The missing heritability of familial colorectal cancer

Schubert, S.A.; Morreau, H.; Miranda, N.F.C.C. de; Wezel, T. van

Mutagenesis, 2020, 35, 221–231
doi:10.1093/mutage/gez027
Review
Advance Access publication 12 October 2019

Review

The missing heritability of familial colorectal cancer

Stephanie A. Schubert, Hans Morreau, Noel F. C. C. de Miranda and Tom van Wezel*

Department of Pathology, Leiden University Medical Center, Leiden University, Albinusdreef 2, 2333 ZA Leiden, The Netherlands

*To whom correspondence should be addressed. Tel: +31 71 5266629; Fax: +31 71 5266952; Email: t.van_wezel@lumc.nl

Received 20 June 2019; Accepted 5 September 2019.

Abstract

Pinpointing heritability factors is fundamental for the prevention and early detection of cancer. Up to one-quarter of colorectal cancers (CRCs) occur in the context of familial aggregation of this disease, suggesting a strong genetic component. Currently, only less than half of the heritability of CRC can be attributed to hereditary syndromes or common risk loci. Part of the missing heritability of this disease may be explained by the inheritance of elusive high-risk variants, polygenic inheritance, somatic mosaicism, as well as shared environmental factors, among others. A great deal of the missing heritability in CRC is expected to be addressed in the coming years with the increased application of cutting-edge next-generation sequencing technologies, routine multigene panel testing and tumour-focussed germline predisposition screening approaches. On the other hand, it will be important to define the contribution of environmental factors to familial aggregation of CRC incidence. This review provides an overview of the known genetic causes of familial CRC and aims at providing clues that explain the missing heritability of this disease.

Introduction

Colorectal cancer (CRC) is a common disease; it is the third most frequent cancer in men and the second most frequent in women worldwide (1). Family history is a major risk factor for the development of CRC (2,3). Familial aggregation of this disease is observed in up to 25% of the cases (4–6). In addition, the risk for colorectal carcinomas is increased in individuals with a personal or family medical history of colorectal polyps (2,7). Thus, CRC syndromes have historically been divided into polyposis (adenomatous, hamartomatous, mixed and serrated) and non-polyposis syndromes, based largely on the number and histology of the colorectal polyps. However, their clinical presentation can be highly variable and overlapping. Luckily, molecular tumour features can aid in the diagnosis of hereditary CRC syndromes caused by underlying germline DNA repair defects, as seen in Lynch syndrome (LS) and MUTYH-associated polyposis (MAP). In both instances, tumour analysis helped identify the aetiology of these syndromes; microsatellite instability (MSI) resulted in the identification of the DNA mismatch repair (MMR) gene *MSH2* as the first causal gene for LS, and base-excision repair (BER)-deficiency pointed towards biallelic *MUTYH* mutations as the molecular basis of disease in a subset of colorectal adenomatous polyposis patients. Thus, the screening of somatic features in cancer tissues has already proven to be crucial in the identification of novel predisposition genes in the past and is likely to play a large role in routine onco-genetic diagnostics in the coming years with the widespread implementation of next-generation sequencing (NGS) methods in the molecular pathology. Currently, tumours can be classified more extensively on the basis of somatic alterations than previously possible, focussing on somatic mutations in genetic hotspot regions, tumour mutation burden (TMB) (i.e. the number of somatic events), and somatic mutational signatures, describing genome-wide mutation patterns generated by specific underlying mutational processes (8). NGS approaches analysing germline (i.e. leucocyte-derived) DNA have been less successful in identifying novel familial CRC genes than originally anticipated (9,10). Nevertheless, they have played a key role in the discovery of some predisposition genes for CRC and polyposis, e.g. *POLE*, *POLD1*, *NTHL1* and *MSH3*; however, these syndromes only explain a minority of the familial cases (11).

Currently, only up to one-quarter of all the familial CRCs can be attributed to the known Mendelian-inherited syndromes, suggesting

that a substantial portion of its heritability still remains unexplained (12). In the polyposis subgroup, germline diagnostic approaches are more successful as less than 20% of the cases remain unexplained after screening the known polyposis-associated genes (13). Strikingly, no genetic cause can be identified for MMR-deficiency in 60% of the MMR-deficient CRC cases, the so-called 'suspected Lynch syndrome' or 'Lynch-like syndrome' patients (14,15). Finally, in the MMR-proficient non-polyposis familial CRC subgroup, in spite of enormous efforts in the last years, no causal genes have been confirmed. The following review is an overview of the genetic causes of familial CRCs and proposes explanations for the missing heritability of this heterogenous disease.

Known CRC and polyposis syndromes

Approximately 5% of the CRCs occur as a consequence of well-defined inherited syndromes (16) (Table 1). The majority of these cases are caused by CRC or polyposis syndromes, which can be either dominantly or recessively inherited, while a small fraction can be attributed to rare multi-tumour syndromes with an increased risk for CRC (12,17).

Lynch syndrome

LS (OMIM#120435) is the most frequent form of hereditary CRC, estimated to account for almost 3% of all CRCs in the general population (18,19). This syndrome also predisposes to endometrial, ovarian, stomach, small bowel and urinary tract cancers (20). It is caused by heterozygous mutations in one of the DNA MMR genes, *MLH1*, *MSH2*, *MSH6* and *PMS2*. Furthermore, deletion of *EPCAM* resulting in transcriptional inactivation of *MSH2* and heritable *MLH1* promoter hypermethylation can cause LS (21–23). Notably, biallelic germline mutations in one of the MMR genes, known as constitutional MMR-deficiency syndrome (CMMRD; OMIM#276300), predispose to a large spectrum of childhood and adolescent cancers and adenomatous polyposis (24). The MMR system broadly consists of two major components: MutS and MutL (25). MutS heterodimers (MSH2/MSH6 or MSH2/MSH3) recognise mismatches and small insertions/deletions (indels). MutL heterodimers (MLH1/PMS2, MLH1/PMS1 or MLH1/MLH3) act as endonucleases after complexing with MutS. In general, the clinical phenotype of *MLH1* or *MSH2* mutation carriers is more severe than *PMS2* or *MSH6* mutation carriers, likely due to the existence of functional redundancy within the MutS and MutL heterodimers (26,27). MMR-deficiency in LS-associated neoplasms is usually the result of a second hit in the respective MMR gene, frequently leading to loss of protein expression and a microsatellite instability-high (MSI-H) phenotype (28). Approximately 15% of the sporadic CRCs also show this hypermutator phenotype, mainly due to somatic promoter hypermethylation of the *MLH1* gene (29). The MSI-H hypermutator phenotype is associated with a high mutation load (≥10 mutations/Mb) with a preferential accumulation of C:G>T:A and C:G>A:T mutations and thousands of small insertions/deletions, particularly at microsatellite sequences (9,30). This mutational process is thought to underlie multiple mutational signatures identified

Table 1. Colorectal cancer-associated genes

Genes	Syndrome	Abbr.	Inheritance	Proportion of colorectal cancer	Tumour genetics	COSMIC mutational signatures v3[a]
MLH1 *MSH2* *MSH6* *PMS2*	Lynch syndrome (Constitutional mismatch repair-deficiency)	LS (CMMRD)	Dominant (Recessive)	~3% (Unknown)	MSI-H, hypermutated	SBS6, SBS15, SBS21, SBS26, SBS44, DBS7, DBS10, ID7, ID1[c], ID2[c]
APC	(Attenuated) Familial adenomatous polyposis	AFAP/FAP	Dominant	~1%		
MUTYH	MUTYH-associated polyposis	MAP	Recessive	<1%		SBS36
POLE *POLD1*	POLE/POLD1-associated tumour syndrome[b]	PPATS	Dominant	Unknown	Ultra-hypermutated	SBS10a, SBS10b, SBS14, SBS20, DBS3
NTHL1	NTHL1-associated tumour syndrome	NATS	Recessive	Unknown		SBS30
SMAD4 *BMPR1A*	Juvenile polyposis	JPS	Dominant	<1%		
STK11	Peutz-Jeghers syndrome	PJS	Dominant	<1%		
GREM1	Hereditary mixed polyposis syndrome	HMPS	Dominant	<1%		
MSH3	MSH3-associated polyposis		Recessive	Unknown	EMAST	
RNF43	Serrated polyposis syndrome	SPS	Dominant	Unknown	CIMP	
PTEN	PTEN hamartoma tumour syndrome	PHTS	Dominant	<1%		
TP53	Li-Fraumeni syndrome	LFS	Dominant	Unknown		

[a]https://cancer.sanger.ac.uk/cosmic/signatures/(8).
[b]Also known as polymerase proofreading-associated polyposis (PPAP).
[c]More than 10 000 mutations per tumour.
SBS, single base substitution signature; DBS, double base substitution signature; ID, insertion and deletion signature.

in CRCs or other tumours: single base substitution mutational signatures SBS6, SBS15, SBS21, SBS26 and SBS44, double base substitutions DBS7 and DBS10 and insertion and deletion signature ID7 and very large numbers (>10 000) of ID1 and ID2 mutations (8).

Familial adenomatous polyposis

The most prevalent polyposis syndrome and the second most common cause of hereditary CRC, explaining approximately 1% of the CRCs, is familial adenomatous polyposis (FAP; OMIM#175100), caused by heterozygous mutations in the *APC* gene (19). While FAP exhibits an autosomal dominant pattern of inheritance, approximately 25% of affected individuals have no family history and present with *de novo* mutations (31,32). FAP is a severe polyposis syndrome associated with the development of hundreds to thousands of colorectal adenomas at early age and a lifetime risk of CRC of nearly 100% (16). Attenuated FAP (AFAP) is a milder form of adenomatous polyposis with fewer adenomas (<100) at a later age, and is typically caused by *APC* mutations located at the proximal and distal ends of the gene or exon 9 (33,34). Both FAP and AFAP patients have an increased risk for the development of extra-colonic polyps (16). APC is a large tumour suppressor (2843 amino acids and 312 kDa) involved in the Wnt signalling pathway, which functions by negatively regulating the β-catenin oncoprotein (35). APC is somatically inactivated in about 80% of the sporadic CRCs and is believed to be one of the initiating steps of tumorigenesis (36). Therefore, the pathogenesis of FAP tumours seems to mimic their sporadic counterparts (37).

