IJC
International Journal of Cancer

Comprehensive screening for mutations associated with colorectal cancer in unselected cases reveals penetrant and nonpenetrant mutations

Cornelia Kraus[1*], Tilman T. Rau[2*], Philipp Lux[3*], Katharina Erlenbach-Wünsch[2], Sabine Löhr[1], Mandy Krumbiegel[1], Christian T. Thiel[1], Robert Stöhr[2], Abbas Agaimy[2], Roland S. Croner[3], Michael Stürzl[3], Werner Hohenberger[3*], Arndt Hartmann[2*] and André Reis[1*]

[1] Institute of Human Genetics, University Hospital Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
[2] Institute of Pathology, University Hospital Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
[3] Department of Surgery, University Hospital Erlangen, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

Germline mutation testing in patients with colorectal cancer (CRC) is offered only to a subset of patients with a clinical presentation or tumor histology suggestive of familial CRC syndromes, probably underestimating familial CRC predisposition. The aim of our study was to determine whether unbiased screening of newly diagnosed CRC cases with next generation sequencing (NGS) increases the overall detection rate of germline mutations. We analyzed 152 consecutive CRC patients for germline mutations in 18 CRC-associated genes using NGS. All patients were also evaluated for Bethesda criteria and all tumors were investigated for microsatellite instability, immunohistochemistry for mismatch repair proteins and the *BRAF*V600E* somatic mutation. NGS based sequencing identified 27 variants in 9 genes in 23 out of 152 patients studied (18%). Three of them were already reported as pathogenic and 12 were class 3 germline variants with an uncertain prediction of pathogenicity. Only 1 of these patients fulfilled Bethesda criteria and had a microsatellite instable tumor and an *MLH1* germline mutation. The others would have been missed with current approaches: 2 with a *MSH6* premature termination mutation and 12 uncertain, potentially pathogenic class 3 variants in *APC*, *MLH1, MSH2, MSH6, MSH3* and *MLH3*. The higher NGS mutation detection rate compared with current testing strategies based on clinicopathological criteria is probably due to the large genetic heterogeneity and overlapping clinical presentation of the various CRC syndromes. It can also identify apparently nonpenetrant germline mutations complicating the clinical management of the patients and their families.

Cancer Genetics

Twin studies estimated a hereditary background in colorectal cancer (CRC) in up to 35% of cases.[1] In about 5% of all patients, CRC is of genetic origin with a causative germline mutation. The diagnosis of hereditary cancer syndrome has significant implications for the medical management of CRC patients and their families.[2–4] Lynch syndrome or hereditary

Key words: next generation sequencing, gene panel, colorectal cancer, unbiased

Abbreviations: aFAP: attenuated Familial Adenomatous Polyposis; CRC: colorectal cancer; CWS1: Cowden syndrome; DNA: deoxyribonucleic acid; ESP: NHLBI Exome Sequencing Project (Exome Variant Server); FAP: Familial Adenomatous Polyposis; FFPE: formalin fixed paraffin embedded; GD: Grantham Deviation; GV: Grantham Variation; HDGC: hereditary diffuse gastric cancer; HNPCC: hereditary non-polyposis colorectal cancer; IARC: International Agency for Research on Cancer; IHC: immunohistochemistry; JPS: Juvenile polyposis syndrome; LFS1: Li–Fraumeni syndrome; LOVD: Leiden Open Variant Database; MAF: minor allele frequency; MAP: MYH-associated polyposis; MAPP-MMR: multivariate analysis of protein polymorphisms-mismatch repair; MLPA: multiplex ligation-dependent probe amplification; MMR: mismatch repair; MSI: microsatellite instability; MSS: microsatellite stable; NGS: next generation sequencing; PCR: polymerase chain reaction; PJS: Peutz–Jeghers syndrome; PolyPhen-2: prediction of functional effects of human nsSNPs, rB1: revised Bethesda criterium 1; rB2: revised Bethesda criterium 2; rB3: revised Bethesda criterium 3; rB4: revised Bethesda criterium 4; rB5: revised Bethesda criterium 5; SIFT: Sorting Intolerant from Tolerant; VUS: variants of uncertain significance

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: ELAN-Fonds; Grant number: 11.08.30.1; Grant sponsor: "Verein zur Förderung des Tumorzentrums" Friedrich Alexander Universität Erlangen-Nürnberg; Grant sponsor: German Federal Ministry of Education and Research (BMBF, Polyprobe-Study); Grant numbers: 01ES0807 and 01ES1001

*C.K., T.T.R., P.L., W.H., A.H. and A.R. contributed equally to this work

DOI: 10.1002/ijc.29149

History: Received 31 Mar 2014; Accepted 25 July 2014; Online 20 Aug 2014

Correspondence to: Cornelia Kraus, Institute of Human Genetics, Friedrich-Alexander-Universität Erlangen Nürnberg, Schwabachanlage 10, 91054 Erlangen, Germany, Tel.: +[49-9131-8522318], Fax: +[49–9131-8523232], E-mail: cornelia.kraus@uk-erlangen.de

What's new?

It's important to find out whether a colorectal tumor has arisen spontaneously or from an inherited mutation, but only those patients whose tumors match clinical criteria for a hereditary CRC syndrome get screened for germline mutations. Thus, many familial tumors may not be identified as such. This study aimed to find out whether screening newly diagnosed colorectal tumors without regard for histology would identify more hereditary disease. They found that an unbiased screening using next generation sequencing (NGS) did indeed identify more germline mutations than the traditional method; of 3 mutations discovered, 2 would have been missed by current strategies based on clinicopathological presentation. NGS does identify non-penetrant mutations, though, which could be problematic for use with patients.

nonpolyposis colorectal cancer (HNPCC) is the only CRC syndrome for which a molecular functional test from tumor material is available. Microsatellite instability (MSI) and/or immunohistochemistry (IHC) testing of tumors has emerged as a sensitive tool to identify individuals who develop CRC as a result of highly penetrant mismatch repair (MMR) gene mutations. Testing is currently confined to patients fulfilling clinical criteria such as the revised Bethesda criteria, which rely heavily on a positive family history or young age at disease manifestation. These criteria probably underestimate the actual disease incidence.[5,6] All studies to date addressing this question preselected index patients based on either clinical criteria or tumor pathology.[7–11]

Historically, the search for germline mutations in at risk individuals focused on gene mutations associated with highly penetrant disease phenotypes. This stepwise approach is an expensive strategy with limited capture rate and leading to underestimation of familial cases.[12] Next Generation sequencing (NGS) now offers the possibility of simultaneously screening a large set of genes, both with high- and low-penetrance, in a short time frame and at moderate cost. Therefore, we established a customized NGS gene panel simultaneously interrogating 18 high- and low-penetrance genes previously reported to harbor germline mutations associated with CRC[13,14] and used it to investigate all consecutive CRC cases seen in a single university hospital during a period of 18 months. Simultaneously, all relevant clinicopathological criteria, tumor MMR deficiency and somatic *BRAF* mutation were assessed and germline mutation screening was performed in an interdisciplinary approach.

