Gene network analysis of the transcriptome impact of methylated microRNAs on oral squamous cell carcinoma

Emilia Gabriela Avram[1,2,B,D], Ioana Alexandra Moatar[1,3,D], Viktorian Miok[4,C], Flavia Baderca[5,C], Corina Samoila[3,6,B],
Anda Alexa[3,C], Ioana Nicoleta Andreescu[7,C], Angela Podariu[8,A], Catalin Marian[3,6,E], Ioan Ovidiu Sirbu[3,6,A,D,F]

[1] Doctoral School, Victor Babeș University of Medicine and Pharmacy, Timișoara, Romania

[2] Department of Maxillofacial Surgery, Faculty of Dentistry, "Vasile Goldiş" Western University of Arad, Romania

[3] Department of Biochemistry & Pharmacology, Victor Babeș University of Medicine and Pharmacy, Timișoara, Romania

[4] Department of Informatics & Medical Biostatistics, Victor Babeș University of Medicine and Pharmacy, Timișoara, Romania

[5] Department of Microscopic Morphology, Victor Babeș University of Medicine and Pharmacy, Timișoara, Romania

[6] Center for Complex Network Science, Victor Babeș University of Medicine and Pharmacy, Timișoara, Romania

[7] Department of Microscopic Morphology – Genetics, Center of Genomic Medicine, Victor Babeș University of Medicine and Pharmacy, Timișoara, Romania

[8] Department of Preventive Dentistry, Community and Oral Health, Victor Babeș University of Medicine and Pharmacy, Timișoara, Romania

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation;
D – writing the article; E – critical revision of the article; F – final approval of the article

Address for correspondence
Ioana Alexandra Moatar
E-mail: moatar.alexandra@umft.ro

Abstract

Background. Oral squamous cell carcinoma (OSCC) is one of the most common head and neck squamous cell tumors. MicroRNAs and DNA methylation, as epigenetic mechanisms, regulate the expression of oncogenes and tumor suppressor genes, contributing to the carcinogenic development. However, the current knowledge on the genetic and epigenetic landscape of OSCC is still limited.

Objectives. To assess the transcriptomic impact of microRNAs found to be methylated through Infinium genome-wide methylation profiling of archived OSCC tissues, and to analyze their biological role using gene network analysis.

Materials and methods. We used the Infinium array-based methylation assay to assess the genome-wide methylation status at the single-CpG-site level of DNA purified from archived OSCC tissue samples. After quality control, filtering out poorly performing probes and normalization of data, we identified the differentially methylated microRNA loci. We performed a literature-based analysis of OSCC transcriptomic data to identify the predicted target genes for each microRNA, followed by individual network and pathway enrichment analyses.

Results. The analysis of Infinium methylation array data revealed 1469 differentially hypomethylated loci, 4 of which were of interest, namely hsa-microRNA-124-3, hsa-microRNA-24-1, hsa-microRNA-769, and hsa-microRNA-4500. Network and pathway enrichment analyses revealed multiple pathways modulated through DNA methylation-microRNA expression axes.

Conclusions. We describe the transcriptomic impact of 4 differentially methylated microRNAs in OSCC tissues samples and discuss their role in the pathology of OSCC. These results may contribute to a better understanding of how epigenetic mechanisms such as DNA methylation and microRNAs cooperate to impact the development of OSCC.

Key words: DNA methylation, microRNAs, network analysis, oral squamous cell carcinoma

1232

E.G. Avram et al. Methylated microRNAs in OSCC

Background

Oral squamous cell carcinoma (OSCC) is one of the most common head and neck squamous cell carcinoma (HN-SCC) tumors, with over 350,000 new cases annually.[1] Despite recent therapeutic progress, the cure rate of OSCC is relatively low and closely correlated with the tumor stage at diagnosis. Moreover, in recent years, there has been a change in the incidence of OSCC, with an increased occurrence in young people.[2]

Oral cancer is less studied than other cancers, and there is currently a lack of bona fide diagnostic and prognostic biomarkers. Similarly to other solid cancers,[3–5] OSCC has a multifactorial and multigenic character. Its development is driven by various intricate genetic and epigenetic mechanisms which play important roles in the initiation and development of these tumors.[6]

The DNA methylation and non-coding RNAs are 2 of the most studied epigenetic mechanisms impacting global gene expression.[7] The DNA methylation is essential for several important cellular mechanisms, including regulation of gene expression, differentiation, genome imprinting, and stability.[8] The accumulation of alterations to genome methylation leads to the destabilization of cellular homeostasis. The aberrant DNA hypo- and hypermethylation patterns are critical players in tumorigenesis by promoting the expression of oncogenes and silencing tumor suppressor genes.[9–11]

MicroRNAs are small non-coding RNAs involved in the post-transcriptional regulation of approx. 60% of human genes. Multiple microRNAs have been associated with various aspects of oral cancer biology, including onset, metastasis, recurrence, prognosis, and survival.[12–14] The interest for understanding the regulation of microRNA expression in the context of other epigenetic processes has been growing steadily, and there is a need to understand the global impact of microRNAs on cancer cell transcriptomics.

Network analysis approaches, such as protein–protein interaction networks, have the potential to add biological meaning to large-scale gene datasets such as those obtained from high-throughput RNA sequencing, DNA methylation assays or microRNA target gene analysis.[15] Combining both biological networks and pathway enrichment is meaningful for predicting complex disease mechanisms, disease biomarkers, novel gene interactions, and drug design.[16] An example of this is the integrative analysis of DNA methylation and microRNA expression data that has been conducted for hepatocellular carcinoma.[17] However, a network analysis that integrates these 2 epigenetic mechanisms in the context of OSCC is lacking. Therefore, we investigated the potential interaction between these 2 epigenetic mechanisms, shedding light on unknown mechanisms that might contribute to the development of OSCC.

Objectives

The aim of the current study was to evaluate the impact of differentially methylated microRNAs in the OSCC transcriptome. Therefore, we used Infinium array-based methylation assay to analyze the genome-wide methylation status of microRNA loci in a set of OSCC-archived samples and correlated the results with publicly available microRNA expression data. We retrieved the OSCC transcriptome data from the Gene Expression Omnibus (GEO) database and used miRWalk v. 3.0 prediction algorithms (http://mirwalk.umm.uni-heidelberg.de/) to predict the gene targets of the differentially methylated microRNAs. The commonly predicted gene sets were analyzed further using network clustering coupled to Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis, in order to predict the biological impact of the differentially methylated microRNAs.

Materials and methods

Sample collection

The present case-control study was reviewed and the analysis of the provided formalin-fixed paraffin-embedded (FFPE) tissue samples was approved by the Ethics Committee of Victor Babeș University of Medicine and Pharmacy, Timișoara, Romania (approval No. SCUM TM 9414/30.05.2018). Due to the retrospective design of the current study and patient anonymity, the ethical review board determined that informed consent was not required. We used 13 FFPE samples from male OSCC patients (demographic characteristics are presented in Table 1); all samples were assessed and graded by 2 independent, trained pathologists. Adjacent tissues, both cancerous and normal, were dissected out of every FFPE block and pooled into 4 tumor (T) and 4 normal (N) samples for downstream methylation analysis.

