


RESEARCH ARTICLE

Open Access



# Identification of genetic variants for clinical management of familial colorectal tumors

Mev Dominguez-Valentin<sup>1\*</sup> , Sigve Nakken<sup>1</sup>, H el ene Tubeuf<sup>2,3</sup>, Daniel Vodak<sup>1</sup>, Per Olaf Ekstr om<sup>1</sup>, Anke M. Nissen<sup>4,5</sup>, Monika Morak<sup>4,5</sup>, Elke Holinski-Feder<sup>4,5</sup>, Alexandra Martins<sup>2</sup>, P al M oller<sup>1,6,7</sup> and Eivind Hovig<sup>1,8,9</sup>

## Abstract

**Background:** The genetic mechanisms for families who meet the clinical criteria for Lynch syndrome (LS) but do not carry pathogenic variants in the mismatch repair (MMR) genes are still undetermined. We aimed to study the potential contribution of genes other than MMR genes to the biological and clinical characteristics of Norwegian families fulfilling Amsterdam (AMS) criteria or revised Bethesda guidelines.

**Methods:** The Hereditary Cancer Biobank of the Norwegian Radium Hospital was interrogated to identify individuals with a high risk of developing colorectal cancer (CRC) for whom no pathogenic variants in MMR genes had been found in routine diagnostic DNA sequencing. Forty-four cancer susceptibility genes were selected and analyzed by using our in-house designed TruSeq amplicon-based assay for targeted sequencing. RNA splicing- and protein-dedicated in silico analyses were performed for all variants of unknown significance (VUS). Variants predicted as likely to affect splicing were experimentally analyzed by resorting to minigene assays.

**Results:** We identified a patient who met the revised Bethesda guidelines and carried a likely pathogenic variant in *CHEK2* (c.470 T > C, p.I157T). In addition, 25 unique VUS were identified in 18 individuals, of which 2 exonic variants (*MAP3K1* c.764A > G and *NOTCH3* c.5854G > A) were analyzed in the minigene splicing assay and found not to have an effect on RNA splicing.

**Conclusions:** Among high-risk CRC patients that fulfill the AMS criteria or revised Bethesda guidelines, targeted gene sequencing identified likely pathogenic variant and VUS in other genes than the MMR genes (*CHEK2*, *NOTCH3* and *MAP3K1*). Our study suggests that the analysis of genes currently excluded from routine molecular diagnostic screens may confer cancer susceptibility.

**Keywords:** Lynch syndrome, Gene panel testing, *CHEK2*, RNA splicing mutations

## Background

Hereditarily represents a major cause of colorectal cancer (CRC) with at least 20% of the cases estimated to develop due to genetic factors and about 5% being linked to inherited variants in cancer-predisposing genes [1–4]. Currently, patients with CRC are referred to germline mismatch repair (MMR) testing based on the identification of high-risk phenotypic features (i.e. early age of onset, family history, clinical criteria), but beyond microsatellite instability (MSI) and MMR

immunohistochemistry (IHC) testing for Lynch syndrome (LS), no systematic approach to hereditary risk assessment exists [5].

LS is caused by a defective MMR system due to presence of germline defects in at least one of the MMR genes, *MLH1*, *MSH2*, *MSH6*, *PMS2* or to deletions of the 3' portion of the *EPCAM* gene [6]. LS is clinically classified according to the Amsterdam (AMS) criteria and/or the Bethesda guidelines, both relying in clinical information and family history. The Bethesda guidelines also take into account the MSI signature characteristic of MMR-deficient tumors [7–10]. LS patients have an increased lifetime risk of CRC (70–80%), endometrial cancer (50–60%), stomach cancer (13–19%), ovarian

\* Correspondence: [mev\\_dy@yahoo.com](mailto:mev_dy@yahoo.com)

<sup>1</sup>Department of Tumor Biology, Institute for Cancer Research, Oslo University Hospital, Oslo, Norway

Full list of author information is available at the end of the article



cancer (9–14%), cancers of the small intestine, the biliary tract and brain as well as carcinoma of the ureters and renal pelvis [11].

However, a high proportion of cases who meet the clinical criteria for LS (~60%) do not carry pathogenic variants in the MMR genes and have been reported as familial colorectal cancer type X (FCCTX) or Lynch-like syndrome (LLS) according to their MSI status [12–16]. The genetic mechanisms are undetermined in the majority of these families [14].

DNA sequencing (DNA-seq) studies using multi-gene panels have reported that as much as ~18% of patients diagnosed with CRC below the age of 50 years have pathogenic variants in several genes that are not traditionally associated with CRC (*ATM*, *CHEK2*, *BRCA1*, *BRCA2*, *CDKN2A* and *PALB2*) [5, 17]. Notably, there is a need to determine whether these variants contribute to hereditary CRC risk via the combination of low- and moderate-penetrance susceptibility alleles [5, 17, 18].

Given the high frequency and wide spectrum of pathogenic variants, it has been suggested that genetic counseling and testing with a multigene panel should be considered for all patients with early-onset CRC [17, 19–23]. Importantly, the identification of high-risk CRC patients is a major issue, because morbidity and mortality from CRC and extracolonic cancers in these patients and their relatives can be decreased by early screening and intensive surveillance [19, 24–26].

In an effort to discover inherited genetic variants that influence biological and clinical characteristics of familial CRC developed in unrelated high-risk patients, who previously tested negative for pathogenic variants in MMR genes, we examined 44 cancer associated genes using next generation sequencing (NGS), and applied minigene-based assay to analyze the impact of a subset of genetic variants on RNA splicing.

## Methods

### Study population

The Hereditary Cancer Biobank of the Norwegian Radium Hospital was used to identify unrelated high-risk CRC individuals from families that fulfilled the AMS criteria or the revised Bethesda guidelines [7–10, 27]. By the standard diagnostic clinical techniques, all study subjects were demonstrated not to carry pathogenic variants or large genomic rearrangements in MMR genes (*MLH1*, *MSH2*, *MSH6* or *PMS2*).

Ethical approval for the study was granted by the Norwegian Data Inspectorate and Ethical Review Board (ref 2015/2382). All examined patients signed an informed consent for their participation in the study.

### Targeted sequencing

Genomic DNA was isolated from peripheral blood samples and targeted sequencing was carried out using a TrueSeq amplicon based assay v.1.5 on a MiSeq apparatus, as previously described [28, 29]. The 44-gene panel used in this study includes genes associated with cancer predisposition as described in a prior study [28, 29].

### Sequencing data analysis

Paired-end sequence reads were aligned to the human reference genome (build GRCh37) using the BWA-mem algorithm (v.0.7.8-r55) [30]. The initial sequence alignments were converted to BAM format and subsequently sorted and indexed with SAMtools (v.1.1) [30]. Genotyping of single nucleotide variants (SNV) and short indels was performed by GATK's HaplotypeCaller. Filtering of raw genotype calls and assessment of callable regions/loci were done according to GATK's best practice procedures, as described more detailed previously [28].

Variants were annotated using ANNOVAR (version November 2015) [31] and were queried against a range of variant databases and protein resources, namely dbSNP (build 147) [32], 1000 Genome Project phase3 [33], Exome Aggregation Consortium (ExAC) (<http://exac.broadinstitute.org>, accessed August 2015) [34], Genome Aggregation Database (gnomAD) (<http://gnomad.broadinstitute.org>, accessed October 2017) [34], Norwegian Germline Variations Database (<http://norgene.no/vcf-miner/>, accessed October 2017), ClinVar (May 2016) [35], UniProt Knowledgebase (release March 2016) [36] and the Pfam protein domain database (v29, December 2015) [28, 37].

### Nomenclature and classification of genetic variants

The nomenclature guidelines of the Human Genome Variation Society (HGVS) were used to describe the detected genetic variants [38]. The recurrence of the identified variants was established by interrogating four databases (in their latest releases as of November 2016): the Leiden Open Variation Database (LOVD), the Universal Mutation Database (UMD), ClinVar and the Human Gene Mutation Database (HGMD). The variants were classified according to the 5-tier classification system into the following categories: class 5 (pathogenic), class 4 (likely pathogenic), class 3 (uncertain variants or variants of unknown significance, VUS), class 2 (likely not pathogenic) and class 1 (not pathogenic) [3].

### In silico analyses of VUS

Two types of bioinformatics methods were used to predict the impact of selected variants on RNA splicing. First, we used MaxEntScan (MES) and SSF-like (SSFL) to predict variant-induced alterations in 3' and 5' splice site strength, as described by Houdayer et al. 2012 [39],

except that here both algorithms were interrogated by using the integrated software tool Alamut Batch version 1.5, (Interactive Biosoftware, <http://www.interactive-bio-software.com>). For prediction of variant-induced impact on exonic splicing regulatory elements (ESR), we resorted to  $\Delta$ tESRseq- [40],  $\Delta$ HZeI- [41], and SPANR-based [42] as described by Soukarieh et al. [43]. Score differences ( $\Delta$ ) between variant and wild-type (WT) cases were taken as proxies for assessing the probability of a splicing defect. More precisely, we considered that a variant mapping at a splice site was susceptible of negatively impacting exon inclusion if  $\Delta$ MES $\geq$ 15% and  $\Delta$ SSFL $\geq$ 5% [39], whereas an exonic variant located outside the splice sites was considered as a probable inducer of exon skipping if negative  $\Delta$  scores (below the thresholds described below) were provided by all the 3 ESR-dedicated in silico tools. We chose the following thresholds:  $< -0.5$  for  $\Delta$ tESRseq-,  $< -10$  for  $\Delta$ HZeI-, and  $< -0.5$  for SPANR-based scores. In addition, we evaluated the possibility of variant-induced de novo splice sites by taking into consideration local changes in MES and SSFL scores. In this case, we considered that variants located outside the splice sites were susceptible of creating a competing splice site if local MES scores were equal to or greater than those of the corresponding reference splice site for the same exon.

In silico protein impact predictions of missense variants were performed with Align-GVGD (the VUS were predicted as deleterious when the values were from C35 or higher), SIFT, and MAPP using Alamut Batch version 1.4.4 (Interactive Biosoftware) and additionally with PolyPhen-2 and MutationTaster [44–48].

