Chapter from the book *Carcinogenesis*
Downloaded from: http://www.intechopen.com/books/carcinogenesis
Chapter 1

Synergistic Effects of Low-Risk Variant Alleles in Cancer Predisposition

Francesca Duraturo, Raffaella Liccardo, Angela Cavallo, Marina De Rosa and Paola Izzo

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/55417

1. Introduction

It has long been known that cancer can be the result of a genetic predisposition. About 5% of total cancers are associated with known Mendelian susceptibility; in these cancer types the clinical manifestations of disease are due to mutations in high-risk alleles, with a penetrance usually at least of 70%. However, there are many tumors in which the cause of hereditary predisposition can not be explained as the Mendelian syndromes. For colorectal cancer (CRC), for example, about 30% of cases are thought to be due to inherited susceptibility, which only in part can be explained by the known Mendelian inheritance, as FAP, MAP and Lynch syndrome [1]. Breast cancer has a similar gap between Mendelian and overall genetic risk. For prostate cancer, the risk is even higher, as very few cases are attributable to high-risk alleles. This gap needs to be filled by studies to identify predisposition alleles that explain the cases of hereditary tumors for which no association with gene variants has been found, so far [2].

With the advent of high-throughput technology it is now possible to analyze a great number of polymorphic variants in large cohorts of cases and controls. These studies have been used successfully by many groups leading to the identification of a large number of rare variant alleles in patients with an inherited risk of cancer [3, 4]. The simultaneous presence of rare genetic variants in the same patient might contribute in a cooperative manner to increase the risk of tumor development. Another problem is represented by variants of unknown significance (VUSs) within the cancer predisposition highly penetrant genes. These variants are usually missense or silent changes which are generally rather uncommon or rare and thus of doubtful clinical relevance, that make troublesome the genetic counseling for these cancer families. The interpretation of these variations is not easy and requires the combination of different analytical strategies to get a proper assessment of their
pathogenicity [5]. In some cases, VUSs make a more substantial overall contribution to cancer risk than the well-assessed severe Mendelian variants. It is also possible that the simultaneous presence of some polymorphisms and VUSs in cancer predisposition genes that behave as low-risk alleles, might contribute in a cooperative manner to increase the risk of hereditary cancer [6]. Therefore, current literature data suggest that a significant proportion of the inherited susceptibility to relatively common human diseases may be due to the addition of the effects of a series of low frequency variants of different genes, probably acting in a dominant and independent manner, with each of them conferring a moderate but even detectable increase in the relative cancer-risk.

Our studies are concerned with the molecular basis of the Lynch syndrome, which is commonly associated with mutations in mismatch repair (MMR) genes, MLH1 and MSH2. However, mutations in these genes do not account for all Lynch syndrome families. In our experience we have also identified germ-line genetic variants in the other MMR genes, called minor MMR genes: MSH6, PMS2, MLH3 and MSH3. We have shown that several patients were carriers of at least two genomic variants within the “minor” genes or a VUS in a major gene associated to a genetic variant in minor genes. We therefore speculate that the association between weak alleles in the MMR genes could determine the onset of the tumor.

2. Hereditary cancer syndromes

Over 200 hereditary cancer susceptibility syndromes have been described, the majority of which are inherited in an autosomal dominant manner. Although many of these are rare syndromes, they are thought to account for at least 5–10% of all cancer, amounting to a substantial burden of morbidity and mortality in the human population (Figure 1).

![Figure 1](image.png)

**Figure 1.** The majority of most common cancers are sporadic, 5–10% are inherited and arise due to highly penetrant germ-line mutations. An additional 10–15% are referred to as ‘familial’ and may be caused by the interaction of low-penetrance genes, gene–environment interactions, or both.

While characterized by their markedly increased risk of malignancy, these syndromes often predispose to benign tumors and generalized disease, as in Cowden syndrome (CS) and the
multiple endocrine neoplasias [7]. When the benign and malignant manifestations are considered together, many of these syndromes show almost complete penetrance by age 70. An inherited cancer susceptibility is suspected in families with the following characteristics: two or more relatives with the same type of cancer on the same side of the family; several generations affected; earlier ages of cancer diagnosis than what is typically seen for that cancer type; individuals with multiple primary cancers; the occurrence of cancers in one family, which are known to be genetically related (such as breast and ovarian cancer, or colon and uterine cancer); and the occurrence of nonmalignant conditions and cancer in the same person and/or family [8]. In table 1 are reported the more frequent hereditary cancer

