-670.
6. Herrero, R., et al., *Human papillomavirus and oral cancer: the International Agency for Research on Cancer multicenter study.* Journal of the National Cancer Institute, 2003. **95**(23): p. 1772-1783.
7. Chung, C.H., et al., *Genomic alterations in head and neck squamous cell carcinoma determined by cancer gene-targeted sequencing.* Annals of oncology, 2015. **26**(6): p. 1216-1223.
8. Fossum, C.C., et al., *Characterization of the oropharynx: anatomy, histology, immunology, squamous cell carcinoma and surgical resection.* Histopathology, 2017. **70**(7): p. 1021-1029.
9. Tumban, E., *A current update on human papillomavirus-associated head and neck cancers.* Viruses, 2019. **11**(10): p. 922.
10. Spence, T., et al., *HPV Associated Head and Neck Cancer.* Cancers (Basel), 2016. **8**(8).
11. Chaturvedi, A.K., et al., *Human papillomavirus and rising oropharyngeal cancer incidence in the United States.* Journal of clinical oncology, 2011. **29**(32): p. 4294.
12. Gillison, M.L., et al., *Epidemiology of Human Papillomavirus-Positive Head and Neck Squamous Cell Carcinoma.* J Clin Oncol, 2015. **33**(29): p. 3235-42.
13. Argiris, A., et al., *Head and neck cancer.* The Lancet, 2008. **371**(9625): p. 1695-1709.
14. Arantes, L.M.R.B., et al., *Serum, plasma and saliva biomarkers for head and neck cancer.* Expert review of molecular diagnostics, 2018. **18**(1): p. 85-112.
15. Suh, Y., et al., *Clinical update on cancer: molecular oncology of head and neck cancer.* Cell death & disease, 2015. **5**(1): p. e1018.
16. Liederbach, E., et al., *The national landscape of human papillomavirus-associated oropharynx squamous cell carcinoma.* International journal of cancer, 2017. **140**(3): p. 504-512.
17. Chow, L.Q., *Head and neck cancer.* New England Journal of Medicine, 2020. **382**(1): p. 60-72.
18. Lindel, K., et al., *Human papillomavirus positive squamous cell carcinoma of the oropharynx.* Cancer, 2001. **92**(4): p. 805-813.
19. Sturgis, E.M. and P.M. Cinciripini, *Trends in head and neck cancer incidence in relation to smoking prevalence.* Cancer, 2007. **110**(7): p. 1429-1435.

20. Danaei, G., et al., *Causes of cancer in the world: comparative risk assessment of nine behavioural and environmental risk factors.* The Lancet, 2005. **366**(9499): p. 1784-1793.
21. Jemal, A., et al., *Global cancer statistics.* CA: a cancer journal for clinicians, 2011. **61**(2): p. 69-90.
22. Goldenberg, D., et al., *Habitual risk factors for head and neck cancer.* Otolaryngology—Head and Neck Surgery, 2004. **131**(6): p. 986-993.
23. Marur, S., et al., *HPV-associated head and neck cancer: a virus-related cancer epidemic.* The lancet oncology, 2010. **11**(8): p. 781-789.
24. Young, D., et al., *Increase in head and neck cancer in younger patients due to human papillomavirus (HPV).* Oral oncology, 2015. **51**(8): p. 727-730.
25. Gillison, M.L., et al., *Evidence for a causal association between human papillomavirus and a subset of head and neck cancers.* Journal of the National Cancer Institute, 2000. **92**(9): p. 709-720.
26. Chaturvedi, A.K., et al., *Worldwide trends in incidence rates for oral cavity and oropharyngeal cancers.* Journal of clinical oncology, 2013. **31**(36): p. 4550.
27. Ang, K.K., et al., *Human papillomavirus and survival of patients with oropharyngeal cancer.* New England Journal of Medicine, 2010. **363**(1): p. 24-35.
28. Lindquist, D., et al., *Human papillomavirus is a favourable prognostic factor in tonsillar cancer and its oncogenic role is supported by the expression of E6 and E7.* Molecular Oncology, 2007. **1**(3): p. 350-355.
29. Chaturvedi, A.K., et al., *Incidence trends for human papillomavirus-related and-unrelated oral squamous cell carcinomas in the United States.* Journal of clinical oncology, 2008. **26**(4): p. 612-619.
30. Sturgis, E.M. and P.M. Cinciripini, *Trends in head and neck cancer incidence in relation to smoking prevalence: an emerging epidemic of human papillomavirus‐associated cancers?* Cancer: Interdisciplinary International Journal of the American Cancer Society, 2007. **110**(7): p. 1429-1435.
31. Mehanna, H., et al., *Prevalence of human papillomavirus in oropharyngeal and nonoropharyngeal head and neck cancer—systematic review and meta‐analysis of trends by time and region.* Head & neck, 2013. **35**(5): p. 747-755.
32. Panarese, I., et al., *Oral and Oropharyngeal squamous cell carcinoma: prognostic and predictive parameters in the etiopathogenetic route.* Expert review of anticancer therapy, 2019. **19**(2): p. 105-119.
33. Schache, A.G., et al., *HPV-related oropharynx cancer in the United Kingdom: an evolution in the understanding of disease etiology.* Cancer research, 2016. **76**(22): p. 6598-6606.
34. Stransky, N., et al., *The mutational landscape of head and neck squamous cell carcinoma.* Science, 2011. **333**(6046): p. 1157-1160.
35. Braakhuis, B.J., et al., *Genetic patterns in head and neck cancers that contain or lack transcriptionally active human papillomavirus.* Journal of the national cancer institute, 2004. **96**(13): p. 998-1006.
36. Fakhry, C., et al., *Improved survival of patients with human papillomavirus–positive head and neck squamous cell carcinoma in a prospective clinical trial.* Journal of the National Cancer Institute, 2008. **100**(4): p. 261-269.
37. Andl, T., et al., *Etiological involvement of oncogenic human papillomavirus in tonsillar squamous cell carcinomas lacking retinoblastoma cell cycle control.* Cancer research, 1998. **58**(1): p. 5-12.

38. Hafkamp, H.C., et al., *A subset of head and neck squamous cell carcinomas exhibits integration of HPV 16/18 DNA and overexpression of p16INK4A and p53 in the absence of mutations in p53 exons 5–8.* International journal of cancer, 2003. **107**(3): p. 394-400.
39. Klussmann, J.P., et al., *Expression of p16 protein identifies a distinct entity of tonsillar carcinomas associated with human papillomavirus.* The American journal of pathology, 2003. **162**(3): p. 747-753.
40. Slebos, R.J., et al., *Gene expression differences associated with human papillomavirus status in head and neck squamous cell carcinoma.* Clinical cancer research, 2006. **12**(3): p. 701-709.
41. Martinez, I., et al., *Identification of differentially expressed genes in HPV-positive and HPV-negative oropharyngeal squamous cell carcinomas.* European journal of cancer, 2007. **43**(2): p. 415-432.
42. Slebos, R.J., et al., *Proteomic analysis of oropharyngeal carcinomas reveals novel HPV-associated biological pathways.* International journal of cancer, 2013. **132**(3): p. 568-579.
43. Carninci, P., et al., *The transcriptional landscape of the mammalian genome.* science, 2005. **309**(5740): p. 1559-1563.
44. Birney, E. and S.A.D.R.G. TR, *Gingeras EH Margulies Z. Weng M. Snyder ET Dermitzakis RE Thurman et al. 2007. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project.* Nature. **447**: p. 799-816.
45. Esteller, M., *Non-coding RNAs in human disease.* Nature reviews genetics, 2011. **12**(12): p. 861.
46. Calin, G.A., et al., *Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia.* Proceedings of the national academy of sciences, 2002. **99**(24): p. 15524-15529.
47. Reinhart, B.J., et al., *The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans.* nature, 2000. **403**(6772): p. 901.
48. Cheng, L.-C., M. Tavazoie, and F. Doetsch, *Stem cells: from epigeneticsto microRNAs.* Neuron, 2005. **46**(3): p. 363-367.
49. Zhang, B., X. Pan, and T.A. Anderson, *MicroRNA: a new player in stem cells.* Journal of cellular physiology, 2006. **209**(2): p. 266-269.
50. Tanno, B., et al., *Silencing of endogenous IGFBP-5 by micro RNA interference affects proliferation, apoptosis and differentiation of neuroblastoma cells.* Cell death and differentiation, 2005. **12**(3): p. 213.
51. Brenner, S., *The genetics of Caenorhabditis elegans.* Genetics, 1974. **77**(1): p. 71-94.
52. Lee, R.C., R.L. Feinbaum, and V. Ambros, *The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14.* cell, 1993. **75**(5): p. 843-854.
53. Pasquinelli, A.E., et al., *Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA.* Nature, 2000. **408**(6808): p. 86.
54. MacRae I., et al., *Structural basis for double-stranded RNA processing by Dicer.* Science, 2017. **311**: p. 195-198.
55. Bohnsack, M.T., K. Czaplinski, and D. GÖRLICH, *Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs.* Rna, 2004. **10**(2): p. 185-191.

