VICTOR BABEȘ UNIVERSITY OF MEDICINE
AND PHARMACY TIMIȘOARA
FACULTY OF DENTAL MEDICINE
DEPARTMENT I

GABRIELA-EMILIA AVRAM

PhD THESIS

EPIGENETIC MARKERS IN THE PATHOLOGY OF ORAL CANCER

A B S T R A C T

Scientific Coordinator
PROF. UNIV. DR. ANGELA CODRUȚA PODARIU

T i m i ș o a r a
2 0 2 2

TABLE OF CONTENTS

SPECIAL PART

INTRODUCTION

Oral cavity cancer, along with other cancers, has been and still is one of the great challenges in medicine.

Oral squamous cell cancer is the predominant tumor among oral cancers, estimated at 90%, and is considered the most aggressive and widespread tumor. Despite progress and acquisitions in the management of oral squamous cell cancer based on etiopathogenesis data, it has been estimated that 45% of patients have a poor prognosis due to the possibility of local recurrence or metastasis.

Previous and current molecular genetic studies are trying to clarify the mechanisms underlying the malignant process. What something is well shown are the findings of gene mutation, gene amplification, and chromosomal restructuring as being present in most oral cancer patients.

Although many researchers have conducted studies to explain the molecular signatures involved in the transition from the normal epithelial state to the premalignant state in oral carcinoma, one area that is still partially deciphered is that of epigenetic changes. Although epigenetic changes have now been linked to oral tumor initiation and progression, they have only been partially characterized.

Targeting the etiological heterogeneity of oral squamous cell carcinoma (OSCC) and understanding the mechanisms underlying the difference in specific methylation patterns, I aimed to study the role of smoking in triggering the tumor process by identifying possible epigenetic changes at the DNA level in the oral epithelial cell that could be caused by tobacco as well as identifying and evaluating epigenetic markers involved in the pathology of squamous cell carcinoma of the oral cavity.

THE GENERAL PART

1. THE ORAL CAVITY NEOPLASM

Oral tumors account for about 85% of head and neck neoplasms and are one of the severest conditions, with reduced prognosis coming from the fact that it is usually discovered in the late stage of the disease due to lack of clinical signs in the early stages or location in the posterior areas, barely visible, of the mouth.

1.1. CLINICAL AND MORPHO-PATHOLOGICAL CONSIDERATIONS

Oral cavity cancers are histopathologically diverse and include carcinomas, sarcomas, and melanomas. Of all oral cavity cancers, 90% are squamous cell carcinomas.

Clinically, oral squamous cell carcinoma (OSCC) presents different features influenced by etiology, location, and stage of the neoplastic process. In the initial stages, the most common form, regardless of location, is ulceration, and usually traumatic. In the advanced stage the form mixed with ulcerative or lump appearance. In ventral side of the tongue and the floor of the mouth, 40% of patients have early lymph node involvement, and, in T4, this percentage goes up to 90%.

In addition, special attention should be addressed to the morpho-pathological changes in the tumor tissue. The poor cell differentiation leads to increased recurrence and decreased survival. The presence of other cell types in the tumor tissue influences the

development of the pathological process by the action they exert on the tumor cells themselves. Cancer-associated fibroblasts (CAF) control initiation, progression, angiogenesis, and metastasis by interacting with tumor cells through tumor growth factors (TGF), cytokines, or exosome (mRNA, DNA, proteins, and miRNA). Tumor-induced T lymphocytes and granulocytes are associated with increased tumor proliferation, metastasis, and reduced survival.

1.2. ETIOLOGICAL FACTORS INVOLVED IN THE PATHOGENESIS OF OSCC.

1.2.1. BIOLOGICAL FACTORS

1.2.1.1. Human papillomavirus (HPV)

The most commonly detected HPV virus types found in the oral cavity, both in healthy patients and those with benign or malignant lesions, are types 6,11,16,18, 31, 33, 35, 52, 58.

HPV infection of the oral mucosa results in the virus's localization to the basal layer where viral oncoprotein, particularly E6 and E7, influence cell proliferation and DNA synthesis. In HPV-positive oral tumors, a series of chromosomal aberrations, such as amplifications at 3q, 16q, and 20q, were also found. Transcriptional profiling of patients with head and neck SCC revealed the influence of HPV presence on miRNA expression. Besides these changes in patients with HPV-positive SCC, they also found changes in methylation of both the viral and cellular genomes.

