The Journal of Molecular Diagnostics, Vol. 15, No. 3, May 2013

the Journal of Molecular Diagnostics
jmd.amjpathol.org

Evaluating the Effect of Unclassified Variants Identified in MMR Genes Using Phenotypic Features, Bioinformatics Prediction, and RNA Assays

Lucia Pérez-Cabornero,* Mar Infante,* Eladio Velasco,* Enrique Lastra,† Cristina Miner,* and Mercedes Durán*

From the Cancer Genetics Laboratory, Institute of Biology and Molecular Genetics, University of Valladolid, Valladolid, Spain; and the Department of Oncology,† Hospital General Yagüe, Burgos, Spain*

CME Accreditation Statement: This activity ("JMD 2013 CME Program in Molecular Diagnostics") has been planned and implemented in accordance with the Essential Areas and policies of the Accreditation Council for Continuing Medical Education (ACCME) through the joint sponsorship of the American Society for Clinical Pathology (ASCP) and the American Society for Investigative Pathology (ASIP). ASCP is accredited by the ACCME to provide continuing medical education for physicians.

The ASCP designates this journal-based CME activity ("JMD 2013 CME Program in Molecular Diagnostics") for a maximum of 48 AMA PRA Category 1 Credit(s)™. Physicians should only claim credit commensurate with the extent of their participation in the activity.

CME Disclosures: The authors of this article and the planning committee members and staff have no relevant financial relationships with commercial interests to disclose.

Accepted for publication February 1, 2013.

Address correspondence to Mercedes Durán, M.D., Cancer Genetics, IBGM-CSIC-UVA, C/Sanz y Forés 3, 47003, Valladolid, Spain. E-mail: merche@ibgm.uva.es.

Lynch syndrome is caused by mutations in one of the mismatch-repair system (MMR) genes. A major difficulty in diagnosis and management of Lynch syndrome is the existence of unclassified genetic variants (UVs) with unknown clinical significance, especially mutations with new descriptions and missense-type nucleotide substitutions. We evaluated the pathogenicity of 20 such mutations (6 in *MLH1*, 4 in *MSH2*, and 7 in *MSH6*) found in Spanish patients suspected of Lynch syndrome. The UVs were tested for evidence of MMR defect in tumor samples and were evaluated for co-occurrence with a pathogenic mutation, the cosegregation of the variant with the disease; where sufficient data were available, *in silico* resources at the protein level and mRNA analysis were used to assess the putative effect on the splicing mechanism. To evaluate the frequency of these UVs in the general population, a case–control study was also performed. Five variants were identified with similar frequencies in both cases and controls, suggesting a nonpathogenic effect in patients. In contrast, abnormal splicing mutations were detected in a high proportion of patients [3/20 (15%)]. In this study, we classified 15 of the 20 UVs: six variants with strong evidence of pathogenicity and nine variants that should be considered neutral variants. Clinical significance of the other five remains unknown. *(J Mol Diagn 2013, 15: 380–390; http://dx.doi.org/10.1016/j.jmoldx.2013.02.003)*

Hereditary non–polyposis colorectal cancer (HNPCC) or Lynch syndrome is the most frequent autosomal dominant colorectal cancer susceptibility syndrome caused by mutations inactivating one of the genes of the mismatch-repair system (MMR), most frequently *MLH1* (OMIM 120436, NM_000249.3) and *MSH2* (OMIM 609309; NM_000251.1), and less often in *MSH6* (OMIM 600678; NM_000179.2) and *PMS2* (OMIM 600259; NM_000535).[1–4] The phenotype of tumors from these patients is characterized by widespread microsatellite instability (MSI) and loss of protein expression from the affected enzyme, which can be detected by immunohistochemical staining (IHC). This syndrome is characterized by a high risk of early onset of colorectal cancer and several other, extracolonic malignant tumors, especially endometrial cancer in women.[5]

Most of the genetic defects in the human MMR genes responsible for Lynch syndrome are a result of point mutations and small insertions and deletions[6] that truncate

Supported by the regional government of Castile and León, Spain. L.P.-C. was supported by a predoctoral FPI fellowship from the regional government of Castile and León.

http://dx.doi.org/10.1016/j.jmoldx.2013.02.003

and inactivate MMR genes. Mutations in three of these MMR genes (*MLH1*, *MSH2*, and *MSH6*) account for the majority of the patients with Lynch syndrome.[7] Nonetheless, there are unclassified genetic variants (UVs) with unknown clinical significance; these are nucleotide substitutions (generally not truncating missense type) whose clinical interpretation can be difficult when detected in a family with suspected Lynch syndrome. In the MMR genes, these mutations account for the most common type of mutation detected, and missense-type mutations occur in 24% of all unique variants detected in *MLH1*, 17% in *MSH2*, and 27% in *MSH6*, as estimated from the Mismatch Repair Genes Variant Database[8] (*http://www.med.mun.ca/MMRvariants*, last accessed June 30, 2012).

The pathogenicity of a sequence variation is classically determined based on several different lines of evidence, such as the cosegregation of the mutation with the cancer phenotype and the MSI and IHC status of the tumors of the mutation carriers,[9,10] co-occurrence (in trans) with deleterious mutations, the determination of the variant frequency in unaffected controls, amino acid polarity or size, and evolutionary conservation of the residue. However, the clinical phenotype of a nontruncating mutation may vary within different families, and cosegregation data are not always available. Functional assays have therefore been developed to clarify the activity of nontruncating MMR gene mutations.

The analysis of a number of missense variants has shown that, rather than causing changes to a single amino acid as classically predicted, many variants are instead associated with defects in RNA splicing; some silent mutations have also been shown to cause splice defects.[11–13] In recent years, this problem been addressed by a number of investigators.[14–19] To determine whether a change in the sequence may be a cause of disease or not, pathogenicity can be evaluated based on numerous criteria and using different approaches. Thus, for example, Barnetson et al[9] applied a qualitative, point-based, integrative analysis of UVs; of 23 initially unclassified *MLH1* or *MSH2* missense variants, they classified 11 as benign and 2 as pathogenic. Kansikas et al[20] recently proposed a three-step assessment model.

Nonetheless, classification of some of these described substitutions indicates conflict between the clinicopathological data set of the family carriers and contemporaneous data from either functional assay results or *in silico* approaches. We found different interpretations of the same variant, relative to its pathogenicity, in the MMR Gene Unclassified Variants Database (*http://www.MMRuv.info*, last accessed June 30, 2012).[21] Description of more carrier families is therefore needed to help clarify the role of a given UV in susceptibility to Lynch syndrome.

In the present study, we performed a structured assessment of the pathogenicity of all ambiguous variants identified in our series of colorectal cancer cases, using a set of complementary approaches: phenotypic features in the families (co-occurrence with a pathogenic mutation, cosegregation with the disease, tumor MSI, and DNA mismatch repair protein expression analysis), bioinformatics assessment of the functional consequence of the amino acid change, mRNA analysis, and frequency in a control population (case–control comparisons). The ability to determine the likelihood that a given UV contributes to the disease phenotype is likely to have beneficial consequences for management of the patient.

Patients and Methods

Patients and Mutation Screening

The 159 index cases were from unrelated families referred for MMR mutation analysis under the Junta de Castilla y León Cancer Genetic Counseling Program for the years 2007 to 2010. Informed consent was obtained from the subjects or their parents.

