Identification of novel germline mutations in hereditary colorectal cancer patients and characterization of somatic alterations in their tumors

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der

Universität Basel

von

Jian Zhang

aus Nanjing, Volksrepublik China

Basel, 2008

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Prof. Hansjakob Müller

Prof. Gerhard Christofori

Prof. Markus Affolter

In memeory of my truly loved father, outstanding scientist,

Professor Lingyuan zhang

TABLE OF CONTENTS

Page

TABLE OF CONTENTS ..... **i**

ACKNOWLEDGEMENT ..... **iv**

ABBREVIATIONS ..... **v**

ABSTRACT ..... **vii**

CHAPTERS

1. General Introduction 1

1.1 Cancer and colorectal cancer 1
1.2 Hereditary colorectal cancer 2
1.2.1 Predisposition to colorectal cancer without pre-existing polyposis -Hereditary Nonpolyposis Colorectal Cancer (HNPCC) 3
1.2.2 Predispositions to colorectal cancer with pre-existing polyposis - Familial Adenomatous Polyposis (FAP) and MYH Associated Polyposis (MAP) 7

1.3 Tumorigenesis in colorectal cancer 9
1.4 Aims of the thesis 15

2. Methods 16

2.1 DNA extraction 16
2.2 RNA extraction 17
2.3 Microsatellite marker analysis 17
2.4 Immunohistochemistry (IHC) 19
2.5 Denaturing High Performance Liquid Chromatography (dHPLC) 19
2.6 Direct DNA sequencing 21
2.7 Quantitative Multiplex PCR Amplification (QMPA) assay 22
2.8 Multiplex Ligation-dependent Probe Amplification (MLPA) assay 23
2.9 Long Range PCR 25
2.10 RT-PCR 25
2.11 Protein truncation test (PTT) 26

3. Predisposition to colorectal cancer without pre-existing polyposis- Hereditary Nonpolyposis Colorectal Cancer 27

3.1 Evaluation of different screening techniques to detect large genomic rearrangements in *MSH2* and *MLH1* 28
3.1.1 Abstract 28
3.1.2 Introduction 28
3.1.3 Patients and methods 30
3.1.4 Results 32
3.1.5 Discussion 40
3.1.6 References 43

3.2 Gene conversion is a frequent mechanism of inactivation of the wild- type allele in cancers from *MLH1/MSH2* deletion carriers 45
3.2.1 Abstract 45
3.2.2 Introduction 45
3.2.3 Patients and methods 46
3.2.4 Results and discussion 49
3.2.5 References 58

3.3 A *de novo MLH1* germline mutation in a 31 year old colorectal cancer patient 60
3.3.1 Abstract 60
3.3.2 Introduction 60
3.3.3 Results and discussion 61
3.3.4 References 67

4. Predispositions to colorectal cancer with pre-existing polyposis - Familial Adenomatous Polyposis (FAP)) and MYH Associated Polyposis (MAP) . 68

4.1 Detailed genetic analysis of the *APC* locus reveals complex early pathways of tumorigenesis in attenuated familial adenomatous polyposis 69
4.1.1 Abstract 69
4.1.2 Introduction 69
4.1.3 Patients and methods 72
4.1.4 Results 75
4.1.5 Discussion 81
4.1.6 References 90

4.2 Prevalence of *MYH* germline mutations in Swiss *APC* mutation-negative polyposis patients 92
4.2.1 Abstract 92
4.2.2 Introduction 92
4.2.3 Patients and methods 93

4.2.4 Results 95
4.2.5 Discussion 99
4.2.6 References 102

5. General Discussion 104

REFERENCES 112

APPENDIX I 126

APPENDIX II 130

APPENDIX III 131

CURRICULUM VITAE 135

PUBLICATIONS AND CONFERENCES 136

DECLARATION OF INDEPENDENCE 137

ACKNOWLEDGEMENTS

The present work was carried out under the supervision of Prof. Hansjakob Müller and PD Dr. med.et phil. II Karl Heinimann.

I am deeply grateful to all people who supported me and whom I have worked with during this time. I would like to thank for all the help and support from the following list of the people:

Firstly and foremost, I am greatly indebted Professor Hansjakob Müller for his support, scientific vision, and encouragement.
I extend this gratitude to my direct supervisor PD Karl Heinimann for his support, scientific vision, valuable discussions and excellent plan for the research projects.

Special thanks also go for the PhD father Professor Markus Affolter and Co-referee Gerhard Christofori for their advices and for the role they played on my PhD committee.

I would also like to warmly thank our family study specialist Marianne Häusler and excellent technicians Michèle Attenhofer, Sibylle Bertschin and Thomas Woodtli, for their great technical assistance. Special thanks go for Dr. Martina Plasilova, Judith Luz, Dr. Katrin Bühler and Dr. Anna Russell. It has been lovely to work with you and thank you for sharing the experiences with me.
I would also like to acknowledge and heartily thank Dr. Pierre Hutter, Dr. Olive Sieber and Stephanie Segditsas for the support information and data confirmation.

My personal appreciation and regret go to my father who passed away few years ago, for his trust, scientific spirit and encouragement left for me.
Another personal appreciation has to go to my husband Dr. Peter Kozlik and my family, for their truly unconditional love and supports to help me through the tough time and go until the completion of the thesis during my pregnancy.

Through the times I have spent at Matternstrasse 28, Center of Biomedicine, I have made lots of special friends. I thank all of them to share the working experiences and scientific knowledge with me and thank all the support they gave to me.

ABBREVIATIONS

AAR	Amino Acid Repeats
AC	Amsterdam Criteria
AFAP	Attenuated Familial Adenomatous Polyposis
APC	Adenomatous Polyposis Coli
BER	Base Excision Repair
BG	Bethesda Guidelines
CHRPE	Congenital Hypertrophy of the Retinal Pigment Epithelium
CIN	Chromosomal Instability
CRC	Colorectal Cancer
DGGE	Denaturing Gradient Gel Electrophoresis
dHPLC	denaturing High Performance Liquid Chromatography
Dox	Doxicyclin
FAP	Familial Adenomatous Polyposis
8-oxoG	7,8 -dihydro-8-oxo-guanine
HE	Hematoxylen-Eosin stain
HNPCC	Hereditary Nonpolyposis Colorectal Cancer
IGFIIR	Insulin-like growth factor II receptor
IHC	Immunohistochemistry
LINE	Long Interspersed Nuclear Elements
LOH	Loss of Heterozygosity
MAP	MYH associated polyposis
MCR	Mutation Cluster Region

MIN	Multiple Intestinal Neoplasia
MLPA	Multiplex Ligation dependent Probe Amplification
MLH1	Human MutL Homologue 1
MLH3	Human MutL Homologue 3
MMR	Mismatch Repair
MSH2	Human MutS Homologue 2
MSH3	Human MutS Homologue 3
MSH6	Human MutS Homologue 6
MSI	Microsatellite Instability
MSI-H	Microsatellite Instability-High
MSI-L	Microsatellite Instability-Low
MYH	Human *MutY* homologue
PMS1	Human Post Meiotic Segregation 1
PMS2	Human Post Meiotic Segregation 2
PCR	Polymerase Chain Reaction
PTT	Protein Truncation Test
QMPA	Quantitative multiplex PCR Amplification
SD	Standard Deviation
SNP	Single Nucleotide Polymorphism
SSCP	Single Strand Conformation Analysis
TCF-4	T cell factor- 4
TGFβRII	Transforming growth factor- β receptor type II
TSG	Tumor suppressor gene

ABSTRACT

Colorectal cancer has been reported as the third leading cause of cancer related death in the world. About 5-10% of colorectal cancers are due to an inherited predisposition.

This thesis focuses on investigating the prevalence of large genomic rearrangements and other types of germline mutations in novel cancer susceptibility genes in two major hereditary colorectal cancer syndromes, hereditary nonpolyposis colorectal cancer (HNPCC) and familial adenomatous polyposis (FAP). Furthermore, the second somatic mutations were characterized in the tumors from DNA mismatch repair (HNPC) and *APC* gene mutation carriers (FAP) to address the mechanism(s) of tumorigenesis in these two syndromes. All these investigations aim to understand tumor initiation and progression in hereditary colorectal cancer syndromes in order to enable early and reliable presymptomatic diagnosis of a person at increased risk and offer optimal medical management to prevent cancer.

HNPCC is an autosomal dominantly inherited cancer predisposition syndrome caused by germline mutations in DNA mismatch repair (MMR) genes. Prescreening methods are routinely applied to detect MMR gene sequence alterations, but inevitably miss large genomic rearrangements. Here, two novel PCR-based methods to study gene dosage were introduced in 35 MLH/MSH2 HNPCC patients in whom no mutation could be identified by conventional screening methods. These methods are QMPA (quantitative multiplex PCR amplification) and MLPA (multiplex ligation dependent probe amplification). Three patients were found to carry large deletions by QMPA and MLPA. In 1 patient, however, QMPA yielded a false positive result. Both methods, QMPA and MLPA appear to be of comparable sensitivity albeit with different specificity. Since the QMPA technique is difficult to set up and to standardize the PCR conditions, the MLPA assay is better suited to routinely search for large genomic rearrangements.

The investigations subsequently continued to detect the frequency and nature of LOH as second, somatic event in tumors from *MLH/MSH2* germline deletion carriers. MLPA technique was applied to analyze 18 cancer specimens from two independent sets of Swiss and Finnish *MLH1/MSH2* deletion carriers. Results revealed that somatic deletions identical to the ones in the germline occur frequently (55%) in CRCs and that this type of loss of the wild type allele is also present in extracolonic HNPCC associated tumors. Chromosome specific marker analysis implies that loss of the wild type allele predominantly occurs through locus restricted recombination events, i.e. gene conversion, rather than mitotic recombination or deletion of the respective gene

locus. The same investigation was carried on a 31 years old colorectal cancer patient who carries *de novo* mutation (c.666dupA) in the *MLH1* gene. The tumor analysis of this patient showed a similar somatic mutation mechanism to the large genomic deletion carriers.

Prior to our analysis of the somatic hits in the attenuated form of familial adenomatous polyposis (AFAP), earlier investigations had shown that in classical FAP the "two hits" in the *APC* (Adenomatosis polyposis coli) gene are not occurring randomly but are in fact interdependent. AFAP is clinically characterized by fewer than 100 adenomatous polyps in the colorectum and presents with a milder phenotype compared to classical FAP. *APC* mutations in AFAP patients are typically located in the very 5' and 3' gene regions as well as in the alternatively spliced region of exon 9. In a collaborative effort we investigated the somatic alterations in 235 tumors of 35 AFAP patients. Adenomas of AFAP patients were often found to actually exhibit 'three hits' at the *APC* gene that mostly result in loss of the allele carrying the germline *APC* mutation. We assume that this actually leads to an optimization of the beta-catenin level, hence positively regulating the Wnt signal.

Recently, bi-allelic germline mutations in the base excision repair gene MutY homologue (*MYH*) have been associated with an autosomal recessively inherited predisposition to multiple colorectal adenomas. They are also referred to as *MYH*-associated polyposis (MAP). Here, we assessed the prevalence of *MYH* germline alteration in 79 unrelated polyposis patients in whom no *APC* mutation could be detected. The aims of the study were i) to assess the *MYH* mutation carrier frequency among Swiss APC mutation negative patients and (ii) to identify phenotypic differences between *MYH* mutation carriers and *APC / MYH* mutation negative polyposis patients. dHPLC and direct genomic DNA sequencing were applied to screen for mutation. Overall, 7 biallelic and 9 monoallelic MYH germline mutation carriers were identified. 1 out of 10 classical polyposis and 6 out of 35 attenuated polyposis patients carried biallelic *MYH* alterations, 2 of which represent novel gene variants (p.R 171q and p.R 231H). On the basis of our finding and earlier reports, *MYH* mutation screening should be considered if all of the following criteria are fulfilled: (1) presence of classical or attenuated polyposis coli, 2) absence of a pathogenic APC mutation, and 3) a family history compatible with an autosomal recessive mode of inheritance.

CHAPTER 1

1. General Introduction

1.1 Cancer and Colorectal Cancer (CRC)

Cancer is a common and devastating disease and remains one of the most serious public health issues. Despite recent progress in its treatments, so far few types are curable. Thus, cancer is under intense research because of the high prevalence and severe consequences leading to death. Most research aims to apply the knowledge about cancer in order to allow early diagnosis and understanding the mechanism of tumor development. The vast majority of cancers are considered to be sporadic and not primarily due to an inherited susceptibility, but there are 5-10 % of cancers are caused by inherited genetic changes [1, 2]. With the new technologies of molecular biology, it is possible to assess disease risk and guide genetic screening to prevent cancer and achieve successful cancer treatments.

Colorectal cancer is a disease characterized by the development of malignant cells in the lining or epithelium of the colon and rectum (*http://www.cancer.gov*). So far, an estimated 145,290 new cases are diagnosed worldwide, and 56,600 deaths from colorectal cancer will occur every year [3]. It is the third leading cause of cancer related death in most Western countries [4, 5]. Genetic researchs provided the possibility to identify persons at high risk for colorectal cancer because of an inherited predisposition to develop this malignancy [6]. Regular endoscopy is used as screening method to detect the presence of polyps at an early stage, thus prevent CRC development and detecting and treating early-stage cancers can lower the mortality rate for colorectal cancer [7, 8].

About 20% of CRC patients have a family history of colorectal cancer that suggests the influence of genetic factors [5]. Five-6% of CRCs occur are due to an inherited genetic predisposition. It is clear that other non-identified genes and background genetic factors contribute to the development of colorectal cancer, in conjunction with non-genetic risk factors, e.g. environmental factors, diet, etc [9].

1.2 Hereditary Colorectal Cancers

Hereditary CRCs can be divided into two distinct categories: Predisposition to colon cancer without and with pre-existing polyposis.

[1] Predisposition to colon cancer without pre-existing polyposis syndromes

Include hereditary nonpolyposis colon cancer (HNPCC) syndrome, Muir-Torre syndrome and Turcot syndrome (associated with glioblastoma) [10-13]. These syndromes are mainly caused by MMR gene mutations.

[2] Predisposition to colon cancer with pre-existing polyposis

Include adenomatous polyposis and harmartomatous polyposis syndromes. Adenomatous polyposis syndromes include familial adenomatous polyposis syndrome (FAP) and its variant attenuated familial adenomatous polyposis (AFAP) [14, 15] [16], *MYH*-associated polyposis (MAP), Gardner syndrome (variant of FAP) [17, 18]. In harmatomatous polyposis syndromes, the intestinal hamatomatous polyps are obligatory components of four inherited conditions: Peutz-Jeghers syndrome, Juvenile Polyposis syndrome, Cowden disease and its variant, the Bannayan-Ruvalcaba-Riley syndrome.

The genes associated with the above syndromes have been well identified (see Table 1.1).

Table 1.1 Genes associated with risk of different Hereditary Colorectal Cancer Syndromes

Syndrome	**Genes**	**Mean age at diagnosis**
Familial Adenomatous Polyposis (FAP)	*APC*	Age 40y [19]
Hereditary Nonpolyposis Colorectal Cancer (HNPCC)	Mismatch Repair (MMR) Genes, e.g..*MLH1*,*MSH2* ,etc	Age 42y [20, 21]
MYH associated polyposis (MAP)	*MYH*	Age 46y [10]
Juvenile Polyposis syndrome	*SMAD4, PTEN, BMPR1A*	Age 35y [24, 25]
Cowden Syndrome and Bannayan-Ruvalcaba-Riley syndrome	*PTEN*	Age 38-46y [25-27]

1.2.1 Predisposition to Colon Cance r without Pre-Existing Polyposis - Hereditary Nonpolyposis Colorectal Cancer (HNPCC)

HNPCC originally called "cancer family syndrome" or "Lynch syndrome" [28], accounts for about 2% to 5% of all colorectal cancers. Most HNPCC patients do not have numbers of polyps. It is caused by germline mutation in one of the mismatch repair genes (MMR) and is inherited in an autosomal dominant fashion. The mean age of cancer diagnosis is approximately 42 years (see Table 1.1) compared with 64 years in sporadic colorectal cancer [6, 29]. About 70% of HNPCC cancers develop in the proximal side of colon, coecum to transverse colon [30].

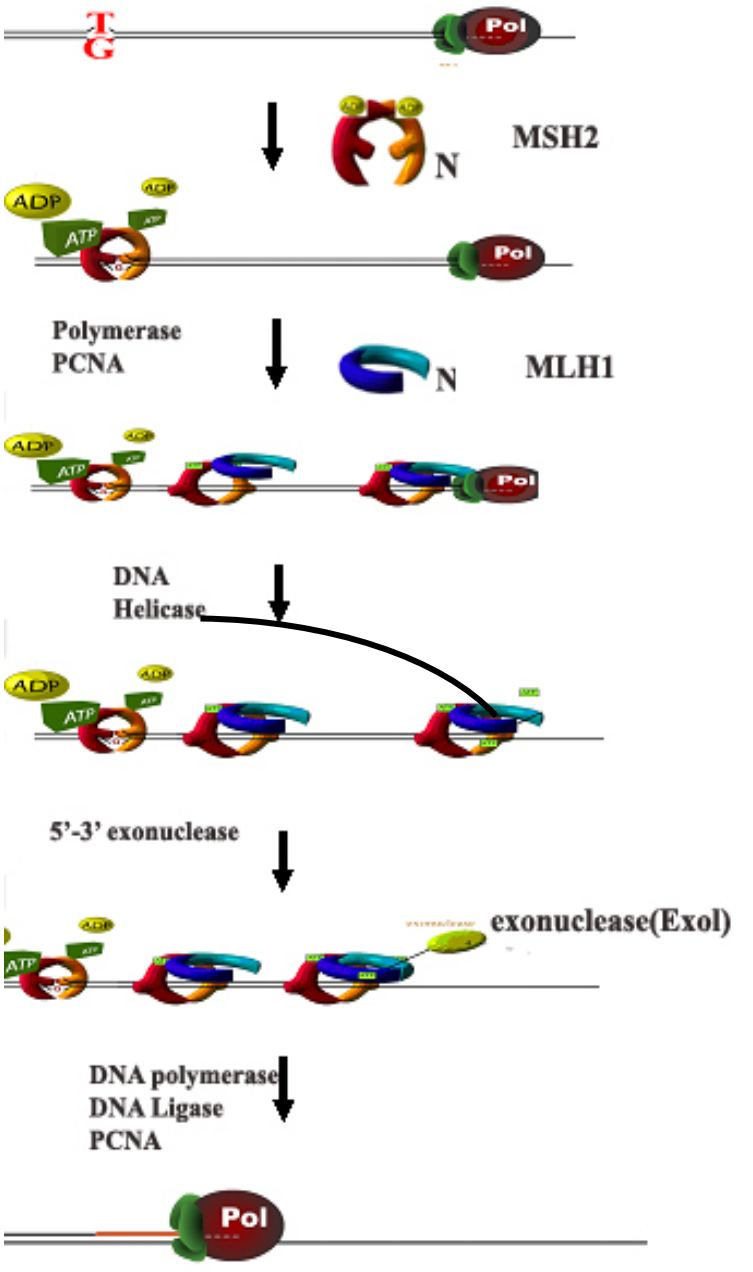
Associated extracolonic cancers are often found in HNPCC patients. These include cancer of the endometrium, stomach, ovary, ureter and renal pelvis, hepatobiliary tract, brain (Turcot-syndrome); sebaceous gland (Muir-Torre syndrome) as well as carcinoma of the small bowel [31-34].

Genetic testing of HNPCC

To define HNPCC families, the Amersterdam criteria I were established with the goal to identify the molecular basis of this disease on common clinical framework by the International collaborative group (ICG) in 1990 [35]. 10 years later, in order to improve the diagnosis of HNPCC clinically, ICG developed revised criteria - Amersterdam Criteria II to appreciate the HNPCC families who carry germline mutation of mismatch repair genes and have extracolonic cancers but do not meet the Amsterdam Criteria I [36]. In 1996, the Bethesda Guidelines were proposed for selection of patients whose tumors should be tested for microsatellite instability (MSI) [37]. This guideline was updated recently to so called revised Bethesda Guidelines [38, 39].

HNPCC results from germline mutation of one of several DNA mismatch repair (MMR) genes [11, 13]. These genes are essential to maintain the fidelity of DNA during replication. Mutation of these genes results in deficient DNA mismatch repair activity. These genes include *MLH1* (human mutL homolog 1), *MSH2* (human mutS homolog 2) [40], *PMS1* and *PMS2* (human postmeiotic segregation 1 and 2) [11] and *MSH6* (human mutS homolog 6) .

In HNPCC patients with a MMR gene mutation, 90% of mutations are found in *MLH1* and *MSH2* and approximately 10% patients carry *MSH6* mutations [41]. Clinical features of HNPCC are tightly related with the



The mismatched nucleotides were introduced into the newly replicated strand.

MSH (MSH2/MSH6, MSH2/MSH3) heterodimers recognize the mismatch loop formed by ATP and ADP.

MLH (MLH/PMS) protein interacts with ATP-bound MSH clamps and signal is transferred.

Helicase recognizes and replace the incised DNA strand

Single Strand DNA (ssDNA) is captured.

Mismatched nucleotides containing strands are elimated by exonucleases (EXO I)

Replication complex re-synthesis excised strand, error was repaired

Figure 1.1: Molecular Switch Model of Mismatch Repair [45]

mutation of these MMR genes (see Table 1.2). These MMR genes form different heterodimers (see Table 1.2) to participate in the mismatch repair process (see Figure 1. 1). All MMR gene mutation carriers are at a 50% risk of passing the altered gene to their offspring according to the mechanism of autosomal dominant inheritance [42] .

Understanding the basic function of these MMR genes is essential to better understand the mechanisms of HNPCC development and develop methods for detecting these gene mutations [43]. Basically, the primary function of the MMR pathway is responsible for the recognition and correcting of the mispairing of DNA nucleotides bases and the insertions or deletions that are frequently present during normal replication. It is essential to maintain fidelity of genomic DNA [44]. Haploinsufficient cells have normal or nearly normal repair activity, but inactivation of both alleles of MMR genes will result in loss of DNA repair activity [21].

Table 1.2*:* Clinical features associated with germline mutations in the MMR genes associated with a predisposition to HNPCC

Gene	Chromosome Locus	Heterodimer	Phenotypic features of HNPCC	Total numbers of Mutations *
MLH1	3p21.3[12]	MutL homologue 1 protein. Interact With PMS2, MLH3, PMS1, MLH2	Typical HNPCC, 30% of mutations are of the missense type whose phenotypic manifestations may vary [49] [30].	409
MSH2	2p22-p21[42]	MutS homolgues protein 2. Interact with MSH3 and MSH6	Typical HNPCC. Patients have more extracolonic cancer than in MLH1 mutations carriers. Is also the major gene underlying Muir-Torre syndrome[30].	337
MSH6	2p16[50]	MutS homologues 6 protein. Interact with MSH2	Typical or atypical HNPCC. Late CRC onset, frequent occurrence of endometrial cancer, distal location of colon cancers and low degree of MSI in tumors [51].	81
PMS2	7p22[11]	Human postmeiotic segregation 2 protein. Interact with MLH1.	Typical or atypical HNPCC. The penetrance of mutations may vary [11]	11
MLH3	14q24.3[52]	MutL homologue 3. Interact with MLH1.	Majority are missense mutations. Atypical HNPCC.	11

			Characteristic as distal location of colorectal cancers [52]	

*These data are extracted from Human Gene Mutation Database at the Institute of Medical Genetics in Cardiff (http://www.hgmd.org), May, 2006

This hypermutable state within the cell has been shown by the insertion or deletion of monoucleotide, dinucleotide, or trinucleotide base pair repeats in the microstatellite tracts in the tumor DNA [46]. Microsatellite sequences are short repetitive sequences throughout the genome [47]. When these sequences are not replicated correctly and not repaired by the MMR proteins, this is called microsatellite instability (MSI). MSI can be detected in around 90% of colorectal cancers from individuals with HNPCC. It has been suggested that mutations in the human mismatch repair genes are responsible for the MSI of the HNPCC tumors [48].

Based on all the knowledge of MMR genes, immunohistochemistry (IHC) and MSI analysis are the first round molecular testing performed in the tumor of HNPCC patients [29, 53]. IHC is a simple assay to screen the protein expression of MMR genes. Loss of expression in any of these proteins suggests germline mutation analysis [54, 55]. If the tumor exhibits MSI, germline mutation will be considered [56].

Commercial sequence testing is available to search for mutations in *MLH1* and *MSH2*. Clinical and cost consideration may guide testing strategies. MLPA and multiplex PCR, southern blot are the methods applied to detect genomic deletion or duplication after sequencing fails to detect the mutation [57]. Once a genetic alteration has been identified in a HNPCC family, the same alteration is easier to be tested for in other affected family members [58].

1.2.2 Predisposition to colon cancer with pre-existing polyposis – Familial adenomatous polyposis (FAP) and MYH associated polyposis (MAP)

FAP is one of the most clearly defined disorders, characterized by hundreds to thousands of polyps in the colon and rectum, which usually develop during late childhood or early adult life [59]. Extracolonic manifestations are variably present, such as osteromas, epidermoid cysts, desmoids, congenital hypertrophy of retinal pigment epithelium (CHRPE) and other cancers [60, 61]. Attenuated FAP (AFAP) is characterized by the presence of fewer less than 100 adenomatous polyps and later clinical manifestation [62]. FAP and AFAP are comparatively rare, representing about 0.5-1% of all CRCs [63].

Genetic testing of FAP and AFAP

FAP is due to mutations of the *adenomatous polyposis coli (APC*) gene. The *APC* gene is located at chromosome 5q21 and encompasses 15 exons. Exon15 comprises 75% of the coding sequence and is also the position where mutations most commonly occur.

The *APC* gene is a tumor suppressor gene which encodes a multifunction protein of 2843 amino acids (Figure 1.2). It is involved into the Wnt singalling pathway. The aberrant activation of Wnt pathway is considered as a major oncogenic mechanism for many tumor types. In FAP, the mutation of *APC* leads to the degredation of ß-catenin [64]. About 682 different disease associated germline mutatios in the *APC* gene have aleady been described (www.hgmd.org). The clinical features of FAP appear to be generally associated with the location of the mutation in the *APC* gene and the type of mutation (framshift and missense mutation or large deletion). In Table 1.2, the major mutations and their related phenotypes are summarized. Genotype-phenotype correlations are useful in increasing the accuracy and effectiveness of screening, surveillance and treatment [65, 66], e.g. mutation at codon 1309 (a deletion of AAAG in the 1309 codon) is the most frequently observed mutation (in 10% of FAP patients) and associated with severe colonic polyposis. (Table 1.3 and Figure 1.2)

Table 1.3: Mutation site and their phenotype in *APC* gene [67]

Codons 457 and 1444	**CHRPE (congenital hypertrophy of the retinal pigment epithelium)**
Codons 1250-1464(mutation cluster region)	**Severe FAP, develop >5000 polyps**
Codon 1309	**Severe phenotype**
Codons 1403-1578	**Desmoid**
5' and 3'end, exon 9	**Attenuated polyposis, develop <100 polyps**

Although an *APC* mutation is responsible for most of FAP families, there are some families that display the phenotype of classical FAP or AFAP syndrome without *APC* mutations. MAP (MutYH-associated polyposis) is a recently described colorectal adenoma and carcinoma predisposition syndrome that is associated with inherited mutations of the human MutY

homologue gene (*MYH)*. It is associated with 10-100 polyps. MAP is inherited in an autosomal recessive manner [10].

MYH is a base-excision-repair gene, which encodes a monofunctional BER glycosidase that is capable of correcting oxidative DNA damage. Failure to correct this damage can lead to the formation of 8-oxoG, causing an increase in G:C/T:A transversions. It has been reported that germline *MYH* mutations cause approximately 1-3% of all unselected colorectal cancers [68, 69]. All these findings have important implications for accurate genetic testing of the patients without *APC* germline mutation who have less than 100 adenomas.

Genetic testing for *APC* and *MYH* alterations are performed on leukocyte-derived DNA of the patients. There are several methods applied so far, e.g. direct sequencing, mutation screening with single strand conformational polymorphism (SSCP), denaturing gradient gel electrophoresis (DGGE), protein truncation test (PTT), denaturing high performance liquid chromatography (dHPLC) and multiplex ligation-dependent probe amplification (MLPA).

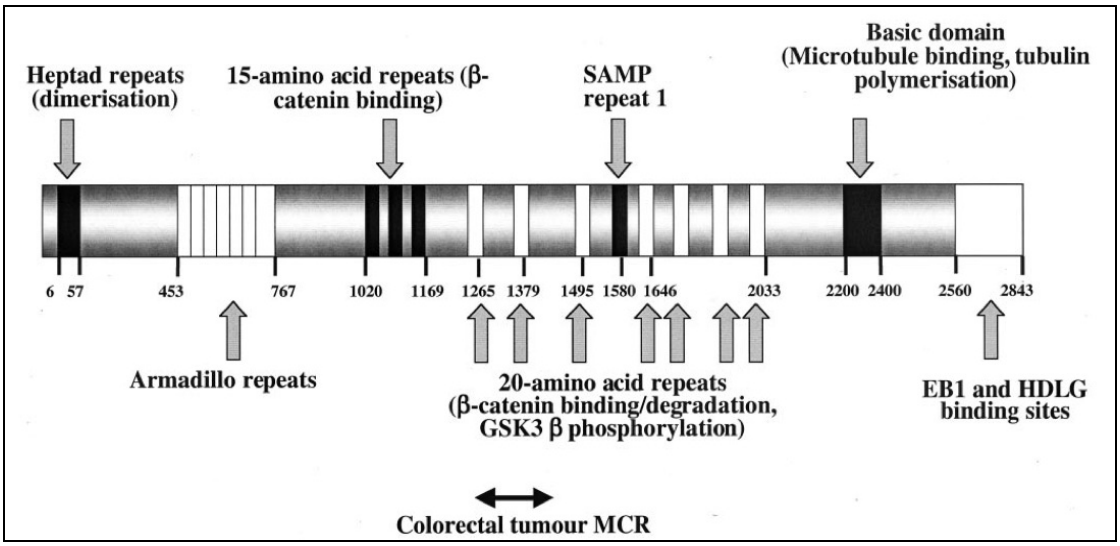


Figure 1.2: **Structural features of the *APC* protein.** Most of the mutations in *APC* occur in the mutator cluster region (MCR) and create truncated proteins. The truncated proteins contain ASEF and β-catenin binding sites in the armadillo-repeat domain but looses the β-catenin regulatory activity which is located in the 20-amino acids repeat domain. Somatic mutations are selected more frequently in FAP patients with germ-line mutations outside of the MCR [72] [73].

Since the majority of *APC* mutations result in the formation of a truncated *APC* protein product, the PTT is the first screening method for genetic testing. With rigorous PTT testing and the use of other screening methods, 90% of mutations can be detected in classical FAP [70]. If an *APC*

pathogenic mutation is detected in the index patient, the same *APC* mutation will be found in all affected family members [71]. If there is no *APC* mutation found in the classical FAP and AFAP phenotype, *MYH* mutation screening is performed within in these patients [58].

1.3 Tumorigenesis of Colorectal Cancers

Colorectal cancer develops as a result of the pathologic transformation of normal colonic epithelium to adenomas of progressively larger size and ultimately to an invasive cancer. Fearon and Vogelstein proposed a multistep progression model in 1990 (adenoma-carcinoma sequence model) [74]. This multistep progression requires years and is accompanied by a number of genetic alterations in tumor suppressor genes and oncogenes, which contribute to the development of the malignant phenotype [75]. A morphological transition corresponding to the genetic mutations from normal colonic mucosa to a benign tumor (adenoma) and to a malignant carcinoma can be observed [74]. (Figure 1. 3)

At least two pathways leading to colon cancer development are identified. They are "gatekeeper" and "caretaker" pathways (Figure 1. 4), which are initiated by "gatekeeper" genes and "caretaker" genes[76].

"Caretaker" genes and "Gatekeeper" genes were distinct by Kinzler and Vogelstein in the determination of cancer in 1997 [75]. "Gatekeeper" genes directly regulate the growth of cells and "caretaker " gene are the genes controlling cell proliferation and cell apotosis directly.

Caretaker pathway

Caretaker genes are the genes controlling cell proliferation and cell apotosis indirectly. Therefore, in the pathway initiated by mutations in caretaker genes, neoplasia occurs indirectly. Inactivation of caretakers leads to genetic instability that results in an elevated mutation rate of all genes, including gatekeeper genes [75, 77]. Accumulation of genetic alterations in other genes that directly control cell apoptosis or cell death will further promote tumor progression. Known caretaker genes include mismatch repair (MMR) genes which cause HNPCC [78].

Hence in HNPCC, patients have inherited a mutant allele of a caretaker (MMR) gene. Then a subsequent somatic mutation of the normal allele inactivates the MMR system in the cell. When the cell accumulates mutations of the MMR genes and other growth controlling genes, tumor

formation is promoted. MMR inactivation causes the infidelity of replication of repeated sequences (microsatellites) in tumor, microsatellite instability (MSI) is the first hallmark of the HNPCC. The HNPCC tumors also arise from adenomatous polyps (but very few or even without polyps), based on the tumorigenesis model, these polyps contain K-ras mutation and "gatekeeper" gene mutation, e.g *APC* mutations[79]. But several target gene like *TGFβRII* , *IGFIIR , PTEN, BLM , TCF-4,Bax* have been found somatically affected in gastrointestinal tumors [80-82].

Gatekeeper Pathway

Gatekeepers are the genes that directly regulate the growth of tumors by inhibiting cell growth or promoting cell death. It is assumed that each cell type has only one or a few gatekeepers [79].

In the majority of CRC, the Wnt-involved *APC* gene serves as the gatekeeper gene. It is one of early and frequently mutated genes in CRC [83]. Inactivation of *APC* gene will cause unbalanced cell growth, i.e., the cell birth rate is over that of cell death, and then the tumors begin to grow [75].

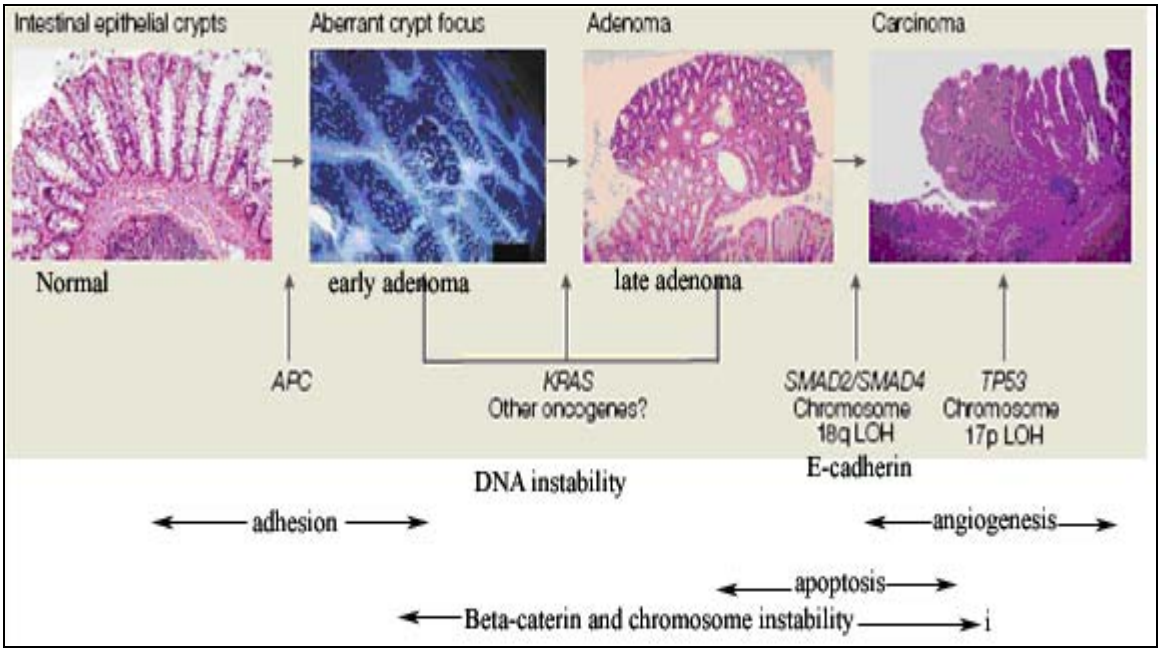


Figure 1.3: Pathways that control colorectal tumorigenesis. Mutations in the *APC*/β-catenin pathway initiate the neoplastic process through microscopic aberrant crypt foci, resulting in small benign tumors (adenomas). As these tumors progress, mutations in other growth-controlling pathway genes (such as K-Ras, B-Raf, PI3K, or p53) accumulate and adenomas become carcinomas, which eventually metastasize. The process is accelerated by mutations in caretaker genes [72].

Based on the adenoma-carcinoma tumorigenesis model (Figure 1. 3), oncogene mutations (e.g, K-ras and C-myc) are often required for tumor progression after the *APC* mutation. In general 50% of all colon cancers show K-ras mutations at the early stages of tumor progression. Their mutation frequency decreased during progress [84].

When adenoma formation is initiated by *APC* gene, it is promoted to grow faster to a large adenoma. Other tumor suppressor genes like *SMAD2/SMAD4* and *DCC* (*deleted in colon cancer*) start to be involved in the progress. *DCC* was lost in 50% of late adenomas and carcinomas but not in intermediate adenomas[85]. Studies have shown that inactivation of *SMAD4* gene resulted in more malignant adenomas with extensive stromal proliferation and invasive growth [85]. *DCC* gene and *SMAD2/SMAD4* are all located on chromosome 18. Thus, loss of activity of one or more genes on Chr18 does appear to be an important step in tumor development.

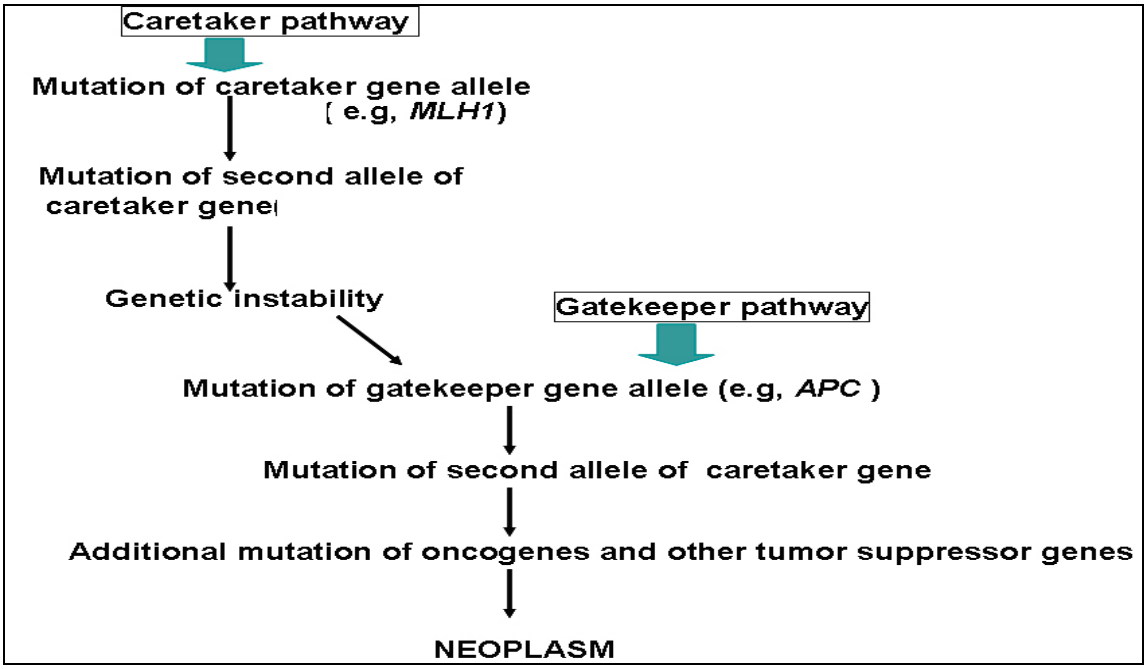


Figure 1.3: Model of genetic alteration in the development of colorectal cancer [74,76]

When finally a P^{53} mutation (on chromosome 17q) occurs, the balance on cellular proliferation and apoptosis is lost due to the failure of cellular apoptosis. At this point, it is assumed cells

accumulated all the genetic alterations and the adenoma progresses to carcinoma accompanied by chromosome instability and aneuploidy [86].

In colorectal cancer, FAP and 85% of sporadic CRCs followed this pathway [87]. Both FAP and sporadic carcinogenesis accumulate mutations by the "adenoma-carcinoma sequence" [88].

The tumorigenesis model and pathways discussed above are believed to contain the backbone of genetic alteration in the majority of sporadic CRCs.

"Two- hit" hypothesis in colorectal cancers

Nonetheless, gatekeepers and caretakers are all tumor suppressor genes in both pathways. The tumor suppressor genes followed Knudson's "two-hit" hypothesis to initiate the tumor growth. In hereditary cancers, tumor suppressor genes (TSG) carry a germline mutation, so it usually only requires a second somatic mutation for tumorigenesis, while in non-hereditary cancer (sporadic cancers), two somatic mutations need to be in the same somatic cell to inactive TSG in order to initiate tumor formation[89]. (Figure 1. 4)

This hypothesis was first developed for retinoblastoma tumors. Later it was found that most dominantly inherited cancers followed this hypothesis. Studies have been shown that this second somatic event may arise by a variety of molecular mechanisms, for example new intragenic mutations, gene deletions, chromosomal loss or somatic recombination [89, 90].