<table>
<thead>
<tr>
<th>Syndrome</th>
<th>MIM#*</th>
<th>Gene(s)</th>
<th>Population incidence</th>
<th>Penetrance^b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cowden syndrome</td>
<td>158350</td>
<td>PTEN</td>
<td>1/200 000</td>
<td>90–95%</td>
</tr>
<tr>
<td>Familial adenomatous polyposis (FAP or MAP)</td>
<td>175100</td>
<td>APC, MYH</td>
<td>1/8000</td>
<td>&lt;100%</td>
</tr>
<tr>
<td>Hereditary breast–ovarian cancer syndrome</td>
<td>113705, 600185</td>
<td>BRCA1 and BRCA2</td>
<td>1/500 to 1/1000</td>
<td>Up to 85%</td>
</tr>
<tr>
<td>Hereditary diffuse gastric cancer</td>
<td>137215</td>
<td>CDH1</td>
<td>Unknown, rare</td>
<td>90%</td>
</tr>
<tr>
<td>Hereditary nonpolyposis colon cancer</td>
<td>114500</td>
<td>MLH1, MSH2, MSH6, MLH3, PMS2</td>
<td>1 in 400</td>
<td>90%</td>
</tr>
<tr>
<td>Juvenile polyposis syndrome</td>
<td>174900</td>
<td>MADH4 (SMAD4), BMPR1A</td>
<td>1/100 000</td>
<td>90–100%</td>
</tr>
<tr>
<td>Li–Fraumeni syndrome</td>
<td>151623</td>
<td>TP53</td>
<td>Rare</td>
<td>90–95%</td>
</tr>
<tr>
<td>Multiple endocrine neoplasia type 1</td>
<td>131100</td>
<td>MEN1</td>
<td>1/100 000</td>
<td>95%</td>
</tr>
<tr>
<td>Multiple endocrine neoplasia type 2</td>
<td>171400, 162300</td>
<td>RET</td>
<td>1/30 000</td>
<td>70–100%c</td>
</tr>
<tr>
<td>Peutz–Jeghers syndrome (PJS)</td>
<td>175200</td>
<td>LKB1 (STK11)</td>
<td>1/200 000</td>
<td>95–100%</td>
</tr>
<tr>
<td>Retinoblastoma, hereditary (RB)</td>
<td>180200</td>
<td>RB</td>
<td>1/13 500 to 1/25 000</td>
<td>90%</td>
</tr>
<tr>
<td>von Hippel–Lindau (VHL)</td>
<td>193300</td>
<td>VHL</td>
<td>1/36 000</td>
<td>90–95%</td>
</tr>
</tbody>
</table>

* MIM numbers beginning with 1 indicate autosomal dominant inheritance; those beginning with 6 are autosomal loci or phenotypes entered into the catalogue after May 1994. ^ Penetrance estimates are up until age 70 years, include both malignant and benign features and with the exception of MEN2, describe clinical penetrance. § By biochemical testing (pentagastrin-stimulated calcitonin levels) is 95–100% by age 70.

Table 1. Highly penetrant cancer syndromes
syndromes that are associated with mutations in high penetrance alleles. Because of phenotypic variability, age-related penetrance, and gender-specific cancer risks, however, many families with an inherited cancer syndrome will not meet these criteria. Furthermore, because cancer is relatively common in the general population, it is possible to have a chance clusterings of the same or related cancers within a family. These familial clusterings are most likely due to low-penetration alleles that are more common than mutations in high penetrant alleles. Thus, they will potentially account for a larger proportion of cancer in the general population than the mendelian classic syndromes. For colorectal cancer (CRC), for example, Mendelian syndrome includes FAP, MAP and Lynch syndrome.

However, about 30% of the variation in CRC risk is thought to be due to inherited susceptibility, which only in part can be explained by the known Mendelian inheritance [2]. Breast cancer has a similar gap between Mendelian and overall genetic risk and for prostate cancer the risk is even higher, as very few cases are attributable to high-risk alleles. It is that gap which must be filled by studies to identify cancer predisposition alleles in the general population [9]. Localization and characterization of low-penetrance alleles are the focus of much research, but the challenges are great due to the multi-factorial nature of cancer and the underlying genetic heterogeneity.

2.1. High-throughput technology for detection of the multiple alleles associated to cancer predisposition

The history of human genetics has focused on mapping regions of the genome that can explain part or all of a disease or human trait.

The ‘rare variant hypothesis’ proposes that a significant proportion of the inherited susceptibility to relatively common human diseases may be due to the sum of the effects of a series of low frequency variants of a variety of different genes, perhaps dominantly and/or independently acting, each conferring a moderate but detectable increase in relative risk [2]. Regardless, there is good supporting evidence that rare variants will often have stronger effects on cancer risk than common variants. This evidence is based on several works whose purpose was to determine whether evaluating rare single-nucleotide polymorphism (SNPs) in case-control association studies could help to identify causal SNPs for common diseases. The sources of data of these works were generally the International HapMap Project and the SeattleSNPs project and they suggest that slightly deleterious SNPs subjected to weak purifying selection are major players in genetic control of susceptibility to common diseases, including cancer. These results suggests that studies with large sample sizes (5000 and higher) targeting SNPs will be a better strategy to identify causal disease SNPs [10]. Instead, genome wide association studies (GWAS) have emerged as an important tool for discovering regions of the genome that harbor uncommon genetic variants that confer risk for complex tumors, whose nature is probably polygenic [11]. These variants include single nucleotide variants (SNVs) or single nucleotide polymorphisms (SNPs), small insertions and deletions and structural genomic variants.
One of the fundamental elements for the success of GWAS is represented by a large collection of biospecimens in case-control and cohort studies so as to have a high degree of reliability of results. The first approaches in this regard were based on technologies such as the Denaturing High Performance Liquid Chromatography (DHPLC) and classical sequencing analysis, that provide a high degree of analytical sensitivity and specificity. However, the new challenge in the field of biotechnology has surely been to make the techniques increasingly automated in order to process multiple samples simultaneously and especially more quickly.