DNA and RNA were purified from peripheral blood lymphocytes by using QIAamp DNA and RNA blood mini kits (Qiagen, Iberia SL, Madrid, Spain; Valencia, CA), respectively. To screen for *MLH1*, *MSH2*, or *MSH6* point mutations, we used a method developed in our laboratory,[22] with validation for MMR genes as recently described.[23]

Evaluation of Clinicopathological Features

For each UV under study, detailed information was gathered on family history, tumor characteristics (IHC, MSI), cosegregation of the variant with disease in families, and co-occurrence with other pathological or UVs.

Somatic *BRAF* Gene Mutation and Methylation Analysis

MLH1 promoter hypermethylation was tested using methylation-sensitive multiplex ligation-dependent probe amplification (MS-MLPA; MRC-Holland, Amsterdam, the Netherlands). *BRAF* p.V600E hot-spot mutation (exon 15) was directly sequenced in both directions in tumor samples of all patients; this assay is characterized by a sensitivity of 10% to 20%. Primers were BRAF-Ex15-forward 5′-TCATAATGCTTGCTCTGATAGGA-3′ and BRAF-Ex15-reverse 5′-GGCCAAAAATTTAATCAGTGGA-3′.

Immunohistochemistry of MLH1, MSH2, and MSH6 Proteins

Tumor immunostaining was analyzed in all patients by a pathologist at the General Yagüe Hospital, Burgos, Spain. Tumor cells were judged to be deficient for protein expression only if they lacked staining in a sample in which normal tissue and stromal cells were stained. If no immunostaining of normal tissue could be demonstrated, the results were considered not evaluable. In brief, a BOND-III stainer system (Leica Biosystems, Barcelona, Spain; Wetzlar, Germany) was used with NCL-L-PMS2, PA0610-MLH1, PA0048-MSH2, and PA0597-MSH6 antibodies (Leica Biosystems), according to the manufacturer's protocol.

Determining the Frequency of UVs in a Control Population

A panel of 478 controls (278 men, 200 women) was selected from the Spanish National DNA bank. These control subjects had no personal histories of cancer and had diagnoses unrelated to the variables of interest. The screening method used was the same as with the case samples, and MMR gene heteroduplex analysis by capillary array electrophoresis[23] was performed.

Protein Effect Prediction Programs

To assess whether the variant amino acids had been evolutionarily conserved across a number of phylogenetically diverse species, we used the T-Coffee tool (*http://tcoffee.vital-it.ch/cgi-bin/Tcoffee/tcoffee_cgi/index.cgi*). The alignment of homologous MLH1, MSH2, and MSH6 proteins was performed using the BLAST tool (*http://blast.ncbi.nlm.nih.gov/Blast.cgi*). The following sequences were used for this purpose: *Homo sapiens*, *Bos taurus*, *Rattus norvegicus*, *Mus musculus*, *Gallus gallus*, *Xenopus tropicalis*, *Danio rerio*, *Drosophila melanogaster*, and *Saccharomyces cerevisiae*. The PolyPhen polymorphism phenotyping tool (*http://genetics.bwh.harvard.edu/pph2/index.shtml*, last accessed June 30, 2012)[24] uses information from protein structure databases and three-dimensional structure databases to predict the effects on the secondary structure of the protein, interchain contacts, and functional sites. PolyPhen scores effects as 1 = benign, 2 = possibly damaging, and 3 = probably damaging. The bioinformatics tool SIFT (the name stands for "sorting intolerant from tolerant"; *http://sift.bii.a-star.edu.sg*, last accessed June 30, 2012) was also used. This tool predicts whether an amino acid substitution affects protein function, based on sequence homology and the physical properties of the amino acid.[25] SIFT can be applied to nonsynonymous polymorphisms, and the tool provides a binary classification: tolerated versus not tolerated (and therefore predicted to affect protein function). The Align-GVGD tool (*http://agvgd.iarc.fr*, last accessed June 30, 2012)[26] combines biophysical characteristics of amino acids and proteins using multiple sequence alignments to predict where reversal substitutions in the genes of interest fall into a spectrum from enriched to deleterious to neutral. The output from Align-GVGD is an ordered series of grades ranging from C65 (most likely deleterious) to C0 (most likely neutral).

Splicing Prediction Programs

Mutant and normal sequences were analyzed using two bioinformatics tools to identify potential splicing mutations. Disruption/creation of splice sites was evaluated with NNSplice (version 0.9; *http://www.fruitfly.org/seq_tools/splice.html*, last accessed June 30, 2012).[27] Analysis of putative splicing regulator elements was done with ESEfinder (version 3.0; *http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home*, last accessed June 30, 2012).[28] This tool is a web-based resource that facilitates rapid analysis of exon sequences to identify putative exonic splicing enhancers (ESEs) responsive to the human SR proteins SF2/ASF, SC35, SRp40, and SRp55, and to predict whether exonic mutations disrupt such elements.

RT-PCR

The synthesis of complementary DNA (cDNA) was performed with a high-capacity cDNA reverse transcription kit (Life Technologies−Applied Biosystems, Foster City, CA) using DNase-treated RNA in the presence of random primers. The nucleotide sequences used to confirm the splicing alteration of the mutation c.2634G>A (p.E878D) were *MSH2*-Ex14-forward 5′-CGATGGATTTGGGTTAG-CAT-3′ and *MSH2*-Ex16-reverse 5′-AGGGCAQTTTGTT-TCACCTTG-3′.

Results

Prevalence of UVs in MMR Genes

Mutational screening of MMR genes was performed in samples from 159 cases of suspected Lynch syndrome. We identified 20 missense mutations or new intronic or synonymous alterations that were categorized of uncertain relevance for causing cancer. These variants are the focus of the present report.

Clinicopathological features and molecular findings for the index patients of these variants are presented in Table 1. Mutation nomenclature is based on GenBank reference sequence NM_000249.3 (*MLH1*); NM_000251.1 (*MSH2*), and NM_000179.2 (*MSH6*). Nucleotide numbering reflects cDNA, with +1 corresponding to the A of the ATG translation initiation codon in the reference sequence, according to Human Genome Variation Society nomenclature for the description of sequence variants (*http://www.hgvs.org/mutnomen*, last accessed June 30, 2012).

The 20 UVs considered here have been described previously: 9 in *MLH1* (45%), 4 in *MSH2* (20%), and 7 in *MSH6* (35%). We present new descriptions for 11 of these variants (3 in *MLH1*, 3 in *MSH2*, and 5 in *MSH6*) (Table 1); the remainder appear in the Leiden Open Variation Database (LOVD 3.0; last accessed June 30, 2012) (LOVD 2.0; *http://www.lovd.nl/2.0*, last accessed June 30, 2012). Their pathogenicity, however, is still under discussion, because the data provided by different authors are contradictory. Thus the return to value in the present study using the system described here.

Considered by type of mutation, the UVs included in the present study consist of 6 synonymous or intronic variants and 14 missense variants (Table 1). The coding variants that we studied are dispersed across different domains of the MLH1, MSH2, and MSH6 polypeptides (Figure 1).

Evaluation of Intronic and Synonymous Variants

Two of the detected UVs are synonymous variants: *MLH1* c.2088C>G (p.T696T) and *MSH6* c.2400T>C (p.V800V). In our series, both variants cosegregated with a mutation previously described as pathological, so the clinicopathological features obtained from these variants seem to be produced by the one that is deleterious. Moreover, neither of them seemed to significantly alter the elements involved in splicing tested by bioinformatics tools (VA-199 and VA-22 families; Table 1).

