ISSN 2304-3415, Russian Open Medical Journal
2024. Volume 13. Issue 3 (September). Article CID e0308
DOI: 10.15275/rusomj.2024.0308

Original article

Germline variants of the MMR/*EPCAM* genes in Russian patients with Lynch syndrome

Aleksey S. Tsukanov [1], Aleksey A. Barinov [1], Vitaly P. Shubin [1], Anna N. Loginova [1], Dmitry Yu. Pikunov [1], Dmitry G. Shakhmatov [1,2], Yuri A. Shelygin [1,2], Sergey I. Achkasov [1]

[1] Ryzhikh National Medical Research Centre for Coloproctology, Moscow, Russia
[2] Russian Medical Academy of Continuing Professional Development, Moscow, Russia

Received 14 December 2023, Revised 26 March 2024, Accepted 18 June 2024

Abstract: *Objective* — Lynch syndrome is the most common hereditary cancer syndrome caused by pathogenic variants in the MMR/*EPCAM* genes. Our goal was to study the germline variants of these genes in the largest sample of patients with Lynch syndrome in Russia.

Methods — The study included data on pathogenic and likely pathogenic variants in the MMR/*EPCAM* genes collected from the local Registry of Hereditary Colorectal Cancer of Ryzhikh National Medical Research Centre for Coloproctology (RNMRCC). We analyzed all available data on 141 probands with Lynch syndrome: 78 men and 63 women aged 21 to 80 years treated at RNMRCC from 2012 through 2023.

Results — The numbers of pathogenic and likely pathogenic variants detected in the key genes were as follows *MLH1* (69 probands), *MSH2* (57), *MSH6* (10), *EPCAM* (3), *PMS2* (1), and *PMS1* (1). Of the 141 germline variants, 17 (12.1%) were detected for the first time: *MLH1* (8 probands) *MSH2* (8), and *MSH6* (1). Thirteen (9.2%) of the detected variants were associated with large deletions/duplications. The most frequent pathogenic variants in the *MLH1* gene were c.1852_1854del (9 probands), c.350C>T (4), and c.1459C>T (3). The most frequent pathogenic variants in the *MSH2* gene were c.942+3A>T (9 probands) and c.1288A>T (3).

Conclusion — Our results allowed establishing the frequency and spectrum of different types of germline variants in the MMR/*EPCAM* genes, which helped optimize the algorithm for selecting and diagnosing Russian patients with Lynch syndrome.

Keywords: Lynch syndrome, colorectal cancer, microsatellite instability, MMR/*EPCAM* genes.

Cite as Tsukanov AS, Barinov AA, Shubin VP, Loginova AN, Pikunov DYu, Shakhmatov DG, Shelygin YuA, Achkasov SI. Germline variants of the MMR/*EPCAM* genes in Russian patients with Lynch syndrome. *Russian Open Medical Journal* 2024; 13: e0308.

Correspondence to Dmitriy Yu. Pikunov. Address: 2 Salyama Adilya St., Moscow 123423, Russia. Phone: +79161823228; E-mail: pikunov.gnck@mail.ru.

Introduction

Lynch syndrome is a disease with an autosomal dominant type of inheritance characterized by the development of malignant tumors of various organs, the cause of which is a germline pathogenic variant in one of the genes of the DNA mismatch repair system (MMR system) or in the *EPCAM* gene. The frequency of Lynch syndrome in Europeans is approximately 1:300 people, which makes it the most common among all hereditary oncological syndromes [1]. Carriers of pathogenic variants can develop tumors of the colon, uterus, urinary system, stomach, etc. DNA diagnostics of blood relatives of patients with Lynch syndrome is extremely important, since lifelong clinical monitoring of carriers of the pathogenic variant significantly reduces both the incidence of malignant neoplasms and mortality caused by them [2].

Lynch syndrome causes only 3% of all cases of colorectal cancer [3], and the process of its diagnosis includes several stages. Initially, a sample of a colon tumor should be examined for microsatellite instability (MSI). This phenomenon occurs in approximately 12% of patients with sporadic colorectal cancer and in all Lynch-associated tumors [4]. Only if it is detected, a further search for a germline variant in the MMR system genes is carried out [2]. These genes were mapped in the 1990s in families of patients with predominantly hereditary colon cancer: *MLH1* (OMIM:120436), *MSH2* (OMIM:609309), *MSH6* (OMIM:600678), *PMS2* (OMIM:600259), *PMS1* (OMIM:600258), etc. It is crucial to mention the *EPCAM* gene (OMIM:185535), the product of which is involved in the adhesion of epithelial cells and also takes part in proliferation. This gene is located immediately before the *MSH2* gene, and large deletions of the 3' end of the *EPCAM* gene lead to epigenetic hypermethylation of the *MSH2* gene promoter, causing the development of Lynch syndrome [5].

The Human Gene Mutation Database Professional (HGMD Pro) 2023.1 contains more than 3,500 described unique pathogenic variants in the MMR/*EPCAM* genes. At the same time, its data are constantly being updated, since previously undescribed variants are found in different patient samples from different populations.

This publication is dedicated to the study of germline variants in the MMR and *EPCAM* genes in Russian patients with Lynch syndrome.

Material and Methods

Patient selection

An observational single-center study aimed at identifying and studying patients with Lynch syndrome was conducted from 2012 through 2023 at the Ryzhikh National Medical Research Centre for Coloproctology (RNMRCC) of the Russian Federation Ministry of Healthcare (Moscow, Russia). The study material included data on 141 patients with a genetically confirmed diagnosis of Lynch syndrome from the local Registry of Hereditary Colorectal Cancer at RNMRCC. Among them, there were 78 men and 63 women aged 21 to 80 years. Informed consent was obtained from all patients included in the study. This study complied with the ethical principles of the Declaration of Helsinki and was approved by the Ethics Committee at RNMRCC.

