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.