The method becoming more widely used is high-throughput sequencing, which allows a massive study of DNA. This is a system able to obtain more than 400,000 different readings in a single stroke of about 8 hours. The operating principle is based on clonal amplification of DNA \textit{in vitro} by emulsion PCR and on a protocol of pyrosequencing that, unlike the classic method of Sanger, is based on the detection of pyrophosphate released by the incorporation of a nucleotide during DNA synthesis. In high-throughput sequencing 454 instrumentation, the sample may be any DNA larger than 1500 base pairs (genomic DNA or portions, cDNAs and large amplicons). The sequences obtained are analysed, properly aligned and oriented in contigs from the sequencer software, according to the shotgun and paired-end strategy. The accuracy of the data obtained is measured in terms of "coverage", that is based on the average number of times that each is accessed (read). This technology, therefore, is able to ensure high accuracy of the results (> 99.5%), thanks also to the careful management of the enormous amount of bioinformatics sequences obtained, which minimizes the production of raw redundant data. This feature, coupled with the extraordinary speed of processing, which makes the method also more economical than the classic automated sequencer, allows the user to analyze and quantify at the same time a large amount of samples. Therefore, the sequencer ultra-massive is an extremely versatile technique for a large number of applications such as resequencing and de novo assembly of entire genomes, and the massive sequencing of amplicons.

This latter approach is now widely applied, for example, for the identification of rare variants that presumably contribute in a synergistic way and in association with other factors predisposing to the development of complex genetic diseases characterized by genetic heterogeneity. This technology therefore offers a great contribution to the studies of Genome Wide Association, because it allows quick identification of the allele frequencies of SNPs in population studies, and to analyze a given target gene in multiple genomes, or a panel of target genes in a single patient, even at the level of gene expression (transcriptome analysis) [4]. However, the high number of next generation sequencing information requires accurate statistically studies. The threshold value for discovery has been established at a high level, known as genome-wide significance, which serves two dual purposes [12]. First, it needs careful consideration of the power to detect the effect sizes expected to be observed in the study. Second, the high bar of genome wide significance protects against the probability of a false-positive finding. The latter is critical because GWAS are discovery
tools that point investigators toward long arduous follow-up studies for unraveling the underlying biology and the pursuit of markers for risk assessment [11, 13]. However, the common cancer alleles detected by GWAS account for only 10% of the familial relative risk of disease.

2.2. Variants of unknown significance in hereditary cancer predisposition genes

Variants of unknown significance (VUS) within the cancer predisposition genes could be responsible for cancer development, in particular when associated with another VUS or SNPs. The influence of these variants on the development of cancer is often difficult to predict [5, 14]. Several criteria have been established for the characterization of these phenotypic variants, particularly for the missense variants [15, 16]; these criteria included the co-segregation of the variant with the disease and the presence/absence of variation in the healthy population. However, these criteria are not always pursued to establish the pathogenetic significance of these variants [17, 18].

Segregation analysis is not always practicable, since, often the families are small or part of family members is reluctant to participate to molecular investigation. Population studies to exclude the polymorphic nature of the variant is often laborious. Recent studies have revealed new strategies to classify the VUS as pathogenic. These strategies include “in silico” analysis, using computational programs such as PolyPhen (Polymorphism Phenotyping) and SIFT (Sorting Intolerant From Tolerant) to assess whether the VUS missense type falls into a phylogenetically conserved domain and/or makes changes to the physical-chemical properties of proteins [19-21].

The program Human Splicing Finder (HSF) [22, 23], which simultaneously uses a set of matrices already available on the network is useful to predict the effects of missense, silent and intronic variants on the signals of splicing and to identify regulators motifs associated with the processing of the mRNA. However, the results of the computational accuracy have a predictive value of about 80% and, therefore, do not always reflect the functional consequences of the variant in vivo. Several papers suggest to combine the results from several bioinformatics approaches especially those based on amino acid conservation status, to increase the predictive value of about 10% [19, 24].

