Pathogenicity, functional significance and clinical phenotype of mismatch repair gene MSH2 variants found in cancer patients

Saara Ollila

Division of Genetics
Department of Biological and Environmental Sciences
Faculty of Biosciences
University of Helsinki

Academic dissertation

To be publicly discussed, with the permission of
The Faculty of Biosciences, University of Helsinki,
in the auditorium 1041 of the Biocenter II, Viikinkaari 5, Helsinki,
on the 19th of September 2008 at 12 o'clock noon.

Supervisor Docent Minna Nyström
Department of Biological and Environmental Sciences
University of Helsinki, Finland

Reviewers Professor Klaus Elenius
Department of Medical Biochemistry and Molecular Biology
University of Turku, Finland

Docent Helmut Pospiech
Leibniz Institute for Age Research
Fritz Lipmann Institute, Jena, Germany

Opponent Professor Torben F. Ørntoft
Department of Clinical Biochemistry
Aarhus University Hospital, Denmark

ISSN 1795-7079
ISBN 978-952-10-4914-9
ISBN 978-952-10-4915-6 (pdf)

Helsinki 2008, Helsinki University Printing House

Nothing in life is to be feared. It's just to be understood.

Marie Curie

CONTENTS

LIST OF ORIGINAL PUBLICATIONS 6
ABBREVIATIONS 7
ABSTRACT 9
INTRODUCTION 11
REVIEW OF THE LITERATURE 13
Cancer 13
Cancer genetics 14
Oncogenes and tumour suppressor genes 14
Hereditary cancer 15
DNA maintenance 17
Origins of mutagenesis 17
DNA repair pathways 17
Mismatch repair 22
MMR in *Escherichia coli* 23
MMR in eukaryotes 24
Hereditary nonpolyposis colorectal cancer 31
Genetics of HNPCC 31
Clinical characteristics of HNPCC patients 33
Muir-Torre syndrome 35
Nontruncating mutations in HNPCC 35
Frequency of nontruncating mutations in MMR genes 35
Functional analysis of nontruncating MMR gene variants 36
AIMS OF THE PRESENT STUDY 39
MATERIALS AND METHODS 40
MSH2 mutations and associated families (I-IV) 40
Functional assays (I-IV) 43
Protein expression and coimmunoprecipitation (I-IV) 43
Protein purification (III-IV) 45
Mismatch repair assays (I-IV) 46
Bandshift assays (III-IV) 49

RESULTS 51
Effects of the studied mutations on MSH2 expression, purification and interaction with MSH6 (II-IV) 51
Effects of MSH2 mutations on mismatch repair and binding and release of mismatches (I-IV) 52
Correlation between the functional assays and clinical phenotypes of mutation carriers (I-IV) 57
DISCUSSION 59
MSH2 L187P and C697F predispose to HNPCC and MTS 60
No evidence for MSH2 N127S and G322D -linked cancer predisposition 61
The majority of the studied MSH2 variants inactivated MMR 62
Decreased protein expression was associated with MMR deficiency in N-terminal MSH2 missense mutations 62
Pathogenic mutations in the ATPase domain mostly interfered with mismatch binding or release 64
MSH2 G674A displayed reduced capability to interact with MSH6 65
Phenotypic characteristics of the mutation carriers correlated with the functional data 66
IHC as a screening tool for HNPCC 67
Classification of nontruncating MSH2 mutations 67
CONCLUSIONS 69
FUTURE PROSPECTS 70
ACKNOWLEDGEMENTS 71
REFERENCES 74

LIST OF ORIGINAL PUBLICATIONS

I Ollila S, Fitzpatrick R, Sarantaus L, Kariola R, Ambus I, Velsher L, Hsieh E, Andersen MK, Raevaara TE, Gerdes AM, Mangold E, Peltomäki P, Lynch HT, Nyström M. The importance of functional testing in the genetic assessment of Muir-Torre syndrome, a clinical subphenotype of HNPCC. *Int J Oncol*. 2006 Jan;28(1):149-53.

II Ollila S, Sarantaus L, Kariola R, Chan P, Hampel H, Holinski-Feder E, Macrae F, Kohonen-Corish M, Gerdes AM, Peltomäki P, Mangold E, de la Chapelle A, Greenblatt M, Nyström M. Pathogenicity of MSH2 missense mutations is typically associated with impaired repair capability of the mutated protein. *Gastroenterology*. 2006 Nov;131(5):1408-17.

III Ollila S, Dermadi Bebek D, Greenblatt M, Nyström M. Uncertain pathogenicity of MSH2 variants N127S and G322D challenges their classification. *Int J Cancer* 2008 Aug 1;123(3):720-4.

IV Ollila S, Dermadi Bebek D, Jiricny J, Nyström M. Mechanisms of pathogenicity in human MSH2 missense mutants. *Human Mutation,* in press.

ABBREVIATIONS

ADP	Adenosine diphosphate
ATM	Ataxia-Telangiectasia mutated
ATP	Adenosine triphosphate
ATR	ATM and Rad3-related
APC	*Adenomatous polyposis coli*
BER	Base excision repair
cDNA	Complementary DNA
CHK1, 2	Checkpoint kinase 1, 2
CRC	Colorectal cancer
CSA, B	Cockayne syndrome A, B
C-terminus	Carboxy terminus
DSBR	Double-strand break repair
dsDNA	Double-stranded DNA
EC	Endometrial carcinoma
E. coli	*Escherichia coli*
EXO1	Exonuclease 1
FAP	Familial adenomatous polyposis coli
FPLC	Fast protein liquid chromatography
GGR	Global genome repair
HNPCC	Hereditary nonpolyposis colorectal cancer
IDL	Insertion / deletion loop
IHC	Immunohistochemistry
InSiGHT	International Society for Gastrointestinal Hereditary Tumors
IR	Ionizing radiation
LOH	Loss of heterozygosity
NE	Nuclear protein extract
NER	Nucleotide excision repair
NHEJ	Non-homologous end joining
Ni-NTA	Nickel-nitrilotriacetic acid
MLH1, 3	MutL homolog 1, 3

MMR	Mismatch repair
MNNG	*N*-methyl-*N*'-nitro-*N*-nitrosoguadinine
MSH2, 3, 6	MutS homolog 2, 3, 6
MSI	Microsatellite instability
MSS	Microsatellite stable
N-terminus	Amino terminus
PCNA	Proliferating cell nuclear antigen
PMS1, 2	Human postmeiotic segregation increased homolog 1, 2
PMSF	Phenylmethylsulfonyl fluoride
RFC	Replication factor C
RPA	Replication protein A
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
Sf 9	*Spodoptera frugiperda* 9
SIFT	Sorting intolerant from tolerant
SNP	Single nucleotide polymorphism
ssDNA	Single-stranded DNA
TCR	Transcription-coupled repair
TE	Total protein extract
WT	Wild type
XP	Xeroderma pigmentosum

ABSTRACT

Hereditary nonpolyposis colorectal cancer (HNPCC) is a hereditary cancer syndrome, which manifestates with high penetrance in early middle age, mainly with colorectal and endometrial tumours. Susceptibility for HNPCC is dominantly inherited with germline defects in the mismatch repair (MMR) genes *MLH1*, *MSH2*, *MSH6* and *PMS2*. While a truncating mutation in one of these genes leads to deficient MMR, thereby promoting genetic instability and tumour formation, a nontruncating mutation can either be a completely neutral variation or lead to a highly increased cancer risk and HNPCC. The phenotypic effects of nontruncating mutations are impossible to predict based on genetic evidence alone. The correct determination of the pathogenicity of different mutations is, however, very important, as the verification of the causative mutation enables genetic counselling and surveillance of mutation carriers, which has been shown to lead to significantly lowered mortality.

The most frequent nontruncating mutations are missense mutations, which alter only one amino acid in the protein. Unlike in *MLH1*, where missense mutations have been characterised extensively, functional studies on nontruncating *MSH2* mutations are rarer. *MSH2* is the second most commonly mutated HNPCC susceptibility gene and defects in it account for 39% of all identified HNPCC mutations. Seventeen percent of all identified *MSH2* variations are of the missense type. The aim of this PhD thesis was to gather functional evidence on the pathogenicity of patient-derived nontruncating *MSH2* variants. We assessed the functionality of 18 mutations and correlated the site of the mutation to the biochemical and phenotypic effects of the mutated protein. The proteins corresponding to the original genetic *MSH2* variants were expressed and purified. The expression level, MMR efficiency, interaction with MSH6, mismatch binding, and mismatch release capabilities of the protein variants were studied. The results of the functional assays were compared to the clinical characteristics of the mutation carriers.

Twelve of the studied eighteen mutations were found to exhibit severe defects in the functional assays, supporting the hypothesis that these mutations were the underlying cause of the cancer phenotype in mutation carriers. In addition, two mutations reduced but did not abolish the function of the protein. Four mutations showed no or only minor defect in the assays. The characterisation of the biochemical defects revealed different

mechanisms through which the pathogenic effects were mediated. The majority of the MMR-deficient mutations which were located in the amino-terminal domains of the MSH2 polypeptide demonstrated defects in the protein expression level. Most of the carboxy-terminal mutations, situated in the ATPase domain, had an impact on the ability of the protein to bind or release mismatched DNA. When comparing the biochemical data to the tumour phenotype, a significant correlation between the functional deficiency *in vitro* and lack of expression of the corresponding protein in the tumour tissue was observed.

The analyses demonstrated that the location of the mutation may affect not only the biochemistry of MMR but also the phenotype of *MSH2* mutation carriers. This study significantly contributed to the knowledge of *MSH2*-associated HNPCC tumorigenesis, thereby facilitating the diagnostics and counselling of the associated families. In addition, the study confirmed and supplemented the prevailing knowledge of the biochemical functions and characteristics of different MSH2 domains.

INTRODUCTION

In 1993, the molecular background of a familially clustered cancer syndrome, hereditary nonpolyposis colorectal cancer (HNPCC), was revealed to be associated with germline mutations in genes encoding DNA mismatch repair (MMR) proteins (Peltomäki *et al.* 1993a). Due to the great clinical importance of HNPCC, this breakthrough led to intense research on MMR. To date, MMR is well characterised and germline defects in four MMR genes, *MLH1*, *MSH2*, *MSH6* and *PMS2*, have been shown to predispose to HNPCC (Peltomäki 2005, Woods *et al.* 2007).

Tumorigenesis in HNPCC results from genetic instability, which reflects the loss of the postreplicative DNA repair activity displayed by MMR. This can be observed as the microsatellite instability (MSI), which is the hallmark of MMR-deficient tumours (Aaltonen *et al.* 1993). The most typical tumours in HNPCC syndrome are tumours of the colon, rectum and endometrium, whereas other types, such as hepatobiliary, small bowel, gastric, ovarian and brain cancers occur more rarely yet being more frequent than in the general population (Watson, Riley 2005). The average age of cancer onset in HNPCC is about 45 years, whereas most sporadic colorectal cancers have a typical onset some twenty years later (Lynch, de la Chapelle 1999a). HNPCC has a very high penetrance, and the lifetime risk of developing cancer in *MLH1* and *MSH2* mutation carriers is close to 100%. The penetrance is somewhat lower in *MSH6* and *PMS2* mutation carriers (Peltomäki 2005).

Thanks to intensive research and highly developed cancer surveillance systems, a large proportion of the HNPCC related malignancies, especially the colorectal ones, can be removed already at an early stage. Thus, in countries, such as Finland, where genetic counselling and cancer surveillance are efficient, HNPCC-related mortality is low (Mecklin *et al.* 2007). However, the efficient screening and counselling of HNPCC patients can only be applied if the predisposing mutation is characterised.

There are several factors that make HNPCC diagnostics challenging. Colorectal cancer (CRC) is the third most common cancer in the Western world. It accounts for 10% of all diagnosed cancers, thus affecting up to 150 000 people in the US and 2 500 people in Finland in a year (Jemal *et al.* 2008, www.cancerregistry.fi). HNPCC accounts for only

2 – 3% of all CRC cases (Lynch, de la Chapelle 2003, Salovaara *et al.* 2000). There are no clear clinical features separating hereditary from sporadic CRC. Traditionally, HNPCC diagnostics has been done using information on the familial clustering and the early age at onset as criteria (Vasen *et al.* 1991, Vasen *et al.* 1999). This approach, however, leaves many HNPCC cases unnoticed, when information about family members is lacking or when the family is too small to fulfil the diagnostic criteria. MSI and immunohistochemistry (IHC) studies on genetic instability and MMR protein expression in the tumour, respectively, give indications as to the MMR defect, but as such do not provide evidence of its heredity. Therefore, mutational analysis is a prerequisite for reliable diagnostics. The considerable sizes of the four predisposing genes make mutational analysis laborious. Furthermore, all genetic variations found in MMR genes are not associated with cancer predisposition, creating further challenges to HNPCC diagnostics.

Missense mutations, which lead to single amino acid alterations, and small in-frame insertions and deletions, may change the structure of the protein only slightly and associate with functional MMR and no increased cancer predisposition. Alternatively, they can inactivate MMR and lead to HNPCC. When reliable data of the co-segregation of the cancer phenotype and genetic variation is not available, the pathogenicity and phenotypic outcome of a nontruncating variant is impossible to predict. In those cases, only functional analysis can aid in determining the pathogenicity of the mutation.

The aim of this PhD thesis was to study the pathogenicity, functional significance and clinical phenotype of nontruncating variants in the *MSH2* gene. *MSH2* is the second most common predisposing gene for HNPCC, and the studied mutations were found from cancer patients. Our findings demonstrate that most of the studied mutations indeed affect MMR, and that the pathogenicity of the mutation is mediated through different mechanisms, depending on the location of the mutation in the MSH2 protein. The results of this study facilitate the genetic counselling of all carriers of the studied mutations, especially those whose pathogenicity was ascertained by our work, confirming the HNPCC diagnosis.

REVIEW OF THE LITERATURE

CANCER

The human body is composed of nearly 10^{14} cells. In order to maintain the appropriate homeostasis of an individual, the differentiation, division and death of all those cells must occur in a highly controlled manner, and failures in this regulation may lead to the formation of tumours. Tumours are characterised by uncontrolled growth of cells, resulting in cell division at an abnormal time or rate or in an abnormal space. Cancer is a group of diseases characterised by malignant tumours, which are differentiated from the benign ones by their ability to invade adjacent tissues.

Cancer is the second most prevalent cause of death in Western countries. It develops slowly and thus affects mainly elderly people. Therefore, cancer incidence is higher in countries where life expectancy in general is high. In Finland, about 27 000 cancers were diagnosed in the year 2006 (Finnish Cancer Registry, www.cancerregistry.com), and the estimate number of cancers in the US for 2008 is nearly 1 500 000 (Jemal *et al.* 2008). In males, the most prevalent cancers occur in the prostate, lung or bronchus, colorectum, and urinary bladder. In females, breast cancer is the most prevalent, followed by colorectal, lung and uterine cancer (Jemal *et al.* 2008, American Cancer Society, www.cancer.org). The same cancer types are prevalent in all Western countries. In total, more than one-third of the population develops a cancer at some point in life, and the general survival rate 5 years from cancer diagnosis is about 66%. (Jemal *et al.* 2008, www.cancer.org.) Because of its high incidence, and vast effects on society both in the form of human suffering and costs to health care, cancer research is one of the most intensive focuses of study in modern biology. The aim of the research is to understand the processes of cancer development and, thereafter, be able to diagnose the tumours earlier, apply more efficient cancer therapy, and, eventually, to be able to prevent tumour formation.

Cancer genetics

Cancer in general is characterised as atypical cell growth resulting from abnormalities in cellular regulation. Most of the abnormalities are consequences of alterations in DNA, the molecule which holds the information on how the cell is built and maintained. Cancer results from the accumulation of defects in genes which regulate cellular homeostasis and growth. For a cell population to become cancerous, it needs to fulfil several requirements which are not met by normal cells: self-sufficiency in growth signals, non-responsiveness to anti-growth signals, avoidance of apoptosis and senescence, formation of vasculature, and capacity for tissue invasion and metastasis (Hanahan, Weinberg 2000). Furthermore, evading the body's own immune response has in recent years been show to be an important step in tumour pathogenesis (Drake, Jaffee & Pardoll 2006). Fulfilling all these conditions requires that several genetic changes take place. Therefore, malignant transformation is believed to occur sequentially through a process where certain genetic alterations give cells a growth advantage, allowing them to expand more efficiently as compared to normally regulated cells, and subsequently acquire more alterations. This can be seen as Darwinian evolution in the cell population: the most efficiently growing cells survive best. (Weinberg, 2007)

Oncogenes and tumour suppressor genes

Genes which participate in tumorigenic processes are divided into two main classes: oncogenes and tumour suppressor genes. In general, proto-oncogenes possess growth-promoting effects under normal circumstances. Transformation of proto-oncogenes to oncogenes can occur through activating mutations, increased expression, or gene amplification. Thus, oncogenes increase the cell's growth potential by gene activation. Their tumour-promoting effect is dominant, already affecting cell growth with one altered allele. Oncogenes are typically genes which encode players in signal transduction pathways, such as growth factor receptor tyrosine kinases (*e.g.* epidermal growth factor receptors, ErbB1-4), signal transduction molecules (Ras, Raf), transcription factors (Myc) or anti-apoptotic proteins (Bcl-2). Many oncogenes, such as *Src*, *Ras*, and *Myc*, have been identified *via* their viral homologs, which have been found to promote tumorigenesis upon viral infection (Diehl, Keller & Ignatoski 2007).

Tumour suppressor genes possess growth-limiting functions. In their case, the growth advantage is acquired by gene inactivation, and in most cases both alleles need to be inactivated before the effect takes place. Therefore, their effect in tumour progression is recessive. Tumour suppressor genes are further divided into so-called gatekeepers, whose function is to regulate the cell division, and caretakers, which look after the integrity of DNA. Typical gatekeepers are for example *retinoblastoma (Rb)* and *INK4a*, which regulate the G1-S cell cycle checkpoint; pro-apoptotic genes of the *BAX* family; and regulators of growth-promoting molecular pathways, such as a*denomatous polyposis coli* (*APC*) (Sherr 2004). Caretaker genes encode proteins which participate in the maintenance of DNA. Absence of this action leads to increased mutagenesis and therefore an increased occurrence of subsequent alterations in proto-oncogenes and tumour suppressors. Reflecting the importance of DNA integrity, to date over 100 proteins with a role in DNA maintenance have been described (Christmann *et al.* 2003), and defects in many of those are connected to cancer formation. This will be discussed in detail in later chapters.

Hereditary cancer

Despite being considered a disease of the genome, the great majority of cancers are not hereditary. However, several rare syndromes characterised by the familial inheritance of cancer predisposition in a (near-) Mendelian manner have been identified. These inherited cancer syndromes are very important areas of study mainly because of two reasons. Firstly, identification of the genetic component predisposing to cancer in a family allows diagnosis and surveillance of the other mutation carriers, and leads to relief from the fear of a high cancer risk in non-carriers (Aktan-Collan *et al*. 2000). Secondly, inherited cancer syndromes provide starting points for understanding the genetic components involved in the regulatory pathways which, when altered, may contribute to cancer formation. Thus, information derived from studies concerning hereditary cancers can be applied to the management of all cancers (Fearon 1997).

In hereditary cancer syndromes, the resulting tumour usually develops at an earlier age as compared to the corresponding sporadic cancers, reflecting the skipping of one step in the chain of somatic mutations needed for tumour development. This skip is a result of a germline alteration, usually in a tumour suppressor gene. The altered gene can predispose to cancer in a dominant or recessive mode. However, also the dominantly

inherited cancer syndromes are believed to act in a recessive manner at the cellular level, elucidating the requirement for a somatic mutation in the second allele. This “second hit” -hypothesis was first presented in the context of retinoblastoma, a cancer of the eye. Retinoblastoma is inherited through an inactivating mutation in one allele of the tumour suppressor gene *Rb*, which plays an important role in regulation of the G1-S cell cycle checkpoint. Based on his observations of retinoblastoma patients, Alfred G. Knudson created his famous model for the formation of a hereditary cancer (Knudson 1971). In the two-hit hypothesis, Knudson proposed that both copies of the *Rb* tumour suppressor gene have to be inactivated in a cell before the cell acquires a growth advantage. In familial cases, one allele is inactivated already in the germline, and when the second copy is lost by somatic inactivation, tumorigenesis is initiated. In sporadic retinoblastoma, both alleles need to be somatically inactivated. Later, this hypothesis has been expanded to many tumour suppressor-associated cancer syndromes (Knudson 1996).

The most common hereditary cancers associated with germline defects in tumour suppressor genes are listed in Table 1. The oncogenes associated with familial cancer are *RET*, *MET*, and *CDK4*, which, when mutated in the germline, predispose to multiple endocrine neoplasia type 2A and 2B, hereditary papillary renal cell carcinoma, and familial melanoma syndromes, respectively (Marsh, Zori 2002).

Table 1. The most common hereditary cancers associated with germline defects in tumour suppressor genes. (Fearon 1997, Marsh, Zori 2002)

Gene	Syndrome	Primary tumour	Function of the gene(s)
APC	Familial adenomatous polyposis	Colorectal cancer	β-catenin regulation
ATM	Ataxia telangiectasia	Lymphoma	DNA damage response
BLM	Bloom's syndrome	Solid tumours	DNA helicase
BRCA1	Familial breast and ovarian cancer	Breast and ovarian cancer	DNA damage response
BRCA2	Familial breast cancer	Breast cancer	DNA damage response
CDKN2	Familial melanoma	Melanoma	Cell cycle regulation
FANC1-12	Fanconi anemia	Leukemia	DNA crosslink repair
LKB1	Peutz-Jeghers syndrome	Gastrointestinal tract cancer	Serine-threoninen kinase
MLH1, MSH2, MSH6, PMS2	Hereditary nonpolyposis colorectal cancer	Colorectal cancer	Mismatch repair
NF1	Neurofibromatosis type 1	Neurofibromas	RAS regulator
NF2	Neurofibromatosis type 2	Acoustic neuromas, meningiomas	Cell adhesion and cytoskeleton
p53	Li-Fraumeni syndrome	Sarcomas, breast cancer	DNA damage response
PRKAR1A	Carney complex syndrome	Pituitary adenoma	cAMP pathway
PTCH	Nevoid basal cell carcinoma syndrome	Basal cell skin cancer	Hedgehog signalling receptor
PTEN	Cowden disease	Breast and thyroid cancer	Tyrosine phosphatase
RB1	Familial retinoblastoma	Retinoblastoma	Cell cycle regulation
SMAD4	Juvenile polyposis coli	Colorectal cancer	TGF-β signalling mediator
VHL	Von Hippel-Lindau syndrome	Renal cancer	Fibronectin matrix assembly
WT1	Wilms tumor	Paediatric kidney tumours	Transcriptional regulation
XPA-G	Xeroderma pigmentosum	Skin cancer	Nucleotide excision repair
XPV	Xeroderma pigmentosum	Skin cancer	Translesion synthesis

DNA MAINTENANCE

Origins of mutagenesis

If cancer is considered a disease resulting from cumulative genetic alterations, then how do these alterations come about? DNA, as well as other molecules in the cell, is at all times exposed to several damaging agents, which alter its chemical features. DNA is the guidebook for building all other cellular molecules and it exists only in two functional sets in each diploid cell. Both copies are necessary. DNA molecules are irreplaceable, and therefore most sensitive to damage. The sources of DNA-damaging agents can be endogenous, originating from the cell's own metabolism, or exogenous, deriving from outside of the body. The most significant source of endogenous DNA damage is reactive oxygen species, which are unavoidable byproducts of oxidative metabolism. Exogenous DNA damage is caused *e.g.* by ionising radiation (IR), such as X-rays or the high-energy radiation resulting from radioactive decay, UV radiation from the sun, and chemical carcinogens, such as those derived from tobacco smoke or food. These agents cause a wide variety of chemical modifications in DNA. Importantly, also spontaneous chemical reactions, such as deaminations, depurinations and depyrimidations take place frequently, destabilizing DNA even in the absence of any particular genotoxic stress.

