DNA Methylome Distinguishes Oropharyngeal and Oral Cancer from Oral Lesions and Healthy Oral Mucosa

Milutin Gašperov Nina (nmilutin@irb.hr)
Institut Ruder Boskovic
Ivan Sabol
Institut Ruder Boskovic
Ksenija Božinović
Institut Ruder Boskovic
Emil Dediol
Klinicka Bolnica Dubrava
Marinka Mravak-Stipetić
Sveuciliste u Zagrebu Stomatoloski fakultet
Danilo Licastro
CBM Scrl
Simeone Dal Monego
CBM Scrl
Magdalena Grce (grce@irb.hr)
Institut Ruder Boskovic https://orcid.org/0000-0001-6178-8418

Research

Keywords: DNA methylation, head and neck squamous cell carcinoma (HNSCC), potentially premalignant oral lesions, healthy oral mucosa, human papillomavirus (HPV)

Posted Date: April 3rd, 2020

DOI: https://doi.org/10.21203/rs.3.rs-17778/v1

Abstract

Background: There is a strong need to find new, good biomarkers of head and neck squamous cell carcinoma (HNSCC), because of the prognoses and high mortality of patients. The aim of this study was to identify the potential biomarkers in HNSCC that have differences in their DNA methylome and potentially premalignant oral lesions, in comparison to healthy oral mucosa.

Methods: In this study 32 oral samples were tested: 9 healthy oral mucosae, 13 HNSCC and 10 potentially premalignant oral lesions for DNA methylation by Infinium MethylationEPIC BeadChip.

Results: Our findings showed a panel of genes significantly hypermethylated in their promoters or specific sites in HNSCC samples in comparison to healthy oral samples, which are mainly oncogenes, receptor and transcription factor genes, or genes included in cell cycle, transformation, apoptosis and autophagy. A group of hypomethylated genes in HNSCC, in comparison to healthy oral mucosa, is mainly involved in the host immune response and transcriptional regulation. The results also showed significant differences in gene methylation between HNSCC and potentially premalignant oral lesions, as well as differently methylated genes that discriminate between oral lesions and healthy mucosa.

Conclusions: The given methylation panels point to the possibly early diagnostics of HNSCC as well as potentially premalignant oral lesions and its possible implication in clinical practice.

Introduction

Head and neck squamous cell carcinoma (HNSCC), which encompasses tumors of the oral and nasal cavities, paranasal sinuses, pharynx, and larynx, is the 6th most frequent malignancy in the world, with over 650,000 new cases diagnosed each year (1). Although the overall survival of patients with oral cancer has improved during the past 20 years, it has only improved marginally; mainly due to the advanced clinical stage at diagnosis and the high rates of treatment failure associated with this advanced disease (2).

The most important risk factors identified so far for HNSCC are excessive tobacco (3,4) and alcohol consumption (5–7) together with high-risk types of human papillomavirus (HPV) (8–10). Although the global incidence of HNSCC is declining, the incidence of HPV related HNSCC, especially oropharyngeal and oral squamous cell carcinoma (OPSCC, OSCC, respectively) is rapidly increasing over the last few decades (11). Recent findings emphasize the importance of epigenetic changes, such as DNA methylation and alterations including micro RNAs (miRNA), in HNSCC progression and implicate the very role of tobacco and alcohol (12) as well as HPV (13) in those changes.

The HNSCCs are one of the cancer types with the worst prognosis and with high mortality of patients, hence, there is a strong need to find new biomarkers of this disease (14). The most appropriate biomarkers would be those pointing out changes on cellular level before carcinoma can be detected or even before carcinoma occurrence. The epigenetic biomarkers, such as methylated genes could efficiently

point to changes before cancers can be clinically detected and help us to better understand tumorigenesis and hopefully improve cancer treatment and prevention (15,16). Some of these potential biomarkers could also differentiate between the groups of potentially premalignant oral lesions that show possible premalignant transformation, such as oral lichen planus (OLP) and oral lichenoid lesions (OLL), whose treatment is different from each other despite their high clinical and histopathological similarities (17–19). Namely, OLP is a chronic immunological mucocutaneous disorder of unknown etiology, while OLL is usually of known etiology, being a lichenoid contact stomatitis (20).

DNA methylation is the most studied epigenetic change in human diseases, especially cancer because it is apparently stable under most storage conditions, even as histological preparations (21). Altered DNA methylation is one of the possible factors associated even with the HNSCC development. The focus of this study was to explore DNA methylation changes that are significantly deregulated in HNSCC samples, particularly OPSCC and OSCC, and potentially premalignant oral lesions (such as OLP and OLL) in comparison to healthy oral mucosa. Identifying DNA methylome differences by means of the Infinium MethylationEPIC BeadChip array (Illumina) on the level of methylated genes, gene promoters and individual CpG sites between HNSCC, potentially premalignant oral lesions, and healthy oral mucosa enables us to suggest novel potential biomarkers for the identification of HNSCC and potentially premalignant oral lesions.

Materials And Methods

Study group

Healthy oral mucosa samples were collected from healthy subjects in the School of Dental Medicine, University of Zagreb, Croatia, during regular process of teeth extraction from 2010 to 2017. Oral samples of potentially premalignant oral lesions, OLP and OLL, were taken cytologically in the School of Dental Medicine, University of Zagreb, Croatia, from 2008 to 2016. HNSCC samples were collected in the Clinical Hospital Dubrava, Zagreb, Croatia, from 2014 to 2018. The study group comprised 32 oral samples: 9 healthy oral mucosa, 10 potentially premalignant oral lesions (8 OLP and 2 OLL), and 13 HNSCC (6 oropharyngeal cancers and 7 oral cancers). The median age among patients with HNSCC was 57 years, while amongst patients with potentially premalignant oral lesions and control healthy mucosa was slightly lower, 43 and 53 years, respectively. There were 10 men and 3 women with HNSCC, 4 men and 6 women with potentially premalignant oral lesions, and 4 men and 5 women with healthy oral mucosa and without drinking and smoking history. Fresh samples were collected with the cytobrush, stored in appropriate buffers for further analysis, HPV testing and DNA methylation analysis.

DNA preparation

The extracted DNA from oral specimens was processed without initial knowledge of patients' data. DNA was isolated using the BioRobot EZ1 (Qiagen) system according to the manufacturer's instructions. After DNA extraction, the purified DNA was dissolved in 50–100 µl of tri-distillate sterile water and stored at −20 °C until further analysis. The quality and integrity of the samples were evaluated on a

NanoPhotometer (Implen), and samples with the ratio A260/280 between 1.7–1.9 were included in the study (22).

HPV detection and typing

HPV testing are previously described (23,24). Briefly, three sets of consensus primers for HPV detection were used: PGMY09/PGMY11, L1C1/L1C2-1/L1C2-2 and GP5+/GP6+. The quality of the isolated DNA was confirmed by amplification of the β-globin gene using PC04/GH20 primers in a multiplex PCR with PGMY primers. Type-specific (TS) primers for HPV types 6/11, 16, 18, 31, 33, 45, 52, and 58 were used for HPV typing according to Milutin-Gasperov et al. (23). Aliquots of each PCR product (10 µl) were analyzed by a 2% agarose gel electrophoresis and stained with Midori Green Advance dye. The amplified products were visualized by UV irradiation of the gels using the UVItec Cambridge (Alliance 4.7) imaging system. HPV positive samples that were not positive for TS-PCR but positive for consensus primer amplification were defined as undetermined HPV type.

Methylation array

Infinium MethylationEPIC BeadChip array (Illumina), which integrates a total of 863,904 CpG loci, together with 2932 non-CpG loci and 59 single nucleotide polymorphisms (SNPs), superseded the HM450 array, while still containing more than 90% of the original HM450 probes. Additional probes included in the new version of the array greatly increased the power of this microarray to study enhancer/regulatory regions (25). Briefly, approximately 1–2 µg of DNA from cancer and oral samples were modified with sodium bisulfite using the EZ DNA Methylation Kit (Zymo Research) and then purified according to the manufacturer's instructions. After bisulfite treatment, 180–200 ng DNA was subjected to whole genome amplification (WGA) and enzymatic digestion with the Infinium MethylationEPIC BeadChip kit reagents. The hybridization of the samples on the BeadChips and washing procedures followed the standard manufacturer's protocol. The iScan System (Illumina) was used to read the BeadChips.

Data processing and statistical analysis

Raw data obtained by iScan readout was imported to and analyzed within R using ChAMP (version 2.9.10) (26) and RnBeads packages (version 1.10.7) (27). Briefly, data was imported to ChAMP pre-processed, and normalized with the PBC method (28). Subsequently, Combat method (29) within ChAMP was used to adjust for batch effects. Resulting normalized and batch-corrected rnb.set was imported to the RnBeads package for subsequent exploratory and differential methylation analysis and customized visualization. RnBeads uses limma method (30) for differential methylation assessment between pairs of groups and herein the calculations were performed while adjusting for patient age and gender as covariates. In addition to analyzing differential methylation between groups on individual CpG site level, all analyses were performed on predefined gene and promoter levels by selecting appropriate region.types options within the RnBeads package.

Differentially methylated gene promoters and individual CpG sites

The resulting differentially methylated promoter or CpG site lists were further filtered by selecting only promoters or sites with false discovery rate (FDR) adjusted P-values ≤ 0.05 and a more stringent differential methylation value of ≥|0.44| (between cancer and control healthy tissues, cancer and potentially premalignant oral lesions, and potentially premalignant oral lesions and control healthy tissues). The filtered tables contain the information on chromosome locations, and relation to any nearby CpG islands. Promoter level data additionally contains information about number of included CpG sites and average GC content for the region.