### Cell-based minigene splicing assays

In order to determine the impact of selected exonic variants on splicing, we performed functional assays based on the comparative analysis of the splicing pattern of WT and mutant reporter minigenes, as follows. First, genomic regions containing the exon of interest (internal exons only) and at least 150 nucleotides of the flanking introns were amplified by PCR [49] using patients' DNA as template and primers indicated in Additional file 1: Table S1. Next, representative minigenes were created by inserting the PCR-amplified fragments into a previously linearized pCAS2 vector [43]. All constructs were sequenced to ensure that no unwanted mutations had been introduced into the inserted fragments during PCR or cloning. Then, WT and mutant minigenes were transfected into HeLa cells grown in 12-well plates (at ~70% confluence) using the FuGENE 6 transfection reagent (Roche Applied Science). Twenty-four hours later, total RNA was extracted using the NucleoSpin RNA II kit (Macherey Nagel) and, the minigenes' transcripts were analyzed by semi-quantitative RT-PCR using the

OneStep RT-PCR kit (Qiagen), as previously described [43]. The sequences of the RT-PCR primers are shown in Additional file 1: Table S1. Later, RT-PCR products were separated by electrophoresis on 2.5% agarose gel containing EtBr and visualized by exposure to UV light under saturating conditions using the Gel Doc XR image acquisition system (Bio-Rad), followed by gel-purification and Sanger sequencing for proper identification of the minigenes' transcripts. Finally, splicing events were quantitated by performing equivalent fluorescent RT-PCR reactions followed by capillary electrophoresis on an automated sequencer (Applied Biosystems), and computational analysis by using the GeneMapper v5.0 software (Applied Biosystems).

## Results

### Clinical characteristics and family history

Upon querying the Hereditary Cancer Biobank of the Norwegian Radium Hospital for cases that fulfill the AMS and/or the revised Bethesda guidelines, we identified 34 unrelated potential high-risk CRC individuals who did not carry pathogenic variants in MMR genes. The median age at first CRC diagnosis was 51.5 years (range: 34–86 years).

Pedigree information showed that 13 (38%) families fulfilled the AMS I and/or II criteria and the revised Bethesda guidelines while 21 (62%) met the revised Bethesda guidelines only (Table 1). Fifteen (44%) patients had tumors with MSI and/or MMR IHC data available, of which 2 (13%) were MSI-high and/or MMR deficient. Clinical, family and tumor data information is detailed in Table 1.

### Germline findings

Given that the families that fulfilled the AMS criteria and/or the Bethesda guidelines did not carry pathogenic variants in the MMR genes, we hypothesized that other genes could be implicated in the genetic determinism of these phenotypes.

In order to pursue this hypothesis, we collected DNA samples from all probands and performed high-throughput sequencing of a panel of 44 cancer-associated genes. For the 34 samples, mean depth of coverage ranged from 127 to 507 with the fraction of target bases with coverage  $\geq$ 25 ranging from 80% to 93. The NGS results revealed that each individual carried an average of 26 SNV (between 19 and 33 per individual) in the set of 44 cancer susceptibility genes, most of which were common polymorphisms (allele frequency  $\geq$  1% in the general population) according to the ExAC database, and some being classified as benign or likely benign (class 1 or class 2) according to either ClinVar or the American College of Medical Genetics and Genomics (ACMG) guidelines [35, 50] (Table 2).

**Table 1** Summary of International Classification of Diseases, 9th Revision (ICD9), gender, age at diagnosis, clinical criteria and tumor molecular characteristics of the familial CRC families

Patient_ID	Gender	ICD9 diagnosis (age)	AMS criteria	Revised Bethesda	Tumor molecular characteristics
3222	F	CC (54), Hyperplastic polyp (55/61/62/63/65), BC (70)	0	Y	MMR IHC proficient
3308	F	CC (43), BC (51/52)	0	Y	MMR IHC proficient
3387	F	BC (40), OC (70), CC (80)	0	Y	MMR IHC proficient
3426	M	MM (39)	I & II	Y	na
4932	F	CC (34), EC (40), Hyperplastic polyp (43), BT (46)	I & II	Y	na
5324	F	M (52), CC (59), SMC (na), BC (72)	0	Y	na
6174	F	Hyperplastic polyp (63/67), BC (65)	I & II	Y	MMR IHC proficient
6977	F	TC (66)	0	Y	MMR IHC proficient
9876	F	M (45), BC (54)	0	Y	na
9998	F	Hyperplastic polyp (45), CC (45)	II	Y	MMR IHC proficient
10,675	F	BC (51), Hyperplastic polyp (59), TC (60)	II	Y	na
12,954	F	Hyperplastic polyp (69), ML (70)	II	Y	na
13,072	M	Hyperplastic polyp (63/64/65), CC (65/67)	0	Y	na
14,930	F	Hyperplastic polyp (86), CC (86)	0	Y	MMR IHC proficient
18,843	F	BC (44), CC (49), SMC (na)	0	Y	na
19,411	M	PC (70)	0	Y	MSH6 IHC deficient
19,673	F	BC (40/42)	II	Y	na
20,612	F	Hyperplastic polyp (59/65), EC (70)	0	Y	na
21,368	F	OC (62)	0	Y	na
22,295	F	Hyperplastic polyp (53), M (58)	0	Y	na
23,761	F	Hyperplastic polyp (40/42/44), BC (50)	0	Y	na
23,910	F	M (43), Hyperplastic polyp (49), BC (63), BT (63)	0	Y	na
24,140	F	CC (45/67), BC (56)	0	Y	na
24,447	F	BC (57/66), CC (66)	0	Y	MLH1/PMS2 IHC deficient and MSI
11,705	F	THC (53), KC (53/63)	II	Y	MMR IHC proficient
12,673	F	OC (23), SMC (36), RC (62)	II	Y	na
13,393	M	RC (48), CST (58)	I & II	Y	MMR IHC proficient and MSS
14,963	F	Hyperplastic polyp (69), BC (62)	0	Y	na
19609 <sup>a</sup>	F	CC (42), M (44), BC (57)	0	Y	MMR IHC proficient
22,953	F	BC (53)	II	Y	na
24,789	F	CC (43), RC (65), BC (72)	I & II	Y	MMR IHC proficient and MSS
25,167	M	CC (55)	I & II	Y	MMR IHC proficient
5597	M	Hyperplastic polyp (53/54), SC (55), KC (62)	0	Y	MMR IHC proficient
8913	F	Hyperplastic polyp (59), BC (61), TC (69)	0	Y	na

CRC colorectal, ICD9 diagnosis International Classification of Diseases, 9th Revision, CC colon cancer, BC breast cancer, AMS Amsterdam criteria, 0 not fulfill the AMS criteria, Y yes, MMR mismatch repair, IHC immunohistochemistry, MSI microsatellite instability, MSS microsatellite stable, na not available, OC ovary cancer, MM multiple myeloma, EC endometrial cancer, BT brain tumor, M melanoma, SMC other malignant of the skin, TC trachea, bronchus, lung cancer, ML malignant neoplasms of lymphoid, PC prostate cancer, THC thyroid cancer, KC kidney cancer, RC rectum cancer, CST malignant neoplasm of connective and soft tissue, SC stomach cancer

<sup>a</sup> Patient carrying *CHEK2* c.470 T > C, p.I157T

Importantly, we identified a likely pathogenic variant in a moderate-penetrance gene (*CHEK2* c.470 T > C, p.I157T) in a female patient diagnosed with colon cancer at 42 years, melanoma at 44 years and BC at 57 years with a proficient IHC MMR profile and fulfilling the revised Bethesda guidelines (Patient 19,609) (Table 1).

The *CHEK2* c.470 T > C has been classified as pathogenic according to the ACMG guidelines [51], and has a lower allele frequency ( $1.89 \times 10^{-3}$ ) in the Norwegian population, compared to the non-Finnish European population ( $5.4 \times 10^{-3}$ ) (<http://norgene.no/vcf-miner/> and gnomAD database, respectively) [34, 35, 50]. The variant

**Table 2** Characterization of germline variants found among Norwegian familial CRC individuals

Patient_ ID	VUS (Class 3)	Benign or Likely Benign variants (Class 1 or 2)	Polymorphisms	Total variants/patient
3222	<i>NOTCH3</i> NM_000435: c.5854G > A, p.V1952 M (rs115582213) <sup>b</sup> <i>POLE</i> NM_006231: c.3046G > A, p.V1016 M (rs147692158)	<i>ATM</i> NM_000051: c.5071A > C, p.S1691R (rs1800059) <i>BRCA1</i> NM_007300: c.5019G > A, p.M1673I (rs1799967) <i>PALB2</i> NM_024675: c.2993G > A, p.G998E (rs45551636) <i>PALB2</i> NM_024675: c.2014G > C, p.E672Q (rs45532440)	rs459552, rs659243, rs2240308, rs1799966, rs16942, rs16941, rs169547, rs4986764, rs1805107, rs506504, rs832582, rs5868032, rs1042821, rs3219484, rs1044009, rs152451, rs2228006, rs1805321, rs4796033, rs1042522, rs861539, rs13125836	28
3308	<i>NBN</i> NM_002485: c.1720 T > A, p.L574I (rs142334798) <i>POLE</i> NM_006231: c.4523G > A, p.R1508H (rs142508245)	<i>BARD1N</i> M_000465: c.1075_1095del, p.L359-P365del(LPECCSSP (rs28997575)	rs459552, rs659243, rs11528010, rs144848, rs169547, rs4986764, rs1805107, rs506504, rs702689, rs832582, rs5868032, rs1799977, rs1042821, rs3219484, rs1805794, rs1044009, rs2228006, rs5744934, rs5744751, rs4796033, rs1042522	24
3387	na	<i>CDKN2A</i> N M_000077: c.442G > A, p.A148T (rs3731249)	rs459552, rs659243, rs1801516, rs2240308 rs2070094, rs2229571, rs11528010, rs144848, rs169547, rs4986764, rs1805107, rs506504, rs1126497, rs702689, rs832582, rs5868032, rs3219489, rs1044009, rs2228006, rs1805321, rs5744934, rs1042522	23
3426	na	na	rs459552, rs659243, rs1801516, rs2240308, rs2070094, rs2229571, rs1048108, rs11528010, rs144848, rs169547, rs4986764, rs1805107, rs506504, rs1126497, rs702689, rs832582, rs5868032, rs1805794, rs1044009, rs2228006, rs1805321, rs1042522, rs13125836	23
4932	<i>NOTCH3</i> NM_000435: c.5854G > A, p.V1952 M (rs115582213) <sup>a</sup> <i>STK11</i> NM_00045:c.841C > A, p.P281T (rs377208033)	na	rs459552, rs659243, rs1801516, rs2240308, rs2229571, rs11528010, rs1799966, rs16942, rs16941, rs799917, rs144848, rs169547, rs1805107, rs506504, rs12642536, rs702689, rs832582, rs5868032, rs3219489, rs1805794, rs1044009, rs2228006, rs1805321, rs861539	26
5324	na	na		19

