Functional Characterization of MutS Homologue Mismatch Repair Proteins and their Variants

Jukka Kantelinen

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

Academic Dissertation

To be presented for public examination with the permission of the Faculty of Biological and Environmental Sciences of the University of Helsinki in the Walter Hall in the EE-building, Agnes Sjöberg street 2, Helsinki, on the 4th of May 2012 at 12 o'clock noon.

Supervisors

Professor Minna Nyström, PhD
Division of Genetics
Department of Biosciences
Faculty of Biological and Environmental Sciences
University of Helsinki, Finland

Docent Reetta Kariola, PhD
Division of Genetics
Department of Biosciences
Faculty of Biological and Environmental Sciences
University of Helsinki, Finland

Reviewers

Docent Esa Kuismanen, PhD
Division of Biochemistry and Biotechnology
Department of Biosciences
Faculty of Biological and Environmental Sciences
University of Helsinki, Finland

Docent Minna Pöyhönen, MD, PhD
HUSLAB Department of Clinical Genetics
Helsinki University Central Hospital
Helsinki, Finland
and
Department of Medical Genetics
Haartman Institute
Faculty of Medicine
University of Helsinki, Finland

Opponent

Professor Hannes Lohi, PhD
Department of Veterinary Biosciences,
Faculty of Veterinary Medicine
and
Department of Medical Genetics
Faculty of Medicine
Haartman Institute
and
Research program for Molecular Medicine
Faculty of Medicine
and
Folkhälsan Research Center,
University of Helsinki, Finland

ISSN 1799-7372
ISBN 978-952-10-7922-1 (paperback)
ISBN 978-952-10-7923-8 (PDF, http://ethesis.helsinki.fi)

Cover layout by Anita Tienhaara; cover photo by Jukka Kantelinen
Helsinki 2012, Unigrafia Oy

“Palapelin kokosin paksuin rukkasin,
ihmetellen minne kaikki palat hukkasin”

- *Eppu Normaali*

CONTENTS

LIST OF ORIGINAL PUBLICATIONS **5**
ABBREVIATIONS **7**
SUMMARY **8**
INTRODUCTION **10**
REVIEW OF THE LITERATURE **11**
DNA replication and mismatch repair (MMR) **11**
DNA replication and replicative polymerases 11
Overview of DNA mismatch repair 12
Human MutS homologues 13
Other proteins and their functions in human MMR 14
In vitro model of MMR 15
Link between DNA replication and MMR 18
Lynch syndrome **18**
Clinical and tumor pathological characteristics of LS **19**
Family cancer history of a putative LS family 19
Microsatellite instability and loss of MMR protein in a tumor 20
MMR gene regulation changes 22
Variations of uncertain significance in MMR genes **22**
MMR gene variations 22
In silico predictions 23
Homozygous and compound heterozygous mutations in MMR genes 23
Functional analyses of MMR gene variants 25
Pathogenicity assessment of MMR variations 27
AIMS OF THE STUDY **29**
MATERIALS AND METHODS **30**
Study material **30**
VUS in studies I, II and III 30
Summary of the methods **33**
Cell lines and antibodies 33
Heteroduplexes used in the in vitro MMR assays 35
RESULTS AND DISCUSSION **36**
Mismatch repair analyses of VUS and their pairs (I, II and III) **36**
Clinical and tumor pathological data of the VUS carriers (II, III) 36
Concomitant effect of the VUS pairs to MMR deficiency (III) 36
Pathogenic MSH6 VUS (II, III) 40
The functional analysis of an MSH3 variation (I) 42
Substrate specificity and activity of MutS homologue wild type proteins and replicative polymerases in MMR (I, IV) **44**
A strong role of MutSβ in dinucleotide loop repair (I, IV) 44
α and ε polymerases have a minor role in mismatch repair in vitro (IV) 45
CONCLUSIONS AND FUTURE PROSPECTS **50**
ACKNOWLEDGEMENTS **52**
REFERENCES **53**

LIST OF ORIGINAL PUBLICATIONS

The present thesis is based on the following original articles, which will be referred to by the Roman numerals in the text (**I-IV**):

I **Kantelinen J***, Kansikas M*, Korhonen MK, Ollila S, Heinimann K, Kariola R, Nyström M. 2010. MutSβ exceeds MutSα in dinucleotide loop repair. Br J Cancer 102:1068-1073.

II **Kantelinen J**, Hansen TV, Kansikas M, Krogh LN, Korhonen MK, Ollila S, Nyström M, Gerdes AM, Kariola R. 2011. A putative Lynch syndrome family carrying *MSH2* and *MSH6* variants of uncertain significance -functional analysis reveals the pathogenic one. Fam Cancer 10:515-20.

III **Kantelinen J**, Kansikas M, Candelin S, Hampel H, Smith B, Kariola R, Nyström M. Mismatch repair deficiency caused by two inherited MSH2 and/or MSH6 variations found in a cancer patient. Submitted.

IV Itkonen HM, **Kantelinen J**, Vaara M, Parkkinen S, Hemmerich P, Schlott B, Grosse F, Nyström M, Syväoja JE, Pospiech H. Physical and functional interactions between human replicative DNA polymerases and MSH proteins. Submitted.

* Equal contribution

Statement of my contribution on published articles:

I I participated in experimental planning and was responsible for laboratory experiments together with Minttu Kansikas. Laboratory work included: site directed mutagenesis, substrate preparation, cell culture and nuclear extract preparation, western blotting, and *in vitro* MMR testing. In addition, I produced the wild type and variant total extracts and performed the variant testing. I also participated in writing the manuscript.

II I participated in planning the experiments, writing the manuscript, and was responsible for the laboratory work including mutagenesis, production of variant protein, and functional analyses.

III I participated in experimental planning and was responsible for laboratory experiments. Laboratory work included: site directed mutagenesis, the substrate productions, cell culture and nuclear extract productions, western blotting and *in vitro* MMR testing. I carried out all of the phases of the laboratory work. I also participated in writing the manuscript.

IV I participated in experimental planning and laboratory analyses of the neutralization and depletion experiments and the *in vitro* MMR assay. I also participated in writing the manuscript.

ABBREVIATIONS

ACI/II	Amsterdam criteria I/II
cDNA	Complementary DNA
ChIP	Chromatin immunoprecipitation
CRA	Colorectal adenoma
CRC	Colorectal cancer
dsDNA	Double stranded DNA
E. coli	*Escherichia coli*
EC	Endometrial cancer
EPCAM	Epithelial cell adhesion molecule
EXO1	Exonuclease 1
GGc	Glycerol gradient centrifugation
HA	Hydroxylapatite
HMGB1	High mobility protein group B1
HNPCC	Hereditary nonpolysis colorectal cancer
IDL	Insertion-deletion loop
IHC	Immunohistochemistry
InSiGHT	International Society for Gastrointestinal Hereditary Tumors
LS	Lynch Syndrome
MAPP	Multivariate analysis of protein polymorphism
MLH	MutL homolog
MLPA	Multiplex ligation dependent probe amplification
MMR	Mismatch repair
MSH	MutS homolog
MSI	Microsatellite instability
MSI-H	Microsatellite high
MSS	Microsatellite stable
NE	Nuclear extract
nt	Nucleotide
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PMS	Post meiotic segregation
Pol α	Polymerase alpha
Pol δ	Polymerase delta
Pol ε	Polymerase epsilon
PolyPhen	Polymorphism phenotyping
RFC	Replication factor C
RPA	Replication protein A
Sf9	*Spodoptera frugiperda 9*
SIFT	Sorting intolerant from tolerant program
ssDNA	Single strand DNA
TE	Total extract
TGFβ2	Transforming growth factor β receptor II
VUS	Variant of uncertain signifigance
WT	Wild type

SUMMARY

Lynch syndrome (LS) is one of the most common hereditary cancer syndromes and may lead to cancer development, mainly in colon or in endometrium, for 20 years earlier than in general population. LS is an autosomal dominantly inherited disorder, associated with the malfunction of a highly conserved postreplicative DNA mismatch repair (MMR) mechanism and germline mutations at least in four different MMR genes, *MLH1*, *MSH2*, *MSH6*, and *PMS2*. The MMR genes *MSH3* and *MLH3* have also been linked to LS but their roles are less clear. To be able to offer an appropriate follow-up and genetic counseling to LS families and their mutation carriers, they must be diagnosed, which usually starts by studying the cancer history in the family and the tumor phenotype of the index patient followed by mutation search and pathogenicity assessment of found variations. The amino acid substitutions, deletions, and insertions, which change only one amino acid in a protein structure, do not necessarily destroy protein and therefore their pathogenicity is difficult to interpret. Furthermore, in rare cases, an individual can carry two variations either in the same or different MMR genes, which further complicates the pathogenicity assessment.

DNA MMR corrects mismatches arising mainly during DNA replication. DNA synthesis in each cell division is carried out by three major replicative DNA polymerases (Pols) α, δ and ε and their incomplete proofreading activity together with malfunction in MMR leads to the accumulation of mismatches in the genome leading to genomic instability and cancer.

MutS homologue (MSH) MMR proteins form the DNA mismatch recognizing factors MutSα (MSH2/MSH6) and MutSβ (MSH2/MSH3). One aim in the present study was to analyze the substrate efficiencies of MutSα and MutSβ by using the functional *in vitro* MMR assay with different substrates and cell lines. The target substrates of MutSα and MutSβ have already been widely studied. However, the extent of their functional redundancy and clinical substance remains unclear. Here, our results show that although MutSα alone seems to be responsible for the mismatch and one nucleotide loop repair, MutSα and MutSβ have functional redundancy in two nucleotide loop repair and MutSβ even seems to exceed MutSα in that. The finding is clinically relevant since such a strong role in two nucleotide loop repair indicates *MSH3* deficiency in tumors with low dinucleotide and no mononucleotide repeat instability.

The second aim in the study was to functionally characterize a possible compound effect of 9 pairs of variants of unknown significance (VUS) found in cancer patients. Four variant pairs were shown to be proficient while one VUS, *MSH2* c.380A>G was individually assessed proficient but in a pair with another VUS deficient. Thus, our results suggest that two inherited MMR gene variations in a cancer patient may have a concomitant contribution to MMR deficiency. Moreover, the role of this frequently reported MMR gene VUS *MSH2* c.380A>G is especially interesting, since its concomitant defect with another variant could finally explain its recurrent occurrence in colorectal cancer patients. Three *MSH6* VUS were

shown to cause MMR deficiency individually. Furthermore, one separately studied *MSH3* variation was shown to be proficient in MMR.

The third aim was to study the role of replicative polymerases α and ε in MMR. Here, we demonstrate a proliferating cell nuclear antigen independent interaction between replicative DNA polymerases and MSH proteins MSH2 and MSH6 by co-purification as well as by conventional and chromatin immunoprecipitation. Chromatin recruitment but not the release of MSH2 appears to depend on DNA replication. The novel interaction provides a potential mechanism for replication-dependent strand discrimination during MMR. In addition, we showed that polymerases of the replication fork have a functional role in human MMR. Our data, suggesting that MSH2 and MSH6 physically interact with Pols δ and Pol α, are in accordance with models where MSH proteins are continuously loaded onto chromatin in a replication-dependent manner and persist on DNA that has already completed replication.

INTRODUCTION

Mismatch repair mechanism (MMR) is responsible for maintaining genomic stability by repairing errors produced mainly by DNA polymerase that form during DNA replication (Jiricny 2006a). Cells lacking functional MMR are not able to repair DNA mismatches which leads to accumulation of mutations in DNA. Mutations can be in important genes regions, such as tumor suppressor genes (Umar *et al.* 2004). Inactivation of these genes is associated with tumor development and progression. Mutations of MMR genes, especially *MHL1*, *MSH2,* and *MSH6* are linked to one of the most prevalent dominantly inherited cancer syndromes, known as Lynch Syndrome (LS) (Peltomäki *et al.* 1993; Peltomäki 2005), while no predisposing mutations have so far been found in the *MSH3*. Typical LS mutation carriers develop cancer in the colon or endometrium in middle age, on average 20 years earlier than than individuals without inherited susceptibility. International Amsterdam criteria (Vasen *et al.* 1991; Vasen *et al.* 1999) have been developed to facilitate and unify LS diagnosis in all over the world enabling indispensable counselling, follow-up and treatments for LS mutation carriers.

MMR proteins form functional heterodimer complexes, such as MutSα (MSH2 and MSH6) and MutSβ (MSH2 and MSH3), which recognize small DNA errors in newly replicated DNA (Jiricny 2006a). Most inherited mutations in MMR genes destroy protein structure and their pathogenicity is easy to assess. Evaluation of the effect of variations in protein/heterodimer structure can be tricky in cases where only one aminoacid is changed in the protein structure, however, a small group of cancer patients have inherited more than one variation either in the same or different MMR genes. The evaluation of the concomitant effect of such small changes is challenging. Several functional assays have been developed to study the molecular basis of the MMR mechanism (Ou *et al.* 2007), and to facilitate the evaluation of small individual and concomitant changes.

The main function of polymerases is to duplicate DNA when cells divide. This process requires several DNA polymerases such as α, δ and ε, which synthesize DNA with high fidelity. Fidelity and accuracy of DNA duplication cannot be achieved without functional MMR. Thus these two mechanisms are interacting with each others. DNA polymerase δ is proposed to take part in the MMR mechanism by replacing incorrect DNA (Longley *et al.* 1997). Other polymerases α and ε, however, are not yet known to take part in MMR.

In this PhD work MutSα and MutSβ substrate specificities and efficiencies are investigated and the concomitant effects of 9 different *MSH2/MSH2*, *MSH6/MSH6* or *MSH2/MSH6* mutation pairs identified in cancer patients are functionally characterized. In addition, one novel *MSH3* variation found in a putative LS patient was analysed. Finally it was studied, whether and how replicative DNA polymerases α, δ, and ε are interacting with the MMR mechanism.

REVIEW OF THE LITERATURE

DNA replication and mismatch repair (MMR)

DNA replication and replicative polymerases

Precise maintenance of genetic information is essential for every organism. High fidelity DNA replication is carried out by the replicative DNA polymerases α, δ and ε (DNA pols α, δ and ε), however low fidelity in DNA replication leads to the accumulation of mutations, which may predispose humans to diseases such as cancer (Jiricny 2006a; McCulloch and Kunkel 2008).

Family B DNA polymerases α, δ and ε are responsible for DNA synthesis in proliferating cells (Figure 1). DNA Pol ε is believed to synthesize the leading strand DNA in a largely continuous fashion (Fukui *et al.* 2004; Pursell *et al.* 2007), while the Okazaki fragments on the lagging strand are synthesized and maturated predominantly by DNA Pol δ (Jin *et al.* 2001; Jin *et al.* 2003; Fukui *et al.* 2004; Garg and Burgers 2005; Nick McElhinny *et al.* 2007). DNA Pol α is the eukaryotic protein which carries DNA primase activity and is responsible for primer synthesis both on the leading strand and on each Okazaki fragment of the lagging strand (Lehman and Kaguni 1989). It synthesizes the first 10 nucleotides of each approximately 250-nucleotide long Okazaki fragments on the lagging strand, constituting about 2 % of human DNA (McCulloch and Kunkel 2008).

Several *in vitro* studies indicate that proofreading improves replication fidelity even more than 100 fold depending on the mismatch, the sequence context and the polymerase. DNA Pols δ and ε synthesize the bulk of the human genome with high fidelity and processivity (Kunkel and Bebenek 2000; Hubscher *et al.* 2002; Beard and Wilson 2003; Garg and Burgers 2005; Hsieh and Yamane 2008; McCulloch and Kunkel 2008). DNA Pol α possesses poor processivity and synthesizes DNA with lower fidelity compared to DNA Pols δ and ε (Garg and Burgers 2005). DNA Pol α lacks proofreading activity and has an error rate of 10^{-4} to 10^{-5} (Hsieh and Yamane 2008; McCulloch and Kunkel 2008). Pol δ is suggested to be responsible for pol α error proofreading (McCulloch and Kunkel 2008), by removing RNA-DNA primers synthesized by DNA Pol α mostly during the maturation process (Pavlov *et al.* 2006).

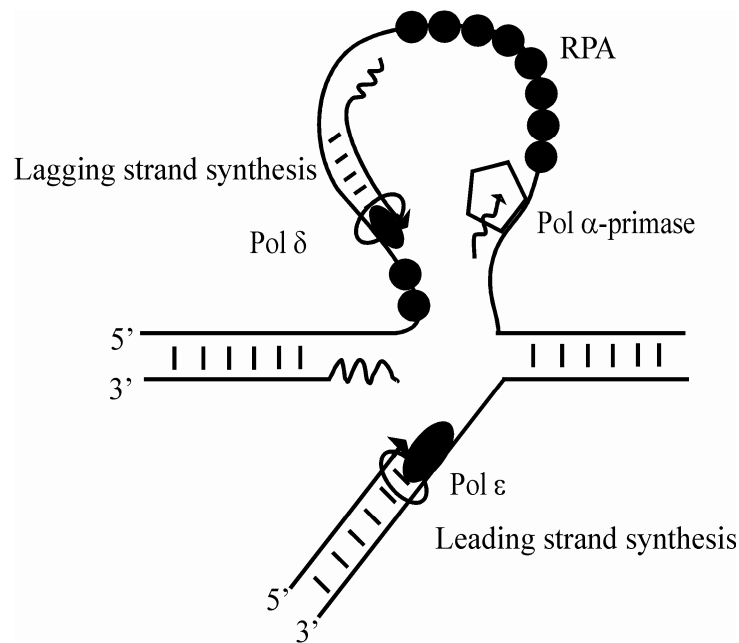


Figure 1. Replicative polymerases, replication protein A (RPA) and their functions in DNA replication. Primary function of Pol ε is the leading strand synthesis (Fukui *et al.* 2004; Pursell *et al.* 2007). Pol δ synthesizes and maturates Okazaki fragments on the lagging strand. Primer synthesis is carried out by DNA Pol α (Lehman and Kaguni 1989). Figure is modified from McCulloch *et al.* 2008.

Overview of DNA mismatch repair

The mismatch repair mechanism is best described in bacteria *Escherichia coli* and in yeast *Saccharomyces cerevisiae*. The first indications of prokaryotic MMR mechanism were reported in the early 1980s (Lu *et al.* 1983; Lu *et al.* 2005), and a few years later the prokaryotic MMR was reconstituted by using purified proteins (Lahue *et al.* 1989). A very similar repair process was later found in eukaryotic cells (Holmes *et al.* 1990). Intensive research during the last thirty years has demonstrated that the basics of the MMR mechanism are very similar in prokaryotes and eukaryotes, indicating similar main protein functions. The main difference is that in the eukaryotic MMR mechanism, MutS and MutL homologue proteins form functional heterodimer complexes with each others, while in prokaryotes the functional complexes are homodimers (Li 2008).

The MMR system increases DNA replication fidelity by repairing postreplicative DNA errors, such as single nucleotide mismatches and small insertion-deletion loops (IDLs) (Palombo *et al.* 1996; Jiricny 2006a; Pavlov *et al.* 2006). Postreplicative errors are suggested to arise due to replication slippage (Ellegren 2004). The highly conserved DNA repair mechanism protects DNA from errors, which arise in approximately every 10^6-10^7 new bases (Kolodner and Marsischky 1999). MMR defects may lead to a 1000 fold decrease in replication fidelity in mammals (Hsieh and Yamane 2008). Studies on the MMR partial

reaction have demonstrated that the DNA error correction system can be divided into three main phases: error detection, strand excision, and synthesis of a new strand.

Human MutS homologues

Five MutS homologues (MSH), *MSH2*, *MSH3*, *MSH4*, *MSH5,* and *MSH6,* are recognized in human cells. Three MSH proteins form two different heterodimers MutSα (MSH2 and MSH6), and MutSβ (MSH2 and MSH3), which participate in MMR. The third heterodimer consists of MSH4 and MSH5 and is not known to have a function in MMR (Drummond *et al.* 1995; Palombo *et al.* 1995; Acharya *et al.* 1996; Bocker *et al.* 1999). The MutSα protein complex recognizes base-base mismatches and small IDLs of 1-5 unpaired nucleotides, whereas MutSβ mainly recognizes two or more unpaired nucleotides in MMR (Umar *et al.* 1994; Drummond *et al.* 1995; Palombo *et al.* 1995; Acharya *et al.* 1996; Drummond *et al.* 1997; Gradia *et al.* 1997; Genschel *et al.* 1998; Zhang *et al.* 2005). MutSα and MutSβ possess partially overlapping functions. MutSβ has a high binding affinity to 2 nt (nucleotide) or more IDLs but in contrast, a very low affinity to simple base/base mispairs (Palombo *et al.* 1995; Acharya *et al.* 1996; Palombo *et al.* 1996), while MutSα has been shown to bind and repair both base/base mispairs and IDLs (Drummond *et al.* 1995; Palombo *et al.* 1996; Genschel *et al.* 1998).

Both MSH2 and MSH6 proteins can be divided into five functional domains: 1) mismatch binding domain (amino acids 1-124 in MSH2 and 362-518 in MSH6), 2) connector domain (amino acids 125-297 in MSH2 and 519-717 in MSH6), 3) lever domains (amino acids 300-456 and 554-619 in MSH2 and 718-934 and 1009-1075 in MSH6) , 4) clamp domains (457-553 in MSH2 and 935-1008 in MSH6), and 5) ATPase domain (620-855 in MSH2 and 1076-1355 in MSH6) (Warren *et al.* 2007). The MSH proteins are ATPases that possess the Walker ATP-binding motif, which contains the highly conserved polypeptide sequence (Jiricny 2006a). The lesion specificity is believed to lie within the *MSH3/MSH6*-specific sequences, which differ notably between them (Owen *et al.* 2009). The process through which ADP-ATP exchange occurs on MSH2 seems to be dependent on the protein it forms a complex with: MSH6 requires ATP stabilization, whereas MSH3 requires ATP hydrolysis, both of which are dependent on specific lesion binding (Owen *et al.* 2009).