Patients and Methods

Patients

We prospectively investigated all patients who underwent primary CRC surgery in the Department of Surgery, University Hospital Erlangen due to a primary histologically confirmed colorectal carcinoma during a period of 18 months. Exclusion criteria were a history of chronic inflammatory bowel disease (ulcerative colitis and Crohn's disease) and surgery for metastatic disease or local tumor recurrence. All patients underwent a standardized clinical evaluation using Bethesda and Amsterdam criteria. This study was approved by the ethics committee of the Medical Faculty of the University Erlangen-Nuernberg (Re.No. 4515, 07/27/2011). All patients provided written informed consent. A blood sample was obtained for DNA extraction and tumor material was analyzed for all patients. Detailed demographic data is provided in Supporting Information Table S1.

Tumor expression of MMR proteins

The expression of MLH1, MSH2, MSH6 and PMS2 was assessed by IHC on 3-μm formalin fixed paraffin embedded (FFPE) tissue sections on a Ventana Benchmark autostainer (Ventana Medical Systems, Tucson, AZ). Antibodies were MLH1 (Clone M3640, Dako, Glostrup, Denmark), MSH2 (clone G219–1129, Cell Marque, Rocklin, CA), MSH6 (clone 44/MSH6, BD Transduction Laboratories, Franklin Lakes, NJ) and PMS2 (clone EP51, Dako, Glostrup, Denmark). Adjacent normal tissue served as internal control. Staining was evaluated by two experienced GI pathologists (TTR, KEW).

Tumor MSI, *BRAF*, *KRAS* and *MLH1* methylation analysis

Molecular pathology analyses were performed on tumor DNA and corresponding normal mucosal DNA extracted from microdissected 5-μm sections of FFPE tumor tissue with the Qiagen FFPE DNA Kit (Qiagen, Hilden, Germany). In case of neoadjuvant therapy primary biopsy material was used with respect to marginal tumor content or conspicious germline mutations. Two batches of 5 microsatellite markers were assessed as recommended in previous studies (initial BAT25, BAT26, D2S123, D17S250, D5S346 and confirmative BAT40, MYCL1, D18S58, D10S197, D13S153).[15–17] The degree of MSI for this study was scored as stable (MSS, 0-1 markers), low instable (MSI-low, 2-4 markers), and highly instable (MSI-high, 5-10 markers). *BRAF* and *KRAS* mutational status was assessed with corresponding pyrosequencing kits (Qiagen, Hilden, Germany). *MLH1* methylation status was quantified after bisulfite treatment and corresponding pyrosequencing (Bisulfite and pyrosequencing kits, Qiagen, Hilden, Germany).

NGS and MLPA analysis

Genomic DNA was extracted according to standard procedures with an automated chemagic MSM I system (Perkin Elmer, Baesweiler, Germany). A customized Ion AmpliSeq

Table 1. CRC associated genes analyzed with NGS

Gene	Gene acession number	Coding sequence covered	Associated disease
MLH1	NM_000249.3	100%	HNPCC
MSH2	NM_000251.1	98.54%	HNPCC
MSH6=GTBP	NM_000179.2	99.24%	HNPCC
PMS2[1]	NM_000535.5	0	HNPCC
PMS1	NM_000534.1	92.35%	HNPCC
MLH3	NM_001040108.1	97.52%	HNPCC
MSH3	NM_002439.4	96.53%	HNPCC
EXO1	NM_0060274.4	99.41%	Possibly HNPCC
TGFBR2	NM_003242.5	96.51%	HNPCC
TGFBR1	NM_004612.2	92.26%	Possibly HNPCC
APC	NM_000038.4	99.51%	FAP
MUTYH	NM_001128425.1	100%	MAP
STK11	NM_000455.4	100%	PJS
MADH4/(SMAD4)	NM_005359.5	98.55%	JPS
BMPR1A	NM_004329.2	99.87%	JPS
TP53	NM_00054546.5	96.63%	LFS1
PTEN	NM_000314.4	98.60%	CWS1
CDH1	NM_004360.3	100%	HDGC

[1]Analyzed by Sanger sequencing.
Abbreviations: CWS1: Cowden syndrome; FAP: Familial adenomatous polyposis; HDGC: hereditary diffuse gastric cancer; HNPCC: Hereditary nonpolyposis colorectal cancer; JPS: Juvenile polyposis syndrome; LFS1: Li–Fraumeni syndrome; MAP: MYH-associated polyposis; PJS: Peutz–Jeghers syndrome.

Panel using Ion AmpliSeq™ Designer (Life technologies, Carlsbad) with 200 bp amplicons and a 5 bp "padding" at the ends of exons was developed for 18 CRC associated genes (Table 1) and used to generate target amplicon libraries. The design consisted of 49 KB of total sequence, 492 amplicons and covered 97% of targeted bases. Briefly, 10 ng of DNA derived from peripheral blood or FFPE tissue was used to prepare barcoded libraries using IonXpress barcoded adapters (Life technologies, Carlsbad USA). Barcoded libraries were combined when possible to a final concentration of 20 pM. Templated beads were generated and after enrichment sequencing was performed on the Ion Torrent PGM following the recommended protocol. Reads were aligned to the reference human genome sequence (hg19) using Torrent Suite 3.4.2. Data was analyzed with the SeqNext module of the Sequence Pilot software (JSI medical systems GmbH, Kippenheim, Germany). *PMS2* was analyzed by conventional Sanger Sequencing since the large family of pseudogenes in the human genome precluded a specific multiplex amplification.

Additionally, multiplex ligation-dependent probe amplification (MLPA) analyses were performed using the SALSA MLPA kits P003 (MLH1/MSH2/EPCAM) and P072 (MSH6/EPCAM) (MRC Holland, Amsterdam, The Netherlands) according to the manufacturer's protocol. Separation of the fragments was performed on an ABI 3100 genetic analyzer (Life technologies, Carlsbad, CA) and allele dosage was assessed using the MLPA module of the Sequence Pilot software (JSI medical systems, Kippenheim, Germany).

In silico analysis

To predict the potential role of the identified nonsynonymous germline variants on protein function we used 4 web-based algorithms: SIFT,[18] Polyphen-2,[19] Align-GVGD[20] for *MLH1*, *MSH2* and *MSH6* missense mutation and additionally, MAPP-MMR[21] for *MLH1* and *MSH2* missense mutations. Splice site mutations were tested by two different web-based splicing effect prediction tools, BDGP: Splice Site Prediction by Neural Network[22] and NetGene2 Server.[23,24] Variant frequencies were compared with European-American and African-American control samples from the Exome Variant Server online database[25] to exclude rare polymorphisms.