Approximately 1/3 of HPV-associated cancers are co-infected with other viruses such as Epstein-Barr (EBV), herpes simplex (HSV), cytomegalic (CMV), and human immunodeficiency virus (HIV).

1.2.1.2. Human herpes virus

The human herpes virus is a family of viruses that includes herpes simplex and Epstein-Barr viruses. Infection with EBV and HSV is very common among the population, affecting about 9 out of 10 adults. The involvement of EBV in the pathogenesis of SCC is questionable. Jiang et al. (2015) in an in vitro cell line study investigates the role of viral coinfection in the tumor process. The results of this study suggest the involvement of HPV/EBV coinfection in increasing invasiveness, not cell proliferation.

Data on the involvement of HSV-1 in the pathogenesis of oral SCC are not very conclusive, although older studies show its presence in the cancer cell without detection in the normal mucosa of the same patient. The low percentage of HSV-1 positive SCC cases detected suggests that it does not play an important role in the pathogenesis of oral SCC.

1.2.2. VICIOUS HABITS

1.2.2.1. Tobacco

Tobacco, regardless of the form of ingestion (smoking or chewing), is associated with the development of cancer of the jugal mucosa, gums, and lips.

On the structural level, tobacco causes epithelial disorganization by breaking cellular attachment and reducing the number of cell layers, as well as by altering the shape and size of cells. The effect of tobacco is not limited to these structural changes, but is also involved in disturbing the balance between cell proliferation and cell death.

Since they considered epigenetic modifications to be involved in the initial stages of the tumor process, the impact of tobacco on the epigenome with consequences in carcinogenesis was also considered. Using the tongue epithelium of smokers, researchers found increase methylation of CpG islands that can reach up to 50% in the case of MRE11A and PMS2 genes. Differences in methylation pattern between smokers and former smokers, and in methylation pattern concerning the number of cigarettes smoked were also found in case of non-coding RNAs.

1.2.2.2 Alcohol

It is considered that 70% of oral cancer patients are drinkers. Saad et al. (2015) identify 8 microRNAs (miR-30a, miR-101, miR-675, miR-934, miR-1266, miR-3164, miR-3178, and miR-3690) associated with alcohol consumption in squamous cell carcinomas of the head and neck. Besides these, miR-375 and miR-21 were also found to have increased expression in alcohol-associated oral SCCs.

2. EPIGENETIC MARKERS

2.1. GENERAL CHARACTERISTICS

Epigenetic mechanisms are key elements of developmental processes and contribute to genetic diversity. Disturbances of these mechanisms or in the markers involved in these mechanisms can lead to pathological phenomena, including cancer. The main epigenetic processes and markers are DNA methylation, histone modifications such as acetylation and methylation, and non-coding RNA.

2.2. EPIGENETIC MECHANISMS

2.2.1 DNA METHYLATION

Methylation is a fundamental process of normal development, and any change in it results in the transmission of a signal at the chromatin level that will cause changes in transcription. Predominantly DNA methylation occurs in the cytosine preceding a guanine (CG regions) but is also found in the non-CG clusters (CA, CT). Promoter methylation reflected in reduced expression of genes involved in DNA repair, tumor suppressor genes, or those involved in invasion and metastasis processes is a hallmark of tumorigenesis. Unlike promoter hypermethylation, hypermethylation at the gene body level is associated with increased gene expression, and hypomethylation is associated with both increased and decreased gene expression. In contrast, intergenic methylation has the role of inactivating the expression of potentially pathogenic genetic elements (transposons or viral elements) to protect the structural and functional integrity of the genome. Global hypomethylation is the consequence of a reduced methylation pattern at the transposable elements rather than at the gene body. Reduced methylation characterized by the appearance of epimutations is associated with the chromosomal instability found in cancers.

2.2.2 DNA HYDROXYMETHYLATION

Alteration of DNA methylation patterns occurs by the appearance of new methylated regions and by the demethylation of existing ones. Oxidation of 5mC by TET family enzymes produces 5 hydroxymethylcytosines (5hmC) as the first product, which, upon further oxidation, forms 5 formylcytosine (5fC) and 5 carboxycytosine (5caC). The pattern of 5hmC in the gene is specific to the type of tissue from which it originates. The overall rank of 5-hmC is affected by the degree of differentiation and cellular maturity.

