Germline Genetics in Colorectal Cancer Susceptibility and Prognosis

Amanda Ewart Toland
The Ohio State University, USA

1. Introduction

Population-based studies indicate that approximately 35% of an individual’s risk of developing colorectal cancer (CRC) is due to inherited genetic factors (Lichtenstein et al. 2000). Indeed, approximately 50,000 individuals diagnosed with CRC in the United States each year will have at least one other family member with CRC (Kaz & Brenthall, 2006). Classically, genetic susceptibility to CRC is described as three types: Low, moderate and high risk. In reality, the risk of developing colorectal cancer due to genetic factors exists on a continuum from very low to very high risk (Figure 1). In addition to colon cancer susceptibility, genetic variants are likely to play a role in response to therapy and prognosis of colon cancer. This Chapter will provide an overview of the current knowledge in genetic susceptibility to hereditary and non-hereditary CRC, the complexities and issues around the identification of germline genetic risk factors, and the current and future use of genetic information in the clinic.

High penetrance risk includes inheritance of mutations in genes which segregate in a Mendelian fashion in families and confer a high lifetime risk of disease. Hereditary cancer syndromes are those in which a mutation confers a high lifetime risk of developing CRC. Several syndromes have been described. The two most familiar CRC syndromes are Familial Adenomatous Polyposis (FAP) and Lynch Syndrome (LS) (Table 1). Moderately-penetrant mutations are mutations or polymorphic variants which can manifest as colon cancer clustering in families but without a clear cancer syndrome or inheritance pattern. Low-penetrance variants are those which are present in a reasonably high frequency in the general population, but have small influences on risk and are not themselves sufficient for the development of colon cancer.

2. Hereditary colorectal cancer syndromes

Hereditary colorectal cancer syndromes are those in which an inherited or de novo germline mutation confers a high lifetime risk of developing CRC. Approximately 5% of all CRC diagnoses are thought to be due to highly penetrant mutations (Bodmer, 2006; de la Chapelle, 2004). These familial mutations were the first germline genetic alterations to be discovered to be important for CRC risk. Several syndromes have been described. They can be subdivided into syndromes with adenomatous polyps, those with hamartomatous polyps and syndromes with polyps of mixed histopathology (Table 1). Syndromes that present with
a few or many adenomatous polyps include FAP, LS and MUTYH-associated polyposis (MAP) (Table 1). The hamartomatous polyp syndromes include Cowden Syndrome, Juvenile Polyposis and Peutz-Jeghers Syndrome. Not all genes contributing to hereditary CRC have been identified and characterized. It is likely as genome-wide and exonic sequencing become more common that additional rare mutations leading to hereditary colorectal cancer will be discovered.

Fig. 1. Genes and variants associated with CRC risk. The allele frequency and lifetime risk of genes associated with familial colorectal cancer (red) and variants associated with a small increase in lifetime risk of CRC (blue) are illustrated. (The Y179C variant is associated with a low risk of CRC when individuals only carry one mutated \textit{MUTYH} allele.)

One common feature of all of these CRC syndromes is that the age of diagnosis of cancer tends to be much earlier than that in the general population. Typically colorectal cancer occurs 10-20 years earlier in individuals with these syndromes than in the general population. In individuals who carry a mutation in one of these hereditary CRC genes there is also considerable increased risk, up to 99%, of developing colon cancer. Individuals with most of these syndromes have detectable polyps prior to the onset of colorectal cancer; however many probands that were not undergoing regular screening are brought to attention following a diagnosis of CRC.

A clinical diagnosis of a specific cancer syndrome can be made based on an extensive family history in combination with histological and pathological information about the number and type of polyps. More definitive diagnoses are made by genetic testing coordinated through a genetic counsellor or medical geneticist of an affected proband, or, in the case of LS, analysis of tumours for loss of one of the LS-related proteins. Even though individually these syndrome are rare, proper management and diagnoses can impact the incidence and
mortality related to CRC. Because mutations leading to these genes confer a high lifetime risk of CRC, knowledge of one’s family history and mutation status influences medical management and can significantly reduce the incidence of CRC.

<table>
<thead>
<tr>
<th>Syndrome</th>
<th>CRC Risk</th>
<th>Unique characteristics</th>
<th>Other features</th>
<th>Gene(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HNPCC</td>
<td>20-75%*</td>
<td>Microsatellite instability of tumors</td>
<td>Endometrial, gastric, ovarian, small bowel, biliary &amp; urothelial ca</td>
<td>MSH1, MLH2, PMS2, MSH6 EPCAM</td>
</tr>
<tr>
<td>FAP</td>
<td>99%</td>
<td>Hundreds of polyps</td>
<td>Desmoids, hepatoblastoma &amp; papillary thyroid ca, CHRPE</td>
<td>APC</td>
</tr>
<tr>
<td>Peutz-Jeghers</td>
<td>39%</td>
<td>Hamartomatous polyps</td>
<td>Gastric, breast &amp; ovarian ca, sex cord tumors, mucocutaneous hyperpigmentation</td>
<td>LKB1 (STK11)</td>
</tr>
<tr>
<td>MAP</td>
<td>80%</td>
<td>Recessive inheritance, many adenomatous polyps</td>
<td>Two common mutations in individuals with Northern European ancestry: Y179C and G396D</td>
<td>MUTYH (MYH)</td>
</tr>
<tr>
<td>Hereditary Mixed Polyposis</td>
<td>ND</td>
<td>Inflammatory &amp; metaplastic polyps</td>
<td></td>
<td>BMPR1A</td>
</tr>
<tr>
<td>Cowden</td>
<td>None</td>
<td>Hamartomatous polyps</td>
<td>Macrocephaly, benign &amp; malignant thyroid, breast &amp; uterine neoplasms</td>
<td>PTEN</td>
</tr>
<tr>
<td>Juvenile Polyposis</td>
<td>39%</td>
<td>Hamartomatous polyps</td>
<td>GI polyps, gastric ca</td>
<td>SMAD4, BMPR1A,</td>
</tr>
</tbody>
</table>

Table 1. Hereditary Colon Cancer and Polyposis Syndromes. ca, cancer; CHRPE, congenital hypertrophy of the retinal pigment epithelium; ND, not determined; GI, gastrointestinal; *CRC risk depends on which gene is mutated.