All intronic variants (*MLH1* c.545+40C>A, c.306+5G>A, and *MSH6* c.457+53insTG) have been analyzed for effects on splicing machinery, to account for their putative pathogenicity. *MLH1* c.545+40C>A and *MSH6* c.457+53insTG create a new enhancer for an SRp protein without relevant implications in the splicing, because the variants are relatively deeply intronic (VA-95 and VA-194 families; Table 1). However, the mutation in *MLH1* c.306+5G>A creates an alternative 5′ donor site, causing an in-frame deletion of five nucleotides. The sequence of the deleted transcript obtained from RT-PCR was reported previously.[23] This mutation appears in three families (VA-44, VA-67, and VA-167), and here we present the tumor analysis and the cosegregation results (Table 1).

Evaluation of Missense Variants by Genes

The clinicopathological features shown by individuals carrying a missense UV mutation were extensively detailed, to assess whether families express a Lynch syndrome phenotype. The results obtained from the promoter methylation study were not relevant and therefore were excluded from further analysis.

MLH1

Of the 11 families carrying missense variants of unknown significance, only 3 families fulfilled the Amsterdam criteria for Lynch syndrome. One is the family with the mutation c.1820T>A (p.L607H), in this instance, despite the strong family history, the variant did not cosegregate with the disease. The second is the family with the c.1865T>A (p.L622H) variant. The third, in one of the six families carrying the variant c.1852_1853AA>GC (p.K618A), did not cosegregate with the disease (family VA-47). The remaining eight families were positive by Bethesda criteria; tumors from five of these families were tested and were found to be were microsatellite stable (MSS).

Of the six different missense variants found in the *MLH1* gene, only one mutation, the variant c.1852_1853AA>GC (p.K618A), was found in more than one unrelated family. In addition, one tumor from a carrier of the p.K618A variant (family VA-55) also carried the BRAF p.V600E mutation in the tumor tissue. This *BRAF* gene mutation is indicative of sporadic cancer.

We also assessed control population frequencies. The c.1217G>A, c.1852_1853AA>GC, and c.2146G>A variants were all identified with similar frequencies in both cases and controls; however, the c.1574G>A and c.1820T>A variants were not identified in the control population (Table 2).

MSH2

We found three missense variants in the *MSH2* gene: c.2651T>G (p.I884S), c.2634G>A (p.E878D), and c.1661G>A (p.S554T). The families carrying these variants met the Amsterdam criteria; the tumor phenotype was unstable (MSI) and had deficient expression of the MSH2−MSH6 heterodimer. The variants cosegregated with the disease. These data indicate a possible deleterious role of the variant in the absence of other pathogenic mutations in the families.

MSH6

As for the missense variants in the *MSH6* gene, we note that two of them, c.2633T>C (p.V878A) and c.431G>T (p.S144I), have also a pathogenic mutation in the *MSH2* gene (the mutation c.431G>T with deletion of exon 7 and c.2633T>C with *MSH2* c.229_230delAG). The tumors from carriers showed deficient expression of the proteins MSH2/MSH6, because of the presence of a deleterious mutation. The other three mutations are new descriptions. Two of them, c.3425C>T (p.T1142M) and c.4004A>C (p.E1335A), cosegregated with the disease; in the third variant, c.98G>C (p.R33P), only the proband was analyzed. Worthy of note is the early age of cancer onset in the families carrying the gene *MSH6*, with a mean age of 38.4 years (range, 27 to 57 years).

Among the variants tested also in the control population, the c.431G>T and c.2633T>C variants were identified with similar frequencies in both cases and controls; the remaining variants tested (c.98G>C, c.3425C>T, and c.4004A>C) were not identified in the control population (Table 2).

In Silico and *in Vitro* Analyses of Missense Substitutions

Bioinformatics evaluation was performed using several approaches (Table 2). The SIFT analysis predicted that 6/14 (42.85%) of the missense changes would not be tolerated. The PolyPhen assessment, however, predicted that 2/14 (14.3%) of these changes were likely to be pathogenic, and Align-GVGD assessment yielded a maximum score for pathogenicity (C65) for 4/9 (44.4%) nonsynonymous coding changes tested.

Evaluations from splicing prediction programs were relevant in two missense mutations for which the proband cases were tested at mRNA level. The NNSplice program predicted the elimination of the 5′ donor site (original donor site with value 0.79) of the mutation c.2634G>A in exon 15 of *MSH2*. This mutation is a priori a missense type, in which a glutamic acid residue is substituted by an aspartic acid residue at position 878 of the protein. RT-PCR analysis showed splicing alterations caused by the complete deletion

Table 1 Clinicopathologic data for 20 UVs of the *MLH1*, *MSH2*, and *MSH6* Genes Identified in Colorectal Cancer Cases

			Index case		Tumor analysis				
Mutation	Family (*n/N*)*	Cosegregation with disease	Cancer (onset†)	Criteria	*BRAF*	MSI	IHC	Co-occurrence with other pathogenic or UV mutations	Other cancers in family (onset†)
MLH1									
c.306+5G>A	VA-44 (5/5)	Yes	CRC (31)	Amst	WT	MSI-H	NE	No	FDR: colon (50); SDR: colon (39)
	VA-67 (3/3)	Yes	CRC (59)	Amst	WT	MSI-H	MLH1 deficient	No	FDR: endo (49), colon (66/72), SDR: endo (50), gastric (50)
	VA-167 (1/1)	Yes	CRC (37)	Amst	NA	NA	NA	No	FDR: colon (54); SDR: colon (56), colon (68)
c.545+40C>A	VA-95 (2/2)	Yes	CRC (42)	Amst	WT	MSS	Normal	MSH6: c.4003A>C; (p.E335A)	FDR: colon (69); SDR: colon (75); endo (45)
c.1217G>A; p.S406N	VA-140 (1/1)	NE	CRC (39)	Beth	WT	MSS	Normal	No	FDR: colon (65); gastric (70)
c.1574G>A; p.S505N	VA-30 (1/1)	NE	CRC (52)	Beth	WT	MSS	NE	No	FDR: meln (80), HL (53); SDR: BC (56)
c.1820T>A; p.L607H	VA-152 (1/2)	No	Polyps (40)	Amst	NA	NA	NA	No	FDR: colon (48), polyps (39), BC (41); SDR: BC (45)
c.1865T>A; p.L622H	VA-179 (1/1)	NE	CRC (37)	Amst	NA	NA	NA	No	Not available
c.1852_1853AA>GC; p.K618A	VA-47 (1/3)	No	CRC (45)	Amst	NA	NA	NA	No	FDR: colon (48), polyps (39), BC (41); SDR: coloNEsoph (42/72), colon (72), colon (35), colon (45), BC (37), BC (43)
	VA-55 (1/2)	NE	CRC (43)	Beth	V600E	MSS	MLH1 deficient	No	No other cancers in the family
	VA-82 (1/1)	NE	CRC/CRC (74/74)	Beth	NA	NA	NA	No	FDR: coloNThyroid (81/86); SDR: gastric (70), colon (70)
	VA-108 (1/1)	NE	CRC (44)	Beth	NA	NA	NA	No	FDR: ovariaNColon (57/57); SDR: Blad (25), colon (48)
	VA-112 (1/1)	NE	CRC (39)	Beth	WT	MSS	NE	No	FDR: colon (75)
	VA-121 (1/1)	NE	CRC (48)	Beth	WT	MSS	NE	No	No other cancers in the family
c.2088C>G; p.T696T	VA-199 (1/1)	NE	Biliary/CRC (50/52)	Amst	NA	NA	NA	MSH2: c.2470C>T; (p.Gln824X)	FDR: colon (36), ovarian (47), colon (40); SDR: CNS (63), gastric (52), colon/urethral (70/80)
c.2146G>A; p.V716M	VA-70 (1/1)	NE	CRC (52)	Beth	WT	MSI-H	MLH1 -	MLH1 promoter methylation	No other cancers in the family
MSH2									
c.212-5delT	VA-43 (1/1)	NE	CRC (72)	Amst	NA	NA	NA	No	Not available
c.1661G>A; p.S554T	VA-6 (7/46)	Yes	CRC (47)	Amst	NA	MSI-H	MSH2/MSH6 deficient	No	FDR: endo/colon/Blad (42/50/60), colon (45), endo (45); SDR: gastric/prostate (40/64)
c.2634G>A; p.E878D	VA-191 (1/1)	NE	Endo (40)	Amst	NA	MSI-H	MSH2/MSH6 deficient	No	FDR: endo (30); SDR: colon/ovariaNColon (29/39/63), endo (30), endo/colon (46/51), endo (30)
c.2651T>G; p.I884S	VA-8 (2/2)	Yes	CRC (20)	Amst	NA	MSI-H	MSH2/MSH6 deficient	No	FDR: ovarian (45), endo (45).
MSH6									
c.98G>C; p.R33P	VA-121 (1/1)	NE	CRC (30)	Beth	WT	MSS	NE	No	No other cancers in the family

