Josip Juraj Strossmayer University of Osijek

University of Dubrovnik

Ruđer Bošković Institute

University Postgraduate Interdisciplinary Doctoral Study of

Molecular biosciences

Ksenija Božinović

IDENTIFICATION OF SPECIFIC EPIGENETIC BIOMARKERS IN HPV-POSITIVE HEAD AND NECK SQUAMOUS CELL CARCINOMA

PhD thesis

Osijek, 2018

ACKNOWLEDGEMENT

For the continuous support during my PhD study, I would like to express my sincere gratitude to my mentors and advisers, Dr Magdalena Grce and Dr Ivan Sabol, as well as for their patience, motivation, and immense knowledge. I could not have imagined having a better advisers and mentors for my PhD study.

Besides my mentors, I would like to thank the members of my thesis committee, Dr Vjekoslav Tomaić, Prof Vera Cesar and Prof Sanja Kapitanović, for their insightful comments and encouragement, but also for the hard questions, which motivated me to widen my research from various perspectives.

My gratitude especially goes to Dr Ruth Tachezy for providing some study samples and for her fruitful advices during my PhD study.

My sincere thanks also goes to Dr Zoran Tadić, Dr Mirna Golemović and Dr Srđan Verstovšek, who provided me an opportunity to join their teams as a Graduate student and provided me with my first scientific work experience. They gave me a foundation for my future research pathway, which motivated me to start my PhD program and continue pursuing a scientific career. Without their support, it would not be possible to conduct this research.

I am grateful to my colleagues from the Laboratory of Molecular Virology and Bacteriology, Division of Molecular Medicine, Ruđer Bošković Institute, Zagreb, and to all my friends from the Institute and wider for their stimulating discussions, and for all the fun we have had in the last four years. I will never forget our lunch and coffee breaks, which made every single day better.

I am especially grateful to my family, my parents Stjepan and Slavica, and to my sister Nikolina for supporting me spiritually throughout my PhD research as well as writing this thesis, and in my life in general. They are the reason why I am here. Thank you for making me feel confident and helping me believe in myself.

Last but not the least, special thanks to my best friend and my soul mate, Krešo for making me a better person every day. I would not be complete without him.

BASIC DOCUMENTATION CARD

Josip Juraj Strossmayer University of Osijek **PhD thesis**
University of Dubrovnik
Ruđer Bošković Institute
University Postgraduate Interdisciplinary Doctoral Study of
Molecular biosciences

Scientific Area: Interdisciplinary Area of Science
Scientific Fields: Basic Medical Sciences and Biology

IDENTIFICATION OF SPECIFIC EPIGENETIC BIOMARKERS IN HPV-POSITIVE HEAD AND NECK SQUAMOUS CELL CARCINOMA

Ksenija Božinović

Thesis has been performed at the Laboratory of molecular virology and bacteriology, at the Division of Molecular Medicine, Ruđer Bošković Institute

Supervisor 1. Magdalena Grce, PhD, senior scientist
Supervisor 2. Ivan Sabol, PhD, research associate

Short abstract: Head and neck squamous cell carcinoma (HNSCC) is broadly categorized into two groups: HPV-positive and HPV-negative tumors. The aim of this study is to find specific biomarkers, which could improve diagnostic, prognostic and therapeutic methods in both groups. Epigenetic deregulation studies in cancer samples can offer such biomarkers. Therefore, we performed a whole-genome methylation and miRNA profiling in fresh tumor samples, and assessed survival analysis in archival and fresh tumor samples to obtain a clear picture of the disease development.

Number of pages: 151
Number of figures: 44
Number of tables: 14
Number of references: 301
Original in: English
Key words: HNSCC, oropharyngeal, oral, HPV, methylation, miRNA

Date of the thesis defense: September 25th, 2018
Reviewers: 1. Vjekoslav Tomaić, PhD
2. Vera Cesar, PhD
3. Sanja Kapitanović, PhD
Substitute: Jelena Knežević, PhD

Thesis deposited in: National and University Library in Zagreb, Ul. Hrvatske bratske zajednice 4, Zagreb; City and University Library of Osijek, Europska avenija 24, Osijek; Josip Juraj Strossmayer University of Osijek, Trg sv. Trojstva 3, Osijek

TEMELJNA DOKUMENTACIJSKA KARTICA

Sveučilište Josipa Jurja Strossmayera u Osijeku **Doktorska disertacija**
Sveučilište u Dubrovniku
Institut Ruđer Bošković
Sveučilišni poslijediplomski interdisciplinarni doktorski studij
Molekularne bioznanosti

Znanstveno područje: Interdisciplinarne područje znanosti
Znanstvena polja: Temeljne medicinske znanosti i Biologija

OTKRIVANJE SPECIFIČNIH EPIGENETIČKIH BILJEGA U PAPILOMA VIRUSOM POZITIVNIM TUMORIMA PLOČASTOG EPITELA GLAVE I VRATE

Ksenija Božinović

Disertacija je izrađena u: Laboratoriju za molekularnu virologiju i bakteriologiju Zavoda za molekularnu medicinu Instituta Ruđer Bošković

Mentor 1: dr. sc. Magdalena Grce, znanstvena savjetnica
Mentor 2: dr. sc. Ivan Sabol, znanstveni suradnik

Kratki sažetak doktorske disertacije: Grupa tumora pločastog epitela glave i vrata dijeli se u dvije skupine: tumore pozitivne te negativne na HPV. Cilj ovog istraživanja je pronaći specifične biljege, kako bi se unaprijedile dijagnostičke, prognostičke i terapeutske metode. Istraživanja epigenetičkih promjena u tumorima omogućuju odabir takvih biljega te je stoga napravljena analiza metiliranja cijelog genoma i profiliranje miRNA u svježim uzorcima tumora te analiza preživljenja na svježim i arhivskim tumorskim uzorcima, kako bi se dobila detaljnija slika razvoja bolesti.

Broj stranica: 151
Broj slika: 44
Broj tablica: 14
Broj literaturnih navoda: 301
Jezik izvornika: engleski
Ključne riječi: orofaringealni HNSCC, tumor usne šupljine, HPV, metiliranje, miRNA

Datum obrane: 25. rujna, 2018.
Stručno povjerenstvo za obranu: 1. doc. dr. sc. Vjekoslav Tomaić, znanstveni suradnik
2. prof. dr. sc. Vera Cesar, znanstvena savjetnica
3. prof. dr. sc. Sanja Kapitanović, znanstvena savjetnica
Zamjena: doc.dr.sc. Jelena Knežević, znanstvena suradnica

Rad je pohranjen u: Nacionalnoj i sveučilišnoj knjižnici Zagreb, Ul. Hrvatske bratske zajednice 4, Zagreb; Gradskoj i sveučilišnoj knjižnici Osijek, Europska avenija 24, Osijek; Sveučilištu Josipa Jurja Strossmayera u Osijeku, Trg sv. Trojstva 3, Osijek

ABBREVIATIONS

AJCC - American Joint Committee on Cancer
APS - adenosine 5' phosphosulfate
ASCUS - Atypical Squamous Cells of Undetermined Significance
ATP - adenosine tri-phosphate
CC - Cervical Cancer
CCD - Charge Coupled Device
CDC - Centers for Disease Control and Prevention
CIN - Cervical Intraepithelial Neoplasm
CpG - Cytosine followed by guanine
DNMT - DNA MethylTransferases
E(1-7) - Early region of the HPV genome encoding E1-E7 proteins
ELISA - Enzyme-Linked ImmunoSorbent Assay
ENCODE - ENCyclopedia Of DNA Elements database
FANTOM5 - Functional Annotation of The Mammalian Genome Database
FFPE - Formalin Fixed Paraffin Embedded
HNC - Head and Neck Cancer
HNSCC - Head and Neck Squamous Cell Carcinoma
HSIL - High-grade Squamous Intraepithelial Lesion
HPLC-UV - High Performance Liquid Chromatography-UltraViolet
HPV - Human PapillomaVirus
ICD - International Classification of Diseases (codes)
KEGG - Kyoto Encyclopedia of Genes and Genomes
L (1-2) - Late Region (in HPV genome, encodes L1 and L2 proteins)
LCR - Long Controlling Region (in HPV genome)
LC MS - Liquid Chromatography coupled with tandem Mass Spectrometry
LINE - Long Interspersed Nuclear Elements
LSIL - Low-grade Squamous Intraepithelial Lesion
miR - microRNA; miRNA
MSM - Men having Sex with Men
MS- PCR - Methyl-Specific Polymerase Chain Reaction
n – number
N – normal control
NGS - Next Generation Sequencing
O+ - Oral HPV-positive cancer
O– - Oral HPV-negative cancer
OP+ - Oropharyngeal HPV-positive cancer
OP– - Oropharyngeal HPV-negative cancer
OPSCC - Oropharyngeal Squamous Cell Carcinoma
OSCC - Oral Squamous Cell Carcinoma
PCR - Polymerase Chain Reaction
PV - PapillomaVirus
QC - Quality Control
RT-qPCR - Reverse Transcription Quantitative PCR
PPi - pyrophosphate
pRB - tumor suppressor
p16 ($p16^{INK4a}$) - tumor suppressor
p53 - tumor suppressor
TCGA - The Cancer Genome Atlas
TEX - Tris-HCl -EDTA-SDS
TNM - Tumor-Node-Metastasis stage
URR - Upstream Regulatory Region (in HPV genome)
VLP - Virus Like Particle
WASR - World Age Standardized Rate
WGM – Whole-Genome Methylation
y – year(s)

CONTENT

1 INTRODUCTION

1.1 Human papillomavirus – HPV

Nobel Prize in Medicine 2008 was awarded to prof. Harald zur Hausen for his discovery to human papillomaviruses (HPVs) causing cervical cancer. He reviewed a detailed historical timeline of HPV identification as a separate risk factor in human cancers in 2009 (Harald zur Hausen 2009). Briefly, HPV was first considered as possible carcinogenic in 1934, when studies revealed that rabbit papillomavirus (PV) is capable of transforming rabbit epithelial cells (Rous and Beard 1935), while experiments trying to establish a relationship between HPV infections and cervical cancer were initiated in 1972 (zur Hausen 1977). In 1976, Meisels and Fortin postulated that koilocytes found in cervical smears of patients with lesions represent the cytopathogenic change of a HPV infection (Meisels and Fortin 1976). For a long time, HPV was considered responsible for cancer of anogenital area exclusively, until 1983, when demonstration of HPV antigens in premalignant lesions of the oropharynx provided the first hints for a possible role of HPV infections in oral squamous cell carcinomas (Syrjänen et al. 1983). Thus, the first reports of specific HPV types in tongue and other oropharyngeal carcinomas appeared in 1985 (Löning et al. 1985; de Villiers et al. 1985). All those studies indicated that some members of the PV family are important human carcinogens, which warranted increased studies on HPV role in human malignancies.

HPVs are small non-enveloped DNA viruses with circular double-stranded genome of approximately 8 kb contained in a protein capsid. Viral particles are icosahedral with diameter of approximately 55 nm (**Figure 1**), and the capsid is composed of 72 pentameric capsomers (IARC 2012). Up to date, more than 200 HPV types have been identified, of which around 60 types infect the anogenital tract (Bernard et al. 2010).

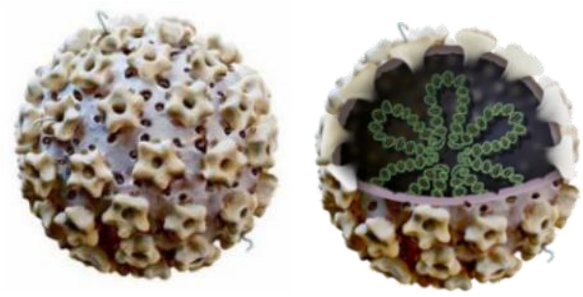

Figure 1. HPV viral components.
Adapted from www.dochandal.com/hpv-oral-cancer/.

All PVs belong to the *Papillomaviridae* family with 16 different genera (**Figure 2**), of which the Alpha genus contains HPVs associated with infection in the mucosal (non-keratinized), while beta and gamma are associated with infection specifically in the cutaneous (keratinized) stratified squamous epithelia (IARC 2007).

HPV types are phylogenetically classified based on pairwise nucleotide sequence identity within the highly conserved L1 gene, and distinct types are defined by differences of at least 10% at the nucleotide level (de Villiers et al. 2004). Based on the ability of transforming infected cells into malignant, HPVs are further classified as (i) carcinogenic to humans (Group 1), (ii) probably carcinogenic to humans (Group 2A), (iii) possibly carcinogenic to humans (Group 2B), (iii) not classifiable as to its carcinogenicity to humans (Group 3), and (iv) probably not carcinogenic to humans (Group 4) (IARC 2018). However, the primary epidemiological classification sorted HPV types into low- and high-risk HPV types (lr- and hr-HPV, respectively) based on how often a particular type was found in cancer, with lr-HPV types found sporadically, and hr-HPV found often in cancers (Jacobs et al. 1995; de Villiers et al. 2004; Bernard et al. 2010). For simplicity, in this study, this classification will be used to differentiate HPV types.

The structure of the HPV genome (**Figure 3**) consists of 3 regions; upstream regulatory region (URR; often called long control region; LCR), the early (E) and the late (L) regions. The E and L regions refer to the phase in the viral life cycle where the genes are usually expressed (Riemer et al. 2010; IARC 2012). The HPV genome, depending on the viral type, encodes approximately eight open-reading frames (ORFs) that are all transcribed from a single DNA strand. The early region encodes proteins E1, E2, E4, E5, E6 and E7, while the late region encodes already mentioned L1 and L2 proteins. The major L1 capsid protein represents 80% of total capsid proteins and is necessary for attachment to the host cell. On the other side, the minor L2 capsid protein is crucial for the efficient transfer of HPV DNA to infected cells. When overexpressed, L1 can form virus-like particles (VLPs) alone, or together with L2 protein. Therefore, these properties of capsid proteins have been used as the basis to create vaccines against HPV infections (Zheng and Baker 2006; Schiffman et al. 2016).

Viral transcription and replication is completely dependent of the host epithelial cell differentiation from the basal to the upper squamous layer (**Figure 4**). The viral genome replication begins in the lower epithelial layers, where cells are in the S-phase but ceases once the cells lose their ability to express the S-phase proteins (Doorbar 2005). The viral cycle may present a productive and a non-productive infection; the first is characterized by the production of virions, while the second only by the expression of viral E6 and E7 oncoproeteins, feature generally found in precancerous and cancer cells (Doorbar 2005).

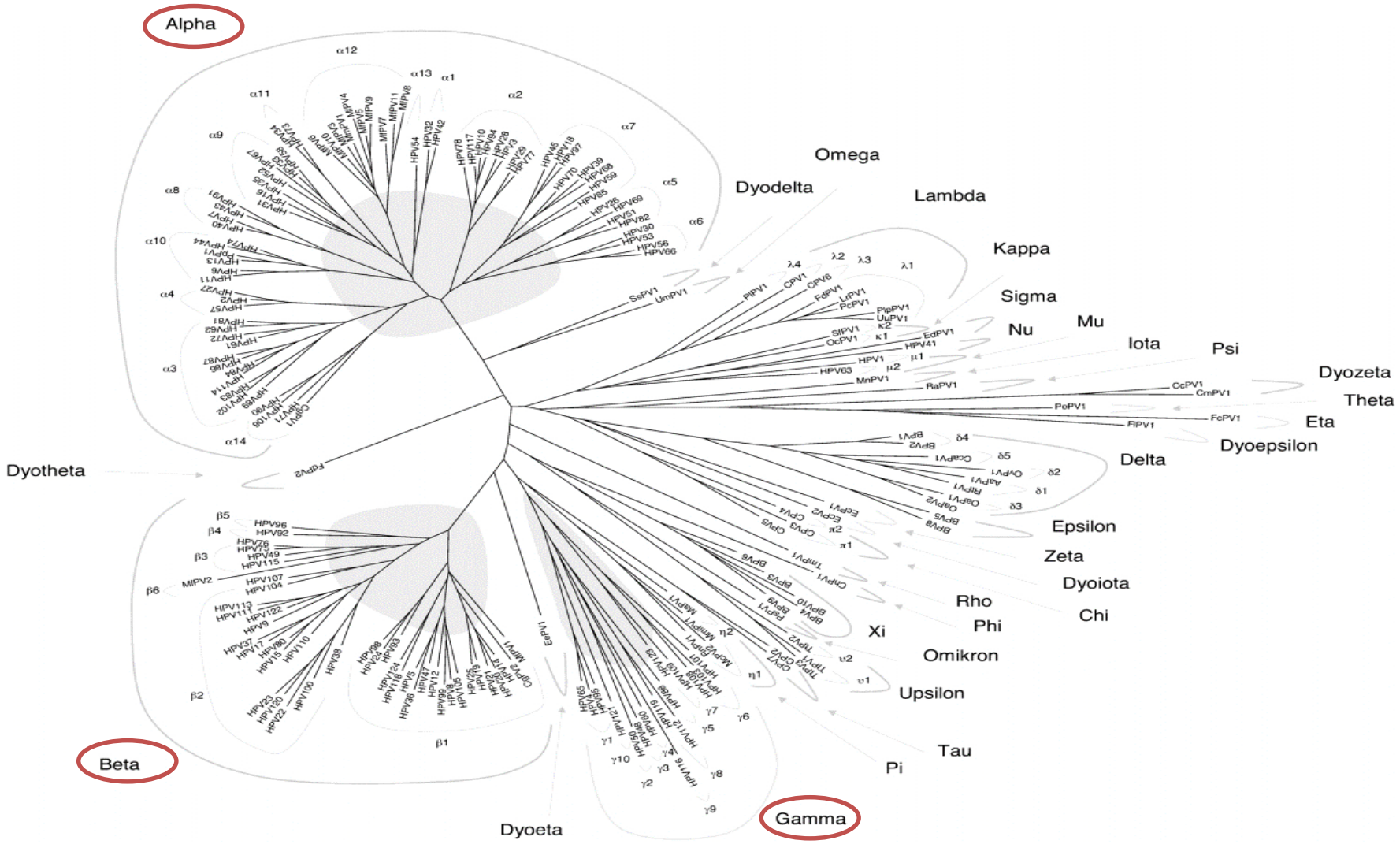

Figure 2. Classification of Papillomaviruses according to genotype and genera. *Adapted from Bernard et al. 2010*.

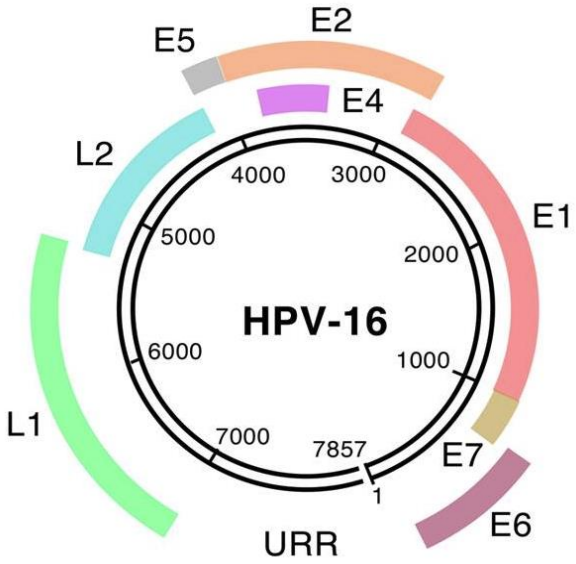

Figure 3. HPV-16 genome. *Adapted from Riemer et al. 2010.*

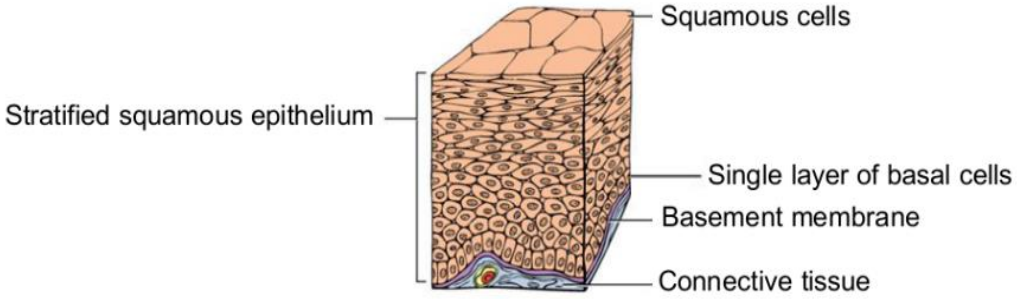

Figure 4. Stratified squamous epithelium. *Adapted from Essentials of Human Anatomy Body Tissues, www.slideplayer.com/slide/5108744/.*

In recent years, there have been a lot of studies investigating HPV proteins (Venuti et al. 2011; Bellanger et al. 2011; Egawa et al. 2012; McBride 2013; Doorbar et al. 2015; de Freitas et al. 2017), which often thoroughly describe their function. Briefly, of all HPV proteins, the most important are certainly the three oncoproteins, E5, E6 and E7, and of the three, E6 and E7 have been investigated thoroughly.

The oncoproteins E6 and E7 cooperate in order to provide an environment suitable for viral

DNA replication, mainly by overcoming cellular apoptotic processes (Yim and Park 2005; Tomaić 2016; Taberna et al. 2017). E6 and E7 bind to two very important tumor suppressors, p53 and pRB, respectively. Binding of E6 to p53 leads to p53 degradation, alteration of cell cycle regulation, apoptosis resistance, and chromosomal instability. Simultaneously, E7 binds to pRB leading to pRB degradation, re-entry into S-phase of the cell cycle, p16 (also known as $p16^{INK4a}$) overexpression, and also chromosomal instability. Finally, the synergistic effect of E6 and E7 leads to cell immortalization. The third HPV oncoprotein, E5 has been linked with the regulation of cell proliferation, growth-factor signaling pathways, as well as immune escape.

The regulation of HPV gene expression is controlled by cellular and viral transcription factors, different promoters, differential splicing, differential transcription termination signals and the stability of different viral mRNAs (Bernard 2002). Most of the regulatory mechanisms are controlled by protein factors that are bound to cis-responsive elements in the URR of the virus. The URRs of most mucosal HPVs range in size from 800 to 900 bp and have a similar organization of cis-responsive elements (IARC 2007).

The E1 protein possess DNA helicase activity, hence its binding to the viral origin of replication is necessary for recruiting the cellular DNA-replication machinery and driving viral replication. The E2 protein possess three major functions: (i) to regulate the expression levels of other viral genes, (ii) to recruit E1 to the viral origin, and (iii) to partition viral genomes into daughter cells during mitosis of infected cell. Although the function of the E4 protein is still obscure, it is the most abundantly expressed viral protein, and has been associated with crucial processes such as aiding viral DNA replication and viral release (IARC 2012).

1.1.1 Epidemiology of HPV infection

HPV infection in humans can cause benign and malignant proliferations (**Figure 5**); depending on the type of virus infecting the skin or oral and genital mucosa. Thus, lr-HPV-6 and -11 (Alpha-PV 10 species) show significantly different tissue tropism, HPV-6 being more common in genital warts, HPV-11 more common in laryngeal papillomas. The hr-HPV16, and related "strains", HPV-31 and -35 (Alpha-PV 9 species) are associated with cervical cancer, although the association between HPV-16 and cancer is significantly stronger than that between either HPV-31 or HPV-35 (Munoz et al. 2003). Moreover, HPV-16 was characterized as HPV type with uniquely powerful carcinogenicity and uniquely associated with tumors of the oropharyngeal region (Munoz et al. 2003; IARC 2007).

As previously pointed, it is known that HPV behavior is strongly correlated with taxonomic categories (IARC 1995, 2007; Chen et al. 2011). Moreover, HPV carcinogenicity has been studied most thoroughly for cervical cancer (CC), and so will be described in detail in this section. HPVs have been found in cervical cancer in 99.7% of cases (Walboomers et al. 1999), as well to low-grade squamous intraepithelial lesions (LSIL) and high-grade squamous

intraepithelial lesions (HSIL), which are precancerous lesions of the cervix (Bosch et al. 2002).

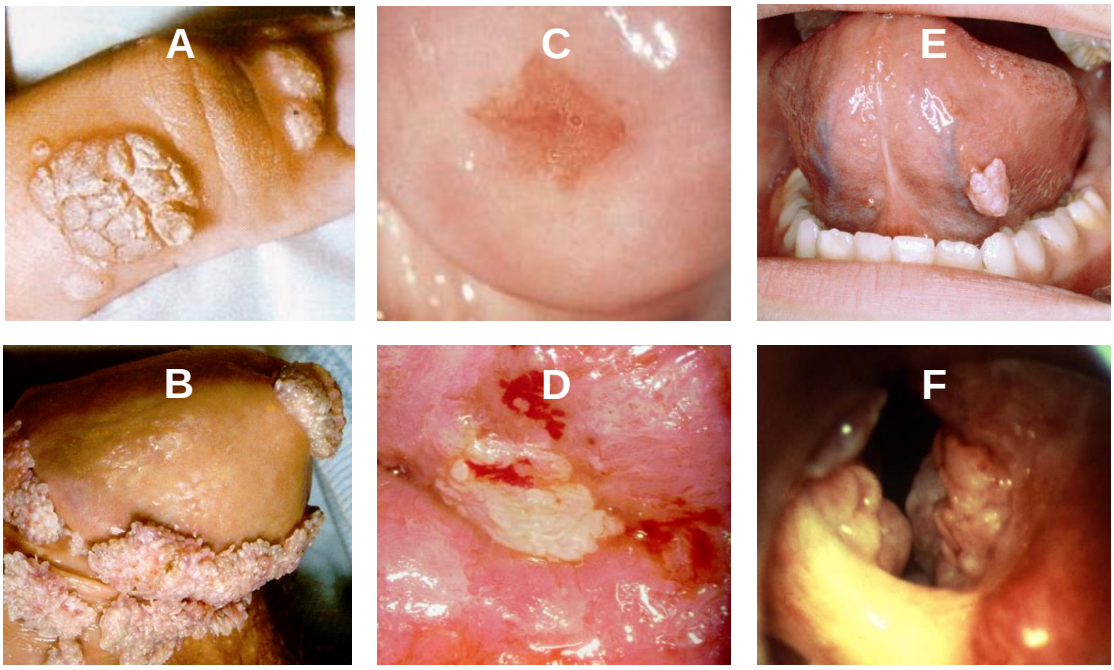

Figure 5. Clinical feature of skin (A, B) and mucosa (C-F) and HPV-associated lesions. Healthy tissue: C) healthy cervix. **Benign proliferation: A)** skin wart located on the finger, **B)** genital wart (*condyloma acuminatum*) located on the penis glans, and **E)** oral papilloma located on the base of tongue. **Malignant proliferation: D)** cancer of the cervix, and **F)** oropharyngeal cancer. (*Pictures provided by the courtesy of Dr Magdalena Grce and Dr Ruth Tachezy*)

The HPV infection is common worldwide (**Figure 6**; Crow 2012), and is mainly transmitted through sexual contact. Most people are infected with HPV shortly after the exposure. According to the WHO data (Cutts et al. 2007), at least half of all sexually active individuals will acquire HPV at some point in their lives, whereas at least 80% of women will acquire an HPV infection by age 50, and some may be repeatedly infected (CDC 2018; Cutts et al. 2007). In 1997, in western countries, such as United States, it was estimated that 10% of the population had an active HPV infection, 4% had an infection that has caused cytological abnormalities, and an additional 1% had infection causing genital warts (Koutsky 1997).

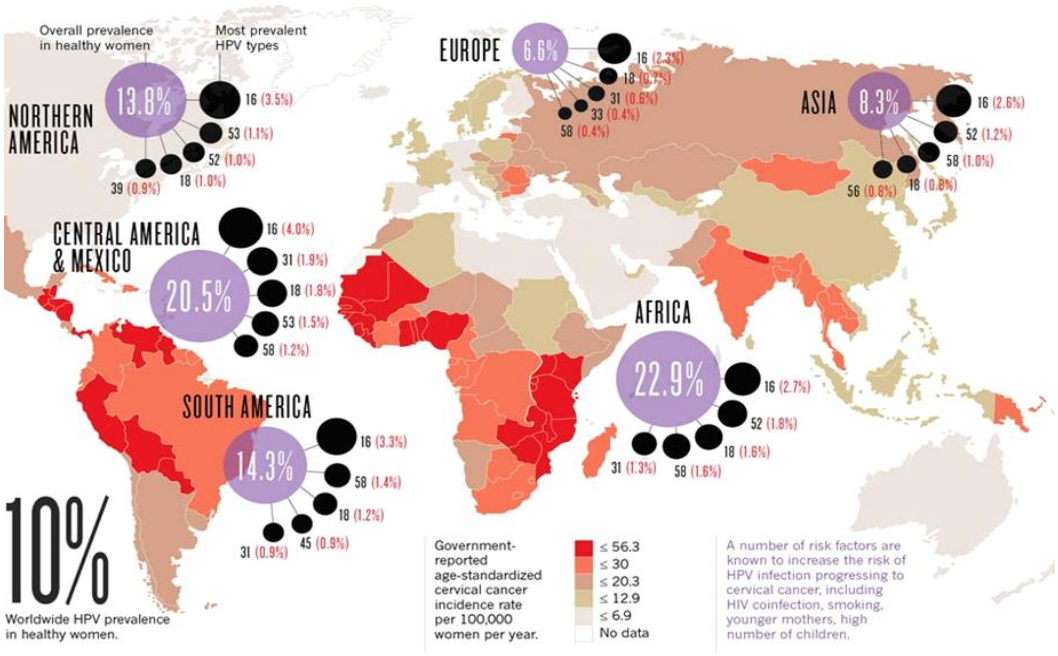

Figure 6. Worldwide HPV distribution by regions. *Adopted from Crow 2012*.

Besides HPV infection in anogenital area, HPV infection is also frequent in head and neck region, mostly in oral and oropharyngeal area (IARC 2012). For instance, in 2014, CDC reports that the prevalence of any oral HPV-type was the highest among Caucasian male adults aged 18–69 years, while the prevalence of any genital hr-HPV for adults aged 18 to 59 years was 45.2% and 25.1% in men, and 39.9% and 20.4% in women, respectively (CDC 2017).

In the epidemiological study on Croatian women (Sabol et al. 2017), 4.432 fresh cervical specimens and 35 archival formalin-fixed, paraffin embedded (FFPE) tissue of cervical cancer specimens were tested for the presence of HPV-DNA. The highest age-specific HPV-prevalence was in the group 18-24 years, which decreased continuously with age regardless of the cytological diagnosis. The prevalence of hr-HPV types significantly increased with the severity of cervical lesions, where HPV-16 was the most commonly type found with a prevalence (with or without another HPV-type) of 6.9% in normal cytology, 15.5% in atypical squamous cells of undetermined significance (ASCUS), 14.4% in LSIL, 33.3% in HSIL, and 60.9% in cervical cancer specimens. Particularly, worldwide HPV type distribution in cervical cancer is shown in **Table 1** (de Sanjose et al. 2010), and highlighted are the most common HPV-types associated with cervical cancer in Croatia.

In the study of Grahovac *et al.* (Grahovac et al. 2012) on 581 men attending the clinic for sexually transmitted diseases (STDs), HPV DNA was detected in 27,4% cases. The study showed that external genitals are more likely to be HPV-positive (25.3%) than the urethra (9.8%), indicating a high risk transmission of oncogenic HPV type between sexual partners.

However, the study of Giuliano *et al.* (Giuliano et al. 2015) shows that women transmit HPV more often to male partners than vice versa.

When the infection is caused by mucosal lr-HPV types, of which HPV-6 and -11 are the most prevalent, there is a great probability for the development of genital warts, mostly on the cervix, vagina, vulva and anus in women, and on the penis, scrotum and anus in men (Lacey, Lowndes, and Shah 2006; Cutts et al. 2007; IARC 2007). They also cause epithelial growths over the vocal cords of children and adults (juvenile respiratory papillomatosis or recurrent respiratory papillomatosis) that require recurrent surgical intervention.

Most HPV infections of the cervix are asymptomatic and more than 90% are cleared within 2 years (Moscicki et al. 2006). Nevertheless, the degree of protection and duration of immunity after the infection, although occurring in a limited number of people, are still not well understood (Beachler et al. 2016). The most prominent problem regarding HPV infection, is that hr-HPVs cause a heavy burden of cancer with more than 600,000 cancers attributed to HPVs worldwide in 2008 of which 236,000 deaths are estimated to be caused by cervical cancer alone each year (GBD 2015).

Infections with lr-HPV genotypes can cause benign genital warts (*condyloma acuminata*) or common warts, depending on the type of the lr-HPV. If the infection occurs with cutaneous lr-HPVs, most notably with HPV-1, -2, -3 and -4, infection can lead to the development of common warts, in general on palms (palmer warts, papillomas) and feet (plantar warts), which are very common among children (Cutts et al. 2007; Egawa and Doorbar 2017).

Table 1. HPV type distribution in cervical cancer by region. *From de Sanjose et al. 2010*

HPV types	Total n = 8977	Europe n = 2058	North America n = 160	Central South America n = 3404	Africa n = 544	Asia n = 2641	Oceania n = 170
HPV-6	10 (<1%)	3 (<1%)	-	3 (<1%)	-	1 (<1%)	3 (2%)
HPV-11	2 (<1%)	-	-	1 (<1%)	-	1 (<1%)	-
HPV-16	**5439 (61%)**	**1348 (66%)**	**115 (72%)**	**2015 (59%)**	**259 (48%)**	**1597 (60%)**	**100 (59%)**
HPV-18	**918 (10%)**	**150 (7%)**	**11 (7%)**	**309 (9%)**	**123 (23%)**	**295 (11%)**	**34 (20%)**
HPV-26	31 (<1%)	3 (<1%)	-	17 (<1%)	-	11 (<1%)	-
HPV-30	31 (<1%)	5 (<1%)	-	14 (<1%)	3 (<1%)	9 (<1%)	-
HPV-31	**335 (4%)**	**69 (3%)**	**5 (3%)**	**166 (5%)**	**10 (2%)**	**80 (3%)**	**1 (<1%)**
HPV-33	**345 (4%)**	**117 (6%)**	**5 (3%)**	**119 (3%)**	**8 (1%)**	**92 (3%)**	**3 (2%)**
HPV-34	6 (<1%)	1 (<1%)	1 (<1%)	3 (<1%)	-	1 (<1%)	-
HPV-35	175 (2%)	46 (2%)	-	72 (2%)	27 (5%)	27 (1%)	4 (2%)
HPV-39	143 (2%)	27 (1%)	2 (1%)	76 (2%)	3 (1%)	31 (1%)	3 (2%)
HPV-39*	3 (<1%)	1 (<1%)	-	1 (<1%)	1 (<1%)	-	-
HPV-42	3 (<1%)	3 (<1%)	-	-	-	-	-
HPV-44	1 (<1%)	1 (<1%)	-	-	-	-	-
HPV-45	**528 (6%)**	**80 (4%)**	**9 (6%)**	**230 (7%)**	**54 (10%)**	**146 (6%)**	**9 (5%)**
HPV-51	114 (1%)	28 (1%)	2 (1%)	53 (2%)	13 (2%)	19 (<1%)	-
HPV-52	**253 (3%)**	**40 (2%)**	**5 (3%)**	**91 (3%)**	**14 (3%)**	**101 (4%)**	**1 (<1%)**
HPV-53	24 (<1%)	10 (<1%)	1 (<1%)	9 (<1%)	-	1 (<1%)	3 (2%)
HPV-56	75 (<1%)	32 (2%)	1 (<1%)	20 (<1%)	4 (<1%)	18 (<1%)	-
HPV-58	**203 (2%)**	**27 (1%)**	**3 (2%)**	**67 (2%)**	**4 (<1%)**	**102 (4%)**	**-**
HPV-59	95 (1%)	15 (<1%)	-	42 (1%)	1 (<1%)	36 (1%)	-
HPV-61	1 (<1%)	-	-	1 (<1%)	-	-	-
HPV-66	7 (<1%)	2 (<1%)	-	2 (<1%)	2 (<1%)	1 (<1%)	-
HPV-67	26 (<1%)	3 (<1%)	-	13 (<1%)	-	10 (<1%)	-
HPV-68	58 (<1%)	13 (<1%)	-	20 (<1%)	1 (<1%)	25 (<1%)	-
HPV-68**	31 (<1%)	4 (<1%)	-	17 (<1%)	1 (<1%)	3 (<1%)	6 (4%)
HPV-69	7 (<1%)	-	-	6 (<1%)	1 (<1%)	-	-
HPV-70	9 (<1%)	1	-	3 (<1%)	-	5 (<1%)	-
HPV-73	43 (<1%)	16 (<1%)	-	14 (<1%)	1 (<1%)	12 (<1%)	-
HPV-74***	2 (<1%)	-	-	-	-	1 (<1%)	-
HPV-82	6 (<1%)	-	-	4 (<1%)	-	2 (<1%)	-
HPV-91	1 (<1%)	1	-	-	-	-	-
HPV-X[#]	52 (<1%)	10 (<1%)	0	15 (<1%)	13 (2%)	12 (<1%)	2 (1%)

N (%) data are based on the upper estimate attribution of multiple HPV types. *HPV type 39, 68, or 73. **HPV type 68 or 73. ***One case in the total attributable to infection with multiple HPV types, and no case with exclusively HPV 74. [#]HPV undetermined

1.1.2 HPV life cycle

HPVs are highly epitheliotropic, meaning they establish productive infections only within stratified epithelia of the skin, the anogenital tract and head and neck area (de Villiers et al. 2004; Conway and Meyers 2009). Many hr-HPV types have been found in cancer, while lr-HPVs are found rarely, or not at all (Doorbar et al. 2015; Egawa and Doorbar 2017). HPV tends to escape immune response in order to maintain in the host cell and complete the life cycle. As long as HPV continues to replicate its genome in low copy number, oncogenes E6 and E7 will express, and maintain suitable environment for the continuous cell replication. Once HPV starts to over-express only E6 and E7, HPV life cycle gets terminated, while overly expressed oncoproteins E6 and E7 become enough for the development of cancer *in situ* and maintaining of the cancer. If cancer cells cross the basement membrane, cancer becomes metastatic (Bissell and Radisky 2001; NCI 2018). The viral life cycle is linked to the differentiation of the infected stratified epithelia (**Figure 7**) and is thought to be initiated by the infection of basal epithelial cells, preferably at the site of epithelial injury (Doorbar et al. 2012).

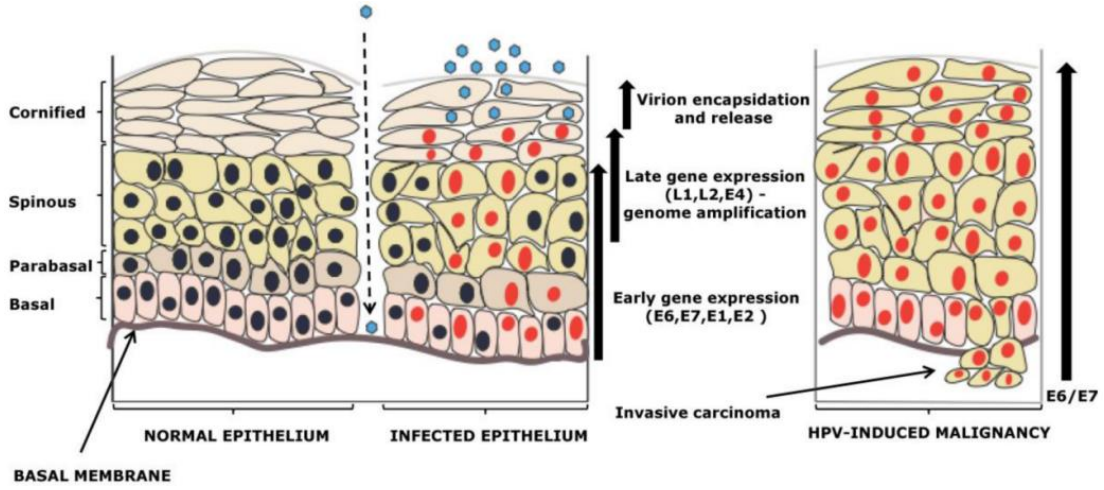

Figure 7. HPV life cycle and cancer development. *Adopted from* Tomaić 2016*.*

Basal cells are in fact multipotent stem cells that have high proliferative potential, enabling HPV to successfully replicate and establish its genome in a low copy number (Kajitani et al. 2012; Moody 2017). The ability of HPVs to establish their genome in basal cells depends on the E1 (Egawa et al. 2012), E2 (Longworth and Laimins 2004), E6 (Münger et al. 2004) and in some cases E7 (Thomas et al. 1999) proteins. In normal conditions, when basal cells undergo division, the daughter cell loses contact with the basement membrane and migrates into the suprabasal compartment, withdraws from the cell cycle and initiates differentiation process (**Figure 7**). However, in HPV-positive human keratinocytes and cervical epithelial cells, the

suprabasal cells continue to express markers for cell proliferation (Tomaić 2016). Moreover, HPV-16 E7 oncoprotein has been shown to be necessary and sufficient to induce suprabasal DNA synthesis; this is possible due to degradation of pRB via E7, while E6-mediated p53 degradation prevents cells to enter apoptosis (Tomaić 2016). In addition, the E5 oncoprotein contributes quantitatively to this property, while also contributing to immune evasion (Venuti et al. 2011; IARC 2012; de Freitas et al. 2017). Within this suprabasal compartment, cells support the amplification of the viral genome, expression of capsid genes and assembly of progeny virus (IARC 2012; Tommasino 2014; Tomaić 2016). The E4 protein, which is detected preferentially in the differentiated compartment of infected tissue, is required for viral DNA amplification and expression of the L1 capsid gene (Doorbar 2013). Finally, encapsidation of HPV DNA within capsids to generate progeny viruses within the terminally differentiated cell compartment is quantitatively dependent on L2 protein (Holmgren et al. 2005). The L2 protein is also required for the infectivity of HPV-16 (Yang et al. 2003) and HPV-31 (Holmgren et al. 2005) virions. It is worth mentioning that HPV can establish its genome and finish life cycle successfully only in cases when it exists in the nucleus as an episome. Once HPV genome integrates into the host genome, it loses parts of its genomic material, and often it is unable to express all proteins necessary for propagation. The only two proteins that continue to aberrantly express are E6 and E7 oncoproteins. Therefore, viral life cycle is replication defective, but E6 and E7 continue to drive and maintain transformation of the cell (**Figure 8**), which finally leads to cancer development (Grce, Sabol, and Milutin Gašperov 2012b; Hu et al. 2015; Morgan, DiNardo, and Windle 2017). Manipulation of E6 and E7 on both transcriptional and translational levels finally lead to senescence and/or apoptosis of the infected cells (Tomaić 2016). The integration process and the differences between cancer with integrated and episomal HPV genome will be described in more details in the following chapter.

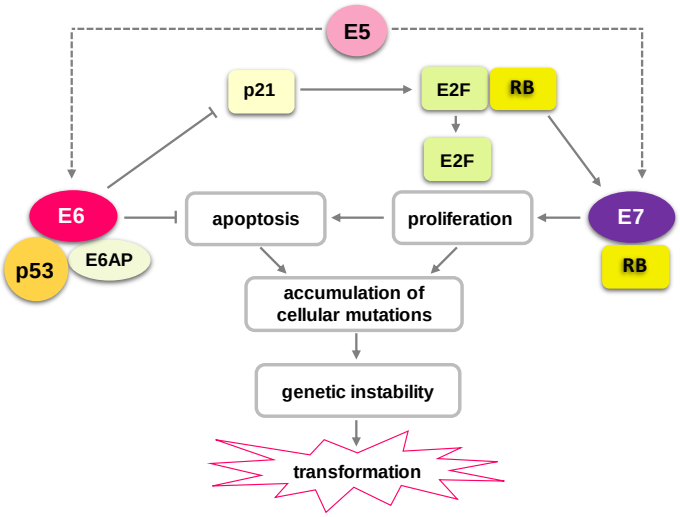

Figure 8. Cell transformation via HPV oncoproteins E6, E7 and E5.

Adopted from Grce, Sabol, and Milutin Gašperov 2012b

1.1.3 Malignancies associated with HPV infection

Table 2 shows the burden of hr-HPV related diseases in Croatia (Grce, Sabol, and Milutin Gašperov 2012a). It is obvious that cervical cancer is the most important HPV-related disease although other HPV-related cancers and benign proliferation (genital warts and laryngeal papillomatosis) represent a significant public health burden. Highlighted is HPV-16, which accounts for the majority of HPV-associated malignancies.

Table 2. Burden of hr-HPV related diseases. *Adapted from Grce, Sabol, and Milutin Gašperov 2012a*.

Disease (ICD)	World annual estimates in 2008	Croatian annual estimate in 2008	HPV contribution	Most common HPV types
Cervical cancer (C53)	529,000 cases (WASR 15.2 per 100,000) 274,000 deaths (WASR 7.8 per 100,000)	360 cases (15.6 per 100,000)	85%	**HPV-16 (61%)** HPV-18 (10%) HPV-31 (4%) HPV-33 (4%) HPV-35 (2%) HPV-45 (6%) HPV-52(3%) HPV-58 (2%)
SCC	~ 90% of all C53 cases		87%	**HPV-16 (62%)** HPV-18 (8%) HPV-45 (5%)
Adenocarcinoma	~ 10% of all C53 cases		62%	**HPV-16 (50%)** HPV-18 (32%) HPV-45 (12%)
Vulvar cancer (C51)	30,000 cases	61 cases (2.7 per 100,000)	40%	**HPV-16 (32%)** HPV-18 (4%)
Vaginal cancer (C52)	15,000 cases	16 cases (0.7 per 100,000)	70%	**HPV-16 (54%)** HPV-18 (8%)
Anal cancer (C21)	30,400 cases (15,900 women & 14,500 men)	19 cases (10 women & 9 men; 0.4 per 100,000)	97%	**HPV-16 (75%)** HPV-18 (3%)
Penile cancer (C60)	26,300 cases	28 cases (1.3 per 100,000)	45%	**HPV-16 (60%)** HPV-18 (13% HPV-6/11 (8%)
Oropharyngeal cancers (C01, C09, C10)	61,500 cases (12,600 women & 48,900 men)	147 cases (127 men & 20 women) (1.3 per 100,000; 2.1 & 0.6 per 100,000 in men & women, respectively)	47% in OP carcinomas and 11% in oral cavity carcinomas; 64% in female and 42% in male cases	**HPV-16 (90%)** in OP carcinomas; HPV-16 (96%) in oral carcinomas

ICD diagnosis code (NIH 2018): C01 base of tongue, C09 tonsils, C10 oropharynx, C21 anus anal canal, C51 vulva, C52 vagina, C53 cervix, C60 penis; WASR = world age standardized rate; SCC = squamous cell carcinoma; OP = oropharyngeal

All HPV genotypes that are known to be involved in cervical cancer development belong to the alpha genus. HPV-16 and HPV-18 have been classified as cervical carcinogens since 1995 (IARC 1995). HPV-31 and HPV-33 were categorized as probably carcinogenic, and in 2005, the group of cervical carcinogens was expanded to include 13 hr-HPV types: HPV-16, -18, -31, -33, -35, -39, -45, -51, -52, -56, -58, -59 and HPV-68 (IARC 2007; Bouvard et al. 2009). The main epidemiological criteria used for the classification of an HPV type as a carcinogen is finding the HPV genotype as a single infection in a cervical scrape or biopsy specimen in a woman with cervical cancer. Unfortunately, this might sometimes be misleading, since colposcopic biopsies can fail to sample the critical cells. Furthermore, the contamination of scrapes and biopsies from precancerous stages of cervical cancer, lower-grade lesions that often surround cancers, can lead to the detection of types other than the causal one (IARC 2012).

HPV-16 is by far the most prevalent carcinogenic HPV type (**Table 2**), associated with approximately 50% of all cervical cancers, the majority of other HPV-related anogenital cancers (60%), and for more than 80% of HPV-associated head and neck squamous cell carcinoma (HNSCC) (de Sanjose et al. 2010; Ndiaye et al. 2013; Serrano et al. 2015; Taylor et al. 2016). Recent advances in high-throughput next-generation sequencing have led to new discoveries in HPV genomic research (Cullen et al. 2015; K. Liu et al. 2017). These findings provide opportunities for further investigation at the intersection of molecular biology and epidemiology that could enhance our molecular understanding of HPV-related carcinogenesis. It is now well known that persistent HPV infection is highly necessary in order to establish malignant transformation in cervical epithelium and it is considered as the main risk factor in HPV-associated cancerogenesis (IARC 2012). In the context of HPV-associated cervical cancer, malignancy is established in the following ways: first the progressive histopathological changes of cervical epithelium include the loss of terminal differentiation, which leads to a cellular state that cannot support the full viral life cycle, and second, the circular HPV genome, which normally resides in nucleus as an episome, often becomes integrated into the host genome, which becomes disrupted and its replication defective (IARC 2012; Tomaić 2016). Analysis from The Cancer Genome Atlas study shows that HPV integration occurs in >80% of HPV-positive cervical cancers (TCGA 2017). Of these, 76% of HPV-16 positive samples have integrated HPV, whereas integration is evident in all HPV-18 positive samples. In HPV-positive oropharyngeal SCC, the incidence of viral integration is lower, and many tumors have either episomal or mixed episomal and integrated viral DNA (Matovina et al. 2009; Parfenov et al. 2014; Vojtechova et al. 2015). The rate of HPV integration in other anogenital cancers is not as well documented, but one study reports that almost 80% of anal carcinomas contain integrated HPV; however, the vast majority of these samples also contained episomal genomes (Valmary-Degano et al. 2013). Whether the virus itself drives this integration event or whether it reflects random recombination events is still not completely clear (IARC 2012; Liu et al. 2015; Holmes et al. 2016). Nevertheless, the two

consequences of the HPV integration are supporting malignant transformation via constitutive upregulation of E6 and E7 and it creates selective growth advantage over cells that harbor the viral genome as an episome (Herdman et al. 2006).

Integration events that are found in cervical cancer lead to the selective overexpression of E6 and E7 (Jeon, Allen-Hoffmann, and Lambert 1995; Jeon and Lambert 1995), which is a hallmark of cervical cancers. In a recent review by Oyervides-Muñoz *et al.* (Oyervides-Muñoz et al. 2018), new oncogenic mechanisms have been explained, suggesting that there are evidence that HPV integration is not random, and that it often affects preferably those genes that are being constantly expressed during DNA transcription, DNA repair factors and transcription regulation factors, to establish carcinogenesis (Liu et al 2016). Furthermore, they emphasized that the progression to cancer could be explained by the viral DNA integration into tumor suppressor genes, which inactivates those genes and it would lead to an uncontrolled growth (Zhao et al. 2016). Evidence suggests that HPV has specific integration points, in or close to fragile sites or hot spots, where it could alter expression patterns of the affected genes (Matovina et al. 2009; Liu et al. 2015). Finally, HPV could be involved in transcription mechanisms or hyperactive epigenetic spots, which could facilitate its integration in those active genes. Therefore, HPV integration is proposed as a driver mutation for cervical cancer progression, while overexpression of E6/E7 certainly contribute to the induction of cell malignancy.

In addition, cervical cancer typically follows age of infection by decades (Schiffman and Wentzensen 2013). HPV DNA and RNA transmitted at young ages usually become undetectable, with no sensitive serological assay existing to measure HPV exposure (Gravitt and Winer 2017). About 10–30% of women with detectable HPV DNA exhibit definite cytological abnormalities, depending on the HPV type, and DNA test (Kovacic et al. 2006). Most HPV infections clear within 1–2 years, however, estimates of duration of infection for individual HPV types vary from study to study, and depend not only on the statistical methods used, but also on the accuracy of the HPV DNA detection methods (IARC 2012). Persistent HPV infection is a prerequisite for the development of high-grade precancerous lesions (cervical intraepithelial neoplasia grade 3, CIN3) and cervical cancer, but for epidemiological purposes there is no consensus on the definition of persistent infection (IARC 2012; Hosaka et al. 2013; Sudenga and Shrestha 2013). In some cases CIN3 can develop very quickly, within 2–3 years following the HPV exposure, especially in young women which reproductive tract is still immature (Saslow et al. 2012). The most abundant carcinogenic genotypes, HPV-16 and HPV-18 are more common among cervical squamous carcinomas than cytologically normal women or even in LSIL, with HPV-18 being more common in adenocarcinomas (Bulk et al. 2006; Lin, Franceschi, and Clifford 2018). The less abundant, but still with high carcinogenic potential, are six additional types HPV-45, -31, -33, -35, - 52, and -58, which have some regional variation in the etiology of cervical cancer (de Sanjose et al. 2010). For example, HPV-52 and -58 are relatively more prevalent in Asia than

in other regions, HPV-33 is most clearly prevalent in Europe, while HPV-45 in Africa (**Table 1**).

Regarding cervical cancer, it is apparent that trends in Croatia follow trends in most European countries (**Figure 9**); the world age-standardized incidence and mortality rates (per 100,000 women-years) of cervical cancer in Croatia from 1968 to 2014 have been slowly declining (CNCR 2016).

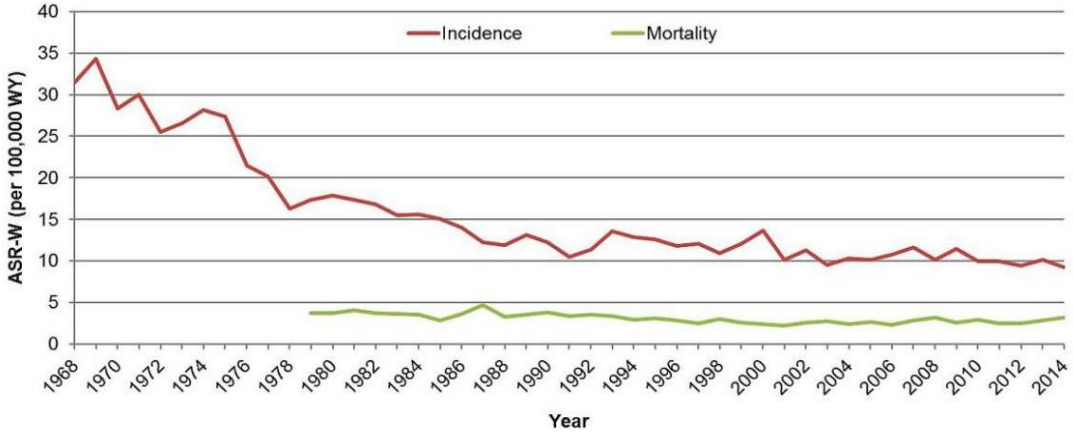

Figure 9. World age-standardized incidence and mortality rates (per 100,000 women-years) of cervical cancer in Croatia. *Adopted from CNCR 2016.*

In our recent epidemiological study on genital HPV infection (Sabol et al. 2017) the highest age-specific HPV-prevalence in Croatia was in the group 18–24 years, which decreased continuously with age regardless of the cytological diagnosis. The prevalence of hr-HPV types significantly increased with the severity of cervical lesions, while HPV-16 was the most common type found with a prevalence of 6.9% in normal cytology, 15.5% in atypical squamous cells of undetermined significance, 14.4% in LSIL, 33.3% in HSIL, and 60.9% in cervical cancer specimens. Moreover, the most common types associated with cervical cancer after the HPV-16 (60.9%) were found to be HPV-33 and -45 (10.9% in both cases), followed by HPV-18 (8.7%), -31 and -58, which were found in 2.2% of cervical cancers.

