Received: 24 February 2018 | Revised: 21 May 2018 | Accepted: 22 May 2018

DOI: 10.1002/jccs.201800072



JOURNAL OF THE CHINESE CHEMICAL SOCIETY

ARTICLE

# Network based identification of different mechanisms underlying pathogenesis of human papilloma virus-active and human papilloma virus-negative oropharyngeal squamous cell carcinoma

Maryam Iman[1] | Zahra Narimani[2,3] | Iman Hamraz[4,5] | Ebrahim Ansari[2]

[1]Chemical Injuries Research Center, Systems Biology and Poisonings Institute, Baqiyatallah University of Medical Sciences, Tehran, Iran

[2]Department of Computer Science and Information Technology, Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan, Iran

[3]Research Center for Basic Sciences and Modern Technologies (RBST), Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan, Iran

[4]Department of pharmaceutics, Faculty of Pharmacy, Baqiyatallah University of Medical Sciences, Tehran, Iran

[5]Department of biostatistics, Faculty of medicine, Arak university of medical sciences, Arak, Iran

**Correspondence**
Maryam Iman. Chemical Injuries Research Center, Systems Biology and Poisonings Institute, Baqiyatallah University of Medical Sciences, Molla-Sadra Street, Tehran, P.O. Box: 19945-581, Iran.
Email: iman1359@yahoo.com; imanm@bmsu.ac.ir; m-iman@alumnus.tums.ac.ir
Zahra Narimani. Department of Computer Science and Information Technology, Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan 45137-66731, Iran.
Research Center for Basic Sciences and Modern Technologies (RBST), Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan 45137-66731, Iran.
Email: narimani@iasbs.ac.ir

**Funding information**
Baqiyatallah University of Medical Sciences

A different molecular mechanism underlying human papilloma virus (HPV)-negative and HPV-active pathogenesis is responsible for better response to therapies in HPV-associated oropharyngeal squamous cell carcinoma (OPSCC). In this study, we aim to provide an insight into molecular basis underlying this distinction and introduce possible targeted therapies for each phenotype. Using weighted gene co-expression network analysis (WGCNA), our aim was to identify not only differentially expressed genes but also significant coexpressed gene modules responsible for genotype and phenotype distinctions between HPV-active and HPV-negative samples. Recognizing differentially expressed genes in each module indicates key regulators that may be ignored in an analysis only based on differential gene expression study. Two modules are investigated in detail in our analysis, related to JAK–STAT dysregulation in HPV-negative samples, and disruption of cell fate commitment possibly induced by overexpression of BCL2 is observed in the HPV-active cohort. The existence of differentially expressed oncogenes and potential miRNA role is investigated in our analysis. The other significant module related to keratinization, keratinocyte differentiation, and intermediate filament cytoskeleton organization was discovered in the resulting co-expression network. A considerable number of genes was downregulated in HPV-active samples in the relative module, postulating the impairment of cytoskeleton-related gene expression caused by HPV intervention.

**KEYWORDS**

BCL2, biological module discovery, co-expression network, HPV-associated OPSCC, JAK–STAT pathway, keratinization, weighted gene co-expression analysis

## 1 | INTRODUCTION

Human papilloma virus (HPV) is a group of heterogeneous viruses with more than 200 identified subtypes, among which 15 are associated with high tumourigenic potential. These are nonenveloped double-stranded DNA viruses with an affinity to epithelium layer of mucosa. The association between high-risk subtypes and cervical cancer, oropharyngeal squamous cell carcinoma (OPSCC) as well as some rare cases of anogenital cancers has been well demonstrated. It has been demonstrated that more than 90% of HPV-positive OPSCC are induced by high-risk HPV16.[1] It has also been demonstrated that, in spite of the reduction in the overall incidence of head and neck squamous cancers in the past decade, the rate of HPV-positive OPSCC has increased worldwide.[2] HPV is sexually transmitted, and it has been indicated that the prevalence of uncontrolled sexual behaviors, especially oral sex, during the last decades is

responsible for the increasing incidence of HPV-positive OPSCC. HPV inactivates two tumor suppressor proteins, P53 and retinoblastoma (RB), that leads to cell cycle arrest, and as a consequence, the progression of the virus into the cell cycle and genomic integration into the host cells become possible.[3] It has been shown that two viral oncoproteins, E6 and E7, are responsible for the observed transformation in the epithelium cells through inactivation of p53 and Rb, respectively. Inactivation of Rb, in turn, results in the overexpression of p16, which is considered a biomarker,[4,5] and the detection of serum antibodies against HPV, for example, E6 antibodies, in the prediagnostic patients is of great predictive value for HPV positive OPSCC.[6,7]

Patients recognized with HPV-associated oropharyngeal cancers (OPC) have been observed to result in a better response to therapies, and the prognosis for HPV-associated OPC is also better compared to the other OPC cases [8]. This has led to the classification of OPC into HPV-positive and HPV-negative cases. Based on viral load and oncogene expression, the HPV-positive groups are further classified into two subgroups: HPV-active and HPV-inactive [9]. Race has also been speculated to affect the diversity of disease stage and survival, for example, different behavior in prevalence, mortality rate, and HPV status has been observed between European American (EA) and African American (AA) cases.[10,11] Transformation of keratinocytes from epithelia observed in HPV-infected patients is the result of distorting cell cycle-regulatory pathways by interference of viral oncoproteins causing OPC (as well as anogenital cancer) progression[12]; however, the exact procedure by which the keratinocytes are transformed is still under study, and the molecular mechanism underlying HPV-related OPSCC is not clearly identified.[13]

Studies so far have been focused on gene expression pattern differences between HPV-negative and HPV-positive patients. Considering gene co-expression patterns can improve previous results as it takes into account intergene relationships, which can lead to finding modules of genes taking part in biological processes responsible for malignant progression of OPC, also demonstrating molecular mechanisms underlying differences between OPC classes. In this work, we focused on two groups: HPV-active, and HPV-negative patients.

Our main goals are to find specific biomarkers for discriminating different types of OPSCC (HPV active and negative) and to propose potential targeted therapies specific for each group. Considering these two groups (HPV-active and HPV-negative), our study will focus only on biological processes underlying the "distinction" of the specified groups (and therefore giving insight to targeted therapy of OPSCC subtypes based on HPV activity); for example, EGFR or PI3K/AKT pathways, which are known as two important pathways (and therefore targets for therapies) in the pathogenesis of head and neck cancers, are not highlighted in our analysis because their behavior has not been observed to be significantly different in these groups.[14,15] This study is composed of two main analyses: gene co-expression network analysis and consensus network analysis. Co-expression network analysis indicates genetic modules extracted by comparing HPV-active and HPV-negative expression profiles. Gene Ontology enrichment analysis of these modules demonstrates significant biological processes correlated with HPV-activity status, suggesting potential different targeted therapies for each disease subtype. Consensus analysis indicates relationships between co-expression modules in different classes. Targets of miRNAs with known dysregulation in HPV-related OPSCC are also identified in the relative modules, and the possible contribution of identified miRNAs is investigated in the pathogenesis of relative tumors.

