Kent Academic Repository
Full text document (pdf)

Citation for published version

Lechner, Matt and Chakravarthy, Ankur R. and Walter, Vonn and Masterson, Liam and Feber, Andrew and Jay, Amrita and Weinberger, Paul M. and McIndoe, Richard A. and Forde, Cillian T. and Chester, Kerry and Kalavrezos, Nicholas and O'Flynn, Paul and Forster, Martin and Jones, Terry M. and Vaz, Francis M. and Hayes, D. Neil and Fenton, TR (2018) Frequent HPV-independent

DOI

https://doi.org/10.1016/j.oraloncology.2018.06.006

Link to record in KAR

https://kar.kent.ac.uk/67260/

Document Version

Author's Accepted Manuscript

Frequent HPV-independent p16/INK4A overexpression in head and neck cancer

Matt Lechner[a,b,#], Ankur R. Chakravarthy[a,#], Vonn Walter[c,d,#,], Liam Masterson[e], Andrew Feber[a], Amrita Jay[f], Paul M. Weinberger[g], Richard A. McIndoe[g], Cillian T. Forde[b], Kerry Chester[a], Nicholas Kalavrezos[b], Paul O'Flynn[b], Martin Forster[a,b], Terry M. Jones[h], Francis M. Vaz[b], D. Neil Hayes[c,*] & Tim R. Fenton[i,*]

[#] Equal contribution

[a] UCL Cancer Institute, University College London, 72 Huntley Street, London, WC1E 6DD, United Kingdom.
[b] Head and Neck Centre, University College London Hospitals NHS Trust, Euston Road, London, NW1 2PG, United Kingdom.
[c] Department of Public Health Sciences, Penn State College of Medicine, Hershey, PA, US
[d] UNC Lineberger Comprehensive Cancer Center, School of Medicine, UNC-Chapel Hill, NC, US
[e] Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, United Kingdom.
[f] Department of Histopathology, University College London Hospitals NHS Trust, Euston Road, London, NW1 2PG, United Kingdom.
[g] Center for Biotechnology and Genomic Medicine, Department of Otolaryngology, and Georgia Regents Cancer Center, Georgia Regents University, Augusta, GA, US
[h] Department of Molecular and Clinical Cancer Medicine, University of Liverpool, Liverpool, United Kingdom.
[i] School of Biosciences, University of Kent, Canterbury, CT2 7NJ, United Kingdom.

*Correspondence to hayes@med.unc.edu or t.fenton@kent.ac.uk

Word Count: 3082

Abstract:

Objectives:

p16^{INK4A} (p16) is the most widely used clinical biomarker for Human Papillomavirus (HPV) in head and neck squamous cell cancer (HNSCC). HPV is a favourable prognostic marker in HNSCC and is used for patient stratification. While p16 is a relatively accurate marker for HPV within the oropharynx, recent reports suggest it may be unsuitable for use in other HNSCC subsites, where a smaller proportion of tumors are HPV-driven.

Materials & Methods:

We integrated reverse phase protein array (RPPA) data for p16 with HPV status based on detection of viral transcripts by RNA-seq in a set of 210 HNSCCs profiled by The Cancer Genome Atlas project. Samples were queried for alterations in *CDKN2A*, and other pathway genes to investigate possible drivers of p16 expression.

Results:

While p16 levels as measured by RPPA were significantly different by HPV status, there were multiple HPV (-) samples with similar expression levels of p16 to HPV (+) samples, particularly at non-oropharyngeal subsites. In many cases, p16 overexpression in HPV (-) tumors could not be explained by mutation or amplification of *CDKN2A* or by *RB1* mutation. Instead, we observed enrichment for inactivating mutations in the histone H3 lysine 36 methyltransferase, *NSD1* in HPV (-) / p16-high tumors.

Conclusions:

RPPA data suggest high p16 protein expression in many HPV (-) non-oropharyngeal HNSCCs, limiting its potential utility as an HPV biomarker outside of the oropharynx. HPV-independent overexpression of wild-type p16 in non-oropharyngeal HNSCC may be linked to global deregulation of chromatin state by inactivating mutations in *NSD1*.

Introduction:

Human papillomavirus (HPV), in particular the HPV16 subtype, represents a major independent risk factor for oropharyngeal carcinoma (OPSCC), a group of head and neck squamous cell cancers (HNSCC) affecting the tonsils and base of tongue [1-3]. This is a rapidly evolving cancer in the Western world, with rising trends particularly in the young [4, 5]. The most commonly used biomarker for HPV (+) HNSCC in the clinical setting at present is p16^{INK4A} (p16), encoded by the *CDKN2A* gene (cyclin-dependent kinase (CDK) inhibitor 2A). p16 is overexpressed in HPV (+) HNSCC [6] and cervical cancer [7], and plays a key role in cell cycle control, exerting its function as a cyclin-dependent kinase (CDK) inhibitor by inactivating the CDKs that phosphorylate the retinoblastoma tumour suppressor protein (pRB), leading to a block of the cell cycle [8-10]. E7, one of two main HPV oncoproteins, blocks pRB's repressive effect on the E2F transcription factors, leading to the transcription of S-phase genes and aberrant cell proliferation [11]. *CDKN2A* transcription is strongly upregulated in cells expressing E7 [12-15], with recent findings revealing that this is via an epigenetic mechanism involving histone H3 K27 demethylation by KDM6B, independent of pRB [16, 17]. p16 upregulation in tumour cells can therefore be a marker of active HPV involvement since it is dependent upon E7 activity. This distinguishes it from HPV DNA detection in isolation, which indicates the presence of the virus, but does not provide evidence for viral oncogene expression and hence a driver role in the tumor.

p16 expression detected using immunohistochemistry (IHC) exhibits a relatively high sensitivity and specificity as a single test of HPV activity and therefore status in oropharyngeal squamous cell cancer (OPSCC) [18]. Combined with E6 DNA qPCR it has been shown to have 97% sensitivity and 94% specificity (compared with the RNA qPCR "gold standard") and to be the best discriminator of favourable outcome [19], adding prognostic value. The recently published UICC 8th edition TNM classification institutes changes to the staging of oropharyngeal cancer in particular, as a reflection of this [20]. p16 is the most widely used marker to determine HPV status and has been employed in many trials [21, 22] to stratify treatment of patients with OPSCC, usually without the concomitant use of HPV-ISH or E6 qPCR. In an attempt to define primary site of origin, p16 overexpression detected in FNAC samples in the context of carcinoma of unknown primary is also recommended in the TNM 8th edition [20, 23] and is routinely tested in non-OPSCC

in some centres outside the US [24]. The College of American Pathologists recommends high-risk HPV testing for all new oropharyngeal squamous cell carcinoma patients, but it is not routinely recommended for head and neck carcinomas outwith the oropharynx [25].

