Functional analysis and carrier detection of mismatch repair gene mutations

Jianghua Ou
Functional analysis and carrier detection of mismatch repair gene mutations

The studies described in this thesis were supported by: the Dutch Cancer Society (grant No. RUG2002-2678), the European Community (FP6-2004-LIFESCIHEALTH-5, proposal No 018754)

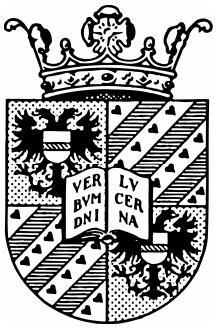
The printing of this thesis was supported by the Dutch Cancer Society (KWF), University of Groningen, the graduate school for drug exploration GUIDE, Mutagene X (Piscataway, NJ 08855 USA)

Cover design: Jianghua Ou and GrafiMedia

Printed by: GrafiMedia, Facilitair Bedrijf RuG, Groningen

RIJKSUNIVERSITEIT GRONINGEN

Functional analysis and carrier detection of mismatch repair gene mutations



Proefschrift

ter verkrijging van het doctoraat in de
Medische Wetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. F. Zwarts,
in het openbaar te verdedigen op
woensdag 27 februari 2008
om 16.15 uur

door

Jianghua Ou

geboren op 28 januari 1969
te XinJiang, China

Promotores:	Prof. dr. R.M.W. Hofstra
	Prof. dr. J.H. Kleibeuker
Copromotor:	Dr. R.H. Sijmons

Beoordelingscommissie:	Prof. dr. L.J. Rasmussen
	Prof. dr. H. Morreau
	Prof. dr. C.H.C.M. Buys

In silence, in steadiness, in severe abstraction,
let him hold by himself, add observation to observation,
patient of neglect, patient of reproach,
and bide his own time, happy enough if he can satisfy himself alone,
that this day he has seen something truly
(Ralph Waldo Emerson, American thinker).

Paranimfen: Lei Ke
Paul Jager

CONTENT

Chapter 1

Introduction

Section 1 Lynch syndrome

1.1 Lynch syndrome: introduction

Fundamentally cancer is a genetic disorder, meaning that a cell needs genetic changes to confer a selective growth advantage before becoming a cancer cell. Such genetic alterations mostly occur somatically, however, mutations through the germline might result in a high cancer risk and predispose to early-onset cancer (1). Colorectal cancer (CRC) is one of the most prevalent cancer types in the Western world, with Lynch syndrome, also known as hereditary nonpolyposis colorectal cancer (HNPCC), as the most common type of hereditary CRC.

Lynch syndrome was first described in 1913 by Alfred Warthin, a pathologist who described the family of his seamstress (2). That it was an inherited disease, with an autosomal dominant pattern of inheritance, was rediscovered in the mid-1960s by Henry Lynch (3). He also noticed that in addition to colon cancer, extracolonic malignancies occurred. As Lynch did recognise the syndrome, it was and is still called Lynch syndrome. In 1991, the international collaborative group on hereditary nonpolyposis colorectal cancer (ICG-HNPCC) formulated the Amsterdam criteria to clinically define the disease and to make *research* on this subject better comparable (4). The families meeting some but not all of these Amsterdam (II) criteria often are classified as “suspected for Lynch syndrome” or, more briefly, “Lynch syndrome suspected”. In particular for genetic screening of these suspected patients additional selection criteria have been developed as this group of patients includes many patients that can “genetically” be classified as Lynch syndrome (*e.g.* have a germline pathogenic mutation in one of the MMR genes). Berends *et al*. (5) reviewed the attempts of various groups to define new criteria to select patients for mutation analysis.

Estimations of the proportion of clinical cases with Lynch syndrome of the total of colorectal cancers, based on the Amsterdam criteria, varied from less than 1% in Finland to 13% in the United Kingdom (6;7).

1.2 Genetic predispositions causing Lynch syndrome

As mentioned, a cell needs multiple genetic alterations before it becomes a cancer cell. It has been proposed that cells need to acquire a “mutator phenotype” at an early stage of tumourigenesis to enable the accumulation of multiple genetic aberrations (8). This hypothesis was proven for Lynch syndrome in the early nineties. It was the finding of numerous somatic mutations in the tumours of patients with Lynch syndrome in combination

with linkage analysis that resulted in the identification of inherited heterozygous defects in two genes encoding proteins that function in the maintenance of genomic stability, the DNA mismatch repair (MMR) genes *MLH1* and *MSH2*. Later other MMR genes, *PMS2* and *MSH6*, were identified as less frequently involved genes for this cancer syndrome (as reviewed in (9;10)).
Recently, several population-based studies concluded that genetically proven Lynch syndrome, c.q. families having mutations in *MLH1* and *MSH2*, accounts for 0.3% to 2.7% of all colorectal cancer cases (11;12).

1.3 Mismatch repair system

The MMR system comprises an evolutionary conserved set of proteins that recognises and removes replication errors that are not immediately corrected by the proofreading activities of replicative DNA polymerases, thereby providing a secondary system of proofreading. This system corrects not only single base-pair mismatches but also small insertion/deletion loops (IDLs) that result from slippage of the replicative polymerases while replicating microsatellites consisting of mono- or di-nucleotide repeats. The ability of many species to repair mismatches in double-stranded DNA has been well documented. The first critical step in this process is the recognition of the mismatched DNA. In the best documented mismatch repair system, the one in *Escherichia coli*, the MutS protein accomplishes this. A homodimer of the MutS protein binds to the site of a mismatch in double-stranded DNA and, with the cooperation of the MutL and the MutH proteins, it targets a section around the mutated DNA strand before removal. Other proteins complete the repair process: the section of DNA that is targeted is removed and degraded, the remaining single stranded DNA gap is filled in using the complementary strand as a template, followed by ligation resulting in a repaired section of double-stranded DNA without mismatches.
In the human MMR system the mismatch recognition function is fulfilled by a heterodimeric protein complex composed of two MutS Homologous (MSH) proteins. The major MSH dimer consists of MSH2 and MSH6, also called MutSα. This complex is able to bind to base-base mismatches and IDLs. A minor and partially redundant mismatch binding heterodimer (MUTSβ) consists of the MSH2 and MSH3 proteins and this complex binds to larger IDLs (1). In the mismatch repair process MutS heterodimers recognise mismatches in the DNA. Binding to these mismatches induces a conformational change for the MutS heterodimer, which allows the uptake of ATP and the signalling towards the MutL heterodimers, and thus subsequently initiating a downstream repair event.
Heterodimers composed of human MutL Homologous (MLH) proteins, MLH1, PMS2 and MLH3, interact with the MSH2-containing DNA-protein complex and coordinate downstream repair events. These downstream

events involve different proteins such as exonucleases (*e.g.* EXO1) and DNA polymerases (13;14). The heterodimer MLH1-PMS2 (MUTLα) interacts with both MSH2-MSH6 and MSH2-MSH3, while the heterodimer MLH1-MLH3 (MUTLγ), appears restricted to interactions with only MSH2-MSH3 (15;16).
Another human MutL homologue, PMS1, has also been suggested to function in mismatch repair but a biochemical role for this protein in mismatch repair has not been demonstrated (17). Lipkin *et al* (15) have suggested that PMS1 as well as PMS2 and MLH3 may exhibit functional redundancy.
In addition to its indispensable role in removing misincorporations, the repair proteins are involved in or have an influence on other processes, such as apoptosis and proliferation (18;19). MMR is also involved in determining cellular sensitivity to a number of DNA damaging agents. Some of these drugs are used as chemotherapeutics, *e.g.* the S_N1 DNA methylating chemotherapeutics procarbazine, dacarbazine and temozolomide, the cross-linking chemotherapeutic cisplatin, and the immunosuppressant azathioprine (for reviews on this subject see (20-22)).

1.4 Lynch syndrome and MMR

As mentioned, the majority (>90%) of germline mutations found in Lynch syndrome patients affect one of two MMR genes, *MSH2* and *MLH1* (9). Five to ten percent of Lynch syndrome families carry a germline mutation in the *MSH6* gene (23). *PMS2* mutations were believed to be very rare in Lynch syndrome (24), although recently the number of reported mutations is rising (25).
Two other MMR genes have been screened for mutations, namely *MLH3* and *EXO1* (26-28). Germline mutations were indeed found in these genes, but at a low frequency and mostly representing missense mutations that cannot be classified directly as pathogenic. These genes are therefore not believed to be key players in Lynch syndrome (29-33). We postulate that mutations in these genes might modify the penetrance of the disease rather than being causative. There have been no reports on germline mutations of *MSH3* in Lynch syndrome (9).

Section 2 Identifying patients with MMR gene mutations

2.1 Identifying patients with MMR gene mutations: introduction

Identifying MMR mutation carriers is important as it opens avenues for families to undergo presymptomatic testing. Identifying MMR mutations greatly facilitates carriers as they can enrol in cancer surveillance programs, making it possible to detect tumours in early stages and offering the possibility of presymptomatic diagnosis in Lynch syndrome families.

Equally important, non-mutation carriers can be dismissed from these screening programs. Mutation detection is, however, hampered by the facts that the disease is heterogeneous, mutations are spread all over the genes, hardly any founder mutation is identified, and the fact that mutation detection is costly and labour-intensive. Rapid and accurate mutation detection in these genes is a technical challenge. So far, there is no single technique allowing the detection of all types of mutations that have been identified in Lynch syndrome families. Rather than testing all patients with colorectal cancer (or any other type of cancer that could be Lynch syndrome associated), criteria have been introduced to select patients that have relatively high chances of being MMR gene mutation carriers for further molecular analysis. Clinical as well as non-clinical criteria can be used for this preselection.

2.2 Clinical criteria

In 1991, the International Collaborative Group on HNPCC (ICG-HNPCC) proposed a set of criteria primarily for research purpose to clinically define Lynch syndrome and to make Lynch syndrome-related research more comparable (4) These criteria, also called the Amsterdam I criteria, define Lynch syndrome as follows: occurrence of histologically CRC in at least three relatives (one of whom is a fist-degree relative of the other two) in at least two successive generations, and in one of the relatives with colorectal cancer age at diagnosis is less than 50 years. Familial adenomatous polyposis should be excluded. The Amsterdam I criteria have contributed enormously to uniformity in the clinical diagnosis of Lynch syndrome. However, it should be noted, that the Amsterdam I criteria fail to acknowledge the contribution of extracolonic cancers and lead to an underdiagnosis of the syndrome. So in 1999, the Amsterdam II criteria were formulated (34), which also take in account this extracolonic cancers.

Again it is argued that also the Amsterdam II criteria are too stringent, as a significant proportion of families carrying MMR germline mutations do not fulfil these criteria and will be excluded from cancer surveillance programs when the Amsterdam criteria for selection for mutation analysis are applied (35;36).

To overcome this disadvantage of the Amsterdam criteria, several groups have proposed less restrictive clinical criteria, including the revised Bethesda criteria, the Japanese criteria, and the Mount Sinai criteria (9;37-40). These criteria are dependent on both family history and clinical pathological characteristics of the tumour and on the association with extracolonic cancers. Of these the Bethesda criteria the best validated and thus are most used. Table 1 lists these criteria. Patients (or tumours from patients) meeting any of the revised Bethesda criteria, as well as tumours from family members, should be tested for microsatellite instability (MSI) and when proven positive, tested for germline mutations in the MMR genes.

Table 1. Revised Bethesda Guidelines for selecting Lynch syndrome related tumours that should be tested for MSI (for MSI see chapter 2.3).

Tumours from individuals should be tested for MSI in the following situations:
1. Colorectal cancer diagnosed in a patient who is less than 50 years of age.
2. Presence of synchronous or metachronous colorectal or other HNPCC-associated tumours,* regardless of age.
3. Colorectal cancer with the MSI-H histology‡ diagnosed in a patient who is less than 60 years of age.
4. Colorectal cancer and one or more first-degree relatives with an HNPCC-related tumour, with one of the cancers being diagnosed under age 50 years.
5. Colorectal cancer diagnosed in two or more first- or second-degree relatives with HNPCC-related tumours, regardless of age.

*Hereditary nonpolyposis colorectal cancer (HNPCC)-related tumours include colorectal, endometrial, stomach, ovarian, pancreas, ureter and renal pelvis, biliary tract, and brain (usually glioblastoma as seen in Turcot syndrome) tumours, sebaceous gland adenomas and carcinomas, keratoacanthomas in Muir–Torre syndrome, and carcinoma of the small bowel (48). ‡Presence of tumour infiltrating lymphocytes, Crohn's-like lymphocytic reaction, mucinous/signet-ring differentiation, or medullary growth pattern.

2.3 Microsatellite instability (MSI) analysis:

MMR deficiency results in the accumulation of unrepaired mutations. Unrepaired mutations are easily detected by looking at repetitive sequences, the so-called microsatellites, as these sequences increased or decreased in length. This phenomenon is referred to as microsatellite instability (MSI). As determining MSI is simple and because the large majority (85-90%) of tumours from Lynch syndrome families show MSI (called MSI-high or MSI-H), MSI testing has been put forward as the selection criterion for mutation analysis.

As mentioned, the MSI-H phenotype is caused by the loss of MMR function. This lack of MMR function is caused by the inactivation of both alleles of the same MMR gene. In Lynch syndrome patients one allele is inactivated by a germline event, the second, wild type, allele is inactivated in the tumour by a somatic event resulting in physical or functional loss. In sporadic colorectal cancer patients two somatic events have to take place. In most sporadic cases the mechanism of bi-allelic inactivation is *MLH1*

promoter hypermethylation (41). As the overall percentage of Lynch syndrome tumours among all MMR deficient tumours is quite small, it is still under debate whether population-based MSI screening in all colorectal carcinomas is effective.

2.4 Immunohistochemical staining (IHC) of MMR proteins

The examination of tumour tissues for the loss of expression of a MMR protein by immunohistochemical staining (IHC) is another method to prescreen tumours for a MMR deficient phenotype. Tumour tissues that demonstrate an absence of nuclear staining in the presence of a positive staining in the normal cells in the same section are presumed to have bi-allelic loss (mutations) of the corresponding MMR gene. IHC analysis is now routinely applied for the detection of MSH2, MLH1, MSH6 and PMS2. Interestingly, loss of only MSH6 and loss of only PMS2 can be observed whereas loss of MSH2 is always accompanied by subsequent loss of MSH6 and similarly loss of MLH1 is always accompanied by loss of PMS2. It is widely accepted that this concomitant loss of MMR proteins is related to the stability of the proteins, a stability that is depending on heterodimer formation (42;43). Loss of MSH2 alone or loss of MSH6 alone or loss of PMS2 alone is an indication for the presence of a germline mutation in the respective gene. Loss of MLH1 in IHC can be caused either by a *MLH1* germline mutation plus a second somatic *MLH1* mutation on the other allele or rarely by germline hypermethylation of one allele plus somatic hypermethylation of the second allele (in the tumour) (44) or by biallelic somatic *MLH1* promoter hypermethylation in the tumour, as seen in most sporadic cases.

According to recent large studies (45-49), IHC is at least equally valuable as MSI analysis for detecting potential germline mismatch repair gene mutation carriers. Hayashi and colleagues showed that the absence of expression of MLH1 or MSH2 had 100 percent specificity for predicting a MSI-H phenotype (42/42). The sensitivity of IHC for detecting a MSI-H tumour was 93.3 percent (42/45). The predictive value of normal expression of both of these proteins for predicting the MSS/MSI-L status was 98.8 percent. (50).

2.5 Identification of germline line mutation in a non-tumour tissue based way

As mentioned, tumour tissues from Lynch syndrome mostly show MSI-H and negative immunohistochemical staining(s) due to a loss of a MMR protein. When tumour tissues are not available, selection based on these parameters can not be performed. Development of a pre-screening method based on non-tumour tissues would therefore in these cases be very useful. Fields and co-worker (51) describe such a test. They tested Epstein Bar Virus (EBV) transformed lymphoblastoid cell lines from MMR gene

mutation carriers on heterozygous MMR protein loss.This strategy is based on the hypothesis that individuals harbouring a truncated germline mutation will show a 50% reduction of the mutated MMR protein. It is this reduction that can be quantified by immunoblotting leading to the identification of MMR germline mutation carriers. The Fields method makes use of EBV-transformed lymphoblastoid cell lines. Making such cell lines is costly and labour and time intensive. Whether this approach can be used on lymphoblasts directly, is yet unclear.

Section 3 Pathogenicity of UVs

3.1 Assessing pathogenicity of mutations in MMR –genes

Finding a DNA variant is one, deciding whether this variant is pathogenic, in other words, is contributing to the disease phenotype, is something else. The decision is easily made for mutations that result in premature termination of translation and thus in loss of function. For mutations not leading to such a premature termination of translation this however is less straightforward. For sure this is the case for mutations that give rise to single amino acid substitutions (~10% of *MSH2*, ~30% of *MLH1*, and ~37% of *MSH6* mutations), or small in-frame deletions (9). These mutations, with an unclear pathogenic nature, are often referred to as unclassified variants (UVs). The question, whether a UV contributes to the disease phenotype or merely represents a rare polymorphism, constitutes a major problem with obvious direct clinical consequences.

Several criteria (52-54) are generally applied, usually in combination, to assess the possible pathogenicity of a missense mutation: (1) de novo appearance of a UV; (2) segregation of the UV with the disease within pedigrees; (3) absence of the UV in control individuals; (4) a change of amino acid polarity or size; (5) occurrence of the amino acid change in a domain which is evolutionary conserved between species and/or shared between proteins belonging to the same protein family; (6) effect of the UV in a functional assay or in an animal model. In case of cancer genetics one might add (7) loss of the non-mutated allele in the tumour (usually by a large deletion, loss of heterozygosity, LOH) and (8) loss of protein expression in the tumour. In case of Lynch syndrome MSI of the tumour is used as well.

Previous inclusion of the UV in the increasing number of disease-specific mutation databases is also often considered. However, as most databases do not add any information besides the fact that the mutation was identified in a patient we recommend not to use this criterion. Several of the above mentioned criteria are sometimes difficult to score. Furthermore, one should be critical on some (53). For instance point 3, segregation of the mutation with the disease, can be caused by linkage disequilibrium of the UV with the real un-identified mutation. LOH and negative IHC and in case of Lynch

syndrome MSI are no direct evidence and results obtained might again be due to an unidentified mutation in the same gene or in its regulatory sequences.
We, and many with us, therefore believe that final proof should come from fulfilment of a combination of criteria including criterion 6, the effect of a mutation in a functional assay or in an animal model.

3.2 Functional assays for MMR proteins

Most functional assays try to measure a specific effect which a possibly causative mutation (UV) has on the function of the mutated MMR protein. Assays have been applied to measure the capacity of two MMR-proteins to bind to each other (dimer-formation). These include pull down assays and yeast two hybrid assays or transient expression of the protein in a relevant MMR-deficient cell line to determine stability of the protein complex (55). Other assays have been applied to determine the intrinsic mismatch nucleotide dependent ATPase activity of MSH protein heterodimers (56) and as one might also anticipate that trafficking or the subcellular localisation of the mutated MMR protein is disturbed, assays have been developed to determine the subcellular localisation of the mutant proteins (57).
Besides testing specific functions of the repair proteins, it seems logical to determine the repair capacity of the mutated protein as a whole. Assays have been set up in which a MMR-deficient yeast strain or human cell line were complemented with the missing protein. This is achieved by transient transfections of expression vectors harbouring the corresponding yeast or human CDNA (wild type or mutant variant)(58-60). Another approach is to restore MMR by adding to protein extracts of MMR deficient cell lines mutant MMR proteins (a cell free system) (59;61).

Section 4 Aim of the thesis

Identification of MMR-gene mutation carriers and enrolling mutation carriers in surveillance programs for the early detection of tumours, has shown to increase survival (62). Present prescreening methods are based on either family history (*e.g.* Amsterdam criteria) or tumour analysis. Unfortunately, tumours may not be available for analysis and the same is true for normal tissue (blood is commonly used) from affected relatives. We therefore tried to develop a lymphocyte-based screening method to predict the presence of particular mismatch repair gene mutations in patients (in case their or their relatives, tumours would be unavailable) or in healthy relatives (in case tumours as well as normal tissue DNA from their affected relatives would be unavailable) from families suspected for Lynch syndrome.

Although a substantial proportion of the nontruncating mutations/variants group might be pathogenic, their clinical importance is as yet uncertain. Microsatellite instability and immunohistochemistry for MMR-gene expression proved to be only partially helpful in deciding on pathogenicity. Classification of the pathogenicity of nontruncating mutations is important in itself for scientific reasons, but has also obvious consequences for patient selection for mutation analysis (*e.g.* the presence of an MMR protein coded by a gene with a missense mutation does not prove normal function of that protein) and for genetic counseling and presymptomatic testing options. This thesis aims at determining the clinical relevance of nontruncating variants in the DNA mismatch repair genes (*MLH3* and *MSH6*) by different kinds of functional assays. It is also imperative that the interpretation of these variants should be coordinated and that this information is made available to as many qualified investigators and clinicians as possible. We therefore decided to compile a database containing all the available data on published MMR-UVs that have been functionally tested so far and that can be updated in the future.

Section 5 Outline of the thesis

In this thesis we describe the development of a (blood) leukocyte-based screening method to identify mismatch repair gene mutations in individuals that might be carrying a MMR gene mutation. We describe how we try to use quantitative protein and RNA levels to identify such germline mutation carriers (**Chapter 2**).
In **Chapter 3** we review the various functional assays described for the Lynch syndrome-associated MMR proteins. We, furthermore, report on the outcomes of these tests on UVs identified in patients diagnosed with or suspected of having Lynch syndrome. The database used in chapter 3 is described in more details in **Chapter 4**. The database we constructed is filled with all details concerning functional assays previously published for Lynch syndrome associated MMR mutant proteins.
To evaluate the pathogenicity of missense mutations found previously by our group we have set up functional assays for both MSH6 and MLH3. In **Chapter 5**, functional assays are described that evaluate the pathogenicity of 11 MLH3 UVs. These mutations were previously found in patients suspected of Lynch syndrome. In **Chapter 6**, a functional assay is described to evaluate the pathogenicity of 5 MSH6 UVs found in patients suspected of Lynch syndrome. In **Chapter 7** we discuss our findings and give recommendation for future studies.

Reference List

(1) Chung DC, Rustgi AK. The hereditary nonpolyposis colorectal cancer syndrome: genetics and clinical implications. *Ann Intern Med* 2003;**138**(7):560-70.

(2) Warthin, A. S. Heredity with reference to carcinoma. Arch.Intern.Med. 12, 546-555. 1913. Ref Type: Generic

(3) Lynch, H. T. Hereditary factors in cancer: study of two large midwestern kindreds. Arch.Intern.Med. 117, 206-212. 1966. Ref Type: Generic

(4) Vasen HF, Mecklin JP, Khan PM *et al*. The International Collaborative Group on Hereditary Non-Polyposis Colorectal Cancer (ICG-HNPCC). *Dis Colon Rectum* 1991;**34**(5):424-5.

(5) Berends MJ, Wu Y, Sijmons RH *et al*. Clinical definition of hereditary non-polyposis colorectal cancer: a search for the impossible? *Scand J Gastroenterol Suppl* 2001;(234):61-7.

(6) Mecklin JP, Jarvinen HJ, Peltokallio P. Cancer family syndrome. Genetic analysis of 22 Finnish kindreds. *Gastroenterology* 1986;**90**(2):328-33.

(7) Houlston RS, Collins A, Slack J *et al*. Dominant genes for colorectal cancer are not rare. *Ann Hum Genet* 1992;**56 (Pt 2)**:99-103.

(8) Loeb LA. Mutator phenotype may be required for multistage carcinogenesis. *Cancer Res* 1991;**51**(12):3075-9.

(9) Peltomaki P, Vasen H. Mutations associated with HNPCC predisposition -- Update of ICG-HNPCC/INSiGHT mutation database. *Dis Markers* 2004;**20**(4-5):269-76.

(10) de la Chapelle A. Genetic predisposition to colorectal cancer. *Nat Rev Cancer* 2004;**4**(10):769-80.

(11) Salovaara R, Loukola A, Kristo P *et al*. Population-based molecular detection of hereditary nonpolyposis colorectal cancer. *J Clin Oncol* 2000;**18**(11):2193-200.

(12) Katballe N, Christensen M, Wikman FP *et al*. Frequency of hereditary non-polyposis colorectal cancer in Danish colorectal

cancer patients. *Gut* 2002;**50**(1):43-51.

(13) Tishkoff DX, Boerger AL, Bertrand P *et al*. Identification and characterization of *Saccharomyces cerevisiae* EXO1, a gene encoding an exonuclease that interacts with MSH2. *Proc Natl Acad Sci U S A* 1997;**94**(14):7487-92.

(14) Tran PT, Simon JA, Liskay RM. Interactions of Exo1p with components of MutLalpha in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 2001;**98**(17):9760-5.

(15) Lipkin SM, Wang V, Jacoby R *et al*. MLH3: a DNA mismatch repair gene associated with mammalian microsatellite instability. *Nat Genet* 2000;**24**(1):27-35.

(16) Flores-Rozas H, Kolodner RD. The *Saccharomyces cerevisiae* MLH3 gene functions in MSH3-dependent suppression of frameshift mutations. *Proc Natl Acad Sci U S A* 1998;**95**(21):12404-9.

(17) Marsischky GT, Filosi N, Kane MF *et al*. Redundancy of *Saccharomyces cerevisiae* MSH3 and MSH6 in MSH2-dependent mismatch repair. *Genes Dev* 1996;**10**(4):407-20.

(18) Jiricny J. The multifaceted mismatch-repair system. *Nat Rev Mol Cell Biol* 2006;**7**(5):335-46.

(19) Shin KH, Han HJ, Park JG. Growth suppression mediated by transfection of wild-type hMLH1 in human cancer cells expressing endogenous truncated hMLH1 protein. *Int J Oncol* 1998;**12**(3):609-15.

(20) Buermeyer AB, Wilson-Van Patten C, Baker SM *et al*. The human MLH1 cDNA complements DNA mismatch repair defects in Mlh1-deficient mouse embryonic fibroblasts. *Cancer Research* 1999;**59**(3):538-41.

(21) Schofield MJ, Hsieh P. DNA mismatch repair: Molecular mechanisms and biological function. *Annual Review of Microbiology* 2003;**57**:579-608.

(22) Stojic L, Brun R, Jiricny J. Mismatch repair and DNA damage signalling. *DNA Repair (Amst)* 2004;**3**(8-9):1091-101.

(23) Berends MJ, Wu Y, Sijmons RH *et al*. Molecular and clinical characteristics of MSH6 variants: an analysis of 25 index carriers of

a germline variant. *Am J Hum Genet* 2002;**70**(1):26-37.

(24) Lynch HT, Lemon SJ, Karr B *et al*. Etiology, natural history, management and molecular genetics of hereditary nonpolyposis colorectal cancer (Lynch syndromes): genetic counseling implications. *Cancer Epidemiol Biomarkers Prev* 1997;**6**(12):987-91.

(25) Hendriks YM, Jagmohan-Changur S, van der Klift HM *et al*. Heterozygous mutations in PMS2 cause hereditary nonpolyposis colorectal carcinoma (Lynch syndrome). *Gastroenterology* 2006;**130**(2):312-22.

(26) Wu Y, Berends MJ, Post JG *et al*. Germline mutations of EXO1 gene in patients with hereditary nonpolyposis colorectal cancer (HNPCC) and atypical HNPCC forms. *Gastroenterology* 2001;**120**(7):1580-7.

(27) Wu Y, Berends MJ, Sijmons RH *et al*. A role for MLH3 in hereditary nonpolyposis colorectal cancer. *Nat Genet* 2001;**29**(2):137-8.

(28) Sun XM, Zheng L, Shen BH. Functional alterations of human exonuclease 1 mutants identified in atypical hereditary nonpolyposis colorectal cancer syndrome. *Cancer Research* 2002;**62**(21):6026-30.

(29) Alam NA, Gorman P, Jaeger EEM *et al*. Germline deletions of EXO1 do not cause colorectal tumors and lesions which are null for EXO1 do not have microsatellite instability. *Cancer Genetics and Cytogenetics* 2003;**147**(2):121-7.

(30) Liu HX, Zhou XL, Liu T *et al*. The role of hMLH3 in familial colorectal cancer. *Cancer Research* 2003;**63**(8):1894-9.

(31) Hienonen T, Laiho P, Salovaara R *et al*. Little evidence for involvement of MLH3 in colorectal cancer predisposition. *International Journal of Cancer* 2003;**106**(2):292-6.

(32) Jagmohan-Changur S, Poikonen T, Vilkki S *et al*. EXO1 variants occur commonly in normal population: Evidence against a role in hereditary nonpolyposis colorectal cancer. *Cancer Research* 2003;**63**(1):154-8.

(33) Thompson E, Meldrum CJ, Crooks R *et al*. Hereditary non-polyposis colorectal cancer and the role of hPMS2 and hEXO1 mutations. *Clin Genet* 2004;**65**(3):215-25.