(*table continues*)

Table 1 (*continued*)

Mutation	Family (*n/N*)*	Cosegregation with disease	Index case		Tumor analysis			Co-occurrence with other pathogenic or UV mutations	Other cancers in family (onset†)
			Cancer (onset†)	Criteria	*BRAF*	MSI	IHC		
c.431G>T; p.S144I	VA-4 (8/12)	Yes	CRC (36)	Amst	NA	MSI-H	MSH2/MSH6 deficient	MSH2: exon7 del	FDR: colon (32), panc (40); SDR: colon (35), colon/Blad (40/40), colon (41), colon (39)
	VA-94 (1/1)	NE	BC (52)	Beth	NA	NA	NA	No	FDR: small intestine (40), coloNColon (69/73),
c.457+53insTG	VA-194 (1/1)	NE	CRC (41)	Beth	NA	MSI-H	MLH1/MSH6 deficient	No	No other cancers in the family
c.2400T>C; p.V800V	VA-22 (7/12)	NE	CRC/panc (45/61)	Amst	NA	MSI-H	MSH2/MSH6 deficient	MSH2: c.229_230delAG	FDR: colon (53); SDR: colon (75); colon (50); colon (55); colon (60); gastric (70)
c.2633T>C; p.V878A	same family	Yes			NA				
c.3425C>T; p.T1142M	VA-85 (2/2)	Yes	Polyps (27)	Beth	NA	NA	NA	No	FDR: polyps (61)
c.4004A>C; p.E1335A	VA-95 (2/2)	Yes	CRC (42)	Amst	WT	MSS	Normal	MLH1: c.545+40C>A	FDR: colon (69); SDR: colon (75); endo (45)

Variants with new descriptions are highlighted in bold.

*Identifier for the family carrying a UV. *n*, number of affected carriers; *N*, total number of relatives tested.

†Age at onset (years).

Amst, fulfilled Amsterdam criteria; BC, breast cancer; Beth, fulfilled Bethesda criteria; Blad, bladder; CNS, central nervous system; CRC, colorectal cancer; del, deletion; endo, endometrial; esoph, esophageal; FDR, first-degree relatives; HL, Hodgkin's lymphoma; meln, melanoma; MSI, microsatellite instability; MSI-H, microsatellite instability, high; MSS, microsatellite stability; NA, not available; NE, not evaluable; panc, pancreatic; SDR, second-degree relatives; WT, wild type.

of exon 15 of *MSH2* (Figure 2). The loss of this exon would result in a loss of reading frame at the protein level and a premature stop codon. Also, we evaluated the *in silico* effect of the transversion in the *MSH2* gene (c.1661G>A; p.S554T), which affects the last nucleotide of exon 10, producing a priori a missense mutation. In this case, it was found that the change of guanine to adenine washes site 5′ splice natural (value 0.60), using a new cryptic donor site at position c.1580 (value 0.56), producing deletion of 81 nucleotides in mRNA. Analysis on the RNA level confirmed the result predicted by the computer program and characterized in a previous work.[23]

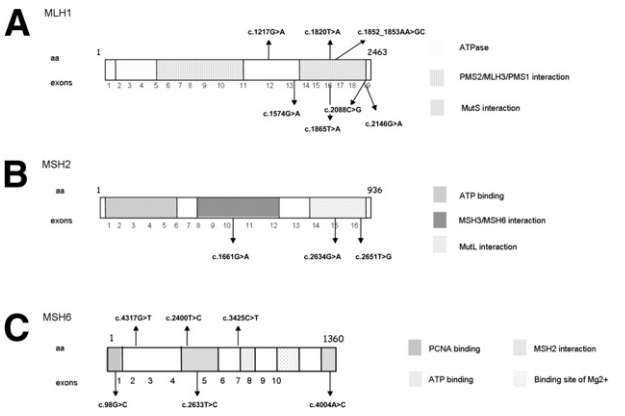

Figure 1 Schematic illustration of *MLH1* (**A**), *MSH2* (**B**), and *MSH6* (**C**), showing the main known functional domains and location of the studied coding variations. aa, amino acids.

Discussion

The present data indicate that the prevalence of variants of unknown effect in MMR genes is high, accounting for 34% of all mutations detected in our population. This fact underscores the importance of the study. The characterization of the alterations of unknown significance is one of the important aspects of the present study, and such characterization has also been the target of other recent studies.[14,17,18]

To understand the importance of UVs, we take into account several variables such as phenotypic features, bioinformatics prediction, and RNA analysis to evaluate the degree of support for pathogenicity for all variants listed (Table 2). At the end of this work, we could distinguish six distinct categories: synonymous, neutral, reclassified, splicing, pathogenic, and UVs.

Although intronic and synonymous variants may suggest low or absent clinical significance, these should nonetheless be dissected, to rule out possible effects on splicing.[13,19] Here, we have presented new descriptions for five variants of this type: *MLH1* c.545+40C>A and c.2088C>G, *MSH2* c.212-5delT, and *MSH6* c.457+53insTG and c.2400T>C. The two synonymous variants, *MLH1* c.2088C>G and *MSH6* c.2400T>C, co-occur with a pathogenic mutation (Table 2). The two deep intronic mutations (*MLH1* c.545+40C>A and *MSH6* c.457+53insTG) were detected in patients who met the Bethesda criteria and had an MSS tumor; evaluation of the clinicopathological features and bioinformatics results allowed us to classify these variants as polymorphisms.