Importantly, it has been shown that MSH6 (domain 1) is responsible for specific mismatch binding in the MutSα complex, while MSH2 makes contact with DNA in an unspecific manner (Dufner *et al.* 2000; Warren *et al.* 2007). It has also been proposed, however, that the mismatch binding site of MSH2 is involved in MutSβ mediated MMR and that the DNA-binding mode of MutSβ varies depending on the loop size (Lee *et al.* 2007; Dowen *et al.* 2010; Tseng *et al.* 2011). The mismatch binding domain in MSH6 includes a conserved Phe-*X*-Glu motif which is responsible for specific DNA mispair interaction both in prokaryotic MutS complex and in eukaryotic MutSα (Lamers *et al.* 2000; Obmolova *et al.* 2000; Warren *et al.* 2007). However, MutSβ does not repair base-base mispairs or 1 nucleotide IDLs, even

if the Phe-*X*-Glu motif can be found in MSH3 (Genschel *et al.* 1998; Tseng *et al.* 2011). Instead MutSβ may participate in the processing of the trinucleotide repeat expansions, since MutSβ has been shown to display identical biochemical and biophysical activity when interacting with a $(CAG)_n$ hairpin and a mismatch (Tian *et al.* 2009).

Other proteins and their functions in human MMR

The MMR mechanism requires a set of different proteins to accomplish the entire repair process. MutLα is a heterodimer complex composed of MutL homologue (MLH) proteins MLH1 and PMS2. The primary function of MutLα is to interact with the MutSα complex and increase the detection sensitivity to mispairs and to act as a mediator with the other components needed in MMR (Jiricny and Nyström-Lahti 2000). MutSα and MutLα together form the so called tertiary complex, which increases heteroduplex specificity (Constantin *et al.* 2005; Zhang *et al.* 2005). It is suggested that MutLα carries the endonuclease activity and therefore is capable of cut starting points to exonucleases (Jiricny 2006b). The endonuclease property of MutLα is suggested to be activated by replication factor C (RFC) and proliferating cell nuclear antigen (PCNA) (Kadyrov *et al.* 2006). Another heterodimer complex, MutLγ, consists of MLH1 and MLH3 proteins. It seems to only take part in MMR if MutLα is not available, and even then, it only partially replaces its error-correction function (Cannavo *et al.* 2005; Korhonen *et al.* 2008).

Proliferating cell nuclear antigen is a sliding clamp, which participates in DNA replication and repair, and interacts with MSH2, MSH3, MSH6 and MLH1 (Umar *et al.* 1996; Gu *et al.* 1998; Clark *et al.* 2000; Bowers *et al.* 2001; Kleczkowska *et al.* 2001; Essers *et al.* 2005). The main function of PCNA in MMR is the initiation and DNA re-synthesis. In addition, PCNA has two roles affecting MutLα function; first, the clamp is required for endonuclease activation and second, PCNA determines the strand direction of MutLα incision (Pluciennik *et al.* 2010). PCNA also assists MutSβ and MutSα proteins in mispair localisation during new strand synthesis (Lau and Kolodner 2003) and mediates MutSβ and MutLα interaction (Iyer *et al.* 2010). PCNA seems to be required for the repair reaction which proceeds to the 3' nick directed but not to the 5' nick directed (Guo *et al.* 2004). The primary function of RFC is in MutLα activation and to load PCNA on DNA (Kolodner and Marsischky 1999; Tainer *et al.* 2010).

Replication protein A (RPA) has multiple roles in MMR. RPA binds to nicked heteroduplex DNA to protect ssDNA from binding to itself during excision and it facilitates the excision (Ramilo *et al.* 2002; Dzantiev *et al.* 2004; Zhang *et al.* 2005). Phosphorylation of RPA decreases its DNA binding affinity, however the phosphorylation level of RPA varies during different stages of the repair process. RPA is not phosphorylated during excision, which increases its binding affinity to DNA. Conversely, while Pol δ creates a new strand, RPA is phosphorylated, stimulating re-synthesis (Guo *et al.* 2006).

The function of an exonuclease 1 (EXO1) has also been shown to be critical in MMR since its inactivation increases misincorporation frequency to 10^{-3} to 10^{-6} depending on the

sequence context. Although, the reconstituted MMR system requires only EXO1 (Constantin *et al.* 2005; Zhang *et al.* 2005), a weak mutator phenotype in EXO1 null mice suggests other yet unknown exonuclease activities (Amin *et al.* 2001; Wei *et al.* 2003) . The main function of exonuclease 1 in MMR is DNA and mismatch excision (Jiricny 2006a) and its interaction partners are MLH1 and MSH2 (Tishkoff *et al.* 1997; Schmutte *et al.* 1998; Tishkoff *et al.* 1998; Amin *et al.* 2001; Nielsen *et al.* 2004; Tran *et al.* 2004).

In addition to the described proteins above, MMR may need some other protein activities, for example high mobility protein group B1 (HMGB1), but their clear role in MMR has not yet been shown. HMGB1 interacts with MSH2 and MSH6 *in vitro* (Yuan *et al.* 2004) and is likely to function as an excision stimulator in MMR. However, HMGB1 may be unnecessary when RPA is present (Zhang *et al.* 2005). Finally, DNA ligase I fills the remaining nicks during MMR (Zhang *et al.* 2005).

In vitro model of MMR

A reconstituted *in vitro* MMR reaction has been accomplished in two separate studies by using purified human proteins (Figure 2, Page 17) (Constantin *et al.* 2005; Zhang *et al.* 2005). These studies were performed using slightly different combinations of purified proteins, which repaired the substrates in a test tube. Bidirectional reactions have also been examined by using substrates, which include a single strand nick either in the 5' or 3'direction from the mismatch (Constantin *et al.* 2005). The first of these studies reported that *in vitro* MMR requires the activity of a total of 7 components: MutSα, MutLα, RPA, EXO1, PCNA, RFC and Pol δ (Constantin *et al.* 2005). Then it was reported that the repair of a substrate with a 5' nick would not require MutLα activity, while substrate with a 3' nick does (Constantin *et al.* 2005; Zhang *et al.* 2005; Jiricny 2006a). The second reconstitution study, which mainly concentrated on the IDL repair mechanism (Zhang *et al.* 2005), demonstrated otherwise similar results to the study of Constantin *et al*. (2005), but suggested that HMGB1 activity was also needed.

The results of *in vitro* mismatch repair studies have been combined to a model that describes MMR in detail (Figure 2, Page 17) (Jiricny 2006a). According to this model, the mismatch repair process is initiated by the binding of the mismatch recognition factor MutSα or MutSβ to the mispair, followed by assembly of the repairosome by MutLα (Constantin *et al.* 2005; Zhang *et al.* 2005). In *Escherichia coli* (*E. coli*), strand discrimination is methyl directed (Lu *et al.* 1983; Lu *et al.* 2005), whilst in the *in vitro* model for eukaryotic MMR, the strand discrimination is directed by a pre-existing nick, since it is not yet known how strand discrimination is determined in eukaryotic cells (Gradia *et al.* 1997; Gradia *et al.* 1999). Three different models for MMR in eukaryotic cells have been proposed: 1) a sliding clamp model (Gradia *et al.* 1999) or 2) a translocation model (Allen *et al.* 1997), and 3) a stationary model (Junop *et al.* 2001). The most recent evidence supports the sliding clamp model (Zhang *et al.* 2005), where the MSH protein heterodimer plays a key role as a protein

complex that encircles DNA and allows it to pass freely through the hole in its centre. The sliding clamp is activated when it recognizes and binds to a mismatch or IDL, which changes its conformation. In the presence of mismatched DNA, MSH heterodimer is wrapped around DNA, followed by ADP/ATP exchange, which changes its conformation and leads to its release from the mismatch site while leaving the heterodimer closed (Jiricny and Nyström-Lahti 2000).

ATP commitment to the tetramer complex causes a conformational change, after which it slides upstream in the 5' direction, reaching first RFC which detaches DNA and then EXO1, which binds to the complex. Using purified proteins from human cell extracts it was discovered that a nick in the 3' or 5' direction from the error directs excision (Genschel and Modrich 2003; Dzantiev *et al.* 2004). According to the simplest model for *in vitro* 5' → 3' strand excision, only MutSα, EXO1, and RPA proteins were needed. The study suggests that an activation of EXO1 does not require MutLα, although MutLα increases error dependence (Constantin *et al.* 2005). This is, however, inconsistent with requirements in an *in vitro* MMR assay, where a substrate with a 5' nick also requires MutLα or MutLγ activity (Nyström-Lahti *et al.* 2002; Raevaara *et al.* 2005; Korhonen *et al.* 2008). When the *in vitro* MMR reaction includes MutLα, MutSα, EXO1, RPA, PCNA, and RFC proteins, excision is carried out regardless of direction of the nick. Furthermore, it was shown more recently that MutLα-complex may have an endonuclease activity enabling the starting point for exonucleases (Jiricny 2006b; Kadyrov *et al.* 2006; Kadyrov *et al.* 2007), while excision is carried out by EXO1. Repair in the 3' direction does not seem to occur in the same way as into the 5' direction, however, here, the RFC protein does not release from DNA during the process, but by binding to the 3' side of the nick, it blocks the EXO1 protein preventing its movement in the wrong direction (Kolodner and Marsischky 1999; Jiricny 2006a). In addition, RFC may also activate PCNA during MMR (Constantin *et al.* 2005).

MutLα also plays a crucial role in termination of the excision, since without it, the excision process would continue beyond the mismatch. The function of RPA is to stabilize the single-stranded gap during the excision and resynthesis. Resynthesis is carried out by pol δ (Longley *et al.* 1997), which is connected to PCNA (Kolodner and Marsischky 1999).

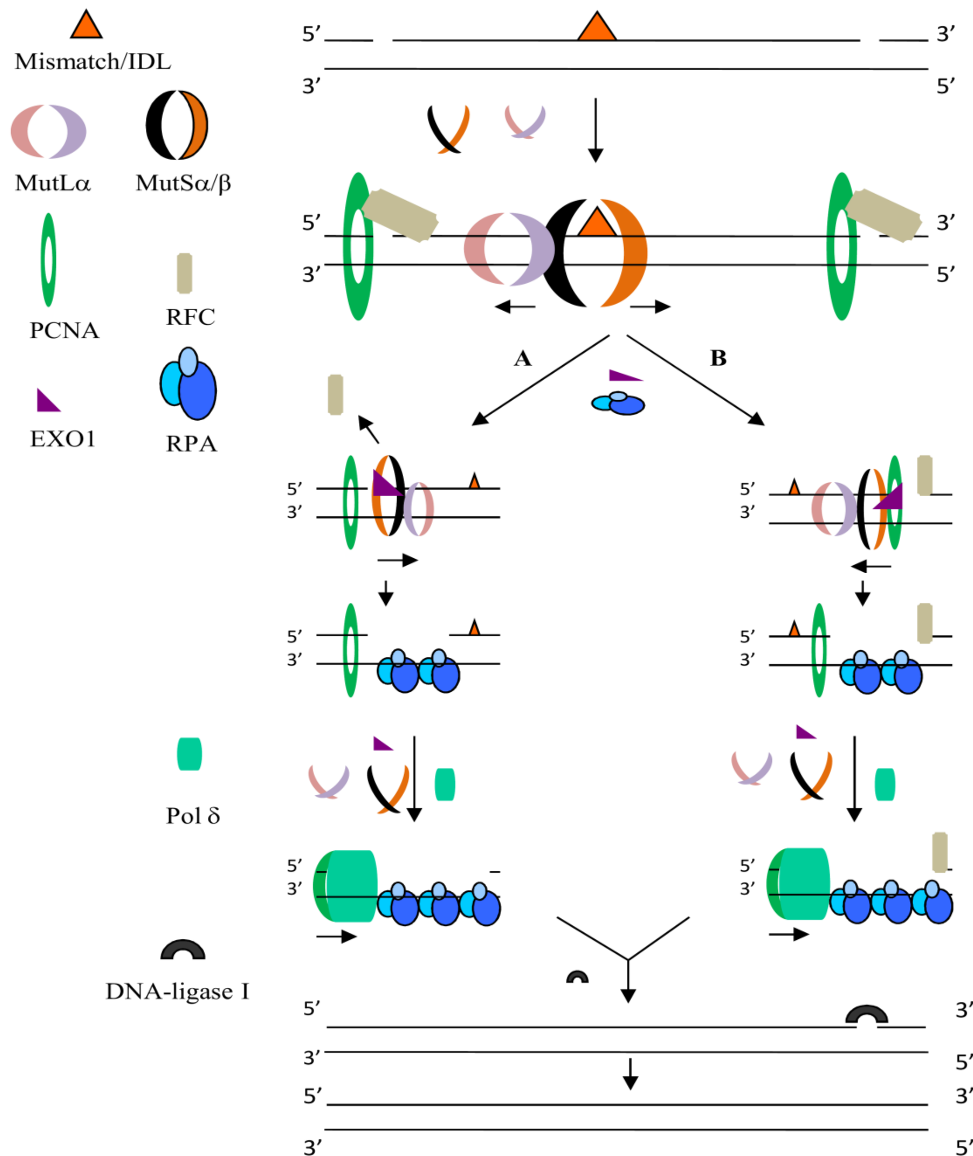


Figure 2. Reconstituted *in vitro* MMR mechanism using purified proteins (Constantin *et al.* 2005; Zhang *et al.* 2005; Jiricny 2006a). Abbreviations: insertion-deletion loop (IDL), heterodimer complex of MSH2 and MSH6 (MutSα), heterodimer complex of MSH2 and MSH3 (MutSβ), heterodimer complex of MLH1 and PMS2 (MutLα), proliferating cell nuclear antigen (PCNA), replication factor C (RFC), exonuclease 1 (EXO1), replication protein A (RPA), polymerase delta (pol δ) and DNA-ligase I. MutSα or MutSβ recognises and binds to an error (IDL or mismatch), MutLα binds to MutSα/β, followed by ATP binding. A conformational change in the tetramer releases it from the error, after which it starts to slide along the DNA molecule, reaching a single strand brake (nick). A and B describe the repair processes that have the break in the 5' (A) or 3' (B) direction from the error, respectively. **A)** The tetramer complex slides to upstream from the error and releases RFC molecule, which is bound to the 5' nick. Tetramer releases RFC and binds to EXO1, which actively removes the incorrect strand. After excision, PCNA and Pol δ synthesize a new strand and DNA ligase I seals the nick. **B)** The tetramer complex slides to downstream of the mismatch and achieve the PCNA protein, which is committed to the 3' direction, EXO1 binds the complex and its activation leads to strand degradation between the nick and the error. RFC prevents degradation to proceed to the wrong direction. Finally, pol δ/PCNA fills the gap and DNA ligase I seals the nick. Figure is modified from Jiricny 2006a.

Link between DNA replication and MMR

The accuracy of DNA replication in the absence of a functional MMR is in the range of 10^{-7} to 10^{-8} base misincorporations per genome (Kunkel and Bebenek 2000; Hsieh and Yamane 2008; McCulloch and Kunkel 2008). This high fidelity is achieved mainly by a strict nucleotide selection at the base incorporation step and by a proofreading exonuclease activity of DNA Pols δ and ε (Kunkel and Bebenek 2000; Beard and Wilson 2003; Hsieh and Yamane 2008). Although the main function of polymerases is to replicate DNA strands during a cell division, another important function is to take part in DNA repair mechanisms. Pol delta (Pol δ) is supposed to be a key player in MMR, where the task is to resynthesize the excised strand (Longley *et al.* 1997).

MMR activity is the highest during the S phase of the cell cycle (Schröering *et al.* 2007) but it is unclear how MMR proteins are recruited to newly replicated DNA. It has been proposed that MMR proteins are linked to the replication machinery especially via PCNA, so that MutSα is physically attached to it (Kleczkowska *et al.* 2001). On the other hand, Schröering et al. (2007) have shown that MSH proteins are recruited to chromatin after inhibition of replication, which led them to suggest that MMR proteins might be recruited to chromatin directly by replicative DNA polymerases, instead of other replication fork associated factors. However, no rigid evidence was provided for that suggestion. A mechanistic link between the replication forks and MSH complex might provide information on unfaithfully replicated DNA strands. The nicks are supposed to occur frequently during the replication of Okazaki fragments. However, the leading strand is synthesized in a more continuous manner (Fukui *et al.* 2004; Chilkova *et al.* 2007; Pursell *et al.* 2007). DNA Pol δ has been shown to catalyze the DNA synthesis step in human nuclear extracts (Longley *et al.* 1997). The role of other replicative DNA polymerases has not been addressed, although genetic studies on *Saccharomyces cerevisiae* suggest that exonuclease activities of both Pol δ and Pol ε participate in the MMR excision process (Harfe and Jinks-Robertson 2000; Schofield and Hsieh 2003).

Lynch syndrome

A large advancement was made in the understanding of the molecular and genetic background of Lynch syndrome (LS) (hereditary non-polyposis colorectal cancer syndrome, HNPCC; MIM # 120435) in 1993, when the first susceptibility gene *MSH2* was mapped to the chromosome 2 by a genome-wide search and linkage analysis (Leach *et al.* 1993; Peltomäki *et al.* 1993). This was closely followed by the mapping of the second susceptibility gene, *MLH1*, to chromosome 3 (Lindblom *et al.* 1993; Papadopoulos *et al.* 1994). Nowadays it is known that germline mutations at least in four different MMR genes, *MLH1*, *MSH2*, *MSH6*, and *PMS2,* predispose to LS. *MLH1* and *MSH2* are the most commonly mutated genes in LS and 70-85 % of all reported mutations are found in these genes. The *MSH6* mutations account for about 10 % of LS cases (Peltomäki and Vasen 2004), while only few predisposing mutations are reported in *PMS2* (www.insight-group.org). Inherited mutations

have also been reported in the MutL homologue, *MLH3* gene but their predisposition to LS is not confirmed (Wu *et al.* 2001). In contrast, no mutations have so far been found in the MutS homologue gene, *MSH3*, even if the dominant mutator effect (accumulation of the mutations in the genome), is connected to MSH3 deficiency (Risinger *et al.* 1996; de la Chapelle 2004).

Although, cancer is not generally inherited, it is always a genetic disease. In Finland, 1600 patients are diagnosed with colon cancer each year, accounting for 5-6 % of all cancer cases (Association of the Nordic Cancer Registries, http://www-dep.iarc.fr/NORDCAN/FI/frame.asp). Of the 1600 cancer patiens, 5–8 % are expected to have LS (Peltomäki *et al.* 2001; Peltomäki 2005), one of the most common cancer syndromes. Lynch syndrome shows high penetrance and an incidence of 1:1000 in the general population (Umar *et al.* 2004). Each year, the number of new colorectal cancer (CRC) cases in the whole world is over 1 million, and about 3 % of them belong to LS families; in other words, one for every 35 patients (1/35) who suffer CRC also have LS (Hampel *et al.* 2008; Lynch *et al.* 2009).

Lynch syndrome is an autosomal dominantly inherited disorder, associated with the malfunction of a highly conserved postreplicative DNA MMR mechanism (de la Chapelle 2004). MMR genes behave like tumor suppressor genes; mutations in both alleles are needed to inactivate the gene and start tumorigenesis (Knudson 1971). Thus in LS, only cancer susceptibility are inherited and LS mutation carriers need a somatic loss of the wild type MMR gene allele to result in defective MMR and progression of tumorigenesis. The main features that distinguish LS from sporadic CRC are an earlier average age of cancer onset (45 years vs. 65 years) and microsatellite instability (MSI) in tumors (Lynch *et al.* 2009).

Clinical and tumor pathological characteristics of LS

Family cancer history of a putative LS family

To be able to offer an appropriate follow-up and genetic counselling to LS families and their mutation carriers, they must be diagnosed. To facilitate LS diagnostics, the first international criteria were already proposed twenty years ago, known as the Amsterdam criteria I (ACI) (Vasen *et al.* 1991), which then were later revised and modified producing the Amsterdam criteria II (ACII) (Vasen *et al.* 1999) and Bethesda guidelines (Table 1) (Umar *et al.* 2004). The first criteria, ACI, which included only colon cancers, were found to be too strict and, therefore, later revised to include also some extracolonic cancers frequent in LS patients (ACII) (Vasen *et al.* 1999). The Amsterdam criteria are based on typical LS characteristics such as cancers present in successive generations indicating dominant inheritance, low age of cancer onset compared to the general population, quite often simultaneous/consecutive tumors in a patient, typical tumor spectrum including colorectal and endometrial tumors, and an MSI phenotype in a tumor.

LS patients carrying *MLH1* or *MSH2* mutations often display typical clinical and tumor pathological features of the syndrome, (Lynch *et al.* 2008). Contrary to typical LS characteristics, mutations in the *MSH6* gene are often associated with a later age of cancer onset and low or no MSI in tumors (Wu *et al.* 1999; Wagner *et al.* 2001; Berends *et al.* 2002; Hendriks *et al.* 2004)

Table 1. International criteria used in LS diagnostics

Amsterdam criteria I	**Amsterdam criteria II**
At least 3 relatives with histologically verified colorectal cancer: 1. One is a first-degree relative of the other two; 2. At least two successive generations affected; 3. At least one of the relatives with colorectal cancer diagnosed at <50 years. of age; 4. Familial adenomatous polyposis has been excluded.	At least 3 relatives with an hereditary nonpolyposis colorectal cancer-associated cancer (colorectal cancer, endometrial, stomach, ovary, ureter/renal pelvis, brain, small bowel, hepatobiliary tract, and skin [sebaceous tumors]): 1. One is a first-degree relative of the other two; 2. At least two successive generations affected; 3. At least one of the hereditary nonpolyposis colorectal cancer-associated cancers should be diagnosed at <50 years. of age; 4. Familial adenomatous polyposis should be excluded in any colorectal cancer cases; Tumors should be verified whenever possible.
Bethesda guidelines	
1. Colorectal cancer diagnosed in a patient who is less than 50 years of age. 2. Presence of synchronous or metachronous colorectal, or other HNPCC-associated tumors, a regardless of age. 3. Colorectal cancer with the MSI-High histology diagnosed in a patient who is less than 60 years of age. 4. Colorectal cancer or HNPCC-associated tumor diagnosed under age 50 years in at least one first-degree relative. 5. Colorectal cancer or HNPCC-associated tumor diagnosed at any age in two first- or second-degree relatives. Criteria 4 and 5 have been reworded to clarify the Revised Bethesda Guidelines.	