56. Lund, E., et al., *Nuclear export of microRNA precursors.* Science, 2004. **303**(5654): p. 95-98.
57. Grishok, A., et al., *Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing.* Cell, 2001. **106**(1): p. 23-34.
58. Hutvágner, G., et al., *A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA.* Science, 2001. **293**(5531): p. 834-838.
59. Zhang, H., et al., *Single processing center models for human Dicer and bacterial RNase III.* cell, 2004. **118**(1): p. 57-68.
60. Harden, M.E. and K. Munger, *Perturbation of DROSHA and DICER expression by human papillomavirus 16 oncoproteins.* Virology, 2017. **507**: p. 192-198.
61. Iwasaki, S., et al., *Hsc70/Hsp90 chaperone machinery mediates ATP-dependent RISC loading of small RNA duplexes.* Molecular cell, 2010. **39**(2): p. 292-299.
62. Peng, Y. and C.M. Croce, *The role of MicroRNAs in human cancer.* Signal Transduction and Targeted Therapy, 2016. **1**: p. 15004.
63. Krol J., Loedige I., and Filipowic W., *The widespread regulation of microRNA biogenesis, function and decay.* Nature Reviews Genetics, 2010. **11**: p. 597-610.
64. Gantier, M.P., et al., *Analysis of microRNA turnover in mammalian cells following Dicer1 ablation.* Nucleic acids research, 2011. **39**(13): p. 5692-5703.
65. Baskerville, S. and D.P. Bartel, *Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes.* Rna, 2005. **11**(3): p. 241-247.
66. Herrero, R., et al., *Human papillomavirus and oral cancer: the International Agency for Research on Cancer multicenter study.* J Natl Cancer Inst, 2003. **95**(23): p. 1772-83.
67. Gillison, M.L., et al., *Epidemiology of human papillomavirus–positive head and neck squamous cell carcinoma.* Journal of clinical oncology, 2015. **33**(29): p. 3235.
68. CDC. *Centre For Disease Control and Prevention (CDC), HPV-Associated Cancer Statistics*. 2016 3rd September 2021]; Available from: http://www.cdc.gov/cancer/hpv/statistics/
69. Brianti, P., E. De Flammineis, and S.R. Mercuri, *Review of HPV-related diseases and cancers.* New Microbiol, 2017. **40**(2): p. 80-85.
70. Syrjänen, K., et al., *Prevalence, incidence, and estimated life-time risk of cervical human papillomavirus infections in a nonselected Finnish female population.* Sexually transmitted diseases, 1990. **17**(1): p. 15-19.
71. Koutsky, L., *Epidemiology of genital human papillomavirus infection.* The American journal of medicine, 1997. **102**(5): p. 3-8.
72. Bruni, L., et al., *Global estimates of human papillomavirus vaccination coverage by region and income level: a pooled analysis.* Lancet Glob Health, 2016. **4**(7): p. e453-63.
73. Zur Hausen, H., *Papillomavirus infections—a major cause of human cancers.* Biochimica et Biophysica Acta (BBA)-Reviews on Cancer, 1996. **1288**(2): p. F55-F78.

74. Calin, G.A., et al., *Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia.* Proc Natl Acad Sci U S A, 2002. **99**(24): p. 15524-9.
75. Reinhart, B.J., et al., *The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans.* Nature, 2000. **403**(6772): p. 901-6.
76. Lee, R.C., R.L. Feinbaum, and V. Ambros, *The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14.* Cell, 1993. **75**(5): p. 843-54.
77. Bagga, S., et al., *Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation.* Cell, 2005. **122**(4): p. 553-63.
78. Bartel, D.P., *MicroRNAs: Genomics, Biogenesis, Mechanism, and Function.* Cell, 2004. **116**(2): p. 281-297.
79. Friedman, J.M. and P.A. Jones, *MicroRNAs: critical mediators of differentiation, development and disease.* Swiss Med Wkly, 2009. **139**(33-34): p. 466-72.
80. O'Brien, J., et al., *Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation.* Front Endocrinol (Lausanne), 2018. **9**: p. 402.
81. Peng, Y. and C.M. Croce, *The role of MicroRNAs in human cancer.* Signal Transduct Target Ther, 2016. **1**: p. 15004.
82. Tran, N., et al., *MicroRNA expression profiles in head and neck cancer cell lines.* Biochem Biophys Res Commun, 2007. **358**(1): p. 12-7.
83. Barker, E.V., et al., *microRNA evaluation of unknown primary lesions in the head and neck.* Mol Cancer, 2009. **8**: p. 127.
84. Ramdas, L., et al., *miRNA expression profiles in head and neck squamous cell carcinoma and adjacent normal tissue.* Head Neck, 2009. **31**(5): p. 642-54.
85. Hui, A.B., et al., *Comprehensive MicroRNA profiling for head and neck squamous cell carcinomas.* Clin Cancer Res, 2010. **16**(4): p. 1129-39.
86. Tran, N., et al., *Potential role of micro-RNAs in head and neck tumorigenesis.* Head Neck, 2010.
87. Johnson, S.M., et al., *RAS is regulated by the let-7 microRNA family.* Cell, 2005. **120**(5): p. 635-47.
88. Moratin, J., et al., *MicroRNA expression correlates with disease recurrence and overall survival in oral squamous cell carcinoma.* Journal of Cranio-Maxillofacial Surgery, 2019. **47**(3): p. 523-529.
89. Gong, S., et al., *The Expression and Effection of MicroRNA-499a in High-tobacco Exposed Head and Neck Squamous Cell Carcinoma: A Bioinformatic Analysis.* Frontiers in oncology, 2019. **9**: p. 678.
90. Chen, L., et al., *Prediction of radiotherapy response with a 5-microRNA signature-based nomogram in head and neck squamous cell carcinoma.* Cancer medicine, 2018. **7**(3): p. 726-735.
91. Olatunji, I., *Potential application of tumor suppressor microRNAs for targeted therapy in head and neck cancer: A mini-review.* Oral oncology, 2018. **87**: p. 165-169.
92. Wald, A.I., et al., *Alteration of microRNA profiles in squamous cell carcinoma of the head and neck cell lines by human papillomavirus.* Head Neck, 2011. **33**(4): p. 504-12.
93. Miller, D.L., et al., *Identification of a human papillomavirus–associated oncogenic miRNA panel in human oropharyngeal squamous cell carcinoma*

validated by bioinformatics analysis of the cancer genome atlas. The American journal of pathology, 2015. **185**(3): p. 679-692.
94. Hui, A.B., et al., *Potentially prognostic miRNAs in HPV-associated oropharyngeal carcinoma.* Clinical cancer research, 2013. **19**(8): p. 2154-2162.
95. Mirghani, H., et al., *Comparative analysis of micro-RNAs in human papillomavirus–positive versus–negative oropharyngeal cancers.* Head & neck, 2016. **38**(11): p. 1634-1642.
96. Pyeon, D., et al., *Fundamental differences in cell cycle deregulation in human papillomavirus–positive and human papillomavirus–negative head/neck and cervical cancers.* Cancer research, 2007. **67**(10): p. 4605-4619.
97. Lajer, C., et al., *Different miRNA signatures of oral and pharyngeal squamous cell carcinomas: a prospective translational study.* British journal of cancer, 2011. **104**(5): p. 830-840.
98. Lajer, C., et al., *The role of miRNAs in human papilloma virus (HPV)-associated cancers: bridging between HPV-related head and neck cancer and cervical cancer.* British journal of cancer, 2012. **106**(9): p. 1526-1534.
99. Gao, G., et al., *A microRNA expression signature for the prognosis of oropharyngeal squamous cell carcinoma.* Cancer, 2013. **119**(1): p. 72-80.
100. Quabius, E.S., et al., *miRNA-expression in tonsillar squamous cell carcinomas in relation to HPV infection and expression of the antileukoproteinase SLPI.* Papillomavirus Research, 2017. **4**: p. 26-34.
101. Mason, D., et al., *Human papillomavirus 16 E6 modulates the expression of miR-496 in oropharyngeal cancer.* Virology, 2018. **521**: p. 149-157.
102. Hill, M., et al., *Developing a virus-microRNA interactome using cytoscape.* MethodsX, 2020. **7**: p. 100700.
103. Lapa, R.M.L., et al., *Integrated miRNA and mRNA expression analysis uncovers drug targets in laryngeal squamous cell carcinoma patients.* Oral oncology, 2019. **93**: p. 76-84.
104. Yoon, A.J., et al., *MicroRNA-based risk scoring system to identify early-stage oral squamous cell carcinoma patients at high-risk for cancer-specific mortality.* Head & Neck, 2020.
105. Zhao, R., et al., *Comprehensive analysis of the whole coding and non-coding RNA transcriptome expression profiles and construction of the circRNA–lncRNA co-regulated ceRNA network in laryngeal squamous cell carcinoma.* Functional & integrative genomics, 2019. **19**(1): p. 109-121.
106. Wang, Y., et al., *Molecular mechanisms and prognostic markers in head and neck squamous cell carcinoma: a bioinformatic analysis.* International Journal of Clinical and Experimental Pathology, 2020. **13**(3): p. 371.
107. Wang, J., et al., *The molecular differences between human papillomavirus-positive and -negative oropharyngeal squamous cell carcinoma: A bioinformatics study.* American Journal of Otolaryngology--Head and Neck Medicine and Surgery, 2019. **40**(4): p. 547-554.
108. Wald, A.I., et al., *Alteration of microRNA profiles in squamous cell carcinoma of the head and neck cell lines by human papillomavirus.* Head & neck, 2011. **33**(4): p. 504-512.
109. Luo, W., et al., *E2F1-miR-20a-5p/20b-5p auto-regulatory feedback loop involved in myoblast proliferation and differentiation.* Scientific reports, 2016. **6**: p. 27904.

110. Qi, J.-C., et al., *miR-20b-5p, TGFBR2, and E2F1 Form a Regulatory Loop to Participate in Epithelial to Mesenchymal Transition in Prostate Cancer.* Frontiers in Oncology, 2020. **9**: p. 1535.
111. Li, X., et al., *miR-150 inhibits proliferation and tumorigenicity via retarding G1/S phase transition in nasopharyngeal carcinoma.* International journal of oncology, 2017. **50**(4): p. 1097-1108.
112. Yang, L., et al., *miR-181b promotes cell proliferation and reduces apoptosis by repressing the expression of adenylyl cyclase 9 (AC9) in cervical cancer cells.* FEBS letters, 2014. **588**(1): p. 124-130.
113. Zheng, W., et al., *miR-31 functions as an oncogene in cervical cancer.* Archives of gynecology and obstetrics, 2015. **292**(5): p. 1083-1089.
114. Tong, F., et al., *HPV+ HNSCC-derived exosomal miR-9 induces macrophage M1 polarization and increases tumor radiosensitivity.* Cancer letters, 2020. **478**: p. 34-44.
115. Park, S., et al., *MiR-9, miR-21, and miR-155 as potential biomarkers for HPV positive and negative cervical cancer.* BMC cancer, 2017. **17**(1): p. 1-8.
116. Chapman, B.V., et al., *MicroRNA-363 targets myosin 1B to reduce cellular migration in head and neck cancer.* BMC cancer, 2015. **15**(1): p. 1-10.
117. Sun, Q., et al., *Dysregulated miR-363 affects head and neck cancer invasion and metastasis by targeting podoplanin.* The international journal of biochemistry & cell biology, 2013. **45**(3): p. 513-520.
118. Park, H., et al., *Dysregulated microRNA expression in adenocarcinoma of the uterine cervix: clinical impact of miR-363-3p.* Gynecologic oncology, 2014. **135**(3): p. 565-572.
119. Wishart, D.S., et al., *DrugBank: a comprehensive resource for in silico drug discovery and exploration.* Nucleic Acids Res, 2006. **34**(Database issue): p. D668-72.
120. Wishart, D.S., et al., *DrugBank 5.0: a major update to the DrugBank database for 2018.* Nucleic acids research, 2018. **46**(D1): p. D1074-D1082.
121. Seiwert, T., et al., *OPTIMA: a phase II dose and volume de-escalation trial for human papillomavirus-positive oropharyngeal cancer.* Annals of Oncology, 2019. **30**(2): p. 297-302.
122. Rigalli, J.P., et al., *Human papilloma virus (HPV) 18 proteins E6 and E7 up-regulate ABC transporters in oropharyngeal carcinoma. Involvement of the nonsense-mediated decay (NMD) pathway.* Cancer letters, 2018. **428**: p. 69-76.
123. Brown, C.J., et al., *A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome.* Nature, 1991. **349**(6304): p. 38-44.
124. Brannan, C.I., et al., *The product of the H19 gene may function as an RNA.* Molecular and cellular biology, 1990. **10**(1): p. 28-36.
125. Mercer, T.R., et al., *Specific expression of long noncoding RNAs in the mouse brain.* Proceedings of the National Academy of Sciences, 2008. **105**(2): p. 716-721.
126. Sun, M. and W.L. Kraus, *From discovery to function: the expanding roles of long noncoding RNAs in physiology and disease.* Endocrine reviews, 2015. **36**(1): p. 25-64.
127. Jarroux, J., A. Morillon, and M. Pinskaya, *History, discovery, and classification of lncRNAs*, in *Long Non Coding RNA Biology*. 2017, Springer. p. 1-46.