Table 2 Evaluation for Pathogenicity of *MLH1*, *MSH2*, and *MSH6* UVs Identified in Colorectal Cancer Cases

	Case–control study				Effect on protein					Effect on RNA		
	Index cases		Controls		Prediction *in silico*					Prediction *in silico*		
Mutation	*n/N*	Freq	*n/N*	Freq	SIFT tolerated*	PolyPhen score†	Conserved (T-Coffee)‡	Align-GVGD§	Functional assay (Refs)	NNSplice scores¶	ESEs altered (ESEFinder‖)	RNA defect (Refs)
MLH1												
c.306+5G>A	3/159	0.019	NA	NA	NE	NE	NE	NE	NT	0/0.94	NA	Del 5 bp, exon 3
c.545+40C>A	1/159	0.006	NA	NA	NE	NE	NE	NE	NT	NC	SC35(+)	NT
c.1217G>A p.S406N	1/159	0.006	1/266	0.004	Yes	2	Yes	C0	NP ([29,30,31])	NC	NC	NT
c.1574G>A p.S505N	1/159	0.006	0/524	0	Yes	1	No	C45	NT	NC	NC	NT
c.1820t>a p.l607h	1/159	0.006	0/266	0	Yes	2	No	C65	NP ([29])	NC	NC	None ([19])
c.1865T>A p.L622H	1/159	0.006	NA	NA	No	3	No	C65	P ([32])	NC	Sp55(+)	
c.1852_1853AA>GC p.K618A	6/159	0.038	14/612	0.023	Yes	1	No	C65	NP ([37,38]); P ([34]); I ([38])	NC	NC	None ([19,36])
c.2088C>G p.T696T	1/159	0.006	NA	NA	NE	NE	NE	NE	NT	NC	SF2/ASF(+), SC35(−)	NT
c.2146G>A p.V716M	1/159	0.006	3/428	0.007	No	2	No	C0	NP ([34])	NC	SC35(−)	None ([13])
MSH2												
c.212-5DelT	1/159	0.006	0/425	0	NE	NE	NE	NE	NT	NC	NC	NT
c.1661G>A p.S554T	1/159	0.006	NA	NA	Yes	1	No	C55	NT	0/0.54	NA	Del 81 bp, exon 10
c.2634G>A p.E878D	1/159	0.006	NA	NA	Yes	1	No	C35	NT	0.79/0	NA	Skip exon 15
c.2651T>G p.I884S	1/159	0.006	n.a	n.a	No	3	No	C65	NT	NC	SF2/ASF(+)	NT
MSH6												
c.98G>C p.R33P	1/159	0.006	0/350	0	Yes	2	No	NA	NT	NC	SF2/ASF(+), SF2/ASF (IgM-BRCA1) (+)	NT
c.431G>T p.S144I	2/159	0.013	4/372	0.011	No	2	No	NA	NP ([37,38]); P ([34]); I ([38])	NC	SC35(−), Rp4(+)	NT
c.457+53insTG	1/159	0.006	NA	NA	NE	NE	NE	NE	NT	NC	Sp35(+)	NT
c.2400T>C p.V800V	1/159	0.006	NA	NA	NE	NE	No	NE	NT	NC	NC	NT
c.2633T>C p.V878A	1/159	0.006	5/501	0.010	Yes	1	No	NA	P ([38])**	NC	SF2/ASF (IgM-BRCA1) (+)	NT
c.3425C>T p.T1142M	1/159	0.006	0/347	0	No	3	Yes	C65	NT	NC	Sp55(+)	NT
c.4004A>C p.E1335A	1/159	0.006	0/347	0	No	1	No	NA	NT	NC	NC	NT

Variants with new descriptions are highlighted in bold.

*For SIFT: Yes, change tolerated; No, change not tolerated

†PolyPhen scoring: 1 = benign; 2 = possibly damaging; 3 = probably damaging.

‡Yes, conserved; No, not conserved. Conservation across phylogenetically diverse species: *H. sapiens, B. taurus, R. norvegicus, M. musculus, G. gallus, X. tropicalis, Da. rerio, Dr. melanogaster*, and *S. cerevisiae*.

§Align-GVGD yields an ordered series of grades ranging from C65 (most likely deleterious) to C0 (most likely neutral).

¶NNSplice score changes: wild type/mutant.

‖ESEs altered according to ESEFinder. (+), creation; (−), elimination.

**Pathogenic due to deficiencies in the protein interaction with ADP.

Del, deletion; ESE, exonic splicing enhancer; Freq, frequency; I, inconclusive; NA, not analyzed; NC, no change; NE, not evaluable; NT, not tested functionally; NP, not pathogenic; P, pathogenic; Refs, references.

Our evaluation classified *MLH1* c.1852_1853AA>GC and c.2146G>A variants as probably neutral. The variant c.1852_1853AA>GC affects a tract of three repeated lysines, located at codons 616 to 618 of the MLH1 protein. The c.1852_1853AA>GC variant is one of the mutations that has been described in patients worldwide.[41] This variant has been classified as deleterious and neutral, as is reflected in the MMR missense database (*http://www.MMRuv.info*). In the present study, it has been identified in six unrelated families, five of which met the Bethesda criteria, with cancer onset age between 39 and 70 years. The phenotypes of the tumor samples analyzed showed MSS tumors with or without loss of MLH1 protein expression. Also, in one case, we detected the *BRAF* V600E mutation in tumor tissue, thus excluding the possibility of the tumor being associated with Lynch syndrome.[42] On the other hand, promoter methylation analysis has limitations as a predictor of mutation status. Our control population analysis indicated that this c.1852_1853delAAinsGC variant occurs at a higher frequency in controls than in patients tested (0.038 versus 0.023). In general, UVs occur at relatively low frequencies, so this finding would need to be verified with a larger panel

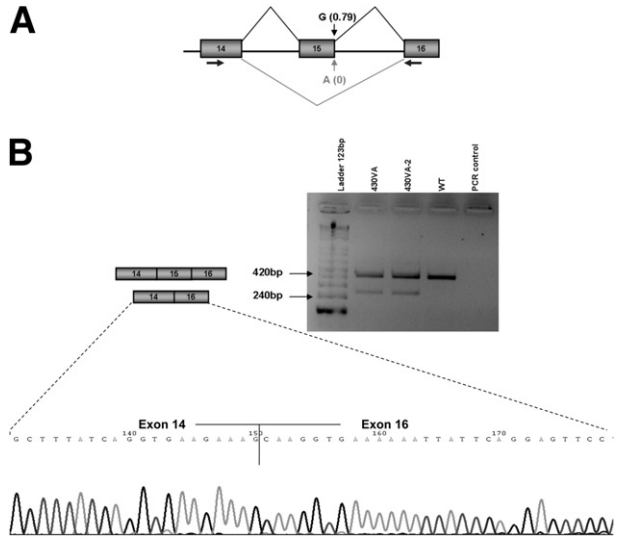

Figure 2 Altered splicing that produces the complete exclusion of exon 15 of *MSH2* c.2634 G>A. **A:** Exon skipping as predicted by NNSplice; the primers in *MSH2* exons 14 and 16 were used to amplify cDNA from an individual with the missense change. **B:** RT-PCR products are shown with the normal and the deleted transcript; the sequence that supports fusion between exons 14 and 16 of *MSH2* is included.

of control samples. Functionally, this variant acts in a similar way to the wild protein, although some functional studies have shown that the substitution affects the interaction between MLH1 and the PMS2 proteins.[33,39] On the other hand, another study, with similar characteristics, suggests that this involvement is not pathogenic.[31,35] Therefore, we have classified this variant as probably neutral. Recently, a study in a Spanish population, analyzing the pathogenicity of this variant in a comprehensive manner, supports our neutral result.[43]

We have reclassified two variants, *MSH6* c.431G>T and c.2633T>C, that were previously classified as probably pathogenic in patients with colon and endometrial cancer.[44] Our conclusion was based on several criteria. First, the c.431G>T mutation was observed in tumors with MSI and was not identified in 399 controls, and it showed loss of MMR function in a study of yeast (although the result was not reproduced in human cells *in vitro*).[40,45] For the c.2633T>C variant, cosegregation with affected individuals has been described within a family with MSI.[46,47] On the other hand, in a study genotyping 2000 individuals,[9] similar frequencies of substitutions were found in both cases (*MSH6* c.431G>T and c.2633T>C) and controls, leading to the classification as neutral variants. Additionally, the variant c.2633T>C is listed in the SNP database (rs2020912, T>C; *http://www.ncbi.nlm.nih.gov/snp*, last accessed June 30, 2012) with an allele frequency of 0.008 in the European population. In the present study, both mutations co-occur with a clearly pathogenic mutation in the *MSH2* gene. They were also present in our control population, with a similar frequency to that of cases for the variant p.S144I, and even slightly higher than that established in cases (0.006 versus 0.01) for the c.2633T>C variant. The presence of a UV in the control population is more significant than its absence,[48] because the fact that a UV (generally of low prevalence) has not been found in controls could be an artifact of underestimation. These circumstances are sufficient to determine that these UVs are not associated with the disease in our population.

