



Clinical implications of genetic testing in familial intermediate and late-onset colorectal cancer

Malene Djursby¹ · Thomas van Overeem Hansen^{1,9} · Karin A. W. Wadt^{1,9} · Majbritt Busk Madsen² · Lukas Adrian Berchtold² · Charlotte Kvist Lautrup^{3,5} · Sara Markholt³ · Uffe Birk Jensen³ · Lotte Nylandsted Krogh⁴ · Malene Lundsgaard⁵ · Anne Marie Gerdes^{1,9} · Mef Nilbert^{6,7} · Christina Therkildsen^{6,8}

Received: 6 May 2022 / Accepted: 23 June 2022 / Published online: 29 July 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

The genetic background of familial, late-onset colorectal cancer (CRC) (i.e., onset > age 50 years) has not been studied as thoroughly as other subgroups of familial CRC, and the proportion of families with a germline genetic predisposition to CRC remains to be defined. To define the contribution of known or suggested CRC predisposition genes to familial late-onset CRC, we analyzed 32 well-established or candidate CRC predisposition genes in 75 families with late-onset CRC. We identified pathogenic or likely pathogenic variants in five patients in *MSH6* ($n = 1$), *MUTYH* (monoallelic; $n = 2$) and *NTHL1* (monoallelic; $n = 2$). In addition, we identified a number of variants of unknown significance in particular in the lower penetrant Lynch syndrome-associated mismatch repair (MMR) gene *MSH6* ($n = 6$). In conclusion, screening using a comprehensive cancer gene panel in families with accumulation of late-onset CRC appears not to have a significant clinical value due to the low level of high-risk pathogenic variants detected. Our data suggest that only patients with abnormal MMR immunohistochemistry (IHC) or microsatellite instability (MSI) analyses, suggestive of Lynch syndrome, or a family history indicating another cancer predisposition syndrome should be prioritized for such genetic evaluations. Variants in *MSH6* and *MUTYH* have previously been proposed to be involved in digenic or oligogenic hereditary predisposition to CRC. Accumulation of variants in *MSH6* and monoallelic, pathogenic variants in *MUTYH* in our study indicates that digenic or oligogenic inheritance might be involved in late-onset CRC and warrants further studies of complex types of inheritance.

Introduction

Heredity is estimated to contribute to ~20% of colorectal cancers (CRC) and covers a complex genetic landscape (Valle 2017) but only ~5% of CRC cases have been explained by a high-risk pathogenic germline variant. The clinical Amsterdam I criteria (Vasen et al. 1991) were

established to identify families with a high risk of germline variants resulting in autosomal dominant hereditary CRC. However, the later detection of the mismatch repair (MMR) genes *MLH1*, *MSH2/EPCAM*, *MSH6* and *PMS2* has dichotomized the Amsterdam I positive families into two groups: those with MMR-deficient tumors linked to Lynch Syndrome (LS) and those with MMR-proficient

✉ Malene Djursby
malene.djursby@regionh.dk

¹ Department of Clinical Genetics, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

² Center for Genomic Medicine, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

³ Department of Clinical Genetics, Aarhus University Hospital, Aarhus, Denmark

⁴ Department of Clinical Genetics, Odense University Hospital, Odense, Denmark

⁵ Department of Clinical Genetics, Aalborg University Hospital, Aalborg, Denmark

⁶ Clinical Research Centre, Copenhagen University Hospital, Amager and Hvidovre Hospital, Copenhagen, Denmark

⁷ Institute of Clinical Sciences, Division of Oncology and Pathology, Lund University, Lund, Sweden

⁸ HNPCC Register, Gastro Unit, Copenhagen University Hospital, Amager and Hvidovre Hospital, Copenhagen, Denmark

⁹ Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

Familial Colorectal Cancer Type X (FCCTX) syndrome with unknown genetic causes (Carethers and Stoffel 2015). Over the past decade, pathogenic variants accounting for only a small fraction of the FCCTX families have been identified in genes such as *POLE*, *POLD1*, and *RPS20* (Palles et al. 2013; Nieminen et al. 2014). Well-established cancer genes such as *BRCA1*, *BRCA2*, *TP53* and *ATM* have also been proposed in CRC predisposition among individuals with suspicion of a hereditary predisposition to CRC (Yurgelun et al. 2015b; Espenschied et al. 2017; Rosenthal et al. 2017; AlDubayan et al. 2018; LaDuca et al. 2019). However, the vast majority of FCCTX families still remains genetically undiagnosed (Palles et al. 2013; Nieminen et al. 2014; Garre et al. 2015; Yurgelun et al. 2015a, b; Broderick et al. 2017; AlDubayan et al. 2018; Dominguez-Valentin et al. 2018; Thompson et al. 2020).

In Denmark, data on families with suspected hereditary CRC have been systematically collected in the national Hereditary Nonpolyposis Colorectal Cancer (HNPCC) registry over the past 30 years resulting in a database with a unique dataset that covers the entire country. A subgroup of particular interest is families with a family history of CRC indicative of autosomal dominant inheritance who—unlike the Amsterdam I positive families—only include cases of CRC with onset above 50 years. In Denmark, these families are referred to as ‘Late-onset CRC’ (Lindberg et al. 2016), though there is no international consensus on defining this disease entity. Since no international criteria defining families with onset of CRC > 50 years exist, both ‘late-onset CRC’ (sometimes defined as onset after age 50 or after age 70) and ‘intermediate onset CRC’ (sometimes used to define the onset of CRC between the ages 50 and 70 years) have been used, which challenges scientific studies of this group of families (Antelo et al. 2012; Pillozzi et al. 2015; Brandariz et al. 2018; Álvaro et al. 2019; Arriba et al. 2019; Archambault et al. 2019).