128. Luo, X., et al., *Long non-coding RNA implicated in the invasion and metastasis of head and neck cancer: possible function and mechanisms.* Molecular cancer, 2018. **17**(1): p. 14.
129. Yan, X., et al., *Comprehensive genomic characterization of long non-coding RNAs across human cancers.* Cancer cell, 2015. **28**(4): p. 529-540.
130. Peng, L., et al., *LncRNAs: key players and novel insights into cervical cancer.* Tumor Biology, 2016. **37**(3): p. 2779-2788.
131. Gascoigne, D.K., et al., *Pinstripe: a suite of programs for integrating transcriptomic and proteomic datasets identifies novel proteins and improves differentiation of protein-coding and non-coding genes.* Bioinformatics, 2012. **28**(23): p. 3042-3050.
132. Gendrel, A.-V. and E. Heard, *Fifty years of X-inactivation research*. 2011, Company of Biologists.
133. Quinn, J.J. and H.Y. Chang, *Unique features of long non-coding RNA biogenesis and function.* Nature Reviews Genetics, 2016. **17**(1): p. 47.
134. Ma, L., V.B. Bajic, and Z. Zhang, *On the classification of long non-coding RNAs.* RNA biology, 2013. **10**(6): p. 924-933.
135. Fang, Y. and M.J. Fullwood, *Roles, functions, and mechanisms of long non-coding RNAs in cancer.* Genomics, proteomics & bioinformatics, 2016. **14**(1): p. 42-54.
136. Wang, K.C. and H.Y. Chang, *Molecular mechanisms of long noncoding RNAs.* Molecular cell, 2011. **43**(6): p. 904-914.
137. Pandey, R.R., et al., *Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation.* Molecular cell, 2008. **32**(2): p. 232-246.
138. Poliseno, L., et al., *A coding-independent function of gene and pseudogene mRNAs regulates tumour biology.* Nature, 2010. **465**(7301): p. 1033.
139. Yoon, J.-H., et al., *Scaffold function of long non-coding RNA HOTAIR in protein ubiquitination.* Nature communications, 2013. **4**(1): p. 1-14.
140. Gupta, R.A., et al., *Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis.* Nature, 2010. **464**(7291): p. 1071.
141. Wang, P., Z. Ren, and P. Sun, *Overexpression of the long non-coding RNA MEG3 impairs in vitro glioma cell proliferation.* Journal of cellular biochemistry, 2012. **113**(6): p. 1868-1874.
142. Kogo, R., et al., *Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers.* Cancer research, 2011. **71**(20): p. 6320-6326.
143. Quagliata, L., et al., *90 HOXA13 AND HOTTIP EXPRESSION LEVELS PREDICT PATIENTS'SURVIVAL AND METASTASIS FORMATION IN HEPATOCELLULAR CARCINOMA.* Journal of Hepatology, 2013. **58**: p. S39-S40.
144. Nohata, N., M.C. Abba, and J.S. Gutkind, *Unraveling the oral cancer lncRNAome: Identification of novel lncRNAs associated with malignant progression and HPV infection.* Oral oncology, 2016. **59**: p. 58-66.
145. Pasic, I., et al., *Recurrent focal copy-number changes and loss of heterozygosity implicate two noncoding RNAs and one tumor suppressor gene at chromosome 3q13. 31 in osteosarcoma.* Cancer research, 2010. **70**(1): p. 160-171.
146. Chung, S., et al., *Association of a novel long non-coding RNA in 8q24 with prostate cancer susceptibility.* Cancer science, 2011. **102**(1): p. 245-252.

147. Huarte, M., et al., *A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response.* Cell, 2010. **142**(3): p. 409-419.
148. Prensner, J.R., et al., *Transcriptome sequencing across a prostate cancer cohort identifies PCAT-1, an unannotated lincRNA implicated in disease progression.* Nature biotechnology, 2011. **29**(8): p. 742.
149. Rasmussen, S., et al., *A Case of HPV-Associated Oropharyngeal Squamous Cell Carcinoma with Block-Like, Partial Loss of p16 Expression.* Head and Neck Pathology, 2022: p. 1-6.
150. Simoens, C., et al., *Accuracy of high-risk HPV DNA PCR, p16((INK4a)) immunohistochemistry or the combination of both to diagnose HPV-driven oropharyngeal cancer.* BMC Infect Dis, 2022. **22**(1): p. 676.
151. Venuti, A. and F. Paolini, *HPV detection methods in head and neck cancer.* Head and neck pathology, 2012. **6**(1): p. 63-74.
152. Liu, G., et al., *A prognostic 5-lncRNA expression signature for head and neck squamous cell carcinoma.* Scientific reports, 2018. **8**(1): p. 1-13.
153. Corrà, F., et al., *The network of non-coding RNAs in cancer drug resistance.* Frontiers in oncology, 2018. **8**: p. 327.
154. Yang, Z., et al., *Long noncoding RNAs in the progression, metastasis, and prognosis of osteosarcoma.* Cell death & disease, 2016. **7**(9): p. e2389-e2389.
155. Gupta, R.A., et al., *Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis.* Nature, 2010. **464**(7291): p. 1071-1076.
156. Loewen, G., et al., *Functions of lncRNA HOTAIR in lung cancer.* Journal of hematology & oncology, 2014. **7**(1): p. 90.
157. Ishibashi, M., et al., *Clinical significance of the expression of long non-coding RNA HOTAIR in primary hepatocellular carcinoma.* Oncology reports, 2013. **29**(3): p. 946-950.
158. Tang, Q. and S.S. Hann, *HOTAIR: an oncogenic long non-coding RNA in human cancer.* Cellular Physiology and Biochemistry, 2018. **47**(3): p. 893-913.
159. Li, J., et al., *Clinicopathological and prognostic significance of long noncoding RNA MALAT1 in human cancers: a review and meta-analysis.* Cancer cell international, 2018. **18**(1): p. 109.
160. Yu, X., et al., *NEAT 1: A novel cancer-related long non-coding RNA.* Cell proliferation, 2017. **50**(2): p. e12329.
161. Li, S., et al., *Long noncoding RNA GAS5 suppresses triple negative breast cancer progression through inhibition of proliferation and invasion by competitively binding miR-196a-5p.* Biomedicine & Pharmacotherapy, 2018. **104**: p. 451-457.
162. Lucafo, M., et al., *Long noncoding RNA GAS5: a novel marker involved in glucocorticoid response.* Current molecular medicine, 2015. **15**(1): p. 94-99.
163. He, Y., et al., *Potential applications of MEG3 in cancer diagnosis and prognosis.* Oncotarget, 2017. **8**(42): p. 73282.
164. Zou, Y., et al., *Long non-coding PANDAR as a novel biomarker in human cancer: A systematic review.* Cell proliferation, 2018. **51**(1): p. e12422.
165. Zhou, Q., et al., *Long non coding RNA XIST as a prognostic cancer marker–A meta-analysis.* Clinica Chimica Acta, 2018. **482**: p. 1-7.
166. Wang, H., et al., *Long noncoding RNA ANRIL as a novel biomarker of lymph node metastasis and prognosis in human cancer: a meta-analysis.* Oncotarget, 2018. **9**(18): p. 14608.

167. Imam, H., et al., *The lncRNA NRON modulates HIV-1 replication in a NFAT-dependent manner and is differentially regulated by early and late viral proteins.* Scientific reports, 2015. **5**(1): p. 1-10.
168. Carnero, E., et al., *Long noncoding RNA EGOT negatively affects the antiviral response and favors HCV replication.* EMBO reports, 2016. **17**(7): p. 1013-1028.
169. Kopczyńska, M., et al., *PRINS lncRNA Is a new biomarker candidate for HPV infection and prognosis of head and neck squamous cell carcinomas.* Diagnostics, 2020. **10**(10): p. 762.
170. Sharma, S. and K. Munger, *The role of long noncoding RNAs in human papillomavirus-associated pathogenesis.* Pathogens, 2020. **9**(4): p. 289.
171. He, H., et al., *Human papillomavirus E6/E7 and long noncoding RNA TMPOP2 mutually upregulated gene expression in cervical cancer cells.* Journal of virology, 2019. **93**(8): p. e01808-18.
172. Zhang, Y., et al., *The expression and significance of lncRNA HOST2 and microRNA let-7b in HPV-positive cervical cancer tissues and cell lines.* European review for medical and pharmacological sciences, 2019. **23**(6): p. 2380-2390.
173. Barr, J.A., et al., *Long non-coding RNA FAM83H-AS1 is regulated by human papillomavirus 16 E6 independently of p53 in cervical cancer cells.* Scientific reports, 2019. **9**(1): p. 1-11.
174. Yang, M., et al., *Long noncoding RNA CCHE1 promotes cervical cancer cell proliferation via upregulating PCNA.* Tumor Biology, 2015. **36**(10): p. 7615-7622.
175. Sharma, S. and K. Munger, *Expression of the cervical carcinoma expressed PCNA regulatory (CCEPR) long noncoding RNA is driven by the human papillomavirus E6 protein and modulates cell proliferation independent of PCNA.* Virology, 2018. **518**: p. 8-13.
176. Naghashi, N. and S. Ghorbian, *Clinical important dysregulation of long non-coding RNA CCHE1 and HULC in carcinogenesis of cervical cancer.* Molecular biology reports, 2019. **46**(5): p. 5419-5424.
177. Zhang, J., et al., *Long noncoding RNA MEG3 is downregulated in cervical cancer and affects cell proliferation and apoptosis by regulating miR-21.* Cancer biology & therapy, 2016. **17**(1): p. 104-113.
178. Zhang, J. and Y. Gao, *Long non-coding RNA MEG3 inhibits cervical cancer cell growth by promoting degradation of P-STAT3 protein via ubiquitination.* Cancer cell international, 2019. **19**(1): p. 1-10.
179. Gibb, E.A., et al., *Long non-coding RNAs are expressed in oral mucosa and altered in oral premalignant lesions.* Oral oncology, 2011. **47**(11): p. 1055-1061.
180. Feng, L., et al., *Transcriptome analysis reveals differentially expressed lncRNAs between oral squamous cell carcinoma and healthy oral mucosa.* Oncotarget, 2017. **8**(19): p. 31521.
181. Arunkumar, G., et al., *Expression profiling of long non-coding RNA identifies linc-RoR as a prognostic biomarker in oral cancer.* Tumor Biology, 2017. **39**(4): p. 1010428317698366.
182. Shen, Z., et al., *Long non-coding RNA profiling in laryngeal squamous cell carcinoma and its clinical significance: potential biomarkers for LSCC.* PLoS One, 2014. **9**(9): p. e108237.