In addition to chemical modifications, which alter the structure of DNA bases, faulty insertions of structurally perfect bases occur rarely but steadily in the course of DNA replication. Both types of mutagenesis promote tumorigenesis by altering the properties of functionally important genes. Because of the extreme importance of DNA stablility, and the vast spectrum of lesions which destabilize it, several repair pathways and damage responses have evolved to maintain the integrity of DNA (Reviewed *e.g.* in Rouse, Jackson 2002, Christmann *et al.* 2003, Hakem 2008). Supporting the idea of increased mutability leading to cancer formation, inborn defects in many of these DNA repair systems lead to a predisposition to hereditary cancer syndromes.

DNA repair pathways

Nucleotide excision repair

A link between DNA repair and cancer was first established when it was shown that cells of xeroderma pigmentosum (XP) patients, who suffered from sensitivity to sunlight and a predisposition to cancer, were unable to repair DNA lesions after exposure to UV light

(Cleaver 1968). The XP syndrome is inherited in an autosomal recessive manner in genes named *XPA-G* and *V*. These XP-associated genes consist of components of a specific DNA repair system called nucleotide excision repair (NER). NER defects are also associated with a variety of segmental progeria syndromes, connecting DNA repair defects not only to cancer but also to ageing (Andressoo, Hoeijmakers & Mitchell 2006).

NER recognises and repairs a variety of DNA adducts, which cause distortions to the DNA helix, such as UV irradiation-induced pyrimide dimers and 6-4-photoproducts, and bulky adducts caused by chemical mutagens. NER is functionally divided to two distinct pathways, transcription-coupled repair (TCR) and global genome repair (GGR) (Reviewed in Fousteri, Mullenders 2008, Shuck, Short & Turchi 2008). In TCR, the proteins Cockayne syndrome A (CSA) and CSB are required for lesion recognition, which occurs when the elongating RNA polymerase II gets blocked at the site of DNA damage. Therefore, TCR is limited to the template strand of actively transcribed regions of DNA. In GGR, the lesion recognition component is hHR23B/XPC, and GGR repairs DNA without strand bias. Following lesion recognition, the NER machinery shares the same components in both subpathways. The transcription factor IIH (TFIIH) complex is recruited to the site of the lesion and the XPB and XPD helicase subunits of TFIIH unwind the DNA. XPA outlines the site of repair and assembles the remaining essential NER machinery on the site. The defective strand is incised by endonucleases XPG and XPF/ERCC1 at the 5' and 3' ends of the lesion, respectively, and the resulting gap is filled by DNA polymerases and the backbone sealed by DNA ligase I. In total, NER reaction involves over 25 distinct enzymes (Aboussekhra *et al.* 1995).

Base excision repair

Base excision repair (BER) is mainly responsible for the recognition and correction of oxidised and alkylated bases, resulting from cellular metabolic events and IR, and the correction of abasic sites resulting from spontaneous depurination and depyrimidation events. BER also addresses DNA bases arising from deamination reactions, which for example convert cytosine into uracil, giving rise to C→T / G→A transitions if not corrected. Furthermore, BER recognises some DNA mispairs, such as G•T mispairs, which result from the above-mentioned cytosine deamination (Hegde, Hazra & Mitra 2008). The BER machinery is initiated by glycosylases, each of which recognizes a specific type of DNA lesion. For example, OGG1 and OGG2 recognize oxidised bases,

TDG detects T and U in T•G and U•G mispairs, UDG uracil, and MYH adenine in 8-oxo-G•A mispairs. The glycosylases detach the incorrect base from the deoxyribose backbone of the DNA molecule, leaving behind an abasic site. Abasic site endonuclease (APE) then cuts the strand to be repaired 5' from the abasic site, and DNA polymerase β inserts the correct base to the site of repair. The DNA strand is sealed by DNA ligase. In addition, so-called long-batch BER, which removes and resynthesizes 4 – 7 bases around the lesion, has been described (Frosina *et al.* 1996). Defects in BER are also connected to cancer predisposition, as biallelic mutations in MYH can lead to multiple colorectal adenomas and carcinomas (Al-Tassan *et al.* 2002, Sieber *et al.* 2003).

Double-strand break repair

DNA double-strand breaks (DSBs) are generated for example by IR or oxidative damage. They can also form due to the collapse of the replication fork when the replication machinery encounters single-strand breaks or damaged bases. DSBs are a very severe form of DNA damage, and even one such break can cause cell death (Rich, Allen & Wyllie 2000). Unrepaired DSBs or incorrect repair leads to chromosome fusions, deletions and translocations, which are typical rearrangements in cancer cells (Jackson 2002). Two mechanisms are responsible for double-strand break repair (DSBR): homologous recombination (HR) and non-homologous end joining (NHEJ).

HR is an error-free repair system, which processes DNA breaks using the intact identical sister chromatid or, more rarely, the homologous chromosome as the template to rescue the DSB and to construct an intact DNA molecule. Therefore, HR takes place mainly in the S or G2 phases of the cell cycle, when the sister chromatid is available. Also, HR is believed to account for the processing of most if not all DSBs associated with replication fork collapse, because in those cases only one free dsDNA end emerges, making it impossible for the classical NHEJ pathway to repair the lesion (see below).

The MRN complex, consisting of MRE11, Rad50 and NBS1 proteins, is believed to process the free DNA ends in HR to create single-stranded overhangs, whose ends are then bound by Rad52 (Stasiak et al. 2000). RPA coats the single-stranded regions; and Rad51 forms nucleoprotein filaments on ssDNA to promote strand exchange. Rad52 and Rad54 promote the homology search and strand-exchange events of Rad51-coated ssDNA with the complementary DNA strand. Strand invasion is followed by branch migration, gap filling and resolving of the intermediate structures to give rise to two intact DNA

molecules (reviewed *e.g.* in Helleday 2003, Li, Heyer 2008). Among the various other proteins involved in HR are BRCA1 and BRCA2, mutations in which predispose to familial breast and ovarian cancer (Fackenthal, Olopade 2007).

HR is, in addition to functions in DSBR, also involved in the processing of intrastrand crosslinks (ICLs), which are detrimental DNA lesions leading to the blockage of transcription and replication (Dronkert, Kanaar 2001). In response to ICLs, proteins of the Fanconi Anemia (FA) pathway are important for the initiation of Rad51-mediated HR. Patients carrying mutations in genes involved in this pathway (altogether 12 identified *FANC* genes) are prone to cancers, such as acute myeloid leukemia and squamous cell carcinoma, and the hallmark of FA patient cells is sensitivity to DNA intrastrand crosslinking agents, such as Mitomycin C (Patel, Joenje 2007).

The NHEJ pathway is used for DSBR in the G0 and G1 phases of the cell cycle, when the sister chromatid templates for HR are not available. In the NHEJ reaction, the MRN complex processes the free DNA ends, followed by DNA end binding by Ku (Ku70-Ku80 complex). Then, Ku binds the DNA-dependent protein kinase catalytic subunit (DNA-PKcs), forming an enzyme complex called DNA-PK. DNA-PK activates a complex of XRCC4 and ligase IV, which link and ligate the broken DNA ends together (reviewed in Weterings, Chen 2008). NHEJ is error-prone, as deletions occur due to the degradation of the DNA ends in the search for microhomology before the two DNA ends can be joined. An exception, however, are the cases where the two DNA ends have complementary overhangs, such as when the DNA break is induced by nucleases in V(J)D recombination or class-switch recombination, both important processes in the production of antibodies, and both of which use the NHEJ machinery for DNA strand reattachment (Lieber *et al.* 2004).

Translesion synthesis

Translesion synthesis (TS) is an error-prone mechanism, which uses specific TS polymerases (*e.g.* pol η, pol κ, pol ι and Rev1) to replicate the DNA strand past lesions which block the progress of the replicating high-fidelity polymerases δ and ε. This activity is called lesion bypass. TS polymerases insert bases opposite to the damaged nucleotides with low fidelity, resulting in frequent mis-insertions, which promote mutagenesis (McCulloch, Kunkel 2008). Although it introduces replication errors in DNA, the TS pathway can circumvent more severe conditions, such as double-strand breaks, which

occur when replication forks collapse. In addition to lesion bypass, TS polymerases are also active in some DNA repair pathways, such as HR, NER and BER (Kawamoto *et al.* 2005, Ogi, Lehmann 2006, Prasad *et al.* 2003), and defects in them have been connected to cancer susceptibility in mice and man (Dumstorf *et al.* 2006, Lin *et al.* 2006, Broughton et al. 2002).

DNA damage response pathway

In ideal cases, when DNA damage is detected, the damage is repaired fast and with high fidelity to ensure DNA integrity and continuation of the cell cycle. However, in some cases the process is slow or not possible, and the cell cycle has to be arrested until DNA repair is complete. This activity is mediated by specific signalling cascades, which activate the DNA damage response (DDR) pathway. DDR co-operates with DNA repair and contributes to enhanced repair and the activation of cell cycle checkpoints. Alternatively, if the damage persists, DDR directs the affected cell to apoptosis or senescence (Rouse, Jackson 2002).

DDR sensors, which detect damage to DNA, are probably the proteins of DNA repair pathways which recognise and bind to their specific target lesions. If the problem persists, DDR is activated. The central proteins mediating the DDR signals are the phosphatidyl-inositol 3-kinase (PI3K) -like protein kinases Ataxia-Telangiectasia mutated (ATM) and ATM and Rad3-related (ATR). ATM is activated mainly in response to DSBs, whereas ATR has a more diverse variety of activators (Abraham 2001). Activation of these kinases leads to the phosphorylation of their downstream targets, which include the signal transducers checkpoint kinase 2 (CHK2) and CHK1, and the common DNA damage response signalling protein p53. These phosphorylation cascades lead to *e.g.* H2AX histone phosphorylation and the accumulation of repair factors such as the MRE11-RAD50-NBS1 (MRN) complex at the site of the lesion, cell cycle checkpoint activation, increased transcription or posttranslational modification of DNA repair factors, and eventually, if the problem persists, cell death (Rouse, Jackson 2002). Germline alterations in the DDR pathway genes lead to cancer syndromes, such as ataxia telangiectasia (the mutated gene is *ATM*) (Savitsky *et al.* 1995), Nijmegen breakage syndrome (*NBS1*) (Matsuura *et al.* 1998), and Li-Fraumeni syndrome (*p53*) (Malkin *et al.* 1990, Srivastava *et al.* 1990). Moreover, p53 is sporadically inactivated in about 50% of cancers, emphasizing the extreme importance of the DDR pathway (Soussi *et al*. 2006).

MISMATCH REPAIR

In the course of DNA replication, it is estimated that despite efficient proofreading, the replicating polymerase makes an insertion mistake every 10^6-10^7 bases it incorporates into nascent DNA (Kolodner, Marsischky 1999). Mismatch repair (MMR) is the DNA repair machinery responsible for correcting these errors. The most common mispair is G•T, which causes only a slight DNA strand distortion (Hunter *et al.* 1987), and therefore is the most likely mispair to be ignored by the polymerase's proofreading activity. Another common type of error arises during the replication of repetitive sequences such as common adenine mononucleotide repeats or CA-dinucleotide repeats, so-called microsatellites. During the replication of these sequences, the two DNA strands occasionally detach and renature, giving rise to extrahelical unpaired nucleotides (insertion-deletion loops, IDLs) (Kunkel 1993). MMR screens along the postreplicative DNA and corrects the mismatches and IDLs, thereby reducing the spontaneous mutation rate by a further two to three orders of magnitude (Modrich, Lahue 1996).

In addition to their best characterised function in monitoring postreplicative DNA, MMR proteins are also involved in many other cellular processes, which are briefly described here. For example, the MMR system recognises a variety of DNA lesions caused by *e.g.* alkylating agents, 6-thioguanine, and cisplatin, and mediates cell cycle checkpoint activation and apoptosis (Karran 2001). MMR also plays a role in somatic hypermutation, which occurs in B lymphocytes after antigen stimulation. There, MutSα is believed to recognise the G•U mispairs caused by activation-induced cytidine deaminase (AID), mediate the excision of the U containing strand, and recruit error-prone translesion polymerases to fill the single-stranded gap (Peled *et al.* 2008). The MMR activity in somatic hypermutation leads to mutations primarily in A•T base pairs, whereas base excision repair glycosylases and replication of G•U mispairs leads to mutations in G•C base pairs (Rada *et al.* 1998). Another function of MMR proteins is to suppress recombination of similar but not identical, homeologous sequences (Surtees, Argueso & Alani 2004). On the other hand, large triplet repeat expansions which are associated with many neurodegenerative diseases, such as myotonic dystrophy and Huntington's disease, seem to be dependent on active MMR (Manley *et al.* 1999, Savouret *et al.* 2004).

Despite the variety of activities played by MMR, the main function and focus of this work is the repair of DNA mispairs arising during DNA replication. The fundamental difference between MMR and the DNA damage repair pathways dealing with chemically

altered DNA is that in MMR, no malformed DNA bases are involved, thus creating the dilemma of which strand to degrade and which one to use as a template. Therefore, the repair has to be directed to the newly synthesised strand, which by default contains the incorrect nucleotide.

MMR in *Escherichia coli*

MMR was first described in prokaryotes and the reaction was reconstituted *in vitro* already in 1989 (Lahue, Au & Modrich 1989). Mismatch recognition in *E. coli* is performed by a homodimer of two MutS-proteins. Mismatch-bound MutS complex is then bound by another protein, the homodimeric MutL. The MutS-MutL-DNA complex activates MutH which functions as a latent endonuclease and the strand discrimination sensor. The strand discrimination is based on the transient absence of methylation at the GATC-sites in the nascent strand, where deoxyadenine (DAM) methylase adds methyl groups about 2 minutes after DNA synthesis. MutH incises the DNA in the vicinity of the mismatch by the closest unmethylated GATC-site (Grilley, Griffith & Modrich 1993). DNA is unwound by DNA helicase II (MutU), allowing exonucleases, such as ExoI, RecJ, ExoVII or ExoX to excise the incorrect strand past the mismatch. The resulting gap is filled by DNA polymerase III and sealed by DNA ligase (Burdett *et al.* 2001, Modrich, Lahue 1996). Additional proteins required for the reaction include single-strand binding protein (SSB), which coats the ssDNA gap resulting from exonuclease activity; β-clamp protein, which possibly recruits MutS to mismatches and is required for the processivity of DNA polymerase III; and γ-δ Complex, which loads the β-clamp onto DNA (Kunkel, Erie 2005). The outline of the *E. coli* MMR is depicted in Figure 1.

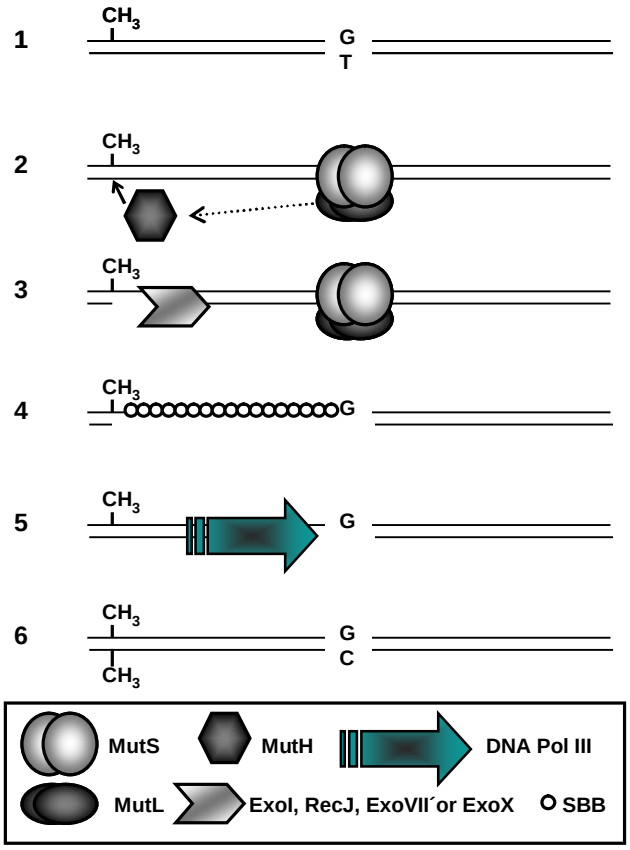

Figure 1. Mismatch repair in *E.coli*. (1) A G-T mispair has escaped the replicative polymerase's proofreading activity. (2) A dimer of MutS binds to the mismatch and attracts dimeric MutL to the site, and the endonuclease MutH is activated. MutH incises the newly synthesised strand in the vicinity of the closest unmethylated GATC site. (3) Exonucleases degrade the nascent strand until the mismatched DNA has been removed. (4). The resulting single-stranded gap is bound by single-stranded protein (SSB). (5) DNA polymerase III fills the gap using the intact strand as a template. (6) The mismatch has been corrected, sealed by ligase, and the new strand methylated by DAM methylase.

MMR in eukaryotes

In eukaryotes, the MMR reaction and involved proteins are highly similar to their prokaryotic counterparts, albeit with some differences. A great deal of work has been done in yeast, contributing significantly to the knowledge of eukaryotic MMR we have today (Reviewed in Fishel, Kolodner 1995). In this work, the main focus is on the human system, but the yeast (*Saccharomyces cerevisiae*) MutS and MutL homologues are shortly introduced.

In humans, there are altogether 5 MutS homologues, of which MutS Homologues MSH2, MSH3 and MSH6 have been associated with MMR (Drummond *et al.* 1995, Palombo *et al.* 1996), whereas MSH4 and MSH5 function in meiosis (Bocker *et al.* 1999).

The corresponding protein homologues are also found in yeast. Furthermore, a sixth MutS homologue, Msh1, not found in mammals, is reported to function in yeast mitochondrial MMR (Reenan, Kolodner 1992). In contrast to the prokaryotic proteins, which act as homodimers, MutS and MutL homologues function in eukaryotes as heterodimers. The mispair recognition is done by a MutS homologue heterodimer, but the exact dimer composition depends on the type of lesion. Base-base mismatches and small (<2 bp) IDLs are recognised by a complex of MSH2 and MSH6 (MutSα), whereas an MSH2-MSH3 (MutSβ) complex recognises only IDLs (Acharya *et al.* 1996, Palombo *et al.* 1996). Thus, MSH6 and MSH3 play partially redundant roles in MMR.

The human MutL Homologues are MLH1, MLH3, PMS1 (Post-Meiotic Segregation 1) and PMS2. In yeast, the closest homologue of human PMS2 is Pms1, whereas yeast Mlh2 corresponds to human PMS1 (Wang, Kleckner & Hunter 1999). In human MMR, the heterodimer of MLH1 and PMS2 (MutLα) is the most important MutL complex, but also the MLH1-MLH3 (MutLγ) complex is able to repair base-base mismatches *in vitro*, and is suggested to act as a backup for MutLα (Cannavo *et al.* 2005, Korhonen *et al.* 2007). In yeast, the main MutL homolog is Mlh1-Pms1, and the Mlh1-Mlh3 complex has been reported to participate in the repair of >3 bp insertion-deletion loops (Flores-Rozas, Kolodner 1998). The roles of human MLH1-PMS1 -complex (MutLβ) (Raschle et al. 1999) and yeast Mlh2 (Wang, Kleckner & Hunter 1999) are uncertain. The eukaryotic MutS and MutL are presented in Table 2.

Table 2. E.coli, yeast, and human MutS and MutL homologues.

E. coli	Yeast	Human	Function in MMR
MutS	Msh1	-	Mitochondrial MMR
	Msh2	MSH2	Mismatch and IDL recognition
	Msh3	MSH3	IDL recognition
	Msh4	MSH4	(Meiotic recombination)*
	Msh5	MSH5	(Meiotic recombination)*
	Msh6	MSH6	Mismatch and small IDL recognition
MutL	Mlh1	MLH1	MMR assembly
	Mlh3	MLH3	MMR assembly and endonuclease (backup?)
	Pms1	PMS2	MMR assembly and endonuclease
	Mlh2	PMS1	?

*These proteins do not have a function in DNA repair

Although several eukaryotic MutS and MutL homologues have been identified, homologues for the endonuclease MutH have not been found. The excision of the

incorrect strand in eukaryotes has been suggested to begin from the free DNA end associated with the progression of the replication fork, either at the 3' or 5' end of the Okazaki fragment in the lagging strand, or the 3' end of the leading strand (Modrich, Lahue 1996). Given the lack of GATC methylation in eukaryotic DNA, the strand discontinuity could also account for the strand discrimination signal. Alternatively, strand discrimination might be directed through the interaction of MMR proteins with the replisome-associated proliferating cell nuclear antigen (PCNA) (Bowers *et al.* 2001, Umar *et al*. 1996), which is the processivity factor for replicative DNA polymerases. Recently, however, the significant finding that MutLα possesses a cryptic endonuclease activity was reported (Kadyrov *et al.* 2006). This activated endonuclease introduces several incisions primarily on the 5' side of the mismatch in the MMR reaction. The activity is disturbed by inactivating mutagenesis in the identified endonuclease sequence motif in the C-terminus of the PMS2 subunit. This endonuclease (DQHA(X)(2)E(X)(4)E) sequence motif is also present in the prokaryotic endonuclease MutH. Human MLH3, but not PMS1, contains the motif, supporting the interpretation that MutLγ, but not MutLβ, plays a role in MMR.

Eukaryotic MMR has also been reconstituted *in vitro* (Constantin *et al.* 2005, Zhang *et al.* 2005). The MMR substrate mimicking the replication error was a double-stranded DNA plasmid with a mismatch and a single-strand nick either 3' or 5' of the mismatch. Such breaks direct the MMR reaction to the correct DNA strand in human cell extracts (Holmes, Clark & Modrich 1990, Thomas, Roberts & Kunkel 1991b). The factors required for the 5' reaction (nick situated 5' of the mismatch) were MutSα; PCNA; replication factor C (RFC), which loads PCNA on the DNA; Exonuclease I (EXOI), which excises the newly synthesised DNA strand; single-strand binding protein RPA, which binds to the ssDNA gap resulting from the ExoI activity; and polymerase δ, which fills the gap using the intact strand as a template. Surprisingly, only the 3' reaction required MutLα (Constantin *et al.* 2005). Another study reported that also the protein HMGB1 was needed for the reaction (Zhang *et al.* 2005).

The lack of MutLα requirement in 5'-directed MMR reconstitution is as yet not understood, as lack of MLH1 is the classical cause for MMR deficiency both *in vitro* and *in vivo* (Li, Modrich 1995, Lindblom *et al.* 1993), and MutLα is believed to be an indispensible molecular matchmaker in the MMR reaction (Jiricny, Nyström-Lahti 2000). The recent finding that MutLα is an endonuclease which provides the 5' break for excision initiation for 5'-3' exonuclease ExoI (Kadyrov *et al.* 2006) partially explains the

dispensability of MutLα in the MMR reconstitution with the 5'-nicked heteroduplex. Why the MMR of 5'-nicked substrates does not occur in MLH1-deficient cell extracts (*e.g.* Nyström-Lahti *et al.* 2002, Raevaara *et al.* 2005) is yet to be clarified. The proteins required for MMR reconstitution in prokaryotes and eukaryotes are listed in Table 3. The overview of the current model of eukaryotic MMR is depicted in Figure 2.

Table 3. *E.coli* and Human MMR proteins.