Modified from Lynch *et al.* 2008

Microsatellite instability and loss of MMR protein in a tumor

Microsatellites are short repetitive regions in the genome. Approximately 15 % of the human genome is repetitive DNA including interspersed repeats, LINE (long interspersed elements) and SINE (short interspersed elements) elements, and tandemly repeated elements such as satellite, microsatellite, and minisatellite elements. Microsatellites or short tandem repeats (STR'S) account for 3 % of the genome. The microsatellite repetitive unit length is 1-6 base pairs and whole satellite length is up to 100 kb. It is estimated that human DNA contains approximately 500 000 microsatellites locating mostly in introns, but also in promoter areas, exons, and untranslated terminal regions. Dinucleotide is the most common form of a repeat unit, followed by mono- and tetranucleotide repeats. Trinucleotide repeats are less frequently

seen. The most frequent dinucleotide repeat unit is $(CA)_n$. Repetitive DNA is prone to mutations and especially to deletions and insertions, which lead to satellite length variation. Length variation in microsatellites is called microsatellite instability (Bennett 2000; Subramanian *et al.* 2003).

MSI is a typical feature of LS tumors (Aaltonen *et al.* 1993; Ionov *et al.* 1993; Thibodeau *et al.* 1993). Initially, researchers used a wide range of different markers in MSI studies. In 1997, in a meeting in Bethesda, researchers reviewed the results collected from instability studies to develop the most suitable panel of markers to recognize MSI from tumors and consequently created the "Bethesda guidelines" to diagnose LS (Table 1) (Umar *et al.* 2004; Lynch *et al.* 2008). That marker panel, which consists of two mononucleodide markers (BAT25, BAT26) and three dinucleotide markes (D2S123, D5S346, D17S250) differentiating CRC tumors with MSI-high (MSH-H), MSI-low (MSI-L) and MSI-stable (MSS) phenotypes, has since then been widely used.

Since MMR deficient cells cannot repair DNA replication slippage errors, MMR deficient tumors are strongly associated with microsatellite instability. About 90 % of LS tumors show MSI (Aaltonen *et al.* 1993; Ionov *et al.* 1993; Thibodeau *et al.* 1993). However, the degree and type of MSI differs from low (at least one marker shows instability) to high, when at least 40 % of markers are unstable (Umar *et al.* 2004) and between mono-, di-, tri-, and tetranucleotide instability or elevated microsatellite alterations at selected tetranucleotide repeats (EMAST) (Peltomäki and Vasen 2004; Plaschke *et al.* 2004; Haugen *et al.* 2008). The MLH1 and MSH2 deficient tumors show typically high MSI and in both mono- and dinucleotide repeats, whereas in MSH6 deficient tumors, the level of MSI is generally lower (Bhattacharyya *et al.* 1995; Papadopoulos *et al.* 1995), and it has been recently shown that mononucleotide markers have a high sensitivity to detect *MSH6* mutation carriers (de la Chapelle and Hampel 2010). MSH6 deficient cells are not able to repair single base mismatches, while they retain proficiency to repair two, three, and four base loops (Drummond *et al.* 1995; Risinger *et al.* 1996; Umar *et al.* 1997) thus, causing only mononucleotide repeat instability in tumors (Wagner *et al.* 2001; Plaschke *et al.* 2004; de la Chapelle and Hampel 2010). EMAST and also low dinucleotide repeat instability have recently been associated with MSH3 deficiency, both in tumor cell lines and in sporadic colorectal tumors, while no *MSH3* mutations are found in LS (Haugen *et al.* 2008).

Loss of MMR protein expression is connected to MMR deficiency and, other than MSI, it is another typical feature of LS tumors. A mutation in a MMR gene leads to a lack of a respective protein in a tumor. Immunohistochemical (IHC) staining of MLH1, MSH2, MSH6 and PMS2 proteins is a widely used, low cost, and sensitive diagnostic method to study MMR deficiences. IHC is based on protein identification by a specific antibody and thus, detects both the MMR deficiency and the deficient protein. Sometimes a loss of one MMR protein also leads to the degradation of its counterpart. For instance, it has been shown that a lack of MSH2 protein in a tumor leads to degradation of MSH6, while only in rare cases MSH6 loss causes degradation of MSH2 (Bedeir and Krasinskas 2011). Together IHC and MSI analyses form a sensitive tool to diagnose LS tumors. In summary, the data of cancers in

a family and results of MSI and IHC analyses form a first step in LS diagnostics. When LS is suspected, studies continue with mutation search.

MMR gene regulation changes

A mutation in an LS susceptibility gene leads to MMR malfunction, accumulation of mutations across the entire genome, and to genome instability. MSI may lead to inactivation of tumor suppressor genes and activation of oncogenes, which contain a repetitive region as a target site such as the genes *TGFβRII* (transforming growth factor β receptor II), *Bax*, *MSH3,* and *MSH6* (Umar *et al.* 2004). When a positive MSI phenotype and/or a loss of MMR protein(s) is found in a tumor, LS needs to be confirmed by mutation search. Unfortunately, in 50 % of the suspected LS cases an inherited predisposing mutation cannot be found in known MMR genes (Umar *et al.* 2004). One reason for that is that not all MSI positive tumors are LS tumors. Although, the MSI phenotype is a hallmark of LS tumors, it is also present in approximately 10-15 % of sporadic CRCs and in some other extracolonic cancers as well. In contrast to LS, in sporadic tumors MSI is typically caused by epigenetic silencing of the *MLH1* gene via promoter hypermethylation (Veigl *et al.* 1998). The epigenetic silencing of *MLH1* complicates the LS diagnosis, since together with an MSI phenotype, it also leads to a loss of the MLH1 protein in a tumor.

It has been suggested that large genomic rearrangements in *MSH2* and *MLH1* are relatively common (Nakagawa *et al.* 2003). Recently, several groups have reported that germline EPCAM (Epithelial Cell Adhesion Molecule) deletions may also cause LS (Kovacs *et al.* 2009; Ligtenberg *et al.* 2009; Niessen *et al.* 2009) since 3' EPCAM deletions cause the methylation of the *MSH2* promoter and hence silencing of the neighbouring *MSH2* gene. Furthermore, patients showing 3' EPCAM mutations almost exclusively represent colon cancer (Lynch *et al.* 2011), although cancer risk is dependent on the size and location of the deleted EPCAM region (Kempers *et al.* 2011; Lynch *et al.* 2011). Such EPCAM deletion may explain MSI and loss of the MSH2 protein in tumors, although patients have no mutations in *MSH2* (Lynch *et al.* 2011).

Variations of uncertain significance in MMR genes

MMR gene variations

Approximately 97 % of all reported LS germline mutations are found in three different MMR genes, *MLH1*, *MSH2,* and *MSH6* (de la Chapelle 2004; Woods *et al.* 2007). Altogether 1515 germline variations have been reported in the database; 659 *MLH1* (44% of the all identified MMR gene variations), 595 *MSH2* (39%), 216 *MSH6* (14%), and 45 *PMS2* (3%) (Woods *et al.* 2007; http://www.insight-group.org/). Regardless of the MMR gene in question,

truncating MMR gene mutations, which destroy protein structure, are generally considered to be disease causing. The amino acid substitutions, deletions, and insertions, which change only one amino acid in a protein structure, do not necessarily destroy protein and therefore their pathogenicity is difficult to interpret. A variant of uncertain significance (VUS), also known as an unclassified variant, is an alteration in a gene sequence whose association with the disease risk is unknown (Goldgar *et al.* 2008). Today's focus in international collaborative research for gastrointestinal hereditary tumors, is to functionally characterize all VUS found in suspected LS families and thus distinguish non-pathogenic variants from pathogenic ones facilitating gene testing and genetic counseling in these families (Couch *et al.* 2008).

In silico predictions

The most effective and quick studies to predict the functional effects and assess pathogenicity of individual variations are done *in silico*. These computational analyses mainly identify conserved areas of a gene through multiple sequence alignment analyses across numerous species, and thereafter, deduce possible functional defects caused by the variation. Due to their high sensitivity and specificity (Tavtigian *et al.* 2008), the best *in silico* prediction algorithms, which have been chosen to analyze the possible effects of the individual MMR gene variations, are sorting intolerant from tolerant (SIFT) (Ng and Henikoff 2001) (http://sift.jcvi.org/), the multivariate analysis of protein polymorphism (MAPP-MMR) (Stone and Sidow 2005; Chao *et al.* 2008) (http://mendel.standford.edu/SidowLab/), and polymorphism phenotyping (PolyPhen-2 (version 2.1.0; HumDiv)) (Adzhubei *et al.* 2010) (http://genetics.bwh.harvard.edu/pph/). The MAPP-MMR algorithm is not, however, compatible for *MSH6* VUS predictions. Unfortunately, the pathogenicity caused by two or more MMR gene variations in one carrier is also impossible so far to predict *in silico* and requires much more complicated and laborious functional studies.

Homozygous and compound heterozygous mutations in MMR genes

The susceptibility to LS is generally associated with one inherited mutated MMR gene allele, however, in rare cases, an individual can carry two MMR gene mutations. The studies on phenotype-genotype correlations in individuals who have inherited homozygous or biallelic compound heterozygous MMR gene mutations has recently been comprehensively summarized (Felton *et al.* 2007; Wimmer and Etzler 2008; Durno *et al.* 2010). An individual who has inherited a monoallelic MMR mutation has an increased susceptibility to cancer, while carriers of homozygous or biallelic compound heterozygous mutations, which severely damage the protein structure and function, develop hematological and brain malignancies during the first or second decade of life (Felton *et al.* 2007). In addition, in contrary to LS, the normal tissue of mutation carriers is also MMR deficient because of the constitutional deficiency (Durno *et al.* 2010). Furthermore, distribution of predisposing mutations among the MMR genes differs. Compared to approximately 90 % of heterozygous mutations in typical LS, only 14 % of biallelic MMR gene mutations predisposing to gastrointestinal

cancers are located in *MLH1* and *MSH2* (de la Chapelle 2004; Durno *et al.* 2010). The type and site of a mutation may also effect the age of cancer onset and tumor spectrum. Generally, the more severe the mutation is, the more severe is the phenotype (Felton *et al.* 2007). Approximately, two thirds (2/3) of biallelic mutations are of nonsense, frameshift, or large deletion types of mutations, and one third (1/3) are missense, splice site, or in-frame deletion mutations (Durno *et al.* 2010). The family history of a biallelic mutation carrier does not typically fulfill ACI/II and suggest LS. Surprisingly, almost 60 % of patients with biallelic MMR gene mutations have no past medical or family history suggestive of a hereditary cancer syndrome.

A small group of cancer patients have inherited two missense variations (pathogenic or VUS) either in a same MMR gene allele (monoallelic), in different MMR gene alleles (biallelic), or in different MMR genes. So far, altogether 18 different inherited pairs of MMR gene missense variations have been reported (Table 2).

Table 2. Heterozygous MMR gene missense variations found as pairs in cancer patients.

Gene(s)	Variation I	Variation II	Cancer type/ index patient	Age of onset	Reference
MLH1/ MLH1	p.Ser44Phe	p.Ala441Thr	BC	35	(Hackman *et al.* 1997)
MLH1/ MLH1	p.Lys618Ala	c.606-2A>G	CRC/Sarcoma	35/65	(Liu *et al.* 1999)
MLH1/ MLH1	p.Val722Ile	c.1039 -8T>A	CRC	30	(Christensen *et al.* 2009)
MLH1/ MLH1	p.Lys618Ala	p.Arg659Gln	CRC	32	(Raevaara *et al.* 2005)
MSH2/ MSH2	p.Asn127Ser	p.Ala328Pro	CRC	65	(Samowitz *et al.* 2001)
MSH2/ MSH2	p.Glu205Gln	p.Val367Ile	PC	59	(Gargiulo *et al.* 2009)
MSH2/ MSH2	p.Gly322Asp	p.Asp487Glu	EC	57	(Hampel *et al.* 2006)
MSH2/ MSH6	p.Ile145Met	p.Arg1095His	CRC	65/74	(Kariola *et al.* 2003)
MSH2/ MSH6	p.Ile145Met	p.Leu1354Gln	CRC	53	(Kariola *et al.* 2003)
MSH2/ MSH6	p.Val923Glu	p.Ser1188Asn	CRC	70	(Ollila *et al.* 2006)
MSH6/ MSH6	p.Cys765Trp	p.Val878Ala	CRC	31	(Plaschke *et al.* 2006)
MSH6/ MSH6	p.Leu435Pro	p.Val878Ala	EC	59	(Hampel *et al.* 2006)
MSH2/MLH3	p.Glu198Gly	p.Trp1276Arg	CRC	29	(Liu *et al.* 2003)
MSH6/MLH3	p.Val878Ala	p.Glu1451Lys	CRC	45	(Wu *et al.* 2001)
MLH1/MSH2	p.Glu460Ala	p.Met663fs	CRC	53	(Christensen *et al.* 2008)
MLH1/MSH2	p.Thr117Met	p.Gly322Asp	CRC	39	(Lee *et al.* 2005)
PMS2/ PMS2	p.Ser46Asn	p.Ser46Ile	CRA	4	(Jackson *et al.* 2008)
MSH2/MLH1/ MSH6	p.Met688Val	p.Thr117Met p.Ala1339Val	CRC	46	(Christensen *et al.* 2008)

Whether and how a variation pair affects a carrier's clinical phenotype and cancer risk depends on the variations' locations and pathogenicity: if 1) only one of the variations is pathogenic, a carrier is a typical LS mutation carrier; 2) the two variations are in the same gene but in both parental alleles and both are pathogenic, a carrier resembles a homozygous mutation carrier (Felton *et al.* 2007); 3) neither of the two variations is pathogenic (Kariola *et al.* 2003), a carrier is not predisposed to cancer more than general population; or 4) variations, which are non-pathogenic individually, may increase pathogenicity together

(Martinez and Kolodner 2010). A family pedigree including cancer data and mutation carriers plays a key role when studying whether variations are located in the same or different parental alleles. If two inherited variations are located in a same gene allele, or in different MMR genes, a carrier still has one wild type (WT) allele of the gene(s), whereas if they are located in the same gene but in different alleles, a carrier has no wild type allele of the gene in a constitutive genome (Figure 3). This may impose a difference in cancer susceptibility if both variations are pathogenic.

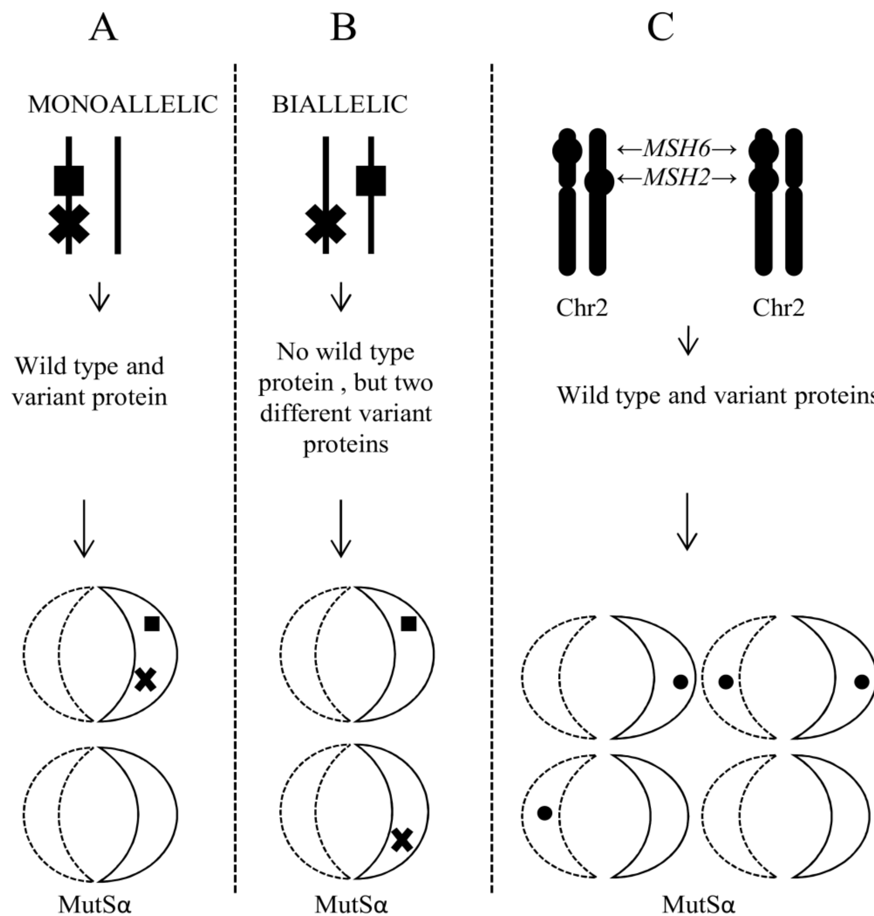


Figure 3. Description of protein products and heterodimer alternatives (here MutSα) when two different variations locate in a same gene (*MSH2* or *MSH6*) A) in one allele, B) in different alleles, or C) in different genes (*MSH2* and *MSH6*).

Functional analyses of MMR gene variants

When an inherited MMR gene variation has been identified, it is necessary to determine whether it is pathogenic or not. This chapter summarizes the functional assays, which have been developed and used by different research groups to test the functional significance of MMR variations. Functional assays can be divided into two groups according to their aims; to those, which study specific function of the variant protein, and to those, which evaluate

MMR repair capacity as a complete process (Ou *et al.* 2007). The latter group can be further divided into *in vivo* and *in vitro* assays. Functional analyses, which measure the repair capacity of a variant protein, are the most applicable to a general assessment of pathogenicity, while the other functional tests may give specific information about the causes of repair deficiency and pathogenicity.

Yeast-based functional assays have been developed based on the fact that the MMR system is evolutionarily well conserved (Ou *et al.* 2007). Functional yeast assays have been used to study variations found in CRC patients by creating a corresponding change in the yeast MMR gene, because human proteins are not functional in yeast (Shcherbakova and Kunkel 1999; Gammie *et al.* 2007). In this assay, the variant proteins, which are not able to complement the absence of a wild type protein, cause a strong mutator phenotype and are interpreted as pathogenic.

Another functional assay performed in yeast (*Saccharomyces cerevisiae*) is based on the observation that the expression of the human MLH1 WT protein in yeast, prevents its functional MMR, while the mutated human MMR protein does not interfere with yeast MMR (Shimodaira *et al.* 1998). This phenomenon is described as a dominant mutator effect. If the analyzed protein causes the dominant mutator phenotype, the variant is interpreted as pathogenic (Shimodaira *et al.* 1998; Drotschmann *et al.* 1999).

Three slightly different *in vitro* complementation assays have been published (Marra *et al.* 1998; Nyström-Lahti *et al.* 2002; Drost *et al.* 2010). These assays are based on previous studies, which have shown that cell extracts are capable to repair mismatches in artificial substrates. The first reported *in vitro* MMR assay used a bacteriophage derived heteroduplex that contains a base pair error in the *LacZ α*-complementation gene as a substrate (Marra *et al.* 1998). Cytoplasmic protein extract (CE), which lacks the MMR protein in question, is supplemented with recombinant MMR protein (WT or variant) produced in insect cells and incubated together with the heteroduplex. After the reaction, the heteroduplex is transformed into *E. coli* cells and grown together with the α-complementation bacterial strain. Functional (repaired) *LacZ* gene increases the number of blue colonies, which indicates a successful repair.

The assay above was modified by complementing MMR deficient human nuclear extracts (Nyström-Lahti *et al.* 2002). This *in vitro* MMR assay has been especially utilized in determining the pathogenicity of the MMR gene variations found in putative LS patients. The equivalent missense variations are constructed with a PCR-based site-directed mutagenesis method to cDNA and the recombinant MMR proteins produced in a suitable host, such as *Spodopthera frugiperda* 9 insect cells. The ability of a variant protein to repair heteroduplex substrates is studied by a complementation assay. If a variant protein restores a repair capacity of the deficient nuclear protein extract, which lacks the examined MMR protein, repaired heteroduplexes can be cleaved by appropriate restriction enzymes (Lahue *et al.* 1989; Holmes *et al.* 1990; Nyström-Lahti *et al.* 2002).

A cell free assay was recently published, which is very similar to the *in vitro* MMR assay described above (Drost *et al.* 2010). The main differences between these two assays are in protein production and repair detection. Here, variant proteins are produced by the *in vitro* transcription-translation system and successful repair is measured from signals coming from fluorescently labelled heteroduplexes.

To mention some functional assays, which do not assess MMR capacity as a complete process but study some specific parts of it, EMSA (electrophoretic mobility shift assay) is a method, which can be used to study mismatch binding and releasing activity of MutSα and MutSβ protein complexes (Clark *et al.* 1999; Drotschmann *et al.* 1999; Heinen *et al.* 2002; Ollila *et al.* 2008a; Ollila *et al.* 2008b). In this method, studied variant proteins are produced either in yeast or other suitable host, after which the proteins are purified. The test is based on the fact that MMR complexes will bind and release to oligonucleotides that contain a mismatch when they are incubated together. The binding can detected by running the samples in a native polyacrylamide gel where free oligonuclotides move faster than the ones bound to protein complexes. Variations, which affect the normal binding or releasing of complexes, are presumed to be pathogenic.