It was understandable that people who inherit an inactivated copy of a tumor suppressor gene had a higher risk of developing the associated form(s) of cancer than people born with two normal copies, as postulated in two-hit model. Indeed, it was shown that in the tumors of these predisposed patients, the remaining wild-type copy of the tumor suppressor gene was lost, a process referred to as loss of heterozygosity (LOH) [90]. LOH leads to either deletion of the tumor suppressor locus or "reduction to homozygosity" (two alleles occur to be identical without net loss of genetic material) [91, 92]. Later studies confirmed that this concept is also suitable for other tumor suppressor genes.

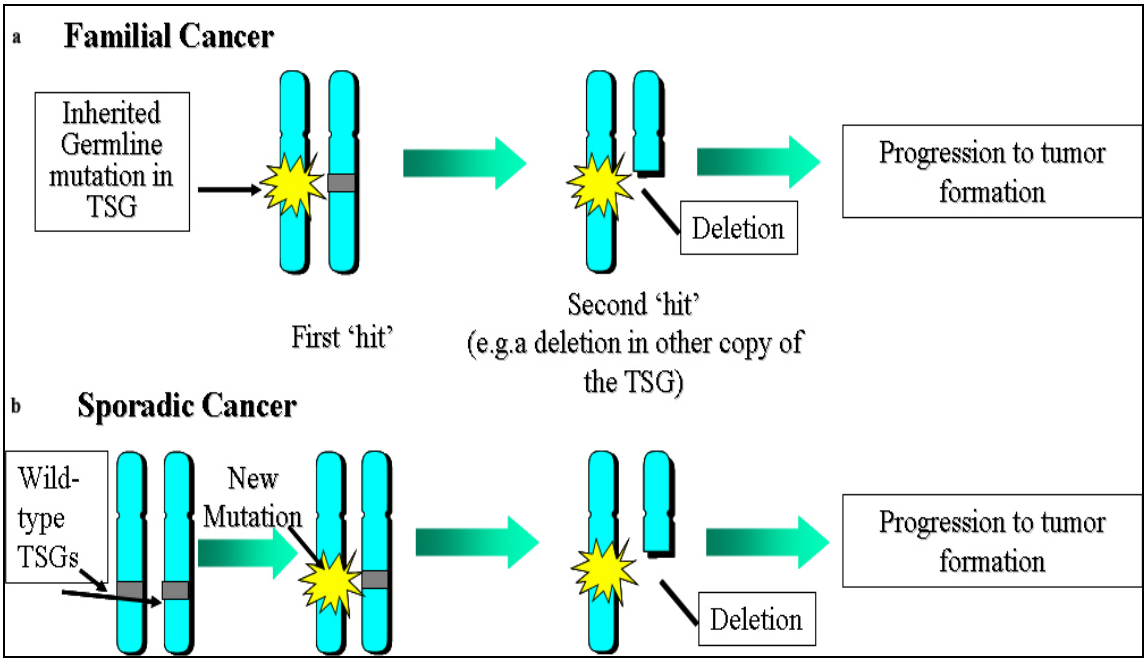


Figure 1.4: Knudson's two-hit hypothesis for tumorigenesis involving a tumor suppressor gene (TSG) One pair of chromosomes is depicted, with one TSG (the normal gene (grey), the mutated gene (yellow star), and deletion of the gene (absence) are shown. (a) In familial cancer, Individuals inherited a germline mutation of the TSG as first 'hit' in every cell and require only one subsequent 'hit' in a cell to initiate a cancer (b) Normal individuals have two normal copies of the TSG, so two independent 'hits' (mutations) are required in the same somatic cell to initiate a cancer [90].

In FAP syndrome, in agreement with Knudson's "two-hit" hypothesis, inactivation of both *APC* alleles can be detected in most intestinal tumors at early stages of tumor development [93]. However, detailed mutation analysis of tumors from patients with FAP and *APC* min mice has shown an interesting result: The position and type of the second hit in FAP polyps depends on the localization of the *APC* germline mutation. This is claimed in a two-hit model, in which any LOH mutation would result in tumor formation. It showed that the dependence between germline mutation and the resulting spectrum of somatic mutations that successfully lead to tumor formation is more complex than suggested previously. Most of time the somatic mutation and germline mutation are linked to the multi-function region of the *APC* gene. This multifunction region contains three 15-amino-acid repeats and seven 20-amino-acid repeats (AAR). which act as the binding domain of ß-catenin and are crucial for downregulating ß-catenin [94-96]. Somatic mutation analysis of polyps from different FAP patients also showed how *APC* is inactivated and starts tumor formation in association with the activation of ß-catenin signaling rather than at the

complete loss of regulatory function of *APC* within this signaling pathway. This is called the "just right" model [91] . Therefore, in FAP the somatic mutation of *APC* depends on its germline defect. Additional specific subset of somatic mutations will successfully lead to tumor formation in the colon and rectum [97] .

Compared to the good understanding of the *APC* gene in FAP patients, little is known about the second hit and the molecular mechanisms of the malignant tumor initiation and progression in HNPCC with MMR gene mutations. In large, the ability of defective MMR genes to cause HNPCC appears to follow the "second hit" hypothesis, in which germline mutations confer predisposition but need a second hit for tumor initiation.

Generally, heterozygous mutants of mismatch repair genes are still mismatch repair proficient [98-100]. However, when the wild type allele of the gene is also lost through somatic events (second hit), the tumor will progress[101]. This leads to replication errors (RER) in the short repeat sequence, that is why we believe microsatellite instability is caused by the somatic inactivation of the corresponding second mismatch repair allele ("second hit") [46]. There are several mechanisms possibly responsible for the inactivation of the mismatch repair genes, e.g, point mutations, allelic losses as well as epigenetic processes such as aberrant methylation of cytosine and guanine rich promoter regions (CpG islands). Previous reports have shown that LOH is frequently found in tumors from HNPCC patients with germline MMR mutations [102]. But in 2001, Kruse etal have shown that in Muir-Torre syndrome which is caused by *MSH2* germline mutation, loss of heterozygosity is not the preferred model of somatic inactivation of the second MSH2 allele. Therefore, it remains unclear which somatic inactivation mechanisms account for tumor initiation in patients with known MMR germline mutations.

1.4 Aims of this thesis

In this thesis, we investigated the frequence and nature of large genomic rearrangements in MMR mutation negative patients (Chapter3.1), the prevalence of germline mutations of *MYH* inAPC mutation negative polyposis patients (Chapter 4.2). Subsequently, we did a detailed investigation of the in somatic alterations in cancers from HNPCC and AFAP patients to characterize the second hit and third hits (Chapter 3.2, Chapter 3.3, Chapter 4.1) in order to understand the mechanism of the tumor initiation and progression in colorectal cancer.

CHAPTER 2

2. Methods

2.1 DNA extraction

DNA extraction from peripheral blood

Genomic DNA was isolated from EDTA blood applying the salting-out procedure described by Miller et al (103). Briefly, 10 ml of blood were mixed with 30 ml of EL buffer (155mM NH_4Cl, 10 mM $KHCO_3$, 1 mM EDTA, pH 7.4) and left on ice for 15 minutes. The lysate was centrifuged at 2000 rpm for 10 minutes, washed twice with EL buffer and the intact leukocyte pellet resuspended in NL buffer (10 mM Tris-HCl, pH 8.2, 400 mM NaCl, 2 mM Na_2EDTA, 1% SDS and 200µg/ml proteinase K) and incubated overnight at 37°C. The next day, 1 ml of 6M NaCl was added and vigorously shaken followed by centrifugation to remove cellular proteins. The supernatant containing the DNA was placed in a fresh tube and the DNA precipitated with ethanol. The resulting DNA pellet was washed with 70% ethanol, dried briefly, and then suspended in 1 ml of TE buffer (10 mM Tris.HCl, pH 7.5, 0.1 M EDTA) for over night until all the pellet dissolved(103) .

DNA extraction from paraffin-embedded formalin-fixed tissue

After histopathological classification of hematoxylin/eosin-stained, formalin-fixed tissue blocks, a representative portion of the tumor (adenoma or carcinoma) with an average tumor contents of $\geq$ 70% was scraped off and DNA extraction performed according to the Qiaamp tissue kit' protocol (QIAGEN,Basel, Switzerland). Briefly, 180µl of buffer ATL and 20µl proteinase K (20 ng/ µl) were added to each tumor sample, which then were incubated overnight at 55ºC for digestion, until the tissue was completely lysed. Next day, 200µl of buffer AL were added and incubated at 70ºC for 10 min, followed by the addition of 210µl of ethanol (100%) and mixed thoroughly by vortexing. Then mixture was transferred into Qiaamp spin column and centrifuged at 10'000 rpm for 5 min. After having discarded the filtrate, 250µl of washing buffer AW were added 2 times and centrifuged at full speed (14'000 rpm). Finally, the DNA was eluted twice with 50µl -200µl of buffer AE.

Quantitation of genomic DNA

Measure DNA concentration by Eppendor Biophotometer (Eppendorf AG, basel, Switzerland). Quantify DNA by diluting 5µl DNA into 55µl distilled water (1:12 dilution). An absorbance of 1 unit 260nm corresponds to 50µg DNA/µl.

2.2 RNA extraction

Total RNA was isolated by Qiagen RNeasy Mini Kit from Heparin blood of patients according to the protocol supplied by the manufacturer (QIAGEN, Basel, Switzerland). Collect blood cells (not over maximum $1x10^{7)}$ by centrifugation in 350µl RLT buffer (lysis buffer), disrupt and homogenize the samples to break down genomic DNA and reduce viscosity of the lysate. After centrifuged Lysate for 3 min at 14000rpm, supernatant is transferred to the 350 µl 70% ethanol. 700 µl of sample was applied to an RNase mini spin column. The column was centrifuged at maximum speed. 700RW1 buffer is added to the RNeasy column, in cubate column for 5 minutes, and centrifuge it. RPE buffer 500 was added to RNeasy column following, and also centrifuge at maximum speed to wash the membrane. 30ul DEPC (RNase free) water was added as elution buffer into column, samples were collected by centrifuging it for 5min at maximum speed.

Quantitation of total RNA

Measure DNA concentration by Eppendor Biophotometer (Eppendorf AG, basel, Switzerland). Quantify RNA by diluting 5 µl RNA in 55 µl DEPC water (1:12 dilution). An absorbance of 1 unit at 260nm corresponds to 40µg RNA/ml.

2.3 Microsatellite Marker Analysis

Microsatellite Marker analysis of HNPCC tumors

For MSI analysis, genomic DNA and tumor DNA were investigated using a panel of microsatellite markers. 11 microsatellite markers were applied for analysis. They are located at different chromosome corresponding to different genes (*MSH2, MSH6, MLH1,c-kit ,3-beta-HSD,APC*). They are Marker DS123, D2S2227, D2s2369 and BAT 26; D3s1597, D3s3611, D3s3594 and D3s 3601; BAT 25; BAT 40 and D5S346. All these markers shared similar PCR amplification process. Within PCR amplification, 50-100ng of genomic DNA and tumor DNA

were mixed with 15µl true allele mix (Applied Biosystems, Rotkreuz, Switzerland), PCR reaction is performed on Eppendorf Mastercycle machine(Eppendorf AG, basel, Switzerland). The PCR program was initiated by 94°C 12 minute to active hot-start Tag polymerase and denature the template, followed by 10 cycles at 94°C 15 seconds, annealing (detail see appendix 1) 15 seconds, 72°C 30 seconds, another 20 cycles were performed at 89°C 15seconds, annealing 15 seconds , 72°C 15 seconds, with final cycle at 72°C, 6 min. Primer sequences, product length and the labeling dye of these primers are shown in table 1 of appendix 1.

Subsequently, 2 µl of PCR products mixed with 18 µl deionized formamide (Applied Biosystems, Rotkreuz, Switzerland), 0.5 µl ROX500 size standard was added and the mixture was loaded onto an ABI PRISM 310 sequencing machine using the POP4 polymer (PE Applied Biosystems, USA). Analysis was performed by Genescan software and Genotyper 2.5 software. MSI was determined with respect to the number of microsatellite markers displaying allelic expansions or contractions. The interpretation of the presence of MSI, defined as the occurrence of novel alleles, followed the NCI workshop's recommendations (59): MSS: all the markers are stable; MSI-Low: >0- $\leq$30% markers are unstable; MSI-high: >30% of markers are unstable. Tumor samples from HNPCC patients were included as positive controls. Loss of heterozygosity (LOH) was defined as a >50% reduction in relative intensity of one allele compared to the other (104, 105).

Loss of heterozygosity analysis of the APC gene and MMR genes

In the case of germline nonsense mutations and large deletion in *APC* and MMR genes, loss of heterozygosity (LOH, allelic loss) analysis was performed using microsatellite markers: D5S346, D5S299 and D5S82 , D5S318,MBC, DS123, D2S2227 ,D2S2369, D3S1597, D3S3594, D3S 3601 (see primer sequence in appendix 1), which map on the location of these genes. In the case of germline (and somatic) frameshift mutations, LOH analysis was performed using oligonucleotide primers which encompassed the germline insertion/deletion, which was then used to access allelic loss. Standard methods of fluorescence-based genotyping on the ABI310 sequencer were used. Allelic loss was scored at any informative marker if the area under one allelic peak in the tumor was reduced by more than 50% relative to the other allele, after correction for the relative peak areas of the alleles found in constitutional DNA of the same patient.

2.4 Immunohistochemistry (IHC)

After MSI analysis, all tumors with MSI were analyzed by immunohistochemistry (IHC) experiment in Zurich (Dr. Giancarlo Marra, the Institute of Molecular Cancer Research, University of Zurich, Switzerland). Four micrometer serial sections from paraffin blocks were mounted on silanized slides, deparaffinized and rehydrated. Antigen retrieval was obtained by heating the sections in a pressure cooker at 120°C for 2 min in 10mM citrated-buffered solution (pH 6.0). DAKO peroxidase blocking reagent and goat serum were sequencially used to suppress nonspecific staining due to endogenous peroxidase activity and nonspecific binding of antibodies, respectively. Incubations with primary monoclonal antibodies were performed as follows: anti*MSH2*: 24 hours at 4C with Ab NA26 (Oncogene Research), 1μg/ml; antiMSH6:2 hours at RT with Ab G70220 (Transduction Laboratories), 4μg/ml; anti *MLH1*: 1 hour at RT with Ab 13271A (PharMingen), 1.2 μg/ml; anti-hPMS2: 24 hours at 4C with Ab 65861A (PharMingen), 3 μg/ml. After washing, anti-mouse secondary antibodies conjugated to peroxidase labelled polymer (DAKO EnVision+kit) were applied for 30 min at RT, and the peroxidase activity was developed by incubation with 3.3, diaminobenzidine (DAB) chromogen solution (DAKO). Sections were then counterstained slightly with hematoxylin.

2.5 Denaturing High Performance Liquid Chromatography (dHPLC)

The dHPLC method was developed primarily as pre-screening method in the identification of sequence variations in a number of disease genes (106, 107). dHPLC is based on the detection of heteroduplexes in short segments of DNA by ion-reversed phase high performance liquid chromatography (108, 109).Partial heat denaturation within an acetonitrile gradient leads to the separation of the DNA strands, resulting in the formation of hybrid wild type/mutant heteroduplexes. These heteroduplexes have a reduced column retention time and hence an altered mobility compared to their homoduplex counterpart. The big advantages of the dHPLC method include low cost, the use of automated instrumentation and the speed of the analysis (2.4 minutes to 5 minutes per sample). This technique has been successfully employed in the detection of mutations and polymorphisms in the Y chromosome, exons from the factor IX and neurofibromatosis type 1 genes10, rearranged transforming (RET), cystic fibrosis transmembrane conductance regulator (CFTR) and phosphatase and tensin homologue on chromosome (*PTEN*) genes (108) , *BRCA1* and *BRCA2* (110) and *MLH1* and *MSH2* (111).

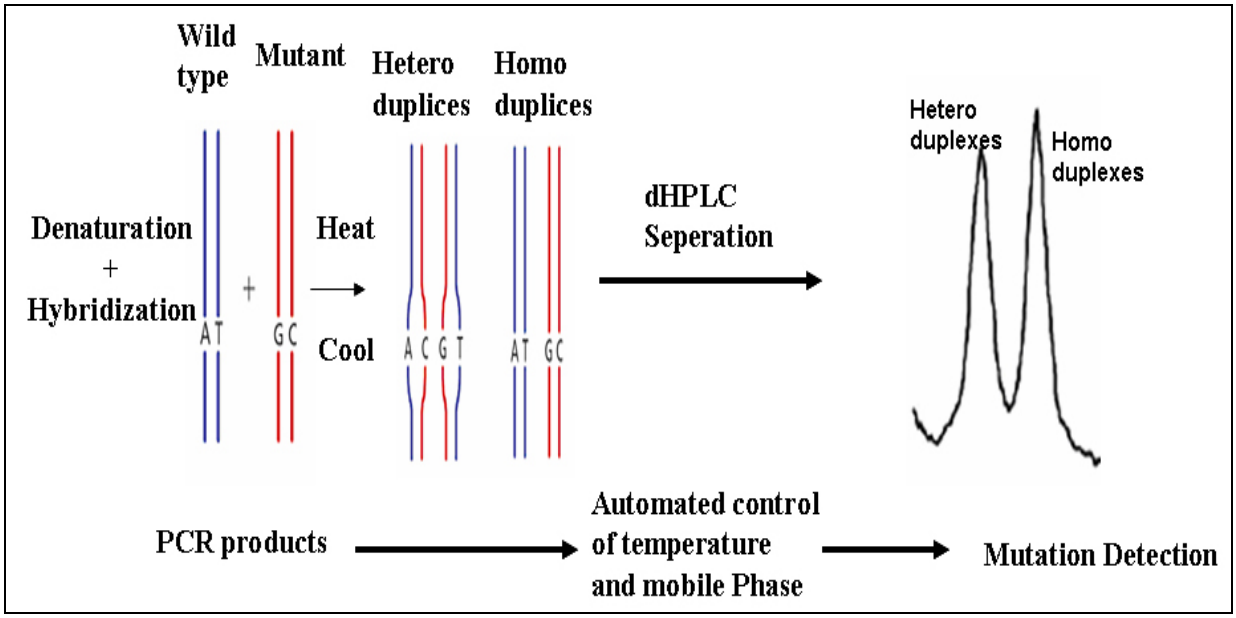


Figure 2.1: Outline of dHPLC method (109).

Screening of Mutation Cluster Region (MCR) of APC gene by dHPLC

The *APC* protein contains 2,843 residues with several structural motifs. The mutation cluster region is the region that encompass from codon 1250 to codon 1560. It contains two of the most commonly found pathogenic mutations, 5-bp deletions creating stop codons at positions 1061 and 1309. This region hosts three Armadillo repeats (15– and 20–amino acid repeats). The 20 amino acid repeats are important for *APC* mediatie β-catenin degradation (112).

DNA samples were amplified for mutations in the tumor DNA of *APC* attenuated mutation carrier by PCR based methods. Because the difficulties of tumor DNA amplification, we designed 12 primers to cover the whole MCR region with short PCR products. The final products were applied on the highly sensitive WAVE 3500HT dHPLC (Transgenomic, Crewe, UK). Melting temperatures for dHPLC were predicted by the Wavemaker software version 4.1.42 (Transgenomic). The different elution profiles were observed, in comparison to control samples (negative control samples and positive control samples) run in parallel. 12 different PCR products of *APC* MCR region were denatured by different melting temperature (see appendix 2). This method was also applied for *MYH* mutation analysis, see melting temperature of different exons of *MYH* in appendix 2.

2.6 Direct DNA Sequencing

After IHC screening or dHPLC screening, direct sequencing was applied to screen entire coding region of all the suspective MMR gene. DNA sequencing was also applied for the sample, which showed different patterns in dHPLC screening of *APC* MCR region and *MYH* coding region. PCR products were purified using the QIAquick PCR Purification kit (Qiagen, Basel, Switzerland). The sequencing reaction was performed using the Big Dye Teminator Cycle Sequencing kit (Applied Biosystems, Rotkreuz, Switzerland), according to the manufacturers' guidelines. Following purification using the DyeEx 2.0 Spin Kit (Qiagen, Basel, Switzerland) sequencing products were analysed on an ABI PRISM 310 Sequencing machine (Applied Biosystems, Rotkreuz, Switzerland). All the mutations identified in all genes were confirmed by sequencing in both, forward and reverse directions, and from at least 2 independent PCR products. Appendix 1 showed all the primers applied for the directly DNA seqeuencing.

2.7 Quantitive Multiplex PCR Ampilification (QMPA) detection

To amplify all 16 exons of *MSH2* and all 19 exons of *MLH1*, we motified the primer pairs and PCR conditions according to reference (113). Primers length and sequence are descriped in appendix 1. All the upstream primers were labeled with the fluorescent dye 6-FAM at their 5'end. Seven groups of multiplex PCR reactions were performed. Each group contained six pairs of primers encompassing both genes (Table 2. 1). Some primer pairs were present in two or more multiplex reactions as controls. The final volume of each multiplex PCR was 12.5 µl, containing 50ng of template DNA, 10 pmol of each pair of primers, dNTPs (final concentration 0.2 mM) and 0.25µl units Taq DNA polymerase (Invitrogen, Karlsruhe, Germany). Because of the different annealing temperatures of the primers, we performed PCR in two steps: the first step comprised 10 cycles starting with an annealing temperature corresponding to the highest Tm value of the primer set followed by a decrease of 1ºC/cycle.The second step comprised 10 cycles at a constant annealing temperature equal to the lowest Tm value of this primer set (detailed information in Table 2. 1). PCR was performed on Eppendorf Mastercycle PCR machine (Eppendorf AG, Basel, switzerland). Then 2µl PCR products were mixed with 18µl of deionized formamide plus 0.5µl ROX500 size standard. The mixture was analyzed on ABI 310 DNA Sequencer machine with POP4 polymer. (Applied Biosystems, Rotkreuz, Switzerland).

Table 2.1. Primer concentration and annealing temperature of QMPA

Group	***MSH2*** **volume of primer µl (50pm / µl)**	***MLH1*** **volume of primer µl (50pm / µl)**	**Annealing temperature**	**Cycles**
Group 1	exon3 (0.15+0.15) exon5 (0.1+0.1) exon9 (0.15+0.15) exon13 (0.2+0.2) exon16 (0.15+0.15)	exon 8 (0.15+0.15)	59°C in cycle 1 51°C in cycle 2	12(cycles) 12(cycles)
Group 2	exon 2 (0.3+0.3) exon 15 (0.1+0.1)	exon 1 (0.05+0.05) exon 8 (0.075+0.075) exon 10 (0.1+0.1) exon 16 (0.1+0.1)	58°C in cycle 1 51°C in cycle 2	12(cycles) 12(cycles)
Group 3	exon 6 (0.3+0.3)	exon 5 (0.15+0.15) exon 6 (0.2+0.2) exon 10 (0.2+0.2) exon 18 (0.2+0.2) exon 19 (0.15+0.15)	55°C in cylce 1 51°C in cycle 2	12(cycles) 12(cycles)
Group 4	exon 3 (0.15+0.15) exon 14 (0.15+0.15)	exon 3 (0.2+0.2) exon 4 (0.15+0.15) exon 11 (0.15+0.15) exon 12 (0.1+0.1)	57°C in cycle1 51°C in cycle 2	12(cycles) 12(cycles)
Group 5	exon 1 (0.2+0.2) exon 10 (0.2+0.2) exon 14 (0.15+0.15)	exon 2 (0.1+0.1) exon 7 (0.05+0.05) exon 14 (0.05+0.05)	55°C in cycle1 51°C in cycle2	12(cycles) 12(cycles)
Group 6	exon 2 (0.2+0.2) exon 4 (0.2+0.2) exon 7 (0.15+0.15) exon 12 (0.2+0.2)	exon 7 (0.1+0.1) exon 13 (0.15+0.15)	58°C in cylce1 52°C in cycle 2	12(cycles) 12(cycles)
Group 7	exon 8 (0.2+0.2) exon 11 (0.2+0.2)	exon 2 (0.2+0.2) exon 9 (0.1+0.1) exon 15 (0.2+0.2) exon 17 (0.1+0.1)	56°C in cylce1 51°C in cycle2	12(cycles) 12(cycles)

Detection of Genomic Deletions

Detection of genomic deletion was based on the comparison of the peak areas of different exons amplified in a multiplex PCR. All the samples were analyzed by Genescan software and the Peak areas were calculated by Genotyper software 2.5 (Applied Biosystems) and exported to an excel spread sheet. For evaluation of the different exons of a multiplex PCR group, one peak was taken as a reference (Pr), the other five peaks as intended peaks (Pi). In a first step the ratio between peak areas (Pi/Pr) was evaluated. To calculate the copy number, the ratio Pi/Pr was further divided by the ratio of the same exons obtained in negative control samples (Ci/Cr). A value around 1.0 of the (Pi/Pr) ./(Ci/Cr) ratio was regarded as absence of deletion; a value reduced to 0.5 was interpreted as a heterozygous deletion of the intended exon. Since experimental conditions do not allow differentiation of deletions from duplications when one of the two genes is entirely involved, an additional multiplex PCR that also amplifies fragments from a third gene was performed in such cases (see example in chapter 3.1).

2.8 Multiplex ligation-dependent probe amplification (MLPA) detection

For the detection of aberrant copy numbers in the *MLH1* and *MSH2* genes in constitutional (leukocyte-derived and tumor DNA), the *SALSA P003 MLH1 / MSH2 test* MLPA kit (MRC Holland, Amsterdam, The Netherlands) was used (114).The kit contains probes for the 16 exons of *MSH2* and the 19 exons of *MLH1* as well as 7 probes located on different chromosomes as controls. DNA samples from 2 known germline deletion carriers (*MLH1 exon1_10del; MSH2 exon8_15del*) as well as from 10 healthy probands were used to confirm the sensitivity and specificity of the method. Each mutation was confirmed on a second, independently drawn blood sample from the respective patient.

The MLPA reaction contains three steps to complete the reaction: Probe hybridization, Ligation and PCR amplification. All the reactions were performed on Eppendorf Mastercycle PCR machine (Eppendorf AG, Basel, switzerland).PCR products were analyzed on ABI PRISM310 sequencing machine in POP 4 polymer with 47 cm capillary. (See detail in Appendix III)

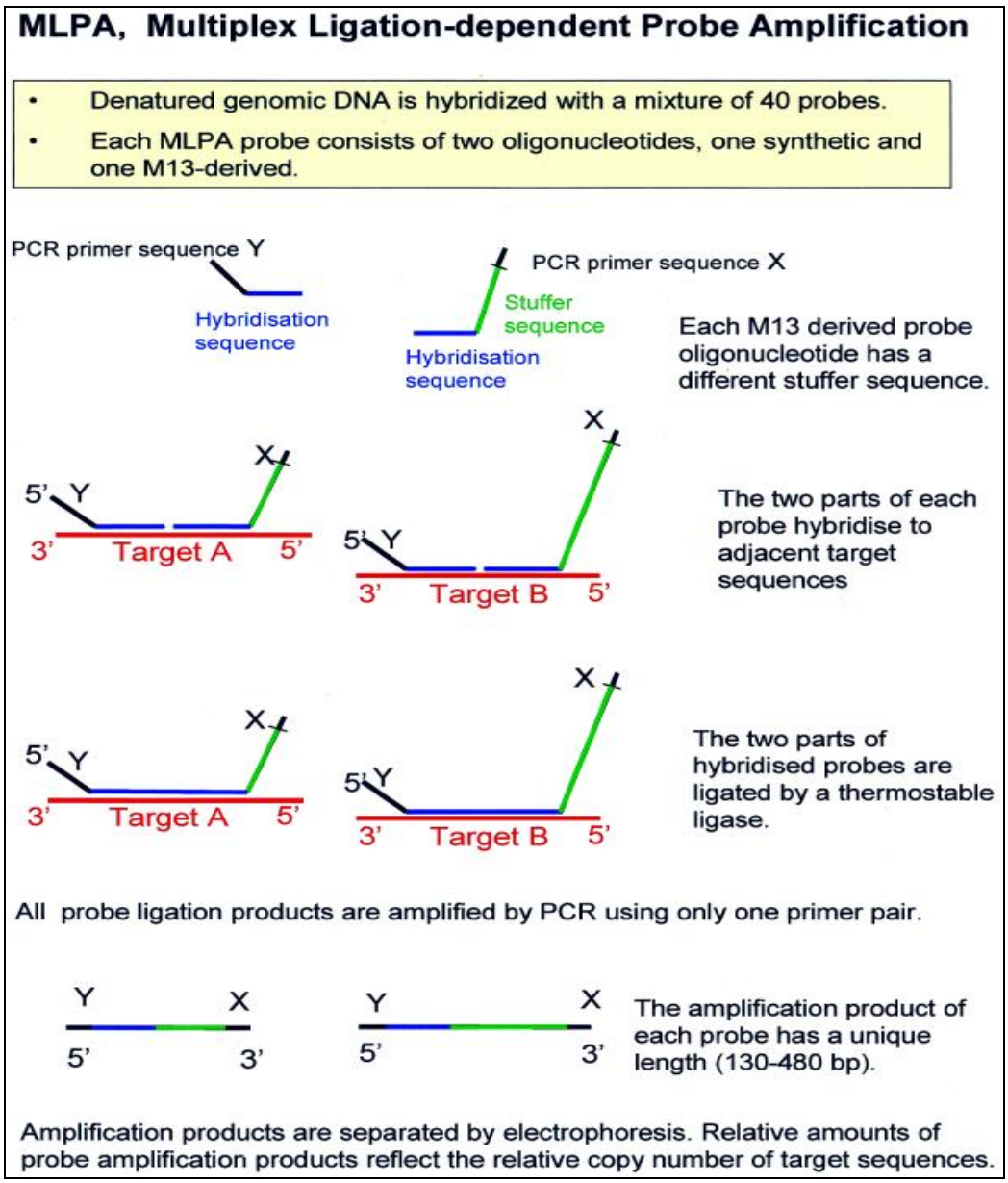


Figure 2.2: Outline of the MLPA method (114).

Statistic Analysis of MLPA products

Fragment analysis was performed on an ABI PRISM 310 sequencing machine with POP4 polymer. The results were analysed using the Genescan and Genotyper software (Applied Biosystems, Rotkreuz, Switzerland) to identify the specific peak representing the respective exons and control loci. Peak areas and heights were then exported to a Microsoft Excel spreadsheet and calculations performed according to the method described by Taylor et al.;

http://leedsdna.info/science/dosage/REX-MLPA/REX-MLPA.htm). Fragments with high standard deviation (≥15%) were omitted from further analysis. An average dosage ratio close to 1 is expected for individuals with two copies, whereas values close to 0.5 indicate loss of one copy. In tumor-derived DNA samples, inevitably containing some degree of contaminating normal tissue, values of ≤0.3 implied loss of both copies.

MLPA results, which indicated a germline or a somatic deletion were independently confirmed in at least one additional, independent experiment as well as independently drawn blood samples if available. All apparently single exon deletions were screened by direct DNA sequencing to exclude sequence variations within the ligation-probe binding site which can mimic single exon deletions (115, 116).

2.9 Long Range PCR

Long-Range PCR on genomic DNA was used to confirm the deletions uncovered by multiplex PCR with the Expand High Fidelity PCR System (Roche, Diagnostics GbmH, Mannheim, Germany). Primers located in intron 6 and intron 9 of *MLH1* (see mutiplex PCR primer appendix 1) were applied. PCR was performed according to the manufacturers recommendations with some small modifications: samples were denaturated at 95°C for 5 min followed by 10 cycles of denaturation at 95°C for 45 sec, annealing/extension at 62°C for 10 min, then 30 cycles of denaturation at 95°C for 45 sec, annealing at 55°C for 50 sec and extension at 68°C for 10 min, and a final extension at 68°C for 15 min. PCR products were separated on a 1% agarose gel and visualized by ethidium bromide staining.

2.10 RT-PCR

cDNA amplification

cDNA was amplified by using Qiagene One Step RT-PCR kit.
About 50ng to 3mg of total RNA were reverse transcribed into complementary DNA with 5 µl 10 x RT-PCR buffer(1x buffer: 10mM/L Tris, 50mM/L KCl, and 0.2mg/ml BAS, pH 8.5), 5 µl of 10mMdNTP mix (5mM each dNTP) and 2ul random primer (10 µM), 10 units RNA inhibitor, 3 µl Reverse Transcriptase 600U/ µl. The procedure was completed by heating the samples for 2 hours at 37°C. Polymerase chain reaction (PCR) amplifications were performed in 50 µl total volumes on an Eppendorf Mastercycle (Eppendorf AG, Basel, Switzerland).

2.11 Protein Truncation Test (PTT)

After cDNA synthesis, PCR amplifications were performed in 50 µl total volumes as following: 100ng cDNA, 0.2U Taq (Gibco/PWO, Gibco USA/Boehringer Mannheim, USA), 2.5 µM each dNTP, 5mM MgCl2, 10x reaction buffer (1x buffer: 10mM/L Tris, 50mM/L KCl, and 0.2mg/ml BAS, pH 8.5) and 0.5 µM of each primer. PTT primer sequences for *MLH1*, *MSH2* were carefully designed (see appendix 1), and used to amplify each gene into two overlapping segments of different size . The cycling conditions were as follows: 94°C-4 min. for 1 cycle, 94°C-45 secs , 55°C /55°C/56°C-1 min. (for *MSH2*, *APC* and *MLH1*, respectively), and 72°C-3 mins for 45 cycles, and 72°C-10 mins for 1 cycle on Eppendorf Mastercycle (Eppendorf AG, basel, Switzerland).

PCR products were first evaluated on a 1% agarose gel. Subsequently, the PTT was run by adding 4 µl PCR product to 6 µl PTT Mix (200 µl TNT T7 coupled Reticulocyte Lysate System, 8 µl RNasin, 16 µl TNT reaction buffer, 16 µl S^{35} Methionine) and heating for 60 minuates at 30°C. The reaction was stopped with 10 µl of 1x sodium dodecyl sulfate (SDS) sample buffer. Subsequently, the products were loaded onto a 12% SDS polyacrylamide gel and run for 110 minuatess at 35 mA. The gels were then fixed (10% glacial acetic acid, 30% methanol) for one hour and dried for 45 minutes at 80°C before exposure on a Biomax film (Kodak, Rochester, NY). Results were analyzed and compared to healthy control in parallel.

CHAPTER 3

Predisposition to Colorectal Cancer

without Pre-existing Polyposis :

Hereditary Nonpolyposis Colorectal Cancer - (HNPCC)

CHAPTER 3.1

3.1 Evaluation of different screening techniques to detect large genomic rearrangements in *MSH2* and *MLH1*

3.1.1 Abstract

Large genomic rearrangements in the mismatch repair genes *MSH2* and *MLH1* are estimated to account for up to 27% of all mutations in patients with hereditary nonpolyposis colorectal cancer (HNPCC). Since large genomic deletions are missed by direct DNA sequencing, two novel methods were recently introduced to overcome this limitation: i) the semi-quantitative multiplex PCR assay (QMPA), ii) the multiplex ligation dependent probe amplification (MLPA) assay. Whereas the first method divides all 35 exons of *MSH2* and *MLH1* into 7 separate groups for PCR multiplexing, MLPA amplifies up to 45 sequences simultaneously.

We tested both methods on 35 Swiss patients clinically suspected of HNPCC, in whom no germline mutation could be identified by direct DNA sequencing. Twenty-one of them presented with microsatellite instability in their cancers, 17 of which showed immunohistochemical loss of either *MSH2* or *MLH1*.

Both QMPA assay and MLPA readily identified the deletions in the control samples. Novel *MLH1* germline deletions spanning exons 7 to 9 as well as a novel *MSH2* deletion encompassing exons 7 and 8 were detected by both methods. The mutations were found to segregate with disease and were further characterized by RT-PCR and long-range PCR. An additional *MSH2* deletion detected by QMPA could not be confirmed by other methods.

In conclusion, we have identified two novel large genomic deletions in *MSH2* and *MLH1*. Four deletion carriers were identififed by QMPA, and three of them could be confirmed by MLPA. Both methods, QMPA and MLPA, appear to be of comparable sensitivity albeit with different specificity.

3.1.2 Introduction

Hereditary nonpolyposis colorectal cancer is an inherited cancer syndrome caused by mutations in mismatch repair genes (1, 2), with the majority of mutations being detected in *MLH1* and *MSH2* (3-5). Somatic inactivation of the remaining copy leads to cancer development. Several kinds of germline mutations have been detected in HNPCC, including truncating, frameshift, splicing or missense mutations (mutation database, www.hgmd.org). The presence of many

deletions and other rearrangements in these genes can not be detected by common mutation screening methods, e.g. heteroduplex analysis, denaturing gradient gel electrophoresis (DGGE), single strand conformational analysis (SSCP), and direct DNA sequencing (6, 7). Alternatively the Protein Truncation Test (PTT) and cDNA amplification are methods able to identify intragenic exon spanning deletions. However, due to alternatively spliced sites and nonsense mediated decay, large deletions could be missed by both techniques (8). In addition, mRNA material is not always available for investigation. Southern blot has been the gold standard to demonstrate genomic deletions in *MSH2* were much more prevalent than previously thought (9), but this method requires large amounts of DNA (10μg). To fill this detection gap, several PCR based gene dosage measurement techniques have been developed recently. With these methods large DNA rearrangements were detected in several cancer predisposition genes at a frequency of around 4-15% (10, 11).

Large germline deletions within the mismatch repair genes *MSH2* and *MLH1* account for a significant proportion (up to 27%) of all deleterious mutations of these genes which are associated with HNPCC syndrome (12, 13). The QMPA method is a simple and reliable means of screening for such alterations (14). With this method, several PCR reactions cover the 35 exons of *MLH1* and *MSH2*. We have modified the primers and PCR conditions for this multiplex PCR protocol, compared to those previously published. The method is based on semi-quantitative PCR of all the exons. Thus, a deletion of one gene can reliably be detected by using exons of the other gene as a reference. Two large deletions in *MSH2* and two large deletions in the *MLH1* gene could be detected by this method. These results were confirmed by other methods including MLPA (Multiplex Ligation-Dependent Probe Amplification). MLPA is a new and high-resolution method for detecting copy number variation in genomic sequences (15). It has been reported to be a robust assay, and offers several advantages over existing techniques by existing reports (16). Many diagnostic genetics laboratories are therefore adopting this as a routine method for gene dosage analysis of genes such as BRCA1, BRCA2, and the mismatch repair genes in preference to other techniques.

There is a clear clinical need for simple and reliable means of screening for these rearrangements. Importantly, for each new molecular re-arrangement thus detected, it is desirable to devise a simple PCR based diagnostic method to search for the mutation in family members at risk . We therefore evaluated the two methods for quantitative analysis, which could complement routine screening for mutations of MMR genes.

Our aim was to compare these two techniques QMPA and MLPA, in terms of their sensitivity and specificity to detect copy number variations.

For this we screened genomic DNA of 35 mutation negative HNPCC patients by QMPA and MLPA. We identified 4 deletions by QMPA, three of which were confirmed by MLPA. In addition,we were able to determine the breakpoint in one of the deletion carriers.

3.1.3 Patients and Methods

Patients

A total of 35 Swiss patients with clinically diagnosed HNPCC were screened for germline mutations in the *MLH1* and *MSH2* genes. No germline mutations were detected in any of these patients by direct DNA sequencing. (Table 3.1. 1)

The diagnostic criteria applied were the Amsterdam Criteria I and the Bethesda Guidelines. Twenty-one patients showed microsatellite instability in their CRCS. Seventeen of which showed also immunohistochemical loss of either *MSH2* or *MLH1*. The mean age of CRC diagnosis was 47 years. The tumor status was MSI-high in 17 patients, and MSI–low in 4 patients. Two positive controls with known genomic deletions status (contributed by the Human Genetics research group from the University of Bonn, Germany) were used to validate the techniques. DNA of 10 healthy individuals was used as negative controls.

Table 3.1.1: Clinical and molecular features of the 35 MLH1/MSH2 mutation-negative HNPCC patients investigated for genomic rearrangements.

Family	Age at diagnose	IHC gene	MSI	Criteria	Sex
1676	43	MLH1	MSI-Low	ACI	f
1806	61	MLH1	MSI-High	ACI	m
1739	69	MLH1	MSI-High	none	f
1754	48	MLH1	MSI-Low	AC I	m
1781	53	MLH1	MSI-High	AC I	f
1806	61	MLH1	MSI-High	AC I	m
1739	76	MLH1	MSI-High	No criteria	f
2055	36	MLH1	MSI-Low	BG	f
2068	70	MLH1	MSI-High	none	f
2064	83	MLH1	MSI-High	none	m
1671	55	MSH2	MSI-High	No criteria	m
1750	51	MSH2	MSI-High	ACI	m
1835	63	MSH2	MSI-High	ACI	f
1804	35	MSH2	MSI-High	ACI	m
1833	36	MSH2	MSI-High	ACI	m
1942	39	MSH2	MSI-High	BG	m
2081	43	MSH2	MSI-High	BG	m
1672	49	nd	MSI-Stable	AC I	m
1645	48	nd	MSI-Stable	No criteria	m
1692	33	nd	MSI-Low	AC I	m
1703	54	nd	MSI-Stable	AC I	f
1716	26	nd	MSI-Stable	BG	m
1722	38	nd	MSI-Stable	BG	f
1776	35	nd	MSI-High	BG	f
1809	35	nd	MSI-Stable	BG	f
1815	39	nd	MSI-Stable	AC I	m
1817	38	nd	MSI-Stable	AC I	m
1826	74	nd	MSI-Stable	AC I	f
1831	79	nd	MSI-Stable	AC I	m
1844	39	nd	MSI-Stable	BG	m
1865	33	nd	MSI-Stable	BG	f
1857	31	nd	MSI-High	AC I	f
1885	44	nd	MSI-High	AC I	f
1895	19	nd	MSI-Stable	BG	f
1903	33	nd	MSI-Stable	BG	f

Abbreviations: CRC denotes colorectal cancer; MSI: microsatellite instability;IHC, immunohistochemically assessed loss of expression of respective protein; ACI, Amsterdam criteria I; BG, Bethesda guidelines. f: female; m: male; nd=not determined

Methods

DNA extraction: See Chapter 2 general methods **2.1**

QMPA: See Chapter 2 general methods **2.7**

Primers: See appendix I

Detection of Genomic Deletions: See Chapter 2 general methods **2.7**

MLPA: See Chapter 2 general methods **2.8**

RT-PCR: See Chapter 2 general methods **2.10**

Long Range PCR: See Chapter 2 general methods **2.9**

3.1.5 Results

With QMPA, the 16 exons of *MSH2* and the 19 exons of *MLH1* were amplified simultaneously in seven multiplex PCR reaction groups followed by fragement analysis on an ABI 310 genetic analyzer. Chromatograms were generated and peak heights and areas evaluated by Genotyper 2.5 software (*Applied Biosystems*) for each multiplex PCR group. Although there is a good correlation between peak height and peak areas, the peak area proved to be the more reliable parameter for calculations.

