# ANALYSIS OF *hMLH1* GERMLINE MUTATIONS IN THREE TURKISH HEREDITARY NONPOLYPOSIS COLORECTAL CANCER KINDREDS



A THESIS SUBMITTED TO

THE DEPARTMENT OF MOLECULAR BIOLOGY AND GENETICS

AND

THE INSTITUTE OF ENGINEERING AND SCIENCE OF

BİLKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF MASTER OF SCIENCE



By

CEMALİYE AKYERLİ

July, 1998

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Assist. Prof. Marie D. Ricciardone

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Assist. Prof. Uğur B. Yavuzer

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Prof. Dr. Ufuk Gündüz

Approved for the Institute of Engineering and Science

Director of Institute of Engineering and Science
Prof. Dr. Mehmet Baray

## ABSTRACT

ANALYSIS OF *hMLH1* GERMLINE MUTATIONS IN THREE TURKISH HEREDITARY NONPOLYPOSIS COLORECTAL CANCER KINDREDS

Cemaliye Akyerli
M.S. in Molecular Biology and Genetics
Supervisor: Assist. Prof. Marie D. Ricciardone
July 1998, 96 Pages

Hereditary nonpolyposis colorectal cancer (HNPCC) is one of the most common genetic diseases in Western world. It is a clinical syndrome characterized by an inherited predisposition to early onset colorectal and an increased incidence of other cancers. The disease is caused by a germline defect in one of five human DNA mismatch repair genes, *hMLH1*, *hMSH2*, *hPMS1*, *hPMS2*, and *hMSH6*. Defects in *hMLH1* and *hMSH2* account for the majority of mutations found in HNPCC families. In this study, a variety of mutation detection methods were used to identify *hMLH1* germline mutations in three Turkish HNPCC kindreds.

Restriction enzyme analysis of genomic DNA was used to analyze five members of an HNPCC family with a previously described G884C mutation. The genotypes of all five individuals were determined by *Dde*I digestion and the results were confirmed by DNA sequence analysis. *Hph*I restriction enzyme analysis was used to analyze twenty-nine members of an unrelated HNPCC family for a previously identified A1652C mutation. Genotypes were determined for all individuals and the results were confirmed by DNA sequence analysis. Both of these restriction enzyme analyses are reliable, cost-effective methods that can be used in mutation screening programs for family members who request genetic counseling.

Single strand conformation polymorphism analysis (SSCP) was used to screen for unknown germline mutations. Nine DNA samples with defined mutations in the *hMLH1* gene were analyzed using several gel formulations and electrophoretic conditions to determine the most sensitive protocols. These protocols were then used for routine mutation detection.

In a third HNPCC family, for whom no mutation has yet been defined, the complete coding sequence of the *hMLH1* gene was screened by SSCP. Two exons, 7 and 15, showed an altered mobility compared to control sequences. The nucleotide sequence of these two exons was determined by automated fluorescence DNA sequence analysis. The differential mobility observed for exon 15 appears to be due to an intonic polymorphism in the control sample. Preliminary results for exon 7 show no difference between proband and control nucleotide sequences. Thus, the DNA mismatch repair defect in this kindred appears not to be in *hMLH1*. Further studies will focus on the analysis of *hMSH2*.

# ÖZET


KALITSAL POLİPOZ OLMAYAN KOLOREKTAL KANSERLİ ÜÇ TÜRK AİLESİNDE, *hMLH1* GENİNDEKİ EŞEY HÜCRESİ MUTASYONLARININ ANALİZİ

Cemaliye Akyerli
Moleküler Biyoloji ve Genetik Yüksek Lisans
Tez Yöneticisi: Yardımcı Doçent Dr. Marie D. Ricciardone
Temmuz 1998, 96 Sayfa

Kalıtsal polipoz olmayan kolorektal kanseri, Batı ülkelerinde en sık rastlanan genetik hastalıklardan biridir. Bu klinik sendrom, erken yaşta ortaya çıkan kolorektal kanseri ve diğer kanserlerin ortaya çıkmasını artıran kalıtsal yatkınlıkla karakterizedir. Beş, insan yanlış eşleşme DNA tamir genlerinden (*hMLH1*, *hMSH2*, *hPMS1*, *hPMS2*, ve *hMSH6*) birinde olan eşey hücresi bozukluğu, bu hastalığın ortaya çıkmasına neden olur. Kalıtsal polipoz olmayan kolorektal kanserli ailelerde bulunan mutasyonların çoğunluğu, *hMLH1* ve *hMSH2* genlerinde görülmektedir. Bu çalışmada, kalıtsal polipoz olmayan kolorektal kanserli üç Türk ailesinde bulunan eşey hücresi mutasyonlarının tanımlanması için çeşitli mutasyon tarama metodları kullanılmıştır.

Genomik DNA'nın restriksiyon enzim analizi, daha önceden tanımlanmış G884C mutasyonu taşıyan bir kalıtsal polipoz olmayan kolorektal kanserli ailenin beş ferdinin incelenmesinde kullanılmıştır. Bu beş bireyin genotipleri, *Dde*I enzim analizi ile belirlenmiş ve sonuçlar DNA dizi analizi ile doğrulanmıştır. Daha önceden tanımlanmış A1652C mutasyonunu taşıyan diğer bir kalıtsal polipoz olmayan kolorektal kanserli ailenin yirmidokuz ferdi, *Hph*I restriksiyon enzim analizi ile tanımlanmıştır ve sonuçlar DNA dizi analizi ile doğrulanmıştır. Her iki analiz de genetik danışma isteyen bu ailelerin mutasyon taramasında, güvenilir ve hesaplı tekniklerdir.

Tek iplikçikli yapısal çeşitlilik analizi, bilinmeyen eşey hücresi mutasyonlarının taranmasında kullanılmıştır. En hassas protokollerin belirlenmesi için, daha önceden mutasyonları tanımlanmış dokuz DNA örneği, çeşitli jel içeriği ve elektroforez koşullarında çalışılmıştır. Bu protokoller, daha sonra rutin mutasyon tarama çalışmalarında kullanılmıştır.

Mutasyonu henüz tanımlanmamış olan üçüncü bir kalıtsal polipoz olmayan kolorektal kanserli ailenin, *hMLH1* geninin bütün kodlayıcı dizisi, bu metodla taranmıştır. Ekson 7 ve 15, kontrol diziye göre farklılık göstermiştir. Bu iki eksonun nükleotid dizileri, floresan otomatik DNA analizi ile belirlenmiştir. Ekson 15 deki farklılığın, kontrol örnekteki intronik bir polimorfizimden kaynaklandığı görülmektedir. Ekson 7 için ilk veriler, probanla kontrol nükleotid sekans arasında bir farklılık olmadığını göstermektedir. Sonuç olarak, bu ailedeki yanlış eşleşme DNA tamir bozukluğunun hMLH1 geninden kaynaklanmadığı düşünülmektedir. İleriki çalışmalar, *hMSH2* geninin incelenmesi üzerinde yoğunlaşacaktır.

**To my grandmother Hatice M. Ömer**

**For as long as there is memory, you will live in my heart**

## ACKNOWLEDGEMENT

It is my pleasure to express my deepest gratitude to my advisor Assist. Prof. Marie D. Ricciardone for her guidance, encouragement, laboratory discipline and invaluable efforts throughout my thesis work. I should also add that she is perfect in optimizing experiments.

I wish to express my thanks to Prof. Mehmet Öztürk for his support and suggestions and also for providing cell lines from his old laboratory. I would like to thank Assoc. Prof. Tayfun Özçelik for his help and encouragement.

I would like to address my special thanks to Birsen Cevher for her friendship, sharing her experience with me with patience and also for the automated sequencing. I would like to thank biologist Lütfiye Mesci for her unconditioned help and also to Hilal Özdağ for helping me whenever I needed. You are a wonderful friend.

A special thanks goes to my housemate Arzu for her closeness that makes us more like family than friends. The God gave me a wonderful friendship with you that I am very thankful for it.

I want to thank Tolga and Tuba for everything that make me glad we are friends. Both of you will always hold a very special place in my heart.

Special thanks to Reşat for taking care of me like my big brother. Sharing friendship with you is sweeter than ever.

I would like to thank Burçak, Buket, Emre Sayan, Berna, Neco, Çağla, Alper, Hani and Anastasia for extending their helping hands whenever I needed.

I should thank to all friends in laboratory for their warm friendships, suggestions and help.

Thank you very much my dearest friend Harun for your understanding. You have never hesitated helping and listening to me. What really makes you special is simply being you. I appreciate your friendship.

My sincere thanks to my parents for their unconditioned support and interest and thanks to my sister Hatice for giving a great support, encouragement and making life more enjoyable. I couldn't have a nicer sister not even if I'd picked you out myself. Having a sister like you who just couldn't be loved more means so much.

# TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

## ABBREVIATIONS

| | |
|---|---|
| A | absorbance |
| APS | ammonium persulfate |
| ATP | adenine triphosphate |
| bisacrylamide | N, N, methylene bis-acrylamide |
| bp | base pair |
| cDNA | complementary DNA |
| dATP | adenosine deoxyribonucleoside triphosphate |
| dCTP | cytosine deoxyribonucleoside triphosphate |
| $ddH_2O$ | deionized water |
| ddNTP | dideoxynucleotide triphosphate |
| dGTP | guanosine deoxyribonucleoside triphosphate |
| DNA | deoxyribonucleic acid |
| DNase | deoxyribonuclease |
| dNTP | deoxynucleotide triposphate |
| dTTP | thymine deoxyribonucleoside triphosphate |
| EDTA | ethylenediaminetetraacetic acid |
| EtBr | ethidium bromide |
| EtOH | ethanol |
| g | gram |
| HNPCC | hereditary nonpolyposis colorectal cancer |
| kb | kilobase |
| M | molar |
| MBq | million becquerel |
| min | minute |
| ml | milliliter |
| mM | milimolar |
| µl | microliter |
| MMR | mismatch repair |
| mRNA | messengerRNA |
| NaOAc | sodium acetate |
| ng | nanogram |
| nm | nanometer |
| PCR | polymerase chain reaction |
| rpm | revolution per minute |
| RNA | ribonucleic acid |
| SDS | sodium dodecyl sulphate |
| SSCP | single strand conformation polymorphism |
| TBE | tris-boric acid-EDTA |
| TEMED | N,N,N,N-tetramethyl-1,2 diaminoethane |
| U | unit |

| | |
|---|---|
| UV | ultraviolet |
| v | volt |
| v/v | volume for volume |
| µCi | microCurie |
| µg | microgram |
| µl | microliter |

# 1. Introduction

## 1.1. Colorectal Cancer

Colorectal cancer is a significant cause of morbidity and mortality in Western populations. Approximately 50% of the Western population develops a colorectal tumor by age 70 and in approximately 10% of these individuals, the tumor will become malignant (Kinzler and Vogelstein, 1996). Colorectal cancer develops as a result of the pathologic transformation of normal colonic epithelium to an adenomatous polyp and ultimately an invasive cancer. This transformation is a multistep progression that requires years, possibly decades, and depends on a number of recently characterized genetic alterations.

Genetic alterations that confer a proliferative advantage on a specific cell and lead to cancer occur within three classes of genes: (1) proto-oncogenes, which when mutated (oncogenes), promote uncontrolled cell growth; (2) tumor suppresor genes, which when mutated, fail to regulate cell proliferation; and (3) DNA repair genes, which when mutated, fail to ensure fidelity of DNA replication, eventually leading to mutations in proto-oncogenes and tumor suppressor genes (Gryfe et al., 1997). Telomerase activity, which maintains the integrity of the chromosome ends and thereby immortalizes the cell, has been detected in almost all cancers, including colorectal cancer (Gryfe et al., 1997). Other genetic alterations that allow transformed colorectal epithelial cells to escape cell cycle arrest or apoptosis have also been recognized (Gryfe et al., 1997). In addition, hypomethylation or

hypermethylation of DNA sequences may alter gene expression without nucleic acid mutation (Gryfe et al., 1997).

Epidemiologic studies strongly suggest that the diet can influence colorectal cancer incidence. Lipids are thought to be among the critical dietary components because a higher rate of colorectal cancer has been associated with diets containing large amounts of red meat. Moreover, it has been shown that nonsteroidal anti-inflammatory drugs, that inhibit the cyclooxygenases that metabolize the lipid arachidonic acid, can prevent tumor formation and even cause existing colorectal tumors to regress (reviewed in Kinzler and Vogelstein., 1996).

Although most colorectal cancers are sporadic cancers that result from the accumulation of multiple genetic changes, some colorectal cancers have a hereditary genetic factor. The two most common familial colorectal cancers are familial adenomatous polyposis (FAP) and hereditary nonpolyposis colorectal cancer (HNPCC).

### 1.1.1. Familial Adenomatous Polyposis (FAP)

FAP is a rare, autosomal dominant inherited syndrome caused by germline mutations of the adenomatous polyposis coli (APC) tumor-suppressor gene located on chromosome 5q21-22. It is associated with almost complete penetrance and has an estimate germline frequency of 1 in 10,000 in the general population (Gryfe et al., 1997). The disease is characterized by the development of hundreds to thousands of small benign polyps (adenomas) in the second or third decade of life. Left untreated, these polyps develop into large malignant cancers (carcinomas) at an average age of 44 years.

Patients with germline mutations of APC do not necessarily develop colorectal cancer but have a much greater risk compared to general population. Somatic inactivation of the wild-type APC allele is the first event and the rate-limiting step in FAP colorectal carcinogenesis. Mutation leading to deregulation of the K-*ras* proto-oncogene is also thought to be an early event in colon cancer formation. Loss of heterozygosity on chromosome 18q and consequent loss of tumor suppressor genes occurs later in the sequence of development from adenoma to carcinoma. Finally, mutation of the p53 tumor suppressor gene on chromosome 17p appears to be a late phenomenon in colorectal carcinogenesis, which probably allows the growing tumor with multiple genetic alterations to evade cell cycle arrest and apoptosis (reviewed in Kinzler and Vogelstein, 1996).

### 1.1.2. Hereditary Nonpolyposis Colorectal Cancer (HNPCC)

HNPCC is a relatively common autosomal dominant disease. It affects 1 in 200 to 1,000 individuals and accounts for 3% of all colorectal cancers (Boland, 1998). HNPCC patients have an increased risk of colorectal cancer, that is distinguished from sporadic colorectal cancer by a younger age of onset (mid-40s). The risk of colorectal cancer in HNPCC individuals (penetrance) has been estimated at approximately 80% (Gryfe et al., 1997). In addition to colorectal cancers, HNPCC individuals have an increased incidence of other cancers, such as endometrium, ovary, stomach, small intestine, hepatobiliary system, kidney and ureter. Two subsets of families have been described: Lynch syndrome I families show only colorectal cancers and Lynch syndrome II families show tumors in other organs. Families with colorectal cancer are designated as HNPCC families if they fulfil the following criteria defined by the International Collaborative Group on HNPCC (Amsterdam criteria):

(1) three or more relatives with histologically verified colorectal cancer, one of whom is a first-degree relative of the other two, (2) colorectal cancer affecting at least two generations, and (3) one or more colorectal cancer cases diagnosed before age 50 (Lynch et al., 1993).

Cancers associated with HNPCC tend to have a high degree of genomic instability, which is manifested as alterations in the lengths of simple repeat elements or microsatellites (Ionov et al., 1993; Thibodeau et al., 1993; Peltomaki et al., 1993b). This genomic instability is called microsatellite instability. A similar phenomenon had also been observed in bacteria and yeast with mutant DNA mismatch repair genes (Levinson and Gutman, 1987; Strand et al., 1993). These observations suggested that human homologs of the DNA mismatch repair genes might be involved in HNPCC. To date, defects in five DNA mismatch repair genes have been linked to HNPCC: *hMSH2* (Fishel et al., 1993; Leach et al., 1993), *hMLH1* (Lindblom et al., 1993; Bronner et al., 1994), *hPMS1* (Papadopoulos et al., 1994), *hPMS2* (Nicolaides et al., 1994) and *hMSH6* (Akiyama et al., 1997). Because cancer develops in HNPCC when the DNA mismatch repair system fails, the gene products that function in eukaryotic mismatch repair are of great interest.

## 1.2. DNA Mismatch Repair

The term "mismatch repair" was initially coined to refer to a cellular activity capable of recognizing abnormal base pairs and correcting the sequence on one strand to restore a normal A•T or G•C pairing (Rhyu, 1996). This activity was also found to correct stretches of unpaired bases that result from insertion or deletion of nucleotides on one of the two DNA strands. Both prokaryotic and eukaryotic cells are capable of repairing mismatched base pairs in their DNA.

Mismatched base pairs in DNA can arise by several processes (Friedberg et al., 1995). A significant source of mismatched bases is DNA replication errors. Occasionally an incorrect nucleotide is incorporated into the DNA strand being synthesized. While the majority of these misincorporations are excised by the DNA polymerase proofreading 3'-5' exonuclease, approximately 1 in $10^9$ errors remain. The DNA mismatch repair system can repair approximately 99.5% of the mutations that escape proofreading, thus, decreasing the error rate to 1 in $10^{12}$ base pairs (Boland, 1998). In this case, the correct base is located in the parental strand of the newly replicated DNA and correction of the mismatch helps maintain the fidelity of the genetic information.

Another source of mismatched base pairs is heteroduplex formation between two homologous, but slight different, DNA molecules during recombination. Mismatches can also result when hairpins form between imperfect palindromes. Mismatched base pairs can also arise when deamination of 5-methylcytosine converts a G•5-mC base pair to a G•T base pair. Finally mismatched base pairs result when base analogs or chemically modified derivatives of normal bases are incorporated into DNA.

DNA mismatch repair plays a major role in two cellular processes: (1) repair of errors made during DNA replication or resulting from chemical damage to DNA and (2) processing of recombination intermediates that may yield new configurations of genetic markers. More recent studies have suggested that mismatch repair may also be crucial for: (3) regulation of recombination events between divergent DNA sequences that could result in different types of genetic instability; (4) nucleotide excision repair responsible for repair of physical/chemical damage to DNA and (5) participation in a cell cycle checkpoint control system by recognizing certain types of

DNA damage and triggering cell cycle arrest or other responses to DNA damage (reviewed in Kolodner, 1996).

### 1.2.1. DNA Mismatch Repair in Prokaryotes

The best understood DNA mismatch repair system is the methyl-directed mismatch repair system of *Escherichia coli*, a system that has been completely reconstituted using purified enzymes (Lahue et al., 1989). Mismatch repair is tightly coupled with DNA replication, so that mismatches formed during DNA replication are repaired using the methylated parental strand as template, resulting in a reduction of misincorporation errors. The methylation signals specifying the parental strand can be located a considerable distance from the actual mismatch and thus, the excision tracts associated with this pathway can be large, $10^3$ bp or more, so the system is often referred to as long-patch DNA mismatch repair.

The *E. coli* MutHLS system repairs a broad spectrum of mispaired bases. It recognizes and repairs all single-base mispairs except C•C (Kolodner, 1996). It also repairs small insertion / deletion mispairs, although it may not efficiently recognize insertion / deletion mispairs that have more than 4 unpaired bases. Strand discrimination is determined by recognition of $N^6$-methylation of the adenine residue in the GATC palindrome. Immediately after DNA replication, the newly-synthesized daughter-strand DNA is undermethylated relative to the parental strand. This difference in methylation state between parental and daughter strands just behind the replication fork permits discrimination between the two strands.

Repair is initiated by binding of MutS protein to a mismatch. MutL subsequently binds to MutS and activates MutH, which then nicks the unmethylated strand of DNA at hemimethylated GATC sites immediately 5' to the guanine residue

(Figure 1). The incised strand is then displaced by DNA helicase II and excised from the nick to the mispaired base by one of the single-stranded DNA exonucleases (Exonuclease I [3'exonuclease activity], Exonuclease VII [both 3'and 5'exonuclease activities] or RecJ exonuclease [5'exonuclease activity]) depending on whether the nicked is 5' or 3' to the mispair. Resynthesis of the DNA strand is mediated by DNA polymerase III, single-strand DNA-binding proteins and DNA ligase (reviewed in Kolodner, 1996, Eshleman et al., 1996).

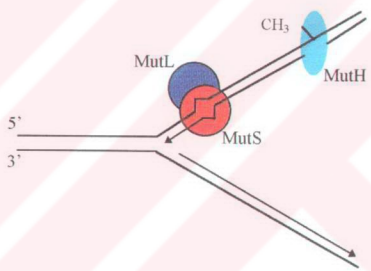



**Figure 1. *E. coli* Mut HLS DNA mismatch repair system at replication fork.**

Repair is initiated when MutS recognizes and binds the mismatch. Subsequently, MutL binds MutS and activates MutH, which then nicks the unmethylated strand of DNA at hemimethylated GATC sites. The unmethylated strand is then excised from the nick to the mispair. A new DNA strand is synthesized in the resulting gap. (Adapted from Kolodner, 1996.)

### 1.2.2. DNA Mismatch Repair in Eukaryotes

The overall mechanism of DNA mismatch repair has been highly conserved in evolution and eukaryotes have a mismatch repair system like the *E. coli* MutHLS system. In general, repair in eukaryotes involves heterodimeric protein complexes, rather than the single proteins or homodimers employed by bacteria (Lindahl et al., 1997). However, the precise mechanisms of mismatch recognition, identification of the incorrect DNA strand, excision and replacement of the mismatched DNA segment is not yet well understood. As in prokaryotes, eukaryotic mismatch repair systems appear to play important roles in the maintenance of genetic fidelity during DNA replication, genetic recombination, and genome stability. In addition, they appear to be important for preventing the appearance of certain types of cancers.