The Amsterdam Criteria II or Bethesda Guidelines were employed to select patients with suspected Lynch syndrome from 2012 to 2014. Subsequently, from 2014 to 2023, the following more effective selection criteria developed at RNMRCC were used:

1. Colorectal cancer in a patient under 43 years of age (sensitivity: 88.9%; specificity: 82.9%);

2. In addition to colorectal cancer, two or more cases of Lynch syndrome-associated cancer in the patient or in his/her blood relatives in the same line, regardless of age (sensitivity: 100%; specificity: 64.7%).

Table 1. The spectrum of pathogenic and likely pathogenic variants in the *MLH1* gene

HGVS c.DNA NM_000249.4:	*HGVS protein*	*Consequence*	*Pts*	*New*	*Status*	*Pathogenic criteria*
c.2T>G	p.Met1?	start_lost	1			
c.100G>T	p.Glu34Ter	stop_gained	1	YES	P	PVS1, PM2, PP4 (supporting)
c.114C>G	p.Asn38Lys	missense_variant, splice_region_variant	1			
c.117-2A>G	-	splice_acceptor_variant	1			
c.139_140insAT	p.Ile47AsnfsTer4	frameshift_variant	1	YES	P	PVS1, PM2, PP4 (moderate)
c.160_166del	p.Gly54Ter	frameshift_variant	1	YES	P	PVS1, PM2, PP4 (moderate)
c.187G>A	p.Asp63Asn	missense_variant	1			
c.207+2T>A	-	splice_donor_variant	1			
c.298C>T	p.Arg100Ter	stop_gained	2			
c.299G>C	p.Arg100Pro	missense_variant	1			
c.306+5G>A	-	splice_donor_variant, intron_variant	2			
c.322_335del	p.Ser108CysfsTer9	frameshift_variant	1			
c.346dup	p.Thr116AsnfsTer6	frameshift_variant	1			
c.350C>A	p.Thr117Lys	missense_variant	1			
c.350C>T	p.Thr117Met	missense_variant	4			
c.444_450del	p.Gln149ArgfsTer9	frameshift_variant	1	YES	LP*	PVS1, PM2
c.445dup	p.Gln149ProfsTer23	frameshift_variant	1	YES	P	PVS1, PM2, PP4 (supporting)
c.454-13A>G	-	splice_acceptor_variant, intron_variant	1			
c.546-2A>G	-	splice_acceptor_variant	1			
c.677G>A	p.Arg226Gln	missense_variant, splice_region_variant	2			
c.677G>T	p.Arg226Leu	missense_variant, splice_region_variant	2			
c.694G>T	p.Gly232Ter	stop_gained	1	YES	P	PVS1, PM2, PP4 (moderate)
c.947del	p.Phe316SerfsTer51	frameshift_variant	1	YES	P	PVS1, PM2, PP4 (supporting)
c.1072del	p.Glu358ArgfsTer9	frameshift_variant	1			
c.1225C>T	p.Gln409Ter	stop_gained	1			
c.1459C>T	p.Arg487Ter	stop_gained	3			
c.1520dup	p.Leu507PhefsTer8	frameshift_variant	1			
c.1652A>C	p.Asn551Thr	missense_variant	1			
c.1668-1G>C	-	splice_acceptor_variant	1			
c.1731G>A	p.Ser577=	splice_region_variant, synonymous_variant	2			
c.1731G>T	p.Ser577=	splice_region_variant, synonymous_variant	1			
c.1756G>C	p.Ala586Pro	missense_variant	1			
c.1783_1784del	p.Ser595TrpfsTer14	frameshift_variant	1			
c.1852_1854del	p.Lys618del	inframe_deletion	9			
c.1896+1G>C	-	splice_donor_variant	1			
c.1896+1G>T	-	splice_donor_variant	1			
c.1921dup	p.Leu641ProfsTer4	frameshift_variant	1			
c.1949T>A	p.Leu650Ter	stop_gained	1	YES	LP*	PVS1, PM2
c.1990-2A>G	-	splice_acceptor_variant	1			
c.2038T>C	p.Cys680Arg	missense_variant	1			
c.2041G>A	p.Ala681Thr	missense_variant	1			
c.2059C>T	p.Arg687Trp	missense_variant	1			
c.2073_2074del	p.Ser692Ter	frameshift_variant	1			
c.2103+1G>C	-	splice_donor_variant	2			
c.2219del	p.Ile740ThrfsTer43	frameshift_variant	1			

* family data are unavailable. LP, likely pathogenic; P, pathogenic.

Table 2. The spectrum of pathogenic and likely pathogenic variants in the *MSH2* gene

