

16. Ashour M. H., Hajjar W. M., Ishaq M., Alamassi M., Saleh W. I. [et al.] Pulmonary hydatid cysts: the naturally occurring models for rupture. *Asian Cardiovascular and Thoracic Annals*. 2016;24(7):670-675.
<https://doi.org/10.1177/0218492316658374>
17. Kocaman O. H., Günendi T., Dere O., Dörterler M. E., Boleken M. E. Pulmonary Hydatid Cyst in Children: A Single-Institution Experience. *Cureus*. 2022;14(7):e26670.
<https://doi.org/10.7759/cureus.26670>

About authors:

Aydemirov Artur Nasirovich, MD, PhD, Professor, Head of the Department of Hospital Surgery;
tel.: +79624479697; e-mail: aaidemirov@mail.ru; ORCID: 0000-0002-3320-8436

Minaev Sergey Viktorovich, MD, PhD, Professor, Head of the Department of Pediatric Surgery;
tel.: +79624507653; e-mail: sminaev@yandex.ru; ORCID: 0000-0002-8405-6022

Rubanov Maria Fedorovna, postgraduate student of the Department of Pediatric Surgery;
tel.: +79289232328; e-mail: maryrubanova@mail.ru; ORCID: 0000-0001-5168-6004

Grigorova Alina Nikolaevna, PhD, Assistant of the Department of Surgical Diseases of Children;
tel.: +79633877244; e-mail: alina.mashchenko@mail.ru; ORCID: 0000-0001-5020-232X

Gerasimenko Igor Nikolaevich, MD, Associate Professor of the Department of Pediatric Surgery;
tel.: +79187704217; e-mail: igor9551@yandex.ru; ORCID: 0000-0003-3003-612X

Timofeev Sergey Ivanovich, PhD, Associate Professor of the Department of Surgery;
tel.: +79148537178; e-mail: timofeev_si@mail.ru; ORCID: 0000-0002-5808-0686

© Group of authors, 2023
UDC 616.311-002-006:575.113
DOI – <https://doi.org/10.14300/mnnc.2023.18034>
ISSN-2073-8137

ANALYSIS OF THE FUNCTION OF THE GENES WITH THE HIGHEST NUMBER OF GERMINAL MUTATIONS IN PATIENTS WITH LEUKOPLAKIA AND CANCER OF THE ORAL MUCOSA

Karpuk N. A.¹, Rubnikov S. P.², Mazur O. Ch.³, Zhyltsov I. V.¹,
Karpuk I. Yu.¹, Sirak S. V.³, Shchetinin E. V.³, Mikhalenko A. P.⁴

¹ Vitebsk State Medical University, Republic of Belarus

² Belarusian State Medical University, Minsk, Republic of Belarus

³ Stavropol State Medical University, Russian Federation

⁴ Institute of Genetics and Cytology of the Belarusian National Academy of Sciences, Minsk, Republic of Belarus

АНАЛИЗ ФУНКЦИИ ГЕНОВ С НАИБОЛЬШИМ КОЛИЧЕСТВОМ ГЕРМИНАЛЬНЫХ МУТАЦИЙ У ПАЦИЕНТОВ ПРИ ЛЕЙКОПЛАКИИ И РАКЕ СЛИЗИСТОЙ ОБОЛОЧКИ РОТОВОЙ ПОЛОСТИ

Н. А. Карпук¹, С. П. Рубникович², О. Ч. Мазур³, И. В. Жильцов¹,
И. Ю. Карпук¹, С. В. Сирак³, Е. В. Щетинин³, Е. П. Михаленко⁴

¹ Витебский государственный медицинский университет, Республика Беларусь

² Белорусский государственный медицинский университет, Минск, Республика Беларусь

³ Ставропольский государственный медицинский университет, Российская Федерация

⁴ Институт генетики и цитологии НАН Беларуси, Минск, Республика Беларусь

To date, a small fraction of the mutations associated with a high risk of developing neoplasms of the oral mucosal (OM) have been described. An important consideration is the determination of germ genetic variants related to the development of leukoplakia (LOM) and squamous cell carcinoma (SCCOM). The function of the genes with the highest number of bacterial mutations was carried out in 24 patients with LEO and 24 patients with SCCOM.

It was found that the structure of the identified mutations was dominated by variants of the MUC3A, MUC4, MUC12, and MUC16 genes responsible for the synthesis of the mucin glycoprotein family (44.7 % and 41.2 % of germinal mutations in patients with LOM and SCCOM, respectively). Insufficient production or a decrease in the functional activity of mucins is a trigger factor for both the development of keratosis and malignant degeneration of epitheliocytes. An infrequent germline

mutation p.K416E of the CHEK2 gene, found in 29.2 % of patients with LOM and 25 % of patients with SCCOM, is associated with the development of both diseases with a high probability (RR 822.67 in LOM and RR 705.15 in SCCOM). Thus, the obtained data contribute to developing test systems PCR and NGS to predict the course of these diseases and personalize therapy.