**Table 2** Characterization of germline variants found among Norwegian familial CRC individuals (Continued)

Patient_ ID	VUS (Class 3)	Benign or Likely Benign variants (Class 1 or 2)	Polymorphisms	Total variants/patient
			rs459552, rs659243, rs2240308, rs2070094, rs2229571, rs144848, rs169547, rs4986764, rs1805107, rs506504, rs1126497, rs12642536, rs832582, rs1799977, rs3219489, rs1044009, rs2228006, rs1805321, rs5744934	
6174	na	<i>PMS2</i> NM_000535: c.1454C > A, p.T485 K (rs1805323)	rs459552, rs659243, rs2240308, rs2070094, rs2229571, rs1048108, rs1799966, rs16942, rs16941, rs799917, rs4986850, rs169547, rs4986764, rs1805107, rs506504, rs1126497, rs12642536, rs702689, rs832582, rs5868032, rs1799977, rs3219484, rs1805794, rs1044009, rs2228006, rs1805321, rs1042522, rs861539	29
6977	na	<i>BARD1</i> NM_000465: c.1075_1095del, p.L359-P365delLPECSPP (rs28997575) <i>BARD1</i> NM_000465: c.1972C > T, p.R658C (rs3738888) <i>BRCA2</i> NM_000059: c.9976A > T, p.K3326X (rs11571833)	rs459552, rs659243, rs1801516, rs2240308, rs1799966, rs16942, rs16941, rs799917, rs144848, rs169547, rs4986764, rs1805107, rs506504, rs1126497, rs12642536, rs702689, rs832582, rs3219489, rs1044009, rs2228006, rs5744934, rs1042522, rs861539, rs28908468	27
9876	<i>PSMC3IP</i> NM_016556: c.136G > A, p.V46 M (rs757057684) <i>RAD51B</i> NM_133509: c.1063G > A, p.A355T (rs61758785)	<i>RAD51D</i> NM_002878: c.698A > G, p.E233G (rs28363284)	rs459552, rs659243, rs2240308, rs2070094, rs2229571, rs169547, rs4986764, rs1805107, rs506504, rs702689, rs832582, rs5868032, rs1799977, rs1042821, rs1805794, rs1044009, rs2228006, rs1726801	21
9998	na	<i>MSH6</i> NM_000179: c.2633 T > C, p.V878A (rs2020912)	rs459552, rs659243, rs1801516, rs2240308, rs2070094, rs2229571, rs1048108, rs144848, rs169547, rs4986764, rs1805107, rs506504, rs1126497, rs12642536, rs702689, rs832582, rs5868032, rs1799977, rs1044009, rs2228006, rs1805321, rs5744751, rs1042522	24
10,675	na	<i>PMS2</i> NM_000535: c.1531A > G, p.T511A (rs2228007)	rs459552, rs659243, rs1801516, rs2070094, rs2229571, rs11528010, rs1799966, rs16942, rs16941rs799917, rs144848, rs169547,	26



**Table 2** Characterization of germline variants found among Norwegian familial CRC individuals (*Continued*)

Patient_ ID	VUS (Class 3)	Benign or Likely Benign variants (Class 1 or 2)	Polymorphisms	Total variants/patient
12,954	<i>MUTYH</i> NM_012222: c.812G > A, p.R271Q (rs149866955) <i>RAD51C</i> NM_058216: c.790G > A, p.G264S (rs147241704)	<i>PMS2</i> NM_000535: c.1789A > T, p.T597S (rs1805318)	rs4986764, rs1805107, rs506504, rs1126497, rs702689, rs832582, rs5868032, rs1799977, rs1805794, rs1044009, rs2228006, rs10254120, rs1042522 rs459552, rs659243, rs2240308, rs2229571, rs144848, rs169547, rs4986764, rs1805107, rs506504, rs12642536, rs702689, rs832582, rs5868032, rs1044009, rs2228006, rs1805321, rs1726801, rs4796033, rs1042522	22
13,072	<i>BRCA1</i> NM_007300: c.4315C > T, p.L1439F (rs781260818)	<i>PMS2</i> NM_000535: c.1531A > G, p.T511A (rs2228007)	rs459552, rs659243, rs1801516, rs2240308, rs2070094, rs2229571, rs11528010, rs1799966, rs16942, rs16941, rs799917, rs144848, rs169547, rs4986764, rs1805107, rs506504, rs1126497, rs12642536, rs702689, rs832582, rs5868032, rs1799977, rs1805794, rs1044009, rs2228006, rs5744751, rs1042522	29
14,930	na	<i>PALB2</i> NM_024675: c.925A > G, p.I309V (rs3809683)	rs459552, rs659243, rs2240308, rs2070094, rs2229571, rs1048108, rs1799966, rs16942, rs16941, rs799917, rs4986850, rs169547, rs4986764, rs1805107, rs506504, rs1126497, rs12642536, rs702689, rs832582, rs5868032, rs1799977, rs3219489, rs1805794, rs1044009, rs152451, rs2228006, rs1805321, rs5744934, rs1042522	30
18,843	<i>MSH6</i> NM_000179.2: c.2195G > A, p.R732Q (rs749746725)	<i>BRCA1</i> NM_007300: c.3119G > A, p.S1040 N (rs4986852)	rs459552, rs659243, rs2240308, rs2070094, rs2229571, rs11528010, rs144848, rs169547, rs4986764, rs1805107, rs506504, rs1126497, rs12642536, rs832582, rs5868032, rs1799977, rs3219489, rs1044009, rs2228006, rs1805321, rs5744934, rs5744751, rs1042522, rs861539	26
19,411	na	<i>BARD1</i> NM_000465: c.1972C > T, p.R658C (rs3738888) <i>PALB2</i> NM_024675: c.1010 T >	rs459552, rs659243, rs2240308, rs2070094, rs2229571, rs1799966, rs16942, rs16941, rs799917, rs4986850,	28

**Table 2** Characterization of germline variants found among Norwegian familial CRC individuals (*Continued*)

Patient_ ID	VUS (Class 3)	Benign or Likely Benign variants (Class 1 or 2)	Polymorphisms	Total variants/patient
19,673	<i>AXIN2</i> NM_004655: c.344A > G, p.N115S (rs370257532)	<i>APC</i> NM_001127510: c.7504G > A, p.G2502S (rs2229995) <i>PMS2</i> NM_000535: c.1454C > A, p.T485 K (rs1805323)	C, p.L337S (rs45494092) rs144848, rs169547, rs1805107, rs506504, rs1126497, rs12642536, rs702689, rs832582, rs5868032, rs1805794, 1,044,009, rs2228006, rs1805321, rs5744751, rs1042522, rs861539 rs459552, rs659243, rs2240308, rs2070094, rs2229571, rs1799966, rs16942, rs16941, rs799917, rs144848, rs169547, rs4986764, rs1805107, rs506504, rs702689, rs832582, rs5868032, rs1042821, rs1805794, rs1044009, rs2228006, rs1805321, rs4796033, rs1042522, rs861539, rs3218536	29
20,612	na	<i>CDKN2A</i> NM_000077: c.442G > A, p.A148T (rs3731249)	rs459552, rs659243, rs2240308, rs2070094, rs2229571, rs11528010, rs1799966, rs16942, rs16941, rs799917, rs169547, rs4986764, rs1805107, rs506504, rs12642536, rs702689, rs5868032, rs1799977, rs1805794, rs1044009, rs2228006, rs1805321, rs10254120, rs5744934, rs4796033, rs1042522, rs861539	28
21,368	<i>MAP3K1</i> NM_005921: c.764A > G, p.N255S (rs56069227)	<i>ATM</i> NM_000051: c.2572 T > C, p.F858 L (rs1800056)	rs459552, rs659243, rs2240308, rs2229571, rs11528010, rs1799950, rs169547, rs1805107, rs506504, rs1126497, rs702689, rs832582, rs5868032, rs1799977, rs1044009, rs2228006, rs1805321, rs5744934, rs4796033, rs1042522	22
22,295	na	<i>BRCA1</i> NM_007300: c.5019G > A, p.M1673I (rs1799967) <i>BRIP1</i> NM_032043: c.890A > G, p.K297R (rs28997570)	rs459552, rs659243, rs2240308, rs2070094, rs2229571, rs1048108, rs1799966, rs16942, rs16941, rs799917, rs144848, rs169547, rs1805107, rs506504, rs12642536, rs702689, rs832582, rs5868032, rs1799977, rs1042821, rs3219489, rs1805794, rs1044009, rs2228006, rs1805321, rs5744934, rs1042522, rs861539	30
23,761	na	na	rs459552, rs659243, rs2240308, rs2070094, rs2229571, rs11528010, rs1799966, rs16942,	26



**Table 2** Characterization of germline variants found among Norwegian familial CRC individuals (Continued)

Patient_ ID	VUS (Class 3)	Benign or Likely Benign variants (Class 1 or 2)	Polymorphisms	Total variants/patient
			rs16941, rs799917, rs1799950, rs144848, rs169547, rs4986764, rs1805107, rs506504, rs1126497, rs702689, rs832582, rs5868032, rs1799977, rs1044009, rs2228006, rs1805321, rs1802683, rs1042522	
23,910	na	<i>BRCA2</i> NM_000059: c.6100C > T, p.R2034C (rs1799954) <i>MSH2</i> NM_000251: c.965G > A, p.G322D (rs4987188) <i>MSH6</i> NM_000179: c.2633 T > C, p.V878A (rs2020912) <i>PALB2</i> NM_024675: c.2794G > A, p.V932 M (rs45624036) <i>PMS2</i> NM_000535: c.1454C > A, p.T485 K (rs1805323) <i>BARD1</i> NM_000465: c.1670G > C, p.C557S (rs28997576)	rs459552, rs659243, rs2240308, rs2070094, rs2229571, rs11528010, rs169547, rs4986764, rs1805107, rs506504, rs1126497, rs702689, rs832582, rs5868032, rs1799977, rs1042821, rs1805794, rs2228006, rs10254120, rs1042522, rs13125836, rs3218536	28
24,140	na	<i>BRCA1</i> NM_007300: c.5019G > A, p.M1673I (rs1799967) <i>PMS2</i> NM_000535: c.1531A > G, p.T511A (rs2228007)	rs459552, rs659243, rs2240308, rs2070094, rs2229571, rs1048108, rs1799966, rs16942, rs16941, rs799917, rs144848, rs169547, rs1805107, rs506504, rs1126497, rs702689, rs832582, rs5868032, rs1799977, rs1042821, rs3219489, rs1805794, rs1044009, rs2228006, rs10254120, rs5744751, rs1042522	29
24,447	<i>CHEK2</i> NM_007194: c.74 T > C, p.V25A (rs587780188)	<i>NOTCH3</i> NM_000435: c.3399C > A, p.H1133Q (rs112197217)	rs459552, rs659243, rs1801516, rs2240308, rs2070094, rs2229571, rs1799966, rs16942, rs16941, rs799917, rs169547, rs1805107, rs506504, rs1126497, rs702689, rs832582, rs5868032, rs1805794, rs1044009, rs2228006, rs1805321, rs1042522	24
11,705	<i>ATM</i> NM_000051: c.4375G > A, p.G1459R (rs145667735) <i>MSH2</i> NM_000251: c.1284C > G, p.H428Q (rs776034412)	<i>MSH2</i> NM_000251: c.965G > A, p.G322D (rs4987188) <i>PMS2</i> NM_000535: c.1454C > A, p.T485 K (rs1805323)	rs459552, rs659243, rs1801516, rs2229571, rs11528010, rs169547, rs4986764, rs1805107, rs506504, rs1126497, rs702689, rs832582, rs5868032, rs1799977, rs3219489, rs1805794, rs1044009, rs2228006, rs1805321, rs5744751, rs1042522, rs861539, rs13125836	27
12,673	na	<i>ATM</i> NM_000051: c.2572 T > C,	rs459552, rs659243, rs2240308, rs2070094,	26