Pyrosequencing validation of methylation

Pyrosequencing assays were developed for the following genes: SPRR3, FBXO2, TRDC, and LAIR2. The PCR and sequencing primers were designed using the PyroMark Assay Design software, version 2.0.1.15 (Qiagen) to assess particular CpG sites of depicted genes from the Infinium MethylationEPIC BeadChip array analysis. All primers were purchased from Macrogen (Macrogen Korea). The primer sequences, amplicon sizes, and the optimal annealing temperatures are indicated in Table A1. The analysis was performed on 4 control healthy and 4 HNSCC tissues (2 HPV positive and 2 HPV negative), which were already tested by the Infinium MethylationEPIC BeadChip array. Briefly, approximately 500 ng of extracted DNA was used for the bisulfite treatment performed with the EZ DNA Methylation Kit, according to the instructions by the manufacturer, and eluted in 20 µl elution buffer (Zymo Research). The PCR reactions were performed according to the PyroMark PCR protocol (Qiagen) in a total volume of 30 µl. Briefly, 0.10 µmol/L of each primer, 1.5 mM $MgCl_2$, PyroMark PCR Master mix (Qiagen), Coral Load (Qiagen) and 50 ng of bisulfite treated template DNA were added to PCR reaction and performed in a thermocycler (Veriti, 96 Well Thermal Cycler, Applied Biosystems). The program was as follows: initial denaturation of 1 min at 95 °C, followed by 45 cycles of 30 s denaturation at 95 °C, specific annealing temperature for each primer pair (Table A1) for 30 s and extension for 30 s at 72 °C with the final extension for 10 min at 72 °C. Pyrosequencing was performed using a PyroMark Q24 Reagent Kit and a PyroMark Q24 system (Qiagen) as described previously by Mikeska et al. (31). The nucleotide addition order was optimized by the PyroMark Q24 Software (Qiagen) and the results were automatically analyzed using the same software. Percentage of methylation for each CpG island between the two sample groups (cancer vs. controls) was compared and P-values determined using t-test.

Gene set enrichment analysis

The list of differentially methylated gene promotor regions was assessed to determine whether the affected genes are enriched for specific sets of functions or pathways. However, for analysis were selected only those regions with assigned RefGene names indicating nearby or overlapping genes. To make the analysis more stringent, only promotors with at least 2 CpG sites were included. The analysis was done using WebGestalt functional enrichment analysis web tool (32). Methylation data was explored with all 3 different analysis approaches available, over-representation enrichment analysis (ORA), gene set enrichment analysis (GSEA) and network topology-based analysis (NTA). For ORA and GSEA analysis, the gene ontology – biological process (no-redundant) and KEGG pathway databases were chosen. For ORA, reference gene set was set to the whole genome, since many differentially methylated regions were

related to miRNA and other non-coding sequences. For GSEA, gene promotors were ranked according to the Log2 of the mean difference and this data supplied in addition to the gene symbol. For NTA analysis, the data from TCGA-HNSC project was used to create the network using network retrieval and prioritization approach. Differentially methylated genes were used to retrieve the relevant subnetworks. Enrichment of gene ontology biological process categories was then assessed for the retrieved subnetworks. Unless indicated otherwise, default parameters were used.

External validation of differentially methylated gene promoters in HNSCC vs. control tissue

Our gene promoters' methylation findings were compared to TCGA Illumina HISeq RNAseq data of TCGA-HNSC project through Wanderer (http://maplab.imppc.org/wanderer/), an interactive viewer to explore DNA methylation and gene expression data in human cancer (33). The RNAseq estimation of expression for top fifteen hypermethylated and top fifteen hypomethylated gene promoters in our results were visualized in Wanderer. The TCGA dataset included 497 tumor and 43 normal tissue samples. In the cases where the direction of expression change did not correspond with our methylation change, we visualized complementary TCGA in Illumina 450K DNA methylation array results for the same genes.

Ethics approval and consent to participate

This study was approved by the Ethical Board of the Ruđer Bošković Institute, the Ethical Board of the Clinical Hospital Dubrava, and the School of Dental Medicine, University of Zagreb. The study is in line with the Helsinki Declaration of 1975 (DoH/Oct2008). An informed consent to participate in the study was obtained from each participant.

Results

HPV status

HPV DNA was found in 9 of 32 samples, of which 5 cancer samples, 3 potentially premalignant oral lesions samples and 1 control sample (healthy mucosa) (Figure A1). HPV 16 was found in all five HPV positive cancer samples. HPV 58 was found in one OLP sample positive, while undetermined HPV types were found in one OLP, one OLL and one control sample.

The overall DNA methylation findings

After pre-processing, normalization and batch correction of Infinium MethylationEPIC BeadChip data in ChAMP, 679851 probes were retained for analysis by RnBeads package. The inclusion criteria was the value of the mean difference across all sites in a region (mean.mean.diff), the highest and the lowest values from the Illumina assay data. For every set of data, the list of first 15 genes, the best significantly differentiated between groups of samples in methylation, i.e. hyper- or hypomethylated, is presented. The

whole list of genes, besides a large number of defined genes includes not annotated (NA) genes, pseudogenes and RNA genes (could be provided upon request).

Within the exploratory analysis, the principal component analysis (PCA) showed that all samples clustered within 3 distinct clusters: cancer, lesions and controls (Fig. 1). One healthy oral mucosa sample of 32 oral samples analyzed by Infinium MethylationEPIC BeadChip, was automatically excluded during filtering and normalization steps (Figs. 1, 2, 3 and S1).

Unsupervised hierarchical clustering analysis based on all investigated methylation sites across the genome and all methylation values showed good clustering of cancer samples on one side, however control samples and potentially premalignant oral lesions were clustered together on the other side (Fig. 2A). The situation was similar when visualizing only the top 1,000 most variable positions (Fig. 2B). As expected, global cancer hypomethylation can be seen in both figures as more CpG sites exhibit lower methylation values in cancer cluster. The heatmap representation of samples clustering for the top fifteen hypermethylated and the top fifteen hypomethylated CpG gene sites in cancer tissue compared to control healthy tissue also showed good clustering of cancer samples on one side and control samples on the other, while most of the potentially premalignant oral lesions were clustered between those two groups (Fig. 3).

Differentially methylated gene promoters in HNSCC tissue compared to control tissue

The top fifteen genes significantly hypermethylated in their promoters in cancer tissues in comparison to healthy tissues are: GPRC5D, TMPRSS11B, PIAS2, ARG1, SRPK2, AADACL2, RGPD4, SPRR3, DEGS1, TXNDC8, SH3TC1, ZPLD1, FBXO2, ATG16L1 and GRHL1 (Table 1). The FDR adjusted P-value was < 0.05 and difference in average methylation value between 0.78 and 0.44. The hypermethylated genes are mostly involved in different cellular enzymatic reactions and autophagy.

The top fifteen genes that were significantly hypomethylated in their promoters in cancer tissues in comparison to healthy tissues are: TRBC2, DGAT2, ALG1L, PDE4D, TRDC, DNAJC6, IGKV3-20, TMEM150B, LAIR2, UBQLN3, ANKFN1, MS4A1, CCT8L2, SPOCK1 and IGHV4-39 (Table 1). The FDR adjusted P-value was < 0.05 and difference in average methylation value between − 0.80 and − 0.61. The hypomethylated genes are mainly involved in immune response.

Differentially methylated CpG sites in HNSCC tissue compared to control tissue

From the complete list of differentially methylated CpG sites, only those falling within or near defined genes were selected, while all other sites for which biologic relevance couldn’t be evaluated were excluded (Fig. 3). Thus, the top fifteen significantly hypermethylated sites in cancer tissues in comparison to healthy tissues are: LMBR1L, CDH1, EIF6, C16orf70, ETNK2, C11orf73, ADARB2, GAB1, ITPR3, WDR61, PGAP2, DDX10, DGKH, RAB40C and BEAN1 genes (Table 2). The FDR adjusted P-value was < 0.05 and

the difference between mean methylation values across sites between 0.93 and 0.89. The hypermethylated genes are mostly involved in translation processes and cellular growth, transformation and proliferation.

The top fifteen genes significantly hypomethylated on different sites across the genome (5'UTR, 3'UTR, TSS1500, TSS200, 1st exon, exon body) in cancer tissues in comparison to healthy tissues are: ATXN1, PPP2R2C, CCR6, RAB37, DUSP27, ZNF521, SLC6A17, SPIN1, CXCR1, SPTBN1, NBAS, NRG3, COL5A1, CDX1 and BATF3 (Table 2). The FDR adjusted P-value was < 0.05 and the difference between mean methylation values across sites between − 0.96 and − 0.89. The hypomethylated genes are mostly involved in transcriptional and immune regulation.

Aberrant methylation in potentially premalignant oral lesions

The aberrant methylation in gene promoters and CpG sites within defined genes in HNSCC tissue compared to potentially premalignant oral lesions are shown in Tables 3 and 4, respectively, while the aberrant methylation in gene promoters and CpG sites in oral lesions, OLP and OLL, compared to healthy tissue are shown in supplement material (Tables A2 and A3, respectively).

The top fifteen genes significantly hypermethylated in their promoters that could distinguish HNSCC from potentially premalignant oral lesions are: RAD51B, BARX2, SLC5A10/FAM83G, NINL, NSMCE2, PGAP2, INO80C, IL34, ZNF516, GFOD2, PARD3, MCEE, POLM, ASPG and TBC1D2 (Table 3). The top fifteen genes significantly hypomethylated in their promoters in HNSCC compared to potentially premalignant oral lesions are: ART4, EPB41L3, ESRRG, ENPP1, GNG7, PAPSS2, NGEF, HIPK4, GPR158, GSG1L, SMPD3, GDF2, RERE, CDH13 and HS3ST4 (Table 3).