## 2 | MATERIALS AND METHOD

### 2.1 | Dataset

Gene expression data of OPSCC patients, GSE55542,[10] were obtained from Gene Expression Omnibus (GEO). The dataset consists of overall 36 samples: 12 HPV-active cases (of which only 1 case is AA, and the rest are EAs), 8 HPV-inactive cases (of which 4 belong to AA and 4 belong to EA races), and 16 HPV-negative case (consisting of 8 AA and 8 EA cases). The main study by Tomar et al.[10] includes 65 oral or oropharyngeal fresh frozen tissue samples (from AA and EA patients treated in South Carolina between 2010 and 2012) from which microarray data are provided after image analysis (using Agilent Technologies platform—Feature Extractor Software version 10.7.3.1). The resulting data are then background corrected and log2 transformed. The dataset contains 38 oropharyngeal tumor samples as described previously in this section, which is used in the current study. As reported by the authors, mortar and pestle systems (CryoGrinder™, CryoCooler™, OPS diagnostics LLC) and TRIzol® Reagent (Ambion®, by Life Technologies) are used to extract total RNA and DNA from the samples.[10] The goal of the main study was to study differential gene expression patterns in HPV-related tumors in AA and EA patients; the study showed 10% of AA tumors (out of total 65) and 39% of EA tumors were HPV active (HPV DNA-positive and also E7 mRNA expressed). The study demonstrated that HPV-inactive tumors represent a different expression pattern from both HPV-active and HPV-negative groups in oral and oropharyngeal tumors. The analysis in the original study focuses on differential gene expression, and a network-based approach has not been used to compare the groups.

A subset of 12 HPV-active cases and 16 HPV-negative cases (total 28) of all oropharyngeal tumors was selected for the current study. The original dataset was normalized. In the pairwise analysis and network reconstruction between

these two groups, we normalized the subsample data again using Variance Stabilizing Normalization library.[16] The data were preprocessed to remove genes with missing values and outlier samples using a hierarchical clustering approach. No outlier was detected in the samples. For network reconstruction, we limited our study to 5,000 probes having larger values of coefficient of variation among all samples for each study. In our study, we focus on OPC samples only, and our goal is to convey a thorough analysis (based on both differential expression and network-based analysis) in order to better understand the similarities and differences in the two group. Network-based analysis indicates the important biological processes responsible for the differences between the two groups and can explain the different pattern of survival and response to treatments in the two group.

### 2.2 | Construction of gene co-expression network

Methods based on correlation,[17] information theory-based techniques,[18–20] and Bayesian Network-based methods[21,22] present approaches used to construct gene networks. A gene co-expression network is a network in which genes are represented by nodes, and an edge between two genes represents high correlation between the expression values of the corresponding (two) genes. Pearson correlation is one possible method for computing the gene–gene correlations in such a network. In a correlation-based gene co-expression network, the correlation values are usually compared to a prespecified threshold, and only edges with high confidence (correlation value above the threshold) are considered to be present in the network. Choosing the threshold is itself one of the challenges of network inference methods. The benefit of using correlation to construct such networks benefit includes having low computational cost. Here, we used the R package Weighted Gene Co-expression Network Analysis (WGCNA) for inferring gene networks. WGCNA uses a soft-thresholding power (a power to which the similarity values are raised in order to compute the similarity matrix), which is selected such that the resulting generated network has a scale-free topology.[23]

### 2.3 | Clustering of co-expression network

The adjacency matrix generated by WGCNA is converted to a generalized version of the topological overlap matrix[24] in order to reduce the effect of noise. A matrix of dissimilarities is then computed from the topological overlap matrix, which is further used as input in the hierarchical clustering algorithm, aiming to generate gene modules. As the generated network is constructed from the samples belonging to two different states (e.g., HPV-positive and HPV-negative), the edges in the co-expression network are representative of correlations responsible for differences in the two states. Modules in such a network represent the groups of genes with high collaboration, which are candidates for genes belonging to biological processes underlying the phenotypical differences across samples. Gene modules finally undergo gene ontology enrichment using the DAVID functional annotation tool.[25] Analyzing biological process and function results from GO demonstrates the relevancy and statistical validity of each module.

### 2.4 | Consensus analysis

Much literature focuses on studying co-expression modules. While studying these modules demonstrates significant biological processes active in certain conditions, the relationship between these modules is of biological importance and should be studied.

The goal of consensus analysis is to discover if intermodule relationships (inter-module correlations) is preserved among the two states (HPV-negative and HPV-active). For that purpose, the adjacency matrices for the HPV-negative and HPV-active sample are extracted separately. These adjacency matrices, which are in fact topological overlap matrices or TOMs,[24] are further used in order to construct a consistent adjacency matrix (consensus_TOM, in which each element is the minimum of the corresponding elements in $TOM_{HPV\text{-}negative}$ and $TOM2_{HPV\text{-}active}$). Note that the modules extracted here are the ones constructed by the clustering of the consensus network (the network represented by consensus_TOM) and are therefore different from the modules extracted in the "module detection and functional enrichment" section.

The gene expression profile of each module is finally summarized and represented by an eigengene, and the eigengene network in which the co-expression values between every two eigengene represents module correlation information (maintaining the sign of co-expression value) can be used to investigate intermodule relationships.

Here, we constructed eigengene networks [26] for different datasets mentioned in previous section. Eigengene networks are meta-networks representing the similarities between gene modules. Modules that are shared between the two datasets (e.g., networks shared between HPV-active and HPV-inactive conditions) are called consensus modules. Comparing consensus modules in the two conditions allows one to survey if the relationship between two modules is preserved under the two conditions or not. This analysis discloses information about possible network rewiring. Consensus analysis is conducted using the WGCNA package.

## 3 | RESULTS AND DISCUSSION

Our computational result consists of two sections; first, we present the results regarding modules discovered in c-expression networks, and then, the consensus analysis results are presented. The third part of this section (i.e., modules overview and association with HPV activity state) focuses on biological analysis of the discovered modules and the relevancy of the results.

### 3.1 | Module detection and functional enrichment

In this study, we used a subset of dataset GSE55542, including 12 samples belonging to the HPV-active category (GSM1338930-GSM1338941) and 16 samples from the HPV-negative category (GSM1338950-GSM1338965). The samples were normalized using the VSN method,[16] and 5,000 probes with highest coefficient of variation values were used for network reconstruction. After clustering the network using hierarchical tree clustering (hclust function of WGCNA package), 19 modules were detected. The 19 modules were further merged into 13 modules by comparing the distance between module eigengenes. Each module is represented (and named) by means of a color name. The genes not assigned to any module in our analysis are grouped into a single module called the "gray" module, which is not considered in any further analysis. The clustering dendogram is provided in File S1; the module colors (before and after merging) are represented along with the dendogram.