Validation of the QMPA

Before screening the patients, we tested the reproducibility of the assay with DNA samples from 10 healthy controls and from 2 patients with known exon deletions in *MLH1* and *MSH2* in five consecutive experiments. Finally, the ratios between peaks areas of five different exons compared to that of several reference exons were calculated.
In Table 3.1.2, we present an example of the validation tests: The relative ratios obtained from 10 healthy controls displayed a similar standard deviation (SD) of <15%, and were comparable to SD values obtained from patients DNA samples. The known deletions of exons 1-10 in *MLH1* gene (deletion control 1) and exons 8-15 in *MSH2* (deletion control 2) were reproducibly detected by this assay (Table 3.1. 3).

Quantification detection of QMPA

The detection of genomic deletions was based on the comparison of on the peak areas of different exons simultaneously amplified in a multiplex PCR from healthy controls and patients. Peak

areas were obtained by Genotyper Analysis software 2.5 and exported to an excel sheet. Methods for calculation: within one reaction group which contains six exons from the *MLH1* or *MSH2* gene, one peak from a control exon (The exons were chosen from either *MLH1* or *MSH2*) was taken as a reference (Pr), then the other five peaks were taken as intended peaks (Pi) to obtain the Pi/Pr value.The 10 healthy control samples were calculated in the same manner to get Ci/Cr ratio. The ratio Pi/Pr was further divided by the ratio obtained from the control samples Ci/Cr. If the value (Pi/Pr)/(Ci/Cr) equals 1.0, this means absence of a deletion; a value of around 0.5 was interpreted as a heterozygous deletion of the intended exon. If the value was 2 to 2.5, this indicated a deletion in the reference exon or duplication in the intended exon (Table 3.1. 3).

Base on this calculation rules, patient 1817 and 1835 showed deletion in *MSH2* exon8 (reference). Patient 1806 showed a deletion in *MLH1* exon 9 (see Table 3.1. 3). Standard deviations of the experiments were fewer than 15%.

Table 3.1.2: Relative Ratios of QMPA products (Pi/Pr) from different exons in ten healthy controls

Control No.	*MLH1* ex 17	*MLH1* ex 9	*MLH1* ex 15	*MLH1* ex 2	*MSH2* ex 11	*MSH2* ex 8
1	0.86	0.62	0.82	1.16	0.95	1
2	0.69	0.48	0.71	1.06	1.04	1
3	0.77	0.55	0.8	1.15	1.04	1
4	0.69	0.69	0.73	1.07	1.07	1
5	0.7	0.66	0.73	0.95	1.05	1
6	0.94	0.79	0.88	0.99	0.94	1
7	0.79	0.48	0.78	0.98	0.99	1
8	0.8	0.73	0.82	0.99	1.03	1
9	0.85	0.74	0.82	1.13	0.96	1
10	0.82	1.01	0.94	1.06	0.95	1
mean±	0.79	0.67	0.8	1.05	1.01	1
SD	0.08	0.11	0.07	0.07	0.05	0
SD%	10.37	12.78	8.71	7.06	4.63	0

*Ratio is calculated from Ci/Cr (intended peak area/reference peak area), in this group, *MSH2* is taken as internal reference.

Table 3.1.3: Deletion analysis example for MLH1 exon 9 by QMPA in 23 patients

Patient ID.	MLH1 exon9 (Pi)	MSH2 exon8 (Pr)	(PiPr)/(Ci/Cr)	Intepretation
Deletion control 1	23478	38277	**0.32**	**Deletion in Pi**
Deletion control 2	48354	12257	**2.03**	**Deletion in Pr**
1676	49226	19924	1.03	
1692	18626	9889	0.97	
1739	45882	22247	1.06	
1750	44868	20677	1.12	
1781	40553	17255	1.21	
1804	48213	21552	1.15	
1806	25277	29698	**0.44**	**Deletion in Pi**
1817	45086	9327	**2.49**	**Deletion in Pr**
1833	42263	19551	1.11	
1835	40053	10873	**1.9**	**Deletion in Pr**
1857	44714	18398	1.25	
1885	47481	20908	1.17	
1942	56410	27501	1.06	
1934	46809	19223	1.25	
1971	40907	17468	1.21	
2068	45314	24565	0.95	
1957	47391	20720	1.18	
2079	42356	17590	1.24	
2055	36582	18886	1	
2064	48805	20215	1.24	
2065	45314	24565	0.76	
2081	47391	20720	1.24	
2227	42356	17590	1.05	
Control mean value	50566**(Ci)**	26050**(Cr)**	**Ci/Cr=0.944**	

*Peak area was given by Genotyper 2.5 software; Pi and Pr, intended or reference peak areas in patients (internal reference is MSH exon8 peak areas); Ci and Cr, intended or reference peak areas in control. Deletion control 1 carries *MLH1* exon1 to 10 deletions. Deletion Control 2 carries a *MSH2* exon8 to 15 deletions.

Results of deletion screening in 35 patients

As an example, Figure 3.1. 1 shows a deletion detected by QMPA, Table 3.1. 4 shows calculation of relative copy number. After QMPA screening, four large deletions were identified: Two patients (patient 1817 and patients 1835) having *MSH2* exons 7 and 8 deleted. *MLH1* exons 7 and 9 were deleted in patient 1806 , patient 1676 carries an *MLH1* exon 13 deletion.

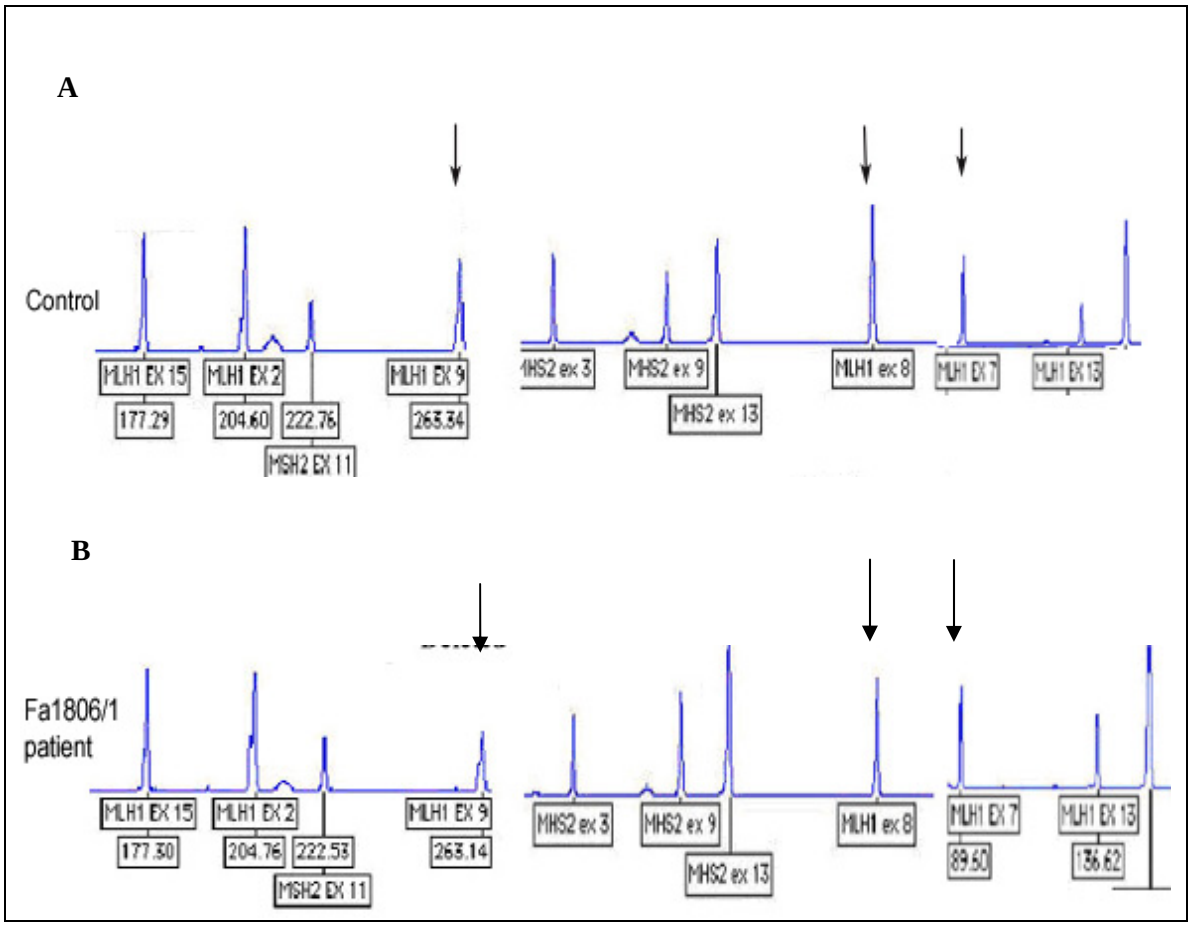


Figure 3.1. 1: Peak patterns of QMPA PCR products (**A panel**: control samples, **B panel**: patient 1806). These patterns were generated using primers of three different multiplex groups. Comparing control vs. patient, peak area ratios for exons 7, 8 and 9 of MLH1 were calculated. (Table 3.1.3)

Table 3.1.4: Relative Values of QMPA products (PiPr/CiCr) from *MLH1* exons 1-11
Deletion control 1 carries a deletion of *MLH1*exons 1 to 10
1806/1 and 1806/2 are two affected family members from family 1806
Controls 1 to 10 are healthy control subjects. Values below 0.5 indicated exonic deletions.

Patient ID	exon1	exon2	exon3	exon4	exon5	exon6	exon7	exon8	exon9	exon10	exon11
Deletion Control 1	**0.53**	**0.48**	**0.56**	**0.48**	**0.49**	**0.41**	**0.39**	**0.38**	**0.46**	**0.39**	0.78
1806/01	1.06	0.95	1.15	0.97	1.09	0.99	**0.51**	**0.47**	**0.49**	1.21	1.06
1806/02	1	0.96	1.14	0.96	1.05	1.07	**0.5**	**0.46**	**0.47**	1.11	0.88
Control 1	0.99	1.01	0.99	0.97	0.96	0.97	0.99	0.96	0.92	0.91	0.97
Control 2	1.44	1.47	1.13	1	1.21	1.33	1.12	1.23	1.27	1.49	1.46
Control 3	0.92	0.93	1.05	0.92	0.98	0.98	1.03	0.97	0.9	0.95	1.04
Control 4	1.01	0.97	1.09	1.02	0.94	0.97	0.95	0.9	0.88	0.83	0.86
Control 5	1.1	1.03	1	1.04	0.86	0.93	0.92	0.8	0.76	0.77	0.93
Control 6	1.03	1.03	0.94	1.1	0.87	0.93	0.99	0.88	0.89	0.71	0.91
Control 7	1.01	1.02	1.04	0.99	0.98	1.02	1.01	1	0.96	0.94	1.01
Control 8	0.97	0.96	1.02	1.01	0.96	0.98	0.99	0.97	0.95	0.91	0.98
Control 9	0.92	0.93	0.89	1.01	1.09	0.96	1	1.09	1.21	1.17	1.03
Control10	0.98	0.96	0.96	1.02	0.95	0.95	0.95	0.92	0.96	0.9	0.98

Confirmation of the deletions

cDNA confirmation

The deletion of exons 7 to 9 in the *MLH1* gene was successfully confirmed by cDNA amplification (Figure 3.1. 2). The sequencing result for patient 1806 cDNA showed the precise breakpoint location at *MLH1* codon 454 (exon 5) and codon 884 (exon11). Due to alternative spliced sites in *MLH1* exon 6, exon 9 and exon 10 (18), exons 7 and 9 to 10 were spliced out in healthy control; in patient 1806, the break point was found to be located at the end of exon 5 (Figure 3.1. 2) and the start of exon 11. The breakpoint identification of *MSH2* exon7_8 deletion patients could not be assessed, because no patients mRNA was available for analysis.

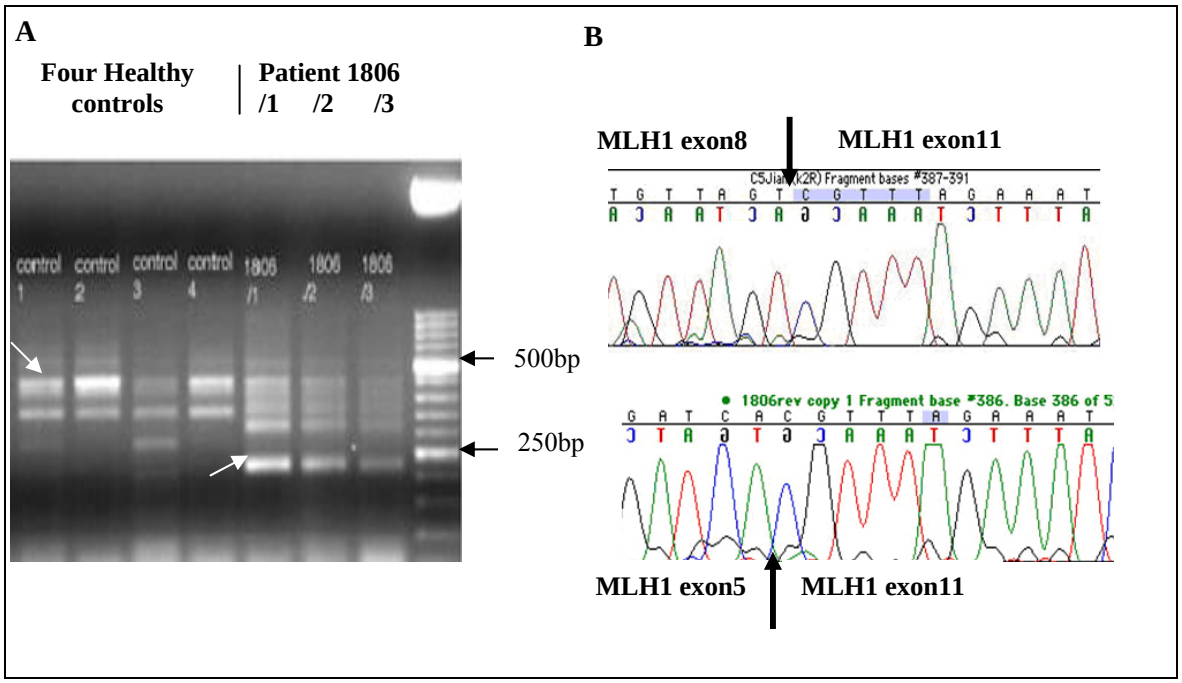


Figure 3.1.2: Determination of deletion breakpoints in patient 1806
A: cDNA amplification (*MLH1* exon 5 to 11) of healthy controls and affected members from family 1806. Because of alternative splicing of exons 6 and exons 9-10, the amplification products are smaller than expected (660bp). Note the 250 bp fragment only present in affected members from family 1806. **B**: Compared to the healthy negative control sample displaying the direct transition from exon 8 to 11 (**exons - 6, 9 and 10 are spliced out**), patient 1806-1, having deleted exons 7 to 9, shows joining of exon 5 to 11.

Long Range PCR confirmation

Long Range PCR was also applied to confirm the deletion of *MLH1* exons 7 to 9 from genomic DNA. The patients who carry the novel *MLH1* deeltion (exon7_9 del) showed a 10kb PCR product (Figure 3.1. 3). The products were amplified by primers located at *MLH1* exon 6, and downstream primer located at *MLH1* exon 10. The presence of a 7.8 kb product suggests the products with presence of the exon7 to 9 deletions.

Long range PCR could not performed to confirm the deltion of MSH2 exon7_8 deletion because *MSH2* exons 7 and 8 host very large intronic sequence (13kb and 15kb).

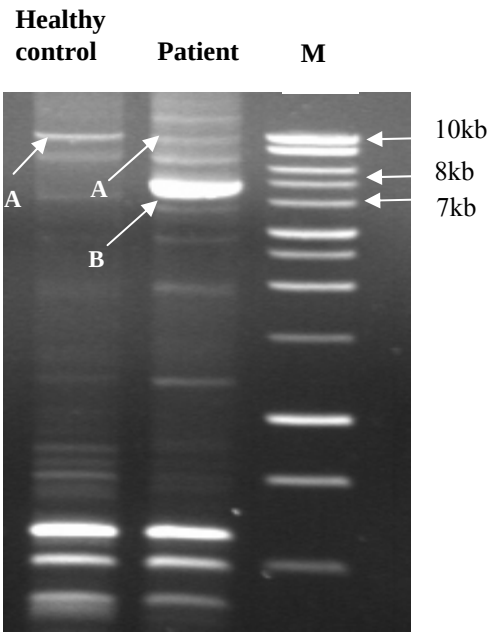


Figure 3.1.3: Confimation of deletion by Long Range PCR from genomic DNA
A: arrow indicate PCR products without deletion.
B: arrow indicate PCR products generated by exon deletion of MLH1 exon7 to exon9

MLPA

We also applied the MLPA method to re-screen all patients. The MLPA assay proved robust and more reliable regarding peak area distribution (Figure 3.1. 4) compared to QMPA and lower standard deviations (Table 3.1. 5). Changing the size standard profile in GENESCAN software allowed us to align the amplification patterns of controls and patients for direct comparision (Figure 3.1. 4). Genotyper 2.5 software was applied to obtain the peak area values of the chromatogram. The MLPA method was first evaluated by screening ten healthy controls with standard deviations (SD) below 0.15.

The result of the MLPA was confirmed in three of the deletion carriers. Patient 1817 whose *MSH2* exon 7 to 8 deletion was detected by QMPA showed no deletion in *MSH2* by MLPA. (Table 3.1. 5)

Table 3.1.5: Gene dosage analysis by MLPA: Typical dosage result showing a deletion of *MLH1* exon7 to exon9 in patient 1806.But in patient 1817, there is no deletion found in *MSH2* exon7 to exon 8.

Patient 1817			**Patient 1806**		
	Ave			**Ave**	
Category	**DR**	**SD**	**Category**	**DR**	**SD**
con 10p11	1.07	0.11	con 10p11	0.89	0.13
con 10p14	0.91	0.1	con 10p14	1.27	0.11
con 17q21	1.09	0.05	con 17q21	1.02	0.14
con11p12	0.81	0.13	con11p12	0.99	0.14
con11p13	0.98	0.08	con11p13	0.88	0.13
con.5q31	1.32	0.09	con.5q31	0.85	0.12
con4q25	1.03	0.12	con4q25	1.12	0.09
MSH2 ex5	1.17	0.11	*MLH1*.EX5	0.98	0.14
MSH2 ex6	1.13	0.14	*MLH1*.EX6	0.99	0.14
***MSH2* ex7**	**1.19**	**0.06**	***MLH1*.EX7**	**0.55**	**0.08**
***MSH2* ex8**	**1.01**	**0.11**	***MLH1*.EX8**	**0.51**	**0.07**
MSH2 ex09	1.45	0.09	***MLH1*.EX9**	**0.46**	**0.07**
MSH2 ex10	1	0.1	*MLH1*.EX10	1.15	0.12
MSH2 ex11	0.97	0.15	*MLH1*.EX11	0.91	0.13

*The average Dosage Ratio (DR) is the mean of the dosage ratio. This dosage ratio is calculated like: Peak area (each exon fragment)/Peak area (control fragment). Values close to 1 are expected for individuals with two copies of the test fragment and close to 0.5 for individuals with loss of one copy.

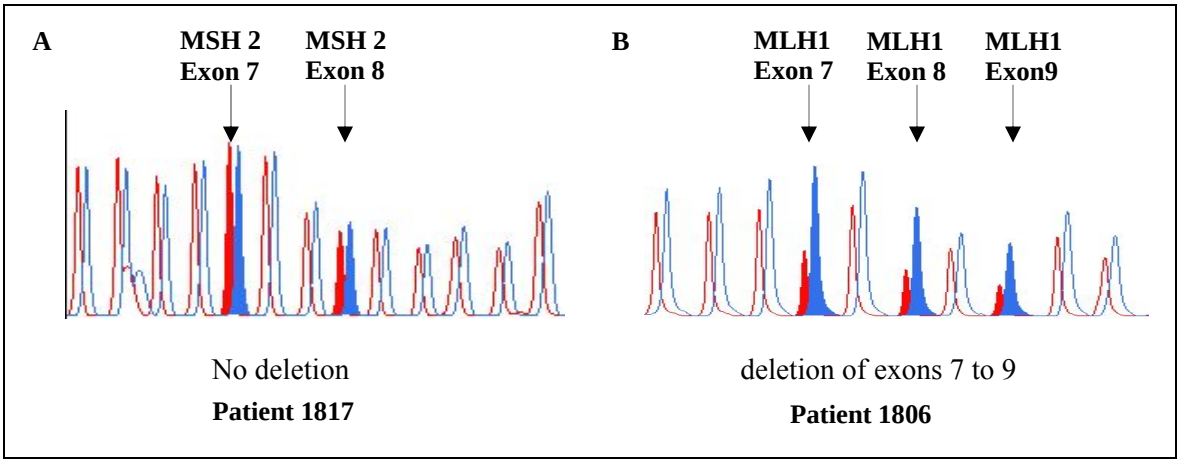


Figure 3.1.4: Peak profile of MLPA products, Red peaks indicated patient, Blue peaks indicated healthy control, **A: Patient 1817 and healthy control comparison pattern, B: Patient 1806 and healthy control comparison pattern**. Peak area calculation is showed in Table 3.1. 5.

3.1.5 Discussion

With mutation detection methods such as heteroduplex analysis, DGGE, SSCP, or dHPLC (6), it is not possible to uncover large exon deletions at the level of genomic DNA. Southern blot is the gold standard to detect large genomic deletions in DNA mismatch repair genes (9). But as a routine application, this method is time consuming and also requires large amounts of DNA. All of these disadvantages limit its value in a routine diagnostic setting. Other methods like RNA based sequencing and PTT are able to detect intragenic deletions (8). Large deletions however might extend over the location of the primers used in RT- PCR and will therefore fail to yield a PCR product. So deletion of an entire gene can hardly be detected by this method. Another disadvantage of RNA based methods is that large deletions can also be missed due to alternative splicing (8). Recently two different multiplex PCR assays have been introduced, they are QMPA and MLPA (19). Here we have compared sensitivity and specificity of these methods in detecting large genomic deletions in *MLH1* and *MSH2* gene.

The detection of genomic deletions by QMPA, requires three principles to be followed in order to get reliable results: 1) PCR reactions must be performed within the period of the exponential amplification. The quantification correlates with the quantity of DNA template copies. 2) The primers for different exons of *MLH1* and *MSH2* within the same group should be reliably amplified. 3) Each multiplex PCR reaction has seven different primers with different melting temperatures (Tm). In order to get efficient annealing temperatures, each primer concentration has to be well optimized to balance the individual primer specific PCR efficiency. As a rule of thumb: The higher the Tm values of primers, the lower the concentration of the respective primer pair.

It is important however, to take the following consideration into account. The group of primers working in the same QMPA has to be designed carefully to avoid interference between primers. The QMPA assay presented here is characterized by stability and sensitivity. It can be applied for simultaneous detection of genomic deletions in both *MSH2* and *MLH1* genes. In the case of rearrangements involving the entire *MSH2* or *MLH1* gene, an additional internal control of another gene has to be used in order to differentiate between deletion or duplication of one of these two genes.

Because primer concentrations have to be modified within the same group to balance QMPA efficiency, other conventional methods after QMPA screening must be applied to confirm the results in order to avoid the false positive and negative results.

MLPA has gained growing reputation in genetic diagnostic laboratories due to its simplicity, relative low cost, low DNA consumtion, capacity for reasonably high throughput and robustness (20). With the MLPA, we were able to amplify products covering all 35 exons of *MLH1* and *MSH2* together plus additional seven chromosomal controls in one single PCR. Chromosomal controls have been well selected as internal controls to evaluate every amplification. They located at different chromosomes in addition to the chromosomes on which *MLH1/MSH2* are located.

The MLPA reaction relies on the probe to hybridize to the exact and unique location of the respective exonic sequence. The hybridization sites of *MLH1* and *MSH2* MLPA kit are carefully picked to avoid a possible polymorphisms iin these two genes. Because they all share the same universal PCR primers, the primer concentrations do not have to be modified in the MLPA assay.

Using the MLPA assay, we were able to confirm the two deletions in the *MLH1* and the one in *MSH2*. One deletion was false-positive by QMPA, but due to unavailability of mRNA and technical reasons (>30kb intronic sequence), we could not do further investigation for this.

In conclusion, both methods, QMPA and MLPA can readily identify the deletions in the deletion control samples, albeit with variable specificity. The QMPA technique, however is difficult to set and standardize the PCR conditions in order to obtain reproducible results. The MLPA method, in contrast, proved easy to use (one step amplification) and gave fast and highly reproducible results.

3.1.6 References

1. Lynch, DA, Lobo, AJ, Sobala, GM, Dixon, MF, and Axon, AT Failure of colonoscopic surveillance in ulcerative colitis. Gut, 1993; *34*(8)*:* 1075-1080.
2. de la Chapelle, A and Peltomaki, P Genetics of hereditary colon cancer. Annu Rev Genet, 1995; *29*(329-348.
3. Wijnen, J, Khan, PM, Vasen, H*, et al.* Majority of h*MLH1* mutations responsible for hereditary nonpolyposis colorectal cancer cluster at the exonic region 15-16. Am J Hum Genet, 1996; *58*(2)*:* 300-307.
4. Miyaki, M, Konishi, M, Tanaka, K*, et al.* Germline mutation of MSH6 as the cause of hereditary nonpolyposis colorectal cancer. Nat Genet, 1997; *17*(3)*:* 271-272.
5. Lynch, HT and Lynch, JF The Lynch syndromes. Curr Opin Oncol, 1993; *5*(4)*:* 687-696.
6. Wijnen, J, Vasen, H, Khan, PM*, et al.* Seven new mutations in h*MSH2*, an HNPCC gene, identified by denaturing gradient-gel electrophoresis. Am J Hum Genet, 1995; *56*(5)*:* 1060-1066.
7. Lynch, HT and de la Chapelle, A Genetic susceptibility to non-polyposis colorectal cancer. J Med Genet, 1999; *36*(11)*:* 801-818.
8. Jakubowska, A, Gorski, B, Kurzawski, G*, et al.* Optimization of experimental conditions for RNA-based sequencing of *MLH1* and *MSH2* genes. Hum Mutat, 2001; *17*(1)*:* 52-60.
9. Wijnen, JT, Vasen, HF, Khan, PM*, et al.* Clinical findings with implications for genetic testing in families with clustering of colorectal cancer. N Engl J Med, 1998; *339*(8)*:* 511-518.
10. Wagner, HJ, Scott, RS, Buchwald, D, and Sixbey, JW Peripheral blood lymphocytes express recombination-activating genes 1 and 2 during Epstein-Barr virus-induced infectious mononucleosis. J Infect Dis, 2004; *190*(5)*:* 979-984.
11. Peelen, T, van Vliet, M, Bosch, A*, et al.* Screening for BRCA2 mutations in 81 Dutch breast-ovarian cancer families. Br J Cancer, 2000; *82*(1)*:* 151-156.
12. Wang, Y, Friedl, W, Sengteller, M*, et al.* A modified multiplex PCR assay for detection of large deletions in *MSH2* and *MLH1*. Hum Mutat, 2002; *19*(3)*:* 279-286.
13. Casey, G, Lindor, NM, Papadopoulos, N*, et al.* Conversion analysis for mutation detection in *MLH1* and *MSH2* in patients with colorectal cancer. Jama, 2005; *293*(7)*:* 799-809.
14. Charbonnier, F, Raux, G, Wang, Q*, et al.* Detection of exon deletions and duplications of the mismatch repair genes in hereditary nonpolyposis colorectal cancer families using

multiplex polymerase chain reaction of short fluorescent fragments. Cancer Res, 2000; *60*(11)*:* 2760-2763.

15. Schouten, JP, McElgunn, CJ, Waaijer, R, Zwijnenburg, D, Diepvens, F, and Pals, G Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. Nucleic Acids Res, 2002; *30*(12)*:* e57.
16. Taylor, CF, Charlton, RS, Burn, J, Sheridan, E, and Taylor, GR Genomic deletions in *MSH2* or *MLH1* are a frequent cause of hereditary non-polyposis colorectal cancer: identification of novel and recurrent deletions by MLPA. Hum Mutat, 2003; *22*(6)*:* 428-433.
17. Gille, JJ, Hogervorst, FB, Pals, G*, et al.* Genomic deletions of *MSH2* and *MLH1* in colorectal cancer families detected by a novel mutation detection approach. Br J Cancer, 2002; *87*(8)*:* 892-897.
18. Charbonnier, F, Martin, C, Scotte, M, Sibert, L, Moreau, V, and Frebourg, T Alternative splicing of *MLH1* messenger RNA in human normal cells. Cancer Res, 1995; *55*(9)*:* 1839-1841.
19. Charbonnier, V, Morrison, BM, Jr., Paye, M, and Maibach, HI An open assay model to induce subclincal non-erythematous irritation. Contact Dermatitis, 2000; *42*(4)*:* 207-211.
20. Bunyan, DJ, Eccles, DM, Sillibourne, J*, et al.* Dosage analysis of cancer predisposition genes by multiplex ligation-dependent probe amplification. Br J Cancer, 2004; *91*(6)*:* 1155-1159.

CHAPTER 3.2

3.2 Gene Conversion Is a Frequent Mechanism of Inactivation of the Wild-Type Allele in Cancers from MLH1/MSH2 Deletion Carriers

Zhang J, Lindroos A, Ollila S, Russell A, Marra G, Mueller H, Peltomaki P, Plasilova M, Heinimann K. (Cancer Res 2006; (66)2: 659-64)

3.2.1 Abstract

Hereditary nonpolyposis colorectal cancer (HNPCC) is an autosomal dominantly inherited cancer predisposition syndrome caused by germline mutations in DNA mismatchrepair genes, predominantly *MLH1* and *MSH2*, with large genomic rearrangements accounting for 5 to 20% of all mutations. Although crucial to the understanding of cancer initiation, little is known about the second, somatic hit in HNPCC tumorigenesis, commonly referred to as loss of heterozygosity (LOH). Here we applied a recently developed method, multiplex ligation dependent probe amplification (MLPA), to study *MLH1/MSH2* copy number changes in 16 unrelated Swiss HNPCC patients, whose cancers displayed microsatellite instability and loss of MLH1 or MSH2 expression, but in whom no germline mutation could be detected by conventional screening. The aims of the study were i) to determine the proportion of large genomic rearrangements among Swiss *MLH1*/*MSH2* mutation carriers and ii) to investigate the frequency and nature of LOH as second, somatic event in tumors from *MLH1*/*MSH2* germline deletion carriers. Large genomic deletions were found to account for 4.3% and 10.7% of *MLH1* and *MSH2* mutations, respectively. MLPA analysis of 18 cancer specimens from two independent sets of Swiss and Finnish *MLH1*/*MSH2* deletion carriers revealed that somatic mutations identical to the ones in the germline occur frequently in colorectal cancers (6/11; 55%) and are also present in extracolonic HNPCC-associated tumors. Chromosome-specific marker analysis implies that loss of the wild-type allele predominantly occurs through locus-restricted recombinational events, i.e. gene conversion, rather than mitotic recombination or deletion of the respective gene locus.

3.2.2 Introduction

Hereditary nonpolyposis colorectal cancer (HNPCC) is an autosomal dominantly inherited cancer predisposition syndrome characterized by the occurrence of early onset colorectal carcinoma (CRC) as well as a defined spectrum of extracolonic tumors, such as cancers of the endometrium and renal pelvis (1). HNPCC is caused by germline mutations in DNA mismatch repair (MMR) genes, predominantly *MLH1* and *MSH2*, with 5 to 20% of mutations being large genomic

rearrangements missed by conventional mutation screening techniques (2-4). Recently, the multiplex ligation dependent probe amplification (MLPA) method has been introduced to assess DNA copy number changes semiquantitatively (5). This method requires considerably less DNA (50 to 200 ng) than conventional Southern blotting (5 to 10 μg) and the short recognition sequence of the probes moreover allows to determine copy number changes in partially degraded DNA, such as DNA from formalin-fixed cancer tissue. According to Knudson's "two-hit" hypothesis a second, somatic mutation, inactivating the wild-type allele and commonly referred to as loss of heterozygosity (LOH), is required for tumorigenesis to start (6). Subsequent mismatch repair deficiency leads to accumulation of replication errors, mainly at short repetitive DNA sequences, in the tumor cell and gives rise to the molecular hallmark of HNPCC, microsatellite instability (MSI). Although crucial to the understanding of cancer initiation, only scarce data are available on the nature of the second hit in HNPCC tumors (7). Here we applied the MLPA technique on 16 unrelated Swiss HNPCC patients, whose cancers displayed MSI and loss of MLH1 or MSH2 expression, but in whom no germline mutation could be detected by conventional DNA sequence. The study aimed i) to determine the proportion of large genomicrearrangements in our set of Swiss *MLH1*/*MSH2* mutation carriers and ii) to investigate the frequency and nature of LOH as second, somatic event in HNPCC tumorigenesis in cancers from *MLH1*/*MSH2* germline deletion carriers.

3.2.3 Patients and Methods

Sixteen unrelated Swiss patients referred to the division of Medical Genetics because of clinically suspected HNPCC syndrome were included in this study. The cancers from these patients had been found to display microsatellite instability and loss of MLH1 or MSH2 expression (Table 3.2. 1). Since no pathogenic germline mutation could be detected in *MLH1* or *MSH2* by conventional DNA sequencing, all patients were investigated for the presence of large genomic rearrangements in their germline. Subsequently, the presence/absence of the identified *MLH1*/*MSH2* germline deletion was assessed in 11 cancers as well as in 7 cancers from known Finnish germline deletion carriers (8, 9). Written informed consent was obtained from all patients included in the study.

Table 3.2.1: Clinical and molecular features of the 16 MLH1/MSH2 mutation-negative HNPC patients investigated for genomic rearrangements. CRC denotes colorectal cancer; MSI: microsatellite instability; IHC: immunohistochemical protein loss; f: female; m: male; ACI: Amsterdam Criteria I; BG: Bethesda guidelines.

Patient ID	Sex	Age at diagnosis	Criteria	CRC site	MSI	IHC	Exon(s) deleted
1676/1	f	43	ACI	Sigmoid	MSI-Low	MLH1	exon 13
1806/1	m	61	ACI	Transverse	MSI-High	MLH1	exon 7 to 9
2055/1	f	36	BG	Sigmoid	MSI-Low	MLH1	no deletion detected
2079/1	m	65	BG	Caecum	MSI-Low	MLH1	no deletion detected
2065/1	m	68	none	Transverse	MSI-High	MLH1	no deletion detected
1739/1	f	69	none	Transverse	MSI-High	MLH1	no deletion detected
2068/1	f	70	none	Sigmoid	MSI-High	MLH1	no deletion detected
2064/1	m	83	none	Ascendens	MSI-High	MLH1	no deletion detected
2264/1	f	39	ACI	Ascendens	MSI-High	MSH2	exon 8 to 16
2227/1	m	59	BG	Caecum	MSI-High	MSH2	exon 8 to 11
1835/1	f	63	ACI	Descendens	MSI-High	MSH2	exon 7 to 8
1804/1	m	35	ACI	Transverse	MSI-High	MSH2	no deletion detected
1833/1	m	36	ACI	Caecum	MSI-High	MSH2	no deletion detected
1942/1	m	39	BG	Sigmoid	MSI-High	MSH2	no deletion detected
2081/1	m	43	BG	Ascendens	MSI-High	MSH2	no deletion detected
1750/1	m	51	ACI	Ascendens	MSI-High	MSH2	no deletion detected

Abbreviations: CRC denotes colorectal cancer; MSI: microsatellite instability;IHC, immunohistochemically assessed loss of expression of respective protein; ACI, Amsterdam criteria I; BG, Bethesda guidelines. f: female; m: male;

DNA extraction from peripheral blood and tumor tissue.

DNA from peripheral blood was isolated by using a salting-out procedure described by Miller et al. (10). Prior to DNA extraction from tumor tissue, histopathologic classification of H&E stained, formalin-fixed tissue blocks was carried out and a, representative portion of the tumor with an average tumor content of ≥70% was scraped off. DNA extraction was done according to the QIAamp tissue kit protocol (Qiagen, Basel, Switzerland).

Analysis of MSI and MMR protein expression.

Based on the recommendations of the National Cancer Institute workshop on MSI, a panel of five microsatellite loci (BAT25, BAT26, D5S346, D17S250, and D2S123) was used to assess MSI (11). The presence or absence of four MMR proteins (MLH1, MSH2, MSH6, and PMS2) in the tumor was examined by standard immunohistochemical techniques (12).

Analysis of LOH.

LOH, also referred to as "allelic loss", was investigated using the following flanking polymorphic microsatellite markers: D3S1304, D3S1263, D3S2338, D3S1266, D3S1277, D3S1300, D3S1566, and D3S1278 for the *MLH1* locus; and D2S168, D2S165, D2S367, D2S391, D2S337, D2S2110, D2S286, D2S2333, and D2S347 for the *MSH2* locus. LOH was scored at any informative marker if the area under one allelic peak in the tumor was reduced by >50% relative to the other allele, after correcting for the relative peak areas in leukocyte-derived constitutional DNA (13). PCR conditions for LOH analysis are available from the authors upon request. PCR products were analyzed on an ABI PRISM 310 Genetic Analyzer (Applied Biosystems, Rotkreuz, Switzerland).

MLPA

For the detection of aberrant copy numbers in the *MLH1* and *MSH2* genes in constitutional, leukocyte-derived, and tumor DNA, the SALSA P003 *MLH1/MSH2* test MLPA kit (MRC Holland, Amsterdam, the Netherlands) was used (5). The kit contains probes for the 16 exons of *MSH2* and the 19 exons of MLH1 as well as seven probes located on different chromosomes as controls. DNA samples from two known germ line deletion carriers (MLH1 exon1_10del and MSH2 exon8_15del) as well as from 10 healthy probands were used to confirm the sensitivity and specificity of the method. Each mutation was confirmed on a second, independently drawn, blood sample from the respective patients. All reactions were carried out according to the manufacturer's protocol. Fragment analysis was done on an ABI 310 capillary sequencer and results were analyzed using the Genescan and Genotyper software (Applied Biosystems) to identify the specific amplicons representing the respective exons and control loci. Peak areas and heights were then exported to a Microsoft Excel spreadsheet and calculations were done according to the method described by Taylor et al. (14). Fragments with high SD (≥20%) were omitted from further analysis. An average dosage quotient close to 1 is expected for individuals with two copies, whereas values close to 0.5 indicate loss of one copy. In tumor-derived DNA samples, inevitably containing some degree of contaminating normal tissue, values ≤0.3 implied loss of both copies. MLPA results indicative of a germ line or a somatic deletion were independently confirmed in at least one additional, independent experiment as well as independently drawn blood samples, and, where available, cDNA was used to assess the individual break points. All apparently single exon deletions were screened by direct DNA sequencing to exclude sequence variations within the ligation-probe binding site which can mimic single exon deletions (15, 16).

Statistical analysis.

Statistical comparison of patients' features, encompassing phenotypic characteristics (gender, age at diagnosis, etc.), and mutational status, was done using the χ^2 and Fisher's exact test for categorical variables, or Student's *t* test for continuous variables, with all of the probabilities reported as two-tailed *P* values, considering $P < 0.05$ to be statistically significant.

3.3.4 Results and Discussion

In this study 16 unrelated Swiss HNPCC patients without identified pathogenic germline mutation in *MLH1* or *MSH2* were screened for the presence of large genomic rearrangements using the MLPA assay (Table 3.2. 1 and Figure 3.2. 1, Figure 3.2. 2 ,Figure 3.2. 3). Three (38%) out of 8 index patients whose colorectal cancers (CRC) showed high microsatellite instability (MSI) and loss of the MSH2 protein were found to harbour 3 different genomic deletions in *MSH2*. Among the 8 index patients whose tumours had lost MLH1 expression, 2 (25%) were found to carry large genomic rearrangements in the *MLH1* gene. In view of the late age at diagnosis (>68 years) as well as an inconspicuous family history in 4 of the *MLH1* mutation negative patients, it is likely that these cancers are actually sporadic in origin, due to hypermethylation of the *MLH1* promoter rather than due to an inherited germline mutation (17). With mRNA available for further study, the consequence of the *MLH1* deletion encompassing exons 7 to 9 (family 1806) could be assessed. By cDNA sequencing we found this genomic deletion to result in direct joining of exon 5 to exon 11 (c.454_884del) leading to a frameshift and a first premature stop codon 11 amino acids downstream. Taken together, large genomic deletions account for 10.7% (3/28) of *MSH2* and for 4.3% (2/47) of *MLH1* mutations in our set of mutation-positive Swiss HNPCC families (n=75). The remaining mutation-negative patients are currently under intensive study for mutations in regulatory regions and/or epimutations in these MMR genes.

