Cover Page

Universiteit Leiden



The handle http://hdl.handle.net/1887/24909 holds various files of this Leiden University dissertation

Author: Drost, Mark
Title: Development of in vitro and in vivo functional assays to enable diagnosis of variants of uncertain significance in the common cancer predisposition Lynch syndrome
Issue Date: 2014-03-26

DEVELOPMENT OF *IN VITRO* AND *IN VIVO* FUNCTIONAL ASSAYS TO ENABLE DIAGNOSIS OF VARIANTS OF UNCERTAIN SIGNIFICANCE IN THE COMMON CANCER PREDISPOSITION LYNCH SYNDROME

Mark Drost

Cover Design, Layout & Printing: Off Page, Amsterdam, www.offpage.nl

The cover displays the human MutSα (MSH2/MSH6) crystal structure. This structure was downloaded from the Protein Databank (http://www.rcsb.org/pdb/; PBD ID PD0951) and described in Warren JJ et al., Mol Cell 26:579–592.

ISBN: 978-94-6182-407-3

Printing of this thesis was supported financially by Biolegio BV and Promega BV.

© Copyright 2014 by Mark Drost. All rights reserved.

No part of this book may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, without prior permission of the author, or when appropriate, of the publisher of the represented published articles.

DEVELOPMENT OF *IN VITRO* AND *IN VIVO* FUNCTIONAL ASSAYS TO ENABLE DIAGNOSIS OF VARIANTS OF UNCERTAIN SIGNIFICANCE IN THE COMMON CANCER PREDISPOSITION LYNCH SYNDROME

Proefschrift

ter verkrijging van de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 26 maart 2014
klokke 16:15 uur

door

Mark Drost

geboren te Delft
in 1983

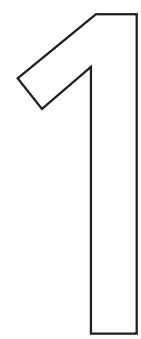
PROMOTIECOMMISSIE

Promotor:	Prof. Dr. L.H.F. Mullenders	
Co-promotor:	Dr. N. de Wind	
Overige leden:	Prof. Dr. P. Devilee	
	Prof. Dr. C.J. van Asperen	
	Prof. Dr. R. Sijmons	Universitair Medisch Centrum Groningen
	Prof. Dr. H. te Riele	Nederlands Kanker Instituut

TABLE OF CONTENTS

Chapter 1	General Introduction: The Genetics and Biochemistry of DNA Mismatch Repair and Implications for Carcinogenesis	7
Chapter 2	Pathological Assessment of Mismatch Repair Gene Variants in Lynch Syndrome: Past, Present, and Future	31
Chapter 3	A Rapid and Cell-free Assay to Test the Activity of Lynch Syndrome-associated *MSH2* and *MSH6* Missense Variants	51
Chapter 4	A Cell-free Assay for the Functional Analysis of Variants of the Mismatch Repair Protein MLH1	71
Chapter 5	Inactivation of DNA Mismatch Repair by Variants of Uncertain Significance in the *PMS2* Gene	93
Chapter 6	Genetic Screens to Identify Pathogenic Mismatch Repair Gene Variants in the Common Cancer Predisposition Lynch Syndrome	103
Chapter 7	A Genetic Screen to Identify Inactivating Variants in the *Msh6* Gene Suggests Defective Mismatch Binding as a Major Mechanism of Pathogenicity	127
Chapter 8	General Discussion and Future Perspectives	143
Appendix	Summary	157
	Samenvatting	161
	List of Publications	165
	Curriculum Vitae	167
	Acknowledgements	169

General Introduction: The Genetics and Biochemistry of DNA Mismatch Repair and Implications for Carcinogenesis



Genome Stability

DNA repair pathways protect cells from genomic instability, and the organism from aging and cancer. To preserve the information encoded by the DNA of a cell it is essential to prevent mutations. Mutations can arise as a consequence of unrepaired exogenous and endogenous damage to DNA. A plethora of mechanisms is available to cells to either repair or tolerate DNA damages (reviewed in [Hoeijmakers, 2009; Jackson and Bartek, 2009; Ciccia and Elledge, 2010]. Though DNA damage is a major threat to the genome, its integrity can also be threatened by DNA replication. The fidelity of DNA replication is ensured by several factors. First, the main replicative polymerases δ and ε display a high degree of nucleotide selectivity, misincorporating at a frequency of only approximately 10^{-4} to 10^{-5} [Iyer et al., 2006; Arana and Kunkel, 2010]. Then, the proofreading activities of polymerases δ and ε can remove most misincorporations, increasing replication fidelity to 10^{-7} per basepair per replication round. Residual misincorporations are detected and corrected by the DNA mismatch repair (MMR) system. The orchestrated action of these pathways ensures extremely accurate replication fidelity [Arana and Kunkel, 2010], and protects organisms from cancer development [Hanahan and Weinberg, 2000, 2011].

DNA Mismatch Repair

MMR is an evolutionary conserved postreplicative DNA repair system that corrects (spontaneous) replication errors, such as base-base mispairs and insertion/deletion loops (IDLs; reviewed in [Hsieh and Yamane, 2008; Li, 2008; Jiricny, 2013]). Besides its role in replication fidelity, MMR has also been implicated in homologous recombination, in somatic hyper mutation, in triplet repeat expansion, crosslink repair and in the response to various DNA damaging agents such as methylating agents, oxidative agents, ionizing radiation and nucleotide analogs [Stojic et al., 2004; McMurray, 2010; Deans and West, 2011; Jiricny, 2013 and references therein].

During the repair of replication errors in eukaryotes, the MMR pathway is initiated by the recognition of a mismatch or IDL by either the heterodimeric MutSα, consisting of the homologous MSH2 and MSH6 proteins, or the MutSβ heterodimer consisting of MSH2 and MSH3. These early factors in MMR are named after their bacterial homolog MutS (MSH=MutS Homolog) that was first described over 40 years ago [Cox et al., 1972]. The MutSα complex is responsible for the recognition of base-base mispairs and small IDLs, whereas MutSβ is responsible for the redundant repair mainly of larger IDLs. Each subunit of MutSα and MutSβ consists of five major domains, which have been referred to as a mismatch binding domain, a connector domain, a lever domain, a clamp domain and an ATPase domain (Figure 1A and B; [Warren et al., 2007; Gupta et al., 2012]). These domain names are analogous to the domains described in the bacterial MutS [Lamers et al., 2000; Obmolova et al., 2000]. The term ”mismatch binding domain” is, in fact, a misnomer for MSH2 and MSH3, as only the mismatch binding in MSH6 actually binds mismatches. Both the MutSα and the MutSβ complex very much resembe a Θ-like structure. The (mis) matched DNA is in the upper channel, the lower channel is empty, and MSH2 and MSH6 line the sides along the long axis (Figure 1C; [Warren et al., 2007; Gupta et al., 2012]).

The MutSα heterodimer is recruited onto chromatin through a PWWP motif in the N-terminus of MSH6, by interacting with the histone mark H3K56me3 [Li et al., 2013]. In addition, MutSα complexes are targeted to newly synthesized DNA by being a constitutive part of replication factories in a manner depending on an interaction with PCNA (Figure 2; [Iyer et al., 2008; Hombauer et al., 2011a]). Also, MMR proteins only function when expressed during replication and not after, in support of the coupling between replication and MMR [Hombauer et al., 2011b]. Upon mismatch binding, MSH6 makes contact with the mismatch through a conserved Phe-X-Glu motif in the mismatch binding domain in its N-terminus. The phenylalanine residue is inserted into the DNA helix and stacks against the mispaired nucleotide, which results in kinking of the DNA by 60° [Lamers et al., 2000; Obmolova et al., 2000; Warren et al., 2007]. MSH3 lacks a Phe-X-Glu motif but it contacts the backbone of the DNA through several amino acids its N-terminus [Gupta et al., 2012].

All MutS homologs are ATPases that carry a Walker ATP motif, a common motif among DNA repair proteins, in their C-terminal ATPase domains. They are members of the ATP Binding Cassette (ABC) transporter ATPase superfamily (Figure 1A; reviewed in [Holland and Blight, 1999]). Each MutS homolog heterodimer has two composite nucleotide binding sites, one on each subunit. Both contain a Walker A and Walker B motif from one subunit, and an ABC signature motif from the other subunit. The binding sites are nonidentical: in MutSα, MSH6 has high affinity for ATP, whereas MSH2 has a higher affinity for ADP [Antony and Hingorani, 2003; Antony et al., 2006]. After the MutSα complex binds a mismatch through its mismatch binding domain, the complex exchanges ADP for ATP and undergoes an ATP-induced conformational change, converting the heterodimer into a sliding clamp [Gradia et al., 1999], which can freely move along the DNA helix contour [Gorman et al., 2007].

After DNA mismatch binding MutSα (or MutSβ) recruits the MutLα heterodimer (consisting of MLH1-PMS2; Figure 2). This complex is thought to act as a matchmaker protein between the MutSα mismatch recognizing complex and the downstream exonuclease. Both MLH1 and PMS2 contain an N-terminal ATPase domain and a C-terminal dimerization domain, connected by a linker (Figure 1C). MLH1 (Mut L Homolog 1) belongs to the superfamily of GKHL (gyrase II/Hsp90/histidine kinase/MutL) ATPases (Figure 1C and D). MutLα is partially redundant with another heterodimer, MutLγ (consisting of MLH1-MLH3). However, MutLγ cannot fully substitute for MutLα, at least *in vitro* [Cannavo et al., 2005; Korhonen et al., 2008]. The MutLα complex helps to stabilize the ternary DNA·MutSα·MutLα complex. Like MutSα, MutLα has (weak) ATPase activity and undergoes sequential conformational changes after ATP binding (Figure 1C; [Sacho et al., 2008]), but the relevance of these changes is still very poorly understood.

Besides functioning as a matchmaker, MutLα is very important because of the endonuclease activity within the C-terminus of the PMS2 (Postmeiotic segregation increased 2) subunit. To direct MMR-mediated excision to the newly synthesized strand, MMR requires a strand discrimination signal that *in vitro* is a nick [Lahue et al., 1989; Constantin et al., 2005]. It has been proposed that lagging strand MMR is directed, at least in part, by 3' or 5' ends of Okazaki fragments [Pavlov et al., 2003]. MMR at the leading strand, in turn, partly depends

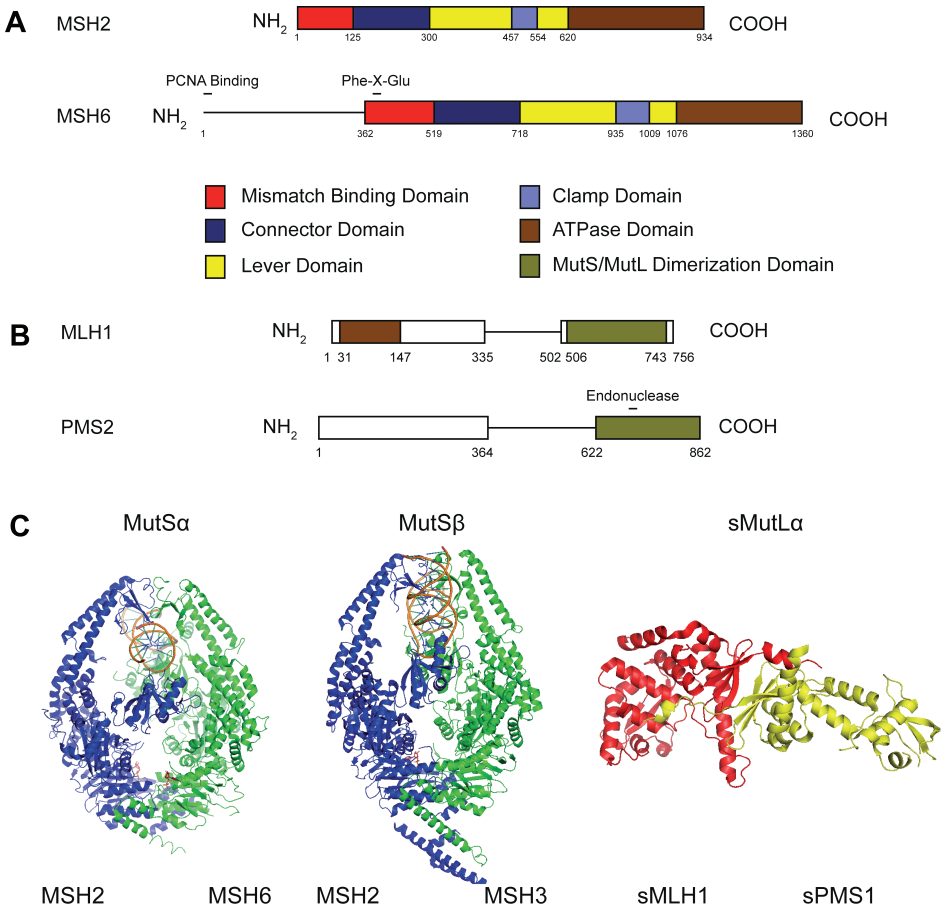


Figure 1. *Domains and crystal structures of mismatch repair proteins and complexes. (A) Schematic representation of domains in the primary protein structures of MSH2 and MSH6. Numbering reflects amino acid numbering. Domains are as published [Warren et al., 2007]. Of note, the mismatch binding domain in MSH2 does not bind mismatches (see main text). PCNA binding: PCNA binding motif [Hombauer et al., 2011a]. Phe-X-Glu: Mismatch binding motif. (B) As A, but for MLH1 and PMS2. Domains have been described in [Raevaara et al., 2005; Gueneau et al., 2013]. Endonuclease: Endonuclease motif [Kadyrov et al., 2006]. (C) Crystal structures of the human MutSα (MSH2/MSH6; [Warren et al., 2007]), human MutSβ (MSH2/MSH3; [Gupta et al., 2012]) and S. cerevisiae MutLα (MLH1/PMS1, only the C-terminal domain is shown as only this part has been crystallized, now [Gueneau et al., 2013]).*

on small gaps induced after the removal of ribonucleotides from the nascent strand by RNase H2, which are fortuitously incorporated during replication by the leading strand polymerase ε [Ghodgaonkar et al., 2013; Lujan et al., 2013]. In addition to these RNase H2-induced gaps, additional signals may serve as strand a discrimination, such as the 3' terminus of a leading strand. The 3' terminus at the nascent strand dictates the polarity of the processivity factor Proliferating Cell Nuclear Antigen (PCNA) on DNA, which in turn determines the directionality of the nicking activity of MutLα on DNA [Peña-Diaz and Jiricny, 2010; Pluciennik

et al., 2010]. The endonuclease activity of the PMS2 subunit introduces new nicks on the nascent strand, 5' of the mismatch [Kadyrov et al., 2006]. Such nicks serve as entry point for Exonuclease 1 (EXO1), a 5'-3' exonuclease that carries out excision of the misincorporation-containing DNA strand. *In vitro* reconstitution of MMR depends on EXO1 [Dzantiev et al., 2004; Constantin et al., 2005] but, oddly enough, *Exo1*$^{-/-}$ cells have a less severe phenotype than *Msh2*$^{-/-}$ cells [Wei et al., 2003], suggesting that other exonucleases or mechanisms may substitute for EXO1 function. It has recently been shown that MMR, at least *in vitro*, can also occur independently from EXO1 through strand displacement by Polδ (Figure 2; [Kadyrov et al., 2009]). In addition, EXO1 is preferentially required for lagging strand MMR and not for leading strand MMR [Hombauer et al., 2011a], suggesting mechanistically distinct MMR events that differentially require EXO1. As the lagging strand provides an ideal substrate both for strand displacement and EXO1-mediated degradation, this latter finding is surprising.

Once the misincorporation-containing DNA strand is excised, the single stranded DNA patch is protected by the single-strand DNA binding protein Replication Protein A (RPA) and the strand is resynthesized by Polδ in the presence the clamp loader Replication Factor

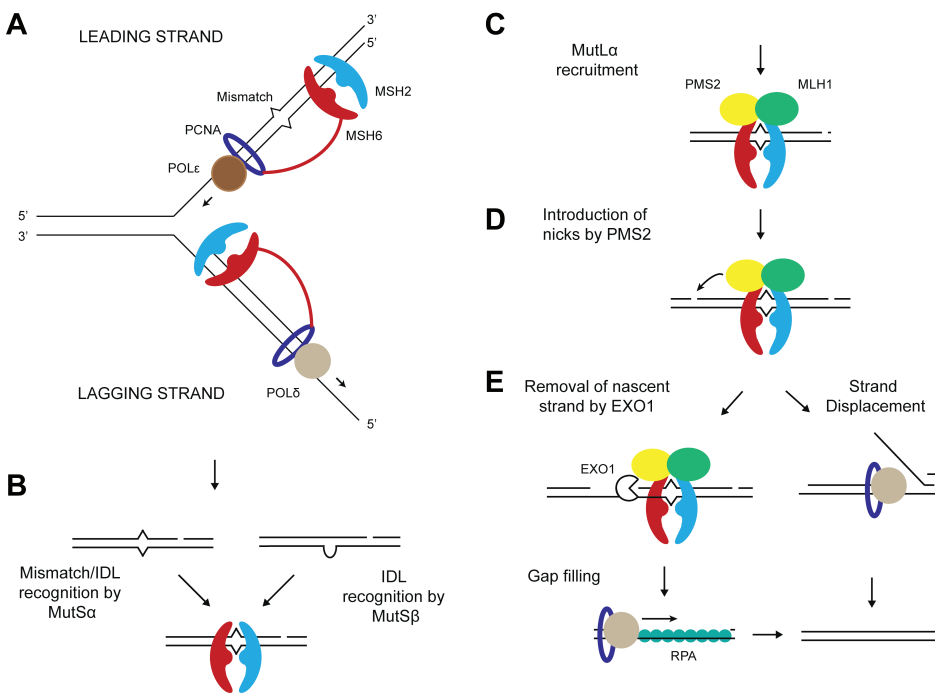


Figure 2. *Sequential steps of canonical MMR. (A) MutSα is a constitutive part of replication factories through an interaction between the N-terminus of MSH6 and PCNA. (B) Mismatches are recognized by MutSα (left panel), while IDLs are recognized mostly by MutSβ. (C) Once a substrate is bound, the MutLα heterodimer is recruited. (D) The PMS2 subunit introduces nicks 5' to the mismatch, creating a substrate for either strand displacement by POLδ or for EXO1-mediated strand removal. After the former, the flap is removed by a flap endonuclease. (E) After the latter, the resulting gap is filled by RFC, PCNA and POLδ, the nick is ligated, and repair is completed.*

C (RFC) that loads PCNA (Figure 2; [Longley et al., 1997; Dzantiev et al., 2004; Constantin et al., 2005]). Additional proteins possibly involved in MMR are HMGB1 [Yuan et al., 2004; Genschel and Modrich, 2009], RFX [Zhang et al., 2008] and PARP1 [Liu et al., 2011], but their roles are more elusive.

Discrimination by MutSα between matched and mismatched DNA *in vitro* is only ~5-30 fold [Hays et al., 2005] and of all mismatches, only the G·T mismatch is bound well by MutSα in Electrophoretic Mobility Shift Assays [Jiricny et al., 1988]. However, in *in vitro* MMR assays, all mismatches are repaired when embedded in perfectly paired DNA [Hays et al., 2005]. Repair of the C·C mismatch is less efficient, but this mismatch is very efficiently removed by the proofreading activity of the replicative polymerases. Overall, the mechanistic basis of detecting mismatches in a large excess of matched DNA is still poorly understood. Mismatch discrimination may be a very complex process, ensured by several (kinetic) mechanisms and/or additional MMR factors [Labazi et al., 2009]. Multiple (ATP-controlled) conformational changes in MutSα may contribute to specificity. Also, the ATP-induced sliding along the DNA may provide only a short time window for MutLα to bind, providing a "kinetic control" of heteroduplex binding. Furthermore, ATP-induced conformational changes in MutLα may also contribute to mismatch specificity [Sacho et al., 2008].

Mutator Phenotype and Microsatellite Instability

MMR increases replication fidelity by 100 to 1000 fold to one replication error per 10^9-10^{10} nucleotides, and in this way protects cells from spontaneous mutations [Iyer et al., 2006]. Vice versa, the inactivation of MMR leads to a mutator phenotype in which spontaneous mutation rates are highly elevated. These elevated mutation rates affect the whole genome, including protein-encoding genes.

MMR deficiency also results in microsatellite instability (MSI). MSI refers to size changes, often shrinkage, of genomic repetitive sequences (such as $[CA]_N$, or $[A]_N$; reviewed in [Boland and Goel, 2010]). This hyper mutable phenotype is caused by low processivity of the replicative polymerases on microsatellites. During replication of these loci, polymerases often have to be reloaded, a process that frequently induces IDLs. These IDLs are a substrate for MMR, which aids in the maintenance of the stability of microsatellites [Strand et al., 1993].

Tolerance of Genotoxic Agents

Besides recognizing mismatches and IDLs, MutSα is also capable of detecting and binding modified DNA bases, such as those induced by oxygen radicals [Mazurek et al., 2002; Macpherson et al., 2005], methylating agents [Duckett et al., 1996; Rasmussen and Samson, 1996] or UV (but only when mispaired [Hoffman et al., 2005]). Cells deficient for MMR are tolerant to various genotoxic agents, such as the methylating agents *N*-Methyl-*N*-nitroso-*N'*-nitroguanidine (MNNG) and Temozolomide, but also to nucleotide analogs such as 6-thioguanine (6-TG; [Swann et al., 1996]) and Fluorouracil (5-FU; [Meyers et al., 2003]).

Of these drugs, the response of MMR proficient and deficient cells to the methylating agent MNNG has received much attention. This agent gives rise to O^6-methylguanines (amongst

others), a highly mutagenic lesion due to its efficient basepairing with thymine [Warren et al., 2006]. The finding that O^6-methylguanines are the toxic lesions is substantiated by the fact that the detoxifying protein methylguanine-DNA methyltransferase (MGMT) protects the cells against MNNG-induced apoptosis [Kaina et al., 2007]. MMR deficient cells are up to a 100 fold more tolerant to MNNG compared with MMR proficient cells [Goldmacher et al., 1986; Kat et al., 1993; Stojic et al., 2005]. The investigation of the MMR-mediated response to MNNG has yielded two major models, that are not mutually exclusive, to explain these phenotypes. In the first model, generally referred to as the direct signaling model, the MMR-mediated response to methylating agents may be regulated in a direct manner, in the absence of MMR-mediated excision. *In vitro* the signaling kinase ATR/ATRIP, but not RPA, is preferentially recruited to O^6meG·T mismatches on a covalently closed circular substrate in a MutSα and MutLα dependent manner. Also, ATR is activated to phosphorylate its target CHK1 in the presence of O^6-meG·T mismatches and MMR proteins *in vitro* in the absence of excision, suggesting direct activation of signaling proteins (Figure 3, right panel; [Yoshioka et al., 2006]). This may be further supported by cellular data showing interactions between MutSα and ATR [Liu et al., 2009; Pabla et al., 2011], TopBP1, Claspin and CHK1, and between MutLα and TopBP1 and Claspin [Liu et al., 2009]. Liu et. al also showed binding of ATR, TopBP1 and CHK1, but not RPA, to chromatin after MNNG treatment in a MutSα and MutLα dependent manner. Oddly enough, this paper only shows recruitment of these factors but not their activation nor contribution in cell cycle responses. An interesting experiment would be to find/generate mutant MMR proteins that do not bind the signaling proteins, and see whether these interactions are required to induce apoptosis in response to MNNG.

In case the direct signaling model is true, one would predict that DNA damage signaling and apoptosis are induced in the first cell cycle after treatment. In contrast to this, DNA damage signaling, G2 arrest and apoptosis are induced in the second cell cycle after treatment [Stojic et al., 2004; Mojas et al., 2007]. This finding is the basis of a second model, referred to as the futile cycling model. Since the O^6-meG cannot properly form a Watson-Crick basepair with any other nucleotide, basepairing of O^6-meG will often yield an MMR substrate. However, since MMR is directed only towards the newly synthesized strand, not the O^6-meG in the parental strand, but its mispairing partner in the nascent strand is removed. Polymerase-mediated gap filling of this structure yields yet another MMR substrate, inducing multiple rounds of MMR [York and Modrich, 2006]. The notion that actual excision of the mismatch is required is supported by the fact that O^6-meG induced cell death requires (the catalytic function of) EXO1 [Klapacz et al., 2009; Schaetzlein et al., 2013]. After several attempts of futile repair, iterative MMR-induced gaps opposite the O^6-meG are believed to be carried over to a subsequent S-phase, in which they becomes toxic DNA intermediates that will induce double-strand breaks when replicated (Figure 3, left panel; [Stojic et al., 2005; Mojas et al., 2007]). The idea that double strand breaks are the inducer of cell death in response to methylating agents is supported by the finding that both ATM [Debiak et al., 2004] and BRCA1 [Yamane et al., 2007] are involved in the cellular response to methylation damage and that *S. cerevisiae* is relatively tolerant to MNNG-induced cell death due to efficient recombination [Cejka et al., 2005].

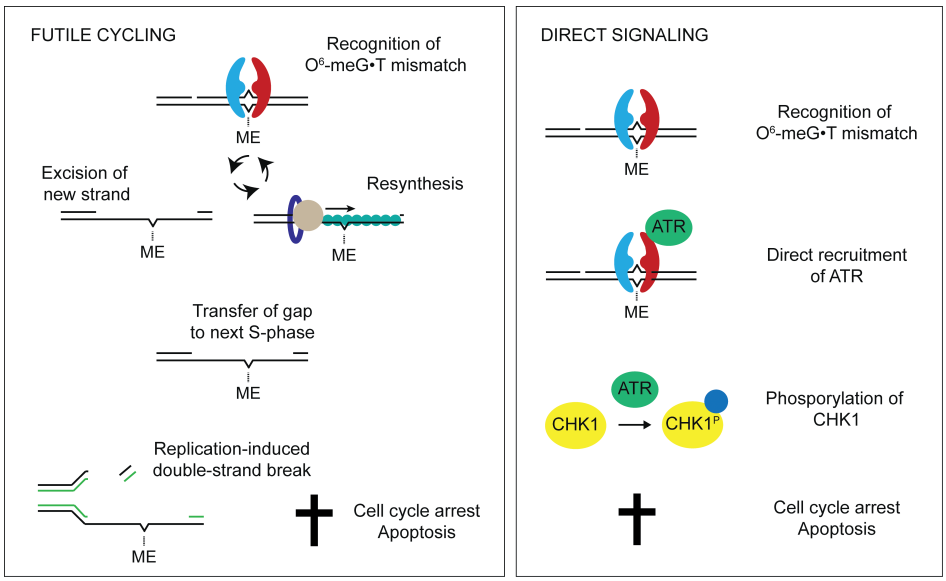


Figure 3. *Possible mechanisms of the MMR-mediated response to methylating agents. Left panel: MutSα binds the mismatched, methylated basepair and triggers excision. Since the methylated nucleotide is in the parental strand, it is not excised and gap-filling creates a novel MMR substrate, which induces rounds of futile cycling. Gapped intermediates are transferred to the next cell cycle, in which they are replicated and generate a double-strand break, which triggers apoptosis. Right panel: MutSα binds the mismatched, methylated basepair and directly recruits ATR, which phosphorylates its target CHK1, leading to cell cycle arrest and apoptosis.*

Cancer in Mammals

The accumulation of mutations in DNA is proposed to be a driving force for cancer development [Hanahan and Weinberg, 2000, 2011], and this is reflected in mouse models of MMR. These models have directly linked loss of MMR to cancer predisposition, and help to understand the etiology and tissue distribution of tumours (reviewed in [Wei et al., 2002; Edelmann and Edelmann, 2004; Taketo and Edelmann, 2009]).

In terms of cancer predisposition, the divergent phenotypes of *Msh2*$^{-/-}$, *Msh3*$^{-/-}$ and *Msh6*$^{-/-}$ mice reflect the functional redundancy between MutSα and MutSβ. *Msh2*$^{-/-}$ mice die within a year due to strong cancer predisposition [de Wind et al., 1995, 1999]. The majority of these mice develop T-cell lymphomas. *Msh6*$^{-/-}$ mice are also tumor prone, but less than their *Msh2*$^{-/-}$ counterparts, as 75% of mice succumb within a year [Edelmann et al., 1997; de Wind et al., 1999]. *Msh3*$^{-/-}$ mice have no apparent phenotype, but in an *Msh6*$^{-/-}$ background, Msh3 loss accelerates tumorigenesis [de Wind et al., 1999].

Interestingly, inactivation of either *Mlh1* or *Pms2* leads to infertility [Baker et al., 1996; Edelmann et al., 1996]. Mlh1-deficient spermatocytes exhibit high levels of prematurely separated chromosomes and arrest in the first division of meiosis, suggesting involvement of Mlh1 in meiotic crossing over. *Mlh1*$^{-/-}$ animals are, like the *Msh2*$^{-/-}$ mice, highly cancer prone. *Pms2*$^{-/-}$ mice succumb to cancer later in life, suggesting functional redundancy of

Pms2 with other heterodimeric Mlh1 partners [Prolla et al., 1998]. Mouse knockouts of the exonuclease involved in MMR, Exo1, are predisposed to cancer but not as strong as *Msh2*$^{-/-}$ or *Mlh1*$^{-/-}$ mice. As mentioned before, this suggests that other exonucleases or mechanisms may substitute for Exo1 function [Wei et al., 2003].

The mouse and human MMR systems are very similar. This is reflected by the fact that the human MLH1 cDNA is able to complement MMR defects in mouse *Mlh1*$^{-/-}$ fibroblasts [Buermeyer et al., 1999]. While MMR deficient mice rapidly develop cancer, mice heterozygous for an MMR gene do not. Such mice are MMR proficient and have to stochastically lose their remaining wild type allele through Loss of Heterozygosity (LOH), and subsequently accumulate mutations for mutation-driven carcinogenesis. Due to the relatively short lifespan of mice, there likely is insufficient time for these events. In contrast, humans heterozygous for an MMR gene suffer from a cancer susceptibility, called Lynch syndrome.

Lynch Syndrome

Lynch syndrome (LS) is a hereditary cancer predisposition syndrome characterized by an increased risk of cancers, mainly of the gastrointestinal and genitourinary tracts (reviewed in [Lynch and de la Chapelle, 2003; Lagerstedt Robinson et al., 2007; Boland and Lynch, 2013]). LS is caused by inherited heterozygous mutations in one of the MMR genes (see below). Since the acquisition of a single affected allele through the germline is considered to cause LS, it is inherited in an autosomal dominant fashion. However, since LS patients require a second hit at a cellular level to induce a phenotype, LS is considered recessive cellularly. It is the stochastic loss of the second, wild type, allele (Loss of Heterozygosity, LOH) that inactivates MMR in a subset of cells. This leads to an accumulation of mutations and the accelerated onset of cancer.

Genetics of Lynch Syndrome

Lynch syndrome is typically caused by mutations in *MSH2, MLH1*, *MSH6* or *PMS2* [Fishel et al., 1993; Leach et al., 1993; Bronner et al., 1994; Nicolaides et al., 1994; Edelmann et al., 1997]. Mutations in *MLH1* and *MSH2* account for the majority of Lynch syndrome (42% and 33%, respectively; [Plazzer et al., 2013]). This is in line with the genetics of MMR, where *MSH2* and *MLH1* are major players in the pathway. *MSH6* and *PMS2* are less important players in LS due to their redundancy with other MMR factors, and cover 18 and 7% of the MMR gene mutations, respectively [Gruber, 2006, Plazzer et al., 2013]. This distribution of mutations is probably subject to an ascertainment bias, as historically, highly penetrant families have been diagnosed first, leading to an overrepresentation of mutations in *MSH2* and *MLH1*.

Besides the four core MMR genes, other genes such as *MLH3* and *EXO1* have also been implicated in LS, but their roles are less clear or controversial [Wu et al., 2001b; Peltomäki and Vasen, 2004]. Some *MLH3* mutations are found in families with atypical LS features [Liu et al., 2003], while other studies were unable to identify a correlation between LS and *MLH3* [Lipkin et al., 2000]. *EXO1* mutations were also found in families with atypical LS [Wu et al., 2001a].

In the last few years, evidence for epimutations as a cause of LS has emerged. The *MSH2* gene can be inactivated *in cis* due to deletions of the 3' region of its upstream gene *TACSTD1*, which

encodes the EpCAM protein. In tissues expressing the mutated *TACSTD1* gene, transcription extends into *MSH2*, thereby resulting in methylation and silencing of the *MSH2* promoter [Ligtenberg et al., 2009]. Carriers of such a mutation have an LS-like phenotype since they are heterozygous for an MMR gene-inactivating mutation. However, since MMR deficiency is limited to EpCAM expressing tissues, their tumor spectrum tends to be different (Table 1). As for *MLH1*, Mendelian inheritance of a constitutional *MLH1* epimutation was shown to be induced by a c.-27C>A nucleotide variant within the 5' UTR [Hitchins et al., 2011].

Characteristics of Lynch Syndrome

Lynch syndrome is noteworthy for its tumor spectrum, including cancers of the gastrointestinal tract such as colorectal, stomach, pancreas, biliary tract and small bowel cancer, but also endometrial, ovarian, ureter/renal pelvis, brain and/or sebaceous gland cancers. The risk of developing such tumors depends on both the tumor type as well as the affected gene (Table 1; [Stoffel et al., 2009; Bonadona et al., 2011; Engel et al., 2012]). The age of onset of colorectal cancer (CRC) for LS patients is on average 45 years old (versus 63 years in the general population).

A hallmark of MMR deficient cancer is MSI, which is found in more than 80% of all Lynch syndrome-related tumors [de la Chapelle and Hampel, 2010]. MSI is typically found in tumors deficient for MSH2 or MLH1. MSH6 deficient tumors typically show lower levels of MSI, due to the functional redundancy of MutSα with MutSβ. Though most microsatellites are found within noncoding regions, some are found within tumor-suppressor genes such as *TGFβRII* and *IGF2R*, involved in growth, *BAX*, involved in apoptosis, and the DNA repair genes *BLM*, *MRE11*, *RAD50*, *MSH3* and *MSH6* [Giráldez et al., 2008]. *TGFβRII* for example, is inactivated in ~80% of MMR-deficient tumors [Markowitz et al., 1995;

Table 1. Risk at Lynch syndrome-related tumors per genotype, as decribed in several studies.

	Cumulative Risk, % (95% Conf. Interval)[a]				
	Gene				
Cancer subtype	MLH1	MSH2	MSH6	PMS2	EPCAM
Colorectal	41 (25-70) [b]	48 (30-77) [b]	12 (8-22) [b]	20 (11-34) [d]	75 (65-85) [c]
	79 (68-90) [c]	77 (64-90) [c]	50 (38-62) [c]		
Endometrial	54 (20-80) [b] 33 (15-51) [c]	21 (8-77) [b] 51 (33-69) [c]	16 (8-32) [b] 34 (20-48) [c]	15 (6-35) [d]	12 (0-27) [c]
Ovarian	20 (1-65) [b]	24 (3-52) [b]	1 (0-3) [b]	NA	NA
Stomach	6 (0.2-17) [b]	0.2 (0-10) [b]	0 [b]	NA	NA
Urothelium	0.2 (0-2.6) [b]	2.2 (0.6-8) [b]	0.7 (0-2.1) [b]	NA	NA
Small bowel	0.4 (0.1-3) [b]	1.1 (0-5) [b]	0 [b]	NA	NA
Biliary tract	1.9 (0-15) [b]	0.02 (0-0.2) [b]	0 [b]	NA	NA

NA: No data available. [a] Cumulative cancer risk at 70 years of Lynch syndrome associated cancers according to the mutated gene, [b] [Bonadona et al., 2011], [c] [Kempers et al., 2011], [d] [Senter et al., 2008].

Alhopuro et al., 2012]. MSI in these genes may be a bystander effect of MMR deficiency, but may also contribute to carcinogenesis. The MSI phenotype is also considered typical for the development of LS tumors, which are usually (near) diploid [Boland and Goel, 2010; de la Chapelle and Hampel, 2010]. This is in contrast with most sporadic CRCs that are aneuploid, and consistent with the hypothesis that MMR-deficient tumors are not driven by gross chromosomal instability (called the CIN phenotype; [Hanahan and Weinberg, 2000, 2011]) but by their intrinsic mutator phenotype.

In contrast to the knockout mice LS patients do not develop lymphomas, and develop tumors only in a subset of tissues. This seems a paradox, since MMR is active in all dividing cells. Several properties of MMR-deficient cells, together with the microenvironment of the colon, may explain the tumor tropism in LS patients [Chao and Lipkin, 2006; Hofstra et al., 2008]. First, the exposure to mutagens from bile and food in the proximal colon [Medina-Arana et al., 2012] may increase the mutation or LOH rate of the remaining wild type MMR allele in cells heterozygous for an MMR gene [Borgdorff et al., 2005]. In light of this, LS (or MSI) tumors are often found in the right-sided (proximal) colon, while CIN tumors are found more distally/left sided [Jass, 2004]. This can be explained by the mutagenicity of products secreted by the gallbladder, which is proximal of the colon. Second, MMR deficient cells are tolerant to various DNA damaging agents, giving these cells a survival advantage over neighboring cells that allows their clonal outgrowth [Toft et al., 1999]. In addition, the mutator phenotype of these MMR deficient cells is boosted by the hypermutability to the mutagens they are exposed to [Glaab et al., 1998; Toft et al., 1999]. The high proliferation rate of the epithelial cells of the colon provides an opportunity for mutations to accumulate. All of these factors may contribute to the tumor spectrum of LS patients [Kloor et al., 2012].

Diagnosis of Lynch Syndrome

Recognition of genetic susceptibility in suspected LS patients is critical for the clinical management of these patients [Vasen et. al, 2013]. LS patients comply to one of a specified set of criteria named the Revised Bethesda Criteria (Figure 4). These criteria are:

- CRC before 50 years of age.
- Presence of synchronous, metachronous colorectal or another LS-associated tumor (colorectal, endometrial, stomach, ovarian, pancreas, ureter/renal pelvis, biliary tract, small bowel, brain and/or sebaceous gland), regardless of age.
- CRC with MSI-H histology in a patient below 60 years of age.
- CRC diagnosed in one or more first-degree relatives with an LS-related tumor, one of which is diagnosed before the age of 50.
- CRC diagnosed in two or more first or second-degree relatives with LS-related tumors, regardless of age.

Once an individual complies to one of these criteria, that individual is clinically diagnosed as an LS patient. Even though the Revised Bethesda criteria are still commonly used, they lack sensitivity to diagnose less penetrant alleles, especially in genes such as *MSH6* and

PMS2 [Sjursen et al., 2010; Moreira et al., 2012]. Besides the Besthesda criteria, alternative prediction models for MMR gene mutations have been described but these seem to perform similarly [Win et. al, 2013]. Several studies have shown that sensitivity can be increased by screening all CRC cases (independent of the Bethesda criteria) for MMR gene mutations, but this is not yet cost-effective [Hampel et al., 2005; Moreira et al., 2012].

Diagnosis of LS is completed only by pinpointing a genetic defect in an MMR gene in these patients, therefore, a clinical diagnosis *per se* does not formally confirm the syndrome. Even so, once LS is clinically diagnosed in the index patient, the index patient and all direct relatives enroll in interval screening procedures. These include colonoscopies every one to two years, to be able to detect colorectal carcinomas or their precursor lesions at an early stage. Such surveillance effectively reduces morbidity and mortality [Lagerstedt Robinson et al., 2007]. In addition, annual gynecological examination, vaginal ultrasound and testing for the tumor biomarker CA-125 in blood is recommended for women, due to the increased risk of endometrial carcinomas [Chen et al., 2007]. It is only a genetic diagnosis which enables the presymptomatic diagnosis of LS, and will enable the differentiation between family members with normal or increased cancer risk. As a consequence, unaffected relatives can be liberated from the burden associated with screening.

Genetic testing in suspected LS patients is a stepwise process, which starts by testing the proband for MMR gene mutations. This is usually preceded by prescreening methods such as MSI testing and immunohistochemistry (IHC; Figure 4). As a pre-screening before genetic testing, nuclear expression of MMR proteins in tumors is visualized by antibody staining. IHC has the advantage that it indicates the MMR gene most eligible for sequencing analysis. The absence of expression of one of the MMR proteins is suggestive for a germline mutation within its corresponding gene and will indicate its sequencing. Only the loss of MSH6 or PMS2 can be visualized by itself, whereas loss MSH2 is always accompanied by MSH6 loss, and MLH1 loss is always accompanied by PMS2 loss [Hendriks et al., 2006]. This feature has been attributed to the loss of protein stability due to loss of dimer formation [de Wind et al., 1999; Hendriks et al., 2006]. Therefore, loss of either MSH6 or PMS2 implies a mutation in its respective gene, loss of both MSH2 and MSH6 implies an MSH2 gene defect and loss of both MLH1 and PMS2 implies an MLH1 defect (Figure 4). IHC is considered a superior predictor of MMR gene mutations, with a sensitivity of 88-90% and a specificity of 84-99% [Niessen et al., 2006].

In case IHC is inconclusive, MSI is considered a good secondary selection criterion for MMR gene mutation screening, as MSI is seen in most of the LS-associated tumors [de la Chapelle and Hampel, 2010]. The sensitivity of MSI for an MMR gene mutation is 82% and the specificity is 70% [Niessen et al., 2006]. MSI is typically tested by PCR-amplifying five microsatellites: mononucleotide repeats BAT-25 and BAT-26 and dinucleotide repeats D2S123, D5S346 and D17S250 [Umar et al., 2004; Boland and Goel, 2010; de la Chapelle and Hampel, 2010]. Recently, though, a panel of the mononucleotide repeats BAT-25, BAT-26, NR-21, NR-24 and NR-27 was put forward as an alternative, as testing for dinucleotide repeats is relatively insensitive [Goel et al., 2010]. The size of these microsatellites is assessed after PCR amplification and fragment analysis.

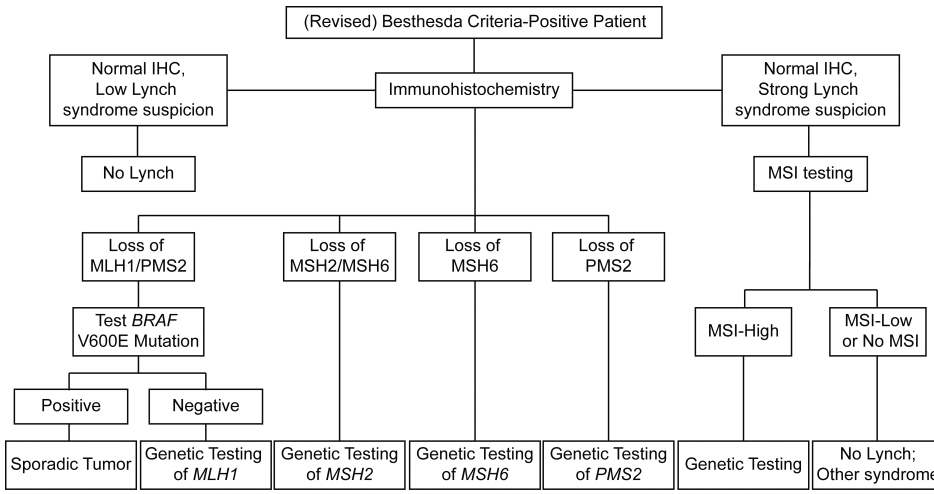


***Figure 4**. The Lynch syndrome diagnosic tree. Adapted from [Coetzee et al., 2013].*

Since not only LS tumors, but also 10-15% of all sporadic tumors express MSI, MSI is not highly specific for LS. MSI in sporadic cases is mostly a consequence of hypermethylation of the 5'CpG island of the *MLH1* promoter, leading to transcriptional silencing [Herman et al., 1998]. To screen against such tumors, MSI testing is often accompanied by testing for the *BRAF* V600E mutation. The presence of this mutation is associated with *MLH1* silencing but not with LS, therefore, wild type *BRAF* together with MSI suggests a germline MMR defect (Figure 4; [Domingo et al., 2005]).

Implications of Lynch Syndrome Diagnosis

As described, confirmed LS patients benefit from tight surveillance programs. In addition, LS patients benefit from chemoprevention, such as daily aspirin intake which significantly reduces the incidence of cancer [Burn et al., 2011]. Both smoking and a high body-mass index (BMI) increase CRC risk in LS patients [van Duijnhoven et al., 2013]. Even though consumption of red meat and alcohol increase the risk of sporadic CRC, such an association has not been shown for LS patients [van Duijnhoven et al., 2013].

LS patients benefit from stratified therapy [Hewish et al., 2010]. For instance, CRC is treated with adjuvant chemotherapy that includes 5-FU, amongst others [Devaud and Gallinger, 2013]. Unfortunately, 5-FU only improves survival of microsatellite-stable tumors and not of MSI tumors [Ribic et al., 2003], in agreement with functional data that shows that MMR-deficient cells are tolerant to 5-FU [Meyers et al., 2003; Liu et al., 2008]. Therefore, LS patients do not benefit from standard treatment regimens. Novel approaches to treat MMR deficient cancers are mostly based on synthetic lethality. Silencing of PTEN-induced putative kinase 1 (PINK1) [Martin et al., 2011] or polymerase β (POLβ) [Martin et al., 2010] is synthetically lethal with MMR deficiency, as is treatment with the drug methotrexate [Martin

et al., 2009]. All of these treatments lead to an increase in oxidative stress, suggesting that oxidative DNA-damaging agents may be utilized to selectively target MMR-deficient cancers [Brierley and Martin, 2013]. Of note, methotrexate induces amplification of the *DHFR* locus and of the adjacent *MSH3*. This increases MSH3 protein levels, which sequesters MSH2 into MutSβ and decreases MutSα levels, impairing MMR [Marra et al., 1998; Chang et al., 2002]. Methotrexate selectively retards MMR deficient (tumor) cells, but as an adverse effect may also induce MMR deficiency in normal somatic cells. In addition, the high levels of oxidative stress in these cells may be toxic, but could also further boost the mutator phenotype in the absence of MMR. Additional (mouse) experiments are required to validate the therapeutic potential of PINK1/POLß silencing and/or methotrexate treatment.

In addition to the mentioned therapies, MMR-deficient tumors may be treated by a selective gene therapy. This method is based on combining a harmless prodrug with a gene encoding a suicide enzyme. The gene is out-of-frame, resulting from a microsatellite within the gene. Only in MMR-deficient (tumor) cells the microsatellite size decreases, the suicide gene is rendered in frame and the suicide enzyme is produced, converting the prodrug into a toxic metabolite that kills the cells [Ferrás et al., 2009].

Variants of Uncertain Significance

Once a variant MMR allele has been identified, its relevance to the disease phenotype needs to be determined. In several classes of mutations, such as deletions, frameshifts or truncating mutations, this is achieved relatively easy. These mutant proteins usually lose their function and/or are degraded by the proteasome. For mutations that do not lead to protein truncations, such as silent mutations, missense mutations, intron variants or small in-frame insertion/deletions, classification is not straightforward. These variants are often referred to as Variants of Uncertain Significance (VUS). The lack of classification of these VUS precludes diagnosis for carriers and their relatives. Therefore, when a suspected Lynch syndrome patient carries a VUS, all first-degree relatives currently enroll lifelong periodic screening, irrespective of their mutation status [Castells et al., 2009]. Due to the inability to diagnose such variants, genetic testing is often not offered to the proband's relatives. The uncertainty of their disease status and the screening itself pose a physical and psychological burden on these families, but also a burden on the preventive health care apparatus.

Currently, 34% of all variants in MMR genes comprise missense mutations [Plazzer et al., 2013]. Inactivating missense mutations may currently remain under detected. The current LS diagnosis protocol only implies an MMR gene for genetic testing once loss of expression of the respective protein is shown in IHC, even though inactivating missense mutations do not necessarily destabilize a protein [Drost et al., 2013]. Now that DNA sequencing technologies are becoming cheaper and more widely applied, the stringency at which MMR genes are selected for sequencing may lower, and the incidence of MMR gene VUS is believed to further increase [Heinen, 2010; Rasmussen et al., 2012]. The development and use of robust and accessible approaches to identify pathogenic VUS will enable the implementation of targeted preventive and curative healthcare for carriers of pathogenic

MMR gene VUS. Meanwhile, unaffected relatives of such carriers can be liberated from the burden associated with the uncertain pathogenicity of the VUS while unwarranted inflow into preventive healthcare systems is reduced [Hennink et al., 2013].

Classification of VUS

VUS may be classified using a large set of criteria, such as:

- Segregation of the VUS with disease within affected families,
- *De novo* appearance of a VUS,
- Presence of the VUS in individuals with a second, obviously pathogenic, MMR gene defect arguing against the pathogenicity of the VUS,
- Absence of the VUS in control sets,
- Change in amino acid polarity or size,
- Presence of the VUS at a predicted splice site.

Unfortunately, such criteria cannot always be applied, since family- or segregation data are often absent, or families are too small to calculate segregation. Even if such correlations can be made, they do not provide definitive proof of pathogenicity. Segregation of a VUS, for example, only suggests that the allele segregates with disease. Therefore, additional genetic factors on that allele, other than the VUS, may cause cancer predisposition. Definite proof of pathogenicity can be provided by assays that test the functional activity of the VUS.

AIM OF THIS THESIS

The inability to assess the pathogenicity of VUS in DNA Mismatch Repair genes is a bottleneck in LS diagnosis. Lack of pathogenic assessment precludes diagnosis for carriers and their relatives, precludes targeted preventive and curative healthcare for carriers of pathogenic VUS and clutters preventive healthcare systems. For these reasons, there is a need for convenient functional assays for the diagnostic assessment of MMR gene VUS. This thesis focuses on the development such assays, while making use of the well-established genetics and biochemistry of the MMR pathway.

The thesis starts with a discussion of recent advances in functional analysis of Variants of Uncertain Significance in DNA Mismatch Repair genes (**Chapter 2;** [Rasmussen et al., 2012]). In **Chapter 3**, I describe the development of a functional assay based on measuring MMR activity of *MSH2* and *MSH6* VUS *in vitro* that allows for a rapid assessment of pathogenicity [Drost et al., 2012]. A similar assay is applied for VUS in *MLH1* [Drost et al., 2010] and *PMS2*, which is discussed in **Chapters 4** and **5**, respectively. **Chapter 6** describes a different approach for the functional analysis of MMR gene VUS. Here, a genetic screen is developed and experimentally validated that systematically identifies residues essential for *in vivo* function of *Msh2* [Drost et al., 2013]. Finally, in **Chapter 7** a similar genetic screen is applied to *Msh6*, and the generated cell lines are used to study mismatch binding and intramolecular signaling in MutSα.

REFERENCES

Alhopuro P, Sammalkorpi H, Niittymäki I, Biström M, Raitila A, Saharinen J, Nousiainen K, Lehtonen HJ, Heliövaara E, Puhakka J, Tuupanen S, Sousa S et al. 2012. Candidate driver genes in microsatellite-unstable colorectal cancer. Int J Cancer 130: 1558-1566.

Antony E, Hingorani MM. 2003. Mismatch recognition-coupled stabilization of Msh2-Msh6 in an ATP-bound state at the initiation of DNA repair. Biochemistry 42: 7682-7693.

Antony E, Khubchandani S, Chen S, Hingorani MM. 2006. Contribution of Msh2 and Msh6 subunits to the asymmetric ATPase and DNA mismatch binding activities of Saccharomyces cerevisiae Msh2-Msh6 mismatch repair protein. DNA Repair 5: 153-162.

Arana ME, Kunkel TA. 2010. Mutator phenotypes due to DNA replication infidelity. Semin Cancer Biol 20: 304-311.

Baker SM, Plug AW, Prolla TA, Bronner CE, Harris AC, Yao X, Christie DM, Monell C, Arnheim N, Bradley A, Ashley T, Liskay RM. 1996. Involvement of mouse Mlh1 in DNA mismatch repair and meiotic crossing over. Nat Genet 13: 336-342.

Boland CR, Goel A. 2010. Microsatellite instability in colorectal cancer. Gastroenterology 138: 2073-2087.e3.

Boland CR, Lynch HT. 2013. The History of Lynch Syndrome. Fam Cancer 12:147-157.

Bonadona V, Bonaïti B, Olschwang S, Grandjouan S, Huiart L, Longy M, Guimbaud R, Buecher B, Bignon YJ, Caron O, Colas C, Noguès C et al. 2011. Cancer risks associated with germline mutations in MLH1, MSH2, and MSH6 genes in Lynch syndrome. JAMA 305: 2304-2310.

Borgdorff V, van Hees-Stuivenberg S, Meijers CM, de Wind N. 2005. Spontaneous and mutagen-induced loss of DNA mismatch repair in Msh2-heterozygous mammalian cells. Mutat Res 574: 50-57.

Brierley DJ, Martin SA. 2013. Oxidative stress and the DNA mismatch repair pathway. Antioxid Redox Signal 18: 2420-2428.

