The handle http://hdl.handle.net/1887/62614 holds various files of this Leiden University dissertation.

**Author:** Jansen, A.M.L.
**Title:** Discovery of genetic defects in unexplained colorectal cancer syndromes
**Issue Date:** 2018-05-29
Chapter 1

General introduction
and thesis outline
General introduction

Colorectal cancer (CRC) is the third most common cancer in men and second most common in women, with an incidence of 1.4 million cases worldwide in 2012. An estimated 25-35% of all CRCs have heritable components, either pathogenic variants in high-risk CRC susceptibility genes (3-5%) or a positive family history (20-30%). The underlying genetic cause in the patients without penetrant mutations but with family history is not well understood, but is expected to be a combination of environmental and inherited genetic factors with common, low-penetrant genetic alterations. With large genome-wide association studies (GWAS) more of these CRC susceptibility loci are being identified. These loci, often single-nucleotide polymorphisms (SNPs), slightly increase colorectal cancer risk. With information about the combined risk of multiple of these SNPs a personal CRC-risk profile can be created. An effort to calculate personal cancer risk scores (with a polygenic risk model) by combining risk scores of multiple (moderate) cancer susceptibility loci is already being done for breast cancer, and could be a possibility for colorectal cancer, if more CRC susceptibility loci are mapped.

Colorectal carcinomas usually start as benign polyps that grow from normal colonic mucosa. The progression of normal colonic epithelial cells to adenocarcinomas usually follows the classical progression of precursor lesions with somatic, genetic and epigenetic changes. These changes often confer a growth advantage leading to clonal expansion of the altered cells. This process is better known as the adenoma-carcinoma sequence and typically spans over 15 years. One fundamental aspect of the tumorigenesis process is the acquisition of genomic instability, which can be present in one of these forms: microsatellite instability (MSI), chromosomal instability (CIN) or CpG island methylator phenotype (CIMP). MSI is caused by defects in the mismatch repair system, a characteristic of Lynch Syndrome tumors. CIN is described to be caused by a combination of oncogene activation (e.g. KRAS and PICK3CA) and tumor suppressor gene inactivation (e.g. APC, TP53 and SMAD4). Over 80% of adenomas and CRCs are found to have inactivating APC variants. These pathogenic variants result in Wnt signaling activation, a key early event in CRC tumorigenesis. CIMP is a subset of CRCs that result from epigenetic changes and that are characterized by the inactivation of multiple tumor suppressor genes and other tumor-related genes.