Other studies complemented “in silico” analysis to a direct study of the mRNA, to confirm or rule out the effects of splicing variants [25, 26]. In addition, many recent literature data emphasize the importance of developing functional assays in vitro and in vivo to assess the effects of VUS on specific biological functions [18]. All studies conducted so far show that none of the above criteria, including functional assays, is an indicator of pathogenicity, if considered individually; it is necessary that most of these strategies are used in combination with each other so that they can lead to a correct evaluation pathogenicity of numerous variant data.
2.3. The simultaneous presence of low-risk alleles increases the risk of hereditary cancer: review of literature data

Genome-wide association studies in cancer based on high-throughput sequencing approaches have already identified over 150 regions associated with two dozen specific cancers, such as breast, prostate and colorectal cancer, providing new insights into common mechanisms of carcinogenesis. Since each region confers a small contribution to the cancer risk, it is daunting to consider any single nucleotide polymorphism as a clinical test, rather one should think about the synergistic action of different SNP as well as the environmental factor [11, 27]. These studies allowed researchers to identify large susceptibility chromosomal regions for many unrelated cancers. For example, the 8q24 region harbor multiple cancer susceptibility SNP loci associated with prostate cancer, colorectal cancer and precancerous colorectal adenomas, and bladder cancer risk; these loci affect genes such as MYC oncogene and the prostate stem cell antigen gene (PSCA) [11, 28].

Another common cancer susceptibility chromosomal region is the 5p15.33; in this region common variants in the TERT-CLPTM1L have been identified by GWAS in association with the prostate, uterine cervix and skin cancers [11]. TERT is an attractive candidate gene, because it encodes the reverse transcriptase component of the telomerase, a gene that is critical for telomere replication and stabilization by controlling telomere length. TERT promotes epithelial proliferation and telomere maintenance has been implicated in the progression from KRAS-activated adenoma to adenocarcinoma in a murine model. There is additional evidence for its association with bladder, prostate, uterine cervix and skin cancers [11]. Moreover, phenotypic heterogeneity in the breast cancer, such as merging estrogen receptor negative and positive cases, has been need to identify other loci that might contribute to different phenotypes. Preliminary GW analysis has shown that a subset of the discovered loci may be specific to ER-pos breast cancer while select loci could be more important for ER-neg breast cancer [29]. Similar studies have identified an association between coding variants in CASP8 gene and breast cancer [30]. CASP8 belongs to many key pathways, including p53 signaling, apoptosis, and cancer [31]. The decreased risk for breast cancer with CASP8 Asp302His was revealed in an another recent association study [32]. Others proposed that rare variants within the double strand break repair genes CHEK2, BRIP1 and PALB2 predispose to breast cancer [33].

Other large studies have identified 31.7% of the novel gene-variant breast cancer significant associations between 145 variants analyzed. A large GWAS conducted with East Asian women provided convincing evidence for an association with a novel independent susceptibility locus located at 6q25.1, near the TAB2 gene (TGF-beta activated kinase 1). Furthermore this study shows that genetic variants in the ESR1 gene (estrogen receptor 1) may be related to breast cancer risk [34]. A recent study of populations conducted by Smith et al. [35] has pointed out that the simultaneous presence of mutations in the TP53 gene and single nucleotide polymorphisms (SNPs) in genes belonging to different repair systems such as complex BER, NER, MMR and DSBR (Double-Strand Break Repair) is associated with earlier age of onset of breast cancer (<50 years), thus suggesting the idea of an additive or multiplicative effect.
In prostate cancer, there are at least 35 distinct loci harboring common susceptibility alleles identified by GWAS that could distinguish between aggressive and non-aggressive disease, but other studies are required [36]. These analyses were conducted in both European and Asian populations [37]. Moreover, a fine mapping of a region of chromosome 11q13 showed a complex genomic architecture characterized by multiple independent signals contributing to prostate cancer risk. This study further annotates common and uncommon variants across this region. In particular, a variant in the promoter of the MSMB gene on chromosome 10q13, is known to have influence in the gene expression, and in the protein PSP94 (prostate secretory protein 94) levels, showing significant association with prostate cancer. This chromosomal region was extensively resequenced and it is possible that a neighboring gene, the androgen receptor coactivator (NCOA4), could also be a candidate gene for analysis [38]. Moreover, GWAS for chromosomal 19q13.33 region, that harbors the gene responsible for the prostate serum antigen (PSA), suggested that variants in this gene, including a nonsynonymous SNP, could contribute to both prostate carcinogenesis and PSA levels [39].

A large GWAS conducted in several populations (European Americans and African Americans) showed that genetic associations by race are modified by interactions between individual SNPs and prostate cancer and that significance of particular GWAS “hits” is not the same between racial groups. This study highlights the need to conduct GWAS and GWAS replication studies in a variety of racial groups in order to gain a more complete understanding of differences in risk alleles by race and in order to study gene-gene and gene-environment interactions [40]. A similar study conducted in two European populations suggested a list of SNP–SNP interactions that can be followed in other confirmation studies to explore the etiology of prostate cancer [41].