2.1 Hereditary non-polyposis colorectal cancer/lynch syndrome

Hereditary non-polyposis colorectal cancer (HNPCC), also known as Lynch syndrome (LS), is the most common hereditary CRC syndrome associated with a strong family history of colon cancer. Approximately 2-4% of all colon cancers are due to LS (Hampel et al., 2008). Although individuals with LS frequently have adenomatous polyps, individuals with this syndrome do not have polyposis. In addition to a lifetime risk of approximately 50-80% for colorectal cancer, there is an increased risk of ovarian, endometrial and gastric cancers (Jasperson et al., 2010). Since the genes have been identified, some conditions which were once thought to be distinct (e.g. Muir-Torre syndrome which is characterized by familial...
CRC with sebaceous neoplasms and Turcot syndrome which is characterized by CRC with glioblastomas) are variants of LS. LS is inherited in an autosomal dominant fashion and is associated with mutations in four genes important in mismatch repair (MMR): *MLH1*, *MSH2*, *PMS2* and *MSH6*. Tumours arising in individuals with germline mutations in these genes, which are important for DNA repair, typically exhibit microsatellite instability which is sometimes used clinically to aid in a diagnosis. The most commonly mutated genes in LS are *MLH1* and *MSH2*. Recently, deletions of the 3’end of *EPCAM*, a gene mapping 5’ of the *MSH2* gene, have been found to give rise to LS by causing methylation of the *MSH2* gene in about 6% of LS cases (Niessen et al., 2009a).

Criteria based on family history, Amsterdam I, Amsterdam II and Bethesda criteria, were developed in order to identify families for further evaluation of Lynch syndrome. Amsterdam I criteria, the first described, includes three first degree relatives with CRC, two or more generations affected, one family member with CRC diagnosed under the age of 50 and FAP must be ruled out (Vasen et al., 1991). Amsterdam II criteria are the same as Amsterdam I except that endometrial, small bowel or other LS-related cancers can be substituted for CRC (Vasen et al., 1999). Revised Bethesda criteria include one CRC diagnosed under the age of 50, one CRC under the age of 60 with evidence of microsatellite instability, CRC or LS-related cancer in at least one first degree relative under the age of fifty or CRC or LS-related cancers in at least two first or second degree relatives at any age (Umar et al., 2004). In addition, three online prediction programs MMRpredict, PREMM, and MMRpro have been developed to identify families with LS. Sensitivities of the online models and revised Bethesda criteria are about 75% with a range of specificity from 50-60%; the Amsterdam II criteria have a lower sensitivity of 37.5%, but a better specificity of 99% (Tresallet et al., 2011). Many hospitals have begun to screen all colon cancer cases by immunohistochemistry for loss of the four MMR proteins or for microsatellite instability regardless of family history which has a higher sensitivity than the family history based models (Hampel et al., 2008). Individuals whose tumours show absence of *MLH1*, *MSH2*, *PMS2* and *MSH6* and/or microsatellite instability are referred for genetic evaluation of LS and, in some cases, additional testing to rule out somatic events specific to the tumours which lead to loss of the proteins.

The risk and spectrum of cancer depends on which LS gene is mutated. The lifetime colon cancer risk is 97% for males and 53% for females with germline *MLH1* mutations. Endometrial cancer risk associated with *MLH1* mutations ranges from 25 to 33% (Ramsoekh et al., 2009; Stoffel et al., 2009). The lifetime risk of CRC in male *MSH2* mutation carriers is 52% and is 40% for females. There is a 44-49% risk of endometrial cancer for female *MSH2* mutation carriers (Stoffel et al., 2009). About 10% of Lynch syndrome families have a mutation in *MSH6* (Talseth-Palmer et al., 2010). The estimated risk of colorectal cancer in individuals with a *MLH6* mutation is lower than some of the other mutations at 30-61% and the risk of endometrial cancer is higher with 65-70% of females developing endometrial cancer by the age of 70 (Talseth-Palmer et al., 2010; Ramsoekh et al., 2009). *PMS2* mutations are less frequently the cause of LS accounting for only 2 to 14% of LS cases (Senter et al., 2008; Niessen et al., 2009b; Talseth-Palmer et al., 2010). The cumulative risk by the age of 70 of developing a colon cancer in mono-allelic *PMS2* mutation carriers is 15-20%, endometrial cancer is 15% and other Lynch-related cancers is from 25-32% (Senter et al., 2008). Bi-allelic mutations in the MMR genes have been observed and lead to a severe phenotype with childhood brain tumours, leukaemia, and LS-associated tumors known as Constitutional MMR Deficiency (Senter et al., 2008; Felton et al., 2007; Wimmer & Kratz, 2010). *EPCAM*
deletions show a comparable risk of colon cancer to MLH1 and MSH2 mutation carriers, but a decreased risk of endometrial cancer. In families with EPCAM deletions there is a 75% risk of developing cancer by the age of 70 and a 12% risk of endometrial cancer (Kempers et al., 2011). As the risks of colon, endometrial and other cancers are high, regardless of the gene, individuals with LS are recommended to follow more intensive cancer screening guidelines.