Keywords: genes, germinal mutations, leukoplakia, cancer, oral mucosa

К настоящему моменту описана малая часть мутаций, ассоциированных с высоким риском развития новообразований слизистой оболочки ротовой полости (СОРП). Важным является определение герминальных генетических вариантов, ассоциированных с развитием лейкоплакий (ЛСОРП) и плоскоклеточного рака (ПРСОРП). Проведен анализ функции генов с наибольшим количеством герминальных мутаций у 24 пациентов с ЛСОРП и у 24 больных с ПРСОРП. Установлено, что в структуре выявленных мутаций преобладают варианты генов MUC3A, MUC4, MUC12 и MUC16, ответственных за синтез семейства гликопротеидов-муцинов (44,7 % и 41,2 % герминальных мутаций у пациентов с ЛСОРП и с ПРСОРП соответственно). Недостаточная продукция либо снижение функциональной активности муцинов являются триггерным фактором как развития кератоза, так и злокачественного перерождения эпителиоцитов. Нечастая герминальная мутация p.K416E гена CHEK2, выявленная у 29,2 % пациентов с ЛСОРП и у 25 % пациентов с ПРСОРП, с высокой вероятностью (ОР 822,67 при ЛСОРП и ОР 705,15 при ПРСОРП) ассоциирована с развитием обоих заболеваний. Таким образом, полученные данные способствуют разработке ПЦР и NGS-тест-систем для прогнозирования течения данных заболеваний и персонализации проводимой терапии.

Ключевые слова: гены, герминальные мутации, лейкоплакия, рак, слизистая оболочка ротовой полости

For citation: Karpuk N. A., Rubnikov S. P., Mazur O. Ch., Zhyltsov I. V., Karpuk I. Yu., Sirak S. V., Shchetin E. V., Mikhaleenko A. P. ANALYSIS OF THE FUNCTION OF THE GENES WITH THE HIGHEST NUMBER OF GERMINAL MUTATIONS IN PATIENTS WITH LEUKOPLAKIA AND CANCER OF THE ORAL MUSCOSA. *Medical News of North Caucasus*. 2023;18(2):155-161. DOI – <https://doi.org/10.14300/mnnc.2023.18034>

Для цитирования: Карпук Н. А., Рубников С. П., Мазур О. Ч., Жильцов И. В., Карпук И. Ю., Сирак С. В., Щетинин Е. В., Михаленко А. П. АНАЛИЗ ФУНКЦИИ ГЕНОВ С НАИБОЛЬШИМ КОЛИЧЕСТВОМ ГЕРМИНАЛЬНЫХ МУТАЦИЙ У ПАЦИЕНТОВ ПРИ ЛЕЙКОПЛАКИИ И РАКЕ СЛИЗИСТОЙ ОБОЛОЧКИ РОТОВОЙ ПОЛОСТИ. *Медицинский вестник Северного Кавказа*. 2023;18(2):155-161. DOI – <https://doi.org/10.14300/mnnc.2023.18034>

CI – confidence interval
DNA – deoxyribo nucleic acid
LOM – leukoplakia of the oral mucosa
NGS – Next-Generation Sequencing
OM – oral mucosa

OR – odds ratio
PCR – polymerase chain reaction
RR – risk ratio
SCCOM – squamous cell carcinoma of the oral mucosa
SNP – single nucleotide polymorphism

To date, only a tiny fraction of mutations associated with a high risk of developing oral mucosa disease (OM) have been described. The primary laboratory standard for diagnosing dysplasia OM is the histological method, which does not allow you to track the onset of malignant degeneration, is invasive, and has a low rate of diagnostic tests [1]. In this regard, it seems relevant to study the frequency of specific mutations among patients with OM diseases (the so-called «genetic landscape» of OM diseases). There are relatively few studies on this topic. Thus, there are indications that the formation of OM leukoplakia (LOM) may be associated with mutations in the DKC1 gene; such LOM is congenital [2]. Another study noted that mutations in the BRCA1 and BRCA2 genes are often detected in the tissue of LOM [3]. It was pointed out that spotted LOM may be associated with pathogenic variants of the NFKB gene [4]. Various studies have also noted that LOM may be related to mutations in the TP53 gene [5], p16INK4a and p14ARF genes [6], the CTC1 gene (also associated with congenital forms of LOM) [7], etc. In general, it draws attention is drawn to the lack of studies on the molecular genetic pathogenesis of LOM, the small sample size in most studies, as well as the diversity and ambiguity of the interpretation of the results.

The number of studies devoted to the molecular genetics of OM squamous cell carcinoma (SCCOM) is also tiny. According to the published results of such studies, SCCOM may be associated with mutations in the genes of the NOTCH family [8], MCM2 (with concomitant increased expression of this gene) [9], TP53 (a pathogenic mutation of this gene TP53Arg72Pro is described) [10], in the FBXL5, UGT2B15, UGT2B28,

KANSL1, GSTT1, and DUSP22 genes [11], in the RAS gene family (Haras, Kras, and N-ras) [12], in the FAT1 and COL9A1 genes (genetic variants rs28647489 and rs550675, respectively) [13]. The same deficiencies characterize these studies as those of NCN genetics – they are few and scattered, the number of patients studied is small, and the results are contradictory.