(34) Vasen HF, Watson P, Mecklin JP *et al*. New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC. *Gastroenterology* 1999;**116**(6):1453-6.

(35) Lamberti C, Kruse R, Ruelfs C *et al*. Microsatellite instability-a useful diagnostic tool to select patients at high risk for hereditary non-polyposis colorectal cancer: a study in different groups of patients with colorectal cancer. *Gut* 1999;**44**(6):839-43.

(36) Syngal S, Fox EA, Eng C *et al*. Sensitivity and specificity of clinical criteria for hereditary non- polyposis colorectal cancer associated mutations in MSH2 and MLH1. *J Med Genet* 2000;**37**(9):641-5.

(37) Rodriguez-Bigas MA, Boland CR, Hamilton SR *et al*. A National Cancer Institute Workshop on Hereditary Nonpolyposis Colorectal Cancer Syndrome: meeting highlights and Bethesda guidelines. *J Natl Cancer Inst* 1997;**89**(23):1758-62.

(38) Bapat BV, Madlensky L, Temple LK *et al*. Family history characteristics, tumor microsatellite instability and germline MSH2 and MLH1 mutations in hereditary colorectal cancer. *Hum Genet* 1999;**104**(2):167-76.

(39) Fujita S, Moriya Y, Sugihara K *et al*. Prognosis of hereditary nonpolyposis colorectal cancer (HNPCC) and the role of Japanese criteria for HNPCC. *Jpn J Clin Oncol* 1996;**26**(5):351-5.

(40) Umar A, Boland CR, Terdiman JP *et al*. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *J Natl Cancer Inst* 2004;**96**(4):261-8.

(41) Jones PA, Laird PW. Cancer epigenetics comes of age. *Nat Genet* 1999;**21**(2):163-7.

(42) Buermeyer AB, Deschenes SM, Baker SM *et al*. Mammalian DNA mismatch repair. *Annu Rev Genet* 1999;**33**:533-64.

(43) Raschle M, Marra G, Nystrom-Lahti M *et al*. Identification of hMutLbeta, a heterodimer of hMLH1 and hPMS1. *J Biol Chem* 1999;**274**(45):32368-75.

(44) Gonzalo V, Castellvi-Bel S, Castells A. MLH1 germ-line epimutations: is there strong evidence of its inheritance?

Gastroenterology 2007;**133**(3):1042-4.

(45) Southey MC, Jenkins MA, Mead L *et al*. Use of molecular tumor characteristics to prioritize mismatch repair gene testing in early-onset colorectal cancer. *J Clin Oncol* 2005;**23**(27):6524-32.

(46) Hampel H, Frankel WL, Martin E *et al*. Screening for the Lynch syndrome (hereditary nonpolyposis colorectal cancer). *N Engl J Med* 2005;**352**(18):1851-60.

(47) Pinol V, Castells A, Andreu M *et al*. Accuracy of revised Bethesda guidelines, microsatellite instability, and immunohistochemistry for the identification of patients with hereditary nonpolyposis colorectal cancer. *JAMA* 2005;**293**(16):1986-94.

(48) Stormorken AT, Bowitz-Lothe IM, Noren T *et al*. Immunohistochemistry identifies carriers of mismatch repair gene defects causing hereditary nonpolyposis colorectal cancer. *J Clin Oncol* 2005;**23**(21):4705-12.

(49) Niessen RC, Berends MJ, Wu Y *et al*. Identification of mismatch repair gene mutations in young patients with colorectal cancer and in patients with multiple tumours associated with hereditary non-polyposis colorectal cancer. *Gut* 2006;**55**(12):1781-8.

(50) Hayashi T, Arai M, Ueno M *et al*. Frequency of immunohistochemical loss of mismatch repair protein in double primary cancers of the colorectum and stomach in Japan. *Dis Colon Rectum* 2006;**49**(10 Suppl):S23-S29.

(51) Fields JZ, Gao Z, Gao Z *et al*. Immunoassay for wild-type protein in lymphocytes predicts germline mutations in patients at risk for hereditary colorectal cancer. *J Lab Clin Med* 2004;**143**(1):59-66.

(52) Syngal S, Fox EA, Li C *et al*. Interpretation of genetic test results for hereditary nonpolyposis colorectal cancer: implications for clinical predisposition testing. *JAMA* 1999;**282**(3):247-53.

(53) Hofstra RM, Osinga J, Buys CH. Mutations in Hirschsprung disease: when does a mutation contribute to the phenotype. *Eur J Hum Genet* 1997;**5**(4):180-5.

(54) Cotton RG, Scriver CR. Proof of "disease causing" mutation. *Hum Mutat* 1998;**12**(1):1-3.

(55) Raevaara TE, Korhonen MK, Lohi H *et al*. Functional significance and clinical phenotype of nontruncating mismatch repair variants of MLH1. *Gastroenterology* 2005;**129**(2):537-49.

(56) Heinen CD, Wilson T, Mazurek A *et al*. HNPCC mutations in hMSH2 result in reduced hMSH2-hMSH6 molecular switch functions. *Cancer Cell* 2002;**1**(5):469-78.

(57) Brieger A, Plotz G, Raedle J *et al*. Characterization of the nuclear import of human MutLalpha. *Mol Carcinog* 2005;**43**(1):51-8.

(58) Drotschmann K, Clark AB, Tran HT *et al*. Mutator phenotypes of yeast strains heterozygous for mutations in the MSH2 gene. *Proc Natl Acad Sci U S A* 1999;**96**(6):2970-5.

(59) Holmes J, Jr., Clark S, Modrich P. Strand-specific mismatch correction in nuclear extracts of human and Drosophila melanogaster cell lines. *Proc Natl Acad Sci U S A* 1990;**87**(15):5837-41.

(60) Trojan J, Zeuzem S, Randolph A *et al*. Functional analysis of hMLH1 variants and HNPCC-related mutations using a human expression system. *Gastroenterology* 2002;**122**(1):211-9.

(61) Umar A, Boyer JC, Kunkel TA. DNA loop repair by human cell extracts. *Science* 1994;**266**(5186):814-6.

(62) Jarvinen HJ, Aarnio M, Mustonen H *et al*. Controlled 15-year trial on screening for colorectal cancer in families with hereditary nonpolyposis colorectal cancer. *Gastroenterology* 2000;**118**(5):829-34.

Chapter 2

Can the presence of germline mismatch repair (MMR) gene mutations be predicted from the level of MMR proteins in leukocyte cultures?

Jianghua Ou[1], Paul Jager[1], Helga Westers[1], Anne-Marie van der Kevie[1], Renée C. Niessen[1], Karin Ansink[1], Erica Gerkes[1], Jan H. Kleibeuker[2], Robert M.W. Hofstra[1], Rolf H. Sijmons[1]

(1) Department of Genetics and (2) Gastroenterology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands

Manuscript in preparation

Abstract

Background

Identifying a MMR gene mutation in persons or patients suspected for Lynch syndrome, also referred to as hereditary non-polyposis colorectal cancer (HNPCC), is costly and time consuming. Screening for mutations is therefore only performed in individuals at high risk and pre-screening of tumour tissue by immunohistochemistry staining is preferred to direct mutation analysis. When no tumour material is available, all Lynch syndrome-associated MMR genes need to be screened for mutations in patients and this diagnostic approach is even less efficient if only DNA from asymptomatic relatives is available as often multiple individuals need to be screened. In search of an alternative pre-screening approach we have tested a method that quantifies the amounts of MMR proteins in leukocyte cultures. It is based on the hypothesis that (1) in dividing cells MMR genes are unregulated and, therefore, the products of these genes are detectable, and that (2) individuals with a germline truncating mutation in a particular MMR gene show a reduced level of protein encoded by that gene.

Methods

Western blotting was performed on protein extracts isolated from uncultured and shortly cultured leukocytes (3 days of culturing). Carriers of truncating *MLH1* or *MSH2* mutations or patients carrying in-frame deletions (5 patients for each gene) and healthy controls were studied. Densitometric quantification of the protein bands corresponding to MLH1 and MSH2 was performed. Moreover, quantitative *MLH1* and *MSH2* mRNA analysis was performed.

Results

The MLH1 and MSH2 protein products could be detected in the cultured leukocytes and not in the uncultured cells. However, the mutated gene in proven mutation carriers could not be predicted from densitometric analysis of the MLH1 and MSH2 protein bands seen on the Western blots. Furthermore, our analysis showed that there is a wide variation in the expression levels of MLH1 and MSH2 both on the mRNA and the protein level in patients and controls.

Conclusion

Shortly cultured leukocytes express easily detectable levels of MLH1 and MSH2 protein. Because of the wide variation in protein levels observed within and between the groups of mutation carriers and controls, these levels cannot be used to predict the presence of a germline truncating *MLH1* or *MSH2* mutation with the current method.

Introduction

Lynch syndrome, also referred to as hereditary non-polyposis colorectal cancer (HNPCC), is one of the most common hereditary cancer syndromes. It is an autosomal dominant disorder characterized clinically by increased risks to develop colorectal and endometrial cancer and a range of other tumours (1;2). It is caused by germline mutations in one of the DNA mismatch repair (MMR) genes *MLH1*, *MSH2*, *MSH6* and *PMS2* (3-5). Lynch syndrome-associated tumours usually show microsatellite instability (MSI) as a manifestation of their MMR-deficiency (6;7), and loss of immunohistochemical (IHC) staining for the MMR protein encoded by the corresponding mutated MMR gene (8-10).
Screening for germline MMR gene mutations in patients who have a personal or family history suggestive of Lynch syndrome is costly and time consuming. Screening for mutations is therefore usually restricted to those individuals that have tumours featuring the above mentioned characteristics.
If, however, tumours are unavailable then this type of pre-screening to select for MMR gene mutation analysis is not an option and affected individuals are then usually offered mutation analysis. In some families suggestive of having Lynch syndrome, neither tumour material nor DNA from affected relatives is available for analyses. In those cases, MMR gene mutation analysis can be offered to (a number of) unaffected close relatives of the affected patients, but this approach is even less efficient than in affected patients with untested tumours. A pre-screening method in normal tissue might therefore be a welcome addition to the existing selection techniques. For this reason we wondered if measuring expression levels of the MMR genes or of the encoded proteins in blood leucocytes could help predicting the presence of truncating germline MMR gene mutations, as mutation carriers would be expected to produce less, theoretically 50%, of the corresponding MMR protein in their dividing normal cells. Although it has been recently demonstrated that not all autosomal genes normally have bi-allelic expression, this does appear to be the rule for the *MLH1* gene in lymphoblastoid cells (11). To the best of our knowledge, only a single paper has reported on such an approach. Fields *et al*. (12) demonstrated reduced levels of MMR protein in EBV transformed lymphoblastoid cell lines of known MMR gene mutation carriers. However, making EBV transformed cell lines is costly and time consuming. In the ideal situation we would like to analyse leukocytes directly. However, as MMR protein levels are possibly low or not present in non-dividing cells, we also opted for a short culture procedure as routinely used in cytogenetic laboratories. Using protein extracts and RNA isolated from both uncultured and cultured leukocytes we investigated whether differences in expression of the MMR proteins or RNA levels between known MMR gene mutation

carriers and controls could be determined. Furthermore, we studied whether the afore mentioned procedures could be used in our routine diagnostics for the identification of mutation carriers.

Material and methods

Patients and Controls

Individuals included in this study are either proven MMR mutation carriers or healthy controls. We included 5 carriers of *MLH1* and 5 of a *MSH2* pathogenic germline mutation. As a control group we included 10 anonymous healthy individuals. Of these controls, blood from 5 individuals was used for uncultured leukocyte isolation and blood from 10 individuals for leukocyte culturing. The mutations tested are listed in Table 1. Among these mutations several deletion-mutations are present, all of which are considered pathogenic as either a whole exon or multiple exons are deleted or functional assays have been performed that showed pathogenicity (3 bp deletion in *MLH1*) (13).

Leukocyte culturing

Ten ml of heparinized whole peripheral blood was drawn from each of the selected individuals. For each individual 8 cultures were set up. Each culture was started by adding 0.5ml of heparinized whole blood to 10ml of RPMI 1640 medium (Invitrogen, Breda, the Netherlands) in a T25 flask with 5% Fetal Calf Serum (Invitrogen, Breda, the Netherlands) and 0.1485 mg /ml of Phyto Haemaglutinin (Remel Inc. Lenexa, USA). The cultures were incubated at 37ºC in 5% CO_2 atmosphere for 72 hours. Directly before and after culturing, the leukocytes were counted as an index on divisional activity. Out of these cultures leukocytes were isolated using a modified hypotonic method (protocol available upon request).

Protein isolations

Six of the eight T25 flasks of leukocyte isolations were used for protein isolations. Leukocyte pellets were lysed for 1 hour at 4ºC in 500 µl CelLytic ^{TM} M (mammalian cell lysis reagent; Sigma, Saint Louis, U.S.A) with one protease inhibitor tablet was added (1x final concentration; Roche, Basel, Switzerland). Lysates were stored at -80ºC.

RNA isolation

Total RNA was isolated out of two T25 flasks. The leukocyte pellets were added to 1 ml RNAbee (AMS Biotechnology Europe Ltd, Abingdon, Oxon, UK). Next, total RNA was isolated according to the manufacturer's instructions and cleaned with the RNeasy MinElute Cleanup Kit (Qiagen Benelux B.V. Netherlands, Venlo, the Netherlands) to remove genomic DNA

contamination. RNA quality was evaluated with an Agilent 2100 Bioanalyzer (Agilent Technologies Netherlands B.V., Amstelveen, the Netherlands). First strand cDNA was prepared from 1 μg of total cellular RNA using random hexadeoxynucleotide primers and first strand cDNA synthesis ready to use beads (Amersham, GE Healthcare Europe GmbH, Diegem, Belgium).

Western blot analysis for the detection of MMR protein

100 μg of total protein extracts from uncultured or cultured leukocytes was separated on denaturing 6% SDS polyacrylamide gels and blotted onto nitrocellulose membranes (Bio-Rad, Veenendaal, the Netherlands). The following antibodies were used: anti-MSH2 (Catalog No. NA27, Calbiochem, La Jolla, USA; immunogen: a C-terminal fragment of human MSH2; 1:500 dilution); anti-MLH1 (clone G168-728, Pharmingen, Breda, the Netherlands; immunogen: full-length protein; 1:500 dilution); anti-mouse IgG HRP-conjugated secondary antibody (Pierce Biotechnology, Rockford, USA). Signal visualization was performed with the Supersignal West Dura Extended Duration Substrate kit (Pierce Biotechnology, Rockford, USA). ECL signals were visualized with a G:BOX Chemi HR16 automated image analyser (SYNGENE, Cambridge, UK) and unsaturated blot images were obtained by using the Genesnap software (SYNGENE, Cambridge, UK). Densitometric analyses of the signals was performed using the GENETOOLS software of the image analyzer

Quantitative PCR

Quantitative PCR was performed in a volume of 25 μl containing 1x iQ™ SYBR Green supermix (Bio-Rad laboratories B.V, Veenendaal, the Netherlands). SYBR Green fluorescence was measured real time during PCR (IQ 6 PCR, Bio-Rad Laboratories B.V, Veenendaal, the Netherlands). Quantitative PCR was performed in triplicate for *MLH1* and *MSH2* primer sets and the mean value was used as relative quantification. For normalization *TBP* (TATA box binding protein) was used. The primers used for *MLH1* are: 5'-TGAGGTGAATTGGGACGAAGA-3' and 5'GTGTGAGCGCAAGGCTTTATAGA-3'; the primers used for *MSH2* are: 5'-GCTAAACAGAAAGCCCTGGAAC-3 and 5'-ACCTTGGACAGGAACTCCTGAAT-3'; the primers used for *TBP* are: 5'-GCCCGAAACGCCGAATAT-3' and 5'-CCGTGGTTCGTGGCTCTCT-3'.

Results

Culturing leukocytes

The leukocyte number was counted directly before culturing and after culturing. The count number of leukocytes directly before culturing was approximately 2-3 x10^9/ml. After 3 days of culturing the number of cells had doubled (between 4 and 6 x10^9 /ml).

Western blotting of MLH1 and MSH2 in control samples

Western blotting on protein extracts from uncultured leukocytes from controls revealed no detectable signal for the MLH1 or MSH2 proteins (Figure 1). Western blotting of protein extracts isolated from cultured leukocytes from controls did show detectable levels of the MLH1 and MSH2 proteins (Figure 1).

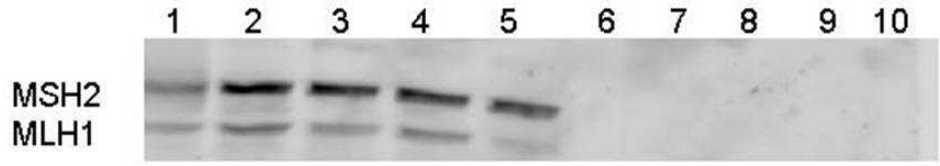


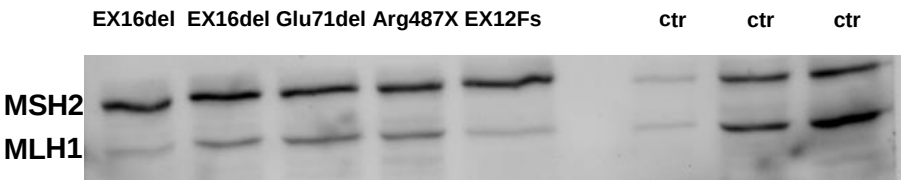
Figure 1. Western-blot analysis for MLH1 and MSH2 in Leukocytes from control individuals. Lanes 1-5 are cultured leukocytes (three days of culturing), lanes 6-10 are loaded with protein extracts form uncultured blood leukocytes. The blot was probed with MLH1 and MSH2.

MLH1 mutation carriers

Figure 2 shows that the MLH1 protein levels of 5 proven *MLH1* mutation carriers are decreased. The amounts of MLH1 for the 5 mutation carriers differed considerably. These samples (lanes 1-5) belong to individuals with pathogenic *MLH1* mutations (lanes 1 & 2: EX16del; lane 3 p.Glu71del; lane 4 p.Arg487X; lane 5 p.EX12FS, designated name details see table 1). As we expect that the ratio between the MLH1 and MSH2 protein should be constant (for proper repair) we used MSH2 for normalisation. We calculated the MSLH1/MSH2 ratio (see table 1 and figure 2 Beneath the MLH1/MSH2 ratios are given). The ratio data show that these ratios vary (even between individuals with the same mutation) in both mutation carriers and controls.

Table 1. Immunoassay for MMR ratios in cultured leukocytes from 10 mutation carriers. *MSH2/MLH1

Patient	Gene	designated name	Mutation	Densitometric ratio MLH1/MSH2
1	MLH1	Ex16del	deletion of exon 16 (c.1732-?_1896+?del)	0.16
2	MLH1	Ex16del	deletion of exon 16 (c.1732-?_1896+?del)	0.26
3	MLH1	E71del	Glu71del or c.211-213delGAA, in frame	0.71
4	MLH1	Arg487X	Arg487X, truncating mutation	0.43
5	MLH1	Ex12fs	c.1224_1258del,exon 12 frameshift	0.17
6	MSH2	EX5Fs	c.922_923insAGCAGTCA frameshift	1.25*
7	MSH2	EX5Fs	c.922_923insAGCAGTCA frameshift	5.83*
8	MSH2	EX5Fs	c.922_923insAGCAGTCA frameshift	5.71*
9	MSH2	EX4_7del	c.646-?_1276+?del, frameshift	1.03*
10	MSH2	EX12_16del	MSH2 C.1760-?_2802+?del, frameshift	1.4*



MLH1/MSH2 0.16 0.26 0.71 0.43 0.17 0.65 0.84 1.88

Figure 2. Western-blot analysis for MLH1 and MSH2 in cultured leukocytes from *MLH1* mutation carriers (lane 1- 5) and in 3 healthy control samples (lane 7-9). The ratio between MLH1 and MSH2 (MLH1/MSH2) is given below the figure.

MSH2 mutation carriers

Figure 3 shows no significant differences for the MSH2/MLH1 ratio between MSH2 mutation carriers and healthy controls. In this experiment we used MLH1 for normalisation. As for MLH1, large variation in the level of expression among MSH2 mutation carriers or the healthy control group is observed.

That the expression level of this protein is variable is further strengthened by the results obtained from three individuals harbouring the same MSH2 mutation (lanes 1-3 Figure 3).

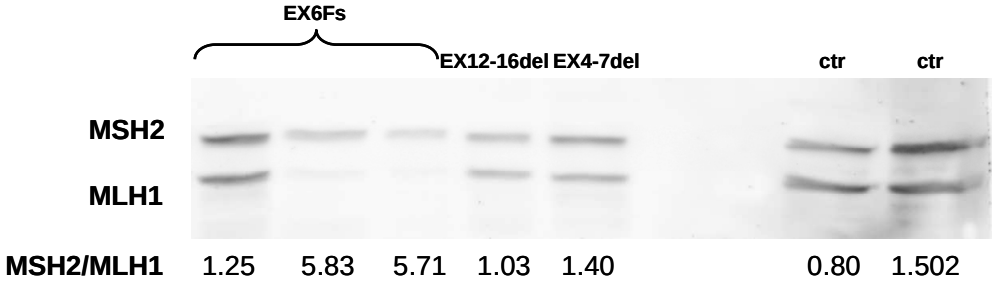


Figure 3. Western-blot analysis for MLH1 and MSH2 in cultured leukocytes from MSH2 mutation carriers (lane 1- 5) and in 2 healthy control samples (lane 8-9). The ratio between MSH2 and MLH1 (MSH2/MLH1) is given below the figure.

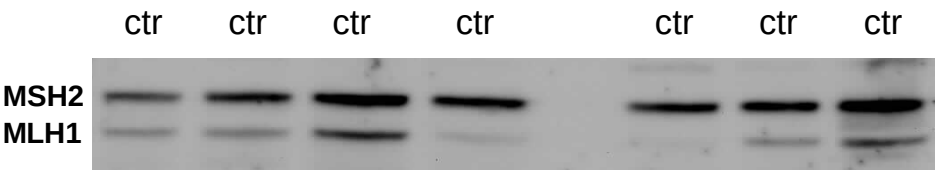
Additional controls for Western blotting

As we did observe large deviations in the levels of protein expression of cultured leukocytes from both the *MLH1* and *MSH2* mutation carriers and the 3 controls, we analysed an additional 7 controls. The data obtained (Figure 4a) again illustrates that large variations in the eventual protein levels exist between individuals.

Quantitative PCR on controls

Quantitative PCR (Figure 4b) shows that the *MLH1* expression levels in 7 control samples, when normalised to the housekeeping gene *TBP*, show slight differences in expression, as all expression levels are within a 0.4 fold range. For *MSH2* there was an about 5 fold difference between the highest levels and the lowest levels of expression. When we determined the *MLH1/MSH2* ratios we also observed a 4.6 fold difference between the highest levels and the lowest levels of expression.

A.



B.

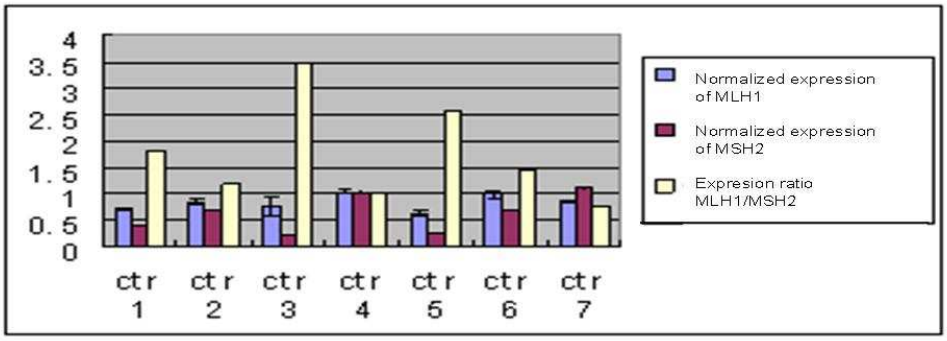


Figure 4. A. Western-blot analysis for MLH1 and MSH2 in cultured leukocytes from 7 healthy control samples. B. Quantitative real time PCR analysis of *MLH1* and *MSH2* mRNA level in 7 healthy controls.

Discussion

In contrast to measurements in uncultured leukocytes, MLH1 and MSH2 proteins could easily be detected in leukocytes cultured for 3 days according to a culture method routinely used in cytogenetics laboratories. However, due to a large variation in the eventual results within and between the groups of mutation carriers and controls in the level of these proteins, no reliable prediction could be made for the mutation status. Several factors could have contributed to the large intersample variability. These reasons could be technical. Western blotting is a semi-quantitative technique. Saturation of signals limits proper quantification. The variation might also be caused by differences in the functioning of the nonsense mediated RNA decay pathway (NMD pathway). The NMD pathway has been shown to degrade transcripts harbouring disease causing nonsense or frameshift mutations (14;15). However, NMD efficiency may vary considerably between different cell types and the treatment of those cell types (16-18). A wide variation in the abundancy of the MMR encoded protein between different individuals may be physiological. Yu *et al*. (19) recently observed

up to 4761.5 fold differences in MLH1 mRNA expression between normal colon tissues. Similar physiological differences may apply to other tissues, including leukocytes and might be the best explanation for our findings. Whatever the reason for the wide variation in protein levels within and between the groups of mutation carriers and controls and between cultures may be, presently, we cannot recommend this method as a practical clinical tool.

Alternative pre-selection methods have been reported. Fields *et al.*. (12) demonstrated reduced levels of MMR protein in EBV transformed lymphoblastoid cell lines of known MMR gene mutation carriers. They also found large variation in the ratios between MSH2 and MLH1, as protein ratios range from 0.19 to 0.71 in truncated *MSH2* germline mutation carriers. However, detailed data was not shown for their control group, so it is difficult to weigh these data with respect to the predictive power of this test for mutation status. In addition, the facility for making EBV-transformed cell lines may not be available in every lab. To the best of our knowledge this method has not been reported by other investigators or introduced in clinical laboratories as a patient selection tool for MMR gene mutation analysis.
Another approach is the demonstration at the mRNA level of expression of one or both alleles of a particular MMR gene. This technique requires the presence of a polymorphism that can distinguish between the two alleles of each of the Lynch syndrome associated genes. This approach has been shown to work in some cases (18). However one needs informative coding markers in all MMR genes which might turn out difficult. In addition, NMD efficiency depends on the position of a mutation and therefore allelic imbalance may differ in all mutation carriers (18). Together, this limits the use of this approach in larger groups of Lynch syndrome suspected patients.

In summary, we have demonstrated easily detectable levels of MLH1 and MSH2 proteins in shortly cultured leukocytes from *MLH1* and *MSH2* mutation carriers and controls. However, these protein levels could not reliably predict mutation status. There is a clinical need for a method to select patients (or their relatives in case DNA from patients is unavailable) for mutation analysis when tumour tissue is unavailable. Presently, no such method for practical use on a larger scale is available. In the future this may pose less of a problem, as solid phase sequencing is getting cheaper and sequencing the Lynch syndrome-associated genes in a single experiment is becoming available. Therefore mutation analysis will become more affordable and can be applied to larger groups of individuals with less need for molecular pre-selection methods.

Acknowledgments

This work was supported by: the Dutch Cancer Society (grant No. RUG2002-2678), the European Community (FP6-2004-LIFESCIHEALTH-5, proposal No 018754).

Reference List

(1) Lynch HT, Krush AJ. Cancer family "G" revisited: 1895-1970. *Cancer* 1971;**27**(6):1505-11.

(2) Mecklin JP, Jarvinen HJ. Tumor spectrum in cancer family syndrome (hereditary nonpolyposis colorectal cancer). *Cancer* 1991;**68**(5):1109-12.

(3) Lynch HT, de la CA. Genetic susceptibility to non-polyposis colorectal cancer. *J Med Genet* 1999;**36**(11):801-18.

(4) Peltomaki P, Vasen H. Mutations associated with HNPCC predisposition -- Update of ICG-HNPCC/INSiGHT mutation database. *Dis Markers* 2004;**20**(4-5):269-76.

(5) Peltomaki P, Vasen HF. Mutations predisposing to hereditary nonpolyposis colorectal cancer: database and results of a collaborative study. The International Collaborative Group on Hereditary Nonpolyposis Colorectal Cancer. *Gastroenterology* 1997;**113**(4):1146-58.

(6) Thibodeau SN, Bren G, Schaid D. Microsatellite instability in cancer of the proximal colon. *Science* 1993;**260**(5109):816-9.

(7) Parsons R, Li GM, Longley MJ *et al*. Hypermutability and mismatch repair deficiency in RER+ tumor cells. *Cell* 1993;**75**(6):1227-36.

(8) Debniak T, Kurzawski G, Gorski B *et al*. Value of pedigree/clinical data, immunohistochemistry and microsatellite instability analyses in reducing the cost of determining hMLH1 and hMSH2 gene mutations in patients with colorectal cancer. *Eur J Cancer* 2000;**36**(1):49-54.

