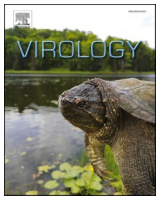
Virology 600 (2024) 110220

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/virology



The lncRNA and miRNA regulatory axis in HPV16-positive oropharyngeal cancers

Dayna Sais [a,*], Meredith Hill [b], Fiona Deutsch [a], Phuong Thao Nguyen [c], Valerie Gay [d], Nham Tran [a,**]

[a] School of Biomedical Engineering, Faculty of Engineering and Information Technology, The University of Technology Sydney, Australia
[b] Graduate School of Biomedical Engineering, University of New South Wales, Australia
[c] Transdisciplinary School, The University of Technology Sydney, Australia
[d] School of Electrical and Data Engineering, Faculty of Engineering and Information Technology, The University of Technology Sydney, Australia

ARTICLE INFO

Handling Editor: Jasmine Tomar

Keywords:
Oropharyngeal cancer (OPC)
Human papillomavirus (HPV) type 16
Long ncRNAs
microRNA and messenger RNA interactomes
RNA regulation

ABSTRACT

The global rise of oropharyngeal cancers (OPC) associated with the human papillomavirus (HPV) type 16 necessitates a deeper understanding of their underlying molecular mechanisms. Our study utilised RNA-sequencing data from The Cancer Genome Atlas (TCGA) to identify and analyse differentially expressed (DE) long non-coding RNAs (lncRNAs), microRNAs (miRNAs), and messenger RNAs (mRNAs) in HPV16-positive OPC, and to elucidate the interplay within the lncRNA/miRNA/mRNA regulatory network. We revealed 1929 DE lncRNAs and identified a significant expression shift in 37 of these, suggesting a regulatory 'sponge' function for miRNAs that modulate cellular processes. Notably, the lncRNA Linc00911 exhibited decreased expression in HPV16-positive OPC, a change directly attributable to HPV oncogenes E6 and E7 as confirmed by RT-qPCR in cell lines and patient samples. Our comprehensive analysis presents an expansive landscape of ncRNA-mRNA interactions, offering a resource for the ongoing pursuit of elucidating the molecular underpinnings of HPV-driven OPC.

1. Introduction

Oropharyngeal cancers, a subtype of head and neck cancers (HNCs), typically develop as squamous cell carcinomas (SCC). These cancers originate in areas such as the tonsils, base of the tongue, soft palate, and the pharyngeal wall (Tumban, 2019). The prevalence of oropharyngeal cancers (OPC) linked to human papillomavirus 16 (HPV16) is increasing, making it one of the most prevalent HPV-associated cancers in Western countries (Gooi et al., 2016; Chaturvedi et al., 2011). Unlike other subtypes of HNC, HPV16-positive OPCs are distinct and are considered a separate clinical group (Gillison et al., 2000).

Human papillomavirus 16 is characterised by two key viral oncogenes, E6 and E7; their combined activity influences various cellular pathways, ultimately contributing to tumour development (Moody et al., 2010). HPV16 and its oncogenes can also impact distinct RNA regulators with the non-coding RNAs (ncRNAs), further impacting the molecular landscape of the disease (Casarotto et al., 2020; Sais et al., 2021).

Non-coding RNAs (ncRNAs) are critical regulators of gene expression, playing a pivotal role in maintaining normal cellular functions. Disrupting the levels of these ncRNAs is often associated with diseased states (Nemeth et al., 2024). In the context of HPV16-positive cancers, several studies have focused on how the virus affects microRNA (miRNA/miRs) expression. These miRNAs, which are short (20–22 nucleotides) single-stranded RNAs, bind to the 3′ untranslated region (UTR) of their target messenger RNA (mRNA) to suppress mRNA expression (Bartel, 2004; Ambros, 2004). In HPV16-positive OPC, there is a distinct pattern of miRNA expression compared to HPV16-negative OPC (Sais et al., 2021). This aberrant miRNA expression is linked to various cancer hallmarks, including cell cycle progression, apoptosis, and metastasis (Goodall et al., 2021; Dhawan et al., 2018). Long non-coding RNAs (lncRNAs), another type of ncRNA, are RNA transcripts over 200 nucleotides long without coding potential. These transcripts can interact with RNA and DNA through base pairing or form structures for protein binding and sequestration (Zhang et al., 2019). Regarding OPC, there is limited understanding of the mechanisms of how lncRNAs' can interact

* Corresponding author.
** Corresponding author.
E-mail addresses: Dayna.Sais@uts.edu.au (D. Sais), Nham.Tran@uts.edu.au (N. Tran).

https://doi.org/10.1016/j.virol.2024.110220
Received 3 June 2024; Received in revised form 18 August 2024; Accepted 2 September 2024
Available online 5 September 2024

and regulate other RNA families.

Both the long non-coding and microRNAs are often viewed as independent regulators. However, recent findings suggest that these two types of ncRNAs can interact, introducing a new layer of complexity to gene regulation. LncRNAs, for example, can act as "sponges" for miRNAs. By binding to miRNAs, lncRNAs prevent them from regulating their target mRNAs, effectively directing lncRNAs as positive regulators of mRNA transcription (Yang et al., 2022). This interaction, known as the competitive endogenous RNA (ceRNA) network and has been observed in various cancers (Xia et al., 2014; Gao et al., 2019), primarily involves lncRNAs using their miRNA response elements (MREs) near the 3′ end, which are complementary to Argonaute (Ago) binding sites (Jalali et al., 2013).

The relationship between long non-coding RNAs and microRNAs forms a broad RNA regulatory network that can influence the entire transcriptome. This interaction is particularly relevant to OPC, where human papillomavirus (HPV16) is implicated in 50–70% of cases (Chaturvedi et al., 2011; Schache et al., 2016). Of these, up to 85% are associated with the high-risk HPV type 16 variant (Castellsagué et al., 2016). To investigate the influence of HPV, especially the HPV16 variant, on lncRNA and miRNA interactions in OPC, we conducted a detailed analysis.

Our research analysed RNA-sequencing data from The Cancer Genome Atlas (TCGA), identifying differentially expressed lncRNAs, miRNAs, and mRNAs in cases of HPV16-positive OPC. We applied a comprehensive bioinformatics approach to investigate the specific roles of lncRNAs, particularly their function as miRNA sponges. We then validated our findings using samples from OPC patients. Subsequently, we constructed an interactome of lncRNA/miRNA/mRNA to highlight potential new pathways involved in OPC transformation. These insights could enhance our understanding of the molecular pathways associated with OPC and potentially uncover new networks and mechanisms in HPV16-driven cancers.

2. Methods

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 downloaded RNA-seq BAM files of 49 patients, all of which were primary oropharyngeal tumours (tonsil, base of tongue, oropharynx) from the HNC cohort. The OPC was also filtered for those tested for HPV16 by ISH or identified to contain HPV16 by detecting HPV E6/E7 via RNA sequencing done by TCGA (Network, 2015); all other HPV types were excluded from this study. In this study, 34 samples were identified as HPV16-positive and 15 as HPV16-negative. Samples were annotated according to their Barcode ID and available clinical information, including tumour site, tumour stage, sex, HPV status, age, clinical stage and alcohol and smoking history (Supplementary Table 1). MiRNA-sequencing data was also downloaded as raw count files from the TCGA. Total RNA-seq data was 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 (DE) analysis of individual mRNAs, lncRNAs and miRNAs was carried out 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 DE mRNA, lncRNAs and miRNAs between HPV16-positive and HPV16-negative 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 Log2 Fold Change 5 (Log2FC5) cut-off was applied to the lncRNAs and mRNAs, and a Log2 Fold Change 2 (Log2FC2) cut-off was applied to the miRNAs. The methodology for TCGA RNA-sequencing data analysis is described in Fig. 1.

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

To identify lncRNA-miRNA sponging interactions, three online tools were utilised, Diana tools LncBase 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 $\geq$150 and a minimum free energy less than −20Cal/mol to identify reliable targets (John et al., 2004).