MUTYH-associated polyposis

Another possible diagnosis for patients presenting with 10–100 adenomas is MAP (OMIM#608456), which accounts for less than 1% of all CRCs (19). The lifetime risk of CRC development in MAP patients is ~80%, while their risk for extra-colonic cancers seems only slightly increased (16,38,39). MAP is an autosomal recessive disorder caused by biallelic mutations in the *MUTYH* gene (40,41). MUTYH is a BER protein, involved in the cell's protection and repair mechanism against one of the most common oxidative DNA lesions, 8-oxoguanine (8-oxoG) (42). MUTYH suppresses tumorigenesis by repairing mutagenic G:C>T:A transversions, but also by inducing MUTYH-dependent cell death (40,43). Consequently, MUTYH-defective tumours show an excess of G:C>T:A transversions, which are characteristic of mutational signature SBS36 (8,44). In addition, the detection of a c.34G>T transversion at codon 12 of *KRAS* in colorectal tumours can point to the diagnosis of MAP (45). Furthermore, the widespread occurrence of chromosomal copy-neutral loss of heterozygosity (cnLOH) is another peculiar feature of MAP CRCs (46).

NTHL1-associated tumour syndrome

NTHL1 encodes a DNA glycosylase that participates in BER and recessively inherited mutations in this gene cause a polyposis and CRC syndrome (OMIM#616415) estimated to be at least five times less frequent than MAP (47,48). More extensive clinical characterisation of biallelic mutation carriers has prompted a name change from NTHL1-associated polyposis (NAP) to the broader NTHL1-associated tumour syndrome (NATS), as these patients are also at increased risk of developing brain, breast and some other cancer types (10,47). Similar to MAP tumours, the BER-deficiency in NATS tumours results in a unique, clearly distinct mutational signature composed of C:G>T:A transitions at non-CpG sites, designated as mutational signature SBS30 (8,49).

POLE/POLD1-associated tumour syndrome

A small fraction of familial CRC patients are carriers of a heterozygous germline mutation in the exonuclease domain of the DNA polymerase subunits *POLE* (OMIM#615083) or *POLD1* (OMIM#612591); however, their exact prevalence remains unknown (50,51). Originally identified in adenomatous polyposis patients, carriers have also been found amongst non-polyposis CRC patients with MMR-proficient or MMR-deficient tumours (52,53). Their tumour spectrum has recently been extended to include additional neoplasms, such as endometrial, ovarian, brain and pancreatic cancer (54–56). The original name coined for this syndrome is 'polymerase proofreading-associated polyposis (PPAP)'; however, we believe '*POLE/POLD1*-associated tumour syndrome' would be a more appropriate name due to the recently expanded tumour spectrum and the absence of colonic polyposis in some patients. The genes *POLE* and *POLD1* code for the catalytic subunit of DNA polymerase ε (epsilon) and δ (delta), involved in DNA replication of the leading and lagging strands, respectively (57). Tumours with pathogenic *POLE* or *POLD1* mutations, either germline or somatic, show an ultra-hypermutator phenotype with the number of somatic mutations exceeding the 100 mutations/Mb and suggestive of a deficiency of the proofreading capacity of these polymerases (9,58). POLE-mutant tumours especially show an excess of C:G>A:T and C:G>T:A mutations, characteristic of mutational signatures SBS10a and SBS10b, and enrichment of DBS3 in some tumours (8,30). Thus far, no clear mutational signatures have been associated with *POLD1* exonuclease domain mutations. Interestingly, a combined deficiency in MMR function and POLE or POLD1 proofreading is relatively common and results in the generation of unique mutational signatures SBS14 and SBS20, respectively (8).

Rare polyposis syndromes

A fraction of the familial CRC risk can be explained by rare autosomal dominant hamartomatous polyposis syndromes, occurring with approximately one-tenth of the frequency of adenomatous polyposis syndromes (59). Two of the most common are juvenile polyposis (JPS; OMIM#174900; *SMAD4* and *BMPR1A*) and Peutz-Jeghers syndrome (PJS; OMIM#175200; *STK11*) (60–63). In addition, patients with hereditary mixed polyposis syndrome (HMPS; OMIM#601228; *GREM1*) present with polyps of different histological type, including hamartomatous polyps (59,64). These syndromes have distinct clinical phenotypes, histological features, frequencies and location of polyps, organ-specific manifestations and predispositions for the development of other malignancies (16,60).

In addition, in very rare cases, serrated polyposis syndrome (SPS) can be attributed to heterozygous (likely) pathogenic mutations in the *RNF43* gene (OMIM#617108; 7 families), encoding an E3 ubiquitin ligase that acts as a Wnt inhibitor (65,66). A common feature of the serrated pathway is the occurrence of CpG island methylator phenotype (CIMP), which was also observed in more than 75% of the RNF43 syndrome-associated colonic lesions (66). Another very rare genetic cause of adenomatous polyposis was reported in one study, where biallelic mutations in the MMR gene *MSH3* were identified in two families with unexplained adenomatous polyposis, resulting in MSH3-associated polyposis syndrome (OMIM#617100) (67). These tumours showed loss of MSH3 expression and a high degree of microsatellite instability in tetranucleotide repeats (EMAST). MSH3-deficiency has previously been shown to induce this mutator phenotype (68,69). The exact prevalence and penetrance of

heterozygous mutations in *RNF43* and biallelic mutations in *MSH3* still remain unknown and warrant further research.

Multi-tumour syndromes

Currently, approximately three-quarters of familial CRC cases cannot be explained by germline mutations in known CRC-associated genes (12). Both classical testing strategies and multigene panel tests in CRC cases have uncovered pathogenic germline variants in genes associated with inherited syndromes not primarily linked with an increased risk for CRC.

PTEN hamartoma tumour syndrome

Heterozygous germline mutations in the tumour suppressor gene *PTEN* can cause PTEN hamartoma tumour syndrome (PHTS; OMIM#158350), a collection of rare syndromes with overlapping clinical presentations (70). The phenotypic spectrum of PHTS includes Cowden syndrome (CS), Bannayan-Riley-Ruvalcaba syndrome (BRRS), Proteus syndrome (PS) and Proteus-like syndrome (PSL) (70). CS and BRRS are considered variable phenotypic presentations of the same hamartomatous polyposis syndromes (71). CS is the best-described phenotype of PHTS and is classically characterised by the development of multiple hamartomas and characteristic dermatological manifestations, but also with an increased risk of breast, endometrial, thyroid and CRC (60,72). BRRS patients are generally affected by gastrointestinal hamartomatous polyps, lipomas, haemangiomas and macrocephaly, but a high cancer risk has been less well-documented (60,71).

Li-Fraumeni syndrome

Another multi-tumour syndrome with an increased risk for CRC is Li-Fraumeni syndrome (LFS; OMIM#151623), caused by heterozygous mutations in the *TP53* gene. This is a very severe hereditary cancer syndrome with a lifetime cancer risk of 70–100% (73). Patients are predisposed to various and multiple early-onset cancers, mainly bone and soft tissue sarcomas, breast cancer and brain cancer (74). Between 4–14% of the Li-Fraumeni syndrome patients are diagnosed with CRC, mostly before 50 years of age (74–76). Furthermore, recent germline screening efforts in large cohorts of early-onset CRC patients frequently (0.1–1%) detected *TP53* mutations in these patients (19,77,78).

Other hereditary tumour syndromes

Recently, additional genes causing hereditary tumour syndromes not classically associated with CRC have also been implicated in CRC predisposition (11). For example, hereditary breast cancer genes (OMIM#11448; e.g. *BRCA1*, *BRCA2*, *CHEK2*, *ATM* and *PALB2*) might confer a low-to-moderate risk for CRC (19,79). Rare *CHEK2* variants have been associated with a 2-fold increased risk of CRC (80). *CHEK2* is considered a low-to-moderate risk gene in multi-organ cancer susceptibility (OMIM+604373), including breast, thyroid, prostate, colon and kidney tumours (81). The *CHEK2* mutation c.1100delC is primarily associated with an increased risk for breast cancer, but it seems especially frequent in families with both hereditary breast and CRC combined (82,83). Furthermore, CRC risk associations in *BRCA1* (OMIM#60437) and *BRCA2* (OMIM#612555) carriers have yielded contradictory results with risk increases up to 5-fold (84,85). However, a recent meta-analysis found a 1.5-fold higher risk of CRC in *BRCA1* mutation carriers, but none in *BRCA2* mutation carriers (86). Other genes linked to CRC risk include breast and pancreatic cancer genes *ATM* (OMIM*607585) and *PALB2* (OMIM*610355), and melanoma and pancreatic cancer gene *CDKN2A* (OMIM#60719) (19,79,87). It is likely that with the routine implementation of multigene panel tests in the clinic more phenotypic overlaps will be observed between traditionally distinct cancer predisposition syndromes, resulting in the recognition of additional CRC risk genes in the coming years.