CRC with onset after age 70 has been characterized by a high female proportion, increased levels of microsatellite instability (MSI)—often due to somatic *MLH1* promoter methylation—recurrent *BRAF* mutations, right-sided colon cancer and multiple primary cancers (Pillozzi et al. 2015; Brandariz et al. 2018; Álvaro et al. 2019). Very few studies have investigated the ‘intermediate onset’ CRC between age 50–70 years. One study has shown that this group of patients share characteristics, such as gender and prognosis, with the early onset CRC patients (i.e. onset < 50 years) but also have several features reminiscent of CRCs that develop above age 70. Furthermore, a gradual correlation was found with more familial accumulation of CRC and HNPCC/LS-related cancer in cases with early-onset CRCs than in patients with late-onset CRC, and the proportion of apparently sporadic cases increased with higher age (43% of early onset CRC cases, 47% of intermediate-onset CRC cases and 78% of

late-onset CRC cases). These characteristics suggest that a proportion of the patients with intermediate onset of CRC have a monogenic etiology (Arriba et al. 2019).

We aimed to estimate the contribution of pathogenic variants in established or candidate CRC-predisposing genes in Danish late-onset CRC families with a strong family history of CRC with onset between age 50 and age 70.

Materials and methods

Patient selection

The families in the HNPCC Register have been reported to the national Danish HNPCC register by geneticists and surgeons since 1991. In December 2016, 324 families fulfilled the Danish ‘late-onset CRC’ criteria and 299 families fulfilled the ‘late-onset CRC suspected’ criteria. The ‘late-onset CRC’ criteria are equal to the Amsterdam I criteria except that all cases of CRC have been diagnosed ≥ 50 years. The ‘late-onset CRC suspected’ criteria covers families, who for several reasons do not meet the ‘late-onset CRC’ criteria and include: (A) Late-onset CRC criteria but with only one affected generation; (B) Two CRCs and one high grade adenoma instead of the third CRC; (C) Two CRCs in a small family; (D) Two CRCs and one HNPCC-related cancer; (E) At least three CRCs but one generation is unaffected/skipped.

All pedigrees from the 623 families of interest (324 ‘late-onset CRC’ families and 299 ‘late-onset CRC suspected’ families) were manually evaluated and the families with the highest potential (based on the following characteristics) of having a monogenic, high-risk variant were selected: Accumulation of CRC at a younger age, i.e. between age 50 to age 70 years (but preferably between age 50 years to age 60 years), a large number of affected individuals with CRC and also with other types of cancer (of any kind but preferably HNPCC/LS related) and with a clear dominant inheritance pattern. Available DNA from either a living affected individual with either cancer or colonic adenomas or a deceased, affected individual was an inclusion criterion. This resulted in 182 families eligible for inclusion, 163 from the ‘late-onset CRC’ group of families and 19 from the ‘late-onset CRC suspected’ group of families.

A total of 220 individuals from the 182 families were selected and invited to participate from January 1, 2017 to April 30, 2019. All living participants received written information about the project, and all individuals included received genetic counseling at the local Department of Clinical Genetics (Copenhagen, Aarhus, Aalborg, Odense or Vejle) either by phone or by physical consultation. In total, 168 individuals responded, of which 107 accepted the invitation. Dropouts were reported for 29 individuals due to

severe illness ($n=2$), hearing loss resulting in the inability to attend a telephone conversation ($n=1$), worries or doubts ($n=8$), other reasons ($n=1$) or because they did not answer any of the three independent phone calls in which the genetic counseling should be planned ($n=17$). A total of 78 patients were included from the following subgroups: ‘Late-onset CRC’ ($n=65$), late-onset CRC suspected ($n=13$).

Following oral and written information, the participants signed a written consent. Blood samples were taken at the nearest hospital for genetic analyses. The study was approved by the Ethical and Scientific Committee of the Capitol Region (H-4-2014-050).

NGS sequencing

The DNA extraction and sequencing methods used in here have been described previously (Djursby et al. 2020). Briefly, germline DNA purified from lymphocytes from peripheral whole blood samples was fragmented to an average size of 400 bp (Covaris S2 AFA ultrasonicator). Using biotinylated oligos provided through Roche NimbleGen (Roche, Basel, Switzerland) target DNA sequences were captured including exons and exon/intron boundaries (± 50 bp). 1400 ng of genomic DNA was used to prepare the library using KAPA library DNA adaptors (Roche Diagnostics, Basel, Switzerland). The target fragments were sequenced using paired-end next generation sequencing (NGS) (MiSeq, Illumina, USA), and data was processed using a standard clinical GATK-based pipeline (v. 4.1.0.0 suite) (McKenna et al. 2010) and germline variants were called (Poplin et al. 2017).

Gene panel analysis

The gene panel analysis included 32 well-established CRC predisposing genes, other well-established cancer predisposing genes and suggested CRC candidate genes; a complete list can be found in Supplementary Table S1. In addition to *PMS2*, the gene panel also target *PMS2* pseudogenes. Therefore, specific care is taken classifying variants in this gene, and variants classified as class 3, 4, and 5 with a variant allele frequency below 0.35 is therefore also analyzed by long-range PCR analysis.

Ingenuity variant analysis

The called variants were filtered using Ingenuity Variant Analysis (IVA; <https://www.qiagen.com/>). In the first step, we selected variants for further classification based on the following criteria: call quality > 20 , read depth > 10 and a variant allele fraction (VAF) > 15 . In the second step, we only selected variants with a minor allele frequency (MAF) $< 0.5\%$ (i.e., a population frequency $< 0.5\%$) in gnomAD (<https://gnomad.broadinstitute.org/>) unless a variant

was an established pathogenic variant. In the third step, we selected variants fulfilling the following four Ingenuity settings: 1) Exonic variants and intronic variants ± 20 bp AND 2) the variant should be disease-associated based on computed American College of Medical Genetics and Genomics (ACMG) Guidelines classification in Ingenuity or listed in HGMD or ClinVar, OR 3) The variant should be in a gene associated with gain of function, which was established in the literature OR 4) the variant should result in a frameshift/in-frame indel/missense variant or start/stop codon change, or be predicted deleterious by having a CADD score > 20 , or the variant should be predicted to cause splice site loss up to 10 bp into an intron.