(9) Paraf F, Gilquin M, Longy M *et al*. MLH1 and MSH2 protein immunohistochemistry is useful for detection of hereditary non-polyposis colorectal cancer in young patients. *Histopathology* 2001;**39**(3):250-8.

(10) Marcus VA, Madlensky L, Gryfe R *et al*. Immunohistochemistry for hMLH1 and hMSH2: a practical test for DNA mismatch repair-deficient tumors. *Am J Surg Pathol* 1999;**23**(10):1248-55.

(11) Gimelbrant A, Hutchinson JN, Thompson BR *et al*. Widespread monoallelic expression on human autosomes. *Science* 2007;**318**(5853):1136-40.

(12) Fields JZ, Gao Z, Gao Z *et al*. Immunoassay for wild-type protein in lymphocytes predicts germline mutations in patients at risk for hereditary colorectal cancer. *J Lab Clin Med* 2004;**143**(1):59-66.

(13) Raevaara TE, Timoharju T, Lonnqvist KE *et al*. Description and functional analysis of a novel in frame mutation linked to hereditary non-polyposis colorectal cancer. *J Med Genet* 2002;**39**(10):747-50.

(14) Mendell JT, Sharifi NA, Meyers JL *et al*. Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. *Nat Genet* 2004;**36**(10):1073-8.

(15) Lewis BP, Green RE, Brenner SE. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc Natl Acad Sci U S A* 2003;**100**(1):189-92.

(16) Linde L, Boelz S, Neu-Yilik G *et al*. The efficiency of nonsense-mediated mRNA decay is an inherent character and varies among different cells. *Eur J Hum Genet* 2007;**15**(11):1156-62.

(17) Jeganathan D, Fox MF, Young JM *et al*. Nonsense-mediated RNA decay in the TSC1 gene suggests a useful tool pre- and post-positional cloning. *Hum Genet* 2002;**111**(6):555-65.

(18) Tournier I, Raux G, Di FF *et al*. Analysis of the allele-specific expression of the mismatch repair gene MLH1 using a simple DHPLC-Based Method. *Hum Mutat* 2004;**23**(4):379-84.

(19) Yu J, Mallon MA, Zhang W *et al*. DNA repair pathway profiling and microsatellite instability in colorectal cancer. *Clin Cancer Res* 2006;**12**(17):5104-11.

Chapter 3

Functional analysis helps to clarify the clinical importance of unclassified variants in DNA mismatch repair genes

Jianghua Ou,[1] Renée C. Niessen,[1] Anne Lützen,[2] Rolf H. Sijmons,[1] Jan. H. Kleibeuker,[3] Niels de Wind,[4] Lene Juel Rasmussen,[2] and Robert M.W. Hofstra[1]

[1]Department of Genetics, University Medical Center Groningen, Groningen, the Netherlands; [2]Department of Science, Systems and Models, Roskilde University, Denmark; [3]Department of Gastroenterology, University Medical Center Groningen, Groningen, the Netherlands; [4]Department of Toxicogenetics, Leiden University Medical Center, Leiden, the Netherlands

Human Mutation 2007;28(11):1047-1054

Abstract

Hereditary nonpolyposis colorectal cancer (HNPCC) or Lynch syndrome is caused by DNA variations in the DNA mismatch repair (MMR) genes MSH2, MLH1, MSH6 and PMS2. Many of the mutations identified result in premature termination of translation and thus in loss-of-function of the encoded mutated protein. These DNA variations are thought to be pathogenic mutations. However, some patients carry other DNA mutations, referred to as unclassified variants (UVs), which do not lead to such a premature termination of translation and it is not known whether these contribute to the disease phenotype or merely represent rare polymorphisms. This is a major problem which has direct clinical consequences. Several criteria can be used to classify these UVs, such as: whether they segregate with the disease within pedigrees, are absent in control individuals, show a change of amino acid polarity or size, provoke an amino acid change in a domain which is evolutionary conserved and/or shared between proteins belonging to the same protein family, and their effects in a functional assay or animal model.
In this review we discuss the various functional assays reported for the HNPCC-associated MMR proteins and the outcomes of these tests on UVs identified in patients diagnosed with or suspected of having HNPCC. We conclude that a large proportion of MMR UVs are likely to be pathogenic, suggesting that missense variants of MMR proteins do indeed play a role in HNPCC.

Introduction

Hereditary nonpolyposis colorectal cancer (HNPCC) or Lynch syndrome is an inherited disorder caused by germline mutations in DNA mismatch repair (MMR) genes. Most of the mutations found in HNPCC patients affect one of two MMR genes, MSH2 or MLH1 (Peltomaki P., *et al.*) while a small fraction of HNPCC families carry a germline mutation in the MSH6 gene [Berends *et al*., 2002]. PMS2 mutations were believed to be very rare in HNPCC [Lynch *et al*., 1997], however, recently the number of reported mutations has been rising [Hendriks *et al*., 2006].

Mismatch Repair System

The mutated genes found in HNPCC patients all encode proteins that are part of the MMR system, which comprises an evolutionary conserved set of proteins to recognize and remove replication errors that escape the proofreading activities of replicative DNA polymerases, thereby providing a back-up system for the polymerases' proofreading. The MMR system corrects not only mispaired nucleotides but also small insertion/deletion

loops resulting from slippage of the replicative polymerases during replication of microsatellites consisting of mono- or di-nucleotide repeats.
The ability of many species to repair mismatches in double-stranded DNA has been well documented. The first critical step in this process is the recognition of the mismatched DNA. In the *Escherichia coli* MMR system - still the best documented - the MutS protein accomplishes this initial step. A homodimer of the MutS protein binds to the site of a mismatch in double-stranded DNA and, with the help of the MutL and MutH proteins, targets a section around the mutated strand before removal. Other proteins complete the repair process: the target section of DNA is removed and degraded by exonucleases, the remaining single-strand DNA gap is filled by polymerase using the complementary strand as a template, followed by ligation that results in a repaired section of double-stranded DNA without mismatches.
In the human MMR system the mismatch recognition function is fulfilled by a heterodimeric protein complex composed of two MutS Homologous (MSH) proteins. The major MSH dimer consists of MSH2 and MSH6, also called MutSα. This complex is able to bind to base-base mismatches and insertion/deletion loops. A minor and partially redundant mismatch binding heterodimer (MUTSβ) consists of the MSH2 and MSH3 proteins and this complex binds to larger insertion/deletion loops [Chung and Rustgi, 2003]. In the mismatch repair process, MutSα recognizes mismatches in DNA. Binding to this mismatch induces a conformational change for MutSα, which allows ATP to bind, attract the MLH1/PMS2 heterodimer and then initiate the downstream repair event. The repair proteins are, however, also involved in, or at least have an influence, on other processes, such as apoptosis and proliferation [Shin *et al*., 1998; Jiricny, 2006].
Heterodimers composed of human MutL Homologous (MLH) proteins - MLH1, PMS2 and MLH3 - interact with the MSH2-containing DNA-protein complex and coordinate the downstream repair events. These downstream events involve different proteins such as exonucleases (*e.g.* EXO1) and DNA polymerases [Tishkoff *et al*., 1997; Tran *et al*., 2001]. The heterodimer MLH1-PMS2 (MutLα) interacts with both MSH2-MSH6 and MSH2-MSH3, while the heterodimer MLH1-MLH3 (MutLγ), appears to be restricted to interact with only MSH2-MSH3 [Flores-Rozas and Kolodner, 1998; Lipkin *et al*., 2000a]. Another human MutL homologue, PMS1, has been suggested to also have a role in mismatch repair, but a biochemical role for this protein has not been demonstrated in mismatch repair [Marsischky *et al*., 1996]. Lipkin *et al*. [2000] suggested that PMS1, PMS2 and MLH3 may exhibit functional redundancy.
In addition to its indispensable role in removing misincorporations, MMR is also involved in determining cellular sensitivity to a number of DNA damaging agents. Some of these agents are used as chemotherapeutic drugs (procarbazine, dacarbazine and temozolomide methylate S_N1 DNA, cisplatin is cross-linking, and azathioprine is an immunosuppressant) (for

recent reviews on this subject see [Buermeyer *et al*., 1999; Schofield and Hsieh, 2003; Stojic *et al*., 2004]).

Disease Associated Mutations

DNA variations that result in premature termination of translation, and thus in loss-of-function, give rise to HNPCC and are thought to be pathogenic mutations. For DNA variants that do not lead to such a premature termination of translation, this is however less straightforward. This is certainly the case for DNA variants that give rise to single amino acid substitutions (~10% of MSH2, ~30% of MLH1, and ~37% of MSH6 variants), or small- or large in-frame deletions (~10% of both MSH2 and MLH1 variants see http://www.insight-group.org). These DNA variations, with an unclear pathogenic nature, are often called unclassified variants (UVs).

Whether an UV contributes to the disease phenotype or merely represents a rare polymorphism constitutes a major problem with direct clinical consequences for HNPCC patients [Syngal *et al*., 1999; Cravo *et al*., 2002]. Pathogenic mutation carriers have a high cancer risk and should therefore be monitored regularly to detect neoplasias at the earliest possible stages. Presymptomatic testing can be offered to all (adult) family members who carry pathogenic mutations and those who are not carriers can be released from cancer surveillance programs.

Several criteria [Hofstra *et al*., 1997; Cotton and Scriver, 1998; Syngal *et al*., 1999] can be applied, usually in combination, to assess the possible pathogenicity of a UV: (1) *de novo* appearance, (2) segregation with the disease within pedigrees, (3) absence in control individuals, (4) a change of amino acid polarity or size, (5) occurrence of the amino acid change in a domain which is evolutionary conserved between species and/or shared between proteins belonging to the same protein family, and (6) effect in a functional assay or in an animal model. In the case of cancer genetics, we could add: (7) loss of the non-mutated allele by a large deletion (loss-of-heterozygosity, LOH), and (8) loss of protein expression. In the case of HNPCC, microsatellite instability of the tumour is sometimes used as well. Previous inclusion of the UV in the increasing number of disease-specific mutation databases is also often considered a valid argument for pathogenicity but, since most databases do not add any information besides the fact that the mutation was found in a patient, we do not recommend using this criterion. Several of the above criteria can be difficult to score and some may be unreliable [Hofstra *et al*., 1997]. For instance point (3), the mutation segregating with the disease can be caused by the UV being in linkage disequilibrium to the true, as yet unidentified mutation. LOH and negative immunohistochemistry and, in the case of HNPCC, microsatellite instability do not provide direct evidence, and again the results obtained might be due to an unidentified mutation in the same gene

or in its regulatory sequences.
We therefore believe that we should look for final proof via criterion 6, the effect of a mutation in a functional assay or in an animal model.
Our aim was to review the functional assays reported for the HNPCC-associated MMR proteins and to discuss the outcomes of these tests on UVs found in patients diagnosed with HNPCC or suspected of having it.

Functional Assays for MMR Proteins

Most functional assays are designed to measure a specific effect that potentially causative mutations (UVs) may have on the biological or biochemical function of the mutated MMR protein.
A properly functional MMR system needs to perform a wide range of steps: produce proteins; transport to the nucleus; bind different protein-complexes at the site of the DNA mismatch all of which may result in repaired mismatches. It should therefore not come as a surprise that many of these steps have been made the focus of functional analysis. Such assays determine the capacity of two MMR-proteins to form protein-protein complexes (dimer formation) and commonly used assays include pull-down and yeast two-hybrid assays, or transient MMR gene expression in relevant MMR-deficient cell lines to determine stabilization of the endogenous or exogenous binding partner [Raevaara *et al*., 2005„]. Some assays have also been applied to reveal ADP↔ATP cycling by MSH protein dimers [Heinen *et al*., 2002]. It is also possible that intracellular trafficking or the subcellular localization of the mutated MMR protein is disturbed, so that determining the subcellular localization of the mutant protein in the cell seems an obvious course of action [Nielsen *et al*., 2004; Brieger *et al*., 2005].
Besides testing distinct properties of the repair proteins, it seems logical to determine whether the mutated protein can restore repair capacity in MMR-deficient cells or cell extracts. Assays have been developed in which MMR-deficient yeast or human cells were complemented with an exogenous expressed MMR gene or genes and the overall repair capacity could be measured [Holmes, Jr. *et al*., 1990; Drotschmann *et al*., 1999; Trojan *et al*., 2002]. Table 1 shows a classification of the MMR functional assays.

Table 1. Classification of Mismatch Repair (MMR) Functional Assay

Classification of functional assays	Subclassification	Specific functional assay [#]	Feature	Disadvantages
Group I Tests for a specific biological or biochemical function of a MMR protein	Measures the capacity of two MMR proteins binding to each other (heterodimer formation)	Yeast two-hybrid assay (**IA**) [#]	Measures the physical interaction (*in vivo*) between two MMR proteins in yeast	Tests only one specific feature of the MMR protein (heterodimer formation) Tests only mutations involved in the physical interaction False-positive or -negative problem
		Glutathione S-transferase (GST) pull-down assay (**IB**)	Measures the physical interaction (*in vitro*) between two MMR proteins	Tests only one specific feature of the MMR protein Tests only mutations involved in the physical interactions False-negative results
		Transient expression of MMR genes in relevant MMR deficient cell line (**IC**)	Measures protein stability (*in vivo*) between two MMR proteins	Tests only one specific feature of the MMR protein (stability) High concentrations of protein might be toxic
	Heterodimer binding to mismatched DNA	DNA mobility shift assays (gel shift assay) (**ID**)	Measures the capacity of two MMR proteins (a heterodimer) to recognize and bind mismatched DNA (*in vitro*)	Tests only one specific feature of the MMR protein (DNA binding) The label for visualization could potentially interfere with the DNA-protein interactions
		Binding to immobilized DNA (**1D**)	Measures the capacity of MMR heterodimers to recognize and bind mismatched DNA on the surface of a biosensor *(in vitro)*	Tests only one specific feature of the MMR protein (DNA binding)
	ATP↔ADP cycling	ATP binding, ATP→ADP exchange, ATP-induced dissociation from a mismatch, etc.	Tests ATPase catalytic efficiency, ATP binding, ATP→ADP exchange, conformational change in the presence of ATP and efficiency of loading multiple sliding clamps on to a circular DNA strand	Only tests on the MutS proteins (restricted to MSH2)
	MMR protein subcellular localization	Localization experiments (**IF**)	Expression of fluorescent MMR proteins in mammalian cells to localize the distribution of these proteins in the cell	Overexpression of MMR proteins might interfere with proper localization Overexpression of MMR proteins might be toxic
Group I Tests for MMR repair capacity as a complete process	Functional assays using yeast **(IIA**)	Expression of mutant yeast MMR genes in haploid yeast strains (**IIA-1**)	Expression of mutant human or corresponding yeast MMR in yeast strains to monitor the repair capacity as a whole (based on the homology of human and yeast MMR proteins)	Only tests mutations occurring in evolutionary conserved regions between human and yeast The use of a heterologous system may result in artifacts
		Expression of mutant human MMR genes in haploid yeast strains (**IIA-2**)		
		Expression of mutant yeast MMR genes in diploid yeast strains (**IIA-3**)		
	Functional assays using cell-free systems	Cell-free *in vitro* MMR assays (**IIB**)	Tests in vitro the repair of mismatched DNA by protein extracts. Mostly baculovirus infected insect cell extracts are used to complement MMR-deficient cell extract	The ratio between the substrate and the proteins added to *in vitro* MMR assay is critical. Does not measure defects in splicing, protein folding and stability, intracellular localization, etc.
	Functional assays using human cell lines	Cell based *in vitro* MMR functional assay using a human expression system (**IIC**)	Expression of mutant human MMR genes in homologous human cell lines to monitor the repair capacity as a whole	Human cell lines are sometimes difficult to transfect Expression level (and ratios) are critical May not be sensitive to defects in splicing, intracellular localization, etc. Overexpression of MMR proteins might be toxic

Group I. Tests for a specific biological or biochemical function of an MMR protein

Yeast Two-Hybrid Assay (referred to as Assay IA in Tables 1-5)

The yeast two-hybrid assay can be used to measure the physical interaction between two proteins. In this assay one of the proteins of interest (*e.g.* MSH2) is fused to a DNA-binding domain of a transcription factor (*e.g.* GAL4), whereas its interacting partner (*e.g.* MSH6) is fused to the DNA activation domain of the same transcription factor. If the two proteins of interest interact, the DNA binding domain and activation domain are brought into each other's proximity upon which they are able to activate the transcription of a reporter gene or genes (*e.g.* His3, or LacZ). In 2003, for instance, Kondo *et al*. [2003,] used this assay to show that 18 of 23 (78%) of the tested MLH1 missense mutations disrupt proper protein binding and might therefore be considered as causative for HNPCC.

GST Pull-Down Assay (Assay IB)

The glutathione S-transferase (GST)-fusion interaction (pull-down) assay is a different approach to measuring the physical interaction between two proteins. It is an *in vitro* technique that consists of a fusion-tagged 'bait' protein for which a binding partner (*i.e*. the 'prey') is being sought. In most cases, a GST-tagged bait protein is bound to an immobilized glutathione support, after which a labeled potential interacting partner, the prey (*in vitro* transcribed and translated) is added. If interaction occurs, the labeled protein can be visualized. Guerrette *et al*. [1999] used this assay to show that 9 of 11 (82%) MLH1 UVs displayed reduced binding to PMS2.

Both the yeast two-hybrid assays and the GST pull-down assays are based on the physical interaction between two MMR proteins, such as MSH2 and MSH6. Both systems can only analyze variants that interfere with proper binding and will not, therefore, detect all the possible functional defects of the mutant protein being investigated.

Transient Expression of MMR Genes in Relevant MMR-Deficient Cell Lines (Assay IC)

Some of the MMR proteins need binding with an MMR protein partner to stabilize their expression. Brieger *et al*. [2002] showed that the MSH6 protein became detectable in LOVO cells (cells carrying homozygous mutations in MSH2 resulting in destabilization of MSH6), after transfection with wild-type MSH2. In contrast, the MSH6 protein was not detectable in LOVO cells when some mutant MSH2 proteins were expressed (like hMSH2 Cys697Arg). These data indicate a defect in MSH2-MSH6 interaction due to the MSH2 UVs.

DNA Mobility Shift Assays (Assay 1D)

The capacity of MSH-dimers (MSH2-MSH6 or MSH2–MSH3) to bind to mismatched DNA can be measured directly by using an affinity sensor or by electrophoretic mobility shift assay (EMSA). This assay is based on the observation that complexes of protein and mismatch-containing DNA oligonucleotides migrate through a non-denaturing polyacrylamide gel more slowly than free oligonucleotides. EMSA has been used to test the binding capacity to mismatched oligonucleotides of MutSα proteins extracted from baculovirus or from yeast strains overexpressing these genes, with or without a missense mutation [Clark *et al*., 1999; Drotschmann *et al*., 1999]. Alternatively, a biosensor can be used to measure mismatch binding of MSH2 UVs [Heinen *et al*., 2002].

The MutS ADP-ATP Cycle (Assay 1E)

Apart from testing the capacity of the heterodimer to bind mismatched DNA, assays for investigating other biochemical behaviors of MutSα mutants have been set up. An example is the mismatch binding-dependent ADP→ATP exchange which triggers a conformational change of MutS, resulting in a so-called 'sliding clamp' on the DNA. Heinen *et al*. [Heinen *et al*., 2002] examined seven missense mutations for mismatch binding (see 1D) and for: (a) intrinsic ATPase activity, (b) ATP-binding activity, (c) ADP→ATP exchange, and (d) the efficiency of loading multiple MutS heterodimers on to circular DNA as a measure of sliding clamp formation. These assays showed that 6 of the 7 missense mutations cause a reduced molecular switch function of the mutant MSH2-MSH6 compared with the wild-type MSH2-MSH6 complex [Gradia *et al*., 1997; Fishel, 2001].

Protein Localization Experiments (Assay IF)

Mutations in MMR proteins can affect other functional aspects of the proteins, including transcription, splicing and translation, protein stability, post-translational modification, and subcellular localization. Clearly, MMR proteins need to be transported to the nucleus to reach their target, the damaged DNA. The regulation of the subcellular localization of MMR proteins, although important for their function, is still poorly understood. It has been shown that the regulation of MMR protein localization depends on (a) specific sequences in the MMR proteins that interact with factors known to be involved in the nuclear import of proteins; as well as on (b) sequences involved in translocation of the protein to the nucleus, and on (c) specific protein-protein interactions in the MMR complexes [Brieger *et al*., 2005]. Subcellular localization of mutant MMR proteins can be investigated by immunostaining or by expressing fluorescent fusion proteins in mammalian cells and subjecting them to fluorescent microscopy. Raevaara *et al*. [Raevaara *et al*., 2005] found that an MLH1 mutant (D63E) was indeed unable to translocate to the nucleus.

Group II. Tests of MMR repair capacity as a complete process

Functional Assays using Yeast (Assays IIA1-IIA3)

Yeast-based functional assays have been developed based on the fact that the MMR system is evolutionary conserved. By replacing an evolutionary conserved amino acid with its mutated counterpart, we should be able to find out whether this replacement indeed has functional consequences. The system has been used mainly for analyzing mutations in MLH1, but there are limits to its use. As the human and the yeast MLH1 protein only share significant homology in the N- and C-termini, only mutations located in these regions can be analyzed. At least three different strategies have been used: (a) expression of the mutant yeast MMR gene in a haploid yeast strain deficient for the same MMR gene; (b) expression of a human MMR gene in a haploid WT yeast strain: this approach makes use of the dominant negative effect that wild-type human MMR proteins have in these cells; and (c) expression of a mutant yeast MMR gene in a diploid yeast strain containing only one functional allele of the same MMR gene. Although defects found in yeast-based functional assays are indicative of an UV's pathogenic nature, we cannot excluded the fact that functional differences between yeast and human MMR may cause artificial results.

a. Expression of Mutant Yeast MMR Genes in Haploid Yeast Strains (Assay IIA-1)

Shcherbakova *et al*. [1999] introduced six missense mutations into the yeast mlh1 gene. All six were analogous to mutations identified in HNPCC patients. The mutants were integrated and expressed in a wild-type haploid strain (by homologous recombination), and when the spontaneous mutation rates of the mlh1 mutants were compared with a mlh1-deleted mutant they showed a mutation rate similar to that of the mlh1-deleted yeast haploid strain, for all six mutants. Shcherbakova *et al*. therefore concluded that all six MLH1 missense mutations were likely to be pathogenic.

b. Expression of Mutant Human MMR Genes in Haploid Yeast Strains (Assay IIA-2)

Clark *et al*. [2000] examined the mutation rate in an msh2 mutant yeast strain. They showed that expression of the human MMR proteins, alone or in combination, did not reduce the mutation rate in this mutated yeast strain. Thus, the human genes are not functional in yeast. Furthermore, expressing the individual human proteins in a wild-type yeast strain did not induce an increased mutation rate. On the other hand, expressing both wild-type human MSH2 and MSH6 in a wild-type yeast stain did increase the mutation rate enormously (4000-fold). Co-expression of human MSH2 and MSH3 elevated the mutation rate (only) 5-fold, reflecting the minor role of the latter complex in MMR. The authors showed that the human MMR proteins bind to mismatched DNA, suggesting that *in vivo* the human

MutSα and MutSβ complexes compete with the yeast MMR heterodimers, thereby preventing correction of replication errors. This is called a 'dominant mutator effect'.
In another study Shimodaira *et al*. [1998] showed that human MLH1 in MMR-proficient yeast behaved differently to the human MSH2 described by Clark *et al*. [2000], as human MLH1 on its own was sufficient to cause a dominant mutator phenotype. Both authors showed that certain MLH1 missense mutations identified in HNPCC patients, could partially or completely abolish the dominant negative mutator effect indicative of the pathogenic nature of the mutant MLH1 genes.

c. Expression of Mutant Yeast MMR Genes in Diploid Yeast Strains (Assay IIA-3)

Drotschmann et al. [1999] showed that in yeast the combination of any of five mutant yeast *MSH2* genes (under their natural promoter) and a wild-type yeast *MSH2* gene resulted in a mutator phenotype. This likely reflects a dominant mutator effect exerted by the mutant MutS dimers: as an example, introduction of the *MSH2* G693A mutant into a diploid yeast strain containing only one functional *MSH2* allele dramatically elevated the mutation rate (74-fold). The authors concluded that this result might explain why the corresponding heterozygous mutation in this patient was associated with early onset of colon cancer.
Several methods have been applied to evaluate the mutation rates in yeast strains containing mutant MMR proteins, most of which make use of reporter genes. These reporter genes serve as selection markers; when mutated they lose or regain their function and can be used to determine the MMR capacity of the added mutant MMR protein [Tran *et al*., 1997; Polaczek *et al*., 1998; Shimodaira *et al*., 1998].

Cell-Free *in vitro* MMR Assays (Assays IIB)

To measure the DNA repair capacity of a mutant MMR protein, protein extracts from a MMR defective cell are complemented with the missing (mutant) MMR protein [Holmes, Jr. *et al*., 1990; Li and Modrich, 1995]. Repair of artificial substrates containing a mismatch can be determined by the use of appropriate restriction enzymes that uniquely cleave repaired molecules. Optimal repair is seen when both proteins of a heterodimer complex are added. Li *et al*. [1995], for instance, showed that protein extracts of H6 cells that were deficient for MLH1 could regain their repair function when complemented with the human MutL homologues MLH1 and PMS2. Another example was reported by Nystrom-Lahti *et al*. [2002] who tested several MLH1 mutations found in HNPCC patients. The proteins were expressed in SF9 insect cells transfected with recombinant baculoviruses encoding mutant MLH1 cDNAs, followed by testing of the functional activity of the mutant proteins *in vitro*. They found that all the

mutants but one showed loss-of-function. One missense mutation MLH1 S93G was still functional in this *in vitro* MMR assay, although the missense mutation segregated with the disease. The latter mutant may be defective in processes like mRNA splicing or intracellular transport of the protein, but these processes are not assayed in the procedure used.

Cell-Based *In vitro* MMR Functional Assay using a Human Expression System (IIC)

A functional assay described by Trojan *et al*. [2002] is based on transient transfection of a MLH1 mutant together with wild-type PMS2 in the (human) 293T cell line, a cell line that is MLH1-deficient due to *MLH1* promoter methylation (resulting in degradation of endogenous PMS2). Extracts from transfected cells were incubated with bacteriophage M13 DNA containing a mismatch in the LacZ α-complementation domain. After incubation, these heteroduplexes were introduced into suitable *E. coli* cells and their repair efficiency was measured by the number and color of the resulting plaques. This system enabled pathogenic mutations in MLH1 to be detected.

If we compare functional assays based on the use of human cell lines and human proteins with those that are yeast-based, the human systems have several advantages: they permit the study of protein/protein interaction of any two known components of the human MMR system, and testing of all mutations, irrespective of their evolutionary conservation status [Nystrom-Lahti *et al*., 2002] or position in the protein [Shimodaira *et al*., 1998; Nystrom-Lahti *et al*., 2002]. Human systems can give answers to questions of not only protein interaction but also on the repair procedure as a whole [Nystrom-Lahti *et al*., 2002].

Functionally Tested MMR UVs

The mutation database set up by the International Society for Gastrointestinal Hereditary Tumours (www.insight-group.org) gives an overview of most of the MMR mutations published. It contains information on more than 400 mutations found in over 700 HNPCC kindreds [Peltomaki and Vasen, 2004]. However, the database does not provide details about any functional assays that may have been performed. Particularly we need additional information for the UVs to be able to decide whether or not these are pathogenic. We therefore decided to compile a database containing all the available data on published MMR-UVs that have been functionally tested. Tables 2-5 show the UVs and functional assays performed for human MLH1, MSH2, MSH6 and PMS2, respectively.

Several points should be kept in mind: (1) There is a clear bias in the mutations tested. The majority of the tested mutations are missense mutations that were identified in patients with an MSI-H tumour. (2)

Sometimes clinical data (and MSI data) are not given, which makes genotype-phenotype correlations impossible. It is therefore impossible to answer questions such as whether pathogenic missense mutations always give rise to MSI-H tumours. (3) The type of functional assays used has changed over the years. Before 1999 almost all functional assays were yeast-based or used pull-down assays, and most of them were used to study only MLH1 and MSH2 variants. After 1999 more and more publications describe *in vitro* MMR assays or (human) expression systems for analyzing repair capacity as well as the subcellular localization of the mutant proteins [Trojan *et al*., 2002; Raevaara *et al*., 2005]. This trend reflects the notion that mismatch repair is involved in different biochemical events and a good functional assay should permit the whole repair process to be measured, including mechanisms such as gene expression, mRNA splicing, protein stability and protein localization [Trojan *et al*., 2002; Raevaara *et al*., 2005]. Since 2002 MSH6 UVs have also been analyzed.
From the studies presented in Tables 2, we can conclude that a high percentage of tested UVs is likely to be pathogenic. For MLH1 49/70 (70%) of the UVs were considered to be pathogenic mutations in different functional assays. For MSH2 this was 25/35 (71%), while for MSH6 it was only 1/8 (13%). However, we cannot exclude the fact that these percentages may be skewed due to the ascertainment bias described.