It can be assumed that there are regional features of genetic variants associated with the development of LOM and SCCOM. Knowledge of such options would allow the development of PCR and NGS testing systems to identify clinically relevant genetic options, which in turn would expand and complement existing protocols to assist patients with OM diseases, first of all, in terms of early diagnosis of these diseases and prediction of their course and outcome. In addition, the analysis of individual mutations in genes responsible for different degrees of epithelial dysplasia OM will allow for to choice of personalized therapy modes for patients, avoid the prescription of obviously ineffective drugs, and make it possible to achieve maximum treatment effectiveness.

The aim of the study was to analyze the function of genes with the highest number of germinal (congenital) mutations in patients with LOM and SCCOM.

Material and Methods. In total, the study included 24 patients with a morphologically verified diagnosis of LOM with grade 1 squamous intraepithelial neoplasia of the epithelium (fifteen men, nine women). The average age of patients was 59 years (from 42 until 72 years old, 95 % CI: 57–65 years). In all cases, there was a flat form of LOM, the most common in the population.

Also included in the study were 24 patients with an established diagnosis of SCCOM (13 men, 11 women). The average age of patients was 60.5 years (from 38 to

75 years old, 95 % CI: 55–65 years). In all cases, there was a primary tumor; also, 100 % of patients were diagnosed with SCCOM.

During the study, 24 blood samples from patients with LOM and 24 blood samples from patients with SCCOM were collected. DNA was isolated using the QIAamp DNA FFPE Tissue Kit (Qiagen, Germany). Peripheral blood from a vein was taken from the cubital vein in the morning on an empty stomach into 10 ml vacuum tubes with K2 EDTA, after which plasma was obtained from it by centrifugation, which was subjected to deep freezing for subsequent storage ($t = -80^{\circ}\text{C}$).

All operations for preparing DNA libraries for sequencing are performed step by step in strict accordance with the instructions for use supplied by the manufacturer (Illumina, Inc., USA) to the Illumina Nextera DNA Exome reagent kit for exome sequencing [14, 15].

Whole exome sequencing of samples was performed to look for congenital (germinal) genetic variants («mutations») that are highly likely to be associated with the development of LOM and SCCOM. These genetic variants are present in 100 % of body cells, which is why they can be detected in leukocyte DNA, and 20-fold read coverage of ≥ 85 % of sequenced sequences provided by the NexteraDNA Exome kit (Illumina, Inc., USA) is sufficient for their reliable identification and documentation.

Bioinformatics analysis of results of exome DNA sequencing using specialized software packages Illumina BaseSpace and Galaxy Project and following current methodological recommendations [16, 17].

Statistical data processing was performed using specialized software packages STATISTICA 12(StatSoft Inc., USA) and MedCalc 18.9.1 (MedCalcSoftware Ltd., Belgium). The central trend and spread of the values of the analyzed quantitative indicators were described in the form of median-quartile characteristics: median, 25th, and 75th quartiles. Comparison of categorical variables was performed using criterion χ^2 , and Fisher's exact test, and the detection of statistical significance of differences in quantitative characteristics was carried out using the Mann – Whitney U-test. Spearman's correlation analysis (Spearman's Rho), as well as logistic regression analysis. The regression analysis included indicators with a significance level of $p < 0.05$. To assess the influence of individual genetic variants on the probability of developing the pathology under study, odds ratios (OR) and risk ratios (RR), as well as their 95 % confidence intervals (CI), were calculated. In all cases, the revealed patterns were con-

sidered statistically significant at a significance level of $p < 0.05$, while the optimal level of significance, generally recognized among bioinformaticians and unambiguously indicating the presence of a relationship between a genetic variant and a phenotype, was $p \leq 5 \times 10^{-8}$ [18].

Results and Discussion. Bioinformatics analysis of sequencing results identified genetic variants («mutations/polymorphisms») in almost every one of the 24,918 human exome genes that can be sequenced using the Illumina Nextera DNA Exome kit. The total number of identified unique genetic variants was 124475 in the group of patients with LOM and 152767 in the group of patients with SCCOM (average of 11688 mutant genes and ≈ 33600 genetic variants in each patient studied). At the same time, the vast majority of identified variants of nucleotide sequences had no clinical significance, significantly affecting neither protein synthesis nor protein function.

However, most variants of nucleotide sequences have been found in relatively few genes. In patients with LOM, these were (in ascending order of the number of genetic variants found) MAP 2 K 3 genes (72 variants on average), DNAH 5 (315 variants), HSPG 2 (383 variants), OBSCN (731 variants), SYNE 1 (733 variants), HLA – DRB 1 (766 variants), HLA – DQB 1 (842 variants), TTN (911 variants), AHNAK 2 (1029 variants), HLA – A (1038 variants), PDE 4 DIP(1210 variants), MUC 12 (1293 variants), MUC 3 A (1561 variants), MUC 4 (1680 variants), MUC 16 (1955 variants) – a total of 14519 variants in 15 genes.