2.2 Familial adenomatous polyposis
Familial adenomatous polyposis (FAP) is an autosomal dominant CRC syndrome that accounts for less than 1% of all CRC diagnoses (Burt, 2000). In classic FAP, individuals develop hundreds to thousands of adenomatous polyps beginning in the early to mid-teen years. By age 35, 95% of individuals with FAP have polyps and the penetrance of colorectal cancer associated with this disease is over 90% (Petersen et al., 1991) which is why prophylactic colectomy is recommended in individuals who have this syndrome. FAP is caused by mutations in the adenomatous polyposis coli (APC) gene, although a high percentage (12-20%) of individuals with a clinical diagnosis of classical FAP do not have identifiable mutations in APC or another polyposis gene, MUTYH, (de la Chapelle, 2004; Filipe et al., 2009). Other features of FAP include small bowel adenomas, gastric polyposis, congenital hypertrophy of the retinal pigment epithelium (CHRPE), and desmoids (Allen & Terdiman, 2003). Desmoids cause significant mortality in FAP despite their non-malignant nature. Attenuated FAP is a milder form of this disease in which there are fewer polyps and a later onset of disease. Two other variants of FAP are Turcot syndrome which includes FAP and central nervous syndrome tumors, primarily medulloblastomas, and Gardner syndrome which includes soft-tissue tumours, osteomas and dental abnormalities.

2.3 MUTYH-associated polyposis
MUTYH-associated polyposis (MAP) is an autosomal recessive syndrome conferring a 43 to nearly 100% lifetime risk of CRC (Farrington et al., 2005; Lubbe et al., 2009). Penetrance for cancer in bi-allelic mutation carriers is estimated to be 20% at 40 years and 43% at 50 years of age (Lubbe et al., 2009). Individuals typically present with 10-100 adenomatous polyps, although some bi-allelic mutation carriers do not have any polyps on screening (Farrington et al., 2005; Nielsen et al., 2007). Polyposis of the duodenum can also be observed. Between 24 and 56% of FAP and attenuated-FAP families lacking mutations in APC have been found to carry bi-allelic mutations in MUTYH, suggesting that mutations in the two genes account for a significant proportion of familial polyposis (Nielsen et al., 2007; Gomez-Fernandez et al., 2009). Two common MUTYH mutations comprising 80-85% of disease causing mutations in Caucasians of Northern European ancestry are Tyr179Cys and Gly396Asp (previously known as Y165C and G382D; Al-Tassan et al., 2002). Importantly, 4% of bi-allelic mutation carriers will not have either of the two common mutations (Goodenberger et al., 2011). The mutation frequency of these mutations varies between populations and other founder and relatively frequent mutations have been identified (Gomez-Fernandez et al., 2009).

2.4 Other adenomatous polyposis syndromes
A handful of case reports of mutations leading to unique or rare familial presentation of CRC exist which may explain a small proportion of polyposis families that do not have APC or MUTYH germline mutations. One recent description is of homozygous mutations in BUB1B leading to CRC (Rio Frio et al., 2010). This gene has not been extensively tested in
polyposis families so it is unknown if it will contribute much to the overall risk of familial adenomatous polyposis. Mutations in the \textit{AXIN2} gene are associated with tooth agenesis-colorectal cancer syndrome in a large Finnish family (Lammi et al., 2004), but mutations in this gene do not appear to account for a large proportion of hereditary CRC.

\subsection*{2.5 Familial Colorectal Cancer Type X}
About half of the families that have a strong-family history of colorectal cancer suggestive of LS by Amsterdam or Bethesda criteria have no evidence of mismatch repair deficiency or loss of any of the HNPCC related proteins in tumours (de la Chapelle & Lynch, 2003). To be classified as Familial Colorectal Cancer Type X, families must meet Amsterdam I criteria and have no evidence of MMR deficiency. A closer look at these pedigrees shows that they tend to have fewer individuals diagnosed with cancer, their cancers are less likely to look like those in LS and their average age of diagnosis is older than those in families with MMR mutations (Jass et al., 1995; Lindor et al. 2005). The genes contributing to Familial Colorectal Cancer Type X are as yet unknown.

\subsection*{2.6 Peutz-Jeghers syndrome}
Peutz-Jeghers syndrome is a rare autosomal dominant condition with an estimated incidence of 1 in 200,000 births first described by Peutz in 1921. It is characterized by childhood onset of hamartomatous polyps in the gastrointestinal tract and by mucocutaneous pigmentation of the lips and buccal mucosa. Mutations in \textit{LKB1 (STK11)} are the only known cause of Peutz-Jeghers syndrome. \textit{LKB1} mutations have been found in 50-94\% of individuals with classic features of this disorder indicating that there may be locus heterogeneity (Boardman et al. 2000; Volikos et al., 2006; Aretz et al., 2005). The penetrance for GI polyps in this syndrome is 100\%. There is also a 76-85\% lifetime risk of cancer which includes lifetime risks of 40\% for colon cancer, 30-60\% for gastric cancer, 15-30\% for small intestinal cancer and 11-35\% for pancreatic cancer (Hearle et al., 2006, van Lier et al., 2010; van Lier et al., 2011). Breast and gynaecological cancers can be seen at high frequencies. There is a high mortality associated with this syndrome with a median age of death at 45 years of life, mostly due to cancer or bowel intussusceptions (van Lier et al., 2011).

