Understanding the regulation of endogenous mutagenesis.

AMRUTA MUKUND SHRIKHANDE

DPHIL BIOCHEMISTRY

UNIVERSITY OF SUSSEX

SUBMITTED SEPTEMBER 2018

Declaration

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another University for the award of any other degree.

Signature:……

Acknowledgements

First of all, I would like to thank my supervisor Prof. Eva Hoffmann for her support and guidance throughout my PhD. I am deeply grateful for her being so patient and encouraging through all these years and helping me in all possible ways.

I would also like to thank my co-supervisor Dr Francis Pearl, and my thesis committee members Prof Simon Ward and Prof Mark O'Driscoll for their valuable feedback. I am also very thankful to Dr Javier Peña Diaz for all the scientific discussion, encouraging words and providing reagents.

I am also very grateful to all past and present members of the Hoffmann lab. Special thanks to Jacob for teaching me yeast techniques, Jenny for all the fantastic English lessons, Andy for always answering all my random and scientific questions, and Xavi for always being so helpful with everything.

I would also like to thank my friends Sarika, Lorenza, Signe, Daniela and Philipp for their help throughout my PhD and making life in Copenhagen much brighter!

Finally, I would like to thank my family and friends back home, for their patience and encouragement. Thank you Aai, Baba and Didi, I could have not done this without you being so supportive and encouraging through all these years!

UNIVERSITY OF SUSSEX

AMRUTA MUKUND SHRIKHANDE

DPHIL BIOCHEMISTRY

UNDERSTANDING THE REGULATION OF ENDOGENOUS MUTAGENESIS

SUMMARY

There are a number of cellular mechanisms conserved across eukaryotic organisms that prevent or reduce spontaneous mutation rates and maintain genomic stability. Inactivation of processes that reduce the spontaneous mutation rate, can lead to carcinogenesis and ageing. DNA mismatch repair (MMR) proteins increase the fidelity of DNA replication by several orders of magnitude by correcting mismatches and frameshift mutations generated during DNA replication. Mutations in MMR genes are associated with Lynch Syndrome, a cancer predisposition syndrome characterized by elevated mutation rates.

In order to identify gene deletions which, show additive effect when combined with lower MMR activity (hypomorphic MMR), a genome-wide screen was carried out in *S. cerevisiae*. The genome-wide screen identified more than 163 deletion strains (3.2% of the total number of strains screened), whose deletion resulted in increased mutation rates in MMR compromised cells which suggests their role in compensating for defective MMR. The 163 genes are enriched in multiple biological processes such as DNA repair, DNA replication etc. The screen also revealed 543 deletion strains (10.6% of the total strains screened), which showed a decrease in mutation rate in hypomorphic MMR cells, suggesting mutation rate can also go down when MMR is not functional in cells. The genes resulting in lower mutation rate were enriched in biological processes, including protein metabolism and transport, general cell metabolism and growth. Interestingly, deletion of genes involved in regulation of transcription coupled repair resulted in lower mutation rate, suggesting their potential role in driving mutagenesis.

List of abbreviation

6-TG	6-thioguanine
bp	base pair
CAN	canavanine
CMT	Charcot-Marie-Tooth
CRC	colorectal cancer
CS	Cockayne syndrome
DEF1	RNAPII DEgradation Factor
dNTP	deoxynucleoside triphosphate
DSB	double strand break
E. coli	Escherichia coli
EUROSCARF	European *S. cerevisiae* archive- functional analysis
EXO1	exonuclease 1
ExoVII	exonuclease VII
ExoX,	exonuclease X
FAN1	Fanconi-associated nuclease 1
GFP	green fluorescent protein
GGR	global nucleotide excision repair
H2O2	hydrogen peroxide
HPRT	hypoxanthine phosphoribosyl transferase
HR	homologous recombination
HT	hypoxanthine–thymidine
IDLs	insertion deletion loops
LS	Lynch syndrome

LYS	Lysine
MLH	MutL protein homolog
MMR	Mismatch repair
MRE11	Meiotic recombination 11
MSH	MutS protein homolog
MSI	microsatellite instability
NER	nucleotide excision repair
NPC	Nuclear pore complex
NTP	nucleoside triphosphate
ORF	Open reading frame
PCNA	Proliferating Cell Nuclear Antigen
RFC	Replication factor C
RNA	Ribonucleic acid
S. cerevisiae	*Saccharomyces cerevisiae*
SSB	Single strand break
TCR	Transcription coupled repair
TFIIH	transcription factor II H
TLS	translesion synthesis
TREX-2	TRanscription-EXport complex
UDGs	uracil-DNA glycosylases
UDGs	uracil-DNA glycosylases
UV	Ultraviolet radiation
XP	xeroderma pigmentosum
Δ	Deletion mutant

Index

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

1 Chapter 1 Introduction

1.1 Mutagenesis:

DNA within all living cells suffer constant structural and chemical modifications as a result of exposure to various exogenous chemical or physical agents, such as heat and ionizing radiation, as well as a result of various cellular processes (e.g., cellular respiration, replication, and DNA demethylation)(Smith, 1992). Even though a majority of these modifications seem like small alterations to the DNA structure, e.g., modifying the DNA base by addition of an oxygen or methyl group, if not repaired they often result in serious consequences like cancer or cell death (Olinski et al., 2018). On the other hand, mutagenesis plays a critical role in driving evolution and in promoting variability in immunoglobulin, thus it is important to understand the processes that regulate the spontaneous mutation rate of cells (Smith, 1992).

Originally, exogenous sources were considered as the major sources of DNA damage in cells, but over the years the importance and existence of abundant types of endogenous sources of DNA damage have been recognized. It is highly likely that even without exposure to mutagens every cell harbours at least 50,000 endogenous DNA lesions ever day (Lindahl, 1993).

During replication, the major DNA replicative polymerases Pol δ and Pol ε only introduce non-complementary nucleotides approximately once in every 10,000,000 base pairs replicated due to their intrinsically high specificity and proof reading activity (Reha-Krantz, 2010). This mutation rate is further lowered

to 1 in 10^{-9} to 10^{-10} per nucleotide replicated by a specialized mismatch repair pathway (see section 1.2) (Kunkel & Erie, 2015). However, some DNA lesions are caused by translesion polymerases utilized during the error-prone replication in response to existing lesion on a template stand, known as translesion synthesis (TLS). TLS utilizes specialized polymerases which have large active site and therefore can accommodate damaged or distorted templates however have lower replication fidelity (Goodman & Woodgate, 2013). Moreover, along with polymerase fidelity, the balance of cellular nucleoside triphosphate (NTP) and deoxynucleoside triphosphate (dNTP) pools can greatly influence the mutation rate (Chabes et al., 2003). Studies have shown that some replicative polymerases like pol ε have very low selectivity for dNTPs over rNTPs, which can cause incorporation of rNTPs in the newly synthesized DNA. This incorporation may result in DNA backbone cleavage or strand slippage in the following rounds of replication (Nick McElhinny et al., 2010). In a rare event during replication, U:A mispairing can occur by accidental insertion of uracil instead of thymidine by the replicative polymerase, which leads to the formation of an abasic site upon excision of the uracil by uracil-DNA glycosylases (UDGs) (Olinski et al., 2018). Similarly, spontaneous deamination of cytosine can cause a mutagenic U:G mispairing (Olinski et al., 2018).

The movement of the replication fork can be impaired during replication by the lack of available dNTPs or various physical obstacles such as secondary DNA structures or damaged DNA bases, resulting in a stalled replication fork (Gaillard et al., 2015; Lambert & Carr, 2013). The restarting of replication then depends on the type of obstacle as well as other factors such as the location

and timing of the stalled fork (Lambert & Carr, 2013). If the fork is unable to restart replication, it may collapse which then requires the intervention of a secondary pathway for its processing (Branzei & Foiani, 2007; Paulsen & Cimprich, 2007). Even though we do not know the exact mechanism for the mutagenicity of collapsed replication fork, replicative helicases may continue their action even after the collapse. This continued activity may then lead to the accumulation of ssDNA that is then susceptible to other mutagenic processes (Lambert & Carr, 2013).

Another source of mutations that can lead to significant DNA damage is oxygen-derived species such as superoxide radicals, hydrogen peroxide, singlet oxygen and hydroxyl radicals, together known as reactive oxygen species (Waris & Ahsan, 2006). Along with various environmental factors, cellular respiration is a major source of endogenous ROSs (Olinski et al., 2018). These ROSs constantly alter cellular DNA, both nuclear and mitochondrial, and lead to a variety of types of DNA damage—e.g., a range of oxidized purines and pyrimidines, alkali labile sites, single strand breaks (Breen & Murphy, 1995; Cooke et al., 2003; Dizdaroglu et al., 2002; Jaruga et al., 2004; Rodriguez et al., 1998). All aerobic cells contain ROSs; therefore to protect cells against the oxidative damage all aerobic organisms have established a series of defences. This defence include buffering of ROSs with antioxidant compounds, for example vitamin A or vitamin C in humans (Davies, 2000). Further aerobic organisms are also capable of synthesizing enzymes that can disarm ROS. One of the most common examples is enzymes of super oxide dismutase family. These enzymes can convert the superoxide anion free radical (O^-_2) into

molecular oxygen and hydrogen peroxide (H_2O_2) (McCord & Fridovich, 1969). Disruption of the balance between antioxidants and reactive oxygen species is cytotoxic and can lead to many diseases (Phaniendra et al., 2015).

1.2 Mismatch Repair Pathway.

The accidental observation of higher mutation rates than wild type in some strains of *E.coli* and Salmonella in 1950 provoked various large screens for identification of mutators. Strains were either checked for increased reversion of auxotrophic markers or increased mutation to antibiotic resistance (Liberfarb & Bryson, 1970)

Mismatched base pairs during DNA replication are a unique challenge to repair, as they are comprised entirely of undamaged DNA and are therefore very difficult to distinguish. The proof reading activity of replicative DNA polymerases allows for the removal of mispaired nucleotides from the 3' end of the newly synthesized strand. The misalignment of primer and template causes a dynamic barrier to polymerase, however if the end of the primer is correctly annealed it can escape proof reading by the polymerase (Jiricny, 2013). These errors which escape proofreading are then subject to DNA mismatch repair (MMR), which ultimately keeps the replication error rate less than 10^{-9} to 10^{-10} per nucleotide (Kunkel & Erie, 2015).

Table 1.1 Proteins associated with Mismatch repair		
E.coli	Eukaryotes	Function
MutS	Msh2-Msh6 (MutSα)	recognition of base-base mispairs or 1-2base IDL
	Msh2-Msh3 (MutSβ)	recognition of more than 2 base IDL
MutL	Mlh1-Pms2 (MutLα) Mlh1-Pms1 in *S. cerevisiae*	co-ordinates multilple steps of MMR
	Mlh1-Mlh2 (MutLβ)	
	Mlh1-Mlh3 (MutLγ)	mainly required for Meiosis but might play a role in IDL suppression
MutH	none	incises the newly synthesized unmethylated GATC strand
γ-δ complex	RFC complex	required for PCNA loading
β-clamp	PCNA	required during strand re-synthesis, also important for stand discrimination
Exo I	Exo1	Excision of mismatch containing strand
Exo X	none	
RecJ	none	
ExoVII	none	
DNA pol III	DNA pol δ	strand re-synthesis
SSB	RPA	protecting single stranded DNA
DNA ligase	DNA ligase	ligation

1.2.1 MMR pathway in *E. coli*

MMR in bacteria has been extensively studied and reconstituted *in vitro* using purified proteins by the Modrich group (Grilley et al., 1993). The three main steps in MMR are recognition, excision and gap-filling DNA synthesis. In *E. coli,* MutS recognizes mismatches generated during replication (S. S. Su et al., 1988). The interaction of MutS with the β-clamp accessory protein may help in the recognition of mismatches by MutS. Unlike the eukaryotic homolog of MutS, in bacteria MutS can function as dimers or tetramers (Bjornson et al., 2003).

After recognition of a mis-pair by MutS; the MutS complex undergoes a conformational change and rapidly slide along the DNA to recruit MutL (Gradia et al., 1999; Hingorani, 2016). MutL complex is then required for recruitment and activation of MutH. The recruitment of MutH is dependent on ATP-binding activity but not ATP hydrolysis (Au et al., 1992; Ban & Yang, 1998). The interaction between the C-terminal domain of MutH and the N-terminal domain of MutL can activate the endonuclease activity of MutH partially (Heinze et al., 2009). However for the complete activation ATP hydrolysis and interaction with mismatch bound MutS-MutL complex is required (Au et al., 1992). To correct the newly synthesized strand, this complex first searches for the nearest hemi-methylated GATC sequence (Langle-Rouault et al., 1987). Next, MutH incises the newly synthesized unmethylated GATC strand in an ATP dependant manner (Bruni et al., 1988; Modrich, 1989). The resulting nick is then used as an entry site by DNA exonucleases and helicases to remove the nascent strand that contains the mismatches (Bruni et al., 1988). In bacteria, depending on the location of the nick, multiple 5'-3' and 3'-5' exonucleases (e.g., Exo1, ExoVII,

ExoX, and RecJ) can then remove the DNA mismatch (Burdett et al., 2001). Finally, DNA polymerase III accurately resynthesizes the excised strand and DNA ligation, thus signalling the completion of MMR.

Figure 1.1: The Mismatch repair pathway in *E.Coli*
A) The pathway begins with the recognition of mismatch or IDLs by MutS homodiamer and MutH monomers binding to hemi-methylated GATC sites.

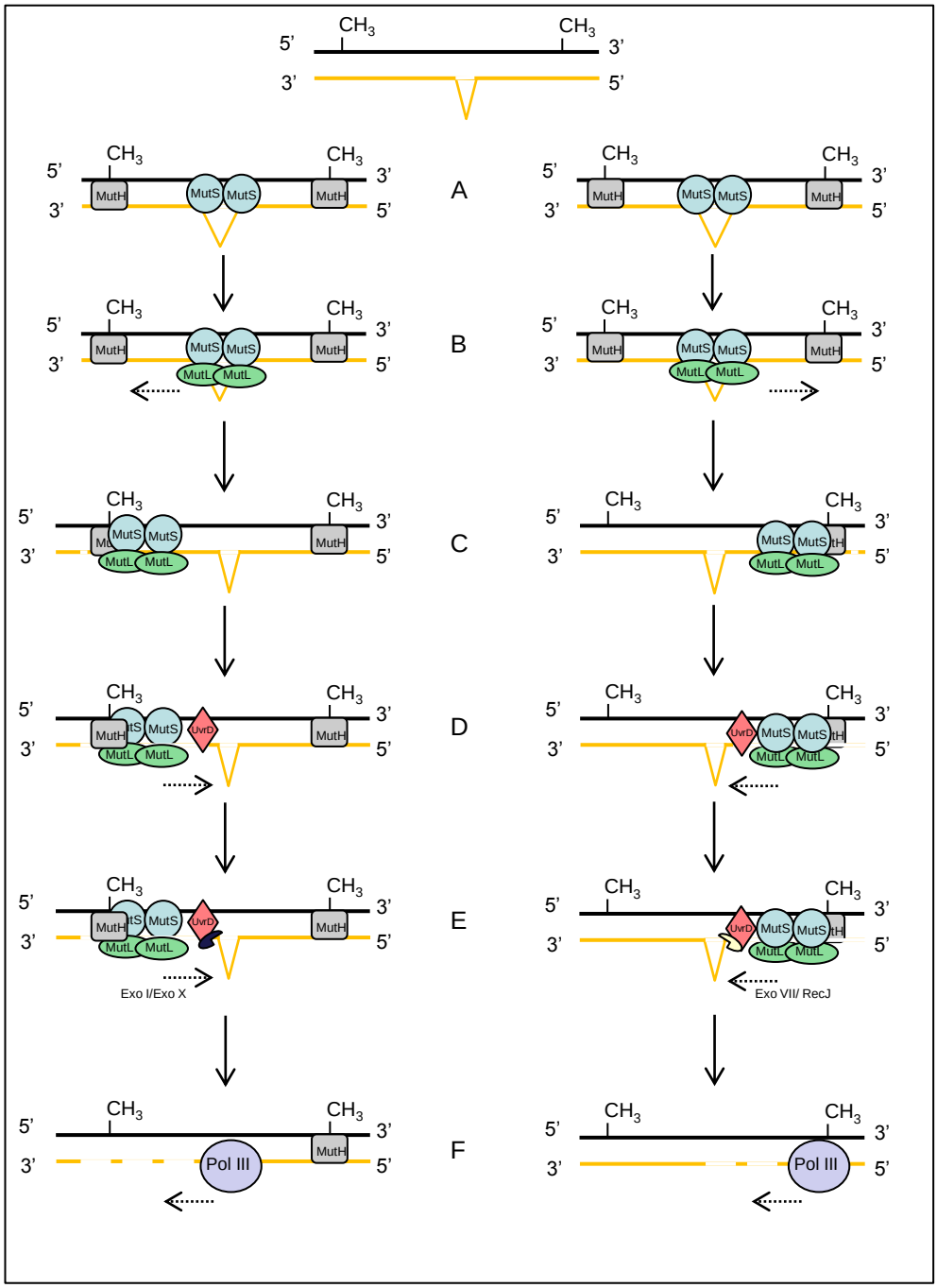
B) MutL homodiamer is then recruited by MutS which form a tetrameric complex that moves along the DNA in search of nearest MutH.

C) The MutS-MutL then activates the endonuclease activity of MutH which is required to create a nick in the unmethylated strand.

D) UvrD helicase then unwinds the nascent strand.

E) Excision of the nascent strand is carried out by Exonuclease I/Exonuclease X in the 3' to 5' direction or Exonuclease VII or RecJ in 5' to 3' direction.

F) The strand is then re sythesised and ligated by replication machinary.


CH3
CH3
5'
3'
3'
5'
A
B
C
D
E
F
MutH
MutS
MutL
UvrD
Exo I/Exo X
Exo VII/ RecJ
Pol III

1.2.2 MMR pathway in Eukaryotes

1.2.2.1 Recognition

The mechanism of mismatch repair (MMR) is highly conserved from *E. coli* to eukaryotes. In eukaryotes, the MMR pathway begins with the recognition of structural aberrations to DNA caused by mismatched nucleotide base pairs (Kunkel & Erie, 2005; G. M. Li, 2008). Two heterodimers are responsible for mismatch recognition. Firstly, MutSα is a heterodimer of hMsh2 and hMsh6 and is responsible for the recognition of base-base mispairs and small insertion-deletion loops (IDLs) of 1-2 nucleotides (Harfe & Jinks-Robertson, 2000; McCulloch et al., 2003). MutSβ is a heterodimer of hMsh2 and hMsh3 and is responsible for the recognition of larger IDLs. Mutsα and MutSβ are both ATPases and play critical roles in the initiation of the MMR pathway (Kunkel & Erie, 2005; G. M. Li, 2008) Binding of mismatch leads to an ADP to ATP exchange within the ABC-ATPase domains of Msh2 and Msh6, which causes a conformational change in that converts it to a sliding clamp capable of translocating along the DNA backbone *in vitro* (Gradia et al., 1999). In absence of the mismatch, these proteins are removed from the DNA in an ATP dependent manner (Acharya et al., 2003; Gradia et al., 1999; Mendillo et al., 2005). (Figure1.2).

After mismatch recognition by MutSα or MutSβ, one of the MutL heterodimer complexes (MutLα, MutLβ, or MutLγ) relocates to the site of DNA damage. Thus far, studies have identified four human MutL proteins: hMlh1, hMlh3, hPms1 and hPms2 (Kunkel & Erie, 2005). The hMutLα heterodimer consist of

hMLH1 and hPMS2, and is required for co-ordinating multiple steps in MMR (Kunkel & Erie, 2005). hMutLγ consists of hMLH1 and hMLH3, and is thought to be involved in repairing a subset of IDLs. hMutLγ also plays a critical role in the repair of programmed double strand breaks in meiosis (G. M. Li, 2008). As of yet, no role has been identified for the hMutLβ heterodimer of hMLH1-hPMS1 (Kunkel & Erie, 2005; G. M. Li, 2008)

1.2.2.2 Strand discrimination signal

In order to maintain replication fidelity, it is crucial that the removal of a mismatch occurs on the newly synthesized DNA strand. In *E. coli*, the brief lack of methylation at adenines in a d(GATC) sequence on newly replicated DNA, acts as a strand discrimination signal for MutH (Langle-Rouault et al., 1987). As eukaryotic cells do no express a protein analogous to MutH, what causes strand discrimination in eukaryotes is still unclear. *In vitro* studies using mammalian cell extracts have shown strand specific excision initiated by a nick located within a few hundred base pairs either 5' or 3' of the mismatch site (Pavlov et al., 2003). As the lagging-strand synthesis occurs discontinuously, mismatches arising during lagging strand replication should have a 5′ terminus no further away than 250 bases (Pavlov et al., 2003). The gap formed can then be used by Exo1 as an entry site. This model is further supported by the observation that the deletion of exo1 leads to a higher mutation rate when combined with a deficiency in lagging strand polymerase, than when combined with a deficiency in leading strand polymerase (Liberti et al., 2013). Additionally, studies involving repair of 8-oxo-guanine (8-OxoG) in DNA show a more efficient repair rate on

the lagging strand due to a higher density of strand discrimination signals (Pavlov et al., 2003).

Leading strand synthesis occurs in a continuous manner, but studies suggest that the 3' termini of a replication fork on the leading strand could be sufficient to provide the strand discrimination signal (Constantin et al., 2005). An *in vitro* study suggests that after recognition of the 3' nick, MutLα, which possess bidirectional intrinsic endonuclease activity, makes an incision 5' to the mismatch and Exo1 performs 5' to 3' excision from the MutLα incision site (Kadyrov et al., 2009). This requires PCNA and the clamp loader RFC (G. M. Li, 2008). This suggestion is supported by the observation that endonuclease activity of MutLα was found to be activated and directed to the nick-containing strand by PCNA loaded on the nick with specific orientation, consequently creating nicks at 5' site to the mismatches for the entry of EXO1 (Pluciennik et al., 2010).

RnaseH2 defective yeast cells have elevated mutation rate when combined with error-prone Polε and MMR deficiency (Lujan et al., 2013). Previously it is been reported that, DNA polymerases incorporate rNTP into the elongating DNA strand during replication (Nick McElhinny et al., 2010). These are repaired by ribonucleotide excision repair, which is initiated by RnaseH2 to introduce a nick at the ribonucleotide site (Sparks et al., 2012). This nick can be utilized for strand discrimination signal in leading strand as Polε introduces one rNTP into the elongating DNA strand for every 1250dNTPs (Nick McElhinny et al., 2010). This model was further confirmed by a study involving human cell nuclear

extract; it showed that the MMR utilize ribonucleotide in nick-free heteroduplex to perform correct MMR in vitro (Ghodgaonkar et al., 2013). However, the DNA sequence influence the incorporation of rNTPs significantly, thus it is possible that the distance between the nick and mismatch is quite large and cannot be utilized to support efficient MMR. Further as RnaseH2 null yeast strain showed weaker mutator phenotype as compared to complete loss of MMR, suggest that the contribution of nicks generated by RnaseH2 to the MMR in eukaryotic cells is minor (Lujan et al., 2013).

It is possible that, nicks/gaps generated on nascent strand by other DNA repair pathway during replication can also be utilized by MMR as the strand discrimination signal. Oxidation of the guanines in the genome is a frequent event and up to 1,000 to 100,000 8oxo-G residues can occur depending on the cell types (Burrows & Muller, 1998; Gedik et al., 2005). A study involving repair of 8oxo-G on the template strand showed that the nick generated during the repair can be used by MMR (Repmann et al., 2015).

1.2.2.3 EXO1-dependent excision

Exonuclease 1 (Exo1) carries out excision of mis-incorporated nucleotides. Exo1 belongs to the Rad2 family of exonucleases, it can initiate excision at a pre-existing nick and possesses 5′ → 3′ double stranded DNA exonuclease and flap endonuclease activities (Szankasi & Smith, 1992). The observation of increased recombination rates between intragenic markers (hyper-rec phenotype) when *EXO1* is mutated suggested a role of *EXO1* in MMR

(Szankasi & Smith, 1995). The heterodimer of MutLα has a latent endonuclease activity. PCNA, which is either available due to on-going replication or can also be recruited by Msh2–Msh6 and loaded by RFC, at the nicks in the vicinity of the mispair can activate the Mlh1–Pms1 endonuclease activity on the newly synthesized strand. MutSα/β- MutLα further diffuses along the DNA to introduce several nicks on nascent (daughter) DNA strand (Kadyrov et al., 2009). These nicks are then used as an entry site by Exo1, which is recruited and activated via interaction with Msh2 and Mlh1 (Nielsen et al., 2004; Tran et al., 2004). Exo1 excises the DNA in a 5′ → 3′ direction, termination of excision can occur due to a weak processivity of Exo1 caused by absence of the Mlh1–Pms1 or Msh2–Msh6 complexes (Langle-Rouault et al., 1987).

1.2.2.4 EXO1-independent excision

Unlike the multiple exonucleases in bacteria, Exo1 is the only known endonuclease that participates in MMR of eukaryotic cells. The involvement of Exo1 in the excision step of the human strand-specific mismatch repair reaction was confirmed by an in vitro assay for mismatch-provoked excision. However, the mutator phenotype of complete deletion of Exo1 in both *S cerevisiae* and mouse cells is weaker compared to that of complete deletion of Msh2 or Mlh1 (Hsieh & Yamane, 2008) also like other MMR genes including MLH1, MSH2, MSH6 and PMS2, deficiency of EXO1 in patients is not associated with LS (Jagmohan-Changur et al., 2003; Thompson et al., 2004). Therefore it is likely that at least one Exo1 independent pathway is involved in eukaryotic MMR pathway (G. M. Li, 2008). Finally, *in vitro* MMR assay using mice ES cell extracts and human purified proteins, showed strong residual MMR activity in

the absence of Exo1 confirming the idea that MMR can be still carried out in absence of Exo1 (Kadyrov et al., 2009; Wei et al., 2003). A genome wide screen was carried out in yeast, to screen for mutations that can induce an additive effect on mutation in combination with deletion of *EXO1*, but no potential candidates or alternative exonucleases were identified (Amin et al., 2001). However, the result is inconclusive considering the combinational depletion of *EXO1* and alternative exonuclease may lead to lethality or the participation of multiple exonucleases in Exo1-independent excision pathway leads to no obvious phenotype when only one backup exonuclease is depleted.

In human cells, inactivation of a 3'-5' exonuclease and single strand DNA endonuclease, *MRE11,* can induce microsatellite instability (MSI); Deletion of *MRE11* also down regulated MMR activity in a 3' nicked substrate. These two observations suggest a potential role of *MRE11* in MMR (Vo et al., 2005). Recently, a study involving *MRE11*, *Artemis* and *FAN1* in Exo1-independent MMR in MEF cells, revealed that inactivation of any one gene results in a weak or no change in MMR activity, whereas depletion of all four simultaneously, lead to complete loss of MMR like phenotype indicating that multiple exonucleases may participate in the Exo1-independent MMR pathway (Desai & Gerson, 2014).

The third subunit of Pol δ of yeast, Pol32 interacts with PCNA and participate in the error bypass synthesis (Huang et al., 2000; Hughes et al., 1999). A nonsense mutation in Pol32 can lead to additive mutation rate when combined with *EXO1* depletion. During elongation, Pol δ can to continue the extension

from the 5’ end of the DNA/RNA primer by displacing it; inactivation of Pol32 compromises in the strand displacement activity of Polymerase (Stith et al., 2008). Based on an observation using purified human MMR proteins, Modrich group has suggested one Exo1-independent MMR pathway in which the elimination of mismatches in DNA is executed by the strand-displacement activity of Pol δ (Kadyrov et al., 2009) but this model could not be confirmed using Exo1-deficient MEF nuclear extracts to perform the in vitro MMR assay.

Another proposal about Exo1-independent MMR pathway suggests role of the eukaryotic RecQ helicases in the removal of error containing DNA dependent on the nicks incised by MutL α (Song et al., 2010).Currently, five human RecQ homologues, including RECQL1, BLM, WRN, RECQL4 and RECQL5,RecQ helicases are shown to play role in DNA replication, recombination and repair through catalytically unwinding double-strand DNA (Bachrati & Hickson, 2008; Brosh & Bohr, 2007). The observations like, some cases of patients having a disease caused by deficiency of WRN, namely Werner Syndrome have also shown deficiency in MMR. However direct evidence of such interaction is still available. In summary, it is possible that in absence of Exo1, one or multiple of these factors in combination like alternative exonucleases, strand displacement synthesis and helicases, can compensate for lack of Exo1 dependent activity.

1.2.2.5 Resynthesis

The intact template strand is protected from nuclease degradation by RPA. Some *in vitro* studies show that RPA is essential for MMR and it can promote

excision termination in a MutLα-dependent manner (Y. Zhang et al., 2005). The DNA strand is resynthesized by DNA polymerase-δ and polymerase-ε in the presence of PCNA, RFC and RPA (Rasmussen et al., 2012) (Figure1.2).

1.2.3 Association of Mismatch Repair in with replication

Strand discrimination is essential for maintaining genomic integrity during MMR. In *E.coli*, strand discrimination is accomplished by the brief lack of methylation of adenine on the nascent strand; however, the methylation takes place soon after replication suggesting that MMR and replication are coordinated. The observations that MMR requires pre-existing nicks and several replication proteins, like PCNA for the efficient repair, indicated that MMR is coupled with replication (Iyer et al., 2006). Even though MMR proteins like Msh2-Msh3 and MS2-Msh6 directly interact with PCNA, the loss of binding between either Msh6 and PCNA or Msh3 and PCNA does not cause complete loss of MMR This indicates that interaction of PCNA with either of these proteins is not essential for MMR function (Clark et al., 2000; Flores-Rozas et al., 2000; Shell et al., 2007). Later, cell cycle specific studies using fluorescently tagged MMR and replication proteins showed a co-localization of the MMR recognition complex with replication machinery, supporting the idea that MMR is coupled with replication. Comparison of mutation rates of strains, in which *MSH6* expression is either restricted to S-phase or in G2/M phase, by fusing it with a cell-cycle phase-specific cyclin, revealed that loss of *MSH6* in S-phase leads higher mutation rate compared to G2/M phase, further highlighting association of MMR

with S-phase (Hombauer, Campbell, et al., 2011; Hombauer, Srivatsan, et al., 2011).

Figure 1.2: The Mismatch repair pathway in Eukaryotes.

A): The pathway begins with the recognition of mismatch or IDLs by MutS complex.

B): MutS complex then recruits the MutL complex which is involved in co-ordination of multiple steps in MMR. What signal directs repair of the newly synthesized strand is unclear but it is thought that strand discontinuities arising during replication or Okazaki fragments might direct the MMR proteins to repair newly synthesized strand.

C): Excision of the newly synthesized strand is done by Exo1 which needs pre-existing nicks which could be available during replication or due to activities of other DNA repair pathways or by processing by RnaseH2 etc.

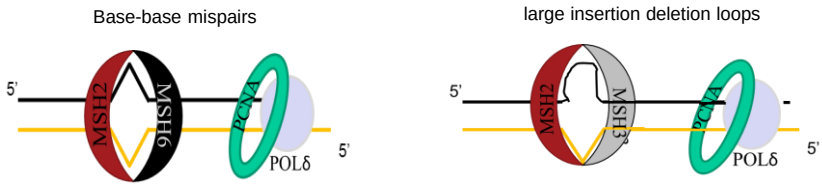
D) In absence of pre-existing nicks, it is thought that MutLα creates a nick 5' of the mismatch in the daughter strand because of its intrinsic endonuclease activity which can be used by Exo1 for excision.

E) Excision of the newly synthesized strand by Exo1 in 5' to 3' direction.

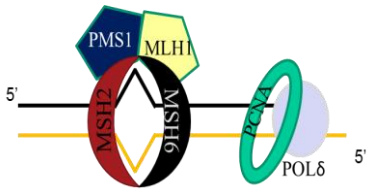
F) MutLα endonuclease introduces multiple breaks into the incised DNA strand and DNA polymerase δ carries out synthesis-driven displacement of mismatch and heteroduplex repair from these multiple incisions.

G and H): Resynthesis and ligation is carried out by DNA polymerase-δ or polymerase-ε in presence of PCNA, the clamp loader RFC and RPA and is ligated by ligase I.

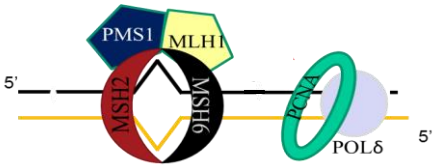
A) Recognition



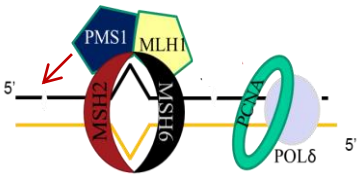
B) MutLα Recruitment

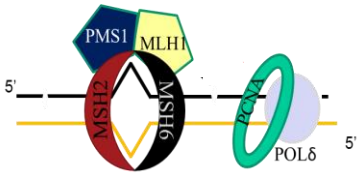


C) Pre-existing nicks

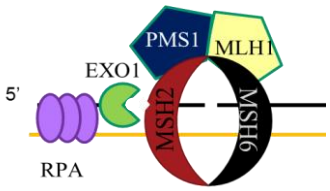


D) MutL dependent nicks

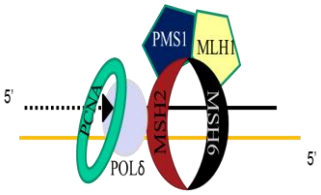




E) 5' to 3' error removal

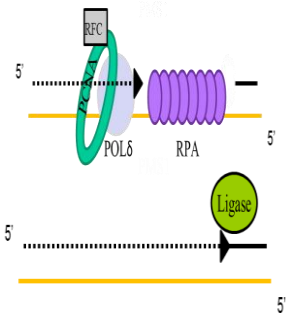


F) Exo1 independent error removal



G) Re-synthesis

H) Ligation



1.3 Other factors affecting efficiency of MMR

PCNA (Proliferating Cell Nuclear Antigen) is an integral part of replication and MMR pathway. PCNA is required for DNA synthesis after excision of the mispair, and plays a critical role in multiple upstream repair steps that leads to efficient functioning of MMR pathway (Chen et al., 2013; Flores-Rozas et al., 2000; Liberti et al., 2013; Pluciennik et al., 2010). PCNA belongs to the structurally and functionally conserved family of DNA sliding clamps. They form ring-shaped complexes (homodimers in eubacteria, homotrimers in eukaryotes and T4 bacteriophage, heterotrimers in archaea) with pseudohexameric symmetry, which encircle the DNA and are able to slide freely in both directions (Gulbis et al., 1996; Kelman & O'Donnell, 1995; Krishna et al., 1994; Moarefi et al., 2000).

Mutants of PCNA which can disrupt the DNA-PCNA interaction fail to initiate MMR (Giannini et al., 2002). In addition, mutations that disrupt interaction of PCNA with human MLH1 cause a severe mutator phenotype (Franchitto et al., 2003; Wen et al., 2008).Furthermore, interaction of PCNA with MSH6 or MSH3 during the initiation step of MMR, enhances the affinity of MutS complexes for DNA mismatches thereby improving MMR efficiency significantly (Vo et al., 2005). Studies using human cell extracts have shown that PCNA can interact with Exo1, this helps in recruitment of Exo1 also improves efficiency of MMR (Desai & Gerson, 2014; Huang et al., 2000). PCNA is also responsible for activating the endonuclease activity of MutLα in mismatch- containing heteroduplex DNA in Exo1 independent repair (Zou & Elledge, 2003).