E. coli	Human	Function
MutS	MutSα, MutSβ	Mismatch recognition
MutL	MutLα, MutLγ	Repairosome assembly Endonuclease in eukaryotes
MutH	-	Endonuclease
DNA helicase II (MutU)	-	DNA helicase
ExoI, RecJ, ExoVII	ExoI (and others?)	Exonuclease
SSB	RPA (and HMGB1)	Single-strand gap protection
DNA pol III	DNA pol δ	Polymerase
β-clamp	PCNA	Polymerase processivity factor
γ-δ Complex	RFC	Processivity factor loader

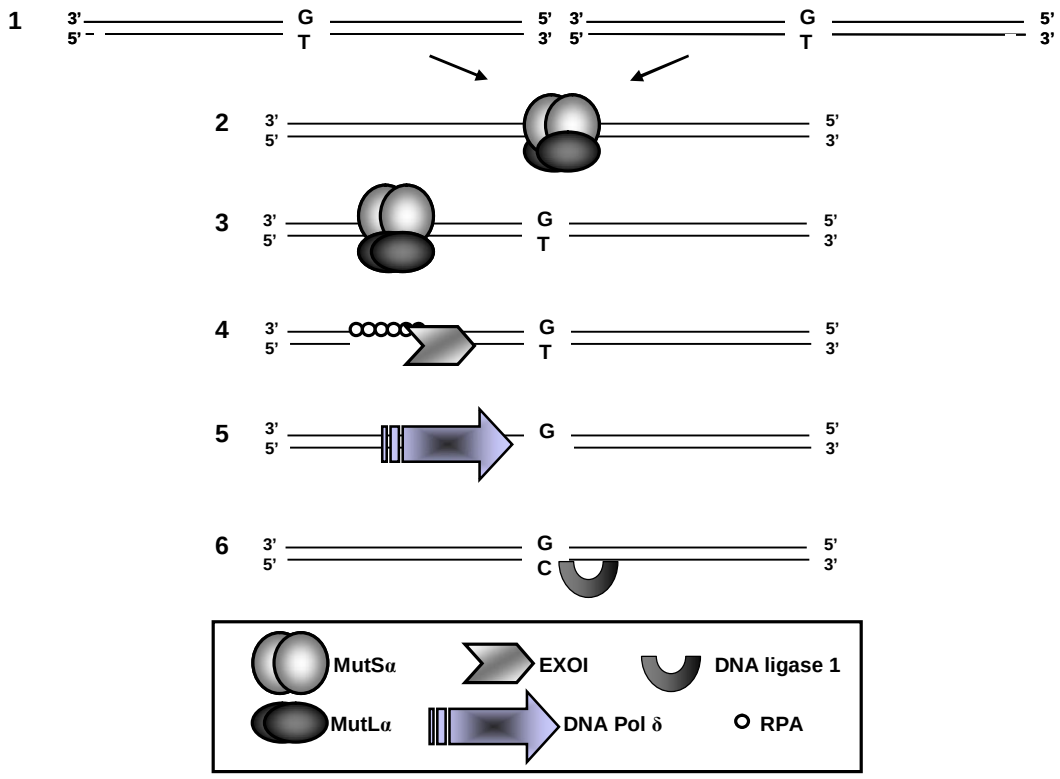

Figure 2. A model for eukaryotic MMR. (1) Strand discontinuity, either 3' or 5' from the mismatch, serves as the strand discrimination signal in the MMR reaction. How the strand discrimination signal is communicated to the site of nicking is unclear. (2) MutSα binds to the mismatch and recruits MutLα to the site. (3) MutSα (possibly bound to MutLα) leaves the site of mismatch in search of strand discontinuity. MutLα makes several incisions in the repairable strand in the vicinity of the mismatch. (4) RPA covers the resulting ssDNA. (5) DNA polymerase δ fills the gap and (6) DNA ligase I ligates the remaining nick. According to (Jiricny 2006, Kadyrov et al. 2006).

Role of MutSα in MMR

When screening postreplicative DNA, MutSα is suggested to be physically attached to the replication machinery by an MSH6-mediated contact to polymerase processivity factor PCNA (Kleczkowska *et al.* 2001). Upon encountering a misintegrated base or an IDL, MutSα binds to the mismatch. It then recruits MutLα to the site, and releases the mismatch by sliding from the site along the DNA in order to allow the subsequent repair process to take place (Blackwell *et al.* 1998, Gradia *et al.* 1999). MutSα changes conformational states from general DNA sliding to mismatch binding and downstream signalling mode by alternating the binding of adenine nucleotides in its two subunits, both of which possess an ATP-binding and hydrolysis domain in their carboxy terminus (Warren *et al.* 2007). As these ATP / ADP switches control the DNA binding activities of the heterodimer, the nucleotide binding and hydrolysis activities of MutSα are vital, and mutations in the ATPase domains of MSH2 and MSH6 have been shown to inactivate functional MMR (Dufner *et al.* 2000, Iaccarino *et al.* 1998). To date, the exact mode of MutSα translocation along DNA remains uncertain. The two favoured models are known as the "sliding clamp" model (Gradia, Acharya & Fishel 1997a) and the "active translocation" model (Blackwell *et al.* 1998). These models differ in terms of the energy requirement for the movement along DNA – the former suggests that several MutSα clamps diffuse stochastically in both directions from the mismatch until they find the strand discontinuity signal, and the latter proposes that the translocation is ATP hydrolysis-driven. However, both models agree that MutSα binds mismatches in an ADP-bound state, and that switching of ADP to ATP mediates a conformational change in the molecule, allowing movement along DNA. Whether the whole MutSα-MutLα ternary complex, which forms upon the mismatch, or MutSα alone actually slides along the DNA remains unknown.

Structure of MutSα

While crystal structures of prokaryotic MutS dimers have been available since 2000 (Lamers *et al.* 2000b, Obmolova *et al.* 2000), the human MutSα structure was solved only recently (Warren *et al.* 2007). As already demonstrated by the prokaryotic structures, the two MutS subunits are organised asymmetrically. The human structure confirmed the previous observations made by mutagenesis experiments that MSH6 is the mismatch binding monomer of MutSα (Dufner *et al.* 2000). According to both the prokaryotic MutS and human MutSα crystal structures, both subunits of the complex are divided into five

functional subunits: a DNA-binding domain (domain 1), a connector domain (domain 2), a lever domain (domain 3), a clamp domain (domain 4) and an ATPase domain (domain 5). Furthermore, the extreme C-terminus in both monomers contains a helix-turn-helix-motif, which stabilizes the ATPase domains of the MutSα subunits (Lamers *et al.* 2000a, Obmolova *et al.* 2000, Warren *et al.* 2007). The crystal structure of MutSα is shown in Figure 3.

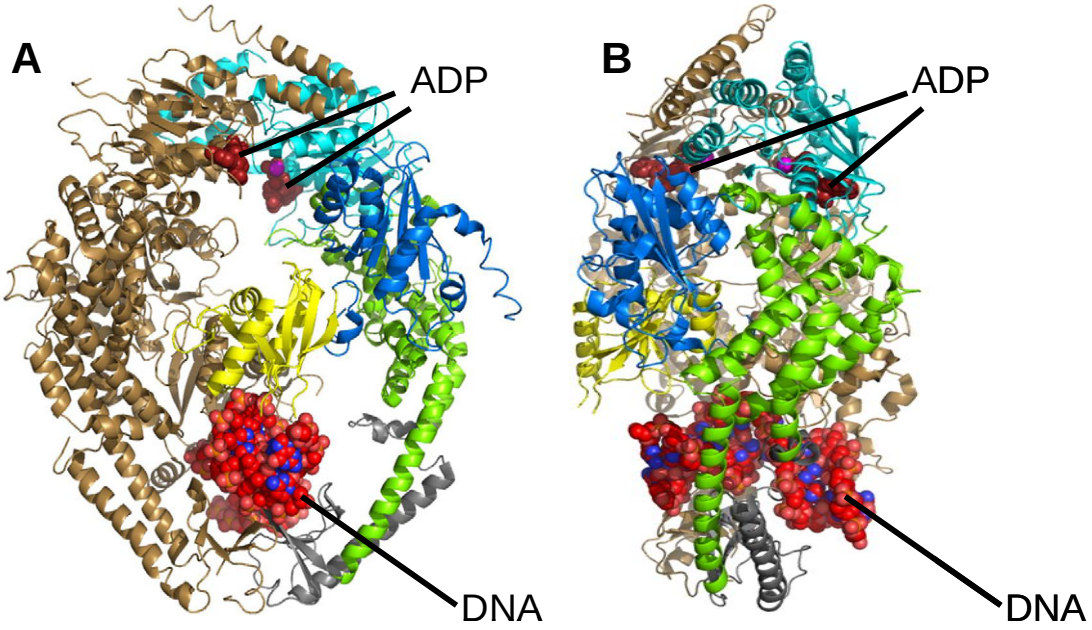

Figure 3. The crystal structure of human MutSα on mismatched DNA. A. The MSH6 subunit of the MSH2-MSH6 heterodimer is displayed on the left, and coloured with light brown. The MSH2 subunit is on the right and its separate functional domains are coloured differentially. Yellow: domain I, DNA binding domain; dark blue: domain II, connector domain; green: domain III, lever domain; grey: domain IV, clamp domain; light blue: domain V, ATPase domain. B. The same structure rotated 90°. The crystal structure represents MutSα bound to a G•T mismatch, and both monomers carry ADP. According to Warren et al. 2007.

The DNA-binding domain of MSH2 makes an unspecific DNA contact in the vicinity of the mismatch, while MSH6 is responsible for the actual binding to the mismatch. The connector domain connects the DNA-binding subunit to the rest of the MutSα heterodimer, and is responsible for the intramolecular interactions and allosteric signalling between different protein domains. The lever domain is a large domain which connects the ATP-binding subunit to the clamp domain, which makes unspecific DNA contacts. It is believed to mediate signals between the ATP- and DNA binding parts of the protein and to communicate the structural transformation messages. The ATP-binding /

hydrolysis subunit modulates the conformation of the protein dimer by binding either no nucleotide, ADP or ATP. As the ATP-binding sites can be occupied by different ligands, the two sites can exist in several different combinations. Because of this, it has been difficult to exclusively determine the nucleotide-binding states of MSH2 and MSH6 in different stages of the MMR reaction (Gradia, Acharya & Fishel 2000, Warren *et al.* 2007).

Human MutSα has been crystallised bound to four different DNA lesions (Warren *et al.* 2007). The G•T mismatch and the unpaired T nucleotide represent replication errors, the classical MMR substrates. As mentioned above, MutSα is known to be involved in other cellular pathways, for example somatic hypermutation and the response to alkylating damage. A yeast study proposed that the conformation of MutSα, bound to its substrate, is determined by the pathway it is employed in (Drotschmann *et al.* 2004). To address this issue in humans, human MutSα complex was also crystallised bound to two other structures: a G•U mispair, reflecting somatic hypermutation, and an O-6-Methyl-G•T, a lesion resulting from alkylating damage. All four substrates were bound similarly, indicating that although MutSα plays a role in several cellular processes, the different signalling does not represent differences in the substrate binding. Instead, varying downstream factors are more likely to mediate the variety of responses (Warren *et al.* 2007).

Defective MMR

As in other DNA repair pathways, also defective MMR leads to genetic instability. The main role of MMR is to correct postreplicative errors in DNA, and MMR deficiency gives rise to point mutations and, in particular, variety in the length of short repetitive sequences, microsatellites. This variety results from unrepaired IDLs, and is called microsatellite-instability (MSI). Due to the accumulation of MSI and other replication errors in the genome, MMR deficiency eventually leads to cancer. MMR defects are frequently found in sporadic tumours, and inherited MMR deficiency leads to hereditary nonpolyposis colorectal cancer (HNPCC), also called Lynch syndrome (Lynch, de la Chapelle 1999a).

HEREDITARY NONPOLYPOSIS COLORECTAL CANCER

HNPCC is a relatively common hereditary cancer syndrome, accounting for approximately 2-3% of all colorectal cancers (CRCs) (Lynch, de la Chapelle 2003, Salovaara *et al.* 2000). The susceptibility to cancer is inherited in an autosomal dominant manner. HNPCC was first reported already in 1913, when Aldred S. Warthin described a family with a hereditary occurrence of gastric cancer (Lynch, Krush 1971). In the 1960s and 1970s, attention was redrawn to the syndrome by Henry T. Lynch. He characterised several hereditary cancer families suffering from familial colorectal and some extracolonic cancers, mainly endometrial tumours (Lynch, Smyrk & Lynch 1998). Thus, the existence of a familial CRC syndrome was characterised already long before its genetic basis was revealed. The syndrome was called hereditary nonpolyposis colorectal cancer to differentiate it from other known hereditary CRC syndromes, such as familial adenomatous polyposis (FAP), juvenile polyposis and Peutz-Jeghers syndrome, all of which are characterised by the occurrence of numerous polyps in the large intestine (for a recent review on colorectal polyposis syndromes, see *e.g.* Jass 2008).

Genetics of HNPCC

The connection between germline defects in MMR genes and HNPCC was established when the first susceptibility genes *MSH2* (Leach *et al.* 1993, Peltomäki *et al.* 1993b) and *MLH1* (Lindblom *et al.* 1993, Papadopoulos *et al.* 1994) were found and mutations in them were shown to segregate with cancer in HNPCC families. Furthermore, MSI, resulting from defective repair of IDLs, was found to be the hallmark of HNPCC tumours (Aaltonen *et al.* 1993). To date, inherited mutations in *MLH1*, *MSH2*, *MSH6* and *PMS2* have been shown to predispose to HNPCC, whereas the role of *MLH3* is elusive and *MSH3* and *PMS1* most likely do not participate in cancer predisposition.

To date, over 1500 different variants have been identified in the four HNPCC genes. By February 2007, 659 unique variants in *MLH1* (44% of all identified MMR gene variants), 595 in *MSH2* (39%), 216 in *MSH6* (14%) and 45 in *PMS2* (3%) had been published (Woods *et al.* 2007). The most typical alterations found in MMR genes are missense mutations and insertions / deletions. Splice site, silent and nonsense variations are somewhat less frequent. Mutations are not clustered in hot spots. Exon 17 in *MLH1* and exon 11 in *MSH2* are the most frequently mutated, if the number of variations is

correlated with the length of the exon (Woods *et al.* 2007). Founder mutations, affecting several families in a typical geographical area, are rare. A common splice-site mutation in *MSH2* intron 5 has been found in several countries, for example in the US and England (Froggatt *et al.* 1999), and the deletion of exons 1 to 6 is a founder mutation in the US (Clendenning *et al.* 2008). In Finland, a splice-site mutation in MLH1 exon 6 and a deletion of MLH1 exon 16 account for the majority of HNPCC mutations (Nyström-Lahti *et al.* 1995). The MSH2 missense mutation A636P is present in about one-third of HNPCC cases in Ashkenazi Jews (Guillem *et al.* 2003, Guillem *et al.* 2004).

Due to the large amount of identified genetic variations in MMR genes, attempts have been made to collect the information into internet databases to distribute it to HNPCC researchers and clinicians. The first HNPCC mutation database was established and maintained by the International Society for Gastrointestinal Hereditary Tumors (InSiGHT) (www.insight-group.org). This database relies on entries of original data from investigators and therefore only includes information provided by the depositor. Recently, a significant contribution to HNPCC mutation compilation has been made by Michael Woods and colleagues, who have assembled all published MMR mutations in one database (Woods *et al.* 2007, www.med.mun.ca/MMRvariants).

According to the two-hit hypothesis, HNPCC is inherited dominantly but, as in the case of many hereditary cancers, the tumorigenesis requires the inactivation of the second allele (Knudson 1996). This second hit can occur for example through promoter hypermethylation, loss of heterozygosity, or gene conversion (Yuen *et al.* 2002, Zhang *et al.* 2006). Thus, the first allele being absent in the germline, the second hit inactivates MMR. This leads to the failed correction of IDLs and therefore to MSI. MSI, then, affects several genes by altering their reading frame. Among the most often reported MSI target genes are *TGFßRII*, *BAX*, and *IGFIIR*, which all contain mononucleotide repeats in the coding sequence (Markowitz *et al.* 1995, Rampino *et al.* 1997, Souza *et al.* 1996), and act as suppressors of cellular growth (*TGFßRII*, and *IGFIIR*) or as proapoptotic proteins (*BAX*). Also the MMR genes *MSH3* and *MSH6*, and a number of others, have been described as MSI target genes (Duval, Hamelin 2002, Malkhosyan *et al.* 1996).

Clinical characteristics of HNPCC patients

HNPCC-related CRCs have some typical characteristics, although none of those allow reliable discrimination from sporadic CRC. The average age of HNPCC onset is about 45 years, in contrast to sporadic cancers, which appear some 20 years later (Lynch, *de la* Chapelle 1999b, Peltomäki, Gao & Mecklin 2001). In HNPCC, tumours are situated mainly in the proximal colon, and multiple synchronous and metachronous tumours are common. HNPCC patients have better prognosis than sporadic CRC patients, and the tumours have typical histological features, such as tumour-infiltrating lymphocytes and mucinous differentiation (Umar *et al.* 2004). HNPCC is also characterised by the frequent occurrence of various extracolonic tumours, mainly in the endometrium, small intestine, hepatobiliary tract, stomach and skin (Muir-Torre syndrome, see below). The penetrance of the cancer phenotype in *MSH2* or *MLH1* mutation carriers has been estimated to be close to 100%, whereas *MSH6* mutation carriers have slightly reduced and *PMS2* mutation carriers significantly lower penetrance (Peltomäki 2005). The risk of CRC in MMR mutation carriers is estimated to be around 80% by the age of 70 years, with females having a somewhat lower risk than males. Endometrial cancer is even more common than CRC in females, showing about 50 – 60% penetrance (Aarnio *et al.* 1999a, Vasen *et al.* 1996). The life-time risk of other extracolonic cancers is estimated to be between 2 and 10% (Watson, Lynch 2001).

Several clinical criteria have been introduced to unify the international practice of HNPCC diagnostics and to identity the HNPCC families from the frequent sporadic CRCs. The first diagnostic criteria, the Amsterdam criteria (AC), were created 1991 (Vasen *et al.* 1991). AC are based on the young age of onset and familial occurrence of CRC. These criteria were later modified to include the extracolonic tumours of the HNPCC tumour spectrum in the diagnostic guidelines (ACII), (Vasen *et al.* 1999). The Amsterdam criteria are specific and only rarely identify false positive cases, but they are not sensitive and many HNPCC families will be missed if these are used as the single criterion. Therefore, the Bethesda guidelines were established to identify the HNPCC families who, due to *e.g.* small family size or insufficient information, were not found with AC (Rodriguez-Bigas *et al.* 1997). These criteria made use of the MSI phenotype associated with HNPCC tumours. Also the Bethesda criteria have been later modified (Umar *et al.* 2004). Both revised criteria are detailed in Table 4.

Table 4. The diagnostic criteria for HNPCC.

Amsterdam Criteria II (All of following conditions fulfilled)	**Revised Bethesda guidelines** (Any of the following conditions fulfilled)
•At least three relatives with HNPCC-associated cancer (in colorectum, endometrium, small bowel, ureter, renal pelvis)	•CRC diagnosed in a patient before age 50
•One should be a first degree relative of the other two	•Presence of synchronous, metachronous CRC or other HNPCC associated tumours, regardless of age
•At least two affected generations	•CRC with the MSI-high histology* diagnosed in a patient before age 60
•At least one member diagnosed before age 50	•CRC diagnosed in one or more first-degree relatives with an HNPCC-associated tumour, with one of the tumours diagnosed before age 50
•FAP should be excluded	
•Tumours should be verified by pathological examination	•CRC diagnosed in two or more first- or second degree relatives, with HNPCC-associated tumours, regardless of age

*For definition of MSI status, see text below.

Microsatellite-instability and immunohistochemistry in HNPCC diagnostics

Management of the cancer families and the applied tumour therapy differs between HNPCC and sporadic CRCs. Therefore, phenotypic features helping the diagnostics of HNPCC are very important to indicate the requirement for mutation analysis. Phenotypic characteristics of HNPCC tumours include MSI and loss of expression of an MMR protein, reflecting the underlying MMR defect.

MSI analysis is informative for establishing whether the tumour is MMR-deficient or not, but is not specific to HNPCC, due to the frequent somatic inactivation of the MLH1 promoter in sporadic CRCs (Thibodeau *et al.* 1998). However, MSI analysis is often used as a rough screening method for MMR defect. In MSI testing, a general panel of 5 microsatellite markers are used. If two or more of those are unstable, the tumour is classified MSI-high (MSI-H) (Boland *et al.* 1998).

Immunohistochemical (IHC) assessment of the expression of MMR proteins in tumours is another widely used method for detecting MMR deficiency. Lack of expression gives a good indication of the MMR factor behind the MSI phenotype. This holds especially true for MSH2 (Mangold *et al.* 2005), whereas lack of MLH1 expression is often due to promoter hypermethylation. On the other hand, the expression of a protein does not always indicate that it is functional (Mangold *et al.* 2005, Raevaara *et al.* 2005). It is also to be noted that MMR proteins function as heterodimers, and some monomers are not stable without their partners (Chang *et al.* 2000). Thus, if MSH2 is absent, also MSH6 staining is negative, due to the instability of MSH6 without MSH2. The same is true for PMS2: without MLH1, PMS2 degrades. Solid knowledge of the expression

profiles of all four genes gives a good idea where to look for the germline mutation (Hampel *et al.* 2005).

Muir-Torre syndrome

Muir-Torre syndrome (MTS) is a rare subtype of HNPCC. A recent report showed that about 9% of individuals with HNPCC also exhibit MTS (South *et al.* 2008). In addition to tumours of the HNPCC spectrum, the patients also display skin tumours, predominantly sebaceous gland tumours or keratoacanthomas (Cohen *et al.* 1995, Schwartz *et al.* 1989). Both skin and visceral tumours in MTS display high MSI (Kruse, Ruzicka 2004a), reflecting the underlying MMR defect. Potentially, a skin tumour with a diagnosed germline MMR defect could be of strong clinical importance, allowing the diagnosis of MTS already before the visceral cancer manifestates. This would greatly contribute to early surveillance and cancer prevention in the mutation carrier.

By 2006, a total of 41 MMR gene mutations linked to MTS had been reported. 38 of these mutations were situated in *MSH2* and only three in *MLH1*, suggesting a strong bias of MTS syndrome associating with only one MMR gene, *MSH2* (Bapat *et al.* 1999, Kruse, Ruzicka 2004b, Mangold *et al.* 2004, Ponti *et al.* 2005). This is in contrast to HNPCC, where *MLH1* is the most frequently mutated predisposing gene (Peltomäki *et al.* 2005). In addition, only three of the MTS-linked mutations were of the missense type (7%), in contrast to the fact that 22% of all reported MMR gene variations and 17% of *MSH2* variations are missense mutations (www.med.mun.ca/MMRvariants). Some recent systematic studies suggest, however, that *MLH1* mutations could play a more frequent role in MTS than previously thought (Ponti *et al.* 2006, South *et al.* 2008), and that also *MSH6* mutations are involved (Mangold *et al.* 2007, Murphy *et al.* 2008).

NONTRUNCATING MUTATIONS IN HNPCC

Frequency of nontruncating mutations in MMR genes

A major problem in the diagnosis and management of HNPCC is the frequent occurrence of nontruncating mutations. As in all genes, a point mutation, which changes one amino acid in the polypeptide, may either have a harmful (or beneficial) effect on protein function or not affect the function at all. Thus, when encountering a nontruncating

mutation in an MMR gene in a putative HNPCC family, the interpretation can be very difficult, especially if the co-segregation of the mutation and cancer phenotype cannot be confirmed. In MMR genes, nontruncating variations are the most common type of reported mutations (Woods *et al.* 2007), with missense mutations alone accounting for 24% of *MLH1*, 17% of *MSH2*, and 27% of unique *MSH6* variations. Of the reported *PMS2* mutations, missense variants account for nearly 50% (http://www.med.mun.ca/MMRvariants). The pathogenicity of a sequence variation is classically determined based on the conservation status and biochemical significance of the amino acid change, segregation of the mutation with the cancer phenotype, and MSI and IHC status of the tumours of the mutation carriers (Barnetson *et al.* 2007, Genuardi *et al.* 1999). However, the clinical phenotype of a nontruncating mutation may vary within different families, and segregation data is not always available. Therefore, functional assays have been developed to clarify the activity of nontruncating MMR gene mutations. The following chapters summarize the techniques used in the functional characterisation of MMR gene defects.