Besides cervical cancer, HPV has been associated with 80% of other mucosal anogenital and 50-60% of head and neck cancers (Cutts et al. 2007; Taberna et al. 2017), with HPV-16 accounting for the majority of cases (**Table 2**). The worldwide age-standardized incidence rates (per 100,000) of cervical and other anogenital cancer that are attributed to HPV in 2012 are presented in **Figure 10** (Globocan 2012).

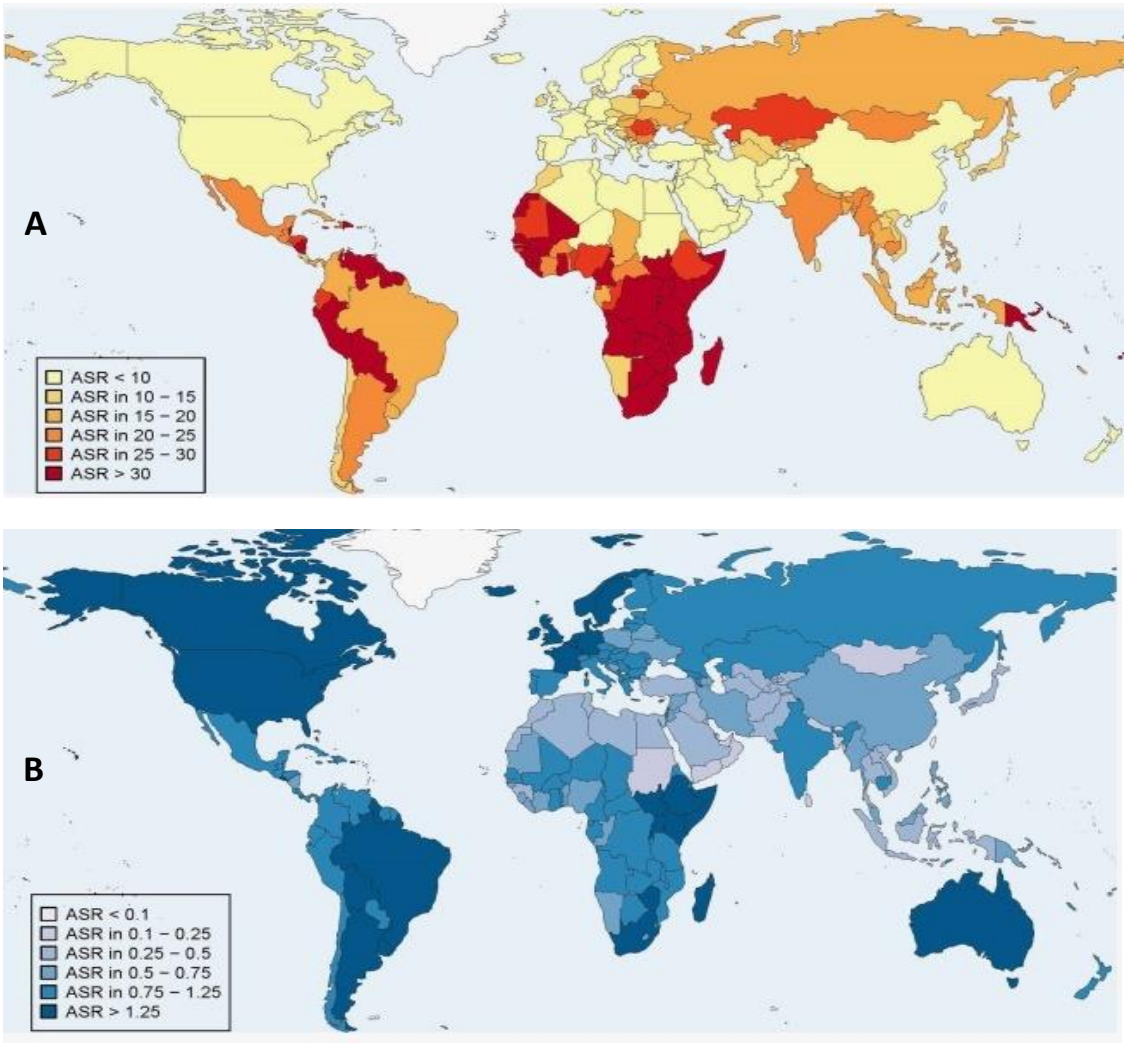

Figure 10. Worldwide age-standardized incidence rates of anogenital cancer attributable to HPV in 2012. A) Cervical cancer, women. B) Anogenital (vulvar, vaginal, anal and penile) cancer, both genders. *Adopted from Globocan 2012.*

Studies by de Martel *et al.*, IARC, and the HPV Centre give a detailed epidemiological overview of HPV related malignancies, and in the following paragraphs key findings are summarized (de Martel et al. 2012; IARC 2012; HPV Centre 2017).

Briefly, cancer of the vulva is rare among women worldwide, while about 60% of all cases occur in more developed countries. One third of tumors is more common in younger women, and often associated with HPV DNA detection. These tumors appear to share the epidemiological factors of cervical cancer. Cancer of the vagina is also a rare cancer, and can be of different cell origin: epithelial, mesenchymal or lymphoid. Epithelial tumors are the most common and strongly related to HPV, with HPV-16 being associated in more than 50% of the cases.

Cancer of the anus is a relatively rare disease in the general population, but increasing

incidence is being reported in men and women in western countries over a period of 20 years (van der Zee et al. 2013). Women have higher incidences of anal cancer than men. Penile cancer incidence rates strongly correlate with those of cervical cancer. It is rare and most commonly affects men aged 50-70 years, while incidence rates are higher in less developed countries, accounting for up to 10% of male cancers in some parts of Africa, South America and Asia.

Besides anogenital cancers, the fraction of head and neck cancers, HNSCC associated with HPV infection specifically, varies between different studies according to (i) the accuracy in the determining anatomical site of tumor origin, (ii) the competing effect of tobacco smoking and alcohol intake, and (iii) the quality of tissue biopsies and HPV-testing protocols used (IARC 2012). As head and neck cancer is the main model for this study, this type of HPV related malignancy will be presented in greater detail in the next section.

Moreover, a number of cutaneous HPVs have been discovered in skin cancers in patients who have epidermodysplasia verruciformis (EV), such as HPV-5 and HPV-8, and these types are also found in both non-melanoma skin cancers and normal skin (IARC 2007). Moreover, both HPV-5 and HPV-8 are considered as cofactors for HPV-induced cancer of the skin, but only in the presence of ultraviolet (UV) damage (IARC 2007). In addition, *in vitro* analyses have shown that HPV-38 E6 bind to p53, similar to mucosal hr-HPV types, but in contrast to those types, they appear to stabilize p53 and not degrade it (Tomaić 2016). This could indicate that cutaneous HPV type might also be classified as lr-HPV and hr-HPV types, but prior to that, more studies are necessary for the better understanding of the viral mechanism. In addition, Human Papillomavirus and Related Diseases Report from 2017 (HPV Centre 2017) provides key information on HPV-related statistics for Croatia on: cervical cancer, other anogenital cancers and head and neck cancers. According to the report, the incidence rates (per 100,000) of HPV related malignancies in Croatia (2003-2007) was as follows: 7.3 for squamous cervical carcinoma, 0.3 for both male and female anal cancer, 1.2 for vulvar cancer, 0.3 for vaginal cancer, 0.7 for penile cancer, 6 for male head and neck cancer and 0.8 for female head and neck cancer.

In conclusion, HPV-associated cancers represent a higher burden in less developed countries than in developed countries, where they have successfully implemented primary cancer prevention measures through HPV vaccination, and secondary cancer prevention measures through population-based screening (mostly cervical cytology). Those preventive measures significantly decreased the incidence of cervical cancer in many developed countries (IARC 2005). HPV testing as a primary screening test followed by cytology triage have been considered and implemented in several western countries, although HPV testing has been widely used as a secondary test for triage of borderline cytology and as a follow-up test after treatment of severe cervical lesions, which is an addition to conventional cytological screening (Grce and Davies 2008; Grce 2009). However, in HNSCC the situation is quite different because no effective screening strategy has been identified (Dixit et al. 2015). There

is also compelling evidence of retrospective studies suggesting that HPV-related head and neck cancers (HNCs) are more frequently cured than those caused by tobacco (Vatca et al. 2014). Furthermore, using FFPE tumor samples in retrospective studies creates the possibility of increasing sample size available for studies and including samples during a longer time period.

1.2 Head and neck cancer

HNC is the sixth most common malignancy worldwide, predominantly arising within the mucosal linings of the upper aerodigestive tract (Argiris et al. 2008), which holds important functional roles for respiration, speech and swallowing. Most HNC develop from squamous cell epithelia (HNSCC), which accounts for 95% of head and neck carcinoma (Vigneswaran and Williams 2014). HNSCC often gets diagnosed in the late phase, when it is difficult to treat the tumor, while both surgical and non-surgical treatment modalities achieve a cure for only 50% of HNSCC patients within a 5-year period (Fakhry et al. 2014; Leemans, Snijders, and Brakenhoff 2018). The main risk factors for HNSCC development are smoking and excessive alcohol consumption. In recent years, however, the role of HPV has emerged, particularly in oropharyngeal cancer (Albers et al. 2017; Taberna et al. 2017). HNSCCs are characterized according to their primary site of origin, with most common sites being the oral cavity, the oropharynx, the hypopharynx, the larynx, and the sinonasal tract (Yan et al. 2010; R. Li, Agrawal, and Fakhry 2015). The identification of the accurate site of tumor origin is of high importance, as misclassification can have serious implications on not only the prognosis, but also the choice of treatment modality and subsequently survival (Li, Agrawal, and Fakhry 2015).

In 2015, The Cancer Genome Atlas (TCGA) consortium published a comprehensive molecular catalogue on HNSCC (TCGA 2015). The TCGA study revealed that HNSCC lacked predominant gain-of-function mutations in oncogenes, whereas an essential role for epigenetics in oncogenesis has become apparent. Furthermore, study by Masuda et al. (2016) emphasizes that HNSCC seems to be an epigenetic disease (Masuda, Wakasaki, and Toh 2016), rather than genetic. Moreover, it is known that the immune system plays a key role in the development, establishment and progression of HNSCC (Ferris 2015). So, the HNSCC is considered as an immunosupressive disease, since HNSCC evade the immune system through different mechanisms. Ferris (2015) reports some of these mechanisms, such as manipulation of tumor immunogenicity, production of immunosupressive mediators, and promotion of immunomodulatory cell types (Ferris 2015). Also, it seems like through these mechanisms, HNSCC can influence the microenvironment and exploit immune system to enable angiogenesis, metastasis and tumor growth. Interestingly, HPV-positive and HPV-negative HNSCC use different mechanisms in the immune escape. Certainly, a greater understanding on mechanisms of immune escape in HNSCC provide the basis for the improvement of therapy and patients outcome.

1.2.1 Epidemiology of head and neck cancer

Globally, HNC accounts for approximately 550,000 cases per year (Keck et al. 2015). The worldwide age-standardized incidence rates (per 100,000) of head and neck cancer

(oropharynx, oral cavity and larynx) that are attributed to HPV in 2012 are presented in **Figure 11** (Globocan 2012). In Croatia, 896 new cases of HNC were estimated in 2015 (CNCR 2016), while oropharyngeal cancer (C01, C09, C10) alone was estimated to 1.3 per 100,000, of which 2.1 per 100,000 was in men and 0.6 per 100,000 was in women (**Table 2**) (Grce, Sabol, and Milutin Gašperov 2012a). Smoking and alcohol consumption have been presumed as the main risk factors for developing oral and oropharyngeal tumors (Argiris et al. 2008; Keck et al. 2015). Recent estimates of tobacco usage by the Centers for Disease Control (CDC 2011) reported that 45.3 million adults in the USA are currently cigarette smokers, comprising 19.3% of the adult population (CDC 2011), while in Croatia, the percentage gets up to 30% of the population (Padjen et al. 2012). However, even though HPV has for a long time been considered responsible for exclusively anogenital cancers, for the first time in 1983, the presence of hr-HPV was demonstrated in HNCs (Syrjänen et al. 1983). In the late '90s of the 20th century, investigating the role of HPV in carcinogenesis became increasingly important (Taberna et al. 2017).

In western countries, tobacco and alcohol consumption-induced HNCs, such as cancer of the oral cavity, hypopharynx and larynx are declining for the past twenty years, mainly due to reduction in smoking. In contrary, tongue and oropharyngeal cancers are becoming more prevalent, especially in younger individuals (Albers et al. 2017; Taberna et al. 2017), which is in line with the increasing incidence of the hr-HPV in oral/oropharyngeal sites (Keck et al. 2015).

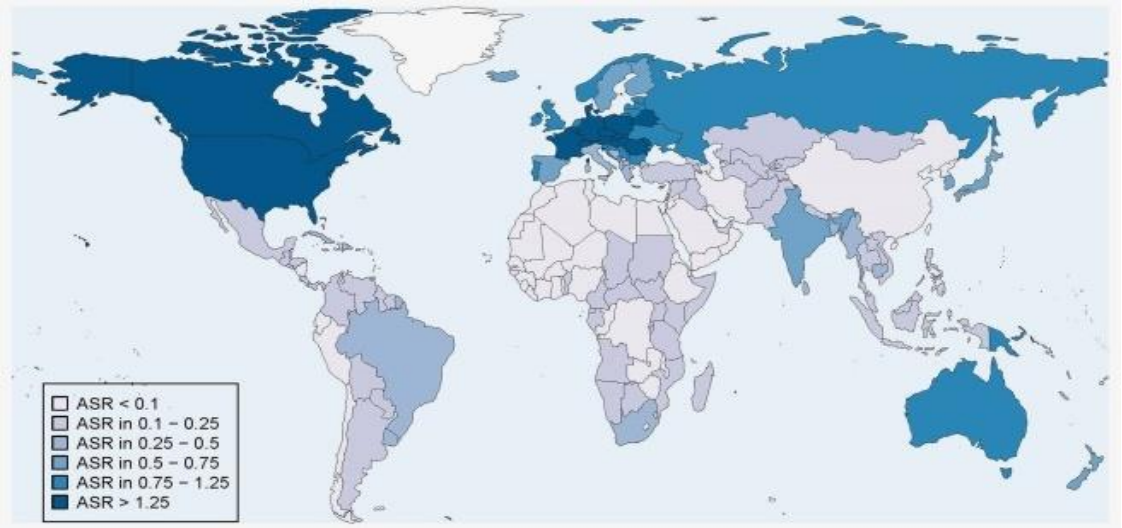

Figure 11. Worldwide age-standardized incidence rates of head and neck cancer (oropharynx, oral cavity and larynx) attributable to HPV in 2012, both genders.
Adopted from Globocan 2012.

1.2.2 Clinical classification of head and neck squamous cell carcinoma – HNSCC

HNSCC are classified according to the anatomical site of origin, TNM staging and grade of the disease. Anatomical site classifies HNSCC to tumors of oral cavity, pharynx (nasopharynx, oropharynx, hypopharynx) and larynx (IARC 2012).

The nasopharynx, according to ICD diagnosis (NIH 2018) code C11.9, is the extension to the oropharynx (C10.9). The border between these two anatomical spaces is a horizontal plane drawn through the soft palate.

There are eight sites of the oral cavity, which are potential sites for the origin of HNSCC (IARC 2012) and include the following: the upper (C00.0) and the lower (C00.1) vermillion lips; the buccal mucosa (C06.0); the lower alveolar ridge (C03.1); the upper alveolar ridge (C03.0); the retromolar trigone mucosa (C06.2); the floor of mouth mucosa (C04.9), and also lining the undersurface of the tongue (C02.1); the hard palate mucosa (C05.0); and the oral tongue (C02.9) (Li, Agrawal, and Fakhry 2015). Taking into account the oral cavity regions is critical when assigning a tumor to an oral cavity or oropharyngeal origin, as the posterior limit of many oral cavity regions continues with the oropharyngeal site.

The seven primary sites of the oropharynx include: the palatine tonsils (ICD C09.0); the base of tongue (C01.9); the oral surface of the soft palate and uvula (C05.2); the posterior pharyngeal wall (C10.3); the lateral pharyngeal walls (C10.2); the mucosa of the anterior and posterior tonsillar pillars (C09.1); and the glossotonsillar sulcus (C09.9). HPV-associated cancer of the oropharynx originates mostly in the lymphoid tissue of the tonsils and base of tongue (Li, Agrawal, and Fakhry 2015; Taberna et al. 2017).

Four sites of the hypopharynx include: the pyriform sinuses (C12.9); the lateral pharyngeal walls (C13.8); the posterior pharyngeal wall (C13.2); and the postcricoid region (C13.0) (IARC 2012; Li, Agrawal, and Fakhry 2015).

The larynx is a complex structure that can be broadly divided into glottic (C32.0), supraglottic (C32.1) and subglottic (C32.2) regions. Tumors of the base of tongue may invade the laryngeal region, potentially rendering anatomical classification of tumor origins inaccurate (Li, Agrawal, and Fakhry 2015).

According to the HPV Information Centre report (HPV Centre 2017) and the study of the Croatian HNC survey (Znaor et al. 2013), the most prevalent cancer type in Croatia was pharyngeal cancer, including C09-10 and C12-14 with the age-standardized incidence rate (per 100,000 patients) estimated to 6.0 for male and 0.8 for women. The age-standardized mortality rate (per 100,000 patients) for men was 4.7, and 0.3 for women. For pharyngeal cancer incidence in Southern Europe, rates (per 100,000 patients) were as follows: 3.4 for men and 0.5 for women, while the mortality rate for men was 1.8 and 0.3 for women. In addition, the incidence rate (per 100,000 patients) for the base of tongue, tonsil and other oropharyngeal sites for males in Croatia were 1.7, 1.8, 1.4, and for women 0.2, 0.2 and 0.1,

respectively.

Besides anatomical site, the most commonly used model for the classification of the severity of HNC malignancy is the TNM staging model. This model assigns a numerical status regarding tumor size and location (T1, -2, -3, -4), degree of lymph node involvement (N0, -1, -2, -3), and the presence (M-1) or absence of distant metastasis (M0). The 8th Edition of the American Joint Committee on Cancer (AJCC) staging manual includes the major changes in staging of the oropharyngeal cancer (AJCC 2018), which is in contrast to the 7th Edition of the AJCC guideline, where staging of tumor is more severe. The study of Cramer et al. (2018) evaluated the new staging system in order to validate this shift in classification (Cramer et al. 2018); they concluded that the 8th Edition of the AJCC guideline have profoundly improved staging of HPV-positive oropharyngeal cancer, and seem to demonstrate survival discrimination. In our study, we have also evaluated survival analysis based on both the 7th and 8th Editions of the AJCC guideline.

The grade of the disease (lat. *gradus*) will provide information on how different the cancer cells are from those of the normal cells. The cancer is graded on a scale from 1 to 3. Grade 1 (or low grade) contains well differentiated cancer cells, that are slightly different from the normal cells. In grade 1, cancer cells do not proliferate as much as in grade 2, which is also known as moderately differentiated, and in grade 3, known as poorly differentiated, with highest rates of proliferation in disorganized and unbalanced pattern (Singhi et al. 2015; Maley et al. 2017). More broaden classification of HNSCC is certainly according to the HPV positivity, which distinguish HPV-positive from etiologically different HPV-negative HNSCC. Clinically, HPV is detected by *in situ* hybridization (ISH) of p16 immunohistochemistry used as a surrogate marker for HPV. The protein p16 (p16^{INK4a}) is a tumor suppressor protein, that in humans is encoded by the CDKN2A (cyclin-dependent kinase inhibitor 2A) gene, and if more than 70% of cells express p16 protein, the sample is considered HPV-positive (Li, Agrawal, and Fakhry 2015). For the HPV testing, it is now known that p16 staining is not a sufficient biomarker, since the presence of p16 has been detected even in HPV-negative HNSCC (Stephen et al. 2013; Albers et al. 2017). Interestingly, it was found that p16 positive, but HPV-negative HNSCC have some common characteristics with HPV-positive HNSCC with favorable prognosis (Albers et al. 2017). Hence, having data on p16 positivity should definitely be of interest. For more accurate HPV diagnosis, and for distinguishing HPV driven tumors from HPV passenger in HNSCC, it is necessary to perform HPV DNA together with HPV RNA testing (Monsonego et al. 2011; Mehanna et al. 2013; Maura L. Gillison et al. 2015).

1.2.3 Characterization of HPV associated and HPV-negative HNSCC

The HPV-16 is now established as crucial carcinogenic in a subset of HNSCC in numerous geographic regions around the world (Maura L. Gillison et al. 2015; Taberna et al. 2017). HPV-positive HNSCCs have genetic alterations that are indicative of HPV E6 and E7

oncoprotein function (Maura L. Gillison 2004), hence they are characterized by wild-type p53 (Gillison et al. 2000; Harriet C. Hafkamp et al. 2003), wild-type p16, and infrequent amplification of cyclin D (Ragin et al. 2006), whereas the opposite is correlated for HPV-negative HNSCC (Taberna et al. 2017). Moreover, p53 mutations in HPV-negative HNSCC are associated with tobacco and alcohol use (Urashima et al. 2013). HPV-positive HNSCCs also differ from HPV-negative HNSCCs in their genetic patterns (TCGA 2015). These observations provide support for at least two separate pathways for the multistage carcinogenesis of HNSCC: one driven primarily by the mutagenic effects of tobacco and alcohol (HPV-negative), and the other driven by HPV-mediated transformation (HPV-positive HNSCC). Although the clinical and molecular-genetic characteristics of HPV-positive HNSCC and HPV-negative HNSCC clearly differ, it has been unclear whether the risk factors for HPV-positive HNSCC are similar to those risk factors for HPV-negative cancers (Gillison et al. 2008).

Oral and oropharyngeal HPV is usually found in younger population, and related with higher oral sex practice (Gillison et al. 2015). HPV-positive HNSCC are often of oropharyngeal origin, with better prognosis and rare p53 mutations, usually with low levels of tobacco smoking and alcohol consumption (Gillison et al. 2015). Unlike HPV-positive HNSCC, HPV-negative tumors are found in elderly population and often with worse prognosis, without preferable origin, with frequent p53 mutations and having long history of tobacco and alcohol consumption. It is known that the incidence of HPV-negative HNSCC in the United States has been declining (Osazuwa-Peters et al. 2017), presumably due to a reduced prevalence of tobacco smoking (IARC 2012; Li, Agrawal, and Fakhry 2015). However, over the past few decades, there has been a rise in HPV-positive oropharyngeal cancers (Mehanna et al. 2013). The proportion of oropharyngeal cancers that are HPV-related in the developing countries is now approximately 70%, which is a substantial increase from previous incidence rates (Gillison et al. 2015; Osazuwa-Peters et al. 2017; Taberna et al. 2017). Even though these two groups are etiologically different, the treatment remains the same (Dok and Nuyts 2016). However, the treatment could be optimized to each group of patients.

As already mentioned, the main treatment modalities for HNSCC are surgical excision, chemo- and radiotherapy, and these treatment regimens can indispose patients, both functionally and cosmetically, while deteriorating their quality of life (IARC 2012; Li, Agrawal, and Fakhry 2015; Taberna et al. 2017). Hence, identifying more specific biomarkers could overcome limitations in understanding the disease mechanism, as well as improving treatment for each group of patients (Taberna et al. 2017; IARC 2012; Boscolo-Rizzo, Pawlita, and Holzinger 2016).

1.3 Epigenetic modifications regulating gene expression

Epigenetic modifications play a major role in diverse biological processes, such as cell cycle, cell signaling, differentiation, stemness, etc. (Ozkul and Galderisi 2016; Kelly and Gatie 2017). The modifications occur at the level of chromatin structure and organization (Liyanage et al. 2014). The DNA is tightly packaged in the nucleus in higher-order chromatin structures, with the basic repeating unit being the nucleosome, in which the DNA molecule is wrapped around a core of the histone proteins (Annunziato 2008). The histone core contains histones H2A, H2B, H3 and H4 (Mariño-Ramírez et al. 2005). The epigenetic modification regulating gene expression can affect histones and DNA itself. Chemical modifications of histones include enzymatic methylation, acetylation, phosphorylation, ubiquitination, and sumoylation (Hashimoto, Vertino, and Cheng 2010), while DNA is affected with methylation. These marks function as signals during various chromatin-based events, and play crucial role such as platforms for the recruitment, assembly or retention of chromatin-associated factors (Rossetto, Avvakumov, and Côté 2012).

The addition of an acetyl group (—CH3CO) on a histone is performed by lysine acetyltransferases, whereas its removal is performed by histone deacetylases. Acetylation of histones generally allows for an active state of chromatin (termed euchromatin), exposing DNA and allowing transcription to occur, while the methylation of histones may be either an activating or a silencing mark, depending on the specific amino acid affected (Liyanage et al. 2014).

Methylation is a process of adding the methyl group (— CH3) to a specific site in the histone, or the DNA. There is a continuous “cross-talk” among the histone post-translational modifications and DNA methylation/demethylation machinery. Histone H3 trimethylated at lysine 4 is most known as an active mark (H3K4me3), which is located at the 5' end of genes (Liyanage et al. 2014).

All four histone tails can be phosphorylated by addition of a phosphate group (PO_3^{-4}) via number of protein kinases, and dephosphorylated by phosphatases (Rossetto, Avvakumov, and Côté 2012). Phosphorylation of H2A(X) is the best known histone modification that plays a major role in DNA damage response (Rossetto et al. 2010). Furthermore, histone phosphorylation is not only important in DNA repair, but also plays a crucial role in transcription and chromatin compaction during cell division and apoptosis.

Histone ubiquitination is also an important modification with histones H2A and H2B been the two most abundant ubiquitinated proteins in the nucleus (Cao and Yan 2012). This modification plays critical roles in many processes in the nucleus, including transcription, maintenance of chromatin structure, and DNA repair.

Histone H4 can be modified by small ubiquitin-related modifier (SUMO) family proteins (Shiio and Eisenman 2003). Studies are suggesting that histone sumoylation mediates gene

silencing through recruitment of histone deacetylase and heterochromatin protein 1 (Shiio and Eisenman 2003; Nathan, Sterner, and Berger 2003).

In differentiated mammalian cells, the principal epigenetic tag found in DNA is the covalent attachment of a methyl group to the 5th carbon atom in DNA bases within the repetitive cytosine and guanine bases. However, recent findings suggest that non-CpG sites can also be methylated, depending on the main function of the targeted genes (Jin, Li, and Robertson 2011). DNA methylation can be also mediated via small non-coding miRNA molecules, regulating gene expression in miRNA-dependent manner (Hu et al. 2014).

1.3.1 DNA methylation

DNA methylation is a process of adding the methyl group (**Figure 12A**) at C5 position (Dricu et al. 2012), mainly of cytosine residues in CpG dinucleotide sequences (referred to as CpG islands), and is carried out by a family of enzymes called DNA methyltransferases (DNMTs) (Lister et al. 2009; Jin and Robertson 2013). There are at least three DNMTs, DNMT1, DNMT3a and DNMT3b, which are required for the establishment and maintenance of DNA methylation patterns. In human DNA, 5-methylcytosine is found in approximately 1.5% of genomic DNA (Lister et al. 2009), and in somatic cells it occurs almost exclusively in the context of paired symmetrical methylation of a CpG site. An exception to this is seen in embryonic stem (ES) cells, where a substantial amount of 5-methylcytosine is observed in non-CpG islands (Ramsahoye et al. 2000). In the bulk of genomic DNA, most CpG sites are heavily methylated, while CpG islands in germ-line tissues that are located near gene promoters remain unmethylated, thus, allowing gene expression to occur (Lister et al. 2009). When a CpG island in the promoter region of a gene is methylated, expression of the gene is repressed (**Figure 13**).

DNA demethylation, the removal of a methyl group is equally important as DNA methylation. The demethylation process is necessary for epigenetic reprogramming of genes and is also directly involved in many important disease mechanisms including tumor progression (Kim and Costello 2017). Decrease in global DNA methylation, also referred as DNA hypomethylation is caused by demethylation due to a variety abnormal states and has been proposed as a molecular marker in multiple biological processes such as cancer (Ehrlich 2009; Milutin Gašperov et al. 2014). Moreover, increase in global DNA methylation, also known as DNA hypermethlyation is well studied and described both in normal and aberrant conditions. In a normal conditions, it serves as a major event in silencing mobile transposable elements in the human genome, hence protecting the genome from chromosomal instability (Wilson, Power, and Molloy 2007), while in cancer, hypermethylation is often affecting promoters of tumor suppressor genes, which is associated with final inactivation of those genes. Knowing that hypermethylation of tumor suppressors (Frigola et al. 2005; Estécio and Issa 2011) and hypomethylation of transposons induce genomic instability and cell transformation, DNA

methylation seems to be promising in putative translational studies. Moreover, epigenetically deregulated genes may serve as potential biomarkers, hence, decoding the human epigenome is valuable and highly informative.

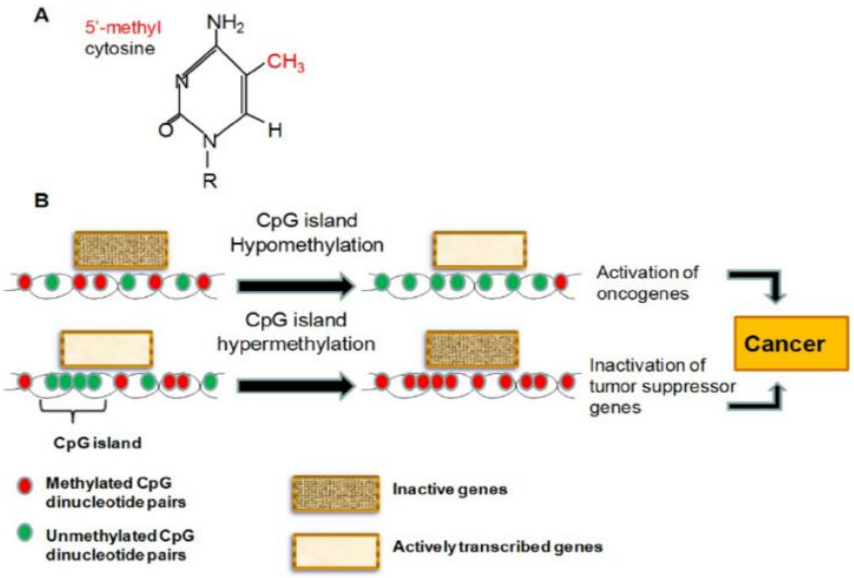

Figure 12. DNA methylation and cancer development. A) Structure of 5'-methyl cytosine; **B)** Role of the DNA hypo- and hypermethylation in the cancer development.
Adopted from Dricu et al. 2012.

To date, most commonly used methods for DNA methylation analysis are high-throughput technologies such as high performance liquid chromatography-ultraviolet (HPLC-UV), liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS), enzyme-linked immunosorbent assay (ELISA) based methods, next generation sequencing, LINE elements pyrosequencing, and methylation arrays. Among low-throughput methods for DNA methylation analysis are methyl specific polymerase chain reaction (MS-PCR, MSP), bisulfite sequencing and pyrosequencing of the specific genes.

The DNA methylation array method is based on hybridization of the sample DNA to the specific oligonucleotide probes in the microarray chip, covering more than 850,000 CpG sites in human genome in the most recent version of the commercially available assays. On the other hand, pyrosequencing interrogates many CpG sites within a single PCR amplicon and gives detailed information on the quantity and the site of methylation in the human genome. Regardless of the throughput of the method, all methods for analyzing methylation depend on detection of single-nucleotide polymorphisms (SNPs), which are artificially created at CpG sites through sodium bisulfite ($NaHSO_3$) modification. The principle of bisulfite modification (Delaney, Garg, and Yung 2015) is presented in **Figure 13**. After deamination, treating genomic DNA with sodium bisulfite selectively converts cytosine to uracil. However,

5-methylcytosine is protected from deamination, so cytosine cannot be converted to uracil, and the CG sequence is preserved in downstream reactions (Delaney, Garg, and Yung 2015; Li and Tollefsbol 2011). Bisulfite conversion is the mandatory step for most of the methods that are used for detection and investigation of DNA methylation.

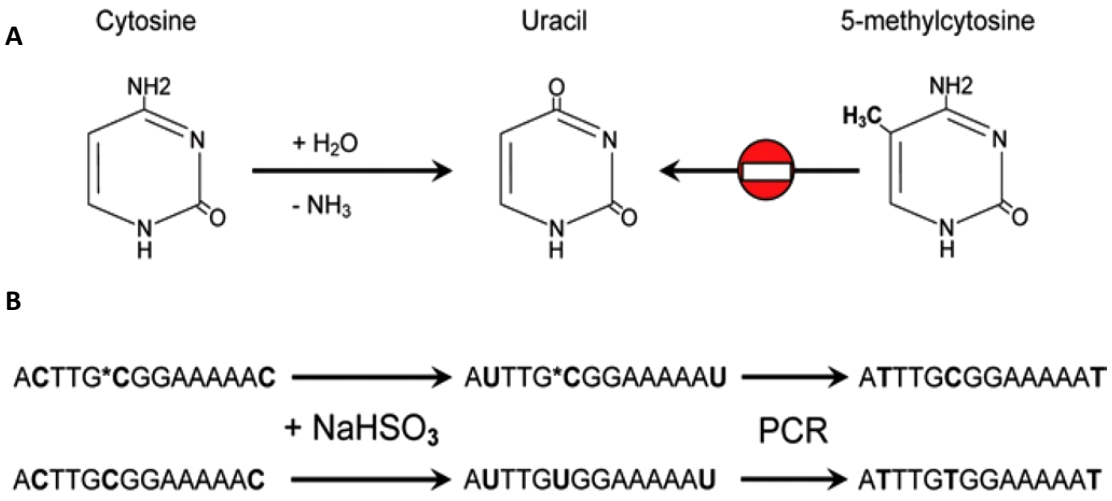

Figure 13. Principle of the bisulfite conversion. A) Deamination of cytosine to uracil is prevented by methylation of the 5-carbon position of cytosine, while cytosine without methyl group is successfully converted to uracil. **B)** Methylated (above) and unmethylated (below) CpG-containing DNA undergoes bisulfite conversion; methylated cytosines (*C) are unchanged while unmethylated cytosines are converted to uracil. Following PCR, the cytosine is retained, while uracil is converted to thymine. Pyrimidines involved in bisulfite conversion are in bold. *Adopted from Delaney, Garg, and Yung 2015.*

Pyrosequencing is based on the "sequencing by synthesis" principle (**Figure 14A-C**), in which the sequencing is performed by detecting the nucleotide incorporated by a DNA polymerase. The reaction relies on the light detection after pyrophosphate is released (Delaney, Garg, and Yung 2015). Detailed principle of pyrosequencing method is as follows.

Firstly, a segment of bisulfite converted DNA segment is amplified in a PCR reaction with reverse primer being biotinylated (**Figure 14A**). The strand with incorporated biotin serves as the pyrosequencing template. Therefore, after the denaturation step, the biotinylated single-stranded PCR amplicon is isolated and allowed to hybridize with a sequencing primer. The hybridized primer and single-stranded biotinylated template are further incubated with the enzymes DNA polymerase, ATP sulfurylase, luciferase, and apyrase, as well as the substrates adenosine 5' phosphosulfate (APS) and luciferin. After the first deoxyribonucleotide triphosphate (dNTP) is added to the reaction, DNA polymerase catalyzes addition of the dNTP to the sequencing primer, when it is complementary to the

base in the template strand. Each incorporation event is accompanied by the release of pyrophosphate (PPi) in an equal quantity to the amount of incorporated nucleotide. After the PPi is released, ATP sulfurylase converts PPi to ATP (**Figure 14B**) in the presence of adenosine 5' phosphosulfate (APS), and the ATP drives the luciferase-mediated conversion of luciferin to oxyluciferin which generates visible light in amounts that are proportional to the amount of ATP. The light produced in this reaction is detected by a charge coupled device (CCD) camera and is seen as a peak in the raw data output, which is called pyrogram. The height of each peak represents the light signal and is proportional to the number of incorporated nucleotides. Apyrase, which is a nucleotide-degrading enzyme, continuously degrades unincorporated nucleotides and ATP (**Figure 14C**). When degradation is complete, another nucleotide is added. Hence the addition of dNTPs is performed sequentially. As the process continues, the complementary DNA strand is built up and the nucleotide sequence is determined from the signal peaks in the pyrogram. Final analysis gives information on the total methylation percentage in investigated CpG island. Therefore, this principle makes pyrosequencing very informative, valuable and a quantitative method for investigating DNA methylation.

Besides pyrosequencing, MSP is a method of assessing the methylation status of virtually any group of CpG sites within a CpG island. It can detect if the investigated region of the gene is methylated or unmethylated, but unlike in pyrosequencing, the information of the methylation quantity is lacking (Herman et al. 1996). Furthermore, MSP is not a reaction analyzed in a “real time” manner, but is still sensitive, specific and informative. The principle of MSP is based on amplifying a bisulfite converted DNA in a PCR reaction using specific primers designed to bind to methylated, or unmethylated CpG sites. The presence of the PCR amplicon is then visualized under UV illumination and DNA intercalating dyes after gel electrophoresis.

In this study, we used DNA methylation array as the high-throughput method, and pyrosequencing as the validation method, while MS-PCR was used as the quality control for bisulfite conversion, necessary for the DNA methylation array and pyrosequencing.

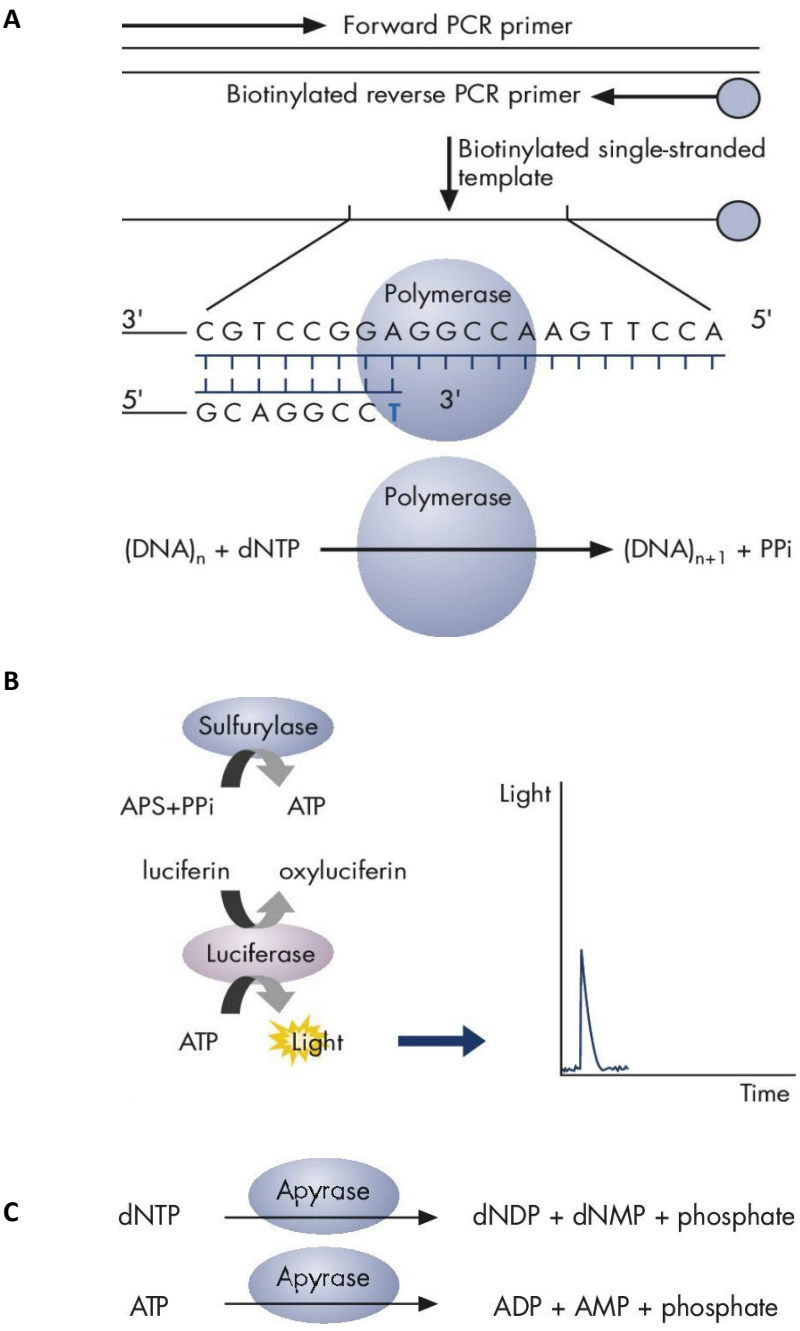

Figure 14. Principle of pyrosequencing. A) The hybridized sequencing primer and single-stranded biotinylated template are incubated with the enzyme mix and the substrates. Addition of dNTP to the sequencing primer is accompanied by the release of a pyrophosphate (PPi). **B)** ATP sulfurylase converts PPi to ATP in the presence of substrate APS. ATP drives the conversion of luciferin to oxyluciferin, which generates visible light, detected by CCD sensors and presented as pyrogram. **C)** Apyrase continuously degrades unincorporated nucleotides and ATP. When degradation is complete, another nucleotide is added. *Adapted from www.qiagen.com/us/.*

1.3.2 Expression of miRNAs

Apart from DNA methylation, alterations in DNA promoter regions, repetitive DNA sequences, and regulation of DNA expression is also mediated via the small non-coding RNAs such as microRNAs (miRNAs; miR). The miRNAs are highly conserved, single-stranded and about 22 nucleotides long molecules known to play an important role in a variety of biological processes, including development, cell proliferation, and differentiation (Wang, Wang, and Huang 2015), as well as in tumor suppression (Suzuki et al. 2012). Prediction indicates that miRNA account for 1-5% of the human genome and regulate at least 30% of protein-coding genes (Macfarlane and Murphy 2010). The miRNA can bind both DNA and mRNA, but their main function is binding target mRNA to prevent protein production by one of two distinct mechanisms (Macfarlane and Murphy 2010; Orang, Safaralizadeh, and Kazemzadeh-Bavili 2014): cleavage of target mRNA with subsequent degradation or inhibition of translation. To date, more than 2000 distinct miRNAs molecules have been identified within the human genome (Hammond 2015). Although little is currently known about the miRNA specific targets, it is evident that miRNA plays a crucial role in the regulation of gene expression, controlling diverse cellular and metabolic pathways (MacFarlane). Mature miRNA is generated (**Figure 15**) through two-step cleavage of primary miRNA (pri-miRNA), which incorporates into the effector complex RNA-induced silencing complex (RISC) (Winter et al. 2009). The miRNA functions by base-pairing with target mRNA to negatively regulate its expression. The level of complementarity between the miRNA and mRNA target determines which silencing mechanism will be employed; cleavage of target messenger RNA (mRNA) with subsequent degradation or inhibition of translation (Fukaya and Tomari 2012; Peng and Croce 2016).

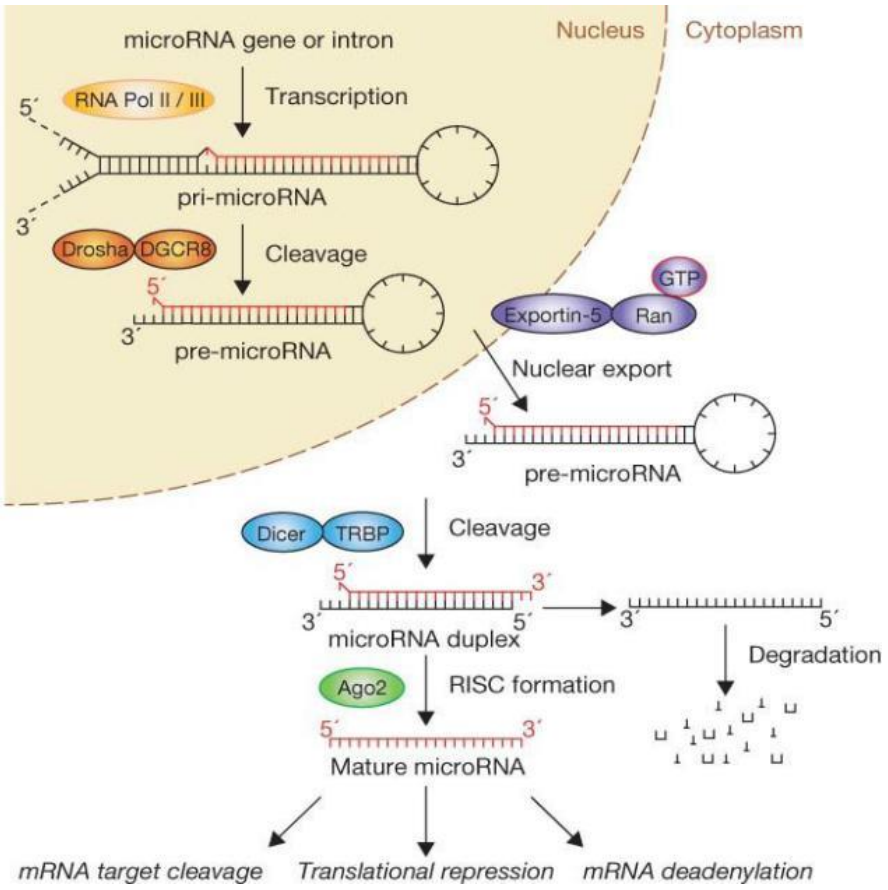

Figure 15. Biogenesis of miRNA. *Adopted from Winter et al. 2009.*

Many studies have demonstrated strong correlation between dysregulation of miRNA expression and cancer development through various mechanisms, including amplification or deletion of miRNA genes, abnormal transcriptional control of miRNAs, dysregulated epigenetic changes and defects in the miRNA biogenesis machinery (Peng and Croce 2016). It is also noted that miRNAs may function as either oncogenes or tumor suppressors (**Figure 16**) under certain conditions (Costa and Pedroso de Lima 2013). Moreover, dysregulated miRNAs have been shown to affect the hallmarks of cancer, including sustaining proliferative signaling, evading growth suppressors, resisting cell death, promoting metastasis and inducing angiogenesis. Compelling number of studies have identified miRNAs as potential biomarkers for human cancer diagnosis, prognosis and therapeutic targets or tools, which still needs further investigation and validation (Lan et al. 2015). It has also been noted that miRNAs play a significant role in HNSCC, while some changes are highly specific to the underlying risk factor (John et al. 2013). Thus, investigating pathways that are deregulated in cancer via miRs might certainly help understand the development and progression of cancer.

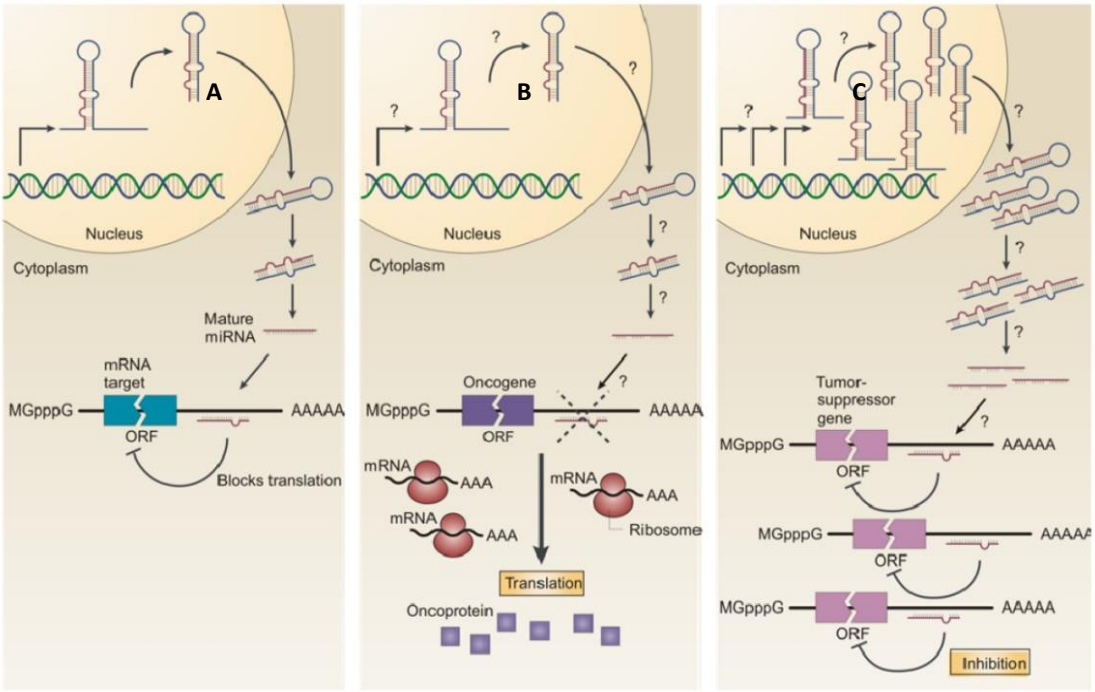

Figure 16. Biological roles of miRNA in **A)** normal tissue; **B)** as tumor suppressors, and **C)** as oncogene. The last scenario can lead to cancer development.

Adapted from Costa and Pedroso de Lima 2013.

1.3.3 Epigenetic changes in cancer development

Access to DNA by various cellular machineries is granted by a series of epigenetic modifications that change the high-order structure of chromatin. These epigenetic processes often entail dynamic chemical modifications to DNA or to the histones (Bannister and Kouzarides 2011; Liyanage et al. 2014). It is apparent that dysregulation of epigenetic modification can lead to malignant transformation and cancer development. In many cancers there has been noted epigenetic deregulation in tumor suppressors and oncogenes, where tumor suppressors get silenced via epigenetic modifications, while many oncogenes get activated (Sharma, Kelly, and Jones 2010). Furthermore, genetic material such as mobile transposons, that are normally silenced via epigenetics, become activated in cancer, enhancing genomic instability and promoting the cancer (Barros and Offenbacher 2009; Sharma, Kelly, and Jones 2010). Also, there have been studies on epigenetic deregulation in many genes that regulates the cell cycle, leading to aberrant DNA synthesis and malignant development (Herceg and Hainaut 2007). In addition, it is well established that some types of cancers, such as HNSCC are considered to be significantly driven by epigenetic deregulation (Masuda, Wakasaki, and Toh 2016), emphasizing the crucial role of epigenetic reprogramming in cancer development.

2 AIMS OF THE STUDY

Until now, numerous studies have tried to identify specific and more sensitive genetic and epigenetic biomarkers, which could contribute to a better understanding of the mechanisms of HPV-positive and HPV-negative HNSCCs to improve diagnostic and therapeutic treatment methods. However, a consensus on specific biomarkers has not yet been achieved. An integrative analysis of different epigenetic mechanisms is imperative, as it could help enlighten the process of HNSCC development, which could consequently improve treatment options for at least HPV-positive patients.

Therefore, the specific aims of this study are:

1. To evaluate the prevalence of the HPV types in archival HNSCC samples.
2. To establish the prevalence HPV types in fresh and prospectively collected control and HNSCC samples.
3. To determine HPV activity by evaluating E6 mRNA expression in HPV-positive fresh and archival HNSCC samples.
4. To evaluate the quantitative changes in the methylome (methylation profiling) and the miRnome (miRNA profiling) in fresh HNSCC samples in comparison to control samples.
5. To ascertain the statistical significance of possible changes in the miRnome in all samples considering the HPV presence.
6. To ascertain the statistical significance of possible changes in the methylome in all samples considering the HPV presence.
7. To identify the signaling pathways and the genes that have been epigenetically changed in HNSCC.
8. To integrate together data on methylome and miRnome considering key target genes.
9. To propose the most plausible epigenetic biomarkers identifying HPV-positive HNSCC.

3 PATIENTS, MATERIAL AND METHODS

3.1 Study population

Fresh samples from adult men and women, age 18-75 years were included in the study (n = 73). Sample groups consisted of volunteering healthy specimens (control group; n = 9) and patients (n = 61) that have been diagnosed with oral/oropharyngeal cancer in Croatian population with additional oropharyngeal cancer samples (n = 3) provided from Charles University, Vestec, Czech Republic. In order to participate in the study participants had to give their Informed consent. In addition, in order to get a clearer picture on HNSCC development, archival FFPE HNSCC samples (n = 115) over the 13-years period have been included in the study for the retrospective analysis of HNSCC incidence, HPV association, HPV activity and survival data on a Croatian population. Fresh tumor samples, collected from 2015 until 2018 were used for the miRnome (n = 16) and methylome (n = 10) screening of the population, and compared to a control group (n = 3 in miRnome, and n = 6 in methylome analysis). The study was approved by Bioethical Board of the Ruđer Bošković Institute (IRB BEP-3748/2-2014), Ethical Board of the Clinical Hospital Centre Zagreb (EP KBCZ 8.1-14/47-2, 02/21-JG) and Ethical Board of the Clinical Hospital Dubrava (EP KBD 10.06.2014).

3.1.1 Archival samples

FFPE tumor tissue from 115 patients with oral and oropharyngeal cancer have been included in this study (**Table 5**). Patients have been treated at the University Hospital Centre Zagreb between 2002 and 2015. The University Hospital Centre Zagreb is the largest hospital center in Croatia, treating patients from the whole country, which is important, since the study covered population from different geographical areas. Classification of the tumors were as follows; tumors of the tongue were categorized in the oral group, while those found on the base of tongue, soft palate, lateral wall of oropharynx, posterior pharyngeal wall and tonsils were classified as oropharyngeal tumors, according to ICD-O-3 code (NIH 2018). Furthermore, the study is in line with the Helsinki Declaration of 1975 as revised in 1983. For survival data purposes, available medical records and patients' vital status, including previous p16 staining were obtained from the hospital information system and the Croatian National Cancer Registry located at the Croatian Institute of Public Health (CNCR 2016). Malignancies were clinically staged according to the 7th Edition of the American Joint Committee on Cancer (AJCC) and the Union for International Cancer Control (UICC) TNM Classification of Malignant tumors (Edge and Compton 2010). The pathologic classification was considered more accurate and was therefore used, except in cases were the data on pathological TNM have been missing, so the clinical classification was used instead. For survival analysis, both

7th and 8th editions (Lydiatt et al. 2017) have been used and compared. Patients' survival time in months was calculated from the date of earliest diagnosis to registered time of death (all causes of death) or December 31st, 2017. Survival analysis was analyzed in MedCalc (v 11.4.2) using Kaplan-Meier survival curves. Logrank test was used to compare survival curves and provide information on statistical significance. The FFPE blocks had previously undergone routine pathological diagnosis and have been tested for p16 prior the HPV testing, using appropriate PCR primers (**Table 3**).