HGVS c.DNA NM_000251.3:	*HGVS protein*	*Consequence*	*Pts*	*New*	*Status*	*Pathogenic criteria*
c.347_350del	p.Asp116GlyfsTer57	frameshift_variant	1			
c.354T>G	p.Tyr118Ter	stop_gained	1			
c.388_389del	p.Gln130ValfsTer2	frameshift_variant	1			
c.571_573del	p.Leu191del	inframe_deletion	1			
c.741dup	p.Lys248GlnfsTer8	frameshift_variant	1	YES	P	PVS1, PM2, PP4 (supporting)
c.792+2T>C	-	splice_donor_variant	1			
c.942+3A>T	-	splice_donor variant, intron_variant	9			
c.942G>A	p.Gln314=	splice_region_variant, synonymous_variant	1			
c.970_971del	p.Gln324ValfsTer8	frameshift_variant	1			
c.989T>C	p.Leu330Pro	missense_variant	2			
c.1119del	p.Arg373SerfsTer39	frameshift_variant	1			
c.1125_1126insAT	p.Leu376IlefsTer37	frameshift_variant	1			
c.1170del	p.Ala391ProfsTer21	frameshift_variant	1			
c.1174A>T	p.Lys392Ter	stop_gained	1	YES	P	PVS1, PM2, PP4 (moderate)
c.1204C>T	p.Gln402Ter	stop_gained	1			
c.1221_1222del	p.Tyr408SerfsTer8	frameshift_variant	1			
c.1231_1234delATAAinsCT	p.Ile411LeufsTer5	frameshift_variant	1	YES	P	PVS1, PM2, PP4 (moderate)
c.1255C>T	p.Gln419Ter	stop_gained	1			
c.1280_1378dup	p.Lys427_Asp459dup	inframe_insertion	1	YES	LP	PM2, PM4, PP4 (moderate)
c.1288A>T	p.Lys430Ter	stop_gained	3			
c.1386+1G>T	-	splice_donor_variant	2			
c.1538T>C	p.Leu513Pro	missense_variant	1	YES	LP	PP3, PM2, PS3, PP4 (supporting), BP1
c.1566C>G	p.Tyr522Ter	stop_gained	1			
c.1699A>T	p.Lys567Ter	stop_gained	1			
c.1786_1788del	p.Asn596del	inframe_deletion	1			
c.1861C>T	p.Arg621Ter	stop_gained	2			
c.1968C>A	p.Tyr656Ter	stop_gained	1			
c.1968C>G	p.Tyr656Ter	stop_gained	1			
c.1979A>G	p.Asp660_Thr668del	missense_variant	1		LP	PS3, PM2, PP5, PP4 (moderate), BP1
c.2038C>T	p.Arg680Ter	stop_gained	2			
c.2086C>T	p.Pro696Ser	missense_variant	2			
c.2231T>G	p.Leu744Ter	stop_gained	1			
c.2266_2267del	p.Thr756LeufsTer30	frameshift_variant	1	YES	P	PVS1, PM2, PP4 (moderate)
c.2287G>C	p.Ala763Pro	missense_variant	1			
c.2397_2398delTCinsA	p.Asn799LysfsTer13	frameshift_variant	1	YES	P	PVS1, PM2, PP4 (supporting)
c.2407dup	p.Thr803AsnfsTer6	frameshift_variant	1			
c.2455_2458dup	p.Gly820GlufsTer5	frameshift_variant, splice_region_variant	1	YES	P	PVS1, PM2, PP4 (supporting)
c.2633_2634del	p.Glu878AlafsTer3	frameshift_variant, splice_region_variant	1			

LP, likely pathogenic; P, pathogenic.

Laboratory methods

The molecular genetic study, conducted from 2012 to 2018, included a primary search for MSI in colorectal cancer samples using fragment analysis by capillary electrophoresis on an ABI PRISM 3500 sequencer (Applied Biosystems, USA) using NR21, NR24, NR27, BAT25, BAT26 markers. When MSI was detected in the tumor, a search for pathogenic variants of the *MLH1*, *MSH2* and *MSH6* genes was carried out in DNA isolated from patients' peripheral blood lymphocytes. For this purpose, polyacrylamide gel electrophoresis was performed on a Sequi-Gen GT Sequencing Cell (BIO RAD, USA). In case of differences in the electrophoretic pattern of the fragments of the patient's genes under study from the control samples, they were sequenced using the Sanger method on an ABI PRISM 3500 device. In the absence of pathogenic variants in the *MLH1/MSH2/MSH6* genes, a panel of genes, including the MMR genes, was sequenced using the Junior 454 instrument (Roche, Switzerland) in accordance with the manufacturer's protocol. All identified hereditary variants were confirmed by the Sanger method [6].

The molecular genetic study, conducted from 2019 through 2023, included an initial search for MSI in the tumor using fragment analysis. When MSI was detected, germline pathogenic variants in the *MLH1* and *MSH2* genes were immediately searched for by Sanger sequencing on an ABI PRISM 3500 instrument. Large rearrangements were also searched by Multiplex Ligation-Dependent Probe Amplification (MLPA) in the *MLH1*, *MSH2*, *MSH6*, and *EPCAM* genes (SALSA MLPA Probemix P003 *MLH1/MSH2*; SALSA MLPA Probemix P072 *MSH6*) according to the manufacturer's protocol. In the absence of pathogenic variants, whole-exome sequencing was performed; 100 ng of total genomic DNA was used to prepare paired-end enriched libraries using a tagmentation-based protocol and xGen Exome Research Panel v2 probes. Sequencing was performed on a NextSeq 550 platform (Illumina, USA) with a read length of 2*75 bp with an average coverage of 100x. The resulting reads were mapped to the reference GRCh38/hg38 genome using the BWA [7] and SAMtools [8] algorithms with recalibration. Obtained via the GATK [9] and Deepvariant [10] software, SNV variants were annotated using the databases (dpSNP v155, GnomAD v 2.1.1, ExAC v0.3.1). CNV analysis was conducted using the CODEX2 package [11]. Whole-exome sequencing was also performed for archival DNA samples in which pathogenic variants had not been previously identified. Variants identified by whole-exome sequencing were confirmed by the Sanger method. To establish the pathogenicity of previously undescribed variants, DNA diagnostics of patient relatives, routine MMR immunohistochemistry of tumor samples and pathogenicity

assessment according to the ACMG and CanVIG criteria for MMR genes were performed [12, 13].

Results

The conducted molecular genetic study allowed diagnosing Lynch syndrome in 141 Russian patients. Previously undescribed germline variants were found in 17 of 141 probands (12.1%). All of them were analyzed in accordance with the most recent recommendations for the interpretation of variants detected in cancer predisposition genes (CanVIG) and a special supplement for the evaluation of variants in MMR genes [14].