Bronner CE, Baker SM, Morrison PT, Warren G, Smith LG, Lescoe MK, Kane M, Earabino C, Lipford J, Lindblom A. 1994. Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. Nature 368: 258-261.

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

Burn J, Gerdes AM, Macrae F, Mecklin JP, Moeslein G, Olschwang S, Eccles D, Evans DG, Maher ER, Bertario L, Bisgaard ML, Dunlop MG et al. 2011. Long-term effect of aspirin on cancer risk in carriers of hereditary colorectal cancer: an analysis from the CAPP2 randomised controlled trial. Lancet 378: 2081-2087.

Cannavo E, Marra G, Sabates-Bellver J, Menigatti M, Lipin 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 Res 65: 10759-10766

Castells A, Castellví-Bel S, Balaguer F. 2009. Concepts in familial colorectal cancer: where do we stand and what is the future? Gastroenterology 137: 404-409.

Cejka P, Mojas N, Gillet L, Schär P, Jiricny J. 2005. Homologous recombination rescues mismatch-repair-dependent cytotoxicity of S(N)1-type methylating agents in S. cerevisiae. Curr Biol 15: 1395-1400.

Chang CL, Marra G, Chauhan DP, Ha HT, Chang DK, Ricciardiello L, Randolph A, Carethers JM, Boland CR. 2002. Oxidative stress inactivates the human DNA mismatch repair system. Am J Physiol, Cell Physiol 283: C148-54.

Chao EC, Lipkin SM. 2006. Molecular models for the tissue specificity of DNA mismatch repair-deficient carcinogenesis. Nucleic Acids Res 34: 840-852.

Chen LM, Yang KY, Little SE, Cheung MK, Caughey AB. 2007. Gynecologic cancer prevention in Lynch syndrome/hereditary nonpolyposis colorectal cancer families. Obstet Gynecol 110: 18-25.

Ciccia A, Elledge SJ. 2010. The DNA damage response: making it safe to play with knives. Mol Cell 40: 179-204.

Coetzee E, Algar U, Goldberg P. 2013. New Insights into Lynch Syndrome Diagnosis. DNA Alterations in Lynch Syndrome (Book).

Constantin N, Dzantiev L, Kadyrov FA, Modrich PL. 2005. Human mismatch repair: reconstitution of a nick-directed bidirectional reaction. J Biol Chem 280: 39752-39761.

Cox EC, Degnen GE, Scheppe ML. 1972. Mutator gene studies in Escherichia coli: the mutS gene. Genetics 72: 551-567.

de la Chapelle A, Hampel H. 2010. Clinical relevance of microsatellite instability in colorectal cancer. J Clin Oncol 28: 3380-3387.

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

de Wind N, Dekker M, Claij N, Jansen L, van Klink Y, Radman M, Riggins G, van der Valk M, van't Wout K, te Riele H. 1999. HNPCC-like cancer predisposition in mice through simultaneous loss of Msh3 and Msh6 mismatch-repair protein functions. Nat Genet 23: 359-362.

Deans AJ, West SC. 2011. DNA interstrand crosslink repair and cancer. Nat Rev Cancer 11: 467-480.

Debiak M, Nikolova T, Kaina B. 2004. Loss of ATM sensitizes against O6-methylguanine triggered apoptosis, SCEs and chromosomal aberrations. DNA Repair 3: 359-368.

Devaud N, Gallinger S. 2013. Chemotherapy of MMR-deficient colorectal cancer. Fam Cancer 12: 301-306.

Domingo E, Niessen RC, Oliveira C, Alhopuro P, Moutinho C, Espín E, Armengol M, Sijmons RH, Kleibeuker JH, Seruca R, Aaltonen LA, Imai K et al. 2005. BRAF-V600E is not involved in the colorectal tumorigenesis of HNPCC in patients with functional MLH1 and MSH2 genes. Oncogene 24: 3995-3998.

Drost M, Lutzen A, Van Hees S, Ferreira D, Calleja F, Zonneveld J, Nielsen F, Rasmussen L, de Wind N. 2013. Genetic screens to identify pathogenic gene variants in the common cancer predisposition Lynch syndrome. Proc Natl Acad Sci USA 110: 9403-9408.

Drost M, Zonneveld JB, Van Dijk L, Morreau H, Tops CM, Vasen HFA, Wijnen J, De Wind N. 2010. A cell-free assay for the functional analysis of variants of the mismatch repair protein MLH1. Hum Mutat 31: 247-253.

Drost M, Zonneveld JB, van Hees S, Rasmussen LJ, Hofstra RM, De Wind N. 2012. A rapid and cell-free assay to test the activity of lynch syndrome-associated MSH2 and MSH6 missense variants. Hum Mutat 33: 488-494.

Duckett DR, Drummond JT, Murchie AI, Reardon JT, Sancar A, Lilley DM, Modrich P. 1996. Human MutSalpha recognizes damaged DNA base pairs containing O6-methylguanine, O4-methylthymine, or the cisplatin-d(GpG) adduct. Proc Natl Acad Sci USA 93: 6443-6447.

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

Edelmann L, Edelmann W. 2004. Loss of DNA mismatch repair function and cancer predisposition in the mouse: animal models for human hereditary nonpolyposis colorectal cancer. American journal of medical genetics Part C, Seminars in medical genetics 129: 91-99.

Edelmann W, Cohen PE, Kane M, Lau K, Morrow B, Bennett S, Umar A, Kunkel T, Cattoretti G, Chaganti R, Pollard JW, Kolodner RD et al. 1996. Meiotic pachytene arrest in MLH1-deficient mice. Cell 85: 1125-1134.

Edelmann W, Yang K, Umar A, Heyer J, Lau K, Fan K, Liedtke W, Cohen PE, Kane MF, Lipford JR, Yu N, Crouse GF et al. 1997. Mutation in the mismatch repair gene Msh6 causes cancer susceptibility. Cell 91: 467-477.

Engel C, Loeffler M, Steinke V, Rahner N, Holinski-Feder E, Dietmaier W, Schackert HK, Goergens H, von Knebel Doeberitz M, Goecke TO, Schmiegel W, Buettner R et al. 2012. Risks of less common cancers in proven mutation carriers with lynch syndrome. J Clin Oncol 30: 4409-4415.

Ferrás C, Oude Vrielink JA, Verspuy JW, te Riele H, Tsaalbi-Shtylik A, de Wind N. 2009. Abrogation of Microsatellite-instable Tumors Using a Highly Selective Suicide Gene/Prodrug Combination. Mol Ther 17: 1373-1380

Fishel R, Lescoe MK, Rao MR, Copeland NG, Jenkins NA, Garber J, Kane M, Kolodner R. 1993. The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. Cell 75: 1027-1038.

Genschel J, Modrich P. 2009. Functions of MutLalpha, replication protein A (RPA), and HMGB1 in 5'-directed mismatch repair. J Biol Chem 284: 21536-21544.

Ghodgaonkar MM, Lazzaro F, Olivera-Pimentel M, Artola-Borán M, Cejka P, Reijns MA, Jackson AP, Plevani P, Muzi-Falconi M, Jiricny J. 2013. Ribonucleotides misincorporated into DNA act as strand-discrimination signals in eukaryotic mismatch repair. Mol Cell 50: 323-332.

Giráldez MD, Castellví-Bel S, Balaguer F, Gonzalo V, Ocaña T, Castells A. 2008. Lynch syndrome in colorectal cancer patients. Expert Rev Anticancer Ther 8: 573-583.

Glaab WE, Risinger JI, Umar A, Barrett JC, Kunkel TA, Tindall KR. 1998. Resistance to 6-thioguanine in mismatch repair-deficient human cancer cell lines

correlates with an increase in induced mutations at the HPRT locus. Carcinogenesis 19: 1931-1937.

Goel A, Nagasaka T, Hamelin R, Boland CR. 2010. An optimized pentaplex PCR for detecting DNA mismatch repair-deficient colorectal cancers. PLoS ONE 5: e9393.

Goldmacher VS, Cuzick RA, Thilly WG. 1986. Isolation and partial characterization of human cell mutants differing in sensitivity to killing and mutation by methylnitrosourea and N-methyl-N'-nitro-N-nitrosoguanidine. J Biol Chem 261: 12462-12471.

Gorman J, Chowdhury A, Surtees JA, Shimada J. 2007. Dynamic Basis for One-Dimensional DNA Scanning by the Mismatch Repair Complex Msh2-Msh6. Mol Cell 28: 359-370

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. Mol Cell 3: 255-261.

Gruber S. 2006. New Developments in Lynch Syndrome (Hereditary Nonpolyposis Colorectal Cancer) and Mismatch Repair Gene Testing. Gastroenterology 130: 577-587.

Gueneau E, Dherin C, Legrand P, Tellier-Lebegue C, Gilquin B, Bonnesoeur P, Londino F, Quemener C, Le Du MH, Márquez JA, Moutiez M, Gondry M et al. 2013. Structure of the MutLα C-terminal domain reveals how Mlh1 contributes to Pms1 endonuclease site. Nat Struct Mol Biol 20: 461-468.

Gupta S, Gellert M, Yang W. 2012. Mechanism of mismatch recognition revealed by human MutSβ bound to unpaired DNA loops. Nat Struct Mol Biol 19: 72-78.

Hampel H, Frankel WL, Martin E, Arnold M, Khanduja K, Kuebler P, Nakagawa H, Sotamaa K, Prior TW, Westman J, Panescu J, Fix D et al. 2005. Screening for the Lynch syndrome (hereditary nonpolyposis colorectal cancer). N Engl J Med 352: 1851-1860.

Hanahan D, Weinberg RA. 2000. The hallmarks of cancer. Cell 100: 57-70.

Hanahan D, Weinberg RA. 2011. Hallmarks of cancer: the next generation. Cell 144: 646-674.

Hays JB, Hoffman PD, Wang H. 2005. Discrimination and versatility in mismatch repair. DNA Repair 4: 1463-1474.

Heinen CD. 2010. Genotype to phenotype: analyzing the effects of inherited mutations in colorectal cancer families. Mutat Res 693: 32-45.

Hendriks Y, De Jong A, Morreau H, Tops CM, Vasen HFA, Wijnen J, Breuning M, Brocker-Vriends A. 2006. Diagnostic Approach and Management of Lynch Syndrome (Hereditary Nonpolyposis Colorectal Carcinoma): A Guide for Clinicians. CA: A Cancer Journal for Clinicians 56: 213-225.

Hennink SD, Hofland N, Gopie JP, van der Kaa C, de Koning K, Nielsen M, Tops C, Morreau H, de Vos Tot Nederveen Cappel WH, Langers AM, Hardwick JC, Gaarenstroom KN et al. 2013. Value-based healthcare in Lynch syndrome. Fam Cancer 12: 347-354.

Herman JG, Umar A, Polyak K, Graff JR, Ahuja N, Issa JP, Markowitz S, Willson JK, Hamilton SR, Kinzler KW, Kane MF, Kolodner RD et al. 1998. Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. Proc Natl Acad Sci USA 95: 6870-6875.

Hewish M, Lord CJ, Martin SA, Cunningham D, Ashworth A. 2010. Mismatch repair deficient colorectal cancer in the era of personalized treatment. Nat Rev Clin Oncol 7: 197-208.

Hitchins MP, Rapkins RW, Kwok CT, Srivastava S, Wong JJ, Khachigian LM, Polly P, Goldblatt J, Ward RL. 2011. Dominantly inherited constitutional epigenetic silencing of MLH1 in a cancer-affected family is linked to a single nucleotide variant within the 5'UTR. Cancer Cell 20: 200-213.

Hoeijmakers JH. 2009. DNA damage, aging, and cancer. N Engl J Med 361: 1475-1485.

Hoffman PD, Wang H, Lawrence CW, Iwai S, Hanaoka F, Hays JB. 2005. Binding of MutS mismatch repair protein to DNA containing UV photoproducts, "mismatched" opposite Watson--Crick and novel nucleotides, in different DNA sequence contexts. DNA Repair 4: 983-993.

Hofstra RM, Spurdle AB, Eccles D, Foulkes WD, De Wind N, Hoogerbrugge N, Hogervorst FB, Group IARCUGVW. 2008. Tumor characteristics as an analytic tool for classifying genetic variants of uncertain clinical significance. Hum Mutat 29: 1292-1303.

Holland IB, Blight MA. 1999. ABC-ATPases, adaptable energy generators fuelling transmembrane movement of a variety of molecules in organisms from bacteria to humans. J Mol Biol 293: 381-399.

Hombauer H, Campbell C, Smith C, Desai A, Kolodner R. 2011a. Visualization of Eukaryotic DNA Mismatch Repair Reveals Distinct Recognition and Repair Intermediates. Cell 147: 1040-1053.

Hombauer H, Srivatsan A, Putnam C, Kolodner R. 2011b. Mismatch Repair, But Not Heteroduplex Rejection, Is Temporally Coupled to DNA Replication. Science 334: 1713-1716.

Hsieh P, Yamane K. 2008. DNA mismatch repair: molecular mechanism, cancer, and ageing. Mech Ageing Dev 129: 391-407.

Iyer RR, Pluciennik A, Burdett V, Modrich PL. 2006. DNA mismatch repair: functions and mechanisms. Chem Rev 106: 302-323.

Iyer RR, Pohlhaus TJ, Chen S, Hura GL, Dzantiev L, Beese LS, Modrich PL. 2008. The MutSalpha-proliferating cell nuclear antigen interaction in human DNA mismatch repair. J Biol Chem 283: 13310-13319.

Jackson SP, Bartek J. 2009. The DNA-damage response in human biology and disease. Nature 461: 1071-1078.

Jass JR. 2004. Role of the pathologist in the diagnosis of hereditary non-polyposis colorectal cancer. Dis Markers 20: 215-224.

Jiricny J. 2013. Postreplicative mismatch repair. Cold Spring Harbor Perspectives in Biology 5.

Jiricny J, Su SS, Wood SG, Modrich P. 1988. Mismatch-containing oligonucleotide duplexes bound by the E. coli mutS-encoded protein. Nucleic Acids Res 16: 7843-7853.

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

Kadyrov FA, Genschel J, Fang Y, Penland E, Edelmann W, Modrich P. 2009. A possible mechanism for exonuclease 1-independent eukaryotic mismatch repair. Proc Natl Acad Sci USA 106: 8495-8500.

Kaina B, Christmann M, Naumann S, Roos WP. 2007. MGMT: key node in the battle against genotoxicity, carcinogenicity and apoptosis induced by alkylating agents. DNA Repair 6: 1079-1099.

Kat A, Thilly WG, Fang WH, Longley MJ, Li GM, Modrich P. 1993. An alkylation-tolerant, mutator human cell line is deficient in strand-specific mismatch repair. Proc Natl Acad Sci USA 90: 6424-6428.

Kempers MJ, Kuiper RP, Ockeloen CW, Chappuis PO, Hutter P, Rahner N, Schackert HK, Steinke V, Holinski-Feder E, Morak M, Kloor M, Büttner R et al. 2011. Risk of colorectal and endometrial cancers in EPCAM deletion-positive Lynch syndrome: a cohort study. Lancet Oncol 12: 49-55.

Klapacz J, Meira LB, Luchetti DG, Calvo JA, Bronson RT, Edelmann W, Samson LD. 2009. O6-methylguanine-induced cell death involves exonuclease 1 as well as DNA mismatch recognition in vivo. Proc Natl Acad Sci USA 106: 576-581

Kloor M, Huth C, Voigt AY, Benner A, Schirmacher P, von Knebel Doeberitz M, Bläker H. 2012. Prevalence of mismatch repair-deficient crypt foci in Lynch syndrome: a pathological study. Lancet Oncol 13: 598-606.

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

Labazi M, Jaafar L, Flores-Rozas H. 2009. Modulation of the DNA-binding activity of Saccharomyces cerevisiae MSH2-MSH6 complex by the high-mobility group protein NHP6A, in vitro. Nucleic Acids Res 37 :7581-7589

Lagerstedt Robinson K, Liu T, Vandrovcova J, Halvarsson B, Clendenning M, Frebourg T, Papadopoulos N, Kinzler KW, Vogelstein B, Peltomäki P, Kolodner RD, Nilbert M et al. 2007. Lynch syndrome (hereditary nonpolyposis colorectal cancer) diagnostics. J Natl Cancer Inst 99: 291-299.

Lahue RS, Au KG, Modrich PL. 1989. DNA mismatch correction in a defined system. Science 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.

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.

Li F, Mao G, Tong D, Huang J, Gu L, Yang W, Li GM. 2013. The Histone Mark H3K36me3 Regulates Human DNA Mismatch Repair through Its Interaction with MutSα. Cell 153: 590-600.

Li GM. 2008. Mechanisms and functions of DNA mismatch repair. Cell Res 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 et al. 2009. Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3' exons of TACSTD1. Nat Genet 41: 112-117.

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

Liu A, Yoshioka K, Salerno V, Hsieh P. 2008. The mismatch repair-mediated cell cycle

checkpoint response to fluorodeoxyuridine. J Cell Biol 105: 245-254.

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

Liu Y, Fang Y, Shao H, Lindsey-Boltz L, Sancar A, Modrich P. 2009. Interactions of human mismatch repair proteins MutS and MutL with proteins of the ATR-Chk1 pathway. J Biol Chem 285: 5974-5982

Liu Y, Kadyrov FA, Modrich P. 2011. PARP-1 enhances the mismatch-dependence of 5'-directed excision in human mismatch repair in vitro. DNA Repair 10: 1145-1153.

Longley MJ, Pierce AJ, Modrich PL. 1997. DNA Polymerase δ Is Required for Human Mismatch Repair in Vitro. J Biol Chem 272: 10917-10921

Lujan SA, Williams JS, Clausen AR, Clark AB, Kunkel TA. 2013. Ribonucleotides are signals for mismatch repair of leading-strand replication errors. Mol Cell 50: 437-443.

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

Macpherson P, Barone F, Maga G, Mazzei F, Karran P, Bignami M. 2005. 8-oxoguanine incorporation into DNA repeats in vitro and mismatch recognition by MutSalpha. Nucleic Acids Res 33: 5094-5105.

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

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

Martin SA, Hewish M, Sims D, Lord CJ, Ashworth A. 2011. Parallel high-throughput RNA interference screens identify PINK1 as a potential therapeutic target for the treatment of DNA mismatch repair-deficient cancers. Cancer Res 71: 1836-1848.

Martin SA, Mccabe N, Mullarkey M, Cummins R, Burgess DJ, Nakabeppu Y, Oka S, Kay E, Lord CJ, Ashworth A. 2010. DNA polymerases as potential therapeutic targets for cancers deficient in the DNA mismatch repair proteins MSH2 or MLH1. Cancer Cell 17: 235-248.

Martin SA, McCarthy A, Barber LJ, Burgess DJ, Parry S, Lord CJ, Ashworth A. 2009. Methotrexate induces oxidative DNA damage and is selectively lethal to tumour cells with defects in the DNA mismatch repair gene MSH2. EMBO Mol Med 1: 323-337.

Mazurek A, Berardini M, Fishel R. 2002. Activation of human MutS homologs by 8-oxo-guanine DNA damage. J Biol Chem 277: 8260-8266.

McMurray CT. 2010. Mechanisms of trinucleotide repeat instability during human development. Nat Rev Genet 11: 786-799.

Medina-Arana V, Delgado L, Bravo A, Martín J, Fernández-Peralta AM, González-Aguilera JJ. 2012. Tumor spectrum in lynch syndrome, DNA mismatch repair system and endogenous carcinogens. J Surg Oncol 106: 10-16.

Meyers M, Hwang A, Wagner MW, Bruening AJ, Veigl ML, Sedwick WD, Boothman DA. 2003. A role for DNA mismatch repair in sensing and responding to fluoropyrimidine damage. Oncogene 22: 7376-7388.

Mojas N, Lopes M, Jiricny J. 2007. Mismatch repair-dependent processing of methylation damage gives rise to persistent single-stranded gaps in newly replicated DNA. Genes Dev 21: 3342-3355.

Moreira L, Balaguer F, Lindor N, de la Chapelle A, Hampel H, Aaltonen LA, Hopper JL, Le Marchand L, Gallinger S, Newcomb P, Haile R, Thibodeau SN et al. 2012. Identification of Lynch syndrome among patients with colorectal cancer. JAMA 308: 1555-1565.

Nicolaides NC, Papadopoulos N, Liu B, Wei YF, Carter KC, Ruben SM, Rosen CA, Haseltine WA, Fleischmann RD, Fraser CM. 1994. Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. Nature 371: 75-80.

Niessen RC, Berends MJ, Wu Y, Sijmons RH, Hollema H, Ligtenberg MJ, de Walle HE, de Vries EG, Karrenbeld A, Buys CH, van der Zee AG, Hofstra RM et al. 2006. Identification of mismatch repair gene mutations in young patients with colorectal cancer and in patients with multiple tumours associated with hereditary non-polyposis colorectal cancer. Gut 55: 1781-1788.

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.

Pabla N, Ma Z, Mcilhatton M, Fishel R, Dong Z. 2011. hMSH2 Recruits ATR to DNA Damage Sites for Activation during DNA Damage-induced Apoptosis. J Biol Chem 286: 10411-10418.

Pavlov YI, Mian IM, Kunkel TA. 2003. Evidence for preferential mismatch repair of lagging strand DNA replication errors in yeast. Curr Biol 13: 744-748.

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

Peña-Diaz J, Jiricny J. 2010. PCNA and MutLα: partners in crime in triplet repeat expansion? Proc Natl Acad Sci USA 107: 16409-16410.

Plazzer JP, Sijmons RH, Woods MO, Peltomäki P, Thompson B, Den Dunnen JT, Macrae F. 2013. The InSiGHT database: utilizing 100 years of insights into Lynch Syndrome. Fam Cancer 12: 175-180.

Pluciennik A, Dzantiev L, Iyer RR, Constantin N, Kadyrov FA, Modrich P. 2010. PCNA function in the activation and strand direction of MutLα endonuclease in mismatch repair. Proc Natl Acad Sci USA 107: 16066-16071.

Prolla TA, Baker SM, Harris AC, Tsao JL, Yao X, Bronner CE, Zheng B, Gordon M, Reneker J, Arnheim N, Shibata D, Bradley A et al. 1998. Tumour susceptibility and spontaneous mutation in mice deficient in Mlh1, Pms1 and Pms2 DNA mismatch repair. Nat Genet 18: 276-279.

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 et al. 2005. Functional significance and clinical phenotype of nontruncating mismatch repair variants of MLH1. Gastroenterology 129: 537-549.

Rasmussen LJ, Heinen CD, Royer-Pokora B, Drost M, Tavtigian S, Hofstra RM, De Wind N. 2012. Pathological assessment of mismatch repair gene variants in Lynch syndrome: Past, present, and future. Hum Mutat 33: 1617-1625.

Rasmussen LJ, Samson LD. 1996. The Escherichia coli MutS DNA mismatch binding protein specifically binds O(6)-methylguanine DNA lesions. Carcinogenesis 17: 2085-2088.

Ribic CM, Sargent DJ, Moore MJ, Thibodeau SN, French AJ, Goldberg RM, Hamilton SR, Laurent-Puig P, Gryfe R, Shepherd LE, Tu D, Redston M et al. 2003. Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer. N Engl J Med 349: 247-257.

Sacho EJ, Kadyrov FA, Modrich PL, Kunkel TA, Erie DA. 2008. Direct Visualization of Asymmetric Adenine Nucleotide-Induced Conformational Changes in MutLalpha. Mol Cell 29: 112-121.

Schaetzlein S, Chahwan R, Avdievich E, Roa S, Wei K, Eoff RL, Sellers RS, Clark AB, Kunkel TA, Scharff MD, Edelmann W. 2013. Mammalian Exo1 encodes both structural and catalytic functions that play distinct roles in essential biological processes. Proc Natl Acad Sci USA 110: E2470-9.

Senter L, Clendenning M, Sotamaa K, Hampel H, Green J, Potter JD, Lindblom A, Lagerstedt K, Thibodeau SN, Lindor NM, Young J, Winship I et al. 2008. The clinical phenotype of Lynch syndrome due to germ-line PMS2 mutations. Gastroenterology 135: 419-428.

Sjursen W, Haukanes BI, Grindedal EM, Aarset H, Stormorken A, Engebretsen LF, Jonsrud C, Bjørnevoll I, Andresen PA, Ariansen S, Lavik LA, Gilde B et al. 2010. Current clinical criteria for Lynch syndrome are not sensitive enough to identify MSH6 mutation carriers. J Med Genet 47: 579-585.

Stoffel E, Mukherjee B, Raymond VM, Tayob N, Kastrinos F, Sparr J, Wang F, Bandipalliam P, Syngal S, Gruber SB. 2009. Calculation of Risk of Colorectal and Endometrial Cancer Among Patients With Lynch Syndrome. Gastroenterology 137: 1621-1627

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

Stojic L, Cejka P, Jiricny J. 2005. High doses of SN1 type methylating agents activate DNA damage signaling cascades that are largely independent of mismatch repair. Cell Cycle 4: 473-477.

Stojic L, Mojas N, Cejka P, di Pietro M, Ferrari S, Jiricny J. 2004. Mismatch repair-dependent G2 checkpoint induced by low doses of SN1 type methylating agents requires the ATR kinase. Genes Dev 18: 1331-1344.

Strand M, Prolla TA, Liskay RM, Petes TD. 1993. Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. Nature 365: 274-276.

Swann PF, Waters TR, Moulton DC, Xu YZ, Zheng Q, Edwards M, Mace R. 1996. Role of postreplicative DNA mismatch repair in the cytotoxic action of thioguanine. Science 273: 1109-1111.

Taketo MM, Edelmann W. 2009. Mouse models of colon cancer. Gastroenterology 136: 780-798.

Toft NJ, Winton DJ, Kelly J, Howard LA, Dekker M, te Riele H, Arends MJ, Wyllie AH, Margison GP, Clarke AR. 1999. Msh2 status modulates both apoptosis and mutation frequency in the murine small intestine. Proc Natl Acad Sci USA 96: 3911-3915.

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

van Duijnhoven FJ, Botma A, Winkels R, Nagengast FM, Vasen HF, Kampman E. 2013. Do lifestyle factors influence colorectal cancer risk in Lynch syndrome? Fam Cancer 12: 285-293.

Vasen HF, Blanco I, Aktan-Collan K, Gopie JP, Alonso A, Aretz S, Bernstein I, Bertario L, Burn J, Capella G, Colas C, Engel C et al. 2013. Revised guidelines for the clinical management of Lynch syndrome (HNPCC): recommendations by a group of European experts. Gut 62: 812–823

Warren JJ, Forsberg LJ, Beese LS. 2006. The structural basis for the mutagenicity of O(6)-methyl-guanine lesions. Proc Natl Acad Sci USA 103: 19701-19706.

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

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

Wei K, Kucherlapati R, Edelmann W. 2002. Mouse models for human DNA mismatch-repair gene defects. Trends in molecular medicine 8: 346-353.

Win AK, Macinnis RJ, Dowty JG, Jenkins MA. 2013. Criteria and prediction models for mismatch repair gene mutations: a review. J Med Genet 50:785–793

Wu Y, Berends MJ, Post JG, Mensink RG, Verlind E, Van Der Sluis T, Kempinga C, Sijmons RH, van der Zee AG, Hollema H, Kleibeuker JH, Buys CH et al. 2001a. Germline mutations of EXO1 gene in patients with hereditary nonpolyposis colorectal cancer (HNPCC) and atypical HNPCC forms. Gastroenterology 120: 1580-1587.

Wu Y, Berends MJ, Sijmons RH, Mensink RG, Verlind E, Kooi KA, Van Der Sluis T, Kempinga C, van Der Zee AG, Hollema H, Buys CH, Kleibeuker JH et al. 2001b. A role for MLH3 in hereditary nonpolyposis colorectal cancer. Nat Genet 29: 137-138.

Yamane K, Schupp JE, Kinsella TJ. 2007. BRCA1 activates a G2-M cell cycle checkpoint following 6-thioguanine-induced DNA mismatch damage. Cancer Res 67: 6286-6292.

York SJ, Modrich PL. 2006. Mismatch Repair-dependent Iterative Excision at Irreparable O6-Methylguanine Lesions in Human Nuclear Extracts. J Biol Chem 281: 22674-22683

Yoshioka K, Yoshioka Y, Hsieh P. 2006. ATR kinase activation mediated by MutSalpha and MutLalpha in response to cytotoxic O6-methylguanine adducts. Mol Cell 22: 501-510.

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

Zhang Y, Yuan F, Wang D, Gu L, Li GM. 2008. Identification of RFX as a novel mismatch repair stimulatory factor. J Biol Chem 283: 12730-12735

Pathological Assessment of Mismatch Repair Gene Variants in Lynch Syndrome: Past, Present, and Future

Lene Juel Rasmussen, Christopher D. Heinen, Brigitte Royer-Pokora, Mark Drost, Sean Tavtigian, Robert M.W. Hofstra, and Niels de Wind

2

Adapted from Rasmussen et. al 2012 Hum Mutat 33:1617–1625

ABSTRACT

Lynch syndrome (LS) is caused by germline mutations in DNA mismatch repair (MMR) genes and is the most prevalent hereditary colorectal cancer syndrome. A significant proportion of variants identified in MMR and other common cancer susceptibility genes are missense or noncoding changes whose consequences for pathogenicity cannot be easily interpreted. Such variants are designated as "variants of uncertain significance" (VUS). Management of LS can be significantly improved by identifying individuals who carry a pathogenic variant and thus benefit from screening, preventive, and therapeutic measures. Also, identifying family members that do not carry the variant is important so they can be released from the intensive surveillance. Determining which genetic variants are pathogenic and which are neutral is a major challenge in clinical genetics. The profound mechanistic knowledge on the genetics and biochemistry of MMR enables the development and use of targeted assays to evaluate the pathogenicity of variants found in suspected patients with LS. We describe different approaches for the functional analysis of MMR gene VUS and propose development of a validated diagnostic framework. Furthermore, we call attention to common misconceptions about functional assays and endorse development of an integrated approach comprising validated assays for diagnosis of VUS in patients suspected of LS.

INTRODUCTION

Mismatch Repair Genes and Cancer

Lynch syndrome (LS; also often named hereditary nonpolyposis colorectal cancer, HNPCC) is an autosomal dominant hereditary cancer syndrome caused by heterozygous germ line mutations in one of the major DNA mismatch repair (MMR) genes MSH2 (MIM# 609309) and MLH1 (MIM# 120436), but also MHS6 (MIM# 600678) and PMS2 (MIM# 600259) [Li, 2008; Peltomäki, 2001, 2003]. Sporadic somatic mutation or loss of heterozygosity of the wild-type allele leaves the cell with a defective MMR system. Because MMR repairs misincorporations during replication, its loss results in the accumulation of spontaneous mutations (a mutator phenotype) that are thought to drive, or contribute to, the neoplastic transformation process. The majority of MMR gene mutations that currently are detected in patients with LS cause truncations or deletions and can be securely classified as pathogenic. However, missense mutations comprise a significant proportion of all mutations (30–40%) [de La Chapelle, 2004] and the functional consequences of many of these so-called variants of uncertain significance (VUS) are not clear. It is a major and urgent challenge to develop validated tools to securely classify MMR VUS, whose prevalence is expected to rise sharply with the advent of next-generation sequencing.

DNA Mismatch Repair

MMR is a postreplicative DNA repair system that corrects errors made by the DNA polymerases (Pol) δ or ε during replication of the bulk genome, including DNA base-base mispairs, caused by accidental misincorporation, and insertion/deletion loops (IDL) that arise during slippage of polymerases at simple sequence repeats (also called microsatellites) [Kunkel and Erie, 2005; Li, 2008]. MMR thereby protects the cells from spontaneous mutations, including microsatellite instability (MSI). The MMR pathway is initiated by the recognition of a mismatch or IDL by either the MutSα heterodimer, consisting of the MSH2 and MSH6 proteins (Fig. 1), or the MutSβ heterodimer consisting of MSH2 and MSH3. MutSα is responsible for the recognition of base-base mispairs and small IDLs whereas MutSβ is responsible for the redundant repair of IDLs. Upon DNA mismatch recognition MutSα (or MutSβ) recruits the MutLα heterodimer (consisting of MLH1-PMS2). MutLα binds several MMR proteins and modulates their activity in a mismatch-dependent manner [Kunkel and Erie, 2005; Li, 2008]. PMS2 introduces a nick in the daughter strand, 5' or 3' of the mismatch [Kadyrov et al., 2006]. This nick serves as entry point for the 5'-exonuclease EXO1 that carry out excision of the misincorporation-containing DNA strand. It should be noted that MMR can occur in the absence of EXO1 [Kadyrov et al., 2009]. Replication Protein A (RPA) protects the MMR excision intermediate from nuclease degradation. The excised DNA strand is resynthesized by Pol δ or ε in the presence of PCNA, the clamp loader RFC and RPA [Kunkel and Erie, 2005; Li, 2008].

A Diagnostic Tree for Assessing the Pathogenicity of VUS in MMR Genes

To develop a widely applicable and standardized procedure for assessing the pathogenicity of VUS in MMR genes, we have previously proposed the development of a well-defined and hierarchical three-step diagnostic tree (see Fig. 2) [Couch et al., 2008]. The first step is

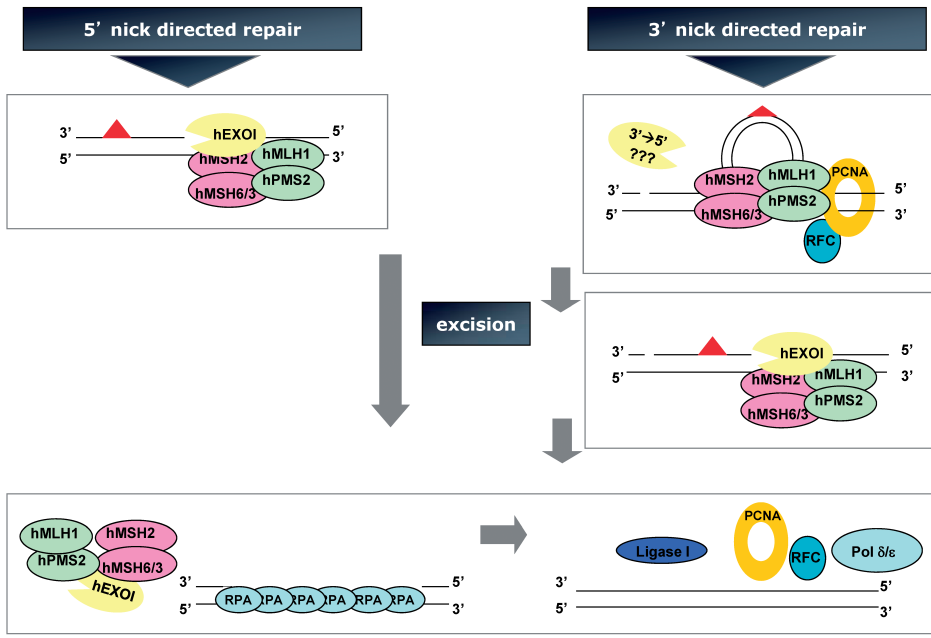


Figure 1. *Steps in MMR. The replicative DNA polymerase misincorporates a nucleotide during DNA replication. The MSH2–MSH6 heterodimer recognizes and binds to the mismatch, followed by mismatch validation by ADP → ATP exchange. This complex recruits the MLH1–PMS2 heterodimer that, in turn, exchanges ADP for ATP. The latent endonuclease activity of PMS2 introduces a nick in the daughter strand, 5' of the misincorporation. The nick serves as entry point for the EXO1 exonuclease that degrades a patch of the daughter strand that includes the misincorporation. The remaining single-stranded DNA gap is covered by the single-strand DNA binding protein RPA and filled by the replicative DNA polymerase.*

part of the current diagnostic procedure of patients with LS, and involves the analysis of the tumor phenotype in suspected individuals by immunohistochemistry and/or MSI analysis, followed by testing for a mutation in the MMR genes (Fig. 2, Step 1). We proposed that, in case a VUS is found, the diagnostic procedure continues with a number of assays that largely are based on the profound knowledge of the biology of MMR. Step 2 (Fig. 2) may comprise any or all of a number of assays, such as in silico assessment of missense substitution severity with respect to cross-species sequence conservation at the relevant position of the protein, in silico prediction of the effect of the underlying nucleotide change on mRNA spicing, and analysis of the activity of the VUS protein in a comprehensive in vitro MMR reaction (see below). The combined Steps 1 and 2 will enable the classification of many VUS in a clinical diagnostics laboratory, as was recently demonstrated in a pilot study [Kansikas, 2011]. If results from the analyses in Steps 1 and 2 are inconclusive, the VUS remains unclassified. Such variants require an additional layer of diagnostic assays (Step 3; Fig. 2). Detailed cell-biological, biochemical, and biophysical analysis of MutSα and MutLα and their prokaryotic homodimeric homologs MutS and MutL, combined with the elucidation of their crystal structures, has revealed important information about the biochemistry of the MMR pathway.

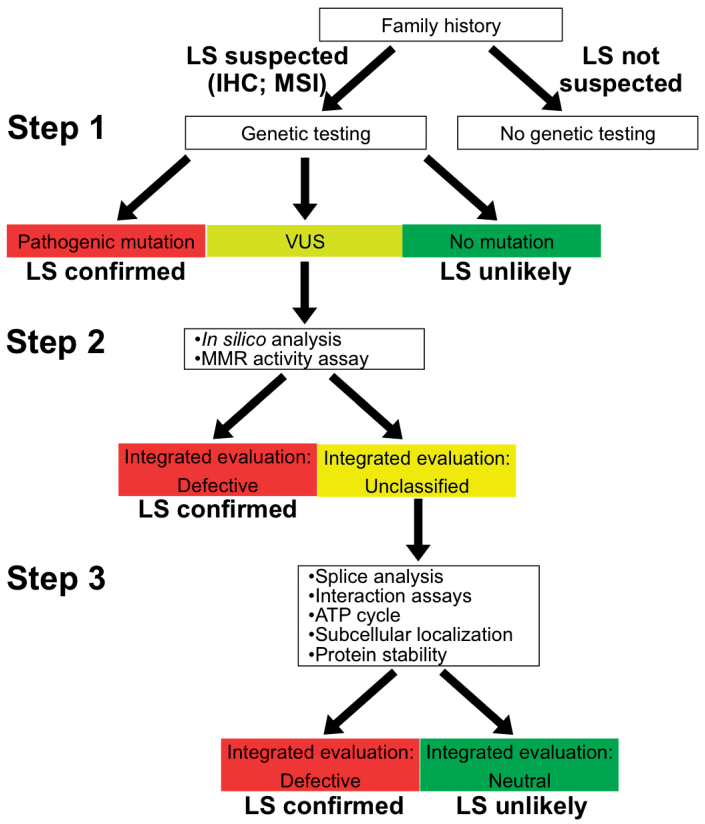


Figure 2. *Proposed decision tree for the diagnostic assessment of mismatch repair protein VUS.*

Based on this body of work, specific biochemical and cell-biological characteristics of the MMR protein VUS can be investigated in Step 3. Furthermore, depending on the predictions of the in silico analyses of Step 2, splicing of the variant mRNA can be assessed in Step 2.

In addition, most, if not all, assays described here investigate aspects of protein function, but VUS that do not cause any functional defect in the protein may still cause aberrations in mRNA splicing or stability or in lowered translation of the mutant protein or aberrant localization in vivo. It is important to combine results from different functional assays (looking at different functional aspects or at function as a whole), in silico analyses of the variants (e.g., conservation of the amino acid, severity of amino acid change) and indirect evidence such as segregation and pathological features [Hofstra et al, 2008]. Individual assays have been developed that may contribute to Steps 2 and 3 of this diagnostic decision tree. However, most of these assays investigate only restricted aspects of MMR proteins, are not widely applicable, have not been validated and are currently not in use in the clinic. Nevertheless, the recent development of second-generation assays hold promise for the development of a widely applicable diagnostic procedure. The Step 1 assays include

diagnostic criteria to start detailed tumor analysis of potential patients with LS. In the following section, we will focus on available Step 2 and Step 3 assays.

Step 2 Functional Assays In Silico Assays

The classification of VUS can be aided by computational analyses [Tavtigian et al., 2008]. Two types of potential problems need to be examined essentially independently: potential dysfunction of the protein caused by amino acid substitutions or small in-frame insertion or deletion mutations, and splicing defects caused by underlying changes to the nucleotide sequence. Algorithms that attempt to predict dysfunction of the protein often use evolutionary conservation as a measure for residue importance, severity of the amino acid substitution relative to the evolutionarily tolerated range of variation at its position, and, in some algorithms, structural features of wild-type and variant proteins [Arnold et al., 2009; Chao et al., 2008; Olatubosun et al., 2012; Tavtigian et al., 2008]. For these programs, three key issues are (1) to build really high-quality protein multiple sequence alignment of the four MMR genes, which helps to reduce the noise in measures of sequence conservation and evolutionarily tolerated range of variation, (2) to calibrate the output of the programs so that their output can be interpreted as a probability in favor of pathogenicity, and (3) to explicitly determine whether the outputs of any pairs or sets of programs can be combined to give more accurate predictions than the most accurate individual program. Similarly, splice site prediction programs have been developed that investigate the possibility of faulty splicing [Reviewed in Spurdle, 2010]. These in silico analyses make use of the relatively well-understood sequence requirements of the consensus splice donor and splice acceptor sites, both to predict whether sequence variants will damage function of the wild-type splice sites and to predict whether they might create de novo splice sites. Sequence variation in exonic splice enhancers also plays an important role [Tournier et al., 2008], but systematic identification functional splice enhancers and the ability to predict the effect of sequence variation on their functional lags behind analysis of donor and acceptor sites. However, it is important to keep in mind that neither in silico prediction of protein dysfunction nor prediction of splice defects have by themselves the predictive values required for counseling individuals carrying the variants.

Splicing Assays

It has been realized only recently that a significant fraction of VUSs in the MMR genes disrupt splicing, rather than protein activity per se [Arnold et al., 2009; Sharp et al., 2004; Tournier et al., 2008]. For this reason, VUS that, on the basis of their position within the gene (close to splice sites) and on in silico predictions, are suspected to affect splicing should be tested for splice defects. In such a splicing assay, the VUS is PCR-amplified from patient genomic DNA with flanking sequences, and subsequently cloned into a splicing reporter minigene. Splicing effects are then evaluated after transfection of the minigene in HeLa cells, followed by RNA isolation and RT-PCR analysis. It is also possible to analyze the RNA from patient lymphocytes directly for splicing alterations, culturing the cells for a short time in the presence of nonsense-mediated RNA decay inhibitors.

Functional Assays Using Yeast

Yeast-based assays for the functional analysis of MMR gene VUS are based on the fact that MMR is evolutionarily conserved [Ou et al., 2007]. In one assay, evolutionarily conserved amino acids in the yeast protein are replaced with its human mutated counterpart, and it is investigated whether this replacement has functional consequences in the yeast cells in vivo. As the human and yeast MLH1 proteins share only significant homology in the NH2- and COOH-termini, mutations located only in these conserved regions can be analyzed in this way.

Yeast functional assays were mainly used for analyzing mutations in MLH1. At least three different strategies have been used: (1) expression of the mutant yeast MMR gene in a haploid yeast strain deficient for the same MMR gene [Shcherbakova and Kunkel, 1999; Wanat et al., 2007]; (2) expression of a human MMR gene in a haploid wild-type yeast strain: this approach makes use of the dominant-negative effect that wild-type human MMR proteins have in these cells [Clark et al., 2000; Shimodaira et al., 1998; Takahashi et al., 2007]; and (3) expression of a mutant yeast MMR gene in a diploid yeast strain containing only one functional allele of the same MMR gene [Drotschmann et al., 1999a].

Several methods have been applied to evaluate the spontaneous mutation rates in various yeast strains expressing mutant MMR proteins, all of which make use of reporter genes that are either located in the yeast chromosome or expressed extra-chromosomally. These reporter genes serve as selection markers; when mutated, they lose or regain their function and can be used to determine the MMR capacity of the introduced mutant MMR protein [Polaczek et al., 1998; Shimodaira et al., 1998; Takahashi et al., 2007; Tran et al., 1997; Wanat et al., 2007]. The reporter genes vary in the length of repeat sequences located in the gene. Such repeat sequences can form loops of different sizes, which are substrates for MMR and can, therefore, influence the results. In addition, a difference can be observed when the mutation is introduced in the yeast gene at corresponding positions, that is, present as a single copy gene at the corresponding locus in the chromosome or when the human variants are expressed from a high copy number plasmid. Furthermore, the human proteins expressed in yeast might not be modified in the same way, as they would have been when expressed in human cells. Such modifications could be phosphorylation of the MMR proteins. Lastly, the strain background can affect the outcome, making the assay less or more sensitive. In summary, although yeast assays are performed in living cells in vivo there are limits to their use. The results should always be correlated with the function of a wild-type protein and, when possible, a known pathogenic variant. In this way, quantitative differences can be determined [Hardt et al., 2011]. However, yeast assays cannot easily be set up in diagnostic laboratories and special expertise in yeast genetics is required.

Cell-Based In Vitro MMR Functional Assays Using a Human Expression System

The largest advantage of using human cell lines and mammalian proteins over yeast-based assays is the ability to test all VUS, irrespective of their evolutionary conservation, inside the cell. Such studies have tested the ability of transfected variant MLH1 or MSH2 cDNAs to complement phenotypes of MLH1-defective A2780 ovarian carcinoma cells [Blasi

et al., 2006] or the MSH2-deficient endometrial tumor cell line Hec59 [Mastrocola and Heinen, 2010], respectively. In the complemented cells, the MMR status was investigated by measuring the cellular response to methylating agents (MMR-deficient cells have acquired tolerance to these agents [Li, 2008]), the spontaneous mutation rate at the endogenous HPRT gene, repair of an exogenously added mismatch-containing GFP plasmid, and by measuring microsatellite instability. A challenge when testing VUS function in mammalian cells is the choice of cell line. Most cell-based studies use a human cancer cell line that has lost expression of the endogenous MMR protein. These models provide an MMR-deficient background for testing the ability of a variant MMR gene to restore repair function compared to wild type. However, these transformed lines have likely undergone multiple rounds of mutation and selection, creating a variable and uncertain genetic background that may affect the performance of the altered MMR protein. Another disadvantage is that the expression of the variant MMR gene is ectopic, and therefore unregulated. The expressed gene may be toxic to the cells and, therefore, these approaches may be susceptible to artefacts. In a recent study, the MMR status of four MSH2 VUS was investigated after their introduction in the endogenous Msh2 gene in mouse embryonic stem (ES) cells, using oligonucleotide-directed gene targeting. This avoids the possible artefacts of unregulated expression and involves a non-transformed cell line, although in these cells both Msh2 alleles contained the VUS whereas in LS only one of the alleles carries a VUS [Wielders et al., 2011].

Cell-Free MMR Assays

The activity of variant MMR proteins can be measured in in vitro assays that study the canonical MMR process as a whole. In one approach, MLH1 variants are transfected together with wild-type PMS2 in the 293T cell line [Trojan et al., 2002] or the HCT116 line [Takahashi et al., 2007] (both human cell lines lackMLH1 and, consequently, PMS2). MMR efficiency is measured by incubating bacteriophage M13 DNA, containing a mismatched basepair in the LacZ α-complementation domain, in cytoplasmic cell extracts from these lines. After incubation, these heteroduplexes are introduced into Escherichia coli cells, and the repair efficiencies are measured by the number and color of the resulting plaques [Thomas et al., 1991].

In another approach, extracts from cell lines lacking a defined MMR protein are complemented with the missing MMR protein [Holmes et al., 1990; Li and Modrich, 1995]. The complemented extracts are added to an artificial substrate that carries a mismatch within a restriction site, interfering with its cleavage. Proficiency in repair results in recreation of the restriction site, which can be determined by simple digestion and gel electrophoresis. Complementation of the repair defect after the addition of a variant MMR protein to this reaction allows the assessment of its activity. Lack of complementation suggests pathogenicity of the variant. This principle has been used in several publications that tested variants in MLH1 [Kosinski et al., 2010; Nyström-Lahti et al., 2002; Raevaara et al., 2005], MSH2 [Ollila et al., 2006], and MSH6 [Kariola et al., 2002]. The variant proteins were expressed in Sf9 insect cells, after transfection with recombinant baculoviruses

encoding the variant MMR protein. In this assay, the amount of Sf9-produced protein frequently exceeds physiological amounts of the protein, and the presence of an excess of the variant in the MMR reaction may mask a partial loss of activity, yielding a false-negative (i.e., non-pathogenic) result [Kansikas et al., 2011].

The assays described above have more practical and technical issues, including the need of molecular cloning and the production of recombinant proteins in cell culture. Furthermore, their labor intensiveness and costs preclude their wide applicability and independent validation. Some of the latter problems were recently tackled by the development of a rapid next-generation assay for testing the integral activity of variants in MLH1 [Drost et al., 2010], MSH2 and MSH6 [Drost et. al, 2012]. In these modified assays, variant MMR genes are produced using a site-directed mutagenic PCR procedure, and variant proteins are produced by expression of these PCR fragments in a commercially available in vitro expression kit. Together with their respective (wild-type) heterodimeric partner, these variants can directly be used in an in vitro complementation assay, similar to that described above, albeit the substrate is fluorescent here, which facilitates the quantification of the repair efficiency. Because this procedure results in non-saturating amounts of variant proteins, variants that also result in a moderate reduction of activity may be diagnosed.

To accurately determine pathogenicity in these in vitro complementation assays, one needs to determine the relationship between the MMR efficiency in vitro and cancer penetrance in carriers. Importantly, it should be realized that most of these assays are not capable of detecting splicing defects, poor expression or degradation in vivo or a lack of nuclear import, which may result in false-negative/positive results. The validation of in vitro complementation assays, using a large panel of well-studied variants, could especially aid in the classification of variants displaying low to intermediate levels of repair. Especially when the functional assays are inconsistent or display wild-type activity of a VUS that the in silico assessments predicted to be pathogenic, additional analytical tools, including Step 3 assays may be required.

Step 3 Functional Assays Interaction assays: yeast two-hybrid assay

The Yeast 2 hybrid (Y2H) assay can be used to measure the physical interaction between two proteins in vivo in the yeast cell. In this assay, one of the proteins of interest (e.g., MSH2) is fused to a DNA binding domain of a transcription factor (e.g., GAL4), whereas its interacting partner (e.g., MSH6) is fused to the transcription activation domain of the same transcription factor. If the two proteins of interest interact, the DNA-binding activation domains are brought into each other's proximity upon which they activate the transcription of a reporter gene or genes (e.g., His3 or LacZ) [Kondo et al., 2003].

For a proper interpretation of the results in Y2H analysis it is important to analyze the expression level of the hybrid MLH1-lexA (or MSH2-Gal4) protein in the yeast strain. It is known that changing a single amino acid can have effects on protein folding, resulting in targeted degradation. Therefore, Western blot analysis of protein extracts from transformed yeast cells, using an internal reference control protein to determine the level of the variant protein, is essential. For example, in the study by Hardt et. al. (2011) seven variants showed

low expression levels in yeast and four of these displayed essentially no interaction. These should be interpreted with caution, as a low protein level would also result in a low level of interaction. Therefore, for a safe interpretation of Y2H results it is advisable to use results only for those variants that are expressed at a level between 50% and 100% in comparison to the wild-type protein [Gammie et al., 2007].

Interaction Assays: GST Pull-Down Assay

The glutathione S-transferase (GST)-fusion interaction (pulldown) assay is a different approach to measure the physical interaction between two proteins. It is an in vitro technique that consists of a fusion-tagged "bait" protein for which a binding partner (i.e., the "prey") is being sought. In most cases, a GST-tagged bait protein is bound to an immobilized glutathione support, after which a labeled potential interacting partner, the prey (in vitro transcribed and translated) is added. If interaction occurs, the labeled protein can be visualized. [Guerrette et al. 1998, 1999] used this assay to show that 9 of 11 MLH1 VUSs displayed reduced binding to PMS2. The disadvantage of this assay is that the IVTT proteins are all expressed at the same level and protein stability due to degradation in cells cannot be evaluated. In addition, the fusion protein (GST-PMS2) is produced in bacteria and, therefore, does not harbour all protein modifications that might be important for a normal interaction. Also, the tag (GST) may influence dimerization of the proteins and generate false positive/negative results.

Both the Y2H assays and the GST pull-down assays are based on the physical interaction between two MMR proteins, such as MSH2 and MSH6, or MLH1 and PMS2. Both systems can only analyze variants that interfere with proper binding and will not, therefore, detect all the possible functional defects of the mutant protein being investigated. A major difference is that Y2H assays are done in live (albeit yeast) cells and the GST pull down is entirely an in vitro assay. In addition, the fusion protein needed for the Y2H assay might affect the interaction.