Gene details and module membership information are presented in File S2. Each module is identified with a color name. The extent to which a gene belongs to each module is reported in the mentioned file such that, for each gene, two values is available per module: module membership (column name: MM. < moduleColor>) and *p*-value reported for module membership (column name: p.MM. < moduleColor>).

For each module, an eigengene [26] is computed using the moduleEigengenes function of WGCNA package. The module eigengene is a representative of the corresponding module, and the eigengene network can be used in order to study the relationship between co-expression network modules and also to study the relationship of modules to traits. The value computed for each eigengene is correlated with the expression level of genes in the relevant module (preserving the sign), and the module–trait relationship can be computed by applying a correlation function, for example, Pearson correlation coefficient, with module eigengene vector and trait vector as input parameters (moduleTraitCor = cor(MEs, datTraits, use = "p"), in which the module TraitCor is a vector representing the correlation value of each trait to each module, which is later visualized by means of a heatmap; MEs is the matrix in which each row represents a sample and columns specify module eigengene list, datTraits is the list of trait information (filled by 0 and 1 s representing HPV-negative or HPV-active state of each sample), and use = "p" is a parameter for selecting Pearson correlation coefficient.

The heatmap plot shown in Figure 1 demonstrates the module–trait correlation along with the *p*-value assigned to each correlation value.

As an be observed from the heatmap plot, there are four modules (brown, orange, green-yellow, pink) with an absolute trait correlation above 0.75 and very low *p*-value; there are also three modules (tan, magenta, purple) with an absolute trait correlation between 0.7 and 0.3 and a *p*-value less than .05. List of probes existing in each module can be found in Supplementary file Sup2-geneInfo-NegVsActive.csv. In the discussion, we focus our investigations on the first four modules (brown, orange, green-yellow and pink), which are more relevant to the HPV state of the disease in terms of correlation coefficient value and also biological relevancy. The modules are discussed in the order of their absolute correlation value. Having more absolute correlation value means that the corresponding module has more relevancy to the trait.

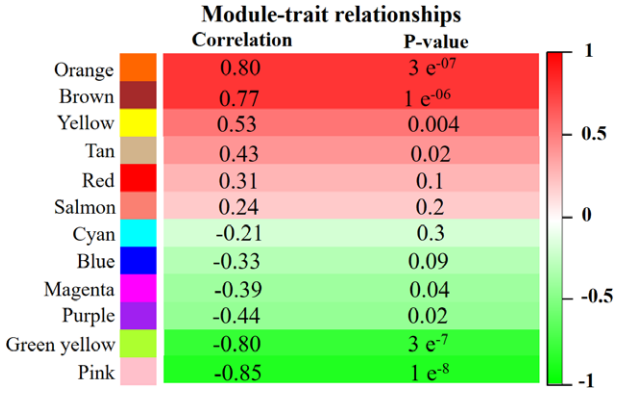


**FIGURE 1** Heatmap plot demonstrating relationship between modules and trait information (HPV-negative, HPV-active); for each module, the correlation value along with a *p*-value is reported. The correlation is reported with regard to HPV-negative trait (i.e., positive correlation means the genes included in the corresponding module are overexpressed in HPV-negative samples in comparison to HPV-active, and negative values means the regarding module genes are overexpressed in HPV-active samples)

The pink module consists of 534 genes, including ZNF541, GRIN2C, CELF4, TCP11, LOC375196, CDKN2C, LOC254559, SMC1B, STAG3, KLHL35, C19orf57, SYCP2, FLJ45482, RANBP17, KIF25, SYCE2, CHDH, TCAM1P, CBX5, DDX25, and TLX2, which are significantly overexpressed in HPV-active samples, and their overexpression in HPV-relevant OPSCC has also been observed in other studies. MYB, MYCN, and other overexpressed genes in the pink module have been reported as miRNA targets. GO enrichment analysis reveals that genes in this module are related to biological processes such as cell developmental process and cell fate commitment (relevant genes such as CDSN, CYP26B1, IL20, KRT16, KRT17, LCE1B, LCE1C, LCE2A, SPRR2G, SPRR4, and etc.). Other genes such as SYCP2, TCAM1P, and STAG3 are among other significantly overexpressed genes in the pink module, which are shown to have an effect on the progression of disease in HPV-positive patients. BCL2, as an apoptosis regulator, is another important gene within the pink module.

The next module we focus on is the brown module, containing 219 genes that are mostly upregulated in HPV-negative samples. This module is associated with GO biological process terms such as epidermis development, skin development, keratinization and keratinocyte differentiation,

intermediate filament cytoskeleton organization, regulation of tyrosine phosphorylation of STAT protein, and regulation of JAK–STAT cascade. The brown module contains several oncogenes (such as ABL2, RAB38, RAB3B, MAFA, and THRB); miRNA target genes (such as F2RL1, AREG, CAV1, CYP26B1, GJA1, IL31RA, SERPINB5, SPRR2G, SNAI2, THRB); and other relevant genes such as KRT6A, KRT6C, KRT9, PKP1, ALOX12B, and KRT1.

The green-yellow and orange modules, consisting of 219 and 34 genes, respectively, are the next two modules highly related to HPV status. Orange-contains genes such as NT5E, RBM44, AIG1, AMIGO2, MPP6, CD209, RNU2–2, and PPP1R36, which are overexpressed in HPV-negative samples. The genes within the orange module are mostly known to have a role in metastasis and cell–cell adhesion. Finally, the green-yellow module consists of 219 genes (such as RAB42, RAB36, KCNB2, XKR4, KELCDKN2A), which are overexpressed in HPV-active samples. Activation of I-kappaB kinase/NF-kappaB signaling and activation of NF-kB signaling are relevant biological processes significant in the green-yellow module. A thorough discussion on these modules is provided in the section 'Modules overview and association with HPV activity state'.

The result of GO functional enrichment analysis (for GO term biological process) of pink and brown modules is represented in Tables 1 and 2, respectively. A summary of enrichment analysis of all modules is available in File S3.