The three different types of genomic instability are a result of heritable factors, environmental factors and random mistakes during normal DNA replication. Known heritable factors often increase the number of variants per replication of the cell by disabling correct proofreading, or by affecting one of the multiple DNA repair pathways present in the cell. Besides genetic variants directly affecting protein function, CRC susceptibility can occur through other forms of transcriptional silencing. The best known transcriptional silencer is epigenetic promoter methylation, described in MLH1, MGMT, APC and P16/CDKN2A, but transcriptional silencing can also occur through microRNAs (miRNAs). MiRNAs are small nucleotide sequences that participate in the regulation of cell differentiation, cell cycle progression, and apoptosis. Dysregulation of miRNAs has been shown to play a role in CRC tumorigenesis. In an effort to create one consensus method to classify CRC subtypes, a large international consortium classified CRCs based on gene expression profiles. Four consensus molecular subtypes (CMS) were defined: CMS1: MSI, hypermutated and strong immune activation, CMS2: epithelial with Wnt and Myc signaling activation, CMS3: epithelial with metabolic dysregulation and CMS4: TGF-b activation, stromal invasion and angiogenesis. This classification was based on gene expression profiles and not on underlying genetic causes.
<table>
<thead>
<tr>
<th>Syndrome</th>
<th>Abbr.</th>
<th>Genes</th>
<th>Pattern of inheritance</th>
<th>Prevalence</th>
<th>Phenotype</th>
<th>Screening</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CRC syndromes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lynch Syndrome</td>
<td>LS</td>
<td>MLH1, MSH2, MSH6, PMS2, EPCAM</td>
<td>Dominant</td>
<td>~2.5%</td>
<td>MMR-/MSI early onset CRC. Freq: EC, SBC, StC, OC, BlC</td>
<td>Colonoscopy every 1-2 years (start age 20-25).</td>
</tr>
<tr>
<td>Familial colorectal cancer type X</td>
<td>FCCTX</td>
<td>SEMA4A, BMPR1A</td>
<td>Unknown</td>
<td>Unk</td>
<td>MMR+ Amsterdam positive families.</td>
<td>Colonoscopy every 3-5 years (start 5-10 years before earliest age of onset in family).</td>
</tr>
<tr>
<td><strong>Polyposis syndromes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Familial adenomatous polyposis</td>
<td>FAP</td>
<td>APC</td>
<td>Dominant</td>
<td>~1%</td>
<td>100-1000 colorectal adenomas at early age. If untreated, CRC.</td>
<td>Colonoscopy every 1-2 years (start age 10-14).</td>
</tr>
<tr>
<td>MUTYH associated polyposis</td>
<td>MAP</td>
<td>MUTYH</td>
<td>Recessive</td>
<td>0.3-1%</td>
<td>10-100 colorectal adenomas, with possible CRC.</td>
<td>Colonoscopy every 2 years (start age 18-20)</td>
</tr>
<tr>
<td>Polymerase proofreading associated polyposis</td>
<td>PPAP</td>
<td>POLE/POLD1</td>
<td>Dominant</td>
<td>Unk</td>
<td>Early onset CRC with ultramutated phenotype, often also polyps</td>
<td>Colonoscopy every 1-2 years (start age 20-25). EC screening from age 40 for female POLD1 carriers.</td>
</tr>
<tr>
<td>NTHL1-associated polyposis</td>
<td>NAP</td>
<td>NTHL1</td>
<td>Recessive</td>
<td>Unk</td>
<td>CRC with polyposis and frequent BrC, EC, SkC, DuoC and PrC.</td>
<td>Unknown</td>
</tr>
<tr>
<td>MSH3-associated polyposis</td>
<td></td>
<td>MSH3</td>
<td>Recessive</td>
<td>Unk</td>
<td>Unk</td>
<td>Unknown</td>
</tr>
<tr>
<td>Juvenile polyposis</td>
<td>JPS</td>
<td>SMAD4, BMPR1A</td>
<td>Dominant</td>
<td>&lt;1%</td>
<td>Spherical, lobulated polyps at young age (mean age 18.5 years)</td>
<td>Colonoscopy every 3 years, starting at age 15.</td>
</tr>
<tr>
<td>Peutz-Jeghers Syndrome</td>
<td>PJS</td>
<td>STK11</td>
<td>Dominant</td>
<td>&lt;1%</td>
<td>Hamartomatous polyps and dark brown pigm. on lips, hands and feet. Freq: StC, BrC, OC, LuC</td>
<td>LB surveillance every 3 years (start age 18), UGI endoscopy every 3 years from age 25 and mammography from age 25.</td>
</tr>
<tr>
<td>PTEN hamartoma tumor syndrome</td>
<td>PTHS</td>
<td>PTEN</td>
<td>Dominant</td>
<td>&lt;1%</td>
<td>Hamartomatous polyps in GI tract, macrocephaly, papillomas and fibromas</td>
<td>Colonoscopy every 5 years from age 35, yearly: thyroid (from age 18) and breast examination (from age 25-30).</td>
</tr>
</tbody>
</table>

*Abbr = abbreviation, MMR = mismatch repair, CRC = colorectal cancer, EC = endometrium cancer, SBC = Small Bowel Cancer, StC = Stomach cancer, OC = ovariunm cancer, BlC = Bladder cancer, SkC = skin cancer, DuoC = duodenum cancer, BrC = breast cancer, PrC = prostate cancer, LuC = lung cancer, LB = large bowl, UGI = upper gastrointestinal, Unk = unknown, MSI = microsatellite instability*
Familial syndromes

A number of different hereditary colorectal cancer and polyposis syndromes have been defined on the basis of distinct clinical, pathological and molecular characteristics. Each of these has been linked to, or even named after, (a) specific gene(s) (Table 1), but the genotype-phenotype connection has become considerably more complicated in recent years. The most common familial colorectal cancer syndromes are discussed separately below.

Lynch Syndrome

Lynch Syndrome is the most common form of hereditary CRC, accounting for 2-5% of all CRCs in the general population, and is caused by heterozygous pathogenic germline variants in one of the mismatch repair (MMR) genes, MLH1, MSH2, MSH6 and PMS2. In addition, in approximately 20-25% of patients with immunohistochemical MSH2 loss but without pathogenic MSH2 variants a germline deletion of the 3' end of the EPCAM gene is found, resulting in allele-specific hypermethylation and transcriptional inactivation of MSH2 that is directly upstream from EPCAM. The function of the MMR pathway is to check DNA replication fidelity and repair DNA mismatches that occur due to replication errors. This is necessary for maintaining genome stability. Although functioning in the same cellular pathway, the MMR proteins form distinct heterodimers. The MutS complex (MSH2/MSH6 or MSH2/MSH3) is responsible for recognition of mismatches and small insertions/deletions (indel). The MutL complex (MLH1/PMS2, MLH1/PMS1 or MLH1/MLH3) is responsible for forming a MutS/MutL/DNA complex, for endonuclease activity and for termination of mismatch-provoked excision. Because MLH1 and MSH2 can heterodimerize with multiple proteins, these proteins are shown to be essential MMR components, while MSH6, PMS2, PMS1, MSH3 and MLH3 are important but partially redundant. When MMR ability is lost, cells develop a ‘mutator’ phenotype characterized by a 100-1000 times increase in mutation rate. Microsatellites, repetitions of small DNA sequences, are more mutation-prone and become unstable if the MMR system is defective. This mutational signature known as microsatellite instability (MSI) is characteristic of MMR-deficient tumors. Other characteristics of MMR-deficient tumors are a high density of tumor infiltrating lymphocytes (TILs) and a proximal location in the colon. TILs are related to a specific antigen-driven immune response, have been described to be activated and to have a cytotoxic nature and are associated with improved prognosis.