The top fifteen genes significantly hypermethylated on different sites across the genome that could distinguish HNSCC from potentially premalignant oral lesions are: EIF6, KANSL1, DDX10, AP2A1, RAB40C, GAB1, ERGIC1, SNX14, PIGU, ARAP1, LMTK2, BEAN1, AP1S3, CDH1, and RYBP (Table 4). The top fifteen genes significantly hypomethylated on different sites across the genome in HNSCC compared to potentially premalignant oral lesions are: FAM69A, ATP6V0A1, LBP, WDR25, SH3RF3, NINJ2, RAB37, CXCR1, SPTBN, RHOH, GRIK5, KLRD1, TENM2, FAM69A, and ITK (Table 4).

The top fifteen genes significantly hypermethylated (SLC5A10, TBC1D2, SH3BP5L, VANGL1, DLEC1, TGOLN2, CTBP2, PPP1CB, VPS52, MEPCE, HDAC4, ARAP1, TCF20, NDUFS7, and GATAD2A) and the top fifteen genes significantly hypomethylated in their promoters (ART4, ENPP1, GNG7, PKD1L3, PLXNC1, CAMK2B, CACNA1S, SCGB1D1, VPS13D, DLGAP4, LRP1B, COL2A1, SLC24A3, TBC1D8 and ABCC8) that could distinguish potentially premalignant oral lesions from healthy oral mucosa are shown in Table A2. The top fifteen genes significantly hypermethylated (GRIP1, MTMR10, RBM47, MPHOSPH9, FOXK1, SNX3, CIT, ZBTB38, DRD3, SPPL3, ZNF407, ADAMTSL1, GNAT3, L3MBTL3 and EEPD1) and the top fifteen genes significantly hypomethylated (PHACTR1, MARCH8, PPP1R1B, HDAC4, IL22RA2, CAMKK2, INPP5D, CSGALNACT1, GTDC1, IGSF3, HELZ, DEFA4, AK5, LHFPL2 and STK10) on different sites across

the genome in potentially premalignant oral lesions in comparison to healthy oral mucosa are shown in Table A3.

Pyrosequencing validation panel

Pyrosequencing was performed on a subset of samples tested by Infinium MethylationEPIC BeadChip array (Illumina) for 4 gene promoters, namely SPRR3, FBXO2 (hypermethylated in HNSCC tissue vs. control healthy tissue; Table 1), TRDC and LAIR2 (hypomethylated in HNSCC tissue vs. control healthy tissue; Table 1) tested on 4 cancer samples and 4 control samples, each. The selection criteria were the role of these genes in biological processes as well as the findings in previous studies. SPRR3 (Small Proline Rich Protein 3) and FBXO2 (F-Box Protein 2) being largely investigated; SPRR3 is involved in cornification, epidermis development, squamous cell differentiation and peptide cross linking (34), while FBXO2 is involved in negative regulation of cell proliferation, cellular protein modification and protein ubiquitination (35). In addition, the hypomethylated genes in cancer are mostly involved in immune response; TRDC (T Cell Receptor Delta Constant) being involved in recognizing foreign antigens, which have been processed as small peptides and bound to major histocompatibility complex (MHC) molecules at the surface of antigen presenting cells (36). LAIR-2 (Leukocyte Associated Immunoglobulin Like Receptor 2) related pathways belongs to innate immune system, and class I MHC mediated antigen processing, and presentation and immunoregulatory interactions (36).

For pyrosequencing validation, six amplifying PCR reactions (SPRR3-1, SPRR3-2, TRDC-1, TRDC-2, LAIR2-1, LAIR2-2, FBXO2-1) with 6 sequencing primers have been performed to cover 4 CpG sites for SPRR3, 4 CpGs for FBXO2, 2 CpGs for TRDC and 5 CpGs for LAIR2 gene. The overall pyrosequencing data for tested CpGs were in agreement with the methylation array data (Figure A2). However, statistical significance was only reached between HNSCC and the control samples in CpG1 and CpG3 of SPRR3 gene ($P = 0.01$ in both cases) and CpG1 of FBXO2 gene ($P = 0.01$).

Gene set enrichment data

Differentially methylated gene promotor regions were analyzed using WebGestalt functional enrichment analysis web tool determine whether the affected genes are enriched for specific sets of functions or pathways. Methylation data was explored with three different analysis approaches available, over-representation enrichment analysis (ORA), gene set enrichment analysis (GSEA) and network topology-based analysis (NTA). The data are presented in supplement material (Figures A7, A8, A9 and A10). Enriched GO PB terms associated with differentially methylated gene promoters on network topology is also shown in supplement material (Figure A11 and Table A5) where the top 20 enriched GO categories are presented.

External database validation

Our gene promoters' methylation findings were compared to TCGA Illumina HISeq RNAseq data of TCGA-HNSC project. The RNAseq estimation of expression for top fifteen hypermethylated and top fifteen hypomethylated gene promoters in our results were visualized in Wanderer, an interactive viewer. The

TCGA dataset included 497 tumor and 43 normal tissue samples. There was a good agreement between our gene promoters' methylation data and the gene expression data (Table A4, Figures A3, A4, A5 and A6). Out of the total of top fifteen hypermethylated gene promoters in our study, ten were found to be either under-expressed or hypermethylated in TCGA cancer cases, as expected, while one had no measurable expression and only a single CpG site in Illumina 450K DNA methylation array. From the top fifteen hypomethylated gene promoters in our study, twelve were also found to be either over-expressed or hypomethylated in TCGA data, with the remaining three lacking annotated data or probes in Illumina 450K DNA methylation array.

Discussion

The aim of this study was to investigate DNA methylome in HNSCC and potentially premalignant oral lesions, as well as in healthy oral tissue, and to identify the best genes that are differentially methylated in gene promoters or specific sites among those groups of samples. We found that components of different cellular pathways are differently methylated in HNSCC in comparison to healthy oral tissue as well as potentially premalignant oral lesions.

Surprisingly, we could not observe any grouping of samples in accordance with their HPV status, thus subsequent analysis focused only on the sample origin. The lack of HPV specific differences could possibly be explained by a limited number of HPV positive samples (9 of 32). Furthermore, HPV is known to be more associated with oropharyngeal tumors than oral cavity cancerogenesis (37). Another possible explanation is the particularity of the Croatian population where smoking and drinking are almost equally present in HPV positive and HPV negative OPSCC patients shown in our previous study (38). On the other hand, the study of Lechner et al. with Infinium HumanMethylation450 BeadChips (Illumina) showed unsupervised clustering over the methylation variable positions of samples in accordance with HPV status. Nevertheless, they showed that HPV positive tumors are heterogeneous what led to the identification of a candidate CpG island methylator phenotype in a sub-group of HPV positive tumors (39).

Herein, the top fifteen genes with significant promoter hypermethylation in cancer tissues in comparison to control healthy tissues, GPRC5D, TMPRSS11B, PIAS2, ARG1, SRPK2, AADACL2, RGPD4, SPRR3, DEGS1, TXNDC8, SH3TC1, ZPLD1, FBXO2, ATG16L1, and GRHL1 are mostly involved in different cellular enzymatic reactions and in autophagy (Table 1). For example, the expression of SPRR3 (Small Proline Rich Protein 3) was found associated with tumor cell proliferation and invasion in glioblastoma multiforme. Liu et al. (2013) found, contrary to our findings, that SPRR3 hypomethylation was associated with the clinical outcome in glioblastoma multiforme patients (40). In anatomically more similar context, SPRR3 was frequently downregulated in OPSCC where it probably suppresses tumorigenicity (41). In our study, we selected the promoter of the SPRR3 and FBXO2 genes for validation by pyrosequencing and found that both methods agree on the direction of methylation deregulation, which is hypermethylation of the gene promotor.

The top fifteen genes in cancer tissues that were found in this study to be significantly hypomethylated in their promoters in comparison to control healthy tissues (TRBC2, DGAT2, ALG1L, PDE4D, TRDC, DNAJC6, IGKV3-20, TMEM150B, LAIR2, UBQLN3, ANKFN1, MS4A1, CCT8L2, SPOCK1, and IGHV4-39) are mainly involved in immune response, i.e. IGHV4-39 (antigen recognition gene), IGKV3-20 (immunoglobulin receptor binding gene), LAIR2 (innate immune response gene), MS4A1 (differentiation of B cells gene), TRBC2 and TRDC (both T cell receptor genes). Indeed, the HNSCC are known for their immune-suppressive character allowing tumor evasion and escape from immune surveillance, which probably can be associated with the methylation of immune-response related genes (42). Here again, from the list of genes with hypomethylated promoters we selected LAIR2 and TRDC for validation and, as expected, both gave comparable results on pyrosequencing.

The top fifteen significantly hypermethylated genes, LMBR1L, CDH1, EIF6, C16orf70, ETNK2, C11orf73, ADARB2, GAB1, ITPR3, WDR61, PGAP2, DDX10, DGKH, RAB40C, and BEAN1 on different gene sites (mostly in 5'UTR and body) in cancer tissues in comparison to control healthy oral tissues are mostly involved in translational processes and cellular growth, along with transformation and proliferation. Among them, CDH1, ETNK2, ADARB2 and RAB40C are found to be aberrantly methylated in different cancers (43–46). For instance, altered methylation levels of CDH1 (Cadherin 1), whose loss contributes to cancer progression by increasing proliferation, invasion, and/or metastasis are recorded in oral cavity (43), oral (47) and in cervical cancer (48). The study of Strzelczyk et al. (43) reported a significantly higher methylation level of CDH1 in tumor tissues compared to surgical margins (57% vs. 25% $p < 0.001$) in patients with oral cavity cancer. The meta-analysis of the gene promoter hypermethylation in oral cancer, that included 29 studies of which 13 were about CDH1 methylation, showed significant correlation of CDH1 hypermethylation with oral cancer risk (47). Moreover, in the meta-analysis of Liu et al. (48) on patients with cervical carcinoma, CDH1 promoter methylation was significantly higher in cancer than in cervical intraepithelial neoplasia lesions and healthy cervical tissues.