**Table 2** Characterization of germline variants found among Norwegian familial CRC individuals (Continued)

Patient_ ID	VUS (Class 3)	Benign or Likely Benign variants (Class 1 or 2)	Polymorphisms	Total variants/patient
		p.F858 L (rs1800056) <i>PMS2</i> NM_000535: c.1454C > A, p.T485 K (rs1805323)	rs2229571, rs1048108, rs1799950, rs169547, rs4986764, rs1805107, rs506504, rs1126497, rs12642536, rs702689, rs832582, rs5868032, rs1799977, rs1044009, rs2228006, rs1805321, rs1802683, rs4796033, rs1042522, rs861539	
13,393	<i>NBN</i> NM_002485.4: c.643C > T, p.R215W (rs34767364)	<i>BARD1</i> NM_000465: c.1972C > T, p.R658C (rs3738888) <i>BRIP1</i> NM_032043: c.577G > A, p.V193I (rs4988346) <i>PMS2</i> NM_000535: c.1531A > G, p.T511A (rs2228007) <i>ATM</i> NM_000051: c.4258C > T, p.L1420F (rs1800058) <i>NOTCH3</i> NM_000435.2: c.3058G > C, p.A1020P (rs35769976) <i>NOTCH3</i> NM_000435: c.3547G > A, p.V1183 M (rs10408676)	rs459552, rs659243, rs2240308, rs2070094, rs2229571, rs1799966, rs16942, rs16941, rs799917, rs169547, rs4986764, rs1805107, rs506504, rs702689, rs832582, rs5868032, rs1805794, rs1044009, rs2228006, rs1805321, rs1042522	28
14,963	<i>PALB2</i> NM_024675: c.232G > A, p.V78I (rs1515726085)	<i>PALB2</i> NM_024675: c.2590C > T, p.P864S (rs45568339) <i>STK11</i> NM_000455: c.1062C > G, p.F354 L (rs59912467)	rs659243, rs1801516, rs2240308, rs2070094, rs2229571, rs11528010, rs1799966, rs16942, rs16941, rs799917, rs1799950, rs144848, rs169547, rs4986764, rs1805107, rs506504, rs1126497, rs12642536, rs702689, rs5868032, rs1799977, rs3219489, rs1805794, rs1044009, rs2228006, rs1805321, rs5744751, rs1042522, rs861539, rs3218536	33
19,609 <sup>a</sup>	na	<i>BRCA2</i> NM_000059: c.4258G > T, p.D1420Y (rs28897727) <i>POLE</i> NM_006231: c.2083 T > A, p.F695I (rs5744799)	rs459552, rs659243, rs2240308, rs2229571, rs1799966, rs16942, rs16941, rs799917, rs144848, rs169547, rs4986764, rs1805107, rs506504, rs702689, rs832582, rs5868032, rs1799977, rs3219484, rs1805794, rs1044009, rs2228006, rs1805321, rs5744934, rs1042522, rs13125836	28
22,953	<i>NOTCH3</i> NM_000435: c.5208G > C, p.E1736D (rs200331646) <i>MSH2</i> NM_000251: c.128A > G, p.Y43C (rs17217723) <i>RAD51B</i> NM_133510: c.515 T > G, p.L172 W (rs34094401)	<i>BRCA2</i> NM_000059: c.2971A > G, p.N991D (rs1799944) <i>CDH1</i> NM_004360: c.1774G > A, p.A592T (rs35187787)	rs459552, rs659243, rs2240308, rs2229571, rs11528010, rs1799966, rs16942, rs16941, rs799917, rs169547, rs4986764, rs1805107, rs506504, rs12642536, rs702689, rs832582, rs5868032, rs1042821, rs3219489, rs1805794,	29

**Table 2** Characterization of germline variants found among Norwegian familial CRC individuals (*Continued*)

Patient_ ID	VUS (Class 3)	Benign or Likely Benign variants (Class 1 or 2)	Polymorphisms	Total variants/patient
24,789	APC NM_001127510: c.4334C > T, p.T1445I (rs760686348) PALB2 NM_024675: c.1250 C > A, p.S417Y (rs45510998)	BARD1 NM_000465: c.1972C > T, p.R658C (rs3738888) PALB2 NM_024675: c.2993G > A, p.G998E (rs45551636) PALB2 NM_024675: c.2014G > C, p.E672Q (rs45532440) POLE NM_006231: c.776G > A, p.R259H (rs61732929) NOTCH3 NM_000435: c.3399C > A, p.H1133Q (rs112197217)	rs1044009, rs2228006, rs1805321, rs1726801 rs459552, rs659243, rs2240308, rs2229571, rs1799966, rs16942, rs16941, rs799917, rs169547, rs4986764, rs1805107, rs506504, rs1126497, rs12642536, rs702689, rs832582, rs5868032, rs1799977, rs3219489, rs1044009, rs152451, rs2228006, rs1042522	30
25,167	NBN NM_002485.4: c.643C > T, p.R215W (rs34767364)	ATM NM_000051: c.2119 T > C, p.S707P (rs4986761)	rs459552, rs659243, rs2240308, rs2229571, rs11528010, rs169547, rs4986764, rs1805107, rs506504, rs12642536, rs702689, rs832582, rs5868032, rs1799977, rs1805794, rs2228006, rs1805321, rs1726801, rs4796033, rs1042522	22
5597	MAP3K1 NM_005921: c.2816C > G, p.S939C (rs45556841)	ATM NM_000051: c.5071A > C, p.S1691R (rs1800059)	rs459552, rs659243, rs1801516, rs2070094, rs2229571, rs11528010, rs1799966, rs16942, rs16941, rs799917, rs4986850, rs144848, rs169547, rs4986764, rs1805107, rs506504, rs1126497, rs12642536, rs702689, rs832582, rs5868032, rs1042821, rs1805794, rs1044009, rs2228006, rs1805321, rs1042522	29
8913	RAD51B NM_133509: c.1063G > A, p.A355T (rs61758785) EPCAM NM_002354: c.267G > C, p.Q89H (rs146480420)	PMS2 NM_000535: c.1454C > A, p.T485K (rs1805323)	rs459552, rs659243, rs2240308, rs2070094, rs2229571, rs1048108, rs11528010, rs1799966, rs16942, rs16941, rs799917, rs144848, rs169547, rs4986764, rs1805107, rs506504, rs1126497, rs702689, rs832582, rs5868032, rs1805794, rs1044009, rs2228006, rs5744751, rs1042522, rs861539, rs13125836	30

<sup>a</sup> Recently classified as Benign by ACMG Guidelines, 2015<sup>b</sup> Patient ID carrying CHEK2 c.470 T > C, p.I157T

is reported in ClinVar as “conflicting interpretations of pathogenicity, risk factor” (Variation ID: 5591). When the revised Bethesda guidelines were considered, the mutation detection rate was thus 4.8% (1/21).

Overall, 25 unique VUS were found in 18 out of the 34 patients (Table 2). The detected VUS were distributed among 17 different genes: *MAP3K1* (in 2 patients), *NBN* (in 3 patients), *NOTCH3* (in 3 patients), *RAD51B* (in 3 patients), *MSH2* (in 2 patients), *PALB2* (in 2 patients), *POLE* (in 2 patients) and the remaining were found in *APC*, *ATM*, *AXIN2*, *BRCA1*, *CHEK2*, *EPCAM*, *MSH6*, *MUTYH*, *RAD51C* and *STK11* (Table 2). The minor allele frequency (MAF) values of these variants were very low or no frequency data have been reported.

#### Protein and splicing-dedicated in silico analyses

The 25 unique VUS were analyzed by using five in silico prediction tools with different underlying algorithms to estimate the impact of the variants on the structure and function of the corresponding proteins.

Concordances between the 5 prediction tools were found for 2 out of the 25 VUS, suggesting a potentially damaging effect on protein level for the variants: *MUTYH* c.812G > A (p.R271Q) and *MSH2* c.128A > G (p.Y43C) (Table 3). In the other hand, 6 out of 25 VUS were consistently predicted as benign: *NBN* c.1720 T > A (p.L574I),

*BRCA1* c.4315C > T (p.L1439F), *MAP3K1* c.764A > G (p.N255S), *CHEK2* c.74 T > C (p.V25A), *PALB2* c.232G > A (p.V78I) and *APC* c.4334C > T (p.T1445I). Discrepancies were pronounced for the variants in the *POLE* ( $n = 2$ ), *STK11*, *MAP3K1*, *PSMC3IP*, *RAD51C*, *MSH6*, *AXIN2*, *MSH2*, *NBN*, *NOTCH3*, *RAD51B*, *PALB2* and *EPCAM* genes (Table 3).