Finally several papers report numerous GWAS for colorectal cancer, identifying a total of 16 new susceptibility loci for colorectal cancer. SNPs both in common genes as MMR genes and in other novel loci as SMAD7 and MYC seem to associate with different clinical outcomes [42], or different pharmacological responses [43]. Moreover, GWAS for chromosomal 20p12.3 region, a site bereft of genes or predicted protein-encoding transcripts, suggested that particular SNP in this region could contribute to colorectal cancer progression. Interestingly, the bone morphogenetic protein 2 (BMP2) maps 342 kb telomeric to this locus, which is an initiator of BMP signaling by binding to its corresponding receptors. BMP signaling can suppress the Wnt pathway to ensure a balanced control of intestinal stem cell self-renewal. As reflected by earlier studies, mutations of BMP pathway have been described in juvenile polyposis, an inherited syndrome that predisposes to CRC. Considering all this information, it has been speculated that this locus might alter the BMP signaling transduction by the effect on BMP2 and thus affect CRC incidence [44].

A different GWAS study assessed a set of single-nucleotide polymorphisms (SNPs) near 157 DNA repair genes in three studies on colorectal cancer (CRC). Although no individual SNP showed evidence of association, the set of SNPs as a whole was associated with colorectal cancer risk, in particular the MLH1 promoter SNP -93G>A (rs1800734) and rare variants in
CHEK2 (I157T and possibly del1100C) [45]. Numerous GWAS data for susceptibility cancer specially for colorectal cancer have been the subject of several functional studies to demonstrate the effective association and to test the hypothesis of a synergistic effect between low risk allelic variants.

In a recent study on the genome of yeast, it has been shown that the weak alleles of MMR complex cause a weak mutator phenotype, but when these interact with each other cause a strong mutator phenotype. In this work, 11 SNPs and 14 missense variants of doubt pathogenetic meaning, previously identified in these genes, have been studied. The mutator effect of these variants both individually and in combination with each other was assayed by testing complementation, in selective media for the amino acids lysine and tyrosine, and for resistance to canavanine [46]. Finally, Demogines et al. [47] have used yeast strains, that differed in terms of geographic and environmental factors, to demonstrate that the association of polymorphic variants, identified in the MMR genes MLH1 and PMS1, affecting the same or different genetic loci, may act as modifiers intra- or inter-gene and this phenomenon may play a role in both the penetrance of the colorectal disease (mutator phenotype) and in the process of evolutionary adaptation (genomic compatibility).

3. The Lynch syndrome

In this chapter we report the results of our studies on detection of mutations in MisMatch Repair (MMR) genes as responsible for Lynch syndrome. Because many patients with hereditary cancer syndrome did not show mutations in high penetrance genes, we speculate that association of several low penetrance alleles could determine a genetic predisposition to cancer development.

Colon cancer is a multifactorial disease. It’s caused by enviromental factors, nutritional factors and genetic predisposition. Our studies are related to the genetic susceptibility of colon cancer, in particular the molecular basis of Lynch syndrome (Hereditary Non Polyposis colorectal cancer, HNPCC). The Lynch Syndrome is one of the syndromes of hereditary cancer with higher incidence in the population [48]. It has an autosomal dominant transmission and occurs in two forms: as Lynch I with an early age of occurrence (25% at 50 years and 70–80% within 70 years), predilection for the proximal colon (60–80%), and high rates of metachronous colorectal cancer (30% at 10 years and 50% at 15 years from the first tumor); and Lynch II, has the same characteristics but also extracolonic tumors involving the uterus (25–60%), ovaries (8–14%), stomach (13%), and urinary tract (4%) (Figure 2).

This syndrome accounts for 5–15% of all colorectal cancers, although the true incidence is unknown, confounded by incomplete penetrance (<80%), rapid progression of adenoma to carcinoma (<5 years), development of extracolonic neoplasms, and the inter- and, occasionally, intra-familiar heterogeneity of the lesions [49]. In Lynch syndrome, the adenomas have the same frequency as in sporadic cases, but a more rapid progression to carcinoma. Due to the deficiency in DNA-repair genes, adenomas accumulate mutations
about three times faster than in sporadic disease. These mutations occur predominantly in microsatellite DNA sequences, a condition defined as microsatellite instability (MSI), which are more susceptible to errors in these genes replication because of their repetitive nature. The microsatellite sequences are also present in very important colorectal cancer tumorigenesis genes, thus the accumulation of errors in these genes determine rapid cellular proliferation. MSI is present in over 90% Lynch cases [50]. The clinical diagnosis of Lynch syndrome is performed upon the Amsterdam Criteria (Tab. 2). However, the Amsterdam Criteria do not identify up to 30% of potential Lynch syndrome carriers [51].