#### 1.2.2.1. DNA Mismatch Repair in *Saccharomyces cerevisiae*

In *Saccharomyces cerevisiae*, there are at least six proteins, Msh1-Msh6, which show a high degree of amino acid similarity with the bacterial MutS proteins (Kolodner, 1996). Three of these proteins, Msh2, Msh3 and Msh6, function in a eukaryotic MutHLS-like mismatch repair pathway. There are two different pathways of Msh2-dependent mismatch repair (Figure 2): one that is primarily specific for single-base substitution mispairs and requires a Msh2•Msh6 complex, and a second that is primarily specific for insertion / deletion mispairs and requires either a Msh2•Msh3 complex or a Msh2•Msh6 complex. The homologs of bacterial MutL protein in *S. cerevisiae* are Mlh1 and Pms1.

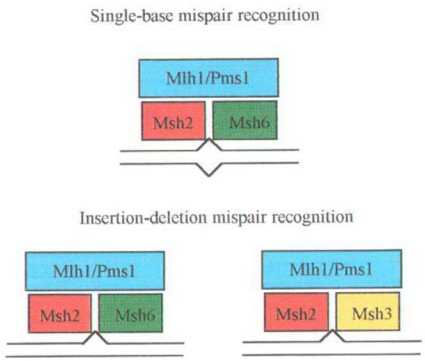



**Figure 2. Model for mismatch recognition in *S. cerevisiae*.**
Complexes between Msh2 and either Msh3 or Msh6 interact with a single base substitution mispair or an insertion / deletion mispair. Exactly which protein - Msh2, Msh3 or Msh6 - actually interacts with the mispaired base is not known. The Mlh1•Pms1 complex interacts with the mispair recognition complex. (Adapted from Kolodner, 1996.)

### 1.2.2.2. DNA Mismatch Repair in Human Cells

The initial steps of correction can be classed as primary and secondary recognition events (Lindahl et al., 1997). Primary recognition and binding of mismatched DNA (Figure 3) is carried out by homologs of the *E. coli* MutS protein, hMSH2, hMSH6 (GTBP) and hMSH3, which can associate to form two different heterodimers. The hMutSα heterodimer contains hMSH2•hMSH6 and recognizes single base mispairs, single base loops and two base loops in repeated dinucleotide sequences. The hMutSβ heterodimer contains hMSH2•hMSH3 and preferentially binds two, three and four base loops.

The secondary recognition event involves the hMutLα heterodimer, which is composed of two homologs of the *E. coli* MutL protein, hMLH1•hPMS2. The

hMutLα heterodimer binds both the hMutSα-DNA and hMutSβ-DNA complexes (Figure 3). Cells with inactivating mutations in either of these components are completely mismatch repair-deficient. Addition of the hMutLα heterodimer to these defective cell extracts is sufficient to restore mismatch repair activity (Lindahl et al., 1997). Following mismatch recognition, the DNA strand with the incorrect base is excised and a new DNA strand is synthesized by DNA polymerase using the nonmutated strand as a template. Finally, DNA ligase seals the gaps and completes repair.

Eukaryotic cells lack d(GATC) methylation and, accordingly, no eukaryotic homologue to the MutH endonuclease MutH has been identified. Some other method of distinguishing the strands presumably exists. One hypothesis is that strand gaps between the Okazaki fragments on the daughter strand may direct the strand specificity of mismatch repair (Boland, 1998). Recently the DNA replication protein proliferating cell nuclear antigen (PCNA) was found to associate with hMLH1 and hMSH2 and participate in an early step of mismatch repair (possibly in strand discrimination) (Umar et al., 1996). PCNA also interacts with the exonuclease complex DNase IV/FEN1 and may stimulate nicking to facilitate removal of the mismatched DNA segment (Chen et al., 1996). Thus, PCNA may have multiple roles in mismatch repair.

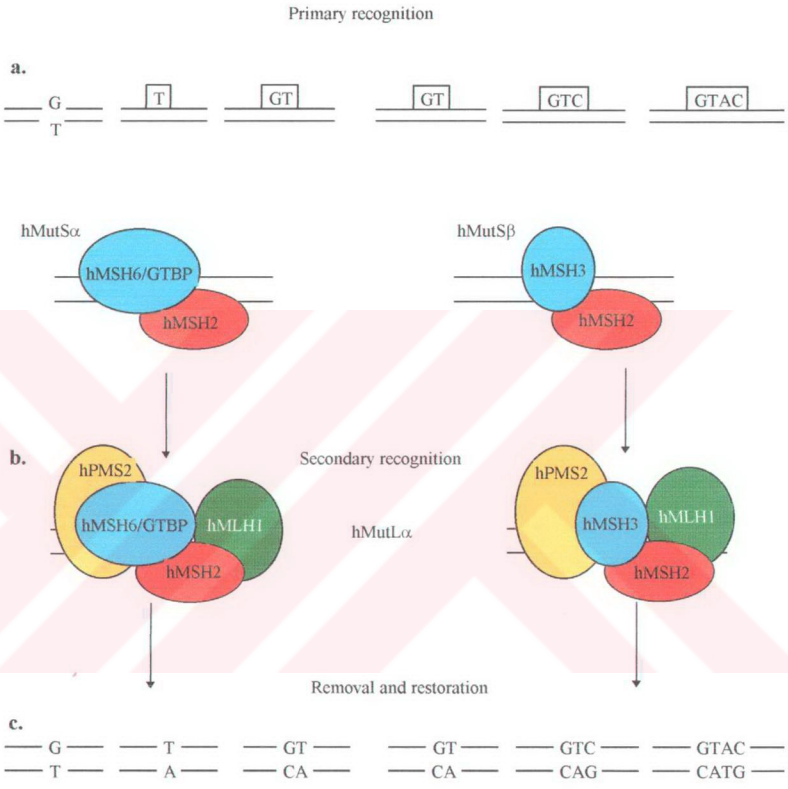



**Figure 3. Early steps in human DNA mismatch repair.**
a. Primary recognition: the heterodimer hMutSα (hMSH2•hMSH6) preferentially recognizes single base mispairs, single base loops, and two base loops in repeated dinucleotide sequences, while the heterodimer hMutSβ (hMSH2•hMSH3) preferentially binds two, three and four base loops. b. Secondary recognition: the MutLα heterodimer is recruited by the hMutSα-DNA and hMutSβ-DNA complexes. The stretch of DNA containing the mismatch is excised, resynthesized and ligated to complete repair. (Adapted from Lindahl et al., 1997.)

### 1.3. Microsatellite Instability in HNPCC

Microsatellite loci are short units (one to five base pairs) of DNA that are tandemly repeated. The repeat nature of microsatellites makes these DNA sequences particularly prone to mutation due to unequal crossing-over during genetic recombination and/or slippage of DNA polymerase during DNA replication (reviewed in Friedberg et al., 1995). During genetic recombination events, if the DNA repeat sequences misalign, the resulting recombinants will have different lengths. During DNA replication, the primer and template DNA strands can transiently disassociate and the realign in a different configuration. If the unpaired bases are in the primer strand, continued synthesis results in an insertion event and elongation of the microsatellite. If the unpaired bases are in the template strand, continued synthesis results in a deletion event and shortening of the microsatellite. When the DNA mismatch repair system is intact, these errors are usually repaired so genome stability is maintained.

Researchers working with *E. coli* showed that defects in *mutS* and *mutL* increased the frequency of frameshift mutations in poly $(GT)_n$ tandem repeats about 13-fold (Levinson and Gutman, 1987). In a similar manner, researchers studing eukaryotic mismatch repair in *S. cerevisiae* demonstrated that mutations in *Pms1*, *Mlh1* and *Msh2* caused a 100-700-fold increase in $(GT)_n$ tract instability (Strand et al., 1993). On the basis of these observations, researchers concluded that an intact DNA mismatch repair system was important for the stability of these short repeat sequences.

In 1993 three research groups looking for loss of heterozygosity in colorectal tumors found an unusual somatic mutation -- the insertion/deletion of simple repeat elements or microsatellites -- in 12-15% of colorectal tumors (Ionov et al., 1993;

Thibodeau et al., 1993; Peltomaki et al., 1993b). Hundreds of thousands of microsatellites are present throughout the human genome, usually in intronic regions of DNA. Multiple alleles with varying numbers of repeat units exist and the loci are highly polymorphic within the population. For each microsatellite locus, individuals have two alleles, often with different numbers of repeat units. Microsatellite lengths are the same within an individual's normal tissues. The observance of widespread variation in microsatellite lengths in HNPCC tumors was termed replication error (RER) phenotype or microsatellite instability (MI).

### 1.4. Identification of Human DNA Mismatch Repair Genes

Based on these observations of microsatellite instability in HNPCC tumors and previous observations in bacteria and yeast, it was predicted that the HNPCC defect would be a mutation in one of the mismatch repair genes (Aaltonen et al., 1993). Linkage analysis of several large HNPCC families showed that the disease was linked to chromosomes 2 (Peltomaki et al., 1993a) and 3 (Lindblom et al., 1993).

#### 1.4.1. Human Homologs of MutS

The first HNPCC gene *hMSH2* was identified using both the candidate gene approach (Fishel et al., 1993) and positional cloning (Leach et al., 1993). Fishel et al. (1993) used degenerate PCR primers designed on amino acid sequences conserved between bacteria and yeast to amplify MutS homologs from colon cancer cell lines. PCR products of expected size were identified, cloned, sequenced and shown to encode a predicted amino acid sequence with homology to MutS. Physical mapping localized the *hMSH2* gene to chromosome 2. Leach et al. (1993) used polymorphic markers to define a 0.8-Mb interval containing the HNPCC locus. They then showed

that a MutS homolog mapped within this interval. Both groups demonstrated the existence of germline mutations that altered the predicted protein product and cosegregated with the disease in HNPCC families.

### 1.4.2. Human Homolgs of MutL

After the human homologs of mutS gene of bacteria and yeast were found to have mutations for HNPCC, investigators searched for other human mismatch repair genes. Two groups simultaneously reported the cloning of the human MutL homolog, *hMLH1* (Papadopoulos et al., 1994; Bronner et al., 1994). Papadopoulos et al. (1994) surveyed a larged database of expressed sequence tags (ESTs) and found 3 human mismatch repair genes related to the bacterial MutL gene. One gene, *hMLH1*, was mapped to chromosome 3p21.3 by fluorescence in situ hybridization. The other two genes were more similar to the yeast mutL homolog, PMS1, and were therefore denoted *hPMS1* and *hPMS2*. Bronner et al. (1994) used degenerate PCR primers designed on amino acid sequences conserved between bacteria and yeast to isolate a human MutL homolog that also mapped to chromosome 3p. Deleterious *hMLH1* germline mutations were demonstrated to cosegregate with disease in HNPCC families (Papadopoulos et al., 1994; Bronner et al., 1994).

### 1.4.3. Homology Alignment of DNA Mismatch Repair Genes

Homology alignment of the *E. coli* mutL (SwissProt: P23367), *S. cerevisiae* MLH1 (SwissProt: P38920) and human MLH1 proteins (SwissProt: P40692) was performed using the Multiple Sequence Alignment program

(http://www.ibc.wustl.edu./ibc/msa.html). As seen in Figure 4, the amino terminal sequences of the proteins have the highest homology. The human and yeast MLH1 proteins show 41% identity.

The same alignment program was used to align *E. coli* mutS (SwissProt: U29579), *S. cerevisiae* MSH2 (SwissProt: M84170) and human MSH2 proteins (SwissProt: P43246). As seen in Figure 5, the carboxy-terminal sequences have the highest homology.

In Figure 4 and Figure 5, the identical amino acids are green shaded and similar amino acids are yellow shaded.

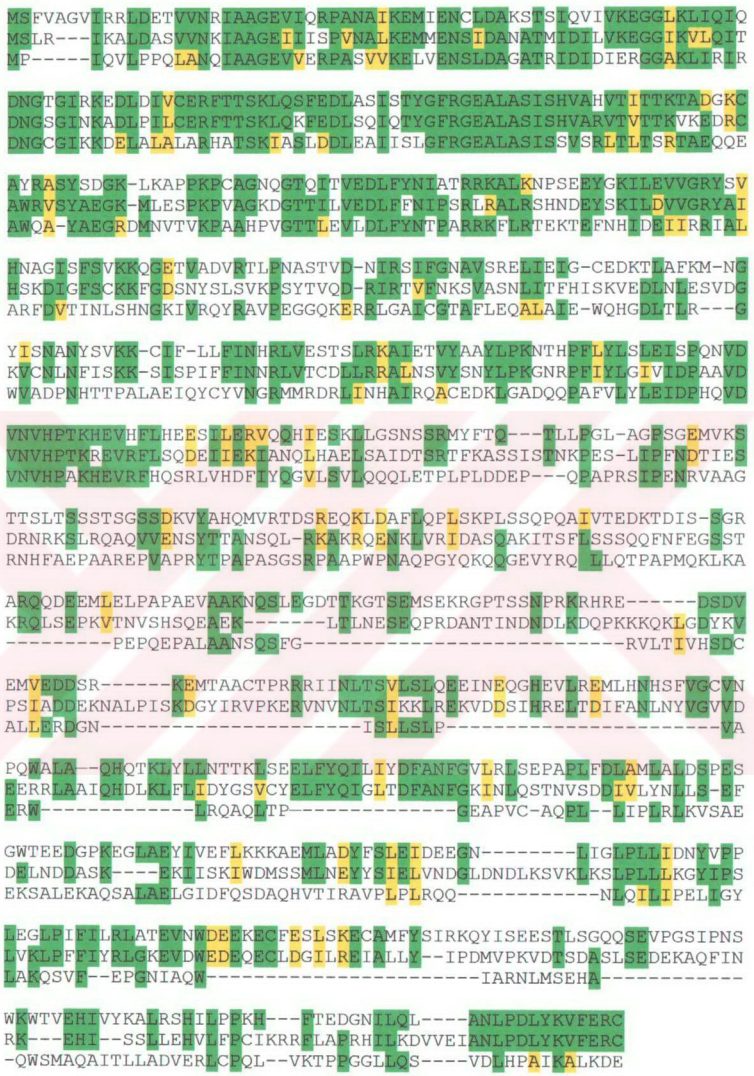


**Figure 4. Homology alignment of MutL homologs.**
Alignment was performed using the Multiple Sequence Alignment program (http://www.ibc.wustl.edu./ibc/msa.html). Row 1: human MLH1 sequence (SwissProt: P40692); Row 2: *S. cerevisiae* MLH1 sequence (SwissProt: P38920); Row 3: *E. coli* mutL sequence (SwissProt: P23367).

**Figure 5. Homology alignment of MutS homologs.**
Homology alignment was performed using the Multiple Sequence Alignment program (http://www.ibc.wustl.edu./ibc/msa.html). Row 1: human MSH2 sequence (SwissProt: P43246); Row 2: *S. cerevisiae* MSH2 sequence (SwissProt: M84170); Row 3: *E. coli* mutS sequence (SwissProt: U29579).

### 1.5. DNA Mismatch Repair Gene Defects in HNPCC

Germline mutations have been identified in five of the six mismatch repair genes believed to play a role in DNA mismatch repair in humans (Table 1): *hMSH2* (Fishel et al., 1993; Leach et al., 1993), *hMLH1* (Papadopoulos et al., 1994; Bronner et al., 1994), *hPSM2* (Nicolaides et al., 1995), *hPMS1* (Nicolaides et al., 1995) and *hMSH6* (Akiyama et al., 1997). Mutations in *hMSH2* and *hMLH1* account for the majority of HNPCC kindreds, while mutations in *hPMS2*, *hPMS1* and *hMSH6* are rare (Papadopoulos and Lindblom, 1997).

**Table 1: DNA Mismatch Repair Genes Implicated in HNPCC**

| *H. sapiens* Gene | Chromosome Location | *E. coli* Homolog | Mutation Frequency[a,b,c] |
|---|---|---|---|
| *hMSH2* | 2p16 | MutS | 38 |
| *hMSH3* | 5q | MutS | 0 |
| *hMSH6* | 2p16 | MutS | 1.5 |
| *hMLH1* | 3p21 | MutL | 59 |
| *hPMS1* | 2q31 | MutL | 0.5 |
| *hPMS2* | 7p22 | MutL | 1.0 |

[a]Human Gene Mutation Database http://www.uwcm.ac.uk/uwcm/mg/
[b]International Coll. Group on HNPCC http://www.nfdht.nl/database/mlh1/htm
[c]Papadopoulos and Lindblom, 1997
Total number of different mutations (n) = 224.

#### 1.5.1. Location of Mutations in *hMLH1* and *hMSH2*

The *hMLH1* gene covers approximately 58 kb of genomic DNA, contains 19 exons and codes for 756 amino acids (2268 bp). Mutations in *hMLH1* are dispersed throughout the coding region (Table 2). Two mutations, common in the Finnish population, are due to founder effects (Nystrom-Lahti et al., 1995). One, a 3.5 kb

genomic deletion that results in deletion of exon 16, was detected in thirty-four Finnish families. The other, a G→T transversion in the splice acceptor of exon 6 that results in truncation of the protein, was detected in ten Finnish families. There appears to be one potential hotspot for mutations in exon 16, where an AAG (Lys) deletion was observed in eleven families worldwide.

The *hMSH2* gene covers approximately 73 kb of genomic DNA, contains 16 exons and codes for 934 amino acids (2802 bp). Mutations in *hMSH2* are also dispersed throughout the coding region with the exception of two potential hotspots. The first hotspot is the splice donor site for exon 5, where a, A→T transition mutation causes and inframe deletion of exon 5; this mutation has been detected in 22 families worldwide. The second hotspot is in exon 12, where an AAT (Asn) deletion has been reported in eight families worldwide.

### 1.5.2. Types of Mutations in *hMLH1* and *hMSH2*

The types of mutations reported for *hMLH1* (Table 4) and *hMSH2* (Table 5) are also heterogeneous and includes missense mutations, nonsense mutations, splicing mutations, small insertions and small deletions. The majority of mutations reported for both *hMLH1* and *hMSH2* result in truncation of the protein product (67% and 80%, respectively). The predominant type of mutations for *hMLH1* was single base substitutions (65%). Mutations in *hMSH2* resulted both from single base substitutions (45%) and small insertions/deletions (55%).

**Table 2: *hMLH1* Mutation Frequency**

| Exon | No. of Mutations [a,b,c] | Frequency (%) |
|---|---|---|
| 1 | 12 | 5.2 |
| 2 | 11 | 4.8 |
| 3 | 4 | 1.7 |
| 4 | 8 | 3.5 |
| 5 | 3 | 1.3 |
| 6 | 11 | 4.8 |
| 7 | 3 | 1.3 |
| 8 | 11 | 4.8 |
| 9 | 11 | 4.8 |
| 10 | 7 | 3.0 |
| 11 | 11 | 4.8 |
| 12 | 10 | 4.3 |
| 13 | 12 | 5.2 |
| 14 | 13 | 5.7 |
| 15 | 5 | 2.2 |
| 16 | 63 | 27.4 |
| 17 | 13 | 5.7 |
| 18 | 5 | 2.2 |
| 19 | 17 | 7.4 |
| Total | 230 | 100 |

[a]The Human Gene Mutation Database http://www.uwcm.ac.uk/uwcm/mg/
[b]International Coll. Group on HNPCC http://www.nfdht.nl/database/mlh1/htm
[c]Papadopoulos and Lindblom, 1997

**Table 3: *hMSH2* Mutation Frequency**

| Exon | No. of Mutations [a,b,c] | Frequency (%) |
|---|---|---|
| 1 | 3 | 2.4 |
| 2 | 3 | 2.4 |
| 3 | 8 | 6.4 |
| 4 | 2 | 1.6 |
| 5 | 26 | 20.8 |
| 6 | 7 | 5.6 |
| 7 | 12 | 9.6 |
| 8 | 8 | 6.4 |
| 9 | 2 | 1.6 |
| 10 | 7 | 5.6 |
| 11 | 5 | 4.0 |
| 12 | 25 | 20.0 |
| 13 | 6 | 4.8 |
| 14 | 3 | 2.4 |
| 15 | 6 | 4.8 |
| 16 | 2 | 1.6 |
| Total | 125 | 100 |

[a]The Human Gene Mutation Database http://www.uwcm.ac.uk/uwcm/mg/
[b]International Coll. Group on HNPCC http://www.nfdht.nl/database/msh2/htm
[c]Papadopoulos and Lindblom, 1997

**Table 4: Mutation types in *hMLH1***

| Mutation Type | Number of Mutations |
|---|---|
| Nucleotide substitution (missense) | 44 |
| Nucleotide substitution (nonsense) | 15 |
| Nucleotide substitution (splicing) | 27 |
| Small deletion | 25 |
| Small insertion | 18 |
| Small insertion / deletion | 2 |
| Complex rearrangement | 2 |
| Total | 133 |

**Table 5: Mutation types in *hMSH2***

| Mutation Type | Number of Mutations |
|---|---|
| Nucleotide substitution (missense) | 17 |
| Nucleotide substitution (nonsense) | 16 |
| Nucleotide substitution (splicing) | 5 |
| Small deletion | 38 |
| Small insertion | 8 |
| Gross insertion / duplication | 1 |
| Total | 85 |

## 1.6. Mouse Models for HNPCC

Mice carrying disruptions in DNA mismatch repair genes can provide insights into HNPCC tumorigenesis. Knock-out mice for *Mlh1* and *Msh2* are viable but they show a high degree of genomic instability (de Wind et al., 1995; Reitmair et al., 1995). By one year of age, the majority of the knock-out mice have developed lymphomas, intestinal adenomas and adenocarcinomas (Baker et al., 1996; Prolla et al., 1998; Reitmair et al., 1996). Both male and female mismatch repair-deficient mice are sterile (Bedell et al., 1997). Male *Mlh1*-deficient mice are infertile because their spermatocytes fail to progress beyond pachytene stage of meiosis (Baker et al., 1996; Edelmann et al., 1996). Further studies with these animal models may help define the role of DNA mismatch repair genes in cancer susceptibility.