Several methods to study protein-protein interactions have been developed. For example, Glutathione S-transferase (GST) fusion protein interaction method, in which studied proteins are produced *in vitro,* has been successfully used to examine how MMR gene variations affect protein-protein interactions. In this assay, the variations, which interfere with interaction, are interpreted as pathogenic, while the variations, which do not affect the interaction, remain unclear (Guerrette *et al.* 1998; Guerrette *et al.* 1999). A co-immunoprecipitation method, or antibody pull down method, has also been used to study the interaction of protein patners in heterodimer complexes (Kariola *et al.* 2002; Nyström-Lahti *et al.* 2002; Ollila *et al.* 2006; Korhonen *et al.* 2008). Here, MutLα or MutSα protein complexes are incubated with an antibody recognizing one of the proteins and after protein collection (pulling down), all the proteins in precipitation are analyzed by western blot (WB). The interaction is not disturbed if both protein partners are present in equal amounts as in the wild type protein complex. The third widely used interaction analysis is a yeast two hybrid method which measures physical interaction (*in vivo*) between two MMR proteins in yeast (Kondo *et al.* 2003). Together with different MMR, mismatch binding, and protein-protein interaction assays, there are several other available functional analysis such as subcellular localization assays (Raevaara *et al.* 2005; Gammie *et al.* 2007), which can be used together with the clinical and tumor pathological data to determine the pathogenicity of MMR gene variations.

Pathogenicity assessment of MMR variations

A mutation is a change in DNA sequence, which has an effect on the expression and/or function of a gene, whereas a polymorphism is a variation without harmful effect and usually frequently present also in the healthy population. It is important for genetic counselling and follow-up treatments of carriers in LS or putative LS families that the functional significance

and pathogenicity of their variations is understood (Ou *et al.* 2007; Ou *et al.* 2008). Truncating and splice site mutations are generally shown to cause a complete lack of protein or produce a severely damaged protein. The challenge is in the interpretation of missense mutations, where only one amino acid is changed in a protein structure. Most of these variations are not easy to interpret and often require several functional analyses (Cotton and Scriver 1998; Syngal *et al.* 1999; Nyström-Lahti *et al.* 2002). Hundreds of VUS have already been found and reported in MMR genes and the number is rapidly increasing (www.insight-group.org). The wide variety of clinical phenotypes in CRC families further complicates pathogenicity assessments and LS diagnostics. Non-truncating MMR gene alterations often associate with atypical clinical phenotypes with a later age of cancer onset and low or no MSI in tumors (Wu *et al.* 1999; Wagner *et al.* 2001; Berends *et al.* 2002; Hendriks *et al.* 2004); since different alterations in a same functional domain and even in a same codon in a MMR gene can cause a complete elimination or, in contrast, little to no effect on protein function (Ellison *et al.* 2001; Raevaara *et al.* 2005). Without pathogenicity assessments based on validated biochemical analyses, the increasing number of VUS will reach a bottleneck stage in LS diagnostics

Overall, the effects of MMR gene-mutations can be expected to fall into six broad classes: interference of DNA binding, loss of ATPase activity, loss of allosteric communication between DNA and ATP binding sites, loss of protein-protein interactions with downstream effectors, loss of MSH2-MSH6 interaction, and general loss of protein stability (Warren *et al.* 2007). A recently established database focusing on missense mutations and small in-frame deletions in MMR genes (www.mmruv.info) includes results of functional and/or *in silico* data. At present, the database contains information for 573 variants, 345 in *MLH1*, 186 in *MSH2*, 20 in *MSH6*, 11 in *PMS2*, and 11 in *MLH3*. Based on the functional data, more than half of these variants seem to be pathogenic, underlining the clinical importance of functional studies (Ou *et al.* 2008).

To facilitate the interpretation of VUS, a decision tree for the *in vitro* analysis of variants in MMR genes in suspected Lynch syndrome cases has been proposed (Couch *et al.* 2008). This model includes three major steps. When family history and/or young age of cancer onset predicts LS and results of immunohistochemical, and MSI analyses suggest MMR deficiency in a tumor tissue, genetic testing is carried out (Step 1, sequencing of relevant MMR genes). When a missense variation (VUS) is identified, the second step (Step 2) dedicated to analyses of splice aberrations is (Spurdle *et al.* 2008), *in silico* predictions of pathogenicity (Tavtigian *et al.* 2008) and *in vitro* MMR. If the variant protein retains activity in the *in vitro* MMR assay, more specific functional studies, such as MMR protein stability, protein-protein interactions, and protein subcellular localisation studies, need to be carried out in Step 3. Deficiency in any of these *in vitro* assays indicates LS. Recently, a study to verify the three step model in pathogenicity assessment performed by using 74 MMR gene VUS (37 *MLH1*, 26 *MSH2*, 11 *MSH6*) demonstrated how surprisingly appropriate the model for LS diagnostics is (Kansikas *et al.* 2011).

AIMS OF THE STUDY

The objective of the study was to functionally characterize MutS homologue mismatch repair proteins and their variants (aim 1 and 2). Moreover, the *in vitro* MMR assay was used to functionally evaluate the role of different replicative polymerases, especially pol α and ε in MMR (aim 3).

Specific aims:

1) To determine MutSα and MutSβ substrate specificities and MMR efficiencies **(I)**
2) The assess the cause of pathogenicity in cases where two inherited MMR gene variations are found in cancer patients **(II, III)**
3) To study the role of replicative polymerases α and ε in MMR **(IV)**

MATERIALS AND METHODS

Descriptions of materials and methods in detail can be found in original articles (**I-IV**).

Study material

VUS in studies I, II and III

Studies **I**, **II**, and **III** consist of 9 VUS pairs affecting MMR genes *MSH2* (NM_000251.1) and/or *MSH6* (NM_000179.2). Nucleotide numbering reflects cDNA numbering with +1 corresponding to the A of the ATG translation initiation codon in the reference sequence, according to www.hgvs.org/mutnomen. The initiation codon is codon 1. Three VUS pairs are in the *MSH2* gene c.380A>G/c.982G>C; (p.Asn127Ser/p.Ala328Pro) (study **III**), c.613G>C/c.1099G>A; (p.Glu205Gln/p.Val367Ile) (study **III**), c.965G>A/c.1461C>G; (p.Gly322Asp/p.Asp487Glu) (study **III**), two in the *MSH6* gene c.1304T>C/c.2633T>C; (p.Leu435Pro/p.Val878Ala) (study **III**), c.1754T>C/c.2030G>C; (p.Leu585Pro/p.Ser677Thr) (study **III**), and in four pairs there is one VUS in the *MSH2* gene and one in the *MSH6* gene in the same patient c.2726A>T;/c.2633T>C; (p.Lys909Ile/p.Val878Ala) (study III), c.435T>G/c.3284G>A; (p.Ile145Met/p.Arg1095His) (study **III**), c.435T>G/c.4061T>A; (p.Ile145Met/p.Leu1354Gln) (study **III**), p.Val923Glu/p.Ser1188Asn (study **II**). Of these, the VUS pairs p.Lys909Ile/p.Val878Ala and p.Leu585Pro/p.Ser677Thr have not been reported before. In addition, the VUS c.2386C>T (p.Arg796Trp) in *MSH3* was included in study **I**. The alterations, age of cancer onset, and tumor pathological data of the VUS carriers are collected in Table 3. Locations of the variants in the MSH2 and MSH6 functional domains are shown in the Figure 4, and the pedigrees of the families are presented in Figure 5.

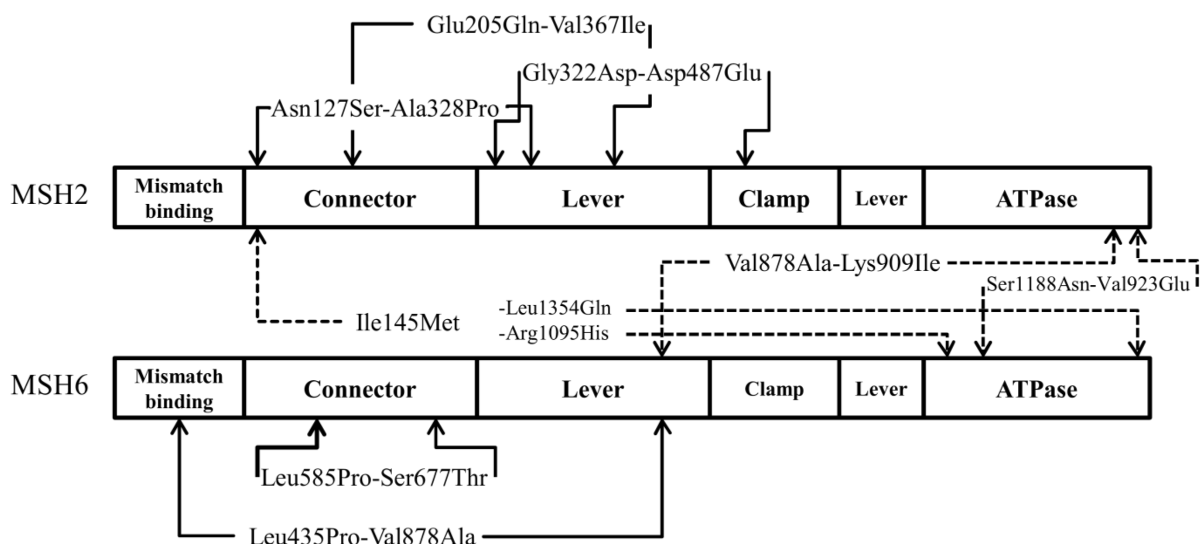


Figure 4. Locations of the studied VUS in the MutSα functional domains. Functional domains adapted from Warren *et al.* 2007.

Table 3. Data of studied VUS and their carriers

VUS and VUS pairs	Nucleotide change	Protein variations	Protein domain[a]	VUS carrier							Reference
				Pedigree/family member	Tumor type[b]	Age of onset	MSI[c]	IHC[d]			
								MLH1	MSH2	MSH6	
MSH2/2	c.380A>G c.982G>C	p.Asn127Ser p.Ala328Pro	Connector Lever	N/A	CRC	65	MSI-H	N/A	N/A	N/A	(Samowitz *et al.* 2001)
	c.613G>C c.1099G>A	p.Glu205Gln p.Val367Ile	Connector Lever	Fig. 5A	PC	59	N/A	N/A	N/A	N/A	(Gargiulo *et al.* 2009)
	c.965G>A c.1461C>G	p.Gly322Asp p.Asp487Glu	Lever Clamp	N/A	EC	57	MSI-L	-	+	- (5%)	(Hampel *et al.* 2006)
MSH2/6	c.435T>G c.3284G>A	p.Ile145Met p.Arg1095His	Connector ATPase	Fig. 5B	CRC	65/74	MSI-H	+	+	+	(Kariola *et al.* 2003)
	c.435T>G c.4061T>A	p.Ile145Met p.Leu1354Gln	Connector ATPase	Fig. 5C	CRC	53	MSI-H	+	-	-	(Kariola *et al.* 2003)
	c.2726A>T c.2633T>C	p.Lys909Ile[e] p.Val878Ala[e]	ATPase Lever	Fig. 5D	CRC	79	MSI-H	-	+	+	**III**
	c.2768T>A c.3563G>A	p.Val923Glu p.Ser1188Asn	ATPase ATPase	Fig. 5E/a	CRC	56	N/A	-	+	±	(Ollila *et al.* 2006)
	c.2768T>A c.3563G>A	p.Val923Glu p.Ser1188Asn	ATPase ATPase	Fig. 5E/b	CRC	70	MSS	+	±	-	(Ollila *et al.* 2006)
	c.2768T>A c.3563G>A	p.Val923Glu p.Ser1188Asn	ATPase ATPase	Fig. 5E/c	CRA	37	MSI-H	N/A	N/A	N/A	(Ollila *et al.* 2006)
	c.2768T>A c.3563G>A	p.Val923Glu p.Ser1188Asn	ATPase ATPase	Fig. 5E/d	BCC	52	N/A	N/A	N/A	N/A	(Ollila *et al.* 2006)
MSH6/6	c.1304T>C c.2633T>C	p.Leu435Pro p.Val878Ala	Mismatch binding Lever	Fig. 5F	EC	59	MSI-H	+	+	-	(Hampel *et al.* 2006)
	c.1754T>C c.2030G>C	p.Leu585Pro[e] p.Ser677Thr[e]	Connector Connector	Fig. 5G	CRC	38	MSI-H	+	+	- (eqv.)	**III**
MSH3	c.2386C>T	p.Arg796Trp	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	**I**

[a]According to Warren *et al.* 2007. [b]BCC, Basocellular carcinoma; CRC, colorectal cancer; EC, endometrial cancer; PC, pancreatic cancer. [c]Microsatellite instability: MSI-H, ≥3 markers indicating MSI; MSI-L, ≤2 markers indicating MSI. [d]Immunohistochemistry: +, present; -, absent; ± reduced or heterogenous expression; N/A, Data not available; Eqv., Equivocal. [e]Mutation, MSI and IHC analysis were performed as described Hampel's work (Hampel *et al.* 2005).

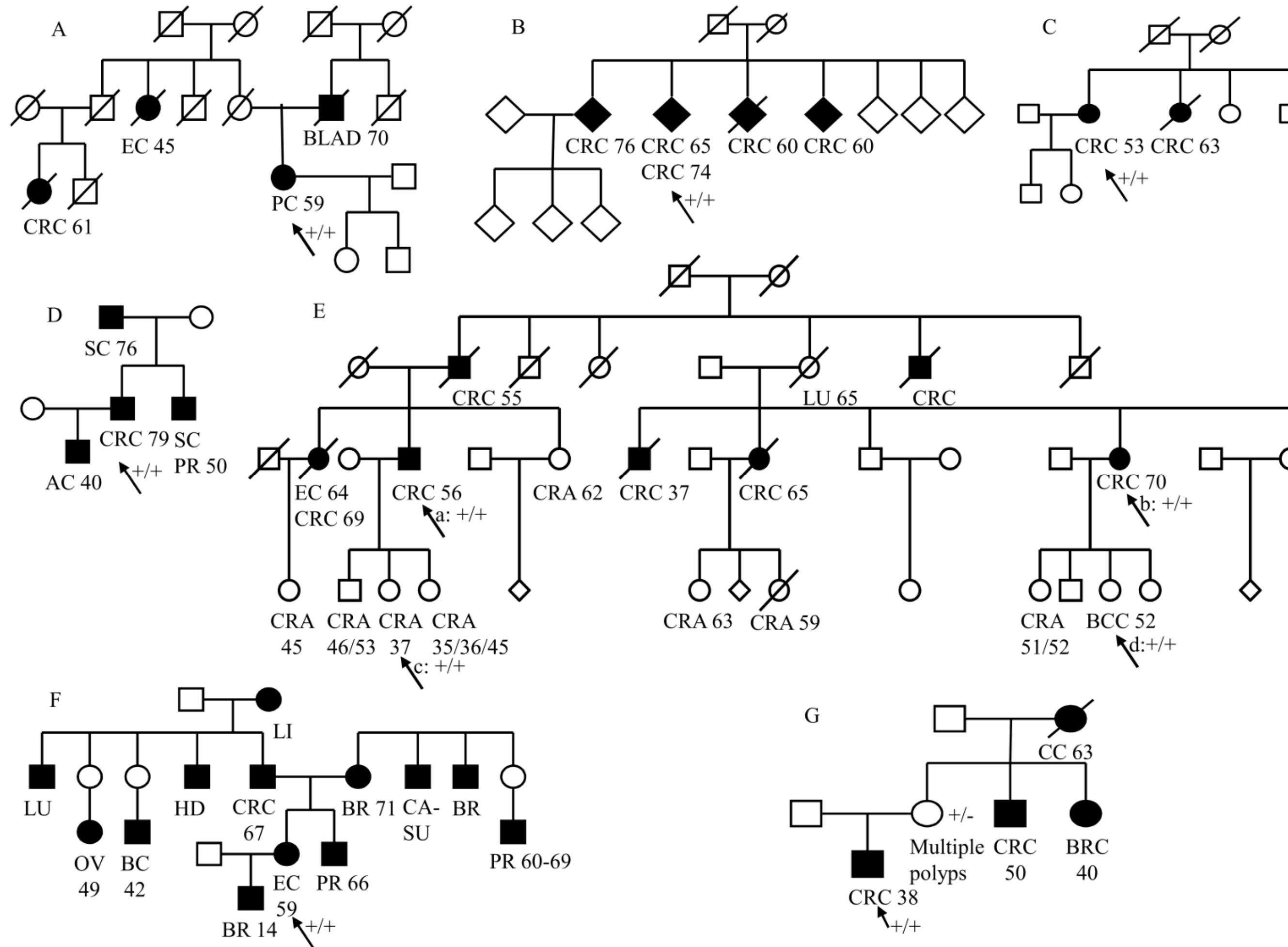


Figure 5. Pedigrees of the families where the proband carriers two VUS. Arrow and +/+; carrier of two VUS, +/-; carrier has only one VUS. Tumor types and ages at cancer onset are marked. Abbreviations of the tumor types: AC; anal cancer, BC; bone cancer, BCC; Basocellular carcinoma, BLAD; bladder cancer, BR; brain cancer, BRC; breast cancer, CASU; unknown cancer, CC; cervical cancer, CRA; colorectal adenoma, CRC; colorectal cancer, EC; endometrial cancer, HD; Hodking's lymphoma, LI; liver cancer, LU; lung cancer, OV; ovarian cancer, PC; pancreatic cancer, PR; prostata cancer, SC; stomach cancer.

Summary of the methods

All the methods used in the studies are listed in Table 4.

Table 4. Methods used in the studies

METHOD	Additional information	Source/reference	Used in
Variant protein production			
PCR: Mutagenesis; QuikChange®lightning Site-directed mutagenesis	RefSeg: NM 000251.1 (*MSH2*), NM 002439.2 (*MSH3*), NM 000179.2 (*MSH6*)	Stratagene	I, II, III
Production of recombinant baculoviruses	Bac-to-bac	(Nyström-Lahti *et al.* 2002)	I, II, III
Production of recombinant proteins	Bac-to-bac, Table 5.	(Kariola *et al.* 2002; Ollila *et al.* 2006; Ollila *et al.* 2008b)	I, II, III
Protein analysis			
Western blotting	Table 6.	(Kariola *et al.* 2002; Ollila *et al.* 2006; Ollila *et al.* 2008b)	I, II, III
3D analysis		(Berman *et al.* 2000; Warren *et al.* 2007; Holm and Rosenström 2010)	III
Functional testing of proteins			
In vitro MMR assay		(Kariola *et al.* 2002; Nyström-Lahti *et al.* 2002; Ollila *et al.* 2006)	I, II, III, IV
Nuclear protein extraction		(Holmes *et al.* 1990; Alvino *et al.* 2006)	I, II, III, IV
Heteroduplex preparation	Figure 6.	(Lahue *et al.* 1989)	I, II, III, IV
Neutralization	Table 6.	-	IV
Immunodepletion	Table 6.	(Tanaka *et al.* 1982)	IV

Cell lines and antibodies

Different cell lines were used for protein production and nuclear protein extraction (Table 5). Furthermore, a panel of antibodies was used for protein detection in the western blot analyses as well as for immunodepletion and neutralization assays (Table 6).