Interaction Assays: Transient Expression of MMR Genes in MMR-Deficient Cell Lines

Some of the MMR proteins require binding with their protein partner for stabilization. Brieger et al. (2002) showed that the MSH6 protein became detectable in LoVo cells (cells carrying homozygous mutations in MSH2 resulting in destabilization of MSH6), after transfection with wild-type MSH2. In contrast, the MSH6 protein was not detectable in LoVo cells when some mutant MSH2 proteins were expressed (such as MSH2 Cys697Arg). These data indicate a defect in MSH2–MSH6 interaction due to the MSH2 VUS. The same assay works for MLH1 variants when transfected in MLH1 defective HCT116 cells. Only wild-type MLH1, or stable proteins capable of normal protein–protein interaction, increase the stability of PMS2. The advantage of this assay is that it is relatively easy to perform. However, PMS2 stabilization is observed only after 72 hr, which corresponds to the time when the level of the transfected protein starts to decrease. Therefore, a disadvantage is that variants that result in a low level of protein cannot be tested.

DNA Mismatch-Binding Assays

The ability of VUS to affect MSH2–MSH6 mismatch binding has been assessed through the use of electrophoretic mobility shift assay (EMSA) [Clark et al., 1999; Drotschmann et al., 1999b; Lützen et al., 2008; Ollila et al., 2008]. More recently, the effect of missense mutations on MSH2–MSH6/MSH2–MSH3 heterodimer mismatch interaction has been examined using surface plasmon resonance (SPR) [Cyr and Heinen, 2008; Cyr et al., 2012; Heinen et al., 2002; Hess et al., 2002, 2006]. SPR measures the interaction between small molecules in real time. The majority of VUS-containing MSH2– MSH6 heterodimers examined bind G/T mismatches more readily than homoduplex DNA; however, some display a decreased affinity for mismatches compared to wild-type proteins that may affect overall MMR function [Clark et al., 1999; Cyr and Heinen, 2008; Cyr et al., 2012; Drotschmann et al., 1999b; Heinen et al., 2002; Hess et al., 2002, 2006; Lützen et al., 2008; Ollila et al., 2008]. Although they enable the precise kinetic analysis of MMR protein binding to mismatch-containing DNA oligonucleotides, SPR and EMSA are unlikely to become widely applicable assays for MMR protein variants due to their complexity and technical requirements.

Assays Based on the MutS ADP-ATP Cycle

In addition to mismatch recognition, MSH2–MSH6 performs mismatch verification, which prevents it from initiating repair at non-mismatched DNA. This kinetic "proofreading" activity is mediated by the DNA binding induced exchange of ADP for ATP in a Walker-type ATPase domain that is distally located from the mismatch-binding domain. In case of mismatch binding, the ATP-bound form of MSH2–MSH6 is stabilized, and it undergoes a conformational change into a DNA-embracing clamp that allows binding of MLH1–PMS2. The ability of the MSH2–MSH6 heterodimer to coordinate mismatch binding with adenosine nucleotide processing is essential for MMR [Fishel, 1998; Kunkel and Erie, 2005; Modrich, 2006]. Thus, in addition to affecting mismatch binding, the VUS may affect ATP binding and/or hydrolysis. In vitro assays that examine the steps of this molecular mechanism have been used to test MSH2 and MSH6 VUS [Cyr et al., 2012; Cyr and Heinen, 2008; Heinen et al., 2002; Lützen et al., 2008; Ollila et al., 2008]. A VUS that disrupts ATP-stimulated mismatch release or mismatch-stimulated ATPase activity would likely not be fully functional in vivo, strongly suggesting an effect of the VUS on pathogenicity. The major strength of these assays is that they provide a more specific description of how a VUS disrupts MMR function. Although this may not be immediately relevant to the clinician who seeks to simply determine whether the patient should be managed as an LS patient or not, it will be very useful to basic and translational researchers who wish to better understand the mechanism of MMR and how it prevents mutations.

Protein Localization Experiments

Clearly, MMR proteins need to be transported to the nucleus to reach their target, the DNA. The regulation of the subcellular localization of MMR proteins, although important for their function, is still poorly understood [Knudsen et al., 2009]. It has been shown that the regulation

of MMR protein localization depends on (1) specific sequences in the MMR proteins that interact with factors known to be involved in the nuclear import of proteins, (2) sequences involved in translocation of the protein to the nucleus, and (3) specific protein–protein interactions in the MMR complexes [Brieger et al., 2005; Knudsen et al., 2007]. Subcellular localization of mutant MMR proteins can be investigated by immuno-staining or by expressing fluorescent fusion proteins in mammalian cells and investigation of their localization by fluorescent microscopy. A major strength of these assays is that they provide a more specific description of how a VUS disrupts MMR function. However, the assays are technically difficult and the results must be analyzed carefully in relation to proper controls since both expression levels and nature of fluorescent fusion proteins play a major role for the results [Brieger et al., 2012].

MMR proteins may be co-imported by each other. Co-import of excision repair proteins refers to the situation where a protein is imported via the nuclear localization signal (NLS) of its binding partner. Such complex formation between proteins in the cytoplasm seems to have a large effect on their subcellular localization; however, it is not clear whether this protein "piggy-backing" works as a backup mechanism or if it is essential for import regulation in vivo. Nonetheless, it ensures that the stoichiometry of the nuclear repair complexes is sustained [Knudsen et al., 2009]. MMR proteins utilize co-import as a mechanism for increasing nuclear accumulation. Some studies have suggested that MSH2– MSH6 complex formation affects nuclear translocation of MSH2, as MSH6-deficient cells display an increased cytoplasmic level of MSH2 compared to MSH6-proficient cells [Christmann and Kaina, 2000]. The importance of MSH2–MSH6 localization is illustrated in patients with LS carrying specific amino acids substitutions in MSH2. Two MSH2–LS mutants (P622L and C697F) display decreased nuclear accumulation compared to wild-type protein [Lützen et al., 2008].

MLH1–PMS2 complex formation affects the nuclear translocation of both MLH1 and PMS2 in human and mice [Mohd et al., 2006; Wu et al., 2003]. Each subunit contains a functional NLS sequence but mutations in the NLS sequence of either MLH1 or PMS2 lead to impaired nuclear localization of both subunits [Brieger et al., 2005; Wu et al., 2003] suggesting that heterodimers form in the cytoplasm prior to import into the nucleus and may be required for efficient nuclear localization. Indeed, recent reports show that human MLH1 and PMS2 as well as yeast Msh2 are imported into the nucleus in the absence of their dimerization partners, but that transport of human MLH1–PMS2 and yeast Mlh1–Pms1 are synergistically enhanced by heterodimerization [Hayes et al., 2009; Leong et al., 2009]. The importance of MMR complex formation is also illustrated in patients with LS, carrying specific amino acids substitutions in MLH1. One missense mutation MLH1 (R659P), not located within the NLS sequence, was found to abolish the interaction with PMS2 in vitro, and also display decreased nuclear localization [Raevaara et al., 2005].

General Discussion MMR and the DNA Damage Response

Published studies addressing the functional effects of MMR VUS have focused almost exclusively on the role of the MMR proteins in DNA repair. However, MMR proteins are involved in other processes that may also have implications for tumorigenesis [Heinen et al.,

2002]. Studies of MMR-deficient human cancer cell lines as well as embryonic fibroblasts from Msh2, Mlh1, and Msh6 knockout mice suggest that the MMR proteins are necessary for the activation of cell cycle checkpoints and apoptosis in response to certain DNA damaging agents [Aebi et al., 1996; Branch et al., 1995; Cejka et al., 2003; de Wind et al., 1998; Kat et al., 1993; Stojic et al., 2004; Toft et al., 1999; Wu et al., 1999; Zhang et al., 1999]. Whether the DNA repair functions of the MMR proteins are required for this damage response is still debated [Jiricny, 2006; Peña-Diaz and Jiricny, 2012; Stojic and Jiricny, 2004]. Thus, it cannot be excluded that certain VUSs affect either the repair or the damage response function while leaving the other function intact. However, expression of four MSH2 VUSs in an MSH2-null cancer cell line revealed that loss of repair and of DNA damage responses coincided in the tested variants [Mastrocola and Heinen, 2010]. In a separate study, four MSH2 variants were introduced into the endogenous locus in mouse embryonic stem cells and tested for different MMR functions including damage responses and a mutator phenotype, a measure for repair [Wielders et al., 2011]. Similar to the study in human cancer cells, the VUS that maintained the damage response function also did not display a mutator phenotype. These studies support the hypothesis that repair and the induction of cell cycle responses are linked properties of the MMR pathway. In contrast, mouse studies revealed missense mutations in Msh2, Msh6, Mlh1, and Pms2 that disrupted DNA repair function, but not the damage-induced apoptotic response suggesting that these functions may be separable [Avdievich et al., 2008; Lin et al., 2004; Marinovic-Terzic et al., 2008; van Oers et al., 2010; Yang et al., 2004]. Determining the effect of VUSs on the MMR-dependent DNA damage response also may be important for predicting the response to therapy. Multiple studies have suggested that patients with MMR-deficient tumors are not as responsive to some commonly used chemotherapies that function by damaging DNA [Carethers et al., 2004; Ribic et al, 2003; Watanabe et al., 2000]. Biochemical studies suggest that the MMR proteins may be involved in the mechanism by which these drugs kill cells [Aebi et al., 1996; Meyers et al., 2005].

Technical Limitations of Functional Assays

The type of functional assays used has changed over the years. Before 1999, almost all functional assays were yeast-based or used pulldown assays, and most of them were used to study only MLH1 and MSH2 variants. After 1999, more and more publications describe in vitro MMR assays or (human) expression systems for analyzing repair capacity as well as the subcellular localization of the mutant proteins. Unfortunately, many of the assays or methods described above require specialized expertise, laboratory facilities, and trained scientific personnel, and this currently precludes their widespread adoption and independent validation. Possibly, the biggest problem of functional assays is the fact that currently for hardly any of the assays the reproducibility, sensitivity, and specificity have been determined. This is reflected by the many inconsistencies in the database that lists all MMR gene VUS and results of functional assays on these VUS (www.mmruv.info). As long as a test is not validated, results should be utilized with great caution, and one might even argue that functional assays cannot be used to classify variants with any confidence in the absence of other (segregation

and clinical/ genetic) data. When an assay result predicts a pathogenic nature of a variant one should realize that it is of the utmost importance that the proper positive and negative controls were included in the study. For analysis of transfected proteins, the expression level is very important as high levels of these proteins might be toxic or could disturb normal functioning of the overexpressed protein and, conversely, overexpression may mask a partial loss of activity. On the other hand, under expression may result in false negative results as the too little protein may also affect MMR function [Cejka et al., 2003].

A further challenge when interpreting the results from the biochemical assays is determining how moderate alterations in activity translate into disease phenotype. At the moment, there is no obvious way to address this important issue and, therefore, it should be given high priority in future to develop methods to correlate intermediate activity of MMR proteins to LS. In addition, examining the same VUS in multiple assays will allow researchers to correlate in vitro biochemical function with in vivo function. Examplifying these problems, four MSH2 VUS were recently expressed in an MSH2-null human cancer cell line to study their ability to restore cellular MMR function [Mastrocola and Heinen, 2010]. Two of these VUS, D167H and K393M, were previously determined to have intermediate effects on MSH2-MSH6 ATP/ADP processing [Heinen et al., 2002]. Both variants restored cellular MMR function to near wild-type levels suggesting that the biochemical defects associated with these VUS are not sufficient to contribute to disease. However, as noted above, the partial restoration of function may have been a consequence of inexact expression levels of the transfected variant genes. The MSH2 P622L VUS, alternatively, had more dramatic effects on in vitro biochemical function [Guerrette et al., 1999; Heinen et al., 2002; Lutzen et al., 2008], which is consistent with the clear MMR defects observed when this variant is expressed in yeast [Drotschmann et al., 1999a; Gammie et al., 2007; Polaczek et al., 1998] or mammalian cell culture systems [Lutzen et al., 2008; Mastrocola and Heinen, 2010]. In addition, most, if not all, assays described here investigate aspects of protein function, but VUS that do not cause any functional defect in the protein may still cause aberrations in mRNA splicing or stability or in lowered translation of the mutant protein or aberrant localization in vivo. It is important to combine results from different functional assays (looking at different functional aspects or at function as a whole), in silico analyses of the variants (e.g., conservation of the amino acid, relative severity of amino acid change) and indirect evidence such as segregation and pathological features [Hofstra et al., 2008].

FUTURE PERSPECTIVES

To achieve a widely applicable, integrated, validated, and cost-effective diagnostic procedure for the diagnosis of MMR gene VUS, as proposed in Figure 2, further work is needed to select and validate a set comprising the most predictive and feasible assays. An integral part of the diagnostic procedure, furthermore, is the development and use of a computer algorithm that, based on prior calibration, stratifies and integrates the results from all assays, and calculates the probability in favor of pathogenicity.

Whether the outcome of such a validated diagnostic procedure should be used for clinical decisions is a question that needs to be addressed by clinicians, clinical geneticists, and patient advocates and probably will depend on whether individuals are comfortable using the results for clinical decisions knowing that they may include errors. If that is the case, one needs to make sure that the target audience understands all the limitations in the assays, is aware of all possible pitfalls, and understands the significance of each class used in the classification. In addition, genetic counselors and clinicians should receive specific training and teaching aids to help the patient make informed decisions.

In conclusion, only the tight integration of genetic, pathology, in silico, in vitro and, possibly, in vivo data will provide reliable pathological assessment of any gene implicated in cancer predisposition. The thorough knowledge of the pathology and genetics of LS and of the biochemistry of MMR may enable to provide a paradigm for the diagnostic assessment of cancer-associated gene variants in the age of personalized genomics.

REFERENCES

Aebi S, Kurdi-Haidar B, Gordon R, Cenni B, Zheng H, Fink D, Christen RD, Boland CR, Koi M, Fishel R, Howell SB. 1996. Loss of DNA mismatch repair in acquired resistance to cisplatin. Cancer Res 56:3087–3090.

Arnold S, Buchanan DD, Barker M, Jaskowski L, Walsh MD, Birney G, Woods MO, Hopper JL, JenkinsMA, BrownMA, Tavtigian SV, Goldgar DE, Young JP, Spurdle AB. 2009. Classifying MLH1 and MSH2 variants using bioinformatic prediction, splicing assays, segregation, and tumor characteristics. Hum Mutat 30:757–770.

Avdievich E, Reiss C, Scherer SJ, Zhang Y, Maier SM, Jin B, Hou Jr H, Rosenwald A, Riedmiller H, Kucherlapati R, Cohen PE, Edelmann W, Kneitz B. 2008. Distinct effects of the recurrent Mlh1 G67R mutation on MMR functions, cancer, and meiosis. Proc Natl Acad Sci USA 105:4247–4252.

Branch P, Hampson R, Karran P. 1995. DNA mismatch binding defects, DNA damage tolerance, and mutator phenotypes in human colorectal carcinoma cell lines. Cancer Res 55:2304–2309.

Brieger A, Plotz G, Hinrichsen I, Passmann S, Adam R, Zeuzem S. 2012. C-terminal fluorescent labeling impairs functionality of DNA mismatch repair proteins. PLoS One 7:e31863.

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

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

Carethers JM, Smith EJ, Behling CA, Nguyen L, Tajima A, Doctolero RT, Cabrera BL, Goel A, Arnold CA, Miyai K, Boland CR. 2004.Use of 5-fluorouracil and survival in patients with microsatellite-unstable colorectal cancer. Gastroenterology 126:394– 401.

Cejka P, Stojic L, Mojas N, Russell AM, Heinimann K, Cannavo E, di Pietro M, Marra G, Jiricny J. 2003. Methylation-induced G(2)/M arrest requires a full complement of the mismatch repair protein hMLH1. EMBO J 22:2245–2254.

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). Hum Mutat 29:852–860.

Christmann M, Kaina B. 2000. Nuclear translocation of mismatch repair proteins MSH2 and MSH6 as a response of cells to alkylating agents. J Biol Chem 275:36256– 36262.

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 Res 27:736–742.

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

Couch FJ, Rasmussen LJ, Hofstra R, Monteiro AN, Greenblatt MS, De Wind N, Group IARCUGVW. 2008. Assessment of functional effects of unclassified genetic variants. Hum Mutat 29:1314–1326.

Cyr JL, Brown GD, Stroop J, Heinen CD. 2012. The predicted truncation from a cancer associated variant of the MSH2 initiation codon alters activity of the MSH2-MSH6 mismatch repair complex. Mol Carcinog 51:647–658.

Cyr JL, Heinen CD. 2008. Hereditary cancer-associated missense mutations in hMSH6 uncouple ATP hydrolysis from DNA mismatch binding. J Biol Chem 283:31641– 31648.

de La Chapelle A. 2004. Genetic predisposition to colorectal cancer. Nat Rev Cancer 4:769–780.

De Wind N, Dekker M, van Rossum A, van der ValkM, te Riele H. 1998. Mousemodels for hereditary non-polyposis colorectal cancer. Cancer Res 58:248–255.

Drost M, Zonneveld JB, Van Dijk L, Morreau H, Tops CM, Vasen HFA, Wijnen J, De Wind N. 2010. A cell-free assay for the functional analysis of variants of the mismatch repair protein MLH1. Hum Mutat 31:247–253.

Drost M, Zonneveld JB, van Hees S, Rasmussen LJ, Hofstra RM, de Wind N. 2012. A rapid and cell-free assay to test the activity of lynch syndrome-associated MSH2 and MSH6 missense variants. Hum Mutat 33:488–494.

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

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

Fishel R. 1998. Mismatch repair, molecular switches, and signal transduction. Genes Dev 12:2096–2101.

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.

Guerrette S, Acharya S, Fishel R. 1999. The interaction of the human MutL homologues in hereditary non-polyposis colon cancer. J Biol Chem 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 non-polyposis colorectal cancer. Mol Cell Biol 18:6616–6623.

Hardt K, Heick SB, Betz B, Goecke T, Yazdanparast H, K¨uppers R, Servan K, Steinke V, Rahner N, Morak M, Holinski-Feder E, Engel C, M¨oslein G, Schackert HK, von Knebel Doeberitz M, Pox C, Peter Propping; German HNPCC consortium, Hegemann JH, Royer-Pokora B. 2011. Missense variants in hMLH1 identified in patients from the German HNPCC consortium and functional studies. Fam Cancer 10:273–284.

Hayes AP, Sevi LA, Feldt MC, Rose MD, Gammie AE. 2009. Reciprocal regulation of nuclear import of the yeast MutSα DNA mismatch repair protein Msh2 and Msh6. DNA repair 8:739–751.

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

Hess MT, Gupta RD, Kolodner RD. 2002. Dominant Saccharomyces cerevisiae msh6mutations cause increased mispair binding and decreased dissociation from mispairs by Msh2-Msh6 in the presence of ATP. J Biol Chem 277:25545–25553.

Hess MT, Mendillo ML, Mazur DJ, Kolodner RD. 2006. Biochemical basis for dominant mutations in the Saccharomyces cerevisiae MSH6 gene 10.1073/pnas.0510078103. PNAS 103:558–563.

Hofstra RM, Spurdle AB, Eccles D, Foulkes WD, de Wind N, Hoogerbrugge N, Hogervorst FB; IARC Unclassified Genetic Variants Working Group. 2008. Tumor characteristics as an analytic tool for classifying genetic variants of uncertain clinical significance. Hum Mutat 29:1292–1303.

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

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

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

Kadyrov FA, Genschel J, Fang Y, Penland E, Edelmann W, Modrich P. 2009. A possible mechanism for exonuclease 1-independent eukaryotic mismatch repair. Proc Natl Acad Sci USA 106:8495–8500.

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

Kariola R, Raevaara TE, Lonnqvist KE, Nystrom-Lahti M. 2002. Functional analysis of MSH6

mutations linked to kindreds with putative hereditary non-polyposis colorectal cancer syndrome. Hum Mol Genet 11:1303–1310.

Kat A, Thilly WG, Fang WH, Longley MJ, Li GM, Modrich P. 1993. An alkylation tolerant, mutator human cell line is deficient in strand-specific mismatch repair. Proc Natl Acad Sci USA 90:6424–6428.

Knudsen NØ, Andersen SD, Lützen A, Nielsen FC, Rasmussen LJ. 2009. Nuclear translocation contributes to regulation of DNA excision repair activities. DNA Repair 8:682–689.

Knudsen NØ, Nielsen FC, Vinther L, Bertelsen R, Holten-Andersen S, Liberti SE, Hofstra R, Kooi K, Rasmussen LJ. 2007. Nuclear localization of human DNA mismatch repair protein exonuclease 1 (hEXO1).Nucleic Acids Res 35:2609–2619.

Kondo E, Suzuki H, Horii A, Fukushige S. 2003. A yeast two-hybrid assay provides a simple way to evaluate the vast majority of hmlh1 germ-line mutations. Cancer Res 63:3302–3308.

Kosinski J, Hinrichsen I, Bujnicki JM, Friedhoff P, Plotz G. 2010. Identification of Lynch syndrome mutations in the MLH1-PMS2 interface that disturb dimerization and mismatch repair. Hum Mutat 31:975–982.

Kunkel TA, Erie DA. 2005. DNA mismatch repair. Annu Rev Biochem 74:681–710.

Leong V, Lorenowicz J, Kozij N, Guarne N. 2009. Nuclear import of human MLH1, PMS2 and MutLα: redundancy is the key. Mol Carcinog 48:742–750.

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

Li GM, Modrich P. 1995. Restoration of mismatch repair to nuclear extracts of H6 colorectal tumor cells by a heterodimer of human MutL homologs. Proc Natl Acad Sci USA 92:1950–1954.

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

Lutzen A, de Wind N, Georgijevic D, Nielsen F, Rasmussen L. 2008. Functional analysis of HNPCC-related missense mutations in MSH2. Mutat Res 645:44–55.

Marinovic-Terzic I, Yoshioka-Yamashita A, Shimodaira H, Avdievich E, Hunton IC, Kolodner RD, Edelmann W, Wang JY. 2008. Apoptotic function of human PMS2 compromised by the non-synonymous single-nucleotide polymorphic variant R20Q. Proc Natl Acad Sci USA 105:13993–13998.

Mastrocola A, Heinen C. 2010. Lynch syndrome-associated mutations in MSH2 alter DNA repair and checkpoint response functions in vivo. Hum Mutat 31:E1699– E1708.

Meyers M, Wagner MW, Mazurek A, Schmutte C, Fishel R, Boothman DA. 2005. DNA mismatch repair-dependent response to fluoropyrimidine-generated damage. J Biol Chem 280:5516–5526.

Modrich P. 2006. Mechanisms in eukaryotic mismatch repair. J Biol Chem 281:30305– 30309.

Mohd AB, Palama B, Nelson SE, Tomer G, Nguyen M, Huo X, Buermeyer AB. 2006. Truncation of the C-terminus of human MLH1 blocks intracellular stabilization of PMS2 and disrupts DNA mismatch repair. DNA Repair 5:347–361.

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 non-polyposis colon cancer. Genes Chromosomes Cancer 33:160–167.

Olatubosun A, Väliaho J, Härkönen J, Thusberg J, VihinenM. 2012. PON-P: Integrated predictor for pathogenicity of missense variants. Hum Mutat 33:1166–1174.

Ollila S, Dermadi Bebek D, Jiricny J, Nystrom M. 2008. Mechanisms of pathogenicity in human MSH2 missense mutants. Hum Mutat 29:1355–1363.

Ollila S, Sarantaus L, Kariola R, Chan P, Hampel H, Holinski-Feder E, Macrae F, Kohonen-Corish M, Gerdes AM, Peltomaki P, Mangold E, de la Chapelle A, Greenblatt M, Nystrom M. 2006. Pathogenicity of MSH2 missense mutations is typically associated with impaired repair capability of the mutated protein. Gastroenterology 131:1408–1417.

Ou J, Niessen RC, Lützen 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. Hum Mutat 28:1047–1054.

Peltomäki P. 2001. Deficient DNA mismatch repair: a common etiologic factor for colon cancer. Hum Mol Genet 10:735–740.

Peltomäki P. 2003. Role of DNA mismatch repair defects in the pathogenesis of human cancer. J Clin Oncol 21:1174–1179.

Pena-Diaz J, Jiricny J. 2012. Mammalian mismatch repair: error-free or error-prone? Trends Biochem Sci 37:206–214.

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

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

Ribic CM, Sargent DJ, Moore MJ, Thibodeau SN, French AJ, Goldberg RM, Hamilton SR, Laurent-Puig P, Gryfe R, Shepherd LE, Tu D, Redston M, Gallinger S. 2003. Tumor microsatellite-instability status as a predictor of benefit from fluorouracilbased adjuvant chemotherapy for colon cancer. N Engl J Med 349:247–257.

Sharp A, Pichert G, Lucassen A, Eccles D. 2004. RNA analysis reveals splicing mutations and loss of expression defects in MLH1 and BRCA1. Hum Mutat 24:272. Shcherbakova PV, Kunkel TA. 1999. Mutator phenotypes conferred by MLH1 overexpression and by heterozygosity for mlh1 mutations. Mol Cell Biol 19:3177–3183.

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

Spurdle AB. 2010. Clinical relevance of rare germline sequence variants in cancer genes: evolution and application of classification models. Curr Opin Genet Dev 20:315– 323.

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

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

Tavtigian SV, Greenblatt MS, Lesueur F, Byrnes GB. 2008. IARC unclassified genetic variants working group in silico analysis of missense substitutions using sequence-alignment based methods. Hum Mutat 29:1327– 1336.

Toft NJ, Winton DJ, Kelly J, Howard LA, Dekker M, te Riele H, Arends MJ, Wyllie AH, Margison GP, Clarke AR. 1999. Msh2 status modulates both apoptosis and mutation frequency in the murine small intestine. Proc Natl Acad Sci USA96:3911– 3915.

Tournier I, Vezain M, Martins A, Charbonnier F, Baert-Desurmont S, Olschwang S, Wang Q, Buisine MP, Soret J, Tazi J, Frébourg T, Tosi M. 2008. A large fraction of VUS of the mismatch repair genesMLH1 and MSH2 is associated with splicing defects. Hum Mutat 29:1412–1424.

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

van Oers JM, Roa S, Werling U, Liu Y, Genschel J, Hou Jr H, Sellers RS, Modrich P, Scharff MD, Edelmann W. 2010. PMS2 endonuclease activity has distinct biological functions and is essential for genome maintenance. Proc Natl Acad Sci USA 107:13384–13389.

Wanat JJ, Singh N, Alani E. 2007. The effect of genetic background on the function of Saccharomyces cerevisiae mlh1 alleles that correspond to HNPCC missense mutations. Hum Mol Genet 16:445–452.

Watanabe Y, Haugen-Strano A, Umar A, Yamada K, Hemmi H, Kikuchi Y, Takano S, Shibata Y, Barrett JC, Kunkel TA, Koi M. 2000. Complementation of an hMSH2 defect in human colorectal carcinoma cells by human chromosome 2 transfer. Mol Carcinog 29:37–49.

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. Hum Mutat 32:389–396.

Wu J, Gu L, Wang H, Geacintov NE, Li GM. 1999. Mismatch repair processing of carcinogen-DNA adducts triggers apoptosis. Mol Cell Biol 19:8292–8301.

Wu X, Platt JL, Cascalho M. 2003. Dimerization of MLH1 and PMS2 limits nuclear localization of MutLalpha. Mol Cell Biol 23:3320–3328.

Yang G, Scherer SJ, Shell SS, Yang K, Kim M, Lipkin M, Kucherlapati R, Kolodner RD, Edelmann W. 2004. Dominant effects of an Msh6 missense mutation on DNA repair and cancer susceptibility. Cancer Cell 6:139–150.

Zhang H, Richards B, Wilson T, Lloyd M, Cranston A, Thorburn A, Fishel R, Meuth M. 1999. Apoptosis induced by overexpression of hMSH2 or hMLH1. Cancer Res 59:3021–3027.

3

A Rapid and Cell-free Assay to Test the Activity of Lynch Syndrome-associated *MSH2* and *MSH6* Missense Variants

Mark Drost, José B.M. Zonneveld, Sandrine van Hees, Lene Juel Rasmussen, Robert M.W. Hofstra and Niels de Wind

Adapted from Drost et. al 2012 Hum Mutat 33:488–494

ABSTRACT

Lynch Syndrome (LS) is an autosomal dominant disorder that predisposes to colon, endometrial and other cancers. LS is caused by a heterozygous germline mutation in one of the DNA mismatch repair (MMR) genes. A significant proportion of all mutations found in suspected LS patients comprises single amino acid alterations. The pathogenicity of these Variants of Uncertain Significance (VUS) is difficult to assess, precluding diagnosis of carriers and their relatives. Here we present a rapid cell-free assay to investigate MMR activity of *MSH2* or *MSH6* VUS. We used this assay to analyze a series of *MSH2* and *MSH6* VUS, selected from the Leiden Open Variation Database. Whereas a significant fraction of the *MSH2* VUS has lost MMR activity, suggesting pathogenicity, the large majority of the *MSH6* VUS appears MMR proficient. We anticipate that this assay will be an important tool in the development of a comprehensive and widely applicable diagnostic procedure for LS-associated VUS.

INTRODUCTION

Lynch Syndrome (LS) is an autosomal dominant cancer syndrome that predisposes to a variety of malignancies, including colon, gastric, brain, ovarian and urological cancers (MIM# 120435)[Lagerstedt Robinson et al., 2007]. The syndrome is caused by mutations in one of the DNA Mismatch Repair (MMR) genes, most frequently *MSH2* (MIM# 609309), *MSH6* (MIM# 600678) or *MLH1* (MIM# 120436). In a significant proportion of individuals suspected of LS, missense mutations or small in-frame deletions are detected in MMR genes (18% in *MSH2*, 38% in *MSH6* and 32% in *MLH1* [Peltomäki and Vasen, 2004]). Since the causality of such variants is difficult to assess, they are often referred to as Variants of Uncertain Significance (VUS) [Ou et al., 2007]. The lack of classification of these VUS precludes LS diagnosis for carriers and their relatives. For this reason, there is a need for convenient functional assays that may play a key role in the diagnostic assessment of suspected LS patients carrying an MMR gene VUS.

The MMR pathway repairs spontaneous misincorporations during replication. In eukaryotic cells, the major MMR pathway is initiated by a heterodimer that recognizes mismatched basepairs, called MutSα and consisting of MSH2 and MSH6 [Hsieh and Yamane, 2008]. This leads to the recruitment of the downstream repair components MutLα (composed of MLH1 and PMS2) and EXO1. These enzymes will nick and then degrade the newly synthesized DNA, which is followed by resynthesis, usually incorporating the correct nucleotide. The entire MMR reaction can also be performed *in vitro* in wild type (WT) human cell extracts [Holmes et al., 1990], but also in cell extracts lacking defined MMR proteins, complemented with purified or overexpressed protein [Li and Modrich, 1995]. Consequently, the ability of an MMR VUS protein to complement such a MMR-deficient cell extract may contribute to the assessment of its pathogenicity [Ou et al., 2007; Kansikas et al., 2011]. Unfortunately, the need for laborious and cost-intensive molecular cloning and cell culture has precluded the wide applicability and independent validation of these assays.

Recently, we described a cell-free assay for the functional analysis of VUS in the *MLH1* gene that avoids these drawbacks [Drost et al., 2010]. Here, we report on the development of a similar assay for the functional analysis of VUS in *MSH2* and *MSH6*. Using this assay we have tested a series of VUS in *MSH2* and *MSH6* that have been found in suspected LS patients. In contrast to many VUS in *MSH2*, most VUS in *MSH6* demonstrate no defect in MMR.

METHODS

In Vitro MMR Assays

In vitro MMR assay conditions were as described (Fig. 1A) [Wang and Hays, 2006; Drost et al., 2010]. Reactions were carried out in a total volume of 25µl containing 75µg nuclear extract (NE). NE was prepared from LoVo colon cancer cells, that lack both MSH2 and MSH6, as described [Holmes et al., 1990; Jiricny, 2003]. The NE was complemented with 12µl dimerized variant MutSα (see below) and 100 ng pJHGT3'lnFAM as a mismatch-containing substrate. After 40 minutes of incubation at 37°C, substrate DNA was purified using the MinElute kit (Qiagen,

Germantown, MD). The substrate was then digested with *Hin*DIII and *Bsr*BI (Fermentas). One-fifth (2µl) of digested substrate was mixed with 8µl Hi-Di Formamide containing 0.2µl GeneScan-500 ROX size standards (Applied Biosystems) and fragment analysis was performed on a 3730 DNA analyzer (Applied Biosystems) at the Leiden Genome Technology Center. Data was analyzed using PeakScan Software (Applied Biosystems). Repair levels are calculated by dividing the height of the MMR-specific peak by the total fluorescent signal.

Recombinant purified wild type MutSα, produced in *Sf9* insect cells, was a gift from Prof. Titia Sixma, The Netherlands Cancer Institute.

Generation of MSH2 and MSH6 VUS by PCR and *In Vitro* Expression of MMR VUS Proteins

The human *MSH2* (Genbank NM_000251.1) and *MSH6* (Genbank NM_000179.2) cDNAs were cloned into the pCITE4a expression vector (Novagen, Beeston, UK). These constructs were used as templates in a PCR procedure to recreate VUS, including the vector-derived T7 promoter and CITE-sequences that are required for efficient protein expression *in vitro*. In brief, overlapping 5' and 3' mutant fragments of the gene were generated in separate, parallel, PCR reactions (Fig. 1B). In a subsequent PCR the two mutant fragments were joined and re-amplified. Oligonucleotides (Supp. Table S1 and S2) were ordered at Biolegio (Nijmegen, The Netherlands). All PCRs were performed in 10µl reactions, containing 0.4 U of Pfx Platinum Polymerase (Invitrogen, Carlsbad, CA, USA). Proper introduction of the mutation was confirmed by direct sequencing of all full-length PCR fragments.

After amplification, PCR fragments were purified using the PCR Purification Kit (Qiagen, Germantown, MD, USA) and eluted in 40µl TE. These purified fragments were then used for *in vitro* protein expression in the TnT T7 Quick Coupled *in vitro* Transcription/Translation (IVTT) System (Promega, Madison, WI, USA), in the presence of PCR Enhancer. Expression levels were verified by producing parallel ^{35}S-Methionine-labeled reactions according to the manufacturer's instructions, which were analyzed by SDS-polyacrylamide gel electrophoresis and Phosphorimaging (Perkin Elmer, Waltham, USA). The WT heterodimeric partner of the variant proteins was produced by large-scale *in vitro* expression from the pCITE4A clones. MutSα for *in vitro* activity testing was generated by allowing *in vitro*-expressed wild type or VUS MSH2 and MSH6 (1:1 v/v), to dimerize for 20 minutes at room temperature.

Western Blotting

After addition of Laemmli SDS sample buffer, samples were separated by 8% SDS-polyacrylamide gel electrophoresis and transferred to Hybond-C membranes (GE Healthcare, Little Chalfont, UK). Membranes were incubated with a rabbit α-MLH1 antibody (H300, Santa Cruz Biotechnology, Santa Cruz, CA) or a mouse antibody against MSH2 (Ab-2, Calbiochem). After incubation with peroxidase-conjugated secondary antibodies (Bio-Rad, Hercules, US), proteins were visualized by enhanced chemiluminescence detection. Signals were quantified using Optiquant software (Packard Instrument Company, Meriden, USA).

Generation of Cloned VUS

As controls, a subset of the VUS was also introduced into *MSH2* or *MSH6* expression plasmids (derived from pCITE4A), using site-directed mutagenesis (Quick Change Site-Directed Mutagenesis, Stratagene, LaJolla, USA). Amplifications were carried out using Pfx Platinum Polymerase. The mutant plasmids were propagated in *E. coli*. The sequences of the resulting plasmids were verified by direct sequencing of the variant *MSH2* or *MSH6* open reading frames.

RESULTS

A Cell-Free Assay that Measures the Repair Activity of MSH2 and MSH6 Proteins

We have recently described a rapid and completely cell-free assay for the functional analysis of VUS in *MLH1*, which may become instrumental in the assessment of their pathogenicity [Drost et al., 2010]. Here, we describe a similar approach for the functional analysis of VUS in *MSH2* and *MSH6*. The assay relies on the use of an engineered plasmid containing a G·T mismatch and a linked fluorescent label. The mismatch is embedded in a *Hin*DIII restriction site, making this site refractory to cutting. Cleavage of the *Hin*DIII site is possible only after MMR-mediated repair of the G·T mismatch, resulting in the generation of a 75 base pair fluorescently-labeled restriction fragment that can be identified and quantified on an automated fragment analyzer (Fig. 1A).

To enable the efficient construction of a series of *MSH2* or *MSH6* VUS we developed a two-stage PCR procedure, analogous to the generation of VUS in *MLH1* (Fig. 1B) [Drost et al., 2010]. This PCR procedure was optimized using non-mutant primers, to recreate wild type *MSH2* and *MSH6* genes. The variant proteins encoded by these PCR-amplified genes were expressed in a commercially available reticulocyte lysate-based transcription/translation kit. Analysis of proteins expressed in the presence of ^{35}S-Methionine demonstrated similar expression levels for MSH2 and MSH6 (Fig. 1C). Subsequently, identical volumes of the *in vitro* expressed MSH2 and MSH6 proteins were heterodimerized and added to an MSH2- and MSH6-deficient LoVo NE under conditions of the cell-free assay. To relate the quantity of *in vitro* produced MutSα with the physiological quantity of the dimer a second reaction was set up, containing NE of MMR-proficient HeLa cells. The concentration of MSH2 in both reactions was then determined by Western blotting. The concentration of *in vitro* expressed MSH2 in the reconstituted LoVo NE was slightly lower than the concentration of endogenous MSH2 in the reaction containing WT HeLa NE (Fig. 1D). This indicates that the *in vitro* expression procedure yields MSH2 or MSH6 proteins in near-physiological quantities.

To enable the detection of subtle defects in the activities of the VUS proteins, the amount of MSH2 and MSH6 in the assay should not be saturating [Drost et al., 2010; Kansikas et al., 2011]. Complementation of a NE of LoVo cells by the addition of increasing amounts of purified MutSα, produced in *Sf9* insect cells, resulted in saturation at 58% repair of the substrate (Fig. 1E). Complementation of the LoVo NE with *in vitro* expressed MSH2 and MSH6 resulted in repair of 42% of the substrate (Fig. 1E). This implies that the amount of *in vitro* expressed MSH2/

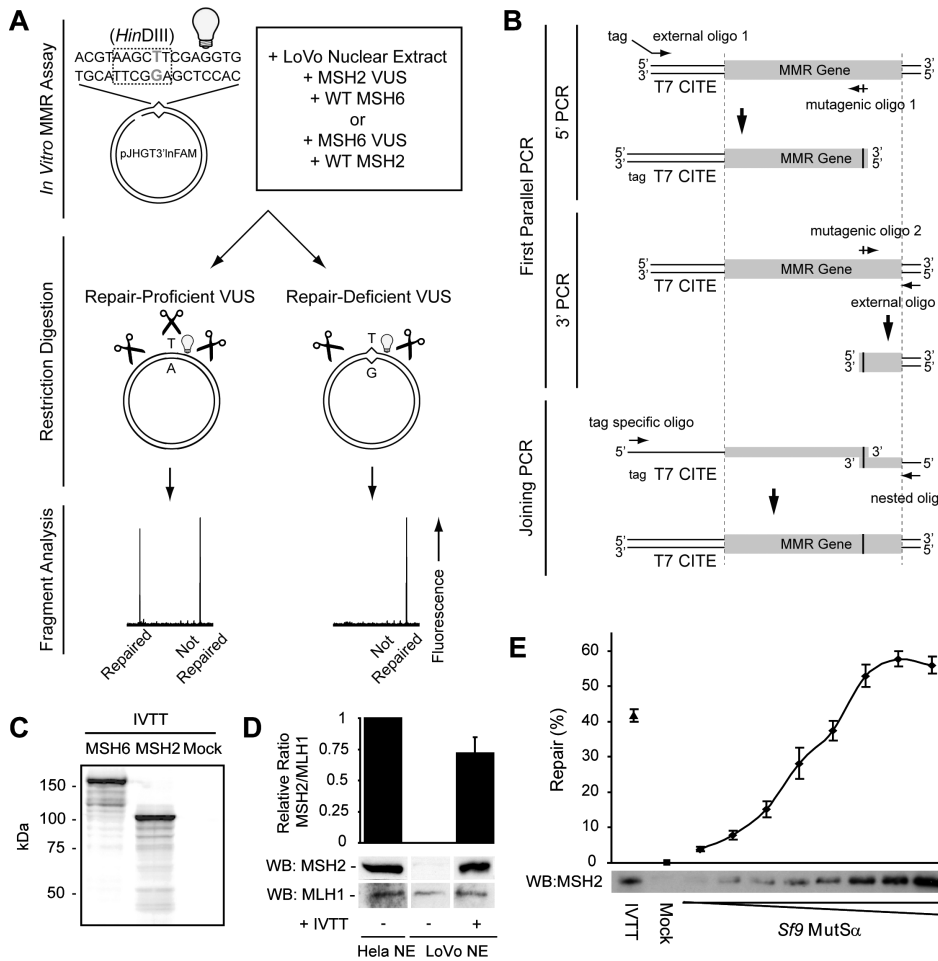


Figure 1. *Outline of the cell-free assay. (A) Flow scheme of the cell-free assay. Fluorescently (FAM)-labeled substrate pJHGT3'lnFAM is incubated in LoVo NE, complemented with an in vitro produced MutSα heterodimer, in which either MSH2 or MSH6 is a VUS. After incubation, the substrate is purified and cleaved with BsrBI and HinDIII. Repair products are analyzed by automated fragment analysis and subsequently quantified. The light bulb represents the fluorescent FAM label. (B) PCR design and primer positions for generating 5' and 3' VUS "half" fragments in parallel reactions and the final, full-length fragment, produced in a second 'joining' PCR. Vertical bar: position of the introduced mutation. Primer sequences can be found in Supplementary Tables 1 and 2. (C) In vitro expression of ^{35}S-Methionine-labeled MSH2 and MSH6, visualized after gel electrophoresis and Phosphorimaging. (D) Western blot analysis of MMR protein levels within the repair reaction. Repair assay reactions were assembled and a fraction was subjected to Western blotting. The relative amount of MSH2 versus endogenous MLH1 is determined. Error bar: S.E.M. (E) In vitro repair activity of in vitro produced wild type MutSα (IVTT), and of increasing amounts of purified Sf9-expressed MutSα, added to LoVo NE. Data are shown as mean±S.E.M. of 3 independent experiments. Bottom: Western blot showing MSH2 protein levels in the repair assay reactions.*

MSH6 is limiting for MMR. Importantly, the addition of *in vitro* expressed MSH2 or MSH6 alone to LoVo NE did not result in detectable repair (see also below). This confirms that neither the *in vitro* expression mixture, nor the NE contained measurable amounts or either wild type protein, which might interfere with testing VUS proteins. Combined, these results indicate that a complementation assay that comprises completely *in vitro* produced protein may be a feasible approach to test the functional activity of MSH2 and MSH6 VUS.

MMR Activities of MSH2 and MSH6 VUS Measured Using a Cell-Free Assay

To investigate the use of the cell-free assay in the functional assessment of MSH2 and MSH6 VUS, identified in suspected LS patients, we determined the repair activity of 25 VUS in *MSH2* and 20 VUS in *MSH6*. An overview of these VUS and their associated pathology data, *in silico* analysis and references can be found in Table 1. For validation purposes, we included MSH2-P622L, MSH2-A636P and MSH6-T1219I as repair-deficient controls and MSH2-G332D, MSH2-A834T, MSH6-S144I, MSH6-P1087R and MSH6-P1087T as repair-proficient controls. The MMR status of these control variants has been assessed in first-generation assays (see Table 1 for references). As an additional repair-deficient control we produced MSH6-G1139S, which alters an essential residue in the ATPase domain of MSH6 [Woods et al., 2005]. The remaining VUS for both genes were selected in an unbiased fashion and were predominantly derived from Dutch and Danish suspected LS patients.

All *MSH2* and *MSH6* VUS proteins were generated by the two-stage PCR procedure, followed by their expression *in vitro*. The *in vitro* expressed VUS proteins were similar in yield despite occasional variation in the quantities of PCR products (Fig. 2A-D). Sequence analysis, performed on all PCR-generated VUS, demonstrated the absence of contamination with wild type gene (Fig. 2E and 2F). All VUS proteins were dimerized with their wild type partner to recreate variant MutSα proteins, and these were tested in the *in vitro* MMR assay. The repair activity of WT MSH2 or WT MSH6 alone, or of all established repair-deficient heterodimerized controls was 16% or less, relative to wild type (Fig 3A and B). In contrast, all repair-proficient controls displayed a repair activity that was significantly higher than the MMR-deficient controls. This result indicates that the assay has sufficient resolution to distinguish repair-proficient from repair-deficient MMR VUS proteins. As a tentative cutoff for MMR proficiency we assumed a repair level that was significantly higher (Student's one-tailed T-test, $p<0.05$) than the average repair efficiency of the repair-deficient controls. The remaining VUS displayed variable repair levels but, notably, MSH2-Y165D, MSH2-N596del, MSH2-G674R, MSH2-K675A, MSH2-C697F, MSH2-S723F and MSH2-G759E were considered repair-deficient by these criteria (Fig. 3A). Surprisingly, all tested MSH6 VUS were repair-proficient (Fig. 3B).

To exclude that a low level of contamination of mutant PCR fragments with wild type *MSH2* or *MSH6* had confounded the results of the assay, we also investigated the repair efficiency of a representative selection of the VUS after cloning and propagation in *E. coli*. Repair levels of these cloned MSH2 or MSH6 VUS were not significantly different from those of their counterparts produced from the mutant PCR fragments (Supp. Fig S1). Importantly, the classification of all cloned VUS was identical to that from the PCR-derived

Table 1. *In silico* data, pathology data and references for all VUS tested in this work.

	Mutation			Pathology [b]				*In silico* analysis		Functional Studies [f]	
					IHC						
	Protein	DNA [a]	*In vitro* MMR	MSI	MSH2	MSH6	References [c]	MAPP-MMR [d]	PolyPhen [e]	Deficient	Proficient
MSH2	T44M	c.131C>T	+					3.92	1.96		
	A45V	c.134C>T	+					4.04	0.186		
	Y165D	c.493T>G	-					8.75	1.49		
	D167H	c.499G>C	+	L			1	9.62	2.272	1[g]	1[g], 12, 13
	E177H	c.529_531delinsCAC [h]	+					3.23	2.257		
	A272V	c.815C>T	+	L/H	+	+	2	3.54	0.712		2, 14
	A305T	c.913G>A	+	L			3	3.55	1.657		1
	G322D	c.965G>A	+	L	+	+	4	3.24	1.395		14, 15,16
	P385L	c.1154C>T	+					5.34	2.846		
	L390F	c.1168C>T	+					5.7	0.949		
	F519L	c.1555T>C	+					3.63	0.605		
	N596del	c.1786_1788delAAT	-					NA	NA		17
	P622L	c.1865C>T	-	H	-	-	1	26.92	3.379	1, 12, 13, 15	17
	A636P	c.1906G>C	-	H	-	+/-	2	8.73	1.154	2, 14, 18	
	H639R	c.1916A>G [h]	+					3.63	3.352	1[g]	1[g]
	G674R	c.2020G>C	-	H	-		1	36.52	2.739	1	
	K675A	c.2023_2025delinsGCC [h]	-					26.45	2.657		
	C697F	c.2090G>T	-	H	-	+	2	11.52	3.676	1, 2, 14	
	S723F	c.2168C>T	-	H	-		5	21.57	2.771		
	G759E	c.2276G>A	-					35.36	2.739		
	L805V	c.2413C>G	+					2.62	0.334		
	A834T	c.2500G>A	+	H	-	-	2	7.91	1.732		1, 2, 14
	C843G	c.2527T>G [h]	+					4.49	1.816		
	S860L	c.2579C>T	+	L			6	3.13	1.665		
	E886G	c.2657A>G	+					4.10	1.11		

MSH6	A20V	c.59C>T	+					Not available for this protein	1.219			
	A25S	c.73G>T	+						0.841			
	S144I	c.431G>T	+	L/H	+	+	7		1.883			7
	A326V	c.977C>T	+						0.158			
	L396V	c.1186C>G	+	L			8		0.782			
	M492V	c.1474A>G	+						1.628			
	S503C	c.1508C>G	+	L	+	+	6		1.802			
	Q522R	c.1565A>G	+	L/H	+	+	9		0.539			
	K610N	c.1830G>T	+						0.007			
	Y850C	c.2549A>G	+	H	+	+/-	9		2.711			
	V878A	c.2633T>C	+	L	+		9		1.54			
	R976H	c.2927G>A	+						2.102			
	A1021D	c.3062C>A	+	L	+	+	9		1.89			
	D1026Y	c.3076G>T	+						2.084			
	P1087R	c.3260C>G	+						2.213			7
	P1087S	c.3259C>T	+	L	+	+	10		1.16			
	P1087T	c.3260C>A	+						2.045			7
	G1139S	c.3415G>A	-	H	-	-	11		2.002			
	T1219I	c.3656C>T	-	H	+	+	9		2.28	18		
	T1225M	c.3674C>T	+						2.129			

[a] Nucleotide numbering reflects cDNA numbering with +1 corresponding to the A of the translation initiation codon in the GenBank reference sequence, amino acid numbering is based on the cDNA with +1 corresponding to the translation initiation codon. *MSH2* (Genbank NM_000251.1) and *MSH6* (NM_000179.2). VUS are registered in the database of the International Society for Gastointentinal Hereditary Tumours (http://chromium.liacs.nl/LOVD2/colon_cancer/home.php) and/or the Newfoundland MMR Genes Variant Database (http://www.med.mun.ca/MMRvariants/)[Woods et al., 2007]. [b] MSI=Microsatellite Instability. H=High, L=Low, IHC=Immunohistochemistry. [c] References: 1 [Lützen et al., 2008], 2 [Ollila et al., 2006], 3 [Wijnen et al., 1997], 4 [Barnetson et al., 2008], 5 [Lucci-Cordisco et al., 2006], 6 [De Jong, 2004], 7 [Kariola et al., 2002], 8 [Kolodner et al., 1999], 9 [Berends et al., 2002], 10 [Niessen et al., 2006], 11 [Woods et al., 2005], 12 [Belvederesi et al., 2008], 13 [Mastrocola and Heinen, 2010], 14 [Ollila et al., 2008], 15 [Wielders et al., 2011], 16 [Ollila et al., 2008], 17 [Guerrette et al., 1998], 18 [Yang et al., 2004]. [d] MAPP-MMR [Chao et al., 2008] scores were calculated from http://mappmmr.blueankh.com/. Values >4.55 are considered pathogenic. NA=Not Available. [e] PolyPhen [Ramensky et al., 2002] scores were calculated from http://genetics.bwh.harvard.edu/pph/. Values >2 are considered pathogenic. NA=Not Available. [f] For references that contain more experiments than an *in vitro* MMR assay alone, only the result for the *in vitro* MMR assay has been included in this table. [g] From Reference 1, only the mismatch binding assay is included in this table. These variants have intermediate phenotypes in the referred manuscripts. VUS have been previously described at the protein level only. Variants have been redesigned at cDNA level.

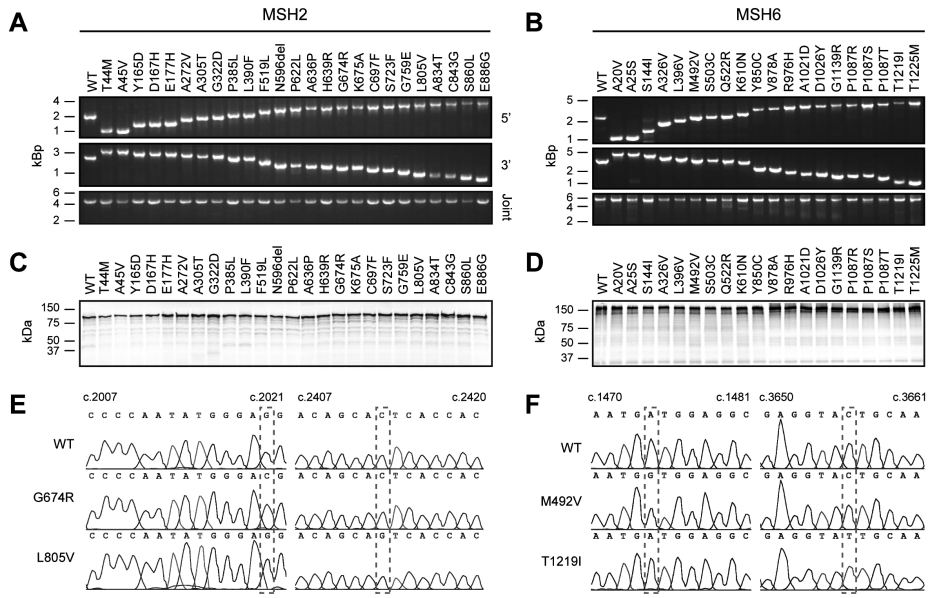


Figure 2. *Overview of the production of MSH2 and MSH6 VUS by PCR and their subsequent in vitro expression. (A) Representative agarose gel showing 5' and 3' 'half', and joint, PCR products of all tested MSH2 VUS. (B) As A, but for MSH6 VUS. (C) Representative expression of ^{35}S-Methionine-labeled MSH2 VUS, visualized after SDS-PAGE gel electrophoresis and autoradiography. (D) As C, but for MSH6 VUS. (E) Example of sequence chromatograms of two randomly chosen PCR-generated full-length MSH2 VUS fragments and a WT control. G674R carries a G→C transversion at position 2020 of the open reading frame; L805V carries a C→G transversion at position 2413, as indicated by the dotted box. No contamination with wild type gene is detected. (F) As E, but for MSH6 VUS.*

VUS (Table 1 and data not shown). This result excludes that the detected MMR proficiency results from carry-over of the PCR-based VUS proteins with wild-type contaminant.