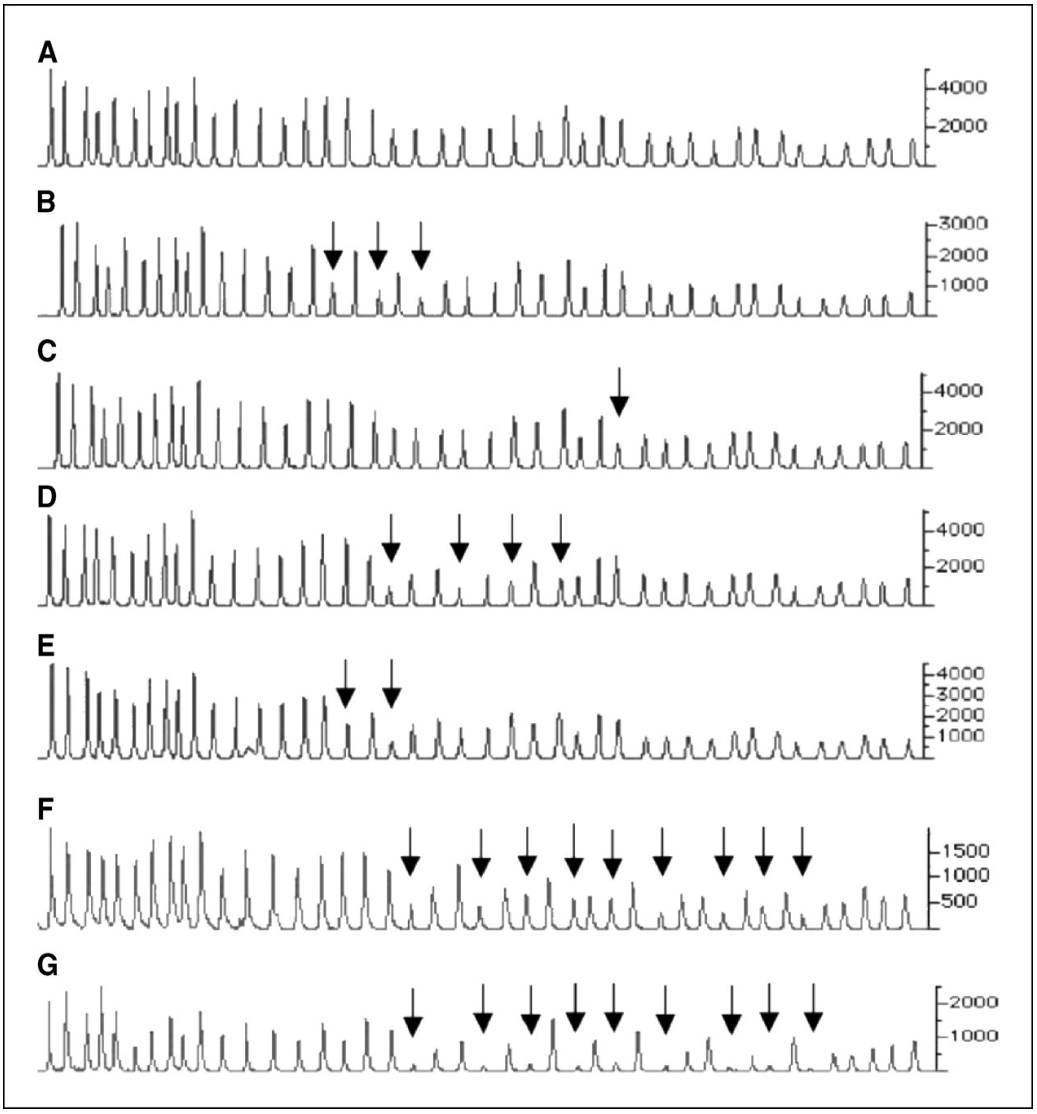


Figure 3.2.1: MLPA electropherograms of *MLH1/MSH2* germline deletion carriers. **A**: healthy control proband; **B**: *MLH1* exon7_9del (patient 1806); **C**: *MLH1* exon13del (patient 1676); **D**: *MSH2* exon8_11del (2227); **E**: *MSH2* exon7_8del (patient 1835); **F:** *MSH2* exon8_16del (patient 2264); **G**: colorectal cancer from patient 2264 indicating a homozygous *MSH2* exon8_16del. Arrows denote deleted exons.

In order to further substantiate the high frequency of cancers harbouring a somatic deletion identical to the one in the germline, 7 additional tumors (4 CRCs, 2 endometrial and1 stomach

cancer) from Finnish HNPCC patients carrying an *MSH2* or *MLH1* germline deletion were investigated (Table 3.2. 2b). Two out of 4 CRCs were found to carry identical, biallelic deletions which were absent in the remaining cancers. Thus, although based on a arguably small number of cases, our findings from two independent sets of patients indicate that the occurrence of somatic deletions identical to the ones in the germline are a frequent event in HNPCC related colorectal tumorigenesis (6 out of 11 CRCs; 55%). Remarkably, 9 none of the Swiss or Finnish tumor specimens showed evidence for large somatic deletions encompassing the entire, respective gene locus.

Overall, genomic deletion carriers were statistically significantly later diagnosed of CRC (median, 59 years; IQR, 19.5; $n = 5$) compared with index patients carrying "conventional" *MLH1/MSH2* mutations (median, 40 years; IQR, 14.0; $n = 70$; $P < 0.006$). Four out of six (67%) deletion carriers (Table 3.2. 2) had, in addition to their CRC, developed extracolonic cancers: patient 1806/3, an ovarian cancer at age 52; patient 1835/1, an endometrial cancer at age 61; patient 2227/1, a urothelial carcinoma at age 60 and an astrocytoma WHO grade 3 at age 67; and patient 2264/1, a duodenal and an endometrial cancer at age 41 and 48, respectively. Compared with other index mutation carriers in our HNPCC database, the frequency of extracolonic HNPCC-associated cancers was statistically significantly increased (67% versus 20%; $P < 0.03$).

With regard to other phenotypic properties (e.g., site of CRC or tumor stage) no further significant differences were observed between the two groups. With one exception, i.e., family 2227, in which HNPCC-related cancers were confined to one generation only, all genomic deletion carriers met the Amsterdam criteria I (Table 3.2. 1). Because the phenotypic observations might have an effect on counseling and management of *MLH1/MSH2* germ line deletion carriers, it is important that these findings are confirmed on larger patient sets.

Table 3.2.2. Histologic, anatomic, and molecular features of 18 cancers from Swiss and Finnish HNPCC patients carrying large genomic deletions

(A) Six Swiss HNPCC patients carrying large genomic deletions

Patient ID	Cancer type	Cancer site	MSI	Gene	Germ line deletion	Somatic deletion
1806/1	adenocarcinoma	transverse colon	high	*MLH1*	exons 7-9	exons 7-9
	adenocarcinoma	transverse colon	high	*MLH1*	exons 7-9	exons 7-9
1806/3	adenocarcinoma	ascending colon	high	*MLH1*	exons 7-9	exons 7-9
	adenocarcinoma	right ovary	high	*MLH1*	exons 7-9	absent
1676/1	adenocarcinoma	sigmoid colon	low	*MLH1*	exon 13	absent
1835/1	adenocarcinoma	descending colon	high	*MSH2*	exons 7-8	absent
	adenocarcinoma	endometrium	high	*MSH2*	exons 7-8	absent
2227/1	adenocarcinoma	cecum	high	*MSH2*	exons 8-11	exon 11
	urothelial carcinoma	left kidney	low	*MSH2*	exons 8-11	exon 11
	astrocytoma WHO grade 3	frontal brain	high	*MSH2*	exons 8-11	exons 8-11
2264/1	adenocarcinoma*	ascending colon	high	*MSH2*	exons 8-16	exons 8-16

(B) Seven Finnish HNPCC patients carrying large genomic deletions

Patient ID	Cancer type	Cancer site	MSI	Gene	Germ line deletion	Somatic deletion
36:1	adenocarcinoma	colon	high	*MLH1*	exons 1-2	exons 1-2
4:4	adenocarcinoma	stomach	high	*MLH1*	exons 3-5	absent
4:5	Adenocarcinoma	transverse colon	high	*MLH1*	exons 3-5	absent
11:12	adenocarcinoma	endometrium	high	*MLH1*	exon 16	absent
1:39	adenocarcinoma	endometrium	high	*MLH1*	exon 16	absent
1:32	adenocarcinoma	transverse colon	high	*MLH1*	exon 16	exon 16
76:1	adenocarcinoma	sigmoid colon	high	*MSH2*	exon 8	absent

* Tubulovillous adenoma with central adenocarcinoma.
† Adenocarcinoma of the intestinal type.

Following the identification of the *MLH1/MSH2* germ line deletion carriers, we applied microsatellite marker analysis and the MLPA assay to the cancer specimens of these patients to gain further insight into the frequency and the nature of the second, somatic mutational event, commonly referred to as LOH, involved in HNPCC tumorigenesis. A total of 11 formalin-fixed

cancers from six genomic deletion carriers (seven CRCs, one ovarian, one endometrial, one kidney cancer, and one astrocytoma) were available for investigation.

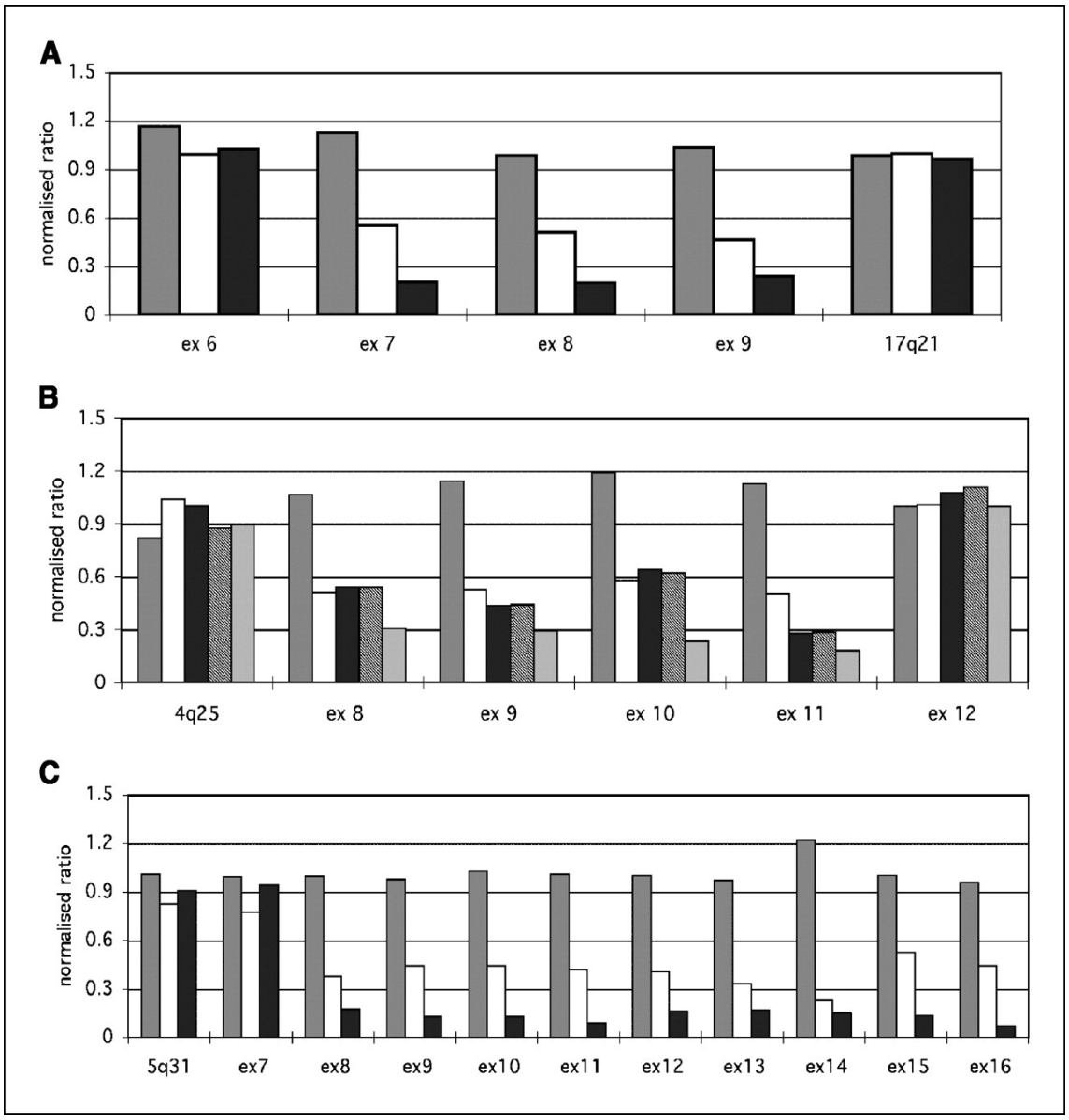


Figure 3.2. 2. Gene dosage analysis in leukocyte- and tumor-derived DNA from *MLH1/MSH2* germ line deletions carriers. *A,* healthy control (*gray*), patient 1806 carrying the *MLH1* exon7_9del mutation (*white*), and his colorectal cancer (*black*). *B,* healthy control (*gray*), patient 2227 carrying a MSH2 exon8_11del mutation (*white*), his colorectal (*black*), urothelial (*shaded*), and brain cancer (*light gray*). *C,* healthy control (*gray*), patient 2264 carrying a MLH1 exon 8_16del mutation (*white*) and his colorectal cancer (*black*).

As depicted in Table 3.2. 2A, MLPA analysis revealed that four (57%) out of seven CRCs, as well as one astrocytoma, actually harbor somatic deletions identical to the ones identified in the germ line (three *MLH1* and two *MSH2*) and evidenced by an average decrease in gene dosage

(normalized ratio) from 0.45 (±0.09 SD; germ line) to 0.18 (±0.07 SD; tumor; Fig. 2). One colorectal and one urothelial carcinoma showed loss of one exon only. No copy number changes were detected in the remaining four tumors (two CRCs, one ovarian, and one endometrial cancer).

In order to further substantiate the high frequency of cancers harboring a somatic deletion identical to the one in the germ line, seven additional tumors (four CRCs, two endometrial, and one stomach cancer) from Finnish HNPCC patients carrying an *MSH2* or *MLH1* germ line deletion were investigated (Table 3.2. 2B). Two out of four CRCs were found to carry identical, biallelic deletions which were absent in the remaining cancers . Thus, although based on an arguably small number of cases, our findings from two independent sets of patients indicate that the occurrence of somatic deletions identical to the ones in the germline is a frequent event in HNPCC-related colorectal tumorigenesis (6 out of 11 CRCs; 55%). Remarkably, none of the specimens from the Swiss or Finnish tumors showed evidence of large somatic deletions encompassing the entire, respective, gene locus.

With regard to extracolonic cancers (n = 7), only one tumor, a grade 3 astrocytoma (patient 2227/1), was found to carry an identical, biallelic deletion. Intriguingly, two other cancers (cecum and kidney) from this patient, with an exon 8 to 11 germ line deletion, harbored an identical single exon deletion (exon 11; Table 3.2. 2A). A false-positive MLPA result could be excluded by directly sequencing the ligation-probe binding site for exon 11. Both introns 10 and 11 of *MSH2* comprise several *Alu* repeats of the AluSx subfamily, which have been shown to be involved in genomic rearrangements of MMR genes (18).

In order to distinguish between the possible mechanisms leading to homozygosity of the germ line mutation in the tumor, i.e., loss of the chromosome harboring the wild-type allele followed by chromosomal reduplication, mitotic recombination, or gene conversion (19), we investigated eight highly polymorphic short tandem repeat markers flanking the gene loci on the respective chromosome (chromosome 2 for *MSH2*, chromosome 3 for *MLH1*). As depicted in Supplementary Table 3.2. S3 (supporting online material), none of the tumors showed allelic loss at the markers flanking the respective gene locus. Therefore, loss of the chromosome carrying the wild-type allele followed by chromosomal reduplication as well as mitotic recombination per se, extending from the gene locus to the telomere, could be ruled out as the underlying mechanism responsible for loss of the wild-type allele. Hence, these findings indicate that rather a locus-restricted event, i.e., gene conversion, has occurred in all cancers which are homozygous for the germ line mutation.

This might, at least in part, be explained by the presence of short interspersed nuclear elements, particularly *Alu* repeats, which have been shown in several studies to be involved in germ line *MLH1/MSH2* locus rearrangements (3, 8, 18, 20). The recurrent pattern of somatic deletions identical to the germ line indicate that it is possible that *Alu*-mediated gene conversion at the *MLH1* and *MSH2* loci is frequently occurring in HNPCC-associated tumors and that it is not restricted to CRC only, as evidenced by our findings on kidney and brain cancer specimens. Moreover, inactivation of the wild-type allele by gene conversion seems to be independent of the type of germ line mutation.[5]

If gene conversion is indeed the predominant mechanism which leads to inactivation of the wild-type allele, the LOH frequency at the *MLH1* or *MSH2* loci would be expected to be low. This is supported by data from Kruse et al., who observed LOH in only one out of nine skin tumors from eight unrelated Muir-Torre patients with *MSH2* mutations (7). Clearly, future studies on larger sets of HNPCC patients carrying *MLH1/MSH2* mutations are needed to conclusively establish the frequency of gene conversion events in colorectal as well as extracolonic cancers. These investigations should also help to elucidate the mechanistic role of *Alu*-mediated recombination in the generation of the second, somatic mutation in tumors from MMR gene mutation carriers.

In conclusion, large genomic deletions in *MLH1/MSH2* were found to account for 4.3% and 10.7% of *MLH1* and *MSH2* mutations, respectively, in our set of Swiss HNPCC families. Deletion carriers were statistically significantly later diagnosed of CRC and exhibited more extracolonic cancers when compared with "conventional" *MLH1/MSH2* mutation carriers. Analysis of cancer specimens from two independent sets of Swiss and Finnish *MLH1/MSH2* deletion carriers revealed (*a*) that somatic deletions identical to the ones in the germ line occur frequently (55%) in CRCs and (*b*) that this type of inactivation of the wild-type allele is also present in extracolonic HNPCC-associated tumors. Chromosome-specific marker analysis implies that loss of the wild-type allele predominantly occurs through locus-restricted recombinational events, i.e., gene conversion, rather than mitotic recombination or deletion of the respective gene locus.

Acknowledgments

We thank all patients for their participation in this study as well as their respective doctors for contributing clinical information and tumour specimens. We also thank Sibylle Bertschin, Marianne Haeusler, Ritva Haider and Thomas Woodtli for excellent technical assistance as well as Waltraut Friedl and Elisabeth Mangold for the *MLH1/MSH2* deletion control samples

Supplementary Table 3.2. 3: Loss of heterozygosity (LOH) analysis in cancers from *MSH2* (a) and *MLH1* (b) germline deletion carriers displaying either somatic deletions identical to the ones in the germline or loss of a single exon only (2227/1CRC and UC). * refers to the physical distance from the respective gene locus; CRC:colorectal cancer, UC: urothelial cancer, AC: astrocytoma, MSI:microsatellite instability, na: repeatedly failed to amplify, ni: notinformative.

a)

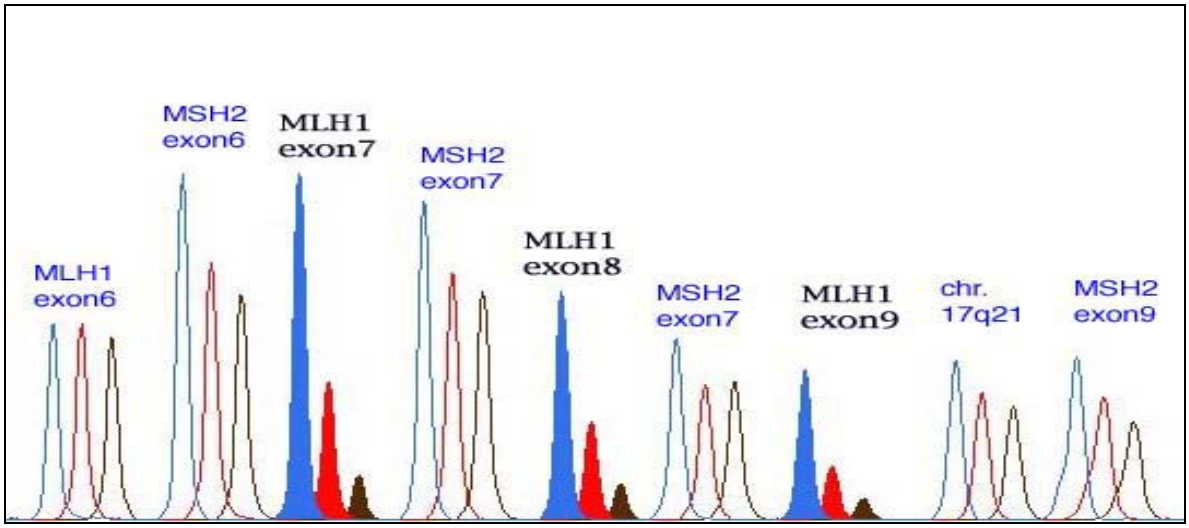
microsatellite marker	distance (Mb)*	chromosomal band	2227/1 CRC	2227/1 UC	2227/1 AC	2264 CRC
D2S168	36.2	2p25.1	no LOH	LOH	na	ni
D2S165	19.0	2p23.2	no LOH	LOH	LOH	ni
D2S367	13.2	2p22.3	ni	ni	ni	ni
D2S391	1.2	2p21	no LOH	no LOH	no LOH	no LOH
***MSH2* (2p21)**						
D2S337	14.0	2p15	ni	ni	ni	ni
D2S2110	25.4	2p13.2	no LOH	no LOH	no LOH	na
D2S2333	37.8	2p11.2	ni	ni	ni	na
D2S347	76.3	2q14.3	ni	ni	ni	na

b)

microsatellite marker	distance (Mb)*	chromosomal band	1806/1 CRC1	1806/1 CRC2	1806/3 CRC
D3S1304	30.1	3p26.1	LOH	no LOH	no LOH
D3S1263	25.5	3p25.3	MSI	no LOH	na
D3S2338	20.2	3p24.3	no LOH	MSI	MSI
D3S1266	9.1	3p24.1	no LOH	no LOH	no LOH
D3S1277	2.4	3p23	no LOH	no LOH	no LOH
***MLH1* gene locus (3p22.3)**					
D3S1300	23.4	3p14.2	no LOH	MSI	no LOH
D3S1566	33.3	3p13	no LOH	no LOH	no LOH
D3S1278	79.5	3q13.31	no LOH	MSI	MSI

Supplementary Figure 3.2.3: MLPA GENESCAN electropherograms of *Fa1806(MLH1* exon7_9 Deletion).

Blue: Indicated Normal control. Red: Fa 1806(MLH1 exon7_9 deletion) from constitution DNA
Dark:colorectal cancer from patient 1806 indicating a homozygous *MSH2* exon8_16del.
Simple line denote non-deletion exons and chromosome control. Filled field denote exons that deleted.



3.2.5 References

1. Lynch HT, de la Chapelle A, Hereditary colorectal cancer, N.Engl.J.Med.2003;348:919-932.

2. Wang Y, Friedl W, Lamberti C et al, frequent occurrence of large genomic deletions in MSH2 and MLH1 genes, Int.J.Cancer. 2003;103:636-641.

3. Wijnen J, van der Klift H, Vasen H et al., MSH2 genomic deletions are a frequent cause of HNPCC, Nat. Genet. 1998;20:326-328.

4. Peltomaki P, Lynch syndrome genes, Fam. Cancer. 2005;4:227-232.

5. Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G., Relativequan- tification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification., Nucleic Acids Res, 2002;30:e57.

7. Kruse R, Rutten A, Hosseiny-Malayeri HR et al., "Second hit" in sebaceous tumors from Muir-Torre patients with germline mutations in MSH2: allele loss is not the preferred mode of inactivation., J.Invest Dermatol 2001;116:463-465

8. Nystrom-Lahti M, Kristo P, Nicolaides NC et al., Founding mutations and Alu-mediated recombination in hereditary colon cancer., Nat. Med. 1995;1:1203-1206

9. Holmberg M, Kristo P, Chadwicks RB et al., Mutation sharing, predominant involvement of the MLH1 gene and description of four novel mutations in hereditary nonpolyposis colorectal cancer., Mutations in brief no. 144. Hum. Mutat. 1998;11:482.

10. Miller SA, Dykes DD, Polesky HF., A simple salting out procedure for extracting DNA from human nucleated cells, Nucleic Acids Res. 1988;16:1215.

11. Boland CR, Thibodeau SN, Hamilton SR et al., A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer., Cancer. Res. 1998;58:5248-5257

12. Truninger K, Menigatti M, Luz J et al, Immunohistochemical analysis reveals high frequency of PMS2 defects in colorectal cancer., Gastroenterology 2005;128:1160- 1171

13. Canzian F, Salovaara R, Hemminki A et al., Semiautomated assessment of loss of heterozygosity and replication error in tumors., Cancer Res .1996;56:3331-3337.

14. Taylor CF, Charlton RS, Burn J, Sheridan E, Taylor GR., Genomic deletions in MSH2 or MLH1 are a frequent cause of hereditary non-polyposis colorectal cancer: identification of novel and recurrent deletions by MLPA, Mutat. 2003;22:428-433.

15. Wehner M, Mangold E, Sengteller M et al., Hereditary nonpolyposis colorectal cancer: pitfalls in deletion screening in MSH2 and MLH1 genes., Eur. J. Hum. Genet. 2005;13:983-986.

16. Nakagawa H, Hampel H, de la Chapelle A., Identification and characterization of genomic rearrangements of MSH2 and MLH1 in Lynch syndrome (HNPCC) by novel techniques, Hum .Mutat. 2003;22:258.

17. Raedle J, Schaffner M, Esser N et al., Frequency of the Amsterdam criteria in a regional German cohort of patients with colorectal cancer., Z. Gastroenterol 2002;40:561-568

18. van der Klift H, Wijnen J, Wagner A et al., Molecular characterization of the spectrum of genomic deletions in the mismatch repair genes MSH2, MLH1, MSH6, and PMS2 responsible for hereditary nonpolyposis colorectal cancer (HNPCC)., Genes. Chromosomes Cancer 2005;44:123-138

19. Tischfield JA, Loss of heterozygosity or: how I learned to stop worrying and love mitotic recombination., Am. J. Hum.Genet 1997;61:995-999.

20. Charbonnier F, Olschwang S, Wang Q et al., MSH2 in contrast to MLH1 and MSH6 is frequently inactivated by exonic and promoter rearrangements in hereditary nonpolyposis colorectal cancer., Cancer Res. 2002;62:848-853

CHAPTER 3.3

3.3 A *de novo* MLH1 germline mutation in a 31 year old colorectal cancer patient

Martina Plasilova, Jian Zhang, Roberta Okhowat, Giancarlo Marra, Markus Mettler,
Hansjakob Mueller and Karl Heinimann

Genes Chromosomes Cancer. 2006 Dec;45(12):1106-10

3.3.1 Abstract

Hereditary non-polyposis colorectal cancer (HNPCC) is an autosomal dominantly inherited cancer predisposition syndrome caused by mutations in DNA mismatch repair (MMR) genes, predominantly MSH2 and MLH1. Here we report the first proven de novo germline mutation in MLH1 (c.666dupA) identified in a 31 year old colorectal cancer patient carrying the alteration in all three germ layers (heterozygous) as well as in his colon cancer (homozygous). The mutation was absent in both biological parents and all sibs available. Despite extensive segregation analysis the parental origin of c.666dupA could not be conclusively determined, representing either a single mutational event in a parental ger cell or (maternal) gonadal mosaicism. Though rare, consequential application of the Bethesda guidelines for genetic testing should allow the clinician to readily identify colorectal cancer patients carrying de novo MMR gene mutations.

3.3.1 Introduction

Hereditary non-polyposis colorectal cancer (HNPCC [MIM 114500]) is an autosomal dominantly inherited cancer predisposition syndrome which accounts for approximately 5% of all colorectal cancers (1). The disease is caused by germline mutations in DNA mismatch repair (MMR) genes, predominantly MSH2 and MLH1. Mutation carriers develop, mainly right-sided, colorectal cancer (CRC) at a mean age of 43 years with CRCs typically exhibiting microsatellite instability (MSI) ,the molecular hallmark of MMR deficiency.

To identify individuals with HNPCC who should be tested for MSI the Bethesda guidelines have been introduced (2, 3). Both, the former and the revised Bethesda guidelines allow investigating individuals with apparently sporadic CRC diagnosed before age 45 and 50 years, respectively. According to the literature, *de novo* germline mutations in MMR genes are exceptionally rare which stands in clear contrast to other hereditary CRC syndromes, such as familial adenomatous polyposis coli, juvenile polyposis and Peutz-Jeghers syndrome, where 25% to 29% of germline mutations in the respective genes are thought to have arisen *de novo* (4-6). Although recurrent *de*

novo mutations in MSH2 have been inferred by haplotype analysis, only one patient with a proven *de novo* germline mutation (in MSH2) has been reported to date (7, 8). Here we present conclusive evidence for a *de novo* germline mutation in the MLH1 gene.

3.3.3 Results and Discussion

Following colonoscopy because of blood in the stool, 31 year old male patient 2247/1 was diagnosed of an invasive adenocarcinoma located at the left colonic flexure (pT3, G3 pN0) and a left-sided hemicolectomy was performed. The detailed family history did not reveal any further first or second degree family member afflicted with cancer. The patient's family originated from Italy and the parents, aged 52 and 57, were not related to each other. Except for the patient, all family members including 4 more siblings were healthy (Figure 3.3. 1). Since the patient fulfilled the Bethesda guidelines, DNA from formalin-fixed colorectal cancer tissue (>70% tumor contents) was investigated for the presence of MSI using the recommended NCI panel of microsatellite markers 2. All markers were found to display novel alleles, corresponding to a MSI-high status in the tumor. Subsequent immunohistochemical analysis for the presence of MMR proteins (MLH1, MSH2, PMS2, MSH6 and MSH3) in the CRC revealed concomitant loss of expression of *MLH1* and PMS2. Consequently, bi-directional DNA sequencing of the coding sequence of *MLH1* (GenBank no. NM_0000249.2; primer sequences available upon request) identified a novel mutation in exon 8, c.666dupA, which results in a frameshift leading to a first premature stop codon at position 225 (p.Ser225X; Figure 3.3. 2a). The mutant protein is expected to lack the MLH/PMS interaction domains 9.

Except for the patient, all family members including 4 more siblings were healthy (Figure 3.3. 1). Since the patient fulfilled the Bethesda guidelines, DNA from formalin-fixed colorectal cancer tissue (≥70% tumor contents) was investigated for the presence of MSI using the recommended NCI panel of microsatellite markers (2). All markers were found to display novel alleles, corresponding to a MSI-high status in the tumor. Subsequent immunohistochemical analysis for the presence of MMR proteins (MLH1, MSH2, PMS2, MSH6 and MSH3) in the CRC revealed concomitant loss of expression of MLH1 and PMS2. Following bi-directional DNA sequencing of the coding sequence of MLH1 (GenBank no.NM_0000249.2;) identified a novel mutation in exon 8, c.666insA, which results in a frameshift leading to a first premature stop codon at position 225 (p.Ser225X; Figure 3.3. 2a). The mutant protein is expected to lack the MLH/PMS interaction domains (9).

To confirm paternity, a panel of 9 highly polymorphic short tandem repeat markers (AmpFlSTR Profiler kit; Applied Biosystems, Rotkreuz, Switzerland) were assessed in the patient and his parents. No inconsistency between the parental and the patient's alleles was observed and the segregation pattern was according to mendelian inheritance (data not shown). These results conclusively show that the MLH1 mutation c.666insA has indeed occurred *de novo*.

Following genetic counselling of the family, the carrier status in two sisters of the patient, aged 27 and 23 years, was determined. The c.666insA mutation t could be identified in none of them, a finding which was in each case confirmed on two independently drawn blood samples (data not shown).

To determine on which parental chromosome the mutation had occurred, ten single nucleotide polymorphisms (SNPs) intragenic of MLH1 (rs9311149, rs4647215, rs4234259, rs4647250, rs4647260, rs1558528, rs2286939, rs655045, rs2286940 and rs2241031) as well as 7 polymorphic microsatellite markers on chromosome 3 (D3S1597, D3S3611, D3S2338, D3S1277, D3S1300, D3S1566, D3S1278) were assessed in all available family members (Figure 3.3. 1). SNPs were informative in the father only and showed that the paternal chromosome transmitted to the patient was also present in one of his sisters. Marker analysis revealed that one of the maternal chromosomes was only present in the patient but in none of his sibs. Assuming gonadal mosaicism, it is therefore conceivable that the c.666insA mutation may actually have arisen on the maternal chromosome.

In order to substantiate this assumption we assessed the presence of loss of heterozygosity (LOH) in the patient's CRC. Since inactivation of the wild-type allele in the tumor is frequently associated with loss of the polymorphic marker alleles on the wild-type chromosome, identification of LOH at these markers could help to determine the parental chromosome carrying the mutation. Direct sequencing of exon 8 in the tumor DNA identified the c.666dupA mutation in a nearly homozygous state (Figure 3.3.s 2b and 3e). Subsequent multiplex ligation-dependent probe amplification (MLPA) analysis showed that both *MLH1* gene copies were present in the tumor (data not shown). Although MSI hampered marker analysis in the cancer, no allelic loss was observed at any of the informative markers (D3S1277 and D3S2338), indicating that loss of the wild-type allele had occurred through a locus-restricted recombinational event. As the patient was not informative for any SNP within the *MLH1* gene locus we could not further determine the parental chromosome harbouring the c.666dupA mutation.

In order to substantiate this assumption we assessed the presence of loss of heterozygosity (LOH) in the patient's CRC. Since inactivation of the wild-type allele in the tumor is frequently associated with loss of the polymorphic marker alleles on the wild-type chromosome, identification of LOH at these markers could help to determine the parental chromosome carrying the mutation. Direct sequencing of exon 8 in the tumor DNA identified the c.666insA mutation in a nearly homozygous state (Figure 3.3. 2b). Subsequent multiplex ligation-dependent probe amplification (MLPA) analysis showed that both MLH1 gene copies were present in the tumor (data not shown). Although MSI hampered marker analysis in the cancer, no allelic loss was observed at any of the informative markers (D3S1277 and D3S2338), indicating that loss of the wild-type allele had occurred through a locus-restricted recombination event. As the patient was not informative for any SNP within the MLH1 gene locus we could not further determine the parental chromosome harbouring the c.666insA mutation. Since the mutation was present in tissues of endodermal (colorectal cancer) and mesodermal (blood leukocytes, colonic smooth muscle) origin, it is unlikely that the mutational event happened postzygotically leading to somatic mosaicism in the patient. Thus, the c.666insA mutation either represents a single mutational event in a parental germ cell or (maternal) gonadal mosaicism as indicated by segregation analysis.

With regard to published data and for reasons unknown, the overall *de novo* mutation frequency in MLH1 and MSH2 appears to be very low. In our group of Swiss HNPCC index patients, *de novo* mutations may represent approximately 2% (1/47) of all MLH1 germline mutations identified. Though very rare, application of the Bethesda guidelines for genetic testing should nevertheless allow the clinician to identify CRC patients carrying *de novo* MMR gene mutations.

Acknowledgments

We thank all family members for their participation in this study as well as the pathologists for contributing tumour specimens. We also thank Sibylle Bertschin, Nemya Boesch, Marianne Haeusler, Ritva Haider and Thomas Woodtli for excellent technical assistance. This research was supported by grants from the Swiss National Science Foundation (no. 3200-067571) and the Swiss Cancer League / Oncosuisse (no. 01358-03-2003).

Figures

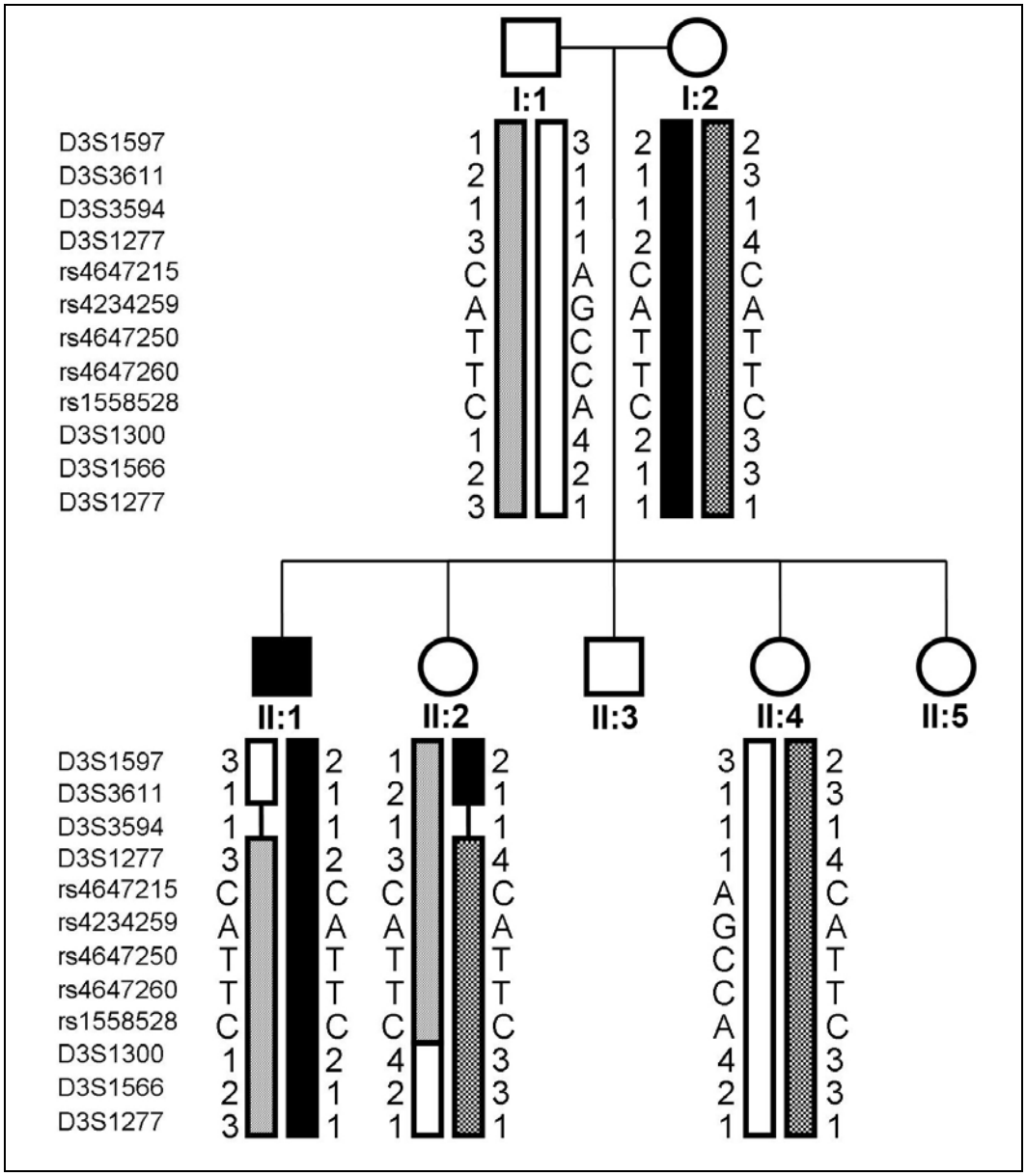


Figure 3.3.1: Pedigree of family 2247 depicting the individual haplotypes in patient 2247/1 (II:1) and his first degree relatives available for microsatellite marker and SNP analysis.