Pathogenic and likely pathogenic variants in the *MLH1* gene were revealed in 64 patients, including eight previously undescribed variants (*Table* 1).

In the *MSH2* gene, 53 pathogenic/likely pathogenic variants were detected, including 8 undescribed before (*Table* 2).

It is necessary to consider in more detail the three most difficult to interpret germline variants in the *MSH2* gene: c.1280_1378dup, c.1538T>C and c.1979A>G.

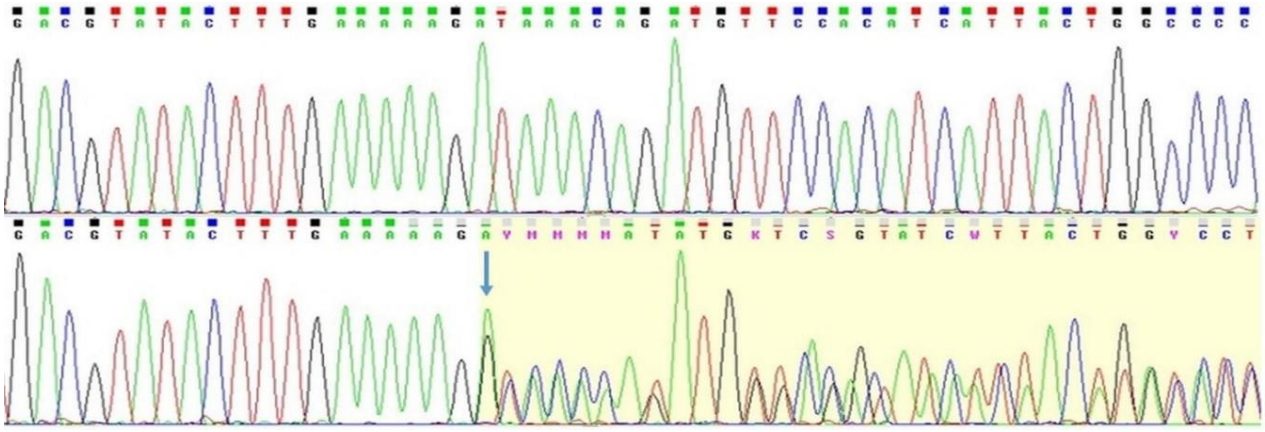

Figure 1. Results of sequencing of the cDNA region of the *MSH2* gene: A, control sample; B, patient with c.1979A>G variant (the arrow indicates the beginning of the deletion).

Table 3. Pathogenic and likely pathogenic variants in the *MSH6*, *PMS1*, and *PMS2* genes

HGVS c.DNA NM_000179.3:	*HGVS protein*	*Consequence*	*Pts*	*New*	*Status*	*Pathogenic criteria*
c.742C>T	p.Arg248Ter	stop_gained	1			
c.1815_1816del	p.Lys606AsnfsTer33	frameshift_variant	1			
c.2234T>A	p.Ile745Asn	missense_variant	1			
c.2764C>T	p.Arg922Ter	stop_gained	1			
c.3103C>T	p.Arg1035Ter	stop_gained	1			
c.3202C>T	p.Arg1068Ter	stop_gained	1			
c.3311_3312del	p.Phe1104TrpfsTer3	frameshift_variant	1			
c.3577G>T	p.Glu1193Ter	stop_gained	1	YES	P	PVS1, PM2, PP4 (supporting)
c.3931G>T	p.Glu1311Ter	stop_gained	1			
HGVS c.DNA NM_000534.5:	HGVS protein	Consequence	Pts	New	Status	Pathogenic criteria
c.829C>T	p.Arg277Ter	stop_gained	1			
HGVS c.DNA NM_000535.7:	HGVS protein	Consequence	Pts	New	Status	Pathogenic criteria
c.1144+1G>A	-	splice_donor_variant	1			

P, pathogenic.

Table 4. Large rearrangements in the MMR/*EPCAM* genes

Gene	HGVS g.DNA (GRCh38/hg38)	Consequence	Pts	Affected exons
MLH1	NC_000003.12 g.(?_36993535)_(36996675_37000995)del	large del	1	del 1,2
MLH1	NC_000003.12 g.(?_36993535)_(37008815_37011825)del	large del	1	del 1-6
MLH1	NC_000003.12 g.(37020405_37025875)_(37028825_37040215)del	large del	1	del 12,13
MLH1	NC_000003.12 g.(37028825_37040215)_(37040215_37042275)del	large del	1	del 14
MLH1	NC_000003.12 g.(37040215_37042275)_(37050745_?)del	large del	1	del 15-19
MSH2	NC_000002.12 g.(47386757_47402797)_(47429847_47445577)del	large del	1	del 1-7
MSH2	NC_000002.12 g.(47416337_47429847)_(47429847_47445577)dup	large_dup	1	dup7
MSH2	NC_000002.12 g.(47470977_47475107)_(47478327_47480787)del	large del	1	del 12-14
MSH2	NC_000002.12 g.(47480977_47485107)_(47492767_?)del	large del	1	del 12-16
MSH6	NC_000002.12 g.(47521307_47782907)_(47783237_47790957)del	large del	1	del 1
EPCAM	NC_000002.12 g.(47373987_47385157)_(47390107_47400377)del	large del	2	del 8,9
EPCAM-MSH2	NC_000002.12 g.(?_47373987)_(47402977_47521307)del	large del	1	del 3-9_del 1

The first proband was found to have a 99-nucleotide insertion: c.1280_1378dup (p.Lys427_Asp459dup), which did not result in a frameshift. The patient's family history was as follows: his father had rectal cancer at the age of 60 years, and his father's mother had colon cancer at the age of 70 years. This proband was diagnosed with sebaceous carcinoma of the skin in the scapular region at the age of 36 years, and with synchronous/metachronous multiple primary malignancies of the sigmoid and cecum at the age of 48 years. MMR immunohistochemical examination of colon tumor biopsies revealed a dMMR phenotype with loss of MSH2/MSH6 proteins, and MSI was detected by PCR and fragment analysis. Variant c.1280_1378dup was classified by us as probably pathogenic based on the sum of the following criteria: PM2, PM4 and PP4 (moderate evidence).