2.2.3 HISTONE METHYLATION

Histone methylation comprises the addition of a methyl group from S-adenosylmethionine (SAM) to the N-terminus of histones. Like DNA methylation, histone methylation is reversible. Histone methylation (H) is a dynamic process with different consequences for regulating transcription, chromatin, and DNA repair depending on the position and methylation pattern of lysine (K) or arginine (R).

2.3 NON-CODING RNA

Non-coding RNAs are a class of RNAs that do not ultimately result in translation into a protein.

2.3.1. CLASSIFICATION OF NON-CODING RNA MOLECULES

Non-coding RNAs are classified according to their strand length into short non-coding RNA (sncRNA) and long non-coding RNA (lncRNA).

2.3.1.1 Short non-coding RNA

The main short non-coding RNA molecules involved in epigenetic processes are miRNA, siRNA, and piRNA. Their role is complex, and they are involved both in regulating messenger RNA (mRNA) transcription (by inhibiting its expression or degradation) and in DNA rearrangement and chromatin-mediated inhibition of gene expression.

MicroRNA is involved in stem cell reproduction and differentiation. MicroRNAs also play a role in regulating the expression of other microRNAs. Besides their involvement in biological processes, about 30% of miRNAs are also involved in tumor processes. Studies have shown that miRNAs act on several genes and may have oncogenic or tumor suppressor effects.

siRNA is one of the most studied non-coding RNAs in vivo or in vitro experiments to determine therapeutic strategies in cancer. Many studies have shown that the administration of specific siRNAs alters the expression of genes involved in the genesis, development, and local and distant spread of tumor processes. Besides direct inhibition, siRNA may also influence patient response to treatment in the case of drug resistance.

piRNAs are found almost exclusively in germ cells from adult testicular tissue, and fetal ovarian tissue. Although piRNA shows increased expression in germ cells without being detected in the somatic cell, recent studies have identified piRNA with increased expression in cancers.

2.3.1.2 Long non-coding RNA

The function of lncRNA is to regulate the expression of tumor suppressor genes or oncogenes. Some mechanisms of regulation of genes located intrachromosomally (cis) are transcription factor binding, chromatin condensation, and gene methylation, and those on other chromosomes (trans) include modulation of factors involved in transcription and interaction with chromatin remodeling complexes. Like other non-coding RNAs, lncRNA can induce cellular changes that trigger the tumor process.

2.3.2 CLINICAL IMPLICATIONS OF NON-CODING RNA IN ORAL CANCER

Since studies have shown that non-coding RNAs have tissue specificity, their involvement in physiological and pathological processes occurring in oral epithelial cells has also been considered. The most studied non-coding RNAs are miRNAs. In recent years, many studies have highlighted the involvement of miRNA in the initiation of OSCC.

Researchers have also observed changes in methylation in microRNAs associated with oral SCC. Methylation of the miR-137 promoter is one factor influencing the reduced expression of this microRNA in oral SCC. They have also detected changes in methylation pattern in oral SCC in miR-127, miR-205, and the miR-200 family correlated or not with their gene expression at the tissue level.

The latest research in oral malignant pathology has also focused on long non-coding RNA. Long non-coding RNA has a different expression in tumor tissue than normal epithelial tissue in the oral cavity and, like microRNA, can have oncogenic and tumor suppressor gene roles. It is involved in the oral cavity's proliferation, migration, and invasion of tumor processes. Studies to detect the mechanisms changing lncRNA expression in OSCC have only identified gene hypermethylation in MEG3.

SPECIAL PART

3. CHANGES IN ORAL MUCOSAL DNA RELATED TO TOBACCO SMOKE EXPOSURE – STUDY 1

3.1 PURPOSE OF THE STUDY

This study aims to identify epigenetic DNA changes in the cells of the jugal mucosa related to exposure to tobacco products (cigarettes) in subjects who smoke and evaluate them compared to non-smokers.

3.2 MATERIAL AND METHOD

3.2.1 STUDY PARTICIPANTS

The subjects enrolled in this study were volunteers. The total number of participants who agreed to take part in the study was 47, divided into 3 groups:
I. Smoking group: 19
II. Non-smoking group: 22
III. Former smokers: 6

3.2.2 COLLECTION OF BIOLOGICAL SAMPLES

We performed jugal mucosal cell collection using sterile OmniSwab sterile buccal swabs.