The first fifteen genes that were significantly hypomethylated on different sites across the genome in cancer tissues in comparison to control healthy tissues, ATXN1, PPP2R2C, CCR6, RAB37, DUSP27, ZNF521, SLC6A17, SPIN1, CXCR1, SPTBN1, NBAS, NRG3, COL5A1, CDX1, and BATF3 are mostly involved in transcriptional and immune regulation. Among this group of genes, aberrantly methylated in other human cancers were CCR6 in oral cancer (49) and chronic lymphocytic leukemia (50), RAB37 in lung cancer (51), ZNF521 in breast cancer (52), and CDX1 in gastric cancer (53), esophageal SCC (54) and in colon cancer (55). The genes involved in immune regulation could belong to the tumor-infiltrating immune cells or tumor-infiltrating lymphocytes, which are often associated with better clinical outcomes. Thus, the aberrantly methylated gene CCR6 (C-C Motif Chemokine Receptor 6), which regulates the migration and recruitment of dendritic and T cells during inflammatory and immunological responses, was also found in human OSCC (49). Lee et al. (49) concluded that hypomethylation of this gene may play an important role in the recruitment or retention of CCR6 + Treg cells into OSCC inflammatory microenvironment at the early stage of tumor progression. In addition, genome-wide DNA methylation analysis of chronic lymphocytic leukemia patients in comparison to healthy donors identified differently methylated CCR6 gene, among other immune regulatory genes (50). In addition, in their study, Kim et al.

presented that the majority of hypomethylated gene sets identified across multiple cancer (breast, lung cancer, colorectal, myeloma, glioblastoma, ovarian, kidney and stomach cancer) studies were immune-related, suggesting DNA methylation-driven cancer cell invasion and tumorigenesis across various types of cancer (56).

The external validation of our top thirty differentially methylated gene promoters in HNSCC vs. control tissue with gene expression data in human cancer through Wanderer, an interactive viewer, gave a very good agreement. In summary, the majority of hypermethylated gene promoters in HNSCC in our study (10 of 15) were found to be either under-expressed or hypermethylated in TCGA cancer cases. In addition, from the top 15 hypomethlylated gene promoters in our study, 12 were also found to be either over-expressed or hypomethylated in TCGA data.

Of particular interest in HNSCC diagnostic, clinical prognosis and/or risk assessment could be the methylation of CDH1, which was also previously described as possible biomarker for the early detection and treatment of HNSCC (43,47,57,58). Herein, we found the CDH1 gene to be significantly hypermethylated on specific sites in the genome (body) on high second place in cancer tissues in comparison to control tissues. The same gene (CDH1) is also among the top fifteen genes that are significantly hypermethylated on different sites across the genome in cancer tissues compared to lesions. The CDH1 gene encodes E-cadherin, a classical cadherin of the cadherin superfamily that is involved in mechanisms regulating cell-cell adhesions, mobility and proliferation of epithelial cells. It is recognized as a tumor suppressor gene; loss of function of this gene is thought to contribute to cancer progression by increasing proliferation, invasion, and/or metastasis (36). Hence, we showed herein that hypermethylation on specific CpGs within the CDH1 gene could be a good biomarker of HNC and a possible option to distinguish HNSCC from potentially premalignant oral lesions and from healthy oral mucosa as well.

Another two gene that are present in the top fifteen most significantly hypermethylated in gene promoter regions in cancer tissues compared to lesions, and in lesions compared to control healthy tissues are the SLC5A10 (Solute Carrier Family 5 Member 10) and the TBC1D2 (TBC1 Domain Family Member 2) gene. The SLC5A10 gene is a member of the sodium/glucose transporter family, while the TBC1D2 gene acts as GTPase-activating protein for RAB7A, involved in cadherin degradation and cell-cell adhesion. Notably, the two out of three genes, whose hypermethylation may be of particular importance in HNSCC diagnostic, CDH1 and TBC1D2, are involved in cadherin regulation of cell-cell adhesion. Suppression of cadherins in HNSCC leads to cells escape from contact-dependent growth, which develop migratory phenotype with low differentiation stage, suggesting that cadherins contribute to the transformation steps (59). The two genes from the group, SLC5A10 and TBC1D2, could also be considered as possible good methylation biomarkers to distinguish oral potentially premalignant lesions from healthy oral tissue.

Unexpectedly, the overlap of significant findings on CpG site and gene promoter levels in the whole study was non-existent, probably because most of the top-rated promoters included only one or rarely few sites

in the analysis. Further, there is no evidence in the literature on this issue to conform or refute these observations.

Using WebGestalt functional enrichment analysis web tool we assessed gene enrichment for specific sets of functions or pathways and networking. Indeed, over-representation enrichment analysis (ORA) of gene ontology non-redundant biological processes for differentially methylated gene promoters presented implication of mostly immune response and cellular defense response pathways, as well as cell-cell adhesion.

The current study is the first to implement Infinium MethylationEPIC BeadChip whole genome methylation assay on a well-defined set of clinical samples encompassing the whole possible spectrum from healthy tissue to cancer. To our knowledge this is the first such study focused on HNSCC, oral lesions and healthy tissue together. In addition, the power of the study relies on prospectively collected fresh samples with minimum delays between sample collection and processing. However, for that reason the limitation of the study might be the possibility that infiltrating immune cells could be present in tumor tissues. Another strength of the study was the simultaneous microarray testing in the same analysis of different tissues, cancer, oral lesions and healthy tissue. On the other hand, the study was limited by anatomical differences in sample material, namely both healthy and potentially premalignant oral lesions samples were mostly derived from oral cavity, where potentially premalignant oral lesions usually originate, while cancer samples included both oral, and oropharyngeal cancer. Another possible limitation was the age of participants as cancer usually develops later in life, while the average age of controls and patients with potentially premalignant oral lesions was lower (43 vs. 53 years). We attempted to adjust for this by including age as covariate. Overall, our study has demonstrated significant overlap with current knowledge, which together with successful validation of the data by pyrosequencing confirms the reliability of the underlying data and strengthens its results.

Conclusion

The presented methylation clustering shows that the potentially premalignant oral lesions (OLL and OLP) are more closely related to healthy mucosa than to the HNSCC although differences between groups exist. The identified panels of hypermethylated and hypomethylated genes, which differentiates HNSCC samples from oral potentially premalignant lesions and healthy mucosa could be clinically a useful tool for early cancer diagnosis and prognosis. Specific genes that could be considered as HNSCC DNA methylation biomarkers belong to the groups of receptor genes, transcription factors, genes involved in adhesion and transport reactions, as well as genes related to immune response. Thus, in HNSCC hypermethylated CDH1 gene, involved in cell-cell adhesion, could be considered as a good biomarker for distinguishing cancer tissues from potentially premalignant oral lesions and from healthy oral mucosa. In addition, hypermethylated gene promoters of SLC5A10, involved in transport, and TBC1D2, involved in cell-cell adhesion, could be also good biomarkers for distinguishing HNSCC from lesions as well as potentially premalignant oral lesions from healthy oral tissues.

Declarations

Ethics approval and consent to participate

This study was approved by the Ethical Board of the Ruđer Bošković Institute, the Ethical Board of the Clinical Hospital Dubrava, and the School of Dental Medicine, University of Zagreb.

Consent for publication

Consent for publication is obtained from each participant in the written form.

Availability of data and material

All data generated or analyzed during this study are included in this manuscript and the supplementary information file.

Competing interests

The authors have no competing interest to declare.

Funding

This work was supported by funds from the Croatian Science Foundation (Grant code IP-2013-11-4758; Epigenetic changes in head and neck squamous cell carcinoma - Epic-HNSCC), which had no influence on the content of the manuscript. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

Authors' contributions

MG and NMG designed the study. NMG was responsible for the methylation analyses and has written the manuscript. IS has done the bioinformatic analyses and contributed to the writing of the manuscript. KB was a major contributor for samples processing, HPV testing and pyrosequencing analyses. ED and MMS were responsible for samples collection and patient data processing. DL has done interpretation of Illumina iScan data. SDM performed the Illumina chip analyses. MG conceived the study, ensured funding, and revised the manuscript. All authors have read and approved the final version of the manuscript.

Acknowledgements

The authors are thankful to the clinicians that were involved in the collection of tissue samples. The authors are also grateful to Mrs Jasminka Golubić Talić for her technical assistance, and to Dr Nino Sinčić and Jure Krasić for enabling the pyrosequencing analysis. Special thanks to Dr Nathaniel Edward Bennett Saidu for manuscript editing and critical review of the study.