DNA purification

Genomic DNA was extracted from the FFPE tissue samples using the QIAamp DNA FFPE Tissue Kit (Qiagen, Hilden, Germany), according to the manufacturer's instructions. Paraffin was dissolved in 1 mL of xylene and removed. Afterwards, the samples were lysed using 180 µL of Buffer ATL under denaturing conditions with 20 µL of proteinase K digestion. The samples were then incubated at 90°C to reverse the formalin crosslinking. To obtain RNA-free genomic DNA, 2 µL of RNAse A were added. The genomic DNA was bonded to the membrane using a mixture of Buffer AL and ethanol. In the washing steps, the residual contaminants were washed away using freshly reconstituted Buffer AW1 and AW2, according to the manufacturer's instructions. Finally, the DNA

Table 1. Demographic and clinical characteristics of the patients included in study

Characteristic	Number	Percentage (%)
Patients enrolled	13	100
Mean age (SD) [years]	54.7 (8.03)	15.4
Range	41–67	100
Localization		
Retromolar	1	7.69
Tongue	5	38.46
Alveolar	1	7.69
Floor of the mouth	2	15.38
Oral cavity	3	23.07
Jugal	1	7.69
Pathological TNM		
Early (I–II)	9	69.23
Advanced (III–IV)	4	30.77
Stage		
I	2	15.38
II	3	23.07
IV	8	61.53
Recurrence		
Yes	2	66.67
No	11	64.7
Keratinization status		
Yes	11	84.61
No	2	15.38
Cell differentiation		
G2	12	92.30
G3	1	7.70

SD – standard deviation; TNM – tumor-nodule-metastasis staging.

was eluted in 50 µL of nuclease-free water. The quality and purity of the extracted DNA were measured using a spectrophotometer (NanoDrop™; Thermo Fisher Scientific, Wilmington, USA). The DNA purity of the samples, as measured using the 260/280 ratio, had an average of 1.89 ±0.18.

Genome-wide DNA methylation in the well-differentiated OSCC samples

We used the Infinium 450K array-based methylation assay (Illumina Inc., San Diego, USA) to quantitatively assess the genome-wide methylation status at single-CpG-site level, using the DNA purified from the pooled FFPE samples. The 1-µg genomic DNA samples were incubated in sodium bisulfite buffer (Sigma-Aldrich, St. Louis, USA) in a SimpliAmp Thermal Cycler (Applied Biosystems, Waltham, USA) with alternating cycles of 95°C and 60°C. The resultant DNA was bound again to a membrane, washed, incubated in a desulphonation buffer

(Zymo Research, Orange, USA), and eluted with 20 µL of DNA elution buffer (Zymo Research). The quality and purity of the treated DNA was measured using a spectrophotometer (Nanodrop Technologies, Wilmington, USA). The yields were in the range of 50–100 ng/µL.

After bisulfite conversion, whole-genome methylation patterns were assayed using the Infinium® HumanMethylation450 BeadChip Kit (Illumina Inc.), following the manufacturer's protocol. Briefly, the raw data were obtained in intensity data file format (*.idat) by scanning the processed chips using an iScan reader (Illumina Inc.). Each point of methylation data was presented with fluorescent signals from the methylated (M) and unmethylated (U) alleles. The methylation status of each CpG site was measured based on the ratio of fluorescent signals from one allele relative to the sum of both methylated and unmethylated alleles (β value) using GenomicRatioSet (Illumina Inc.). Probes with a detection p-value cutoff >0.05 were excluded from the subsequent analysis.

To minimize the variation between and within the samples, we applied the preprocess quantile normalization methon, and then performed a PCA analysis and generated multidimensional scaling (MDS) plots to visualize and explore the differences between OSCC and normal tissues.

The differentially methylated microRNA gene loci with an adjusted p-value < 0.05 were selected for further analysis.

Statistical analyses

Illumina methylation *.idat files were imported into R programming language (R Foundation for Statistical Computing, Vienna, Austria) using the minfi function. The quality control of the imported samples was evaluated by calculating the detection p-value for every CpG island in each sample, which is indicative for the quality of the signal. Since the samples from our dataset were all derived from male donors, it was not necessary to filter the sites from X and Y chromosomes. Furthermore, we have removed the probes where common single nucleotide polymorphisms (SNPs) affected the CpG islands. Once the data were normalized and filtered, we visualized the results using MDS plots. In the final step, we calculated the β (percentage of methylation) and M ($\log_2$ ratios of the intensities of methylated compared to unmethylated probes) values.

The differentially methylated microRNAs loci were identified by a two-tailed, heteroscedastic Student's t-test, being ranked according to their Bonferroni adjusted p-values (p = 0.05 as a cutoff).

Cross validation with literature analysis

To determine whether the differentially methylated microRNAs retrieved from our genome-wide DNA methylation analysis had been previously validated through

experimental studies, we chose to perform a literature search in PubMed. In contrast to Web of Science and Google Scholar, PubMed is a human-curated database, that retrieves only peer-reviewed journal articles that fit into the searchable subject matter, with reliable sorting, reproducibility and reportable results.[18] The key words combinations used in the PubMed literature search were: microRNAs AND oral cancer, microRNAs AND oral squamous cell carcinoma. We filtered out reviews, meta-analyses and articles that did not relate to OSCC and microRNAs.

Network analysis

Target predictions (3'UTR, 5'UTR and coding sequence (CDS) region) for hsa-miR-24-1, hsa-miR-124-3, hsa-miR-769, and hsa-miR-4500 were computed with the miRWalk v. 3.0[19] machine learning algorithm (Bonferroni adjusted p-value ≤0.05 as a cutoff), using the following criteria: 1) microRNA binding probability >0.99 and 2) microRNAs concurrently predicted by at least 1 other algorithm (miRDB or TargetScan). The microRNA target prediction datasets were cross-referenced with the differentially expressed genes (DEGs) retrieved after GEO2R analysis (Benjamini–Hochberg adjusted false discovery rate (FDR) <0.05) of the GSE138206 OSCC transcriptome dataset. The common genes of microRNA targets that were differentially expressed in OSCC transcriptomes were then subjected to KEGG pathway enrichment analysis and cluster analysis on the STRING platform[20] using the following criteria: adjusted FDR value of 0.05 as cutoff and high confidence interaction score set at 0.700.

Prediction of CpG islands

MethPrimer platform (http://www.urogene.org/cgi-bin/methprimer/methprimer.cgi) was used for designing the bisulfite sequencing polymerase chain reaction (BSP PCR) primers and for predicting the CpG islands associated with the 4 microRNAs, using the 2-kb DNA sequence upstream of the microRNA loci as input data. For BSP

primer selection and CpG island prediction, the general parameters have been set according to the recommendations by Li and Dahiya.[21]

Results

The β values represent the ratio between the methylated probe intensity and the overall intensity. The β-value statistics result in a number between 0 and 1, where 0 indicates that all copies of the CpG islands from analyzed samples are completely unmethylated, and a value of 1 indicates that every copy of CpG sites is methylated.[22] In our case, the β-values results are situated close to 0 and are have higher density for tumor samples in contrast to normal samples. Conversely, the M-values are calculated as the $\log_2$ ratio of methylated compared to unmethylated probes. The M-values close to 0 indicate that CpG sites are approximately half-methylated, negative values mean that there are more unmethylated CpG sites than methylated ones, and positive M-values mean the opposite.[22] In our case, the highest density plot of M-values close to 0 was created for the tumor group in contrast to the normal group. Therefore, the analysis of the Infinium methylation array β-values and M-values indicated that there is a higher degree of hypomethylation globally in the tumor group than in the control group (Fig. 1).

To find the statistically significant genes that encode for microRNAs, we filtered the 440,364 gene loci assayed on the Infinium methylation array based on an adjusted p-value < 0.05. A total of 1469 loci were obtained, of which only 4 proved to encode microRNAs, namely hsa-miR-124-3, hsa-miR-769, hsa-miR-24-1, and hsa-miR-4500. Compared to controls, the microRNA loci from OSCC samples are hypermethylated, with log fold change (logFC) values between 3.21 and 4.99. The genomic locations of all 4 microRNAs are presented in Table 2.