3.2.3 DNA ISOLATION

Isolation of genomic DNA from the sampled cells was performed using the specific buccal swab kit QIAamp DNA Mini kit (Qiagen, Hilden, Germany) following the protocol specifications.

3.2.4. SPECTROPHOTOMETRIC ANALYSIS OF THE DNA

Analysis of the quantity and quality of DNA was performed using the NanoDrop Microvolume Spectrophotometer.

3.2.5 DETECTION OF THE GLOBAL DNA METHYLATION IN ORAL EPITHELIAL CELLS FROM THE JUGAL MUCOSA.

Detection of the level of global DNA methylation in oral epithelial tissue was performed with the MethylFlash Methylated DNA Fluorometric Quantification Kit following the specifications in the brochure.

3.2.6 DETECTION OF THE OVERALL PERCENTAGE OF HYDROXYMETHYLATION OF DNA FROM ORAL EPITHELIAL CELLS OF THE JUGAL MUCOSA.

As for the determination of global DNA methylation, for the quantification of the level of hydroxymethylation, we used the fluorometric ELISA method using the hydroxymethylation specific kit Hydroxymethylated DNA Fluorometric Quantification Kit.

Determination of the absolute methylation and hidroxymethylation value was performed using the generated standard curve and the calculation formulas in the protocol. Since the samples were in duplicates, the relative fluorescence value (RFU) of each sample was calculated by averaging the RFU value of the duplicates.

3.2.7 STATISTICAL ANALYSIS

The results of the analyses were expressed as Mean and SD (Standard Deviation). After checking the normality of the distribution using the D'Agostini & Pearson Omnibus Test, the percentage of methylation and hydroxymethylation was compared between groups by using the unpaired t-test or Mann-Whitney U test. The value of $p< 0.05$ was considered statistically significant.

3.3 RESULTS

3.3.1 QUANTIFICATION OF EXTRACTED DNA

The spectrophotometric analysis of extracted DNA measured at A260/A280 showed a DNA purity between 1.09 and 3

3.3.2 GLOBAL DNA METHYLATION ANALYSIS

Forty-seven subjects (25 men and 22 women) aged between 26 and 89 participated in this study. Average daily cigarette consumption was not significantly different in smokers compared to former smokers.

Analyzing the data obtained from this analysis, we observed that although the methylation level is lower in smokers (3.1%) and former smokers (2.1%) than in non-smokers (4.1%) the difference is not statistically significant. The situation is different if we consider the intensity of smoking (number of cigarettes/day and duration of smoking in years).Those who smoked over 20 cigarettes/day, as well as those who smoked for over 20 years, show a significant increase in DNA methylation levels ($p=0.0294$ and $p= 0.0273$, respectively). Because of the small number of study participants in this stratified analysis, former smokers were not included.

3.3.3 GLOBAL DNA HYDROXYMETHYLATION ANALYSIS

Because of the small number of former smokers taking part in the study, the comparative analysis for global DNA hydroxymethylation was made only between smokers and non-smokers.

Through this analysis, we identified a low level of overall DNA hydroxymethylation in smokers compared to non-smokers. Non-smokers had more than twice the percentage of overall DNA hydroxymethylation than smokers.

3.4. DISCUSSIONS

This pilot study explored the impact of smoking on global DNA methylation and hydroxymethylation in oral epithelial cells.

Recent studies have identified a relationship between smoking and extensive changes in DNA methylation levels (global hypomethylation and hypermethylation at the promoter level of many genes).

Zeilinger et al. (156) examining blood samples observed hypomethylated and hypermethylated CpG islands in smokers compared to non-smokers, without being able to identify a precise direction and amplitude of global DNA methylation levels under the influence of smoking. Our findings show a 25% decrease in methylation status in smokers. The decrease in methylation level was more pronounced in former smokers (about 50%). This may be the consequence of the cumulated impact of the age of the subjects in this group.

In order to identify a relationship between the intensity of smoking and the global level of methylation, we divided the subjects into two categories determined by the number of years the patient smoked and the number of cigarettes smoked per day. We considered those who smoked over 20 cigarettes per day and those who smoked for over 20 years as

heavy smokers. Consistent with the results reported by De Araujo Costa et al. (85) we found a significant increase (twice) in overall methylation levels in heavy smokers versus light smokers.