Results

Study cohort characteristics

A total of 190 patients with CRC initially fulfilled inclusion criteria but 27 (15%) patients refused consent including 3 patients who were previously diagnosed with a hereditary CRC (2 HNPCC, 1 MYH-associated polyposis, MAP). Further 11 patients were excluded for various other reasons (*e.g.*, emergency surgery or language barrier). Of the remaining 152 cases, 20 (13%) participated but refused receiving any information regarding the results of the study. Patients were 104 (68%) males and 48 (32%) females with a median age of

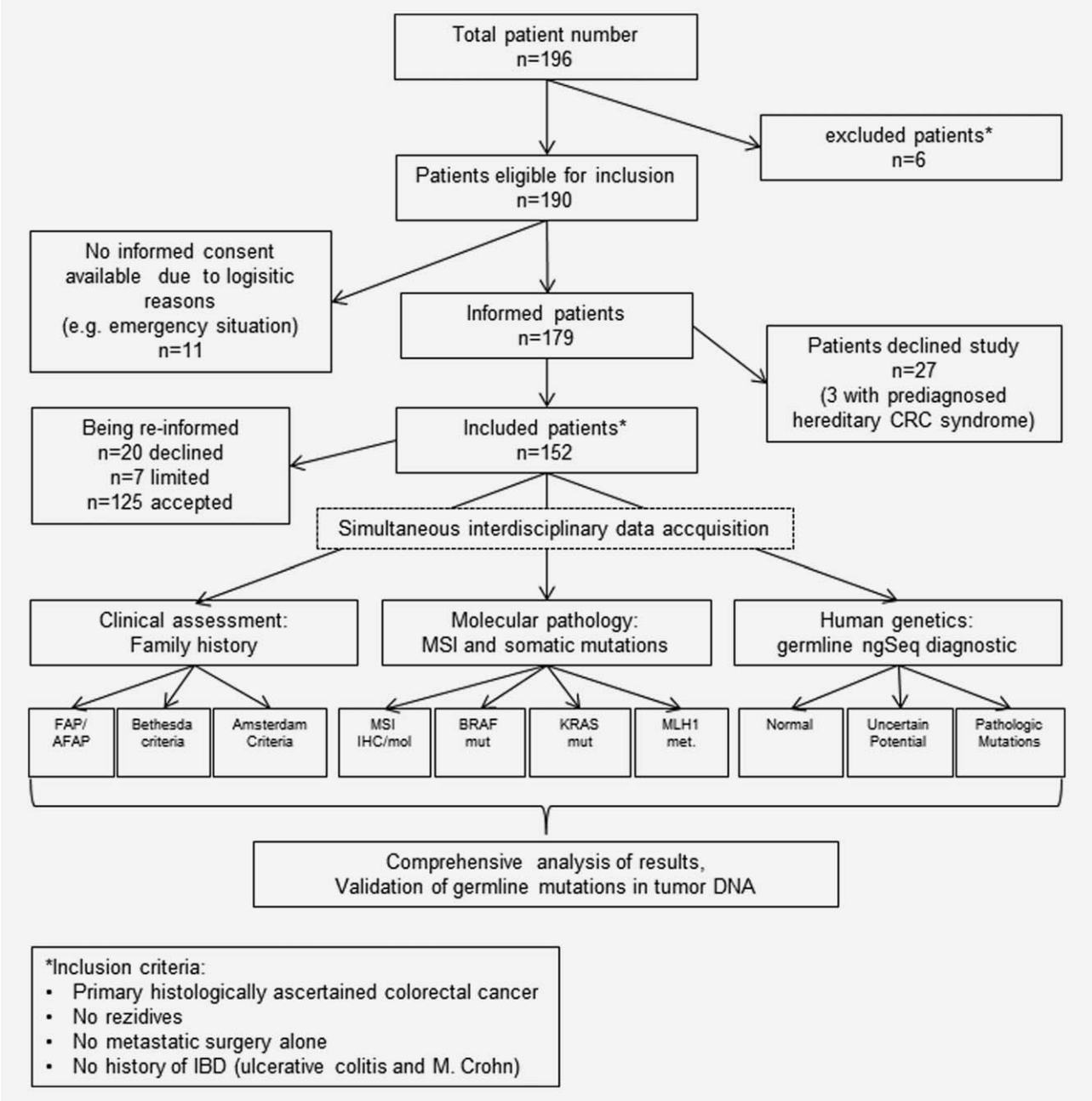

Figure 1. Overview of study design. Numbers of patients included and excluded into the study are shown. Interdisciplinary data acquisition combined clinical, molecular pathologic and genetic analyses. All included patients were analyzed for germline mutations in 18 CRC associated genes.

68 and a mean age of 66.5 (range 26–88); 81 (54%) patients presented with colonic cancer (coecum 12; ascending colon 19; right flexure 4; transverse colon 8, left flexure 2, descending colon 7, sigmoid 29) and 70 (46%) with rectal cancer (Supporting Information Table S1). One patient had two simultaneous carcinomas of the descending colon and the rectum; 38 of the 152 patients fulfilled at least one of the revised Bethesda criteria (Table 3) while only 2 fulfilled Amsterdam criteria (Fig. 1).

Pathological and MSI findings

Classical morphology was used as initial screening for a possible hereditary background of CRC. There were 14 patients with morphological subtypes of CRC linked to HNPCC syndrome. No patient met the number of polyps necessary for a polyposis syndrome. IHC for MMR protein showed intact nuclear expression in 145 cases. Protein loss was observed in 7 patients and was confined to MLH1 and PMS2. Functional molecular MSI testing was confirmative in these cases. MLH1

Table 2. Pathology

Microsatellite status	Histomorphology	MMR-IHC	Interpretation	MLH1 promoter methylation	BRAF mutation
MSI: n = 7	Serrated CRC: 1/7	MLH1: 4/7	HNPCC	1/7 not methyl.	1/7 WT
	Signet ring cell CRC: 0/7	MSH2: 0/7	Sporadic MSI-H	6/7 hypermeth.	5/7 mut
	Medullary CRC: 1/7	MSH6: 0/7			1/7 WT
	Mucinous CRC: 0/7	PMS2: 7/7			
MSS: $n = 145$	Serrated CRC: 5/145	MLH1: 0/145	MSS with MLH1 meth.	3/3 hypermeth	3/3 WT
	Signet ring cell CRC: 1/145	MSH2: 0/145	MSS NOS	135/138[1] not methyl	4/145 mut
	Medullary CRC: 1/145	MSH6: 0/145			141/145 WT
	Mucinous CRC: 5/145	PMS2: 0/145			
Total number	14/152 (9.2%)	7/152 (4.6%)		9/145[1] (6.2%)	9/152 (5.9%)
Chi-Square	$p = 0.0697$	$p < 0.001$		$p < 0.001$	$p < 0.001$

Tumor specific characterization as MSI or MSS tumor. Dependency from known parameters like histomorphology, MMR status, MLH1 methylation analysis and BRAF status were calculated with Chi-square tests. Interestingly MLH1 promoter hypermethylation corresponded better to sporadic MSI tumor than BRAF mutation, but 7 specimens[1] could not be analyzed.