Variant classification

After variant filtering, the remaining variants were classified based on the guidelines from the American College of Medical Genetics and Genomics and the Association for Molecular Pathology using a 5 tier system: 1 = Not pathogenic/no clinical significance, 2 = Likely not pathogenic/little clinical significance, 3 = Uncertain significance/variant of uncertain significance (VUS), 4 = Likely Pathogenic, 5 = Pathogenic (Plon et al. 2008; Richards et al. 2015). We used different tools, programs and databases for variant interpretation, such as Alamut Visual (<https://www.interactive-bioinformatics.com>), ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>), gnomAD, InSiGHT (<https://www.insight-group.org/>), COSMIC (<https://www.sanger.ac.uk/tool/cosmic/>) and PubMed (<https://pubmed.ncbi.nlm.nih.gov/>). Selected variants including all variants listed in Tables 2 and 3 have been manually inspected using the Integrative Genomics Viewer (<https://software.broadinstitute.org>) (Robinson et al. 2011).

Results

A clinical description of the participants can be found in Table 1. The 78 participants came from 75 families; 70 of the participants had developed CRC, 4 had another cancer (breast cancer ($n=1$), breast cancer and leukemia ($n=1$), endometrial cancer ($n=2$)) and 4 only presented with colorectal adenomas. MMR protein IHC analysis had previously been performed in 69% of the patients with findings of loss of expression in 12 (15%) individuals but at the same time retained normal expression in another family member. In total, 54% of the participants had MMR proficient tumors based on IHC/MSI data. This corresponded well with previous MMR gene analysis, which had been performed in 67% of the patients. In all patients tested, the MMR gene analysis had not detected pathogenic variants.

Gene panel analyses identified 342 variants among the 32 genes studied. After filtering, as described in the methods

Table 1 Clinical characteristics of the participants and their families

Number of participants included	78
Number of families included	75
Late-onset CRC	65 (83%)
Late-onset CRC suspected	
3 CRC in one generation (group A)	4 (5%)
2 CRC and 1 adenoma (group B)	2 (3%)
2 CRC + 1 HNPCC-related cancer (group D)	5 (6%)
3 CRC with one unaffected FDR link (group E)	2 (3%)
Gender	
Women	39 (50%)
Men	39 (50%)
Diagnosis (of the participants)	
CRC < 60 years	44 (56%)
CRC ≥ 60 years	26 (33%)
Other cancer	4 (5%)
No cancer (but colorectal adenomas)	4 (5%)
Earliest CRC diagnosis in the family, mean (range)	56 (50–68)
Highest number of adenomas in the family, mean (range)	3 (0–17)
Tumor type and localization in the participants	
Colon, left	20 (26%)
Colon, right	17 (22%)
Rectum	29 (37%)
Colon, unspecified	4 (5%)
Breast	2 (3%)
Endometrial cancer	2 (3%)
Only adenomas	4 (5%)
Previous MSI result (in the participants)	
Microsatellite instable	1 (1%)
Microsatellite stable	3 (4%)
Not available	74 (95%)
Previous MMR IHC result (in the participants)	
Deficient	12 (15%)*
Proficient	42 (54%)
Not available	24 (31%)
Previous MMR gene test (in the participants)	
MMR wildtype	52 (67%)
Not available	26 (33%)

*All had another relative with an MMR proficient CRC

section, 97 variants with an allele frequency < 0.5% (i.e. a population frequency < 1%) in 55 patients were identified. In 23 patients, no variants were identified, 21 patients had 1 variant, 22 patients had 2 variants, 6 patients had 3 variants, 5 patients had 4 variants, and 1 patient had 6 variants.

Among the 97 variants, 4 were pathogenic or likely pathogenic (Table 2).

Pathogenic and likely pathogenic variants

In one person (#1) we identified a frameshift variant in *MSH6*, c.3261dupC, p.(Phe1088Leufs*5). The variant has previously been reported several times (Lavoine et al. 2015; Espenschied et al. 2017; Haraldsdottir et al. 2017) and is classified as pathogenic by InSight (<https://www.insight-group.org/variants/databases/>). In our study, the participant was diagnosed with two LS-associated tumors (synchronous endometrial and rectal cancer) at age 52 years. IHC analysis performed > 10 years ago as part of the clinical evaluation showed normal expression of all four MMR proteins in the rectal tumor and IHC analysis in the endometrial cancer revealed loss of MLH1 and PMS2 expression; data on whether the *MLH1* promoter was methylated is not available. However, following the abnormal IHC result, *MLH1* screening was performed at time of diagnosis without identification of a pathogenic variant. Segregation analysis has not been performed in the family. Still, a second-degree relative with CRC at age 54 had previously had IHC analysis performed with normal expression of all four MMR proteins. Since the IHC analysis in the second degree relative was performed more than 10 years ago at a time of less optimized IHC performance (Palomaki et al. 2009), we repeated the IHC analysis. We demonstrated loss of MSH6 expression based on IHC staining in two independent laboratories. The tumor also showed a high level of microsatellite instability (MSI-H). Unfortunately, we do not know whether the second degree relative also carried the *MSH6* variant.