183. Fang, Z., et al., *Increased expression of the long non-coding RNA UCA1 in tongue squamous cell carcinomas: a possible correlation with cancer metastasis.* Oral surgery, oral medicine, oral pathology and oral radiology, 2014. **117**(1): p. 89-95.
184. Huang, Z.-L., et al., *Long non-coding RNA MEG3 induces cell apoptosis in esophageal cancer through endoplasmic reticulum stress.* Oncology reports, 2017. **37**(5): p. 3093-3099.
185. Kong, X.-p., et al., *The expression and functional role of a FOXC1 related mRNA-lncRNA pair in oral squamous cell carcinoma.* Molecular and cellular biochemistry, 2014. **394**(1-2): p. 177-186.
186. Guo, F., et al., *MALAT1 is an oncogenic long non-coding RNA associated with tumor invasion in non-small cell lung cancer regulated by DNA methylation.* International journal of clinical and experimental pathology, 2015. **8**(12): p. 15903.
187. Zhou, X., et al., *Long non coding RNA MALAT1 promotes tumor growth and metastasis by inducing epithelial-mesenchymal transition in oral squamous cell carcinoma.* Scientific reports, 2015. **5**: p. 15972.
188. Liang, J., et al., *MALAT 1 induces tongue cancer cells' EMT and inhibits apoptosis through Wnt/β-catenin signaling pathway.* Journal of oral pathology & medicine, 2017. **46**(2): p. 98-105.
189. Salyakina, D. and N.F. Tsinoremas, *Non-coding RNAs profiling in head and neck cancers.* NPJ genomic medicine, 2016. **1**(1): p. 1-12.
190. Song, L., et al., *Association of lnc-IL17RA-11 with increased radiation sensitivity and improved prognosis of HPV-positive HNSCC.* Journal of cellular biochemistry, 2019.
191. Cui, J., et al., *An integrated nomogram combining lncRNAs classifier and clinicopathologic factors to predict the recurrence of head and neck squamous cell carcinoma.* Scientific reports, 2019. **9**(1): p. 1-11.
192. Kolenda, T., et al., *EGOT lncRNA in head and neck squamous cell carcinomas.* Polish Journal of Pathology, 2019. **69**(4): p. 356-365.
193. Tang, H., et al., *Salivary lncRNA as a potential marker for oral squamous cell carcinoma diagnosis.* Molecular medicine reports, 2013. **7**(3): p. 761-766.
194. Wang, P., et al., *Long noncoding RNA NEAT1 promotes laryngeal squamous cell cancer through regulating miR-107/CDK6 pathway.* Journal of experimental & clinical cancer research, 2016. **35**(1): p. 1-11.
195. Xu, Y.-z., et al., *The long noncoding RNA FOXCUT promotes proliferation and migration by targeting FOXC1 in nasopharyngeal carcinoma.* Tumor Biology, 2017. **39**(6): p. 1010428317706054.
196. Kong, X.-p., et al., *The expression and functional role of a FOXC1 related mRNA-lncRNA pair in oral squamous cell carcinoma.* Molecular and cellular biochemistry, 2014. **394**(1): p. 177-186.
197. Qian, Y., et al., *Upregulation of the long noncoding RNA UCA1 affects the proliferation, invasion, and survival of hypopharyngeal carcinoma.* Molecular Cancer, 2017. **16**(1): p. 68.
198. Yu, J., et al., *Upregulated long non-coding RNA LINC00152 expression is associated with progression and poor prognosis of tongue squamous cell carcinoma.* J Cancer, 2017. **8**(4): p. 523-530.
199. Chen, M., X. Xu, and H. Ma, *Identification of oncogenic long noncoding RNAs CASC9 and LINC00152 in oral carcinoma through genome-wide comprehensive analysis.* Anti-Cancer Drugs, 2019. **30**(4): p. 356-362.

200. Zheng, X., et al., *LncRNA LINC00152 promotes laryngeal cancer progression by sponging miR-613.* Open Medicine, 2020. **15**(1): p. 240-248.
201. Liu, J., et al., *Decreased expression of pseudogene PTENP1 promotes malignant behaviours and is associated with the poor survival of patients with HNSCC.* Scientific Reports, 2017. **7**(1): p. 41179.
202. Ji, Y., et al., *Long noncoding RNA MEG3 decreases the growth of head and neck squamous cell carcinoma by regulating the expression of miR-421 and E-cadherin.* Cancer Medicine, 2020. **9**(11): p. 3954-3963.
203. Lin, L., X. Liu, and B. Lv, *Long non-coding RNA MEG3 promotes autophagy and apoptosis of nasopharyngeal carcinoma cells via PTEN up-regulation by binding to microRNA-21.* Journal of Cellular and Molecular Medicine, 2021. **25**(1): p. 61-72.
204. Demers, G.W., et al., *Abrogation of growth arrest signals by human papillomavirus type 16 E7 is mediated by sequences required for transformation.* Journal of virology, 1996. **70**(10): p. 6862-6869.
205. Halbert, C., G. Demers, and D. Galloway, *The E7 gene of human papillomavirus type 16 is sufficient for immortalization of human epithelial cells.* Journal of virology, 1991. **65**(1): p. 473-478.
206. Khan, M.R., et al., *The p53-inducible long noncoding RNA TRINGS protects cancer cells from necrosis under glucose starvation.* The EMBO journal, 2017. **36**(23): p. 3483-3500.
207. Khoury, S. and N. Tran, *qPCR multiplex detection of microRNA and messenger RNA in a single reaction.* PeerJ, 2020. **8**: p. e9004.
208. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method.* methods, 2001. **25**(4): p. 402-408.
209. Network, C.G.A., *Comprehensive genomic characterization of head and neck squamous cell carcinomas.* Nature, 2015. **517**(7536): p. 576.
210. Shannon, P., et al., *Cytoscape: a software environment for integrated models of biomolecular interaction networks.* Genome research, 2003. **13**(11): p. 2498-2504.
211. Smoot, M.E., et al., *Cytoscape 2.8: new features for data integration and network visualization.* Bioinformatics, 2010. **27**(3): p. 431-432.
212. Calderone, A., L. Castagnoli, and G. Cesareni, *Mentha: a resource for browsing integrated protein-interaction networks.* Nature methods, 2013. **10**(8): p. 690.
213. Calderone, A., L. Licata, and G. Cesareni, *VirusMentha: a new resource for virus-host protein interactions.* Nucleic acids research, 2014. **43**(D1): p. D588-D592.
214. Stark, C., et al., *BioGRID: a general repository for interaction datasets.* Nucleic acids research, 2006. **34**(suppl_1): p. D535-D539.
215. Huang, D.W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.* Nature protocols, 2008. **4**(1): p. 44.
216. Huang, D.W., B.T. Sherman, and R.A. Lempicki, *Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists.* Nucleic acids research, 2008. **37**(1): p. 1-13.
217. Agarwal, V., et al., *Predicting effective microRNA target sites in mammalian mRNAs.* elife, 2015. **4**: p. e05005.

218. Chou, C.-H., et al., *miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions.* Nucleic acids research, 2017. **46**(D1): p. D296-D302.
219. Wong, N. and X. Wang, *miRDB: an online resource for microRNA target prediction and functional annotations.* Nucleic acids research, 2014. **43**(D1): p. D146-D152.
220. Kent, W.J., et al., *The human genome browser at UCSC.* Genome research, 2002. **12**(6): p. 996-1006.
221. Garcia, M., et al., *Global Cancer Facts & Figures 2007.* American Cancer Society. Atlanta, GA, 2007.
222. Dobrossy, L., *Epidemiology of head and neck cancer: magnitude of the problem.* Cancer Metastasis Rev, 2005. **24**(1): p. 9-17.
223. Gillison, M.L., et al., *Evidence for a causal association between human papillomavirus and a subset of head and neck cancers.* J Natl Cancer Inst, 2000. **92**(9): p. 709-20.
224. Attner, P., et al., *The role of human papillomavirus in the increased incidence of base of tongue cancer.* Int J Cancer, 2010. **126**(12): p. 2879-84.
225. Dahlgren, L., et al., *Human papillomavirus is more common in base of tongue than in mobile tongue cancer and is a favorable prognostic factor in base of tongue cancer patients.* Int J Cancer, 2004. **112**(6): p. 1015-9.
226. Li, W., et al., *Absence of human papillomavirus in tonsillar squamous cell carcinomas from Chinese patients.* Am J Pathol, 2003. **163**(6): p. 2185-9.
227. Stein, A.P., et al., *Prevalence of Human Papillomavirus in Oropharyngeal Cancer: A Systematic Review.* Cancer J, 2015. **21**(3): p. 138-46.
228. Mehanna, H., et al., *Prevalence of human papillomavirus in oropharyngeal and nonoropharyngeal head and neck cancer--systematic review and meta-analysis of trends by time and region.* Head Neck, 2013. **35**(5): p. 747-55.
229. Anantharaman, D., et al., *Combined effects of smoking and HPV16 in oropharyngeal cancer.* Int J Epidemiol, 2016. **45**(3): p. 752-61.
230. Munger, K., et al., *The E6 and E7 genes of the human papillomavirus type 16 together are necessary and sufficient for transformation of primary human keratinocytes.* J Virol, 1989. **63**(10): p. 4417-21.
231. Dyson, N., et al., *The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product.* Science, 1989. **243**(4893): p. 934-7.
232. Howley, P.M., et al., *Molecular mechanisms of transformation by the human papillomaviruses.* Princess Takamatsu Symp, 1989. **20**: p. 199-206.
233. Scheffner, M., et al., *The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53.* Cell, 1990. **63**(6): p. 1129-36.
234. Martinez, I., et al., *Human papillomavirus type 16 reduces the expression of microRNA-218 in cervical carcinoma cells.* Oncogene, 2008. **27**(18): p. 2575-82.
235. Wang, X., et al., *Aberrant expression of oncogenic and tumor-suppressive microRNAs in cervical cancer is required for cancer cell growth.* PLoS One, 2008. **3**(7): p. e2557.
236. Wang, X., et al., *Oncogenic HPV infection interrupts the expression of tumor-suppressive miR-34a through viral oncoprotein E6.* RNA, 2009. **15**(4): p. 637-47.