3.1.2 Fresh samples

For this study, fresh buccal swabs of healthy oral mucosa (n = 6), fresh tissue of non-malignant tonsils (n = 3) and fresh HNSCC tissue (n = 61) have been collected for the nucleic acid extraction purposes. **Tables 6** and **7** contain all relevant biological (for the patients and the control group, respectively) and all clinical and histopathological patient characteristics. For the analysis of specific miRs in HNSCC samples (**Table 7**) as well as methylome analysis, HNSCC patients with primary oral (n = 9) and oropharyngeal (n = 10) tumors treated at the Clinic of maxillofacial surgery of the Clinical Hospital Dubrava between 2015 and 2018, were enrolled. Survival analysis was also assessed on this group of samples (n = 53), however, since the total time period post treatment is less than 5 years, the overall five-year survival could not be calculated. For this reason, patients' survival time in months was calculated from the date of earliest diagnosis to registered time of death (all causes of death) or last registered follow-up. Available medical records on patients' vital status, tumor information and life style habits were obtained from the hospital information system. Malignancies were clinically staged according to the 7th Edition of the AJCC and the UICC TNM Classification of Malignant tumors (Edge and Compton 2010). The pathologic classification was used, except in cases for patients that have not been qualified for surgery and have been treated with radiotherapy (n = 5), so the clinical classification was used instead. Medical records on the complete patient therapy was not obtained, hence, the overall survival based on type of therapy was not assessed. Survival analysis was analyzed in MedCalc (v 11.4.2) using Kaplan-Meier survival curves. Logrank test was used to compare survival curves and provide information on statistical significance.

3.1.2.1 Control samples

Fresh control samples were consisted of two differently collected biological material, based on the study purposes. For the miRnome analysis, RNA from 3 non-malignant tonsils, which was provided by the collaborating group led by dr. Ruth Tachezy (Department of Genetics and Microbiology, Faculty of Science, Charles University, Vestec, Czech Republic) was considered adequate, while for the methylome analysis, buccal swabs from 6 healthy adult

men and women, age 27-58 was used. All control participants (n = 9) had no record of HPV infection in the oral mucosa and preferably no/low level of tobacco and alcohol intake, with only 2 that have been active smokers at the time of sample collection. Fresh oral swab samples (n = 6) have been collected at Ruđer Bošković Institute and processed for further nucleic acid analysis.

3.1.2.2 HNSCC tissue

A total of 65 adult men and women, with diagnosed oral and oropharyngeal cancer (**Table 6**) have been enrolled in the miRnome analysis, of which 10 (**Table 9**) have been also included in the whole-genome methylation (WGM) analysis. In the microRNA analysis, four patients were excluded because 2 tumors were recurrent and 2 were not planocellular carcinoma, thus making the total of 61 patients included in the study. Three additional fresh HPV-positive (HPV+) tonsil carcinoma, which have been kindly provided by the courtesy of dr. Ruth Tachezy, were included in the study, and their miRNA profile analyzed together with controls (n = 3) and fresh samples (n = 16) from the University Clinical Center Zagreb.

3.2 Nucleic acid isolation

3.2.1 DNA isolation

DNA was isolated from all sample groups for purposes of HPV genotyping and methylation analyses, preferably immediately, or within 2 weeks upon sampling. The method of DNA extraction depended on the sample properties (FFPE *vs.* fresh).

3.2.1.1 DNA isolation from FFPE samples

The FFPE blocks containing HNSCC samples have been cut on the microtome (**Figure 17)**. Serial sections of 10 µm were obtained into two separate 2 mL tubes (Eppendorf, Germany). Depending on the tissue size, approximately 5 to 7 sections from each FFPE block have been collected for an adequate nucleic acid isolation. DNA was isolated with NucleoSpin® Tissue kit (Machery-Nagel, Germany), according to the respective manufacturer's protocol. Each cube was cut with a clean scalpel and the microtome was cleaned using 70% ethanol after each sample in order to avoid cross contamination. Extracted DNA was quantified using a NanoPhotometer (Implen, Germany).

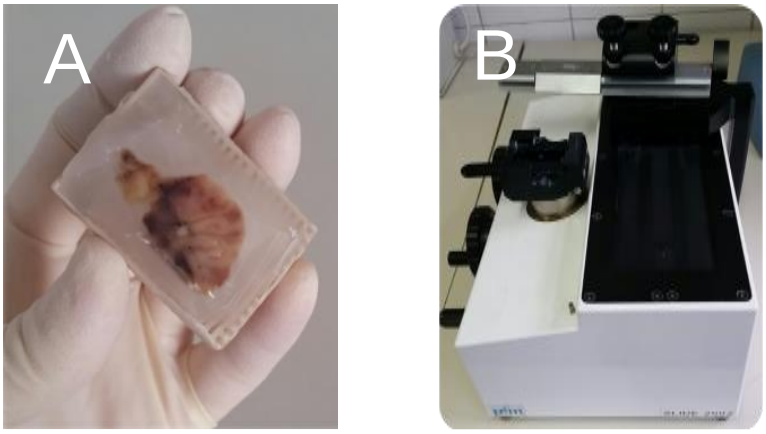

Figure 17. Archival cancer sample; A) FFPE HNSCC sample, and **B)** microtome used for cutting the tissue.

3.2.1.2 DNA isolation from fresh samples

Control swabs have been collected using a cytobrush (Ri. Mos, Italy) and incubated overnight in 700 μL of DNA extraction TEX media (1M Tris-HCl, 0.1M EDTA, 10% SDS) at 37°C, while shaking (600 rpm). Samples have been processed the next day, and the DNA was isolated using an in-house DNA salting method (Grce et al. 2000). Extracted DNA was quantified using NanoPhotometer (Implen, Germany).

Approximately 60-80 mg of fresh tumor tissue has been taken immediately upon surgical excision performed in Clinical Hospital Dubrava and preserved in 2 mL tubes (Eppendorf, Germany), each containing 700 mL of G2 lysis buffer (Qiagen, USA). Samples were stored at +4°C and delivered to the laboratory at the end of the day. In the laboratory, samples were stored at -20°C for up to 2 weeks before DNA isolation. Prior DNA isolation, approximately 10-30 mg of tumor have been cut into smaller pieces using a sterile scissor in order to facilitate DNA isolation from cells. DNA was isolated with an EZ1 Biorobot using EZ1 DNA tissue kit (Qiagen, USA) following the manufacturer's protocol. Extracted DNA was also quantified using NanoPhotometer (Implen, Germany).

3.2.2 RNA isolation

Total RNA was isolated from all samples using the appropriate kit. The method depended upon sample properties (FFPE *vs.* fresh) and further experiments. RNA from FFPE samples was used for identifying viral activity of HPV (E6*I mRNA detection), while total RNA from fresh samples was used for both viral activity analysis and miRNA profiling. Furthermore, precautions to avoid cross-contamination were taken.

3.2.2.1 RNA isolation from FFPE samples

Approximately 5-7 10 µm serial sections of each tissue block were obtained for RNA isolation. RNeasy FFPE Kit (Qiagen, USA) was used for RNA isolation from 45 HPV+ samples, according to the manufacturer's instructions. Extracted RNA was quantified using a NanoPhotometer (Implen, Germany). One sample contained only small portions of the tumor tissue which was not adequate for RNA isolation, hence the RNA concentration was very low. Low tumor material limited quality and the quantity of RNA necessary for reverse transcription experiment.

3.2.2.2 RNA isolation from fresh samples

As with the DNA, fresh tumor tissue has been taken immediately upon surgical excision and preserved in 2 mL microcentrifuge tubes (Eppendorf, Germany) containing RNA later (Ambion, USA) solutions for preserving RNAs. Samples were stored at +4°C and delivered to the laboratory at the end of the day. In the laboratory, samples were stored at -20°C for up to 2 weeks before RNA isolation. Total RNA was isolated using miRNeasy mini kit (Qiagen, USA) following the manufacturer's instructions. Total RNA quality and quantity were analyzed on a NanoPhotometer (Implen, Germany). Prior to further experiments, RNA was ran on a 1% agarose (Sigma Aldrich, USA) gel electrophoresis using appropriate apparatus (BioRad Laboratories, USA), stained with MidoriGreen (Nippon Genetics Europe, Germany) and visualized on a UVITEC Imager (Cleaver scientific, United Kingdom). Furthermore, RNA quality was assessed by Agilent Bioanalyser RNA6000 Nano kit (Agilent Technologies, USA) to determine RIN number.

3.3 HPV genotyping

3.3.1 HPV detection in DNA from FFPE samples

The quality of the extracted DNA was validated by PCR amplification with primers for β-actin (Lesnikova et al. 2010) gene generating a 99 bp fragment (**Table 3**). HPV DNA detection was performed with a PCR method in a thermocycler (2720 Thermal Cycler, Applied Biosystems, USA), using short primers suitable for FFPE tissue analysis, GP5/6 (~142bp amplicon) and SPF-10 (~65bp amplicon) (Snijders et al. 1990; de Sanjose et al. 2010). Thus, primers amplifying short fragments were chosen in order to avoid false negative results, due to high DNA degradation of FFPE samples (Dedhia et al. 2007). The CaSki cell line DNA, containing integrated HPV-16 type (Pattillo et al. 1977) was used as positive control, while negative control reactions contained all PCR reagents except of DNA. All standard precautions for avoiding cross contamination were followed. Ten µl of PCR products were ran on a 3%

agarose (Sigma) gel electrophoresis using the appropriate apparatus (BioRad Laboratories, USA), stained with MidoriGreen (Nippon Genetics, Europe) and visualized on a UVITEC Imager (Cleaver scientific, United Kingdom). A sample was considered as HPV+ if either GP or SPF PCR was positive and the results valid if β-actin PCR was successful.

3.3.2 HPV genotyping on DNA from fresh samples

The presence of HPV DNA was assessed using 4 types of consensus (PGMY, GP5/6+, LC and SPF) and one type specific PCR primer pair (HPV-16) as described previously (Milutin-Gašperov et al. 2007). All PCR reactions have been performed in a thermocycler (2720 Thermal Cycler, Applied Biosystems, USA). Briefly, 50 ng of sample DNA was amplified by PGMY, GP5/6+ and/or SPF10 and LC primers (**Table 3**) and with primers specific for HPV-16 to establish HPV presence. To determine sample adequacy, human β-globin specific PCR was performed as previously described (Milutin-Gašperov et al. 2007). CaSki cell line DNA containing HPV-16 was used as positive control, while negative control reactions contained all PCR reagents except of DNA. All standard precautions for avoiding cross contamination were followed. Ten µl of PCR products were ran on a 2% agarose (Sigma Aldrich, Germany) gel electrophoresis using the appropriate apparatus (BioRad Laboratories, USA), stained with MidoriGreen (Nippon Genetics Europe, Germany) and visualized on a UVITEC Imager (Cleaver scientific, United Kingdom). A sample was considered as HPV+ if either of the performed PCRs was positive and the results valid if β-globin PCR was successful. Samples with any discrepancy in HPV testing were further reanalyzed using INNO-LiPA HPV Genotyping Extra (Fujirebio, Japan) according to the manufacturer's instructions.

Table 3. PCR primers used for the detection of HPV DNA, E6*I mRNA and the quality control testing of the isolated DNA (β-actin PCR)

	Primer	Sequence	Reference
β-actin	6999-7018	5'CCACACTGTGCCCATCTACG3'	(Lesnikova et al. 2010)
	7097-7072	5'AGGATCTTCATGAGGTAGTCAGTCAG3'	
β-globin	GH20	5'GAAGAGCCAAGGACAGGTAC3'	(Vossler, Forbes, and Adelson 1995)
	PC04	5'CAACTTCATCCACGTTCACC3'	
PGMY	PGMY11-A	5'GCACAGGGACATAACAATGG3'	(P. Gravitt et al. 2000)
	PGMY11-B	5'GCGCAGGGCCACAATAATGG3'	
	PGMY11-C	5'GCACAGGGACATAATAATGG3'	
	PGMY11-D	5'GCCCAGGGCCACAACAATGG3'	
	PGMY11-E	5'GCTCAGGGTTTAAACAATGG3'	
	PGMY09-F	5'CGTCCCAAAGGAAACTGATC3'	
	PGMY09-G	5'CGACCTAAAGGAAACTGATC3'	
	PGMY09-H	5'CGTCCAAAAGGAAACTGATC3'	
	PGMY09-Ia	5'GCCAAGGGGAAACTGATC3'	
	PGMY09-J	5'CGTCCCAAAGGATACTGATC3'	
	PGMY09-K	5'CGTCCAAGGGGATACTGATC3'	
	PGMY09-L	5'CGACCTAAAGGGAATTGATC3'	
	PGMY09-M	5'CGACCTAGTGGAAATTGATC3'	
	PGMY09-N	5'CGACCAAGGGGATATTGATC3'	
	PGMY09-Pa	5'GCCCAACGGAAACTGATC3'	
	PGMY09-Q	5'CGACCCAAGGGAAACTGGTC3'	
	PGMY09-R	3'CGTCCTAAAGGAAACTGGTC3'	
	HMB01b	3'GCGACCCAATGCAAATTGGT3'	
GP5/6	GP5	5'TTTGTTACTGTGGTAGATAC3'	(Snijders et al. 1990)
	GP6	5'GAAAAATAAACTGTAAATCA3'	
SPF-10	SP1A	5'TAATACGACTCACTATAGGGCICAGGGICACAATAATGG3'	(Kleter et al. 1998)
	SP1B	5'TAATACGACTCACTATAGGGCICAGGGICATAACAATGG3'	
	SP1CD	5'TAATACGACTCACTATAGGGCICARGGICATAATAATGG3'	
	SP2BD	5'ATTTAGGTGACAACTATAGGTIGTATCIACWACAGTAACAAA3'	
HPV-16	F HPV-16	5'CCCAGCTGTAATCATGCATGGAGA3'	(Cheng 1995)
	R HPV-16	5'GTGTGCCCATTAACAGGTCTTCCA3'	
LC	L1C1	5'CGTAAACGTTTTCCCTATTTTTTT3'	(Yoshikawa et al. 1991)
	L1C2-1	5'TACCCTAAATACTCTGTATTG3'	
	L1C2-2	5'TACCCTAAATACCCTATATTG3'	
E6*I	F HPV16 E6	5'TTACTGCGACGTGAGGTGTA3'	(Smeets et al. 2007)
	R HPV16 E6	5'GGAATCTTTGCTTTTTGTCC3'	

3.4 E6 oncogene expression

For assessing HPV activity in HPV+ samples, the most abundant spliced form of HPV-16 oncogene E6, E6*I mRNA was analyzed. Tumor sample containing spliced E6*I mRNA form (PCR band of ~86 bp) was considered truly HPV+, meaning that the HPV-16 was active and most likely the main cause of cancer development. Samples that contained only full length 16 E6 mRNA (~248 bp) was considered HPV+, but inactive, meaning that the HPV-16 is not necessarily the main driving force in cancer development.

3.4.1 Reverse transcription (RT) on RNA from FFPE samples

Out of 45 samples tested positive for HPV DNA, 2 have been excluded due to samples being recurrent tumor, leaving a total of 43 samples included in the E6 oncogene expression analysis. RNAse-free DNAse digestion was performed to further limit and exclude any carryover DNA contamination. Briefly, 1 µg of RNA was reverse transcribed using a QuantiTect Reverse Transcription kit (Qiagen, USA) according to the manufacturer's instructions. Generated cDNA was used as a DNA template for the E6 cDNA PCR. One sample with very low RNA concentration was not adequately reverse transcribed, and was excluded from the E6*I analysis.

3.4.2 RT on RNA from fresh samples

HPV RNA analysis was performed on HPV DNA (14 in total) positive samples. RNAse-free DNAse digestion was performed to further limit and exclude DNA contamination. Briefly, 1 µg of RNA was reverse transcribed using a QuantiTect Reverse Transcription kit (Qiagen, USA) according to the manufacturer's instructions. Generated cDNA was used as a DNA template for the E6 cDNA PCR.

3.4.3 Amplification of E6 cDNA by polymerase chain reaction – PCR

For both FFPE and fresh HNSCC samples, the most abundant splice variant of the HPV-16 E6 open reading frame (E6*I) was detected by PCR (Smeets et al. 2007). Ten µl of PCR amplicons (~86bp) were ran on a 3% agarose (Sigma Aldrich, Germany) gel electrophoresis using the appropriate apparatus (BioRad Laboratories, USA), stained with MidoriGreen (Nippon Genetics Europe, Germany) and visualized on a UVITEC Imager (Cleaver scientific, United Kingdom). CaSki cell line cDNA was used as positive control, and the negative control contained all PCR reagents except cDNA in both cases. Suitability of cDNA for the

amplification was confirmed by β-actin PCR.

3.5 MicroRNA profiling

3.5.1 Preparation of miR libraries

Before the preparation of miR libraries, total RNA was analyzed on an Agilent 2100 BioAnalyzer (Agilent Technologies, USA) and RIN scores were high (~40% of samples having RIN between 8 and 9.5, others above 7 in almost all cases, with only one sample had 6). To supplement the samples, the analysis included 3 additional HPV+ tonsil carcinoma samples, collected at the University hospital Motol (Prague, Czech Republic). Three normal tonsil samples, also collected at University hospital Motol (Prague, Czech Republic) were additionally included. In total, 22 next generation sequencing (NGS) libraries were constructed from tumor samples using TrueSeq Small RNA Library prep kit (Illumina) following the manufacturer's protocol. For the library pooling, index pools A (1-12) and B (13-22) were used. Agilent 2100 Bioanalyzer (Agilent Technology, USA) was used for quality control of the indicated steps as recommended by the manufacturer. Libraries were quantified and visualized on the same machine.

3.5.2 Next generation sequencing – NGS of miR

Library sequencing was performed on the NextSeq 500 sequencer (Illumina, USA) using NextSeq 500 Mid output kit (Illumina, USA), according to the manufacturer's protocol. For a reliable data, raw sequences have been trimmed of adapter sequences using Illumina FastQC Basespace App (Illumina, USA) and by selecting TrueSeq Small RNA adapter sequences from the relevant app menu.

3.5.3 Bioinformatic analysis of NGS findings

NGS data have been organized in the subgroups, according to the tumor location and HPV status. Sequencing data were analyzed using the Illumina Small RNA Basespace App v1.0.1 (Illumina, USA) to determine significantly different miRNA expression between the groups. The automated pipeline in Basespace uses Bowtie algorithm to align reads against the reference databases to determine counts, which are then assessed for differential expression using DESeq2. Further subgroup analysis were performed by importing Illumina Small RNA Basespace App (Illumina, USA) count data into R studio (v 1.1.383), which was used to interface with R (v 3.4.2.) and perform miRNA differential expression using DEseq2 (v 1.18.1) (Love, Huber, and Anders 2014).

3.5.4 Validation of NGS data by RT quantitative PCR – RT-qPCR

For technical validation of the NGS experiment, we have selected 10 miRNAs that were found to be differentially expressed on the NGS analysis (miR-9-5p, 21-3p, 31-5p, 100-5p, 145-5p, 27a-5p, 34a-5p, 143-3p, 218-5p, 222-3p). Assays were selected in that manner to cover both over expressed and under expressed miRNAs, that were already investigated thoroughly and are well known to be dysregulated in the literature, or with discrepant literature findings (Avissar et al. 2009; Lajer et al. 2011, 2012; G. Gao et al. 2013; Wan et al. 2017). TaqMan Advanced miRNA synthesis (Applied Biosystems, USA) kit was used to convert isolated RNA to cDNA following the manufacturer's protocol for the same samples tested on NGS. Following conversion, 5 µL of diluted cDNA was analyzed by RT-qPCR on a StepOnePlus™ Real-Time PCR System (Thermo Fischer Scientific, USA) using TaqMan Advanced miRNA single tube assays (Applied Biosystems, USA), following the manufacturer's instructions. The three normal tonsil samples were pooled in equal concentration prior the cDNA synthesis to be used as normal reference. For generating Ct values, StepOne Software (v 2.3) was used (Applied Biosystems, USA).

Assays for miR-16-5p and 191-5p were evaluated as internal reference control (manufacturer's recommendation) as well as miR-181a-5p that showed very low intra-sample variation in our NGS experiment. Calculations were performed using each of the 3 referent miRs individually and as an average of all 3 values. Since the calculations gave very similar results, the final comparisons were performed using miR-16-5p for simplicity. The fold changes were calculated using the standard $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen 2001). Briefly, mean replicate Ct values of referent miR-16-5p for each sample were subtracted from mean replicate Ct values of individual miRs in the same sample. Subsequently, the ΔCt value for the control pool sample was subtracted from the ΔCt value of each cancer sample to obtain ΔΔCT value. The fold change was calculated by the $2^{-\Delta\Delta Ct}$ formula. The statistical difference was tested by t-test on ΔCt values of each miR compared to average ΔCt control miR values within each sample subgroup.

For the validation of potential relevant miRs in head and neck cancer, additional clinical samples, which have not been tested by NGS miRNA sequencing (n = 45) were tested with RT-qPCR individual assays in the same way as for the technical validation. The following miRNAs were selected: miR-9-5p, 21-3p, 29a-3p, 100-5p, 106b-5p, 143-3p, 145-5p, and 199b. Same reference control miRNAs were used as for technical validation. For the fold change calculations, combined sample pool of normal samples was used as the referent sample. Basic data analysis was performed in Microsoft excel and statistical testing in Medcalc (v 11.4.2).

3.6 Methylation profiling

MethylationEPIC BeadChip array was used for the genome-wide methylation profiling studies. Array enables quantitative interrogation of methylation sites across the whole genome and offer high-throughput analysis with a low sample input quantity of 250 ng DNA (Illumina microarrays; Illumina 2018).

3.6.1 Bisulfite conversion of DNA

Prior the methylation analysis, DNA from 16 samples (**Table 8**) has been converted by bisulfite using an EZ DNA Methylation Kit (Zymo Research, USA) according to manufacturer's instructions. Briefly, 500 ng of DNA sample was used to get 15 μL of bisulfite converted DNA. Alternative incubation conditions were applied due to using Illumina Infinium® Methylation Assay (Illumina, USA), hence the conditions in thermocycler (2720 Thermal Cycler, Applied Biosystems, USA) were as follows: 16 cycles of 95 C° for 30 s and 50 C° for 60 min, then hold at 4 C°. Conversion was confirmed as successful by performing MSP using specific primers for amplifying unmethylated *hTERT-1* gene. All samples tested positive, indicating successful bisulfite conversion and were credible for the main whole genome methylation experiment.

3.6.2 Whole genome DNA methylation analysis – Illumina

Methylation analysis of generated methylation libraries (TruSeq Methyl Capture EPIC Library Prep, USA) from 16 human bisulfite converted DNA samples was performed using the Infinium MethylationEPIC BeadChip kit WG-317-1001 (16 samples) according to the manufacturer's instructions (Illumina, USA). This kit targets more than 900,000 cytosine positions on the human genome at single-nucleotide resolution and covers: CpG islands in gene promoters; CpG sites outside of CpG islands, non-CpG methylated sites identified in human stem cells (CHH sites); differentially methylated sites identified in tumor *vs.* normal; FANTOM5 (functional annotation of the mammalian genome database; (FANTOM 2018) enhancers, which provide information on transcriptome for every primary cell type (FANTOM web resource); ENCODE (encyclopedia of DNA elements database) open chromatin and enhancers, to analyze functional elements that act at the protein and RNA levels, and regulatory elements that control cells and circumstances in which a gene is active (ENCODE web source; (ENCODE 2018); DNase hypersensitive sites and miRNA promoter regions. The EPIC Beadchip contains most of the content (over 90%) of the previous HumanMethylation450K BeadChip.

MethylationEPIC BeadChip array 3 day workflow included the following steps; (i) library

preparation, which included PCR-free kit using Infinium I assay chemistry technologies (Illumina, USA); (ii) array processing and scanning on the array scanner (iScan System, Illumina, USA), and (iii) data analysis (Illumina 2018). First and second step were performed in a collaboration with CBM (Cluster in BioMedicine, Trieste, Italy), performing Illumina array and scanning data on iScan System, while the sample preparation and bioinformatics analysis were done at the Ruđer Bošković Institute, Zagreb.

Briefly, after the bisulfite treatment, 200 ng DNA was subjected to the whole genome methylation analysis (WGM) and enzymatic digestion with reagents provided within the Infinium MethylationEPIC BeadChip kit (Illumina, USA). The hybridization of the samples on the BeadChips and washing procedures followed the standard procedures obtained from Illumina. An Illumina iScan System (Illumina, CA, USA) was used to read the BeadChips.

3.6.3 Bioinformatic analysis of Illumina findings

WGM data have been analyzed using R studio (v 1.1.383). It was used to interface with the software framework R (v 3.4.2.) and RnBeads package (v 1.2.2.) for the normalization and differential analysis of methylation data.

3.6.4 Validation of DNA methylation biomarkers

For the validation of genome-wide methylation data, we chose the pyrosequencing method. Four gene promoters that showed to be significantly hypomethylated (*TRDC* and *LAIR2*) and hypermethylated (*SPRR3* and *FBXO2*) after genome-wide methylation analysis have been validated. Pyrosequencing was performed on 4 tumor samples (2 HPV+ and 2 HPV−) and 4 normal controls, due to limited DNA material. Approximately 500 ng extracted DNA was used for the bisulfite treatment performed with the EZ DNA Methylation-Lightning Kit (Zymo Research, USA), according to the manufacturer's instructions. The efficiency of bisulfite conversion was assessed using MSP containing specific primers for methylated form of *SIGLEC12* gene. Forward primer sequence was 5'TGTTGATAATGTAGAAGTTCGTGAC3', and reverse primer 5'ACCAATAACCATAAACTAAATCGAA3'. The PCR was performed in thermocycler (Veriti, 96 Well Thermal Cycler, Applied Biosystems, USA) and the program was as follows: initial denaturation step of 5 min at 95°C, followed by 45 cycles of 30 s denaturation at 95°C, annealing of primers at 61°C for 30 s and extension for 50 s at 72°C and one cycle of final extension for 10 min at 72°C. PCR amplicon product was 117 bp. All samples tested positive, indicating successful bisulfite conversion and showing that the samples were suitable for pyrosequencing.

3.6.4.1 Choice of specific primers

Specific primers for the selected genes were designed using the PyroMark Assay Design software v.2.0.1.15. (Qiagen, USA). After the WGM analysis, the promoter region that showed most significantly different methylation pattern between the cancer and normal samples, for each gene was assessed using the UCSC Genome Browser (UCSC 2018) and inserted in the program. Those specific regions revealed the following CpG sites: 15 CpG for gene *SPRR3*, 46 for *FBXO2*, 17 for *TRDC* and 36 for *LAIR2*. The aim was to cover as many CpG sites as possible using the amplifying (with reverse primer being biotinylated) and sequencing primers, which are necessary for the pyrosequencing reaction, while passing the quality criteria. The inclusion criteria for the primer selection was a quality score (Q) higher than 60. Regions with repetitive A and T bases within the primer binding sites were avoided, while regions with 50% of GC bases were preferable for the primer binding sites. Furthermore, only 1 variable region (CpG site) was allowed per primer binding sites. The PCR and sequencing primers were designed to assess the same promoter analyzed by the Infinium MethylationEPIC BeadChip array. In **Table 4**, the primer sequences, amplicon sizes, and the optimal annealing temperatures are indicated.

Table 4. Designed primers (HPLC purified) used in pyrosequencing validation.

Primer ID	Primer sequence	Ta (C°)	Amplicon size
TRDC F	GAATGATTTAGGAGGTAGAGTTTGT		
TRDC R*	ACCTCCAATCACTTCAAACTTCAT	63 C°	132 bp
TRDC S	AGGAGGTAGAGTTTGTA		
SPRR3-1 F	TAGTGTATTGTTTGGAAGGTAGT		
SPRR3-1 R*	CCATTCAACTACTTCTTCCTACT	57 C°	286 bp
SPRR3-1 S	ATAATTGGTTTTTTGATTTTTTTAA		
SPRR3-2 S**	TTTTTTATATAGGGAAATATTG	57 C°	286 bp
LAIR2-1 F	TGTGGTTTTGGTTTTTGTGTAAG		
LAIR2-1 R*	CTTCAATCAAACCCAAAATTCATCCT	57 C°	194 bp
LAIR2-1 S	TGGTTTTTGTGTAAGAGT		
LAIR2-2 S**	TGGGGTTTGAGAGAT	57 C°	194 bp
FBXO2 F	AGATGGGTATGGTGGTATTTG		
FBXO2 R*	CTAACCTCCAATACCCACTTCTATC	55 C°	253 bp
FBXO2 S	GGTGGTATTTGTTTGTAAT		

*5’ biotinylated primer; **amplifying primers (F and R) used in pyrosequencing reaction are the same for both sequencing primers

3.6.4.2 Pyrosequencing

The pyrosequencing analysis was performed on samples which were already tested by Infinium MethylationEPIC BeadChip array. PCR reactions on bisulfite converted DNA were performed according to the PyroMark PCR protocol (Qiagen, USA) in a total volume of 30 µl. PCR consisted 0.10 µmol/L of each primer (Macrogen, South Korea), 1.5 mM $MgCl_2$, PyroMark PCR Master mix (Qiagen, USA), Coral Load (Qiagen, USA) and 50 ng of bisulfite treated template DNA. The PCR was performed in a thermocycler (Veriti, 96 Well Thermal Cycler, Applied Biosystems, USA) and the program was as follows: initial denaturation step of 1 min at 95°C, followed by 45 cycles of 30 s denaturation at 95°C, specific annealing Ta for 30 s and extension for 30 s at 72°C and one cycle of final extension for 10 min at 72°C.

PCR amplicons from all PCR reactions have been visualized using agarose gel electrophoresis, under UV light. PCR amplicons have been stored at -20 C°, until the pyrosequencing was ready to be performed. Pyrosequencing was performed at the University of Zagreb School of Medicine, Department of Medical Biology, Zagreb, using a PyroMark Q24 Reagent Kit and a PyroMark Q24 system (Qiagen, USA) as described previously by Mikeska *et al.* (Mikeska et al. 2011). The nucleotide addition order was optimized by the PyroMark Q24 Software (Qiagen, USA) and the results were automatically analyzed using the same software. After the analysis, percentage of methylation per each investigated CpG has been analyzed in LibreOffice (v5.1.6.2). Percentage of methylation for each CpG island between the two sample groups (cancer *vs.* normal controls) was compared and *P* values determined using t-test.

4 RESULTS

4.1 HPV types associated with HNSCC

Three types of molecular studies, HPV detection (including HPV-16 transcriptional activity), microRNA and methylation profiling, have been performed on the 2 groups of tumor samples: archival FFPE HNSCC samples and fresh HNSCC samples. Archival samples have been adequate for the HPV detection and HPV-16 transcriptional activity study (E6*I mRNA detection). Fresh tumors were used in HPV detection and HPV-16 transcriptional activity study as well, but also for the whole epigenome analysis (miRNA and methylation). Fresh HNSCC samples were better suited for miRnome and methylome profiling due to the better quality of the isolated nucleic acids. During the process of fixing and embedding the tumor tissue in paraffin, DNA and RNA often get degraded, which is why we chose to perform that analysis only on fresh tumor tissue. Moreover, specific primers used in HPV-type specific PCR can often produce false negative results, due to the DNA degradation, which is why appropriate primers should be used when dealing with the DNA from FFPE tumor samples. This is the reason why specific HPV types have been determined only in fresh control and cancer samples, while in archival samples we have detected only short amplicons of mucosal HPV-consensus regions (GP5+/6+ and SPF).

4.1.1 HPV presence in archival samples

A total of 119 FFPE tumor samples have been collected from the KBC Zagreb and the DNA isolated as described in the Materials and methods section. After excluding 1 sample that was not found in the database, and three FFPE samples that belonged to the same patient, the total number of samples that were included in the study was 115. Most of the patients were older, with median age of 59. Medical documentation was missing for a subset of patients for life style habits (53.1% missing data on smoking and drinking), 42.7% for TNM stage, 2.6% of tumor grade and 26.9% on p16 positivity. Regarding documented life style habits, data were presented as active smoker, former, nonsmoker, while for drinking habits as heavy drinking, occasionally or does not drink. Of all patients with documented overall tumor stage and grade, most of the patients were diagnosed with late stage tumors (52 of 66; 78.8%), while the percentage of grade 1, 2 and 3 was almost equally distributed (26%, 35% and 37%, respectively). All samples tested positive for the β-actin PCR, which was used as an internal control of the sample quality.

After confirming DNA was adequate for subsequent experiments, two consensus PCRs have been performed for the detection of HPV DNA: SPF10 and GP5/6+ (**Table 3**). Results were as

follows: out of 115 samples, 42 (36.5%) were tested SPF positive (**Figure 18**) and 19 (16.5%) were GP positive, while only 16 (14%) samples tested positive in both PCR reactions. Sample was considered HPV DNA positive if either of the two PCR was positive, leaving the total of 45 HPV DNA positive samples. Overall HPV prevalence and patient characteristics are presented in **Table 5**. Briefly, for the HPV activity data, most abundant spliced form of HPV-16 oncogene E6, E6*I mRNA has been assessed from HPV DNA positive samples. Thirteen samples were positive on E6*I mRNA, while 32 expressed only full length E6. Based on the HPV DNA and RNA analysis, patients were categorized in the following groups: HPV-positive (+) group contained samples tested positive for both HPV DNA and E6*I mRNA; HPV inactive group contained samples tested positive for HPV DNA, but negative for E6*I mRNA; while HPV-negative (–) group contained samples tested negative for both HPV DNA and any mRNA. Samples corresponding to positive control result were considered positive. Only results with no amplification in negative control reaction were considered truly positive and PCR of any suspect batches was completely repeated. Positive control used in HPV genotyping was DNA isolated from HPV-16 positive CaSki cell line.

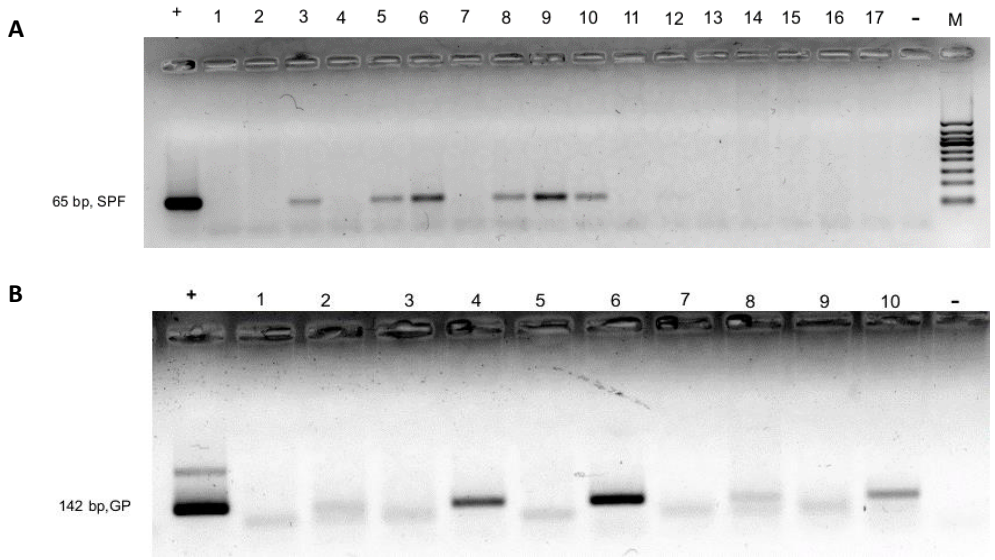

Figure 18. HPV DNA detection in FFPE samples. Representative PCRs using A) SPF and **B)** GP5/6+ primers. A: + = positive control; 1-17: random tumor FFPE samples; – = negative control; M = GelPilot50 (Qiagen, USA), SPF amplicon is ~65 bp. B: + = positive control; 1-10: tumor FFPE samples; - = negative control. GP5/6+ amplicon is ~142 bp.

Regarding tumor location, most of the HPV+ are located in oropharynx and arise from tonsil and the base of tongue (53.8% and 30.8% respectively), which is the same case in HPV inactive group, with 50% of tonsil and 34.4% of base of tongue tumors. In the HPV– group, prevalence of tumors in mentioned regions is slightly lower, but still predominates, with

42.9% in tonsils, and 28.6% in the base of tongue. What is evident from **Table 5**, is that oral cancer is more prevalent in HPV− group, accounting for 14.2% of tongue tumors, while in HPV+ and HPV inactive group, the prevalence was 7.7% and 9.3% respectively. Men are most prevalent in all three groups (62.2% in HPV+, 81.3% in HPV inactive and 87.1% in HPV− group).

For a relatively large number of patients, p16 status was unknown (26.9% overall; 30.8% in HPV+, 31.3% in HPV inactive and 24.3 in HPV−), and marker p16 does not seem to correlate to the HPV testing (r = -0.09 for the comparison with HPV DNA, and r = 0.18 for the comparison with RNA). There is high discrepancy within the groups, where in HPV+ group only 15.4% p16 positive samples were in concordance with HPV PCR results. There is very high percentage of p16 negative samples in HPV+ and HPV inactive groups (53.8% and 43.7%, respectively), while there are 31.5% p16 positive samples in HPV− group (**Table 5**).

Table 5. Patient characteristics of the study population for archival FFPE samples

Variable			HPV DNA+ RNA+	HPV DNA+ RNA–	HPV–	Total
			n = 13 (%)	n = 32 (%)	n = 70 (%)	n = 115 (%)
Gender	M		9 (69.2)	26 (81.3)	61 (87.1)	96 (83.5)
	F		4 (30.8)	6 (18.7)	9 (12.9)	19 (16.5)
Tumor location*	OP	Base/tongue	4 (30.8)	11 (34.4)	20 (28.6)	35 (30.4)
		Tonsil	7 (53.8)	16 (50)	30 (42.9)	53 (46.1)
		Soft palate	0	0	7 (10)	7 (6.1)
		Pharynx	1 (7.7)	2 (6.3)	3 (4.3)	6 (5.2)
	O	Tongue	1 (7.7)	3 (9.3)	10 (14.2)	14 (12.2)
Age group	< 45y		1 (7.7)	1 (3.1)	6 (8.6)	8 (7)
	45-64y		9 (69.2)	21 (65.6)	42 (60)	72 (62.6)
	> 65y		3 (23.1)	10 (31.3)	22 (31.4)	35 (30.4)
	Median age		58	59	59	59
Smoking	Active		1 (7.7)	9 (28.2)	18 (25.7)	28 (24.3)
	Former		0	2 (6.3)	9 (12.9)	11 (9.6)
	Doesn't smoke		3 (23.1)	6 (18.7)	6 (8.6)	15 (13)
	Unknown		9 (69.2)	15 (46.8)	37 (52.8)	61 (53.1)
Drinking	Heavy		0	3 (9.3)	6 (8.6)	9 (7.8)
	Occasionally		2 (15.4)	7 (21.9)	20 (28.6)	29 (25.2)
	Doesn't drink		2 (15.4)	7 (21.9)	7 (10)	16 (13.9)
	Unknown		9 (69.2)	15 (46.9)	37 (52.8)	61 (53.1)
Tumor stage 7th ed.	T	1	1 (7.7)	5 (15.6)	5 (7.1)	11 (9.6)
		2	2 (15.4)	9 (28.2)	13 (18.6)	24 (20.8)
		3	2 (15.4)	5 (15.6)	7 (10)	14 (12.2)
		4	2 (15.4)	2 (6.3)	13 (18.6)	17 (14.7)
		Unknown	6 (46.1)	11 (34.3)	32 (45.7)	49 (42.7)
	N	0	0	8 (25)	13 (18.6)	21 (18.2)
		1	2 (15.4)	6 (18.7)	3 (4.3)	11 (9.6)
		2	5 (38.5)	7 (21.9)	19 (27.1)	31 (26.9)
		3	0	0	3 (4.3)	3 (2.6)
		Unknown	6 (46.1)	11 (34.4)	32 (45.7)	49 (42.7)
Overall Stage	Early	I	0	2 (6.3)	3 (4.3)	5 (4.3)
		II	0	4 (12.4)	5 (7.1)	9 (7.8)
		III	3 (23.1)	7 (21.9)	5 (7.1)	15 (13)
	Late	IVa	4 (30.8)	8 (25)	21 (30)	33 (28.7)
		IVb	0	0	3 (4.3)	3 (2.6)
		IVc	0	0	1 (1.5)	1 (0.9)
	Unknown		6 (46.1)	11 (34.4)	32 (45.7)	49 (42.7)
Grade	1		4 (30.8)	7 (21.9)	19 (27.1)	30 (26.1)
	2		2 (15.4)	13 (40.6)	25 (35.7)	40 (34.7)
	3		7 (53.8)	12 (37.5)	23 (32.9)	42 (36.6)
	Unknown		0	0	3 (4.3)	3 (2.6)
Therapy**	Surgery		3 (23.1%)	10 (31.3)	14 (20%)	27 (23.5%)
	Surgery + RT		0	12 (37.5)	16 (22.9%)	28 (24.3%)
	Surgery + CRT		3 (23.1%)	1 (3.1)	10 (14.3%)	14 (12.2%)
	RT		1 (7.7%)	3 (9.4)	2 (2.9%)	6 (5.2%)
	CT		0	0	1 (1.4%)	1 (0.9%)
	CRT		0	0	4 (5.7%)	4 (3.5%)
	Unknown		6 (46.1%)	6 (18.7)	23 (32.8%)	35 (30.4%)
p16	Positive		2 (15.4)	8 (25)	22 (31.5)	32 (27.8)
	Negative		7 (53.8)	14 (43.7)	31 (44.2)	52 (45.3)
	Unknown		4 (30.8)	10 (31.3)	17 (24.3)	31 (26.9)

*OP = oropharyngeal tumor, O = oral tumor; **RT = radiotherapy, CRT = chemoradiotherapy, CT = chemotherapy

4.1.2 HPV presence in prospectively collected fresh samples

Total of 65 fresh tumor samples have been collected from KBD Dubrava and the DNA isolated as previously described. After excluding 4 samples as 2 were recurrent tumors and 2 were not planocellular carcinoma, the total number of patients included in the study was 61. As with archival samples, most of the patients were older, with median age of 62. Medical documentation was complete for all samples. Regarding life style habits, data were not presented in detailed as with the FFPE samples, hence there was only information whether the patient was smoking (S) or drinking (D), being smoker/drinker (SD) and if the patient self-declared as a non-smoker and non-drinker (NSND). According to available data most prospectively enrolled patients in the study have been SD (35 of 61; 57.4%). As with the archival tumors, most of patients were diagnosed with late stage tumors (45; 73.8%), while the percentage of grade 1 and 2 were almost equally distributed (37.7% and 41.1%, respectively), with only 3 patients having tumors of grade 3 (13.1%). In case when the tumor was inoperable data on tumor grade was missing (5; 8.1%).

All samples tested positive for the β-globin PCR, which served as an internal control for each sample quality. After confirming DNA was adequate for the next experiments, 4 consensus PCRs and 1 HPV type-specific PCR have been performed for the detection of HPV DNA (**Figure 19**): PGMY, nested GP5/6+ from PGMY amplicon, SPF10, LC and HPV-16 type specific amplification to detect the presence of HPV type most often found in HNSCC (**Table 3**). Results were as follows: out of 61 samples, 5 samples (8.2%) were tested PGMY positive, 12 (19.6%) were GP positive, 5 were (8.2%) were SPF10 positive, 2 were positive to LC (3.3%), and 5 samples tested positive in HPV-16 PCR (8.2%). Samples corresponding to the positive control result were considered positive. Only results with no amplification in the negative control reaction were considered. Positive control used in HPV genotyping was DNA isolated from HPV-16 positive CaSki cell line. Overall, a sample was considered HPV DNA positive if it was positive to either of these PCRs; however, samples (n = 14) that showed any inconsistency in results, were further analyzed using the INNO-LiPA HPV Genotyping Extra Assay, in order to avoid any misclassifications of the samples (**Figure 20**). After INNO-LiPA PCR, the total of HPV DNA positive samples was 14 (23%). Hybridization was successful in all samples, and results were as follows: out of 14 samples, 9 (64.3%) were tested positive only for HPV-16 positive, with one being positive for both HPV-16 and HPV-18 types, (7.1%), 2 patients have been positive to HPV-18 (14.3%), while one sample showed weak positivity for mucosal HPV type of an unknown risk and one was HPV– (7.1% in both cases). HPV-16 type was most common, being found in 71.4% of samples.

For the HPV-16 E6*I mRNA analysis, only HPV+ samples (n = 14) have been examined (**Figure 21**). Six samples tested positive for E6*I mRNA, while 8 expressed only full length E6 mRNA. E6*I mRNA analysis results in fresh tumors will be presented in a later chapter in more details. Based on the HPV DNA and RNA analysis, patients were accordingly categorized in

the following groups (**Table 6**): HPV DNA and E6*I mRNA positive; HPV DNA positive, but negative for E6*I mRNA; while HPV− group contained samples tested negative for both HPV DNA and E6*I mRNA.

Overall, 14 samples (22.9%) tested positive for HPV DNA, of which 6 (42.9%) were categorized into HPV DNA and RNA+ and 8 (57.1%) into HPV DNA+ and RNA −. Again, most were males in all three groups, HPV DNA and RNA +, HPV DNA + RNA − and HPV− (83.4%, 87.5% and 74.5%, respectively). The most prevalent type of cancer within the HPV RNA+ group arose from tonsil (50%), within HPV RNA − from tongue (50%), and in HPV DNA and RNA − group from gingiva (25.5%). In fresh tumor samples, oral site of tumor was more prevalent (46; 75.4%), with tumors arising from several sites, gingiva (15; 24.6%), floor of mouth (12; 19.7%), tongue (10; 16.4%), retromolar area (7; 11.4%) and buccal mucosa (2; 3.3%). In HPV RNA+ group, tumors were equally found in oropharyngeal squamous cell carcinoma (OPSCC) and oral squamous cell carcinoma (OSCC), with (50% *vs.* 50%, respectively), while in HPV RNA − group tumors were more prevalent in oral region, with 25% HPV+ OPSCC and 75% HPV+ OSCC.

In summary, of all 61 patients, after the analysis of HPV status and patients' biological and clinical (type of tumor) data, 16 samples were chosen for the NGS testing and 10 for the methylation analysis. For the validation method of NGS data, RT-qPCR was performed on all 61 samples. HPV results were as follows:

- 12 out of 61 samples (19.6%) used for RT-qPCR tested HPV+; 9 samples were HPV-16 DNA positive, 2 (3.3%) samples tested positive for HPV-18 DNA, and one sample was positive for both HPV-16 and HPV-18 types.
- 9 out of 16 cancer samples (56.3%) used for NGS analysis were HPV+ (**Table 5**), and all samples tested positive for HPV-16 DNA.
- 6 out of 10 (60%) cancer samples used for the WGM analysis were HPV+ (**Table 5**), and all samples tested positive for HPV-16 DNA, while 4 samples were HPV−.
- 2 out of 4 (50%) cancer samples used for the pyrosequencing (**Table 10**) were HPV+ and have been positive to HPV-16.

Of all the control samples used as a reference sample in methylome and miRnome analysis, results were as follows:

Three appropriate non-malignant tonsil tissue samples have been used as a control for NGS and RT-qPCR analysis (**Table 7**). All samples tested negative for the presence of HPV DNA, and were similar to tumor samples in terms of age and the type of tissue. High prevalence of HPV associated samples was in fact found in tonsils, and occurred in adults (50-70y in HPV− group and 30-50 y in HPV+ group). All 3 samples were collected from healthy 59 years old adults surgically operated for non-malignant reasons.

Six fresh swabs from healthy oral mucosa have been chosen as control for the methylation analysis (**Table 9**), based on the HPV status and biological characteristics. None of the samples tested positive for the presence of HPV DNA. HPV presence was assessed in control samples in the same manner as with fresh tumor HNSCC. Median age of subjects was 43.5 years.

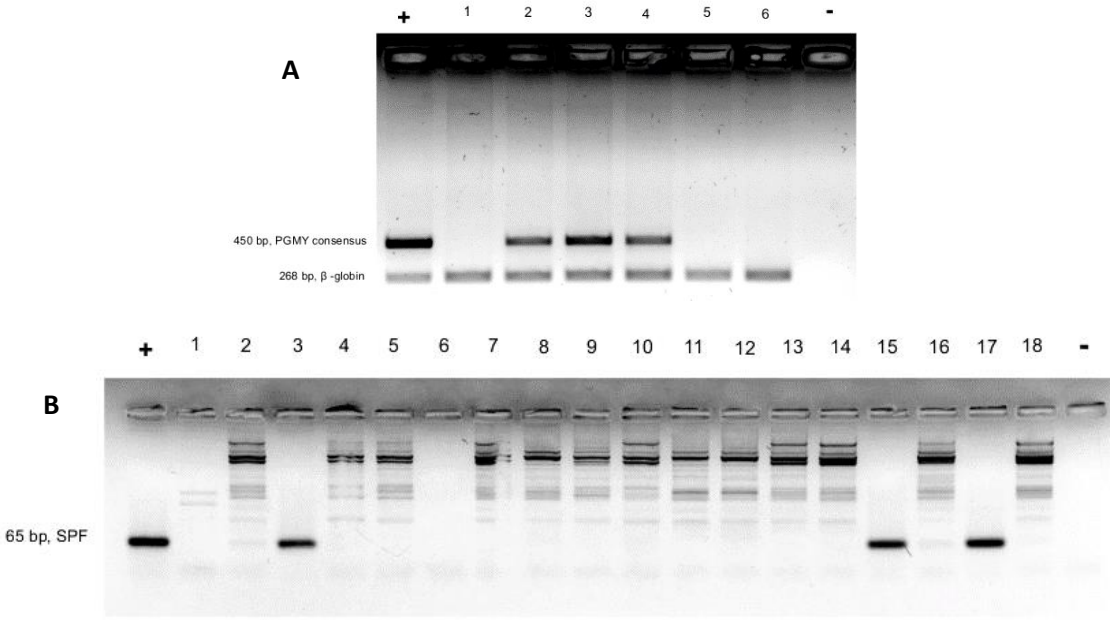

Figure 19. HPV DNA detection in fresh HNSCC samples. Representative PCRs using **A)** PGMY and β-globin and **B)** SPF10 primers. A: + = positive control; lanes 1-6: fresh tumor samples; - = negative control; amplified mucosal consensus PGMY is ~450 bp, and internal control β-globin amplicon is 268 bp; B: + = positive control; lanes 1-18: fresh tumor samples; - = negative control. SPF amplicon is ~65 bp.

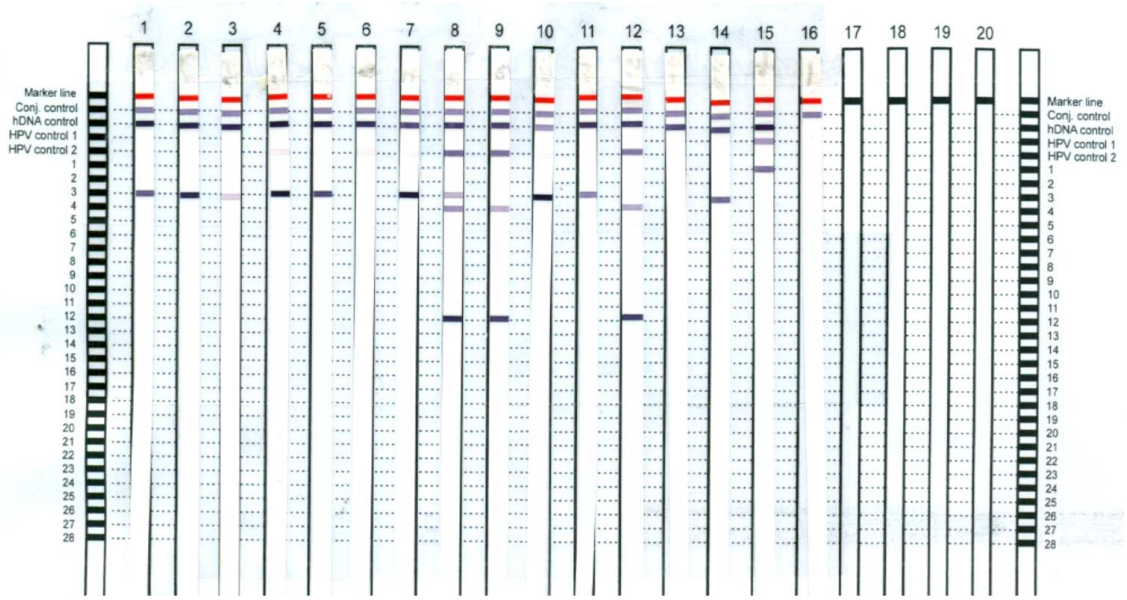

Figure 20. HPV genotyping in fresh samples using INNO-LiPA HPV Genotyping Extra Assay. Lanes 1-14 represent HNSCC fresh samples, lane 15 is a positive control strip, and lane 16 represents negative control strip. First and last lanes represent marker lanes, provided to decipher HPV genotype according to the hybridization pattern.

Table 6. Patient characteristics of the study population for fresh cancer samples

Variable			HPV DNA+ RNA+	HPV DNA+RNA–	HPV–	Total
			n = 6 (%)	n = 8 (%)	n = 47 (%)	n = 61 (%)
Gender	M		5 (83.4)	7 (87.5)	35 (74.5)	47 (77.1)
	F		1 (16.6)	1 (12.5)	12 (25.5)	14 (22.9)
Tumor Location*	OP	Base/tongue	0	1 (12.5)	6 (12.8)	7 (11.5)
		Tonsil	3 (50)	1 (12.5)	2 (4.3)	6 (9.8)
		OP wall	0	0	2 (4.3)	2 (3.3)
	O	Gingiva	1 (16.67)	2 (25)	12 (25.5)	15 (24.6)
		Mouth/floor	1 (16.67)	0	11 (23.4)	12 (19.7)
		Tongue	0	4 (50)	6 (12.7)	10 (16.4)
		Retromolar	1 (16.66)	0	6 (12.7)	7 (11.4)
		Buccal mucosa	0	0	2 (4.3)	2 (3.3)
Age group	<45y		1 (16.6)	1 (12.5)	0	2 (3.3)
	45-64y		2 (33.4)	4 (50)	26 (55.4)	32 (52.4)
	>65y		3 (50)	3 (37.5)	21 (44.6)	27 (44.3)
	Median		62.5	59	62	62
Life style factors**	NSND		2 (33.4)	3 (37.5)	17 (36.1)	22 (36)
	S		0	1 (12.5)	3 (6.3)	4 (6.6)
	SD		4 (66.6)	4 (50)	27 (57.4)	35 (57.4)
Tumor stage 7th ed.	T	1	1 (16.6)	1 (12.5)	4 (8.5)	6 (9.8)
		2	0	2 (25)	16 (34)	18 (29.5)
		3	0	3 (37.5)	12 (25.6)	15 (24.6)
		4	5 (83.4)	2 (25)	15 (31.9)	22 (36)
	N	0	2 (33.3)	4 (50)	21 (44.7)	27 (44.3)
		1	2 (33.3)	1 (12.5)	14 (29.8)	17 (27.9)
		2	2 (33.4)	2 (25)	7 (14.9)	11 (18)
		3	0	1 (12.5)	5 (10.6)	6 (9.8)
Overall stage	Early	I	0	0	2 (4.2)	2 (3.3)
		II	0	2 (25)	12 (25.6)	14 (22.9)
		III	1 (16.6)	1 (12.5)	12 (25.6)	14 (22.9)
	Late	IVa	5 (83.4)	4 (50)	16 (34)	25 (41.1)
		IVb	0	1 (12.5)	5 (10.6)	6 (9.8)
Tumor grade	1		0	4 (50)	19 (40.4)	23 (37.7)
	2		4 (66.6)	2 (25)	19 (40.4)	25 (41.1)
	3		1 (16.7)	2 (25)	5 (10.6)	8 (13.1)
	Unknown		1 (16.7)	0	4 (8.6)	5 (8.1)

*NSNS = no smoking, no drinking; S = smoking; SD = smoking and drinking; **OP = oropharyngeal tumor; O = oral tumor

4.2 Oncogene expression in HPV-positive HNSCC

HPV transcriptional activity was determined by the expression of the most prevalent spliced form E6*I of the oncogenic E6 mRNA encoded by the HPV-16 type. We chose to asses HPV-16 spliced E6 mRNA as HPV-16 predominates in HNSCC accounting for more than 80% cases (Taberna et al. 2017). Using primers for the E6*I mRNA, PCR was performed on only HPV DNA positive samples and the results are presented in **Figures 21 and 23** for archival samples and fresh samples, respectively. Samples containing spliced E6*I mRNA form resulted in PCR amplification of approximately 86 bp amplicon, while samples that contained unspliced (full length) HPV-16 E6 mRNA had approximately 248 bp amplicon.