Approximately 15% of colorectal cancers display the MSI phenotype of whom the majority (>85%) are sporadic and result from somatic MLH1 promoter hypermethylation. These sporadic MLH1 methylated tumors commonly occur at relatively advanced age, and typically do not show a family history of CRC. MLH1 methylated cancers often carry the somatically acquired BRAF V600E variant. While BRAF testing has low specificity, it is still used in some centers as a pre-screening method to select cancers with methylated MLH1 promoters. Pathogenic germline MMR variants and somatic MLH1 hypermethylation are usually described to be mutually exclusive, although rare cases of LS have been described with somatic promoter hypermethylation as a genetic ‘second hit’. Furthermore, though rare, germline MLH1 promoter hypermethylation has been described in young CRC patients with MLH1 promoter hypermethylated tumors. Inheritance of a constitutional epimutation has been previously described in at least three unrelated families.
LS is inherited in an autosomal dominant fashion, with an average cumulative risk of developing CRC at the age of 70 years ranging from 34-67% for MLH1 and MSH2 mutation carriers, 62, 63, 22-69% for MSH6 mutation carriers, 64, 65, and 11-20% for PMS2 mutation carriers. 66, 67 Additionally, female mutation carriers have an average cumulative risk of developing endometrial cancer of 31-60% for MLH1 and MSH2 mutation carriers, 68-71, 44-70% for MSH6 mutation carriers, 64, 65, 71 and 12-15% for PMS2 carriers. 66, 67 Other cancer types, including small bowel, stomach, pancreas, ovary and bladder cancer occur, but less frequently. 63-69, 72 Recent studies also indicate an increased risk for prostate and breast cancer. 73-75

For patients carrying a pathogenic MMR variant, colonoscopy is recommend every 1 to 2 years starting at ages 20 to 25 years to reduce the incidence and mortality of this tumor. 76, 77 Previous studies have shown that MMR-deficient tumors have a better clinical prognosis and possibly a better response to novel regiments such as immunotherapy. 33, 34, 78, 79

In the past, a set of criteria (Amsterdam Criteria) were used to clinically diagnose Lynch Syndrome families. 80 These criteria considered age of onset, number of CRC patients within a family and the relation between the affected family members (one should at least be a first-degree relative of the other). 80 However, even after inclusion of other LS-associated non-colorectal cancers (Amsterdam II Criteria), these diagnostic criteria were found to lack sensitivity and specificity in diagnosing Lynch Syndrome. 81, 82 In 1996 the Bethesda guidelines were introduced to identify individuals who should receive genetic testing for Lynch Syndrome. 83 These guidelines advised to screen patients fulfilling one of the following criteria: (1) individuals in families that meet Amsterdam Criteria, (2) patients with two Lynch-associated cancers, (3) patients with CRC and a first-degree relative (FDR) with an LS-associated cancer before age 45, (4) patients with right-sided undifferentiated or signet-ring cell type CRC before age 45 or (5) patients with adenomas before age 40. 83, 84 The Bethesda guidelines showed a high sensitivity (96%), an improvement on the previously used Amsterdam Criteria, but still a low specificity (27%). 85

The introduction of the MSI analysis and immunohistochemical staining led to the revised Bethesda guidelines and a more sensitive method to determine MMR-deficiency. 86-89 Immunohistochemical staining of tissue slides from formalin-fixed paraffin-embedded tumors is used to determine the presence and location of a protein. The first step, tissue preparation, includes deparaffinization of the tissue slides, blocking of endogenous enzymatic activity and (heat-induced) antigen retrieval. After antigen retrieval specific primary and then secondary antibodies are added to the slides. Antibodies are detected with a chromogenic reaction, in which an enzyme label conjugated to the antibody reacts with a substrate to yield a colored precipitate. Horseradish peroxidase (HRP) is often used for this reaction, and the precipitating substrate DAB shows a typical brown-colored precipitate at the protein localization site. Finally, the slide is counterstained, commonly with hematoxylin, a compound creating a blue color. Brown-colored precipitate will indicate the presence of the target protein at a specific location, while only the blue counterstain will be seen when the protein is absent (see Figure 1).

