Characterization of MSH2 variants by endogenous gene modification in mouse embryonic stem cells

Eva Wielders, Rob J Dekker, Ian Holt, Glenn E Morris, Hein Te Riele

Human Mutation

Characterization of MSH2 variants by endogenous gene modification in mouse embryonic stem cells

Journal:	*Human Mutation*
Manuscript ID:	humu-2010-0353.R1
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	29-Nov-2010
Complete List of Authors:	Wielders, Eva; The Netherlands Cancer Institute, Molecular Biology Dekker, Rob; The Netherlands Cancer Institute, Molecular Biology Holt, Ian; Wolfson Centre for Inherited Neuromuscular Disease, RJAH Orthopaedic Hospital Morris, Glenn; Wolfson Centre for Inherited Neuromuscular Disease, RJAH Orthopaedic Hospital te Riele, Hein; The Netherlands Cancer Institute, Molecular Biology
Key Words:	Lynch Syndrome, Variants of uncertain significance, DNA mismatch repair, Genetic counseling, oligonucleotide-directed gene modification

Characterization of *MSH2* variants by endogenous gene modification in mouse embryonic stem cells

Running title: Analysis of unclassified *MSH2* variants

Eva A.L. Wielders[1], Rob J. Dekker[1], Ian Holt[2], Glenn E. Morris[2] and Hein te Riele[1,*]

[1] The Netherlands Cancer Institute, Division of Molecular Biology, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands

[2] Wolfson Centre for Inherited Neuromuscular Disease, RJAH Orthopaedic Hospital, Oswestry, Shropshire, SY10 7AG, UK and Institute for Science and Technology in Medicine, Keele University, Keele, Staffordshire, ST5 5BG, UK

[*]corresponding author

Phone: 31-20-512 2084

Fax: 31-20-669 1383

Email: h.t.riele@nki.nl

Abstract

Mutations in the mismatch repair gene *MSH2* underlie hereditary non-polyposis colorectal cancer (Lynch syndrome). While disruptive mutations are overtly pathogenic, the implications of missense mutations found in sporadic colorectal cancer patients or in suspected Lynch syndrome families are often unknown. Adequate genetic counseling of mutation carriers requires phenotypic characterization of the variant allele. We present a novel approach to functionally characterize *MSH2* missense mutations.

Our approach involves introduction of the mutation into the endogenous gene of murine embryonic stem cells (ESC) by oligonucleotide-directed gene modification, a technique we recently developed in our lab. Subsequently, the mismatch repair capacity of mutant ESC is determined using a set of validated functional assays.

We have evaluated four clinically relevant MSH2 variants and found one to completely lack mismatch repair capacity while three behaved as wild-type MSH2 and can therefore be considered as polymorphisms.

Our approach contributes to an adequate risk assessment of mismatch repair missense mutations. We have also shown that oligonucleotide-directed gene modification provides a straightforward approach to recreate allelic variants in the endogenous gene in murine embryonic stem cells. This approach can be extended to other hereditary conditions.

Keywords: DNA mismatch repair/Genetic counseling/Lynch syndrome/Oligonucleotide-directed gene modification/Variants of uncertain significance

Introduction

Hereditary non-polyposis colorectal cancer (HNPCC/Lynch syndrome) is characterized by early onset cancer of the gastrointestinal and genitourinary tract, in particular colorectal and endometrial cancer (Lynch and de la Chapelle, 1999). In rare cases, sebaceous gland tumors (Muir-Torre syndrome) and glioblastoma multiforme (Turcot's syndrome) are seen (Lynch and de la Chapelle, 1999). HNPCC-related tumors typically show miscrosatellite instability (MSI) indicative for defective DNA mismatch repair (MMR). The MMR system has three functions that help sustaining genomic stability (Jiricny, 2006): *(1)* recognition and repair of mis- and unpaired bases that can spontaneously arise by erroneous DNA replication; *(2)* suppression of recombination between homologous but not identical DNA sequences (homeologous recombination); *(3)* induction of apoptosis in response to certain DNA damaging chemotherapeutics such as methylating agents, 6-thioguanine (6-TG) and cisplatin. MMR is initiated by MSH2/MSH6 or MSH2/MSH3 heterodimers. These complexes bind to mismatches, small loops of unpaired bases and to several DNA adducts. Subsequently, MLH1/PMS2 or MLH1/MLH3 are recruited to further enable the repair process. Inherited mutations in *MSH2* and *MLH1*, and to a lesser extent in *MSH6* and *PMS2*, have been found to underlie HNPCC/Lynch syndrome. Dependent on the type of mutation, the lifetime risk of developing these cancers is 70-80% and 40-60%, respectively, which is up to tenfold increased compared to the general population (Lindor, et al., 2006).

While large deletions, premature stop codons and frameshifts completely abrogating gene function are obviously pathogenic, the effects of missense mutations affecting only a single amino acid are more difficult to interpret. Given their frequent occurrence in MMR genes in (suspected) HNPCC cases (but also in genes involved in other pathologies), a method to determine the functional implications of such mutations is urgently needed. A

recent survey of MMR gene mutation databases has identified 136 *MSH2* missense mutations of which 13 were deleterious, 8 neutral and 115 unclassified, *i.e.*, with unknown pathological consequences (Chao, et al., 2008). Small family size and incomplete penetrance often preclude reliable assessment of co-segregation of the variant allele with disease. This complicates judgment of the pathological implications of these alleles, which are termed "variants of uncertain significance" (VUS) or "unclassified variants" (UV).

To shed light on the pathogenicity of VUS, a number of approaches are being used. Computer algorithms have been developed to estimate the severity of an amino acid substitution based on evolutionary conservation and the physicochemical consequences of the substitution (Chan, et al., 2007; Chao, et al., 2008). However, this approach requires validation by functional studies. In most of the studies that have been carried out to date, variant and wild-type proteins are ectopically expressed and compared with respect to stability, localization, protein interactions, mismatch binding, ATPase activity or the capacity to sustain an *in vitro* MMR reaction. Also the ability of the ectopically expressed variant to rescue the mutator phenotype of a MMR deficient strain or cell line is sometimes studied (Ou, et al., 2007). A limitation of such assays is that altered MMR protein activity can be masked by unique features of the ectopic expression of the plasmid driven mutant cDNAs, such as overexpression, altered mRNA processing, or the absence of correct epigenetic regulation of transcription. Furthermore, predicting the effect of altered biochemical characteristics of a mutant protein on the functionality of the MMR system is not straightforward. Moreover, MMR functions such as anti-recombination and induction of the DNA damage response are usually not addressed.

To overcome these limitations, we present here a novel approach to analyze the functional implications of *MSH2* VUS. Our system consists of three steps: *(1)* the codon

substitution is introduced into one allele of the endogenous *Msh2* gene in mouse embryonic stem cells (ESCs) by oligonucleotide-directed gene modification ('oligo targeting') (Aarts, et al., 2006); *(2)* cells are made homozygous for the mutation; *(3)* the phenotype is studied by functional assays that address main MMR functions. The great advantage of this approach is that variant alleles are expressed from the endogenous locus. This is crucial as both higher and lower levels of MMR proteins are notorious for affecting MMR capacity (Claij and Te Riele, 2002; Zhang, et al., 1999). Furthermore, this approach allows us to study three MMR functions that are relevant for maintaining the integrity of the genome. Here we have studied four MSH2 VUS that have been found in (suspected) HNPCC families

Materials and methods

Generation of codon substitutions in *Msh2*

The oligo targeting procedure was carried out in *Msh2*$^{+/+}$ or *Msh2*$^{+/hyg}$ ESCs, essentially as described (Aarts, et al., 2006). Briefly, we transfected ESCs with a single-stranded DNA oligonucleotide (35 to 40 residues, see Fig. 1A and Supplementary Figure 1A), after establishing transient down regulation of MSH2. Transfected cells were expanded and plated at 5000 cells/well on 96-well plates. Mutation specific PCR was used to identify pools containing mutated cells. First, an 800-1000 bp fragment containing the modified codon was amplified using genomic DNA isolated from 2.5 x 10^4 ESCs, 1.25 U *Taq* polymerase, 1 x PCR buffer containing 1.6 mM $MgCl_2$, 12.5 pmol of each primer and 0.2 mM dNTPs in a total volume 25 µl. After a denaturing step of 95°C for 5 min, amplification was conducted in 30 cycles of 30 s at 95°C, 1 min at 60°C and 1.5 min at 72°C followed by a final elongation step of 10 min at 72°C. Next, 1 µl of this product was used in nested PCR with mutation

specific primers, resulting in 200-400 bp products. Other than a lower annealing temperature (56°C) and a shorter annealing time (1 min) the PCR program was identical.

A positive pool was subcloned using limiting dilution in pools of 1000, 100 and 1 cell/well. Once a clonal cell line was established, we isolated RNA and made cDNA using an *Msh2* specific RT primer (5'-gagccggagcctttatcc-3'). We then amplified a cDNA fragment containing the modified codon using nested PCR. The resulting PCR product was cloned into the pGEM®-T Easy vector (promega) and sequenced using vector primers T7 and SP6. Primer sequences are available upon request.

Duplication of the targeted allele

We targeted the heterozygous mutant cell lines with a *Pim1-neo* targeting construct (te Riele, et al., 1990). For each cell line, six $Pim1^{+/neo}$ clones were selected and plated on a 6-well plate each, at a density of 1 x 10^5 cells/well. The next day we added G418 at final concentrations of 5, 6, 7, 8, 9 and 10 mg/ml. Medium was refreshed every 3-4 days for 2-3 weeks after which colonies were picked and expanded. These colonies were screened for duplication of the $Pim1^{neo}$ allele by Southern blot analysis (te Riele, et al., 1990) or by PCR using primer pairs Pim1-F 5'-atcaactccctggcccacct-3' and Pim1-R 5'-ggcctggctcaccatcaaag-3'.