Another proband with the hereditary variant c.1538T>C (p.Leu513Pro) was diagnosed with colon cancer at the age of 43 years. His grandfather died of esophageal cancer at the age of 65 years. The results of MMR immunohistochemical examination of the colon tumor revealed a dMMR phenotype with loss of MSH2/MSH6 proteins, and MSI was detected by PCR and fragment analysis. DNA diagnostics revealed the variant c.1538T>C in the *MSH2* gene, leading to the replacement of nonpolar leucine with nonpolar proline. This variant is absent from population samples, it is present in a conserved region among all vertebrates (PhyloP100way 7.252), while functional studies (MAVE 2.43) and over 10 bioinformatic predictors (REVEL 0.943) indicate a high probability of pathogenic significance in this variant.

Consequently, the c.1538T>C variant was classified as likely pathogenic using the following criteria: PP3, PM2, PS3, PP4 (supporting), BP1.

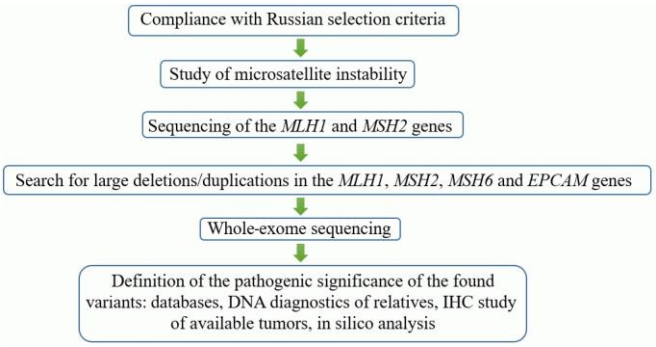

Figure 2. Algorithm of diagnosing Lynch syndrome.

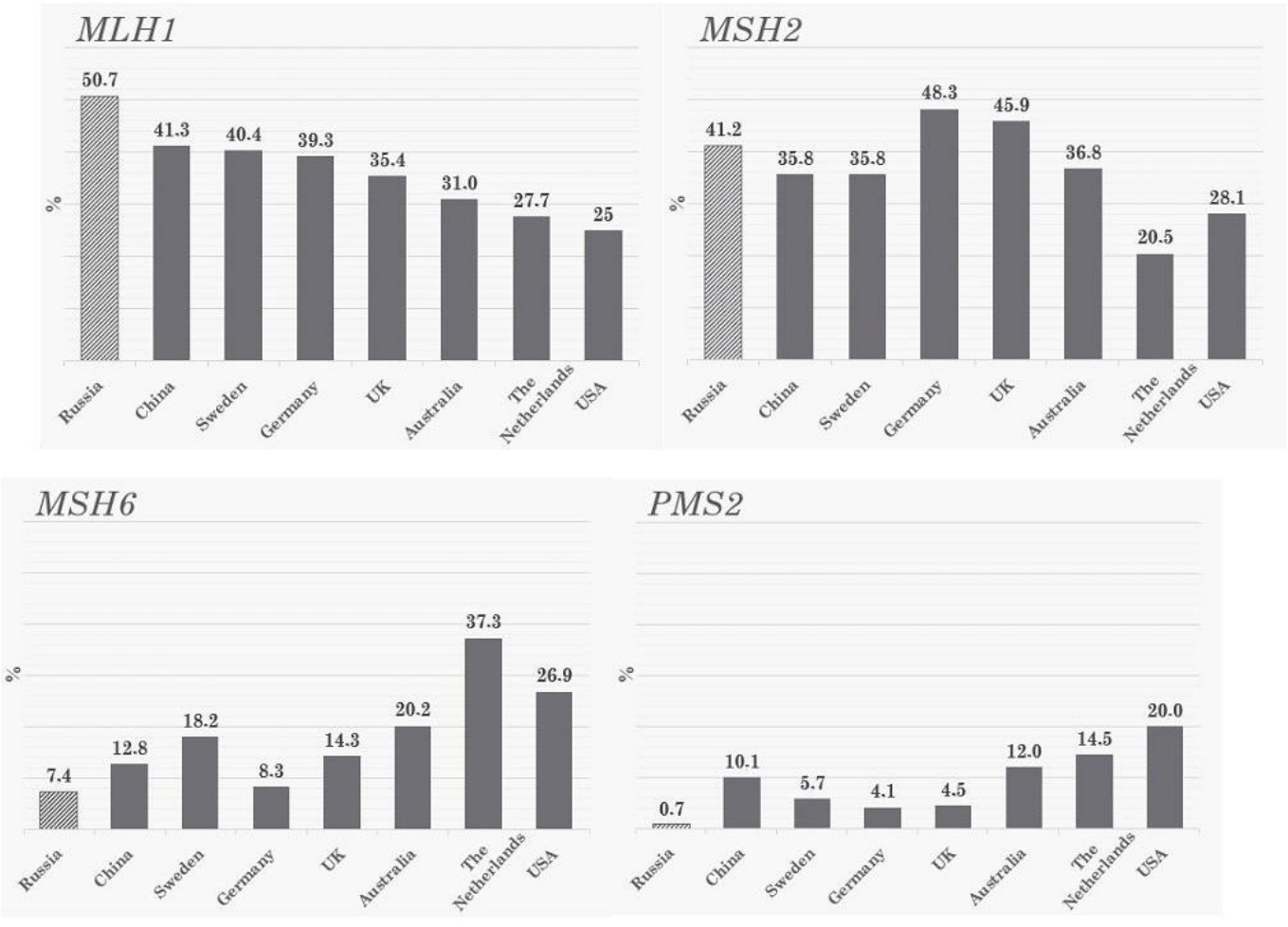

Figure 3. The frequency of germline variants in the MMR genes in Lynch syndrome patients in different countries [18-22].