Each of our DE lncRNAs (Log2FC $\geq$ 5) were analysed using each of these databases or tools, and their miRNA targets were identified. The lncRNA-miRNA targets were filtered to only include those that were DE (miRNAs with a Log2FC $\geq$ 2) in our TCGA RNA-sequencing dataset. The lncRNA-miRNA targets were filtered again to include miRNAs with an inverse DE to the lncRNAs. The mRNA targets of these miRNAs were then identified using MiRTarBase (Huang et al., 2020), a database that describes experimentally validated interactions. The DE miRNAs-mRNA targets were filtered for those with an inverse DE pattern in the TCGA RNA-seq dataset (mRNAs with a Log2FC $\geq$ 2). The DE lncRNAs and their DE miRNA targets were imported into the network mapping program Cystoscope software v3.9. The DE miRNA-DE mRNA targets were also added to this interactome for visualisation and network analysis.

2.3. Gene ontology and KEGG analysis

To gain a biological understanding of those gene sets statistically significantly associated with HPV16-positive OPC, we conducted pathway enrichment and disease association analyses using various programs. The BiNGO (Maere et al., 2005) cystoscope software was utilised for gene ontology analysis and g:Profiler (https://biit.cs.ut.ee/gprofiler/gost) for KEGG analysis.

2.4. Real-time quantitative-PCR (qPCR) and statistical analysis

RNA harvested from three populations of immortalised Human Foreskin Keratinocytes (iHFK) were a gift from Prof Karl Munger (Tufts University). RNA harvested included control cells and HPV16 E6, E7 and E6/E7 expressing iHFKs, all created with recombinant retroviruses (Demers et al., 1996; Halbert et al., 1991; Sharma et al., 2020).

RNA from these cells were used to validate DE lncRNAs identified in RNA sequencing results.

Expression of genes and lncRNAs was determined via RT-qPCR. Complementary DNA was generated with the Hi-Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific, USA) with 500 ng of total RNA input. The cDNA was diluted in a 1:4 ratio for a 5.0 μL qPCR reaction. All qPCR reactions (including non-RT and no template controls) were conducted in triplicate using the StepOne PCR systems (Applied BioSystems, Australia) at the recommended cycle conditions using the TaqMan Universal PCR Master Mix, No AmpErase UNG (2x) (Thermo Fisher Scientific, USA) or the PowerUp SYBR green master Mix (Thermo Fisher Scientific, USA).

TaqMan assays were employed to determine the expression of the target genes HPV16 E6 and HPV16 E7 (Deutsch et al., 2024), and the endogenous reference gene for these assays was β2M (Hs00187842_m1) labelled with a VIC™ probe. The SYBR green assay was utilised to determine the expression levels of lncRNAs identified in RNA sequencing data (custom design), and the endogenous reference gene for these assays was GAPDH. All primer sequences for SYBR green assays are described in Supplementary Table 2. Only LINC00911 yielded reliable qPCR results after optimisation. Mean Ct values were normalised

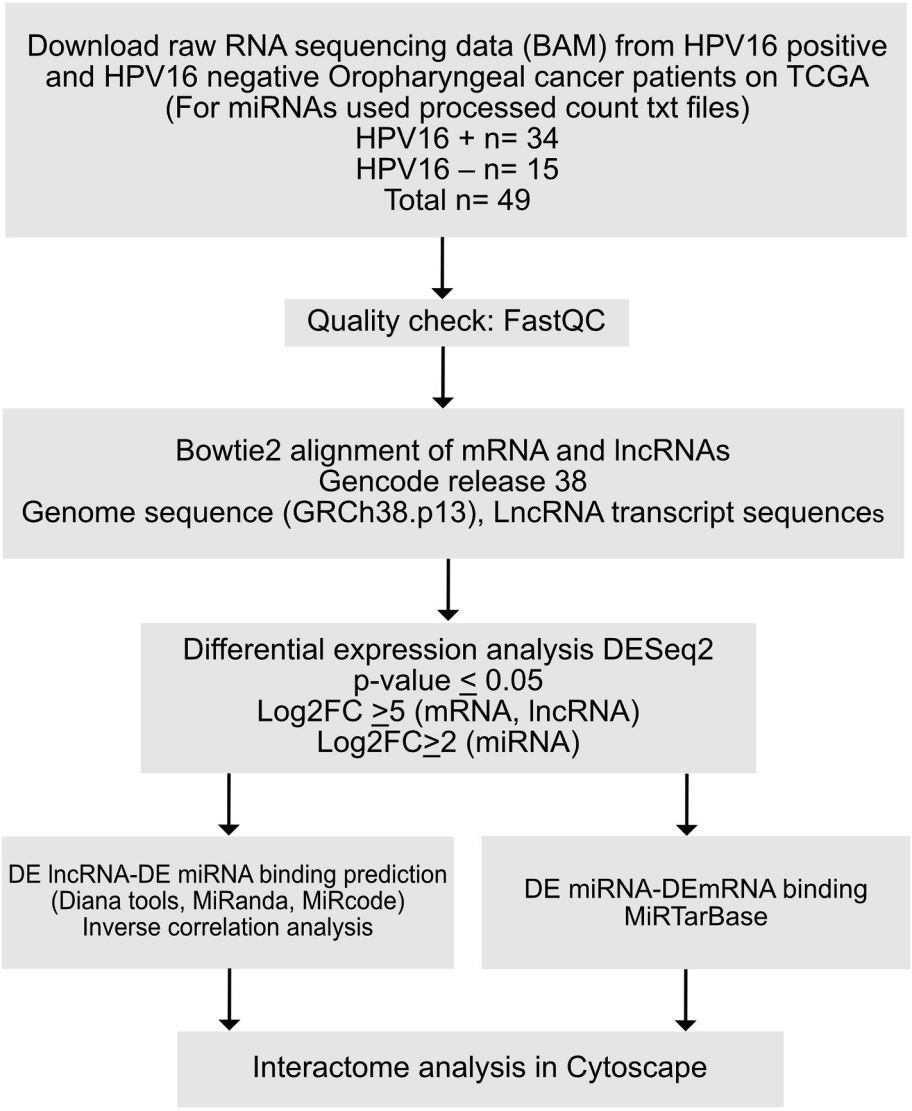


Fig. 1. Analysis flow chart: Flow diagram representing the methodology for processing RAW RNA-seq data from TCGA and development of lncRNA-miRNA-mRNA interactome.

using β2M for TaqMan and GAPDH for SYBR. The relative expression level was calculated using $2^{-\Delta\Delta CT}$ (Livak et al., 2001) and presented as fold change. A two-tailed Student T-test evaluated the significance between gene expression levels.

3. Results

3.1. Identifying differentially expressed lncRNAs, miRNAs and mRNAs between HPV16-positive and HPV16-negative oropharyngeal cancers

We utilised the data from 49 OPC samples deposited in the TCGA database. These were RNA sequencing data generated by the Illumina HiSeq 2000. Samples were filtered for HPV16-positive (n = 34) or HPV16-negative status (n = 15). Differential expression analysis using Deseq2, for lncRNA revealed 1929 transcripts to be differentially expressed (DE) between the HPV16-positive to the HPV16-negative tumours. Of these, 1176 were upregulated, and 753 (p-value of 0.05 or less) were downregulated in the HPV16-positive tumours (Fig. 2A) and (Supplementary Table 4).

To identify the lncRNA transcripts most affected by HPV16, we utilised a Log2 Fold Change (Log2FC) threshold of 5. This analysis identified 37 lncRNAs with differential expression in HPV16-positive OPC compared to HPV16-negative samples. Among these, 23 were upregulated and 14 downregulated in HPV16-positive tumours (Fig. 2B and Supplementary Table 4). The most significantly upregulated lncRNA was the newly identified ENSG00000282816.1, exhibiting a Log2FC of 6.5. This RNA acts as the antisense to the target NEFH (neurofilament heavy chain). The lncRNA ENSG00000231683.6, which is antisense to GCLA (glutamine-cysteine ligase catalytic subunit), showed the highest downregulation with a Log2FC of −7.4. Among the 37 most DE lncRNAs, there are 10 antisense lncRNAs, 13 long intergenic non-coding RNAs, and 14 novel transcripts that have not yet been characterised.

We then proceeded to assess the expression profiles of microRNAs and mRNAs in the same OPC dataset. In this part of our study, we identified 161 miRNAs (Fig. 3A and B) and 9836 mRNAs (Fig. 4A and B) that were DE in HPV16-positive compared to HPV-negative tumours.