We subsequently screened resulting $Pim1^{neo/neo}$ colonies for duplication of the mutated *Msh2* allele. For the $Msh2^{G322D/G322D}$ cell line, we used mutation specific restriction enzyme analysis. RNA was isolated from $Msh2^{+/+}$, $Msh2^{G322D/+}$, and the $Pim1^{neo/neo}$ cell lines. cDNA was made using the *Msh2* specific RT primer and used to amplify a cDNA fragment containing the modified codon. 400 ng of purified PCR product was digested with 4 units of *Btg1* restriction enzyme (NEB) for 2 hours at 37°C.

In case of the *Msh2*$^{P622L/P622L}$ cell line, the loss of the *Msh2*^{hyg} allele was confirmed using Southern blot analysis (de Wind, et al., 1995). As neither southern blot or mutation specific restriction enzyme analysis was possible for the *Msh2*$^{Y103C/Y103C}$ and the *Msh2*$^{M688I/M688I}$ cell lines, we used sequencing to verify duplication of the *Msh2* mutation. We amplified the genomic DNA of the wild-type, heterozygous and homozygous mutant cell lines and performed a sequencing reaction on the purified PCR product. Primer sequences are available upon request.

Western blot analysis

Cells were lysed in a buffer containing 150 mM NaCl, 50 mM Hepes pH 7.5, 5 mM EDTA, 0.1% NP-40, 5 mM NaF, 0.5 mM vanadate, 20 mM β-glycerolphosphate and 1 tablet complete protease inhibitor cocktail (Roche) per 50 ml. Protein extracts from $1.5 \cdot 10^5$ ES cells were separated by 3-8% Tris-Acetate gels (NuPAGE®) using the NuPAGE® electrophoresis system and transferred to nitrocellulose membrane. We used rabbit polyclonal antibodies as primary antibodies to detect MSH2 (de Wind, et al., 1998) (1:500), MSH6 (de Wind, et al., 1999) (1:500) and MLH1 (sc-581, Santa Cruz Biotechnology) (1:500) and mouse monoclonal antibodies to detect MSH3 (1:50) (I. Holt and G. Morris, submitted for publication) and γ-Tubilin (GTU-88, Sigma-Aldrich). Peroxidase-conjugated goat anti-rabbit IgG and goat anti-mouse IgG (BioSource International) were used as a secondary antibody. Signals were visualized with enhanced chemiluminescence.

Immunoprecipitations

$1 \cdot 10^7$ cells of the four mutant cell lines, *Msh2*$^{Y103C/Y103C}$, *Msh2*$^{G322D/G322D}$, *Msh2*$^{P622L/P622L}$, Msh2$^{M688I/M688I}$, and the control cell lines *Msh2*$^{+/+}$, *Msh2*$^{hyg/hyg}$, *Msh2*$^{low/hyg}$ were lysed in 1 ml ELB (50 mM HEPES pH 7.5, 150 mM NaCl, 6 % glycerol, 0.5 % NP-40 and Complete

EDTA-free protease inhibitor cocktail (Roche Applied Science, Indianapolis, IN)) for 30 min at 4 ºC. Lysates were cleared by centrifugation (20,800 x g, 10 min, 4º C). Immunoprecipitations were performed by incubating 700 µl cleared lysate with 2.2 µg MSH2 antibody (ab70270, Abcam, Cambridge, UK) for 4 h at 4 ºC under continuous rotation. Subsequently, 70 µl of protein G Sepharose (GE Healthcare, Piscataway, NJ) was added and incubation was continued for 16 h at 4 ºC. The beads were washed six times with 1 ml ELB at 4 ºC and resuspended in 20 µl 1 x NuPAGE LDS sample buffer (Invitrogen, Carlsbad, CA) for Western blot analysis. Lysates of wild-type ES cells incubated with protein G but without MSH2 antibody served as controls.

Hprt mutation assay and microsatellite instability assay

Single cells of the four mutant cell lines, $Msh2^{Y103C/Y103C}$, $Msh2^{G322D/G322D}$, $Msh2^{P622L/P622L}$, $Msh2^{M688I/M688I}$, and the control cell lines $Msh2^{+/+}$, $Msh2^{hyg/hyg}$, $Msh2^{low/hyg}$, were expanded to 10^9 cells. Each expanded clone was plated on three 150mm gelatin coated tissue culture plates at 1.5 x 10^6 cells/plate. The next day, 6-TG was added at a final concentration of 10 µg/ml. After 10 days, the resistant colonies were counted.

In addition, of each of the expanded cultures we generated about 60 subclones and isolated genomic DNA. The length of three different microsatellite markers (D18Mit19, D7Mit17, D14Mit15) was analyzed by PCR analysis.

Homologous recombination assay

We used targeting constructs 129Rb-pur and Balb/cRb-pur that differ about 0.6% at the nucleotide level (Claij and Te Riele, 2002). The targeting and subsequent analysis was performed in the $Msh2^{Y103C/Y103C}$, $Msh2^{G322D/G322D}$, $Msh2^{P622L/P622L}$, $Msh2^{M688I/M688I}$ and $Msh2^{low/hyg}$ cell lines as described (te Riele, et al., 1992).

Sensitivity to MNNG and 6-TG

The four mutant cell lines, *Msh2*$^{Y103C/Y103C}$, *Msh2*$^{G322D/G322D}$, *Msh2*$^{P622L/P622L}$, *Msh2*$^{M688I/M688I}$, and the control cell lines, *Msh2*$^{+/+}$, *Msh2*$^{hyg/hyg}$, *Msh2*$^{low/hyg}$, were plated onto irradiated mouse embryonic fibroblast feeder layers at a density of 500 cells/1.8 cm^2. The next day we treated the cells for 1h with 0-40 μM MNNG or 6-TG and after 4 days we counted the number of surviving colonies. In the case of MNNG exposure, cells were cultured in the presence of 40 μM O^6-benzylguanine, starting from 1 hour prior to MNNG exposure until counting of colonies. O^6-benzylguanine inhibits the removal of methyl groups from the O^6 position of guanine by endogenous O^6-methylguanine-methyltransferase activity.

Quantitative real time PCR

ES cells of the four mutant cell lines, *Msh2*$^{Y103C/Y103C}$, *Msh2*$^{G322D/G322D}$, *Msh2*$^{P622L/P622L}$, *Msh2*$^{M688I/M688I}$, and the control cell lines, *Msh2*$^{+/+}$, *Msh2*$^{hyg/hyg}$, *Msh2*$^{low/hyg}$ were grown to semi-confluency and total RNA was isolated using the RNeasy minikit (Qiagen). 1 μg of RNA was used for first strand cDNA synthesis using random primers and SuperScript II Reverse Transcriptase (Invitrogen). The standard amplification program of an ABI Prism 7000 sequence detection System (Applied Biosystems, Nieuwerker a/d IJssel, The Netherlands) was used for quantitative real-time PCR. The amplification was performed with 15-30 ng cDNA, 0.4 μM of forward primer MSH2-2109F (AATCGGGTGTTTTGTGCCC) and reverse primer MSH2-2171R (CGAGCAAGGATGCAATCCA), and 12.5 μl of 2x SYBR Green PCR Master Mix (Applied Biosystems) in a total volume of 25 μl. The final mouse Msh2 mRNA levels were normalized to mouse β-actin mRNA levels which were measured as described above using forward primer ACTINB-31F

(GCTTCTTTGCAGCTCCTTCGT) and reverse primer ACTINB-94R (ATATCGTCATCCATGGCGAAC).

Results

We have selected four VUS of clinical relevance, MSH2-Y103C, MSH2-G322D, MSH2-P622L and MSH2-M688I. The MSH2-Y103C mutation was identified by Myriad Genetics, Inc. and is listed in the InSiGHT database. Its yeast equivalent was studied in an ectopic expression system where it conferred an increased resistance to cisplatin without a mutator phenotype (Clodfelter, et al., 2005). The MSH2-G322D mutation has been found in suspected HNPCC families, but also in the control population. The latter is indicative for a polymorphism, however, a patient homozygous for this mutation developed colorectal cancer (CRC) at the age of 19 (Froggatt, et al., 1996). Yeast studies using ectopic expression of the equivalent MSH2 mutation indicated protein functionality to depend on expression levels, labeling it as an "efficiency polymorphism" (Drotschmann, et al., 1999; Ellison, et al., 2001). The third mutation, MSH2-P622L, showed perfect co-segregation with disease in a large HNPCC family (Leach, et al., 1993). Yeast studies have suggested this mutant protein to be defective in the repair of mismatches but capable of inducing a DNA damage response upon cisplatin exposure (Clodfelter, et al., 2005). MSH2-P622L may therefore represent an MSH2 'separation of function' mutant. Finally, MSH2-M688I has been found in families with suspected HNPCC in several Asian studies (Banno, et al., 2003; Nomura, et al., 2000; Yuan, et al., 1998) but also in healthy Japanese individuals, suggesting this is a polymorphism (Banno, et al., 2004). Studies of the yeast equivalent MSH2-M707I revealed a mild mutator

phenotype and a loss of binding to MLH1, PMS1, EXOI and POL30 in a yeast-two-hybrid system, whereas binding to MSH6 and MSH3 was found to be intact (Gammie, et al., 2007).

Oligo targeting was used to introduce the mutations into the endogenous *Msh2* gene of murine ESCs following the procedure described by Aarts et al. (2006). For the P622L mutation, we targeted the mutation into $Msh2^{+/hyg}$ ESCs (Fig. 1A and D), carrying a hygromycin-resistance gene (*hyg*) inserted into one *Msh2* allele (de Wind, et al., 1995). Since the *hyg* insertion is close to Proline 622 in exon 12, we could easily select a clone carrying the P622L substitution in the wild-type allele by mutation-specific PCR (see Supplementary Fig. 1A for the structure of the wild-type and $Msh2^{hyg}$ alleles). However, in general, it is less straightforward to ascertain modification of the $Msh2^{+}$ rather than the $Msh2^{hyg}$ allele and therefore, the other three mutations were introduced in $Msh2^{+/+}$ cells (Fig. 1A, B, C and E). This approach necessitated a second step, which is rendering cells homozygous for the mutant allele. To this aim, we made use of a previously described method to duplicate a *neo*-labeled chromosome by selecting for cells resistant to high concentrations of G418 (Mortensen, et al., 1992). First, a highly efficient targeting vector was used to insert *neo* into the *Pim1* gene, located centromeric of *Msh2* on chromosome 17 (te Riele, et al., 1990). Subsequently, several targeted clones were exposed to 5-10 mg/ml G418 and resistant colonies were screened for concomitant duplication of the *neo* marker and the *Msh2* mutant allele. Southern blot analysis or PCR was used to assess the *Pim1* status (not shown) and mutation specific restriction enzyme analysis or sequencing was used for *Msh2* (Fig. 2A, B and C). The same method was applied to generate a cell line expressing MSH2-P622L from both chromosomes. Here duplication of the mutated allele was verified by Southern blot analysis (loss of $Msh2^{hyg}$ allele) (Fig. 2D). This procedure allowed us to rapidly obtain homozygous mutant cell lines that were subsequently used for functional analysis.