Importantly, non-tumor cells present in the sample should always show staining, since the target protein is still present in these cells. This allows these cells to be used as a positive control, i.e. to confirm that lack of staining is due to loss of protein and not due to technical artifacts.
For MSI analysis the National Cancer Institute microsatellite panel, consisting of at least 5 microsatellite markers, is recommended, although commercially available mononucleotide panels are routinely used. Tumors are characterized to be MSI-low (MSI-L) if one marker shows instability, MSI-high (MSI-H) if two or more markers show instability and MSI-stable (MSS) when no markers show instability.

The revised Bethesda guidelines advise genetic testing of all CRCs with an age of onset younger than 50, as well as all MSI-H CRCs before 60 years. Furthermore, patients with one FDR with an LS-associated CRC before 50, or two or more first or second degree relatives regardless of age should also be tested for genetic variants. Combining revised Bethesda criteria with MSI analysis and IHC was found to result in a sensitivity of 82% and a specificity of 98%. However, more recent studies advocate routine molecular screening of patients under 55 or even under 70, regardless of Bethesda criteria. This molecular screening is then often combined with BRAF V600E or MLH1 promoter hypermethylation testing to exclude sporadic MLH1 loss.

Figure 1: Colon high grade villous adenoma, IHC of the four MMR proteins

Unexplained suspected Lynch Syndrome

A 2014 study estimated that up to 60% of MMR-deficient colorectal cancers do not carry germline MMR variants, nor can be explained by somatic MLH1 promoter hypermethylation. These patients are referred to as ‘suspected Lynch Syndrome’ (sLS) or ‘Lynch-like Syndrome’ (LLS), and failure to determine the underlying (genetic) cause of disease has a major impact on the clinical management of these patients. Three potential reasons for MMR-deficient and/or MSI-H cancers of sLS patients discussed in literature are (1) missed variants in the MMR genes, (2) biallelic somatic inactivation of the MMR genes or (3) variants in other genes that can drive MSI. These possible explanations are comprehensively discussed in Chapter 7 (concluding remarks). The cancer risk in families with sLS is found to be lower than that of families with LS but higher than that of families with sporadic CRCs and more research is needed into the potential (genetic) causes of these CRCs.

Currently, many high-throughput screening efforts are being done to find the genetic cause in these sLS families, resulting in many (MMR) variants of uncertain clinical significance (VUS). These are the variants for which evidence is lacking to classify them as either (likely) benign or (likely) pathogenic. Characterization of MMR variants is done according to the standardized five-tiered scheme of the International Society for Gastrointestinal Hereditary Tumors (InSiGHT). According to this scheme, variants can be classified to be not pathogenic (class 1), likely not pathogenic (class 2), uncertain (class 3), likely pathogenic (class 4) or pathogenic (class 5). While the scheme provides clear clinical guidelines for classes 1, 2, 4, and 5, many variants are assigned to class 3 for lack of good classification evidence, and clinical impact of these variants remains uncertain. In addition to clinical data (such as family history, cosegregation, immunohistochemistry, etc.), functional tests, such as minigene splicing assays or in vitro MMR assays, may help to interpret the clinical impact of these variants.

Constitutional MMR-deficiency

Patients with homozygous or compound heterozygous variants in the MMR genes show a different phenotype than classical LS patients, known as constitutional MMR-deficiency syndrome (CMMRD). CMMRD patients develop a diverse spectrum of childhood cancers, including CRC but also hematological and brain malignancies. Another characteristic of CMMRD is café-au-lait maculae (CALM). Most CMMRD families have homozygous/compound heterozygous PMS2 variants, but families with biallelic MLH1 or MSH6 variants have also been described. The mean age at diagnosis in patients with CMMRD is 7-9 years for brain tumors, 16 years for CRC and 5-12 years for hematological cancers. Recommended surveillance consists of MRI scanning of the brain starting at the age of 2 years at an interval of 6-12 months and colonoscopies every 6 months from the age of 8 years. Siblings of CMMRD patients have a 25% risk of having the same genotype, and 50% chance of carrying a heterozygous MMR variant with an increased risk for LS-associated tumors in adulthood. CMMRD is a severe disorder with a large spectrum of cancers. Depending on the type of pathogenic variant patients can have a severe phenotype with brain tumors in early childhood, or a milder phenotype with later age of onset. Surveillance will aid in early detection of tumors and guide proper treatment, but most patients will die from cancers in early childhood.
**Muir-Torre**