## 1.7. HNPCC Tumorigenesis

Development of both sporadic and inherited colorectal cancer is a multistep process that requires several genetic changes. Hereditary colorectal cancers develop at a much younger age than sporadic colorectal cancers because these indivdiuals are born with germline mutation that predisposes to cancer development (Tannergard et al., 1997). The germline defect in HNPCC individuals is a germline mutation in one of the genes involved in DNA mismatch repair. Heterozygous individuals have apparently normal DNA mismatch repair activity (Tannergard et al., 1997). However, somatic mutation of the wild-type allele results in loss of this ability and accumulation of characteristic mutations, such as single base mispairs and length alterations in homopolymeric tracts (Hemminki et al., 1994).

Kinzler and Vogelstein (1997) have classified genes involved in tumorigenesis as either gatekeepers or caretakers. Cell cycle regulatory genes are termed

gatekeepers because mutational inactivation of these genes results in the initiation of tumorigenesis. DNA mismatch repair genes are termed caretakers because they protect the integrity of the genome. One of the first manifestations of defective DNA mismatch repair is the characteristic microsatellite instability that is seen in HNPCC tumors. (Herfarth et al., 1997). Microsatellite instability can result from mutations in any of the genes required for mismatch repair (Boyer et al., 1995). This microsatellite instability is one sign of the genome-wide instability or "mutator phenotype" (Loeb, 1991; Parsons et al., 1993) that accelerates tumor progression.

hMLH1 and hMSH2 proteins are localized in the nucleus and are highly expressed in the epithelium of the digestive tract (Fink et al., 1997). However, mismatch repair genes are believed to be expressed in all proliferating cells (de Wind et al., 1998). Therefore, it is surprising that HNPCC tumorigenesis occurs predominantly in the colon. This tissue specificity can be explained in several ways (de Wind et al., 1998): (1) oncogenic mutations in mismatch repair-deficient cells may accumulate more rapidly in tissues with a high cell turnover and in tissues exposed to mutagens; (2) proto-oncogenes or tumor suppressor genes that control growth and differentiation of susceptible tissues may have mononucleotide tracts that are hot spots for mutations in mismatch repair-deficient cells; and (3) mismatch repair-deficient cells may have a growth advantage with respect to mismatch repair-proficient cells . In fact, tumors isolated from HNPCC individuals have been found to have mutations in mononucleotide tracts within several growth regulating genes, including the transforming growth factor-β type II receptor (*TGFβRII*) (Markowitz et al., 1995), the insulin-like growth factor type II receptor (*IGFIIR*) (Souza et al., 1996), and *BAX* (Rampino et al., 1997).

Individuals from families at high risk for colorectal cancer should be offered genetic counseling, predictive molecular testing and, when indicated, endoscopic surveillance at appropriate intervals. Early detection of mutations in these tumor susceptibility genes would allow more effective monitoring of individuals at risk.

### 1.8. Mutation Screening

Because the *hMLH1* and *hMSH2* are long genes (58 kb and 73 kb, respectively) with many exons (19 and 16, respectively), it is very difficult to detect mutations. Detection of unknown mutations can involve DNA sequences analysis of thousands of bases and would require large amounts of money and time. This has lead to the development of many techniques that can be used in research and clinical laboratories to screen populations for unknown mutations, as well as to detect known mutations.

An optimal mutation detection technique would (1) be fast, (2) be able to screen large stretches of DNA with high sensitivity and specificity, (3) not involve expensive or elaborate instrumentation, (4) not require toxic or dangerous compounds and (5) provide information about the location and nature of the mutation. Unfortunately, no single procedure yet described possesses all of these attributes.

#### 1.8.1. Protein Truncation Test

The protein truncation test (PTT) rapidly detects mutations that interrupt the reading frames of genes. Templates are generated by PCR using cDNA synthesized by reverse transcription of mRNA (RT-PCR). A promoter is incorporated into forward primer. After in vitro transcription and translation assays are performed, the

protein product is analyzed by gel electrophoresis. If there is a truncating mutation, the protein will have lower molecular weight compared to the normal protein.

### 1.8.2. RT-PCR

mRNA is isolated from blood and RT-PCR is performed by using exonic primers. After that, the products are analyzed on the gel. If there is a mutation, the size of PCR product may be smaller or the PCR product may not be detected if there is no mRNA transcript. This method is very useful for detecting exonic deletions that might not be detected using genomic DNA.

### 1.8.3. Heteroduplex Analysis (HA)

Complementary single-stranded DNA derived from alleles that differ in sequence will include mismatched base pairs (heteroduplexes) when allowed to anneal. Heteroduplex molecules are formed by denaturing DNA at 95°C and then allowing the single stands to anneal by cooling to room temperature slowly. Double-stranded heteroduplex molecules may show an altered electrophoretic migration in non-denaturing gels compared to homoduplexes of either allele. The main advantage of HA is its simplicity, but its sensitivity is about 80-90% (Grompe, 1993).

### 1.8.4. Single-Strand Conformation Polymorphism Analysis (SSCP)

The SSCP method relies on the fact that single-strand DNA molecules in solution under certain conditions have a defined secondary structure. In principle, when a single-stranded DNA molecule is placed in a non-denaturing solution, it will fold in a sequence specific manner. If one of the bases is changed, the molecule is likely to fold into a different shape and size. During non-denaturing gel

electrophoresis, the different shaped molecules are likely to move at different rates so that when electrophoresis is stopped, their positions will be different. This differential mobility can allow one to distinguish normal and mutant alleles. Because there are two complementary mutant strands, each mutant allele has two chances of being detected.

The most clear advantage of SSCP is the simplicity of the method. It can be done without labeling. The mutation detection rate depends on the length of the fragments. As the length of fragments increase, the sensitivity decreases. SSCP is sensitive only for 200-300 bp fragments and mutation detection rate is 80-90%. Variation of gel concentration, gel composition, and/or electrophoresis parameters may improve resolution and increase the mutation detection rate.

### 1.8.5. DNA Sequence Analysis

DNA sequencing refers to direct determination of the nucleotide sequence of a DNA fragment. Because DNA sequencing defines both the location and the nature of the change, it is the confirmatory step of any mutation screening method.

Two sequencing methods are used: the Maxam-Gilbert chemical cleavage method and the Sanger dideoxy chain termination method. In the Maxam-Gilbert method, a segment of DNA is labeled at one end and the phoshodiester bonds between specific bases are cleaved by using specific chemicals. This generates a series of labeled fragments, the lengths of which depend on the distance of the destroyed base from the labeled end of the molecule. The sets of labeled fragments are run side by side on an polyacrylamide gel that separates DNA fragments according to size and the gel is autoradiographed. The pattern of bands on the X-ray film is analyzed to determine the nucleotide sequence.

In the Sanger method, the DNA strand to be sequenced along with a labeled primer, is divided into four DNA polymerase reactions, each containing one of four dideoxynucleotriphosphates (ddNTPs). Incorporation of a dideoxynucleotide terminates DNA synthesis of that molecule. The concentration of the reactants is adjusted to allow some chain termination to occur at each base. The resultant labeled fragments are then separated by size by electrophoresis in a denaturing polyacrylamide gel. After autoradiography, the nucleotide sequence is determined by the pattern of the bands.

Automated sequencing relies on incorporation of fluorescent molecules during DNA extension reactions. Fluorescent dye labels can be incorporated using either 5'-dye labeled primers (dye primers) or 3'-labeled dideoxynucleotide triphosphates (ddNTPs) (dye terminators). With dye terminator labeling, each of the four ddNTPs is tagged with a different fluorescent dye. Thus, the growing chain is simultaneously terminated and labeled with dye that corresponds to that base.

During electrophoresis, the DNA bands are excited by a laser light as they migrate past a set point in the gel. Each dye emits light at a different wavelength. Therefore, all four colors can be detected and distinguished in a single gel lane. The data is collected and analyzed by sophisticated software. This strategy improves sequencing accuracy because it eliminates problems caused by variations in electrophoretic mobility from lane to lane. It also increases the number of templates that can be analyzed on a single gel. However, despite improvements in sequencing chemistry and in the software for heterozygote detection, direct automated sequencing of PCR products does not guarantee complete mutation detection and requires time-consuming and costly optimizations.

### 1.8.6. Restriction Enzyme Analysis

Restriction endonucleases recognize short DNA sequences and cleave double stranded DNA at specific sites within or adjacent to the recognition sequences. The recognition sequences are generally, but not always, 4 to 6 nucleotides in length and are usually characterized by palindromic sequences. In palindromic sequences the recognition site sequence is the same on each DNA strand when read 5'→3'. Some restriction enzymes cleave at the axis of symmetry yielding "blunt" ends. Others make staggered cleavages yielding "sticky" ends. Restriction enzyme cleavage is accomplished by incubating the enzyme with the DNA in appropriate reaction conditions. The amounts of enzyme and DNA, the buffer and ionic concentrations, and the temperature and duration of reaction vary depending upon the specific application.

## 1.9. Aim and Strategy

This project aims to detect germline mutations in the DNA mismatch repair gene, *hMLH1*, in Turkish patients with hereditary non-polyposis colorectal cancer. *hMLH1* is the gene most frequently mutated in HNPCC. SSCP will be used for detection of unknown gene mutations. Restriction enzyme analysis will be used for detection of known gene mutations that occur in the enzyme recognition site. Automated sequencing will be used for confirmation of results. Identification of the germline mutations in Turkish individuals may provide insights into the kinds of disease mutations in the Turkish population. This might allow a reliable and cost-effective screening procedure to be used to identify individuals at high risk of developing colorectal cancer. These individuals could then be closely monitored to detect tumors at an early stage when treatment is most effective.

# 2. Materials and Methods

## 2.1. Materials

### 2.1.1. Patient samples

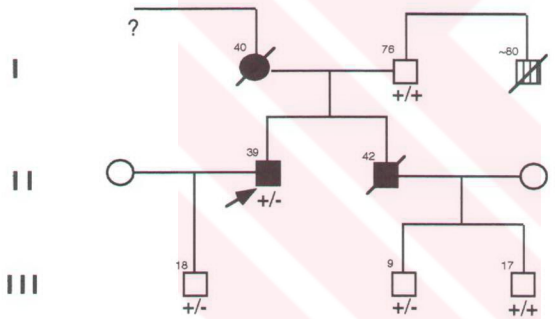
HNPCC patients were referred to Bilkent University, Faculty of Science, Molecular Biology Department (Ankara, Turkey) by collaborating physicians at Marmara University (İstanbul, Turkey), Hacettepe University (Ankara, Turkey), Akdeniz University (Antalya, Turkey) and Çukurova University (Adana, Turkey). Blood samples were collected in tubes containing EDTA. HNPCC family pedigrees are shown in Figures 6-10.

### 2.1.2. Cell Lines

Four lymphoblastoid cell lines with defined mutations in hMLH1 were provided by M. Öztürk from IARC, Lyon, France.

### 2.1.3. Oligonucleotides

Primers used in polymerase chain reactions and cycle sequencing reactions were synthesized on the Beckman Oligo 1000M DNA synthesizer (Beckman Instruments Inc., Fullerton, CA, USA) at Bilkent University, Faculty of Science, Department of Molecular Biology and Genetics (Ankara, Turkey). The nucleotide sequences of the primers used for the analysis of the *hMLH1* gene are given in Table 6.




■ : HNPCC
▥ : Prostate cancer

**Figure 6:** Pedigree of HNPCC 1 Family

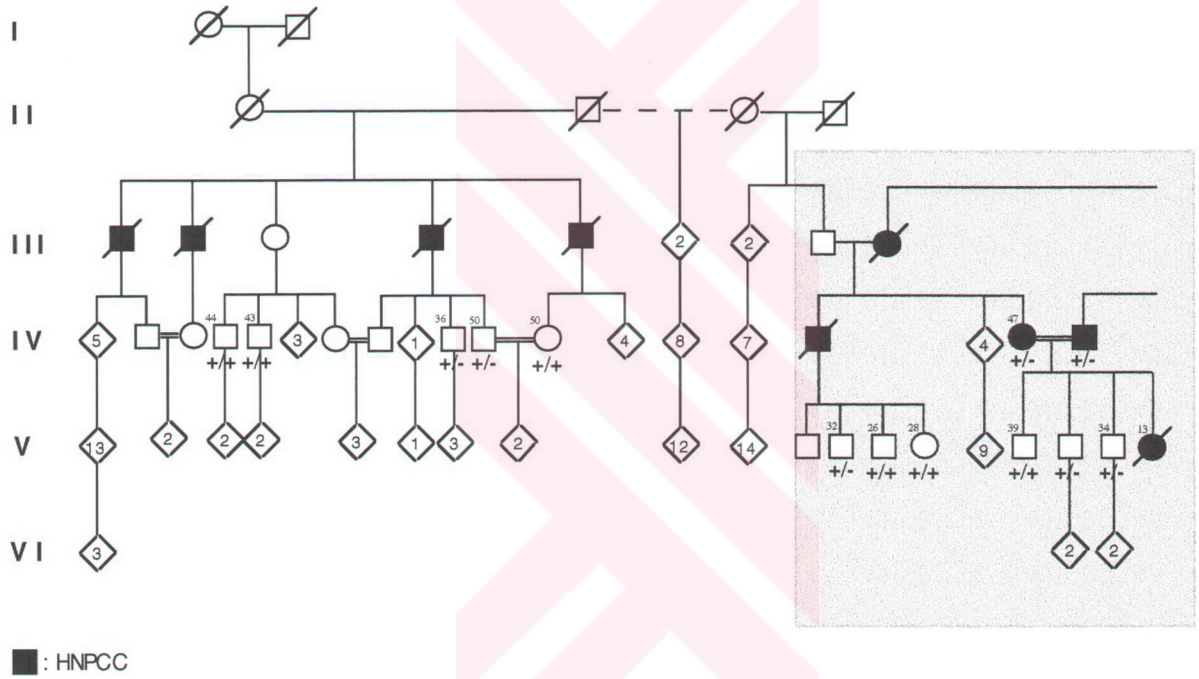



**Figure 7a:** Pedigree of HNPCC 2 Family
The shaded region is common in Figure 7a and 7b.

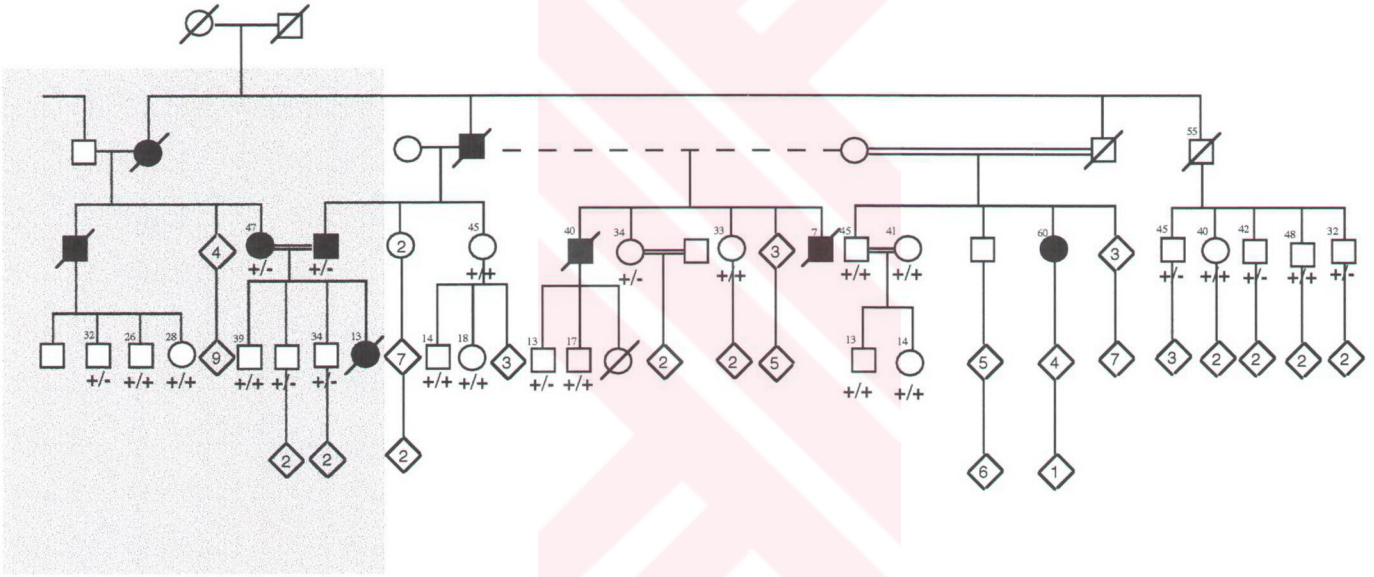
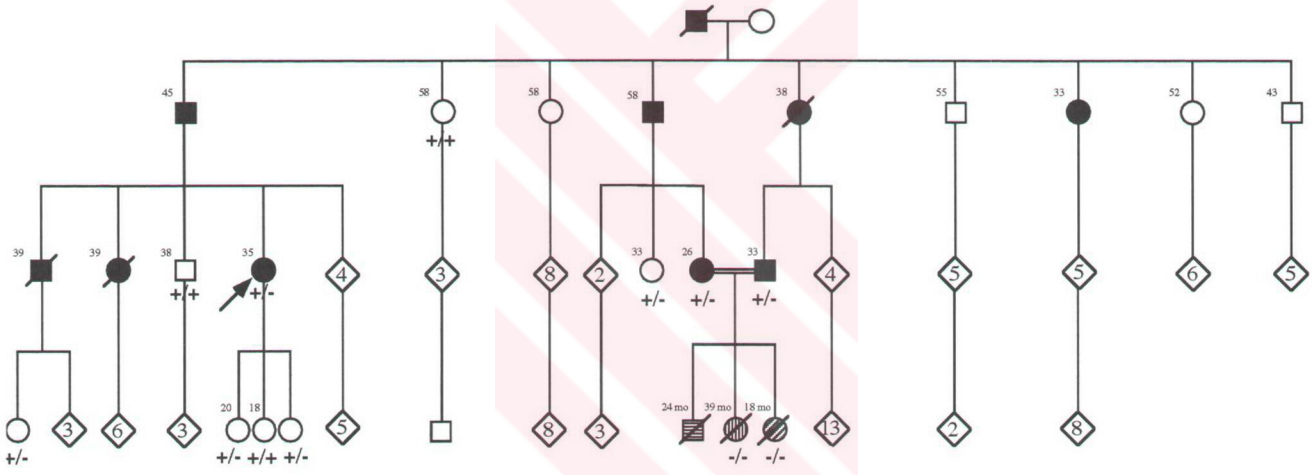



**Figure 7b:** Pedigree of HNPCC 2 Family
The shaded region is common in Figure 7a and 7b.




HNPCC
Acute Myelomonocytic Leukemia + Unconfirmed Neurofibromatosis
Non-Hodgkin Lymphoma + Unconfirmed Neurofibromatosis
Chronic Myeloid Leukemia or Chronic Myelomonocytic Leukemia + Neurofibromatosis