Functional analysis of nontruncating MMR gene variants

Functional assays aim to investigate how a nontruncating mutation affects the biological and biochemical behaviour of a protein variant as compared to the wild type (WT). Recently, data concerning the published functional assays on MMR genes has been collected (Ou *et al.* 2007) and listed in a database (www.mmrmissense.net). With functional testing, one can show that the observed genetic variation really reflects the observed phenotype. In the case of MMR proteins, the functional assays can be divided in two major classes: the ones which measure the success of an MMR reaction, and the ones which monitor one specific function of the MMR protein in question. The former ones can be conducted *in vivo* in yeast assays, or *in vitro* using mammalian cell extracts. The latter ones measure activities such as heterodimer subunit interaction, DNA binding, or subcellular localization. The two types of MMR assays complement each other, as the success of an MMR reaction gives information about the functionality of a given variant, but does not elucidate the reasons behind the putative pathogenicity. Recently, also several computational methods have been developed to assess the tolerability of MMR gene amino acid substitutions.

In vivo MMR assays in yeast

Yeast is an optimal model organism to assay MMR functions *in vivo* due to the conservation between human and yeast MMR systems and the facility of yeast-based techniques. Two main approaches are in use. Mutations corresponding to patient-derived MMR gene mutations can be constructed in homologous positions in the yeast genome and the mutation rate caused by the mutant protein can be determined (Shcherbakova, Kunkel 1999, Drotschmann *et al.* 1999, Gammie *et al*. (2007). Alternatively, the fact, that human MMR proteins are able to bind yeast MMR factors and block the intrinsic MMR activity in yeast, has been exploited. WT and functionally intact mutations introduced to a WT yeast cause a mutator phenotype, whereas non-functional human MMR proteins fail to do so (Clark *et al.* 1999, Shimodaira *et al.* 1998, Clark *et al.* 1999, Takahashi *et al.* 2007). The two methods can be used to complement each other (Drotschmann, Clark & Kunkel 1999).

Although important tools in assessment of the activity of putative HNPCC-related MMR mutations, yeast MMR assays harbour the problem that they always rely on the homology between human and yeast proteins, allowing only the conserved amino acids to be tested. However, Ellison *et al*. (Ellison, Lofing & Bitter 2001) developed an assay where they used yeast-human MLH1 hybrid proteins in a yeast context.

In vitro MMR assays in cell lysates

The homology limitation and cross-species difference problems related to yeast assays can be overcome by using homologous human *in vitro* assays. These human systems are not limited to conserved amino acids, and all reactions make use of human proteins. These assays are based on mismatched DNA substrates, which mimick the repairable cellular DNA. These substrates are incubated with human cell extracts. The detection of successful MMR is based on the *in vitro* correction of the mismatch, resulting either in a restriction site (Lahue, Au & Modrich 1989, Nyström-Lahti *et al.* 2002) or in a change of reading frame in the *lacZ* reporter gene (Thomas, Roberts & Kunkel 1991a).

Assays measuring a specific function of MMR proteins

As both MutSα and MutLα function as dimers, one way to study their functionality is to assess whether the dimer subunits are able to interact with each other. This has been addressed by several methods: GST pull-down assays, yeast two-hybrid assays, and co-

immunoprecipitation assays. (Guerrette *et al.* 1998, Guerrette, Acharya & Fishel 1999, Nyström-Lahti *et al.* 2002, Kondo *et al.* 2003, Kariola *et al.* 2004, Raevaara *et al.* 2005). DNA-binding experiments (bandshift assays or electrophoretic mobility shift assays) are based on the *in vitro* binding of purified MutSα proteins to labelled heteroduplex oligonucleotides. DNA-binding experiments have revealed defective binding to heteroduplex oligomers with mutated human and yeast MutSα (Clark *et al.* 1999, Drotschmann, Clark & Kunkel 1999, Heinen *et al.* 2002). Despite of being functional in an *in vitro* MMR assay, it is possible that *in vivo* the mutated protein is never localised to the nucleus or that its expression levels are reduced. Indeed, both expression and localization have been studied and found to be defective in context of HNPCC-derived MMR mutations (Raevaara *et al.* 2005, Gammie *et al.* 2007).

In silico *prediction algorithms*

One way to differentiate between nonpathogenic and pathogenic missense variants is to use computational algorithms. They are based on comparative sequence or protein structure analysis. Mostly used are PolyPhen (http://coot.embl.de/PolyPhen) and Sorting Intolerant From Tolerant (SIFT) (Ng, Henikoff 2003) (http://blocks.fhcrc.org/sift/SIFT.html). They are based on searching for similar sequences against a database (e.g. SWISS-PROT/TrEMBL), and aligning the protein sequences from different species to address the conservation of each amino acid. Alternatively, users can enter a pre-aligned set of sequences. Probabilities are calculated for each amino acid substitution and those smaller than a chosen cut-off value are predicted to be deleterious. The validation of computer-based methods requires the simultaneous use of in silico and functional analyses of missense variations.

AIMS OF THE PRESENT STUDY

The main aim of this PhD study was to investigate the pathogenicity of nontruncating patient-derived mutations in *MSH2* to facilitate HNPCC diagnostics in families associated with these variations. The specific aims were:

1. To assess whether nontruncating mutations in *MSH2*, found in HNPCC and Muir-Torre families, cause MMR deficiency (I, II)

2. To clarify the biochemical defect underlying the observed MMR deficiencies (IV)

3. To assess the functionality of two frequently occurring *MSH2* variants and to estimate their connection to cancer predisposition based on *in vitro* biochemical data and literature searches (III)

4. To determine whether clinical characteristics of the mutation carriers could be correlated with the results of the biochemical analyses, and with the location of the mutations in the *MSH2* polypeptide (II, IV)

MATERIALS AND METHODS

MSH2 MUTATIONS AND ASSOCIATED FAMILIES (I-IV)

The mutations included in this study consisted of 18 nontruncating *MSH2* mutations, which are discussed throughout this work by referring to the amino acid change in the corresponding MSH2 residue (GenBank accession number AH003235, version U41206.1 to U41220.1). The studied variants were dispersed across different domains of the MSH2 polypeptide, but were somewhat clustered in the amino (N) -terminal connector domain and in the adenosine triphosphatase (ATPase) domain at the C –terminus (Figure 3, p. 29 and Figure 4).

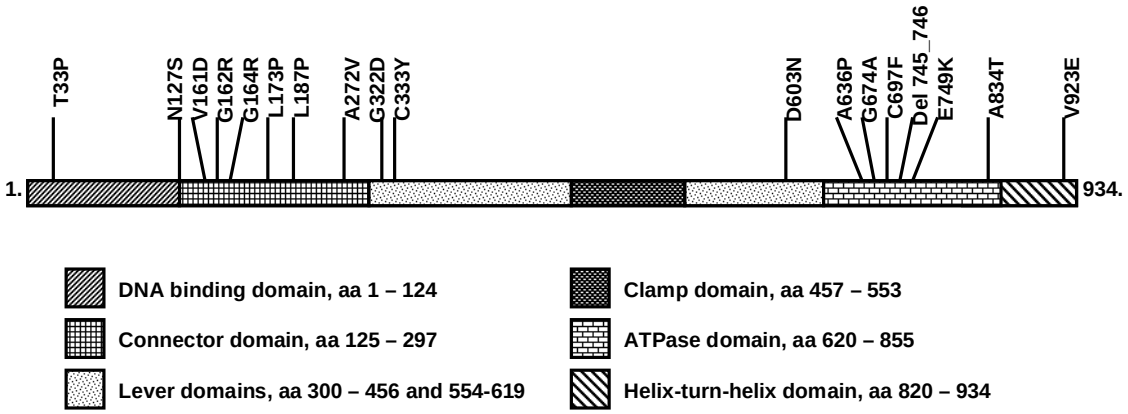

Figure 4. The schematic representation of the studied mutations along the MSH2 polypeptide. According to Lamers *et al.* 2000b, Obmolova *et al.* 2000, Warren *et al.* 2007.

Fifteen of the mutations were found in putative HNPCC families and came to be studied through international collaborations, due to the pathogenicity of an identified nontruncating *MSH2* variant being uncertain (T33P, V161D, G162R, G164R, L173P, L187P, A272V, C333Y, D603N, A636P, C697F, Del 745-746, E749K, A834T and V923E) (II). Some of these mutations have also been reported in databases by other research groups (www.insight-group.org). The genetic and clinical characteristics of the associated families, such as MSI and IHC studies, which were assessed by our collaborators, are shown in Table 5. The L187P and A272V mutations were found in two, and the D603N, A636P, and C697F mutations in three separate families. In addition to

HNPCC, mutations L187P and C697F were also associated with families displaying Muir-Torre syndrome, thus exhibiting concurrent skin and visceral tumours.

Table 5. Genetic and clinical characteristics of the putative HNPCC families associated with 15 nontruncating MSH2 mutations.

MSH2 variant	Nucleotide change in cDNA	Index patient: Age at onset (y) / tumor site	All affected patients[1] / mean age at onset (y)	Amsterdam Criteria I / II	MSI[2] status	Immunohistochemistry[3]		
						MSH2	MSH6	MLH1
T33P	c. 97 A>C	45 / EC	2 / 48	-	High	+	+	+
V161D	c. 482 T>A	52 / CRC	3 / 53	+	High	-	ND	ND
G162R	c. 484 G>A	56 / EC	6 / 52	+	High	-	+/-	+
G164R	c. 490 G>A	39 / CRC	7 / 39	+	ND	-	ND	+
L173P	c. 518 T>C	36 / CRC	9 / 45	+	High[*A]	-	ND	+
L187P	c. 560 T>C	41 / CRC	5 / 42	+	High	-	ND	+
L187P	c. 560 T>C	42 / CRC	11 / 48	+	High[*B]	-	-	+
A272V	c. 815 C>T	41 / CRC	3 / 40	+	Low[*C]	+	ND	+
A272V	c. 815 C>T	41 / CRC	1 / 41	-	High	+	+	+
C333Y	c. 998 G>A	41 / CRC	2 / 41	-	ND	ND	ND	ND
D603N	c. 1808 G>A	50 / EC	2 / 27	-	High	-	-	+
D603N	c. 1808 G>A	46 / EC	5 / 55	+	Stable	-	-	+
D603N	c. 1808 G>A	38 / CRC	1 / 38	-	High	ND	ND	ND
A636P	c. 1906 G>C	42 / CRC, 44 / EC	2 / 44	+	High	-	+/-	+
A636P	c. 1907 G>C	43 / CRC	1 / 43	-	High	-	-	+
A636P	c. 1908 G>C	36 / EC	5 / 49	+	High	-	ND	ND
C697F	c. 2090 G>T	27 /CRC	5 / 45	+	High	-	ND	+
C697F	c. 2090 G>T	33 / CRC	3 / 38	+	High	-	ND	+
C697F	c. 2090 G>T	40 / EC	3 / 49	+	High	-	+	+
Del745-746	c. Del 2235-2240	39 / CRC	4 / 42	+	High	-	-	+
E749K	c. 2245 G>T	29 / CRC	7 / 29	+	High[*D]	+	ND	+/-
A834T	c. 2500 G>A	28 / CRC	3 / 39	+	High	-	-	+
V923E[4]	c. 2768 T>A	70 / CRC	6 / 58	+	High	+/-	-	+

For references, please refer to the Original Article II, from which these data are derived. Each line corresponds to one family. EC, endometrial cancer; CRC, colorectal cancer; ND; no data. [1]Number of affected patients with HNPCC tumours [2]MSI analysis was carried out using the Bethesda panel (Markers BAT-25, BAT-26, D2S123, D5S346, D17S250 or in some cases D18S69). Two or more unstable markers were considered as MSI-high. In families marked with an asterisk (*) MSI was examined with other markers: [*A] D2S123, D2S136, D6S470, D16S663 (unstable) and HBA1 (stable). [*B] Bat26 and Mdf15 (unstable). [*C] TP53-Dint (unstable) D8S254, NM23, D18S35, D5S346, TP53-Penta, D2S123, D1S2883, D3S1611, D7S501 (stable). [*D] D2S123, D16S663 (unstable) D5S346, HBA1, D18S35 (stable). [3] MSI and immunohistochemistry were analysed on the primary tumour from the index patient (for exceptions, see II). [4] The index person carries two mutations, *MSH2* V923E and *MSH6* S1188N.

In addition to the 15 mutations listed above, *MSH2* N127S and G322D were studied because of their frequent occurrence both in published CRC families and healthy individuals and, thus, their contradictory classifications in literature and databases. By functional analysis, we wanted to clarify their pathogenicity status. The published data on the clinical characteristics of these variants and their occurrence in the healthy population is collected in Table 6.

Table 6. The published clinical and population data of the MSH2 N127S (c.380 A>G) and MSH2 G322D (c.965 G>A) variants.

MSH2 variant	Tumor/age[a]	Other cancers in family	MSI[b]	IHC[c] / MSH2	Variant found in healthy controls	Another MMR gene mutation found in a mutation carrier
N127S	HNPCC[d]	+	ND	ND	-	-
N127S	CRC<50, EC<50	+	+	ND	-	MSH2 A328P
N127S	CRC	ND	+	ND	-	-
N127S	CRC 40	+	+	ND	-	MLH1 frameshift c 1877
N127S	CRC<60	-	+	ND	-	MSH2 N108N, MLH1 IVS15-5T>C
N127S	PC 71, BI C 78	-	ND	ND	ND	-
N127S	No cancer	+	ND	ND	ND	-
N127S	CRC31	+	+	-	ND	MSH2 E422X
N127S	CRC34	+	+	-	ND	MSH2 E422X
G322D	ND	ND	ND	ND	+	-
G322D	CRC	+	ND	ND	ND	-
G322D	CRC36, EC45	ND	ND	ND	+	-
G322D	CRC19	ND	ND	ND	+	G322D homozygote
G322D	HNPCC	+	ND	ND	+	-
G322D	CRC40	-	ND	ND	ND	-
G322D	CRC	+	ND	ND	+	-
G322D	CRC	+	-	ND	+	-
G322D	CRC<50	+	+	-	ND	-
G322D	ND	+	ND	ND	ND	MSH2 Q518X
G322D	CRC	-	+	ND	+	-
G322D	CRC	+	+	-	-	MSH2 Q518X
G322D	CRC	-	-	+	-	-
G322D	CRC	ND	ND	ND	+	-
G322D	CRC and EC	ND	+	ND	ND	-
G322D	ND	ND	ND	ND	+	-
G322D	CRC39	+	+	ND	ND	MLH1 T117M
G322D	CRC	+ / -	ND	ND	-	-
G322D	EC49	+	+	+ / -	+	MLH1 D203N[e], MSH2 frameshift
G322D	CRC	+	ND	ND	+	-

Every row represents one published study. For references, see Original Article III. CRC, colorectal cancer; EC, endometrial cancer; PC, pancreatic cancer; BI C, biliary tract cancer; ND, no data. a: The associated cancer type and, when available, age at onset of the index person. b: Microsatellite instability (MSI) found at least in one tumour of a mutation carrier. c: Immunohistochemical (IHC) analysis of MSH2 protein expression in tumour tissue. d: Unspecified cancer belonging to the HNPCC spectrum. e: Somatic mutation.

The mutation G674A (c.2021 G>C) is located in a conserved residue of the MSH2 ATPase domain and was reported to inactivate MMR, but to still be functional in apoptosis signalling in mouse (Lin et al. 2004). G674A was included in the study because we wanted to characterise the functional properties of the corresponding human protein. Another reason was that an HNPCC-associated germline mutation (G674D) has been reported in the same amino acid residue (Raedle et al. 2001).

The sites of all studied mutations are shown in mapped in the crystal structure of MutSα in Figure 5.

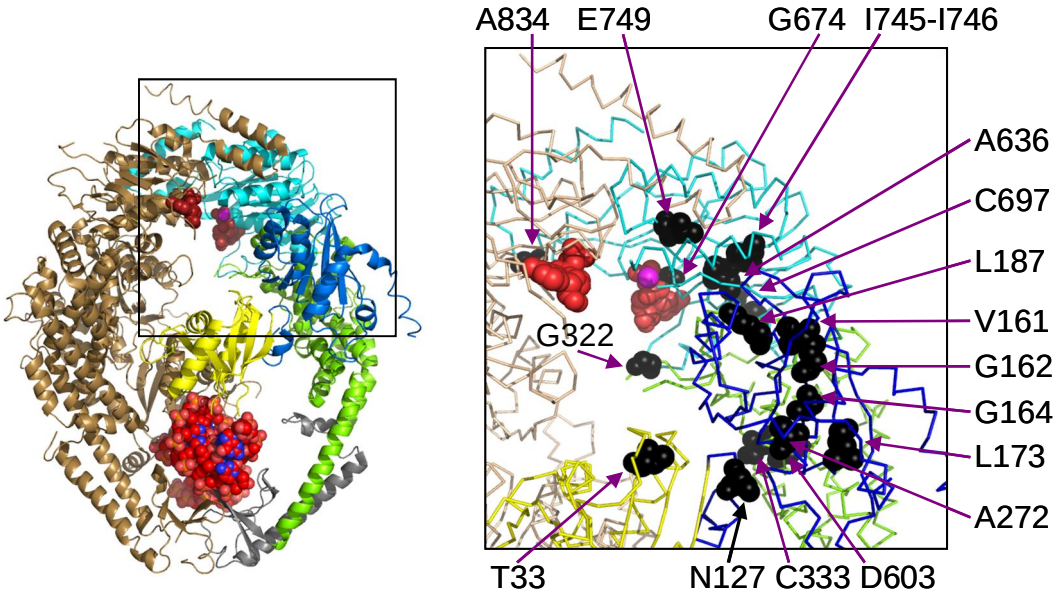

Figure 5. Mapping of analysed MSH2 residues in the crystal structure of MutSα. The MSH6 subunit of the MSH2-MSH6 heterodimer is coloured with light brown. The functional domains of the MSH2 subunit are coloured differentially. Yellow: domain I, DNA binding domain; dark blue: domain II, connector domain; green: domain III, lever domain; grey: domain IV, clamp domain; light blue: domain V, ATPase domain. DNA is shown in red and blue in the lower part of the figure, ADP molecules in red on the upper part. Right: a close-up ribbon structure of the area containing the studied mutations. According to Warren *et al*. 2007.

FUNCTIONAL ASSAYS (I-IV)

Protein expression and coimmunoprecipitation (I-IV)

Site-directed mutagenesis and production of baculovirus expression vectors

All used complementary DNAs (cDNAs) were derived from Professor Josef Jiricny's laboratory in University of Zürich, Switzerland. WT *MSH2* cDNA was cloned into the pFastBac1 plasmid (Invitrogen) between the vector's *Bam*HI and *Xho*I, WT *MSH6* cDNA between its *Bam*HI and *Xho*I, and WT *MSH3* cDNA between its *Xho*I and *Xma*I restriction sites. The *MSH6* cDNA construct included a polyhistidine (His_6) tag at the N-terminus of the MSH6 protein. The expression constructs for the *MSH2* variants were created using site-directed mutagenesis as detailed in the original articles.

Production of expression vectors for human cell expression

For protein expression in human cells, WT and mutated *MSH2* cDNAs were cloned from pFastBac1 into the expression vector pDsRed2-N1 (BD Biosciences) between the *Sac*I

and *Not*I restriction sites, so that the red fluorescent protein was replaced by the insert. The WT *MSH6* cDNA was cloned from pFastBac1 into the expression vector pEGFP-N1 (BD Biosciences) between the *Bam*HI and *Not*I restriction sites, replacing the enhanced green fluorescent protein (EGFP) gene. The resulting constructs expressing MSH2 (WT or mutated) and MSH6 (WT) were named p*MSH2*-N1 and p*MSH6*-N1, respectively.

Baculoviral expression of MutSα variants

The recombinant proteins were produced in *Spodoptera frugiperda* 9 (*Sf*9) insect cells using the Bac-to-Bac baculovirus expression system (Invitrogen). The cDNAs of WT *MSH6*, WT *MSH3* and WT and mutant *MSH2* were transferred to baculovirus vectors *via* a transposon-mediated reaction in DH10Bac *E. coli* cells (Invitrogen). The bacmid DNAs were isolated from bacterial cultures and the baculovirus DNAs, containing the desired cDNA inserts, were used to transfect *Sf9* cells. The secreted baculoviruses were collected after 3 days and amplified in *Sf*9 cells for 5 days. For protein production, *Sf9* cells were co-infected with *MSH2* and *MSH6* baculoviruses, since the functional MutSα-complex requires both proteins and MSH6 is unstable without MSH2 (Chang *et al.* 2000, Marra *et al.* 1998). For control experiments, cells were coinfected with WT *MSH2* and WT *MSH3* baculoviruses for MutSβ dimer production. The total protein extracts (TEs) including the heterodimeric MutSα or MutSβ were extracted in lysis buffer (25 mM Hepes, 2 mM β-mercaptoethanol, 0.5 mM spermidine, 0.15 mM spermine, 0.5 mM phenylmethylsulfonyl fluoride (PMSF), and 2 x Complete protease inhibitor mixture (Roche)). 10% glycerol and 110 mM NaCl were added and the protein extracts were rotated for 30 min at +4°C and then centrifuged for 50 min at 13200 *g*. The soluble protein fractions, containing the recombinant MutSα proteins, were aliquoted and stored at -80°C.

MutSα expression in human cells

Production of recombinant MutSα (heterodimer of MSH2 and MSH6) variants in the LoVo human colon adenocarcinoma cell line (MSH2-/-) (American Type Culture Collection) was performed as follows: a total of 200 000 cells were seeded (in one well of a 6-well plate), and transfected after 24 hours with 2 μg of p*MSH2*-N1 (WT or mutant) and 2 μg p*MSH6*-N1 vectors using 8 μL of Tfx-20 transfection reagent (Promega). After 48 hours from transfection, the cells were collected by trypsinisation. The total protein content was extracted by incubating the cells for 25 min on ice in 50 μl of cold extraction

buffer (50 mM Tris-HCl pH 8.0, 350 mM NaCl, 0.5% Nonidet-P40, 1 x complete protease inhibitor mixture). The suspension was centrifuged at 13200 *g* for 5 min at 4°C, after which the supernatant, containing the desired recombinant proteins, was collected.

Western blot and coimmunoprecipitation analyses

The expression levels and correct sizes of the recombinant proteins were examined by sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) and Western blot analyses. The protein complexes were run in 6% SDS-PAGE gels, blotted to nylon membranes and detected with anti-MSH2 (MSH2 Ab-2, NA27, Calbiochem, 0.4 μg/ml in insect cell and 0.1 μg/ml in human cell expression) and anti-MSH6 (MSH6/GTBP, Clone 44, BD Transduction laboratories, 0.17 μg/ml in insect cell and 0.5 μg/ml in human cell expression) antibodies. The amount of naturally expressed β-tubulin protein (clone 5H1, anti-β-tub; BD Biosciences, 0.5 μg/ml) was used as loading control when the expression levels of MutSα variants produced in LoVo cells were compared**.**

The interactions of MSH2 variants with their counterpart MSH6-WT were studied by combined coimmunoprecipitation and Western blot analysis. For immunoprecipitation, 30 – 100 μg of *Sf9* TEs were adjusted to contain similar amounts of recombinant proteins. The protein extracts were rotated for 3 hours (III, IV) or overnight (II) on a rotating wheel with 1 μg of anti-MSH6 antibody (see above) in RIPA buffer (150 mM NaCl, 1% Nonidet P-40, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris-Cl pH 8.0). A total of 30 μl of protein A/G agarose beads (SC2003, Santa Cruz Biotechnology) were added and the rotation was continued for 1 (III, IV) or 3 (II) hours. The agarose beads, containing the precipitated antibody-protein complexes, were collected by centrifugation and washed three times with RIPA buffer. The precipitated complexes were detected by Western blot (see above) and the amounts of mutated MSH2 proteins in the precipitates were compared to WT MSH2.

