Recent Discoveries in the Genetics of Familial Colorectal Cancer and Polyposis

Laura Valle

Hereditary Cancer Program, Catalan Institute of Oncology, IDIBELL, Hospital de Llobregat, Barcelona, Spain

The development of genome-wide massively parallel sequencing, ie, whole-genome and whole-exome sequencing, and copy number approaches has raised high expectations for the identification of novel hereditary colorectal cancer genes. Although relatively successful for genes causing adenomatous polyposis syndromes, both autosomal dominant and recessive, the identification of genes associated with hereditary non-polyposis colorectal cancer has proven extremely challenging, mainly because of the absence of major high-penetrance genes and the difficulty in demonstrating the functional impact of the identified variants and their causal association with tumor development. Indeed, most, if not all, novel candidate non-polyposis colorectal cancer genes identified so far lack corroborative data in independent studies. Here we review the novel hereditary colorectal cancer genes and syndromes identified and the candidate genes proposed in recent years as well as discuss the challenges we face.

Keywords: Cancer Predisposition; Hereditary Non-polyposis Colorectal Cancer; Adenomatous Polyposis; Serrated Polyposis; Mixed Polyposis.

Colorectal cancer (CRC) is the third most common cancer in men (10.0% of all cancers) and the second most common in women (9.2%) worldwide. Most CRCs arise as a consequence of somatic genomic events that disrupt key cellular processes in individual colonic epithelial cells. The vast majority of CRCs develop from preexisting polyps; dysplasia is the true precursor lesion. Therefore, removal of polyps results in decrease of CRC incidence. Family history of cancer is one of the strongest predictors of CRC risk; this risk is higher with increasing number of affected relatives and when CRC occurs at young age. Crude estimates indicate that 20%–25% of all CRC patients have at least 1 relative affected with the disease, which may be explained by shared genetic and/or environmental factors.

It has been estimated that approximately 3%–6% of all CRC patients carry germline mutations associated with syndromic hereditary CRC. This genetic predisposition to CRC has been classically associated with germline mutations or epimutations in the DNA mismatch repair (MMR) genes MLH1, MSH2, MSH6, and PMS2 for non-polyposis cases and in APC and MUTYH (recessive inheritance) for adenomatous colonic polyposis. Other CRC predisposing syndromes, characterized by the presence of hamartomatous polyps, are caused by mutations in SMAD4, BMPR1A, STK11, and PTEN (Figure 1). Despite this knowledge, much of the genetic predisposition to CRC remains unexplained. This missing heritability may be in part multifactorial, ie, caused by the conjunction of moderate-risk or low-risk genetic variants, possibly in conjunction with environmental or lifestyle risk factors. It has been estimated that CRC low-penetrance variants, including the ones that have not been yet identified, may explain at most 5%–10% of the heritability to CRC. Recently, a model created to accurately determine the risk for CRC on the basis of common genetic CRC susceptibility loci showed that the accumulation of risk variants is significantly associated with increased risk of CRC in individuals with a family history of CRC.

The identification of a bona fide germline pathogenic mutation that causes the increased risk and aggregation of CRC in a family has clear consequences in the clinical management of its members; therefore, important efforts are being invested to identify novel genes that explain this predisposition, in particular in families with clear dominant inheritance of the disease. Here we aim to provide a review of the latest advances and discuss current challenges and future perspectives in this field.

Current Approaches for the Identification of Novel Hereditary Colorectal Cancer Genes

Before the development of high-throughput sequence capture methods and next-generation sequencing technologies, hereditary cancer studies were mainly based on genome-wide linkage analysis of large individual pedigrees or multiple pedigrees, followed by positional cloning and

Abbreviations used in this paper: CMMRD, constitutional mismatch repair deficiency; CRC, colorectal cancer; kb, kilobase; MMR, DNA mismatch repair; MSI, microsatellite instability; SPS, serrated polyposis syndrome; WES, whole-exome sequencing; WGS, whole-genome sequencing; WHO, World Health Organization.
study of somatic analysis of mutations, which led to the identification of the previously mentioned hereditary CRC genes.\textsuperscript{12-14} However, despite the enormous efforts made after the identification of the most prominent hereditary CRC genes, linkage analyses followed by positional cloning and/or sequencing of the genes (coding regions) located within the candidate linkage peaks seemed to be unable to identify additional causal genes,\textsuperscript{15-27} implying large heterogeneity, oligogenic or polygenic modes of inheritance, or unconventional mechanisms of gene inactivation among other possibilities.