\subsection*{2.7 Cowden syndrome}
Cowden syndrome is an autosomal dominant syndrome with features of skin lesions, macrocephaly, thyroid manifestations and hamartomatous polyps of the GI tract. It is caused by mutations in the \textit{PTEN} gene. Although this disorder and an allelic disorder Bannayan-Ruvalcaba-Riley Syndrome are both characterized by many hamartomatous polyps, there is no clear increased risk of colon cancer associated with Cowden syndrome.

\subsection*{2.8 Juvenile Polyposis}
Hereditary juvenile polyposis (JP) is defined as the presence of 10 or more juvenile polyps. These polyps are primarily hamartomatous. The typical age of diagnosis is between the ages of 5 and 15 years (Merg & Howe, 2004). Most children are brought to medical attention because of colorectal bleeding. The risk of CRC associated with JP varies from 20-50\% depending on the study and the gene which is mutated. (Handra-Luca et al., 2005). In addition to CRC, there is a significant risk of upper GI cancers. JP is inherited as an autosomal dominant syndrome. Multiple genes have been implicated in this disorder. The
majority of mutations in individuals with JP have been found in SMAD4 and BMPR1A. Individuals with BMPR1A mutations have a higher risk of cancers including those of the stomach, pancreas and small bowel. Mutations have also been found in PTEN, but these may be misdiagnosed cases of Cowden syndrome. There have been reports of SMAD4 mutations in families with features of both juvenile polyposis and hereditary hemorrhagic telangiectasia (Gallione et al., 2004). Not all individuals with features of JP have identifiable mutations implicating additional as yet unidentified genes (Handra-Luca et al., 2005).

2.9 Hereditary mixed polyposis
Hereditary mixed polyposis (HMP) is an autosomal dominant condition characterized by polyps of mixed histology including adenomatous, hyperplastic and atypical juvenile types. Mutations in BMPR1A have been found in a proportion of families presenting with polyps of mixed type (Cheah et al., 2009). Despite the observation that families with both juvenile polyps and hereditary mixed polyposis can have mutations in BMPR1A, families with HMP are less likely to have juvenile type polyps and have an older age of diagnosis in adulthood (Merg & Howe, 2004). A locus for HMP, called CRAC1 or HMPS, has been mapped to chromosome 15q13-q14 in multiple Ashkenazi Jewish families, but the gene has not yet been identified (Jaeger et al., 2008).

2.10 Hyperplastic polyposis (HPP)
Hyperplastic polyposis syndrome (HPP) is not yet well defined, but is characterized by multiple or large hyperplastic or serrated polyps and an association with an increased risk of CRC. The range of polyps has been described from 5 to over 100 and the pathology of the polyps can be diverse. HPP is often diagnosed in the fifth through seventh decade of life. The frequency of CRC in individuals with HPP ranges from 25-50% (Lage et al., 2004; Kalady et al., 2011). About 30% of individuals with HPP have a family history of CRC. The inheritance pattern of HPP is not well defined, but a few characterized families show possible recessive inheritance (Young & Jass, 2006). A germline mutation in EPHB2 was identified in an individual who had more than 100 hyperplastic polyps, but EPHB2 mutations have not been observed in other HPP cases (Kokko et al., 2006). Thus, the causal genes for most individuals with HPP have yet to be identified.

2.11 Rare cancer predisposition syndromes and risk of colon cancer
Whereas many hereditary cancer syndromes have specific cancers which occur at greater frequency than the general population, a few syndromes have elevated risks of many different types of cancer. Li-Fraumeni syndrome (LFS) is a rare autosomal dominant inherited condition caused by germline mutations in TP53. The classical types of cancer seen in individuals with LFS include breast cancer, sarcomas, brain tumors, leukemia and adrenal cortical tumours; however, there is also an increased risk of colon cancer of 2.8-fold over the general population (Ruijs et al., 2010).

3. Moderate risk alleles
Familial clusters of colon cancer account for approximately 20% of all CRC cases, however, most of these cases will not be due to the known CRC syndromes (Burt et al., 1990). A familial cluster is multiple individuals within families who have a similar presentation, but
no clear inheritance pattern of disease transmission. The risk of colon cancer is increased to individuals who have a relative with CRC or adenomas; first-degree relatives of affected individuals have a two- to three-fold increase in risk and when more than one first-degree relative is affected the risk increases to nearly four-fold (Butterworth et al., 2006; Taylor et al. 2010). To date, moderate-risk alleles (ORs of 3-5) have not been identified. It is possible that some families exhibiting clustering of CRC may have multiple low-penetrance alleles which work synergistically to increase risk.

4. Low-penetrance risk alleles

The majority of genetic risk for CRC in the population is likely to be due to low-penetrance susceptibility alleles which act with other low-penetrance variants and the environment. A debate in the field is whether most of the genetic risk will be due to common variants with low effects and allele frequencies greater than 1% or rare or unique variants with low to moderate effects (Bodmer, 2006). Historically, variants conferring an increased risk of CRC in the general population have been identified through cohort or population-based case/control studies looking at candidate genes, but recent genome-wide association studies (GWAS) have been quite successful in identifying well-replicated variants conferring risk. Whereas a great many studies have identified positive associations with some of these genes, the vast majority have not been consistently replicated. Lack of replication does not mean in all cases that the initial association is faulty, but could be due to differences in populations leading to differences in allele frequencies or linkage disequilibrium, environmental exposures, study design or underpowered replicate studies. Whereas the low-penetrance variants identified to date are not particularly predictive for CRC risk on their own, several direct-to-consumer genetic testing companies include some of these variants in their analysis of genomic risks of common disease. We will highlight some of the different types of variants which have been identified through multiple types of studies as showing evidence of contribution to CRC risk.