To determine whether our DNA methylation array results are consistent with other published findings, we searched the GEO database the key words "OSCC" and "DNA methylation". We retrieved 9 datasets, of which GSE123781 and
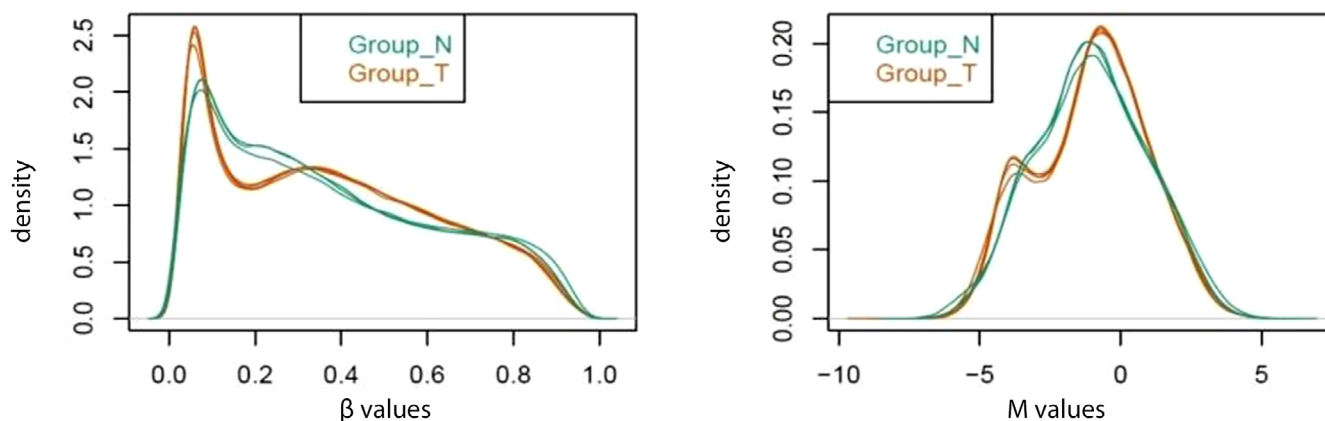

Fig. 1. Distribution of calculated β- and M-values for the oral squamous cell carcinoma (OSCC) tumor and normal groups

Table 2. Differentially methylated microRNAs found in oral squamous cell carcinoma (OSCC) tissues

Chr (strand)	Position	Name	Island name	UCSC ref gene	Gencode name	logFC	adj. p-value
Chr 20(−)	61809724	cg19267861	chr20:61806254-61810867	TSS200	miR124-3; CDH4	4.9	0.0225
Chr 20(−)	61809035	cg02065637	chr20:61806254-61810867	TSS1500	miR124-3; CDH4	3.9	0.0269
Chr 19(+)	46521569	cg08426951	chr19:46518283-46520080	Body; TSS1500	PPP5D15; miR769	3.2	0.0422
Chr 13(+)	88323607	cg26110710	chr13:88323569-88324640	TSS1500	SLITRK5; miR4500	3.9	0.0429
Chr 9(−)	97848222	cg00532885	chr9:97848140-97848525	TSS200; 3'UTR	miR24; C9orf3	3.7	0.0493

TSS – transcription start site; FC – fold change.

Table 3. Common methylated gene loci found in similar Gene Expression Omnibus (GEO) datasets

GSE data	Cases vs. controls	Method	Gene name	adj. p-value	log FC	RangeSTART	RangeEND
GSE 123781	15 vs. 18	Infinium Human Methylation 450K	cg19267861	7.78E-11	0.29	61809724	61809847
			cg26110710	8.13E-03	0.13	88323607	88323730
GSE 87053	11 vs. 10		cg19267861	8.61E-04	0.20	61809724	61809847
			cg02065637	2.37E-02	0.13	61809035	61809158

GSE – list of expression profiles conducted for the experiment (test, controls); FC – fold change.

Table 4. Calculation of CpG island prediction and sequence features for each microRNA

microRNA	Island number	Island size [bp]	Island start	Island end	GC (%)	Obs/Exp ratio
microRNA-124-3	island 1	1855	3	1857	50	0.6
	island 2	175	1862	2036	50	0.6
microRNA-24-1	island 1	186	813	998	50	0.7
	island 2	119	1105	1223	50	0.7
microRNA-4500	island 1	159	55	213	50	0.7
	island 2	146	357	502	50	0.7
	island 3	179	527	705	50	0.7

bp – base pairs; GC – guanine-cytosine; Obs/Exp – observed/expected CpG CG islands ratio.

GSE87053 contained the 4 loci of interest and were further selected for GEO2R analysis. We found that 3 out of the 4 gene loci were significantly hypermethylated, namely cg19267861 (the locus entailing hsa-miR-124-3), cg02065637 (the locus containing hsa-miR-124-3) and cg26110710 (the locus coding for SLITRK5 and hsa-miR-4500) (Table 3).

We used MethPrimer to design BSP PCR primers and predict the presence of CpG islands upstream of genes encoding the 4 microRNAs (Fig. 2, Table 4). The CpG island prediction identified the presence of CpG islands upstream of hsa-miR-124-3 (Fig. 2A) and hsa-miR-769 (Fig. 2B), but not for hsa-miR-4500 and hsa-miR-24-1.

Five sets of BSP primers with their sequences were retrieved for all 4 microRNA molecules, each of them consisting of 2 pairs of forward and reverse primers (Supplementary File 1).

Our PubMed search for microRNAs associated with OSCC led to a total of 189 microRNAs, of which 101 are downregulated, and 88 are upregulated in cancer tissues compared to normal tissues (Supplementary File 2,3). Of note, hsa-miR-24-1 was found to be upregulated,[23] while hsa-miR-769[24] and hsa-miR-124-3[25] were downregulated in OSCC samples compared to controls. To date, no published data link hsa-miR-4500 to oral cancer.

Next, we used miRWalk v. 3.0 to predict the targets of the 4 microRNAs. By filtering out the genes with a binding probability below 0.99, we retrieved 4147 unique predicted targets for hsa-miR-24-1, 6084 predicted targets for hsa-miR-124-3, 743 targets for hsa-miR-4500, and 7256 targets for hsa-miR-769.

To understand the biological significance of these predicted targets in the specific context of oral cancer, we cross-referenced them with the differentially expressed genes in the GSE138206 OSCC dataset (after GEO2R analysis, adjusted p-value < 0.05). The GEO2R analysis of the GSE138206 dataset retrieved 2289 genes with dysregulated expression, of which 1490 were downregulated (FC < −0.538) and 799 were upregulated (FC > 1.0). Since hsa-miR-24-1 has a documented upregulated expression in OSCC,[23] we chose to compare its predicted target genes with the GSE138206 list of downregulated genes, and found 261 putative common targets. The hsa-miR-769[24] and hsa-miR-124-3[25] have both previously been reported to be downregulated in OSCC. Therefore, we compared their