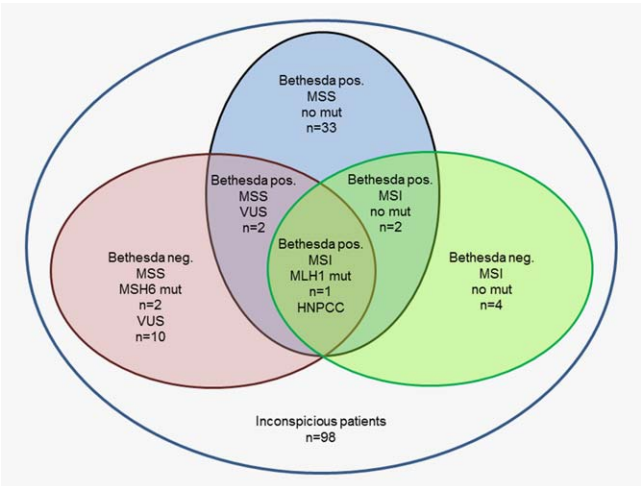

Figure 2. Venn diagram of study results of the disciplinary approaches. The three disciplinary approaches identified different groups of patients being suspicious for a hereditary CRC: Clinical (blue), tumor pathological (green) and genetic (red) investigation. Only one patient was identified by all three approaches (center) and was diagnosed with Lynch syndrome (HNPCC). The other patients with MSI tumors, both Bethesda positive and negative, had sporadic cancers as shown by epigenetic MLH1 promotor inactivation. Out of the Bethesda positive patients, two had a class 3 variant in *MSH2* although the tumor was sporadic. Genetic analysis alone identified two patients with class 5 *MSH6* truncating mutations, 1 with a previously reported *APC mutation* and further 9 uncertain, potentially pathogenic class 3 variants in *APC*, *MLH1, MSH2, MSH6, MSH3* and *MLH3*. Some of the 33 patients fulfilling at least one of the Bethesda criteria might carry predisposing genetic factors elsewhere in the genome.

promoter methylation identified 6 out of 7 patients with MMR protein loss as having a cancer of sporadic origin. MLH1 methylation analysis seems to be more sensitive than *BRAF**V600E mutation, which only identified 5 out of the 7 patients. However, the aggressive bisulfite treatment led to a higher dropout rate in MLH1 promoter analysis (145 cases evaluable). Detailed data are outlined in Table 2.

Germline DNA-variants

NGS based mutation screening uncovered 27 unique sequence variants (Table 4). These included 2 small frameshift deletions or insertions, 1 nonsense mutation, 21 missense and 3 splice site mutations. Four variants (*APC*: p.Arg414Cys, *MLH1*: p.Gly67Arg, *MSH6*: p.Tyr977* and p.Asp1171Glufs*5) were previously published as pathogenic,[26–29] whereat the *APC*: p.Arg414Cys mutation is more recently reassessed as an unclassified variant.[30] 8 variants have been already reported by the InSiGHT Colon Cancer Gene Variant Database for MMR and other colon cancer susceptibility genes hosted by the Leiden Open Variation Database (LOVD).[31–35] If available, classification of the variants was performed according to the information provided by this database. To further characterize the remaining alterations all variants were categorized based on their predicted effect on the mRNA and amino acid levels and defined as deleterious if they resulted in a premature termination codon. Missense and splice-site variants were analyzed using various web-based algorithms as described in the material and methods section and variant frequencies were compared with European-American and African-American control samples from the Exome Variant Server online database to exclude rare polymorphisms. Finally, the criteria of the International Agency for Research on Cancer (IARC) classification system was used to assign these variants to the five-class IARC system.[36–38] Based on the above mentioned criteria 3 variants (*MLH1*: p.Gly67Arg, *MSH6*: p.Tyr977* and p.Asp1171Glufs*5) were classified as class 5 mutation (pathogenic), 12 were classified as class 3 mutation (uncertain), 10 as class 2 mutation (likely not pathogenic) and 2 as class 1 mutation mutation (not pathogenic) (Table 4). The mean age of these patients was 66 years (range 42–84 years), with only one patient younger than 50 years at the time of cancer resection (Table 4). Three patients fulfilled at least one of the Bethesda criteria but only one with an *MLH1* missense mutation p.Gly67Arg showed MSI and loss of MLH1 by IHC in the tumor (Table 4).

Table 3. Summary Bethesda positive patients

Group	*n*	ID	Mean age of diagnosis	rB1	rB2	rB3	rB4	rB5
1	1	2[1]	33	+	+	−	+	+
2	1	128	68	+	+	−	−	−
3	1	20	49	+	−	+	−	−
4	1	46	67	−	+	−	+	−
5	2	60, 71	72 (70–74)	−	+	−	−	+
6	2	47,64	63 (56–70)	−	−	−	+	+
7	10	1, 16, 33,[2] 40, 57, 91, 104, 108, 135, 141	44 (26–50)	+	−	−	−	−
8	5	14, 25,[2] 87, 95, 140	74 (64–82)	−	+	−	−	−
9	4	73, 103, 139, 146,	66 (55–80)	−	−	+	−	−
10	2	41, 54	68 (65–71)	−	−	−	+	−
11	9	15, 28, 39, 43, 55, 61, 76, 117, 131	67 (51–85)	−	−	−	−	+

[1]Class 5 MLH1 mutation.
[2]Class 3 MSH2 mutation.
Abbreviations: Revised Bethesda criteria. rB1: Colorectal cancer diagnosed in a patient who is <50 years of age; rB2: Presence of synchronous, metachronous colorectal or other HNPCC associated tumors, regardless of age; rB3: Colorectal cancer with MSI-H histology diagnosed in a patient who is < 60 years of age; rB4: Individuals with colorectal cancer and one or more first-degree relatives with an Lynch-syndrome associated tumor, with one of the cancers being diagnosed under age 50 years; rB5: Individuals with colorectal cancer and two or more first- or second-degree relatives with Lynch-syndrome related tumors, regardless of age.