A similar situation was observed in patients with SCCOM. In this group, genetic variants were most often found in the genes MAP 2 K 3 (176 variants), MUC 17 (206 variants), SYNE 1 (596 variants), PKD 1 L 2 (625 variants), AHNAK 2 (838 variants), HLA – DRB 1 (856 variants), HLA – B (888 variants), HLA – DQB 1 (1033 variants), HLA – A (1057 variants), PDE 4 DIP (1214 variants), HLA – DQA 1 (1280 variants), MUC 12 (1299 variants), MUC 4 (1433 variants), MUC 3 A (1539 variants), MUC 16 (1864 variants) – a total of 14904 variants of nucleotide sequences also in 15 genes.

Patients with LOM and SCCOM had the highest number of genetic variants found in the same genes, indicating no significant differences in the localization of germline mutations in these patient groups. An assessment of the statistical significance of differences in the incidence of identified variants of nucleotide sequences in identical genes between groups of patients with LOM and SCCOM is shown in Table 1.

Table 1

Statistical significance of differences in the frequency of occurrence of genetic variants in various exome genes between groups of patients with LOM and SCCOM

Gene	LOM		SCCOM		Stat. significance of differences according to the criterion χ^2 , p
	Number of genetic variants	% of total genetic variants (n = 124 475)	Number of genetic variants	% of total genetic variants (n = 152 767)	
MAP2K3	72	0.058	176	0.12	<0.0001 *
SYNE1	733	0.59	596	0.39	<0.0001 *
HLA-DRB1	766	0.62	856	0.56	0.059
HLA-DQB1	842	0.68	1033	0.68	0.99
AHNAK2	1029	0.83	838	0.55	<0.0001*
HLA-A	1038	0.83	1057	0.69	<0.0001*
PDE4DIP	1210	0.97	1214	0.79	<0.0001*
MUC12	1293	1.04	1299	0.85	<0.0001 *
MUC3A	1561	1.25	1539	1.01	<0.0001 *
MUC4	1680	1.35	1433	0.94	<0.0001 *
MUC16	1955	1.57	1864	1.22	<0.0001 *

* The difference is statistically significant.

As follows from Table 1, in general, the genetic variants listed above are significantly more common in patients with LOM than in the group of patients with SCCOM. However, the difference in absolute and relative rates of mutation occurrence between representatives of these groups is small, and the statistical significance of these differences stems from the large number of genetic variants identified, making this difference almost certainly not of clinical significance.

It is noteworthy that the genes with the largest (>40 %) number of variants in both groups of patients encode the synthesis of various mucin proteins that play an essential role in the formation of protective mucous barriers on epithelial surfaces and are involved in the renewal and differentiation of the epithelium. Insufficient production or a decrease in the functional activity of mucins can lead to chronic damage to the epithelial cells of the OM, which,

in turn, can serve as a trigger factor for both keratosis and malignant transformation of epitheliocytes [19–21].

All identified genetic variants were analyzed for known or probable pathogenicity. A total of 608 types of variants of nucleotide sequences with potential or proven pathogenicity due to the development of oncological diseases have been identified.

A list of the most pathogenic genetic variants found is shown in Table 2. In total, 12 such mutations were identified in the studied sample (1 genetic variant was detected in 2 patients, so the total number of variants is 11). The table shows that the most pathogenic genetic variants lead to the formation of stop codons or changes in splicing sites. In the first case, protein synthesis on mRNA suddenly stops. In the second case, one or more introns remain in the mature mRNA, producing abnormal proteins.

Table 2

The most pathogenic germline mutations identified in patients with LOM and SCCOM

Chromosome	Position	REF	ALT	Codons	Consequences	Global frequency (gnomAD)	Influence per phenotype	Frequency in the studied sample
2	219747090	C	A	tgC/tgA	stop gain	0.0006207	high	0.547
8	41834751	C	A	Gaa/Taa	stop gain	0.0002866	high	0.500
8	101725018	C	T		splice acceptor variant	0.0006185	high	1,000
8	143763531	G	A	tGg/tAg	stop gain	0.0208	high	0.833
9	80537095	G	T	taC/taA	stop gain	0.0009184	high	0.444
9	102595061	C	T	Cga/Tga	stop gain	0	high	0.250
10	75135885	G	A	Cga/Tga	stop gain	0.0002461	high	0.581
10	124339150	C	T	Cga/Tga	stop gain	0.001116	high	0.481
13	50205029	G	A	tGg/tAg	stop gain	0.009666	high	0.571
16	50813574	A	G		splice acceptor variant	0.000004	high	0.667
22	29130652	G	A	Cag/Tag	stop gain	0.0001414	high	0.125

Notes: REF – nucleotide in the reference genome, ALT – nucleotide in the same position of the genome of the analyzed sample, «Stop gain» – formation of a stop codon, «Splice acceptor variant» – formation of an abnormal splicing variant.

In the group of patients with LOM, four highly pathogenic genetic variants were identified. In the group of patients with SCCOM, eight highly pathogenic genetic variants were identified (the difference is not statistically significant: p (χ^2)=0.18; OR=2.5 at 95 % CI 0.64–9.82; RR=2 at 95 % CI 0.69–5.76). The number of identified pathogenic genetic variants in both groups was almost the same (502 in patients with LOM, 513 in the group of patients with SCCOM, the difference is not statistically significant).