Protein purification (III-IV)

Fast protein liquid chromatography

WT MutSα and the variants T33P, A636P, E749K, A834T and V923D were purified with fast protein liquid chromatography (FPLC). The TE was first loaded to the Heparin HiTrap column (Amersham Biosciences) and eluted with a NaCl gradient from 300 to

1000 mM. The fractions containing MutSα were loaded onto a MonoQ anion exchange column (Amersham Biosciences) and eluted with a NaCl gradient from 150 to 550 mM. The MutSα-containing fractions were pooled and dialyzed at +4°C for 2 hours against 25 mM HEPES (pH 7.6), 1 mM EDTA, 110 mM NaCl, 10% sucrose, 3 mM DTT, 0.5 mM PMSF and 1μg/ml leupeptin. The purified proteins were aliquoted, frozen with liquid nitrogen and stored at -80°C. The WT MutSα purified with FPLC is hereafter referred to as WT-FPLC.

Ni-NTA affinity chromatography

The variants N127S, A272V, V161D, G162R, G164R, L173P, L187P, G322D, C333Y, D603N, G674A, C697F, Del 745-746, and WT MutSα (WT-his), were purified with Ni-NTA (nitrilotriacetic acid) agarose matrix (Qiagen), which binds the $(His)_6$-tag located at the N-terminus of the MSH6 subunit of MutSα. This approach has previously been successfully used to purify recombinant MutSα (Gradia, Acharya & Fishel 1997b), and being fast and simple to perform, it was chosen as the method of purification for the majority of the proteins. For purification of the MutSα complexes, 100 μl of Ni-NTA matrix (Qiagen) was used for every 1 ml of TE. The matrix was equilibrated with PBS and mixed with TEs. The mixtures were rotated at +4°C for 2 hours and loaded into 1.5 ml polypropylene columns (Qiagen). The MutSα-bound matrix was washed with wash buffer (25mM HEPES, 300 mM NaCl, 20mM imidazole, 1 μg/ml leupeptin, 1x complete EDTA free, 0.5 mM PMSF) and MutSα was eluted with an increasing imidazole concentration. The MutSα-containing fractions were pooled, dialysed and aliquoted as described above.

The purities of all different MutSα preparations, purified with both methods, were compared by SDS-PAGE and coomassie staining, and the concentrations were assessed with the Bradford assay, using bovine serum albumin (BSA) as a standard.

Mismatch repair assays (I-IV)

Preparation of MMR assay substrates

The DNA substrates used in the *in vitro* MMR experiments were pGEM 13Zf+ (Promega) plasmid-derived double-stranded circular heteroduplex DNA molecules, which contained a G•T mismatch or an extrahelical T insertion within the *Bgl*II restriction site, and a single-strand nick 370 bp 5’ of the mismatch site in the repairable strand (Figure 6). The

plasmids were a kind gift from Professor Josef Jiricny. The substrates were named 5'G•T and 5'IDL1, respectively. Both of these are efficiently recognised by MutSα and have been shown to be functional MMR substrates *in vitro* (Thomas, Roberts & Kunkel 1991a). The mismatch and nick were designed to give rise to a complete double-stranded *Bgl*II restriction site upon a successful mismatch repair reaction. For linearisation of the plasmid, a *Bsa*I digestion site, 1360 bp from the *Bgl*II site, was used.

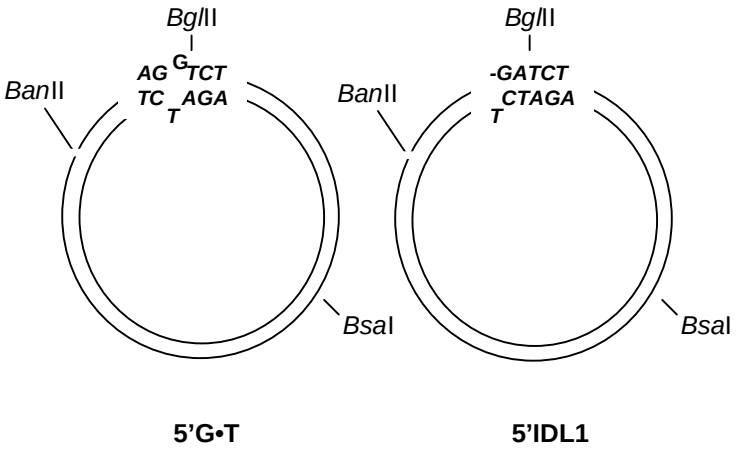

Figure 6. The substrates used in the in vitro MMR assays. Left: 5'G•T substrate. The uncut (bottom) strand contains a complete BglII restriction site (5'AGATCT), whereas the nicked strand has G in the place of A, resulting in a G•T mismatch. Right: 5'IDL1 substrate. The top strand carries a deleted A in the BglII site, creating a 1bp loop in the bottom strand. BanII was used to linearize the top strand for a 5' oriented nick 370 bp from the BglII site. Upon correction of the mismatch, using the bottom strand as a template, a complete BglII site emerges. BsaI cuts both repaired and unrepaired plasmids 1360 bp from the BglII site, while BglII cuts only the repaired molecules. Thus, restriction analysis can be used in the evaluation of the repair efficiency.

The substrates were constructed of three nearly identical pGEM plasmids, of which one (pGEM-TA) contained a complete *Bgl*II restriction site (5'AGATCT), whereas pGEM-delA carried a deletion of A (5'-GATCT) and pGEM-CG carried a G instead of T (5'GGATCT) in its *Bgl*II site. The circular strand in the substrate, possessing the complete restriction site, was amplified from pGEM-TA as single-stranded (ss) DNA with bacteriophage M13KO7 and isolated with standard procedures. The top strand, derived from pGEM-GC for 5'G•T or from pGEM-delA for 5'IDL1 substrate, was amplified as dsDNA and digested with *Ban*II restriction enzyme to achieve double-stranded DNA which was linearized 370 bp 5' from the *Bgl*II site. The circular ssDNA (from pGEM-TA) and linear dsDNA (either from pGEM-GC or pGEM-delA) were annealed with a 10-fold molar excess of ssDNA to maximize the yield of the desired circular heteroduplex molecule, minimizing the self-annealing of the linear dsDNA. The excess ssDNA and the

linear double-stranded homoduplex DNA were purified with Benzoylated Naphthoylated DEAE- (BND) Cellulose resin (Sigma) and Plasmid-safe DNAse V (Epicentre), respectively. Details of the MMR assay substrate preparation are available in (Baerenfaller, Fischer & Jiricny 2006).

Preparation of nuclear extracts

Nuclear protein extracts (NE) were prepared because MMR occurs in the nucleus and nuclear extracts contain all the proteins needed for the reaction. LoVo cells were used due to their intrinsic deficiency in MSH2 expression. First, ~5 x 10^8 cells were collected, washed with PBS and swollen in hypotonic buffer (20 mM Hepes, pH 7.9, 5 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM PMSF, 1 x complete EDTA-free protease inhibitor cocktail (Roche), 0.25 μg/ml aprotinin, 0.7 μg/ml pepstatin, 0.5 μg/ml leupeptin) on ice. The cell membranes were disrupted using a syringe with a narrow gauge (No. 27) needle. The nuclei were collected by centrifugation and the resulting pellet was suspended in 1/5 volumes of cold extraction buffer (25 mM Hepes/KOH (pH 7.5), 10% sucrose, 1 mM PMSF, 0.5 mM DTT, 1 μg/ml leupeptine). The salt concentration was adjusted to 155 mM with 5 M NaCl. The mixture was rotated at 4°C for 1 hour and the nuclear debris was pelleted by centrifugation. The soluble nuclear protein fraction was dialysed for 2 hours against 25 mM Hepes/KOH (pH 7.5), 50 mM KCl, 0.1 mM EDTA (pH 8.0), 10% sucrose, 1 mM PMSF, 2 mM DTT, 1 μg/ml leupeptine. The dialysed extract was centrifuged and the supernatant aliquoted and frozen in liquid nitrogen. The NEs were stored at -80°C.

MMR assays

The mismatch repair assays were used to compare the repair activity of the mutated MutSα proteins to that of the WT (I-IV), and to demonstrate the activity of the protein preparations after purification (IV). 75 μg of MSH2-deficient LoVo NE was incubated with either 4 – 12 μg of *Sf*9 TE, including adjusted amounts of recombinant MutSα, or 1 μg of purified recombinant MutSα (III-IV). The reaction contained 100 ng of 5'G•T or 5'IDL1 substrate plasmid, 20 mM Tris-Cl (pH 7.6), 110 mM KCl, 5 mM $MgCl_2$, 1 mM glutathione, 50 μg/ml BSA, 0.1 mM of each dNTP, and 1.5 mM ATP. Successful mismatch repair completed the *Bgl*II restriction site. The repair efficiencies were calculated after a combined *Bsa*I-*Bgl*II digestion and agarose gel electrophoresis by

comparing the intensity of the double-digested bands (repaired) to the *Bsa*I single-digested bands (unrepaired).

Bandshift assays (III-IV)

Preparation of oligomers

For bandshift assays, two double-stranded 38bp DNA oligomers were constructed: one which carried a G•T mispair, and one which was otherwise identical but had an A•T base pair in the corresponding position. The nucleotide sequence of the invariable strand was 5' TTT CTG ACT TGG ATA CCA **T**CT ATC TAT CTA TAA AAT AT 3', and the complementary strand carried either A or G in the indicated (bolded) position. The sequences were synthesised as ssDNA by Oligomer. 10 pmol of the invariable (T) strand was labelled with ^{32}P in a reaction which contained 10 pmol of ^{32}P-γ-ATP, 100 mM DTT, 1X polynucleotide kinase (PNK) buffer and 7,5 units of PNK (Promega). The reaction was incubated in 37°C for 45 min, after which further 5 units of PNK was added, and incubation was continued for 30 min. The labelled ssDNA oligomers were purified using ProbeQuant G-50 Micro Columns (Amersham Biosciences). The complementary ssDNAs were annealed in 1:1 ratio by incubating them for 5 min in 95°C, followed by 5 min in 37°C and 30 min in RT. The double-stranded, labelled oligomers were stored in +4°C.

Mismatch binding

Bandshift assays were used to assess the ability of the MutSα variants to bind heteroduplexed DNA. Corresponding homoduplexes were used as nonspecific DNA controls. 350 ng (65 nM in total) of each MutSα variant was incubated in 10% glycerol, 100 mM KCl, 25 mM Hepes-KOH (pH 7.5), 1 mM DTT, 0.5 mM $MgCl_2$, 0.1 mM ADP, 75 μg/ml BSA, and 60 ng poly-d(I-C) (Amersham Biosciences) for 20 minutes at 37°C with 25 femtomoles of ^{32}P-labelled 38 bp double-stranded DNA oligomers, which contained either a central G•T mismatch (GT heteroduplex) or an A•T base pair in the corresponding position (AT homoduplex).

The binding was visualized by running the reactions in 5% nondenaturing acrylamide gels, which were dried and exposed to a phosphoscreen (Fujifilm BAS-1500). The MutSα-bound oligomers migrate slower than unbound DNA in the gel, thus giving rise to a bandshift. The data was quantified with the TINA software, version 2.09 (OY

Tamro AB), and calculated as percentages of MutSα-bound oligomers of the total amount of labelled DNA.

Mismatch release

MutSα dissociates from DNA mismatches as a consequence of a conformational change, which follows adenine triphosphate (ATP) uptake, when ATP replaces an ADP molecule in the ATP-binding motif of MutSα subunits. To study the ability of MutSα variants to undergo ADP-ATP exchange and to dissociate from G•T mismatches, ATP was added to the bandshift reactions at different concentrations (0.5, 1 and 2 mM) after 10 minutes of incubation, followed by another 10 minutes at 37°C. The concentrations were chosen based on published data, where the addition of 1 mM ATP results in clear but not 100% loss of MutSα mismatch binding (Dufner *et al.* 2000). The fraction of MutSα-bound DNA after ATP addition was calculated as a percentage of the binding level in the absence of ATP in all experiments.

Statistical analysis

All bandshift experiments were repeated at least four times. The binding data were analyzed in SPSS, version 12.0.1. by one-way ANOVA, followed by a Tukey post hoc test. The level of statistical significance was set at 0.05.

RESULTS

Effects of the studied mutations on MSH2 expression, purification and interaction with MSH6 (II-IV)

Baculovirus expression revealed variance in the protein levels (II, IV)

We expressed WT and all mutated MSH2 proteins variants with MSH6 WT in *Sf*9 cells. Variable expression levels were observed (Table 7, page 56). The MSH2 variants T33P, N127S, A272V, G322D, A636P, G674A, E749K, A834T and V923E were expressed in amounts similar to WT MSH2. Although V161D, G162R, G164R, L173P, L187P, C333Y, D603N, and Del 745-746 were also expressed, their expression levels were <5-fold lower as compared to the wild-type protein. C697F showed an intermediate level of expression (~30% of the WT). The amount of MSH6 in the extracts was proportional to the levels of MSH2, supporting the published observations indicating that MSH6 is stable only in complex with MSH2 (Chang *et al.* 2000, Marra *et al.* 1998).

Human cell expression showed no differences in protein levels (II)

To study the stability of the *MSH2* alterations in human cells, we transiently expressed 15 patient-derived MSH2 variants (T33P, V161D, G162R, G164R, L173P, L187P, A272V, C333Y, D603N, C697F, Del 745-746, E749K, A834T, and V923E) together with WT MSH6 (MutSα) in the LoVo (*MSH2*$^{-/-}$) cells. Surprisingly, in human cell expression, all the studied variants showed comparable amounts of MSH2 to the WT protein (Table 7). However, in general the human cell expression system did not produce enough protein for functional studies. Therefore, we used the proteins produced in *Sf*9 cells for all subsequent assays.

Protein purification was successful with efficiently expressed MSH2 variants (III, IV)

Both purification methods, FPLC and Ni-NTA, resulted in highly pure WT and most mutated MutSα complexes (T33P, N127S, A272V, G322D, A636P, E749K, A834T, and V923E). However, adequate purity was not obtained for the variants present in low amounts in the *Sf9* extracts (V161D, G162R, G164R, L173P, L187P, C333Y, D603N, and Del 745-746) (see above). It is therefore possible that these variants are structurally

imperfect, leading to either low expression or insolubility in the *Sf*9 protein extract preparations and to further degradation during the purification process. These variants were omitted from the bandshift assays, which required purified proteins. One variant (C697F) showed an intermediate level of expression and purity and was included in all experiments (Table 7).

MSH2 G674A impaired the MSH2/MSH6 interaction (II-IV)

We performed combined coimmunoprecipitation and Western blot analysis to study the effect of the mutations on the interaction of the MutSα subunits, MSH2 and MSH6. MSH2 G674A was observed to impair the interaction, while all other variants interacted with MSH6 similarly to WT MSH2 (Table 7). Extracts derived from cells expressing the MutSβ complex (MSH2 WT / MSH3 WT heterodimer) gave no detectable signals in the immunoprecipitation assay, confirming that only MSH6 antibody-bound protein complexes were present in the immunoprecipitate (see Original Article IV, Figure 3).

Effects of MSH2 mutations on mismatch repair and binding and release of mismatches (I-IV)

Twelve mutations were MMR-deficient and one showed decreased activity in the in vitro MMR assays (I-IV)

We tested the ability of the recombinant MutSα variants to complement the MMR-defective LoVo extracts in repairing of G•T or IDL1 mispairs *in vitro* and to verify the activity of MutSα after purification (Table 7). By titrating the amount of the WT MutSα TE in the assay, we concluded that the the amount of the variants which were less efficiently produced in the *Sf*9 than the WT was still sufficient for MMR analysis. Altogether 12 out of 18 MutSα variants appeared completely defective in the MMR assay (V161D, G162R, G164R, L173P, L187P, C333Y, D603N, G674A, A636P, C697F, Del 745-746, and E749K). T33P showed a reduced MMR efficiency, and the variants N127S, A272V, G322D, A834T, and V923E showed no deviation from WT in their MMR activity. The WT MutSα was active in MMR assays after purification with both the applied methods, FPLC and Ni-NTA, indicating that the protein complexes remained functional during the purification process.

Mismatch binding was impaired in two and altered in three variants (III-IV)

The bandshift assays were applied to determine the ability of MutSα variants to bind G•T heteroduplex oligonucleotides *in vitro* (Table 7). First, the binding levels of differentially purified MutSα WT extracts were assessed, and both were found to be active in mismatch binding. His-tag-purified MutSα bound G•T mismatches with lower efficiency (17.8 ± 7.6%) than FPLC-purified MutSα (31.6 ± 12.2%). Therefore, the binding efficiency of the analysed variants was always compared to the WT purified with the same method. Variants T33P, N127S, A272V, G322D, G674A, E749K, and A834T recognised and bound G•T mismatches at levels comparable to MutSα WT. The variants A272V, G674A, and V923E showed retained, but compared to the corresponding WT, slightly reduced binding to heteroduplex oligomers (A272V: 9.3 ± 7.6%, G674A: 11.2 ± 5.1%, V923E: 15.7 ± 6.1%). The A636P variant was deficient in DNA binding (2.3 ± 0.5%), displaying only weak background binding, similar as to homoduplex DNA. Also the C697F variant failed to bind the heteroduplex oligomers (1.8 ± 0.4%). The loss of binding activity was not due to interfering proteins in the C697F extract, as mixed with WT MutSα the C697F extract did not interfere with mismatch binding (data not shown). When statistical analysis was applied to the data, only A636P, C697F and V923E showed statistically significant loss of binding activity ($p<0.01$). The figures showing the mismatch binding levels are displayed on the Original Articles III and IV.

Mismatch release was abnormal with MSH2 G674A and E749K mutations (IV)

To study the ability of the MutSα variants to release mismatched DNA upon ATP uptake, ATP was added to the bandshift reaction (Table 7). Only the mismatch binding-proficient variants (T33P, N127S, A272V, G322D, G674A, E749K, A834T and V923E) and the WT proteins were included in the experiment. The DNA binding efficiencies after ATP addition were quantified and calculated as percentages of the maximum amount of protein-bound DNA (*i.e.* binding in the absence of ATP). The WT MutSα dissociated efficiently already at the lowest used ATP concentration (0.5 mM), where the binding was reduced to 12.9 ± 6.8% (WT-his) or 12.4 ± 6.2% (WT-FPLC) of the maximum. Adding 1 mM ATP further decreased the binding to 7.2 ± 3.2% (WT-his) and 6.1 ± 2.9% (WT-FPLC), and 2 mM ATP to 7.1 ± 4.3% (WT-his) and 5.0 ± 3.0% (WT-FPLC). The variants N127S, A272V, and A834T dissociated from DNA as efficiently as WT protein, whereas the variants T33P, G322D, and V923E showed slightly reduced release compared to WT

protein (T33P: 0.5 mM: 21.7 ± 4.5%, 1 mM: 16.8 ± 6.7%, 2 mM: 13.3 ± 5.0%, G322D: 0.5 mM: 23.8 ± 23.7%, 1 mM: 15.3 ± 14.0%, 2 mM: 9.6 ± 7.5%, V923E: 0.5 mM: 29.0 ± 13.3%, 1 mM: 16.1 ± 9.6%, 2 mM: 9.7 ± 8.5%). G674A and E749K displayed reduced mismatch release at all ATP concentrations, E749K dissociating most aberrantly at highest ATP concentration as compared to the WT protein (G674A: 0.5 mM: 65.5 ± 21.8%, 1 mM: 47.0 ± 19.5%, 2 mM: 33.6 ± 23.1%; E749K: 0.5 mM: 30.8 ± 21.0%, 1 mM: 21.2 ± 15.0%, 2 mM: 21.1 ± 16.3%). E749K and G674A showed statistically significant impairment of mismatch release (G674A with all ATP concentrations, $p<0.01$, and E749K with 1mM and 2 mM ATP, $p<0.05$). The figures of mismatch release are shown in the Original Article IV.

Determination of pathogenicity in MSH2 variants

The final interpretation of the pathogenicity of a variation was made based on both functional and clinical data so, that the variants which showed a clear functional defect and were associated with typical HNPCC phenotype, were interpretated pathogenic. Variants with no or only minor functional defects and variable clinical data were interpretated non-pathogenic. All studied mutations, together with the location of the mutation, fulfilment of AC, IHC result, results of functional tests, and the overall interpretation of their pathogenicity are collected in Table 8. 12 out of 18 studied MSH2 variants displayed severe defects in the functional assays and were interpreted as pathogenic. In addition, two displayed slight defects two assays, and their interpretation was left inconclusive. They may have a milder or no phenotypic effect. Four variants were functional in all assays or had only a minor alteration in one, and were interpreted as non-pathogenic.

SIFT analysis compared to functional results of MSH2 missense variants (II, III)

In our dataset, SIFT predictions on tolerability of MSH2 amino acid substitutions were performed by our collaborators Marc Greenblatt and Phil Chan from the University of Vermont. The SIFT predictions were compared to the pathogenicity of the variants, which was finally determined using both functional and clinical data. SIFT correctly predicted the outcome (tolerated or deleterious, corresponding to non-pathogenic and pathogenic) of thirteen analysed amino acid substitutions (Table 7). Only N127S and A636P SIFT predictions were not confirmed by the functional data. For variants T33P and V923E, the

interpretation of their pathogenicity was not clear, due to milder clinical phenotype and reduced but not absent functionality in MMR assays. T33P was predicted non-tolerated by SIFT. For V923E, only one (otherwise best fitted) comparative sequence alignment predicted it to be tolerated, while alternative alignments predicted it to be deleterious (II). The results of all functional assays and the SIFT prediction scores are summarised in Table 7.

Table 7. Results of all functional analyses conducted with MSH2 variants

Amino acid change	Protein domain[a]	Human cell expression[b]	Baculovirus expression[b]	MSH6 Interaction	Purification	MMR assay	Mismatch binding	Mismatch release	SIFT prediction / cut-off score[c]
T33P	1 DNA binding	+	+	+	+	+/-	+	+ (-)	Deleterious / 0.02
N127S	2 Connector	NA	+	+	+	+	+	+	Deleterious / 0.01
V161D	2 Connector	+	-	+	-	-	NA	NA	Deleterious / 0.00
G162R	2 Connector	+	-	+	-	-	NA	NA	Deleterious / 0.00
G164R	2 Connector	+	-	+	-	-	NA	NA	Deleterious / 0.02
L173P	2 Connector	+	-	+	-	-	NA	NA	Deleterious / 0.00
L187P	2 Connector	+	-	+	-	-	NA	NA	Deleterious / 0.00
A272V	2 Connector	+	+	+	+	+	+ (-)	+	Tolerated / 0.79
G322D	3 Levers	NA	+	+	+	+	+	+ (-)	Tolerated / 0.53
C333Y	3 Levers	+	-	+	-	-	NA	NA	Deleterious / 0.00
D603N	3 Levers	+	-	+	-	-	NA	NA	Deleterious / 0.01
A636P	5 ATPase	+	+	+	+	+	-	NA	Tolerated / 0.20
G674A	5 ATPase	NA	+	-	+	-	+ (-)	-	Deleterious / 0.00
C697F	5 ATPase	+	+/-	+	+/-	-	-	NA	Deleterious / 0.00
Del745-756	5 ATPase	+	-	+	-	-	NA	NA	NA
E749K	5 ATPase	+	+	+	+	-	+	-	Deleterious / 0.00
A834T	5 ATPase	+	+	+	+	+	+	+	Tolerated / 0.10
V923E	5 Helix-turn-helix	+	+	+	+	+	+ (-)	+ (-)	Tolerated / 0.89

All functional assays except for the human cell expression are based on the recombinant proteins produced with the baculovirus system. NA, not assessed. a, the functional domain in which the mutated amino acid is located. b, +: expression approximately at the level of WT, -: expression level <20% of WT. +/-: expression level about 30% of WT. c, SIFT cut-off score < 0,05 was interpretated as deleterious.

