

# Next Generation Sequencing in Molecular Diagnosis of Lynch Syndrome – a Pilot Study Using New Stratification Criteria

Ivana Kašubová<sup>1</sup>, Veronika Holubeková<sup>1</sup>, Katarína Janíková<sup>1,3</sup>, Barbora Váňová<sup>1,2</sup>, Zuzana Sňahničanová<sup>1</sup>, Michal Kalman<sup>3</sup>, Lukáš Plank<sup>3</sup>, Zora Lasabová<sup>1,2,\*</sup>

## ABSTRACT

The development of the new technologies such as the next-generation sequencing (NGS) makes more accessible the diagnosis of genetically heterogeneous diseases such as Lynch syndrome (LS). LS is one of the most common hereditary form of colorectal cancer. This autosomal dominant inherited disorder is caused by deleterious germline mutations in one of the mismatch repair (MMR) genes – MLH1, MSH2, MSH6 or PMS2, or the deletion in the EPCAM gene. These mutations eventually result in microsatellite instability (MSI), which can be easily tested in tumor tissue. According to the actual recommendations, all patients with CRC that are suspect to have LS, should be offered the MSI testing. When the MSI is positive, these patients should be recommended to genetic counseling. Here we report a pilot study about the application of NGS in the LS diagnosis in patients considered to have sporadic colorectal cancer. The inclusion criteria for the NGS testing were MSI positivity, BRAF V600E and MHL1 methylation negativity. We have used 5 gene amplicon based massive parallel sequencing on MiSeq platform. In one patient, we have identified a new pathogenic mutation in the exon 4 of the MSH6 gene that was previously not described in ClinVar, Human Gene Mutation Database, Ensembl and InSight databases. This mutation was confirmed by the Sanger method. We have shown that the implementation of new criteria for colorectal patients screening are important in clinical praxis and the NGS gene panel testing is suitable for routine laboratory settings.

## KEYWORDS

sporadic colorectal cancer; microsatellite instability; Lynch syndrome; MMR genes; next generation sequencing

## AUTHOR AFFILIATIONS

<sup>1</sup> Division of Oncology, Biomedical Center Martin, Comenius University in Bratislava, Jessenius Faculty of Medicine in Martin, Slovakia

<sup>2</sup> Department of Molecular Biology, Comenius University in Bratislava, Jessenius Faculty of Medicine in Martin, Slovakia

<sup>3</sup> Department of Pathological Anatomy, Slovakia, Comenius University in Bratislava, Jessenius Faculty of Medicine University Hospital in Martin, Slovakia

\* Corresponding author: Biomedical Center Martin JFM CU, Malá Hora 4C, 03601 Martin, Slovakia; e-mail: lasabova@jfm.uniba.sk

Received: 15 May 2018

Accepted: 10 September 2018

Published online: 12 December 2018

Acta Medica (Hradec Králové) 2018; 61(3): 98–102

<https://doi.org/10.14712/18059694.2018.125>

© 2018 The Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

## INTRODUCTION

Colorectal cancer (CRC) is a multifactorial disease with strong genetic background. Regarding inheritance, we differentiate between sporadic, familial and hereditary forms. One of the most common hereditary forms is Lynch syndrome (LS). LS is autosomal dominant inherited disorder caused by deleterious germline mutations in one of the mismatch repair (MMR) genes – MLH1, MSH2, MSH6, or PMS2, and the deletion within EPCAM gene or by epimutations in MLH1 gene (1, 2). Early onset of the disease and multiple tumors with microsatellite instability (MSI) are typical in the affected patients. When the patients are suspected for LS according to Amsterdam criteria II (AM) or Revised Bethesda guidelines (RBG) and are positive for microsatellite instability (MSI), testing for germline mutations in candidate genes is recommended after genetic counseling (3). However, it is widely accepted that these diagnostic criteria are suboptimal for the detection of LS. The detection of MSI and immunohistochemistry (IHC) of MMR proteins have to be used prior to genetic testing (3). In the first published study that used MSI as the primary screening method followed by IHC staining for MMR protein expression in 1066 patients, 19.5% expressed MSI and approximately 2.2% harbored mutations causing LS (4). Moreover, other authors reported that MLH1 promoter hypermethylation and BRAF V600E mutation distinguishes the hereditary non – polyposis colon cancer from sporadic MSI-H positive colon cancer (5) and can be used in patient stratification. Because LS is under-recognized in the population, a universal tumor screening (UTS) approach for LS, based on MSI testing and/or IHC for expression of MMR proteins, is recommended by several professional societies (6). Recently, the sensitivity and specificity of MSI and/or IHC testing for the identification of LS was evaluated. The authors concluded that this is an effective screening test for the identification of LS suspect patients (7). The recommendations are also included within the methodical guidelines by the Slovak Association of Medical Genetics (<http://www.sslg.sk/index.php/dokumenty/metodicke-pokyny/90-metodicke-usmernenie-lynch>).