Somatic variants in tumor DNA

To further characterize tumor status we analyzed the same gene panel used for germline variant testing also in tumor DNA from the 11 patients with predictive causative mutations, the 12 patients with VUS and 5 patients without any germline mutation. In all tumor samples the initial germline mutation was verified. None of the tumor samples showed loss-of-heterozygosity. Overall we identified 45 somatic mutations in 19 tumor samples clustering in 4 genes. The most affected gene was *APC* (24 mutations) followed by *TP53* (16 mutations), *SMAD4* (4 mutations), and *CTNNB1* (one mutation). All 24 *APC* mutations were loss of function mutations (Supporting Information Table S2). We found no difference in tumor mutation profile between patients with or without germline mutation. The mutation profiles identified in the tumor samples are consistent with a nonhypermutated cancer mutation profile as proposed by The Cancer Genome Atlas Network.[39] In 2 cases, one with a germline *APC* missense mutation and one with a VUS, we found a somatic nonsense second hit mutation suggestive of an attenuated familial adenomatous polyposis (FAP). Interestingly, these cases did not show polyposis, compatible with the variable spectrum of attenuated FAP with older age of onset and fewer to no polyps.

Discussion

Making a diagnosis of a hereditary cancer syndrome has significant implications for the medical management of CRC patients and their families. To our best knowledge, all studies to date addressing this question preselected their study group based on either clinical features or histopathological findings due to the difficulties of a comprehensive mutation testing with conventional technologies.[7–11] NGS now offers the possibility of a comprehensive approach without any preselection. We therefore compared current clinicopathological approaches with an unbiased NGS based mutation screening. This self-designed gene panel contained 18 high- and low-penetrant CRC genes and was applied to 152 consecutive CRC cases seen in a single university hospital during a period of 18 months (Fig. 1).

With our customized NGS gene panel we were able to sequence 98% of the targeted sequence of 17 out of the 18 associated CRC genes with an average depth of 50 fold. Only PMS2 was not amenable to this approach due to the large family of PMS2 pseudogenes in the human genome precluding a multiplex amplification. This gene was therefore investigated by conventional Sanger sequencing. With this comprehensive approach we identified 27 variants in 9 genes in 23 (18%) of the 152 patients studied (Table 4). To ascertain the significance of this surprisingly high number of unique sequence variants we determined the DNA-variant frequency in the same 9 genes from 359 in-house noncancer patients, who received whole exome sequencing. These control individuals showed significantly less unique variants (25/359; 7%; $p = 0.0003$, Fishers exact test), none of which was protein truncating. This suggested that the majority of the 27 variants detected in the study group may indeed be associated with CRC. Of these variants, 3 (*MLH1*: p.Gly67Arg in patient 2, *MSH6*: p.Tyr977* in patient 5 and p.Asp1171Glufs*5 in patient 115) could clearly be classified as class 5 mutations (pathogenic) whereas 12 were classified as class 3 (uncertain; Table 4).

Table 4. Summary patients with unique mutations

ID	Age	Sex	rB	ICH	MS-status	BRAF * V600E	Gene	cDNA-change	AA-change	Consequence	ESP (MAF %)	PolyPhen-2	SIFT	MAPP-MMR	Align-GVGD	IARC Classification
2	53	F	+	MLH1 loss	MSI	WT	MLH1	c.199G > A[1]	p.Gly67Arg[1]	Missense	0	PrD	NT	–	C65	Class 5 (InSiGHT)
5	74	M	–	normal	MSS	WT	MSH6	c.2931C > G[2]	p.Tyr977*[2]	Nonsense	0.0077	–	–	–		Class 5 (InSiGHT)
115	56	M	–	normal	MSS	WT	MSH6	c.3513_3514delTA[3]	p.Asp1171Glufs*5[3]	Frame-shift	0	–	–	–		Class 5 (InSiGHT)
51	84	F	–	normal	MSS	WT	APC	c.1240C > T[4]	p.Arg414Cys[4]	Missense	0.0384	PrD	NT	–		Class 3
80	59	W	–	normal	MSS	WT	APC	c.1775T > C	p.Leu592Ser	Missense	0	PrD	NT	–		Class 3
138	73	M	–	normal	MSS	WT	MLH1	c.2174G > A	p.Arg725His	Missense	0	PrD	NT	3.03	C0	Class 3 (InSiGHT)
9	52	M	–	normal	MSS	WT	MLH3	c.871A > T	p.Asn291Tyr	Missense	0	PrD	NT	–		Class 3
25	82	M	+	normal	MSS	WT	MSH2	c.157G > A	p.Ala53Thr	Missense	0	PrD	T	5.74	C0	Class 3
132	68	M	–	normal	MSS	WT	MSH2	c.1601G > A	p.Arg534His	Missense	0	PrD	NT	8.18	C25	Class 3
129	57	M	–	normal	MSS	WT	MSH2	c.2187G > A	p.Met729Ile	Missense	0	PrD	NT	21.99	C0	Class 3[5] (InSiGHT)
33	42	M	+	normal	MSS	WT	MSH2	c.2211-6C > A		Splicing?	0	–	–	–		Class 3
125	69	W	–	normal	MSS	WT	MSH3	c.2785A > T	p.Ile929Phe	Missense	0	PrD	NT	–		Class 3
130	77	M	–	normal	MSS	WT	MSH3	c.3130+3A > G		Splicing?	0	–	–	–		Class 3
4	68	F	–	normal	MSS	WT	MSH6	c.4062_4065dupGACT[5]	p.Leu1356Aspfs*4[5]	Frame-shift	0	–	–	–		Class 3
48	70	W	–	normal	MSS	WT	MSH6	c.3758T > A	p.Val1253Glu	Missense	0.0154	PrD	NT	–	C65	Class 3
37	78	M	–	normal	MSS	WT	APC	c.5026A > G	p.Arg1676Gly	Missense	0.0154	B	T	–		Class 2
37	78	M	–	normal	MSS	WT	APC	c.7399C > A	p.Pro2467Thr	Missense	0.0077	PoD	T	–		Class 2
72	52	M	–	normal	MSS	WT	APC	c.3245C > G	p.Thr1082Ser	Missense	0	B	T	–		Class 2
8	67	M	–	normal	MSS	WT	MSH6	c.1624C > G	p.Leu542Val	Missense	0	B	T	–	C0	Class 2
10	80	M	–	normal	MSS	WT	MSH6	c.3986C > T	p.Ser1329Leu	Missense	0.0077	B	T	–	C0	Class 2 (InSiGHT)
72	52	M	–	normal	MSS	WT	PMS1	c.174_175delinsTA	p.Glu59Lys	Missense	0.2614	B	NT	–		Class 2
26	54	W	–	normal	MSS	WT	CDH1	c.2329G > A	p.Asp777Asn	Missense	0.0077	PrD	T	–		Class 2
150	76	M	–	normal	MSS	WT	MLH3	c.1004C > T	p.Thr335Ile	Missense	0	B	T	–		Class 2
147	74	M	–	normal	MSS	WT	MSH2	c.1986G > C	p.Gln662His	Missense	0	B	T	2.42	C0	Class 2
48	70	W	–	normal	MSS	WT	STK11	c.1211C > T	p.Ser404Phe	Missense	0.0416	B	NT	–		Class 2
74	64	M	–	normal	MSS	WT	MSH6	c.3802-44dupT		Splicing?	7.3494	–	–	–		Class 1
9	52	M	–	normal	MSS	WT	APC	c.7862C > G	p.Ser2621Cys	Missense	0.3308	PoD	T	–		Class 1