Moreover, PCNA can enhance the DNA strand displacement activity of Polδ, which may play a role in Exo1- independent repair.

Table 1.2: PCNA interactions with Mismatch repair proteins			
	MMR complex	Interaction	Reference
Recognition step	Msh2-Msh6	PCNA interacts with Msh6 by a PIP-box motif at the N terminus of Msh6 ; this interaction is important for increased mispair binding efficiency and efficient strand discrimination	(Flores-Rozas et al., 2000)
Strand discrimination	Mlh1-Pms1	PCNA interacts with MutLα by the interdomain connector loop in Mlh1, this interaction is important for activation of the latent endonuclease and is critical in Exo1 independent MMR	(Goellner et al., 2014; Lee & Alani, 2006)
Exo1 dependent excision	EXO1	PCNA interacts with the PIP-box motif of Exo1 and this interaction is required for recruitment of exo1 to replicating DNA	(Liberti et al., 2011)

1.3.1 Accessory factors affecting MMR efficiency:

In order to repair the mismatch efficiently, it is crucial that MMR machinery get the required access to the mismatch. Chromatin structure can either facilitate or hamper this access thereby greatly influencing the MMR efficiency and genomic stability (Loyola & Almouzni, 2004). Using a model of a DNA substrate containing the Xenopus 5S rDNA nucleosome Javaid et al showed that replication-associated acetylation modification H3K56A cenhances nucleosome disassembly by hMSH2-hMSH6 to facilitate the downstream MMR events (Javaid et al., 2009). Besides, it was also shown that upon recognition of a mismatch, MutSα can prevent CAF-I-dependent histone H3–H4 remodelling in order protect the newly synthesized DNA strand from excessive degradation by MMR machinery (Schopf et al., 2012).

Histone methylation also has an active role in regulation of mismatch repair efficiency. For example a heterotrimeric remodelling complex called RFX enhances in vitro MMR activity. Using biochemical studies it has been demonstrated that trimethylated histone H3 lysine 36 (H3K36me3) interacts with the PWWP domain of MutS in a cell cycle dependent mannaer and this interaction is important in recruitment of MutS (F. Li et al., 2013). Cells which are deficient in the H3K36 tri methyltransferase SETD2 display higher mutation rate and microsatellite instability (MSI) which is a hallmark of MMR-deficient cells (Li, 2013; Dhayalan, 2010). This is further supported by the observation that, over expression of H3K36me2/3 demethylases, KDM4A-C that leads to depletion of H3K36me3 results in disruption of Msh6 localization with chromatin and is also associated with increase in mutation rate and MSI (Awwad & Ayoub,

2015). However as the PWWP sequence is not conserved and is not present in the yeast proteins it still needs to be investigated if it is a higher eukaryotic feature. Another study have suggested that in hypoxic stressed cells, *MLH1* transcription is down regulated via histone deacetylation, implying that histone acetylation status also affects MMR activity (Mihaylova et al., 2003).

1.4 Mismatch repair proteins in other DNA repair pathways

1.4.1 Mismatch repair proteins and double strand break repair

Double-strand breaks (DSBs) are a common type of DNA lesion which arise due to an exposure to DNA damaging agents such as radiation and certain chemicals, as well as through cellular processes, such as DNA replication and repair. Additionally, in meiotic cells intently create DSBs which triggers recombination based repair to ensure normal chromosome segregation (de Massy, 2013). Various other cellular process like mating-type switching in yeast (Haber, 2012), or T-cell receptor formation in T-lymphocytes, and immunoglobulin class switching in B-lymphocytes (Soulas-Sprauel et al., 2007). also involve programmed induction of DSBs. In eukaryotes DSBs are usually repaired by, one of the three pathways, homologous recombination (HR), non-homologous end-joining (NHEJ) and alternative end-joining which is mainly bases on single strand annealing (SSA).

The HR and SSA pathway both rely on searching for homology followed by recombination based repair. During these Msh2 and Msh3 by are responsible for removal of the non-complementary 3' tails; in order for the homology regions to anneal together. The nuclease Rad1/Rad10 is also required during this process to remove the flap (Fishman-Lobell & Haber, 1992; Paques & Haber, 1997; Sugawara et al., 1997). MMR proteins are also required for reducing the unusual recombination events between non-homologous DNA sequences. A study involving *E. coli* has shown that mismatches in homeologous dsDNA are substrate for MutS and MutL, which inhibit the strand invasion by binding the

mismatch and further signals UvrD helicase, which stops the recombination event (Tham et al., 2013).

1.4.2 Mismatch repair and base excision repair

Cellular processes, such as oxidative stress, can modify nucleotides present in both from the DNA or the nucleotide pool (e.g., oxidation or alkylation etc.). These damaged bases are mainly removed by Base Excision Repair (BER). BER is initiated by a DNA glycosylase, which removes the damage base, by an incision to the DNA backbone which leads to formation of an apurinic/apyrimidinic site (AP site). This AP site is then removed by the endonuclease APE1, and then DNA is re-synthesized by Pol β (Svilar et al., 2011).

In murine embryonic stem cells or fibroblasts cells deficient in Msh2, the efficiency of BER in repairing one of its common substrate 8-oxodG is affected suggesting a possible role of Msh2 in BER (Colussi et al., 2002; DeWeese et al., 1998). Another study in yeast has also shown that Mutsα can lower the G:C to A:T transversions resulting from A/8-oxodG mispairs indicating role of Mutsα in repair of 8-oxodG (Ni et al., 1999).

1.4.3 MMR in interstrand crosslink (ICL) repair

Exposure to external sources like platinum based chemotherapeutic compounds or endogenous crosslinking agents like aldehydes produced during metabolism, can result in nucleotide bases on opposing DNA strands of a DNA helix form a covalent bond called as interstrand crosslink (ICL). Repair of such ICLs is a challenging process and require a combined action of several repair pathways (Noll et al., 2006).

ICL repair is initiated by recognition of the complex containing NER and/or Fanconi anemia (FA) factors. In order to resolve the ICL, one of the strands is incised on both 5' and 3' and the gap is re-synthesized and ligated. Further removal of the ICL from the other strand is carried out by NER proteins (Noll et al., 2006; Shukla et al., 2013). Early findings in eukaryotic cells have suggested that Mutsβ (Msh2-Msh3) can recognise ICL *in vitro* and is required in the incision step in repairing the ICL (Wu et al., 2005; N. Zhang et al., 2002) [83]. This is supported by further study describing that MSH2 deficient human cells exhibit hypersensitivity to psoralen ICLs (Wu et al., 2005).

1.5 MMR and Lynch syndrome:

Heterozygous germ line mutations in MMR genes are associated with Lynch syndrome, also known as Hereditary Non Polyposis Colorectal Cancer (HNPCC). Lynch Syndrome is a dominantly inherited autosomal disorder which accounts for 4-6% of the total colorectal cancer burden in the population (Lynch et al., 1991).The syndrome is heterogeneous and can be divided into two clinical subsets: Lynch I and Lynch II. Lynch I is characterised by early age onset of colonic cancer, whereas Lynch II colonic and extra colonic cancer sites are observed (Lynch et al., 1991). Lynch I patients have a higher risk of colorectal cancer (CRC) (approximately 80%) when compared to the general population (approximately 5.5%) (Cancer.Net. 2013).

According to the InSiGHT database around 38% of Lynch syndrome patients have mutations in *MLH1*, 32% in *MSH2, PMS2* accounts for 19% and remaining are associated with *MSH6*, *MLH3* or *MSH3.* As the predisposing germline mutation inactivates one of the MMR genes, inactivation of the remaining wild-type allele either by somatic mutations or promoter methylation can lead to inefficient MMR that allows accumulation of mutations during DNA replication thereby accelerating the onset of cancer. Recently, inactivation of *MSH2* expression by deletion of an upstream gene *EPCAM* has gained attention as this leads LS like phenotype. The deletion of the two most 3' exons of *EPCAM* caused transcription of *EPCAM* to extend into *MSH2*, which lead to silencing of *MSH2* promoter, since the silencing of *MSH2* expression is limited to *EPCAM* expressing tissue, the tumour spectrum for these patients is different. (Peltomaki, 2003); Bignami et al., 2003)

Disruption of MMR increases the rate of spontaneous mutagenesis and cause microsatellite instability (MSI). Microsatellites are the repeated DNA sequences this makes them difficult to replicate as during replication the polymerase can slip during this region and lead to insertion deletion loops (Figure 1.3). Tumours that are deficient in *MSH2* or *MLH1* display high MSI, whereas patients with mutations in *MSH6* or *PMS2* usually show lower penetrance and later onset of cancer. One possible explaination of high MSI for msh2 deficient tumours is that MSH2 is part of both the recognition complexes MutS α/β.

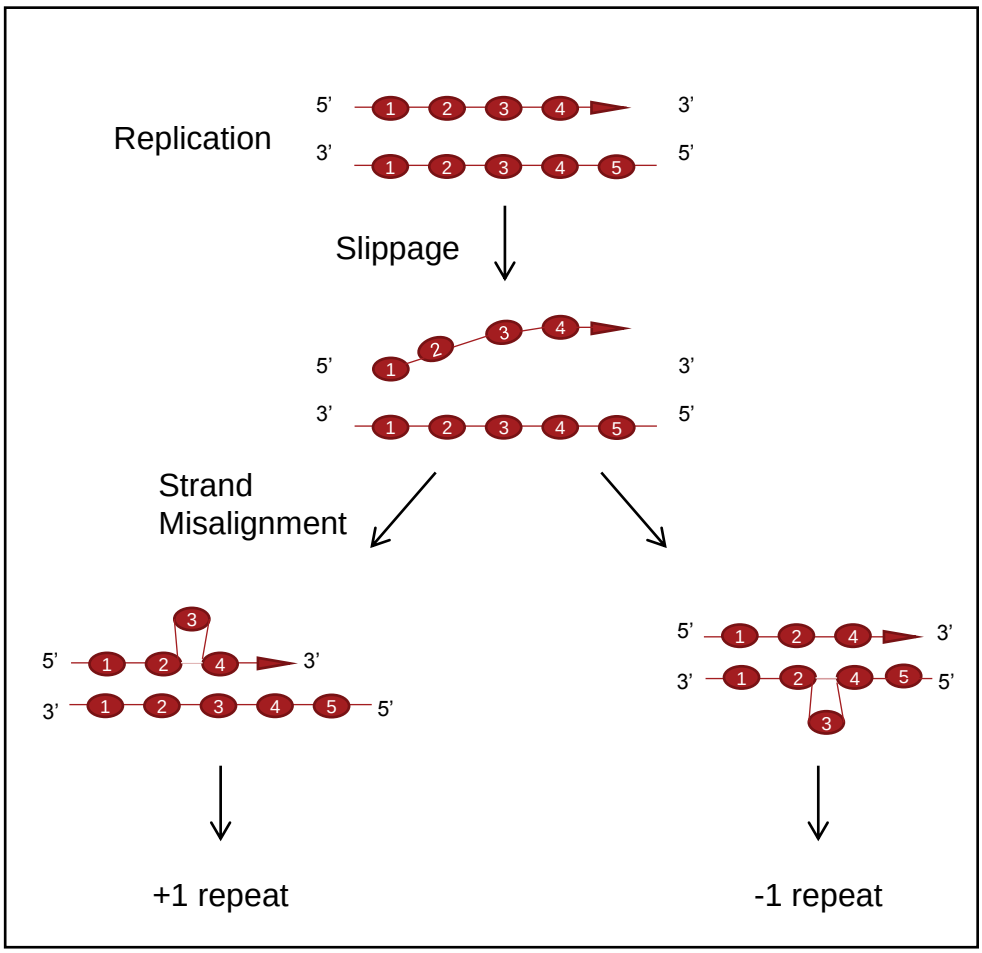


Figure1.3: MSI caused due to polymerase slippage during DNA replication.
During DNA replication of repetitive DNA sequences; dissociation of primer from template strand is common. Strand mis-allignment during re-annealing may result in either in a displaced loop of DNA on the daughter strand (left side of figure) or in a displaced loop on the partenal strand (right side of figure). If the errors formed are not repaired, the microsatellite on the newly synthesized daughter strand will have an additional (left side) or missing (right side) repeat.

There is not a clear genotype –phenotype correlation in LS patients. Thus the reason why cancer spectrum caused by LS in associated with these tissues is still unidentified. One possible explanation is that the combination of a MMR-deficiency is causes mutation in one of the tumour suppressors genes that have high expression level in highly proliferative cells, could lead to onset of the cancer. For example, *TGFBR2* is inactivated in around 80% of MMR deficient tumours. In humans *TGFBR2* encodes a protein transforming growth factor beta type 2 receptor, which is highly expressed in gastrointestinal epithelial cells (Chao & Lipkin, 2006). The gene includes a repeat sequence of ten adenines in the coding region, which could be a target for MSI in MMR deficient cells (Alhopuro et al., 2012; Markowitz et al., 1995). GI cells are usually considered fast proliferating cells with a turnover of 3-5 days, which increases the probability of accumulation of mutations in these cells. For cancer to develop in LS a second hit is needed to abolish MMR, this is likely caused by a constant exposure of a mutagenic agent on these cells such as mutagens from bile and food (Medina-Arana et al., 2012). In addition, MMR deficient cells are significantly tolerant to many DNA damaging agents, which can help them to outgrow the neighbouring cells in stressed environment (Toft et al., 1999).

Individuals with LS have significantly higher likelihood of developing cancers. A simple procedure of adenoma removal can significantly decrease the likelihood of developing CRC and therefore can significantly improve survival. In 1991, an elaborated criterion (Amsterdam I) was constructed for identification of families who are likely to have LS. Over the years with increasing knowledge of the disorder, criteria have been revised and guidelines have been introduced. These guidelines are known as Amsterdam criteria II (Vasen et al., 1999), the

Bethesda guidelines (Boland et al., 1998) and the revised Bethesda guidelines (Umar et al., 2004). As opposed to the Amsterdam criteria which only depended upon the familial history, Bethesda guidelines also take MSI into consideration (Table1.3).

Table1.3: Revised Bethesda guidelines

1) Tumours from individuals should be tested for MSI if:
   CRC is diagnosed in a patient before 50 years of age.

2) Irrespective of the age of a patient, if presence of synchronous, metachronous colorectal or other HNPCC-associated tumours is observed

3) CRC with the MSI-H histology is diagnosed in a patient before 60 years of age.

4) CRC is diagnosed in one or more first-degree relatives with an HNPCC-related tumour, with one of the cancers being diagnosed under age 50 years.

5) CRC is diagnosed in two or more first- or second-degree relatives with HNPCC-related tumors, regardless of age.

For diagnosis of LS, tumour tissue is analysed for expression level of MMR proteins by immunohistochemistry (IHC) followed by sequencing of the respective gene to find the possible pathogenic mutation. MSI screening is carried out if HIS is inconclusive. Nearly 80%of CRC LS tumours have MSI (de la Chapelle & Hampel, 2010). In order to promote consistency across MSI testing, the Bethesda guidelines suggested a panel of markers to test for MSI. Based on the results of MSI testing, tumours can be classified as “MSI-H” (when 40% of the markers were positive for MSI), “MSI-L” (when less than 40%

of the markers were positive for MSI) or “MSS” (when none of the markers showed instability).

In case IHC is inconclusive, MSI screening is thought to be a good secondary selection criterion. When MSI was found to be one of the hallmarks of LS the markers used for the identification of MSI were often chosen based on the protocol used at the local clinic. In the first Bethesda Guidelines from 1998 five markers were suggested as a standard panel to promote consistency across MSI tests. Moreover the classification of MSI was also suggested to be sub grouped into MSI high when 40% of the markers showed MSI and MSI low when less than 40% were positive and MSS when none of the markers showed instability. If a tumour was termed MSI-L or MSS it was recommended to apply additional markers (Boland et al., 1998). However approximately 80% of adenomas associated with LS CRC show MSI-H phenotype, whereas only 20%-30% of endometrial cancers exhibit MSI-H phenotype (Hampel et al., 2006). Surprisingly a small portion of Lynch syndrome-related tumours with MSI do not show any evidence of MMR deficiency (Shia, 2008).

LS tumours are not the only ones that have MSI. Around 10-15% of sporadic cancers also show MSI (de la Chapelle & Hampel, 2010; Hsieh & Yamane, 2008). MSI in sporadic CRC are often caused by somatic events affecting both alleles of a MMR gene. The single base-pair mismatches caused by the faulty DNA replication result in point mutations, whereas IDLs result in frame-shift mutations that can lead to a downstream nonsense mutation; this results in

production of a truncated, non-functional protein. The sequential accumulation of such mutations can lead in development of cancer (Goel & Boland, 2010).

1.6 Nucleotide excision repair (NER) pathway:

Nucleotide excision repair (NER) is a highly conserved and versatile DNA repair pathway and is composed of two sub-pathways: global genome NER (GGR) and transcription coupled nucleotide excision repair (TCR). NER is responsible for repairing a variety of different DNA lesions; however, the primary targets are severely distorting DNA lesions, such as intra-strand crosslinks caused by UV-induced pyrimidine dimers (Batty et al., 2000; Sancar, 1996). NER involves the stepwise action of around thirty proteins that function in several sub-complexes at the site of damage (Araujo et al., 2001; Guzder et al., 1996; Volker et al., 2001). In *S. cerevisiae,* the presences of active high molecular weight NER protein complexes, even in undamaged cells, have suggested that NER might take place through a pre-assembled repair complex (Rodriguez et al., 1998; Svejstrup et al., 1995). Briefly, the pathway consists of two single-stranded incisions on both sides of a DNA lesion and removal of the damaged oligonucleotide, thus creating a 24–32 nucleotide gap that is filled by DNA polymerase and ligated (Figure 2A).

The first sub-pathway, global genome NER (GGR), is responsible for repairing DNA lesions throughout the genome. The efficiency of GGR is not the same across the genome due to the potential effects of the chromatin environment on GGR machinery (Feng et al., 2003). In *S. cerevisiae,* recognition of DNA lesions is carried out by Rad14 (human XPA), Rad4-Rad23 (human XPC/HHR23B), RPA and a complex of Rad7-Rad16. Next the DNA duplex around the lesion is unwound by the helicase domains of the basal transcription/repair factor TFIIH, Rad3 (human XPD) and Rad25 (human XPB)

(Schiestl & Prakash, 1988; Schiestl & Prakash, 1990). To enable the removal of the damage containing oligonucleotide, the nucleases Rad1-Rad10 (human XPF/ERCC1) and Rad2 (human XPG) incise the strand on the 5′ and the 3′ side of the lesion, respectively (Prakash & Prakash, 2000). The nuclease complex of Rad1-Rad10 lacks the specificity required for binding to UV lesions, thus the ability of the complex to associate with Rad14 is crucial for the targeting the nuclease to DNA lesion (Guzder et al., 2006). Finally, the strand is resynthesized by DNA polymerase and ligated together to complete the repair process.

Table 1.4: Proteins associated with Nucleotide excision repair and transcription coupled repair			
NER	Rad1	XPF	Forms a complex with Rad10 and makes the 5' incision to the lesion
	Rad10	ERCC	Forms a complex with Rad1 and makes 5' incision
	Rad2	XPG	Makes 3' incision to the lesion
	Rad3	XPD	
	Rad4	XPC	Forms a complex with Rad23 and is involved in recogninition of lesion
	Rad23	HR23B	Forms a complex with Rad4 and is involved in recogninition of lesion
	Rad7	DDB1	Forms a complex with Rad16 involved in recogninition of lesion
	Rad16	DDB2	Forms a complex with Rad7 involved in recogninition of lesion
	Rad14	XPA	involved in recognition of lesion also helps in stabilization
	DNA pol δ	DNA pol δ	strand re-synthesis
	DNA pol ε	DNA pol ε	strand re-synthesis
	PCNA	PCNA	required during strand re-synthesis
	Cdc9	Ligase I	ligation
TCR	Rad26	Csb	recrutment of TCR regulatory and NER proteins
	Rpb9	Rbb9	required for Rad26 independent repair
	Rad28	CSA	possible role in complex stability

Figure 1.4: The Nucleotide excision repair and Transcription coupled repair.
A) The GGR pathway repairs severely distorting DNA lesions in the DNA

B) In GGR recognition of the lesion is carried out by a complex of Rad4-Rad23, RPA and a complex of Rad7-Rad16. Due to the chromatin-remodelling activity, these complex can open the helix ~10 bp which allow efficient recognition.

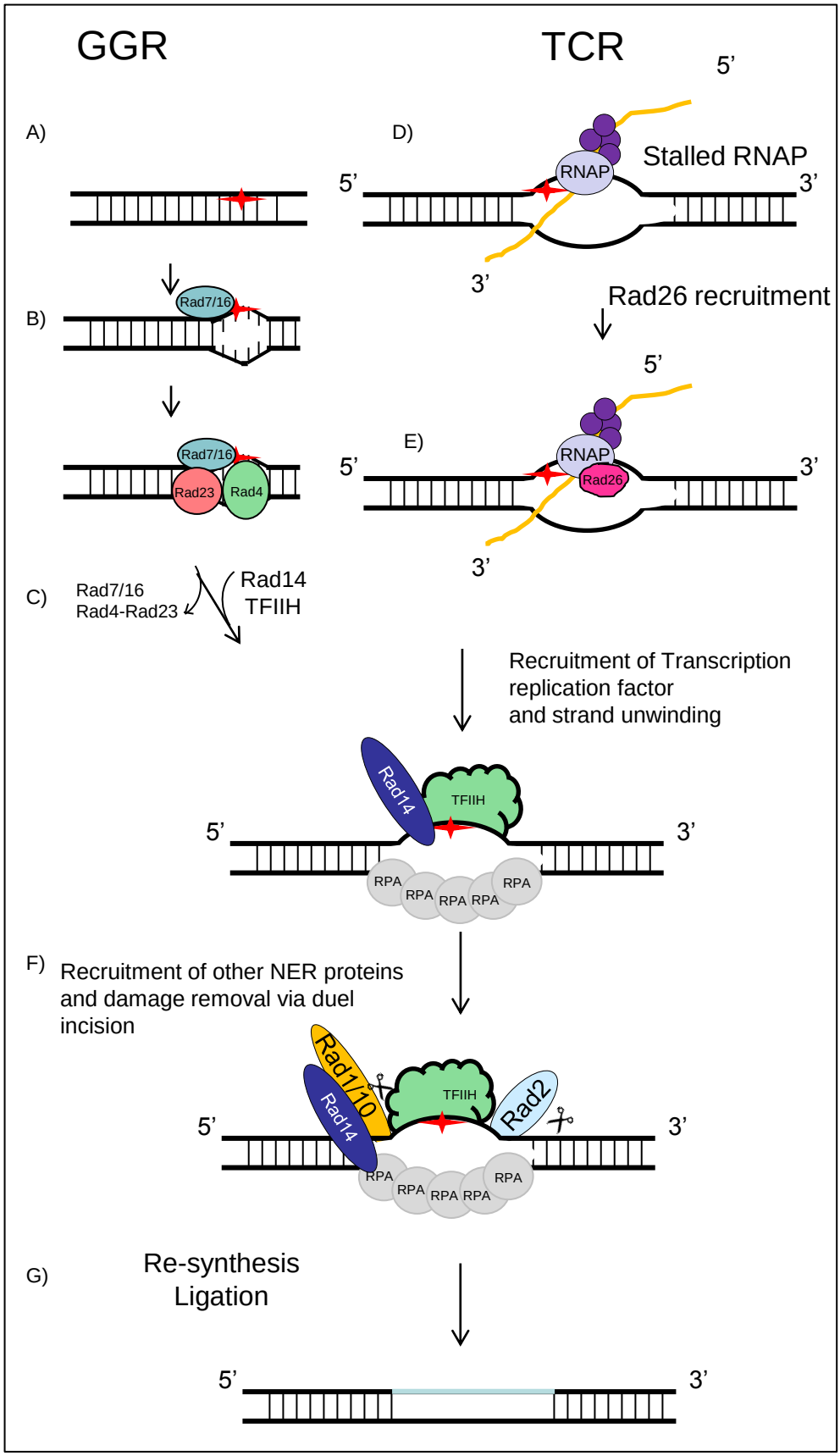
C) Rad14, RPA and TFIIH are recruited which form a pre-incision complex that is important for verification of the lesion and it further unwinds the DNA. During this step Rad4-Rad23 and Rad7-Rad16 are released from the DNA

D) The TCR pathway repairs DNA lesions arising in transcribed strand of transcriptionally active genes.

E) In TCR stalled polymerase at the lesion site during transcription is responsible for recruiting Rad26 which then coordinates the recruitment of TFIIH and other NER complexes.

F) For the damage removal, the structure-specific nucleases Rad1-Rad10 and Rad2 incise the strand on the 5′ and the 3′ side of the lesion respectively.

G) The strand is then re-synthesized and ligated with replication machinery.


GGR
TCR
A)
B)
Rad7/16
Rad7/16
Rad23
Rad4
C)
Rad7/16
Rad4-Rad23
Rad14
TFIIH
D)
5'
RNAP
Stalled RNAP
5'
3'
3'
Rad26 recruitment
5'
E)
RNAP
Rad26
5'
3'
3'
Recruitment of Transcription
replication factor
and strand unwinding
Rad14
TFIIH
5'
3'
RPA
RPA
RPA
RPA
RPA
F)
Recruitment of other NER proteins
and damage removal via duel
incision
Rad1/10
Rad14
TFIIH
Rad2
5'
3'
RPA
RPA
RPA
RPA
RPA
G)
Re-synthesis
Ligation
5'
3'

1.7 Transcription coupled nucleotide excision repair (TCR)

The other sub-pathway of NER is transcription coupled nucleotide excision repair (TCR), which is responsible for preferential repair of DNA alteration in the transcribed strand of active genes (Mullenders et al., 1991). The main difference between TCR and GGR is in the recognition step and how a stalled RNAP II recruits NER machinery to the lesion site. After the recruitment of the gap-opening helicase factor TFIIH and other factors, the damage is repaired in a similar manner with the outcome being the restoration of the normal nucleotide sequence. TCR is conserved from prokaryotes to eukaryotes and mutations in TCR can lead to a severe hereditary disorder called Cockayne's syndrome (Hanawalt, 2001). Even though, the exact mechanism is not yet known, it is thought that a similar preferential repair of the transcribed strand might also occur in ribosomal DNA genes that are transcribed by RNA polymerase I (Christians & Hanawalt, 1994; Verhage et al., 1996)

In *S. cerevisiae,* arrested RNAPII at the site of DNA damage initiates the TCR pathway. During normal transcription, TFIIS can rescue the stalled RNAPII via backtracking; however, it is not always possible in the presence of a DNA lesion and RNAPII must be degraded to prevent the blockage of later transcription (Hanawalt & Spivak, 2008). Two proteins in *S. cerevisiae* are responsible for removal of RNAPII–Def1 and Rad26–by forming a complex with it, which leads to RNAPII ubiquitylation and degradation (Woudstra et al., 2002). Def1, which is also important during elongation, can interact with structural components if the RNAPII complex, such as the subunit Rpb9 and the C-terminal repeat domain (CTD) of Rpb1, the largest subunit of RNAPII (Gaillard et al., 2009). Rad26, with

the help transcription factor TFIIH, coordinates the next step of lesion repair by recruiting NER proteins along with other TCR regulatory proteins like the THO complex, Rad28, Paf1 and Thp1-Sac3 (Gaillard et al., 2009).

Additionally, Rad26, which is a part of the SNF2 sub-family of DNA helicases, is possibly responsible for unwinding the DNA to ensure accessibility for other TCR repair factors (Laine & Egly, 2006; van Gool et al., 1994). However, Rad26 is not essential for TC-NER and an alternate pathway is available that relies on the nonessential subunit of RNAPII, Rpb9 (S. Li et al., 2006). Even though the repair pathways are overlapping in their mechanisms, they differ in the substrate specificity. For example, Rad26 is mainly required for the repair of sequences transcribed at low levels, whereas Rpb9 is required in TCR of highly transcribed genes (S. Li & Smerdon, 2002). Surprisingly, the double mutant *rad26*Δ *rpb9*Δ is moderately UV-sensitive, suggesting that there could be yet another alternate mechanism associated with TCR in yeast.

In *S. cerevisiae,* the functional homolog of human CSA is Rad28. Unlike in mammalian cells, yeast Rad28 mutants are proficient in TCR and Rad28 is not required for cell survival or recovery of RNA synthesis after UV radiation (Laine & Egly, 2006). However, it is possible that Rad28 is required for maintaining the integrity of repair proteins or protecting repair proteins from ubiquitination (Bhatia et al., 1996; Reagan & Friedberg, 1997). Some other factors actively involved in mRNA metabolism and transport are also associated with TCR in yeast. Mutants in any of the THO, Thp1-Sac3, or Spt4/5 complexes, cause TCR deficiency due to their roles in mRNP biogenesis and transport (Gaillard et al.,

2007). Further studies are required to understand the links between transcription elongation, mRNA processing and TCR.

1.8 Transcription associated Mutagenesis

Unlike replication that is active only during S-phase, transcription is active throughout the cell cycle, therefore, making it a significant threat to genomic stability. Importantly, transcription and replication can interfere with each other by competing for the same DNA template, resulting in mutations and DNA damage, known as transcription-associated mutagenesis (Jinks-Robertson & Bhagwat, 2014) (Bradford,).

From prokaryotes to eukaryotes the association of transcription and mutagenesis has been established. It is thought that the movement of the transcription machinery along DNA cause changes in DNA topology or can encounter the replication machinery, which leads to genomic instability known as transcription associated mutagenesis or transcription associated recombination (Choudhury et al., 2007).

During transcription, only one DNA strand gets copied at a time, leaving the non-transcribed strand (NTS) in a transiently single-stranded state thus making it more susceptible to the various exogenous and endogenous DNA damage (Gnatt et al., 2001). The other possible deleterious effect of transcription is formation of R-loops; which are essentially a three-stranded RNA/DNA hybrid structure, which are formed as a consequence of annealing of transcript with the template strand. Within an R-loop, the NTS is exposed to DNA damaging agents but is also prone to formation of secondary structures that would be harmful for various other cellular processes. In prokaryotes, the transcription is coupled with translation that reduces the formation of R-loops. Even in

eukaryotes where these processes are not, linked processing of m-RNA reduces R-loop formation (Gowrishankar & Harinarayanan, 2004; X. Li & Manley, 2006). Importantly, during transcription the accessibility of the transcriptional machinery is facilitated by positive and negative supercoiling, which can introduce DNA damage (Liu & Wang, 1987).

Yeast strains deficient in Hpr1, component of the THO/TREX complex highlight the role of transcription in inducing recombination phenotype. These strains have 1000- fold increase in recombination between direct repeats caused due to direct movement of RNAPII complexes through the substrate (Aguilera & Klein, 1988; Prado et al., 1997). Studies in *E. coli* and yeast have demonstrated that during transcription multiple types of base substitutions are increases (Klapacz & Bhagwat, 2002; Lippert et al., 2011). Further it was also showed that in E.coli, transcription increases up to 4 fold increase in frequency of cytosines deamination to uracils (Beletskii & Bhagwat, 1996)

2 Chapter 2 - Materials and Methods

2.1 Materials

2.1.1 Bacterial Media

Bacterial cells were grown at 37˚C in Lysogeny Broth (LB) (1% w/v bacto-tryptone, 0.5% w/v yeast extract, 0.5% w/v NaCl, pH 7.0) or Luria agar (1% w/v bacto-tryptone, 0.5% w/v yeast extract, 0.5% w/v NaCl, pH 7, 2% w/v agar was added before autoclaving. For the selection of plasmids appropriate antibiotics were added according to the plasmid

2.1.2 Yeast Media

Yeast cells were grown in one of the 3 mentioned media (table2.1).For the selection of auxotrophic markers, appropriate amino acid dropout solution was used according to the concentrations mentioned in table2.2;for antibiotic selection appropriate antibiotics were added according to the concentrations mentioned in table2.3. for solid media 1% agar was added before autoclaving.

Table2.1 Yeast media composition.		
Media	Composition	Used for
YEPEG	Yeast extract 1% Bacto-peptone 2% Glycerol 2% Succinate 1% Adenine 0.5 mM pH 5.5 post autoclave ethanol 2%	Selection of cells with functional mitochondria
YPD	Yeast extract 1% Bacto-peptone 2% D-glucose 2% pH 6.5 Ready mix from of YPD Formedium (CCM0110) was resuspended in water and supplemented with Adenine 0.5 mM	growth
Drop-out	YNB without amino acids 0.17% D-glucose 2.0% Amino acid supplement pH 7.25	Auxotrophic selection
Canavanine	YNB without amino acids 0.17% D-glucose 2.0% Amino acid supplement pH 7.25 post autoclave add Canavanine (60µg/mL)	Mutation rate assay

Media components were mixed according to the final concentration mentioned in the table

Table2.2: The final concentrations of amino acids in dropout media	
Amino acid	final concentration mg/L
Adenine	10
L-Arginine	50
L-Aspartic acid	80
L-Histidine HCl	20
L-Isoleucine	50
L-Leucine	100
L-Lysine HCl	50
L-Methionine	20
L-Phenylalanine	50
L-Threonine	100
L-Tryptophan	50
L-Tyrosine	50
Uracil	20
Valine	140

The drop out media was made by excluding the required amino acid from the mixture, commercially available dropout powders were used to get the final concentrations mentioned in the table

Table2.3: Formedium product number of dropout solutions	
Dropout powder	Formedium product numbers
Arginine droput	DCS0059
Leucine dropout	DCS0099
Lysine dropout	DCS0109
Tryptophan dropout	DCS0149
Uracil dropout	DCS0169
complete medium	DCS0019

2.2 Bacterial methods

2.2.1 Plasmid extraction:

For the plasmid extraction, cells were grown overnight at 37˚C by shaking at 200rpm in LB. Plasmid was extracted using QIAprep Spin Miniprep Kit (Cat No: 27104) with the standard protocol provided by Qiagen.

2.2.2 Storage of bacterial cells:

Bacterial cells were grown overnight at 37˚C by shaking at 200rpm in 5 ml LB. Next day cells were centrifuged and re-suspended in a screw-top tube containing 1 ml 30% glycerol and stored at -80˚C.