**Figure 8:** Pedigree of HNPCC 3 Family

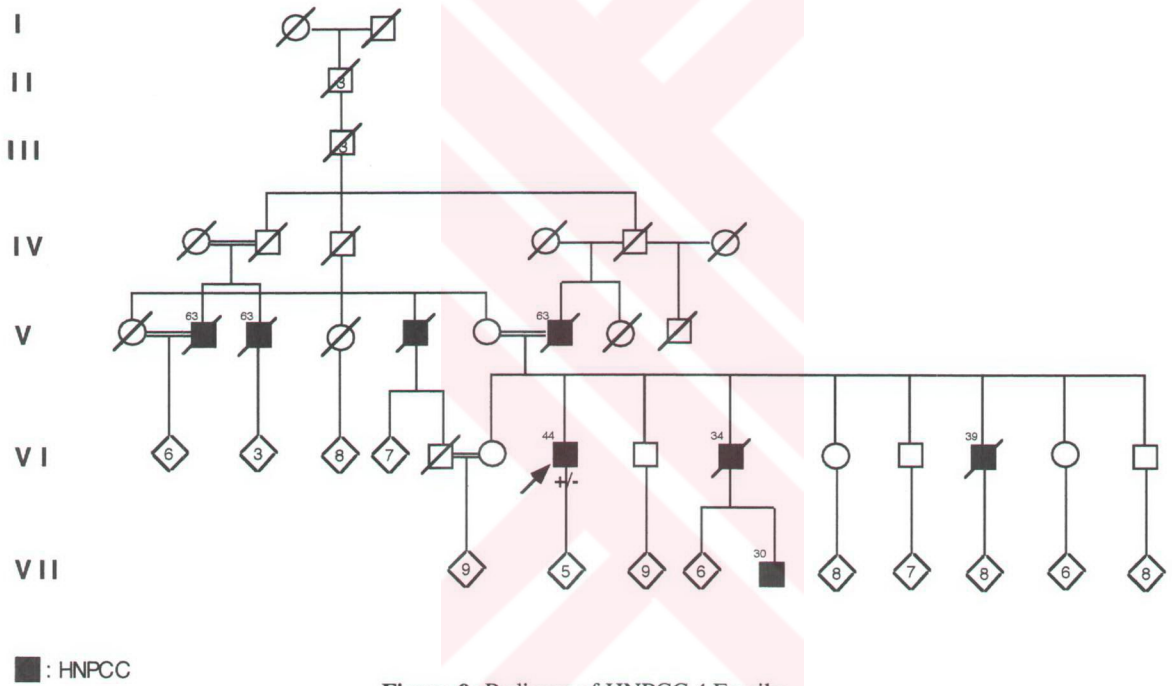



**Figure 9:** Pedigree of HNPCC 4 Family

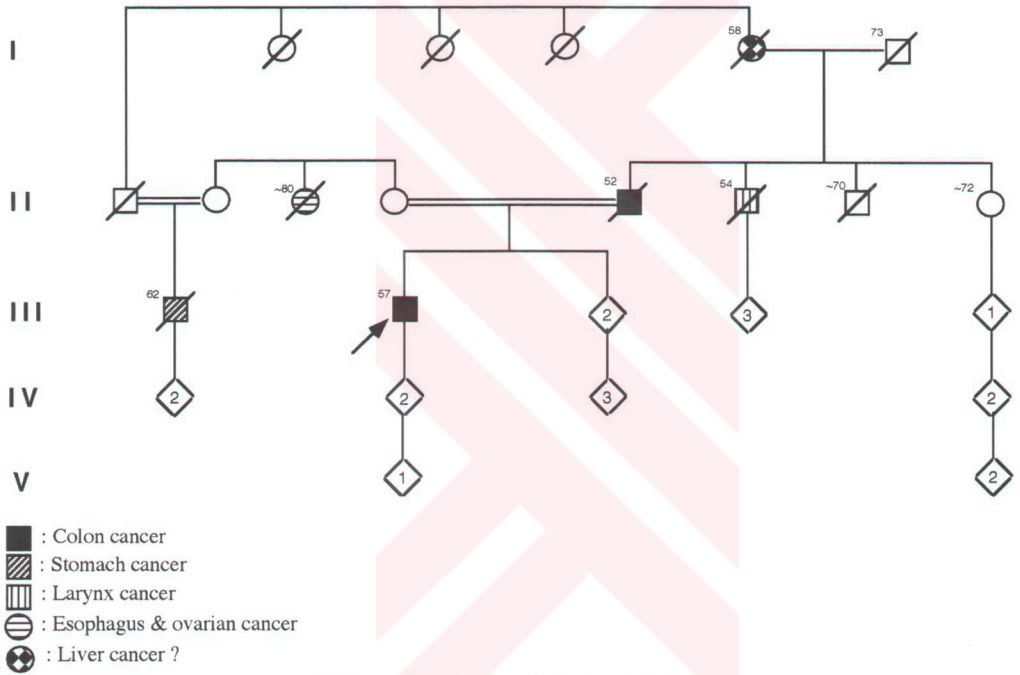



**Figure 10:** Pedigree of HNPCC 5 Family

**Table 6: Sequences of *hMLH1* primers**

| Exons | Name | Sequence (5'→3') |
|---|---|---|
| MLH01 F | GA 281 | CACTGAGGTGATTGGCTGAA |
| MLH01 R | GA 282 | CCGTTAAGTCGTAGCCCTTA |
| MLH02 F | GA 283 | GTACATTAGAGTAGTTGCAGAC |
| MLH02 R | GA 284 | CAGAGAAAGGTCCTGACTC |
| MLH03 F | GA 152 | GAGATTTGGAAAATGAGTAAC |
| MLH03 R | GA 153 | ACTAACAAATGACAGACAATG |
| MLH04 F | GA 285 | CTTTCCCTTTGGTGAGGTGA |
| MLH04 R | GA 286 | TACTCTGAGACCTAGGCCCA |
| MLH05 F | GA 156 | CCCTTGGGATTAGTATCTAT |
| MLH05 R | GA 157 | TACTCTCCCATGTACCATTC |
| MLH06 F | GA 287 | GGGTTTTATTTTCAAGTACTTCTATG |
| MLH06 R | GA 288 | CAGCAACTGTTCAATGTATGAGCAC |
| MLH07 F | GA 289 | GTGTGTGTTTTTGGCAACTC |
| MLH07 R | GA 290 | CCTTATCTCCACCAGCAAAC |
| MLH08 F | GA 291 | CTCAGCCATGAGACAATAAATCC |
| MLH08 R | GA 292 | GGTTCCCAAATAATGTGATGG |
| MLH09 F | GA 164 | TGGATGGATGAATGGACAGG |
| MLH09 R | GA 165 | GGATTTCCAATGTGGTTCTT |
| MLH10 F | GA 333 | TGAATGTACACCTGTGACCTCACC |
| MLH10 R | GA 334 | GAGGAGAGCCTGATAGAACATCTG |
| MLH11 F | GA 337 | CTTTTTCTCCCCCTCCCACTA |
| MLH11 R | GA 338 | TCTGGGCTCTCACGTCT |
| MLH12 F | GA 339 | AATTATACCTCATACTAGC |
| MLH12 R | GA 340 | GTTTTATTACAGAATAAAGGAGG |
| MLH13 F | GA 343 | TGCAACCCACAAAATTTGGC |
| MLH13 R | GA 344 | CTTTCTCCATTTCCAAAACC |
| MLH14 F | GA 293 | TGGTGTCTCTAGTTCTGG |
| MLH14 R | GA 294 | CATTGTTGTAGTAGCTCTGC |
| MLH15 F | GA 172 | TGTCTCATCCATGTGTCAGG |
| MLH15 R | GA 173 | GCGGTCAGTTGAAATGTCAG |
| MLH16 F | GA 345 | CATTTGGATGCTCCGTTAAAGC |
| MLH16 R | GA 346 | CACCCGGCTGGAAATTTTATTTG |
| MLH17 F | GA 176 | GAAAGGCACTGGAGAAATGG |
| MLH17 R | GA 177 | CCGAAATGCTTAGTATCTGC |
| MLH18 F | GA 178 | AAGTAGTCTGTGATCTCCGT |
| MLH18 R | GA 179 | AAGATGTATGAGGTCCTGTC |
| MLH19 F | GA 180 | CAGGACACCAGTGTATGTTG |
| MLH19 R | GA 181 | AAGAACACATCCCACAGTGC |

### 2.1.4. Chemicals and Reagents

The chemicals and reagents used in this study were purchased from the following suppliers:

| Reagent | Supplier |
|---|---|
| Acrylamide | Sigma, St. Louis, MO, USA |
| Acetic acid | Carlo Erba, Milano, Italy |
| Agarose | Sigma, St. Louis, MO, USA |
| Ammonium persulfate | Sigma, St. Louis, MO, USA |
| Bisacrylamide | Sigma, St. Louis, MO, USA |
| Blue dextran | Sigma, St. Louis, MO, USA |
| Boric acid | Sigma, St. Louis, MO, USA |
| Bromophenol blue | Sigma, St. Louis, MO, USA |
| Cellulose acetate filters | Costar, Cambridge, MA, England |
| Chloroform | Carlo Erba, Milano, Italy |
| Disodium ethylenediamine tetraacetate | Carlo Erba, Milano, Italy |
| Ethanol | Merck, Frankfurt, Germany |
| Ethidium bromide | Sigma, St. Louis, MO, USA |
| Ficoll Type 400 | Sigma, St. Louis, MO, USA |
| Formamide | Sigma, St. Louis, MO, USA |
| Glycerol | Carlo Erba, Milano, Italy |
| Isoamyl alcohol | Carlo Erba, Milano, Italy |
| MDE gel solution | FMC BioProducts, Rockland, ME, USA |
| NuSieve 3:1 Agarose | FMC BioProducts, Rockland, ME, USA |
| Silver nitrate | Sigma, St. Louis, MO, USA |
| Phenol | Carlo Erba, Milano, Italy |
| Proteinase K | Appligene-Oncor, USA |
| QIAquick PCR purification kit | Qiagen, Chatsworth, CA, USA |
| Sodium acetate | Carlo Erba, Milano, Italy |
| Sodium chloride | Sigma, St. Louis, MO, USA |
| Sodium dodecyl sulfate (SDS) | Sigma, St. Louis, MO, USA |
| Sodium hydroxide | Sigma, St. Louis, MO, USA |
| TEMED | Carlo Erba, Milano, Italy |
| Tris HCl | Sigma, St. Louis, MO, USA |
| Trisodium citrate | Sigma, St. Louis, MO, USA |
| Urea | BioRad, Hercules, CA, USA |
| Xylene cyanol | Sigma, St. Louis, MO, USA |
| ϕx174 *Hae*III DNA Marker (0.5 mg/ml) | MBI Fermentas Inc., Amherst, NY, USA |
| ϕx174 *Hinf*I DNA Marker (0.5 mg/ml) | MBI Fermentas Inc., Amherst, NY, USA |
| $^{32}$P-dCTP (10 MBq/0.025 ml) | Orbital, Hungary |
| $^{33}$P-dATP (10 mCi/ml) | Amersham, Buckinghamshire, England |
| | Orbital, Hungary |

### 2.1.5. Restriction enzymes

The restriction enzymes used in this study, together with their recognition and cleavage sites, are listed in Table 7. The enzymes were purchased from the designated suppliers and used according to the manufacturers' instructions. The composition of the recommended buffers supplied with the restriction enzymes is also listed in Table 7.

**Table 7: Restriction Enzymes Used for Mutation Detection**

| Restriction Enzyme | Recognition Site | Buffer (1X) | Supplier |
|---|---|---|---|
| *Dde*I (6U/µl) | 5'-C↓TNAG-3'<br>3'-GANT↑C-5' | NEBuffer 3<br>100 mM NaCl<br>50 mM Tris-HCl, pH 7.9<br>10 mM $MgCl_2$<br>1 mM Dithiothreitol | Biolabs<br>Beverly, MA, USA |
| *Hph*I (5U/µl) | 5'-GGTGA(N)$_8$↓ -3'<br>3'-CCACT(N)$_7$↑ -5' | Buffer B<br>10 mM Tris-HCl, pH 7.5<br>10 mM $MgCl_2$<br>0.1 mg/ml BSA | MBI Fermentas Inc.<br>Amherst, NY, USA |

### 2.1.6. PCR materials

PCR kits were obtained from MBI Fermentas Inc. (Amherst, NY, USA). Kits contained *Thermus aquaticus* DNA polymerase (5U/µl), 10X PCR buffer (100 mM Tris-HCl (pH 8.8 at 25°C), 500 mM KCl, 0.8% Nonidet P40, 25 mM $MgCl_2$ solution and 10 mM dNTP mix (one ml of 10 mM dNTP solution contains 10 µmol each of dATP, dCTP, dGTP, dTTP). PCR reactions were performed in 0.2 ml Thermowell™

tubes (Corning Costar Corp., Cambridge, MA, England) using the Gene Amp PCR system 9600 (Perkin Elmer, Foster City, CA, USA).

### 2.1.7. Sequencing Materials

Cycle sequencing was performed using the ABI PRISM™ Ready Reaction Dye Terminator Cycle Sequencing Kit (ABI, Perkin Elmer, Foster City, CA, USA). The kit contained terminator premix with A-dye terminator, C-dye terminator, G-dye terminator, and T-dye terminator; dITP, dATP, dCTP and dTTP; Tris-HCl (pH 9.0); $MgCl_2$; thermal stable pyrophosphatase; and AmliTaq DNA polymerase, FS (8U/µl). Each kit also contained a pGEM®3Zf(+) control template (0.2 µg/µl) and -21 M13 forward primer (0.8 pmol/µl). Cycle sequencing reactions were performed in the Gene Amp PCR system 9600. Electorphoresis was performed using the 377 Sequencer (ABI, Perkin Elmer, Foster City, CA, USA).

### 2.1.8. Standard Solutions and Buffers

Acrylamide: bisacrylamide stock solution (40%) (75:1)

39.5 g acrylamide
0.53 g bisacrylamide
The volume was adjusted to 100 ml by adding $ddH_2O$.

Acrylamide: bisacrylamide stock solution (50%) (99:1)

49.5 g acrylamide
0.5 g bisacrylamide
The volume was adjusted to 100 ml by adding $ddH_2O$.

Acrylamide: bisacrylamide stock solution (40%) (19:1)

38 g acrylamide
2 g bisacrylamide
The volume was adjusted to 100 ml by adding $ddH_2O$

Agarose gel loading buffer (6X)

15% ficoll
0.05% bromophenol blue
0.05% xylene cyanol

Developer solution

1.5% NaOH
0.1% formaldehyde

Extraction buffer

10 mM Tris HCl, pH 8.0
10 mM EDTA, pH 8.0
0.5% SDS

Fixative solution

10% ethanol
0.5% acetic acid

Silver nitrate solution

0.1% silver nitrate

Sequencing loading buffer

5 parts deionized formamide
1 part EDTA / blue dextran
25 mM EDTA (pH 8.0)
50 mg/ml blue dextran

SSC (20X)

3 M NaCl
0.3 M trisodium citrate, pH 7.0

SSCP gel loading buffer (6X)

95% formamide
10 mM NaOH
0.25% bromophenol blue
0.25% xylene cyanol

TE Buffer

10 mM Tris HCl pH 8.0
1 mM EDTA

Tris-boric acid-EDTA (TBE) (10X) (1 L)

108 g Tris HCl
55 g boric acid
20 ml 0.5 M EDTA
q.s. 1000 ml $ddH_2O$

## 2.2. Methods

### 2.2.1. DNA isolation from whole blood specimens

Blood can be stored at 4°C for a maximum of five days. Blood was frozen in 700 µl aliquots in 1.5 ml eppendorf tubes at -80°C for at least one day before starting DNA isolation.

Blood was thawed and 800 µl of 1X SSC was added and mixed by vortexing. Then, it was centrifuged in a microfuge (Heraeus instruments, Biofuge, Osterode, Germany) at 13,000 rpm for 1 minute. The supernatant was removed without disturbing the cell pellet and discarded into disinfectant. Then 1.4 ml 1X SSC was added and the tube was vortexed briefly to resuspend the cell pellet. Again, it was centrifuged in a microfuge at 13,000 rpm for 1 minute and the supernatant was removed, avoiding the pellet. Cell pellet could be washed several times with 1X SSC if necessary.

Next 800 µl extraction buffer (10 mM Tris HCl pH 8.0, 10 mM EDTA pH 8.0, 0.5% SDS) and 10 µl proteinase K (20 mg/ml $ddH_2O$) were added. The tube was vortexed briefly to resuspend the cell pellet. The suspension was incubated at 56°C for at least 1 hour. Incubation could be done overnight if necessary to dissolve the cell pellet.

The DNA was then extracted with 400 µl phenol / chloroform / isoamyl alcohol (25:24:1) and vortexed for 60 seconds. This step must be carried out the

fume hood. The tube was spun in a microfuge for 5 minutes at 13,000 rpm. The upper aqueous layer (~700 µl) was removed and placed in a new tube. If the DNA supernatant was sticky or if the interface was not clear after this step, the supernatant is not removed. An additional extraction step was performed with 350 µl phenol / chloroform / isoamyl alcohol. The recovered supernatant was separated into two or more tubes (350 µl per tube).

The DNA was then precipitated from the suspension by adding 35 µl NaOAc (3M, pH 5.2) and 700 µl ice-cold absolute ethanol (EtOH) were added to each tube, mixed by inversion and placed at -20°C for 30 minutes. The tubes were spun in a microfuge for 15 minutes at 13,000 rpm. The alcohol was removed and the pellet was washed with 1.0 ml room temperature 70% EtOH. The tubes were spun in a microfuge for 5 minutes at 13,000 rpm. All the alcohol was removed with a micropipette and the tubes were left open on the bench (~30 min) to allow the EtOH to evaporate. The DNA was solubilized in 200 µl TE (pH 8.0) by incubating at 56°C for at least 1 hour. Incubation was done overnight if necessary to solubilize the pellet. The DNA was then stored at -20°C.

The concentration and purity of the double stranded DNA was determined on the Beckman Spectrophotometer Du 640 (Beckman Instruments Inc., Fullerton, CA, USA) using the Beckman Instruments Du Series 600 Spectrophotometer software program. Absorbance readings were taken at wavelengths of 260 nm and 280 nm. The $A_{260}$ allows calculation of the concentration of nucleic acid in the sample. An optical density value of one corresponds to approximately 50 µg/ml of double stranded DNA. The $A_{260}/A_{280}$ ratio provides an estimate of the purity of the nucleic acid. A pure preparation of DNA will have $A_{260}/A_{280}$ ratio between 1.8 and 2.0. If

there is contamination with protein or phenol, the $A_{260}/A_{280}$ ratio will be significantly less than the values given above and accurate quantitation of the amount of nucleic acid will not be possible.

DNA was also checked by horizontal agarose gel electrophoresis to verify that it was high molecular weight. A 1.0% agarose minigel with 1X TBE was prepared. Ethidium bromide (1µl/ml) was incorporated into the gel. DNA samples were loaded into the sample wells and the gel was run at 80V. After the run, the DNA was visualized under UV transilluminator.

### 2.2.2. Polymerase Chain Reaction (PCR)

PCR is a rapid procedure for in vitro enzymatic amplification of a specific segment of DNA. There are three distinct events in PCR: template denaturation, primer annealing and DNA synthesis. Template DNA was denatured by heating the reaction to 95-96°C. After denaturation the primers were allowed to hybridize to their complementary single-stranded target sequences. The temperature of this step depends on the homology of the primers for the target sequence as well as the base composition of the oligonucleotides. The last step was the extension of the oligonucleotide primer by the thermostable polymerase. Traditionally, this step was carried out at 72°C. Usually, the larger the template, the longer the time required for a proper extension.

In general, 50-100 ng DNA was used as a template. A 50 µl PCR reaction contained 5 µl 10X PCR buffer (final concentration 1X PCR buffer), 1.5-3.0 mM $MgCl_2$, 200 µM dNTP, 10 pmol forward primer, 10 pmol reverse primer and 1 U Taq polymerase. The volume was adjusted to 50 µl by adding $ddH_2O$. $MgCl_2$ concentration was optimized for each primer pair by setting up a series of PCR

reactions using a range of $MgCl_2$ concentrations. Table 8 lists the appropriate $MgCl_2$ concentration for each exon.

For radioactive PCR, the reaction volume was 25 µl and contained 2.5 µl 10X PCR buffer, 1.5-3.0 mM $MgCl_2$, 10 µM dNTP, 10 pmol forward primer, 10 pmol reverse primer, 1 U Taq polymerase and 1 µCi $^{33}$P-dATP. In general, 50-100 ng DNA was used as a template. The volume was adjusted to 25 µl by adding $ddH_2O$.

Amplification was performed in the GeneAmp PCR with the following parameters: initial denaturation at 95°C for 3 min; 30 cycles of 95°C for 30 sec (denaturation), 55°C for 30 sec (annealing), 72°C for 30 sec (extension); and a final extension at 72°C for 10 min. After the PCR cycles were completed, the tubes were held at 4°C for at least 5 minutes or until removal.

**Table 8: Optimum $MgCl_2$ concentrations for *hMLH1* exon PCR**

| Exons | Primers | [$MgCl_2$] mM |
|---|---|---|
| Exon 1 | GA 281/282 | 1.5-3.0 |
| Exon 2 | GA 283/284 | 1.5-3.0 |
| Exon 3 | GA 152/153 | 3.0 |
| Exon 4 | GA 285/286 | 3.0 |
| Exon 5 | GA 156/157 | 3.0 |
| Exon 6 | GA 287/288 | 1.5-3.0 |
| Exon 7 | GA 289/290 | 1.5-3.0 |
| Exon 8 | GA 291/292 | 1.5-3.0 |
| Exon 9 | GA 164/165 | 1.5-3.0 |
| Exon 10 | GA 333/334 | 3.0 |
| Exon 11 | GA 337/338 | 3.0 |
| Exon 12 | GA 339/340 | 3.0 |
| Exon 13 | GA 343/344 | 3.0 |
| Exon 14 | GA 293/294 | 1.5-3.0 |
| Exon 15 | GA 172/173 | 3.0 |
| Exon 16 | GA 345/346 | 3.0 |
| Exon 17 | GA 176/177 | 1.5-3.0 |
| Exon 18 | GA 178/179 | 1.5-3.0 |
| Exon 19 | GA 180/181 | 1.5-3.0 |

### 2.2.3. Restriction Enzyme Digestion

Restriction enzyme digestions with *Dde*I and *Hph*I were performed in 20 µl reaction volumes using the reaction buffer recommended by the manufacturer. One unit of enzyme was sufficient to digest approximately 30-50 ng PCR product. The PCR template was quantitated by agarose gel electrophoresis using a DNA size marker of known concentration. The incubation temperature was 37°C for both enzymes. After digestion, heat inactivation was performed at 65°C.

After incubation the cut and uncut PCR fragments were analyzed by agarose gel electrophoresis. DNA size markers were used to calculate the sizes of the bands by drawing a standard curve for the particular marker used. Electrophoresis was performed using 4% NuSieve 3:1 agarose at 5 V/cm in 1X TBE for 2.5 hours. After the electrophoresis, the gel was stained in EtBr (1 µg/ml) for 20 minutes and then destained by two 15-minute washes with distilled water.

### 2.2.4. Heteroduplex Analysis (HA)

HA was performed using the EC 160 (E-C Apparatus Corp., Holbrook, NY, USA) electrophoresis apparatus and the Power-PAC 3000 power supply (BioRad, Hercules, CA, USA) with the following acrylamide and MDE gel formulations:

**6% acrylamide (100 ml)**

15 ml acrylamide 19:1
75 ml $ddH_20$
10 ml 10X TBE
400 µl 10% APS
40 µl TEMED

**MDE gel (75 ml)**

28 ml MDE gel (2X)
36.4 ml $ddH_20$
4.5 ml 10X TBE
5.6 ml Glycerol
400 µl 10% APS
40 µl TEMED

All gel components except for APS and TEMED were mixed and just before pouring 10% APS and TEMED were added. The gel was allowed to polymerize for at least 2 hours. The standard running buffer was 0.6X TBE for MDE gel and 1X TBE for acrylamide gel.