1236

E.G. Avram et al. Methylated microRNAs in OSCC
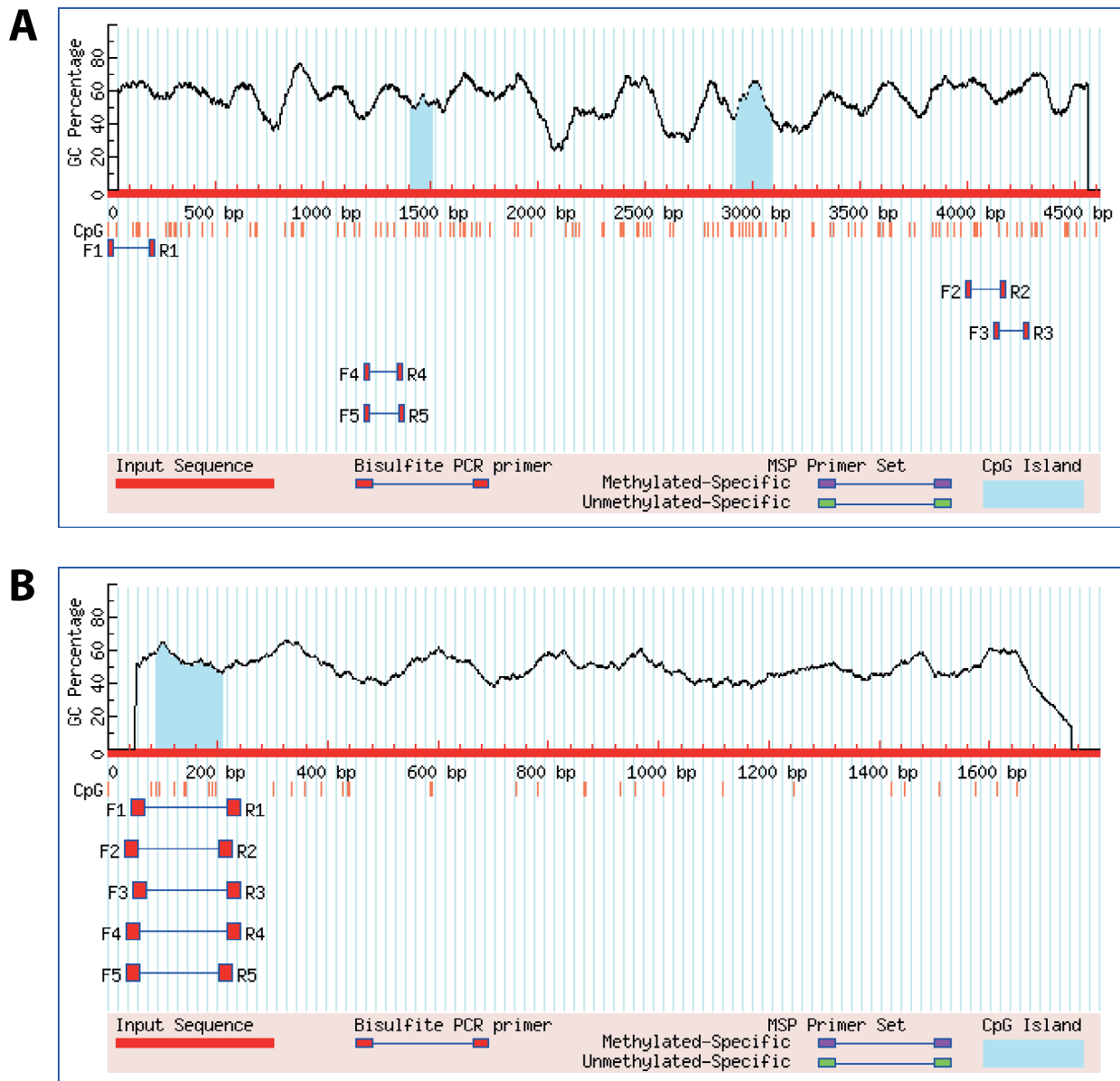
Fig. 2. Visualization of CpG islands predictions results. A. Primers and sequence features such as guanine-cytosine (GC) content percentage, CpG islands (highlighted in light blue) and CpG site for the 2-kb upstream genomic sequence of hsa-microRNA-124-3; B. Primers and sequence features for the 2-kb upstream genomic sequence of hsa-microRNA-769. CpG islands (highlighted in light blue)

predicted targets with the list of upregulated target genes from the GSE138206 dataset and found 274 putative common targets for hsa-miR-124-3 and 298 for hsa-miR-769. To date, there are no published data linking hsa-miR-4500 to OSCC; therefore, we compared its predicted targets with both the upregulated genes (and found 22 common genes) and downregulated genes (and found 51 common genes).

The STRING functional analysis of the hsa-miR-24-1 putative targets led to a protein–protein interaction (PPI) network characterized by 259 nodes, 68 edges and a PPI enrichment p-value of $3.75e^{-12}$, for which no significant signaling pathway could be found after KEGG functional enrichment analysis.

The STRING functional analysis of the hsa-miR-124-3 putative target genes led to a PPI network characterized by 266 nodes, 562 edges and a PPI enrichment p-value below $1.0e^{-16}$. The average local clustering coefficient of the network is 0.359 with an average node degree of 4.23. The cluster analysis using the Markov Cluster Algorithm (MCL) method (inflation parameter of 4) led to the identification of 3 clusters functionally enriched through KEGG pathway analysis (referred to as Cluster 1, 2 and 3) containing 23, 12 and 7 genes, respectively. (Fig. 3, Table 5). Cluster 1 (23 genes) is enriched with genes involved in viral infections, and the local clustering coefficient is high with a value of 0.84 and an average node degree of 10.8. Cluster 2 (local clustering coefficient = 0.769, average node degree = 7.83) includes 12 genes related to protein digestion and absorption, ECM-receptor interaction pathway and focal adhesion.

Cluster 3 contains 7 genes significantly associated with the p53 signaling pathway, with the local clustering coefficient of 0.867 and an average node degree of 2. The high values of the local clustering coefficients suggest that members within these clusters are involved in similar pathways and functions.