Correlation between the functional assays and clinical phenotypes of mutation carriers (I-IV)

In general, the functional and clinical data correlated well. In summary, *MSH2* mutations V161D, G162R, G164R, L173P, L187P, C333Y, D603N, A636P, G675A, C697F, Del 745-746, and E749K were defective in at least two functional assays and displayed phenotypic characteristics which supported interpreting them as pathogenic. T33P showed reduced activity in the MMR assay. V923E exhibited normal MMR activity but reduced mismatch binding. However, as the functional evidence on their pathogenicity was not obvious, and the clinical data pointed to a milder than typical HNPCC phenotype, the interpretation of T33P and V923E was left inconclusive. 10 out of 11 pathogenic mutations which had available family data were derived from families which fulfilled AC I or II, although D603N and A636P were also found from AC negative families (Table 5, p. 41 and Table 8). In contrary, 3 out of 4 mutations, which were MMR proficient and showed no or only minor changes in mismatch binding or release, N127S, A272V and G322D, were associated with variable clinical backgrounds and interpretated as non-pathogenic. In the family carrying the non-pathogenic A834T mutation, an additional mutation, deletion of exon 8 of MSH2, was found after completion of the functional studies. This deletion segregated with the cancer phenotype and likely to be the cause for HNPCC in this family (E. Mangold, personal communication).

Loss or reduction of MSH2 expression in a tumour, as assessed by IHC, was evident in 9 out of 10 studied tumours associated with a pathogenic mutation, proving the loss of MSH2 expression to be highly indicative of a predisposition to HNPCC. On the other hand, MSH2 protein was present in three out of four tumours associated with the functional variants. All variants, which were unstable in the in vitro expression system, also showed loss of MSH2 in the tumour. E749K was the only variation, which was MMR deficient and concluded to be pathogenic, but still expressed MSH2 normally.

Table 8. The clinical and functional characteristics and overall interpretation of the studied MSH2 mutations.

MH2 mutation	Clinical data		Functional data				Interpretation
	AC[a]	MSH2 IHC[b]	Repair[c]	Expression	Binding	Release	
T33P	-	+	+ (-)	+	+	+(-)	Inconclusive[f]
N127S[d]	-/+	-/+	+	+	+	+	Non-pathogenic
V161D	+	-	-	-	NA	NA	Pathogenic
G162R	+	-	-	-	NA	NA	Pathogenic
G164R	+	-	-	-	NA	NA	Pathogenic
L173P	+	-	-	-	NA	NA	Pathogenic
L187P	+	-	-	-	NA	NA	Pathogenic
A272V[d]	-/+	+	+	+	+(-)	+	Non-pathogenic
G322D[d]	-/+	-/+	+	+	+	+(-)	Non-pathogenic
C333Y	-	NA	-	-	NA	NA	Pathogenic
D603N[d]	-/+	-	-	-	NA	NA	Pathogenic
A636P	-/+	-	-	+	-	NA	Pathogenic
G674A	NA	NA	-	+	+(-)	-	Pathogenic
C697F	+	-	-	+ (-)	-	NA	Pathogenic
Del745-756	+	-	-	-	NA	NA	Pathogenic
E749K	+	+	-	+	+	-	Pathogenic
A834T[e]	+	-	+	+	+	+	Non-pathogenic
V923E	+	-	+	+	+ (-)	+(-)	Inconclusive[f]

NA: Not assessed. Clinical data is presented in detail in Original Article II. a. Fulfilment of Amsterdam Criteria (AC) I or II. +: fulfilled, -: not fulfilled, -/+: AC status varies between associated families. b. Immunohistochemical analysis of MSH2 protein expression in a tumour. +: MSH2 expressed, -: MSH2 not expressed, -/+: IHC status varies between reported cases. c. in vitro MMR capability of the mutation. +: functional, -: deficient, + (-): reduced. d. Mutation associated with variable clinical characteristics. e. In this family, another deleterious MSH2 mutation was later found and confirmed as the underlying cause of cancer predisposition. f. Variant was functional but showed reduced efficiency in one assay, suggesting non-pathogenicity or partial pathogenicity.

DISCUSSION

In hereditary nonpolyposis colorectal cancer (HNPCC), frequent surveillance significantly reduces the mortality of the mutation carriers (Aarnio *et al.* 1999b, Mecklin *et al.* 2007) but identification of the family members at risk requires reliable molecular diagnosis. If a found mutation is nontruncating, the diagnosis becomes complicated, as nontruncating alterations appear to be associated with a wide variety of clinical phenotypes, ranging from normal to highly increased cancer risk (Raevaara *et al.* 2005). Therefore, reliable establishment of the functional consequence of a found genetic variation is very important and facilitates the assessment of the risk of cancer. In this study, 18 nontruncating *MSH2* mutations, which had been found in cancer patients, were investigated to gain insight into their participation in HNPCC tumorigenesis, and to clarify the mechanism by which they interfere with MMR. The patients displayed a variety of phenotypes, ranging from typical HNPCC and Muir-Torre syndrome (MTS) characteristics to milder cancer phenotypes in terms of age at onset or tumour penetrance. We assessed the expression, MMR activity, mismatch binding, mismatch release, and subunit interaction capabilities of the mutated MSH2 proteins corresponding to the identified HNPCC-associated *MSH2* mutations.

The analyses revealed severe defects in 12 out of 18 studied variants in the functional assays, and, supported by the clinical data, the respective mutations were interpreted as the causative reason for the cancer phenotype in the mutation carriers. The associated families were typical HNPCC families, which mostly fulfilled the diagnostic AC. Two variants showed retained but clearly reduced activity in at least one assay, suggesting milder pathogenicity. In the case of T33P, only two affected individuals were identified, and V923E was associated with an abnormally high age of onset. The determination of pathogenicity regarding these two variants was left inconclusive in this work. Four variants were completely functional or displayed only a minor, statistically insignificant alteration in one assay, suggesting that these are non-pathogenic. The carriers of these 4 variants displayed a variety of phenotypes, ranging from healthy individuals to CRC phenotypes, which were, however, often associated with atypical characteristics, such as MSH2 expression in the tumour. This suggests that the cancers in these mutation carriers were either not hereditary, or due to other, unidentified MMR gene defects. In

case of patients carrying the A834T variation, the latter was shown to be true. After completion of the functional analyses, an exonic deletion in MSH2 found and shown to cosegregate with the cancer phenotype.

MSH2 L187P and C697F predispose to HNPCC and MTS

Muir-Torre syndrome (MTS) is a condition characterised by the metachronous or synchronous occurrence of skin and visceral cancers (Cohen *et al.* 1995, Schwartz *et al.* 1989), and due to its molecular etiology being connected to germline MMR gene mutations, it is considered a subtype of HNPCC (Kruse *et al.* 1998). The functionality of MSH2 L187P and C697F were studied because of their connection to both HNPCC and MTS families and the uncertainty regarding the role of their variations in cancer pathogenesis. At the time of the study (I), only 3 out of 41 MMR gene mutations which were reported in literature to be linked to MTS were of the missense type. This was in contrast to HNPCC-associated MMR mutations, of which about 22% are missense variations (www.insight-group.org). None of the missense mutations associated with MTS had been functionally assessed in homologous human systems. Thus, the involvement of these alterations in the tumorigenesis of combined skin and visceral cancers remained unverified. Therefore, in order to validate the connection of *MSH2* missense alterations to the syndrome and to ensure the molecular defect underlying the cancer phenotype in the associated families, the functional analysis of the two putatively MTS predisposing variations was considered of high importance. Molecular analysis would then allow the diagnosis of MTS in patients carrying these germline alterations and exhibiting skin malignancies already before the manifestation of a visceral cancer. The analysis revealed that missense mutations L187P and C697F in MSH2 completely inactivate MMR, thereby being the underlying cause of the MSI phenotype and tumour formation in the described families. Our data demonstrate that, albeit rarely, nontruncating mutations do underlie the MTS phenotype as well as HNPCC. We suggest that an MSI-positive skin lesion combined with molecular analysis revealing a missense mutation in an MMR gene should lead to the functional analysis of the given variation to ensure a solid diagnosis.

No evidence for MSH2 N127S and G322D -linked cancer predisposition

N127S (c.380 A>G) and G322D (c.965 G>A) are two of the most frequently occurring *MSH2* variants. They are classified inconsistently in the literature (Table 6, page 42), and conclusive evidence of their role in cancer predisposition is missing or contradictory. In some databases they are listed as single nucleotide polymorphisms (SNPs) occurring with a frequency of 0 – 9,2% (N127S) and 0 – 6,5% (G322D) (www.ensembl.org, www.genome. utah.edu/genesnps/), and as pathogenic variations in others (www.insight-group.org, www.missense.org). These contradictory reports and the absence of functional data on the human proteins led us to study the functionality of MSH2 N127S and G322D to provide support either for their predisposing or non-predisposing roles in HNPCC tumorigenesis. The MMR assay, co-immunoprecipitation and bandshift assays did not reveal differences in their functionality as compared to WT MSH2. In G322D, the dissociation from mismatches was observed to be slightly reduced, but this was not statistically significant.

However, as several reports found both N127S and G322D to be associated with HNPCC families, we performed a systematic literature search to gain a broader view on the occurrence of these mutations. Furthermore, we looked for reports providing evidence for additive effects of other MMR mutations found in N127S and G322D carriers. Indeed, we found that both N127S and G322D frequently coexist with other MMR gene mutations (Table 6), and that there is evidence suggesting that N127S may have an additive effect in a family where a truncating *MSH2* mutation is the primary cause of HNPCC. In the study of Tanyi *et al.* (2006), the occurrence of the truncating mutation together with N127S was observed to lower the age of cancer onset, as compared to the carriers of the truncating mutation only. Carriers of only N127S were asymptomatic. This suggests that N127S on its own does not predispose to cancer, but, in combination with a deleterious mutation, it may have an additive effect.

Based on literature reports, the non-pathogenicity of G322D is more strongly established than that of N127S. In a recent population study, G322D was even reported to occur more frequently in the healthy than in the CRC patient population (Barnetson *et al.* 2007). However, the reports of cancer patients where G322D is the only found MMR variant (Table 6) rule out the possibility that this variation could be strictly classified as clinically non-relevant. Our interpretation is that G322D on its own does not cause MMR malfunction, and thus most likely does not promote colorectal tumorigenesis.

The majority of the studied MSH2 variants inactivated MMR

From previous reports by our research group it was evident that a significant proportion of *MLH1* and *MSH6* missense mutations do not inactivate MMR (Kariola *et al.* 2002a, Kariola *et al.* 2004, Raevaara *et al.* 2003, Raevaara *et al.* 2005). In *MLH1*, many variants were functional in the MMR assay but were shown to be unstable, and some variations inhibited the correct nuclear transport of the protein (Raevaara *et al.* 2005). In *MLH1*, 10 out of 34, and in *MSH6*, 9 out of 11 studied mutations did not show any functional defects, and the clinical phenotypes of the cancer patients carrying these non-pathogenic mutations were variable (Kariola *et al.* 2002b, Kariola *et al.* 2004, Raevaara *et al*. 2005). In *MSH2*, we found that 12 of the 18 investigated mutations (V161D, G162R, G164R, L173P, L187P, C333Y, D603N, A636P, G674A, C697F, Del745-746, and E749K) abolished MMR totally, and one (T33P) caused a moderate reduction in the MMR activity. Five mutations (N127S, A272V, G332D, A834T, and V923E) had similar repair activity to WT MSH2. All mutations with absent MMR activity were interpreted as pathogenic, while the pathogenicity of T33P remained questionable.

Decreased protein expression was associated with MMR deficiency in N-terminal MSH2 missense mutations

In order to assess the mechanisms behind MMR deficiency, more detailed analyses in the characterisation of the MSH2 variants were applied. To assess their mismatch-binding and release capabilities, the MSH2 variants were re-expressed with his-tagged WT MSH6 and purified. Notably, the variation in expression levels which occurred already in the first round of protein expression (II) was repeated in the second round (IV), and led to difficulties in purification of some of the variants. Therefore, we concluded that the variants V161D, G162R, G164R, L173P, L187P, C333Y, D603N, and Del 745-746 cause problems in the expression of MSH2 or its stability. C697F was expressed in amounts lower than the WT but higher than the other unstable variants, and was classified as moderately expressed. Seven out of nine poorly expressed variants localised to either the N-terminal connector or lever domains of the MSH2 protein.

The connector domain connects the DNA-binding subunit of MutSα to the rest of the MutSα heterodimer. It is responsible for the intramolecular interactions and allosteric signalling between different protein domains (Warren *et al.* 2007). (Figure 3, p.29). The

lever domain connects the ATPase domain to the clamp domain, which makes unspecific DNA contacts. It is believed to communicate signals defining the conformation of the protein between the ATP- and DNA-binding parts of the protein. Due to the stabilising and connecting roles of connector and lever domains, it is not surprising that many mutations situated in those domains led to protein conformation or stability problems and, thereby, to defective MMR. The only unstable variants situated outside these domains were C697F and Del 745-746. These mutations are located in the ATPase domain. C697F displayed moderate stability, and was purified to reasonable purity. The Del 745-746 mutation is a two amino acid deletion and therefore more severe than a missense mutation. Thus, it is also likely to have a larger effect on the conformation of the protein.

In agreement with our results, reduced stability of MSH2 proteins has been also reported to be the most frequent explanation for MMR deficiency in a recent yeast study (Gammie *et al.* 2007). There, 54 HNPCC-associated missense mutations in MSH2 were constructed in cognate positions in yeast Msh2 and functionally analysed. 50% of the mutations, predominantly situated in the connector and lever domains of MSH2, were shown to be associated with reduced stability.

To investigate the expression of the mutant proteins in human cells, 15 variants (T33P, V161D, G162R, G164R, L173P, L187P, A272V, C333Y, D603N, A636P, C697F, Del 745-746, E749K, A834T and V923E) were transiently expressed in *MSH2*-deficient human colon carcinoma LoVo cells. Surprisingly, all 15 variants were expressed in similar amounts as WT MSH2. The reason for this discrepancy between insect and human cell expression remains unclear, but probably reflects the different conditions within the different cell types and the differences in the mode of overexpression in the separate systems. However, the half-life of the studied proteins was not assessed, and thus, even though the human cells expressed them as WT, the stability may still be affected. The occasionally observed clonal variation in baculovirus expression is unlikely to be the reason affecting the varying protein levels, as similar results were obtained in two completely independent rounds of expression. Due to difficulties in transfecting other MSH2-negative human cell lines, we were not able to exclude the possibility that the observed behaviour is unique for LoVo cells.

Pathogenic mutations in the ATPase domain mostly interfered with mismatch binding or release

The mutations studied in this work were distributed along the length of the MSH2 polypeptide, with some clustering observed in the connector and ATPase domains (Figure 4, p. 40). As discussed, the likely reason for the pathogenicity of the N-terminal mutations in the connector domain was the low expression or instability of the resulting MSH2 protein. In contrast, only 2 out of 7 studied mutations (C697F and Del 745-746) in the ATPase domain showed reduced expression. Instead, A636P, G674A, C697F, and E749K showed clearly impaired mismatch binding or release capabilities, as assessed by bandshift analysis. V923E, which is situated in the helix-turn-helix (H-T-H) motif in the extreme C-terminus of MSH2, exhibited reduced mismatch binding. As the H-T-H motif is involved in the ATPase function by stabilising the ATPase domains in MutSα (Warren *et al.* 2007), V923E is discussed along with the ATPase site mutations. Furthermore, A272V and G674A showed minor impairment of DNA binding, and T33P, G322D and V923E in mismatch release. These slight alterations were not statistically significant.

Since A636P and C697F were completely deficient in mismatch binding, their mismatch release activities could not be measured, and the mismatch binding problems were concluded to be the causative reason for their pathogenicity. V923E showed reduced binding, and although it was MMR-proficient, the weaker binding could indicate a subtle MMR defect, too mild to be detected in the *in vitro* assay. However, a mild defect could explain the relatively high average age of cancer onset in the V923E family (58 years, Table 5, p. 41). However, in the absence of an *in vitro* MMR defect, the interpretation of V923E remains unconclusive. G674A and E749K bound mismatches similarly to the WT, but their mismatch release was impaired. These two mutations affect the most conserved amino acid residues in the superfamily of ABC transporter ATPase domains, to which the ATPase domains of MSH2 and MSH6 belong (Locher 2004). G674A is located in the Walker A and E749K in the Walker B domain, two of the most essential regions of the ATPase domain. As proper ATP processing is vital for MutSα function in MMR, these mutations cause malfunctions in both mismatch binding and release. Defective mismatch binding is likely to abolish all MMR-mediated functions, as MutSα needs to recognise the abnormal base to initiate MMR or any other related process, such as apoptosis signalling, another important activity of MMR proteins.

Interestingly, it has been proposed that unlike mismatch binding, the ATP processing activity is not required for apoptosis signalling mediated by MMR proteins. In three mouse models for *MSH2, MSH6* and *MLH1*, a point mutation introduced into the respective ATPase domains inactivates MMR, but the cells of the animals undergo apoptosis in response to alkylating agents similarly to WT cells (Avdievich *et al.* 2008, Lin *et al.* 2004, Yang *et al.* 2004). The same phenomenon has been observed in yeast MutS homologs, where certain ATPase mutants inactivated MMR but remained sensitive to cisplatin, unlike the Msh2 / Msh6-negative strains (Drotschmann *et al.* 2004). Therefore, missense mutations in the ATPase site may be of clinical interest due to the remaining apoptosis signalling activity, and might result in unaltered responses to certain anti-tumour therapies, such as cisplatin or temozolomide, which in general are not effective in MMR-defective tumours (Stojic, Brun & Jiricny 2004).

MSH2 G674A displayed reduced capability to interact with MSH6

Only one protein variant, G674A, interfered with the MSH2/MSH6 interaction in the co-immunoprecipitation assay, whereas D603N and V923E, which are in the MSH2-MSH6 interaction regions (Guerrette *et al.* 1998) did not show a decrease in the interaction. This is probably explained by the fact that MSH2 and MSH6 have two distinct interaction sites (codons 378-625 and 875-934 in *MSH2*) and it has been suggested that the loss of function in one region would not result in complete loss of heterodimer formation (Guerrette *et al.* 1998). Thus, it is possible that even if some of the mutations interfered with dimerisation, MSH2 and MSH6 would not be completely detached and would thus precipitate together. G674A affects the Walker A domain of MSH2 ATPase site, which makes direct contacts to the ABC transporter unit of MSH6 (Warren *et al.* 2007). The amino acid substitution in the α-helix may be sufficient to adversely affect the interaction of the two proteins. Notably, the interaction problem became evident only in the co-immunoprecipitation assay, where a stringent RIPA buffer is used. Under more gentle conditions the dimer remained stable, as demonstrated by the successful purification of MutSα-G674A (IV).

Phenotypic characteristics of the mutation carriers correlated with the functional data

Most of the families displaying the pathogenic *MSH2* mutations fulfilled the international AC for HNPCC, showed an early age at cancer onset, as well as high MSI and loss of MSH2 protein in tumours (Table 5). Concerning the pathogenic mutations, the C333Y, one D603N family and two A636P families did not fulfil the ACI/II. However, D603N and A636P are also connected to AC-positive families, and the functional data support the pathogenicity of these mutations. In C333Y, data of only two individuals was available, but they both had early-onsed CRC. In the case of T33P, the MMR activity was reduced but not absent, and the family did not fulfil AC, nor did it show deficient MSH2 expression. However, the MSI status was high. We concluded that the reduced but partially retained MMR activity correlates with the milder clinical phenotype in the family as compared to families connected to fully pathogenic variations. However, as the IHC staining for MSH2 protein carrying T33P mutation was positive, the cancer accumulation may also be unrelated to this variant. The variant V923E functioned similarly to WT MSH2 in the MMR assay but displayed lowered mismatch binding. Although the family carrying this mutation fulfils the AC, the average age at onset is abnormally high for HNPCC (58 years, Table 5).

Regarding the non-pathogenic variations, tumours in the A272V mutation carriers showed MSH2 expression, supporting our interpretation of non-pathogenicity. However, the tumours displayed an MSI phenotype. MLH1 was expressed normally in the tumour of the A272V carrier, so MLH1 somatic inactivation by promoter hypermethylation, which is a frequent cause of MSI in colorectal tumours, is also unlikely. The mutation A834T, also proficient in the MMR assay, has been found in the healthy control population (Genuardi *et al.* 1999). However, the clinical phenotype of the mutation carriers showed an early age at onset as well as high MSI and loss of MSH2 in the tumour. As discussed previously, (p. 60 – 61), an additional, truncating mutation (deletion of exon 8) in MSH2 was found in the A834T family after the functional studies were conducted (E. Mangold, personal communication). This suggests that this deletion, rather than A834T, was the causative reason of HNPCC phenotype in the investigated family. The phenotypic characteristics of the N127S and G322D mutation carriers were discussed in detail above.

IHC as a screening tool for HNPCC

In our set of MMR-deficient *MSH2* variants, MMR deficiency was mainly associated with typical HNPCC characteristics, among them loss of MSH2 protein in the respective tumours. Only one tumour, which was associated with the pathogenic mutations, E749K, showed MSH2 expression in IHC analyses. One, T33P was only mildly defective in the MMR assay, and thus it is possible that it is not the reason for cancer predisposition in the family. Unlike T33P, E749K showed complete MMR deficiency, defective mismatch release and a remarkably young age at onset (29 years) in the mutation carrier, reflecting the strong mutator effect caused by the mutation.

E749K is located in a highly conserved Walker B domain in the MSH2 ATPase region and it was shown in our analyses to impair ATP-provoked mismatch release. Thus, we suggest that the mode of MMR impairment connected to this mutation is the abnormally strong binding to mismatches, which may lead to defects in the recruitment of the downstream MMR machinery, which is known to require ATP hydrolysis (Dufner *et al.* 2000, Iaccarino *et al.* 1998). The above mode of pathogenicity may explain the retained expression of the MSH2 protein in the tumour. This result demonstrates, that certain ATPase mutations in *MSH2* may be, although pathogenic, not identifiable by IHC analysis. The finding is also supported by the report of the missense mutation T1217D in the ATPase domain of mouse *MSH6*. This mutation is MMR-deficient, but expresses MSH6 protein in the tumour (Yang *et al.* 2004), although pathogenic *MSH6* mutations usually associate with loss of protein expression (Hampel *et al.* 2005). Unfortunately, data for MSH2 expression in the tumours of MSH2 G674A mice, or patients carrying the MSH2 G674D mutation, are not available.

In conclusion, our data suggests that some ATPase site mutations in MSH2 are not identifiable with IHC, although overall the correlation between the absence of staining for MSH2 and pathogenicity of the missense mutation was notable and supports IHC analysis as a sensitive method in HNPCC diagnostics with MSH2 mutations.

Classification of nontruncating MSH2 mutations

In light of investigations by us and others, it seems evident that the prevailing system of classification of nontruncating MMR (and other genetic) variants to pathogenic and non-pathogenic is an oversimplification and does not reflect the real situation where a missense

mutation may compromise the function of the protein in a significantly different way than a truncating mutation. For example, some missense mutations in *MLH1* lead to mild HNPCC phenotypes in regard to cancer penetrance or age at onset (Peltomäki, Gao & Mecklin 2001, Raevaara *et al.* 2005). Also, low levels of MLH1 are enough for MMR, but not sufficient for MLH1-mediated cell cycle arrest (Cejka *et al.* 2003). Accordingly, shortage of the mutated protein, rather than its MMR deficiency, has been postulated as the mechanism of pathogenicity in an HNPCC family with a nontruncating MLH1 mutation (Raevaara *et al.* 2004). Furthermore, many MMR gene variants have been reported as both pathogenic and neutral (www.insight-group.org). This suggests that other MMR gene variations in the affected individuals, or the genetic background in general, may in combination with the nontruncating MMR gene variant raise the risk of cancer.