While the M688I and the G322D allele produced normal protein levels, the levels of the Y103C and the P622L mutant protein were markedly decreased (Fig. 3), the latter in agreement with data obtained in yeast (Gammie, et al., 2007) and with immunohistochemical stainings for MSH2 on tumor material of patients carrying the P622L mutation (Southey, et al., 2005; Ward, et al., 2002). In none of the mutant cell lines decreased mRNA levels were detected using quantitative-PCR (Supplementary Fig. 2). Sequencing of the *Msh2*^{P622L} transcript revealed no additional mutations or splicing aberrancies (not shown). To exclude the possibility that low protein level was the result of poor codon usage, we reconstructed the most severely affected mutant, P622L, with a more prevalent leucine codon (CTG), having a codon usage of 40.0 versus 13.3 for the first mutant (TTG) (www.kazusa.or.jp/codon) (Supplementary Fig. 1A and B). Western blot analysis showed a similar reduction in the MSH2 protein levels for both *Msh2*$^{P622L/hyg}$ cell lines, ruling out an influence of codon usage (Supplementary Fig. 1C).

We next assessed the levels of the MSH2 binding partners MSH3 and MSH6. In the two mutant lines with normal MSH2 level, we detected normal MSH3 and MSH6 protein levels (Fig. 3). Since the level of MSH3 and MSH6 depends on their interaction with MSH2, this indicates normal binding of mutant MSH2 to these proteins. Somewhat remarkably MSH6 level was also near to normal in the MSH2-Y103C cell line despite lower levels of mutant MSH2. However, MSH3 level appeared to be twofold reduced. In homozygous *Msh2*$^{P622L/P622L}$ cells, the MSH6 level was reproducibly somewhat higher than in *Msh2*$^{hyg/hyg}$ cells, indicating that also for this mutant the MSH2/MSH6 interaction was intact, consistent with previous findings (Guerrette, et al., 1998). To verify these observations, we performed an MSH2 immunoprecipitation experiment (Supplementary Fig. 3). This confirmed normal binding of MSH2-Y103C, MSH2-G322D and MSH2-M688I to MSH3 and MSH6. Furthermore, we found the residual MSH2-P622L protein to co-precipitate MSH6

(Supplementary Fig. 3). However, binding of this mutant MSH2 protein to MSH3 seemed to be abolished.

MMR protein levels can greatly influence mismatch repair functions. For example, we found that tenfold-reduced wild-type MSH2 protein level in *Msh2*$^{low/hyg}$ cells we previously generated, resulted in a defective DNA damage response while repair of mismatches and prevention of homeologous recombination were intact (Claij and Te Riele, 2002). Since the MSH2/MSH6 protein levels in *Msh2*$^{low/hyg}$ ESCs were similar to those in homozygous *Msh2*$^{P622L/P622L}$ cells, we used this cell line for comparison to identify possible additional effects caused by this mutation.

To examine the mutator phenotype of mutant cells, single ESCs were expanded to 10^9 cells. Subsequently, the frequency of unrepaired slippage errors at three different dinucleotide repeats (MSI) was measured in 60 individual cell clones. Additionally, the expanded cell cultures were exposed to 6-thioguanine (6-TG) to measure the frequency of mutational inactivation of the *Hprt* gene. When we tested *Msh2*$^{Y103C/Y103C}$, *Msh2*$^{G322D/G322D}$ and *Msh2*$^{M688I/M688I}$ cells in the MSI and *Hprt* assays, none or very few mutations were detected (Fig. 4). In contrast, both assays showed a strong mutator phenotype for the *Msh2*$^{P622L/P622L}$ mutant cells, similar to *Msh2*$^{hyg/hyg}$ cells (Fig. 4). Because *Msh2*$^{low/hyg}$ cells did not reveal a strong mutator phenotype, this effect was apparently not caused by the reduction in protein levels, indicating that the P622L substitution abolished repair of mismatches.

To assess a second function of MMR, suppression of homeologous recombination, we compared the targeting efficiency of two *Rb* targeting vectors containing a puromycin resistance cassette in wild-type and *Msh2* mutant cells (Claij and Te Riele, 2002). One vector was derived from the 129 mouse strain, being 100% homologous to our ES cells. The other was derived from the Balb/c genomic sequence and 99.4% homologous. Targeting by the

latter was strongly impaired in MMR proficient cells, while *Msh2*$^{hyg/hyg}$ cells allowed targeting of both vectors with equal efficiency (de Wind, et al., 1995). As shown in Figure 5, the MSH2-Y103C, MSH2-G322D and MSH2-M688I proteins suppressed incorporation of the non-identical construct as effectively as wild-type MSH2, in contrast to the MSH2-P622L protein, which allowed effective recombination of both constructs.

A third function of MMR is related to the toxicity of certain DNA damaging compounds, typically methylating agents and 6-thioguanine (6-TG). In this assay, ESCs were plated at low density and exposed for one hour to various concentrations of either N-methyl-N'-nitro-N-nitrosoguanidine (MNNG) or 6-TG. After four days, surviving colonies were counted. *Msh2*$^{Y103C/Y103C}$, *Msh2*$^{G322D/G322D}$ and *Msh2*$^{M688I/M688I}$ cells displayed a strong sensitivity to both agents, similarly to wild-type cells (Fig. 6A and B). In contrast, *Msh2*$^{P622L/P622L}$ cells were as resistant to MNNG as *Msh2*$^{hyg/hyg}$ and *Msh2*$^{low/hyg}$ cells. *Msh2*$^{P622L/P622L}$ cells were also resistant to 6-TG, similar to *Msh2*$^{hyg/hyg}$ cells (Fig. 6A and B). Since *Msh2*$^{low/hyg}$ cells displayed an intermediate sensitivity to 6-TG, this result indicated that the MSH2-P622L mutant was intrinsically defective in mediating cell death in response to MNNG and 6-TG.

Discussion

Here we present a novel approach to evaluate the functional consequences of missense mutations in the MMR gene *MSH2*, which makes use of oligonucleotide-directed gene modification in mouse embryonic stem cells. Our results classify the P622L mutation as deleterious as it was defective in all functional assays conducted. This is in agreement with results from computational modeling (Chao, et al., 2008) and clinical observations such as early age of onset and perfect co-segregation of mutation and disease. Also, MSI and the

absence of MSH2/MSH6 staining using immunohistochemistry (IHC) in tumors is in accordance with our data (Chialina, et al., 2006; Jenkins, et al., 2002; Leach, et al., 1993; Southey, et al., 2005). Similar to studies in yeast expressing the equivalent mutant protein, yMSH2-P640L (Drotschmann, et al., 1999; Polaczek, et al., 1998), we observed a strong mutator phenotype in *Msh2*$^{P622L/P622L}$ ESCs. However, with respect to the response to DNA damaging agents, our study and a study in yeast appeared contradictory. We found ESCs expressing MSH2-P622L mutant protein to be resistant to methylating agents, whereas yeast cells expressing yMSH2-P640L protein were sensitive to cisplatin (Clodfelter, et al., 2005). This discrepancy may be explained by differences in protein level as in yeast a high level of yMSH2-P640L mutant protein was obtained due to ectopic expression, whereas we observed that the level of MSH2-P622L mutant protein expressed from the endogenous locus was markedly reduced. We observed this reduction in two independent mutants, excluding an effect of codon usage. Since the level of mutant mRNA was not affected and no additional mutations were detected in the transcript, we envisage that the P622L substitution renders MSH2 unstable, likely due to steric hindrance of the leucin residue. The remaining MSH2-P622L protein was able to stabilize MSH6, indicating normal complex formation, and this complex seemed to be capable of binding MLH1 (Supplementary Fig. 3). The phenotype of the residual MSH2-P622L/MSH6 complex was not merely caused by low protein level as *Msh2*$^{P622L/P622L}$ cells were defective in all three functional assays whereas *Msh2*$^{low/hyg}$ cells, producing wild-type protein at similarly low levels, were defective only in the DNA damage assay. One may argue that the MSH2-P622L mutant protein fails to elicit sensitivity to MNNG due to low protein level. However, this is unlikely as a similar low level of wild-type MSH2 did partially sensitize cells to 6-TG whereas the MSH2-P622L mutant protein did not. The intrinsic defect caused by the P622L substitution may also impede nuclear localization

of MSH2-P622L (Lutzen, et al., 2008), contributing to its inability to sustain mismatch repair.

Also MSH2-Y103C was present at reduced level, albeit higher than MSH2-P622L. As we found normal mRNA levels and no evidence for aberrant splicing (similar to the human equivalent (Lastella, et al., 2006)), the Y103C substitution likely affected protein stability. This may be due to loss of a hydrogen bond and the creation of a vacuum in a hydrophobic core. Apparently, this did not affect the affinity for MSH6 as we observed similar protein level as in wild-type cells. Most importantly, the MMR capacity of *Msh2*$^{Y103C/Y103C}$ cells was virtually identical to that of *Msh2*$^{+/+}$ cells in all assays in accordance with the absence of a mutator phenotype in yeast expressing the equivalent yMSH2-Y109C (Clodfelter, et al., 2005). A recent report also showed normal MMR capacity of yMSH2-Y109C/MSH6, although reduced MMR activity of the yMSH2-Y109C/MSH3 was detected (Martinez and Kolodner, 2010). Interestingly, we did notice a slightly reduced level of MSH3, which may be indicative for reduced MSH2-Y103C/MSH3 activity. However, since MSH2-Y103C-expressing cells showed normal MMR activity and sensitivity to methylating agents, and MSH3 deficiency did not predispose to tumorigenesis in mice (de Wind, et al., 1999), we believe that this mutation is unlikely to cause cancer predisposition.