Table 2: Percentage of UVs considered pathogenic after functional testing

Gene	Total number of UVs tested	Outcome functional assays		% pathogenic UVs
		Non-pathogenic	Pathogenic	
MLH1	70	21	49	70%
MSH2	35	10	25	71%
MSH6	8	7	1	13%
PMS2	3		3	100%

Although most data are consistent, there are some discrepancies between the results reported from different functional assays. For example, K618A (MLH1) (expressed *in vitro*) was examined for its ability to interact with PMS2 in a GST pull-down experiment [Guerrette *et al*., 1999]. It was shown that this UV gave rise to a significant loss-of-interaction (>85%). But the same UV was tested by another group [Raevaara *et al*., 2005] using a cell-free *in vitro* MMR system and they found that it behaved similarly to wild-type MLH1. Besides this UV, there were six other missense mutations for which discrepancies between assays were reported (see Supplementary Tables S1-S4, contradictory results marked with **; available on line at

http://www.interscience.wiley.com/jpages/1059-7794/suppmat. It is interesting to note that three out of the seven contradictions showed loss-of-function in a protein-protein interaction assay (IB), or functional assays using yeast (IIA), but no loss-of-function in cell-free *in vitro* MMR assays (IIB). As the repair function seems to be intact, we could argue that this assay would likely be reliable. However, although *in vitro* MMR assays analyze the repair process as a whole, the ratio between the substrate and the proteins added to the *in vitro* MMR assay is critical and such a conclusion may not be justified [Trojan *et al*., 2002].
Although there may be some bias in the data we chose to include, the tables clearly show that missense mutations do play an important role in HNPCC. Apart from this being important news for patients and their families, it should also make us reconsider the inclusion criteria for mutation analysis, as some of them are based on finding truncating mutations. For example, to date, we generally consider a combination of a positive family history, an MSI-H status, and negative staining for an MMR protein in immunohistochemical analysis as predictors and inclusion criteria for MMR mutation analysis [Reyes *et al*., 2002; Umar *et al*., 2004]. Clearly, missense mutation carriers would be excluded as the immunohistochemical staining would very likely be positive. Furthermore, some publications have shown that certain missense mutations (like E578G in MLH1) do not correlate with MSI [Liu *et al*., 1999].
A database containing detailed information on all the UVs tested can be found on http://www.mmrmissense.org.

Limitations of Functional Assays

All the assays described here investigate aspects of MMR protein function, but UVs that do not cause any functional defect in the protein may still cause MMR defects, *e.g.*, as a consequence of aberrations in mRNA splicing or stability [McVety *et al*., 2006]. Thus, we must emphasize that even if no obvious defects are found using a functional assay, pathogenicity can still not be excluded. Inother words, UVs that do not display a clear functional defect should not directly be considered to be polymorphisms.
Also, when assay results predict a pathogenic nature of a variant, one should realize that it is of the utmost important that the proper positive and negative controls were included in the study. Besides these controls the expression level of the transfected proteins should also be kept in mind as high levels of these proteins might be toxic or could disturb normal functioning of the overexpressed protein. Further, when only part of the protein is expressed, this might also lead to findings that do not reflect reality (which is not the case in the assays described here).

Possibly the biggest problem of functional assays is the fact that the sensitivity and specificity has been determined for very few of the assays. As long as a test is not validated, results should be utilized with great caution and one might even argue that functional assays cannot be used to classify variants with any confidence in the absence of other (genetic) data. Finally, as already mentioned, DNA variants in the coding sequence might give rise to aberrant splicing as previously shown and this should be kept in mind (*e.g.*, as described by Auclair *et al*. [2006] and Pagenstecher *et al*. [2006]).These limitation clearly show that refinement of the functional assays are necessary to solve these dilemmas.

Conclusion

How to best screen the UVs for function is still under debate. At the moment, yeast and human cell systems are used, and different biochemical properties can be determined. Since pathogenic non-truncating alterations in MMR proteins may interfere with different biochemical mechanisms, it seems reasonable to combine different functional assays, especially those employing human homologous expression systems and those that elucidate different MMR functions.

Reference List

Belvederesi L, Bianchi F, Loretelli C, Gagliardini D, Galizia E, Bracci R, Rosati S, Bearzi I, Viel A, Cellerino R, Porfiri E. 2006. Assessing the pathogenicity of MLH1 missense mutations in patients with suspected hereditary nonpolyposis colorectal cancer: correlation with clinical, genetic and functional features. Eur J Hum Genet 14:853-859.

Berends MJ, Wu Y, Sijmons RH, Mensink RG, van der ST, Hordijk-Hos JM, de Vries EG, Hollema H, Karrenbeld A, Buys CH, van der Zee AG, Hofstra RM, Kleibeuker JH. 2002. Molecular and clinical characteristics of MSH6 variants: an analysis of 25 index carriers of a germline variant. Am J Hum Genet 70:26-37.

Brieger A, Plotz G, Raedle J, Weber N, Baum W, Caspary WF, Zeuzem S, Trojan J. 2005. Characterization of the nuclear import of human MutLalpha. Mol Carcinog 43:51-58.

Brieger A, Trojan J, Raedle J, Plotz G, Zeuzem S. 2002. Transient mismatch repair gene transfection for functional analysis of genetic hMLH1 and hMSH2 variants. Gut 51:677-684.

Buermeyer AB, Wilson-Van Patten C, Baker SM, Liskay M. 1999. The human MLH1 cDNA complements DNA mismatch repair defects in Mlh1-deficient mouse embryonic fibroblasts. Cancer Research 59:538-541.

Chung DC, Rustgi AK. 2003. The hereditary nonpolyposis colorectal cancer syndrome: genetics and clinical implications. Ann Intern Med 138:560-570.

Clark AB, Cook ME, Tran HT, Gordenin DA, Resnick MA, Kunkel TA. 1999. Functional analysis of human MutS alpha and MutS beta complexes in yeast. Nucleic Acids Research 27:736-742.

Clark AB, Valle F, Drotschmann K, Gary RK, Kunkel TA. 2000. Functional interaction of proliferating cell nuclear antigen with MSH2-MSH6 and MSH2-MSH3 complexes. J Biol Chem 275:36498-36501.

Cotton RG, Scriver CR. 1998. Proof of "disease causing" mutation. Hum Mutat 12:1-3.

Cravo M, Afonso AJ, Lage P, Albuquerque C, Maia L, Lacerda C, Fidalgo P, Chaves P, Cruz C, Nobre-Leitao C. 2002. Pathogenicity of missense and splice site mutations in hMSH2 and hMLH1 mismatch repair genes: implications for genetic testing. Gut 50:405-412.

Drotschmann K, Clark AB, Kunkel TA. 1999a. Mutator phenotypes of common polymorphisms and missense mutations in MSH2. Curr Biol 9:907-910.

Drotschmann K, Clark AB, Tran HT, Resnick MA, Gordenin DA, Kunkel TA. 1999. Mutator phenotypes of yeast strains heterozygous for mutations in the MSH2 gene. Proc Natl Acad Sci U S A 96:2970-2975.

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

Flores-Rozas H, Kolodner RD. 1998. The *Saccharomyces cerevisiae* MLH3 gene functions in MSH3-dependent suppression of frameshift mutations. Proc Natl Acad Sci U S A 95:12404-12409.

Gradia S, Acharya S, Fishel R. 1997. The human mismatch recognition complex hMSH2-hMSH6 functions as a novel molecular switch. Cell 91:995-1005.

Guerrette S, Wilson T, Gradia S, Fishel R. 1998. Interactions of human hMSH2 with hMSH3 and hMSH2 with hMSH6: examination of mutations found in hereditary nonpolyposis colorectal cancer. Mol Cell Biol 18:6616-6623.

Guerrette S, Acharya S, Fishel R. 1999. The interaction of the human MutL homologues in hereditary nonpolyposis colon cancer. J Biol Chem 274:6336-6341.

Heinen CD, Wilson T, Mazurek A, Berardini M, Butz C, Fishel R. 2002. HNPCC mutations in hMSH2 result in reduced hMSH2-hMSH6 molecular switch functions. Cancer Cell 1:469-478.

Hendriks YM, Jagmohan-Changur S, van der Klift HM, Morreau H, van Puijenbroek M, Tops C, van Os T, Wagner A, Ausems MG, Gomez E, Breuning MH, Brocker-Vriends AH, Vasen HF, Wijnen JT. 2006. Heterozygous mutations in PMS2 cause hereditary nonpolyposis colorectal carcinoma (Lynch syndrome). Gastroenterology 130:312-322.

Hofstra RM, Osinga J, Buys CH. 1997. Mutations in Hirschsprung disease: when does a mutation contribute to the phenotype. Eur J Hum Genet 5:180-185.

Holmes J, Jr., Clark S, Modrich P. 1990. Strand-specific mismatch correction in nuclear extracts of human and Drosophila melanogaster cell lines. Proc Natl Acad Sci U S A 87:5837-5841.

Iaccarino I, Marra G, Dufner P, Jiricny J. 2000. Mutation in the magnesium binding site of hMSH6 disables the hMutSalpha sliding clamp from translocating along DNA. J Biol Chem 275:2080-2086.

Jager AC, Rasmussen M, Bisgaard HC, Singh KK, Nielsen FC, Rasmussen LJ. 2001. HNPCC mutations in the human DNA mismatch repair gene hMLH1 influence assembly of hMutLalpha and hMLH1-hEXO1 complexes. Oncogene 20:3590-3595.

Jiricny J. 2006. The multifaceted mismatch-repair system. Nat Rev Mol Cell Biol 7:335-346.

Kariola R, Raevaara TE, Lonnqvist KE, Nystrom-Lahti M. 2002a. Functional analysis of MSH6 mutations linked to kindreds with putative hereditary non-polyposis colorectal cancer syndrome. Hum Mol Genet 11:1303-1310.

Kolodner RD, Tytell JD, Schmeits JL, Kane MF, Gupta RD, Weger J, Wahlberg S, Fox EA, Peel D, Ziogas A, Garber JE, Syngal S, Anton-Culver H, Li FP. 1999a. Germ-line msh6 mutations in colorectal cancer families. Cancer Res 59:5068-5074.

Kondo E, Suzuki H, Horii A, Fukushige S. 2003. A Yeast Two-Hybrid Assay Provides a Simple Way to Evaluate the Vast Majority of hMLH1 Germ-Line Mutations. Cancer Res 63:3302-3308.
Li GM, Modrich P. 1995. Restoration of mismatch repair to nuclear extracts of H6 colorectal tumor cells by a heterodimer of human MutL homologs. Proc Natl Acad Sci U S A 92:1950-1954.

Lipkin SM, Wang V, Jacoby R, Banerjee-Basu S, Baxevanis AD, Lynch HT, Elliott RM, Collins FS. 2000. MLH3: a DNA mismatch repair gene associated with mammalian microsatellite instability. Nat Genet 24:27-35.

Liu T, Tannergard P, Hackman P, Rubio C, Kressner U, Lindmark G, Hellgren D, Lambert B, Lindblom A. 1999. Missense mutations in hMLH1 associated with colorectal cancer. Hum Genet 105:437-441.

Lynch HT, Lemon SJ, Karr B, Franklin B, Lynch JF, Watson P, Tinley S, Lerman C, Carter C. 1997. Etiology, natural history, management and molecular genetics of hereditary nonpolyposis colorectal cancer (Lynch syndromes): genetic counseling implications. Cancer Epidemiol Biomarkers Prev 6:987-991.

Marsischky GT, Filosi N, Kane MF, Kolodner R. 1996. Redundancy of *Saccharomyces cerevisiae* MSH3 and MSH6 in MSH2-dependent mismatch repair. Genes Dev 10:407-420.

McVety S, Li L, Gordon PH, Chong G, Foulkes WD. 2006. Disruption of an exon splicing enhancer in exon 3 of MLH1 is the cause of HNPCC in a Quebec family. J Med Genet 43:153-156.

Nielsen FC, Jager AC, Lutzen A, Bundgaard JR, Rasmussen LJ. 2004. Characterization of human exonuclease 1 in complex with mismatch repair proteins, subcellular localization and association with PCNA. Oncogene 23:1457-1468.

Nystrom-Lahti M, Perrera C, Raschle M, Panyushkina-Seiler E, Marra G, Curci A, Quaresima B, Costanzo F, D'Urso M, Venuta S, Jiricny J. 2002. Functional analysis of MLH1 mutations linked to hereditary nonpolyposis colon cancer. Genes Chromosomes Cancer 33:160-167.

Ollila S, Sarantaus L, Kariola R, Chan P, Hampel H, Holinski-Feder E, Macrae F, Kohonen-Corish M, Gerdes AM, Peltomaki P, Mangold E, de la Chapelle A, Greenblatt M, Nystrom M. 2006. Pathogenicity of MSH2 Missense Mutations Is Typically Associated With Impaired Repair Capability of the Mutated Protein. Gastroenterology 131:1408-1417.

Pang Q, Prolla TA, Liskay RM. 1997. Functional domains of the *Saccharomyces cerevisiae* Mlh1p and Pms1p DNA mismatch repair proteins and their relevance to human hereditary nonpolyposis colorectal cancer-associated mutations. Mol Cell Biol 17:4465-4473.

Peltomaki P, Vasen H. 2004. Mutations associated with HNPCC predisposition -- Update of ICG-HNPCC/INSiGHT mutation database. Dis Markers 20:269-276.

Polaczek P, Putzke AP, Leong K, Bitter GA. 1998. Functional genetic tests of DNA mismatch repair protein activity in *Saccharomyces cerevisiae*. Gene 213:159-167.

Raevaara TE, Korhonen MK, Lohi H, Hampel H, Lynch E, Lonnqvist KE, Holinski-Feder E, Sutter C, McKinnon W, Duraisamy S, Gerdes AM, Peltomaki P, Kohonen-Ccorish M, Mangold E, Macrae F, Greenblatt M, de la Chapelle A, Nystrom M. 2005. Functional significance and clinical phenotype of nontruncating mismatch repair variants of MLH1. Gastroenterology 129:537-549.

Reyes CM, Allen BA, Terdiman JP, Wilson LS. 2002. Comparison of selection strategies for genetic testing of patients with hereditary nonpolyposis colorectal carcinoma: effectiveness and cost-effectiveness. Cancer 95:1848-1856.

Schofield MJ, Hsieh P. 2003. DNA mismatch repair: Molecular mechanisms and biological function. Annual Review of Microbiology 57:579-608.

Shcherbakova PV, Kunkel TA. 1999. Mutator phenotypes conferred by MLH1 overexpression and by heterozygosity for mlh1 mutations. Mol Cell Biol 19:3177-3183.

Shimodaira H, Filosi N, Shibata H, Suzuki T, Radice P, Kanamaru R, Friend SH, Kolodner RD, Ishioka C. 1998. Functional analysis of human MLH1 mutations in *Saccharomyces cerevisiae*. Nat Genet 19:384-389.

Shin KH, Han HJ, Park JG. 1998. Growth suppression mediated by transfection of wild-type hMLH1 in human cancer cells expressing endogenous truncated hMLH1 protein. Int J Oncol 12:609-615.

Stojic L, Brun R, Jiricny J. 2004. Mismatch repair and DNA damage signalling. DNA Repair (Amst) 3:1091-1101.

Syngal S, Fox EA, Li C, Dovidio M, Eng C, Kolodner RD, Garber JE. 1999. Interpretation of genetic test results for hereditary nonpolyposis colorectal cancer: implications for clinical predisposition testing. JAMA 282:247-253.

Tishkoff DX, Boerger AL, Bertrand P, Filosi N, Gaida GM, Kane MF, Kolodner RD. 1997. Identification and characterization of *Saccharomyces cerevisiae* EXO1, a gene encoding an exonuclease that interacts with MSH2. Proc Natl Acad Sci U S A 94:7487-7492.

Tran H, Degtyareva N, Gordenin D, Resnick MA. 1997. Altered replication and inverted repeats induce mismatch repair-independent recombination between highly diverged DNAs in yeast. Mol Cell Biol 17:1027-1036.

Tran PT, Simon JA, Liskay RM. 2001. Interactions of Exo1p with components of MutLalpha in *Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A 98:9760-9765.

Trojan J, Zeuzem S, Randolph A, Hemmerle C, Brieger A, Raedle J, Plotz G, Jiricny J, Marra G. 2002a. Functional analysis of hMLH1 variants and HNPCC-related mutations using a human expression system. Gastroenterology 122:211-219.

Umar A, Boland CR, Terdiman JP, Syngal S, de la Chapelle A, Ruschoff J, Fishel R, Lindor NM, Burgart LJ, Hamelin R, Hamilton SR, Hiatt RA, Jass J, Lindblom A, Lynch HT, Peltomaki P, Ramsey SD, Rodriguez-Bigas MA, Vasen HF, Hawk ET, Barrett JC, Freedman AN, Srivastava S. 2004. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. J Natl Cancer Inst 96:261-268.

Yuan ZQ, Gottlieb B, Beitel LK, Wong N, Gordon PH, Wang Q, Puisieux A, Foulkes WD, Trifiro M. 2002a. Polymorphisms and HNPCC: PMS2-MLH1

protein interactions diminished by single nucleotide polymorphisms. Hum Mutat 19:108-113.

Chapter 4

A Database to Support the Clinical Interpretation of Human Mismatch Repair Gene Variants

Jianghua Ou [1,2], Renée Niessen[1], Robert M.W. Hofstra[1], Rolf H. Sijmons[1]

(1) Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands; (2) Third Department of Tumor Surgery, XinJiang Tumor Hospital, XinJiang Medical University, People's Republic of China

Submitted (on invitation)

Abstract

Germline mutations in the mismatch repair (MMR) *genes MLH1, MSH2, MSH6 or PMS2* can cause Lynch syndrome. This syndrome, also known as hereditary nonpolyposis colorectal cancer (HNPCC), is an autosomal dominantly inherited disorder predominantly characterized by colorectal and endometrial cancer. Truncating MMR gene mutations generally offer a clear handle for genetic counselling and allows for pre-symptomatic testing. In contrast, the clinical implications of most missense mutations and small in-frame deletions detected in patients suspected of having Lynch syndrome are unclear. We have constructed an online database (www.mmrmissense.org) for information on the results of functional assays and other findings that may help in classifying these MMR gene variants. Ideally, such mutations should be clinically classified by a broad expert panel rather than by the individual database curators. In addition, MMR gene mutation databases could be interlinked or combined to increase user-friendliness and avoid unnecessary overlap between them. Within the community of the International Society for Gastrointestinal Hereditary Tumours (InSiGHT), several avenues are now being explored to help reach these goals.

Introduction

Lynch syndrome and the need for missense mismatch repair gene mutation databasing. Germline mutations in the mismatch repair (MMR) genes *MLH1*, *MSH2*, *MSH6* or *PMS2* can cause Lynch syndrome, also known as hereditary nonpolyposis colorectal cancer (HNPCC). Lynch syndrome is an autosomal dominantly inherited cancer syndrome and is associated with a strongly increased risk of developing colorectal and endometrial cancer and, to a lesser extent, a range of other tumours including cancer of the small bowel, stomach, ovaries, renal pelvis, ureter, brain, and sebaceous glands. Lynch syndrome associated tumours are characterized by DNA mismatch repair deficiency which can be demonstrated by the presence of microsatellite instability in the vast majority of cases. Usually, these tumours have physically or functionally lost the wild type MMR allele of the gene mutated in the germline, which can be demonstrated as absence of immunohistochemical staining of the MMR protein in question [1]. The detection of truncating MMR gene mutations in patients suspected of having Lynch syndrome generally offers a good clinical handle for diagnosis and genetic counselling and allows for pre-symptomatic testing in relatives. In contrast, the clinical implications of most missense mutations and small in-frame deletions are unclear (unclassified variants, UVs). A significant proportion of DNA variations found in Lynch syndrome (suspected) patients

are such UVs: 32%, 18%, 38% and 87% for *MLH1*, *MSH2*, MSH6 and MLH3, respectively [2]. Investigators and the clinicians confronted with these UVs would be assisted by online resources that store information on these types of variants. MMR variant databases already exist and the MMR Genes Variant Database (www.med.mun.ca/MMRvariants), the InSiGHT mutation database (www.insight-group.org) and the Human Gene Mutation Database (www.hgmd.org) are probably the most widely consulted for these genes. Their developers and curators do an impressive job of cataloguing published as well as, depending on the database in question, unpublished MMR gene variants. However, these databases do not report in great detail the functional aspects and other findings that may help in clinically classifying these MMR variants. Therefore, we set out to build an online database that particularly focuses on these details.

Methods

We searched the English literature through Entrez PubMed (www.ncbi.nlm.nih.gov/sites/entrez) using sets of keywords to identify publications on the functional analysis of human mismatch repair gene mutations and selected for missense mutations and small in-frame deletions. The reference lists of publications found through this approach were searched for additional relevant papers. Details presented in the selected papers were collected on the type of mutation, the observed clinical phenotype, family history, results of tumour analysis with respect to MSI testing and immunohistochemical staining for the MMR proteins, frequency of the variants detected in controls, segregation of the mutation within the family, evolutionary conservation, type of functional assays used and outcome of the analysis using those assays. Only those mutations with reported results of functional testing were subsequently included in the database. Variant nomenclature was used as proposed by the Human Variome Project (report by Dr.J.den.Dunnen, available at www.hgvs.org/mutnomen). Reference sequences used were as listed in Genbank (www.ncbi.nlm.nih.gov/Genbank) entries NM_000249 for *MLH1*, NM_000251 for MSH2, NM_000179 for *MSH2* and NM_000535 for *MSH6*, respectively. Depending on the information available, each of the variants was labelled with a certain degree of pathogenicity, ranging from non-pathogenic to pathogenic. This labelling was given for research purposes only at this stage, using our expert opinion rather than any formal algorithm. The data thus collected were stored first in an offline Microsoft Access (© Microsoft Corp., Richmond, USA) relational database and subsequently ported to a Microsoft SQL Server (© Microsoft Corp., Richmond, USA) online environment. Interfaces for data retrieval by external users and for

editing the database were programmed using ASP.NET Access (© Microsoft Corp., Richmond, USA).

Results

At the time of submission the database contained information on 69 variants for *MLH1*, 35 for *MSH2*, 8 for *MSH6* and 3 for *PMS2*. The functional studies and their outcome published for these mutations have recently been reviewed in detail elsewhere [3]. Based on the functional data, more than half of these UVs probably is pathogenic, underlining the clinical importance of studying these UVs.
The database can be searched at www.mmrmissense.org through a user interface as shown in figure 1. The results and print lay-out can be sorted in many ways by simply clicking on the column headers. The number of functional tests and clinical reports are shown for each of the individual variants. Details can be displayed by clicking on the “Show Details” button (example shown in figure 2). New data can be added to the database by the curators through a range of editing tools, but also by external users (after a check by the curators) through a special online form.

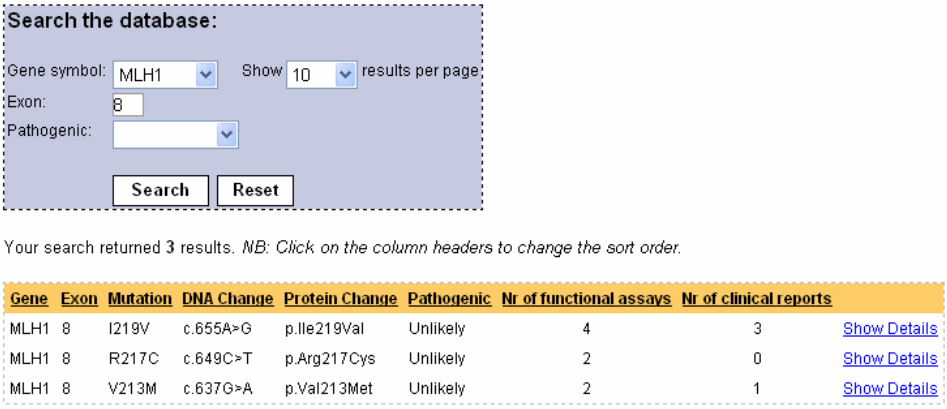


Figure 1. MMR missense database user interface

Figure 2. Mutation details

Details of mutation: P28L

Gene	Mutation	DNA Change	Protein Change	Exon	Position	Change	Classification	Further classification	Pathogenic	Polarity Change	Conservation	Domain
MLH1	P28L	c.83C>T	p.Pro28Leu	1			missense mutation		Likely	No	Yes	

Functional assays

Author	Year	Journal	Type	Methods	Result	Outcome	Comments
Shcherbakova PV	1999	Molecular and Cellular Biology: 3177-3183	functional assays using yeast (IIA)	haploid yeast containing mutation	- increased mutation rate in a lys+ reporter gene compared to WT	pathogenic	compared the rate of Lys+,His+ or Canr in haploid yeast strain containing this mutation with that of wild type or MLH1 deleted mutant ; showed that the mutant retains partial MLH1 function
Raevaara TE	2005	Gastroenterology: 537-549	cell free in vitro MMR assays (IIB)	in vitro MMR assay + Western blot+localization	- reduced repair efficiency compared to WT (2% versus 42%) (G.T heteroduplex) - decreased expression in 293T cells	pathogenic	localization of PMS2 with MLH1: decreased
Kondo E	2003	Cancer Research: 3302-3308	yeast two hybrid assay (IA)	yeast two hybrid assay	- reduced binding to hPMS2	pathogenic	
		UNPUBLISHED	reporter assays in yeast	dominant negative effect	DNE in 1 out of 3 tests	Inconclusive	hMLH1 was transfected in S.Cerevisiae and exerts a dominant negative effect over yeast MLH1. LacZ, GFP and ADE2 reporter assays.
		UNPUBLISHED	qwreqwe	rqwer	rqwer	Inconclusive	rqwre

Clinical reports of this mutation in patients and controls

Author	Year	Journal	Patient	Segregation	Controls	MSI	IHC	Notes	Amsterdam II	Bethesda (revised)
Shcherbakova PV	1999	Molecular and Cellular Biology:3177-3183	no patient information available							
Raevaara TE	2005	Gastroenterology:537-549	all affected patients; mean age:37			high	NA		+	+

Discussion

We have constructed an online database for information on functional assays and other findings that may help in classifying missense variants and small in-frame deletions of the Lynch-syndrome associated MMR genes. It is important that the pathogenicity of these variants is understood for the purpose of genetic counseling and medical management of the families in which these variants have been detected. Classifying these variants is notoriously difficult. Functional assays may produce contradictory results; segregation may not easily be studied due to lack of DNA from affected relatives in Lynch syndrome suspected families; number of control chromosomes analyzed in regionally/ethnically matching populations may be small; and reported clinical and tumour phenotypes may lack detail. For many MMR gene UVs data are incomplete as was recently shown by Lucci-Cordisco *et al.* [4].

Before evidence can be weighed and variants classified, data should be available for that purpose and many challenges from the point of databasing exist in this respect. Keeping up to date with published reports requires effective and efficient literature mining techniques. For this purpose, we are in contact with researchers from the department of Computer Science and Software Engineering at the National ICT Australia Victoria Research Laboratory of the University of Melbourne who may be able to improve our methods of mining relevant literature. Getting unpublished data in a central database is another challenge. Several incentives should be developed to stimulate local laboratories or national laboratory professional societies to submit their data to international online databases, for example inclusion as electronic publication in PubMed for mutations submitted to databases. National, regional or local legislation and rules with respect to privacy and medical confidentiality may prohibit clinicians and DNA laboratories from submitting detailed phenotypic information on their patients and their families with UVs or other types of mutations. Another issue is interpreting the co-occurrence of multiple MMR gene UVs in individual patients. We and others have observed several of these co-occurrences of missense mutations in series of patients tested because of suspected Lynch syndrome. Most information in databases is listed for single variants and adjustments should be made to allow information storage and retrieval for combinations of these variants. Needless to say, functional analysis of these combinations of variants is a challenge in itself.

Methods are needed to weigh available evidence in a structured way. One step in this process would be comparative analysis of the available evidence. For example, computational methods based on comparative

sequence and or protein structure to classify UVs are not necessarily in agreement. Several studies have demonstrated that *in silico* predictions of MLH1 and MSH2 splicing defects can be unreliable and should be complemented by *in vivo* studies (which may even reveal tissue-dependent splicing) whenever possible[5,6]. If the *in silico* predictions are concordant, however, then predictive value is much improved, as has been recently been demonstrated by Chan *et al.* who compared the outcome of these type of computational methods with the outcome of functional analysis for MLH1 and MSH2 [7].