The conventional molecular testing for pathogenic mutations in LS is focusing on the detection of single nucleotide variants (SNVs) and the screening for deletions or duplications by the Sanger sequencing and the multiplex ligand probe dependent amplification (MLPA), respectively. The dideoxysequencing according Sanger is based on a chain-termination using fluorescently labeled dideoxynucleotides and the capillary electrophoresis, which makes it to a robust and reliable, but also laborious and expensive method. The main limitation is the number of sequenced base pairs in the one sequencing run (8). The next-generation sequencing (NGS) based on massive parallel sequencing has revolutionized the molecular diagnostics in numerous clinical settings by their capacity to sequence several millions basepairs in one sequencing run. NGS technology substantially increased the number of sequenced genes, lowered the costs and enhanced the sensitivity of the mutation detection. A new 5-tiered nomenclature for estimation of variant's pathogenicity was introduced, which is very helpful especially in the case

of missense substitutions. According this classification, the variants are defined as non-pathogenic (class 1), likely non-pathogenic (class 2), variants of uncertain clinical significance (VUS, class 3), likely pathogenic (class 4) and pathogenic (class 5) (9). Nowadays, the application of gene panels is a standard approach for the molecular diagnosis of genetically heterogeneous diseases, including LS, because all candidate genes can be sequenced in one run (10).

In order to identify the germline mutations, the massive parallel sequencing on Miseq platform was applied with the panel of mismatch repair genes – the MLH1, MSH2, MSH6, PMS2 and 3'UTR of EPCAM gene. The aim of our study was the implementation of NGS in order to identify mutations in MMR genes in patients without family history considered to have sporadic colorectal cancer, which are positive for MSI-H and negative for both, BRAF V600E and MLH1 hypermethylation.

## MATERIALS AND METHODS

### SPECIMENS

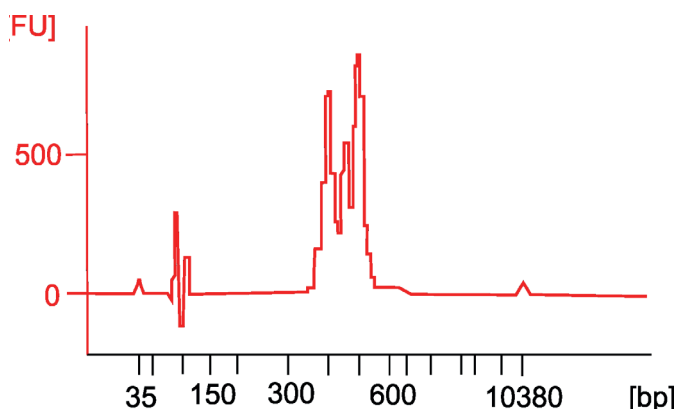
Blood samples were obtained from 4 patients with colorectal carcinoma suspected to have the LS. The patients were positive for MSI and negative for BRAF V600E and also negative for MLH1 methylation, as described previously (11, 12, 13). The formalin-fixed and paraffin-embedded (FFPE) sections were analyzed using IHC for the presence of the MMR protein expression (MLH1, MSH2, MSH6, and PMS2) according the standard protocol for IHC staining (13). All tested patients were recommended for a genetic counseling. The study was approved by the Ethical Committee and the patients signed informed consent. The peripheral blood cells were isolated by the DNeasy Blood and Tissue kit (Qiagen, USA), according to the manufacturer's protocol. The concentration of DNA was measured using Qubit™ dsDNA BR Assay Kit on Qubit 2.0 Fluorometer (Thermo Fisher Scientific, USA) and diluted to a working concentration of 10 ng/μl.