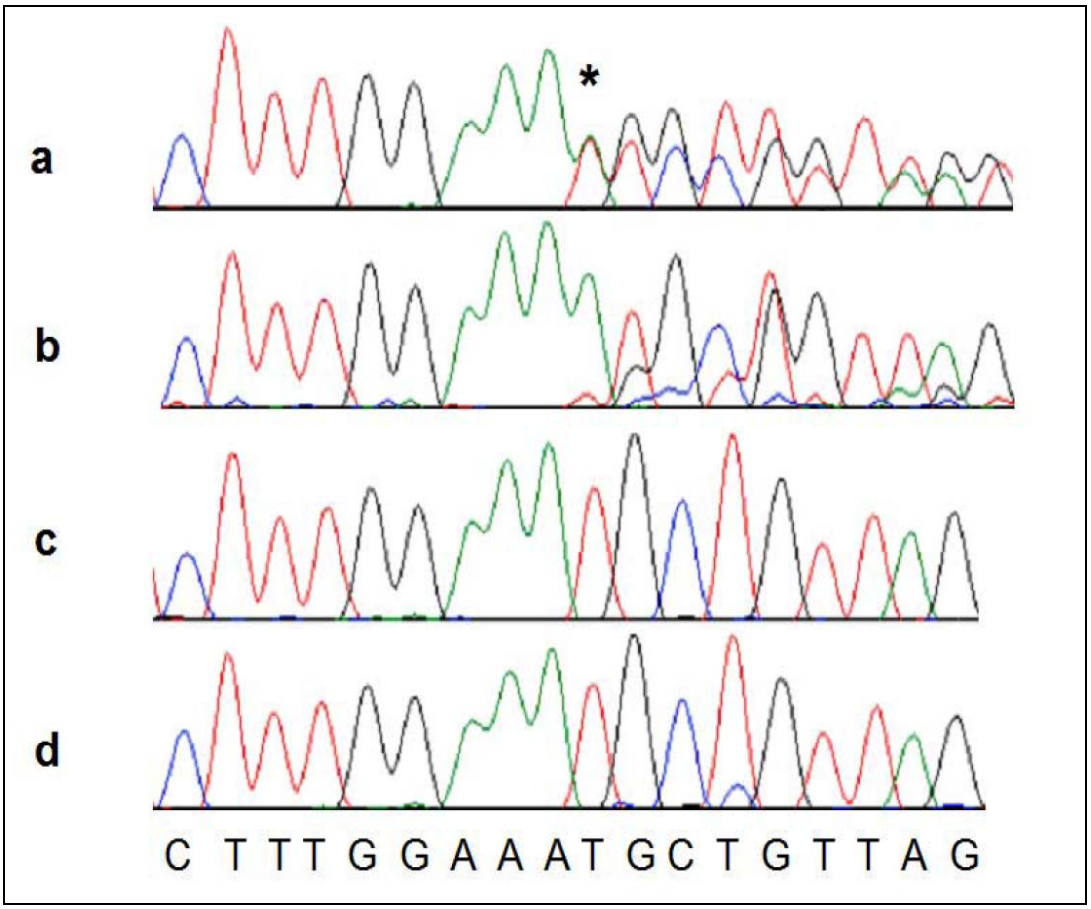


Figure 3.3.2:**Fragment length analysis of *MLH1* exon 8 in different tissues.** Sequencing electropherograms demonstrating the c.666dupA germ line mutation in *MLH1* (*) in a heterozygous state in leukocyte-derived DNA (a) and in a nearlyhomozygous state in tumor-derived DNA (b) from patient 2247/1. Panels c and d depict the wild-type sequence present in leukocyte-derived DNA from the patient's father (c) and mother (d)

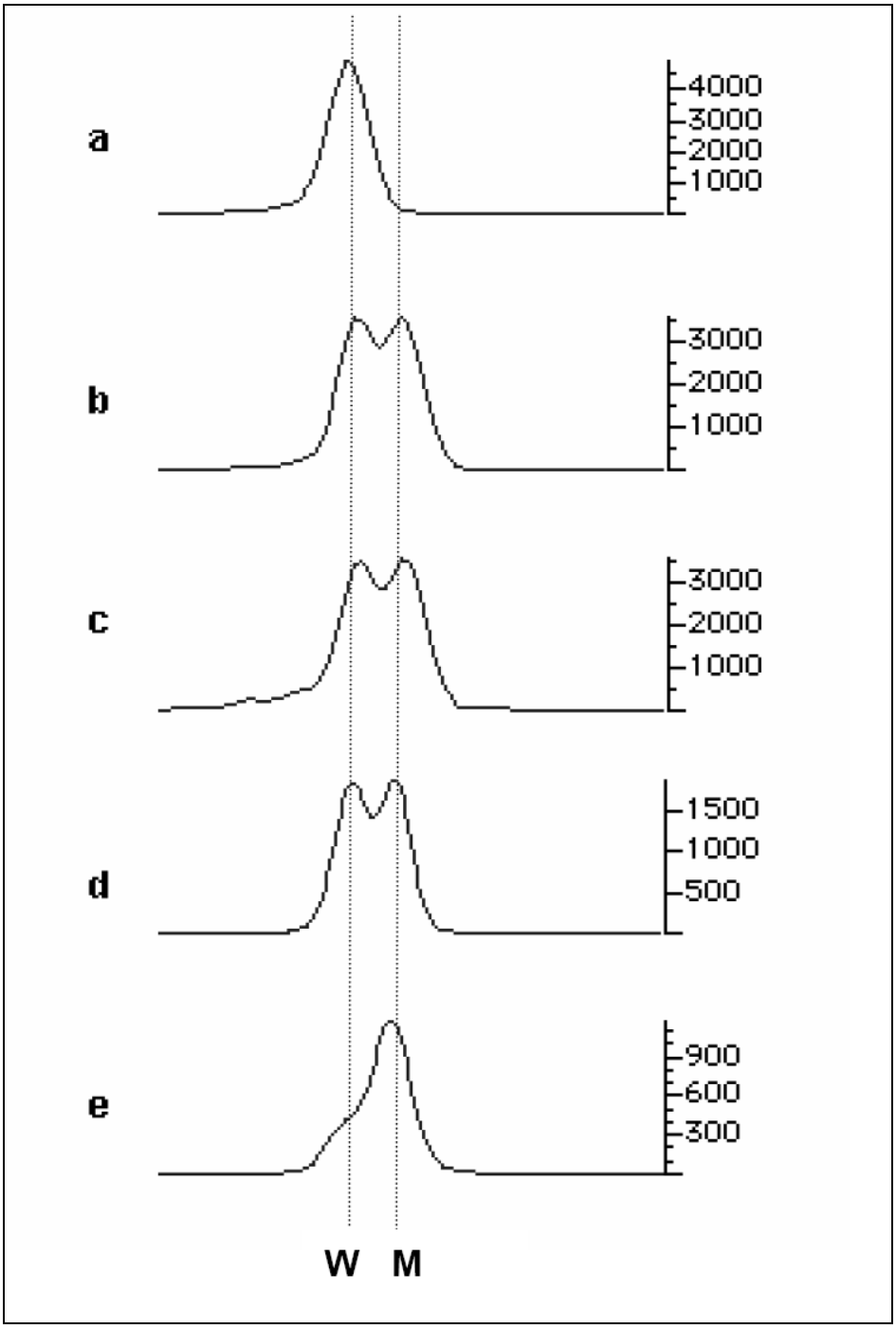


Figure 3.3.3: **Fragment length analysis of *MLH1* exon 8 in different tissues.**
a): leukocyte-derived DNA from a healthy proband displaying only the wild-type allele (w); b) to e): patient 2247/1 carrying the c.666dupA mutation (m) in DNA samples from peripheral blood leukocytes (b), hair follicles (c), sperms (d) and colon cancer (e).

3.3.3 References

1. Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Eshleman JR, Burt RW, et al. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. Cancer Res 1998;58(22):5248-57.
2. Umar A, Boland CR, Terdiman JP, Syngal S, de la Chapelle A, Ruschoff J, et al. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. J Natl Cancer Inst 2004;96(4):261-8.
3. Aretz S, Uhlhaas S, Caspari R, Mangold E, Pagenstecher C, Propping P, et al. Frequency and parental origin of de novo APC mutations in familial adenomatous polyposis. Eur J Hum Genet 2004;12(1):52-8.
4. Schreibman IR, Baker M, Amos C, McGarrity TJ. The hamartomatous polyposis syndromes: a clinical and molecular review. Am J Gastroenterol 2005;100(2):476-90.
5. Westerman AM, Entius MM, Boor PP, Koole R, de Baar E, Offerhaus GJ, et al. Novel mutations in the LKB1/STK11 gene in Dutch Peutz-Jeghers families. Hum Mutat 1999;13(6):476-81.
6. Kraus C, Kastl S, Gunther K, Klessinger S, Hohenberger W, Ballhausen WG. A proven de novo germline mutation in HNPCC. J Med Genet 1999;36(12):919-21.
7. Desai DC, Lockman JC, Chadwick RB, Gao X, Percesepe A, Evans DG, et al. Recurrent germline mutation in MSH2 arises frequently de novo. J Med Genet 2000;37(9):646-52.
8. Lynch HT, de la Chapelle A. Hereditary colorectal cancer. N Engl J Med 2003;348(10):919-32.
9. Boland CR, Fishel R. Lynch syndrome: form, function, proteins, and basketball. Gastroenterology 2005;129(2):751-5.

CHAPTER 4

Predispositions to Colorectal Cancer with Pre-existing polyposis: Familial Adenomatous Polyposis (FAP) and MYH Assoicated Polyposis (MAP)

CHAPTER 4.1

4.1 Disease severity and genetic pathways in attenuated familial adenomatous polyposis vary greatly but depend on the site of the germline mutation.

Sieber OM, **Segditsas S**, **Knudsen AL**, **Zhang J**, **Luz J**, **Rowan AJ**, **Spain SL**, **Thirlwell C**, **Howarth KM**, **Jaeger EE**, **Robinson J**, **Volikos E**, **Silver A**, **Kelly G**, **Aretz S**, **Frayling I**, **Hutter P**, **Dunlop M**, **Guenther T**, **Neale K**, **Phillips R**, **Heinimann K**, **Tomlinson IP**. (in collaboration with Molecular and Population Genetics Laboratory, Cancer Research UK, London Research Institute, London, UK.)

Gut. 2006 Oct;55(10):1440-8.

4.1.1 Abstract

Attenuated familial adenomatous polyposis (AFAP) is associated with germline mutations in the 5', 3' and exon 9 of *APC*. These mutations probably encode a limited amount of functional APC protein. **Methods and Results.** We found that colonic polyp number varies greatly among AFAP patients, but members of the same family tended to have more similar disease severity. 5'-mutants generally had more polyps than the other patients. We analysed somatic *APC* mutations/LOH in 235 tumours from 35 patients (16 families) with a variety of AFAP-associated germline mutations. Like two previous studies of individual kindreds, we found bi-allelic changes ('third hits') in some polyps. We found that the 'third hit' probably initiated tumorigenesis. Somatic mutation spectra were similar in 5'- and 3'-mutant patients, often resembling classical FAP. In exon 9-mutants, by contrast, 'third hits' were more common. Most 'third hits' left three 20-amino acid repeats (20AARs) on the germline mutant *APC* allele, with LOH (or proximal somatic mutation) of the wild-type allele; but some polyps had loss of the germline mutant, with mutation leaving one 20AAR on the wild-type allele. **Conclusions.** We propose that mutations, such as nt4661insA, that leave three 20AARs are preferentially selected *in cis* with some AFAP mutations, because the residual protein function is near-optimal for tumorigenesis. Not all AFAP polyps appear to need 'three hits', however. AFAP is phenotypically and genetically heterogeneous. In addition to effects of different germline mutations, modifier genes may be acting on the AFAP phenotype, perhaps influencing the quantity of functional protein produced by the germline mutant allele.

4.1.2 Introduction

Classical familial adenomatous polyposis (FAP) is caused by germline mutations in the adenomatous polyposis coli (*APC*) gene between codons 178 and 1580. FAP patients typically develop hundreds to thousands of adenomatous polyps in the colon and rectum by the third decade of life. If left untreated, one or more adenomas progress to carcinoma by 45 years of age. Extra-colonic features, such as polyps of the upper gastrointestinal tract, desmoid tumours and osteomas, are also common. Attenuated FAP (AFAP or AAPC) patients generally present with a lower number (<100) of colorectal adenomas by their fourth decade and have a later age of onset of colorectal cancer (mean age 55 years) (1-3). In some AFAP patients, extra-colonic features have been reported to be infrequent ((4)),although other AFAP patients – such as those with hereditary desmoid disease – have severe extra-colonic disease (5, 6). AFAP is associated with germline mutations in specific regions of the *APC* gene (Figure 4.1. 1): the 5'-end (codons 1-177, exons 1-4); the 3'-end (distal to codon 1580); and the alternatively spliced region of exon 9 (codons 311-408) (3, 7, 8). The molecular mechanism(s) underlying these genotype-phenotype associations for *AP* remains largely unknown.

APC is a tumour suppressor gene and almost all mutations truncate the protein or take the form of allelic loss (loss of heterozygosity, LOH). Several genetic studies of colorectal adenomas from FAP patients have shown that somatic *APC* mutations are dependent on the position of the germline *APC* mutation (Figure 4.1. 1) (9-11). The APC protein contains seven 20-amino acid repeats (20AARs) which are involved in degrading the transcriptional cofactor beta-catenin and hence negatively regulate Wnt signalling. In colorectal polyps, germline mutations between codons 1285 and 1378 leave only one 20AAR intact and are strongly associated with somatic loss of the wild-type *APC* allele. LOH usually occurs through mitotic recombination, thus leaving two identical alleles and a total of two 20AARs in the tumour cell (12). FAP patients who carry germline mutations before codon 1285 (no 20AARs) tend to have somatic mutations which leave one or, more commonly, two 20AARs in the protein. Finally, patients with germline mutations after codon 1398 (two or three 20AARs) tend to have somatic mutations before codon 1285. The same associations are also found in sporadic colorectal tumours (13). This interdependence of 'first' and 'second hits' shows that selective constraints on *APC* mutations are active and that an optimum level of beta-catenin mediated signalling must be achieved for the tumour cell to grow (10). There is no reason to expect that AFAP polyps are not subject to the same selection for optimal Wnt signalling as other colorectal adenomas.

The 'first hit-second hit' associations can explain why FAP patients with germline *APC* mutations between codons 1285 and 1378 have particularly severe colorectal disease, because the associated allelic loss occurs at a higher spontaneous frequency than the somatic truncating mutations selected in other FAP patients (9). Conversely, the milder disease in AFAP patients may be explained if the mutations required to give the polyp cell a strong selective advantage are difficult to acquire. Spirio *et al* (1) studied colorectal tumours from a single AFAP family with a germline *APC* mutation in the 5'-end of the gene (codon 142FS). About 12% of their polyps showed loss of the germline mutant allele, implying that this was a 'third hit' subsequent to a mutation on the germline wild-type allele. Furthermore, a large proportion (36%) of the truncating somatic mutations detected were 1bp insertions at an A6-tract between nucleotides 4661-4666 (codons 1554-1556). Spirio *et al* (1) concluded that germline mutations in the 5' region of *APC* encode proteins that retain residual activity, owing to alternative splicing or initiation of translation. Somatic mutations would be required not only to inactivate the wild-type allele, but also to reduce the residual activity of the mutant germline allele. Su *et al* (14) studied 9 adenomas from an AFAP family with agermline mutation (R332X) in exon 9. They found 'third hits', including loss of the germline mutant allele and 4661insA, and showed the latter to occur on the germline mutant chromosome. The APC isoprotein lacking exon 9 retained at least partial ability to down-attenuation of the phenotype. Su *et al* (14) suggested that exon 9-mutant AFAP patients develop more tumours than the general population because the germline mutant *APC* allele could be inactivated by a broad spectrum of somatic mutations, including some, such as nt4661insA, that would not normally affect an wild-type *APC* allele. The existing studies only analysed single families, but established the important principle that 'third hits' can occur in AFAP. These 'third hits' could be LOH or mutation at codon 4661. In this study, we analysed a larger number of AFAP families with the following aims

- to search for phenotypic differences among AFAP families, both between and within kindreds with mutations in each of the three AFAP-associated regions of *APC*
- to determine whether the two families reported were typical of AFAP
- to find out the somatic *APC* mutation spectrum in AFAP patients with 3'-mutations and to compare this with the other AFAP-associated regions of *APC*
- to find out why 4661insA is such a common 'third hit'
- to delineate the pathways of somatic *APC* mutation in AFAP, with emphasis on whether polyps end up with the optimal genotype as predicted by studies of classical FAP
- to determine whether 'three hits' are always needed in AFAP.

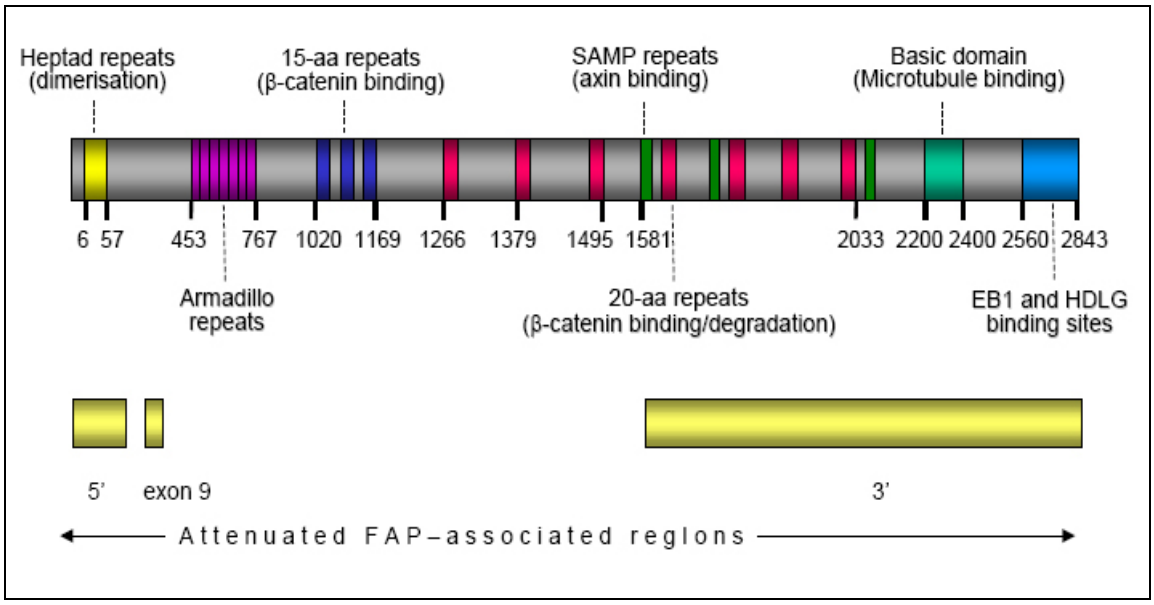


Figure 4.1.1. Representation of the APC protein comprising important functional domains and showing regions of the protein germline mutation of which are associated with AFAP

4.1.3 Patients and Methods

We contacted Polyposis Registries in the United Kingdom, Switzerland, Germany and Denmark with a request to study colorectal tumours from AFAP patients with characterized germline APC mutations in the 5'- or 3'-regions of the gene (codons 1-177 and 1580-2843) or in the alternatively spliced region of exon 9 (codons 311-408). In total, 235 fresh-frozen or formalin-fixed, paraffin-embedded colorectal tumours were obtained from 35 individuals in 16 families. All patients gave written informed consent. 231 of the tumours were colorectal adenomas, almost all of tubular morphology and with a median diameter of 3mm (range=1- 17mm); four tumours were colorectal carcinomas (median diameter=5mm, range=2-20mm). 30 tumours were from 6 AFAP patients from 5 families with germline APC mutations in the 5' region of the gene (G126X, 141FS, Q163X, 170FS, 173FS). 79 tumours were from 10 AFAP patients from 5 families, each of which carried the relatively common R332X nonsense mutation in the alternatively spliced region of APC exon 9. 126 tumours were from 19 AFAP patients from 6 families with germline APC mutations in the 3'-region of the gene (1597FS, 1738FS, 1919FS, 1943FS, 1982FS, 2078FS). Clinical details (APC germline mutation, gender, age at presentation, polyp count) were obtained and are being analysed as part of a larger study of phenotype in AFAP (A.L.Knudsen, in preparation); numbers of polyps analysed per patient are summarised in Table 4.1.1. Paired normal tissue was available for all patients. H&E-stained sections were

prepared from each tumour to confirm the presence of at least 60% neoplastic tissue. DNA was extracted from tumour and normal tissue using standard methods.

Table 4.1.1. Characteristics of 35 patients with germline *APC* mutations in the three AFAP-associated regions (5', exon 9 and 3'; codons 1-177, 311-408 and >1580).
FS = frameshift; n/a = not available; * from reference(8); ** from reference(10)

Patient ID	Germline *APC* mutation	Gender	Age at presentation	Polyp count	Polyps analysed
AFX MK	G126X	M	36	834	6
DFAP48	141FS	F	56	2	1
AVC.III.2	Q163X	M	51	2100	4
554.iv.2/1112	170FS	F	32	1357	11
554.iii.2	170FS	M	50	1077	5
1464/1	173FS	M	39	"multiple"	3
673.iii.3/1132	R332X	F	49	200-300	6
1571.ii.2	R332X	F	68	5	5
578.AA	R332X	F	n/a	n/a	4
578.FPL	R332X	F	27	20	6
578.iii.9	R332X	M	41	n/a	17
578.iv.1	R332X	F	43	130	14
578.iv.4	R332X	F	27	n/a	9
578.iv.7	R332X	M	n/a	n/a	15
DFAP16	R332X	M	52	"multiple"	1
DFAP81	R332X	M	32	<100	2
344-40	1597FS	M	47	99	4
344-44	1597FS	M	43	50	3
01/266	1738FS	F	53	29	19
2233/3	1919FS	n/a	n/a	n/a	15
MD2976	1919FS	F	43	>100	21
77-11	1943FS	M	66	500	24
77-12	1943FS	M	56	500	2
77-40	1943FS	M	39	500	2
1460/6	1982 FS	F	65	>100	5
1460/28	1982 FS	F	33	8	1
1460/42	1982 FS	F	33	>100	4
1460/88	1982 FS	M	48	<100	1
1489/10	1982 FS	F	29	n/a	1
1624/04	1982 FS	F	40	<100	3
S73119	2078FS	F	52	<100	1
DW20284	2078FS	F	n/a	n/a	1
J42424	2078FS	F	n/a	n/a	3
L12562	2078FS	F	60	"numerous"	3
110.2.vi	2078FS	F	n/a	33	14

Mutation screening

All samples were screened for somatic *APC* mutations using fluorescence single strand conformational polymorphism (SSCP) analysis on the ABI3100 sequencer (details available from authors upon request). Fresh-frozen samples were screened between codons 1 and 1779. Owing to the limiting quantity of DNA, formalin-fixed, paraffin-embedded samples were screened between codons 1220 and 1603, an area encompassing the somatic mutation cluster region and extending beyond the first SAMP repeat involved in axin binding.(15)Samples with

bandshifts on SSCP analysis were sequenced in both forward and reverse orientations from a new PCR product.

Cloning

We wished to determine the phase of somatic *APC* mutations with respect to the germline wild-type or mutant allele, but the quality of DNA available from archival tumours was insufficient to allow long-range PCR amplification. We therefore identified a SNP (nt4479 A>G) within *APC* which was close enough to most somatic mutations of interest to be PCRamplified, and which was informative and linked to the disease-causing mutation. After amplification of a region encompassing the somatic *APC* mutation and the SNP, the PCR product was cloned and multiple clones were sequenced using the pGEM-T Easy Vector System II (Promega).

Loss of heterozygosity analysis

In the case of germline nonsense mutations in *APC*, loss of heterozygosity (LOH, allelic loss) analysis was performed using three microsatellite markers, D5S346, D5S421 and D5S656, which map close to *APC*. Where linkage information was available for the microsatellites studied, the allele targeted by the allelic loss was assigned as germline mutant or wild type. Where no linkage information was available, the allele targeted was determined by inspection of the sequencing electropherogram in constitutional and tumour DNA for the region containing the mutation. In the case of germline (and somatic) frameshift mutations, LOH analysis was performed using oligonucleotide primers which encompassed the germline insertion/deletion, which was then used as a pseudo-polymorphism for assessing loss. Standard methods of fluorescence-based genotyping on the ABI3100 sequencer were used. Allelic loss was scored at any informative marker if the area under one allelic peak in the tumour was reduced by more than 50% relative to the other allele, after correction for the relative peak areas of the alleles found in constitutional DNA of the same patient.

Multiplex ligation-dependent probe amplification (MLPA) analysis and real-time quantitative multiplex (RQM-)PCR

MLPA analysis to determine the copy number of the *APC* promoter and individual exons was performed on polyps with allelic loss at *APC* using the Salsa MLPA kit P043 APC (MRCHolland) according to manufacturer's instructions. RQM-PCR to determine the copy number of *APC* exon 14 (normalised against human serum albumin (*Alb*) exon 12) was performed as previously described. (16) The assay has previously been shown to be sensitive for tumour samples containing less than 30% contaminating normal tissue. (10)

4.1.4 Results

Overall phenotypic assessment

We have previously shown that disease severity (number of colorectal adenomas) in classical FAP patients varies considerably independent of the germline mutation, but that family members tend to have similar severities of disease .(17) We searched the published literature (details available from authors) for all patients who had germline mutations in the AFAP-associated regions of *APC* and with precisely-reported colorectal polyp counts at presentation. We then combined these data with our own. Patient age had no significant effect on polyp number. We then tested for familial aggregation of disease severity and found good evidence for this, both when all families were considered together ($p<0.00001$, Kruskal-Wallis test) and when families with germline mutations in the three AFAP-associated regions of *APC* were analysed separately ($p=0.0002$, $p=0.045$, $p=0.0005$ respectively, Kruskal-Wallis test). Whilst some effects of local clinical practice are possible, such strong associations are unlikely to result from systematic errors in polyp counting. We then calculated the median polyp count for each family irrespective of size, and tested whether this varied among the three groups with

Somatic mutations in tumours of patients with AFAP-associated germline APC mutations

Given that our data showed aggregation of disease severity within families, it became more likely that the individual kindreds analysed by previous studies (1, 14) had provided only a partial description of the genetic pathways of tumorigenesis in AFAP. We first screened colorectal tumours from 5'-mutant patients for somatic *APC* changes (Supplementary Table 4.1.1). We found truncating somatic mutations in 9 of 30 (30%) adenomas. Similar to adenomas from classical FAP patients with germline mutations before codon 1285 (9-11), all of the truncating mutations left either one or two 20AARs in the protein. Just two of the adenomas (7%) harboured a detected 'third hit', each in the form of loss of the germline mutant allele (Supplementary Table 4.1.1). Our results were consistent with those reported by Albuquerque *et al* (10) on the polyps of a single 5' mutant-patient, but differed from those of Spirio *et al* (1) in that we found no mutations at nucleotides 4661-6 or at any other site after the third 20AAR. It was notable that while most of the patients of Spirio *et al* (1) had presented with attenuated polyposis, the patient of Albuquerque *et al* had been reported to have about 100 adenomas and most of our 5'-mutant patients had presented with a classical FAP phenotype (Table 4.1.1). The family of Spirio *et al* (1) cannot therefore be considered representative of all patients with mutations in the AFAP-associated 5' region of *APC*.

For patients with exon 9 germline mutations, we found truncating somatic mutations in 47/79 (59%) adenomas (Table 4.1.2, Supplementary Table 4.1.2). Of the total of 50 truncating mutations, 33 (66%) were nt4661insA at codon 1554, and this change was always present on the germline mutant allele where assignment was possible. (An uncharacterised defect in DNA mismatch repair as a cause for this observation was excluded by analysing the microsatellite marker BAT26.) Three other mutations leaving three 20AARs (at codons 1518, 1530 and 1537) were found. LOH was found in 13/79 (16%) adenomas; this affected the wildtype allele in 9 cases and the mutant allele in 4 cases. Thirty-one (39%) adenomas had evidence of 'thirds hits', either two detected somatic changes or a single identified somatic change on the germline mutant allele. The data allowed three main genetic pathways to be identified in the exon 9-mutant patients' polyps with evidence of 'third hits' (Table 4.1.2):

(i) mutation leaving three 20AARs on germline mutant allele, plus loss of the wildtype allele;
(ii) mutation leaving three 20AARs on germline mutant allele, with undetectable mutation of the wildtype allele (most likely towards the 5' end of the gene, which could not be screened in all polyps, and leaving zero 20AARs);
(iii) mutation leaving one 20AAR on the wildtype allele plus loss of the germline mutant allele.

Table 4.1.2. Numbers of tumours with evidence of 'third hits' (somatic mutation of germline mutant allele) at *APC* in exon 9- and 3'mutant patients' polyps
20AAR1 = truncating mutation before first 20AAR, *etcetera*. Note that these are minimum estimates of the true frequency, not only because we could not screen the entire gene for mutations in small archival polyps, but also because it was not possible to assign all mutations to the germline mutant or wildtype allele.

Somatic mutation on germline wildtype allele ('second hit')	Somatic mutation on germline mutant allele ('third hit')	No. 5'-mutant patients	No. exon 9-mutant patients	No. 3'-mutant patients
LOH	20AAR3	0	6	0
LOH	20AAR2	0	0	2
20AAR0	20AAR3	0	2	1
20AAR1	20AAR3	0	2	0
20AAR0	LOH	0	0	1
20AAR1	LOH	2	4	3
20AAR2	LOH	0	0	2
Not found	20AAR3	0	20	0
Not found	20AAR2	0	2	2
Not found	LOH	0	0	9

For patients with 3' germline mutations, we found truncating somatic mutations in 35/126 (28%) adenomas (Table 4.1.2, Supplementary Table 4.1.3). Of the total of 36 truncating mutations, only 2 (6%) were nt4661insA. Three other mutations leaving three 20AARs (at codons 1537, 1576 and 1570) were found. LOH was found in 30/126 (23%) adenomas, equally affecting the wildtype and mutant alleles. Twenty (16%) adenomas had either two detected somatic changes or an identified somatic change of the germline mutant allele. There was no clear tendency for different families to acquire different somatic mutations (Supplementary Table 4.1.3). The data only allowed one consistent genetic pathway to be identified in the 3'-mutant patients' polyps with evidence of 'third hits', namely a mutation leaving one (or two) 20AARs on the germline wildtype allele, plus loss of the mutant allele (Table 4.1.2).

Comparison between somatic mutations in the three groups of patients

The somatic mutation spectra of the 5'- and 3'-mutant patients' tumours did not differ significantly from each other as regards: (i) proportion of mutations leaving one, two or three

20AARs (p=0.074, χ^2_2 test); (ii) overall LOH frequency (2/9 *versus* 30/126, p=0.64); and (iii) proportion of tumours with detected 'third hits' or an identified somatic change on the germline mutant allele (2/9 *versus* 20/126, p=0.45). However, whilst exon 9-mutant patients had a similar frequency of LOH (13/79, 22%, p=0.14) to the other patients, germline exon 9 mutants had a higher frequency of mutations that left three 20AARs (36/50 *versus* 5/45, p<0.001, χ^2_2 test) and a higher frequency of tumours with detected 'third hits' (31/79 *versus* 22/156, p<0.001, χ^2_22 test). In large part, these associations reflected the fact that nt4661insA was particularly common in the exon 9-mutant patients' tumours and exclusively targeted the mutant germline allele. Overall, our data were consistent with a large proportion of polyps in the 5'- and 3'-mutant patients developing along the 'classical' FAP pathway, their polyps showing similar somatic mutations to individuals with germline mutations which leave zero 20AARs (9). Exon 9-mutant patients were, however, significantly different from the other two groups of patients. Although not all nt4 661insA mutations could be assigned to a germline allele, if we made the reasonable assumption that all of these mutations were on the germline mutant allele, over half of all tumours from exon 9-mutant patients had 'third hits' (Supplementary Table 4.1.2). These differences could not readily be explained by features such as the size or dysplasia of the tumours analysed, which did not differ significantly among the three patient groups (details not shown).

Mechanism of LOH

We tested the possibility that different LOH events (for example, those involving the germline wild type and mutant alleles) were caused by different mechanisms, such as mitotic recombination and deletion, which resulted in different gene dosages and functional consequences. However, none of 17 tumours with allelic loss (10 with mutant LOH and 7 with wild-type LOH) showed copy number changes in the *APC* promoter region or exons using MLPA analysis. We selected for RQM-PCR analysis 10 further tumours (2 with mutant LOH and 8 with wild-type LOH) with mean LOH ratios below 0.3 (indicating that contamination with normal tissue was low enough not to confound the detection of deletion, (10) but, all adenomas showed copy number values between 0.79 and 0.97, consistent with diploid *APC* copy number and LOH by mitotic recombination.

Early pathways of tumorigenesis in AFAP polyps with 'three hits'

Our data, combined with previous findings (1, 14), showed that a substantial proportion of AFAP adenomas have acquired two somatic *APC* changes, one targeting the germline wildtype and one

the germline mutant allele. Consideration of the order in which these somatic changes occur and their respective effects on tumour growth and has important implications for determining the molecular genetic mechanism underlying AFAP. In AFAP adenomas, initiation of tumour growth might require all 'three hits' to be present in the tumour cell of origin ('kick-start' model). In this case, the two somatic mutations could occur in either order without functional consequence. The 'kick-start' model implies that a mechanism exists which results in an increase of the *intrinsic* or *effective* mutation rate in order to explain the relatively high frequency of such tumours as compared to the general population. Alternatively, the 'second hit' - necessarily involving the germline wild-type allele - might be sufficient for early adenoma growth, with the 'third hit' (involving the germline mutant allele) being required for subsequent tumorigenesis prior to clinical presentation. This 'stepwise' model postulates that mutation of the germline wild-type allele induces limited clonal expansion (thereby increasing the *effective* mutation rate), and is followed by mutation of the germline mutant allele to give an optimal *APC* genotype.

APC mutation data from individual adenomas can be used to distinguish between these possibilities, because the 'kick-start' and 'stepwise' models are expected to leave distinct footprints as regards the proportion(s) of somatic mutant allele(s), since these proportions depend on the order in which the somatic changes have occurred and some residual adenoma with 'two hits' is expected in the 'stepwise' case (Figure 4.1. 2).

Consider, for example, polyps with loss of the germline wild-type allele and a somatic insertion/deletion mutation on the germline mutant allele. We can measure two ratios of relative allelic dosage, one for the germline mutation and the other for the truncating somatic mutation, and use these to estimate the proportion of each allelotype in the tumour. Furthermore, we can calculate the expected values of these ratios by predicting the proportion of somatic mutant allele expected in the tumour under different models of tumorigenesis (Figure 4.1. 2). By comparing the observed proportion of the somatic mutant allele with that expected, we can determined whether the 'stepwise' or 'kickstart' model fits better (see Figure 4.1. 2 for details). Similarly, observed and expected allele proportions can be determined for adenomas with one somatic insertion/deletion mutation on the germline wildtype allele and loss of the germline mutant allele (Figure 4.1. 2). For tumours with two truncating somatic mutations, the expected ratio of the two mutant alleles under each model can be compared to the ratio measured directly by cloning a PCR product

Table 4.1.3: Observed and expected frequencies of somatic mutant *APC* alleles in AFAP polyps with 'three hits' for 'kick-start' and 'stepwise' models of tumorigenesis. α_{gl} = proportion of germline wild-type allele in polyp; β_{gl} = proportion of germline mutant allele in polyp; β_{som} = proportion of somatic mutant allele in polyp. Observed α_{gl} β_{gl} frequencies were determined from LOH ratios. Observed β_{som} frequencies were similarly determined from LOH ratios generated by PCR amplification of a region encompassing the somatic insertion/deletion and subsequent Genescan analysis (using constitutional DNA from patients with germline mutations identical to the somatic change for normalisation). The observed 'third' to 'second hit' ratio for polyp was determined by sequencing 58 clones of a PCR product encompassing both somatic changes.

A. Polyp with one somatic insertion/deletion and loss of the germline wild-type allele.

Model	Allele frequencies	Polyp 1	Polyp 2	Polyp 3	Polyp 4	Polyp 5	Polyp 6	Polyp 7	X^2_6
	Observed α_{gl}	0.21	0.27	0.21	0.29	0.30	0.26	0.28	
	Observed β_{som}	0.64	0.52	0.49	0.52	0.33	0.40	0.27	
'Kick-start', LOH first	Expected $\beta_{som}=[(1-2\alpha_{gl})/2]$	0.29	0.23	0.29	0.21	0.20	0.24	0.22	1.59
'Kick-start', LOH second	Expected $\beta_{som}=(1-2\alpha_{gl})$	0.58	0.46	0.58	0.42	0.40	0.48	0.44	0.14
'Stepwise'	Expected minimum β_{som}*	0.20	0.20	0.20	0.20	0.20	0.20	0.20	2.72
	Expected maximum $\beta_{som}=[(1-2\alpha_{gl})/2]$	0.29	0.23	0.29	0.21	0.20	0.24	0.22	1.59

* Assuming that the somatic insertion/deletion can be detected if it comprises >20% of all alleles in the polyp.

B Polyp with one somatic insertion/deletion and loss of the germline mutant allele.

Model	Allele frequencies	Polyp 8	Polyp 9	Polyp 10	X^2_3
	Observed β_{gl}	0.33	0.10	0.24	
	Observed β_{som}	0.41	0.83	0.61	
'Kick-start' (LOH must be second)	Expected $\beta_{som}=(1-2\beta_{gl})$	0.34	0.80	0.52	0.03
'Stepwise'	Expected minimum $\beta_{som}=(1-2\beta_{gl})$	0.34	0.80	0.52	0.03
	Expected maximum $\beta_{som}=(1-\beta_{gl})$	0.67	0.90	0.76	0.14

C. Two truncating somatic mutations.

Model	'Third' to 'second hit' ratios	Polyp 11	X^2_1
	Observed	27:31	
'Kick-start'	Expected	1:1	0.03
'Stepwise'	Expected minimum	1:4	14.4
	Expected maximum	1:1	0.03

* Assuming a detection sensitivity of 20% for the somatic 'third hit'.

encompassing both changes, sequencing multiple clones and counting how many times each allele is represented (Figure 4.1. 2).

For seven tumours with loss of the germline wild-type allele and a somatic insertion/deletion mutation, the observed and expected proportions of the somatic mutant allele were very similar to those expected under the 'kick-start' model, assuming that the 'second hit' was the insertion/deletion (on the somatic mutant allele) and the 'third hit' was the allelic loss (Table 4.1.3). Similar results in favour of a 'kick-start' model were obtained for three adenomas with

one somatic insertion/deletion and loss of the germline mutant allele, and for one adenoma with two truncating somatic mutations (Table 4.1.3).

4.1.5 Discussion

Our analysis of a relatively large set of AFAP families has shown complexity in the phenotype and early genetic pathways of tumorigenesis. The two previous analyses of somatic *APC* mutations in AFAP each focussed on single families, one with a germline mutation in the 5' region of the gene (1) and the other with a mutation in exon 9 (14). These two studies unequivocally provided the important and original finding that 'three hits' - that is, two somatic mutations, including loss or mutation of the germline mutant allele - can occur in AFAP tumours. The restricted size of the two studies meant, however, that they were unable to provide further conclusions.

We have found that patients with germline *APC* mutations in the 5' and 3' regions of the gene or the alternatively spliced region of exon 9 have a highly variable large-bowel phenotype, in that the number of colorectal adenomas varies from almost none to the hundreds or thousands of lesions found in classical FAP (3). Although assessment methods necessarily differ among clinical centres, our analysis shows that patients with 5' *APC* mutations (codons 1-177) are likely to have a more severe phenotype phenotype than those with mutations in exon 9 or the 3' end of the gene (>codon 1580). Phenotypic severity also tends to be similar within families, suggesting that restricting analyses to single kindreds may not provide accurate assessment of AFAP patients.

Our study has confirmed that 'three hits' at *APC* often occur in AFAP adenomas. In such polyps, the 'third hit' appears to be required for the initiation of tumorigenesis. Although 'third hits' might occur at loci other than *APC,* we have previously found no mutations at beta-catenin in AFAP polyps (unpubl. data). In polyps with 'three hits' from exon 9-mutant and 3'-mutant patients, we have been able to identify specific combinations of *APC* mutations which tend to occur. Exon 9 is alternatively spliced in all normal and neoplastic tissues which we have examined (not shown). The combinations of *APC* mutations almost certainly produce a near-optimal level of Wnt signalling, comparable with those found in classical FAP (9). Some of the combinations – such as R332X-nt4661insA/LOH – strongly suggest that the tumour has developed as a result of the functional effects of the germline mutant allele, but other combinations of mutations – such as truncating mutation leaving one 20AAR on the wildtype

with LOH of the germline mutant – might simply be indicative of a 'sporadic' tumour occurring on the background of AFAP.

In our families, 'third hits' were much rarer in 5'- and 3'-mutant patients than in the exon 9 mutants. These former families' somatic mutations usually - but not always - resembled those of classical FAP patients who have germline mutations before the first 20AAR of the *APC* protein. In many ways, this is the result which would be predicted were the 5' or 3' mutations simply to cause absent or non-functional protein. 5' *APC* mutations probably produce a small amount of partially functional APC through use of an internal ribosome entry site (IRES) at codon 184 (18). 3'-mutant proteins have been reported as being unstable (19), although the reasons for this are unknown. It is entirely plausible that the levels of functional APC protein vary among individuals with both 5' and 3' mutations, for example as a result of modifier alleles. Thus, for an adenoma to form, some patients would tend to require 'third hits' and others would not. The family of Spirio *et al* (1), for example, may have been relatively efficient at use of the IRES. Formal testing of this hypothesis *in vivo* would require an exceptionally large, unselected series of tumours and patients.

Our analysis of exon 9-mutant cases further provides further evidence to show that not all AFAP patients are the same. 'Third hits' were common in these patients' tumours. There was a markedly increased frequency of mutations which left three 20AARs on the germline mutant allele, particularly – but not exclusively - at nt4661, which appears to be a relatively hypermutable site. Our view differs somewhat from that of Su *et al* (14), who proposed that insAnt4661 mutations were over-represented in AFAP polyps because both 'strong' and 'weak' mutations were sufficient to severely reduce function of the exon 9- mutant allele. We suggest that mutations leaving three 20AARs on the germline mutant allele are common because the resulting allelotype R332X-4661insA gives a near-optimal genotype, taking into account loss of the germline wildtype allele and alternative splicing of exon 9. Variation in splicing efficiency – again through modifier allele action - could explain phenotypic variability in exon 9-mutant AFAP, but it appears that many of these patients produce sufficient functional protein by splicing out exon 9 that 'third hits' are necessary in most polyps.

The reason why AFAP patients develop fewer polyps than classical FAP patients is evident, in that 'three hits' are often needed to produce the near-optimal genotype. We do not, however, claim that all polyps from patients with AFAP-associated *APC* mutations require 'three hits'. Even allowing for the imperfections of mutation screening and LOH analysis in archival specimens, we were able to analyse the fresh-frozen adenomas comprehensively and found many

without ‘three hits’. Moreover, several polyps from our patients had somatic mutations which would have been predicted from a ‘two hit’ model of optimal Wnt signalling. Currently, we cannot explain why in a single patient, some polyps seem to require ‘three hits’ and others do not, but it is possible that ‘third hits’ at other loci can substitute for *APC* mutation. Another possibility is that selective constraints on the diminished APC function needed for tumorigenesis are ‘just right’ (1, 10) at some times, but weaker at others, for example during development or when tissue is undergoing repair.