The STRING functional analysis of the hsa-miR-769 targets identified a PPI network characterized by 290 nodes, 581 edges, and a PPI enrichment p-value below $1.0e^{-16}$. The average local clustering coefficient of the network was 0.357 and the average node degree was 4.01. The MCL cluster analysis of the network also resulted in 3 clusters functionally enriched in KEGG pathways, containing 25, 13 and 8 genes, respectively (Fig. 4, Table 6). Cluster 1 contains 25 genes that were also enriched in viral infection pathways, with a local clustering coefficient of 0.809 and an average node degree of 10.6. Cluster 2 (local clustering coefficient = 0.79, average node degree = 7.54) includes 13 genes which are most significantly enriched in protein digestion and absorption, ECM receptor interaction and AGE-RAGE signaling pathways. Cluster 3 contains 8 genes significantly enriched in AGE-RAGE signaling pathway, ECM-receptor interaction and cancer-related pathways. The local clustering coefficient for Cluster 3 was 0.893, with an average node degree of 2.5.
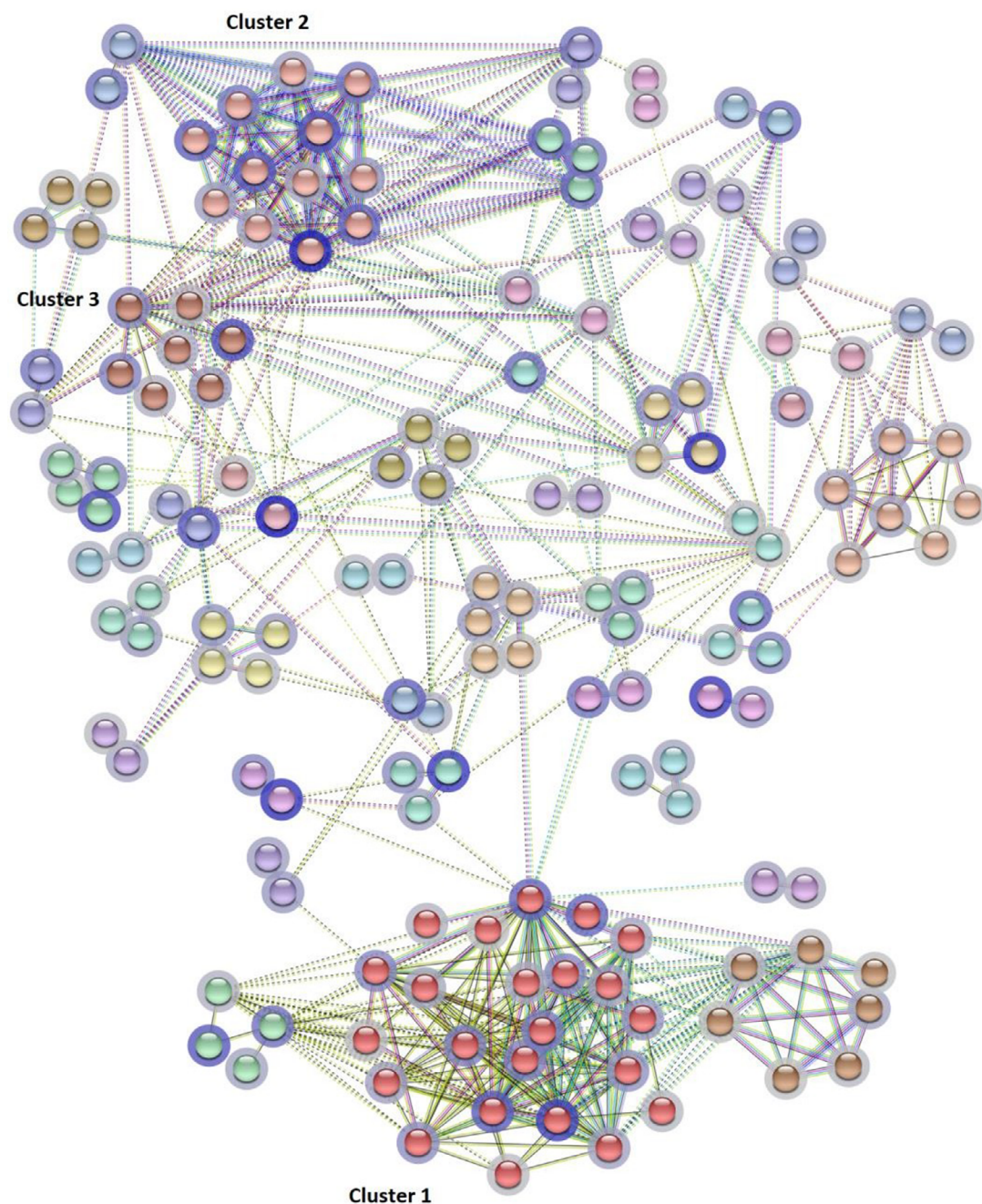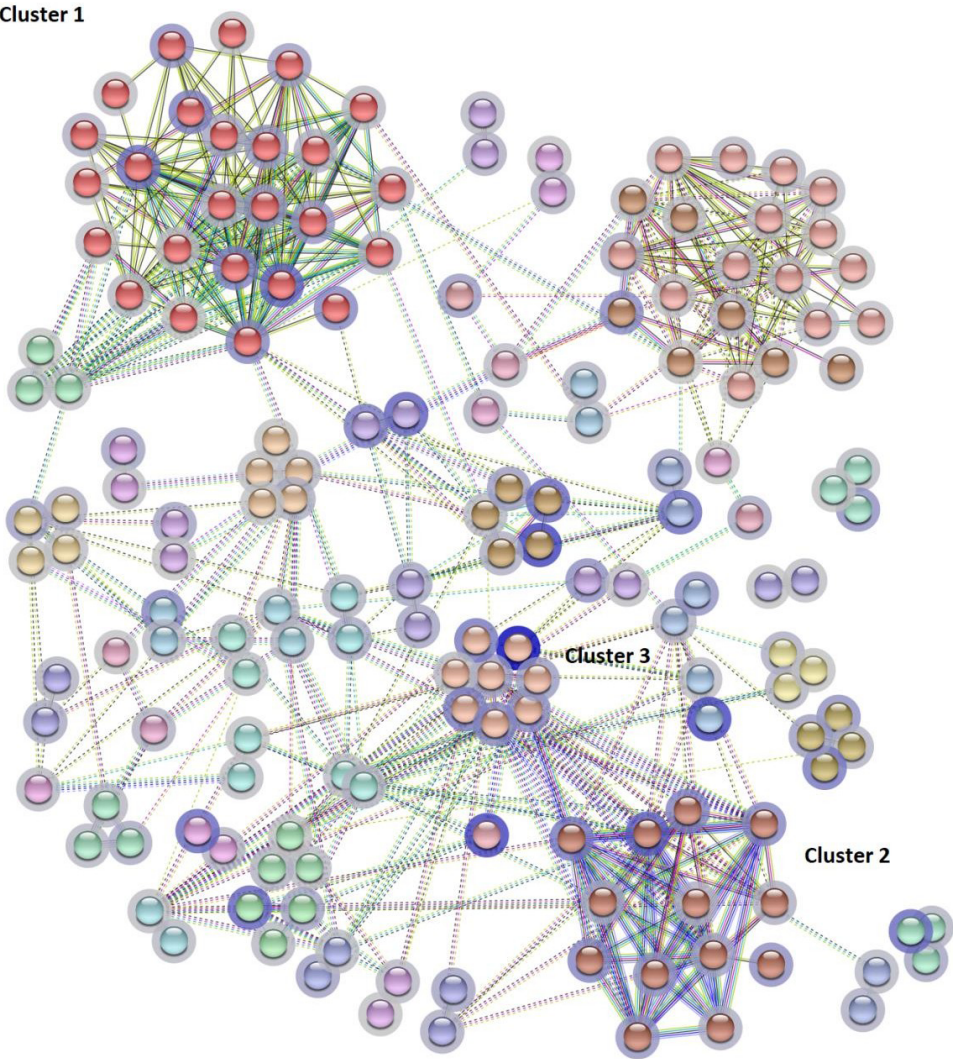

Fig. 3. Cluster analysis (STRING) of the hsa-miR-124-3 target genes in GSE138206 dataset network. The most relevant clusters are the (1) red (Cluster 1), (2) fire brick (Cluster 2) and (3) salmon (Cluster 3). The solid and dotted lines represent connection within the same and different cluster, respectively. Different colors indicate different types of interactions (cyan – databases, pink – experimental, blue – gene co-occurrence, khaki – text mining, black – co-expression, light blue – protein homology). The halo color intensities are based on the fold change values of the submitted genes

1238

E.G. Avram et al. Methylated microRNAs in OSCC

Table 5. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways potentially impacted by hsa-miR-124-3

Cluster	Term ID	Term description	Count in network	Strength	FDR
Cluster 1	hsa05160	hepatitis C	8 of 156	1.64	3.01E-09
	hsa05164	influenza A	8 of 165	1.62	3.01E-09
	hsa05162	measles	7 of 138	1.63	2.62E-08
	hsa05169	Epstein–Barr virus infection	7 of 193	1.49	1.87E-07
	hsa04621	NOD-like receptor signaling pathway	6 of 174	1.47	3.30E-06
	hsa05168	herpes simplex virus1 infection	7 of 479	1.09	5.39E-05
	hsa05165	human papillomavirus infection	4 of 325	1.02	0.0258
	hsa04217	necroptosis	3 of 149	1.23	0.0304
	hsa05161	hepatitis B	3 of 159	1.21	0.0325
	hsa04630	JAK-STAT signaling pathway	3 of 160	1.2	0.0325
Cluster 2	hsa04974	protein digestion and absorption	9 of 100	2.17	2.67E-16
	hsa04512	ECM-receptor interaction	3 of 88	1.74	0.0035
	hsa04510	focal adhesion	3 of 198	1.39	0.0245
Cluster 3	hsa04115	p53 signaling pathway	3 of 72	2.07	0.00063

FDR – false discovery rate.