In the present study, a high proportion (15%) of UVs affected the splicing mechanism. With the present study, we propose a procedure to detect the effect of changes in the splicing exons and introns based on bioinformatics analysis and verification of the effect of altering the splicing mechanism by evaluating the mRNA level by RT-PCR. It is necessary to assess the nature of the altered transcripts produced in each case (Table 2). Do they alter the reading frame, and are they likely to be degraded by nonsense-mediated decay? Or, if not, what changes are caused at the protein level, which amino acids are excluded or included in the protein, and what might its functional importance may be?

With our approach, we detected a high incidence of splicing mutations in MMR genes, representing 15% (3/20) of all analyzed UVs. These mutations (c.306+5G>A in *MLH1* and c.1661G>A and c.2634G>A in *MSH2*) were located in highly conserved regions involved in regulating splicing. The *MLH1* c.306+5G>A mutation creates an alternative 5′ donor site causing an in-frame deletion of five nucleotides. The predicted effect of the previously reported *MLH1* variant c.306+5G>A on splicing was confirmed.[23,49]

MSH2 c.1661G>A is a new variant; its effect is elimination of the cryptic donor site and subsequent activation of a new splice site, which is located 81 nucleotides upstream.[23] The *MSH2* c.2634G>A mutation produces skipping of exon 15. This mutation has been described in a German population (MSH2_00823 in the LOVD database), but functional studies supporting its pathogenicity are lacking. It should be noted that two of these mutations are missense type (c.1661G>A, c.2634G>A). However, we now assume that the probable deleterious effect of all these variants is located in the region involved in the definition of the ends of the exons, mainly the 5′ donor splice site and the 3′ acceptor splice site.

Drawing on our experience, we propose a procedure based on bioinformatics analysis and verification of the effect of altering the splicing mechanism by evaluating the mRNA level by RT-PCR, in diagnostic practice. Such an approach should be able to detect the effect of changes in the splicing exons and introns, subject to the availability of mRNA of patients.

We have classified *MLH1* c.1865T>A, *MSH2* c.2651T>G, and *MSH6* c.3425C>T variants as probably pathogenic. All three have been classified as C65 substitutions by the Align-GVGD program,[50] for which the criteria are very strict; almost all variants here occur at an evolutionarily invariant position. In addition, *MLH1* c.1865T>A was found in a family meeting the Amsterdam criteria in our cohort, but also in another 12 Spanish families, and Borras et al[32] characterized it as a founder mutation. Borras et al[32]

concluded that the c.1865T>A variant causes defects in MLH1 expression and stability. In the present study, the *MSH2* c.2651T>G was found in one family (VA-8) that met the Amsterdam criteria (index case at age 20 years). The tumor probands phenotype was unstable and the expression of MSH2−MSH6 heterodimer was deficient. This variant cosegregated with the disease and was not found in the control population. This mutation is described in a French population study reporting a tumor bearer, although the MSI is associated with normal IHC expression, arguing that this missense mutation is likely to affect the functioning of the protein but not its transcript.[49,51] Additionally, this variant has been analyzed by *in silico* multivariate analysis, but the results remain ambiguous.[51] The data set indicates a possible deleterious role of the variant. However, we cannot exclude the possibility that other undetected mutations may be linked to the c.2651T>G variant in this family. In this case, functional studies could clarify the involvement of this variant protein activity.

The *MSH6* c.3425C>T affects a highly conserved amino acid located in the ATP binding domain of the MSH6 protein (Figure 1C); this mutation was detected in a patient who fulfilled the Bethesda criteria. The index case developed polyps at the age of 27 years. In addition, his mother (who developed polyps at age 61) also carried the alteration. No tumor tissue was available for analysis of MSI. However, it has been reported that the tumor phenotype, in the patients who carried variants in *MSH6* gene, shows different degrees of microsatellite instability. This variant was not detected in the Spanish control population; *in silico* studies are consistent and indicate a strong pathogenic role of this change at the level of protein functionality. Because the description of this variant is new, and because the size of the family studied is limited, biochemical study of the variant might determine the effect on function of the protein, and our research group is therefore preparing the necessary protocols for such functional analysis.

The remaining UVs, which are of ambiguous clinical relevance in the present study, are *MLH1* c.1217G>A, c.1574G>A, and c.1820T>A and *MSH6* c.98G>C and c.4004A>C. The *MLH1* c.1217G>A and c.1820T>A variants likely have a neutral role, based on the *in silico* results. Functional studies have not been performed, and the cosegregation evaluation does not allow confirmation. It has not been possible to classify the variant *MLH1* c.1574G>A, a newly described variant; it is not located on any important protein domain (Figure 1A) and it has not been detected in controls.

The family size for the *MSH6* UV probands is small. These variants have been described in only one family each, and so we can only speculate about its possible functional effect. The c.4004A>C variant is located in the domain of union with *MSH2*. Functional assays to assess the ability of heterodimerization of the mutant protein could be enlightening.

The integrated assessment model used in the present study is a useful tool for evaluating the pathogenicity of the variations in MMR genes, especially the new-description mutations. We detected 20 UVs in our series, and only 5 (25%) of these remained unclassified by the method that we propose here. Determining the role of UVs is relevant, because decisions about colon and gynecological prevention screening depend on the interpretation of these mutations as pathogenic or neutral. The combination of all these approaches, along with the study of these variants in a large number of specific-population control samples, should help predict whether these variants contribute to the disease phenotype or merely represent rare polymorphisms. It is known that inherited predisposition to disease is often linked to reduced activity of a disease-associated gene product. Sometimes a change in the nucleotide sequence is found for which there is not enough information to decide whether it affects the function of the gene product and, by consequence, influences the cancer risk. Especially in the case of genes that strongly affect risk, such noninformative results can be a source of anxiety to individuals and their offspring, who cannot use such information from genetic testing to modify behavior or lifestyle or to make important clinical decisions that may, in many cases, involve prophylactic surgery.

We propose that the continuum model is most likely to be applied where multiple, cancer-promoting mutations have relatively small, additive effects, either through well-established additive germline predisposition alleles or through a largely hypothetical situation in which cancer may have acquired several minidriver somatic mutations, each with weaker effects than classical tumor suppressors or fully activated oncogenes.[52]

The present findings suggest that most of the intronic and synonymous variants are not pathogenic, and that a high proportion of missense variants disrupt the splicing mechanism. This integrated evaluation can be applied in similarly sized independent series of missense substitutions in these genes or in other similar genes (eg, *BRCA1/2* in breast cancer). Molecular diagnostics has become an essential component in clinical decision-making, and better classification of tumors by molecular methods can help realize personalized cancer medicine and prevention.[53]

Finally, we believe that, in the future, biological and biochemical functional assays may help clarify the activity of these variants. Currently, our research group is preparing protocols and yeast cell cultures to assess the functional consequences of new-description variants of our population. To do so, we apply such techniques as double-hybrid (protein−protein interaction) localization studies with fluorescent tagging and expression studies of the protein (mutant versus wild). The establishment of reference laboratories is necessary for such studies.