Using TCGA data, we have recently demonstrated an HPV-dependent etiology for approximately 4% of non-oropharnygeal HNSCCs (non-OPSCC) [26]. Whilst HPV-targeted therapies such as therapeutic vaccines [27], and potentially, deintensification might be expected to benefit these patients, accurately identifying them requires a robust HPV assay suitable for use in both OPSCC and non-OPSCC. Recent studies suggest that in non-OPSCC, where the prevalence of HPV is lower, the diagnostic performance of p16 is inferior [28] but more data from larger sample sets are required to confirm these initial findings. We hypothesized that rather than being false positives in the IHC assay, p16 protein may be over-expressed due to HPV-independent mechanisms in HPV (-) non-OPSCC. To explore this hypothesis in a cohort for which comprehensive molecular data are available, we interrogated a subset of 210 HNSCCs (11 OPSCC and 199 non-OPSCC) profiled by TCGA in which p16 protein expression was analyzed by reverse-phase protein array (RPPA) [29].

Materials and Methods:

Data analysis:

R statistical software v3.2.2 (R Core Team [30]) was used for the analysis of the data. In summary, we used reverse phase protein array (RPPA) data from head and neck cancer samples from the TCGA in combination with HPV status based on the detection of viral transcripts retrieved from RNA-seq analysis to examine p16INK4A protein expression in a set of 210 HNSCCs (Supplemental table 1), 14 of which are HPV (+). RPPA values were correlated against IHC in a selected subset of cases where both measures were available. In addition and in an attempt to understand the mechanism underlying p16 expression the TCGA data were further interrogated for a variety of alterations in *CDKN2A*, and other pathway genes including *RB1* using matched TCGA exome sequence and copy number alteration data. In an exploratory analysis, HPV (-) - p16 high samples were interrogated for alternative pathway alterations, including a novel association with mutations of the gene *NSD1*, which might explain elevated protein levels. In detail, RSEM-based gene expression quantification measurements (n = 520) and automated mutation calls (n = 508) were

downloaded from the Legacy Archive of the Genomic Data Commons (https://portal.gdc.cancer.gov/legacy-archive/search/f), while thresholded gene-level DNA copy number calls were downloaded from the Broad Institute's Firehose GDAC (https://gdac.broadinstitute.org/). RPPA-based measurements of p16 abundance (n = 210) were obtained from the cBioPortal (http:// http://www.cbioportal.org/), and HPV status was determined by RNA-seq analysis, as previously described [26], and associated with clinical data (Breakdown by status and site in supplementary data). Figures are based on log2(RSEM + 1) values. Scatterplots and boxplots based on gene and protein expression were created. Differences in expression by group were tested using a Wilcoxon Rank Sum test and visualization of expression values was carried out using the ggplot2 package. Copy number data as inferred by TCGA GISTIC analysis and mutation status were retrieved from the cBio portal [31] for cancer genomics (imported from the latest run of Broad's Firehose GDAC). Non-oropharyngeal HPV (-) HNSCCs showed similar expression levels across anatomical subsites and were therefore considered a single group. To search for molecular correlates, p16 levels (RPPA) were regressed against level 3 RNA-seq data using limma-voom, with genes expressed at less than 1 count in at least half the dataset filtered before model fitting. Data were quantile normalized and significantly associated genes were identified at FDR < 0.05. Associations with protein abundances were established by regressing p16 levels against RPPA estimates for other protein abundances (Figure 3B), and significantly associated proteins were again defined at FDR < 0.05. Pathway analysis was carried out using IPA (Ingenuity Pathway Analysis) using experimentally confirmed interactions in human cells/tissue.

p16 IHC staining and HPV-ISH

p16 IHC staining was carried out in the Bond Autostainer (Leica Microsystems Inc, Norwell, MA, USA) according to the manufacturer's IHC protocol. Slides were put in a 60° oven to remove excess paraffin. Slides were then placed in the autostainer and dewaxed in Bond Dewax solution (AR9222) and hydrated in Bond Wash solution (AR9590). Antigen retrieval was performed for 30 min at 100 °C in Bond-Epitope Retrieval solution 1 (pH 6.0, AR9961). Slides were then incubated with p16 antibody (mouse monoclonal anti-p16 antibody (MAB4133), Chemicon International Company/Millipore Corporation, Temecula, CA, USA) for 15 min. Antibody detection was performed using the Bond Polymer Refine Detection System (DS9800). Stained slides were dehydrated and coverslips added.

Immunohistochemical staining was performed in the Translational Pathology Lab at UNC. After completion of IHC, slides are stored at room temperature and a virtual scanned copy of all slides was obtained using the Aperio brand scanner and software. p16 immunostaining and HPV-ISH of archival FFPE tissue was performed on the the in-house samples (National Research Ethics Service Committee of East of England; reference number 12/EE/44; Supplemental Figure 1).

Results:

Elevated p16 levels in HPV (-) tumors at non-oropharyngeal sites:

By integrating RPPA (Reverse Phase Protein Array) data for a number of proteins, including p16 which was made available for 210 tumors (11 OPSCC, 199 non-OPSCC, **Supplemental Table 1**) as part of the recent TCGA analysis of HNSCC [29], with assignment of HPV status based on E6/E7 transcript detection by RNA-seq [31, 32], we assessed the relationship between p16 protein expression and HPV in this cohort (see **Supplementary Table 2** for clinical and pathological details of the analysed TCGA HNSCC cohort). Although p16 protein levels were significantly elevated in HPV (+) samples (Wilcoxon's rank sum test, *P*-value = 1.149e-7), multiple HPV (-) samples displayed similar p16 expression levels to those seen in the HPV (+) samples (**Figure 1**). Using the lower limit of p16 expression in HPV (+) tumors as the threshold for a positive test (dashed line in **Figure 1**) we determined the sensitivity, specificity, positive and negative predictive values in the overall population to be 100%, 79%, 27% and 100% respectively. The particularly unfavorable performance of p16 in non-oropharyngeal cancers for determining HPV status (ppv = 18% among the non-OPSCC samples in this dataset) has been previously demonstrated and reviewed for IHC [33], but to our knowledge this is the first dataset to document this by integrated genomic analysis, proving that the poor performance is not just a function of technical aspects of IHC reagents or staining interpretation. We considered the possibility that high p16 expression in HPV (-) tumors might be confined to a specific anatomic site but found high p16 expression in at least 15% of samples from each of the non-OPSCC sites, including 9 of 49 layngeal, 9 of 28 oral cavity, 4 of 20 floor of mouth and 16 of 45 oral tongue SCCs.

Integration with mutation and copy number data for recurrently altered G1/S checkpoint pathway components

We next considered the possibility that alterations in gene copy number for *CDKN2A*, such as amplifications might explain elevations in gene expression. Additionally, we considered that aberrant signaling through other mutations in the CDK4/6 pathway might be associated with elevated p16 protein expression. We therefore conducted linear regression of p16 protein levels as measured by RPPA against mutation and copy number alteration data for *RB1*, *CCND1*, *TP53, CDK4, CDK6* and *CDKN2A,* HPV status and site (oropharynx versus non-oropharynx). Of these factors, only HPV status (p=1.03e-07) and missense mutations (p=0.0205) or copy number alterations (5.68e-14) in *CDKN2A* itself were associated with p16 protein levels (**Figure 2 and Supplementary Table 2**). The strong association between *CDKN2A* copy number and p16 protein levels is expected due to *CDKN2A* deletions causing loss of p16 in HPV (-) disease (see samples with GISTIC copy number of -2 in Figure 2A). While accumulation of non-functional *CDKN2A* due to missense mutation may therefore explain its overexpression in a subset of samples, wild-type *CDKN2A* is observed in 24 cases with expression of p16 equivalent to that seen in the HPV (+) tumors (**Figure 1**). Four of these tumors harbor *CDKN2A* amplifications, thus in 20 cases we observe high p16 expression in the absence of missense mutations or genomic amplifications.