![Lifetime risk (%)](image)

**Figure 2.** Lifetime Risk of development of cancer associated with Lynch Syndrome.
All of the following must apply for a putative diagnosis of HNPCC to be made in a family

<table>
<thead>
<tr>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>There are at least three relatives with an HNPCC-associated cancer (large bowel, endometrium, small bowel, ureter, or renal pelvis, although not including stomach, ovary, brain, bladder, or skin)</td>
</tr>
<tr>
<td>One affected person is a first-degree relative of the other two</td>
</tr>
<tr>
<td>At least two successive generations are affected</td>
</tr>
<tr>
<td>At least one person was diagnosed before the age of 50 years</td>
</tr>
<tr>
<td>Familial adenomatous polyposis has been excluded</td>
</tr>
<tr>
<td>Tumors have been verified by pathologic examination</td>
</tr>
</tbody>
</table>

Table 2. Amsterdam Criteria I and II

For this reason, in some patients with colon cancer, as suggested by the Bethesda guidelines (Table 3) [52], it is possible to analyse microsatellite instability in colon tumor specimens, to identify the inefficiency of DNA mismatch repair complex. If there is microsatellite instability, there is a higher likelihood for a Lynch syndrome diagnosis.

Tumors from any of the following should be tested for MSI and then positive patients should continue for MMR testing

<table>
<thead>
<tr>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individuals with cancer in families that meet the Amsterdam Criteria</td>
</tr>
<tr>
<td>Individuals with two HNPCC-associated cancers, including synchronous and metachronous CRC or associated extracolonic cancers</td>
</tr>
<tr>
<td>Individuals with CRC and a first-degree relative with CRC and/or HNPCC-related extracolonic cancer and/or a colorectal adenoma diagnosed at age &lt; 40 years</td>
</tr>
<tr>
<td>Individuals with CRC or endometrial cancer diagnosed at age &lt; 45 years</td>
</tr>
<tr>
<td>Individuals with right-sided CRC with an undifferentiated pattern (solid or cribriform) on histopathology diagnosed at age &lt; 45 years</td>
</tr>
<tr>
<td>Individuals with signet-ring–cell-type CRC diagnosed at age &lt; 45 years</td>
</tr>
<tr>
<td>Individuals with adenomas diagnosed at age &lt; 40 years</td>
</tr>
</tbody>
</table>

Table 3. Bethesda Guidelines for MSI Testing

Germ-line mutations in the MLH1 and MSH2 genes account for a majority of families with Lynch Syndrome. The majority of research into mutations has focused on MLH1 and MSH2, however mutations in these two gene are not present in many patients. So far, 10% of mutations in MMR genes have been identified in the MSH6 gene and a total of 5% in MLH3 and PMS2 and very recently germ-line mutations in the MSH3 gene [53]. These genes are defined as “minor MMR genes” because they have redundant functions in mismatch repair in replication. It is known that as well as being involved in mismatch repair in replication,
the MMR system also has other functions [54], such as: DNA damage response, diversification of antibody, promotion of meiotic crossover. In these functions the “minor” MMR genes play an important role.

3.1. Results of mutation detection analysis in MMR genes

Recently, several studies have shown that association of low penetrance alleles could determine a genetic predisposition to cancer development [46,47]. For this reason, we studied 63 Lynch families recruited from various health centres in Campania (Southern Italy). Of these, forty families met the Amsterdam criteria and twenty-three patients with high microsatellite instability (MSI-H) met the Bethesda guidelines, in which no pathogenetic germline mutations were identified in MLH1 and MSH2 genes. We performed detection mutation analysis in each minor MMR gene (MSH6, MLH3, PMS2 and MSH3) by DHPLC. All samples exhibiting abnormal DHPLC profiles were analyzed by directed sequencing (Figure 3). In our studies we have identified overall 65 genetic variants in these “minor” MMR genes.

![Figure 3. A) Chromatogram and B) electropherogram of the missense mutation c.2732 T>G (Leu>Trp) in MSH3 gene.](image)

The analysis of the damaged point mutations at the structural level is considered to be very important to understand the functional activity of the protein concerned. For this purpose we used the server PolyPhen (bibl), which is available at http://coot.embl.de/PolyPhen/, for missense mutations identified in this study. Moreover, we also used the bioinformatic analysis for the silent and intronic variants.

These variants were analyzed by the software “Human Splicing Finder”, a tool to predict the effects of mutations on splicing signals or to identify splicing motifs in any human sequence. Most of these variants result in a polymorphism, which, however, can cause phenotypic variability, affecting the accuracy and efficiency of the protein function [24]. Interestingly, several patients were carriers of at least two genomic variants within the “minor” genes or a VUS in a major gene associated with a genetic variant in minor genes (Table 4).

Recently, the effect of polymorphisms and missense mutations in human MMR genes was studied in a *Saccharomyces cerevisiae*-based system. A number of weak alleles of MMR genes and MMR gene polymorphisms that are capable of interacting with other weak alleles of MMR genes to produce strong polygenic MMR defects, have been identified [46]. A similar situation found in our studies might support the hypothesis that weak MMR gene alleles are
capable of polygenic interactions with other MMR gene alleles that might lead to tumour progression in Lynch syndrome.