4.1 Candidate-gene studies

The studies to assess the risk of DNA variants, mainly single nucleotide polymorphisms (SNPs), have been association case/control studies or cohort studies testing SNPs in genes with relevant biological function for CRC. Many such candidate gene studies for CRC risk have been completed. Some of these have been replicated in one or two studies, but few have stood up to repeated replication studies or meta-analyses. A meta-analysis of 50 published CRC association studies for common alleles in 13 genes found three variants: APC I1307K, a HRAS1 repeat variant and MTHFR 677V were convincingly associated with modest CRC risk in the general population (Houlston & Tomlinson, 2001). Other genes with variants showing CRC risk in multiple studies include NAT1, NAT2 and TGFBRII (Burt, 2010). It is possible that some of these are real associations, but are population-specific or depend upon gene-environment interactions present only in certain individuals.

Candidate genes for CRC case/control studies have been chosen in a variety of ways. Variants and genes studied include common variants in high-risk genes, genes in pathways believed to be important in CRC and genes in pathways linked to environmental factors associated with CRC. One strategy which has been under-utilized is to map loci for cancer susceptibility in the mouse and then test these genes/loci for cancer risk in human
4.2 Variants of high-risk genes

Once genes for hereditary cancer syndromes began to be identified, researchers hypothesized that common variants in these genes contribute to cancer risk in the general population. Studies on common variants of almost all hereditary CRC predisposition genes have been assessed, but only a handful of variants in these high risk genes have been determined to contribute to sporadic CRC risk.

4.2.1 Carriers of MUTYH mutations

Bi-allelic mutations in MUTYH lead to MAP as described above. Several studies have looked at the cancer risks associated with mono-allelic mutations in MUTYH. The range of cancer incidence associated with the Y179C allele is between 1.27 to 2-fold (Table 2, Jones et al., 2009; Tenesa et al., 2006; Theodoratou et al., 2010). As a result, colon adenomas or cancer may be seen in multiple generations in families with MAP.

4.2.2 APC I1307K

One frequently described modest-risk allele is the I1307K variant in the APC gene. This variant is seen in approximately 6% of individuals of Ashkenazi Jewish (AJ) ancestry and 28% of AJ individuals with a family history of CRC (Laken et al., 1997). Carriers of the I1307K allele have 1.5 to 2-fold increased risk of CRC compared to individuals who are non-carriers (Table 2, Dundar et al., 2007; Gryfe et al., 1999). The variant itself is not thought to change function of the APC gene, however, the change results in a stretch of eight consecutive adenosines. During replication this polyadenosine track is thought to have increased risk of somatic mutation due to polymerase slippage. Addition or loss of a single nucleotide then results in a frame-shift and non-functional APC gene (Laken et al., 1997). As the age of onset of CRC related to this polymorphism does not appear to differ from sporadic CRC, cancer screening beyond general population recommendations is not typically done (Petersen et al., 1999).

4.2.3 Bloom’s syndrome mutation carriers

Bloom’s syndrome is a rare autosomal recessive condition characterized by abnormal rates of sister chromatid exchange, growth deficiency and an increased incidence of multiple types of cancer. One in 107 individuals of AJ ancestry carries a founder mutation, designated as BlmAsh, in the Bloom’s syndrome gene, BLM (Li et al., 1998). Early studies suggested a 2-fold increase in colon cancer risk in BlmAsh carriers, but this has not held up in subsequent studies (Gruber et al, 2002; Cleary et al., 2003).

4.2.4 CHEK2

CHEK2 is a gene important in response to DNA damage. Studies have demonstrated increased risk of breast cancer with an 1100delC polymorphism but have been inconclusive.
for the role of the 1100del C variant in CRC risk. However, another variant, I157T, has been associated with Lynch-syndrome like cancers and confers an increased risk of 1.4 to 2-fold of CRC (Kilpivaara et al., 2006).

4.3 Genome-wide association studies (GWAS)
With technological advances allowing the genotyping of up to millions of SNPs, the ability to interrogate the entire genome without bias has led to the identification of SNPs with reproducible, but small associations with cancer risk. The strength of these studies is that very large numbers of samples were used, large independent replication studies have been completed and very low p-values were required to meet genome-wide significance. About 18 SNPs have been identified to date (Table 2). Interestingly, although many map near genes, none of them fall within coding regions of genes suggesting that these SNPs may play a role in gene regulation. Despite the confidence that these are “real”, the variants identified through GWAS to date explain a very small percentage of the overall risk of CRC ascribed to genetics. One computational assessment estimates that there may be as many as 170 low-penetrance variants which contribute to CRC risk (Tenesa and Dunlop, 2009).

4.3.1 8q24 and rs6983267
Two of the first GWAS studies for CRC identified a variant, rs6983267, on chromosome 8q24 which was significantly associated with CRC risk (Tomlinson et al. 2007; Haiman et al. 2007). The effects were modest with ORs ranging from of 1.14 to 1.24. Additional variants on 8q24 including rs7014346, rs783728, and rs10505477 were also identified in subsequent screens (Tenesa et al. 2008; Poynter et al. 2007). Several groups have replicated these findings and show a consistent effect of the rs6983267 variant. This SNP falls into a gene-poor region on 8q24 with the closest gene, cMYC, 335 kb away. One study showed that the rs6983267 variant falls within an enhancer element and alleles differentially bind a WNT-related transcription-factor 7-like 2 (TCF7L2). However, correlation with expression of MYC has not been detected, and the exact role of this SNP in colon cancer development has yet to be definitively determined (Pomerantz et al. 2009; Tuupanen et al. 2009).