237. Melar-New, M. and L.A. Laimins, *Human papillomaviruses modulate expression of microRNA 203 upon epithelial differentiation to control levels of p63 proteins.* J Virol, 2010. **84**(10): p. 5212-21.
238. Lewis, B.P., C.B. Burge, and D.P. Bartel, *Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets.* Cell, 2005. **120**(1): p. 15-20.
239. Xie, X., et al., *Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals.* Nature, 2005.
240. Griffiths-Jones, S., et al., *miRBase: tools for microRNA genomics.* Nucleic Acids Res, 2008. **36**(Database issue): p. D154-8.
241. Bartel, D.P., *MicroRNAs: target recognition and regulatory functions.* Cell, 2009. **136**(2): p. 215-33.
242. Khoury, S. and N. Tran, *Circulating microRNAs: potential biomarkers for common malignancies.* Biomark Med, 2015. **9**(2): p. 131-51.
243. Zhang, X., et al., *Alterations in miRNA processing and expression in pleomorphic adenomas of the salivary gland.* Int J Cancer, 2009. **124**(12): p. 2855-63.
244. Zhang, X., et al., *Regulation of the tumour suppressor PDCD4 by miR-499 and miR-21 in oropharyngeal cancers.* BMC Cancer, 2015. **16**(1): p. 86.
245. Wald, A.I., et al., *Alteration of microRNA profiles in squamous cell carcinoma of the head and neck cell lines by human papillomavirus.* Head Neck, 2010.
246. Hong, A.M., et al., *Squamous cell carcinoma of the oropharynx in Australian males induced by human papillomavirus vaccine targets.* Vaccine, 2010. **28**(19): p. 3269-72.
247. Li, W., et al., *The expression of key cell cycle markers and presence of human papillomavirus in squamous cell carcinoma of the tonsil.* Head Neck, 2004. **26**(1): p. 1-9.
248. Lyons, J.G., et al., *Snail up-regulates proinflammatory mediators and inhibits differentiation in oral keratinocytes.* Cancer Res, 2008. **68**(12): p. 4525-30.
249. Cardinali, M., et al., *Tyrosine phosphorylation as a marker for aberrantly regulated growth-promoting pathways in cell lines derived from head and neck malignancies.* Int J Cancer, 1995. **61**(1): p. 98-103.
250. Pang, E., et al., *Radiosensitization of oropharyngeal squamous cell carcinoma cells by human papillomavirus 16 oncoprotein E6*I.* International Journal of Radiation Oncology, 2010. **Accepted for publication 2nd July, 2010. doi:10.1016/j.ijrobp.2010.06.028**.
251. White, J.S., et al., *The influence of clinical and demographic risk factors on the establishment of head and neck squamous cell carcinoma cell lines.* Oral Oncol, 2007. **43**(7): p. 701-12.
252. Saeed, A.I., et al., *TM4: a free, open-source system for microarray data management and analysis.* Biotechniques, 2003. **34**(2): p. 374-8.
253. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.* Methods, 2001. **25**(4): p. 402-8.
254. Kim, S.K., et al., *A gene expression map for Caenorhabditis elegans.* Science, 2001. **293**(5537): p. 2087-92.
255. Beaudenon, S. and J.M. Huibregtse, *HPV E6, E6AP and cervical cancer.* BMC Biochem, 2008. **9 Suppl 1**: p. S4.
256. Tran, N., B.R. Rose, and C.J. O'Brien, *Role of human papillomavirus in the etiology of head and neck cancer.* Head Neck, 2007. **29**(1): p. 64-70.

257. Song, S., et al., *Human papillomavirus types 16 E6 and E7 contribute differently to carcinogenesis.* Virology, 2000. **267**(2): p. 141-50.
258. Yamada, T., et al., *Biologic activity of human papillomavirus type 16 E6/E7 cDNA clones isolated from SiHa cervical carcinoma cell line.* Virus Genes, 1995. **10**(1): p. 15-25.
259. Antinore, M., et al., *The human papillomavirus type 16 E7 gene product interacts with and trans-activates the AP1 family of transcription factors.* The EMBO journal, 1996. **15**(8): p. 1950-1960.
260. Taylor, H.E., et al., *Sterol regulatory element-binding protein 2 couples HIV-1 transcription to cholesterol homeostasis and T cell activation.* Journal of virology, 2011. **85**(15): p. 7699-7709.
261. Huang, D.W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.* Nature protocols, 2009. **4**(1): p. 44-57.
262. Li, W., et al., *Human papillomavirus positivity predicts favourable outcome for squamous carcinoma of the tonsil.* Int J Cancer, 2003. **106**(4): p. 553-8.
263. Sethi, N., et al., *MicroRNAs and head and neck cancer: reviewing the first decade of research.* Eur J Cancer, 2014. **50**(15): p. 2619-35.
264. Longworth, M.S., R. Wilson, and L.A. Laimins, *HPV31 E7 facilitates replication by activating E2F2 transcription through its interaction with HDACs.* EMBO J, 2005. **24**(10): p. 1821-30.
265. Saito, S., et al., *Deregulation and mislocalization of the cytokinesis regulator ECT2 activate the Rho signaling pathways leading to malignant transformation.* J Biol Chem, 2004. **279**(8): p. 7169-79.
266. Slebos, R.J., et al., *Gene expression differences associated with human papillomavirus status in head and neck squamous cell carcinoma.* Clin Cancer Res, 2006. **12**(3 Pt 1): p. 701-9.
267. Iyoda, M., et al., *Epithelial cell transforming sequence 2 in human oral cancer.* PLoS One, 2010. **5**(11): p. e14082.
268. Lal, A., et al., *miR-24 Inhibits cell proliferation by targeting E2F2, MYC, and other cell-cycle genes via binding to "seedless" 3' UTR microRNA recognition elements.* Molecular cell, 2009. **35**(5): p. 610-625.
269. Lajer, C., et al., *The role of miRNAs in human papilloma virus (HPV)-associated cancers: bridging between HPV-related head and neck cancer and cervical cancer.* British journal of cancer, 2012. **106**(9): p. 1526-1534.
270. Horton, J.D., J.L. Goldstein, and M.S. Brown, *SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver.* The Journal of clinical investigation, 2002. **109**(9): p. 1125-1131.
271. Horton, J.D., et al., *Combined analysis of oligonucleotide microarray data from transgenic and knockout mice identifies direct SREBP target genes.* Proceedings of the National Academy of Sciences, 2003. **100**(21): p. 12027-12032.
272. Goldman, M.J., et al., *Visualizing and interpreting cancer genomics data via the Xena platform.* Nature biotechnology, 2020. **38**(6): p. 675-678.
273. HLA-B, N.L., *Integrated genomic and molecular characterization of cervical cancer.* Nature, 2017. **543**: p. 16.
274. Varkonyi-Gasic, E., et al., *Protocol: a highly sensitive RT-PCR method for detection and quantification of microRNAs.* Plant methods, 2007. **3**(1): p. 12.
275. Bommer, G.T. and O.A. MacDougald, *Regulation of lipid homeostasis by the bifunctional SREBF2-miR33a locus.* Cell metabolism, 2011. **13**(3): p. 241-247.

276. Najafi-Shoushtari, S.H., et al., *MicroRNA-33 and the SREBP host genes cooperate to control cholesterol homeostasis.* Science, 2010. **328**(5985): p. 1566-1569.
277. Marquart, T.J., et al., *miR-33 links SREBP-2 induction to repression of sterol transporters.* Proceedings of the national academy of sciences, 2010. **107**(27): p. 12228-12232.
278. Zhong, C., et al., *SREBP2 is upregulated in esophageal squamous cell carcinoma and co-operates with c-Myc to regulate HMGCR expression.* Molecular medicine reports, 2019. **20**(4): p. 3003-3010.
279. Jiang, Y.-Y., et al., *Targeting super-enhancer-associated oncogenes in oesophageal squamous cell carcinoma.* Gut, 2017. **66**(8): p. 1358-1368.
280. Wang, Y., C. Liu, and L. Hu, *Cholesterol regulates cell proliferation and apoptosis of colorectal cancer by modulating miR-33a-PIM3 pathway.* Biochemical and biophysical research communications, 2019. **511**(3): p. 685-692.
281. Wang, H., et al., *miR-33a promotes glioma-initiating cell self-renewal via PKA and NOTCH pathways.* The Journal of clinical investigation, 2014. **124**(10): p. 4489-4502.
282. Karatas, O.F., et al., *miR-33a is a tumor suppressor microRNA that is decreased in prostate cancer.* Oncotarget, 2017. **8**(36): p. 60243.
283. Karataş, Ö.F. and M. Ittmann, *MicroRNA-33a levels do not correlate with the expression of its host gene SREBF2 and its isoforms in prostate cancer cell lines.* The European Research Journal, 2018. **4**(2): p. 79-84.
284. Baselga-Escudero, L., et al., *Resveratrol and EGCG bind directly and distinctively to miR-33a and miR-122 and modulate divergently their levels in hepatic cells.* Nucleic acids research, 2013. **42**(2): p. 882-892.
285. Kim, Y.K. and V.N. Kim, *Processing of intronic microRNAs.* The EMBO journal, 2007. **26**(3): p. 775-783.
286. Sun, Y., et al., *Transcriptome integration analysis in hepatocellular carcinoma reveals discordant intronic miRNA-host gene pairs in expression.* International journal of biological sciences, 2017. **13**(11): p. 1438.
287. He, C., et al., *Young intragenic miRNAs are less coexpressed with host genes than old ones: implications of miRNA–host gene coevolution.* Nucleic acids research, 2012. **40**(9): p. 4002-4012.
288. Gao, X., et al., *Enemy or partner: relationship between intronic micrornas and their host genes.* IUBMB life, 2012. **64**(10): p. 835-840.
289. He, C., et al., *Young intragenic miRNAs are less coexpressed with host genes than old ones: implications of miRNA-host gene coevolution.* Nucleic Acids Res, 2012. **40**(9): p. 4002-12.
290. Schanen, B.C. and X. Li, *Transcriptional regulation of mammalian miRNA genes.* Genomics, 2011. **97**(1): p. 1-6.
291. Veronese, A., et al., *Mutated β-catenin evades a microRNA-dependent regulatory loop.* Proceedings of the National Academy of Sciences, 2011. **108**(12): p. 4840-4845.
292. Ma, N., et al., *Coexpression of an intronic microRNA and its host gene reveals a potential role for miR-483-5p as an IGF2 partner.* Molecular and cellular endocrinology, 2011. **333**(1): p. 96-101.
293. Qiao, Y., et al., *MiR-483-5p controls angiogenesis in vitro and targets serum response factor.* FEBS letters, 2011. **585**(19): p. 3095-3100.