DISCUSSION

The assessment of pathogenicity to an MMR gene VUS provides certainty to carriers. Conversely, relatives who do not carry the pathogenic mutation are liberated from the burden of surveillance, a procedure that effectively reduces morbidity and mortality in individuals diagnosed with LS [Jarvinen et al., 2009]. With the advent of personalized genomics the number of identified MMR gene VUS will likely increase, further emphasizing the need for a validated and standardized diagnostic procedure [Heinen, 2009]. In this work, we present a reproducible, rapid and cell-free assay for the functional analysis of VUS in *MSH2* and *MSH6*. The substrate, NE and reagents can be prepared and stored in large quantities. Patient-derived material is not required, PCR enzymes and the *in vitro* transcription/translation kit are commercially available, and the assay only uses common laboratory equipment. The amount of *in vitro* recreated variant MutSα protein is limiting and therefore the assay may be suited to detect subtle losses-of-function of the VUS.

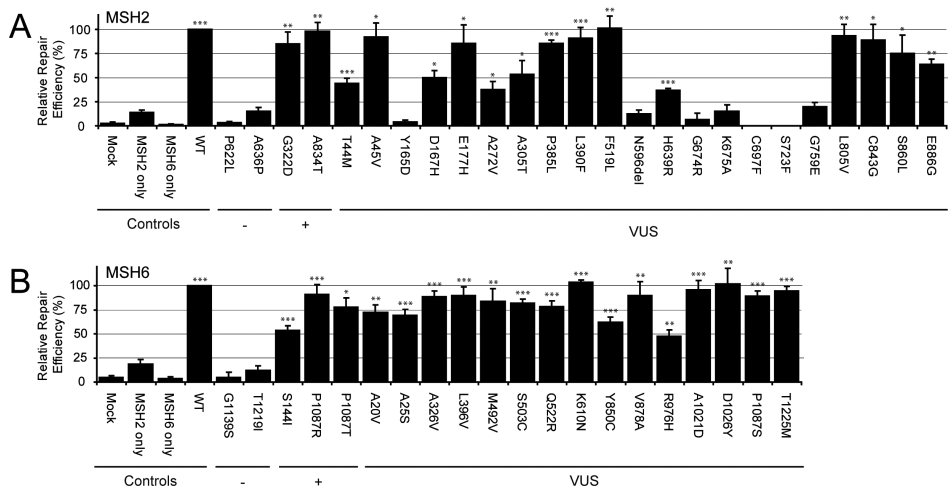


Figure 3. *MMR activity of MSH2 and MSH6 VUS as measured in the cell-free assay. (A) Relative repair efficiencies for MSH2 VUS. (-): Repair-deficient controls, (+): Repair-proficient controls. Results are shown as mean±S.E.M. of 3-4 fully independent experiments for all VUS and >6 experiments for controls. Mock: Mock expression. Asterisks: Significantly higher than repair-deficient controls. *=p<0.05, **=p<0.01, ***=p<0.001 (Student's one-tailed t-test). For VUS MSH2-C697F and MSH2-S723F no repair was detected in all experiments. (B) As A, but for MSH6 VUS.*

Based on these characteristics we infer that the assay may become a key ingredient in a comprehensive diagnostic procedure for LS-associated VUS.

All MSH2 VUS tested here were also analyzed *in silico* with PolyPhen, as well as with MAPP-MMR, a tailored alignment-based algorithm [Ramensky et al., 2002; Chao et al., 2008]. As MAPP-MMR is not available for MSH6, the MSH6 VUS were tested only with PolyPhen. Results from these algorithms correspond with our functional data in 83% and 73% of all cases, respectively (Table 1). Our functional results are consistent with previously published activity assays on the mammalian VUS. VUS MSH2-H639R demonstrates low repair levels in our assay, albeit reproducibly and significantly higher than controls. This is in agreement with minor mismatch-binding defects [Lützen et al., 2008], supporting the notion that the cell-free complementation assay can be used to detect subtle defects of VUS. However, the assay described here requires more validation to classify VUS displaying low to intermediate levels of repair. Also, the validation of this assay using a much larger panel of well-studied variants should prove its specificity and sensitivity. Ultimately, this may allow the determination of the relation between repair efficiency *in vitro* and cancer penetrance in suspected LS patients. These steps may be important for the future integration of the assay in a Bayesian integrated evaluation of MMR gene VUS [Goldgar et al., 2008].

Remarkably, besides the repair-deficient controls we have included, none of the *MSH6* VUS were repair-deficient, in line with other functional studies [Martinez and Kolodner, 2010; Kansikas et al., 2011]. Consequent to its redundancy with a minor MSH3-dependent MMR pathway, inactivating mutations in *MSH6* may be less penetrant, and therefore

less prevalent, in suspected LS patients than *MSH2* mutations. Suspected LS families are usually identified using the Amsterdam or Bethesda criteria that are biased towards high disease penetrance. Indeed, these criteria were recently shown to lack sensitivity to identify *MSH6* mutation carriers [Sjursen et al., 2010]. Alternatively, a lack of pathogenicity may be caused by a bias in MSH6 VUS databases. Pathogenic VUS may be underrepresented in the relatively small absolute number of MSH6 VUS currently registered, and the relative amount of pathogenic MSH6 VUS may increase along with the total number of VUS. Also, we cannot exclude that some VUS that show normal activity *in vitro* are non-functional *in vivo*, e.g. resulting from defects in splicing, stability or intracellular localization. Thus, although loss of activity in an *in vitro* MMR assay may be a strong indication of pathogenicity, wild type activity does not exclude pathogenicity of the VUS. For this reason, the latter class of MMR gene VUS might require additional analyses, including splice analysis [Tournier et al., 2008] and *in vivo* stability and localization assays [Lützen et al., 2008].

Together with our previous work describing a similar cell-free assay for the functional assessment of VUS in *MLH1* [Drost et al., 2010] rapid, reproducible and cell-free assays are now available for the analysis of VUS in all major genes, implicated in LS. We anticipate that these assays will be important tools for the development of a comprehensive diagnostic procedure for LS-associated VUS [Couch et al., 2008; Kansikas et al., 2011].

ACKNOWLEDGEMENTS

This work was supported by grants from the Nuts/Ohra foundation (grant SNO-T-03-06), the Dutch Digestive Foundation (grant MWO 05-16) and the European Union (grant FP6-018754). We thank Piya Temviriyanukul and Jacob G. Jansen for critical reading.

GRANT SUPPORT

This work was supported by grants from the Nuts/Ohra foundation (grant SNO-T-03-06), the Dutch Digestive Foundation (grant MWO 05-16) and the European Union (grant FP6-018754).

REFERENCES

Barnetson RA, Cartwright N, van Vliet A, Haq N, Drew K, Farrington S, Williams N, Warner J, Campbell H, Porteous ME, Dunlop MG. 2008. Classification of ambiguous mutations in DNA mismatch repair genes identified in a population-based study of colorectal cancer. Hum Mutat 29: 367-374.

Belvederesi L, Bianchi F, Galizia E, Loretelli C, Bracci R, Catalani R, Amati M, Cellerino R. 2008. MSH2 missense mutations and HNPCC syndrome: pathogenicity assessment in a human expression system. Hum Mutat 29: E296-E309.

Berends MJW, Wu Y, Sijmons RH, Mensink RGJ, Van Der Sluis T, Hordijk-Hos JM, De Vries EGE, Hollema H, Karrenbeld A, Buys CHCM, Van Der Zee AGJ, Hofstra RMW, 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.

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). Hum Mutat 29: 852-860.

Couch FJ, Rasmussen LJ, Hofstra R, Monteiro AN, Greenblatt MS, De Wind N, Group

IARCUGVW. 2008. Assessment of functional effects of unclassified genetic variants. Hum Mutat 29: 1314-1326.

De Jong A. 2004. Microsatellite Instability, Immunohistochemistry, and Additional PMS2 Staining in Suspected Hereditary Nonpolyposis Colorectal Cancer. Clinical Cancer Research 10: 972-980.

Drost M, Zonneveld JB, Van Dijk L, Morreau H, Tops CM, Vasen HFA, Wijnen J, De Wind N. 2010. A cell-free assay for the functional analysis of variants of the mismatch repair protein MLH1. Hum Mutat 31: 247-253.

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

Goldgar DE, Easton DF, Byrnes GB, Spurdle AB, Iversen ES, Greenblatt MM. 2008. Genetic Evidence and Integration of Various Data Sources for Classifying Uncertain Variants Into a Single Model. Hum Mutat 29:1265-1272

Heinen CD. 2010. Genotype to phenotype: Analyzing the effects of inherited mutations in colorectal cancer families. Mutat Res 693: 32-45.

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

Hsieh P, Yamane K. 2008. DNA mismatch repair: molecular mechanism, cancer, and ageing. Mech Ageing Dev 129: 391-407.

Jarvinen H, Renkonen-Sinisalo L, Aktan-Collan K, Peltomaki P, Aaltonen L, Mecklin J. 2009. Ten Years After Mutation Testing for Lynch Syndrome: Cancer Incidence and Outcome in Mutation-Positive and Mutation-Negative Family Members. Journal of Clinical Oncology 27: 4793-4797.

Jiricny J. 2003. DNA repair defects in colon cancer. Curr Opin Genet Dev 13: 61-69.

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

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. Hum Mol Genet 11: 1303-1310.

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

Lagerstedt Robinson K, Liu T, Vandrovcova J, Halvarsson B, Clendenning M, Frebourg T, Papadopoulos N, Kinzler KW, Vogelstein B, Peltomäki P, Kolodner RD, Nilbert M, Lindblom A. 2007. Lynch syndrome (hereditary nonpolyposis colorectal cancer) diagnostics. J Natl Cancer Inst 99: 291-299.

Li GM, Modrich PL. 1995. Restoration of mismatch repair to nuclear extracts of H6 colorectal tumor cells by a heterodimer of human MutL homologs. Proc Natl Acad Sci USA 92: 1950-1954.

Lucci-Cordisco E, Boccuto L, Neri G, Genuardi M. 2006. The use of microsatellite instability, immunohistochemistry and other variables in determining the clinical significance of MLH1 and MSH2 unclassified variants in Lynch syndrome. Cancer Biomark 2: 11-27.

Lützen A, de Wind ND, Georgijevic D, Nielsen FC, Rasmussen LJ. 2008. Functional analysis of HNPCC-related missense mutations in MSH2. Mutat Res 645: 44-55.

Martinez S, Kolodner RD. 2010. Functional analysis of human mismatch repair gene mutations identifies weak alleles and polymorphisms capable of polygenic interactions. Proc Natl Acad Sci USA 107: 5070-5075.

Mastrocola A, Heinen C. 2010. Lynch syndrome-associated mutations in MSH2 alter DNA repair and checkpoint response functions in vivo. Hum Mutat 31: E1699-E1708.

Niessen RC, Sijmons RH, Ou J, Olthof SG, Osinga J, Ligtenberg M, Hogervorst FB, Weiss M, Tops CM, Hes F, de Bock GH, Buys CH, Kleibeuker J, Hofstra RM. 2006. MUTYH and the mismatch repair system: partners in crime? Hum Genet 119: 206-211.

Ollila S, Dermadi Bebek D, Greenblatt M, Nyström M. 2008. Uncertain pathogenicity of MSH2 variants N127S and G322D challenges their classification. Int J Cancer 123: 720-724.

Ollila S, Sarantaus L, Kariola R, Chan P, Hampel H, Holinskifeder E, Macrae F, Kohonencorish M, Gerdes A, Peltomaki P. 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, Lützen A, Sijmons RH, Kleibeuker JH, de Wind ND, Rasmussen LJ, Hofstra RM. 2007. Functional analysis helps to clarify the clinical importance of unclassified variants in

DNA mismatch repair genes. Hum Mutat 28: 1047-1054.

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

Ramensky V, Bork P, Sunyaev S. 2002. Human non-synonymous SNPs: server and survey. Nucleic Acids Res 30:3894–3900.

Sjursen W, Haukanes BI, Grindedal EM, Aarset H, Stormorken A, Engebretsen LF, Jonsrud C, Bjørnevoll I, Andresen PA, Ariansen S, Lavik LA, Gilde B, Bowitz-Lothe IM, Maehle L, Møller P. 2010. Current clinical criteria for Lynch syndrome are not sensitive enough to identify MSH6 mutation carriers. J Med Genet 47: 579-585.

Tournier I, Vezain M, Martins A, Charbonnier F, Baert-Desurmont S, Olschwang S, Wang Q, Buisine MP, Soret J, Tazi J, Frébourg T, Tosi M. 2008. A large fraction of unclassified variants of the mismatch repair genes MLH1 and MSH2 is associated with splicing defects. Hum Mutat 29: 1412-1424.

Wang H, Hays JB. 2006. Construction of MMR plasmid substrates and analysis of MMR error correction and excision. Methods Mol Biol 314: 345-353.

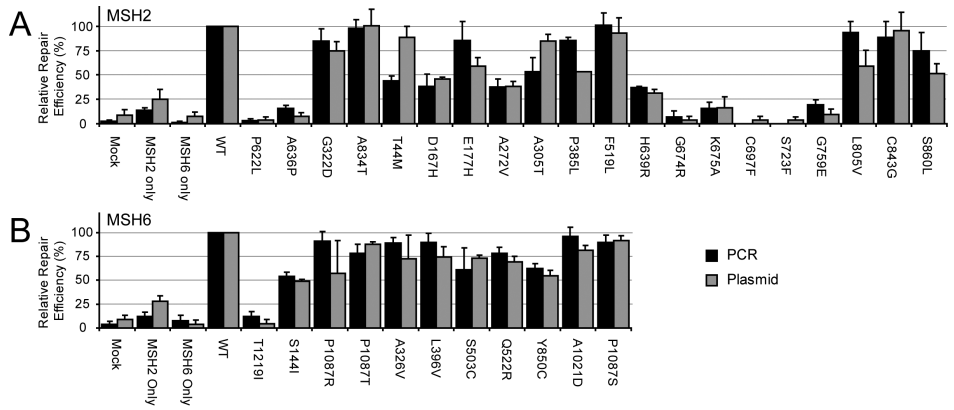
Wielders EA, Dekker RJ, Holt I, Morris GE, te Riele H. 2011. Characterization of MSH2 variants by endogenous gene modification in mouse embryonic stem cells. Hum Mutat 32: 389-396.

Wijnen J, Khan PM, Vasen H, van der Klift H, Mulder A, van Leeuwen-Cornelisse I, Bakker B, Losekoot M, Møller P, Fodde R. 1997. Hereditary nonpolyposis colorectal cancer families not complying with the Amsterdam criteria show extremely low frequency of mismatch-repair-gene mutations. American Journal of Human Genetics 61: 329-335.

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. Hum Mutat 28: 669-673.

Woods MO, Hyde AJ, Curtis FK, Stuckless S, Green JS, Pollett AF, Robb JD, Green RC, Croitoru ME, Careen A, Chaulk JA, Jegathesan J, McLaughlin JR, Gallinger SS, Younghusband HB, Bapat BV, Parfrey PS. 2005. High frequency of hereditary colorectal cancer in Newfoundland likely involves novel susceptibility genes. Clin Cancer Res 11: 6853-6861.

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



Supplementary Figure S1. *Comparison of MMR activity of VUS expressed in vitro from PCR products or from their cloned counterparts. (A) Relative repair efficiencies for MSH2 VUS, expressed from either PCR products (black bars, results are as in Fig. 3) or from plasmids (grey bars). Results are shown as mean±S.E.M. of 3-4 experiments. Mock: Mock expression. (B) As A, but for MSH6 VUS.*

Supplementary Table S1. Oligonucleotide primer sequences used for the generation of all MSH2 VUS using two-stage PCR amplification.

Mutant [a]	Forward/ Reverse [b]	Sequence 5'-3'
WT[d]	For	TCCCAGATCTTAACCGACTTGCCAAGAAGT
	Rev	ACTTCTTGGCAAGTCGGTTAAGATCTGGGA
T44M	For	CGACCGGGGCGACTTCTATA**T**GGCGCACGGCGAGGACG
	Rev	CGTCCTCGCCGTGCGCC**A**TATAGAAGTCGCCCCGGTCG
A45V	For	GGCGACTTCTATACGG**T**GCACGGCGAGGAC
	Rev	GTCCTCGCCGTGC**A**CCGTATAGAAGTCGCC
Y165D	For	GTTGGAGTTGGG**G**ATGTGGATTCCATACAG
	Rev	CTGTATGGAATCCACAT**C**CCCAACTCCAAC
D167H	For	GGAGTTGGGTATGTG**C**ATTCCATACAGAGG
	Rev	CCTCTGTATGGAAT**G**CACATACCCAACTCC
E177H	For	AGAGGAAACTAGGACTGTGT**C**A**C**TTCCCTGATAA
	Rev	TTATCAGGGAA**G**T**G**ACACAGTCCTAGTTTCCTCT
A272V	For	GCAGTTTCATCACTGTCTG**T**GGTAATCAAG
	Rev	CTTGATTACC**A**CAGACAGTGATGAAACTGC
A305T	For	GGATATT**A**CAGCAGTCAGAGCCCTTAACC
	Rev	GGTTAAGGGCTCTGACTGCTG**T**AATATCC
G322D	For	CTGTTGAAGATACCACTG**A**CTCTCAGTCTC
	Rev	GAGACTGAGAG**T**CAGTGGTATCTTCAACAG
P385L	For	TTCGTCGATTCC**T**AGATCTTAACCGACTTG
	Rev	CAAGTCGGTTAAGATCT**A**GGAATCGACGAA
L390F	For	TCCCAGATCTTAACCGA**T**TTGCCAAGAAGT
	Rev	ACTTCTTGGCAA**A**TCGGTTAAGATCTGGGA
F519L	For	AACTGGATTCCAGTGCACAG**C**TTGGATATT
	Rev	AATATCCAA**G**CTGTGCACTGGAATCCAGTT
N596del	For	GCAGACACTC*GATGTGTTAGCTCAGCTAGA*
	Rev	*TCTAGCTGAGCTAACACATC*GAGTGTCTGC
P622L	For	CCTGTTCCATATGTACGAC**T**AGCCATTTTG
	Rev	CAAAATGGCT**A**GTCGTACATATGGAACAGG
A636P	For	ATTATATTAAAA**C**CATCCAGGCATGCTTGT
	Rev	ACAAGCATGCCTGGATG**G**TTTTAATATAAT
H639R	For	CATCCAGGC**G**TGCTTGTGTTGAAGTTCAAG
	Rev	CTTGAACTTCAACACAAGCA**C**GCCTGGATG
G674R	For	ACTGGCCCCAATATGGGA**C**GTAAATCAAC
	Rev	TGTTGATTTAC**G**TCCCATATTGGGGCCAGT
K675A	For	ATATGGGAGGT**GCC**TCAACATATATTCGAC
	Rev	GTCGAATATATGTTGA**GGC**ACCTCCCATAT

Supplementary Table S1. *Continued*

Mutant [a]	Forward/ Reverse [b]	Sequence 5'-3'
C697F	For	TTTGTGCCAT**T**TGAGTCAGCAGAAGTGTCC
	Rev	GGACACTTCTGCTGACTCA**A**ATGGCACAAA
S723F	For	GGAGTCT**T**CACGTTCATGGCTGAAATGTTG
	Rev	CAACATTTCAGCCATGAACGTG**A**AGACTCC
G759E	For	TCTACCTACGATG**A**ATTTGGGTTAGCATGG
	Rev	CCATGCTAACCCAAAT**T**CATCGTAGGTAGA
L805V	For	ATGTCACAGCA**G**TCACCACTGAAGAGACCT
	Rev	AGGTCTCTTCAGTGGTGA**C**TGCTGTGACAT
A834T	For	TTCATGTTGCAGAGCTT**A**CTAATTTCCCTA
	Rev	TAGGGAAATTAG**T**AAGCTCTGCAACATGAA
C843G	For	TCCCTAAGCATGTAATAGAG**G**GTGCTAAAC
	Rev	GTTTAGCAC**C**CTCTATTACATGCTTAGGGA
S860L	For	TTCAGTATATTGGAGAAT**T**GCAAGGATATG
	Rev	CATATCCTTGC**A**ATTCTCCAATATACTGAA
E886G	For	GGTGAAAAAATTATTCAGG**G**GTTCCTGTCC
	Rev	GGACAGGAAC**C**CCTGAATAATTTTTTCACC
Common Primers		
External oligo 1	For	AATTAACCCTCACTAAAGGTGGTGCACTCTCAGTACAATC
External oligo 2	Rev	AGGCTTGTACATATTGTCGT
Tag Specific oligo	For	AATTAACCCTCACTAAAGG
Nested oligo	Rev	TTAGGTGACACTATAGAATAC

[a] Amino acid numbering is based on the cDNA with +1 corresponding to the translation initiation codon.
[b] "Forward" refers to mutagenic oligo 2 in Figure 1B, "Reverse" refers to mutagenic oligo 1 in Figure 1B.
[c] Bold, underlined font: VUS introducing nucleotide.
[d] Wild type primers do not introduce a mutation and have been chosen randomly.
[e] See Figure 1B for the positions of the external primers.

Supplementary Table S2. Oligonucleotide primer sequences used for the generation of all MSH6 VUS using two-stage PCR amplification.

Mutant [a]	Forward/ Reverse [b]	Sequence 5'-3' [c]
WT [d]	For	GACTCCAGAAATGATGGAGGCACGATGTAG
	Rev	CTACATCGTGCCTCCATCATTTCTGGAGTC
A20V	For	CCGGCGCTGAGTGATG**T**CAACAAGGCCTCG
	Rev	CGAGGCCTTGTTG**A**CATCACTCAGCGCCGG
A25S	For	GAGTGATGCCAACAAGGCCTCG**T**CCAGGG
	Rev	CCCTGG**A**CGAGGCCTTGTTGGCATCACTC
S144I	For	ACAAGGGGCTGGGTTA**T**CAAAAGGCTTTTA
	Rev	TAAAAGCCTTTTG**A**TAACCCAGCCCCTTGT
A326V	For	CAGCCACCAAACAAG**T**AACTAGCATTTCATCAG
	Rev	CTGATGAAATGCTAGTT**A**CTTGTTTGGTGGCTG
L396V	For	GCATCTACA**G**TCTATGTGCCTGAGGATTTC
	Rev	GAAATCCTCAGGCACATAGA**C**TGTAGATGC
M492V	For	CCAGAAATG**G**TGGAGGCACGATGTAGAAAG
	Rev	CTTTCTACATCGTGCCTCCA**C**CATTTCTGG
S503C	For	GATGGCACATATAT**G**CAAGTATGATAGAGTGGTGAG
	Rev	CTCACCACTCTATCATACTTG**C**ATATATGTGCCATC
Q552R	For	CATTACCAAGGGTACAC**G**GACTTACAGTGTGCTGGAAGG
	Rev	CCTTCCAGCACACTGTAAGTC**C**GTGTACCCTTGGTAATG
K610N	For	GGAAACTAAAACAATTCTAAA**T**AGTTCATTGTCC
	Rev	GGACAATGAACT**A**TTTAGAATTGTTTTAGTTTCC
Y850C	For	GTATGAAGAAACTACAT**G**CAGCAAGAAGAAG
	Rev	CTTCTTCTTGCTG**C**ATGTAGTTTCTTCATAC
V878A	For	GATCATGGAAGAAG**C**TGCTGATGGTTTTAAG
	Rev	CTTAAAACCATCAGCA**G**CTTCTTCCATGATC
R976H	For	GTAGGAACC**A**TTACCAGCTGGAAAT
	Rev	ATTTCCAGCTGGTAA**T**GGTTCCTAC
A1021D	For	CTAATCTCATAAATG**A**TGAAGAACGGAGGGATG
	Rev	CATCCCTCCGTTCTTCA**T**CATTTATGAGATTAG
D1026Y	For	GCTGAAGAACGGAGG**T**ATGTATCATTGAAG
	Rev	CTTCAATGATACAT**A**CCTCCGTTCTTCAGC
P1087R	For	GAAGATACCCC**C**CGCTTCTTAGAGCTTAAAG
	Rev	CTTTAAGCTCTAAGAAGCG**G**GGGGTATCTTC
P1087S	For	GAAGATACCCCC**T**CCTTCTTAGAGCTTAAAG
	Rev	CTTTAAGCTCTAAGAAGG**A**GGGGGTATCTTC
P1087T	For	GAAGATACCCCC**A**CCTTCTTAGAGCTTAAAG
	Rev	CTTTAAGCTCTAAGAAGG**T**GGGGGTATCTTC

Supplementary Table S2. *Contiuned*

Mutant [a]	Forward/ Reverse [b]	Sequence 5'-3' [c]
G1139S	For	CCAAATATGGGG**A**GCAAGTCTACGCTTATG
	Rev	CATAAGCGTAGACTTGC**T**CCCCATATTTGG
T1219I	For	GATGAATTAGGAAGAGGTA**T**TGCAACATTTGATG
	Rev	CATCAAATGTTGCA**A**TACCTCTTCCTAATTCATC
T1225M	For	CATTTGATGGGA**T**GGCAATAGCAAATGCAG
	Rev	CTGCATTTGCTATTGCC**A**TCCCATCAAATG

For explanation of [a], [b], [c], and [d], see Supplementary Table S1.
Common primers used for *MSH6* are identical to those used for *MSH2*.

A Cell-free Assay for the Functional Analysis of Variants of the Mismatch Repair Protein MLH1

4

Mark Drost, José B.M. Zonneveld, Linda van Dijk, Hans Morreau, Carli M. Tops, Hans F.A. Vasen, Juul T. Wijnen and Niels de Wind

Adapted from Drost et. al 2010 Hum Mutat 31:247-253

The methodology discussed in this chapter has contributed to additional publications/manuscripts:

Andersen SD, Liberti SE, Lützen A, **Drost M**, Bernstein I, Nilbert M, Dominguez M, Nyström M, Van Overeem Hansen T, Christoffersen JW, Jäger AC, de Wind N, Nielsen FC, Tørring PM, Rasmussen LJ. Functional characterization of MLH1 missense variants identified in Lynch Syndrome patients. *Hum Mutat 2012 33:1647–1655*

Farrel MP, Hughes DJ, **Drost M**, Cummings RJ, Fletcher TA, Meany MA, Andrews EJ, Kay EW, de Wind N, Power DG, Green AJ, Gallagher DJ. Multivariate analysis of MLH1 c.1664T>C (p.Leu555Pro) mismatch repair gene variant demonstrates its pathogenicity. *Fam Cancer 2013 12:741-747*

Parsons MT, Whiley PJ, Beesley J, **Drost M**, de Wind N, Thompson BA, Marquart L, Hopper JL, Jenkins MA; Australasian Colorectal Cancer Family Registry, Brown MA, Tucker K, Warwick L, Buchanan DD, Spurdle AB. Consequences of germline variation disrupting the constitutional translational initiation codon start sites of MLH1 and BRCA2: Use of potential alternative start sites and implications for predicting variant pathogenicity. *Mol Carcinog 2014 in press*

Dominguez Valentin M, **Drost M**, Therkildsen C, Rambech E, Ehrencrona H, de Wind N, Nilbert M, Rasmussen LJ. Functional Implications of the p.Cys680Arg mutation in the MLH1 Mismatch Repair Protein. *Under revision*

ABSTRACT

The hereditary colon and endometrium cancer predisposition Lynch Syndrome (also called HNPCC) is caused by a germ-line mutation in one of the DNA mismatch repair (MMR) genes. A significant fraction of the gene alterations detected in suspected Lynch Syndrome patients is comprised of amino acid substitutions. The relevance for cancer risk of these variants is difficult to assess as currently no time- and cost-effective, validated and widely applicable functional assays for the measurement of MMR activity are available. Here we describe a rapid, cell-free and easily quantifiable MMR activity assay for the diagnostic assessment of variants of the MLH1 MMR protein. This assay allows the parallel generation and functional analysis of a series of variants of the MLH1 *in vitro* using readily available, or pre-prepared, reagents. Using this assay we have tested 26 MLH1 variants and of these, 15 had lost activity. These results are in concordance with those obtained from first-generation assays and with *in silico* and pathology data. After its multifocal technical and clinical validation this assay could have great impact for the diagnosis and counseling of carriers of an MLH1 variant and their relatives.

INTRODUCTION

Lynch Syndrome (or Hereditary Non-Polyposis Colorectal Cancer, HNPCC; OMIM# 120435 and 609310) is an autosomal dominant disorder that predominantly predisposes to colon and endometrial cancer [Vasen et al., 2007; Stoffel et al., 2009]. Lynch Syndrome is caused by a germ-line mutation in one of the DNA mismatch repair (MMR) genes. In a Lynch syndrome patient loss of the wild-type allele leads to MMR deficiency which results in a spontaneous mutator phenotype that drives rapid carcinogenesis [Li, 2008].

A definitive diagnosis of Lynch syndrome can only be made by finding a causative germ-line mutation in an MMR gene [Vasen et al., 2007; Stoffel et al., 2009]. Additionally, the identification of a causative mutation allows the presymptomatic screening of affected relatives [Vasen et al., 2007]. Of all mutations found in suspected Lynch Syndrome patients, ~15% in MSH2 (OMIM# 609309), ~30% in MLH1 (OMIM# 120436) and ~40% in MSH6 (OMIM# 600678) give rise to single amino acid alterations [Peltomäki and Vasen, 2004; Lagerstedt Robinson et al., 2007] of which the pathological significance is often unclear (so-called Variants of Uncertain Significance, or VUS). Lack of classification of these variants precludes diagnosis for carriers and their relatives. Therefore, when a suspected Lynch syndrome patient carries a VUS, all first-degree relatives currently enroll lifelong periodic screening, irrespective of their mutation status [Castells et al., 2009]. Screening poses a physical and psychological burden on these families, but also a burden on the preventive health care apparatus. To enable correct diagnosis and to avoid unnecessary screening it is of great relevance to develop a widely applicable, validated and standardized approach for the diagnostic assessment of VUS in MMR genes.

The MMR pathway corrects misincorporations arising during DNA replication and thereby prevents the accumulation of spontaneous mutations [Hsieh and Yamane, 2008; Li, 2008]. MMR is initiated by the recognition of a misincorporation by the heterodimeric MutSα protein, consisting of the MSH2/MSH6 subunits. This triggers the recruitment of the MLH1/PMS2 heterodimeric protein (called MutLα) and of Exonuclease 1 that removes the misincorporation-containing DNA strand. Removal is directed towards the misincorporation-containing daughter strand by the presence of a strand discontinuity ('nick'), 5' or 3' of the mismatch. The excised DNA strand is subsequently resynthesized.

Functional MMR activity of a VUS is considered an important diagnostic criterion [Couch et al., 2008; Goldgar et al., 2008; Plon et al., 2008]. *In vitro* MMR activity assays have been developed [Li and Modrich, 1995] and in such a complementation assay the inability of a variant to restore MMR activity to a MMR-deficient cell extract almost certainly assigns pathogenicity to the VUS [Ou et al., 2007; Couch et al., 2008]. However, the requirement for molecular cloning, cell culture and the use of bacteria precludes their independent validation and wide applicability, complicating diagnosis and screening of suspected patients and their relatives.

Here, we present the analysis of *in vitro* MMR activity of 26 MLH1 VUS using a completely cell-free MMR complementation assay. The results from this assay are consistent with those from other functional assays and with *in vitro* predictions of protein function. We infer that this

assay is amenable to independent validation and applicable in molecular diagnostic laboratories and therefore could greatly contribute to the diagnostic classification of MLH1 VUS.

METHODS

Substrate preparation

The plasmid pUC19CPD has been used for the generation of MMR substrates [Wang and Hays, 2006] and was modified by introducing additional recognition sites for nicking endonucleases. The resulting plasmid was named pJH and, after the introduction of a mismatch (see below), allows to measure MMR activity at each DNA strand, directed by the presence of a nick either at the 5' or the 3' side of the mismatch (Figure 1A and Supp. Figure S1). The substrate for MMR was prepared from pJH, essentially as described [Wang and Hays, 2006], with modifications. Briefly, a 26 basepair (bp) single-stranded DNA stretch between two Nt.BstNBI sites was removed by heating and annealing to an excess of complementary oligonucleotide. The resulting gapped pJH was purified on Sephacryl S-400 HR colums (GE Healthcare, Piscataway, NJ). Phosphorylated oligonucleotide G/T-Fam, containing a mismatching nucleotide and an internal 6-FAM fluorescent label, was annealed into the gapped plasmid and ligated. After purification the resulting mismatched substrate was nicked at the bottom strand, 138 bp 3' to the mismatch with Nb.BsmI. Finally, the substrate, named pJHGT3'lnFAM (Figure 1B, top panel), was purified, aliquoted and frozen until use. All enzymes were purchased from New England Biolabs.

Preparation of Nuclear Extracts

HCT116 MutLα-deficient colon cancer cells [Jiricny, 2003] were grown at 37°C in RPMI-1640 (Gibco, Paisley, UK) supplemented with 10% FCS (PAA, Pasching, Austria), penicillin, streptomycin and Pyruvate (Gibco). Cells were harvested in log-phase and nuclear extracts were prepared as described [Holmes et al., 1990]. These extracts can be prepared in large quantities, distributed, and stored for prolonged periods.

Generation of MLH1 mutants by PCR and *in vitro* Expression of MMR VUS Proteins

The human *MLH1* (GenBank NM_000249.3) and *PMS2* (GenBank U14658.1) cDNAs were cloned into the pCITE4a vector (Novagen, Beeston, UK). The *MLH1* plasmid was used to generate *MLH1* mutants, containing vector-derived T7 promoter- and CITE sequences required for efficient expression *in vitro* in a two-step PCR procedure (Figure 2A-C). In the first round of PCR, overlapping mutant 5' and 3' *MLH1* fragments were generated separately. In a second round of amplification, the 5' and 3' overlapping mutant *MLH1* fragments were pooled and amplified to generate the full-length mutant *MLH1* gene. For all PCR reactions we used Pfx Platinum polymerase (Invitrogen). All oligonucleotide primer sequences can be found in Supp. Table S1. Nucleotide numbering reflects cDNA numbering with +1 corresponding to the A of the translation initiation codon in the GenBank reference sequence. Amino acid numbering is based on the cDNA with +1 corresponding to the translation initiation codon.

MLH1-mutant PCR fragments were directly used for *in vitro* protein expression in the TnT Quick Coupled Transcription/Translation Kit (Promega, Madison, WI) in the presence of PCR Enhancer*, according to the manufacturer's instructions. To quantify expression levels, proteins were labeled with ^{35}S-Methionine in parallel reactions. Alternatively, protein expression was investigated by Western blotting using rabbit α-MLH1 (Santa Cruz Biotechnology, Santa Cruz, CA, 1:1000) or mouse α-PMS2 (BD Pharmagen, 1:1.000). Proteins were visualized by enhanced chemiluminescence detection. Signals were quantified using OptiQuant software.

Detailed protocols for all procedures are available upon request.

In vitro MMR Complementation Assays

After estimation of quantities of expressed VUS MLH1 proteins by Western blotting, the proteins were diluted to similar concentrations in expression mix. This was followed by dimerization with *in vitro* produced wild type PMS2 protein in expression mix (1:1 vol/vol), for 30 minutes at room temperature.

Assay reactions were performed in a total volume of 25 µl containing 75 µg of HCT116 nuclear extract. The extract was complemented with 12 µl of *in vitro* produced MutLα and 100 ng of substrate pJHGT3'lnFAM. Reaction conditions for MMR were as described [Wang and Hays, 2006]. After the MMR reaction and purification using the MinElute kit (Qiagen, Germantown, MD), the substrate was digested with HinDIII and BsrBI (both Fermentas). One-fifth (2µl) of digested substrate was mixed with 8µl Hi-Di Formamide containing 0.2µl GeneScan-500 ROX size standards (Applied Biosystems) and fragment analysis was performed on a 3730 DNA analyzer (Applied Biosystems) at the Leiden Genome Technology Center. Data was analyzed using PeakScan Software (Applied Biosystems). Repair levels are calculated by dividing the height of the MMR-specific peak by the total fluorescent signal.

Generation of cloned mutants

Most mutations were additionally introduced into *MLH1*, cloned in pCITE4a, by site directed mutagenesis (Quick Change Site-directed Mutagenesis, Stratagene, LaJolla, CA). Primer sets (Biolegio, Nijmegen, The Netherlands) were designed (Supp. Table S2) and PCR was performed using Pfx Platinum polymerase (Invitrogen). PCR products were ligated and used to transform *E. coli*. Resulting plasmids were verified by direct sequencing of the entire mutant *MLH1* insert.

In silico analysis

All MLH1 VUS tested in this work were subjected to *in silico* analyses [Tavtigian et al., 2008] using the the Sorting Intolerant From Tolerant (SIFT) [Ng and Henikoff, 2003] and the MAPP-MMR [Chao et al., 2008] algorithms.

RESULTS

Construction of fluorescent heteroduplex substrate for MMR assays

For the current study we constructed a T/G mismatch-containing substrate, called pJHGT3'lnFAM (Figure 1B, top panel). MMR is directed towards the G-containing (bottom)

strand by a nick at that strand, 3' from the mismatch. Repair of the T/G mismatch to T/A by MMR creates a recognition site for HinDIII, within a 174 bp BsrBI restriction fragment. This generates a 75 bp fluorescent diagnostic fragment (Figures 1B and 1C, bottom panels). A 174 bp BsrBI fragment represents the unrepaired substrate.

A functional assay using *in vitro*-produced MutLα

Wild type *MLH1* and *PMS2* were cloned into pCITE4A. The *MLH1* clone was used as a template to amplify by PCR the open reading frame of the gene, together with sequences required for *in vitro* transcription/translation. MLH1 and PMS2 were expressed from the PCR fragment and from the expression vector, respectively, as judged by W blotting (Figure 1D, left panel). Importantly, the reticulocyte lysate was devoid of detectable rabbit MLH1 or PMS2 that might give false-positive results. Both proteins were synthesized in approximately equimolar amounts as judged by ^{35}S-Methionine levels in both proteins, using a parallel expression reaction (Figure 1D, right panel). Next, MLH1 and PMS2 were used to complement an MLH1-deficient HCT116 nuclear extract for the repair of substrate pJHGT3'lnFAM (Figure 1B). Using this assay, nearly half of the substrate was repaired (Figure 1C, left panel). This result is similar to the repair efficiency in other assays that use non-fluorescent substrates and *in vivo* produced MutLα [Raevaara et al., 2005; Plotz et al., 2006]. Importantly, mock-expression reactions and reactions containing only MLH1 or PMS2 did not result in significant repair (Figure 1C, right panel, see also Figure 3). Moreover, the MMR protein stochiometry in our MMR assay mimics the *in vivo* stochiometry, as evidenced by comparing MLH1:MSH6 ratios between an MMR complementation assay and wild type cells (Figure 1E). Taken together, these results indicate that our cell-free assay is suited for testing MLH1 activity.

Production of *MLH1* mutants in a two-step PCR procedure

To create *MLH1*-mutant genes suited for transcription/translation *in vitro* without the need of prior cloning we developed a two-step PCR-based method to produce 26 *MLH1* mutants together with regulatory sequences (Table 1, Figure 2A-D, Supp. Table S1 and Methods). This procedure was successful for all mutants, although final DNA yields varied somewhat (Figure 2E). Sequencing of the *MLH1*-mutant fragments showed no visible contamination with wild type input *MLH1* (Figure 2F and Supp. Figure S2). To investigate the PCR fidelity more extensively, a PCR product of one of the mutants was cloned and ten independent clones were fully sequenced. This revealed no mutations, additional to the desired one (not shown). We conclude that the PCR-based approach to generate mutant *MLH1* genes for expression *in vitro* is efficient and free of detectable artifacts.

MMR activities of *in vitro* produced *MLH1* VUS

The panel of 26 *MLH1* mutants included the three putative polymorphisms and four presumed pathogenic mutants (for references see Table 1). All mutants are described in the database of the International Society for Gastrointestinal Hereditary Tumours [Peltomäki and Vasen, 2004] (http://chromium.liacs.nl/LOVD2/colon_cancer) and most of them also in the MMR Gene Unclassified Variants Database [Ou et al., 2008] (http://www.mmrmissense.net).

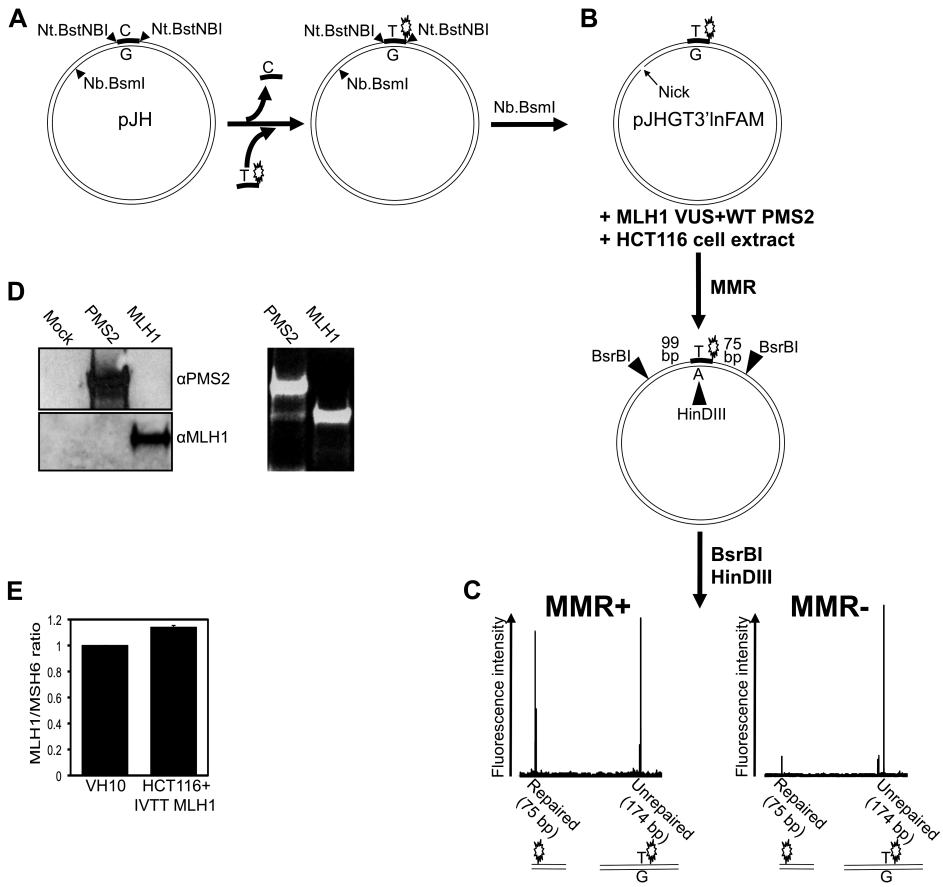


Figure 1. *Outline of the cell free MMR complementation assay. (A). Plasmid pJH. Bold line: single-stranded DNA fragment that can be removed by cleavage with Nt.BstNBI and heating. To generate the substrate pJHGT3'lnFAM, the latter fragment is replaced with a mismatch- and fluorescent label (star)-containing fragment. Then, the bottom strand is nicked at Nb.BsmI to provide strand specificity to the MMR reaction. See Supp. Figure S1 for more details. (B). Flow-scheme of the MMR assay. Top panel: MMR substrate, derived from pJH. To the substrate in vitro expressed MutLα-variant protein and MutLα-deficient HCT116 cell extract are added in the presence of cofactors. After incubation, substrate is purified and cleaved with BsrBI and HinDIII (bottom panel). The latter enzyme only cleaves when the mismatch is repaired, resulting in a diagnostic fluorescent fragment of 75 bp. Numbers: sizes of restriction sites. (C). Examples of automated fragment analysis after a MMR assay. Left panel: Fragment analysis after testing an MMR-proficient variant. Right panel: Example of a fragment analysis chromatogram after testing of a mock MutLα expression. (D). Left panel: Western blot showing PMS2 and MLH1 proteins produced in vitro from PCR-amplified genes, as well as a mock expression. Right panel: ^{35}S-Methionine-labeled PMS2 and MLH1, expressed in vitro from cloned PMS2, and from PCR-amplified MLH1 genes. (E). Quantification of relative MLH1:MSH6 ratios in VH10 nuclear extracts and in a MMR assay, complemented with in vitro synthesized MutLα.*

An overview of pathology data is shown in Table 1. Mutant *MLH1* PCR fragments (Figure 2E) were used to produce proteins *in vitro* (Figure 2G). No relation between quantities of the input PCR fragments and the amount of expressed protein was observed, and incidental

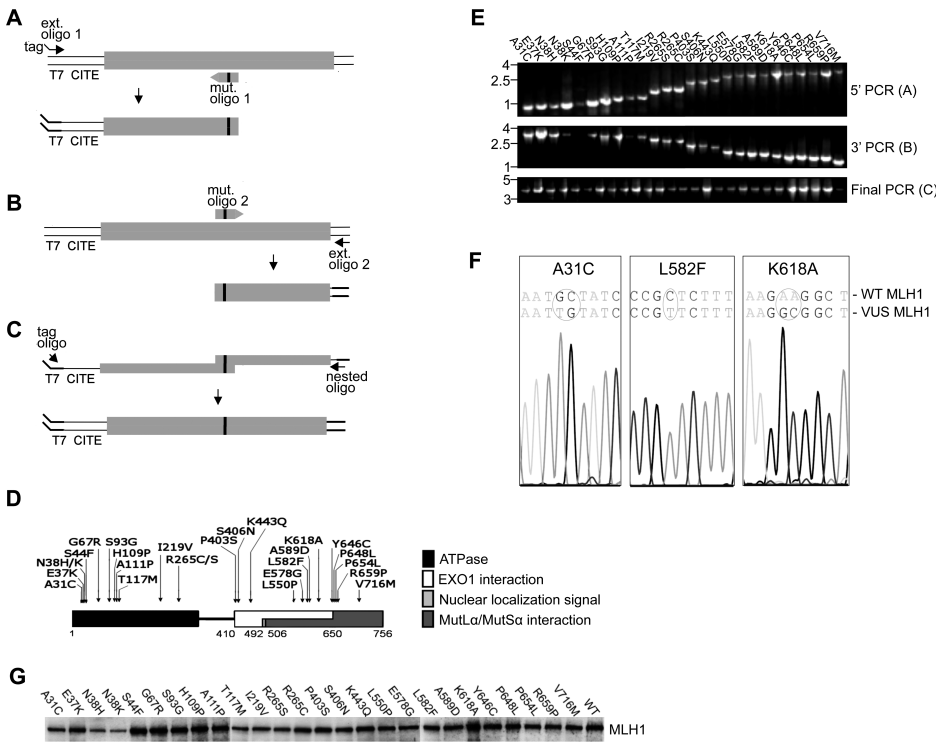


Figure 2. *In vitro generation of MLH1-variant genes and proteins. (A,B). PCR design and primer positions for generating 5' and 3' MLH1-mutant 'half' fragments, respectively. (C). The 5' and 3' mutant MLH1 fragments are joined in a second PCR step, resulting in mutant MLH1, preceded by regulatory sequences. (D). Functional domains of MLH1 [Raevaara et al., 2005; Wanat et al., 2006], with all 26 constructed and tested VUS. Amino acid numbering is based on the cDNA with +1 corresponding to the translation initiation codon. (E). Representative agarose gel picture of the PCR-based generation of 26 MLH1-mutant PCR fragments. Top panel: 5' mutant 'half' fragments, middle panel: 3' mutant 'half' fragments, bottom panel: full-length mutant MLH1 fragments. Left: size markers (kbp). (F). Example of sequencing results from three randomly chosen PCR-generated full-length MLH1-mutant fragments. Top sequence: wild type sequence Bottom sequence: mutant sequence. The desired nucleotide alterations are encircled. See Supp. Figure 2 for a color version. (G). Representative Western blot of all 26 in vitro produced MLH1-variant proteins, expressed directly from mutagenic PCR-generated templates.*

differences in expression levels were not reproducible between experiments (not shown). Next, approximately equal amounts of each protein were tested for MMR activity. The putative polymorphisms all repaired more than 84% of wild type, whereas the presumed pathogenic mutants all repaired less than 20% (Figure 3, black bars and Supp. Figure S3). This difference allows us to selectively identify pathogenic *MLH1* mutants. In the other 19 *MLH1* mutants we found a large variation in MMR activity (Figure 3). As a tentative cutoff for MMR deficiency we assumed a repair level that was not significantly higher (Student's two-tailed T-test, $p<0.05$) than the average repair efficiency of the presumed pathogenic mutants. In support, all repair deficient VUS were predicted to be deleterious using SIFT analysis and most also by MAPP-MMR, an algorithm focused at MMR gene variants (Table 1).

Repair efficiencies of cloned mutants mirror repair efficiencies of PCR-based mutants

To further investigate the robustness of the PCR product-based expression system, we investigated the activity of MLH1 VUS, expressed after their cloning. Although the activities of cloned VUS reflected those of PCR-generated VUS (Figure 3), repair levels for the MMR-proficient cloned VUS were somewhat higher than those of the PCR-based VUS, due to the presence of MMR-inhibitory contaminants in the PCR-based expressions (not shown). Nevertheless, the generation of *MLH1* VUS in a two-step PCR followed by their direct expression provides a rapid and reliable alternative for cloning.

MMR VUS classification does not critically depend on protein levels

As the expression levels between individual variant proteins was subject to some random variation (see above) we wanted to investigate whether the MMR assay is tolerant to variations in MLH1 levels. To this end, wild type MLH1, MMR-proficient VUS E578G and MMR-deficient VUS G67R were expressed from PCR products (Figure 2G). Then, the reactions were diluted up to twofold such that each contained a similar concentration of MLH1. During further serial dilution, both the wild type protein and E578G retained activity that discriminates them from G67R (Figure 4). Remarkably, at increasing dilutions the activity of E578G was more affected than of wild type MLH1, indicating that the mutation may confer partial loss of function. This result shows that the assay is tolerant to stochastic variations in *in vitro* expression levels of the MLH1 variants.

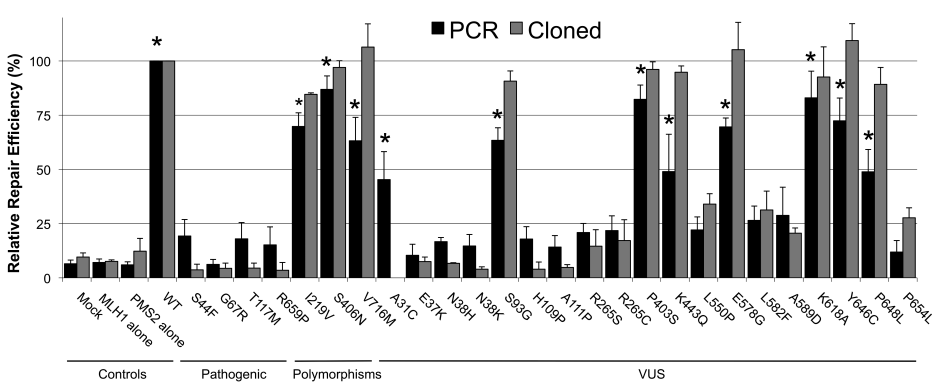


Figure 3. *MMR activity as measured in the in vitro MMR complementation assay. MMR activity assays of PCR fragment-based MLH1-variant proteins (black bars) and of plasmid clone-based MLH1 variant proteins (grey bars). Pathogenic: Putative pathogenic Lynch Syndrome VUS, Polymorphisms: Putative MLH1 polymorphisms. Results are mean±S.E.M. of at least three completely independent (from PCR to MMR assay) experiments. *: Significantly higher repair activity as compared to the four putatively pathogenic variants (two-tailed T test, p<0.05). The repair activity of wild type MLH1 was set at 100%. Representative examples of fragment analysis diagrams for each variant can be found in Supp. Figure S3.*

Table 1. *In silico* data, pathology data and references for all MLH1 VUS tested in this work.