Molecular correlates of p16 expression in HPV(+)- HNSCC.

Having interrogated known G1/S checkpoint components, we next used global gene expression (RNA-seq) data to search for genes whose expression is associated with p16 protein levels. Analysis of RNA-seq data identified a set of 29 genes whose expression was linearly associated with p16 RPPA protein levels (FDR < 0.05, **Supplemental Figure 2**). The *CDKN2A* gene by itself was also transcriptionally upregulated in these tumours, suggesting higher levels were driven transcriptionally. The scatterplot of p16 versus gene expression of *CDKN2A* demonstrated the high correlation that would be expected (Spearman ρ = 0.63, p < 2.2e-16), confirming the consistency of orthogonal measures of the biomarker (**Figure 3B**). Pathway analysis of this gene set using Ingenuity Pathway Analysis however failed to uncover significant functional associations. Analysis of RPPA data identified two proteins (Yes Associated Protein (YAP) and c-Jun N-terminal Kinase 2 (JNK2 / MAPK9), whose expression was positively associated with that of p16 (FDR < 0.05). Interestingly, high levels of nuclear YAP1 have previously been reported in HPV (+) OPSCC

cells [34].

Increased expression of wild-type *CDKN2A* transcripts in *NSD1* mutant, HPV- HNSCC.

To help validate the utility of RPPA and gene expression as measures of the clinically used p16 IHC, we reviewed the limited representative stained samples available from the TCGA dataset that were definitively HPV (-), p16-high by RPPA, and elevated for *CDKN2A*. In all available cases, the p16 IHC confirmed abundant staining typical of a positive test such as shown in the example (**Figure 3A**) from the UNC TCGA cohort.

Together with our observation of strong p16 IHC staining in two HPV (-) OPSCC samples from an independent cohort collected at University College London Hospital, UK (Supplemental Figure 1), these cases indicate a real upregulation of p16 protein in a subset of HPV (-) HNSCC. Having exhausted canonical candidate CDK4/6 pathway alterations and likewise largely excluded supervised gene expression biomarkers as an explanation for elevations in p16, we entertained exploratory candidates. Interestingly the tumour shown in **Figure 3A** that displayed a high p16 expression harboured an inactivating mutation in the histone H3 lysine 36 (H3K36) methyltransferase, Nuclear Receptor Binding SET Domain Protein 1 (*NSD1*), as first reported by TCGA [29]. We then characterized p16 and *CDKN2A* expression as a function of *NSD1* mutation in the full TCGA dataset, noting a clear trend for *NSD1* mutations at the highest range of *CDKN2A* gene expression. Stratifying by copy number, HPV status and *NSD1* mutational status across the expanded TCGA cohort for which *CDKN2A* transcript data are also available (n=494), *CDKN2A* is frequently wild-type and highly expressed in HPV (-) / *NSD1* mutant tumours (**Figure 3C**). Furthermore, we observed significant enrichment for *NSD1* mutations in *CDKN2A* wild-type tumors (OR = 3.8; p-value = 0.01489, Chi square test, df =1), consistent with a scenario in which *NSD1* mutations may contribute to, and/or allow cells to tolerate high levels of functional p16 protein.

Discussion

Although p16 is the most commonly used clinical biomarker for HPV (+) head and neck

cancer, our analysis would suggest specific limitations that need to be observed. There is a subset of HPV (-) HNSCCs where expression of p16 could lead to misidentification as HPV (+). A subset of the p16 (+) HPV(-) samples have p16 missense mutations, which likely account for accumulation of inactive p16, although insufficient samples with IHC were available to us to verify that such mutants are detected by IHC. However, in other tumours, specific molecular alterations presumably cause, and allow the cells to tolerate, high expression of the wild-type p16 protein.

We have provided ample and parallel data to confirm that the presence of p16 is not simply an artifact of inaccurate IHC, but is in fact present at high levels comparable to HPV (+) samples (we acknowledge the potential limitation that the number of HPV+ samples, compared with HPV- samples, is lower in our dataset. However, in line with previous reports these show significantly increased p16 expression, compared with the other groups). *RB1* mutations are not associated with high p16 levels in our analysis, consistent with recent data showing that HPV E7 induces p16 through upregulation of the KDM6 family H3K27 demethylases, independent of its inhibitory effects on pRb [15]. Our finding that wild-type *CDKN2A* is frequently expressed in *NSD1* mutant / HPV- tumours suggests another epigenetic mechanism (possibly the genome-wide accumulation of repressive histone (H3K27 trimethylation) marks observed upon loss of NSD1 in embryonic stem cells [35] through which it can be induced and tolerated in HPV (-) cells. It is the significant enrichment of *NSD1* mutations and not a result of many other mutated genes such as *RB1*, as previously indicated, which seem to contribute to, and/or allow cells to tolerate high levels of functional p16 protein. The potential link between *NSD1* and p16 will require further examination in cell-based models and larger HNSCC cohorts. Regardless of the underlying mechanisms, it is clear that p16 protein is expressed to the degree seen in HPV (+) HNSCC in a significant proportion of HPV (-), non-oropharyngeal HNSCCs. One intriguing question is whether the HPV (-) HNSCCs that express high levels of wild-type p16 actually become dependent on elevated p16 expression for their survival, as has recently been described for HPV (+) cells [17]. If so, this would point towards a potentially targetable vulnerability shared between HPV (+) and HPV (-) tumors with high p16 expression.

These findings are of clinical relevance. Current therapies for head and neck cancer often carry a significant risk of long-term morbidity, including enterostomy feeding tube dependence, osteoradionecrosis, pharyngo-oesophageal fibrosis leading to dysphagia,

xerostomia, dental failure, hypothyroidism, vascular destruction with associated late occurring strokes and carotid stenosis as well as treatment induced tumours. All these effects/risks are directly related to radiation therapy – therefore reducing the dose exposure is of great therapeutic interest. The de-escalation of therapeutic intensity in HPV (+) disease is currently being tested in clinical trials, such as RTOG 1016, De-ESCALaTE, Quarterback, TROG-12.01, ADEPT, ECOG 3311 and PATHOS (recently reviewed in [36]).