In two patients, we identified heterozygosity for pathogenic or likely pathogenic *MUTYH* variants. The first patient (#2) had a rare variant, c.1550_1551delinsAG, p.(Cys517*). The variant is not listed in gnomAD but is reported once in

Table 2 Pathogenic or likely pathogenic variants

Family	Gene	Transcript variant	Protein variant	Variant type	CADD score	Variant classification	Cancer history (age)
#1	<i>MSH6</i>	c.3261dup	p.(Phe1088Leufs*5)	Frameshift	24.6	5	CRC (52), endometrial cancer (52)
#4	<i>NTHL1</i>	c.268C>T	p.(Q90*)	Nonsense	35.0	5	CRC (62)
#5	<i>NTHL1</i>	c.268C>T	p.(Q90*)	Nonsense	35.0	5	CRC (51)
#3	<i>MUTYH</i>	c.1187G>A	p.(Gly396Asp)	Missense	34.0	5	CRC (50)
#2	<i>MUTYH</i>	c.1550_1551delinsAG	p.(Cys517*)	Nonsense	n/a	4	CRC (62), prostate cancer (67), possibly sarcoma

Table 3 Variants of unknown significance in the MMR genes

Family	Gene	Transcript variant	Protein variant	CADD Score	Highest freq. (population) [‡]	Allele Count [#]	InSight classification	IHC data	Comments
#8	<i>MSH6</i>	c.628-7C>A	p.(?)	<10	0.015% (NFE)	21 (0)	n/a	n/a	Patient with 6 adenomas (not cancer). A relative had tumor <i>MLH1</i> promoter hypermethylation and normal <i>MSH2/MSH6</i> expression
#9	<i>MSH6</i>	c.806C>G	p.(Thr269Ser)	12.8	0.0029% (NFE)	3 (0)	3	n/a	
#6	<i>MSH6</i>	c.3100C>G	p.(Arg1034Gly)	25.3	0.00099% (NFE)	1 (0)	n/a	IHC: Normal (2013)	The patient also has another <i>MSH6</i> VUS
#10	<i>MSH6</i>	c.3260C>A	p.(Pro1087His)	25.2	0.051% (AMR)	33 (0)	n/a	IHC: Normal (2016)	
#6	<i>MSH6</i>	c.3541G>A	p.(Asp1181Asn)	33.0	0.0026% (NFE)	3 (0)	n/a	IHC: Normal (2013)	The patient also has another <i>MSH6</i> VUS
#11	<i>MSH6</i>	c.3674C>T	p.(Thr1225Met)	31.0	0.014% (NFE)	21 (0)	3	n/a	Drost et al.: MMR function intact
#14	<i>PMS2</i>	c.2276-7C>A		11.0	0.00098% (NFE) [§]	1 (0)	n/a	n/a	

References: Drost et al. (2012)

[‡]The highest frequency in gnomAD (in exomes and genomes combined, or exomes alone if no WGS data was available) in non-cancer population. AMR, Latino/admixed American; NFE, European (Non-Finnish)

[#]Allele count: Number of alleles with the variant in gnomAD. The number in parentheses refers to the number of homozygous individuals

[§]Frequency data in gnomAD was uncertain

ClinVar as pathogenic. Since the variant is located in the last exon of the gene, where a new stop gain will not cause nonsense mediated RNA decay, the variant is predicted to create a premature stop codon most likely resulting in a truncated protein lacking the last 33 amino acids. The patient had metachronous sigmoid and prostate cancer at 62 and 67 years, respectively, and > 10 colonic adenomas, of which several had high-grade neoplasia. He has most likely also had a chondrosarcoma based on medical reports. The family history showed another two cases of intermediate onset colon cancer in close relatives. Segregation analysis has not been performed. Since functional data are unavailable, we consider the variant likely pathogenic (class 4). The second patient (#3) with rectal cancer at age 50 and a single adenoma, had the previously reported, pathogenic *MUTYH* variant c.1187G>A, p.(Gly396Asp). The family history revealed several cases of CRC and other cancers.

In two patients (#4 and #5), we detected a nonsense variant, c.268C>T, p.(Gln90*) variant in *NTHL1*. The variant

is located in exon 2, hence the mRNA molecule is predicted to be degraded by nonsense mediated mRNA decay. Both patients were heterozygous carriers. In the first family, there were four cases of CRC in two generations indicating autosomal dominant inheritance. In the second family, there were three cases of CRC in two generations and a distant relative with early-onset breast cancer. The p.(Gln90*) variant causes *NTHL1*-associated polyposis (NAP) in homozygous or compound heterozygous carriers and is considered pathogenic (class 5) (Weren et al. 2015).

Mismatch repair variants

We identified a total of 15 MMR gene variants in *MLH1* ($n=2$), *MSH6* ($n=9$, including the pathogenic c.3261dupC, p.(Phe1088Leufs*5) variant described above) and *PMS2* ($n=4$).

Based on frequency data from gnomAD, variant type and in silico data, seven variants in *MLH1*, *MSH6* and

PMS2 were classified as benign/likely benign (see supplementary Table S2). A total of seven variants were classified as VUS (Table 3). In all families with an MMR VUS, the person tested had either normal expression of all four MMR proteins or MMR protein expression data were unavailable.

Variants of unknown significance in other known CRC predisposing genes

We detected 10 variants in *APC*. Based on frequency data from gnomAD, variant type and in silico data, nine variants were classified as benign/likely benign (see supplementary Table S2). The remaining *APC* variant was classified as a VUS: c.3891T>G, p.(Asp1297Glu). This variant is not listed in gnomAD and has previously been reported to ClinVar twice as a VUS. None of the patients with *APC* variants had polyposis.

In *AXIN2*, the missense variant, c.1882C>T, p.(Arg628Trp) was detected. The variant has been identified 48 times in gnomAD with a pop-max frequency of 0.060% in the Finnish population. The variant is predicted to strengthen a possible existing cryptic splice site by four out of five in silico splice prediction programs in Alamut. It is unknown if there have been dental abnormalities, which has been reported to associate with *AXIN2* defects, in the family. Functional data on splicing are unavailable and we have classified the variant as a VUS.