The HA reaction was performed by heating 15 µl PCR product at 95°C for 5 minutes in the Perkin Elmer GeneAmp PCR 9600 system to denature the double stranded DNA. The denatured DNA samples were allowed to anneal at room temperature for 30 minutes. Then 3µl 6X loading buffer (15% ficoll, 0.05% xylene cyanol, 0.05% bromophenol blue) was added to each tube and 4µl was loaded onto the gel. Electrophoresis was performed at 5W constant power at room temperature (25°C).

After electrophoresis, the DNA bands were visualized by silver staining. The gel was washed two times with $ddH_2O$ for one minute, then treated two times with fixative solution for 3 minutes and again washed with $ddH_2O$. In the next step the gel was stained for 10 minutes with 0.1% silver nitrate solution. The silver nitrate solution was discarded and the gel was rapidly washed twice with $ddH_2O$. Finally, the gel was treated with ice-cold developer solution for about 30 minutes. Once the DNA bands were observed, the gel was placed into 7.5% acetic acid solution for 10 minutes to stop the reaction.

### 2.2.5. Single-Strand Conformation Polymorphism (SSCP)

Three different SSCP gel formulations were evaluated: 10% 75:1 acrylamide: bisacrylamide, 6% 75:1 acrylamide: bisacrylamide and 10% 99:1 acrylamide: bisacrylamide. All gels contained 10% glycerol, a neutral compound that alters the

strand separation and increases the resolution. The acrylamide gel formulations are given below:

| **10% acrylamide (100 ml)** | **6% acrylamide (100 ml)** |
|---|---|
| 25 ml acrylamide 75:1 or 99:1 | 15 ml acrylamide 75:1 |
| 59 ml $ddH_20$ | 69 ml $ddH_20$ |
| 6 ml 10X TBE | 6 ml 10X TBE |
| 10 ml glycerol | 10 ml glycerol |
| 400 µl 10% APS | 400 µl 10% APS |
| 40 µl TEMED | 40 µl TEMED |

All the gel components except APS and TEMED were combined and filtered using a Nalgene apparatus with 0.2 µm cellulose acetate filters (Corning Costar Corp., Cambridge, MA, England) and a Hetomaster jet water pump (Heto-Holten A/S, Allerød, Denmark). Vacuum was applied for 5 minutes in order to degas the acrylamide solution. Immediately before pouring the gel 10% APS and TEMED were added. The gel was allowed to polymerize for at least 2 hours. The standard running buffer was 0.6X TBE. For gels run on the Genomyx system, the upper buffer was 0.6X TBE and the lower buffer was 1X TBE.

Three gel electrophoresis apparatuses were used: EC 160 and EC 175 (E-C Apparatus Corp., Holbrook, NY, USA) and Genomyx (Genomyx, Foster City, CA, USA). A Power-PAC 3000 power supply (BioRad, Hercules, CA, USA) was used with the EC 160 and EC 175. A comparison of the gel sizes and electrophoresis conditions used with each is given in Table 9.

The SSCP reaction was performed by mixing 2 µl PCR product with 18 µl SSCP loading buffer (95% formamide, 10 mM NaOH, 0.25% bromophenol blue, 0.25% xylene cyanol), heating at 95°C for 2 minutes and putting the tubes on ice for 5 minutes. The samples were then immediately loaded onto the gel. Electrophoresis

was performed at constant power (5, 20 or 45W) depending on the gel apparatus, at room temperature (25°C) or at 4°C. After electrophoresis, the gel was removed from glass plates using Whatman 3MM paper and dried on the SGD 2000 slab gel dryer (Savant, Holbrook, NY, USA) at 80°C for 2 hours. Radioactive bands were detected by autoradiography or phosphoimaging.

**Table 9: Comparison of Gel Apparatuses used for SSCP Analysis**

| | EC 160 | EC 175 | Genomyx |
|---|---|---|---|
| Gel Size | 33 cm x 42 cm | 15 cm x 28 cm | 35 cm x 65 cm |
| Well Capacity | 4-6 µl | 20 µl | 4 µl |
| Power | 45 W | 5 W | 20W |
| Voltage | 1500-2200 V | 400-1000 V | 1300-2000 V |
| Current | 20-30 mA | 5-13 mA | - |
| Temperature | 4°C | 4°C and 25°C | 25°C |

For autoradiography, the dried gel was placed into a cassette with X-ray film (Kodak, Rochester, NY, USA) and an intensifying screen. The cassette was placed at -80°C for exposure. After the exposure, the film was developed using a Hyperprocessor developer (Amersham, Buckinghamshire, England).

Phosphoimaging was performed using the Molecular Imager® System, GS-525 (BioRad, Hercules, CA, USA). The dried gel was placed into the cassette and a foil protector placed on top. A clean, erased image screen was then placed over the gel. After exposure, the screen was placed in the scanner and the data captured using Multi-Analyst software. This method allows one to adjust the contrast and brightness for the best band resolution.

### 2.2.6. DNA Sequence Analysis

DNA sequence analysis of *hMLH1* exon 15 was performed for samples 98-05 and 97-630 using primers GA 172 (forward primer) for the sense strand sequence and GA 173 (reverse primer) for the antisense strand sequence. The primers for *hMLH1* exon 7 GA 289 (forward primer) for the sense strand sequence and GA 290 (reverse primer) for the antisense strand sequence. The PCR template was quantitated by agarose gel electrophoresis using a DNA size marker of known concentration. Cycle sequencing reactions were set up in 0.2 Thermowell$^{TM}$ tubes (Corning Costar Corp., Cambridge, MA, England) as follows:

**Cycle sequencing reaction**

| | |
|---|---|
| 3-6 µl | Template (10-30 ng/µl PCR product) |
| 1 µl | Primer (3.2 pmol) |
| 4 µl | Terminator ready reaction mix |
| q.s. | $ddH_20$ |
| 20 µl | |

Cycle sequencing reactions were performed in either the Perkin Elmer GeneAmp PCR 9600 system or 2400 thermal cyclers according to the following protocols:

**Cycle Sequencing Protocol**

| **Sense Strand (GA 172)** | | **Antisense Strand (GA 173)** | |
|---|---|---|---|
| 98°C x 20 sec | | 98°C x 20 sec | |
| 98°C x 10 sec | | 96°C x 10 sec | |
| 50°C x 5 sec | 30 cycles | 50°C x 5 sec | 30 cycles |
| 60°C x 4 min | | 60°C x 4 min | |
| 4°C hold | | 4°C hold | |

Sequencing products (20 µl) were precipitated by adding 68 µl cold absolute EtOH and 7 µl 7.5 M ammonium acetate and placing on ice for 15 min. The DNA

was pelleted by centrifuged at 13,000 rpm for 15 min, washed with 250 µl cold 70% EtOH and air dried. The DNA pellet was resuspended in 6 µl sequencing loading buffer and 2 µl were loaded on the gel.

Polyacrylamide gel electrophoresis was performed using standard conditions on a 377 Prism sequencer (ABI). The standard sequencing gel contained 4% acrylamide/bisacrylamide (19:1), 7 M urea, and 1X TBE:

**4% sequencing gel solution (50 ml)**

| | |
|---|---|
| 19 g | urea dissolved in 25 ml $ddH_2O$ |
| 5 ml | 40% acrylamide / bisacrylamide (19:1) |
| 5 ml | 10X TBE |
| q.s. | $ddH_20$ (to 50 ml) |
| 250 µl | 10% APS |
| 34 µl | TEMED |

Electrophoresis conditions were 29 W, 18 mA, 1.65 kV and the temperature was 50°C. Electrophoresis was carried out for approximately 5-7 hours depending on the maximum fragment length.

The data was collected by ABI data collection software during electrophoresis. After the run finished, the collected data was automatically analyzed by ABI sequencing analysis software. The analyzed files were then imported into the Factura program to identify the unambiguous sequences. These analyzed sequences were then imported into the Sequencing Navigator for alignment with the reference DNA sequence.

# 3. Results

## 3.1. DNA Isolation

Genomic DNA was isolated from patient blood using a phenol chloroform extraction method. All extracted DNAs were evaluated quantitatively and qualitatively by UV spectrophotometry and agarose gel electrophoresis. An example is shown in Figure 11. Although the sample in lane 1 is high molecular weight DNA, the $A_{260}/A_{280}$ ratio was 1.3, which indicated poor quality. The extraction was repeated but the same results were obtained, probably because the blood sample had partially coagulated. The samples in lanes 2-4 all have high molecular weight DNA of good quality. Their $A_{260}/A_{280}$ ratios were ~1.7. The concentrations of these DNA samples are 100, 50, 182 and 86 μg/ml, respectively.

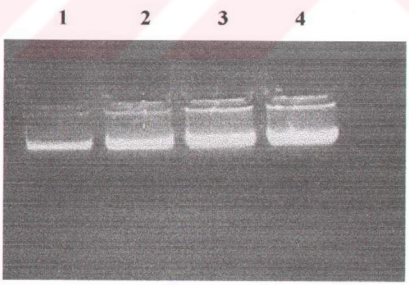



**Figure 11: Analysis of extracted genomic DNA.**
Samples were electrophoresed through a 1% agarose gel at 8V/cm for 1 hour. Lane 1: sample 96-16, Lane 2: sample D, Lane 3: sample MR1, Lane 4: sample MR2.

### 3.2. Polymerase Chain Reaction

All PCR products were evaluated by agarose gel electrophoresis. The samples shown in Figure 12 were amplified by *hMLH1* exon 10 primers, GA 333 and GA 334. The expected size of the PCR product was 227 bp and this was confirmed by comparison of the observed DNA band with the size marker. The negative control sample has no DNA band indicating that the PCR mixture contained no contaminating DNA.

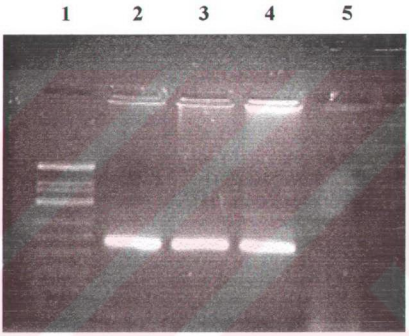



**Figure 12: Analysis of PCR products.**
Samples were electrophoresed through a 2% agarose gel at 8V/cm for 35 min. Lane 1: ∅X174 DNA *Hin*fI marker, Lane 2: sample 96-16, Lane 3: sample 96-17, Lane 4: sample 96-19, Lane 5: negative control.

### 3.3. Detection of G884C *hMLH1* mutation by *Dde*I digestion

A G884C point mutation was previously identified in the proband of the HNPCC 1 kindred (Wang et al., 1997). This point mutation is a missense mutation that results in a Ser→Thr amino acid substitution at codon 295. The mutation occurs within the recognition site of the restriction enzyme *Dde*I and results in a loss of the restriction site. Thus, digestion of the *hMLH1* exon 10 PCR product with *Dde*I might

provide a simple, cost-effective assay for identifying other members of the HNPCC 1 kindred who inherited the same mutation.

*Dde*I has two recognition sites in the wild type *hMLH1* exon 10 DNA sequence. The PCR amplified fragment is 227 bp. After *Dde*I digestion the expected fragment sizes are 132 bp, 69 bp and 26 bp (schematically depicted in Figure 13). When there is a G884C mutation, one *Dde*I site is lost. In a heterozygous individual the expected fragment sizes are 201 bp, 132 bp, 69 bp and 26 bp (Figure 13).

Amplified *hMLH1* exon 10 PCR products were incubated with *Dde*I in the recommended buffer at 37°C for 2 hours. As a control, amplified *hMLH1* exon 10 PCR products were also incubated in the same *Dde*I buffer except without the enzyme. After digestion the DNA samples were analyzed by agarose gel electrophoresis. A standard curve for the ∅X174 DNA *Hin*fI fragments was drawn showing the distance migrated (mm) vs the size (logarithm bp). The sizes of the *Dde*I-digested DNA fragments were then calculated using this curve (data not shown).

The *Dde*I digestion results for individual 96-18 are shown in Figure 14. From the digestion profile it is obvious that individual 96-18 is heterozygous. Both the 96-18 sample and the control sample show a single band when *Dde*I is not added to the reaction buffer. However, after *Dde*I digestion there are three visible bands in the 96-18 sample but only two visible bands in the control sample (the 26 bp fragment is too small to detect with this agarose concentration). The sizes of the *Dde*I-digested bands in the 96-18 sample were calculated to be 200, 132 and 65 bp. The sizes of the *Dde*I-digested bands in the control sample were calculated to be 132 bp and 65 bp. These values are very close to the expected values. Thus, one can conclude that individual 96-18 is heterozygous for the G884C mutation.

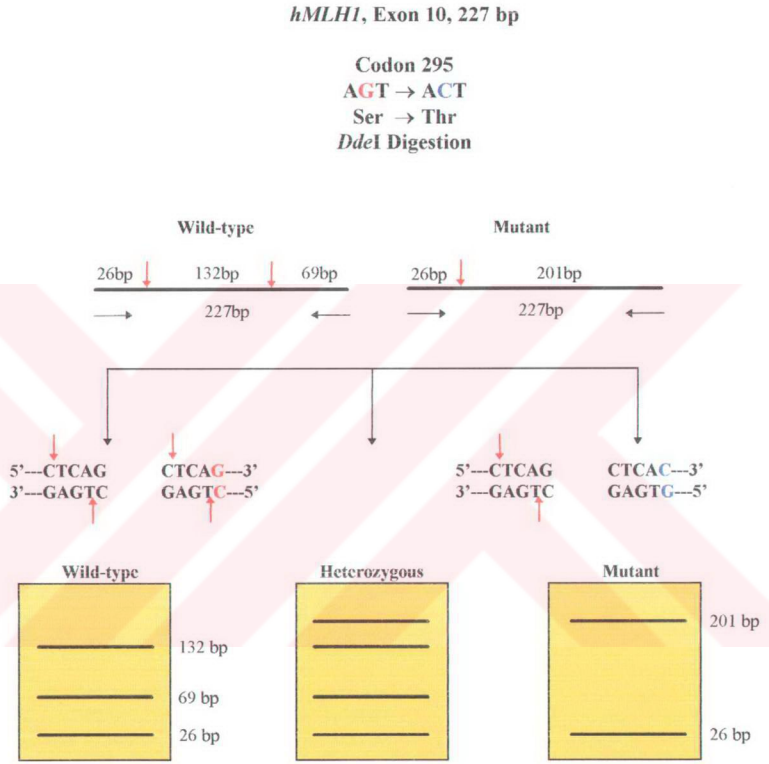



**Figure 13: Expected *Dde*I fragment sizes for wild-type, heterozygous and homozygous mutant individuals.**
Red arrows indicate the recognition / cleavage sites of *Dde*I.

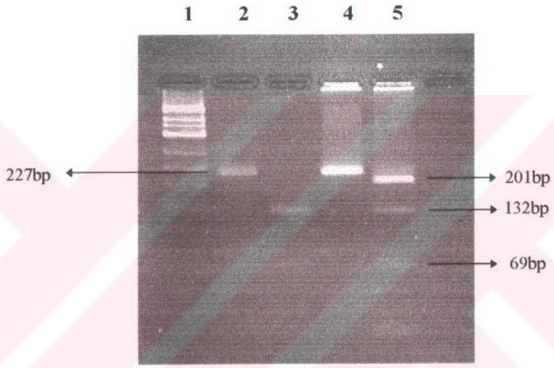



**Figure 14: *DdeI* digestion profile.**
DNA fragments were electrophoresed in 4% NuSieve 3:1 agarose at 5V/cm for 2.5 hours. Lane 1: ∅X174 DNA *Hinf* I marker, Lane 2: undigested control, Lane 3: *Dde*I-digested control, Lane 4: undigested 96-18, Lane 5: *Dde*I-digested 96-18.

DNA samples of five family members from the HNPCC 1 kindred (Figure 6) were analyzed using this method. Initially, conclusive results were obtained for only three of the five family members; two DNA samples were only partially digested with *Dde*I. Increasing the increased amount of enzyme and/or the time of incubation did not help. However, complete digestion was obtained when the PCR products were purified with QIAquick PCR purification kit (QIAGEN, Chatsworth, CA, USA) prior to *Dde*I digestion. The results are presented in Table 10. Two individuals (96-17 and 96-19) have two wild-type alleles. Three family members (96-15, 96-16 and 96-18) are heterozygous for the mutant allele. These results were confirmed DNA sequence analysis (performed by B. Cevher).

**Table 10: Detection of G884C *hMLH1* mutation by *Dde*I digestion**

| Family Member | Genotype *Dde*I digestion | Phenotype | Age |
|---|---|---|---|
| 96-15 | + / - | colon cancer | 39[a] |
| 96-16 | + / - | no disease | 18 |
| 96-17 | + / + | no disease | 75 |
| 96-18 | + / - | no disease | 9 |
| 96-19 | + / + | no disease | 17 |

[a]Age at diagnosis.

### 3.4. Detection of A1652C *hMLH1* mutation by *Hph*I digestion

An A1652C point mutation was previously identified in the proband of the HNPCC 2 kindred (Wang et al., 1997). This point mutation is a missense mutation that results in a Asn→Thr amino acid substitution at codon 551. This mutation creates a recognition site for the restriction enzyme *Hph*I. Because *Hph*I cleaves the mutant sequence but does not cleave the wild-type sequence, digestion of the *hMLH1* exon 14 PCR product with *Hph*I might provide a simple, cost-effective assay for identifying other members of the HNPCC 2 kindred who inherited the same mutation.

*Hph*I has no recognition site in the wild type *hMLH1* exon 14 DNA sequence. The PCR amplified fragment is 254 bp. The A1652C mutation introduces an *Hph*I site. In a heterozygous individual the expected fragment sizes are 254 bp, 183 bp and 71 bp (schematically depicted in Figure 15).

Amplified *hMLH1* exon 14 PCR products were incubated with *Hph*I in the recommended buffer at 37°C for 2 hours. As a control, amplified *hMLH1* exon 14 PCR products were also incubated in the same *Hph*I buffer, except without the enzyme. After digestion the DNA samples were analyzed by agarose gel electrophoresis. A standard curve for the ∅X174 DNA *Hin*fI fragments was drawn showing the distance migrated (mm) vs the size (logarithm bp). The sizes of the *Hph*I-digested DNA fragments were then calculated using this curve (data not shown).

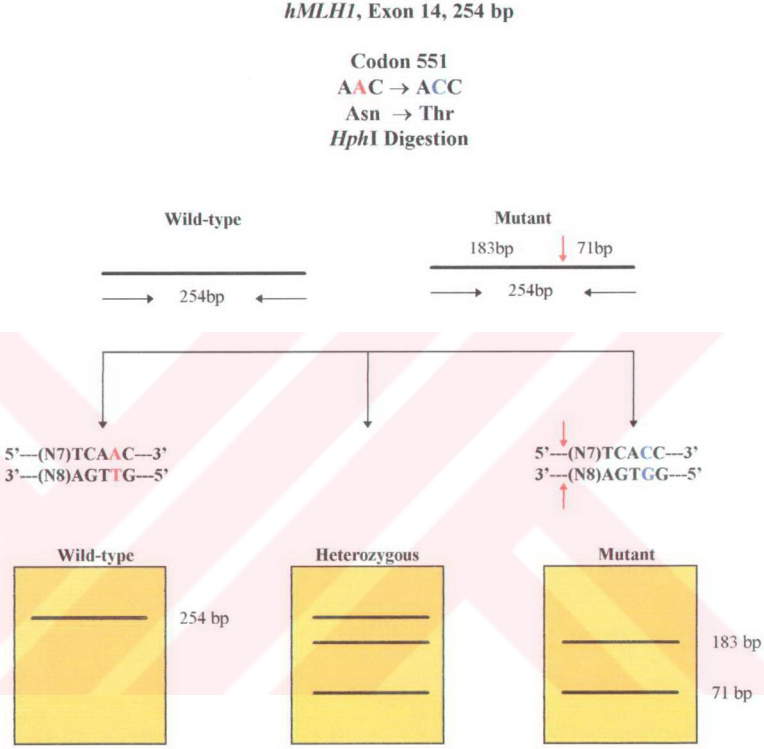



**Figure 15: Expected *Hph*I fragment sizes for wild-type, heterozygous and homozygous mutant individuals.**
Red arrows indicate the recognition / cleavage site of *Hph*I.

The *Hph*I digestion results for individual 97-160 are shown in Figure 16. From the digestion profile it is obvious that individual 97-160 is heterozygous. Both the 97-160 sample and the control sample show a single band when *Hph*I is not added to the reaction buffer. However, after *Hph*I digestion there are three bands visible in the 97-160 sample but only one band visible in the control sample. The sizes of the *Hph*I -digested bands in the 97-160 sample were calculated to be 251, 182 and 70 bp. The size of the *Hph*I -digested band in the control sample was calculated to be 251 bp. These values are very close to the expected values. Thus, one can conclude that individual 97-160 is heterozygous for the A1652C mutation.

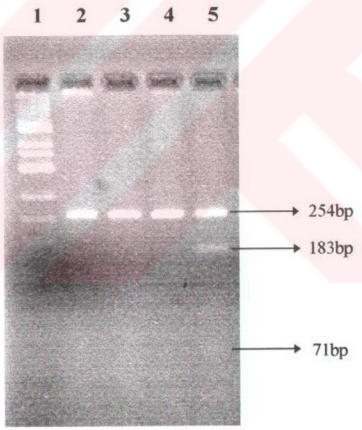



**Figure16: *Hph*I digestion profile.**
DNA fragments were electrophoresed in 4% NuSieve 3:1 agarose at 5V/cm for 2.5 hours. Lane 1: ∅X174 DNA *Hinf* I marker, Lane 2: undigested control, Lane 3: *Hph*I-digested control, Lane 4: undigested 97-160, Lane 5: *Hph*I-digested 97-160.