2.3 Yeast methods

2.3.1 Vegetative growth conditions

Yeast strains were stored at -80˚C as glycerinates. To ensure presence of functional mitochondria cells were woken up on YEPEG media (1% w/v succinate, 1% w/v yeast extract, 2% w/v bacto-Peptone, 2% v/v glycerol, 500 µM adenine, pH 5.5. 2% w/v agar was added before autoclaving and 2% v/v ethanol was added after autoclaving). Cells were then streaked for single colonies on YPD plate (commercially available YPD medium was used, Formedium CCM0105 supplemented with 500 µM adenine, pH 6.5) (1% w/v yeast extract, 2% w/v bacto-peptone, 2% w/v glucose, 500 µM adenine, pH 6.5 then 2% w/v agar was added before autoclaving). For the liquid cultures single colonies were grown by shaking at 200rpm at 30˚C in YPD liquid (commercially available YPD medium was used, Formedium CCM0405 supplemented with 500 µM adenine, pH 6.5)

2.3.2 Sporulation conditions

To enhance sporulation efficiency, S288C diploids were grown overnight on a pre-sporulation media (1% w/v yeast extract, 2% w/v bacto-peptone, 6% w/v glucose, 500 µM adenine, pH 6.5 then 2% w/v agar was added before autoclaving). Next day cells were transferred to 1% KAC (without raffinose and amino acids) plates or 5 ml liquid. As the S288C strains require more time to sporulate compared to other strains for example SK1, plates were kept at 30˚C incubator for 3 days and liquids cultures were kept at 30˚C shaking at 200rpm for 2 days.

2.3.3 LiAc Yeast transformation

The lithium acetate method was used for yeast transformation (Gietz et al., 1995; Gietz and Schiestl, 2007). Yeast strains were grown shaking at 30˚C in 5 ml YPD liquid. When they reached stationary phase (next day), cells were diluted 10 times in fresh 5 ml YPD liquid. When cells reached an OD600 of 0.8 (usually after 3 hours) cells were centrifuged at 4000rpm for 5 minutes. The cell pellet was washed twice with 1 ml sterile distilled water. Pellet was re-suspended in 1 ml of sterile 0.1M LiAC and divided in two halves (one to be used as negative control for transformation). Cells were again centrifuged at 13000rpm and supernatant was removed. To each of the pellet, 240µl PEG (Sigma P4338), 36 µl of 1M LiAc (Sigma L4158), 50 µl of activated salmon sperm DNA (2mg/ml) (Sigma D1626) and ~1 µg DNA in a 50 µl volume (or 50 µl sterile distilled water for the negative control) were added without mixing until end. (The salmon sperm DNA was activated by boiling at 95˚C) After addition of all the reagents cells were gently mixed by pipetting, and incubated at 30˚C for 30 mins. Cells were then transferred to 42˚C and kept for 30mins. After heat shock 1 ml sterile distilled water was added and cells were centrifuged at 4000rpm. When selecting for prototrophy, cells were resupended in 500 µl sterile distilled water and plated on appropriate media plates. When selecting for drug-resistance cells were re-supended in 1ml YPD without any antibiotic and were incubated for 3 hours at 30 ˚C. Afterwards cells were again centrifuged at 4000rpm as earlier and were re-supended in 500µl sterile distilled water and were plated on appropriate media plates. All the plates were

incubated at 30 ˚C for 3 days and checked for the presence of colonies. Transformations were also verified by PCR.

2.3.4 PCR based gene deletion

In order to replace the gene of interest with the selectable marker, a PCR amplification of Longtine plasmid (Longtine et al, 1998) was carried out. Primers used for this amplification were specially designed in a way that they will have nearly 45 bp homology to the regions immediately outside the gene of interest and nearly 20bp homology with the marker cassette.Further for the gene deletion, this amplified DNA fragment having homology was used as the DNA for LiAc transformation (method 6). The homology between gene of interest and the cassette helps the recombination allowing the replacement of the gene with the marker cassette, which can be easily selected on the marker specific plates. Transformations were also verified by junction PCR.

2.3.5 Generating diploid strains

To generate required diploid yeast strains, specific haploid strains with opposite mating types were mixed in small but equal amounts on a YPD plate. The plate was further incubated for minimum of 5 hours at 30˚C. Diploids were either selected by replica plating them on plates having both selectable markers or by picking zygotes using a microscope as mentioned in method 9.

2.3.6 Selecting zygotes

For the strains with lower sporulation efficiency especially S288C, to achieve greater sporulation zygotes were pulled before transferring to pre-sporulation media. For this after 4 hours of mating of haploids, using a micromanipulator microscope zygotes were isolated and placed on a clear space on the same YPD plate and grown at 30˚C for 2 days.

2.3.7 Dissection of tetrads

In order to dissect tetrads, diploids were generated as mentioned in method 8. Depending on the strain background, diploids were either directly transferred to 5ml of 1%KAC liquid (W303) or transferred to the pre-sporulation media for at least 16 hours (S288C) and then to 5 ml of 1% KAC liquid. After 2 days of shaking at 200rpm at 30˚C, cells were checked for presence of tetrads under a light microscope. Further samples were centrifuged, and cell-pellet was re-suspended in 100µL of dissecting buffer (10 mM EDTA, 1M Sorbitol and 10mM NaH2PO4), and 5µL of 10 mg/ml Zymolyase (20T), cells were incubated at 37˚C for 30mins. Further 400µL of dissecting buffer was added after 30mins, and samples were used for dissection or stored at 4˚C. For dissection, small amount of cell-wall digested sample was streaked on flat and thin YPD plate. Further with the help of a microscope and micromanipulator, each ascospore was separated from the tetrad and moved to a different location, all the 4 spores were aligned horizontally in a line of 4. Plates were incubated at 30˚C for 2 days for spores to germinate, and then transferred to appropriate selective media plates to select for the required markers.

2.3.8 Spot test:

Strains were grown overnight at 30˚C in 5ml liquid media, shaking at 200rpm. Next day cell density was measured in terms of OD600. A suspension was made by diluting cells equivalent to 0.6 OD600. 10µl of the suspension was added per spot. 4 spots were prepared for each sample on the required plate for example YEPEG, Canavanine containing plate etc. Next day plates were checked and images were taken using the Gel Doc XR+ Imager.

2.3.9 *CAN1* forward mutation assay (Qualitative assay):

The strains were grown overnight in the respective media. Next day the optical density at 600nm (OD600) was measured and adjusted to 0.6. 10 µl of the cell suspension was loaded on Canavanine plates also on arginine drop out plates to be considered as control. These plates were incubated at 30˚C for 3 days for visual assessment of mutation spectra. Canavanine resistant colonies were visually assessed compared with the wild type. Images were captured using Gel Doc XR+ Imager.

2.3.10 Fluctuation test

Strains were streaked for single colonies and plates were incubated at 30° C for three days. 5 independent colonies were inoculated in 5ml complete media (or selectable media for plasmid in case strains have plasmid) and grown until they reach stationary phase at 30° C shaking at 200 rpm. Cells were spun down at 1,258 × g for 5 minutes and re-suspended in 800µL of sterile water. The volume of each cell pellet was carefully noted after re-suspension in water with the help

of pipette. This neat solution was then appropriately diluted for different strains to achieve countable numbers of colonies. 100 µl solutions were plated out on complete plates and canavanine plates. After incubation for three days, the number of colonies were noted. Mutation rate for each strain was calculated using the equation R_o= M (1.24 + ln M),

where R_o is the median number of canavanine-resistant colonies,

M is the average number of canavanine-resistant colonies per culture. The M value was determined by interpolation using Goal Seek analysis in Microsoft Excel and then it was used in the second equation:

R= M/N,

where N is the average number of cells per culture and R is the mutation rate. An average mutation rate was calculated from mutation rates obtained from the three independent experiments.

2.3.11Growth curves:

Strains were grown overnight in YPD (till they reach stationary phase). They were then counted using Haemocytometer and 2×10^6 cells were inoculated in Complete and Canavanine containing media. The optical density at 600nm (OD600) was measured every hour till 12 hours.

2.3.12Mutagenesis Screen

The EUROSCRAF library Genes were created using a PCR-based gene deletion strategy. The gene of interest was replaced by mitotic recombination with the KanMX cassette, which confers resistance to kanamycin, allowing the cells to grow on G418 containing media. A tertiary library of the 543 deletion strains which showed decreased mutation rates compared to the wild type in a hypomorphic MMR background is created. This library was then crossed with a strain in which *MLH1* is replaced by NATmx4 cassette, which confers resistance to NAT. Diploids are then selected on YPD +NAT +G418 plates. Selected diploids are then replica plated on 6% YPD plate (pre sporulation media) and incubated overnight at 30°C and then transferred to 1% KAC plates next day. Cells are incubated for 3 days at 30°C and then transferred to YPD +NAT +G418 + cyclohexamide. As the cyh-resistant haploid strain will have relatively higher resistance to cyclohexamide over a CYH2/cyh2 heterozygous diploids concentration of 8µg/ml cyclohexamide was used to select only haploids along with having both G418 and NAT resistance. These haploid cells are then transferred to Arginine dropout plate to select for the cells which don't have a defect in Arginine synthesis pathway, cells are incubated overnight at 30°C and then transferred to Canavanine plates and incubated for another 3 days for visual assessment of mutation spectra (Fig2.1).

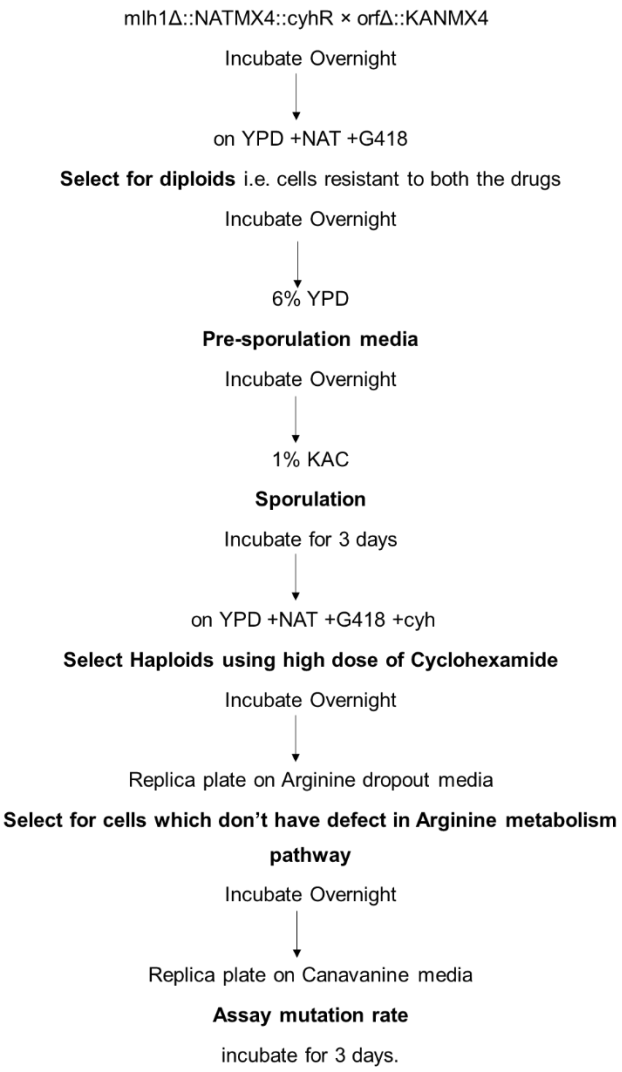


Fig 2.1 Schematic representation for selection of haploids having gene of interest and *mlh1* deleted.

To get the haploids having strain of interest deleted together with *MLH1*, parent strains where each parent containing one of the gene deleted with one marker cassette were crossed on YPD. In order to get better sporulation efficiency, first diploids were selected using a media containing both antibiotics corresponding to the markers of parents, thus making only diploids having both the markers can grow , further cells were transferred to pre-sporulation media and then to 1%KAC to get spores; haploid cells were then selected using cyclohexamide resistance along with the both parents specific antibiotics; these haploids were futher selected for growth in absence of arginine and then transferred to canavanine media to assay mutation rate.

2.3.13Storage of yeast strains

Yeast cells were grown overnight at 30˚C by shaking at 200rpm in 5 ml YPD. Next day cells were centrifuged and re-suspended in a screw-top tube containing 1 ml 30% glycerol and stored at -80˚C.

2.3.14Colony PCR:

Colony PCR was used for verification of gene deletion, a fresh colony or a small amount of a fresh colony (around 1×10^3 to1×10^5cells) was re-suspended in 100µl on NaOH and boiled for 5 minutes at 95 °C. The samples were then centrifuged and cooled on ice. 1µl of the suspension was used as a template DNA and PCR was carried out same as general PCR (method 2.2.15). Samples were stored at 4 °C up to 1 month.

2.3.15General PCR:

PCR was used to amplify DNA for cloning and verification of gene deletion. Two different polymerases PrimeSTAR HS DNA Polymerase (for gene amplification) and DreamTaq™ (for verification of gene deletion) polymerase were used. Reaction mixture was prepared according to following table:

Component	Volume/ Reaction (µl)	Final concentration
10x DreamTaq™ Buffer	2	1x
dNTP mix	2	200 µM each
DreamTaq™polymerase	0.3	1.5units
Oligo mix	2	0.1µM (each oligo)
Template	1	40ng
Distilled water	10.7	
Total	20	

Component	Volume/ Reaction (µl)	Final concentration
10x Primestar Buffer	4	1x
dNTP mix	1.6	200 µM each
Primestar polymerase	0.4	1 unit
Oligo mix	1	0.05 µM each (oligo)
Template	1	1ng
Distilled water	12	
Total	20	

All the reactions were carried out using Eppendorf Mastercycler RP Gradient-S

35 cycles with the conditions as follows:

Initial denaturation	95°C	5 min
Denaturation	95°C	30 s
Annealing	55°C	30 s
Extension	72°C	1 minute/kb of amplification
Final extension	72°C	5 min

Samples were then kept at 4 °C, and run on 1% agarose gel for verification.

2.3.16 Agarose Gel Electrophoresis:

Agarose gel electrophoresis was used to verify the size of DNA fragments after various experiments like PCR, restriction digestion etc. Depending on the fragment size and distinction efficiency required the agarose percentage of the gel was selected. For example, for smaller fragments higher percentage of agarose was used. For standard separation 1% agarose-TAE gel containing 5 µM ethidium bromide was used. Gel was run in 1× TAE buffer, with a voltage of ~4 V/cm between the electrodes (typically 60-100 V). Gel was visualized using Bio-Rad ChemiDoc imaging system.

2.3.17 Protein Extraction using NaOH

Cells were grown overnight in 5ml liquid media. Cells were then centrifuged, and cell pellets were resuspended in 100µl of 0.2M NaOH which was then

incubated at room temperature for 5 minutes; centrifuged and pellets were resuspended in 100µl of SDS-PAGE sample buffer (0.06M Tris +β mercaptoethanol). Samples were boiled for 5 minutes at 95°C in a heat block and then centrifuged at full speed for 2 minutes at 4°C.

2.3.18Protein Extraction using TCA

Cells were grown overnight in 5ml liquid media. After checking if all the samples have similar OD_{600}, 2 ml of cell culture was harvested. Cells were centrifuged at 4000 rpm at 4 °C for 10 minutes, in 14 ml polypropylene tubes. Cell pellets were re-suspended in 500 µl cold 20% TCA and transferred to 2 ml Ribolyser tubes. Tubes were then centrifuged at 4°C for 10 minutes at 10,000 rpm in a cooling bench-top centrifuge to discard 20% TCA, and re-suspended in 200µl cold 10% TCA for immediate processing. Acid-washed glass beads (425– 600µm, Sigma) were added and tubes were ribolysed three times in Ribolyser set to 6.5 m/s for 60 sec each time with 5 minute resting of the tubes on ice between each cycle. The supernatant was collected from the beads and placed into a fresh 2 ml Eppendorf tube. The beads were washed 3 times (once using 200µl of cold 10% TCA and a additional two times with 400µl cold 10% TCA) vortexing thoroughly during each wash and the supernatant was collected. Tubes containing supernatant were centrifuged at 5000 rpm for 10 minutes, at 4 °C. The cell pellet was thoroughly re-suspended in 200µl 4x sample buffer (0.25 M Tris-HCl pH 6.8, 8% SDS, 10% ß-mercaptoethanol, 30% glycerol, 0.02% bromophenol blue) and 35µl of 1.5 M Tris-HCl pH 8.8 to neutralise the acidity. Samples were boiled at 95°C for 5 minutes and immediately centrifuged at max-

speed for 2 minutes, at room temperature. The supernatant was collected and stored at -20°C for future use.

2.3.19Western Blotting / Western blot analyses

The primary antibodies used in this study were: anti-hMLH, 1:500 in BSA-PBS-Tween; PGK1, 1:2500 PBS-Tween. The samples were separated on 8% SDS–PAGE gel and transferred to a membrane by semi-dry transfer using Trans-Blot® SD Semi-Dry Transfer Cell. The membranes were blocked with 5% BSA-PBS-tween for 30 min, incubated overnight with primary antibodies, washed three times with PBS-tween for 10 min each, incubated with the secondary antibody (anti-mouse IgG, 1:5000 in PBS-tween) for 60 min and , washed three times with PBS-tween for 10 min each, and proteins were detected using ECL (Thermo scientific, # 1859698).

2.4 Mammalian Methods

2.4.1 Cell lines used

The ideal system to assess spontaneous mutation rate in absence of MMR would be using cells in which the expression of *hMLH1* can be tightly regulated for example by with the TetOff system. However the previously characterized, human embryonic kidney cell line 293T with TetOff system for *hMLH1*, has more than one copy of X-chromosome and therefore of *HPRT* gene, this makes it difficult to detect mutation rates using the *HPRT* assay (Cejka, 2003).

Instead, I chose to use HCT116 cells, a human colorectal carcinoma cell line that has a mutation in *MLH1* gene making them MMR deficient. As a control for functional MMR, the HCT116 cell line with extra chromosome 3 for correction of *MLH1* defect was used (HCT116 + chr3). The use of HCT116 cells with added chromosome 3 has been used extensively in the field, however, it does have the disadvantage that not only is hMLH1 expressed, but so are nearly 1,800 other genes. Thus, the endogenous mutation rates may be affected by over expression of nearly 1,800 other genes.

2.4.2 The hypoxanthine phosphoribosyl transferase (*HPRT*) assay

For the detection and quantification of mutation frequencies in mammalian cells, various assays targeting different genetic loci can be used. In our lab we started optimizing *HPRT* assay, because of its similarity with *CAN1* assay with respect to the type of mutations that can be detected using these assays. The *HPRT* gene is located on an X-chromosome. As large deletions in X-chromosomes are lethal this mutation assay only allows detection of small changes such as point mutations and exon deletions.

The *HPRT* gene encodes for Hypoxanthine-guanine phosphoribosyl transferase that is a key enzyme of the purine salvage pathway. Cells having functional *HPRT* can incorporate 6-thioguanine (6-TG), which is a toxic analogue of guanine leading to cell death, however mutational inactivation of *HPRT* can lead to cell survival which can be quantified as number of colonies formed in presence of 6-TG (Fig.2.2). Before analysing the mutational frequency of a cell it is important to cleanse cells for previously accumulated mutations in *HPRT* locus. Nucleotides are synthesized either by endogenous pathway which require dihydrofolate reductase, or a salvage pathway which require functional *HPRT*. To cleanse the cells having a *HPRT* mutation, cells are treated with hypoxanthine-aminopterin-thymidine medium (Bhatia et al., 1996). This medium contains an aminopterin, which blocks the endogenous synthesis of nucleotides by inhibiting dihydrofolate reductase. This causes cells to rely on the salvage pathway for synthesis of nucleotides, and as a consequence, the cells that

already contain a mutated *HPRT* gene die. During the cleansing step the media is also supplemented with hypoxanthine and thymidine, which are required for the salvage pathway (Fig 2.3). After cleansing the population of cells that contain pre-existing *HPRT* mutations, we treated cells with the desired siRNA and cultured for at least 10 days before plating on 6-TG containing medium to quantify mutation frequencies.

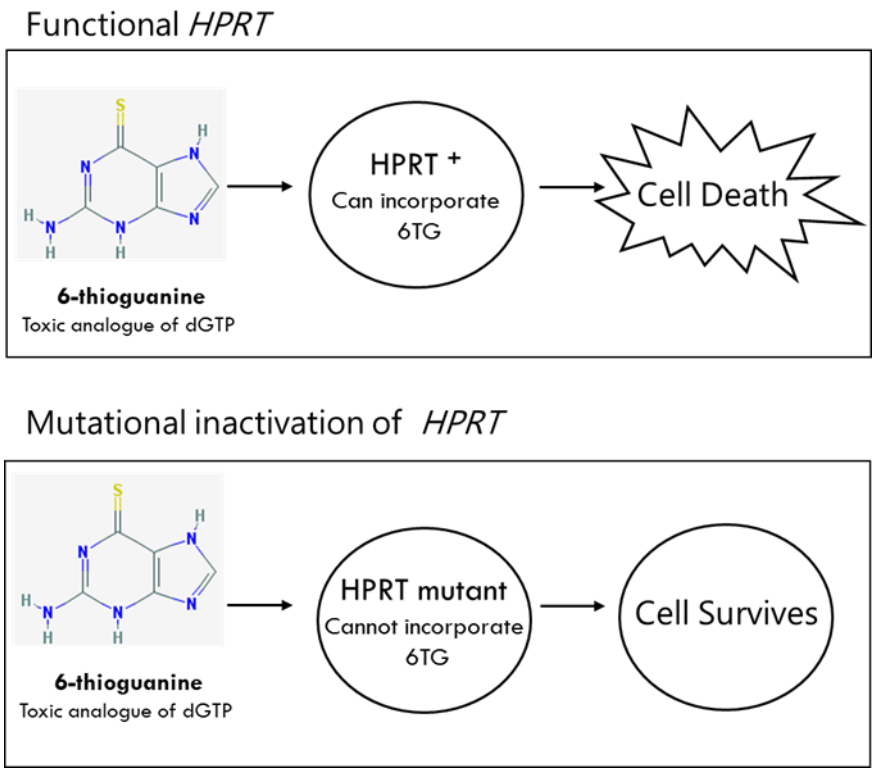


Figure 2.2: The hypoxanthine phosphorybosyl transferase (*HPRT*) mutation assay.
HPRT proficient cells incorporate 6-thioguanine which is the toxic analogue of dGTP into the DNA and die, however *HPRT* deficient cells not incorporate this toxic analogue into their DNA and they survive.

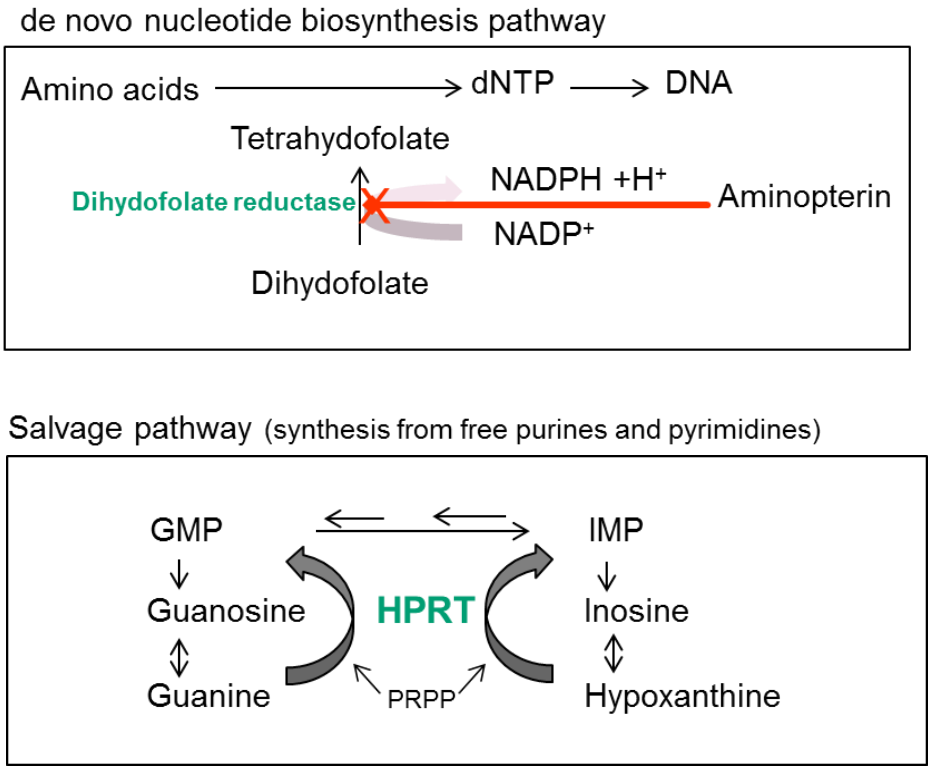


Figure 2.3: Cleansing of *HPRT* mutants

A) To 'cleanse' the previously occurred mutations, cells are treated with hypoxanthine–aminopterin–thymidine (Bhatia et al.) media, which inhibits the endogenous de novo nucleotide biosynthesis pathway by inhibiting dihydofolate reductase, forcing cells to use salvage pathway for dNTP synthesis, cells that are incapable of using this pathway can no longer divide and undergo cell death.

B) In absence of guanine**,** Salvage pathway for synthesis of nucleotides require hypoxanthine, which can be converted to guanine by *HPRT* and then inserted into DNA, however cells having mutation cannot synthesis guanine and thus die.

2.4.3 Cell maintenance

MMR deficient human cell lines HCT116 (MLH1-defective),and were maintained at 37 °C, 5% CO2 in Dulbecco's Modified Eagle Medium (DMEM) (Gibeco 21969) supplemented with 10% foetal calf serum (v/v) (standard quality from PAA,A11-201), 1% penicillin- streptomycin (v/v) (10,000 units/ml Penicillin; 10,000µg/ml Streptomycin) (Gibeco 15140), and 1% L-glutamate (v/v) (Gibeco 25030).

293Tα cells, containing the Tet-Off expression system for *hMLH1* expression were maintained at 37 °C, 5% CO_2 in Dulbecco's Modified Eagle Medium (DMEM) (Gibeco 21969) supplemented with 10% foetal calf serum (v/v) (standard quality from PAA, A11-201), 1% penicillin- streptomycin (v/v) (10,000 units/ml Penicillin; 10,000 µg/ml Streptomycin, Gibeco 15140), 1% L-glutamate (v/v, Gibeco 25030),100 µg/ml Zeocin (Invitrogen R250-0), 300 µg/ml Hygromycin (Sigma-Aldrich-H7772) and 50 ng/ml doxycycline (Sigma-Aldrich-D9891) for induction.

2.4.4 siRNA-induced gene silencing

Cells were trypsinized and the number of viable cells was counted using a haemocytometer. For western blot analysis, 1×105 cells were plated per well. The plates were incubated overnight at 37 °C, 5% CO2, to allow attachment. Cells were then treated with 25nM siRNA, and then subsequently incubated for 48hrs. at 37 °C, 5% CO2, with siRNA and antibiotics/ drug free media. After 42 hours, cells were trypsinized and the number of viable cells was counted using a haemocytometer. Cell lysates were prepared by sonication (3 times 10 sec each) in 8M urea lysis buffer (200µl).

siRNA used in this study

	siRNA sequence	siRNA name
MCM3AP	GGCGGCUCAGAAACAAGAC	MCM3AP_1
MCM3AP	GUUCAUGGGAGAUGAAGGC	MCM3AP_2

3 Chapter 3 - Genome wide screen to understand regulation of mutagenesis in absence of mismatch repair

3.1 Introduction

Eukaryotic organisms maintain genomic stability via a number of conserved cellular mechanisms, which prevent or reduce spontaneous mutations. Inactivation of these processes can lead to increased carcinogenesis and ageing. Unfortunately, our current understanding of the regulation of mutagenesis via these pathways is insufficient to guide the development of more effective approaches for cancer prevention and early detection. One pathway of interest is DNA mismatch repair (MMR), which increases the fidelity of DNA replication by correcting mismatches and frame shift mutations generated primarily during DNA replication. Using various mutation rate assays, the deletion of genes within the MMR pathway has been shown to significantly increase mutagenesis.

A previous post-doctoral fellow, Dr. Judith Offman carried out a genome-wide screen in *S. cerevisiae* to identify genes that have a minor effect on MMR activity when alone, but when combined cause a significant decrease in MMR activity (hypomorphic MMR). For example, compromising MMR in *S. cerevisiae* through a deficiency in the ATP hydrolytic activity of Mlh1 (mlh1- E31) leads to a minor increase in the mutation rate; however, when combined with a deletion of

the exonuclease *EXO1*, the mutation rate is similar to a complete loss of MMR (Tran et al., 2001) (Figure 3.2). The absence of MMR activity also allowed us to identify novel substrates for the MMR pathway, such as metabolism. In the screen, spontaneous mutation rate was qualitatively measured using a *CAN1* forward mutation rate assay. The assay takes advantage of the *S. cerevisiae CAN1* gene, which encodes an arginine permease responsible for transport of arginine across the membrane. The mutational inactivation of *CAN1* results in resistance to canavanine (a toxic analogue of arginine), as cells cannot import canavanine from the external medium, thus allowing mutants to grow on canavanine containing medium (Grenson, 1966) (Figure 3.1A).

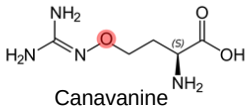
In the primary screen, 4,847 deletion strains were screened for altered mutation rate. In this screen 1,478 strains influenced the resistance to canavanine (Figure 3.1B). All 1,478 strains from the primary screen were re-examined for *CAN1* forward mutation rate using four independent transformants (Figure 3.1C). After two rounds of screening, 163 deletion strains (3.2% of the total number of strains screened) showed increased mutation rates when compared to compromised MMR activity controls. Surprisingly, 543 deletion strains (10.6% of the total strains screened) showed a decrease in mutation rate. For my project, I validated the screen to evaluate the effect of hypomorphic MMR in yeast deletion stains and investigate the role of MMR in shaping the mutagenesis landscape in *S. Cerevisiae*.

Figure 3.1: Genome-wide screen for mutants that affect mutation rates in hypomorphic MMR background.
A) *CAN1* forward mutation assay. *CAN1* encodes for an arginine permease which transports arginine and its toxic analogue canavanine across the membrane; mutational inactivation of *CAN1* blocks transport of canavanine thus cells can grow on canavanine containing media; and mutation rate can be assessed qualitatively as well as quantitatively

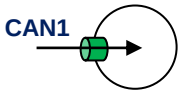
B) In the primary screen, the entire library was transformed with *hMLH1* expression plasmid. *hMLH1* interferes with yeast MMR resulting in the dominant mutator phenotype. Transformants were selected on leucine omission media and replica plated onto canavanine plate to assay *CAN1* forward mutation rates. In the primary screen 1,478 strains were identified that influence the resistance to canavanine.

C) In the secondary screen, four independent transformants of the strains of interest identified from the primary screen were re-patched in and then replica plated onto canavanine plates. The arrows show increased (Frederico et al.)(Frederico et al.) or decreased (blue) mutation rates compared to the wild type in a hypomorphic MMR background. The secondary screen revealed 543 deletion strains, which showed decreased and 163 deletion strains, which showed increased mutation rates compared to the wild type in a hypomorphic MMR cells (WT +*hMLH1*). Control wild type with only plasmid is also shown (WT +vector) (exp ID: E61_as02).

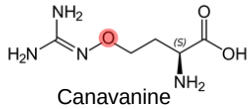
A) **Functional *CAN1* gene**



Canavanine
Toxic **analogue** of Arginine



Mutational inactivation of *CAN1* gene

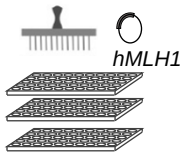


Canavanine
Toxic **analogue** of Arginine

Colony formation

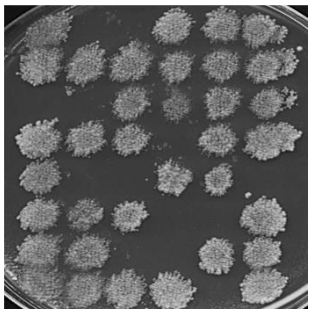
B) **Primary Screen**

disruption of yeast MMR by *hMLH1* expression



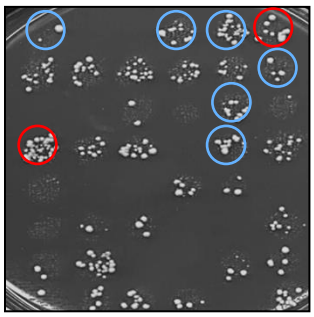
~4800 deletion Strains

Selection of transformants



Leu-

Mutation rate assay



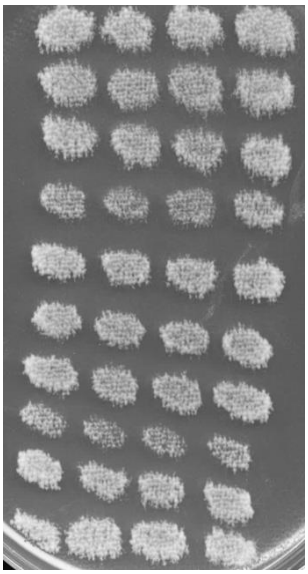
Canavanine

C) **Secondary Screen**

Independent transformants

1 2 3 4

WT + *hMLH1*
rrg1Δ
pet100Δ
ydr090c1Δ
ste5Δ
mtq2Δ
msp1Δ
mdm34Δ
spt4Δ
WT+ vector

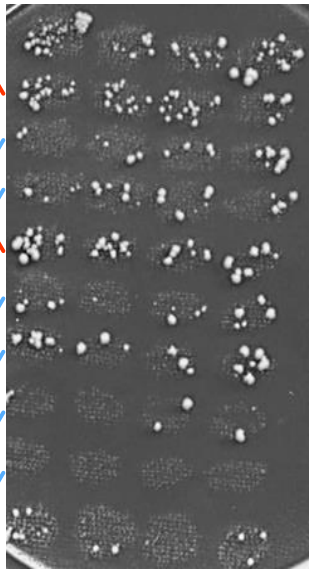


Leu-

Mutation rate assay

1 2 3 4

WT + *hMLH1*
rrg1Δ ↑
pet100Δ ↓
ydr090c1Δ ↓
ste5Δ ↑
mtq2Δ ↓
msp1Δ ↓
mdm34Δ ↓
spt4Δ ↓
WT+ vector



Canavanine

3.2 Expression of *hMLH1* in yeast can be used to mimic loss of ATP hydrolytic activity of Mlh1.

The genomic screen was carried out in *S. cerevisiae* cells with reduced MMR activity. To reduce the MMR activity in in *S. cerevisiae* cells, a human mismatch repair gene *MLH1* (*hMLH1*) was expressed in yeast using a low copy plasmid. The low-copy *hMLH1* expression vector, pCI-ML10, is a CENARS LEU2 vector which expresses *hMLH1* derived from wild-type *hMLH1* cDNA and leads to a dominant mutator phenotype in yeast (Shimodaira, 1998). Even though the exact mechanism of this mutator phenotype is not known, based on the high degree of sequence similarity between human and yeast Mlh1 protein one likely explanation, is that it could form non-functional complexes with yeast MMR proteins that lowers the efficiency of yeast MMR. This strategy for inducing a hypomorphic MMR phenotype was used as it was a high throughput screen and it was faster to express *hMLH1* using a plasmid than creating or crossing strains with the entire yeast deletion library of around 4800 deletion strains. Therefore, it was important to assess quantitatively that expression of *hMLH1* can be used to mimic lower MMR activity in yeast such as observed in hydrolytic mutant MLH1-E31A.