However, as the current study suggests, there is some risk in the use of p16 as the biomarker for HPV status, whether in the oropharynx or other anatomic location. Before selection of the most appropriate therapy, either escalation for high risk patients or de-escalation for lower risk HPV (+) cases, patients deserve a more accurate assay such as directly for the presence of HPV (by e.g. ISH, qPCR, RNA FISH or qRT-PCR). This is of particular relevance in view of the recently published 8th edition TNM classification which institutes changes to the staging of oropharyngeal cancer in particular, as a reflection of this, which distinguishes between e.g. nodal status of HPV(+) and HPV(-) OPSCC, but p16 staining or other forms of HPV testing are not stipulated except for FNAC specimens from metastatic neck nodes [20]. This appears particularly true for non-OPSCC. The American Society of Pathologists does not recommend the routine testing of nonsquamous oropharyngeal carcinomas or nonoropharyngeal carcinomas for HPV [25]. In summary, the decision for any de-escalation of therapeutic intensity must not be made on a p16 staining result alone, especially in view of potential extension of current de-escalation trials to other head and neck subsites [28, 37, 38].

In summary, we show that HPV (-) HNSCCs at non-oropharyngeal subsites frequently express levels of p16 protein seen in HPV (+) oropharyngeal tumors. We show that p16 expression is not due to genomic alterations in *RB1* or other components of the G1/S cell cycle checkpoint, nor can it be explained by mutations in p16 (*CDKN2A*) itself. Instead we find an association with mutations in the histone methyltransferase, *NSD1*, which (like p16 expression) have been linked to improved prognosis in HNSCC [39]. Our observations reinforce the need for alternative HPV biomarkers to complement p16, particularly beyond the oropharynx, and suggest a hitherto unexplored connection between *NSD1* mutation and p16 upregulation in HPV (-) HNSCC.

Acknowledgments

References

[1] D'Souza G, Kreimer AR, Viscidi R, Pawlita M, Fakhry C, Koch WM, et al. Case-control study of human papillomavirus and oropharyngeal cancer. N Engl J Med. 2007;356:1944-56.

[2] Tran N, Rose BR, O'Brien CJ. Role of human papillomavirus in the etiology of head and neck cancer. Head Neck. 2007;29:64-70.

[3] Mendenhall WM, Logan HL. Human Papillomavirus and Head and Neck Cancer. Am J Clin Oncol. 2009:Epub ahead of Print.

[4] Hunter KD, Parkinson EK, Harrison PR. Profiling early head and neck cancer. Nat Rev Cancer. 2005;5:127-35.

[5] Warnakulasuriya S. Global epidemiology of oral and oropharyngeal cancer. Oral Oncol. 2009;45:309-16.

[6] Pyeon D, Newton MA, Lambert PF, den Boon JA, Sengupta S, Marsit CJ, et al. Fundamental differences in cell cycle deregulation in human papillomavirus-positive and human papillomavirus-negative head/neck and cervical cancers. Cancer Res. 2007;67:4605-19.

[7] Lesnikova I, Lidang M, Hamilton-Dutoit S, Koch J. p16 as a diagnostic marker of cervical neoplasia: a tissue microarray study of 796 archival specimens. Diagn Pathol. 2009;4:22.

[8] Stone S, Jiang P, Dayananth P, Tavtigian SV, Katcher H, Parry D, et al. Complex structure and regulation of the P16 (MTS1) locus. Cancer Res. 1995;55:2988-94.

[9] Liggett WH, Jr., Sidransky D. Role of the p16 tumor suppressor gene in cancer. J Clin Oncol. 1998;16:1197-206.

[10] Rocco JW, Sidransky D. p16(MTS-1/CDKN2/INK4a) in cancer progression. Exp Cell Res. 2001;264:42-55.

[11] Sherr CJ. The INK4a/ARF network in tumour suppression. Nature reviews Molecular cell biology. 2001;2:731-7.

[12] Xiong Y, Zhang H, Beach D. Subunit rearrangement of the cyclin-dependent kinases is associated with cellular transformation. Genes Dev. 1993;7:1572-83.

[13] Tam SW, Shay JW, Pagano M. Differential expression and cell cycle regulation of the cyclin-dependent kinase 4 inhibitor p16Ink4. Cancer Res. 1994;54:5816-20.

[14] Hara E, Smith R, Parry D, Tahara H, Stone S, Peters G. Regulation of p16CDKN2 expression and its implications for cell immortalization and senescence. Mol Cell Biol. 1996;16:859-67.

[15] Li Y, Nichols MA, Shay JW, Xiong Y. Transcriptional repression of the D-type cyclin-dependent kinase inhibitor p16 by the retinoblastoma susceptibility gene product pRb. Cancer Res. 1994;54:6078-82.

[16] McLaughlin-Drubin ME, Crum CP, Munger K. Human papillomavirus E7 oncoprotein induces KDM6A and KDM6B histone demethylase expression and causes epigenetic reprogramming. Proc Natl Acad Sci U S A. 2011;108:2130-5.

[17] McLaughlin-Drubin ME, Park D, Munger K. Tumor suppressor p16INK4A is necessary for survival of cervical carcinoma cell lines. Proc Natl Acad Sci U S A. 2013;110:16175-80.

[18] Pannone G, Rodolico V, Santoro A, Lo Muzio L, Franco R, Botti G, et al. Evaluation of a combined triple method to detect causative HPV in oral and oropharyngeal squamous cell carcinomas: p16 Immunohistochemistry, Consensus PCR HPV-DNA, and In Situ Hybridization. Infect Agent Cancer. 2012;7:4.

[19] Schache AG, Liloglou T, Risk JM, Filia A, Jones TM, Sheard J, et al. Evaluation of human papilloma virus diagnostic testing in oropharyngeal squamous cell carcinoma: sensitivity, specificity, and prognostic discrimination. Clin Cancer Res. 2011;17:6262-71.

[20] Huang SH, O'Sullivan B. Overview of the 8th Edition TNM Classification for Head and Neck Cancer. Curr Treat Options Oncol. 2017;18:40.

[21] Gilbert J, Schell MJ, Zhao X, Murphy B, Tanvetyanon T, Leon ME, et al. A randomized phase II efficacy and correlative studies of cetuximab with or without sorafenib in recurrent and/or metastatic head and neck squamous cell carcinoma. Oral Oncol. 2015;51:376-82.

[22] Hoff CM, Lassen P, Eriksen JG, Hansen HS, Specht L, Overgaard M, et al. Does

transfusion improve the outcome for HNSCC patients treated with radiotherapy? - results from the randomized DAHANCA 5 and 7 trials. Acta Oncol. 2011;50:1006-14.

[23] Park GC, Lee M, Roh JL, Yu MS, Choi SH, Nam SY, et al. Human papillomavirus and p16 detection in cervical lymph node metastases from an unknown primary tumor. Oral Oncol. 2012;48:1250-6.

[24] Duncan LD, Winkler M, Carlson ER, Heidel RE, Kang E, Webb D. p16 immunohistochemistry can be used to detect human papillomavirus in oral cavity squamous cell carcinoma. J Oral Maxillofac Surg. 2013;71:1367-75.