Fig. 4. Cluster analysis (STRING) of the hsa-miR-769 target genes in GSE138206 dataset network. The most relevant clusters are (1) red (Cluster 1), (2) fire brick (Cluster 2) and (3) salmon (Cluster 3). The solid and dotted lines represent connection within the same and different cluster, respectively. Different colors indicate different types of interactions (cyan – databases, pink – experimental, blue – gene co-occurrence, khaki – text mining, black – co-expression, light blue – protein homology)

Table 6. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways potentially impacted by hsa-microRNA-769

Cluster	Term ID	Term description	Count in network	Strength	FDR
Cluster 1	hsa05169	Epstein–Barr virus infection	8 of 193	1.51	3.34E-08
	hsa05160	hepatitis C	7 of 156	1.55	1.73E-07
	hsa05164	influenza A	7 of 165	1.52	1.73E-07
	hsa05162	Measles	6 of 138	1.53	1.88E-06
	hsa05168	herpes simplex virus 1 infection	8 of 479	1.12	5.40E-04
	hsa04621	NOD-like receptor signaling pathway	5 of 174	1.35	0.00015
	hsa05165	human papillomavirus infection	5 of 325	1.08	0.0025
	hsa05167	Kaposi sarcoma-associated herpes virus infection	4 of 187	1.22	0.0040
	hsa04625	C-type lectin receptor signaling pathway	3 of 102	1.36	0.0118
	hsa04217	necroptosis	3 of 149	1.2	0.0312
	hsa05161	hepatitis B	3 of 159	1.17	0.0341
Cluster 2	hsa04974	protein digestion and absorption	10 of 100	2.18	1.95E-18
	hsa04512	ECM-receptor interaction	3 of 88	1.71	0.0045
	hsa04933	AGE-RAGE signaling pathway in diabetic complications	3 of 98	1.66	0.0045
	hsa05146	amoebiasis	3 of 100	1.65	0.0045
	hsa04611	platelet activation	3 of 122	1.57	0.0047
	hsa04926	relaxin signaling pathway	3 of 128	1.55	0.0047
	hsa04510	focal adhesion	3 of 198	1.36	0.0136
	hsa05165	human papillomavirus infection	3 of 325	1.14	0.0494
Cluster 3	hsa04933	AGE-RAGE signaling pathway in diabetic complications	3 of 98	2.00	1.61E-05
	hsa04512	ECM-receptor interaction	3 of 88	1.92	0.00090
	hsa05222	small cell lung cancer	3 of 92	1.90	0.00090
	hsa05146	amebiasis	3 of 100	1.87	0.00090
	hsa04926	relaxin signaling pathway	3 of 128	1.76	0.0011
	hsa05200	pathways in cancer	4 of 517	1.28	0.0018
	hsa04510	focal adhesion	3 of 198	1.57	0.0028
	hsa05165	human papillomavirus infection	3 of 325	1.35	0.0103
	hsa04151	PI3K-Akt signaling pathway	3 of 350	1.32	0.0114

FDR – false discovery rate.

Due to the limited number of hsa-miR-4500 target genes found in the OSCC transcriptome, the resulting PPI network was very poor and lacked statistical significance. Therefore, this was not pursued further.

Discussion

Oral squamous cell carcinoma (OSCC) is one of the most common types of HNSCC. Therefore, it is critical to gain a better understanding of the epigenetic mechanisms that underlie OSCC initiation, development, evolution, and response to therapy.[26] In patients with OSCC, increased methylation status of microRNA loci has been observed compared to controls, suggesting that this parameter may serve as a predictor for malignant transformations.[27] In the present study, we used archived FFPE samples to explore the OSCC genome-wide methylation profile of microRNA loci.

Our analysis led to the identification of 4 microRNA loci significantly hypermethylated in OSCC tissue compared to normal adjacent tissue, namely hsa-miR-24, hsa-miR-769, hsa-miR-124, and hsa-miR-4500.

The hsa-miR-24 is one of the most heavily studied microRNAs in cancer, being associated with tumor initiation, progression, metastasis, and response to therapy in a wide array of malignant pathologies.[28] In our study, the hsa-miR-24 loci were significantly hypermethylated both at the 3'UTR site and 200 bp upstream of the transcription start site (TSS). There is evidence that if DNA methylation occurs within the 3'UTR region, it can both positively and negatively impact the transcription.[29] Therefore, the hypermethylation status of the hsa-miR-24-1 locus may still correlate with the reported upregulated expression of miR-24-1/-3. Of note, in OSCC, the mature miR-24-1/-3 is overexpressed not only in malignant tissues[23] (where it correlates with the staging of the disease) but also in the patients' saliva, serum and

plasma, qualifying it as a possible diagnostic and predictive biomarker.[30–32]

We also found that the hsa-miR-124-3 locus is methylated both distally (1500 bp upstream, TSS1500) and in close proximity (200 bp upstream, TSS200) to the TSS, which suggests gene silencing. This is also consistent with reports showing the downregulation of hsa-miR-124-3 in an OSCC cell line[25] and animal models of oral cancer.[33] The KEGG pathway enrichment analysis results for the hsa-miR-124-3 gene targets in OSCC were related to viral infections in the most populated gene cluster. Hepatitis C, Epstein–Barr virus, herpes simplex 1 and human papillomavirus are among the most significant viruses associated with this set of genes. The links between the etiologic causes of OSCC and viruses have long been investigated.[34–36] Hepatitis C virus was demonstrated to be involved in the development of oral lichen planus, a chronic inflammatory condition that increases the risk of OSCC.[37] Epstein–Barr virus,[38] herpes simplex virus 1[39] and human papillomavirus[40] have also been investigated in the promotion of the malignant transformation of cells in OSCC. In Cluster 2, pathways such as protein digestion and absorption, and ECM-receptor interaction pathway proved to be the most relevant ones. The presence of ECM-receptor interaction pathway associated with the Cluster 2 gene set confirms previous literature findings reporting the involvement of hsa-miR-124-3 in the ECM receptor interaction pathway through its direct repression of ITGB1 in OSCC cell lines.[25]

Regarding hsa-miR-769, we found the hypermethylation of both 1500 bp upstream of the TSS and within the miR-locus. Our data are consistent with the documented down-regulation of hsa-miR-769 expression in OSCC tissues, since hypermethylation of its TSS may block the transcription.[24] The KEGG pathway enrichment analysis of Cluster 1 identified Epstein–Barr virus infection, hepatitis C and influenza A as the most significant pathways associated with hsa-miR-769 target genes that were differentially expressed in OSCC. In this case, the ECM-receptor interaction was also retrieved from among the significant pathways in Clusters 2 and 3. Interestingly, this pathway has previously been predicted to be associated with OSCC.[41] A new entry for this set of genes is the AGE-RAGE signaling pathway retrieved for Clusters 2 and 3. This pathway was initially reported to be responsible for the initiation of diabetic complications.[42] The receptor for advanced glycation end products (RAGE) molecule, a constituent of this pathway, has been reported to be a key factor that accelerates tumor progression and metastasis in various malignancies.[43–45] Regarding OSCC, there have been several investigations that aimed to evaluate the role and expression of RAGE receptor.[46–48] In a study that evaluated 74 OSCC samples, the RAGE receptor was reported to be a prognostic factor, and was significantly associated with the depth of invasion and local recurrence.[46] Moreover, a particular role of RAGE in OSCC pathogenesis was confirmed in an in vitro study,

where this receptor proved to be activated in a paracrine manner by the extracellular molecule high-mobility group box 1 (HMGB1) that was secreted by the OSCC cells. This activation promotes tumor progression and stimulates bone destruction.[47] Furthermore, RAGE proved to be significantly decreased by evodiamine, a novel anti-tumor drug, inhibiting proliferation, invasion and angiogenesis of OSCC both in vitro and in vivo.[48] Finally, the gene that encodes the RAGE receptor is predicted to be one of the hsa-miR-769-5p targets both in CDS and 5'UTR positions, according to the miRWalk v. 3.0 database.[19]

Finally, for hsa-miR-4500, there is no evidence so far concerning its expression in OSCC. When performing the KEGG pathway functional enrichment analysis for the commonly predicted target genes of this microRNA, there was no significant pathway associated with it due to a limited number of resulting genes.