Two out of the 25 VUS were bioinformatically predicted to affect RNA maturation by potentially modifying splicing signals (Table 3). More specifically, according to our in silico results, *NOTCH3* c.5854G > A (identified in Patients 3222 and 4932) was predicted to potentially induce exon 32 skipping by alteration of exonic splicing regulatory elements, whereas *MAP3K1* c.764A > G (detected in Patient 21,368) was predicted to introduce a deletion of the first 131 nucleotides of exon 3 (r.634\_764del) due to the creation of a putative new acceptor splice site. Skipping of *NOTCH3* exon 32 would produce a transcript with a frameshift deletion of 98 nucleotides (*NOTCH3* r.5816\_5913del), potentially leading to the production of a carboxy-terminally truncated NOTCH3 protein p.(Lys1940Glyfs\*14). The *MAP3K1* r.634\_764del transcript would be expected to be degraded by nonsense mediated decay and/or result in a very short MAP3K1 protein p.(Val212Leufs\*45). The *NOTCH3* c.5854G > A was identified in two patients (Patients 3222 and 4932) that fulfilled the revised

Bethesda guidelines and AMS criteria, respectively while the *MAP3K1* c.764A > G (Patient 21,368) in a patient which family fulfilled the revised Bethesda guidelines (Table 1).

#### Minigene splicing assays

Because patient RNA was not available, we decided to experimentally assess the impact of these 2 variants (*NOTCH3* c.5854G > A and *MAP3K1* c.764A > G) might have on RNA splicing by performing cell-based minigene splicing assays.

As shown in Fig. 1 we found that *NOTCH3* c.5854G > A and *MAP3K1* c.764A > G did not modify the splicing pattern of the minigenes' transcripts. These data thus disagree with the in silico predictions and suggest that either the exon 32 of *NOTCH3* and the exon 3 of *MAP3K1* are refractory to splicing mutations (the predictions thus being incorrect) or that the minigenes used in our study do not fully reproduce the splicing pattern of the mutant exons in *NOTCH3* and *MAP3K1* bona fide transcripts (the predictions being eventually correct). Complementary studies using RNA from *NOTCH3* c.5854G > A and *MAP3K1* c.764A > G carriers need to be performed to verify the pertinence of these results.

#### Discussion

The major unexpected finding in our Norwegian high-risk CRC cohort was the detection of a likely pathogenic variant in *CHEK2* (c.470 T > C, p.I157T), a moderate-penetrance gene not traditionally associated with CRC, in an individual with a LS-évocative personal/family history and a high number of Class 3 variants in BC- and CRC- associated genes. Interestingly, the *CHEK2* (c.470 T > C, p.I157T) has an allele frequency of  $1.89 \times 10^{-3}$  in the Norwegian population (<http://norgen.no/vcf-miner/>), and is reported in ClinVar as having conflicting interpretations of pathogenicity/being a risk factor (Variation ID: 5591). Importantly, there is no systematic classification for most of the genetic variants found by NGS, and, in more general terms, the impact of low- to moderate-penetrance pathogenic variants with respect to clinical management is not fully understood [52]. Co-segregation or case-control studies for further evaluation will be key in understanding whether such germline variant may have a modifying effect, since we do not yet have evidence-based guidelines for the majority of these genes.

On the other hand, *CHEK2* germline variants have been described to confer an elevated risk of BC (relative risk = 3.0) [53]. However, the presence of pathogenic variants in *CHEK2* is not frequently associated with cancer in high-risk BC families, prompting speculation that there may be several low-penetrance or moderate-penetrance BC risk genes segregating independently

**Table 3** In silico data obtained for the variants of unknown significance (VUS) identified in our study of familial CRC individuals

Selected variants (VAR)	Reference splice site-dedicated analyses										Cryptic splice site-dedicated analyses				ESR-dedicated analyses				Protein-dedicated analyses			
	Patient ID	Genomic position (GRCh37)	Gene	Exon	Nucleotide change (cNomen)	Predicted protein change (pNomen)	Nearest reference splice site	MES scores		SSFL scores		Potential local splice effect	Local MES scores	$\Delta$ ESRseq	$\Delta$ Hzei	$\Delta$ $\Psi$	AGVGD	SIFT	MAPP	PolyPhen-2	Mutation Taster	
								WT	Var vs WT	WT	Var vs WT											WT
3222 & 49 32	chr 191 5273 335 C>T	<b>NOTCH3</b>	<b>32</b>	<b>c585 4G&gt;A</b>	<b>p.V1 952M</b>	39 3'	11.5	11.5 0	89.2	89.2 0	-	-	-1.78 501	-11. 15	-0.89	Deleterious	C15	bad	probably damaging	disease causing		
3222		<i>POLE</i>	25	c304 6G>A	p.V10 16M	-15 5'	9.1	9.1 0	82.5	82.5 0	-	-	-0.74 478	1	-0.43	Deleterious	C0	bad	benign	disease causing		
3308	chr 8909 65597 A>T	<i>NBN</i>	11	c172 0T>A	p.L5 74I	-126 5'	8.0	8.0 0	82.4	82.4 0	-	-	1.97 427	33 38	0.18	Tolerated	C0	good	benign	polymorphism		
4932	chr 21333 219838 C>T	<i>POLE</i>	35	c452 3G>A	p.R1 508H	-29 5'	7.9	7.9 0	73.8	73.8 0	-	-	-0.600 279	-3.17	-0.08	Tolerated	C0	good	benign	disease causing		
5597	chr 912 21318 C>A	<i>STK11</i>	6	c841 C>A	p.P2 81T	-22 5'	6.0	6.0 0	79.9	79.9 0	-	-	-0.17 437	7.7	-0.72	Tolerated	C0	good	benign	disease causing		
9876	chr 17:40 729 320 C>T	<i>MAP3K14</i>	14	c281 6C>G	p.S9 39C	447 3'	12.0	12.0 0	100.0	100.0 0	-	-	-0.4 86 881	-16.1	0	Deleterious	C0	good	benign	polymorphism		
9876 & 89 13	chr 17:40 729 320 C>T	<i>PSM C3IP</i>	3	c136 G>A	p.V4 6M	1 3'	12.5	11.7 -6	87.5	83.6 -4	-	-	2.04 13	74 24	-0.3	Deleterious	-	-	possibly damaging	disease causing		
129 54	chr 1:45 7979 50_C>T	<i>RAD 51B</i>	11*	c106 3G>A	p.A3 55T	27 3'	11.8	11.8 0	80.2	80.2 0	-	-	-1.24 035	-50. 64	-	Deleterious	-	-	benign	polymorphism		
129 54	chr 1:45 7979 50_C>T	<i>MUTYH</i>	10	c812 G>A	p.R2 71Q	33 3'	9.5	9.5 0	86.8	86.8 0	-	-	-2.31 042	0.88	0.09	Deleterious	C35	bad	possibly damaging	disease causing		
130 72	chr 17:56 787304 G>A	<i>RAD 51C</i>	5	c790 G>A	p.G2 64S	-48 5'	8.6	8.6 0	75.4	75.4 0	New Acceptor Site?	0 2.5	-1.71 397	-593.1	-0.21	Tolerated	C0	good	benign	disease causing		
130 72	chr 17:41	<i>BRCA1</i>	12	c431 5C>T	p.L1 439F	-43 5'	6.6	6.6 0	85.2	85.2 0	-	-	-0.26 1141	8. 83	-2. 67	Tolerated	C0	good	benign	polymorphism		

**Table 3** In silico data obtained for the variants of unknown significance (VUS) identified in our study of familial CRC individuals (Continued)

Patient ID	Selected variants (VAR)	Reference splice site-dedicated analyses				Cryptic splice site-dedicated analyses				ESR-dedicated analyses		Protein-dedicated analyses													
		Genomic position (GRCh37)	Gene	Exon	Nucleotide change (cNomen)	Predicted protein change (pNomen)	Nearest reference splice site	Distance (nt)	Type (3' or 5' ss)	MES scores		SSFL scores		Potential local splice effect	Local MES scores	ΔHzei	ΔΨ	ΔESRseq	AGVGD	SIFT	MAPP	PolyPhen-2	Mutation Taster		
										WT	Var	WT	Var											WT	Var
234463																									
G>A																									
188	chr 2:48	<i>MSH6</i>	4**	c.219 5G>A	p.R7 32Q	-978	5'	8.9	8.9	0	81.6	81.6	0	-	-	-0.80	5417	-13.	0	C0	Tolerated	good	benign	disease causing	
027317																									
G>A																									
19	chr 17:6	<i>AXIN2</i>	2	c.344 A>G	p.N1 155	460	3'	11.1	11.1	0	93.3	93.3	0	New Acceptor Site?	11.1	7.2	-0.16	2849	-2.	-0.05	C0	Tolerated	good	benign	disease causing
3554395																									
T>C																									
21	chr 5:561	<i>MAP3K1</i>	3	c.764 A>G	p.N2 555	-71	5'	7.5	7.5	0	78.5	78.5	0	New Acceptor Site?	4.7	8.8	-1.	18	6.7	-0.04	C0	Tolerated	good	benign	polymorphism
55672																									
A>G																									
24	chr 22:2	<i>CHEK2</i>	2	c.74 T>C	p.V2 5A	80	3'	1.7	1.7	0	85.5	85.5	0	-	-	-0.1	66	36.	-0.07	C0	Tolerated	good	benign	polymorphism	
9130																									
636																									
A>G																									
11	chr 11:10	<i>ATM</i>	29	c.437 5G>A	p.G14 59R	-62	5'	8.9	8.9	0	87.5	87.5	0	-	-	0.5	52	-27.	-0.09	C15	Deleterious	bad	probably damaging	disease causing	
8160																									
467																									
G>A																									
133	chr 2:47	<i>MSH2</i>	8	c.128 4C>G	p.H4 28Q	8	3'	10.1	10.1	0	87.3	87.3	0	-	-	0.7	03	22.	0.05	C0	Tolerated	good	possibly damaging	disease causing	
672694																									
C>G																									
93	chr 8:90	<i>NBN</i>	6	c.64 3C>T	p.R2 15W	59	3'	6.2	6.2	0	86.8	86.8	0	-	-	-1.00	071	-16.	-0.09	C0	Tolerated	good	probably damaging	polymorphism	
88																									
167																									
G>A																									
149	chr 16:2	<i>PALB2</i>	4**	c.232 G>A	p.V78I 36D	21	3'	10.0	10.0	0	90.3	90.3	0	-	-	-0.7	83	-29.	0.8	C0	Tolerated	good	benign	polymorphism	
364																									
7635																									
C>T																									
229	chr 19:1	<i>MOT1</i>	29	c.5208 G>C	p.E17 36D	9	3'	7.8	7.8	0	84.2	84.2	0	-	-	0.0	50	-10.	0.22	C35	Deleterious	bad	benign	disease causing	
527																									
8214																									
C>G																									
1789																									