**TABLE 1** GO functional enrichment of pink module

| GO term | *p*-value |
|---|---|
| Negative regulation of cell development | 4/14E−06 |
| Cell fate specification | 2/34E−05 |
| Embryo development | 3/88E−05 |
| Regulation of nervous system development | 8/12E−05 |
| Negative regulation of developmental process | 1/77E−04 |
| Peripheral nervous system development | 2/68E−04 |
| Cell differentiation | 2/81E−04 |
| Positive regulation of developmental process | 3/73E−04 |
| Regulation of neuron differentiation | 4/01E−04 |
| Regulation of neurogenesis | 4/20E−04 |
| Single-organism process | 4/22E−04 |
| Negative regulation of cell differentiation | 6/78E−04 |
| Synapse assembly | 0/00106 |
| Regulation of cell development | 0/00110 |
| Glial cell differentiation | 0/00130 |
| Drug metabolic process | 0/00131 |
| Developmental process | 0/00148 |
| Calcium-independent cell–cell adhesion via plasma membrane cell-adhesion molecules | 0/0019 |
| Cell fate commitment | 0/0020 |

**TABLE 2** GO functional enrichment of the brown module

| GO term | *p* value |
|---|---|
| Epidermis development | 5/44E−22 |
| Skin development | 2/32E−18 |
| Keratinization | 7/20E−18 |
| Keratinocyte differentiation | 1/85E−15 |
| Epithelium development | 2/31E−13 |
| Cell differentiation | 2/57E−06 |
| Anatomical structure development | 5/18E−7 |
| Multicellular organism development | 2/28E−05 |
| Single-organism process | 2/65E−05 |
| Establishment of skin barrier | 4/56E−05 |
| Single-organism developmental process | 5/84E−05 |
| Chemical homeostasis | 6/92E−05 |
| Regulation of water loss via skin | 7/10E−5 |
| Intermediate filament cytoskeleton organization | 1/15E−04 |
| Regulation of tyrosine phosphorylation of STAT protein | 1/40E−04 |
| Actin filament-based movement | 1/44E−04 |
| Tyrosine phosphorylation of STAT protein | 1/92E−04 |
| Regulation of JAK–STAT cascade | 3/31E−04 |
| Regulation of STAT cascade | 3/31E−04 |
| Positive regulation of tyrosine phosphorylation of STAT protein | 6/73E−04 |
| Multicellular organismal process | 7/71E−04 |
| JAK–STAT cascade | 8/25E−04 |
| Positive regulation of tyrosine phosphorylation of Stat3 protein | 9/20E−04 |
| STAT cascade | 9/95E−04 |

### 3.2 | Consensus analysis

As mentioned in the Methods section, the goal of consensus analysis is to extract the same modules on two different datasets (e.g., data regarding patients' different HPV status) and investigate the intermodule preservation in the two datasets. Note that consensus modules are extracted independent of the modules extracted in the previous section, and having the same name does not mean any correlation between the two sets of modules.

Figure 2 demonstrates the result of consensus analysis on HPV-negative and HPV-active samples.

Significant biological processes in the turquoise module are related to immune system process, leukocyte activation, lymphocyte activation, T cell activation and aggregation, and positive regulation of immune system process. For the pink module, the significant processes include xenobiotic metabolic process, glutathione metabolic process, cellular response to jasmonic acid stimulus, and oxidation–reduction process. Studying interplay between these modules indicates certain network rewiring in HPV-active phenotype in contrast to HPV-negative. Full enrichment analysis is provided in File S4.

### 3.3 | Modules overview and association with HPV activity state

In this section, we focus on comparing modules constructed from a gene expression dataset of HPV-active and HPV-negative samples. Patterns of co-expression, possibly responsible for the different molecular mechanism

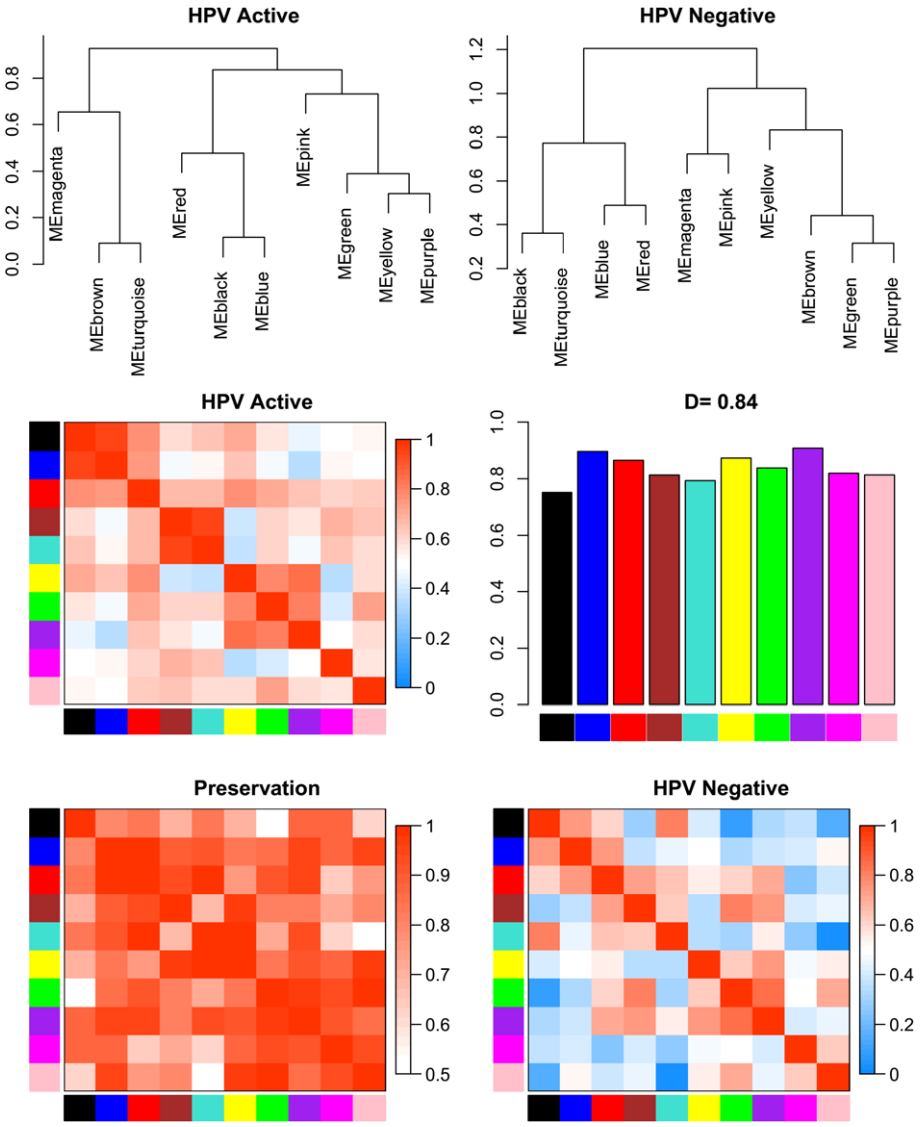


**FIGURE 2** Ten consensus modules are detected in HPV-active and HPV-negative samples. The preservation plot depicts low preservation of intermodule correlations in several cases (discussed in the text). The bar plot represents mean preservation of each module in two cohorts

underlying the distinction of the phenotypes, can highlight gene co-expression modules, which can further be investigated in our module enrichment and functional analysis. Significantly differentially expressed genes in each module are provided in Files S5–S8. In order to find the important differentially expressed genes in modules (reported in File S5–S8), first, the top 1,000 significantly differentially expressed probes among all 5,000 probes are extracted. Probes within these 1,000 ones belonging to each module are considered important, differentially expressed probes in each module. R package, Limma,[27] was used to calculate log2 expression fold change values.