<table>
<thead>
<tr>
<th>PATIENTS</th>
<th>MSH6</th>
<th>PMS2</th>
<th>MLH3</th>
<th>MSH3</th>
<th>PHENOTYPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>9525</td>
<td>ex4  c.2633 T&gt;C (Val&gt;Ala)</td>
<td>ex14 c.2324 A&gt;G (Asn&gt;Ser)</td>
<td>ex1 c.2530 C&gt;T (Pro&gt;Ser) c.2533 T&gt;C (Ser&gt;Pro)</td>
<td>IVS7-9 T&gt;C</td>
<td>AM+</td>
</tr>
<tr>
<td>013</td>
<td>ex6  c.665G&gt;C (Ser&gt;Thr)</td>
<td>IVS6 +16A&gt;G</td>
<td>ex1 c.2533 T&gt;C (Ser&gt;Pro)</td>
<td>NO AM MSI-H</td>
<td></td>
</tr>
<tr>
<td>103</td>
<td>ex5  c.3261_62insC (Phe&gt;stop)</td>
<td></td>
<td>ex1 c.2533 T&gt;C (Ser&gt;Pro)</td>
<td>ex12 c.1860G&gt;A (Asp&gt;Asn)</td>
<td>NO AM later onset MSI-H</td>
</tr>
<tr>
<td>423</td>
<td></td>
<td>IVS12-4G&gt;A</td>
<td>ex1 c.2530 C&gt;T (Pro&gt;Ser) c.2533 T&gt;C (Ser&gt;Pro)</td>
<td>AM+ later onset MSI-L</td>
<td></td>
</tr>
<tr>
<td>015</td>
<td>ex5  c.3295_97delTT (Ile&gt;stop)</td>
<td></td>
<td>ex1 c.666 G&gt;A (Lys) c.2191 G&gt;T (Val&gt;Phe) c.2533A&gt;G (Ser&gt;Gly)</td>
<td></td>
<td>AM+ MSI-H</td>
</tr>
<tr>
<td>210</td>
<td>ex4  c.2941 A&gt;G (Ile&gt;Val)</td>
<td>IVS6+16A&gt;G</td>
<td>ex13 c.2324 T&gt;C (Phe)</td>
<td>IVS6-64 C&gt;T</td>
<td>AM+</td>
</tr>
<tr>
<td>211</td>
<td>ex4  c.2941 A&gt;G (Ile&gt;Val)</td>
<td>IVS12-4 G&gt;A</td>
<td></td>
<td>IVS6-64 C&gt;T</td>
<td>AM+</td>
</tr>
<tr>
<td>416</td>
<td>ex11 c.1714C&gt;A (Thr&gt;Lys)</td>
<td>ex 1 c.2027G&gt;A (Arg&gt;Lys)</td>
<td></td>
<td>IVS6-64 C&gt;T</td>
<td>AM+ MSI-H</td>
</tr>
<tr>
<td>504*</td>
<td></td>
<td></td>
<td></td>
<td>ex4 c.693G&gt;A (Pro) ex20 c.2732 T&gt;G (Leu&gt;Trp)</td>
<td>AM+ MSI-H</td>
</tr>
</tbody>
</table>

Table 4. Patients carrying variants in several MMR genes: MSH6, PMS2, MSH3, MLH3; *the patient shows also the UV in MSH2 gene (c.984 C>T)

In detail, we report the case of a Lynch family with mutations in several MMR genes. The index case of family 504 (II-5 in Figure 4), who had developed an adenocarcinoma of the left colon at the age of 34 years, an adenocarcinoma of the right colon at the age of 53 years and
a new malignancy of the colon at 59 years of age, show two mutations in MSH3 gene, the c.2732 T>G in exon 20 and c.693 G>A in exon 4, and an UV within the MSH2 gene, the c.984 C>T in exon 6. The PolyPhen in silico analysis showed that the missense variant in MSH3 might alter the function of the protein, because it falls into a highly conserved region in different species, while the silent variant, analyzed by HSF could affect the splicing process.

To elucidate whether the mutation was associated with the disease in this family, we analysed another eight members. These variants was found in a brother of the index case, with the same phenotype. Instead, another brother (II-8 in Fig. 4) showed only a variant in the MSH2 gene and no genetic variants in the MSH3 gene. This patient had developed a polyp of the colon at 47 years of age. Today he is 59 years old, undergoes regular colonoscopy and so far has not presented other polyps. In the third generation (Fig. 4), we analysed four affected family members. Subjects III-1 and III-2, in Figure 4, showed a silent variant in MSH3 and a variant in MSH2; both subjects showed an early-onset right colon tumor. Subjects III-3 and III-4, in Figure 4, the sons of our proband, developed colon cancer at 36 years of age and a tubular adenoma of the colon at 34 years of age, respectively. Both subjects showed a silent variant in MSH2 and a missense variant and a silent variant in MSH3. The MSI analysis performed on DNA extracted from tumour tissues of patients II-5 and III-3 showed an MSI-H status. Thus, both subjects presented a strong mutator phenotype, probably due to an additive effect by several variants that leads to inefficiency of the MMR complex. The other family members analysed showed only one mutation in the MSH3 gene and they do not present a typical phenotype of Lynch syndrome (Tab.4). Therefore, it is clear that all subjects in this family with the Lynch phenotype showed the c.984T allele of MSH2 and a germ-line variant in the MSH3 gene (a missense and/or silent variant).