Muir-Torre syndrome (MTS) is an autosomal dominant skin condition characterized by sebaceous gland tumors or keratoacanthoma, with colorectal, endometrial, urological or upper gastrointestinal neoplasms.\(^{110-113}\) Clinical evidence suggests there might be two types of MTS, one with MMR-proficient/MSS and one with MMR-deficient/MSI-H tumors.\(^{110}\) The latter is possibly a clinical variant of Lynch Syndrome, and is often caused by pathogenic variants in the *MSH2* gene.\(^{110, 112-114}\) Identical *MSH2* variants have been found in LS- and MTS-patients, and a possible explanation for the different phenotype could be that the *MSH2* variant in MTS-patients co-segregates with variants in other modulator genes involved in skin carcinogenesis or that inactivation of *MSH2* could result in molecular changes in other genes.\(^{110, 113, 114}\) For the MMR-proficient type of MTS the genetic cause is still unknown, but biallelic *MUTYH* germline variants have been implicated as one of the possible genetic causes.\(^{115, 116}\)

**Familial colorectal cancer type X**

Approximately half of the families positive for the Amsterdam I criteria carry MMR-proficient and MSS colorectal cancers.\(^{117-119}\) These families are classified as familial colorectal cancer type X (FCCTX) and the molecular mechanism underlying these tumors is not well understood.\(^{117-120}\) Specific clinical features of FCCTX (compared to LS) include absence of endometrial cancers and high prevalence of rectal cancers, lower incidence of CRC and diagnosis at a higher mean age (57.3 years, compared to 49.7 for LS).\(^{119-124}\) FCCTX cancers have not (yet) been linked to one specific gene.\(^{123}\) Recent studies indicate *SEMA4A* and *BMPRIA* as possible underlying genetic causes, explaining a small percentage of FCCTX cases.\(^{125, 126}\) Other candidate genes are *CENPE, KIF24, GALNT12, ZNF367, GABBR2* and *BMP4*, but evidence for the involvement of these genes in FCCTX-tumorigenesis is lacking.\(^{120, 124}\) One study suggests that FCCTX is not a single entity, but rather a name for a combination of multiple entities.\(^{123}\) Surveillance recommendations currently include colonoscopies every 3-5 years, starting 5-10 years before the earliest age of onset in the family.\(^{118, 124}\)

**Familial adenomatous polyposis**

Familial adenomatous polyposis (FAP) is a rare autosomal dominant disease accounting for ~1% of colorectal cancers.\(^{24, 127}\) FAP is caused by germline heterozygous variants in the *APC* gene and is associated with the development of hundreds to thousands of colorectal adenomas at an early age.\(^{24, 128}\) On average, cancers develop a decade after the first appearance of adenomas and the average age of CRC diagnosis if left untreated is 39 years.\(^{4, 128, 129}\) *De novo* pathogenic *APC* variants are responsible for approximately 25% of FAP patients.\(^{4, 130-132}\) *APC* is a tumor suppressor gene involved in the Wnt signaling pathway, which functions by negatively regulating the β-catenin oncoprotein.\(^{128, 130}\) The 310 kDa protein has four β-catenin binding domains, and seven domains involved in binding and down-regulating β-catenin.\(^{4, 130-133}\) In absence of APC β-catenin accumulates in the nucleus and interacts with factors that upregulate transcription of genes involved in cell cycle, proliferation, differentiation, migration, apoptosis and progression.\(^{130}\) Notably, when *APC* is inactivated in the tumor, β-catenin overexpression is still kept in check by unknown mechanisms. Apparently, some residual downregulation is of great importance for tumor formation.\(^{133}\) In addition, APC stabilizes microtubules and loss of APC leads to chromosomal instability, defective chromosome segregation and aberrant mitosis.\(^{130}\)
Somatic inactivation of APC is a common molecular event in sporadic colorectal cancer and is present in about 80% of sporadic colorectal cancers. The majority of pathogenic APC variants are truncating variants, either nonsense (26%), small insertions (10%) or small deletions (46%).

FAP can be present in the classical, more severe form, or as attenuated FAP (AFAP) a milder form with fewer adenomas (<100) and a later age of onset. The severity of FAP is associated with the location of the APC variant within the APC gene. Variants resulting in an AFAP phenotype are located before codon 157, after codon 1595 and in the alternative spliced region of exon 9. Variants located in the DNA binding domain of the APC gene are described to lead to a severe type of FAP (>thousands of adenomas). The location of the variant however, is not the only predictive factor of the severity of polyposis. APC variants in a mosaic fashion are described to lead to an (attenuated) form of polyposis, irrespective of the variant’s location. The severity of polyposis in patients with mosaic APC variants depends on timing and origin of the mutation. Patients with FAP should be examined by colonoscopy every 1-2 years, beginning at age 10-14. Once adenomas are detected, annual follow-up is recommended. Management of FAP includes endoscopic polypectomy and surgery. Colectomy should be considered when more than 20 adenomas develop, when adenomas >1 cm are found, or when advanced histology (ulcerated, high grade dysplasias) appears.