The most significantly related to the trait module is the pink module (Figure 1), with a correlation of −0.85 and a *p*-value of 1e−8. This module consists of the gene SYCP2 (with high module membership of 0.8). Increased expression level of SYCP2 in HPV-related OPSCC is reported in previous studies[28,29] and is also observed in our differential gene expression analysis (logFC 6.35, adj.*p*-value 8.9e−7). It also has been shown that early-stage HPV-positive oropharyngeal carcinoma can be predicted by deregulation of SYCP2 [30]. Another gene in pink module is TCAM1P (testicular cell adhesion molecule 1) with overexpression in HPV-active samples (logFC 7.02 and adj.*p*-value of 6.54e−6), and STAG3 (stromal antigen 3) with overexpression in HPV-related samples (logFC 5.56, adj.*p*-val 2.04e−7) is also a member of the pink module (both with module membership > 0.65). All three genes are known as markers for both cervical- and HPV-related head and neck cancers.[31] Top upregulated genes in HPV-active phenotype include: ZNF541, GRIN2C, CELF4, TCP11, LOC375196, CDKN2C, LOC254559, SMC1B, STAG3, KLHL35, C19orf57, SYCP2, FLJ45482, RANBP17, KIF25, SYCE2, CHDH, TCAM1P, CBX5, DDX25, and TLX2. Overexpression of several of these genes has been observed in different studies.[32,33]

GO functional enrichment analysis conducted on this module demonstrates terms related mostly to the cell developmental process. The list of genes in the pink module

associated with cell fate commitment is provided in Table 3. Within these genes, TLX3 is overexpressed in HPV-active samples with a significant logFC of 6.14 (adj.*p*-value of 3e −5). BCL2 is a regulator for apoptosis, [34] but it is not considered an oncogene itself; however, in the presence of a proto-oncogene, they together may lead to uncontrolled cell development. An overexpression of BCL2 with a logFC of 4.12 and adjusted *p*-value of .0003 is observed in HPV-active samples (although this is not among the top 100 differentially expressed genes in our analysis). The role of BCL2 overexpression in apoptotic resistance and tumorigenesis in lymphoma, leukemia, and also solid tumors (ovarian, colorectal, and also head and neck cancers) has been reported in several previous studies.[35–39] Downregulation of miR-143 and miR-126 observed in HPV-positive OPSCC samples[13,40] may be possible causes for upregulation of BCL2 in HPV-active samples compared to HPV-negatives. BCL2 targeted therapy has been proposed to be effective for head and neck squamous cell carcinoma.[41] Due to our observation, significant overexpression of BCL2 for HPV-active samples (compared to HPV-negative ones) suggest that HPV-active OPSCC is a better candidate for therapies targeting BCL2; therefore, drugs inhibiting BCL2 are expected to be more effective in HPV-active OPSCC tumors compared to HPV-negative OPSCC ones.

The MYB proto-oncogene is one of the oncogenes present in the pink module with an overexpression of 4.5 logFC (adjusted *p*-value .0004). MYCN, with an overexpression of 4.4 (adjusted *p*-value of .0008), belonging to pink the module is the target of downregulated miRNA-126.[13] The role of MYCN in the progression of the cell life cycle and induction of proliferation is known in different cancer types (neuroblastoma, breast cancer, lung cancer, Wilms' tumor).[42] Overexpression of RAB26, a member of the RAS oncogene family and another member of the pink module, with logFC of 3.52 and adjusted *p*-value of .00025 is observed. RAB9B, which is not differentially expressed in HPV-active samples, is also another member of the pink module.

**TABLE 3** List of genes belonging to the pink module that are associated with cell fate commitment

| Gene symbol | Description |
|---|---|
| *BCL2** | Apoptosis regulator |
| *DMRTA2** | DMRT like family A2 |
| *DSCAML1** | DS cell adhesion molecule like 1 |
| *EYA2* | EYA transcriptional coactivator and phosphatase 2 |
| *ISL1* | ISL LIM homeobox 1 |
| *POU4F1** | POU class 4 homeobox 1 |
| *SOX2* | SRY-box 2 |
| *SOX6* | SRY-box 6 |
| *SOX8** | SRY-box 8 |
| *TLX3** | T-cell leukemia homeobox 3 |
| *FOXA1* | Forkhead box A1 |
| *IL23A* | Interleukin 23 subunit alpha |
| *MESP1** | Mesoderm posterior bHLH transcription factor 1 |
| *PAX6** | Paired box 6 |
| *PTCH1* | Patched 1 |

*Genes differentially expressed with logFC > 2 and adjusted *p*-value < .01.

Table 4 summarizes information regarding the pink module's genes related to the regulation of cell proliferation/development, which are targeted by miRNAs reported as differentially expressed in HPV-active samples (compared to HPV-negative). Information regarding miRNA target genes was obtained from mirtarbase.[43]

Another module significantly related to state module is the brown module in which most of the genes were upregulated in HPV-negative samples. GO terms in detail consist of: epidermis development, skin development, keratinization and keratinocyte differentiation, epidermal cell differentiation, epithelium development, anatomical structure development, establishment of skin barrier, intermediate filament cytoskeleton organization, regulation of tyrosine phosphorylation of STAT protein, actin filament-based movement, actin-mediated cell contraction, tyrosine phosphorylation of STAT protein, regulation of JAK–STAT cascade, regulation of STAT cascade, and positive regulation of tyrosine phosphorylation of Stat3 protein.

Simultaneously prompting cell immortalization and decreased expression of cytokeratin induced by HPV has been postulated by several studies.[44] The list of genes from the brown module belonging to the keratinocyte differentiation GO biological process is provided in Table 5. Differentially expressed genes with logFC > 2 and *p*-value < .01 are specified with an asterisk sign. Other genes playing role in intermediate filament cytoskeleton organization (such as KRT6A, KRT6C, KRT9* (logFC −2.7 and adj *p*-value of 0.02), PKP1, ALOX12B, KRT1 * (logFC −3.7, 0.04)) and other keratins like KRT75 * (logFC −3.5, 0.01) are also present in the brown module. The brown module also contained five oncogenes: ABL2, RAB38 (member of RAS oncogene family), RAB3B (member of RAS oncogene family), MAFA, and THRB (thyroid hormone receptor, beta). ABL2 (logFC 3.6), RAB38 (logFC 3.6), RAB3B (logFC −3.3), and MAFA (logFC 2.9) are significantly differentially expressed in HPV-negative samples. All adjusted *p*-values are below 0.005. Overexpressed ABL2 (ABL proto-oncogene 2), MAFA (MAF Bzip transcription factor A), and RAB3B (member RAS oncogene family) play a role in the positive regulation of cellular process.