The genetic analysis of colorectal tumours from patients with germline mutations inAFAP-associated regions of *APC*, by this study and others, has revealed a novel mechanism underlying the genotype-phenotype association in this tumour syndrome, namely an requirement for ‘three hits’ in at least some AFAP adenomas. This finding mustbe viewed in the framework of the model of optimal combinations of *APC* mutations, ratherthan simple loss of protein function. More than one different combination of *APC* mutationscan provide near-optimal Wnt signalling in AFAP. However, not all AFAP patients are thesame. Given that assembling a very large series of AFAP patients is extremely difficult, it isnot easy to decide on what is the ‘typical’ AFAP phenotype or somatic genotype. In the seven families with 5’ *APC* mutations studied to date ((1, 10) and this study), about 15-20% of polyps seem to acquire ‘three hits’, but only Spirio *et al* (1) found a high frequency of nt4661insA. In the six 3’-mutant families studied (all from this study), the frequency of ‘third hits’ seems similar to that of the 5’-mutants. Six exon 9-mutant families have been studied (14) and this study) and almost all of these show evidence of a high frequency of ‘third hits’– we estimate a minimum of 50% in our study. In addition, there appear to be genetic factors apart from the germline *APC* mutation that influence disease severity, as evidenced by the tendency for polyp numbers to be similar within families. The phenotypic and somatic molecular heterogeneity in AFAP means that clinical management of patients with AFAP associated mutations must be empirical. Accurate prediction of phenotype may only be possible when factors, such as modifier genes, that influence genetic pathways and diseas eseverity are identified.

Acknowledgements

We are grateful to the Equipment Park, Cancer Research UK, to the patients taking part in the study, and to several collaborating Geneticists and Histopathologists without whose help the study could not have been performed.

Supplementary tables and figures

Supplementary Table 4.1.1. Somatic *APC* mutations and allelic loss in tumours from AFAP patients with 5' germline mutations. All tumours with mutation or LOH are shown. FS = frameshift; LOH = loss of heterozygosity; wt = germline wild-type allele; mut = germline mutant allele, where this assignment was possible

Patient ID	Germline mutation	Somatic mutation	Type of somatic change	20AARs in somatic mutant allele	LOH
DFAP48	141FS	1309FS	3927del5bp	1	NL
AVC.III.2	Q163X	1305FS	3914delC	1	NL
AVC.III.2	Q163X	1309FS	3927del5bp	1	NL
554.iii.2	170FS	Q1378X	4132C>T	1	NL
554.iii.2	170FS	1398FS	4192delAG	2	NL
554.iii.2	170FS	1439FS	4316delC	2	NL
554.iv.2	170FS	1462FS	4386delGA	2	NL
AFX MK	G126X	S1356X	4067C>G	1	LOH mut
AVC.III.2	Q163X	1309FS	3927del5bp	1	LOH mut

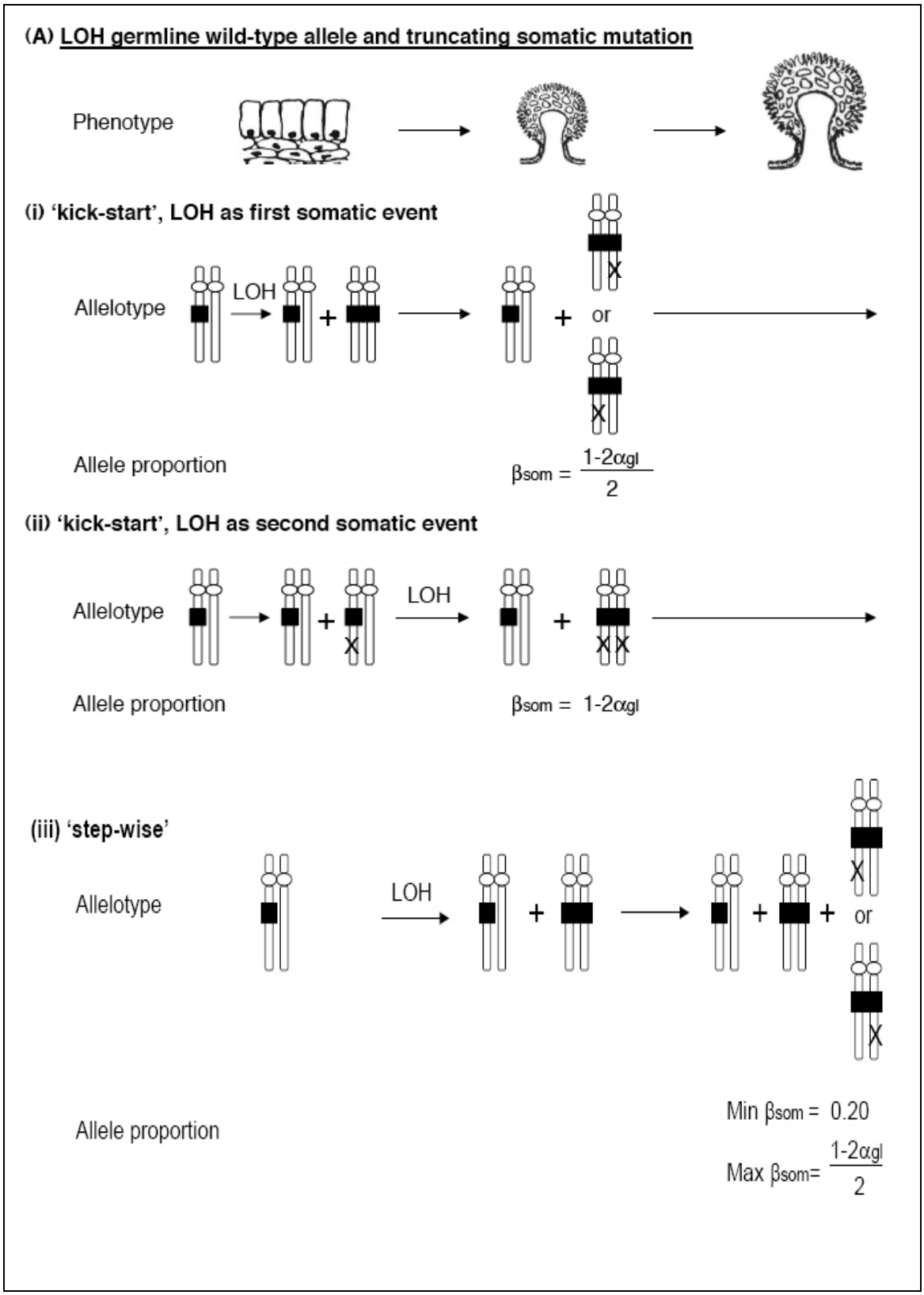
Supplementary Table 4.1.2. Somatic *APC* mutations and allelic loss in 79 tumours from AFAP patients with germline mutations in the alternatively spliced region of exon 9.
See Supplementary Table 4.1.1 for abbreviations.

Patient ID	Somatic mutation(s)	Type(s) of somatic change(s)	20AARs in somatic mutant allele(s)	LOH
673.iii.3	K792X	2374A>G	0	NL
578.iv.4	1335FS	4004del31bp	1	NL
1571.ii.2	1462FS	4386delGA	2	NL
578.iv.1	1505FS	4514dup7bp	2	NL
578.iv.1	K1469X	4405C>T	2	NL
578.iv.4	1518FS wt	4552delA	3	NL
578.iv.7	E1265X ; 1462FS	3793G>T; 4386del4bp	0 ; 2	NL
578.FPL	1462FS mut	4386delGA	2	NL
578.iv.7	1462FS mut	4386delGA	2	NL
578.iii.9	1372FS ; 1554FS	4114delG; 4661insA	1 ; 3	NL
578.iv.4	1372FS ; 1554FS	4117delC; 4661insA	1 ; 3	NL
578.AA	1554FS mut	4661insA	3	NL
578.FPL	1554FS mut	4661insA	3	NL
578.FPL	1554FS mut	4661insA	3	NL
578.FPL	1554FS mut	4661insA	3	NL
578.iii.9	1554FS mut	4661insA	3	NL
578.iii.9	1554FS mut	4661insA	3	NL
578.iii.9	1554FS mut	4661insA	3	NL
578.iii.9	1554FS mut	4661insA	3	NL
578.iii.9	1554FS mut	4661insA	3	NL
578.iv.4	1530FS mut	4589insA	3	NL
578.iv.4	1554FS mut	4661insA	3	NL
578.iv.4	1554FS mut	4661insA	3	NL
578.iv.4	1554FS mut	4661insA	3	NL
578.iv.7	1554FS mut	4661insA	3	NL
578.iv.7	1554FS mut	4661insA	3	NL
578.iv.7	1554FS mut	4661insA	3	NL
578.iv.7	1554FS mut	4661insA	3	NL
1571.ii.2	1554FS	4661insA	3	NL
1571.ii.2	1537FS	4611delAG	3	NL
578.iv.1	1554FS	4661insA	3	NL
578.iv.1	1554FS	4661insA	3	NL
578.iv.1	1554FS	4661insA	3	NL
578.iv.1	1554FS	4661insA	3	NL
578.iv.1	1554FS	4661insA	3	NL
578.iv.1	1554FS	4661insA	3	NL
578.iv.4	1554FS	4661insA	3	NL
DFAP81	1554FS	4661insA	3	NL
578.iv.1	none detected			LOH wt
578.iv.1	none detected			LOH wt
578.iv.1	none detected			LOH wt
578.iii.9	1554FS mut	4661insA	3	LOH wt
578.iii.9	1554FS mut	4661insA	3	LOH wt
578.iv.7	1554FS mut	4661insA	3	LOH wt
578.iv.7	1554FS mut	4661insA	3	LOH wt
578.iv.1	1554FS mut	4661insA	3	LOH wt
578.iv.7	1554FS mut	4661insA	3	LOH wt
578.FPL	1394FS	4182delTA	1	LOH mut
578.iii.9	E1286X	3856G>T	1	LOH mut
578.iii.9	S1315X	3944C>A	1	LOH mut

Supplementary Table 4.1.3. Somatic *APC* mutations and allelic loss in 126 adenomas from AFAP patients with 3' germline mutations. See Supplementary Table 4.1.1 for abbreviations.

Patient ID	Germline mutation	Somatic mutation(s)	Type(s) of somatic change(s)	20AARs in somatic mutant allele(s)	LOH
MD2976	1919FS	R283X	847C>T	0	NL
MD2976	1919FS	R283X	847C>T	0	NL
MD2976	1919FS	R283X	847C>T	0	NL
77-11	1943FS	E1374X	4120G>T	1	NL
DW20284	2078FS	Q1338X	4012C>T	1	NL
J42424	2078FS	1300FS	3902insC	1	NL
L12562	2078FS	Q1338X	4012C>T	1	NL
110.2.vi	2078FS	1354FS	4061delTT	1	NL
01/266	1738FS	1441FS	4323delA	2	NL
01/266	1738FS	1414FS	4242insT	2	NL
01/266	1738FS	1424FS	4270delC	2	NL
01/266	1738FS	E1408X	4222G>T	2	NL
01/266	1738FS	R1435X	4304A>T	2	NL
01/266	1738FS	R1435X	4304A>T	2	NL
2233/3	1919FS	1398FS	4192delAG	2	NL
2233/3	1919FS	1407FS	4219delAG	2	NL
2233/3	1919FS	R1435X	4303A>T	2	NL
77-11	1943FS	1490FS	4468insG	2	NL
77-11	1943FS	Q1406X	4216C>T	2	NL
110.2.vi	2078FS	1403FS	4209insCT	2	NL
1460/6	1982 FS	1537FS	4611delAG	3	NL
1460/6	1982 FS	1462FS mut	4386del4bp	2	NL
110.2.vi	2078FS	1485FS mut	4455delT	2	NL
MD2976	1919FS	1129FS wt; 1554FS mut	3387delT; 4661insA	0 ; 3	NL
344-44	1597FS	1554FS	4661insA	3	NL
01/266	1738FS	1576FS	4726delG	3	NL
77-11	1943FS	1570FS	4709del6bp	3	NL
2233/3	1919FS	none detected			LOH wt
2233/3	1919FS	none detected			LOH wt
2233/3	1919FS	none detected			LOH wt
77-11	1943FS	none detected			LOH wt
77-11	1943FS	none detected			LOH wt
77-11	1943FS	none detected			LOH wt
77-11	1943FS	none detected			LOH wt
77-12	1943FS	none detected			LOH wt
77-12	1943FS	none detected			LOH wt
77-40	1943FS	none detected			LOH wt
1460/6	1982 FS	none detected			LOH wt
1460/88	1982 FS	none detected			LOH wt
1624/04	1982 FS	none detected			LOH wt
77-11	1943FS	1462FS mut	4386delGA	2	LOH wt
77-11	1943FS	1429FS mut	4286delA	2	LOH wt
77-11	1943FS	1234FS	3701ins8bp	0	LOH mut
77-11	1943FS	none detected			LOH mut
77-11	1943FS	none detected			LOH mut
1460/6	1982 FS	none detected			LOH mut
1460/42	1982 FS	none detected			LOH mut
1460/42	1982 FS	none detected			LOH mut
1460/42	1982 FS	none detected			LOH mut
1624/04	1982 FS	none detected			LOH mut
J42424	2078FS	none detected			LOH mut
J42424	2078FS	none detected			LOH mut
1460/42	1982 FS	1357FS	4069delGG	1	LOH mut
1489/10	1982 FS	1354FS	4060delAA	1	LOH mut
1624/04	1982 FS	E1322X	3964G>T	1	LOH mut
01/266	1738FS	1431FS	4294delC	2	LOH mut
MD2976	1919FS	1493FS	4479delG	2	LOH mut

§


(A) LOH germline wild-type allele and truncating somatic mutation
Phenotype
(i) 'kick-start', LOH as first somatic event
Allelotype
LOH
or
Allele proportion
βsom = 1-2αgl / 2
(ii) 'kick-start', LOH as second somatic event
Allelotype
LOH
Allele proportion
βsom = 1-2αgl
(iii) 'step-wise'
Allelotype
LOH
or
Allele proportion
Min βsom = 0.20
Max βsom= 1-2αgl / 2

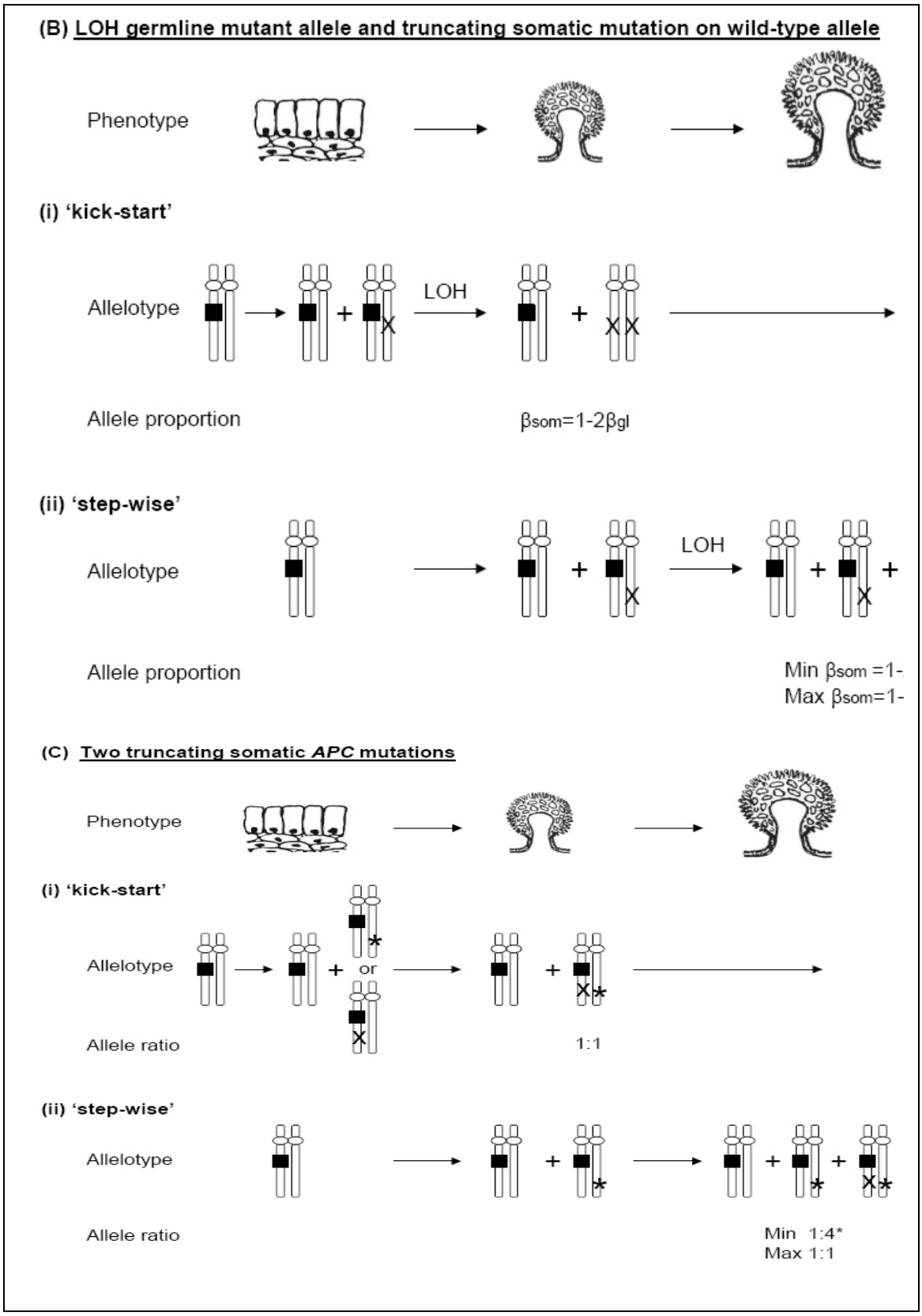

(B) LOH germline mutant allele and truncating somatic mutation on wild-type allele
Phenotype
(i) 'kick-start'
Allelotype
LOH
Allele proportion
βsom=1-2βgl
(ii) 'step-wise'
Allelotype
LOH
Allele proportion
Min βsom =1-
Max βsom=1-
(C) Two truncating somatic APC mutations
Phenotype
(i) 'kick-start'
Allelotype
or
Allele ratio
1:1
(ii) 'step-wise'
Allelotype
Allele ratio
Min 1:4*
Max 1:1

Figure 4.1. 2. Pathways of tumorigenesis in AFAP polyps with 'three hits'.

The Figure 4.1.2. illustrates the possible sequences in which somatic mutations/allelic loss may occur in AFAP polyps with 'three hits', as well as the possible functional effects of these changes. The expected proportions β_{gl}, β_{som} and β_{gl} are shown. ■= germline mutation; X, * = truncating somatic mutation; LOH = loss of heterozygosity by mitotic recombination; β_{som} = proportion of somatic mutant allele in polyp; α_{gl} = proportion of germline wild-type allele in polyp; β_{gl} = proportion of germline mutant allele in polyp **(A)** Loss of the germline wild-type allele and truncating somatic mutation In a 'kick-start' model these changes can occur in either order (i) or (ii) and tumour growth ensues once both somatic changes have occurred; in a 'step-wise' model loss of the germline wild-type allele leads to limited clonal expansion and is followed by the truncating somatic mutation which promotes further tumour growth. **(B)** Loss of the germline mutant allele and truncating somatic mutation In both the 'kick-start' (i) and the 'step-wise' (ii) model the truncating somatic mutation precedes loss of the germline mutant allele, but in the 'kick-start' model, both changes are required for tumour growth. **(C)** Two truncating somatic mutations In a 'kick-start' model (i), these changes can occur in either order and tumour growth ensues once both somatic changes have occurred; in a 'step-wise' model (ii), somatic mutation of the germline wild-type allele causes limited clonal expansion and is followed by somatic mutation of the germline mutant allele which promotes further tumour growth. For each model of scenarios **(A)** and **(B),** the expected proportion of the somatic mutant allele (β_{som}) in the polyp can be determined from the proportions of the germline wild-type (α_{gl}) or mutant (β_{gl}) allele as shown. α_{gl} and β_{gl} can be estimated from the LOH ratio. For scenario (C) the expected ratio of the two somatic alleles is 1:1 for the 'kick-start model', but lies between 1:4 and 1:1 for the 'step-wise' model with the minimum estimate (*) assuming a mutation detection sensitivity of 20%.

4.1.6 Reference

1. Spirio LN, Samowitz W, Robertson J, Robertson M, Burt RW, Leppert M, et al. Alleles of APC modulate the frequency and classes of mutations that lead to colon polyps. Nat Genet 1998;20(4):385-8.

2. Hernegger GS, Moore HG, Guillem JG. Attenuated familial adenomatous polyposis: an evolving and poorly understood entity. Dis Colon Rectum 2002;45(1):127-34; discussion 134-6.

3. Knudsen AL, Bisgaard ML, Bulow S. Attenuated familial adenomatous polyposis (AFAP): a review of the literature. Fam Cancer 2003;2(1):43-55.

4. Rozen P, Samuel Z, Shomrat R, Legum C. Notable intrafamilial phenotypic variability in a kindred with familial adenomatous polyposis and an APC mutation in exon 9. Gut 1999;45(6):829-33.

5. Eccles DM, van der Luijt R, Breukel C, Bullman H, Bunyan D, Fisher A, et al. Hereditary desmoid disease due to a frameshift mutation at codon 1924 of the APC gene. Am J Hum Genet 1996;59(6):1193-201.

6. Scott RJ, Froggatt NJ, Trembath RC, Evans DG, Hodgson SV, Maher ER. Familial infiltrative fibromatosis (desmoid tumours) (MIM135290) caused by a recurrent 3' APC gene mutation. Hum Mol Genet 1996;5(12):1921-4.

7. Young J, Simms LA, Tarish J, Buttenshaw R, Knight N, Anderson GJ, et al. A family with attenuated familial adenomatous polyposis due to a mutation in the alternatively spliced region of APC exon 9. Hum Mutat 1998;11(6):450-5.

8. Soravia C, Berk T, Madlensky L, Mitri A, Cheng H, Gallinger S, et al. Genotypephenotype correlations in attenuated adenomatous polyposis coli. Am J Hum Genet 1998;62(6):1290-301.

9. Lamlum H, Ilyas M, Rowan A, Clark S, Johnson V, Bell J, et al. The type of somatic mutation at APC in familial adenomatous polyposis is determined by the site of the germline mutation: a new facet to Knudson's 'two-hit' hypothesis. Nat Med 1999;5(9):1071-5.

10. Albuquerque C, Breukel C, van der Luijt R, Fidalgo P, Lage P, Slors FJ, et al. The 'just-right' signaling model: APC somatic mutations are selected based on a specific level of activation of the beta-catenin signaling cascade. Hum Mol Genet 2002;11(13):1549-60.

11. Crabtree M, Sieber OM, Lipton L, Hodgson SV, Lamlum H, Thomas HJ, et al. Refining the relation between 'first hits' and 'second hits' at the APC locus: the 'loose fit' model and evidence for differences in somatic mutation spectra among patients. Oncogene 2003;22(27):4257-65.

12. Sieber OM, Heinimann K, Gorman P, Lamlum H, Crabtree M, Simpson CA, et al. Analysis of chromosomal instability in human colorectal adenomas with two mutational hits at APC. Proc Natl Acad Sci U S A 2002;99(26):16910-5.
13. Rowan AJ, Lamlum H, Ilyas M, Wheeler J, Straub J, Papadopoulou A, et al. APC mutations in sporadic colorectal tumors: A mutational "hotspot" and interdependence of the "two hits". Proc Natl Acad Sci U S A 2000;97(7):3352-7.
14. Su LK, Barnes CJ, Yao W, Qi Y, Lynch PM, Steinbach G. Inactivation of germline mutant APC alleles by attenuated somatic mutations: a molecular genetic mechanism for attenuated familial adenomatous polyposis. Am J Hum Genet 2000;67(3):582-90.
15. Fearnhead NS, Britton MP, Bodmer WF. The ABC of APC. Hum Mol Genet 2001;7:721-33.
16. Sieber OM, Lamlum H, Crabtree MD, Rowan AJ, Barclay E, Lipton L, et al. Wholegene APC deletions cause classical familial adenomatous polyposis, but not attenuated polyposis or "multiple" colorectal adenomas. Proc Natl Acad Sci U S A 2002;99(5):2954-8.
17. Crabtree MD, Tomlinson IP, Hodgson SV, Neale K, Phillips RK, Houlston RS. Explaining variation in familial adenomatous polyposis: relationship between genotype and phenotype and evidence for modifier genes. Gut 2002;51(3):420-3.
18. Heppner Goss K, Trzepacz C, Tuohy TM, Groden J. Attenuated APC alleles produce functional protein from internal translation initiation. Proc Natl Acad Sci U S A 2002;99(12):8161-6.
19. van der Luijt RB, Meera Khan P, Vasen HF, Breukel C, Tops CM, Scott RJ, et al. Germline mutations in the 3' part of APC exon 15 do not result in truncated proteins and are associated with attenuated adenomatous polyposis coli. Hum Genet 1996;98(6):727-34.

CHAPTER 4.2

4.2 Prevalence of MYH germline mutations in Swiss APC mutation-negative polyposis patients

(Russell A, Zhang J, Luz J, Hutter P, Chappuis P, Berthod C, Maillet P, Mueller H, Heinimann K. (published in Int. J. Cancer: 118, 1937–1940 2006) (Equally contributed)

4.2.1 Abstract

In 10-30% of patients with classical familial adenomatous polyposis (FAP) and up to 90% of those with attenuated (<100 colorectal adenomas; AFAP) polyposis, no pathogenic germline mutation in the adenomatous polyposis coli (APC) gene can be identified (APC mutation-negative). Recently, biallelic mutations in the base excision repair gene MYH have been shown to predispose to a multiple adenoma and carcinoma phenotype. This study aimed to (i) assess the MYH mutation carrier frequency among Swiss APC mutation-negative patients and (ii) identify phenotypic differences between MYH mutation carriers and APC/MYH mutation-negative polyposis patients. Seventy-nine unrelated APC mutation-negative Swiss patients with either classical (n = 18) or attenuated (n = 61) polyposis were screened for germline mutations in MYH by dHPLC and direct genomic DNA sequencing. Overall, 7 (8.9%) biallelic and 9 (11.4%) monoallelic MYH germline mutation carriers were identified. Among patients with a family history compatible with autosomal recessive inheritance (n = 45), 1 (10.0%) out of 10 classical polyposis and 6 (17.1%) out of 35 attenuated polyposis patients carried biallelic MYH alterations, 2 of which represent novel gene variants (p.R171Q and p.R231H). Colorectal cancer was significantly ($p < 0.007$) more frequent in biallelic mutation carriers (71.4%) compared with that of monoallelic and MYH mutation-negative polyposis patients (0 and 13.8%, respectively). On the basis of our findings and earlier reports, MYH mutation screening should be considered if all of the following criteria are fulfilled: (i) presence of classical or attenuated polyposis coli, (ii) absence of a pathogenic APC mutation, and (iii) a family history compatible with an autosomal recessive mode of inheritance.

4.2.2 Introduction

Familial adenomatous polyposis (FAP; OMIM entry no.175100) is an autosomal dominantly inherited colorectal cancer (CRC) predisposition caused by germline mutations in the adenomatous polyposis coli (*APC*) gene and characterized by the development of hundreds to

thousands of adenomatous polyps throughout the intestinal tract.[1] Attenuated FAP (AFAP) represents a clinical variant of classical FAP, associated with multiple (<100) colorectal adenomas and caused by mutations in the most 5′ or 3′ regions of *APC* or in the alternatively spliced region of exon 9(1,2,3). With routine screening techniques failing to detect pathogenic *APC* germline mutations in 10-30% of classical FAP patients and in up to 90% of AFAP patients,(4) investigations about the role of other polyposis predisposition genes are topical.

Recently, Al Tassan *et al.* demonstrated that biallelic germline mutations in the human homologue of the base excision repair gene *MutY* (*MYH*) cause a phenotype of multiple colorectal adenomas and carcinomas, thus, describing for the first time an autosomal recessively inherited CRC predisposition(5,6). The DNA glycosylase MYH removes adenines from mispairs with 8-oxoguanine that occur during replication of oxidized DNA. Failure to correct these mispairs consequently leads to G:C⟶T:A transversion mutations, a typical "footprint" of oxidative DNA damage.(7)The observation of an excess of transversion mutations in tumors eventually led to the discovery of *MYH*-associated polyposis (MAP). A number of studies have already been conducted in attempts to establish the extent to which germline mutations in the *MYH* gene may contribute to individuals with an AFAP phenotype(6)(8)(9). As a result, biallelic *MYH* germline mutations have been attributed to ~1-3% of all unselected CRC patients. This nation-wide study aimed to (*i*) assess the frequency of *MYH* mutation carriers in 79 unrelated Swiss patients presenting with either classical or attenuated polyposis and in whom no pathogenic *APC* germline mutation could be identified and (*ii*) to identify phenotypic differences between biallelic mutation carriers, monoallelic mutation carriers and *APC/MYH* mutation-negative patients.

4.2.3 Patients and Methods

Patients

This nation-wide study investigated 79 ostensibly unrelated Swiss index patients referred between 1994 and 2004 to either the Research Group Human Genetics, Division of Medical Genetics, Basel, or the Unit of Genetics, Institut Central des Hôpitaux Valaisans, Sion, Switzerland, because of classical (≥100 polyps, n = 18) or multiple adenomas/attenuated (5-99 polyps) FAP (AFAP; n = 61). In all patients, no germline *APC* mutation could be identified by means of the protein truncation test, single strand conformation polymorphism or direct DNA sequencing (patients thereafter referred to as *APC* mutation-negative).[10] Forty-five patients displayed a family history compatible with autosomal recessive inheritance; in the remainder there was either evidence for vertical transmission or no detailed family history available. In

addition, 100 control Swiss individuals were enrolled so as to establish the carrier frequency of previously reported *MYH* variants as well as novel mutations of unknown pathogenic significance in unaffected individuals. Informed consent for the study was obtained from all individuals investigated. Patients were considered as anonymous cases, and the results of the various genetic analyses were independently checked by at least 2 assessors.

DNA extraction

Genomic DNA was isolated from EDTA blood, using methods previously described by Miller *et al* (11). Briefly, 10 ml blood was mixed with 30 ml of EL buffer (155 mM NH_4Cl, 10 mM $KHCO_3$, 1 mM EDTA, pH 7.4) and left on ice for 15 min. The lysate was centrifuged at 2000*g* for 10 min, washed twice with EL buffer and the intact leukocyte pellet resuspended in NL buffer (10 mM Tris.HCl, pH 8.2, 400 mM NaCl, 2 mM Na_2EDTA, 1% SDS and 200 µg/ml proteinase K) and incubated overnight at 37ºC. The next day, 1 ml of 6 M NaCl was added and vigorously shaken, followed by centrifugation to remove cellular proteins. The supernatant containing the DNA was placed in a fresh tube and the DNA precipitated with ethanol. The resulting DNA pellet was washed with 70% ethanol, dried briefly and then suspended in 1 ml of TE buffer (10 mM Tris.HCl, pH 7.5, 0.1 M EDTA).

MYH mutation analysis

In 57 (72%) patients (15 FAP and 42 AFAP), the entire *MYH* coding sequence was analyzed by direct DNA sequencing. An additional 22 patients (3 FAP and 19 AFAP) were exclusively screened for mutations in exons 7 and 13 in which the most common pathogenic mutations in the Caucasian population, p.Y165C and p.G382D, occur. Each time a heterozygous *MYH* mutation was identified, the entire gene was subsequently analyzed by direct DNA sequencing (exons 2, 5, 8 and 12) and dHPLC (exons 1, 3, 4, 6, 9, 10, 11, 14, 15, 16) to identify/exclude the presence of a second germline mutation.Exon specific primer pairs were used to amplify the 16 exons of *MYH* (HUGO ID: MUTYH; Genbank accession no. NM_012222), including the respective exon-intron boundaries (primer sequences and PCR conditions available from the authors upon request). Twenty-five microliters of PCR reaction mixture contained 50 ng of genomic DNA, 10 pmol of each primer and a PCR mastermix at 1.5 mM $MgCl_2$, according to the manufacturer's instructions (Invitrogen, Basel, Switzerland). All PCR reactions were done on a Hybaid OmnE thermocycler (Catalys AG, Wallisellen, Switzerland).

dHPLC

As a prescreening method to detect DNA sequence changes, dHPLC was performed using the 3500HT WAVE nucleic acid fragment analysis system (Transgenomic, Crewe, UK). Melting temperatures for dHPLC were predicted by the Wavemaker software version 4.1.42 (Transgenomic) (dHPLC melting temperatures available from the authors upon request). Where different elution profiles were observed in comparison with the control samples run in parallel, direct DNA sequencing was performed to characterize the nature of the sequence alteration.

Direct DNA sequencing

For DNA sequencing, PCR products were purified using the QIAquick PCR Purification kit (Qiagen, Basel, Switzerland). The sequencing reaction was performed using the Big Dye Teminator Cycle Sequencing kit (Applied Biosystems, Rotkreuz, Switzerland), according to the manufacturers' guidelines. Following purification using the DyeEx 2.0 Spin Kit (Qiagen, Basel, Switzerland), sequencing products were analyzed on an ABI PRISM 310 Genetic Analyser (Applied Biosystems). Germline mutations identified in *MYH* were confirmed by sequencing in both, forward and reverse, directions, and from at least 2 independent PCR products. Germline mutations p.Y165C and p.G382D were independently confirmed by restriction enzyme digests, using *Ila*I and *Bgl*II, respectively.

Statistical analysis

Statistical comparison of patients' features, encompassing phenotypic characteristics (gender, age at diagnosis, polyp number, extracolonic manifestations, family history) and mutational status was performed using the χ^2 and Fisher's exact test for categorical variables, or Student's *t*-test for continuous variables, with all of the probabilities reported as two-tailed ps, considering a *p* value of <0.05 to be statistically significant.

4.2.4 Results

Seventy-nine *APC* mutation-negative Swiss polyposis patients from the Basel ($n = 58$) and Sion ($n = 21$) medical genetic centers were investigated for the presence of *MYH* germline alterations. Twenty-three percent of the individuals were referred because of suspected classical FAP ($n = 18$), whilst the majority exhibited an attenuated or multiple adenoma phenotype ($n = 61$).

Table 4.2.1. Phenotypic Features and Germline Mutations Identified in *MYH* Mutation Carriers[1]

						MYH	
Patient ID	**Sex**	**Age**	**Polyp No.**	**CRC**	**Extracolonic disease**	**1st Mutation**	**2nd Mutation**
Biallelic *MYH* mutation carriers							
1775/01	M	38	<100	Yes	Yes	p.G84fs	p.W138_M139ins IW
1828/01	F	42	<100	Yes	No	p.Y165C	p.Y165C
1859/01	M	33	<100	No	No	p.Y165C	p.Y165C
2013/01	M	50	<100	Yes	No	p.G382D	p.G382D
2073/01	F	60	≈ 50	No	No	p.Y165C	p.R171Q
2184/01	M	48	>100	Yes	No	p.G382D	p.G382D
2185/01	M	48	≈ 74	Yes	No	p.Y165C	p.R231H
Monoallelic *MYH* mutation carriers							
1384/01	F	20	Multiple	No	Yes	p.G382D	None detected
1665/01	F	54	>100	No	No	p.I209V	None detected
2145/01	M	40	≈70	No	No	p.Y165C	None detected
2243/01	M	49	≈50	No	No	p.Y165C	None detected
2261/01	F	69	>100	No	No	p.Y165C	None detected
DFAP 17	F	34	≈ 20	No	Yes	p.G382D	None detected
DFAP 82	M	58	>100	No	No	p.G382D	None detected
DFAP 99	F	63	≈43	No	No	p.G382D	None detected
SA 453	M	41	5	No	No	p.G382D	None detected

[1] Patient 1775/01 has previously been reported by Sieber et al.(8)

MYH mutation analysis

The complete coding sequence of the *MYH* gene was investigated in 57 index patients. In addition, 22 patients were screened for alterations in exons 7 and 13, which harbor the most common pathogenic mutations, p.Y165C and p.G382D. Overall, 7 (8.9%) biallelic and 9 (11.4%) monoallelic *MYH* germline mutation carriers were identified. According to the clinical classification, 1 (5.6%) out of 18 FAP and 6 (9.8%) out of 61 AFAP patients harbored a biallelic *MYH* mutation. If only individuals with a family history compatible with autosomal recessive inheritance were considered ($n = 45$), 10.0% (1/10) of patients with classical polyposis and 17.1% (6/35) of AFAP patients harbored biallelic *MYH* germline mutations (Table 4.2.I).

In addition to the mutations p.Y165C and p.G382D, which accounted for 43% and 29% of mutant alleles in the biallelic patients, respectively, 2 novel alterations were detected in AFAP patients compound heterozygote for p.Y165C/p.R171Q and p.Y165C/p.R231H (Figs. 1*a* and 1*b*). One FAP patient, who was found to be a compound heterozygote with a p.G84fs/p.W138_M139insIW mutation, has been previously reported by Sieber *et.a l* (8). The healthy parents of this individual were available for investigation and were found to be heterozygous carriers of the p. W138_M139insIW and the p.G84fs alteration, respectively. Although the pathogenicity of p.R171Q and p.R231H remains to be established by functional studies, such gene alterations were not observed in 200 chromosomes from Swiss control samples. Furthermore, the amino acid positions are evolutionary highly conserved across distantly related species (*E. coli, S. pombe*, mouse, rat and human).

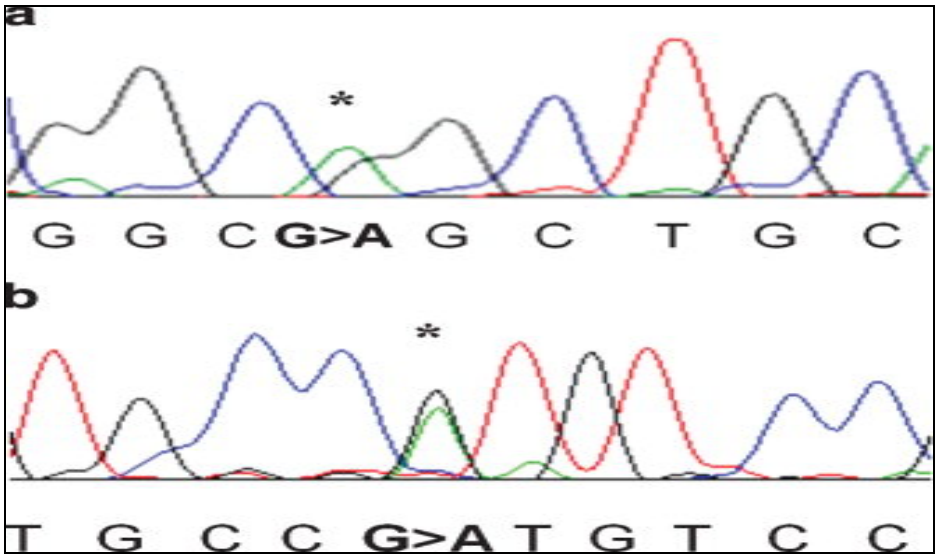


Figure 4.2.1. Sequencing chromatograms displaying the 2 novel MYH germline variants marked by an asterisk (*). (*a*) c.512G>A, p.RI7IQ (heterozygote); (*b*) c.693G>A, p.R231H (heterozygote).

Nine patients were identified as monoallelic *MYH* mutation carriers, with the p.G382D mutation present in 5 (56%) of them (Table 4.2.1). In the remaining 63 (80%) patients, no pathogenic *MYH* mutations could be identified. The previously described polymorphisms in exon 2 (c.64G > A; p.V22M) and exon 12 (c.972G > C; p.Q324H) were detected with allele frequencies of 6% and 17%, respectively, similar to that of a Swiss control sample population (200 chromosomes) assessed in parallel (2% p.V22M and 12% p.Q324H).

Genotype-phenotype comparisons

The phenotypic features of the 7 biallelic *MYH* mutation carriers are depicted in Table 4.2.1, with one of them displaying classical FAP. In 5 (71%) patients, CRC had been diagnosed at a median age of 48 years (IQR 10.5, range 33-60 years), with 3 of them located at proximal to the splenic flexure. The family history in all biallelic mutation carriers corresponded to an autosomal recessive mode of inheritance. Remarkably, in 3 out of 11 siblings of patient 2073/01 (p.Y165C/p.R171Q) a CRC had been diagnosed at a median age of 51 years (range 49-54). Except for patient 1775, in whom duodenal adenomas had been detected, no apparent extracolonic disease manifestations were observed in the other biallelic mutation carriers.

Among the 9 monoallelic *MYH* mutation carriers (Tables 4.2.1 and 4.2.2), 4 patients (no. 1384/01, 2243, DFAP17 and DFAP 82) had siblings with either CRC or polyps reported. With respect to extracolonic disease manifestations, a facial lipoma was observed in patient DFAP17 and a duodenal adenocarcinoma at age 20 in patient 1384/01.