In the last decade, the rapid development of massively parallel sequencing–based approaches and genome-wide copy number techniques, associated with the decrease in their economic cost, restored hope for the identification of additional hereditary cancer genes. Among the approaches most commonly used for the identification of causal mutations in a genome-wide manner are whole-genome sequencing (WGS) and whole-exome sequencing (WES) and genome-wide, usually array-based, scanning of copy number alterations. These approaches are performed in isolated high-risk families (large pedigrees) or in multiple families or probands (often with an early age of cancer onset as an indicator of genetic predisposition) to identify a shared causal gene. Moreover, in some instances, combination of the above-mentioned methodologies, such as WES/WGS and linkage data, and/or combination of germline and somatic analyses are used to optimize the process.

**Novel Hereditary Colorectal Cancer Genes**

By using the above-mentioned approaches several novel genes have been proposed as responsible for hereditary CRC cases, in some instances also associated with polyposis. For some of the identified genes, the evidence gathered to date is robust, and their testing has been included in routine genetic diagnostics, whereas for others, the identification of additional pathogenic mutations in high-risk families is mandatory to provide the required evidence to consider the study of the gene in the clinical setting.

**Genes Identified Through Agnostic Approaches**

**Novel hereditary colorectal cancer genes.** POLE and POLD1: polymerase proofreading-associated polyposis. By using a combination of WGS/WES and linkage analysis in probands with more than 10 adenomas by age 60 but no germline mutations in APC, MUTYH, or the MMR genes, Palles et al.\textsuperscript{28} identified mutations in the proofreading (exonuclease) domains of DNA polymerase epsilon (POLE; MIM# 174762) and delta (POLD1; MIM# 174761) in individuals/families with multiple colorectal adenomas and CRC, observing high risk of endometrial cancer in female POLD1 mutation carriers. The study of additional series of familial CRC and/or polyposis cases has provided conclusive evidence of the causal role of germline polymerase-proofreading mutations in the predisposition to CRC and polyposis, allowing a better definition of the syndrome and of its associated phenotype.\textsuperscript{29-37}

Data gathered point to a highly penetrant autosomal syndrome characterized by attenuated or oligo-adenomatous colorectal polyposis, CRC, gastroduodenal (mostly duodenal) adenomas, and probably brain tumors. Moreover, female POLD1 mutation carriers are at very high risk of endometrial cancer and possibly at moderate risk of breast tumors.\textsuperscript{28,29,31,32,35} The presence of other tumor types has also been reported in some families,\textsuperscript{28,30,32,33,37} fitting with a defect in a mechanism of correction of DNA errors. The polymerase-proofreading–associated phenotype may appear as non-germline CRC syndrome,\textsuperscript{29,35} and in some instances, presence of mismatch repair defects in the tumors has been reported.\textsuperscript{31,34}

Tumors developed in the context of polymerase-proofreading mutations, both germline and somatic,
show a hypermutated phenotype, leading to more than a million base substitutions in some tumors, with a tumor mutational spectrum characterized by increased proportion of G:C to T:A and A:T to C:G transversions.30,39

Similar to microsatellite instability (MSI), the presence of somatic mutations in POLE has been recently associated with favorable prognosis in sporadic tumors,40 but whether this also applies to tumors arising in germline mutation carriers remains unsolved.

**NTHL1-associated adenomatous polyposis.** By performing WES in 51 adenomatous polyposis and CRC patients from 48 families, Weren et al41 identified 3 Dutch families with homozygous truncating mutations in the *NTHL1* gene (MIM# 616415), fitting with a recessive disorder. All 3 families carried the same *NTHL1* mutation in homozygosis, c.268C>T (p.Gln90*).