To test the effect of *hMLH1* expression in exonuclease deficient strain *exo1*Δ, and hydrolytic activity deficient strain *MLH1-E31A,* deletion strains were transformed with either *hMLH1* or vector only, and the mutation rates were assessed using *CAN1* forward mutation rate assay. In WT cells, expression of

hMLH1 leads to increase in mutation rate compared to vector only, but is lower than *mlh1*Δ, suggesting that there is still some MMR activity left (Figure 3.2). For the hydrolytic activity mutant, *MLH1-E31A* strain too*,* expression of *hMLH1* causes increase in mutation rate compared to vector only and results in complete loss of MMR. For the *exo1*Δ, expression of *hMLH1* leads to complete loss of MMR and increases mutation rate similar to *mlh1*Δ. This observation suggests that expression of *hMLH1*, can be used to mimic loss of *MLH1* hydrolytic activity; and thereby can allow us to identify downstream targets in this pathway.

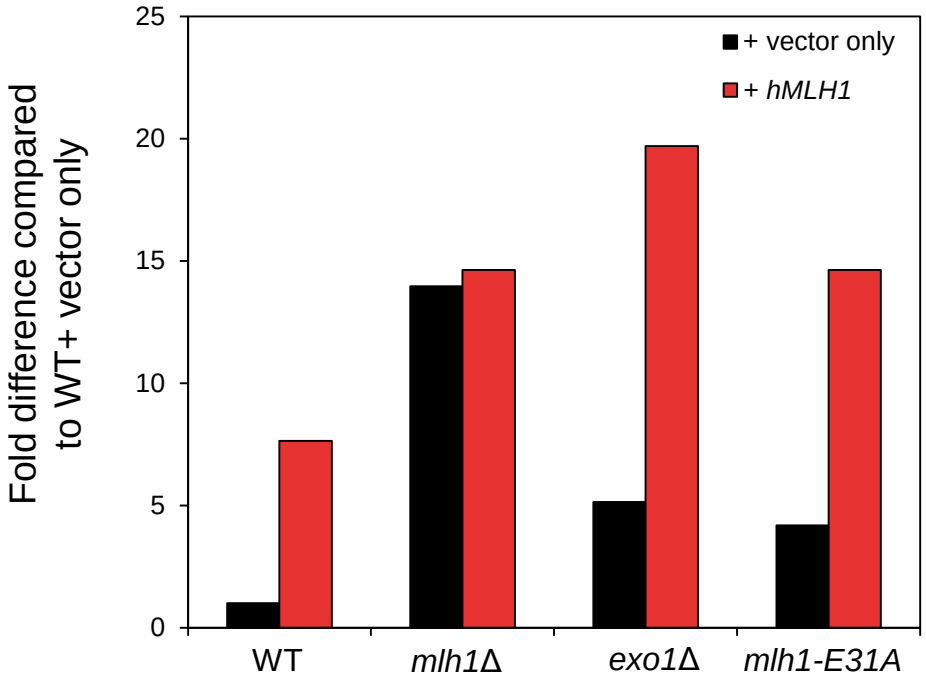


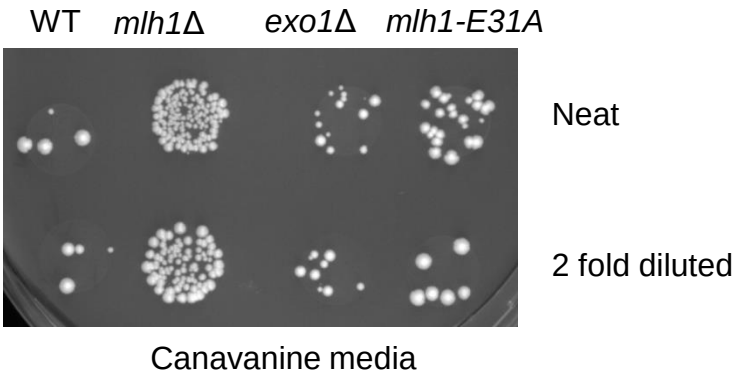
Figure 3.2: Expression of *hMLH1* in yeast disrupts Mismatch repair and can be used to asses additive effect

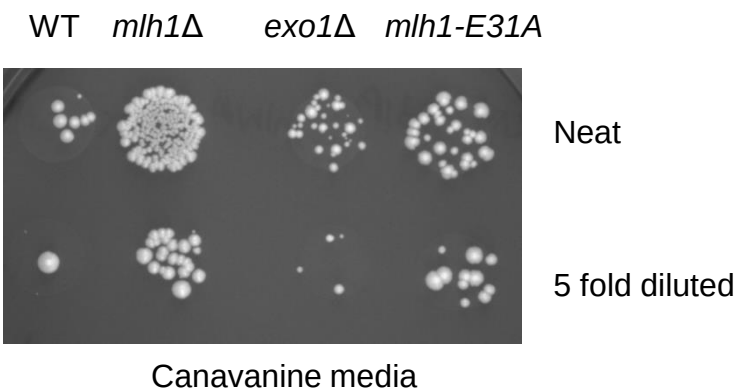
Fold differences relative to WT with vector only; for *CAN1* forward mutation assay when either vector only or *hMLH1* is expressed in mismatch repair yeast strains. Mutation rates were determined using method of median, five colonies per gene deletion were used. Black bars represent gene of interest deletion with vector only. red bars represent gene of interest deletions with *hMLH1*. (exp ID: E58_as01)

3.3 Five fold difference in *CAN1* forward mutation rate can be detected qualitatively.

In the mutagenesis screen, deletion library strains were transformed with *hMLH1* and assessed for fold difference in mutation rate qualitatively. Therefore it was important to test the detection power for *CAN1* forward mutation rate assay.

To observe fold difference qualitatively WT, *mlh1*Δ, *exo1*Δ and *MLH1-E31A* strain were diluted two-fold, five-fold, and ten-fold and plated on canavanine containing media. Mutation rates were qualitatively assessed difference detection. Fig 3.3 shows that, up to five fold difference can be detected using the *CAN1* assay, and is sufficient to pick up genes that when combined with hypomorphic mlh1-E31A or *hMLH1* increase the mutation rate to *mlh1Δ* levels.





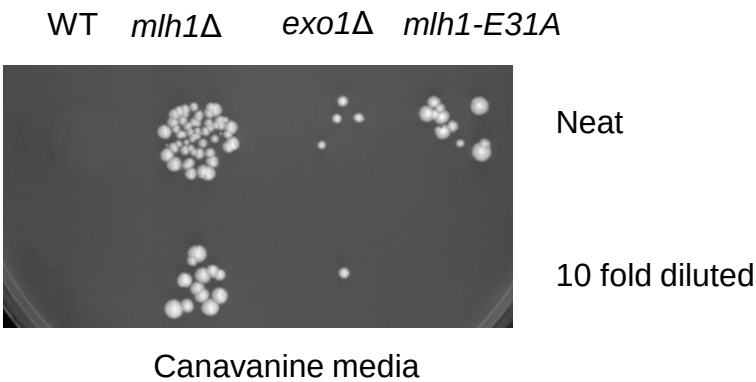


Figure 3.3 Difference of 5 fold in mutation rate detection can be detected qualitatively using *CAN1* forward mutation rate assay

As the primary screen was carried out qualitatively, different fold dilutions of WT, *mlh1*Δ, exo1Δ and *mlh1-E31A* were plated on canavanine media and mutation rates were assessed qualitatively.

3.4 In the screen, the influence on mutation rate is due to expression of *hMLH1*.

In the genome wide screen to reduce MMR activity, deletion strains were transformed with *hMLH1* low-copy expression plasmid (*LEU2*). This strategy allowed us to screen strains in hypomorphic MMR background, as expression of human protein does not abolish yeast MMR; with further advantage of being experimentally simpler and easier to assess the phenotype (Shimodaira et al., 1998). It is therefore important to first ensure that *hMLH1* is expressed in the strains which showed influence on mutation rates in the previous rounds of the screen.

To verify the expression of *hMLH1*, western blot, was carried out for 20 strains which showed lower mutation rate compared to WT when MMR activity was lower. These strains were randomly selected, and mutation rate and protein expression was verified for the same strains. An 80KDa protein band of hMlh1 is observed in strains containing the plasmid; whereas strains with the empty vector lack the band (Figure 3.4 B and D).

In *CAN1* forward mutation rate assay 18 out of 20 strains showed lower mutation rate compared to WT when MMR activity was lower due to expression of *hMLH1*(Figure 3.4A and C). As the *hMLH1* plasmid is marked with leucine marker, a leucine dropout plate was used as control to ensure presence of plasmid and normal growth for *CAN1* forward assay (Figure 3.4A and C).

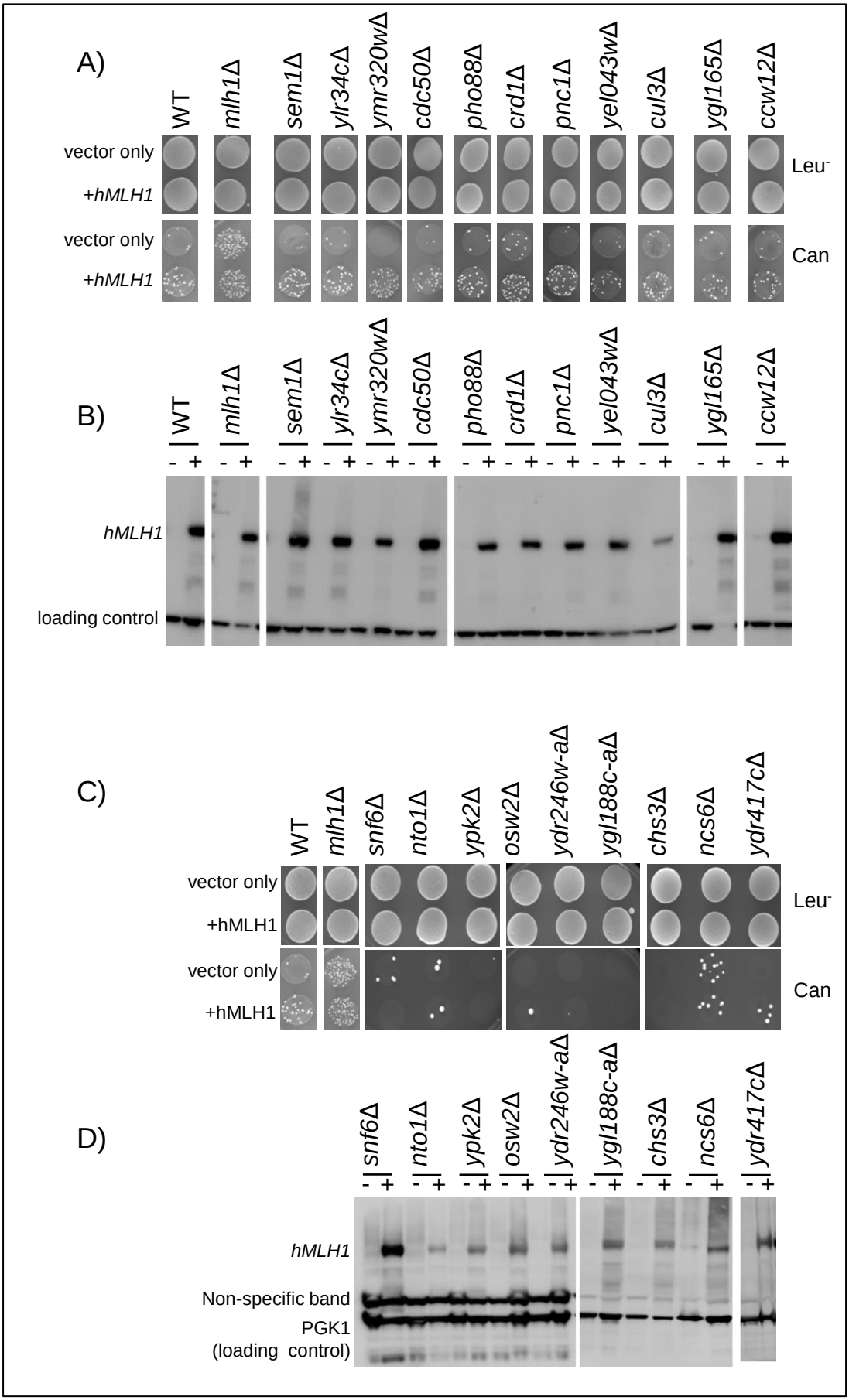
Thus it can be confirmed that *hMLH1* is expressed in these samples and the counterintuitive decrease in mutation rates compared to the wild type may be associated with *hMLH1* expression.

Figure 3.4: Confirmation of *hMLH1* expression for strains having lower mutation rate compared to WT when MMR is attenuated.
A) Qualitative representation of *CAN1* forward mutation rate assay, when either only vector or *hMLH1* is expressed in gene of interest

B) Western blot of the same strains from figure A where samples were extracted by TCA protein extraction (materials and methods 22) and probed with antibody anti-hMLH1, 1:500 in PBS-Tween, protein size around 80KDa. and anti PGK1 (as loading control), 1:25000 in PBS- Tween size around 50KDa.

C) Qualitative representation of CAN1 forward mutation rate assay, when either only vector or hMLH1 is expressed in gene of interest

D) Western blot of the same strains from figure A where samples were extracted by TCA protein extraction (materials and methods 22) and stained with antibody anti-hMLH1, 1:500 in PBS-Tween, protein size around 80KDa. and anti PGK1 (as loading control), 1:25000 in PBS- Tween size around 50KDa.


A)
WT
mlh1Δ
sem1Δ
ylr34cΔ
ymr320wΔ
cdc50Δ
pho88Δ
crd1Δ
pnc1Δ
yel043wΔ
cul3Δ
ygl165Δ
ccw12Δ
vector only
+hMLH1
Leu-
Can
B)
hMLH1
loading control
C)
vector only
+hMLH1
Leu-
Can
D)
snf6Δ
nto1Δ
ypk2Δ
osw2Δ
ydr246w-aΔ
ygl188c-aΔ
chs3Δ
ncs6Δ
ydr417cΔ
hMLH1
Non-specific band
PGK1
(loading control)

3.5 *CAN1* Forward mutation rate of reported gene deletions confirms the results of the screen.

In order to validate the screen further, *CAN1* mutation rates were calculated using fluctuation test (Method 2.3.10). Total 25 strains were used to calculate mutation rate in presence and absence of *hMLH1* plasmid. From the gene deletion strains which showed lower mutation rate in the screen eight strains were selected randomly, and mutation rate was calculated. For the strains which showed increased mutation rate in previous rounds of screening, Barbara Domanska carried out fluctuation tests for 17 genes associated with replication and repair were selected namely *RNR4*, *OGG1*, *ELG1*, *MGS1*, *MMS2*, *PSY3*, *RAD27*, *RAD34*, *RAD5*, *RNH201*, *TRM2*, *RRM3*, *RTT107*, *SHU1*, *SLX4*, and *TEL1*.

CAN1 forward mutation assay was calculated using by the method of median (Lea & Coulson, 1949). Mutation rates confirm that the deleted genes influence spontaneous mutation rate in MMR deficient background (Figure 3.5 and Figure 3.6) (representative images are shown in figure 9.1). In the figure the error bars represent experimental variation between three independent fluctuation tests. Out of the 17 genes associated with replication, deletion of 14 genes showed increase in mutation rate in presence of *hMLH1*; However deletion of *RNR4*, *MGS1* and *RAD34* did not increase mutation rate in presence of *hMLH1* when compared to WT with *hMLH1.* From remaining 14 genes five genes *RNH201, TRM2, RRM3, RTT107* and *TEL1* showed to have a multiplicative effect in the absence of functional MMR; however did not increase the mutation rate higher

than *mlh1*Δ , this is suggest that they might compensate for loss of MMR activity, this will be investigated further in chapter 4.

Mutation rates for the eight gene deletions previously shown to have a lower mutation accumulation compared to wild type expressing *hMLH1* are lower than WT with *hMLH1;* this confirms the previous findings from the screen. As the fluctuations test for these strains was only carried out once (using five independent colonies) there are no error bars. This lower mutation rate is unexpected but is important to understand as it suggests that these genes could have a potential role in rescuing the mutator phenotype of *MLH1 or* mutants which are defective in the Mismatch repair system.

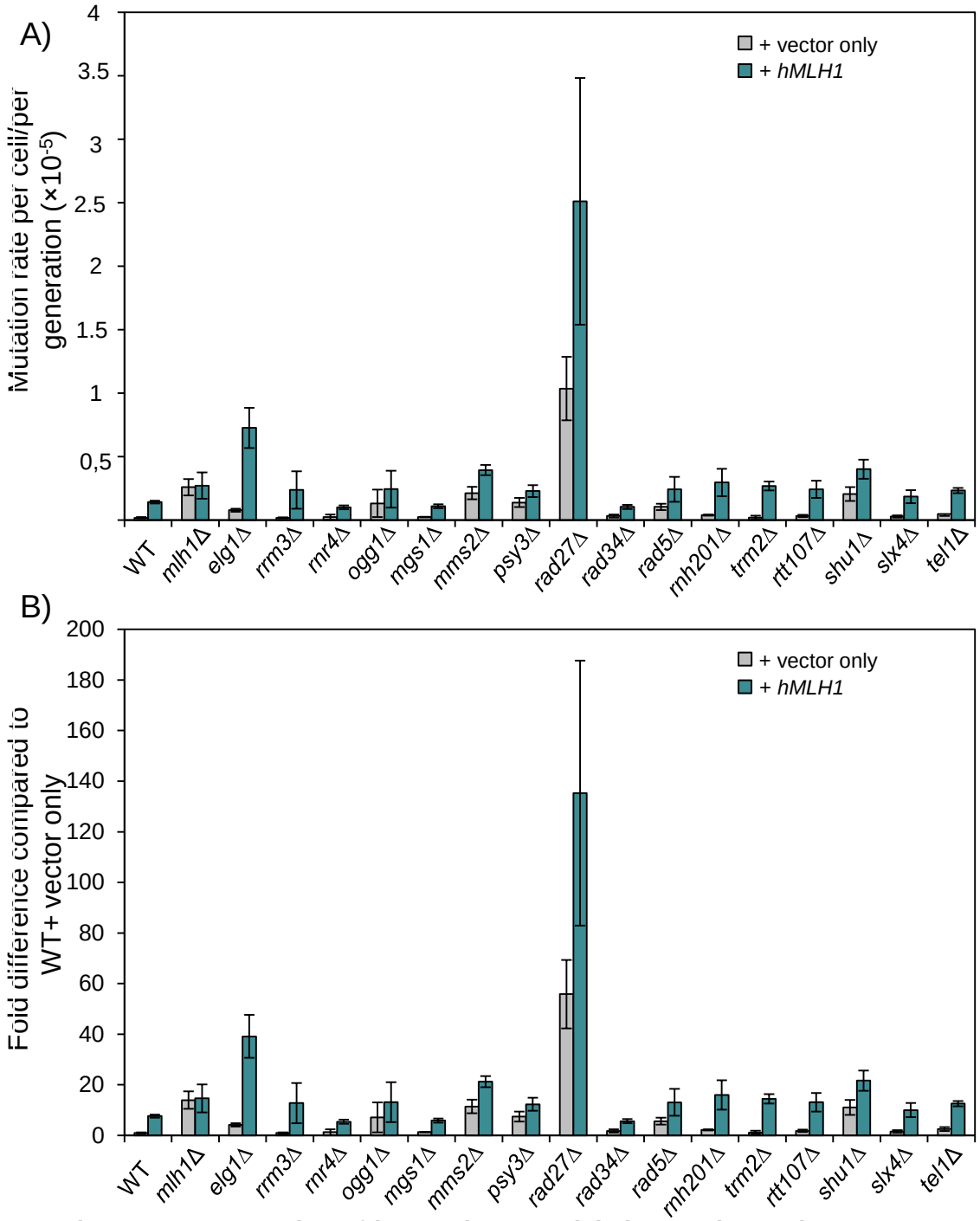


Figure 3.5: Expression of *hMLH1* in yeast deletion strains and assessment of mutations using *CAN1* forward mutation rate assay for deletion strains causing higher mutation rate compared to WT.

A) Average mutation rates per cell for deletion strains using *CAN1* forward mutation rate assay when either vector only or *hMLH1* is expressed. Data was obtained by Barbara Domanska. Mutation rates were determined using method of median, 3 independent experiments were performed each time seven colonies per gene deletion were used. Greay bars represent gene of interest deletion with vector only. green bars represent gene of interest deletions with *hMLH1*.

B) Fold differences relative to WT with vector only. grey bars represent gene of interest deletion with vector only. green bars represent gene of interest deletions with *hMLH1*. Data was obtained by Barbara Domanska (exp ID: E58_as01)

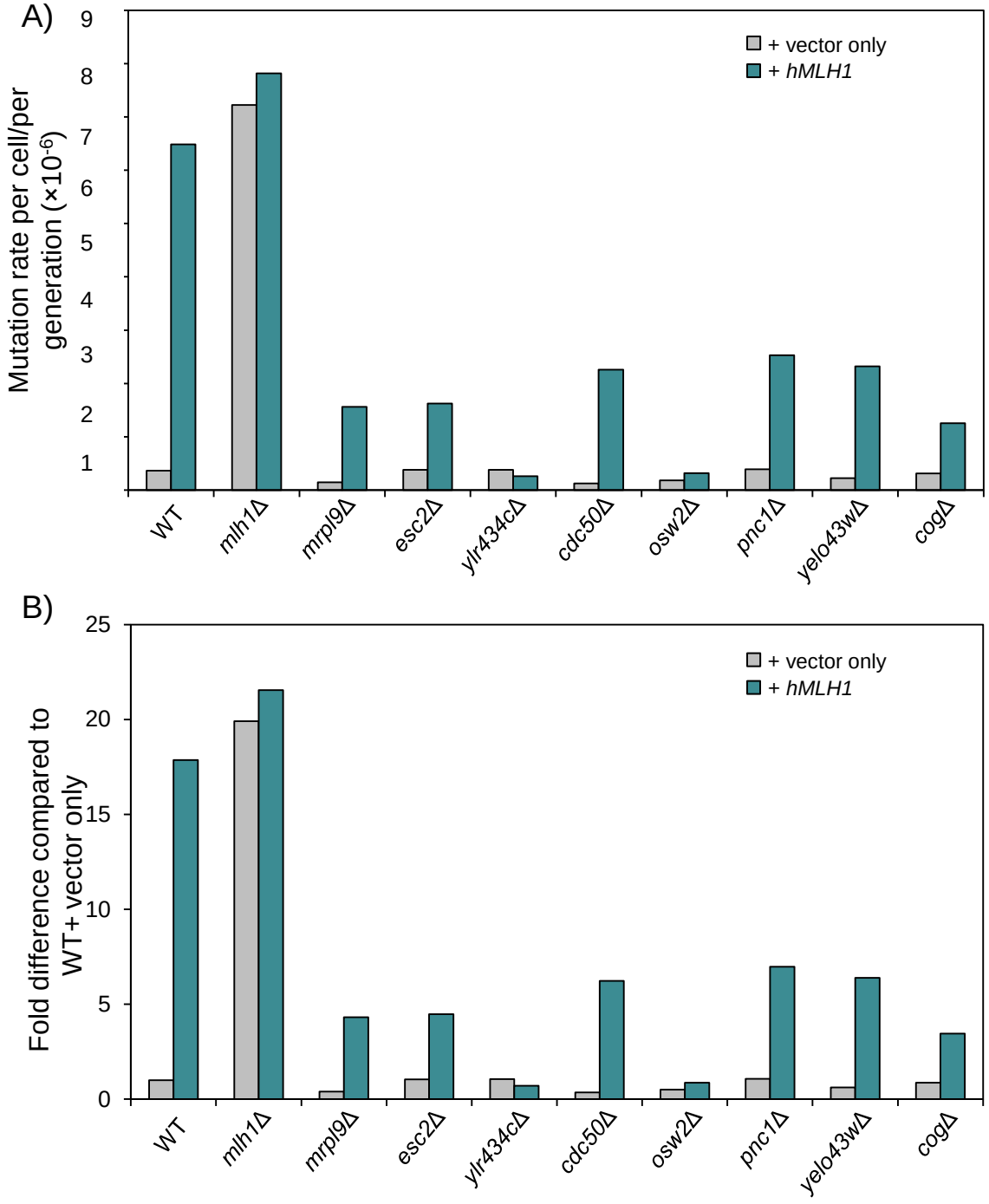


Figure 3.6: Expression of *hMLH1* in yeast deletion strains and assessment of mutations using *CAN1* forward mutation rate assay for deletion strains causing lower mutation rate compared to WT.

A) Average mutation rates per cell for deletion strains using *CAN1* forward mutation rate assay when either vector only or *hMLH1* is expressed. Mutation rates were determined using method of median, five colonies per gene deletion were used. bars represent gene of interest deletion with vector only. green bars represent gene of interest deletions with *hMLH1*.

B) Fold differences relative to WT with vector only. Black bars represent gene of interest deletion with vector only. green bars represent gene of interest deletions with *hMLH1*. (exp ID: E58_as01)

3.6 Identification of enriched biological processes amongst deletions resulting in a changed mutation rate

With the help of Database for Annotation, Visualization and Integrated Discovery (DAVID), biological processes enriched amongst the tertiary library genes were identified (Huang da et al., 2009). The DAVID database can be used to classify genes based on their biological process, cellular component or molecular function.

The screen revealed 163 deletion strains (3.2% of the total number of strains screened), with increased mutation rates in the compromised MMR background, including genes like *TRM2* and *RRM3* which showed a synthetically elevated phenotype similar to complete loss of MMR. The 163 genes that showed an increase in mutation rate are enriched in multiple biological processes like DNA repair, DNA replication (Fig3.7) (Table 9.3).

The 543 genes were found to be enriched in 97 biological processes. And when searched for specific pathways, Endocytosis and Ribosome pathways are shown to be enriched the most. The biological processes are further clustered manually by tracing the parent GO term and clearing of the redundant GO terms according to their function. Figure 3.8 represents the main biological processes enriched in gene deletions resulting in decreased mutation rate (Table 9.4).

When further analysed there was a sub-cluster of four genes which belong to transcription coupled repair (TCR) (*TFB5*, *SAC3*, *DEF1* and *THO2*). This cluster

was significant, as in yeast a total of 16 genes are involved in TCR, out of which only 12 were present in the deletion library and 5 out of 12 showed slow growth when deleted. Biologically this is interesting as this suggests that transcription and MMR could be associated. This hypothesis will be tested in chapter 4.

Collectively this genome-wide screen will help us better understand the processes affecting spontaneous mutagenesis in combination with MMR. The screen also revealed that mutation rate can also go down when MMR is compromised in combination with some genes. This finding will be useful in understanding the relationship of mutagenesis and cancer particularly for the sporadic as well as lynch syndrome tumours where MMR is disrupted.

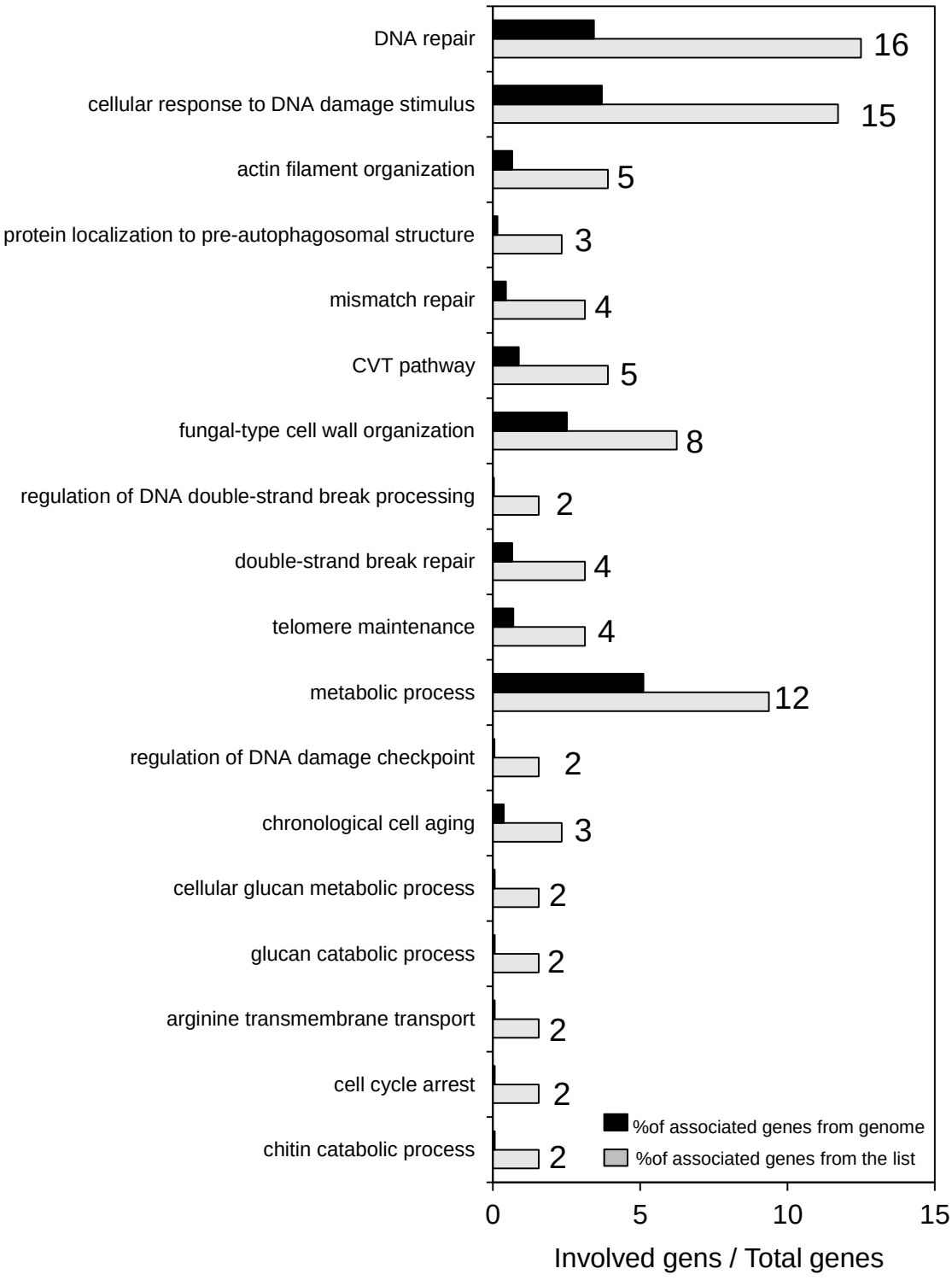


Figure 3.7: Biological processes enriched amongst deletions resulting in increased mutation rate

DAVID database was used to classify genes based on their biological process (exp ID: E80_as02_01).

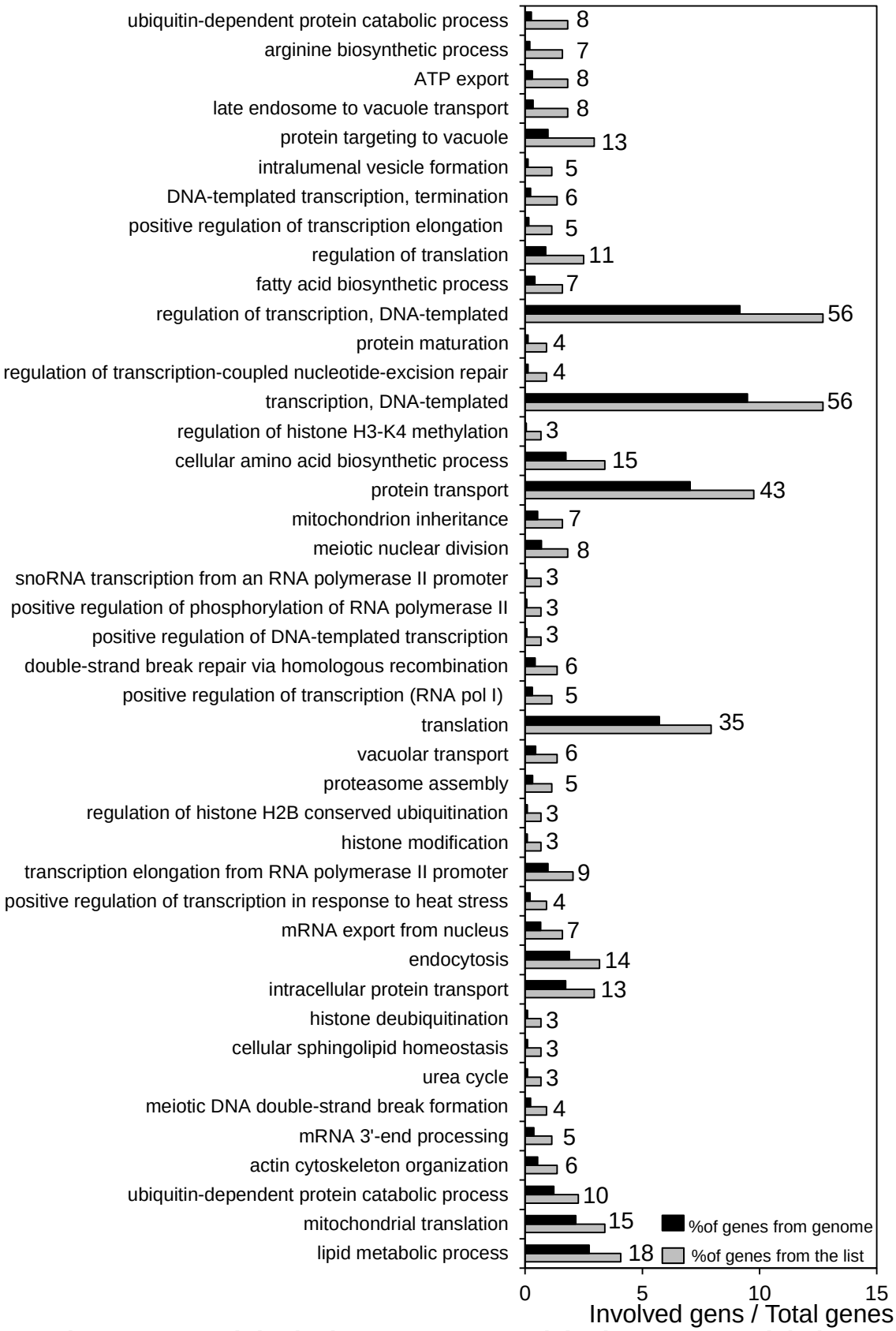


Figure 3.8: Biological processes enriched amongst deletions resulting in increased mutation rate

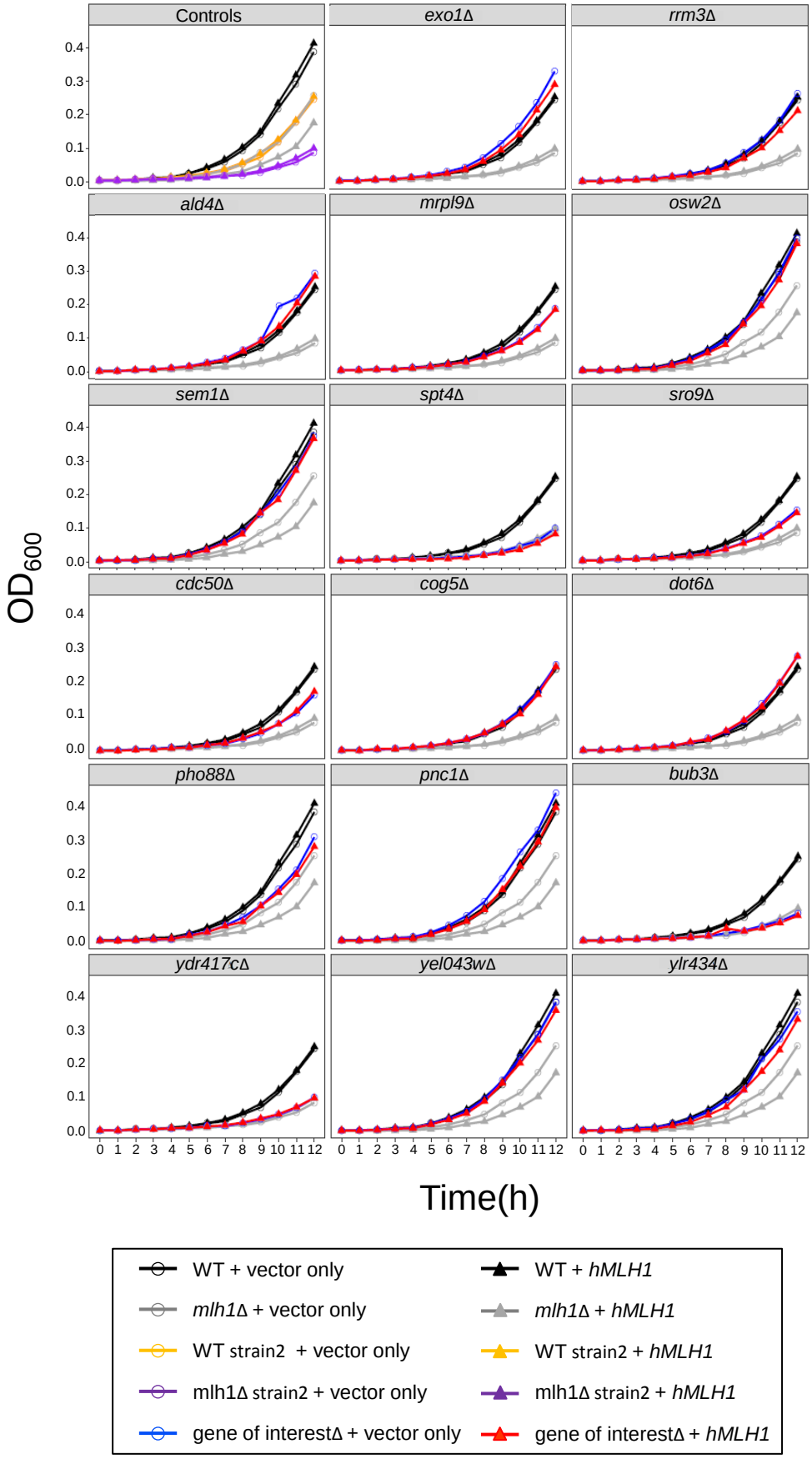
DAVID database was used to classify genes based on their biological process, (exp ID: E80_as02_01).