We classify the G322D mutation as a polymorphism as it retained functionality in all assays. Although this mutation is generally regarded as a polymorphism, clinical, computational and functional analyses have shown contradictory results. The mutation has been found both in CRC patients and in healthy controls with an allele frequency of 0.012 (http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?rs=4987188). Statistical analysis of 3,785 CRC cases and matched controls has not revealed an association between the MSH2-G322D mutation and an increased CRC risk (Chao, et al., 2008). However, besides microsatellite-stable also microsatellite-unstable tumors have been found in *MSH2*^{G322D} carriers, indicative

of a MMR defect (Lucci-Cordisco, et al., 2006). In some cases IHC failed to detect MSH2, although in the majority of tumors in which only the G322D mutation was found, MSH2 protein levels were normal (Lucci-Cordisco, et al., 2006). In yeast, the functionality of the equivalent yMSH2-G317D mutant protein appeared to depend on high ectopic expression levels (Drotschmann, et al., 1999; Ellison, et al., 2001). It should be noted however that Drotschmann *et al.* judged the alignments insufficient to establish unequivocally that Gly317 in yMSH2 is the functional equivalent of Gly322 in hMSH2 (Drotschmann, et al., 1999). In mouse MSH2, Gly322 is in the middle of a stretch of 30 amino acids that is 100% identical to the human sequence. We demonstrate here that the MSH2-G322D protein expressed from its endogenous locus behaved like wild-type MSH2. We therefore envisage that MSI and negative MSH2 IHC occasionally observed in tumors from $MSH2^{G322D}$ carriers was due to another MMR defect.

Finally, our approach also classifies the MSH2-M688I mutation as a polymorphism as MSH2-M688I showed normal interaction with MSH3, MSH6 and MLH1 and $Msh2^{M688I/M688I}$ cells were MMR proficient in all assays conducted. This was also concluded by Banno *et al.* who detected the mutation in 3 out of 179 healthy individuals (Banno, et al., 2004). The mutation has been found in suspected HNPCC families, however, no co-segregation data has been published (Banno, et al., 2003; Nomura, et al., 2000; Yuan, et al., 1998). Our observations did not confirm computational modeling using MAPP-MMR, which scored this mutation as pathogenic (Chao, et al., 2008). In yeast, the equivalent MSH2-M707I caused an intermediate mutator phenotype, which was tenfold lower than in *Msh2*-deficient cells. The interaction between yMSH2-M707I and MSH3 and MSH6 appeared normal, however, interaction with MLH1, PMS1, EXOI and POL30 could not be detected in a yeast-2-hybrid assay (Gammie, et al., 2007). Thus, substitutions of evolutionary-conserved

amino acids in distantly-related species may not necessarily have identical phenotypic consequences.

With the advance of massive sequencing of the human genome, there is an increasing need for the phenotypic characterization of allelic variants of disease-causing genes. The procedure we present here is the first that makes use of oligo targeting, a novel and effective technique to directly mutate endogenous genes in ESCs. Our study demonstrates that oligo targeting can be efficiently used to study the effects of missense mutations in genes involved in familial disease. A prominent feature of our approach is that we recreate the human situation by mutating the endogenous gene, thereby ensuring normal regulation of expression. Notably, this allows us to study the phenotypic consequences of the mutation at the cellular level using three functional assays each assessing a different MMR function. Furthermore, in cases where the functional implications are less clear, or when a mutation only partially affects MMR capacity, mice can be generated from mutant ESCs to investigate whether attenuated MMR activity promotes tumorigenesis *in vivo*. Our approach will directly contribute to the correct risk assessment of MMR missense mutations found in (suspected) HNPCC families. This will facilitate proper counseling and treatment, especially in cases where altered drug response is observed. Our procedure can be adapted to provide similar insights into the effects of missense mutations found in other genes involved in genetic disease.

Acknowledgements

We thank our colleagues Marieke Aarts, Sietske Bakker, Koen van de Wetering and Rob Wolthuis for discussions and valuable comments on the manuscript and Titia Sixma for help with structural analysis. We thank Romana Nijman and Ron Schackmann for technical

assistance and Roelof Pruntel for assistance with microsatellite analysis. This work was supported by the Dutch Cancer Society (grant number NKI 2004-3084).

References

Aarts M, Dekker M, de Vries S, van der Wal A, te Riele H. 2006. Generation of a mouse mutant by oligonucleotide-mediated gene modification in ES cells. Nucleic Acids Res 34(21):e147.

Banno K, Susumu N, Hirao T, Yanokura M, Hirasawa A, Aoki D, Udagawa Y, Sugano K, Nozawa S. 2003. Identification of germline MSH2 gene mutations in endometrial cancer not fulfilling the new clinical criteria for hereditary nonpolyposis colorectal cancer. Cancer Genet Cytogenet 146(1):58-65.

Banno K, Susumu N, Nozawa S, Sugano K. 2004. Met688Ile and Leu390Phe of the MSH2 gene are not functional mutations, but polymorphisms in Japanese individuals. Cancer Genet Cytogenet 155(1):92.

Chan PA, Duraisamy S, Miller PJ, Newell JA, McBride C, Bond JP, Raevaara T, Ollila S, Nystrom M, Grimm AJ, Christodoulou J, Oetting WS, Greenblatt MS. 2007. Interpreting missense variants: comparing computational methods in human disease genes CDKN2A, MLH1, MSH2, MECP2, and tyrosinase (TYR). Hum Mutat 28(7):683-93.

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):852-60.

Chialina SG, Fornes C, Landi C, de la Vega Elena CD, Nicolorich MV, Dourisboure RJ, Solano A, Solis EA. 2006. Microsatellite instability analysis in hereditary non-polyposis colon cancer using the Bethesda consensus panel of microsatellite markers in the absence of proband normal tissue. BMC Med Genet 7:5.

Claij N, Te Riele H. 2002. Methylation tolerance in mismatch repair proficient cells with low MSH2 protein level. Oncogene 21(18):2873-9.

Clodfelter JE, Gentry MB, Drotschmann K. 2005. MSH2 missense mutations alter cisplatin cytotoxicity and promote cisplatin-induced genome instability. Nucleic Acids Res 33(10):3323-30.

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):321-30.

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):359-62.

de Wind N, Dekker M, van Rossum A, van der Valk M, te Riele H. 1998. Mouse models for hereditary nonpolyposis colorectal cancer. Cancer Res 58(2):248-55.

Drotschmann K, Clark AB, Kunkel TA. 1999. Mutator phenotypes of common polymorphisms and missense mutations in MSH2. Curr Biol 9(16):907-10.

Ellison AR, Lofing J, Bitter GA. 2001. Functional analysis of human MLH1 and MSH2 missense variants and hybrid human-yeast MLH1 proteins in Saccharomyces cerevisiae. Hum Mol Genet 10(18):1889-900.

Froggatt NJ, Joyce JA, Evans DG, Lunt PW, Koch DJ, Ponder BJ, Maher ER. 1996. MSH2 sequence variations and inherited colorectal cancer susceptibility. Eur J Cancer 32A(1):178.

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(2):707-21.

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(11):6616-23.

Jenkins MA, Baglietto L, Dite GS, Jolley DJ, Southey MC, Whitty J, Mead LJ, St John DJ, Macrae FA, Bishop DT, Venter DJ, Giles GG, Hopper JL. 2002. After hMSH2 and hMLH1--what next? Analysis of three-generational, population-based, early-onset colorectal cancer families. Int J Cancer 102(2):166-71.

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

Lastella P, Surdo NC, Resta N, Guanti G, Stella A. 2006. In silico and in vivo splicing analysis of MLH1 and MSH2 missense mutations shows exon- and tissue-specific effects. BMC Genomics 7:243.

Leach FS, Nicolaides NC, Papadopoulos N, Liu B, Jen J, Parsons R, Peltomaki P, Sistonen P, Aaltonen LA, Nystrom-Lahti M, Guan X-Y, Zhang J, Meltzer PS, Yu J-W, Kao F-T, Chen DJ, Cerosaletti KM, Keith Fournier RE, Todd S, Lewis T, Leach RJ, Naylor SL, Weissenbach J, Mecklin J-P, Järvinen H, Petersen GM, Hamilton SR, Green J, Jass J, Watson P, Lynch HT, Trent JM, de la Chapelle A, Kinzler KW, Vogelstein B. 1993. Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. Cell 75(6):1215-25.

Lindor NM, Petersen GM, Hadley DW, Kinney AY, Miesfeldt S, Lu KH, Lynch P, Burke W, Press N. 2006. Recommendations for the care of individuals with an inherited predisposition to Lynch syndrome: a systematic review. JAMA 296(12):1507-17.

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(1-2):11-27.

Lutzen 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.

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

Martinez SL, Kolodner RD. 2010. Functional analysis of human mismatch repair gene mutations identifies weak alleles and polymorphisms capable of polygenic interactions. Proc Natl Acad Sci U S A 107(11):5070-5.

Mortensen RM, Conner DA, Chao S, Geisterfer-Lowrance AA, Seidman JG. 1992. Production of homozygous mutant ES cells with a single targeting construct. Mol Cell Biol 12(5):2391-5.

Nomura S, Sugano K, Kashiwabara H, Taniguchi T, Fukayama N, Fujita S, Akasu T, Moriya Y, Ohhigashi S, Kakizoe T, Sekiya T. 2000. Enhanced detection of deleterious and other germline mutations of hMSH2 and hMLH1 in Japanese hereditary nonpolyposis colorectal cancer kindreds. Biochem Biophys Res Commun 271(1):120-9.