In most cases (269 in the group of patients with LOM, 283 in the group of patients with SCCOM), the described genetic variants were detected in single patients. A relatively small number of pathogenic variants coincided in several patients from the studied sample. All of these genetic variants were moderately pathogenic (few highly pathogenic genetic variants were found, and only one of them was detected in two patients at the same time).

None of the identified moderately pathogenic genetic variants occurred in every patient in the study sample: the most common mutations were observed in a maximum of eight patients with LOM (33.3 % of the total size of this group) and seven patients with SCCOM (29.2 %,

respectively) (Table 3). Thus, it cannot be argued that any of the pathogenic genetic variants found in our study is unambiguously associated with developing the studied diseases in patients.

Nevertheless, for all found genetic variants, the average frequency of their occurrence in the human population was estimated, for which information was used from the international database of genetic variants gnomAD [22] (Table 3). Almost all of the pathogenic genetic variants we have identified are predictably very common (according to the gnomAD classification) since they occur in >0.5 % of human individuals (from 0.83 % to 4.97 %). In both groups of the studied sample of patients, these genetic variants were statistically significantly more common than in the human population as a whole (RR from 3.55 to 21.94, while the lower limit of the 95 % confidence interval exceeded 1 for all calculated RR values). Still, the indicated excess is unlikely to be of clinical significance due to the small size of the studied sample ($n = 24$ for both groups of patients), which does not exclude the influence of unpredictable random factors on the result, for example, non-random selection of patients in the sample or the distribution of specific alleles among the population of the Republic of Belarus or even Vitebsk regions (Table 3).

Table 3

The most common pathogenic genetic variants (all with moderate pathogenicity) in groups of patients with LOM and SCCOM, as well as an assessment of the degree of influence of these genetic variants on the likelihood of developing malignant transformation of the epithelium

Option	LOM with D1, n (%)	SCCOM, n (%)	Frequency in the population, % (gnomAD)	Effect of mutation on the phenotype of transl. proteins	Risk ratios (RR) and odds ratios (OR) (95 % CI)
p.A306G	4 (16.7)	–	3.78	Amino acid replacement Ala → Gly	RR: 4.41 (1.80-10.79) OR: 5.09 (1.74-14.90)
p.D109N	4 (16.7)	–	1.14	Amino acid substitution Asp → Asn	RR: 14.62 (5.96-35.86) OR: 17.35 (5.92-50.86)
p.E304K	4 (16.7)	–	4.57	Amino acid substitution G Glu → Lys	RR: 3.65 (1.49-8.93) OR: 4.18 (1.43-12.23)
p.G2502S	4 (16.7)	–	1.97	Amino acid substitution G Gly → Ser	RR: 8.46 (3.45-20.73) OR: 9.95 (3.40-29.16)
p.G613V	4 (16.7)	–	4.17	Amino acid substitution Gly → Val	RR: 4.00 (1.63-9.78) OR: 4.60 (1.57-13.54)
p.H604Y	4 (16.7)	–	4.17	Amino acid substitution His → Tyr	RR: 4.00 (1.63-9.78) OR: 4.60 (1.57-13.54)
p.N697S	4 (16.7)	4 (16.7)	4.70	Amino acid substitution A Asn → Ser	RR: 3.55 (1.45-8.68) OR: 4.06 (1.39-11.87)
p.Q356R	4 (16.7)	4 (16.7)	4.65	Amino acid substitution Gln → Arg	RR: 3.58 (1.46-8.77) OR: 4.10 (1.40-12.01)
p.S1024T	4 (16.7)	–	3.46	Amino acid substitution Ser → Thr	RR: 4.82 (1.97-11.79) OR: 5.58 (1.91-16.34)
p.K1019E	5 (20.8)	5 (20.8)	2.23	Amino acid substitution Lys → Glu	RR: 9.34 (4.28-20.41) OR: 11.54 (4.30-30.94)
p.R909K	5 (20.8)	6 (25.0)	1.40	Amino acid substitution Arg → Lys	LOM with D1: RR: 14.88 (6.81-32.54) OR: 18.54 (6.91-49.74) SCCOM: RR: 17.86 (8.91-35.80) OR: 23.48 (9.30-59.27)
p.Y673C	5 (20.8)	–	0.95	Amino acid substitution Tyr → Cys	RR: 21.94 (10.03-48.03) OR: 27.46 (10.22-73.73)
p.K416E	7 (29.2)	6 (25.0)	0.036	Amino acid substitution Lys → Glu	LOM with D1: RR: 822.67 (396.98-1704.85) OR: 1161.00 (445.58-3025.11) SCCOM: RR: 705.15 (320.37-1552.08) OR: 939.86 (346.45-2549.73)
p.R83Q	8 (33.3)	–	2.59	Amino acid substitution Arg → Gln	RR: 12.87 (7.30-22.71) OR: 18.81 (8.04-44.00)
p.S4010P	8 (33.3)	–	2.63	Amino acid substitution Ser → Pro	RR: 12.67 (7.19-22.35) OR: 18.51 (7.91-43.30)
p.E999G	–	4 (16.7)	0.83	Amino acid substitution Glu → Gly	RR: 20.08 (8.18-49.30) OR: 23.90 (8.15-70.12)
p.H144R	–	4 (16.7)	2.10	Amino acid substitution His → Arg	RR: 7.94 (3.24-19.44) OR: 9.33 (3.18-27.31)
p.R218C	–	4 (16.7)	4.11	Amino acid substitution Arg → Cys	RR: 4.06 (1.66-9.93) OR: 4.67 (1.59-13.66)
p.R34859Q	–	4 (16.7)	1.25	Amino acid substitution Arg → Gln	RR: 13.33 (5.44-32.69) OR: 15.80 (5.39-46.31)
p.V443I	–	4 (16.7)	0.98	Amino acid substitution Val → Ile	RR: 17.01 (6.93-41.74) OR: 20.22 (6.98-59.29)
p.V5L	–	4 (16.7)	3.92	Amino acid substitution Val → Leu	RR: 4.25 (1.74-10.41) OR: 4.90 (1.67-14.35)
p.A154G	–	5 (20.8)	3.38	Amino acid substitution Ala → Gly	RR: 6.16 (2.82-13.46) OR: 7.52 (2.81-20.16)
p.W1037S	–	5 (20.8)	4.61	Amino acid substitution Trp → Ser	RR: 4.52 (2.07-9.86) OR: 5.44 (2.03-14.59)
p.A402E	–	6 (25.0)	1.68	Amino acid substitution Ala → Glu	RR: 14.86 (7.43-29.83) OR: 19.51 (7.73-49.24)
p.P363L	–	7 (29.2)	4.97	Amino acid substitution Pro → Leu	RR: 5.87 (3.14-10.96) OR: 7.87 (3.26-19.00)
p.T87I	–	7 (29.2)	2.38	Amino acid substitution Thr → Ile	RR: 12.25 (6.56-22.89) OR: 16.88 (6.99-40.77)