The third patient was found to have the c.1979A>G variant, which was previously reported to have conflicting data regarding its association with Lynch syndrome. In this regard, we performed a functional study at the mRNA level, which confirmed the loss of nine amino acids (p.Asp660_Thr668del) (*Figure* 1). This allowed applying the PS3 criterion. In addition, a modified PP4 (moderate evidence) criterion was applied to this variant, since the patient's pedigree showed an aggravated family history: the father's mother died of colon cancer at the age of 65 years; the father was diagnosed with metachronous colorectal cancer at the age of 69 years, 70 years and 72 years; the father's sister was diagnosed with colon cancer at the age of 40 years, and the patient per se was diagnosed with metachronous multiple primary cancer of the cecum at the age of 35 years, descending colon at the age of 46 years, and a sebaceous carcinoma of the nose skin at the age of 45 years. Hence, this variant was classified as likely pathogenic based on the sum of the criteria: PS3, PM2, PP5, PP4 (moderate), BP1.

Eleven germline variants were revealed in the genes *MSH6*, *PMS1*, *PMS2*, including one new variant (*Table* 3).

MLPA method made it possible to find large deletions/duplications in 13 probands with Lynch syndrome (*Table* 4).

Discussion

Identification of patients with Lynch syndrome is an extremely difficult problem for two reasons. The first one is the issue of patient selection, and it lies in the fact that Lynch syndrome causes only 3% of colon cancer [3], which indicates the impossibility of conducting complete genetic diagnostics in all patients with colorectal cancer due to its high cost and time expenditures. At the same time, the lack of timely diagnosis of Lynch syndrome leads to incorrect treatment tactics and the exclusion of examination of the patient's blood relatives. Also, the sensitivity of known patient selection guidelines, such as Amsterdam Criteria II and Bethesda Guidelines, is far from 100%, since they fail to identify 28% of patients with Lynch syndrome [15]. The second reason is genetic. There is a need to study a large number of genes to identify a pathogenic variant in patients with suspected Lynch syndrome: *MLH1*, *MSH2*, *MSH6*, *PMS2*, *PMS1*, *EPCAM* and some other. At the same time, various molecular genetic methods used in different laboratories allow reducing the cost and time of DNA diagnostics. Such methods include, in particular, electrophoresis (SSCP, conformation-sensitive, etc.) and HRM analysis (the sensitivity and specificity of which does not allow identifying all possible pathogenic variants). As a result of using the method of high throughput sequencing of a significant number of genes, a huge number of germline variants can be identified, often with previously unknown missense mutations and intron substitutions, the pathogenicity of which is not always possible to establish.

To solve the first problem (patient selection), after the first two years of clinical genetic testing, we developed our own criteria, the sensitivity of which significantly exceeded the sensitivity values of the revised Bethesda and Amsterdam II guidelines [16].

Solving the genetic problem took much longer. First of all, we found that the conformation-sensitive electrophoresis we used did not allow detecting single-nucleotide pathogenic variants in approximately 7% of patients, which led to a complete rejection of this method and its replacement with Sanger sequencing for all exons of the *MLH1* and *MSH2* genes. With the advent of next-generation sequencing, we initially used a gene panel to search for mutations, but later decided to use whole-exome sequencing, which allowed us to immediately find some of the large rearrangements. The MLPA method was used to identify the remaining large deletions/duplications. Finally, as one of the most important methods for determining the pathogenic significance of new variants (in addition to DNA diagnostics of relatives and the use of prognostic programs), we chose MMR immunohistochemical examination of tumor samples not only from the patient, but also from the patient's relatives with oncological diseases. Thus, the period for developing the final diagnostic algorithm for Lynch syndrome was almost seven years (*Figure* 2).

The results obtained using this algorithm allow stating with confidence that at the moment we can identify the most complete spectrum of pathogenic variants of the MMR/*EPCAM* genes in almost all of our patients suffering from Lynch syndrome. Moreover, as an example of its high efficiency, we would like to cite a unique instance of using the developed algorithm. As a result, in one of the examined patients we were able to genetically diagnose two severe hereditary syndromes at once: Lynch syndrome (gene *MSH6*: NM _000179.2:c.742C>T, p.Arg248Ter) and Diamond-Blackfan anemia (deletion of the chromosome 15 locus with the capture of the interval 82662932–84816747bp, including the complete sequence of the *RPS17* gene). It should be noted that the estimated frequency of such patients was only 1 case per 480 million people, while such cases have not been previously described at all [17].

Unfortunately, not all laboratories of the world currently diagnose the genes *EPCAM*, *PMS1*, etc. Therefore, we can compare our own data with most other results only for four main genes: *MLH1*, *MSH2*, *MSH6* and *PMS2* (*Figure* 3).