4.2.1 E6 mRNA in FFPE HNSCC samples

All HNSCC samples that have been positive for HPV DNA (n = 45), except of the two recurrent tumors, have been included in the E6 mRNA analysis, which makes 43 samples in total. RNA was extracted and successfully converted to cDNA. **Figure 21A**, shows the detection of amplified β-actin gene, which served as the quality control of the RT reaction. From 43 samples, only one sample tested negative for the β-actin gene. That sample was the sample containing a small portion of tumor tissue and low RNA concentration and was excluded from further analysis. In **Figure 21B** are shown samples expressing the spliced E6*I form, full length and both forms.

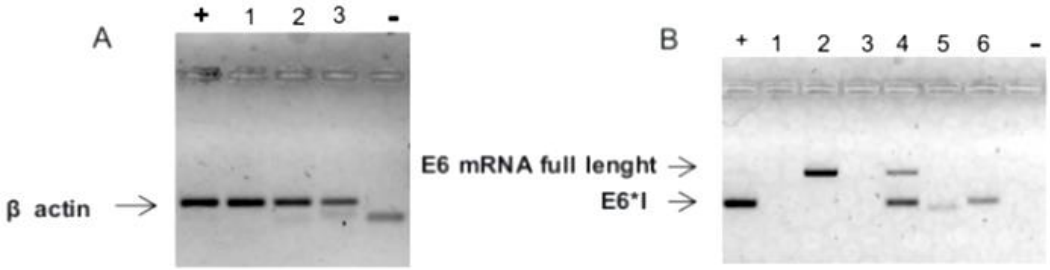

Figure 21. HPV activity analysis. Representative PCRs using **A)** amplified β-actin gene which served as a quality control of cDNA. + = positive control, lanes 1-3 = selected tumor FFPE samples, – = negative control. PCR product of β-actin gene is 99 bp. **B)** PCR product of the E6 mRNA (full length) and E6*I (spliced) amplified from HPV+ patients' cDNA. + = positive control, lanes 1-6 = selected tumor FFPE samples, – = negative control

Transcriptionally active HPV is associated with 11.3% (13 out of 115), while inactive HPV is associated with 27.8% (32 out of 115) archival HNSCC samples. Examination of different forms of HPV RNA (full length and E6*I) in tumors showed increased number of samples with E6 expression in relation to tumor grade, but without any statistical significance (Chi-Square, *P* = 0.564). (**Figure 22**). When analyzing correlation of the tumor stage with the total HPV status, there was a weak positive correlation between the presence of the spliced E6*I mRNA and the tumor stage, but without statistical significance (r = 0.24; Chi-Square, *P* = 0.818). When analyzing the correlation with smoking and drinking habits, there was no linear relationship between the two variables (r = 0.01). Furthermore, there is no statistical correlation between the age of patients and the HPV activity, with r = -0.12 (data not shown).

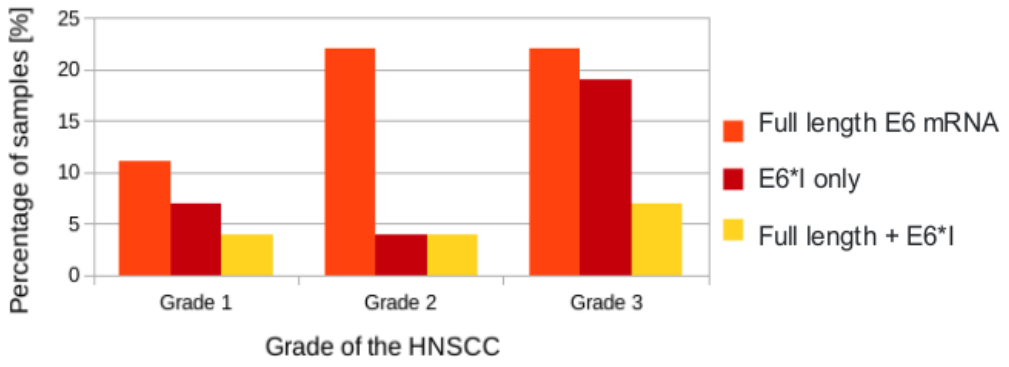

Figure 22. Percentage of the full length E6*I spliced variant and both forms in HPV+ FFPE samples depending on the HNSCC grade

4.2.2 E6 mRNA in freshly collected HNSCC samples

From a total of 14 HPV DNA positive samples, 2 samples were HPV-18 positive, and thus were not included in the HPV-16 E6*I analysis, leaving 12 samples in total. All 12 samples that were positive for HPV-16 have been reverse transcribed and analyzed for the HPV-16 E6*I mRNA presence (**Figure 23**). Six samples tested positive for spliced E6*I, while others had full length E6 mRNA amplified in the PCR, hence, in fresh cancer samples, active HPV was found in 9.8% (6 out of 61), while inactive HPV was associated with 13.1% (8 out of 61) HNSCC.

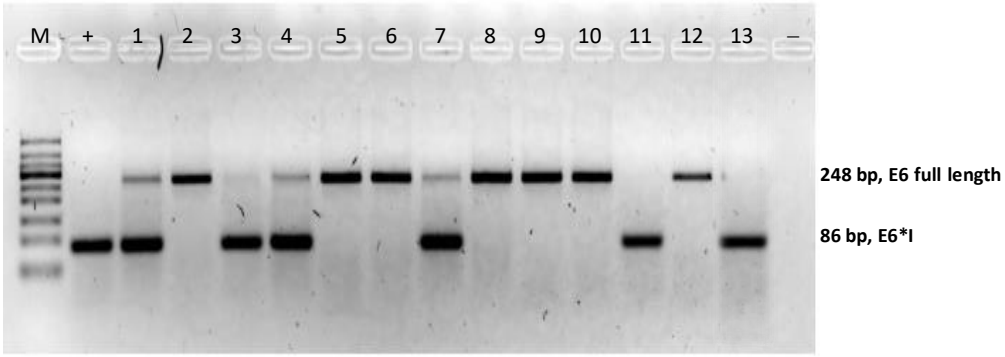

Figure 23. E6 mRNA (full length) and E6*I spliced form, amplified from HPV+ fresh cDNA. M = GelPilot50 marker (Qiagen, USA); + = positive control; lanes 1-13 = fresh tumor samples; – = negative control.

4.3 Survival analysis

Overall, HPV analysis (HPV DNA presence and oncogene expression analysis) in HNSCC (both archival and fresh cancer samples) in the Croatian population provided an important data on cancer characteristics in our setting. The characteristics for overall FFPE and fresh tumor samples are as follows: majority of patients were males (83.5% and 76.6%, respectively), and most of the patients were treated for advanced stage disease (78 and 75%). On average, the patients were older in both set of samples, regardless of HPV positivity, hence in the FFPE setting, median age was 59y (24-91y), while in fresh setting 62y (31-85y). After comparison of the three groups regarding the HPV status, there was no statistical difference in terms of gender (Chi-square, $P = 0.99$ in FFPE and $P = 0.94$ in fresh tumors) and age (t-test, $P = 0.89$ in FFPE, $P = 0.74$ in fresh) between the three groups of patients (HPV DNA and RNA+, HPV DNA+ RNA– and HPV DNA–) in both archival and fresh set of samples. Most of the patients were smokers and drinkers, with an overall history of smoking for FFPE 24.3% and for fresh 57.4%, while 33% of archival samples stated heavily and occasionally alcohol consume, and 57.4% report drinking on daily base in fresh setting.

When comparing life style habits in the groups based on HPV status, results were as follows: in archival samples, out of 13 HPV+ samples, there were no heavy drinkers and one was an active smoker (7.7%) with only 3 (23%) nonsmokers and 2 (15%) drinkers. In addition, within the 32 HPV transcriptionally inactive samples, there were 9 (28.2%) active smokers and 6 (19%) nonsmokers, while 3 (9%) heavy drinkers and 7 (22%) nondrinkers. Out of 70 HPV– samples, 18 (26%) patients have been declared as active smokers, and only 6 (8%) nonsmokers, while 6 (8%) heavy drinkers and 7 (10%) non-drinkers. As presented in **Table 5**, overall the most common type of therapy in the FFPE subset was the modality of surgery and radiotherapy (24.3%), followed by surgery only (23.5%) and surgery with the addition of chemo- and radiotherapy (12.2%). Within patients with inoperable cancer, 5.2 % patients have been treated with radiotherapy only, 3.5% with the combination of both chemo- and radiotherapy, while only 1 patient (0.9%) was treated with chemotherapy only.

After comparing cancer samples in the Croatian population based on patients' biological, socioepidemiological, clinical and histopathological characteristics, we aimed to analyze overall survival analysis in both settings.

4.3.1 Survival analysis on archival FFPE samples

Not all HNSCC cases were suitable for the survival analysis. Overall, 2 HPV+ and 7 HPV– samples have been excluded, due to samples being recurrent tumors, while for 9 samples the patient's follow-up data was unavailable. In addition, 5 samples have been excluded from survival analysis due to receiving only palliative treatment. Furthermore, the disease specific mortality data was absent. Thus, all-cause mortality outcomes were assessed for a total of 92

patients. Median follow-up for up to December 31st, 2017 was 23.83 months, and overall 5-year survival was 32.0%. Kaplan-Meyer survival curves are presented in **Figures 24-26** according to patients' biological characteristics, clinical parameters, type of therapy and two different tumor staging approaches. Moreover, samples have been classified using combined risk level model (Ang et al. 2010), which takes into account different patient aspects, as suggested by Ang *et al.* (2010). The original method included classification by HPV status, smoking and T and N stages for HPV+ and HPV– tumor samples, respectively. In this study, smoking was disregarded for grouping patients according to risk factors, due to high proportion of missing data (53.1% unknown smoking history).

Evidently, gender and alcohol consumption does not seem to influence survival (P = 0.25 and P = 0.096, respectively). Only two variables (**Figure 24**) which seems to influence overall survival are the age at the time of diagnosis and smoking history; however, the difference was not significant (P = 0.12 and P = 0.50, respectively). Surprisingly, HPV status and p16 have no influence on survival. HPV-positive and HPV-negative patients, regardless of the method for the HPV detection (DNA, RNA or p16) had very similar survival (P = 0.522 and P = 0.641, respectively).

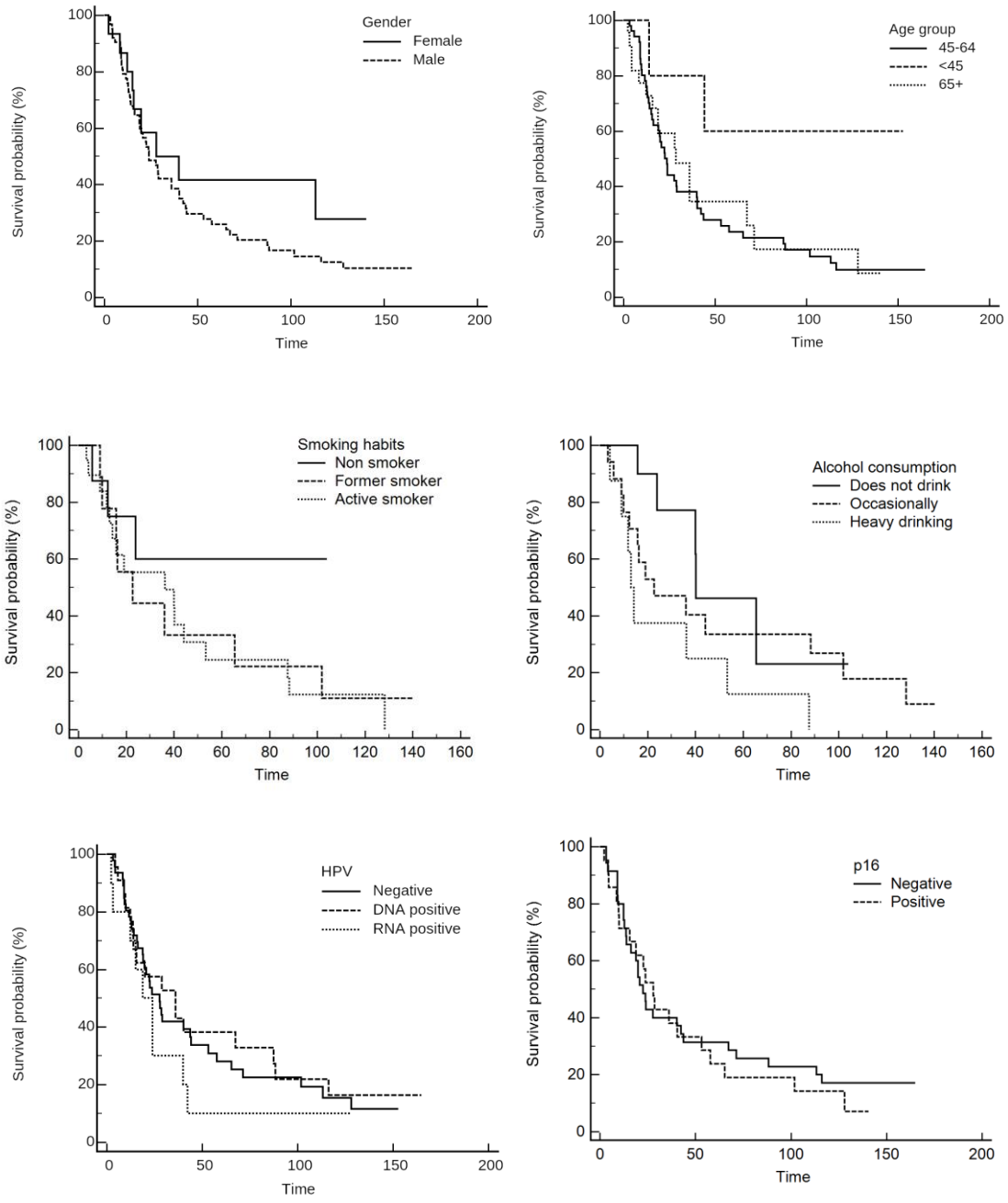

Figure 24. Overall survival (in months) of HNSCC patients based on gender, age group, life style habits (smoking history and alcohol use), HPV positivity and p16 data.

Furthermore, of clinical parameters (**Figure 25**), it is evident that N stage significantly affected the overall survival (P = 0.05), while T stage had no significan effect on survival (P = 0.115). In addition, for more accurate calculations, the survival rate of patients based on 7th and the most recent 8th edition of AJCC and the UICC cancer staging was compared. After applying the 8th edition, where staging of HPV+ cancers is less severe, the difference in survival between the different tumor stages is more evident. Nevertheless, those tumors with late stage (III and IV) had worse survival than those with early stage (I and II) in both cases, and the difference was significant after applying both staging editions (P = 0.011 and P = 0.009, respectively).

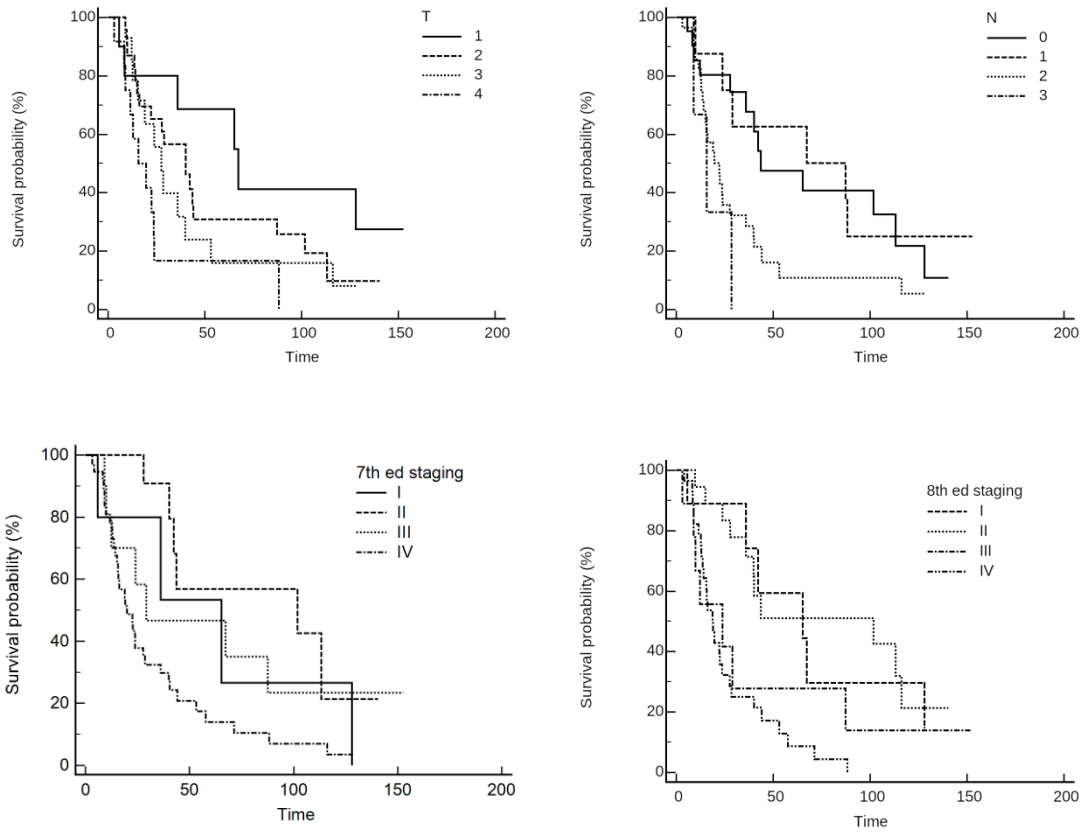

Figure 25. Overall survival (in months) of HNSCC patients based on the T and N stage, 7th edition and 8th edition of tumor staging

After analyzing the tumor grade and the type of therapy (**Figure 26**), it is evident that tumor grade had no influence on survival (*P* = 0.707), while therapy was significantly affecting overall survival of HNSCC patients (*P* = 0.046). Moreover, classifying patients based on the combined risk level has reached statistical significance (HPV status and T/N stage; *P* = 0.0079).

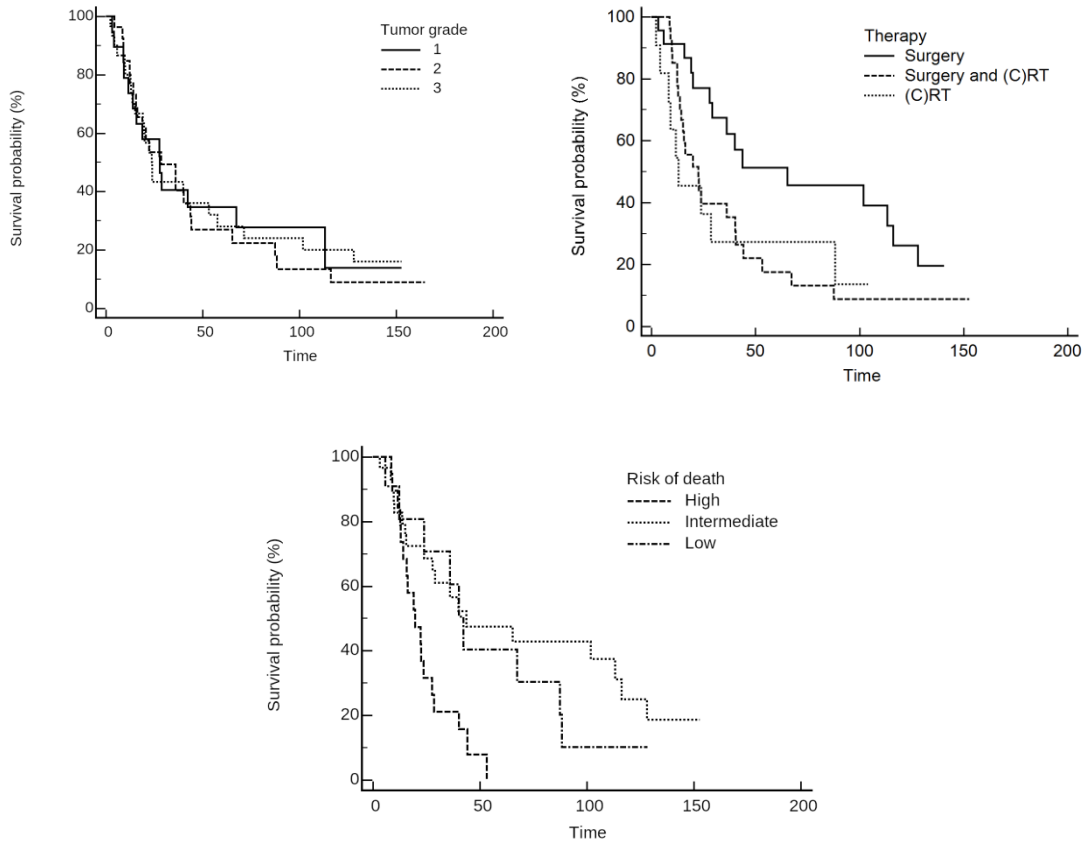

Figure 26. Overall survival (in months) of HNSCC patients based on the tumor grade, type of therapy and combined risk level (HPV status and T/N stage)

4.3.2 Survival analysis on fresh samples

As with the retrospective survival analysis of FFPE, not all fresh tumor samples were suitable for the overall survival analysis. For survival analysis, 8 patients with recurrent tumors have been excluded, leaving a total number of 53 included patients with oral and oropharyngeal cancer, that have been diagnosed between 2015 and 2018. As already mentioned, patients' survival time in months was calculated from the date of earliest diagnosis to registered time of death (all causes of death) or last registered follow-up. Kaplan-Meyer survival curves are presented in **Figures 27-29** according to patients' biological characteristics, clinical tumor parameters, tumor staging and combined risk level. Since samples have been collected from the Clinic of maxillofacial surgery of the Clinical Hospital Dubrava, medical records of other types of therapy (besides surgery) were not available. Also, for patients with overly large tumor proportions, tumors were inoperable and adequate only for radiotherapy, hence, pathological data were missing, and only clinical TNM status was used. For the tumor staging, the 7th edition of AJCC was used. In addition, there was no data on p16 positivity, therefore survival analysis based on the type of therapy and p16 positivity was done.

After assessing survival of patients based on patients' biological characteristics (**Figure 27**), there was an indication that survival might be affected by gender and age group, however, statistical significance was not reached (P = 0.456, P = 0.139, respectively). In addition, among 5 deceased patients, there was only 1 patient who was younger than 45 years, hence influencing survival curves. Regarding life style history (smoking and alcohol intake), there was no missing data, however, the data were less informative, giving the only basic information if patient is smoking or drinking or not. Detailed smoking or drinking history was unavailable i.e. how long the patients smoked/consumed alcohol, how many packages of cigarettes per day, what sort of alcoholic beverages have been consumed, or if the patients are former smokers/drinkers. After the analysis, patients who smoked and consumed alcohol showed better survival, than those who were declared as non-smokers and non-drinkers, but there was no statistical significance (P = 0.258, P = 0.444, respectively).

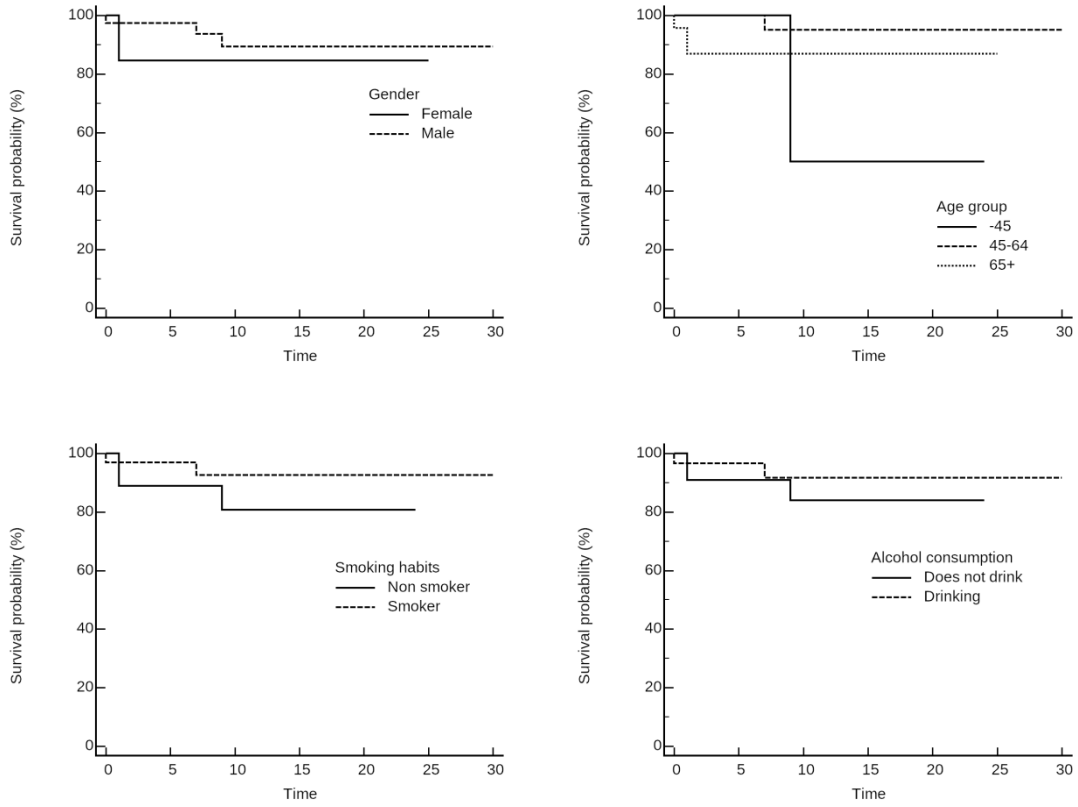

Figure 27. Overall survival (in months) of HNSCC patients based on gender, age group, and life style habits (smoking history and alcohol use).

Histopathological data on tumor grade, T, N, M status and information on tumor invasion were obtained and survival analysis based on those parameters have been performed. **Figure 28** represents survival based on HPV positivity (HPV DNA and HPV RNA testing), tumor grade and T, N status. Even though survival curves might indicate that HPV status (P = 0.294), tumor grade (P = 0.576) and T status (P = 0.156) influence patients' survival, statistical significance was reached only for the influence of N status (P = 0.005) on survival.

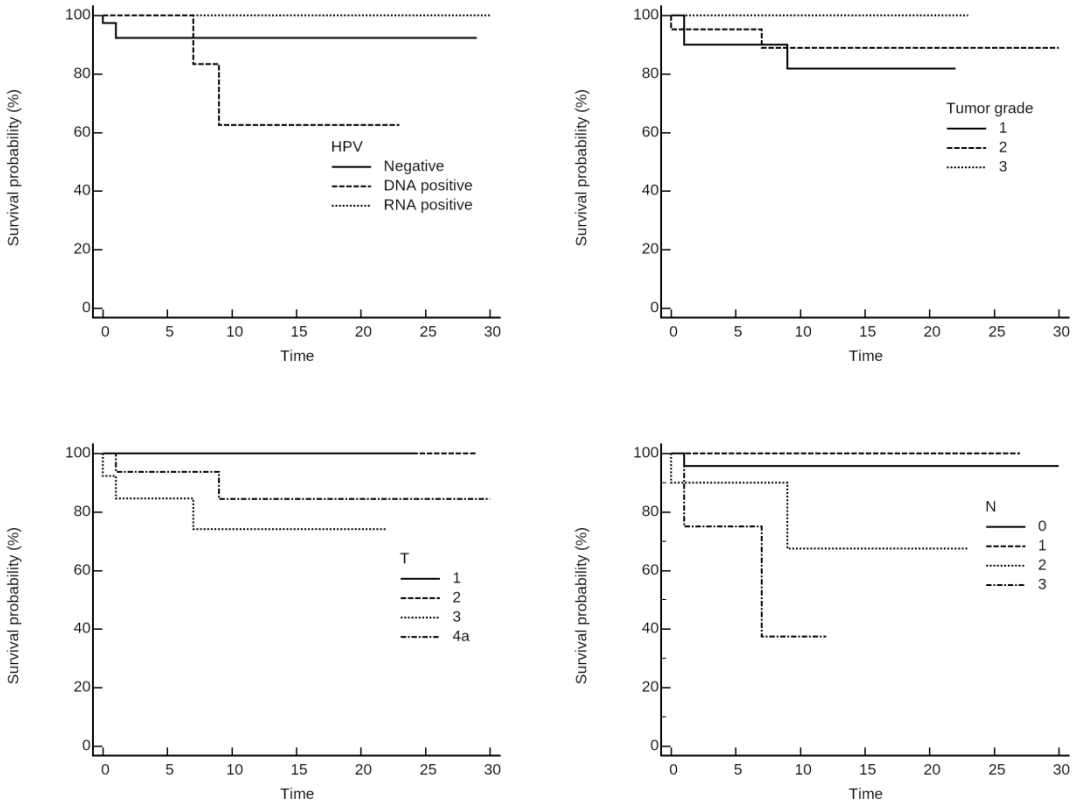

Figure 28. Overall survival (in months) of HNSCC patients based on the HPV positivity, tumor grade, T and N stage.

As with the archival samples, fresh tumor samples have been classified into high, intermediate or low risk tumors, based on HPV status, smoking history and T and N stages (combined risk level (Ang et al. 2010). Since there was no missing data on patients' smoking habits, smoking was also included according to the original classification. However, after assessing survival based on the tumor stage, tumor invasion and risk of death, analysis showed no statistical significance (P = 0.474, P = 0.168, P = 0.116, respectively).

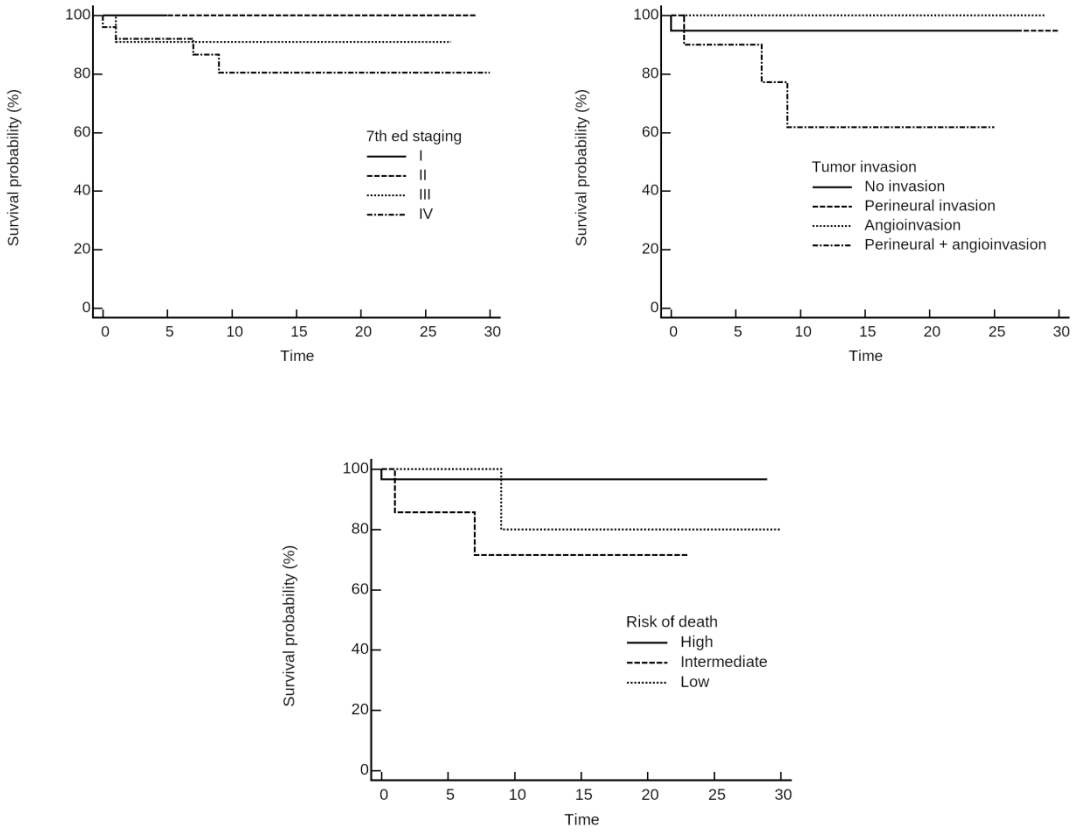

Figure 29. Overall survival (in months) of HNSCC patients based on the 7th edition of staging, tumor invasion and combined risk level (HPV status, smoking history and T/N stage)

4.4 Epigenetic changes in HNSCC and control samples

Epigenetic modification studies have been performed on fresh samples solely due to the suboptimal nucleic acid quality after formalin fixation in FFPE tissues. A total of 61 fresh oral/oropharyngeal HNSCC samples, in addition with 3 oropharyngeal cancers from tonsils, provided from dr. Ruth Tachezy (Czech Republic) have been included in the miRNA profiling and have been compared to 3 normal controls (nonmalignant tonsils).

A total of 10 HNSCC and 6 buccal swabs from healthy adults have been included in the whole-methylation profiling. A total of 16 HNSCC samples have been used in pyrosequencing experiment. Results are presented in the following sections.

4.4.1 MicroRNA profiles in HNSCC and control samples

Characteristics of all samples included in the miRNA analysis are presented in **Table 7**. In total, together with 61 prospectively collected fresh HNSCC samples from KBD Dubrava, Zagreb, total number of samples adequate for miRNA profiling was 64. Quality of the RNA isolated from samples was assessed on an agarose gel electrophoresis (**Figure 30**) and on Agilent 2100 Bioanalyzer (**Figure 31**).

Table 7. Characteristic of fresh HNSCC and normal controls included in NGS study of miRs

NGS Index	Sample Group	Gender	Age (years)	Tumor stage	Tumor grade	Smoking/ Drinking
1	O+	M	59	IVa	2	1
2	O-	M	61	II	2	1
3	OP+	M	66	IVa	2	1
4	OP+	F	32	III	?	0
5	O-	M	66	IVa	2	1
6	O-	M	61	III	2	1
7	OP-	M	53	IVb	/	1
8	O-	M	52	II	2	1
9	O-	M	53	IVc	1	1
10	OP-	F	55	III	1	1
11	O+	M	31	IVa	1	0
12	OP-	M	56	IVa	2	1
13	O-	M	56	III	1	1
14	OP+	M	64	IVb	2	1
15	O+	M	69	IVa	3	1
16	OP-	F	57	IVa	2	1
17	N	F	59	0	0	1
18	N	F	59	0	0	1
19	N	F	59	0	0	1
20	OP+	F	73	IVb	2	0
21	OP+	M	59	IVb	2	1
22	OP+	M	55	IVa	3	1

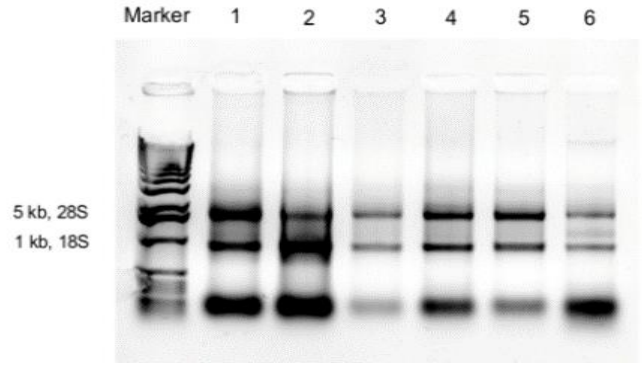

Figure 30. Agarose gel electrophoresis of the RNA isolated from fresh tumor samples. RNA from 9 HNSCC fresh tumors was visualized on a 1% agarose gel. M = DNA Molecular Weight Marker IX (Roche Diagnostics; Germany); lanes 1-6 = RNA isolated from 6 selected tumor samples.

Expected RNA peaks, as shown in **Figure 31A** are also visible in our representative sample (**Figure 31B),** which shows very high quality of the extracted RNA, with RIN number of 8.80. As expected, visible peaks come from miRNAs, 18s rRNA and 28s rRNA.

Briefly, total RNA was ligated with RNA adapters on both 5' and 3' ends that enable further amplification of tagged molecules with specific primers. After following the manufacturer's instructions in the process of miR library generation, specific miR band (together with adapters approximately 150 bp in size) were cut out from the gel and purified, and the quality of DNA from amplified miR libraries was also assessed on Agilent 2100 Bioanalyzer (**Figure 32**). Electropherograms of the generated DNA libraries showed specific miRNA peaks and the data were presented qualitatively and quantitatively, which was necessary in order to prepare dilutions for further NGS analysis. To analyze miRNA differences in expression profiles, a total of 22 next generation sequencing libraries were made from: 6 HPV+ (DNA and RNA) and 4 HPV– oropharyngeal (OP+ and OP–) cancer samples; 3 HPV+ (DNA and RNA) and 6 HPV– oral cancer samples (O+ and O–); and 3 normal tonsil tissue samples. Complete sample annotation table is provided **Table 7**.

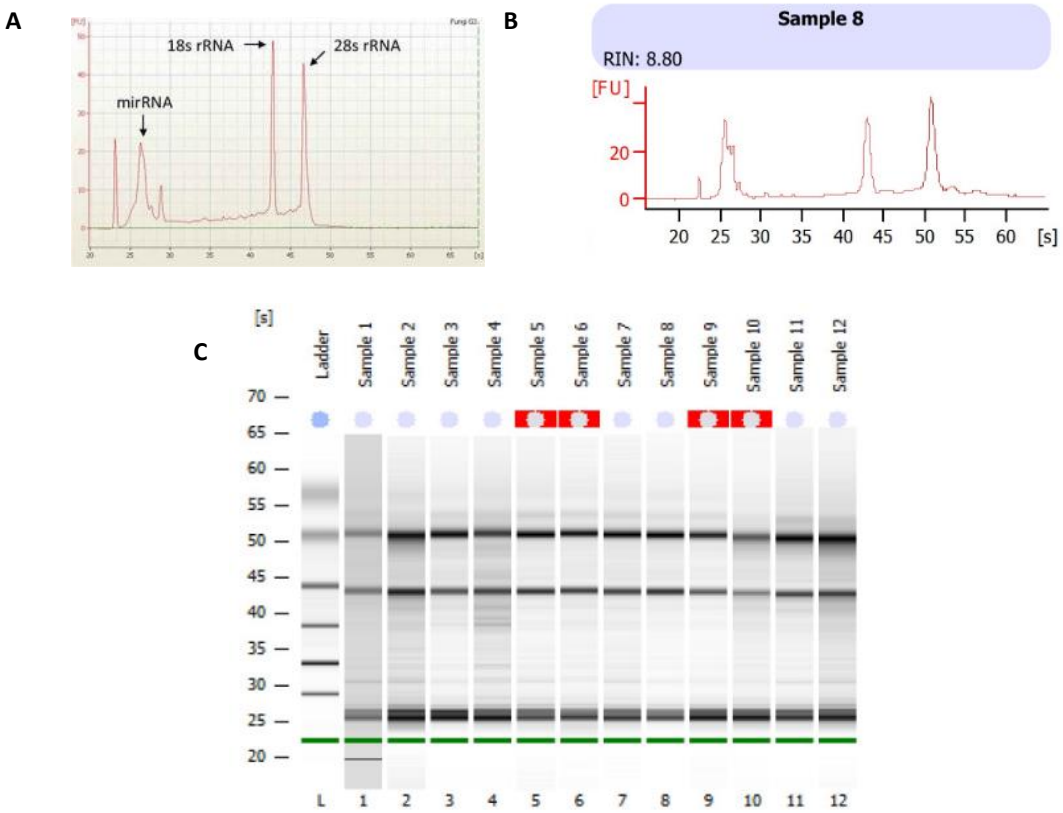

Figure 31. Total RNA analysis performed on the Agilent 2100 Bioanalyzer.
A) Specific RNA peaks (NORGEN 2018); **B)** Electropherogram of the RNA extracted from fresh HNSCC tissue; **C)** Chip electrophoresis of the RNA extracted from 12 HNSCC samples.

After miRNA libraries quality and quantity control, NGS experiment was performed on NextGen 500 machine. The run data was accessed through Basespace platform (Illumina, USA). Results were of high quality: the sequencing run generated approximately 110 million reads that passed quality control (QC) filter. A total of 96.4% bases had a score of Q30 or higher. Raw sequence reads were trimmed of remaining adapter sequences using the FastQC app on the Basespace platform, while preliminary statistical analysis was performed with the SmallRNA app, which was used to obtain alignments, counts, normalized counts and to perform initial DESeq2 differential expression analysis. Raw sequencing read data was deposited to the ArrayExpress (www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-7030. Subsequently, count data was imported to the R software and analyzed using the DESeq2 package to obtain customized plots. Oropharyngeal HPV+ (OP+) samples, oral HPV+ (O+), oropharyngeal HPV– (OP–) and oral HPV– (O–) tissues have been analyzed for the differential expression in comparison to normal control non-malignant tissue. As shown in **Figure 33,** most of the sequenced reads arise from small RNAs that are ~22 bp in size, with miRNAs being mostly sequenced in each sample. This analysis also revealed high percentage

of sequenced isomiRs. Sample distance and PCA plots indicated that one sample behaved as outliers and have been excluded from subsequent analysis.

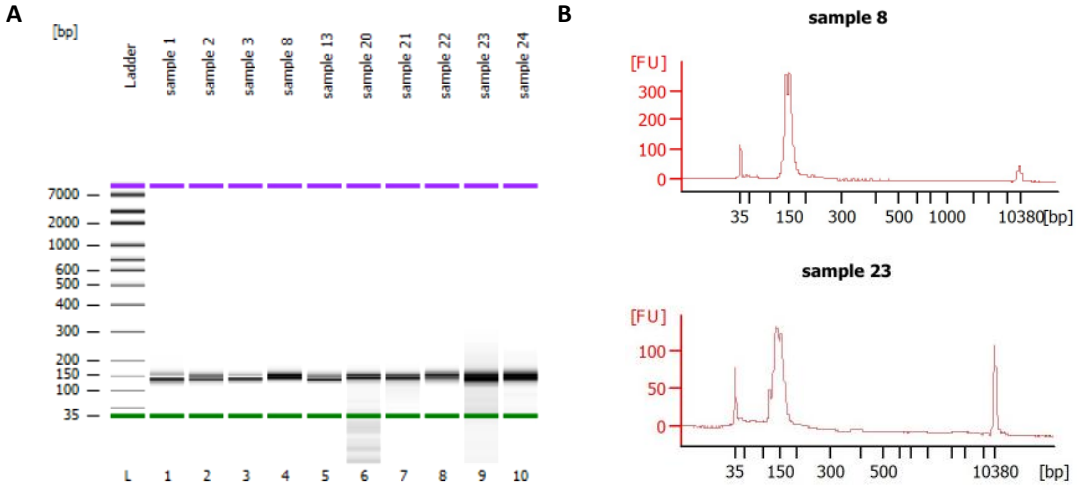

Figure 32. DNA from generated miR libraries analyzed on the Agilent 2100 Bioanalyzer. A) Electrophoresis of the generated DNA libraries from 10 samples;
B) Electropherograms of the generated DNA libraries shown on two samples.

Final analysis revealed 552 different unique miRNA sequences with adjusted *P* value < 0.05, which were significantly deregulated in 1119 comparisons of all four sample groups. Several most significantly relevant miRs that have been deregulated in cancer *vs.* normal are presented in **Table 8**. However, of the 552 unique sequences only 108 fully corresponded to the mature miRNA sequences, while the rest were corresponding to isomiRs (**Figure 33**), either with sequence mismatches or with slightly different sizes.

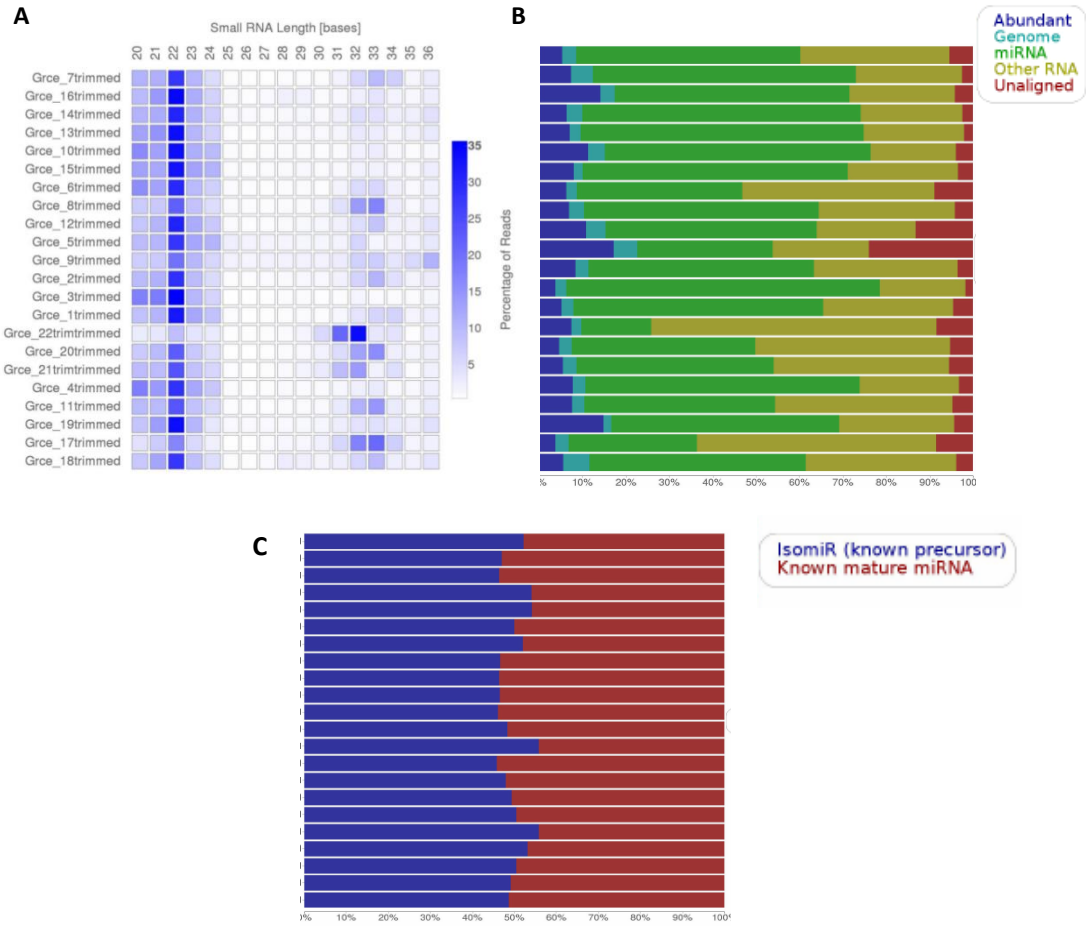

Figure 33. Preliminary quality control statistics of NGS results. A) Small RNA length distribution; **B)** Hits by category; **C)** miRNA hits by category.

Based on the final results of miR profiling, samples did not form distinguishing clusters on heatmap (**Figure 34A**) or the principal component analysis (PCA) plots (**Figure 34B**), however, clear separation of normal control samples is evident. In addition, after analyzing oropharyngeal cancer samples *vs.* control samples (**Figures 35A, 35B**), and oral cancer *vs.* controls separately (**Figures 35C, 35D**), relevant subgroups become slightly more separated on the same heatmap and PCA plots, but more efficiently in oropharynx. In this way miR profiling could separate HPV+ and HPV– cancer tissues at least in oropharyngeal cancer. Comparison of the significantly deregulated miRs in each sample subgroup have been performed using Venny 2.1 tool (**Figure 36**). Of all significantly deregulated miRNAs only 77 were specific to oropharyngeal HPV+ samples, while 3 shared with HPV+ oral cancer group. However, of those 80 significantly deregulated miRNAs, found in HPV+ tumors, only 16 corresponded to mature miRNAs (miR-9-5p, 25-5p, -29a-3p, -29b-3p, -34a-5p, -93-5p, 106b-5p, -133a-5p, -133a-3p, -139-5p, -140-5p, -147b, -208b-3p, 210-5p, 328-3p, -1307-3p)

while the rest corresponded to different isomiRs of 42 individual miRNAs.

Table 8. Subset of differentially expressed miRs in HNSCC samples *vs.* normal controls

miR ID	FC	*P* value
miR-9-5p	9.051	0.0002
miR-27b-5p	5.398	0.0020
miR-26b-5p	-1.961	0.0001
miR-25-5p	9.031	0.0008
miR-96-5p	6.704	0.0002
miR-31-5p	8.164	0.0004
miR-29c-3p	-2.532	0.0072
miR-29a-3p	-2.021	0.0079
miR-222-3p	2.846	0.0046
miR-221-3p	4.031	0.0196
miR-21-3p	9.563	0.0000
miR-191-5p	1.706	0.0268
miR-187-3p	6.900	0.0076
miR-151a-3p	2.697	0.0107
miR-143-3p	-2.810	0.0237
miR-145-5p	-4.179	0.0076
miR-100-5p	-3.274	0.0006

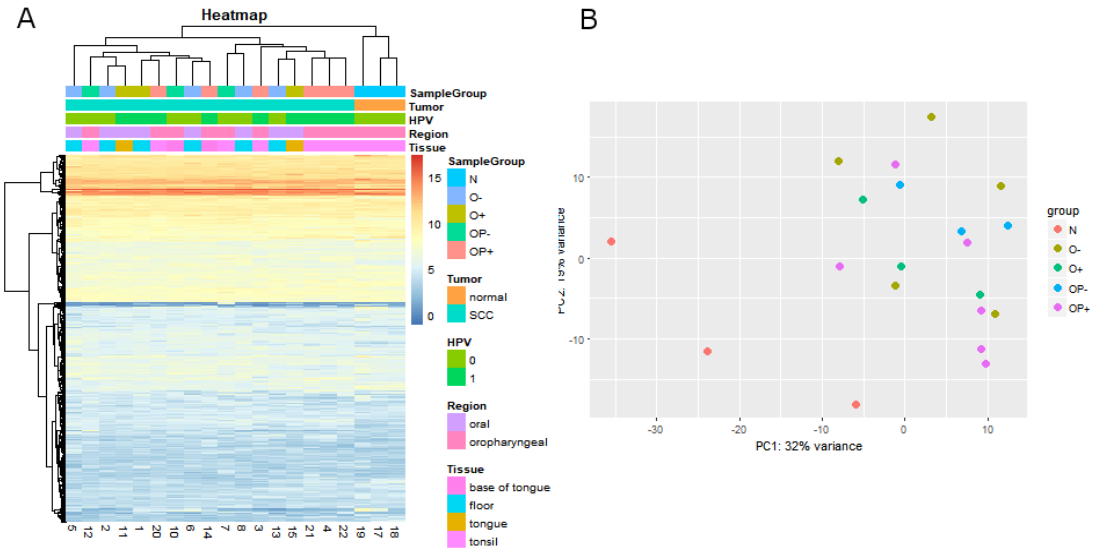

Figure 34. Clustering of the study population on heatmap (A) and PCA (B). SCC = squamous cell carcinoma; N = control normal (non-malignant tonsil samples); O– = oral HPV-negative cancer samples; O+ = oral HPV-positive cancer samples; OP– = oropharyngeal HPV-negative cancer samples; OP+ = oropharyngeal HPV-positive cancer samples.