References

1. Globocan. Globocan 2012 - Home [Internet]. 2012 [cited 2018 Jun 13]. Available from: http://globocan.iarc.fr/Default.aspx
2. Arantes LMRB, De Carvalho AC, Melendez ME, Lopes Carvalho A. Serum, plasma and saliva biomarkers for head and neck cancer. Expert Rev Mol Diagn. 2018;18.
3. Alexandrov LB, Ju YS, Haase K, Loo PV, Martincorena I, Nik-Zainal S, et al. Mutational signatures associated with tobacco smoking in human cancer. Science [Internet]. 2016 [cited 2018 Sep 12];354:618–22. Available from: http://science.sciencemag.org/content/354/6312/618
4. Burris JL, Studts JL, DeRosa AP, Ostroff JS. Systematic Review of Tobacco Use After Lung or Head/Neck Cancer Diagnosis: Results and Recommendations for Future Research. Cancer Epidemiol Biomarkers Prev [Internet]. 2015 [cited 2018 Sep 12];24:1450–61. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4592460/
5. Kawakita D, Matsuo K. Alcohol and head and neck cancer. Cancer Metastasis Rev [Internet]. 2017 [cited 2018 Sep 12];36:425–34. Available from: https://doi.org/10.1007/s10555-017-9690-0
6. Hashibe M, Brennan P, Benhamou S, Castellsague X, Chen C, Curado MP, et al. Alcohol drinking in never users of tobacco, cigarette smoking in never drinkers, and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. J Natl Cancer Inst. 2007;99:777–89.
7. Hashibe M, Brennan P, Chuang S-C, Boccia S, Castellsague X, Chen C, et al. Interaction between tobacco and alcohol use and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. Cancer Epidemiol Biomarkers Prev. 2009;18:541–50.
8. Marur S, D'Souza G, Westra WH, Forastiere AA. HPV-associated Head and Neck Cancer: A Virus-related Cancer Epidemic – A Review of Epidemiology, Biology, Virus Detection and Issues in Management. Lancet Oncol [Internet]. 2010 [cited 2018 Sep 12];11:781–9. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5242182/
9. D'Souza G, Dempsey A. The role of HPV in head and neck cancer and review of the HPV vaccine. Prev Med [Internet]. 2011 [cited 2018 Sep 12];53:S5–11. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3287051/
10. Kobayashi K, Hisamatsu K, Suzui N, Hara A, Tomita H, Miyazaki T, et al. A Review of HPV-Related Head and Neck Cancer. Journal of Clinical Medicine [Internet]. 2018 [cited 2018 Sep 12];7:241. Available from: http://www.mdpi.com/2077-0383/7/9/241
11. Shah A, Malik A, Garg A, Mair M, Nair S, Chaturvedi P. Oral sex and human papilloma virus-related head and neck squamous cell cancer: a review of the literature. Postgrad Med J. 2017;
12. Ghantous Y, Schussel JL, Brait M. Tobacco and alcohol-induced epigenetic changes in oral carcinoma: Current Opinion in Oncology [Internet]. 2018 [cited 2018 Sep 12];1. Available from: http://Insights.ovid.com/crossref?an=00001622-900000000-99265

13. Wilson GA, Lechner M, Köferle A, Caren H, Butcher LM, Feber A, et al. Integrated virus-host methylome analysis in head and neck squamous cell carcinoma. Epigenetics [Internet]. 2013 [cited 2018 Sep 12];8:953–61. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3883772/
14. Sanderson RJ, Ironside JAD. Squamous cell carcinomas of the head and neck. BMJ [Internet]. 2002 [cited 2018 Sep 12];325:822–7. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1124330/
15. Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. Nature Reviews Genetics [Internet]. 2002 [cited 2018 Sep 12];3:415–28. Available from: https://www.nature.com/articles/nrg816
16. Herceg Z, Hainaut P. Genetic and epigenetic alterations as biomarkers for cancer detection, diagnosis and prognosis. Mol Oncol [Internet]. 2007 [cited 2018 Sep 12];1:26–41. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5543860/
17. van der Meij EH, Mast H, van der Waal I. The possible premalignant character of oral lichen planus and oral lichenoid lesions: a prospective five-year follow-up study of 192 patients. Oral Oncol. 2007;43:742–8.
18. Casparis S, Borm JM, Tektas S, Kamarachev J, Locher MC, Damerau G, et al. Oral lichen planus (OLP), oral lichenoid lesions (OLL), oral dysplasia, and oral cancer: retrospective analysis of clinicopathological data from 2002–2011. Oral Maxillofac Surg [Internet]. 2014 [cited 2015 Nov 6];19:149–56. Available from: http://link.springer.com/article/10.1007/s10006-014-0469-y
19. Dudhia BB, Dudhia SB, Patel PS, Jani YV. Oral lichen planus to oral lichenoid lesions: Evolution or revolution. J Oral Maxillofac Pathol [Internet]. 2015 [cited 2018 May 16];19:364–70. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4774291/
20. Mravak-Stipetić M, Sabol I, Kranjčić J, Knežević M, Grce M. Human Papillomavirus in the Lesions of the Oral Mucosa According to Topography. PLoS One [Internet]. 2013 [cited 2018 Sep 13];8. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3726768/
21. Ambrosi C, Manzo M, Baubec T. Dynamics and Context-Dependent Roles of DNA Methylation. Journal of Molecular Biology [Internet]. 2017 [cited 2018 Sep 12];429:1459–75. Available from: http://www.sciencedirect.com/science/article/pii/S0022283617300839
22. Manos MM, Ting Y, Wright DK, Levwis AJ, Broker TR, Wolinsky SM. The use of polymerase chain reaction amplification for the detection of genital human papillomaviruses. U: Cancer Cells. 7:; 1989. page 209–14.
23. Milutin-Gasperov N, Sabol I, Halec G, Matovina M, Grce M. Retrospective study of the prevalence of high-risk human papillomaviruses among Croatian women. Coll Antropol [Internet]. 2007 [cited 2012 Jun 15];31 Suppl 2:89–96. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17598510
24. Milutin Gašperov N, Sabol I, Planinić P, Grubišić G, Fistonić I, Ćorušić A, et al. Methylated Host Cell Gene Promoters and Human Papillomavirus Type 16 and 18 Predicting Cervical Lesions and Cancer. PLoS ONE. 2015;10:e0129452.

25. Pidsley R, Zotenko E, Peters TJ, Lawrence MG, Risbridger GP, Molloy P, et al. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. Genome Biology [Internet]. 2016 [cited 2018 May 22];17. Available from: http://genomebiology.biomedcentral.com/articles/10.1186/s13059-016-1066-1
26. Morris TJ, Butcher LM, Feber A, Teschendorff AE, Chakravarthy AR, Wojdacz TK, et al. ChAMP: 450k Chip Analysis Methylation Pipeline. Bioinformatics [Internet]. 2014 [cited 2018 Sep 7];30:428–30. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3904520/
27. Assenov Y, Müller F, Lutsik P, Walter J, Lengauer T, Bock C. Comprehensive Analysis of DNA Methylation Data with RnBeads. Nat Methods [Internet]. 2014 [cited 2018 Sep 7];11:1138–40. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4216143/
28. Dedeurwaerder S, Defrance M, Calonne E, Denis H, Sotiriou C, Fuks F. Evaluation of the Infinium Methylation 450K technology. Epigenomics. 2011;3:771–84.
29. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007;8:118–27.
30. Smyth GK, Michaud J, Scott HS. Use of within-array replicate spots for assessing differential expression in microarray experiments. Bioinformatics. 2005;21:2067–75.
31. Mikeska T, Felsberg J, Hewitt CA, Dobrovic A. Analysing DNA Methylation Using Bisulphite Pyrosequencing. Epigenetics Protocols [Internet]. Humana Press; 2011 [cited 2018 May 23]. page 33–53. Available from: https://link.springer.com/protocol/10.1007/978-1-61779-316-5_4
32. Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. Nucleic Acids Res [Internet]. [cited 2019 May 28]; Available from: https://academic.oup.com/nar/advance-article/doi/10.1093/nar/gkz401/5494758
33. Díez-Villanueva A, Mallona I, Peinado MA. Wanderer, an interactive viewer to explore DNA methylation and gene expression data in human cancer. Epigenetics Chromatin [Internet]. 2015 [cited 2019 May 28];8. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4480445/
34. Chen BS, Wang MR, Cai Y, Xu X, Xu ZX, Han YL, et al. Decreased expression of SPRR3 in Chinese human oesophageal cancer. Carcinogenesis. 2000;21:2147–50.
35. Zhang H-J, Tian J, Qi X-K, Xiang T, He G-P, Zhang H, et al. Epstein-Barr virus activates F-box protein FBXO2 to limit viral infectivity by targeting glycoprotein B for degradation. PLoS Pathog. 2018;14:e1007208.
36. GeneCards - Human Genes | Gene Database | Gene Search [Internet]. [cited 2018 Aug 22]. Available from: https://www.genecards.org/
37. Taberna M, Mena M, Pavón MA, Alemany L, Gillison ML, Mesía R. Human papillomavirus-related oropharyngeal cancer. Ann Oncol [Internet]. 2017 [cited 2018 Jan 25];28:2386–98. Available from: https://academic.oup.com/annonc/article/28/10/2386/3868408
38. Božinović K, Sabol I, Rakušić Z, Jakovčević A, Šekerija M, Lukinović J, et al. HPV-driven oropharyngeal squamous cell cancer in Croatia — Demography and survival. PLoS One [Internet].

2019 [cited 2019 Feb 27];14. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6358070/