### NGS SEQUENCING ON MISEQ

#### NGS library preparation

For the target enrichment of coding regions and the flanking intronic sequences, all exons of MLH1, MSH2, MSH6, PMS2 and 3'UTR of EPCAM gene were amplified using HNPCC MASTR Plus kit (Multiplicom, Belgium) in five multiplex PCRs. The multiplex PCR was performed using 50 ng of genomic DNA per reaction. The amplicons (84) cover four MMR genes and 3'UTR of the EPCAM gene. The PCR products were purified with Agencourt AM Pure XP beads (Beckman Coulter, USA) and diluted 1/1000. The diluted PCR products were ligated with molecular identifiers MID from Illumina Miseq kit (1–48) (Multiplicom, Belgium). In a universal PCR, 2 μl of the ligated PCR products and 48 μl of universal master mix reaction were used. The PCR products (469bp) were analyzed with the High Sensitivity DNA 1000 kit (Agilent technologies, USA) (Fig. 1) as described previously (14). The concentration of obtained tagged amplicons was determined with Qubit™ dsDNA HS Assay kit (Thermo Fisher Scientific, USA), after repeated

purification of PCR products with Agencourt AM Pure XP beads (Beckman Courter, CA). The amplicon's libraries were diluted to 10 nM, pooled equimolar at a 2 nmol/L concentration and denatured.



**Fig. 1** Electrophoresis on the chip after universal PCR showing successful library preparation. The amplicon size after universal PCR should be between 350 to 550 bp in blood samples. Electrophoresis was prepared at Agilent 2100 Bioanalyzer (Agilent).

#### NGS

Amplicon pool was processed by bridge amplification and sequenced by Miseq Reagent Nano kit v2 (500 cycles) (Illumina, USA) with 500Mb capacity and the read length  $2 \times 250$  bp on Miseq (Illumina, USA) platform. Analysis was conducted according to the manufacturer's instructions. The coverage of the region of interest (ROI) was indicated by the manufacturer as 100%.

#### Bioinformatic evaluation

The NGS generated for each cluster a forward and a reverse read. Read1 and Read2 are defined as read pairs. The sequencing run was valid when both Read1 and Read2 had at least 220 bp. The bioinformatic evaluation comprised of three parts - alignment of obtained reads to .bam files (Assembly GRCh37), secondary data analysis getting the variant annotation in the .vcf files and tertiary data analysis resulting in interpretation of variants. The sequence alignment was performed with MiSeq Reporter Software v2.5.1.3 (Illumina, USA) and the .vcf files were generated with Illumina Variant Studio Software v 3.0 (Illumina, USA). Sequences were annotated starting with the first nucleotide that corresponds for the first A in the first translated amino acid in coding reference cDNA. The cDNA reference sequences for MHL1, MSH2, MSH6 and PMS2 were NM\_000249.3, NM\_000251.2, NM\_000179.2 and NM\_000535.5, respectively. The potential pathogenic variant was compared with free accessible database ClinVar, Human Gene Mutation Database (<http://www.hgmd.cf.ac.uk/ac/all.php>), Ensembl ([https://www.ensembl.org/Homo\\_sapiens/Transcript/ProtVariations/](https://www.ensembl.org/Homo_sapiens/Transcript/ProtVariations/)), the InSight databases (<https://www.insight-group.org/variants/databases/>) and published literature. The presence of pathogenic variant was confirmed by Sanger sequencing.

#### SANGER SEQUENCING VALIDATION

The forward primer MSH6 F and the reverse primer MSH6 R were used for the Sanger method (Table 1). The PCR products were placed into a thermal cycler with the following program: 95 °C denaturation for 10 min, cycling was repeated 30× with 95 °C for 30 sec, annealing at 60 °C for 30 sec and extension at 72c for 30 sec. The PCR products were purified using NucleoSpin Gel and PCR Clean-up (Macherey-Nagel, Germany). After purification, the sequencing PCR reaction was performed using BigDye Terminator v 1.1 Cycle Sequencing Kit (ThermoFisher Scientific, USA) according to the manufacturer's instructions. The sequenced fragments were purified with SigmaSpin Sequencing Reaction Clean-up (Sigma-Aldrich, USA). The purified products were denatured in deionized Hi-Di Formamid (Thermo Fisher Scientific, USA) and sequenced on the ABI 3500 Genetic Analyzer (Applied Biosystems, USA). The data were analyzed using Chromas Pro Software (Technelysium Pty Ltd, Australia).