Common CRC risk loci

The 'common-disease common-variant' model attributes some genetic susceptibility of common diseases to the inheritance of many risk alleles common in the population, which individually confer low-risk increases, but combined they can predispose an individual to a substantial risk (88). For CRC predisposition, approximately 100 common low-risk loci have been identified in genome-wide association studies (GWASs) to date, and validated risk loci include chromosome 8q24.21, 10q22.3, 12p13.13 and 15q13.3, associated with a 1.04- to 1.23-fold increased risk (89–91). Some of the genes located in these regions have already been linked to CRC aetiology, such as components of the transforming growth factor beta (TGF-β) signalling pathway, including *GREM1*, *SMAD3*, *SMAD7*, *SMAD9*, *BMP2*, *BMP4* and *RHPN2* (92). The TGF-β signalling pathway is implicated in tumorigenesis due to its biological role in cell proliferation, differentiation, migration and apoptosis and the presence of a transcriptional signature characteristic of TGF-β activation in a large group of colorectal tumours with a poor prognosis (consensus molecular classification (CMS) 4 subgroup) (93,94). In addition, rare pathogenic germline mutations in the TGF-β genes *SMAD4*, *BMPR1A* and *GREM1* are responsible for hereditary polyposis syndromes. The combined risk of inheriting multiple low-risk alleles can be estimated using risk prediction models, as so-called polygenic risk scores (PRS). More research has been conducted on PRS in individual clinical management for breast cancer; however, recent studies have hinted at a possible implementation in the clinic for familial CRC patients (89,90,95). Enrichment of low-risk loci has also been observed in familial CRC research (96). The combined contribution of low-risk variants to the familial risk is estimated to be up to 15% (90,91,97,98). These studies also indicate that more research is needed to identify additional uncommon susceptibility alleles through expanded sample sizes and increased ethnical diversity. Furthermore, research has mainly focussed on the identification of low-risk single nucleotide variants (SNVs), while the characterisation of common gross chromosomal structural variations and their influence on cancer risks remains largely unexplored (99,100).

Candidate variants for CRC

Comprehensive genetic screens were intensively applied in the last decade to discover novel genetic causes of CRCs and were able to identify some promising candidates; however, it is likely that additional culprits will be identified in the future.

Monogenic inheritance

Up to half of the missing heritability of CRC might be caused by still elusive high-risk monogenic variants (101). NGS approaches have specifically boosted the identification of candidate genes in smaller families or families suspected of lower penetrance mutations, in which genetic linkage analysis failed to identify causal loci. To date, multiple candidate genes for hereditary CRC and polyposis

have been proposed, including genes encoding proteins involved in DNA repair, DNA damage response and Wnt signalling (11). The most promising candidate genes, identified in two or more studies, are *GALTN12* (11 families), *BUB1* (6 families), *RPS20* (3 families) and *AXIN2* (3 families) (102–112). Additional candidate genes include *FAN1* (3 families), *LRP6* (2 family), *SETD6* (1 family) and *MCM9* (1 family), among others, identified in single whole-exome sequencing studies (113–116). Furthermore, a combination of homozygosity mapping and sequencing strategies identified *MIA3* as a novel candidate gene (117). The problem with the identification of novel cancer predisposition genes using family-based genetic screening approaches is the reproducibility in validation cohorts. A systematic review of eight candidate genes was only able to validate variants in one gene (*RPS20*), but not in the other genes (*FANCM*, *FAN1*, *TP53*, *BUB1*, *BUB3*, *LRP6* and *PTPN12*) (106). The study by Chubb *et al.*, where 16% of familial CRC cases could be attributed to high-risk mutations, found that the majority of the variants were identified in the most prevalent CRC-associated genes such as the MMR genes and *APC* (118). These studies refute the existence of novel high-risk susceptibility genes that are recurrently mutated in unrelated populations and they point towards variants which will only explain the hereditary predisposition for individuals within few or single (extended) families.

Polygenic inheritance

The co-inheritance of multiple rare low-to-moderate risk alleles likely cause a substantial proportion of hereditary cancers where high-risk variants cannot explain familial segregation of disease (88). However, these variants are difficult to identify with the current research approaches focussing on either common low-risk variants or rare high-risk variants in the familial context. To date, few rare moderate-risk variants have been identified in CRC-associated genes (16,80). The increased risk for monoallelic carriers of pathogenic *MUTYH* variants is still under debate, although available data suggests an at most 3-fold increased risk of CRC (119). The *APC* c.I1307K variant, carried by ~6% of Ashkenazi Jews, has been shown to increase both the risk of colorectal adenomas and cancer by approximately 2-fold (120). Furthermore, a digenic inheritance model, were the genotype at two loci explains the phenotype, was proposed once in an early-onset CRC patient carrying heterozygous germline variants in two genes involved in the repair of oxidised DNA lesions (*MUTYH* and *OGG1*) (121,122). Additionally, the co-occurrence of *MSH6* and *MUTYH* within CRC patients points towards a synergistic effect of monoallelic variants in both genes (123,124).

Missed genetics

The presently implemented diagnostic screening approaches are missing and misclassifying some pathogenic genetic alterations. Broader diagnostic screening approaches and technological improvements of NGS platforms are likely to aid in the identification of these unexplored causal variants; however, to what extent these variants can explain the missing heritability of CRC is currently unknown.

Somatic mosaicism

Part of the heritability of CRC can already be explained by somatic mosaicism, the phenomenon in which variants are only present in a fraction of an organism's cells. The first report of somatic mosaicism in CRC patients was two decades ago in apparent *de novo* FAP patients (125). Current estimates show that 20% of *de novo* FAP cases have somatic *APC* mosaicism (126,127). Depending on the embryonic stage at which a somatic mutation occurs, i.e. whether it affects gonadal tissues, the mutation can be transmitted to the offspring (128,129). Screening strategies which include DNA derived from multiple tissues are more sensitive and specific than leucocyte DNA analysis alone in the identification of mosaic *APC* mutations (128). Little research has been performed on the prevalence and penetrance of mosaic mutations in other CRC-associated genes. However, more mosaic patients are likely to be identified in the future, as tumour-focussed predisposition screening approaches will probably be more widely applied in clinical diagnostics. Thus far, two families have been described with a mosaic MMR mutation, one with *MSH2* and the other with *MLH1* mosaicism (130,131).

Epigenetic inheritance

Constitutional epimutations are epigenetic defects that cause disruption of gene expression. These epimutations are present in all somatic tissues and can occur due to aberrant promoter hypermethylation in the absence or presence of underlying genetic changes, known as primary and secondary epimutations, respectively. The existence of transmission of primary epimutations, i.e. transgenerational epigenetic inheritance, is still under debate in humans, due to intergenerational DNA methylation erasure (132). The inheritance of an epimutation was first described in a LS family with constitutional *MSH2* promoter hypermethylation (133). In this family and multiple additional patients *EPCAM* deletions were subsequently identified, which cause the atypical epigenetic state of *MSH2*, classifying them as secondary epimutations (21,22). In addition, constitutional epimutations of *MLH1* explain 2–3% of the suspected LS patients with MLH1-deficient tumours (134). Most of these mutations are considered primary epimutations (seemingly occurring independently of any DNA sequence changes), and cause a severe LS phenotype in carriers, but do not appear to be transmitted to the offspring (135,136). There are three reports suggesting non-Mendelian transmission of a primary epimutation, indicating that the inheritance of a primary epimutation might be possible in extremely rare cases (135–137). On the other hand, in most families suspected of a heritable constitutional *MLH1* epimutation an underlying genetic defect has been identified, and therefore undiscovered genetic changes may still be present in the cases suspected of transgenerational epigenetic inheritance (135,138). Next to *MLH1* other tumour suppressor genes are frequently somatically silenced by promoter hypermethylation in tumours, and constitutional epimutations of these genes, such as *BRCA1* and *RB1*, might also predispose to cancer (139,140). It is conceivable that CRC predisposition in some patients may be caused by primary or secondary constitutional epimutations, and this warrants further exploration.

Technological limitations

For several decades, the gold standard for diagnosing CRC syndromes was Sanger sequencing of the genes matching the clinical phenotype of the patient under investigation. In recent years, short-read NGS technologies (e.g. Illumina and Ion Torrent) have been implemented in routine germline diagnostic testing, improving the detection of pathogenic mutations by screening several genes simultaneously (141). These approaches focus on exonic genomic regions and it is therefore conceivable that uncharacterised pathogenic variants are missed in regions outside the coding sequence, such as intergenic region, promoter sequences, intronic regions or

untranslated regions (UTRs). Although, whole gene capture of the MMR genes in suspected LS patients (negative for MMR mutations after diagnostic screening) was not able to identify any meaningful variants, as none of the numerous intronic variants was functionally tested (53). The strengths of the short-read sequencing technologies lie in the identification of single nucleotide variants (SNV) and small insertion/deletions (indels), while structural variant identification (large duplications, large deletions, translocations and inversions) remains challenging but not impossible. This can be illustrated by the study from Franch-Expósito *et al.* where a 400 kb duplication was identified by whole-exome sequencing analysis (142). Also, paralogous sequence variant discrimination can be complicated with the current sequencing analyses, e.g. identification of pathogenic variants in *PMS2*, one of the genes associated with LS, is hampered by multiple pseudogenes (143). Furthermore, mutation identification is troublesome in high GC-content regions and repetitive sequences, i.e. short tandem repeat (STR) and variable number tandem repeat (VNTR). In the future, the routine implementation of novel sequencing technologies such as long-read sequencing [e.g. PacBio SMRT sequencing and Oxford Nanopore Technology (ONT)] will likely overcome some of these limitations (144). In addition, an alternative for DNA diagnostic strategies is transcriptome sequencing [RNA sequencing (RNA-seq)], which screens for expressed germline variants, splice-site variants, chimeric gene fusions and, possibly, (epi)mutations affecting gene expression (145). Nevertheless, a combination of these approaches will most likely be necessary to elucidate the remaining heritability of CRC.