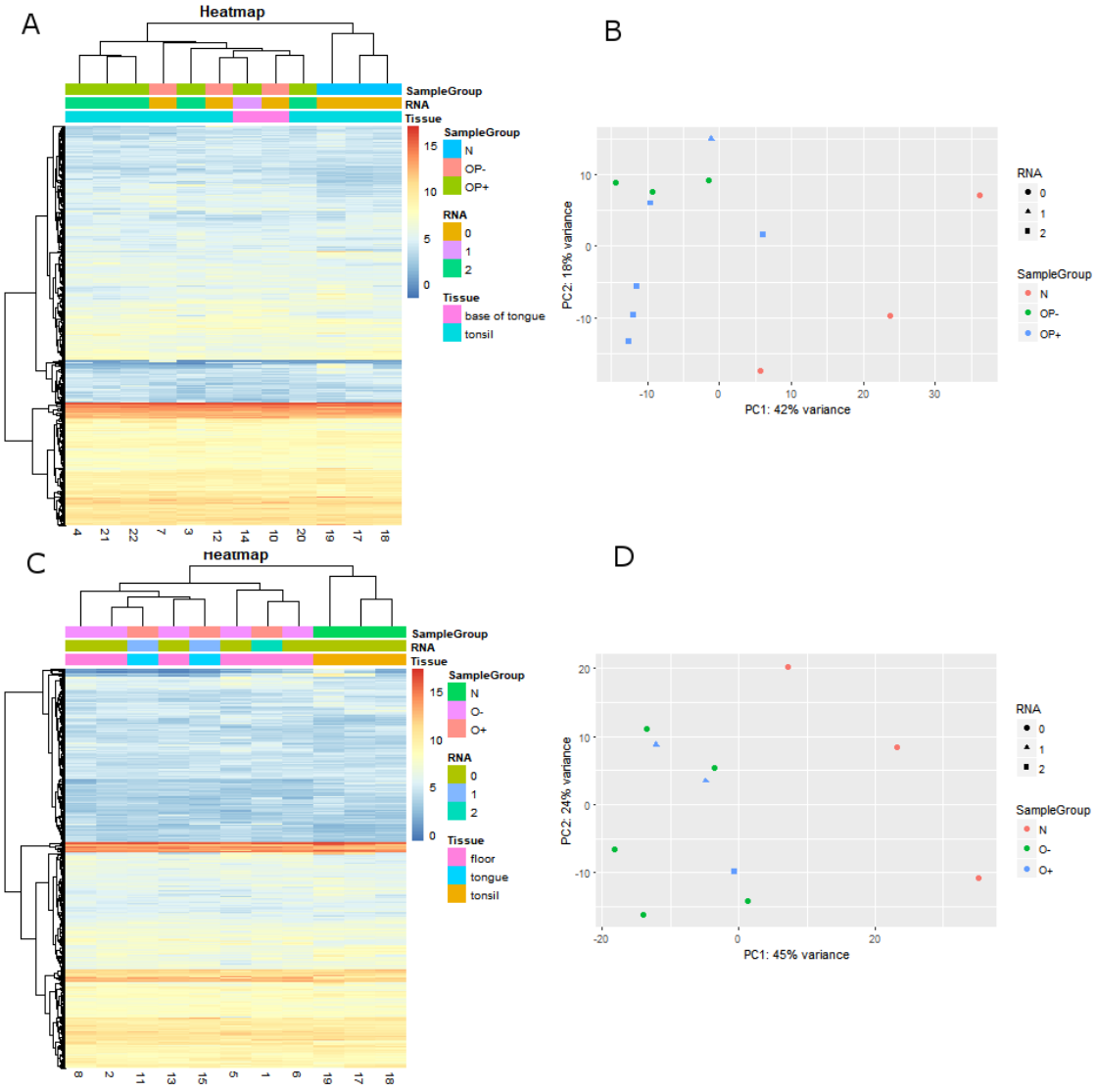

Figure 35. Clustering of the samples divided into oropharyngeal (OP) and oral (O) subgroups (panels AB, and CD) using heatmap (A & C) and PCA (B & D) plots. SCC = squamous cell carcinoma; N = control normal (non-malignant tonsil samples); OP- = oropharyngeal HPV-negative cancer samples; OP+ = oropharyngeal HPV-positive cancer samples; O– = oral HPV-negative cancer samples; O+ = oral HPV-positive cancer samples

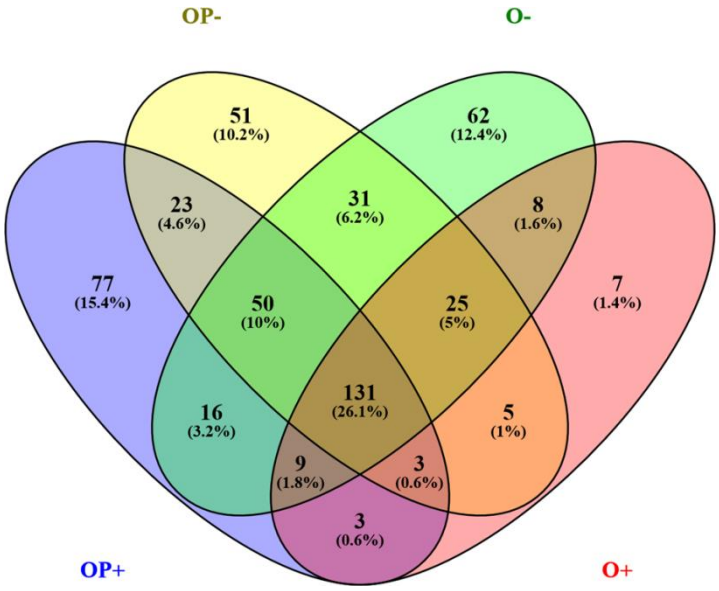

Figure 36. Venn diagram of significantly differentially deregulated miRNA sequences across sample groups. OP+ = oropharyngeal HPV-positive cancer samples; OP- = oropharyngeal HPV-negative cancer samples; O+ = oral HPV-positive cancer samples; and O- = oral HPV-negative cancer samples.

For technical validation of the miRNA identified in NGS sequencing, we have chosen several miRNAs widely known to be deregulated in cancer (two upregulated miR-21-3p, miR-31-5p and one downregulated miR-100-5p) as well as less studied miRs (upregulated miR-9-5p, -27a-5p, -34a-5p, -222-3p and downregulated miR-143-3p, -145-5p, -218-5p). The average Ct values of the three miRs (miR-16-5p, -191-5p and -181a-5p) were considered as the reference for the fold change calculations. Specific miRNAs with potential implications in HPV+ HNSCC have been selected and their levels of expression assessed in the whole set of clinical samples. The following miRNAs were chosen: miR-9-5p, -29a-3p, -100-5p, -106b-5p, -143-3p, -145-5p and -199b-5p. Furthermore, globally relevant miR-21 was included in the analysis as an internal control, since miR-21 is widely identified as an oncogene miRNA, hence it is overly expressed in almost all cancer types. Same referent miRS were used as before. The analysis confirmed NGS results in the oropharyngeal HPV+ subset for mir-9 (P = 0.0015), mir-21 (P<0.0001), miR-100 (P = 0.0026), miR-144 (P = 0.047), miR-145 (P = 0.0048) and miR-199b (P = 0.048) (**Figure 37**). However, miR-29a (P = 0.065) and miR-106b (P = 0.902) were not identified as significant in the complete pool of samples.

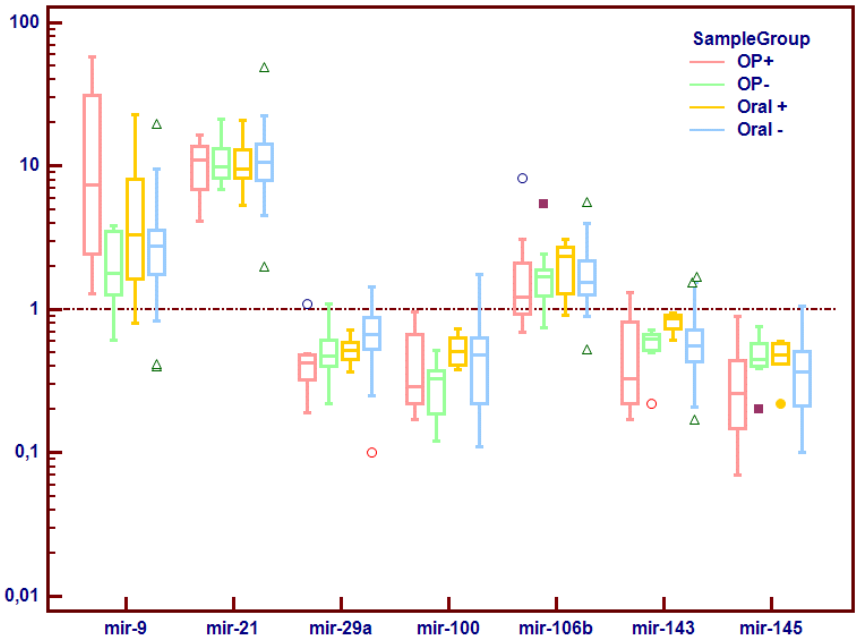

Figure 37. qRT-PCR analysis of the selected miRNAs in all clinical samples. OP+ = oropharyngeal HPV-positive cancer samples; OP– = oropharyngeal HPV-negative cancer samples; Oral+ = oral HPV-positive cancer samples; and Oral – = oral HPV-negative cancer samples. Horizontal referent line is set at fold change 1 and indicates relative expression of the normal tissue (used as control). Median, interquartile range and extreme values are plotted as box and whisker plots, while outliers are indicated as individual markers.

4.4.2 DNA methylation in HNSCC and control samples

Characteristics of samples included in genome-wide methylation analysis are presented in **Table 9**. Samples were categorized in old (>50 years) or young age group, and according to HPV status in three groups as already described. Samples positive for HPV DNA and E6*I are noted as 2, those positive for HPV DNA but E6*I negative as 1 and those negative for both HPV DNA and RNA were classified as 0. Median years is 56.5y, while the average is 51y (27 to 73). Most were males (11; 68.8%), and old group was predominant with 11 samples (68.8%). Prior the genome-wide methylation analysis and pyrosequencing, bisulfite conversion has been tested using MSP amplifying unmethylated *hTERT1* gene (Dessain et al. 2000) and methylated *SIGLEC12* (Milutin Gašperov et al. 2014), respectively, as presented in **Figure 38A-B**. All samples that tested positive for *hTERT1* unmethylated and *SIGLEC12* methylated gene were considered suitable for further analysis.

Table 9. Characteristics of samples selected for the WGM analysis.

Sample ID	Sample group	Gender	Age years	Age group*	Location	HPV** status
C1	cancer	M	59	old	oral	2
C2	cancer	M	31	young	oral	1
C3	cancer	M	66	old	oropharyngeal	2
C4	cancer	F	32	young	oropharyngeal	2
C5	cancer	M	73	old	oral	2
C6	cancer	M	61	old	oral	0
C7	cancer	F	55	old	oropharyngeal	0
C8	cancer	M	56	old	oral	0
C9	cancer	M	64	old	oropharyngeal	1
C10	cancer	M	61	old	oral	0
N11	normal	M	53	old	oral	0
N12	normal	F	29	young	oral	0
N13	normal	F	57	old	oral	0
N14	normal	M	58	old	oral	0
N15	normal	F	34	young	oral	0
N16	normal	M	27	young	oral	0

*Those with age below 50 years have been classified into young age group; **HPV status: 0 = HPV-negative, 1 = HPV DNA positive, 2 = HPV DNA and RNA positive

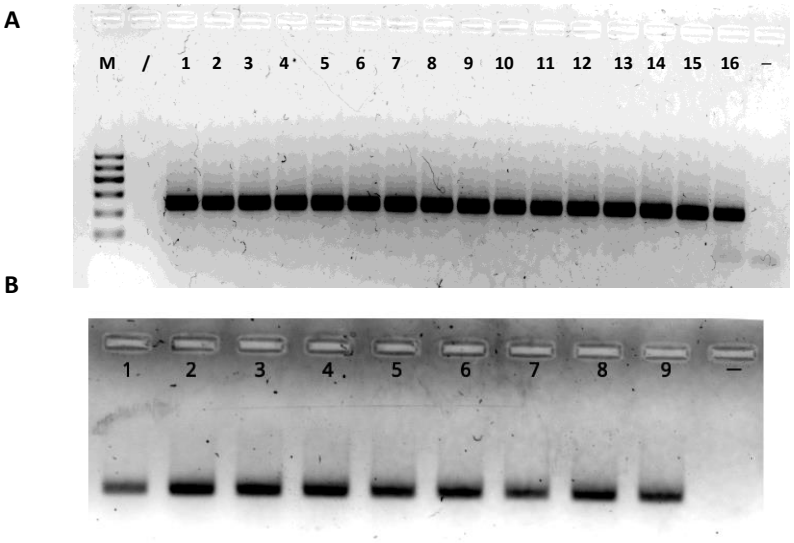

Figure 38. Methyl specific PCR of bisulfite converted DNA served as a quality control from **A)** 16 samples used in whole-genome methylation analysis, *hTERT1* amplicon of 260 bp; M = gel pilot 100 bp marker (Qiagen, USA); / = empty well, 1-10 = cancer samples, 11-16 = normal controls, – = negative control, and **B)** 8 samples used in pyrosequencing analysis, *SIGLEC12*-M amplicon of 117 bp; 1 = methylated control, 2-5 = normal controls, 6-9 = cancer samples, – = negative control

4.4.2.1 Whole genome DNA methylation profiling

Methylome of the cancer samples and normal control samples has been compared and analyzed with the RNBeads software (v 1.2.2.). Findings of DNA methylation profiling are presented in **Figures 39-41** and **Tables 10** and **11**. After excluding 2 cancer samples and 1 normal sample, due to unreliable data as determined by preliminary processing and the quality control steps, total number of samples that qualified for the main analysis was 13. Out of ~ 900,000 total probes used in the array Greedycut algorithm of the RNBeads package found and excluded 49,432 unreliable probes, leaving a total number of probes to 867,926. Final analysis included 22,360,170 reliable measurements. In addition to the CpG sites, there were four sets of genomic covered region: (i) genome tiling regions, which covered 237,795 regions in a dataset, (ii) Ensembl genes (Ensembl Genes 75), which covered 33,658 regions, (iii) promoter regions of Ensembl genes (version Ensembl Genes 75), which covered 42,890 and (iv) CpG island track of the UCSC Genome browser, that covered 25,766 regions in a dataset. Differential methylation analysis was conducted on site and region level according to the sample groups specified in the analysis, and it confirmed 120,901 differentially methylated sites in cancer *vs.* normal. Differential methylation measure (FDR) was adjusted to P value < 0.05.

Scatter plot in **Figure 39** is showing a good grouping of samples with the cancer samples separating at one side and normal samples on the other. Samples showed clustering based on the methylation pattern also on heatmap (**Figure 40**), but after selecting sites and regions with the highest variance across all samples, the difference in methylation pattern between the two groups is more obvious (**Figure 41**). The most significantly methylation differentiated gene promoters between two groups of samples, according to Δβ-value are indicated in **Table 10** (hypermethylated gene promoters in cancer samples *vs.* normal samples) and **Table 11** (hypomethylated gene promoters in cancer samples *vs.* normal samples). Total number of CpGs within specific gene promoter that have been analyzed in the WGM analysis and total number of CpGs within the gene promoter region are also presented in **Tables 10** and **11**. In both cases of hyper- and hypomethylated genes in cancer, DNA methylation difference (Δβ) higher than |0.459| was considered statistically relevant, and those genes were presented.

Table 10. Hypermethylated gene promoter regions in cancer tissues compared with normal tissues

Gene name	Description	No. of analyzed CpGs in WGM	No. of CpGs in promoter region	Δβ	Adjusted *P* fdr value
GPRC5D	G Protein-Coupled Receptor Class C Group 5 Member D	1	32	0.775845	6.49E-07
TMPRSS11B	Transmembrane Protease, Serine 11B	2	14	0.756323	2.04E-08
PIAS2	Protein Inhibitor Of Activated STAT 2	1	18	0.664443	1.22E-06
ARG1	Arginase 1	2	14	0.639505	2.50E-07
SRPK2	SRSF Protein Kinase 2	1	26	0.621438	1.30E-06
AADACL2	Arylacetamide Deacetylase Like 2	1	7	0.588406	7.38E-07
RGPD4	RANBP2-Like And GRIP Domain Containing 4	1	96	0.575957	4.71E-06
SPRR3	Small Proline Rich Protein 3	4	15	0.529901	2.44E-07
DEGS1	Delta 4-Desaturase, Sphingolipid 1	2	42	0.494578	2.04E-08
FBXO2	F-Box Protein 2	1	46	0.459531	0.000236

Table 11. Hypomethylated gene promoter regions in cancer tissues compared with normal tissues

Gene name	Description	No. of analyzed CpGs in WGM	No. of CpGs in promoter region	Δβ	Adjusted *P* fdr value
TRBC2	T-Cell Receptor Beta Constant 2	1	25	-0.79652	2.48E-07
DGAT2	Diacylglycerol O-Acyltransferase 2	1	14	-0.69618	2.48E-07
ALG1L	ALG1, Chitobiosyldiphosphodolichol Beta-Mannosyltransferase Like	1	37	-0.69448	3.29E-05
PDE4D	Phosphodiesterase 4D	1	10	-0.68484	1.06E-05
TRDC	T-Cell Receptor Delta Constant	1	17	-0.67205	1.46E-05
DNAJC6	DnaJ Heat Shock Protein Family (Hsp40) Member C6	1	9	-0.6704	1.63E-06
IGKV3-20	Immunoglobulin Kappa Variable 3-20	2	13	-0.66367	1.71E-05
TMEM150B	Transmembrane Protein 150B	1	54	-0.65928	8.13E-05
LAIR2	Leukocyte Associated Immunoglobulin Like Receptor 2	1	36	-0.64482	1.81E-05
UBQLN3	Ubiquilin 3	4	9	-0.64015	2.57E-06

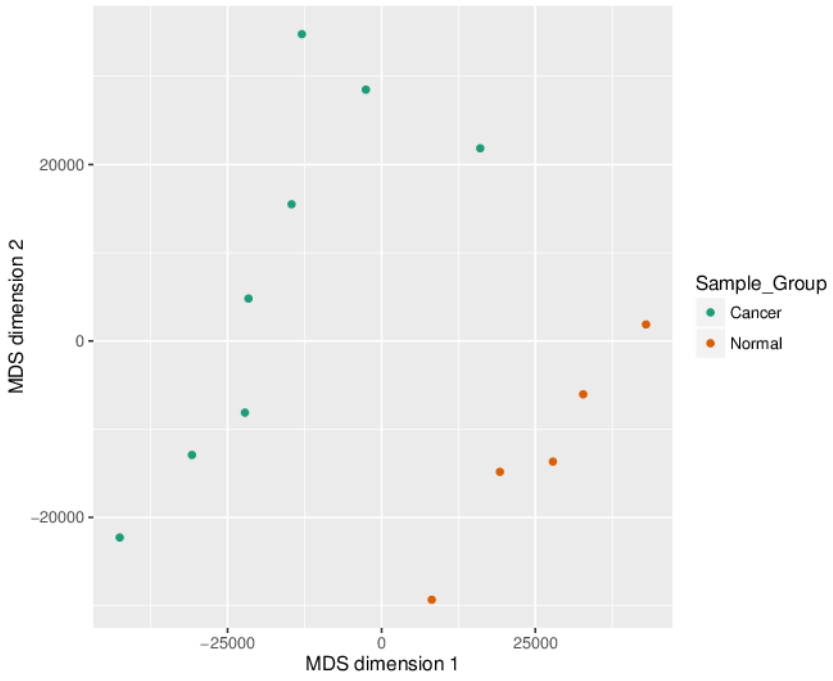

Figure 39. Scatter plot showing samples after performing Kruskal's non-metric multidimensional scaling

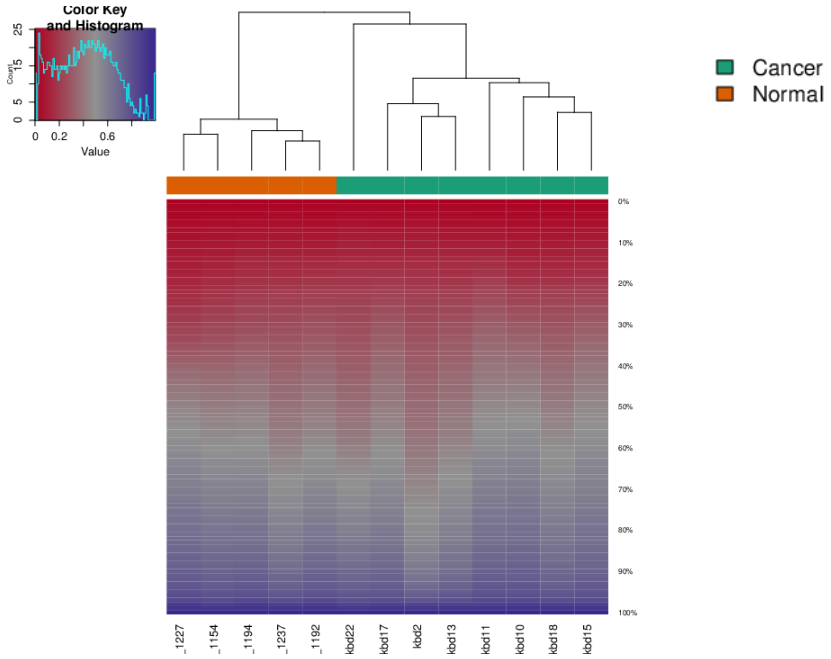

Figure 40. Hierarchical clustering of samples based on all methylation values. The heatmap displays methylation percentiles per sample.

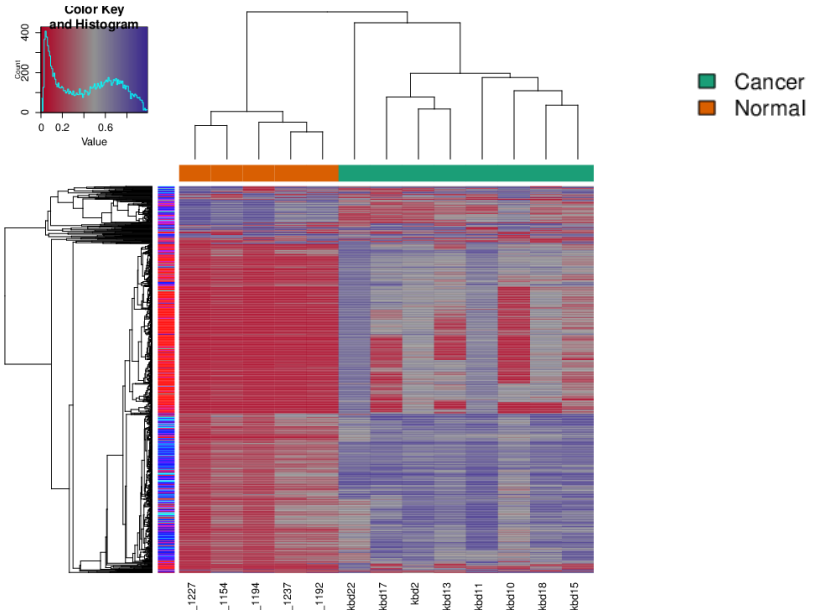

Figure 41. Hierarchical clustering of samples based on all methylation values. The heatmap displays only selected sites/regions with the highest variance across all samples.

4.4.2.2 Validation of potential DNA methylation biomarkers

Validation of whole genome methylation data was performed on the same set of samples (**Table 9**) that have been analyzed in Infinium MethylationEPIC BeadChip Array (Illumina, USA). Of 13 samples that passed the quality control filter of RnBeads analysis, 4 cancer samples and 1 normal control have been excluded due to insufficient amount of DNA necessary for the bisulfite conversion, leaving a total of 4 cancer samples and 4 normal controls for pyrosequencing analysis. All selected samples have been successfully bisulfite converted and prepared for pyrosequencing (**Figure 38B**)

A total of 10 top hypermethylated genes found in cancer, Small Proline Rich Protein 3 (SPRR3; $\Delta\beta$ = 0.529) and F-Box Protein 2 (FBXO2; $\Delta\beta$ = 0.459) have been chosen for pyrosequencing validation. In the same way, of total 10 top hypomethylated genes found in cancer, T-Cell Receptor Delta Constant (TRDC; $\Delta\beta$ = -0.672) and Leukocyte Associated Immunoglobulin Like Receptor 2 (LAIR2; $\Delta\beta$ = -0.644) have been chosen for pyrosequencing validation.

The selection criteria included the role of the genes in biological processes and the number of previous studies that investigated those genes. Specific primers used in pyrosequencing experiment, which have been designed in the PyroMark Assay Design software are presented in **Table 4**. Primers were chosen based on the inclusion and exclusion criteria that were

already mentioned in the Material and methods section. The quality score was above 70 for all primers, except for TRDC-2, where the score was 68. 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 gene *SPRR3*, 2 CpG for *TRDC*, 5 for *LAIR2* and 4 CpG for *FBXO2*. After confirming successful PCR amplification (**Figure 42**), PCR amplicons were ready for the pyrosequencing. Results are presented in **Figures 43 and 44**.

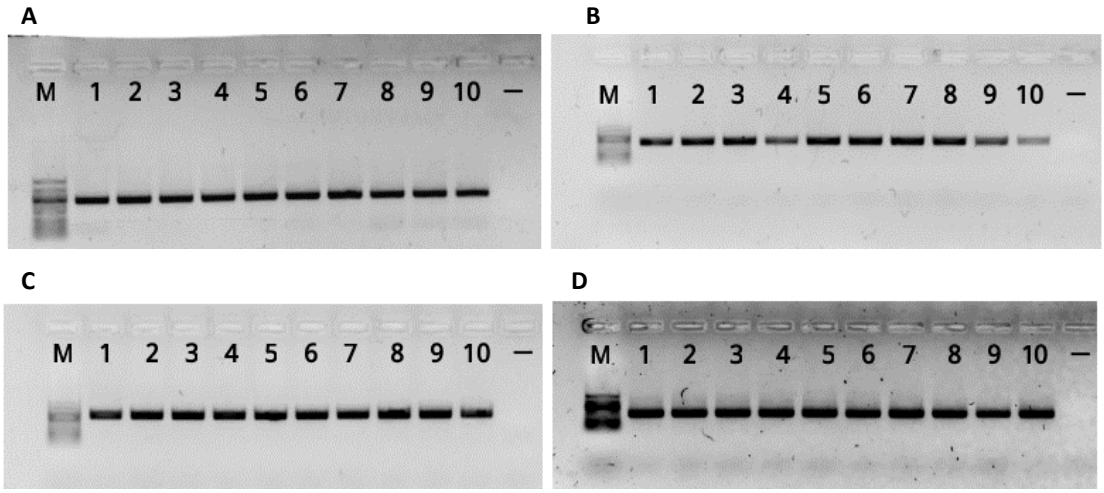

Figure 42. PyroMark **PCR amplicons of genes A) *SPRR3,* B) *FBXO2*, C) *TRDC* and D) *LAIR2* amplified for the analysis by pyrosequencing.** M = GelPilot 50 bp marker, lanes 1-4 = normal controls, lanes 5-8 = cancer samples, lane 9 = unmethylated control, lane 10 = methylated control, –= negative control

As presented in **Figures 43** and **44**, pyrosequencing results confirmed WGM analysis. Unfortunately, due to low number of samples included in the pyrosequencing validation, statistical significance was only reached between cancer and controls in 1st and 3rd CpG islands of the *SPRR3* gene (P = 0.01 in both cases) and 1st CpG of the *FBXO2* (P = 0.01) gene.

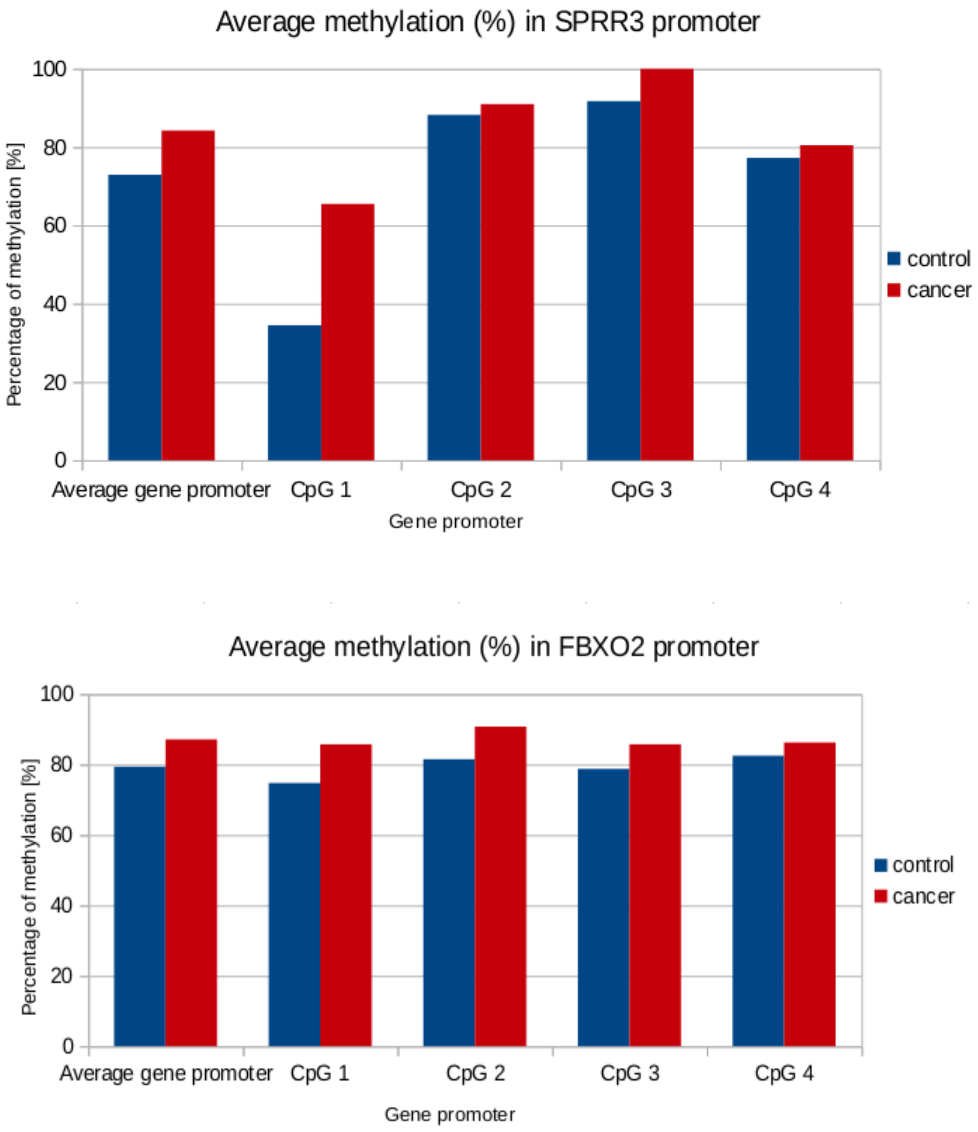

Figure 43. Methylation of average and specific CpG sites in the promoters of hypermethylated genes *SPRR3* and *FBXO2,* analyzed by pyrosequencing.

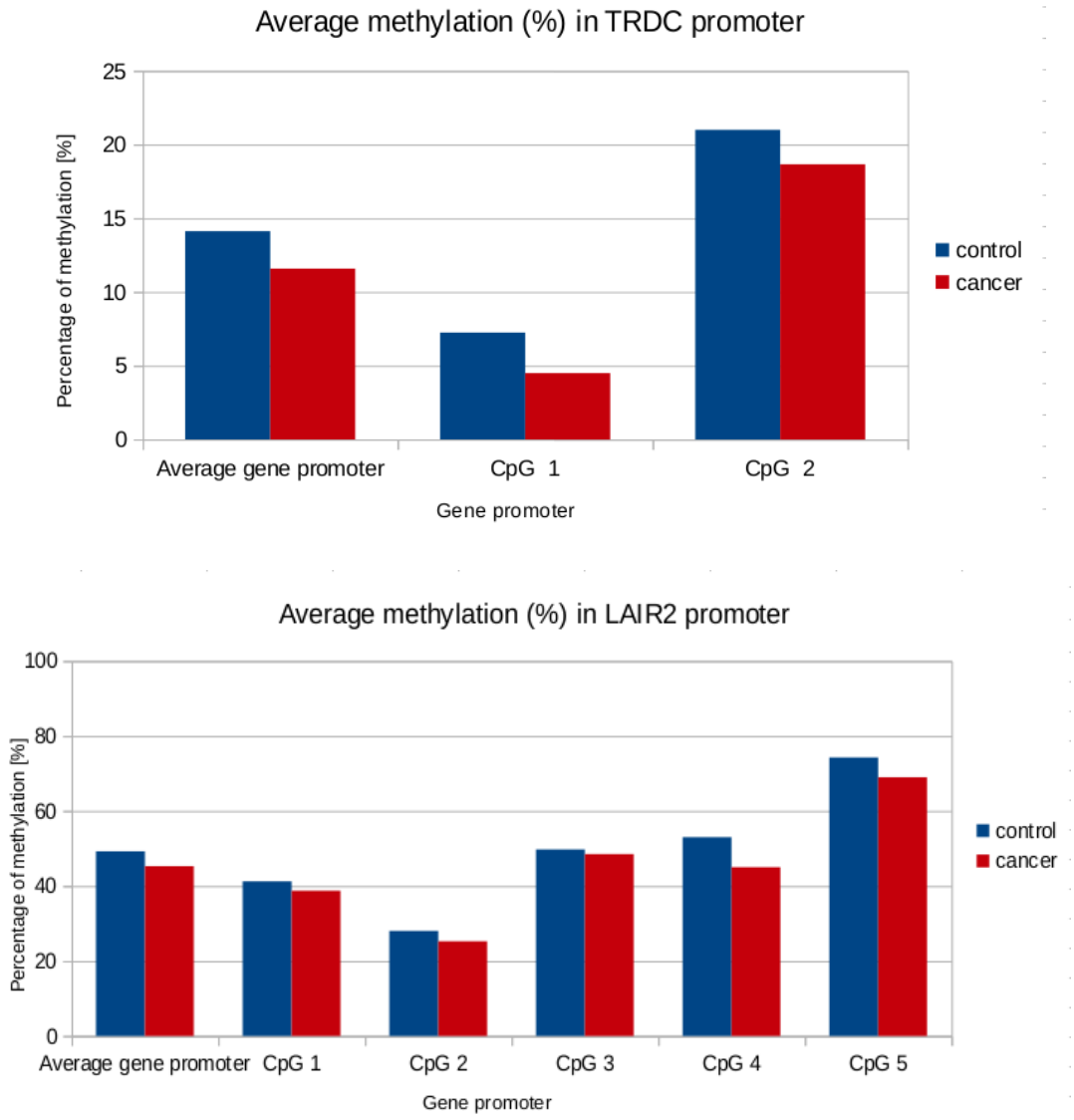

Figure 44. Methylation of average and specific CpG sites in the promoters of hypomethylated genes *TRDC* and *LAIR2,* analyzed by pyrosequencing.

4.4.3 Key epigenetic changes in HNSCC

4.4.3.1 Epigenetically affected genes in HNSCC

Firstly, miRs that were statistically deregulated in cancer and validated by qRT-PCR (hsa-miR-9, -21, -29a, -100, -106b, -143, -145, -199b) have been entered into publicly available database miRDB (www.mirdb.org) and all target genes of those miRs have been identified. Results are presented in **Table 12**. Database miRDB identified target genes of those miRs, and only the top 5 genes were presented. Genes that were found to be statistically differentially methylated in the WGM analysis are in bold.

Secondly, top 20 differentially methylated genes in cancer (top 10 hyper- and top 10 hypomethylated) identified in the WGM analysis were entered into the miRDB and miRTarBase databases (www.mirtarbase.mbc.nctu.edu.tw/php/index.php) and all miRs that are targeting those genes have been identified. All miRs found to target those genes were compared with the list of miRs found to be deregulated in our experiment. Genes with no miRs found to target those genes in either database were not entered in **Table 12.** In addition, only miRs that matched the list of sequenced miRNA in our study are presented in **Table 12**. Those miRs that were targeting at least 2 genes identified in our study and those miRs that were matching miRs sequenced in NGS experiment are in bold.

Table 12. Selection of significantly deregulated miRs and differentially methylated genes and their targets.

miRNA analysis			
miR	**No. target genes**	**Top 5 genes**	**DataBase**
miR-9	545	**SNX25**, **MDGA2, ENPEP, CTNNA1**, MESDC1	miRDB
miR-21	485	**ZNF326**, **MAP3K1**, FYTTD1, AKAP11, **STK38L**	miRDB
miR-29	632	BRWD3, **COL3A1**, **ERCC6**, **DGKH**, GATAD2B	miRDB
miR-100	27	KBTBD8, **HS3ST2**, ZZEF1, **MTOR**, **MBNL1**	miRDB
miR-106b	855	**PTPN4**, ARID4B, **EPHA4**, PKD2, **PDCD1LG2**	miRDB
miR-143	375	**GXYLT1**, VASH1, **ITM2B**, **ATP10A**, IGFBP5	miRDB
miR-145	495	ABCE1, MPZL2, DAB2, **KCNA4**, ABHD17C	miRDB
miR-199b	330	**ARHGAP12**, **HAPLN1**, SOS2, **NPAS2**, **ARHGAP21**	miRDB
Methylation analysis			
Gene (methylation status in WGM)	**No. miRs targeting**	**miRs found by NGS experiment (expression status in cancer)**	**DataBase**
SPRR3 (↑)	8	**miR-335** (↑)	miRDB, miRTarBase
FBXO2 (↑)	14	miR-3065 (↑)	miRDB, miRTarBase
TRDC (↓)	1	**miR-335** (↑)	miRTarBase
LAIR2 (↓)	4	**miR-335** (↑)	miRDB, miRTarBase
DGAT2 (↓)	22	miR-218 (↓), miR-3133 (↑), miR-2467 (↑), **miR-3664** (↓), miR-873 (↑)	miRTarBase
PDE4D (↓)	52	**miR-335** (↑)**, miR-7** (↑), miR-497 (↓), miR-424(↑), **miR-31** (↑)**,** miR-3120 (↑)	miRTarBase
DNAJC6 (↓)	12	**miR-26b** (↓), miR-486 (↓)	miRTarBase
UBQLN3 (↓)	1	**miR-26b** (↓)	miRTarBase
PIAS2 (↑)	49	**miR-3664** (↑), miR-3140 (↓), miR-2116 (↓)	miRTarBase
ARG1 (↑)	3	**miR-7** (↑)	miRTarBase
SRPK2 (↑)	12	**miR-21** (↑)	miRTarBase
RGPD4 (↑)	44	miR-548 (↓)	miRTarBase

4.4.3.2 Epigenetically affected signaling pathways in HNSCC

Finally, to evaluate the potential role of miRNAs that were significantly differentially expressed exclusively in HPV+ subsets, all such miRNAs with at least 100 normalized counts on average were entered into the Diana tools miRPath (v3.0) (www.diana.imis.athena-innovation.gr/ DianaTools/index.php?r=mirpath/index). The analysis indicated that at least 62 KEGG pathways appear to be significantly associated with genes those miRs interact with. The same approach was used for the analysis of cell signaling pathways in HPV- cancer subsets and it revealed 88 KEGG pathways. Only cancer-relevant KEGG pathways, along with their rank in HPV+ and HPV-negative cancer subsets have been presented in **Table 13.** Of particular interest was the pathway viral carcinogenesis (ranked second with *P* value 7.90E-08) in the HPV+ group, while in the HPV-negative group, it was ranked fifth (*P* = 2.88E-09). Pathway "Proteoglycans in cancer" seems to be highly involved in both HPV+ and HPV-, with first rank in both cases (P = 9.40E-12, P = 2.34E-16, respectively). Moreover, several pathways have shown large differences between HPV+ and HPV- HNSCC subsets, such as the "TGF-β" pathway which is at 48th place in HPV+ (*P* = 0.00659) and 12th place in HPV-negative subset, with large *P* value difference (*P* = 1.34E-06). In addition, it seems like pathways that are more affected in HPV-positive group were "Endocytosis", "Transcriptional misregulation in cancer", "Pancreatic cancer", "Adherens junction" and "N-Glycan biosynthesis" than in HPV-negative HNSCC (ranked 4^{th} *vs.* 14^{th} by *P* value, 5^{th} *vs.* 28^{th}, 6^{th} *vs.* 25^{th}, 9^{th} *vs.* 26^{th}, and 11^{th} *vs.* 55^{th}, respectively). In contrary, HPV-negative tumors seem to be more strongly associated with "Pathways in cancer", "HIF-1 signaling pathway", "Ubiquitin mediated proteolysis", "TGF-beta signaling pathway", and "Cell cycle pathway" (ranked 7^{th} *vs.* 14^{th} in HPV-positive group, 16^{th} *vs.* 61^{st}, 4^{th} *vs.* 40^{th}, 12^{th} *vs.* 48^{th}, respectively).

Table 13. Signaling pathway analysis of miRs deregulated in HNSCC.

KEGG pathway	HPV-positive		HPV-negative	
	Rank	*P* value	Rank	*P* value
Proteoglycans in cancer	1	9.40E-12	1	2.34E-16
Viral carcinogenesis	2	7.90E-08	5	2.88E-09
Hippo signaling pathway	3	1.10E-07	8	6.53E-07
Endocytosis	4	1.36E-07	14	2.07E-06
Transcriptional misregulation in cancer	5	1.36E-07	28	2.62E-05
Pancreatic cancer	6	1.36E-07	25	1.08E-05
Adherens junction	9	2.14E-07	26	1.66E-05
Colorectal cancer	10	2.79E-07	22	5.99E-06
N-Glycan biosynthesis	11	7.31E-07	55	0.003471
Prion diseases	13	8.17E-07	37	0.000605
Pathways in cancer	14	8.17E-07	7	9.2E-08
Glioma	15	1.51E-06	24	9.26E-06
Chronic myeloid leukemia	16	1.87E-06	23	5.99E-06
Cell cycle	18	2.59E-06	2	5.18E-11
ECM-receptor interaction	19	3.19E-06	13	1.71E-06
Renal cell carcinoma	20	8.47E-06	6	2.03E-08
Non-small cell lung cancer	21	1.01E-05	32	0.000182
p53 signaling pathway	22	1.43E-05	20	4.57E-06
Focal adhesion	23	1.63E-05	21	4.57E-06
Protein processing in endoplasmic reticulum	24	2.42E-05	3	9.79E-11
Central carbon metabolism in cancer	26	8.55E-05	42	0.001893
Prostate cancer	28	0.000109	10	1.11E-06
Other types of O-glycan biosynthesis	29	0.00015	49	0.002373
Regulation of actin cytoskeleton	31	0.000257	NA	NA
DNA replication	33	0.00037	43	0.001985
Melanoma	34	0.000472	59	0.006071
Bladder cancer	35	0.000472	57	0.004419
Estrogen signaling pathway	36	0.000491	60	0.006203
Endometrial cancer	37	0.000604	27	2.62E-05
Small cell lung cancer	38	0.000635	19	3.57E-06
Ubiquitin mediated proteolysis	40	0.001667	4	5.33E-10
Thyroid cancer	41	0.002487	45	0.001985
FoxO signaling pathway	42	0.002915	35	0.00043
MAPK signaling pathway	45	0.004836	82	0.040009
Thyroid hormone signaling pathway	47	0.00632	11	1.11E-06
TGF-beta signaling pathway	48	0.00659	12	1.34E-06
Acute myeloid leukemia	50	0.007006	15	3.09E-06
mTOR signaling pathway	52	0.009948	39	0.000605
PI3K-Akt signaling pathway	53	0.018919	71	0.017726
TNF signaling pathway	57	0.024399	51	0.002373
HIF-1 signaling pathway	61	0.031673	16	3.15E-06

Furthermore, similar analysis was done with the signaling pathways implicated with differentially methylated genes, but with a slightly different approach. For the analysis, a list of 10 hyper- and 10 hypomethylated genes was entered into the Reactome database (www.reactome.org/) and DAVID (www.david.ncifcrf.gov/) tool, and KEGG pathways have been analyzed. WGM analysis provided information on difference between cancer *vs.* normal controls, and HPV status was not analyzed as a separate factor, hence, **Table 14** presents signaling pathways implicated with genes whose methylation is deregulated in cancer, regardless the HPV positivity. Only statistically significant ($P < 0.05$) pathways have been presented. After entering hypermethylated genes, the Reactome database identified signaling pathways that are involved in sumoylation processes (6 pathways of 8), "Urea cycle" (0.0107) and "Sphingolipid de novo biosynthesis" ($P = 0.0466$). For hypomethylated genes, Reactome database identified signaling pathways involved in secretory mechanisms and endocytosis ("Lysosome Vesicle Biogenesis" and "Golgi Associated Vesicle Biogenesis"; $P = 0.0329$ and $P = 0.0494$, respectively), metabolism ("Acyl chain remodeling of DAG and TAG", "Triglyceride biosynthesis", "Triglyceride metabolism"; $P = 0.0071$, P = 0.0125, $P = 0.0337$, respectively) and in neurotransmitter regulation ("DARP-32", $P = 0.0214$).

The bioinformatic database DAVID revealed different signaling pathways to those of Reactome. After entering a list of hypermethylated genes, a list of KEGG pathways was identified. Some of signaling pathways have also been identified through miRNA profile of HNSCC: "ECM-receptor interaction", "Viral carcinogenesis", "Proteoglycans in cancer", "Adherens junction", "Focal adhesion", "Hippo signaling pathway", "p53 signaling pathway", "Pathways in cancer", "Cell cycle", "Regulation of actin cytoskeleton" and "Transcriptional misregulation in cancer", ranked by *P* value, starting from the most strongly associated signaling pathway. The signaling pathways that have not been found in the miRNA analysis are "Fatty acid biosynthesis", "Lysine degradation", "Hepatitis B" and "Neurotrophin signaling pathway", also ranked according to P value. Signaling pathways associated with hypomethylated list of genes revealed implications in immune system ("Complement activation, classical" and "Regulation of immune response", $P = 0.035$, $P = 0.05$, respectively) and endocytosis ("Receptor-mediated endocytosis", $P = 0.05$).

Table 14. Signaling pathway analysis of differentially methylated genes in HNSCC.

Signaling pathway	Rank	*P* value	Database (Category)
Hypermethylated genes			
SUMOylation of intracellular receptors	1	5.07E-04	**Reactome (Biologic Process)**
Urea cycle	2	0.0107	
SUMO E3 ligases SUMOylate target proteins	3	0.0143	
SUMOylation	4	0.0153	
SUMOylation of transcription factors	5	0.0214	
SUMOylation of ubiquitinylation proteins	6	0.0435	
Sphingolipid de novo biosynthesis	7	0.0466	
SUMOylation of transcription cofactors	8	0.0466	
ECM-receptor interaction*	1	1.06E-39	**DAVID (KEGG pathway)**
Fatty acid biosynthesis	2	8.8E-19	
Viral carcinogenesis*	3	1.7E-08	
Lysine degradation	4	1.1E-07	
Fatty acid metabolism	5	4.2E-06	
Proteoglycans in cancer*	6	7.2E-06	
Adherens junction*	7	2.5E-05	
Focal adhesion*	8	6.8E-05	
Fatty acid elongation	9	0.0002	
Hepatitis B	10	0.0004	
Hippo signaling pathway*	11	0.0020	
p53 signaling pathway*	12	0.0021	
Pathways in cancer*	13	0.0021	
Neurotrophin signaling pathway	14	0.0231	
HTLV-I infection	15	0.0395	
Cell cycle*	16	0.0452	
Regulation of actin cytoskeleton*	17	0.0452	
Transcriptional misregulation in cancer*	18	0.0452	
Hypomethylated genes			
Acyl chain remodeling of DAG and TAG	1	0.0071	**Reactome (Biologic Process)**
Triglyceride biosynthesis	2	0.0125	
DARPP-32 events	3	0.0214	
Immunoregulatory interactions between a Lymphoid and a non-Lymphoid cell	4	0.0280	
Lysosome Vesicle Biogenesis	5	0.0329	
Triglyceride metabolism	6	0.0337	
Golgi Associated Vesicle Biogenesis	7	0.0494	
Complement activation, classical	1	0.035	**David (KEGG pathway)**
Regulation of immune response	2	0.05	
Receptor-mediated endocytosis	3	0.05	

*signaling pathways associated also with miRNA deregulated in our study

5 DISCUSSION

5.1 HPV analysis in FFPE HNSCC samples

Detection of HPV genotypes in cervical cancer and precursor lesions strongly demonstrated the causal relationship of particular HPV genotypes to specific cervical changes (de Sanjose et al. 2010; Sabol et al. 2017). Furthermore, the use of FFPE samples creates the possibility to include a high number of samples during a longer time period, making these type of samples very valuable and informative for many different large-scale studies. Thus, we conducted a retrospective study, where we have assessed the presence of HPV (positivity of HPV DNA) and its transcriptional activity (positivity of HPV-16 E6*I spliced form) in HNSCC samples collected within a 13-year period. The identification of HPV in HNSCC FFPE samples is of high importance, due to different etiology of the disease, and there is a strong indication that the treatment could be tailored for a specific group of patients, HPV-positive and HPV-negative (Van Doorslaer, Chen, and McBride 2016; Dok and Nuyts 2016).

We successfully extracted and analyzed HPV DNA from 115 HNSCC samples, mostly of oropharyngeal localization, collected at the Clinical hospital center Zagreb, which is the largest hospital in Croatia where patients from the whole country are referred for treatment. Following HPV DNA detection, 39% samples tested positive, which is in line with the previous reports on similar cancer samples (Boscolo-Rizzo, Pawlita, and Holzinger 2016; Mehanna et al. 2013). Boscolo-Rizzo *et al.* highlighted that the presence of HPV DNA in tissue biopsies is not always sufficient to attribute cancer to HPV, and suggests performing a combination of the two different tests, i.e. HPV DNA testing and testing for the presence of the E6 mRNA (Boscolo-Rizzo et al. 2013). Moreover, other studies also suggested (Robinson et al. 2012; Seiwert 2013) that the presence of both HPV DNA and RNA in tissue biopsies should be assessed, since inactive HPV could be just a passenger, and not the cancer driver. Therefore, they suggest the usage of mRNA assays to be more reliable in resolving the HPV involvement in cancer development and suggest direct viral involvement in carcinogenesis (Butel 2000; Viarisio, Gissmann, and Tommasino 2017).

Investigating full length E6 mRNA and its splicing forms could be used to indicate HPV activity. There are 4 alternative spliced forms of the E6 mRNA (McFarlane et al. 2015), with the most abundant form in advanced cancer being E6*I form (Graham and Faizo 2017). Therefore, in our study we have examined the presence of the E6*I mRNA in HPV-positive samples and bypassed possible misclassifications of the tumors, as proposed by Boscolo-Rizoo *et al.* (2013). We have successfully extracted and analyzed HPV RNA from 43 HPV DNA positive samples. The HPV-16 E6*I mRNA analysis revealed that 73.3% of HPV DNA positive samples had active E6 oncogene transcription.

We categorized the patients into three groups combining data on HPV DNA and RNA analysis:

HPV (true) positive (DNA and E6*I positive), HPV inactive (DNA and full length 16 E6 positive) and HPV (DNA) negative group. However, our data have not shown significant differences between HPV DNA and HPV RNA positive groups regarding age, gender, tumor stage, tumor grade, life style habits and overall survival *(P* = 0.353, *P* = 0.331, *P* = 0.195, *P* = 0.151, *P* = 0.303, *P* = 0.523, respectively). This could be due to a limited number (n = 45) of analyzed FFPE samples.

The p16 marker is considered as a good surrogate marker for the HPV infection (Lesnikova et al. 2009; El-Naggar and Westra 2012; Larsen et al. 2014); however, the association of hr-HPV with specific cancer sites plays a crucial role. There are studies that already proved high concordance of p16 and HPV status in cervical epithelial neoplasia and cervical cancer (Wang et al. 2004; L. Hu et al. 2005; Zhang et al. 2007), where HPV is associated with 99% of the cases. In HNC, the association of HPV infection with cancer is lower (40-50%), and the correlation of HPV and p16 varies. However, the p16 is a good surrogate marker for only oropharyngeal HNSCC according to several studies (Lewis et al. 2010; Prigge et al. 2016; Albers et al. 2017). Anyhow, this certainly suggests that p16 is not a suitable replacement for HPV testing in the current setting, which is in line with other studies (Albers et al. 2017; Lechner et al. 2018), in which they stated that the p16 is a relatively accurate marker for HPV within the oropharynx, while it can be unsuitable for use in other HNSCC subsites, where a smaller proportion of tumors are HPV-driven. Moreover, there are also studies where the presence of p16 has been detected in HPV-negative oropharyngeal cancer (Stephen et al. 2013). The literature also suggests that p16 positive, but HPV-negative HNSCC share some common characteristics, with HPV-positive HNSCC, such favorable prognosis (Stephen et al. 2013; Albers et al. 2017; Taberna et al. 2017).

In our retrospective study, the correlation between p16 immunostaining and HPV data was very low with only 15.4% of active HPV cases being positive for p16. The correlation coefficients were r = -0.09 for the comparison with HPV DNA and r = 0.18 for the comparison with HPV RNA. However, there is statistically significant association of p16 with older age (*P* = 0.0385). Even though p16 positive and younger patients are usually associated with good prognosis (Albers et al., 2017), in this case, the p16 association with age showed the opposite. When oral samples are excluded from the statistical analysis, and because p16 is considered unreliable for this subgroup of cancer, the significance of this association is lost (*P* = 0.052). The presence of p16 was therefore considered not a reliable surrogate for active HPV involvement in HNSCC development in this study.

Comparison of HPV RNA positive, HPV DNA positive and HPV-negative HNSCC based on gender shows an obvious predominance of males regardless the HPV positivity. Furthermore, there was no significant difference between the age groups (*P* = 0.338), which is in contrast with other studies (Taberna et al. 2017; Boscolo-Rizzo et al. 2017). Even though the incidence of HNSCC was considered mostly associated with men in this study, currently, there are indications that the incidence of HNSCC in males is declining, while it is increasing in females

(Boscolo-Rizzo et al. 2018).

Considering the tumor location, we found more cases of HPV-positive tumors in the oropharynx than in the oral cavity, which was expected, as most of the cancer (>70%) arises from the oropharyngeal origin (Dok and Nuyts 2016; Taberna et al. 2017). Presented in **Table 4**, the proportion of HPV-positive, HPV inactive and HPV-negative oral tumors was low (7.7%, 9.3% and 14.2%, respectively). Furthermore, when we compared solely specific tumor sites, a very similar pattern in all three groups are again seen, which was contrary to other studies (Taberna et al. 2017). The authors reported a non preferable site of origin in HPV-negative tumors, while confirming the predilection of HPV to the oropharyngeal region (Taberna et al. 2017; Osazuwa-Peters et al. 2017; Hussein et al. 2017); in these studies, the most HPV-positive samples were found in the base of tongue and tonsils. We have confirmed those findings in our study, with the most common sites of tumor location being the base of tongue and tonsils in all three groups, HPV RNA-positive, HPV DNA-positive and HPV-negative (30.8%, 34.4% and 28.6% for base of tongue, respectively, and 53.8, 50% and 42.9% for tonsils respectively). For these HPV-positive cancers this is already expected, and it has already been reported in different studies (Vigneswaran and Williams 2014; Gillison et al. 2015; Ramqvist et al. 2015; Taberna et al. 2017). Moreover, cancers at those particular regions, tonsils and base of tongue are the most responsible for the rising incidence of HNSCC in many Western countries.

Dahlstrom et al. (2015) pointed out that HPV-positive HNSCC primarily affects younger males and females with higher socioeconomical status and higher oral sex practice (Dahlstrom et al. 2015). Therefore, sexual behavior as an independent risk factor for HNSCC development can be potentially equally responsible for HPV infection in both genders (Boscolo-Rizzo et al. 2017). Furthermore, it could be assumed from our study population, that the life style (smoking habits and alcohol consumption) of the individuals could affect the male/female ratio, since smoking and alcohol consumption in Croatia have been apparently more associated with males (Goel and Budak 2007). This could possibly explain the observed higher differences in gender distribution (in favor of men) in HPV-negative, than in HPV-positive cancers (69.2% *vs.* 30.8% in HPV-positive, 81.3% *vs.* 18.7% in HPV inactive and 87.1% *vs.* 12.9% in HPV-negative group, for males *vs.* females, respectively). Moreover, in the study of Anantharaman *et al.* (2016) the authors investigated a combined effect of smoking and HPV-16 in oropharyngeal cancer, and they reported that the prevalence of oropharyngeal cancer increases with smoking for both HPV-positive and HPV-negative group (Anantharaman et al. 2016). In addition, not just smoking poses the problem, but also alcohol intake, since it is also known to increase cancer risk (Marron et al., 2012; Huang et al, 2017).

Smoking and drinking habits of our study population (**Table 4**) could potentially explain why there are similar patterns in oropharyngeal cancer, regardless of HPV-positivity. In our study, of the patients with available data, only 13% never smoked and 14% never consumed alcohol

with the majority still being active smokers and at least moderate drinkers. This observation is in line with the study of Descamps *et al.* (2016) in which they showed almost no effect of HPV on patients' outcome in a population of heavy tobacco and alcohol consumers (Descamps et al. 2016). They emphasized that the life style habits, like smoking and drinking has the greatest impact on survival. In addition, Hafkamp *et al*. (2008) reported that HPV-positive tumors show a favorable prognosis as compared to those with HPV-negative tumors, even though tobacco use was still the strongest prognostic indicator (Hafkamp et al. 2008).