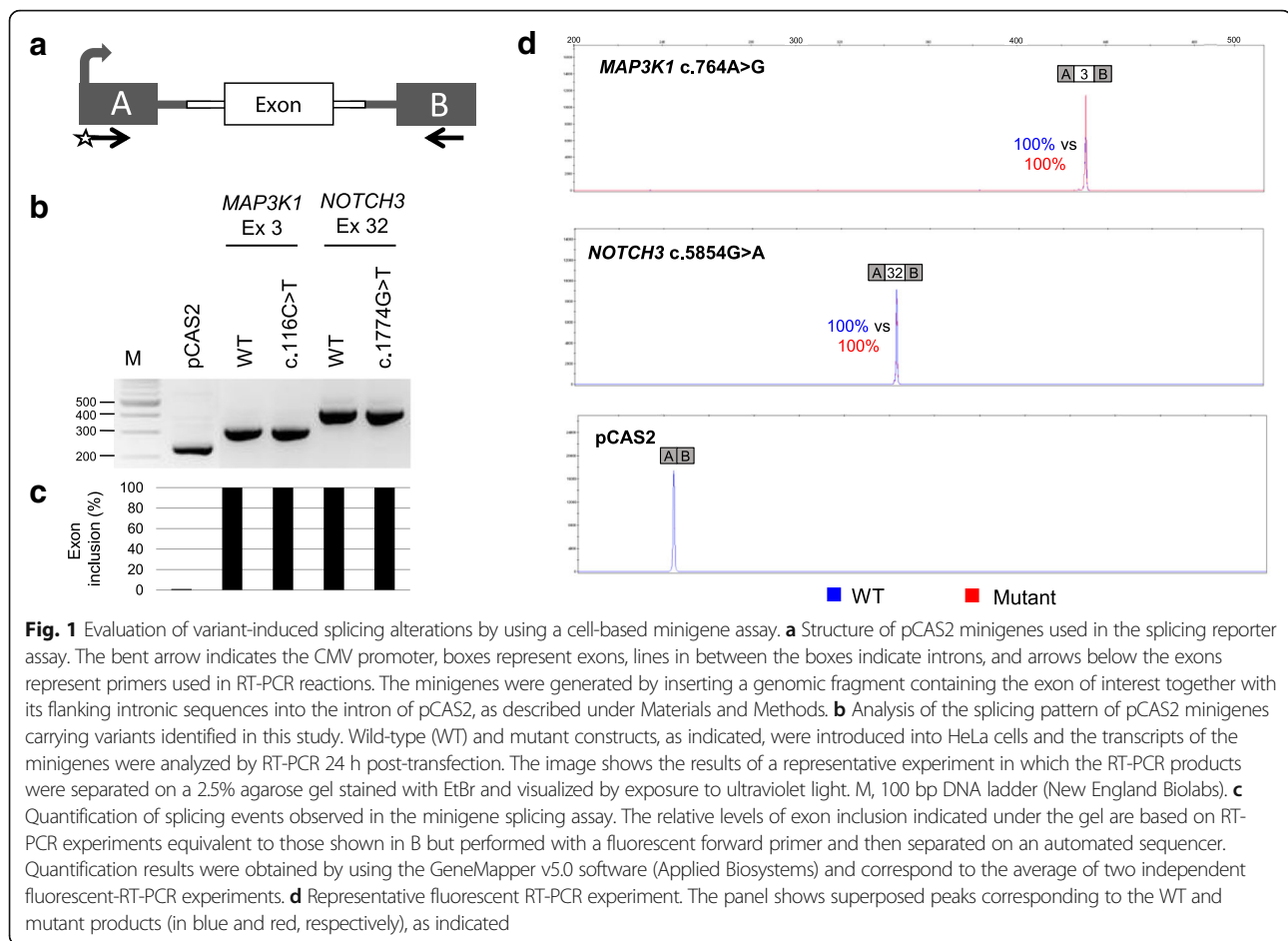


**Table 3** In silico data obtained for the variants of unknown significance (VUS) identified in our study of familial CRC individuals (Continued)

Selected variants (VAR)	Reference splice site-dedicated analyses				Cryptic splice site-dedicated analyses			ESR-dedicated analyses		Protein-dedicated analyses															
	Patient ID	Genomic position (GRCh37)	Gene	Exon	Nucleotide change (cNomen)	Predicted protein change (pNomen)	Nearest reference splice site	Distance (nt)	Type (3' or 5' ss)	MES scores	WT vs VAR	VAR vs WT	SSFL scores	Potential local splice effect	Local MES scores	WT	VAR	ΔESRseq	ΔHzei	ΔΨ	AGVD	SIFT	MAPP	PolyPhen-2	Mutation Taster
247	chr 247		<i>MSH2</i>	1*	c.128 A>G	p.Y43C	-84	5'	10.1	10.1	0	90.6	90.6	0	0	1.2	1.50425	4.7	2.11	C55	Deleterious	bad	probably damaging	disease causing	
630	chr 46		<i>RAD51B</i>	6	c.515 T>G	p.L172W	-58	5'	9.5	9.5	0	83.7	83.7	0	-	-	1.6	67.	1.31	C0	Deleterious	bad	possibly damaging	polymorphism	
458	chr 835		<i>APC</i>	15*	c.4334 C>T	p.T1445I	2376	3'	7.5	7.5	0	93.6	93.6	0	-	-	-1.61	-90.	-	C0	Tolerated	good	benign	polymorphism	
2648	chr 623		<i>PALB2</i>	4**	c.1250 C>A	p.S417Y	-435	5'	8.9	8.9	0	87.5	87.5	0	-	-	-0.5	-22.	0.05	C0	Deleterious	bad	probably damaging	disease causing	
89	chr 617		<i>EPCAM</i>	3	c.267 G>C	p.Q89H	83	3'	7.0	7.0	0	90.1	90.1	0	-	-	-0.40	-28.	0	C0	Tolerated	bad	possibly damaging	disease causing	
	chr 601																237	93							
	chr 029																								
	chr G>C																								

In order to predict the biological impact of the 25 VUS, RNA splicing- and protein-dedicated bioinformatics analyses were performed as described under Materials and Methods. The stars indicate exons that could not be tested in our minigene assay, either because of their terminal position (\*, 1st or last exon) or because of their large size (\*\*). Results shown in bold were considered as predictive of a potential variant-induced negative biological effect. MES, MaxEntScan; SSFL, Splice Site Finder-Like; nt, nucleotide; 3' or 5'ss, 3' splice site or 5' splice site; ESR, exonic splicing regulators; AGVD, align-GVGD (C0, C15, C25, C35, C45, C55, or C65 with C65 most likely to interfere with function and C0 least likely), SIFT, Sorting Intolerant From Tolerant (tolerated or deleterious), MAPP, Multivariate Analysis of Protein Polymorphism (good or bad), PolyPhen-2, Polymorphism Phenotyping v2 (benign, possibly damaging or probably damaging), MutationTaster (polymorphism or disease causing), CRC, colorectal cancer.

In order to predict the biological impact of the 25 VUS, RNA splicing- and protein-dedicated bioinformatics analyses were performed as described under Materials and Methods. The stars indicate exons that could not be tested in our minigene assay, either because of their terminal position (\*, 1st or last exon) or because of their large size (\*\*). Results shown in bold were considered as predictive of a potential variant-induced negative biological effect. MES, MaxEntScan; SSFL, Splice Site Finder-Like; nt, nucleotide; 3' or 5'ss, 3' splice site or 5' splice site; ESR, exonic splicing regulators; AGVD, align-GVGD (C0, C15, C25, C35, C45, C55, or C65 with C65 most likely to interfere with function and C0 least likely), SIFT, Sorting Intolerant From Tolerant (tolerated or deleterious), MAPP, Multivariate Analysis of Protein Polymorphism (good or bad), PolyPhen-2, Polymorphism Phenotyping v2 (benign, possibly damaging or probably damaging), MutationTaster (polymorphism or disease causing), CRC, colorectal cancer.



within these families [23, 54, 55]. Co-segregation analyses may add clues in our understanding whether this germline variant is implicated in CRC predisposition. Finally, we did not find pathogenic variants in *POLE* in our cohort, which is in contrast to what has been described in families with high burden of CRC adenomas and carcinomas in addition to extra-colonic cancers [56].

According to the Prospective LS Database (PLSDB), a total of 125 Norwegian families had a demonstrated pathogenic variant in either *MLH1* ( $n = 21$ ), *MSH2* ( $n = 52$ ), *MSH6* ( $n = 36$ ), or *PMS2* ( $n = 16$ ) [25]. On the other hand, a large portion of high-risk CRC families without pathogenic variant in MMR or *EPCAM* genes may be explained by a polygenic model involving a combination of multiple genomic risk factors, including the effect of either low-penetrance susceptibility alleles [57], high-penetrance genes which have not been tested, or the effect of environmental factors. In addition, emerging data suggest that CRC cases negative for pathogenic MMR variants may contain a significantly higher number of copy-neutral loss of heterozygosity (cnLOH) regions, some located within well-known oncogenes and tumor

suppressor genes, compared to cases of sporadic CRC [58]. These genomic variations, which were not investigated in this study, may provide an additional explanation for high-risk CRC phenotypes without MMR or *EPCAM* pathogenic variants.

Recent NGS studies described the presence of heterozygous pathogenic *BRCA1/2* or *APC* variants as well as biallelic *MUTYH* alterations in individuals with clinical features resembling those of LS [5, 22]. More precisely, those studies reported that 7% of patients with CRC carried pathogenic variants in non-LS genes, including 1.0% with *BRCA1/2* mutations, and nearly two thirds of probands with high-penetrance non-LS mutations lacked clinical histories suggestive of their respective syndromes [5].

From 34 high-risk CRC individuals, our NGS panel testing identified one patient that carried a pathogenic variant in a gene with reportedly moderate penetrance. Our finding is in line with the mutation frequency (6%) in non-LS cancer susceptibility genes for individuals undergoing LS genetic testing [21] and 4% of patients with BC tested negative for *BRCA1/2* genes [23]. Our results may have implications for an appropriate genetic

counseling and follow-up of the patients and family members.