In terms of hydroxymethylation, the results showed a considerably lower level of global DNA hydroxymethylation than that of global DNA methylation. It agrees with preceding investigations of 5-mC and 5-hmC levels in human tissue. We also found that smokers had a significantly reduced level of overall DNA hydroxymethylation compared to non-smokers, a finding not impacted by the age of the subjects.

3.5. CONCLUSIONS

This research determines the global DNA methylation and hydroxymethylation in the smoking-exposed epithelial cells from the oral cavity. We found a decrease of 25% in methylation status in smokers and a decrease of 50% in former smokers. Regarding hydroxymethylation status, we also found a reduced level in smokers.

4. IDENTIFICATION AND EVALUATION OF EPIGENETIC MARKERS INVOLVED IN THE PATHOLOGY OF SQUAMOUS CELL CARCINOMA OF THE ORAL CAVITY - STUDY 2

4.1 PURPOSE OF THE STUDY

This study aimed to investigate genomic methylation changes associated with oral squamous cell carcinoma (OSCC) pathology.

The main objectives of this study are:

- Identify the genome-wide methylation profile in oral cancer samples.
- Evaluation and validation of methylation profiling at selected sites associated with oral mucosal cancer pathology
- Assessment of the expression of specific microRNA molecules to establish a correlation between the expression level of these molecules and the methylation profile of specific sites in the genome
- In silico bioinformatics analysis of data got from the investigation of methylation of the genome to generate the expression profile of the differently methylated genes identified and their association with specific signaling pathways or biological processes.

4.2 MATERIAL AND METHOD

4.2.1 MATERIAL USED IN THE STUDY

To determine the degree of global methylation with OSSC, we used 13 pairs of paraffin blocks containing tumor tissue (T) and adjacent normal tissue (N) from OSCC male patients got from the archive of the Department of Pathology of the Municipal Emergency Clinical Hospital of Timisoara.

4.2.2 INCLUSION/EXCLUSION CRITERIA

Inclusion criteria:

- Histopathologically confirmed diagnosis of oral squamous cell carcinoma (OSCC).

- For each paraffin block with tissue representing oral SCC, there should be a paraffin block with adjacent non-infiltrated tumor tissue.

Exclusion criteria:

- Insufficient tissue in the paraffin block
- Association with dysplastic or lichen planus lesions
- Cases with HPV cytopathic effect.

4.2.3 VALIDATION AND SAMPLING.

After selecting paraffin blocks that corresponded to the inclusion criteria, they were examined in order to eliminate samples with one or more exclusion criteria. After validation of the samples, tissue fragments were collected from each paraffin block included in the study.

4.2.4 ISOLATION OF GENOMIC DNA FROM PARAFFIN SAMPLES.

QIAamp DNA FFPE Tissue Kit was used for DNA isolation following the specifications in protocol.

4.2.5 SPECTROPHOTOMETRIC ANALYSIS OF THE OBTAINED DNA

Analysis of the quantity and quality of DNA obtained was performed using the NanoDrop Microvolume Spectrophotometer.

4.2.6 IDENTIFICATION OF GENOME-WIDE METHYLATION PROFILE IN CANCER SAMPLES VERSUS ADJACENT NORMAL TISSUE

For the determination of CpG loci showing differential methylation in cancer samples versus adjacent normal tissue, we performed Next Generation Sequencing (NGS) using the Infinium Human Methylation 450k BeadChip Kit.

4.2.7 STATISTICAL ANALYSIS

Bioinformatics analysis of the data was performed using the Bioconductor software programs minfi and limma from the R statistical program (www.R-project.org).

4.2.8 IDENTIFICATION OF DIFFERENTIALLY METHYLATED MICROARN LOCI

We analyzed all differentially methylated genes (adjusted p-value ≤ 0.05) to detect the presence of microRNAs. After detection of differentially methylated microRNAs, we identified the number of CpG loci generating differential methylation within the detected microRNAs.

4.2.9 CPG ISLANDS PREDICTION

We used the MethPrimer platform to calculate the prediction of CpG islands at these loci, to detect CpG islands capable of generating primers specific for methylation PCR and primers specific for bisulfite conversion-based PCR.

4.2.10. CROSS-VALIDATION WITH LITERATURE ANALYSIS.

The literature data on microRNA expression in oral squamous cell carcinoma required for this analysis were obtained using PubMed.