Applying a Log2FC cut-off of 2 for miRNAs to identify those significantly influenced by HPV16, we discovered 32 miRNAs to be DE (Fig. 3B). Among these, 14 miRNAs were upregulated, including miR-125b(Log2FC 3.9), miR-6510(Log2FC 3.8), miR-20b(Log2FC 3.5), and miR-499a(Log2FC 3.4). In comparison, 18 miRNAs were downregulated, with the most significantly downregulated being miR-516a-1 (Log2FC −4.6), miR-116a-2(Log2FC −4.4), miR-519a-1(Log2FC −4.0), and miR-522(Log2FC −3.0).

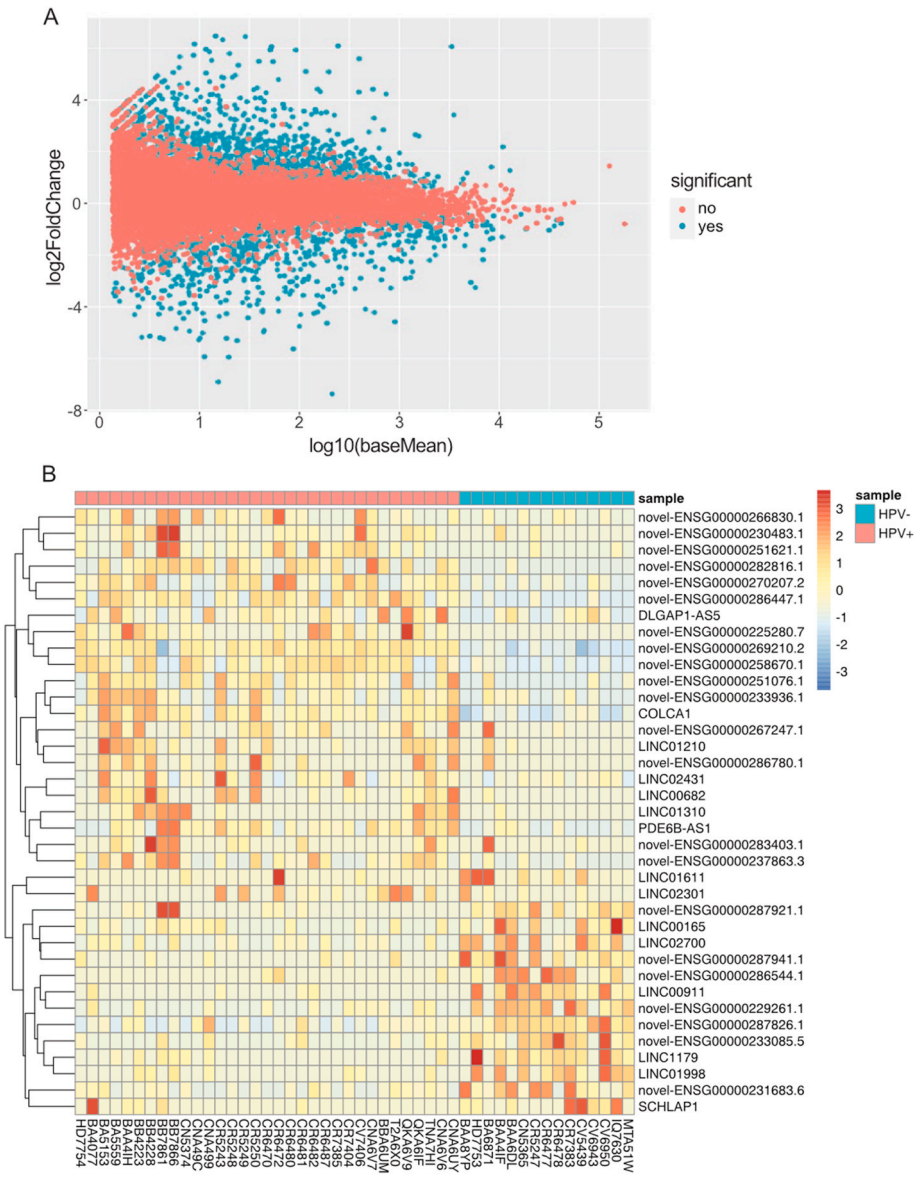


Fig. 2. Differentially expressed lncRNAs between HPV16-positive and HPV-negative TGCA OPC patients. A) MA plot representing the differences between HPV16-positive and HPV-negative OPC patients, comparing Log2 Fold change on the Y axis and Log10 of the mean average normalised counts. B) A constructed heat map showing the differentially expressed lncRNAs between HPV-positive (pink bar) and HPV-negative (blue bar) oropharyngeal tumours. 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 Log2FC5.

A key observation from our analysis was that many of the down-regulated miRNAs were part of the miRNA cluster on chromosome 19 (C19MC), one of the largest miRNA clusters containing 46 miRNAs. This includes 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, and miR-516a-2.

For mRNAs, using the same Log2FC cut-off of 5 to identify those most significantly affected by HPV16, we found 60 mRNAs to be DE (Fig. 4B), with 39 upregulated and 21 downregulated. The most significantly upregulated mRNAs were FEZF2(Log2FC 22.4), GBX1(Log2FC 8.5), NKX2-4(Log2FC 8.1), and ZPBP2(Log2FC 7.9). Whilst the most significantly downregulated mRNAs were PRR9(Log2FC −7.1), FOLR3 (Log2FC −6.6), CNHB1(Log2FC 6.2), and LCE1A(Log2FC −6.1). Additionally, our findings revealed that one of the mRNAs, NKX2-4, which was upregulated, had an associated antisense lncRNA, ENSG00000225280.7, also differentially expressed in HPV16-positive OPC tumours.

3.2. RT-qPCR validation of Linc00911 using HPV16 E6 and E7 expressing iHFKS

To investigate whether the expression of the dysregulated lncRNAs in HPV-positive OPC patients was directly influenced by HPV16 E6 and E7, we used iHFKs expressing either E6, E7, or both (Supplementary Fig. 1). We decided to focus on the top four upregulated (novel-ENSG00000231683.6, Linc00682, Linc02301, and novel-ENSG00000282816.1) and downregulated (novel-ENSG00000231683.6, Linc00911, Linc01179, and novel-ENSG00000233085.5).

Primers were designed for all targets and tested across various cell lines (SCC090, iHFK and UMSCC22B). However, only the Linc00911 primers could amplify the expected amplicon (Supplementary Fig. 2). Using this set of primers, we evaluated the relative expression levels of Linc00911 in iHFKs overexpressing HPV16 E6 and E7, Fig. 5A. We found that iHFKs expressing either E7 alone or the E6/E7 combination showed