Ou J, Niessen RC, Lutzen A, Sijmons RH, Kleibeuker JH, de Wind N, Rasmussen LJ, Hofstra RM. 2007. Functional analysis helps to clarify the clinical importance of unclassified variants in DNA mismatch repair genes. Hum Mutat 28(11):1047-54.

Polaczek P, Putzke AP, Leong K, Bitter GA. 1998. Functional genetic tests of DNA mismatch repair protein activity in Saccharomyces cerevisiae. Gene 213(1-2):159-67.

Southey MC, Jenkins MA, Mead L, Whitty J, Trivett M, Tesoriero AA, Smith LD, Jennings K, Grubb G, Royce SG, Walsh MD, Barker MA, Young JP, Jass JR, St John DJ, Macrae FA, Giles GG, Hopper JL. 2005. Use of molecular tumor characteristics to prioritize mismatch repair gene testing in early-onset colorectal cancer. J Clin Oncol 23(27):6524-32.

te Riele H, Maandag ER, Berns A. 1992. Highly efficient gene targeting in embryonic stem cells through homologous recombination with isogenic DNA constructs. Proc Natl Acad Sci U S A 89(11):5128-32.

te Riele H, Maandag ER, Clarke A, Hooper M, Berns A. 1990. Consecutive inactivation of both alleles of the pim-1 proto-oncogene by homologous recombination in embryonic stem cells. Nature 348(6302):649-51.

Ward R, Meldrum C, Williams R, Mokany E, Scott R, Turner J, Hawkins N, Burgess B, Groombridge C, Spigelman A. 2002. Impact of microsatellite testing and mismatch repair protein expression on the clinical interpretation of genetic testing in hereditary non-polyposis colorectal cancer. J Cancer Res Clin Oncol 128(8):403-11.

Yuan Y, Han HJ, Zheng S, Park JG. 1998. Germline mutations of hMLH1 and hMSH2 genes in patients with suspected hereditary nonpolyposis colorectal cancer and sporadic early-onset colorectal cancer. Dis Colon Rectum 41(4):434-40.

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(13):3021-7.

Figure legends

Figure 1. Generation of *Msh2* mutant ESC lines by oligo targeting.

(A) *Msh2* targeting oligonucleotides (upper case) hybridized to their complementary genomic sequence (lower case); mismatching bases in the oligonucleotides are shown in red. **(B)** Sequence analysis of $Msh2^{Y103C/+}$cDNA, **(C)** $Msh2^{G322D/+}$ cDNA, **(D)** $Msh2^{P622L/hyg}$ cDNA and **(E)** $Msh2^{M688I/+}$ cDNA.

Figure 2. Generation of homozygous *Msh2* mutant ESC lines

(A) Sequence analysis of $Msh2^{+/+}$, $Msh2^{Y103C/+}$ and $Msh2^{Y103C/Y103C}$ genomic DNA. **(B)** Sequence analysis of $Msh2^{+/+}$, $Msh2^{M688I/+}$ and $Msh2^{M688I/M688I}$ genomic DNA. **(C)** Restriction enzyme analysis of $Msh2^{G322D}$ mutant cDNAs. Only the mutant sequence is cut by the *Btg1* enzyme resulting in two bands of 374 and 248 bp. m indicates a 100 bp DNA marker. **(D)** Southern blot analysis of the $Msh2^{P622L/hyg}$ (left lanes of each panel) and $Msh2^{P622L/P622L}$ (right lanes of each panel) cell lines, showing duplication of the $Pim1^{neo}$ allele in the left panel and loss of the $Msh2^{hyg}$ allele in the right panel, upon exposure to 9 mg/ml G418. The position of both genes on chromosome 17 is indicated.

Figure 3. Western blot analysis of mutant cell lines and controls. Whole cell lysates were analyzed for MSH2, MSH3 and MSH6. γ-Tubulin was used as a loading control. Hyg and pur indicate knockout alleles.

Figure 4. Mutator phenotype in *Msh2* mutant ESC lines.

Black bars show the average percentage of unstable microsatellites (left Y-axis) as measured in 60 colonies for 3 different markers. The grey bars show the average number of 6-TG

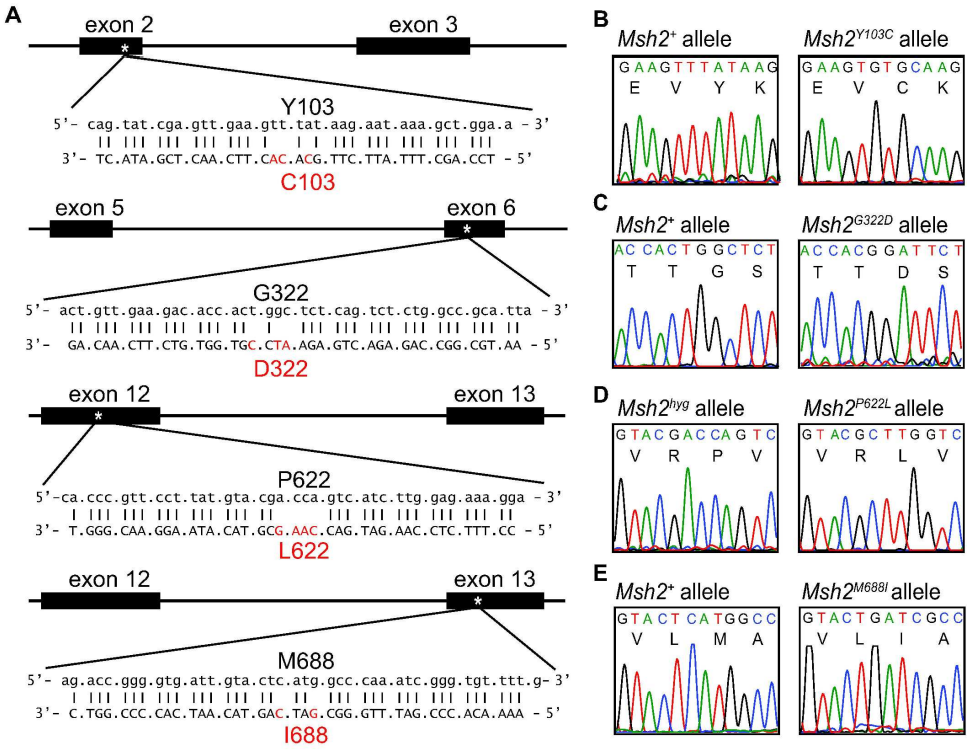
resistant colonies per 10^6 plated cells (right Y-axis). Error bars show standard errors, measured over 3-4 experiments per cell line.

Figure 5. Prevention of homeologous recombination in *Msh2* mutant ESC lines.

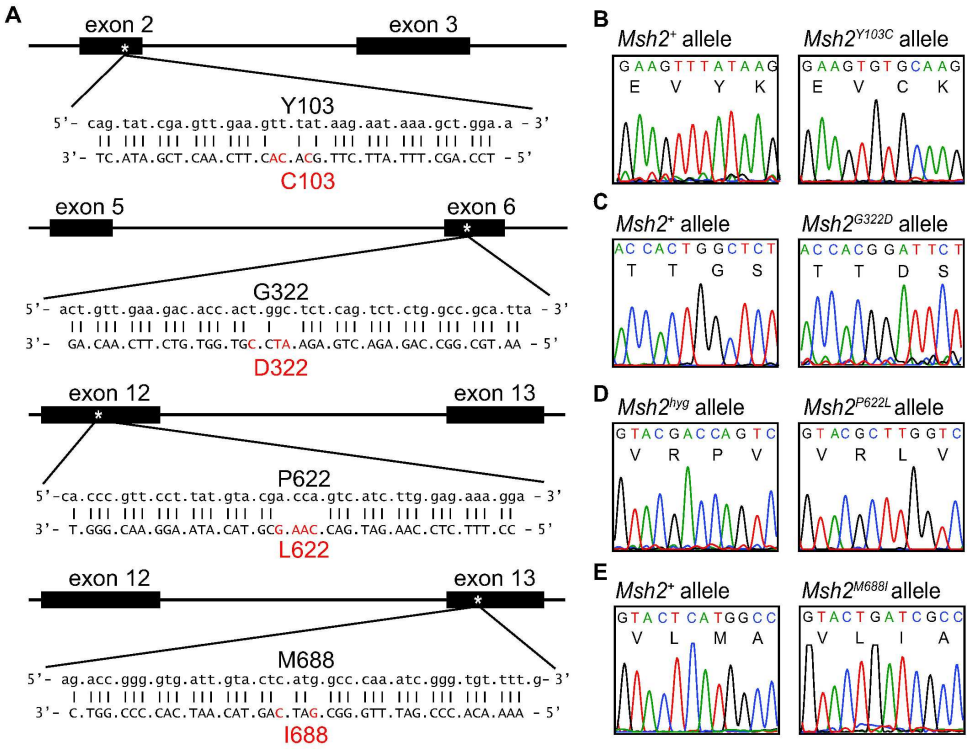
Targeting efficiencies are shown in mutant and control cell lines for the 129 (black bars) and the Balb/c (white bars) derived *Rb* targeting constructs. Targeting efficiencies in $Msh2^{+/+}$ and $Msh2^{hyg/hyg}$ are taken from (de Wind, et al., 1995) and shown as controls.

Figure 6. Toxicity of methylating agents in *Msh2* mutant ESC lines.