4.3.2 GWAS variants in the Bone Morphogenic Protein (BMP)/Transforming Growth Factor Beta (TGFβ) pathway
Multiple variants identified by GWAS (rs4444235, rs4939827, rs4779584, rs961253 rs10411210, rs4925386) are located near genes involved in BMP and/or TGFβ signalling (Tenesa & Dunlop, 2009). BMP proteins are positive regulators of the Wnt pathway which have long been known to be important in CRC tumorigenesis. TGFβ is a master signalling molecule controlling many processes important in cancer and cancer suppression. There is considerable overlap and interaction between the BMP and TGFβ pathways. Germline mutations in SMAD4, BMPR1A and APC are associated with specific hereditary colon cancer syndromes. Whereas the exact effect of these SNPs on the nearby BMP/TGFβ signalling genes is not determined, location and number of these SNPs suggest that perturbation of these pathways may be critical for CRC risk in the general population.

4.4 Population specific risk factors
One of the caveats to many of the candidate gene and GWAS studies for CRC risk to date is that they have been predominantly completed in Caucasian populations. The rs6983267
<table>
<thead>
<tr>
<th>SNP</th>
<th>Locus</th>
<th>Nearby genes</th>
<th>OR</th>
<th>Type of Study</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rs6691170</td>
<td>1q41</td>
<td>DUSP10</td>
<td>1.06</td>
<td>GWAS</td>
<td>Houlston et al., 2010</td>
</tr>
<tr>
<td>Rs6687758</td>
<td>1q41</td>
<td>DUSP10</td>
<td>1.09</td>
<td>GWAS</td>
<td>Houlston et al., 2010</td>
</tr>
<tr>
<td>Rs10936599</td>
<td>3q26.2</td>
<td>MYNN</td>
<td>0.93</td>
<td>GWAS</td>
<td>Houlston et al., 2010</td>
</tr>
<tr>
<td>Rs16892766</td>
<td>8q23.3</td>
<td>EIF3H</td>
<td>1.27</td>
<td>GWAS</td>
<td>Tomlinson et al., 2008</td>
</tr>
<tr>
<td>Rs6983267</td>
<td>8q24</td>
<td>Gene desert</td>
<td>1.21-1.23</td>
<td>GWAS</td>
<td>Tomlinson et al. 2007; Xiong et al. 2011</td>
</tr>
<tr>
<td>Rs7014346</td>
<td>8q24</td>
<td>POU5F1P1, DQ515897</td>
<td>1.19</td>
<td>GWAS</td>
<td>Tenesa et al., 2008</td>
</tr>
<tr>
<td>Rs719725</td>
<td>9p24</td>
<td>Several</td>
<td>1.46</td>
<td>GWAS</td>
<td>Poynter et al. 2007</td>
</tr>
<tr>
<td>Rs10795668</td>
<td>10p14</td>
<td>FLJ3802842</td>
<td>1.23</td>
<td>GWAS</td>
<td>Tomlinson et al., 2008; Xiong et al., 2011</td>
</tr>
<tr>
<td>Rs3802842</td>
<td>11q23</td>
<td>Several</td>
<td>1.11-1.29</td>
<td>GWAS</td>
<td>Pittman et al., 2008; Tenesa et al., 2008; Xing et al., 2011</td>
</tr>
<tr>
<td>Rs11169552</td>
<td>12q13.13</td>
<td>LARP4, DIP2B</td>
<td>0.92*</td>
<td>GWAS</td>
<td>Houlston et al., 2010</td>
</tr>
<tr>
<td>Rs7136702</td>
<td>12q13.13</td>
<td>LARP4, DIP2B</td>
<td>1.06</td>
<td>GWAS</td>
<td>Houlston et al., 2010</td>
</tr>
<tr>
<td>Rs4444235</td>
<td>14q22.2</td>
<td>BMP4</td>
<td>1.11</td>
<td>GWAS</td>
<td>Houlston et al., 2008</td>
</tr>
<tr>
<td>Rs4779584</td>
<td>15q13.3</td>
<td>CRAC1/GREMI</td>
<td>1.23</td>
<td>GWAS</td>
<td>Jaeger et al., 2008</td>
</tr>
<tr>
<td>Rs9929218</td>
<td>16q22.1</td>
<td>CDH1</td>
<td>0.91*</td>
<td>GWAS</td>
<td>Houlston et al. 2008</td>
</tr>
<tr>
<td>Rs4939827</td>
<td>18q21</td>
<td>SMAD7</td>
<td>1.17</td>
<td>GWAS</td>
<td>Tenesa et al., 2008; Xiong et al., 2011</td>
</tr>
<tr>
<td>Rs10411210</td>
<td>19q13.1</td>
<td>RHPN2</td>
<td>0.87*</td>
<td>GWAS</td>
<td>Houlston et al., 2008</td>
</tr>
<tr>
<td>Rs961253</td>
<td>20p12.3</td>
<td>BMP2</td>
<td>1.12-1.37</td>
<td>GWAS</td>
<td>Xiong et al., 2011</td>
</tr>
<tr>
<td>Rs4925386</td>
<td>20q13.3</td>
<td>LAMA5</td>
<td>0.93</td>
<td>GWAS</td>
<td>Houlston et al., 2010</td>
</tr>
<tr>
<td>I1307K</td>
<td>5q21-22</td>
<td>APC</td>
<td>1.5-2.0</td>
<td>Candidate</td>
<td>Dundar et al., 2007; Gryfe et al., 1999</td>
</tr>
<tr>
<td>Y179C (mono-allelic)</td>
<td>1p34.1</td>
<td>MUTYH</td>
<td>1.27-2.0</td>
<td>Candidate</td>
<td>Tenesa et al., 2006; Theodoratou et al., 2010</td>
</tr>
</tbody>
</table>