When we investigated differences in the three HPV-defined groups in terms of the age of diagnoses, surprisingly, the Chi-square test showed no significant differences (mean1 = 58y; mean2 = 59y; mean3 = 59y; tStat = -0.76; $P = 0.224$; $\alpha = 0.05$). This suggest that our study does not show a typical higher incidence of HPV-positive HNSCC in younger population, like it was shown in the study of Hussein *et al.* (2017) where they reported an increased incidence of patients worldwide with HPV-associated HNSCC at younger age (defined as <45 years old) (Hussein et al. 2017). Moreover, Wagner *et al.* (2017) reported a similar median age in HPV-positive and HPV-negative groups, which is in line with our study, but in their study HPV was the most important predictor of survival, which is not the case in our study (Wagner et al. 2017). All those data lead us to presume that the effect of HPV presence in the Croatian population might be overshadowed by other risk factors, such as the prevailing tobacco and alcohol intake. To evaluate this hypothesis, we assessed all available relevant patient data such as age of diagnosis, life style habits, TNM status, tumor grade, HPV presence, p16 positivity and the type of therapy to determine survival of the patients. Overall, survival analysis of the HPV DNA positive (HPV-positive and HPV-inactive) and HPV-negative (**Figure 22**) showed a small, but no statistical difference in survival rate between the two groups, which was unexpected. The overall survival was only 32%, however, it is worth to mention that this could be influenced by the lack of disease-specific mortality data. The use of all-cause mortality might also be masking the effects of other variables, which failed to stratify patient risk of death. The only variable that showed a strong significant influence on survival was the N stage ($P = 0.05$; **Figure 25**). However, the combined risk stratification as proposed by Ang *et al.* (2010), originally including HPV status, smoking habits (which was excluded in our statistical analysis due to partial documentation) and TNM stage successfully classified patients in low-, intermediate- and high-risk of death groups within the first 5-years ($P = 0.0079$; **Figure 26**). Even though patients in high-risk group have the worst survival, separation was not as good, since patients in the intermediate-risk group had better survival, than those in the low-risk group. Moreover, due to missing data on T and N stage, not all patients could be classified according to the Ang *et al.* model, which could also affect survival.

Interestingly, the study by Nygård *et al.* (2012) reported a period in Norway (1981-1995) in which the survival of HPV-positive HNSCC was worse than the survival of HPV-negative

HNSCC (Nygård et al. 2012). However, in more recent years (1996-2007) there was a dramatic shift for the survival of HPV-positive patients. Furthermore, in their population, the median age of diagnosis decreased from 63.2 to 59.8 years for HPV-positive but remained unchanged at 66 years for HPV-negative patients. In this study population, decreasing of median age was not observable, since in the time period from 2002-2015 the median age of HPV-positive patients was 58 years, while in the period from 2015-2018 (prospectively collected samples; **Table 5**) the median age was even higher (62.5-years); in HPV-negative patient the median age also increased from 59 to 62 years.

As already mentioned, the lack of the positive impact of HPV on overall survival could be attributed, not only to life style habits, but possibly to the treatment nature for the HPV-positive patients as well (**Table 4**). Currently, many studies have been made on the treatment de-escalation for HPV-driven oropharyngeal cancer (Mirghani and Blanchard 2018). Moreover, it was previously suggested that HPV might affect chemo/radio sensitivity (Lassen et al. 2014; Marcu 2016; Dok and Nuyts 2016). However, in our study group, only 9 patients have a favorable therapy, survival data, and HPV DNA and RNA positivity, which might be insufficient to significantly shift the survival curves. Moreover, our study population was from the time period when modalities of the therapies based on HPV positivity have not yet been implemented in the clinics. Today the situation is different, and the studies are investigating the best option for the HPV-positive patients. In our study population, surgery followed by radiation was the most common type of therapy in all groups, while the second most used choice of therapy was surgery alone. Surgery together with chemo- and radiotherapy was the third most common type of therapy. In our study, there were no difference in choice of the therapy between HPV-positive and HPV-negative patients (P = 0.0685). In contrast, the study of Broglie *et al.* reported better prognosis for non-smokers HPV-positive oropharyngeal tumors when treated with surgery alone (Broglie et al. 2017), while Hong *et al.* and Lassen *et al.* found increased sensitivity to radiotherapy in HPV-positive group (Hong et al. 2010; Lassen et al. 2014), which could possibly mean that radiotherapy is a favorable type of treatment for the HPV-associated cancer. Furthermore, tumors at the lower stage of development are usually treated with surgery alone, while advanced tumors are treated with non-surgical or in combination with non-surgical methods, such as radio- and/or chemotherapy according to the American Cancer Society (ACS 2018). Since we found that tumor stage significantly influences survival (P = 0.0107), it is expected that therapies associated with a particular stage could also have different outcomes.

Regarding the tumor grade, we did not find significant differences between groups (P = 0.152); however, there is indication that the tumors of higher grade might be more associated with HPV-positive tumors. For example, the study of Vokes *et al.* (2015) reported that HPV-positive cancers are prone to be of higher grade, poorly differentiated, hence more sensitive to radiotherapy, unlike HPV-negative cancers with higher differentiation and lower grade (Vokes et al. 2015). In this study, there is a trend in HNSCC grade difference based on

the HPV activity (**Figure 22**). The highest percentage of HPV active (E6*I positive) samples were found to be in grade 3 cancer. This is also in line with the study of Olmedo-Nieva *et al.* (2018) performed on the cervical cancer model (Olmedo-Nieva et al. 2018). Moreover, it was already shown that the alternative splicing depends on the molecular signaling (Shin and Manley 2004). Rosenberger *et al.* (2010) showed that the splice variant is regulated via EGF signaling (Rosenberger et al. 2010). It is certainly of interest to investigate whether there is a correlation between the disease grade and the presence of the E6*I mRNA form in HNC, since this type of study has not yet been investigated. Our statistical analysis showed again no significant correlation between the grade and the presence of splicing form (P = 0.151); however, there is a possible trend and potential correlation (**Figure 22**), which might be revealed on a larger number of samples.

Missing medical records was a strong limitation to our study. This is mainly due to the retrospective nature of the study, where adequate detailed medical documentation was not always available. Sometimes detailed follow-up was impossible because the patients attend other hospitals after the initial treatment, which is why the medical records were supplemented by the data from the Croatian Cancer Registry (CCR). The CCR collects some of the relevant information on cancer patients irrespective of where the patient was referred and also contains survival data on a national level. Unfortunately, there were other limitations when using FFPE, such as DNA degradation, due to the tumor fixation process in paraffin, as was stressed out by Dietrich *et al.* (Dietrich et al. 2013). We, however, bypassed this particular limitation by using small specifically designed primers for the detection of HPV (Kleter et al. 1998) in highly degraded FFPE samples as it was emphasized in the study of de Sanjose *et al.* (de Sanjose et al. 2010). Therefore, we have maximally reduced the possibility of getting false negative results. Furthermore, low tumor material in some FFPE samples presents potential limitation, as it can contribute to poor RNA concentration and consequently limits the RT reaction. Thus, one sample was excluded from the study since it was not possible to repeat the nucleic acid extraction. Such problems were also discussed in the study by Roberts *et al.* (2009), where they emphasized the importance of the RNA quality that was still critical for reliable analysis of FFPE samples; hence, having low tumor material for adequate RNA isolation was a crucial problem in RNA analysis (Roberts et al. 2009).

Overall, this retrospective study provides the baseline relevant data on HNSCC patients in Croatia, since this population shows differences in the nature of the disease but show no distinctive features of HPV-positive patients, as seen in other similar studies mentioned previously. In other words, this current data indicates that in our population, HPV should not yet be considered as a critical favorable prognostic biomarker, as observed in Western populations. Moreover, better designed studies with higher number of patients with detailed follow-ups are needed for clearer clarification of disease differences in order to implement appropriate treatment.

5.2 Analysis of miRNA profiling in fresh HNSCC samples

Identifying the key miRs that would preferably distinguish HPV-negative from HPV-positive HNSCC has posed a big challenge to many researchers including us. Even though miRNA profiling could be performed using larger scale retrospective FFPE samples, analysis of miRnome in this present study was performed using only fresh cancer samples. Vojtechova *et al.* investigated systematic comparison of the miRNA expression profiles between paired fresh and archival FFPE tumors from tonsil, and found that only 27-38% of the differentially deregulated miRNAs overlapped between the two source systems (Vojtechova et al. 2017). They emphasized that for an accurate comparison of the miRNA expression profiles from published studies, it is important to use the same type of clinical material and to test and select the best-performing normalization method for data analysis. This indicates that there might be a problem of choosing the best type of samples, as well as choosing the best method for the analysis. Each sample sources have advantages and disadvantages. While fresh samples contribute to better RNA quality, they are relatively rarely available, especially in countries such as Croatia, where incidence of HNSCC is very low. On the other hand, FFPE offer higher pool of relatively easy available samples, but with obviously degraded RNA. An important fact that needs to be considered when discussing miRNA analysis studies, is that there is a high discrepancy regarding different studies, and there is no clear consensus on the significance of the individual miRNAs, which could serve as potential biomarkers.

The study of Lubov *et al.* highlights miRNAs profiling results is often problematic due to the small sample size, as well as to biological variations among tumors and non-standardized assays for miR detection (Lubov et al. 2017). This might also explain the high discrepancy in the literature. Hence, to add to the current knowledge, we assessed the miRNA profiling in Croatian population in order to identify plausible miRs that might play a role in developing and/or maintaining HNSCCs in the Croatian population and those that resemble Croatian life style habits and genetic structure. The whole genome miRNA analysis was performed on fresh cancer samples collected at the KBD Dubrava, Zagreb; we have successfully collected 65 oral and oropharyngeal tumors over the 2015-2018 period, of which only 61 qualified for the study on miRNA profiling, and only 53 cancer samples qualified for the survival study because only patients with primary cancer were included. As with the archival samples, patients were categorized in three main groups regarding the HPV DNA and RNA status (**Table 5**). The majority of cancers were of oral origin (60%, tongue, floor of mouth, buccal mucosa, gingiva and retromolar region), with males being predominant in all three groups. The majority of participants declared as being smokers (77%). The overall HPV prevalence was 23%, which is lower compared to other Western countries where the HPV presence is shown in 50% of cases (Mehanna et al. 2013). As expected, most of the HPV DNA positive samples arose from HPV-16 (12/14; 85.7%), with 2 samples being HPV-18 positive, which is in line with expected predominance of >80% of HPV-16 in HNSCC (Gillison et al. 2008).

Within the HPV RNA positive (HPV active) group of patients, the highest incidence of cancer was in tonsils (50%), which was also expected (Taberna et al. 2017), while in HPV DNA positive (HPV inactive group) was found in tongue cancer (50%). HPV-negative cancer most often occurred on gingiva (25%), followed by floor of mouth (23%). Altogether, approximately a quarter of cancers were of oropharyngeal origin (23%; base of tongue, tonsil and posterior pharyngeal wall). Furthermore, a high concordance of HPV E6*I mRNA presence and HPV-16 DNA was found in only 2 cases (16.7%) where HPV was not active, while usually only 40% are shown to be active (Jung et al. 2010).

There is almost an identical median age in HPV-positive, HPV-inactive and HPV-negative patients (62.5, 59, 62, respectively), which is surprising, considering numerous studies reporting significantly younger population among HPV-positive HNSCC (Syrjänen 2010; Keck et al. 2015; Taberna et al. 2017). Furthermore, besides the age, it seems that HPV RNA-positive, HPV DNA-positive and HPV-negative patients do not differ in smoking and drinking habits (66%, 50% and 57%, respectively). Moreover, the survival analysis (**Figure 27**) confirmed that there was no significant difference between HPV-positive (HPV DNA and HPV RNA group) and HPV-negative group in terms of gender, age group and smoking and drinking habits (P = 0.4563, P = 0.1393, P = 0.2582, and P = 0.4443, respectively). Nevertheless, it should be taken into consideration that survival analysis could be under the influence of low sample number involved in the study (53, of which only 5 patients died) and short follow-up time (less than 5 years).

Some differences could be seen regarding the tumor stage, as 100% of HPV RNA-positive, 75% of HPV DNA-positive and 47% of HPV-negative patients were diagnosed in late stage. These parameters appear to differ from other studies, since HPV group is considered to be of younger population, with better survival and no/minimum drinking or smoking intake (Gillison et al. 2008; O'Rorke et al. 2012; Boscolo-Rizzo, Pawlita, and Holzinger 2016; Taberna et al. 2017); however, this is probably due to particularities of the Croatian population already shown in the previous retrospective study. So, smoking in Croatia still poses a serious problem (Padjen et al. 2012), which together with alcohol consumption can significantly increase cancer risk. Hence, Croatian and similar populations are at a great risk for HNSCC and HPV appears to play a smaller role. This could potentially explain similarities of tumor samples regarding age, TNM stage, and could also explain the high prevalence of oral tumors in our study, since smoking and drinking are most prevalent risk factors for oral cancer development (Taberna et al. 2017). Survival analysis (**Figure 28**) again confirmed no significant differences between the three groups based on HPV status, tumor grade and tumor stage (P = 0.2942, P = 0.5764, P = 0.4741, respectively). The only significant variable that influenced survival was found to be the N status (P = 0.0057). Finally, three cancer groups have been compared based on tumor invasion and combined risk of death (**Figure 29**); even though survival curves indicate some influence on survival, the statistical significance has not been reached (P = 0.1689 and P = 0.1167, respectively). This was not expected, since

the effect of perineural invasion on overall survival in HNSCC has already been demonstrated (Huyett et al. 2017). Moreover, like with tumor invasion, combined risk of death showed no statistical significance, but survival curves indicate that there might be a trend. As stated earlier, it is possible that survival is strongly affected by the patient short follow-up time and unequal distribution between alive and deceased patients (48 *vs.* 5, respectively). This should be taken into considerations since it might influence survival curve slopes and lead to false conclusions. Moreover, since it was not possible to calculate 5-years survival, this survival analysis at this point is only preliminary.

After examining HPV-positive and HPV-negative cancer samples based on patients' biological, clinical and histopathological characteristics, and after assessing the overall survival based on specific factors, we were interested to see if we could find any significant differences between the two groups based on miRNA profiling. During the course of the study, we have thoroughly reviewed current literature and focused on all miRNAs found in each reviewed study, instead of focusing only on those selected by the authors. Unfortunately, this approach has also failed to increase the overlap of literature data. Thus, after thorough literature reviewing, we chose a subset of samples for the high-throughput miRNA analysis, performed by NGS sequencing. A total of 22 samples passed the criteria for library generation (**Table 6**). Selection of samples was according to the tumor site and the HPV status. HPV RNA inactive samples were considered as HPV-negative in this case. Therefore, we included 4 different sample groups, HPV-positive oral squamous cell carcinoma (HPV+OSCC; O+), HPV-negative oral squamous cell carcinoma (HPV-OSCC; O−), HPV-positive oropharyngeal squamous cell carcinoma (HPV+OPSCC; OP+) and HPV-negative oropharyngeal squamous cell carcinoma (HPV-OPSCC; OP−). This classification could certainly get better understanding of miR deregulation in head and neck cancer. Numerous studies by now have already performed miRNA profiling, but the information on specific cancer sites as a distinctive factor, or the HPV status, was not always taken into consideration (Avissar et al. 2009; Lajer et al. 2011, 2012; G. Gao et al. 2013; Wan et al. 2017).

The main NGS sequencing experiment detected significant differential regulation of 1,172 unique miRNA sequences across comparisons (some have been presented in T**able 7**). However, the majority were in fact isomiR sequences (**Figure 33**). To which extent do these sometimes very abundant forms affect results of other studies where they cannot be distinguished from proper targets, is yet unknown. Certainly, there are studies that point out the importance of isomiRs in gene regulation (Cloonan et al. 2011; Telonis et al. 2017; Guo et al. 2016). Moreover, Guo et al. emphasized that some isomiRs have been proven as functional small RNAs by associating with target mRNAs and influencing miR stability or effectiveness (Guo et al. 2016). Up to date, there is only one study using hybridization techniques that reports the isomiRs influence in HNC (Saito et al. 2013), where only miR/isomiR 196 was sequenced, but not the high-throughput whole-genome sequencing. In our study, we distinguished miRs from isomiRs, but certainly, more thorough study on

isomiRs effect in HNSCC is necessary.

Another outcome of the miRNA NGS profiling was the apparent inability of this method to clearly differentiate 4 specific subgroups of samples (**Figure 34**). Only normal samples could clearly be distinguished according to the respective microRNA profile. This indicates potential lack of homogeneity of subgroup profiles despite careful selection of available samples. Furthermore, the samples clustering strongly depends on a tumor site with clustering of oropharyngeal subset being better than oral subset (**Figure 35**). This is in line with other studies (Miller et al. 2015; Vojtechova et al. 2016) indicating that miRNA profiling can distinguish more precisely specific tumor groups in oropharyngeal cancer than in other tumor sites, such as oral cancer. This also supports the fact that HPV is more associated with oropharyngeal cancers, rather than with tumors of other origin. Deeper analysis also revealed that HPV inactive samples cluster similarly to HPV-negative samples, which could indicate that active HPV form, and not just the presence of HPV DNA, is highly necessary for driving the cancer in a particular direction.

The NGS data have been validated with RT-qPCR on a total of 61 tumor samples. The fold changes from all miRs that have been validated and compared with fold changes from the NGS experiment showed no significant difference (P = 0.142). Hence, our findings are in concordance and reliable, showing specific miRNA profile based on the tumor site and the overall HPV status (**Figure 34**). This is in line with the findings of other studies (Q. Huang et al. 2002; Chung et al. 2004) where the authors report how it is crucial to specify the tumor site, since HNC is considered to be of an extremely heterogeneous nature. The expression analysis using RT-qPCR data were also in line with the majority of similar studies (Vojtechova et al. 2016), where miR-21 was upregulated in all 4 groups, regardless of the tumor site and the HPV status, while miR-9 was found to be significantly upregulated in OP+ (P = 0.0015), which is in line with the studies of Sethi *et al.* (Sethi et al. 2014). The other miR found to be potentially associated with OP+ is miR-143, which was downregulated in cancers, but with slight predominance in this group. Nevertheless, the statistical significance was not reached (P = 0.068). Moreover, among eight miR-s chosen for clinical testing, only miR-9 and miR-21 showed significant upregulation compared to controls, while for upregulated miR-106b, the difference in the FC is smaller between cancer and the control. Of eight validated miRs, miR-29, -100, -143, -145 and -199 all have been downregulated in all 4 groups.

The mir-21 was intensively investigated, with reports on miR-21 as highly expressed in the larynx, pharynx and tonsils when compared to healthy tissue as well as in blood from OSCC patients when compared to normal healthy blood controls (Chang et al. 2008; Cao et al. 2013; Zhang et al. 2016). Furthermore, mRi-21 is proposed to serve as a potential biomarker for the oral tongue carcinoma, especially (Zujian Chen et al. 2017), while in the study of Arantes *et al.* the overexpression of miR-21 has been associated with patients that consume high levels of alcohol (Arantes et al. 2016). Extensive review of the literature revealed that miR-21 appeared to be highly involved in the immune system, which could potentially explain the

high prevalence of deregulated miR-21 in various cancers. It is known that miR-21 plays a crucial role in the switch from pro-inflammatory into anti-inflammatory response, allowing cancer cells to more successfully evade immune response (Sheedy 2015). Since HNSCC is an immunosuppressive disease (Ferris, 2015), it is possible that miR-21 plays a crucial role in HNSCC. The question remains if this is enough for being identified as a potential biomarker in the disease development. For this study, we have bypassed miRs that were not cancer-specific, hence miR-21 definitely did not meet the criteria for the identification of potential biomarkers, even though it might play a crucial role in HNSCC development. Moreover, the question arises if upregulation of miR-21 is the cause or the consequence of the cancer cell signaling.

As stated, most evidently all other miRs, except of miR-9 showed no specificity to any group. One of the most prominent miRs that have been significantly associated with HPV-positive OPSCC is certainly miR-9. Moreover, deregulation of miR-9 has been proven in many types of cancers, and recently, high association of miR-9 upregulation with HPV-positive HNSCC has been confirmed several studies (Sass et al. 2015; Spence et al. 2016; Tuna and Amos 2016; Husain and Neyaz 2017). Song *et al.* reported that miR-9 promotes tumor metastasis by repressing E-cadherin in esophageal SCC (Song et al. 2014). Given that HPV inhibits cell differentiation, it is possible to explain the association of miR-9 and HPV-positive HNSCC. Wang *et al.* also confirmed the potential correlation of miR-9 and E-cadherin *via* TGF-β1, but in lung cancer, which could mean that miR-9 also plays a role in the immune system (Wang et al. 2017). They reported that TGF-β1 plays an important role in the epithelial–mesenchymal transition (EMT) of epithelial cancers, including non-small cell lung cancer. In our study, TGF-β1 is significantly associated with the HPV-negative group (**Table 13**), meaning that miR-9 might affect other signaling pathways.

The analysis of deregulated miRNA in HPV-positive and HPV-negative HNSCC using KEGG pathway database revealed the main signaling pathways involved in the disease development, which could narrow the choice of the most reliable miRs as potential biomarkers. As expected, viral carcinogenesis is highly associated with HPV-positive tumors (*P* value ranked 2; while for HPV-negative, *P* value ranked 5). The TGF-β pathway was shown to be more associated with HPV-negative cancers. TGF-β is a cytokine that is important for maintaining Tregs, which are necessary for immunosuppression (Ansa-Addo et al. 2017). This could possibly mean, that one of the mechanisms for evading immune response from the host in HPV-negative cancers is *via* TGF-beta signaling pathways, while HPV-positive cancers use other mechanisms. Ferris suggested that HPV-positive HNSCC, evades immune response by gaining T-cell tolerance to the persistent HPV infection, and production of low genome copy numbers in the basal layer of epithelium (Ferris 2015). Another immunity-involved signaling pathway in our analysis is the mTOR pathway, ranking 39th (*P* = 0.000605) in the HPV-negative group, and 52nd (*P* = 0.009948) in the HPV-positive group, which is in line with the study of Marques *et al.* Where the mTOR pathway proteins are presented as highly

associated with poor overall survival (Marques et al. 2016). This could suggest that the mTOR pathway plays a role in maintaining cancer in the HPV-negative group, as this group has worse survival (Taberna et al. 2017). In addition, miRNAs are involved in several signaling pathways that play crucial roles in immune response (**Table 13**). It is known that HNC is considered as an immunosupressive disease, and Tregs have been highly associated with HNSCC (Ferris 2015). This supports the idea to investigate the direct role of each specific miRNA that we validated in qRT-PCR, and see if miRs deregulated in our study have, among others, possible role in immune escape, and possibly responsible for maintaining the cancer with huge impact on clinical outcome. The potential role of miR-100 in immune response was described in the study of Negi *et al.*; they suggest that altered expression and editing of miR-100 regulates immunosupressive Treg differentiation (Negi et al. 2015). All those data support the hypothesis of Masuda *et al.* and Ferris *et al.* (Ferris 2015; Masuda, Wakasaki, and Toh 2016) that HNSCC might be an epigenetic and an imunosupresive disease.

The mir-29 was found to be highly associated with epigenetic factors in cancer development (Sethi et al. 2014). For instance, miR-29 was shown to be overexpressed in B-cell lymphoma, due to reduced histone deacetylation, while downregulated in lung cancer, by hypermethylation of the promoter, disabling the polymerase to bind to the DNA. This suggests a potential interplay of the methylation and miRNA roles in tumor development, and that it should not be always considered as two separate mechanisms. The study of Steiner *et al.* suggests that miR-29a regulates helper T-cell differentiation by repressing multiple target genes (Steiner et al. 2011). Some studies confirmed the tumor suppressive role of miR-29a in HNSCC (Fukumoto et al. 2016), and Wang *et al.* (Wang et al. 2018) even proposed miR-29a as a potential therapeutic target.

For miR-106b, it is known that its deregulation i associated with many cancer types, including HNC (Y. Gao et al. 2016). Even though it was not significantly associated with HPV-positive cancer in our study, many studies proved that upregulation of miR-106b in laryngeal carcinoma (Cai, Wang, and Bao 2011; Cheng et al. 2017). Regarding function, studies showed that miR-106b plays a role in controlling cell cycle (Ivanovska et al. 2008), interplay with signaling pathways, which control proliferation and induction of stem cell-like phenotype in cancers (Lu et al. 2017), while some studies also showed an implication of miR-106b in the immune response (Xiao and Rajewsky 2009). The study of Lu *et al.* suggest that miR-106b mediates the constitutive activation of Wnt/β-catenin signaling in renal cancer (Lu et al. 2017), while Cioffi *et al.* reports that clustering of miR-25-93-106b regulates tumor metastasis and immune evasion via modulation of CXCL12 and PD-L1 (programmed death-1) (Cioffi et al. 2017). Moreover, the PD-1 receptor and its ligands PD-L1 and PD-L2 are known to be significantly involved in T-cell regulation (Müller et al. 2017), hence PD-L1 expressing cancer cells have the ability to efficiently evade the host immune system. Interestingly, Ferris reported an increased PD-L1 expression in HPV-positive tumors and increased PD-1 expression in cytotoxic T-lymphocytes, and suggests this being the mechanism by which

HNSCC cells, especially in HPV-positive patients evade immune response (Ferris 2015). Moreover, it has been stated the clinical efficacy of US Food and Drug Administration approved monoclonal antibodies targeting PD-1 (Ferris 2015). Potentially, miR-106b could serve as biomarker, but first it should be evaluated on a bigger pool of samples, which will enable more thorough investigation of its potential association with HNSCC.

Regarding miR-143 and miR-145, it has been proven that they are among the best examples of tumor suppressor miRNAs (Raisch 2013; Darfeuille-Michaud, and Nguyen 2013), and many studies reported a correlation of miR-143 and miR-145 downregulation with poor prognosis (Slaby et al. 2007; Schepeler et al. 2008). In particular, miR-143 is involved in the inhibition of oncogene KRAS expression (Chen et al. 2009), while the miR-145 inhibits tumor growth and angiogenesis by directly targeting kinases that are activated by mTOR in colorectal cancer (Xu et al. 2012). Studies of those miRs are not so extensive in the head and neck area, but there are reports that investigated functions of those miRs particularly in HNC (Sethi et al. 2014; Bufalino et al. 2015; Sun and Zhang 2017).

In respect of miR-199b, it has been stated that the lymph node metastasis (N stage) or perineural invasion is associated with low miR-199b levels in HNSCC (Sousa et al. 2016). Sousa *et al.* also suggested miR-199b, among others, as a potential prognosis marker and therapeutic target (Sousa et al. 2016). Interestingly, miR-199b was found to be involved in inducing autophagic death of endometrial carcinoma cells by targeting the mTOR pathway (Cai et al. 2017), which also propose that this miR, as other miRs plays an important role in the immune system. Nevertheless, more studies are necessary in order to get a better understanding of the mechanisms and the importance of specific miRs included in HNSCC development. What we assume, is that some of those miRs certainly qualify for potential biomarkers, which could help enlighten mechanisms of cancer development, maintenance and progression. Our study provides important results which could contribute to the overall miRNA studies in HNSCC.

5.3 Methylation profiling in fresh HNSCC samples

The whole genome methylation study was performed on 16 HNSCC samples, of which 13 showed high quality and performance on the Infinium MethylationEPIC BeadChip kit, so their methylome has been further analyzed. Tumor and control cells have expectedly formed distinguished clusters based on their methylation profile (**Figure 39**). Furthermore, the fine clustering could not be seen clearly when the heatmap displayed methylation percentiles per sample (**Figure 40**), but when the heatmap displayed only selected sites and regions with the highest variance across all samples (**Figure 41**), clear clustering was evident. Before the selection of control samples, it was important to get a population as similar as possible to the cancer group, in order to get the most reliable results. Hence, the study group consisted of adult males (males were predominant; 69%) and females with a median age of 56.5 years, with the older adults being predominant (11 samples; 69%). Moreover, it is well known that age affects the methylation pattern in the human genome (Reynolds et al. 2014; Dongen et al. 2016), due to the fact that methylome is subjected to both genetic and environmental effects. Therefore, by matching the age of cancer patients with normal controls, we minimized potential differences which are not cancer related.

The most significantly hypermethylated gene promoters in HNSCC in comparison to normal buccal samples are mostly receptors and genes included in relevant cells functions (**Table 9**). The other significant group of gene promoters affected by methylation belongs mainly to the genes included in the immune response and they were found to be hypomethylated in cancer samples in comparison to normal samples (**Table 10**). Furthermore, most of the hypomethylated genes were genes that play a crucial roles in the immune system (**Table 10**), which is in line with an earlier study by Milutin Gašperov et al, in which the authors elaborated the possible importance of the timing of activation of the immune system by demethylation of specific genes in cervical cancer (Milutin Gašperov et al. 2014). Therein, all significantly hypomethylated genes in cervical cancer tissue *vs.* normal tissue were identified by the Illumina Infinium HumanMethylation450K BeadChip method, an earlier version of the assay used by the current study. Moreover, the authors reported that the most strongly correlated genes based on their function were immune effectors' process (*AIM2, BST2, BTN3A3, and IL12RB1*) and response to virus related genes (*AIM2, BST2, and IL12RB1*). Thus, they hypothesized that the activation of these genes through demethylation is probably triggered by HPV oncogenes (Milutin Gašperov et al. 2014).

In the current study, we performed a whole genome analysis of the methylation profile in HPV-positive and HPV-negative HNSCC and we found the top 10 hypomethylated genes to be *TRBC2, DGAT2, ALG1L, PDE4D, TRDC, DNAJC6, IGKV3-20, TMEM150B, LAIR-2* and *UBQLN3*, which were thus activated in cancer. Among that group of genes two were chosen for validation by pyrosequencing: *TRDC* and *LAIR2*. The T-Cell Receptor Delta Constant (TRDC)

has not yet been reported as hallmark of solid cancer, nor has it been investigated in epigenetic deregulation of HNSCC; this is the first investigation in such manner. Leukocyte immunoglobulin-like receptors (LILR) belong to the family of receptors possessing extracellular immunoglobulin domains and within this family, the most prominent are leukocyte-associated immunoglobulin-like receptor 1 and 2 (LAIR-1 and LAIR-2). *LAIR-1* is broadly expressed on the majority of immune cells, however, the biological role of LAIR in solid tumors has yet to be elucidated, unfortunately, the number of studies of both, LAIR-1 and LAIR-2 is insufficient. However, there is a study by Wang *et al.* (2016), in which they used immunohistochemical staining analysis, in order to determine the expression of LAIR-1 in human cervical cancer cells and in normal-adjacent tissue (Wang et al. 2016). Their results indicated that the expression of *LAIR-1* in cervical tissue was higher compared with that in noncancerous tissue. This could potentially mean that HPV activity might influence this gene. Unfortunately, no such study has been performed for *LAIR-2* in the same setting, nor in HNSCC.

Genes that have been hypermethylated in our study, Small Proline Rich Protein 3 (*SPRR3*) and F-Box Protein 2 (*FBXO2*) have been largely investigated. SPRR3 is a protein involved in cornification, epidermis development, keratinocyte differentiation and peptide cross linking, while FBXO2 is involved in negative regulation of cell proliferation, in cellular protein modification, protein ubiquitination etc. (GCS 2018). Moreover, SPRR3 is a member of the SPRR family of cornified envelope precursor proteins, and it is a marker for terminal squamous cell differentiation. Since it is known that the HPV-positive HNSCC tend to be of high-grade, and less differentiated than HPV-negative cancers, it is possible that there is an interplay between HPV and *SPRR3* in those tumors. In addition, the study of Lehr *et al.* (2004) proved that infection with HPV in genital area alters the expression of the small proline rich proteins 2 and 3 (Lehr et al. 2004). Moreover, the study of Jeon *et al.* (2004) performed global gene expression profiles of HPV-positive HNSCC cell lines and showed underexpression of *SPRR3* (Jeon et al. 2004). Furthermore, it was also proven that the F-box proteins play an important role in the epigenetic regulation of cancer, mediated through ubiquitination-dependent and -independent manner (Shen and Spruck 2017).

Genes that have been activated in cancer *via* hypomethylation are genes mostly involved in immune response, with TRDC 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, hence it is crucial in effective immune responses. Moreover, among the LAIR-2 protein related pathways are the innate immune system, and class I MHC mediated antigen processing and presentation. Since those four genes encode proteins involved in crucial biological roles and were not thoroughly studied in methylation analysis, we chose to investigate the exact methylation pattern in several CpG sites in their promoters. Unfortunately, our pyrosequencing experiments had some limitations. Briefly, only 8 samples of a total 16 samples that have been used in the WGM

studies were available for additional pyrosequencing experiments, therefore, statistical analysis was limited. Further, pyrosequencing is a method for detecting many CpG islands in close proximity, hence it was not possible to include CpG sites that were located at greater distances or in particular genomic settings due to primer design limitations. In particular, some areas consisted of extremely densely located CpG sites, which was another limited factor in the study since it was not preferred to have more than one variable region at the primer annealing area. Nevertheless, in the comparative studies, pyrosequencing has been shown to be among the most accurate and reproducible technologies for locus-specific DNA methylation analysis and has become a widely used tool for the validation of DNA methylation changes identified in genome-wide studies (Busato et al. 2018).

Pyrosequencing results revealed similar methylation status of those 4 gene promoters*; i.e. TRDC* and *LAIR2* hypomethylated and *SPRR3* and *FBXO2* hypermethylated (**Figure 43** and **Figure 44**), but due to low sample numbers, statistical significance has been reached only in *SPRR3* and *FBXO2* genes (P = 0.01 in both genes). In addition, since probes used in WGM analysis do not cover all CpG sites in the gene promoter region (**Tables 9** and **10**), and due to the limitation factors in pyrosequencing primer design, it was not possible to analyse the exact CpG sites analyzed in the WGM experiment. This might be another reason why it was not possible to reach statistical significance after pyrosequencing.

Nevertheless, data presented on graphs represents the expected trend, with *TRDC* and *LAIR2* being hypomethylated in cancer, while *SPRR3* and *FBXO2* have been hypermethylated in cancer. Moreover, after analyzing the average methylation in these 4 validated genes regarding HPV positivity, results were as follows: 85.3% in HPV-positive cancer *vs*. 83.1% in HPV-negative for *SPRR3* gene, 87% *vs.* 85.4% for *FBXO2*, 43.3% *vs.* 47.3% for *LAIR2* and for *TRDC* gene, the average methylation was 7.2% in HPV-positive and 13.8% in HPV-negative samples. These results indicate that there might be correlation between the methylation status and HPV positivity. Certainly, for that analysis, these tests should be performed on a bigger pool of samples.

5.4 Integration of miRnome and methylome data

We integrated all obtained data on epigenetic modifications in HNSCC and identified genes and signaling pathways implicated in cancer development. This analysis could point out the potential epigenetic biomarkers of the disease, especially in HPV-positive cancers. Integration of the miRnome and methylome was performed using publicly available databases: DAVID (www.david.ncifcrf.gov), miRPath (www.mpd.bioinf.uni-sb.de), miRDB (www.mirdb.org), miRTarBase (www.bio.tools/mirtarbase), and Reactome (www.reactome.org). The aim was to identify the interplay, if possible, between miRNA deregulation and differently methylated genes found by the miRNA NGS and the WGM experiments.

Validated miRs (hsa-miR-9-5p, -21-3p, -29a-3p, -100-5p, -106b-5p, -143-3p, -145-5p, -199b-5p) were analyzed in miRDB and their target genes were identified (**Table 12**). Even though there was no overlap with the top 20 deregulated genes identified in the methylation studies (**Tables 9** and **10)**, some of miR's target genes have been also statistically differentially methylated in the WGM study (highlighted in **Table 12**). Moreover, there was a possible correlation between the signaling pathways found in the miRNA profiling study and genes found in the methylation study. One of the top 5 target genes of miR-9, which was found to be significantly associated with HPV-positive tumors, is the Sorting Nexin 25 (*SNX25*) gene, which was also significantly differentially methylated in the WGM studies (P = 0.034). This gene may be involved in several stages of intracellular trafficking (Gene cards, Human Gene Database; www.genecards.org). This is of interest, since signaling pathways involved in endocytosis and protein processing in the endoplasmic reticulum (ER) were found in both cancer groups (**Table 13**), but endocytosis being more associated with HPV-positive cancer (ranked 4^{th} *vs.* 14^{th}) and protein processing in the ER to HPV-negative cancer (ranked 3^{rd} *vs.* 24^{th}). In addition, after analyzing the signaling pathways of genes that were differentially methylated in cancers (**Table 14**), it is evident that hypomethylated genes are associated with endocytosis (lysosome vesicle biogenesis, Golgi associated vesicle biogenesis, and receptor-mediated endocytosis). Moreover, a study by Huang and Chen reports SNX25 to enhance TGF-β receptor degradation in lysosomes independent of ubiquitination (Huang and Chen 2012), meaning that this gene is also implicated in the TGF-β signaling pathway, which was found to be more associated with HPV-negative cancer (**Table 13**).

Another target gene of miR-9 that was identified is the MAM Domain Containing Glycosylphosphatidylinositol Anchor 2 (*MDGA2*) gene, which is involved in signaling pathways like metabolism of proteins, post-translational modifications, and cell-cell interactions (www.genecards.org). Moreover, the lysine degradation signaling pathway, and the sumoylation (post-translational modification) were significantly associated with

hypermethylated genes (**Table 14**), which implies the importance of these signaling pathways in cancer development and maintenance.

Another miR of interest, miR-29a targets the Bromodomain And WD Repeat Domain Containing 3 (*BRWD3*) gene, which is thought to have a chromatin-modifying function, and may play a role in transcription (www.genecards.org). The study of Li and Grandis also reports *BRWD3* as commonly mutated in HPV-negative cancers (Li and Grandis 2015).

The target gene of miR-29 is the diacylglycerol kinase eta (*DGKH*) gene, which is involved in regulating intracellular concentrations of diacylglycerol (DAG) (National Center for Biotechnology Information – NCBI; www.ncbi.nlm.nih.gov). In the present study, Acyl chain remodeling of DAG and TAG is significantly (P = 0.0071) associated with hypomethylated genes (**Table 14**). Interestingly, fatty acid biosynthesis and metabolism are also associated with hypermethylated genes, indicating the importance of this signaling pathways involved in HNSCC. Luo *et al.* (2017) emphasized the importance of lipid metabolism in cancer progression, since it is known that cancer cells frequently display fundamentally altered cellular metabolism (Luo et al. 2017).

Target genes of other important miRNAs such as miR-106b are implicated in signaling pathways that might play a crucial role in HNSCC. As stated before, the PD-1 receptor and its ligands, PD-L1 and PD-L2 are known to be significantly involved in T-cell regulation (Müller et al. 2017), and cancer cells expressing PD-L1 have the ability to efficiently evade the host immune system. The Programmed Cell Death 1 Ligand 2 (*PDCD1LG2*) gene is a target gene of miR-106b (**Table 12**), and it was also statistically differentially methylated in cancer *vs.* controls in the WGM analysis (P = 0.0039). Moreover, as already stated, increased *PD-L1* expression was associated with HPV-positive cancer (Ferris 2015) as well as increased *PD-1* expression in cytotoxic T-lymphocytes, suggesting that this might be the mechanism through which HPV-positive HNSCC patients evade immune response. In addition, the signaling pathways implicated in the immune system (**Table 14**) are significantly associated with hypomethylated genes such as the complement activation, a classical pathway gene (P = 0.035), and regulation of immune response genes (P = 0.05).

The signaling pathways of miR-199b targeting genes also correlate with our methylation study. For instance, one of its top 5 target genes, Hyaluronan And Proteoglycan Link Protein 1 (*HAPLN1*) gene stabilizes the aggregates of proteoglycan monomers with hyaluronic acid in the extracellular cartilage matrix. As seen from the signaling pathway analysis, extracellular matrix (ECM)-receptor interaction is strongly associated with hypermethylated genes (P = 1.06E-39), as well as the proteoglycans in cancer signaling pathway (P = 7.2E-06). Moreover, this pathway was ranked as 1st in both HPV-positive and HPV-negative cancer (**Table 13**).

After we analyzed miRs targeting the most differentially methylated genes in cancer *vs.* normal controls, the analysis revealed several miRs that were targeting these genes. We were interested if some of those miRs were found in the NGS analysis (**Table 12**). In addition,

there were several miRs that were targeting more than one gene such as hsa-miR-7-5p, hsa-miR-21-5p, hsa-miR-26b-5p, hsa-miR-31-3p, hsa-miR-335-5p, and hsa-miR-3664-3p; only miR-26b-5p and miR-3664 have been downregulated, while others were upregulated.

Finally, after analyzing all the data and investigating any potential overlap between methylome and miRnome of our HNSCC patients, we were able to propose several key genes and miRNA ($P > 0.05$) and identify the signaling pathways that are found to be of high importance in HNSCC occurrence. Therefore, as NGS findings indicate, miR-9 was found to be highly associated with HPV-positive patients ($P < 0.001$, what was also confirmed by qRT-PCR ($P = 0.0015$) (**Figure 37**). In addition, miR-9 was found to be associated with cell differentiation, cellular trafficking and endocytosis, which was also found to be more associated with HPV-positive tumors by the integration analysis (**Table 13**). Another miR often linked to HPV-positive patients is miR-106b. Even though this miR has not been confirmed by qRT-PCR ($P = 0.902$), there is still a trend of higher expression in HPV-positive OPSCC and HPV-positive OSCC. Another fact that supports this statement is that miR-106b target the *PDCD1LG2* gene, which has also been associated with HPV-positive HNSCC in other studies (Lajer et al. 2012; Sethi et al. 2014; Miller et al. 2015).

Another miRNA potentially involved in HPV-positive HNSCC, according to the NGS findings and integration data analysis is certainly miR-335. Even though this miRNA has not been selected for validation by qRT-PCR, the NGS results indicate upregulation of miR-335 in HPV-positive tonsil cancer (FC = 4.21), which is also in line with the study of Vojtechova et al. (Vojtechova et al. 2016). Moreover, miR-335 targets the *PDE4D* gene, which was hypomethylated in the WGM study. This could also indicate that hypomethylation of *PDE4D* is associated with HPV-positive cancer. Another target of miR-335 is the *SPRR3* gene, that was hypermethylated in our study. This gene presents a potential biomarker for HPV-positive cancer since it is a marker for terminal squamous cell differentiation (GCS 2018). Since HPV prevents differentiation of infected cells, it is possible that these genes are hypermethylated in HPV-positive cancers. In addition, the *LAIR2* gene might be indicative in HPV-positive cancers, due to studies of activation (hypomethylation) of *LAIR1* in HPV-positive cervical cancers (Yue Wang et al. 2016). Since, we confirmed the activation (hypomethylation) of the *LAIR2* gene in HNSCC (**Figure 44**), and since this gene is also a target of miR-335, it is highly possible that *LAIR2* is associated with HPV-positive tumors. Unfortunately, there are no other studies confirming this hypothesis, so we intend to investigate further this association on a larger number of HNSCC.

Regarding HPV-negative HNSCC, data suggest that miR-29a might be a suitable biomarker. Again, qRT-PCR did not confirm any significance in this subset, but it seems more associated with HPV-negative tumors based on the NGS findings (**Figure 37**), and it is also in line with the study of Vojtechova *et al.* (Vojtechova et al. 2016). Moreover, miR-29a targets the *BRWD3* gene, which is already reported as commonly mutated in HPV-negative cancers (Li and Grandis 2015). In respect to miR-199b, it is reported that the N stage and perineural

invasion were shown to be associated with low miR-199b levels in HNSCC (Sousa et al. 2016). Furthermore, miR-199b was found to be involved in inducing autophagic death through targeting the mTOR pathway (Cai et al. 2017), which could indicate involvement of miR-199b in HPV-negative tumors as this miRNA was shown to be deregulated therein.

Overall, this comprehensive study provides the analysis on HNSCC in the Croatian population on the HPV status in archival and fresh cancer samples, and epigenetic changes including miRNA and DNA methylation profiling. Our findings showed potential epigenetic biomarkers that could be used in diagnostic, prognostic and therapeutic approaches not only on Croatian population, but also on other populations with similar characteristics.

6 CONCLUSIONS

This study on head and neck cancer could be summarized into several key points that are related to the HPV status and the epigenetic changes:

1. HPV was present in 39% (45/115) of archival FFPE tumors.

2. HPV was present in 23% (14/61) of fresh tumor samples, with HPV-16 being the most prevalent type found in 12 HPV samples (86%), of which one was with both HPV-16 and HPV-18 (7%). Besides HPV-16, there were two HPV-18 positive samples (14%).

3. From 12 HPV-16 positive fresh tumor samples 6 (50%) were E6 transcriptionally active, while 13 of 45 (28%) HPV-positive archival tumor samples were E6 transcriptionally active.

4. The analysis of miRnome and methylome of cancer samples compared to normal controls, revealed different patterns and specific sample clustering.

5. Statistical analysis identified differences between cancer and normal samples regarding miRnome profiles; miR profiling showed 552 different unique miRNA sequences ($P < 0.05$), of which, miR-9, -21, -29a, -100, -106b, -143, -145, and -199b have been validated in qRT-PCR. Only miR-9 was significantly associated with HPV positivity by both methods, NGS and qRT-PCR ($P = 0.0004$, $P = 0.00156$, respectively). The miR-335 determined by NGS but not validated by qRT-PCR was significantly ($P = 0,0006$) associated with HPV positivity.

Whole genome DNA methylation profiling revealed 120,901 differentially methylated sites in cancer *vs.* normal ($P < 0.05$). Statistical analysis determined differences between cancer and normal samples regarding the methylome in the top 20 genes (10 hypermethylated and 10 hypomethylated in cancer) of which four genes (*FBXO2, LAIR2, SPRR3, TRDC*) have been validated by pyrosequencing. The average methylation of the hypermethylated *SPRR3* gene between HPV-positive and HPV-negative samples was 85.3% and 83.1%, respectively, while average methylation for the hypomethylated gene *LAIR2* was 43.3% in HPV-positive, while 47.3% in HPV-negative samples, indicating a possible association with HPV positivity.

6. MiRnome and methylome data revealed key implicated cell signaling pathways, that have been statistically associated with HPV-positive and HPV-negative HNSCC. Most epigenetically deregulated cell pathways significantly associated with both HPV-positive and HPV-negative group were signaling pathways involved in endocytosis, differentiation, extracellular matrix receptor interaction, epithelial cell-cell interaction, lipid metabolism, and immune response. The HPV-positive group was mostly affected by endocytosis ($P = 1.36E-07$), adherens junctions ($P = 2.14E-0$), and N-Glycan biosynthesis ($P = 7.31E-07$), while for HPV-negative group, cell cycle pathway, ubiquitin-mediated proteolysis, TGF-β and mTOR signaling seem to be more implicated in cancer ($P = 5.18E-1$, $P = 5.33E-10$, $P = 1.34E-06$, $P = 0.000605$,

respectively). Genes that have been significantly epigenetically changed in HNSCC are hypomethylated *TRBC2*, *DGAT2*, *ALG1L*, *PDE4D*, *TRDC*, *DNAJC6*, *IGKV3-20*, *TMEM150B*, *LAIR2*, and *UBQLN3*, while among the hypermethylated genes, genome-wide methylation analysis revealed *GPRC5D*, *TMPRSS11B*, *PIAS2*, *ARG1*, *SRPK2*, *AADACL2*, *RGPD4*, *DEGS1*, *SPRR3*, *FBXO2* ($P < 0.05$ for all 20 genes).

7. The integration analysis of miRnome and methlyome data revealed several overlapping target genes deregulated in HNSCC, that could contribute to the overall analysis: *SNX25*, *MDGA2*, *BRWD3*, *PDCD1LG2*, *HAPLN1*, *PDE4D*, *SPRR3* and *LAIR-2*.

8. After complete integration analysis of the genes deregulated in head and neck cancer (miRnome and methlyome) and signaling pathways that are implicated in this type of cancer, we propose several potential epigenetic biomarkers in HPV-positive HNSCC to be evaluated by more focused and extensive studies on a larger cohort of samples: miR-9, miR-335, *PDE4D*, *SPRR3* and *LAIR2.*

7 REFERENCES

ACS. 2018. “Treatment Options for Oral Cavity and Oropharyngeal Cancer by Stage. American Cancer Society.” 2018. https://www.cancer.org/cancer/oral-cavity-and-oropharyngeal-cancer/treating/by-stage.html.

AJCC. 2018. “AJCC - Cancer Staging Manual. Https://Cancerstaging.Org/References-Tools/Deskreferences/Pages/Default.Aspx.” 2018. https://cancerstaging.org/references-tools/deskreferences/Pages/default.aspx.

Albano, Pia Marie, Dana Holzinger, Christianne Salvador, Jose Orosa, Sheryl Racelis, Modesty Leaño, Danilo Sanchez, et al. 2017. “Low Prevalence of Human Papillomavirus in Head and Neck Squamous Cell Carcinoma in the Northwest Region of the Philippines.” *PLoS ONE* 12 (2). https://doi.org/10.1371/journal.pone.0172240.

Albers, Andreas E., Xu Qian, Andreas M. Kaufmann, and Annekatrin Coordes. 2017. “Meta Analysis: HPV and P16 Pattern Determines Survival in Patients with HNSCC and Identifies Potential New Biologic Subtype.” *Scientific Reports* 7 (1): 16715. https://doi.org/10.1038/s41598-017-16918-w.

Anantharaman, Devasena, David C Muller, Pagona Lagiou, Wolfgang Ahrens, Ivana Holcátová, Franco Merletti, Kristina Kjærheim, et al. 2016. “Combined Effects of Smoking and HPV16 in Oropharyngeal Cancer.” *International Journal of Epidemiology* 45 (3): 752–61. https://doi.org/10.1093/ije/dyw069.

Ang, K. Kian, Jonathan Harris, Richard Wheeler, Randal Weber, David I. Rosenthal, Phuc Felix Nguyen-Tân, William H. Westra, et al. 2010. “Human Papillomavirus and Survival of Patients with Oropharyngeal Cancer.” *New England Journal of Medicine* 363 (1): 24–35. https://doi.org/10.1056/NEJMoa0912217.

Annunziato, Anthony T. 2008. “DNA Packaging: Nucleosomes and Chromatin | Learn Science at Scitable.” 2008. https://www.nature.com/scitable/topicpage/dna-packaging-nucleosomes-and-chromatin-310.

Ansa-Addo, Ephraim A., Yongliang Zhang, Yi Yang, George S. Hussey, Breege V. Howley, Mohammad Salem, Brian Riesenberg, et al. 2017. “Membrane-Organizing Protein Moesin Controls Treg Differentiation and Antitumor Immunity via TGF-β Signaling.” *The Journal of Clinical Investigation* 127 (4): 1321–37. https://doi.org/10.1172/JCI89281.

Arantes, Lidia Maria Rebolho Batista, Ana Carolina Laus, Matias Eliseo Melendez, Ana Carolina de Carvalho, Bruna Pereira Sorroche, Pedro Rafael Martins De Marchi, Adriane Feijó Evangelista, Cristovam Scapulatempo-Neto, Luciano de Souza Viana, and André Lopes Carvalho. 2016. “MiR-21 as Prognostic Biomarker in Head and Neck Squamous Cell Carcinoma Patients

Undergoing an Organ Preservation Protocol." *Oncotarget* 8 (6): 9911–21. https://doi.org/10.18632/oncotarget.14253.

Argiris, Athanassios, Michalis V. Karamouzis, David Raben, and Robert L. Ferris. 2008. "Head and Neck Cancer." *The Lancet* 371 (9625): 1695–1709.

Avissar, Michele, Brock C. Christensen, Karl T. Kelsey, and Carmen J. Marsit. 2009. "MicroRNA Expression Ratio Is Predictive of Head and Neck Squamous Cell Carcinoma." *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 15 (8): 2850–55. https://doi.org/10.1158/1078-0432.CCR-08-3131.

Bannister, Andrew J., and Tony Kouzarides. 2011. "Regulation of Chromatin by Histone Modifications." *Cell Research* 21 (3): 381–95. https://doi.org/10.1038/cr.2011.22.

Barros, S.P., and S. Offenbacher. 2009. "Epigenetics." *Journal of Dental Research* 88 (5): 400–408. https://doi.org/10.1177/0022034509335868.

Beachler, Daniel C., Gwendolyne Jenkins, Mahboobeh Safaeian, Aimée R. Kreimer, and Nicolas Wentzensen. 2016. "Natural Acquired Immunity Against Subsequent Genital Human Papillomavirus Infection: A Systematic Review and Meta-Analysis." *The Journal of Infectious Diseases* 213 (9): 1444–54. https://doi.org/10.1093/infdis/jiv753.

Bellanger, Sophie, Chye Ling Tan, Yue Zhen Xue, Sébastien Teissier, and Françoise Thierry. 2011. "Tumor Suppressor or Oncogene? A Critical Role of the Human Papillomavirus (HPV) E2 Protein in Cervical Cancer Progression." *American Journal of Cancer Research* 1 (3): 373–89.

Bernard, Hans-Ulrich. 2002. "Gene Expression of Genital Human Papillomaviruses and Considerations on Potential Antiviral Approaches." *Antiviral Therapy* 7 (4): 219–37.

Bernard, Hans-Ulrich, Robert D Burk, Zigui Chen, Koenraad van Doorslaer, Harald zur Hausen, and Ethel-Michele de Villiers. 2010. "Classification of Papillomaviruses (PVs) Based on 189 PV Types and Proposal of Taxonomic Amendments." *Virology* 401 (1): 70–79. https://doi.org/10.1016/j.virol.2010.02.002.

Bissell, Mina J., and Derek Radisky. 2001. "Putting Tumours in Context." *Nature Reviews Cancer* 1 (1): 46–54. https://doi.org/10.1038/35094059.

Bosch, F. Xavier, A. Lorincz, N. Muñoz, C. J. L. M. Meijer, and K. V. Shah. 2002. "The Causal Relation Between Human Papillomavirus and Cervical Cancer." *Journal of Clinical Pathology* 55 (4): 244–65.

Boscolo-Rizzo, Paolo, Annarosa Del Mistro, Francesco Bussu, Valentina Lupato, Lorena Baboci, Giovanni Almadori, Maria Cristina Da Mosto, and Gaetano Paludetti. 2013. "New Insights into Human Papillomavirus-Associated Head and Neck Squamous Cell Carcinoma." *Acta Otorhinolaryngologica Italica* 33 (2): 77–87.

Boscolo-Rizzo, Paolo, Carlo Furlan, Valentina Lupato, Jerry Polesel, and Elisabetta Fratta. 2017. "Novel Insights into Epigenetic Drivers of Oropharyngeal Squamous Cell Carcinoma: Role of HPV and Lifestyle Factors." *Clinical Epigenetics* 9 (1). https://doi.org/10.1186/s13148-017-0424-5.

Boscolo-Rizzo, Paolo, Michael Pawlita, and Dana Holzinger. 2016. "From HPV-Positive towards HPV-Driven Oropharyngeal Squamous Cell Carcinomas." *Cancer Treatment Reviews* 42 (January): 24–29. https://doi.org/10.1016/j.ctrv.2015.10.009.