Already published as disease causing.
[1]Tannergard *et al*. 1995.
[2]Cederquist *et al*. 2004.
[3]Plaschke *et al*. 2004.
[4]Nishisho *et al*. 1991
[5]InSiGHT: concluded pathogenity: pathogenic.
[6]In InSiGHT a different base pair change is annotated (c.2187G > T) but leading to the same aa-change.
Abbreviations: Age: age at recruitment; F: female; M: Male; rB: revised Bethesda criteria; +: positive; –: negative; B: benigne; class 1: not pathogenic; class 2: likely not pathogenic; class 3: uncertain; class 5: pathogenic; ESP: NHLBI Exome Sequencing Project (Exome Variant Server); IARC: International Agency for Research on Cancer; IHC: Immunohistochemistry; InSiGHT: InSiGHT Colon Cancer Gene Variant Database for MMR and other colon cancer susceptibility genes hosted by the Leiden Open Variation Database; MAF: minor allele frequency; MAPP-MMR: multivariate analysis of protein polymorphisms-mismatch repair (Available for MLH1 and MSH2 only); MSI: microsatellite instable; MSS: microsatellite stable; NT: not tolerated; PoD: possibly damaging; PolyPhen-2: prediction of functional effects of human nsSNPs; PrD: probably damaging; SIFT: Sorting Intolerant from Tolerant; T: tolerated; WT: wildtype.

To confirm the pathogenicity of these variants, we investigated DNA of the corresponding tumors for further somatic mutations using the same gene panel. The fragmented tumor DNA from FFPE samples was perfectly suited for the NGS analyses probably because of the short-length of the NGS amplicons, with detection and success rates similar to genomic DNA. We investigated tumor samples of the 23 patients with germline variants as well as 5 tumors from individuals without such variants confirming the unique germline variant in all tumors in the heterozygous state, indicating no loss-of-heterozygosity. In all tumor samples with an MMR germline variant no second hit mutation could be identified. In contrast, two tumors, one with a class 3 *APC* germline variant (p.Arg414Cys in patient 51) and one with a class 1 *APC* germline variant (p.Ser2621Cys in patient 9), respectively, did harbor a second-hit *APC* mutation, suggesting that these individuals might have attenuated FAP without the necessary number of polyps for a clinicopathological diagnosis. The APC variant p.Arg414Cys has been initially described as a germline mutation of probable pathogenicity.[26] More recently, this variant was reported in a patient with attenuated FAP. However, in this case, the tumor had lost the allele carrying the p.Arg414Cys mutation questioning its pathogenicity.[30] In our patient, though, the p.Arg414Cys mutation was also present in the tumor and in addition a second truncating mutation was identified, indicating that it may indeed be associated with aFAP (Supporting Information Table S2). Altogether, we identified 45 somatic mutations mainly clustering in 4 genes (24 *APC*, 16 *TP53*, 4 *SMAD4* and 1 *CTNNB1*; Supporting Information Table S2). The mutation profile of all but one was typical for a nonhypermutated mutation profile as described by The Cancer Genome Atlas Network.[39] In agreement with the sequencing results, these tumors showed MSS and normal results on IHC. Altogether, our data indicate that these tumors were not MMR deficient and thus the MMR germline variants identified were nonpenetrant in these cases.

We found a high concordance rate when comparing tumor pathology and both germline (genomic) and somatic (tumor) DNA sequencing (Fig. 2). Five MSI tumors were *BRAF**V600E positive and did not harbor any germline mutation but exhibited a hypermutated tumor mutation profile indicative of a sporadic MMR deficient tumor. In one patient with an MSI tumor but neither a germline nor *BRAF* mutation we found a high level of *MLH1* promoter methylation, suggesting a sporadic origin of this cancer as well (Table 2). In contrast, patient 2 with an MSI tumor without *BRAF* mutation and with loss of MLH1 by IHC presented a previously reported *MLH1* germline mutation (p.Gly67Arg)[29] confirming the diagnosis of Lynch syndrome in this patient (Table 4). Thus our results confirm that the detection of the *BRAF**V600E mutation in an MSI tumor is a good predictor of a sporadic MMR deficient tumor.[40–42] *MLH1* promoter methylation might serve as a substitute for *BRAF* mutation analysis, but has a higher dropout rate due to the aggressive bisulfite treatment of DNA (Table 2).

Altogether, besides the *MLH1* p.Gly67Arg mutation identified in the classical Lynch-syndrome patient we found 2 additional class 5 and 10 class 3 germline variants with a prediction of pathogenicity in 12 patients, both in classical high- and low-penetrant MMR genes as well as in *MLH3*, whose association with Lynch syndrome is controversial (Table 4).[43] Interestingly, none of the tumors of these patients showed any evidence for MMR deficiency, neither on molecular pathology nor by somatic DNA-sequencing, suggesting that these tumors arose independently from a possible genetic predisposition for Lynch syndrome. Nevertheless the premature protein truncating mutations found in *MSH6* are highly suspicious of being pathogenic, two mutations p.Tyr977* (patient 5) and p.Asp1171Glufs*5 (patient 115) have even been previously associated with Lynch syndrome.[27,28] The third mutation p.Leu1356Aspfs*4 (patient 4) is located in the last coding exon shortening the protein only by two aminoacids. In the InSiGHT Colon Cancer Gene Variant Database (LOVD) a similar duplication c.4064_4065insGTCA is classified as class 3 mutation but was concluded as being pathogenic [patient data (#0021205)]. Therefore, the pathogenicity of this variant remains uncertain. *MSH6* has been proposed as a low-penetrance MMR gene where not all tumors exhibit MMR deficiency.[44–46] Nevertheless, such premature termination mutations in *MSH6* have been reported to increase lifetime risk for CRC in men to 44% and in women to 20% by age 80 and for endometrial carcinoma to 44% by age 80.[47] Therefore we conclude that although the tumors were unrelated to the *MSH6* mutation, at least two of these patients should be considered as predisposed for Lynch syndrome and should be monitored accordingly. It is conceivable, that the 10 class 3 missense and splice site mutations identified in the MMR genes are nonpathogenic, although they affect evolutionary highly conserved positions and were not reported as common polymorphisms in the large variant databases interrogated (Table 4). No further family members neither affected nor unaffected were available for segregation analysis. Interestingly, two of the patients, patient 25 with a *MHS2* missense variation (p.Ala53Thr) and patient 33 with a *MSH2* splice site variation (c.2211-6C>A), fulfilled one of the revised Bethesda criteria. Patient 25 was 82 years of age and presented two Lynch-syndrome associated cancers and therefore fulfilled the revised Bethesda criterium 2. Patient 33 was 42 years of age fulfilling the revised Bethesda criterium 1. Unfortunately, he refused further investigations. Therefore, further clarification of the class 3 MMR gene variations would require functional *in-vitro* testing, which is not routinely performed and was beyond the scope of this study.