Note. All odds and risks ratios given in this table are statistically significant (the lower limit of the 95 % confidence interval for all indicated ratios is >1).

However, the analysis identified the genetic variant p.K416E, which occurs in the human population «not too often» according to the gnomAD classification: its detection frequency averages 0.036 %. The specified genetic variant in both clinical groups of the studied sample was much more common than on average in the population: RR 822.67 (396.98–1704.85) in LOM and RR 705.15 (320.37–1552.08) in SCCOM, i. e., the found mutation occurs in the studied diseases 700–800 times more often than the average in the population, which may be of clinical significance. It is reasonable to assume that the p.K416E mutation is associated with a relatively high probability of developing both LOM and SCCOM.

The p.K416E genetic variant is a point SNP mutation (mononucleotide polymorphism); it occurs in the CHEK 2 gene (checkpoint kinase 2) and leads to the replacement of lysine by glutamic acid in exon 12 (out of 16) of the corresponding protein (CHK2). Lysine is an amino acid with strong main properties due to the ϵ -amino group. Glutamic acid, being an organic acid, has properties due to the free carboxyl group. The properties of these amino acids are so different that the replacement of Lys and Glu should cause significant conformational and, therefore, functional changes in the protein.

The CHK2 (serine-threonine protein kinase) protein encoded by the CHEK 2 gene is a critical cell cycle checkpoint regulator and a putative tumor suppressor. It contains a fork-headed Thr68 domain activated by phosphorylation in response to DNA damage. When activated, the CHK 2 protein inhibits CDC25C phosphatase, thereby preventing the cell from entering mitosis and stabilizing the p53 tumor suppressor protein, resulting in cell cycle arrest in G1. In addition, this protein interacts with and phosphorylates BRCA1, allowing BRCA1 to repair DNA integrity after damage. Mutations in the CHEK 2 gene are associated with Li-Fraumeni syndrome, a highly penetrating familial cancer phenotype commonly associated with germline mutations in TP53. In addition, mutations in the CHEK 2 gene are believed to predispose to sarcomas, breast cancer, and brain tumors [23].

Thus, it can be assumed that the p.K416E germline mutation of the CHEK 2 gene, which is relatively rare in the human population and was detected in a significant number of patients in the studied sample (in 29.2 % of patients with LOM and 25 % of patients with SCCOM), is associated with a high probability and the development of malignant transformation of the epithelium in both diseases. This mutation leads to disruption of the structure and function of the CHK 2 protein, which is one of the crucial components of the p53 signaling pathway and regulates the process of nuclear DNA repair after damage, which is why various pathological variants of the CHEK 2 gene are associated with the development of a number of oncological diseases [24].