Boscolo-Rizzo, Paolo, Manuel Zorzi, Annarosa Del Mistro, Maria Cristina Da Mosto, Giancarlo Tirelli, Carlotta Buzzoni, Massimo Rugge, Jerry Polesel, and Stefano Guzzinati. 2018. "The Evolution of the Epidemiological Landscape of Head and Neck Cancer in Italy: Is There Evidence for an Increase in the Incidence of Potentially HPV-Related Carcinomas?" *PLoS ONE* 13 (2). https://doi.org/10.1371/journal.pone.0192621.

Bouvard, Véronique, Robert Baan, Kurt Straif, Yann Grosse, Béatrice Secretan, Fatiha El Ghissassi, Lamia Benbrahim-Tallaa, et al. 2009. "A Review of Human Carcinogens--Part B: Biological Agents." *The Lancet Oncology* 10 (4): 321–22.

Broglie, Martina A., Sandro J. Stoeckli, Rafael Sauter, Philippe Pasche, Antoine Reinhard, Laurence de Leval, Gerhard F. Huber, et al. 2017. "Impact of Human Papillomavirus on Outcome in Patients with Oropharyngeal Cancer Treated with Primary Surgery." *Head & Neck* 39 (10): 2004–15. https://doi.org/10.1002/hed.24865.

Bufalino, Andreia, Nilva K. Cervigne, Carine Ervolino de Oliveira, Felipe Paiva Fonseca, Priscila Campioni Rodrigues, Carolina Carneiro Soares Macedo, Lays Martin Sobral, et al. 2015. "Low MiR-143/MiR-145 Cluster Levels Induce Activin A Overexpression in Oral Squamous Cell Carcinomas, Which Contributes to Poor Prognosis." *PLoS ONE* 10 (8). https://doi.org/10.1371/journal.pone.0136599.

Bulk, S., J. Berkhof, N. W. J. Bulkmans, G. D. Zielinski, L. Rozendaal, F. J. van Kemenade, P. J. F. Snijders, and C. J. L. M. Meijer. 2006. "Preferential Risk of HPV16 for Squamous Cell Carcinoma and of HPV18 for Adenocarcinoma of the Cervix Compared to Women with Normal Cytology in The Netherlands." *British Journal of Cancer* 94 (1): 171–75. https://doi.org/10.1038/sj.bjc.6602915.

Busato, Florence, Emelyne Dejeux, Hafida El Abdalaoui, Ivo Glynne Gut, and Jörg Tost. 2018. "Quantitative DNA Methylation Analysis at Single-Nucleotide Resolution by Pyrosequencing®." *Methods in Molecular Biology (Clifton, N.J.)* 1708: 427–45. https://doi.org/10.1007/978-1-4939-7481-8_22.

Butel, Janet S. 2000. "Viral Carcinogenesis: Revelation of Molecular Mechanisms and Etiology of Human Disease." *Carcinogenesis* 21 (3): 405–26. https://doi.org/10.1093/carcin/21.3.405.

Cai, J., Y. Zhang, S. Huang, M. Yan, J. Li, T. Jin, and S. Bao. 2017. "MiR-100-5p, MiR-199a-3p and MiR-199b-5p Induce Autophagic Death of Endometrial Carcinoma Cell through Targeting MTOR."

International Journal of Clinical and Experimental Pathology 10 (9): 9262–72.

Cai, Kemin, Yu Wang, and Xueli Bao. 2011. "MiR-106b Promotes Cell Proliferation via Targeting RB in Laryngeal Carcinoma." *Journal of Experimental & Clinical Cancer Research : CR* 30 (1): 73. https://doi.org/10.1186/1756-9966-30-73.

Cantrell, S. C., B. W. Peck, G. Li, Q. Wei, E. M. Sturgis, and L. E. Ginsberg. 2013. "Differences in Imaging Characteristics of HPV-Positive and HPV-Negative Oropharyngeal Cancers: A Blinded Matched-Pair Analysis." *American Journal of Neuroradiology* 34 (10): 2005–9. https://doi.org/10.3174/ajnr.A3524.

Cao, Jian, and Qin Yan. 2012. "Histone Ubiquitination and Deubiquitination in Transcription, DNA Damage Response, and Cancer." *Frontiers in Oncology* 2: 26. https://doi.org/10.3389/fonc.2012.00026.

Cao, Pengyu, Liang Zhou, Jin Zhang, Fengyun Zheng, Huijun Wang, Duan Ma, and Jie Tian. 2013. "Comprehensive Expression Profiling of MicroRNAs in Laryngeal Squamous Cell Carcinoma." *Head & Neck* 35 (5): 720–28. https://doi.org/10.1002/hed.23011.

CDC. 2011. "Current Cigarette Smoking Among Adults — United States, 2011." 2011. https://www.cdc.gov/mmwr/preview/mmwrhtml/mm6144a2.htm.

CDC, Centers for Disease Control and Prevention. 2017. "HPV Statistics | CDC." 2017. https://www.cdc.gov/std/hpv/stats.htm.

CDC, Centers for Disease Control and Prevention. 2018. "Genital HPV Infection - CDC Fact Sheet.," 3.

Chang, Steven S., Wei Wen Jiang, Ian Smith, Luana M. Poeta, Shahnaz Begum, Chad Glazer, Shannon Shan, William Westra, David Sidransky, and Joseph A. Califano. 2008. "MicroRNA Alterations in Head and Neck Squamous Cell Carcinoma." *International Journal of Cancer. Journal International Du Cancer* 123 (12): 2791–97. https://doi.org/10.1002/ijc.23831.

Chen, X., X. Guo, H. Zhang, Y. Xiang, J. Chen, Y. Yin, X. Cai, et al. 2009. "Role of MiR-143 Targeting KRAS in Colorectal Tumorigenesis." *Oncogene* 28 (10): 1385–92. https://doi.org/10.1038/onc.2008.474.

Chen, Zigui, Mark Schiffman, Rolando Herrero, Rob DeSalle, Kathryn Anastos, Michel Segondy, Vikrant V. Sahasrabuddhe, Patti E. Gravitt, Ann W. Hsing, and Robert D. Burk. 2011. "Evolution and Taxonomic Classification of Human Papillomavirus 16 (HPV16)-Related Variant Genomes: HPV31, HPV33, HPV35, HPV52, HPV58 and HPV67." *PLoS ONE* 6 (5). https://doi.org/10.1371/journal.pone.0020183.

Chen, Zujian, Tianwei Yu, Robert J. Cabay, Yi Jin, Ishrat Mahjabeen, Xianghong Luan, Lei Huang, Yang Dai, and Xiaofeng Zhou. 2017. "MiR-486-3p, MiR-139-5p, and MiR-21 as Biomarkers for the Detection of Oral Tongue Squamous Cell Carcinoma." *Biomarkers in Cancer* 9 (January): 1–8.

https://doi.org/10.4137/BIC.S40981.

Cheng, G. 1995. “Divergent Human Papillomavirus Type 16 Variants Ar... [J Infect Dis. 1995] - PubMed Result.” 1995. http://www.ncbi.nlm.nih.gov/pubmed?term=Divergent%20human%20papillomavirus%20type%2016%20variants%20are%20serologically%20cross-reactive.

Cheng, Jinzhang, Junjun Chen, Zonggui Wang, Dan Yu, and Yuanzhang Zu. 2017. “The Functional Role of MicroRNAs in Laryngeal Carcinoma.” *Open Life Sciences* 12 (1). https://doi.org/10.1515/biol-2017-0054.

Chung, Christine H, Joel S Parker, Gamze Karaca, Junyuan Wu, William K Funkhouser, Dominic Moore, Dale Butterfoss, et al. 2004. “Molecular Classification of Head and Neck Squamous Cell Carcinomas Using Patterns of Gene Expression.” *Cancer Cell* 5 (5): 489–500. https://doi.org/10.1016/S1535-6108(04)00112-6.

Cioffi, Michele, Sara M Trabulo, Mireia Vallespinos, Deepak Raj, Tony Bou Kheir, Meng-Lay Lin, Julfa Begum, et al. 2017. “The MiR-25-93-106b Cluster Regulates Tumor Metastasis and Immune Evasion via Modulation of CXCL12 and PD-L1.” *Oncotarget* 8 (13): 21609–25. https://doi.org/10.18632/oncotarget.15450.

Cloonan, Nicole, Shivangi Wani, Qinying Xu, Jian Gu, Kristi Lea, Sheila Heater, Catalin Barbacioru, et al. 2011. “MicroRNAs and Their IsomiRs Function Cooperatively to Target Common Biological Pathways.” *Genome Biology* 12 (12): R126. https://doi.org/10.1186/gb-2011-12-12-r126.

CNCR. 2016. “Croatian National Cancer Registry. Cancer incidence and mortality in Croatia. Zagreb: Croatian Institute of Public Health.” 2016. https://www.hzjz.hr/.

Conway, M.J., and C. Meyers. 2009. “Replication and Assembly of Human Papillomaviruses.” *Journal of Dental Research* 88 (4): 307–17. https://doi.org/10.1177/0022034509333446.

Costa, Pedro M., and Maria C. Pedroso de Lima. 2013. “MicroRNAs as Molecular Targets for Cancer Therapy: On the Modulation of MicroRNA Expression.” *Pharmaceuticals* 6 (10): 1195–1220. https://doi.org/10.3390/ph6101195.

Cramer, John D., Kate E. Hicks, Alfred W. Rademaker, Urjeet A. Patel, and Sandeep Samant. 2018. “Validation of the Eighth Edition American Joint Committee on Cancer Staging System for Human Papillomavirus-Associated Oropharyngeal Cancer.” *Head & Neck* 40 (3): 457–66. https://doi.org/10.1002/hed.24974.

Crow, James Mitchell. 2012. “HPV: The Global Burden.” *Nature* 488 (7413): S2-3. https://doi.org/10.1038/488S2a.

Cullen, Michael, Joseph F. Boland, Mark Schiffman, Xijun Zhang, Nicolas Wentzensen, Qi Yang, Zigui Chen, et al. 2015. “Deep Sequencing of HPV16 Genomes: A New High-Throughput Tool for Exploring the Carcinogenicity and Natural History of HPV16 Infection.” *Papillomavirus Research (Amsterdam, Netherlands)* 1 (December): 3–11. https://doi.org/10.1016/j.pvr.2015.05.004.

Cutts, FT, S Franceschi, S Goldie, X Castellsague, S de Sanjose, G Garnett, WJ Edmunds, et al. 2007. "Human Papillomavirus and HPV Vaccines: A Review." *Bulletin of the World Health Organization* 85 (9): 719–26. https://doi.org/10.2471/BLT.06.038414.

Dahlstrom, Kristina R., Diana Bell, Duncan Hanby, Guojun Li, Li-E Wang, Qingyi Wei, Michelle D. Williams, and Erich M. Sturgis. 2015. "Socioeconomic Characteristics of Patients with Oropharyngeal Carcinoma According to Tumor HPV Status, Patient Smoking Status, and Sexual Behavior." *Oral Oncology* 51 (9): 832–38. https://doi.org/10.1016/j.oraloncology.2015.06.005.

Dedhia, Pratiksha, Shivraj Tarale, Gargi Dhongde, Rashmi Khadapkar, and Bibhu Das. 2007. "Evaluation of DNA Extraction Methods and Real Time PCR Optimization on Formalin-Fixed Paraffin-Embedded Tissues." *Asian Pacific Journal of Cancer Prevention: APJCP* 8 (1): 55–59.

Delaney, Colin, Sanjay K. Garg, and Raymond Yung. 2015. "Analysis of DNA Methylation by Pyrosequencing." *Methods in Molecular Biology (Clifton, N.J.)* 1343: 249–64. https://doi.org/10.1007/978-1-4939-2963-4_19.

Descamps, Géraldine, Yasemin Karaca, Jérôme R Lechien, Nadège Kindt, Christine Decaestecker, Myriam Remmelink, Denis Larsimont, et al. 2016. "Classical Risk Factors, but Not HPV Status, Predict Survival after Chemoradiotherapy in Advanced Head and Neck Cancer Patients." *Journal of Cancer Research and Clinical Oncology* 142 (10): 2185–96. https://doi.org/10.1007/s00432-016-2203-7.

Dessain, S. K., H. Yu, R. R. Reddel, R. L. Beijersbergen, and R. A. Weinberg. 2000. "Methylation of the Human Telomerase Gene CpG Island." *Cancer Research* 60 (3): 537–41.

Dietrich, Dimo, Barbara Uhl, Verena Sailer, Emily Eva Holmes, Maria Jung, Sebastian Meller, and Glen Kristiansen. 2013. "Improved PCR Performance Using Template DNA from Formalin-Fixed and Paraffin-Embedded Tissues by Overcoming PCR Inhibition." *PLoS ONE* 8 (10). https://doi.org/10.1371/journal.pone.0077771.

Dixit, Ronak, Joel L. Weissfeld, David O. Wilson, Paula Balogh, Pamela Sufka, Jill M. Siegfried, Jennifer R. Grandis, and Brenda Diergaarde. 2015. "Incidence of Head and Neck Squamous Cell Carcinoma among Subjects at High-Risk for Lung Cancer: Results from the Pittsburgh Lung Screening Study." *Cancer* 121 (9): 1431–35. https://doi.org/10.1002/cncr.29189.

Dok, Rüveyda, and Sandra Nuyts. 2016. "HPV Positive Head and Neck Cancers: Molecular Pathogenesis and Evolving Treatment Strategies." *Cancers* 8 (4). https://doi.org/10.3390/cancers8040041.

Dongen, Jenny van, Michel G. Nivard, Gonneke Willemsen, Jouke-Jan Hottenga, Quinta Helmer, Conor V. Dolan, Erik A. Ehli, et al. 2016. "Genetic and Environmental Influences Interact with Age and Sex in Shaping the Human Methylome." *Nature Communications* 7 (April): 11115. https://doi.org/10.1038/ncomms11115.

Doorbar, J. 2005. “The Papillomavirus Life Cycle.” *Journal of Clinical Virology* 32: 7–15.

Doorbar, John. 2013. “The E4 Protein; Structure, Function and Patterns of Expression.” *Virology* 445 (1–2): 80–98. https://doi.org/10.1016/j.virol.2013.07.008.

Doorbar, John, Nagayasu Egawa, Heather Griffin, Christian Kranjec, and Isao Murakami. 2015. “Human Papillomavirus Molecular Biology and Disease Association.” *Reviews in Medical Virology* 25 Suppl 1 (March): 2–23. https://doi.org/10.1002/rmv.1822.

Doorbar, John, Wim Quint, Lawrence Banks, Ignacio G. Bravo, Mark Stoler, Tom R. Broker, and Margaret A. Stanley. 2012. “The Biology and Life-Cycle of Human Papillomaviruses.” *Vaccine* 30, Supplement 5 (November): F55–70. https://doi.org/10.1016/j.vaccine.2012.06.083.

Dricu, Anica, Stefana Oana Purcaru, Raluca Budiu, Roxana Ola, Daniela Elise Tache, and Anda Vlad. 2012. “Epigenetic Alteration of Receptor Tyrosine Kinases in Cancer | IntechOpen.” 2012. /books/dna-methylation-from-genomics-to-technology/epigenetic-alteration-of-receptor-tyrosine-kinases-in-cancer.

Edge, Stephen B., and Carolyn C. Compton. 2010. “The American Joint Committee on Cancer: The 7th Edition of the AJCC Cancer Staging Manual and the Future of TNM.” *Annals of Surgical Oncology* 17 (6): 1471–74. https://doi.org/10.1245/s10434-010-0985-4.

Egawa, Nagayasu, and John Doorbar. 2017. “The Low-Risk Papillomaviruses.” *Virus Research*, Human Papillomaviruses, 231 (March): 119–27. https://doi.org/10.1016/j.virusres.2016.12.017.

Egawa, Nagayasu, Tomomi Nakahara, Shin-ichi Ohno, Mako Narisawa-Saito, Takashi Yugawa, Masatoshi Fujita, Kenji Yamato, Yukikazu Natori, and Tohru Kiyono. 2012. “The E1 Protein of Human Papillomavirus Type 16 Is Dispensable for Maintenance Replication of the Viral Genome.” *Journal of Virology* 86 (6): 3276–83. https://doi.org/10.1128/JVI.06450-11.

Ehrlich, Melanie. 2009. “DNA Hypomethylation in Cancer Cells.” *Epigenomics* 1 (2): 239–59. https://doi.org/10.2217/epi.09.33.

El-Naggar, Adel K., and William H. Westra. 2012. “P16 Expression as a Surrogate Marker for HPV-Related Oropharyngeal Carcinoma: A Guide for Interpretative Relevance and Consistency.” *Head & Neck* 34 (4): 459–61. https://doi.org/10.1002/hed.21974.

ENCODE. 2018. “ENCODE: Encyclopedia of DNA Elements – ENCODE.” 2018. https://www.encodeproject.org/.

Estécio, Marcos R.H., and Jean-Pierre J. Issa. 2011. “Dissecting DNA Hypermethylation in Cancer.” *Febs Letters* 585 (13): 2078–86. https://doi.org/10.1016/j.febslet.2010.12.001.

Fakhry, Carole, Qiang Zhang, Phuc Felix Nguyen-Tan, David Rosenthal, Adel El-Naggar, Adam S. Garden, Denis Soulieres, et al. 2014. “Human Papillomavirus and Overall Survival after Progression of Oropharyngeal Squamous Cell Carcinoma.” *Journal of Clinical Oncology: Official*

Journal of the American Society of Clinical Oncology 32 (30): 3365–73. https://doi.org/10.1200/JCO.2014.55.1937.

FANTOM. 2018. “FANTOM.” 2018. http://fantom.gsc.riken.jp/.

Ferris, Robert L. 2015. “Immunology and Immunotherapy of Head and Neck Cancer.” *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology* 33 (29): 3293–3304. https://doi.org/10.1200/JCO.2015.61.1509.

Freitas, Antonio Carlos de, Talita Helena Araújo de Oliveira, Marconi Rego Barros, and Aldo Venuti. 2017. “HrHPV E5 Oncoprotein: Immune Evasion and Related Immunotherapies.” *Journal of Experimental & Clinical Cancer Research : CR* 36 (May). https://doi.org/10.1186/s13046-017-0541-1.

Frigola, Jordi, Xavier Solé, Maria F. Paz, Victor Moreno, Manel Esteller, Gabriel Capellà, and Miguel A. Peinado. 2005. “Differential DNA Hypermethylation and Hypomethylation Signatures in Colorectal Cancer.” *Human Molecular Genetics* 14 (2): 319–26. https://doi.org/10.1093/hmg/ddi028.

Fukaya, Takashi, and Yukihide Tomari. 2012. “MicroRNAs Mediate Gene Silencing via Multiple Different Pathways in Drosophila.” *Molecular Cell* 48 (6): 825–36. https://doi.org/10.1016/j.molcel.2012.09.024.

Fukumoto, Ichiro, Naoko Kikkawa, Ryosuke Matsushita, Mayuko Kato, Akira Kurozumi, Rika Nishikawa, Yusuke Goto, et al. 2016. “Tumor-Suppressive MicroRNAs (MiR-26a/b, MiR-29a/b/c and MiR-218) Concertedly Suppressed Metastasis-Promoting LOXL2 in Head and Neck Squamous Cell Carcinoma.” *Journal of Human Genetics* 61 (2): 109–18. https://doi.org/10.1038/jhg.2015.120.

Gao, Ge, Hiram A. Gay, Rebecca D. Chernock, Tian R. Zhang, Jingqin Luo, Wade L. Thorstad, James S. Lewis, and Xiaowei Wang. 2013. “A MicroRNA Expression Signature for the Prognosis of Oropharyngeal Squamous Cell Carcinoma.” *Cancer* 119 (1): 72–80. https://doi.org/10.1002/cncr.27696.

Gao, Yuzhe, Huajian Chen, Qing Ni, Qi Jia, and Shiqiong Su. 2016. “Effect of MiR-106b on the Proliferation, Invasion, and Migration of Breast Cancer MCF-7 Cells,” 6.

GBD. 2015. “Global, Regional, and National Age–Sex Specific All-Cause and Cause-Specific Mortality for 240 Causes of Death, 1990–2013: A Systematic Analysis for the Global Burden of Disease Study 2013.” *The Lancet* 385 (9963): 117–71. https://doi.org/10.1016/S0140-6736(14)61682-2.

GCS. 2018. “GeneCards - Human Genes | Gene Database | Gene Search.” 2018. https://www.genecards.org/.

Gillison, Maura L. 2004. “Human Papillomavirus-Associated Head and Neck Cancer Is a Distinct

Epidemiologic, Clinical, and Molecular Entity." *Seminars in Oncology* 31 (6): 744–54.

Gillison, Maura L., Anil K. Chaturvedi, William F. Anderson, and Carole Fakhry. 2015. "Epidemiology of Human Papillomavirus-Positive Head and Neck Squamous Cell Carcinoma." *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology* 33 (29): 3235–42. https://doi.org/10.1200/JCO.2015.61.6995.

Gillison, Maura L., Gypsyamber D'Souza, William Westra, Elizabeth Sugar, Weihong Xiao, Shahnaz Begum, and Raphael Viscidi. 2008. "Distinct Risk Factor Profiles for Human Papillomavirus Type 16-Positive and Human Papillomavirus Type 16-Negative Head and Neck Cancers." *Journal of the National Cancer Institute* 100 (6): 407–20. https://doi.org/10.1093/jnci/djn025.

Gillison, M.L., W.M. Koch, R.B. Capone, M. Spafford, W.H. Westra, L. Wu, M.L. Zahurak, et al. 2000. "Evidence for a Causal Association between Human Papillomavirus and a Subset of Head and Neck Cancers." *Journal of the National Cancer Institute* 92 (9): 709–720.

Giuliano, Anna R., Alan G. Nyitray, Aimée R. Kreimer, Christine M. Pierce Campbell, Marc T. Goodman, Staci L. Sudenga, Joseph Monsonego, and Silvia Franceschi. 2015. "EUROGIN 2014 Roadmap: Differences in HPV Infection Natural History, Transmission, and HPV-Related Cancer Incidence by Gender and Anatomic Site of Infection." *International Journal of Cancer. Journal International Du Cancer* 136 (12): 2752–60. https://doi.org/10.1002/ijc.29082.

Globocan. 2012. "Globocan 2012 - Home." 2012. http://globocan.iarc.fr/Default.aspx.

Goel, Rajeev K, and Jelena Budak. 2007. "Smoking Patterns in Croatia and Comparisons with European Nations," 6.

Graham, Sheila V., and Arwa Ali A. Faizo. 2017. "Control of Human Papillomavirus Gene Expression by Alternative Splicing." *Virus Research*, Human Papillomaviruses, 231 (March): 83–95. https://doi.org/10.1016/j.virusres.2016.11.016.

Grahovac, Blaženka, Anka Dori, Željka Hruškar, Ita Hadžisejdi, and Maja Grahovac. 2012. "Human Papillomavirus Infection in Croatian Men: Prevalence and HPV Type Distribution." *Sexually Transmitted Infections*, 16.

Gravitt, Patti E., and Rachel L. Winer. 2017. "Natural History of HPV Infection across the Lifespan: Role of Viral Latency." *Viruses* 9 (10). https://doi.org/10.3390/v9100267.

Gravitt, PE, CL Peyton, TQ Alessi, CM Wheeler, F. Coutlee, A. Hildesheim, MH Schiffman, DR Scott, and RJ Apple. 2000. "Improved Amplification of Genital Human Papillomaviruses." *Journal of Clinical Microbiology* 38 (1): 357–361.

Grce, M, and P Davies. 2008. "Human Papillomavirus Testing for Primary Cervical Cancer Screening." Expert Review of Molecular Diagnostics 8 (5): 599–605. https://doi.org/10.1586/14737159.8.5.599.

Grce, M, K Husnjak, M Skerlev, J Lipozencić, and K Pavelić. 2000. “Detection and Typing of Human Papillomaviruses by Means of Polymerase Chain Reaction and Fragment Length Polymorphism in Male Genital Lesions.” *Anticancer Research* 20 (3B): 2097–2102.

Grce, Magdalena. 2009. “Primary and Secondary Prevention of Cervical Cancer.” *Expert Review of Molecular Diagnostics* 9 (8): 851–57. https://doi.org/10.1586/erm.09.64.

Grce, Magdalena, Ivan Sabol, and Nina Milutin Gašperov. 2012a. “Burden and Prevention of HPV Related Diseases: Situation in Croatia.” *Periodicum Biologorum* 114 (2): 175–86.

Grce, Magdalena, Ivan Sabol, and Nina Milutin Gašperov. 2012b. “The Transforming Properties of Human Papillomavirus Oncoproteins.” *Periodicum Biologorum* 114 (4): 479–487.

Guo, Li, Jiafeng Yu, Tingming Liang, and Quan Zou. 2016. “MiR-IsomiRExp: A Web-Server for the Analysis of Expression of MiRNA at the MiRNA/IsomiR Levels.” *Scientific Reports* 6 (March): 23700. https://doi.org/10.1038/srep23700.

Hafkamp, H. C., J. J. Manni, A. Haesevoets, A. C. Voogd, M. Schepers, F. J. Bot, A. H. N. Hopman, F. C. S. Ramaekers, and Ernst-Jan M. Speel. 2008. “Marked Differences in Survival Rate between Smokers and Nonsmokers with HPV 16-Associated Tonsillar Carcinomas.” *International Journal of Cancer* 122 (12): 2656–64. https://doi.org/10.1002/ijc.23458.

Hafkamp, Harriet C., Ernst J.M. Speel, Annick Haesevoets, Fredrik J. Bot, Winand N.M. Dinjens, Frans C.S. Ramaekers, Anton H.N. Hopman, and Johannes J. Manni. 2003. “A Subset of Head and Neck Squamous Cell Carcinomas Exhibits Integration of HPV 16/18 DNA and Overexpression of P16INK4A and P53 in the Absence of Mutations in P53 Exons 5-8.” *International Journal of Cancer* 107 (3): 394–400. https://doi.org/10.1002/ijc.11389.

Hammond, Scott M. 2015. “An Overview of MicroRNAs.” *Advanced Drug Delivery Reviews* 87 (June): 3–14. https://doi.org/10.1016/j.addr.2015.05.001.

Hashimoto, Hideharu, Paula M. Vertino, and Xiaodong Cheng. 2010. “Molecular Coupling of DNA Methylation and Histone Methylation.” *Epigenomics* 2 (5): 657–69. https://doi.org/10.2217/epi.10.44.

Hausen, H. zur. 1977. “Human Papillomaviruses and Their Possible Role in Squamous Cell Carcinomas.” *Current Topics in Microbiology and Immunology* 78: 1–30.

Hausen, Harald zur. 2009. “Papillomaviruses in the Causation of Human Cancers — a Brief Historical Account.” *Virology* 384 (2): 260–65. https://doi.org/10.1016/j.virol.2008.11.046.

Herceg, Zdenko, and Pierre Hainaut. 2007. “Genetic and Epigenetic Alterations as Biomarkers for Cancer Detection, Diagnosis and Prognosis.” *Molecular Oncology* 1 (1): 26–41. https://doi.org/10.1016/j.molonc.2007.01.004.

Herdman, M. Trent, Mark R. Pett, Ian Roberts, William O. F. Alazawi, Andrew E. Teschendorff,

Xiao-Yin Zhang, Margaret A. Stanley, and Nicholas Coleman. 2006. “Interferon-Beta Treatment of Cervical Keratinocytes Naturally Infected with Human Papillomavirus 16 Episomes Promotes Rapid Reduction in Episome Numbers and Emergence of Latent Integrants.” *Carcinogenesis* 27 (11): 2341–53. https://doi.org/10.1093/carcin/bgl172.

Herman, J.G., J.R. Graff, S. Myöhänen, B.D. Nelkin, and S.B. Baylin. 1996. “Methylation-Specific PCR: A Novel PCR Assay for Methylation Status of CpG Islands.” *Proceedings of the National Academy of Sciences* 93 (18): 9821.

Holmes, Allyson, Sonia Lameiras, Emmanuelle Jeannot, Yannick Marie, Laurent Castera, Xavier Sastre-Garau, and Alain Nicolas. 2016. “Mechanistic Signatures of HPV Insertions in Cervical Carcinomas.” *NPJ Genomic Medicine* 1: 16004. https://doi.org/10.1038/npjgenmed.2016.4.

Holmgren, Sigrid C., Nicole A. Patterson, Michelle A. Ozbun, and Paul F. Lambert. 2005. “The Minor Capsid Protein L2 Contributes to Two Steps in the Human Papillomavirus Type 31 Life Cycle.” *Journal of Virology* 79 (7): 3938–48. https://doi.org/10.1128/JVI.79.7.3938-3948.2005.

Hong, A. M., T. A. Dobbins, C. S. Lee, D. Jones, G. B. Harnett, B. K. Armstrong, J. R. Clark, et al. 2010. “Human Papillomavirus Predicts Outcome in Oropharyngeal Cancer in Patients Treated Primarily with Surgery or Radiation Therapy.” *British Journal of Cancer* 103 (10): 1510–17. https://doi.org/10.1038/sj.bjc.6605944.

Hosaka, Masayoshi, Hiromasa Fujita, Sharon JB Hanley, Takayuki Sasaki, Yozo Shirakawa, Mitsuharu Abiko, Masataka Kudo, et al. 2013. “Incidence Risk of Cervical Intraepithelial Neoplasia 3 or More Severe Lesions Is a Function of Human Papillomavirus Genotypes and Severity of Cytological and Histological Abnormalities in Adult Japanese Women.” *International Journal of Cancer* 132 (2): 327–34. https://doi.org/10.1002/ijc.27680.

HPV Centre. 2017. “HPV INFORMATION CENTRE.” 2017. http://www.hpvcentre.net/.

Hu, Lulin, Ming Guo, Zhi He, Justin Thornton, Larry S. McDaniel, and Michael D. Hughson. 2005. “Human Papillomavirus Genotyping and P16INK4a Expression in Cervical Intraepithelial Neoplasia of Adolescents.” *Modern Pathology: An Official Journal of the United States and Canadian Academy of Pathology, Inc* 18 (2): 267–73. https://doi.org/10.1038/modpathol.3800290.

Hu, Wangxiong, Tingzhang Wang, Jianhong Xu, and Hongzhi Li. 2014. “MicroRNA Mediates DNA Methylation of Target Genes.” *Biochemical and Biophysical Research Communications* 444 (4): 676–81. https://doi.org/10.1016/j.bbrc.2014.01.171.

Hu, Zheng, Da Zhu, Wei Wang, Weiyang Li, Wenlong Jia, Xi Zeng, Wencheng Ding, et al. 2015. “Genome-Wide Profiling of HPV Integration in Cervical Cancer Identifies Clustered Genomic Hot Spots and a Potential Microhomology-Mediated Integration Mechanism.” *Nature Genetics* 47 (2): 158–63. https://doi.org/10.1038/ng.3178.

Huang, Fei, and Ye-Guang Chen. 2012. “Regulation of TGF-β Receptor Activity.” *Cell & Bioscience*

2 (March): 9. https://doi.org/10.1186/2045-3701-2-9.

Huang, Qiang, Guo Pei Yu, Steven A. McCormick, Juan Mo, Bhakti Datta, Manoj Mahimkar, Philip Lazarus, Alejandro A. Schäffer, Richard Desper, and Stimson P. Schantz. 2002. "Genetic Differences Detected by Comparative Genomic Hybridization in Head and Neck Squamous Cell Carcinomas from Different Tumor Sites: Construction of Oncogenetic Trees for Tumor Progression." *Genes, Chromosomes & Cancer* 34 (2): 224–33. https://doi.org/10.1002/gcc.10062.

Husain, Nuzhat, and Azfar Neyaz. 2017. "Human Papillomavirus Associated Head and Neck Squamous Cell Carcinoma: Controversies and New Concepts." *Journal of Oral Biology and Craniofacial Research* 7 (3): 198–205. https://doi.org/10.1016/j.jobcr.2017.08.003.

Hussein, Aisha A., Marco N. Helder, Jan G. de Visscher, C. René Leemans, Boudewijn J. Braakhuis, Henrica C. W. de Vet, and Tymour Forouzanfar. 2017. "Global Incidence of Oral and Oropharynx Cancer in Patients Younger than 45 Years versus Older Patients: A Systematic Review." *European Journal of Cancer (Oxford, England: 1990)* 82: 115–27. https://doi.org/10.1016/j.ejca.2017.05.026.

Huyett, Phillip, Mark Gilbert, Lijun Liu, Robert L. Ferris, and Seungwon Kim. 2017. "A Model for Perineural Invasion in Head and Neck Squamous Cell Carcinoma." *Journal of Visualized Experiments: JoVE*, no. 119 (January). https://doi.org/10.3791/55043.

IARC. 1995. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Human Papillomaviruses*. Vol. 64. Lyon, France: International Agency for Research on Cancer, World Health Organization.

IARC, WHO. 2005. *IARC Handbooks of Cancer Prevention Volume 10: Cervix Cancer Screening*. Vol. 10. Lyon, France: IARC Press.

IARC, WHO. 2007. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans Volume 90 Human Papillomaviruses*. Vol. 90. Lyon, France: Iarc Press.

IARC, WHO. ed. 2012. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Volume 100 B, Biological Agents: This Publication Represents the Views and Expert Opinions of an IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, Which Met in Lyon, 24 February - 03 March 2009*. Lyon: IARC.

IARC, WHO. 2018. "IARC Monographs - Classification of Carcinogens." 2018. http://monographs.iarc.fr/ENG/Classification/index.php.

Illumina. 2018. "Methylation Arrays | Interrogate Methylation Sites across the Genome." 2018. https://www.illumina.com/techniques/microarrays/methylation-arrays.html.

Ivanovska, Irena, Alexey S. Ball, Robert L. Diaz, Jill F. Magnus, Miho Kibukawa, Janell M. Schelter, Sumire V. Kobayashi, et al. 2008. "MicroRNAs in the MiR-106b Family Regulate P21/CDKN1A and Promote Cell Cycle Progression." *Molecular and Cellular Biology* 28 (7): 2167–74.

https://doi.org/10.1128/MCB.01977-07.

Jacobs, M V, A M de Roda Husman, A J van den Brule, P J Snijders, C J Meijer, and J M Walboomers. 1995. “Group-Specific Differentiation between High- and Low-Risk Human Papillomavirus Genotypes by General Primer-Mediated PCR and Two Cocktails of Oligonucleotide Probes.” *Journal of Clinical Microbiology* 33 (4): 901–5.

Jeon, Geoung A., Ju-Seog Lee, Vyomesh Patel, J. Silvio Gutkind, Snorri S. Thorgeirsson, Eun Cheol Kim, In-Sun Chu, Panomwat Amornphimoltham, and Myung Hee Park. 2004. “Global Gene Expression Profiles of Human Head and Neck Squamous Carcinoma Cell Lines.” *International Journal of Cancer* 112 (2): 249–58. https://doi.org/10.1002/ijc.20399.

Jeon, S, B L Allen-Hoffmann, and P F Lambert. 1995. “Integration of Human Papillomavirus Type 16 into the Human Genome Correlates with a Selective Growth Advantage of Cells.” *Journal of Virology* 69 (5): 2989–97.

Jeon, S., and P. F. Lambert. 1995. “Integration of Human Papillomavirus Type 16 DNA into the Human Genome Leads to Increased Stability of E6 and E7 MRNAs: Implications for Cervical Carcinogenesis.” *Proceedings of the National Academy of Sciences of the United States of America* 92 (5): 1654–58.

Jin, Bilian, Yajun Li, and Keith D. Robertson. 2011. “DNA Methylation.” *Genes & Cancer* 2 (6): 607–17. https://doi.org/10.1177/1947601910393957.

Jin, Bilian, and Keith D. Robertson. 2013. “DNA Methyltransferases, DNA Damage Repair, and Cancer.” *Advances in Experimental Medicine and Biology* 754: 3–29. https://doi.org/10.1007/978-1-4419-9967-2_1.

John, Keziah, Jennifer Wu, Bing-Wei Lee, and Camile S. Farah. 2013. “MicroRNAs in Head and Neck Cancer.” *International Journal of Dentistry* 2013. https://doi.org/10.1155/2013/650218.

Jung, Alain C., Jenny Briolat, Régine Millon, Aurélien de Reyniès, David Rickman, Emilie Thomas, Joseph Abecassis, Christine Clavel, and Bohdan Wasylyk. 2010. “Biological and Clinical Relevance of Transcriptionally Active Human Papillomavirus (HPV) Infection in Oropharynx Squamous Cell Carcinoma: HPV Associated with HNSCC.” *International Journal of Cancer* 126 (8): 1882–94. https://doi.org/10.1002/ijc.24911.

Kajitani, Naoko, Ayano Satsuka, Akifumi Kawate, and Hiroyuki Sakai. 2012. “Productive Lifecycle of Human Papillomaviruses That Depends Upon Squamous Epithelial Differentiation.” *Frontiers in Microbiology* 3 (April). https://doi.org/10.3389/fmicb.2012.00152.

Keck, Michaela K., Zhixiang Zuo, Arun Khattri, Thomas P. Stricker, Christopher D. Brown, Matin Imanguli, Damian Rieke, et al. 2015. “Integrative Analysis of Head and Neck Cancer Identifies Two Biologically Distinct HPV and Three Non-HPV Subtypes.” *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 21 (4): 870–81.

https://doi.org/10.1158/1078-0432.CCR-14-2481.

Kelly, Gregory M., and Mohamed I. Gatie. 2017. “Mechanisms Regulating Stemness and Differentiation in Embryonal Carcinoma Cells.” *Stem Cells International* 2017. https://doi.org/10.1155/2017/3684178.

Kim, Mirang, and Joseph Costello. 2017. “DNA Methylation: An Epigenetic Mark of Cellular Memory.” *Experimental & Molecular Medicine* 49 (4): e322. https://doi.org/10.1038/emm.2017.10.

Kleter, B., L. J. van Doorn, J. ter Schegget, L. Schrauwen, K. van Krimpen, M. Burger, B. ter Harmsel, and W. Quint. 1998. “Novel Short-Fragment PCR Assay for Highly Sensitive Broad-Spectrum Detection of Anogenital Human Papillomaviruses.” *The American Journal of Pathology* 153 (6): 1731–39. https://doi.org/10.1016/S0002-9440(10)65688-X.

Koutsky, L. 1997. “Epidemiology of Genital Human Papillomavirus Infection.” *The American Journal of Medicine* 102 (5A): 3–8.

Kovacic, M. B., P. E. Castle, R. Herrero, M. Schiffman, M. E. Sherman, S. Wacholder, A. C. Rodriguez, et al. 2006. “Relationships of Human Papillomavirus Type, Qualitative Viral Load, and Age with Cytologic Abnormality.” *Cancer Research* 66 (20): 10112–19. https://doi.org/10.1158/0008-5472.CAN-06-1812.

Lacey, Charles J. N., Catherine M. Lowndes, and Keerti V. Shah. 2006. “Chapter 4: Burden and Management of Non-Cancerous HPV-Related Conditions: HPV-6/11 Disease.” *Vaccine* 24 Suppl 3 (August): S3/35-41. https://doi.org/10.1016/j.vaccine.2006.06.015.

Lajer, C B, E Garnæs, L Friis-Hansen, B Norrild, M H Therkildsen, M Glud, M Rossing, et al. 2012. “The Role of MiRNAs in Human Papilloma Virus (HPV)-Associated Cancers: Bridging between HPV-Related Head and Neck Cancer and Cervical Cancer.” *British Journal of Cancer* 106 (9): 1526–34. https://doi.org/10.1038/bjc.2012.109.

Lajer, C. B., F. C. Nielsen, L. Friis-Hansen, B. Norrild, R. Borup, E. Garnæs, M. Rossing, et al. 2011. “Different MiRNA Signatures of Oral and Pharyngeal Squamous Cell Carcinomas: A Prospective Translational Study.” *British Journal of Cancer* 104 (5): 830–40. https://doi.org/10.1038/bjc.2011.29.

Lan, Huiyin, Haiqi Lu, Xian Wang, and Hongchuan Jin. 2015. “MicroRNAs as Potential Biomarkers in Cancer: Opportunities and Challenges.” Research article. BioMed Research International. 2015. https://doi.org/10.1155/2015/125094.

Larsen, C. Grønhøj, M. Gyldenløve, D. H. Jensen, M. H. Therkildsen, K. Kiss, B. Norrild, L. Konge, and C. von Buchwald. 2014. “Correlation between Human Papillomavirus and P16 Overexpression in Oropharyngeal Tumours: A Systematic Review.” *British Journal of Cancer* 110 (6): 1587–94. https://doi.org/10.1038/bjc.2014.42.

Lassen, Pernille, Hanne Primdahl, Jørgen Johansen, Claus A. Kristensen, Elo Andersen, Lisbeth J. Andersen, Jan F. Evensen, Jesper G. Eriksen, and Jens Overgaard. 2014. "Impact of HPV-Associated P16-Expression on Radiotherapy Outcome in Advanced Oropharynx and Non-Oropharynx Cancer." *Radiotherapy and Oncology* 113 (3): 310–16. https://doi.org/10.1016/j.radonc.2014.11.032.

Lechner, Matt, Ankur R. Chakravarthy, Vonn Walter, Liam Masterson, Andrew Feber, Amrita Jay, Paul M. Weinberger, et al. 2018. "Frequent HPV-Independent P16/INK4A Overexpression in Head and Neck Cancer." *Oral Oncology* 83 (August): 32–37. https://doi.org/10.1016/j.oraloncology.2018.06.006.

Leemans, C. René, Peter J. F. Snijders, and Ruud H. Brakenhoff. 2018. "The Molecular Landscape of Head and Neck Cancer." *Nature Reviews. Cancer* 18 (5): 269–82. https://doi.org/10.1038/nrc.2018.11.

Lehr, Elizabeth, Daniel Hohl, Marcel Huber, and Darron Brown. 2004. "Infection with Human Papillomavirus Alters Expression of the Small Proline Rich Proteins 2 and 3." *Journal of Medical Virology* 72 (3): 478–83. https://doi.org/10.1002/jmv.20011.

Lesnikova, Iana, Marianne Lidang, Stephen Hamilton-Dutoit, and Jørn Koch. 2009. "P16 as a Diagnostic Marker of Cervical Neoplasia: A Tissue Microarray Study of 796 Archival Specimens." *Diagnostic Pathology* 4 (July): 22. https://doi.org/10.1186/1746-1596-4-22.

Lesnikova, Iana, Marianne Lidang, Steven Hamilton-Dutoit, and Jørn Koch. 2010. "Rapid, Sensitive, Type Specific PCR Detection of the E7 Region of Human Papillomavirus Type 16 and 18 from Paraffin Embedded Sections of Cervical Carcinoma." *Infectious Agents and Cancer* 5 (January): 2. https://doi.org/10.1186/1750-9378-5-2.

Lewis, James S., Wade L. Thorstad, Rebecca D. Chernock, Bruce H. Haughey, James H. Yip, Qin Zhang, and Samir K. El-Mofty. 2010. "P16 Positive Oropharyngeal Squamous Cell Carcinoma: An Entity With a Favorable Prognosis Regardless of Tumor HPV Status." *The American Journal of Surgical Pathology* 34 (8). https://doi.org/10.1097/PAS.0b013e3181e84652.

Li, Hua, and Jennifer R Grandis, eds. 2015. *Chapter 8. Mutational Profile of HPV-Positive HNSCC*. Springer International Publishing. //www.springer.com/la/book/9783319210995.

Li, Ryan, Nishant Agrawal, and Carole Fakhry. 2015. "Anatomical Sites and Subsites of Head and Neck Cancer." In *HPV and Head and Neck Cancers*, 1–11. Head and Neck Cancer Clinics. Springer, New Delhi. https://doi.org/10.1007/978-81-322-2413-6_1.

Li, Yuanyuan, and Trygve O. Tollefsbol. 2011. "DNA Methylation Detection: Bisulfite Genomic Sequencing Analysis." *Methods in Molecular Biology (Clifton, N.J.)* 791: 11–21. https://doi.org/10.1007/978-1-61779-316-5_2.

Lin, Chunqing, Silvia Franceschi, and Gary M Clifford. 2018. "Human Papillomavirus Types from

Infection to Cancer in the Anus, According to Sex and HIV Status: A Systematic Review and Meta-Analysis." *The Lancet Infectious Diseases* 18 (2): 198–206. https://doi.org/10.1016/S1473-3099(17)30653-9.

Lister, Ryan, Mattia Pelizzola, Robert H. Dowen, R. David Hawkins, Gary Hon, Julian Tonti-Filippini, Joseph R. Nery, et al. 2009. "Human DNA Methylomes at Base Resolution Show Widespread Epigenomic Differences." *Nature* 462 (7271): 315–22. https://doi.org/10.1038/nature08514.

Liu, Kang, Guiqin Song, Xiaoyan Zhu, Xiaolin Yang, Yuewu Shen, Wan Wang, Guidong Shi, et al. 2017. "Association between ALDH2 Glu487Lys Polymorphism and the Risk of Esophageal Cancer." *Medicine* 96 (16). https://doi.org/10.1097/MD.0000000000006111.

Liu, Ying, Zheming Lu, Ruiping Xu, and Yang Ke. 2015. "Comprehensive Mapping of the Human Papillomavirus (HPV) DNA Integration Sites in Cervical Carcinomas by HPV Capture Technology." *Oncotarget* 7 (5): 5852–64. https://doi.org/10.18632/oncotarget.6809.

Liyanage, Vichithra R. B., Jessica S. Jarmasz, Nanditha Murugeshan, Marc R. Del Bigio, Mojgan Rastegar, and James R. Davie. 2014. "DNA Modifications: Function and Applications in Normal and Disease States." *Biology* 3 (4): 670–723. https://doi.org/10.3390/biology3040670.

Longworth, Michelle S., and Laimonis A. Laimins. 2004. "Pathogenesis of Human Papillomaviruses in Differentiating Epithelia." *Microbiology and Molecular Biology Reviews* 68 (2): 362–72. https://doi.org/10.1128/MMBR.68.2.362-372.2004.

Löning, Thomas, Hans Ikenberg, Jürgen Becker, Lutz Gissmann, Ilsetraut Hoepfer, and Harald zur Hausen. 1985. "Analysis of Oral Papillomas, Leukoplakias, and Invasive Carcinomas for Human Papillomavirus Type Related DNA." *Journal of Investigative Dermatology* 84 (5): 417–20. https://doi.org/10.1111/1523-1747.ep12265517.

Lu, Jun, Jin-Huan Wei, Zi-Hao Feng, Zhen-Hua Chen, Yong-Qian Wang, Yong Huang, Yong Fang, et al. 2017. "MiR-106b-5p Promotes Renal Cell Carcinoma Aggressiveness and Stem-Cell-like Phenotype by Activating Wnt/β-Catenin Signalling." *Oncotarget* 8 (13): 21461–71. https://doi.org/10.18632/oncotarget.15591.

Lubov, Joshua, Mariana Maschietto, Iman Ibrahim, Alex Mlynarek, Michael Hier, Luiz Paulo Kowalski, Moulay A. Alaoui-Jamali, et al. 2017. "Meta-Analysis of MicroRNAs Expression in Head and Neck Cancer: Uncovering Association with Outcome and Mechanisms." *Oncotarget* 8 (33): 55511–24. https://doi.org/10.18632/oncotarget.19224.

Luo, Xiangjian, Can Cheng, Zheqiong Tan, Namei Li, Min Tang, Lifang Yang, and Ya Cao. 2017. "Emerging Roles of Lipid Metabolism in Cancer Metastasis." *Molecular Cancer* 16 (April). https://doi.org/10.1186/s12943-017-0646-3.

Lydiatt, William M., Snehal G. Patel, Brian O'Sullivan, Margaret S. Brandwein, John A. Ridge, Jocelyn C. Migliacci, Ashley M. Loomis, and Jatin P. Shah. 2017. "Head and Neck Cancers-Major

Changes in the American Joint Committee on Cancer Eighth Edition Cancer Staging Manual." *CA: A Cancer Journal for Clinicians* 67 (2): 122–37. https://doi.org/10.3322/caac.21389.

Macfarlane, Leigh-Ann, and Paul R. Murphy. 2010. "MicroRNA: Biogenesis, Function and Role in Cancer." *Current Genomics* 11 (7): 537–61. https://doi.org/10.2174/138920210793175895.

Maley, Carlo C., Athena Aktipis, Trevor A. Graham, Andrea Sottoriva, Amy M. Boddy, Michalina Janiszewska, Ariosto S. Silva, et al. 2017. "Classifying the Evolutionary and Ecological Features of Neoplasms." *Nature Reviews. Cancer* 17 (10): 605–19. https://doi.org/10.1038/nrc.2017.69.

Marcu, Loredana G. 2016. "Future Treatment Directions for HPV-Associated Head and Neck Cancer Based on Radiobiological Rationale and Current Clinical Evidence." *Critical Reviews in Oncology/Hematology* 103 (July): 27–36. https://doi.org/10.1016/j.critrevonc.2016.05.002.

Mariño-Ramírez, Leonardo, Maricel G Kann, Benjamin A Shoemaker, and David Landsman. 2005. "Histone Structure and Nucleosome Stability." *Expert Review of Proteomics* 2 (5): 719–29. https://doi.org/10.1586/14789450.2.5.719.

Marques, Ana Elizia Mascarenhas, Silvia Taveira Elias, André Luís Porporatti, Rogerio Moraes Castilho, Cristiane Helena Squarize, Graziela De Luca Canto, and Eliete Neves Silva Guerra. 2016. "MTOR Pathway Protein Immunoexpression as a Prognostic Factor for Survival in Head and Neck Cancer Patients: A Systematic Review and Meta-Analysis." *Journal of Oral Pathology & Medicine: Official Publication of the International Association of Oral Pathologists and the American Academy of Oral Pathology* 45 (5): 319–28. https://doi.org/10.1111/jop.12390.

Marron, Manuela, Paolo Boffetta, Henrik Møller, Wolfgang Ahrens, Hermann Pohlabeln, Simone Benhamou, Christine Bouchardy, et al. 2012. "Risk of Upper Aerodigestive Tract Cancer and Type of Alcoholic Beverage: A European Multicenter Case-Control Study." *European Journal of Epidemiology* 27 (7): 499–517. https://doi.org/10.1007/s10654-012-9699-1.

Martel, Catherine de, Jacques Ferlay, Silvia Franceschi, Jérôme Vignat, Freddie Bray, David Forman, and Martyn Plummer. 2012. "Global Burden of Cancers Attributable to Infections in 2008: A Review and Synthetic Analysis." *The Lancet. Oncology* 13 (6): 607–15. https://doi.org/10.1016/S1470-2045(12)70137-7.

Masuda, Muneyuki, Takahiro Wakasaki, and Satoshi Toh. 2016. "Stress-Triggered Atavistic Reprogramming (STAR) Addiction: Driving Force behind Head and Neck Cancer?" *American Journal of Cancer Research* 6 (6): 1149–66.

Matovina, Mihaela, Ivan Sabol, Goran Grubisić, Nina Milutin Gasperov, and Magdalena Grce. 2009. "Identification of Human Papillomavirus Type 16 Integration Sites in High-Grade Precancerous Cervical Lesions." *Gynecologic Oncology* 113 (1): 120–27. https://doi.org/10.1016/j.ygyno.2008.12.004.

McBride, Alison A. 2013. "The Papillomavirus E2 Proteins." *Virology* 445 (0): 57–79.

https://doi.org/10.1016/j.virol.2013.06.006.

McFarlane, Melanie, Alasdair I. MacDonald, Andrew Stevenson, and Sheila V. Graham. 2015. “Human Papillomavirus 16 Oncoprotein Expression Is Controlled by the Cellular Splicing Factor SRSF2 (SC35).” *Journal of Virology* 89 (10): 5276–87. https://doi.org/10.1128/JVI.03434-14.

Mehanna, Hisham, Tom Beech, Tom Nicholson, Iman El-Hariry, Christopher McConkey, Vinidh Paleri, and Sally Roberts. 2013. “Prevalence of Human Papillomavirus in Oropharyngeal and Nonoropharyngeal Head and Neck Cancer--Systematic Review and Meta-Analysis of Trends by Time and Region.” *Head & Neck* 35 (5): 747–55. https://doi.org/10.1002/hed.22015.

Meisels, A., and R. Fortin. 1976. “Condylomatous Lesions of the Cervix and Vagina. I. Cytologic Patterns.” *Acta Cytologica* 20 (6): 505–9.

Mikeska, Thomas, Jörg Felsberg, Chelsee A. Hewitt, and Alexander Dobrovic. 2011. “Analysing DNA Methylation Using Bisulphite Pyrosequencing.” In *Epigenetics Protocols*, 33–53. Methods in Molecular Biology. Humana Press. https://doi.org/10.1007/978-1-61779-316-5_4.

Miller, Daniel L., J. Wade Davis, Kristen H. Taylor, Jeff Johnson, Zonggao Shi, Russell Williams, Ulus Atasoy, James S. Lewis, and M. Sharon Stack. 2015. “Identification of a Human Papillomavirus–Associated Oncogenic MiRNA Panel in Human Oropharyngeal Squamous Cell Carcinoma Validated by Bioinformatics Analysis of The Cancer Genome Atlas.” *The American Journal of Pathology* 185 (3): 679–92. https://doi.org/10.1016/j.ajpath.2014.11.018.

Milutin Gašperov, Nina, Sanja A. Farkas, Torbjörn K. Nilsson, and Magdalena Grce. 2014. “Epigenetic Activation of Immune Genes in Cervical Cancer.” *Immunology Letters* 162 (2 Pt B): 256–57. https://doi.org/10.1016/j.imlet.2014.09.019.

Milutin-Gašperov, N, I Sabol, G Halec, M Matovina, and M Grce. 2007. “Retrospective Study of the Prevalence of High-Risk Human Papillomaviruses among Croatian Women.” *Collegium Antropologicum* 31: 89–96.

Mirghani, Haitham, and Pierre Blanchard. 2018. “Treatment De-Escalation for HPV-Driven Oropharyngeal Cancer: Where Do We Stand?” *Clinical and Translational Radiation Oncology* 8 (January): 4–11. https://doi.org/10.1016/j.ctro.2017.10.005.

Monsonego, Joseph, Michael G. Hudgens, Laurent Zerat, Jean-Claude Zerat, Kari Syrjänen, Philippe Halfon, Fabrice Ruiz, and Jennifer S. Smith. 2011. “Evaluation of Oncogenic Human Papillomavirus RNA and DNA Tests with Liquid-Based Cytology in Primary Cervical Cancer Screening: The FASE Study.” *International Journal of Cancer* 129 (3): 691–701. https://doi.org/10.1002/ijc.25726.

Moody, Cary A. 2017. “Mechanisms by Which HPV Induces a Replication Competent Environment in Differentiating Keratinocytes.” *Viruses* 9 (9). https://doi.org/10.3390/v9090261.

Morgan, Iain M., Laurence J. DiNardo, and Brad Windle. 2017. “Integration of Human

Papillomavirus Genomes in Head and Neck Cancer: Is It Time to Consider a Paradigm Shift?" *Viruses* 9 (8). https://doi.org/10.3390/v9080208.