294. von Knebel Doeberitz, M., et al., *Inhibition of tumorigenicity of cervical cancer cells in nude mice by HPV E6-E7 anti-sense RNA.* International journal of cancer, 1992. **51**(5): p. 831-834.
295. Arvanitis, D.A. and D.A. Spandidos, *Deregulation of the G1/S phase transition in cancer and squamous intraepithelial lesions of the uterine cervix: a case control study.* Oncology Reports, 2008. **20**(4): p. 751-760.
296. Yu, Y., et al., *LncRNA SNHG16 induces the SREBP2 to promote lipogenesis and enhance the progression of pancreatic cancer.* Future Oncology, 2019. **15**(33): p. 3831-3844.
297. Li, X., et al., *SREBP-2 promotes stem cell-like properties and metastasis by transcriptional activation of c-Myc in prostate cancer.* Oncotarget, 2016. **7**(11): p. 12869-84.
298. Wen, Y.-A., et al., *Downregulation of SREBP inhibits tumor growth and initiation by altering cellular metabolism in colon cancer.* Cell death & disease, 2018. **9**(3): p. 1-14.
299. Sitarz, K., et al., *The impact of HPV infection on human glycogen and lipid metabolism–a review.* Biochimica et Biophysica Acta (BBA)-Reviews on Cancer, 2022. **1877**(1): p. 188646.
300. Krycer, J.R., et al., *The Akt–SREBP nexus: cell signaling meets lipid metabolism.* Trends in Endocrinology & Metabolism, 2010. **21**(5): p. 268-276.
301. Sitarz, K., et al., *Dual Switch in Lipid Metabolism in Cervical Epithelial Cells during Dysplasia Development Observed Using Raman Microscopy and Molecular Methods.* Cancers, 2021. **13**(9): p. 1997.
302. Munir, R., et al., *Lipid metabolism in cancer cells under metabolic stress.* British journal of cancer, 2019. **120**(12): p. 1090-1098.
303. Igal, R.A., *Stearoyl-CoA desaturase-1: a novel key player in the mechanisms of cell proliferation, programmed cell death and transformation to cancer.* Carcinogenesis, 2010. **31**(9): p. 1509-1515.
304. Yeung, C.L.A., et al., *Human papillomavirus type 16 E6 suppresses microRNA-23b expression in human cervical cancer cells through DNA methylation of the host gene C9orf3.* Oncotarget, 2017. **8**(7): p. 12158.
305. Wong, A.C. and J.E. Rasko, *Splice and Dice: Intronic microRNAs, Splicing and Cancer.* Biomedicines, 2021. **9**(9): p. 1268.
306. Monteys, A.M., et al., *Structure and activity of putative intronic miRNA promoters.* Rna, 2010. **16**(3): p. 495-505.
307. Marsico, A., et al., *PROmiRNA: a new miRNA promoter recognition method uncovers the complex regulation of intronic miRNAs.* Genome biology, 2013. **14**(8): p. 1-23.
308. Sannigrahi, M.K., et al., *Role of host miRNA hsa-mir-139-3p in hpv-16–induced carcinomas.* Clinical cancer research, 2017. **23**(14): p. 3884-3895.
309. Jiménez-Wences, H., O. Peralta-Zaragoza, and G. Fernández-Tilapa, *Human papilloma virus, DNA methylation and microRNA expression in cervical cancer.* Oncology reports, 2014. **31**(6): p. 2467-2476.
310. O'Donnell, K.A., et al., *c-Myc-regulated microRNAs modulate E2F1 expression.* Nature, 2005. **435**(7043): p. 839-843.
311. Woods, K., J.M. Thomson, and S.M. Hammond, *Direct regulation of an oncogenic micro-RNA cluster by E2F transcription factors.* Journal of Biological Chemistry, 2007. **282**(4): p. 2130-2134.

312. Bueno, M.J., et al., *Multiple E2F-induced microRNAs prevent replicative stress in response to mitogenic signaling.* Molecular and cellular biology, 2010. **30**(12): p. 2983-2995.
313. Petrocca, F., et al., *E2F1-regulated microRNAs impair TGFβ-dependent cell-cycle arrest and apoptosis in gastric cancer.* Cancer cell, 2008. **13**(3): p. 272-286.
314. Bhan, A., M. Soleimani, and S.S. Mandal, *Long noncoding RNA and cancer: a new paradigm.* Cancer research, 2017. **77**(15): p. 3965-3981.
315. Jiang, Y., et al., *The role of MALAT1 correlates with HPV in cervical cancer.* Oncology letters, 2014. **7**(6): p. 2135-2141.
316. Shen, C.-J., Y.-M. Cheng, and C.-L. Wang, *LncRNA PVT1 epigenetically silences miR-195 and modulates EMT and chemoresistance in cervical cancer cells.* Journal of drug targeting, 2017. **25**(7): p. 637-644.
317. Lai, S.Y., et al., *Long noncoding RNA SNHG12 modulated by human papillomavirus 16 E6/E7 promotes cervical cancer progression via ERK/Slug pathway.* Journal of cellular physiology, 2020. **235**(11): p. 7911-7922.
318. Sharma, S. and K. Munger, *Expression of the long noncoding RNA dino in human papillomavirus-positive cervical cancer cells reactivates the dormant TP53 tumor suppressor through ATM/CHK2 signaling.* MBio, 2020. **11**(3): p. e01190-20.
319. Raudvere, U., et al., *g: Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update).* Nucleic acids research, 2019. **47**(W1): p. W191-W198.
320. Maddocks, O.D., et al., *Serine starvation induces stress and p53-dependent metabolic remodelling in cancer cells.* Nature, 2013. **493**(7433): p. 542-546.
321. Hu, W., Z. Feng, and A.J. Levine, *The regulation of multiple p53 stress responses is mediated through MDM2.* Genes & cancer, 2012. **3**(3-4): p. 199-208.
322. Lee, S., et al., *A nucleocytoplasmic malate dehydrogenase regulates p53 transcriptional activity in response to metabolic stress.* Cell Death & Differentiation, 2009. **16**(5): p. 738-748.
323. Lin, C.J., et al., *Head and neck squamous cell carcinoma cell lines: established models and rationale for selection.* Head & Neck: Journal for the Sciences and Specialties of the Head and Neck, 2007. **29**(2): p. 163-188.
324. Friedman, J., et al., *Deficient TP53 expression, function, and cisplatin sensitivity are restored by quinacrine in head and neck cancer.* Clinical Cancer Research, 2007. **13**(22): p. 6568-6578.
325. Yeudall, W.A., et al., *Functional characterization of p53 molecules expressed in human squamous cell carcinomas of the head and neck.* Molecular Carcinogenesis: Published in cooperation with the University of Texas MD Anderson Cancer Center, 1997. **18**(2): p. 89-96.
326. van Harten, A.M., et al., *Characterization of a head and neck cancer-derived cell line panel confirms the distinct TP53-proficient copy number-silent subclass.* Oral oncology, 2019. **98**: p. 53-61.
327. Calderone, A., L. Licata, and G. Cesareni, *VirusMentha: a new resource for virus-host protein interactions.* Nucleic acids research, 2015. **43**(D1): p. D588-D592.
328. Li, X.L., et al., *Long noncoding RNA PURPL suppresses basal p53 levels and promotes tumorigenicity in colorectal cancer.* Cell reports, 2017. **20**(10): p. 2408-2423.

329. Zhang, A., M. Xu, and Y.-Y. Mo, *Role of the lncRNA–p53 regulatory network in cancer.* Journal of molecular cell biology, 2014. **6**(3): p. 181-191.
330. Barsotti, A.M., et al., *p53-Dependent induction of PVT1 and miR-1204.* Journal of Biological Chemistry, 2012. **287**(4): p. 2509-2519.
331. Bullock, A.N. and A.R. Fersht, *Rescuing the function of mutant p53.* Nature Reviews Cancer, 2001. **1**(1): p. 68-76.
332. Menendez, D., A. Inga, and M.A. Resnick, *The Biological Impact of the Human Master Regulator p53 Can Be Altered by Mutations That Change the Spectrum and Expression of Its Target Genes.* Molecular and Cellular Biology, 2006. **26**(6): p. 2297-2308.
333. van Houten, V.M., et al., *Mutated p53 as a molecular marker for the diagnosis of head and neck cancer.* The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland, 2002. **198**(4): p. 476-486.
334. Zhou, G., Z. Liu, and J.N. Myers, *TP53 mutations in head and neck squamous cell carcinoma and their impact on disease progression and treatment response.* Journal of cellular biochemistry, 2016. **117**(12): p. 2682-2692.
335. Sarvagalla, S., S.P. Kolapalli, and S. Vallabhapurapu, *The two sides of YY1 in cancer: a friend and a foe.* Frontiers in oncology, 2019. **9**: p. 1230.
336. Li, M., et al., *ATF6 as a transcription activator of the endoplasmic reticulum stress element: thapsigargin stress-induced changes and synergistic interactions with NF-Y and YY1.* Molecular and cellular biology, 2000. **20**(14): p. 5096-5106.
337. Potluri, V., et al., *Transcriptional repression of Bim by a novel YY1-RelA complex is essential for the survival and growth of Multiple Myeloma.* PloS one, 2013. **8**(7): p. e66121.
338. Baritaki, S., et al., *Overexpression of VEGF and TGF-β1 mRNA in Pap smears correlates with progression of cervical intraepithelial neoplasia to cancer: implication of YY1 in cervical tumorigenesis and HPV infection.* International journal of oncology, 2007. **31**(1): p. 69-79.
339. He, G., et al., *YY1 is a novel potential therapeutic target for the treatment of HPV infection-induced cervical cancer by arsenic trioxide.* International Journal of Gynecologic Cancer, 2011. **21**(6).
340. Demers, G.W., C.L. Halbert, and D.A. Galloway, *Elevated wild-type p53 protein levels in human epithelial cell lines immortalized by the human papillomavirus type 16 E7 gene.* Virology, 1994. **198**(1): p. 169-174.
341. Jones, D.L. and K. Münger, *Analysis of the p53-mediated G1 growth arrest pathway in cells expressing the human papillomavirus type 16 E7 oncoprotein.* Journal of virology, 1997. **71**(4): p. 2905-2912.
342. Gooi, Z., J.Y. Chan, and C. Fakhry, *The epidemiology of the human papillomavirus related to oropharyngeal head and neck cancer.* The Laryngoscope, 2016. **126**(4): p. 894-900.
343. Mason, D., K. Munger, and N. Tran, *The dynamic interactome of microRNAs and the human papillomavirus in head and neck cancers.* Current Opinion in Virology, 2021. **51**: p. 87-95.
344. Zhang, X., et al., *Mechanisms and Functions of Long Non-Coding RNAs at Multiple Regulatory Levels.* Int J Mol Sci, 2019. **20**(22).
345. Jalali, S., et al., *Systematic transcriptome wide analysis of lncRNA-miRNA interactions.* PloS one, 2013. **8**(2): p. e53823.
346. Xia, T., et al., *Long noncoding RNA associated-competing endogenous RNAs in gastric cancer.* Scientific reports, 2014. **4**(1): p. 1-7.