Table 5. Cell lines used in the studies

Cell lines	Description	Protein deficiency[a]	Source[b]	Used in
GP5d	Human colorectal adenocarcinoma cells	MSH2, MSH3, MSH6, MLH1	ECACC	I
HCT116	Human colorectal adenocarcinoma cells	MLH1, MSH3	ATCC	I
HeLa	Cervix	-	ATCC	I
HeLa	Cervix	-	M. Frilander, UH	II, III, IV
LoVo	Human colorectal adenocarcinoma cells	MSH2, MSH3, MSH6	ATCC	I, II, III
Sf9	Insects cells (*spodoptera frugiperda*)	-	Gibco BRL	I, II, III

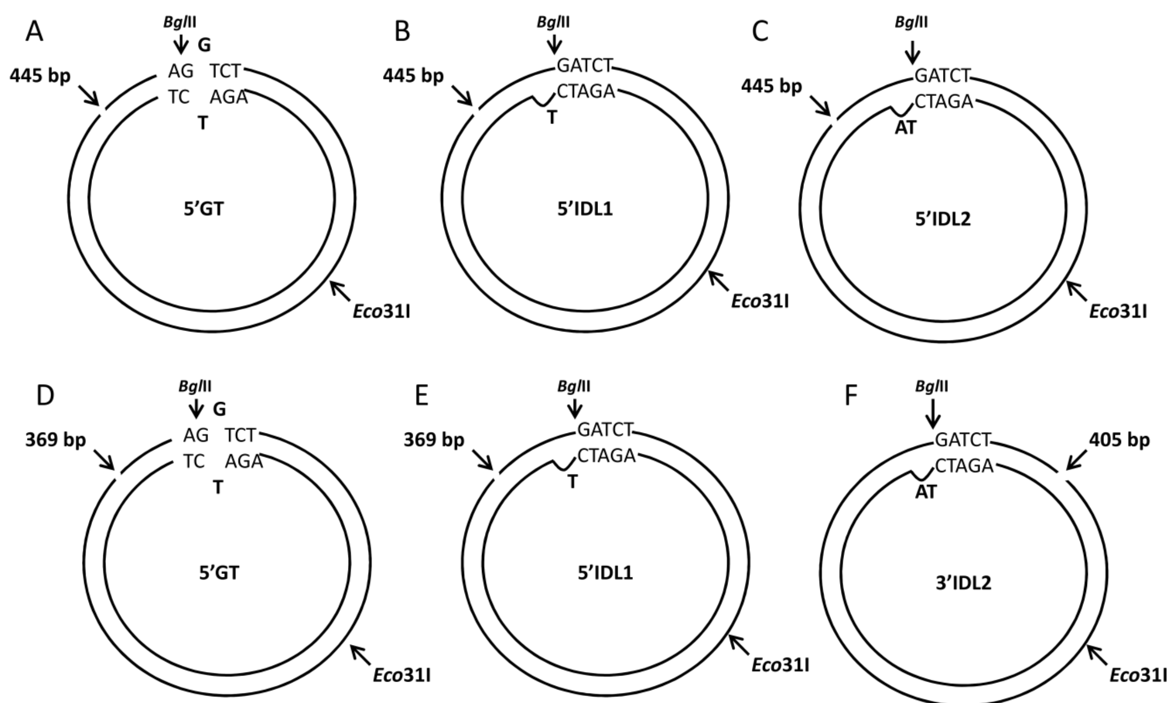
[a](Cannavo *et al.* 2005), [b]ATCC; American Type Culture Collection, ECACC; European Collection of Cell Culture

Table 6. Antibodies used in the studies

Antigen	Antibodies used in western blotting	Source	Used in
α-tubulin	DM1A, 0,2 mg/ml	Sigma	I
MLH1	Clone 168-15, 0,5 mg/ml	BD Biosciences/Pharmingen	I
MSH2	MSH2-Ab1, NA-26, 0,2 mg/ml	Calbiochem	I, II, III
MSH3	M94120, 250 mg/ml	BD Transduction, Laboratories	I
MSH6	Clone 44, 0,02 mg/ml	BD Transduction Laboratories	I, II, III
PMS2	Ab-1, 0,2 mg/ml	Calbiochem/Oncogene Research	I
Antigen	**Antibodies used in neutralisation and immunodepletion**	**Source/Reference**	**Used in**
MSH6	Clone 44, 0,02 mg/ml	BD Transduction Laboratories	IV
Polα	SJK-287-38	(Tanaka *et al.* 1982)	IV
Polα	SJK-132-20	ATCC CRL-1640, protein G, (Tanaka *et al.* 1982)	IV
Polε	K18	(Pospiech *et al.* 1999)	IV

Heteroduplexes used in the in vitro MMR assays

Figure 6 represents substrates which were used in the *in vitro* MMR assays. Four different heteroduplex constructs were prepared: GT mismatch (5'GT), a single (5'IDL1), and two nucleotide IDLs (5'IDL2 and 3'IDL2). The protocol for heteroduplex preparation is described in Lahues's work (Lahue *et al.* 1989). The heteroduplex DNA is a circular molecule (3193 bp long) with a single-strand nick upstream or downstream from the site of the error. The 5'IDL1 contains a deletion of 1 nt (delA) and 5'- and 3'IDL2 contains a deletion of 2 nt (delAT) in the unnicked strand. The GT mismatch was created by replacing an adenine with guanine maintaining a thymine on the complementary strand.



Substrate	A, B, C	D, E	E	E, C, F
Used in:	**I**	**II**	**III**	**IV**

Figure 6. The substrates used in the *in vitro* MMR, neutralization, and immunodepletion experiments. The bottom strand of the heteroduplex contains a whole *Bgl*II -restriction site, whereas the top strand contains an error. GT heteroduplex were created by replacing adenine to guanine (A, D). IDLs were created by deleting 1 nt (B, E) or 2 nt (C, F) from the top strand. Three different restriction enzymes were used to create a nick either in the 5' or 3' position: *Dra*III created a 5' nick 445 bp from a mismatch or IDL (A, B, C), *Ban*II created a 5' nick 369 bp from an error (D, E), and A*fl*III created a 3' nick 405 bp from the an IDL (F).

RESULTS AND DISCUSSION

In the text, the original figures from the articles (**I-IV**) are cited by using the number of the original article and the Figure.

Mismatch repair analyses of VUS and their pairs (I, II and III)

Clinical and tumor pathological data of the VUS carriers (II, III)

As seen in Table 3, clinical data such as the tumor type, the age of onset, the MSI status, and the immunohistochemical (IHC) analysis of the tumors may represent the clinical phenotype caused by the coexistence of the two VUS. Excluding the carrier of *MSH6* p.Leu585Pro/p.Ser677Thr, all other compound VUS carriers have a relatively late age of cancer onset compared to that typically associated with LS. However, the tumor types such as CRC and endometrial cancer affecting most of the carriers as well as the high MSI status seen in cases and the protein expression status in their tumors, affecting most of the carriers, belong to the typical LS tumor spectrum. However, excluding the patient carrying the pair MSH6 p.Leu585Pro/p.Ser677Thr, whose family data showed that the two variations were inherited from different parents, and the family pedigrees did not reveal the type of inheritance of the two VUS (Figure 5). Altogether, the families of 4 VUS pair carriers fulfill the LS criteria (ACII), families of 3 carriers and do not fulfill, and of 2 families, the data is not available.

Concomitant effect of the VUS pairs to MMR deficiency (III)

This study was designed to mimic two different modes of inheritance, where the variations are either inherited in a same allele (Figure 3A), in different alleles (Figure 3B) or in different genes (Figure 3C). The produced heterodimer protein complexes included either one VUS (**III, Fig 1a**), or a VUS pair (**III, Fig 1b**), in one of the partners (MSH2 or MSH6) together with its wild type (WT) partner (MSH6 or MSH2, respectively) or one VUS in both partners (MSH2 and MSH6) (**III, Fig 1c**). The amount of variant protein total extract (TE) to be used in the *in vitro* MMR assay was determined by western blot analysis by adjusting the amount of its wild type heterodimerization partner in MutSα to be equal to that in the MutSα WT complex. Three types of pairs were tested, MSH2/MSH2, MSH2/MSH6 and MSH6/MSH6. Protein variant molecules were functionally analysed separately and together with the other VUS found in the same patient. The results of *in silico* predictions, expression stability in *Sf*9 cells and MMR capability of the studied variants and variant pairs, are shown in Table 7, in page 43. *In silico* alignment analyses assess each VUS individually, while the functional

analysis with the MMR assay allows the assessment of the potential concomitant effect caused by two VUS in a carrier.

By comparing the relative repair efficiencies (**III, Figure 2 e**), of the MSH2/MSH2 VUS pairs p.Asn127Ser/p.Ala328Pro, significantly decreased repair efficiency was demonstrated when compared to that of the MutSα WT complex. When individually assayed, p.Asn127Ser was able to correct the mismatches as WT (22%, STD ±2% and 19%, STD ±5%, respectively) (p=0.25), whereas the repair efficiency of p.Ala328Pro seems to be decreased (14%, STD ±3%), although not significantly (p=0.11) (**III, Figure 2 a, e**). A statistically significant decrease in repair efficiency was seen in experiments, when the effect of these VUS pairs were tested in the same molecule, MSH2 p.Asn127Ser-p.Ala328Pro (12%, STD ±4%) and WT (19%, STD ±5%) (p=0.04) (**III, Figure 2 a, e**), or in different molecules, MSH2 p.Asn127Ser+p.Ala328Pro (14%, STD ±4%) and WT (21%, STD ±1%) (p=0.02) (**III, Figure 2 d, e**), while keeping the total amount of complementing recombinant protein at the level of MutSα WT.

A plausible concomitant contribution to MMR deficiency can be suggested for the MSH2 pair p.Asn127Ser/p.Ala328Pro. Although when individually assayed with the optimal amount of *Sf*9 total extract, MSH2 p.Asn127Ser indicates proficiency, by halving its amount whilst maintaining the total recombinant MutSα amount in the assay at the same level as that of MutSα WT. MSH2 p.Asn127Ser cannot complement the deficiency caused by MSH2 p.Ala328Pro. Instead, their concomitant presence in the assay, either in same or different heterodimers, slightly increases the MMR deficiency. According to our previous experiments, when the MMR activity of MMR deficient extract is complemented with different amounts of WT extract, its optimal amount can be reduced at least by a factor of 10 without a notable reduction in the repair efficiency (Raevaara *et al.* 2003).

Thus, rather than haplo-insufficiency the reason for the concomitant deficiency of MSH2 p.Asn127Ser/p.Ala328Pro is a functional defect in both. Although, p.Asn127Ser is a rare variation, and is generally assessed as a nonpathogenic variation based on several different functional studies and healthy phenotype in many mutation carriers, it is among the most frequently reported VUS in CRC (Hampel *et al.* 2006; Ollila *et al.* 2008a). Thus, together with the previous data, the present study implies an extremely subtle MMR defect, which may not dominantly predispose to cancer but together with another inherited MMR gene variation like with a truncating mutation in *MSH2* (c.1264G>T, Glu422Stop) as was reported by Tanyi's work (Tanyi *et al.* 2008), the *MSH2* c.380A>G (p.Asn127Ser) seems to increase the cancer risk. In fact, Tanyi and colleagues demonstrated that this could decrease the age of cancer onset into the early thirties.

The original report of the CRC patient carrying the *MSH2* variations c.380A>G and c.982G>C (Samowitz *et al.* 2001), unfortunately does not show the data of the other mutation carriers and cancers in the family to reveal if the latter variation could already alone predispose to cancer. The amino acid change in VUS *MSH2* c.982G>C is predicted to be deleterious by Polyphen and MAPP-MMR alignment analyses and neutral by SIFT, while VUS *MSH2* c.380A>G is predicted deleterious by Polyphen and SIFT, but not MAPP-MMR.

The mutations were mapped to the structure of the MutSα complex (Warren *et al.* 2007). MSH2 and MSH6 have a similar domain architecture, consisting of 1) a mismatch binding domain (1-125/362-519), 2) a connector domain (125-300//519-718), 3) a lever domain (300-347, 554-620/718-935, 1009-1076), 4) a clamp (457-554/935-1009), and 5) an ATPase domain (620-934/1076-1360) (Warren *et al.* 2007). The substitution p.Ala328Pro maps to a long helix in the lever domain (**III; Figure 4**). The substitution p.Asn127Ser is located in the hinge between domains 1 and 2. The distance between the substituted amino acids is fairly long (24 Å), and therefore their non-additive effect is due to indirect interactions. The overall structure of the MutSα complex is an oval with allosteric communication between the DNA and ATP binding sites (Warren *et al.* 2007). Crystal structures with multiple substrates and normal mode analysis suggest that conserved domain motion is important for allostery (Warren *et al.* 2007; Mukherjee *et al.* 2009). More particularly, domains 3 and 5 move together as a unit while domain 1 moves a lot during the catalytic cycle (cf. Figure 7 in (Warren *et al.* 2007)). The introduction of a proline into a helix by the substitution p.Ala328Pro is expected to cause a kink in the helix of the lever domain. This could affect the allosteric communication between the DNA and ATP binding domains, as the p.Ala328Pro mutation alone was shown to mildly impair activity. The substitution p.Asn127Ser alone in the hinge did not impair activity, so we may assume that all rotation states of domain 1 during the catalytic cycle remain accessible. The energy landscape, though, may be altered due to perturbations of the hydrogen bonds formed by asparagine versus serine side chains. It is not impossible that a different transition path may be favoured in the presence of p.Asn127Ser. The non-additive impairment seen in the double mutant would also then be explained by blockage of the alternative transition path between rotation states for domain 1 in the presence of p.Ala328Pro. However, as long as the knowledge about the interactions and function of MutSα heterodimer molecules (one or more) with other players in the repair complex is under debate, this kind of concomitant contribution cannot be verified by biochemical experiments and rather serves as an example of how tricky the interpretation of the pathogenicity *in vitro* can be.

When the ultimate aim in clinical work is to obtain a classification of the MMR VUS based on probability of being pathogenic as was proposed by using five probability classes from definitely pathogenic, to not pathogenic, or of no clinical significance (Plon *et al.* 2008), even small differences in repair capability such as is seen between MSH2 p.Ala328Pro alone and together with p.Asn127Ser become important, even if the validation assays and their cut offs for decision making have not yet been determined. Although the critical level of needed repair capability *in vivo* depends on the circumstances, the concomitant defect of MSH2 p.Asn127Ser/p.Ala328Pro, which nearly halves the repair capability of MutSα WT, is most probably a cause of pathogenicity in the carrier. Especially interesting is the role of *MSH2* c.380A>G, whose effect leading to even slightly increased cancer risk could finally explain discrepancies in its repeatedly analyzed proficiency albeit the frequent occurrence in CRC patients.

By comparing the relative repair efficiencies (Table 7; **III, Figure 2 e**) of the other assayed MSH2/MSH2 VUS pairs, p.Gly322Asp/p.Asp487Glu also shows a significantly decreased

repair efficiency, whereas MSH2 p.Glu205Gln/p.Val367Ile do not interfere with repair capability (Table 7; **III, Figure 2 a**) when compared to that of the MutSα WT complex. The decrease in repair efficiencies was significant when the MSH2 p.Gly322Asp and p.Asp487Glu proteins were tested individually (10%, STD ±1% and 10%, STD ± 2%, respectively) (p=0.02), and as a pair in the same molecule MSH2 p.Gly322Asp-p.Asp487Glu (8%, STD ±3%) (p=0.02) (**III, Figure 2 a, e**) or in different molecules MSH2 p.Gly322Asp+p.Asp487Glu (12%, STD ±3%) (p=0.004) (**III, Figure 2 d, e**). Thus, the pair p.Gly322Asp/p.Asp487Glu differs from MSH2 p.Asn127Ser/p.Ala328Pro, which shows significantly decreased repair capability only as a VUS pair. Although, different locations of the two VUS (either in the same or different molecules) did not significantly affect their repair efficiency, the repair efficiencies of MSH2 p.Asn127Ser/p.Ala328Pro and MSH2 p.Gly322Asp/p.Asp487Glu were lowest when constructed into the same allele suggesting stronger impairment on MSH2, MutSα, and its function in a repair complex.

Variation p.Gly322Asp is among the most frequently reported VUS in CRC (Hampel *et al.* 2006; Ollila *et al.* 2008a), although the majority of the published data discusses p.Gly322Asp as a neutral polymorphism (Ollila *et al.* 2008a; Martinez and Kolodner 2010), it has also been hypothesized to be a low penetrance allele, supported by the functional analyses conducted with yeast assays (Drotschmann *et al.* 1999; Ellison *et al.* 2001). Irrespective of our previous study, where the purified p.Gly322Asp variant did not show MMR deficiency (Ollila *et al.* 2008a), and the recent studies, where enhancer screens with the yeast homolog msh2 p.Gly317Asp did not yield enhancer mutations and mouse ECS cells did not show a phenotype defect (Martinez and Kolodner 2010; Wielders *et al.* 2011), here the protein extract with over expressed MSH2 p.Gly322Asp individually and as a pair with p.Asp487Glu shows a statistically significant decrease in repair efficiency. It is possible that the purification process has at least partly excluded structurally damaged heterodimers suggesting that the *in vitro* MMR assay performed with total extract is more reliable. Overall, it seems that results vary a lot dependent on the assay and assay path used for pathogenicity assessments. Recent quite comprehensive study of the effect of MSH2 p.Gly322Asp on MMR in mouse embryonic stem cells (ESC) demonstrated that expressed from its endogenous locus it behaved like wild-type MSH2 (Wielders *et al.* 2011). The critical difference between that and the present study is that the functionality of MSH2 variant is assessed in undifferentiated stem cells or in human cell extracts originated from differentiated cancer cells, respectively. It is still unexplained why in Lynch syndrome a constitutional heterozygous MMR gene mutation predisposes to cancers only in some specific tissues or how a human embryo carrying a homozygous MMR gene mutation can succeed through all replications and recombinations occurring in several cell divisions and not lead to serious consequences until the first or second decade of life. Our results suggest that the *in vitro* MMR assay performed in human based system and thus detecting defective protein function in its own environment and repair machinery may reveal problems not detectable in all other assay models.

Pathogenic MSH6 VUS (II, III)

Although the other studied VUS pairs do not suggest concomitant contribution to MMR deficiency, which as such represents important information to the families, three studied *MSH6* VUS, p.Leu435Pro, p.Leu585Pro, and p.Ser1188Asn, which were previously functionally uncharacterized, were found to be unfunctional in the *in vitro* MMR assay, whereas their partners were fully functional (Table 7; **II, Figure 2b; III, Figure 2c**). Furthermore, *in silico* analysis by Polyphen and SIFT predicts pathogenicity of these *MSH6* variations (Table 7). Consequently, the repair capabilities of the pairs, MSH6 p.Leu435Pro-p.Val878Asn, MSH6 p.Leu585Pro-p.Ser677Thr and MSH2/MSH6 p.Val923Glu-p.Ser1188Asn, including both VUS in the same molecule, are also unfunctional (Table 7; **II, Figure 2b; III, Figure 2b, d, e**). Since the western blot analysis of the MMR deficient proteins p.Leu435Pro and p.Leu585Pro further revealed their instability (Table 7; **III, Figure 3**), the *in vitro* MMR assay was repeated by using an exaggerated amount of the proteins (data not shown). Nevertheless, both variants remained MMR deficient. It was previously observed that a leucine to proline change may affect protein stability and/or function. This is especially true when such change occurs in a MutSα- connector or mismatch binding site as here is the case. Furthermore, it has been reported that substituting leucine to proline in MSH2 leads to protein instability (Ollila *et al.* 2008b). Our results are also compatible with immunohistochemical staining of the tumor showing a lack of MSH6, but presence of MSH2 and MLH1 (Table 3; Figure 5F/G). Hence the assessment of these two VUS as pathogenic mutations is based on several observations: the lack of the MSH6 protein in the tumor tissue, instability of the mutated protein in *Sf9* expression, and MMR deficiency in the *in vitro* MMR assay, as well as deficiency supported by the *in silico* analyses. Furthermore, MSH6 p.Leu435Pro has previously been shown to skip exon 4, already indicating its pathogenicity (Hampel *et al.* 2007). Therefore, there are no additive effects seen when these two mutations are paired with other VUS in MSH6 since they sufficiently disrupt MMR function independently.

Other MSH2/MSH6 variants in pairs, including the novel pair p.Lys909Ile/p.Val878Ala, showed no significant decrease in the repair capability, nor any evidence of compound contribution to MMR deficiency (Table 7; **III, Figure 2 b, d, e**). Production of MSH2 p.Val923Glu and MSH6 p.Ser1188Asn and their pair p.Val923Glu/p.Ser1188Asn was successful in *Sf9* insect cells indicating the stability (Table 7; **II, Figure 2a**). The functional analysis of the two VUS, however, *MSH2* (c.2768T>A, p.Val923Glu) and *MSH6* (c.3563G>A, p.Ser1188Asn) revealed that only p.Ser1188Asn was pathogenic in the assay. The MSH2 p.Val923Glu variant was previously found to be MMR proficient in the *in vitro* MMR assay, and neutral based on a SIFT (Ollila *et al.* 2006), but deleterious in Polyphen (Table 7). Later, the MSH2 p.Val923Glu variant was, however, suggested to have slightly reduced mismatch binding and release capacity compared to the wild type MSH2 protein (Ollila *et al.* 2008b). Here, our results confirm the previous findings and suggest that *MSH6* VUS is the pathogenic mutation in the family. The C-terminal part of the MSH6 polypeptide consists of an evolutionarily highly conserved ABC-ATPase domain between amino acids

1076-1360. In this sequence, amino acids 1180-1186 form a disordered loop structure, which may play a crucial role in ATP binding (Warren *et al.* 2007). Two cancer associated mutations have been reported in the vicinity of that region, the truncating mutation *MSH6* c.3558_3565delTGAAAGTA, p.Gly1186fsX1190, which was detected in a CRC patient at the age of 27 (Pinto *et al.* 2006), and the missense variation, *MSH6* c.3577G>A, p.Glu1193Lys, which was identified in two endometrial cancer patients with late age of onset (59 and 60 years) and poor family history (Kariola *et al.* 2004). Remarkably, both the truncating and the missense variation turned out to be MMR deficient (Kariola *et al.* 2004; Pinto *et al.* 2006), indicating that the region between amino acids 1186-1193 is extremely important in repair function. By supplying evidence, that *MSH6* (c.3563G>A, p.Ser1188Asn) caused complete loss of protein function in the MMR reaction, our results reinforce the impression that this region in MSH6 is particularly important.

Results of mutation analyses as well as tumor analyses including IHC staining of MLH1, MSH2, MSH6 and PMS2 proteins, and MSI results are collected in Table 3. Although, the IHC analysis shows problems in MSH2 expression in two family members (Table 3; Figure 5E/a, b) carrying the variations *MSH2* (c.2768T>A, p.Val923Glu) and *MSH6* (c.3563G>A, p.Ser1188Asn), our results demonstrate that the expression problems of both MSH2 and MSH6 as well as the high MSI phenotype in one of the mutation carriers (Table 3; Figure 5E/b) is rather associated with MSH6 than MSH2 deficiency. Since, immunohistochemical analysis of MMR protein expression cannot distinguish between EPCAM deletion carriers and *MSH2* mutation carriers (Kloor *et al.* 2011), multiplex ligation dependent probe amplification analysis (MLPA) was performed in the carrier 5E/b to exclude a germline EPCAM deletion.

Previously, functionally studied *MSH6* VUS have been associated with low cancer susceptibility (Kariola *et al.* 2004). Here, the clinical features such as the late mean age of cancer onset (59.5 years), not completely lost but reduced or heterogenous expression of MSH6 in the tumor of the mutation carrier (5E/a), and loss of MSH6 expression in the carrier (5E/b), support the MSH6 predisposition. Although, the tumors of both mutation carriers showed expression deficiencies of MSH6 and MSH2, neither MSH6 p.Ser1188Asn nor MSH2 p.Val923Glu, showed expression problems in our *in vitro Sf*9 expression system (Table 7). The explanation for this discrepancy is most probably in the *in vitro* expression system, in which the protein is abundantly expressed under a strong virus promoter. Thus, successful MMR protein production in *Sf9* insect cells does not necessarily mean stability in the tumors and we cannot totally exclude the *MSH2* VUS contribution to cancer predisposition in these patients. However, its nonpathogenicity was supported by both the *in vitro* MMR and *in silico* analyses (Polyphen-2 and SIFT), recently found to be a reliable assay combination to verify pathogenicity/nonpathogenicity of an MMR VUS (Kansikas *et al.* 2011). Finally, in the fourth generation, there are several family members, who do not carry either VUS but still have colorectal adenomas (CRA) at a young age. The fact that other MMR gene mutations including large genomic rearrangements in *MLH1*, *MSH2,* and *MSH6,* were also excluded, and that all four MMR proteins were normally expressed in their

adenomas (Table 3; Figure 5E/c) suggests that predisposition to CRA in these individuals is associated with something other than MMR deficiency.