39. Lechner M, Fenton T, West J, Wilson G, Feber A, Henderson S, et al. Identification and functional validation of HPV-mediated hypermethylation in head and neck squamous cell carcinoma. Genome Med [Internet]. 2013 [cited 2019 Dec 6];5:15. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3706778/
40. LIU Q, ZHANG C, MA G, ZHANG Q. Expression of SPRR3 is associated with tumor cell proliferation and invasion in glioblastoma multiforme. Oncol Lett [Internet]. 2014 [cited 2018 May 30];7:427–32. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3881942/
41. Zhang Y, Feng Y-B, Shen X-M, Chen B-S, Du X-L, Luo M-L, et al. Exogenous expression of Esophagin/SPRR3 attenuates the tumorigenicity of esophageal squamous cell carcinoma cells via promoting apoptosis. International Journal of Cancer [Internet]. [cited 2018 Jul 13];122:260–6. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/ijc.23104
42. Solomon B, Young RJ, Rischin D. Head and neck squamous cell carcinoma: Genomics and emerging biomarkers for immunomodulatory cancer treatments. Semin Cancer Biol. 2018;52:228–40.
43. Strzelczyk JK, Krakowczyk Ł, Owczarek AJ. Aberrant DNA methylation of the *p16, APC, MGMT, TIMP3* and *CDH1* gene promoters in tumours and the surgical margins of patients with oral cavity cancer. Journal of Cancer [Internet]. 2018 [cited 2018 Jun 15];9:1896–904. Available from: http://www.jcancer.org/v09p1896.htm
44. Kim J-S, Kim SY, Lee M, Kim SH, Kim S-M, Kim EJ. Radioresistance in a human laryngeal squamous cell carcinoma cell line is associated with DNA methylation changes and topoisomerase II α. Cancer Biol Ther [Internet]. 2015 [cited 2018 Jun 15];16:558–66. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4622512/
45. Gentilini D, Scala S, Gaudenzi G, Garagnani P, Capri M, Cescon M, et al. Epigenome-wide association study in hepatocellular carcinoma: Identification of stochastic epigenetic mutations through an innovative statistical approach. Oncotarget [Internet]. 2017 [cited 2018 Jun 15];8:41890–902. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5522036/
46. Paz N, Levanon EY, Amariglio N, Heimberger AB, Ram Z, Constantini S, et al. Altered adenosine-to-inosine RNA editing in human cancer. Genome Res [Internet]. 2007 [cited 2018 Aug 23];17:1586–95. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2045141/
47. Wen G, Wang H, Zhong Z. Associations of RASSF1A, RARβ, and CDH1 promoter hypermethylation with oral cancer risk. Medicine (Baltimore) [Internet]. 2018 [cited 2018 Jun 15];97. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5882397/
48. Liu G. CDH1 promoter methylation in patients with cervical carcinoma: a systematic meta-analysis with trial sequential analysis. Future Oncology [Internet]. 2017 [cited 2018 Jun 15];14:51–63. Available from: https://www.futuremedicine.com/doi/abs/10.2217/fon-2017-0267
49. Lee J-J, Kao K-C, Chiu Y-L, Jung C-J, Liu C-J, Cheng S-J, et al. Enrichment of Human CCR6+ Regulatory T Cells with Superior Suppressive Activity in Oral Cancer. The Journal of Immunology

[Internet]. 2017 [cited 2018 Aug 27];199:467–76. Available from: http://www.jimmunol.org/content/199/2/467

50. Wu J, Xu X, Lee E-J, Shull AY, Pei L, Awan F, et al. Phenotypic alteration of CD8+ T cells in chronic lymphocytic leukemia is associated with epigenetic reprogramming. Oncotarget [Internet]. 2016 [cited 2018 Jun 15];7:40558–70. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5130028/
51. Wu C-Y, Tseng R-C, Hsu H-S, Wang Y-C, Hsu M-T. Frequent down-regulation of hRAB37 in metastatic tumor by genetic and epigenetic mechanisms in lung cancer. Lung Cancer [Internet]. 2009 [cited 2018 Jun 15];63:360–7. Available from: http://www.sciencedirect.com/science/article/pii/S0169500208003577
52. Bond HM, Scicchitano S, Chiarella E, Amodio N, Lucchino V, Aloisio A, et al. ZNF423: A New Player in Estrogen Receptor-Positive Breast Cancer. Front Endocrinol (Lausanne) [Internet]. 2018 [cited 2018 Jun 15];9. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5968090/
53. Rivas-Ortiz CI, Lopez-Vidal Y, Arredondo-Hernandez LJR, Castillo-Rojas G. Genetic Alterations in Gastric Cancer Associated with Helicobacter pylori Infection. Front Med (Lausanne) [Internet]. 2017 [cited 2018 Jun 15];4. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5411440/
54. Chen C, Peng H, Huang X, Zhao M, Li Z, Yin N, et al. Genome-wide profiling of DNA methylation and gene expression in esophageal squamous cell carcinoma. Oncotarget [Internet]. 2015 [cited 2018 Jun 15];7:4507–21. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4826222/
55. Hatano Y, Semi K, Hashimoto K, Lee MS, Hirata A, Tomita H, et al. Reducing DNA methylation suppresses colon carcinogenesis by inducing tumor cell differentiation. Carcinogenesis. 2015;36:719–29.
56. Kim JH, Karnovsky A, Mahavisno V, Weymouth T, Pande M, Dolinoy DC, et al. LRpath analysis reveals common pathways dysregulated via DNA methylation across cancer types. BMC Genomics [Internet]. 2012 [cited 2019 Dec 18];13:526. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3505188/
57. Mielcarek-Kuchta D, Paluszczak J, Seget M, Kiwerska K, Biczysko W, Szyfter K, et al. Prognostic factors in oral and oropharyngeal cancer based on ultrastructural analysis and DNA methylation of the tumor and surgical margin. Tumour Biol [Internet]. 2014 [cited 2018 Sep 5];35:7441–9. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4158182/
58. Juodzbalys G, Kasradze D, Cicciù M, Sudeikis A, Banys L, Galindo-Moreno P, et al. Modern molecular biomarkers of head and neck cancer. Part I. Epigenetic diagnostics and prognostics: Systematic review. Cancer Biomarkers [Internet]. 2017 [cited 2017 Sep 19];17:487–502. Available from: http://www.medra.org/servlet/aliasResolver?alias=iospress&doi=10.3233/CBM-160666
59. Wong T-S, Gao W, Chan JY-W. Interactions between E-Cadherin and MicroRNA Deregulation in Head and Neck Cancers: The Potential Interplay. Biomed Res Int [Internet]. 2014 [cited 2019 Dec 17];2014. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4138976/

Tables

Table 1. Aberrant gene promoter methylation in HNSCC tissue compared to control healthy tissue. The list is merged of top fifteen differentially methylated (according to extent of methylation difference value) genes.

Gene name	Function	mean.mean.diff*	comb.p.adj.fdr**
Hypermethylated gene promoters			
GPRC5D	G protein-coupled receptor	0.78	6.49E-07
TMPRSS11B	Transmembrane protease	0.76	2.04E-08
PIAS2	Sumoylation	0.66	1.22E-06
ARG1	Arginase activity	0.64	2.50E-07
SRPK2	Protein kinase	0.62	1.30E-06
AADACL2	Hydrolase activity	0.59	7.38E-07
RGPD4	RNA transport	0.58	4.71E-06
SPRR3	Structural molecule activity	0.53	2.44E-07
DEGS1	Desaturase activity	0.49	2.04E-08
TXNDC8	Oxidoreductase activity	0.48	1.15E-05
SH3TC1	Myelination	0.47	2.05E-06
ZPLD1	Cerebral malformations	0.47	6.17E-06
FBXO2	Ubiquitination	0.46	0.000236
ATG16L1	Autophagy	0.46	0.000306
GRHL1	Transcription factor	0.44	5.12E-07
Hypomethylated gene promoters			
TRBC2	T cell receptor	-0.80	2.48E-07
DGAT2	Acyltransferase activity	-0.70	2.48E-07
ALG1L	Transferase activity	-0.70	3.29E-05
PDE4D	Enzyme binding	-0.68	1.06E-05
TRDC	T cell receptor	-0.67	1.46E-05
DNAJC6	Phosphatase activity	-0.67	1.63E-06
IGKV3-20	Immunoglobulin receptor binding	-0.66	1.71E-05
TMEM150B	Transmembrane protein	-0.66	8.13E-05
LAIR2	Innate immune response	-0.65	1.81E-05
UBQLN3	Protein degradation	-0.64	2.57E-06
ANKFN1	Not known	-0.64	1.71E-07
MS4A1	Differentiation of B-cells	-0.63	3.80E-05
CCT8L2	Channel activity	-0.62	3.59E-06

SPOCK1	Not known	-0.61	5.06E-05
IGHV4-39	Antigen recognition	-0.61	9.44E-07

*mean.mean.diff = the mean difference across all sites in a region; **comb.p.adj.fdr = FDR adjusted combined P-value

Table 2. Aberrant CpG sites methylation in HNSCC tissue compared to control healthy tissue. The list is merged of top fifteen differentially methylated (according to extent of methylation difference value) genes.

Gene name	Function	cg position	mean.mean.diff*	comb.p.adj.fdr**
Hypermethylated CpG sites				
LMBR1L	Probable receptor	cg12348519	0.93	4.59E-08
CDH1	Adhesions, mobility and proliferation	cg08285862	0.92	3.49E-08
EIF6	Initiation of translation	cg09957666	0.92	1.96E-08
C16orf70	Not known	cg03664901	0.92	3.48E-08
ETNK2	Transferase and kinase activity	cg12142497	0.92	5.33E-08
C11orf73	Cellular response to heat stress	cg23450586	0.91	5.01E-09
ADARB2	RNA editing	cg26569590	0.91	2.10E-08
GAB1	Cellular growth, transformation and apoptosis	cg23020414	0.91	2.83E-08
ITPR3	Metabolism and growth	cg05876496	0.91	6.35E-08
WDR61	Transcriptional regulation	cg12339790	0.90	4.35E-08
PGAP2	Protein transport	cg01156876	0.90	1.24E-08
DDX10	RNA helicase	cg18585558	0.90	6.14E-09
DGKH	Kinase activity	cg22899750	0.90	1.09E-07
RAB40C	Oncogene	cg01770948	0.89	2.00E-08
BEAN1	Not known	cg19471156	0.89	5.59E-08
Hypomethylated CpG sites				
ATXN1	Not known	cg07713291	-0.96	2.97E-09
PPP2R2C	Cell growth	cg05805165	-0.93	1.22E-08
CCR6	Immune regulation	cg05094429	-0.92	1.08E-07
RAB37	Oncogene	cg25267982	-0.92	1.29E-08
DUSP27	Phosphatase activity	cg23713934	-0.91	1.94E-07
ZNF521	Transcription factor	cg21830945	-0.91	7.58E-08
SLC6A17	Transporter	cg12072789	-0.90	7.38E-08
SPIN1	Methylated histone binding	cg13554018	-0.90	1.63E-08
CXCR1	Receptor	cg13519373	-0.90	1.22E-08
SPTBN1	Cell shape	cg06149826	-0.89	6.91E-09
NBAS	Golgi to ER transport	cg27424261	-0.89	5.31E-07
NRG3	Ligand	cg10656958	-0.89	9.45E-08
COL5A1	Forming collagen	cg26087052	-0.89	3.96E-08
CDX1	Transcriptional regulation	cg12473781	-0.89	5.77E-08

BATF3	Transcriptional regulation	cg03219362	-0.89	2.10E-08

*mean.mean.diff, the mean difference across all sites in a region; **comb.p.adj.fdr, FDR adjusted combined P-value

Table 3. Aberrant gene promoter methylation in HNSCC tissue compared to oral lesions. The list is merged of top fifteen differentially methylated (according to extent of methylation difference value) genes.