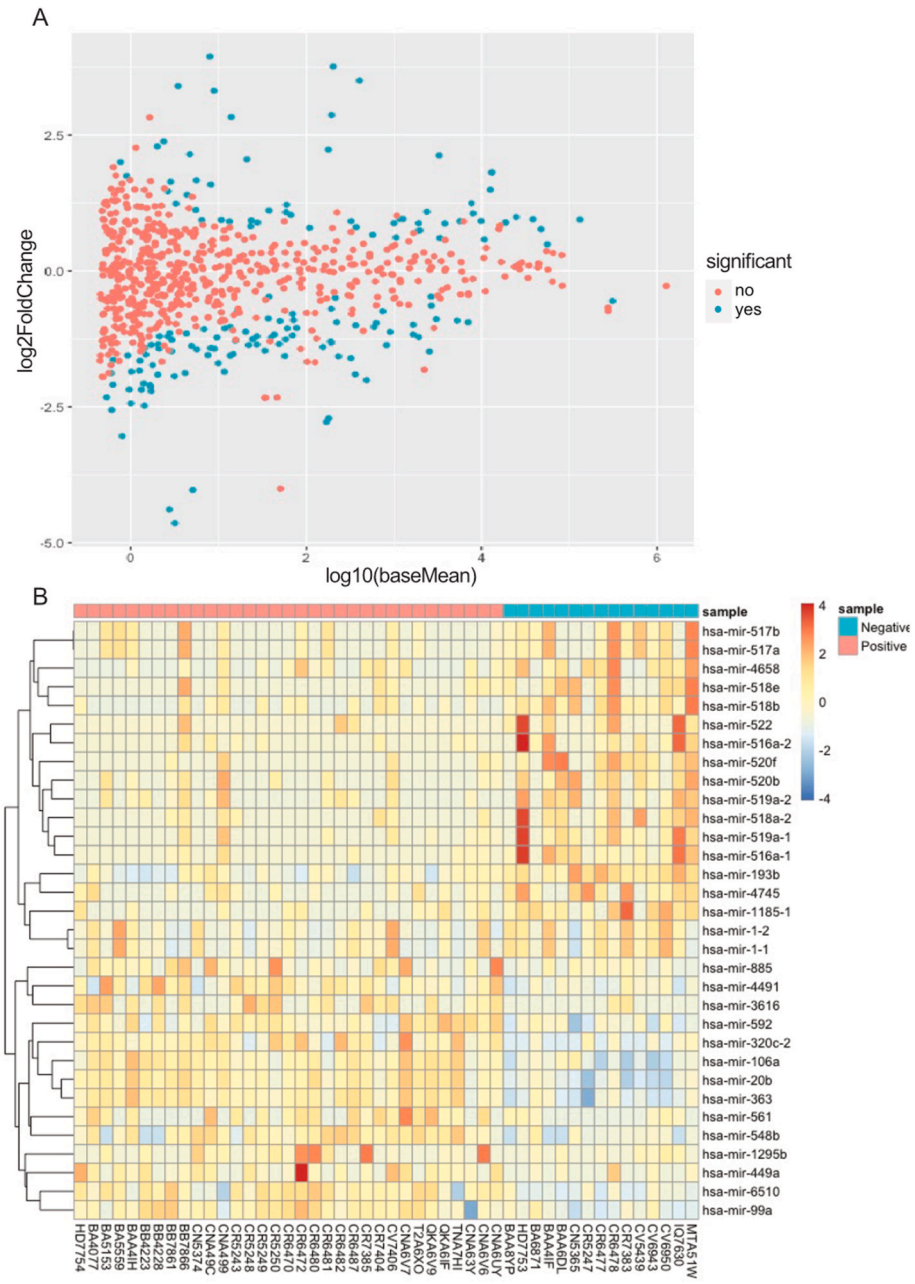


Fig. 3. Differentially expressed miRNAs between HPV16-positive and HPV-negative OPC patients. A) MA plot representing the differences between miRNAs expression in HPV16-positive and HPV-negative OPC patients, comparing Log2 Fold change on the Y axis and Log10 of the mean average normalised counts. B) A constructed heat map displaying the differentially expressed miRNAs between HPV16-positive (pink bar) and HPV-negative (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.

a significant decrease in Linc00911 expression compared to control cells. Although there was a reduction in Linc00911 expression in E6-expressing cells, this change was not statistically significant. We then measured the expression of Linc00911 in a patient cohort of nine HPV16-positive and twelve HPV16-negative samples (Fig. 5B). This patient cohort is previously described in Sais et al. (2018) (Supplementary Table 3). There was a lower RNA expression of Linc00911 in the HPV16-positive samples. To further support this trend, we then mined the TCGA data, and this showed that the normalised counts for Linc00911 were again lower for HPV16-positive patients (Fig. 5C).

3.3. Building an interactome of differentially expressed lncRNAs-miRNAs-mRNAs in HPV16 OPC

We explored the potential of DE lncRNAs to act as regulators or 'sponges' for miRNAs, by employing bioinformatics tools to predict miRNA binding sites on these lncRNAs. We specifically focused on lncRNA-miRNA pairs where the miRNAs showed DE Log2FC $>$ 2 in HPV16-positive OPC (Supplementary Table 5). These interactions were then visualised using the network mapping program Cytoscape software v3.3, creating a distinct interactome for upregulated (Fig. 6A) and downregulated lncRNAs (Fig. 6B).

Of the 37 DE lncRNAs analysed, novel-ENSG00000283403 emerged as having the most DE miRNA targets (12 upregulated and 16

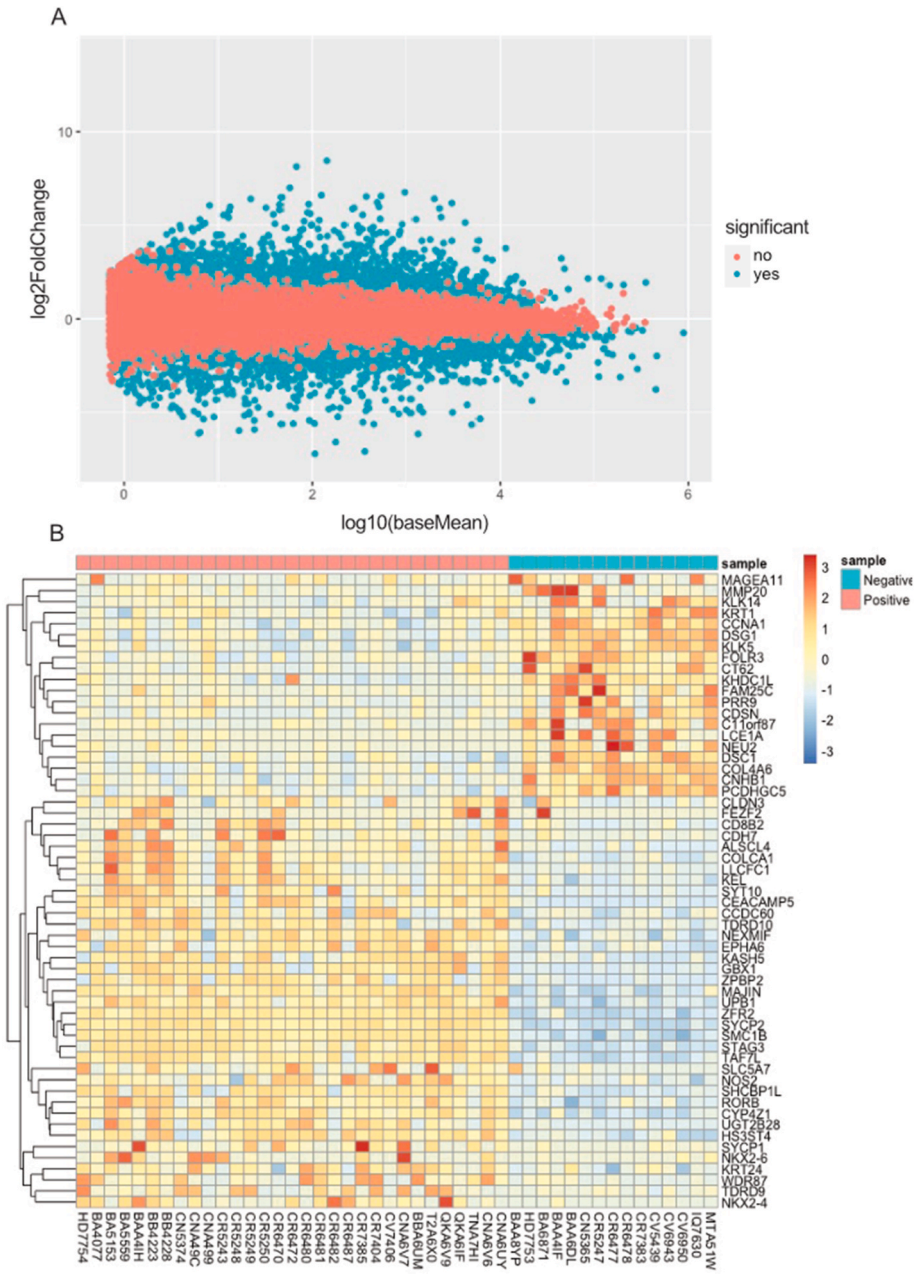


Fig. 4. Differentially expressed mRNAs between HPV16-positive and HPV-negative OPC patients. A) MA plot representing the differences between mRNAs expression in HPV16-positive and HPV-negative OPC patients, comparing Log2 Fold change on the Y axis and Log10 of the mean average normalised counts. B) A constructed heat map displaying the differentially expressed mRNAs between HPV16-positive (pink bar) and HPV-negative (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.

downregulated). When filtering for lncRNA-miRNA interactions with inverse expression patterns, novel-ENSG00000287941 (downregulated) was observed to sponge the highest number of upregulated miRNAs (n = 13), while ENSG00000283403 (upregulated) targeted the most downregulated miRNAs (n = 16). Among the 32 most DE miRNAs, hsa-miR-6510-5p was the most frequently sponged, targeted by 25 different lncRNAs.