We detected a missense variant in *POLE*, c.861T>A, p.(Asp287Glu). The variant is located in the CRC-associated exonuclease domain and has been reported in ClinVar several times, mostly interpreted as a VUS. The variant has a total count of 242 out of 282,842 alleles in gnomAD corresponding to a frequency of 0.09% but is most frequent in non-Finnish Europeans, where the allele frequency is 0.18%.

Variants of unknown significance in candidate CRC predisposing genes

In *BRCA2*, we detected the c.3962A > G, p.(Asp1321Gly), variant in an individual with colon cancer at 55 years. The family history showed several CRCs and other cancer types, albeit not breast cancer or ovarian cancer. The variant is most frequent among non-Finnish Europeans with a pop-max of 0.0048%, and it has been reported to ClinVar several times as a VUS. According to four out of five in silico splice prediction programs in Alamut, the variant creates a new cryptic, strong donor site, however, since functional data is unavailable the significance of the variant is unknown.

Discussion

We have analyzed patients from families with an accumulation of CRC diagnosed after age 50, also referred to as late-onset CRC. To our knowledge, no studies have previously performed large gene panel analyses in a similar group of patients, and the intermediate- and late-onset CRC families are in general not well-characterized genetically.

The patients had predominantly MMR-proficient CRCs (based on IHC/MSI data) and in the majority, previous MMR gene tests had been performed without disease-predisposing variants detected. Consistently, only a single patient was diagnosed with Lynch Syndrome (patient #1, *MSH6*: c.3261dupC, p.(Phe1088Leufs*5)). Although familial CRC with onset > 50 years can be expected to be a heterogeneous group consisting of families with high-risk cancer predisposition syndromes and families with accumulation of CRC caused by combinations of hereditary factors, lifestyle factors and environmental factors, we demonstrate low diagnostic yield considering that several families have multiple cases of CRC and other cancers and onset of disease between age 50 and 60. It is striking that moderate risk variants, which may indicate oligogenic or polygenic inheritance, seemed to play a very limited role herein. However, since all cases of CRC in the families—and participants with CRC – occurred above age 50 and since CRC is a common malignancy that increases in frequency with increasing age, the risk of phenocopies is a concern. We tried to come around this issue by including the youngest affected individual when possible, but still, we cannot rule out that including a phenocopy might have influenced the results.

The families were highly selected based on family history and age at onset (but not IHC/MSI status), and although we identified clinically actionable variants (i.e., the pathogenic variants in *MSH6* and *MUTYH*) in 3/75 families (= 4%), the only high-risk cancer predisposition syndrome identified was the patient with *MSH6*-related LS (~ 1.3% of the cohort). In Denmark, familial clustering of late-onset CRC has been systematically reported to the HNPCC register for three decades (Lindberg et al. 2016). In the group of ‘late-onset CRC’ families in the HNPCC registry, another 6 families have been diagnosed with Lynch Syndrome with pathogenic variants in *MLH1* ($n = 2$), *MSH2* ($n = 1$), *MSH6* ($n = 2$) and *PMS2* ($n = 1$) (data from the HNPCC registry). Including our family, this corresponds to ~ 2% of the families (7 out of 324 families). Based on the data from the current study and unpublished data, it seems reasonable to assume that approximately 2% of the Danish late-onset CRC families have LS. This is much higher than the prevalence of LS in the background

population (which has been estimated to be 1:226 to 1:370 Hampel and De La Chapelle 2013; Haraldsdottir et al. 2017; Win et al. 2017; Grzymalski et al. 2020)) but around the same level as unselected CRC patients.

Thus, genetic testing with gene panel analysis of CRC-associated genes in late-onset CRC families appears to be of limited diagnostic value with a diagnostic hit rate of only 1% to 2% of the patients. This questions whether routine genetic testing is indicated in this group of families. In Denmark, universal screening with MMR protein immunostaining (IHC) is performed routinely on all CRC tumors as part of the diagnostic evaluation of the patient, irrespective of age at onset. Hence, in this context, we suggest that in Danish families, gene panel analysis can be restricted to patients from late-onset CRC families with MMR-deficient tumors. The exception is families with features indicating another cancer predisposition syndrome. This approach depends on reliable IHC and MSI results, which historically have been challenging; however, the methods have improved significantly (Leclerc et al. 2021).

Routine genetic testing of gene panels in late-onset CRC families also depends on economic resources in the healthcare systems, availability of genetic counselling, and access to surveillance once a pathogenic high-risk variant has been identified. In high-income countries with healthcare systems able to perform comprehensive genetic testing and clinical follow-up, it could prove beneficial to offer genetic testing of the MMR genes in selected late-onset families (e.g. families with many cases of CRC between age 50 and age 60 or with MMR-deficient tumors or individuals with ≥ 2 HNPCC/LS-associated cancers) or to offer routine IHC/MSI analysis. In countries with sparse economical resources, genetic counselling and -testing could be reserved for families with a higher probability of having high-risk genetic cancer syndromes, such as families with several affected persons with early onset disease, multiple primary tumors, polyposis or syndromic features.

In our study, the patient with the *MSH6* c.3261dupC, p.(Phe1088Leufs*5) variant developed two classic LS-associated tumors at age 52 but IHC analysis in both tumors (performed > 10 years ago) did not reveal lack of MSH6 expression. In a second degree relative with colon cancer at age 54 years and initially normal IHC analysis, repeated IHC analysis showed lack of MSH6 expression and high levels of MSI (MSI-H) (however, although she likely carried the variant, mutational analysis has not been available). Today, the strategy mentioned above would have identified this family due to both abnormal IHC and MSI analysis and syndrome stigmata (> 1 primary LS-associated tumor). This family demonstrates the risk of diagnostic errors when older IHC analysis is used alone and emphasizes the importance of repeating old IHC analyses or performing MSI analysis.