Variants of unknown significance

A possible drawback of sequencing larger regions of the genome is the identification of more variants of unknown significance (VUS). The International Agency for Research on Cancer (IARC) classification system divides genetic variants into five classes: class 5 (pathogenic), class 4 (likely pathogenic), class 3 (uncertain), class 2 (likely not pathogenic) and class 1 (not pathogenic). For instance, the large majority of the missense variants identified in the MMR genes during LS screening, lack sufficient corresponding clinical data and are therefore classified as VUS (class 3) (146). To address this issue high-throughput functional assays have been developed for the diagnostic assessment of variants identified in suspected LS patients, examples are RNA splicing assay, cell-free MMR assays (CIMBRA), protein heterodimer formation assays (Yeast 2 hybrid assay) and protein localisation experiments (147–150). More recently, the functional assessment of splice-site variants was performed on RNA isolated from formalin-fixed paraffin-embedded tissue (151). Due to the complexity of the data and the large clinical consequences for patients, variant classification will probably need to be performed by a panel of experts based on an integrated assessment of the different functional assays, *in silico* prediction tools, clinical and epidemiological data. The next challenge will be the development and implementation of similar functional assessment protocols used for LS for VUS identified in other CRC-associated genes.

Non-genetic risk factors

In developed countries, there is a steady decline in the incidence of CRC, largely attributed to the widespread introduction of screening programmes, but since the past two to three decades the incidence of early-onset disease (diagnosed <50 years) is rapidly increasing, representing a challenge for the healthcare system as this group of patients consists mostly of sporadic cases with late-stage disease (152,153). As CRC is 3-fold more frequently diagnosed in more developed countries compared to less developed countries, this points towards the influence of dietary patterns, obesity and lifestyle factors on cancer risk (1). The World Cancer Research Fund/American Institute for Cancer Research reports strong evidence for an increased risk of CRC for the consumption of red and processed meat, alcohol intake, obesity and height, while physical activity, consumption of wholegrains, fibres, dairy products and calcium supplements shows strong evidence for a risk decrease (154). Greater adherence to these recommendations, including non-smoking, has been shown to reduce CRC risk (155,156). Mounting evidence suggests that the effect of dietary patterns on CRC risk is, at least partially, mediated by the gut microbiota composition, i.e. microbial diversity and colonisation of specific bacterial strains (157,158). The correlation of CRC with the gut microbiome is a relatively recent discovery, and the effect of the microbiome, but also the use of probiotics and antibiotics, on CRC risk is still intensely researched. Additional, non-environmental, risk factors for CRC include age, male gender and chronic intestinal inflammation, among others (1,159). Patients with a personal history of inflammatory bowel disease (IBD), a term used to describe ulcerative colitis and Crohn disease, have an increased risk for CRC, especially those with uncontrolled disease with severe, longstanding inflammation (160–162). Genetic variations have been identified that are associated with IBD, and these loci may also underlie some of the familial aggregation of CRC (163).

An explanation for part of the missing heritability of CRC might lie in the overestimation of the actual heritability of disease. Shared non-genetic factors amongst close relatives can result in the belief that multiple cancers within the same family are the result of an underlying genetic defect, while, in fact, the cancer predisposition may be explained by exposure to shared environmental factors (159). However, shared environmental factors in adulthood alone do not primarily attribute to familial clustering, as CRC risk estimates between spouses is less than 1.1 (164). Furthermore, in families with exactly two first-degree relatives affected by a given disease there is only a moderate probability of this being due to the segregation of genetic variants (165). In addition, the familial risk might be overestimated in relatives of suspected LS patients. In approximately half of these patients, the MSI-H phenotype can be explained by two pathogenic somatic mutations, highlighting an additional mechanism resulting into MMR-deficiency next to germline mutations, constitutional epimutations and somatic hypermethylation of the *MLH1* promoter (166,167). Risk estimates in these families should, therefore, be based on family history, as an underlying pathogenic germline variant outside the known LS-associated genes cannot be excluded, such as germline mutations in *POLE*, *POLD1* or *MUTYH* (52,168).

Conclusions

Only 25% of all familial CRCs can be explained by rare pathogenic mutations in one of the CRC-associated genes (12,169). The combined effect of common low-risk loci is currently estimated to be up to 15% of the familial risk, but this is likely to increase once less common variants and structural aberrations are included in large-scale analyses (90). Taken together, more than half of the CRC heritability still remains unexplained. Still undiscovered monogenic cancer predisposition genes may explain up to half of the missing heritability of CRC. Cancer risk in additional familial cases might result from the polygenic inheritance of multiple uncharacterised moderate-risk variants. Furthermore, the current diagnostic screening approaches are likely to be missing some of the pathogenic

genetic alterations, such as somatic mosaicism and gross chromosomal variations. Variants of unknown significance (VUS) probably explain some of the remaining CRC heritability. In other cases, the seeming heritability might not be genetic, but may in fact, result from shared environmental risk factors. Future research will likely focus on the integrated analysis of an individual's environmental exposures with his or her genetic background to study gene-environment interactions.

Hopefully, in the coming years broader diagnostic screening approaches and technological improvements can partially elucidated the missing heritability of CRC. Increasing evidence shows phenotypic overlap between CRC and non-CRC syndromes, supporting multigene panel testing over strict germline testing of only syndrome-specific genes, which has the potential to miss a sizeable fraction of the hereditary cancer patients (19,170,171). Lastly, the implementation of mutational signature and TMB analyses in molecular pathology is impending, as its importance as a biomarker for response to novel immunotherapies is becoming more apparent (172,173). A substantial proportion of the genes and candidate genes associated with familial CRC encode proteins involved in DNA repair mechanisms, resulting in distinct somatic mutation patterns in these tumours. Therefore, tumour-focussed germline predisposition screening approaches will likely play an increasing role in oncogenetic diagnostics by aiding in the identification of known and potentially novel hereditary cancer syndromes.

Funding

This work was supported by Dutch Digestive Foundation (grant number MLDS FP13-13 to T.v.W.) and the Fight Colorectal Cancer-Michael's Mission-AACR Fellowship in Young Onset, Late-Stage Colorectal Cancer Research 2015 (grant number 15-40-1645-DEMI to N.F.C.C.d.M.).

Acknowledgements

This article is based upon work from COST Action CA17118, supported by COST (European Cooperation in Science and Technology).
Conflict of interest statement: None declared.

References

1. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A. and Jemal, A. (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA*, 68, 394–424.
2. Johns, L. E. and Houlston, R. S. (2001) A systematic review and meta-analysis of familial colorectal cancer risk. *Am. J. Gastroenterol.*, 96, 2992–3003.
3. Frank, C., Fallah, M., Sundquist, J., Hemminki, A. and Hemminki, K. (2015) Population landscape of familial cancer. *Sci. Rep.*, 5, 12891.
4. Lichtenstein, P., Holm, N. V., Verkasalo, P. K., Iliadou, A., Kaprio, J., Koskenvuo, M., Pukkala, E., Skytthe, A. and Hemminki, K. (2000) Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl. J. Med.*, 343, 78–85.
5. Kerber, R. A., Neklason, D. W., Samowitz, W. S. and Burt, R. W. (2005) Frequency of familial colon cancer and hereditary nonpolyposis colorectal cancer (Lynch syndrome) in a large population database. *Fam. Cancer*, 4, 239–244.
6. Mucci, L. A., Hjelmborg, J. B., Harris, J. R., *et al.* (2016) Familial risk and heritability of cancer among twins in nordic countries. *JAMA*, 315, 68–76.
7. Tuohy, T. M., Rowe, K. G., Mineau, G. P., Pimentel, R., Burt, R. W. and Samadder, N. J. (2014) Risk of colorectal cancer and adenomas in the families of patients with adenomas: a population-based study in Utah. *Cancer*, 120, 35–42.
8. Alexandrov, L. B., Kim, J., Haradhvala, N. J., *et al.*; Network, o.b.o.t.P.M.S.W.G.a.t.I.T.P.-C.A.o.W.G. (2018) The repertoire of mutational signatures in human cancer. *bioRxiv* 322859.
9. Campbell, B. B., Light, N., Fabrizio, D., *et al.* (2017) Comprehensive analysis of hypermutation in human cancer. *Cell*, 171, 1042–1056.e1010.
10. Grolleman, J. E., de Voer, R. M., Elsayed, F. A., *et al.* (2019) Mutational signature analysis reveals NTHL1 deficiency to cause a multi-tumor phenotype. *Cancer Cell*, 35, 256–266.e255.
11. Valle, L., de Voer, R. M., Goldberg, Y., *et al.* (2019) Update on genetic predisposition to colorectal cancer and polyposis. *Mol. Aspects Med*, pii: S0098-2997(19)30004-4. doi:10.1016/j.mam.2019.03.001.
12. Chubb, D., Broderick, P., Frampton, M., *et al.* (2015) Genetic diagnosis of high-penetrance susceptibility for colorectal cancer (CRC) is achievable for a high proportion of familial CRC by exome sequencing. *J. Clin. Oncol.*, 33, 426–432.
13. Mongin, C., Coulet, F., Lefevre, J. H., *et al.* (2012) Unexplained polyposis: a challenge for geneticists, pathologists and gastroenterologists. *Clin. Genet.*, 81, 38–46.
14. Buchanan, D. D., Rosty, C., Clendenning, M., Spurdle, A. B. and Win, A. K. (2014) Clinical problems of colorectal cancer and endometrial cancer cases with unknown cause of tumor mismatch repair deficiency (suspected Lynch syndrome). *Appl. Clin. Genet.*, 7, 183–193.
15. Rodriguez-Soler, M., Perez-Carbonell, L., Guarinos, C., *et al.* (2013) Risk of cancer in cases of suspected lynch syndrome without germline mutation. *Gastroenterology*, 144, 926–932.e921; quiz e913-924.
16. Jasperson, K. W., Tuohy, T. M., Neklason, D. W. and Burt, R. W. (2010) Hereditary and familial colon cancer. *Gastroenterology*, 138, 2044–2058.
17. Martin-Morales, L., Rofes, P., Diaz-Rubio, E., *et al.* (2018) Novel genetic mutations detected by multigene panel are associated with hereditary colorectal cancer predisposition. *PLoS One*, 13, e0203885.
18. Hampel, H., Frankel, W. L., Martin, E., *et al.* (2008) Feasibility of screening for Lynch syndrome among patients with colorectal cancer. *J. Clin. Oncol.*, 26, 5783–5788.
19. Yurgelun, M. B., Kulke, M. H., Fuchs, C. S., *et al.* (2017) Cancer susceptibility gene mutations in individuals with colorectal cancer. *J. Clin. Oncol.*, 35, 1086–1095.
20. Vasen, H. F., Offerhaus, G. J., den Hartog Jager, F. C., Menko, F. H., Nagengast, F. M., Griffioen, G., van Hogezand, R. B. and Heintz, A. P. (1990) The tumour spectrum in hereditary non-polyposis colorectal cancer: a study of 24 kindreds in the Netherlands. *Int. J. Cancer*, 46, 31–34.
21. Niessen, R. C., Hofstra, R. M., Westers, H., *et al.* (2009) Germline hypermethylation of MLH1 and EPCAM deletions are a frequent cause of Lynch syndrome. *Genes Chromosomes Cancer*, 48, 737–744.
22. Ligtenberg, M. J., Kuiper, R. P., Chan, T. L., *et al.* (2009) Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3' exons of TACSTD1. *Nat. Genet.*, 41, 112–117.
23. Gazzoli, I., Loda, M., Garber, J., Syngal, S. and Kolodner, R. D. (2002) A hereditary nonpolyposis colorectal carcinoma case associated with hypermethylation of the MLH1 gene in normal tissue and loss of heterozygosity of the unmethylated allele in the resulting microsatellite instability-high tumor. *Cancer Res.*, 62, 3925–3928.
24. Wimmer, K., Kratz, C. P., Vasen, H. F., *et al.* (2014) Diagnostic criteria for constitutional mismatch repair deficiency syndrome: suggestions of the European consortium 'care for CMMRD' (C4CMMRD). *J. Med. Genet.*, 51, 355–365.
25. Jiricny, J. (2006) The multifaceted mismatch-repair system. *Nat. Rev. Mol. Cell Biol.*, 7, 335–346.
26. Boland, C. R., Koi, M., Chang, D. K. and Carethers, J. M. (2008) The biochemical basis of microsatellite instability and abnormal immunohistochemistry and clinical behavior in Lynch syndrome: from bench to bedside. *Fam. Cancer*, 7, 41–52.
27. Moller, P., Seppala, T. T., Bernstein, I., *et al.* (2018) 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*, 67, 1306–1316.
28. Shia, J. (2008) Immunohistochemistry versus microsatellite instability testing for screening colorectal cancer patients at risk for heredi-