Mutation		Pathology[¶]				*In silico* analysis		References to functional studies[§]			
			IHC					Other MMR assays		Other functional assays	
Protein	DNA	MSI	MLH1	PMS2	*In vitro* MMR	SIFT[¥]	MAPP- MMR[π]	Deficient	Proficient	Deficient	Proficient
A31C	c.91_92GC>TG	NA	+	+	+	0	3.26				
E37K	c.109G>A	H	+	-	-	0	18.32				
N38H	c.112A>C	H	+	+	-	0	14.29				
N38K	c.114C>G	H	+	+	-	0	25.31				
S44F	c.131C>T	H	-	NA	-	0	16.74	1		2, 3, 4	
G67R	c.199G>A	H	-	-	-	0	36.52	1, 5		2, 3, 6, 7	
S93G	c.277A>G	NA	+	+	+	0.06	2.47		1, 5, 8		9
H109P	c.326A>C	H	NA	NA	-	0	9.36				
A111P	c.331G>C	H	-	NA	-	0.01	14.1				
T117M	c.350C>T	NA	-	-	-	0	21.05	1, 10		2, 3, 11	
I219V	c.655A>G	NA	NA	NA	+	0.36	3.51		1, 5		6, 11, 12
R265S	c.793C>A	NA	NA	NA	-	0	21.16			9	
R265C	c.793C>T	H	+	NA	-	0	27.43	1, 10	1, 2	9, 13	

P403S	c.1207C>T	H/L	+	NA	+	0.78	1.8				
S406N	c.1217G>A	H	NA	NA	+	1	1.07		1		3, 9
K443Q	c.1327A>C	NA	-	NA	+	0.65	2.19		5		
L550P	c.1649T>C	H	-	NA	-	0.04	8.88		5		
E578G	c.1733A>G	H	+	NA	+	0.21	4.85	1		2, 3	4
L582F	c.1744C>T	NA	-	-	-	0.02	2.24				
A589D	c.1766C>A	H	-	NA	-	0.01	8.32		5		
K618A	c.1852AA>GC	H	+	NA	+	0.01	5.06		5	3, 4, 6, 9, 13	
Y646C	c.1937A>G	L	+	NA	+	0.02	3.69		5		
P648L	c.1943C>T	H	+	NA	-	0	6.87	1	5		
P654L	c.1961C>T	H	-	NA	-	0	19.2	1	5	3	
R659P	c.1976G>C	H	NA	NA	-	0.01	11.7	1, 5, 8		2, 4	
V716M	c.2146G>A	H	-	NA	+	0.16	2.78		1, 5		

[¶]MSI=Microsatellite Instability. H=High, L=Low, NA=Not Available. IHC=Immunohistochemistry.
[¥]SIFT scores as calculated on the SIFT webpage (http://sift.jcvi.org/). Scores <0.06 are considered pathogenic/not tolerated.
[π]MAPP-MMR [Chao et al., 2008] scores were calculated from http://mappmmr.blueankh.com/. Values >4.55 are considered pathogenic.
[§]References: 1 [Takahashi et al., 2007], 2 [Shimodaira et al., 1998], 3 [Kondo et al., 2003], 4 [Guerrette et al., 1999], 5 [Raevaara et al., 2005], 6 [Blasi et al., 2006], 7 [Avdievich et al., 2008], 8 [Nyström-Lahti et al., 2002], 9 [Wanat et al., 2007], 10 [Plotz et al., 2006], 11 [Spina et al., 2008], 12 [Trojan et al., 2002], 13 [Perera and Bapat, 2008].

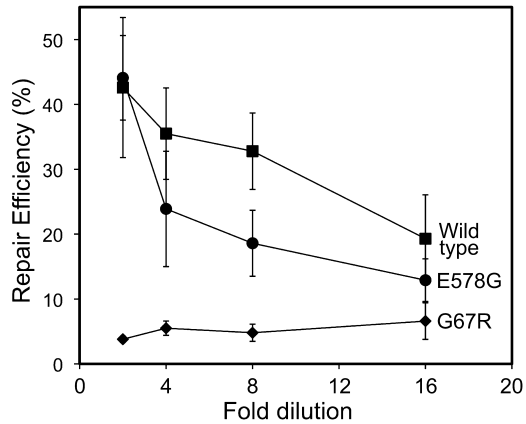


Figure 4. *Limited dependence of MMR activity on MLH1 concentration. To achieve similar protein concentrations, wild type MLH1 and variant G67R (Figure 2G) were initially diluted twofold. After further dilution, MMR activity was tested for all three proteins. Error bars: S.E.M.*

DISCUSSION

Proper genetic counseling of suspected Lynch Syndrome patients requires an accurate diagnosis based on the nature of the germ line mutation found. Cancer risk resulting from non-truncating, and especially missense, variants is difficult to predict. This results in anxiety and, frequently, unnecessary periodic screening in carriers of such a variant, but also in both VUS-carrying and non-carrying relatives. In this work we present a cell-free assay for the functional analysis of MLH1 VUS that enables the rapid identification of pathogenic variants.

Repair deficiency for a VUS in our assay was, with few exceptions, consistent with the *in silico* prediction of pathogenicity and with pathology data (Table 1). Some of the same VUS were analyzed before by others and most *MLH1* VUS that were inactive in our assay were also inactive in other assays. Notable exceptions are VUS R265C that, although classified repair-deficient in three studies [Plotz et al., 2006; Takahashi et al., 2007; Wanat et al., 2007], was classified proficient in another [Trojan et al., 2002]. Additionally, L550P, A589D and P654L were active in another assay [Raevaara et al., 2005] but inactive in yeast-based assays [Takahashi et al., 2007; Wanat et al., 2007]. The latter VUS are located at the protein-protein interaction domain (Figure 2G). *In vitro* binding studies (Andersen et al., personal communication) support a defect for these variants and another study has shown binding defects for other VUS in this region of the protein [Guerrette et al., 1999]. Furthermore, the observed lack of nuclear import of PMS2 by these VUS [Raevaara et al., 2005] supports a dimerization defect, since heterodimerization of MLH1 and PMS2 stimulates nuclear import of MutLα [Knudsen et al., 2009]. Additionally, all four VUS that show discrepant results (R265C, L550P, A589D and P654L) are predicted to be defective by *in silico* analyses and were associated with microsatellite instability, a hallmark of MMR deficiency (Table 1). For these reasons, we conclude that a lack of activity in our assay is in agreement with most other functional data, with pathology data and with *in silico* analysis.

Since a loss of MMR is strongly predictive for Lynch Syndrome, the positive predictive value of our assay is bound to be high. Further validation of our assay seems warranted, using a much larger group of accepted polymorphisms and pathogenic mutations from well-studied patients. This could aid in determining the relation between MMR activity and disease penetrance, enabling to assess pathogenicity of VUS that display a low- to intermediate level of repair. However, as our assay does not address specific *in vivo* MMR defects, such as a splicing, stability or intracellular localization, nor subtle *in vitro* defects, the negative predictive value will probably be lower. As an example, E578G displayed reduction in MMR activity only after its dilution (Figure 4) that may reflect a subtle dimerization defect. We infer that, in addition to a cell-free complementation assay as described here, additional functional data might be used for final diagnosis of VUS that display residual activity. This may include nuclear localization assays [Raevaara et al., 2005], splicing assays [Tournier et al., 2008], protein-protein interaction studies [Guerrette et al., 1999] or stability studies [Perera and Bapat, 2008]. A combination of different analytical approaches, pathological and family data should result in a comprehensive, validated, diagnostic procedure of VUS in MLH1 and other MMR genes [Plon et al., 2008].

In conclusion, with this assay, we have developed a simple, rapid, cost-effective and reliable method that may fill a niche in diagnostic labs [Heinen, 2009]. Cell extracts, substrate and other ingredients can be prepared, stored and distributed in large quantities. The full assay (recreation of the mutants by PCR, *in vitro* expression and *in vitro* MMR assay) can be performed in three to four days, and many mutants can be tested in parallel. The procedure can be used to screen the large amount of mutants already detected, a number that, due to the advent of personalized genomics, is likely to rise sharply. Alternatively, the assay may be used for high-resolution *a priori* VUS analyses as was done for p53 [Kato et al., 2003]. Currently we are developing and testing very similar cell-free assays to test functional activity of VUS in the *MSH2* and *MSH6* MMR genes.

ACKNOWLEDGEMENTS

We acknowledge Prof. John B. Hays and Prof. Josef Jiricny for providing plasmids and Dr. Giancarlo Marra for help. Drs. Anne Lützen, Sofie Dabros Andersen and Prof. Lene Juel Rasmussen are thanked for *MLH1*-mutant plasmids, sharing unpublished data and helpful discussions. We thank Drs. Anastasia Tsaalbi-Shtylik, Cristina Ferrás, Jaap Jansen and Prof. Robert Hofstra for critical reading.

GRANT SUPPORT

This work was supported by grants from the Nuts/Ohra foundation (SNO-T-03-06), the Dutch Digestive Foundation (MWO 05-16) and the European Union (FP6-018754).

REFERENCES

Castells A, Castellví-Bel S, Balaguer F. 2009. Concepts in familial colorectal cancer: where do we stand and what is the future? Gastroenterology 137: 404-409.

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). Hum Mutat 29: 852-860.

Couch FJ, Rasmussen LJ, Hofstra R, Monteiro AN, Greenblatt MS, de Wind N, Group IARCUGVW. 2008. Assessment of functional effects of unclassified genetic variants. Hum Mutat 29: 1314-1326.

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

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

Heinen CD. 2010. Genotype to phenotype: Analyzing the effects of inherited mutations in colorectal cancer families. Mutat Res 693: 32-45.

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

Hsieh P, Yamane K. 2008. DNA mismatch repair: molecular mechanism, cancer, and ageing. Mech Ageing Dev 129: 391-407.

Jiricny J. 2003. DNA repair defects in colon cancer. Curr Opin Genet Dev 13: 61-69.

Kato S, Han S-Y, Otsuka K, Shibata H, Kanamaru R, Ishioka C. 2003. Understanding the function–structure and function–mutation relationships of p53 tumor suppressor protein by high-resolution missense mutation analysis. Proc Natl Acad Sci USA 14: 8424-8429.

Knudsen NØ, Andersen SD, Lützen A, Nielsen FC, Rasmussen LJ. 2009. Nuclear translocation contributes to regulation of DNA excision repair activities. DNA Repair 8: 682-689.

Lagerstedt Robinson K, Liu T, Vandrovcova J, Halvarsson B, Clendenning M, Frebourg T, Papadopoulos N, Kinzler KW, Vogelstein B, Peltomäki P, Kolodner RD, Nilbert M, Lindblom A. 2007. Lynch syndrome (hereditary nonpolyposis colorectal cancer) diagnostics. J Natl Cancer Inst 99: 291-299.

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

Li GM, Modrich P. 1995. Restoration of mismatch repair to nuclear extracts of H6 colorectal tumor cells by a heterodimer of human MutL homologs. Proc Natl Acad Sci USA 92: 1950-1954.

Ng PC, Henikoff S. 2003. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Research 31: 3812-3814.

Ou J, Niessen RC, Lützen 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. Hum Mutat 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. Hum Mutat 29: 1337-1341.

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

Perera S, Bapat B. 2008. The MLH1 variants p. Arg265Cys and p. Lys618Ala affect protein stability while p. Leu749Gln affects Heterodimer Formation. Hum Mutat 994: Online.

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

Plotz G, Welsch C, Giron-Monzon L, Friedhoff P, Albrecht M, Piiper A, Biondi RM, Lengauer T, Zeuzem S, Raedle J. 2006. Mutations in the MutSalpha interaction interface of MLH1 can abolish DNA mismatch repair. Nucleic Acids Res 34: 6574-6586.

Raevaara T, Korhonen M, Lohi H, Hampel H, Lynch E, Lonnqvist K, Holinskifeder E, Sutter C, Mckinnon W, Duraisamy S. 2005. Functional Significance and Clinical Phenotype of Nontruncating Mismatch Repair Variants of MLH1. Gastroenterology 129: 537-549.

Stoffel E, Mukherjee B, Raymond VM, Tayob N, Kastrinos F, Sparr J, Wang F, Bandipalliam P, Syngal S, Gruber SB. 2009. Calculation of

Risk of Colorectal and Endometrial Cancer Among Patients With Lynch Syndrome. Gastroenterology 137: 1621-1627.

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

Tavtigian SV, Greenblatt MS, Lesueur F, Byrnes GB, Group IARCUGVW. 2008. In silico analysis of missense substitutions using sequence-alignment based methods. Hum Mutat 29: 1327-1336.

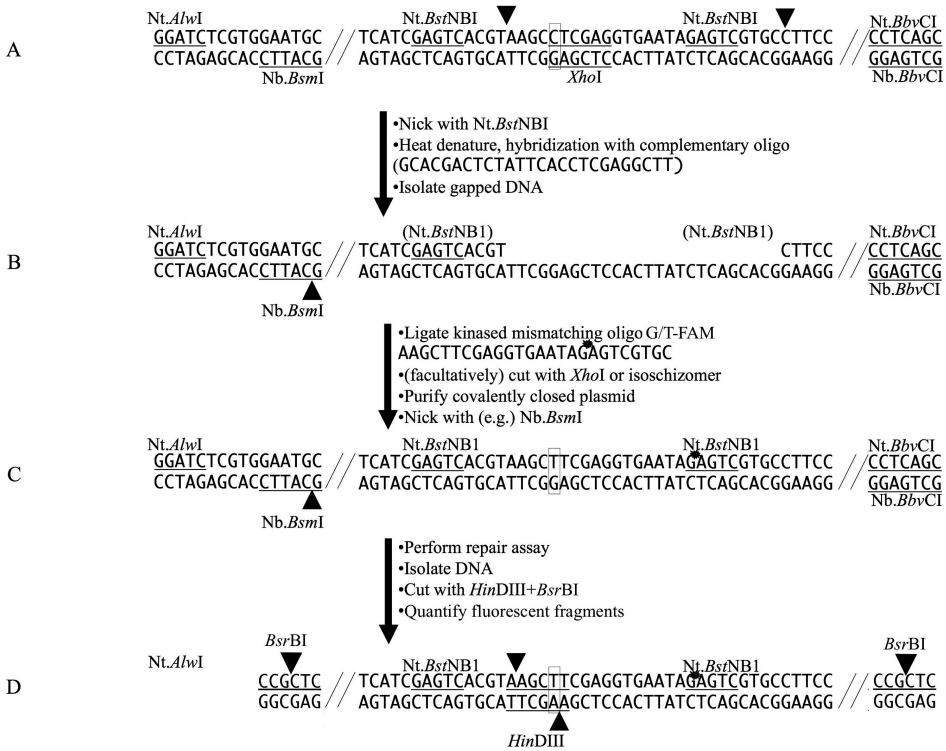
Tournier I, Vezain M, Martins A, Charbonnier F, Baert-Desurmont S, Olschwang S, Wang Q, Buisine MP, Soret J, Tazi J, Frébourg T, Tosi M. 2008. A large fraction of unclassified variants of the mismatch repair genes MLH1 and MSH2 is associated with splicing defects. Hum Mutat 29: 1412-1424.

Trojan J, Zeuzem S, Randolph A, Hemmerle C, Brieger A, Raedle J, Plotz G, Jiricny J, Marra G. 2002. Functional Analysis of hMLH1 Variants and HNPPC-Related Mutations Using a Human Expression System. Gastroenterology 122: 211-219.

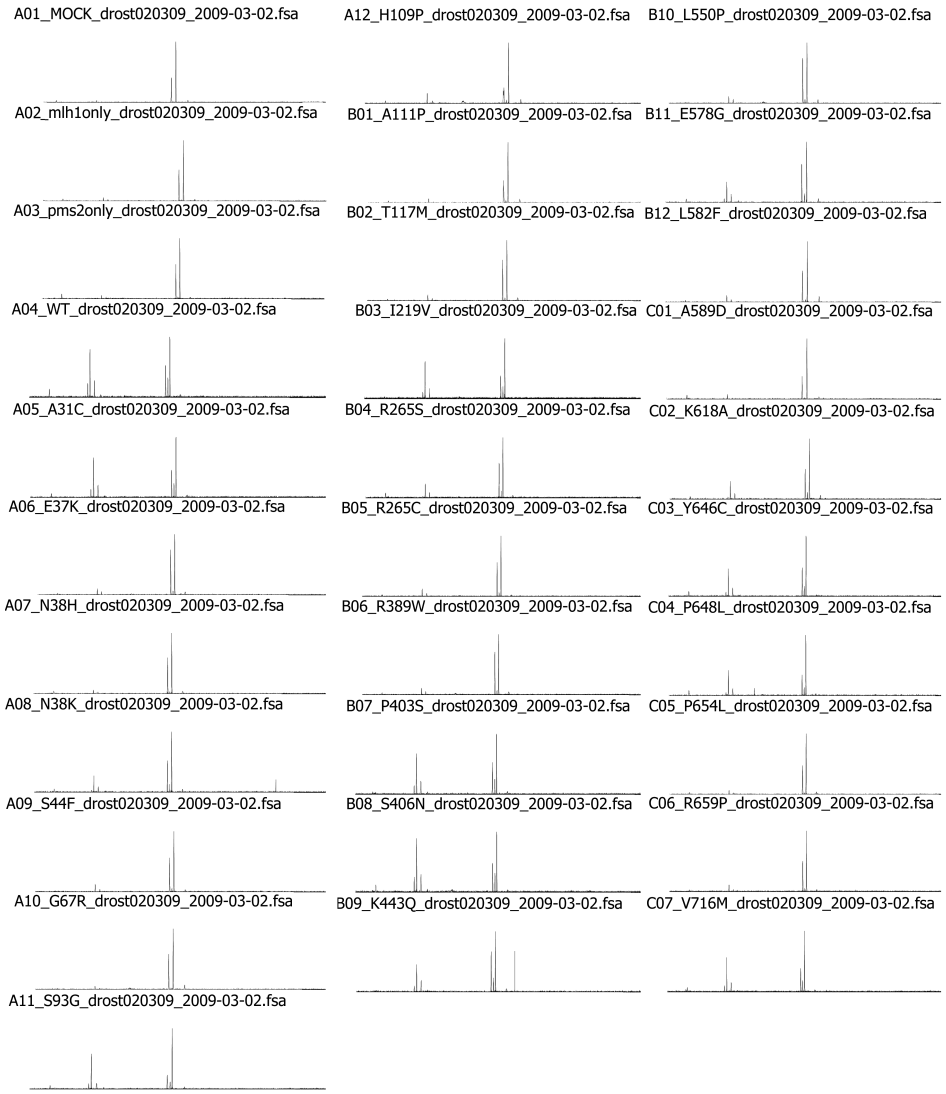
Vasen HFA, Möslein G, Alonso A, Bernstein I. 2007. Guidelines for the clinical management of Lynch syndrome (hereditary non-polyposis cancer). British Medical Journal 44: 353-362.

Wanat JJ, Singh N, Alani E. 2007. The effect of genetic background on the function of Saccharomyces cerevisiae mlh1 alleles that correspond to HNPCC missense mutations. Hum Mol Genet 16: 445-452.

Wang H, Hays JB. 2006. Construction of MMR plasmid substrates and analysis of MMR error correction and excision. Methods Mol Biol 314: 345-353.



Supplementary Figure S1. *Schematic overview of substrate production at the nucleotide level. (A) Relevant sequences of plasmid pJH. (B) A single-stranded DNA stretch of 26 nucleotides between two NtBstNBI nicking endonuclease sites is removed. (C) In the single-stranded DNA gap, an oligonucleotide that introduces a T/G mismatch and in addition contains a fluorescent label (star) is annealed and ligated. Then, the bottom strand is nicked 138 nucleotides 3' of the T/G mismatch with Nb.BsmI. This directs the MMR reaction towards the lower, G-containing strand. (D) MMR-mediated correction of the T/G mismatch results in the generation of a diagnostic HinDIII restriction site. Double cleavage with HinDIII and BsrBI (not shown) results in a diagnostic fluorescent fragment of 75 bp.*



Supplementary Figure S2. *Functional assessment of MLH1 variants in vitro. The Figure represents MMR assays for a mock reaction and 26 MLH1 variants, performed in a single run. See Figure 1 for an explanation. Occasionally, minor ghost peaks are seen.*

Supplementary Table S1. Primer sequences used for the generation of all *MLH1* mutants using the two-step PCR amplification.

Mutant	Forward/ Reverse	Sequence
Variant introducing primers		
WT[1]	For	CCACAACAAGTCTGACCTCGTCTTCTAC
	Rev	GTAGAAGACGAGGTCAGACTTGTTGTGG
A31C	For	ATCCAGCGGCCAGCTAAT**TGT**ATCAAAGAG
	Rev	CTCTTTGAT**ACA**ATTAGCTGGCCGCTGGAT
E37K	For	GCTATCAAAGAGATGATT**A**AGAACTGTTTAGATGC
	Rev	GCATCTAAACAGTTCT**T**AATCATCTCTTTGATAGC
N38H	For	GAGATGATTGAG**C**ACTGTTTAG
	Rev	CTAAACAGT**G**CTCAATCATCTC
N38K	For	AGATGATTGAGAA**G**TGTTTAGATGCAAAATCC
	Rev	GGATTTTGCATCTAAACA**C**TTCTCAATCATCT
S44F	For	GTTTAGATGCAAAA**T**TCACAAG
	Rev	CTTGT**G**AATTTTGCATCTAAAC
G67R	For	AATGGCACC**A**GGATCAGGAA
	Rev	TTCCTGATCC**T**GGTGCCATT
H109P	For	CCAGCATAAGCC**C**TGTGGCTCATGTTACTA
	Rev	TAGTAACATGAGCCACA**G**GGCTTATGCTGG
A111P	For	GCCATGTG**C**CTCATGTTACT
	Rev	AGTAACATGAG**G**CACATGGC
T117M	For	CAA**T**GAAAACAGCTGATGG
	Rev	CCATCAGCTGTTTTC**A**TTG
I219V	For	TCGCTCC**G**TCTTTGGAAATG
	Rev	CATTTCCAAAGA**C**GGAGCGA
R265S	For	CTTCATCAACCAT**A**GTCTGGTAG
	Rev	CTACCAGAC**T**ATGGTTGATGAAG
R265C	For	CATCAACCAT**T**GTCTGGTAG
	Rev	CTACCAGAC**A**ATGGTTGATG
R389W	For	GGTTCGTACAGATTCC**T**GGGAACAGA
	Rev	TCTGTTCCC**A**GGAATCTGTACGAACC
P403S	For	GCCTCTGAGCAAA**T**CCCTGTCCAGTC
	Rev	GACTGGACAGGG**A**TTTGCTCAGAGGC
S406N	For	GCCTCTGAGCAAACCCCTGTCCA**A**TCAGCC
	Rev	GGCTGA**T**TGGACAGGGGTTTGCTCAGAGGC
K443Q	For	CCCTGCTGAAGTGGCTGCC**C**AAAATCAGAG
	Rev	CTCTGATTTT**G**GGCAGCCACTTCAGCAGGG
L550P	For	GTTATACCTTC**C**CAACACCACCAAGC
	Rev	GCTTGGTGGTGTTG**G**GAAGGTATAAC

Supplementary Table S1. *Continued*

Mutant	Forward/ Reverse	Sequence
Variant introducing primers		
E578G	For	AATTTTGGTGTTCTCAGGTTATCGG**G**GCCAGCA
	Rev	TGCTGGC**C**CCGATAACCTGAGAACACCAAAATT
L582F	For	CAGCACCG**T**TCTTTGACCTTGCCATGCTTG
	Rev	CAAGCATGGCAAGGTCAAAGA**A**CGGTGCTG
A589D	For	GACCTTGCCATGCTTG**A**CTTAGATAG
	Rev	CTATCTAAG**T**CAAGCATGGCAAGGTC
K618A	For	AGTTTCTGAAGAAG**GC**GGCTGAGATG
	Rev	CATCTCAGCC**GC**CTTCTTCAGAAACT
Y646C	For	TACCCCTTCTGATTGACAACT**G**TGTG
	Rev	CACA**C**AGTTGTCAATCAGAAGGGGTA
P648L	For	CCTTCTGATTGACAACTATGTGC**T**CCCTTTGG
	Rev	CCAAAGGG**A**GCACATAGTTGTCAATCAGAAGG
P654L	For	GGACTGC**T**TATCTTCATTCT
	Rev	AGAATGAAGATA**A**GCAGTCC
R659P	For	CCTATCTTCATTCTTC**C**ACTAGCCACTGAG
	Rev	CTCAGTGGCTAGT**G**GAAGAATGAAGATAGG
V716M	For	CCTGGAAGTGGACT**A**TGGAACA
	Rev	TGTTCCA**T**AGTCCACTTCCAGG
General primers		
Tag	For	TTAGGTGACACTATAGAATAC
External oligo 1	For	TTAGGTGACACTATAGAATACTCAGGCTGCGCAACTGTTGG
Nested oligo 2	Rev	TCGCCACCTCTGACTTGAGC
pCite 1	Rev	CAGTGGCGATAAGTCGTGTCTTACCG
pCite 2	Rev	CGAACTGAGATACCTACAGCGTGAGCTATG

[1] Wild type primers do not introduce a mutation and have been chosen randomly.
All primers are shown from 5' to 3' polarity
Bold, underlined font: mutant nucleotide.

Supplementary Table S2. Primer sequences for site-directed mutagenesis of cloned *MLH1*.

Mutant	Primer name	Sequence
E37K	Mlh1mut109Sca-for	***t***AC***t***AGTATTCAAGTGATTG
	Mlh1mut109Sca-rev	***ct***TTTTGCATCTAAACAGTTCT**T**AATCATC
N38H	Mlh1mut112XBA-for	G**C**ACTGT***c***TAGATGCAAAATCC
	Mlh1mut112XBA-rev	TCAATCATCTCTTTGATAGC
N38K	Mlh1mut114Bsm1-for	CTATCAAAGAGATGATTGAGAA**G**TGTTTAG
	Mlh1mut114Bsm1-rev	CATT***c***GCTGGCCGCTGGATAAC
S44F	Mlh1mut131Spe-for	AGTATTCAAGTGATTGTTAAAGAGG
	Mlh1mut131Spe-rev	***a***GTG**A**ATTTTGCATCTAAACAG
G67R	Mlh1mut199Kpn-for	***t***ACC**A**GGATCAGGAAAGAAG
	Mlh1mut199Kpn-rev	CCATTGTCTTGGATCTGAATC
H109P	Mlh1mut326Stu-for	GCC**C**TGTGGCTCATGTTACTATTA
	Mlh1mut326Stu-rev	TTATGCTGGCCAA***g***GCCTCA
A111P	Mlh1mut331Stu-for	GCCATGTG**C**CTCATGTTACT
	Mlh1mut331Stu-rev	TTATGCTGGCCAA***g***GCCTCA
R265S	Mlh1mut793SAL rev	C**T**ATGGTTGATGAAGAGTAAGAAGATGCAC
	Mlh1mut793SAL for	TCTGGTAGA***g***TC***g***ACTTCCTTGAGAAAAGC
R265C	Mlh1mut793CysSap1-for	TTCATCAACCA**T**TGTCTGGT
	Mlh1mut793CysSap1-rev	GAG***c***AAGAAGATGCACTTCT
P403S	Mlh1mut1207Hind-for	C**T**TGTCCAGTCAGCCCCAGG
	Mlh1mut1207Hind-rev	***ct***TTTGCTCAGAGGCTGCAG
L582F	Mlh1mut1744Nhe-for	CT***a***GCCTTAGATAGTCCAGAG
	Mlh1mut1744Nhe-rev	CATGGCAAGGTCAAAGA**A**CGGTG
K618A	Mlh1mut1852EcoR1-for	TT***c***CTGAAGAAG**GC**GGCTGAG
	Mlh1mut1852EcoR1-rev	***t***TCAACAATGTATTCAGCAAG
R659P	Mlh1mut1976Hpa1-for	***t***AACTGGGACGAAGAAAAGG
	Mlh1mut1976Hpa1-rev	ACCTCAGTGGCTAGT**G**GAAG

Amino acid numbering is based on the cDNA with +1 corresponding to the translation initiation codon.
Bold, underlined font: mutant nucleotide.
Small, italic font: silent mutations introducing a restriction site for identification of the linked nucleotide substitution
All mutant plasmids used in this study, but not listed in this table have been gently provided from Professor L. J. Rasmussen, Roskilde University, Roskilde, Denmark.

Inactivation of DNA Mismatch Repair by Variants of Uncertain Significance in the *PMS2* Gene

Mark Drost, Hester Koppejan and Niels de Wind

Adapted from Drost et. al 2013 Hum Mutat 34:1477-1480

5

ABSTRACT

Lynch syndrome (LS) is a common cancer predisposition caused by an inactivating mutation in one of four DNA mismatch repair (MMR) genes. Frequently a variant of uncertain significance (VUS), rather than an obviously pathogenic mutation, is identified in one of these genes. The inability to define pathogenicity of such variants precludes targeted healthcare. Here, we have modified a cell-free assay to test VUS in the MMR gene *PMS2* for functional activity We have analyzed nearly all VUS in *PMS2* found thus far and describe loss of MMR activity for five, suggesting the applicability of the assay for diagnosis of LS.

LS predisposes to various cancers, most notably to colon and ovarian cancer [de la Chapelle, 2004]. LS is caused by heterozygous inactivating mutations in one of the MMR genes *MSH2* (MIM# 609309), *MSH6* (MIM# 600678), *MLH1* (MIM# 120436) or *PMS2* (MIM# 600259) [de la Chapelle, 2004]. Inadvertent loss of the second, wild type, allele in somatic cells results in MMR deficiency, which underlies the accumulation of spontaneous genomic mutations and the rapid development of cancer [de la Chapelle, 2004]. Confirmed LS patients enroll in lifelong preventive surveillance programs and may benefit from personalized chemoprevention and chemotherapy [Hewish et al., 2010; Burn et al., 2011]. Unfortunately, genetic diagnosis of LS patients is complicated by the fact that a significant fraction of all MMR gene alterations found are so-called Variants of Uncertain Significance (VUS) [de la Chapelle, 2004]. In the absence of data on the impact of the VUS on gene function, it is often difficult to interpret their pathogenicity. Moreover, the incidence of VUS is believed to increase steeply with the advent of personalized genomics [Rasmussen et al., 2012]. To enable personalized healthcare for carriers of pathogenic variants and to liberate unaffected relatives from the burden associated with the uncertain pathogenicity of the VUS, it is of great importance to develop procedures to evaluate their pathogenicity [Rasmussen et al., 2012].

Of all MMR genes, VUS in *PMS2* have the highest incidence, comprising ~49% of all alterations described in this gene (http://www.med.mun.ca/MMRvariants/statistics.aspx). We have recently described a cell-free assay to measure the functional activity of VUS in the MMR genes *MSH2*, *MLH1* and *MSH6* [Drost et al., 2010, 2012]. To facilitate the assessment of pathogenicity of VUS in *PMS2* we have modified the cell-free assay to analyze their functional activity. In this assay the mutated cDNA is recreated by PCR, followed by *in vitro* transcription/translation of the variant PMS2 protein and of its wild type heterodimeric partner MLH1 (Figure 1*A*). To serve as template for the generation of the variant *PMS2* alleles we used a wild type PMS2 fused to short S and thrombin tags, as this fusion protein displayed higher *in vitro* expression than the native PMS2 (Supp. Figure S1A, compare the first two lanes), while it did not affect its activity (Supp. Figure S1B). The variant MLH1/PMS2 heterodimer is added to an MLH1/PMS2-deficient cell extract and tested for its ability to restore a *Hin*DIII restriction enzyme recognition site that is disrupted by an embedded G·T mismatch (Figure 1*B*). The inability to restore repair of the mismatch is indicative of the pathogenicity of a *PMS2* VUS. All experimental procedures are described in the Supp. Materials & Methods. The assay appeared relatively insensitive to the amount of PMS2 included in the reaction, contributing to its robustness (Supp. Figure S1). The absolute repair efficiency of tagged wild type PMS2 under the conditions used in this assay is 43.5%±4.2 (mean±S.E.M.) and this is highly reproducible (Supp. Figure 1*B*). As the substrate concentration is in excess, absolute repair efficiencies are not a relevant measure for defects of a VUS. For this reason the *in vitro* MMR assay data is expressed as percentage of repair relative to wild type.

To test the applicability of the assay we have determined repair efficiencies of 27 VUS in *PMS2*, representing the large majority of all VUS registered in the Leiden Open Variation Database (LOVD; Table 1). Of these, variant E705K served as an MMR-deficient, pathogenic, control [Deschênes et al., 2007; van Oers et al., 2010]. Additionally, variants

E541K and G857A were included as innocuous polymorphisms, as judged from their high allele frequencies (dbSNP rs2228006 and rs1802683, respectively). *PMS2* variants registered in the LOVD as *PMS2* pseudogene-derived were excluded from analysis. All alleles were recreated by PCR and protein was produced *in vitro* (Figure 1*C*).

As the polymorphic *PMS2* alleles enabled repair activities significantly higher than the known pathogenic control E705K, the assay has sufficient resolution to distinguish repair-proficient from repair-deficient VUS (Figure 1*D*). Variants with repair efficiencies not significantly higher than the pathogenic control (E41A, S46I, S46N and C843Y) were

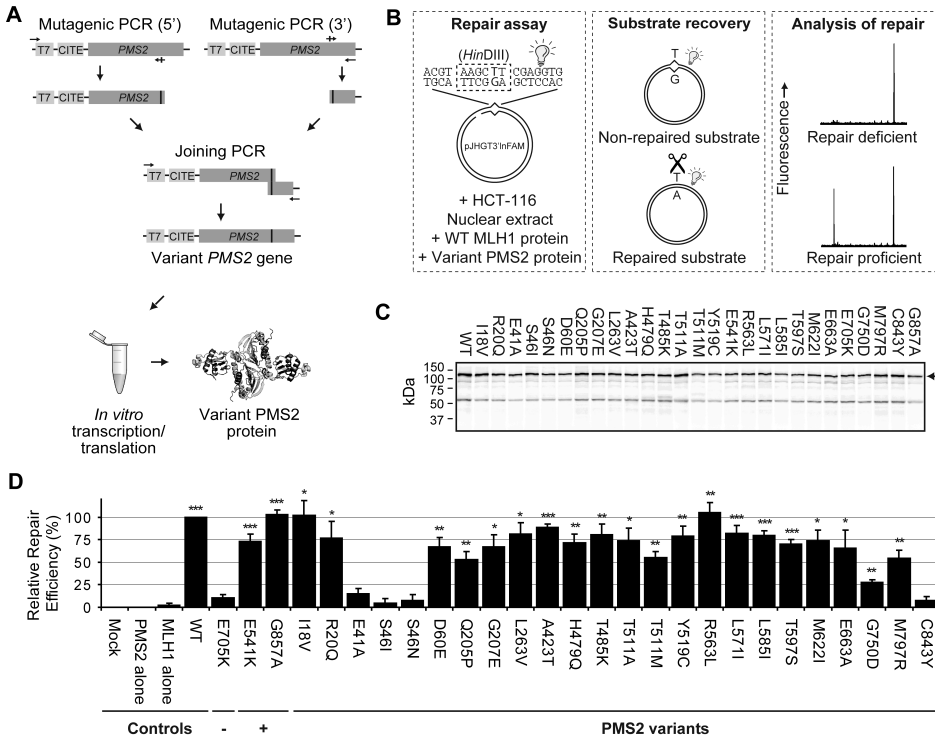


Figure 1. *Mismatch repair activity of PMS2 VUS. (A) Production of variant PMS2 alleles and proteins. All alleles, including template vector-derived T7 promoter and CITE sequences that are required for efficient transcription/ translation in vitro, are generated by two sequential site-specific mutagenic PCR reactions. Variant PMS2 alleles are then used as a template in an in vitro transcription/translation reaction to produce variant PMS2 proteins. (B) Flow scheme of the cell-free assay. Left: Fluorescently labelled (light bulb) substrate pJHGT3'lnFAM is incubated in HCT-116 nuclear extract and in vitro produced heterodimeric variant PMS2/wild type MLH1 protein. Middle: After incubation, the substrate is purified and digested. Right: Repair products are visualized by automated fragment analysis and quantified. (C) Representative expression of 35^S-Methionine-labeled variant PMS2 proteins, visualized after SDS-PAGE gel electrophoresis and autoradiography. Arrow: full-length PMS2 variants. (D)Relative repair efficiencies for PMS2 VUS. (–): Repair-deficient control, (+): Repair-proficient controls. Results are shown as mean±S.E.M. of 3–4 independent experiments for all VUS and >6 experiments for controls. Mock: Mock expression. Asterisks: Significantly higher than repair-deficient control E705K. * p<0.05; ** p<0.01; *** p<0.001 (Student's one-tailed t-test). For the "Mock" and "PMS2 only" reactions, no repair was detected in any of the experiments.*

Table 1. Pathology data and references for *PMS2* VUS tested in this work.

Mutation			Pathology[c]			
				IHC		
Protein[a]	DNA[b]	*In vitro* MMR (This work)	MSI	PMS2	MLH1	References[d]
I18V	c.52A>G	+				1, 2
R20Q	c.59G>A	+	H[3]			1, 2, 3, 4, 5
E41A	c.122A>C	-				5
S46I	c.137G>T	-	H[3, 6]	Neg[3, 6]	Pos[6]	2, 3, 4, 6, 7, 8
S46N	c.137G>A	-				7, 8
D60E	c.180C>G	+				LOVD
Q205P	c.614A>C	+				LOVD
G207E	c.620G>A	+	H[9]	Neg[9]		9
L263V	c.787C>G	+				LOVD
A423T	c.1267G>A	+				10
H479Q	c.1437C>G	+				11, 12
T485K	c.1454C>A	+	H[6]	Neg[6]	Pos[6]	1, 2, 4, 6, 11
T511A	c.1531A>G	+	H[3]			2, 3, 6, 11, 12
T511M	c.1532C>T	+				LOVD
Y519C	c.1556A>G	+				LOVD
E541K	c.1621A>G	+				1, 2, 5, 6
R563L	c.1688G>T	+				1, 2, 8
L571I	c.1711C>A	+				2
L585I	c.1753C>A	+				1
T597S	c.1789A>T	+	L[3]			1, 3, 4, 5, 11
M622I	c.1866G>A	+				1, 2, 4, 6, 8, 11
E663A	c.1988A>C	+				8
E705K	c.2113G>A	-				2, 6, 8, 13, 14
G750D	c.2249G>A	+				8
M797R	c.2390T>G	+				2
C843Y	c.2528G>A	-				8
G857A	c.2570G>C	+				1, 2, 6, 12

[a] Amino acid numbering is based on the PMS2 reference sequence NP_000526.1 with +1 corresponding to the translation initiation amino acid.

[b] Nucleotide numbering reflects cDNA numbering with +1 corresponding to the A of the translation initiation codon of the *PMS2* GenBank reference sequence (NM_000535.5).

[c] MSI=Microsatellite Instability. H=High, L=Low, IHC=Immunohistochemistry. Neg = Negative for staining, Pos = Positive for staining. In case cells are empty, the variant was not tested.

[d] Variants have been selected from the LOVD (http://chromium.liacs.nl/LOVD2/colon_cancer/variants.php?select_db=PMS2&action=view_all&view=Prot_sub). Appropriate references are shown. References: 1 [Hendriks et al., 2006], 2 [Clendenning et al., 2006], 3 [Pastrello et al., 2011], 4 [Thompson et al., 2012], 5 [Borràs et al., 2013], 6 [Nakagawa, 2004], 7 [Jackson et al., 2008], 8 [Senter et al., 2008], 9 [Montazer Haghighi et al., 2009], 10 [Ganster et al., 2010], 11 [Wang et al., 1999], 12 [Basil et al., 1999], 13 [Deschênes et al., 2007], 14 [van Oers et al., 2010].

considered repair deficient and therefore presumably are pathogenic (Figure 1*D*). We conclude that this assay effectively identifies repair-deficient PMS2 variants. Variants such as Q205P, T511M and G750D that display repair efficiencies significantly higher than E705K but are compromised compared to wild type were not classified as repair deficient and we surmise that these variants might be pathogenic with reduced penetrance. Extensive calibration of the assay with clinical data is required to assess pathogenicity of alleles displaying such intermediate repair efficiencies.

For some of these VUS microsatellite instability (MSI), a hallmark of MMR deficient cancer, was previously investigated. Indeed, the MSI of a tumor carrying the S46I allele corresponds with its deficiency in MMR (Table 1). However, the MSI in tumors of carriers of the R20Q, G207E, T485K or T511A alleles is in apparent contrast with their normal MMR activity *in vitro* (Table 1). Possibly these patients carry another, yet unidentified MMR gene defect. Contrariwise, we cannot exclude that these alleles cause a defect that is only apparent *in vivo*. For this reason, all variants that are repair proficient in this assay cannot be classified as neutral, but may require additional analyses such as splicing assays, protein stability assays or nuclear localization assays [Rasmussen et al., 2012]. Ultimately, after the calibration of the assay, it may become a part of an integrated Bayesian analysis that determines pathogenicity of MMR gene VUS, as proposed by us and others [Goldgar et al., 2008; Rasmussen et al., 2012].

ACKNOWLEDGEMENTS

We thank Jacob G. Jansen for his comments on the manuscript. The authors declare no conflict of interest.

GRANT SUPPORT

This work was supported by NIH grant 1R01CA164944-01A1.

REFERENCES

Basil JB, Swisher EM, Herzog TJ, Rader JS, Elbendary A, Mutch DG, Goodfellow PJ. 1999. Mutational analysis of the PMS2 gene in sporadic endometrial cancers with microsatellite instability. Gynecol Oncol 74: 395-399.

Borràs E, Pineda M, Cadiñanos J, Del Valle J, Brieger A, Hinrichsen I, Cabanillas R, Navarro M, Brunet J, Sanjuan X, Musulen E, van der Klift H et al. 2013. Refining the role of pms2 in Lynch syndrome: germline mutational analysis improved by comprehensive assessment of variants. J Med Genet. 50: 552-63.

Burn J, Gerdes AM, Macrae F, Mecklin JP, Moeslein G, Olschwang S, Eccles D, Evans DG, Maher ER, Bertario L, Bisgaard ML, Dunlop MG et al. 2011. Long-term effect of aspirin on cancer risk in carriers of hereditary colorectal cancer: an analysis from the CAPP2 randomised controlled trial. Lancet 378: 2081-2087.

Clendenning M, Hampel H, LaJeunesse J, Lindblom A, Lockman J, Nilbert M, Senter L, Sotamaa K, de la Chapelle A. 2006. Long-range PCR facilitates the identification of PMS2-specific mutations. Hum Mutat 27: 490-495.

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

Deschênes SM, Tomer G, Nguyen M, Erdeniz N, Juba NC, Sepúlveda N, Pisani JE, Liskay RM. 2007. The E705K mutation in hPMS2 exerts recessive, not dominant, effects on mismatch repair. Cancer Lett 249: 148-156.

Drost M, Zonneveld JB, Van Dijk L, Morreau H, Tops CM, Vasen HFA, Wijnen J, De Wind N. 2010. A cell-free assay for the functional analysis of variants of the mismatch repair protein MLH1. Hum Mutat 31: 247-253.

Drost M, Zonneveld JB, van Hees S, Rasmussen LJ, Hofstra RM, De Wind N. 2012. A rapid and cell-free assay to test the activity of lynch syndrome-associated MSH2 and MSH6 missense variants. Hum Mutat 33: 488-494.

Ganster C, Wernstedt A, Kehrer-Sawatzki H, Messiaen L, Schmidt K, Rahner N, Heinimann K, Fonatsch C, Zschocke J, Wimmer K. 2010. Functional PMS2 hybrid alleles containing a pseudogene-specific missense variant trace back to a single ancient intrachromosomal recombination event. Hum Mutat 31: 552-560.

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

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

Hewish M, Lord CJ, Martin SA, Cunningham D, Ashworth A. 2010. Mismatch repair deficient colorectal cancer in the era of personalized treatment. Nat Rev Clin Oncol 7: 197-208.

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

Montazer Haghighi M, Radpour R, Aghajani K, Zali N, Molaei M, Zali M. 2009. Four novel germline mutations in the MLH1 and PMS2 mismatch repair genes in patients with hereditary nonpolyposis colorectal cancer. Int J Colorectal Dis 24: 885-893.

Nakagawa H. 2004. Mismatch Repair Gene PMS2: Disease-Causing Germline Mutations Are Frequent in Patients Whose Tumors Stain Negative for PMS2 Protein, but Paralogous Genes Obscure Mutation Detection and Interpretation. Cancer Res 64: 4721-4727.

Pastrello C, Pin E, Marroni F, Bedin C, Fornasarig M, Tibiletti M, Oliani C, De Leon M, Urso E, Puppa L, Agostini M, Viel A. 2011. Integrated analysis of unclassified variants in mismatch repair genes. Genet Med 13: 115-124.

Rasmussen LJ, Heinen CD, Royer-Pokora B, Drost M, Tavtigian S, Hofstra RM, De Wind N. 2012. Pathological assessment of mismatch repair gene variants in Lynch syndrome: Past, present, and future. Hum Mutat 33: 1617-1625.

Senter L, Clendenning M, Sotamaa K, Hampel H, Green J, Potter JD, Lindblom A, Lagerstedt K, Thibodeau SN, Lindor N, Young J, Winship I et al. 2008. The Clinical Phenotype of Lynch Syndrome Due to Germ-Line PMS2 Mutations. Gastroenterology 135: 419-428

Thompson B, Greenblatt M, Vallee M, Herkert J, Tessereau C, Young E, Adzhubey I, Li B, Bell R, Feng B, Mooney S, Radivojac P et al. 2012. Calibration of Multiple In Silico Tools for Predicting Pathogenicity of Mismatch Repair Gene Missense Substitutions. Hum Mutat 34: 255-265.

van Oers JM, Roa S, Werling U, Liu Y, Genschel J, Hou H, Sellers RS, Modrich P, Scharff MD, Edelmann W. 2010. PMS2 endonuclease activity has distinct biological functions and is essential for genome maintenance. Proc Natl Acad Sci USA 107: 13384-13389.

Wang Q, Lasset C, Desseigne F, Saurin JC, Maugard C, Navarro C, Ruano E, Descos L, Trillet-Lenoir V, Bosset JF, Puisieux A. 1999. Prevalence of germline mutations of hMLH1, hMSH2, hPMS1, hPMS2, and hMSH6 genes in 75 French kindreds with nonpolyposis colorectal cancer. Hum Genet 105: 79-85.

SUPPLEMENTARY MATERIALS

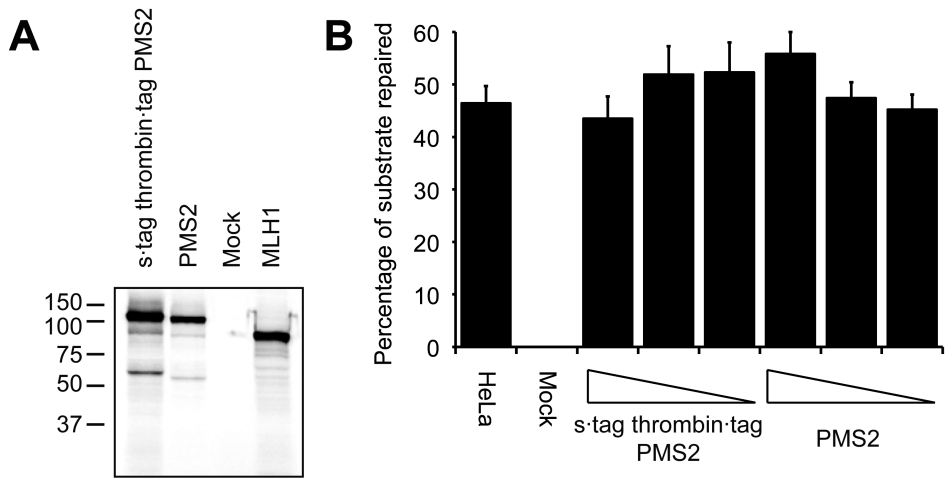
MATERIALS & METHODS

Cloning of PMS2. The human *PMS2* (GenBank NM_000535.5) cDNA was cloned into the pCITE4a expression vector (Novagen, Beeston, UK). *PMS2* cDNA was PCR-amplified using forward primer 5'-CGCGGATCCGCCACCATGGAGCGAGCTGAGAGCTCGAGTAC-3' and reverse primer 5'-GCGCGGCCGCGGTCAGTTCTGAGAAATGACAC-3'. The PCR fragment was cloned into pCITE4a after *Bam*HI and *Not*I (New England Bioland, Ipswich, MA, USA) digestion. The encoded PMS2 protein carries a 40 amino acid N-terminal tail, containing both an s tag and a thrombin tag. This tail increases protein expression without affecting protein activity (Supplementary Figure 1).

Generation of variant PMS2 proteins by PCR and In Vitro Expression. The plasmid construct described above was used as a template in a PCR procedure to recreate variant PMS2 alleles, including the vector-derived T7 promoter and CITE sequences that are required for efficient transcription/translation *in vitro*. PCR reactions were performed as described (Figure 1*A*) [Drost et al., 2012]. In brief, overlapping 5' and 3' PCR fragments were generated in separate, PCR reactions. 5' PCR fragments were generated using a common pCITE forward primer and a reverse primer specific for the VUS. 3' PCR fragments were produced using a forward primer complementary to the reverse primer in the 5' PCR, specific to the VUS, and a common pCITE reverse primer. In a subsequent PCR reaction, the two mutant fragments were joined and reamplified using nested primers. Oligonucleotide sequences are available upon request and were ordered from Biolegio (Nijmegen, The Netherlands). All PCR reactions were performed in a total volume of 10 µl, containing 0.4 U of Pfx Platinum Polymerase (Invitrogen, Carlsbad, CA). Proper introduction of the variant codon was confirmed by direct sequencing of PCR fragments.

Following amplification, all final PCR products were purified using the PCR Purification Kit (Qiagen). These purified fragments were then used for *in vitro* protein expression in the TnT Quick Coupled in vitro Transcription/Translation System (Promega, Madison, WI), in the presence of PCR Enhancer with minor modifications to the manufacturer's protocol. Expression levels were verified by producing parallel ^{35}S-Methionine-labeled reactions according to the manufacturer's instructions. Labeled proteins were analyzed by SDS-polyacrylamide gel electrophoresis and Phosphorimaging (Perkin Elmer, Waltham). Wild type MLH1, the heterodimeric partner of PMS2, was produced by large-scale *in vitro* expression from its pCITE4a clone [Drost et al., 2010]. The PMS2/MLH1 heterodimers were generated by allowing *in vitro*-expressed wild type or variant PMS2 and WT MLH1 (1:1 v/v) to dimerize for 30 minutes at room temperature.

In Vitro MMR Assays. *In vitro* MMR assays were performed as described (Figure 1B) [Drost et al., 2010, 2012]. Reactions are carried out in 25µl containing 75µg of nuclear extract. Nuclear extract was prepared from HCT-116 colon cancer cells that lack both PMS2 and MLH1, as described [Holmes et al., 1990]. Extracts were complemented with 12µl of dimerized PMS2/MLH1 and 100ng of fluorescent substrate pJHGT3'lnFAM. After



Supplementary Figure 1. *The 40 amino acid N-terminal tag on PMS2 does not influence MMR activity. (A) Representative expression of ^{35}S-Methionine-labeled proteins, visualized after SDS-PAGE gel electrophoresis and autoradiography. (B) Repair efficiencies of in vitro produced PMS2 with or without fused N-terminal tags. PMS2 proteins were sequentially diluted twofold and dimerized to MLH1. Repair assays were carried out as described. MMR-proficient HeLa nuclear extract serves as a positive control. No significant differences were found between any of the PMS2 measurements, supporting the robustness of the assay. Results are shown as mean±S.E.M. of 4 independent experiments.*

40 min of incubation at 37°C, substrate DNA was purified using the MinElute kit (Qiagen, Germantown, MD). Then, the substrate was digested with *Hin*DIII and *Bsr*BI (Fermentas, Burlington, ON, Canada). One-fifth (2 µl) of digested substrate was mixed with 7.8 µl Hi-Di Formamide and 0.2 µl GeneScan-500 ROX size standard (Applied Biosystems, Bedford, MA) and fluorescent fragment analysis was performed under standard conditions. Signal was quantified using GeneMarker software (Softgenetics, State College, PA). Repair levels were calculated by dividing the height of the MMR-specific peak by the total fluorescent signal, and then normalized to wild type levels.

SUPPLEMENTARY REFERENCES

Drost M, Zonneveld JB, Van Dijk L, Morreau H, Tops CM, Vasen HFA, Wijnen J, De Wind N. 2010. A cell-free assay for the functional analysis of variants of the mismatch repair protein MLH1. Hum Mutat 31: 247-253.

Drost M, Zonneveld JB, van Hees S, Rasmussen LJ, Hofstra RM, De Wind N. 2012. A rapid and cell-free assay to test the activity of lynch syndrome-associated MSH2 and MSH6 missense variants. Hum Mutat 33: 488-494.