347. Gandhi, M., et al., *The lncRNA lincNMR regulates nucleotide metabolism via a YBX1-RRM2 axis in cancer.* Nature communications, 2020. **11**(1): p. 1-15.
348. Gao, C., et al., *The construction and analysis of ceRNA networks in invasive breast cancer: a study based on The Cancer Genome Atlas.* Cancer Manag Res, 2019. **11**: p. 1-11.
349. Xu, Y., et al., *A ceRNA-associated risk model predicts the poor prognosis for head and neck squamous cell carcinoma patients.* Scientific Reports, 2021. **11**(1): p. 6374.
350. Zhou, H., et al., *Identification novel prognostic signatures for Head and Neck Squamous Cell Carcinoma based on ceRNA network construction and immune infiltration analysis.* Int J Med Sci, 2021. **18**(5): p. 1297-1311.
351. Zhou, R.-S., et al., *Integrated analysis of lncRNA-miRNA-mRNA ceRNA network in squamous cell carcinoma of tongue.* BMC Cancer, 2019. **19**(1): p. 779.
352. Wang, Z., et al., *The exploration of new therapeutic targets for HPV-negative head and neck squamous cell cancer through the construction of a ceRNA network and immune microenvironment analysis.* J Cell Biochem, 2020. **121**(5-6): p. 3426-3437.
353. Senkomago, V., et al., *Human papillomavirus–attributable cancers—United States, 2012–2016.* Morbidity and Mortality Weekly Report, 2019. **68**(33): p. 724.
354. Castellsagué, X., et al., *HPV involvement in head and neck cancers: comprehensive assessment of biomarkers in 3680 patients.* Journal of the National Cancer Institute, 2016. **108**(6): p. djv403.
355. John, B., et al., *Human microRNA targets.* PLoS biology, 2004. **2**(11): p. e363.
356. Huang, H.-Y., et al., *miRTarBase 2020: updates to the experimentally validated microRNA–target interaction database.* Nucleic acids research, 2020. **48**(D1): p. D148-D154.
357. Maere, S., K. Heymans, and M. Kuiper, *BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks.* Bioinformatics, 2005. **21**(16): p. 3448-9.
358. Dias, T.R., et al., *Long non-coding RNAs regulate the hallmarks of cancer in HPV-induced malignancies.* Critical Reviews in Oncology/Hematology, 2021. **161**: p. 103310.
359. Gutschner, T. and S. Diederichs, *The hallmarks of cancer.* RNA Biology, 2012. **9**(6): p. 703-719.
360. Andreasen, S., et al., *An update on head and neck cancer: new entities and their histopathology, molecular background, treatment, and outcome.* APMIS, 2019. **127**(5): p. 240-264.
361. Peltekova, V.D., et al., *Identification of genes expressed by immune cells of the colon that are regulated by colorectal cancer-associated variants.* Int J Cancer, 2014. **134**(10): p. 2330-41.
362. Song, N., et al., *Colorectal cancer susceptibility loci and influence on survival.* Genes, Chromosomes and Cancer, 2018. **57**(12): p. 630-637.
363. Closa, A., et al., *Identification of candidate susceptibility genes for colorectal cancer through eQTL analysis.* Carcinogenesis, 2014. **35**(9): p. 2039-2046.
364. Hitomi, Y., et al., *rs1944919 on chromosome 11q23.1 and its effector genes COLCA1/COLCA2 confer susceptibility to primary biliary cholangitis.* Sci Rep, 2021. **11**(1): p. 4557.

365. Qi-Dong, X., et al., *Development and Validation of a Nine-Redox-Related Long Noncoding RNA Signature in Renal Clear Cell Carcinoma.* Oxidative Medicine and Cellular Longevity, 2020. **2020**: p. 6634247.
366. Jinesh, G.G., E.R. Flores, and A.S. Brohl, *Chromosome 19 miRNA cluster and CEBPB expression specifically mark and potentially drive triple negative breast cancers.* PLoS One, 2018. **13**(10): p. e0206008.
367. Vaira, V., et al., *The microRNA cluster C19MC is deregulated in parathyroid tumours.* Journal of molecular endocrinology, 2012. **49**(2): p. 115.
368. Augello, C., et al., *MicroRNA profiling of hepatocarcinogenesis identifies C19MC cluster as a novel prognostic biomarker in hepatocellular carcinoma.* Liver International, 2012. **32**(5): p. 772-782.
369. Kleinman, C.L., et al., *Fusion of TTYH1 with the C19MC microRNA cluster drives expression of a brain-specific DNMT3B isoform in the embryonal brain tumor ETMR.* Nature genetics, 2014. **46**(1): p. 39-44.
370. Flor, I., et al., *Expression of microRNAs of C19MC in different histological types of testicular germ cell tumour.* Cancer genomics & proteomics, 2016. **13**(4): p. 281-289.
371. Yoshitomi, T., et al., *Restoration of miR-517a expression induces cell apoptosis in bladder cancer cell lines.* Oncol Rep, 2011. **25**(6): p. 1661-1668.
372. Ma, W., et al., *miR-517a is an independent prognostic marker and contributes to cell migration and invasion in human colorectal cancer.* Oncol Lett, 2016. **11**(4): p. 2583-2589.
373. Lu, Y.-C., et al., *MiR-520b as a novel molecular target for suppressing stemness phenotype of head-neck cancer by inhibiting CD44.* Scientific Reports, 2017. **7**(1): p. 2042.
374. An, F., et al., *Subpath analysis of each subtype of head and neck cancer based on the regulatory relationship between miRNAs and biological pathways.* Oncol Rep, 2015. **34**(4): p. 1745-1754.
375. Tian, F., et al., *MicroRNA-519a inhibits the proliferation and promotes the apoptosis of ovarian cancer cells through targeting signal transducer and activator of transcription 3.* Exp Ther Med, 2018. **15**(2): p. 1819-1824.
376. Yang, G., et al., *MicroRNA-522 reverses drug resistance of doxorubicin-induced HT29 colon cancer cell by targeting ABCB5.* Mol Med Rep, 2015. **12**(3): p. 3930-3936.
377. Tang, H., et al., *Linc00221 modulates cisplatin resistance in non-small-cell lung cancer via sponging miR-519a.* Biochimie, 2019. **162**: p. 134-143.
378. Piotrowski, I., et al., *miRNAs as Biomarkers for Diagnosing and Predicting Survival of Head and Neck Squamous Cell Carcinoma Patients.* Cancers, 2021. **13**(16): p. 3980.
379. Lopez, Y.O.N., et al., *Characteristic miRNA expression signature and random forest survival analysis identify potential cancer-driving miRNAs in a broad range of head and neck squamous cell carcinoma subtypes.* Reports of Practical Oncology and Radiotherapy, 2018. **23**(1): p. 6-20.
380. Gungormez, C., et al., *Novel miRNAs as potential biomarkers in stage II colon cancer: microarray analysis.* Molecular Biology Reports, 2019. **46**(4): p. 4175-4183.
381. Miller, D.L., et al., *Identification of a Human Papillomavirus–Associated Oncogenic miRNA Panel in Human Oropharyngeal Squamous Cell Carcinoma Validated by Bioinformatics Analysis of The Cancer Genome Atlas.* The American Journal of Pathology, 2015. **185**(3): p. 679-692.