tary nonpolyposis colorectal cancer syndrome. Part I. The utility of immunohistochemistry. *J. Mol. Diagn.*, 10, 293–300.
29. Cunningham, J. M., Christensen, E. R., Tester, D. J., Kim, C. Y., Roche, P. C., Burgart, L. J. and Thibodeau, S. N. (1998) Hypermethylation of the hMLH1 promoter in colon cancer with microsatellite instability. *Cancer Res.*, 58, 3455–3460.
30. Haradhvala, N. J., Polak, P., Stojanov, P., *et al*. (2016) Mutational strand asymmetries in cancer genomes reveal mechanisms of DNA damage and repair. *Cell*, 164, 538–549.
31. Bisgaard, M. L., Fenger, K., Bulow, S., Niebuhr, E. and Mohr, J. (1994) Familial Adenomatous Polyposis (FAP): frequency, penetrance, and mutation rate. *Hum. Mutat.*, 3, 121–125.
32. Aretz, S., Uhlhaas, S., Caspari, R., Mangold, E., Pagenstecher, C., Propping, P. and Friedl, W. (2004) Frequency and parental origin of de novo APC mutations in familial adenomatous polyposis. *Eur. J. Hum. Genet.*, 12, 52–58.
33. Nielsen, M., Hes, F. J., Nagengast, F. M., *et al*. (2007) Germline mutations in APC and MUTYH are responsible for the majority of families with attenuated familial adenomatous polyposis. *Clin. Genet.*, 71, 427–433.
34. Nieuwenhuis, M. H. and Vasen, H. F. (2007) Correlations between mutation site in APC and phenotype of Familial Adenomatous Polyposis (FAP): a review of the literature. *Crit. Rev. Oncol-Hematol.*, 61, 153–161.
35. Fearnhead, N. S., Britton, M. P. and Bodmer, W. F. (2001) The ABC of APC. *Hum. Mol. Genet.*, 10, 721–733.
36. Rowan, A. J., Lamlum, H., Ilyas, M., Wheeler, J., Straub, J., Papadopoulou, A., Bicknell, D., Bodmer, W. F. and Tomlinson, I. P. (2000) APC mutations in sporadic colorectal tumors: a mutational "hotspot" and interdependence of the "two hits". *Proc. Natl. Acad. Sci. U.S.A.*, 97, 3352–3357.
37. Ma, H., Brosens, L. A.A., Offerhaus, G. J.A., Giardiello, F. M., de Leng, W. W.J. and Montgomery, E. A. (2018) Pathology and genetics of hereditary colorectal cancer. *Pathology*, 50, 49–59.
38. Farrington, S. M., Tenesa, A., Barnetson, R., Wiltshire, A., Prendergast, J., Porteous, M., Campbell, H. and Dunlop, M. G. (2005) Germline susceptibility to colorectal cancer due to base-excision repair gene defects. *Am. J. Hum. Genet.*, 77, 112–119.
39. Vogt, S., Jones, N., Christian, D., *et al*. (2009) Expanded extracolonic tumor spectrum in MUTYH-associated polyposis. *Gastroenterology*, 137, 1976–1985.e1971-1910.
40. Al-Tassan, N., Chmiel, N. H., Maynard, J., *et al*. (2002) Inherited variants of MYH associated with somatic G:C-->T:A mutations in colorectal tumors. *Nat. Genet.*, 30, 227–232.
41. Nielsen, M., Joerink-van de Beld, M. C., Jones, N., Vogt, S., Tops, C. M., Vasen, H. F., Sampson, J. R., Aretz, S. and Hes, F. J. (2009) Analysis of MUTYH genotypes and colorectal phenotypes in patients with MUTYH-associated polyposis. *Gastroenterology*, 136, 471–476.
42. Nakabeppu, Y., Sakumi, K., Sakamoto, K., Tsuchimoto, D., Tsuzuki, T. and Nakatsu, Y. (2006) Mutagenesis and carcinogenesis caused by the oxidation of nucleic acids. *Biol. Chem.*, 387, 373–379.
43. Oka, S. and Nakabeppu, Y. (2011) DNA glycosylase encoded by MUTYH functions as a molecular switch for programmed cell death under oxidative stress to suppress tumorigenesis. *Cancer Sci.*, 102, 677–682.
44. Viel, A., Bruselles, A., Meccia, E., *et al*. (2017) A specific mutational signature associated with DNA 8-oxoguanine persistence in MUTYH-defective colorectal cancer. *EBioMedicine*, 20, 39–49.
45. Aime, A., Coulet, F., Lefevre, J. H., Colas, C., Cervera, P., Flejou, J. F., Lascols, O., Soubrier, F. and Parc, Y. (2015) Somatic c.34G>T KRAS mutation: a new prescreening test for MUTYH-associated polyposis? *Cancer Genetics*, 208, 390–395.
46. Middeldorp, A., van Puijenbroek, M., Nielsen, M., *et al*. (2008) High frequency of copy-neutral LOH in MUTYH-associated polyposis carcinomas. *J. Pathol.*, 216, 25–31.
47. Weren, R. D., Ligtenberg, M. J., Geurts van Kessel, A., De Voer, R. M., Hoogerbrugge, N. and Kuiper, R. P. (2018) NTHL1 and MUTYH polyposis syndromes: two sides of the same coin? *J. Pathol.*, 244, 135–142.
48. Weren, R. D., Ligtenberg, M. J., Kets, C. M., *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.
49. Drost, J., van Boxtel, R., Blokzijl, F., *et al*. (2017) Use of CRISPR-modified human stem cell organoids to study the origin of mutational signatures in cancer. *Science (New York, N.Y.)*, 358, 234–238.
50. Palles, C., Cazier, J. B., Howarth, K. M., *et al*. (2013) Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. *Nat. Genet.*, 45, 136–144.
51. Elsayed, F. A., Tops, C. M.J., Nielsen, M., Ruano, D., Vasen, H. F.A., Morreau, H., Hes, F. J. and van Wezel, T. (2019) Low frequency of POLD1 and POLE exonuclease domain variants in patients with multiple colorectal polyps. *Molecular Genetics & Genomic Medicine*, 7, e603.
52. Elsayed, F. A., Kets, C. M., Ruano, D., *et al*. (2015) Germline variants in POLE are associated with early onset mismatch repair deficient colorectal cancer. *Eur. J. Hum. Genet.*, 23, 1080–1084.
53. Jansen, A. M., van Wezel, T., van den Akker, B. E., *et al*. (2016) Combined mismatch repair and POLE/POLD1 defects explain unresolved suspected Lynch syndrome cancers. *Eur. J. Hum. Genet.*, 24, 1089–1092.
54. Church, D. N., Briggs, S. E., Palles, C., *et al*. (2013) DNA polymerase epsilon and delta exonuclease domain mutations in endometrial cancer. *Hum. Mol. Genet.*, 22, 2820–2828.
55. Rohlin, A., Zagoras, T., Nilsson, S., Lundstam, U., Wahlstrom, J., Hulten, L., Martinsson, T., Karlsson, G. B. and Nordling, M. (2014) A mutation in POLE predisposing to a multi-tumour phenotype. *Int. J. Oncol.*, 45, 77–81.
56. Hansen, M. F., Johansen, J., Bjornevoll, I., Sylvander, A. E., Steinsbekk, K. S., Saetrom, P., Sandvik, A. K., Drablos, F. and Sjursen, W. (2015) A novel POLE mutation associated with cancers of colon, pancreas, ovaries and small intestine. *Fam. Cancer*, 14, 437–448.
57. Nick McElhinny, S. A., Gordenin, D. A., Stith, C. M., Burgers, P. M. and Kunkel, T. A. (2008) Division of labor at the eukaryotic replication fork. *Mol. Cell*, 30, 137–144.
58. Kane, D. P. and Shcherbakova, P. V. (2014) A common cancer-associated DNA polymerase epsilon mutation causes an exceptionally strong mutator phenotype, indicating fidelity defects distinct from loss of proofreading. *Cancer Res.*, 74, 1895–1901.
59. Schreibman, I. R., Baker, M., Amos, C. and McGarrity, T. J. (2005) The hamartomatous polyposis syndromes: a clinical and molecular review. *Am. J. Gastroenterol.*, 100, 476–490.
60. Gammon, A., Jasperson, K., Kohlmann, W. and Burt, R. W. (2009) Hamartomatous polyposis syndromes. *Best Pract. Res. Clin. Gastroenterol.*, 23, 219–231.
61. Howe, J. R., Bair, J. L., Sayed, M. G., Anderson, M. E., Mitros, F. A., Petersen, G. M., Velculescu, V. E., Traverso, G. and Vogelstein, B. (2001) Germline mutations of the gene encoding bone morphogenetic protein receptor 1A in juvenile polyposis. *Nat. Genet.*, 28, 184–187.
62. Howe, J. R., Roth, S., Ringold, J. C., *et al*. (1998) Mutations in the SMAD4/DPC4 gene in juvenile polyposis. *Science (New York, N.Y.)*, 280, 1086–1088.
63. Jenne, D. E., Reimann, H., Nezu, J., Friedel, W., Loff, S., Jeschke, R., Muller, O., Back, W. and Zimmer, M. (1998) Peutz-Jeghers syndrome is caused by mutations in a novel serine threonine kinase. *Nat. Genet.*, 18, 38–43.
64. Jaeger, E., Leedham, S., Lewis, A., *et al*. (2012) Hereditary mixed polyposis syndrome is caused by a 40-kb upstream duplication that leads to increased and ectopic expression of the BMP antagonist GREM1. *Nat. Genet.*, 44, 699–703.
65. Gala, M. K., Mizukami, Y., Le, L. P., Moriichi, K., Austin, T., Yamamoto, M., Lauwers, G. Y., Bardeesy, N. and Chung, D. C. (2014) Germline mutations in oncogene-induced senescence pathways are associated with multiple sessile serrated adenomas. *Gastroenterology*, 146, 520–529.
66. Quintana, I., Mejias-Luque, R., Terradas, M., *et al*. (2018) Evidence suggests that germline RNF43 mutations are a rare cause of serrated polyposis. *Gut*, 67, 2230–2232.
67. Adam, R., Spier, I., Zhao, B., *et al*. (2016) Exome sequencing identifies biallelic MSH3 germline mutations as a recessive subtype of colorectal adenomatous polyposis. *Am. J. Hum. Genet.*, 99, 337–351.
68. Tseng-Rogenski, S. S., Chung, H., Wilk, M. B., Zhang, S., Iwaizumi, M. and Carethers, J. M. (2012) Oxidative stress induces nuclear-to-cytosol shift of hMSH3, a potential mechanism for EMAST in colorectal cancer cells. *PLoS One*, 7, e50616.