[25] Lewis JS, Jr., Beadle B, Bishop JA, Chernock RD, Colasacco C, Lacchetti C, et al. Human Papillomavirus Testing in Head and Neck Carcinomas: Guideline From the College of American Pathologists. Arch Pathol Lab Med. 2018;142:559-97.

[26] Chakravarthy A, Henderson S, Thirdborough SM, Ottensmeier CH, Su X, Lechner M, et al. Human Papillomavirus Drives Tumor Development Throughout the Head and Neck: Improved Prognosis Is Associated With an Immune Response Largely Restricted to the Oropharynx. J Clin Oncol. 2016;34:4132-41.

[27] Yang A, Farmer E, Wu TC, Hung CF. Perspectives for therapeutic HPV vaccine development. J Biomed Sci. 2016;23:75.

[28] Chung CH, Zhang Q, Kong CS, Harris J, Fertig EJ, Harari PM, et al. p16 protein expression and human papillomavirus status as prognostic biomarkers of nonoropharyngeal head and neck squamous cell carcinoma. J Clin Oncol. 2014;32:3930-8.

[29] Cancer Genome Atlas N. Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature. 2015;517:576-82.

[30] R-Core-Team. A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria2015.

[31] Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012;2:401-4.

[32] Parfenov M, Pedamallu CS, Gehlenborg N, Freeman SS, Danilova L, Bristow CA, et al. Characterization of HPV and host genome interactions in primary head and neck cancers. Proc Natl Acad Sci U S A. 2014;111:15544-9.

[33] Combes JD, Franceschi S. Role of human papillomavirus in non-oropharyngeal head and neck cancers. Oral Oncol. 2014;50:370-9.

[34] Alzahrani F, Clattenburg L, Muruganandan S, Bullock M, MacIsaac K, Wigerius M, et al. The Hippo component YAP localizes in the nucleus of human papilloma virus positive oropharyngeal squamous cell carcinoma. J Otolaryngol Head Neck Surg. 2017;46:15.

[35] Streubel G, Watson A, Jammula SG, Scelfo A, Fitzpatrick DJ, Oliviero G, et al. The H3K36me2 Methyltransferase Nsd1 Demarcates PRC2-Mediated H3K27me2 and H3K27me3 Domains in Embryonic Stem Cells. Mol Cell. 2018;70:371-9 e5.

[36] Masterson L, Moualed D, Liu ZW, Howard J, Dwivedi R, Benson R, et al. De-escalation treatment protocols for HPV associated oropharyngeal squamous cell carcinoma: a systematic review and meta-analysis of current clinical trials. . Eur J Cancer. 2014;doi: 10.1016/j.ejca.2014.07.001.

[37] Wang H, Sun R, Lin H, Hu WH. P16INK4A as a surrogate biomarker for human papillomavirus-associated oropharyngeal carcinoma: consideration of some aspects. Cancer science. 2013;104:1553-9.

[38] Lingen MW, Xiao W, Schmitt A, Jiang B, Pickard R, Kreinbrink P, et al. Low etiologic fraction for high-risk human papillomavirus in oral cavity squamous cell carcinomas. Oral Oncol. 2013;49:1-8.