The results leave ~99% of the patients genetically unsolved and the most obvious explanation is a multifactorial predisposition to CRC. Another explanation could be the coinheritance of few low-penetrant variants each conferring a low or moderate risk but with a combined synergistic effect, referred to as digenic or oligogenic inheritance. There is not yet clear evidence of digenic or oligogenic inheritance in CRC but few cases suggesting these types of inheritance have been reported (Morak et al. 2011; Ciavarella et al. 2018; Schubert et al. 2020) and several cases of patients with the accumulation of pathogenic variants have also been published (Whitworth et al. 2016; Stradella et al. 2019). Other possible explanations include epigenetic changes or variants in other candidate genes not included in our gene panel. A perhaps less likely explanation is genetic variants in genes not yet identified or associated with CRC. Over the past years, many research groups have searched for new CRC predisposing genes. Although new CRC-predisposing genes have been identified such as *RPS20*, they appear to account only for a very low proportion of families with familial CRC (Nieminen et al. 2014; Chubb et al. 2016; Valle et al. 2019; Djursby et al. 2020). A weakness in our study is the lack of copy number variation (CNV) analysis. As CNVs account for ~10% of pathogenic variants, with a large difference between genes, it is a possibility that this disease-causing mechanism has been overlooked (LaDuca et al. 2019).

As expected, we identified a multitude of VUS, mainly missense variants, including *MSH6* VUS ($n=6$) but also interesting variants in *BRCA2*, c.3962A>G, p.(Asp1321Gly), and *AXIN2*, c.1882C>T, p.(Arg628Trp), which might affect splicing. Variants in *MSH6* have been shown to confer a significantly lower risk of CRC than *MLH1* and *MSH2* pathogenic variants (Møller et al. 2017). Since the late-onset CRC families have a less severe phenotype as compared to the Amsterdam I/II positive families (at least when it comes to age at onset) or LS families caused by pathogenic *MLH1* or *MSH2* variants, it would not be surprising if a larger proportion of the late-onset CRC families had *MSH6* variants. Intriguingly, the published cases with possible digenic/oligogenic inheritance have also included variants in *MUTYH* and *MSH6*. While biallelic pathogenic variants in *MUTYH* confer a very high lifetime risk of developing CRC, monoallelic variants are associated with a ~twofold increased lifetime risk of developing CRC (Win et al. 2011). In addition to several VUS in *MSH6*, we also identified two patients with monoallelic pathogenic variants in *MUTYH* making it tempting to speculate that these patients could have other (yet unidentified) variants which cause a digenic or oligogenic predisposition to CRC. In two patients, the recurrent truncating *NTHL1* variant c.268C>T, p.(Gln90*) were detected and, like *MUTYH*, *NTHL1* is also involved in base excision repair. In contrast to *MUTYH*, monoallelic variants in *NTHL1* has not been associated with an increased risk of CRC, and large population-based

studies estimating the risk of colorectal cancer and polyposis in monoallelic *NTHL1* carriers are needed to clarify this matter (Elsayed et al. 2020; Salo-mullen et al. 2021).

An interesting perspective is the use of polygenic risk scores (PRS) in the prediction of a personal cancer risk estimation. At present, PRS and models aiming at predicting cancer risk by combining PRS, personal factors (such as gender and age), family history and lifestyle factors as listed above are only moderately accurate (Frampton et al. 2016; Frampton and Houlston 2017; Weigl et al. 2018; Archambault et al. 2020). Further studies exploring the role of PRS in late-onset CRC patients are warranted. Another interesting perspective is use of tumor sequencing and mutational signatures such as the COSMIC mutational signatures (Alexandrov et al. 2013). Tumor testing in several affected individuals in a family could help identifying relatives with the same mutational signature and could therefore also aid in excluding likely phenocopies from further analyses. As certain mutational signatures are associated with specific underlying germline alterations, this approach could also aid in identifying relevant genes or specific pathways in a family.

Based on our data and unpublished data (as referred above), it appears that LS and other CRC predisposition syndromes only play a minor role in families with aggregation of CRC with onset of all cases above age 50. Given the high burden of CRC and other cancers in these families, hereditary factors seem likely to play a role, though shared environmental factors and lifestyle could also be important. Further studies elucidating the genetic mechanisms involved in late-onset CRC families, including calculation of polygenic risk scores, is warranted.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00439-022-02470-9>.

Acknowledgements We thank the families that participated in this project.

Funding Financial support was granted from the Danish Cancer Association (A12.599 and R90-A5807), Rigshospitalets Forskningspuljer (no grant number), Anna og Hans Steffensens Fond til støtte for Dansk Kraeftforskning (grant 2189001), Aase og Ejnar Danielsens Fond (10-001917) and A. P. Møller Fonden: Fonden til Laegevidenskabens Fremme (no grant number).