Like *MUTYH*, a known recessive adenomatous polyposis gene, *NTHL1* encodes a glycosylase involved in base excision repair, the primary pathway for the repair of oxidative DNA damage. In comparison with the MUTYH protein, NTHL1 targets a broader range of DNA lesions,40 and the tumors developed by *NTHL1* mutation carriers show an increase in C:G>T:A transitions rather than G>C>A>T transversions,41 as has been observed in double knockout mice.43

Despite the recent description of the syndrome and the publication of only 4 families carrying biallelic *NTHL1* mutations, the *NTHL1*-associated phenotype may be not only characterized by the presence of attenuated adenomatous polyposis and CRC but represents a multifocal tumor syndrome41,44 whose precise tumor spectrum remains to be defined.

**GREM1: mixed polyposis in Ashkenazi Jewish families.** Hereditary mixed polyposis syndrome (MIM# 601228) is an unusual disease associated with polyps of multiple and mixed morphologies, including serrated lesions, Peutz-Jeghers polyps, juvenile polyps, conventional adenomas, and CRC, in the absence of any identifiable extracolonic features.45 Between 1999 and 2008 linkage analyses performed in Ashkenazi families with hereditary mixed polyposis allowed to map a mixed polyposis gene (also known as *CRAC1*) to chromosome 15q13.3 and to identify a shared haplotype within the region, suggesting a single founder mutation.46–48 However, no novel, potentially pathogenic mutations in the genes located in the region were identified. The use of a custom oligonucleotide array to search for copy number variation in the region identified the presence of a heterozygous single-copy duplication of about 40 kilobases (kb) centered on chr15:30.77 Mb present in the affected members of the studied families. The change was a simple tandem tail-head duplication with the insertion of a 30 base pair sequence of unknown origin and no homology to known sequences between the duplions. The duplication extended from intron 2 of *SGS5* to a site just upstream of the *GREM1* CpG island and is associated with greatly increased, allele-specific *GREM1* expression.49 50

To date, the *GREM1* 40-kb upstream duplication has only been found in the Ashkenazi Jewish descendants of a single founder. However, the phenotype is not restricted to mixed polyposis, which has led to recommend genetic testing of the *GREM1* founder duplication to all Ashkenazi Jewish subjects with multiple colorectal polyps and those fulfilling the criteria for non-polyposis CRC.50

Recently, a disease-causing tandem-repeat duplication of 16 kb in the regulatory region of *GREM1* (7.7 kb upstream of the gene) was identified in a non-Ashkenazi family with attenuated polyposis with some indications of polypl morphology similar to a juvenile and a metaplastic type.51 Likewise, a duplication of the whole *GREM1* gene was identified in a single early-onset CRC patient without features of mixed polyposis.52 Moreover, it has been demonstrated that a common *GREM1* variant affecting an enhancer, rs16969681, is also associated with CRC susceptibility, conferring approximately 20% differential risk in the general population.52

**RNF43-associated serrated polyposis.** Serrated polyposis is a clinically defined syndrome with multiple serrated polyps in the colorectum and increased CRC risk. The true prevalence of serrated polyposis syndrome (SPS), as defined by the World Health Organization (WHO) criteria, is unclear because of the risk of bias across studies, but it is likely to be below 0.09% as derived from primary colonoscopy screening programs.53 Moreover, the risk to develop CRC (approximately 1.9% in 5 years), lower than it had been first estimated, largely depends on presence of serrated polyps containing dysplasia, advanced adenomas, and/or combined WHO criteria 1 and 3.54,55

For years, researchers have unsuccessfully tried to identify the genetic cause(s) of this clinical entity by agnostic approaches or studying genes that cause other colonic polyposis syndromes.56 By carrying out WES in 20 unrelated subjects with multiple sessile serrated adenomas (16 fulfilled the WHO criteria of SPS), Gala et al57 identified germline mutations in genes that regulate senescence, *ATM, PIF1, TEL20, XAF1*, and *RBL1*, in 5 patients and nonsense mutations in *RNF43*, a regulator of ATM/ATR DNA damage response, in 2 of the 20 studied patients (ages at diagnosis, 51 and 52). The protein encoded by *RNF43* is a RING-type E3 ubiquitligase, which is thought to negatively regulate Wnt signaling.58