MUTYH-associated polyposis

Besides FAP, another possible diagnosis for patients presenting with 10-100 adenomas is MUTYH-associated polyposis (MAP). MAP is an autosomal recessive disorder caused by biallelic variants in the base excision repair gene MUTYH, accounting for approximately 0.3 - 1% of all CRCs. The base excision repair (BER) pathway has an important role in preventing variants associated with oxidative damage. Reactive oxygen species (8-oxo-7,8-dihydro-2'-deoxyguanosine or 8-oxodG) can be incorporated in the DNA through direct oxidation of guanine or via incorporation of 8-oxodGTP from the nucleotide pool. DNA polymerase incorporates adenosine opposite 8-oxodG, leading to G:C>T:A transversions. The function of MUTYH within this pathway is to scan the daughter strand after replication and to remove adenosine residues mispaired with guanose or 8-oxoG.

Pathogenic germline variants in the MUTYH gene were first detected in 2002 in one family in which 11 tumors from 3 affected siblings were screened for somatic APC variants. In these 11 adenomas and carcinomas 18 inactivating somatic variants were found, of which the majority (n=15) were G:C>T:A variants. This was taken as a strong indication of a BER defect, previously described in yeast. Sequencing of leukocyte DNA for variants in the BER-genes MUTYH, OGG1 and MTH, led to the discovery of pathogenic compound heterozygous variants in MUTYH. Another hallmark of these tumors is the KRAS c.34G>T variant, found in 64% of MAP colorectal cancers.

MAP is characterized by a greatly increased risk of lifetime colorectal cancer (43-100%), often in combination with colonic adenomas. Monoallelic variants in MUTYH are present in 1-2% of the general population and the cancer risk of these heterozygous carriers is still under debate. Surveillance for MAP consists of colonoscopy every two years starting at age 18-20 years.
Polymerase proofreading associated polyposis
Recently, variants in the exonuclease domains of POLE and POLD1 genes have been described to be associated with colorectal carcinomas (CRC), endometrial cancer (EC) and colorectal polyposis.\textsuperscript{158-165} POLE and POLD1 are the genes that encode for the catalytic subunit of polymerases ε and δ, involved in DNA replication of the lagging and leading strand respectively. The exonuclease domain provides proofreading capabilities essential for maintenance of replication fidelity.\textsuperscript{158-165} The mean age of onset of CRC is 40.7 years for POLE mutation carriers, and 35.9 years for POLD1 carriers.\textsuperscript{165} In a small percentage of POLE/POLD1 mutation carriers brain tumors are diagnosed.\textsuperscript{163-165} This variable phenotype has been coined polymerase proofreading associated polyposis (PPAP), a syndrome with high penetrance and dominant inheritance.\textsuperscript{162-165} Interestingly, in contrast to the classical tumor development model, only a minority of tumors are found to have loss of the wildtype allele, or sustain other variants that could act as a ‘second hit’.\textsuperscript{158-160, 162} Somatic and germline variants in POLE/POLD1 are believed to account for 3% of all CRCs and 7% of ECs.\textsuperscript{160-164, 166} PPAP tumors are often MSS but MSI-H PPAP tumors have been described, where the MMR-deficiency supposedly resulted from somatic secondary MMR variants.\textsuperscript{160, 167, 168} Guidelines recommend colonoscopies every 1-2 years starting at age 20-25, combined with endometrial cancer screening at age 40 for POLD1 female carriers.\textsuperscript{165}

Pathogenic POLE and POLD1 variants have been described as inherited (PPAP) and somatically acquired, both leading to an ‘ultramutated’ phenotype with a variant incidence exceeding 100 variants/Mb.\textsuperscript{158-164} While the majority of somatic variants are C:G>T:A variants, a particular increase in G:C>T:A transversions are characteristic of POLE/POLD1 mutated tumors, with an elevated TCT>TAT and TCG>TTG mutational pattern.\textsuperscript{158-163} Germine or somatic POLE/POLD1 mutated tumors are significantly more immunogenic with increased lymphocyte infiltration and cytotoxic T-cell marker expression, and have a favorable prognosis.\textsuperscript{169, 170}