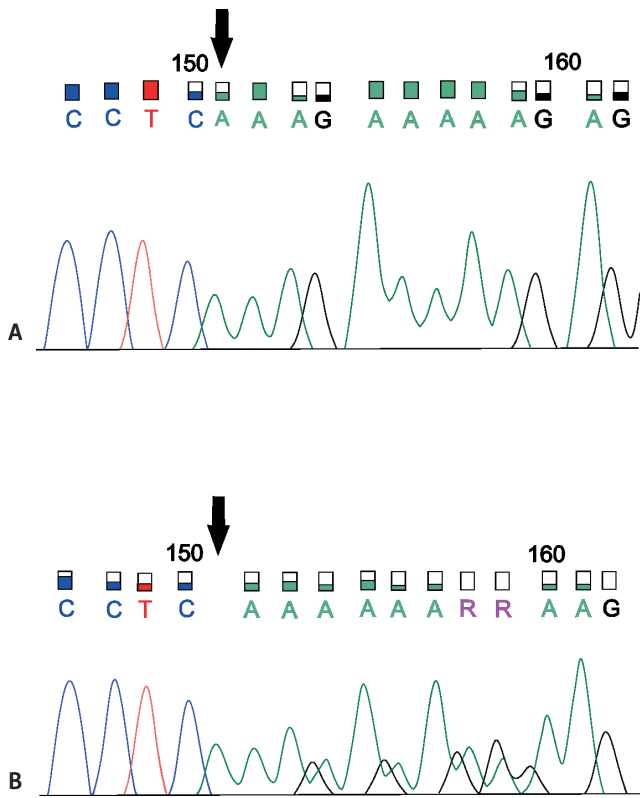
**Tab. 1** Primer sequences for the sequencing of the region of interest in the exon 4 of MSH6 gene and the size of the final PCR product.

Primer	Sequence	Annealing	Size (bp)
MSH6 F	GCACGAGTGGAACAGACTGA	60 °C	339
MSH6 R	TTGTACTGGGGGATAGTGTGC	60 °C	

#### RESULTS

Using the HNPCC MASTR Plus Kit (Multiplicom, Belgium) and NGS, we identified pathogenic variants in two patients. The value of the Q30 score for run was 78.46%. We identified 95.53% of reads and only 4.47% of reads remained undetermined. In the first sample, we detected a 4 bp deletion in the exon 4 of the MSH6 gene c.1627\_1630AAAAG (Assembly GRCh37). This mutation is within the protein coding sequence resulting into a frameshift at the protein level (p. Glu544Lysfs\*26). The termination codon lies 26 nucleotide downstream of the deletion and was confirmed by the Sanger method (Fig. 2). This mutation was not recorded in any database we mentioned in the section Methods. A very similar pathogenic mutation NM\_000179.2 (MSH6):c.1632\_1635delAAAA (p.Lys545Argfs) was reported in the ClinVar database (assembly GRCh38), in Ensembl (rs267608064) and InSight (15, 16) and was assigned into the class 5 of the 5-tiered scale. Moreover, the sequencing of DNA obtained from the peripheral blood of patients with Muir-Torre syndrome identified a new 4 bp deletion within MSH6 gene c.1634\_1637delAAGA. This mutation causes a premature stop gain 24 codons downstream, is considered pathogenic (17) and registered in HGMD, ClinVar and Ensembl. Therefore, we consider the mutation c.1627\_1630AAAAG (p.Glu544Lysfs) pathogenic. Our result was confirmed also by the certified commercial diagnostic laboratory, which recommended also testing of asymptomatic descendants of the patient. In the second sample, we identified a pathogenic stop gain variant in MSH2 gene c.1030C>T (Gln344Ter). We were not able to verify this by

the Sanger sequencing. According to a five-tiered classification of the International Agency for Research on Cancer (on Assembly GRCh37), this variant has high impact on the MSH2 gene and we recommended a verification by a commercial laboratory, which result is unknown to us. In other two samples tested, we detected variants, which were classified as benign.



**Fig. 2** A – Validation of the mutation using Sanger sequencing with deletion 1627\_1630del AAAG (Glu544Lysfs\*26); B – reference DNA sequence.