Knockout of this gene had been demonstrated to contribute to an intestinal polyposis phenotype in mice.59 Subsequently, WES performed in a serrated polyposis family, including 2 affected and 2 non-affected individuals, revealed a stop-gain novel mutation in *RNF43* shared by the affected family members and absent in the healthy ones.60 Recently, another WES project involving 4 families (6 individuals) with serrated polyposis identified a splice-site mutation in *RNF43* that resulted in a truncating protein product.54 The mutation was identified in 3 siblings in the family, 2 with serrated polyposis diagnosed at age 65 (number of polyps >100,
predominantly sessile serrated adenomas) and at age 64
(number of polyps >20 sessile serrated adenomas or
hyperplastic polyps) and the third with rectal cancer and
a hyperplastic polyp at age 49. Two children of the latter
subject, also carriers of the RNF43 mutation, showed
several sessile serrated adenomas or hyperplastic polyps
at colonoscopy examination at 35 and 37 years. Four
additional family members screened for the mutation
resulted in non-carriers, and all had normal colonos-
copies. A total of 16 serrated polyps, 5 adenomas, and 1
tumor from the 5 identified mutation carriers could be
analyzed, all showing somatic inactivation of the wild-
type allele, thus confirming the tumor suppressor
nature of RNF43.61

In total, germline RNF43 mutations have been iden-
tified in 12.5% (2/16)57 and 25% (1/4)61 of unrelated
patients satisfying the WHO criteria of SPS, thus sup-
porting the need to include mutation screening of RNF43
in routine germline testing, which is something that has
not yet been done because of the recent (2016) publi-
cation of the second and conclusive study. Interestingly,
in vitro and preclinical evidence has shown that RNF43
mutant tumors are sensitive to Wnt secretion inhib-
itors,62–64 opening a potent therapeutic window for
RNF43 mutation carriers.

Novel candidate genes. In addition to the genes
mentioned above, additional candidates have been pro-
posed with more or less compelling evidence of causality
for hereditary CRC (Table 1). Among the genes with
strongest evidence of association with hereditary CRC
are FAN1,65 RPS20,66 PTPN12,67 LRP6,67 BUB1 and
BUB3,68 FOCAD,69 and the constitutional epigenetic
silencing of PTPRJ.70 FAN1, BUB1, and BUB3 are involved
in DNA damage response and genetic instability, and
LRP6 is a component of the Wnt-Fzd-LRP5-LRP6 com-
plex that triggers β-catenin signaling. RPS20 encodes a
ribosomal protein, FOCAD, a focal adhesion protein, and
PTPN12 and PTPRJ, protein tyrosine phosphatases
(GeneCards; www.genecards.org). In addition to these
genes, the study of relatively large numbers of families/
probands has allowed the identification of several
candidate genes,71–76; however, additional evidence
needs to be gathered to discard or confirm their causal
role in CRC predisposition. For the SEMA4A gene, first
described as a strong candidate for hereditary non-
polyposis CRC,79 a subsequent study failed to validate
the variation in SEMA4A as a determinant of CRC risk.80
In conclusion, the confirmation of the diagnostic value of
all these candidate genes depends on the publication of
additional supporting and conclusive evidence.

Genes Identified Through Candidate
Gene Approaches

The implication of known hereditary CRC genes in
DNA repair processes or in well-known relevant path-
ways for colorectal carcinogenesis has motivated
researchers to assess the causal role of genes related to
these mechanisms and pathways in CRC predisposition.
In the case of AXIN2, a component of the Wnt pathway,
there is convincing evidence of its pathogenicity for
attenuated familial adenomatous polyposis with or
without features of ectodermal dysplasia.81–83

Despite the efforts made to elucidate their causal role,
the evidence gathered to date is still insufficient or
contradictory to include OGG1,84–86 NUDT1,84,85 BMP4,87
and EPHB88 in routine genetic diagnostics. On the other
hand, AMER1, UNC5C, and GALNT12 have been discarded
as high-penetrance genes, although a role as a moderate-
risk/low-risk gene cannot be ruled out.89–93