69. Carethers, J. M., Koi, M. and Tseng-Rogenski, S. S. (2015) EMAST is a form of microsatellite instability that is initiated by inflammation and modulates colorectal cancer progression. *Genes*, 6, 185–205.
70. Eng, C. (2003) PTEN: one gene, many syndromes. *Hum. Mutat.*, 22, 183–198.
71. Marsh, D. J., Kum, J. B., Lunetta, K. L., *et al.* (1999) PTEN mutation spectrum and genotype-phenotype correlations in Bannayan-Riley-Ruvalcaba syndrome suggest a single entity with Cowden syndrome. *Hum. Mol. Genet.*, 8, 1461–1472.
72. Pilarski, R. (2009) Cowden syndrome: a critical review of the clinical literature. *J. Genet. Couns.*, 18, 13–27.
73. Chompret, A., Brugieres, L., Ronsin, M., *et al.* (2000) P53 germline mutations in childhood cancers and cancer risk for carrier individuals. *Br. J. Cancer*, 82, 1932–1937.
74. Bougeard, G., Renaux-Petel, M., Flaman, J. M., *et al.* (2015) Revisiting li-fraumeni syndrome from TP53 mutation carriers. *J. Clin. Oncol.*, 33, 2345–2352.
75. Rengifo-Cam, W., Shepherd, H. M., Jasperson, K. W., Samadder, N. J., Samowitz, W., Tripp, S. R., Schiffman, J. D. and Kohlmann, W. (2018) Colon pathology characteristics in Li-Fraumeni syndrome. *Clin. Gastroenterol. Hepatol.*, 16, 140–141.
76. Wong, P., Verselis, S. J., Garber, J. E., Schneider, K., DiGianni, L., Stockwell, D. H., Li, F. P. and Syngal, S. (2006) Prevalence of early onset colorectal cancer in 397 patients with classic Li-Fraumeni syndrome. *Gastroenterology*, 130, 73–79.
77. Yurgelun, M. B., Masciari, S., Joshi, V. A., *et al.* (2015) Germline TP53 mutations in patients with early-onset colorectal cancer in the colon cancer family registry. *JAMA Oncology*, 1, 214–221.
78. Stoffel, E. M., Koeppe, E., Everett, J., Ulintz, P., Kiel, M., Osborne, J., Williams, L., Hanson, K., Gruber, S. B. and Rozek, L. S. (2018) Germline genetic features of young individuals with colorectal cancer. *Gastroenterology*, 154, 897–905.e891.
79. Pearlman, R., Frankel, W. L., Swanson, B., *et al.* (2017) Prevalence and spectrum of germline cancer susceptibility gene mutations among patients with early-onset colorectal cancer. *JAMA Oncology*, 3, 464–471.
80. Ma, X., Zhang, B. and Zheng, W. (2014) Genetic variants associated with colorectal cancer risk: comprehensive research synopsis, meta-analysis, and epidemiological evidence. *Gut*, 63, 326–336.
81. Cybulski, C., Gorski, B., Huzarski, T., *et al.* (2004) CHEK2 is a multiorgan cancer susceptibility gene. *Am. J. Hum. Genet.*, 75, 1131–1135.
82. Meijers-Heijboer, H., van den Ouweland, A., Klijn, J., *et al.* (2002) Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. *Nat. Genet.*, 31, 55–59.
83. Meijers-Heijboer, H., Wijnen, J., Vasen, H., *et al.* (2003) The CHEK2 1100delC mutation identifies families with a hereditary breast and colorectal cancer phenotype. *Am. J. Hum. Genet.*, 72, 1308–1314.
84. Phelan, C. M., Iqbal, J., Lynch, H. T., *et al.* (2014) Incidence of colorectal cancer in BRCA1 and BRCA2 mutation carriers: results from a follow-up study. *Br. J. Cancer*, 110, 530–534.
85. Risch, H. A., McLaughlin, J. R., Cole, D. E., *et al.* (2001) Prevalence and penetrance of germline BRCA1 and BRCA2 mutations in a population series of 649 women with ovarian cancer. *Am. J. Hum. Genet.*, 68, 700–710.
86. Oh, M., McBride, A., Yun, S., Bhattacharjee, S., Slack, M., Martin, J. R., Jeter, J. and Abraham, I. (2018) BRCA1 and BRCA2 gene mutations and colorectal cancer risk: systematic review and meta-analysis. *J. Natl. Cancer Inst.*, 110, 1178–1189.
87. AlDubayan, S. H., Giannakis, M., Moore, N. D., *et al.* (2018) Inherited DNA-repair defects in colorectal cancer. *Am. J. Hum. Genet.*, 102, 401–414.
88. Fletcher, O. and Houlston, R. S. (2010) Architecture of inherited susceptibility to common cancer. *Nat. Rev. Cancer*, 10, 353–361.
89. Huyghe, J. R., Bien, S. A., Harrison, T. A., *et al.* (2019) Discovery of common and rare genetic risk variants for colorectal cancer. *Nat. Genet.*, 51, 76–87.
90. Schmit, S. L., Edlund, C. K., Schumacher, F. R., *et al.* (2019) Novel common genetic susceptibility loci for colorectal cancer. *J. Natl. Cancer Inst.*, 111, 146–157.
91. Law, P. J., Timofeeva, M., Fernandez-Rozadilla, C., *et al.* (2019) Association analyses identify 31 new risk loci for colorectal cancer susceptibility. *Nat. Commun.*, 10, 2154.
92. Tenesa, A. and Dunlop, M. G. (2009) New insights into the aetiology of colorectal cancer from genome-wide association studies. *Nat. Rev. Genet.*, 10, 353–358.
93. David, C. J. and Massague, J. (2018) Contextual determinants of TGFbeta action in development, immunity and cancer. *Nat. Rev. Mol. Cell Biol.*, 19, 419–435.
94. Guinney, J., Dienstmann, R., Wang, X., *et al.* (2015) The consensus molecular subtypes of colorectal cancer. *Nat. Med.*, 21, 1350–1356.
95. Michailidou, K., Lindstrom, S., Dennis, J., *et al.* (2017) Association analysis identifies 65 new breast cancer risk loci. *Nature*, 551, 92–94.
96. Middeldorp, A., Jagmohan-Changur, S., van Eijk, R., *et al.* (2009) Enrichment of low penetrance susceptibility loci in a Dutch familial colorectal cancer cohort. *Cancer Epidemiol. Biomarkers Prev.*, 18, 3062–3067.
97. Dunlop, M. G., Tenesa, A., Farrington, S. M., *et al.* (2013) Cumulative impact of common genetic variants and other risk factors on colorectal cancer risk in 42,103 individuals. *Gut*, 62, 871–881.
98. Jiao, S., Peters, U., Berndt, S., *et al.* (2014) Estimating the heritability of colorectal cancer. *Hum. Mol. Genet.*, 23, 3898–3905.
99. Kuiper, R. P., Ligtenberg, M. J., Hoogerbrugge, N. and Geurts van Kessel, A. (2010) Germline copy number variation and cancer risk. *Curr. Opin. Genet. Dev.*, 20, 282–289.
100. Audano, P. A., Sulovari, A., Graves-Lindsay, T. A., *et al.* (2019) Characterizing the major structural variant alleles of the human genome. *Cell*, 176, 663–675.e619.
101. Win, A. K., Jenkins, M. A., Dowty, J. G., *et al.* (2017) Prevalence and penetrance of major genes and polygenes for colorectal cancer. *Cancer Epidemiol. Biomarkers Prev.*, 26, 404–412.
102. Evans, D. R., Venkitachalam, S., Revoredo, L., *et al.* (2018) Evidence for GALNT12 as a moderate penetrance gene for colorectal cancer. *Hum. Mutat.*, 39, 1092–1101.
103. Guda, K., Moinova, H., He, J., *et al.* (2009) Inactivating germ-line and somatic mutations in polypeptide N-acetylgalactosaminyltransferase 12 in human colon cancers. *Proc. Natl. Acad. Sci. U.S.A.*, 106, 12921–12925.
104. Clarke, E., Green, R. C., Green, J. S., Mahoney, K., Parfrey, P. S., Younghusband, H. B. and Woods, M. O. (2012) Inherited deleterious variants in GALNT12 are associated with CRC susceptibility. *Hum. Mutat.*, 33, 1056–1058.
105. Nieminen, T. T., O'Donohue, M. F., 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.e595.
106. Broderick, P., Dobbins, S. E., Chubb, D., Kinnersley, B., Dunlop, M. G., Tomlinson, I. and Houlston, R. S. (2017) Validation of recently proposed colorectal cancer susceptibility gene variants in an analysis of families and patients-a systematic review. *Gastroenterology*, 152, 75–77.e74.
107. Lammi, L., Arte, S., Somer, M., Jarvinen, H., Lahermo, P., Thesleff, I., Pirinen, S. and Nieminen, P. (2004) Mutations in AXIN2 cause familial tooth agenesis and predispose to colorectal cancer. *Am. J. Hum. Genet.*, 74, 1043–1050.
108. Marvin, M. L., Mazzoni, S. M., Herron, C. M., Edwards, S., Gruber, S. B. and Petty, E. M. (2011) AXIN2-associated autosomal dominant ectodermal dysplasia and neoplastic syndrome. *Am. J. Med. Genet. A*, 155a, 898–902.
109. Rivera, B., Perea, J., Sanchez, E., Villapun, M., Sanchez-Tome, E., Mercadillo, F., Robledo, M., Benitez, J. and Urioste, M. (2014) A novel AXIN2 germline variant associated with attenuated FAP without signs of oligondontia or ectodermal dysplasia. *Eur. J. Hum. Genet.*, 22, 423–426.
110. de Voer, R. M., Geurts van Kessel, A., Weren, R. D., *et al.* (2013) Germline mutations in the spindle assembly checkpoint genes BUB1 and BUB3 are risk factors for colorectal cancer. *Gastroenterology*, 145, 544–547.