**DISCUSSION**

The development of massive parallel sequencing allowed to perform whole-genome, whole-exome and gene panel sequencing in different clinical settings. The method of choice for a lot of clinical applications including the diagnostics of LS, is the use of gene panels with the defined number of candidate genes known to be mutated in this particular disease. NGS technologies provide useful tools for the detection of single-nucleotide variants (SNV) in many genes simultaneously (10). The other advantage of NGS is that it requires, in contrast to traditional sequencing methods, very low input of DNA/RNA (18). In our study, we used 250 ng of DNA for the sequencing of 84 amplicons. Nowadays, the conventional Sanger sequencing is used mostly for the confirmation of NGS results. Here we report the application of NGS in diagnostics of patients with Lynch syndrome using the commercially available gene panel with four MMR genes (MLH1, MSH2, MSH6 and PMS2) and 3' UTR of the EPCAM gene. This work was a part of the study, which has developed a patient's stratification algorithm for LS. A similar algorithm is already

recommended by the Slovak Association of Medical Genetics (<http://www.sslg.sk/index.php/dokumenty/metodicka-pokyny/90-metodicke-usmernenie-lynch>). We have used the gene panel included in the HNPCC MASTR Plus Kit (Multiplicom, Belgium) and MiSeq NGS platform. We identified two pathogenic variants, one of them was confirmed by Sanger method by us and also in a commercial laboratory. The identified mutation c.1627\_1630AAAG (p.Glu544Lysfs) is a deletion of 4 bp resulting in a downstream termination of protein synthesis. It was not described in any database, which we used for a result interpretation. However, similar 4 bp deletion resulting in preterm stop codon was described in this gene region involving codons 543 and 544 in patients with LS (15, 16) and was classified as pathogenic (class 5) in the database of the InSight consortium (9). Another similar variant was identified in codon 544 in DNA of a patient with Muir-Torre syndrome (17). Therefore we consider this variant pathogenic.

The mutations in the MSH6 gene comprise 18% of mutations in MMR genes predisposing to LS ([www.insight-group.org](http://www.insight-group.org)). Compared with pathogenic variants in MLH1 and MSH2 genes, germline variants in MSH6 gene are mostly associated with mild and variable phenotypes. The onset of the disease symptoms among mutation carriers is after age 50 and the incidence of colorectal cancer is lower than in other patients with LS. The patients often fulfill neither the AM nor RBG (19, 20). Our patient had no family history and colorectal cancer was diagnosed at age 64. One explanation of this is that MSH6 mutation could cause functional redundancy of MSH6 protein. This protein can be particularly substituted by MSH3 in the heteromeric MutSα complex (21).

The NGS gene panel testing allows rapid identification of mutations in genetically heterogeneous diseases as we have shown in our small pilot study. There are several targeted gene panels for inherited cancer syndrome that are commercially available and are best suited for the routine laboratory praxis.

**CONCLUSION**

Here we report the application of targeted NGS in the diagnostics of Lynch syndrome in patients considered to have sporadic CRC and stratified using MSI test, BRAF V600E determination and MLH1 methylation screening. We have identified a new pathogenic variant in a patient without family history of the disease. Moreover, we have shown the implementation of the MSI test, the evaluation of BRAF V600E and MLH1 methylation in patients considered to have sporadic CRC are important in clinical praxis. Our aim was to address the need for more clearly stratification of patients suspected for LS. We recommend utilization of the MSI screening strategy for all patients with colorectal cancer, not only for patients with positive family history, and subsequent NGS sequencing in MSI-H, BRAF V600E and MLH1 methylation negative cases. It would be also appropriate to include this mutation c.1627\_1630AAAG (p.Glu544Lysfs) into further studies.

## ACKNOWLEDGMENTS

We thank MUDr. Jana Kršiaková for obtaining informed consents from patients and professional consultations. This work was supported by the Biomedical Center Martin (IMTS 26220220187), which is co-financed from EU sources, and by the projects of the Slovak Research and Development Agency no. APVV-14-0273, APVV-16-0066 and VEGA 1/0380/18.