Patients belonging to other families showed mutations in several MMR genes; however, for these families it wasn’t possible to perform segregation analysis of mutations with disease because no other family members were available for the analysis. In conclusion, several germ-line variants have been identified in several MMR genes using a DHPLC procedure; a method robust, automated, highly sensitive, fast, feasible and particularly useful for high-throughput analyses.

On the basis of this study, it is conceivable to hypothesize a model in which these genetic variants behave as low-risk alleles that contribute to the risk of colon cancer in Lynch families, mostly together with other low-risk alleles of other MMR genes. Therefore, if our assumptions are correct, these studies may indicate a novel inheritance model in the Lynch syndrome, and might suggest that the risk alleles identified to date represent just the tip of an iceberg of risk variants likely to include hundreds of modest effects and possibly thousands of very small effects. This could pave the way toward new diagnostic perspectives. Moreover, The same situation could occur in other forms of hereditary cancer and it may explain the large number of cases remained unresolved as well as the phenotypic heterogeneity that characterizes all hereditary cancer syndromes.
Figure 4. Pedegree of 504 family [53]. Symbols and abbreviations used are denoted as fellow: Arrows, analysed members of family; black symbol, colorectal cancer or cancer associate with HNPCC; gray symbols, adenomas or cancer not associated with HNPCC; CRC, colorectal cancer; Br, brain cancer; GU, gastric ulcer; BL, bladder cancer; Bre, breast cancer; TA, tubular adenoma. Number next to diagnosis denote age at oneset; I not detected.

<table>
<thead>
<tr>
<th>PATIENT</th>
<th>c.984C&gt;T EX6 MSH2</th>
<th>c.693G&gt;A EX4 MSH3</th>
<th>c.2732T&gt;G EX20 MSH3</th>
</tr>
</thead>
<tbody>
<tr>
<td>I-3</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>II-3</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>II-4</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>II-5</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>II-6</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>II-8</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>II-9</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>III-1</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>III-2</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>III-3</td>
<td>YES</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>III-4</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
</tr>
</tbody>
</table>

Table 5. Genotypes of analysed patients; the patients are identified with number of pedigree (Fig.4).
4. Further research

The Lynch syndrome is associated mainly with germ-line mutations in MSH2 and MLH1 genes. However, mutational analysis of these two genes do not always provide informative results for genetic counseling of patients with a clinical diagnosis strongly predisposing to cancer development. Therefore, these subjects are considered candidates with simultaneous molecular analysis of all MMR genes. For this reason, high-throughput sequencing could be considered as an analytical approach that adapts better to clarify the molecular basis for each subject with a significant colorectal cancer history. In the future, these new technologies will enable faster identification of the molecular basis of cancer; it will improve the genotype-phenotype correlations the purpose of implementing a clinical treatment more personalized.

5. Conclusions

A field of biology where the “high-throughput technologies” is now widely applied is certainly the genetics of cancer for identification of constitutive and somatic mutations of putative genes associated with hereditary predisposition to cancer, particularly for those diseases characterized by genetic heterogeneity. Nowadays, we are witnessing a revolution in oncologic medicine, and the hope is that an increasing understanding of genetics will one day unlock the potential of personalized medicine. Clinical cancer genetics has traditionally been associated with risk estimation. Genome-wide germ-line mutation analysis will result in the identification of new cancer-associated alleles across the spectrum of risk. This may in time permit more precise estimation of development cancer risk. The new genetics will bridge the gap between germ-line and somatic genetics; prior analysis of the genetic makeup of the person and their tumour at time of diagnosis will be needed in order to tailor therapy. Central to this endeavour will be the increasing use of next-generation sequencers as whole cancer genomes become unravelled, revealing critical pathways that drive tumour progression and resistance. In the future these new technologies will enable faster identification of the molecular basis of cancer and thus improve the genotype-phenotype correlations, in order to implement more personalized monitoring and clinical treatment.

Nomenclature


Author details

Francesca Duraturo, Raffaella Liccardo, Angela Cavallo, Marina De Rosa and Paola Izzo
Department of Molecular Medicine and Medical Biotechnologie, University of Naples Federico II, Italy
Acknowledgement

Ministero Salute - Ricerca Oncologica - RECAM-2006-353005; PRIN 2007-prot. 2007EN8F7T-004

6. References