382. Cui, X., et al., *miR-106a Regulates Cell Proliferation and Autophagy by Targeting LKB1 in HPV-16-Associated Cervical Cancer.* Mol Cancer Res, 2020. **18**(8): p. 1129-1141.
383. Zhang, C., et al., *DGCR8/miR-106 Axis Enhances Radiosensitivity of Head and Neck Squamous Cell Carcinomas by Downregulating RUNX3.* Front Med (Lausanne), 2020. **7**: p. 582097.
384. Emmett, S.E., et al., *MicroRNA expression is associated with human papillomavirus status and prognosis in mucosal head and neck squamous cell carcinomas.* Oral Oncol, 2021. **113**: p. 105136.
385. Hui, A.B., et al., *Potentially prognostic miRNAs in HPV-associated oropharyngeal carcinoma.* Clin Cancer Res, 2013. **19**(8): p. 2154-62.
386. Szekerczés, T., et al., *Increased miR-20b Level in High Grade Cervical Intraepithelial Neoplasia.* Pathol Oncol Res, 2020. **26**(4): p. 2633-2640.
387. Cheng, Y., et al., *Human papillomavirus E6-regulated microRNA-20b promotes invasion in cervical cancer by targeting tissue inhibitor of metalloproteinase 2.* Mol Med Rep, 2017. **16**(4): p. 5464-5470.
388. Shen, C., et al., *Variant of SNP rs1317082 at CCSlnc362 (RP11-362K14.5) creates a binding site for miR-4658 and diminishes the susceptibility to CRC.* Cell Death Dis, 2018. **9**(12): p. 1177.
389. Chen, P.P., et al., *LncRNA SNHG12 promotes proliferation and epithelial mesenchymal transition in hepatocellular carcinoma through targeting HEG1 via miR-516a-5p.* Cell Signal, 2021. **84**: p. 109992.
390. Ota, T., et al., *Complete sequencing and characterization of 21,243 full-length human cDNAs.* Nature Genetics, 2004. **36**(1): p. 40-45.
391. Yakushina, V.D., et al., *Long noncoding RNA landscapes specific to benign and malignant thyroid neoplasms of distinct histological subtypes.* Scientific Reports, 2021. **11**(1): p. 1-13.
392. Zhou, J.-Y., et al., *Analysis of microRNA expression profiles during the cell cycle in synchronized HeLa cells.* BMB reports, 2009. **42**(9): p. 593-598.
393. Licata, L., et al., *MINT, the molecular interaction database: 2012 update.* Nucleic acids research, 2012. **40**(D1): p. D857-D861.
394. Kerrien, S., et al., *The IntAct molecular interaction database in 2012.* Nucleic acids research, 2012. **40**(D1): p. D841-D846.
395. Salwinski, L., et al., *The database of interacting proteins: 2004 update.* Nucleic acids research, 2004. **32**(suppl_1): p. D449-D451.
396. Chautard, E., et al., *MatrixDB, the extracellular matrix interaction database.* Nucleic acids research, 2010. **39**(suppl_1): p. D235-D240.
397. White, E.A., et al., *Systematic identification of interactions between host cell proteins and E7 oncoproteins from diverse human papillomaviruses.* Proceedings of the National Academy of Sciences, 2012. **109**(5): p. E260-E267.
398. White, E.A., et al., *Comprehensive analysis of host cellular interactions with human papillomavirus E6 proteins identifies new E6 binding partners and reflects viral diversity.* Journal of virology, 2012. **86**(24): p. 13174-13186.
399. Weiss, R., *Tumour-inducing viruses.* British Journal of Hospital Medicine, 2016. **77**(10): p. 565-568.
400. Moore, P.S. and Y. Chang, *Why do viruses cause cancer? Highlights of the first century of human tumour virology.* Nature reviews cancer, 2010. **10**(12): p. 878.

401. De Martel, C., et al., *Global burden of cancers attributable to infections in 2008: a review and synthetic analysis.* The lancet oncology, 2012. **13**(6): p. 607-615.
402. Barabási, A.L., N. Gulbahce, and J. Loscalzo, *Network medicine: a network-based approach to human disease.* Nat Rev Genet, 2011. **12**(1): p. 56-68.
403. Martínez-Noël, G., et al., *Identification of MicroRNAs That Stabilize p53 in Human Papillomavirus-Positive Cancer Cells.* Journal of Virology, 2022. **96**(4): p. e01865-21.
404. Jung, H.M., B.L. Phillips, and E.K. Chan, *miR-375 activates p21 and suppresses telomerase activity by coordinately regulating HPV E6/E7, E6AP, CIP2A, and 14-3-3ζ.* Molecular cancer, 2014. **13**(1): p. 1-16.
405. Fang, X.-N., et al., *Comprehensive analysis of competitive endogenous RNAs network associated with head and neck squamous cell carcinoma.* Scientific reports, 2018. **8**(1): p. 1-13.
406. Farooq, Q.U.A., et al., *Inferring Virus-Host relationship between HPV and its host Homo sapiens using protein interaction network.* Sci Rep, 2020. **10**(1): p. 8719.
407. Zheng, Z.-M. and X. Wang, *Regulation of cellular miRNA expression by human papillomaviruses.* Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms, 2011. **1809**(11): p. 668-677.
408. Harden, M.E., et al., *Modulation of microRNA-mRNA Target Pairs by Human Papillomavirus 16 Oncoproteins.* mBio, 2017. **8**(1): p. e02170-16.
409. Corona-Meraz, F.-I., et al., *Ageing influences the relationship of circulating miR-33a and miR-33b levels with insulin resistance and adiposity.* Diabetes and Vascular Disease Research, 2019. **16**(3): p. 244-253.
410. Harden, M.E. and K. Munger, *Human papillomavirus 16 E6 and E7 oncoprotein expression alters microRNA expression in extracellular vesicles.* Virology, 2017. **508**: p. 63-69.
411. Berti, F.C.B., et al., *From squamous intraepithelial lesions to cervical cancer: Circulating microRNAs as potential biomarkers in cervical carcinogenesis.* Biochimica et Biophysica Acta (BBA) - Reviews on Cancer, 2019. **1872**(2): p. 188306.
412. Ramqvist, T., et al., *Studies on human papillomavirus (HPV) 16 E2, E5 and E7 mRNA in HPV-positive tonsillar and base of tongue cancer in relation to clinical outcome and immunological parameters.* Oral Oncology, 2015. **51**(12): p. 1126-1131.
413. Ren, S., et al., *HPV E2, E4, E5 drive alternative carcinogenic pathways in HPV positive cancers.* Oncogene, 2020. **39**(40): p. 6327-6339.
414. Baedyananda, F., A. Chaiwongkot, and P. Bhattarakosol, *Elevated HPV16 E1 Expression Is Associated with Cervical Cancer Progression.* Intervirology, 2017. **60**(5): p. 171-180.
415. Liao, S., et al., *Human papillomavirus 16/18 E5 promotes cervical cancer cell proliferation, migration and invasion in vitro and accelerates tumor growth in vivo.* Oncol Rep, 2013. **29**(1): p. 95-102.
416. Liu, C., et al., *HPV16 early gene E5 specifically reduces miRNA-196a in cervical cancer cells.* Scientific Reports, 2015. **5**(1): p. 7653.
417. Beaudenon, S. and J.M. Huibregtse, *HPV E6, E6AP and cervical cancer.* BMC Biochemistry, 2008. **9**(1): p. S4.

418. Jones, D.L., D.A. Thompson, and K. Münger, *Destabilization of the RB tumor suppressor protein and stabilization of p53 contribute to HPV type 16 E7-induced apoptosis.* Virology, 1997. **239**(1): p. 97-107.
419. Sharma, S., K. Munger, and L. Banks, *KDM6A-Mediated Expression of the Long Noncoding RNA DINO Causes TP53 Tumor Suppressor Stabilization in Human Papillomavirus 16 E7-Expressing Cells.* Journal of Virology, 2020. **94**(12): p. e02178-19.
420. Li, C., et al., *Antiviral activity of ISG15 against classical swine fever virus replication in porcine alveolar macrophages via inhibition of autophagy by ISGylating BECN1.* Vet Res, 2020. **51**(1): p. 22.
421. Mathieu, N.A., et al., *HERC5 and the ISGylation Pathway: Critical Modulators of the Antiviral Immune Response.* Viruses, 2021. **13**(6).
422. Zhang, J., et al., *Cancer specific long noncoding RNAs show differential expression patterns and competing endogenous RNA potential in hepatocellular carcinoma.* PloS one, 2015. **10**(10): p. e0141042.
423. Zhou, X., J. Liu, and W. Wang, *Construction and investigation of breast-cancer-specific ceRNA network based on the mRNA and miRNA expression data.* IET systems biology, 2014. **8**(3): p. 96-103.
424. Zhou, M., et al., *Construction and analysis of dysregulated lncRNA-associated ceRNA network identified novel lncRNA biomarkers for early diagnosis of human pancreatic cancer.* Oncotarget, 2016. **7**(35): p. 56383.
425. Pan, Y., et al., *Analysis of lncRNA-Mediated ceRNA Crosstalk and Identification of Prognostic Signature in Head and Neck Squamous Cell Carcinoma.* Frontiers in Pharmacology, 2019. **10**.
426. Tommasino, M. and L. Crawford, *Human papillomavirus E6 and E7: proteins which deregulate the cell cycle.* Bioessays, 1995. **17**(6): p. 509-518.
427. Shai, A., et al., *The human papillomavirus E6 oncogene dysregulates the cell cycle and contributes to cervical carcinogenesis through two independent activities.* Cancer research, 2007. **67**(4): p. 1626-1635.
428. Zhang, L., et al., *The role of the PI3K/Akt/mTOR signalling pathway in human cancers induced by infection with human papillomaviruses.* Molecular cancer, 2015. **14**(1): p. 1-13.
429. Li, Y., et al., *Long Noncoding RNA SChLAP1 Accelerates the Proliferation and Metastasis of Prostate Cancer via Targeting miR-198 and Promoting the MAPK1 Pathway.* Oncology research, 2018. **26**(1): p. 131-143.
430. Prensner, J.R., et al., *The long noncoding RNA SChLAP1 promotes aggressive prostate cancer and antagonizes the SWI/SNF complex.* Nature Genetics, 2013. **45**(11): p. 1392-1398.
431. Bai, X., et al., *Long non-coding RNA SChLAP1 regulates the proliferation of triple negative breast cancer cells via the miR-524-5p/HMGA2 axis.* Mol Med Rep, 2021. **23**(6): p. 446.
432. Xia, L., et al., *Oncogenic miR-20b-5p contributes to malignant behaviors of breast cancer stem cells by bidirectionally regulating CCND1 and E2F1.* BMC Cancer, 2020. **20**(1): p. 949.
433. Yang, H., et al., *miR-20b-5p functions as tumor suppressor microRNA by targeting cyclinD1 in colon cancer.* Cell Cycle, 2020. **19**(21): p. 2939-2954.
434. Lin, N., et al., *Identification and validation of a five-lncRNA signature for predicting survival with targeted drug candidates in ovarian cancer.* Bioengineered, 2021. **12**(1): p. 3263-3274.

435. Wang, H., et al., *LncRNAs expression profiling in normal ovary, benign ovarian cyst and malignant epithelial ovarian cancer.* Scientific Reports, 2016. **6**(1): p. 38983.
436. Liu, B., et al., *The regulatory role of antisense lncRNAs in cancer.* Cancer Cell International, 2021. **21**(1): p. 459.
437. Dogan, T., et al., *Role of the E3 ubiquitin ligase RNF157 as a novel downstream effector linking PI3K and MAPK signaling pathways to the cell cycle.* Journal of Biological Chemistry, 2017. **292**(35): p. 14311-14324.
438. Bridges, M.C., A.C. Daulagala, and A. Kourtidis, *LNCcation: lncRNA localization and function.* Journal of Cell Biology, 2021. **220**(2).