[39] Bui N, Huang JK, Bojorquez-Gomez A, Licon K, Sanchez KS, Tang SN, et al. Disruption of NSD1 in head and neck cancer promotes favorable chemotherapeutic responses linked to hypomethylation. Mol Cancer Ther. 2018.
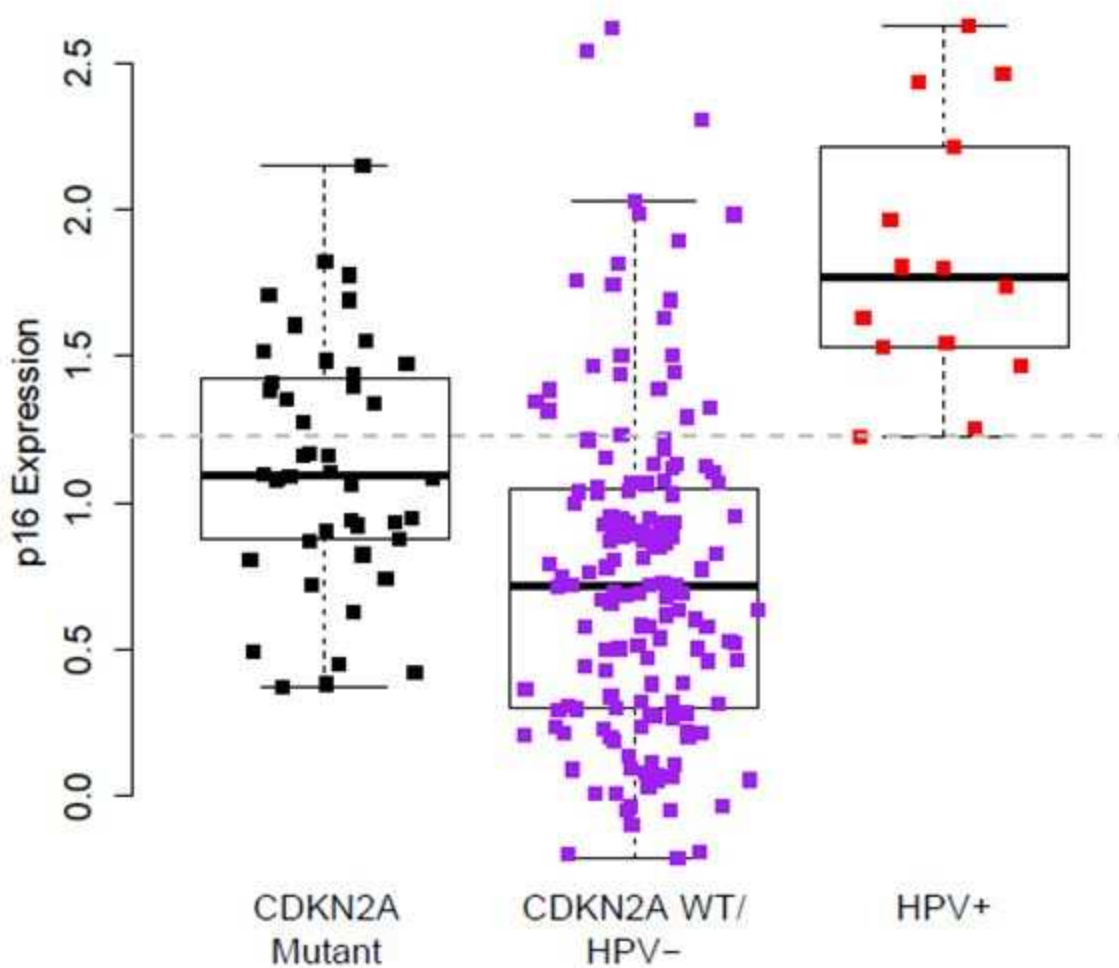
Figure 1: HPV (-) HNSC tumors frequently display elevated p16 levels equivalent to those seen in HPV (+) tumors. RPPA measurements of p16 expression by CDKN2A mutation and HPV status in TCGA HNSC cohort. The dashed line represents the lowest p16 level among the HPV+ samples. [Print in colour]
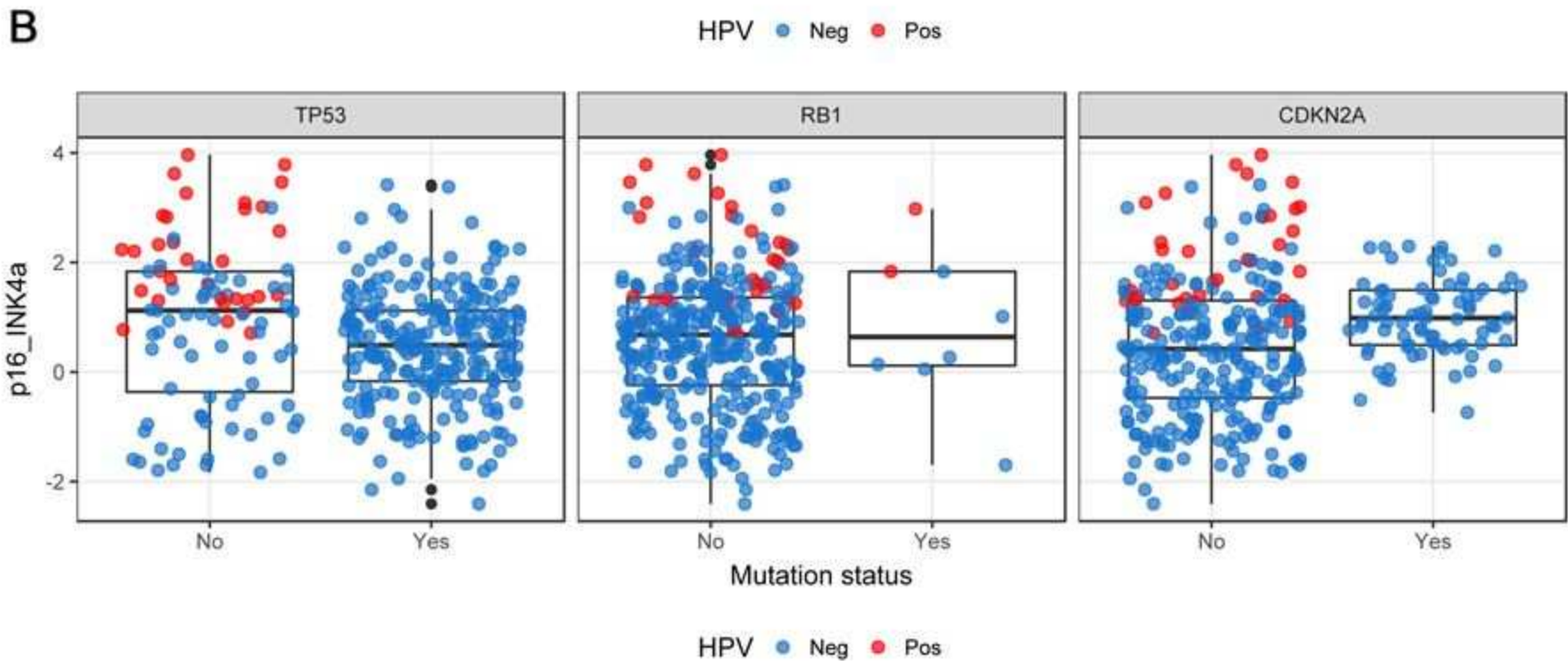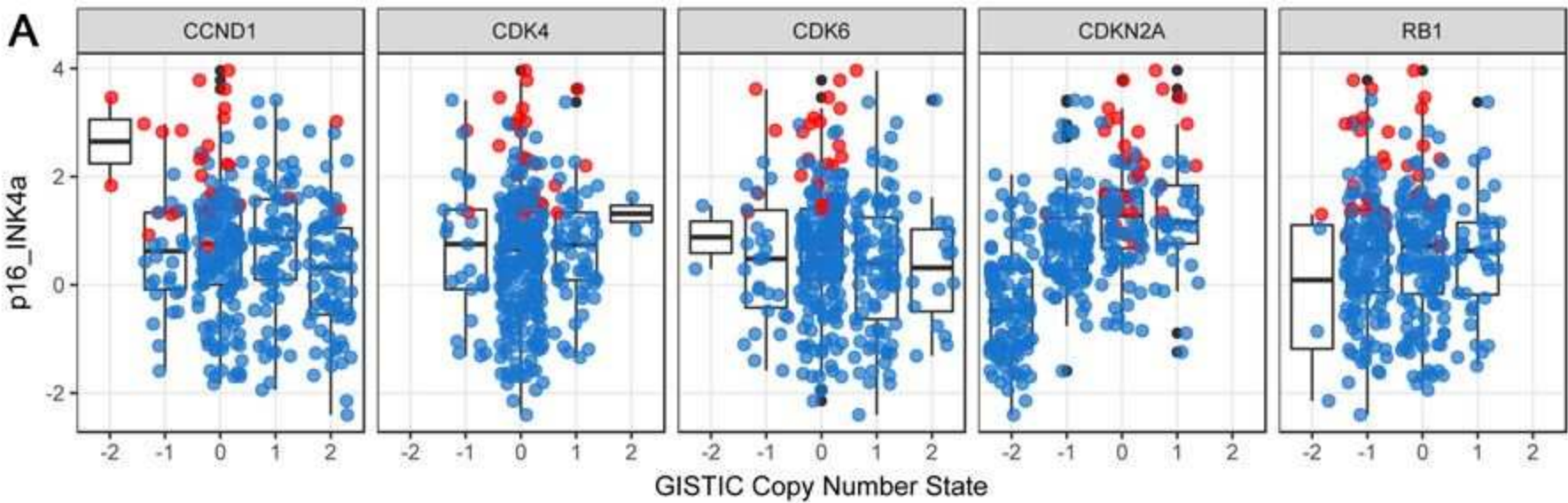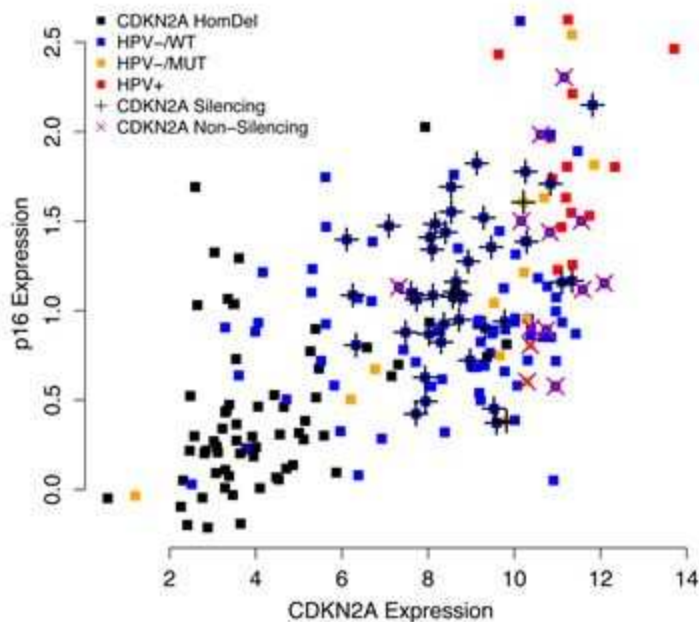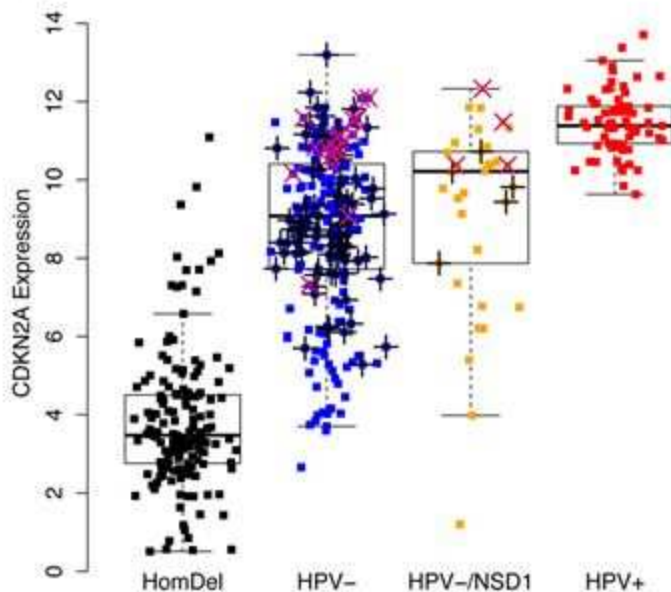
Figure 2: Elevated p16 expression is not associated with mutations or copy number alterations in related cell cycle checkpoint components. Integration of p16 RPPA staining levels with copy number alterations **(A)** or mutations **(B)** by HPV status. [Print in colour]
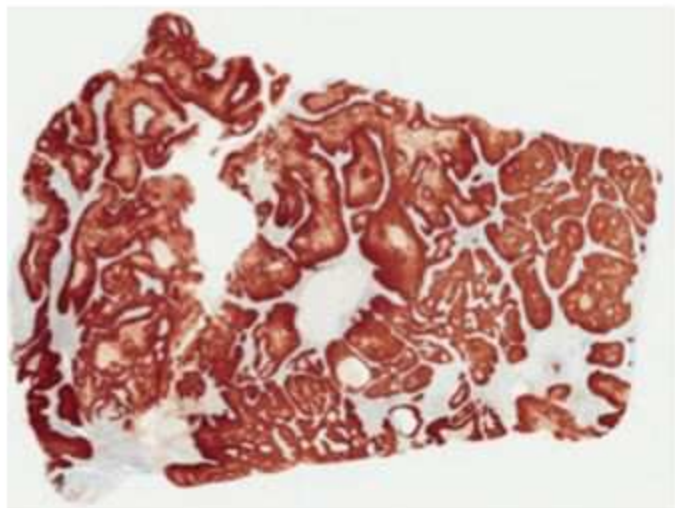
Figure 3: (A) p16 IHC of TCGA-BA-5555. Confirmed HPV (-), *CDKN2A* WT, *NSD1* mutant, elevated CDKN2A mRNA and p16-high by RPPA. **(B)** Scatterplot showing p16 expression assessed by RPPA versus *CDKN2A* gene expression as measured by RNA-seq (Spearman ρ = 0.63, p < 2.2e-16). Shapes and colours of points refer to HPV, *CDKN2A* and *NSD1* status (HomDel = homozygous deletion; HPV-/WT = HPV (-) / *NSD1* WT; HPV+/MUT = HPV (+) / *NSD1* mutant; silencing / non-silencing refers to the predicted effect of *CDKN2A* mutation on protein expression (loss of protein expression). **(C)** Boxplot showing *CDKN2A* expression stratified by *CDKN2A* copy number, HPV status and *NSD1* mutation status (n=494). [Print in colour]