Here, it is clearly demonstrated that when LS is suspected, all the MMR susceptibility genes should be included in mutation analyses and all identified VUS should be functionally assessed. As was seen in this family (Table 3; Figure 5E), when multiple VUS are found in the genes *MSH2* and *MSH6*, both genes located on the same chromosome (chromosome 2), VUS may show a similar segregation pattern thus, complicate the interpretation.

The functional analysis of an MSH3 variation (I)

The *in vitro* MMR assay allows functional analysis of all kinds of missense variations in different MMR genes, if a suitable cell line, which lacks the analyzed MMR protein, is available. Since no LS predisposing mutations have been identified in *MSH3* thus far, the MMR assay was applied here for the first time to test the repair efficiency of an *MSH3* variation c.2386C>T (p.Arg796Trp) found in a putative LS patient (Table 3) (unpublished). By comparing the MSH3 variant´s repair efficiency against MSH3-WT´s capability (**I, Figure 3**), results were obtained to suggest that the variant MSH3-R796W is proficient (p=0.358) (Table 7). Remarkably, the assay itself functioned well, signifying its utility for further *MSH3* testing.

Table 7. The results of *in silico,* protein expression and MMR analyses.

VUS and VUS pairs		*In silico*[a]			Protein expression	*In vitro* MMR	Used HD[d]
		P	S	M			
MSH2/2	p.Asn127Ser	-	-	+	Normal[b]	Normal[b]	5'IDL1
	p.Ala328Pro	-	+	-	Normal	Normal	
	p.Asn127Ser-p.Ala328Pro	N/A	N/A	N/A	Normal	Decreased[c]	
	p.Asn127Ser+p.Ala328Pro	N/A	N/A	N/A	N/A	Decreased	
	p.Glu205Gln	+	+	+	Normal	Normal	
	p.Val367Ile	+	+	+	Normal	Normal	
	p.Glu205Gln-p.Val367Ile	N/A	N/A	N/A	Normal	Normal	
	p.Glu205Gln+p.Val367Ile	N/A	N/A	N/A	N/A	Normal	
	p.Gly322Asp	+	+	+	Normal	Decreased	
	p.Asp487Glu	+	+	+	Normal	Decreased	
	p.Gly322Asp-p.Asp487Glu	N/A	N/A	N/A	Normal	Decreased	
	p.Gly322Asp+p.Asp487Glu	N/A	N/A	N/A	N/A	Decreased	
MSH2/6	p.Ile145Met	+	+	-	Normal	Normal	5'IDL1
	p.Arg1095His	-	+	N/A	Normal	Normal	
	p.Leu1354Gln	-	+	N/A	Normal	Normal	
	p.Ile145Met-p.Arg1095His	N/A	N/A	N/A	Normal	Normal	
	p.Ile145Met+p.Arg1095His	N/A	N/A	N/A	N/A	Normal	
	p.Ile145Met-p.Leu1354Gln	N/A	N/A	N/A	Normal	Normal	
	p.Ile145Met+p.Leu1354Gln	N/A	N/A	N/A	N/A	Normal	
	p.Lys909Ile	-	-	+	Normal	Normal	
	p.Val878Ala	+	+	N/A	Normal	Normal	
	p.Lys909Ile-p.Val878Ala	N/A	N/A	N/A	Normal	Normal	
	p.Lys909Ile+p.Val878Ala	N/A	N/A	N/A	N/A	Normal	
	p.Val923Glu	-	+	N/A	Normal	Normal	5'GT and 5'IDL1
	p.Ser1188Asn	-	-	N/A	Normal	Decreased	
	p.Val923Glu-p.Ser1188Asn	N/A	N/A	N/A	Normal	Decreased	
MSH6/6	p.Leu435Pro	-	-	N/A	Decreased[c]	Decreased	5'IDL1
	p.Val878Ala	+	+	N/A	Normal	Normal	
	p.Leu435Pro-p.Val878Ala	N/A	N/A	N/A	Decreased	Decreased	
	p.Leu585Pro	-	-	N/A	Decreased	Decreased	
	p.Ser677Thr	+	+	N/A	Normal	Normal	
	p.Leu585Pro-p.Ser667Thr	N/A	N/A	N/A	Decreased	Decreased	
MSH3	p.Arg796Trp	N/A	N/A	N/A	Normal	Normal	5'IDL2

[a]Pathogenicity predictions by P, Polyphen-2; S, SIFT; M, MAPP-MMR; +, neutral; -, deleterious. [b]Mutated protein expressed/functioned as wild type. [c]Mutated protein expressed/functioned abnormally. [d]Used heteroduplex substrate in the *in vitro* MMR assay.

Substrate specificity and activity of MutS homologue wild type proteins and replicative polymerases in MMR (I, IV)

A strong role of MutSβ in dinucleotide loop repair (I, IV)

The *in vitro* MMR assay is able to demonstrate differences in substrate specificities, as well as repair efficiencies of MutSα and MutSβ, by using different substrate structures and cell lines. In this work, the MMR protein contents of the used cell lines, LoVo, GP5d, and HCT116, were first analysed by western blot (**I, Figure 1**). This confirmed the absence of MSH2, MSH6, and MSH3, in both LoVo and GP5d NEs, thus making them suitable for substrate specificity and functionality studies of MutSα and MutSβ. The level of PMS2 in GP5d NE was also decreased. HCT116 NE only expressed MSH6 and MSH2, presenting an opportunity to study the substrate specificity and repair efficiency of MutSβ and MutLα with it.

Here, together with different cell lines, three different substrates, 5'GT, 5'IDL1 (delA), and 5'IDL2 (delAT) were used to study MutSα and MutSβ heterodimer complexes. In contrast to MutLα, which is known to be essential to all these substrates, the most effective MutS complex (MutSα or MutSβ) varies depending on the mismatch. The MMR assays with LoVo, GP5d, and HCT116, including various combinations of natural or complemented MutLα, demonstrated that the role of MutSα is obvious in the repair of 5'GT and 5'IDL1 mismatches, whereas an efficient repair of dinucleotide loops requires MutSβ (**I, Figure 2a, b**). Remarkably, all the three tested cell lines demonstrate more efficient dinucleotide repair with MutSβ than with MutSα. In HCT116, the mean repair efficiency was 26% higher (p=0.0014), in LoVo 14% higher (p=0.284), and in GP5d 5% higher (p=0.230) with MutSβ than with MutSα.

The overlapping roles of these heterodimeric complexes have been previously reported (Acharya *et al.* 1996; Genschel *et al.* 1998) generally emphasizing the role of MutSα predominantly for the recognition of base/base mispairs and small IDLs, and MutSβ for the recognition of larger (>2bp) IDLs (Acharya *et al.* 1996; Palombo *et al.* 1996; Zhang *et al.* 2005). Here, the functional redundancy is supported, but contradictory to the previous impression that the repair efficiency of MutSβ exceeded that of MutSα in the repair of dinucleotide loop structures (Genschel *et al.* 1998; Zhang *et al.* 2005). Although, HCT116 expressed a sufficient amount of MutSα to repair GT and IDL1 mismatches, the repair efficiency of IDL2 was increased three fold when the cells were complemented with MutSβ (**I, Figure 2a, b**). The increase in repair efficiencies was also seen in LoVo and GP5d cells when complemented with MutSβ, but here the differences between MutSα and MutSβ were not statistically significant as in HCT116. Neutralization of MutSα (MSH6) by MSH6 antibody reduced the repair of 5'IDL1 by half (**IV, Figure 3A**), confirming the role of MSH6 in the repair of small IDLs (Genschel *et al.* 1998). In contrast, the repair of a dinucleotide loop (3'IDL2) was much less affected by this antibody (**IV, Figure 3B**) irrespective of the site of the nick. Overall, our functional and neutralization assays showed that MutSβ not only participated but exceeded the IDL2 repair efficiency of MutSα.

In cancer diagnostics, the positive MSI phenotype has been a hallmark of LS tumors. However, the levels have varied from high to low or no MSI, and between mono-, di-, tri-, and tetranucleotide repeat instability, dependent on the MMR gene affected, which might be linked to its substrate specificities. HCT116 cells, which are deficient in MLH1 and MSH3, have demonstrated mononucleotide repeat stability after complementation with *MLH1* through the addition of chromosome 3, but a low level of dinucleotide and a high level of tetranucleotide repeat instability remains, caused by deficiency in MSH3. Although, tetranucleotide repeat markers represented a five times higher level of instability than dinucleotide markers, supporting the functional overlap of MutSβ and MutSα in IDL2 repair. Low dinucleotide repeat instability was also caused by defective MutSβ (MSH3) (Haugen *et al.* 2008). The microsatellite instability was indeed reversible by complementing HCT116 cells with both chromosomes 3 and 5, hence expressing both lost proteins MLH1 and MSH3 (Haugen *et al.* 2008). Generally, the MSI marker panel (the Bethesda panel) used in diagnostics includes mono- and dinucleotide markers and in MSI positive cases, *MLH1, MSH2,* and *MSH6* genes are analyzed for mutations. Our results are clinically relevant, emphasizing the importance of MSH3 in dinucleotide loop repair and encouraging *MSH3* mutation analysis, especially when a tumor shows dinucleotide but no mononucleotide repeat instability.

α and ε polymerases have a minor role in mismatch repair in vitro (IV)

The aim of the work was to investigate interactions between MSH (MSH2, MSH3, MSH6) proteins and replicative polymerases (α, δ and ε) in HeLa cell extracts. Interactions were studied from hydroxylapatite-column (HA) fractions, these column separations were part of the purification process. Proteins from the polymerase peak fractions were analysed by mass spectrometry and western blotting (**IV; Figure 1A, B**). The catalytic subunit of DNA Pol δ and mismatch repair protein MSH2 were identified in certain fractions. Since PCNA was removed at an early step of the purification the observed co-purification is not mediated by PCNA (**IV; Figure S2B**). Analysis of the HA fractions revealed that MSH2 was present in highly purified Pol δ fractions at nearly stoichiometric amounts, whereas MSH6 and traces of MSH3, but no MSH2 was detected in Pol α peak fractions (**IV; Figure 1A**). Glycerol gradient centrifugation method (GGc) was used to further purify HA polymerase fractions which were analysed for the presence of mismatch repair proteins by WB (**IV; Figure S3**). After GGc treatment minor amounts of MSH2 co-sedimented with Pol α, while MSH6 migrated more slowly in the gradient, than the polymerase. In the GGc fraction, which includes Pol δ only MSH2 co-sedimented, but almost perfectly. Pol ε fraction was pure from all contaminating proteins, including MSH2 and MSH6. Physical interaction of MSH2, MSH3 and MSH6 proteins with replicative DNA polymerases was further confirmed by reciprocal immunoprecipitation (**IV; Figure 2A, B**).

The functional characterization of these interactions was performed in the *in vitro* MMR assay (**IV; Figure S3A, B**). Neutralization was achieved by incubating HeLa cell extracts

with specific antibodies. Neutralization with antibody against MSH6 (**IV; Table S1**) reduced the repair of 5'IDL1 by 50 % (**IV; Fig. 3A**). This confirms the role of MSH6 in repair of small IDLs (Genschel *et al.* 1998). In contrast, the repair efficiency of a dinucleotide insertion (3'IDL2) was decreased only by 10 % (**IV; Figure 3B**). Antibodies that specifically neutralise the activity of human Pols α and ε have been reported (Tanaka *et al.* 1982; Pospiech *et al.* 1999; Rytkönen *et al.* 2006). Neutralization of the polymerase activity of Pol α and Pol ε significantly inhibited the repair efficiency of the 5'IDL1 substrate by 13 % (p=0.016) and 17 % (p=0.018), respectively (**IV; Figure 3A**). The 5'IDL2 substrate was also inhibited to a similar extent (data not shown). We observed a similar inhibitory effect of Pol ε neutralization on the repair of the 3'IDL2 substrate, decreasing the efficiency of repair by 20 % compared to the complete reaction (p=0.035) (**IV; Figure 3B**), whereas neutralization of Pol α had no effect on the repair of this substrate. Unfortunately, neutralization of Pol δ was not successful (data not shown).

As neutralization of the polymerase activity of Pol α resulted only in small although reproducible decrease in MMR efficiency of 5'IDL1 substrate, the addition of neutralising antibodies would not be enough to completely abolish the polymerase activity. Therefore, the requirement of Pol α in human MMR was further investigated by an immunodepletion assay where Pol α was removed from the HeLa nuclear extract. The depletion was confirmed by WB, in which the amounts of other MMR associated proteins remained constant and β-Tubulin confirmed equal loading (**IV; Figure 4A**). In our *in vitro* MMR analysis HeLa NE depleted for Pol α showed a 10 % decrease (p=0.013) for the comparison of depleted versus complete NEs in the repair of the 5'IDL1 substrate (**IV; Figure 4B**). Moreover, the repair defect caused by Pol α depletion, could be completely reverted by supplementation of the extract with highly purified, recombinant Pol α (**IV; Figure 4B and C**). By using the 3'IDL2 substrate, we observed only a insignificant decrease in the repair efficiency (data not shown). The effects of Pol α depletion are in agreement with those of its neutralization. Taken together, the results of the MMR activity assay corroborate the requirement of aphidicolin-sensitive polymerases in the mismatch repair, as was shown previously (Holmes *et al.* 1990; Thomas *et al.* 1991; Fang and Modrich 1993). While inhibition of Pol α only affected the repair of the substrate with a 5' nick, Pol ε appears to contribute to the repair of both 5'IDL1 and 3'IDL2. The moderate effect of antibodies against Pols α and ε suggests only a minor or redundant role for these enzymes.

After protein interaction and functional studies, the nucleoprotein fraction was used to link MMR to the DNA replication apparatus. Highly purified nucleoprotein fractions from HeLa cells arrested in early S phase or released to progress to late S phase were prepared. The purified the early S phase nucleoprotein fraction contained the licensing factor and helicase component MCM3, replicative Pols α, δ and ε as well as mismatch recognition factors MSH2, MSH3 and MSH6 (**IV; Figure 5A**). The robust increase of MSH2, MSH3 and MSH6 was seen in fractions in the mid-late S phase and replicative Pols α and δ in nucleoprotein were found to moderately increased towards the mid-late S Phase (**IV; Figure 5A, B**). In addition, the level of MCM3 protein was shown to decrease in mid-late S phase (**IV; Figure 5A, B**). Pol ε shows a different, reproducible pattern: chromatin association of the form

moving faster in a gel increases during the S-phase while a form moving more slowly decreases.

The association of the MSH protein complex with components of the replication apparatus *in vivo* was characterized by chromatin immunoprecipitation (ChIP), which specifically precipitates the chromatin-associated MSH2 and MSH6 (**IV; Figure 5C**). ChIP was used to investigate whether components of the replication machinery were present in the MSH2- and MSH6-containing chromatin complexes. The results show that PCNA could be detected in the anti-MSH2 and anti-MSH6 precipitates both at early and late time points, although its signals in the late S Phase were repeatedly weaker (**IV; Figure 5B, C**). MCM3 also co-precipitated with both MSH2 and MSH6 in early and late S Phase. A minor MCM3 signal in control IPs was much weaker than the signal in the specific IP, suggesting that MCM3 is enriched in MSH2/6-containing, chromatin-associated protein complexes. In addition to PCNA and MCM3, we also detected small amounts of the replicative Pols α, δ and ε in the precipitates of nucleoprotein complexes immunoprecipitated by the antibody against MSH6, but not by the unspecific control IgG antibody (**IV; Figure 5C**). We also performed reciprocal immunoprecipitations against Pols α, δ and ε (**IV; Figure 5C, lower panels**). Western blot analyses of the immunoprecipitates for MSH2 and MSH6 confirmed the association between MSH2/6 and replicative DNA polymerases. A strong signal of MSH2 and MSH6 and a milder of MSH3 could be detected in the precipitates of all three replicative DNA polymerases.

Association of MSH2 and MSH6 with replicative DNA polymerases, led to determine if the MMR proteins are particularly enriched at origins sites of DNA replication during the S phase. Quantitative real time PCR analysis of chromatin precipitates for the enrichment of DNA of the well-characterised origins in the *LB2* promoter region and the *MCM4* upstream region suggest that MSH2 associates with the DNA when Orc2 and MCM3 are released from the origin, i.e. after origin firing (**IV; Figure 6**). MSH2, but not Orc2 and MCM3 remains associated with the replicated DNA regions at these loci for several hours. The association of MSH2 reflects well the observed overall increase of MSH proteins in the nucleoprotein in late S phase (**IV; Figure 5A**). Furthermore, for the two genomic regions investigated, the loading of MSH2 appears to follow the origin firing. Once associated with chromatin, MSH2 persists late into S phase at the sites where DNA replication has already been completed.

Immunofluorescence technique was used to gain further insight into the relationship between MMR and DNA replication, we studied the co-localization of MSH2 with respect to Pol δ and PCNA in HeLa cells during S phase. In early S phase HeLa cells both MSH2 and Pol δ display punctuate patterns which partly co-localize (**IV; Figure 7**). However, as S phase progresses the degree of co-localization decreases substantially, peaking after 6 hours. Immunofluorescence data support our results with chromatin immunoprecipitation (ChIP) showing decreasing interaction between MMR and replication machineries during S Phase.

The data analysis of the extracts and reciprocal precipitation suggests a direct interaction between Pol δ and MSH2, and also between Pol ε and MSH6. Furthermore, it can be excluded that the interactions described here were mediated by PCNA, since all precipitates

presented here were essentially free of PCNA. Altogether, immunoprecipitation studies from the nucleoprotein fractions support the results from the conventional precipitations (**IV; Figure 2A, 2B and 5C**) and suggest that MSH complexes are associated with components of the DNA replication apparatus including MCM3, PCNA and the replicative Pols α, δ, and ε, assembled on chromatin. All three DNA polymerases precipitated MSH proteins, while MSH2 and MSH6 only weakly precipitated Pols. It seems that only a small fraction of the chromatin-bound MSH proteins are associated with replicative polymerases and during the progression of S phase more MSH proteins are bound to the chromatin without being associated directly with active replication forks (**IV; Figure 5A and 5C**). Therefore, the direct interactions between MSH factors and replicative DNA polymerases observed here may provide an additional level of regulation of MMR during S phase for the cell.

Using our *in vitro* MMR assay, we were able to confirm that mismatch repair is strongly inhibited by aphidicolin. The moderate reduction in MMR activity after inhibition of Pols α and ε with neutralising antibodies imply that these DNA polymerases are not mandatory for MMR, but may increase the repair efficiency. This is in line with the findings of Longley's work (Longley *et al.* 1997) where the removal of human Pol δ from cell extracts prevented MMR. Pol ε contributed to the repair efficiency of the all three different substrates, while neutralisation of Pol α affected the repair of substrates with a 5'-nick only. One explanation for this kind of substrate specificity may be the fact that Pol α does not possess a 3'-proofreading exonuclease, whereas Pol δ and ε do. Therefore, it is possible that choice of DNA polymerase is not only determined at the re-synthesis step, but already by the mechanism of the excision and the location of the nick relative to the mismatch (Dzantiev *et al.* 2004; Guo *et al.* 2004; Constantin *et al.* 2005).

Current understanding of MMR has been largely derived from *in vitro* assays including reconstitution of strand-specific MMR from highly purified factors (Constantin *et al.* 2005; Zhang *et al.* 2005). However, these *in vitro* assays do not necessarily take into account the cellular context. MMR during the S phase is intertwined with DNA replication, and this aspect is not taken into account by commonly used *in vitro* MMR assays. Cell biological studies have shown that MMR proteins can be found at sites of DNA replication during S phase (Kleczkowska *et al.* 2001; Schröering *et al.* 2007). PCNA is the major candidate for the factor that couples MMR to the replication fork, since disruption of the PCNA-MSH3/6 interaction disrupts association with sites of DNA replication (Clark *et al.* 2000; Kleczkowska *et al.* 2001; Masih *et al.* 2008). Direct association of MSH2 and MSH6 with replicative DNA polymerases may provide an additional recruitment mechanism.

In proliferating cells, MMR proteins become increasingly associated with chromatin as S phase progresses (Schröering *et al.* 2007; Mastrocola and Heinen 2010). In this study, this was also reflected by the augmented level of MSH proteins in the nucleoprotein fraction of HeLa cells in late S phase compared to cells arrested in early S phase (**IV; Figure. 5A**). The analysis of the association of MSH2 with two specific and well-characterised origins of replication indicates that MSH2 is recruited to these sites after origin firing. After origin firing, ORC2 and MCM3 were released from the origins and from neighbouring regions, whereas MSH2 became enriched at these sites. In contrast, the replicative Pols α, δ and ε

became enriched at origin sequences during firing, and rapidly released thereafter, in the same way as MCM3 (communicated by H. Pospiech). Although the recruitment of MSH2 appeared to depend on active DNA replication, MSH2 apparently did not migrate with the replication fork. Rather, it was left behind on the replicated chromatin. In summary, results presented here, suggest the direct interactions between MSH proteins and replicative DNA polymerases described here represent a new, previously unnoticed link between DNA replication and mismatch repair.

CONCLUSIONS AND FUTURE PROSPECTS

Conclusion 1.