Acknowledgments

We thank the National DNA Bank of Salamanca and the Science Foundation AECC. We also thank Noemy Martínez and Lara Hernández for their excellent technical support

and Alan Hynds for his critical reading of this article in manuscript.

References

1. Fishel R, Lescoe MK, Rao MR, Copeland NG, Jenkins NA, Garber J, Kane M, Kolodner R: The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer [Erratum appeared in Cell 1994, 77(1):1 p following 166]. Cell 1993, 75:1027−1038
2. Bronner CE, Baker SM, Morrison PT, Warren G, Smith LG, Lescoe MK, Kane M, Earabino C, Lipford J, Lindblom A, Tannergård P, Bollag RJ, Godwin AR, Ward DC, Nordenskjøld M, Fishel R, Kolodner R, Liskay RM: Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. Nature 1994, 368:258−261
3. Nicolaides NC, Papadopoulos N, Liu B, Wei YF, Carter KC, Ruben SM, Rosen CA, Haseltine WA, Fleischmann RD, Fraser CM, Adams MD, Venter JC, Dunlop MG, Hamilton SR, Petersen GM, de la Chapelle A, Vogelstein B, Kinzler KW: Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. Nature 1994, 371:75−80
4. Wijnen J, van der Klift H, Vasen H, Khan PM, Menko F, Tops C, Meijers Heijboer H, Lindhout D, Møller P, Fodde R: MSH2 genomic deletions are a frequent cause of HNPCC. Nat Genet 1998, 20:326−328
5. Lynch HT, Lynch PM, Lanspa SJ, Snyder CL, Lynch JF, Boland CR: Review of the Lynch syndrome: history, molecular genetics, screening, differential diagnosis, and medicolegal ramifications. Clin Genet 2009, 76:1−18
6. Heinen CD: Genotype to phenotype: analyzing the effects of inherited mutations in colorectal cancer families. Mutat Res 2010, 693:32−45
7. Peltomäki P: Deficient DNA mismatch repair: a common etiologic factor for colon cancer. Hum Mol Genet 2001, 10:735−740
8. Woods MO, Williams P, Careen A, Edwards L, Bartlett S, McLaughlin JR, Younghusband HB: A new variant database for mismatch repair genes associated with Lynch syndrome. Hum Mutat 2007, 28:669−673
9. Barnetson RA, Cartwright N, van Vliet A, Haq N, Drew K, Farrington S, Williams N, Warner J, Campbell H, Porteous ME, Dunlop MG: Classification of ambiguous mutations in DNA mismatch repair genes identified in a population-based study of colorectal cancer. Hum Mutat 2008, 29:367−374
10. Genuardi M, Carrara S, Anti M, Ponz de Leòn M, Viel A: Assessment of pathogenicity criteria for constitutional missense mutations of the hereditary nonpolyposis colorectal cancer genes MLH1 and MSH2. Eur J Hum Genet 1999, 7:778−782
11. Gorlov IP, Gorlova OY, Frazier ML, Amos CI: Missense mutations in hMLH1 and hMSH2 are associated with exonic splicing enhancers. Am J Hum Genet 2003, 73:1157−1161
12. Lastella P, Surdo NC, Resta N, Guanti G, Stella A: In silico and in vivo splicing analysis of MLH1 and MSH2 missense mutations shows exon- and tissue-specific effects. BMC Genomics 2006, 7:243
13. Auclair J, Busine MP, Navarro C, Ruano E, Montmain G, Desseigne F, Saurin JC, Lasset C, Bonadona V, Giraud S, Puisieux A, Wang Q: Systematic mRNA analysis for the effect of MLH1 and MSH2 missense and silent mutations on aberrant splicing. Hum Mutat 2006, 27:145−154
14. Couch FJ, Rasmussen LJ, Hofstra R, Monteiro AN, Greenblatt MS, de Wind N; IARC Unclassified Genetic Variants Working Group: Assessment of functional effects of unclassified genetic variants. Hum Mutat 2008, 29:1314−1326
15. Lucci-Cordisco E, Boccuto L, Neri G, Genuardi M: The use of microsatellite instability, immunohistochemistry and other variables in determining the clinical significance of MLH1 and MSH2 unclassified variants in Lynch syndrome. Cancer Biomark 2006, 2: 11−27
16. Ou J, Niessen RC, Lützen A, Sijmons RH, Kleibeuker JH, de Wind N, Rasmussen LJ, Hofstra RM: Functional analysis helps to clarify the clinical importance of unclassified variants in DNA mismatch repair genes. Hum Mutat 2007, 28:1047−1054
17. Pastrello C, Pin E, Marroni F, Bedin C, Fornasarig M, Tibiletti MG, Oliani C, Ponz de Leon M, Urso ED, Della Puppa L, Agostini M, Viel A: Integrated analysis of unclassified variants in mismatch repair genes. Genet Med 2011, 13:115−124
18. Perera S, Li B, Tsitsikotas S, Ramyar L, Pollett A, Semotiuk K, Bapat B: A novel and rapid method of determining the effect of unclassified MLH1 genetic variants on differential allelic expression. J Mol Diagn 2010, 12:757−764
19. Tournier I, Vezain M, Martins A, Charbonnier F, Baert-Desurmont S, Olschwang S, Wang Q, Buisine MP, Soret J, Tazi J, Frébourg T, Tosi M: A large fraction of unclassified variants of the mismatch repair genes MLH1 and MSH2 is associated with splicing defects. Hum Mutat 2008, 29:1412−1424
20. Kansikas M, Kariola R, Nystrom M: Verification of the three-step model in assessing the pathogenicity of mismatch repair gene variants. Hum Mutat 2011, 32:107−115
21. Ou J, Niessen RC, Vonk J, Westers H, Hofstra RM, Sijmons RH: A database to support the interpretation of human mismatch repair gene variants. Hum Mutat 2008, 29:1337−1341
22. Velasco E, Infante M, Durán M, Esteban-Cardeñosa E, Lastra E, García-Girón C, Miner C: Rapid mutation detection in complex genes by heteroduplex analysis with capillary array electrophoresis. Electrophoresis 2005, 26:2539−2552
23. Perez-Cabornero L, Velasco E, Infante M, Sanz D, Lastra E, Hernández L, Miner C, Duran M: A new strategy to screen MMR genes in Lynch syndrome: hA-CAE. MLPA and RT-PCR 5. Eur J Cancer 2009, 45:1485−1493
24. Sunyaev S, Lathe W 3rd, Bork P: Integration of genome data and protein structures: prediction of protein folds, protein interactions and "molecular phenotypes" of single nucleotide polymorphisms. Curr Opin Struct Biol 2001, 11:125−130
25. Ng PC, Henikoff S: Predicting deleterious amino acid substitutions. Genome Res 2001, 11:863−874
26. Tavtigian SV, Deffenbaugh AM, Yin L, Judkins T, Scholl T, Samollow PB, de Silva D, Zharkikh A, Thomas A: Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. J Med Genet 2006, 43:295−305
27. Reese MG, Eeckman FH, Kulp D, Haussler D: Improved splice site detection in Genie. J Comput Biol 1997, 4:311−323
28. Liu HX, Zhang M, Krainer AR: Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins. Genes Dev 1998, 12:1998−2012
29. Takahashi M, Shimodaira H, Andreutti-Zaugg C, Iggo R, Kolodner RD, Ishioka C: Functional analysis of human MLH1 variants using yeast and in vitro mismatch repair assays. Cancer Res 2007, 67:4595−4604
30. Wanat JJ, Singh N, Alani E: The effect of genetic background on the function of Saccharomyces cerevisiae mlh1 alleles that correspond to HNPCC missense mutations. Hum Mol Genet 2007, 16: 445−452
31. Kondo E, Suzuki H, Horii A, Fukushige S: A yeast two-hybrid assay provides a simple way to evaluate the vast majority of hMLH1 germline mutations. Cancer Res 2003, 63:3302−3308
32. Borràs E, Pineda M, Blanco I, Jewett EM, Wang F, Teulé A, Caldés T, Urioste M, Martínez-Bouzas C, Brunet J, Balmaña J, Torres A, Ramón y Cajal T, Sanz J, Pérez-Cabornero L, Castellví-Bel S, Alonso A, Lanas A, González S, Moreno V, Gruber SB, Rosenberg NA, Mukherjee B, Lázaro C, Capellá G: MLH1 founder mutations with moderate penetrance in Spanish Lynch syndrome families. Cancer Res 2010, 70:7379−7391
33. Belvederesi L, Bianchi F, Loretelli C, Gagliardini D, Galizia E, Bracci R, Rosati S, Bearzi I, Viel A, Cellerino R, Porfiri E: Assessing