111. Mur, P., De Voer, R. M., Olivera-Salguero, R., *et al*. (2018) Germline mutations in the spindle assembly checkpoint genes BUB1 and BUB3 are infrequent in familial colorectal cancer and polyposis. *Mol. Cancer*, 17, 23.
112. Zhang, J. X., Fu, L., de Voer, R. M., *et al*. (2015) Candidate colorectal cancer predisposing gene variants in Chinese early-onset and familial cases. *World J. Gastroenterol.*, 21, 4136–4149.
113. Martin-Morales, L., Feldman, M., Vershinin, Z., Garre, P., Caldes, T. and Levy, D. (2017) SETD6 dominant negative mutation in familial colorectal cancer type X. *Hum. Mol. Genet.*, 26, 4481–4493.
114. Segui, N., Mina, L. B., Lazaro, C., *et al*. (2015) Germline mutations in FAN1 Cause hereditary colorectal cancer by impairing DNA repair. *Gastroenterology*, 149, 563–566.
115. de Voer, R. M., Hahn, M. M., Weren, R. D., *et al*. (2016) Identification of novel candidate genes for early-onset colorectal cancer susceptibility. *PLoS Genet.*, 12, e1005880.
116. Goldberg, Y., Halpern, N., Hubert, A., Adler, S. N., Cohen, S., Plesser-Duvdevani, M., Pappo, O., Shaag, A. and Meiner, V. (2015) Mutated MCM9 is associated with predisposition to hereditary mixed polyposis and colorectal cancer in addition to primary ovarian failure. *Cancer Genetics*, 208, 621–624.
117. Schubert, S. A., Ruano, D., Elsayed, F. A., *et al*. (2017) Evidence for genetic association between chromosome 1q loci and predisposition to colorectal neoplasia. *Br. J. Cancer*, 117, 1215–1223.
118. Chubb, D., Broderick, P., Dobbins, S. E., *et al*. (2016) Rare disruptive mutations and their contribution to the heritable risk of colorectal cancer. *Nat. Commun.*, 7, 11883.
119. Nielsen, M., Morreau, H., Vasen, H. F. and Hes, F. J. (2011) MUTYH-associated polyposis (MAP). *Crit. Rev. Oncol-Hematol.*, 79, 1–16.
120. Laken, S. J., Petersen, G. M., Gruber, S. B., *et al*. (1997) Familial colorectal cancer in Ashkenazim due to a hypermutable tract in APC. *Nat. Genet.*, 17, 79–83.
121. Morak, M., Massdorf, T., Sykora, H., Kerscher, M. and Holinski-Feder, E. (2011) First evidence for digenic inheritance in hereditary colorectal cancer by mutations in the base excision repair genes. *Eur. J. Cancer*, 47, 1046–1055.
122. Schaffer, A. A. (2013) Digenic inheritance in medical genetics. *J. Med. Genet.*, 50, 641–652.
123. Niessen, R. C., Sijmons, R. H., Ou, J., *et al*. (2006) MUTYH and the mismatch repair system: partners in crime? *Hum. Genet.*, 119, 206–211.
124. Giraldez, M. D., Balaguer, F., Caldes, T., *et al*. (2009) Association of MUTYH and MSH6 germline mutations in colorectal cancer patients. *Fam. Cancer*, 8, 525–531.
125. Farrington, S. M. and Dunlop, M. G. (1999) Mosaicism and sporadic familial adenomatous polyposis. *Am. J. Hum. Genet.*, 64, 653–658.
126. Hes, F. J., Nielsen, M., Bik, E. C., Konvalinka, D., Wijnen, J. T., Bakker, E., Vasen, H. F., Breuning, M. H. and Tops, C. M. (2008) Somatic APC mosaicism: an underestimated cause of polyposis coli. *Gut*, 57, 71–76.
127. Aretz, S., Stienen, D., Friedrichs, N., Stemmler, S., Uhlhaas, S., Rahner, N., Propping, P. and Friedl, W. (2007) Somatic APC mosaicism: a frequent cause of familial adenomatous polyposis (FAP). *Hum. Mutat.*, 28, 985–992.
128. Jansen, A. M., Crobach, S., Geurts-Giele, W. R., *et al*. (2017) Distinct patterns of somatic mosaicism in the APC gene in neoplasms from patients with unexplained adenomatous polyposis. *Gastroenterology*, 152, 546–549.e543.
129. Tuohy, T. M. and Burt, R. W. (2008) Somatic mosaicism: a cause for unexplained cases of FAP? *Gut*, 57, 10–12.
130. Sourrouille, I., Coulet, F., Lefevre, J. H., Colas, C., Eyries, M., Svrcek, M., Bardier-Dupas, A., Parc, Y. and Soubrier, F. (2013) Somatic mosaicism and double somatic hits can lead to MSI colorectal tumors. *Fam. Cancer*, 12, 27–33.
131. Geurts-Giele, W. R., Rosenberg, E. H., Rens, A. V., Leerdam, M. E.V., Dinjens, W. N. and Bleeker, F. E. (2019) Somatic mosaicism by a de novo MLH1 mutation as a cause of Lynch syndrome. *Molecular Genetics & Genomic Medicine*, 7, e699.
132. Horsthemke, B. (2018) A critical view on transgenerational epigenetic inheritance in humans. *Nat. Commun.*, 9, 2973.
133. Chan, T. L., Yuen, S. T., Kong, C. K., *et al*. (2006) Heritable germline epimutation of MSH2 in a family with hereditary nonpolyposis colorectal cancer. *Nat. Genet.*, 38, 1178–1183.
134. Peltomaki, P. (2016) Update on Lynch syndrome genomics. *Fam. Cancer*, 15, 385–393.
135. Damaso, E., Castillejo, A., Arias, M. D.M., *et al*. (2018) Primary constitutional MLH1 epimutations: a focal epigenetic event. *Br. J. Cancer*, 119, 978–987.
136. Ward, R. L., Dobbins, T., Lindor, N. M., Rapkins, R. W. and Hitchins, M. P. (2013) Identification of constitutional MLH1 epimutations and promoter variants in colorectal cancer patients from the Colon Cancer Family Registry. *Genet. Med.*, 15, 25–35.
137. Sloane, M. A., Nunez, A. C., Packham, D., Kwok, C. T., Suthers, G., Hesson, L. B. and Ward, R. L. (2015) Mosaic epigenetic inheritance as a cause of early-onset colorectal cancer. *JAMA Oncology*, 1, 953–957.
138. Leclerc, J., Flament, C., Lovecchio, T., *et al*. (2018) Diversity of genetic events associated with MLH1 promoter methylation in Lynch syndrome families with heritable constitutional epimutation. *Genet. Med.*, 20, 1589–1599.
139. Al-Moghrabi, N., Nofel, A., Al-Yousef, N., *et al*. (2014) The molecular significance of methylated BRCA1 promoter in white blood cells of cancer-free females. *BMC Cancer*, 14, 830.
140. Quinonez-Silva, G., Davalos-Salas, M., Recillas-Targa, F., Ostrosky-Wegman, P., Aranda, D. A. and Benitez-Bribiesca, L. (2016) "Monoallelic germline methylation and sequence variant in the promoter of the RB1 gene: a possible constitutive epimutation in hereditary retinoblastoma". *Clinical Epigenetics*, 8, 1.
141. Weiss, M. M., Van der Zwaag, B., Jongbloed, J. D., *et al*. (2013) Best practice guidelines for the use of next-generation sequencing applications in genome diagnostics: a national collaborative study of Dutch genome diagnostic laboratories. *Hum. Mutat.*, 34, 1313–1321.
142. Franch-Exposito, S., Esteban-Jurado, C., Garre, P., *et al*. (2018) Rare germline copy number variants in colorectal cancer predisposition characterized by exome sequencing analysis. *J. Genet. Genomics = Yi chuan Xue Bao*, 45, 41–45.
143. Vaughn, C. P., Robles, J., Swensen, J. J., Miller, C. E., Lyon, E., Mao, R., Bayrak-Toydemir, P. and Samowitz, W. S. (2010) Clinical analysis of PMS2: mutation detection and avoidance of pseudogenes. *Hum. Mutat.*, 31, 588–593.
144. Pollard, M. O., Gurdasani, D., Mentzer, A. J., Porter, T. and Sandhu, M. S. (2018) Long reads: their purpose and place. *Hum. Mol. Genet.*, 27, R234–R241.
145. Byron, S. A., Van Keuren-Jensen, K. R., Engelthaler, D. M., Carpten, J. D. and Craig, D. W. (2016) Translating RNA sequencing into clinical diagnostics: opportunities and challenges. *Nat. Rev. Genet.*, 17, 257–271.
146. Thompson, B. A., Spurdle, A. B., Plazzer, J. P., *et al*. (2014) Application of a 5-tiered scheme for standardized classification of 2,360 unique mismatch repair gene variants in the InSiGHT locus-specific database. *Nat. Genet.*, 46, 107–115.
147. Drost, M., Tiersma, Y., Thompson, B. A., *et al*. (2019) A functional assay-based procedure to classify mismatch repair gene variants in Lynch syndrome. *Genet. Med.*, 21, 1486–1496.
148. Tournier, I., Vezain, M., Martins, A., *et al*. (2008) A large fraction of unclassified variants of the mismatch repair genes MLH1 and MSH2 is associated with splicing defects. *Hum. Mutat.*, 29, 1412–1424.
149. Lutzen, A., de Wind, N., Georgijevic, D., Nielsen, F. C. and Rasmussen, L. J. (2008) Functional analysis of HNPCC-related missense mutations in MSH2. *Mutat. Res.*, 645, 44–55.
150. Rasmussen, L. J., Heinen, C. D., Royer-Pokora, B., Drost, M., Tavtigian, S., Hofstra, R. M. and de Wind, N. (2012) Pathological assessment of mismatch repair gene variants in Lynch syndrome: past, present, and future. *Hum. Mutat.*, 33, 1617–1625.