So far, no predisposing *MSH3* mutations have been found in Lynch Syndrome patients. The present study shows that in contrast to MutSα (MSH2/MSH6), the MutSβ (MSH2/MSH3) heterodimer has no role in mononucleotide loop repair but a strong role in dinucleotide loop repair, suggesting that *MSH3* mutation screening would be important when a tumor shows only dinucleotide MSI. Thus, the generally used Bethesda panel including both mono- and dinucleotide markers may not be the most suitable to diagnose *MSH3* mutation carriers.

Future prospects associated with conclusion 1:

• To clinically verify the observation by screening *MSH3* mutations especially from tumors showing instability only in dinucleotide repeats.

• To study the repair capability of MutSβ by using substrates with even longer than dinucleotide loops and accordingly its possible association with diseases due to expansion of tri- or tetranucleotide repeats.

Conclusion 2.

Our results show that two inherited MMR gene variations in a cancer patient may have a concomitant contribution to MMR deficiency. The role of the frequently reported MMR gene VUS *MSH2* c.380A>G is especially interesting, since its concomitant defect with another variant could finally explain its recurrent occurrence in CRC patients.

Future prospects associated with conclusion 2:

• Whenever LS is suspected, the all MMR susceptibility genes should be included in mutation analyses and the functional significance of all identified VUS should be analyzed.

• Since epidemiologically it might be of more significance to identify low risk variations with high prevalence than high risk rare variations in a population, it could be worth assessing the overall significance of *MSH2* c.380A>G in CRC risk according to Bayes´ theorem.

Conclusion 3.

The present functional study shows that Pols α and ε contribute to but are not essential for human MMR, and results suggest that MSH2 and MSH6 physically interact with Pols δ and Pol α, respectively. This is consistent with models where MSH proteins are continuously

loaded onto chromatin in a replication-dependent manner, and persist on DNA that has already completed replication.

Future prospects associated with conclusion 3:

• To study the Okazaki fragment formations and the participation of different polymerases with the *in vitro* MMR assay by using different types of substrates.

ACKNOWLEDGEMENTS

This work has been carried out during the years 2008-2012 at the Division of Genetics, Department of Biosciences, Faculty of Biological and Environmental Sciences, University of Helsinki. This work was supported by grants from: The Academy of Finland, the Sigrid Juselius Foundation, European Research Council, Finnish Cancer Organizations, The Biocentrum Helsinki Organization, The Research Foundation of the University of Helsinki and Kuopio Naturalists' Society. All the financial support is gratefully acknowledged.

My warmest thanks go to my supervisors Professor Minna Nyström and Docent Reetta Kariola. I would especially like to thank Minna Nyström for giving me the opportunity to work in an interesting research project and for supervision. Their support and encouragement during these years have been very important to me.

Docent Minna Pöyhönen and Docent Esa Kuismanen are acknowledged for reviewing and commenting on this thesis. Dr. Jennifer Rowland and MSc. Minttu Kansikas are acknowledged for skillful revision of the English language of this thesis.

I am thankful to all the people I have worked with over these years. Our research group has been a wonderful place to work and I thank all the former and present members of our research group: Minttu Kansikas, Reetta Kariola, Mari Korhonen, Satu Candelin, Saara Ollila, Laura Sarantaus, Marjaana Pussila, Satu Lowndes, and Denis Dermadi; they all are wonderful people and colleagues. Thank you all for your help, kindness and patience with me. It has been great working with you. Special thanks goes to Minttu Kansikas, Reetta Kariola, Mari Korhonen, Satu Candelin, and Saara Ollila who have helped me uncountable times and of course for their contributions to my articles. I am also very grateful to our collaborators around the world, especially for Docent Helmut Pospiech, MSc. Harri Itkonen and Professor Liisa Holm.

The Division of Genetics, led by Professor Tapio Palva, has provided excellent facilities for research. I thank the people at the Division of Genetics for the warm atmosphere there. Arja Ikävalko, Arja Välimäki, Merja Forsberg and Pekka Heino are acknowledged for their help in many general issues.

Eino Lindholm, to whom I am most grateful for being a true mentor and my best friend. Eino deserves lots of thanks for all the support and encouragement during these years. I want to thank him for the numerous e-mails and debates that have opened up the world to me. My friends outside the laboratory for support and encouragement and many joyful and relaxing moments together.

Olen kiitollinen vanhemmilleni Ritvalle ja Paulille. Erityisen kiitoksen tahdon osoittaa äidilleni Ritvalle rakkaudesta ja tuesta kaikkina hetkinä. Olen paljosta velkaa siskolleni Kaisalle ja pikkuveljelleni Pekalle tuesta, kannustuksesta ja hyväntahtoisuudesta minua kohtaan.

REFERENCES

Aaltonen LA, Peltomäki P, Leach FS, Sistonen P, Pylkkänen L, Mecklin JP, Järvinen H, Powell SM, Jen J, Hamilton SR. 1993. Clues to the pathogenesis of familial colorectal cancer. Science (New York, N.Y.) 260:812-816.

Acharya S, Wilson T, Gradia S, Kane MF, Guerrette S, Marsischky GT, Kolodner R, Fishel R. 1996. hMSH2 forms specific mispair-binding complexes with hMSH3 and hMSH6. Proceedings of the National Academy of Sciences of the United States of America 93:13629-13634.

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. 2010. A method and server for predicting damaging missense mutations. Nature methods 7:248-249.

Allen DJ, Makhov A, Grilley M, Taylor J, Thresher R, Modrich P, Griffith JD. 1997. MutS mediates heteroduplex loop formation by a translocation mechanism. The EMBO journal 16:4467-4476.

Alvino E, Castiglia D, Caporali S, Pepponi R, Caporaso P, Lacal PM, Marra G, Fischer F, Zambruno G, Bonmassar E, Jiricny J, D'Atri S. 2006. A single cycle of treatment with temozolomide, alone or combined with O(6)-benzylguanine, induces strong chemoresistance in melanoma cell clones in vitro: role of O(6)-methylguanine-DNA methyltransferase and the mismatch repair system. International journal of oncology 29:785-797.

Amin NS, Nguyen MN, Oh S, Kolodner RD. 2001. exo1-Dependent mutator mutations: model system for studying functional interactions in mismatch repair. Molecular and cellular biology 21:5142-5155.

Beard WA and Wilson SH. 2003. Structural insights into the origins of DNA polymerase fidelity. Structure (London, England : 1993) 11:489-496.

Bedeir A and Krasinskas AM. 2011. Molecular diagnostics of colorectal cancer. Archives of Pathology & Laboratory Medicine 135:578-587.

Bennett P. 2000. Demystified ... microsatellites. Molecular pathology 53:177-183.

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

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. 2000. The Protein Data Bank. Nucleic acids research 28:235-242.

Bhattacharyya NP, Ganesh A, Phear G, Richards B, Skandalis A, Meuth M. 1995. Molecular analysis of mutations in mutator colorectal carcinoma cell lines. Human molecular genetics 4:2057-2064.

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

Bowers J, Tran PT, Joshi A, Liskay RM, Alani E. 2001. MSH-MLH complexes formed at a DNA mismatch are disrupted by the PCNA sliding clamp. Journal of Molecular Biology 306:957-968.

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

Chao EC, Velasquez JL, Witherspoon MS, Rozek LS, Peel D, Ng P, Gruber SB, Watson P, Rennert G, Anton-Culver H, Lynch H, Lipkin SM. 2008. Accurate classification of MLH1/MSH2 missense variants with multivariate analysis of protein polymorphisms-mismatch repair (MAPP-MMR). Human mutation 29:852-860.

Chilkova O, Stenlund P, Isoz I, Stith CM, Grabowski P, Lundstrom EB, Burgers PM, Johansson E. 2007. The eukaryotic leading and lagging strand DNA polymerases are loaded onto primer-ends via separate mechanisms but have comparable processivity in the presence of PCNA. Nucleic acids research 35:6588-6597.

Christensen LL, Kariola R, Korhonen MK, Wikman FP, Sunde L, Gerdes AM, Okkels H, Brandt CA, Bernstein I, Hansen TV, Hagemann-Madsen R, Andersen CL, Nystrom M, Orntoft TF. 2009. Functional characterization of rare missense mutations in MLH1 and MSH2 identified in Danish colorectal cancer patients. Familial cancer 8:489-500.

Christensen LL, Madsen BE, Wikman FP, Wiuf C, Koed K, Tjonneland A, Olsen A, Syvänen AC, Andersen CL, Orntoft TF. 2008. The association between genetic variants in hMLH1 and hMSH2 and the development of sporadic colorectal cancer in the Danish population. BMC medical genetics 9:52.

Clark AB, Cook ME, Tran HT, Gordenin DA, Resnick MA, Kunkel TA. 1999. Functional analysis of human MutSalpha and MutSbeta complexes in yeast. Nucleic acids research 27:736-742.

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

Constantin N, Dzantiev L, Kadyrov FA, Modrich P. 2005. Human mismatch repair: reconstitution of a nick-directed bidirectional reaction. The Journal of biological chemistry 280:39752-39761.

Cotton RG and Scriver CR. 1998. Proof of "disease causing" mutation. Human mutation 12:1-3.

Couch FJ, Rasmussen LJ, Hofstra R, Monteiro AN, Greenblatt MS, de Wind N, IARC Unclassified Genetic Variants Working Group. 2008. Assessment of functional effects of unclassified genetic variants. Human mutation 29:1314-1326.

de la Chapelle A. 2004. Genetic predisposition to colorectal cancer. Nature reviews.Cancer 4:769-780.

de la Chapelle A and Hampel H. 2010. Clinical relevance of microsatellite instability in colorectal cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 28:3380-3387.

Dowen JM, Putnam CD, Kolodner RD. 2010. Functional studies and homology modeling of Msh2-Msh3 predict that mispair recognition involves DNA bending and strand separation. Molecular and cellular biology 30:3321-3328.

Drost M, Zonneveld JB, van Dijk L, Morreau H, Tops CM, Vasen HF, Wijnen JT, de Wind N. 2010. A cell-free assay for the functional analysis of variants of the mismatch repair protein MLH1. Human mutation 31:247-253.

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

Drummond JT, Genschel J, Wolf E, Modrich P. 1997. DHFR/MSH3 amplification in methotrexate-resistant cells alters the hMutSalpha/hMutSbeta ratio and reduces the efficiency of base-base mismatch repair. Proceedings of the National Academy of Sciences of the United States of America 94:10144-10149.

Drummond JT, Li GM, Longley MJ, Modrich P. 1995. Isolation of an hMSH2-p160 heterodimer that restores DNA mismatch repair to tumor cells. Science (New York, N.Y.) 268:1909-1912.

Dufner P, Marra G, Räschle M, 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. The Journal of biological chemistry 275:36550-36555.

Durno CA, Holter S, Sherman PM, Gallinger S. 2010. The gastrointestinal phenotype of germline biallelic mismatch repair gene mutations. The American Journal of Gastroenterology 105:2449-2456.

Dzantiev L, Constantin N, Genschel J, Iyer RR, Burgers PM, Modrich P. 2004. A defined human system that supports bidirectional mismatch-provoked excision. Molecular cell 15:31-41.

Ellegren H. 2004. Microsatellites: simple sequences with complex evolution. Nature reviews.Genetics 5:435-445.

Ellison AR, Lofing J, Bitter GA. 2001. Functional analysis of human MLH1 and MSH2 missense variants and hybrid human-yeast MLH1 proteins in Saccharomyces cerevisiae. Human molecular genetics 10:1889-1900.

Essers J, Theil AF, Baldeyron C, van Cappellen WA, Houtsmuller AB, Kanaar R, Vermeulen W. 2005. Nuclear dynamics of PCNA in DNA replication and repair. Molecular and cellular biology 25:9350-9359.

Fang WH and Modrich P. 1993. Human strand-specific mismatch repair occurs by a bidirectional mechanism similar to that of the bacterial reaction. The Journal of biological chemistry 268:11838-11844.

Felton KE, Gilchrist DM, Andrew SE. 2007. Constitutive deficiency in DNA mismatch repair. Clinical genetics 71:483-498.

Fukui T, Yamauchi K, Muroya T, Akiyama M, Maki H, Sugino A, Waga S. 2004. Distinct roles of DNA polymerases delta and epsilon at the replication fork in Xenopus egg extracts. Genes to cells : devoted to molecular & cellular mechanisms 9:179-191.

Gammie AE, Erdeniz N, Beaver J, Devlin B, Nanji A, Rose MD. 2007. Functional characterization of pathogenic human MSH2 missense mutations in Saccharomyces cerevisiae. Genetics 177:707-721.

Garg P and Burgers PM. 2005. DNA polymerases that propagate the eukaryotic DNA replication fork. Critical reviews in biochemistry and molecular biology 40:115-128.

Gargiulo S, Torrini M, Ollila S, Nasti S, Pastorino L, Cusano R, Bonelli L, Battistuzzi L, Mastracci L, Bruno W, Savarino V, Sciallero S, Borgonovo G, Nyström M, Bianchi-Scarra G, Mareni C, Ghiorzo P. 2009. Germline MLH1 and MSH2 mutations in Italian pancreatic cancer patients with suspected Lynch syndrome. Familial cancer 8:547-553.

Genschel J, Littman SJ, Drummond JT, Modrich P. 1998. Isolation of MutSbeta from human cells and comparison of the mismatch repair specificities of MutSbeta and MutSalpha. The Journal of biological chemistry 273:19895-19901.

Genschel J and Modrich P. 2003. Mechanism of 5'-directed excision in human mismatch repair. Molecular cell 12:1077-1086.

Goldgar DE, Easton DF, Byrnes GB, Spurdle AB, Iversen ES, Greenblatt MS, IARC Unclassified Genetic Variants Working Group. 2008. Genetic evidence and integration of various data sources for classifying uncertain variants into a single model. Human mutation 29:1265-1272.

Gradia S, Acharya S, Fishel R. 1997. 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, Fishel R. 1999. hMSH2-hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA. Molecular cell 3:255-261.

Gu L, Hong Y, McCulloch S, Watanabe H, Li GM. 1998. ATP-dependent interaction of human mismatch repair proteins and dual role of PCNA in mismatch repair. Nucleic acids research 26:1173-1178.

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

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

Guo S, Presnell SR, Yuan F, Zhang Y, Gu L, Li GM. 2004. Differential requirement for proliferating cell nuclear antigen in 5' and 3' nick-directed excision in human mismatch repair. The Journal of biological chemistry 279:16912-16917.

Guo S, Zhang Y, Yuan F, Gao Y, Gu L, Wong I, Li GM. 2006. Regulation of replication protein A functions in DNA mismatch repair by phosphorylation. The Journal of biological chemistry 281:21607-21616.

Hackman P, Tannergard P, Osei-Mensa S, Chen J, Kane MF, Kolodner R, Lambert B, Hellgren D, Lindblom A. 1997. A human compound heterozygote for two MLH1 missense mutations. Nature genetics 17:135-136.

Hampel H, Frankel W, Panescu J, Lockman J, Sotamaa K, Fix D, Comeras I, La Jeunesse J, Nakagawa H, Westman JA, Prior TW, Clendenning M, Penzone P, Lombardi J, Dunn P, Cohn DE, Copeland L, Eaton L, Fowler J, Lewandowski G, Vaccarello L, Bell J, Reid G, de la Chapelle A. 2006. Screening for Lynch syndrome (hereditary nonpolyposis colorectal cancer) among endometrial cancer patients. Cancer research 66:7810-7817.

Hampel H, Frankel WL, Martin E, Arnold M, Khanduja K, Kuebler P, Clendenning M, Sotamaa K, Prior T, Westman JA, Panescu J, Fix D, Lockman J, LaJeunesse J, Comeras I, de la Chapelle A. 2008. Feasibility of screening for Lynch syndrome among patients with colorectal cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 26:5783-5788.

Hampel H, Frankel WL, Martin E, Arnold M, Khanduja K, Kuebler P, Nakagawa H, Sotamaa K, Prior TW, Westman J, Panescu J, Fix D, Lockman J, Comeras I, de la Chapelle A. 2005. Screening for the Lynch syndrome (hereditary nonpolyposis colorectal cancer). The New England journal of medicine 352:1851-1860.

Hampel H, Panescu J, Lockman J, Sotamaa K, Fix D, Comeras I, LaJeunesse J, Nakagawa H, Westman JA, Prior TW, Clendenning M, de la Chapelle A, Frankel W, Penzone P, Cohn DE, Copeland L, Eaton L, Fowler J, Lombardi J, Dunn P, Bell J, Reid G, Lewandowski G, Vaccarello L. 2007. Comment on: Screening for Lynch Syndrome (Hereditary Nonpolyposis Colorectal Cancer) among Endometrial Cancer Patients. Cancer research 67:9603.

Harfe BD and Jinks-Robertson S. 2000. DNA mismatch repair and genetic instability. Annual Review of Genetics 34:359-399.

Haugen AC, Goel A, Yamada K, Marra G, Nguyen TP, Nagasaka T, Kanazawa S, Koike J, Kikuchi Y, Zhong X, Arita M, Shibuya K, Oshimura M, Hemmi H, Boland CR, Koi M. 2008. Genetic instability caused by loss of MutS homologue 3 in human colorectal cancer. Cancer research 68:8465-8472.

Heinen CD, Goss KH, Cornelius JR, Babcock GF, Knudsen ES, Kowalik T, Groden J. 2002. The APC tumor suppressor controls entry into S-phase through its ability to regulate the cyclin D/RB pathway. Gastroenterology 123:751-763.

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

Holm L and Rosenström P. 2010. Dali server: conservation mapping in 3D. Nucleic acids research 38:W545-9.

Holmes J,Jr, Clark S, Modrich P. 1990. Strand-specific mismatch correction in nuclear extracts of human and Drosophila melanogaster cell lines. Proceedings of the National Academy of Sciences of the United States of America 87:5837-5841.

Hsieh P and Yamane K. 2008. DNA mismatch repair: molecular mechanism, cancer, and ageing. Mechanisms of ageing and development 129:391-407.

Hubscher U, Maga G, Spadari S. 2002. Eukaryotic DNA polymerases. Annual Review of Biochemistry 71:133-163.

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

Iyer RR, Pluciennik A, Genschel J, Tsai MS, Beese LS, Modrich P. 2010. MutLalpha and proliferating cell nuclear antigen share binding sites on MutSbeta. The Journal of biological chemistry 285:11730-11739.

Jackson CC, Holter S, Pollett A, Clendenning M, Chou S, Senter L, Ramphal R, Gallinger S, Boycott K. 2008. Cafe-au-lait macules and pediatric malignancy caused by biallelic mutations in the DNA mismatch repair (MMR) gene PMS2. Pediatric blood & cancer 50:1268-1270.

Jin YH, Ayyagari R, Resnick MA, Gordenin DA, Burgers PM. 2003. Okazaki fragment maturation in yeast. II. Cooperation between the polymerase and 3'-5'-exonuclease activities of Pol delta in the creation of a ligatable nick. The Journal of biological chemistry 278:1626-1633.

Jin YH, Obert R, Burgers PM, Kunkel TA, Resnick MA, Gordenin DA. 2001. The 3'-->5' exonuclease of DNA polymerase delta can substitute for the 5' flap endonuclease Rad27/Fen1 in processing Okazaki fragments and preventing genome instability. Proceedings of the National Academy of Sciences of the United States of America 98:5122-5127.

Jiricny J. 2006a. The multifaceted mismatch-repair system. Nature reviews. Molecular cell biology 7:335-346.

Jiricny J. 2006b. MutLalpha: at the cutting edge of mismatch repair. Cell 126:239-241.

Jiricny J and Nyström-Lahti M. 2000. Mismatch repair defects in cancer. Current opinion in genetics & development 10:157-161.

Junop MS, Obmolova G, Rausch K, Hsieh P, Yang W. 2001. Composite active site of an ABC ATPase: MutS uses ATP to verify mismatch recognition and authorize DNA repair. Molecular cell 7:1-12.

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

Kadyrov FA, Holmes SF, Arana ME, Lukianova OA, O'Donnell M, Kunkel TA, Modrich P. 2007. Saccharomyces cerevisiae MutLalpha is a mismatch repair endonuclease. The Journal of biological chemistry 282:37181-37190.

Kansikas M, Kariola R, Nyström M. 2011. Verification of the three-step model in assessing the pathogenicity of mismatch repair gene variants. Human mutation 32:107-115.

Kariola R, Hampel H, Frankel WL, Raevaara TE, de la Chapelle A, Nyström-Lahti M. 2004. MSH6 missense mutations are often associated with no or low cancer susceptibility. British journal of cancer 91:1287-1292.

Kariola R, Otway R, Lönnqvist KE, Raevaara TE, Macrae F, Vos YJ, Kohonen-Corish M, Hofstra RM, Nyström-Lahti M. 2003. Two mismatch repair gene mutations found in a colon cancer patient--which one is pathogenic? Human genetics 112:105-109.

Kariola R, Raevaara TE, Lönnqvist KE, Nyström-Lahti M. 2002. Functional analysis of MSH6 mutations linked to kindreds with putative hereditary non-polyposis colorectal cancer syndrome. Human molecular genetics 11:1303-1310.

Kempers MJ, Kuiper RP, Ockeloen CW, Chappuis PO, Hutter P, Rahner N, Schackert HK, Steinke V, Holinski-Feder E, Morak M, Kloor M, Buttner R, Verwiel ET, van Krieken JH, Nagtegaal ID, Goossens M, van der Post RS, Niessen RC, Sijmons RH, Kluijt I, Hogervorst FB, Leter EM, Gille JJ, Aalfs CM, Redeker EJ, Hes FJ, Tops CM, van Nesselrooij BP, van Gijn ME, Gomez Garcia EB, Eccles DM, Bunyan DJ, Syngal S, Stoffel EM, Culver JO, Palomares MR, Graham T, Velsher L, Papp J, Olah E, Chan TL, Leung SY, van Kessel AG, Kiemeney LA, Hoogerbrugge N, Ligtenberg MJ. 2011. Risk of colorectal and endometrial cancers in EPCAM deletion-positive Lynch syndrome: a cohort study. The lancet oncology 12:49-55.