It is extremely important to emphasize that in Russia, among patients with Lynch syndrome, the most common hereditary mutations are found in the *MLH1* gene (50.7%). However, such a high frequency of germline variants in one of the MMR genes has not been described in any other country (*Figure* 3). This fact indicates the advisability of initiating molecular genetic testing in Russian patients with the *MLH1* gene, which will significantly reduce the time of DNA diagnostics of every second patient with Lynch syndrome. The combined frequency of pathogenic and likely pathogenic variants in the *MLH1* and *MSH2* genes in Russia is 91.9%, which is also the highest figure among all presented countries; the closest results to ours were obtained in Germany (87.6%) and Great Britain (81.3%) [19]. On the other hand, the minimum frequency of germline variants in Russia is described both in the *MSH6* gene (7.4%) and in the *PMS2* gene (0.7%), the combined frequency of which ranges from 12.4% in Germany to 51.8% in the Netherlands (*Figure* 2) [19, 22]. Nevertheless, the obtained results should be taken into account when conducting DNA diagnostics in Russian patients in order to make it as effective as possible.

Interestingly, among patients with colorectal cancer against the background of Lynch syndrome caused by a pathogenic variant in the *MSH2* gene in Russia, male patients predominated (35:22). At the same time, in cases where Lynch syndrome was caused by pathogenic variants in other genes, the frequency of cancer patients in men and women did not differ: 36 men and 33 women had mutations in the *MLH1* gene, 5 men and 5 women in *MSH6*; and only a total of 3 women and 2 men had mutations in *PMS2*,

PMS1 and *EPCAM* genes. These data may indicate a higher risk of colon cancer in men carrying pathogenic variants in the *MSH2* gene.

It is extremely important to emphasize that the current CanVIG recommendations [13, 14] do not take into account cases of sebaceous skin carcinomas in the family and personal anamnesis of carriers of germinal variants in the *MSH2* gene, while their presence allows diagnosing Muir-Torre syndrome, which is a special variant of Lynch syndrome. We believe that it is appropriate to correct this situation in the updated version of these recommendations.

Conclusion

In conclusion, it should be noted that the analysis of data from Russian patients with Lynch syndrome allowed establishing that pathogenic and likely pathogenic variants are most often found in the *MLH1* and *MSH2* genes (91.9% of cases). Therefore, to conduct the most effective DNA diagnostics in patients suspected of having this syndrome, it is necessary to begin diagnostics with the study of these two genes. The frequency of major rearrangements in the MMR/*EPCAM* genes was 9.2%, which indicates the need to include the MLPA method in routine DNA diagnostics of patients. Germinal variants previously undescribed in the world population were found in 12.1% of probands, demonstrating the presence of population characteristics in Russian patients with Lynch syndrome.

Author contributions

AT: designed and conducted the study, analyzed data, prepared the draft manuscript. AB, VS and AL: performed genetic analysis and analyzed genetic data. DP: collected clinical data, identified patients, critically reviewed the draft manuscript, corrected the translated version of the manuscript. DS: critically reviewed the manuscript. YS and SA: supervised the project, reviewed the manuscript. All authors contributed to the study concept and design, as well as read and approved the final version of the manuscript.

Conflict of interest

The authors declare that no conflicts of interest pertaining to this study.

Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the Ethics Committee at Ryzhikh National Medical Research Centre for Coloproctology and with 1964 Declaration of Helsinki and its later amendments.