Next, we investigated the clinical benefit of our unbiased approach compared with current clinicopathological approaches. A preselection by means of clinical criteria, histopathology or medical genetics would lead to different results regarding patients identified (Fig. 2). Of the 38

patients fulfilling at least one of the revised Bethesda criteria, one patient (patient 2) was confirmed to have Lynch syndrome with a class 5 *MLH1* germline mutation (p.Gly67Arg) and MSI-H in his tumor. This patient fulfilled 4 of the revised Bethesda criteria (Group 1; Table 3). In two patients (25 and 33) fulfilling only one of the revised Bethesda criteria namely rB1 (group 7) and rB2 (group 8) class 3 mutations were identified. The low percentage of Bethesda positive patients harboring a MMR germline mutation maybe due to the fact that the potential of identifying a patient with Lynch syndrome differs significantly depending on the individual clinical criteria[48]; 30 (87%) out of the 38 positive patients fulfilled only one of the revised Bethesda criteria. Of these, 23 (77%) patients fuIfilled only rB1 (group 7), rB3 (group 9) or rB5 (group 11), respectively (Table 3). These groups are reported to have a much lower MMR mutation frequency compared to patients with more than one Lynch syndrome associated tumor in their history (rB2, group 8) or patients with one or more affected first-degree family members and one at young age (rB4, group 10).[48] When considering only MMR-deficient tumors based on MSI and IHC testing, 7 out of the 152 patients would have been initially selected for mutation analysis. Five of them, though, tested positive for the *BRAF**V600E somatic mutation and one had a causative *MLH1* promoter methylation in the tumor excluding a Lynch syndrome. As expected none of them carried a germline mutation. The seventh was the *MLH1* positive Lynch syndrome patient discussed above. The unbiased approach identified 3 cases with class 5 mutations (2%) and 12 class 3 mutations (8%), 12 of which would have been missed with current approaches, that is, Bethesda criteria plus MMR deficiency in the tumor: 2 with a premature termination mutation in *MSH6*, 1 with a previously reported mutation in *APC*, as well as 11 uncertain but nevertheless potentially pathogenic class 3 variants in the *APC* gene and in the MMR genes *MLH1, MSH2, MSH6, MSH3* and *MLH3* (Fig. 2 and Table 4).

Therefore an unbiased genetic mutation detection approach seems more efficient than the currently used clinicopathological approaches followed by sequential gene testing, since it overcomes the problem of genetic heterogeneity and overlap in clinical presentation between the various CRC syndromes. In addition, panel testing is much more economical and less time consuming than sequential testing of candidate genes. The downside of this approach is a considerable increase of nonpenetrant variants and those of unknown significance as already anticipated by Domchek *et al.*[49] The interpretation of these variants can be challenging, as our knowledge of the genotype/phenotype correlation is still limited. To better discern risk-associated variants from neutral polymorphisms the compilation of sequence variants together with phenotypic information such as in the InSiGHT Colon Cancer Gene Variant Database for MMR and other colon cancer susceptibility genes hosted by the Leiden Open Variation Database (LOVD; http://www.insight-group.org/variants/database/)[31–35] is urgently needed. Thus at this stage genetic susceptibility testing should be offered only in the context of clinical trials with appropriate informed consent and follow-up. Although in our study most patients requested being informed about study results, a noticeable number of patients (13%) provided a sample but preferred not to receive any results after pretest counseling. This number is similar to that of the study by Ward *et al.*,[11] reflecting the anxiety that results of susceptibility testing can evoke in patients and their relatives. The work to compile and analyze the data needed to reclassify genetic variants and quantify the magnitude of cancer risks will continue to require multidisciplinary collaborations between clinicians, geneticists, molecular biologists and statisticians.

In conclusion, our study shows that an unbiased approach to genetic mutation detection in CRC is more efficient to identify patients at risk for hereditary disease than currently used clinicopathological approaches followed by sequential gene testing. Due to our limited knowledge of the pathogenicity of genetic variants, interpretation of results can be challenging in some cases.

Acknowledgements

The authors thank all patients for their participation in the study, Karin Erdtmann for excellent assistance in patient recruiting and clinical documentation and Heike Friebel-Stange, Madeleine Demleitner and Angelika Diem for excellent technical assistance.

References

1. Lichtenstein P, Holm NV, Verkasalo PK, et al. Environmental and heritable factors in the causation of cancer–analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med* 2000;343:78–85.
2. Kastrinos F, Stoffel EM. History, genetics, and strategies for cancer prevention in Lynch syndrome. *Clin Gastroenterol Hepatol* 2014;12:715–27.
3. Manolio TA, Chisholm RL, Ozenberger B, et al. Implementing genomic medicine in the clinic: the future is here. *Genet Med* 2013;15: 258–67.
4. Win AK, Macinnis RJ, Dowty JG, et al. Criteria and prediction models for mismatch repair gene mutations: a review. *J Med Genet* 2013;50: 785–93.
5. de la Chapelle A, Palomaki G, Hampel H. Identifying Lynch syndrome. *Int J Cancer* 2009;125: 1492–3.
6. Syngal S, Fox EA, Eng C, et al. Sensitivity and specificity of clinical criteria for hereditary nonpolyposis colorectal cancer associated mutations in MSH2 and MLH1. *J Med Genet* 2000;37: 641–5.
7. Engel C, Rahner N, Schulmann K, et al. Efficacy of annual colonoscopic surveillance in individuals with hereditary nonpolyposis colorectal cancer. *Clin Gastroenterol Hepatol* 2010;8:174–82.
8. Goecke T, Schulmann K, Engel C, et al. Genotype-phenotype comparison of German MLH1 and MSH2 mutation carriers clinically affected with Lynch syndrome: a report by the German HNPCC Consortium. *J Clin Oncol* 2006;24:4285–92.
9. Steinke V, Holzapfel S, Loeffler M, et al. Evaluating the performance of clinical criteria for predicting mismatch repair gene mutations in Lynch syndrome: a comprehensive analysis of 3671 families. *Int J Cancer* 2014;135:69–77.
10. Teutsch SM, Bradley LA, Palomaki GE, et al. The Evaluation of Genomic Applications in Practice and Prevention (EGAPP) Initiative: methods of