Thus, in many cases it is simply not possible to evaluate the phenotypic consequences of a nontruncating variation based on any single parameter. The combined use of tumour studies, segregation analysis, population data, conservation of the amino acids in question, and functional analysis provide the most reliable assessment of pathogenicity. If both population and functional data support the non-pathogenic interpretation, the variant probably does not, at least on its own, contribute to familial cancer. A functionally deficient mutation for which there is positive data showing co-segregation of the mutation and the cancer phenotype can be classified as pathogenic. However, the mutations which occur in healthy individuals and cancer families alike, or which are associated with atypical HNPCC characteristics and display only modest functional defects, could be classified as low penetrance variants.

CONCLUSIONS

The functional studies on 18 nontruncating variants in *MSH2*, performed as described in this PhD thesis, led to following conclusions:

- In Muir-Torre syndrome, skin tumours combined with the identification of a germline MMR gene mutation can provide diagnostic clues for consequent visceral cancers. Nontruncating MMR gene mutations found in skin cancer patients should therefore be functionally assessed.
- Pathogenicity of MSH2 missense mutations is typically associated with an impaired repair capability of the mutated protein.
- MMR deficiency associated with nontruncating mutations in *MSH2* is mediated through different mechanisms. Those include defective protein expression or stability, and impaired mismatch binding or release.
- Reduced protein expression is predominantly associated to mutations in the connector and lever domains of MSH2.
- Mutations affecting the ATPase domain of MSH2 are mostly stable *in vitro*, but display problems in mismatch binding or release.
- Mutations classified as pathogenic based on functional assays associate mainly with typical HNPCC phenotypes in mutation carriers.
- Mutations, which display only mild defects in functional assays, may associate with a milder HNPCC phenotype.
- Variants, which act like the WT in functional assays, are associated with variable clinical phenotypes.
- Most pathogenic mutations associate with a lack of MSH2 protein expression in tumour tissue. However, some pathogenic mutations abrogating the mismatch release but not protein stability may still retain MSH2 expression. Thus, immunohistochemical analysis is insufficient to exclude a MMR defect.

FUTURE PROSPECTS

To supplement the present study, it would be interesting to

- Further characterise the stability of the studied MSH2 variants in different human cell lines and by applying protein half-life measurements
- Study the effects of nontruncating MSH2 mutations on the MMR-mediated response to cytotoxic drugs
- Study the effects of nontruncating MSH2 mutations on the subcellular localisation of the protein
- Characterise the impact of the mutations on MutSα-MutLα ternary complex formation
- Study the mutations in complex with MSH3

ACKNOWLEDGEMENTS

This study has been carried out at the Department of Biological and Environmental Sciences, Division of Genetics, University of Helsinki, during the years 2004 – 2008. Funding for the studies was provided by the Sigrid Juselius Foundation, Academy of Finland, the Finnish Cancer Foundation, the Research Foundation of the University of Helsinki, and the Paulo Foundation. Hereby I wish to thank all the people directly and indirectly contributing to the fact that this thesis came into daylight.

First, I sincerely thank my supervisor, docent Minna Nyström. Minna has been a great example of strong willpower, organised research, and excellent financial management. Along with that, she has always been emphatic, understanding and supportive. And she always shows up so full of energy and happiness! Without Minna's strong devotion to her students, getting my PhD would have been a lot rockier road.

I am thankful to Professor Klaus Elenius and docent Helmut Pospiech for reviewing this PhD thesis. They both provided great ideas on how to improve it, and did it with professional yet friendly way. Jack Leo is acknowledged for correcting the English language of the thesis, and for his efforts with the crystal structure figures.

Our research group has been a good place to work and I warmly thank my colleagues for help, support, and for their patience with my occasional waves of over-excitement (or misery…). Tiina Raevaara is thanked especially for great lunch-break discussions on evolution and homeopathy. Mari Korhonen had to put up with me when we were the only members of the group ☺ - thanks for the uncountable favours you've done for me in the lab! Laura Sarantaus is thanked for her great contribution to my first two papers. Denis Dermadi, with his excitement and energy, was so much fun to do research with; I warmly thank him for his great work. Reetta Kariola is like my big sister in science, I am truly grateful for all her help and for the many good discussions and laughs we have had together. Jukka Kantelinen is a fun guy and good in the lab; it has been a pleasure to get to know him. Also all other present and past members of the group are thanked for all their help and for the good lab spirit.

All our Finnish and international co-workers are thanked for providing the clinical data of the families, for great contributions in writing the manuscripts, and for fruitful

collaboration in general. I wish to especially acknowledge Professor Päivi Peltomäki's group for educative meetings, and Marc Greenblatt for the SIFT work.

While doing research I was also teaching regularly, and I thank my fellow teacher, docent Pekka Heino for making it fun and relaxed. I thank Professor Jim Schröder for our discussions and for his general everyday warmth and friendliness. For technical and practical issues, I would have fallen into despair several times without the two Arjas, Ikävalko and Välimäki. Thank you both! The Division of Genetics, led by Professor Tapio Palva, has been an excellent facility for research. Everybody working in Genetics is thanked for the easy-going working atmosphere, which has made my time here so enjoyable. Special thanks to C.P.O.C. (Christmas Party Organizing Committee)!

Many people have had an impact on my road in the world of science, also before the times in Helsinki. My high school biology teacher Markku Milonoff posed us intriguing questions in class, making us think of our views of life, crystallising in my head the idea that biology is something amazingly fascinating. During my Master studies in the University of Turku, I was treated with warmth, respect and intelligence by professor Marja-Liisa Savontaus. Her way of supporting and encouraging students in the world of genetics impressed me deeply, and I thank her for that. Docents Seppo and Christina Nokkala are warmly acknowledged for the opportunity to start teaching, and for their great support during my years in the Department of Genetics in Turku.

I have had precious opportunities to work also in labs abroad. I thank Karl Heinimann from the University of Basel for supervising my Master thesis and for collaboration thereafter. Professor Josef Jiricny from the University of Zürich is warmly thanked for the research period I had in his laboratory during my PhD studies. His scientific mind and warm personality are exceptional. Big thanks also go to Katja Baerenfaller and Franziska Fischer, who taught me to purify proteins and bind them on radioactive pieces of DNA.

Science would be difficult to manage without a good life also outside the lab. Luckily I have fantastic parents, who have provided me with endless understanding in all my choices, both good and bad. That has always given me a feeling of freedom, yet with constant support. I am also thankful to have an adventurous brother, Erkki, and a fantastic sister, Onerva, and in general a pretty great family.

I thank all my friends for countless great moments, which have been totally unrelated to any scientific issues, but so necessary for sanity and life-enjoyment!

Especially thanked are childhood friends from Valkeakoski: Maija T, Maija H, Mari, Saikku and Anna-Kaisa; coursemates from biology studies: Mervi, Terhi, Reetta, Lauri, Liisa, Antero, Tuula, and Manta; from the big SHO crew: Sampo, Päivi, Reetta, and Teresa; and Sanna and Heljä from England days and forward. I feel privileged to have such cool friends. Maija T and Mervi deserve special thanks for also being of help in scientific issues.

Mouse chromosomes are acknowledged for bringing Jaan-Olle to my life ☺. His genuine happiness and excitement towards science and life in general are amazing, and his love and support are invaluable to me.

Helsinki, August 2008

Saara Ollila

REFERENCES

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

Aarnio, M., Sankila, R., Pukkala, E., Salovaara, R., Aaltonen, L.A., de la Chapelle, A., Peltomäki, P., Mecklin, J.P. and Jarvinen, H.J. 1999a. Cancer risk in mutation carriers of DNA-mismatch-repair genes. *Int. J. Cancer.* 81: 214-218.

Aarnio, M., Sankila, R., Pukkala, E., Salovaara, R., Aaltonen, L.A., de la Chapelle, A., Peltomäki, P., Mecklin, J.P. and Jarvinen, H.J. 1999b. Cancer risk in mutation carriers of DNA-mismatch-repair genes. *Int. J. Cancer.* 81: 214-218.

Aboussekhra, A., Biggerstaff, M., Shivji, M.K., Vilpo, J.A., Moncollin, V., Podust, V.N., Protic, M., Hubscher, U., Egly, J.M. and Wood, R.D. 1995. Mammalian DNA nucleotide excision repair reconstituted with purified protein components. *Cell.* 80: 859-868.

Abraham, R.T. 2001. Cell cycle checkpoint signaling through the ATM and ATR kinases. *Genes Dev.* 15: 2177-2196.

Acharya, S., Wilson, T., Gradia, S., Kane, M.F., Guerrette, S., Marsischky, G.T., Kolodner, R. and Fishel, R. 1996. hMSH2 forms specific mispair-binding complexes with hMSH3 and hMSH6. *Proc. Natl. Acad. Sci. U. S. A.* 93: 13629-13634.

Aktan-Collan, K., Mecklin, J.P., Jarvinen, H., Nyström-Lahti, M., Peltomäki, P., Soderling, I., Uutela, A., de la Chapelle, A. and Kaariainen, H. 2000. Predictive genetic testing for hereditary non-polyposis colorectal cancer: Uptake and long-term satisfaction. *Int. J. Cancer.* 89: 44-50.

Al-Tassan, N., Chmiel, N.H., Maynard, J., Fleming, N., Livingston, A.L., Williams, G.T., Hodges, A.K., Davies, D.R., David, S.S., Sampson, J.R. and Cheadle, J.P. 2002. Inherited variants of MYH associated with somatic G:C-->T:A mutations in colorectal tumors. *Nat. Genet.* 30: 227-232.

Andressoo, J.O., Hoeijmakers, J.H. and Mitchell, J.R. 2006. Nucleotide excision repair disorders and the balance between cancer and aging. *Cell. Cycle.* 5: 2886-2888.

Avdievich, E., Reiss, C., Scherer, S.J., Zhang, Y., Maier, S.M., Jin, B., Hou, H.,Jr, Rosenwald, A., Riedmiller, H., Kucherlapati, R., Cohen, P.E., Edelmann, W. and Kneitz, B. 2008. Distinct effects of the recurrent Mlh1G67R mutation on MMR functions, cancer, and meiosis. *Proc. Natl. Acad. Sci. U. S. A.* 105: 4247-4252.

Baerenfaller, K., Fischer, F. and Jiricny, J. 2006. Characterization of the "mismatch repairosome" and its role in the processing of modified nucleosides in vitro. *Methods Enzymol.* 408: 285-303.

Bapat, B.V., Madlensky, L., Temple, L.K., Hiruki, T., Redston, M., Baron, D.L., Xia, L., Marcus, V.A., Soravia, C., Mitri, A., Shen, W., Gryfe, R., Berk, T., Chodirker, B.N., Cohen, Z. and Gallinger, S. 1999. Family history characteristics, tumor microsatellite instability and germline MSH2 and MLH1 mutations in hereditary colorectal cancer. *Hum. Genet.* 104: 167-176.

Barnetson, R.A., Cartwright, N., van Vliet, A., Haq, N., Drew, K., Farrington, S., Williams, N., Warner, J., Campbell, H., Porteous, M.E. and Dunlop, M.G. 2007. Classification of ambiguous mutations in DNA mismatch repair genes identified in a population-based study of colorectal cancer. *Hum. Mutat.*

Blackwell, L.J., Martik, D., Bjornson, K.P., Bjornson, E.S. and Modrich, P. 1998. Nucleotide-promoted release of hMutSalpha from heteroduplex DNA is consistent with an ATP-dependent translocation mechanism. *J. Biol. Chem.* 273: 32055-32062.

Bocker, T., Barusevicius, A., Snowden, T., Rasio, D., Guerrette, S., Robbins, D., Schmidt, C., Burczak, J., Croce, C.M., Copeland, T., Kovatich, A.J. and Fishel, R. 1999. hMSH5: A human MutS homologue that forms a novel heterodimer with hMSH4 and is expressed during spermatogenesis. *Cancer Res.* 59: 816-822.

Boland, C.R., Thibodeau, S.N., Hamilton, S.R., Sidransky, D., Eshleman, J.R., Burt, R.W., Meltzer, S.J., Rodriguez-Bigas, M.A., Fodde, R., Ranzani, G.N. and Srivastava, S. 1998. A national cancer institute workshop on microsatellite instability for cancer detection and familial predisposition: Development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res.* 58: 5248-5257.

Bowers, J., Tran, P.T., Joshi, A., Liskay, R.M. and Alani, E. 2001. MSH-MLH complexes formed at a DNA mismatch are disrupted by the PCNA sliding clamp. *J. Mol. Biol.* 306: 957-968.

Broughton, B.C., Cordonnier, A., Kleijer, W.J., Jaspers, N.G., Fawcett, H., Raams, A., Garritsen, V.H., Stary, A., Avril, M.F., Boudsocq, F., Masutani, C., Hanaoka, F., Fuchs, R.P., Sarasin, A. and Lehmann, A.R. 2002. Molecular analysis of mutations in DNA polymerase eta in xeroderma pigmentosum-variant patients. *Proc. Natl. Acad. Sci. U. S. A.* 99: 815-820.

Burdett, V., Baitinger, C., Viswanathan, M., Lovett, S.T. and Modrich, P. 2001. In vivo requirement for RecJ, ExoVII, ExoI, and ExoX in methyl-directed mismatch repair. *Proc. Natl. Acad. Sci. U. S. A.* 98: 6765-6770.

Cannavo, E., Marra, G., Sabates-Bellver, J., Menigatti, M., Lipkin, S.M., Fischer, F., Cejka, P. and Jiricny, J. 2005. Expression of the MutL homologue hMLH3 in human cells and its role in DNA mismatch repair. *Cancer Res.* 65: 10759-10766.

Cejka, P., Marra, G., Hemmerle, C., Cannavo, E., Storchova, Z. and Jiricny, J. 2003. Differential killing of mismatch repair-deficient and -proficient cells: Towards the therapy of tumors with microsatellite instability. *Cancer Res.* 63: 8113-8117.

Chang, D.K., Ricciardiello, L., Goel, A., Chang, C.L. and Boland, C.R. 2000. Steady-state regulation of the human DNA mismatch repair system. *J. Biol. Chem.* 275: 29178.

Christmann, M., Tomicic, M.T., Roos, W.P. and Kaina, B. 2003. Mechanisms of human DNA repair: An update. *Toxicology.* 193: 3-34.

Clark, A.B., Cook, M.E., Tran, H.T., Gordenin, D.A., Resnick, M.A. and Kunkel, T.A. 1999. Functional analysis of human MutSalpha and MutSbeta complexes in yeast. *Nucleic Acids Res.* 27: 736-742.

Cleaver, J.E. 1968. Defective repair replication of DNA in xeroderma pigmentosum. *Nature.* 218: 652-656.

Clendenning, M., Baze, M.E., Sun, S., Walsh, K., Liyanarachchi, S., Fix, D., Schunemann, V., Comeras, I., Deacon, M., Lynch, J.F., Gong, G., Thomas, B.C., Thibodeau, S.N., Lynch, H.T., Hampel, H. and de la Chapelle, A. 2008. Origins and prevalence of the american founder mutation of MSH2. *Cancer Res.* 68: 2145-2153.

Cohen, P.R., Kohn, S.R., Davis, D.A. and Kurzrock, R. 1995. Muir-torre syndrome. *Dermatol. Clin.* 13: 79-89.

Constantin, N., Dzantiev, L., Kadyrov, F.A. and Modrich, P. 2005. Human mismatch repair: Reconstitution of a nick-directed bidirectional reaction. *J. Biol. Chem.* 280: 39752-39761.

Diehl, K.M., Keller, E.T. and Ignatoski, K.M. 2007. Why should we still care about oncogenes? *Mol. Cancer. Ther.* 6: 418-427.

Drake, C.G., Jaffee, E. and Pardoll, D.M. 2006. Mechanisms of immune evasion by tumors. *Adv. Immunol.* 90: 51-81.

Dronkert, M.L. and Kanaar, R. 2001. Repair of DNA interstrand cross-links. *Mutat. Res.* 486: 217-247.

Drotschmann, K., Clark, A.B. and Kunkel, T.A. 1999. Mutator phenotypes of common polymorphisms and missense mutations in MSH2. *Curr. Biol.* 9: 907-910.

Drotschmann, K., Clark, A.B., Tran, H.T., Resnick, M.A., Gordenin, D.A. and Kunkel, T.A. 1999. Mutator phenotypes of yeast strains heterozygous for mutations in the MSH2 gene. *Proc. Natl. Acad. Sci. U. S. A.* 96: 2970-2975.

Drotschmann, K., Topping, R.P., Clodfelter, J.E. and Salsbury, F.R. 2004. Mutations in the nucleotide-binding domain of MutS homologs uncouple cell death from cell survival. *DNA Repair (Amst).* 3: 729-742.

Drummond, J.T., Li, G.M., Longley, M.J. and Modrich, P. 1995. Isolation of an hMSH2-p160 heterodimer that restores DNA mismatch repair to tumor cells. *Science.* 268: 1909-1912.

Dufner, P., Marra, G., Raschle, M. and Jiricny, J. 2000. Mismatch recognition and DNA-dependent stimulation of the ATPase activity of hMutSalpha is abolished by a single mutation in the hMSH6 subunit. *J. Biol. Chem.* 275: 36550-36555.

Dumstorf, C.A., Clark, A.B., Lin, Q., Kissling, G.E., Yuan, T., Kucherlapati, R., McGregor, W.G. and Kunkel, T.A. 2006. Participation of mouse DNA polymerase iota in strand-biased mutagenic bypass of UV photoproducts and suppression of skin cancer. *Proc. Natl. Acad. Sci. U. S. A.* 103: 18083-18088.

Duval, A. and Hamelin, R. 2002. Mutations at coding repeat sequences in mismatch repair-deficient human cancers: Toward a new concept of target genes for instability. *Cancer Res.* 62: 2447-2454.

Ellison, A.R., Lofing, J. and Bitter, G.A. 2001. Functional analysis of human MLH1 and MSH2 missense variants and hybrid human-yeast MLH1 proteins in saccharomyces cerevisiae. *Hum. Mol. Genet.* 10: 1889-1900.

Fackenthal, J.D. and Olopade, O.I. 2007. Breast cancer risk associated with BRCA1 and BRCA2 in diverse populations. *Nat. Rev. Cancer.* 7: 937-948.

Fearon, E.R. 1997. Human cancer syndromes: Clues to the origin and nature of cancer. *Science.* 278: 1043-1050.

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

Fishel, R. and Kolodner, R.D. 1995. Identification of mismatch repair genes and their role in the development of cancer. *Curr. Opin. Genet. Dev.* 5: 382-395.

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

Fousteri, M. and Mullenders, L.H. 2008. Transcription-coupled nucleotide excision repair in mammalian cells: Molecular mechanisms and biological effects. *Cell Res.* 18: 73-84.

Froggatt, N.J., Green, J., Brassett, C., Evans, D.G., Bishop, D.T., Kolodner, R. and Maher, E.R. 1999. A common MSH2 mutation in english and north american HNPCC families: Origin, phenotypic expression, and sex specific differences in colorectal cancer. *J. Med. Genet.* 36: 97-102.

Frosina, G., Fortini, P., Rossi, O., Carrozzino, F., Raspaglio, G., Cox, L.S., Lane, D.P., Abbondandolo, A. and Dogliotti, E. 1996. Two pathways for base excision repair in mammalian cells. *J. Biol. Chem.* 271: 9573-9578.

Gammie, A.E., Erdeniz, N., Beaver, J., Devlin, B., Nanji, A. and Rose, M.D. 2007. Functional characterization of pathogenic human MSH2 missense mutations in saccharomyces cerevisiae. *Genetics.* 177: 707-721.

Genuardi, M., Carrara, S., Anti, M., Ponz de Leon, M. and Viel, A. 1999. Assessment of pathogenicity criteria for constitutional missense mutations of the hereditary nonpolyposis colorectal cancer genes MLH1 and MSH2. *Eur. J. Hum. Genet.* 7: 778-782.

Gradia, S., Acharya, S. and Fishel, R. 2000. The role of mismatched nucleotides in activating the hMSH2-hMSH6 molecular switch. *J. Biol. Chem.* 275: 3922-3930.

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

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

Gradia, S., Subramanian, D., Wilson, T., Acharya, S., Makhov, A., Griffith, J. and Fishel, R. 1999. hMSH2-hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA. *Mol. Cell.* 3: 255-261.

Grilley, M., Griffith, J. and Modrich, P. 1993. Bidirectional excision in methyl-directed mismatch repair. *J. Biol. Chem.* 268: 11830-11837.

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

Guerrette, S., Wilson, T., Gradia, S. and 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.

Guillem, J.G., Moore, H.G., Palmer, C., Glogowski, E., Finch, R., Nafa, K., Markowitz, A.J., Offit, K. and Ellis, N.A. 2004. A636P testing in ashkenazi jews. *Fam. Cancer.* 3: 223-227.

Guillem, J.G., Rapaport, B.S., Kirchhoff, T., Kolachana, P., Nafa, K., Glogowski, E., Finch, R., Huang, H., Foulkes, W.D., Markowitz, A., Ellis, N.A. and Offit, K. 2003. A636P is associated with early-onset colon cancer in ashkenazi jews. *J. Am. Coll. Surg.* 196: 222-225.

Hakem, R. 2008. DNA-damage repair; the good, the bad, and the ugly. *EMBO J.* 27: 589-605.

Hampel, H., Frankel, W.L., Martin, E., Arnold, M., Khanduja, K., Kuebler, P., Nakagawa, H., Sotamaa, K., Prior, T.W., Westman, J., Panescu, J., Fix, D., Lockman, J., Comeras, I. and de la Chapelle, A. 2005. Screening for the lynch syndrome (hereditary nonpolyposis colorectal cancer). *N. Engl. J. Med.* 352: 1851-1860.

Hanahan, D. and Weinberg, R.A. 2000. The hallmarks of cancer. *Cell.* 100: 57-70.

Hegde, M.L., Hazra, T.K. and Mitra, S. 2008. Early steps in the DNA base excision/single-strand interruption repair pathway in mammalian cells. *Cell Res.* 18: 27-47.

Helleday, T. 2003. Pathways for mitotic homologous recombination in mammalian cells. *Mutat. Res.* 532: 103-115.

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

Holmes, J.,Jr, Clark, S. and 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.

Hunter, W.N., Brown, T., Kneale, G., Anand, N.N., Rabinovich, D. and Kennard, O. 1987. The structure of guanosine-thymidine mismatches in B-DNA at 2.5-A resolution. *J. Biol. Chem.* 262: 9962-9970.

Iaccarino, I., Marra, G., Palombo, F. and Jiricny, J. 1998. hMSH2 and hMSH6 play distinct roles in mismatch binding and contribute differently to the ATPase activity of hMutSalpha. *EMBO J.* 17: 2677-2686.

Jackson, S.P. 2002. Sensing and repairing DNA double-strand breaks. *Carcinogenesis.* 23: 687-696.

Jarvinen, H.J., Aarnio, M., Mustonen, H., Aktan-Collan, K., Aaltonen, L.A., Peltomäki, P., De La Chapelle, A. and Mecklin, J.P. 2000. Controlled 15-year trial on screening for colorectal cancer in families with hereditary nonpolyposis colorectal cancer. *Gastroenterology.* 118: 829-834.

Jass, J.R. 2008. Colorectal polyposes: From phenotype to diagnosis. *Pathol. Res. Pract.* 204: 431-447.

Jemal, A., Siegel, R., Ward, E., Hao, Y., Xu, J., Murray, T. and Thun, M.J. 2008. Cancer statistics, 2008. *CA Cancer. J. Clin.* 58: 71-96.

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

Jiricny, J. and Nyström-Lahti, M. 2000. Mismatch repair defects in cancer. *Curr. Opin. Genet. Dev.* 10: 157-161.