For BRCA1 and 2 variants, Easton *et al*. recently compared personal and family history of cancer, segregation of the variant in families and co-occurrence with known deleterious mutations, with position of the mutation in functional domains, evolutionary conservation and predicted splice site involvement. Variants predicted to be pathogenic based on the clinical data also were likely to have high conservation, more likely to affect splicing and to be located in particular protein domains. [8]. In another comparative study on BRCA1 mutations, Lovelock *et al.* compared the results from a multifactorial likelihood analysis incorporating evolutionary conservation, segregation in families, co-occurrence with known pathogenic mutations and histopathology with the outcome of functional analysis. This likelihood analysis was shown to improve classification of some variants, whereas conflicting results from functional analysis were present in others. [9]. Approaches like these may prove of value in the classification of MMR gene variants. They may for example help in testing and further developing scoring systems like that of Barnetson *et al*. who recently published a system to classify MLH1, MSH2 and MSH6 UVs [10]. These authors noted lack and contradictory results of functional data and discordant *in silico* predictions of effects on splicing and protein function of UVs that had been detected in a large series of colorectal cancer patients. Most weight in their scoring system is attributed to absence of the UVs in controls, segregation in family members, loss of expression of the relevant MMR gene and presence of tumour microsatellite instability in the tumour. Ideally, unclassified MMR gene variants mutations should be clinically classified by a broad expert panel. Such a panel and professional organization behind it might also carry more weight in directing additional research to fill in the data that are needed to classify particular mutations and possibly apply for funding that is needed to support that research.

Lynch syndrome is not a rare disorder and with the increasing availability of mutation analysis a substantial contribution to the already known pool of unclassified MMR gene variants can surely be expected. The need for clinical classification of these variants will increase likewise. Within the community of the International Society for Gastrointestinal Hereditary

Tumours (InSiGHT) there is a strong motivation to address these issues. InSiGHT activities will include interconnecting or integrating the existing MMR gene databases and bringing researchers and clinicians together to explore the possibility of building an expert panel for MMR gene variant classification. These joint activities and their outcome may help those interested in other disorders and genes to develop their own strategies of addressing the problem of unclassified gene variants.

Acknowledgements
We thank Jan Herman Veldkamp (www.aardworm.com) for his assistance in programming the database. We thank Professor Finlay Macrae, secretary of InSiGHT (www.insight-group.org) for his comments on the manuscript. This work was supported by: the Dutch Cancer Society (grant No. RUG2002-2678), the European Community (FP6-2004-LIFESCIHEALTH-5, proposal No 018754).

References

1. Henry T Lynch, C Richard Boland, Gordon Gong, Trudy G Shaw, Patrick M Lynch, Riccardo Fodde, Jane F Lynch and Albert de la Chapelle. Phenotypic and genotypic heterogeneity in the Lynch syndrome: diagnostic, surveillance and management implications. European Journal of Human Genetics 2006;14: 390–402.

2. Päivi Peltomäki and Hans Vasen. Mutations associated with HNPCC predisposition -- Update of ICG-HNPCC/INSiGHT mutation database. Dis Markers 2004;20(4-5):269-76

3. Jianghua Ou, Renée C. Niessen, Anne Lützen, Rolf H. Sijmons, Jan. H. Kleibeuker, Niels de Wind, Lene Juel Rasmussen, Robert M.W. Hofstra. Functional analysis helps to clarify the clinical importance of unclassified variants in DNA mismatch repair genes. Hum Mutat. 2007;28:1047-54

4. Emanuela Lucci-Cordisco, Luigi Boccuto, Giovanni Neri, Maurizio Genuardi. The use of microsatellite instability, immunohistochemistry and other variables in determining the clinical significance of MLH1 and MSH2 unclassified variants in Lynch syndrome. Cancer Biomark. 2006;2(1-2):11-27

5. Jessie Auclair, Marie Pierre Busine, Claudine Navarro, Eric Ruano, Gilles Montmain, Françoise Desseigne, Jean Christophe Saurin, Christine Lasset, Valérie Bonadona, Sophie Giraud, Alain Puisieux,

Qing Wang. Systematic mRNA analysis for the effect of MLH1 and MSH2 missense and silent mutations on aberrant splicing. Hum Mut 2006 27(2),145-154

6. Patrizia Lastella, Nicoletta Concetta Surdo, Nicoletta Resta, Ginevra Guanti and Alessandro Stella. *In silico* and *in vivo* splicing analysis of MLH1 and MSH2 missense mutations shows exon- and tissue-specific effects. BMC Genomics 2006, 7:243

7. Philip A. Chan, Sekhar Duraisamy, Peter J. Miller, Joan A. Newell, Carole McBride, Jeffrey P. Bond, Tiina Raevaara, Saara Ollila, Minna Nyström, Andrew J. Grimm, John Christodoulou, William S. Oetting, Marc S. Greenblatt. Interpreting missense variants: comparing computational methods in human disease genes CDKN2A, MLH1, MSH2, MECP2, and tyrosinase (TYR). Hum Mutat. 2007 Jul;28(7):683-93

8. Douglas F. Easton, Amie M. Deffenbaugh, Dmitry Pruss, Cynthia Frye, Richard J. Wenstrup, Kristina Allen-Brady, Sean V. Tavtigian, Alvaro N. A. Monteiro, Edwin S. Iversen, Fergus J. Couch, and David E. Goldgar. A Systematic Genetic Assessment of 1,433 Sequence Variants of Unknown Clinical Significance in the BRCA1 and BRCA2 Breast Cancer–Predisposition Genes. Am J Hum Genet 2007;81:873-883

9. Paul K Lovelock, Amanda B Spurdle, Myth T S Mok, Daniel J Farrugia, Sunil R Lakhani, Sue Healey, Stephen Arnold, Daniel Buchanan, kConFab Investigators, Fergus J Couch, Beric R Henderson, David E Goldgar, Sean V Tavtigian, Georgia Chenevix-Trench and Melissa A Brown. Identification of BRCA1 missense substitutions that confer partial functional activity: potential moderate risk variants? Breast Cancer Res. 2007 Nov 26;9(6):R82.

10. Rebecca A. Barnetson, Nicola Cartwright, Annelot van Vliet, Naila Haq, Kate Drew, Susan Farrington, Nicola Williams, Jon Warner, Harry Campbell, Mary E. Porteous, Malcolm G. Dunlop. Classification of ambiguous mutations in DNA mismatch repair genes identified in a population-based study of colorectal cancer. Hum Mutat. 2007 Nov 21; [Epub ahead of print]

Chapter 5

Do MLH3 missense mutations contribute to cancer development? A functional study.

Jianghua Ou[1], Merete Rasmussen[2], Helga Westers[1], Sofie D Andersen[2], Paul O Jager[1], Krista K Kooi[1], Renée C Niessen[1], Bart J Eggen[3], Finn C Nielsen[4], Jan H Kleibeuker[5], Rolf H Sijmons[1], Lene J Rasmussen[2] and Robert M W Hofstra[1]

[1]Department of Genetics and [5]Gastroenterology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands; [2]Department of Science, Systems and Models, Roskilde University, Denmark; [3]Developmental Genetics, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Haren, the Netherlands; [4]Department of Clinical Biochemistry, University Hospital of Copenhagen, Denmark.

submitted

Abstract

Background: So far 18 *MLH3* germline mutations/variations have been identified in familial colorectal cancer cases. Sixteen of these variants are amino acid substitutions of which the pathogenic nature is still unclear. These substitutions are known as unclassified variants or UVs.
Aim: To clarify a possible role for eight of these *MLH3* UVs identified in suspected Lynch syndrome patients, we performed functional tests.
Methods and results: Of eight UV-containing MLH3 mutants we determined the protein expression and stability, protein localisation and interaction of the mutant MLH3 proteins with MLH1. All eight MLH3 UVs gave protein expression levels comparable with wildtype MLH3. Furthermore, the UV-containing proteins were all localised normally in the nucleus and they interacted normally with wildtype MLH1.
Conclusion: Our different functional assays yielded no evidence that the eight *MLH3* UVs tested are the cause of hereditary colorectal cancer, including Lynch syndrome.

Introduction

Mismatch repair (MMR) proteins form a highly conserved group of proteins that play a crucial role in correcting DNA mismatches that have escaped the proofreading activity of DNA polymerases. In the human MMR system the mismatch recognition component is fulfilled by a heterodimeric protein complex composed of two MutS homologous (MSH) proteins. The major MSH-heterodimer consists of MSH2 and MSH6 (MUTSα). This heterodimer is able to recognise and bind to base-base mismatches as well as to small insertion/deletion loops. A minor and partially redundant mismatch recognizing/binding heterodimer called MUTSβ consists of the MSH2 and MSH3 proteins. This protein complex recognises and binds mainly to larger insertion/deletion loops (1). During the mismatch repair process, when MUTSα/MUTSβ recognise DNA mismatches that arise during DNA replication, the protein complex binds to the mismatch, thereby inducing a conformational change of the heterodimer and allowing ATP to bind and activate the protein complex. After activation of the MUTS complex, a heterodimer of two MutL homologous (MLH) proteins will bind to the DNA/protein complex. Two heterodimers composed of either MLH1 or PMS2 (MUTLα) or MLH1 and MLH3 (MUTLγ) are proven interactors with the MSH2-containing DNA-protein complex. These MLH complexes are thought to coordinate the downstream repair events, involving different proteins such as exonucleases (*e.g.* EXO1) and DNA polymerases (2-4). The heterodimer MLH1-PMS2 (MUTLα) interacts with both MUTSα and MUTSβ, while the heterodimer MLH1-MLH3 (MUTLγ), appears to interact

with only MSH2-MSH3. In particular, the role of MLH3 in MMR has been under debate as no MMR involvement could be shown in yeast (5). Chen and co-workers (6), however, showed in mice that Mlh3 deficiency causes microsatellite instability, impaired DNA damage response and increased gastrointestinal tumour susceptibility. Furthermore, cultured mammalian cells, stably expressing a dominant negative truncated human MLH3, showed microsatellite instability (MSI) (7). For an excellent review on the MMR process in humans see (8).
Loss of MMR proteins results in the accumulation of unrepaired mutations. It is therefore not unexpected that mutations in these MMR genes are associated with tumour development. Germline mutations in four MMR genes, namely *MLH1*, *MSH2*, *PMS2*, and *MSH6*, have been identified in the majority of families with hereditary nonpolyposis colorectal cancer or Lynch syndrome (9). Many of the mutations identified result in premature termination of translation and thus in loss-of-function of the encoded mutated protein. This loss of MMR function results in unrepaired mutations in non-coding but also in coding sequences. It is mainly these coding sequence mutations that contribute to tumour development (10).
Recently, we identified nine *MLH3* missense mutations and a *MLH3* frameshift mutation in patients suspected of having Lynch syndrome (11). The missense mutations will be called UVs throughout this paper since it is not yet known whether these DNA variants contribute to disease development. Identifying these DNA variants, in combination with the identification of somatic *MLH3* mutations in three tumours of patients with these DNA variants (11), led us to hypothesise that, besides the four known MMR genes, *MLH3* might also play a role in Lynch syndrome development. However, as all but one of the variations identified so far were UVs, the role of *MLH3* in the development of Lynch syndrome is still under debate. We functionally tested eight identified *MLH3* UVs to see whether there is a possible role for *MLH3* in Lynch syndrome.

Material and methods

Cell lines
HEK293T cells, human embryonic kidney cells that lack *MLH1* and *MLH3* expression due to hypermethylation of the promoter regions of both genes (12), were grown in DMEM, supplemented with 10% foetal bovine serum, 1% penicillin-streptomycin and L-Glutamine (all from Invitrogen, Breda, the Netherlands). HeLa cells (ATCC CCL-2) were grown in DMEM containing 10% FBS and 1% penicillin-streptomycin (all from Gibco, Taastrup, Denmark).

In silico analysis of the unclassified variants in MLH3

A description of the different *MLH3* UVs that were tested and the clinical characteristics of the mutation carriers are given in Table 1 and in Figure 1. Alignments with MLH3 homologous from six other vertebrates were obtained by blasting the complete MLH3 protein. The program used was T-Coffee (http://www.tcoffee.org) (13). The following seven sequences were used in the Blast search: *Homo sapiens, Rattus norvegicus*, *Mus musculus*, *Canis familiaris*, *Bos taurus*, *Gallus gallus*, and *Xenopus tropicalis*.

All eight *MLH3* UVs were further analysed *in silico* for putative functional effects using the polymorphism phenotyping (PolyPhen) algorithm (http://genetics.bwh.harvard.edu/pph). PolyPhen is a web-based algorithm that predicts how an amino acid substitution could possibly affect the structure and function of a human protein using straightforward physical and comparative considerations. Calculations are based on amino acid homology. *In silico* analysis using the PolyPhen algorithm was done on the entire MLH3 protein sequence (including the UVs). Amino acid substitutions are reported as benign, possibly damaging or probably damaging.

Three web-based algorithms were used to predict possible splice defects initiated by the UVs: Netgene2 (http://www.cbs.dtu.dk/services/NetGene2), Splicesitefinder (http://violin.genet.sickkids.on.ca/~ali/splicesitefinder.html), and Splicesite predict (http://www.fruitfly.org/seq_tools/splice.html). As input we gave a fragment containing the exon sequence plus 200 nucleotides upstream and 200 base pairs downstream of the exon. This was done for exons 1, 11 and 12 since the eight UVs tested lie within these exons.

Table 1. Genetic and clinical data of the tested *MLH3* UVs.

Exon	Nucleotide Change	Predicted protein change	Additional mutations	Cancer and age of diagnosis	MSI status	IHC MSH2/MLH1 /MSH6
1	c.70C>G	p.Gln24Glu		CRC 50	H	+/+/+
1	c.1496A>G	p.Asn499Ser		CRC 62 EC 63	H	+/+/+
1	c.1870G>C	p.Glu624Gln		CRC 20	ND	+/+/+
1	c.2449A>G	p.Ser817Gly	*MSH6* IVS9+43 ins10bp	EC 46 OV 46	EC:L OV:H	+/+/-
1	c.2533A>G	p.Ser845Gly		EC	L	ND
1	c.2578delA	p.N860IfsX13		CRC 43	L	+/+/+
1	c.2941G>A	p.Gly981Ser		CRC 24	L	+/+/+
1	c.3020A>G	p.Asn1007Ser		CRC 49	L	+/+/+
11	c.4180G>A	p.Ala1394Thr		CRC 44	ND	+/+/+
12	c.4351G>A	p.Glu1451Lys	*MSH6* c.2633T>C	CRC 45	ND	+/+/+
12			*MSH6* c.651dupT	CRC41	ND	+/+/-*

* loss of MSH6 is due to the truncating mutation identified; none of the families fulfilled the Amsterdam II criteria; CRC=colon cancer, EC=endometrial cancer, OV=ovarian cancer; H=MSI-H, L=MSI-L, ND not determined.

We also ran ESEfinder (http://rulai.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi? process=home), a program that searches for sequences that act as binding sites for four members of the serine/arginine rich family of splicing enhancer proteins. Input sequences are screened for consensus binding sequences for the SR proteins CF2/ASF, SC35, SRp40 and SRp55. Regions with scores above a certain threshold value are predicted to act as SR protein binding sites and thus function as ESEs (exonic splice enhancers). The wildtype (or UV-containing) *MLH3* cDNA sequence (AB 039667) was used as input. The program was run in an exon-by-exon manner.

MLH3 vectors and *MLH3* mutants

To clone the *MLH3* cDNA (wildtype) in pAS2 (a yeast two-hybrid vector containing a GAL4 DNA binding domain) and pACT2 (a yeast two-hybrid vector containing a GAL4 activation domain), *MLH3* was first cloned into

pBluescript (a generous gift from Dr. Steven Lipkin, Departments of Biological Chemistry and Medicine, University of California, Irvine, USA). *MLH3* was PCR-amplified using primers containing the 5' end and the 3' end of the coding sequence coupled to a *Bam*HI (forward) and an *Eco*47III (reverse) restriction site. An extra AG was inserted between the *Bam*HI site and the first codon of *MLH3* (to get *MLH3* in the correct reading frame after subsequent subcloning into pACT2 and pAS2). After cloning this PCR-amplified *MLH3* in pBluescript, the insert was sequenced and subcloned in pACT2 and pAS2 using the *Bam*HI and *Eco*47III restriction sites. Mutations (*i.e.* the eight UVs mentioned in Table 1, one known polymorphic missense variant (p.Ser845Gly) and the identified frameshift mutation (see Table 1)) were introduced in these plasmids using the Stratagene QuickChange XL Mutagenesis Kit (La Jolla, CA, USA). This was done according to the manufacturer's instructions. After mutagenesis the *MLH3* inserts were sequenced to confirm the mutation status of the plasmids.
Wildtype *MLH3* (out of pBluescript) was cloned directly, in-frame with YFP, into pEYFPC1 (Clontech Laboratories, Woerden, the Netherlands) and the constructed YFP-MLH3 vector was used for subcellular localisation studies. Mutations (*i.e*. the eight UVs mentioned in Table 1, one known polymorphic missense variant (p.Ser845Gly) and the truncating mutation we identified (see Table 1)) were introduced in WT-YFP-MLH3 using the Stratagene QuickChange XL Mutagenesis Kit (La Jolla, CA, USA). This was done according to the manufacturer's instructions. After mutagenesis all inserts were verified by sequencing.

Protein expression of MLH3 in the HEK293T cell line

HEK293T cells were seeded in a 6-well plate 24 hours before transfection to ensure 60% confluence on the day of transfection. 4 µg of YFP-MLH3 vector (WT or mutant) was transfected using 10 µl lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA). WT-CFP-MLH1 and WT-YFP-MLH3 were used as positive controls. Transfected HEK293T cells were lysed 48 hours after transfection using a non-denaturing lysis buffer (20mM Tris-HCl, pH 8, 135 mM NaCl, 1.5 mM MgCl2, 1 mM EGTA, 1% Triton X-100, 10% glycerol, protease inhibitors). Protein concentration determinations were performed according to Bradford (14). 75 µg of the cell lysate was loaded and size separated on a 6% SDS-PAGE gel. After separation by SDS-PAGE, proteins were transferred to a nitrocellulose membrane (Bio-Rad, Veenendaal, the Netherlands). MLH3 was detected with specific antibodies (Santa Cruz Biotechnology Inc, Santa Cruz, CA, USA) or with anti-GFP (GeneTex, Inc, San Antonio, TX, USA) and anti-mouse IgG HRP-conjugated secondary antibody (Pierce Biotechnology, Rockford, USA). Beta-tubulin (a housekeeping protein) was used as an internal loading control and detected with an anti-beta-tubulin antibody (Sigma-Aldrich Corp, St. Louis, MO, USA) and anti-mouse IgG HRP-conjugated secondary

antibody (Pierce Biotechnology, Rockford, USA). Signal visualisation was performed with the Supersignal West Dura Extended Duration Substrate kit (Pierce Biotechnology, Rockford, USA).

Subcellular localisation of MLH3

On the day prior to transfection, 100.000 cells were seeded in small glass petri dishes (20 mm in diameter). On the day of transfection, the media were refreshed. For transfection 1 µg of plasmid DNA (also 1 µg in total for the double transfection experiments) and 3 µl Fugene6 per petri dish was used. Fugene6 (3 µl) was mixed with 50 µl optimem and incubated for 5 minutes at room temperature. DNA was added and mixed, and this mixture was incubated for 20 minutes after which it was added to the cells. The cells were then incubated for ~24 hours. After incubation, the transiently expressed, fluorescently labelled proteins were visualised by confocal laser scanning microscopy (Zeiss LSM510, Carl Zeiss MicroImaging GmbH, Jena, Germany).

Yeast two hybrid analysis

To investigate the ability of UV-containing MLH3 proteins to form MLH1-MLH3 heterodimers, yeast two-hybrid assays were performed essentially as described by Rasmussen *et al*. (15). pAS2-MLH1 and pACT2-MLH3 were co-transformed into the *Saccharomyces cerevisiae* strain Y190. Transformants were selected on synthetic dextrose minimal medium (SD) lacking tryptophan, leucine and histine (SD/-Trp/-Leu/-His). At least five colonies were streaked onto an SD/-Trp/-Leu/-His plate that was supplemented with 25 mM 3-amino-1,2,4-triazole(3-AT). Colonies were further tested on SD/-Trp/-Leu/-His/3-AT + 40 µg/ml 5-bromo-4-chloro-3-indolyl-ß-D-galactosidase (X-gal) plates. Blue colonies that can grow on these selection plates express GAL4-tagged proteins that are able to interact. The two-hybrid vectors containing MLH3 or MLH1, in combination with the empty vectors, were used as negative controls.

Results

In silico analysis of the unclassified variants in MLH3

The results of the *in silico* analyses of the MLH3 UVs are presented in Table 2 and Figure 2. Seven of the eight UVs were predicted to be benign based on PolyPhen, a program partially based on conservation. Only the mutation p.Asn499Ser was predicted to be possibly pathogenic based on the PolyPhen algorithm.

A study for the possible splicing effects of the UVs, as determined by Netgene2, Splicesitefinder, Splice Site Predict and ESEfinder, made clear that none of the UVs were likely to cause splice defects. This analysis

included donor and acceptor site changes as well as inactivation of exonic splicing enhancer sequences.

Table 2. *In silico* analyses of the eight tested MLH3 UVs.

MLH3 UVs	Polarity change	Prediction by PolyPhen	Splice defect prediction
p.Gln24Glu (p.Q24E)	P→AP	benign	no
p.Asn499Ser (p.N499S)	P→P	possibly pathogenic	no
p.Glu624Asp (p.E624D)	AP→AP	benign	no
p.Ser817Gly (p.S817G)	P→NP	benign	no
p.Gly981Ser (p.G981S)	NP→P	benign	no
p.Asn1007Ser (p.N1007S)	P→P	benign	no
p.Ala1394Thr (p.A1394T)	NP→P	benign	no
p.Glu1451Lys (p.E1451K)	AP→BP	benign	no

P=polar amino acid (AA), NP=non-polar AA, AP= acidic charged polar AA, BP= basic charged polar AA.

The alignment depicted in Figure 2 shows that three of the eight UVs changed highly conserved amino acids (p.Gln24Glu, p.Glu624Asp, and p.Ala1394Thr), whereas two of the eight UVs changed reasonably well conserved amino acids (p.Asn1007Ala and p.Glu1451Lys).

These *in silico* experiments suggest that the MLH3 UVs are likely to be benign, with the possible exception of p.Asn499Ser.

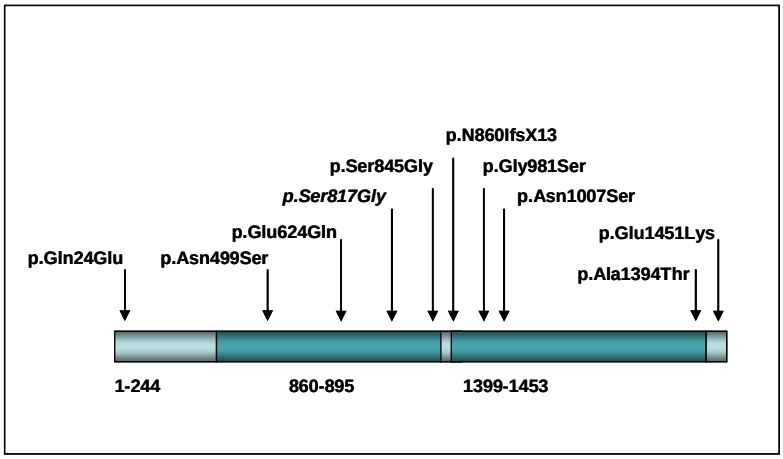


Figure 1. Schematic representation of the MLH3 protein and location of the tested MLH3 UVs and the truncating mutation in the protein. The light blue parts of the protein represent the MLH1 interaction domains.

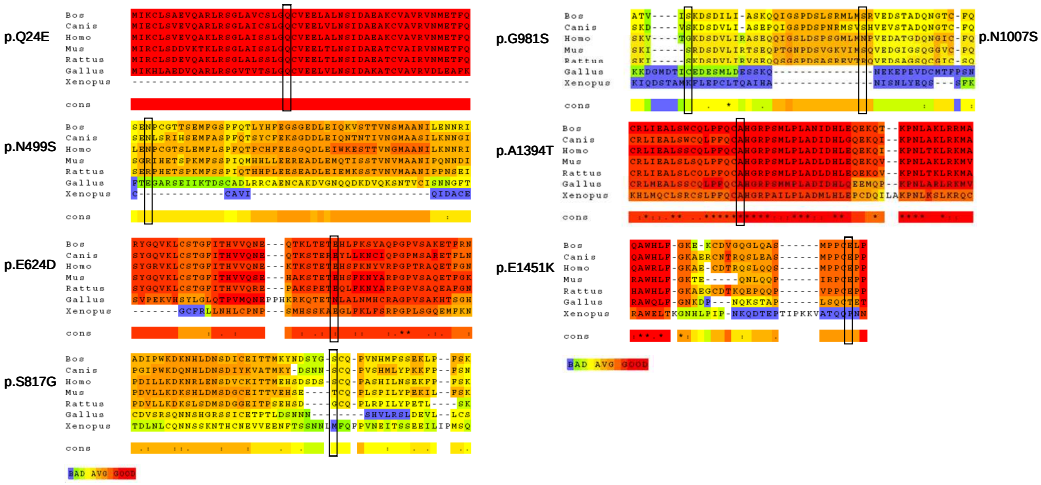


Figure 2. Alignment of the human MLH3 protein with MLH3 proteins of different vertebrates. The alignment includes the (putative) MLH3 proteins of the following organism: *Homo sapiens*, *Bos taurus*, *Canis familiaris*, *Gallus gallus*, *Mus musculus*, *Rattus norvegicus*, and *Xenopus tropicalis*. Only protein regions that contain the tested UVs are presented in this figure. Identical amino acids (*), conserved substitutions (:), or semi-conserved substitutions (.) are marked. The multiple sequence alignment was performed using the T-COFFEE program.

Transient expression of MLH3 in HEK293T cells

To evaluate the stability of the WT- or Mutant-YFP-MLH3 proteins, we transiently expressed Mutant-YFP-MLH3 in HEK293T cells. Notably, HEK293-T cells are deficient for both MLH1 and MLH3 due to hypermethylation of the promoter regions of both genes (12). The wildtype MLH3 protein was stably expressed, even in the absence of MLH1 (see Figure 3) and the MLH3 expression levels of all UV-containing proteins were comparable with the expression level of the WT-YFP-MLH3 (see Figure 3A) and that of one known missense polymorphism (p.Ser845Gly). The MLH3 protein variant with a frameshift mutation, p.Ans860IfsX13, was not detectable (see Figure 3A). This was caused by shortening of the protein leading to a loss of the epitope recognised by the MLH3-antibody. The MLH3 (H-2) antibody is a mouse monoclonal antibody raised against amino acids 1228-1453 of human MLH3. To verify the presence of the truncated protein, we reprobed the blot with an anti-GFP antibody. A protein with the size expected for the truncated protein was detected (around 126 kD), see Figure 3B. These results suggest normal expression of all UV-containing MLH3 proteins.

Subcellular localisation of MLH3

In order to evaluate whether the UV-containing MLH3 proteins were transported correctly into the nucleus, WT-YFP-MLH3 and Mutant-YFP-MLH3 were transfected into HEK293T cells or into HeLa cells. Figure 4 shows representative results of these experiments in HeLa cells (similar results were obtained in HEK293T cells, data not shown). In the HeLa and HEK293T cell lines, being MLH3-deficient and -proficient, respectively, all mutated and WT proteins were mainly localised in the nucleus. Some cytoplasmatic staining was seen, although this was independent of the presence of MLH1. These results suggest that the subcellular localisation of the mutated proteins is normal. The subcellular localisation for the MLH3 protein containing a truncating mutation clearly differs from the UV-containing mutations since, besides the nuclear staining, a strong cytoplasmatic staining was also observed.