Conclusions

1. The most significant number of germinal mutations in patients with both LOM and SCCOM is localized in 19 genes: MAP2K3, DNAH5, HSPG2, OBSCN, SYNE1, HLA-DRB1, HLA-DQA1, HLA-DQB1, HLA-A, HLA-B, PKD1L2, TTN, AHNK2, PDE4DIP, MUC3A, MUC4, MUC12, MUC16, MUC17 (from 72 to 1955 variants per gene). In the structure of the identified mutations, variants of the MUC3A, MUC4, MUC12, and MUC16 genes responsible for the synthesis of the family of mucin glycoproteins predominate (44.7 % and 41.2 % of germinal mutations in patients with LOM and SCCOM, respectively). Insufficient production or reduction of mucin functional activity can lead to chronic damage of SORP epithelial cells, which can serve as a trigger factor for both the development of keratosis and the malignant transformation of epithelial cells.

2. Highly pathogenic germline mutations occur in the exome of the studied patients with SCCOM more often than with LOM (8 and 4 genetic variants, respectively). Still, the difference in their frequency was not statistically significant. An infrequent germline mutation p.K416E of the CHEK 2 gene, found in 29.2 % of patients with LOM and 25 % of patients with SCCOM, is associated with the development of both diseases. This mutation disrupts the structure and function of the CHK 2 protein, which activates the process of nuclear DNA repair after damage and is also an essential component of the p53 signaling pathway.

Disclosures: The authors declare no conflict of interest.

References

1. Villa A., Hanna G. J., Kacew A., Frustino J., Hammerman P. S. [et al.]. Oral keratosis of unknown significance shares genomic overlap with oral dysplasia. *Oral Dis.* 2019;25(7):1707–1714. <https://doi.org/10.1111/odi.13155>
2. Noto Z., Tomihara K., Furukawa K., Noguchi M. Dyskeratosis congenita associated with leukoplakia of the tongue. *Int. J. Oral Maxillofac. Surg.* 2016;45(6):760–763. <https://doi.org/10.1016/j.ijom.2015.12.005>
3. Farah C. S., Jessri M., Bennett N. C., Dalley A. J., Shearston K. D. [et al.]. Exome sequencing of oral leukoplakia and oral squamous cell carcinoma implicates DNA damage repair gene defects in malignant transformation. *Oral Oncol.* 2019;96:42–50. <https://doi.org/10.1016/j.oraloncology.2019.07.005>
4. Govindarajan G. V., Bhanumurthy L., Balasubramanian A., Ramanathan A. A Novel Mutation in the DNA Binding Domain of NFKB Is Associated with Speckled Leukoplakia. *Asian Pac. J. Cancer Prev.* 2016;17(7):3627–3629.
5. Ogmundsdóttir H. M., Hilmarisdóttir H., Astvaldsdóttir A., Jóhannsson J. H., Holbrook W. P. Oral lichen planus has a high rate of TP53 mutations. A study of oral mucosa in iceland. *Eur. J. Oral Sci.* 2002;110(3):192–198. <https://doi.org/10.1034/j.1600-0447.2002.21235.x>
6. Kresty L. A., Mallory S. R., Knobloch T. J., Li J., Lloyd M. [et al.]. Frequent alterations of p16INK4a and p14ARF in oral proliferative verrucous leukoplakia. *Cancer Epidemiol. Biomarkers Prev.* 2008;17(11):3179–3187. <https://doi.org/10.1158/1055-9965.EPI-08-0574>
7. Han E., Patel N. A., Yannuzzi N. A. A unique case of coats plus syndrome and dyskeratosis congenita in a patient with CTC1 mutations. *Ophthalmic Genet.* 2020;41(4):363–367. <https://doi.org/10.1080/13816810.2020.1772315>
8. Mao L. NOTCH mutations: multiple faces in human malignancies. *Cancer Prev. Res. (Phila)*. 2015;8(4):259–261. <https://doi.org/10.1158/1940-6207.CAPR-15-0063>
9. Gouvêa A. F., Santos Silva A. R., Speight P. M., Hunter K., Carlos R. [et al.]. High incidence of DNA ploidy abnormalities and increased Mcm2 expression may predict malignant change in oral proliferative verrucous leukoplakia. *Histopathology.* 2013;62(4):551–562. <https://doi.org/10.1111/his.12036>
10. Zarate A. M., Don J., Secchi D., Carrica A., Costa F. G. [et al.]. Study of the TP53 codon 72 polymorphism in oral cancer and oral potentially malignant disorders in Argentine patients. *Tumor Biol.* 2017;39(5):1010428317699113. <https://doi.org/10.1177/1010428317699113>
11. Ribeiro I. P., Marques F., Barroso L., Rodrigues J., Caramelo F. [et al.]. Genomic profile of oral squamous cell carcinomas with an adjacent leukoplakia or with an