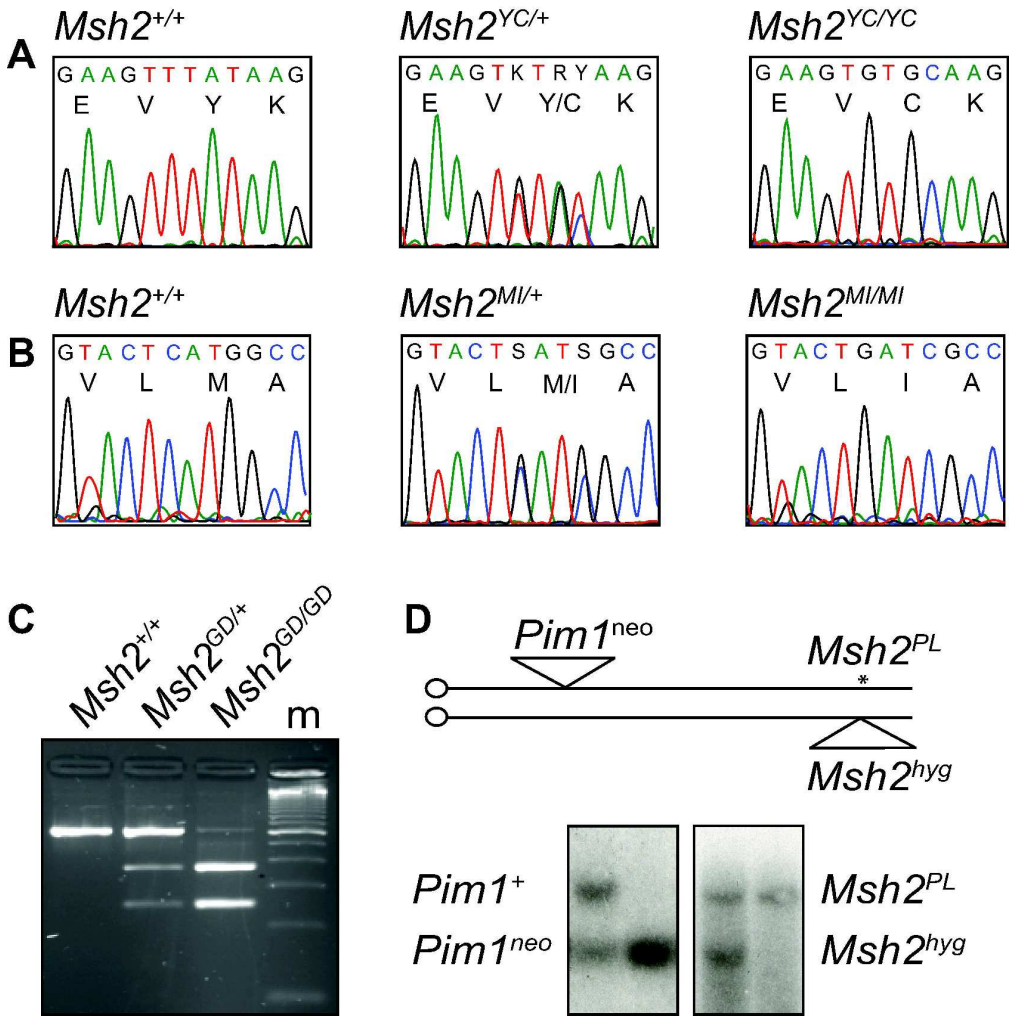
(A) Tolerance of mutant and control cell lines to MNNG (n=4-8). **(B)** Tolerance of mutant and control cell lines to 6-TG (n=3-5). Error bars show standard errors from independent experiments.



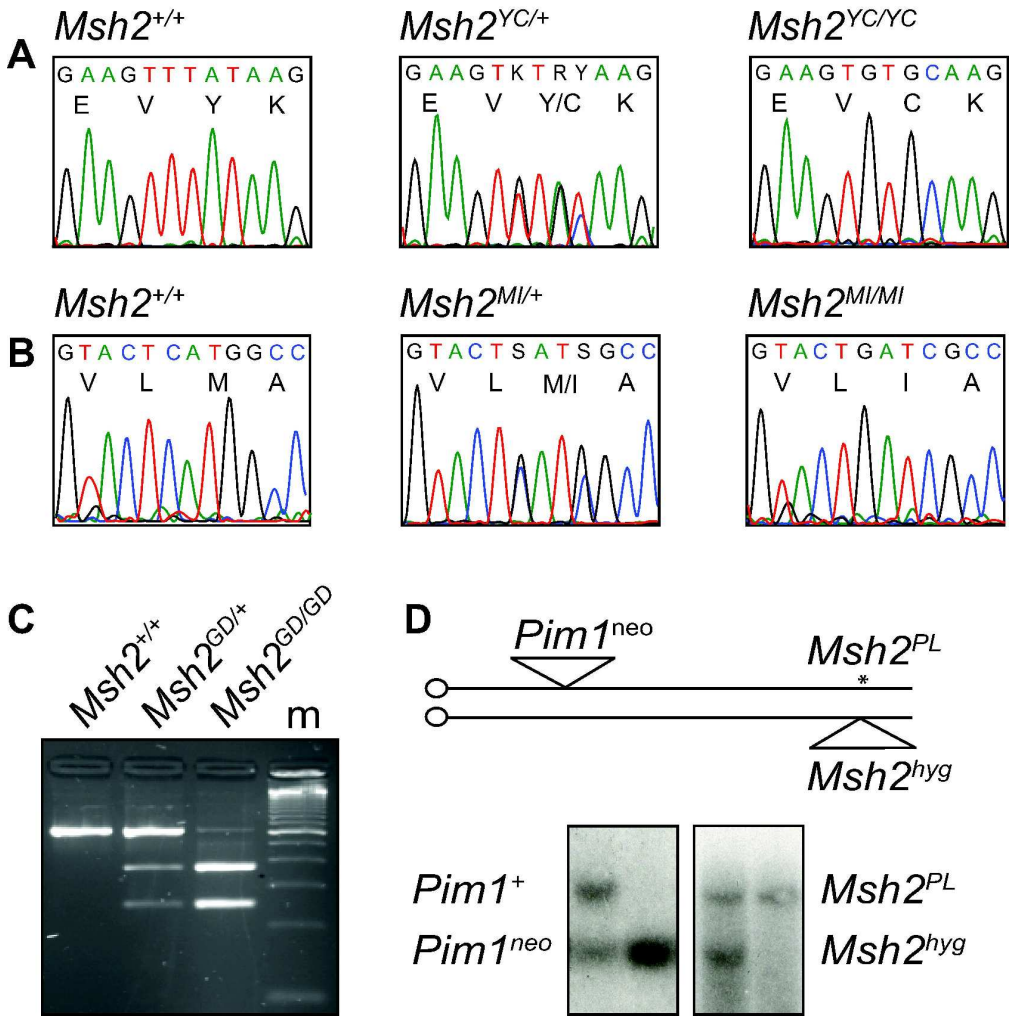
185x138mm (600 x 600 DPI)



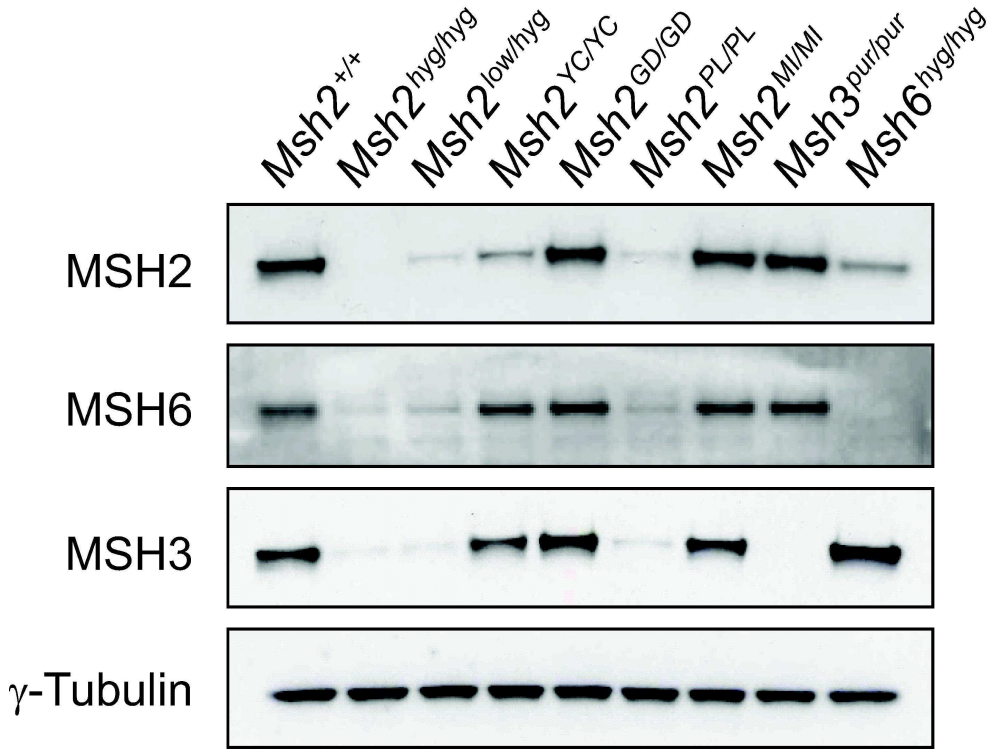
185x138mm (600 x 600 DPI)



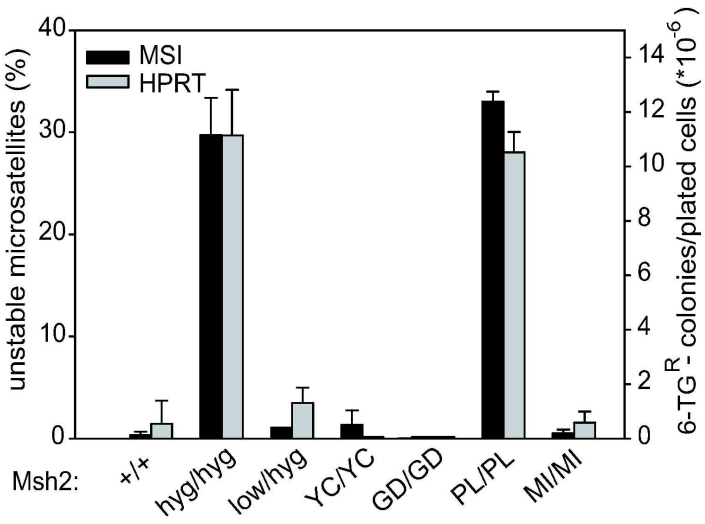
142x145mm (600 x 600 DPI)