Table 4.2.2. Phenotypic Characteristics of Biallelic *MYH* Mutation Carriers, Monoallelic Mutation Carriersand *APC*/*MYH* Mutation-Negative Patients with a Family History Compatible with autosomal Recessive Inheritance

	Biallelic *MYH* mutation carriers (*n* = 7)	**Monoallelic *MYH* mutation carriers (*n* = 9)**	***MYH* mutation-negative patients (*n* = 29)**
Sex			
Male	5 (71)[1]	5 (56)	18 (62)
Female	2 (29)	4 (44)	11 (38)
Clinical classification			
	1 (14)	3 (33)	6 (21)
FAP (≥100 polyps)			
AFAP (<100 polyps)	6 (86)	6 (67)	23 (79)
Age at diagnosis (yr)			
Median	48	49	48
IQR	10.5	20.8	20
Range	33-60	20-69	22-77
Colorectal cancer			
Present	5 (71)	0	4 (14)
Absent	2 (29)	9 (100)	25 (86)
Extracolonic disease			
Present	1 (14)	2 (22)	4 (14)
Absent	6 (86)	7 (78)	25 (86)

FAP:familial adenomatous polyposis; AFAP, attenuated FAP.

[1] Values given in parentheses indicate percentages.

Twenty-nine (46%) out of 63 *MYH* mutation-negative patients had a family history of CRC and/or multiple polyps/polyposis compatible with an autosomal recessive mode of inheritance and could, therefore, be included in the genotype-phenotype analysis (Table 4.2.II). Comparing the phenotypic properties of biallelic *MYH* mutation carriers, monoallelic mutation carriers and *APC/MYH* mutation-negative polyposis patients, colorectal cancer was significantly more frequent in biallelic mutation carriers than in the other subgroups (71% *vs*. 0% and 14%, respectively; χ^2 14.5, $p < 0.001$). Median age at diagnosis was similar between the 3 subgroups (48, 49 and 48 years, respectively). No further statistically significant phenotypic differences with respect to polyp number, age at diagnosis or extracolonic disease was observed.

4.2.5 Discussion

In this nation-wide survey on 79 Swiss *APC* mutation-negative polyposis patients, 9% were found to harbor biallelic ($n = 7$) and 11% monoallelic ($n = 9$) germline mutations in the base excision repair gene *MYH*. Considering only patients with a family history compatible with autosomal recessive inheritance, biallelic *MYH* mutation carriers were observed in 10% (1/10) of patients with classical and in 17% (6/35) of those with attenuated polyposis, respectively. No *MYH* alterations were identified in patients exhibiting a family history suggestive of an autosomal dominant inheritance pattern.

In addition to the most common pathogenic missense mutations, p.Y165C and p.G382D(5)(6)(8)(12). 2 novel alterations in the *MYH* gene p.R171Q and p.R231H were detected. Two hundred control chromosomes, assessed in parallel, did not harbor these missense changes, which proved to be target amino acids highly conserved across 5 distantly related species. Furthermore, whilst p.R171 constitutes part of a 6 helix barrel domain that contains the Helix-Hairpin-Helix motif, p.R231 lies within the alpha-8 helix making up the cluster domain. Together they form part of a DNA binding complex, where 9 lysines and 5 arginines form an electrostatically positive DNA interaction surface (13). Clearly, functional studies are needed to ascertain the pathogenicity of these novel mutations. Moreover, since the parents of the individuals harboring these gene alterations were not available for screening, we cannot exclude the possibility that the mutations in the compound heterozygotes may lie on the same allele.

In our study population, the overall allele frequency of the missense variants p.Y165C and p.G382D amounted to 5.7% (9/158) and 5.1% (8/158), respectively; if only patients with a family history compatible with an autosomal recessive mode of inheritance were considered, the allele frequencies

raised to 10% (9/90) and 8.9% (8/90). In contrast, these alterations were not present in Swiss control samples (0/100), similar to reports on Finnish blood donors (0/424) and healthy British controls (2/100) (5)(12). This further substantiates the view that the frequency of the p.Y165C and p.G382D mutations in the general population is too low to justify large-scale mutation screening (7).

The overall frequencies of biallelic mutation carriers did not significantly differ between patients displaying a classical (5.6%) and those displaying an attenuated (9.8%) FAP phenotype that is similar to reports by Sieber *et al*. who identified biallelic mutations in 7.5% and 5% of patients, respectively (8). The frequency of monoallelic mutation carriers, however, was significantly higher in our study group (11.4% compared with that of 3.9% as reported by Sieber *et al* (8). which may reflect ethnic and geographic differences between the populations studied. Six (86%) out of 7 biallelic *MYH* mutation carriers were found to have less than 100 polyps at the time of diagnosis and 5 (71%) had developed colorectal cancer. Thus, in contrast to initial studies reporting classical disease (>100 adenomas) in all biallelic mutation carriers (6), the *MYH* associated-polyposis phenotype in our patients is predominantly an attenuated one, which is in accordance with recent data from Enholm *et al*.(12) ,who investigated a population-based series of Finnish CRC patients.

On the basis of clinicopathological features, it is virtually impossible to discriminate biallelic from monoallelic *MYH* mutation carriers and *MYH* mutation-negative polyposis patients who have a family history compatible with autosomal recessive inheritance. In all groups, median age at diagnosis did not differ significantly, and the occurrence of extracolonic disease was similar. Colorectal adenocarcinomas, however, were significantly ($p < 0.001$) more frequent among biallelic as compared to that of monoallelic *MYH* mutation carriers and *MYH* mutation-negative polyposis patients.

In conclusion, biallelic *MYH* germline alterations were identified in 15.5% of Swiss *APC* mutation-negative patients with a family history compatible with autosomal recessive inheritance. Biallelic mutation carriers were more frequently observed in AFAP patients compared to those with classical FAP (17% *vs*. 10%). Colorectal cancer was significantly more frequent in biallelic as compared to monoallelic mutation carriers or those without *MYH* alterations. Based on our experience and earlier reports, we suggest that *MYH* mutation screening should be offered to individuals who fulfill all of the following criteria: (*i*) presence of classical or attenuated polyposis, (*ii*) absence of an *APC* germline mutation and (*iii*) a family history compatible with an autosomal recessive mode of inheritance. It remains to be determined within the framework of

international collaborative studies if monoallelic *MYH* mutation carriers, compared to the general population, may actually be at an increased risk for developing colorectal cancer (14).

Acknowledgements

We thank all patients for their participation in this study, as well as their respective doctors for contributing clinical information. We also thank Michele Attenhofer and Marianne Haeusler for excellent technical assistance.

4.2.6 References

1. Fearnhead NS, Britton MP, Bodmer WF. The ABC of APC. *Hum Mol Genet* 2001; **10**: 721-33.
2. Spirio LN, Samowitz W, Robertson J, Robertson M, Burt RW, Leppert M, White R. Alleles of APC modulate the frequency and classes of mutations that lead to colon polyps. *Nat Genet* 1998; **20**: 385-88.
3. Su LK, Barnes CJ, Yao W, Qi Y, Lynch PM, Steinbach G. Inactivation of germline mutant APC alleles by attenuated somatic mutations: a molecular genetic mechanism for attenuated familial adenomatous polyposis. *Am J Hum Genet* 2000; **67**: 582-90.
4. Armstrong JG, Davies DR, Guy SP, Frayling IM, Evans DG. APC mutations in familial adenomatous polyposis families in the Northwest of England. *Hum Mutat* 1997; **10**: 376-380.
5. Al-Tassan N, Chmiel NH, Maynard J, Fleming N, Livingston AL, Williams GT, Hodges AK, Davies DR, David SS, Sampson JR, Cheadle JP. Inherited variants of MYH associated with somatic G:C->T:A mutations in colorectal tumors. *Nat Genet* 2002; **30**: 227-32.
6. Jones S, Emmerson P, Maynard J, Best JM, Jordan S, Williams GT, Sampson JR, Cheadle JP. Biallelic germline mutations in MYH predispose to multiple colorectal adenoma and somatic G:C->T:A mutations. *Hum Mol Genet* 2002; **11**: 2961-67.
7. Marra G, Jiricny J. Multiple colorectal adenomas-is their number up? *N Engl J Med* 2003; **348**: 845-7.
8. Sieber OM, Lipton L, Crabtree M, Heinimann K, Fidalgo P, Phillips RK, Bisgaard ML, Orntoft TF, Aaltonen LA, Hodgson SV, Thomas HJ, Tomlinson IP. Multiple colorectal adenomas, classic adenomatous polyposis, and germ-line mutations in MYH. *N Engl J Med* 2003; **348**: 791-9.
9. Sampson JR, Dolwani S, Jones S, Eccles D, Ellis A, Evans DG, Frayling I, Jordan S, Maher ER, Mak T, Maynard J, Pigatto F, et al. Autosomal recessive colorectal adenomatous polyposis due to inherited mutations of MYH. *Lancet* 2003; **362**: 39-41.
10. Heinimann K, Thompson A, Locher A, Furlanetto T, Bader E, Wolf A, Meier R, Walter K, Bauerfeind P, Marra G, Muller H, Foernzler D, et al. Nontruncating apc germ-line mutations and mismatch repair deficiency play a minor role in apc mutation-negative polyposis. *Cancer Res* 2001; **61**: 7616-22.

11. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988; **16**: 1215.
12. Enholm S, Hienonen T, Suomalainen A, Lipton L, Tomlinson I, Karja V, Eskelinen M, Mecklin JP, Karhu A, Jarvinen HJ Aaltonen LA. Proportion and phenotype of MYH-associated colorectal neoplasia in a population-based series of Finnish colorectal cancer patients. *Am J Pathol* 2003; **163**: 827-32.
13. Guan Y, Manuel RC, Arvai AS, Parikh SS, Mol CD, Miller JH, Lloyd S Tainer JA. MutY catalytic core, mutant and bound adenine structures define specificity for DNA repair enzyme superfamily. *Nat Struct Biol* 1998; **5**: 1058-64.
14. Croitoru ME, Cleary SP, Di Nicola N, Manno M, Selander T, Aronson M, Redston M, Cotterchio M, Knight J, Gryfe R Gallinger S. Association between biallelic and monoallelic germline MYH gene mutations and colorectal cancer risk. *J Natl Cancer Inst* 2004; **96**: 1631-34.

CHAPTER 5

5. General Discussion

In this thesis, our investigation have focused on determining the prevalence of large genomic rearrangements and the germline mutations in novel susceptibility genes within major hereditary colorectal cancers syndromes: hereditary nonpolyposis colorectal cancer (HNPCC), familial adenomatous polyposis (FAP) and *MYH* - Assoicated Polyposis (MAP). In addition, we have characterized the second somatic mutation in tumors from *MMR* and *APC* gene mutation carriers to address the mechanisms of tumor initiation in HNPCC and FAP.

All these investigations are aimed to understand tumor initiation and progression in hereditary colorectal cancer in order to enable early diagnosis and devise optimal medical therapy and prevention of cancers.

Prevalence of large genomic rearrangements in HNPCC

Hereditary Nonpolyposis Colorectal Cancer (HNPCC) is an inherited cancer syndrome caused by a deficiency in the DNA mismatch repair system. The majority of mismatch repair (MMR) gene mutations have been detected in the *MLH1* and *MSH2* genes. Most mutations are substitutions, small insertions and deletions. However, standard methods of mutation analysis do not detect large genomic rearrangements which may account for a significant proportion of *MLH1* and *MSH2* mutations. Two novel methods (QMPA and MLPA) were established and compared by the detection of large deletions in 35 mutation negative Swiss HNPCC patients. Twenty - one of them presented with tumors exhibiting microsatellite instability and 16 of them showing immunohistochemical loss of either the *MLH1* or *MSH2* gene product. Four large genomic deletions were detected by QMPA and three of them could be confirmed by MLPA. The results indicated that genomic deletions account for a substantial fraction of mutations in both the *MLH1* and *MSH2* genes. Two methods applied for large deletion screening, QMPA and MLPA are readily identifying the deletions in the patients, albeit with variable specificity. Compared to MLPA, the QMPA technique is difficult to establish and standardize the PCR conditions to obtain reproducible results. Therefore, MLPA to QMPA is better suited for the routine genetic testing for large genomic rearrangements.

Prevalence of MYH germline mutations in FAP patients

FAP accounts for approximately 0.1-1% of all colorectal cancers. Classical FAP is characterized by the presence of hundreds to thousands of adenomatous colorectal polyps. The majority of

classical FAP patients carry mutations in the *APC* gene. However, in 10-30% of patients with classical FAP and up to 90% of those with attenuated FAP, no pathogenic germline alterations in *APC* gene can be identified. Recently, homozygous mutations in the *MYH* gene have been shown to predispose to a multiple adenoma and carcinoma phenotype. In an attempt to assess the prevalence of *MYH* mutation carriers among mutation negative polyposis patients and to identify possible phenotypic differences between *MYH* mutation carriers and *APC*/*MYH* mutation negative polyposis patients, 79 unrelated *APC* mutation negative Swiss patients were screened for *MYH* by dHPLC and direct DNA sequencing. 9% of them were found to harbor biallelic (n=7) and 11% monoallelic (n=9) germline mutations in the *MYH* gene. Considering only patients with a family history compatible with autosomal recessive inheritance, biallelic *MYH* mutation carriers were observed in 10% of patients with classical and in 17% of those with attenuated polyposis. Two *MYH* mutation hotspots in p.Y165C and p.G382D account for 43% and 29% mutant alleles in the biallelic patients. Biallelic *MYH* germline mutations were identified in 15.5% of Swiss *APC* mutation negative patients with a family history compatible with autosomal recessive inheritance. They were more frequently observed in AFAP patients rather than in those with classical FAP. Colorectal cancer was significantly more frequent in biallelic than in monoallelic mutation carriers or in those without *MYH* alterations. From these study, we suggest that *MYH* mutation screening should be offered to individuals if all of the following criteria are fulfilled: (1) presence of classical or attenuated polyposis and early on-set CRC (2) absence of an *APC* germline mutation, and (3) a family history compatible with an autosomal recessive mode of inheritance.

Characterization of the Somatic mutation in tumors from MMR and APC gene mutation carriers

Generally, in most hereditary cancer predispositions, the first hit is inherited in an autosomal dominant type [80]. According to Knudson's "two hit" hypothesis, in hereditary cancers, only one mutated copy in a given tumor suppressor gene (TSG) is not enough to enable cancer initiation, a second somatic mutation of wild-type allele of TSG is necessary for cancer development [81]. In sporadic cancers, two somatic mutations of TSG need to occur in one somatic cell to initiate cancer development.

Thus the germline mutation carriers get much greater chance than general population to get cancer since they already harbor the first germline mutation in all the cells of the body [90] .

Colorectal cancer is an excellent model to study the genetic mechanisms for tumor initiation and progression [74]. Major genes like *APC* and *MMR* genes involved in major familial colon cancer

syndromes have been well characterized. However, little is known about the second somatic hits in HNPCC and AFAP tumorigenesis.

Characterization of second hit mechanism in Hereditary Nonpolyposis Colorectal Cancer (HNPCC)

Tumor development in HNPCC is believed to be initiated by the loss function of the DNA mismatch repair (MMR) system. DNA mismatch repair deficiency results from both the germline and somatic mutation in the affected MMR gene such as *MLH1 and MSH2* in a cell. It is also known that aberrant promoter methylation of a DNA mismatch repair is associated with loss function of MMR.

In this thesis, we performed a comprehensive analysis of second hit in tumors of well-defined set of Swiss and Finnish HNPCC patients carrying large deletions in *MLH1 or MSH2* gene. We aimed to define its contribution to somatic inactivation of the remaining wildtype allele.

Nine cancers from 5 Swiss *MLH1* or *MSH2* large genomic carriers and 7 tumors from 7 Finnish *MLH1* or *MSH2* deletion carrierrs were investigated. These 16 tumors are: 11 of them are adenocarcinomas of the colon, 2 endometrium cancers, 1 stomach cancer, 1 urothelial carcinoma and 1 astrocytoma (brain tumor). Most of them were exhibiting high MSI. Only two carcinomas (urothelial carcinoma of patient 2227 and colon adenomcarcinoma of patient 1676) showed low microsatellite instability (MSI). Pathological reports were available of all Swiss tumors. Within 5 Swiss large genomic deletion carriers, 5 tumors showed identical somatic mutation to their germline mutation (Chapter 3.2, Table 3.2.2a). 2 Tumors out of 7 Finnish tumors also showed homozygous somatic deletion. One colorectal and one urothelial carcinoma showed loss of one exon only. No large genomic deletions and duplications were detected in the remaining tumors (2 CRCs, 1 ovarian and 1 endometrial cancer). Thus, our findings from two independent sets of patients indicate that homozygous mutation of somatic cell is a frequent event in HNPCC (6 out of 11 CRCs; 55%). Remarkably, none of the Swiss or Finnish tumor specimens showed evidence for large somatic deletions encompassing the entire respective gene locus (LOH).

In addition to these results, we were able to collect data from one sporadic cancer patient who carrys a *de novo* germline mutation (c.666dupA) in the *MLH1* gene. The patient is 31 years old. The tumor developed in the left side of the colon like other sporadic colon cancer. The somatic

mutation that was found in the colorectal carcinoma of this patient was identical to its germline mutation (C.666dup A). MLPA results showed no LOH in his tumor. (Chapter 3.3)
There are several possible mechanisms leading to this homo/hemizygosity mutation in HNPCC tumors. Among them are: loss of the complete wild type allele, complete deletion of the gene locus of the wild type allele, loss of the chromosome harboring the wild-type allele followed by chromosomal reduplication, mitotic recombination, restricted recombination like gene conversion [120]. Multiplex ligation-dependent probe amplication (MLPA) and microsatellite marker analysis flanking the gene loci on the respective chromosome were applied to distinguish between them. The results of those experiments ruled out the following possibilies: loss of the complete wild type allele, complete deletion of the gene locus of the wild type allele, loss of the chromosome harboring the wild-type allele followed by chromosomal reduplication and mitotic recombination except restricted recombination. The data pointed to locus-restricted recombination as the putatitive mechanism.

Locus-restricted recombination, i.e gene conversion is a possible consequence of selection for reduced rates of unequal exchange between repeated DNA sequences for which the copy number is subject to stabilizing selection. The repeated DNA sequences, expecially short interspersed nuclear elements (SINE) like Alu elements may also contribute to hereditary disease including cancer [117]. Alu sequences are 300 base pair long and are classified as short interspersed elements (SINEs) amongst the class of repetitive DNA elements. It is estimated that about 10% of the mass of human genome consists of Alu sequences [119] . Alu insertion has been implicated in several inherited human diseases including cancers. Alu repeat elements have been already shown to be involved in germline *MLH1/MSH2* rearrangements by several studies [118]. Based on this hypothesis, we analyzed the genome sequences of *MLH1*/*MSH2* by repeatmarker program in order to assess the frequency and type of DNA repeats inside (UCSC genome analysis website: http://genome.ucsc.edu/) (Figure 1).

Figure 1 depicts the number of repeat sequences in all introns involved in locus-restricted recombination in our patients. Interestingly, all intronic regions involved in the sequence were regions rich in repeated sequences.

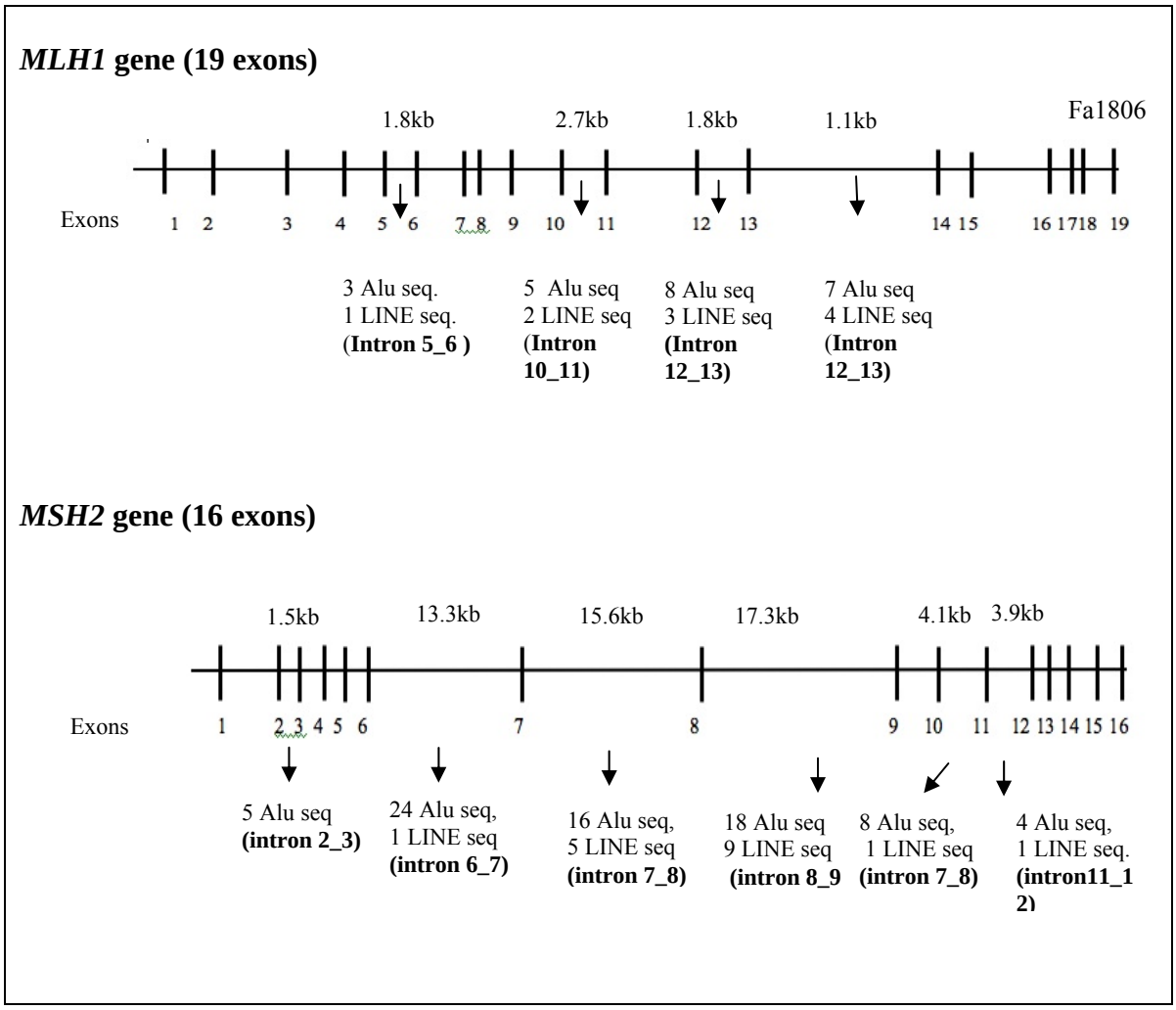


Figure 5.1: Distribution of Alu sequences within the introns of *MLH1* and *MSH2* gene. Intron sequences were analysis by RepeatMarker programe **(***MSH2* gene ensembl ID ENSG00000095002, *MLH1* gene Ensembl ID ENSG00000076242). LINE: Long interspersed nuclear element . Alu sequence: a family of short repeat sequence (<300bp) through the human genome

Unequal recombination often occurs intrachromosomally, resulting in large genomic rearrangements and more complex chromosomal abnormalities. It has been reported that the unequal crossing over mediated by Alu repeats is a possible principle factor in tumor progression through loss of heterozygosity (LOH) and genomic rearrangement [120]. The mechanism implies that high density Alu repeats decreased the sensitivity of base pairing fidelity that would presumably allow recombination to happen between more poorly matched homologous [117]. Our results support the assumption that the somatic mutation identical to the one in the germline in tumors of HNPCC patients is due to locus-restricted recombination (i.e, gene conversion), most likely caused by high density Alu sequences within the introns of the deleted regions.

In additon to Alu repeats, Figure 1 also shows LINE (long interspersed nuclear elements) sequences in the detected intron sequence. It has been suggested that LINE sequences could be another factor contributing to the recombination involved. Even a 10-fold lower copy number of LINE elements compared to Alu sequence would be more than enough to cause recombination. It has been oberserved that LINE/LINE (L/L) recombination is usually involved in larger region [120].

To explain the recombination mechanisms which cause the second hit in tumors from HNPCC patients, the breakpoints of respective deletions have to be determined. Unfortunately, fresh tumor materials were not available to further assess this issue.

In conclusion, our analysis of cancer specimens from two independent sets of Swiss and Finnish *MLH1/MSH2* deletion carriers and analysis of one *de novo* case revealed high frequency of somatic mutation identical to the ones in the germline as a common second hit in CRCs. This type of inactivation of the wild type allele is also considered as a common second hit in extra colonic HNPCC associated tumors. Chromosome specific marker analysis implies that loss of the wild-type allele predominantly occurs through locus-restricted recombination events, i.e., gene conversion rather than mitotic recombination or deletion of the respective gene locus. This was also confirmed by the result of a colorectal cancer from a patient with *de novo* germline mutation (c.666dupA) in *MLH1* gene.

Characterization of somatic hits in attenuated Familial Adenomatous Polyposis (aFAP)

The 'first hit –second hit' association in FAP syndrome has been discussed in several reports [97,122]. In this thesis, we collaborated with nine research groups to do a detailed investigation in AFAP patients whose germline mutation are located at the very 5' end or 3'end of the *APC* gene. In total, 235 tumors from 35 patients (16 families) with a variety of AFAP associated germline mutations were involved in the investigation. A number of methods have been used to detect and confirm the mutations that were found in the tumors of all the patients involved. These analytical methods included dHPLC, direct DNA sequencing, SSCP, MLPA and site restricted cloning.

The study showed that two somatic mutations, including loss or mutation of the germline mutant allele could occur ('three hits') in AFAP tumors. The 'third hit' probably initiated tumorigenesis. We found when the mutation happened at exon 9 of *APC* gene, a 'third hits' is very common.

There are six exon 9 mutant families involved in this study, almost all tumors have been found with the 'third hits'. Most 'third hits' left three 20-amino acid repeats on the germline mutant APC allele with LOH of the wild-type allele. By contrast, the 'third hit' was much more rare in patients with 5' and 3' germline mutation. Around 15-20% 5' APC mutation carriers seem to acquire 'three hits'. The frequency of 'third hits' in the patients with 3' mutation is similar to the patients with 5' mutation.

Overall, these AFAP tumors studies provide an understanding why these patients had fewer polyps than classical FAP patients due to 'three Hits' inactivation of APC gene. In conclusion, the genetic analysis of AFAP patients has revealed a novel mechanism to the genotype phenotype association in their tumor syndrome. For some AFAP adenomas, three hits are needed for tumorigensis.

In summary, the understanding of the interdependence and the mechanisms involved in the acquisition of mutations is essential to our knowledge on tumor initiation and tumor progression in hereditary colorectal cancers as well as in sporadic cancers. This should help us ultimately to identify new potential target areas for the cancer therapy and design new efficient drugs to cure cancer.

The knowledge on germline and somatic mutations may be used in the future to create personalized chemotherapeutic strategies and eventually prevent susceptible individuals from cancer.

REFERENCES

1. Neta, R. (2000). The promise of molecular epidemiology in defining the association between radiation and cancer. Health Phys *79*, 77-84.
2. Albertini, R.J., Nicklas, J.A., and O'Neill, J.P. (1996). Future research directions for evaluating human genetic and cancer risk from environmental exposures. Environ Health Perspect *104 Suppl 3*, 503-510.
3. Society, A.C. (2005). Cancer Facts and Figures 2005. Ga:American Cancer Society.
4. Gryfe, R., Kim, H., Hsieh, E.T., Aronson, M.D., Holowaty, E.J., Bull, S.B., Redston, M., and Gallinger, S. (2000). Tumor microsatellite instability and clinical outcome in young patients with colorectal cancer. N Engl J Med *342*, 69-77.
5. Chang, S.C., Lin, J.K., Yang, S.H., Wang, H.S., Li, A.F., and Chi, C.W. (2006). Relationship between genetic alterations and prognosis in sporadic colorectal cancer. Int J Cancer *118*, 1721-1727.
6. Burt, R.W., Leppert, M.F., Slattery, M.L., Samowitz, W.S., Spirio, L.N., Kerber, R.A., Kuwada, S.K., Neklason, D.W., Disario, J.A., Lyon, E., Hughes, J.P., Chey, W.Y., and White, R.L. (2004). Genetic testing and phenotype in a large kindred with attenuated familial adenomatous polyposis. Gastroenterology *127*, 444-451.
7. Arnold CN, G.A., Blum HE, Boland CR. (2005). Molecular pathogenesis of colorectal cancer: implications for molecular diagnosis. Cancer. *(10)*:2035-2047.
8. Nugent, K.P., Talbot, I.C., Hodgson, S.V., and Phillips, R.K. (1993). Solitary juvenile polyps: not a marker for subsequent malignancy. Gastroenterology *105*, 698-700.
9. Ahmed, F.E. (2003). Colon cancer: prevalence, screening, gene expression and mutation, and risk factors and assessment. J Environ Sci Health C Environ Carcinog Ecotoxicol Rev *21*, 65-131.
10. Leppert, M., Burt, R., Hughes, J.P., Samowitz, W., Nakamura, Y., Woodward, S., Gardner, E., Lalouel, J.M., and White, R. (1990). Genetic analysis of an inherited

predisposition to colon cancer in a family with a variable number of adenomatous polyps. N Engl J Med *322*, 904-908.

11. Nicolaides, N.C., Papadopoulos, N., Liu, B., Wei, Y.F., Carter, K.C., Ruben, S.M., Rosen, C.A., Haseltine, W.A., Fleischmann, R.D., Fraser, C.M., and et al. (1994). Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. Nature *371*, 75-80.
12. Papadopoulos, N., Nicolaides, N.C., Wei, Y.F., Ruben, S.M., Carter, K.C., Rosen, C.A., Haseltine, W.A., Fleischmann, R.D., Fraser, C.M., Adams, M.D., and et al. (1994). Mutation of a mutL homolog in hereditary colon cancer. Science *263*, 1625-1629.
13. Leach, F.S., Nicolaides, N.C., Papadopoulos, N., Liu, B., Jen, J., Parsons, R., Peltomaki, P., Sistonen, P., Aaltonen, L.A., Nystrom-Lahti, M., and et al. (1993). Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. Cell *75*, 1215-1225.
14. Kinzler, K.W., Nilbert, M.C., Su, L.K., Vogelstein, B., Bryan, T.M., Levy, D.B., Smith, K.J., Preisinger, A.C., Hedge, P., McKechnie, D., and et al. (1991). Identification of FAP locus genes from chromosome 5q21. Science *253*, 661-665.
15. Groden, J., Thliveris, A., Samowitz, W., Carlson, M., Gelbert, L., Albertsen, H., Joslyn, G., Stevens, J., Spirio, L., Robertson, M., and et al. (1991). Identification and characterization of the familial adenomatous polyposis coli gene. Cell *66*, 589-600.
16. Slupska MM, B.C., Luther VW,Chiang JH,Wei YF,Miller JH (1996). Cloning and sequencing a human homolog (hMYH) of the Escherichia coli mutY gene whose function is required for the repair of oxidative DNA damage. J Bacteriol *13*, 3885-3892.
17. Pedemonte, S., Sciallero, S., Gismondi, V., Stagnaro, P., Biticchi, R., Haeouaine, A., Bonelli, L., Nicolo, G., Groden, J., Bruzzi, P., Aste, H., and Varesco, L. (1998). Novel germline APC variants in patients with multiple adenomas. Genes Chromosomes Cancer *22*, 257-267.
18. Sieber, O.M., Lamlum, H., Crabtree, M.D., Rowan, A.J., Barclay, E., Lipton, L., Hodgson, S., Thomas, H.J., Neale, K., Phillips, R.K., Farrington, S.M., Dunlop,

M.G., Mueller, H.J., Bisgaard, M.L., Bulow, S., Fidalgo, P., Albuquerque, C., Scarano, M.I., Bodmer, W., Tomlinson, I.P., and Heinimann, K. (2002). Whole-gene APC deletions cause classical familial adenomatous polyposis, but not attenuated polyposis or "multiple" colorectal adenomas. Proc Natl Acad Sci U S A *99*, 2954-2958.

19. Bussey, H.J. (1979). Familial polyposis coli. Pathol Annu *14 Pt 1*, 61-81.
20. Jarvinen, H.J., Aarnio, M., Mustonen, H., Aktan-Collan, K., Aaltonen, L.A., Peltomaki, P., De La Chapelle, A., and Mecklin, J.P. (2000). Controlled 15-year trial on screening for colorectal cancer in families with hereditary nonpolyposis colorectal cancer. Gastroenterology *118*, 829-834.
21. Lindgren, G., Liljegren, A., Jaramillo, E., Rubio, C., and Lindblom, A. (2002). Adenoma prevalence and cancer risk in familial non-polyposis colorectal cancer. Gut *50*, 228-234.
22. Hemminki, A., Avizienyte, E., Roth, S., Loukola, A., Aaltonen, L.A., Jarvinen, H., and de la Chapelle, A. (1998). [A serine/threonine kinase gene defective in Peutz-Jeghers syndrome]. Duodecim *114*, 667-668.
23. Lim, W., Hearle, N., Shah, B., Murday, V., Hodgson, S.V., Lucassen, A., Eccles, D., Talbot, I., Neale, K., Lim, A.G., O'Donohue, J., Donaldson, A., Macdonald, R.C., Young, I.D., Robinson, M.H., Lee, P.W., Stoodley, B.J., Tomlinson, I., Alderson, D., Holbrook, A.G., Vyas, S., Swarbrick, E.T., Lewis, A.A., Phillips, R.K., and Houlston, R.S. (2003). Further observations on LKB1/STK11 status and cancer risk in Peutz-Jeghers syndrome. Br J Cancer *89*, 308-313.
24. Coburn, M.C., Pricolo, V.E., DeLuca, F.G., and Bland, K.I. (1995). Malignant potential in intestinal juvenile polyposis syndromes. Ann Surg Oncol *2*, 386-391.
25. Huang, S.C., Chen, C.R., Lavine, J.E., Taylor, S.F., Newbury, R.O., Pham, T.T., Ricciardiello, L., and Carethers, J.M. (2000). Genetic heterogeneity in familial juvenile polyposis. Cancer Res *60*, 6882-6885.
26. Marsh, D.J., Kum, J.B., Lunetta, K.L., Bennett, M.J., Gorlin, R.J., Ahmed, S.F., Bodurtha, J., Crowe, C., Curtis, M.A., Dasouki, M., Dunn, T., Feit, H., Geraghty, M.T., Graham, J.M., Jr., Hodgson, S.V., Hunter, A., Korf, B.R., Manchester, D., Miesfeldt, S., Murday, V.A., Nathanson, K.L., Parisi, M., Pober, B., Romano, C.,

Eng, C., and et al. (1999). PTEN mutation spectrum and genotype-phenotype correlations in Bannayan-Riley-Ruvalcaba syndrome suggest a single entity with Cowden syndrome. Hum Mol Genet *8*, 1461-1472.

27. Zhou, X.P., Waite, K.A., Pilarski, R., Hampel, H., Fernandez, M.J., Bos, C., Dasouki, M., Feldman, G.L., Greenberg, L.A., Ivanovich, J., Matloff, E., Patterson, A., Pierpont, M.E., Russo, D., Nassif, N.T., and Eng, C. (2003). Germline PTEN promoter mutations and deletions in Cowden/Bannayan-Riley-Ruvalcaba syndrome result in aberrant PTEN protein and dysregulation of the phosphoinositol-3-kinase/Akt pathway. Am J Hum Genet *73*, 404-411.
28. Lynch, H.T., Shaw, M.W., Magnuson, C.W., Larsen, A.L., and Krush, A.J. (1966). Hereditary factors in cancer. Study of two large midwestern kindreds. Arch Intern Med *117*, 206-212.
29. De Jong, A.E., Morreau, H., Van Puijenbroek, M., Eilers, P.H., Wijnen, J., Nagengast, F.M., Griffioen, G., Cats, A., Menko, F.H., Kleibeuker, J.H., and Vasen, H.F. (2004). The role of mismatch repair gene defects in the development of adenomas in patients with HNPCC. Gastroenterology *126*, 42-48.
30. Vasen, H.F., Sanders, E.A., Taal, B.G., Nagengast, F.M., Griffioen, G., Menko, F.H., Kleibeuker, J.H., Houwing-Duistermaat, J.J., and Meera Khan, P. (1996). The risk of brain tumours in hereditary non-polyposis colorectal cancer (HNPCC). Int J Cancer *65*, 422-425.
31. Cavenee, W.K. (2003). The recessive nature of dominance. Genes Chromosomes Cancer *38*, 322-325.
32. Kullmann, F. (2003). [Carcinogenesis and hereditart colon cancers]. Internist (Berl) *44*, 254-255, 258-267.
33. Plaschke, J., Engel, C., Kruger, S., Holinski-Feder, E., Pagenstecher, C., Mangold, E., Moeslein, G., Schulmann, K., Gebert, J., von Knebel Doeberitz, M., Ruschoff, J., Loeffler, M., and Schackert, H.K. (2004). Lower incidence of colorectal cancer and later age of disease onset in 27 families with pathogenic MSH6 germline mutations compared with families with MLH1 or MSH2 mutations: the German Hereditary Nonpolyposis Colorectal Cancer Consortium. J Clin Oncol *22*, 4486-4494.

34. Gazzoli, I., Loda, M., Garber, J., Syngal, S., and Kolodner, R.D. (2002). A hereditary nonpolyposis colorectal carcinoma case associated with hypermethylation of the MLH1 gene in normal tissue and loss of heterozygosity of the unmethylated allele in the resulting microsatellite instability-high tumor. Cancer Res *62*, 3925-3928.

35. Vasen, H.F., Mecklin, J.P., Khan, P.M., and Lynch, H.T. (1991). The International Collaborative Group on Hereditary Non-Polyposis Colorectal Cancer (ICG-HNPCC). Dis Colon Rectum *34*, 424-425.

36. Vasen, H.F., Watson, P., Mecklin, J.P., and Lynch, H.T. (1999). New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC. Gastroenterology *116*, 1453-1456.

37. Rodriguez-Bigas, M.A., Boland, C.R., Hamilton, S.R., Henson, D.E., Jass, J.R., Khan, P.M., Lynch, H., Perucho, M., Smyrk, T., Sobin, L., and Srivastava, S. (1997). A National Cancer Institute Workshop on Hereditary Nonpolyposis Colorectal Cancer Syndrome: meeting highlights and Bethesda guidelines. J Natl Cancer Inst *89*, 1758-1762.

38. Umar, A., Boland, C.R., Terdiman, J.P., Syngal, S., de la Chapelle, A., Ruschoff, J., Fishel, R., Lindor, N.M., Burgart, L.J., Hamelin, R., Hamilton, S.R., Hiatt, R.A., Jass, J., Lindblom, A., Lynch, H.T., Peltomaki, P., Ramsey, S.D., Rodriguez-Bigas, M.A., Vasen, H.F., Hawk, E.T., Barrett, J.C., Freedman, A.N., and Srivastava, S. (2004). Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. J Natl Cancer Inst *96*, 261-268.

39. Macrae, F., and Harris, M. (2005). Re: Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. J Natl Cancer Inst *97*, 936-937; author reply 937-938.

40. Peltomaki, P., and Vasen, H.F. (1997). Mutations predisposing to hereditary nonpolyposis colorectal cancer: database and results of a collaborative study. The International Collaborative Group on Hereditary Nonpolyposis Colorectal Cancer. Gastroenterology *113*, 1146-1158.

41. Peltomaki, P. (2005). Lynch syndrome genes. Fam Cancer *4*, 227-232.

42. Fishel, R., Lescoe, M.K., Rao, M.R., Copeland, N.G., Jenkins, N.A., Garber, J., Kane, M., and Kolodner, R. (1993). The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. Cell *75*, 1027-1038.

43. Grady, W.M., and Markowitz, S.D. (2003). Hereditary colon cancer genes. Methods Mol Biol *222*, 59-83.

44. Chung, D.C., and Rustgi, A.K. (1995). DNA mismatch repair and cancer. Gastroenterology *109*, 1685-1699.

45. Fishel, R. (2001). The selection for mismatch repair defects in hereditary nonpolyposis colorectal cancer: revising the mutator hypothesis. Cancer Res *61*, 7369-7374.

46. Boland, C.R., Thibodeau, S.N., Hamilton, S.R., Sidransky, D., Eshleman, J.R., Burt, R.W., Meltzer, S.J., Rodriguez-Bigas, M.A., Fodde, R., Ranzani, G.N., and Srivastava, S. (1998). A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. Cancer Res *58*, 5248-5257.

47. Weber, J.L., and May, P.E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet *44*, 388-396.

48. Aaltonen, L.A., Peltomaki, P., Leach, F.S., Sistonen, P., Pylkkanen, L., Mecklin, J.P., Jarvinen, H., Powell, S.M., Jen, J., Hamilton, S.R., and et al. (1993). Clues to the pathogenesis of familial colorectal cancer. Science *260*, 812-816.

49. Eshleman, J.R., and Markowitz, S.D. (1996). Mismatch repair defects in human carcinogenesis. Hum Mol Genet *5 Spec No*, 1489-1494.

50. Wijnen, J., de Leeuw, W., Vasen, H., van der Klift, H., Moller, P., Stormorken, A., Meijers-Heijboer, H., Lindhout, D., Menko, F., Vossen, S., Moslein, G., Tops, C., Brocker-Vriends, A., Wu, Y., Hofstra, R., Sijmons, R., Cornelisse, C., Morreau, H., and Fodde, R. (1999). Familial endometrial cancer in female carriers of MSH6 germline mutations. Nat Genet *23*, 142-144.

51. Hendriks, Y.M., Wagner, A., Morreau, H., Menko, F., Stormorken, A., Quehenberger, F., Sandkuijl, L., Moller, P., Genuardi, M., Van Houwelingen, H., Tops, C., Van Puijenbroek, M., Verkuijlen, P., Kenter, G., Van Mil, A., Meijers-Heijboer, H., Tan, G.B., Breuning, M.H., Fodde, R., Wijnen, J.T., Brocker-Vriends, A.H., and Vasen, H. (2004). Cancer risk in hereditary nonpolyposis colorectal cancer due to MSH6 mutations: impact on counseling and surveillance. Gastroenterology *127*, 17-25.