Kadyrov, F.A., Dzantiev, L., Constantin, N. and Modrich, P. 2006. Endonucleolytic function of MutLalpha in human mismatch repair. *Cell.* 126: 297-308.

Kariola, R., Hampel, H., Frankel, W.L., Raevaara, T.E., de la Chapelle, A. and Nyström-Lahti, M. 2004. MSH6 missense mutations are often associated with no or low cancer susceptibility. *Br. J. Cancer.* 91: 1287-1292.

Kariola, R., Raevaara, T.E., Lonnqvist, K.E. and Nyström-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.

Kariola, R., Raevaara, T.E., Lonnqvist, K.E. and Nyström-Lahti, M. 2002b. Functional analysis of MSH6 mutations linked to kindreds with putative hereditary non-polyposis colorectal cancer syndrome. *Hum. Mol. Genet.* 11: 1303-1310.

Karran, P. 2001. Mechanisms of tolerance to DNA damaging therapeutic drugs. *Carcinogenesis.* 22: 1931-1937.

Kawamoto, T., Araki, K., Sonoda, E., Yamashita, Y.M., Harada, K., Kikuchi, K., Masutani, C., Hanaoka, F., Nozaki, K., Hashimoto, N. and Takeda, S. 2005. Dual roles for DNA polymerase eta in homologous DNA recombination and translesion DNA synthesis. *Mol. Cell.* 20: 793-799.

Kleczkowska, H.E., Marra, G., Lettieri, T. and Jiricny, J. 2001. hMSH3 and hMSH6 interact with PCNA and colocalize with it to replication foci. *Genes Dev.* 15: 724-736.

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

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

Kolodner, R.D. and Marsischky, G.T. 1999. Eukaryotic DNA mismatch repair. *Curr. Opin. Genet. Dev.* 9: 89-96.

Kondo, E., Suzuki, H., Horii, A. and 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.

Korhonen, M.K., Raevaara, T.E., Lohi, H. and Nyström, M. 2007. 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.* 17: 351-354.

Kruse, R., Rutten, A., Lamberti, C., Hosseiny-Malayeri, H.R., Wang, Y., Ruelfs, C., Jungck, M., Mathiak, M., Ruzicka, T., Hartschuh, W., Bisceglia, M., Friedl, W. and Propping, P. 1998. Muir-torre phenotype has a frequency of DNA mismatch-repair-gene mutations similar to that in hereditary nonpolyposis colorectal cancer families defined by the amsterdam criteria. *Am. J. Hum. Genet.* 63: 63-70.

Kruse, R. and Ruzicka, T. 2004a. DNA mismatch repair and the significance of a sebaceous skin tumor for visceral cancer prevention. *Trends Mol. Med.* 10: 136-141.

Kruse, R. and Ruzicka, T. 2004b. DNA mismatch repair and the significance of a sebaceous skin tumor for visceral cancer prevention. *Trends Mol. Med.* 10: 136-141.

Kunkel, T.A. 1993. Nucleotide repeats. slippery DNA and diseases. *Nature.* 365: 207-208.

Kunkel, T.A. and Erie, D.A. 2005. DNA mismatch repair. *Annu. Rev. Biochem.* 74: 681-710.

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

Lamers, M.H., Perrakis, A., Enzlin, J.H., Winterwerp, H.H., de Wind, N. and Sixma, T.K. 2000a. The crystal structure of DNA mismatch repair protein MutS binding to a G x T mismatch. *Nature.* 407: 711-717.

Lamers, M.H., Perrakis, A., Enzlin, J.H., Winterwerp, H.H., de Wind, N. and Sixma, T.K. 2000b. The crystal structure of DNA mismatch repair protein MutS binding to a G x T mismatch. *Nature.* 407: 711-717.

Leach, F.S., Nicolaides, N.C., Papadopoulos, N., Liu, B., Jen, J., Parsons, R., Peltomäki, P., Sistonen, P., Aaltonen, L.A. and Nyström-Lahti, M. 1993. Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. *Cell.* 75: 1215-1225.

Li, G.M. and 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.

Li, X. and Heyer, W.D. 2008. Homologous recombination in DNA repair and DNA damage tolerance. *Cell Res.* 18: 99-113.

Lieber, M.R., Ma, Y., Pannicke, U. and Schwarz, K. 2004. The mechanism of vertebrate nonhomologous DNA end joining and its role in V(D)J recombination. *DNA Repair (Amst).* 3: 817-826.

Lin, D.P., Wang, Y., Scherer, S.J., Clark, A.B., Yang, K., Avdievich, E., Jin, B., Werling, U., Parris, T., Kurihara, N., Umar, A., Kucherlapati, R., Lipkin, M., Kunkel, T.A. and Edelmann, W. 2004. An Msh2 point mutation uncouples DNA mismatch repair and apoptosis. *Cancer Res.* 64: 517-522.

Lin, Q., Clark, A.B., McCulloch, S.D., Yuan, T., Bronson, R.T., Kunkel, T.A. and Kucherlapati, R. 2006. Increased susceptibility to UV-induced skin carcinogenesis in polymerase eta-deficient mice. *Cancer Res.* 66: 87-94.

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

Locher, K.P. 2004. Structure and mechanism of ABC transporters. *Curr. Opin. Struct. Biol.* 14: 426-431.

Lynch, H.T. and de la Chapelle, A. 2003. Hereditary colorectal cancer. *N. Engl. J. Med.* 348: 919-932.

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

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

Lynch, H.T. and Krush, A.J. 1971. Cancer family "G" revisited: 1895-1970. *Cancer.* 27: 1505-1511.

Lynch, H.T., Smyrk, T. and Lynch, J.F. 1998. Molecular genetics and clinical-pathology features of hereditary nonpolyposis colorectal carcinoma (lynch syndrome): Historical journey from pedigree anecdote to molecular genetic confirmation. *Oncology.* 55: 103-108.

Malkhosyan, S., Rampino, N., Yamamoto, H. and Perucho, M. 1996. Frameshift mutator mutations. *Nature.* 382: 499-500.

Malkin, D., Li, F.P., Strong, L.C., Fraumeni, J.F.,Jr, Nelson, C.E., Kim, D.H., Kassel, J., Gryka, M.A., Bischoff, F.Z. and Tainsky, M.A. 1990. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science.* 250: 1233-1238.

Mangold, E., Pagenstecher, C., Friedl, W., Fischer, H.P., Merkelbach-Bruse, S., Ohlendorf, M., Friedrichs, N., Aretz, S., Buettner, R., Propping, P. and Mathiak, M. 2005. Tumours from MSH2 mutation carriers show loss of MSH2 expression but many tumours from MLH1 mutation carriers exhibit weak positive MLH1 staining. *J. Pathol.* 207: 385-395.

Mangold, E., Pagenstecher, C., Leister, M., Mathiak, M., Rutten, A., Friedl, W., Propping, P., Ruzicka, T. and Kruse, R. 2004. A genotype-phenotype correlation in HNPCC: Strong predominance of msh2 mutations in 41 patients with muir-torre syndrome. *J. Med. Genet.* 41: 567-572.

Mangold, E., Rahner, N., Friedrichs, N., Buettner, R., Pagenstecher, C., Aretz, S., Friedl, W., Ruzicka, T., Propping, P., Rutten, A. and Kruse, R. 2007. MSH6 mutation in muir-torre syndrome: Could this be a rare finding? *Br. J. Dermatol.* 156: 158-162.

Manley, K., Shirley, T.L., Flaherty, L. and Messer, A. 1999. Msh2 deficiency prevents in vivo somatic instability of the CAG repeat in huntington disease transgenic mice. *Nat. Genet.* 23: 471-473.

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

Marra, G., Iaccarino, I., Lettieri, T., Roscilli, G., Delmastro, P. and Jiricny, J. 1998. Mismatch repair deficiency associated with overexpression of the MSH3 gene. *Proc. Natl. Acad. Sci. U. S. A.* 95: 8568-8573.

Marsh, D. and Zori, R. 2002. Genetic insights into familial cancers-- update and recent discoveries. *Cancer Lett.* 181: 125-164.

Matsuura, S., Tauchi, H., Nakamura, A., Kondo, N., Sakamoto, S., Endo, S., Smeets, D., Solder, B., Belohradsky, B.H., Der Kaloustian, V.M., Oshimura, M., Isomura, M., Nakamura, Y. and Komatsu, K. 1998. Positional cloning of the gene for nijmegen breakage syndrome. *Nat. Genet.* 19: 179-181.

McCulloch, S.D. and Kunkel, T.A. 2008. The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. *Cell Res.* 18: 148-161.

Mecklin, J.P., Aarnio, M., Laara, E., Kairaluoma, M.V., Pylvanainen, K., Peltomäki, P., Aaltonen, L.A. and Jarvinen, H.J. 2007. Development of colorectal tumors in colonoscopic surveillance in lynch syndrome. *Gastroenterology.* 133: 1093-1098.

Modrich, P. and Lahue, R. 1996. Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu. Rev. Biochem.* 65: 101-133.

Murphy, H.R., Armstrong, R., Cairns, D. and Greenhalgh, K.L. 2008. Muir-torre syndrome: Expanding the genotype and phenotype-a further family with a MSH6 mutation. *Fam. Cancer.*

Ng, P.C. and Henikoff, S. 2003. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* 31: 3812-3814.

Nyström-Lahti, M., Kristo, P., Nicolaides, N.C., Chang, S.Y., Aaltonen, L.A., Moisio, A.L., Jarvinen, H.J., Mecklin, J.P., Kinzler, K.W. and Vogelstein, B. 1995. Founding mutations and alu-mediated recombination in hereditary colon cancer. *Nat. Med.* 1: 1203-1206.

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

Obmolova, G., Ban, C., Hsieh, P. and Yang, W. 2000. Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA. *Nature.* 407: 703-710.

Ogi, T. and Lehmann, A.R. 2006. The Y-family DNA polymerase kappa (pol kappa) functions in mammalian nucleotide-excision repair. *Nat. Cell Biol.* 8: 640-642.

Ou, J., Niessen, R.C., Lutzen, A., Sijmons, R.H., Kleibeuker, J.H., de Wind, N., Rasmussen, L.J. and Hofstra, R.M. 2007. Functional analysis helps to clarify the clinical importance of unclassified variants in DNA mismatch repair genes. *Hum. Mutat.*

Palombo, F., Iaccarino, I., Nakajima, E., Ikejima, M., Shimada, T. and Jiricny, J. 1996. hMutSbeta, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA. *Curr. Biol.* 6: 1181-1184.

Papadopoulos, N., Nicolaides, N.C., Wei, Y.F., Ruben, S.M., Carter, K.C., Rosen, C.A., Haseltine, W.A., Fleischmann, R.D., Fraser, C.M. and Adams, M.D. 1994. Mutation of a mutL homolog in hereditary colon cancer. *Science.* 263: 1625-1629.

Patel, K.J. and Joenje, H. 2007. Fanconi anemia and DNA replication repair. *DNA Repair (Amst).* 6: 885-890.

Peled, J.U., Kuang, F.L., Iglesias-Ussel, M.D., Roa, S., Kalis, S.L., Goodman, M.F. and Scharff, M.D. 2008. The biochemistry of somatic hypermutation. *Annu. Rev. Immunol.* 26: 481-511.

Peltomäki, P. 2005. Lynch syndrome genes. *Fam. Cancer.* 4: 227-232.

Peltomäki, P., Aaltonen, L., Mecklin, J.P. and de la Chapelle, A. 1993a. A breakthrough in solving the genetic background of colon cancer. *Duodecim.* 109: 1367-1369.

Peltomäki, P., Aaltonen, L.A., Sistonen, P., Pylkkanen, L., Mecklin, J.P., Jarvinen, H., Green, J.S., Jass, J.R., Weber, J.L. and Leach, F.S. 1993b. Genetic mapping of a locus predisposing to human colorectal cancer. *Science.* 260: 810-812.

Peltomäki, P., Gao, X. and Mecklin, J.P. 2001. Genotype and phenotype in hereditary nonpolyposis colon cancer: A study of families with different vs. shared predisposing mutations. *Fam. Cancer.* 1: 9-15.

Ponti, G., Losi, L., Di Gregorio, C., Roncucci, L., Pedroni, M., Scarselli, A., Benatti, P., Seidenari, S., Pellacani, G., Lembo, L., Rossi, G., Marino, M., Lucci-Cordisco, E. and de Leon, M.P. 2005. Identification of muir-torre syndrome among patients with sebaceous tumors and keratoacanthomas: Role of clinical features, microsatellite instability, and immunohistochemistry. *Cancer.* 103: 1018-1025.

Ponti, G., Losi, L., Pedroni, M., Lucci-Cordisco, E., Di Gregorio, C., Pellacani, G. and Seidenari, S. 2006. Value of MLH1 and MSH2 mutations in the appearance of muir-torre syndrome phenotype in HNPCC patients presenting sebaceous gland tumors or keratoacanthomas. *J. Invest. Dermatol.* 126: 2302-2307.

Prasad, R., Bebenek, K., Hou, E., Shock, D.D., Beard, W.A., Woodgate, R., Kunkel, T.A. and Wilson, S.H. 2003. Localization of the deoxyribose phosphate lyase active site in human DNA polymerase iota by controlled proteolysis. *J. Biol. Chem.* 278: 29649-29654.

Rada, C., Ehrenstein, M.R., Neuberger, M.S. and Milstein, C. 1998. Hot spot focusing of somatic hypermutation in MSH2-deficient mice suggests two stages of mutational targeting. *Immunity*. 9: 135-141.

Raedle, J., Trojan, J., Brieger, A., Weber, N., Schafer, D., Plotz, G., Staib-Sebler, E., Kriener, S., Lorenz, M. and Zeuzem, S. 2001. Bethesda guidelines: Relation to microsatellite instability and MLH1 promoter methylation in patients with colorectal cancer. *Ann. Intern. Med.* 135: 566-576.

Raevaara, T.E., Gerdes, A.M., Lonnqvist, K.E., Tybjaerg-Hansen, A., Abdel-Rahman, W.M., Kariola, R., Peltomäki, P. and Nyström-Lahti, M. 2004. HNPCC mutation MLH1 P648S makes the functional protein unstable, and homozygosity predisposes to mild neurofibromatosis type 1. *Genes Chromosomes Cancer*. 40: 261-265.

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

Raevaara, T.E., Vaccaro, C., Abdel-Rahman, W.M., Mocetti, E., Bala, S., Lonnqvist, K.E., Kariola, R., Lynch, H.T., Peltomäki, P. and Nyström-Lahti, M. 2003. Pathogenicity of the hereditary colorectal cancer mutation hMLH1 del616 linked to shortage of the functional protein. *Gastroenterology*. 125: 501-509.

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

Raschle, M., Marra, G., Nyström-Lahti, M., Schar, P. and Jiricny, J. 1999. Identification of hMutLbeta, a heterodimer of hMLH1 and hPMS1. *J. Biol. Chem.* 274: 32368-32375.

Reenan, R.A. and Kolodner, R.D. 1992. Isolation and characterization of two saccharomyces cerevisiae genes encoding homologs of the bacterial HexA and MutS mismatch repair proteins. *Genetics*. 132: 963-973.

Rich, T., Allen, R.L. and Wyllie, A.H. 2000. Defying death after DNA damage. *Nature*. 407: 777-783.

Rodriguez-Bigas, M.A., Boland, C.R., Hamilton, S.R., Henson, D.E., Jass, J.R., Khan, P.M., Lynch, H., Perucho, M., Smyrk, T., Sobin, L. and Srivastava, S. 1997. A national cancer institute workshop on hereditary nonpolyposis colorectal cancer syndrome: Meeting highlights and bethesda guidelines. *J. Natl. Cancer Inst.* 89: 1758-1762.

Rouse, J. and Jackson, S.P. 2002. Interfaces between the detection, signaling, and repair of DNA damage. *Science*. 297: 547-551.

Salovaara, R., Loukola, A., Kristo, P., Kaariainen, H., Ahtola, H., Eskelinen, M., Harkonen, N., Julkunen, R., Kangas, E., Ojala, S., Tulikoura, J., Valkamo, E., Jarvinen, H., Mecklin, J.P., Aaltonen, L.A. and de la Chapelle, A. 2000. Population-based molecular detection of hereditary nonpolyposis colorectal cancer. *J. Clin. Oncol*. 18: 2193-2200.

Savitsky, K., Bar-Shira, A., Gilad, S., Rotman, G., Ziv, Y., Vanagaite, L., Tagle, D.A., Smith, S., Uziel, T., Sfez, S., Ashkenazi, M., Pecker, I., Frydman, M., Harnik, R., Patanjali, S.R., Simmons, A., Clines, G.A., Sartiel, A., Gatti, R.A., Chessa, L., Sanal, O., Lavin, M.F., Jaspers, N.G., Taylor, A.M., Arlett, C.F., Miki, T., Weissman, S.M., Lovett, M., Collins, F.S. and Shiloh, Y. 1995. A single ataxia telangiectasia gene with a product similar to PI-3 kinase. *Science*. 268: 1749-1753.

Savouret, C., Garcia-Cordier, C., Megret, J., te Riele, H., Junien, C. and Gourdon, G. 2004. MSH2-dependent germinal CTG repeat expansions are produced continuously in spermatogonia from DM1 transgenic mice. *Mol. Cell. Biol.* 24: 629-637.

Schwartz, R.A., Goldberg, D.J., Mahmood, F., DeJager, R.L., Lambert, W.C., Najem, A.Z. and Cohen, P.J. 1989. The muir-torre syndrome: A disease of sebaceous and colonic neoplasms. *Dermatologica.* 178: 23-28.

Shcherbakova, P.V. and Kunkel, T.A. 1999. Mutator phenotypes conferred by MLH1 overexpression and by heterozygosity for mlh1 mutations. *Mol. Cell. Biol.* 19: 3177-3183.

Sherr, C.J. 2004. Principles of tumor suppression. *Cell.* 116: 235-246.

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

Shuck, S.C., Short, E.A. and Turchi, J.J. 2008. Eukaryotic nucleotide excision repair: From understanding mechanisms to influencing biology. *Cell Res.* 18: 64-72.

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

Soussi, T., Ishioka, C., Claustres, M. and Beroud, C. 2006. Locus-specific mutation databases: Pitfalls and good practice based on the p53 experience. *Nat. Rev. Cancer.* 6: 83-90.

South, C.D., Hampel, H., Comeras, I., Westman, J.A., Frankel, W.L. and de la Chapelle, A. 2008. The frequency of muir-torre syndrome among lynch syndrome families. *J. Natl. Cancer Inst.*

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

Srivastava, S., Zou, Z.Q., Pirollo, K., Blattner, W. and Chang, E.H. 1990. Germ-line transmission of a mutated p53 gene in a cancer-prone family with li-fraumeni syndrome. *Nature.* 348: 747-749.

Stasiak, A.Z., Larquet, E., Stasiak, A., Muller, S., Engel, A., Van Dyck, E., West, S.C. and Egelman, E.H. 2000. The human Rad52 protein exists as a heptameric ring. *Curr. Biol.* 10: 337-340.

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

Surtees, J.A., Argueso, J.L. and Alani, E. 2004. Mismatch repair proteins: Key regulators of genetic recombination. *Cytogenet. Genome Res.* 107: 146-159.

Takahashi, M., Shimodaira, H., Andreutti-Zaugg, C., Iggo, R., Kolodner, R.D. and Ishioka, C. 2007. Functional analysis of human MLH1 variants using yeast and in vitro mismatch repair assays. *Cancer Res.* 67: 4595-4604.

Tanyi, M., Olasz, J., Lukacs, G., Csuka, O., Toth, L., Szentirmay, Z., Ress, Z., Barta, Z., Tanyi, J.L. and Damjanovich, L. 2006. Pedigree and genetic analysis of a novel mutation carrier patient suffering from hereditary nonpolyposis colorectal cancer. *World J. Gastroenterol.* 12: 1192-1197.

Thibodeau, S.N., French, A.J., Cunningham, J.M., Tester, D., Burgart, L.J., Roche, P.C., McDonnell, S.K., Schaid, D.J., Vockley, C.W., Michels, V.V., Farr, G.H.,Jr and O'Connell, M.J. 1998. Microsatellite

instability in colorectal cancer: Different mutator phenotypes and the principal involvement of hMLH1. *Cancer Res.* 58: 1713-1718.

Thomas, D.C., Roberts, J.D. and Kunkel, T.A. 1991a. Heteroduplex repair in extracts of human HeLa cells. *J. Biol. Chem.* 266: 3744-3751.

Thomas, D.C., Roberts, J.D. and Kunkel, T.A. 1991b. Heteroduplex repair in extracts of human HeLa cells. *J. Biol. Chem.* 266: 3744-3751.

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

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

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

Vasen, H.F., Mecklin, J.P., Khan, P.M. and Lynch, H.T. 1991. The international collaborative group on hereditary non-polyposis colorectal cancer (ICG-HNPCC). *Dis. Colon Rectum.* 34: 424-425.

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

Vasen, H.F., Wijnen, J.T., Menko, F.H., Kleibeuker, J.H., Taal, B.G., Griffioen, G., Nagengast, F.M., Meijers-Heijboer, E.H., Bertario, L., Varesco, L., Bisgaard, M.L., Mohr, J., Fodde, R. and Khan, P.M. 1996. Cancer risk in families with hereditary nonpolyposis colorectal cancer diagnosed by mutation analysis. *Gastroenterology.* 110: 1020-1027.

Wang, T.F., Kleckner, N. and Hunter, N. 1999. Functional specificity of MutL homologs in yeast: Evidence for three Mlh1-based heterocomplexes with distinct roles during meiosis in recombination and mismatch correction. *Proc. Natl. Acad. Sci. U. S. A.* 96: 13914-13919.

Warren, J.J., Pohlhaus, T.J., Changela, A., Iyer, R.R., Modrich, P.L. and Beese, L.S. 2007. Structure of the human MutSalpha DNA lesion recognition complex. *Mol. Cell.* 26: 579-592.

Watson, P. and Lynch, H.T. 2001. Cancer risk in mismatch repair gene mutation carriers. *Fam. Cancer.* 1: 57-60.

Watson, P. and Riley, B. 2005. The tumor spectrum in the lynch syndrome. *Fam. Cancer.* 4: 245-248.

Weinberg RA. 2007. The Biology of Cancer. Garland Science, Taylor & Francis Group, LLC: USA. 413-416.

Weterings, E. and Chen, D.J. 2008. The endless tale of non-homologous end-joining. *Cell Res.* 18: 114-124.

Woods, M.O., Williams, P., Careen, A., Edwards, L., Bartlett, S., McLaughlin, J.R. and Younghusband, H.B. 2007. A new variant database for mismatch repair genes associated with lynch syndrome. *Hum. Mutat.* 28: 669-673.

Yang, G., Scherer, S.J., Shell, S.S., Yang, K., Kim, M., Lipkin, M., Kucherlapati, R., Kolodner, R.D. and Edelmann, W. 2004. Dominant effects of an Msh6 missense mutation on DNA repair and cancer susceptibility. *Cancer. Cell.* 6: 139-150.

Yuen, S.T., Chan, T.L., Ho, J.W., Chan, A.S., Chung, L.P., Lam, P.W., Tse, C.W., Wyllie, A.H. and Leung, S.Y. 2002. Germline, somatic and epigenetic events underlying mismatch repair deficiency in colorectal and HNPCC-related cancers. *Oncogene.* 21: 7585-7592.

Zhang, J., Lindroos, A., Ollila, S., Russell, A., Marra, G., Mueller, H., Peltomäki, P., Plasilova, M. and Heinimann, K. 2006. Gene conversion is a frequent mechanism of inactivation of the wild-type allele in cancers from MLH1/MSH2 deletion carriers. *Cancer Res.* 66: 659-664.

Zhang, Y., Yuan, F., Presnell, S.R., Tian, K., Gao, Y., Tomkinson, A.E., Gu, L. and Li, G.M. 2005. Reconstitution of 5'-directed human mismatch repair in a purified system. *Cell.* 122: 693-705.