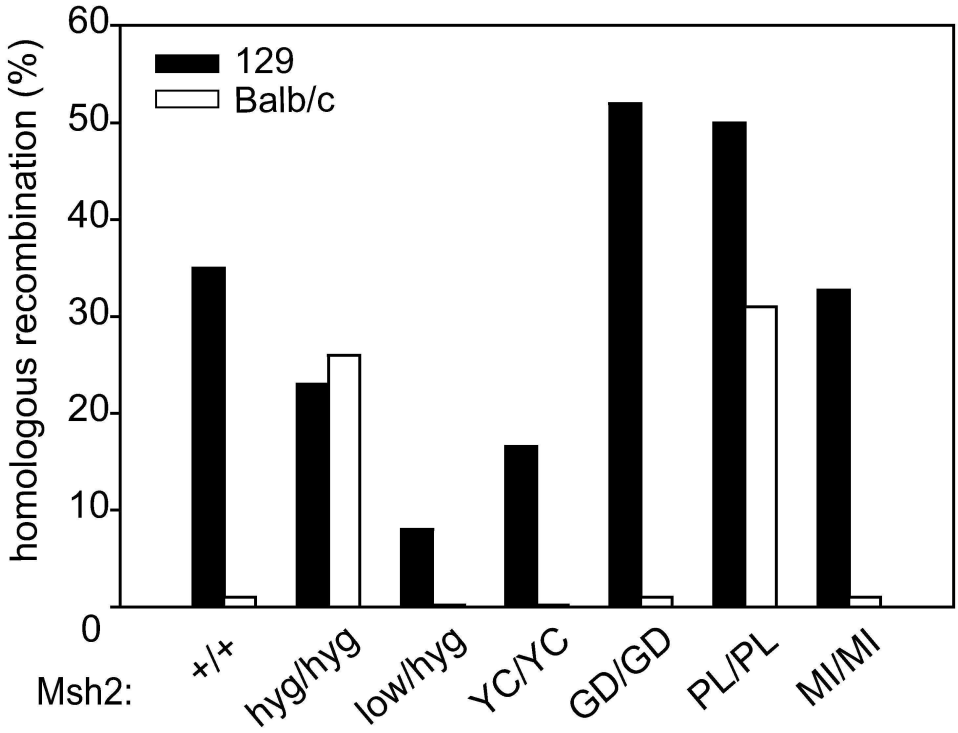
142x145mm (600 x 600 DPI)



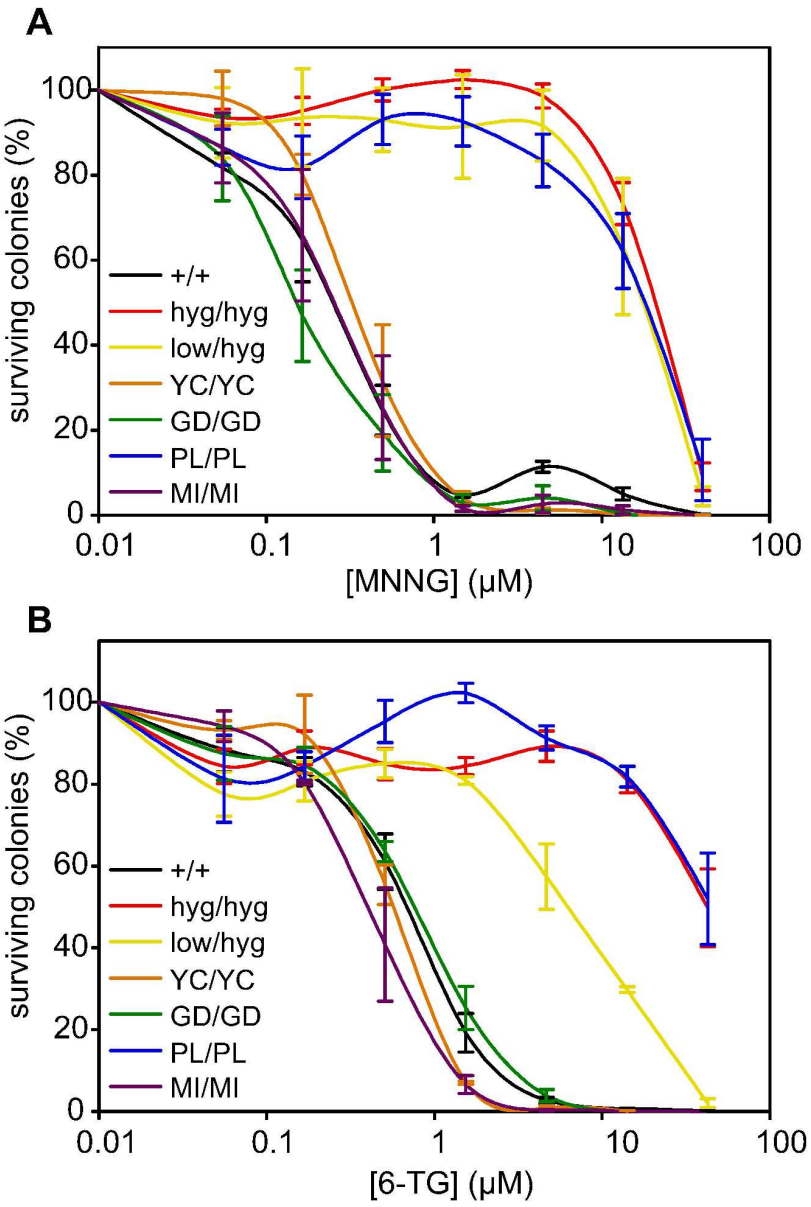
122x92mm (600 x 600 DPI)



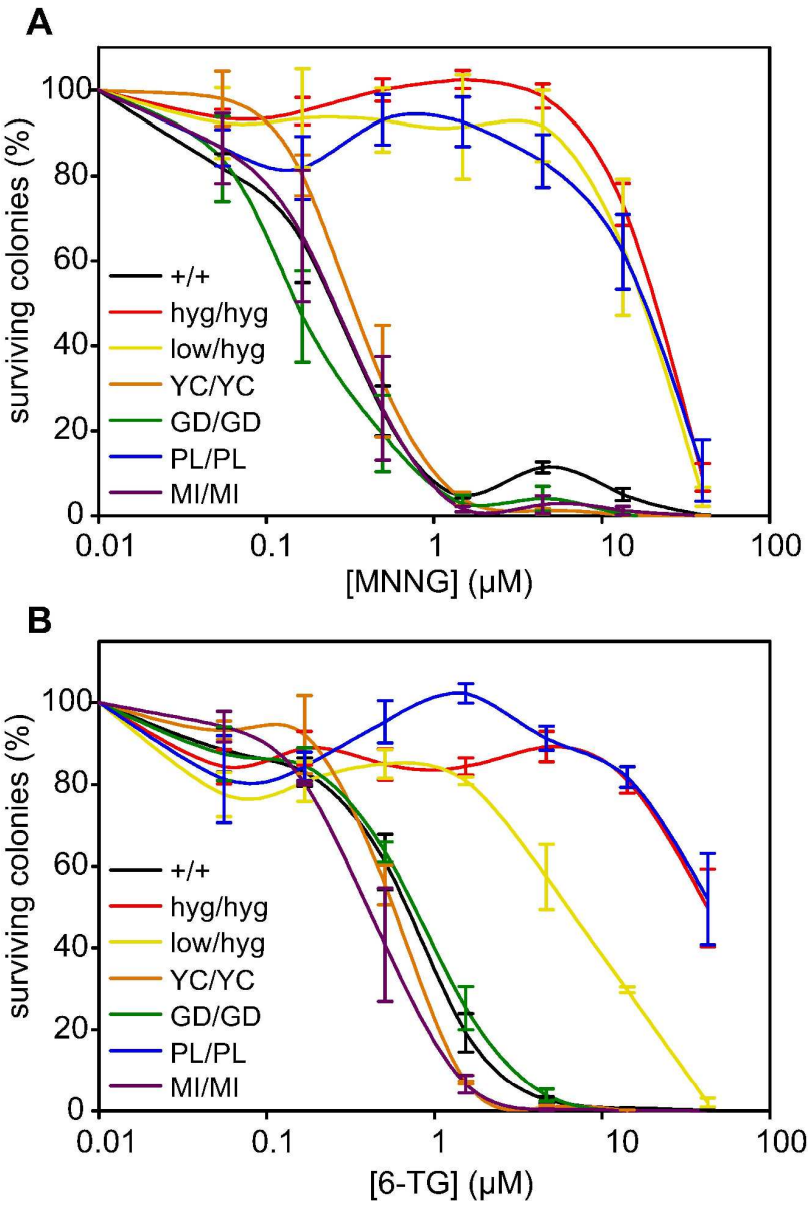
182x94mm (600 x 600 DPI)



129x98mm (600 x 600 DPI)