4.2.11. *IN SILICO* ANALYSIS OF microRNA/ FUNCTIONAL INTEGRATION OF microRNA

4.2.11.1 Prediction of target genes

A first step in assessing microRNA function is the identification of target genes. For faster and more comprehensive identification of target genes, we used the computerized method miRWalk 3.0.

For the prediction of microRNA target genes identified at the 3' UTR, 5'UTR, and CDS transcription sites, we set a Bonferroni adjusted p-value ≤ 0.05. Besides this, we also considered the following criteria:

- MicroRNA binding probability to be at least 0.99
- Validation by at least one other algorithm: miRDB or TargetScan.

4.2.11.2 Functional analysis of target genes

After identifying the set of target genes corresponding to each microRNA type, we compared these datasets with the Differentially Expressed Genes (DEG) obtained after Geo2R analysis of the GSE138206 dataset (False Discovery Rate adjusted by Benjamini and Hochberg correlation procedure FDR ≤0.05). The common target gene sets revealed from this comparison were investigated using KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway enrichment analysis and cluster analysis from the STRING 9.1 platform using an adjusted FDR ≤ 0.05 and high confidence interaction score (0.700).

4.3. OUTCOME

4.3.1 IDENTIFICATION OF GENOME-WIDE METHYLATION PROFILE IN CANCER SAMPLES VERSUS ADJACENT NORMAL TISSUE

The final results of the Infinium 450K array-based Methylation Assay are expressed as beta and M values. The analysis of beta and M values resulting from the investigation of samples using the Illumina 450K bead array reflects that, overall, there is a high degree of hypomethylation in tumor tissue samples compared to normal tissue samples.

Illumina microarray investigation revealed 440365 loci with different methylation, of which 1468 showed significant differences in methylation between cancer and adjacent tissue samples without tumor invasion (adjusted p-value ≤ 0.05).

4.3.2. IDENTIFICATION OF DIFFERENTIALLY METHYLATED MICROARN LOCI

Analyzing the 1469 methylated differential loci, we detected five loci coding for four microRNAs: hsa-miR-124-3, has-miR-24-1, hsa-miR-769, and hsa-miR-4500.

4.3.3 CPG ISLAND PREDICTION

The MethPrimer analysis revealed several issues. CpG island prediction identified CpG islands upstream of hsa-miR-124-3 and hsa-miR-769, but not for hsa-miR-4500 and hsa-miR-24-1. Five sets of BSP primers, with their sequences, were found for all four microRNA; each consisting of two set primers forward and reverse.

4.3.4 CROSS-VALIDATION WITH LITERATURE.

Analysis of the data led to the identification of 191 microRNAs. 101 showed low gene expression, and 88 showed high gene expression in OSCC samples compared to normal tissue. In these lists, we found hsa-miR-24-1with high gene expression and hsa-miR-124-3 and hsa-miR-769 with low gene expression.

4.3.5 IN SILICO ANALYSIS OF microRNA/ FUNCTIONAL INTEGRATION OF microRNA

4.3.5.1 Prediction of target genes

After gene selection, considering the established criteria (binding probability ≤ 0.99 and validation by at least one other algorithm) and duplicate removal, we found 4147 target genes for hsa-miR-24-1, 6084 target genes for hsa-miR-124-3, 743 target genes for hsa-miR-4500 and 7256 target genes for hsa-miR-769.

4.3.5.2 Functional analysis of target genes

The Geo2R analysis selected 2289 genes whose gene expression was altered, 1490 having low gene expression (FC< -0.538) and 799 having high gene expression (FC <1.0).

For the list of common genes obtained by comparing the target genes of hsa-miR-24-1 with the genes with a reduced expression from the GSE138206 dataset, we obtained a protein-protein interaction (PPI) network characterized by 259 nodes and 68 edges. Qualitative

evaluation of the PPI yielded a p of 3.75e-12. KEGG functional analysis revealed no significant biological mechanism associated with this gene set.

The common gene list between hsa-miR-124-3 target genes and genes with increased expression from the GSE138206 dataset generated a PPI image that characterized 266 nodes and 562 edges. Here, the p-value was below 1.0e-16. Cluster analysis using the MCL method led to the identification of three functionally enriched clusters in the KEGG pathways. Cluster 1 is enriched in genes involved in viral infections, cluster 2 in genes involved in protein digestion and adsorption, ECM receptor interaction pathway, and AGE-RAGE signaling pathways, and cluster 3 in genes involved in AGE-RAGE signaling pathways, ECM-receptor interaction, and cancer-associated signaling pathways.