Besides the likely pathogenic *CHEK2* variant, we identified a total of 25 variants in our cohort for which there were not so much data as to their clinical significance. We thus undertook bioinformatics analyses in an attempt to predict the biological impact of these Class 3 variants, both at the RNA and protein level, the ultimate goals being: (i) to discriminate pathogenic from non-pathogenic alterations in this set of variants and (ii) to further pinpoint the genetic determinants of high risk CRC in our cohort. On one hand, our RNA splicing-dedicated bioinformatics evaluation predicted that 2 out of the 25 VUS identified in this study (*NOTCH3* c.5854G>A, p.V1952 M and *MAP3K1* c.764A>G, p.N255S) could potentially affect RNA splicing. These two variants were then experimentally analyzed by performing minigene splicing assay. Our results revealed that neither variant altered the splicing pattern of the representative minigenes, suggesting that they do not affect the splicing of *NOTCH3* or *MAP3K1* transcripts. Additional experiments based on the analysis of RNA from carriers of these variants will be important to verify our minigene results. On the other hand, our protein-dedicated bioinformatics analysis yielded 8 consistent predictions (2 VUS predicted as deleterious and 6 as benign) and several conflicting results that were not explored further.

In this scenario, not only functional tests, but also co-segregation studies will be key to understanding whether the VUS detected in this work are non-pathogenic or otherwise have a causal or a modifying effect. Importantly, we do not yet have evidence-based guidelines for the majority of the genes carrying the VUS identified in this study and, in more general terms, the impact of low- to moderate-penetrance pathogenic variants with respect to clinical management is not fully understood. Most of these variants may in the future be reclassified as deleterious or benign, but in the meantime, they cannot be used to make clinical decisions [59]. Informed (re)-classification of VUS in cancer-associated genes may cater to more appropriate risk-management, and may provide significant clues for the identification of additional patients carrying such uncommon variants.

NGS panel testing may benefit patients with a personal or family history compatible with more than one recognized CRC inherited syndrome. The CRC risk management strategy for these individuals is not yet available and there is a need to identify new high-, moderate-, and low- penetrance gene variants that may affect the risk of CRC or LS-associated tumors in non-MMR pathogenic carriers. The identification of such gene variants in combination with family history may contribute

to more intensive surveillance and improved prevention [23].

## Conclusions

Our study provides information on genetic locus that might possibly be related to cancer susceptibility, demonstrating that genes presently not routinely tested may be important for capturing cancer predisposition in these patients. In addition, we stratified 25 VUS by the use of RNA splicing- and protein-dedicated in silico analyses. Further studies are necessary for making reliable estimates of cancer risk for the VUS found in this study and allowing appropriate genetic counseling for the patients and their relatives.

Surveillance for early cancer detection is essential to ensure optimal survival for patients afflicted with familial cancers. Our findings pinpoint the need of more studies to unravel the mechanisms underlying the development of CRC in high-risk patients and the identifying for new cancer predisposition genes.

## Additional file

**Additional file 1: Table S1.** Primers used in the pCAS2 minigene splicing assay. (DOCX 15 kb)

## Abbreviations

ACMG: American college of medical genetics and genomics; AMS: Amsterdam criteria; BC: Breast cancer; CRC: Colorectal cancer; IHC: Immunohistochemistry; LS: Lynch syndrome; MAF: Minor allele frequency; MMR: Mismatch repair genes; MSI: Microsatellite instability; NGS: Next generation sequencing; PLSDB: Prospective lynch syndrome database; SNV: Single nucleotide variant; VUS: Variants of unclassified significance

## Acknowledgements

We thank the families for their participation and contribution to this study.

## Funding

This work was supported by the Radium Hospital Foundation (Oslo, Norway), Helse Sør-Øst (Norway), the French Association Recherche contre le Cancer (ARC), the Groupement des Entreprises Françaises dans la Lutte contre le Cancer (Gefluc), the Association Nationale de la Recherche et de la Technologie (ANRT, CIFRE PhD fellowship to H.T.) and by the OpenHealth Institute.

## Availability of data and materials

All data generated or analyzed during this study are included in the manuscript.

## Authors' contributions

All authors have taken part in the different steps of the study: MDV, PM and EH designed the study, AM, HT performed in silico splicing predictions and the minigene assays, POE performed validation experiments, MM, AN and EHF performed in silico protein predictions, SN, DV performed the sequence analysis. MDV drafted the manuscript and all have read, revised and approved the manuscript.

## Ethics approval and consent to participate

Ethical approval for the study was granted by the Norwegian Data Inspectorate and Ethical Review Board (ref 2015/2382). All examined patients signed an informed consent for their participation in the study.

**Consent for publication**

Not Applicable.

**Competing interests**

The authors declare that they have no competing interests

**Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Author details**

<sup>1</sup>Department of Tumor Biology, Institute for Cancer Research, Oslo University Hospital, Oslo, Norway. <sup>2</sup>Inserm-U1245, UNIROUEN, Normandie Univ, Normandy Centre for Genomic and Personalized Medicine, Rouen, France. <sup>3</sup>Interactive Biosoftware, Rouen, France. <sup>4</sup>Medizinische Klinik und Poliklinik IV, Campus Innenstadt, Klinikum der Universität München, Ziemssenstr. 1, Munich, Germany. <sup>5</sup>MGZ—Medizinisches Genetisches Zentrum, Munich, Germany. <sup>6</sup>Department of Human Medicine, Universität Witten, Herdecke, Germany. <sup>7</sup>Department of Medical Genetics, Oslo University Hospital, Oslo, Norway. <sup>8</sup>Department of Informatics, University of Oslo, Oslo, Norway. <sup>9</sup>Institute of Cancer Genetics and Informatics, Oslo University Hospital, Oslo, Norway.

Received: 30 June 2017 Accepted: 24 January 2018

Published online: 20 February 2018

**References**

- Llor X, Pons E, Xicola RM, Castells A, Alenda C, Pinol V, Andreu M, Castellvi-Bel S, Paya A, Jover R, et al. Differential features of colorectal cancers fulfilling Amsterdam criteria without involvement of the mutator pathway. *Clin Cancer Res*. 2005;11(20):7304–10.
- Balmana J, Castells A, Cervantes A, Group EGW. Familial colorectal cancer risk: ESMO Clinical Practice Guidelines. *Ann Oncol*. 2010;21(Suppl 5):v78–81.
- Buchanan DD, Tan YY, Walsh MD, Clendenning M, Metcalf AM, Ferguson K, Arnold ST, Thompson BA, Lose FA, Parsons MT, et al. Tumor mismatch repair Immunohistochemistry and DNA MLH1 Methylation testing of patients with endometrial cancer diagnosed at age younger than 60 years optimizes triage for population-level Germline mismatch repair gene mutation testing. *J Clin Oncol*. 2014;32(2):90.
- Patel SG, Ahnen DJ. Familial colon cancer syndromes: an update of a rapidly evolving field. *Curr Gastroenterol Rep*. 2012;14(5):428–38.
- Yurgelun MB, Kulke MH, Fuchs CS, Allen BA, Uno H, Hornick JL, Ukaegbu CI, Brais LK, McNamara PG, Mayer RJ, et al. Cancer susceptibility gene mutations in individuals with colorectal cancer. *J Clin Oncol*. 2017;35(10):JCO2016710012.
- Kuiper RP, Vissers LE, Venkatachalam R, Bodmer D, Hoenselaar E, Goossens M, Haufe A, Kamping E, Niessen RC, Hogervorst FB, et al. Recurrence and variability of germline EPCAM deletions in lynch syndrome. *Hum Mutat*. 2011;32(4):407–14.
- Vasen HF, Mecklin JP, Khan PM, Lynch HT. The international collaborative group on hereditary non-polyposis colorectal cancer (ICG-HNPCC). *Dis Colon Rectum*. 1991;34(5):424–5.
- Moller P, Evans G, Haites N, Vasen H, Reis MM, Anderson E, Apold J, Hodgson S, Eccles D, Olsson H, et al. Guidelines for follow-up of women at high risk for inherited breast cancer: consensus statement from the biomed 2 demonstration Programme on inherited breast cancer. *Dis Markers*. 1999; 15(1–3):207–11.
- RodriguezBigas MA, Boland CR, Hamilton SR, Henson DE, Jass JR, Khan PM, Lynch H, Perucho M, Smyrk T, Sobin L, et al. A National Cancer Institute workshop on hereditary nonpolyposis colorectal cancer syndrome: meeting highlights and Bethesda guidelines. *J Natl Cancer I*. 1997;89(23):1758–62.
- Umar A, Boland CR, Terdiman JP, Syngal S, de la Chapelle A, Ruschoff J, Fishel R, Lindor NM, Burgart LJ, Hamelin R, et al. Revised Bethesda guidelines for hereditary nonpolyposis colorectal cancer (lynch syndrome) and microsatellite instability. *J Natl Cancer I*. 2004;96(4):261–8.
- Kobayashi H, Ohno S, Sasaki Y, Matsuura M. Hereditary breast and ovarian cancer susceptibility genes (review). *Oncol Rep*. 2013;30(3):1019–29.
- Lindor NM. Familial colorectal cancer type X: the other half of hereditary nonpolyposis colon cancer syndrome. *Surg Oncol Clin N Am*. 2009;18(4):637.
- Rodriguez-Soler M, Perez-Carbonell L, Guarinos C, Zapater P, Castillejo A, Barbera VM, Juarez M, Bessa X, Xicola RM, Clofent J, et al. Risk of cancer in cases of suspected lynch syndrome without Germline mutation. *Gastroenterology*. 2013;144(5):926.
- Dominguez-Valentin M, Therkildsen C, Da Silva S, Nilbert M. Familial colorectal cancer type X: genetic profiles and phenotypic features. *Mod Pathol*. 2015;28(1):30–6.
- da Silva FC, Ferreira JRD, Torrezan GT, Figueiredo MCP, Santos EMM, Nakagawa WT, Brianese RC, de Oliveira LP, Begnani MD, Aguiar S, et al. Clinical and molecular characterization of Brazilian patients suspected to have lynch syndrome. *PLoS One*. 2015;10(10):e0139753.
- Amendola LM, Jarvik GP, Leo MC, McLaughlin HM, Akkari Y, Amaral MD, Berg JS, Biswas S, Bowling KM, Conlin LK, et al. Performance of ACMG-AMP variant-interpretation guidelines among nine Laboratories in the Clinical Sequencing Exploratory Research Consortium (vol 98, pg 1067, 2016). *Am J Hum Genet*. 2016;99(1):247.
- Pearlman R, Frankel WL, Swanson B, Zhao W, Yilmaz A, Miller K, Bacher J, Bigley C, Nelsen L, Goodfellow PJ, et al. Prevalence and Spectrum of Germline cancer susceptibility gene mutations among patients with early-onset colorectal cancer. *JAMA Oncol*. 2016;3(4):464–71. <https://doi.org/10.1001/jamaoncol.2016.5194>.
- Picelli S, Lorenzo Bermejo J, Chang-Claude J, Hoffmeister M, Fernandez-Rozadilla C, Carracedo A, Castells A, Castellvi-Bel S, Members of EC-GOGotSGA, Naccarati A, et al. Meta-analysis of mismatch repair polymorphisms within the cogent consortium for colorectal cancer susceptibility. *PLoS One*. 2013;8(9):e72091.
- Ballester V, Boardman L. Next generation multigene panel testing: the next step for identification of hereditary colorectal cancer syndromes? *Gastroenterology*. 2015;149(3):526–8.
- Kurian AW, Hare EE, Mills MA, Kingham KE, McPherson L, Whittemore AS, McGuire V, Ladabaum U, Kobayashi Y, Lincoln SE, et al. Clinical evaluation of a multiple-gene sequencing panel for hereditary cancer risk assessment. *J Clin Oncol*. 2014;32(19):2001–9.
- Yurgelun MB, Masciari S, Joshi VA, Mercado RC, Lindor NM, Gallinger S, Hopper JL, Jenkins MA, Buchanan DD, Newcomb PA, et al. Germline TP53 mutations in patients with early-onset colorectal cancer in the colon cancer family registry. *JAMA Oncol*. 2015;1(2):214–21.
- Yurgelun MB, Allen B, Kaldate RR, Bowles KR, Judkins T, Kaushik P, Roa BB, Wenstrup RJ, Hartman AR, Syngal S. Identification of a variety of mutations in cancer predisposition genes in patients with suspected lynch syndrome. *Gastroenterology*. 2015;149(3):604–13. e620
- Tung N, Lin NU, Kidd J, Allen BA, Singh N, Wenstrup RJ, Hartman AR, Winer EP, Garber JE. Frequency of Germline mutations in 25 cancer susceptibility genes in a sequential series of patients with breast cancer. *J Clin Oncol*. 2016;34(13):1460–8.
- Moller P, Seppala T, Bernstein I, Holinski-Feder E, Sala P, Evans DG, Lindblom A, Macrae F, Blanco I, Sijmons R, et al. Cancer incidence and survival in lynch syndrome patients receiving colonoscopic and gynaecological surveillance: first report from the prospective lynch syndrome database. *Gut*. 2015;66(3):464–72.
- Moller P, Seppala T, Bernstein I, Holinski-Feder E, Sala P, Evans DG, Lindblom A, Macrae F, Blanco I, Sijmons R, et al. Incidence of and survival after subsequent cancers in carriers of pathogenic MMR variants with previous cancer: a report from the prospective lynch syndrome database. *Gut*. 2017; 66(9):1657–64. <https://doi.org/10.1136/gutjnl-2016-311403>.
- Moller P, Seppala TT, Bernstein I, Holinski-Feder E, Sala P, Gareth Evans D, Lindblom A, Macrae F, Blanco I, Sijmons RH, et al. Cancer risk and survival in path\_MMR carriers by gene and gender up to 75 years of age: a report from the prospective lynch syndrome database. *Gut*. 2017; <https://doi.org/10.1136/gutjnl-2017-314057>. [Epub ahead of print]
- Lynch HT, Riley BD, Weissman SM, Coroneo SM, Kinarsky Y, Lynch JF, Shaw TG, Rubinstein WS. Hereditary nonpolyposis colorectal carcinoma (HNPCC) and HNPCC-like families: problems in diagnosis, surveillance, and management. *Cancer Am Cancer Soc*. 2004;100(1):53–64.
- Dominguez-Valentin M, Nakken S, Tubeuf H, Vodak D, Ekstrom PO, Nissen AM, Morak M, Holinski-Feder E, Martins A, Moller P, et al. Potentially pathogenic germline CHEK2 c.319+2T>a among multiple early-onset cancer families. *Familial Cancer*. 2018;17(1):141–53. <https://doi.org/10.1007/s10689-017-0011-0>.
- Dominguez-Valentin M, Evans DGR, Nakken S, Tubeuf H, Vodak D, Ekstrom PO, Nissen AM, Morak M, Holinski-Feder E, Martins A, Møller P, Hovig E. Genetic variants of prospectively demonstrated phenocopies in BRCA1/2