NTHL1-associated polyposis
Due to many high-throughput sequencing efforts, new genes predisposing to familial colorectal syndromes continue to be found. In 2015, whole exome sequencing of 48 families with colorectal cancer and polyps led to the identification of NTHL1-associated polyposis (NAP).\textsuperscript{171} NAP is a recessive disorder caused by biallelic inactivation of the NTHL1 gene.\textsuperscript{171-174} This gene is part of the base excision repair pathway and encodes for the NTHL1 glycosylase which is involved in removing oxidative pyrimidine lesions.\textsuperscript{175} So far, only a few families with NAP have been described, and the prevalence and the exact phenotype remain unknown. Families with NTHL1 variants appear to have a phenotype predominantly consisting of colorectal cancer with adenomatous polyposis, although breast, endometrial, duodenal, skin, prostate and pancreatic cancers have also been described in NAP patients.\textsuperscript{128, 171} Furthermore, the mutational profile of these cancers resembles an MAP phenotype with G:C>T:A transversions.\textsuperscript{171} Currently, only nonsense and splice site variants have been described, often the NTHL1 p.Gln90* hotspot variant.\textsuperscript{171-174}
MSH3-associated polyposis

In 2016 another new genetic underlying cause of unexplained polyposis was detected through whole exome sequencing (WES). After WES on leukocyte DNA from 102 unrelated individuals with unexplained polyposis, two individuals with compound heterozygous MSH3 loss of function variants were found. Tumors of both patients showed high microsatellite instability in di- and tetranucleotides (EMAST) and immunohistochemical loss of MSH3. Loss of MSH3 protein expression was already shown to be frequent in MSI-H tumors due to a microsatellite in MSH3, but no MSH3 germline mutation carrier had been described until 2016.

Hamartomatous polyposis syndromes

Hamartomatous polyposis syndromes are a rare heterogeneous group of autosomal dominant disorders accounting for less than 1% of all hereditary colorectal cancer syndromes. Hamartomatous polyps are the main characteristic of these syndromes. While these polyps are benign they have the potential to become malignant and progress into carcinomas. This progression is through a hamartoma to carcinoma sequence in which stromal elements create a local environment that promotes epithelial dysplasia. The different syndromes within this group all have different clinical phenotypes, each with different frequencies and location of the polyps, distinct organ-specific manifestations, and predispositions for the development of other malignancies. Proper distinction between these syndromes is of great importance for appropriate clinical management.

Juvenile polyposis syndrome

Juvenile polyposis syndrome (JPS) is characterized by the development of multiple gastrointestinal polyps in the colon. These polyps generally vary in size from 5 mm to 50 mm and typically have a spherical, lobulated and pedunculated appearance. JPS presents in the first or second decade of life, with an average age of diagnosis around 18.5 years. Symptoms for JPS can include rectal bleeding, anemia, abdominal pain, constipation or change in stool size, shape or color, though some JPS patients remain asymptomatic. The cumulative lifetime risk for colorectal cancer is 40-70%. If a pathogenic variant is present, surveillance with colonoscopy or endoscopy should start at the age of 15 years and should be repeated every 3 years. In about 20-60% of JPS patients a pathogenic germline variant in SMAD4 or BMPR1A is found. Both genes are involved in the TGF-beta signaling pathway. The majority of pathogenic germline SMAD4 and BMPR1A variants are missense or small deletions and 15% is deletions of one or more exons.

Peutz-Jeghers syndrome

Peutz-Jeghers syndrome (PJS) is characterized by the presence of hamartomatous polyps in the gastrointestinal tract and distinctive mucocutaneous pigmentation. A typical dark blue to dark brown pigmentation is present in 95% of PJS patients and is seen on the vermilion border of the lips, the buccal mucosa, hands and feet. The small intestine is most commonly affected, although polyps can be found in colon, stomach, rectum, bladder, esophagus and gallbladder. The polyps are generally between 5 to 50 mm in diameter. PJS presents in the second or third decade of life. Symptoms for PJS can include rectal bleeding, anemia, bowel obstruction and abdominal pain. The risk of developing any cancer at age 65 is 47-93%, with an especially high risk of developing
stomach, small intestine, colon and breast cancer but also elevated risk of developing cancer in the esophagus, pancreas, lung, uterus and ovaries. \(^{186,189}\) Surveillance of the large bowel is recommended every three years, starting at age 18 and upper gastrointestinal endoscopies are recommended every three years starting at age 25.\(^ {190}\) Notably, since breast cancer risks are comparable to \(BRCA1/BRCA2\) mutation carriers (40-85% lifetime risk), PJS patients are recommended equal surveillance with monthly breast self-examination starting at age 18 and mammography starting at age 25.\(^ {186,189}\)

Pathogenic germline variants in \(STK11\) gene (also known as \(LKB1\)) are found in 30-80% of PJS patients.\(^ {184-186,188-192}\) \(STK11\) is a serine-threonine kinase involved in regulation of cellular proliferation via G1 cell-cycle arrest, in WAS1 signaling and P53 mediated apoptosis.\(^ {185}\) A clinical diagnosis of PJS is made when a patient has at least two PJS polyps or one PJS polyp with a positive family history or mucocutaneous pigmentation.\(^ {185,192}\) The mutation detection rate in patients who meet these criteria is 70-90% and the majority of variants detected are nonsense or frameshift deletions resulting in a truncated protein.\(^ {185,186,191}\)