We then predicted the mRNA targets of these miRNAs using MiRTarBase, a database of experimentally validated miRNA-target interactions. We filtered the DE miRNAs and their mRNA targets for those exhibiting an inverse differential expression pattern in the TCGA RNA-seq dataset (mRNAs with a Log2FC > 2). These DE miRNA and mRNA targets were also included in our interactome for further visualisation and network analysis. Notably, CDK6 was targeted by the highest number of miRNAs (miR-20b-5p, miR-4491, miR-499a, and miR-561-5p), while the upregulated targets NCAN and BEND4 were targeted by the most miRNAs.

Lastly, we examined all mRNAs using the VirusMentha database to determine if any were directly targeted by HPV16 viral genes. We found that only two mRNAs, SLC7A5 (targeted by miR-4757-5p) and CAV1 (targeted by miR-106b-5p and miR-20b-5p), are known targets of the viral gene E5. Interestingly, the HPV16, E6 and E7, did not directly regulate any of the sponged miRNA-mRNA targets in our interactomes. This suggests that the altered expression of these mRNAs might be attributed to the complex interplay of lncRNA/miRNA/mRNA interactions or there is an indirect effect by these two oncogenes.

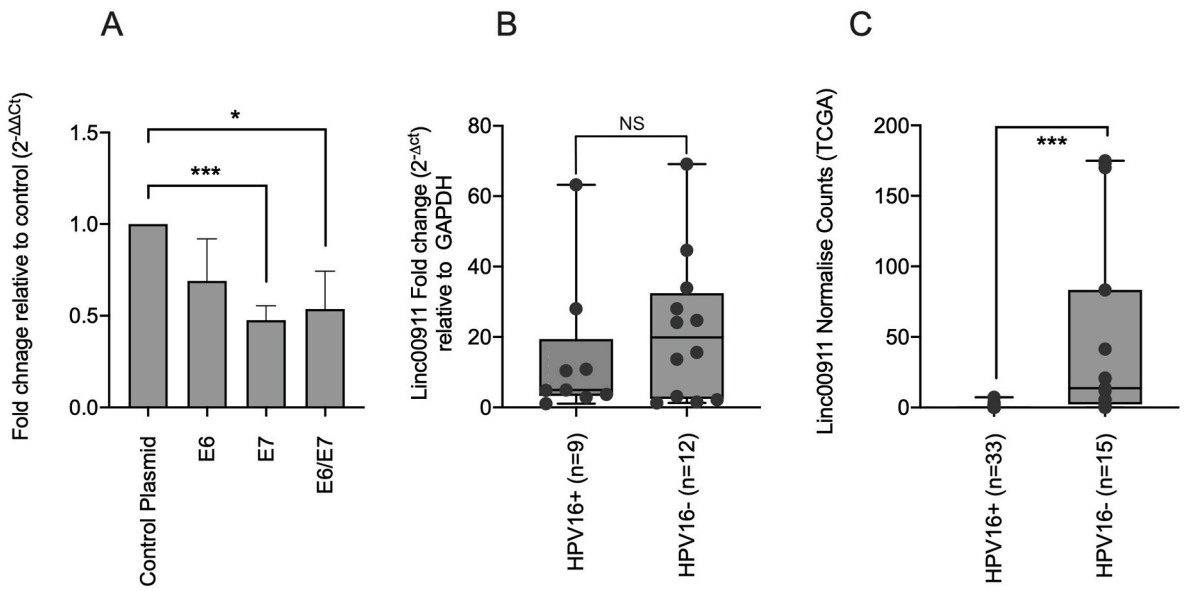


Fig. 5. Quantitative PCR Linc00911 in iHFKs, OPC patient cohort and TCGA patient cohort: A) Linc00911 expression in iHFKs expressing HPV16 E6, E7 alone and E6/E7 combined. B) Linc00911 expression in HPV16-positive (n = 9) OPC tumours compared to HPV-negative tumours (n = 12) in an independent OPC patient cohort. Gene expression was normalised using the calibrator GAPDH. Statistical analysis was performed using a two-tailed student T-test*$p < 0.05$, ***$p < 0.0005$. C) Normalised counts for LINC0911 in the TCGA OPC patient cohort show Linc00911 to be decreased in the HPV16-positive tumours compared to HPV-negative.

Building on our interactome analysis, we refined the interactome from Fig. 6B to exclusively highlight the network associated with Linc00911, a lncRNA whose expression we confirmed to be regulated by HPV16 E6/E7 (Fig. 7). Linc00911 can potentially sponge hsa-miR-20b-5p and hsa-miR-6510-5p, both of which are upregulated in HPV16-positive OPC patients. Notably, hsa-miR-20b-5p targets various mRNAs that are also downregulated in HPV16 OPC.

Several of these genes are also associated with chemotherapy sensitivity in HPV-negative HNC. In HNC cells, silencing of SLCA7a1 and CDK6 increased the cisplatin sensitivity of resistant cells (Roh et al., 2016) and enhanced radiosensitivity (Göttgens et al., 2019), respectively. The silencing of CAV1 also resulted in decreased cell migration and proliferation of HNC cells by promoting ferroptosis (Lu et al., 2022). Patients with low CCND1 levels were more likely to respond to the chemotherapy (Feng et al., 2011). In HPV16-positive HNC, low SLCA7a11 expression levels resulted in a greater sensitivity of HNC to ferroptosis (Hémon et al., 2020).

The downregulation of each of these genes in our HPV16-positive OPC cohort may suggest that the Linc00911/miR-20b/target axis is a mechanism behind the improve response to treatment seen in HPV16-positive HNC patients. This example clearly shows the molecular impact that alterations in a single lncRNA's expression can have, potentially influencing downstream cellular processes.

3.4. Gene ontology and Kegg pathway analysis

To further understand the impact of these by DE ncRNA targets identified from the interactome, we employed a Gene Ontology (GO) enrichment analysis for each network. In the interactome of down-regulated lncRNAs, from Fig. 6B, we predicted sponged miRNA (upregulated) and downregulated mRNA targets. These associated mRNA targets exhibited significant enrichment in key biological processes, including positive regulation of the cell cycle, negative regulation of cellular process and negative regulation of cell population proliferation (Fig. 8A). For the molecular function processes, mRNA targets showed substantial associations with platelet-derived growth factor binding, growth factor receptor binding, and signalling receptor binding (Fig. 8B). Contrastingly, the network encompassing upregulated lncRNAs from Fig. 6A demonstrated a significant association in a single GO biological process, specifically neuron cell-cell adhesion (Fig. 8C).

We then undertook KEGG pathway enrichment analysis specifically for the lncRNA interactome. For the lncRNA that were downregulated (from Fig. 6B), this analysis revealed that the DE-mRNAs were significantly correlated with pathways, including focal adhesion, PI3K-Akt signalling pathway, and Human papillomavirus infection (Fig. 8D). Within the "Human papillomavirus infection" pathway, genes such as CCND1, CDK6, COLA1, and ITGA5 were identified. The upregulated lncRNA interactome did not exhibit any significant correlations to KEGG pathways. These findings from the GO and KEGG pathway analyses suggest a multifaceted and complex interaction network, wherein these lncRNA/miRNA/mRNA axis could contribute to diverse cellular pathways rather than converging on a singular cellular process.

4. Discussion

The human papillomavirus is recognised as a primary etiological factor in HNC. High-risk HPV type 16 is predominantly implicated in OPC. An understanding of the molecular mechanisms induced by HPV16 infection in OPC is needed to uncover possible drivers of tumour development during HPV infection. This study is focused on the impact of HPV16 on lncRNA and miRNAs in OPC.

While the study of ncRNAs in HPV-related HNC is still emerging, our study represents one of the first comprehensive investigations into lncRNAs and miRNAs within HPV-related OPC. This is a significant departure from previous studies, which predominantly addressed the effects of HPV16 within a broader context, often encompassing all HNC subtypes (Nohata et al., 2016; Salyakina et al., 2016). Clinically, understanding these distinctions is essential, as HNCs are diagnosed and treated according to their specific subtypes (Andreasen et al., 2019). Notably, patients with HPV16-positive HNCs often have better prognoses and outcomes, making the study of this subgroup particularly relevant.