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

Genetic Screens to Identify Pathogenic Mismatch Repair Gene Variants in the Common Cancer Predisposition Lynch Syndrome

6

Mark Drost, Anne Lützen, Sandrine van Hees, Daniel Ferreira, Fabienne Calléja, José B.M. Zonneveld, Finn Cilius Nielsen, Lene Juel Rasmussen and Niels de Wind

Adapted from Drost et. al Proc Natl Acad Sci USA. 2013 23:9403-9408

ABSTRACT

In many individuals suspected of the common cancer predisposition Lynch syndrome a variant of unclear significance (VUS), rather than an obviously pathogenic mutation, is identified in one of the DNA mismatch repair (MMR) genes. The uncertainty whether such VUS inactivate MMR, and therefore are pathogenic, precludes targeted healthcare for both carriers and their relatives. To facilitate the identification of pathogenic VUS we have developed an *in cellulo* genetic screen-based procedure for the large-scale mutagenization, identification and cataloging of residues of MMR genes, critical for DNA mismatch repair gene function *in vivo*. When a residue, identified as mutated in an individual suspected of Lynch syndrome, is listed as critical in such a 'reverse diagnosis catalog' there is a high probability that the human VUS is pathogenic. To investigate the applicability of this approach we have generated and validated a prototypic reverse diagnosis catalog for the MMR gene *Msh2* by mutagenizing, identifying and cataloging 26 deleterious mutations in 23 amino acids. Extensive *in vivo* and *in vitro* analysis of mutants listed in the catalog revealed both recessive and dominant-negative phenotypes. Nearly half of these critical residues match with VUS, previously identified in individuals suspected of Lynch syndrome. This aids in the assignment of pathogenicity to these human VUS and validates the approach described here as a diagnostic tool. In a wider perspective, this work provides a model for the translation of personalized genomics into targeted healthcare.

INTRODUCTION

Lynch syndrome is a prevalent, dominantly inherited, predisposition to colon and other visceral cancers, caused by heterozygosity for a deleterious defect in one of the DNA mismatch repair (MMR) genes *MSH2*, *MSH6*, *MLH1* or *PMS2*. Inadvertent loss of the second, wild type, allele in somatic cells results in MMR deficiency. The consequent inability to remove accidental nucleotide misincorporations during DNA replication leads to the accumulation of spontaneous genomic mutations and the rapid development of cancer [Lynch and de la Chapelle, 2003].

Lifelong preventive surveillance is an effective means to reduce cancer incidence and mortality in carriers of a pathogenic MMR gene mutation [Lagerstedt Robinson et al., 2007] . Additionally, carriers may benefit from both targeted chemoprevention [Burn et al., 2011] and chemotherapy [Hewish et al., 2010]. Unfortunately, in a significant fraction of all individuals suspected of Lynch Syndrome a subtle alteration of one of the MMR genes is identified, such as a single amino acid alteration or a mutation of a putative splice site [Peltomäki and Vasen, 2004]. The pathogenicity of these so-called Variants of Unclear Significance (VUS) often cannot be assessed, due to the absence of conclusive clinical and segregation data, and of data on the impact of the mutation on gene function. It is furthermore believed that primary screening for MMR gene mutations in selected populations, a cost-effective approach to improve general health, will cause a further increase in the incidence of MMR gene VUS [Barnetson et al., 2008; Dinh et al., 2011; Heinen et al., 2010]. The development and use of robust and accessible approaches to identify pathogenic VUS will enable the implementation of targeted preventive and curative healthcare for carriers of pathogenic MMR gene VUS. Meanwhile, unaffected relatives of such carriers can be liberated from the burden associated with the uncertain pathogenicity of the VUS while unwarranted inflow into preventive healthcare systems is reduced [Castells et al., 2009].

The genetics and biochemistry of the MMR pathway have been established in detail [Hsieh et al., 2008], which has enabled the development of functional assays to predict pathogenicity of MMR gene VUS [Rasmussen et al., 2012]. Although these assays can be useful tools in the prediction of pathogenicity of VUS they are often laborious and time consuming, whereas defects that specifically affect functions of the variant genes *in vivo*, such as splicing, intracellular localization or stability, may escape detection. To circumvent these issues we have taken an *in vivo*, genetic screen-based, approach. This approach comprises the large-scale mutagenesis and identification of critical amino acids of MMR genes, the analysis of the mutant cell lines and proteins, and the compilation of these data in an annotated 'reverse diagnosis catalog'. Such a catalog enables to rapidly see whether the VUS is at a residue that is critical for gene function *in vivo*, and therefore aids in the assessment of pathogenicity. Here, we describe the construction and extensive validation of a prototypic reverse diagnosis catalog of the core MMR protein Msh2.

RESULTS

A genetic screen to mutagenize and map critical MMR gene residues

We argued that residues, critical for MMR gene function *in vivo*, could be identified by a three-step procedure: (i) Treatment of cultured cells using a wide-spectrum mutagen to randomly induce genetic variants, (ii) Isolation of clones that have lost MMR consequent to mutation of a critical amino acid, (iii) Identification of the causal mutations and their compilation in a reverse diagnosis catalog. Mouse embryonic stem (ES) cells is a suited cell type for such a screen as ES cells are primary diploid cells, though immortal. Furthermore, ES cell lines monoallelic for each of the MMR genes are available [de Wind et al., 1995; de Wind et al., 1999], which requires the inactivation of only the single, wildtype, allele by a mutagenic event. Importantly, mouse and human MMR genes are >95% identical or homologous at the amino acid level.

To generate a pilot reverse diagnosis catalog we used an ES cell line, monoallelic for *Msh2* [$Msh2^{+/-}$; de Wind et al., 1995]. This line was treated with the powerful, wide-spectrum, and relatively non-toxic point mutagen N-ethyl-N-nitrosourea [ENU; Fig. 1A; Borgdorff et al., 2005; Russel et al., 1979]. To isolate cell clones that have become MMR-deficient resulting from a mutagenic event at the monoallelic MMR gene we subsequently exposed cells to a low concentration of the nucleotide analog 6-thioguanine (6TG; Fig. 1B), to which MMR-deficient cells are tolerant [Jansen et al., 1995; Glaab et al., 1998]. Then, unwanted clones that had acquired 6TG tolerance due to mutational inactivation of the X-linked *Hprt* gene, rather than by loss of MMR, were killed by culture in Hypoxanthine-Aminopterine-Thymidine-supplement-containing medium (Fig. 1C). In the remaining clones, the single *Msh2* allele might be lost by either an ENU-induced substitution at a critical residue ($Msh2^{M/-}$ in Fig. 1D), or by loss of heterozygosity [$Msh2^{-/-}$ in Fig. 1D; Borgdorff et al., 2005]. The latter clones were discarded after screening with an *Msh2* allele-specific PCR. The remaining cell clones are expected to harbor an ENU-induced substitution that inactivates a critical residue of the monoallelic *Msh2* gene. Importantly, when wild type ($Msh2^{+/+}$) ES cells were subjected to these treatments no surviving clones were obtained, indicating that this procedure does not result in clones with a simultaneous mutational inactivation of both alleles of any of the MMR genes.

To identify ENU-induced amino acid substitutions in Msh2 in the selected clones we sequenced the complete *Msh2* open reading frame of clones that had retained expression of full-length *Msh2* cDNA, as judged by reverse-transcriptase PCR (Table 1). This resulted in the identification of an amino acid substitution in monoallelic Msh2 in 41 clones. These substitutions alter 23 different amino acids, of which 3 amino acids were represented by two different substitutions in individual clones. These residues were then compiled in a prototypic reverse diagnosis catalog (Tables 2 and S1). Mapping of these amino acids to the MSH2 crystal structure revealed a clustering around the lever, connector and ATPase domains of the protein indicating the relative importance of these domains [Fig. 2A, 2B and S3; Warren et al., 2007]. Conversely, few deleterious mutations were found in the N-terminal region of the Msh2 protein (Fig. 2A). MAPP-MMR [Chao et al., 2008], an alignment-based algorithm, predicted 25 of 26 Msh2 mutants in the reverse diagnosis

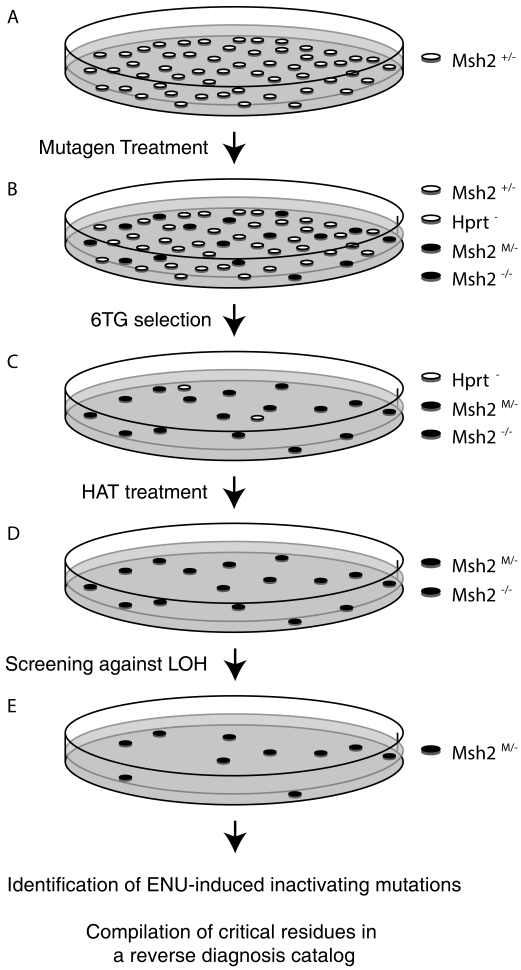


Figure 1. *A genetic screen to identify critical MMR gene residues. (A) An ES cell line, hemizygous for Msh2 (Msh2$^{+/-}$) is subjected to mutagenic treatment with ENU. Preexisting, inadvertent, Msh2$^{-/-}$ cells [Borgdorff et al., 2005] are not displayed here. (B) Cells that have acquired low-level 6TG tolerance, either by loss of heterozygosity at Msh2 (LOH, Msh2$^{-/-}$), by an ENU-induced loss-of-function mutation at a critical residue in the monollelic Msh2 gene (Msh2$^{M/-}$), or by a deleterious mutation at the monoallelic Hprt gene are selected using two brief 6TG selections. (C) Unwanted Hprt-deficient clones are eliminated using HAT-supplemented medium. (D) Unwanted clones that have lost the wild type Msh2 allele by loss of heterozygosity are excluded using an allele-specific PCR. (E) The inactivating mutation in the remaining clones is identified by sequence analysis. A reverse diagnosis catalog is compiled that lists critical residues at Msh2.*

catalog to be deleterious and one to be borderline deleterious, a result that cross-validates both MAPP-MMR and our reverse diagnosis screen (Table 2).

To investigate the applicability of our approach to aid in the identification of pathogenic human MMR gene variants we then aligned the database that lists MSH2 variants, identified in the human population (The Leiden Open Variation Database) with the reverse

Table 1. Generation of the pilot Msh2 Reverse Diagnosis Catalog described here.

Total no. of 6TG and HAT-tolerant clones	1483
Clones that did not display LOH at *Msh2*[1]	426
Clones analyzed for *Msh2* mutations[2]	81
Clones with a single, unique, amino acid substitution[3]	41 (26)
Mutated amino acids	23
Other mutations identified in *Msh2*[4]	10

[1] LOH was assessed by allele-specific PCR.
[2] Only full-length cDNAs were sequenced.
[3] Six substitutions were found in more than one individual clone. Between parentheses: unique substitutions.
[4] Comprises nonsense, frameshift or silent mutations, tandem mutations, or small in-frame deletions.

diagnosis catalog. This revealed that, of the 23 residues listed as critical in the catalog, 10 were previously identified as mutated in humans. Moreover, in 6 out of these 10 cases, the ENU-induced deleterious amino acid substitution was identical to a human VUS (Table 2). This result provides strong support for the applicability of reverse diagnosis catalogs to identify pathogenic human MMR gene variants.

Msh2-mutant cell lines display phenotypic hallmarks of MMR deficiency

To validate the genetic screen experimentally we assembled a validation panel consisting of 13 of the 26 mutant cell lines. These substitutions were chosen because all of these residues are represented in the LOVD, either as identical or different substitution, or as silent mutation (Table 2).

Microsatellite instability is a diagnostic hallmark of cancer in Lynch syndrome [Lynch and de la Chapelle, 2003]. We measured the size of six different microsatellites in 40-50 subclones of each line from the validation panel. Microsatellites were shortened in the large majority of the subclones of each cell line, in support of MMR deficiency (Fig. 3A). The microsatellite size reductions were less pronounced in the cell lines than in Lynch Syndrome-associated cancers [Lynch and de la Chapelle, 2003]. This likely is a consequence of the absence of iterative cycles of expansion and clonal selection, that characterizes carcinogenesis. In addition to microsatellite instability, MMR defects are characterized by the accumulation of spontaneous nucleotide substitutions at protein-encoding genes. Indeed, all cell lines from the validation panel displayed high frequencies of spontaneous *Hprt* mutants, mimicking the *Msh2*-disrupted control cell line (Fig. 3B).

MMR-deficient cells have acquired tolerance of DNA-methylating drugs, a class of agents that includes several chemotherapeutic drugs [Hsieh et al., 2008]. We tested our validation panel of cell lines for tolerance of the prototypic methylating drug N-Methyl-N'-Nitro-N-Nitrosoguanidine (MNNG). Indeed, all cell lines proved as tolerant as the *Msh2*-disrupted control line (Fig. 3C). Cumulatively, these results demonstrate that all Msh2-mutant cell lines from the validation panel have acquired all phenotypic hallmarks of MMR deficiency.

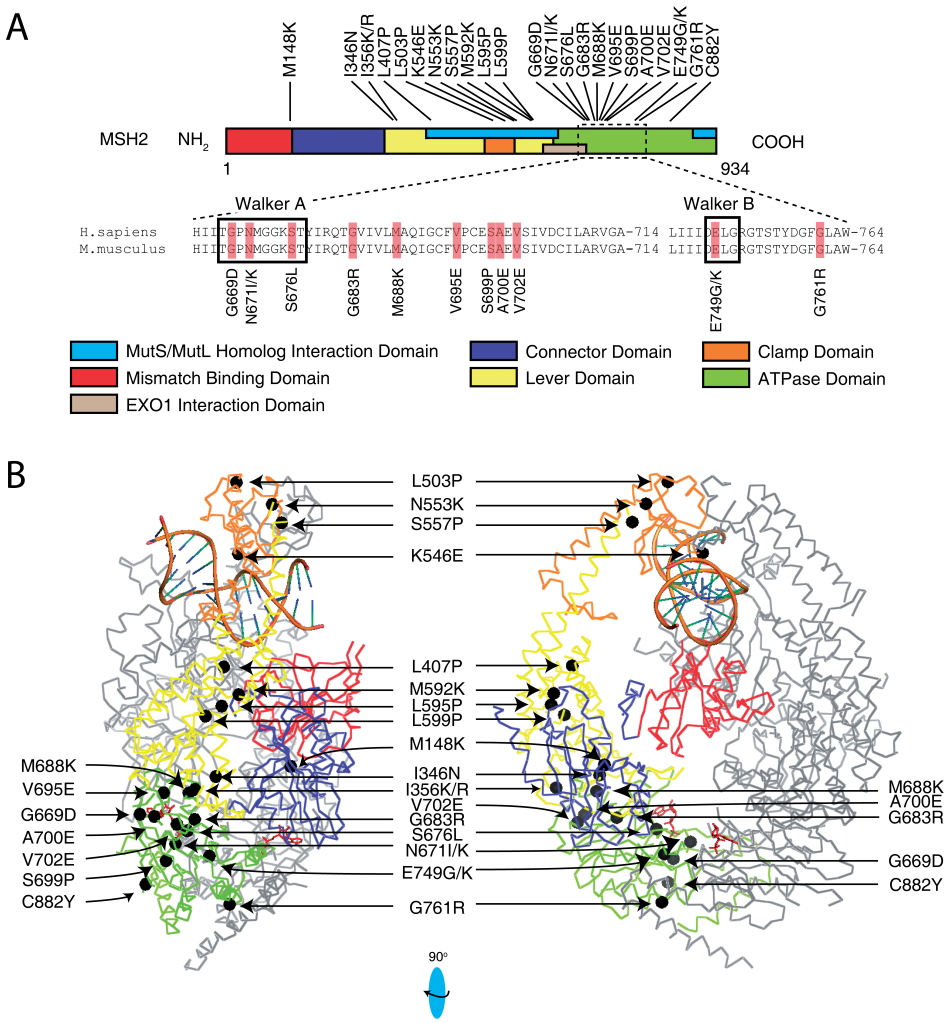


Figure 2. *Critical Msh2 residues mapped in this study. (A) Schematic representation of all identified deleterious amino acid substitutions in Msh2. The colored protein domains are as described [Warren et al., 2007]. It should be noted that the term 'mismatch-binding domain' is a misnomer as the homologous domain in Msh6, rather than Msh2, binds the mismatch. The middle panel of the figure zooms in on an alignment of the Walker A and Walker B ATPase domains of human and mouse Msh2, highlighting residues that were mutated in the screen described here (orange boxes). Numbers reflect amino acid numbering. (B) Mapping of critical Msh2 residues mutated in our screen to the crystal structure of the human MSH2/MSH6 heterodimer. In color: MSH2 (see Fig. 2A for legends), in grey: MSH6. Mutated residues are shown as black spheres. The MutS/MutL and EXO1 interaction domains are not indicated in the crystal structure. Mutants S699P and V695E are obscured in the right panel.*

MMR deficiency is caused by the ENU-induced mutations in Msh2

To provide conclusive evidence of the causality of the Msh2 mutations for MMR deficiency we followed a biochemical approach, employing a cell-free MMR assay that measures

Table 2. Msh2 mutants generated in this study.

Mutation[1]	cDNA		MAPP- MMR[2]	LOVD[3]
M148K	c.443	T>A	3.75	-
I346N	c.1037	T>A	18.3	-
I356K	c.1067	T>A	14.66	-
I356R	c.1067	T>G	16.34	-
L407P	c.1220	T>C	18.56	L407L
L503P	c.1508	T>C	14.91	**L503P**
K546E	c.1636	A>G	23.7	-
N553K	c.1659	C>A	24.75	-
S557P	c.1669	T>C	9.71	T557P
M592K	c.1775	T>A	8.21	M592V
L595P	c.1784	T>C	7.68	-
L599P	c.1796	T>C	17.42	-
G669D	c.2006	G>A	35.86	**G669D/R**
N671I	c.2012	A>T	22.05	**N671K/Y**
N671K	c.2013	T>A	25.31	
S676L	c.2027	C>T	28.64	S676P
G683R	c.2047	G>A	44	**G683R**/W
M688K	c.2063	T>A	32.5	M688R/I
V695E	c.2084	T>A	30.37	-
S699P	c.2095	T>C	5.97	-
A700E	c.2099	C>A	37.91	-
V702E	c.2105	T>A	33.26	V702G
E749K	c.2245	G>A	19.98	**E749K**
E749G	c.2246	A>G	28.23	
G761R	c.2281	G>A	40.89	-
C822Y	c.2465	G>A	6.48	-

[1] In bold: mutants included in the validation panel.
[2] *In silico* MAPP-MMR prediction of pathogenicity. Scores above 4.55 are predicted to be pathogenic, variant M148K is borderline pathogenic (calculated from http://mappmmr.blueankh.com/Impact.php; [Chao et al., 19]).
[3] LOVD: A mutation at the same residue identified in a suspected Lynch syndrome patient, as reported in the Leiden Open Variant Database (LOVD). In bold: identical substitution found in reverse diagnosis catalog and LOVD, allowing to assess pathogenicity to the patient VUS.

the ability of cell extracts to repair a defined G·T mismatch [Drost et al., 2010]. MMR activity was lost in extracts from all lines from the validation panel (Fig. 4A, black bars). To determine causality of the Msh2 mutations for MMR loss we investigated whether addition of recombinant purified wild type human MSH2/MSH6 rescued the MMR defect in extracts from each cell line of the validation panel. Indeed, the defect was significantly complemented

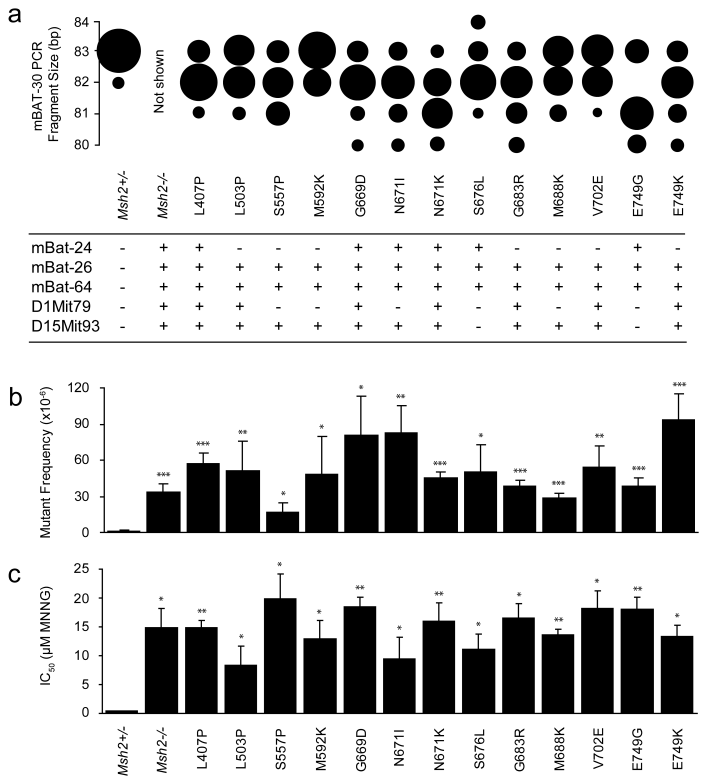


Figure 3. *Analysis of the validation panel. (A) Upper panel: Microsatellite instability analysis on mononucleotide microsatellite mBAT-30. The size of each sphere is proportional to the relative number of subclones with the indicated mBAT-30 PCR fragment length. All mutants from the validation panel display microsatellite instability ($p<0.05$ compared with the $Msh2^{+/-}$ line). Microsatellite size shifts in the $Msh2^{-/-}$ control line are not displayed as in this line the number of cell divisions before subcloning was much higher than in the lines of the validation panel, which makes a quantitative comparison impossible. Lower panel: Microsatellite instability analysis on mononucleotide repeats mBAT-24, mBAT-26 and mBAT-64, and on dinucleotide repeats D1Mit79 and D15Mit93 in the 'patient' ES cell lines. (+): MSI ($p<0.05$ compared with the $Msh2^{+/-}$ line); (-): no MSI. (B) Assessment of spontaneous mutator phenotypes at the genomic Hprt gene in Msh2-mutant ES cell lines from the validation panel. Bars represent Hprt mutant frequencies±S.E.M. *=$p<0.05$, **=$p<0.01$, ***=$p<0.001$ compared with the parental $Msh2^{+/-}$ line. (C) Tolerance to the methylating drug MNNG in Msh2-mutant ES cell lines. Bars represent mean±S.E.M. *=$p<0.05$, **=$p<0.01$, ***=$p<0.001$ compared with the parental $Msh2^{+/-}$ line.*

in 11 cell extracts, confirming that the MMR deficiency is caused by inactivation of Msh2/Msh6, rather than by inadvertent loss of one of the other (biallelic) MMR genes (Fig. 4B, grey bars). Only in two out of the extracts, each derived from a cell line with a different substitution at Msh2 residue N671, the MMR activity was not significantly restored by wild type MSH2/MSH6 (Fig. 4A and see below).

We wanted to directly prove causality for loss of MMR of the mutated Msh2 amino acids. To this aim we used a wild type human MSH2 cDNA as template to mimic the 13 mouse Msh2 mutants from the validation panel by site-directed mutagenic PCR in the

human gene. This was followed by *in vitro* transcription and translation to recreate MSH2 mutant proteins [Drost et al., 2010, 2012]. The resulting proteins were heterodimerized with wild type MSH6 and tested for their ability to restore MMR in extracts from the MSH2/MSH6-deficient human cancer cell line LoVo. None of the 13 mutant MSH2/MSH6 human proteins complemented the LoVo cell extract (Fig. 4B). This result provides direct and conclusive evidence for the causality of each of the ENU-induced amino acid substitutions in Msh2 for MMR-deficiency, also in the human setting. Therefore, this result provides strong support for pathogenicity of the human mutants.

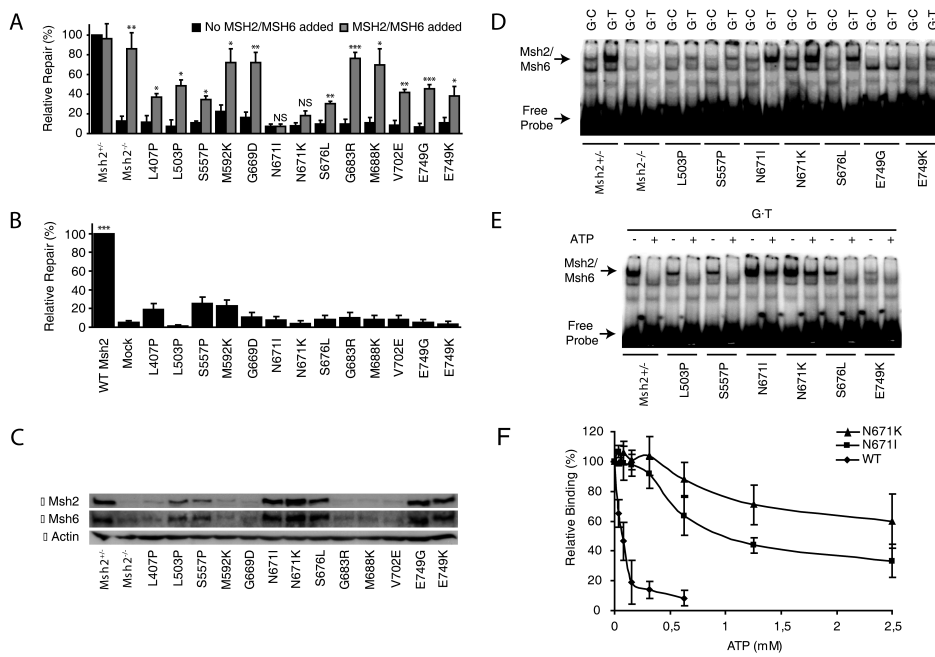


Figure 4. *Biochemical characterization of Msh2-mutant ES cell lines from the validation panel. (A) In vitro repair of a G·T mismatch in mutant cell extracts in the absence (black bars) and presence of purified wild type human MSH2/MSH6 (grey bars). Bars represent mean±SEM. *=p<0.05, **=p<0.01, ***=p<0.001, when comparing repair efficiencies in the presence and absence of recombinant MSH2/MSH6 for each extract. NS=Not significant. (B) MMR activity in in vitro recreated mutant humanized MSH2 proteins. These proteins were tested for the repair of a G·T mismatch after addition to an MSH2/MSH6-deficient human cell extract. Mock: mock Msh2/Msh6 expression. Bars represent mean±S.E.M. *=p<0.05, **=p<0.01, ***=p<0.001 compared to the mock reaction. (C) Western blot analysis of total lysates from mutant ES cell extracts. Actin: β-actin loading control. (D) Binding of control and mutant Msh2/Msh6 cell extracts to a G·T mismatch within a double-stranded oligonucleotide probe in an electrophoretic mobility shift assay. Msh2/Msh6 indicates the Msh2/Msh6·DNA complex. (E) ATP-induced mismatch release of Msh2-mutant cell extracts was tested in an electrophoretic mobility shift assay. ATP (1.25mM) was added after allowing the proteins to bind to the probe. (F) Quantification of ATP-induced mismatch release of Msh2/Msh6, Msh2-N671I/Msh6 or Msh2-N671K/Msh6. The values represent the average from three independent experiments; error bars: S.E.M.*

MMR deficiency is caused by loss of protein expression or by specific biochemical defects

To provide biochemical insights into the MMR defects in the cell lines we investigated levels of mutant Msh2 proteins in the validation panel by Western blotting. This revealed that 6 of the cell lines had lost expression of Msh2 (Fig. 4C), suggesting defective protein folding and/or stability. The heterodimeric partner of Msh2, Msh6, was also destabilized in these 'low protein-level' Msh2 mutants (Fig. 4C), consistent with previous results [de Wind et al., 1999]. By analyzing levels of mutant Msh2 proteins in fractionated cell extracts, we excluded intracellular mislocalization as cause for the MMR defect of the 7 Msh2 mutants that had retained significant protein levels (Fig. S1).

To obtain insights in the biochemical defects of the mutants that had retained protein expression we tested binding to G·T mismatched probes in extracts from these cell lines, using an electrophoretic mobility shift assay. With the exception of mutants N671I and N671K, binding to both the matched and mismatched probes was significantly reduced, suggesting a general defect in DNA binding in those mutants (Fig. 4D).

Substitutions at residue N671 of Msh2 confer a dominant-negative phenotype

After mismatch binding by Msh2/Msh6, binding of ATP to the distal ATPase domains of the protein (Fig. 2) results in the conversion of the mismatch-bound form of Msh2/Msh6 into a repair-competent DNA-embracing clamp that releases the mismatch [Hsieh et al., 2008]. We tested ATP-induced mismatch release in cell extracts from all 6 mutants that had retained low- to normal levels of mismatch binding. ATP released the protein from the mismatch in most of these extracts, which indicates that ATP binding nor intramolecular signaling are affected (Fig. 4E). In contrast, extracts from both mutants at residue N671 displayed a specific defect in the ATP-induced release from the mismatch (Fig. 4E, F). Indeed, N671 is a highly conserved residue at the Walker A (ATP-binding) motif of the ATPase domain [Fig. 2A; Hsieh et al., 2008]. It is not unlikely that the normal mismatch binding of the N671I and N671K mutants, combined with the defect in ATP-induced mismatch release, results in persistent binding of inactive mutant heterodimer at the mismatch, precluding binding of wild type Msh2/Msh6. This result is consistent with the inability of wild type MSH2/MSH6 to restore MMR activity to extracts from both mutant cell lines (Fig. 4A and see above). Complementary to this, we show that the addition of both N671 mutant proteins, but not the recessive G683R mutant protein, to a HeLa cell extract *in vitro* inhibits MMR activity (Fig. S2). Based on these results we infer that the N671I and N671K substitutions confer dominant-negative phenotypes to Msh2.

DISCUSSION

Genetic screens for the identification of pathogenic MMR gene variants

The absence of well-established procedures to classify VUS in genes associated with cancer predisposition is a bottleneck in the translation of personalized genomics into targeted preventive

or curative healthcare. To aid in resolving this problem in Lynch syndrome we have developed a genetic screening procedure to mutagenize and identify critical MMR gene residues. This results in reverse diagnosis catalogs that list critical gene residues, together with the associated mutant cell lines. Given the causal relation of MMR deficiency for Lynch syndrome, a VUS that is listed as deleterious in the reverse diagnosis catalog has a high probability of pathogenicity.

The screens are based on the mutagenic treatment of an ES cell line, monoallelic for one of the MMR genes, followed by selection of clones that have acquired a deleterious substitution of a critical residue of the monoallelic gene. This procedure mimics the somatic inactivation of the wild type allele and the subsequent clonal outgrowth of MMR-deficient tumor cells that precedes cancer development in Lynch syndrome [Liu et al., 1995]. Therefore, the mutant cell lines can be used as models for cancer cell lines from patients carrying the corresponding variant. The generation of reverse diagnosis catalogs to aid in the identification of pathogenic variants in Lynch syndrome is not restricted to *Msh2*, since also cells deficient for the other core MMR genes are tolerant to 6TG [Glaab et al., 1998], and monoallelic ES cell lines for these genes are available. Indeed, we have recently generated a prototypic reverse diagnosis catalog for Msh6 using an identical procedure as described here (in preparation). A comprehensive reverse diagnosis catalog (Tables 2 and S1), linear gene map (Fig. 2A) or structural representation (Fig. 2B and S3), annotated with functional data, has to be generated only once for each MMR gene and will aid to assess pathogenicity to human MMR gene VUS, now and in the future.

In the prototypic screen presented here we have identified 26 cell lines, each carrying a deleterious Msh2 amino acid substitution, in total unveiling 23 amino acids critical for Msh2 function *in vivo.* Ten of these 23 amino acids have previously been identified as mutated in humans (Table 2). In 6 out of these, the amino acid substitution in the individual and in the reverse diagnosis catalog were identical (Table 2). Since these substitutions were proven to be causal for loss of MMR in this work, and since loss of MMR is causal for Lynch syndrome, these data strongly suggest the pathogenicity of these human VUS.

To validate the screen we have extensively tested a panel consisting of 13 of the mutants, including those that were also identified as VUS in humans. All mutant cell lines from the validation panel display all hallmarks of MMR deficiency (Table S1). With the exception of the dominant-negative substitutions at residue N671 (see below) the biochemical MMR defect was rescued by wild type human MSH2/MSH6 in extracts from all mutant mouse ES cell lines. This demonstrates that the mutagenic treatment has not inadvertently inactivated any of the other MMR genes (Fig. 4A). Finally, we have directly proven causality of the ENU-induced Msh2 substitutions in these 13 cell lines for the loss of MMR in a human setting (Fig. 4B). All taken together, these experiments validate the use of the mouse cell-based genetic screen presented here to identify critical MMR gene residues, to use these data for the assignment of pathogenicity to human MMR gene VUS, and to investigate *in vivo* phenotypes of the mutant cell lines as models for tumor cell lines.

Our pilot screen yielded only a single inactivating mutation at the N-terminus of Msh2. This may have different causes, including (i) a bias in the mutagenicity of ENU; (ii) a bias in

the recovery of N-terminal deleterious mutations, e.g. resulting from degradation of mRNA (our mutants were identified by cDNA sequencing) (iii) an intrinsically low density of critical residues in the N-terminal part of Msh2. In our support, deletion of N-terminal 133 amino acids of Msh2 hardly affects MMR in *Saccharomyces cerevisiae*, indicating that this domain is less important for MMR [Lee et al., 2007]. Conversely, the C-terminal ATPase domain of Msh2 is highly conserved and indispensable for gene function, explaining the density of substitutions in our pilot screen [Hsieh et al., 2008]. In the Human Gene Mutation Database (HGMD; https://portal.biobase-international.com), (presumed deleterious) missense variants are ubiquitous in the N-terminal part of MSH2. Possibly, many of these N-terminal substitutions represent hypomorphic mutations whereas our screen selects for complete loss-of-function. Alternatively, as the criteria for most of these classifications are unclear or not stringent, it cannot be excluded that many of these N-terminal (and other) substitutions in MSH2 are classified incorrectly.

Dominant-negative phenotypes of Msh2 variants and cancer predisposition

The extensive analysis of the mutants generated here has allowed the identification of dominant-negative phenotypes of Msh2 mutants N671K and N671I. In families suspected of Lynch syndrome, both identical (N671K) and non-identical (N671Y) substitutions at MSH2 N671 have been described [Chao et al., 2008; Lu et al., 1996]. The identification of dominant-negative mutations is of particular relevance, since these may confer a disease phenotype already in a heterozygous state. This may include the accelerated onset and expanded tropism of cancer. In fact, a family with an MSH2 N671Y mutation displays both an extended spectrum (colon, ovary, breast, esophagus) and a young age of onset of cancer [37 years; Lu et al., 1996]. In addition to substitutions N671I and N671K, several other mutants obtained in our screen have retained Msh2 protein expression (Fig. 4C). From a diagnostic point-of-view this is a significant observation since loss of protein expression frequently is used to direct MMR gene sequencing in individuals suspected of Lynch syndrome [Lynch and de la Chapelle, 2003]. Therefore, even when immunostaining of a tumor is positive for the presence of all MMR proteins, the patient should be analyzed for VUS in the MMR genes.

Approaches to generate comprehensive reverse diagnosis catalogs

In this study we have focused on the identification of ENU-induced mutants that had retained full-length *Msh2* cDNA. For this reason, *Msh2* mutants that cause defects at the RNA level will have escaped identification. Indeed, in many of the clones, full-length cDNA or protein was not detected and these were discarded (Table 1). In the future, deleterious mutations that disrupt transcription, splicing or mRNA stability [Sharp et al., 2004; Tournier et al., 2008; Betz et al., 2010] may be identified by genomic, rather than cDNA, sequencing.

Analysis of the mutation spectrum at *Msh2* in the clones generated here reveals ENU-induced substitutions at both A·T and G·C base pairs (Fig. S4), supporting the notion that ENU is a wide-spectrum mutagen [Russel et al., 1979]. The identification of different ENU-induced substitutions at the same amino acid in individual clones (Table 2) provides an indication of

the relative saturation of the prototypic screen described here. However, since mutagenesis by ENU may be not fully random, the mutagenesis and ensuing identification of all critical gene residues may be hindered. Higher saturation of the reverse diagnosis catalog may be achieved by the use of drugs with a complementary mutagenic activity, such as 4-nitroquinoline N-oxide [Ryu et al., 1999] or 7,12-dimethylbenz[a]anthracene [Manjanatha et al., 1996].

Integration of reverse diagnosis catalogs in Lynch syndrome diagnostics

It is conceivable that the genetic screens proposed here may not be fully saturating, even when performed at a large scale and using multiple mutagens, which results in the absence of critical residues from the reverse diagnosis catalog. In other instances, albeit at the same amino acid, the substitution identified in an individual may be different from that listed in the catalog. Such variants may correctly be assessed using cell-free MMR assays [Drost et al., 2010, 2012]. Conversely, cell-free assays may not correctly identify pathogenic variants when the mutation specifically affects *in vivo* characteristics of the variant protein, such as its stability or intracellular localization, and these may rely on reverse diagnosis catalogs for functional assessment. We envision that, after their stratification with human variants of known pathogenicity, reverse diagnosis catalogs and cell free activity assays will become important components of an integrated diagnostic procedure that also includes independent data sources such as segregation, pathological and *in silico* data [Rasmussen et al., 2012]. Such an integrated diagnostic procedure is a prerequisite to relieve the bottleneck between personalized genomics and targeted healthcare in Lynch syndrome. In the upcoming era of personalized genomics, this approach may provide a paradigm for the analysis of pathogenicity of gene variants in any genetic disease that has a defined underlying cellular and biochemical phenotype.

MATERIALS AND METHODS

Determination of genomic mutator phenotypes, microsatellite instability analysis, methylation tolerance assays and other procedures are described in the Supplementary Methods.

Construction of an Msh2 reverse diagnosis catalog. The outline of the generation of a reverse diagnosis catalog for Msh2 is described in Fig. 1. $5x10^6$ *Msh2*$^{+/-}$ mouse embryonic stem (ES) cells [de Wind et al., 1995] were seeded in a P90 dish on irradiated feeder cells in ES culture medium (Gibco, Carlsbad, CA, USA). To induce mutations, cells were incubated in medium containing 0.2 mg/ml ethylnitrosourea (ENU, Sigma-Aldrich, St. Louis, MI, USA) for 2 hours [Borgdorff et al., 2005], or mock-treated. Cells were propagated for a week to allow loss of wild type *Msh2* transcripts and protein. Then, cells were seeded at a density of $2x10^6$ cells per T90, followed by double treatment, with a seven-day interval, with 40μM 6TG (Sigma-Aldrich) for 4 hours. Inadvertent *Hprt*-deficient cells were eliminated by culturing the clones one week in Hypoxantine-Aminopterine-Thymidine (HAT, Gibco)-supplemented medium. Surviving clones were picked and expanded in 96-well plates. Next, clones that had lost the wild type *Msh2* allele by loss of heterozygosity, rather than by mutational inactivation, were discarded after their identification by allele-specific PCR

on genomic DNA [Borgdorff et al., 2005]. In the remaining clones the status of the *Msh2* mRNA was analyzed by reverse transcriptase-PCR analysis. The *Msh2* open reading frame of full-length cDNA clones was sequenced as described [Borgdorff et al., 2006].

Statistical analyses. For experiments measuring microsatellite instability (Fig. 2A), mutator phenotypes (Fig. 2B) and methylation tolerance (Fig. 2C), p-values were calculated by comparing with the *Msh2*$^{+/-}$ ES cell line, using a one-tailed t-test assuming unequal variance. For the cell-free MMR complementation assay (Fig. 4A), p-values were calculated for every extract individually, comparing repair efficiencies in the absence or presence of exogenous MSH2/MSH6 using a one-tailed t-test assuming unequal variance. For the cell-free MMR assay (Fig. 4B), p-values were calculated compared to the Mock using a one-tailed t-test assuming unequal variance.

ACKNOWLEDGEMENTS

We thank Jacob G. Jansen and Piya Temviriyanukul for critically reading the manuscript. We thank Joyce H.G. Lebbink (Erasmus Medical Center, Rotterdam, The Netherlands) and Meindert H. Lamers (MRC, Cambridge, United Kingdom) for help with crystal structure modeling. We thank Flora Groothuizen and Titia Sixma (The Netherlands Cancer Institute, Amsterdam, The Netherlands) for recombinant wild type MSH2/MSH6.

GRANT SUPPORT

This work was funded by the European Union (grant FP6-018754).

REFERENCES

Arnold S, Buchanan DD, Barker M, Jaskowski L, Walsh MD, Birney G, Woods MO, Hopper JL, Jenkins MA, Brown MA, Tavtigian SV, Goldgar DE et al. 2009. Classifying MLH1 and MSH2 variants using bioinformatic prediction, splicing assays, segregation, and tumor characteristics. Hum Mutat 30(5):757–770.

Barnetson RA, Cartwright N, van Vliet A, Haq N, Drew K, Farrington S, Williams N, Warner J, Campbell H, Porteous ME, Dunlop MG. 2008. Classification of ambiguous mutations in DNA mismatch repair genes identified in a population-based study of colorectal cancer. Hum Mutat 29(3):367–374.

Betz B, Theiss S, Aktas M, Konermann C, Goecke TO, Möslein G, Schaal H, Royer-Pokora B. 2010. Comparative in silico analyses and experimental validation of novel splice site and missense mutations in the genes MLH1 and MSH2. J Cancer Res Clin Oncol 136(1):123–134.

Borgdorff V, van Hees-Stuivenberg S, Meijers CM, de Wind N. 2005. Spontaneous and mutagen-induced loss of DNA mismatch repair in Msh2-heterozygous mammalian cells. Mutat Res 574(1-2):50–57.

Borgdorff V, Pauw B, van Hees-Stuivenberg S, De Wind N. 2006. DNA mismatch repair mediates protection from mutagenesis induced by short-wave ultraviolet light. DNA Repair 5(11):1364–1372.

Burn J, Gerdes AM, Macrae F, Mecklin JP, Moeslein G, Olschwang S, Eccles D, Evans DG, Maher ER, Bertario L, Bisgaard ML, Dunlop MG et al. 2011. Long-term effect of aspirin on cancer risk in carriers of hereditary colorectal cancer: an analysis from the CAPP2 randomised controlled trial. Lancet 378(9809):2081–2087.

Castells A, Castellví-Bel S, Balaguer F. 2009. Concepts in familial colorectal cancer: where do we stand and what is the future? Gastroenterology 137(2):404–409.

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). Hum Mutat 29(6):852–860.

Dinh TA, Rosner BI, Atwood JC, Boland CR, Syngal S, Vasen HF, Gruber SB, Burt RW. 2011. Health benefits and cost-effectiveness of primary genetic screening for Lynch syndrome in the general population. Cancer Prev Res (Phila) 4(1):9–22.

Drost M, Zonneveld JE, 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. Hum Mutat 31(3):247–253.

Drost M, Zonneveld JB, van Hees S, Rasmussen LJ, Hofstra RM, de Wind N. 2012. A rapid and cell-free assay to test the activity of lynch syndrome-associated MSH2 and MSH6 missense variants. Hum Mutat 33(3):488–494.

Glaab WE, Risinger JI, Umar A, Barrett JC, Kunkel TA, Tindall KR. 1998. Resistance to 6-thioguanine in mismatch repair-deficient human cancer cell lines correlates with an increase in induced mutations at the HPRT locus. Carcinogenesis 19(11):1931–1937.

Heinen CD. 2010. Genotype to phenotype: analyzing the effects of inherited mutations in colorectal cancer families. Mutat Res 693(1-2):32–45.

Hewish M, Lord CJ, Martin SA, Cunningham D, Ashworth A. 2010. Mismatch repair deficient colorectal cancer in the era of personalized treatment. Nat Rev Clin Oncol 7(4):197–208.

Hsieh P, Yamane K. 2008. DNA mismatch repair: molecular mechanism, cancer, and ageing. Mech Ageing Dev 129(7-8):391–407.

Jansen JG, Vrieling H, van Teijlingen CM, Mohn GR, Tates AD, van Zeeland AA. 1995. Marked differences in the role of O6-alkylguanine in hprt mutagenesis in T-lymphocytes of rats exposed in vivo to ethylmethanesulfonate, N-(2-hydroxyethyl)-N-nitrosourea, or N-ethyl-N-nitrosourea. Cancer Res 55(9):1875–1882.

Lagerstedt Robinson K, Liu T, Vandrovcova J, Halvarsson B, Clendenning M, Frebourg T, Papadopoulos N, Kinzler KW, Vogelstein B, Peltomäki P, Kolodner RD, Nilbert M, Lindblom A. 2007. Lynch syndrome (hereditary nonpolyposis colorectal cancer) diagnostics. J Natl Cancer Inst 99(4):291–299.

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. J Mol Biol 366(1):53-66

Lu SL, Akiyama Y, Nagasaki H, Nomizu T, Ikeda E, Baba S, Ushio K, Iwama T, Maruyama K, Yuasa Y. (1996) Loss or somatic mutations of hMSH2 occur in hereditary nonpolyposis colorectal cancers with hMSH2 germline mutations. Jpn J Cancer Res 87(3):279–287.

Lynch HT, de la Chapelle A. 2003. Hereditary colorectal cancer. N Engl J Med 348(10):919–932.

Manjanatha MG, Chen JB, Shaddock JG Jr, Harris AJ, Shelton SD, Casciano DA. 1996. Molecular analysis of lacI mutations in Rat2 cells exposed to 7,12-dimethylbenz[a]anthracene: evidence for DNA sequence and DNA strand biases for mutation. Mutat Res 372(1):53–64.

Peltomäki P, Vasen HFA. 2004. Mutations associated with HNPCC predisposition -- Update of ICG-HNPCC/INSiGHT mutation database. Dis Markers 20(4-5):269–276.

Rasmussen LJ, Heinen CD, Royer-Pokora B, Drost M, Tavtigian S, Hofstra RM, de Wind N. 2012. Pathological assessment of mismatch repair gene variants in Lynch syndrome: Past, present, and future. Hum Mutat 33(12):1617–1625.

Russell WL, Kelly EM, Hunsicker PR, Bangham JW, Maddux SC, Phipps EL. 1979. Specific-locus test shows ethylnitrosourea to be the most potent mutagen in the mouse. Proc Natl Acad Sci USA 76(11):5818–5819.

Ryu J, Youn J, Kim Y, Kwon O, Song Y, Kim H, Cho K, Chang I. 1999. Mutation spectrum of 4-nitroquinoline N-oxide in the lacI transgenic Big Blue Rat2 cell line. Mutat Res 445(1):127–135.

Sharp A, Pichert G, Lucassen A, Eccles D. 2004. RNA analysis reveals splicing mutations and loss of expression defects in MLH1 and BRCA1. Hum Mutat 24(3):272.

Tournier I, Vezain M, Martins A, Charbonnier F, Baert-Desurmont S, Olschwang S, Wang Q, Buisine MP, Soret J, Tazi J, Frébourg T, Tosi M. 2008. A large fraction of unclassified variants of the mismatch repair genes MLH1 and MSH2 is associated with splicing defects. Hum Mutat 29(12):1412–1424.

de Wind N, Dekker M, Berns A, Radman M, te Riele H. 1995. Inactivation of the mouse Msh2 gene results in mismatch repair deficiency, methylation tolerance, hyperrecombination, and predisposition to cancer. Cell 82(2):321–330.

de Wind N, Dekker M, Claij N, Jansen L, van Klink Y, Radman M, Riggins G, van der Valk M, van't Wout K, te Riele H. 1999. HNPCC-like cancer predisposition in mice through simultaneous

loss of Msh3 and Msh6 mismatch-repair protein functions. Nat Genet 23(3):359–362.

Warren JJ, Pohlhaus TJ, Changela A, Iyer RR, Modrich PL, Beese LS. 2007. Structure of the human MutSalpha DNA lesion recognition complex. Mol Cell 26(4):579–592.

MATERIALS AND METHODS

Crystal structure modeling. Crystal structures were visualized using PyMOL software (Delano Scientific, San Francisco), using the Protein Database (PDB) file from the published human MSH2/MSH6 crystal structure (MMDB ID: 53938; (1)).

Determination of genomic mutator phenotypes at *Hprt*. Spontaneous mutagenesis at the monoallelic genomic *Hprt* gene was determined essentially as described (2). Briefly, cell lines were cultured in HAT-supplemented medium for two subsequent passages to eliminate any pre-existing *Hprt*-deficient cells, after which cell lines were propagated for three passages in the absence of HAT, to allow the accumulation of spontaneous mutations at *Hprt*. Cells were then continuously cultured in medium containing 60μM 6TG (Sigma-Aldrich) to select for *Hprt*-deficient clones. Of note, MMR-deficient (but *Hprt* proficient) cells do not survive these stringent selection conditions. After 10 days, *Hprt*-deficient colonies were fixed, stained and counted. Counts were corrected for cloning efficiencies.

Microsatellite instability was analyzed by isolating approximately 50 subclones of each *Msh2*-mutant cell line. From these subclones, genomic DNA was isolated by Proteinase K lysis, and microsatellite sizes were determined after amplification by PCR and subsequent fragment analysis, as described previously (3). Data was analyzed using GeneMarker software (Softgenetics, State College, PA).

Methylation tolerance assays. *Msh2*-mutant cells were treated with *N*-Methyl-*N*'-Nitro-*N*-Nitrosoguanidine and simultaneously with O^6-benzylguanine, an inhibitor of the repair enzyme Methyl Guanine Methyl Transferase, as described (4). After three days, adhering cells were counted, survival relative to the untreated cells was calculated, and the IC_{50} was determined.

***In vitro* MMR assays.** MMR assays to test humanized variant MMR proteins were performed as described (5). To test MMR activity in cell extracts from the validation panel, 130 μg of cytoplasmic extract was assayed in the absence or presence of 100ng of recombinant wild type human MSH2/MSH6. (kindly provided by Prof. T. Sixma, The Netherlands Cancer Institute, Amsterdam). The results for these experiments were verified in a second set of experiments using independent batches of cell extracts.

Western blotting. Analysis of Msh2 and Msh6 protein expression by Western blotting was performed as described (4).

Electrophoretic mobility shift assays. Oligonucleotide sequences were as described (6). 170 fmol of double-stranded matched or mismatched (G·T) oligonucleotide, labeled at the 5' end using γ-^{32}P ATP and polynucleotide kinase, was incubated with 20 μg of cytoplasmic extract in 1x DNA binding buffer (12% glycerol, 20 mM HEPES pH 7.9, 100 mM NaCl, 1 mM DTT, and 0.1 mM EDTA, 0.05 μg/μl poly-dIdC and 425 fmol unlabeled matched oligonucleotide) for 20 min. at 37°C in a total volume of 20μl. For adenine nucleotide challenge experiments, ATP was added 10 min after addition of the DNA probe. The reaction mixture was subjected to electrophoresis in a 4% polyacrylamide:bisacrylamide (29:1) gel in 0.5x TBE buffer containing 5% glycerol. The gels were dried, signals were

visualized using a Cyclone Plus Phosphor Imager (PerkinElmer, Waltham, MA, USA) and images were analyzed using OptiQuant software.

SUPPLEMENTARY REFERENCES

Warren JJ et al. (2007) Structure of the human MutSalpha DNA lesion recognition complex. *Mol Cell* 26(4):579–592.

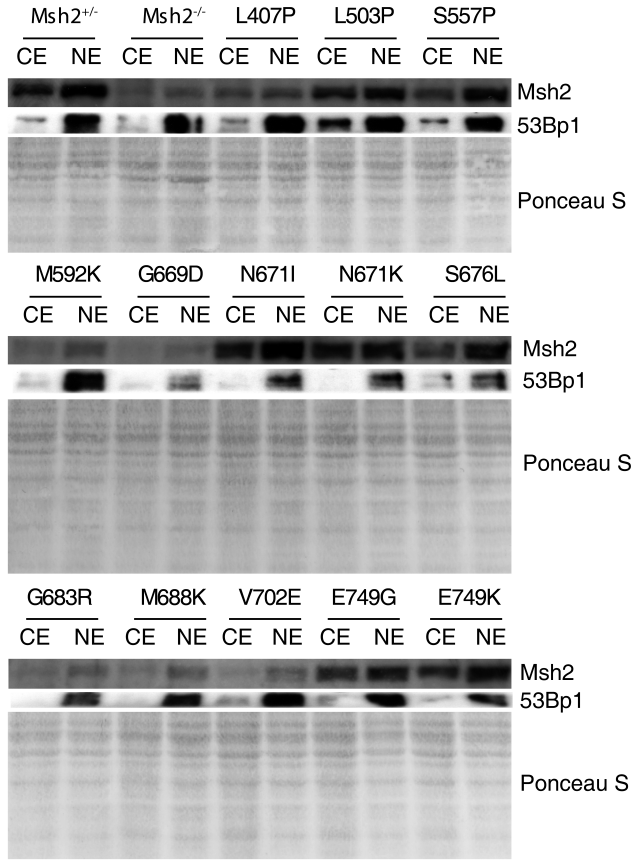
Borgdorff V, Pauw B, van Hees-Stuivenberg S, de Wind N (2006) DNA mismatch repair mediates protection from mutagenesis induced by short-wave ultraviolet light. 5(11):1364–1372.