3.7 In the screen, the expression of *hMLH1* does not affect the growth of the strains.

One possibility for lower mutation rate is that the cells have slower growth leading to a lower accumulation of mutations. To ensure growth of the deletion strain is not the factor influencing mutation rate, growth curves of 17 strains were plotted. For the experiment 15 strains, showing lower mutation rates compared to WT in primary screen and 2 strains having higher mutation rates compared to WT were selected.

Table 3.2.1.1: description of the strains checked for growth.				
Gene name	**Systematic Name**	**Function summary**	**Biological process**	
EXO1	YOR033C	5'-3' exonuclease and flap-endonuclease	cellular response to DNA damage stimulus	Showed increased mutation rate compared to WT+hMLH1in primary screen (used as a control)
RRM3	YHR031C	DNA helicase involved in rDNA replication	cellular response to DNA damage stimulus	
MRPL9	YGR220C	Mitochondrial ribosomal protein of the large subunit	translation	Showed decreased mutation rate compared to WT+ *hMLH1* in primary screen
BUB3	YOR026W	Kinetochore checkpoint WD40 repeat protein	negative regulation of cellular component organization	
SRO9	YCL037C	Cytoplasmic RNA-binding Protein	translation	
DOT6	YER088C	Protein involved in rRNA and ribosome biogenesis	translation	
SPT4	YGR063C	Spt4p/5p (DSIF) transcription elongation factor complex Subunit	regulation of transcription	
COG5	YNL051W	Component of the conserved oligomeric Golgi complex	protein localization	
ALD4	YOR374W	Mitochondrial aldehyde dehydrogenase	organic acid biosynthetic process	
CDC50	YCR094W	Endosomal protein that interacts with phospholipid flippase Drs2p	endocytosis	
SEM1	YDR363W-A	19S proteasome regulatory particle lid subcomplex Component	vesicle-mediated transport	
PNC1	YGL037C	Nicotinamidase that converts nicotinamide to nicotinic acid	regulation of transcription	
OSW2	YLR054C	Protein of unknown function reputedly involved in spore wall Assembly	ascospore wall assembly	
PHO88	YBR106W	Component of alternate ER targeting pathway and may have potential role in SRP-independent targeting of substrates to the ER	protein maturation	
YLR434C	YLR434C	Dubious open reading frame		
YEL043W	YEL043W	Predicted cytoskeleton protein involved in intracellular signalling		

Figure 3.9: Growth curves of different gene deletions strains
Growth curves of different gene deletions strains from Table 3.2.1.1, were plotted with either vector only or h*MLH1* in leucine dropout medium along with control strains which include the S288C WT and *mlh1*Δ strain with either vector only or *hMLH1*. In the primary screen, Strains *exo1*Δ and *rrm3*Δ showed increased mutation rate and the remaining strains showed lower mutation rate.


Controls
exo1Δ
rrm3Δ
ald4Δ
mrpl9Δ
osw2Δ
sem1Δ
spt4Δ
sro9Δ
cdc50Δ
cog5Δ
dot6Δ
pho88Δ
pnc1Δ
bub3Δ
ydr417cΔ
yel043wΔ
ylr434Δ
OD600
0.0
0.1
0.2
0.3
0.4
0 1 2 3 4 5 6 7 8 9 10 11 12
Time(h)
WT + vector only
mlh1Δ + vector only
WT strain2 + vector only
mlh1Δ strain2 + vector only
gene of interestΔ + vector only
WT + hMLH1
mlh1Δ + hMLH1
WT strain2 + hMLH1
mlh1Δ strain2 + hMLH1
gene of interestΔ + hMLH1

The growth curves suggest that even though the strains *bub3*Δ, *spt4*Δ, *sro9*Δ, *cdc50*Δ, *ydr417c*Δ do have a slower growth compared to WT, it is still comparable to *mlh1*Δ*;* suggesting that the growth rate of these deletion strains is not the reason for their lower mutation rate. However, as the number of samples used to plot the curves was only 17; a systematic literature search for slow growth for all of the 523 genes was carried out. The Saccharomyces Genome Database (SGD) (http: //www.yeastgenome.org/) was searched for slow growth phenotype which revealed 2560 genes, of which 329 genes were reported to have lower mutation rate in our screen. In the screen cells were selected on leucine dropout media before transferring to canavanine, SGD was again searched for slower growth on leucine dropout media in particular. As this search did not lead to any result, strains were manually checked for slower growth on leucine dropout media in the images of primary screen and 22 gene deletions were considered to have slower growth on leucine dropout media. When checked for overlap, only 11 genes from this list were present in main SGD slow growth list (Figure 3.10).

Even though the number of genes in the main SGD slow growth genes were very high, their growth in leucine dropout media might differ; for example three genes *MRPL9*, *COG5*, *SEM1* were identified as slow growing, however when checked by plotting growth curve (Figure 3.8), did not have growth defect in leucine dropout media when compared to WT (with or without *hMLH1* expression). Thus it is possible that the observed lower mutation rate in these strains is independent of their mentioned growth defect in SGD.

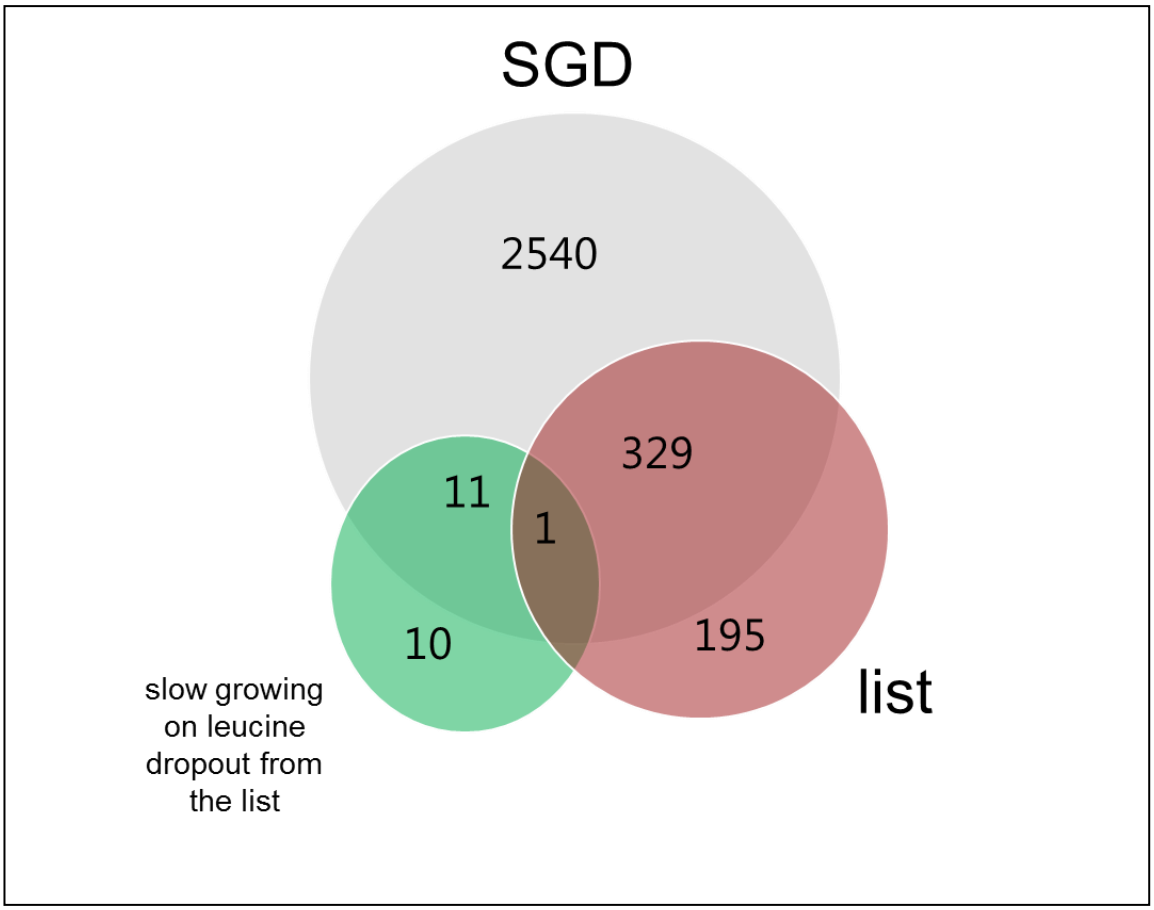


Figure 3.10: Analysis of slow growth genes from the list.

Genes from the list were classified according to the growth defect mentioned in SGD or manually checked using the pictures.

3.8 Effect of complete deletion of yeast MMR on spontaneous mutations.

In the genome wide mutagenesis screen, we identified 543 deletion strains where the mutation rates were reduced compared to wild type, when *hMLH1* was expressed. This observation was surprising both in terms of the number of affected strains as well as the general assumption that mutation rates only increase when MMR is defective. Several reasons could explain the result. For example, if hMLH1 protein is expressed variably, it would cause variation in the interference with yeast MMR and potentially restore wild-type mutation rates. To validate if these deletion strains have lower mutation rates due to their intrinsic function, as opposed to artefacts induced by variably regulation of *hMLH1*, I generated three libraries in which I deleted either *MLH1*, *MSH2* and *PMS1.*

I chose 348 of the 543 deletion strains and generated double mutants by crossing diploids and then sporulating to get haploid mutants (method 2.3.12). The *CAN1* mutation rate assay is based on mutational inactivation of arginine transport to the cells. One trivial explanation for the lowered growth on canavanine medium is therefore that the cells cannot synthesize arginine. Therefore before transferring cells to canavanine plate, cells were first checked for growth on arginine drop-out media, however in our screen we did not identify any slow growing arginine pathway defective cells, but they were still excluded from our further analysis.

Figure 3.11 a, is an example of one of the tertiary library plates on arginine media after crossing with *mlh1*Δ strain (Materials and method 2.3.12). In each plate a spot represents one ORF deletion. The spontaneous mutation rate was compared with *mlh1*Δ also with 10 times diluted *mlh1*Δ. The figure 3.11B shows that the spots of double mutants of gene of interest and *MLH1* show less canavanine resistant colonies as compared to *mlh1*Δ. . In total, of the 348 strains, 341 showed less canavanine resistant colonies. We infer this is due to lowered mutation rates.

Overall, the *CAN1* forward mutation rate assay of haploids having both the gene of interest and MMR defect revealed that the result from preliminary screen was reproducible and we estimated a false discovery rate of only 2%. Thus, of the 543 genes, 530 are expected to lower mutation rates with mismatch repair deficient cells.

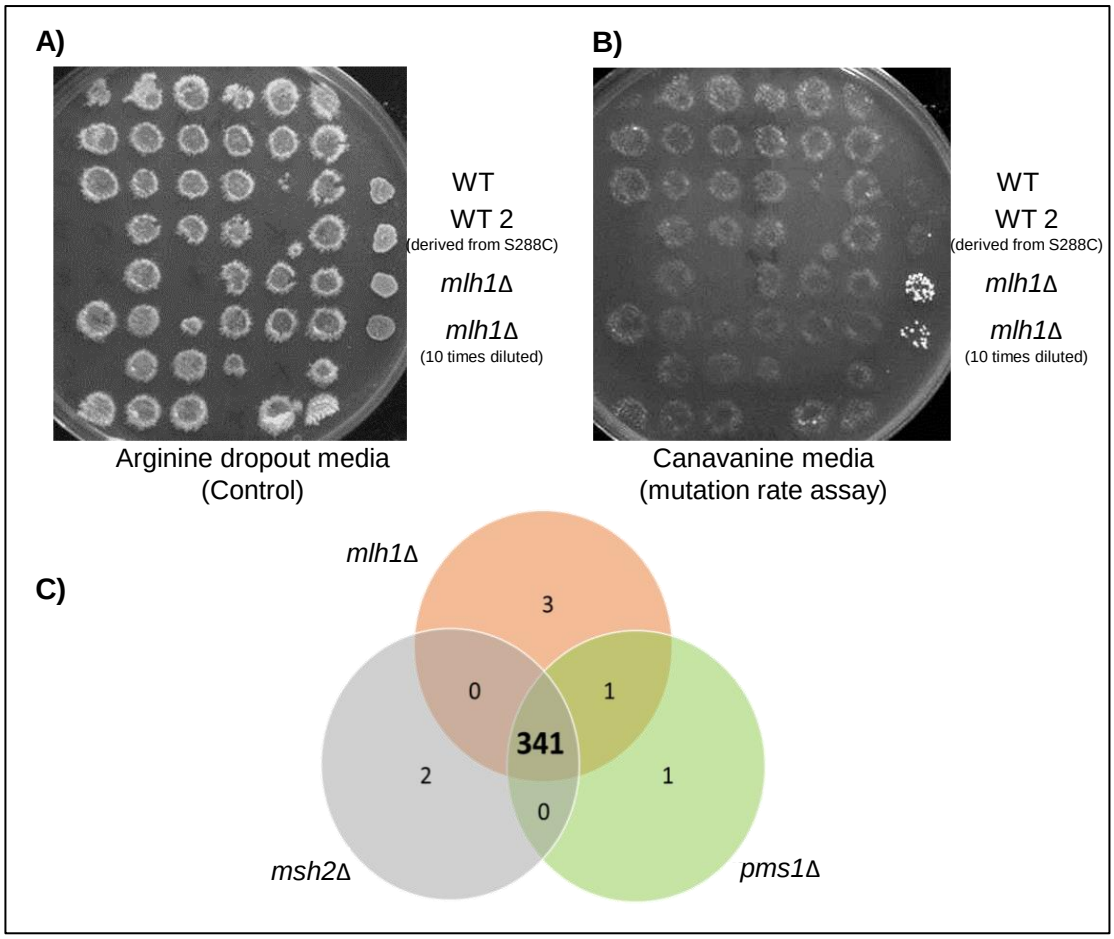


Figure 3.11 Complete deletion of yeast *mlh1* showed lower mutation rates than wild type. A tertiary library of genes which showed lower mutation rate as compared to WT in previous screen was created and crossed with mlh1Δ strain. Haploids were selected for deletion of gene of interest together with *mlh1*, growth of these haploids on Arginine drop-out media was checked first to confirm for no defects in arginine metabolism pathway and then transferred to canavanine plate to assay mutation rates. When compared with the *mlh1*Δ the tertiary library genes showed a decrease in mutation rate, consistent to the results found in preliminary screen.

A) Representative image of 40 deletion strains on Canavanine containing media, to assay mutation rate. Each spot represent a gene deletion.

B) Representative image of 40 deletion strains on Arginine dropout media, to check growth before transferring to canavanine containing plate. Each spot represent a gene deletion.

C) Complete deletion of other genes involved in yeast MMR namely *MSH2* and *PMS1* also showed reproducible data in 341 strains suggesting that the observed lower mutation rate is due to MMR attenuation (exp ID: E38_as01)

3.9 Discussion

MMR is associated with replication forks and is essential for repairing replication errors (Hombauer, Campbell, et al., 2011; Hombauer, Srivatsan, et al., 2011). The screen revealed that approximately 3% the deletion library strains showed an increase in mutation rates and ~11% of the strains showed a decrease in mutation rate compared to the controls with compromised MMR activity. As the screen was assessed qualitatively, the false discovery rate (FDR) for genes that increased mutation rate was calculated using fluctuation tests for 17 deletion strains. The quantitative analysis of these deletion strains showed a FDR of 12%. However, for the genes that showed a decrease in mutation rate, FDR was calculated by a secondary screen of double mutants of the gene-of-interest in combination with one of the key MMR genes (*MLH1*, *MSH2*, or *PMS1*). The secondary screen revealed a FDR of only 2%.

The primary screen was carried of using a deletion library of 4700 non-essential deletion strains; therefore it could not assess the impact of essential genes on mutation rate in absence of MMR. Another limitation of this screen was that the exact mechanism of how *hMLH1* inhibits the endogenous MMR machinery in yeast is unknown. However as the screen was conducted manually expression of *hMLH1* allowed us to increase throughput as compared to compromising the endogenous MMR genes directly. In addition, it also represents a biologically relevant situation and would lead us to find novel interactions that would compensate for lower MMR activity, as opposed to a complete deletion. Furthermore the screen rationale was also quantitatively confirmed by the observation, that expression of *hMLH1* in of *exo1*Δ strains, increases mutation

rate and leads to complete loss of MMR as observed in *mlh1*Δ (section3.2). As the *CAN1* assay is an indirect assay to assess the mutations caused by MMR defect, more direct approaches involving green fluorescent protein (EGFP) gene to quantitatively measure MMR activity in cells could allow us study accurate and direct involvement of MMR in shaping mutagenesis landscape.

Deletions identified in the screen that showed an increased mutation rate were found to be enriched in processes associated with DNA repair, DNA replication and response to DNA damage. This suggests that these processes may function in series in a common pathway or they can function in different pathways but have a common substrate (*e.g.*, *trm2*Δ, *rnh201*Δ, *rrm3*Δ, *rtt107*Δ and *tel1*Δ). These will be investigated further in chapter 5. However, GO enrichment amongst deletions resulting in a decreased mutation rate were mainly associated with mitochondria or general cell metabolism and growth. This suggests that these genes may contribute to an increase in the basal spontaneous mutation rate, thus when they are absent the mutation rate is lower and does not require MMR. A separate mechanism, through which genes associated with metabolism could regulate mutagenesis, is by influencing the redox state of a cell leading to an indirect accumulation of mutations.

The screen provided us an opportunity to understand the involvement of MMR genes in regulating endogenous mutation rates in yeast. Along with confirming previously known interactions, the screen provided new findings including the observation that mutation rates do not always increase in the absence of mismatch repair. This finding is important, as it brings into question the

accuracy of utilizing microsatellite instability (MSI) as a diagnosis criterion for early detection of Lynch Syndrome.

4 Chapter 4 - Identification of new interactors with MMR

4.1 Introduction

The mismatch repair pathway is important for maintaining replication fidelity and repairing chemically modified bases or mis-paired bases formed during replication (Jiricny, 2013). The MMR pathway is highly conserved from prokaryotes to higher eukaryotes and has been reconstituted using purified proteins together with mismatch containing heteroduplexes as substrates (Y. Zhang et al., 2005). As opposed to the four known exonucleases in *E.Coli*, only one exonuclease (Exo1) carries out excision of DNA strands containing a mis-pair in higher eukaryotes; however, surprisingly deletion of *EXO1* does not completely abolish MMR activity as seen with the loss of other MMR proteins (Kadyrov et al., 2009). This mild phenotype, together with the observation that approximately 20% to 30% of Lynch syndrome families with MSI do not have any mutations in known MMR genes, suggest that there are potentially more unidentified regulators of MMR and potentially Lynch syndrome (Rustgi, 2007).

Various attempts have been made to identify novel eukaryotic nucleases that complement the loss of Exo1, however, they have not generated any conclusive results (Amin et al., 2001) most likely due to the presence of many redundant exonucleases. For instance, in *E. coli*, abolishing two of the four exonucleases does not yield a strong mutator phenotype (Harris et al., 1998; Kolodner, 1996; Modrich & Lahue, 1996; Viswanathan & Lovett, 1998), therefore, a similar redundancy system may be in place in eukaryotes. This has been shown in

EXO1-deleted MEF cells where inactivation of exonucleases *MRE11*, *ARTEMIS* or *FAN1* showed either a weak change in MMR activity or no change at all, whereas a combined loss of all three genes showed a complete loss of mismatch repair (Desai & Gerson, 2014). While these three proteins have been identified as redundant for Exo1, it is possible proteins other than exonucleases are required to compensate for the loss of MMR.

In the genome wide screen, 163 genes showed an increased mutation rate when MMR was compromised, suggesting they may be involved in the MMR pathway or compensate for reduced MMR activity. In this chapter I will aim to determine if these genes, originally identified as enriched in biological processes like DNA repair and DNA replication, have additional roles in MMR.

4.2 *RNH201*, *TRM2*, *RRM3*, *RTT107* and *TEL1* may compensate for reduced MMR

The genome wide screen identified 163 gene deletions that increased mutation rate in combination with reduced MMR efficiency (hypomorphic MMR). The hypomorphic MMR was created by expressing *hMLH1*, which interferes with the yeast MMR machinery through the formation of non-functional protein complexes, but does not completely abolish the endogenous MMR (Shimodaira et al., 1998). The hypomorphic MMR system provided an advantage over complete loss of the pathway and helped to identify interactions that could potentially compensate for reduced MMR activity without being lethal. Amongst the 163 genes that showed an increase in mutation rate when MMR was compromised, a majority were associated with DNA replication and repair. To verify the link between MMR and these processes, the *hMLH1* experiment was repeated for the 17 genes involved in DNA repair and replication process from the screen: *RNR4*, *OGG1*, *ELG1*, *MGS1*, *MMS2*, *PSY3*, *RAD27*, *RAD34*, *RAD5*, *RNH201*, *TRM2*, *RRM3*, *RTT107*, *SHU1*, *SLX4*, and *TEL1* (Fig 3.5; conducted by master student Barbara Domanaska).

The *CAN1* forward mutation rate of *rrm3*Δ, *rnh201*Δ, *trm2*Δ, *rtt107*Δ and *tel1*Δ strains transformed with either *hMLH1* or empty vector showed an increase in mutation rate in the presence of *hMLH1* (Figure 4.1 A and B). When calculating the potential additive effect of the gene deletion with the mutation rate of WT+*hMLH1* all five genes showed a synergistic effect, however, expression of *hMLH1* did not further increase the mutation rate in *mlh1*Δ (Figure 5.1C).

Similarly expression of *hMLH1* in these five deletion strains (*rrm3*Δ, *rnh201*Δ, *trm2*Δ, *rtt107*Δ and *tel1*Δ), increased mutation rate higher than the potential additive effect, however, they did not surpass the mutation rate of the *mlh1*Δ strain (Figure 4.1C). This suggests that these genes might have a role in compensating for reduced MMR activity, thus deletion of them in combination with lower MMR efficiency leads to complete loss of MMR but does not increase the mutation rate any further.

Figure 4.1: Mutations in replication or repair associated deletion strains increases synergistically when *hMLH1* is expressed

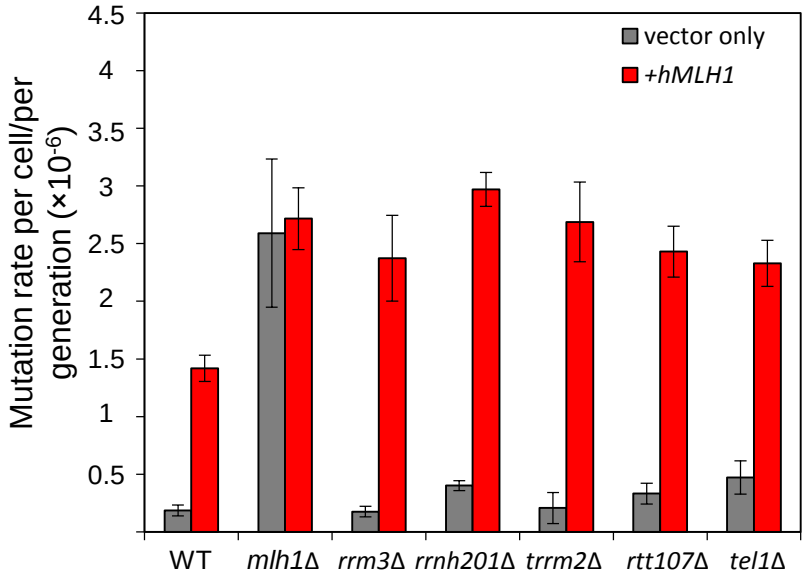
A) Average mutation rates per cell for the deletion strains associated with replication or repair, using *CAN1* forward mutation rate assay. Mutation rates were determined using method of median, theree independent experiments were performed each time seven colonies per gene deletion were used. Grey bars represent deletion strains with vector only. Orange bars represent deletion strains i.e. gene of interest with *hMLH1*.

B) Fold differences relative to WT with vector only. Grey bars represent deletion strains with vector only. Orange bars represent deletion strains i.e. gene of interest with hMLH1

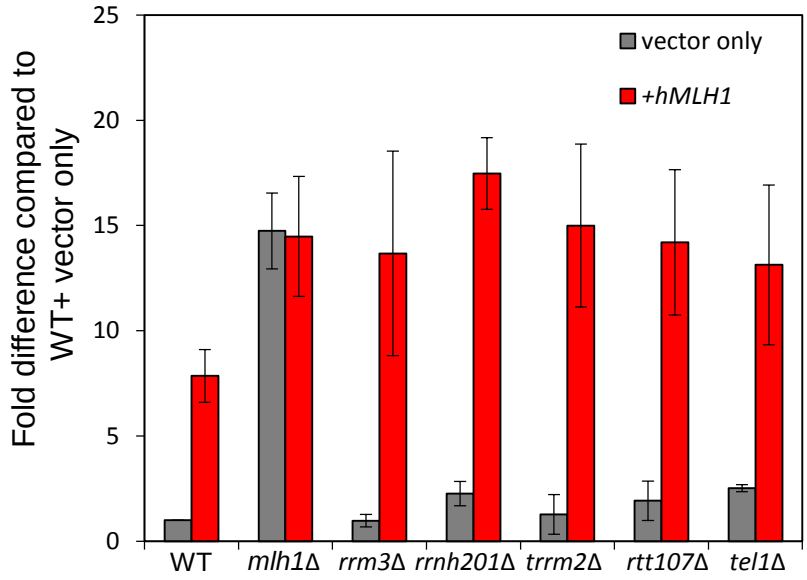
C) Potential additive effect for deletion strains and WT with the black bar represent potential additive effect calculated mannualy by adding mutation rate of gene of interest deletion with vector only and WT with *hMLH1.* Grey bars represent deletion strains with vector only. Orange bars represent deletion strains i.e. gene of interest with *hMLH1*

Data is generated by Barbara Domanska (exp ID:E71_as01).

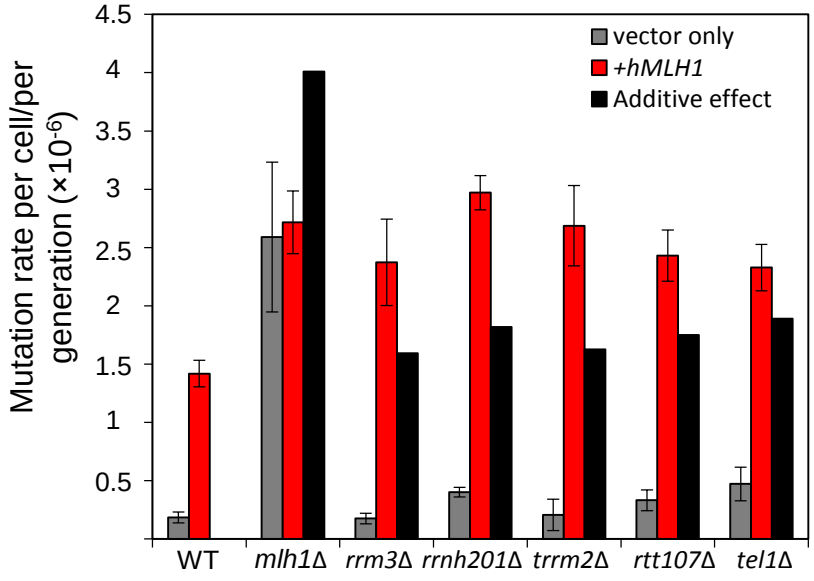
A)


Mutation rate per cell/per generation (×10-6)
4.5
4
3.5
3
2.5
2
1.5
1
0.5
vector only
+hMLH1
WT
mlh1Δ
rrm3Δ
rrnh201Δ
trrm2Δ
rtt107Δ
tel1Δ

B)


Fold difference compared to WT+ vector only
25
20
15
10
5
0
vector only
+hMLH1
WT
mlh1Δ
rrm3Δ
rrnh201Δ
trrm2Δ
rtt107Δ
tel1Δ

C)


Mutation rate per cell/per generation (×10-6)
4.5
4
3.5
3
2.5
2
1.5
1
0.5
vector only
+hMLH1
Additive effect
WT
mlh1Δ
rrm3Δ
rrnh201Δ
trrm2Δ
rtt107Δ
tel1Δ

4.3 Known interaction of *RNH201*, *RRM3*, *TRM2*, *TEL1* and *RTT107* with MMR

In order to further evaluate role of *RNH201*, *RRM3*, *TRM2*, *TEL1* and *RTT107* in MMR, previous known interactions were studied and are summarized below.

RNH201 encodes for the ribonuclease H2 catalytic subunit. It is required for removal of RNA primers during Okazaki fragment synthesis (Frank et al., 1998; Qiu et al., 1999) and has been previously shown to interact with MMR proteins like Exo1, Mlh1, Mlh3, Msh2, Msh3 and Msh6 (Clark et al., 2011; Ghodgaonkar et al., 2013; Shen et al., 2011; X. A. Su & Freudenreich, 2017). *RNase H2* is also required to remove incorrectly added ribonucleotides during replication and to initiate the ribonucleotide excision repair (Sparks et al., 2012), which creates a nick that can provide a strand discrimination signal during MMR (Bruni et al., 1988; Burdett et al., 2001; Mechanic et al., 2000). Inactivation of the human homolog, *RNASEH2A* leads to the neurological disorder called Aicardi-Goutières syndrome [AGS] (Crow et al., 2006).

RRM3 encodes for a 5'-3' DNA helicase required for stability and replication of rDNA (Keil & McWilliams, 1993). It belongs to the PIF1 family of helicases, but effects fork progression in the opposite direction as the 3'-5' PIF helicase (Ivessa et al., 2002; Ivessa et al., 2000); *RRM3* is also required for Ty1 transposition and in relieving replication fork pauses in telomeric regions ((Makovets, 2004; Scholes, 2001). Loss of Rrm3p shows an increase in rDNA breakage and accumulation of rDNA circles (Ivessa et al., 2000). Even though a

direct role of *RRM3* in MMR has not been demonstrated, studies using DNA polymerase mutant strains have associated RRM3 with maintaining replication fidelity (T. T. Schmidt et al., 2017). Synthetic genetic array studies have shown that *RRM3* has a synergistic effect with the deletion of MMR genes *MSH2* and *MSH6* (Collins et al., 2007; Kuzmin et al., 2018). Finally, with the help of two-hybrid analysis RRM3 has been show to directly interact with PCNA in vitro and in vivo (K. H. Schmidt et al., 2002).

TRM2 encodes a tRNA methyltransferase and is involved in tRNA stabilization and maturation (Johansson & Bystrom, 2002). Trm2 is also an endo-exonuclease with single strand endonuclease activity and 5′ to 3′ exonuclease activity (Choudhury et al., 2007). *TRM2* has no known interaction with MMR proteins except with *EXO1*. A cell survival assay after methyl methane sulfonate (MMS) treatment showed a synergistic interaction between *TRM2* and *EXO1* which suggest that, these two gene products might have an overlapping role in their repair (Choudhury et al., 2007).

TEL1 encodes for a protein kinase mainly involved in telomere length regulation ((Lustig, 1986; Mallory, 2000). *TEL1* is a member of the PIK- related kinase family and is associated with regulation of the cell cycle checkpoint in response to DNA damage. *TEL1* is known to have interactions with Exo1 to regulate checkpoint activation in response to double strand breaks (Clerici et al., 2014). Finally, double mutants of *tel1*Δ and *msh2*Δ have a higher rate of gross chromosomal rearrangements than the single mutants (K. H. Schmidt et al., 2006).

RTT107 encodes for a protein necessary for reinitiating replication after repair of alkylating DNA damage (Hanway et al., 2002) . RTT107p is known to interact with Mms22p and Slx4p (Baldwin et al., 2005; Roberts et al., 2006). Deletion of the *RTT107* gene causes hypersensitivity to DNA-damaging agents such as the DNA-alkylating agent MMS. However, to date, there have been no known interaction of *RTT107* with any of the MMR genes.

4.4 *RRM3* may compensate for reduced MMR activity.