DNA samples of 29 family members from the HNPCC 2 kindred (Figure 7) were analyzed using this method. The results are presented in Table 11. Conclusive results were obtained for all family members tested. Seventeen individuals have an *Hph*I digestion profile consistent with wild-type. Twelve individuals have an *Hph*I digestion profile consistent with heterozygosity. DNA sequence analysis confirmed the results for twelve of thirteen samples. Sample 97-184 was reported to be wild-type (Wang et al., 1997). However, repeat DNA sequence analysis of sample 97-184 (performed by B. Cevher) verified that our DNA sample was heterozygous for the mutation.

**Table 11: Detection of A1652C *hMLH1* mutation by *Hph*I digestion**

| Family Member | Genotype *Hph*I digestion | DNA Sequence Confirmation | Phenotype | Age |
|---|---|---|---|---|
| 97-160 | + / - | yes | colon cancer | 47[a] |
| 97-161 | + / - | yes | colon cancer | 33[a] |
| 97-162 | + / + | yes | no disease | 39 |
| 97-163 | + / - | yes | no disease | 34 |
| 97-164 | + / - | yes | no disease | 31 |
| 97-165 | + / + | yes | ovarian cyst | 45 |
| 97-166 | + / - | not done | no disease | 40 |
| 97-167 | + / + | not done | no disease | 40 |
| 97-168 | + / - | not done | no disease | 36 |
| 97-169 | + / + | not done | no disease | 18 |
| 97-170 | + / + | not done | no disease | 14 |
| 97-171 | + / + | yes | no disease | 51 |
| 97-172 | + / + | not done | no disease | 22 |
| 97-173 | + / + | not done | no disease | 18 |
| 97-174 | + / + | yes | no disease | 45 |
| 97-175 | + / + | not done | no disease | 41 |
| 97-176 | + / + | not done | no disease | 26 |
| 97-177 | + / - | not done | no disease | 32 |
| 97-178 | + / + | not done | no disease | 28 |
| 97-179 | + / + | not done | no disease | 48 |
| 97-180 | + / - | not done | no disease | 14 |
| 97-181 | + / + | not done | no disease | 18 |
| 97-182 | + / - | yes | no disease | 32 |
| 97-183 | + / - | yes | no disease | 45 |
| 97-184 | + / - | yes | no disease | 41 |
| 97-185 | + / + | not done | no disease | 43 |
| 97-186 | + / + | not done | no disease | 43 |
| 97-187 | + / - | yes | no disease | 34 |
| 97-188 | + / + | yes | no disease | 33 |

[a]Age at diagnosis.

### 3.5. Heteroduplex Analysis

Four cell lines with known mutations were analyzed by heteroduplex analysis. The samples were run on 6% non-denaturing gel electrophoresis and the bands were observed by silver staining. However, we were not able to detect any migration difference between the mutant PCR products and the control PCR products. This is probably due to the insensitivity of the method with silver staining.

### 3.6. Single Strand Conformation Polymorphism

#### 3.6.1. Detection of known mutations

DNA samples with known mutations were used as controls to define SSCP conditions that would detect both point mutations and insertion/deletion mutations. Use of these samples demonstrated the accuracy and the efficiency of a specific method in mutation screening. The nine DNA samples listed in Table 12 represent three transversion mutations, two transition mutations and one small deletion.

**Table 12: Defined *hMLH1* mutations**

| Exon | Mutation (Type) | Codon Change | Sample |
|---|---|---|---|
| 3 | G293C (transversion) | Gly98Ala | HNPCC 4 |
| 6 | A544G (transition) | Arg182Gly | CL 11872 |
| 8 | A655G (transition) | Ile219Val* | HNPCC 4 |
| 8 | C676T (transition) | Arg226stop | HNPCC 3 |
| 10 | G884C (transversion) | Ser295Thr | HNPCC 1 |
| 14 | A1652C (transversion) | Asn551Thr | HNPCC 2 |
| 15 | G1672T (transversion) | Glu558stop | CL 13481 |
| 16 | 1846-1848 deletion | del Lys 616 | CL 12826 |
| 17 | G1989T (transversion) | Glu663Asp | CL 13158 |

* Polymorphism

The nine DNA samples with known mutations were analyzed using seven SSCP protocols. The ability of the different protocols to detect the known mutations is summarized in Table 13. The three base pair deletion was the easiest mutation to detect; it was detected using five protocols (1, 2, 5, 6 and 7). In terms of the variety of mutations detected, protocol 2 was the most successful, detecting five of the nine mutations. Both protocol 1 and protocol 7 were able to detect multiple mutations (three and two, respectively). Protocols 3, 4, 5 and 6 detected one mutation each. The two A→G transition mutations were not detected using any of these protocols.

**Table 13: Detection of known mutations by SSCP**

| Mutation | Protocol | Apparatus | Acryl % | Acryl:Bis | Power | Temp | Time |
|---|---|---|---|---|---|---|---|
| G293C | 1 | EC 160 | 10% | 99:1 | 45 W | 4°C | 14 hrs |
| A544G | not detected | | | | | | |
| A655G | not detected | | | | | | |
| C676T | 2 | EC 175 | 10% | 99:1 | 5W | 4°C | 30 hrs |
| G884C | 7 | Genomyx | 6% | 75:1 | 20W | 25°C | 10 hrs |
| A1652C | 2 | EC 175 | 10% | 99:1 | 5W | 4°C | 22 hrs |
| | 3 | EC 175 | 10% | 75:1 | 5W | 4°C | 30 hrs |
| G1672T | 1 | EC 160 | 10% | 99:1 | 45W | 4°C | 14 hrs |
| | 2 | EC 175 | 10% | 99:1 | 5W | 4°C | 30 hrs |
| 1846-48 | 1 | EC 160 | 10% | 99:1 | 45W | 4°C | 14 hrs |
| ΔAAG | 2 | EC 175 | 10% | 99:1 | 5 W | 4°C | 30 hrs |
| | 5 | Genomyx | 10% | 99:1 | 20W | 25°C | 15 hrs |
| | 6 | Genomyx | 10% | 75:1 | 20W | 25°C | 13 hrs |
| | 7 | Genomyx | 6% | 99:1 | 20W | 25°C | 10 hrs |
| G1989T | 2 | EC 175 | 10% | 99:1 | 5W | 4°C | 14 hrs |
| | 4 | EC 175 | 10% | 75:1 | 5W | 25°C | 30 hrs |

Sample 97-630 is from an HNPCC individual with a G293C transversion in exon 3 that results in a Gly→Ala substitution at codon 98. When analyzed by SSCP protocol 1, this PCR product showed four DNA bands whereas PCR products from an unrelated HNPCC individual (98-05) and a normal control showed only two DNA bands (Figure 17). This mutation was not detected using any other SSCP protocol.

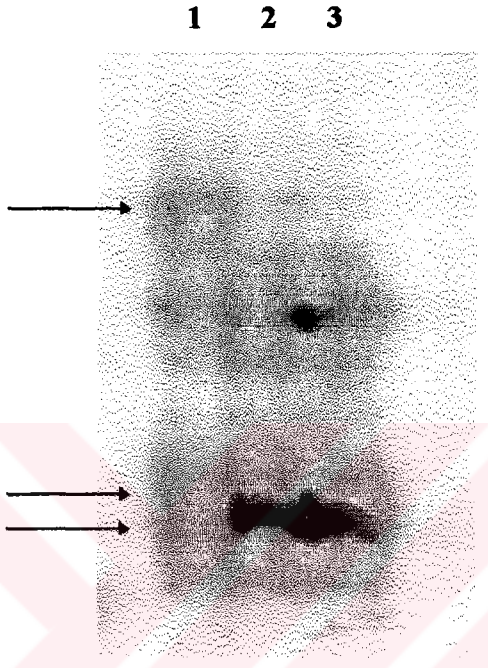



**Figure 17: Detection of G293C transversion [codon 98].**
Lane 1: sample 97-630, Lane 2: sample 98-05, Lane 3: control sample. SSCP gel was 10% 99:1 acrylamide:bisacrylamide. Electrophoresis was performed on the EC 160 at 45W constant power at 4°C for 14 hours. Autoradiography was at -80°C for 5 days. Arrows indicate shifted bands in mutant sample.

Sample 97-195 is from an HNPCC individual with a C676T transition mutation in exon 8 that results in an Arg→stop substitution and protein truncation at codon 226. When analyzed by SSCP protocol 2, this PCR product showed three bands whereas the control PCR product showed only two bands (Figure 18). This mutation was not detected using other SSCP protocols.


**Figure 18: Detection of C676T transition [codon 226].**
Lane 1: sample 97-195, Lane 2: control sample. SSCP gel was 10% 99:1 acrylamide: bisacrylamide. Electrophoresis was performed on the EC 175 at 5W constant power at 4°C for 30 hours. Autoradiography was at -80°C for 5 days. Arrow indicates shifted band in mutant sample.

Sample 96-15 is from an HNPCC individual with a G884C transversion mutation in exon 10 that results in a Ser→Thr substitution at codon 295. When analyzed by SSCP protocol 7, this PCR product showed an altered mobility compared to the control PCR product. Both samples have three bands but the patterns of these bands is different (Figure 19). The uppermost band of sample 96-15 is higher than the uppermost band of the control sample. This mutation was not detected using any other SSCP protocol.

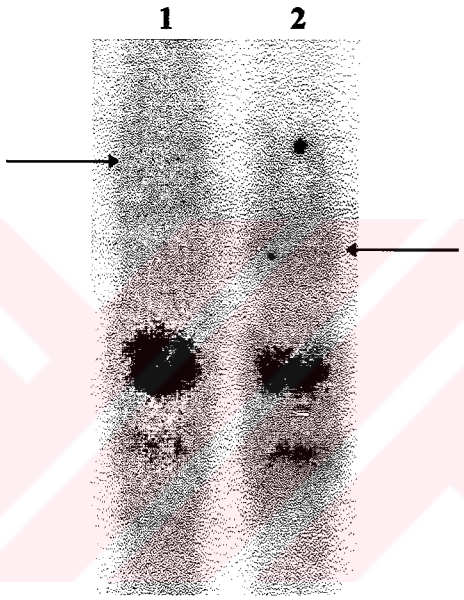



**Figure 19: Detection of G884C transversion [codon 295].**
Lane 1: sample 96-15, Lane 2: control sample. SSCP gel was 6% 75:1 acrylamide: bisacrylamide. Electrophoresis was performed on Genomyx at 20W constant power at 25°C for 10.5 hours. Autoradiography was at -80°C for 8 days. Arrows indicate bands with differential mobility.

Samples 97-160 and 97-161 are from two HNPCC individuals from the same family. The inherited mutation in this family is an A1652C transversion in exon 14 that results in an Asn→Thr substitution at codon 551. When analyzed by SSCP protocol 2, the mutant PCR products displayed two bands, while the control sample showed only one (Figure 20). This mutation was also detected using SSCP protocol 3 (data not shown).

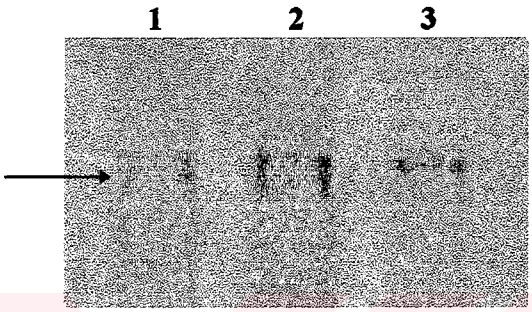



**Figure 20: Detection of A1652C transversion [codon 551].**
Lane 1: sample 97-160, Lane 2: sample 97-161, Lane 3: control sample. SSCP gel was 10% 75:1 acrylamide:bisacrylamide. Electrophoresis was performed on the EC 175 at 5W constant power at 4°C for 22 hours. Autoradiography was at -80°C for 3.5 days. Arrow indicates extra band in mutant samples.

The cell line 13481 has a G1672T transversion mutation in exon 15 that results in a Glu→stop substitution at codon 558. When analyzed by SSCP protocol 2, the mutant PCR product shows four bands, while the control PCR product shows only two bands (Figure 21). This mutation was also detected using protocol 1 (data not shown).

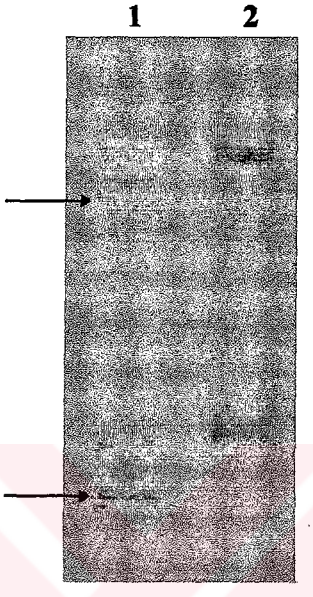



**Figure 21: Detection of G1672T transversion [codon 558].**
Lane 1: cell line 13481, Lane 2: control sample. SSCP gel was 10% 99:1 acrylamide: bisacrylamide. Electrophoresis was performed on the EC 175 at 5W constant power at 4°C for 30 hours. Autoradiography was at -80°C for 5 days. Arrows indicate extra bands seen in mutant sample.

The cell line 12826 has 3 bp deletion in exon 16 that results in deletion of Lys codon 616. This mutation was detected by five SSCP protocols: numbers 1, 2, 5, 6 and 7. The SSCP results using protocol 7 are shown in Figure 22. The mutant PCR product migrates as two distinct single-stranded molecules, while the control PCR product migrates as a single broader band.

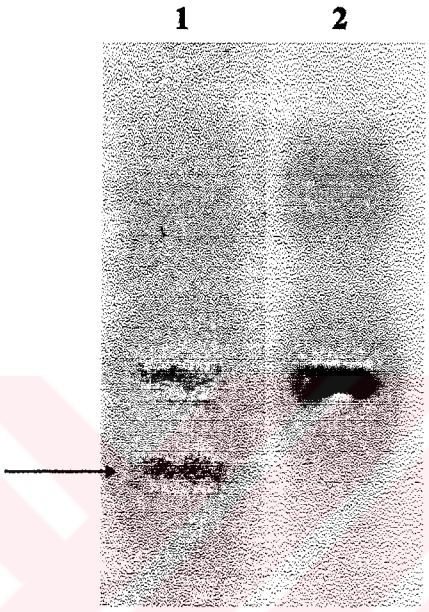



**Figure 22: Detection of AAG deletion [codon 618].**
Lane 1: cell line 12826, Lane 2: control sample. SSCP gel was 6% 75:1 acrylamide: bisacrylamide. Electrophoresis was performed on the Genomyx apparatus at 20W constant power at 25°C for 10.5 hours. Autoradiography was at -80°C for 8 days. Arrow indicates extra band present in mutant sample.

The cell line 13158 has a G1989T transversion mutation in exon 17 that results in a Glu→Asp substitution at codon 663. When analyzed using SSCP protocol 4, the mutant PCR product shows three bands, while the control PCR product shows only two (Figure 23). This mutation was also detected using SSCP protocol 2.

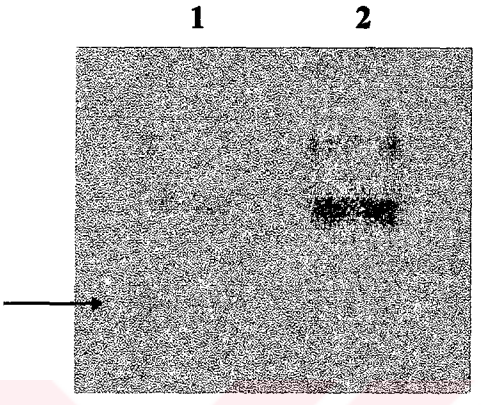



**Figure 23: Detection of G1989T transversion [codon 663].**
Lane 1: cell line 13158, Lane 2: control sample. SSCP gel was 10% 75:1 acrylamide: bisacrylamide. Electrophoresis was performed on the EC 175 at 5W constant power at room temperature for 14 hours. Autoradiography was at -80°C for 3 days. Arrow indicates extra band seen in mutant sample.

The A→G transition mutations in exons 6 and 8 were analyzed using all seven SSCP protocols. Unfortunately, no mobility difference was detected between these samples with known mutations and control samples.

### 3.6.2. Detection of unknown mutations

*hMLH1* mutation screening for the HNPCC 5 proband was carried out using the SSCP protocols previously determined to have good detection capabilities. All 19 exons in both proband and control DNA samples were amplified by radioactive PCR. The PCR products were then analyzed using three SSCP protocols (Table 14). The results are presented in Table 14.

**Table 14: SSCP Mutation Screening for HNPCC 5 proband**

| Exons | Apparatus | Acryl % | Power | Temp | Time | Mobility Difference |
|---|---|---|---|---|---|---|
| 1-19 | EC 160 | 10% 99:1 | 45W | 4°C | 14 hr | exon 7<br>exon 15 |
| 15 | EC 175 | 10% 99:1 | 5W | 4°C | 30 hr | exon 15 |
| 1-19 | Genomyx | 10% 99:1 | 20W | 25°C | 15 hr | exon 15 |

Comparison of the SSCP results for the HNPCC 5 proband with those for normal controls *hMLH1* revealed a differential strand mobility for exon 15 using all three SSCP protocols. As seen in Figure 24, three bands were observed for cell line 13481 (G1672T transversion, exon 15), sample 97-630 (G293C transversion, exon 3), and a normal control sample, whereas only two bands were observed for sample 98-05. Because the migration of sample 98-05 is markedly different from the other samples, this might reflect a DNA sequence difference.

SSCP analysis of *hMLH1* exon 7 also showed differential migration for the proband and the normal control (Figure 25). The control sample has three bands, whereas the 98-05 sample has only two bands. This difference might be due to differences in the DNA sequence.

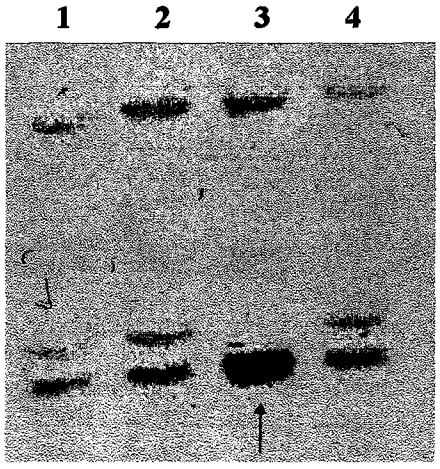



**Figure 24: Mobility shift in *hMLH1* exon 15 of the HNPCC 5 proband.**
Lane 1: cell line 13481, Lane 2: control sample, Lane 3: proband sample 98-05, Lane 4: sample 97-630. SSCP gel was 10% 99:1 acrylamide:bisacrylamide. Electrophoresis was performed on the Genomyx apparatus at 20W constant power at 25°C for 15 hours. Autoradiography was at -80°C for 5 days.

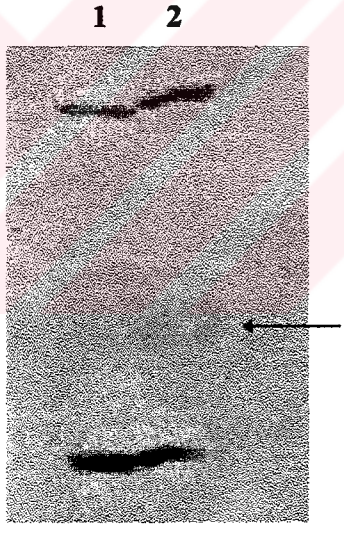



**Figure 25: Mobility shift in *hMLH1* exon 7 of the HNPCC 5 proband.**
Lane 1: sample 98-05, Lane 2: control sample. SSCP gel was 10% 99:1 acrylamide: bisacrylamide. Electrophoresis was performed on the EC 160 apparatus at 45W constant power at 4°C for 14 hours. Autoradiography was at -80°C for 6 days.

### 3.7. DNA Sequence Analysis

#### 3.7.1. *hMLH1* exon 15

Cycle sequencing reactions using template 98-05 DNA, 97-630 DNA and control DNA were carried out using the *hMLH1* exon 15 primers, GA 172 and GA 173 and the ABI PRISM kit. The DNA products were then analyzed by electrophoresis on the 377 DNA Sequencer with ABI Sequence Analysis software. Heterozygous bases were identified using the Factura software program. Finally, the DNA sequence was aligned with the reference sequence using the Sequence Navigator software program (Appendix C). The sample 98-05 exon sequence (nucleotides 89 to 175) was identical to the reference sequence. The only difference occurred at nucleotide 88 in the intron. At this position, the 98-05 DNA sequence was determined to be an A (adenine) residue, while the control DNA sequence was determined to be heterozygous "R" with one allele having an A (adenine) residue at this position and the other allele having a G (guanine) residue at this position (Figure 26). The 97-630 DNA sequence was also heterozygous at this position. Thus, the SSCP migration difference is most likely due to this intronic polymorphism.

#### 3.7.2. *hMLH1* exon 7

The nucleotide sequence of *hMLH1* exon 7 was determined for samples 98-05 and a control sample as described above, except that the exon 7-specific primers GA 289 and GA 290 were used. Until now, only the sense strand sequence has been obtained using the forward primer. Alignment of the 98-05 sequence with the reference sequence (Appendix D) reveals no discrepancy in the region for which clear data was available (nucleotides 51 to 93).