Table 2. Low-penetrance variants associated with CRC Risk

OR, odds ratio; *The common allele is the risk allele.

variant has been replicated in multiple ethnic groups including African-American and Chinese populations with fairly consistent results (Xiong et al., 2011; He et al. 2011). However, rs3802842 on chromosome 11q23 is associated with no increased risk in Japanese populations. A GWAS performed in Japanese CRC cases identified a novel variant, rs7758229 on chromosome 6 which has not been linked with risk in Caucasians (Cui et al., 2011).
These examples illustrate that specific variants identified by GWAS are often markers for the causal, as yet unidentified variant, which may be absent or in different linkage disequilibrium patterns with the identified SNP in other populations. Additional possibilities for the differential risk effects include different frequencies of important interacting variants or population specific gene-environment interactions. When low-penetrance variants become a part of determination of cancer risk, care should be taken to ensure that only variants which have been validated in the ethnic background of the patient be utilized.

4.5 Missing heritability: Additive effects, gene-gene and gene-environmental factors

When GWAS were first utilized they were hailed as the tool by which all low-penetrance variants for disease risk would be identified. While these screens have been successful, the variants identified to date explain less than 10% of the estimated genetic contributions to CRC which makes use of known variants for risk prediction difficult. Several possible explanations for the “missing heritability” exist. One is that much of the genetic risk of CRC will be due to rare or unique low-penetrance variants which are not detectible by population-based GWAS. A second is that synergistic or epistatic gene-gene interactions which are only detectible when taking into account interacting loci will account for the missing heritability. Gene-environmental interactions in which risk is dependent on both a variant and exposure to the environmental risk factor may also play a role. Transgenerational epigenetic effects, epigenetic alterations which are inherited through the germline and/or observed through multiple generations, may also account for some of the missing heritability (Fleming et al., 2008).

As the effects of single variants identified by candidate gene or genome-wide studies have been low, it is important to determine if there are combined effects of carrying multiple risk alleles. One study assessed three SNPs identified through GWAS (rs3802842, rs7014346, rs4939827) and found that carrying all six risk alleles confers a 2.6-fold increased risk of CRC (Tenesa et al., 2008). It is likely that CRC risk could increase if all 18 identified GWAS variants were included in the analysis. Animal models have been instrumental for demonstrating synergistic and epistatic interactions between genetic risk factors (Nagase et al. 2001). Mouse models led to the identification of several CRC susceptibility loci that interact synergistically or epistatically (van Wezel et al., 1996; van Wezel et al. 1999). Thus far, no synergistic or epistatic interactions for CRC risk have been definitively identified in human populations. As computational tools for assessing the data from GWAS studies improve, genetic interactions are likely to be identified as important factors for CRC risk in humans.

Several environmental factors including cigarette smoking, body mass index, polycyclic aromatic hydrocarbons, N-nitroso compounds, and diets high in red meat which increase exposure to heterocyclic amines have been associated with increased risk of CRC (Botteri et al. 2008; Norat et al. 2002; Pischon et al. 2006). Although many studies show contradictory results, interactions between genetic variations and environmental exposures can modify CRC risk. mEH is an enzyme important in xenobiotic activation of tobacco carcinogens. Two variants have been identified in the mEH gene which lead to low or high activity. A meta-analysis of several studies showed that smokers with the mEH3 low metabolizer genotype had lower risk of colon cancer compared to smokers with a mEH3 high metabolizer genotype suggesting that genetic variants can modify the cancer risk associated with smoking (Raimondi et al. 2009). Interactions between dietary factors and genotypes have also been observed for CRC. In one study, individuals who consumed browned red meat...
and had the 751Lys/Lys and 312 Asp/Asp genotypes in the XPD gene were at highest risk of developing CRC (Joshi et al., 2009). Much work remains to be done to fully assess gene-environment interactions for CRC.

### 4.6 Modifier genes for penetrance in high-risk syndromes

The risk of cancer does not reach 100% even for individuals with mutations leading to high-penetrance CRC syndromes like LS. Thus, it has been proposed that even in the context of a mutation, environmental factors and low-penetrance variants may impact cancer risk. To this end, modifier genes, alleles that modify the risk of a mutation, have been sought. Most of the studies to date have been for LS. A variant in CHEK2, 1100delC, has been established as a moderate risk allele for breast cancer. Some families with 1100delC mutations have CRC in addition to breast cancer. Several studies have tested this variant for risk in LS families and some found modest increases in risk in LS carriers who also have the 1100delC variant (Wasielewski et al., 2008) while others have found no increase in risk (Sanchez et al., 2005).

Genes with variants conferring suggestive effects in some studies for age of diagnosis in HNPCC carriers include CCND1, CDKN2A, AURKA, TP53, E2F2, and IGF1 (Talseth et al. 2008; Jones et al., 2004; Chen et al. 2009).

### 5. The use of genetic information for CRC treatment and prognosis

Currently, the use of germline genetic variants is not used routinely for making clinical decisions regarding colorectal cancer therapy. The bulk of research and clinical application has been with somatic tumour mutations. Despite this, germline mutations and variants have been associated with different tumor histopathology and survival outcomes.