Limitations

Due to financial limitations, we were not able to validate our methylation results in additional independent cohorts. Caution must be taken given the small sample size and the fact that the findings might not be entirely transferable to a larger population. Further investigations are needed to validate the correlation between the expression and methylation status of hsa-miR-24-1, hsa-miR-124-3 and hsa-miR-769 in OSCC samples through quantitative real-time (qRT)-PCR.

Conclusions

Returning to the study's central question, we suggest that, in stark contrast with the overall genome DNA hypomethylation, microRNA loci could show increased methylation status, which may translate into differential expression of their mature forms. Of note, one of the identified microRNAs has been proposed as an OSCC biomarker, given its presence in multiple biological fluids. Our bioinformatics analysis also suggests that methylation-dependent modulation of microRNA expression could significantly impact the response pathways to viral infections, the ECM-receptor interaction pathway and AGE-RAGE signaling pathway. These may be relevant for the development, progression and pathogenesis of OSCC.

Supplementary data

The supplementary Tables are available at https://doi.org/10.5281/zenodo.6759729.

The package contains the following files:

Supplementary File 1. Primer picking results for bisulfite sequencing PCR (BSP);

Supplementary File 2. Downregulated microRNAs curated from literature analysis;

Supplementary File 3. Upregulated microRNAs curated from literature analysis.

The references used for this supplementary data are listed at the end of the files.

ORCID iDs
Emilia Gabriela Avram https://orcid.org/0000-0002-2808-5566
Ioana Alexandra Moatar https://orcid.org/0000-0001-7756-8823
Viktorian Miok https://orcid.org/0000-0003-1018-2111
Flavia Baderca https://orcid.org/0000-0002-5755-179X
Corina Samoila https://orcid.org/0000-0003-2468-6564
Anda Alexa https://orcid.org/0000-0003-4364-0254
Ioana Nicoleta Andreescu https://orcid.org/0000-0002-2647-4234
Angela Podariu https://orcid.org/0000-0003-0505-0018
Catalin Marian https://orcid.org/0000-0002-7749-1384
Ioan Ovidiu Sirbu https://orcid.org/0000-0003-1618-9656