The *S. cerevisiae* MutLα complex belongs to the GHL dimeric ATPase superfamily (Ivessa et al., 2002; Ivessa et al., 2000). Consistent with other members of the family of ATPases, like gyrase b and Hsp90, MutLα undergoes ATP-dependent dimerization of the N-terminal ATPase domains (Moarefi et al.). This conformational change may play a role in recruiting downstream effectors, as deactivating these ATPase domains in *MLH1* lead to a mutator phenotype. Previously, it has been shown that deletion of *EXO1* enhances the weak mutator phenotype of the MutLα ‘‘ATPase’’ mutation and leads to a complete loss of MMR (Tran et al., 2001). Importantly, this suggests that the ATP-hydrolysis motif may play a role in *EXO1* independent mismatch repair, possibly by recruiting other factors redundant to Exo1 activity in MMR. The genomic screen presented in this thesis was carried out in *S. cerevisiae* cells with reduced MMR activity, with one of the objectives being to identify novel interactors of MMR that are involved in such compensatory interactions. Therefore, to further identify if *RNH201*, *TRM2*, *RRM3*, or *TEL1* have any such interactions, these genes of interest were deleted in combination with a hydrolytic activity mutant strain of *MLH1*, mlh1E31A.

The *CAN1* forward mutation rate was calculated for deletion strains of RNH201, *TRM2*, *RRM3*, and *TEL1* in combination with the mlh1E31A deletion (Figure 4.2 A). When compared to WT, deletion of the hydrolytic activity of Mlh1 (mlh1E31A) did not lead to a complete loss of MMR as seen in *mlh1*Δ (Fig 4.2B blue bar). The mutation rate was high in *rnh201*Δ when combined with mlh1E31A, but was not significantly higher than the mlh1E31A mutant alone.

Similarly for *trm2*Δ there was no increase in the mutation rate when the combined with mlh1E31. This indicates that Rnh201 and Trm2 do not have a role in compensating for the hydrolytic activity of *Mlh1*. The observed decrease in mutation rate in the case of tel1Δ could be an experimental error and needs to be investigated further, as the colonies formed by *tel1*Δ in combination with mlh1E31A were small and difficult to count, even after five days of incubation, compared to other strains.

In the case of the *rrm3*Δ strain, loss of Mlh1 hydrolytic activity showed a synergistic increase in mutation rate. Even though it is not equivalent to a complete loss of MMR, it does indicate that *RRM3* might have a role in compensating for the hydrolytic activity of Mlh1. Further characterization is needed.

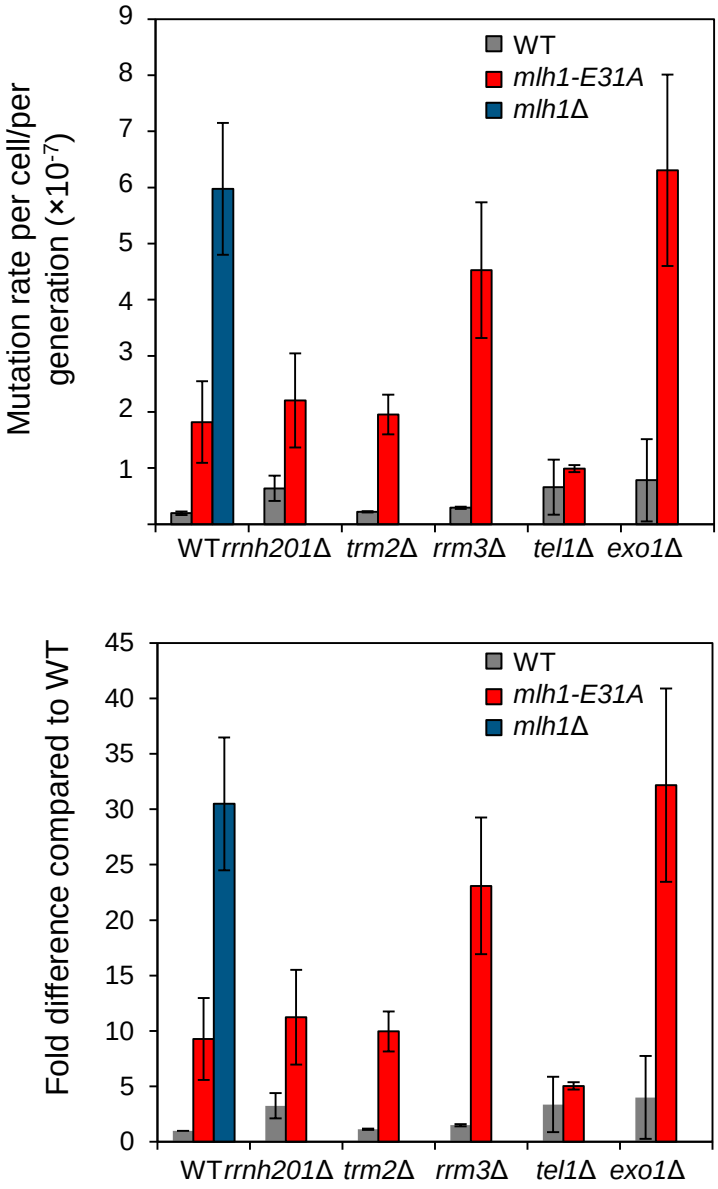


Figure 4.2: *RRM3* deletion strain shows synergistic increase in mutation rate when in hydrolytic activity mutant of *MLH1*

A) Average mutation rates per cell for the deletion strains associated with replication or repair, in combination with mlh1E31A (hydrolytic activity mutant of *MLH1*). *CAN1* forward mutation rates were determined using method of median, 3 independent experiments were performed each time 5 colonies per gene deletion were used. Grey bars represent deletion strains. Red bars represent deletion strains i.e. gene of interest in combination with mlh1E31A (hydrolytic activity mutant of *MLH1*). Purple bar represent *mlh1*Δ

B) Fold differences relative to WT Grey bars represent deletion strains. Orange bars represent deletion strains i.e. gene of interest in combination with mlh1E31A (hydrolytic activity mutant of *MLH1*). Purple bar represent *mlh1*Δ

4.5 Discussion

Disruption of MMR increases the spontaneous mutation rate. In deletion strains, *rnh201*Δ, *trm2*Δ, *rrm3*Δ, *tel1*Δ and *rtt107*Δ, the disruption of MMR showed a synergistic effect. This indicates that these genes may function in a common pathway or have a common substrate. Consistent with this idea, a double mutant of *RRM3* and *MSH6* showed a decreased colony size phenotype (Collins et al., 2007). Additionally, a *TRM2* and *EXO1* double mutant showed a decreased vegetative growth phenotype (Choudhury et al., 2007) indicating that *TRM2* and *EXO1* can complement each other functionally (Choudhury et al., 2007). These interactions may also imply that defects in the human homologues of these genes, together with MMR deficiency, may increase predisposition to carcinogenesis and may be potential targets for treatment of MMR deficient tumours.

To investigate their role in mismatch repair, we assessed the mutation rate of these stains in combination with the *MLH1* ATP hydrolytic mutant, and found only the *RRM3* deletion showed a synergistic increase in mutation rate. This suggests role of Rrm3 in compensating for low MMR activity. These findings are consistent with previous studies associating *RRM3* in maintaining replication fidelity. Even though, *trm2*Δ, *rnh201*Δ and *tel1*Δ did not increase the mutation rate any further in the Mlh1 ATP hydrolytic mutant, a possible role in regulating MMR cannot be ruled out and needs to be further assed using double mutants of other MMR genes or replicative polymerase mutants.

RRM3 is known to interact with replication machinery and is required for progression of replication fork through natural pausing site (Ivessa et al., 2003). One possible way in which Rrm3 is required for MMR is through its helicase activity; which could facilitate the recognition of mismatch in difficult to replicate regions. Further experiments involving tertiary structures as template are required to test this hypothesis; however as *CAN1* is not associated with any tertiary structures it is possible that Rrm3 may participate in MMR differently. As Rrm3 is able to interact with pol δ and PCNA it is possible that Rrm3 is required in Exo1 independent MMR. This hypothesis is of particular interest as the ATP hydrolytic activity of *MLH1* is required to create nicks by MutLα which is critical in Exo1 independent MMR. To test this hypothesis, further experiments will be conducted using double mutants of *EXO1* and *RRM3.*

To summarize, our data indicates a possibility for *RRM3*, *TRM2*, *RNH201* and *TEL1* to have a role in compensating for low MMR activity, however, the exact mechanism through which they act requires further analysis in yeast and mammalian cells.

5 Chapter 5 - The role of genes associated with regulation of transcription coupled repair in spontaneous mutagenesis

5.1 Introduction

Transcription efficiency has a direct impact on multiple cellular processes, such as DNA repair and mutagenesis. Over 20 years ago, it was demonstrated in yeast that an increased transcription level stimulates spontaneous mutagenesis (Datta & Jinks-Robertson, 1995). Recent genome-wide sequencing approaches in mammalian cells have similarly suggested that transcriptionally active regions are more susceptible to the accumulation of mutations than less-active regions (Barlow et al., 2013; Chiarle et al., 2011; Klein et al., 2011).

Transcription coupled repair (TCR), a sub-pathway of nucleotide excision repair (NER), repairs DNA lesions on the transcribed strand of active genes and is initiated when RNAPII is stalled due to strand distorting lesions (Mellon et al., 1987). In yeast, there are 16 genes associated with the regulation of TCR. The deletion library used for our mutation screen contained 12 of these genes, out of which a cluster of four genes showed lower mutation rates in the absence of MMR. These four genes were identified as mutator genes, i.e., in these deletion strains the mutation rate was lower than the rate seen in hypomorphic MMR cells alone (WT + *hMLH1*). The fact that this cluster represents approximately one third of all TCR regulatory genes in the yeast genome suggests a possible link between mutagenesis, and TCR and MMR pathways. The four genes from

this mutator cluster *DEF1*, *TFB5*, *THO2* and *SAC3* do not contribute in the TCR pathway directly, but have shown several indirect effects (Fig 5.1).

Removal of stalled RNA polymerase II (RNAPII) from damage sites, which is critical for repair and resumption of transcription, is triggered by ubiquitylation of the polymerase by a complex of Def1 and Rad26 (Woudstra et al., 2002). Even though Def1 is not required for TCR directly, *DEF1* mutants show inefficient transcript elongation, increased sensitivity to transcription elongation inhibitors such as 6AU, these mutants cannot degrade RNAPII in response to DNA damage, and disruption of abasic site repair on transcribed strands (Owiti et al., 2017). Additionally, *def1*Δ is synthetically sick when combined with mutations in *PIF1*, the DNA helicase that participates in DNA maintenance and replication at DNA breaks (Stundon & Zakian, 2015; Wilson et al., 2013). Further studies have identified roles of Def1 in cytokinesis, double strand break repair in both meiosis and mitosis, and mutagen sensitivity (Jordan et al., 2007).

The other three genes of interest *TFB5*, *THO2*, and *SAC3* all play more indirect roles in the TCR pathway. The mutator gene *TFB5* impacts the efficiency of the TCR pathway by affecting the stability and structure of transcription factor II H (THIIH), which plays an important role in TCR through interactions with both RNAPII and DNA repair (Ranish et al., 2004). In *Saccharomyces cerevisiae*, the TREX-2 (TRanscription-EXport) complex plays an important role throughout mRNA biogenesis—including transcription, processing, and transport (Fischer et al., 2002; Kohler & Hurt, 2007). The mutator protein Tho2 is an integral subunit of the THO/TREX-2 complex and plays a role in mRNA processing by

translocating with RNAPII during transcription. TREX-2 is able to then move mRNA out of the nucleus via the nuclear pore complex (NPC); this process is facilitated by a complex of Thp1 and the mutator protein from the screen Sac3. Sac3 was initially thought to be an actin suppressor (Novick et al., 1989). New studies have identified additional roles in nuclear protein and RNA export (Fischer et al., 2002; Jones et al., 2000; Lei et al., 2003). Additionally, transcription-associated hyper-recombination has been observed due to the inability of mRNA to be processed or to be exported out of the nucleus in *THO2* and *SAC3* mutants, respectively (Gallardo et al., 2003). Further, UV sensitivity studies in yeast have confirmed that *THO2* and *SAC3* deletion impacts the efficiency of TCR (Gaillard et al., 2007).

Collectively, our data indicates that genes associated with regulation of TCR could have an impact on increasing mutations that require MMR pathway dependant repair. In this chapter, I will be further validating the impact of these (*DEF1*, *TFB5*, *THO2* and *SAC3)* genes associated with regulation of TCR on mutation rate .

Figure 5.1: The importance of mutator genes in regulating TCR.
A) Tho2 is a part of THO complex associated with RNAP II and is required for efficient transcription. It is regulates TCR by efficient m-RNA processing.
Sac3 is required for regulating TCR by facilitating the transport of processed m-RNAs through NPC.

B) Stalling of RNAPII due to strand distorting lesions send the signal to initiate TCR.

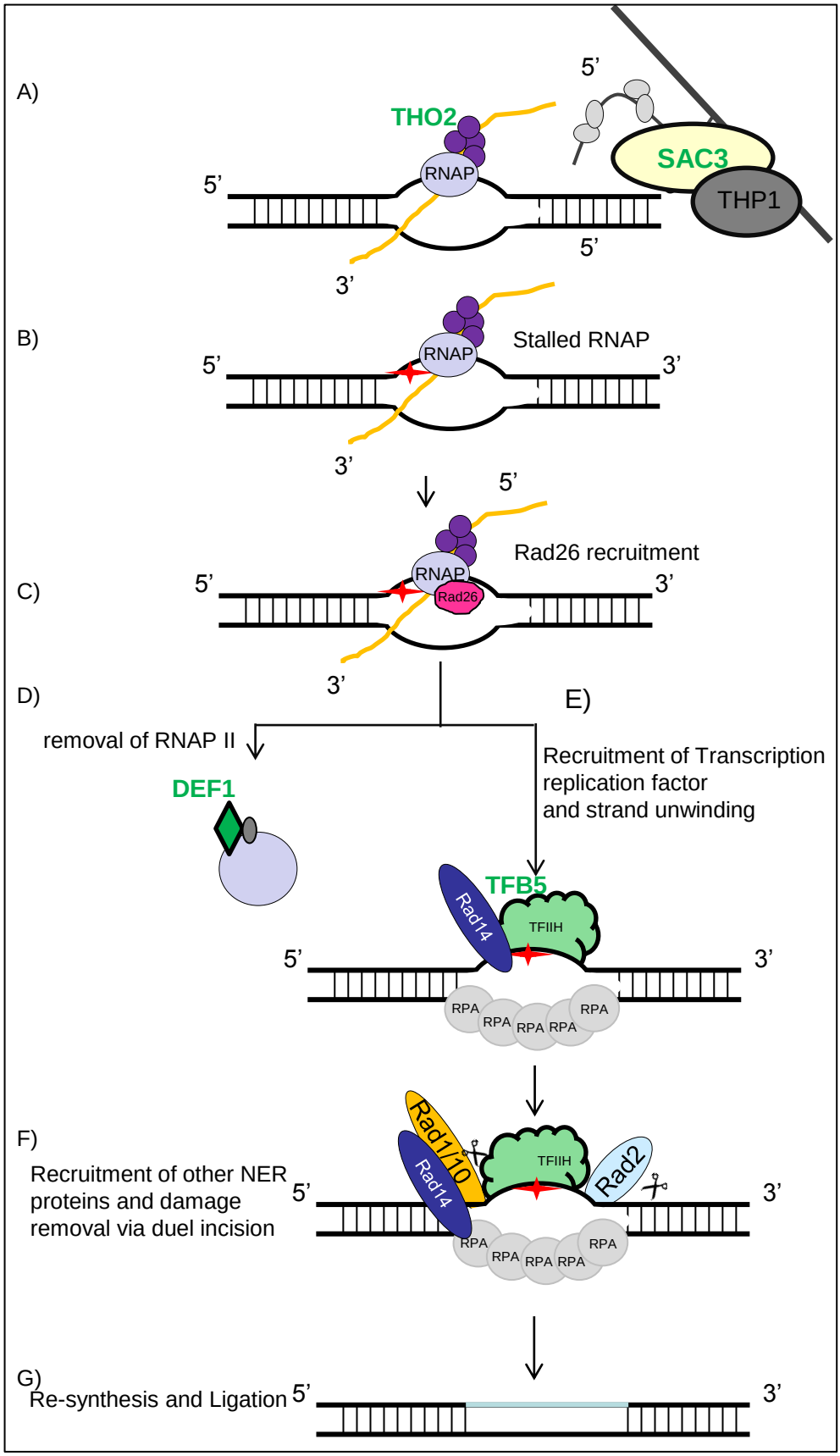
C) Rad26 is recruited to lesion site which is important for the downstream TCR.

D) Def1 is required for the removal of stalled RNAPII from the lesion site to allow access to repair machinery; it is carried out by Def1 and Rad26 complex.

E) Recruitment of transcription factor IIH (TFIIH). TFIIH is required for efficient unwinding of the strand and recruitment of NER machinery. Tfb5 is a subunit of (TFIIH), and it regulates efficiency of TCR by maintaining stability of the complex TFIIH.

G) Recruitment of the structure-specific endonucleases Rad1-Rad10 and Rad2 which carry out the lesion removal by dual incision

H) The strand is then re-synthesis and ligated.


A)
5'
THO2
SAC3
THP1
RNAP
5'
5'
3'
B)
Stalled RNAP
RNAP
5'
3'
3'
5'
C)
Rad26 recruitment
RNAP
Rad26
5'
3'
3'
D)
E)
removal of RNAP II
Recruitment of Transcription
replication factor
and strand unwinding
DEF1
TFB5
Rad14
TFIIH
5'
3'
RPA
RPA
RPA
RPA
RPA
F)
Recruitment of other NER
proteins and damage
removal via duel incision
Rad1/10
Rad14
TFIIH
Rad2
5'
3'
RPA
RPA
RPA
RPA
RPA
G)
Re-synthesis and Ligation
5'
3'

5.2 Mutation rates are lower in TCR deletion strains when compared to mutation rates in *mlh1*Δ.

To validate if genes associated with TCR have a lower mutation rate in the absence of functional MMR, we conducted a fluctuation test to determine the spontaneous mutation rates per cell using the method of median (Lea & Coulson, 1949; Luria & Delbruck, 1943). *def1*Δ, *sac3*Δ, *tfb5*Δ and *tho2*Δ strains with and without *MLH1* were grown until stationary phase and appropriate dilutions were plated on canavanine containing media and complete media. Canavanine resistant colonies were counted and spontaneous mutation rate was calculated (Method 2.3.10).

Through this study we found that deletion of these TCR associated genes did not increase the mutation rate when compared to WT (Fig 5.2A and B). In WT cells, deletion of *MLH1* leads to a 25-fold increase in mutation rate, whereas the mutation rates for *sac3*Δ *mlh1*Δ and *tfb5*Δ *mlh1*Δ were nearly 15-fold higher than WT (Fig 4.2A and B). Surprisingly, deletion of *DEF1* and *THO2* did not result in any canavanine resistant colonies with or without *MLH1* (Fig 5.2A and B), suggesting that there is no accumulation of mutations in *CAN1*. Cells were plated at five times higher concentration than WT and there was still no growth on canavanine media (data not shown).

The mutation rate was increased when *MLH1* was deleted in *tfb5*Δ and *sac3*Δ strains, but was not as high as *mlh1*Δ. However, for *def1*Δ and *tho2*Δ there were no *CAN1* resistant colonies observed when *MLH1* was present or absent.

This suggests that these gene deletions do not lead to accumulation of mutation in the *CAN1* gene and thereby do not lead to colony formation. Additionally, also they might have a role in causing mutations, which are then repaired in MMR dependent way.

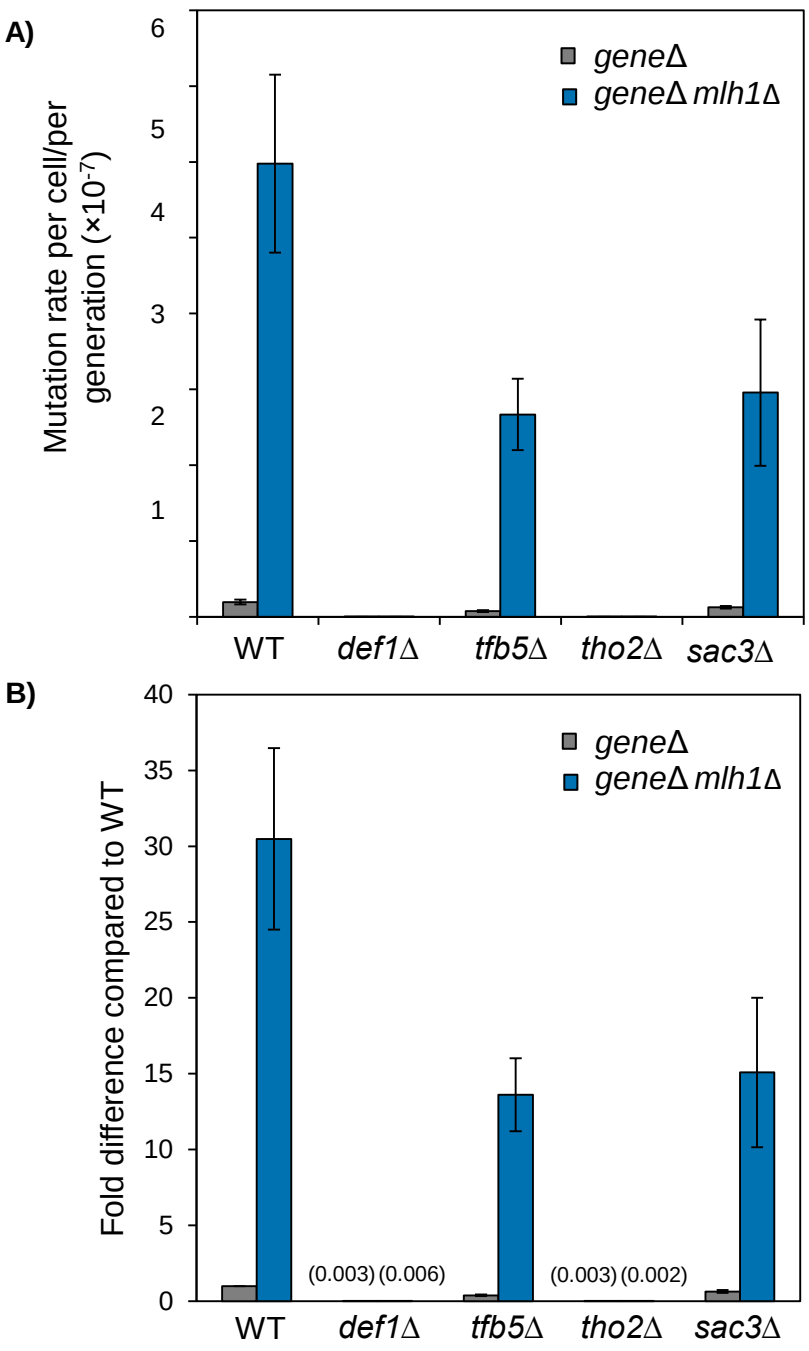


Figure 5.2: Mutation rate in TCR deletion strains is lower than mutation rate in *MLH1* deleted cells.

A) Average mutation rates per cell for TCR deletion was calculated using *CAN1* forward mutation rate. Mutation rates were determined using method of median, 3 independent experiments were performed each time 5 colonies per gene deletion were used. Grey bars represent gene of interest deletion. Blue bars represent double deletions i.e. gene of interest and *MLH1*.

B) Fold differences relative to WT. Grey bars represent gene of interest deletion. Blue bars represent double deletions with gene of interest and *MLH1; fold difference v*alues for *DEF1* and *THO2* deletion strains are mentioned in brackets (exp ID:E71_as01).

5.3 Validation of the lower mutation rate in TCR deletion strains.

5.3.1 Deletion of genes associated with regulation of TCR does not affect growth on canavanine media.

One simple reason for the lack of any canavanine resistant colonies for *tho2*Δ and *def1*Δ could be that these gene deletions simply affect their growth on canavanine containing media. It was therefore important to ensure that cells lacking these genes associated with regulation of TCR can survive in presence of canavanine. To test survival of these deletion strains on canavanine, we deleted the gene of interest along with *CAN1* to disrupt the arginine permease.

Growth of four independent double mutants of *can1*Δ and one of the regulation of TCR cluster gene, i.e. *def1*Δ, *sac3*Δ, *tfb5*Δ and *tho2*Δ, were assessed after 24hrs on complete media and canavanine containing media plates. The regulation of TCR cluster gene single deletion strains and WT strain did not grow on canavanine containing media, whereas the double mutants showed growth patterns similar to *can1*Δ (Fig 5.3.1).

Growth of these double mutants suggest that the observed lower mutation rate in TCR associated mutator deletion strains in the absence of *MLH1* (Fig 4.3A and B) is independent of growth on canavanine media.

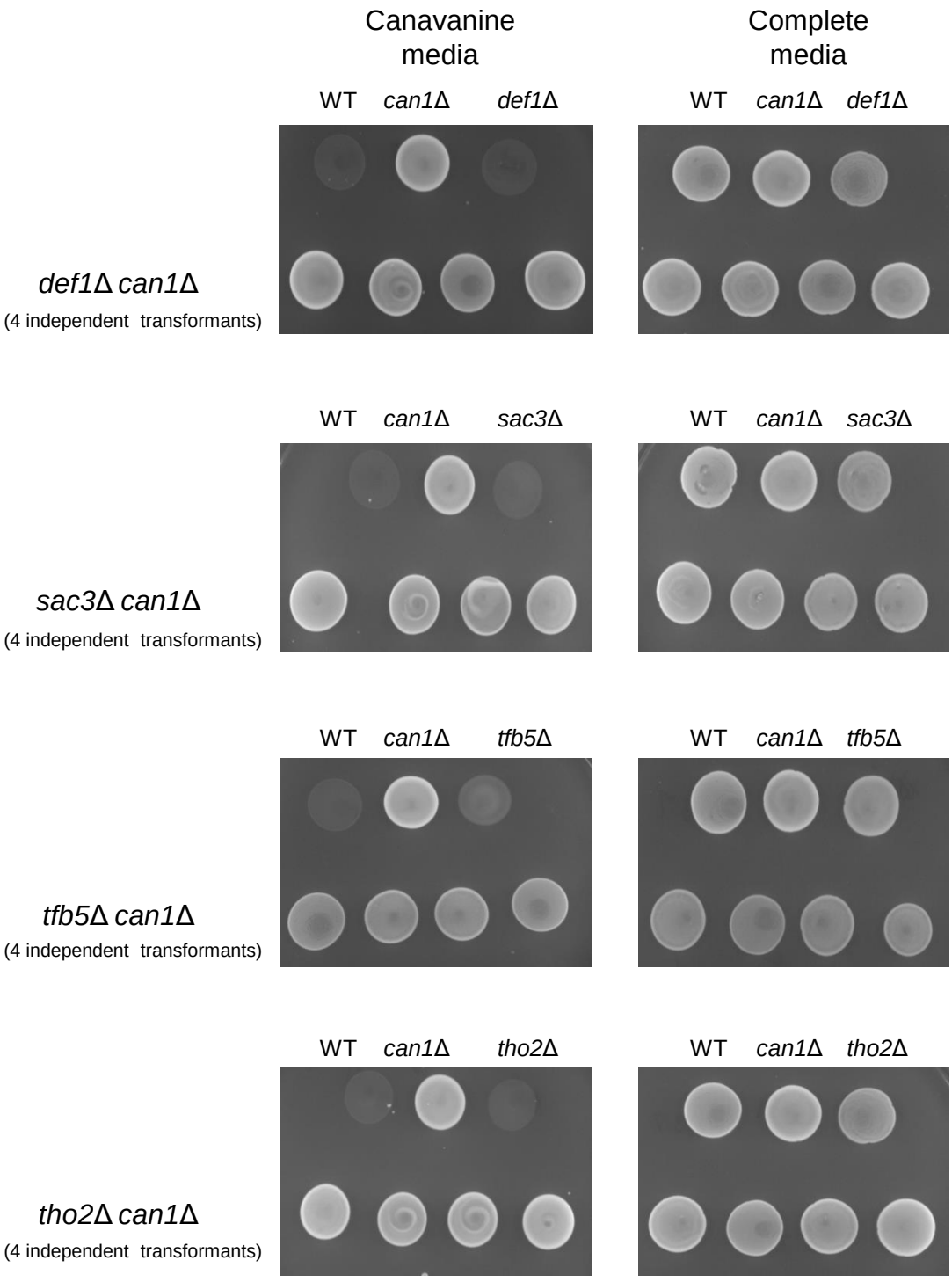


Figure 5.3.1: growth on canavanine media is not affected by deletion of TCR cluster genes.

Growth on canavanine media, when gene of interest is deleted with *CAN1*, each plate has 4 independent transformants for double mutants. WT and single deletion of gene of interest is shown as a negative control and can1 is shown as a positive control (exp ID:E72_as01) .

5.3.2 Lower mutation rate in deletion strains associated with regulation of TCR cluster is not explained by the slow growth phenotype.

Slower growth in the absence of *MLH1* may also influence mutation accumulation; therefore, we plotted growth curves for the regulation of TCR cluster deletion strains with and without *MLH1*. Deletion strains associated with TCR cluster regulation combined with *MLH1* deletion were grown in complete media for 13 hours and OD_{600} was measured every hour. Cells were checked under microscope every four hours to ensure that cell morphology did not change.

When compared to WT and *mlh1*Δ controls, the *tho2*Δ strain did not show a significant growth defect (Fig 5.3.2 A); however *def1*Δ did grow slowly with and without *MLH1* (Fig 5.3.2 A). Strains *sac3*Δ and *tfb5*Δ on their own did have slower growth phenotype compared to WT and *mlh1*Δ, but surprisingly in a double mutant with *mlh1*Δ their growth was similar to both controls (WT and *mlh1*Δ alone). Further investigation is needed to determine why the deletion of *MLH1* rescues the growth defect in these double mutants (Fig 5.3.2 C and D).

As the calculation of mutation rate (Fig 5.2) takes growth in to account, the presence of slow growth phenotypes does not affect our identified mutation rates. The growth curves were primarily used to verify that the cells are healthy and was confirmed in Fig 5.3.2, which shows that deletion of the genes

associated with regulation of TCR cluster is not lethal. This also suggests that the observed lower mutation rate is not due to accelerated cell death.

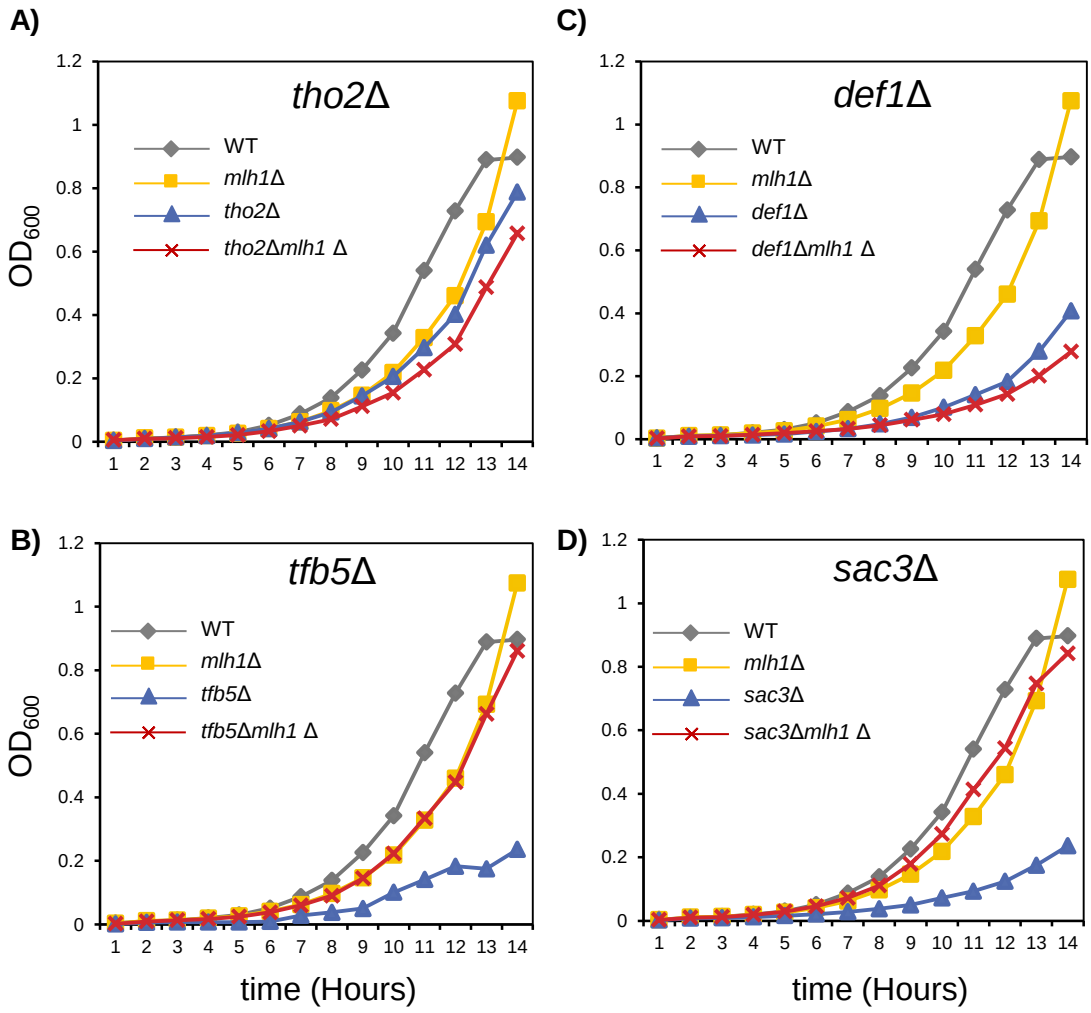


Figure 5.3.2 There is a difference in growth rate for wild type and strains having TCR cluster genes deleted.

Cells were grown in YPD medium for 12 hours and OD_{600} after every hour was measured. Grey line represents WT, yellow like represents *mlh1*Δ, blue line represents TCR cluster gene single deletion and red line represents TCR cluster gene deletion with *mlh1*Δ (exp ID:E76_as01),

A) *sac3*Δ; B) *tfb5*Δ ;C) *def1*Δ; D) *tho2*Δ

5.4 Deletion of genes associated with regulation of TCR in *mlh1*∆ cells leads to lower accumulation of frameshift mutations in repetitive DNA sequences.

Mutations in MMR genes characteristically lead to an increased accumulation of frameshift mutations in repetitive DNA sequences. *CAN1* in cells accumulate all type of mutations, mainly single base pair substitutions, therefore only 1% of mutations identified through the *CAN1* assay are associated with MMR (Fig 5.4 A) (Lang & Murray, 2008). To assess mutations caused by MMR loss, a more specific assay was used that can identify frame shift mutations by calculating the reversion rate in homonucleotide runs using a previously reported lys2::InsE-A14 construct (Tran et al., 1997). To assess mutation rate, fluctuation tests were performed by scoring colonies able to grow on lysine dropout media due to frame shift mutations.