With has-miR-769 and genes with increased expression in the GSE138206 dataset, the protein-protein interaction network shows 290 nodes and 581 edges. The p-value is less than 1.0-16. The MCL cluster analysis of the network also resulted in three functionally enriched clusters in the KEGG pathways. Cluster 1 is enriched in genes involved in viral infections, cluster 2 in genes involved in protein digestion and adsorption, ECM receptor interaction pathway, and focal adhesion, and cluster 3 in genes involved in the p53 signaling pathway.

Because of the low number of common genes found between has-miR-4500 target genes and genes with differential gene expression in the GSE138206 dataset, the protein-protein interaction network does not show statistical significance.

A comparison of the hsa-miR-24-1 and hsa-miR-4500 target-specific gene sets with the reduced expression gene set from the GSE138206 generated a PPI image with 290 nodes, 95 edges, and a p-value of 7.46e-12 for PPI quality assessment. Here, KEGG functional analysis revealed only one biologically significant mechanism, namely beta-Alanine metabolism (FDR of 0.03).

The protein-protein interaction network in the group's case formed by the specific target genes hsa-miR124-3, hsa-miR-769, and hsa-miR-4500, together with the genes with increased gene expression in the GSE138206 dataset, shows 319 nodes, 816 edges with p-value less than 1.0e-16. KEGG functional analysis revealed 36 significant physiological mechanisms involving these genes, the most significant being the ECM receptor interaction mechanism.

4.4. DISCUSSION

Although it is the most common form of oral cancer, squamous cell carcinoma still presents many uncertainties regarding the epigenetic mechanisms involved in its initiation, development, tumor invasion, treatment, progression, and prognosis.

In this study, we analyzed the global methylation profile in OSCC in order to detect CpG loci associated with microRNAs displaying changes in methylation status. This investigation resulted in the detection of four microRNAs (hsa-miR-24-1, hsa-miR-124-3, hsa-miR-769, and hsa-miR-4500) exhibiting hypermethylation at multiple CpG loci in OSCC samples compared to adjacent normal tissue.

In OSCC, hsa-miR-24 is associated with proliferation, migration, and tumor invasion. Previous studies have reported that hsa-miR-24 is present not only in tumor tissue but also in saliva, serum, and plasma of patients with OSCC making it a possible biomarker for the prediction and diagnosis of oral cancer. Also, with this microRNA, we can say that the results obtained agree with those in the literature because hsa-miR-124-3 has been detected to have decreased gene expression in OSCC in both cell lines and animal models. The results of the KEGG functional analysis revealed beside the best-known viral infections HPV, EBV, and HSV, also hepatitis C, hepatitis B, and Influenza A. Hepatitis C is associated with the development of oral conditions considered being risk factors in oral SCC pathology. For cluster 2 this result confirms the findings of previous cell line studies that in OSCC, hsa-miR-124-3 directly suppresses the expression of ITGB19 (integrin subunit beta 1), considered being a key factor in modulating the ECM receptor interaction mechanism.

In the case of hsa-miR-769, the most significant ones in cluster 1 are EBV infection, hepatitis C and influenza A. Besides viral infections, another common mechanism of the two miRNAs is the ECM-receptor interaction. This result agrees with the research of Zhang et al., who reported, using the ”attract” method, that the extracellular matrix receptor interaction mechanism may play a significant role in the development of OSCC. Another mechanism is the AGE-RAGE signaling pathway. A special role of RAGE in the pathogenesis of OSCC has been demonstrated in an in vitro study in which this receptor was shown to be activated due to the paracrine secretion of High-mobility group box 1 (HMGB1) by the oral SCC cells. This activation enhances tumor progression and triggers bone destruction.

As for the fourth microRNA identified in this study, has-miR-4500, there is as yet no information on its expression in OSCC.

Performing KEGG functional analysis of all specific target genes hsa-miR-124-3, hsa-miR-769, and hsa-miR-4500 compared to genes with increased expression in the GSE138206 dataset we found that, again, the most significant is the ECM receptor interaction mechanism. In contrast, KEGG functional analysis of the common genes of hsa-miR-24-1, hsa-miR-4500, and genes with low expression in the GSE138206 dataset revealed only one significant mechanism: beta-Alanine metabolism. Alanine has been detected in both saliva and tumor tissue of patients with head and neck SCC, suggesting its use as a non-invasive biomarker.