Bacher JW, Abdel Megid WM, Kent-First MG, Halberg RB (2005) Use of mononucleotide repeat markers for detection of microsatellite instability in mouse tumors. *Mol Carcinog* 44(4):285–292.

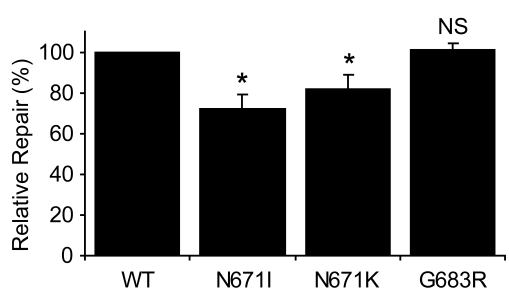
de Wind N et al. (1999) HNPCC-like cancer predisposition in mice through simultaneous loss of Msh3 and Msh6 mismatch-repair protein functions. *Nat Genet* 23(3):359–362.

Drost M et al. (2012) A rapid and cell-free assay to test the activity of lynch syndrome-associated MSH2 and MSH6 missense variants. *Hum Mutat* 33(3):488–494.

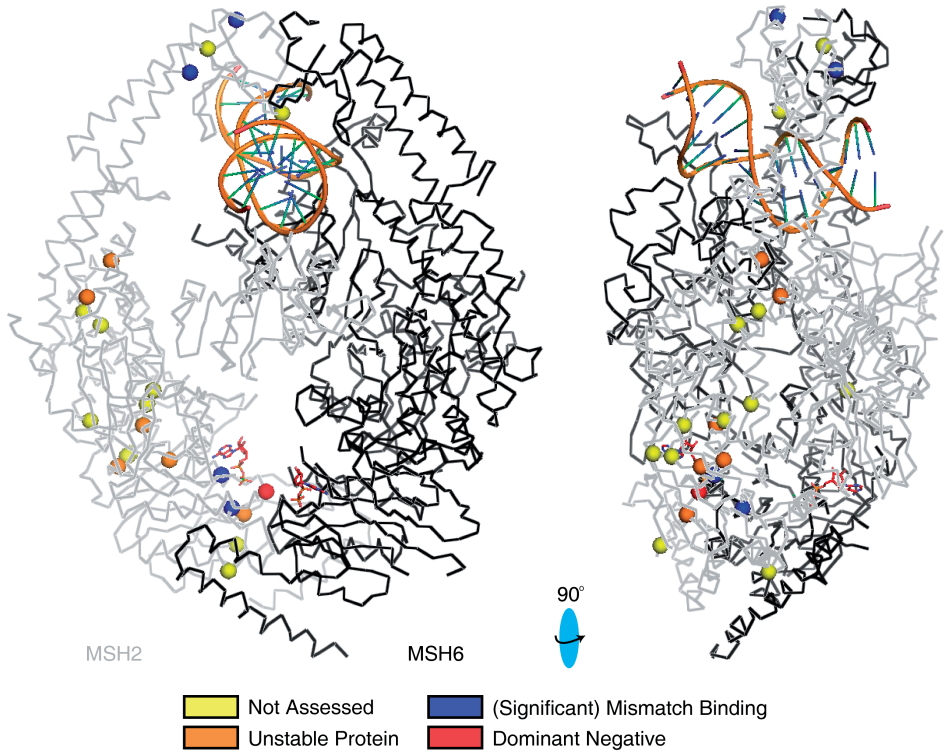
Lützen A, de Wind N, Georgijevic D, Nielsen FC, Rasmussen LJ (2008) Functional analysis of HNPCC-related missense mutations in MSH2. *Mutat Res* 645(1-2):44–55.



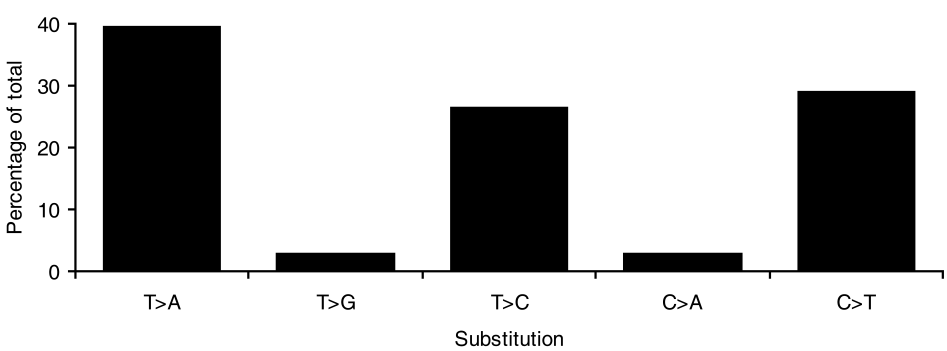
Supplementary Figure 1. *Western blot analysis of cytoplasmic and soluble nuclear extracts (CE and NE, respectively) of mutant cell lines. 53Bp1 is used as a nuclear marker. Ponceau S staining reveals equal loading of extracts.*



Supplementary Figure 2. *Inhibition of in vitro mismatch repair in HeLa cell extracts by in vitro expressed variants. MMR substrate was incubated with 12 µl of reticulocyte lysate containing expressed wild type or variant MSH2/MSH6 in the presence of buffers and cofactors. After 10 min. of incubation at 37°C, 75 µg of HeLa extract was added and the reaction was incubated for an additional 25 min. Reactions were continued as described (5). Bars represent mean±SEM. *=p<0.05 compared to WT. NS=Not significant. Partial, rather than complete, inhibition by the N671 variants was observed since the substrate concentration exceeds the MSH2/MSH6 concentration.*



Supplementary Figure 3. *Mapping of the residues mutated in our screen to the crystal structure of the human MSH2/MSH6 heterodimer. All residues are colored according to their in vitro phenotype. In grey: MSH2, in black: MSH6. Mutated residues are shown as spheres.*



Supplementary Figure 4. *Mutations induced by ENU at the wild type Msh2 allele. This spectrum illustrates the broad mutational capacity of ENU.*

Supplementary Table 1. Prototypic *Msh2* reverse diagnosis catalog.

Mutation[1]	cDNA position	(codon change)	MAPP MMR[2]	Microsatellite Intability	*In vivo* protein level[3]	Spontaneous mutator phenotype
M148K	c.443	ATG>AAG	3.75	ND[5]	ND	ND
I346N	c.1037	ATC>AAC	18.3	ND	ND	ND
I356K	c.1067	ATA>AAA	14.66	ND	ND	ND
I356R	c.1067	ATA>AGA	16.34	ND	ND	ND
L407P	c.1220	CTG>CCG	18.56	+	–	+
L503P	c.1508	CTC>CCC	14.91	+	±	+
K546E	c.1636	AAG>GAG	23.7	ND	ND	ND
N553K	c.1659	AAC>AAA	24.75	ND	ND	ND
S557P	c.1669	TCC>CCC	9.71	+	±	+
M592K	c.1775	ATG>AAG	8.21	+	–	+
L595P	c.1784	CTC>CCC	7.68	ND	ND	ND
L599P	c.1796	CTG>CCG	17.42	ND	ND	ND
G669D	c.2006	GGT>GAT	35.86	+	–	+
N671I[6]	c.2012	AAT>ATT	22.05	+	+	+
N671K[6]	c.2013	AAT>AAA	25.31	+	+	+
S676L	c.2027	TCA>TTA	28.64	+	+	+
G683R	c.2047	GGG>AGG	44	+	–	+
M688K	c.2063	ATG>AAG	32.5	+	–	+
V695E	c.2084	GTG>GAG	30.37	ND	ND	ND
S699P	c.2095	TCG>CCG	5.97	ND	ND	ND
A700E	c.2099	GCA>GAA	37.91	ND	ND	ND
V702E	c.2105	GTG>GAG	33.26	+	–	+
E749G	c.2246	GAG>GGG	28.23	+	+	+
E749K	c.2245	GAG>AAG	19.98	+	+	+
G761R	c.2281	GGG>AGG	40.89	ND	ND	ND
C822Y	c.2465	TGT>TAT	6.48	ND	ND	ND

[1] Amino acid position. [2] *In silico* MAPP-MMR prediction of pathogenicity. Scores above 4.55 are predicted to be pathogenic, variant M148K is borderline pathogenic. [3] +: (near) wild type protein level, ±: reduced protein level. –: no protein. [4] CIMRA: Complete *In vitro* MMR Assay (Fig 4B). [5] ND: not determined (not in the validation panel). [6] Dominant-negative variant.

Methylation tolerance	Repair in cell extracts	Complementation with wt MSH2/MSH6	Activity in CIMRA[4]	Mismatch binding	ATP-induced Mismatch release	Corresponding VUS in LOVD
ND	ND	ND	ND	ND	ND	
ND	ND	ND	ND	ND	ND	
ND	ND	ND	ND	ND	ND	
ND	ND	ND	ND	ND	ND	
+	–	+	–	ND	ND	L407L
+	–	+	–	±	+	L503P
ND	ND	ND	ND	ND	ND	
ND	ND	ND	ND	ND	ND	
+	–	+	–	±	ND	T557P
+	–	+	–	ND	ND	M592V
ND	ND	ND	ND	ND	ND	
ND	ND	ND	ND	ND	ND	
+	–	+	–	ND	ND	**G669D**/R
+	–	–	–	+	–	**N671K**/Y
+	–	–	–	+	–	
+	–	+	–	+	+	S676P
+	–	+	–	ND	ND	**G683R**/W
+	–	+	–	ND	ND	M688R/I
ND	ND	ND	ND	ND	ND	
ND	ND	ND	ND	ND	ND	
ND	ND	ND	ND	ND	ND	
+	–	+	–	ND	ND	V702G
+	–	+	–	–	ND	
+	–	+	–	±	+	E749K
ND	ND	ND	ND	ND	ND	
ND	ND	ND	ND	ND	ND	

A Genetic Screen to Identify Inactivating Variants in the *Msh6* Gene Suggests Defective Mismatch Binding as a Major Mechanism of Pathogenicity

7

Mark Drost, Fabienne Calléja, José Zonneveld, Sandrine van Hees, Anne Lützen, Finn Cilius Nielsen, Lene Juel Rasmussen and Niels de Wind

ABSTRACT

Lynch syndrome (LS) is a hereditary cancer predisposition caused by heterozygous inactivating mutations in one of the DNA mismatch repair (MMR) genes. A significant proportion of all variants found in these genes are so-called Variants of Uncertain Significance (VUS), generally missense mutations. The inability to diagnose such variants hinders presymptomatic diagnosis and precludes both targeted chemotherapy and chemoprevention for carriers of the VUS and their effected relatives.

To enable the diagnosis of such variants, we have developed a procedure to mutagenize, identify and catalog residues of MMR proteins that are essential for its function. Mapping of the essential residues on the protein results in "Reverse Diagnosis Catalogs". In case a VUS from a suspected patient and a residue identified in the catalog overlap, there is a high probability that the VUS is pathogenic.

In this work, we describe the construction and validation of a pilot catalog for the mismatch repair protein Msh6. We show the applicability of this approach for the diagnostic assessment of *MSH6* VUS. In addition, we make use of the generated cell lines to study (intra)molecular signaling in MutSα. We describe a large panel of characterized Msh6 mutants and show that many residues are required for mismatch binding *in vitro*, suggesting that defective mismatch binding is a major mechanism of pathogenicity.

INTRODUCTION

The hereditary cancer predisposition syndrome Lynch syndrome (LS; or Hereditary Nonpolyposis Colorectal Cancer, HNPCC; MIM #s 120435, 609310) is an autosomal dominant disorder that predisposes to various cancer types, including colon and endometrial cancer [Boland and Lynch, 2013]. The syndrome is caused by the germline inactivation of a single allele of one of the DNA mismatch repair (MMR) genes *MSH2*, *MSH6*, *MLH1* or *PMS2* [de la Chapelle, 2004]. The MMR pathway corrects misincorporations or insertion/deletions loops (IDLs) generated spontaneously during DNA replication, and thereby prevents the accumulation of spontaneous mutations (reviewed in [Hsieh and Yamane, 2008; Jiricny, 2013]). Loss of the remaining wild type allele of an MMR gene in LS patients induces cellular MMR deficiency, which leads to a mutator phenotype that accelerates carcinogenesis.

The MSH2 and MSH6 proteins compose the MutSα heterodimer, which is the initial factor in MMR. MutSα forms a θ-like structure where the (mis)matched DNA is in the upper channel, whereas the lower channel is empty, and MSH2 and MSH6 line the sides along the long axis [Warren et al., 2007]. DNA mismatch repair is initiated when MutSα binds a mismatch through a conserved Phe-X-Glu motif in the mismatch binding domain of the MSH6 subunit, proximal to the mismatch [Warren et al., 2007]. Upon DNA binding, an ADP in the more distal ATPase domain is exchanged for an ATP molecule. In case MutSα accidentally binds homoduplex DNA, this ATP is rapidly hydrolyzed, inducing a conformational change that allows the clamp domains (which are at the opposite end of the heterodimer) to open and allow MutSα to fall of the DNA. In case MutSα binds a mismatch ATP hydrolysis is inhibited, the mismatch binding domain is retracted and MutSα is converted into a sliding clamp [Gradia et al., 1999; Antony and Hingorani, 2003], which can freely move along the DNA helix contour [Gorman et al., 2007]. These successive events all rely on long-range intramolecular signaling between distant domains. ATP binding, its hydrolysis as well as the conformational changes are essential for mismatch repair and are thought to play key roles in mismatch discrimination [Hays et al., 2005]. All of these steps contribute to the recruitment of the MutLα heterodimer, which consists of MLH1 and PMS2. PMS2 has endonuclease activity and nicks the nascent strand, 5' to the mismatch [Kadyrov et al., 2006]. Such nicks serve as entry points for the exonuclease EXO1, which degrades the nascent strand, after which the strand is resynthesized and repair is completed [Dzantiev et al., 2004; Constantin et al., 2005].

The genetics and biochemistry of the MMR pathway have been established in detail, and this has been instrumental in the diagnosis of LS. LS diagnosis is only definitive once an inactivating mutation in an MMR gene has been identified. This allows the presymptomatic screening of relatives [Vasen et al., 2013]. Confirmed LS patients enroll in lifelong preventive surveillance programs and benefit from personalized chemoprevention and chemotherapy [Hewish et al., 2010; Burn et al., 2011]. Of all mutations found in the *MSH6* gene, ~37% gives rise to single amino acid substitutions [Plazzer et al., 2013]. Since the causality of such variants is difficult to assess, they are often referred to as Variants of Uncertain Significance (VUS). Lack of classification of such VUS precludes LS diagnosis for the proband and his/

her direct relatives [Sijmons et al., 2013]. For this reason it is of great importance to develop approaches for the diagnostic assessment of VUS in MMR genes [Rasmussen et al., 2012].

We have previously described an *in vivo*, genetic screen-based approach which combines the large-scale mutagenesis and identification of critical amino acids of *Msh2* to generate "reverse diagnosis catalogs" [Drost et al., 2013]. Such a catalog enables us to judge the importance of residues in MMR proteins for its function, and aids in the rapid assessment of pathogenicity of VUS. Furthermore, the missense mutant cell lines that are generated in such a screen may serve as tools to study the mechanisms of MMR. Here, we describe the construction and validation of a reverse diagnosis catalog for the mismatch repair protein Msh6, and show the applicability of such an approach for the diagnostic assessment of *MSH6* VUS. Additionally, we make use of the generated cell lines to study (intra)molecular signaling in MutSα, describe a large panel of characterized mutants and show that most mutated residues are required for mismatch binding *in vitro*.

RESULTS

To generate a prototype Msh6 reverse diagnosis catalog, we applied a stepwise procedure [Drost et al., 2013]. To randomly introduce mutations in *Msh6*, mouse embryonic stem cells (mESC) heterozygous for *Msh6* (*Msh6*$^{+/-}$; [de Wind et al., 1999]) were treated with the point mutagen *N*-ethyl-*N*-nitrosourea (ENU). ENU is a suited mutagen for this screen, as it induces a wide spectrum of point mutations [Jansen et al., 1994] and a relatively low amount of loss-of-heterozygosity (LOH; [Borgdorff et al., 2005]). After ENU treatment, cells that have lost functional MMR are selected using 6-thioguanine (6-TG), which is a nucleotide analog to which MMR-deficient cells are tolerant [Swann et al., 1996]. As 6-TG also selects for *Hprt* mutants, such mutants were selected against by culturing the cells in hypoxanthine-aminopterin-thymidine-supplement-containing medium. We screened against clones that lost functional MMR due to LOH of the *Msh6* locus by means of an allele-specific PCR. The remaining clones presumably have lost functional MMR due to an ENU-induced inactivating nucleotide substitution in the *Msh6* gene.

In this work, we have focused specifically on amino acid substitutions, as such proteins do not only allow us to assess pathogenicity of VUS in *MSH6*, but also to perform functional studies. To identify such substitutions, we selected clones that retained expression of full-length *Msh6* cDNA, as judged by reverse transcription-PCR, and sequenced the complete *Msh6* open reading frame of these clones. This led to the identification of 76 clones that carried a missense substitution, reflecting 51 unique substitutions, affecting 45 residues. All of these were compiled into a reverse diagnosis catalog (Figure 1 and Table 1).

The mapping of the substitutions on the Msh6 protein showed that, in contrast to Msh2 [Drost et al., 2013], substitutions cover all domains (Figure 1). Substitutions in the mismatch binding domain (MBD) of Msh6 are abundant, as >30% of all substitutions found are located here. These substitutions include residue F431, which was shown to contact mismatches directly [Bowers et al., 1999; Dufner et al., 2000; Warren et al., 2007], but also residues involved in non-

Table 1. A Pilot Reverse Diagnosis Catalog of Msh6.

Mutation[¥]	cDNA[*]	LOVD[†]	Mutation[¥]	cDNA[*]	LOVD[†]
E264V	c.791 A>T		**C691R**	c.2071 T>C	**C694R**
L310F	c.927_928 delinsTT		R769W	c.2305 C>T	
W371R	c.1111 T>A		L770P	c.2324 T>C	
W412R	c.1233 T>C			c.2308 T>C	
V423G	c.1268 T>G		L818P	c.2453 T>C	
F431L	c.1291 T>C		L872P	c.2615 T>C	
	c.1293 T>G		S881T	c.2641 T>A	
F431S	c.1292 T>C		I964K	c.2891 T>A	
H436R	c.1307 A>G		Y991S	c.2972 A>C	
D438E	c.1314 T>A		Y1003N	c.3007 T>A	
L448P	c.1343 T>C	**L449P**	T1005I	c.3014 C>T	
H457Y	c.1369 C>T		N1134K	c.3402 T>A	
Q484K	c.1450 C>A		M1135K	c.3404 T>A	
Q484R	c.1451 A>G		**G1137D**	c.3410 G>A	**G1139S**
V507E	c.1520 T>A	V509A	**G1137S**	c.3409 G>A	
V508E	c.1523 T>A		S1139P	c.3415 T>C	
R510G	c.1528 A>G		L1141P	c.3422 T>C	
I515N	c.1544 T>A	**I516N**	L1154R	c.3461 T>G	
G519V	c.1556 G>T		**G1155D**	c.3464 G>A	G1157S
G519D	c.1556 G>A		C1163R	c.3487 T>C	
T520K	c.1559 C>A		L1165P	c.3494 T>C	
G556D	c.1667 G>A		G1176D	c.3464 G>A	
V557D	c.1670 T>A		**D1211G**	c.3632 A>G	D1213V
L636P	c.1907 T>C		**D1211E**	c.3633 T>A	
L678P	c.2033 T>C		**T1217I**	c.3650 C>T	**T1219I**
G683D	c.2048 G>A		**H1246R**	c.3737 A>G	H1248D
V685D	c.2054 T>A				

[¥] Boldface represents the mutants included in the validation panel. Amino acid numbering is based on the Msh6 GenBank reference sequence NP_034960.1 with +1 corresponding to the translation initiation codon.
[*] Nucleotide numbering reflects cDNA numbering with +1 corresponding to the A of the translation initiation codon in the *Msh6* GenBank reference sequence NM_010830.2.
[†] Substitutions reported in patients suspected of LS, as described in the Leiden Open Variant Database (LOVD). In bold: identical substitution found in reverse diagnosis catalog and LOVD, allowing to assess pathogenicity to the patient VUS. Amino acid numbering is based on the MSH6 GenBank reference sequence NP_000170.1 with +1 corresponding to the translation initiation codon.

specific contacts with the DNA backbone such as L448, V507, V508 and R510 [Warren et al., 2007]. Besides from residues in the MBD, residues in other critical domains are overrepresented in our screen, such as those within the Walker A (also known as P-loop) and Walker B motifs

(Figure 1A). Furthermore, residues R769 and L770 were proposed to be important for signal transduction between ATPase and DNA binding domains [Warren et al., 2007]. In addition, we and others have shown the importance of residue T1217 for mammalian [Yang et al., 2004; Drost et al., 2012; Geng et al., 2012] and *S. cerevisiae* MMR [Das Gupta and Kolodner, 2000; Hess et al., 2006]. Finally, functional defects of substitutions at the *S. cerevisiae* equivalents of Q484 [Das Gupta and Kolodner, 2000] and H1246 [Hess et al., 2006; Cyr and Heinen, 2008] were shown before. These observations combined suggest that the Msh6 residues identified in the screen performed here are enriched for critical/important residues.

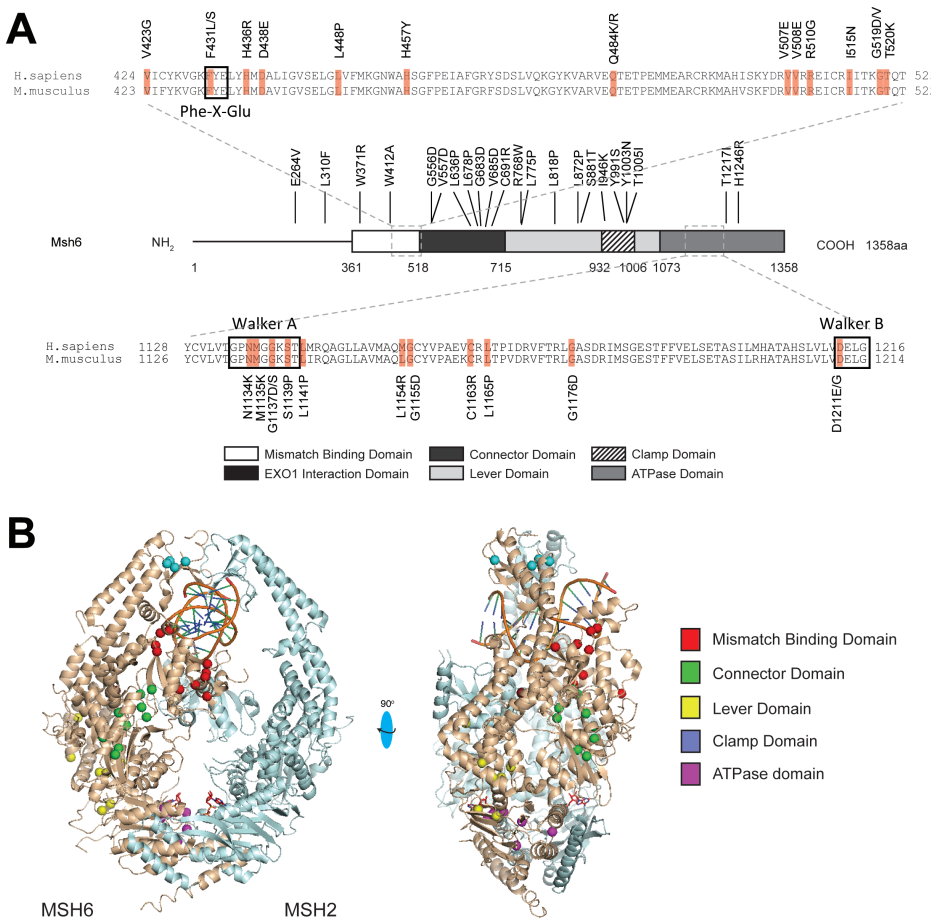


Figure 1. *A Pilot Msh6 Reverse Diagnosis Catalog. (A) A schematic representation of all inactivating missense substitutions in Msh6 identified in this screen. All domains were adapted from [Warren et al., 2007]. The top and bottom of the figure shows an alignment of the MBD and ATPase domains of human and mouse Msh6, highlighting residues that were mutated in the screen described here (orange boxes). Numbers reflect amino acid numbering. (B) Mapping of residues mutated in this screen to the MutSα crystal structure. In wheat color, MSH6; in cyan, MSH2. Mutated residues are shown as spheres. Spheres are colored based on the domain the residue is part of (see legend in figure). The MSH6 crystal structure starts from residue 362, therefore not all residues are depicted.*

Reverse Diagnosis Catalogs Aid To Assign Pathogenicity to *MSH6* VUS

To investigate the applicability of our screen for the diagnosis of VUS in *MSH6*, we aligned the substitutions in our catalog with those in the Leiden Open Variation Database (LOVD; [Fokkema et al., 2011]). This database contains a large collection of variants in MMR genes found in humans, irrespective of their pathogenicity. Of the 45 residues affected in our screen, eight are represented in the LOVD. More precisely, the four substitutions Msh6-L449P, I515N, C691R and T1217I from our screen are the functional equivalents of the human MSH6-L449P, I516N, C694R and T1219I alleles in the LOVD. This result strongly suggests that these human alleles are pathogenic and provides support for the applicability of the Msh6 reverse diagnosis catalog to identify pathogenic human *MSH6* variants.

Reverse Diagnosis Catalog-Derived Cell Lines Show Hallmarks of MMR Deficiency

To validate this screen experimentally, we have tested a validation panel of cell lines for phenotypic hallmarks of MMR deficiency. The panel was established from cell lines that carry substitutions at residues that are represented in the LOVD, either as an identical or different substitution (Table 1). MMR deficient cells often show instability of simple repeat sequences, a phenotype referred to as microsatellite instability (MSI). Even though MSI is considered a diagnostic hallmark of LS cancers [Boland and Goel, 2010], *MSH6* deficient tumors show only low levels of MSI due to functional redundancy of MutSα with another heterodimer, MutSβ. This heterodimer is composed of MSH2 and MSH3 and can partly substitute for MutSα in the repair of IDLs. Higher levels of MSI can only be detected in cells or tumors expressing dominant negative *Msh6* alleles, which interfere with both MutSα and MutSβ function. We tested for MSI in the validation panel by measuring the size of different microsatellites in 40-50 subclones of each line. We detected low levels of MSI in all missense mutant cell lines similar to the *Msh6*$^{-/-}$ control line, suggesting that these lines have an Msh6-like MSI phenotype.

Due to the inability to repair replication errors, MMR deficient cells have a spontaneous mutator phenotype [Jiricny, 2013]. We therefore measured mutant frequencies at the *Hprt* reporter gene. All tested clones displayed a spontaneous mutator phenotype that is significantly higher than the *Msh6*$^{+/-}$ control line. We were unable to measure mutant frequencies in the Msh6-V508E line due to low cloning efficiency of this line. The T1217I allele displayed highly elevated mutation frequencies, which may suggest a dominant-negative phenotype.

These lines are also methylation tolerant, as they were selected for using 6-TG. Therefore, we conclude that the cell lines in this validation panel have all hallmarks of cellular MMR deficiency.

Incidental Lack of Correlation Between *In Vivo* and *In Vitro* Phenotypes

To directly prove causality of the mutated residues in Msh6 for loss of MMR, we used a wild-type human MSH6 cDNA as a template to recreate the mouse Msh6 mutants by site-directed mutagenic PCR in the human gene. This was followed by *in vitro* transcription and translation to create "humanized" MSH6 mutant proteins. These proteins were dimerized with wild type MSH2, and tested for their ability to restore MMR in extracts from the

MSH2/MSH6-deficient human cancer cell line LoVo [Drost et al., 2012]. Out of the 11 alleles tested, 8 mutant proteins were unable to complement the LoVo cell extract (Fig. 3C). This result provides direct and conclusive evidence for the causality of those ENU-induced amino acid substitutions in Msh6 for MMR deficiency. Remarkably, the D1211E, D1211G and H1246R alleles restored MMR to levels not significantly different from wild type. This may imply that these alleles confer MMR defects that are not detectable in the *in vitro* MMR assay, such as in protein stability or in nuclear localization.

Missense Substitutions In Msh6 Typically Impair Mismatch Binding

To gain insights into the mechanism of MMR loss in the missense mutant cell lines we have generated we investigated protein expression levels by Western blotting. This revealed that ~40% of all missense substitutions destabilize Msh6, leading to either a strong reduction or absence of Msh6 protein (not shown). This suggests that in these clones defective protein folding and/or stability is causal for MMR loss.

From the remaining clones that had retained (near) normal Msh6 expression levels, we tested the ability of these mutant proteins to perform the earliest step in MMR: mismatch binding. To do so, we tested the binding of cytoplasmic extracts to a G·T mismatched probe in an electrophoretic mobility shift assay. Of the 29 mutants tested, 21 lost the ability to bind to mismatched DNA (Figure 2A), suggesting that MMR deficiency in these clones is a result of either defective mismatch binding or of defective overall DNA binding. This included the earlier mentioned substitutions at residues F431, V507, V508 and R510 that contact either the mismatch or DNA backbone, but also residues not directly implicated in mismatch/DNA binding such as L770, M1135, G1155D and others (Figure 2A). This result suggests that for these missense mutant proteins, defective mismatch binding is the mechanistic basis for MMR deficiency.

Since cell extracts from mutant lines L310F, S1139P, D1211E, D1211G, T1217I and H1246R are proficient for mismatch binding we tested the ability of these cell extracts to perform a subsequent event in MMR: to form a sliding clamp after ATP binding. Sliding clamp formation leads to a loss of binding to heteroduplex DNA. Of the tested extracts, only T1217I showed a defect in ATP-induced mismatch release, suggesting that this allele has decreased capability of forming sliding clamps. Since L310F, S1139P, D1211E, D1211G and H1246R show normal sliding clamp formation, this suggests that these alleles do not have a (detectable) MMR defect, or a defect downstream of sliding clamp formation, such as in MutLα recruitment.

To investigate whether the mentioned alleles indeed have an MMR defect downstream of ATP binding, and to confirm the MMR defect in the other alleles, we tested a selection of these alleles as "humanized" *MSH6* alleles in the *in vitro* MMR assay described earlier. This experiment showed that the majority of alleles were unable to rescue the *in vitro* MMR defect of LoVo cell extracts, supporting their repair deficiency. In addition, like the previously described alleles D1211E, D1211G and H1246R, allele L310F rescued repair efficiency to wild type levels. As mentioned, this may imply that this allele confers MMR defects that are not detectable in the *in vitro* MMR assay, such as in protein stability or in nuclear localization.

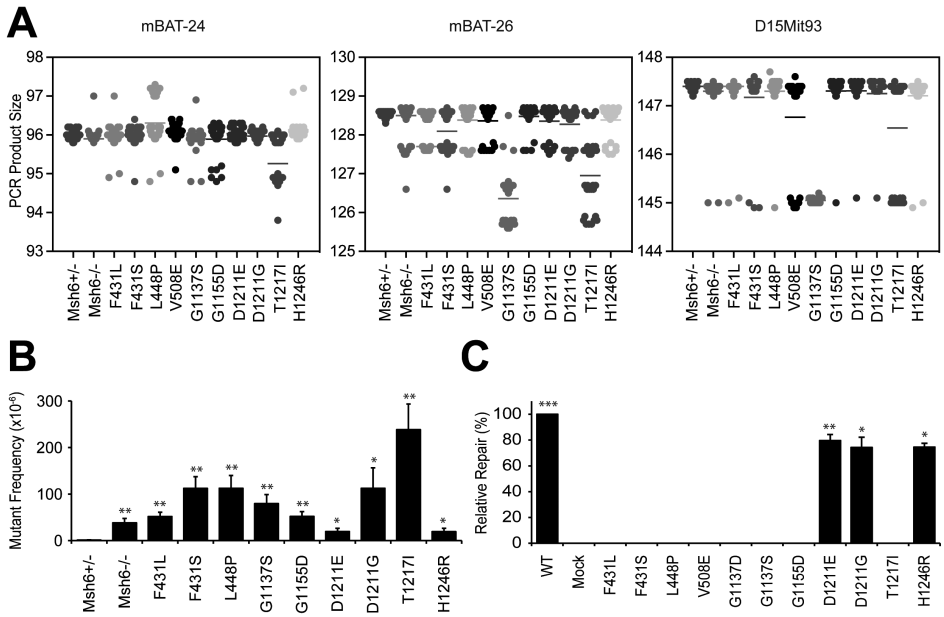


Figure 2. *Cell Lines From The Reverse Diagnosis Catalogs Shows Cellular Hallmarks Of MMR Deficiency. (A) Microsatellite instability analysis of Msh6 missense mutant cell lines. The PCR amplicon size of subclones is depicted for three different microsatellites. (B) Assessment of spontaneous mutator phenotypes at the Hprt gene of Msh6 missense mutant cell lines and controls. Bars represent mean Hprt mutant frequencies ± SEM. *p < 0.05, **p< 0.01, ***p<0.001 compared with the Msh6$^{+/-}$ line. (C) MMR activity of in vitro recreated mutant humanized MSH6 proteins. These proteins were tested for the repair of a G·T mismatch after addition to an MSH2/MSH6-deficient human cell extract. Mock: mock Msh2/Msh6 expression. Bars represent mean ± SEM. *p < 0.05, **p < 0.01, ***p < 0.001 compared with the mock reaction.*

DISCUSSION

Reverse Diagnosis Catalogues As Tools For MSH6 VUS Diagnosis

Assessing pathogenicity of VUS provides certainty to carriers. In contrast, relatives who do not carry the VUS are liberated from the burden of extensive screening procedures. Now that Next Generation Sequencing technologies are becoming more prevalent, also in clinical genetics, the prevalence of identified VUS is likely to increase [Heinen, 2010; Sikkema-Raddatz et al., 2013]. In this work, we describe a genetic screening procedure that identifies critical residues in Msh6, which aids in assessing pathogenicity to *MSH6* VUS. The data from such a screen may integrate in current diagnostics by becoming a part of diagnostic trees for MMR gene VUS [Kansikas et al., 2011; Rasmussen et al., 2012] and/or Bayesian analyses, which integrate various data sources for classifying VUS into a single model [Goldgar et al., 2008].

In our previous study, 10 out of the 23 (~44%) mutated residues we identified in the Msh2 protein were represented in the LOVD [Drost et al., 2013]. In this study, only 8 out of the 45 residues (~18%) in Msh6 are represented in the LOVD. This difference in LOVD representation may reflect an ascertainment bias in the LOVD. Due to redundancy with *MSH3*, inactivating

mutations in *MSH6* may be less penetrant in suspected LS patients and remain under detected. This is also a consequence of the lack of sensitivity of the Besthesda criteria, a set of criteria used to clinically diagnose LS patients, to identify *MSH6* mutation carriers [Sjursen et al., 2010].

Discrepancies Between Methylation Tolerance And MMR Status

Even though all cell lines generated in this study are 6-TG tolerant, we have been unable to show *in vitro* MMR deficiency for several alleles (Figure 3C). Remaining alleles will be tested for *in vitro* repair capacity in the near future. The discrepant results between *in vivo* and *in vitro* repair may have different explanations. Firstly, these alleles may infer *in vivo* MMR defects that are not detected in the *in vitro* assay, such as defects in nuclear localization. Therefore, the repair proficient variants in the *in vitro* assay may reflect false-negatives, strengthening the need for an *in vivo* assay such as the Reverse Diagnosis Catalogs. Indeed, residue L310 was previously implicated in nuclear localization [Gassman et al., 2011].

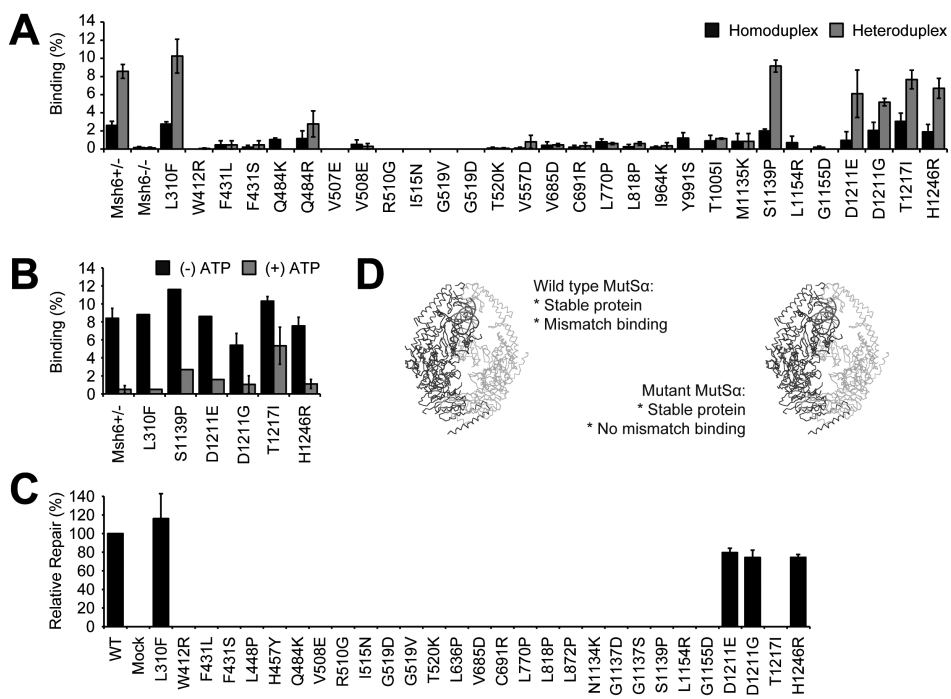


Figure 3. *Defective Mismatch Binding Is A Frequent Cause Of MMR Deficiency Binding of control and mutant Msh6 cell extracts to a G·T mismatch within a double-stranded oligonucleotide probe in an electrophoretic mobility shift assay. Bars represent mean±SEM. ATP-induced mismatch release of Msh6-mutant cell extracts was tested in an electrophoretic mobility shift assay. ATP (1.25 mM) was added after allowing the proteins to bind to the probe. Bars represent mean±SEM. MMR activity of in vitro recreated mutant humanized MSH6 proteins. These proteins were tested for the repair of a G·T mismatch after addition to an MSH2/MSH6-deficient human cell extract. Mock: mock Msh2/Msh6 expression. Bars represent mean ± SEM. *p < 0.05, **p < 0.01, ***p < 0.001 compared with the mock reaction.*

Secondly, these alleles may confer so-called separation-of-function phenotypes, uncoupling apoptosis from repair [Lin et al., 2004; Yang et al., 2004]. A third explanation is that there is only a minor, difficult to measure, drop in MMR efficiency that is sufficient to induce methylation tolerance but insufficient to abolish repair [Claij and te Riele, 2002]. In the latter two cases, the induction of methylation tolerance is still highly relevant for LS diagnosis, since this phenotype alone may predispose to cancer [Lin et al., 2004; Yang et al., 2004].

The Msh6-T1217I Allele Is Not A Separation-Of-Function Allele

The T1217D allele was previously described to be repair deficient, but methylation sensitive and was termed a separation-of-function allele, as it uncouples cell death from repair [Yang et al., 2004]. Since our screen selects for methylation tolerance the finding of a very similar allele, T1217I, is surprising. The T1217D allele is capable of mismatch binding, but incapable of recruiting MutLα [Geng et al., 2012]. This may be related to its weak ability to bind ATP and the subsequent inability to release mismatches and form a sliding clamp [Hess et al., 2006; Geng et al., 2012]. MSH2 or MSH6 alleles harboring ATP-binding defects often confer dominant-negative phenotypes [Das Gupta and Kolodner, 2000; Hess et al., 2002; Yang et al., 2004; Martín-López et al., 2012; Drost et al., 2013]. This is corroborated by cancer predisposition in *Msh6*$^{+/T1217D}$ mice, high microsatellite instability in tumors of *Msh6*$^{T1217D/T1217D}$ mice and inhibition of IDL repair in *in vitro* assays where cell extracts from wild type and *Msh6*$^{T1217D/T1217D}$ cells are mixed [Yang et al., 2004]. The T1217I allele from our screen may also have a dominant negative effect, as judged from high spontaneous mutation frequencies (Figure 2B), (possibly) increased MSI compared to *Msh6*$^{-/-}$ cells (Figure 2A) and a defect in ATP-induced mismatch release (Figure 2B). It is plausible that the persistence of Msh6-T1217D on DNA impedes cellular processes other than MMR, such as transcription and replication. If this is true, one would predict increased spontaneous apoptosis in cells carrying this allele compared to cells either wild type or deficient for *Msh6*, which indeed can be deduced from experimental data [Yang et al., 2004]. Therefore, the published apoptosis after treatment with methylating agents in the *Msh6*$^{T1217D/T1217D}$ cells may not reflect a true separation-of-function, but merely a consequence of decreased overall cellular fitness.

Inactivating Substitutions in Msh6 Cluster in the MDB and ATPase Domain But Are Absent in The N-Terminus

The inactivating substitutions identified in Msh6 are distributed across the protein with hotspots in the mismatch binding and ATPase domains, while mutants in the N-terminus seem underrepresented. These observations may reflect a bias in our screen, but may also represent genuine traits of Msh6. In support of the latter, in *S. cerevisiae*, the N-terminus of MSH6 is not required for nuclear localization, DNA binding, or mismatch discrimination [Shell et al., 2007b]. Thus, it is likely dispensable for methylation sensitivity, and substitutions in this area will not infer methylation tolerance. In addition, its structure is disordered, which likely makes it relatively tolerant to substitutions [Shell et al., 2007b]. Furthermore, its low importance is reflected by its poor evolutionary conservation, as prokaryotic MutS does not contain such an N-terminal tail [Lamers et al., 2000; Obmolova et al., 2000].

The abundancy of inactivating substitutions in the Msh6 MBD is in contrast with the pilot reverse diagnosis catalog we generated for *Msh2*, in which inactivating substitutions in the Msh2 MBD were virtually absent. These findings are supported by the notion that residues in MSH6, but not in MSH2, contact the mismatch [Bowers et al., 1999; Dufner et al., 2000; Warren et al., 2007] and that MSH6 is the monomer that determines substrate specificity [Shell et al., 2007a].

Mismatch Binding and Intramolecular Signaling

The binding of a mismatch by MutSα is one of the earliest events in the recognition of replication errors by DNA Mismatch Repair. As mentioned, previous studies have shown that Msh6 is the monomer that contacts the mismatch. Several well-characterized residues are involved either in directly binding the mismatch or in non-specific DNA contacts. Structural studies have shown that only a subset of residues are in direct contact with the mismatched DNA, especially between residues 400-470 [Warren et al., 2007]. We now show that an unexpectedly high proportion of residues throughout Msh6 are either directly or indirectly involved in mismatch binding (Figure 3c).

When keeping the crystal structure data in mind, the most plausible explanation for the involvement of residues distal from the MBD in mismatch binding is a role either structural, in conformational changes and/or in intramolecular signaling. Since the proteins we have tested are stable in Western blots, there are only minor structural changes expected. Such structural changes may influence mismatch binding in an indirect way. As an example, residues Y1003 and T1005 are directly in front of an long α helix in Msh6 that spans between the clamp and ATPase domains. Mutating these residues may hinder the movement of these domains. In this light, conformational changes in MutSα may not only be important after mismatch binding, but also during or before mismatch binding. Further research is required to elucidate the exact effects of these substitutions on the biochemistry of MMR.

ACKNOWLEDGEMENTS.

We thank Joyce H. G. Lebbink (Erasmus Medical Center) and Meindert H. Lamers (Medical Research Council) for help with crystal structure modelling.

MATERIALS & METHODS

Generation of Reverse Diagnosis Catalogs for Msh6

Reverse diagnosis catalogs were generated as described [Drost et al., 2013]. In brief, $5x10^6$ *Msh6*$^{+/-}$ cells were cultured on irradiated feeder cells in Knockout-DMEM (Gibco). These cells were treated with 0,2 mg/ml ENU (Sigma-Aldrich) for 2h, or mock treated, and subsequently cultured for a week. Then, $2x10^6$ cells were seeded per T90 followed by dual brief 40 μM 6-TG treatments (Sigma-Aldrich) for 4h, in a 7-day interval. Cells were culture in hypoxanthine-aminopterin-thymidine (HAT)-supplement containing medium (Gibco) to select against *Hprt* mutants, and surviving clones were isolated and expanded. These clones were screened

against LOH by an allele specific PCR for MSH6. The full MSH6 ORF of remaining clones expressing full length cDNA was sequenced after RT-PCR and PCR amplification.

Crystal Structure Modeling

Crystal structures were visualized and edited using PyMOL software (Delano Scientific), using the Protein Database file from the MutSa crystal structure (Molecular Modeling Database ID #53938) [Warren et al., 2007].

Determination of Genomic Mutator Phenotypes

Mutant frequencies at the *Hprt* reporter gene were measured as described [Drost et al., 2013]. Briefly, cells were cultured in HAT-containing medium for two passages to eliminate pre-existing *Hprt*-mutants. Cells were then grown in the absence of HAT for three passages to allow spontaneous mutagenesis. *Hprt* mutants were subsequently selected for by a continuous 60 μM 6-TG treatment. 6-TG resistant clones were stained and scored after 14 days, and counts were corrected for cloning efficiencies.

Microsatellite instability was determined by isolating ~50 subclones of each line. Genomic DNA of these subclones was isolated by Proteinase K lysis, and microsatellite sizes were measured after PCR amplification and fragment analysis as described [Bacher et al., 2005]. Data were analyzed using GeneMarker software (Softgenetics).

In Vitro MMR Assays, Western blotting and Electrophoretic Mobility Shift Assays.

All biochemical procedures were carried out as described [Drost et al., 2013].

REFERENCES

Antony E, Hingorani MM. 2003. Mismatch recognition-coupled stabilization of Msh2-Msh6 in an ATP-bound state at the initiation of DNA repair. Biochemistry 42: 7682-7693.

Bacher JW, Abdel Megid WM, Kent-First MG, Halberg RB. 2005. Use of mononucleotide repeat markers for detection of microsatellite instability in mouse tumors. Mol Carcinog 44: 285-292.

Boland CR, Goel A. 2010. Microsatellite instability in colorectal cancer. Gastroenterology 138: 2073-2087.e3.

Boland CR, Lynch HT. 2013. The History of Lynch Syndrome. Fam Cancer , *in press*

Borgdorff V, van Hees-Stuivenberg S, Meijers CM, de Wind N. 2005. Spontaneous and mutagen-induced loss of DNA mismatch repair in Msh2-heterozygous mammalian cells. Mutat Res 574: 50-57.

Bowers J, Sokolsky T, Quach T, Alani E. 1999. A mutation in the MSH6 subunit of the Saccharomyces cerevisiae MSH2-MSH6 complex disrupts mismatch recognition. J Biol Chem 274: 16115-16125.

Burn J, Gerdes AM, Macrae F, Mecklin JP, Moeslein G, Olschwang S, Eccles D, Evans DG, Maher ER, Bertario L, Bisgaard ML, Dunlop MG et al. 2011. Long-term effect of aspirin on cancer risk in carriers of hereditary colorectal cancer: an analysis from the CAPP2 randomised controlled trial. Lancet 378: 2081-2087.

Claij N, te Riele H. 2002. Methylation tolerance in mismatch repair proficient cells with low MSH2 protein level. Oncogene 21: 2873-2879.

Constantin N, Dzantiev L, Kadyrov FA, Modrich PL. 2005. Human mismatch repair: reconstitution of a nick-directed bidirectional reaction. J Biol Chem 280: 39752-39761.

Cyr JL, Heinen CD. 2008. Hereditary cancer-associated missense mutations in hMSH6 uncouple ATP hydrolysis from DNA mismatch binding. J Biol Chem 283: 31641-31648.

Das Gupta R, Kolodner RD. 2000. Novel dominant mutations in Saccharomyces cerevisiae MSH6. Nat Genet 24: 53-56.

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

de Wind N, Dekker M, Claij N, Jansen L, van Klink Y, Radman M, Riggins G, van der Valk M, van't Wout K, te Riele H. 1999. HNPCC-like cancer predisposition in mice through simultaneous loss of Msh3 and Msh6 mismatch-repair protein functions. Nat Genet 23: 359-362.

Drost M, Lutzen A, Van Hees S, Ferreira D, Calleja F, Zonneveld J, Nielsen F, Rasmussen L, de Wind N. 2013. Genetic screens to identify pathogenic gene variants in the common cancer predisposition Lynch syndrome. Proceedings of the National Academy of Sciences 110: 9403-9408.

Drost M, Zonneveld JB, van Hees S, Rasmussen LJ, Hofstra RM, De Wind N. 2012. A rapid and cell-free assay to test the activity of lynch syndrome-associated MSH2 and MSH6 missense variants. Hum Mutat 33: 488-494.

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. J Biol Chem 275: 36550-36555.

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

Fokkema IF, Taschner PE, Schaafsma GC, Celli J, Laros JF, den Dunnen JT. 2011. LOVD v.2.0: the next generation in gene variant databases. Hum Mutat 32: 557-563.

Gassman NR, Clodfelter JE, McCauley AK, Bonin K, Salsbury FR, Scarpinato KD. 2011. Cooperative nuclear localization sequences lend a novel role to the N-terminal region of MSH6. PLoS ONE 6: e17907.

Geng H, Sakato M, DeRocco V, Yamane K, Du C, Erie DA, Hingorani M, Hsieh P. 2012. Biochemical analysis of the human mismatch repair proteins hMutSα MSH2(G674A)-MSH6 and MSH2-MSH6(T1219D). J Biol Chem 287: 9777-9791.

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

Gorman J, Chowdhury A, Surtees JA, Shimada J, Reichman DR, Alani E, Greene EC. 2007. Dynamic basis for one-dimensional DNA scanning by the mismatch repair complex Msh2-Msh6. Mol Cell 28: 359-370.

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. Mol Cell 3: 255-261.

Hays JB, Hoffman PD, Wang H. 2005. Discrimination and versatility in mismatch repair. DNA Repair 4: 1463-1474.

Heinen CD. 2010. Genotype to phenotype: analyzing the effects of inherited mutations in colorectal cancer families. Mutat Res 693: 32-45.

Hess MT, Gupta RD, Kolodner RD. 2002. Dominant Saccharomyces cerevisiae msh6 mutations cause increased mispair binding and decreased dissociation from mispairs by Msh2-Msh6 in the presence of ATP. J Biol Chem 277: 25545-25553.

Hess MT, Mendillo ML, Mazur DJ, Kolodner RD. 2006. Biochemical basis for dominant mutations in the Saccharomyces cerevisiae MSH6 gene. Proc Natl Acad Sci USA 103: 558-563.

Hewish M, Lord CJ, Martin SA, Cunningham D, Ashworth A. 2010. Mismatch repair deficient colorectal cancer in the era of personalized treatment. Nat Rev Clin Oncol 7: 197-208.

Hsieh P, Yamane K. 2008. DNA mismatch repair: molecular mechanism, cancer, and ageing. Mech Ageing Dev 129: 391-407.

Jansen JG, Mohn GR, Vrieling H, van Teijlingen CM, Lohman PH, van Zeeland AA. 1994. Molecular analysis of hprt gene mutations in skin fibroblasts of rats exposed in vivo to N-methyl-N-nitrosourea or N-ethyl-N-nitrosourea. Cancer Res 54: 2478-2485.

Jiricny J. 2013. Postreplicative mismatch repair. Cold Spring Harbor Perspectives in Biology 5:a012633

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

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

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.

Lin DP, Wang Y, Scherer SJ, Clark AB, Yang K, Avdievich E, Jin B, Werling U, Parris T, Kurihara N, Umar A, Kucherlapati R et al. 2004. An Msh2 point mutation uncouples DNA mismatch repair and apoptosis. Cancer Res 64: 517-522.