the EGAPP Working Group. *Genet Med* 2009;11:3–14.
11. Ward RL, Hicks S, Hawkins NJ. Population-based molecular screening for Lynch syndrome: implications for personalized medicine. *J Clin Oncol* 2013;31:2554–62.
12. Stoffel EM, Kastrinos F. Familial Colorectal Cancer, Beyond Lynch Syndrome. *Clin Gastroenterol Hepatol* 2014;12:1059–68.
13. National Cancer Institute. Available at: http://www.cancer.gov/. Accessed on November 2011.
14. .COSMIC: catalogue of somatic mutations. Available at: http://cancer.sanger.ac.uk/cancergenome/projects/cosmic/. Accessed on November 2011.
15. Boland CR, Thibodeau SN, Hamilton SR, et al. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res* 1998;58:5248–57.
16. Jass JR, Walsh MD, Barker M, et al. Distinction between familial and sporadic forms of colorectal cancer showing DNA microsatellite instability. *Eur J Cancer* 2002;38:858–66.
17. Umar A, Boland CR, Terdiman JP, et al. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *J Natl Cancer Inst* 2004;96:261–8.
18. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 2009;4:1073–81.
19. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248–9.
20. Tavtigian SV, Deffenbaugh AM, Yin L, et al. Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. *J Med Genet* 2006;43:295–305.
21. Chao EC, Velasquez JL, Witherspoon MS, et al. Accurate classification of MLH1/MSH2 missense variants with multivariate analysis of protein polymorphisms-mismatch repair (MAPP-MMR). *Hum Mutat* 2008;29:852–60.
22. Reese MG, Eeckman FH, Kulp D, et al. Improved splice site detection in Genie. *J Comput Biol* 1997;4:311–23.
23. Brunak S, Engelbrecht J, Knudsen S. Prediction of human mRNA donor and acceptor sites from the DNA sequence. *J Mol Biol* 1991;220:49–65.
24. Hebsgaard SM, Korning PG, Tolstrup N, et al. Splice site prediction in Arabidopsis thaliana pre-mRNA by combining local and global sequence information. *Nucleic Acids Res* 1996;24:3439–52.
25. Exome Variant Server. Available at: http://evs.gs.washington.edu/EVS/. Accessed on May 2013.
26. Nishisho I, Nakamura Y, Miyoshi Y, et al. Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients. *Science* 1991;253:665–9.
27. Cederquist K, Emanuelsson M, Goransson I, et al. Mutation analysis of the MLH1, MSH2 and MSH6 genes in patients with double primary cancers of the colorectum and the endometrium: a population-based study in northern Sweden. *Int J Cancer* 2004;109:370–6.
28. 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. *Hum Mutat* 2004;23:285.
29. Tannergard P, Lipford JR, Kolodner R, et al. Mutation screening in the hMLH1 gene in Swedish hereditary nonpolyposis colon cancer families. *Cancer Res* 1995;55:6092–6.
30. Schlussel AT, Donlon SS, Eggerding FA, et al. Identification of an APC variant in a patient with clinical attenuated familial adenomatous polyposis. *Case Rep Med* 2014;2014:432324.
31. Plazzer JP, Sijmons RH, Woods MO, et al. The InSiGHT database: utilizing 100 years of insights into Lynch syndrome. *Fam Cancer* 2013;12:175–80.
32. Peltomaki P, Vasen H. Mutations associated with HNPCC predisposition – Update of ICG-HNPCC/INSiGHT mutation database. *Dis Markers* 2004;20:269–76.
33. Peltomaki P, de la Chapelle A. Mutations predisposing to hereditary nonpolyposis colorectal cancer. *Adv Cancer Res* 1997;71:93–119.
34. Ou J, Niessen RC, Lutzen A, et al. Functional analysis helps to clarify the clinical importance of unclassified variants in DNA mismatch repair genes. *Hum Mutat* 2007;28:1047–54.
35. Woods MO, Williams P, Careen A, et al. A new variant database for mismatch repair genes associated with Lynch syndrome. *Hum Mutat* 2007;28:669–73.
36. Plon SE, Eccles DM, Easton D, et al. Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. *Hum Mutat* 2008;29:1282–91.
37. Thompson BA, Greenblatt MS, Vallee MP, et al. Calibration of multiple in silico tools for predicting pathogenicity of mismatch repair gene missense substitutions. *Hum Mutat* 2013;34:255–65.
38. Thompson BA, Spurdle AB, Plazzer JP, et al. Application of a 5-tiered scheme for standardized classification of 2,360 unique mismatch repair gene variants in the InSiGHT locus-specific database. *Nat Genet* 2014;46:107–15.
39. Cancer Genome Atlas N. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 2012;487:330–7.
40. Parsons MT, Buchanan DD, Thompson B, et al. Correlation of tumour BRAF mutations and MLH1 methylation with germline mismatch repair (MMR) gene mutation status: a literature review assessing utility of tumour features for MMR variant classification. *J Med Genet* 2012;49:151–7.
41. Thiel A, Heinonen M, Kantonen J, et al. BRAF mutation in sporadic colorectal cancer and Lynch syndrome. *Virchows Arch* 2013;463:613–21.
42. Zlobec I, Bihl MP, Foerster A, et al. Stratification and prognostic relevance of jass's molecular classification of colorectal cancer. *Front Oncol* 2012;2:7.
43. Liu HX, Zhou XL, Liu T, et al. The role of hMLH3 in familial colorectal cancer. *Cancer Res* 2003;63:1894–9.
44. Akiyama Y, Sato H, Yamada T, et al. Germ-line mutation of the hMSH6/GTBP gene in an atypical hereditary nonpolyposis colorectal cancer kindred. *Cancer Res* 1997;57:3920–3.
45. Plaschke J, Kruppa C, Tischler R, et al. Sequence analysis of the mismatch repair gene hMSH6 in the germline of patients with familial and sporadic colorectal cancer. *Int J Cancer* 2000;85:606–13.
46. Wu Y, Berends MJ, Mensink RG, et al. Association of hereditary nonpolyposis colorectal cancer-related tumors displaying low microsatellite instability with MSH6 germline mutations. *Am J Hum Genet* 1999;65:1291–8.
47. Baglietto L, Lindor NM, Dowty JG, et al. Risks of Lynch syndrome cancers for MSH6 mutation carriers. *J Natl Cancer Inst* 2010;102:193–201.
48. Steinke V, Holzapfel S, Loeffler M, et al. Evaluating the performance of clinical criteria for predicting mismatch repair gene mutations in Lynch syndrome: a comprehensive analysis of 3,671 families. *Int J Cancer* 2014;135:69–77.
49. Domchek SM, Bradbury A, Garber JE, et al. Multiplex genetic testing for cancer susceptibility: out on the high wire without a net? *J Clin Oncol* 2013;31:1267–70.