Among genes in the brown module, CAV1 (caveolin 1, logFC 3.9), EPO (erythropoietin, locFC 3.4), FLRT3 (fibronectin leucine-rich transmembrane protein 3, logFC 4), IFNE (interferon epsilon), IL18 (interleukin 18), IL20 (interleukin 20, logFC 3.6), IL24 (interleukin 24, logFC 3.6), IL31RA (interleukin 31 receptor A), and CLCF (cardiotrophin-like cytokine factor 1, logFC 2.7) belong to the GO cellular process related to regulation of JAK–STAT cascade. Most of these genes are downregulated significantly in

**TABLE 4** Cell proliferation/development-related genes in the pink module targeted by miRNAs known to be dysregulated in HPV-active samples

| Gene | HPV-actvlogFC | Adj. *p*-val | Module membership | miRNA | miRNA in HPV+ |
|---|---|---|---|---|---|
| BCL2 | 4.12 | .0003 | 0.78 | miR-143 | Downregulated |
| CACNB4 | 3.07 | .0002 | 0.59 | miR-26b | Downregulated |
| AMOT | 3.27 | .02 | 0.64 | miR-127-5p mirt/3p | Downregulated |
| ABCA3 | 4.8 | .0009 | 0.72 | miR-409-5p | Downregulated |
| ADAMTS17 | 3.18 | .007 | 0.65 | miR-432-5p / mir-432 | Downregulated |

**TABLE 5** List of genes belonging to the pink module that are associated with cell fate commitment

| Gene symbol | Description |
|---|---|
| CDSN* | ATP binding cassette subfamily A member 12 (ABCA12) corneodesmosin |
| CYP26B1* | Cytochrome P450 family 26 subfamily B member 1 |
| IL20* | Interleukin 20 |
| KRT16* | Keratin 16 |
| KRT17* | Keratin 17 |
| LCE1B | Late cornified envelope 1B |
| LCE1C | Late cornified envelope 1C |
| LCE2A* | Late cornified envelope 2A |
| LCE2B | Late cornified envelope 2B |
| LCE2C* | Late cornified envelope 2C |
| LCE2D* | Late cornified envelope 2D |
| LCE3B* | Late cornified envelope 3B |
| LCE3C* | Late cornified envelope 3C |
| LCE3D* | Late cornified envelope 3D |
| LCE3E* | Late cornified envelope 3E |
| PRR9* | Proline rich 9 |
| SPRR2G* | Small proline rich protein 2G |
| SPRR4* | Small proline rich protein 4 |

*Genes differentially expressed with logFC > 2 and adjusted *p*-value < .01.

HPV-active samples in contrast with HPV-negative ones. The signaling cascade of JAK–STAT is activated (observed in HPV-negative tissue) in response to different growth factor and cytokines [45]. The role of the JAK–STAT signaling pathway is known in proliferation, apoptosis, and oncogenesis and can be activated by ABL oncogenes [46]. ABL2 overexpression observed in HPV-negative samples is a potential cause of observed JAK–STAT dysregulation. The activation of JAK–STAT pathway has been observed in neoplastic tissue, [47–49] and the positive effect of RAS and JAK–STAT inhibitors in inducing apoptosis of cancer cells has been observed [50–53]. According to our analysis, the JAK–STAT pathway is significantly upregulated in HPV-negative samples relative to HPV-active ones. The role of the JAK–STAT pathway in angiogenesis, carcinogenesis, and elevated cell proliferation of head and neck cancer has been studied [54]. Our analysis confirms previous studies on HPV-associated OPSCC and demonstrates the idea that HPV-active tumor progression is conducted via a different mechanism (regulated via genes in the pink module). Drugs targeting ABL2 can be proposed as possible inhibitors for the JAK–STAT pathway, leading to apoptosis induction in HPV-negative OPSCC tumors but, due to our analysis, cannot be beneficial in case of HPV-active samples.

Some upregulated genes (selected from GO enrichment result related to epithelial development) in the brown module with downregulation of the miRNA-targeting gene (upregulation and downregulation in HPV-negatives is of interest) is listed in Table 6. Selected miRNAs are those significantly differentially expressed in HPV-associated OPSCC.[13,40] We suggest these genes (especially F2RL1, AREG, and CAV1 as significantly overexpressed in HPV-negative samples) for further studies. These genes and their targeting miRNAs are listed in Table 6. The two other modules significantly related to HPV positive–negative status are the green-yellow and orange modules, consisting of 219 and 34 genes, respectively. In orange, we recognized no oncogenes. Genes belonging to this module (all overexpressed in HPV-negative samples) are mostly significant in leukocyte cell–cell adhesion, which is also known to have a role in metastasis[55] and act as a biomarker in breast cancer.[56,57] NT5E and CD209 are significantly overexpressed cell–cell adhesion-related genes within this module.

In the green-yellow module, in which all of genes are overexpressed in HPV-active samples, two oncogenes RAB42 (overexpressed with logFC of 3.2, adjusted *p*-value of .013) and RAB36 (logFC of 2.6 and adjusted *p*-value of .03) were present. The green-yellow module suggests activation of I-kappaB kinase/NF-kappaB signaling in the HPV-active cohort. Activation of NF-kB signaling is known as a response to stimuli such as viral infection. Evidently, the pink and brown module represent different pathogenic pathways active in the samples, proposing that the distinction in phenotype and genotypical characteristics between HPV-active and HPV-negative samples and identified genes introduces potential targeted therapies inhibiting unwanted dysregulation in each cohort.

## 4 | CURRENT AND FUTURE DEVELOPMENTS

In this study, the distinct molecular mechanism underlying HPV-active and HPV-negative samples were investigated. Extraction of gene co-expression network resulting from the analysis demonstrated gene modules that have further undergone GO enrichment analysis. The result of enrichment analysis, besides the differential expression analysis and

**TABLE 6** Epithelial development-related genes in the brown module targeted by miRNAs known to be dysregulated in HPV-active samples

| Gene | HPV-negative logFC | Adj. *p*-val | Module membership | miRNA | miRNA in HPV-negative |
|---|---|---|---|---|---|
| F2RL1 | 4.9 | 1.16e−5 | 0.86 | miR-20b, miR-150, mir-93-5p | Downregulated |
| AREG | 4.4 | 1.44e−5 | 0.80 | miR-335 | Downregulated |
| CAV1 | 4.0 | .0006 | 0.82 | miR-107 | Downregulated* |
| CYP26B1 | 2.9 | .021 | 0.73 | miR-195, miR-15b | Downregulated |
| GJA1 | 3.3 | .004 | 0.82 | miR-222-3p | Downregulated |
| IL31RA | 2.8 | .019 | 0.62 | miR-335 | Downregulated |
| SERPINB5 | 2.6 | .033 | 0.78 | miR-335, miR-107 | Downregulated* |
| SPRR2G | 4.7 | .018 | 0.86 | miR-335 | Downregulated |
| SNAI2 | 3.2 | .003 | 0.81 | miR-9 | Downregulated |
| THRB | 3.0 | .013 | 0.66 | miR-15b, miR-320a | Downregulated* |