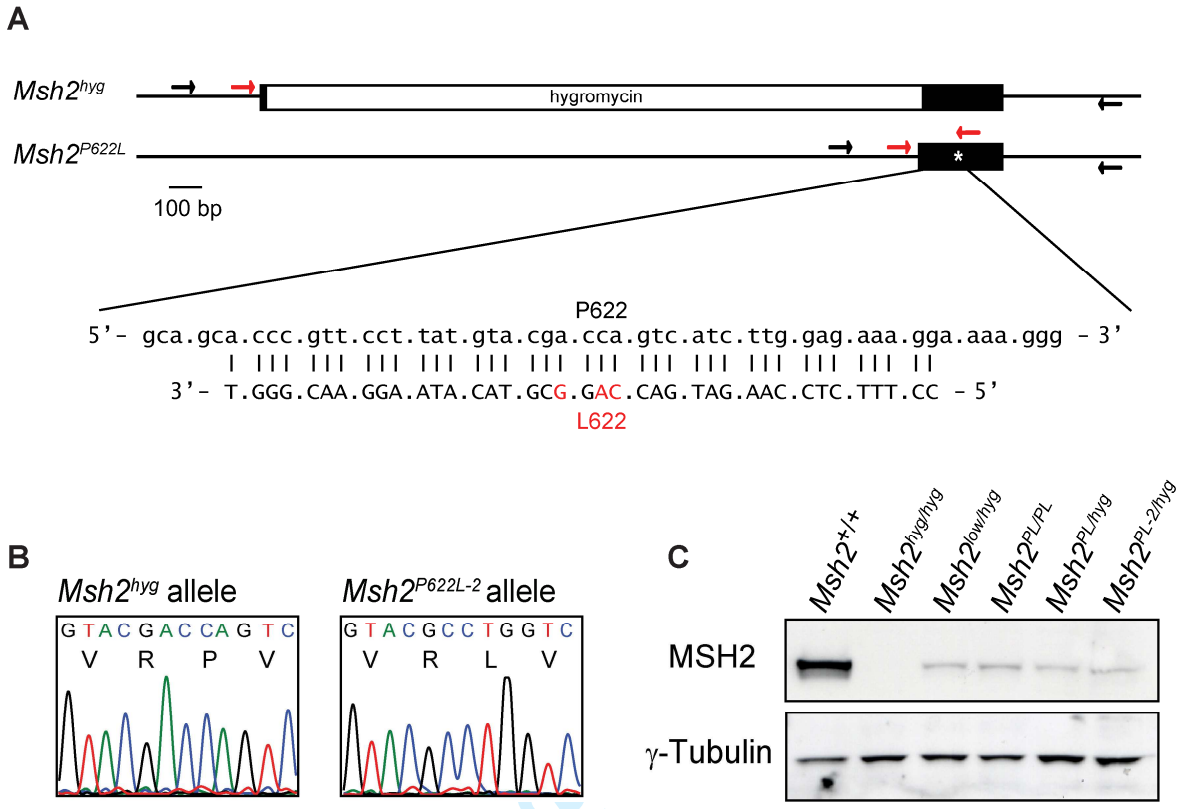


112x167mm (600 x 600 DPI)



112x167mm (600 x 600 DPI)

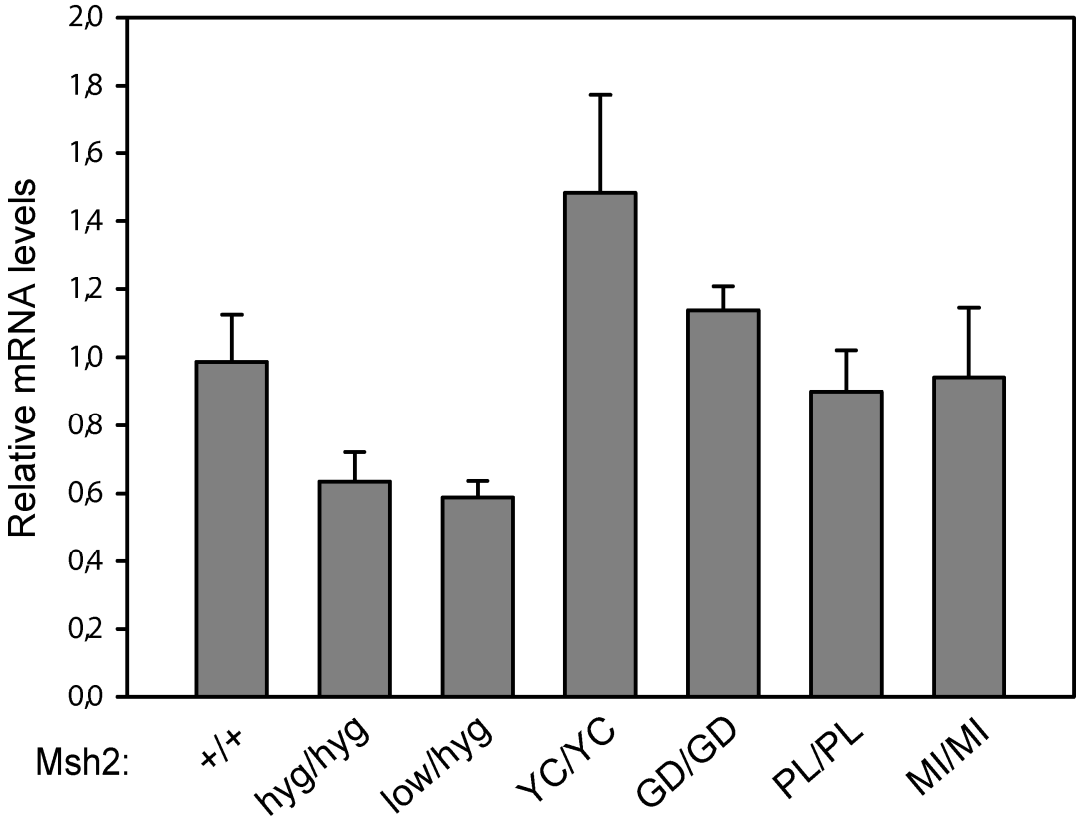
Supplementary Figure 1



Supplementary Figure 1. Generation of a *Msh2*^{P622L} mutant with enhanced codon usage.

(A) The P622L-2 oligonucleotide (upper case) hybridized to its complementary genomic sequence (lower case); mismatching bases in the oligonucleotide are shown in red. Due to the insertion of a hygromycin marker gene in exon 12 of the *Msh2*^{hyg} allele, the mutation specific PCR will only result in a product if the mutation, indicated by a white asterisk in exon 12, is on the *Msh2*$^{+}$ allele. Black arrows indicate the position of the primers used for the first PCR whereas red arrows depict the primers for the nested PCR. **(B)** Sequence analysis of *Msh2* cDNA shows replacement of the proline codon, at position 622, for leucine in the *Msh2*$^{P622L-2}$ mutant. **(C)** Western blot analysis of the *Msh2*^{P622L} mutant cell lines, the *Msh2*$^{P622L-2/hyg}$ cell line and controls. Whole cell lysates were analyzed for MSH2 and γ-Tubulin was used as a loading control.

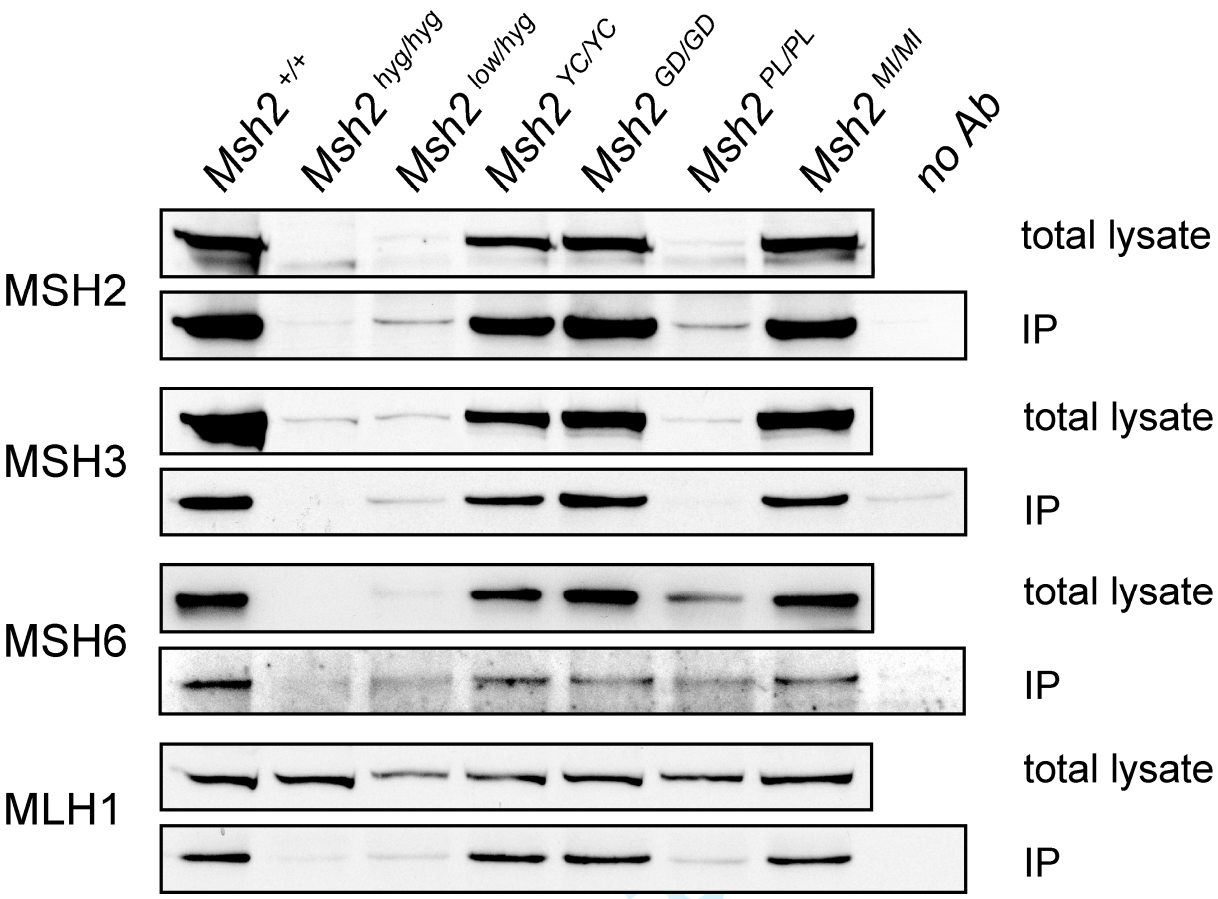
Supplementary Figure 2



Supplementary figure 2. Quantitative real time PCR analysis of mutant *Msh2* mRNAs

Relative mRNA levels of mutant and control *Msh2* mRNA. Hyg indicates a knockout allele. Error bars show standard errors from 2 independent experiments.

Supplementary Figure 3



Supplementary figure 3. MSH2 immunoprecipitations of mutant cell lines and controls

Western blot analysis of total lysate and immunoprecipitated (IP) fractions of MSH2 immunoprecipitations, blotted for MSH2, MSH3, MSH6 and MLH1. Hyg indicates a knockout allele. No Ab indicates the lysate of the wild type cells incubated with protein G but without the MSH2 antibody.