#### 5.1 Low-penetrance risk alleles and tumor histopathology

In addition to playing a role in overall CRC risk, studies indicate that SNPs may impact morphology and type of colon cancer. Some of the low-penetrance SNPs identified through GWAS (rs3802842, rs4939827) have higher risks for rectal cancer than colon cancer (Tenesa et al., 2008). Preliminary studies also suggest differences in risk of necrosis (rs719725), mucin production (rs96153), desmoplastic reaction (rs10441210), Crohn-like lymphocytic reaction (rs6983267, rs4444235) and moderate/well-differentiated histology (rs10795668)(Ghazi et al., 2010). The SNP rs4779584 is associated with reduced risk of death in a Chinese cohort (Xing et al. 2011), but thus far, none of the GWAS-identified variants assessed show an effect on overall survival in Caucasian populations (Tenesa et al., 2010). These studies suggest that variants may impact the development of certain subtypes of colon cancer which provide possible mechanistic insights into colon tumorigenesis and new therapeutic targets.

#### 5.2 Variants to predict response to and off-target effects of cancer therapy

Targeted therapies for colorectal cancer have been developed based on somatic mutations occurring during tumorigenesis. In addition to targeted therapies to somatic mutations, germline variations in enzymes which process more standard chemotherapeutic agents impact prognosis and treatment response. Standard therapies for CRC include 5-fluorouracil and oxaliplatin (FOLFOX). One study looked at the role of germline variants of DNA repair pathways on metastatic CRC patients’ response to FOLFOX (Lamas et al., 2011). One variant in XPD, Lys751Gln, was associated with longer survival, but the numbers in this study were
small. Variants in TS and GSTT1 have been found to be associated with response to both LV5FU2 and FOLFOX (Boige et al. 2010), and two variants in MTHFR (677C>T and 1298 C>A) were associated with better response to FOLFOX (Etienne-Grimaldi et al. 2010). A variant in ERCC2, K751Q was associated with FOLFOX-induced hematologic toxicity (Boige et al. 2010), suggesting that germline variations may also predict CRC treatment side-effects. Since the field of colon cancer pharmacogenomics is still new, it is likely that other variants important in metabolism of CRC therapeutics will be identified.

5.3 Prognosis in tumors with mismatch instability mutations

Colorectal cancers can be divided into different subtypes depending on histology and the presence or absence of specific molecular markers. Treatment and prognosis of these subtypes varies. About 15% of all CRC tumours show evidence of microsatellite instability of which a small proportion are germline mutations (Murphy et al., 2006; Salovaara et al., 2000). A meta-analysis of 32 studies correlating MSI status with clinical outcomes included patients with both germline LS mutations and sporadic MMR defects. Individuals with no evidence of MSI or with MSI-low tumours showed decreased survival (HR=0.67, 95%CI 0.53-0.83) compared to individuals with MSI-high tumours. Polymorphisms in MMR genes are also associated with specific phenotypes. In one study, individuals who carried one or more G alleles of the MLH1 655A>G SNP had a better outcome and less risk of vascular invasion, distant metastasis or recurrence (Nejda et al., 2009). Some studies have documented better survival in individuals with MUTYH-associated colorectal cancer compared to matched controls with colon cancer (Nielsen et al., 2010). Together these data suggest that tumours that have deficits in DNA repair capabilities through germline or somatic mutations show better survival than tumours competent in DNA repair.

6. Conclusion

In summary, just over a third of colorectal cancer risk is thought to be due to inherited genetic factors. A number of mutations in genes have been found to increase CRC in individuals with hereditary cancer syndromes. These syndromes confer a vastly increased risk of CRC over the general population and an earlier age of diagnosis. Individuals with a family history of CRC should be referred to genetics for evaluation of a genetic syndrome and for guidelines for risk-reducing strategies. In addition to highly-penetrant mutations, many variants of small effect sizes have been identified through family-based and genome-wide association based studies. Whereas many of the variants identified through GWAS have been replicated in many studies, the effect size is small, only a small part of the total genetic risk has been identified, and the clinical utility is not established. The use of germline genetic information may be of clinical utility in the prevention and treatment of CRC; yet there is much that remains to be discovered. Technological advances are yielding new insights into genetic susceptibility to CRC on the population level, but we have yet to find the aetiology of most of the genetic risk contributing to CRC.

7. Acknowledgments

This work was funded by the National Institutes of Health/National Cancer Institute (R01 CA-134461-01) and the Ohio State University Comprehensive Cancer Center. We thank Heather Hampel for thoughtful review of this manuscript.
8. References


Germline Genetics in Colorectal Cancer Susceptibility and Prognosis


relative frequency of cancers in different familial phenotypes. *Journal of Medical Genetics, Vol. 47, No.6, pp.421-428, ISSN 1468-6244.*


Colorectal cancer is a common disease, affecting millions worldwide and represents a global health problem. Effective therapeutic solutions and control measures for the disease will come from the collective research efforts of clinicians and scientists worldwide. This book presents the current status of the strides being made to understand the fundamental scientific basis of colorectal cancer. It provides contributions from scientists, clinicians and investigators from 20 different countries. The four sections of this volume examine the evidence and data in relation to genes and various polymorphisms, tumor microenvironment and infections associated with colorectal cancer. An increasingly better appreciation of the complex inter-connected basic biology of colorectal cancer will translate into effective measures for management and treatment of the disease. Research scientists and investigators as well as clinicians searching for a good understanding of the disease will find this book useful.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