Martín-López JV, Barrios Y, Medina-Arana V, Andújar M, Lee S, Gu L, Li GM, Rüschoff J, Salido E, Fishel R. 2012. The hMSH2(M688R) Lynch syndrome mutation may function as a dominant negative. Carcinogenesis 33: 1647-1654.

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.

Plazzer JP, Sijmons RH, Woods MO, Peltomäki P, Thompson B, Den Dunnen JT, Macrae F. 2013. The InSiGHT database: utilizing 100 years of insights into Lynch Syndrome. Fam Cancer 12: 175-180.

Rasmussen LJ, Heinen CD, Royer-Pokora B, Drost M, Tavtigian S, Hofstra RM, De Wind N. 2012. Pathological assessment of mismatch repair gene variants in Lynch syndrome: Past, present, and future. Hum Mutat 33: 1617-1625.

Shell SS, Putnam CD, Kolodner RD. 2007a. Chimeric Saccharomyces cerevisiae Msh6 protein with an Msh3 mispair-binding domain combines properties of both proteins. Proc Natl Acad Sci USA 104: 10956-10961.

Shell SS, Putnam CD, Kolodner RD. 2007b. The N terminus of Saccharomyces cerevisiae Msh6 is an unstructured tether to PCNA. Mol Cell 26: 565-578.

Sijmons RH, Greenblatt MS, Genuardi M. 2013. Gene variants of unknown clinical significance in Lynch syndrome. An introduction for clinicians. Fam Cancer 12: 181-187.

Sikkema-Raddatz B, Johansson LF, de Boer EN, Almomani R, Boven LG, van den Berg MP, van Spaendonck-Zwarts KY, van Tintelen JP, Sijmons RH, Jongbloed JD, Sinke RJ. 2013. Targeted Next-Generation Sequencing can Replace Sanger Sequencing in Clinical Diagnostics. Hum Mutat 34: 1035-1042.

Sjursen W, Haukanes BI, Grindedal EM, Aarset H, Stormorken A, Engebretsen LF, Jonsrud C, Bjørnevoll I, Andresen PA, Ariansen S, Lavik LA, Gilde B et al. 2010. Current clinical criteria for Lynch syndrome are not sensitive enough to identify MSH6 mutation carriers. J Med Genet 47: 579-585.

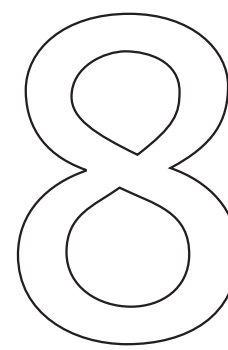
Swann PF, Waters TR, Moulton DC, Xu YZ, Zheng Q, Edwards M, Mace R. 1996. Role of postreplicative DNA mismatch repair in the cytotoxic action of thioguanine. Science 273: 1109-1111.

Vasen HF, Blanco I, Aktan-Collan K, Gopie JP, Alonso A, Aretz S, Bernstein I, Bertario L, Burn J, Capella G, Colas C, Engel C et al. 2013. Revised guidelines for the clinical management of Lynch syndrome (HNPCC): recommendations by a group of European experts. Gut 62: 812-823.

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

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

General Discussion and Future Perspectives



Variants of Uncertain Significance (VUS) pose a major problem not only in Lynch syndrome diagnostics, but in clinical genetics in general [Spurdle, 2010; Rasmussen et al., 2012; Sijmons et al., 2013]. A lack of diagnostic assessment of VUS leaves carriers and both affected and unaffected relatives in uncertainty regarding their disease status. This, combined with the surveillance procedures that such patients and all relatives (both carriers and non-carriers) are enrolled in pose a physical and psychological burden on families, but also a burden on the preventive health care apparatus.

The current record of variants in DNA mismatch repair genes is a reflection of sequencing efforts from the last few decades [Fokkema et al., 2011; Plazzer et al., 2013]. From a diagnostic point-of-view, LS families with many disease cases were probably genetically tested first. This likely has resulted in an ascertainment bias where mainly mutations that completely inactivate MMR and are highly penetrant have been identified, such as nonsense of frameshift mutations. The advent of Next Generation Sequencing technologies has made genetic testing easier, cheaper and faster. This will likely lower the criteria for sequencing, which will further increase the incidence of VUS [Heinen, 2010; Rasmussen et al., 2012; Sikkema-Raddatz et al., 2013]. For these reasons, the development and use of robust and accessible approaches to diagnose pathogenic VUS is becoming increasingly important. This will enable the implementation of preventive and curative healthcare that is targeted only to carriers of pathogenic MMR gene VUS. Meanwhile, unaffected relatives of such carriers can be liberated from the burden associated with the uncertain pathogenicity of the VUS while unwarranted inflow into preventive healthcare systems is reduced.

Integration Of Novel Assays For Functional Analysis In Current Lynch Syndrome Diagnostics

Current LS diagnosis protocols are aimed at determining the inactivating mutation in a DNA MMR gene through genetic testing (Chapters 1 and 2). Once such an inactivating mutation has been found, LS diagnosis is completed. In case a VUS is found, LS diagnosis is halted. To assist geneticists in subsequent diagnostics, we and others have proposed additional steps to the current diagnostic tree (discussed in Chapter 2; [Couch et al., 2008]). We have proposed that, once a VUS has been found, the allele is first tested in an *in vitro* MMR assay (Chapters 3-5) and is simultaneously evaluated using *in silico* analyses [Tavtigian et al., 2008; Thompson et al., 2012b]. As the *in vitro* MMR assay covers (nearly) all aspects of canonical MMR and is relatively quick, this assays seems the most feasible approach. Furthermore, owing to the relative simplicity of this assay it may be implemented in multiple labs globally. Based on preliminary data the approach proposed above appears to be a powerful one for the diagnostic assessment of MMR gene VUS [Kansikas et al., 2011].

To integrate the *in vitro* assay in current LS diagnostics, its output may be integrated in a five-class classification system [Plon et al., 2008]. These classes can replace poorly defined terms such as "VUS", "neutral" and "pathogenic" and may aid clinical management by defining specific recommendations for each category. Those five classes are as follows:

- Class 5 variant - Pathogenic, > 99% probability of pathogenicity
- Class 4 variant - Likely Pathogenic, 95-99% probability of pathogenicity
- Class 3 variant - Uncertain, 5-95% probability of pathogenicity
- Class 2 variant - Likely Neutral, 0.1-5% probability of pathogenicity
- Class 1 variant – Neutral < 0.1% probability of pathogenicity

A five-class classification system can be established by combining multiple criteria for assessing pathogenicity to VUS, including well-established ones such as family history, immunohistrochemistry and microsatellite instability (MSI), but also *in silico* assays [Tavtigian et al., 2008] and the *in vitro* MMR assay [Drost et al., 2010, 2012]. These data sets are then integrated into a single statistical model [Plon et al., 2008]. For the assays integration and validation, the assay should be evaluated prosprectively by large-scale validation efforts that include performing the assay on a large number of substitutions, classified in either class 1 or class 5, without the use of functional assays. Such a validation effort should aid in determining positive and negative predictive values and Likelyhood Ratios, which will eventually help to determine a posterior probability of pathogenicity [Thompson et al., 2012a, 2012b].

The *in vitro* MMR assay in its current status has a binary outcome. MMR variants can be either repair deficient (not significantly different from a repair deficient control) or repair proficient (significantly higher than a repair deficient control). When integrating the *in vitro* MMR assays' output into a five-class classification system, its readout (percentage of repair compared to wild type) could be calibrated to be translated into a probability in favor of pathogenicity and may serve as a constant variable.

The *in vitro* MMR assay could be a powerful tool to diagnose VUS in MMR genes, however, the *in vitro* nature of the assays has its pitfalls. The assay is not able to pick up certain *in vivo* defects and therefore, in case a VUS is repair proficient *in vitro* the diagnostic value of the *in vitro* MMR assay is inconclusive and more specialized assays, such as protein stability [Perera and Bapat, 2008], nuclear localization assays [Raevaara et al., 2005; Andersen et al., 2012; Borràs et al., 2013] and splicing assays [Tournier et al., 2008] are required to assess pathogenicity [Rasmussen et al., 2012]. In this regard, the Reverse Diagnosis Catalogs (RDC) may contribute to diagnostic assessment. The *a priori* knowledge of essential residues in MMR proteins may be valuable information in identifying pathogenic missense mutations. In case a human VUS is identified in the RDC, this would imply a high probability of pathogenicity of that VUS. Once Likelyhood Ratios for the RDC are determined and the RDC is incorporated in a five-class classification system, the identification of a certain substitution in the RDC should serve as strong evidence to classify a VUS as class 4 or 5.

To integrate both the *in vitro* assay and the Reverse Diagnosis Catalogs (RDC), these assays could be cross-validated. The RDC may then serve as a "golden standard" of pathogenic alleles, to which the *in vitro* assay may be calibrated against. In this manner, the RDC are a substitute for patient-derived inactivating alleles. Taking this one step further, the cell lines made in the RDC screens could be used to generate mice to determine penetrance of the alleles, which then can be linked to *in vitro* repair efficiencies.

To continue along this line, the RDCs may serve as a standard not only to validate the *in vitro* assay, but also to validate *in silico* algorithms such as MAPP-MMR [Chao et al., 2008], PolyPhen [Adzhubei et al., 2010] or the combination of both [Thompson et al., 2012b]. Once the RDC protocol is adapted to also identify disrupting intronic variants or splice site mutations (see below), the data from the RDC may also aid in validating *in silico* splice site prediction software [Vreeswijk et al., 2009; Vreeswijk and van der Klift, 2012].

Towards Comprehensive Reverse Diagnosis Catalogs

This thesis describes the generation of pilot RDCs for Msh2 (Chapter 6) and Msh6 (Chapter 7). In order to generate saturated, comprehensive catalogs the methodology described in this thesis may be adapted in several ways.

First, mutations are introduced into the MMR genes using the mutagen N-ethyl-N-nitrosourea (ENU). Even though ENU is a relatively wide-spectrum mutagen, using a single chemical for mutagenesis may not suffice to saturate a RDC. Higher saturation of the catalog could be reached by using additional drugs with a complementary mutagenic activity, such as 4-nitroquinoline *N*-oxide [Ryu et al., 1999] or 7,12-dimethylbenz[a]anthracene [Manjanatha et al., 1996]. ENU is not only a suited mutagen due to its diverse mutagenic activity, but also due to its property to induce only low amounts of loss of heterozygosity (LOH). Using other drugs may induce a higher proportion of LOH events, which will decrease the efficiency of the screen. Such an issue may be circumvented by putting selective pressure against LOH, by targeting a selectable marker directly next to the non-disrupted *Msh2* and *Msh6* alleles and using the generated cell lines for RDC development.

Second, the current protocol selects for MMR deficiency using 6-thioguanine (6-TG). Our experimental data suggests that the selection conditions applied may have been quite harsh, as only fully (*Msh2*$^{-/-}$-like) tolerance of MNNG was measured in all generated cell lines (Chapter 6). In future work, selection conditions could be further optimized. By applying a less stringent selection the protocol could be adjusted as such, so that alleles with intermediate levels of methylation tolerance can be selected for. In this way, the RDC could be extended from completely methylation tolerant, probably highly penetrant alleles, towards less penetrant alleles. In this respect, 6-TG could also be replaced by MNNG as a selection agent. Other have shown that cells tolerant to MNNG can still be proficient in repair [Claij and te Riele, 2002]. This could make MNNG a more appropriate drug to select for less penetrant alleles.

Finally, the RDC presented in this thesis are aimed completely at disruptive missense mutations as only those 6-TG-tolerant clones that retain expression of full-length cDNAs were sequenced. From a clinical geneticists' point-of-view, additional types of VUS are of interest, such as intronic variants or splice site variants. To identify these, one would have to apply methods other than cDNA sequencing, such as sequencing of genomic DNA using a classical PCR and Sanger sequencing approach, or to use Next Generation Sequencing methods.

Functional Assays in Fundamental Research

The assays we have developed will not only allow us to assess pathogenicity of VUS in MMR genes, but can also aid in functional studies of MMR. Especially the combination of both the *in vitro* assay and the RDC may help to identify novel functional residues/motifs in MMR proteins that are important for *in vivo* MMR.

A very interesting allele that shows (relevant) discrepancies between *in vivo* and *in vitro* phenotypes was recently described by Li and co-workers [Li et al., 2013]. The PWWP domain in MSH6 helps to recruit MutSα onto chromatin by interacting with the histone mark H3K56me3. Knockdown of SETD2, the transferase that methylates H3K56, gives an MSI phenotype suggesting that SETD2 and H3K56me3 are upstream of MMR *in vivo*. *MSH6* alleles with mutations in the PWWP motif are MMR deficient *in vivo*, but MMR proficient in the *in vitro* MMR assay, which might be explained by the difference in chromatin states of the DNA substrates. The identification and subsequent characterization of this type of alleles may provide new insights into the biochemistry of MMR. Such alleles may be identified in the RDC, by using a readout for *in vivo* MMR and, in parallel, testing those alleles in the *in vitro* MMR assay. The Msh6 RDC may have already identified such alleles. Msh6 alleles D1211E, D1211G and H1246R are repair proficient in the *in vitro* assay but show hallmarks of *in vivo* MMR deficiency (Chapter 7). These are interesting alleles for future research.

In Chapters 6 and 7 several methods are described for measuring MMR *in vivo*, such as MSI assessment and determining spontaneous mutation frequencies at a reporter gene. An alternative approach would be to introduce a fluorescent reporter such as a GFP, rendered out-of-frame due to a microsatellite preceding its ORF. In MMR deficient cells, this microsatellite will be unstable, rendering *GFP* into frame. Therefore, the amount of GFP proficient cells in a cell population will be a measure of the MMR capacity of the MMR allele(s) expressed in those cells [Koole et al., 2013].

Methylation Tolerance: Futile Cycling Or Direct Signaling?

The mechanism of the MMR-mediated toxicity of methylating agents has been under debate for many years now. The MMR field has been divided into two "camps": One supporting the futile cycling model [Stojic et al., 2004; Mojas et al., 2007] and one supporting the direct signaling model [Yang et al., 2004; Yoshioka et al., 2006]. Even though both models are not mutually exclusive, several arguments are in favor of the futile cycling model. First, apoptosis after methylating agents is induced in the second S-phase after treatment, which suggests that a (DNA) intermediate from the first cell cycle is carried over to the second and only then induces apoptosis [Mojas et al., 2007]. Second, proteins such as ATM [Debiak et al., 2004], RAD51 and BRCA1 protect against apoptosis induced by methylating agents, which suggests that a DNA double-strand break is involved, which would be unexpected with the direct signaling model in mind. Third, methylation-induced apoptosis requires a full complement of MLH1 [Cejka et al., 2003] and the catalytic activity of EXO1 [Schaetzlein et al., 2013]. This suggests that a canonical MMR complex is required to induce apoptosis.

The major pieces of evidence in favor of the direct signaling model are 1) Binding of ATR, TopBP1 and CHK1, but not RPA, to chromatin after MNNG treatment in a MutSα and MutLα dependent manner [Liu et al., 2009], 2) Protein interactions between MutSα and ATR [Liu et al., 2009; Pabla et al., 2011], TopBP1 and CHK1, and between MutLα and TopBP1 [Liu et al., 2009], 3) The phosphorylation of CHK1 in the presence of O^6-meG·T mismatches and MMR proteins *in vitro* in the absence of excision [Yoshioka et al., 2006] and 4) Separation-of-function alleles [Lin et al., 2004; Yang et al., 2004]. The relevance of the described protein interactions and chromatin loading briefly after MNNG is questionable, as ATR is only activated at later timepoints [Stojic et al., 2004] and, even if ATR would be activated at earlier timepoints, there is insufficient signaling to trigger apoptosis [Mojas et al., 2007]. The *in vitro* phosphorylation of CHK1 in the presence of O^6-meG·T mismatches and MMR proteins, but in the absence of excision, is prone to artefacts. The absence of excision is measured in an *in vitro* MMR assay which is quantified on an agarose gel [Yoshioka et al., 2006]. In our experience, quantification on gel is not very sensitive and repair percentages of up to 10% can easily stay underdetected, which may suffice to trigger signaling *in vitro*. Therefore, the phosphorylation of CHK1 may still result from single-stranded DNA and a more sensitive method would be required to rule out excision.

Perhaps the strongest evidence for the direct signaling model is the generation and characterization of two separation-of-function alleles in *Msh2* [Lin et al., 2004] and *Msh6* [Yang et al., 2004]. Both of these alleles confer a phenotype that is characterized by repair deficiency, but methylation sensitivity. It should be noted, though, that both described cell lines show high levels of spontaneous apoptosis. This could be a consequence of the ability of these MMR alleles to bind DNA, but their inability to subsequently release the heteroduplex or the inadvertently bound homoduplex. This may lead to a dominant-negative phenotype, in which the persistently DNA-bound MMR alleles impede other cellular processes such as transcription and replication. Therefore, the proficiency to induce apoptosis may not reflect a true separation-of-function phenotype, but a consequence of decreased cellular fitness. An interesting piece of evidence in favor of the direct signaling model would be an allele that has the opposite phenotype: repair proficiency, but methylation tolerance. The RDC protocol would be a perfect approach to find such an allele, as the screen selects for methylation tolerance and, once a GFP reporter is introduced, allows for a high-throughput screening for repair proficiency. The pilot Msh2 and Msh6 RDCs that we generated (Chapters 6 and 7) argues against the existence of such alleles, as we have only found alleles that are repair deficient *in vivo* as well as methylation tolerant. Conversely, this could be a consequence of the limited numbers of alleles identified so far, and increasing the scale of the RDC may suffice to identify such alleles.

Molecular And Clinical Testing And The Future Of Clinical Genomics

Next-generation sequencing technologies are becoming the primary tool in human genetics [Goldstein et. al, 2013; Katsanis et. al, 2013]. These technologies are reaching the point of being able to detect genetic variations at a high accuracy and reduced cost, which may alter

(personalized) medicine. The magnitude of data could lead to weakly justified claims of causality between genetic variants and disease [Goldstein et. al, 2013]. In this thesis, multiple methods are described that allow to determine causality of DNA variants for the cancer predisposition Lynch syndrome. Such methods may be instrumental in facilitating diagnosis in the upcoming age of personalized medicine, aiding in the future of clinical molecular diagnostics.

REFERENCES

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

Andersen SD, Liberti SE, Lützen A, Drost M, Bernstein I, Nilbert M, Dominguez M, Nyström M, Hansen TV, Christoffersen JW, Jäger AC, De Wind N et al. 2012. Functional characterization of MLH1 missense variants identified in lynch syndrome patients. Hum Mutat 33: 1647-1655.

Borràs E, Pineda M, Cadiñanos J, Del Valle J, Brieger A, Hinrichsen I, Cabanillas R, Navarro M, Brunet J, Sanjuan X, Musulen E, van der Klift H et al. 2013. Refining the role of pms2 in Lynch syndrome: germline mutational analysis improved by comprehensive assessment of variants. J Med Genet 50: 552-563

Cejka P, Stojic L, Mojas N, Russell AM. 2003. Methylation-induced G2/M arrest requires a full complement of the mismatch repair protein hMLH1. EMBO J 22: 2245-2254

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). Hum Mutat 29: 852-860.

Claij N, te Riele H. 2002. Methylation tolerance in mismatch repair proficient cells with low MSH2 protein level. Oncogene 21: 2873-2879.

Couch FJ, Rasmussen LJ, Hofstra R, Monteiro AN, Greenblatt MS, De Wind N, Group IARCUGVW. 2008. Assessment of functional effects of unclassified genetic variants. Hum Mutat 29: 1314-1326.

Debiak M, Nikolova T, Kaina B. 2004. Loss of ATM sensitizes against O6-methylguanine triggered apoptosis, SCEs and chromosomal aberrations. DNA Repair 3: 359-368.

Drost M, Zonneveld JB, Van Dijk L, Morreau H, Tops CM, Vasen HFA, Wijnen J, De Wind N. 2010. A cell-free assay for the functional analysis of variants of the mismatch repair protein MLH1. Hum Mutat 31: 247-253.

Drost M, Zonneveld JB, van Hees S, Rasmussen LJ, Hofstra RM, De Wind N. 2012. A rapid and cell-free assay to test the activity of lynch syndrome-associated MSH2 and MSH6 missense variants. Hum Mutat 33: 488-494.

Fokkema IF, Taschner PE, Schaafsma GC, Celli J, Laros JF, den Dunnen JT. 2011. LOVD v.2.0: the next generation in gene variant databases. Hum Mutat 32: 557-563.

Goldstein DB, Allen A, Keebler J, Margulies EH, Petrou S, Petrovski S, Sunyaev S. 2013. Sequencing studies in human genetics: design and interpretation. Nat Rev Genet. 14: 460-470.

Heinen CD. 2010. Genotype to phenotype: analyzing the effects of inherited mutations in colorectal cancer families. Mutat Res 693: 32-45.

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

Katsanis SH, Katsanis N. 2013. Molecular genetic testing and the future of clinical genomics. Nat Rev Genet. 14: 415-426

Koole W, Schäfer HS, Agami R, van Haaften G, Tijsterman M. 2013. A versatile microsatellite instability reporter system in human cells. Nucleic Acids Res *in press*.

Li F, Mao G, Tong D, Huang J, Gu L, Yang W, Li GM. 2013. The Histone Mark H3K36me3 Regulates Human DNA Mismatch Repair through Its Interaction with MutSα. Cell 153: 590-600.

Lin DP, Wang Y, Scherer SJ, Clark AB, Yang K, Avdievich E, Jin B, Werling U, Parris T, Kurihara N, Umar A, Kucherlapati R et al. 2004. An Msh2 point mutation uncouples DNA mismatch repair and apoptosis. Cancer Res 64: 517-522.

Liu Y, Fang Y, Shao H, Lindsey-Boltz L, Sancar A, Modrich P. 2009. Interactions of human mismatch repair proteins MutS and MutL with proteins of the ATR-Chk1 pathway. J Biol Chem 285: 5974-5982

Manjanatha MG, Chen JB, Shaddock JG, Harris AJ, Shelton SD, Casciano DA. 1996. Molecular analysis of lacI mutations in Rat2 cells exposed to 7,12-dimethylbenz[a]anthracene: evidence for DNA sequence and DNA strand biases for mutation. Mutat Res 372: 53-64.

Mojas N, Lopes M, Jiricny J. 2007. Mismatch repair-dependent processing of methylation damage gives rise to persistent single-stranded gaps in newly replicated DNA. Genes Dev 21: 3342-3355.

Pabla N, Ma Z, Mcilhatton M, Fishel R, Dong Z. 2011. hMSH2 Recruits ATR to DNA Damage Sites for Activation during DNA Damage-induced Apoptosis. J Biol Chem 286: 10411-10418.

Perera S, Bapat B. 2008. The MLH1 variants p.Arg265Cys and p.Lys618Ala affect protein stability while p.Leu749Gln affects heterodimer formation. Hum Mutat 29: 332

Plazzer JP, Sijmons RH, Woods MO, Peltomäki P, Thompson B, Den Dunnen JT, Macrae F. 2013. The InSiGHT database: utilizing 100 years of insights into Lynch Syndrome. Fam Cancer 12: 175-180.

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

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 et al. 2005. Functional significance and clinical phenotype of nontruncating mismatch repair variants of MLH1. Gastroenterology 129: 537-549.

Rasmussen LJ, Heinen CD, Royer-Pokora B, Drost M, Tavtigian S, Hofstra RM, De Wind N. 2012. Pathological assessment of mismatch repair gene variants in Lynch syndrome: Past, present, and future. Hum Mutat 33: 1617-1625.

Ryu J, Youn J, Kim YH, Kwon O, Song Y, Kim H, Cho K, Chang I. 1999. Mutation spectrum of 4-nitroquinoline N-oxide in the lacI transgenic Big Blue Rat2 cell line. Mutat Res 445: 127-135.

Schaetzlein S, Chahwan R, Avdievich E, Roa S, Wei K, Eoff RL, Sellers RS, Clark AB, Kunkel TA, Scharff MD, Edelmann W. 2013. Mammalian Exo1 encodes both structural and catalytic functions that play distinct roles in essential biological processes. Proc Natl Acad Sci USA 110: 2470-2479.

Sijmons RH, Greenblatt MS, Genuardi M. 2013. Gene variants of unknown clinical significance in Lynch syndrome. An introduction for clinicians. Fam Cancer *in press*.

Sikkema-Raddatz B, Johansson LF, de Boer EN, Almomani R, Boven LG, van den Berg MP, van Spaendonck-Zwarts KY, van Tintelen JP, Sijmons RH, Jongbloed JD, Sinke RJ. 2013. Targeted Next-Generation Sequencing can Replace Sanger Sequencing in Clinical Diagnostics. Hum Mutat 34: 1035-1042.

Spurdle AB. 2010. Clinical relevance of rare germline sequence variants in cancer genes: evolution and application of classification models. Curr Opin Genet Dev .

Stojic L, Mojas N, Cejka P, di Pietro M, Ferrari S, Jiricny J. 2004. Mismatch repair-dependent G2 checkpoint induced by low doses of SN1 type methylating agents requires the ATR kinase. Genes Dev 18: 1331-1344.

Tavtigian SV, Greenblatt MS, Lesueur F, Byrnes GB, Group IARCUGVW. 2008. In silico analysis of missense substitutions using sequence-alignment based methods. Hum Mutat 29: 1327-1336.

Thompson B, Goldgar D, Paterson C, Clendenning M, Walters R, Arnold S, Parsons M, Michael D W, Gallinger S, Haile R, Hopper J, Jenkins M et al. 2012a. A Multifactorial Likelihood Model for MMR Gene Variant Classification Incorporating Probabilities Based on Sequence Bioinformatics and Tumor Characteristics: A Report from the Colon Cancer Family Registry. Hum Mutat 34: 200-209.

Thompson B, Greenblatt M, Vallee M, Herkert J, Tessereau C, Young E, Adzhubey I, Li B, Bell R, Feng B, Mooney S, Radivojac P et al. 2012b. Calibration of Multiple In Silico Tools for Predicting Pathogenicity of Mismatch Repair Gene Missense Substitutions. Hum Mutat 34: 255-265.

Tournier I, Vezain M, Martins A, Charbonnier F, Baert-Desurmont S, Olschwang S, Wang Q, Buisine MP, Soret J, Tazi J, Frébourg T, Tosi M. 2008. A large fraction of unclassified variants of the mismatch repair genes MLH1 and MSH2 is associated with splicing defects. Hum Mutat 29: 1412-1424.

Vreeswijk MP, Kraan JN, van der Klift HM, Vink GR, Cornelisse CJ, Wijnen JT, Bakker E, van Asperen CJ, Devilee P. 2009. Intronic variants in BRCA1 and BRCA2 that affect RNA splicing can be reliably selected by splice-site prediction programs. Hum Mutat 30: 107-114.

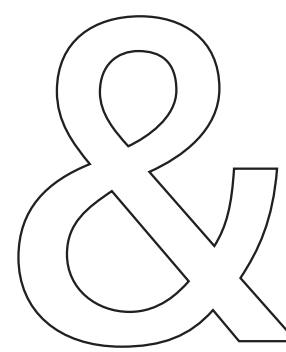
Vreeswijk MP, van der Klift HM. 2012. Analysis and interpretation of RNA splicing alterations in genes involved in genetic disorders. Methods Mol Biol 867: 49-63.

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

Yoshioka K, Yoshioka Y, Hsieh P. 2006. ATR kinase activation mediated by MutSalpha and MutLalpha in response to cytotoxic O6-methylguanine adducts. Mol Cell 22: 501-510.

Appendix

Summary
Samenvatting
List of Publications
Curriculum Vitae
Acknowledgements



SUMMARY

Lynch syndrome (LS) is a hereditary syndrome that predisposes to various cancer types, including colon, endometrial, gastric, brain and urological cancers. The syndrome is caused by heterozygosity for a mutation that inactivates one of the DNA mismatch repair (MMR) genes *MSH2*, *MSH6*, *MLH1* or *PMS2*. In a significant fraction of individuals suspected of LS a missense mutation or small in-frame deletion is identified in one of the MMR genes. Since the causality of such genetic variants is difficult to assess these are often referred to as variants of uncertain significance (VUS). The lack of classification of VUS precludes LS diagnosis. As a consequence, targeted chemoprevention and chemotherapy cannot be applied. In addition, when a suspected Lynch syndrome patient carries a VUS, all first-degree relatives enroll in lifelong periodic screening, irrespective of their mutation status. Screening poses a physical and psychological burden on these families, but also a burden on the preventive health care apparatus. For these reasons it is of great relevance to develop widely applicable, validated, and standardized approaches for the diagnostic assessment of VUS in MMR genes.

This thesis focuses on the development of *in vitro* and *in vivo* functional assays to enable diagnosis of VUS in MMR genes. **Chapter 1** introduces the field of MMR by reviewing recent work on the biochemistry and the genetics of MMR. Furthermore, the genetics, phenotype, diagnosis and the implications of LS are reviewed. The thesis then continues with a discussion of recent advances in functional analysis of VUS in MMR genes (**Chapter 2**).

In **Chapter 3**, an *in vitro* assay is presented that functionally analyzes missense variants in *MSH2* and *MSH6*. Missense variants found in patients suspected of LS are recreated as variant cDNAs in two sequential PCR reactions. The amplicons are then used as templates to produce variant proteins in an in vitro transcription/translation reaction. The variant protein is dimerized with its partner to create a heterodimer, and this dimer is used to complement a cell extract deficient for that heterodimer. This recombinant heterodimer and cell extract mixture are then supplemented with an MMR substrate: a plasmid containing a single, defined G·T mismatch that disrupts a *Hin*DIII restriction site. After the *in vitro* MMR reaction repair of the mismatch to A·T can be assessed by *Hin*DIII digestion. Since the plasmid substrate is fluorescently labeled, accurate quantification of repair products can be performed on a fragment analyzer.

The assay is technically validated by showing its reproducibility, by showing that the relevant proteins are present in (near) stoichiometric quantities and by showing that the assay has sufficient resolution to distinguish known pathogenic variants from polymorphic variants. The *in vitro* expressed variant proteins are not in excess, which should allow the detection of subtle repair defects. In the future, this may allow the diagnosis of variants with intermediate penetrance. The assay does not require molecular cloning, cell culture and use of bacteria. In addition, due to its relative simplicity, the ability to generate reagents in large quantities and due to its relatively high speed, this assay may be suited for integration into clinical diagnostic labs, which would be unique in the field of diagnosis of variants in cancer predisposition genes. We have adapted this assay for the functional analysis of VUS in the MMR genes *MLH1* (**Chapter 4**) and *PMS2* (**Chapter 5**).

Due to the cell-free nature of the assay described above this assay may have some pitfalls. Thus, defects in splicing, in protein stability or nuclear localization are not detected. These drawbacks could lead to false negatives (variants that are repair proficient in the assay, while they are actually pathogenic). To circumvent these issues, we have established an additional protocol to classify VUS in MMR genes. We reasoned that the comprehensive knowledge of residues in MMR proteins that are essential for *in vivo* function could aid in classifying MMR gene VUS. Such knowledge would enable the *a priori* classification of VUS without the need for further functional analysis. To acquire such knowledge we have taken a genetic screen-based approach (**Chapter 6**). Cells, heterozygous for *Msh2* due to targeted disruption of one of the alleles, are treated with a point mutation-inducing mutagen. After this, clones that have acquired an inactivating substitution at the *Msh2* gene are selected for and the complete *Msh2* open reading frame in these clones is sequenced to identify the causal mutation, and thereby the essential residue.

In this chapter we describe the generation and analysis of 26 of the above-mentioned cell lines, all carrying unique amino acid substitutions in Msh2. These substitutions affect 23 residues. We show that 10 of these residues have previously been reported as mutated in humans, and 6 substitutions are identical at the protein level. This supports the applicability of this protocol for the identification of pathogenic missense variants identified in humans. We have further validated this approach by showing that these cell lines have all phenotypes of MMR deficiency including microsatellite instability, a spontaneous mutator phenotype and strong tolerance to S_N1-methylating drugs. Furthermore, we show that these phenotypes are caused by a genuine MMR defect, as extracts from these cell lines cannot repair a mismatch on an artificial substrate. Since this phenotype can be rescued by addition of recombinant wild type Msh2/Msh6, this defect is indeed due to defects in Msh2/Msh6 and not due to defects in other genes/proteins. In addition, the mutations found inactivate MMR in an *in vitro* MMR assay and are therefore causative for the phenotype. Moreover, we describe dominant-negative effects for two substitutions. Such a phenotype could be of particular relevance since in patients these substitutions may already confer cancer predisposition in a heterozygous state. This is corroborated by the phenotype of a carrier of one of these substitutions who shows an extended tumour spectrum (colon, ovary, breast, oesophagus) and a young average age of onset of cancer. Another finding in this work is that in a number of the MMR-deficient variants protein stability is not affected. This is of relevance since genetic testing in LS patients is often preceded by an immunohistochemical analysis of protein expression in the tumor. In case normal protein expression is found, no further genetic screening is performed. Our data suggest that this is not warranted.

In **Chapter 7**, we apply the genetic screening technique described in Chapter 6 to the MMR gene *Msh6*. We have generated 51 cell lines that carry unique amino acid substitutions in Msh6, affecting 45 residues. Similar to Chapter 6, we show that several of these substitutions are also identified in humans, which suggests pathogenicity for these alleles and supports applicability of this assay to identify pathogenic variants in *MSH6*. A selection of these cell lines has been further validated and these lines display an MMR deficient phenotype, apparent from their elevated spontaneous mutation frequency and microsatellite instability.

To gain insights into the mechanism of loss of MMR in these lines we first investigated Msh6 protein expression by Western blotting and found that ~40% of all substitutions destabilize Msh6. The lines that retained normal Msh6 protein expression were tested for their ability to perform mismatch binding. This revealed that, of the 29 mutants tested, 21 have lost the ability to bind mismatched DNA, suggesting that defective mismatch binding is a major mechanism of loss of activity. These mutants included substitutions at residues that have previously been implicated in mismatch binding in crystallographic studies, but also substitutions that have not been implicated in this process. The proteins that do bind mismatches generally are proficient in sliding clamp formation, with the exception of substitution T1217I. This suggests that these substitutions have an MMR defect downstream of sliding clamp formation, such as in MutLα recruitment. We propose a model in which mutations hinder (minor) conformational changes that may be required during or before mismatch binding.

The thesis ends with a summarizing discussion and a reflection on future perspectives (**Chapter 8**). This thesis describes multiple methods that allow a functional analysis of MMR gene VUS. These methods allow to determine pathogenicity of such variants and may be instrumental in facilitating diagnosis in the upcoming age of personalized medicine, helping to shape the future of clinical molecular diagnostics. In addition, these methods provide tools to provide new insights in the genetics and biochemistry of the MMR pathway.

SAMENVATTING

Lynch syndroom (LS) is een erfelijk syndroom dat aanleg veroorzaakt voor het ontwikkelen van verschillende soorten kanker, waaronder dikke darm-, endometrium-, maag-, hersen- en urologische kankers. Het syndroom wordt veroorzaakt door heterozygositeit voor een mutatie die één van de DNA mismatch herstel (MMH) genen *MSH2*, *MSH6*, *MLH1* of *PMS2* inactiveert. In een significante fractie van individuen die worden verdacht van LS wordt een missense mutatie of een kleine in-frame deletie in een MMH gen geïdentificeerd. Omdat de causaliteit van zulke genetische varianten moeilijk te bepalen is wordt er vaak naar deze varianten verwezen als "varianten met onzekere significantie" (VOS). Het uitblijven van classificatie van VOS verhindert LS diagnose. Als gevolg kan er geen gerichte chemopreventie en chemotherapie worden toegepast. Daarnaast zorgt de detectie van een VOS in een LS-verdachte ervoor dat al zijn/haar familieleden in levenslange periodieke screeningsprogramma's terecht komen, ongeacht hun mutatiestatus. Dergelijke screenings vormen een lichamelijke en psychische last voor die families, maar ook voor de preventieve gezondheidszorg. Om deze redenen is het zeer relevant om breed toepasbare, gevalideerde en gestandaardiseerde toepassingen te ontwikkelen om VOS in MMH genen diagnostisch te kunnen onderzoeken.

Dit proefschrift richt zich op het ontwikkelen van *in vitro* en *in vivo* functionele assays om VOS in MMH genen te kunnen diagnosticeren. **Hoofdstuk 1** introduceert het MMH veld door recent gepubliceerd werk over de biochemie en de genetica van MMH te bespreken. Daarnaast wordt de genetica, het fenotype, de diagnose en de implicaties van LS besproken. Het proefschrift gaat dan verder met een bespreking van de recente ontwikkelingen in de functionele analyse van VOS in MMH genen (**Hoofdstuk 2**).

In **Hoofdstuk 3** wordt een *in vitro* assay gepresenteerd voor de functionele analyse van missense varianten in *MSH2* en *MSH6*. Missense varianten die zijn gevonden in patiënten verdacht van LS worden gerecreëerd als variante cDNAs in twee opvolgende PCR reacties. De PCR producten worden dan gebruikt als template in een *in vitro* transcriptie/translatie reactie. De hieruit resulterende eiwitten worden gedimeriseerd met hun partner om een heterodimeer te creëren, en deze dimeer wordt gebruikt om een cel extract te complementeren dat deficiënt is voor die bewuste heterodimeer. Deze mix van cel extract en (variante) heterodimeer wordt gesupplementeerd met een MMH substraat: een plasmide dat een enkele, gedefinieerde G·T mismatch bevat die een *Hin*DIII restrictiesite verstoort. Na de *in vitro* MMH reactie kan reparatie van de mismatch naar A·T worden bepaald door *Hin*DIII digestie. Omdat het plasmide substraat fluorescent is gelabeld kunnen herstel producten nauwkeurig worden gekwantificeerd door middel van fragment analyse.

De assay is technisch gevalideerd door aan te tonen dat de assay reproduceerbaar is, dat de relevante eiwitten in vrijwel stoichiometrische hoeveelheden aanwezig zijn en door aan te tonen dat de assay voldoende resolutie heeft om bekende, pathogene varianten te kunnen onderscheiden van bekende polymorfismen. Het *in vitro* tot expressie gebrachte eiwit is niet in overmaat, wat het detecteren van subtiele herstel defecten mogelijk moet maken. Wellicht leidt dit in de toekomst tot de mogelijkheid om varianten met intermediaire

penetrantie te kunnen diagnosticeren. De assay vereist geen kloneerwerk, geen celkweek en geen bacteriewerk. Door de relatieve eenvoud van de assay, de mogelijkheid de reagentia in grote hoeveelheden te produceren en door de relatieve hoge doorvoersnelheid, voldoet de assay aan de eisen om geïntegreerd te worden in klinisch diagnostische laboratoria, wat uniek zou zijn in het veld van diagnose van varianten in kanker predispositie genen. We hebben de assay tevens aangepast voor de functionele analyse van VOS in de MMH genen *MLH1* (**Hoofdstuk 4**) en *PMS2* (**Hoofdstuk 5**).

Door de cel-vrije aard van de assay die hierboven is beschreven, heeft de assay mogelijke nadelen. Defecten in pre-mRNA splicing, in eiwit stabiliteit of nucleaire eiwitlokalisatie worden niet gedetecteerd. Dit kan leiden tot vals negatieven (varianten die herstel proficient zijn in de assay, terwijl ze werkelijk pathogeen zijn). Om deze problemen te omzeilen hebben we een tweede protocol ontwikkeld om VOS in MMH genen te classificeren. We redeneerden dat de allesomvattende kennis van residuen in MMH eiwitten die essentieel zijn voor *in vivo* eiwitfunctie zou kunnen helpen bij het classificeren van VOS in MMH genen. Dergelijke kennis zou een *a priori* classificatie van VOS mogelijk maken, zonder verdere noodzaak voor functionele analyses. Om die kennis te verwerven hebben we voor een genetische screen als aanpak gekozen (**Hoofdstuk 6**). Cellen, heterozygoot voor *Msh2* door een gerichte verstoring van één van de allelen, worden behandeld met een puntmutatie-inducerende stof. Vervolgens wordt er geselecteerd voor kloons die een inactiverende mutatie in *Msh2* hebben opgelopen, het volledige open leesraam van *Msh2* wordt gesequenced om de causale mutatie te identificeren, en hierbij ook het essentiële residu.

In dit hoofdstuk beschrijven we de generatie en analyse van 26 van de eerder beschreven cel lijnen, die allen unieke aminozuur substituties dragen in Msh2. Deze substituties hebben betrekking op 23 residuen. We laten zien dat 10 van deze residuen als gemuteerd zijn gerapporteerd in mensen, waarvan 6 substituties identiek zijn op eiwitniveau. Dit ondersteunt de toepasbaarheid van dit protocol op de identificatie van pathogene missense substituties in mensen. We hebben de aanpak verder gevalideerd door aan te tonen dat de gegenereerde cellijnen alle fenotypes hebben die passen bij MMH deficiëntie, waaronder microsatelliet instabiliteit, een spontaan mutator fenotype en sterke tolerantie jegens S_N1-methylerende stoffen. Daarnaast tonen we aan dat dit fenotype wordt veroorzaakt door een echt MMH defect, omdat extracten van deze cellijnen een mismatch op een kunstmatig substraat niet kunnen herstellen. Dit defect kan weer gered worden door wild type Msh2/Msh6 aan het extract toe te voegen, wat aangeeft dat het MMH defect wordt veroorzaakt door defecten in Msh2/Msh6 en niet door defecten in andere eiwitten. Omdat de mutaties die we in de lijnen aantonen MMH inactiveren in een *in vitro* assay, zijn deze causatief voor het fenotype. Bovendien beschrijven we dominant-negatieve effecten voor twee substituties. Een dergelijk fenotype zou zeer relevant kunnen zijn, omdat deze substituties in dragers al voor kanker predispositie zou kunnen zorgen in een heterozygote staat. Dit wordt ondersteund door het fenotype van een drager van één van deze mutaties, die een verlengd tumor spectrum vertoont (dikke darm-, eierstok-, borst- en slokdarmkanker) en een jonge leeftijd waarop deze kankers zich manifesteren. Een andere vinding in dit werk is, dat enkele MMH-deficiënte varianten

eiwitstabiliteit behouden. Dit is relevant omdat het genetisch diagnosticeren van LS patiënten vaak vooraf wordt gegaan door een immunohistochemische kleuring van de tumor van de patiënt. Wanneer er normale eiwitexpressie wordt gevonden, wordt er geen verdere genetische test uitgevoerd. Onze data suggereert dat dit ongegrond is.

In **Hoofdstuk 7** passen we de genetische screen beschreven in Hoofdstuk 6 toe op het MMH gen *Msh6*. We hebben 51 cel lijnen gegenereerd die unieke aminozuur substituties bevatten in Msh6, die 45 residuen aantasten. Gelijk aan Hoofdstuk 6 laten we zien dat enkele van deze substituties in mensen zijn geïdentificeerd, wat pathogeniciteit voor deze allelen suggereert en de toepasbaarheid van deze assay voor de identificatie van pathogene varianten in *MSH6* ondersteunt. Een selectie van deze cellijnen is verder gevalideerd en deze lijnen vertonen een MMH deficiënt fenotype, wat duidelijk wordt door de verhoogde spontane mutatie frequentie en microsatelliet instabiliteit.

Om inzichten te verkrijgen in het mechanisme van MMH verlies in deze lijnen, hebben we allereest Msh6 eiwitexpressie onderzocht door middel van Western blotting. Hieruit blijkt dat ~40% van alle substituties Msh6 destabiliseren. Cel extracten van de lijnen met stabiele Msh6 expressie zijn vervolgens getest op het vermogen om mismatches te binden. Dit toonde aan dat, van de 29 mutante extracten die zijn getest, 21 extracten niet in staat zijn om mismatches te binden. Dit suggereert dat een defect in mismatch binding een vooraanstaand mechanisme is voor verlies van MMH activiteit. Het verlies van mismatch binding werd aangetoond voor mutanten met substituties op residuen die eerder zijn geïmpliceerd in mismatch binding in kristallografische studies, maar ook voor residuen waarvan de implicaties in dit proces niet eerder zijn aangetoond. De eiwitten die nog mismatches binden zijn, over het algemeen, proficient in het vormen van een "sliding clamp", substitutie T1217I uitgezonderd. Dit suggereert dat deze substituties een MMH defect hebben na sliding clamp formatie, zoals in het rekruteren van MutLα. We stellen een model voor, waarin mutaties op residuen (kleine) conformatie veranderingen verhinderen die noodzakelijk zijn vóór mismatch binding.

Dit proefschrift eindigt met een discussie en een reflectie op toekomstige perspectieven (**Hoofdstuk 8**). Het proefschrift beschrijft verschillende methoden om VOS in MMH genen functioneel te kunnen analyseren. Deze methodes stellen ons in staat de pathogeniciteit van zulke varianten te bepalen, kunnen diagnose faciliteren in het opkomende tijdperk van gepersonaliseerde gezondheidszorg, en helpen de toekomst te vormen in de klinisch moleculaire diagnostiek. Daarnaast leveren deze methoden het gereedschap om inzichten in de genetica en biochemie van MMH te verkrijgen.

LIST OF PUBLICATIONS

Drost M, Zonneveld JB, Van Dijk L, Morreau H, Tops CM, Vasen HFA, Wijnen J, De Wind N. A cell-free assay for the functional analysis of variants of the mismatch repair protein MLH1. *Hum Mutat 2010 31:247–253.*

Drost M, Zonneveld JB, van Hees S, Rasmussen LJ, Hofstra RM, de Wind N. A rapid and cell-free assay to test the activity of Lynch syndrome-associated MSH2 and MSH6 missense variants. *Hum Mutat 2012 33:488-494.*

Andersen SD, Liberti SE, Lützen A, **Drost M**, Bernstein I, Nilbert M, Dominguez M, Nyström M, Van Overeem Hansen T, Christoffersen JW, Jäger AC, de Wind N, Nielsen FC, Tørring PM, Rasmussen LJ. Functional characterization of MLH1 missense variants identified in Lynch Syndrome patients. *Hum Mutat 2012 33:1647–1655*

Rasmussen LJ, Heinen CD, Royer-Pokora B, **Drost M**, Hofstra RM, de Wind N. Pathological assessment of mismatch repair gene variants in Lynch syndrome: past, present and future. *Hum Mutat 2012 33:1617–1625*

Drost M, Lützen A, van Hees S, Ferreira D, Calléja F, Zonneveld JB, Nielsen FC, Rasmussen LJ, de Wind N. Genetic screens for the identification of pathogenic gene variants in the cancer predisposition Lynch syndrome. *Proc Natl Acad Sci USA 2013 23:9403-9408*

Rasmussen LJ, Heinen CD, Royer-Pokora B, **Drost M**, Tavtigian S, Hofstra RM, de Wind N. Response to: Design of a core classification process for DNA mismatch repair variations of a priori unknown functional significance. *Hum Mutat 2013 34:923-924*

Drost M, Koppejan HJ, de Wind N. Functional analysis of missense variants in the DNA mismatch repair gene *PMS2* found in patients suspected of lynch syndrome. *Hum Mutat 2013 34:1477-1480*

Farrel MP, Hughes DJ, **Drost M**, Cummings RJ, Fletcher TA, Meany MA, Andrews EJ, Kay EW, de Wind N, Power DG, Green AJ, Gallagher DJ. Multivariate analysis of MLH1 c.1664T>C (p.Leu555Pro) mismatch repair gene variant demonstrates its pathogenicity. *Fam Cancer 2013 12:741-747*

Parsons MT, Whiley PJ, Beesley J, **Drost M**, de Wind N, Thompson BA, Marquart L, Hopper JL, Jenkins M, Colon Cancer Family Registry, Brown MA, Tucker K, Warwick L, Buchanan DD, Spurdle AB. Consequences of germline variation disrupting the constitutional translational initiation codon start sites of MLH1 and BRCA2: use of potential alternative start sites and implications for predicting variant pathogenicity. *Mol Carcinog 2014 in press*

Dominguez Valentin M, **Drost M**, Therkildsen C, Rambech E, Ehrencrona H, de Wind N, Nilbert M, Rasmussen LJ. Functional implications of the p.Cys680Arg mutation in the MLH1 mismatch repair protein. *Under revision*

CURRICULUM VITAE

Mark Drost was born in Delft, The Netherlands on April 8th, 1983. He attended the Stanislas College in Delft and received his VWO diploma in 2001. From 2002 to 2006 he attended the higher laboratory school (HLO) in Leiden, focusing on molecular biology. During his studies he performed an internship at the department of Clinical Genetics of the Leiden University Medical Center where he investigated the effect of DNA variants in the BRCA1 and BRCA2 breast cancer predisposition genes on splicing of their pre-mRNAs. In 2006 he acquired his Bsc and started working at the department of Toxicogenetics of the Leiden University Medical Center, in the group of Dr. Niels de Wind. He worked on a project funded by the Dutch Stomach Liver Intestine Foundation that focused on developing an in vitro assay for the functional analysis of missense mutations in DNA mismatch repair genes. In 2008 he started his PhD, under supervision of Dr. Niels de Wind and Prof. Leon Mullenders, on an European Union-funded project, further focusing on the functional analysis of missense mutations in DNA mismatch repair genes, not only in vitro, but also in vivo.

This thesis describes the efforts of this period. Also, Mark spent one week as a guest scientist at the Institute for Molecular Cancer Research in Zürich. In addition to his research, he helped to organize the annual PhD student retreat to Bruges, Belgium, in 2009. He has attended national and international conferences, was awarded a poster prize at the PhD student retreat in 2011 and was selected for oral presentations at the Human Variome Project meeting in Paris in 2010, The Novel Targets in Personalized Cancer Medicine meeting in Antwerp in 2011, and the Responses to DNA Damage meeting in Egmond aan Zee in 2011.

From 2014 he continues his work as a postdoctoral researcher on projects funded by the Dutch Cancer Society and National Institutes of Health that focus on the further development and validation of functional assays for the diagnostic assessment of missense mutations in mismatch repair genes.

ACKNOWLEDGEMENTS

There it is: The thesis. I remember when I was in primary school and my teacher asked us to make a drawing of the future you, depicting your profession. I drew a guy in a white coat holding test tubes. Frankly: In high school, I dropped chemistry once I could. Now I acquire a PhD, with a thesis that has a crystal structure on the cover. Life tends to catch up with you, I guess.

This PhD feels like an absolute highlight, professionally. However, I think in many years from now, I will mainly remember it socially. My PhD kicked off with a PhD workshop in Heidelberg, quickly followed by a Young Scientist Meeting in Porto. I will remember my PhD as a time during which I met lots of great people. To start with my fellow PhD students from the department: Godelieve (the Porto meeting I will never forget), Angela, Aude, Dimitris, Bennie, Wouter, Sophie, Daphne, Agata and Robin. Yes, you truly are great people.

Organizing the MGC PhD Workshop in Bruges was an absolute highlight. Madeleine, Marion, Sjoerd, Dwi, Rudolf, Tommy, Elena en Boet: I had great fun. Organizing a scientific workshop is one thing, but making it such a great "social" success is just as important. I think we succeeded in both.

Many thanks to all of my fellow (former) LUMC PhDs, Antoine, Peter, Rudolf, Elmar, Joris, Patrick, Hester, Steven, Wouter, Kristiaan. Colleagues from Rotjeknor: Michael, Petra, Kasper, Klaas, Tommy: Thanks for giving me a great time. The same goes for lots of fellow scientists who work abroad, who I have met either in Copenhagen, Zürich, Porto, Düsseldorf, Paris or Ventura. And everyone else who I forgot.

To the "de Wind group": It has been a pleasure working with all of you.

Jaap, thanks for all of your (wise and calm!) advice, but in particular all talks in the lab, where we'd discuss models, publications, work discussions, etcetera. Anastasia, you are truly a great presence within a group. Whenever I hear someone humming, I think of you. Sandrine, your light is always shining bright. Poppy, for me, it is interesting to see how two people within one group can do completely different PhDs. You are next! Ana, I hope you'll have a great time in the group. Ewoud, for now I will leave it at "Shabu".

To all former group members: Giel, Cristina, Leone, Bea, Johan, thanks a lot for everything. A special word of gratitude goes to Linda, Daniel, Fabiënne and José, without you this thesis would not have been possible.

Jane, Ivo and Jordi: Thanks for being great roomies.

To all of my students: Bas, Hester, David, Saskia, Max: I have been very lucky to have had you as students, all of you did very well. Hester, your data has greatly contributed to a very nice publication. For David, Saskia and Max, that honor will hopefully follow soon.

Leon, thank you for being my promotor.

Niels, I am very grateful for all the possibilities you have given me and still give me. To me, you are a great boss and supervisor. Working together has always been a pleasure: We always seem on the same track, if not, mere words are needed to correct this.

Aan mijn dertien vaste "Hollandse korfbal vrienden": Ik hoop na mijn promotie nooit meer te horen wanneer ik mijn celletjes weer te eten moet geven.

Aan mijn familie, ouders: Een goed nest is een basis voor alles.

Dear Joyce, my dear little Noud, the two of you mean the world to me.