Kleczkowska HE, Marra G, Lettieri T, Jiricny J. 2001. hMSH3 and hMSH6 interact with PCNA and colocalize with it to replication foci. Genes & development 15:724-736.

Kloor M, Voigt AY, Schackert HK, Schirmacher P, von Knebel Doeberitz M, Blaker H. 2011. Analysis of EPCAM protein expression in diagnostics of Lynch syndrome. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 29:223-227.

Knudson AG,Jr. 1971. Mutation and cancer: statistical study of retinoblastoma. Proceedings of the National Academy of Sciences of the United States of America 68:820-823.

Kolodner RD and Marsischky GT. 1999. Eukaryotic DNA mismatch repair. Current opinion in genetics & development 9:89-96.

Kondo E, Suzuki H, Horii A, Fukushige S. 2003. A yeast two-hybrid assay provides a simple way to evaluate the vast majority of hMLH1 germ-line mutations. Cancer research 63:3302-3308.

Korhonen MK, Vuorenmaa E, Nyström M. 2008. The first functional study of MLH3 mutations found in cancer patients. Genes, chromosomes & cancer 47:803-809.

Kovacs ME, Papp J, Szentirmay Z, Otto S, Olah E. 2009. Deletions removing the last exon of TACSTD1 constitute a distinct class of mutations predisposing to Lynch syndrome. Human mutation 30:197-203.

Kunkel TA and Bebenek K. 2000. DNA replication fidelity. Annual Review of Biochemistry 69:497-529.

Lahue RS, Au KG, Modrich P. 1989. DNA mismatch correction in a defined system. Science (New York, N.Y.) 245:160-164.

Lamers MH, Perrakis A, Enzlin JH, Winterwerp HH, de Wind N, Sixma TK. 2000. The crystal structure of DNA mismatch repair protein MutS binding to a G x T mismatch. Nature 407:711-717.

Lau PJ and Kolodner RD. 2003. Transfer of the MSH2.MSH6 complex from proliferating cell nuclear antigen to mispaired bases in DNA. The Journal of biological chemistry 278:14-17.

Leach FS, Nicolaides NC, Papadopoulos N, Liu B, Jen J, Parsons R, Peltomäki P, Sistonen P, Aaltonen LA, Nyström-Lahti M. 1993. Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. Cell 75:1215-1225.

Lee SC, Guo JY, Lim R, Soo R, Koay E, Salto-Tellez M, Leong A, Goh BC. 2005. Clinical and molecular characteristics of hereditary non-polyposis colorectal cancer families in Southeast Asia. Clinical genetics 68:137-145.

Lee SD, Surtees JA, Alani E. 2007. Saccharomyces cerevisiae MSH2-MSH3 and MSH2-MSH6 complexes display distinct requirements for DNA binding domain I in mismatch recognition. Journal of Molecular Biology 366:53-66.

Lehman IR and Kaguni LS. 1989. DNA polymerase alpha. The Journal of biological chemistry 264:4265-4268.

Li GM. 2008. Mechanisms and functions of DNA mismatch repair. Cell research 18:85-98.

Ligtenberg MJ, Kuiper RP, Chan TL, Goossens M, Hebeda KM, Voorendt M, Lee TY, Bodmer D, Hoenselaar E, Hendriks-Cornelissen SJ, Tsui WY, Kong CK, Brunner HG, van Kessel AG, Yuen ST, van Krieken JH, Leung SY, Hoogerbrugge N. 2009. Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3' exons of TACSTD1. Nature genetics 41:112-117.

Lindblom A, Tannergard P, Werelius B, Nordenskjöld M. 1993. Genetic mapping of a second locus predisposing to hereditary non-polyposis colon cancer. Nature genetics 5:279-282.

Liu HX, Zhou XL, Liu T, Werelius B, Lindmark G, Dahl N, Lindblom A. 2003. The role of hMLH3 in familial colorectal cancer. Cancer research 63:1894-1899.

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

Longley MJ, Pierce AJ, Modrich P. 1997. DNA polymerase delta is required for human mismatch repair in vitro. The Journal of biological chemistry 272:10917-10921.

Lu AL, Clark S, Modrich P. 2005. Methyl-directed repair of DNA base-pair mismatches in vitro. 1983. DNA repair 4:134-8; discussion 133, 139.

Lu AL, Clark S, Modrich P. 1983. Methyl-directed repair of DNA base-pair mismatches in vitro. Proceedings of the National Academy of Sciences of the United States of America 80:4639-4643.

Lynch HT, Lynch JF, Lynch PM, Attard T. 2008. Hereditary colorectal cancer syndromes: molecular genetics, genetic counseling, diagnosis and management. Familial cancer 7:27-39.

Lynch HT, Lynch JF, Snyder CL, Riegert-Johnson D. 2011. EPCAM deletions, Lynch syndrome, and cancer risk. The lancet oncology 12:5-6.

Lynch HT, Lynch PM, Lanspa SJ, Snyder CL, Lynch JF, Boland CR. 2009. Review of the Lynch syndrome: history, molecular genetics, screening, differential diagnosis, and medicolegal ramifications. Clinical genetics 76:1-18.

Marra G, Iaccarino I, Lettieri T, Roscilli G, Delmastro P, Jiricny J. 1998. Mismatch repair deficiency associated with overexpression of the MSH3 gene. Proceedings of the National Academy of Sciences of the United States of America 95:8568-8573.

Martinez SL and Kolodner RD. 2010. Functional analysis of human mismatch repair gene mutations identifies weak alleles and polymorphisms capable of polygenic interactions. Proceedings of the National Academy of Sciences of the United States of America 107:5070-5075.

Masih PJ, Kunnev D, Melendy T. 2008. Mismatch Repair proteins are recruited to replicating DNA through interaction with Proliferating Cell Nuclear Antigen (PCNA). Nucleic acids research 36:67-75.

Mastrocola AS and Heinen CD. 2010. Nuclear reorganization of DNA mismatch repair proteins in response to DNA damage. DNA repair 9:120-133.

McCulloch SD and Kunkel TA. 2008. The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. Cell research 18:148-161.

Mukherjee S, Law SM, Feig M. 2009. Deciphering the mismatch recognition cycle in MutS and MSH2-MSH6 using normal-mode analysis. Biophysical journal 96:1707-1720.

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

Ng PC and Henikoff S. 2001. Predicting deleterious amino acid substitutions. Genome research 11:863-874.

Nick McElhinny SA, Stith CM, Burgers PM, Kunkel TA. 2007. Inefficient proofreading and biased error rates during inaccurate DNA synthesis by a mutant derivative of Saccharomyces cerevisiae DNA polymerase delta. The Journal of biological chemistry 282:2324-2332.

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

Niessen RC, Hofstra RM, Westers H, Ligtenberg MJ, Kooi K, Jager PO, de Groote ML, Dijkhuizen T, Olderode-Berends MJ, Hollema H, Kleibeuker JH, Sijmons RH. 2009. Germline hypermethylation of MLH1 and EPCAM deletions are a frequent cause of Lynch syndrome. Genes, chromosomes & cancer 48:737-744.

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

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

Ollila S, Dermadi Bebek D, Greenblatt M, Nyström M. 2008a. Uncertain pathogenicity of MSH2 variants N127S and G322D challenges their classification. International journal of cancer.Journal international du cancer 123:720-724.

Ollila S, Dermadi Bebek D, Jiricny J, Nyström M. 2008b. Mechanisms of pathogenicity in human MSH2 missense mutants. Human mutation 29:1355-1363.

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. 2006. Pathogenicity of MSH2 missense mutations is typically associated with impaired repair capability of the mutated protein. Gastroenterology 131:1408-1417.

Ou J, Niessen RC, Lutzen A, Sijmons RH, Kleibeuker JH, de Wind N, Rasmussen LJ, Hofstra RM. 2007. Functional analysis helps to clarify the clinical importance of unclassified variants in DNA mismatch repair genes. Human mutation 28:1047-1054.

Ou J, Niessen RC, Vonk J, Westers H, Hofstra RM, Sijmons RH. 2008. A database to support the interpretation of human mismatch repair gene variants. Human mutation 29:1337-1341.

Owen BA, H Lang W, McMurray CT. 2009. The nucleotide binding dynamics of human MSH2-MSH3 are lesion dependent. Nature structural & molecular biology 16:550-557.

Palombo F, Gallinari P, Iaccarino I, Lettieri T, Hughes M, D'Arrigo A, Truong O, Hsuan JJ, Jiricny J. 1995. GTBP, a 160-kilodalton protein essential for mismatch-binding activity in human cells. Science (New York, N.Y.) 268:1912-1914.

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

Papadopoulos N, Leach FS, Kinzler KW, Vogelstein B. 1995. Monoallelic mutation analysis (MAMA) for identifying germline mutations. Nature genetics 11:99-102.

Papadopoulos N, Nicolaides NC, Wei YF, Ruben SM, Carter KC, Rosen CA, Haseltine WA, Fleischmann RD, Fraser CM, Adams MD. 1994. Mutation of a mutL homolog in hereditary colon cancer. Science (New York, N.Y.) 263:1625-1629.

Pavlov YI, Frahm C, Nick McElhinny SA, Niimi A, Suzuki M, Kunkel TA. 2006. Evidence that errors made by DNA polymerase alpha are corrected by DNA polymerase delta. Current biology : CB 16:202-207.

Peltomäki P. 2005. Lynch syndrome genes. Familial cancer 4:227-232.

Peltomäki P, Aaltonen LA, Sistonen P, Pylkkänen L, Mecklin JP, Järvinen H, Green JS, Jass JR, Weber JL, Leach FS. 1993. Genetic mapping of a locus predisposing to human colorectal cancer. Science (New York, N.Y.) 260:810-812.

Peltomäki P, Gao X, Mecklin JP. 2001. Genotype and phenotype in hereditary nonpolyposis colon cancer: a study of families with different vs. shared predisposing mutations. Familial cancer 1:9-15.

Peltomäki P and Vasen H. 2004. Mutations associated with HNPCC predisposition -- Update of ICG-HNPCC/INSiGHT mutation database. Disease markers 20:269-276.

Pinto C, Veiga I, Pinheiro M, Mesquita B, Jeronimo C, Sousa O, Fragoso M, Santos L, Moreira-Dias L, Baptista M, Lopes C, Castedo S, Teixeira MR. 2006. MSH6 germline mutations in early-onset colorectal cancer patients without family history of the disease. British journal of cancer 95:752-756.

Plaschke J, Engel C, Kruger S, Holinski-Feder E, Pagenstecher C, Mangold E, Moeslein G, Schulmann K, Gebert J, von Knebel Doeberitz M, Ruschoff J, Loeffler M, Schackert HK. 2004. Lower incidence of colorectal cancer and later age of disease onset in 27 families with pathogenic MSH6 germline mutations compared with families with MLH1 or MSH2 mutations: the German Hereditary Nonpolyposis Colorectal Cancer Consortium. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 22:4486-4494.

Plaschke J, Linnebacher M, Kloor M, Gebert J, Cremer FW, Tinschert S, Aust DE, von Knebel Doeberitz M, Schackert HK. 2006. Compound heterozygosity for two MSH6 mutations in a patient with early onset of HNPCC-associated cancers, but without hematological malignancy and brain tumor. European journal of human genetics : EJHG 14:561-566.

Plon SE, Eccles DM, Easton D, Foulkes WD, Genuardi M, Greenblatt MS, Hogervorst FB, Hoogerbrugge N, Spurdle AB, Tavtigian SV, IARC Unclassified Genetic Variants Working Group. 2008. Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. Human mutation 29:1282-1291.

Pluciennik A, Dzantiev L, Iyer RR, Constantin N, Kadyrov FA, Modrich P. 2010. PCNA function in the activation and strand direction of MutLalpha endonuclease in mismatch repair. Proceedings of the National Academy of Sciences of the United States of America 107:16066-16071.

Pospiech H, Kursula I, Abdel-Aziz W, Malkas L, Uitto L, Kastelli M, Vihinen-Ranta M, Eskelinen S, Syväoja JE. 1999. A neutralizing antibody against human DNA polymerase epsilon inhibits cellular but not SV40 DNA replication. Nucleic acids research 27:3799-3804.

Pursell ZF, Isoz I, Lundström EB, Johansson E, Kunkel TA. 2007. Yeast DNA polymerase epsilon participates in leading-strand DNA replication. Science (New York, N.Y.) 317:127-130.

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

Raevaara TE, Vaccaro C, Abdel-Rahman WM, Mocetti E, Bala S, Lönnqvist KE, Kariola R, Lynch HT, Peltomäki P, Nyström-Lahti M. 2003. Pathogenicity of the hereditary colorectal cancer mutation hMLH1 del616 linked to shortage of the functional protein. Gastroenterology 125:501-509.

Ramilo C, Gu L, Guo S, Zhang X, Patrick SM, Turchi JJ, Li GM. 2002. Partial reconstitution of human DNA mismatch repair in vitro: characterization of the role of human replication protein A. Molecular and cellular biology 22:2037-2046.

Risinger JI, Umar A, Boyd J, Berchuck A, Kunkel TA, Barrett JC. 1996. Mutation of MSH3 in endometrial cancer and evidence for its functional role in heteroduplex repair. Nature genetics 14:102-105.

Rytkönen AK, Vaara M, Nethanel T, Kaufmann G, Sormunen R, Laara E, Nasheuer HP, Rahmeh A, Lee MY, Syväoja JE, Pospiech H. 2006. Distinctive activities of DNA polymerases during human DNA replication. The FEBS journal 273:2984-3001.

Samowitz WS, Curtin K, Lin HH, Robertson MA, Schaffer D, Nichols M, Gruenthal K, Leppert MF, Slattery ML. 2001. The colon cancer burden of genetically defined hereditary nonpolyposis colon cancer. Gastroenterology 121:830-838.

Schmutte C, Marinescu RC, Sadoff MM, Guerrette S, Overhauser J, Fishel R. 1998. Human exonuclease I interacts with the mismatch repair protein hMSH2. Cancer research 58:4537-4542.

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

Schröering AG, Edelbrock MA, Richards TJ, Williams KJ. 2007. The cell cycle and DNA mismatch repair. Experimental cell research 313:292-304.

Shcherbakova PV and Kunkel TA. 1999. Mutator phenotypes conferred by MLH1 overexpression and by heterozygosity for mlh1 mutations. Molecular and cellular biology 19:3177-3183.

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

Spurdle AB, Couch FJ, Hogervorst FB, Radice P, Sinilnikova OM, IARC Unclassified Genetic Variants Working Group. 2008. Prediction and assessment of splicing alterations: implications for clinical testing. Human mutation 29:1304-1313.

Stone EA and Sidow A. 2005. Physicochemical constraint violation by missense substitutions mediates impairment of protein function and disease severity. Genome research 15:978-986.

Subramanian S, Mishra RK, Singh L. 2003. Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. Genome biology 4:R13.

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

Tainer JA, McCammon JA, Ivanov I. 2010. Recognition of the ring-opened state of proliferating cell nuclear antigen by replication factor C promotes eukaryotic clamp-loading. Journal of the American Chemical Society 132:7372-7378.

Tanaka S, Hu SZ, Wang TS, Korn D. 1982. Preparation and preliminary characterization of monoclonal antibodies against human DNA polymerase alpha. The Journal of biological chemistry 257:8386-8390.

Tanyi M, Olasz J, Kamory E, Csuka O, Tanyi JL, Ress Z, Damjanovich L. 2008. Difficulties in recognizing families with Hereditary Non-polyposis Colorectal Carcinoma. Presentation of 4 families with proven mutation. European journal of surgical oncology : the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology 34:1322-1327.

Tavtigian SV, Greenblatt MS, Lesueur F, Byrnes GB, IARC Unclassified Genetic Variants Working Group. 2008. In silico analysis of missense substitutions using sequence-alignment based methods. Human mutation 29:1327-1336.

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

Thomas DC, Roberts JD, Kunkel TA. 1991. Heteroduplex repair in extracts of human HeLa cells. The Journal of biological chemistry 266:3744-3751.

Tian L, Hou C, Tian K, Holcomb NC, Gu L, Li GM. 2009. Mismatch recognition protein MutSbeta does not hijack (CAG)n hairpin repair in vitro. The Journal of biological chemistry 284:20452-20456.

Tishkoff DX, Amin NS, Viars CS, Arden KC, Kolodner RD. 1998. Identification of a human gene encoding a homologue of Saccharomyces cerevisiae EXO1, an exonuclease implicated in mismatch repair and recombination. Cancer research 58:5027-5031.

Tishkoff DX, Boerger AL, Bertrand P, Filosi N, Gaida GM, Kane MF, Kolodner RD. 1997. Identification and characterization of Saccharomyces cerevisiae EXO1, a gene encoding an exonuclease that interacts with MSH2. Proceedings of the National Academy of Sciences of the United States of America 94:7487-7492.

Tran PT, Erdeniz N, Symington LS, Liskay RM. 2004. EXO1-A multi-tasking eukaryotic nuclease. DNA repair 3:1549-1559.

Tseng Q, Orans J, Hast MA, Iyer RR, Changela A, Modrich PL, Beese LS. 2011. Purification, crystallization and preliminary X-ray diffraction analysis of the human mismatch repair protein MutSbeta. Acta crystallographica.Section F, Structural biology and crystallization communications 67:947-952.

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

Umar A, Boyer JC, Kunkel TA. 1994. DNA loop repair by human cell extracts. Science (New York, N.Y.) 266:814-816.

Umar A, Buermeyer AB, Simon JA, Thomas DC, Clark AB, Liskay RM, Kunkel TA. 1996. Requirement for PCNA in DNA mismatch repair at a step preceding DNA resynthesis. Cell 87:65-73.

Umar A, Koi M, Risinger JI, Glaab WE, Tindall KR, Kolodner RD, Boland CR, Barrett JC, Kunkel TA. 1997. Correction of hypermutability, N-methyl-N'-nitro-N-nitrosoguanidine resistance, and defective DNA mismatch repair by introducing chromosome 2 into human tumor cells with mutations in MSH2 and MSH6. Cancer research 57:3949-3955.

Umar A, Risinger JI, Hawk ET, Barrett JC. 2004. Testing guidelines for hereditary non-polyposis colorectal cancer. Nature reviews.Cancer 4:153-158.

Vasen HF, Mecklin JP, Khan PM, Lynch HT. 1991. The International Collaborative Group on Hereditary Non-Polyposis Colorectal Cancer (ICG-HNPCC). Diseases of the colon and rectum 34:424-425.

Vasen HF, Watson P, Mecklin JP, Lynch HT. 1999. New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC. Gastroenterology 116:1453-1456.

Veigl ML, Kasturi L, Olechnowicz J, Ma AH, Lutterbaugh JD, Periyasamy S, Li GM, Drummond J, Modrich PL, Sedwick WD, Markowitz SD. 1998. Biallelic inactivation of hMLH1 by epigenetic gene silencing, a novel mechanism causing human MSI cancers. Proceedings of the National Academy of Sciences of the United States of America 95:8698-8702.

Wagner A, Hendriks Y, Meijers-Heijboer EJ, de Leeuw WJ, Morreau H, Hofstra R, Tops C, Bik E, Brocker-Vriends AH, van Der Meer C, Lindhout D, Vasen HF, Breuning MH, Cornelisse CJ, van Krimpen C, Niermeijer MF, Zwinderman AH, Wijnen J, Fodde R. 2001. Atypical HNPCC owing to MSH6 germline mutations: analysis of a large Dutch pedigree. Journal of medical genetics 38:318-322.

Warren JJ, Pohlhaus TJ, Changela A, Iyer RR, Modrich PL, Beese LS. 2007. Structure of the human MutSalpha DNA lesion recognition complex. Molecular cell 26:579-592.

Wei K, Clark AB, Wong E, Kane MF, Mazur DJ, Parris T, Kolas NK, Russell R, Hou H,Jr, Kneitz B, Yang G, Kunkel TA, Kolodner RD, Cohen PE, Edelmann W. 2003. Inactivation of Exonuclease 1 in mice results in DNA mismatch repair defects, increased cancer susceptibility, and male and female sterility. Genes & development 17:603-614.

Wielders EA, Dekker RJ, Holt I, Morris GE, te Riele H. 2011. Characterization of MSH2 variants by endogenous gene modification in mouse embryonic stem cells. Human mutation 32:389-396.

Wimmer K and Etzler J. 2008. Constitutional mismatch repair-deficiency syndrome: have we so far seen only the tip of an iceberg? Human genetics 124:105-122.

Woods MO, Williams P, Careen A, Edwards L, Bartlett S, McLaughlin JR, Younghusband HB. 2007. A new variant database for mismatch repair genes associated with Lynch syndrome. Human mutation 28:669-673.

Wu Y, Berends MJ, Mensink RG, Kempinga C, Sijmons RH, van Der Zee AG, Hollema H, Kleibeuker JH, Buys CH, Hofstra RM. 1999. Association of hereditary nonpolyposis colorectal cancer-related tumors displaying low microsatellite instability with MSH6 germline mutations. American Journal of Human Genetics 65:1291-1298.

Wu Y, Berends MJ, Sijmons RH, Mensink RG, Verlind E, Kooi KA, van der Sluis T, Kempinga C, van dDer Zee AG, Hollema H, Buys CH, Kleibeuker JH, Hofstra RM. 2001. A role for MLH3 in hereditary nonpolyposis colorectal cancer. Nature genetics 29:137-138.

Yuan F, Gu L, Guo S, Wang C, Li GM. 2004. Evidence for involvement of HMGB1 protein in human DNA mismatch repair. The Journal of biological chemistry 279:20935-20940.

Zhang Y, Yuan F, Presnell SR, Tian K, Gao Y, Tomkinson AE, Gu L, Li GM. 2005. Reconstitution of 5'-directed human mismatch repair in a purified system. Cell 122:693-705.