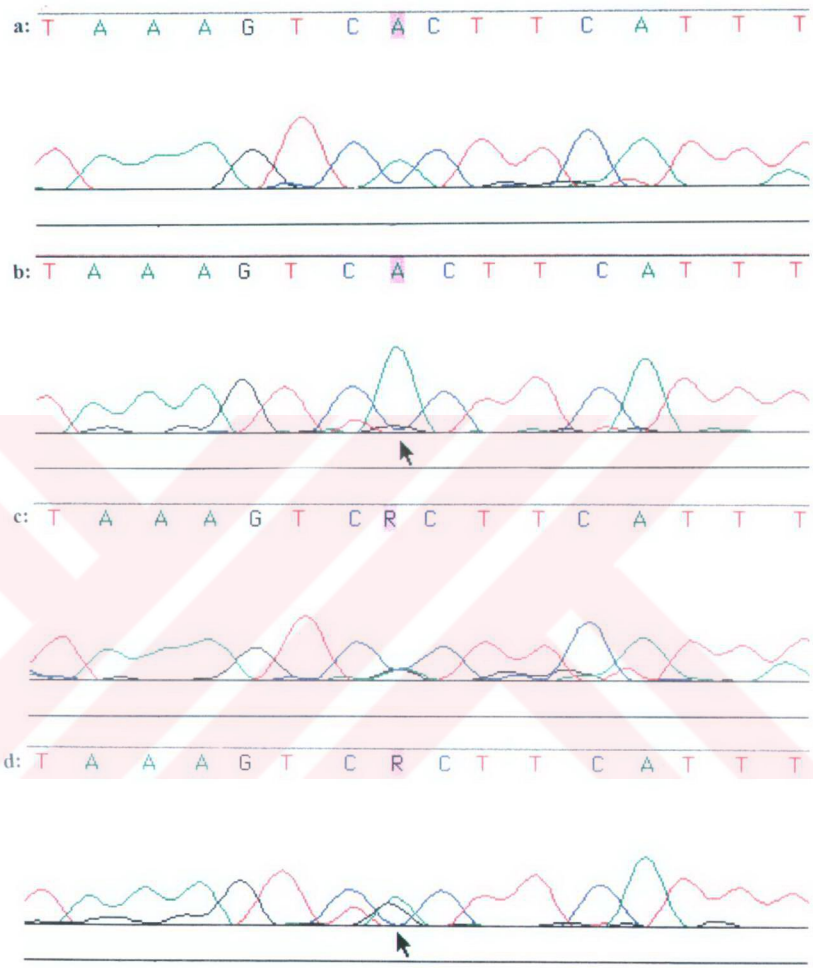



**Figure 26. Intronic polymorphism in *hMLH1* intron 15.**
a. Sample 98-05: DNA sequence obtained with forward primer; b. sample 98-05: DNA sequence obtained with reverse primer; c. control: DNA sequence obtained with forward primer; d. control: DNA sequence obtained with reverse primer. Arrow indicates intronic polymorphism in control DNA sample.

# 4. Discussion

HNPCC is a relatively common hereditary disease that affects approximately 1 in 200 to 1,000 individuals and accounts for 3% of all colorectal cancers (Boland, 1998). Defects in five DNA mismatch repair genes, *hMLH1*, *hMSH2*, *hPMS1*, *hPMS2*, and *hMSH6*, have been linked to HNPCC, with the majority of the mutations occurring in *hMLH1* and *hMSH2*. Individuals with a germline mutation in one of the DNA mismatch repair genes have a much greater risk of developing colorectal cancer.

Detection of mutations in these tumor susceptibility genes is extremely important for HNPCC individuals. If a germline mutation is identified, then preventive screening for colorectal tumors and other tumors typically associated with HNPCC can be done. If a tumor is discovered early, then surgical intervention might be curative. In addition, identification of unaffected individuals is also important. Individuals who do not have the mutation need not undergo unnecessary, invasive tests that otherwise would have been advised based on the family history.

This study proposed to identify possible germline mutations in the DNA mismatch repair gene, *hMLH1*, in HNPCC kindreds. Restriction enzyme analysis was used to analyze known mutations that affected a recognition/cleavage site. Single-strand conformation polymorphism (SSCP) analysis was used to scan the entire *hMLH1* coding region for unknown mutations. Finally, DNA sequence analysis was performed to confirm these results.

Although the Amsterdam criteria are very useful in defining HNPCC kindreds, they are often too restrictive. Germline mutations in mismatch repair genes have been demonstrated in families that did not meet the Amsterdam criteria (Beck et al., 1997b). The HNPCC 5 family does not fulfill the Amsterdam criteria because there are only two first-degree family members affected with colorectal cancer and because neither individual was diagnosed before the age of 50. This particular family was included as a potential HNPCC family because many other family members are affected with other cancers, such as stomach, larynx, esophagus, ovarian and possibly liver cancer. If the criteria are modified to include slightly older individuals with tumors typically found in HNPCC families, then the HNPCC 5 family can be included. Too rigid interpretation of the Amsterdam criteria may eliminate true HNPCC families, while a more relaxed interpretation may increase the identification of potential HNPCC families.

The specificity and sensitivity of the SSCP mutation detection method is not 100%. For example, SSCP is not able to detect whole exon deletions, which occur in some HNPCC families (Cotton, 1997). Moreover, some mutations, particularly point mutations and frameshift mutations are not detected by SSCP (Beck et al., 1997a) and this is a great concern. Although modification of SSCP conditions may increase the detection rate, this necessity of performing multiple screens increases both the cost and the time required for diagnosis.

In the HNPCC 1 family, three heterozygous individuals and two normal individuals were determined by *Dde*I analysis. These results were confirmed by DNA sequence analysis. Two of the heterozygous individuals are very young (9 and 18 years old). If and when the parents (and possible the teenager) are informed, it is

important that the genetic counselor be able to give appropriate psychological support.

The HNPCC 1 defect is a missense mutation G884C that causes a Ser→Thr substitution. The effect of this substitution on the structure and function of the MLH1 protein is not known. However, several observations support the conclusion that this is a pathogenic mutation and not simply a polymorphism. First, no G884C transversion was found in 50 unrelated individuals from Turkish population (Wang et al., 1997). Second, the G884C mutation was identified in tumors of affected family members with the loss of wild-type allele (Wang et al., 1997). The serine to threonine substitution is a conservative substitution. Both amino acids are hydrophilic with uncharged hydroxyl side chains. These amino acids often participate in hydrogen bond formation to other polar molecules. Threonine has longer side chain. Since this mutation is present near a conserved region (Figure 4), it might have an effect on the conformation and, therefore, function of the protein. Serine residues can be potential O-linked glycosylation sites or phosphorylation sites. However, it is not known whether the serine 295 residue has this function.

In the HNPCC 2 family, *Hph*I restriction enzyme analysis determined that 12 out of 29 individuals were heterozygous for the A1652C mutation. This mutation cosegregates with the disease as seen in the family pedigree (Figure 7). There is only one tested heterozygous individual who is older that 50 and does not have colorectal cancer. Two children were diagnosed with colorectal cancer at very early ages (13 and 7). The parents of the 13 year old girl were first cousins in a consanguineous marriage. Both father and mother tested heterozygous for the A1652C mutation, as did two of their three children tested. There is a 25% probability that the 13 year old daughter could have inherited a mutant allele from each parent and have been

homozygous for the mutation. The resulting MLH1-deficiency might have caused her to die at such an early age. This phenomenon has been previously observed in another Turkish HNPCC kindred (manuscript submitted, Ricciardone et al.).

Some of these restriction enzyme results were confirmed for several individuals by sequencing. There was one discrepancy: individual 97-184 was determined to be heterozygous by both *Hph*I restriction enzyme analysis and DNA sequence analysis at Bilkent, but wild-type by DNA sequence analysis in France (Wang et al., 1997). Isolation of DNA from a different sample aliquot, followed by *HphI* restriction enzyme analysis and DNA sequence analysis gave the same results. In order to eliminate any suspicion of sample mix-up, a new blood sample must be obtained from this individual and tested before any genetic counseling can be given.

This A1652C missense mutation results in a Asn→Thr substitution, whose effect on MLH1 structure and function is unknown. However, since no A1652C transversion was found in 50 unrelated individuals from Turkish population (Wang et al., 1997), this, too, is a probably pathogenic mutation and not simply a polymorphism. Asparagine and threonine are both hydrophilic amino acids with uncharged side chains. Asparagine 551 is a potential site (underlined in Figure 4) for glycosylation, a post-translational chemical modification needed for proper folding. Loss of this site might result in a change of protein structure and function.

In the HNPCC 5 family, *hMLH1* exons 15 and 7 showed a difference in migration between the proband and control samples by SSCP analysis. All the other exons were negative by SSCP analysis. In order to determine the cause of this differential mobility DNA sequence analysis was performed. No mutation was observed in the coding sequence of exon 15. However, an intronic polymorphism

was detected in the control DNA samples and this polymorphism most likely accounts for the SSCP differential mobility.

Preliminary DNA sequence analysis of exon 7 suggests that there is no nucleotide change within the region analyzed. The antisense sequence must be obtained using the reverse primer in order to confirm these preliminary results. If a sequence change is found in the coding region of this exon, further studies should be done on other family members to demonstrate that the mutation cosegregates with the disease. Other analyses must be done to determine whether the alteration is a pathogenic mutation or a polymorphism. The nature of the mutation will give some clues. Nonsense mutations and frameshift mutations are usually pathogenic, whereas missense mutations are often polymorphisms.

If there is no mutation in the coding sequence of exon 7, then there appears to be no germline *hMLH1* mutation. However, this conclusion must be tempered by the knowledge that the SSCP protocols that have been used were not capable of detecting all mutations. The three SSCP protocols that were used in this analysis were capable of detecting only seven out of nine (~80%) known mutations.

In order to rule out an *hMLH1* mutation, the remaining exons should be examined by DNA sequence analysis. If no mutation is found in the *hMLH1* DNA sequence, then one of the other mismatch repair genes may be defective. The most obvious candidate would be *hMSH2* because germline mutations in this gene have been identified in a high proportion of HNPCC families.

Samples from the HNPCC 3 and HNPCC 4 individuals were used as controls in the SSCP analysis because the mutations responsible for HNPCC were previously defined. These two mutations do not destroy or add a restriction enzyme recognition site, so restriction enzyme analysis is not a feasible low-cost screening method for

these families. However, now that these mutations have been detected by SSCP, other family members can be screened using this assay rather than DNA sequence analysis.

**Perspectives**

In order to fully characterize the spectrum of mutations in the Turkish population, more HNPCC families must be screened. Based on the results presented here, a mutation screening protocol for Turkish kindreds might include: (1) *Dde*I restriction enzyme analysis for the G884C mutation; (2) *Hph*I restriction enzyme analysis for the A1652C mutation; (3) protein truncation test for the C676T mutation and other nonsense mutations; (4) SSCP analysis using protocols 1, 2, and 7; and finally (5) DNA sequence analysis. Restriction enzyme analysis is efficient and inexpensive. SSCP is a valuable screening method but under optimal conditions sensitivity is only 80% (Cotton, 1997). DNA sequence analysis must be used to confirm a DNA mutation that is suggested by differential mobility with SSCP.

More sensitive mutation screening tests e.g. fluorescent assisted mismatch cleavage analysis (FAMA) need to be developed. In this technique, sense and antisense strands are simultaneously labeled with strand-specific fluorochromes during PCR amplification. Heteroduplexes are allowed to form between wild-type and mutant sequences. Mismatches are cleaved using chemicals or enzymes and the cleaved products are then analyzed and sized on the automated ABI 377 DNA Sequencer using GeneScan software. The main advantages of FAMA are its ability to detect 100% of mutations and to provide accurate positional information. Its ability to screen 1-2 kb lengths of DNA is a significant advantage over DNA sequencing.

Although the missense mutations so far discovered in Turkish HNPCC individuals appear to be pathogenic, a test that could assay the efficiency of mismatch repair by in wild-type and mutant proteins would be useful in determining whether a particular mutation was deleterious or merely a polymorphism. For this assay, colorectal cancer cell lines can be used as controls and +/+, +/- and -/- cell lines can be tested for their ability to repair DNA substrates containing different mismatches (Edelmann et al., 1997).

Additional studies to determine the downstream consequences of these mismatch repair gene mutations also need to be done. Tumors isolated from HNPCC individuals have been found to have mutations in mononucleotide tracts within several growth regulating genes (Markowitz et al., 1995; Souza et al., 1996; Rampino et al., 1997). Genetic analysis of tumors from these HNPCC individuals might yield further clues about other events required for the development of colorectal cancers.

While some of the studies mentioned above are long-term projects, an immediate effect of this study is the availability of relatively simple, cost-effective and reliable diagnostic tests that can be used to assess a number of germline mutations responsible for HNPCC in Turkish kindreds. Presymptomatic screening of germline mutations in DNA mismatch repair genes is clearly warranted for these family members. In addition, these family members must receive genetic counseling that is compassionate, empathetic and combined with appropriate education. Special care must be given to patients' culture, beliefs and traditions.

## REFERENCES


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. et al (1993) Clues to the pathogenesis of familial colorectal cancer. *Science* **260:** 812-816.

Aaltonen L.A., Sankila R., Mecklin J-P., Jarvinen H., Pukkala E., Peltomaki P., de la Chapelle A. (1994) A novel approach to estimate the proportion of hereditary non-polyposis colorectal cancer of total colorectal cancer burden. *Cancer Detection and Prevention* **18:** 57-63.

Aaltonen L.A., Salovaara R., Kristo P., Canzian F., Hemminki A., Peltomaki P., Chadwick R.B., Kaariainen H., Eskelinen M., Jarvinen H., Mecklin J-P., de la Chapelle A. (1998) Incidence of hereditary non-polyposis colorectal cancer and the feasibility of molecular screening for the disease. *N. Engl. J. Med.* **338:** 1481-1487.

Akiyama Y., Sato H., Yamada T., Nagasaki H., Tsuchiya A., Abe R., Yuassa Y. (1997) Germ-line mutation of the hMSH6/GTBP gene in an atypical hereditary non-polyposis colorectal cancer kindred. *Cancer Res.* **57:** 3920-3923.

Arnheim N., Shibata D. (1997) DNA mismatch repair in mammals: role in disease and meiosis. *Curr. Opin. Genet. Devel.*. **7:** 364-370.

Baker S.M., Plug A.W., Prolla T.A., Bronner C.E., Harris A.C., Yao X., Christie D.M., Monel L.C., Arnheim N., Bradley A., Ashley T., Liskay R.M. (1996) Involvement of mouse *Mlh1* in DNA mismatch repair and meiotic crossing over. *Nature Genet.* **13**: 336-343.

Beck N.E., Tomlinson I.P.M., Homfray T., Frayling I., Hodgson S.V., Harocopos C.J., Bodmer W.F. (1997a) Use of SSCP analysis to identify germline mutations in HNPCC families fulfilling the Amsterdam criteria. *Hum. Genet.* **99:** 219-224.

Beck N.E., Tomlinson I.P.M., Homfray T., Hodgson S.V., Harocopos C.J., Bodmer W.F. (1997b) Genetic testing is important in families with a history suggestive of hereditary non-polyposis colorectal cancer even if the Amsterdam criteria are not fulfilled. *Brit. J. Surg.* **84:** 233-237.

Bedell M.A., Largaespada D.A., Jenkins N.A., Copeland N.G. (1997) Mouse models of human disease. Part II: Recent progress and future directions. *Genes and Devel.* **11:** 11-43.

Bellacosa A., Genuardi M., Anti M., Viel A., Ponz de Leon M. (1996) Hereditary non-polyposis colorectal cancer: Review of clinical, molecular genetics, and counselling aspects. *Am. J. Med. Genet.* **62:** 353-364.

Boland C.R. (1998) Hereditary nonpolyposis colorectal cancer. In: Vogelstein B, Kinzler KW, editors. *The Genetic Basis of Human Cancer*. New York: McGraw-Hill, pp. 333-346.

Boyer J.C., Umar A., Risinger J.I., Lipford J.R., Kane M., Yin S., Barrett J.C., Kolodner R.D., Kunkel T.A. (1995) Microsatellite instability, mismatch repair deficiency and genetic defects in human cancer cell lines. *Cancer Res.* **55:** 6063-6070.

Bronner C.E., Baker S.M., Morrison P.T., Warren G., Smith L.G., Lescoe M.K., Kane M., Earabino C., Lipford J., Lindblom A., Tannergärd P., Bollag R.J., Godwin A.R., Ward D.C., Nordenskjold M., Fishel R., Kolodner R., Liskay R.M. (1994) Mutation in the DNA mismatch repair gene homolog *hMLH1* is associated with hereditary non-polyposis colon cancer. *Nature* **368:** 258-261.

Burke W., Petersen G., Lynch P., Botkin J., Daly M., Garber J., Kahn M.J.E., McTiernan A., Offit K., Thomson E., Varricchio C. (1997) Recommendations for follow-up care of individuals with an inherited predisposition to cancer. *JAMA*. **227:** 915-919.

Chen U., Chen S., Saha P., Dutta A. (1996) p21Cip1/Waf1 disrupts the recruitment of human Fen1 by proliferating-cell nuclear antigen into the DNA replication complex. *Proc. Natl. Acad. Sci.* **93:** 11597-11602.

Cotton R.G.H. (1993) Current methods of mutation detection. *Mutation Res,* **285:** 125-144 .

Cotton R.G.H. (1997) Scanning methods. In: *Mutation Detection*. New York: Oxford University Press, pp. 32-127.

Dean M. (1996) SSCP and heteroduplex analysis. In: Landegren U, editor. *Laboratory Protocols for Mutation Detection*. New York: Oxford University Press, pp. 23-26.

de Wind N., Dekker M., Berns A., Radman M., te Riele H. (1995) Inactivation of the mouse *Msh2* gene results in mismatch repair deficiensy, methylation tolerance, hyperrecombination, and predisposition to cancer. *Cell* **82**: 321-330.

de Wind N., Dekker M., Van Rossum A., Van der Valk M., te Riele H. (1998) Mouse models for hereditary non-polyposis colorectal cancer. *Cancer Res.* **58:** 248-255.

Dunlop M.G., Farrington S.M., Carothers A.D., Wyllie A.H., Sharp L., Burn J., Liu B., Kinzler K.W., Vogelstein B. (1997) Cancer risk associated with germline DNA mismatch repair gene mutations. *Hum. Mol. Genet.* **6:** 105-110.

Edelmann W., Cohen P.E., Kane M., Lau K., Morrow B., Bennett S., Umar A., Kunkel T., Cattoretti G., Chaganti R., Pollard J.W., Kolodner R.D., Kucherlapati R. (1996) Meiotic pachytene arrest in MLH1-deficient mice. *Cell* **85**: 1125-1134.

Eshleman J.R., Markowitz S.D. (1996) Mismatch repair defects in human carcinogenesis. *Hum. Mol. Genet.* **5:** 1489-1494.

Fink D., Nebel S., Aebi S., Zheng H., Kim H.K., Christen R.D., Howell S.B. (1997) Expression of the DNA mismatch repair proteins hMLH1 and hPMS2 in normal human tissues. *Brit. J. Cancer* **76:** 890-893.

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

Fishel R., Wilson T. (1997) *MutS* homologs in mammalian cells. *Curr. Opin. Genet. Devel.* **7:** 105-113.

Friedberg E.C., Walker G.C., Siede W. (1995) Mismatch repair. In: *DNA repair and mutagenesis*. Washington (DC): ASM Press, pp. 367-399.

Genuardi M., Anti M., Capozzi E., Leonardi F., Fornasarig M., Novella E., Bellacosa A., Valenti A., Gasbarrini G.B., Roncucci L., Benatti P., Percesepe A., Ponz de Leon M., Coco C., De Paoli A., Valentini M., Boiocchi M., Neri G., Viel A. (1998) *MLH1* and *MSH2* constitutional mutations in colorectal cancer families not meeting the standard criteria for hereditary non-polyposis colorectal cancer. *Int. J. Cancer*. **75:** 835-839.

Grompe M. (1993) The rapid detection of unknown mutations in nucleic acids. *Nature Genet.* **5:** 111-117.

Gryfe R., Swallow C., Bapat B., Redston M., Gallinger S., Couture J. (1997) Molecular biology of colorectal cancer. In: *Current Problems in Cancer*. St. Louis (MO): Mosby-Year Book, Inc., pp. 234-300.

Han H-J., Maruyama M., Baba S., Park J-G., Nakamura Y. (1995) Genomic structure of human mismatch repair gene, *hMLH1*, and its mutation analysis in patients with hereditary non-polyposis colorectal cancer (HNPCC). *Hum. Mol. Genet.* **4:** 237-242.

Harlow E. (1994) An introduction to the puzzle. *Cold Spring Harbor Symposia on Quantitative Biology* **LIX:** 709-723.

Heinimann K., Müller H., Weber W., Scott R.J. (1997) Disease expression in Swiss hereditary non-polyposis colorectal cancer (HNPCC) kindreds. *Int. J. Cancer.* **74:** 281-285.

Hemminki A., Peltomaki P., Mecklin J-P., Jarvinen H., Salovaara R., Nystrom-Lahti M., de la Chapelle A., Aaltonen L.A. (1994) Loss of the wild-type MLLH1 gene is a feature of hereditary nonpolyposis colorectal cancer. *Nature Genet.* **8**: 405-410.

Herfarth K.K-F., Kodner I.J., Whelan A.J., Ivanovich J.L., Bracamontes J.R., Wells S.A., Goodfellow P.J. (1997) Mutations in MLH1 are more frequent than in MSH2 in sporadic colorectal cancers with microsatellite instability. *Genes Chromosomes and Cancer* **18:** 42-49.