Germline and Somatic Biallelic
Inactivation of DNA Mismatch
Repair Genes

The clinical spectrum linked to germline variation in
the MMR genes, initially restricted to monoallelic muta-
tions in Lynch syndrome, keeps expanding. Constitu-
tional mismatch repair deficiency (CMMRD) is a distinct
childhood cancer predisposition syndrome characterized
by the presence of homozygous or biallelic germline
mutations in the MMR genes. Phenotypically, biallelic
carriers may develop hematologic malignancies, brain
tumors, and gastrointestinal cancers as well as café-au-
lait spots and other findings that mimic neurofibroma-
tosis type 1.94–96 In 2014, the CMMRD consortium
reported the genetic and clinical characteristics of a
series of 23 children with CMMRD from 14 families; 18
children from 13 families were genetically characterized,
showing the universal presence of café-au-lait spots and
the extremely high penetrance of cancer because 22 of
the 23 CMMRD children had developed tumors (4 of
them gastrointestinal cancer). Of note, the presence of
gastrointestinal polyposis was reported in half of the
described carriers (9/18) (ages at diagnosis, 9–17.5
years), a clinical trait that had already been observed.95

Several observations suggest that only less penetrant
(mild) MMR gene mutations are viable in homozygosis,
whereas the highly penetrant, such as those causing Lynch
syndrome, are embryonic lethal when they occur in the 2
alleles: (1) the scarcity of Lynch syndrome–related can-
cers in the parents and other relatives of biallelic mutation
 carriers; (2) the fact that many germline mutations re-
ported in CMMRD have not been reported in Lynch syn-
drome; and (3) MSH2 mutations, the most commonly
reported and highly penetrant in Lynch syndrome, are the
least common in CMMRD.95–97

Very recently, WES was carried out in 102 unrelated
patients with adenomatous polyposis in absence of
germline mutations in the classic and novel polyposis
genes, which allowed the identification of 1 case with
biallelic PMS2 mutations and 2 cases with biallelic mu-
tations in MSH3, another DNA MMR gene, following an
autosomal recessive mode of inheritance.98 All 4 MSH3
mutations were deleterious variants (c.1148delA,
<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Original study</th>
<th>Additional evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adenomatous polyposis, multiple tumors</td>
<td>Weren et al, 201541</td>
<td>Rivera et al, 201544, Wimmer et al, 200895, Bakry et al 201496 (gastrointestinal polyposis in CMMRD caused by PMS2, MSH6, MLH1, and MSH2 mutations)</td>
</tr>
<tr>
<td>Adenomatous polyposis, benign and malignant tumors (adult CMMRD)</td>
<td>Adam et al, 201698</td>
<td>Venkatachalam et al, 201151, Laitman et al, 201550, Rohlin et al, 201636, Taupin et al, 201551 (RNF43)</td>
</tr>
<tr>
<td>Mixed polyposis of serrated and juvenile types</td>
<td>Jaeger et al, 201249 (Ashkenazi founder 40-kb duplication)</td>
<td>Yan et al, 201661 (RNF43)</td>
</tr>
<tr>
<td>Serrated polyposis</td>
<td>Gala et al, 201457</td>
<td>Yan et al, 201661 (RNF43)</td>
</tr>
<tr>
<td>Polyposis, CRC, primary ovarian failure</td>
<td>Weren et al, 201549</td>
<td>Weren et al, 201549 (RNF43)</td>
</tr>
<tr>
<td>Adenomatous polyposis</td>
<td>Spier et al, 201672</td>
<td>Spier et al, 201672 (POLE2)</td>
</tr>
<tr>
<td>Hereditary non-polyposis CRC</td>
<td>Nieminen et al, 201446</td>
<td>Smith et al, 2016114 (hereditary pancreatic cancer)</td>
</tr>
<tr>
<td>Hereditary non-polyposis CRC</td>
<td>Segui et al, 201548</td>
<td>Kinnersley et al, 201659 (no validation)</td>
</tr>
<tr>
<td>Early-onset and familial CRC</td>
<td>de Voer et al, 201358</td>
<td>Glavieux et al, 201358, Schulz et al, 201479</td>
</tr>
<tr>
<td>Hereditary non-polyposis CRC</td>
<td>Venkatachalam et al, 201070</td>
<td>Spier et al, 201572 (POLE2)</td>
</tr>
<tr>
<td>Hereditary breast/ovarian cancer, early-onset CRC</td>
<td>Villacis et al, 2016115</td>
<td></td>
</tr>
<tr>
<td>Hereditary non-polyposis CRC</td>
<td>Wei et al, 2015116</td>
<td></td>
</tr>
<tr>
<td>Hereditary non-polyposis CRC</td>
<td>Gylfe et al, 201373</td>
<td></td>
</tr>
<tr>
<td>Hereditary non-polyposis CRC</td>
<td>Esteban-Jurado et al, 201471</td>
<td>Garre et al, 2015107 (BRCA2)</td>
</tr>
<tr>
<td>Early-onset and familial CRC</td>
<td>Tanskanen et al, 201575</td>
<td>Yurgelun et al, 2015108 (BRCA2, BRIP1)</td>
</tr>
<tr>
<td>Hereditary non-polyposis CRC</td>
<td>Esteban-Jurado et al, 201676</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Novel Genes and Candidate Genes of CRC Predisposition Identified Through WES/WGS or Genome-wide CN Approaches
Table 1. Continued