- kindreds. *Hered Cancer Clin Pract.* 2018;15(16):4. <https://doi.org/10.1186/s13053-018-0086-0>.
30. Li L, Chen HC, Liu LX. Sequence alignment algorithm in similarity measurement. *2009 International Forum on Information Technology and Applications.* PRO. 2009;1:453–6.
  31. Auclair J, Vaissiere T, Desseigne F, Lasset C, Bonadona V, Giraud S, Saurin JC, Joly MO, Leroux D, Faivre L, et al. Intensity-dependent constitutional MLH1 promoter methylation leads to early onset of colorectal cancer by affecting both alleles. *Genes Chromosomes Cancer.* 2011;50(3):178–85.
  32. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K: dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* 2001;29(1):308–11.
  33. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, et al. A global reference for human genetic variation. *Nature.* 2015;526(7571):68–74.
  34. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature.* 2016;536(7616):285–91.
  35. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, Maglott DR. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res.* 2014;42(D1):D980–5.
  36. Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang HZ, Lopez R, Magrane M, et al. UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* 2004;32:D115–9.
  37. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heeger A, Hetherington K, Holm L, Mistry J, et al. Pfam: the protein families database. *Nucleic Acids Res.* 2014;42(D1):D222–30.
  38. den Dunnen JT, Antonarakis SE. Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. *Hum Mutat.* 2000;15(1):7–12.
  39. Antoniou AC, Kuchenbaecker KB, Soucy P, Beesley J, Chen XQ, McGuffog L, Lee A, Barrowdale D, Healey S, Sinilnikova OM, et al. Common variants at 12p11, 12q24, 9p21, 9q31.2 and in ZNF365 are associated with breast cancer risk for BRCA1 and/or BRCA2 mutation carriers. *Breast Cancer Res.* 2012;14(1):1–18.
  40. Di Giacomo D, Gaildrat P, Abuli A, Abdat J, Frebourg T, Tosi M, Martins A. Functional analysis of a large set of BRCA2 exon 7 variants highlights the predictive value of hexamer scores in detecting alterations of exonic splicing regulatory elements. *Hum Mutat.* 2013;34(11):1547–57.
  41. Erkelenz S, Hillebrand F, Widera M, Theiss S, Fayyaz A, Degrandi D, Pfeffer K, Schaal H. Balanced splicing at the tat-specific HIV-1 3' ss A3 is critical for HIV-1 replication. *Retrovirology.* 2015;12:29.
  42. Susswein LR, Marshall ML, Nusbaum R, Vogel Postula KJ, Weissman SM, Yackowski L, Vaccari EM, Bissonnette J, Booker JK, Cremona ML, et al. Pathogenic and likely pathogenic variant prevalence among the first 10,000 patients referred for next-generation cancer panel testing. *Genet Med.* 2016;18(8):823–32.
  43. Soukariéh O, Gaildrat P, Hamieh M, Drouet A, Baert-Desurmont S, Frebourg T, Tosi M, Martins A. Exonic splicing mutations are more prevalent than currently estimated and can be predicted by using in Silico tools. *PLoS Genet.* 2016;12(1):1–26.
  44. 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(4):295–305.
  45. 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(7):1073–82.
  46. Stone EA, Sidow A. Physicochemical constraint violation by missense substitutions mediates impairment of protein function and disease severity. *Genome Res.* 2005;15(7):978–86.
  47. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. *Nat Methods.* 2010;7(4):248–9.
  48. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods.* 2014;11(4):361–2.
  49. Gaildrat P, Killian A, Martins A, Tournier I, Frebourg T, Tosi M. Use of splicing reporter minigene assay to evaluate the effect on splicing of unclassified genetic variants. *Methods Mol Biol.* 2010;653:249–57.
  50. Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, Gu B, Hart J, Hoffman D, Hoover J, et al. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* 2016;44(D1):D862–8.
  51. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* 2015;17(5):405–24.
  52. Yadav S, Fulbright J, Dreyfuss H, Reeves A, Campian S, Thomas V, Zakalik D. Outcomes of retesting BRCA-negative patients using multigene panels. *J Clin Oncol.* 2015;33(28):319–28.
  53. Easton DF, Pharoah PD, Antoniou AC, Tischkowitz M, Tavtigian SV, Nathanson KL, Devilee P, Meindl A, Couch FJ, Southey M, et al. Gene-panel sequencing and the prediction of breast-cancer risk. *N Engl J Med.* 2015;372(23):2243–57.
  54. Meijers-Heijboer H, van den Ouweland A, Klijn J, Wasielewski M, de Snoo A, Oldenburg R, Hollestelle A, Houben M, Crepin E, van Veghel-Plandsoen M, et al. Low-penetrance susceptibility to breast cancer due to CHEK2\*(1100delC) in noncarriers of BRCA1 or BRCA2 mutations. *Nat Genet.* 2002;31(1):55–9.
  55. Consortium CBCC-C. CHEK2\*1100delC and susceptibility to breast cancer: a collaborative analysis involving 10,860 breast cancer cases and 9,065 controls from 10 studies. *Am J Hum Genet.* 2004;74(6):1175–82.
  56. Hansen MF, Johansen J, Bjornevoll I, Sylvander AE, Steinsbekk KS, Saetrom P, Sandvik AK, Drablos F, Sjrursen W. A novel POLE mutation associated with cancers of colon, pancreas, ovaries and small intestine. *Familial Cancer.* 2015;14(3):437–48.
  57. Obermeier K, Sachsenweger J, Friedl TW, Pospiech H, Winqvist R, Wiesmuller L. Heterozygous PALB2 c.1592delT mutation channels DNA double-strand break repair into error-prone pathways in breast cancer patients. *Oncogene.* 2016;35(29):3796–806. <https://doi.org/10.1038/ncr.2015.448>.
  58. Villavic RA, Basso TR, Canto LM, Pinheiro M, Santiago KM, Giacomazzi J, de Paula CA, Carraro DM, Ashton-Prolla P, Achatz MI, et al. Rare germline alterations in cancer-related genes associated with the risk of multiple primary tumor development. *J Mol Med (Berl).* 2017;
  59. Pinto P, Paulo P, Santos C, Rocha P, Pinto C, Veiga I, Pinheiro M, Peixoto A, Teixeira MR. Implementation of next-generation sequencing for molecular diagnosis of hereditary breast and ovarian cancer highlights its genetic heterogeneity. *Breast Cancer Res Treat.* 2016;159(2):245–56.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