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

- AlDubayan SH, Giannakis M, Moore ND et al (2018) Inherited DNA-repair defects in colorectal cancer. *Am J Hum Genet* 102:401–414. <https://doi.org/10.1016/j.ajhg.2018.01.018>
- Alexandrov LB, Nik-Zainal S, Wedge DC et al (2013) Signatures of mutational processes in human cancer. *Nature* 500:415–421. <https://doi.org/10.1038/nature12477>
- Álvarez E, Cano JM, García JL et al (2019) Clinical and molecular comparative study of colorectal cancer based on age-of-onset and tumor location: two main criteria for subclassifying colorectal cancer. *Int J Mol Sci*. <https://doi.org/10.3390/ijms20040968>
- Antelo M, Balaguer F, Shia J et al (2012) A High Degree of LINE-1 hypomethylation is a unique feature of early-onset colorectal cancer. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0045357>
- Archambault AN, Su Y-R, Jeon J et al (2019) Cumulative burden of colorectal cancer-associated genetic variants is more strongly associated with early-onset vs late-onset cancer. *Gastroenterology*. <https://doi.org/10.1053/j.gastro.2019.12.012>
- Archambault AN, Su YR, Jeon J et al (2020) Cumulative burden of colorectal cancer-associated genetic variants is more strongly associated with early-onset vs late-onset cancer. *Gastroenterology* 158:1274–1286.e12. <https://doi.org/10.1053/j.gastro.2019.12.012>
- Arriba M, Sánchez C, Vivas A et al (2019) Intermediate-onset colorectal cancer: a clinical and familial boundary between both early and late-onset colorectal cancer. *PLoS ONE* 14:1–12. <https://doi.org/10.1371/journal.pone.0216472>
- Brandariz L, Arriba M, García JL et al (2018) Differential clinicopathological and molecular features within late-onset colorectal cancer according to tumor location. *Oncotarget* 9:15302–15311. <https://doi.org/10.18632/oncotarget.24502>
- Broderick P, Dobbins SE, Chubb D et al (2017) Validation of recently proposed colorectal cancer susceptibility gene variants in an analysis of families and patients—a systematic review. *Gastroenterology* 152:75–77
- Carethers JM, Stoffel EM (2015) Lynch syndrome and Lynch syndrome mimics: the growing complex landscape of hereditary colon cancer. *World J Gastroenterol* 21:9253–9261. <https://doi.org/10.3748/wjg.v21.i31.9253>
- Chubb D, Broderick P, Dobbins SE et al (2016) Rare disruptive mutations and their contribution to the heritable risk of colorectal cancer Daniel. *Nat Commun*. <https://doi.org/10.1038/ncomms11883>
- Ciavarella M, Miccoli S, Prossomariti A et al (2018) Somatic APC mosaicism and oligogenic inheritance in genetically unsolved colorectal adenomatous polyposis patients. *Eur J Hum Genet* 26:387–395
- Djursby M, Madsen MB, Frederiksen JH et al (2020) New pathogenic germline variants in very early onset and familial colorectal cancer patients. *Front Genet* 11:1–16. <https://doi.org/10.3389/fgene.2020.566266>
- Dominguez-Valentin M, Nakken S, Tubeuf H et al (2018) Identification of genetic variants for clinical management of familial colorectal tumors. *BMC Med Genet* 19:1–19. <https://doi.org/10.1186/s12881-018-0533-9>
- Drost M, Zonneveld JBM, van Hees S et al (2012) A rapid and cell-free assay to test the activity of Lynch syndrome-associated MSH2 and MSH6 missense variants. *Hum Mutat* 33:488–494. <https://doi.org/10.1002/humu.22000>
- Elsayed FA, Grolleman JE, Ragunathan A et al (2020) Monoallelic *NTHL1* loss-of-function variants and risk of polyposis and colorectal cancer. *Gastroenterology* 159:2241–2243. <https://doi.org/10.1053/j.gastro.2020.08.042>
- Espenschied CR, LaDuca H, Li S et al (2017) Multigene panel testing provides a new perspective on Lynch syndrome. *J Clin Oncol* 35:2568–2575. <https://doi.org/10.1200/JCO.2016.71.9260>
- Frampton M, Houlston RS (2017) Modeling the prevention of colorectal cancer from the combined impact of host and behavioral risk factors. *Genetics Med* 2017:19. <https://doi.org/10.1038/gim.2016.101>