<table>
<thead>
<tr>
<th>Gene</th>
<th>Phenotype</th>
<th>Original study</th>
<th>Additional evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZRANB1, CDC27, CENPE,</td>
<td>Hereditary non-polypsis CRC</td>
<td>DeRycke et al, 2013</td>
<td></td>
</tr>
<tr>
<td>DDX12, HAUS6/FAM29A,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HIST1H2BE, KIF23, TACC2,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZC3HC1, CTBP2, IRF5,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MED12, RNF111, SF1,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLE1, TLE4, TRIP4, BTNL2,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAGE, CARD8, FANK1,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIR2DL1, KIR2DS4,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KIR3DL3, MASP1, NLRP8,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FANCM, LAM4, PTCHD3,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LAMC3, TRESH, NOTCH3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRC</td>
<td></td>
<td>Smith et al, 2013</td>
<td></td>
</tr>
</tbody>
</table>

*Current diagnostic value.

c.2319-1G>A, c.2760delC, and c.3001-2A>C), and the tumors of the affected carriers showed high MSI of dinucleotides and tetranucleotides and complete loss of the MSH3 protein. In total, 4 MSH3 biallelic carriers were identified, showing an associated phenotype characterized by the presence of colorectal adenomatous polypsis, diagnosed in their 30s in most cases, accompanied by additional benign and malignant lesions in the gastrointestinal tract and extracolonic, such as duodenal adenomas (2/4 biallelic carriers), thyroid adenomas (2/4), intraductal papillomas (2/4), gastric cancer (1/4), and astrocytoma (1/4).98 The phenotype observed in MSH3 biallelic carriers largely resembles that of attenuated familial adenomatous polyposis, although still conserving some features of CMMRD occurring at more advanced age (not in childhood), most probably because MSH3 mutations in heterozygosity, unlike the Lynch syndrome–associated MMR genes, are not considered as disease causing.

In as many as 60%–70% of cases in which Lynch syndrome is clinically and molecularly suspected because of the presence of MMR deficiency in the tumors, genetic testing fails to identify a germline MMR gene mutation.99 These cases have been grouped as Lynch-like. Unlike sporadic microsatellite unstable CRCs, Lynch-like CRCs do not show epigenetic inactivation of the DNA MMR gene MLH1 or mutation in BRAF. We currently know that for 50%–60% of the so-called Lynch-like cases, the MMR deficiency detected in the tumors may be explained by the presence of double somatic hits in MMR genes, most probably representing sporadic cases. Also, germline mutations in other genes that cause impaired DNA error correction, such as MUTYH or the proofreading domains of POLE and POLD1, may in some instances cause MMR deficiency in the tumors because of the somatic alteration of MMR genes.31,34,104–106 Taking into account the presupposed heterogeneity of the so-called Lynch-like entity, in absence of a known germline defect, genetic counseling and clinical surveillance should be guided by the family and/or individual cancer history.