Supplemental Material

Age

Mean = 62.14
Standard deviation = 12.28

Gender

Female - 61
Male - 149

Ethnicity

Hispanic/Latino - 13
Not Hispanic/Latino - 188

Histology

Head and neck squamous cell carcinoma - 210

Site

|	
Alveolar ridge – 6	HPV(-) 4, HPV(+) 2
Base of tongue – 6	HPV(-) 4, HPV(+) 2
Buccal mucosa – 7	HPV(-) 7, HPV(+) 0
Floor of mouth – 24	HPV(-) 24, HPV(+) 0
Hard palate – 3	HPV(-) 2, HPV(+) 1
Hypopharynx – 1	HPV(-) 1, HPV(+) 0
Larynx – 58	HPV(-) 57, HPV(+) 1
Lip – 2	HPV(-) 2, HPV(+) 0
Oral cavity – 37	HPV(-) 36, HPV(+) 1
Oral tongue – 61	HPV(-) 57, HPV(+) 4
Oropharynx – 1	HPV(-) 1, HPV(+) 0
Tonsil – 4	HPV(-) 1, HPV(+) 3

Supplementary Table 1: Details on the head and neck cancer samples of the TCGA cohort for which RPPA (Reverse Phase Protein Array) data were available (n=210).

	ESTIMATE	ST. ERROR	T VALUE	P VALUE	
(INTERCEPT)	0.79435	0.150537	5.277	2.47E-07	***
CDKN2A CNA	0.484486	0.06152	7.875	5.68E-14	***
CDK4 CNA	0.00571	0.105709	0.054	0.957	
CDK6 CNA	-0.00554	0.076725	-0.072	0.9425	
RB1 CNA	0.002373	0.07957	0.03	0.9762	
CCND1 CNA	0.032629	0.058314	0.56	0.5762	
CDKN2A MUTATION	0.336628	0.144547	2.329	0.0205	*
TP53 MUTATION	0.009346	0.157208	0.059	0.9526	
RB1 MUTATION	-0.0677	0.351622	-0.193	0.8474	
HPV POSITIVE	1.504397	0.276084	5.449	1.03E-07	***
SUBSITE - OROPHARYNX	-0.3045	0.206869	-1.472	0.142	