**PTEN hamartoma tumor syndrome**

PTEN hamartoma tumor syndrome (PHTS) is an autosomal dominant disorder caused by germline variants in the tumor suppressor gene \(PTEN\).\(^ {193}\) PHTS encompasses multiple overlapping syndromes as Cowden Syndrome (CS) and Bannayan-Riley-Ruvalcaba syndrome (BRRS).\(^ {177,193}\) \(PTEN\) is a tumor suppressor gene with multiple, but incompletely understood, functions. As a lipid phosphatase it plays a role in the P13K/Akt signaling, involved in G1 cell-cycle arrest and apoptosis. As a protein phosphatase it can regulate cell-survival pathways.\(^ {194-196}\) It furthermore might play a role in cellular migration and focal adhesion.\(^ {194,196}\)

CS is characterized by macrocephaly, mucocutaneous lesions, acral keratosis, papillomas and fibromas, and 30-85% of patients develop hamartomatous polyps in the gastrointestinal tract.\(^ {178,194,197,198}\) CS patients have an increased lifetime risk of developing breast (25-50%), endometrial (5-28%), thyroid (3-17%), colon (9-16%), skin (6%) and renal cancers.\(^ {198,199}\) With a range of possible tumors, management recommendations are broad, including yearly breast examinations from age 25-30, yearly thyroid examination or ultrasound starting at age 18 and colonoscopies every 5 years from age 35.\(^ {197,199}\) BRRS patients develop lipomas, gastrointestinal hamartomatous polyps, macrocephaly, hemangiomas and developmental delay.\(^ {177}\) In 11 – 80% of patients meeting clinical criteria for PHTS, a pathogenic variant in \(PTEN\) is found.\(^ {178,193,194,198,199}\) In patients without a \(PTEN\) variant, pathogenic germline variants in \(SDHB, SDHC, SDHD, AKT, PIK3CA\) as well as hypermethylation of \(KLLN\) are found explaining the PHTS-like phenotype.\(^ {199,200}\)
Thesis outline

Over the last few years, advances have been made in discovering the underlying genetic cause in unexplained CRC and polyposis patients. Three new CRC and polyposis syndromes—polymerase proofreading associated polyposis (PPAP), NTHL1-associated polyposis (NAP) and MSH3-associated polyposis—were discovered in the last three years, and new genes are still being described. The aim of this thesis was to find the underlying genetic cause in a large cohort of unexplained suspected Lynch Syndrome patients and a cohort of unexplained polyposis patients. We hypothesized that the suspected Lynch Syndrome (sLS) patients could be explained by missed variants in the mismatch repair (MMR) genes, by bi-allelic somatic inactivation of the MMR genes or by variants in other susceptibility genes.

Chapter 2 reports a whole gene capture effort in which we screened sLS patients for variants in the exonic- or intronic regions of 15 CRC susceptibility genes, including MLH1, MSH2, MSH6 and PMS2. Chapter 3 describes variants in the polymerase genes POLE and POLD1 in sLS patients. Variants in the exonuclease domain of these genes result in hypermutated tumors. We hypothesize that the MMR-deficiency in these tumors is due to secondary MMR hits resulting from this hypermutated phenotype. Chapter 4 shows a splicing assay to analyse the effect of variants (predicted) to result in splicing. For this assay RNA was isolated from formalin-fixed paraffin-embedded (FFPE) tissue, enabling analysis of somatic variants and variants in patients of which only FFPE is available. Chapter 5 describes a practical guide on detecting and analysing variants in the PMS2 gene in DNA isolated from FFPE. Analysis of this gene is complex due to the presence of many pseudogenes. Chapter 6 focuses on unexplained polyposis patients. Half of unexplained patients with 20-100 adenomas could be explained by a mosaic APC variant, either present in leukocyte and colon, confined to the colon, or only detected in the adenomas but not in normal colonic mucosa of a patient. Finally, concluding remarks and future perspectives are presented in Chapter 7.
References

79. Castro MP, Goldstein N. Mismatch repair deficiency associated with complete remission to combination programmed cell death ligand immune therapy in a patient with sporadic urothelial carcinoma: immunothera-
82. Syngal S, Fox EA, Eng C, et al. Sensitivity and specificity of clinical criteria for hereditary non-polyposis colo-
84. Hampel H, Frankel WL, Martin E, et al. Screening for the Lynch syndrome (hereditary nonpolyposis colorec-
General introduction and thesis outline

169. Domingo E, Freeman-Mills L, Rayner E, et al. Somatic POLE proofreading domain mutation, immune re-