Overlap Among Hereditary Cancer Syndromes

The identification of novel hereditary CRC genes has centralized the research in the field during the last decade. However, the prevalence of pathogenic mutations in the novel genes identified is very small; it is almost insignificant when considering the entire burden of unexplained hereditary cases. The use of next-generation sequencing-based approaches, either for the discovery of novel genes (WGS or WES) or for genetic testing (WES or multi-gene panels), has allowed the identification of hereditary CRC families with germline pathogenic mutations in genes classically associated with other cancer syndromes such as BRCA1 and BRCA2,76,107,108 TP53,109 BARD1,71 or BRIP176,108 or associated with other very distinct CRC/polyposis syndromes. This has been the case for the adenomatous polyposis genes MUTYH and POLD1 and the juvenile polyposis gene BMPRIA, which have been found mutated in hereditary non-polypsis CRC families.29,35,104,106,110,111 These observations support the use of generic gene panels in routine genetic diagnostics of hereditary cancer. Moreover, a broad coverage of genes would also identify clinically misdiagnosed cases, eg, misdiagnosed polyposis types. Nevertheless, the infrequency of these out-of-the-rule exceptions suggests that clinical phenotypes are the best way to prioritize, mainly when coping with variants of unknown significance.

More Than 1000 Exomes of Familial and Early-onset Colorectal Cancer Patients

Attempts have been made to estimate by using agnostic approaches the fraction of CRC cases in which the molecular basis of predisposition can be identified. Recently, Chubb et al78 comprehensively assayed the impact of rare (MAF <1%) germline mutations on CRC risk by analyzing WES data from 1006 early-onset
of familial CRC cases could be explained by already known hereditary CRC genes. Interestingly, for other well-documented CRC genes including MSH6 and PMS2, no significant association was detected, reflecting a more limited contribution to CRC predisposition.

When searching for novel CRC predisposing genes, the analysis was restricted to disruptive (nonsense and frameshift) variants in genes recurrently mutated with biological plausibility. This approach identified IL12RB1, LIMK2, and POLE2 as novel potential candidates. A subsequent analysis considering shared biological processes revealed that the DNA-dependent DNA replication gene set (Gene Ontology: genes involved in DNA replication driven by DNA polymerases) was significantly associated with CRC, driven by disruptive mutations in the MMR genes, but also in POLE2, POT1, and MRE11A. Overall, this study confirms the lack of further major high-penetance susceptibility genes that individually account for >1% of the familial risk.

Concluding Remarks

The application of genome-wide WES, WGS, and CN approaches for the identification of novel hereditary CRC genes has been reasonably successful for polyposis syndromes either dominant (POLE, POLD1, GREM1, and RNF43) or recessive (NTHL1 and MSH3), although there is still a high proportion of polyposes that are not explained by currently known genes. Because POLE, POLD1, and GREM1 mutation screening is already implemented in routine genetic testing and/or multi-gene panels, the study of NTHL1, MSH3, and RNF43 will soon become available. Moreover, genetic counseling and surveillance recommendations for mutation carriers of these newly described genes/syndromes, currently guided by the available clinical phenotypic observations, will be refined with the description and report of additional mutation carriers.

In the case of hereditary non-polyposis CRC, the absence of novel major genes and the anticipated extreme genetic heterogeneity are the major conclusions derived from the research carried out so far. Despite the enormous efforts made in recent years, the yield of WES/WGS for the identification of novel hereditary non-polyposis CRC genes has been low. Moreover, none of the proposed candidates has yet gathered the required evidence and supporting validation to be included in routine genetic diagnostics, an issue on which future/current research should invest efforts and resources.

Adding to these complex circumstances is the interpretation of the identified genetic variants, which hinders the identification and validation of novel genes. First, the human genome harbors many rare variants, most of which may not be clearly associated with a disease phenotype, and second, some of these variants may interplay with other genetic and/or environmental factors, which may in turn determine their expressivity.