Supplementary Table 2: Linear model coefficients from a regression of p16 protein levels against genomic events, HPV status and anatomic subsite. CNA=copy number alteration.
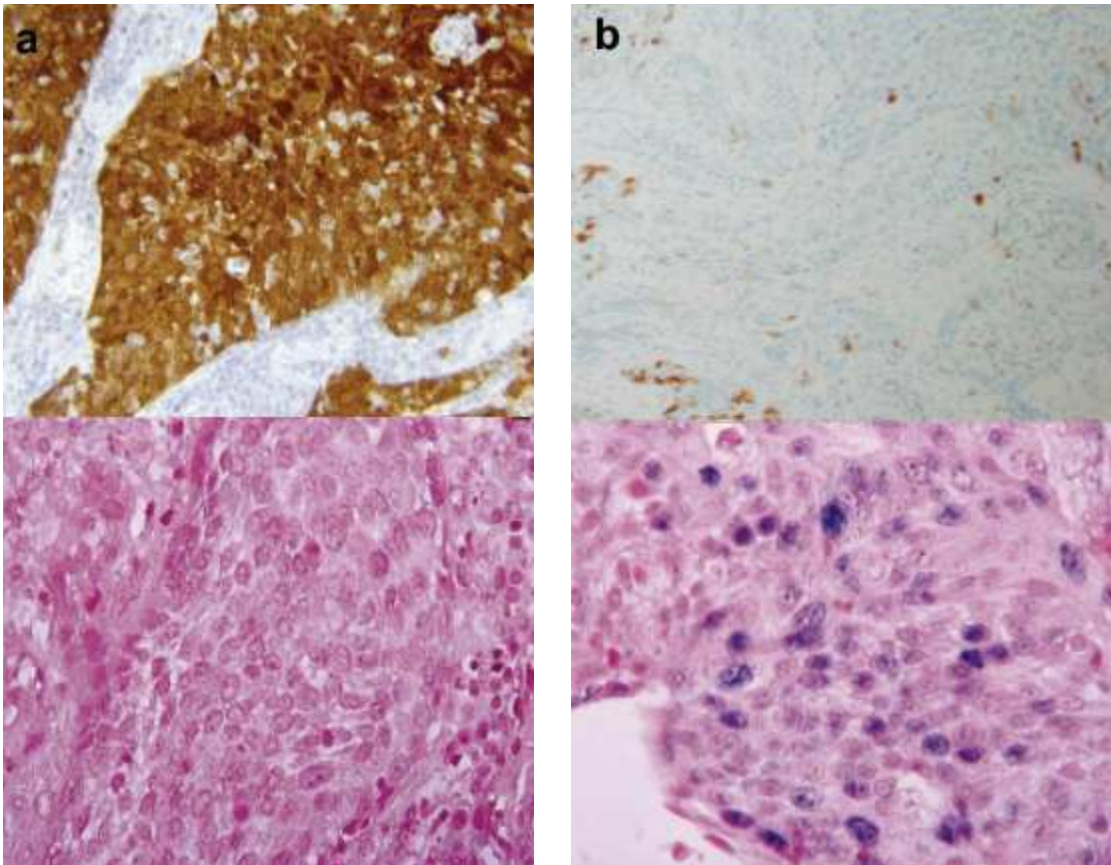
Supplemental Figure 1: Examples of p16 immunohistochemistry and HR HPV-DNA-ISH non-concordant results. Patient A, upper slide, p16 positive tumour by immunohistochemistry showing strong nuclear and cytoplasmic staining (×400); lower slide, HR HPV negative tumour by ISH (×600). Patient B, upper slide shows a p16 negative tumor by immunohistochemistry (×400); lower slide, HR HPV positive tumor by ISH (×600). These results are in line with RT-qPCR results for high-risk HPV RNA, obtained from the same samples. The p16$_{INK4a}$ IHC was carried out on FFPE tissue using a mouse mAb (BD Biosciences, Franklin Lakes, NJ, USA), as described previously (L. Masterson et al., Investigating the role of human papillomavirus in squamous cell carcinoma of the temporal bone. Head Neck Oncol 2013; 5(2): 22). DNA in situ hybridization consisted of a probe directed against high-risk HPV subtypes 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, and 66 (INFORM HPV III; Ventana, Tucson, AZ, USA), as described previously (T. Grogan, Interpretation Guide for Ventana INFORM HPVzr Probes In Situ Hybridization (ISH) Staining of Cervical Tissue; 2006).
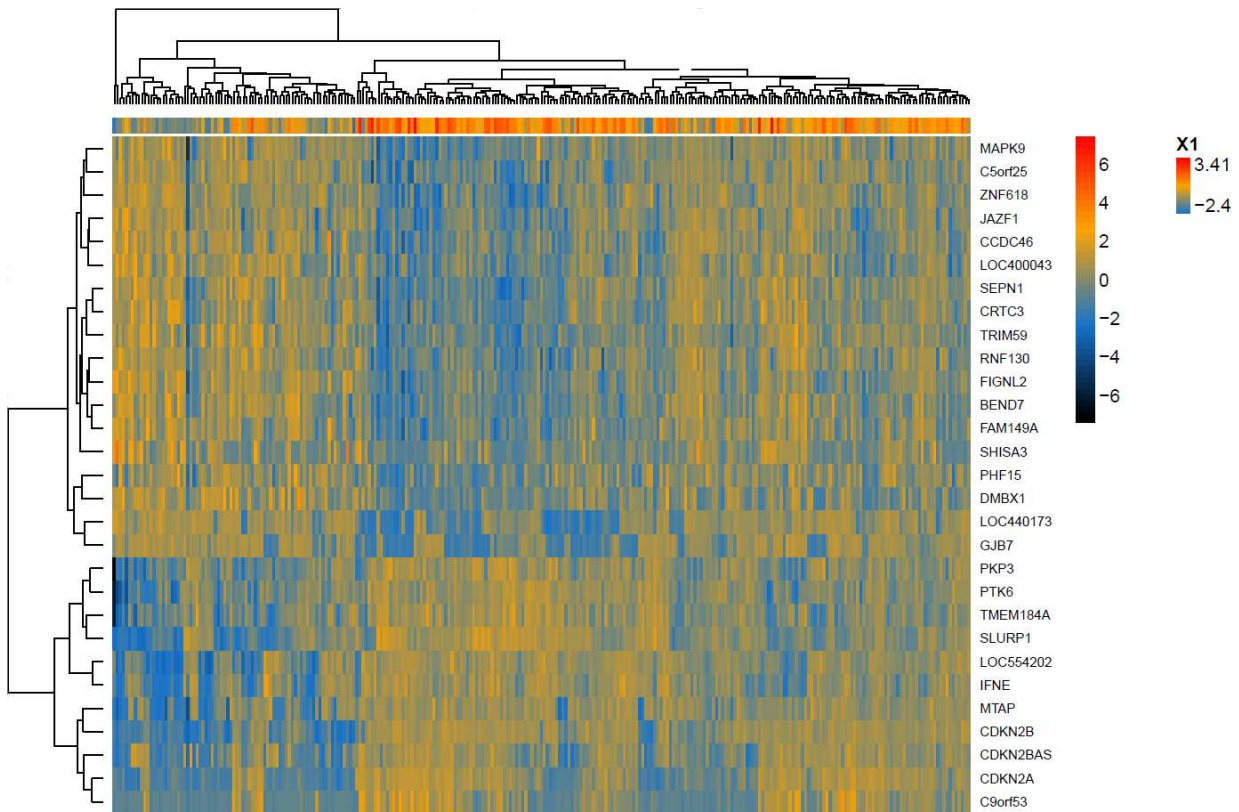
Supplemental Figure 2: Analysis of RNA-seq data using limma-voom identified a set of 29 genes whose expression was linearly associated with p16 RPPA protein levels (FDR < 0.05).