151. Jansen, A. M., van der Klift, H. M., Roos, M. A., van Eendenburg, J. D., Tops, C. M., Wijnen, J. T., Hes, F. J., Morreau, H. and van Wezel, T. (2018) RNA analysis of cancer predisposing genes in formalin-fixed paraffin-embedded tissue determines aberrant splicing. *Eur. J. Hum. Genet.*, 26, 1143–1150.
152. Mauri, G., Sartore-Bianchi, A., Russo, A. G., Marsoni, S., Bardelli, A. and Siena, S. (2019) Early-onset colorectal cancer in young individuals. *Mol. Oncol.*, 13, 109–131.
153. Bailey, C. E., Hu, C. Y., You, Y. N., Bednarski, B. K., Rodriguez-Bigas, M. A., Skibber, J. M., Cantor, S. B. and Chang, G. J. (2015) Increasing disparities in the age-related incidences of colon and rectal cancers in the United States, 1975–2010. *JAMA Surgery*, 150, 17–22.
154. World Cancer Research Fund International/American Institute for Cancer Research. Continuous Update Project Expert Report 2018. *Diet, nutrition, physical activity and colorectal cancer*. Available at dietandcancerreports.org.
155. El Kinany, K., Huybrechts, I., Kampman, E., *et al.* (2019) Concordance with the World Cancer Research Fund/American Institute for Cancer Research recommendations for cancer prevention and colorectal cancer risk in Morocco: a large, population-based case-control study. *Int. J. Cancer*, 145, 1829–1837.
156. Carr, P. R., Weigl, K., Jansen, L., Walter, V., Erben, V., Chang-Claude, J., Brenner, H. and Hoffmeister, M. (2018) Healthy lifestyle factors associated with lower risk of colorectal cancer irrespective of genetic risk. *Gastroenterology*, 155, 1805–1815.e1805.
157. Vipperla, K. and O'Keefe, S. J. (2016) Diet, microbiota, and dysbiosis: a 'recipe' for colorectal cancer. *Food & Function*, 7, 1731–1740.
158. Van Raay, T. and Allen-Vercoe, E. (2017) Microbial interactions and interventions in colorectal cancer. *Microbiol. Spectr.*, 5. doi:10.1128/microbiolspec.BAD-0004-2016
159. Haggar, F. A. and Boushey, R. P. (2009) Colorectal cancer epidemiology: incidence, mortality, survival, and risk factors. *Clinics in Colon and Rectal Surgery*, 22, 191–197.
160. Jess, T., Gamborg, M., Matzen, P., Munkholm, P. and Sorensen, T. I. (2005) Increased risk of intestinal cancer in Crohn's disease: a meta-analysis of population-based cohort studies. *Am. J. Gastroenterol.*, 100, 2724–2729.
161. Eaden, J. A., Abrams, K. R. and Mayberry, J. F. (2001) The risk of colorectal cancer in ulcerative colitis: a meta-analysis. *Gut*, 48, 526–535.
162. Andersen, N. N. and Jess, T. (2013) Has the risk of colorectal cancer in inflammatory bowel disease decreased? *World J. Gastroenterol.*, 19, 7561–7568.
163. McGovern, D. P., Kugathasan, S. and Cho, J. H. (2015) Genetics of inflammatory bowel diseases. *Gastroenterology*, 149, 1163–1176.e1162.
164. Frank, C., Fallah, M., Ji, J., Sundquist, J. and Hemminki, K. (2014) The population impact of familial cancer, a major cause of cancer. *Int. J. Cancer*, 134, 1899–1906.
165. Dudbridge, F., Brown, S. J., Ward, L., Wilson, S. G. and Walsh, J. P. (2018) How many cases of disease in a pedigree imply familial disease? *Ann. Hum. Genet.*, 82, 109–113.
166. Jansen, A. M., Geilenkirchen, M. A., van Wezel, T., *et al.* (2016) Whole gene capture analysis of 15 CRC susceptibility genes in suspected lynch syndrome patients. *PLoS One*, 11, e0157381.
167. Mensenkamp, A. R., Vogelaar, I. P., van Zelst-Stams, W. A., *et al.* (2014) Somatic mutations in MLH1 and MSH2 are a frequent cause of mismatch-repair deficiency in Lynch syndrome-like tumors. *Gastroenterology*, 146, 643–646.e648.
168. Morak, M., Heidenreich, B., Keller, G., Hampel, H., Laner, A., de la Chapelle, A. and Holinski-Feder, E. (2014) Biallelic MUTYH mutations can mimic Lynch syndrome. *Eur. J. Hum. Genet.*, 22, 1334–1337.
169. Aaltonen, L., Johns, L., Jarvinen, H., Mecklin, J. P. and Houlston, R. (2007) Explaining the familial colorectal cancer risk associated with mismatch repair (MMR)-deficient and MMR-stable tumors. *Clin. Cancer Res.*, 13, 356–361.
170. Espenschied, C. R., LaDuca, H., Li, S., McFarland, R., Gau, C. L. and Hampel, H. (2017) Multigene panel testing provides a new perspective on Lynch syndrome. *J. Clin. Oncol.*, 35, 2568–2575.
171. Gallego, C. J., Shirts, B. H., Bennette, C. S., *et al.* (2015) Next-generation sequencing panels for the diagnosis of colorectal cancer and polyposis syndromes: a cost-effectiveness analysis. *J. Clin. Oncol.*, 33, 2084–2091.
172. Allgauer, M., Budczies, J., Christopoulos, P., *et al.* (2018) Implementing tumor mutational burden (TMB) analysis in routine diagnostics-a primer for molecular pathologists and clinicians. *Translational Lung Cancer Research*, 7, 703–715.
173. Ma, J., Setton, J., Lee, N. Y., Riaz, N. and Powell, S. N. (2018) The therapeutic significance of mutational signatures from DNA repair deficiency in cancer. *Nat. Commun.*, 9, 3292.