52. Loukola, A., Vilkki, S., Singh, J., Launonen, V., and Aaltonen, L.A. (2000). Germline and somatic mutation analysis of MLH3 in MSI-positive colorectal cancer. Am J Pathol *157*, 347-352.

53. Loukola, A., de la Chapelle, A., and Aaltonen, L.A. (1999). Strategies for screening for hereditary non-polyposis colorectal cancer. J Med Genet *36*, 819-822.

54. Lindor, N.M., Burgart, L.J., Leontovich, O., Goldberg, R.M., Cunningham, J.M., Sargent, D.J., Walsh-Vockley, C., Petersen, G.M., Walsh, M.D., Leggett, B.A., Young, J.P., Barker, M.A., Jass, J.R., Hopper, J., Gallinger, S., Bapat, B., Redston, M., and Thibodeau, S.N. (2002). Immunohistochemistry versus microsatellite instability testing in phenotyping colorectal tumors. J Clin Oncol *20*, 1043-1048.

55. Cawkwell, L., Gray, S., Murgatroyd, H., Sutherland, F., Haine, L., Longfellow, M., O'Loughlin, S., Cross, D., Kronborg, O., Fenger, C., Mapstone, N., Dixon, M., and Quirke, P. (1999). Choice of management strategy for colorectal cancer based on a diagnostic immunohistochemical test for defective mismatch repair. Gut *45*, 409-415.

56. Thibodeau, S.N., French, A.J., Roche, P.C., Cunningham, J.M., Tester, D.J., Lindor, N.M., Moslein, G., Baker, S.M., Liskay, R.M., Burgart, L.J., Honchel, R., and Halling, K.C. (1996). Altered expression of hMSH2 and hMLH1 in tumors with microsatellite instability and genetic alterations in mismatch repair genes. Cancer Res *56*, 4836-4840.

57. Ainsworth, P.J., Koscinski, D., Fraser, B.P., and Stuart, J.A. (2004). Family cancer histories predictive of a high risk of hereditary non-polyposis colorectal

cancer associate significantly with a genomic rearrangement in hMSH2 or hMLH1. Clin Genet *66*, 183-188.

58. Rodriguez-Bigas, M.A., Vasen, H.F., Pekka-Mecklin, J., Myrhoj, T., Rozen, P., Bertario, L., Jarvinen, H.J., Jass, J.R., Kunitomo, K., Nomizu, T., and Driscoll, D.L. (1997). Rectal cancer risk in hereditary nonpolyposis colorectal cancer after abdominal colectomy. International Collaborative Group on HNPCC. Ann Surg *225*, 202-207.
59. Sampson, J.R., Jones, S., Dolwani, S., and Cheadle, J.P. (2005). MutYH (MYH) and colorectal cancer. Biochem Soc Trans *33*, 679-683.
60. Kimura, H., Yabushita, K., Tsuneda, A., Tazawa, K., Inoue, T., Earashi, M., Maeda, K., Kuroda, Y., Konishi, K., Tsuji, M., and Miwa, A. (1996). Adult colorectal invagination associated with colonic cancer and established by fiberscopy of the sigmoid colon. Endoscopy *28*, 636.
61. McGahan, T.P., and Gilinsky, N.H. (1994). Colonic tumors. Endoscopy *26*, 70-87.
62. Spirio, L., Olschwang, S., Groden, J., Robertson, M., Samowitz, W., Joslyn, G., Gelbert, L., Thliveris, A., Carlson, M., Otterud, B., and et al. (1993). Alleles of the APC gene: an attenuated form of familial polyposis. Cell *75*, 951-957.
63. Jarvinen, H.J. (1992). Epidemiology of familial adenomatous polyposis in Finland: impact of family screening on the colorectal cancer rate and survival. Gut *33*, 357-360.
64. Smits, R., van der Houven van Oordt, W., Luz, A., Zurcher, C., Jagmohan-Changur, S., Breukel, C., Khan, P.M., and Fodde, R. (1998). Apc1638N: a mouse model for familial adenomatous polyposis-associated desmoid tumors and cutaneous cysts. Gastroenterology *114*, 275-283.
65. Soravia, C., Berk, T., Madlensky, L., Mitri, A., Cheng, H., Gallinger, S., Cohen, Z., and Bapat, B. (1998). Genotype-phenotype correlations in attenuated adenomatous polyposis coli. Am J Hum Genet *62*, 1290-1301.
66. Peltomaki, P., and de la Chapelle, A. (1997). Mutations predisposing to hereditary nonpolyposis colorectal cancer. Adv Cancer Res *71*, 93-119.

67. Bertario, L., Russo, A., Sala, P., Varesco, L., Giarola, M., Mondini, P., Pierotti, M., Spinelli, P., and Radice, P. (2003). Multiple approach to the exploration of genotype-phenotype correlations in familial adenomatous polyposis. J Clin Oncol *21*, 1698-1707.

68. Halford, S.E., Rowan, A.J., Lipton, L., Sieber, O.M., Pack, K., Thomas, H.J., Hodgson, S.V., Bodmer, W.F., and Tomlinson, I.P. (2003). Germline mutations but not somatic changes at the MYH locus contribute to the pathogenesis of unselected colorectal cancers. Am J Pathol *162*, 1545-1548.

69. Sieber, O.M., Lipton, L., Crabtree, M., Heinimann, K., Fidalgo, P., Phillips, R.K., Bisgaard, M.L., Orntoft, T.F., Aaltonen, L.A., Hodgson, S.V., Thomas, H.J., and Tomlinson, I.P. (2003). Multiple colorectal adenomas, classic adenomatous polyposis, and germ-line mutations in MYH. N Engl J Med *348*, 791-799.

70. Laken, S.J., Papadopoulos, N., Petersen, G.M., Gruber, S.B., Hamilton, S.R., Giardiello, F.M., Brensinger, J.D., Vogelstein, B., and Kinzler, K.W. (1999). Analysis of masked mutations in familial adenomatous polyposis. Proc Natl Acad Sci U S A *96*, 2322-2326.

71. Giardiello, F.M. (1997). Genetic testing in hereditary colorectal cancer. Jama *278*, 1278-1281.

72. Fodde, R., Smits, R., and Clevers, H. (2001). APC, signal transduction and genetic instability in colorectal cancer. Nat Rev Cancer *1*, 55-67.

73. Groves, C., Lamlum, H., Crabtree, M., Williamson, J., Taylor, C., Bass, S., Cuthbert-Heavens, D., Hodgson, S., Phillips, R., and Tomlinson, I. (2002). Mutation cluster region, association between germline and somatic mutations and genotype-phenotype correlation in upper gastrointestinal familial adenomatous polyposis. Am J Pathol *160*, 2055-2061.

74. Fearon, E.R., and Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. Cell *61*, 759-767.

75. Gryfe, R., Swallow, C., Bapat, B., Redston, M., Gallinger, S., and Couture, J. (1997). Molecular biology of colorectal cancer. Curr Probl Cancer *21*, 233-300.

76. Narayan, S., and Roy, D. (2003). Role of APC and DNA mismatch repair genes in the development of colorectal cancers. Mol Cancer *2*, 41.

77. Li, G.M. (2003). DNA mismatch repair and cancer. Front Biosci *8*, d997-1017.

78. Lynch, H.T., and de la Chapelle, A. (1999). Genetic susceptibility to non-polyposis colorectal cancer. J Med Genet *36*, 801-818.

79. Kinzler, K.W., and Vogelstein, B. (1997). Cancer-susceptibility genes. Gatekeepers and caretakers. Nature *386*, 761, 763.

80. Fernandez-Peralta, A.M., Nejda, N., Oliart, S., Medina, V., Azcoita, M.M., and Gonzalez-Aguilera, J.J. (2005). Significance of mutations in TGFBR2 and BAX in neoplastic progression and patient outcome in sporadic colorectal tumors with high-frequency microsatellite instability. Cancer Genet Cytogenet *157*, 18-24.

81. Calin, G., Herlea, V., Barbanti-Brodano, G., and Negrini, M. (1998). The coding region of the Bloom syndrome BLM gene and of the CBL proto-oncogene is mutated in genetically unstable sporadic gastrointestinal tumors. Cancer Res *58*, 3777-3781.

82. Ouyang, H., Shiwaku, H.O., Hagiwara, H., Miura, K., Abe, T., Kato, Y., Ohtani, H., Shiiba, K., Souza, R.F., Meltzer, S.J., and Horii, A. (1997). The insulin-like growth factor II receptor gene is mutated in genetically unstable cancers of the endometrium, stomach, and colorectum. Cancer Res *57*, 1851-1854.

83. Kopnin, B. (1993). Genetic events responsible for colorectal tumorigenesis: achievements and challenges. Tumori *79*, 235-243.

84. Takayama, T., Katsuki, S., Takahashi, Y., Ohi, M., Nojiri, S., Sakamaki, S., Kato, J., Kogawa, K., Miyake, H., and Niitsu, Y. (1998). Aberrant crypt foci of the colon as precursors of adenoma and cancer. N Engl J Med *339*, 1277-1284.

85. Takaku, K., Oshima, M., Miyoshi, H., Matsui, M., Seldin, M.F., and Taketo, M.M. (1998). Intestinal tumorigenesis in compound mutant mice of both Dpc4 (Smad4) and Apc genes. Cell *92*, 645-656.

86. Arends, J.W. (2000). Molecular interactions in the Vogelstein model of colorectal carcinoma. J Pathol *190*, 412-416.

87. Fearnhead, N.S., Wilding, J.L., and Bodmer, W.F. (2002). Genetics of colorectal cancer: hereditary aspects and overview of colorectal tumorigenesis. Br Med Bull *64*, 27-43.

88. Kinzler, K.W., and Vogelstein, B. (1998). Landscaping the cancer terrain. Science *280*, 1036-1037.

89. Knudson, A.G., Jr. (1971). Mutation and cancer: statistical study of retinoblastoma. Proc Natl Acad Sci U S A *68*, 820-823.

90. Knudson, A.G. (1996). Hereditary cancer: two hits revisited. J Cancer Res Clin Oncol *122*, 135-140.

91. Tomlinson, I.P., Roylance, R., and Houlston, R.S. (2001). Two hits revisited again. J Med Genet *38*, 81-85.

92. Cavenee, W.K., Hansen, M.F., Scrable, H.J., and James, C.D. (1989). Loss of genetic information in cancer. Ciba Found Symp *142*, 79-88; discussion 88-92.

93. Powell, S.M., Zilz, N., Beazer-Barclay, Y., Bryan, T.M., Hamilton, S.R., Thibodeau, S.N., Vogelstein, B., and Kinzler, K.W. (1992). APC mutations occur early during colorectal tumorigenesis. Nature *359*, 235-237.

94. Rubinfeld, B., Souza, B., Albert, I., Muller, O., Chamberlain, S.H., Masiarz, F.R., Munemitsu, S., and Polakis, P. (1993). Association of the APC gene product with beta-catenin. Science *262*, 1731-1734.

95. Miltenberger, R.J., Farnham, P.J., Smith, D.E., Stommel, J.M., and Cornwell, M.M. (1995). v-Raf activates transcription of growth-responsive promoters via GC-rich sequences that bind the transcription factor Sp1. Cell Growth Differ *6*, 549-556.

96. Su, L.K., Vogelstein, B., and Kinzler, K.W. (1993). Association of the APC tumor suppressor protein with catenins. Science *262*, 1734-1737.

97. Albuquerque, C., Breukel, C., van der Luijt, R., Fidalgo, P., Lage, P., Slors, F.J., Leitao, C.N., Fodde, R., and Smits, R. (2002). The 'just-right' signaling model: APC somatic mutations are selected based on a specific level of activation of the beta-catenin signaling cascade. Hum Mol Genet *11*, 1549-1560.

98. Heinen, C.D., Schmutte, C., and Fishel, R. (2002). DNA repair and tumorigenesis: lessons from hereditary cancer syndromes. Cancer Biol Ther *1*, 477-485.

99. Tannergard, P., Liu, T., Weger, A., Nordenskjold, M., and Lindblom, A. (1997). Tumorigenesis in colorectal tumors from patients with hereditary non-polyposis colorectal cancer. Hum Genet *101*, 51-55.

100. Hemminki, K., and Mutanen, P. (2001). Genetic epidemiology of multistage carcinogenesis. Mutat Res *473*, 11-21.

101. Lynch, H.T., Shaw, T.G., and Lynch, J.F. (2004). Inherited predisposition to cancer: a historical overview. Am J Med Genet C Semin Med Genet *129*, 5-22.

102. Papadopoulos, N., and Lindblom, A. (1997). Molecular basis of HNPCC: mutations of MMR genes. Hum Mutat *10*, 89-99.

103. Miller, S.A., Dykes, D.D., and Polesky, H.F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res *16*, 1215.

104. Canzian, F., Salovaara, R., Hemminki, A., Kristo, P., Chadwick, R.B., Aaltonen, L.A., and de la Chapelle, A. (1996). Semiautomated assessment of loss of heterozygosity and replication error in tumors. Cancer Res *56*, 3331-3337.

105. Muller, A., Edmonston, T.B., Dietmaier, W., Buttner, R., Fishel, R., and Ruschoff, J. (2004). MSI-testing in hereditary non-polyposis colorectal carcinoma (HNPCC). Dis Markers *20*, 225-236.

106. Huber, C.G., Oefner, P.J., and Bonn, G.K. (1993). High-resolution liquid chromatography of oligonucleotides on nonporous alkylated styrene-divinylbenzene copolymers. Anal Biochem *212*, 351-358.

107. Huber, C.G., Oefner, P.J., Preuss, E., and Bonn, G.K. (1993). High-resolution liquid chromatography of DNA fragments on non-porous poly(styrene-divinylbenzene) particles. Nucleic Acids Res *21*, 1061-1066.

108. Emmerson, P., Maynard, J., Jones, S., Butler, R., Sampson, J.R., and Cheadle, J.P. (2003). Characterizing mutations in samples with low-level mosaicism by collection and analysis of DHPLC fractionated heteroduplexes. Hum Mutat *21*, 112-115.

109. Lilleberg, S.L., Durocher, J., Sanders, C., Walters, K., and Culver, K. (2004). High sensitivity scanning of colorectal tumors and matched plasma DNA for

mutations in APC, TP53, K-RAS, and BRAF genes with a novel DHPLC fluorescence detection platform. Ann N Y Acad Sci *1022*, 250-256.

110. Young, J., Barker, M., Fraser, L., Walsh, M.D., Spring, K., Biden, K.G., Hopper, J.L., Leggett, B.A., and Jass, J.R. (2002). Mutation searching in colorectal cancer studies: experience with a denaturing high-pressure liquid chromatography system for exon-by-exon scanning of tumour suppressor genes. Pathology *34*, 529-533.
111. Yuan, Z.Q., Gottlieb, B., Beitel, L.K., Wong, N., Gordon, P.H., Wang, Q., Puisieux, A., Foulkes, W.D., and Trifiro, M. (2002). Polymorphisms and HNPCC: PMS2-MLH1 protein interactions diminished by single nucleotide polymorphisms. Hum Mutat *19*, 108-113.
112. Munemitsu, S., Albert, I., Souza, B., Rubinfeld, B., and Polakis, P. (1995). Regulation of intracellular beta-catenin levels by the adenomatous polyposis coli (APC) tumor-suppressor protein. Proc Natl Acad Sci U S A *92*, 3046-3050.
113. Charbonnier, F., Raux, G., Wang, Q., Drouot, N., Cordier, F., Limacher, J.M., Saurin, J.C., Puisieux, A., Olschwang, S., and Frebourg, T. (2000). Detection of exon deletions and duplications of the mismatch repair genes in hereditary nonpolyposis colorectal cancer families using multiplex polymerase chain reaction of short fluorescent fragments. Cancer Res *60*, 2760-2763.
114. Schouten, J.P., McElgunn, C.J., Waaijer, R., Zwijnenburg, D., Diepvens, F., and Pals, G. (2002). Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. Nucleic Acids Res *30*, e57.
115. Wehner, M., Mangold, E., Sengteller, M., Friedrichs, N., Aretz, S., Friedl, W., Propping, P., and Pagenstecher, C. (2005). Hereditary nonpolyposis colorectal cancer: pitfalls in deletion screening in MSH2 and MLH1 genes. Eur J Hum Genet *13*, 983-986.
116. Nakagawa, H., Hampel, H., and de la Chapelle, A. (2003). Identification and characterization of genomic rearrangements of MSH2 and MLH1 in Lynch syndrome (HNPCC) by novel techniques. Hum Mutat *22*, 258.
117. Deininger, P.L., and Batzer, M.A. (1999). Alu repeats and human disease. Mol Genet Metab *67*, 183-193.

118. Charbonnier, F., Olschwang, S., Wang, Q., Boisson, C., Martin, C., Buisine, M.P., Puisieux, A., and Frebourg, T. (2002). MSH2 in contrast to MLH1 and MSH6 is frequently inactivated by exonic and promoter rearrangements in hereditary nonpolyposis colorectal cancer. Cancer Res *62*, 848-853.

119. Grover, D., Mukerji, M., Bhatnagar, P., Kannan, K., and Brahmachari, S.K. (2004). Alu repeat analysis in the complete human genome: trends and variations with respect to genomic composition. Bioinformatics *20*, 813-817.

120. Batzer, M.A., and Deininger, P.L. (2002). Alu repeats and human genomic diversity. Nat Rev Genet *3*, 370-379.

121. Tischfield JA, Loss of heterozygosity or: how I learned to stop worrying and love mitotic recombination., Am. J. Hum.Genet 1997;61:995-999.

122. Lamlum H, Ilyas M, Rowan A, Clark S, Johnson V, Bell J, et al. The type of somatic mutation at APC in familial adenomatous polyposis is determined by the site of the germline mutation: a new facet to Knudson's 'two-hit' hypothesis. Nat Med 1999;5(9):1071-5.

APPENDIX I

Sequences and annealing temperatures of all primers

Table 1: Primer sequences applied for microsatellite marker analysis

Map Gene	Marker Name	Direction	Sequence	Product length	Annealing TmºC	Labelled Dye
MSH2	D2s288	Forward primer	agggccttgctctggatt	276-284	52	6-FAM
		Reverse primer	ggccagtgattgttcccc			
	D2s2227	Forward primer	gacgtgtccatctctgaat	207-221	52	6- FAM
		Reverse primer	gcagtttctcggaataacca			
	D2s 123	Forward primer	aaacaggatgcctgccttta	197 -227	52	6-FAM
		Reverse primer	ggactttccacctatgggac			
	D2s2369	Forward primer	ctgacctgaacttgtgcc	242-256	52	6-FAM
		Reverse primer	tggggctttccacatt			
	Bat 26	Forward primer	tgactacttttgacttcagcc	100-120	55	HEX
		Reverse primer	aaccattcaacatttttaaccc			
MLH1	D3s 1597	Forward primer	agtacaaatacacacaaatgtctc	162-180	50	6-FAM
		Reverse primer	gcaaatcgttcattgct			
	D3s3611	Forward primer	gctacctctgctgagcat	107-137	50	6-FAM
		Reverse primer	tagcaagactgttgggg			
	D3s3594	Forward primer	caatgggctcatcgca	261-279	50	HEX
		Reverse primer	cttggaatagtgggccaga			
	D3s 3601	Forward primer	cagttaccttgatagactggtagtg	239-253	50	6-FAM
		Reverse primer	gagatttagttgactcacccac			
	D3s3589	Forward primer	aagcaatattttctaccactttct	235-245	50	HEX
		Reverse primer	tctgagccaccagcac			
APC	D5s299	Forward primer	gctattctctcaggatcttg	156-182	55	6-FAM
		Reverse primer	gtaagccaggacaagatgacag			
	D5S82	Forward primer	cccaattgtatagatttagaagtc	169-179	55	HEX
		Reverse primer	cccaattgtatagatttagaagtc			
	D5s346	Forward primer	atgaccaccaggtaggtgtatt	215	55	HEX
		Reverse primer	actcactctagtgataaatcggg			
	D5s318	Forward primer	agcagataagacagtattactagtt	96-106	55	6-FAM
		Reverse primer	tctagaggatcttccctctt			
	Bat 25	Forward primer	ccatcggtagaactaatttc	116-128	55	HEX
		Reverse primer	tcgcctccaagaatgtaagt			
	Bat 40	Forward primer	attaacttcctacaccacaac	128	55	6-FAM
		Reverse primer	gtagagcaagaccaccttg			

Table 2: Primer sequences of *MLH1* exons for QMPA

***MLH1* exons**	**Forward primers 5'-3'**	**Reverse primers 5'-3'**	**Product length(bp)**
1	ttcggcggctggacgagac	tcgtagcccttaagtgagc	134
2	tacattagagtagttgcaga	agagaaaggtcctgactc	224
3	taacagaaagaagatctgg	acaatgtcatcacaggagg	157
4	ctttccctttggtgaggtga	attactctgagacctagg	224
5	gattttctcttttccccttggg	acaaagcttcaacaatttactct	187
6	ttctgttcaggtggaggacc	tggactgtacctgccaacaa	112
7	caagcagcaaacttacaaga	accaccaccaactttatgag	175
8	cactacccaatgcctcaacc	acatgattcacgccacagaa	288
9	gggaaggaaccttgtgtttt	cataaaattccctgtgggtg	266
10	gttttgaactggttgctttc	gtggtgtatgggattcactc	247
11	tttgaccactgtgtcatctg	acctgggtgaagtacatcct	271
12	aatccacaacaagtctgacc	aaatgcatcaagcttctgtt	107
13	tttgagtctccaggaagaaa	ctgtagtgccactctgacaa	139
14	ttttgttttgcagttctcc	tagctctgcttgttcacaca	157
15	cttctcccattttgtccc	aaatttcagaagtgaaaagga	177
16	ttgctccttcatgttcttgc	tccaaagagaaatagtctgc	185
17	taccccttctgattgacaac	atctgcttgatcactgacct	92
18	agtctgtgaatctccgttttaga	cagtgtgcatcaaccactgta	210
19	aatcctcttgtgttcaaggc	tgcaggatatttccaatcttc	298

Table 3: Primer sequences of *MSH2* exons of QMPA

MSH2 exons	Forward primer	Reverse primer	product length(bp)
1	ttcgtgcgcttcttcag	acccccctgggtcttgaac	134
2	cttaaggagcaaagaatctgc	atgccacctaccaatcattc	154
3	tatattgtcagcttccattggt	gggtaaaacacattcctttgg	186
4	catcatatcagtgtcttgcac	cttgaggtcctgataaatgtc	221
5	tactgtctgcggtaatcaagt	taaaggtaagggctctgac	131
6	ttaatgagcttgccattct	gagaggctgcttaatccac	151
7	acaagcagcaaacttacaaga	gaccaccaccaactttatgag	175
8	ttccaagtttcaggaaatga	gctttcttaaagtggcctttg	111
9	ctgacttggaaaagaagatgc	cgggcttgtttaaatgacatc	226
10	tttcgtgtaacctgtaaggaa	ggggctatttaacaaatggtg	256
11	agatactttggatatgtttca	ccaggtgacattcagaaca	225
12	tttaggaaatgggttttgaat	atgcctggatgcttttaat	271
13	cattgtggactgcatcttagc	caaagtccacaggaaaacaac	340
14	gctacgatggatttgggttag	tttcccattaccaagttctg	238
15	gvctgtctcttctcatgctgt	tgatagcacttctttgctgct	214
16	tctgtccaaggtgaaacaaa	cccattactgggatttttcac	160

Table 4 : Primer sequences of the *APC* mutation cluster region (MCR). PCR products range between 90-150bp in length. All fragments share the same annealing temperature 60°C.

MCR Fragments	Forward primer 5'-3'	Reverse primer 5'-3'
1	tgcaaagtttcttctattaaccaa	atttaggtgacactattctgcttcctgtgtcgtctg
2	ttcattatcatctttgtcatcagc	atttaggtgacactattggaacttcgctcacaggat
3	gcagaaataaaagaaaagattggaa	atttaggtgacactatctttgtgcctggctgattct
4	ctagaaccaaatccag cagact	atttaggtgacactatgaacatagtgttcaggtggactttt
5	agcgaaatctccctccaaaa	atttaggtgacactatctggcaatggaacgactctc
6	cccactcatgtttagcagatg	atttaggtgacactatgtttgtccagggctatctgg
7	tggaatggtaagtggcattat	atttaggtgacactatcagcagtaggtgctttatttttagg
8	tcctcaaacagctcaaacca	atttaggtgacactatagcatctggaagaacctgga
9	aagcaagctgcagtaaatgct	atttaggtgacactatatggctcatcgaggctca
10	aagtactccagatggattttcttg	atttaggtgacaqctatggctgctctgattctgtttca
11	atgcctccagttcaggaaaa	atttaggtgacactattcaatatcatcatcatctgaatcatc
12	aaaaactattgactctgaaaaggac	atttaggtgacactatggtggaggtaattttgaagca

Table 5: Primer sequences of *MYH* used for *MYH* dHPLC analysis and sequencing.

MYH exons	Forward primers 5'-3'	Reverse primer 5'-3'
1	tgaaggctacctctgggaag	aggagacggaccgcaag
2	ggctgggtctttttgtttca	gggccacaacctagttcctt
3a	actgtgtcccaagaccctgat	ttggtcgtaccagcttagca
3b	agctgaagtcacagccttcc	cacccactgtccctgctc
4	cctccaccctaactcctcatc	aaagtggccctgctctcag
5	caggtcagcagtgtcctcat	gtctgacccatgacccttcc
6	caggtcagcagtgtcctcat	gtctgacccatgacccttcc
7	cgggtgatctctttgacctc	gttcctaccctcctgccatc
8	tcttgagtcttgcactccaatc	aaagtgggggtgggctgt
9	gctaactctttggcccctct	cacccttgttaccccaacat
10	ctgcttcacagcagtgttcc	gacttctcactgccccttcc
11	acactcaaccctgtgcctct	ggaatggggcttctgactg
12a	acttggcttgagtagggttcg	ggctgttccagaacacaggt
12b	gagtggtcaacttccccaga	cacgcccagtatccaggta
13	agggaatcggcagctgag	gctattccgctgctcactta
14	aggcctatttgaaccccttg	caacaaagacaacaaaggtagtgc
15	ccctcacctccctgtcttct	tgttcacccagacattcgtt
16a	actacaaggcctccctccttc	gctgcactgttgaggctgt
16b	gccagcaagtcctggataat	acatagcgagacccccatct

Table 6: PTT primer sequences for *MSH2/MLH1*

MSH2	Forward primer	Reverse Primer	Product length
Fragment 1	ggatcctaatacgactcactatagggagaccaccat ggcggtgcagccgaa	catcctgggcttcttcata tctgttttat	1.7kb
Fragment 2	ggatcctaatacgactcactatagggagaccaccc ttggcttggacctggcaaac	tcaatattaccttcattccattactggg attt	1.3kb

MLH1	Forward primer	Reverse primer	Product length
Fragment 1	ggatcctaatacgactcactatagggagaccaccat gac atctagacgtttcctt	aaatgcatcaagcttctgttc	1.2kb
Fragment 2	ggatcctaatacgactcactatagggagaccaccat ggtgcagcagcacatcgaga	cacagtgcataaataaccat	1.3kb

APPENDIX II:

dHPLC conditions of different genes

Table 7: dHPLC analysis conditions for *MYH* exon. Each exons has two different melting temperatures Tm1 (°C) and Tm2 (°C)

MYH exons	Tm 1(°C)	Tm2(°C)
1	62.7	66.7
2	60.9	63.9
3a	61.7	62.7
3b	62.5	64.5
4	61.4	62.4
5	61.7	63.7
6	60.4	63.4
7	60.7	64.7
8	61.1	66.1
9	60.9	64.9
10	61.5	65.5
11	62.2	64.2
12a	62.5	64.5
12b	62.1	65.1
13	63.6	65.6
14	59.6	63.6
15	59.6	63.6
16a	60.8	62.8
16b	56.3	61.3

Table 8: dHPLC analysis conditions (heteroduplex fragment analysis) for the *APC* mutation cluster region. Each fragment has two different melting temperatures Tm1(°C) and Tm2(°C)

MCR Fragments	Tm1(°C)	Tm2 (°C)
1	54.3	56.3
2	57.6	58.1
3	58	60.5
4	59.4	60
5	59.7	60.5
6	59.1	59.6
7	58.5	60.5
8	58.5	60.5
9	59.2	60.2
10	58	58.8
11	56.1	57.9
12	55.5	57.8

APPENDIX III

Multiplex Ligation dependent Probe Amplification –MLPA

MLPA protocol

1. **DNA-Denaturation** :
   Dilute the DNA-sample (250 ng DNA) with TE to 5 µl.
2. Heat 5 minutes at 98°C; Cool to 25°C before opening the thermal cycler.
3. **Hybridisation of SALSA probes** : Add 1.5 µl SALSA Probe-mix (black cap) add 1.5 µl MLPA buffer (yellow cap). Mix with care. Incubate 1 minute at 95°C,
4. hold on 60°C for 16 hours (over night) .
5. **Ligation reaction** : Reduce the temperature of the thermal cycle to 54°C. add 32 µl Ligase-65 mixt to each sample and mix.
   Mix: Ligase-65 mix (made less than 1 hr. before use and stored on ice) includes 3 µl Ligase-65 buffer A (transparent cap) and 3 µl Ligase-65 buffer B (white cap) ,plus 25 µl Water, 1 µl Ligase-65 (brown cap)
6. Incubate it 15 minutes at 54°C, then heat 5 minutes at 98°C.
7. **PCR amplification** : Mix: 1 µl SALSA PCR-primers (purple cap) and 1 µl SALSA enzyme dilution buffer (blue cap) were added to 5ul MLPA ligation reaction .Then 0.25ul SALSA PCR enzyme and distilled water were added to fill the final volume up to 20 µl .
8. PCR reaction: 95°C 5minutes, 95°C30 seconds; 60°C30 seconds; 72°C 60 seconds, run the program for 35 cycles ,then end the incubation at 72°C for 20 minutes.
9. **Fragment analysis on genetic analyzer**
   Sample preparation: 2µl MLPA PCR products mixed with 12 µl deionized formamide (Applied Biosystems, Rotkreuz, Switzerland) and 0.5 µl ROX 500 (Applied Biosystems, Rotkreuz, Switzerland).The mixture was denatured at 90°C 2 minutes, then puton ice for 2 minutes. The samples were loaded on ABI PRISM 310 sequence analysis machine .
   Sample analysis condition : GC Matrix A

Table 9: Detail information of MLH1/MSH2 salsa probes

	Chromosome	Location

Length (nt)	MLPA MLH1/MSH2 probe	Control	MLH1	MSH2
64-70-76-82	DQ-control bands*			
94	Synthetic Control probe	2q14		
130	Control probe0797-L0463	5q31		
136	Control probe 0981-L0566	10p11		
142	MLH1 probe 0886-L0474		exon 1	
148	MSH2 probe 1027-L0599			exon 1
154	MLH1 probe 1008-L0577		exon 2	
160	MSH2 probe 0906-L0494			exon 2
166	MLH1 probe 0888-L0476		exon 3	
172	MSH2 probe 1029-L0601			exon 3
178	MLH1 probe 0889-L0477		exon 4	
184	MSH2 0908-L0496			exon 4
193	Control probe 0976-L0563	11p13		
202	MLH1 probe 0890-L0478		exon 5	
211	MSH2 probe 0909-L0497			exon 5
220	MLH1 probe 0891-L0479		exon 6	
229	MSH2 probe 0910-L0498			exon 6
238	MLH1 probe 0892-L0480		exon 7	
247	MSH2 probe 0911-L0499			exon 7
256	MLH1 probe 0893-L0481		exon 8	
265	MSH2 probe 0912-L0582			exon 8
274	MLH1 probe 0894-L0482		exon 9	
283	Control probe 0438-L0003	17q21		
292	MSH2 probe 0913-L0583			exon 9
301	MLH1 probe 0895-L0483		exon 10	
310	MSH2 probe 0914-L0584			exon 10
319	MLH1 probe 0896-L0484		exon 11	
328	MSH2 probe 0915-L0503			exon 11
337	MLH1 probe 0897-L0485		exon 12	
346	MSH2 probe 0916-L0504			exon 12
355	MLH1 probe 0898-L0486		exon 13	
364	MSH2 probe 1013-L0575			exon 13
373	Control probe 0681-L0154	4q25		
382	MLH1 probe 0899-L0586		exon 14	
391	MSH2 probe 0918-L0506			exon 14
400	MLH1 probe 0900-0488		exon 15	
409	MSH2 probe 0919-L0585			exon 15
418	MLH1 probe 1009-L0576		exon 16	
427	MSH2 probe 1053-L0627			exon 16
436	MLH1 probe 1030-L0602		exon 17	
445	MLH1 probe 1031-L0603		exon 18	
454	MLH1 probe 0904-L0492		exon 19	
463	Control probe 0979-L0568	10p14		
472	Control probe 0980-L0567	11p12		

Table 10: Sequences detected by the probes and distances between the *MLH1* exons.

Length (nt)	Salsa Probe #	MLH1 exon	Ligationsite	Sequence at Ligation site 3q22.1	Distance to next exon
		startcodon	*22-24*		
142	0886-L0474	exon 1	28-29	AAAATGTCGT-TCGTGGCAGG	3.1 kb
154	1008-L0577	exon 2	206-207	TTCAGATCCA-AGACAATGGC	4.3 kb
166	0888-L0476	exon 3	286-287	CAGTCCTTTG-AGGATTTAGC	3.4 kb
178	0889-L0477	exon 4	350-351	TAAGCCATGT-GGCTCATGTT	2.6 kb
202	0890-L0478	exon 5	450-451	ACCATGTGCT-GGCAATCAAG	1.8 kb
220	0891-L0479	exon 6	487-488	GAGGACCTTT-TTTACAACAT	3.0 kb
238	0892-L0480	exon 7	587-588	ACAATGCAGG-CATTAGTTTC	0.2 kb
256	0893-L0481	exon 8	641-642	TTAGGACACT-ACCCAATGCC	2.3 kb
274	0894-L0482	exon 9	743-744	TAGCCTTCAA-AATGAATGGT	3.0 kb
301	0895-L0483	exon 10	842-843	CCTTGAGAAAAGCCATAGAA	2.8 kb
319	0896-L0484	exon 11	1013-1014	AGCACATCGAGAGCAAGCTC	5.4 kb
337	0897-L0485	exon 12	1305-1306	GCAGCAAGATGAGGAGATGC	2.9 kb
355	0898-L0486	exon 13	1488-1489	CCGAAAGGAAATGACTGCAG	11.4 kb
382	0899-L0586	exon 14	1622-1623	TGGGCTGTGTGAATCCTCAG	2.0 kb
400	0900-0488	exon 15	1715-1716	AGATACTCAT-TTATGATTTT	5.2 kb
418	1009-L0576	exon 16	1790-1791	TGCTTGCCTTAGATAGTCCA	1.0 kb
436	1030-L0602	exon 17	1995-1996	CTTCATTCTTCGACTAGCCA	0.4 kb
445	1031-L0603	exon 18	2092-2091 (Rev.)	CCTCAGATATGTACTGCTTC	1.6 kb
454	0904-L0492	exon 19	2177-2178	TGGAACACATTGTCTATAAA	

Table 11: Sequences detected by the probes and distances between the *MSH2* exons.

Length (nt)	SALSA MLPA probe	MLH1 exon	Ligation site U03911	Partial sequence (20 nt adjacent to ligation site)	Distance to next exon
		startcodon	*4-6*		
190	6343-L5729	exon 1	(-28 - -29)	GTCGCGCATTTTCTTCAACC	0.1 kb
148*	1027-L0599	exon 1	130-131	GGCGACTTCTATACGGCGCA	5.1 kb
172	1029-L0601	exon 3	454-455	GTGGGTGTTAAAATGTCCGC	2.3 kb
184	0908-L0496	exon 4	741-742	GTTGTTGAAAGGCAAAAAGG	1.8 kb
211	0909-L0497	exon 5	821-822	TGTCTGCGGTAATCAAGTTT	2.0 kb
229	0910-L0498	exon 6	996-997	CTTGCTGAATAAGTGTAAAA	13.5 kb
247	0911-L0499	exon 7	1172-1173	TTAACCGACTTGCCAAGAAG	15.7 kb
265	0912-L0582	exon 8	1327-1328	CCTCTTACTGATCTTCGTTC	17.4 kb
292	0913-L0583	exon 9	1414-1415	TTCCTTGTAAAACCTTCATT	3.6 kb
310	0914-L0584	exon 10	1536-1537	CAAACAGATTAAACTGGATT	4.3 kb
328	0915-L0503	exon 11	1710-1709 (rev.)	TTCTTCATATTCTGTTTTAT	4.2 kb
346	0916-L0504	exon 12	1858-1859	CCTGTTCCATATGTACGACC	1.3 kb
364	1013-L0575	exon 13	2074-2075	ATGGCCCAAATTGGGTGTTT	1.9 kb
391	0918-L0506	exon 14	2281-2282	TACGATGGATTTGGGTTAGC	2.4 kb
409	0919-L0585	exon 15	2564-2565	TTGAGGAGTTTCAGTATATT	1.9 kb
427	1053-L0627	exon 16	2640-2641	GTTTCAGCAAGGTGAAAAAA	

CURRICULUM VITAE

Personal Information

Name: Zhang Jian
Age: 37
Marital Status: Married
Nationality: P.R.China

Education and Experiences

PhD candidate: (2002-2006)

Research Group Human Genetics, Department of Medical Genetics, Center for Biomedicine, UKBB, University of Basel, Basel

PhD canditate: (2001-2002)

Institute of Veterinary Biochemistry and Molecular Biology, University of Zurich

Postgraduate Fellowship (1998- 2001)

EPH department, School of Medicine, Yale University

Research Scientist (1997-1998)

Department of Research and Development, Fuxing Company, Shanghai

Master Degree: (1994-1997)

Molecular Cell and Genetics group, Nanjing Normal University (joint student with Fudan University, Shanghai)

Research Technician (1991-1994)

Laboratory of Molecular Genetics, Huadong Medical Institute.

Awards

1. Postgraduate fellowship, Yale University, School of Medicine, 1998
2. Outstanding graduate student, Nanjing Normal University, 1997
3. Outstanding graduate student, Jiangsu Province, 1997
4. Outstanding Thesis, the *fifth china international thesis competition,*1997
5. Fellowship of Chen Yifeng, Nanjing Normal University, 1996

PUBLICATIONS

1. **Zhang J**, Lindroos A, Ollila S, Russell A, Marra G, Mueller H, Peltomaki P, Plasilova M, Heinimann K. Gene conversion is a frequent mechanism of inactivation of the wild-type allele in cancers from *MLH1/MSH2* deletion carriers. Cancer Res, 2006; *66*(2)*:* 659-664.
2. Russell A, **Zhang J**, Luz J, Hutter P, Chappuis P, Berthod CR, Maillet P, Mueller H, Heinimann K. Prevalence of MYH germline mutations in Swiss APC mutation-negative polyposis patients. Int J Cancer, 2006; *118*(8)*:* 1937-1940.(Equally Contributed)
3. **Zhang , J**, Lingyuan Zhang , Xiran Zhang, Jianhua Chai.The high expression of Human beta-nerve growth factor in E.coli. Wei Sheng Wu Xue Bao. 1997 Dec;37(6):429-33. Chinese
4. **Zhang J** , Zhang L, Xiran Z..The biological activity demonstration of recombinant human NGF. Biotechniques in Pharmacology, 21(5): 98-104, 1998.Chinese
5. Sieber OM, Segditsas S, Knudsen AL, **Zhang J**, Luz J, Rowan AJ, Spain SL, Thirlwell C, Howarth KM, Jaeger EE, Robinson J, Volikos E, Silver A, Kelly G, Aretz S, Frayling I, Hutter P, Dunlop M, Guenther T, Neale K, Phillips R, Heinimann K, Tomlinson IP. . Disease severity and genetic pathways in attenuated familial adenomatous polyposis vary greatly but depend on the site of the germline mutation.

   Gut. 2006 Oct;55(10):1440-8.
6. Pasilova M, **Zhang J**, Okhowat R., Marra G., Mettler M., Mueller H.J. Heinimann K., A *de novo* MLH1 germ line mutation in a 31 year old colorectal cancer patient **.** Genes Chromosomes Cancer. 2006 Dec;45(12):1106-10.
7. Luna C, Hoa NT, **Zhang J,** Kanzok SM, Brown SE, Imler JL, Knudson DL, Zheng L. Characterization of three Toll-like genes from mosquito Aedes aegypti. Insect Mol Biol. 2003 Feb;12(1):67-74.
8. Luna C., Wang X., Huang ., **Zhang J,** Zheng L, Characterization of four Toll related genes during development and immune responses in *Anopheles gambia*. Insect Biochemistry and Molecular Biology 32: 1171-1179, 2002

CONFERENCES

Poster presentation

1. **ESHG (European Human Genetics Conference), Prague, May, 2005**
   Title: Second hit analysis in tumors from HNPCC patients carrying novel large genomic deletions in *MLH1/MSH2*
2. **ESHG (European Human Genetics Conference), Munich, June 2004**

   Title: Evaluation of different screening techniques to detect large genomic rearrangements in *MLH1/MSH2*
3. **Second Brain Research Interactive conference: Neuropeptides at the millennium, Miami, Aug, 1999**

   Title: The high expression of Human nerve growth factor and biological activity demonstration in E.coli

DECLARATION OF INDEPENDENCE

I declare that I wrote this thesis "Identification of novel germline mutations in hereditary colorectal cancer patients and characterization of somatic alterations in their tumors " with the help indicated and only handed it in to the faculty of science of the University of Basel and to no other faculty and no other university.