Our bioinformatics analysis revealed 1929 differentially expressed lncRNAs, which were filtered to 37 DE lncRNAs using a Log2 5-fold change cut-off. Among these, the most significantly upregulated was ENSG00000282816.1, and the most downregulated was ENSG00000231683.6, both of which are lncRNAs with previously unreported functions. In line with earlier studies, we observed

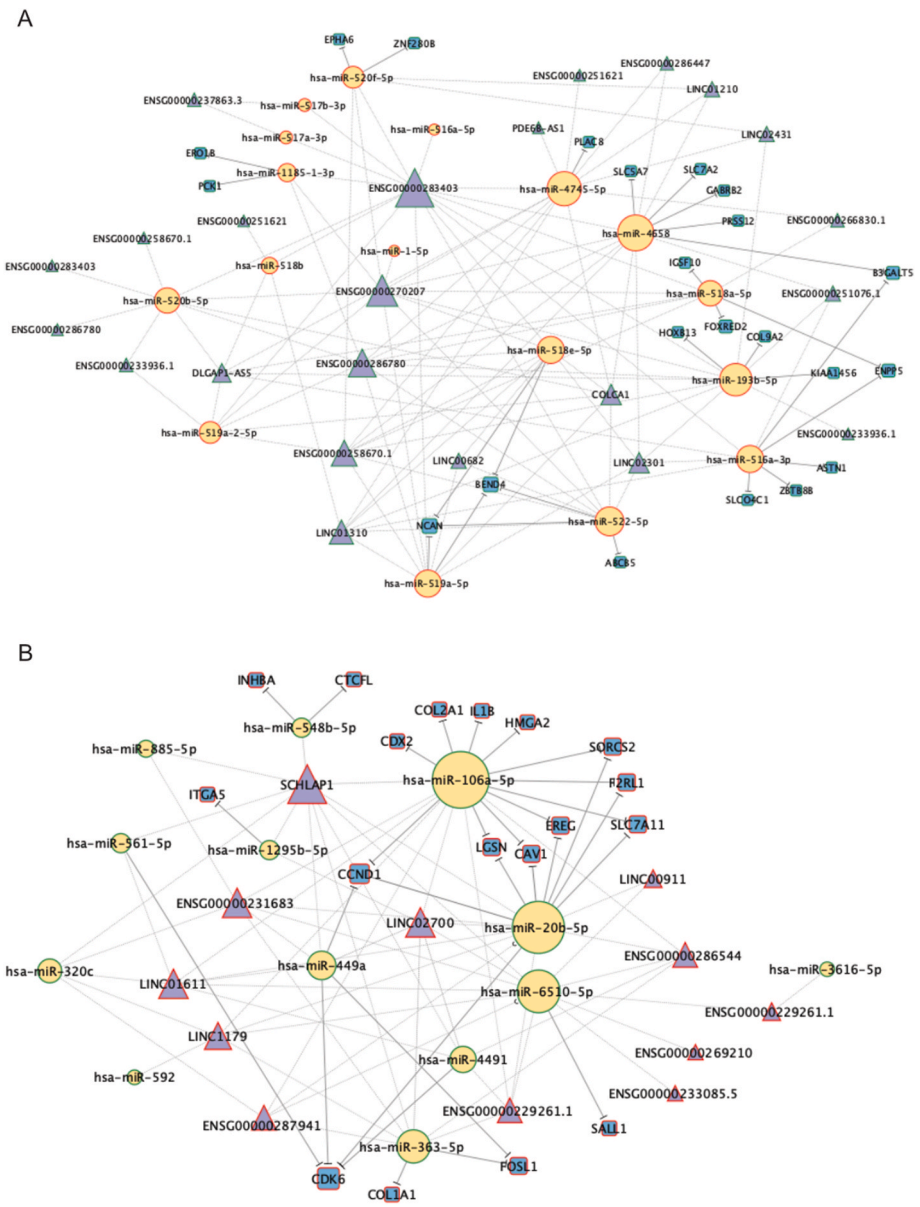


Fig. 6. ceRNA regulatory network in HPV16-positive OPC patients. A) HPV16 upregulated lncRNA and B) HPV16 downregulated lncRNAs and the predicted miRNAs they sponge: This interactome depicts lncRNA that were differentially expressed in HPV16-positive oropharyngeal cancers compared to HPV-negative tumours. The lncRNAs are linked to their predicted target miRNAs with an inverse expression profile in HPV16-positive oropharyngeal cancers. The known miRNA-mRNA target interactions with an inverse expression pattern in HPV16-positive oropharyngeal cancers are also depicted. Triangles represent lncRNAs, circles represent miRNAs, and squares represent mRNA. Green node borders represent upregulated expression, and red represents downregulated expression patterns. The size of the nodes indicates the number of interactions each node has; the larger the size, the more interactions.

upregulation in known lncRNAs such as COLCA1 (Salyakina et al., 2016), PDE6B-AS1 (Salyakina et al., 2016), ENSG00000267247.1 (Salyakina et al., 2016), ENSG00000251076.1 (Salyakina et al., 2016), ENSG00000258670 (Nohata et al., 2016) and ENSG00000233936 (Nohata et al., 2016), and downregulation in ENSG00000233085.5 (Salyakina et al., 2016). However, all other lncRNAs captured using our fold change criterion were unique to this study.

Two primary factors might explain the discovery of novel lncRNAs in our study. Firstly, our analysis is the most up-to-date, utilising a recently updated lncRNA database, which may have contributed to the identification of these novel lncRNAs. Secondly, the DE lncRNAs we identified could be unique to the HPV16-positive OPC subtype. Our focus on OPC solely might have led to the uncovering of these novel DE lncRNAs, which might not have been detected in broader HNC studies. While it's conceivable that other studies might detect some of these lncRNAs with changes less than a 2-fold change, exhaustive lists of all DE lncRNAs were not always provided, limiting the scope for comparative analysis. To standardise the nomenclature of these lncRNAs, we opted to use names approved by the HUGO Gene Nomenclature Committee (HGNC) and Ensembl IDs for novel transcripts. For those interested in the various gene names of these lncRNAs across different databases, we have compiled this information in Supplementary Table 6.

Given that the majority of these DE-lncRNAs are novel lncRNA transcripts and their functions have not been identified, we created a lncRNA/miRNA/mRNA interactome analysis to elucidate possible molecular interactions. The lncRNA ENSG00000287941 sponged the most miRNAs (13 upregulated miRNAs), and ENSG00000283403 sponged the most downregulated miRNAs (16 miRNAs). These two lncRNAs are novel, and their function has not previously been described. Only LINC00165 and ENSG00000225280 did not sponge any of our miRNAs,

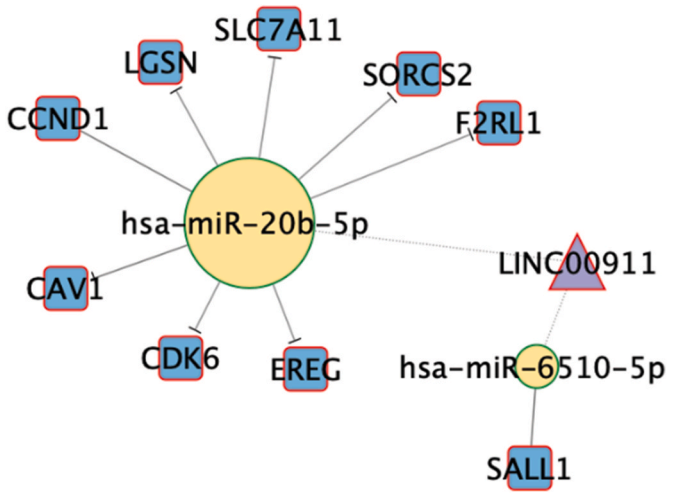


Fig. 7. HPV16 E6/E7 regulated lncRNA Linc00911 and the predicted miRNAs it sponges: This filtered interactome depicts the downregulated lncRNA Linc00911 linked to the two predicted miRNAs it targets, has-miR-20b-5p and has-miR-6510-5p that are also upregulated in HPV16-positive OPC. The miRNAs are also linked to their known targets, downregulated in HPV16-positive OPC. Triangles represent lncRNAs, circles represent miRNAs, and squares represent mRNA. Green node borders represent upregulated expression, and red represents downregulated expression patterns. The size of the nodes indicates the number of interactions each node has, the larger the size, the more interactions.

although they still possess the ability to sponge other miRNAs.