- Frampton MJE, Law P, Litchfield K et al (2016) Implications of polygenic risk for personalised colorectal cancer screening. *Ann Oncol* 27:429–434. <https://doi.org/10.1093/annonc/mdv540>
- Garre P, Martín L, Sanz J et al (2015) BRCA2 gene: A candidate for clinical testing in familial colorectal cancer type X. *Clin Genet* 87:582–587. <https://doi.org/10.1111/cge.12427>
- Grzymiski JJ, Elhanan G, Morales Rosado JA et al (2020) Population genetic screening efficiently identifies carriers of autosomal dominant diseases. *Nat Med* 26:1235–1239. <https://doi.org/10.1038/s41591-020-0982-5>
- Hampel H, De La Chapelle A (2013) How do we approach the goal of identifying everybody with Lynch Syndrome? *Fam Cancer* 12:313–317. <https://doi.org/10.1007/s10689-013-9611-5>
- Haraldsdottir S, Rafnar T, Frankel WL et al (2017) Comprehensive population-wide analysis of Lynch syndrome in Iceland reveals founder mutations in MSH6 and PMS2. *Nat Commun* 8:1–11. <https://doi.org/10.1038/ncomms14755>
- LaDuca H, Polley EC, Yussuf A et al (2019) A clinical guide to hereditary cancer panel testing: evaluation of gene-specific cancer associations and sensitivity of genetic testing criteria in a cohort of 165,000 high-risk patients. *Genet Med* 22:407–415. <https://doi.org/10.1038/s41436-019-0633-8>
- Lavoine N, Colas C, Muleris M et al (2015) Constitutional mismatch repair deficiency syndrome: clinical description in a French cohort. *J Med Genet* 52:770–778. <https://doi.org/10.1136/jmedgenet-2015-103299>
- Leclerc J, Vermaut C, Buisine MP (2021) Diagnosis of lynch syndrome and strategies to distinguish lynch-related tumors from sporadic MSI/DMMR tumors. *Cancers (basel)* 13:1–33. <https://doi.org/10.3390/cancers13030467>
- Lindberg LJ, Ladelund S, Frederiksen BL et al (2016) Outcome of 24 years national surveillance in different hereditary colorectal cancer subgroups leading to more individualised surveillance. *J Med Genet*. <https://doi.org/10.1136/jmedgenet-2016-104284>
- McKenna A, Hanna M, Banks E et al (2010) The Genome analysis toolkit: a mapreduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20:1297–1303. <https://doi.org/10.1101/gr.107524.110.20>
- Møller P, Seppälä T, Bernstien I et al (2017) Cancer incidence and survival in Lynch syndrome patients receiving colonoscopic and gynaecological surveillance: first report from the prospective Lynch syndrome database. *Gut* 66:464–472. <https://doi.org/10.1136/gutjnl-2015-309675>
- Morak M, Massdorf T, Sykora H et al (2011) First evidence for digenic inheritance in hereditary colorectal cancer by mutations in the base excision repair genes. *Eur J Cancer* 47:1046–1055
- Nieminen TT, O'Donohue MF, Wu Y et al (2014) Germline mutation of RPS20, encoding a ribosomal protein, causes predisposition to hereditary nonpolyposis colorectal carcinoma without DNA mismatch repair deficiency. *Gastroenterology* 147:595–598
- Palles C, Cazier J, Howarth K et al (2013) Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. *Nat Genet* 45:136–146
- Palomaki GE, McClain MR, Melillo S et al (2009) EGAPP supplementary evidence review: DNA testing strategies aimed at reducing morbidity and mortality from Lynch syndrome. *Genet Med* 11:42–65. <https://doi.org/10.1097/GIM.0b013e31818fa2db>
- Pilozzi E, Maresca C, Duranti E et al (2015) Left-sided early-onset vs late-onset colorectal carcinoma histologic, clinical, and molecular differences. *Am J Clin Pathol* 143:374–384. <https://doi.org/10.1309/AJCPNOC55IOLXFUD>
- Plon SE, Eccles DM, Easton D et al (2008) Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. *Hum Mutat* 29:1282–1291. <https://doi.org/10.1002/humu.20880>
- Poplin R, Ruano-Rubio V, DePristo MA et al (2017) Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv* 10:7
- Richards S, Aziz N, Bale S et al (2015) 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* 17:405–424. <https://doi.org/10.1038/gim.2015.30>
- Robinson JT, Thorvaldsdóttir H, Winckler W et al (2011) Integrative genome Viewer. *Nat Biotechnol* 29:24–26. <https://doi.org/10.1038/nbt.1754>
- Rosenthal ET, Bernhisel R, Brown K et al (2017) Clinical testing with a panel of 25 genes associated with increased cancer risk results in a significant increase in clinically significant findings across a broad range of cancer histories. *Cancer Genet* 218–219:58–68. <https://doi.org/10.1016/j.cancergen.2017.09.003>
- Salo-mullen EE, Maio A, Mukherjee S et al (2021) Prevalence and characterization of biallelic and monoallelic NTHL1 and MSH3 variant carriers from a pan-cancer patient population. *JCO Precis Oncol* 26:455–465
- Schubert SA, Ruano D, Tiersma Y et al (2020) Digenic inheritance of MSH6 and MUTYH variants in familial colorectal cancer. *Genes Chromosom Cancer* 59:697–701. <https://doi.org/10.1002/gcc.22883>
- Stradella A, Del Valle J, Rofes P et al (2019) Does multilocus inherited neoplasia alleles syndrome have severe clinical expression? *J Med Genet* 56:521–525. <https://doi.org/10.1136/jmedgenet-2018-105700>
- Thompson BA, Snow AK, Koptiuch C et al (2020) A novel ribosomal protein S20 variant in a family with unexplained colorectal cancer and polyposis. *Clin Genet* 97:943–944. <https://doi.org/10.1111/cge.13757>
- Valle L (2017) Recent discoveries in the genetics of familial colorectal cancer and polyposis. *Clin Gastroenterol Hepatol* 15:809–819. <https://doi.org/10.1016/j.cgh.2016.09.148>
- Valle L, de Voer RM, Goldberg Y et al (2019) Update on genetic predisposition to colorectal cancer and polyposis. *Mol Aspects Med* 69:10–26. <https://doi.org/10.1016/j.mam.2019.03.001>
- Vasen HF, Mecklin JP, Khan PM, Lynch HT (1991) The international collaborative group on hereditary non-polyposis colorectal cancer (ICG-HNPCC). *Dis Colon Rectum* 34:424–425
- Weigl K, Chang-claude J, Knebel P et al (2018) Strongly enhanced colorectal cancer risk stratification by combining family history and genetic risk score. *Clin Epidemiol* 10:143–152
- Weren RDA, Ligtenberg MJL, Kets CM et al (2015) A germline homozygous mutation in the base-excision repair gene NTHL1 causes adenomatous polyposis and colorectal cancer. *Nat Genet* 47:668–671
- Whitworth J, Skytte AB, Sunde L et al (2016) Multilocus inherited neoplasia alleles syndrome. *JAMA Oncol* 2:373–379. <https://doi.org/10.1001/jamaoncol.2015.4771>
- Win AK, Cleary SP, Dowty JG et al (2011) Cancer risks for monoallelic MUTYH mutation carriers with a family history of colorectal cancer. *Int J Cancer* 129:2256–2262. <https://doi.org/10.1007/bf02236735>
- Win AK, Jenkins MA, Dowty JG et al (2017) Prevalence and penetrance of major genes and polygenes for colorectal cancer. *Cancer Epidemiol Biomarkers Prev* 26:404–412
- Yurgelun MB, Allen B, Kaldate RR et al (2015a) Identification of a variety of mutations in cancer predisposition genes in patients with suspected lynch syndrome. *Gastroenterology* 149:604–613.e20. <https://doi.org/10.1053/j.gastro.2015.05.006>
- Yurgelun MB, Masciari S, Joshi VA et al (2015b) Germline tp53 mutations in patients with early-onset colorectal cancer in the colon cancer family registry. *JAMA Oncol*. <https://doi.org/10.1001/jamaoncol.2015.0197>

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