In the last column (miRNA in HPV-negative), the asterisk sign means that the opposite state (up/downregulation) has also been reported in the literature, but with lower logFC

investigation of miRNA target genes (for miRNAs showing differential expression patterns in HPV-associated OPSCC), provides a more thorough perspective of the biological processes responsible for difference in phenotype and response to therapy between HPV-negative and HPV-active malignant tissue. This study focuses on the differences between the two phenotypes (and not general characteristics of OPSCC). Consensus analysis represents major intermodule rewiring in the co-expression networks of the two subtypes. The strengths of our analysis are network reconstruction and module extraction and, finally, analyzing each module separately. This approach will keep our attention on genes with an important role in different relevant biological process, whereas applying merely differential expression analysis will ignore many of these genes only because they will not appear on the top of the differentially expressed gene list. Proposed directions for future research include investigation of proposed targeted therapies for each OPSCC subtype and a more detailed study on the proposed gene sets responsible for relative biological process. The study of mechanisms by which HPV affects the proposed pathways is also a direction for future work.

## Conflict of interests

The authors declare no conflict of interest.

## ACKNOWLEDGMENTS

The authors thank Baqiyatallah University of Medical Sciences who supported this study.

## ORCID

*Maryam Iman* http://orcid.org/0000-0002-2467-6262

## REFERENCES

[1] A. K. Chaturvedi, E. A. Engels, R. M. Pfeiffer, B. Y. Hernandez, W. Xiao, E. Kim, B. Jiang, M. T. Goodman, M. Sibug-Saber, W. Cozen, L. Liu, C. F. Lynch, N. Wentzensen, R. C. Jordan, S. Altekruse, W. F. Anderson, P. S. Rosenberg, M. L. Gillison, *J. Clin. Oncol.* **2011**, *29*, 4294.
[2] O. T. Selcuk, *EJENTAS* **2016**, *17*, 127.
[3] E. M. Garland, L. C. Parker, *Physician Assist. Clin.* **2016**, *1*, 465.
[4] G. Qian, Z. Hu, H. Xu, S. Müller, D. Wang, et al., *J. Oral Pathol. Med.* **2015**, *45*, 399–408.
[5] Y. Suh, I. Amelio, T. G. Urbano, M. Tavassoli, *Cell Death Dis.* **2015**, *5*, e1018.
[6] A. R. Kreimer, M. Johansson, T. Waterboer, R. Kaaks, J. Chang-Claude, D. Drogen, A. Tjønneland, K. Overvad, J. R. Quirós, C. A. González, M. J. Sánchez, N. Larrañaga, C. Navarro, A. Barricarte, R. C. Travis, K. T. Khaw, N. Wareham, A. Trichopoulou, P. Lagiou, D. Trichopoulos, P. H. M. Peeters, S. Panico, G. Masala, S. Grioni, R. Tumino, P. Vineis, H. B. Bueno-de-Mesquita, G. Laurell, G. Hallmans, J. Manjer, J. Ekström, G. Skeie, E. Lund, E. Weiderpass, P. Ferrari, G. Byrnes, I. Romieu, E. Riboli, A. Hildesheim, H. Boeing, M. Pawlita, P. Brennan, *J. Clin. Oncol.* **2013**, *31*, 2708.
[7] E. S. Prigge, M. Arbyn, M. von Knebel Doeberitz, M. Reuschenbach, *Int. J. Cancer* **2017**, *140*, 1186.
[8] J. P. Klussmann, J. J. Mooren, M. Lehnen, S. M. Claessen, M. Stenner, et al., *Clin. Cancer Res.* **2009**, *15*, 1779.
[9] Z. Deng, M. Hasegawa, A. Kiyuna, S. Matayoshi, T. Uehara, S. Agena, Y. Yamashita, K. Ogawa, H. Maeda, M. Suzuki, *Head Neck* **2013**, *35*, 800.
[10] S. Tomar, C. A. Graves, D. Altomare, S. Kowli, S. Kassler, S. Tomar, C. A. Graves, D. Altomare, S. Kowli, S. Kassler, N. Sutkowski, M. B. Gillespie, K. E. Creek, L. Pirisi, *Head Neck* **2016**, *38*, E694–E704.
[11] D. P. Zandberg, S. Liu, O. G. Goloubeva, R. Ord, S. E. Strome, M. Suntharalingam, R. Taylor, J. S.Wolf, A. Zimrin, J. E. Lubek, *J. Clin. Oncol.* **2014**, *32*, 6083.
[12] G. Campisi, V. Panzarella, M. Giuliani, C. Lajolo, O. Di Fede, G. Campisi, V. Panzarella, M. Giuliani, C.Lajolo, O. Di Fede, S. Falaschini, C. Di Liberto, C. Scully, L. Lo Muzio, *Int. J. Oncol.* **2007**, *30*, 813.
[13] D. L. Miller, J. W. Davis, K. H. Taylor, J. Johnson, Z. Shi, R. Williams, U. Atasoy, J. S. Lewis, M. S. Stack, *Am. J. Pathol.* **2015**, *185*, 679.
[14] S. Raza, N. Kornblum, V. P. Kancharla, M. A. Baig, A. B. Singh, M. Kalavar., *Recent Pat. Anticancer Drug Discov.* **2011**, *6*, 246.
[15] T. D. Wagner, G. Y. Yang, *Recent Pat. Anticancer Drug Discov.* **2008**, *3*, 76.
[16] W. Huber, A. Von Heydebreck, H. Sültmann, A. Poustka, M. Vingron, *Bioinformatics* **2002**, *18*, S96.
[17] P. Langfelder, S. Horvath, *BMC Bioinformatics* **2008**, *9*, 1.
[18] G. Altay, F. Emmert-Streib, *Biol. Direct* **2011**, *6*, 31.
[19] A. A. Margolin, I. Nemenman, K. Basso, C. Wiggins, G. Stolovitzky, R. Dalla Favera, A. Califano, *BMC Bioinformatics* **2006**, *7*, S7.
[20] C. Olsen, P. E. Meyer, G. Bontempi, *EURASIP J. Bioinform. Syst. Biol.* **2008**, *2009*, 308959.
[21] R. Bonneau, D. J. Reiss, P. Shannon, M. Facciotti, L. Hood, N. S. Baliga, V. Thorsson, *Genome Biol.* **2006**, *7*, R36.
[22] J. Yu, V. A. Smith, P. P. Wang, A. J. Hartemink, E. D. Jarvis, *Bioinformatics* **2004**, *20*, 3594.
[23] B. Zhang, S. Horvath, *Stat. Appl. Genet. Mol. Biol.* **2005**, *4*, 17.
[24] A. M. Yip, S. Horvath, *BMC Bioinformatics* **2007**, *8*, 22.
[25] G. Dennis, B. T. Sherman, D. A. Hosack, J. Yang, W. Gao, H. Lane, R. A. Lempicki, *Genome Biol.* **2003**, *4*, R60.
[26] P. Langfelder, S. Horvath, *BMC Syst. Biol.* **2007**, *1*, 54.