## REFERENCES

- Niessen RC, Hofstra RM, Westers H, Ligtenberg MJ, Kooi K, Jager PO. Germline hypermethylation of MLH1 and EPCAM deletions are a frequent cause of Lynch syndrome. *Genes Chromosomes Cancer* 2009; 48(8): 373–744.
- Peltomäki P. Update on Lynch syndrome genomics. *Familial Cancer* 2016; 15: 385–93.
- Ramsoekh D, Van Leerdam NE, Wagner A, Kuipers EJ. Review article: Detection and management of hereditary non-polyposis colorectal cancer (Lynch syndrome). *Aliment Pharmacol Ther* 2007; 2: 101–11.
- Hampel H, Frankel WL, Martin E, et al. Screening for the Lynch syndrome (hereditary nonpolyposis colorectal cancer). *New England Journal of Medicine* 2005; 352(18): 1851–60.
- McGivern A, Wynter CV, Whitehall VL, et al. Promoter hypermethylation frequency and BRAF mutations distinguish hereditary non-polyposis colon cancer from sporadic MSI-H colon cancer. *Familial Cancer* 2004; 3(2): 101–7.
- Cohen SA, Laurino M, Bowen DJ, et al. Initiation of universal tumor screening for lynch syndrome in colorectal cancer patients as a model for the implementation of genetic information into clinical oncology practice. *Cancer* 2016; 122(3): 393–401.
- Provenzale D, Gupta S, Ahnen DJ, et al. Genetic/familial high-risk assessment: colorectal version 1.2016. *NCCN Clinical Practice Guidelines in Oncology. Journal of the National Comprehensive Cancer Network* 2016; 14(8): 1010–30.
- Dovichi NJ. DNA sequencing by capillary electrophoresis. *Electrophoresis* 1997; 18(12–13): 2393–9.
- Thompson BA, Spurdle AB, Plazzer JP. Application of a five-tiered scheme for standardized classification of 2,360 unique mismatch repair gene variants lodged on the InSight locis specific database. *Nature Genetics* 2014; 46: 107–15.
- Rey JM, Ducros V, Pujol P, et al. Improving mutation screening in patients with colorectal cancer predisposition using next-generation sequencing. *Journal of Molecular Diagnostics* 2017; 19(4): 589–601.
- Kašubová I, Jašek K, Lasabová Z, Kalman M, Plank L. Detekcia mikrosatelitnej instability a mutácie V600E v géne BRAF u pacientov s kolorektálnym karcinómom. In: *Molekulová biológia v medicíne 3. Prehľady, protokoly, aplikácie*. Univerzita Komenského v Bratislave, Jesseniova lekárska fakulta v Martine, ed. Lasabová Z, 2015, 10–17.
- Jasek K, Buzalkova V, Szepe P, Plank L, Lasabova Z. Limits of dideoxysequencing in the detection of somatic mutations in gastrointestinal stromal tumors. *Acta Medica Martiniana* 2015; 15(3): 13–20.
- Jasek K, Buzalkova V, Minarik G, et al. Detection of mutations in the BRAF gene in patients with KIT and PDGFA wild-type gastrointestinal stromal tumors. *Virchows Archives* 2017; 470: 29–36.
- Vazan M, Kasubova I, Vanochova A, Lukac P, Plank L, Lasabova Z. Implementation of microfluidic chip electrophoresis for the detection of B-cell clonality. *Acta Medica Martiniana* 2016; 16: 14–23.
- Plaschke J, Kruger S, Dietmaier W, et al. Eight novel MSH6 germline mutations in patients with familial and nonfamilial colorectal cancer selected by loss of protein expression in tumor tissue. *Human Mutation* 2004; 23(3): 285.
- Steinke V, Rahner N, Morak M, et al. No association between MUYH and MSH6 germline mutations in 64 HNPCC patients. *European Journal of Human Genetics* 2008; 16: 587–92.
- Arnold A, Payne S, Fisher S, et al. An individual with Muir-Torre syndrome found to have a pathogenic MSH6 gene mutation. *Familial Cancer* 2007; 6: 317–21.
- Serratì S, De Summa S, Pilato B, et al. Next-generation sequencing: advances and applications in cancer diagnosis. *Oncology Targets Ther* 2016; 9: 7355–65.
- Hendriks YM, Wagner A, Morreau H, et al. Cancer risk in hereditary nonpolyposis colorectal cancer due to MSH6 mutations: impact on counseling and surveillance. *Gastroenterology* 2004; 127: 17–25.
- Plaschke J, Engel C, Kruger S. Lower incidence of colorectal cancer and later age of disease onset in 27 families with pathogenic MSH6 germline mutations compared with families with MLH1 or MSH2 mutations: the German Hereditary Non-polyposis Colorectal Cancer Consortium. *Journal of Clinical Oncology* 2004; 22: 4486–94.
- Peltomäki P. Lynch syndrome genes. *Familial Cancer* 2005; 4(3): 227–32.