References

1. Haraldsdottir S, Rafnar T, Frankel WL, Einarsdottir S, Sigurdsson A, Hampel H, et al. Comprehensive population-wide analysis of Lynch syndrome in Iceland reveals founder mutations in MSH6 and PMS2. *Nat Commun* 2017; 8: 14755. https://doi.org/10.1038/ncomms14755.
2. Seppälä TT, Latchford A, Negoi I, Sampaio Soares A, Jimenez-Rodriguez R, Sánchez-Guillén L, et al. European guidelines from the EHTG and ESCP for Lynch syndrome: An updated third edition of the Mallorca guidelines based on gene and gender. *Br J Surg* 2021; 108(5): 484-498. https://doi.org/10.1002/bjs.11902.
3. Moreira L, Balaguer F, Lindor N, de la Chapelle A, Hampel H, Aaltonen LA, et al. Identification of Lynch syndrome among patients with colorectal cancer. *JAMA* 2012; 308(15): 1555-1565. https://doi.org/10.1001/jama.2012.13088.
4. Geiersbach KB, Samowitz WS. Microsatellite instability and colorectal cancer. *Arch Pathol Lab Med* 2011; 135(10): 1269-1277. https://doi.org/10.5858/arpa.2011-0035-RA.
5. Tutlewska K, Lubinski J, Kurzawski G. Germline deletions in the EPCAM gene as a cause of Lynch syndrome – literature review. *Hered Cancer Clin Pract* 2013; 11(1): 9. https://doi.org/10.1186/1897-4287-11-9.
6. Shelygin YuA, Achkasov SI, Semenov DA, Sushkov OI, Shakhmatov DG, Romanova EM, et al. Genetic and phenotypic characteristics of 60 Russian families with Lynch syndrome. *Koloproktologia* 2021; 20(3): 35-42. https://doi.org/10.33878/2073-7556-2021-20-3-35-42.
7. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009; 25(14): 1754-1760. https://doi.org/10.1093/bioinformatics/btp324.
8. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. Gigascience 2021; 10(2): giab008. https://doi.org/10.1093/gigascience/giab008.
9. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010; 20(9): 1297-1303. https://doi.org/10.1101/gr.107524.110.
10. Poplin R, Chang PC, Alexander D, Schwartz S, Colthurst T, Ku A, et al. A universal SNP and small-indel variant caller using deep neural networks. *Nat Biotechnol* 2018; 36(10): 983-987. https://doi.org/10.1038/nbt.4235.
11. Jiang Y, Wang R, Urrutia E, Anastopoulos IN, Nathanson KL, Zhang NR. CODEX2: Full-spectrum copy number variation detection by high-throughput DNA sequencing. *Genome Biol* 2018; 19(1): 202. https://doi.org/10.1186/s13059-018-1578-y.
12. Miller DT, Lee K, Abul-Husn NS, Amendola LM, Brothers K, Chung WK, et al. ACMG SF v3.1 list for reporting of secondary findings in clinical exome and genome sequencing: A policy statement of the American College of Medical Genetics and Genomics (ACMG). *Genet Med* 2022; 24(7): 1407-1414. https://doi.org/10.1016/j.gim.2022.04.006.
13. Garrett A, Callaway A, Durkie M, Cubuk C, Alikian M, Burghel GJ, et al. Cancer Variant Interpretation Group UK (CanVIG-UK): An exemplar national subspecialty multidisciplinary network. *J Med Genet* 2020; 57(12): 829-834. https://doi.org/10.1136/jmedgenet-2019-106759.
14. CanVIG-UK Gene Specific Recommendations. https://www.cangene-canvaruk.org/gene-specific-recommendations.
15. Pérez-Carbonell L, Ruiz-Ponte C, Guarinos C, Alenda C, Payá A, Brea A, et al. Comparison between universal molecular screening for Lynch syndrome and revised Bethesda guidelines in a large population-based cohort of patients with colorectal cancer. *Gut* 2012; 61(6): 865-872. https://doi.org/10.1136/gutjnl-2011-300041.
16. Tsukanov AS, Shelygin YuA, Achkasov SI, Frolov SA, Kashnikov VN, Kuzminov AM, et al. Principles of diagnosis and personalized treatment of hereditary colorectal cancer. *Annals of Russian Academy of Medical Sciences* 2019; 74(2): 118-124. https://doi.org/10.15690/vramn1083.
17. Tsukanov AS, Pikunov DYu, Shubin VP, Barinov AA, Kashnikov VN, Shelygin YA, et al. Unique combination of Diamond–Blackfan anemia and Lynch syndrome in adult female: A case report. *Front Oncol* 2021; 11: 652696. https://doi.org/10.3389/fonc.2021.652696.
18. Dong L, Jin X, Wang W, Ye Q, Li W, Shi S, et al. Distinct clinical phenotype and genetic testing strategy for Lynch syndrome in China based on a large colorectal cancer cohort. *Int J Cancer* 2020; 146(11): 3077-3086. https://doi.org/10.1002/ijc.32914.
19. International Mismatch Repair Consortium. Variation in the risk of colorectal cancer in families with Lynch syndrome: A retrospective cohort study. *Lancet Oncol* 2021; 22(7): 1014-1022. https://doi.org/10.1016/S1470-2045(21)00189-3.
20. Lagerstedt-Robinson K, Rohlin A, Aravidis C, Melin B, Nordling M, Stenmark-Askmalm M, et al. Mismatch repair gene mutation spectrum in the Swedish Lynch syndrome population. *Oncol Rep* 2016; 36(5): 2823-2835. https://doi.org/10.3892/or.2016.5060.

21. Coughlin SE, Heald B, Clark DF, Nielsen SM, Hatchell KE, Esplin ED, et al. Multigene panel testing yields high rates of clinically actionable variants among patients with colorectal cancer. *JCO Precis Oncol* 2022; 6: e2200517. https://doi.org/10.1200/po.22.00517.
22. Goverde A. Lynch Syndrome Improving Diagnostics and Surveillance: Thesis. Erasmus University Rotterdam Dissertation. Optima Grafische Communicatie B.V., Netherlands. 2018; 194 p. https://core.ac.uk/download/pdf/158601381.pdf.

Authors:
Aleksey S. Tsukanov – MD, PhD, Principal Researcher, Laboratory of Genetics, Ryzhikh National Medical Research Centre for Coloproctology, Moscow, Russia. https://orcid.org/0000-0001-8571-7462.
Aleksey A. Barinov – Researcher, Laboratory of Genetics, Ryzhikh National Medical Research Centre for Coloproctology, Moscow, Russia. https://orcid.org/0000-0002-1443-960X.
Vitaly P. Shubin – MD, PhD, Senior Researcher, Laboratory of Genetics, Ryzhikh National Medical Research Centre for Coloproctology, Moscow, Russia. https://orcid.org/0000-0002-3820-7651.
Anna N. Loginova – MD, Researcher, Laboratory of Genetics, Ryzhikh National Medical Research Centre for Coloproctology, Moscow, Russia. https://orcid.org/0000-0002-7248-111X.
Dmitry Yu. Pikunov – MD, PhD, Senior Researcher, Department of Coloproctology, Ryzhikh National Medical Research Centre for Coloproctology, Moscow, Russia. https://orcid.org/0000-0001-7040-6979.
Dmitry G. Shakhmatov – MD, PhD, Head of the Department of Colorectal Surgery, Ryzhikh National Medical Research Centre for Coloproctology; Associate Professor, Department of Coloproctology, Russian Medical Academy of Continuing Professional Development, Russian Federation Ministry of Healthcare, Moscow, Russia. https://orcid.org/0000-0001-7964-2126.
Yuri A. Shelygin – MD, PhD, Professor, Scientific Director, Ryzhikh National Medical Research Centre for Coloproctology; Academician, Russian Academy of Sciences; Head of the Department of Coloproctology, Russian Medical Academy of Continuing Professional Development, Russian Federation Ministry of Healthcare, Moscow, Russia. https://orcid.org/0000-0002-8480-9362.
Sergey I. Achkasov – MD, PhD, Professor, Director, Ryzhikh National Medical Research Centre for Coloproctology; Corresponding Member, Russian Academy of Sciences, Moscow, Russia. https://orcid.org/0000-0001-9294-5447.