References

1. Pires FR, Ramos AB, de Oliveira JBC, Tavares AS, da Luz PSR, dos Santos TCRB. Oral squamous cell carcinoma: Clinicopathological features from 346 cases from a single Oral Pathology Service during an 8-year period. *J Appl Oral Sci*. 2013;21(5):460–467. doi:10.1590/1679-775720130317
2. Abdulla R, Adyanthaya S, Kini P, Mohanty V, D'Souza N, Subbannayya Y. Clinicopathological analysis of oral squamous cell carcinoma among the younger age group in coastal Karnataka, India: A retrospective study. *J Oral Maxillofac Pathol*. 2018;22(2):180–187. doi:10.4103/jomfp.JOMFP_16_18
3. Garcia-Martinez L, Zhang Y, Nakata Y, Chan HL, Morey L. Epigenetic mechanisms in breast cancer therapy and resistance. *Nat Commun*. 2021;12(1):1786. doi:10.1038/s41467-021-22024-3
4. Isik A, Isik N, Kurnaz E. Complete breast autoamputation: Clinical image. *Breast J*. 2020;26(11):2265–2266. doi:10.1111/tbj.14072
5. Herceg Z, Paliwal A. Epigenetic mechanisms in hepatocellular carcinoma: How environmental factors influence the epigenome. *Mutat Res*. 2011;727(3):55–61. doi:10.1016/j.mrrev.2011.04.001
6. Magic Z, Supic G, Brankovic-Magic M, Jovic N. DNA Methylation in the pathogenesis of head and neck cancer. In: Dricu A, ed. *Methylation: From DNA, RNA and Histones to Diseases and Treatment*. InTech; 2012: 185–216. doi:10.5772/51169
7. Lopez-Serra P, Esteller M. DNA methylation-associated silencing of tumor-suppressor microRNAs in cancer. *Oncogene*. 2012;31(13): 1609–1622. doi:10.1038/onc.2011.354
8. Moore LD, Le T, Fan G. DNA methylation and its basic function. *Neuropsychopharmacology*. 2013;38(1):23–38. doi:10.1038/npp.2012.112
9. Alam H, Kundu ST, Dalal SN, Vaidya MM. Loss of keratins 8 and 18 leads to alterations in α6β4-integrin-mediated signalling and decreased neoplastic progression in an oral-tumour-derived cell line. *J Cell Sci*. 2011;124(Pt 12):2096–2106. doi:10.1242/jcs.073585
10. Zargoun IM, Bingle L, Speight PM. DNA ploidy and cell cycle protein expression in oral squamous cell carcinomas with and without lymph node metastases. *J Oral Pathol Med*. 2017;46(9):738–743. doi:10.1111/jop.12554
11. Kulis M, Esteller M. DNA methylation and cancer. *Adv Genet*. 2010; 70:27–56. doi:10.1016/B978-0-12-380866-0.60002-2
12. Li X, Fan Q, Li J, Song J, Gu Y. miR-124 down-regulation is critical for cancer associated fibroblasts-enhanced tumor growth of oral carcinoma. *Exp Cell Res*. 2017;351(1):100–108. doi:10.1016/j.yexcr.2017.01.001
13. Qiu H, Huang Z, Huang Z, He Z, You S. miR-22 regulates cell invasion, migration and proliferation in vitro through inhibiting CD147 expression in tongue squamous cell carcinoma. *Arch Oral Biol*. 2016;66:92–97. doi:10.1016/j.archoralbio.2016.02.013
14. Hedbäck N, Jensen DH, Specht L, et al. miR-21 expression in the tumor stroma of oral squamous cell carcinoma: An independent biomarker of disease free survival. *PLoS One*. 2014;9(4):e95193. doi:10.1371/journal.pone.0095193
15. Sanchez R, Mackenzie SA. Integrative network analysis of differentially methylated and expressed genes for biomarker identification in leukemia. *Sci Rep*. 2020;10(1):2123. doi:10.1038/s41598-020-58123-2
16. Liu L, Wei J, Ruan J. Pathway enrichment analysis with networks. *Genes (Basel)*. 2017;8(10):246. doi:10.3390/genes8100246
17. Varghese RS, Barefoot ME, Jain S, et al. Integrative analysis of DNA methylation and microRNA expression reveals mechanisms of racial heterogeneity in hepatocellular carcinoma. *Front Genet*. 2021;12:708326. doi:10.3389/fgene.2021.708326
18. Ossom Williamson P, Minter CIJ. Exploring PubMed as a reliable resource for scholarly communications services. *J Med Libr Assoc*. 2019;107(1):16–29. doi:10.5195/jmla.2019.433
19. Sticht C, De La Torre C, Parveen A, Gretz N. miRWalk: An online resource for prediction of microRNA binding sites. *PLoS One*. 2018; 13(10):e0206239. doi:10.1371/journal.pone.0206239
20. Szklarczyk D, Franceschini A, Wyder S, et al. STRING v10: Protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res*. 2015;43(Database issue):D447–D452. doi:10.1093/nar/gku1003
21. Li LC, Dahiya R. MethPrimer: Designing primers for methylation PCRs. *Bioinformatics*. 2002;18(11):1427–1431. doi:10.1093/bioinformatics/18.11.1427
22. Du P, Zhang X, Huang CC, et al. Comparison of beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics*. 2010;11:587. doi:10.1186/1471-2105-11-587
23. Lin SC, Liu CJ, Lin JA, Chiang WF, Hung PS, Chang KW. miR-24 up-regulation in oral carcinoma: Positive association from clinical and in vitro analysis. *Oral Oncol*. 2010;46(3):204–208. doi:10.1016/j.oraloncology.2009.12.005
24. Zhou Y, Xu XM, Feng Y. miR-769-5p inhibits cancer progression in oral squamous cell carcinoma by directly targeting JAK1/STAT3 pathway. *Neoplasma*. 2020;67(3):528–536. doi:10.4149/neo_2020_190703N582
25. Hunt S, Jones AV, Hinsley EE, Whawell SA, Lambert DW. MicroRNA-124 suppresses oral squamous cell carcinoma motility by targeting ITGB1. *FEBS Lett*. 2011;585(1):187–192. doi:10.1016/j.febslet.2010.11.038
26. Blatt S, Krüger M, Ziebart T, et al. Biomarkers in diagnosis and therapy of oral squamous cell carcinoma: A review of the literature. *J Craniomaxillofac Surg*. 2017;45(5):722–730. doi:10.1016/j.jcms.2017.01.033
27. Dang J, Bian YQ, Sun JY, et al. MicroRNA-137 promoter methylation in oral lichen planus and oral squamous cell carcinoma. *J Oral Pathol Med*. 2013;42(4):315–321. doi:10.1111/jop.12012
28. Wang H, Chen C, Ding K, Zhang W, Hou J. MiR-24-3p as a prognostic indicator for multiple cancers: From a meta-analysis view. *Biosci Rep*. 2020;40(12):BSR20202938. doi:10.1042/BSR20202938
29. Rauluseviciute I, Drabløs F, Rye MB. DNA hypermethylation associated with upregulated gene expression in prostate cancer demonstrates the diversity of epigenetic regulation. *BMC Med Genomics*. 2020;13(1):6. doi:10.1186/s12920-020-0657-6
30. Zhao J, Hu C, Chi J, et al. miR-24 promotes the proliferation, migration and invasion in human tongue squamous cell carcinoma by targeting FBXW7. *Oncol Rep*. 2016;36(2):1143–1149. doi:10.3892/or.2016.4891
31. He L, Ping F, Fan Z, et al. Salivary exosomal miR-24-3p serves as a potential detective biomarker for oral squamous cell carcinoma screening. *Biomed Pharmacother*. 2020;121:109553. doi:10.1016/j.biopha.2019.109553
32. Karimi A, Bahrami N, Sayedyahossein A, Derakhshan S. Evaluation of circulating serum 3 types of microRNA as biomarkers of oral squamous cell carcinoma: A pilot study. *J Oral Pathol Med*. 2020;49(1):43–48. doi:10.1111/jop.12959
33. Yu T, Wang XY, Gong RG, et al. The expression profile of microRNAs in a model of 7,12-dimethyl-benz[a]anthrance-induced oral carcinogenesis in Syrian hamster. *J Exp Clin Cancer Res*. 2009;28(1):64. doi:10.1186/1756-9966-28-64
34. Sugerman P, Shillitoe E. The high risk human papillomaviruses and oral cancer: Evidence for and against a causal relationship. *Oral Dis*. 2008;3(3):130–147. doi:10.1111/j.1601-0825.1997.tb00025.x
35. Steele C, Shillitoe EJ. Viruses and oral cancer. *Crit Rev Oral Biol Med*. 1991;2(2):153–175. doi:10.1177/10454411910020020201
36. Metgud R, Astekar M, Verma M, Sharma A. Role of viruses in oral squamous cell carcinoma. *Oncol Rev*. 2012;6(2):e21. doi:10.4081/oncol.2012.e21
37. Nagao Y, Sata M, Itoh K, Tanikawa K, Kameyama T. Quantitative analysis of HCV RNA and genotype in patients with chronic hepatitis C accompanied by oral lichen planus. *Eur J Clin Invest*. 1996;26(6):495–498. doi:10.1046/j.1365-2362.1996.167314.x
38. Yen CY, Lu MC, Tzeng CC, et al. Detection of EBV infection and gene expression in oral cancer from patients in Taiwan by microarray analysis. *J Biomed Biotechnol*. 2009;2009:904589. doi:10.1155/2009/904589

39. Niller HH, Wolf H, Minarovits J. Viral hit and run-oncogenesis: Genetic and epigenetic scenarios. *Cancer Lett*. 2011;305(2):200–217. doi:10.1016/j.canlet.2010.08.007

40. Kingsley K, Johnson D, O'Malley S. Transfection of oral squamous cell carcinoma with human papillomavirus-16 induces proliferative and morphological changes in vitro. *Cancer Cell Int*. 2006;6:14. doi:10.1186/1475-2867-6-14

41. Zhang G, Bi M, Li S, Wang Q, Teng D. Determination of core pathways for oral squamous cell carcinoma via the method of attract. *J Can Res Ther*. 2018;14(Supplement):S1029–S1034. doi:10.4103/0973-1482.206868

42. Kay AM, Simpson CL, Stewart JA. The role of AGE/RAGE signaling in diabetes-mediated vascular calcification. *J Diabetes Res*. 2016;2016:6809703. doi:10.1155/2016/6809703

43. Wang D, Li T, Ye G, et al. Overexpression of the receptor for advanced glycation endproducts (RAGE) is associated with poor prognosis in gastric cancer. *PLoS One*. 2015;10(4):e0122697. doi:10.1371/journal.pone.0122697

44. Kuniyasu H, Chihara Y, Takahashi T. Co-expression of receptor for advanced glycation end products and the ligand amphoterin associates closely with metastasis of colorectal cancer. *Oncol Rep*. 2003;10(2):445–448. doi:10.3892/or.10.2.445

45. Ishiguro H, Nakaigawa N, Miyoshi Y, Fujinami K, Kubota Y, Uemura H. Receptor for advanced glycation end products (RAGE) and its ligand, amphoterin are overexpressed and associated with prostate cancer development. *Prostate*. 2005;64(1):92–100. doi:10.1002/pros.20219

46. Sasahira T, Kirita T, Bhawal UK, et al. Receptor for advanced glycation end products (RAGE) is important in the prediction of recurrence in human oral squamous cell carcinoma. *Histopathology*. 2007;51(2):166–172. doi:10.1111/j.1365-2559.2007.02739.x

47. Sakamoto Y, Okui T, Yoneda T, et al. High-mobility group box 1 induces bone destruction associated with advanced oral squamous cancer via RAGE and TLR4. *Biochem Biophys Res Commun*. 2020;531(3):422–430. doi:10.1016/j.bbrc.2020.07.120

48. Ren L, Lou Y, Sun M. The anti-tumor effects of evodiamine on oral squamous cell carcinoma (OSCC) through regulating advanced glycation end products (AGE)/receptor for advanced glycation end products (RAGE) pathway. *Bioengineered*. 2021;12(1):5985–5995. doi:10.1080/21655979.2021.1972082