Gene name	Function	mean.mean.diff*	comb.p.adj.fdr**
Hypermethylated gene promoters			
RAD51B	RAD51 Paralog B	0.85	2.53E-08
BARX2	BARX Homeobox 2	0.81	3.94E-08
SLC5A10;FAM83G	Solute Carrier Family 5 Member 10	0.78	2.58E-08
NINL	Ninein Like	0.77	3.96E-08
NSMCE2	NSE2/MMS21 Homolog, SMC5-SMC6 Complex SUMO Ligase	0.76	6.36E-07
PGAP2	Post-GPI Attachment To Proteins 2	0.75	9.48E-08
INO80C	INO80 Complex Subunit C	0.74	1.96E-09
IL34	Interleukin 34	0.74	2.20E-09
ZNF516	Zinc Finger Protein 516	0.73	4.90E-08
GFOD2	Glucose-Fructose Oxidoreductase Domain Containing 2	0.73	1.36E-07
PARD3	Par-3 Family Cell Polarity Regulator	0.73	1.36E-07
MCEE	Methylmalonyl-CoA Epimerase	0.72	2.89E-08
POLM	DNA Polymerase Mu	0.72	3.93E-07
ASPG	Asparaginase	0.71	4.43E-08
TBC1D2	TBC1 Domain Family Member 2	0.71	3.74E-07
Hypomethylated gene promoters			
ART4	ADP-Ribosyltransferase 4 (Dombrock Blood Group)	-0.88	7.92E-11
EPB41L3	Erythrocyte Membrane Protein Band 4.1 Like 3	-0.87	6.18E-11
ESRRG	Estrogen Related Receptor Gamma	-0.86	8.51E-09
ENPP1	Ectonucleotide Pyrophosphatase/Phosphodiesterase 1	-0.86	1.26E-09
GNG7	G Protein Subunit Gamma 7	-0.86	4.75E-09
PAPSS2	3'-Phosphoadenosine 5'-Phosphosulfate Synthase 2	-0.85	4.10E-09
NGEF	Neuronal Guanine Nucleotide Exchange Factor	-0.84	1.87E-09
HIPK4	Homeodomain Interacting Protein Kinase 4	-0.84	6.69E-09
GPR158	G Protein-Coupled Receptor 158	-0.83	9.82E-10
GSG1L	GSG1 Like	-0.83	1.04E-08
SMPD3	Sphingomyelin Phosphodiesterase 3	-0.83	1.64E-08
GDF2	Growth Differentiation Factor 2	-0.83	5.15E-10
RERE	Arginine-Glutamic Acid Dipeptide Repeats	-0.82	2.19E-08

CDH13	Cadherin 13	-0.82	1.81E-10
HS3ST4	Heparan Sulfate-Glucosamine 3-Sulfotransferase 4	-0.82	1.02E-08

*mean.mean.diff, the mean difference across all sites in a region; **comb.p.adj.fdr, FDR adjusted combined P-value

Table 4. Aberrant CpG sites methylation in HNSCC tissue compared to oral lesions. The list is merged of top fifteen differentially methylated (according to extent of methylation difference value) genes.

Gene name	Function	cg position	mean.mean.diff*	comb.p.adj.fdr**
Hypermethylated CpG sites				
EIF6	Eukaryotic Translation Initiation Factor 6	cg09957666	0.91	5.68E-10
KANSL1	KAT8 Regulatory NSL Complex Subunit 1	cg07281649	0.91	1.43E-09
DDX10	DEAD-Box Helicase 10	cg18585558	0.89	3.58E-10
AP2A1	Adaptor Related Protein Complex 2 Alpha 1 Subunit	cg08969148	0.89	8.97E-10
RAB40C	RAB40C, Member RAS Oncogene Family	cg01770948	0.89	1.84E-09
GAB1	GRB2 Associated Binding Protein 1	cg23020414	0.88	5.30E-09
ERGIC1	Endoplasmic Reticulum-Golgi Intermediate Compartment 1	cg07769006	0.88	1.25E-09
SNX14	Sorting Nexin 14	cg03776905	0.88	3.20E-09
PIGU	Phosphatidylinositol Glycan Anchor Biosynthesis Class U	cg09450087	0.88	1.22E-10
ARAP1	ArfGAP With RhoGAP Domain, Ankyrin Repeat And PH Domain 1	cg09010791	0.87	1.56E-09
LMTK2	Lemur Tyrosine Kinase 2	cg05941925	0.87	2.59E-09
BEAN1	Brain Expressed Associated With NEDD4 1	cg19471156	0.87	6.96E-09
AP1S3	Adaptor Related Protein Complex 1 Sigma 3 Subunit	cg25666945	0.87	1.66E-09
CDH1	Cadherin 1	cg08285862	0.87	2.28E-08
RYBP	RING1 And YY1 Binding Protein	cg08086385	0.86	3.11E-10
Hypomethylated CpG sites				
FAM69A	Family With Sequence Similarity 69 Member A	cg22727960	-0.93	7.05E-11
ATP6V0A1	ATPase H+ Transporting V0 Subunit A1	cg19022525	-0.92	9.01E-11
LBP	Lipopolysaccharide Binding Protein	cg18979491	-0.92	3.02E-10
WDR25	WD Repeat Domain 25	cg24211276	-0.91	6.22E-11
SH3RF3	SH3 Domain Containing Ring Finger 3	cg27294813	-0.91	1.01E-09
NINJ2	Ninjurin 2	cg05534515	-0.91	2.74E-12
RAB37	RAB37, Member RAS Oncogene Family	cg25267982	-0.90	1.17E-09
CXCR1	C-X-C Motif Chemokine Receptor 1	cg13519373	-0.90	1.76E-10
SPTBN	Spectrin Beta, Non-Erythrocytic 1	cg06149826	-0.90	1.54E-10
RHOH	Ras Homolog Family Member H	cg15729055	-0.90	1.90E-09
GRIK5	Glutamate Ionotropic Receptor Kainate Type Subunit 5	cg03100024	-0.90	2.47E-09
KLRD1	Killer Cell Lectin Like Receptor D1	cg05377120	-0.90	6.88E-09
TENM2	Teneurin Transmembrane Protein 2	cg26758826	-0.89	3.56E-11

FAM69A	Family With Sequence Similarity 69 Member A	cg05172999	-0.89	3.14E-10
ITK	IL2 Inducible T Cell Kinase	cg12250498	-0.89	3.46E-10

*mean.mean.diff, the mean difference across all sites in a region; **comb.p.adj.fdr, FDR adjusted combined P-value

Figures

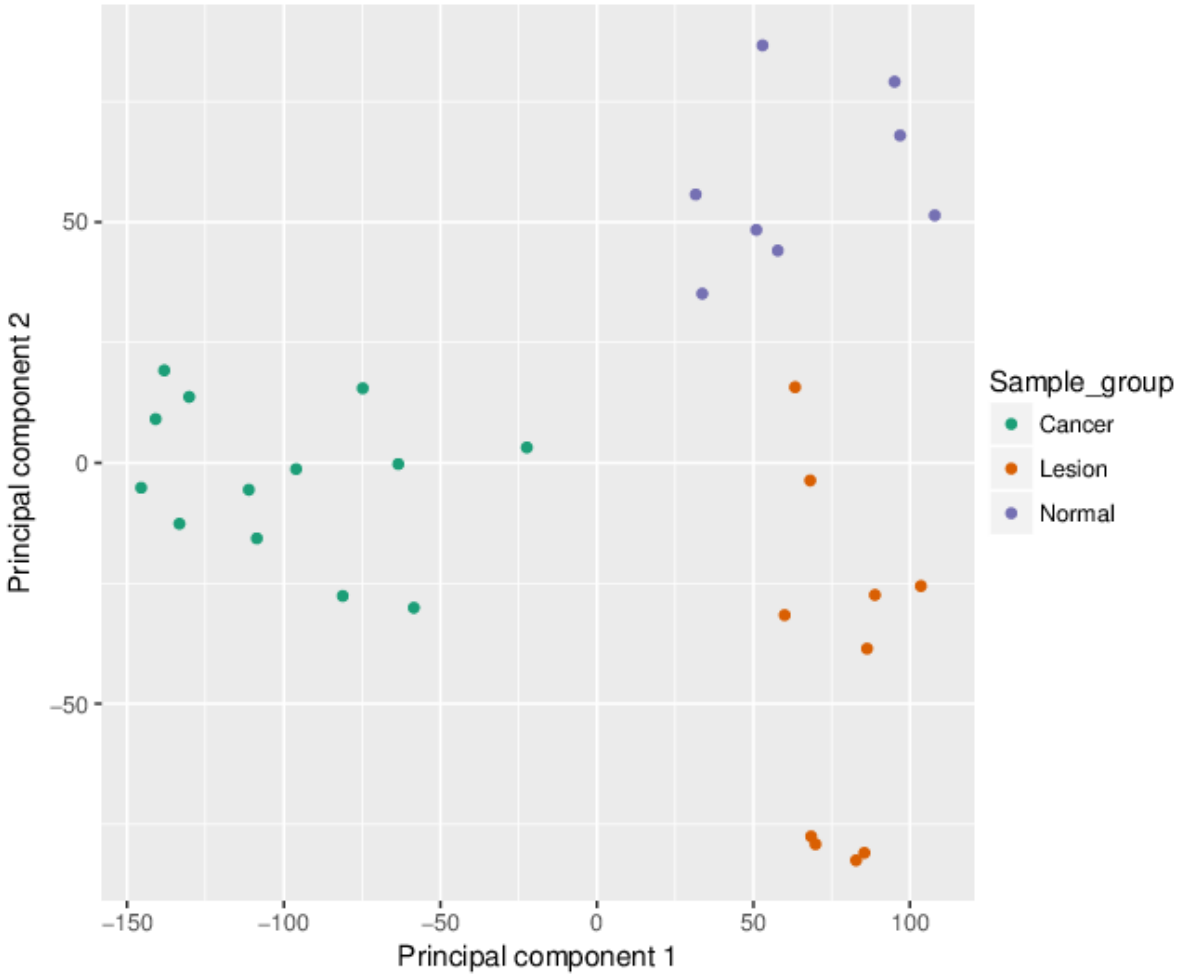

Figure 1

Infinium MethylationEPIC BeadChip findings: samples grouping by the principal component. Green: HNSCC samples, oropharyngeal cancer (n=6) and oral cancer (n=7), orange: oral lesions, OLP (n=8) and OLL (n=2), violet: healthy oral mucosa (n=8).