The *CAN1* assay, which identifies multiple types of mutations repaired through several different pathways, was used to determine the overall mutation rate (Fig5.4 B), whereas the lys2::InsE-A14 construct was used to specifically assess frameshift mutations repaired by MMR (Fig5.4 D). As expected we observed that the mutation rate per cell using the lys2::InsE-A14 assay in all the strains was higher when compared to the *CAN1* assay; therefore, we focused mainly on the lys2::InsE-A14 assay to identify MMR specific mutations in *tfb5*Δ and *sac3*Δ strains (Fig 5.4 B). The reversion rate for *tfb5*Δ and *sac3*Δ calculated using the lys2::InsE-A14 construct was similar to WT. Deletion of *MLH1* increased the reversion rate in *tfb5*Δ and *sac3*Δ, but it was still lower than

mlh1Δ alone (Fig 5.4 C). This data, in combination with the lower overall mutation rate identified via the *CAN1* assay (Fig 5.2), suggest that *tfb5*Δ and *sac3*Δ strains accumulate fewer mutations independent of the type of mutation.

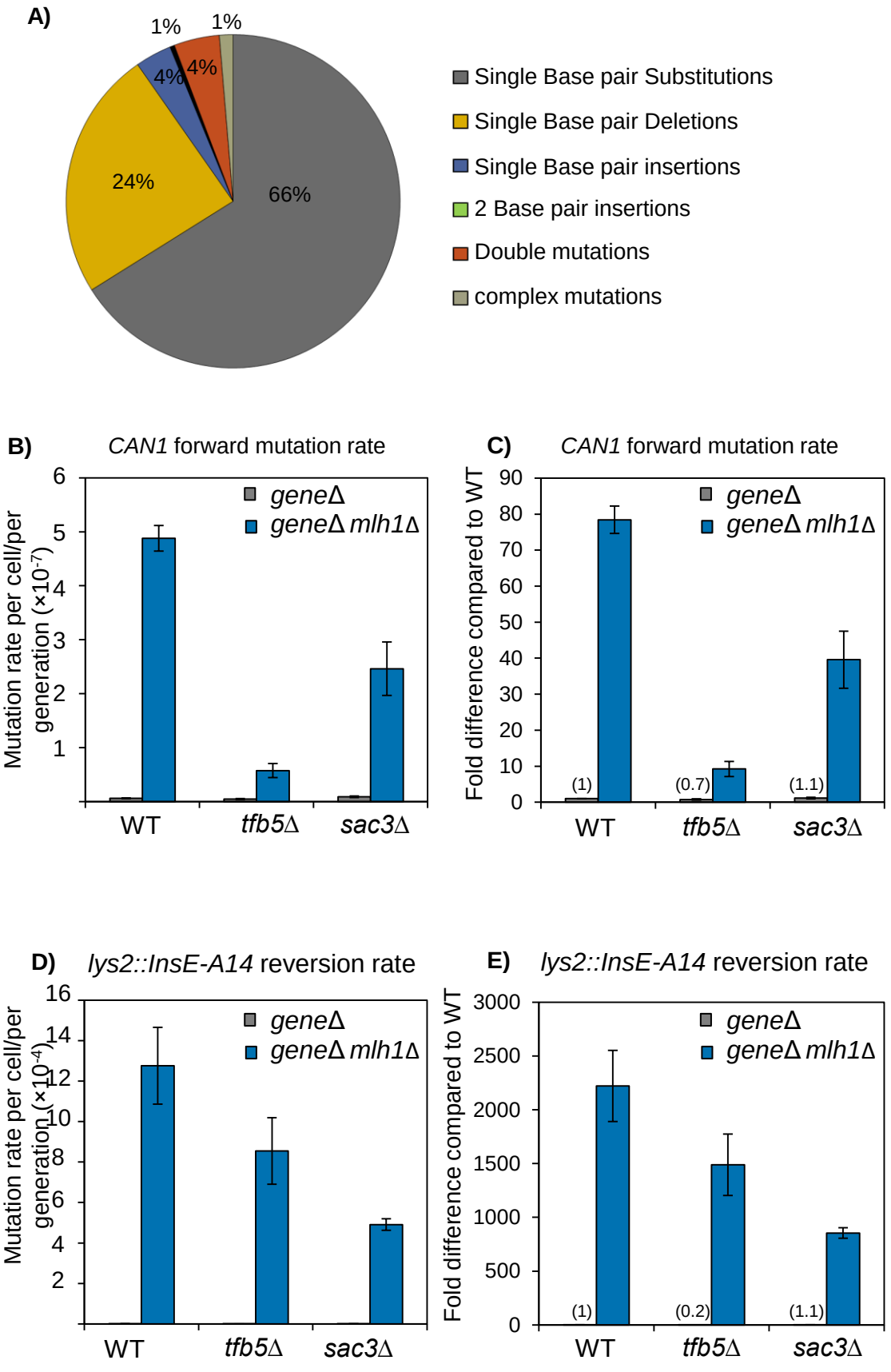
Interestingly for the *tfb5*Δ*mlh1*Δ strain there was a difference in the mutation rate depending on type of mutations assessed which needs to be further investigated. As the transformation using the lys2::InsE-A14 construct for *def1*Δ and *tho2*Δ strains did not present any positive transformants; these strains could not be assessed for frameshift mutations. Therefore future experiments are required to assess the mutations caused by MMR loss in these strains.

Figure 5.4 Frameshift mutations in repetitive DNA sequences are lower in TCR deletion strains when *MLH1* is deleted.

A) Type of mutations accumulated in *CAN1*. Data is obtained from Lang, 2008.

B) Average mutation rates per cell for TCR deletion using *CAN1* forward mutation rate assay. Mutation rates were determined using method of median, grey bars represent gene of interest deletion. blue bars represent double deletions i.e. gene of interest and *MLH1* fold difference values for *TFB5* and *SAC3* deletion strains are mentioned in brackets.

C) Average mutation rates per cell for TCR deletion using Lys-14A reporter assay. Mutation rates were determined using method of median, grey bars represent gene of interest deletion. Blue bars represent double deletions i.e. gene of interest and *MLH1*; fold difference values for *TFB5* and *SAC3* deletion strains are mentioned in brackets


A)
1%
1%
4%
4%
24%
66%
Single Base pair Substitutions
Single Base pair Deletions
Single Base pair insertions
2 Base pair insertions
Double mutations
complex mutations
B)
CAN1 forward mutation rate
Mutation rate per cell/per generation (×10-7)
geneΔ
geneΔ mlh1Δ
WT
tfb5Δ
sac3Δ
C)
CAN1 forward mutation rate
Fold difference compared to WT
(1)
(0.7)
(1.1)
D)
lys2::InsE-A14 reversion rate
Mutation rate per cell/per generation (×10-4)
E)
lys2::InsE-A14 reversion rate
Fold difference compared to WT
(1)
(0.2)
(1.1)

5.5 Deletions of genes essential for TCR for yeast lead to lower accumulation of mutations in *mlh1*△ cells.

Dissection of different NER pathways in yeast indicates that there are two main pathways through which TCR can take place. The first pathway is Rpb9-mediated TCR, which mainly operates in coding regions. The second is Rad26-mediated TCR and operates equally well in both coding regions and in regions upstream of the transcription start site (S. Li & Smerdon, 2002). To assess if the observed lower mutation accumulation in the absence of TCR cluster genes is due to inefficient TCR activity, genes essential for yeast TCR were deleted and the *CAN1* mutation rate measured using fluctuation tests.

When compared to WT, deletion of *RPB9* and *RAD26* did not increase mutation rate (Figure 5.4A and B). One possible reason could be that as there is no external damage induced (one of the main triggers for TCR deletion of these genes), thus the mutation rate does not increase. It has also been previously identified that the TCR and NER pathways are overlapping in yeast, thus the loss of TCR does not create a strong mutator phenotype. In *rad26*Δ*mlh1*Δ double mutants the mutation rate was as high as *mlh1*Δ alone, whereas surprisingly in a *rpb9*Δ*mlh1*Δ strain the mutation rate was increased but not as high as *mlh1*Δ (Fig5.5 A and B).

To further asses if defects in NER can also lead to a similar phenotype, the single-stranded DNA endonuclease complex Rad1-Rad10 was inactivated by deletion of *RAD10*. Rad1-Rad10 is responsible for cleavage of single-stranded

DNA during NER. The deletion of *MLH1* in the *rad10*Δ strain increased mutation as high as *mlh1*Δ (Fig5.5 A and B).

These results suggest that in absence of *MLH1;* deletion of the genes that are identified in our screen do not have the same phenotype as deletion of *RAD26* or *RAD10.* However *rpb9*Δ*mlh1*Δ strain did have lower accumulation of mutations in *CAN1* as compared to *mlh1*Δ indicating a possible role for *RPB9* in mutagenesis. A*s* Rpb9*,* Def1, Sac3, Tfb5 and Tho2 all have a direct interaction with RNAPII it is possible that they influence the transcription associated mutagenesis (TAM) and MMR then could repair these mutations*.* However we have not been able to investigate it any further but future experiments will be carried out to verify impact of these deletions in combination with loss of MMR in transcription associated mutagenesis using inducible promoter assays.

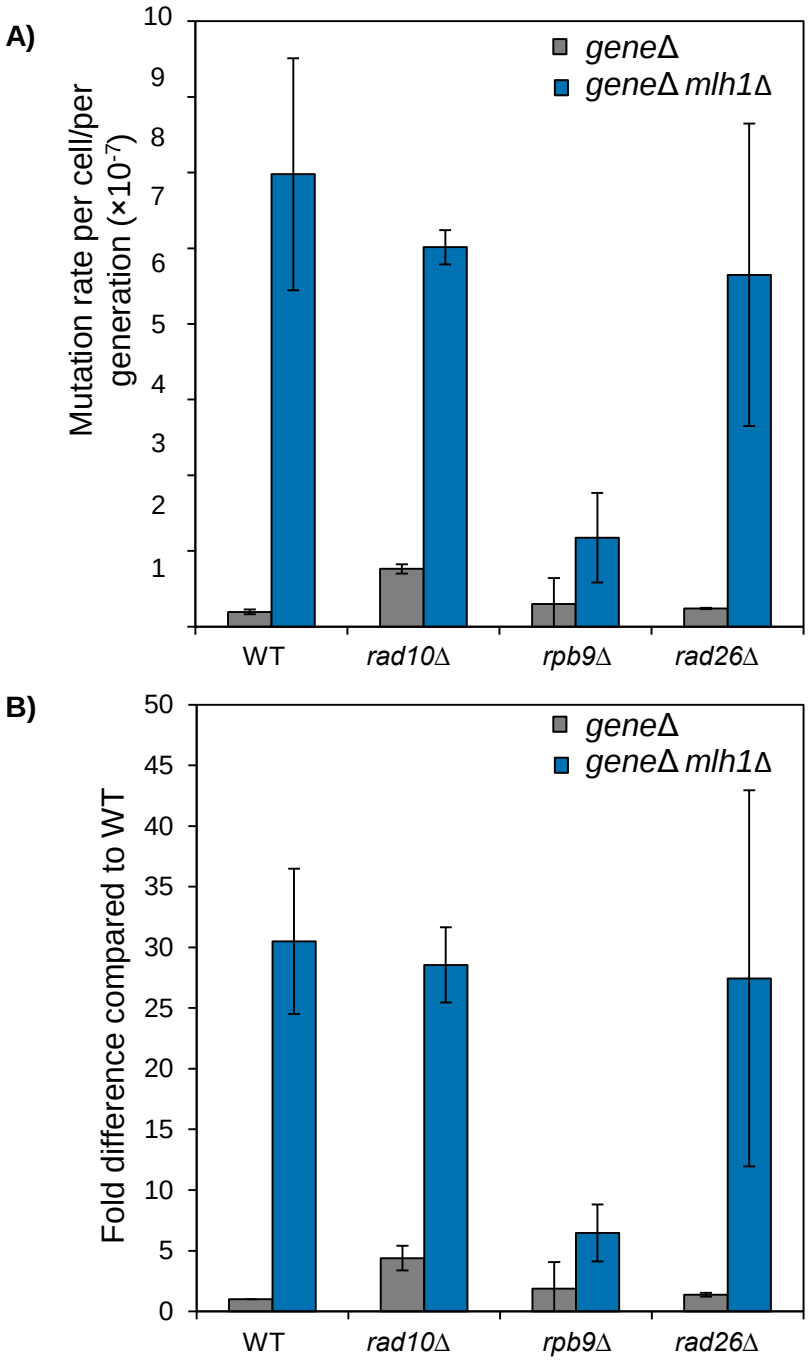


Figure 5.5 Deletions of genes essential for TCR for yeast lead to similar accumulation of mutations in mlh1Δ cells.

A) Average mutation rates per cell for TCR deletion using CAN^r reporter assay. Mutation rates were determined using method of median, 3 independent experiments were performed each time 5 colonies per gene deletion were used. Grey bars represent gene of interest deletion. Blue bars represent double deletions i.e. gene of interest and *MLH1*.

B) Fold differences relative to WT. Grey bars represent gene of interest deletion. Blue bars represent double deletions with gene of interest and *MLH1* (exp ID:E77_as01).

5.6 Discussion

The various DNA repair pathways play an essential role in protecting DNA from endogenous and exogenous DNA damage and preserving genetic integrity (Datta & Jinks-Robertson, 1995); therefore, it is important to understand how these pathways are regulated and how they interact with each other. Previous studies have confirmed that mutagenesis increases with higher transcription, due to the single-stranded nature of the non-transcribed strand of DNA during highly activated transcription. This configuration leads to enhanced accessibility of DNA to endogenous damaging agents, leading to an increase in mutation- or recombination-initiating lesions (Datta & Jinks-Robertson, 1995). Another possible reason behind this outcome is the collision between the replication fork and transcription machinery, which may result in a collapsed replication fork that may be further repaired by error prone recombination.

Our data suggest that even though the deletion of essential genes for TCR such as *rpb9*Δ and *rad26*Δ seem to have the similar accumulation of mutation in absence of MMR gene *MLH1*. The deletion of genes associated with TCR regulation (i.e. *DEF1*, *SAC3*, *TFB5* and *THO2*) seem to accumulate lower mutations compared to *mlh1*Δ alone, thus emphasizing the role of these genes in driving mutagenesis. There are three possible explanations for accumulating lower mutations in the absence of *MLH1* and TCR regulatory genes.

Firstly, the mutator genes associated with TCR could interact physically with *MLH1* and impact the efficiency of MMR and/or replication, thus resulting in a higher mutation rate. Therefore, the absence of *MLH1*, together with one of

these mutator genes, would result in a lower mutation rate. However, the data available for establishing a direct interaction between MMR and TCR proteins is inconsistent. In *S. cerevisiae*, strains deficient in MMR have decreased TCR mediated repair of thymine glycol adducts, but are capable of repairing UV damage (Leadon & Avrutskaya, 1998; Sweder et al., 1996). Further mutations in the heterodimer Msh2-Msh3, in combination with the NER complex Rad1-Rad10, leads to a defect in repair of large insertion-deletion mispairs (Kirkpatrick & Petes, 1997). Consistent with these findings, an in vitro study in Drosophila suggested a possible role of *MEI9* in MMR; a homologue of *S. cerevisiae RAD1* (Bhui-Kaur et al., 1998). However, a study by Sweder et al. showed contradicting findings suggesting that TCR in yeast was unaffected by MMR gene mutations (Sweder et al., 1996). Further, *MSH2* deficient mouse embryonic fibloblasts do not have a defect in TCR due to CPD repair on the faster non-transcribed strand as compared to the transcribed strand of active genes (Mellon et al., 1996). Currently there is no data to support a direct interaction between *DEF1*, *SAC3*, *TFB5* or *THO2* with *MLH1* or any of the MMR genes; therefore, a physical interaction or direct effect on replication fidelity and/or mutagenesis in association with MMR is unlikely.

Secondly, the mutator genes may have a direct mutagenic impact on replication or on basal mutations that are then repaired by MMR. However, previous studies with single deletions of *tho2*Δ and *sac3*Δ have shown contradicting data, suggesting that deletion of these genes actually leads to genomic instability and an increase in recombination (Chavez et al., 2000). For *def1*Δ and *tfb5*Δ deletions there are no direct associations with replication fidelity. One

way to check the involvement of these TCR regulating genes with mutations arising during replication is by combining the defect in replication with these gene deletions. To accumulate replication errors we used three active-site DNA polymerase mutants (*pol1- L868M*, *pol2-M644G*, and *pol3-L612M*), which all display a weak mutator phenotype because their DNA polymerase proofreading activity is not compromised (T. T. Schmidt et al., 2017). Deletion of *MLH1* causes more than 100 fold increase in mutation rate in these three active-site DNA polymerase mutants. In order to make a strain deficient in polymerase activity together with a TCR cluster gene and *MLH1* deletion, transformations were carried out which did not lead to any positive transformants as these triple mutants were all eventually lethal (data not shown). For example, when we attempted to cross the TCR clustered gene *THO2* deletion with a polymerase and *MLH1* double mutant, it led to very few viable spores which could germinate well but died after 4-5 divisions. The reason behind the cell death needs to be further investigated as the double mutants of polymerase active site mutants and *mlh1*Δ, as well as TCR cluster genes and *mlh1*Δ double mutants, are all viable.

Finally, there is the possibility of no direct involvement of MMR and/or TCR in affecting mutation rates. Instead, the mutator genes affect the efficiency of transcription activity in a way that increases transcription associated mutations (e.g., by avoiding collisions between the replication fork and transcription machinery) (Bradford,). Mutations repaired by MMR occur through multiple traditional mechanisms, such as. mis-incorporation of nucleotides during replication or error-prone activity of DNA, but there may still be unidentified

mutations that need MMR activity. Previous studies have shown that single-stranded DNA (ssDNA) is more susceptible to many chemical reactions than double- stranded DNA (dsDNA). For example; spontaneous deamination of cytosine is 140-fold more efficient on ssDNA than on dsDNA (Frederico et al., 1990). Interestingly, in the *E.coli* tac region, transcription causes a 4- to 5-fold increase in C to T mutations (Beletskii & Bhagwat, 1996). Furthermore, transcription associated mutagenesis is higher in a mutant of T7 RNA polymerase, which has a slower elongation rate. This suggests that cytosine deamination in the non-transcribed strand may be dependent on the length of time that the DNA remains open during transcription and elongation (Beletskii et al., 2000). As the genes *THO2* and *SAC3* have a role in processing of mRNA in WT cells, it is possible that they have an active role in extending the time that transcription maintains the open DNA configuratoin during elongation and leads to a higher accumulation of mutations in general.

In summary, our data suggest that in the absence of MMR, deletion strains of the TCR regulatory genes *DEF1*, *TFB5*, *THO2* and *SAC3* accumulate lower mutations. Understanding the exact mutation patterns could further enhance our knowledge of mutagenesis in these strains, but as there were very few canavanine resistant colonies it was challenging to select samples for sequencing and identifying their specific mutations.

6 Chapter 6- Assessment of spontaneous mutation rate in mammalian cells defective in mismatch repair.

6.1 Introduction

In the genome wide mutagenesis screen we identified a cluster of four genes associated with transcription coupled repair (TCR) with potential roles as mutator genes. I.e., deletion of the genes in the strains with hypomorphic MMR lowered the mutation rate this was the case both the *hMLH1* was expressed and when *MLH1* was deleted entirely (Chapter 5). We therefore wanted to investigate whether TCR may also be a cause of endogenous mutagenesis in mammalian cells.

Defects in TCR are associated with a rare, neuro-degeneration disorder known as Cockayne syndrome (CS). CS is categorized into three subtypes pending the severity and onset. CS patients have mutations in RNA polymerase co- factors CSA (*ERCC8*) and CSB (*ERCC6*). CSA and CSB interact with p44, a subunit of TFIIH, in TCR. CSA consists of 398 amino acids and belongs to the WD repeat family. The protein contains seven WD40 motifs and beta propeller structure and forms an ubiquitin ligase complex together with DDB1, Cullin 4A, and Roc1. CSB is a 168 kDA protein (1493 amino acids) that belongs to the SWI2/SNF2 family of DNA-dependent ATPases. The ATPase domain is encoded by seven motifs but the protein has no conventional strand displacement activity (Citterio,

2000). Instead, CSB contains ATP-dependent chromatin remodeling activity as well as strand annealing and exchange activities (Muftuoglu, 2006).

Deficiency of global nucleotide excision repair (GGR) causes xeroderma pigmentosum (XP) which is characterized by increased skin cancer incidences (Bradford, 2011; Kraemer, 1984;Cleaver, 2009). CS and XP patients do have similar symptoms of high photosensitivity, however CS patients have a significantly lower incidence of cancers (Kubota, 2015; Lehmann, 1987 ;Wilson, 2013; Zhang, 2016; Nance, 1992; Cleaver, 2009). One of the most likely reasons is that CS patients have a significantly shorter life span than XP patients. Our observations that TCR contributes towards endogenous mutagenesis in budding yeast could also explain the lower incidence of cancers in CS patients, if this function of TCR is conserved.

Previous studies that assessed the accumulation of mutations after UV irradiation in cells from patients with defects in TCR and NER have shown contrasting results (Parris, 1993; Muriel, 1991; Lin, 1995). However a recent sequencing study revealed that CSA and CSB patient cells accumulate fewer UV induced mutations compared to the NER deficient XP-C cells and wild type cells. This is consistent with our hypothesis that functional TCR is important for generating endogenous mutations. It further suggests that the cytotoxicity observed in CS cells is not regulated by accumulation of mutations (Reid-Bayliss, 2016).

In our screen, we identified four genes from the *S. cerevisiae* TCR pathway. Def1 is required for removal and degradation of RNA polymerase II from the lesion site and there is no known orthologues in humans. Tho2 is conserved from yeast to humans and hTho2 is a subunit of human TREX complex, which is required coupling transcription elongation to mRNA export (Strasser, 2002). Tfb5 is a part of yeast TFIIH complex and is also thought to be part of mammalian TFIIH. Mutations in TFIIH are implicated in causing a DNA repair–deficient form of the tricothiodystrophy disorder called TTD-A2 (Giglia-Mari, 2004; Ranish, 2004). Finally, the fourth gene identified in our screen, Sac3 in yeast, is important for nuclear export of mRNA. Sac3 protein in yeast is associated with several processes that include actin filament-based process, mRNA export from nucleus, and nucleic acid metabolic process. In yeast Sac3 localizes to the transcription export complex 2 and co-localizes with the nuclear pore. The human orthologous of *Sac3* include the *MCM3AP* (minichromosome maintenance complex component 3 associated protein) and its isoform, GANP (germinal-centre associated nuclear protein). The gene encoding MCM3AP is situated entirely within a much larger sequence encoding the GANP gene (Wickramasinghe, 2010). The localization of the human TREX-2 to the NPCs requires the MCM3AP domain of GANP (Jani, 2012). Therefore; to assess the effect of defects in mRNA export and spontaneous mutation rates in mammalian cells we started the experiment with siRNA of *MCM3AP*.

6.2 Mutation frequency for HCT116 cells using *HPRT* assay

HCT116 cells are colon cancer cells that do not express *MLH1*. The mutation frequency of HCT 116 cells was quantified using *HPRT* assay before and after cleansing with HAT media. HCT116 cells with extra chromosome 3 were used as a control for functional *MLH1*.

The cells were initially grown in HAT media for 3 days. After 3 days cells were washed and supplemented with hypoxanthine–thymidine (HT) media for 1 day, so that both *de novo* and salvage pathway for nucleotide biosynthesis will be functional. These cells were then cultured for 10 days to allow accumulation of mutations, and further plated with media containing 0.6 µg/ml 6-Thioguanine to assay mutation frequency. Cells were incubated for 14 days to allow colony formation. Number of colonies was counted after staining with crystal violet and mutation frequency was calculated (mutation frequency = Avg number of colonies per plate/ number of cells seeded (in this case 200,000)*cloning efficiency).

After three independent experiments, it was observed that HCT116 cells and HCT116 +chromosome 3 treated with HAT medium have lower mutation frequency as compared to cells without HAT treatment (Fig. 7.3). This suggests that treatment with aminopterin reduced the pre-existing *HPRT* mutant cells, thus lowering the mutation frequency. These cleansed populations of cells were used for further experiments. The mutation frequency of HCT116+chromosome 3 cells was significantly lower when compared to HCT116 cells, this suggests

that the accumulation of mutation in HCT116 cells is due to loss of functional *MLH1*.

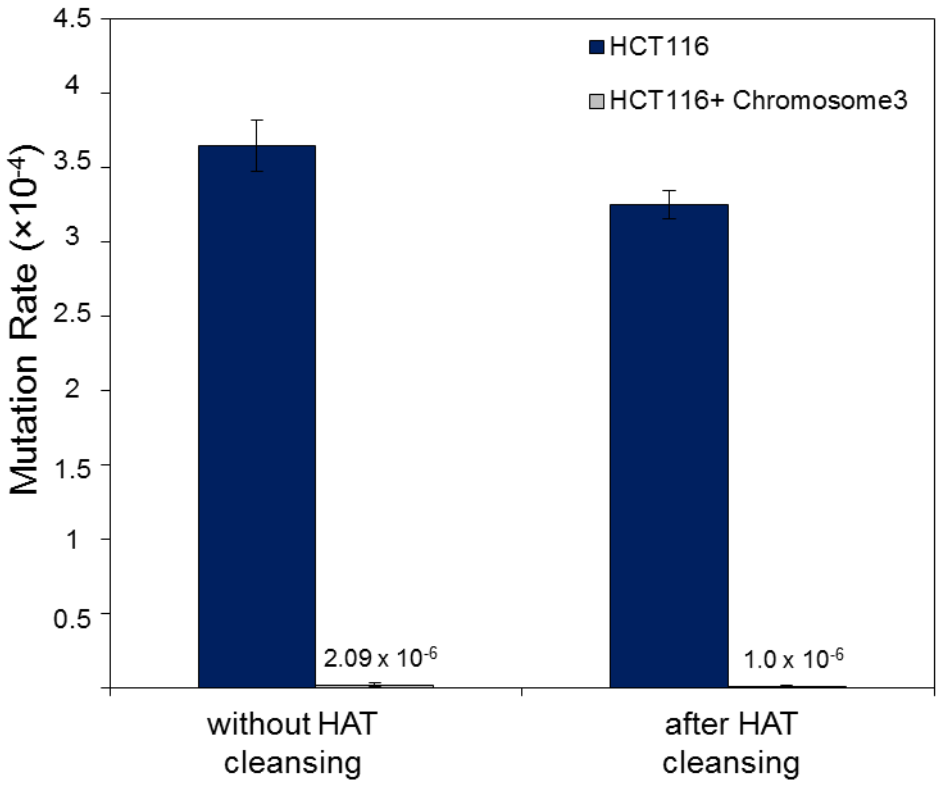


Figure 6.1: Mutation rate of HCT116 and HCT116+ch3 cells after cleansing with HAT media

Average mutation rates per cell for HCT116 and HCT116+ch3 were calculated using *HPRT* reporter assay. 3 independent experiments were performed. Mutation frequency was calculated using the formula mutation frequency = Avg number of colonies per plate/ number of cells seeded (in this case 200,000)*cloning efficiency.

6.3 siRNA transfection of *MCM3AP* does not affect mutation frequency at *HPRT.*

To test if loss of the orthologue of *SAC3*, which we identified in our screen), leads to lower accumulation of mutation in *MLH1-/-* cells, we were treated HCT116 and HCT116-chr3 with siRNA against *MCM3AP.*

The siRNA used for *MCM3AP* knockdown was obtained from already published study, which showed *MCM3AP* specific knockdown (Poole, 2012). The *HPRT* assay was performed similar to mentioned in section 7.3. HCT116 and HCT116 cells with extra chromosome 3 were cleansed for pre-existing HPRT mutations. Cells were then cultured for 7 days to allow accumulation of mutation in absence of both *MCM3AP* and *MLH1*. The transfection with siRNA was repeated every 48 hours to keep *MCM3* expression knocked down. After 7 days cells were plated for colony formation in presence of 6-TG, and mutation frequency was assessed by counting colonies.

The mutation frequency of the cells treated with siRNA for *MCM3*AP was not different that with control siRNA for HCT116 and HCT116 + Chromosome 3 cells (Fig. 7.4). However due to lack of efficient antibody, a western blot could not be performed to assess the expression of *MCM3AP*. Thus it is not possible for now to conclude if the observed phenotype of no difference in mutation accumulation is a result of loss of MCM3AP.

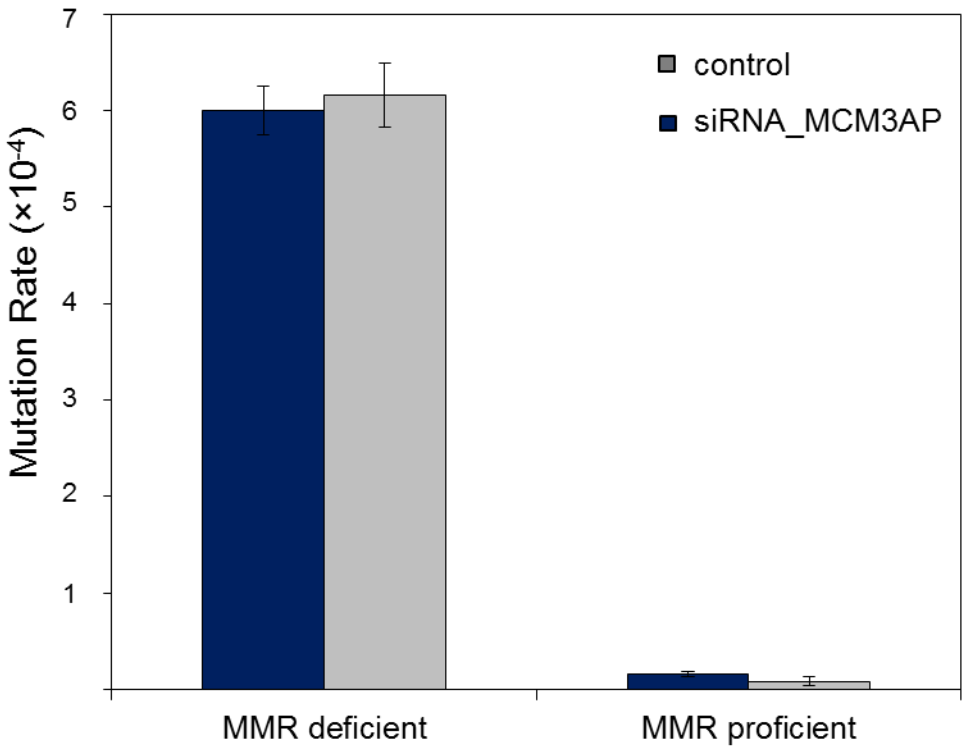


Figure 6.2: Mutations in HCT116 cells after siRNA knockdown of *MCM3AP*

Average mutation rates per cell for HCT116 and HCT116+ch3 were calculated using *HPRT* reporter assay. 3 independent experiments were performed. Mutation frequency was calculated using the formula mutation frequency = Avg number of colonies per plate/ number of cells seeded (in this case 200,000)*cloning efficiency.

6.4 Discussion

In our yeast genome wide screen, we identified a cluster of transcription coupled repair associated genes which show lower mutation rate in absence of mismatch repair. Even though the human orthologous of these yeast genes are not directly associated in TCR, their defects are associated with disorders like tricothiodystrophy disorder TTD-A2 (Giglia-Mari, 2004; Ranish, 2004;Ylikallio, 2017). As the transcription coupled repair in humans is more complex, we wanted to investigate if the orthologues of genes identified have a similar impact on mutation regulation in mammalian cells defective in MMR. In order to assess effect MCM3AP depletion on mutation rates we used MMR defective HCT116 cells and observed no effect on mutation rate using HPRT assay. However as the depletion of expression could not be verified using western blot, it needs further validation in terms of assessing MCM3AP depletion. Phenotypically loss of MCM3AP is associated with UV sensitivity and cell cycle arrest, so HCT116 cells treated with these siRNA will be checked for these phenotypes (Ylikallio, 2017).

The aim of this Chapter was to explore the impact of TCR on endogenous mutation rates in mammalian cells. To this end, I chose two approaches. The first approach is to create isogenic deletion cell lines using CRISPR for the human TCR genes and assess mutation rates using the *HPRT* assay and whole-genome sequencing. For this we are currently collaborating with Hickson lab at University of Copenhagen to generate CSA and CSB lines. The second strategy involves validation of the yeast data to assess mutation rate in mammalian cells in defective MMR background. CRISPR based gene deletion

would be an ideal approach for creating deletion cell lines, as use of siRNA requires transfecting cells every 48 hours which may add variability in mutation rates due to experimental errors. However as such generation of cell lines require a lot of time I started the preliminary experiments with the siRNA based knockdown for orthologous of one of the TCR cluster gene namely, MCM3AP/GANP.

Mutations in MCM3AP are associated with autosomal recessive disorder called as Charcot-Marie-Tooth (CMT) neuropathy which is associated with intellectual disability (Ylikallio, 2017). The majority of the reported mutations associated with CMT accumulate in the conserved Sac3 domain of GANP and MCM3P; thus implicating the defects in mRNA export in pathogenic processes (Karakaya, 2017). I therefore assessed whether reducing expression of MCM3AP would cause a reduction in the mutation rates of a MMR-deficient cell line.

Further studies will be carried out in isogenic mutant cells for TCR, in order to asses if TCR in humans is associated with mutagenesis and to investigate if MMR has any potential role in this TCR regulated mutagenesis.

Cockayne syndrome (CS) is caused by mutations in CSA, CSB, XPB, XPD, or XPG, which are required for transcription-coupled NER. Even though the photosensitivity of the CS patients is similar to XP patients having defect in Global NER, the rate of accumulation of mutation and incidence of cancer differs. Moreover with the help of sequencing, a recent study have presented that the type of mutations which accumulate in TCR deficient cells have a

mutational signature of oxidative DNA-damage as opposed to the UV-induced mutational pattern (Reid-Bayliss, 2016). This suggests that the deficiency of TCR has a different mutational spectrum than NER.

7 Chapter - Discussion

Germline mutations in mismatch repair (MMR) leads to the cancer predisposing condition Lynch syndrome (LS). Correct diagnosis of LS patients is necessary to be able to provide personalized counseling and necessary further genetic testing. The first step in identifying patients with an increased risk of having LS is testing for the expression of microsatellite instability (MSI) and/or MMR proteins; however, these criteria for LS diagnosis has many limitations (Umar et al., 2004). One such limitation is the high degree of heterogeneity between MSI expression and the type of cancer, for example, MSIs are found in 90% of colorectal cancer (CRC) patients and only 20-30% of endometrial cancer patients associated with Lynch syndrome(Hampel et al., 2006). This variability suggests that the relationship between MSI and tumorigenesis is not clearly established. Another limitation in diagnosing LS is the lack of an identifiable pathogenic mutation in approximately one-third of suspected LS cases (Rustgi, 2007). This suggests that an alternative mechanism or gene(s) are associated with causing LS. Both limitations suggest that MMR mutation is not the only source for MSI expression and other influences, such as environmental or polygenic factors, may affect phenotypic expression.

To understand the mutation landscape when MMR activity is decreased, a genome wide mutagenesis screen in yeast was carried out with the preliminary objective to uncover mechanisms that affect the spontaneous mutations in a cell. This would have not been visible in MMR proficient cells as MMR is crucial in repairing the basal spontaneous mutations of a cell. This screen discovered

that around 163 gene deletions increased mutation rate and surprisingly 543 gene deletions decreased mutation rate when deleted in the absences of MMR. Gene ontology analysis identified that majority of metabolic genes lower mutation rate when deleted in mismatch repair deficient background. As discussed in Chapter 3.4, our findings raise the possibility that genetic background and metabolic activity regulate endogenous mutation rates, even when MMR is inactive. Even though the results the screens were reproducible quantitatively (for 17 genes associated with TCR and replication) and the estimated FDR is only 2%, the screen was carried out qualitatively; therefore more direct approaches like using GFP based mismatch repair assays needs to be used to further validate the screen.