- erythroleukoplakia that evolved after the treatment of primary tumor: A report of two cases. *Mol. Med. Rep.* 2017;16(5):6780-6786.
<https://doi.org/10.3892/mmr.2017.7428>
12. Chang S. E., Bhatia P., Johnson N. W., Morgan P. R., McCormick F. [et al.] Ras mutations in United Kingdom examples of oral malignancies are infrequent. *Int. J. Cancer.* 1991;48(3):409-412.
<https://doi.org/10.1002/ijc.2910480318>
13. Chung C. M., Hung C. C., Lee C. H., Lee C. P., Lee K. W. [et al.] Variants in FAT1 and COL9A1 genes in male population with or without substance use to assess the risk factors for oral malignancy. *PLoS One.* 2019;14(1):e0210901.
<https://doi.org/10.1371/journal.pone.0210901>
14. Nextera DNA Exome Reference (1000000039018). Available at: https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/samplepreps_nextera/nextera-dna-exome/nextera-dna-exome-reference-1000000039018-00.pdf. Accessed December 6, 2022.
15. Illumina TruSight Oncology 500 Reference Guide. Available at: https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/trusight/oncology-500/trusight-oncology-500-reference-guide-10000000067621_07.pdf. Accessed May 5, 2022.
16. Kanzi A. M., San J. E., Chimukangara B., Wilkinson E., Fish M. [et al.] Next Generation Sequencing and Bioinformatics Analysis of Family Genetic Inheritance. *Front. Genet.* 2020;11:544162.
<https://doi.org/10.3389/fgene.2020.544162>
17. Fox A. J., Hiemenz M. C., Lieberman D. B., Sukhadia S., Li B. [et al.] Next Generation Sequencing for the Detection of Actionable Mutations in Solid and Liquid Tumors. *J. Vis. Exp.* 2016;(115):52758. <https://doi.org/10.3791/52758>
18. Buzdugan L., Kalisch M., Navarro A., Schunk D., Fehr E., Bühlmann P. Assessing statistical significance in multivariable genome wide association analysis. *Bioinformatics.* 2016;32(13):1990-2000.
<https://doi.org/10.1093/bioinformatics/btw128>
19. Reibel J. Prognosis of oral pre-malignant lesions: significance of clinical, histopathological, and molecular biological characteristics. *Crit. Rev. Oral Biol. Med.* 2003;14(1):47-62. <https://doi.org/10.1177/154411130301400105>
20. Kumar M. H., Sanjai K., Kumarswamy J., Keshavaiah R., Papaiah L. [et al.] Expression of MUC1 mucin in potentially malignant disorders, oral squamous cell carcinoma and normal oral mucosa: An immunohistochemical study. *J. Oral Maxillofac. Pathol.* 2016;20(2):214-218.
<https://doi.org/10.4103/0973-029X.185916>
21. Villa A., Celentano A., Glurich I., Borgnakke W. S., Jensen S. B. [et al.] World Workshop on Oral Medicine VII: Prognostic biomarkers in oral leukoplakia: A systematic review of longitudinal studies. *Oral Dis.* 2019;25(Suppl 1):64-78. <https://doi.org/10.1111/odi.13087>
22. The Genome Aggregation Database (gnomAD). Available at: <https://gnomad.broadinstitute.org>. Accessed July 25, 2022.
23. Boonen R., Wiegant W. W., Celosse N., Vrolijk B., Heijl S. [et al.] Functional Analysis Identifies Damaging CHEK2 Missense Variants Associated with Increased Cancer Risk. *Cancer Res.* 2022;82(4):615-631.
<https://doi.org/10.1158/0008-5472.CAN-21-1845>
24. Offit K., Garber J. E. Time to check CHEK2 in families with breast cancer? *J. Clin. Oncol.* 2008;26(4):519-520.
<https://doi.org/10.1200/JCO.2007.13.8503>

About authors:

Karpuk Natalia Anatolievna, CMSc, Associate Professor, Associate Professor of the Department of General and Orthopedic Dentistry with the course of FAT and RP;
tel.: +375295961083; e-mail: ms.karpuk@mail.ru; ORCID: 0000-0001-9991-7034

Rubnikovich Sergey Petrovich, Corresponding Member, DMSc, Professor, Rector;
tel.: +375296372156; e-mail: rubnikovichs@mail.ru; ORCID: 0000-0002-7450-3757

Mazur Oksana Cheslavovna, Researcher, Laboratory of Environmental Genetics and Biotechnology;
tel.: +375172841918; e-mail: terezia@mail.ru; ORCID: 0000-0002-6093-4548

Zhylytsou Ivan Viktorovich, DMSc, Professor, Head of the Department of Evidence-Based Medicine and Clinical Diagnostics of the Faculty of Advanced Training and Retraining;
tel.: +375297104368; e-mail: zhylytsou@tut.by; ORCID: 0000-0002-4912-2880

Karpuk Ivan Yurievich, DMSc, Professor of the Department of General Dentistry including the course of Prosthodontic Dentistry;
tel.: +375297119736; e-mail: ikarpuk@mail.ru; ORCID: 0000-0001-9991-7035

Sirak Sergey Vladimirovich, MD, PhD, Professor, Head of the Department of Dentistry;
tel.: +78652350551; e-mail: sergejsirak@yandex.ru; ORCID: 0000-0002-4924-5792

Shchetinin Evgeny Vyacheslavovich, MD, PhD, Professor, Head of the Department of Pathophysiology;
tel.: +78652352684; e-mail: ev.cliph@rambler.ru; ORCID: 0000-0001-6193-8746

Mikhalechenko Alena Petrovna, PhD (Biol.), Senior Researcher, Laboratory of Environmental Genetics and Biotechnology;
tel.: +375172841918; e-mail: michalenko75@mail.ru