Among the differentially expressed microRNAs in our study, miR-20b, miR-363, miR-193b, and miR-517a have been previously reported to exhibit similar DE patterns in other HNC studies (Sais et al., 2021). The remaining DE-miRNAs identified in our research appear to be unique for this study. A significant proportion of the downregulated miRNAs belong to the miRNA cluster on chromosome 19 (C19MC), which is one of the largest known miRNA clusters and comprises 46 miRNAs. The altered expression of C19MC has been linked to a variety of cancers (Jinesh et al., 2018; Augello et al., 2012; Flor et al., 2016) with several of these miRNAs functioning as tumour suppressors.

For instance, miR-517a's increased expression in bladder cancer can inhibit cell proliferation and induce apoptosis (Yoshitomi et al., 2011). Similarly, miR-520b, known to be downregulated in HNC cell lines, contributes to heightened tumour growth (Lu et al., 2017), while miR-519a serves as a tumour suppressor in ovarian cancers (Tian et al., 2018). Additionally, miR-518b, downregulated in oesophageal SCC, is linked to cell proliferation (Zhang et al., 2012). Several C19MC miRNAs also play a role in chemosensitivity. We know that miR-522, when overexpressed in DOX-resistant colon cancer cells, can restore DOX sensitivity (Yang et al., 2015). Notably, miR-519a has been observed to enhance cisplatin sensitivity in lung cancer cells caused by its sponging to the lncRNA LINC00221 (Tang et al., 2019). Most of the upregulated lncRNAs (19 out of 23) identified in our study were predicted to regulate twelve miRNAs within the C19MC, indicating the significance of this miRNA cluster in the context of HPV16 infection.

To gain a greater understanding of the impact of these DE-lncRNA/miRNAs interactions, we predicted the known mRNA targets of our DE-miRNAs and filtered these for those with an inverse expression pattern. The miRNAs with the most DE-mRNA targets were the upregulated hsa-miR-106a-5p (11 mRNA targets) and hsa-miR-20b-5p (8 mRNA targets). Their upregulation was also described in other studies in HPV16 OPC (Sais et al., 2021). The downregulated miRNAs with the most targets, has-miR-4658 (5 mRNA targets) and hsa-miR-516a-3p (5 mRNA targets), are unique to our study. Their gene targets, SLOC4C1, B3GALT5, and GABRB2, have all been shown to regulate cell

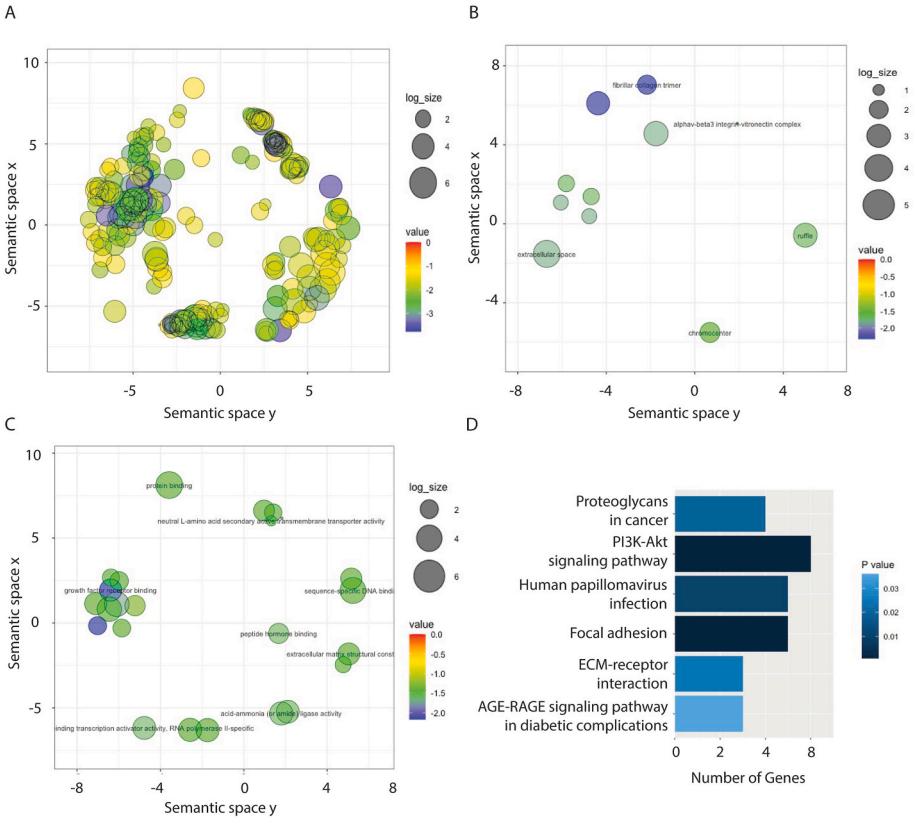


Fig. 8. Gene ontology and KEGG analysis of mRNAs targeted by upregulated miRNA sponged by downregulated lncRNAs in HPV16-positive Oropharyngeal cancers: A) GO terms under Biological process, B) GO terms under Molecular Function, C) GO terms under cellular component, D) KEGG terms.

proliferation, migration and invasion of various cancers (Ban et al., 2017; Hu et al., 2020; Liao et al., 2021). HMGA2 (Li et al., 2022) and EREG (Liu et al., 2020) have also been shown to regulate tumour progression of HNC.

The notion that E6 and E7 could directly affect the expression of specific lncRNAs was measured in iHFK cells overexpressing these oncogenes. The results revealed decreased lncRNA Linc00911 expression in iHFK cells overexpressing HPV16 E6 and E7. This relationship between Linc00911 and HPV16 was also observed in two independent clinical cohorts. Our findings suggest a possible direct regulatory relationship involving E6, E7, and Linc00911 expression. In breast cancer patients, higher levels of Linc00911 were associated with a poorer prognosis (Li et al., 2020). This raises the possibility that in HPV-positive OPC, the downregulation of Linc00911 might contribute to the better prognosis observed in these OPC patients.

One of the caveats of our study is the predictive nature of our analysis and the in-silico parameters and threshold we have set to delineate a convincing narrative. We have set high threshold values of Log2 5-fold increases to provide a greater probability of identifying bona fide genes. The validation of Linc00911 in cells and patient samples, to some extent, supports our bioinformatics approach. We would suggest a degree of judgment when interpreting any bioinformatics prediction, and any conclusions should be made within the context of the disease or regulatory pathway.

In summary, this study has provided a comprehensive analysis of differentially expressed lncRNAs and miRNAs in HPV16-positive OPC. We discovered known and unique ncRNA expression profiles, which were then utilised to construct an interactome to dissect and identify pathways that might be associated with HPV16-driven oncogenesis. What is clear is that both the ncRNA and RNA families are highly connected, forming a complex RNA regulatory network. The analysis yielded several interesting lncRNAs that were dysregulated in HPV16-positive OPC. In particular, the lncRNA Linc00911 may be regulated by HPV16 E6 and E7, as validated in independent patient cohorts. Moreover, Linc00911 could potentially bind to miR-20b impacting the miRNAs downstream targets. Sponging of a single miRNA can lead to a more significant regulator effect, and we suspect the role of sponging miRNAs is more common, and the interplay between these ncRNA families could drive transformation events in Oropharyngeal Cancers.

Data availability statement

All data will be made available upon request to the corresponding authors.

Additional information (including a competing interests statement)

All authors do not have any competing interest.

CRediT authorship contribution statement

Dayna Sais: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Meredith Hill:** Writing – review & editing, Methodology, Formal analysis. **Fiona Deutsch:** Writing – review & editing, Methodology, Formal analysis. **Phuong Thao Nguyen:** Writing – review & editing, Methodology, Formal analysis. **Valerie Gay:** Writing – review & editing, Supervision, Methodology, Formal analysis. **Nham Tran:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.virol.2024.110220.

References

Ambros, V., 2004. The functions of animal microRNAs. Nature 431 (7006), 350–355.

Andreasen, S., et al., 2019. An update on head and neck cancer: new entities and their histopathology, molecular background, treatment, and outcome. APMIS 127 (5), 240–264.

Augello, C., et al., 2012. MicroRNA profiling of hepatocarcinogenesis identifies C19MC cluster as a novel prognostic biomarker in hepatocellular carcinoma. Liver Int. 32 (5), 772–782.

Ban, M.J., et al., 2017. Solute carrier organic anion transporter family member 4A1 (SLCO4A1) as a prognosis marker of colorectal cancer. J. Cancer Res. Clin. Oncol. 143 (8), 1437–1447.