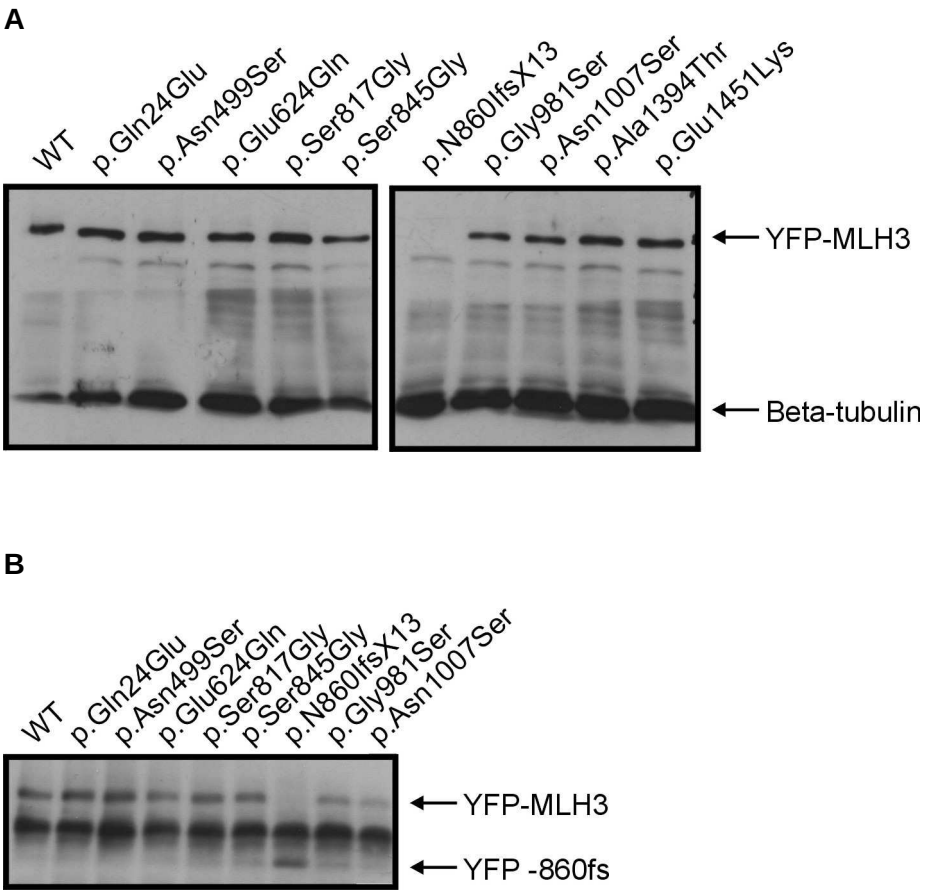


Figure 3. A. Expression of (UV-containing) MLH3 in HEK293T cells. This Western blot, using an MLH3 specific antibody, shows no differences in MLH3 expression when comparing cells transfected with WT-MLH3 and those transfected with UV-containing MLH3. B. Expression of MLH3 in HEK293T cells. This Western blot, using a GFP antibody shows in all but one lane the expected band of 190 kD. In lane 7 (extracts from cells transfected with p.Asn860IfsX13-MLH3) a lower band is visible (126 kD). Data shown are representative of three independent experiments. In this figure, in all lanes an aspecific band of approximately 150 kD is visible.

Yeast two-hybrid analysis

To determine whether MLH1 and UV-containing or wildtype MLH3 are able to form protein dimers *in vivo*, we performed a yeast two-hybrid screen. The screen, as shown in Figure 5, showed the eight MLH3 UVs analysed, one known MLH3 missense polymorphism (p.Ser845Gly), and the MLH3 frameshift mutation identified (see Table 1). All mutated MLH3 proteins were able to interact with WT-MLH1 and this interaction was comparable with WT-MLH3-WT-MLH1 interaction and with that of the known missense

polymorphism. The data suggest that the UV-containing MLH3 proteins can bind to WT-MLH1 *in vivo*.
The control experiments also showed colonies for the combination of the two empty vectors and the combination of pAS2 (empty vector) with pACT2-MLH3. The observed colonies, however, were white whereas the colonies observed for the WT-MLH1 with WT- or MUTANT-MLH3 were all blue, as expected. We also observed interaction between the truncating MLH3 mutation (p.Ans860IfsX13) and WT-MLH1. Unexpectedly, no colonies were seen when testing the interaction between p.MLH1 and p.Glu1415Lys. This could point towards a loss of interaction, but as we did not see normal colonies on the first selection plate, it is more likely that the experiment failed.

Discussion

Protein expression/stability studies
Transfection of mutant and WT-MLH3 in an MLH3-deficient cell line, HEK293-T, showed no difference in protein levels for the UVs tested. Our data also showed that the MLH3 protein is stable without a heterodimeric or other partner, corroborating previous findings (12).

MLH3 localisation studies
To repair mismatches *in vivo*, the MLH3 protein needs to be present in the nucleus. Recent localisation experiments (17) suggested that endogenous MLH3 is mainly localised in the cytoplasm whereas MLH1 and PMS2 are localised in the nucleus in a human MMR-proficient cell line. It was also shown that when MLH3 was transiently expressed in HCT116 (a human cell line deficient for MLH1 and PMS2), the MLH3 protein also localised in the cytoplasm. Only after co-transfection with MLH1, Sunyaev *et al*. (18) saw that MLH3 was partially transported into the nucleus. Furthermore, co-transfection of MLH3 with both MLH1 and PMS2 resulted in a cytoplasmatic localisation of MLH3, suggesting that MLH3 localisation not only depends on MLH1 but is also in competition with PMS2 (18). Our data do not corroborate these findings as we observed that both WT and UV-containing MLH3 proteins were largely located in the nucleus and only partially in the cytoplasm, regardless of whether MLH3 was transfected alone or co-transfected with MLH1 and regardless of the cell line we used for transfection. These results suggest, in contrast to Sunyaev *et al*.'s study, that nuclear MLH3 localisation does not depend on MLH1. When transfecting the vector expressing the truncated MLH3 protein (p.Asn860IfsX13) however, more cytoplasmatic staining can clearly be seen (see Figure 4). The protein, however, is normally not expressed as the naturally occurring unprocessed mRNA containing such a frameshift

mutation will be removed by the nonsense mediated RNA decay pathway.

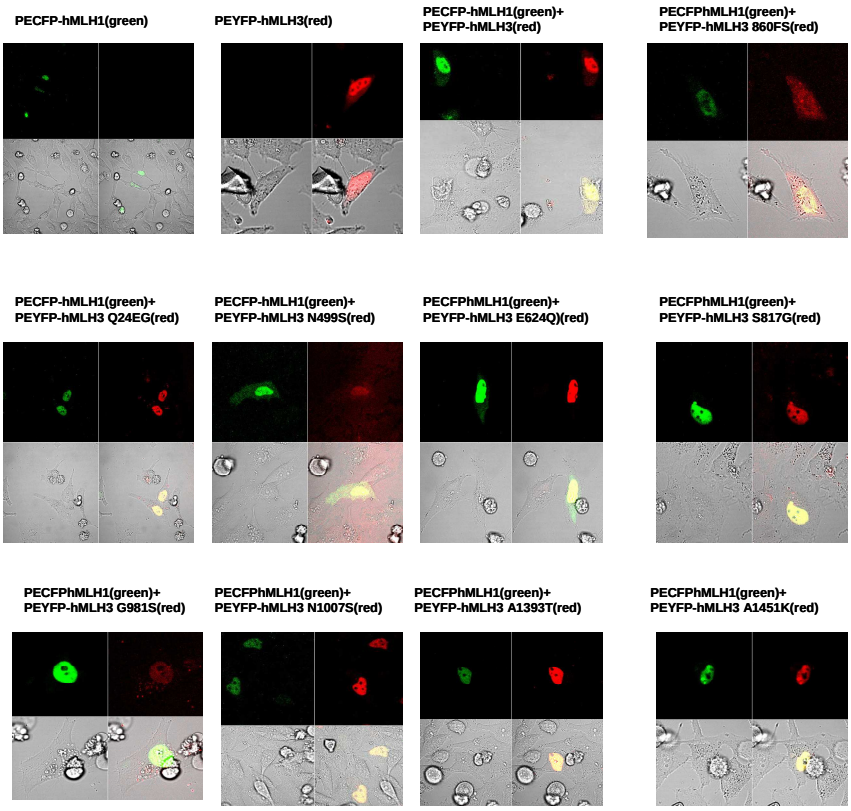


Figure 4. Subcellular localisation of CFP-MLH1 and YFP-MLH3, WT-MLH3, nine UV-containing MLH3s, and one truncating mutation containing MLH3 in HeLa cells.

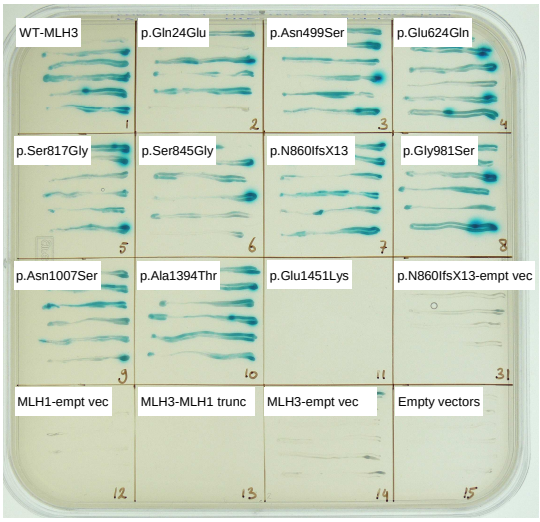


Figure 5. ß-gal assays (yeast two-hybrid assays) for WT-MLH3, nine UV-containing MLH3s, and one truncating mutation containing MLH3. Panels 1-11 show the interaction between MLH3 (with the indicated UV) with WT-MLH1. Panel 12. WT-MLH1 combined with empty vector, Panel 13. MLH1 (with a truncating mutation) combined with WT-MLH3, Panel 14. WT-MLH3 combined with an empty vector, Panel 15. two empty vectors, Panel 31. MLH3 coupled to the GAL4 activation domain instead of the GAL4 DNA binding domain (as is used in all other experiments) combined with empty vector.

Yeast two-hybrid studies

To study whether the identified MLH3 UVs influence the interaction with MLH1 *in vivo*, we tested the ten MLH3 mutants in a yeast two-hybrid system. Yeast two-hybrid studies showed that all the analysed MLH3 mutants were able to interact with WT-MLH1 and that these interactions were comparable with a WT-MLH3-WT-MLH1 interaction. The p.Glu1451Lys MLH3 protein was an exception as we did not obtain any colonies at all, but since we also saw no normal colonies on a minimal selection medium, we assume that the lack of colonies is due to a failed experiment rather than a lack of interaction.

Surprisingly, we also observed interaction between the truncating *MLH3* mutation (p.Asn860IfsX13) and WT-MLH1. A plausible explanation for this would be that a truncated MLH3 protein is produced, as shown in Figure 3B (see Results section, protein expression/stability studies). This truncated protein might still be able to interact with MLH1 to form a heterodimer with MLH1. Notably, the N-terminal MLH1 interaction domain was still present in the truncated form of MLH3, whereas the entire C-terminal MLH1 interaction domain was lacking (Figure 1).

In conclusion, these functional assays do not show a change in the function of the mutated proteins we tested. Not finding a functional defect in these assays does, however, not preclude the *MLH3* UVs being involved in Lynch syndrome. The assays described above investigate only specific aspects of MMR protein functioning. It is conceivable that the UVs tested can cause a functional defect of the protein that was not detectable with the assays used in this study. We cannot therefore fully exclude these UVs from being pathogenic.

Theoretical arguments for pathogenicity of MLH3 UVs

Besides data from functional assays, we also collected theoretical arguments that might help in determining the pathogenic nature of the MLH3 amino acid substitutions identified. We determined conservation in six vertebrates and showed that two of the amino acids that were mutated are highly conserved (Figure 2). We also looked for polarity changes and saw again that several of the UVs give rise to substantial changes in polarity. However, when all these data were combined in a program called PolyPhen we saw that only one of the eight UVs could be considered as possibly causative. The PolyPhen software combines structural, evolutionary and physicochemical properties. Notably, validation of the program showed a proper prediction in only 80% of known deleterious mutations, and thus false-positive or -negative findings can be expected (18,19), However, the prediction results are in line with the functional assays described above, with the exception of p.Asn499Ser, which was predicted to be possibly pathogenic. Our *in silico* analysis on splicing

showed that splicing abnormalities were not predicted for any of the UVs.
Based on the results from the functional assays and the theoretical prediction algorithms, we conclude that we have no convincing evidence that the MLH3 UVs tested are involved in the development of Lynch syndrome.

Is *MLH3* involved in Lynch syndrome?

This leaves us with the question whether mutations in *MLH3* can contribute to the development of Lynch syndrome? Our data do not support any involvement of the *MLH3* UVs identified in Lynch syndrome. Whether or not this can be concluded for the two reported *MLH3* frameshift mutations (11,16) remains unanswered. The fact that the tumours of these patients were MSI-low might be an argument against involvement in tumourigenesis. On the other hand, the finding of somatic mutations, knocking out the second allele in the tumour in three of the nine patients, argues in favour of their involvement (11). Finding MSI-L tumours in these patients should also be no surprise since it was shown that only a small but significant (20%) repair of both G/T mismatches and +1 insertion/deletion loop substrates was observed when MMR-deficient HEK293T nuclear protein extracts were supplemented with high amounts of MUTLγ. This suggests that MUTLγ might play a backup role in human MMR (12). A low activity is also reflected in the presence of low amounts of endogenous MLH3 protein in human cell lines. Semi-quantitative Western analysis of HeLa cells revealed endogenous MLH3 levels 60 times less abundant than in PMS2 and 6 times less abundant than in PMS1 (12).

Conclusion

We analysed the functional significance of eight *MLH3* UVs that had not previously been evaluated in functional assays. Our assays show that the MLH3 UVs are likely to be as functional as the wildtype MLH3 protein, suggesting that *MLH3* is not a major player in Lynch syndrome. However, we cannot fully exclude a role for MLH3 as a modifier in tumourigenesis.

Acknowledgments

This work was supported by: the Dutch Cancer Society (grant No. RUG2002-2678), the European Community (FP6-2004-LIFESCIHEALTH-5, proposal No 018754) and the Danish Cancer Society, Danish Research Council. We thank Jackie Senior for editing the text.

Reference List

(1) Chung DC, Rustgi AK. The hereditary nonpolyposis colorectal

cancer syndrome: genetics and clinical implications. *Ann Intern Med* 2003;**138**(7):560-70.

(2) Tishkoff DX, Boerger AL, Bertrand P *et al*. Identification and characterization of *Saccharomyces cerevisiae* EXO1, a gene encoding an exonuclease that interacts with MSH2. *Proc Natl Acad Sci U S A* 1997;**94**(14):7487-92.

(3) Tran PT, Simon JA, Liskay RM. Interactions of Exo1p with components of MutLalpha in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 2001;**98**(17):9760-5.

(4) Liberti SE, Rasmussen LJ. Is hEXO1 a cancer predisposing gene? *Mol Cancer Res* 2004;**2**(8):427-32.

(5) Harfe BD, Minesinger BK, Jinks-Robertson S. Discrete *in vivo* roles for the MutL homologs Mlh2p and Mlh3p in the removal of frameshift intermediates in budding yeast. *Curr Biol* 2000;**10**(3):145-8.

(6) Chen PC, Dudley S, Hagen W *et al*. Contributions by MutL homologues Mlh3 and Pms2 to DNA mismatch repair and tumor suppression in the mouse. *Cancer Res* 2005;**65**(19):8662-70.

(7) Lipkin SM, Wang V, Jacoby R *et al*. MLH3: a DNA mismatch repair gene associated with mammalian microsatellite instability. *Nat Genet* 2000;**24**(1):27-35.

(8) Jiricny J. The multifaceted mismatch-repair system. *Nat Rev Mol Cell Biol* 2006;**7**(5):335-46.

(9) Peltomaki P, Vasen H. Mutations associated with HNPCC predisposition -- Update of ICG-HNPCC/INSiGHT mutation database. *Dis Markers* 2004;**20**(4-5):269-76.

(10) Perucho M. Correspondence re: C.R. Boland *et al*., A National Cancer Institute workshop on microsatellite instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. Cancer Res., 58: 5248-5257, 1998. *Cancer Res* 1999;**59**(1):249-56.

(11) Wu Y, Berends MJ, Sijmons RH *et al*. A role for MLH3 in hereditary nonpolyposis colorectal cancer. *Nat Genet* 2001;**29**(2):137-8.

(12) Cannavo E, Marra G, Sabates-Bellver J *et al*. Expression of the

MutL homologue hMLH3 in human cells and its role in DNA mismatch repair. *Cancer Res* 2005;**65**(23):10759-66.

(13) Notredame C, Higgins DG, Heringa J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* 2000;**302**(1):205-17.

(14) Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 1976;**72**:248-54.

(15) Rasmussen LJ, Rasmussen M, Lee B *et al*. Identification of factors interacting with hMSH2 in the fetal liver utilizing the yeast two-hybrid system. *In vivo* interaction through the C-terminal domains of hEXO1 and hMSH2 and comparative expression analysis. *Mutat Res* 2000;**460**(1):41-52.

(16) Liu HX, Zhou XL, Liu T *et al*. The role of hMLH3 in familial colorectal cancer. *Cancer Research* 2003;**63**(8):1894-9.

(17) Korhonen MK, Raevaara TE, Lohi H *et al*. Conditional nuclear localization of hMLH3 suggests a minor activity in mismatch repair and supports its role as a low-risk gene in HNPCC. *Oncol Rep* 2007;**17**(2):351-4.

(18) Sunyaev SR, Lathe WC, III, Ramensky VE *et al*. SNP frequencies in human genes an excess of rare alleles and differing modes of selection. *Trends Genet* 2000;**16**(8):335-7.

(19) Ng PC, Henikoff S. Accounting for human polymorphisms predicted to affect protein function. *Genome Res* 2002;**12**(3):436-46.

Chapter 6

Functional analysis of Lynch Syndrome related missense mutations in MSH6

Jianghua Ou[1], Helga Westers[1], Renée C. Niessen[1], Sofie D. Andersen[2], Krista K. Kooi[1], Paul Jager[1], Bart J. Eggen[3], Jan. H. Kleibeuker[4], Rolf H. Sijmons[1], Lene J. Rasmussen[2], Robert M.W. Hofstra[1]

(1) Department of Genetics and (4) Gastroenterology, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands; (2) Department of Science, Systems and Models, Roskilde University, Denmark; (3) Developmental Genetics, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Haren, the Netherlands

Manuscription in preparation

Abstract

Introduction
Inherited pathogenic mutations in the mismatch repair (MMR) genes *MLH1*, *MSH2*, and *MSH6* predispose to Lynch syndrome. Major challenges in Lynch syndrome diagnostics are the DNA variants with an unclear pathogenic nature (unclassified variants, UVs), such as single amino acid substitutions and small or large in-frame deletions. In particular MSH6 UVs account for a substantial proportion of these UVs. This study has been performed to evaluate the pathogenicity of five of such inherited MSH6 UVs found in patients suspected of Lynch syndrome.
Methods
The mutated MSH6 proteins, all containing single amino acid substitutions, were tested for expression and stability in an MSH2/MSH6-deficient cell line (LOVO), for interaction with MSH2 by yeast two-hybrid experiments and for its subcellular localization.
Results
Protein expression of four of the five MSH6 mutants (p.Ser144Ile, p.Ala1021Asp, p.Ala326Val, and p.Thr1219Ile) was significantly decreased after transfection when compared with the expression of wildtype MSH6. Determining MSH6 gene expression by real time PCR showed a high similarity between the protein and the gene expression patterns. This suggests that the low protein expression is caused by low mRNA levels. No effects were observed on protein-protein interactions and the subcellular localization for all five MSH6 UVs was comparable to that of the WT MSH6.
Conclusion
Our data show that four of the five tested MSH6 UV seem to have an influence on gene expression and thereby on protein translation. Why the expression of these four UV-containing proteins *in vivo* is reduced is yet unknown. These UVs might therefore be pathogenic.

Introduction

Lynch syndrome or Hereditary Non-Polyposis Colorectal Cancer (HNPCC) is the most common autosomal dominant inherited colorectal cancer syndrome known. Lynch syndrome is caused by germline mutations in DNA mismatch repair (MMR) genes of which the majority are found in MSH2 and MLH1 (1). Additionally, five to ten percent of Lynch syndrome families carry a germline mutation in the MSH6 gene (2;3), but the number of mutations reported in the MSH6 gene is continuously rising. Besides a lower frequency of mutations in the Amsterdam criteria positive families, the average age of onset compared to MSH2 and MLH1 mutation carriers is higher in MSH6 mutation carriers (4) and therefore germline mutations in

the *MSH6* gene are recognized more frequently in so-called atypical Lynch syndrome or late onset families, families that do not fulfil all the clinical "Amsterdam criteria" (2). Furthermore, in families having germline *MSH6* mutations endometrial cancer is more frequent (5).
The involvement of MMR genes in Lynch syndrome can be made visible by the presence of small insertions or deletions in simple repetitive DNA sequences (also called microsatellite instability high or MSI-H) in tumours from these patients (6). Almost all tumours of Lynch syndrome patients demonstrate MSI. An exception on this can be found among tumours from patients with an *MSH6* mutation (7).
The rarity of *MSH6* germline mutations in Lynch syndrome and the atypical phenotype conferred by the *MSH6* germline mutation can partly be explained by redundancy of its function by another MMR protein. MSH6 forms a protein complex with MSH2, called MUTSα. This protein complex recognizes base-base mismatches and insertion/deletion loops and binds to these. The function of a second complex consisting of MSH2 and MSH3 (called MUTSβ) is partly redundant (8). These data are corroborated by studies on $Msh6^{-/-}$ mice, which have a reduced lifespan and developed a spectrum of tumours, all showing little or no MSI (9). Cellular extracts from $Msh6^{-/-}$ cells were deficient for the repair of single nucleotide mismatches (10). These findings indicated that inactivation of MSH6 can lead to cancer susceptibility and may be associated with cancer predisposition syndrome without showing microsatellite instability.

To date, the International society of Gastrointestinal Hereditary Tumours (InSiGHT) reports on over forty potentially pathogenic MSH6 mutations in their database (http://www.insight-group.org). More than one third of them result in an amino acid substitution of which the pathogenic nature is unclear. These mutations are therefore called unclassified variants (UVs). Only one of them has been proven pathogenic in functional assays (11) and few were shown to segregate in families (12). For an overview of functionally tested MMR UVs see www.mmrmissense.org (11). To address the question of pathogenicity of five of such MSH6 germline UVs, we performed several functional assays to determine protein-protein interactions, subcellular-localisation and protein expression-protein stability.

Material and Methods

Selected UVs
For this study five UVs were selected. These variants were identified in patients with atypical Lynch syndrome. The genetic and clinical characteristics of the UVs and the patients carrying the UVs can be found

in Table 1. Figure 1 shows a schematic representation of the MSH6 protein depicting the functional domains and the position of the UVs in the protein.

Table 1. Clinical genetic features associated with five MSH6 missense variants and their potential relationship with Lynch syndrome MSH6 UVs.
Cancer and age of diagnosis: described as tumour type, age of diagnosis. CRC: colorectal cancer EC:endometrial cancer, OV=ovarian cancer; H=MSI-H, L=MSI-L.

Exon	Nucleotide change	Predicted protein change	Cancer and age of diagnosis	MSI status	IHC MSH2/ MLH1/ MSH6	Amster-dam criteria II
2	c.431G>T	p.Ser144Ile	EC 45	H	+/+/+	-
2	c.431G>T	p.Ser144Ile	CRC 48	L	+/+/+	-
2	c.431G>T	p.Ser144Ile	CRC 49	H	+/-/-	-
4	c.977C>T	p.Ala326Val	CRC 65	H	+/+/+	-
4	c.1565A>G	p.Gln522Arg	CRC 40	L	+/+/+	-
4	c.3062C>A	p.Ala1021Asp	CRC 63	L	+/+/+	-
8	c.3656C>T	p.Thr1219Ile	CRC 37	H	+/+/+	-

In silico analysis of the unclassified variants in MSH6

Alignments with MSH6 homologues from six other vertebrates and yeast were obtained by Blasting the complete MSH6 protein. The program used was T-Coffee (13) (http://www.tcoffee.org). The following eight sequences were used in the Blast search: *Homo sapiens; Rattus norvegicus; Mus musculus; Canis familiaris; Bos taurus; Gallus gallus; Xenopus tropicalis; Saccharomyces cerevisiae*.

All five MSH6 UVs were further analyzed *in silico* for putative functional effects using the Polymorphism Phenotyping (PolyPhen) algorithm (http://genetics.bwh.harvard.edu/pph). PolyPhen is a web-based algorithm that predicts how an amino acid substitution will possibly affect the structure and function of a human protein using straightforward physical and comparative considerations. Calculations are based on amino acid homology. *In silico* analysis using the PolyPhen algorithm was done on the entire MSH6 protein sequence (including the UVs). Amino acid substitutions are reported as benign, possibly damaging or probably damaging.

Three web-based algorithms were used to predict possible splice defects initiated by the UVs: Netgene2 (http://www.cbs.dtu.dk/services/NetGene2), SpliceSiteFinder (http://violin.genet.sickkids.on.ca/~ali/splicesitefinder.html), and Splice Site Predict (http://www.fruitfly.org/seq_tools/splice.html). As input we gave a fragment containing the exon sequence plus 200

nucleotides upstream and 200 basepairs downstream of the exon. This was done for the exons 2, 4 and 8 as all 5 tested UVs are within these exons.

DNA techniques

MSH6 cDNA (wildtype) was cloned into pFastBac1 (a generous gift from Dr. Minna Nystrom-Lathi) between *Bam*HI and *Xho*I. Using this plasmid, wildtype *MSH6* was cloned into pCDNA3.1 and pYFPC1. To subclone *MSH6* WT in pAS2 (a yeast two-hybrid vector containing a GAL4 DNA binding domain) and pACT2 (a yeast two-hybrid vector containing a GAL4 activation domain), we PCR-amplified the gene using pFstBac1-WT-MSH6 as template using primers containing an additional CG between *Bam*HI and the first codon to obtain an in frame GAL4-MSH6 fusion protein.
pAS2-WT-MSH2 cDNA (cloned within a *Nco*I site) was used to subclone the MSH2-WT in pACT2 (in *Nco*I) (pAS2-WT-MSH2 was a generous gift from Dr. Anna Lutzen, University of Roskilde, Denmark).
The mutations p.Ser144Ile, pGln.522Arg, p.Ala1021Asp, p.Ala326Val, and p.Thr1219Ile were introduced into the *MSH6* gene in pFastbac1 and pYFPC1-MSH6 using the QuickChange XL mutagenesis kit from Stratagene (Stratagene European Headquarters, Amsterdam, the Netherlands) following the manufacturers protocols. All mutations were confirmed by DNA sequencing (the whole *MSH6* insert was sequenced).

The pFastbac1-MSH6 containing the p.Ser144Ile, p.Gln522Arg and p.Ala1021Asp were used for subcloning in: pACT2-MSH6 or pAS2-MSH6 using the restriction enzymes *Bln*I and *Bss*HII; in pCDNA3.1-MSH6 using *Bam*HI and *Bbv*CI for Ser144Ile; in pCDNA3.1-MSH6 using *Bam*HI and *Xho*I for Gln522Arg or Ala1021Asp. The p.Ala326Val and p.Thr1219Val mutants were directly introduced in the constructs pAS2-MSH6-WT, pACT2-MSH6-WT and pCDNA3.1-MSH6-WT using site-directed mutagenesis, followed by direct sequencing to confirm the mutations.

Yeast two-hybrid assay

The yeast two-hybrid assay was used to assess the interaction between the MSH6 (WT and mutant) and WT-MSH2 proteins. In short, pAS2-WT-MSH2 and pCAT2-WT-MSH6 or PCAT2-mutant-MSH6 were co-transfected into the *Saccharomyces cerevisiae* strain Y190. In the pACT2 vector *MSH6* is cloned in frame with the activation domain of GAL4 and in pAS2 MSH2 is cloned in frame with the GAL4 DNA binding domain. Furthermore, the pAS2 vector contains a leucine selection marker and pCAT2 contains a tryptophan selection marker. When the proteins are in each others vicinity also histidine can be used as a selection marker. Transformants were selected on synthetic dextrose minimal medium (SD medium) lacking tryptophan, leucine, and histidine. At least five colonies were streaked onto SD plates lacking tryptophan, leucine, and histidine and supplemented with

25 mM 3-amino-1,2,4-triazole(3-AT)(SD/-Trp-leu-His + 25 mM 3-AT) in order to test for protein interaction. Clones capable of growing on SD/-Trp/-leu/-His + 25 mM 3-AT + 40 µg/ml 5-bromo-4-chloro-3-indolyl-ß-D-galactosidase (X-gal) should contain proteins that are capable of assembling a functional GAL4 transcription factor. These were further evaluated for their ß-D-galactosidase activity. *MSH6* with a 1201 base frameshift deletion was used a negative control.

Cell lines and culturing

The cell lines used in this study were HeLa and LOVO. The HeLa cell line was derived from cervical cancer and the cell line is MMR-proficient. LOVO is a human colon adenocarcinoma-derived cell line that is MSH2-deficient. The MSH2-deficiency is caused by a homozygous deletion in the MSH2 gene (exon 3 to exon 8) (14). HeLa cells and LOVO cells were cultured in a humidified atmosphere containing 5% carbon dioxide in MEM (Invitrogen, Breda, the Netherlands) and RPMI 1640 medium (Invitrogen, Breda, the Netherlands), respectively. All media contained 10% fetal calf serum (Invitrogen, Breda, the Netherlands).