[27] G. K. Smyth, M. Ritchie, N. Thorne, J. Wettenhall. in *Bioinformatics and Computational Biology Solutions Using R and Bioconductor. Statistics for Biology and Health*, New York, Springer. **2005**.
[28] I. Martinez, J. Wang, K. F. Hobson, R. L. Ferris, S. A. Khan, *Eur. J. Cancer* **2007**, *43*, 415.
[29] R. J. Slebos, Y. Yi, K. Ely, J. Carter, A. Evjen, R. J. Slebos, Y. Yi, K. Ely, J. Carter, A. Evjen, X. Zhang, Y. Shyr, B. M. Murphy, A. J. Cmelak, B. B. Burkey, J. L. Netterville, S. Levy, W. G. Yarbrough, C. H. Chung, *Clin. Cancer Res.* **2006**, *12*, 701.
[30] L. Masterson, F. Sorgeloos, D. Winder, M. Lechner, A. Marker, S. Malhotra, H. Sudhoff, P. Jani, P. Goon, J. Sterling, *Cancer Sci.* **2015**, *106*, 1568.
[31] D. Pyeon, M. A. Newton, P. F. Lambert, J. A. Den Boon, S. Sengupta, D. Pyeon, M. A. Newton, P. F. Lambert, J. A. den Boon, S. Sengupta, C. J. Marsit, C. D. Woodworth, J. P. Connor, T. H. Haugen, E. M. Smith, K. T. Kelsey, L. P. Turek, P. Ahlquist, *Cancer Res.* **2007**, *67*, 4605.
[32] D. L. Miller, M. D. Puricelli, M. S. Stack, *Biochem. J.* **2012**, *443*, 339.
[33] S. Tomar, University of South Carolina, personal communication (**2013**).
[34] S. J. Korsmeyer, *Trends Genet.* **1995**, *11*, 101.
[35] R. Henriksen, E. Wilander, K. Oberg, *Br. J. Cancer* **1995**, *72*, 1324.
[36] S. E. Herman, A. L. Gordon, A. J. Wagner, N. A. Heerema, W. Zhao, J. M. Flynn, J. J. L. Andritsos, K. D. Puri, B. J. Lannutti, *Blood* **2010**, *116*, 2078.
[37] J. C. Pena, C. B. Thompson, W. Recant, E. E. Vokes, C. M. Rudin, *Cancer* **1999**, *85*, 164.
[38] R. Singh, N. Saini, *J. Cell Sci.* **2012**, *125*, 1568.
[39] D.-p. Zhao, X.-w. Ding, J.-p. Peng, Y.-x. Zheng, S.-z. Zhang, *J Zhejiang Univ. Sci. B* **2005**, *6*, 1163.
[40] C. Lajer, F. Nielsen, L. Friis-Hansen, B. Norrild, R. Borup, C. B. Lajer, F. C. Nielsen, L. Friis-Hansen, B. Norrild, R. Borup, E. Garnæs, M. Rossing, L. Specht, M. H. Therkildsen, B. Nauntofte, S. Dabelsteen, C. von Buchwald, *Br. J. Cancer* **2011**, *104*, 830.
[41] L. V. dos Santos, A. L. Carvalho, *Recent Pat. Anticancer Drug Discov.* **2011**, *6*, 45.
[42] M. Huang, W. A. Weiss, *Cold Spring Harb. Perspect. Med.* **2013**, *3*, a014415.
[43] S.-D. Hsu, F.-M. Lin, W.-Y. Wu, C. Liang, W.-C. Huang, S. D. Hsu, F. M. Lin, W. Y. Wu, C. Liang, W. C. Huang, W. L. Chan, W. T. Tsai, G. Z. Chen, C. J. Lee, C. M. Chiu, C. H. Chien, M. C. Wu, C. Y. Huang, A. P. Tsou, H. D. Huang, *Nucleic Acids Res.* **2010**, *39*, D163.
[44] A. Santoro, G. Pannone, R. Ninivaggi, M. Petruzzi, A. Santarelli, G. M. Russo, S. Lepore, M. Pietrafesa, I. Laurenzana, R. Leonardi, P. Bucci, M. I. Natalicchio, A. Lucchese, S. Papagerakis, P. Bufo, *Infect. Agents Cancer* **2015**, *10*, 46.
[45] Y. Zhu, H. Li, W. Guo, K. Drukker, L. Lan, Y. Zhu, H. Li, W. Guo, K. Drukker, L. Lan, M. L. Giger, Y. Ji, *Sci. Rep.* **2015**, *5*, 17787.
[46] N. N. Danial, P. Rothman, *Oncogene* **2000**, *19*, 2523.
[47] P. Dutta, W. X. Li, *eLS* **2013**.
[48] J. J. O'shea, S. M. Holland, L. M. Staudt, *N. Engl. J. Med.* **2013**, *368*, 161.
[49] S. Thomas, J. Snowden, M. Zeidler, S. Danson, *Br. J. Cancer* **2015**, *113*, 365.
[50] D. F. Calvisi, S. Ladu, A. Gorden, M. Farina, E. A. Conner, J.–. S. Lee, V. M. Factor, S. S. Thorgeirsson, *Gastroenterology* **2006**, *130*, 1117.
[51] P. Khanna, P. J. Chua, B. H. Bay, G. H. Baeg, *Int. J. Oncol.* **2015**, *47*, 1617.
[52] J. M. Schieler, J. O. Henderson, *JSR* **2016**, *5*, 11.
[53] H. Xiong, Z.-G. Zhang, X.-Q. Tian, D.-F. Sun, Q.-C. Liang, Y. J. Zhang, R. Lu, Y. X. Chen, J. Y. Fang, *Neoplasia* **2008**, *10*, 287.
[54] V. Sriuranpong, J. I. Park, P. Amornphimoltham, V. Patel, B. D. Nelkin, J. S. Gutkind, *Cancer Res.* **2003**, *63*, 2948.
[55] G. Bendas, L. Borsig, *Int. J. Cell Biol.* **2012**, *2012*, 676731.
[56] J. A. King, S. F. Ofori-Acquah, T. Stevens, A.-B. Al-Mehdi, O. Fodstad, W. G. Jiang. *Breast Cancer Res.* **2004**, *6*, R478.
[57] V. Kulasingam, Y. Zheng, A. Soosaipillai, A. E. Leon, M. Gion, E. P. Diamandis, *Int. J. Cancer* **2009**, *125*, 9.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Iman M, Narimani Z, Hamraz I, Ansari E. Network based identification of different mechanisms underlying pathogenesis of human papilloma virus-active and human papilloma virus-negative oropharyngeal squamous cell carcinoma. *J Chin Chem Soc*. 2018;1–10. https://doi.org/10.1002/jccs.201800072