the pathogenicity of MLH1 missense mutations in patients with suspected hereditary nonpolyposis colorectal cancer: correlation with clinical, genetic and functional features. Eur J Hum Genet 2006, 14: 853−859
34. Raevaara TE, Korhonen MK, Lohi H, Hampel H, Lynch E, Lönnqvist KE, Holinski-Feder E, Sutter C, McKinnon W, Duraisamy S, Gerdes AM, Peltomäki P, Kohonen-Ccorish M, Mangold E, Macrae F, Greenblatt M, de la Chapelle A, Nyström M: Functional significance and clinical phenotype of nontruncating mismatch repair variants of MLH1. Gastroenterology 2005, 129:537−549
35. Perera S, Bapat B: The MLH1 variants p.Arg265Cys and p.Lys618Ala affect protein stability while p.Leu749Gln affects heterodimer formation. Hum Mutat 2008, 29:332
36. Sharp A, Pichert G, Lucassen A, Eccles D: RNA analysis reveals splicing mutations and loss of expression defects in MLH1 and BRCA1. Hum Mutat 2004, 24:272
37. Kariola R, Raevaara TE, Lönnqvist KE, Nyström-Lahti M: Functional analysis of MSH6 mutations linked to kindreds with putative hereditary non-polyposis colorectal cancer syndrome. Hum Mol Genet 2002, 11:1303−1310
38. Cyr JL, Heinen CD: Hereditary cancer-associated missense mutations in hMSH6 uncouple ATP hydrolysis from DNA mismatch binding. J Biol Chem 2008, 283:31641−31648
39. Guerrette S, Acharya S, Fishel R: The interaction of the human MutL homologues in hereditary nonpolyposis colon cancer. J Biol Chem 1999, 274:6336−6341
40. Kolodner RD, Tytell JD, Schmeits JL, Kane MF, Gupta RD, Weger J, Wahlberg S, Fox EA, Peel D, Ziogas A, Garber JE, Syngal S, Anton-Culver H, Li FP: Germ-line msh6 mutations in colorectal cancer families. Cancer Res 1999, 59:5068−5074
41. Peltomäki P, Vasen H: Mutations associated with HNPCC predisposition−Update of ICG-HNPCC/INSiGHT mutation database. Dis Markers 2004, 20:269−276
42. Domingo E, Laiho P, Ollikainen M, Pinto M, Wang L, French AJ, Westra J, Frebourg T, Espín E, Armengol M, Hamelin R, Yamamoto H, Hofstra RM, Seruca R, Lindblom A, Peltomäki P, Thibodeau SN, Aaltonen LA, Schwartz S Jr: BRAF screening as a low-cost effective strategy for simplifying HNPCC genetic testing. J Med Genet 2004, 41:664−668
43. Castillejo A, Guarinos C, Martinez-Canto A, Barbera VM, Egoavil C, Castillejo MI, Perez-Carbonell L, Sanchez-Heras AB, Segura A, Ochoa E, Lazaro R, Ruiz-Ponte C, Bujanda L, Andreu M, Castells A, Carracedo A, Llor X, Clofent J, Alenda C, Paya A, Jover R, Soto JL: Evidence for classification of c. 1852_1853AA>GC in MLH1 as a neutral variant for Lynch syndrome. BMC Med Genet 2011, 12:12
44. Berends MJ, Wu Y, Sijmons RH, Mensink RG, van der Sluis T, Hordijk-Hos JM, de Vries EG, Hollema H, Karrenbeld A, Buys CH, van der Zee AG, Hofstra RM, Kleibeuker JH: Molecular and clinical characteristics of MSH6 variants: an analysis of 25 index carriers of a germline variant. Am J Hum Genet 2002, 70:26−37
45. Kariola R, Hampel H, Frankel WL, Raevaara TE, de la Chapelle A, Nyström-Lahti M: MSH6 missense mutations are often associated with no or low cancer susceptibility. Br J Cancer 2004, 91: 1287−1292
46. Sánchez de Abajo A, de la Hoya M, Tosar A, Godino J, Fernández JM, Asenjo JL, Villamil BP, Segura PP, Diaz-Rubio E, Caldes T: Low prevalence of germline hMSH6 mutations in colorectal cancer families from Spain. World J Gastroenterol 2005, 11:5770−5776
47. Wijnen J, de Leeuw W, Vasen H, van der Klift H, Møller P, Stormorken A, Meijers-Heijboer H, Lindhout D, Menko F, Vossen S, Möslein G, Tops C, Bröcker-Vriends A, Wu Y, Hofstra R, Sijmons R, Cornelisse C, Morreau H, Fodde R: Familial endometrial cancer in female carriers of MSH6 germline mutations. Nat Genet 1999, 23:142−144
48. Hofstra RM, Spurdle AB, Eccles D, Foulkes WD, de Wind N, Hoogerbrugge N, Hogervorst FB; IARC Unclassified Genetic Variants Working Group: Tumor characteristics as an analytic tool for classifying genetic variants of uncertain clinical significance. Hum Mutat 2008, 29:1292−1303
49. Bécouarn Y, Rullier A, Gorry P, Smith D, Richard-Molard B, Echinard E, Texereau P, Beyssac R, Legoux JL, Lamouliatte H, Frebourg T, Olschwang S, Gilbert B, Venat L, Picot V, Paraf F, Longy M: Value of microsatellite instability typing in detecting hereditary non-polyposis colorectal cancer. A prospective multicentric study by the Association Aquitaine Gastro. Gastroenterol Clin Biol 2005, 29:667−675
50. Tavtigian SV, Greenblatt MS, Lesueur F, Byrnes GB; IARC Unclassified Genetic Variants Working Group: In silico analysis of missense substitutions using sequence-alignment based methods. Hum Mutat 2008, 29:1327−1336
51. Chao EC, Velasquez JL, Witherspoon MS, Rozek LS, Peel D, Ng P, Gruber SB, Watson P, Rennert G, Anton-Culver H, Lynch H, Lipkin SM: Accurate classification of MLH1/MSH2 missense variants with multivariate analysis of protein polymorphisms-mismatch repair (MAPP-MMR). Hum Mutat 2008, 29:852−860
52. Berger AH, Knudson AG, Pandolfi PP: A continuum model for tumour suppression. Nature 2011, 476:163−169
53. Ogino S, Fuchs CS, Giovannucci E: How many molecular subtypes? Implications of the unique tumor principle in personalized medicine. Expert Rev Mol Diagn 2012, 12:621−628