Hutter P., Couturier A., Scott R.J., Alday P., Delozier-Blanchet C., Cachat F., Antonarakis S.E., Joris F., Gaudin M., D'Amato L., Buerstedde J-M. (1996) Complex genetic predisposition to cancer in an extended HNPCC family with an ancestral *hMLH1* mutation. *J. Med. Genet.* **33:** 636-640.

Ionov Y., Peinado M.A., Malkhosyan S., Shibata D., Perucho M. (1993) Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* **363:** 558-561.

Jager A.C., Bisgaard M.L., Myrhoj T., Bernstein I., Rehfeld J., Nielsen F.C. (1997) Reduced frequency of extracolonic cancers in hereditary non-polyposis colorectal families with monoallelic *hMLH1* expression. *Am. J. Hum. Genet.* **61:** 129-138.

Jiricny J. (1994) Colon cancer and DNA repair: Have mismatches met their match? *TIG* **10:** 164-168.

Kinzler K.W., Vogelstein B. (1996) Lessons from hereditary colorectal cancer. *Cell* **87:** 159-170.

Kinzler K.W., Vogelstein B. (1997) Gatekeepers and caretakers. *Nature* **386:** 761-763.

Kolodner R. (1996) Biochemistry and genetics of eukaryotic mismatch repair. *Genes and Devel.* **10:** 1433-1442.

Kolodner R.D. (1995) Mismatch repair: Mechanisms and relationship to cancer susceptibility. *TIBS.* **20:** 397-401.

Kolodner R.D., Hall N.R., Lipford J., Kane M.F., Rao M.R.S., Morrison P., Wirth L., Finan P.J., Burn J., Chapman P., Earabino C., Merchant E., Bishop D.T., Garber J., Bronner C.E., Baker S.M., Warren G., Smith L.G., Lindblom A., Tannegard P., Bollag R.J., Godwin A.R., Ward D.C., Nordenskjold M., Liskay R.M., Copeland N., Jenkins N., Lescoe M.K., Ewel A., Lee S., Griffith J., Fishel R. (1994) Human mismatch repair genes and their association with

hereditary non-polyposis colon cancer. *Cold Spring Harbor Symposia on Quantitative Biology* **LIX:** 331-337.

Lahue R.S., Au K.G., Modrich P. (1989) DNA mismatch correction in a defined system. *Science* **245:** 160-164.

Leach F.S., Nicolaides N.C., Papadopoulos N., Liu B., Jen J., Parsons R., Peltomaki P., Sistonen P., Aaltonen L.A., Nystrom-Lahti M., Yu J-W., Kao F-T., Chen D.J., Cerosaletti K.M., Fournier R.E.K., Todd S., Lewis T., Leach R.J., Naylor S.L., Weissenbach J., Mecklin J-P., Jarvinen H., Petersen G.M., Hamilton S.R., Green J., Jass J., Watson P., Lynch H.T., Trent J.M., de la Chapelle A., Kinzler K.W., Vogelstein B. (1993) Mutations of a *mutS* homolog in hereditary non-polyposis colorectal cancer. *Cell* **75:** 1215-1225.

Leach F.S., Polyak K., Burrell M., Johnson K.A., Hill D., Dunlop M.G., Wyllie A.H., Peltomaki P., de la Chapelle A., Hamilton S.R., Kinzler K.W., Vogelstein B. (1996) Expression of the human mismatch repair gene hMSH2 kn normal and meoplastic tissues. *Cancer Res.* **56**: 235-240.

Levinson G., Gutman A. (1987) High frequencies of short frameshifts in poly-CA/TG tandem repeats borne by bacteriophage M13 in Escherishia coli K-12. *Nucleic Acid Res.* **15:** 5323-5338.

Lindahl T., Karran P., Wood R.D. (1997) DNA excision repair pathways. *Curr. Opin. Genet. Devel.* **7:** 158-159.

Lindblom A., Tannergärd P., Werelius B., Nordenskjold M. (1993) Genetic mapping of a second locus predisposing to hereditary non-polyposis colon cancer. *Nature Genet.* **5:** 279-282.

Liu B., Parsons R.E., Hamilton S.R., Petersen G.M., Lynch H.T., Watson P., Markowitz S., Willson J.K.V., Green J., de la Chapelle A., Kinzler K.W., Vogelstein B. (1994) hMSH2 mutations in hereditary non-polyposis colorectal cancer kindreds. *Cancer Res.* **54:** 4590-4594.

Liu B., Parsons R., Papadopoulos N., Nicolaides N.C., Lynch H.T., Watson P., Jass J.R., Dunlop M., Wyllie A., Peltomaki P., de la Chapelle A., Hamilton S.R., Vogelstein B., Kinzler K.W. (1996) Analysis of mismatch repair genes in hereditary non-polyposis colorectal cancer patients. *Nature Med.* **2:** 169-174.

Liu T., Wahlberg S., Rubio C., Holmerg E., Grönberg H., Lindblom A. (1998) DGGE screening of mutations in mismatch repair genes (*hMSH2* and *hMLH1*) in 34 Swedish families with colorectal cancer. *Clin. Genet.* **53:** 131-135.

Loeb L.A. (1991) Mutator phenotype maybe required for multistage carcinogenesis. *Cancer Res.* **5**: 3075-3079.

Lynch H.T., Drauhard T., Vasen H.F.A., Cavalieri J., Lynch J., Nord S., Smyrk T., Lanspa S., Murphy H., Whelan K.L., Peters J., de la Chapelle A. (1996) Genetic counseling in Navajo hereditary non-polyposis colorecta cancer kindred. *Cancer* **77:** 30-35.

Lynch H.T., Smyrk T., Lynch J.F. (1996) Overview of natural history, pathology, molecular genetics and management of HNPCC (Lynch syndrome). *Int. J. Cancer.* **69:** 38-43.

Lynch H.T., Smyrk T., Lynch J. (1997) An update of HNPCC (Lynch syndrome). *Cancer Genet. Cytogenet.* **93:** 84-99.

Lynch H.T., Smyrk T., Lynch J., Fitzgibbons R., Lanspa S., McGinn T. (1995) Update on the differential diagnosis, surveillance and management of hereditary non-polyposis colorectal cancer. *Eur. J. Cancer* **31A:** 1039-1046.

Lynch H.T., Smyrk T.C., Watson P., Lanspa S.J., Lynch J.F., Lynch P.M., Cavalieri J., Boland R. (1993) Genetics, natural history, tumor spectrum, and pathology of hereditary non-polyposis colorectal cancer: An updated review. *Gasroenterology* **104:** 1535-1549.

Maliaka Y.K., Chudina A.P., Belev N.F., Alday P., Bochkov N.P., Buerstedde J-M. (1996) CpG dinucleotides in the hMSH2 and hMLH1 genes are hotspots for HNPCC mutations. *Hum. Genet.* **97:** 251-255.

Markowitz S., Wang J., Myeroff L., Parsons R., Sun L., Lutterbaugh J., Fan R.S., Zborowska E., Kinzler K., Vogelstein B., Brattain M., Willson J.K.V. (1995) Inactivation of the type II TGF-β receptor in colon cancer cells with microsatellite instability. *Science* **268**: 1336-1338.

Marra G., Boland C.R. (1995) Hereditary non-polyposis colorectal cancer: the syndrome, the genes, and historical perspectives. *J. Natl. Cancer Inst.* **87:** 1114-1125.

Moslein G., Tester D.J., Noralane M.L., Honchel R., Cunningham J.M., French A.J., Halling K.C., Schwab M., Goretzki P., Thibodeau S.N. (1996) Microsatellite instability and mutation analysis of *hMSH2* and *hMLH1* in patients with sporadic, familial and hereditary colorectal cancer. *Hum. Mol. Genet.* **5:** 1245-1252.

Muta H., Noguchi M., Perucho M., Ushio K., Sugihara K., Ochiai A., Nawata H., Hirohashi S. (1996) Clinical implications of microsatellite instability in colorectal cancers. *Cancer* **77:** 265-270.

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., Adams M.D., Venter J.C., Dunlop M.G., Hamilton S.R., Petersen G.M., de al Chapelle A., Vogelstein B., Kinzler K.W. (1994) Mutations of two PMS homologues in hereditary non-polyposis colon cancer. *Nature* **371:** 75-80.

Nystrom-Lahti M., Kristo P., Nicolaides N.C., Chang S.Y., Aaltonen L.A., Moisio A.L., Jarvinen H.J., Mecklin J.P., Kinzler K.W., Vogelstein B., de la Chapelle A., Peltomaki P. (1995) Founding mutations and Alu-mediated recombination in hereditary colon cancer. *Nature Med.* **11**: 1203-1206.

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

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., Venter J.C., Hamilton S.R., Petersen G.M., Watson P., Lynch H.T., Peltomaki P., Mecklin J-P., de la Chapelle A., Kinzler K.W., Vogelstein B. (1994) Mutation of a mutL homolog in hereditary cancer. *Science* **263:** 1625-1629.

Parsons R., Li G-M., Longley M.J., Fang W., Papadopoulos M., Jen J., de al Chapelle A., Kinzler K.W., Vogelstein B., Modrich P. (1993) Hypermutability and mismatch repair deficiency in $RER^+$ tumor cells. *Cell* **75**: 1227-1236.

Peltomaki P., Aaltonen L.A., Sistonen P., Pylkkanen L., Mecklin J.P., Jarvinen H., Green J.S., Jass J.R., Weber J.L., Leach F.S., Petersen G.M., Hamilton S.R., de la Chapelle A., Vogelstein B. (1993a) Genetic mapping of a locus predisposing to human colorectal cancer. *Science* **260:** 810-812.

Peltomaki P., Lothe R.A., Aaltonen L.A., Pylkkanen D., Nystrom-Lahti M., Seruca R., David L., Holm R., Ryberg D., Haugen A. et al (1993b) Microsatellite instability is associated with tumors characterize the hereditary non-polyposis colorectal carcinoma syndrome. *Cancer Res.* **53:** 5853-5855.

Petersen G.M. (1995) Genetic epidemiology of colorectal cancer. *Eur. J. Cancer* **31A:** 1047-1050.

Ponz de Leon M., Sassatelli R., Benatti P., Roncucci R. (1993) Identification of hereditary non-polyposis colorectal cancer in the general population. *Cancer* **71:** 3493-3501.

Prolla T.A., Baker S.M., Harris A.C., Tsao J.L., Yao X., Bronner C.E., Zheng B., Gordon M., Reneker J., Arnheim N., Shibata D., Bradley A., Liskay R.M. (1998) Tumor susceptibility and spontaneous mutation in mice deficient in Mlh1, Pms1 and Pms2 DNA mismatch repair. *Nature Genet.* **18**: 276-279.

Rampino N., Yamamoto H., Ionov Y., Li Y., Sawai H., Reed J.C., Perucho M. (1997) Somatic frameshift mutations in the BAX gene in colon cancers of the microsatellite mutator phenotype. *Science* **275**: 967-969.

Reitmair A.H., Schmits R., Ewel A., Bapat B., Redston M., Mitri A., Waterhouse P., Mittrucker H.W., Wakeham A., Liu B., Thomason A., Griesser H., Gallinger S., Ballhausen W.G., Fishel R., Mak T.W. (1995) MSH2-deficient mice are viable and susceptible to lymphoid tumors. *Nature Genet.* **11**: 64-70.

Reitmair A.H., Redston M., Cai J.C., Chuang T.C., Bjerknes M., Cheng H., Hay K., Gallinger S., Bapat B., Mak T.W. (1996) Spontaneous intestinal carcinomas and skin neoplasms in Msh2-deficient mice. *Cancer Res.* **56**: 3842-3849.

Rhyu M.S. (1996) Molecular mechanisms underlying hereditary non-polyposis colorectal carcinoma. *J. Natl. Cancer Inst.* **88:** 240-250.

Scott R.J., Muller H. (1993) Familial and genetic aspects of colorectal carcinogenesis. *Eur. J. Cancer* **29A:** 2163-2167.

Souza R.F., Appel R., Yin J., Wang S., Smolinski K.N., Abraham J.M., Zou T., Shi Y., Lei J., Cottrell J., Cymes K., Biden K., Simms L., Leggett B., Lynch P.M., Frazier M., Powell S.M., Harpaz N., Sugimura H., Young J., Meltzer S.J. (1996) Microsatellite instability in the insulin-like growth factor II receptor gene in gastrointestinal tumors. *Nature Genet.* **14**: 255-257.

Strand M., Prolla T.A., Liskay R.M., Petes T.D. (1993) Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. *Nature* **365:** 274-276.

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

Thibodeau S.N., Bren G., Schaid D. (1993) Microsatellite instability in cancer of the proximal colon. *Science* **260:** 816-819.

Toribara N.W., Sleisenger M.H. (1995) Screening for colorectal cancer. *N. Engl. J. Med.* **332:** 861-867.

Umar A., Buermeyer A.B., Simon J.A., Thomas D.C., Clark A.B., Liskay R.M., Kunkel T.A. (1996) Requirement for PCNA in DNA mismatch repair at a step preceding DNA resynthesis. *Cell* **87:** 65-73.

Vogelstein B., Kinzler K.W. (1994) Colorectal cancer and the intersection between basic and clinical research. *Cold Spring Harbor Symposia on Quantitative Biology* **LIX:** 517-520.

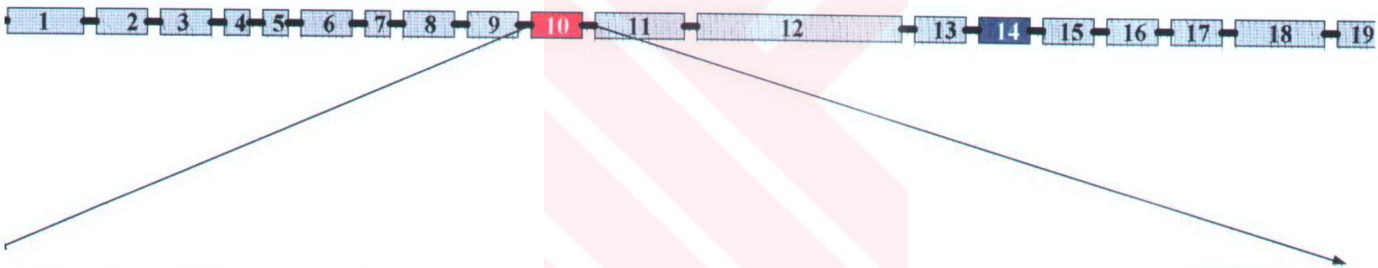
Wang Q., Desseigne F., Lasset C., Saurin J-C., Navarro C., Yağcı T., Keser İ., Bağcı H., Lüleci G., Gelen T., Chayvialle J-A., Puisieux A., Öztürk M. (1997) Germline *hMSH2* and *hMLH1* gene mutations in incomplete HNPCC families. *Int. J. Cancer.* **73:** 831-836.

Wijnen J., Khan P.M., Vasen H., Van der Klift H., Mulder A., Van Leeuwen-Cornelisse I., Bakker B., Losekoot M., Moller P., Fodde R. (1997) Hereditary non-polyposis colorectal cancer families not complying with the Amsterdam criteria show extremely low frequency of mismatch-repair-gene mutations. *Am. J. Hum. Genet.* **61:** 329-335.

Wijnen J., Khan P.M., Vasen H., Menko F., Van der Klift H., Van den Broek M., Van Leeuwen-Cornelisse I., Nagengast F., Meijers-Heijboer E.J., Lindhout D., Griffioen G., Cats A., Kleibeuker J., Varesco L., Bertario L., Bisgaard M-L., Mohr J., Kolodner R., Fodde R. (1996) Majority of *hMLH1* mutations responsible for hereditary non-polyposis colorectal cancer cluster at the exonic region 15-16. *Am. J. Hum. Genet.* **58:** 300-307.

## Human DNA Mismatch Repair Protein Homolog (*hMLH1*) Gene , Exon 10

APPENDIX A.




tagtgggctggaaagtggccacaggtaaaggtgcacctttcttcctggggatgtgatgtgcatatcactacagaaatgtctttcctgag gtgatgtcat gactttgtg tgaatgtacac

tgtgacctcacc cctcaggacagttttga actggttgctttctttttat tgttt ag ATC GTCTGGTAGA ATCAACTTCC TTGAGAAAAGCCATAGA

AC AGTGTATGCA GCCTATTTGCCCAAAAACAC ACACCCATTC CTGTACCTCAG gt aatgtagcaccaaactcctcaaccaag

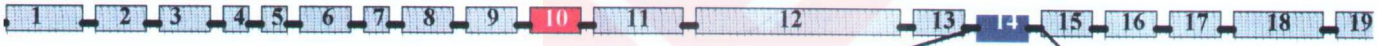
tcacaaggaa cagatgtt ctatcaggct ctcctc tttg aaagagatgagcatgctaatagtacaatcagagtgaatcccatacaccactggcaaaaggatgttctgtcccttct

caagtacaatcagagtgaatcccatacaccactggcaaaaggatgttctgtcccttcttaca

⇨ Known mutaion: G ⟶ C transversion ( at codon 295 : Ser [AGT] ⟶ Thr [ACT] )

## Human DNA Mismatch Repair Protein Homolog (*hMLH1*) Gene , Exon 14

APPENDIX B.




cagttttcaccaggaggctcaaatcaggcnnctttgcttat tggtgtct ctagttctgg tgcctggtgctttggtcaatgaagtggggttggtaggatt ctattacttacctgtttttt

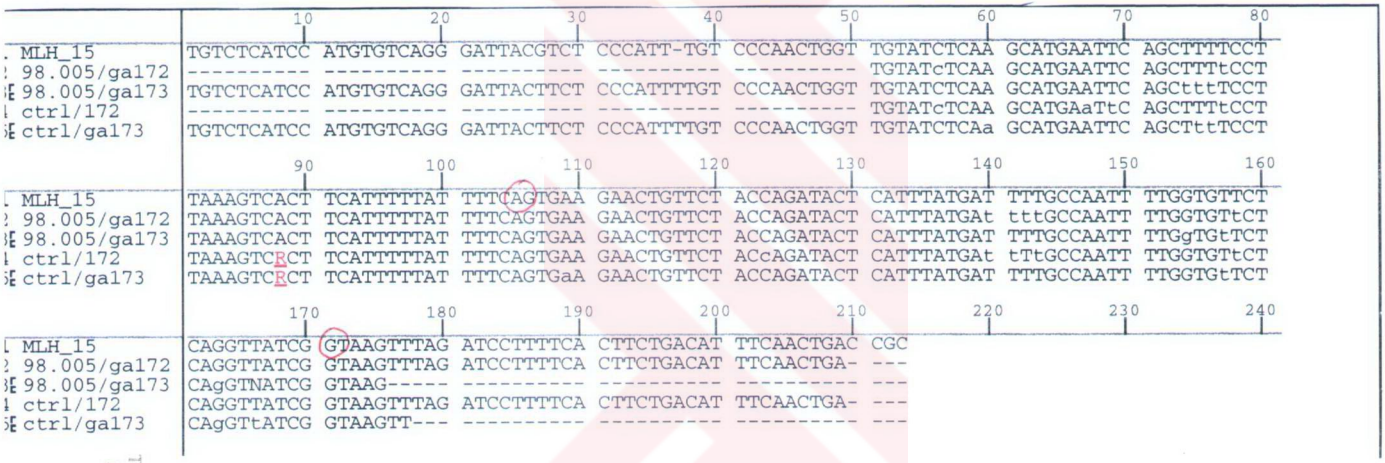
ggttttatttttgttttgc ag TTCTCCGG GAGATGTTGC ATAACCACTC CTTCGTGGGC TGTGTGAATC CTCAGTGGGC

CTTGGCACAGCATCAAACCA AGTTATACCT TCTCA A CACCACCAAGCTTAG gt aaatcagctgagtgtgtgaacaa gcag

agctactaca acaatg gtccagggagcacaggcacaaaagctaaggagag cagcatgaag gtagttgggagggcacaggctttggagtc agcacatgt

⇨ Known mutaion: A ⟶ C transversion ( at codon 551 : Asn [AAC] ⟶ Thr [ACC] )

## APPENDIX C.



## APPENDIX D.


1 MLH_07
2 98.005/289
4
3 Ctrl/GA289
5
GTGTGTGTTT TTGGCAACTC TTTTCTTACT CTTTTGTTTT TCTTTTCCAG GTATTCAGTA CACAATGCAG GCATTAGTTT
---------- ---------- ---------- ---------- ------CCAG GTATTCAsTA CACMATGCAG GCATtAGYTT
DDDDDDDDDD DDDDDDDDDD DDDDDDDDDD DDDDDDDDDD DDDDD---- -------*-- ---*------ ----*--*--
---------- ---------- ---------- ---------- ---------- ----TCASTA CACMaTGCAG GCATTAGYTT
DDDDDDDDDD DDDDDDDDDD DDDDDDDDDD DDDDDDDDDD DDDDDDDDDD DDDD---*-- ---**----- -------*--
CTCAGTTAAA AAAGTAAGTT CTTGGTTTAT GGGGGATGGT TTTGTTTTAT GAAAAGAAAA AAGGGGATTT TTAATAGTTT
CTCAGTTAAA AAAGTAAGTT CTTGGTTTAT GGGGGATGGT TTTGTYTTAT GAAAAGAAAA AAGGGGATTT TTAATAGTTT
---------- ---------- ---------- ---------- -----*---- ---------- ---------- ----------
CTCAGTTAAA AAAGTAAGTt CTTGGYTTAT GGGGGATSGT tTTGTYTTAT GAaAAGAaAn AAGGGGATtT tTAATAGTTT
---------- ---------*  -----*---- --------*-- *----*----- --*----*-* ---------*- *---------
GCTGGTGGAG ATAAGG
GCTGGTGGAG ATA---
---------- ---DDD
GCTGGTGG-- -------
--------DD DDDDDD-