Transient transfection MSH2 and MSH6 in LOVO

LOVO cells were grown to 50-80% confluency in a 6 well plate and co-transfected with 2 µg maxiprep purified pCDNA3.1-MSH2 and 2 µg pCDNA3.1-MSH6 (WT or relevant mutants) plasmid DNA that was mixed with 5 µl lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA). Transfections were performed according to the manufacturer's protocol. After 48 hrs the cells were harvested and lysed in 20mM Tris-HCl, pH 8, 135 mM NaCl, 1.5 mM $MgCl_2$, 1 mM EGTA, 1% Triton X-100, 10% glycerol, supplemented with protease inhibitor tablets in a 1x final concentration (Roche, Basel, Switzerland) for total protein extraction. Protein concentration determinations were performed according to Bradford (15). 100 µg total protein extract was separated on denaturing 6% SDS polyacrylamide gels and blotted onto nitrocellulose membranes (Bio-Rad, Veenendaal, the Netherlands). Signal visualization was performed with the Supersignal West Dura Extended Duration Substrate kit (Pierce Biotechnology, Rockford, USA). The following antibodies were used: anti-MSH2 (Catalog No. NA27, Calbiochem, La Jolla, USA; 1:500 dilution), anti-GTBP (MSH6, clone 44, Transduction Laboratories, Breda, the Netherlands, 1:250 dilution) and MLH1 (clone G168-728, Pharmingen, Breda, the Netherlands, 1:500 dilution) and anti-mouse IgG HRP-conjugated secondary antibody (Pierce Biotechnology, Rockford, USA).

Quantitative PCR analysis of *MSH6* mRNA

LOVO cells were grown to 50-80% confluency in 6 well plate and co-transfected with 2 µg maxi-prep purified pCDNA3.1-MSH2 and 2 µg

pCDNA3.1-MSH6 (WT or relevant mutants) plasmid DNA that was mixed with 5 µl lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA). Transfections were performed according to the manufacturer's protocol. 48 hours after transfection, total RNA was extracted using RNABee (AMS Biotechnology (Europe) Ltd, Abingdon, Oxon, UK).
First strand cDNA was prepared from 1 µg total cellular RNA using random hexadeoxynucleotide primers and first strand cDNA synthesis ready to use beads (Amersham, GE Healthcare Europe GmbH, Diegem, Belgium). PCR was performed in a volume of 25 µl containing 1X Biorad iQ™ SYBR Green supermix (Bio-Rad laboratories B.V, Veenendaal, the Netherlands). SYBR Green fluorescence was measured realtime during PCR (IQ 6 PCR, Bio-Rad Laboratories B.V, Veenendaal, the Netherlands). Quantitative PCR was performed in triplicate for *MSH2* and *MSH6* primer sets and the mean value was used as relative quantification value. For normalization the *TBP* gene (TATA box binding protein) was used.

Localization of the MSH6 mutant proteins in the HeLa cell line
HeLa cells were maintained in DMEM (Invitrogen, Taastrup, Denmark) supplemented with 10%FBS (Invitrogen,Taastrup, Denmark) and 1% penicillin/streptomycin (Invitrogen, Taastrup, Denmark) and grown in 10% CO_2. On the day prior to transfection 100,000 cells are seeded in 20mm (in diameter) glass Petri dishes in medium containing both FBS and penicillin/streptomycin. On the day of transfection the medium is changed to fresh media (still with FBS and penicillin/streptomycin). 0.5 µg pCFP-MSH2 and 0.5 µg pYFP-MSH6 (WT or mutant) that was mixed with 3 µl Fugene6 transfection reagent (Roche, Basel, Switzerland) per dish. Transfections are performed according to the manufacturer's protocol. Cells are then incubated for ~24 hours before visualization of transfection by confocal laser scanning microscopy (Zeiss LSM510, Carl Zeiss MicroImaging GmbH, Jena, Germany).

Results

In silico analysis of the MSH6 UVs
The results of the *in silico* analyses of the MSH6 UVs are presented in Table 2 and Figure 2. These data shows that four of the five UVs change highly conserved amino acids (p.Ser144Ile, p.Glu522Arg, p.Ala1021Asp, and p.Thr1219Ile), whereas 1 of the five UVs changes a reasonably good conserved amino acid (p.Ala326Val).
Two of the five UVs (p.Ala326Val and p.Gln522Arg) are predicted to be benign based on PolyPhen, a program partially based on conservation. The other UVs were predicted to be possibly pathogenic based on the PolyPhen algorithm.

A study for the possible splicing effects of the UVs, as determined by Netgene2, SpliceSitefinder and Splice Site Predict, clarifies that one of the UVs (p.Ser144Ile) could have an alternative splicing due to an additional (occurring) acceptor site. Notably, this additional site has been predicted with the SpliceSiteFinder prediction program only and thus the consequences are not really clear.
These *in silico* experiments suggest that the *MSH6* UVs are likely to be benign with a possible exception for p.Ser144Ile.

Table 2. *In silico* analyses of the five tested MSH6 UVs.
Polarity change on Amino acid substitutions: P=polar amino acid (AA), NP=non-polar AA, AP= acidic charged polar AA, BP= basic charged polar AA

MSH6 UVs	Polarity change	Prediction by PolyPhen	Splice defect prediction
p.Ser144Ile (p.S144I)	P→ NP	Possibly damaging	additional acceptor site predicted
p.Ala326Val (p.A326V)	P→ NP	benign	No change
p.Gln522Arg (p.Q522R)	P→ NP	benign	No change
p.Ala1021Asp (p.A1021D)	NP → AP	Possibly damaging	No change
p.Thr1219Ile (p.T1219I)	P → NP	Probably damaging	No change

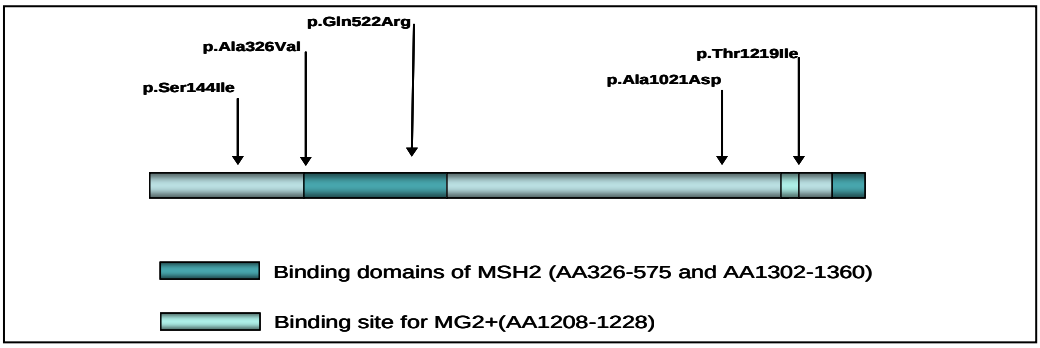


Figure 1. Schematic representation of the location of the five MSH6 missense mutations.

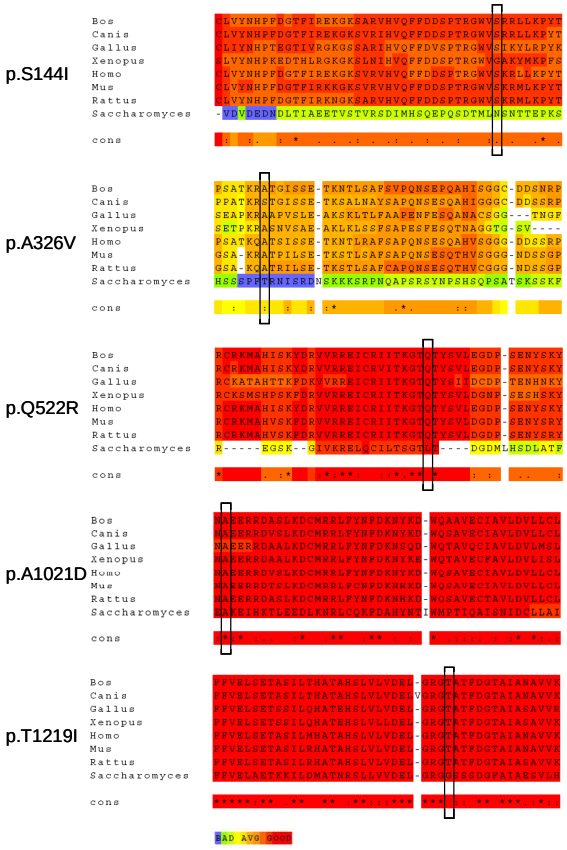


Figure 2. Alignment of the human MSH6 protein with MSH6 proteins of different vertebrates and yeast. The alignment includes the (putative) MSH6 proteins of the following organism: *Homo sapiens, Bos taurus, Canis familiaris, Gallus gallus, Mus musculus, Rattus norvegicus, Xenopus tropicalis,* and *Saccharomyces cerevisiae*. Only protein regions that contain the tested UVs are presented in this figure. Identical amino acids (*), conserved substitutions (:), or semi-conserved substitutions (.) are marked. The multiple sequence alignment was performed by using the T-COFFEE program.

Expression of wild type and mutant MSH6 in MSH2- and MSH6-deficient cells

To evaluate whether the UVs affect protein stability, WT or mutant MSH6 were transiently expressed in combination with MSH2 in LOVO cells. Representative results from 3 independent experiments are presented in Figure 3. Western blot analysis shows a strong reduction of MSH6 protein levels for the UV-containing proteins p.Ala1021Asp, p.Ala326Val, and p.Thr1219Ile when compared to MSH6 wildtype. The p.Ser144Ile mutant shows a slight decrease in MSH6 protein levels. The data suggest that these MSH6 mutant proteins are less stable, that the protein synthesis is lower or that the transcription of the gene is lower. The MSH6 p.Gln522Arg protein was expressed at levels that were comparable to wildtype MSH6 protein levels (see Figure 3).

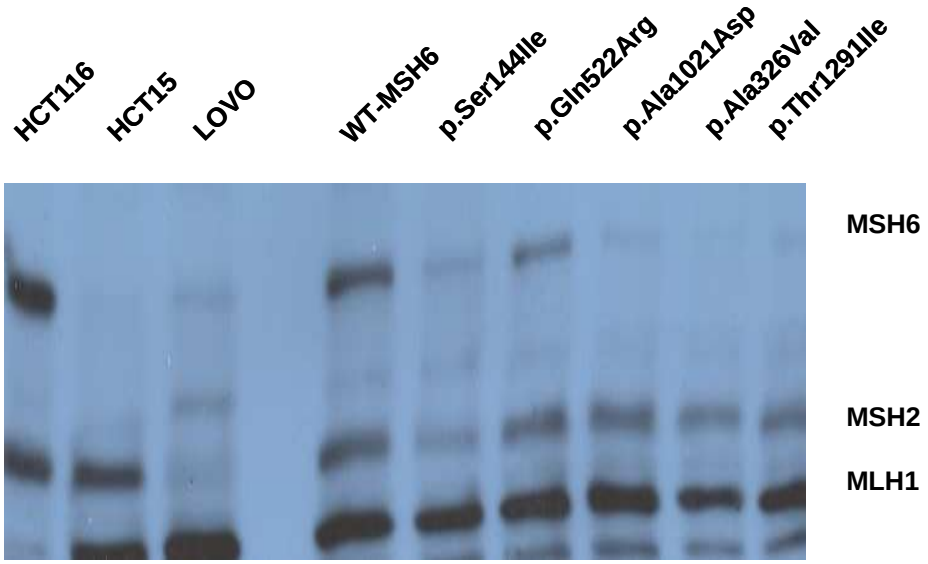


Figure 3. Expression of MSH2 and MSH6 in LOVO cells.
Cells were transiently transfected with pCDNA3.1-MSH2 and with pCDNA3.1-MSH6 (WT or containing one of the UVs). Controls on this blot are an MLH1-deficient cell line (HCT 116) an MSH6-deficient cell line (HCT15) and an MSH2 and MSH6-deficient cell line (LOVO).

Quantitative PCR analysis of *MSH6* mRNA

To differentiate between reduced transcription and reduced translation or protein stability, *MSH6* mRNA levels were quantified in transiently transfected LOVO cells. Quantification of *MSH6* mRNA levels showed significant differences between LOVO cells expressing wildtype *MSH6* or the mutated gene (48 hours after transfection, see Figure 4). In particular the cell lines transfected with the four UV-containing plasmids showed lower levels of the *MSH6* gene expression (p.Ala1021Asp, p.Ala326Val, p.Thr1219Ile, and p.Ser144Ile) indicating that for these four mutants the transcription of the gene is lower when compared to wildtype *MSH6*.

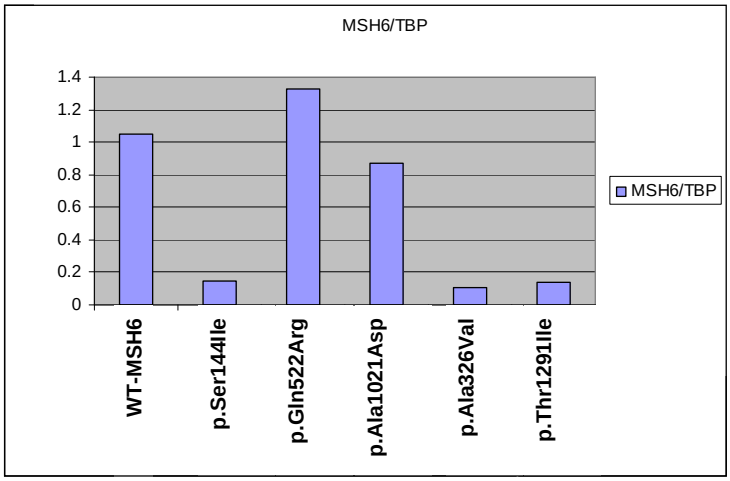


Figure 4. Quantitative PCR-analysis 48 hours after transfection.
Total RNA of transiently transfected LOVO cells expressing the MSH2 wildtype protein and wildtype *MSH6* or one of five UVs were extracted 48 hours after transfection. *MSH6* mRNA levels were quantified using real-time PCR. Y-axis represents the relative mRNA quantity of MSH6 normalized to the house keeping gene *TBP*.

Localization experiments

We investigated the co-localization of CFP-MSH2 and YFP-MSH6 (WT and five UV-containing proteins) in HeLa cells. Figure 5 shows representative results of these tansfections. All five MSH6 mutants localize to the nucleus when transfected together with wildtype MSH2 suggesting normal subcellular localization of all five MSH6 mutants.

Yeast two-hybrid analyses for determining an interaction between MSH2 and MSH6

The yeast two-hybrid screen revealed that all five tested MSH6 UV-containing proteins showed an interaction with WT-MSH2 that was

comparable with the interaction of wildtype MSH6 with wildtype MSH2 (see Figure 5). Unexpectedly, we observed colonies for the combination of the two empty vectors and the combination pAS2 with pACT2-MSH6. The observed colonies, however, were white whereas the colonies observed for the WT-MSH2 with (WT or mutant) MSH6 were all blue (see Figure 6).

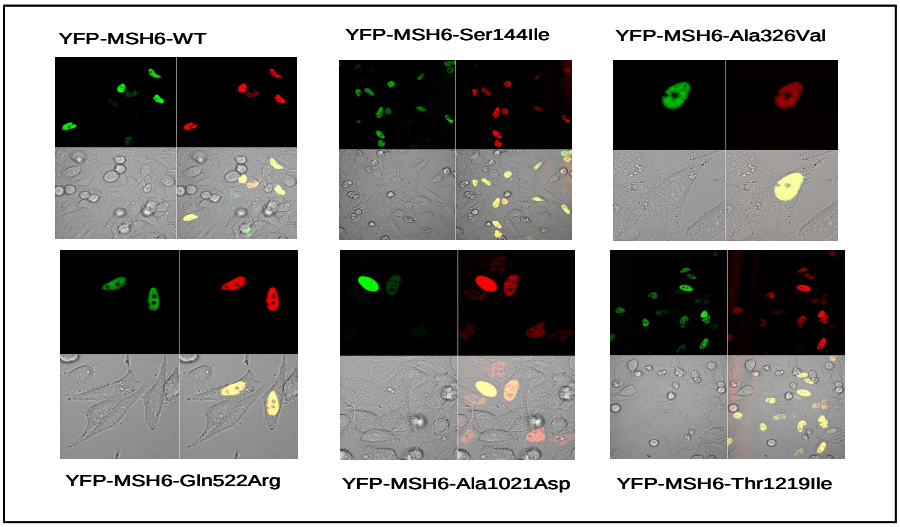


Figure 5. Subcellular localization of CFP-MSH2 and YFP-MSH6 (WT and one of 5 UVs). Nuclear localization of MSH6 proteins can be seen.

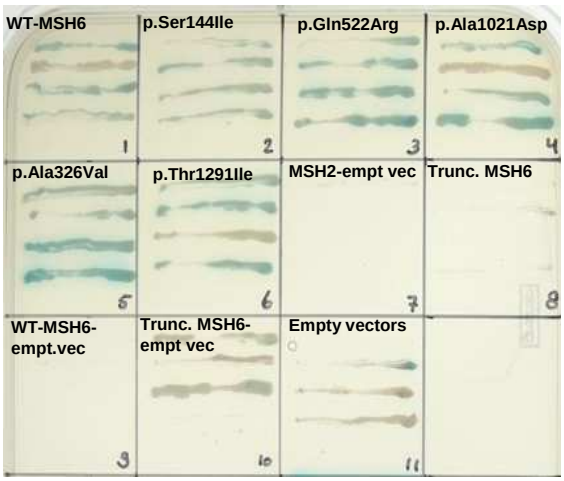


Figure 6.
Yeast two-hybrid assays for wildtype MSH6 and 5 UV-containing MSH6 proteins in combination with the wildtype MSH2 protein. Panels 1-6 show the ß-gal assays for wildtype MSH2 (coupled to the GAL4 DNA binding domain) in combination with MSH6 (wildtype or one of the five UVs coupled to the DNA activation domain). Panels 7-11; negative controls. Panels 7 and 8; MSH2(WT) coupled to the GAL4 DNA binding domain combined with the empty vector containing the GAL4 DNA activation domain or the GAL4 DNA activation domain coupled to a truncated MSH6 protein. Panels 9-10; MSH6(WT) or truncated MSH6 (coupled to the GAL4 DNA activation domain) combined with the empty vector (containing the GAL4 DNA binding domain). Panel 11; the two empty vectors.

Discussion

Yeast two hybrid assays showed that all five UV-containing MSH6 proteins were able to interact with WT-MSH2 and that these interactions are comparable to that between WT-MSH6 and WT-MSH2 showing that these mutations have no influence on the interaction between MSH6 and MSH2 in yeast. Also localization studies showed that the five mutant proteins are present mostly in the nucleus as is the WT-MSH6. When however studying the protein expression and stability *in vivo* (in LOVO cells) we noticed a decreased expression for four of the five UVs (p.Ser144Ile, p.Ala1021Asp, p.Ala326Val and p.Thr1219Ile).
Finding lower levels of the UV-containing proteins might be caused by; a) rapid degradation of the protein; b) lower levels of gene expression; c) a rapid degradation of the mRNA; d) mutations in the expression plasmid that lower the expression.
To determine whether the lower protein expression levels found were caused by lower gene expression we performed quantitative PCR on the transfected cells. These experiments showed that the RNA and protein levels show a positive correlation, low RNA levels relate to low protein levels. This makes it plausible that the observed low proteins levels for four of the five UVs are caused by low gene expression levels. What is causing the low expression levels? It might be that the UVs have an influence of the expression levels. Another option might be the presence of mutations in the vector, *e.g.* in the promoter region of *MSH6*. As however three MSH6 mutated inserts were cloned out of pFastbac1-MSH6, the vector backbones of these three UVs are the same. As one of the three is showing normal expression (p.Gln522Arg) the differences in expression can not be explained by mutations in the vector. Likely, therefore, is an influence of the UVs on the expression level.
In conclusion, it is not unlikely that four of the five tested MSH6 UVs are pathogenic. Some caution, however, might be considered. The tumours mostly show normal expression of MSH6 by immunohistochemical staining. These immunohistochemical stainings are not quantitative so difference between normal expression and low expression might be difficult to see. Moreover, when expressing the YFP-tagged MSH6 proteins in HeLa cells we did see normal expression. Additional experiments are thus needed to confirm pathogenicity. The prediction results are in line with the functional assays as described above, with the exception that the UV p.Gln522Arg is predicted to be benign. Furthermore, our analysis on splicing showed that the mutation c.431G>T could lead to an alternative splicing.

<u>Additional evidence</u>
Previously, the mutation p.Ser144Ile in the MSH6 protein was studied in a yeast-based functional assay (16) and in an *in vitro* MMR assay (17). In the

yeast-based assay, the mutation was thought to be pathogenic. As however the mutation occurs at a position that is not conserved in the yeast MSH6 protein (see figure 2) the data must be taken with great caution. In the *in vitro* assay (17), it was shown that the UV-containing MSH6 protein was as functional as wildtype MutSα. Combining our data with that of Kariola might indicate that, when our data is correct, the pathogenicity of the p.Ser144Ile is not directly linked to MMR function, but rather to lowered expression.
Finding low expression is in agreement with previous observations that certain missense variants were associated with a markedly reduced expression of the corresponding allele in peripheral blood lymphocytes. Furthermore, it was suggested that transcript deregulation or unbalanced germline line expression of missense alleles may represent a relevant mechanism for MMR UVs (18).
Finally, our data suggest that certain MSH6 UVs might be associated with down regulation of *MSH6* mRNA levels in stead of a lower MMR functioning. Further studies are needed to confirm these data.

Acknowledgments

This work was supported by: the Dutch Cancer Society (grant No. RUG2002-2678), the European Community (FP6-2004-LIFESCIHEALTH-5, proposal No 018754).

Reference List

(1) Peltomaki P, Vasen H. Mutations associated with HNPCC predisposition -- Update of ICG-HNPCC/INSiGHT mutation database. *Dis Markers* 2004;**20**(4-5):269-76.
(2) Berends MJ, Wu Y, Sijmons RH *et al*. Molecular and clinical characteristics of MSH6 variants: an analysis of 25 index carriers of a germline variant. *Am J Hum Genet* 2002;**70**(1):26-37.
(3) de la Chapelle A. Genetic predisposition to colorectal cancer. *Nat Rev Cancer* 2004;**4**(10):769-80.
(4) Miyaki M, Konishi M, Tanaka K *et al*. Germline mutation of MSH6 as the cause of hereditary nonpolyposis colorectal cancer. *Nat Genet* 1997;**17**(3):271-2.
(5) Wijnen J, de Leeuw W, Vasen H *et al*. Familial endometrial cancer in female carriers of MSH6 germline mutations. *Nat Genet* 1999;**23**(2):142-4.
(6) Boland CR, Thibodeau SN, Hamilton SR *et al*. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection

and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res* 1998;**58**(22):5248-57.

(7) Wu Y, Berends MJ, Mensink RG *et al*. Association of hereditary nonpolyposis colorectal cancer-related tumours displaying low microsatellite instability with MSH6 germline mutations. *Am J Hum Genet* 1999;**65**(5):1291-8.

(8) Chung DC, Rustgi AK. The hereditary nonpolyposis colorectal cancer syndrome: genetics and clinical implications. *Ann Intern Med* 2003;**138**(7):560-70.

(9) Edelmann W, Yang K, Umar A *et al*. Mutation in the mismatch repair gene Msh6 causes cancer susceptibility. *Cell* 1997;**91**(4):467-77.

(10) Genschel J, Littman SJ, Drummond JT *et al*. Isolation of MutSbeta from human cells and comparison of the mismatch repair specificities of MutSbeta and MutSalpha. *J Biol Chem* 1998;**273**(31):19895-901.

(11) Ou J, Niessen RC, Lutzen A *et al*. Functional analysis helps to clarify the clinical importance of unclassified variants in DNA mismatch repair genes. *Hum Mutat* 2007;**28**(11):1047-54.

(12) Cederquist K, Emanuelsson M, Wiklund F *et al*. Two Swedish founder MSH6 mutations, one nonsense and one missense, conferring high cumulative risk of Lynch syndrome. *Clin Genet* 2005;**68**(6):533-41.

(13) Notredame C, Higgins DG, Heringa J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* 2000;**302**(1):205-17.

(14) Umar A, Boyer JC, Thomas DC *et al*. Defective mismatch repair in extracts of colorectal and endometrial cancer cell lines exhibiting microsatellite instability. *J Biol Chem* 1994;**269**(20):14367-70.

(15) Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 1976;**72**:248-54.

(16) Kolodner RD, Tytell JD, Schmeits JL *et al*. Germ-line msh6 mutations in colorectal cancer families. *Cancer Res* 1999;**59**(20):5068-74.

(17) Kariola R, Raevaara TE, Lonnqvist KE *et al*. Functional analysis of MSH6 mutations linked to kindreds with putative hereditary non-polyposis colorectal cancer syndrome. *Hum Mol Genet* 2002;**11**(11):1303-10.

(18) Curia MC, Palmirotta R, Aceto G *et al*. Unbalanced germ-line expression of hMLH1 and hMSH2 alleles in hereditary nonpolyposis colorectal cancer. *Cancer Res* 1999;**59**(15):3570-5.

Chapter 7

General discussion and future perspectives

Carrier detection in Lynch syndrome

Lynch syndrome, also known as hereditary non-polyposis colorectal cancer (HNPCC), is one of the most common hereditary cancer syndromes known worldwide. It is caused by germline mutations in one of the DNA mismatch repair (MMR) genes, *MLH1*, *MSH2*, *MSH6* or *PMS2* (1). Identifying a germline MMR mutation is important as inclusion of mutation carriers in cancer surveillance programs has proven to be life saving. Identification of mutation carriers, therefore, has been given major attention. As mutation analysis of all MMR genes in all patients with Lynch syndrome-associated tumours is not practical, several selection criteria for mutation screening of the MMR genes have been proposed. These selection criteria are based either on family history or on tumour tissue related assays such as microsatellite instability (MSI) or immunohistochemical staining of the MMR proteins (IHC) (2;3). When tumour tissue from affected patients suspected of having Lynch syndrome is unavailable, MMR gene mutation analysis can be offered to them, or in case DNA from the patients is unavailable as well, to their unaffected family members. However, this approach is time consuming and costly as all known Lynch syndrome genes, often in multiple family members, would need screening. In our experience, the situation of unavailable tumour tissues and DNA from affected relatives is not rare. Therefore, we decided to explore the possibility of a pre-screen method that would allow us to predict the presence of a germline mutated MMR gene in normal tissue of a possible carrier. The concept behind our approach was to detect a (50%) loss of expression (haploinsuffiency) due to this germline mutation. To detect this loss of expression we analyzed gene-expression by quantitative PCR of RNA and protein expression by Western blotting. The data we collected, however, made us conclude that the expression levels of MLH1 and MSH2, the two genes/proteins we analyzed, varied significantly, both within and between the groups of patients and controls. This variation in expression made it impossible to reliably predict the mutation carrier status (Chapter 2). As our proposed method did not work, alternatives should be considered. An option could be measuring allele-specific expression of the MMR genes in lymphocyte or lymphoblast cell lines. The allelic expression method differs from the method we applied as it measures the relative expression of both alleles separately. Variations in the expression levels of the MMR genes as detected by us are therefore not important. The group of Prof. Burn (Newcastle, UK) demonstrated on a poster at the InSIGHT meeting in Yokohama, Japan (2007) that this procedure indeed works. To do so,

however, one needs to be able to distinguish both alleles of a gene. If one wants to use this method for the three most commonly mutated genes (*MSH2*, *MLH1* and *MSH6*) one first needs to identify informative coding polymorphisms in the three genes in the individuals that ask for screening. This clearly is a major drawback of the method.
Another option might be to block the nonsense mediated RNA decay (NMD) pathway in cultured cells and compare the gene or protein expression levels before and after blocking (4;5). An increase of the gene expression or protein levels or an aberrant protein product after NMD inhibition, would indicate the presence of a mutation. A drawback of an NMD based method is that NMD is not able to breakdown messengers with mutations in the first and last exons of genes (6;7).
One might also opt for a total different approach, namely solid phase sequencing (also called medical sequencing or deep sequencing). As this type of sequencing is getting cheaper and cheaper, sequencing all MMR genes in a single experiment is becoming affordable (8;9). These developments make it likely that our present concerns, that the screening of all genes in several individuals of a single family is too costly and too labour intensive, will be no longer valid before long.

Missense mutations in Lynch syndrome
After identifying a DNA variant one needs to decide whether the variant identified is involved in the development of the disease. For truncating mutations the decision is easily made. For missense variants and small in frame deletions or insertions this is far from easy (see chapter 3 review). It is therefore that these variants are called unclassified variants (UVs). A significant proportion of DNA variants found in Lynch syndrome (suspected) patients are such UVs: 32%, 18%, 38% and 87% for *MLH1*, *MSH2*,*MSH6* and *MLH3*, respectively (1). These UVs constitute a major problem in DNA diagnostics for Lynch syndrome.

Database and decision making
Screening the literature made clear that only a small proportion of the MMR UVs identified in Lynch syndrome suspected patients have been functionally tested. Until early 2007 we could find 115 UVs that were functionally characterized. The testresults for the UVs we came across in our literature search are collected in a database that is now online for all those interested (www.mmrmissense.info) (see also Chapters 3 and 4). An interesting finding when analyzing these data was that if the data from these studies indeed are correct, a large proportion of these UVs are probably pathogenic. For *MLH1* 70% (49/70) of the UVs were considered pathogenic by the authors. For *MSH2* this was also 71% (25/35), while for *MSH6* it was only 13% (1/8). However, we cannot exclude that these percentages are subject to ascertainment bias.