Moscicki, Anna-Barbara, Mark Schiffman, Susanne Kjaer, and Luisa L. Villa. 2006. "Chapter 5: Updating the Natural History of HPV and Anogenital Cancer." *Vaccine* 24 (August): S42–51. https://doi.org/10.1016/j.vaccine.2006.06.018.

Müller, Tim, Martin Braun, Dimo Dietrich, Seher Aktekin, Simon Höft, Glen Kristiansen, Friederike Göke, et al. 2017. "PD-L1: A Novel Prognostic Biomarker in Head and Neck Squamous Cell Carcinoma." *Oncotarget* 8 (32): 52889–900. https://doi.org/10.18632/oncotarget.17547.

Münger, Karl, Amy Baldwin, Kirsten M. Edwards, Hiroyuki Hayakawa, Christine L. Nguyen, Michael Owens, Miranda Grace, and KyungWon Huh. 2004. "Mechanisms of Human Papillomavirus-Induced Oncogenesis." *Journal of Virology* 78 (21): 11451–60. https://doi.org/10.1128/JVI.78.21.11451-11460.2004.

Munoz, N., F. X Bosch, S. de Sanjose, R. Herrero, X. Castellsague, K. V Shah, P. J.F Snijders, C. J. Meijer, and others. 2003. "Epidemiologic Classification of Human Papillomavirus Types Associated with Cervical Cancer." *New England Journal of Medicine* 348 (6): 518–528.

Nathan, Dafna, David E. Sterner, and Shelley L. Berger. 2003. "Histone Modifications: Now Summoning Sumoylation." *Proceedings of the National Academy of Sciences* 100 (23): 13118–20. https://doi.org/10.1073/pnas.2436173100.

NCI. 2018. "NCI Dictionary of Cancer Terms." NciAppModulePage. National Cancer Institute. 2018. https://www.cancer.gov/publications/dictionaries/cancer-terms.

Ndiaye, Cathy, Laia Alemany, Yankhoba Diop, Nafissatou Ndiaye, Marie-Joseph Diémé, Sara Tous, Jo Ellen Klaustermeier, et al. 2013. "The Role of Human Papillomavirus in Head and Neck Cancer in Senegal." *Infectious Agents and Cancer* 8 (1): 14. https://doi.org/10.1186/1750-9378-8-14.

Negi, Vinny, Deepanjan Paul, Sudipta Das, Prashant Bajpai, Suchita Singh, Arijit Mukhopadhyay, Anurag Agrawal, and Balaram Ghosh. 2015. "Altered Expression and Editing of MiRNA-100 Regulates ITreg Differentiation." *Nucleic Acids Research* 43 (16): 8057–65. https://doi.org/10.1093/nar/gkv752.

NIH. 2018. "ICD-O-3 Site Codes | SEER Training. Https://Training.Seer.Cancer.Gov/Head-Neck/Abstract-Code-Stage/Codes.Html." 2018. https://training.seer.cancer.gov/head-neck/abstract-code-stage/codes.html.

NORGEN. 2018. "Plant/Fungi Total RNA Purification Kit (Cat. 25800, 31350, 25850) | Norgen Biotek Corp." 2018. https://norgenbiotek.com/product/plantfungi-total-rna-purification-kit.

Nygård, Mari, Bjarte Aagnes, Freddie Bray, Bjørn Møller, and Jon Mork. 2012. "Population-Based Evidence of Increased Survival in Human Papillomavirus-Related Head and Neck Cancer." *European Journal of Cancer (Oxford, England: 1990)* 48 (9): 1341–46.

https://doi.org/10.1016/j.ejca.2012.03.014.

Olmedo-Nieva, Leslie, J. Omar Muñoz-Bello, Adriana Contreras-Paredes, and Marcela Lizano. 2018. “The Role of E6 Spliced Isoforms (E6*) in Human Papillomavirus-Induced Carcinogenesis.” *Viruses* 10 (1). https://doi.org/10.3390/v10010045.

O’Rorke, M. A., M. V. Ellison, L. J. Murray, M. Moran, J. James, and L. A. Anderson. 2012. “Human Papillomavirus Related Head and Neck Cancer Survival: A Systematic Review and Meta-Analysis.” *Oral Oncology* 48 (12): 1191–1201. https://doi.org/10.1016/j.oraloncology.2012.06.019.

Osazuwa-Peters, Nosayaba, Matthew C. Simpson, Sean T. Massa, Eric Adjei Boakye, Jastin L. Antisdel, and Mark A. Varvares. 2017. “40-Year Incidence Trends for Oropharyngeal Squamous Cell Carcinoma in the United States.” *Oral Oncology* 74 (November): 90–97. https://doi.org/10.1016/j.oraloncology.2017.09.015.

Oyervides-Muñoz, Mariel Araceli, Antonio Alí Pérez-Maya, Hazyadee Frecia Rodríguez-Gutiérrez, Gabriela Sofía Gómez-Macias, Oscar Raúl Fajardo-Ramírez, Víctor Treviño, Hugo Alberto Barrera-Saldaña, and María Lourdes Garza-Rodríguez. 2018. “Understanding the HPV Integration and Its Progression to Cervical Cancer.” *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases* 61 (July): 134–44. https://doi.org/10.1016/j.meegid.2018.03.003.

Ozkul, Yusuf, and Umberto Galderisi. 2016. “The Impact of Epigenetics on Mesenchymal Stem Cell Biology.” *Journal of Cellular Physiology* 231 (11): 2393–2401. https://doi.org/10.1002/jcp.25371.

Padjen, Ivan, Marina Dabić, Tatjana Glivetić, Zrinka Biloglav, Dolores Biočina-Lukenda, and Josip Lukenda. 2012. “The Analysis of Tobacco Consumption in Croatia--Are We Successfully Facing the Epidemic?” *Central European Journal of Public Health* 20 (1): 5–10.

Parfenov, Michael, Chandra Sekhar Pedamallu, Nils Gehlenborg, Samuel S. Freeman, Ludmila Danilova, Christopher A. Bristow, Semin Lee, et al. 2014. “Characterization of HPV and Host Genome Interactions in Primary Head and Neck Cancers.” *Proceedings of the National Academy of Sciences of the United States of America* 111 (43): 15544–49. https://doi.org/10.1073/pnas.1416074111.

Pattillo, R. A., R. O. Hussa, M. T. Story, A. C. Ruckert, M. R. Shalaby, and R. F. Mattingly. 1977. “Tumor Antigen and Human Chorionic Gonadotropin in CaSki Cells: A New Epidermoid Cervical Cancer Cell Line.” *Science (New York, N.Y.)* 196 (4297): 1456–58.

Peng, Yong, and Carlo M. Croce. 2016. “The Role of MicroRNAs in Human Cancer.” *Signal Transduction and Targeted Therapy* 1 (January): 15004. https://doi.org/10.1038/sigtrans.2015.4.

Prigge, Elena-Sophie, Marc Arbyn, Magnus von Knebel Doeberitz, and Miriam Reuschenbach. 2016. “Diagnostic Accuracy of P16INK4a Immunohistochemistry in Oropharyngeal Squamous Cell Carcinomas: A Systematic Review and Meta-Analysis.” *International Journal of Cancer* 140 (5):

1186–98. https://doi.org/10.1002/ijc.30516.

Ragin, C. C. R., E. Taioli, J. L. Weissfeld, J. S. White, K. M. Rossie, F. Modugno, and S. M. Gollin. 2006. “11q13 Amplification Status and Human Papillomavirus in Relation to P16 Expression Defines Two Distinct Etiologies of Head and Neck Tumours.” *British Journal of Cancer* 95 (10): 1432–38. https://doi.org/10.1038/sj.bjc.6603394.

Raisch, Jennifer, Arlette Darfeuille-Michaud, and Hang Thi Thu Nguyen. 2013. “Role of MicroRNAs in the Immune System, Inflammation and Cancer.” *World Journal of Gastroenterology : WJG* 19 (20): 2985–96. https://doi.org/10.3748/wjg.v19.i20.2985.

Ramqvist, Torbjörn, Michael Mints, Nikolaos Tertipis, Anders Näsman, Mircea Romanitan, and Tina Dalianis. 2015. “Studies on Human Papillomavirus (HPV) 16 E2, E5 and E7 MRNA in HPV-Positive Tonsillar and Base of Tongue Cancer in Relation to Clinical Outcome and Immunological Parameters.” *Oral Oncology* 51 (12): 1126–31. https://doi.org/10.1016/j.oraloncology.2015.09.007.

Ramsahoye, Bernard H., Detlev Biniszkiewicz, Frank Lyko, Victoria Clark, Adrian P. Bird, and Rudolf Jaenisch. 2000. “Non-CpG Methylation Is Prevalent in Embryonic Stem Cells and May Be Mediated by DNA Methyltransferase 3a.” *Proceedings of the National Academy of Sciences* 97 (10): 5237–42. https://doi.org/10.1073/pnas.97.10.5237.

Reynolds, Lindsay M., Jackson R. Taylor, Jingzhong Ding, Kurt Lohman, Craig Johnson, David Siscovick, Gregory Burke, et al. 2014. “Age-Related Variations in the Methylome Associated with Gene Expression in Human Monocytes and T Cells.” *Nature Communications* 5 (November): 5366. https://doi.org/10.1038/ncomms6366.

Riemer, Angelika B., Derin B. Keskin, Guanglan Zhang, Maris Handley, Karen S. Anderson, Vladimir Brusic, Bruce Reinhold, and Ellis L. Reinherz. 2010. “A Conserved E7-Derived Cytotoxic T Lymphocyte Epitope Expressed on Human Papillomavirus 16-Transformed HLA-A2+ Epithelial Cancers.” *The Journal of Biological Chemistry* 285 (38): 29608–22. https://doi.org/10.1074/jbc.M110.126722.

Roberts, Lisa, Jessica Bowers, Kelly Sensinger, Andrew Lisowski, Robert Getts, and Mark G. Anderson. 2009. “Identification of Methods for Use of Formalin-Fixed, Paraffin-Embedded Tissue Samples in RNA Expression Profiling.” *Genomics* 94 (5): 341–48. https://doi.org/10.1016/j.ygeno.2009.07.007.

Robinson, Max, Andrew Schache, Philip Sloan, and Selvam Thavaraj. 2012. “HPV Specific Testing: A Requirement for Oropharyngeal Squamous Cell Carcinoma Patients.” *Head and Neck Pathology* 6 (Suppl 1): 83–90. https://doi.org/10.1007/s12105-012-0370-7.

Rosenberger, Simone, Johanna De-Castro Arce, Lutz Langbein, Renske D. M. Steenbergen, and Frank Rösl. 2010. “Alternative Splicing of Human Papillomavirus Type-16 E6/E6* Early MRNA Is Coupled to EGF Signaling via Erk1/2 Activation.” *Proceedings of the National Academy of Sciences*

of the United States of America 107 (15): 7006–11. https://doi.org/10.1073/pnas.1002620107.

Rossetto, D., A. W. Truman, S. J. Kron, and J. Cote. 2010. “Epigenetic Modifications in Double-Strand Break DNA Damage Signaling and Repair.” *Clinical Cancer Research* 16 (18): 4543–52. https://doi.org/10.1158/1078-0432.CCR-10-0513.

Rossetto, Dorine, Nikita Avvakumov, and Jacques Côté. 2012. “Histone Phosphorylation.” *Epigenetics* 7 (10): 1098–1108. https://doi.org/10.4161/epi.21975.

Rous, Peyton, and J. W. Beard. 1935. “Carcinomatous Changes in Virus-Induced Papillomas of the Skin of the Rabbit.” *Proceedings of the Society for Experimental Biology and Medicine* 32 (4): 578–80. https://doi.org/10.3181/00379727-32-7770P.

Sabol, Ivan, Nina Milutin Gašperov, Mihaela Matovina, Ksenija Božinović, Goran Grubišić, Ivan Fistonić, Dragan Belci, et al. 2017. “Cervical HPV Type-Specific Pre-Vaccination Prevalence and Age Distribution in Croatia.” *PloS One* 12 (7): e0180480. https://doi.org/10.1371/journal.pone.0180480.

Sanjose, S de, WGV Quint, L Alemany, DT Geraets, JE Klaustermeier, B Lloveras, S Tous, et al. 2010. “Human Papillomavirus Genotype Attribution in Invasive Cervical Cancer: A Retrospective Cross-Sectional Worldwide Study.” *The Lancet Oncology* 11 (11): 1048–56. https://doi.org/10.1016/S1470-2045(10)70230-8.

Saslow, Debbie, Diane Solomon, Herschel W. Lawson, Maureen Killackey, Shalini Kulasingam, Joanna Cain, Francisco A. R. Garcia, et al. 2012. “American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology Screening Guidelines for the Prevention and Early Detection of Cervical Cancer.” *Journal of Lower Genital Tract Disease* 16 (3): 175–204. https://doi.org/10.1097/LGT.0b013e31824ca9d5.

Sass, Steffen, Adriana Pitea, Kristian Unger, Julia Hess, Nikola S. Mueller, and Fabian J. Theis. 2015. “MicroRNA-Target Network Inference and Local Network Enrichment Analysis Identify Two MicroRNA Clusters with Distinct Functions in Head and Neck Squamous Cell Carcinoma.” *International Journal of Molecular Sciences* 16 (12): 30204–22. https://doi.org/10.3390/ijms161226230.

Schepeler, Troels, Jørgen T. Reinert, Marie S. Ostenfeld, Lise L. Christensen, Asli N. Silahtaroglu, Lars Dyrskjøt, Carsten Wiuf, et al. 2008. “Diagnostic and Prognostic MicroRNAs in Stage II Colon Cancer.” *Cancer Research* 68 (15): 6416–24. https://doi.org/10.1158/0008-5472.CAN-07-6110.

Schiffman, Mark, John Doorbar, Nicolas Wentzensen, Silvia de Sanjosé, Carole Fakhry, Bradley J. Monk, Margaret A. Stanley, and Silvia Franceschi. 2016. “Carcinogenic Human Papillomavirus Infection.” *Nature Reviews Disease Primers* 2 (December): 16086. https://doi.org/10.1038/nrdp.2016.86.

Schiffman, Mark, and Nicolas Wentzensen. 2013. “Human Papillomavirus (HPV) Infection and the

Multi-Stage Carcinogenesis of Cervical Cancer." *Cancer Epidemiology, Biomarkers & Prevention : A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology* 22 (4): 553–60. https://doi.org/10.1158/1055-9965.EPI-12-1406.

Seiwert, T. 2013. "Accurate HPV Testing: A Requirement for Precision Medicine for Head and Neck Cancer." *Annals of Oncology* 24 (11): 2711–13. https://doi.org/10.1093/annonc/mdt417.

Serrano, Beatriz, Silvia de Sanjosé, Sara Tous, Beatriz Quiros, Nubia Muñoz, Xavier Bosch, and Laia Alemany. 2015. "Human Papillomavirus Genotype Attribution for HPVs 6, 11, 16, 18, 31, 33, 45, 52 and 58 in Female Anogenital Lesions." *European Journal of Cancer (Oxford, England: 1990)* 51 (13): 1732–41. https://doi.org/10.1016/j.ejca.2015.06.001.

Sethi, Neeraj, Alexander Wright, Henry Wood, and Pamela Rabbitts. 2014. "MicroRNAs and Head and Neck Cancer: Reviewing the First Decade of Research." *European Journal of Cancer (Oxford, England: 1990)* 50 (15): 2619–35. https://doi.org/10.1016/j.ejca.2014.07.012.

Sharma, Shikhar, Theresa K. Kelly, and Peter A. Jones. 2010. "Epigenetics in Cancer." *Carcinogenesis* 31 (1): 27–36. https://doi.org/10.1093/carcin/bgp220.

Sheedy, Frederick J. 2015. "Turning 21: Induction of MiR-21 as a Key Switch in the Inflammatory Response." *Frontiers in Immunology* 6: 19. https://doi.org/10.3389/fimmu.2015.00019.

Shen, Jia, and Charles Spruck. 2017. "F-Box Proteins in Epigenetic Regulation of Cancer." *Oncotarget* 8 (66): 110650–55. https://doi.org/10.18632/oncotarget.22469.

Shiio, Yuzuru, and Robert N. Eisenman. 2003. "Histone Sumoylation Is Associated with Transcriptional Repression." *Proceedings of the National Academy of Sciences* 100 (23): 13225–30. https://doi.org/10.1073/pnas.1735528100.

Shin, Chanseok, and James L. Manley. 2004. "Cell Signalling and the Control of Pre-MRNA Splicing." *Nature Reviews. Molecular Cell Biology* 5 (9): 727–38. https://doi.org/10.1038/nrm1467.

Singhi, Aatur D., Raja R. Seethala, Katie Nason, Tyler J. Foxwell, Robyn L. Roche, Kevin M. McGrath, Ryan M. Levy, James D. Luketich, and Jon M. Davison. 2015. "Undifferentiated Carcinoma of the Esophagus: A Clinicopathological Study of 16 Cases." *Human Pathology* 46 (3): 366–75. https://doi.org/10.1016/j.humpath.2014.11.021.

Slaby, O., M. Svoboda, P. Fabian, T. Smerdova, D. Knoflickova, M. Bednarikova, R. Nenutil, and R. Vyzula. 2007. "Altered Expression of MiR-21, MiR-31, MiR-143 and MiR-145 Is Related to Clinicopathologic Features of Colorectal Cancer." *Oncology* 72 (5–6): 397–402. https://doi.org/10.1159/000113489.

Smeets, Serge J, Albertus T Hesselink, Ernst-Jan M Speel, Annick Haesevoets, Peter J F Snijders, Michael Pawlita, Chris J L M Meijer, Boudewijn J M Braakhuis, C René Leemans, and Ruud H Brakenhoff. 2007. "A Novel Algorithm for Reliable Detection of Human Papillomavirus in Paraffin

Embedded Head and Neck Cancer Specimen." *International Journal of Cancer. Journal International Du Cancer* 121 (11): 2465–72. https://doi.org/10.1002/ijc.22980.

Snijders, P. J., A. J. van den Brule, H. F. Schrijnemakers, G. Snow, C. J. Meijer, and J. M. Walboomers. 1990. "The Use of General Primers in the Polymerase Chain Reaction Permits the Detection of a Broad Spectrum of Human Papillomavirus Genotypes." *The Journal of General Virology* 71 (Pt 1) (January): 173–81. https://doi.org/10.1099/0022-1317-71-1-173.

Song, Ye, Jiangchao Li, Yinghui Zhu, Yongdong Dai, Tingting Zeng, Lulu Liu, Jianbiao Li, et al. 2014. "MicroRNA-9 Promotes Tumor Metastasis via Repressing E-Cadherin in Esophageal Squamous Cell Carcinoma." *Oncotarget* 5 (22): 11669–80.

Sousa, Lucas O., Lays M. Sobral, Camila S. Matsumoto, Fabiano P. Saggioro, Rossana V.M. López, Rodrigo A. Panepucci, Carlos Curti, Wilson A. Silva, Lewis J. Greene, and Andréia M. Leopoldino. 2016. "Lymph Node or Perineural Invasion Is Associated with Low MiR-15a, MiR-34c and MiR-199b Levels in Head and Neck Squamous Cell Carcinoma." *BBA Clinical* 6 (November): 159–64. https://doi.org/10.1016/j.bbacli.2016.11.001.

Spence, Tara, Jeff Bruce, Kenneth W. Yip, and Fei-Fei Liu. 2016. "HPV Associated Head and Neck Cancer." *Cancers* 8 (8). https://doi.org/10.3390/cancers8080075.

Steiner, David F., Molly F. Thomas, Joyce K. Hu, Zhiyong Yang, Joshua E. Babiarz, Christopher D. C. Allen, Mehrdad Matloubian, Robert Blelloch, and K. Mark Ansel. 2011. "MicroRNA-29 Regulates T-Box Transcription Factors and Interferon-γ Production in Helper T Cells." *Immunity* 35 (2): 169–81. https://doi.org/10.1016/j.immuni.2011.07.009.

Stephen, Josena K., George Divine, Kang Mei Chen, Dhananajay Chitale, Shaleta Havard, and Maria J. Worsham. 2013. "Significance of P16 in Site-Specific HPV Positive and HPV Negative Head and Neck Squamous Cell Carcinoma." *Cancer and Clinical Oncology* 2 (1): 51–61. https://doi.org/10.5539/cco.v2n1p51.

Sudenga, S.L., and S. Shrestha. 2013. "Key Considerations and Current Perspectives of Epidemiological Studies on Human Papillomavirus Persistence, the Intermediate Phenotype to Cervical Cancer." *International Journal of Infectious Diseases: IJID : Official Publication of the International Society for Infectious Diseases* 17 (4): e216–20. https://doi.org/10.1016/j.ijid.2012.12.027.

Sun, Xianghui, and Lei Zhang. 2017. "MicroRNA-143 Suppresses Oral Squamous Cell Carcinoma Cell Growth, Invasion and Glucose Metabolism through Targeting Hexokinase 2." *Bioscience Reports* 37 (3). https://doi.org/10.1042/BSR20160404.

Suzuki, Hiromu, Reo Maruyama, Eiichiro Yamamoto, and Masahiro Kai. 2012. "DNA Methylation and MicroRNA Dysregulation in Cancer." *Molecular Oncology*, Cancer epigenetics, 6 (6): 567–78. https://doi.org/10.1016/j.molonc.2012.07.007.

Syrjänen, K, S Syrjänen, M Lamberg, S Pyrhönen, and J Nuutinen. 1983. “Morphological and Immunohistochemical Evidence Suggesting Human Papillomavirus (HPV) Involvement in Oral Squamous Cell Carcinogenesis.” *International Journal of Oral Surgery* 12 (6): 418–24.

Syrjänen, S. 2010. “The Role of Human Papillomavirus Infection in Head and Neck Cancers.” *Annals of Oncology* 21 (suppl 7): vii243–45. https://doi.org/10.1093/annonc/mdq454.

Taberna, M., M. Mena, M. A. Pavón, L. Alemany, M. L. Gillison, and R. Mesía. 2017. “Human Papillomavirus-Related Oropharyngeal Cancer.” *Annals of Oncology* 28 (10): 2386–98. https://doi.org/10.1093/annonc/mdx304.

Taylor, Sylvia, Eveline Bunge, Marina Bakker, and Xavier Castellsagué. 2016. “The Incidence, Clearance and Persistence of Non-Cervical Human Papillomavirus Infections: A Systematic Review of the Literature.” *BMC Infectious Diseases* 16: 293. https://doi.org/10.1186/s12879-016-1633-9.

TCGA. 2015. “Comprehensive Genomic Characterization of Head and Neck Squamous Cell Carcinomas.” *Nature* 517 (7536): 576–82. https://doi.org/10.1038/nature14129.

TCGA. 2017. “Integrated Genomic and Molecular Characterization of Cervical Cancer | Nature.” 2017. https://www.nature.com/articles/nature21386.

Telonis, Aristeidis G., Rogan Magee, Phillipe Loher, Inna Chervoneva, Eric Londin, and Isidore Rigoutsos. 2017. “Knowledge about the Presence or Absence of MiRNA Isoforms (IsomiRs) Can Successfully Discriminate amongst 32 TCGA Cancer Types.” *Nucleic Acids Research* 45 (6): 2973–85. https://doi.org/10.1093/nar/gkx082.

Thomas, J. T., W. G. Hubert, M. N. Ruesch, and L. A. Laimins. 1999. “Human Papillomavirus Type 31 Oncoproteins E6 and E7 Are Required for the Maintenance of Episomes during the Viral Life Cycle in Normal Human Keratinocytes.” *Proceedings of the National Academy of Sciences of the United States of America* 96 (15): 8449–54.

Tomaić, Vjekoslav. 2016. “Functional Roles of E6 and E7 Oncoproteins in HPV-Induced Malignancies at Diverse Anatomical Sites.” *Cancers* 8 (10). https://doi.org/10.3390/cancers8100095.

Tommasino, Massimo. 2014. “The Human Papillomavirus Family and Its Role in Carcinogenesis.” *Seminars in Cancer Biology* 26 (June): 13–21. https://doi.org/10.1016/j.semcancer.2013.11.002.

Tuna, Musaffe, and Christopher I. Amos. 2016. “Next Generation Sequencing and Its Applications in HPVassociated Cancers.” *Oncotarget* 8 (5): 8877–89. https://doi.org/10.18632/oncotarget.12830.

UCSC. 2018. “UCSC Genome Browser Home.” 2018. https://genome.ucsc.edu/.

Urashima, Mitsuyoshi, Takanori Hama, Toshihito Suda, Yutaka Suzuki, Masahiro Ikegami, Chikako Sakanashi, Taisuke Akutsu, et al. 2013. “Distinct Effects of Alcohol Consumption and Smoking on

Genetic Alterations in Head and Neck Carcinoma." *PLoS ONE* 8 (11). https://doi.org/10.1371/journal.pone.0080828.

Valinezhad Orang, Ayla, Reza Safaralizadeh, and Mina Kazemzadeh-Bavili. 2014. "Mechanisms of MiRNA-Mediated Gene Regulation from Common Downregulation to MRNA-Specific Upregulation." *International Journal of Genomics* 2014. https://doi.org/10.1155/2014/970607.

Valmary-Degano, Séverine, Elise Jacquin, Jean-Luc Prétet, Franck Monnien, Benoit Girardo, Francine Arbez-Gindre, Marie Joly, Jean-François Bosset, Bernadette Kantelip, and Christiane Mougin. 2013. "Signature Patterns of Human Papillomavirus Type 16 in Invasive Anal Carcinoma." *Human Pathology* 44 (6): 992–1002. https://doi.org/10.1016/j.humpath.2012.08.019.

Van Doorslaer, Koenraad, Zigui Chen, and Alison A. McBride. 2016. "Detection and Genotyping of Human Papillomaviruses from Archival Formalin Fixed Tissue Samples." *Current Protocols in Microbiology* 43 (November): 14B.9.1-14B.9.20. https://doi.org/10.1002/cpmc.16.

Vatca, M., J.T. Lucas, J. Laudadio, R.B. D'Agostino, J.D. Waltonen, C.A. Sullivan, R. Rouchard-Plasser, et al. 2014. "Retrospective Analysis of the Impact of HPV Status and Smoking on Mucositis in Patients with Oropharyngeal Squamous Cell Carcinoma Treated with Concurrent Chemotherapy and Radiotherapy." *Oral Oncology* 50 (9): 869–76. https://doi.org/10.1016/j.oraloncology.2014.06.010.

Venuti, Aldo, Francesca Paolini, Lubna Nasir, Annunziata Corteggio, Sante Roperto, Maria S. Campo, and Giuseppe Borzacchiello. 2011. "Papillomavirus E5: The Smallest Oncoprotein with Many Functions." *Molecular Cancer* 10 (November): 140. https://doi.org/10.1186/1476-4598-10-140.

Viarisio, Daniele, Lutz Gissmann, and Massimo Tommasino. 2017. "Human Papillomaviruses and Carcinogenesis: Well-Established and Novel Models." *Current Opinion in Virology* 26: 56–62. https://doi.org/10.1016/j.coviro.2017.07.014.

Vigneswaran, Nadarajah, and Michelle D. Williams. 2014. "Epidemiologic Trends in Head and Neck Cancer and Aids in Diagnosis." *Oral and Maxillofacial Surgery Clinics of North America* 26 (2): 123–41. https://doi.org/10.1016/j.coms.2014.01.001.

Villiers, E. M. de, H. Weidauer, H. Otto, and H. zur Hausen. 1985. "Papillomavirus DNA in Human Tongue Carcinomas." *International Journal of Cancer* 36 (5): 575–78.

Villiers, Ethel-Michele de, Claude Fauquet, Thomas R Broker, Hans-Ulrich Bernard, and Harald zur Hausen. 2004. "Classification of Papillomaviruses." *Virology* 324 (1): 17–27. https://doi.org/10.1016/j.virol.2004.03.033.

Vojtechova, Zuzana, Ivan Sabol, Martina Salakova, Jana Smahelova, Jiri Zavadil, Lubomir Turek, Marek Grega, Jan Klozar, Bohumir Prochazka, and Ruth Tachezy. 2016. "Comparison of the MiRNA Profiles in HPV-Positive and HPV-Negative Tonsillar Tumors and a Model System of Human

Keratinocyte Clones." *BMC Cancer* 16 (July). https://doi.org/10.1186/s12885-016-2430-y.

Vojtechova, Zuzana, Ivan Sabol, Martina Salakova, Lubomir Turek, Marek Grega, Jana Smahelova, Ondrej Vencalek, Eva Lukesova, Jan Klozar, and Ruth Tachezy. 2015. "Analysis of the Integration of Human Papillomaviruses in Head and Neck Tumours in Relation to Patients' Prognosis." *International Journal of Cancer* 138 (2): 386–95. https://doi.org/10.1002/ijc.29712.

Vojtechova, Zuzana, Jiri Zavadil, Jan Klozar, Marek Grega, and Ruth Tachezy. 2017. "Comparison of the MiRNA Expression Profiles in Fresh Frozen and Formalin-Fixed Paraffin-Embedded Tonsillar Tumors." *PLOS ONE* 12 (6): e0179645. https://doi.org/10.1371/journal.pone.0179645.

Vossler, J. L., B. A. Forbes, and M. D. Adelson. 1995. "Evaluation of the Polymerase Chain Reaction for the Detection of Human Papillomavirus from Urine." *Journal of Medical Virology* 45 (3): 354–60.

Wagner, Steffen, Shachi Jenny Sharma, Nora Wuerdemann, Jennifer Knuth, Henrike Reder, Claus Wittekindt, and Jens Peter Klussmann. 2017. "Human Papillomavirus-Related Head and Neck Cancer." *Oncology Research and Treatment* 40 (6): 334–40. https://doi.org/10.1159/000477252.

Walboomers, J M, M V Jacobs, M M Manos, F X Bosch, J A Kummer, K V Shah, P J Snijders, J Peto, C J Meijer, and N Muñoz. 1999. "Human Papillomavirus Is a Necessary Cause of Invasive Cervical Cancer Worldwide." *The Journal of Pathology* 189 (1): 12–19. https://doi.org/10.1002/(SICI)1096-9896(199909)189:1<12::AID-PATH431>3.0.CO;2-F.

Wan, Yunxia, Dimitrios Vagenas, Carolina Salazar, Liz Kenny, Chris Perry, Diego Calvopiña, and Chamindie Punyadeera. 2017. "Salivary MiRNA Panel to Detect HPV-Positive and HPV-Negative Head and Neck Cancer Patients." *Oncotarget* 8 (59): 99990–1. https://doi.org/10.18632/oncotarget.21725.

Wang, Hui, Qian Wu, Ying Zhang, Hua-Nan Zhang, Yong-Bin Wang, and Wei Wang. 2017. "TGF-B1-Induced Epithelial–Mesenchymal Transition in Lung Cancer Cells Involves Upregulation of MiR-9 and Downregulation of Its Target, E-Cadherin." *Cellular & Molecular Biology Letters* 22 (November). https://doi.org/10.1186/s11658-017-0053-1.

Wang, Jin-yan, Qian Zhang, Dan-dan Wang, Wei Yan, Huan-huan Sha, Jian-hua Zhao, Su-jin Yang, et al. 2018. "MiR-29a: A Potential Therapeutic Target and Promising Biomarker in Tumors." *Bioscience Reports* 38 (1): BSR20171265. https://doi.org/10.1042/BSR20171265.

Wang, Sophia S., Marcus Trunk, Mark Schiffman, Rolando Herrero, Mark E. Sherman, Robert D. Burk, Allan Hildesheim, et al. 2004. "Validation of P16INK4a as a Marker of Oncogenic Human Papillomavirus Infection in Cervical Biopsies from a Population-Based Cohort in Costa Rica." *Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology* 13 (8): 1355–60.

Wang, Yue, Xueshan Zhang, Fang Miao, Yanning Cao, Jiangnan Xue, Qizhi Cao, and Xiaoshu Zhang.

2016. “Clinical Significance of Leukocyte-Associated Immunoglobulin-like Receptor-1 Expression in Human Cervical Cancer.” *Experimental and Therapeutic Medicine* 12 (6): 3699–3705. https://doi.org/10.3892/etm.2016.3842.

Wang, Yuhan, Jie Wang, and Yuanshuai Huang. 2015. “MicroRNAs as New Biomarkers for Human Papilloma Virus Related Head and Neck Cancers.” *Cancer Biomarkers: Section A of Disease Markers* 15 (3): 213–18. https://doi.org/10.3233/CBM-150464.

Wilson, Ann S., Barbara E. Power, and Peter L. Molloy. 2007. “DNA Hypomethylation and Human Diseases.” *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* 1775 (1): 138–62. https://doi.org/10.1016/j.bbcan.2006.08.007.

Winter, Julia, Stephanie Jung, Sarina Keller, Richard I. Gregory, and Sven Diederichs. 2009. “Many Roads to Maturity: MicroRNA Biogenesis Pathways and Their Regulation.” *Nature Cell Biology* 11 (3): 228–34. https://doi.org/10.1038/ncb0309-228.

Xiao, Changchun, and Klaus Rajewsky. 2009. “MicroRNA Control in the Immune System: Basic Principles.” *Cell* 136 (1): 26–36. https://doi.org/10.1016/j.cell.2008.12.027.

Xu, Qing, Ling-Zhi Liu, Xu Qian, Qi Chen, Yue Jiang, Dan Li, Lihui Lai, and Bing-Hua Jiang. 2012. “MiR-145 Directly Targets P70S6K1 in Cancer Cells to Inhibit Tumor Growth and Angiogenesis.” *Nucleic Acids Research* 40 (2): 761–74. https://doi.org/10.1093/nar/gkr730.

Yan, Wusheng, Ignacio I Wistuba, Michael R Emmert-Buck, and Heidi S Erickson. 2010. “Squamous Cell Carcinoma – Similarities and Differences among Anatomical Sites.” *American Journal of Cancer Research* 1 (3): 275–300.

Yang, Rongcun, Patricia M. Day, William H. Yutzy IV, Ken-Yu Lin, Chien-Fu Hung, and Richard B. S. Roden. 2003. “Cell Surface-Binding Motifs of L2 That Facilitate Papillomavirus Infection.” *Journal of Virology* 77 (6): 3531–41. https://doi.org/10.1128/JVI.77.6.3531-3541.2003.

Yim, Eun-Kyoung, and Jong-Sup Park. 2005. “The Role of HPV E6 and E7 Oncoproteins in HPV-Associated Cervical Carcinogenesis.” *Cancer Research and Treatment: Official Journal of Korean Cancer Association* 37 (6): 319–24. https://doi.org/10.4143/crt.2005.37.6.319.

Yoshikawa, H, T Kawana, K Kitagawa, M Mizuno, H Yoshikura, and A Iwamoto. 1991. “Detection and Typing of Multiple Genital Human Papillomaviruses by DNA Amplification with Consensus Primers.” *Japanese Journal of Cancer Research: Gann* 82 (5): 524–31.

Zee, R. P. van der, O. Richel, H. J. C. de Vries, and J. M. Prins. 2013. “The Increasing Incidence of Anal Cancer: Can It Be Explained by Trends in Risk Groups?” *The Netherlands Journal of Medicine* 71 (8): 401–11.

Zhang, Qi, Louise Kuhn, Lynette A. Denny, Michelle De Souza, Sylvia Taylor, and Thomas C. Wright. 2007. “Impact of Utilizing P16INK4A Immunohistochemistry on Estimated Performance of Three Cervical Cancer Screening Tests.” *International Journal of Cancer* 120 (2): 351–56.

https://doi.org/10.1002/ijc.22172.

Zhang, Xiaoying, Harriet Gee, Barbara Rose, C. Soon Lee, Jonathan Clark, Michael Elliott, Jennifer R. Gamble, et al. 2016. “Regulation of the Tumour Suppressor PDCD4 by MiR-499 and MiR-21 in Oropharyngeal Cancers.” *BMC Cancer* 16 (1). https://doi.org/10.1186/s12885-016-2109-4.

Zhao, Jun-Wei, Fang Fang, Yi Guo, Tai-Lin Zhu, Yun-Yun Yu, Fan-Fei Kong, Ling-Fei Han, Dong-Sheng Chen, and Fang Li. 2016. “HPV16 Integration Probably Contributes to Cervical Oncogenesis through Interrupting Tumor Suppressor Genes and Inducing Chromosome Instability.” *Journal of Experimental & Clinical Cancer Research: CR* 35 (November). https://doi.org/10.1186/s13046-016-0454-4.

Zheng, Z. M, and C. C Baker. 2006. “Papillomavirus Genome Structure, Expression, and Post-Transcriptional Regulation.” *Frontiers in Bioscience: A Journal and Virtual Library* 11: 2286.

Znaor, Tin, Luka Vucemilo, Tomislav Kulis, and Ariana Znaor. 2013. “Incidence and Mortality Trends of Head and Neck Cancer in Croatia in the Period 1988–2008.” *Acta Oto-Laryngologica* 133 (3): 305–12. https://doi.org/10.3109/00016489.2012.743031.

8 SUMMARY

Head and neck squamous cell carcinoma (HNSCC) is the 6th most common malignancy worldwide, with a 5-year survival of only 40-50%. HNSCC are broadly categorized into two groups: human papillomaviruses (HPV) positive and HPV-negative cancer.

The aim of this study is to find more sensitive and specific epigenetic biomarkers, which could enable tailored therapy to the particular group of patients and improve the diagnostic, prognostic and therapeutic approaches.

Whole-genome methylation (WGM) analysis and miRNA profiling was performed in fresh HPV-positive and HPV-negative HNSCC using DNA methylation microarray and Next-Generation Sequencing (NGS). Data from WGM and NGS have been validated using pyrosequencing and qRT-PCR. Survival analysis was performed on both archival and fresh HNSCC.

Transcriptionally active HPV was found in 9.8% fresh HNSCC and 11.3% archival HNSCC. The majority of patients were older males and active smokers treated for advanced stage disease. Retrospective survival analysis failed to show positive influence of HPV activity on patient outcome, while survival analysis on fresh samples showed better survival for patients with transcriptionally active HPV. The statistical analysis determined difference between cancer and normal samples regarding methylome and miRnome profiles; NGS analysis showed 552 different miRNAs ($P < 0.05$), of which only miR-9 was significantly associated with HPV positivity by both NGS ($P = 0.0004$) and qRT-PCR ($P = 0.00156$). The miR-335, determined by NGS but not validated by qRT-PCR was also significantly ($P = 0.0006$) associated with HPV positive HNSCC. WGM profiling revealed 120,901 differentially methylated sites in cancer ($P < 0.05$), and from the top 20 differentially methylated genes four genes (*FBXO2, LAIR2, SPRR3, TRDC*) have been validated by pyrosequencing.

Integration analysis of miRnome and methlyome data revealed several potential epigenetic biomarkers in HPV-positive HNSCC: miR-9, miR-335, *PDE4D*, *SPRR3* and *LAIR2*.

9 PROŠIRENI SAŽETAK

Uvod

Karcinom pločastog epitela glave i vrata (*engl.* Head and Neck Squamous Cell Carcinoma, HNSCC) po učestalosti je šesta najčešća zloćudna bolest širom svijeta, dok je u Hrvatskoj 2015. zabilježeno 896 novih slučajeva. Petogodišnje preživljenje iznosi samo 40-50%. HNSCC se karakterizira prema primarnom anatomskom mjestu nastanka, a najčešća mjesta podrijetla karcinoma su usna šupljina, orofarinks i larinks. Obzirom na prisutnost papiloma virusa čovjeka (*engl*. Human PapillomaVirus, HPV), HNSCC dijelimo u dvije skupine: pozitivne na HPV, koji se javljaju u mlađoj populaciji i koji su uzrokovani HPV-posredovanom transformacijom stanica; te HPV-negativne HNSCC, koji se javljaju u starijoj populaciji i koji su uzrokovani prvenstveno mutagenim učincima duhana i alkohola. Iako se ove dvije skupine etiološki razlikuju, tretman liječenja je trenutno isti, no tedndencija je ka optimizaciji istog za svaku pojedinu skupinu oboljelih.

Cilj istraživanja

Cilj ovog istraživanja je pronaći osjetljivije i specifične biljege, kako bi se terapija prilagodila određenoj skupini oboljelih te unaprijedile dijagnostičke, prognostičke i terapeutske metode. Istraživanja na epigenetičkim promjenama u tumorima kao što su metiliranje DNA te ispoljavanje miRNA, predstavljaju adekvatan odabir takvih specifičnih bioloških biljega.

Pacijenti, materijal i metode

U ovom istraživanju napravljena je analiza metiliranja cijelog genoma i profiliranje miRNA u svježim uzorcima tumora na homogenoj populaciji i patološki dobro definiranim karcinomima. Svi uzorci su kategorizirani u skupine obzirom na prisutnost i dokazanu aktivnost HPV-a. Analizirano je metiliranje DNA u cijelom genomu (*engl.* whole-genome methylation, WGM), koristeći mikročipove (*engl*. microarray), a podaci profiliranja metiliranosti DNA potvrđeni su metodom pirosekvenciranja. MiRNA profiliranje je provedeno tehnologijom sekvenciranja slijedeće generacije (*engl.* Next-generation sequencing, NGS) metode, dok su podaci validirani pomoću kvantitativne reverzno-prepisane lančane reakcije polimerazom (engl. quantitative reverse transcription-polymerase chain reaction; qRT-PCR). Nadalje, provedena je temeljita analiza preživljavanja pomoću svježih i arhivskih uzoraka uklopljenih u parafin (*engl.* formalin fixed paraffin embedded; FFPE) orofaringealnog tumora i tumora usne šupljine kako bi se dobila jasnija slika razvoja bolesti u Hrvatskoj i sličnim populacijama.

Rezultati

Prisutnost transkripcijski aktivnog HPV virusa dokazana je u 9.8% svježih uzoraka tumora te u 11.3% arhivskih uzoraka tumora. U ukupnoj populaciji većinu su sačinjavali muškarci, koji su bili pušači, s medijanom od 60 godina. Tumori su bili u poodmaklom stadiju (TNM kategorija 3 i 4) sa slabo diferenciranim stanicama. U ovom istraživanju nije pronađena razlika između HPV-pozitivnih (HPV DNA i RNA+), HPV-neaktivnih (HPV DNA+ RNA-) i HPV-negativnih pacijenata s obzirom na dob, spol, stil života (pušenje i konzumacija alkohola) i stadij tumora u oba seta uzoraka. Nadalje, retrospektivna analiza preživljenja je pokazala nedostatak pozitivnog utjecaja HPV-a na ukupno preživljenje oboljelih, što je začuđujuće, s obzirom na dobro poznatu korelaciju HPV aktivnosti sa boljim preživljenjem i odgovorom na terapiju. Iako je bilo nemoguće napraviti petogodišnju analizu preživljenja na svježim uzorcima (vrijeme praćenja <5 god.), analiza pokazuje da je najduže preživljenje bilo upravo kod bolesnika s dokazanim transkripcijskim aktivnim HPV-om.

Temeljem analize ispoljavanja miRNA u cijelom genomu, kao i analize metiliranosti genoma, dobivene su razlike između oboljelih naspram zdravih pojedinaca, što je i očekivano.; dobivena je statistička razlika u 552 različito ispoljene miRNA molekule ($P < 0,05$), od kojih su miR-9, -21, -29a, -100, -106b, -143, -145 i -199b validirane pomoću qRT-PCR. Samo se miR-9 pokazao kao značajno povezan s HPV-om koristeći obje metode, NGS i qRT-PCR ($P = 0.0004$, $P = 0.00156$), dok se miR-335 statistički pokazao povezan s HPV-om samo u NGS analizi ($P = 0,0006$), ali nije validiran koristeći qRT-PCR.

Profiliranje metiliranjem DNA cijelog genoma otkrilo je 120.901 različito metiliranih mjesta u genomu oboljelih naspram normala (P <0.05). Statistička analiza utvrdila je najveću razliku u 20 gena (10 hipermetiliranih i 10 hipometiliranih u karcinomu) od kojih su četiri gena (*FBXO2*, *LAIR2*, *SPRR3*, *TRDC*) validirana pirosekvenciranjem. Nakon validacije, rezultati pirosekvenciranja su korelirali sa podacima metiliranja DNA cijelog genoma. Geni *SPRR3* i *FBXO2* bili su pojačano metilirani u raku, dok su *LAIR2* i *TRDC* smanjeno metilirani u raku.

Rasprava i zaključci

Integracijska analiza svih rezultata epigenetičkog profiliranja otkrila je ključne signalne puteve, koji su statistički povezani s HPV-pozitivnim i HPV-negativnim HNSCC. Analiza je dokazala da je većina epigenetski dereguliranih staničnih signalnih puteva, značajno povezanih s HPV-pozitivnim i HPV-negativnim HNSCC, uključena u procese endocitoze, diferencijacije, interakcije receptora s izvanstaničnim matriksom, u procese interakcije epitelnih stanica, metabolizam lipida i imunološki odgovor. Unutar HPV-pozitivne skupine karcinoma identificirani epigenetski deregulirani signalni putevi bili su oni koji sudjeluju u procesima endocitoze (P = 1.36E-07), adherentnim staničnim vezama (P = 2.14E-0) i biosintezom N-glikana (P = 7.31E-07), dok su kod HPV-negativne skupine karcinoma, regulacija staničnog

ciklusa, proteoliza posredovana ubikvitinom, TGF-β i mTOR bili najviše povezani s karcinomom (*P* = 5.18E-1, *P* = 5.33E-10, *P* = 1.34E-06 , *P* = 0.000605).

Integracijska analiza svih rezultata dobivenih analizom metiliranja DNA i ispoljavanja miRNA, otkrila je nekoliko preklapajućih ciljnih gena dereguliranih u HNSCC, što bi moglo doprinijeti ukupnoj analizi ciljnih epigenetičkih biljega: *SNX25*, *MDGA2*, *BRWD3*, *PDCD1LG2*, *HAPLN1*, *PDE4D*, *SPRR3* i *LAIR2*. Nakon potpune integracije analize gena dereguliranih u karcinomima glave i vrata (miRnome i methlyome) i signalnih puteva koji su uključeni u razvoj raka, predlažemo nekoliko potencijalnih epigenetičkih biljega u HPV-pozitivnim HNSCC: miR-9, miR-335, PDE4D, SPRR3 i LAIR2. Navedeni biljezi bi se svakako trebali evaluirati na studiji sa većim brojem uzoraka.

10 CURRICULUM VITAE

I was born on December 6, 1985 in Slavonski Brod, where I finished Classical gymnasium "fra Marijan Lanosović". I studied at the University of Zagreb, Faculty of Science. In 2010 I have graduated Biology and Chemistry and in 2012 Molecular Biology. During my last year of Molecular Biology, I started the Master thesis project "Biological properties of mesenchymal stem cells expanded *in vitro* in media with different human platelet lysate content" at the University Clinical Center Zagreb, under the supervision of Mirna Golemović, PhD, where I worked with mesenchymal stem cells and gained lots of experience and knowledge in research fields as hematopoiesis, cell differentiation and stem cells. From 2012-2014, I have started research on pathophysiology of primary myelofibrosis in Srđan Verstovšek, MD PhD, Laboratory in Houston, Texas. During my stay in Houston, I extended my laboratory skills and gained knowledge in cancer immunology and hematologic malignancies. In 2015 I came back to Croatia and started the PhD program at the Ruđer Bošković Institute, Zagreb, Croatia under supervision of Magdalena Grce, PhD, where I continued working in Cancer Biology, but this time in solid tumors. I implemented and improved my previous skills in immunology and cell differentiation and gained new skills in fields of virology, epigenetics and bioinformatics. My PhD study is conducted within the project "Epigenetic changes in head and neck squamous cell carcinoma – Epic-HNSCC" funded by the Croatian Science Foundation. Up to date, I have participated as co-author in the realization and writing of 3 scientific papers, while 3 are in the process of being published. I have 43 citations and a h-index of 2 (Scholar Google).

Publications:

1. Uri Rozovski, Srdan Verstovsek, Taghi Manshouri, Vilma Dembitz, **Ksenija Božinović**, Kate Newberry, Ying Zhang, Joseph E. Bove, Sherry Pierce, Hagop Kantarjian and Zeev Estrov. *An accurate, simple prognostic model consisting of age, JAK2, CALR, and MPL mutation status for patients with primary myelofibrosis.* **Haematologica**, 102 (2016), 1; 79-84. doi:10.3324/haematol.2016.149765. **IF (2016) = 7.702**

2. Srdan Verstovsek, Taghi Manshouri, Darrell Pilling, Carlos E. Bueso-Ramos, Kate J Newberry, Sanja Prijić, Liza Knez, **Ksenija Božinović**, David M. Harris, Erica L. Spaeth, Sean M. Post, Asha S. Multani, Raajit K. Rampal, Jihae Ahn, Ross L. Levine, Chad J. Creighton, Hagop M. Kantarjian and Zeev Estrov. *Role of neoplastic monocyte-derived fibrocytes in primary myelofibrosis.* **The Journal of experimental medicine**, 213 (2016), 9; 1723-1740. doi:10.1084/jem.20160283. **IF (2016) = 11.991**

3. Ivan Sabol, Nina Milutin Gašperov, Mihaela Matovina, **Ksenija Božinović**, Goran Grubišić, Ivan Fistonić, Dragan Belci, Laia Alemany, Sonja Džebro, Mara Dominis, Mario Šekerija, Sara Tous, Silvia de Sanjosé and Magdalena Grce. *Cervical HPV type-specific pre-vaccination prevalence and age distribution in Croatia.* **PLoS One**, 12 (2017), 7; e0180480-e0180480.

doi:10.1371/journal.pone.0180480. **IF (2017) = 2.806**

4. **Ksenija Božinović**, Ivan Sabol, Antonia Jakovčević, Zoran Rakušić, Mario Šekerija, Juraj Lukinović, Drago Prgomet, Magdalena Grce. Not all HPV-positive HNSCC confer Better Outcome: Retrospective Analysis of Oropharyngeal Cancer Patients. 2018 (submitted for revision)

5. Ivan Sabol, **Ksenija Božinović**, Emil Dediol, Nina Milutin Gašperov, Zuzana Vojtechova, Ruth Tachezy, Magdalena Grce. Genome-wide miRNA profiling reinforces the importance of miR-9 in human papillomavirus associated oral and oropharyngeal head and neck cancer. 2018 (submitted for revision)

6. Nina Milutin Gašperov, Ivan Sabol, **Ksenija Božinović**, Marinka Mravak-Stipetić M, Emil Dediol, Danilo Licastro, Simeone Dal Monego, Magdalena Grce. DNA methylome in head and neck cancer and potentially malignant oral lesions. 2018 (submitted for revision)

Presented posters/abstracts at international conferences:

1. **Ksenija Božinović**, Taghi Manshouri, Sean M. Post, Ying Zhang, Miguel Gallardo, Xiaorui Zhang and Srdan Verstovsek. Altered Expression and Mutation Analysis Of Heterogeneous Nuclear Ribonucleoprotein K In Bone Marrow Of Primary Myelofibrosis Patients. 55th ASH Annual Meeting and Exposition, American Society of Hematology, New Orleans, Louisiana, USA, December 2013.

2. Taghi Manshouri, Zeev Estrov, Liza Knez, Ying Zhang, **Ksenija Bozinovic,** Vilma Dembitz, Hagop M. Kantarjian and Srdan Verstovsek. Bone Marrow Stromal Cell Paracrine Factors from PMF Patients, but NOT Healthy Donors, Protect JAK2V617F Cells from Ruxotinib-Induced Apoptosis. 56th ASH Annual Meeting and Exposition, American Society of Hematology, San Francisco, California, USA, December 2014:

3. Uri Rozovski, Taghi Manshouri, Vilma Dembitz, **Ksenija Božinović**, Sherry Pierce, Hagop M. Kantarjian, Zeev Estrov and Srdan Verstovsek. JAK2, Calr and MPL Mutation Status Predicts the Survival Outcome of Patients with Primary Myelofibrosis. 56th ASH Annual Meeting and Exposition", American Society of Hematology, San Francisco, California, USA, December 2014.

4. Uri Rozovski, Taghi Manshouri, Vilma Dembitz, **Ksenija Božinović**, Sherry Pierce, Hagop Kantarjian, Zeev Estrov and Srdan Verstovsek. Prognostic model based on JAK2, CALR and MPL mutation status and age predicts the survival outcome of patients with primary myelofibrosis. Society of Hematologic Oncology, SOHO Annual Meeting Proceedings", Houston, Texas, USA, September 2014.

5. **Ksenija Božinović,** Jasminka Golubić Talić, Ivan Sabol, Nina Milutin Gašperov, Marjana Knežević, Marinka Mravak Stipetić and Magdalena Grce. Alpha, Beta and Gamma Papillomaviruses Prevalence in Oral Mucosa. DNA Tumor Virus Meeting, Trieste, Italy, July 2015.

6. **Ksenija Božinović,** Taghi Manshouri, Zeev Estrov, Darrell Pilling, Liza Knez, Kate J. Newberry,

Carlos E. Bueso-Ramos, Sean Post and Srdan Verstovsek. Dysregulation of microRNA cluster miR-29 in fibrocytes from patients with JAK2 V617F -positive primary myelofibrosis. Leukemia & Limphoma, Dubrovnik, Croatia, September, 2015.

7. **Ksenija Božinović**, Marinka Mravak-Stipetić, Magdalena Grce and Nina Milutin Gašperov. CCNA1 promoter methylation distinguishes oral lesions. Game of Epigenomics, Dubrovnik, Croatia, April 2016.

8. **Ksenija Božinović**, Ivan Sabol, Emil Dediol, Ruth Tachezy and Magdalena Grce. Differentially expressed miRNAs in fresh tumor samples distinguishes HPV-positive and HPV-negative HNSCC. CLEPSO, 7th International Meeting of the Clinical Epigenetics International Meeting, Düsseldorf, Germany March 2017.

9. Nina Milutin Gašperov, **Ksenija Božinović**, Ivan Sabol, Emil Dediol, Marinka Mravak Stipetić and Magdalena Grce. DNA methylation in healthy oral mucosa and head and neck cancer. CLEPSO, 7th International Meeting of the Clinical Epigenetics International Meeting, Düsseldorf, Germany March 2017.

10. Nina Milutin Gašperov, Ivan Sabol, Emil Dediol, **Ksenija Božinović** and Magdalena Grce. Candidate Methylation Biomarkers for Head and Neck Cancer by Infinium MethylationEPIC BeadChip. 3rd Annual Epigenetics Discovery Congress", London, Great Britain, October 2017.

11. **Ksenija Božinović**, Ivan Sabol, Emil Dediol, Zuzana Vojtechova, Ruth Tachezy and Magdalena Grce. Identification of Differentially Expressed miRNAs in Head and Neck Squamous Cell Carcinoma Implicated in Immune System. 3rd Annual Epigenetics Discovery Congress, London, Great Britain, October 2017.

12. Anamaria Đukić, Jayashree Vijay Thatte, **Ksenija Božinović**, Emil Dediol, Magdalena Grce, Lawrence Banks and Vjekoslav Tomaić. Biochemical Characterization of Naturally Occurring HPV-16 E6 and E7 Variants. Power of viruses, Poreč, Croatia, May 2018.