4.5. CONCLUSIONS

In this study, we identified 4 microRNAs: hsa-miR-24-1, has-miR1-24-3, hsa-miR-769, and hsa-miR-4500 with differential methylation in OSCC samples versus adjacent normal tissue. Although all microRNAs were hypermethylated, the location of the CpG islands that caused this hypermethylation may explain the positive correlation with gene expression with hsa-miR-24-1 or the negative correlation with gene expression with has-miR-124-3, hsa-miR-769.

Exploration of the functional mechanisms of the identified miRNA target genes revealed their involvement in viral infections, the ECM-receptor interaction pathway, protein digestion, and adsorption and the AGE-RAGE signaling pathway. Along with these mechanisms, the exploration of the functional mechanisms of the target genes in hsa-miR-24-1 and hsa-miR-4500 revealed their involvement in the metabolism of beta-alanine.

CONCLUSIONS AND PERSONAL CONTRIBUTIONS

This study aimed to investigate the role of smoking in the initiation of the tumor process by identifying possible epigenetic changes at the DNA level in oral epithelial cells that tobacco can cause and to identify and assess epigenetic markers involved in the pathology of squamous cell carcinoma of the oral cavity.

In this study, I observed that:

- Concerning smoking patients, there is a change in global methylation and global hydroxymethylation levels.
- The methylation level vas decreased in smokers and former smokers compare with never smokers
- The decrease in methylation level was pronounced in former smokers.
- A significant increase (twice) in overall methylation levels in heavy smokers versus light smokers.
- A considerably lower level of global DNA hydroxymethylation than that of global DNA methylation.
- The smokers had a significantly reduced (twice) level of global DNA hydroxymethylation compared to non-smokers

- The age of the patient might influence the methylation level, while with hydroxymethylation this influence is null.

To eliminate biases, we used only epithelial cells harvested from the jugal mucosa and performed fluorometric enzyme-linked immunosorbent assay (ELISA) analysis, a much more sensitive method of detecting 5-mc and 5-hmc levels.

Comparing our results with those of previous studies, we observed that the characteristics of the samples, the collection method, as well as the analysis method used in this study allow us to detect a more appropriate level of 5-mc and 5-hmc in oral epithelial cells.

Investigation of genome-wide methylation profile in cancer samples versus adjacent normal tissue using Infinium 450K array-based Methylation Assay revealed 1468 loci with significant differences in methylation pattern between cancer and adjacent tissue samples (adjusted p-value ≤ 0.05). The distribution of these differentially methylated regions varied, being located both genetically at the promoter (TSS 200 and TSS 1500), 5'UTR, first exon, body, 3'UTR, and intergenic levels. We also observed that loci showed changes in methylation status distributed on all chromosomes.

After investigating the methylation profile of the samples, four microRNAs (hsa-miR-24-1, hsa-miR-124-3, hsa-miR-769, and hsa-miR-4500) were detected to exhibit hypermethylation at multiple CpG loci in OSCC samples compared to adjacent normal tissue. Within the microRNA, loci are mostly positioned at the promoter level (TSS region) but also in 3'UTR.

Regarding the microARN detected in this study, we observed that:

- Hsa-miR-4500 is the first time it has been detected in OSCC cancer samples,
- Hsa-miR-24-1 has been proposed as an epigenetic marker in OSCC.

Our bioinformatics analysis suggests that methylation-dependent modulation of these microRNA expressions might significantly impact the response pathways to viral infections, the ECM-receptor interaction pathway, protein digestion and absorption, and the AGE-RAGE signaling pathway, which could be relevant for the development, progression, and pathogenesis of OSCC.

These are the first studies in Romania that investigated epigenetic markers in OSCC pathology and the influence of environmental factors on the epigenome in the oral mucosa.

I considered that to understand the OSCC etiopathogenesis and to find a method of prevention and eradication of this neoplasm, more studies are required.

As for research directions, further studies are necessary to:

- Validate the influence of smoking on global DNA methylation and hydroxymethylation in oral epithelial cells.
- Validate the implication of hsa-miR-4500 in OSCC pathology.
- Increase the knowledge about the correlation linking the expression and methylation status of hsa-miR-24-1, hsa-miR-124-3, and hsa-miR-769 in OSCC samples.