If indeed these data are true it would have a large impact on the carriers of these variants and their families and on how one should prescreen patients for mutation analysis. We should reconsider the inclusion criteria for mutation analysis, as some of them are based on finding truncating mutations only. For example, to date, a combination of a young age at diagnosis and /or a positive family history, with an MSI-H status and/or negative IHC for an MMR protein are generally considered as predictors and inclusion criteria for MMR mutation analysis (10;11)). Clearly, missense variant carriers would be excluded as in many cases the immuno-histochemical staining would very likely be positive. Furthermore, some publications have shown that certain missense mutations (like E578G in *MLH1*) do not correlate with MSI (12). Validation of the data therefore is crucial.

Functionally Testing UVs

As mentioned, a large proportion of the UVs might be pathogenic, in particular when we consider *MLH1* and *MSH2*. However, for *MSH6* only eight UVs had been functionally tested and of these eight only one was considered pathogenic. As we have found a relative large number of UVs in *MSH6* (13), we performed functional assays on 5 of these. These UVs were identified previously in patients suspected of Lynch syndrome. Although we did find differences in expression levels for some, we have no conclusive evidence that the 5 tested *MSH6* UVs are involved in cancer development (chapter 6). Are these UVs, as most other *MSH6* UVs, not involved in Lynch syndrome? Clearly, the assays used do not show dramatic effects for these UVs in the assays performed. Small effects, however, cannot be excluded. Furthermore, we only tested for specific functions of the proteins and did not yet determine mismatch repair functioning itself. We therefore can not exclude some involvement of these UVs in Lynch syndrome. Might these UVs be weak mutations (mutations with a small effect on the protein function), then the assays used might not detect the possible pathogenic nature of these mutations. Moreover, we know that the penetrance of *MSH6* mutations can be lower than that of mutations in *MSH2* and *MLH1* and we know that the age of onset is generally higher (14). It can therefore not be excluded that these UVs have a small but significant effect. Last but not least, we have identified the co-occurrence of a second MMR gene mutation several times, in particular in patients with *MSH6* UVs (15;16). This might be coincidental but it might also point towards a more complex mode of inheritance.

In addition to the UVs in *MSH6* we also tested eight *MLH3* UVs. As in the tested *MSH6* UVs, we found no evidence that these variants contribute to disease development. Again, we can not totally exclude involvement of the UVs in disease development. For *MLH3,* involvement of the UVs as modifiers for the disease phenotype, rather than being high penetrant

mutations, seems even more likely as normal expression levels of this gene are very low when compared to *MSH2* and *MLH1* (17).

Limitations of the functional assays

All the assays performed and described in this thesis investigate specific aspects of MMR protein functioning. UVs that do not change the functional aspects analyzed may still lead to a MMR defect. Therefore, it should be emphasized that even if no obvious defects are found by such assays, pathogenicity can still not be totally excluded.

On the other hand, when assays predict a pathogenic nature of a variant one should realize that it is of utmost importance that the proper positive and negative controls were included in the study. Besides these controls also the expression level of the transfected proteins should be kept in mind. High levels of these MMR proteins might be toxic or could disturb normal functioning of the overexpressed protein. Further, when only part of the protein is expressed (which is not the case in the assays performed by us and described in this thesis) this might also lead to findings that may not reflect the impact on function of a full-length protein.

Possibly the biggest problem of functional assays is the fact that for hardly any of the assays the sensitivity and specificity has been determined. As long as a test has not been validated properly, results should be utilized with great caution.

Taken together, the outcome of functional testing in its present state cannot always be used in clinical practice as direct proof of (non)pathogencity (see chapter 3 for a review)

The future of functional assays

Clearly, one of the greatest challenges in Lynch syndrome research is the development of more sensitive and well validated functional assays. One might opt, as we did, for different functional assays that each test a specific property of the protein. Another option might be a test that measures the DNA mismatch repair capacity as a whole. Because deficient mismatch repair is considered the source of accumulation of somatic mutations and thereby tumour development in Lynch syndrome, the most logical option would be to first test the repair capacity of the mutated protein. After demonstrating deficient mismatch repair, one might use specific tests to determine the precise effect the UV has on the mutated protein to support and understand the findings. In this approach, this first test would then be crucial. However, considering these repair assays, several things should be kept in mind. To name a few: the *in vivo* and *in vitro* differences that exist for proteins (*in vivo* expressed proteins *e.g.* undergo proper folding and many modifications); the amount of protein might be crucial (it should mimic the real situation) –for example, high levels of these proteins are toxic or

could disturb normal function of the proteins. And last but not least, using non-human test systems might give false positive or false negative results.

At this moment a second major challenge is data interpretation. One surely needs to combine different arguments: determine MMR repair; perform *in silico* analysis of the variants (*e.g.* conservation of the amino acid, kind of amino acid change); use also indirect evidence such as segregation and MSI. And hopefully in the future, based on many tested UVs in validated assays, algorithms can be made that do allow prediction of the function of UVs.

Reference List

(1) Peltomaki P, Vasen H. Mutations associated with HNPCC predisposition -- Update of ICG-HNPCC/INSiGHT mutation database. *Dis Markers* 2004;**20**(4-5):269-76.

(2) Thibodeau SN, Bren G, Schaid D. Microsatellite instability in cancer of the proximal colon. *Science* 1993;**260**(5109):816-9.

(3) Debniak T, Kurzawski G, Gorski B *et al*. Value of pedigree/clinical data, immunohistochemistry and microsatellite instability analyses in reducing the cost of determining hMLH1 and hMSH2 gene mutations in patients with colorectal cancer. *Eur J Cancer* 2000;**36**(1):49-54.

(4) Rossi MR, Hawthorn L, Platt J *et al*. Identification of inactivating mutations in the JAK1, SYNJ2, and CLPTM1 genes in prostate cancer cells using inhibition of nonsense-mediated decay and microarray analysis. *Cancer Genet Cytogenet* 2005;**161**(2):97-103.

(5) Durand S, Cougot N, Mahuteau-Betzer F *et al*. Inhibition of nonsense-mediated mRNA decay (NMD) by a new chemical molecule reveals the dynamic of NMD factors in P-bodies. *J Cell Biol* 2007;**178**(7):1145-60.

(6) Tournier I, Raux G, Di FF *et al*. Analysis of the allele-specific expression of the mismatch repair gene MLH1 using a simple DHPLC-Based Method. *Hum Mutat* 2004;**23**(4):379-84.

(7) Mendell JT, Sharifi NA, Meyers JL *et al*. Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. *Nat Genet* 2004;**36**(10):1073-8.

(8) Rosenthal A, Schwertner S, Hahn V *et al*. Solid-phase methods for sequencing of nucleic acids I. Simultaneous sequencing of different oligodeoxyribonucleotides using a new, mechanically stable anion-exchange paper. *Nucleic Acids Res* 1985;**13**(4):1173-84.

(9) Bentley DR. Whole-genome re-sequencing. *Curr Opin Genet Dev* 2006;**16**(6):545-52.

(10) Reyes CM, Allen BA, Terdiman JP *et al*. Comparison of selection strategies for genetic testing of patients with hereditary nonpolyposis colorectal carcinoma: effectiveness and cost-effectiveness. *Cancer* 2002;**95**(9):1848-56.

(11) Umar A, Boland CR, Terdiman JP *et al*. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *J Natl Cancer Inst* 2004;**96**(4):261-8.

(12) Liu T, Tannergard P, Hackman P *et al*. Missense mutations in hMLH1 associated with colorectal cancer. *Hum Genet* 1999;**105**(5):437-41.

(13) Niessen RC, Berends MJ, Wu Y *et al*. Identification of mismatch repair gene mutations in young patients with colorectal cancer and in patients with multiple tumours associated with hereditary non-polyposis colorectal cancer. *Gut* 2006;**55**(12):1781-8.

(14) Berends MJ, Wu Y, Sijmons RH *et al*. Molecular and clinical characteristics of MSH6 variants: an analysis of 25 index carriers of a germline variant. *Am J Hum Genet* 2002;**70**(1):26-37.

(15) Wu Y, Berends MJ, Sijmons RH *et al*. A role for MLH3 in hereditary nonpolyposis colorectal cancer. *Nat Genet* 2001;**29**(2):137-8.

(16) Niessen RC, Sijmons RH, Ou J *et al*. MUTYH and the mismatch repair system: partners in crime? *Hum Genet* 2006;1-6.

(17) Cannavo E, Marra G, Sabates-Bellver J *et al*. Expression of the MutL homologue hMLH3 in human cells and its role in DNA mismatch repair. *Cancer Res* 2005;**65**(23):10759-66.

Chapter 8

English summary

In this thesis we present a number of studies related to the classification of missense mutations in Lynch syndrome and we tested an alternative, tumourindependent approach for the identification of mismatch repair (MMR) mutation carriers.

Identification of MMR mutation carriers.

Identification of MMR-gene mutation carriers and enrolling mutation carriers in surveillance programs for the early detection of tumours has been shown to remarkably reduce the risk of colon cancer in the identified mutation carriers and of mortality due to colon cancer. Identification of mutation carriers, therefore, has been given considerable attention. However, as mutation analysis of all MMR genes is timeconsuming and costly, several clinical selection criteria and laboratory prescreening methods have been developed which help to decide whether or not a patient should be screened for mutations. The Dutch guidelines for the selection of patients for mutation analysis are based on the age at diagnosis, microsatellite instability screening, immunohistochemical screening of the MMR proteins in the tumour, methylation screening of the *MLH1* promoter and *BRAF* mutation screening. When a tumour is microsatellite instable (MSI-H) and/or immunohistochemistry is negative for MSH6 and/or MSH2, mutation analysis will be performed. When MLH1 is negative by immunohistochemistry, tumours are first tested for a *BRAF* mutation and *MLH1* promoter hypermethylation. Tumours with *MLH1* promoter hypermethylation and a *BRAF* mutation will not be tested for germline mutations as this is evidence that points towards sporadic cases. Tumours without *MLH1* promoter hypermethylation and without a *BRAF* mutation will be screened for a MMR mutation.

When no tumour material (and or no DNA) from a patient or another affected family member is available one generally needs to screen all Lynch associated MMR genes for mutations often in multiple family members. We worked on an approach for which no tumour material is necessary and which is based on the expression levels of the MMR proteins/genes in leukocytes. What we were hoping to develop was a method that could detect the loss of the mutated MMR allele (haploinsuffiency) based on gene- or protein-expression levels. Haploinsuffiency would point directly to the mutated gene (in the germline) thereby saving time and money. We show (Chapter 2) that MMR gene expression in blood leukocytes is extremely low and that a direct analysis on blood leukocyte samples is therefore not possible. However, after a short-term culture of leukocytes the expression level of the MMR genes is

boosted making it possible to measure MMR protein or mRNA levels. Unfortunately, large deviations in expression levels, both at the RNA as well as at the protein level were detected in proven mutation carriers and healthy controls. We therefore had to conclude that this method could not be used for the accurate detection of MMR mutation carriers.

Classification of missense mutations in Lynch syndrome

Finding a DNA variant is one issue, deciding whether a DNA variant is pathogenic, in other words whether it is contributing to the disease phenotype is another. This decision is easily made for mutations that result in premature termination of translation and thus in loss of function. For mutations that do not lead to a truncated protein this decision is much more difficult. This type of DNA variants, consisting mostly of missense mutations, is usually called unclassified variant (UV). These UVs form a significant proportion of the DNA variants found in Lynch syndrome (suspected) patients. They consist of DNA variants that give rise to single amino acid substitutions or small in-frame deletions (~10% of both *MSH2* and *MLH1* mutations). The question whether a UV contributes to the disease phenotype or merely represents a rare polymorphism, constitutes a major problem with obvious direct clinical consequences. In practice, it is difficult to determine pathogenecity. Most evidence is indirect such as segregation data, MSI, immunohistochemistry or absence of the UV in controls. We therefore believe that most proof for pathogenicity should come from functional assays. In chapter 3 all published functional assays applied to study MMR variants and the test outcome are reviewed. This information has been stored in an online database (www.mmrmissense.info) (see chapter 4).

Although the database contains data on over a 100 functionally tested UVs, a much larger set of UVs has never been tested. One subset of mutations never tested are those identified in *MLH3*. Until now, 18 *MLH3* germline mutations/UVs have been identified in colorectal cancer cases suspected of Lynch syndrome. Sixteen of these are UVs, all single amino acid substitutions, for which the pathogenic nature is yet unclear. In chapter 5, we functionally tested 8 *MLH3* missense variants that we had previously identified in Lynch syndrome suspected patients. We determined protein expression and stability, protein localization and interaction of the mutant MLH3 proteins with MLH1. All 8 mutant MLH3 proteins were expressed at levels comparable to wild type MLH3. Furthermore, the mutant proteins all localized normally to the nucleus and they interacted normally with wild type MLH1. We therefore have found no proof that the 8 missense variants tested are involved in colon cancer development and thus in Lynch syndrome.

Another subgroup of UVs that is underrepresented in the set of functionally tested UVs are those in *MSH6.* In chapter 6, functional assays have been

performed to evaluate the pathogenicity of 5 of such inherited *MSH6* UVs found in patients suspected of Lynch syndrome. The mutated MSH6 proteins, all containing single amino acid substitutions, were tested for expression and stability, interaction with MSH2, and for the subcellular localization. It was shown that protein expression of 4 of the 5 MSH6 mutants (S144I, A1021D, A326V, and T1219I) was significantly decreased after transfection when compared with expression of the wild type MSH6. Quantitative PCR analyses of hMSH6 mRNA after transfection showed that possibly this could be due to lower mRNA levels. No effects were observed on MSH2-MSH6 protein-protein interaction and the subcellular localization was normal for all 5 MSH6 missense mutants. Our data show that 4 of the 5 tested *MSH6* UVs seem to have an influence on gene expression and thereby on protein translation. Why the expression of these 4 UVs *in vivo* is reduced is yet unknown. These UVs might therefore be pathogenic. Further studies are needed to confirm our findings.

Chapter 9

Nederlandse samenvatting

In dit proefschrift beschrijven we enkele studies die gericht zijn op het classificeren en karakteriseren van missense mutaties, DNA varianten die op eiwitniveau resulteren in aminozuurveranderingen, in zogenaamde MisMatch Repair (MMR) genen, respectievelijk eiwitten. De bestudeerde mutaties zijn eerder gevonden in patiënten met Lynch syndroom of in patiënten die daarvan werden verdacht. Verder hebben we een methode getest die het mogelijk zou moeten maken dragerschap van een MMR gen mutatie te identificeren (in tumor vrije individuen) zonder direct uitgebreid DNA mutatie onderzoek te hoeven doen.

<u>Identificatie van MMR-gen mutatie dragers</u>

Identificatie van MMR-gen mutatiedragers en het laten participeren van mutatiedragers in preventieve screeningsprogramma's is bewezen succesvol te zijn, omdat het risico op het krijgen van dikkedarmkanker er door sterk afneemt en omdat het dikkedarmkanker gerelateerde sterfterisico sterk daalt. Vanwege deze waarnemingen krijgt het identificeren van mutatiedragers veel aandacht. We kunnen echter niet iedereen screenen op ziekteverwekkende mutaties in alle MMR-genen, dit is duur en kost relatief veel tijd. Daarom gaat de aandacht alleen uit naar mensen met een Lynch syndroom gerelateerde tumor en naar familieleden van dergelijke patiënten. Omdat in deze selecte groep mensen ook maar een klein deel een duidelijk ziekteveroorzakende MMR-gen mutatie heeft, zijn criteria (richtlijnen) opgesteld die moeten helpen bij het selecteren van patiënten en families die in aanmerking komen voor genetisch onderzoek. Deze criteria zijn gebaseerd op familiehistorie, leeftijd van de patiënt en op kenmerken van de tumor. De Nederlandse criteria met betrekking tot de tumorkarakteristieken zijn gebaseerd op: instabiliteit van microsatellieten (repeterende sequenties in het DNA), een kenmerk dat vrijwel altijd samen gaat met het verlies van functie van een MMR-gen; het verlies van een MMR-eiwit in de tumor (mogelijk als gevolg van een mutatie); methylatie van de promotor van één van de MMR-genen (*MLH1*); aanwezigheid van een mutatie in het *B-RAF* gen. Als een tumor microsatellietinstabiliteit laat zien (dit noemen we MSI-H) of als de tumor verlies laat zien van een MMR-eiwit, en dan met name van MSH6 en/of MSH2, wordt mutatieonderzoek van het DNA van de patiënt ingezet. Als de MLH1 eiwitkleuring negatief is of als een tumor promotormethylatie van *MLH1* laat zien en daarboven op ook een *B-RAF* mutatie heeft, dan wordt geen mutatie onderzoek verricht. Dit omdat het dan niet om een erfelijke tumor gaat. Patiënten met tumoren met een negatieve MLH1 kleuring zonder *B-RAF* mutatie worden wel verder onderzocht.

Als echter geen tumormateriaal aanwezig is, of als er geen DNA van een patiënt uit de familie kan worden onderzocht, dan rest slechts het screenen van alle MMR-genen, veelal in meerdere familieleden. Om dit te voorkomen hebben we aan een snelle, relatief goedkope methode gewerkt die het mogelijk zou moeten maken dragers van een MMR-genmutatie te identificeren aan de hand van de expressie van de MMR-genen in bloedcellen (leukocyten). De hypothese is dat dragers van een pathogene mutatie één van de twee kopieën van een gen heeft verloren en dat dit verlies leidt tot een halvering van de hoeveelheid RNA en eiwit. Het vinden van halvering van de expressie van een MMR-gen of van het MMR-eiwit zou richting geven aan de mutatiescreening. Dit zou tijd en geld besparen. In hoofdstuk 2 laten we zien dat de expressie van de MMR-genen extreem laag is in leukocyten. Door de leukocyten echter kort te kweken (3 dagen) gaat de expressie van de MMR-genen sterk omhoog waardoor deze goed meetbaar is. Als we de expressie bekijken, zien we helaas dat de niveaus van expressie erg variabel zijn. We vonden in controles soms zelfs een factor 2 verschil in expressieniveau. We moeten daarom concluderen dat deze methode niet bruikbaar is om dragers van MMR-genmutaties te identificeren.

Classificatie van missense mutaties in Lynch syndroom

Het vinden van een DNA-variant is één, het neme van een beslissing met betrekking tot de pathogeniciteit van de variant, met andere woorden, of de variant ook daadwerkelijk bijdraagt aan het ontstaan van de ziekte, is twee. De beslissing is eenvoudig te maken voor een mutatie die zorgt voor een verkort eiwit en dus voor een verlies van functie. Voor mutaties die niet leiden tot een verkort eiwit is het maken van een beslissing over betrokkenheid bij de ziekte een stuk lastiger. Dergelijke DNA-varianten, meestal DNA-veranderingen die leiden tot één aminozuurverandering of kleine in frame deleties of inserties, worden daarom ook wel unclassified variants, afgekort als UV's, genoemd.

UV's worden vrijwel even vaak gevonden als de duidelijk ziekteverwekkende mutaties en dan vooral in patiënten die verdacht worden van Lynch syndroom (ongeveer 10% van deze patiënten heeft een dergelijke UV). Omdat het niet duidelijk is of de betreffende DNA-variant bijdraagt aan de ziekte kan de UV niet gebruikt worden in de kliniek. Wat kan er worden gedaan om dit wel mogelijk te maken? Indirect bewijs kan worden verkregen door te kijken naar microsatellietinstabiliteit in de tumor van de persoon in wie de UV is gevonden. Ook kan gekeken worden naar de frequentie waarin de UV in controlespersonen voorkomt en zou gekeken kunnen worden naar segregatie van de UV met het optreden van kanker in leden van de familie waarin de UV is gevonden. Direct bewijs wordt hiermee echter niet verkregen. Wij denken daarom dat functionele testen, die wel een direct bewijs geven, hieraan zouden moeten worden

toegevoegd. In hoofdstuk 3 worden alle functionele testen die ooit zijn beschreven voor het testen van UV's in MMR-eiwitten uitgelegd en wordt een overzicht gepresenteerd van de gepubliceerde resultaten. De gegevens zijn online te bekijken op www.mmrmissense.info (Hoofdstuk 4).

De database zoals beschreven in hoofdstuk 4 bevat 115 UV's die zijn getest, echter dit is maar het topje van de ijsberg. De meeste gevonden UV's zijn nog niet getest. Een set van UV's die nog nooit zijn getest, zijn de UV's gevonden in *MLH3*. Tot nu toe zijn 18 *MLH3* UV's beschreven in families verdacht van Lynch syndroom. In hoofdstuk 5 beschrijven we functionele testen voor 8 van de 18 gevonden *MLH3* UV's. We hebben gekeken naar: eiwitexpressie en eiwitstabiliteit, naar de subcellulaire lokalisatie van mutante eiwitten en de interactie van de mutante MLH3-eiwitten met het wildtype MLH1-eiwit. Alle proeven laten zien dat de mutante eiwitten vergelijkbare resultaten laten zien als het wildtype MLH3. We concluderen daarom dat de 8 geteste MLH3 UV's zeer waarschijnlijk geen belangrijke rol spelen in de ontwikkeling van tumoren bij de mensen waarin de UV's zijn gevonden en dat de UV's niet oorzakelijk zijn voor het Lynch syndroom.

Een andere groep UV's die nog maar beperkt zijn onderzocht, zijn de UV's in MSH6. In hoofdstuk 6 hebben we functionele testen uitgevoerd op 5 MSH6 UV's die we eerder gevonden hadden in patiënten die verdacht werden van Lynch syndroom. De 5 missense varianten hebben we getest op expressie en stabiliteit van het mutante eiwit, op interactie met wildtype MSH2 en op de lokalisatie van het mutante eiwit in de cel. Alle eiwit-eiwit interactieproeven en de lokalisatieproeven lieten een normaal beeld zien (vergelijkbaar met wildtype MSH6). In de expressieproeven zagen we echter een verminderde expressie van 4 van de 5 MSH6 mutanten, te weten S144I, A326V, A1021D en T1219I, in vergelijking tot het wildtype MSH6 eiwit. Kwantitatieve PCR op MSH6 na transfectie liet zien dat de verminderde eiwitexpressie samengaat met een verminderde hoeveelheid *MSH6* mRNA. Het lijkt er dus op dat de 4 UV's invloed hebben op de hoeveelheid *MSH6* mRNA en daarmee op de hoeveelheid MSH6 eiwit. De reden waarom de expressie verlaagd is, is nog onduidelijk. Vervolgstudies zijn nodig om hier meer inzicht in te verkrijgen. We kunnen daarom nog geen definitieve uitspraak doen over de pathogeniciteit van deze UV's.

Chapter 10

Acknowledgements

Doing a Ph.D. in a foreign country is an absolute challenge. Four years ago, it was this challenge that motivated me to come to the Netherlands to perform my Ph.D. project. Because of all the sincere help from my colleagues and friends, and the collision of different cultures, this challenge finally turned out to be an exciting and unique journey in my life.

To my supervisor, Prof. Robert Hofstra: I am so grateful to have been one of your Ph.D. students and to have worked closely with you in the past 4 years. I trained as a tumor surgeon in China and made a big switch to become a researcher in Holland, but it has turned out to be a memorable experience to work in your lab. Your inspiration always lit my way, so that I could go further and avoid some of the perils in doing research; your enthusiasm led me through the 'winter time' of my research work; your trust ensured me of the great freedom to try my best; certainly your Dutch Samenvatting has enabled me to finish my dissertation on time.

I am also grateful to my other supervisor, Prof. J.H. Kleibeuker, who gave me many precise and constructive comments and suggestions for my manuscript, and to Dr. Rolf Sijmons, who always ignited insightful sparks to inspire me: your versatile mind reminds me that a researcher's life can be colorful.

Prof. Charles Buys: thank you for your advice and encouragement and attention to my research. Your gift of longjin tea (which you brought back from your journey to China) made me experience a slight nostalgia for my home country. Furthermore, I am very honored by the fact that you are willing to be a member of my reading committee, which of course also goes for Prof. Morreau and Prof. Rasmussen as well.

Dr. Bart Eggen and group members from the Department of Developmental Genetics in Haren: thanks for all your support and help in the initial stage of this project.

To my colleagues in the research group of the Department of Genetics, particularly Frans, Helga and Renée: thanks for all your sincere help and I appreciated all the beneficial discussions with you. You were never too busy to help me. Renée: thanks for the collaboration in our projects and all the fun we had in the squash competition (de Nonsense Mutanten squashteam), and of course thanks for all Dutch words (like bitterbal) that you taught me. To Pieter, Krista Bos, Ludolf, Jan Osinga: I thank all of you

for the time that you spent with me, you were never too busy to teach me lab techniques. My sincere appreciation goes to Paul and Krista Kooi, two technicians working on this project: your fundamental contribution speeded up my research and ensured this project could be finished in good time. I thank Ellen, Klaas and Martin, with whom I have had constructive discussions when I met bottlenecks in my project. To my roommates, Renée, Ana (we had a nice port borrel time), Helga, Gerben, Tjakko and Jihane: you have made the office a happy place to be. Tjakko, Gerben: we had fun discussing Dutch cultural topics (like how to make Haring). To the group of Ph.D. students, Maria, Yunia, Gosia, Agata, Mats, Ivan and Greg: thanks for the cheerful times we were together. I hope we will soon have a few more Ph.D. parties! A special word of thanks goes to Janet and Wendy, two M.Sc. students involved in the subcloning work for this project: your work gave this project a good start.

To the dear members of the DNA diagnosis group, Yvonne, Henny, Jan, Bart, Monique, Jos, Gerwin, Bea, Chantal, Jelko and Annemieke: heartfelt thanks for all your help, I really appreciated the team spirit in the DNA group.

To the cytogenetics group members and geneticists-in-training: Klasien, Inge, Nettie, Nancy, Ellen and Annie-Marie: without your help, it would be impossible to carry out blood culture smoothly. Annie-Marie, Karin and Erica: thanks you for your cooperation on chapter 2.

To Prof. Cisca Wijmenga, our new head of department, and Jackie Senior, who organized the research retreat and SWOT analysis: thanks for your strategic suggestions for career development. Jackie, thanks for all your cordial help and rich culture knowledge when I tried to use ‘standard’ English.

Henk, Mentje, Ria and Bote, whenever I went to your office, I always received a satisfactory reply and help. Bote, thanks for teaching me book binding skills.

To the Department of Oncology, Roelof, Xiangyi Li and Nang Yan, cordial thanks for your assistance when I used facilities in your group. Roelof, my good friend, the times that we were together were so cheerful and you will stay forever in my memory.

To Prof. Lene Rasmussen and group members, Merete, Anne, Sasha, Anne Maria, and Sofie, from the Department of Science, Systems and Models in Denmark. I had a good time when we met up, while I learned

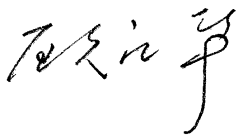
from you and enjoyed your hospitality in such a fairytale country. Merete, are my turnips in your garden still doing well?

To Dr. Niels de Wind, thanks for your useful comments on the review (see chapter 3) and the extra work that now needs to be incorporated into chapter 6 which describes our functional work on the MSH6 mutations.

To Sieta and Kees and your lovely Hidde and Maaike, my family friends, whenever we met difficulties, you are always standing beside me and give me a hand. Your invitation to a Dutch pub or to a Dutch New Year dinner at your home, and your concern for my life in Groningen showed me another side of Dutch culture, its hospitality and home-basis.

To my dear Chinese friends in Groningen, your names will, in the future, remind me of the joys and the sorrows we have shared in Groningen: Yi Liao, Youchun Zhang, Ning Qu and Yijin Ren, Chaohong Zhou and Jianhong Zhang, Weidong Li and Jichi Zheng, Youjun Zhao and Liqiang Qi, Yin Wan and Chao Zhou and TianZhuo Wang, Ping Gao, JinYuan Fu, Liang Zhang. A friend in need is a friend indeed.

I owe my parents and parents-in-law a lot. They have always supported me to their best ability, and their confidence in me has shaped me to face all kinds of challenges in an optimistic way. Qiwei, my lovely son, when I took you to the Netherlands, I was initially worried about the challenges that you would meet: with no Dutch or English language background, no friends. And now, after 3 years in the Netherlands, you have made such achievements that we are very proud of you: you speak fluent Dutch and English but have not stopped learning Chinese - your home language and one of the most difficult languages in the world -, you enjoy your school life, have the courage to face challenges, are open minded and optimistic, and have made a lot of Dutch friends. To my dear wife Xiao Mei, thanks for your love and care in all these years. Life is not easy, but our persistence together is always the best way to handle difficulties. This book is dedicated to all of you.



Jianghua Ou
January 2008
Groningen