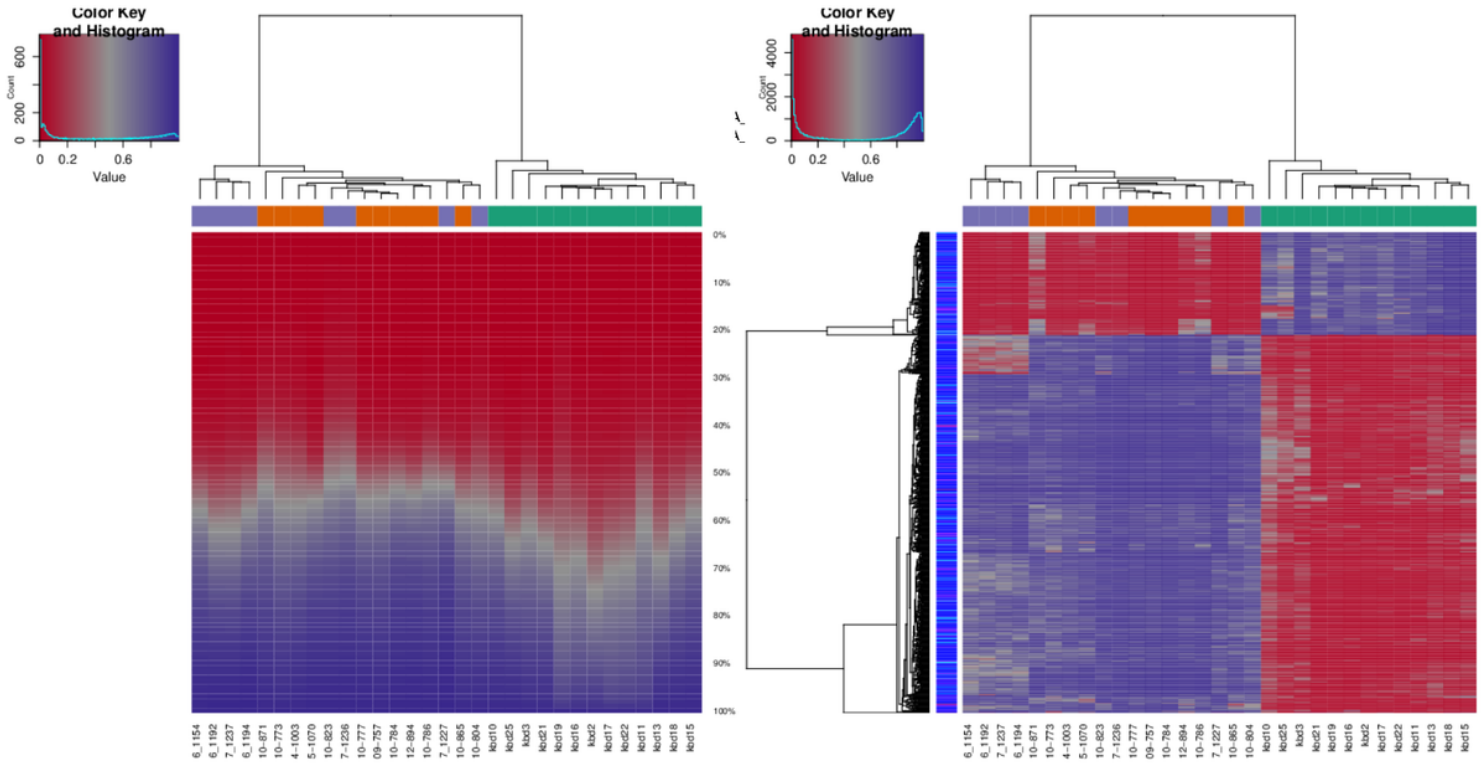

Figure 2

Infinium MethylationEPIC BeadChip findings: hierarchical clustering of samples based on all methylation values. A) The heatmap displays methylation percentiles per sample. B) The heatmap displays only selected sites/regions with the highest variance across all samples. Green: HNSCC samples, oropharyngeal cancer (n=6) and oral cancer (n=7), orange: oral lesions, OLP (n=8) and OLL (n=2), violet: healthy oral mucosa (n=8).

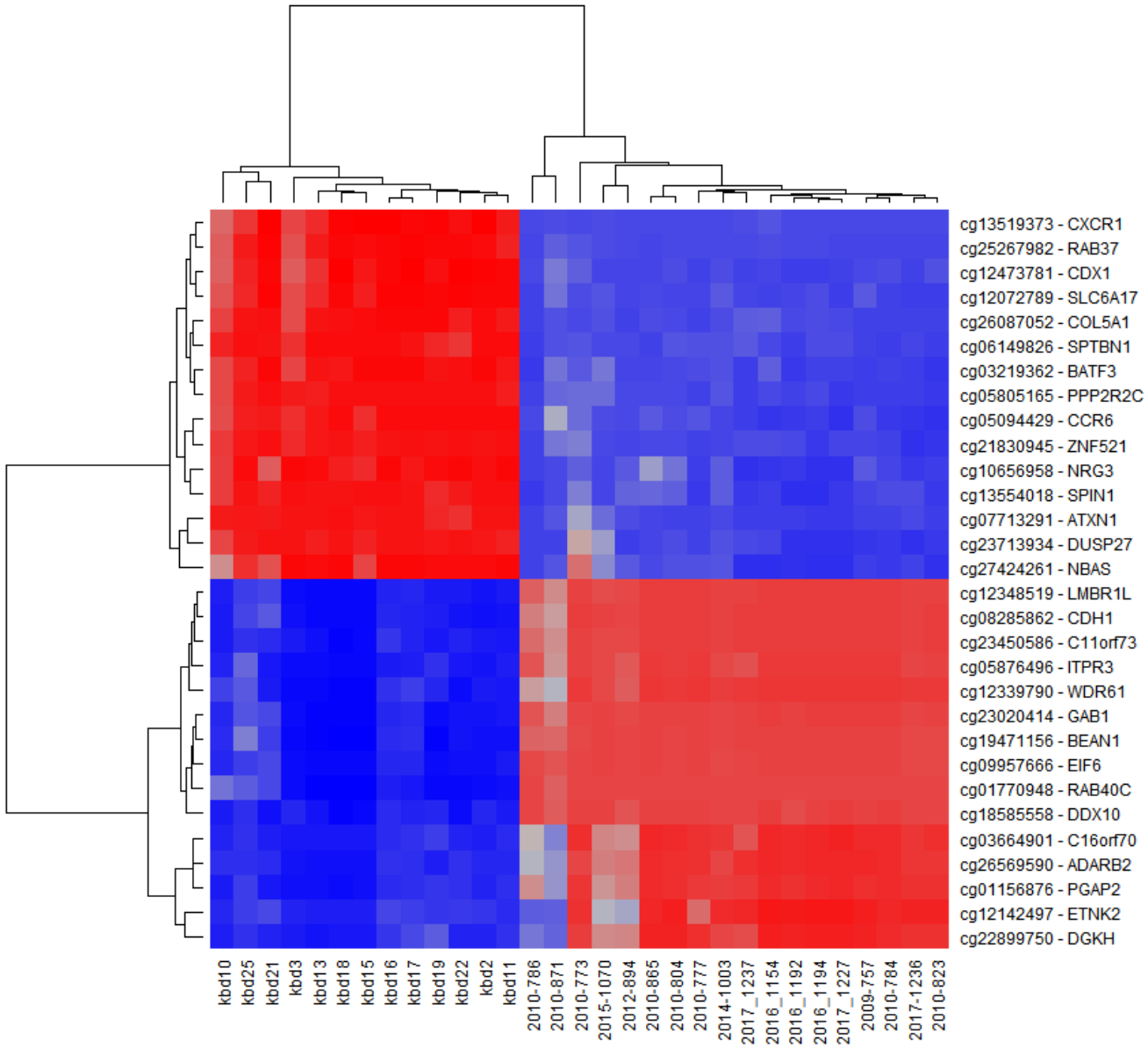

Figure 3

Infinium MethylationEPIC BeadChip findings: heatmap representation of samples clustering for top fifteen hypermethylated (blue) and top fifteen hypomethylated (red) CpG sites in cancer tissue compared to control healthy tissue (according to methylation difference value; Table 2). Left side: HNSCC samples (n=13; kbd10–kbd11); right side: healthy oral mucosa (n=8; 2010-804, 2017-1237–2017-1227, 2017-1236, 2010-823); middle: oral lesions (n=10; 2010-786–2010-865, 2010-777, 2014-1003, 2009-757, 2010-784).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Milutin.GasperovAdditionalfiles.docx