Bartel, D.P., 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116 (2), 281–297.

Casarotto, M., et al., 2020. Beyond MicroRNAs: emerging role of other non-coding RNAs in HPV-driven cancers. Cancers 12 (5), 1246.

Castellsagué, X., et al., 2016. HPV involvement in head and neck cancers: comprehensive assessment of biomarkers in 3680 patients. J. Natl. Cancer Inst. 108 (6), djv403.

Chaturvedi, A.K., et al., 2011. Human papillomavirus and rising oropharyngeal cancer incidence in the United States. J. Clin. Oncol. 29 (32), 4294.

Demers, G.W., et al., 1996. Abrogation of growth arrest signals by human papillomavirus type 16 E7 is mediated by sequences required for transformation. J. Virol. 70 (10), 6862–6869.

Deutsch, F., et al., 2024. Biplex quantitative PCR to detect transcriptionally active human papillomavirus 16 from patient saliva. BMC Cancer 24 (1), 442.

Dhawan, A., et al., 2018. Pan-cancer characterisation of microRNA across cancer hallmarks reveals microRNA-mediated downregulation of tumour suppressors. Nat. Commun. 9 (1), 5228.

Feng, Z., et al., 2011. CCND1 as a predictive biomarker of neoadjuvant chemotherapy in patients with locally advanced head and neck squamous cell carcinoma. PLoS One 6 (10), e26399.

Flor, I., et al., 2016. Expression of microRNAs of C19MC in different histological types of testicular germ cell tumour. Cancer genomics & proteomics 13 (4), 281–289.

Gao, C., et al., 2019. The construction and analysis of ceRNA networks in invasive breast cancer: a study based on the Cancer Genome Atlas. Cancer Manag. Res. 11, 1–11.

Gillison, M.L., et al., 2000. Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. J. Natl. Cancer Inst. 92 (9), 709–720.

Goodall, G.J., Wickramasinghe, V.O., 2021. RNA in cancer. Nat. Rev. Cancer 21 (1), 22–36.

Gooi, Z., Chan, J.Y., Fakhry, C., 2016. The epidemiology of the human papillomavirus related to oropharyngeal head and neck cancer. Laryngoscope 126 (4), 894–900.

Göttgens, E.-L., et al., 2019. Inhibition of CDK4/CDK6 enhances radiosensitivity of HPV negative head and neck squamous cell carcinomas. Int. J. Radiat. Oncol. Biol. Phys. 105 (3), 548–558.

Halbert, C., Demers, G., Galloway, D., 1991. The E7 gene of human papillomavirus type 16 is sufficient for immortalization of human epithelial cells. J. Virol. 65 (1), 473–478.

Hémon, A., et al., 2020. SLC7A11 as a biomarker and therapeutic target in HPV-positive head and neck Squamous Cell Carcinoma. Biochem. Biophys. Res. Commun. 533 (4), 1083–1087.

Hu, X., et al., 2020. Knockdown of SLCO4C1 inhibits cell proliferation and metastasis in endometrial cancer through inactivating the PI3K/Akt signaling pathway. Oncol. Rep. 43 (3), 919–929.

Huang, H.-Y., et al., 2020. miRTarBase 2020: updates to the experimentally validated microRNA–target interaction database. Nucleic Acids Res. 48 (D1), D148–D154.

Jalali, S., et al., 2013. Systematic transcriptome wide analysis of lncRNA-miRNA interactions. PLoS One 8 (2), e53823.

Jinesh, G.G., Flores, E.R., Brohl, A.S., 2018. Chromosome 19 miRNA cluster and CEBPB expression specifically mark and potentially drive triple negative breast cancers. PLoS One 13 (10), e0206008.

John, B., et al., 2004. Human microRNA targets. PLoS Biol. 2 (11), e363.

Li, X.-X., et al., 2020. Identification of long noncoding RNAs as predictors of survival in triple-negative breast cancer based on network analysis. BioMed Res. Int. 2020, 8970340.

Li, Z., et al., 2022. HMGA2-Snai2 axis regulates tumorigenicity and stemness of head and neck squamous cell carcinoma. Exp. Cell Res. 418 (1), 113271.

Liao, Y.-M., et al., 2021. High B3GALT5 expression confers poor clinical outcome and contributes to tumor progression and metastasis in breast cancer. Breast Cancer Res. 23 (1), 5.

Liu, S., et al., 2020. EREG-Driven oncogenesis of head and neck squamous cell carcinoma exhibits higher sensitivity to erlotinib therapy. Theranostics 10 (23), 10589–10605.

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. Methods 25 (4), 402–408.

Lu, Y.-C., et al., 2017. MiR-520b as a novel molecular target for suppressing stemness phenotype of head-neck cancer by inhibiting CD44. Sci. Rep. 7 (1), 2042.

Lu, T., et al., 2022. Caveolin-1 promotes cancer progression via inhibiting ferroptosis in head and neck squamous cell carcinoma. J. Oral Pathol. Med. 51 (1), 52–62.
Maere, S., Heymans, K., Kuiper, M., 2005. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics 21 (16), 3448–3449.
Moody, C.A., Laimins, L.A., 2010. Human papillomavirus oncoproteins: pathways to transformation. Nat. Rev. Cancer 10 (8), 550–560.
Nemeth, K., et al., 2024. Non-coding RNAs in disease: from mechanisms to therapeutics. Nat. Rev. Genet. 25 (3), 211–232.
Network, C.G.A., 2015. Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature 517 (7536), 576.
Nohata, N., Abba, M.C., Gutkind, J.S., 2016. Unraveling the oral cancer lncRNAome: identification of novel lncRNAs associated with malignant progression and HPV infection. Oral Oncol. 59, 58–66.
Roh, J.-L., et al., 2016. Induction of ferroptotic cell death for overcoming cisplatin resistance of head and neck cancer. Cancer Lett. 381 (1), 96–103.
Sais, D., et al., 2018. Human papillomavirus 16 E6 modulates the expression of miR-496 in oropharyngeal cancer. Virology 521, 149–157.
Sais, D., Munger, K., Tran, N., 2021. The dynamic interactome of microRNAs and the human papillomavirus in head and neck cancers. Current Opinion in Virology 51, 87–95.
Salyakina, D., Tsinoremas, N.F., 2016. Non-coding RNAs profiling in head and neck cancers. NPJ genomic medicine 1 (1), 1–12.
Schache, A.G., et al., 2016. HPV-related oropharynx cancer in the United Kingdom: an evolution in the understanding of disease etiology. Cancer Res. 76 (22), 6598–6606.
Sharma, S., Munger, K., 2020. KDM6A-Mediated expression of the long noncoding RNA DINO causes TP53 tumor suppressor stabilization in human papillomavirus 16 E7-expressing cells. J. Virol. 94 (12), e02178, 19.
Tang, H., et al., 2019. Linc00221 modulates cisplatin resistance in non-small-cell lung cancer via sponging miR-519a. Biochimie 162, 134–143.
Tian, F., et al., 2018. 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. 15 (2), 1819–1824.
Tumban, E., 2019. A current update on human papillomavirus-associated head and neck cancers. Viruses 11 (10), 922.
Xia, T., et al., 2014. Long noncoding RNA associated-competing endogenous RNAs in gastric cancer. Sci. Rep. 4 (1), 1–7.
Yang, G., et al., 2015. MicroRNA-522 reverses drug resistance of doxorubicin-induced HT29 colon cancer cell by targeting ABCB5. Mol. Med. Rep. 12 (3), 3930–3936.
Yang, Y., et al., 2022. Integrated analysis of lncRNA-associated ceRNA network in p16-positive and p16-negative head and neck squamous cell carcinoma. Medicine 101 (33), e26120.
Yoshitomi, T., et al., 2011. Restoration of miR-517a expression induces cell apoptosis in bladder cancer cell lines. Oncol. Rep. 25 (6), 1661–1668.
Zhang, M., et al., 2012. miR-518b is down-regulated, and involved in cell proliferation and invasion by targeting Rap1b in esophageal squamous cell carcinoma. FEBS (Fed. Eur. Biochem. Soc.) Lett. 586 (19), 3508–3521.
Zhang, X., et al., 2019. Mechanisms and functions of long non-coding RNAs at multiple regulatory levels. Int. J. Mol. Sci. 20 (22).