The breakthroughs made through the years in the identification of novel CRC/polyposis predisposing genes or candidate genes have made evident the relevance of genes involved in DNA repair, DNA replication, and maintenance of genomic instability. Furthermore, the study of certain ethnicities such as Ashkenazi Jewish or specific populations (eg, Finnish or Dutch) continues to be helpful in the identification of novel genes because of the presence of either founder mutations, as was the case for GREM1, or homozygous carriers of pathogenic mutations, as happened with NTHL1. These findings highlight the importance of combining regional or national studies with data sharing and coordination in international consortia.

The presence of overlapping phenotypes, ie, causal germline mutations in genes that mainly cause predisposition to other tumor types or other phenotypes, has been recurrently reported in recent years, especially with the introduction of generic gene panels for genetic testing in hereditary cancer or the use of WES/WGS for research or diagnostic purposes. Although a priori this phenomenon does not seem to be extremely frequent, it might well explain more uncharacterized familial cases than the novel hereditary CRC gene candidates that are identified so far. The study of additional series together with genotyping of population-based cohorts will provide a more refined assessment of the clinical spectrum and expressivity of the germline variation detected.

Although historically only associated with hereditary non-polyposis CRC (restricted to heterozygous mutations in MSH2, MLH1, MSH6, and PMS2), through the years the clinical and molecular spectra of pathogenic MMR gene mutations have evolved to (1) CMMRD caused by germline biallelic mutations; (2) attenuated adenomatous polyposis in the presence of germline biallelic mutations, including in MSH3 an MMR gene with no effect in heterozygosity; and (3) Lynch-like syndrome when somatically inactivated (2 somatic hits).

The possibility of oligogenic or polygenic inheritance is no longer a theoretical hypothesis but a reality exemplified at one end by the presence of germline pathogenic mutations in more than 1 high-penetance cancer-predisposing gene, the so-called multilocus inherited neoplasia alleles syndrome, and at the other end by the increased cancer risk in carriers of multiple low-penetrance alleles. However, these extreme cases most likely only represent the tip of the iceberg, because the identification and classification of variants/genes of moderate/lower expressivity, which probably exert their effect when co-inherited with others or under specific conditions.
environmental conditions, are one of the most difficult challenges we face in the field. Indeed, the oligogenic or polygenic inheritance, together with the presence of alterations, genetic or epigenetic, in non-coding regulatory regions (expression or splicing regulators, or non-coding RNAs), opens the door to additional studies in the numerous families where neither WES nor WGS focused on variants of evident functional impact has been able to identify a monogenic cause for the familial aggregation of CRC.

**References**


43. Chan MK, Ocampo-Hafalla MT, Vartanian V, et al. Targeted deletion of the genes encoding NTH1 and NEIL1 DNA N-glycosylases reveals the existence of novel carcinogenic oxidative damage to DNA. DNA Repair (Amst) 2009;8:786–794.


50. de Voer RM, Geurts van Kessel A, Weren RD, et al. Germline mutations in the spindle assembly checkpoint genes BUB1 and


85. Dalloso AR, Dolwani S, Jones N, et al. Inherited predisposition to colorectal adenomas caused by multiple rare alleles of MUTYH but not OGG1, NUDT1, NTH1 or NEIL 1, 2 or 3. Gut 2008;57:1252–1255.


95. Wimmer K, Etzler J. Constitutional mismatch repair deficiency syndrome: have we so far seen only the tip of an iceberg? Hum Mutat 2008;147:1308–1316.


Reprint requests
Address requests for reprints to: Laura Valle, PhD, Hereditary Cancer Program, Catalan Institute of Oncology, IDIBELL, Av. Gran Via 199-203, 08908 Hospitalet de Llobregat, Barcelona, Spain. e-mail: lvalle@iconcologia.net; fax: +34-932607466.

Acknowledgments
The author thanks Gabriel Capellá and Marta Pineda for critical review of the manuscript.

Conflicts of interest
The author discloses no conflicts.

Funding
Supported by the Scientific Foundation Asociación Española Contra el Cáncer, the Spanish Collaborative Cancer Research Network (RTICC RD12/0036/0031), and the Government of Catalonia (2014SGR338).