The 163 genes that showed an increase in mutation rate are enriched in biological processes such as DNA repair and DNA replication but may also impact the efficiency of MMR pathway thereby have a possible role in causing Lynch syndrome. Further studies on *trm2*Δ, *rnh201*Δ, *rrm3*Δ, *tel1*Δ in absence of hydrolytic activity of *MLH1* showed that *RRM3* deletion leads to synergistic effect. This suggests a possible role for *RRM3* in compensating for lower MMR activity. It needs to be further investigated if human homologue of *RRM3* can also compensate for reduced MMR activity. Phenotypic enhancement studies have shown the interaction of Rrm3 and Pif1 with polymerase δ (Polδ); hence it would be interesting to understand if these helicases have a possible role in strand displacement activity of Polδ and thereby in MMR.

The enrichment of genes associated with TCR regulation points at a possibility that TCR is a mutagenic process that requires MMR for its repair. Further experiments are required to understand if it is possible that TCR affects mutation rate either by directly introducing errors or by regulation the transcription associated mutagenesis. This also highlights a role of MMR in highly transcribed regions, and suggest that defects in MMR in highly transcribed genes would have higher impact on mutagenesis and can provide a possible explanation for tissue specific low MSI levels. However it needs to be further verified if in mammalian cells we can reduce mutation rate, or see a different mutational spectrum in absence of the homologs of these TCR regulatory genes.

It is known that mutation rate differ across the eukaryotic genome and defects in one of the DNA repair pathway impact mutational landscape differently (Roberts & Gordenin, 2014). For example defects in MMR causes up to 100,000 fold increase in long repetitive DNA (Lujan et al., 2015). However defects in exonuclease domain of replicative polymerase leads to increase indel rates in short run sequences (Lujan et al., 2015). This different mutational pattern also impact the tumour phenotype. The MMR defective tumours are often associated with MSI-H phenotype; whereas somatic mutations in *POLE* are associated with microsatellite stable tumours even after having higher mutation load than MSI positive tumours (Briggs & Tomlinson, 2013). Our data suggests that the loss of MMR in combination with other cellular processes can have a differential impact on the basal mutation rate of a cell. Understanding this mutational landscape

can be used to further increase our knowledge of tissue specific MSI expression and carcinogenesis.

8 References

Acharya, S., Foster, P. L., Brooks, P., & Fishel, R. (2003). The coordinated functions of the E. coli MutS and MutL proteins in mismatch repair. *Mol Cell, 12*(1), 233-246.

Aguilera, A., & Klein, H. L. (1988). Genetic control of intrachromosomal recombination in Saccharomyces cerevisiae. I. Isolation and genetic characterization of hyper-recombination mutations. *Genetics, 119*(4), 779-790.

Alhopuro, P., Sammalkorpi, H., Niittymaki, I., Bistrom, M., Raitila, A., Saharinen, J., Nousiainen, K., Lehtonen, H. J., Heliovaara, E., Puhakka, J., Tuupanen, S., Sousa, S., Seruca, R., Ferreira, A. M., Hofstra, R. M., Mecklin, J. P., Jarvinen, H., Ristimaki, A., Orntoft, T. F., Hautaniemi, S., Arango, D., Karhu, A., & Aaltonen, L. A. (2012). Candidate driver genes in microsatellite-unstable colorectal cancer. *Int J Cancer, 130*(7), 1558-1566. doi:10.1002/ijc.26167

Amin, N. S., Nguyen, M. N., Oh, S., & Kolodner, R. D. (2001). exo1-Dependent mutator mutations: model system for studying functional interactions in mismatch repair. *Mol Cell Biol, 21*(15), 5142-5155. doi:10.1128/MCB.21.15.5142-5155.2001

Araujo, S. J., Nigg, E. A., & Wood, R. D. (2001). Strong functional interactions of TFIIH with XPC and XPG in human DNA nucleotide excision repair, without a preassembled repairosome. *Mol Cell Biol, 21*(7), 2281-2291. doi:10.1128/MCB.21.7.2281-2291.2001

Au, K. G., Welsh, K., & Modrich, P. (1992). Initiation of methyl-directed mismatch repair. *J Biol Chem, 267*(17), 12142-12148.

Awwad, S. W., & Ayoub, N. (2015). Overexpression of KDM4 lysine demethylases disrupts the integrity of the DNA mismatch repair pathway. *Biol Open, 4*(4), 498-504. doi:10.1242/bio.201410991

Bachrati, C. Z., & Hickson, I. D. (2008). RecQ helicases: guardian angels of the DNA replication fork. *Chromosoma, 117*(3), 219-233. doi:10.1007/s00412-007-0142-4

Ban, C., & Yang, W. (1998). Crystal structure and ATPase activity of MutL: implications for DNA repair and mutagenesis. *Cell, 95*(4), 541-552.

Barlow, J. H., Faryabi, R. B., Callen, E., Wong, N., Malhowski, A., Chen, H. T., Gutierrez-Cruz, G., Sun, H. W., McKinnon, P., Wright, G., Casellas, R., Robbiani, D. F., Staudt, L., Fernandez-Capetillo, O., & Nussenzweig, A. (2013). Identification of early replicating fragile sites that contribute to genome instability. *Cell, 152*(3), 620-632. doi:10.1016/j.cell.2013.01.006

Batty, D., Rapic'-Otrin, V., Levine, A. S., & Wood, R. D. (2000). Stable binding of human XPC complex to irradiated DNA confers strong discrimination for damaged sites. *J Mol Biol, 300*(2), 275-290. doi:10.1006/jmbi.2000.3857

Beletskii, A., & Bhagwat, A. S. (1996). Transcription-induced mutations: increase in C to T mutations in the nontranscribed strand during transcription in Escherichia coli. *Proc Natl Acad Sci U S A, 93*(24), 13919-13924.

Bhatia, P. K., Verhage, R. A., Brouwer, J., & Friedberg, E. C. (1996). Molecular cloning and characterization of Saccharomyces cerevisiae RAD28, the yeast homolog of the human Cockayne syndrome A (CSA) gene. *J Bacteriol, 178*(20), 5977-5988.

Bhui-Kaur, A., Goodman, M. F., & Tower, J. (1998). DNA mismatch repair catalyzed by extracts of mitotic, postmitotic, and senescent Drosophila tissues and involvement of mei-9 gene function for full activity. *Mol Cell Biol, 18*(3), 1436-1443.

Bjornson, K. P., Blackwell, L. J., Sage, H., Baitinger, C., Allen, D., & Modrich, P. (2003). Assembly and molecular activities of the MutS tetramer. *J Biol Chem, 278*(36), 34667-34673. doi:10.1074/jbc.M305513200

Boland, C. R., Thibodeau, S. N., Hamilton, S. R., Sidransky, D., Eshleman, J. R., Burt, R. W., Meltzer, S. J., Rodriguez-Bigas, M. A., Fodde, R., Ranzani, G. N., & Srivastava, S. (1998). A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res, 58*(22), 5248-5257.

Branzei, D., & Foiani, M. (2007). Interplay of replication checkpoints and repair proteins at stalled replication forks. *DNA Repair (Amst), 6*(7), 994-1003. doi:10.1016/j.dnarep.2007.02.018

Breen, A. P., & Murphy, J. A. (1995). Reactions of oxyl radicals with DNA. *Free Radic Biol Med, 18*(6), 1033-1077.

Briggs, S., & Tomlinson, I. (2013). Germline and somatic polymerase epsilon and delta mutations define a new class of hypermutated colorectal and endometrial cancers. *J Pathol, 230*(2), 148-153. doi:10.1002/path.4185

Brosh, R. M., Jr., & Bohr, V. A. (2007). Human premature aging, DNA repair and RecQ helicases. *Nucleic Acids Res, 35*(22), 7527-7544. doi:10.1093/nar/gkm1008

Bruni, R., Martin, D., & Jiricny, J. (1988). d(GATC) sequences influence Escherichia coli mismatch repair in a distance-dependent manner from positions both upstream and downstream of the mismatch. *Nucleic Acids Res, 16*(11), 4875-4890.

Burdett, V., Baitinger, C., Viswanathan, M., Lovett, S. T., & Modrich, P. (2001). In vivo requirement for RecJ, ExoVII, ExoI, and ExoX in methyl-directed mismatch repair. *Proc Natl Acad Sci U S A, 98*(12), 6765-6770. doi:10.1073/pnas.121183298

Burrows, C. J., & Muller, J. G. (1998). Oxidative Nucleobase Modifications Leading to Strand Scission. *Chem Rev, 98*(3), 1109-1152.

Chabes, A., Georgieva, B., Domkin, V., Zhao, X., Rothstein, R., & Thelander, L. (2003). Survival of DNA damage in yeast directly depends on increased dNTP levels allowed by relaxed feedback inhibition of ribonucleotide reductase. *Cell, 112*(3), 391-401.

Chao, E. C., & Lipkin, S. M. (2006). Molecular models for the tissue specificity of DNA mismatch repair-deficient carcinogenesis. *Nucleic Acids Res, 34*(3), 840-852. doi:10.1093/nar/gkj489

Chavez, S., Beilharz, T., Rondon, A. G., Erdjument-Bromage, H., Tempst, P., Svejstrup, J. Q., Lithgow, T., & Aguilera, A. (2000). A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects

transcription elongation with mitotic recombination in Saccharomyces cerevisiae. *EMBO J, 19*(21), 5824-5834. doi:10.1093/emboj/19.21.5824

Chen, X., Paudyal, S. C., Chin, R. I., & You, Z. (2013). PCNA promotes processive DNA end resection by Exo1. *Nucleic Acids Res, 41*(20), 9325-9338. doi:10.1093/nar/gkt672

Chiarle, R., Zhang, Y., Frock, R. L., Lewis, S. M., Molinie, B., Ho, Y. J., Myers, D. R., Choi, V. W., Compagno, M., Malkin, D. J., Neuberg, D., Monti, S., Giallourakis, C. C., Gostissa, M., & Alt, F. W. (2011). Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. *Cell, 147*(1), 107-119. doi:10.1016/j.cell.2011.07.049

Choudhury, S. A., Asefa, B., Webb, A., Ramotar, D., & Chow, T. Y. (2007). Functional and genetic analysis of the Saccharomyces cerevisiae RNC1/TRM2: evidences for its involvement in DNA double-strand break repair. *Mol Cell Biochem, 300*(1-2), 215-226. doi:10.1007/s11010-006-9386-1

Christians, F. C., & Hanawalt, P. C. (1994). Repair in ribosomal RNA genes is deficient in xeroderma pigmentosum group C and in Cockayne's syndrome cells. *Mutat Res, 323*(4), 179-187.

Clark, A. B., Lujan, S. A., Kissling, G. E., & Kunkel, T. A. (2011). Mismatch repair-independent tandem repeat sequence instability resulting from ribonucleotide incorporation by DNA polymerase epsilon. *DNA Repair (Amst), 10*(5), 476-482. doi:10.1016/j.dnarep.2011.02.001

Clark, A. B., Valle, F., Drotschmann, K., Gary, R. K., & Kunkel, T. A. (2000). Functional interaction of proliferating cell nuclear antigen with MSH2-MSH6 and MSH2-MSH3 complexes. *J Biol Chem, 275*(47), 36498-36501. doi:10.1074/jbc.C000513200

Clerici, M., Trovesi, C., Galbiati, A., Lucchini, G., & Longhese, M. P. (2014). Mec1/ATR regulates the generation of single-stranded DNA that attenuates Tel1/ATM signaling at DNA ends. *EMBO J, 33*(3), 198-216. doi:10.1002/embj.201386041

Collins, S. R., Miller, K. M., Maas, N. L., Roguev, A., Fillingham, J., Chu, C. S., Schuldiner, M., Gebbia, M., Recht, J., Shales, M., Ding, H., Xu, H., Han, J., Ingvarsdottir, K., Cheng, B., Andrews, B., Boone, C., Berger, S. L., Hieter, P., Zhang, Z., Brown, G. W., Ingles, C. J., Emili, A., Allis, C. D., Toczyski, D. P., Weissman, J. S., Greenblatt, J. F., & Krogan, N. J. (2007). Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. *Nature, 446*(7137), 806-810. doi:10.1038/nature05649

Colussi, C., Parlanti, E., Degan, P., Aquilina, G., Barnes, D., Macpherson, P., Karran, P., Crescenzi, M., Dogliotti, E., & Bignami, M. (2002). The mammalian mismatch repair pathway removes DNA 8-oxodGMP incorporated from the oxidized dNTP pool. *Curr Biol, 12*(11), 912-918.

Constantin, N., Dzantiev, L., Kadyrov, F. A., & Modrich, P. (2005). Human mismatch repair: reconstitution of a nick-directed bidirectional reaction. *J Biol Chem, 280*(48), 39752-39761. doi:10.1074/jbc.M509701200

Cooke, M. S., Evans, M. D., Dizdaroglu, M., & Lunec, J. (2003). Oxidative DNA damage: mechanisms, mutation, and disease. *FASEB J, 17*(10), 1195-1214. doi:10.1096/fj.02-0752rev

Crow, Y. J., Leitch, A., Hayward, B. E., Garner, A., Parmar, R., Griffith, E., Ali, M., Semple, C., Aicardi, J., Babul-Hirji, R., Baumann, C., Baxter, P., Bertini, E., Chandler, K. E., Chitayat, D., Cau, D., Dery, C., Fazzi, E., Goizet, C., King, M. D., Klepper, J., Lacombe, D., Lanzi, G., Lyall, H., Martinez-Frias, M. L., Mathieu, M., McKeown, C., Monier, A., Oade, Y., Quarrell, O. W., Rittey, C. D., Rogers, R. C., Sanchis, A., Stephenson, J. B., Tacke, U., Till, M., Tolmie, J. L., Tomlin, P., Voit, T., Weschke, B., Woods, C. G., Lebon, P., Bonthron, D. T., Ponting, C. P., & Jackson, A. P. (2006). Mutations in genes encoding ribonuclease H2 subunits cause Aicardi-Goutieres syndrome and mimic congenital viral brain infection. *Nat Genet, 38*(8), 910-916. doi:10.1038/ng1842

Datta, A., & Jinks-Robertson, S. (1995). Association of increased spontaneous mutation rates with high levels of transcription in yeast. *Science, 268*(5217), 1616-1619.

Davies, K. J. (2000). Oxidative stress, antioxidant defenses, and damage removal, repair, and replacement systems. *IUBMB Life, 50*(4-5), 279-289. doi:10.1080/713803728

de la Chapelle, A., & Hampel, H. (2010). Clinical relevance of microsatellite instability in colorectal cancer. *J Clin Oncol, 28*(20), 3380-3387. doi:10.1200/JCO.2009.27.0652

de Massy, B. (2013). Initiation of meiotic recombination: how and where? Conservation and specificities among eukaryotes. *Annu Rev Genet, 47*, 563-599. doi:10.1146/annurev-genet-110711-155423

Desai, A., & Gerson, S. (2014). Exo1 independent DNA mismatch repair involves multiple compensatory nucleases. *DNA Repair (Amst), 21*, 55-64. doi:10.1016/j.dnarep.2014.06.005

DeWeese, T. L., Shipman, J. M., Larrier, N. A., Buckley, N. M., Kidd, L. R., Groopman, J. D., Cutler, R. G., te Riele, H., & Nelson, W. G. (1998). Mouse embryonic stem cells carrying one or two defective Msh2 alleles respond abnormally to oxidative stress inflicted by low-level radiation. *Proc Natl Acad Sci U S A, 95*(20), 11915-11920.

Dizdaroglu, M., Jaruga, P., Birincioglu, M., & Rodriguez, H. (2002). Free radical-induced damage to DNA: mechanisms and measurement. *Free Radic Biol Med, 32*(11), 1102-1115.

Feng, Z., Hu, W., Chasin, L. A., & Tang, M. S. (2003). Effects of genomic context and chromatin structure on transcription-coupled and global genomic repair in mammalian cells. *Nucleic Acids Res, 31*(20), 5897-5906.

Fischer, T., Strasser, K., Racz, A., Rodriguez-Navarro, S., Oppizzi, M., Ihrig, P., Lechner, J., & Hurt, E. (2002). The mRNA export machinery requires the novel Sac3p-Thp1p complex to dock at the nucleoplasmic entrance of the nuclear pores. *EMBO J, 21*(21), 5843-5852.

Fishman-Lobell, J., & Haber, J. E. (1992). Removal of nonhomologous DNA ends in double-strand break recombination: the role of the yeast ultraviolet repair gene RAD1. *Science, 258*(5081), 480-484.

Flores-Rozas, H., Clark, D., & Kolodner, R. D. (2000). Proliferating cell nuclear antigen and Msh2p-Msh6p interact to form an active mispair recognition complex. *Nat Genet, 26*(3), 375-378. doi:10.1038/81708

Franchitto, A., Pichierri, P., Piergentili, R., Crescenzi, M., Bignami, M., & Palitti, F. (2003). The mammalian mismatch repair protein MSH2 is required for correct MRE11 and RAD51 relocalization and for efficient cell cycle arrest induced by ionizing radiation in G2 phase. *Oncogene, 22*(14), 2110-2120. doi:10.1038/sj.onc.1206254

Frank, P., Braunshofer-Reiter, C., & Wintersberger, U. (1998). Yeast RNase H(35) is the counterpart of the mammalian RNase HI, and is evolutionarily related to prokaryotic RNase HII. *FEBS Lett, 421*(1), 23-26.

Frederico, L. A., Kunkel, T. A., & Shaw, B. R. (1990). A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy. *Biochemistry, 29*(10), 2532-2537.

Gaillard, H., Garcia-Muse, T., & Aguilera, A. (2015). Replication stress and cancer. *Nat Rev Cancer, 15*(5), 276-289. doi:10.1038/nrc3916

Gaillard, H., Tous, C., Botet, J., Gonzalez-Aguilera, C., Quintero, M. J., Viladevall, L., Garcia-Rubio, M. L., Rodriguez-Gil, A., Marin, A., Arino, J., Revuelta, J. L., Chavez, S., & Aguilera, A. (2009). Genome-wide analysis of factors affecting transcription elongation and DNA repair: a new role for PAF and Ccr4-not in transcription-coupled repair. *PLoS Genet, 5*(2), e1000364. doi:10.1371/journal.pgen.1000364

Gaillard, H., Wellinger, R. E., & Aguilera, A. (2007). A new connection of mRNP biogenesis and export with transcription-coupled repair. *Nucleic Acids Res, 35*(12), 3893-3906. doi:10.1093/nar/gkm373

Gallardo, M., Luna, R., Erdjument-Bromage, H., Tempst, P., & Aguilera, A. (2003). Nab2p and the Thp1p-Sac3p complex functionally interact at the interface between transcription and mRNA metabolism. *J Biol Chem, 278*(26), 24225-24232. doi:10.1074/jbc.M302900200

Gedik, C. M., Collins, A., & Escodd. (2005). Establishing the background level of base oxidation in human lymphocyte DNA: results of an interlaboratory validation study. *FASEB J, 19*(1), 82-84. doi:10.1096/fj.04-1767fje

Ghodgaonkar, M. M., Lazzaro, F., Olivera-Pimentel, M., Artola-Boran, M., Cejka, P., Reijns, M. A., Jackson, A. P., Plevani, P., Muzi-Falconi, M., & Jiricny, J. (2013). Ribonucleotides misincorporated into DNA act as strand-discrimination signals in eukaryotic mismatch repair. *Mol Cell, 50*(3), 323-332. doi:10.1016/j.molcel.2013.03.019

Giannini, G., Ristori, E., Cerignoli, F., Rinaldi, C., Zani, M., Viel, A., Ottini, L., Crescenzi, M., Martinotti, S., Bignami, M., Frati, L., Screpanti, I., & Gulino, A. (2002). Human MRE11 is inactivated in mismatch repair-deficient cancers. *EMBO Rep, 3*(3), 248-254. doi:10.1093/embo-reports/kvf044

Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A., & Kornberg, R. D. (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. *Science, 292*(5523), 1876-1882. doi:10.1126/science.1059495

Goel, A., & Boland, C. R. (2010). Recent insights into the pathogenesis of colorectal cancer. *Curr Opin Gastroenterol, 26*(1), 47-52. doi:10.1097/MOG.0b013e328332b850

Goellner, E. M., Smith, C. E., Campbell, C. S., Hombauer, H., Desai, A., Putnam, C. D., & Kolodner, R. D. (2014). PCNA and Msh2-Msh6 activate an Mlh1-Pms1 endonuclease pathway required for Exo1-independent

mismatch repair. *Mol Cell, 55*(2), 291-304. doi:10.1016/j.molcel.2014.04.034
Goodman, M. F., & Woodgate, R. (2013). Translesion DNA polymerases. *Cold Spring Harb Perspect Biol, 5*(10), a010363. doi:10.1101/cshperspect.a010363
Gowrishankar, J., & Harinarayanan, R. (2004). Why is transcription coupled to translation in bacteria? *Mol Microbiol, 54*(3), 598-603. doi:10.1111/j.1365-2958.2004.04289.x
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*(2), 255-261.
Grenson, M. (1966). Multiplicity of the amino acid permeases in Saccharomyces cerevisiae. II. Evidence for a specific lysine-transporting system. *Biochim Biophys Acta, 127*(2), 339-346.
Grilley, M., Griffith, J., & Modrich, P. (1993). Bidirectional excision in methyl-directed mismatch repair. *J Biol Chem, 268*(16), 11830-11837.
Gulbis, J. M., Kelman, Z., Hurwitz, J., O'Donnell, M., & Kuriyan, J. (1996). Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA. *Cell, 87*(2), 297-306.
Guzder, S. N., Sommers, C. H., Prakash, L., & Prakash, S. (2006). Complex formation with damage recognition protein Rad14 is essential for Saccharomyces cerevisiae Rad1-Rad10 nuclease to perform its function in nucleotide excision repair in vivo. *Mol Cell Biol, 26*(3), 1135-1141. doi:10.1128/MCB.26.3.1135-1141.2006
Guzder, S. N., Sung, P., Prakash, L., & Prakash, S. (1996). Nucleotide excision repair in yeast is mediated by sequential assembly of repair factors and not by a pre-assembled repairosome. *J Biol Chem, 271*(15), 8903-8910.
Haber, J. E. (2012). Mating-type genes and MAT switching in Saccharomyces cerevisiae. *Genetics, 191*(1), 33-64. doi:10.1534/genetics.111.134577
Hampel, H., Frankel, W., Panescu, J., Lockman, J., Sotamaa, K., Fix, D., Comeras, I., La Jeunesse, J., Nakagawa, H., Westman, J. A., Prior, T. W., Clendenning, M., Penzone, P., Lombardi, J., Dunn, P., Cohn, D. E., Copeland, L., Eaton, L., Fowler, J., Lewandowski, G., Vaccarello, L., Bell, J., Reid, G., & de la Chapelle, A. (2006). Screening for Lynch syndrome (hereditary nonpolyposis colorectal cancer) among endometrial cancer patients. *Cancer Res, 66*(15), 7810-7817. doi:10.1158/0008-5472.CAN-06-1114
Hanawalt, P. C. (2001). Controlling the efficiency of excision repair. *Mutat Res, 485*(1), 3-13.
Hanawalt, P. C., & Spivak, G. (2008). Transcription-coupled DNA repair: two decades of progress and surprises. *Nat Rev Mol Cell Biol, 9*(12), 958-970. doi:10.1038/nrm2549
Hanway, D., Chin, J. K., Xia, G., Oshiro, G., Winzeler, E. A., & Romesberg, F. E. (2002). Previously uncharacterized genes in the UV- and MMS-induced DNA damage response in yeast. *Proc Natl Acad Sci U S A, 99*(16), 10605-10610. doi:10.1073/pnas.152264899
Harfe, B. D., & Jinks-Robertson, S. (2000). DNA mismatch repair and genetic instability. *Annu Rev Genet, 34*, 359-399. doi:10.1146/annurev.genet.34.1.359

Harris, R. S., Ross, K. J., Lombardo, M. J., & Rosenberg, S. M. (1998). Mismatch repair in Escherichia coli cells lacking single-strand exonucleases ExoI, ExoVII, and RecJ. *J Bacteriol, 180*(4), 989-993.
Heinze, R. J., Giron-Monzon, L., Solovyova, A., Elliot, S. L., Geisler, S., Cupples, C. G., Connolly, B. A., & Friedhoff, P. (2009). Physical and functional interactions between Escherichia coli MutL and the Vsr repair endonuclease. *Nucleic Acids Res, 37*(13), 4453-4463. doi:10.1093/nar/gkp380
Hingorani, M. M. (2016). Mismatch binding, ADP-ATP exchange and intramolecular signaling during mismatch repair. *DNA Repair (Amst), 38*, 24-31. doi:10.1016/j.dnarep.2015.11.017
Hombauer, H., Campbell, C. S., Smith, C. E., Desai, A., & Kolodner, R. D. (2011). Visualization of eukaryotic DNA mismatch repair reveals distinct recognition and repair intermediates. *Cell, 147*(5), 1040-1053. doi:10.1016/j.cell.2011.10.025
Hombauer, H., Srivatsan, A., Putnam, C. D., & Kolodner, R. D. (2011). Mismatch repair, but not heteroduplex rejection, is temporally coupled to DNA replication. *Science, 334*(6063), 1713-1716. doi:10.1126/science.1210770
Hsieh, P., & Yamane, K. (2008). DNA mismatch repair: molecular mechanism, cancer, and ageing. *Mech Ageing Dev, 129*(7-8), 391-407. doi:10.1016/j.mad.2008.02.012
Huang da, W., Sherman, B. T., & Lempicki, R. A. (2009). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res, 37*(1), 1-13. doi:10.1093/nar/gkn923
Huang, M. E., de Calignon, A., Nicolas, A., & Galibert, F. (2000). POL32, a subunit of the Saccharomyces cerevisiae DNA polymerase delta, defines a link between DNA replication and the mutagenic bypass repair pathway. *Curr Genet, 38*(4), 178-187.
Hughes, P., Tratner, I., Ducoux, M., Piard, K., & Baldacci, G. (1999). Isolation and identification of the third subunit of mammalian DNA polymerase delta by PCNA-affinity chromatography of mouse FM3A cell extracts. *Nucleic Acids Res, 27*(10), 2108-2114.
Ivessa, A. S., Lenzmeier, B. A., Bessler, J. B., Goudsouzian, L. K., Schnakenberg, S. L., & Zakian, V. A. (2003). The Saccharomyces cerevisiae helicase Rrm3p facilitates replication past nonhistone protein-DNA complexes. *Mol Cell, 12*(6), 1525-1536.
Ivessa, A. S., Zhou, J. Q., Schulz, V. P., Monson, E. K., & Zakian, V. A. (2002). Saccharomyces Rrm3p, a 5' to 3' DNA helicase that promotes replication fork progression through telomeric and subtelomeric DNA. *Genes Dev, 16*(11), 1383-1396. doi:10.1101/gad.982902
Ivessa, A. S., Zhou, J. Q., & Zakian, V. A. (2000). The Saccharomyces Pif1p DNA helicase and the highly related Rrm3p have opposite effects on replication fork progression in ribosomal DNA. *Cell, 100*(4), 479-489.
Iyer, R. R., Pluciennik, A., Burdett, V., & Modrich, P. L. (2006). DNA mismatch repair: functions and mechanisms. *Chem Rev, 106*(2), 302-323. doi:10.1021/cr0404794
Jagmohan-Changur, S., Poikonen, T., Vilkki, S., Launonen, V., Wikman, F., Orntoft, T. F., Moller, P., Vasen, H., Tops, C., Kolodner, R. D., Mecklin, J.

P., Jarvinen, H., Bevan, S., Houlston, R. S., Aaltonen, L. A., Fodde, R., Wijnen, J., & Karhu, A. (2003). EXO1 variants occur commonly in normal population: evidence against a role in hereditary nonpolyposis colorectal cancer. *Cancer Res, 63*(1), 154-158.

Jaruga, P., Theruvathu, J., Dizdaroglu, M., & Brooks, P. J. (2004). Complete release of (5'S)-8,5'-cyclo-2'-deoxyadenosine from dinucleotides, oligodeoxynucleotides and DNA, and direct comparison of its levels in cellular DNA with other oxidatively induced DNA lesions. *Nucleic Acids Res, 32*(11), e87. doi:10.1093/nar/gnh087

Javaid, S., Manohar, M., Punja, N., Mooney, A., Ottesen, J. J., Poirier, M. G., & Fishel, R. (2009). Nucleosome remodeling by hMSH2-hMSH6. *Mol Cell, 36*(6), 1086-1094. doi:10.1016/j.molcel.2009.12.010

Jinks-Robertson, S., & Bhagwat, A. S. (2014). Transcription-associated mutagenesis. *Annu Rev Genet, 48*, 341-359. doi:10.1146/annurev-genet-120213-092015

Jiricny, J. (2013). Postreplicative mismatch repair. *Cold Spring Harb Perspect Biol, 5*(4), a012633. doi:10.1101/cshperspect.a012633

Johansson, M. J., & Bystrom, A. S. (2002). Dual function of the tRNA(m(5)U54)methyltransferase in tRNA maturation. *RNA, 8*(3), 324-335.

Jones, A. L., Quimby, B. B., Hood, J. K., Ferrigno, P., Keshava, P. H., Silver, P. A., & Corbett, A. H. (2000). SAC3 may link nuclear protein export to cell cycle progression. *Proc Natl Acad Sci U S A, 97*(7), 3224-3229. doi:10.1073/pnas.050432997

Jordan, P. W., Klein, F., & Leach, D. R. (2007). Novel roles for selected genes in meiotic DNA processing. *PLoS Genet, 3*(12), e222. doi:10.1371/journal.pgen.0030222

Kadyrov, F. A., 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 U S A, 106*(21), 8495-8500. doi:10.1073/pnas.0903654106

Keil, R. L., & McWilliams, A. D. (1993). A gene with specific and global effects on recombination of sequences from tandemly repeated genes in Saccharomyces cerevisiae. *Genetics, 135*(3), 711-718.

Kelman, Z., & O'Donnell, M. (1995). Structural and functional similarities of prokaryotic and eukaryotic DNA polymerase sliding clamps. *Nucleic Acids Res, 23*(18), 3613-3620.

Kirkpatrick, D. T., & Petes, T. D. (1997). Repair of DNA loops involves DNA-mismatch and nucleotide-excision repair proteins. *Nature, 387*(6636), 929-931. doi:10.1038/43225

Klapacz, J., & Bhagwat, A. S. (2002). Transcription-dependent increase in multiple classes of base substitution mutations in Escherichia coli. *J Bacteriol, 184*(24), 6866-6872.

Klein, I. A., Resch, W., Jankovic, M., Oliveira, T., Yamane, A., Nakahashi, H., Di Virgilio, M., Bothmer, A., Nussenzweig, A., Robbiani, D. F., Casellas, R., & Nussenzweig, M. C. (2011). Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. *Cell, 147*(1), 95-106. doi:10.1016/j.cell.2011.07.048

Kohler, A., & Hurt, E. (2007). Exporting RNA from the nucleus to the cytoplasm. *Nat Rev Mol Cell Biol, 8*(10), 761-773. doi:10.1038/nrm2255
Kolodner, R. (1996). Biochemistry and genetics of eukaryotic mismatch repair. *Genes Dev, 10*(12), 1433-1442.
Krishna, T. S., Kong, X. P., Gary, S., Burgers, P. M., & Kuriyan, J. (1994). Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell, 79*(7), 1233-1243.
Kunkel, T. A., & Erie, D. A. (2005). DNA mismatch repair. *Annu Rev Biochem, 74*, 681-710. doi:10.1146/annurev.biochem.74.082803.133243
Kunkel, T. A., & Erie, D. A. (2015). Eukaryotic Mismatch Repair in Relation to DNA Replication. *Annu Rev Genet, 49*, 291-313. doi:10.1146/annurev-genet-112414-054722
Kuzmin, E., VanderSluis, B., Wang, W., Tan, G., Deshpande, R., Chen, Y., Usaj, M., Balint, A., Mattiazzi Usaj, M., van Leeuwen, J., Koch, E. N., Pons, C., Dagilis, A. J., Pryszlak, M., Wang, J. Z. Y., Hanchard, J., Riggi, M., Xu, K., Heydari, H., San Luis, B. J., Shuteriqi, E., Zhu, H., Van Dyk, N., Sharifpoor, S., Costanzo, M., Loewith, R., Caudy, A., Bolnick, D., Brown, G. W., Andrews, B. J., Boone, C., & Myers, C. L. (2018). Systematic analysis of complex genetic interactions. *Science, 360*(6386). doi:10.1126/science.aao1729
Laine, J. P., & Egly, J. M. (2006). When transcription and repair meet: a complex system. *Trends Genet, 22*(8), 430-436. doi:10.1016/j.tig.2006.06.006
Lambert, S., & Carr, A. M. (2013). Impediments to replication fork movement: stabilisation, reactivation and genome instability. *Chromosoma, 122*(1-2), 33-45. doi:10.1007/s00412-013-0398-9
Lang, G. I., & Murray, A. W. (2008). Estimating the per-base-pair mutation rate in the yeast Saccharomyces cerevisiae. *Genetics, 178*(1), 67-82. doi:10.1534/genetics.107.071506
Langle-Rouault, F., Maenhaut-Michel, G., & Radman, M. (1987). GATC sequences, DNA nicks and the MutH function in Escherichia coli mismatch repair. *EMBO J, 6*(4), 1121-1127.
Lea, D. E., & Coulson, C. A. (1949). The distribution of the numbers of mutants in bacterial populations. *J Genet, 49*(3), 264-285.
Leadon, S. A., & Avrutskaya, A. V. (1998). Requirement for DNA mismatch repair proteins in the transcription-coupled repair of thymine glycols in Saccharomyces cerevisiae. *Mutat Res, 407*(2), 177-187.
Lee, S. D., & Alani, E. (2006). Analysis of interactions between mismatch repair initiation factors and the replication processivity factor PCNA. *J Mol Biol, 355*(2), 175-184. doi:10.1016/j.jmb.2005.10.059
Lei, E. P., Stern, C. A., Fahrenkrog, B., Krebber, H., Moy, T. I., Aebi, U., & Silver, P. A. (2003). Sac3 is an mRNA export factor that localizes to cytoplasmic fibrils of nuclear pore complex. *Mol Biol Cell, 14*(3), 836-847. doi:10.1091/mbc.e02-08-0520
Li, F., Mao, G., Tong, D., Huang, J., Gu, L., Yang, W., & Li, G. M. (2013). The histone mark H3K36me3 regulates human DNA mismatch repair through its interaction with MutSalpha. *Cell, 153*(3), 590-600. doi:10.1016/j.cell.2013.03.025

Li, G. M. (2008). Mechanisms and functions of DNA mismatch repair. *Cell Res, 18*(1), 85-98. doi:10.1038/cr.2007.115

Li, S., Ding, B., Chen, R., Ruggiero, C., & Chen, X. (2006). Evidence that the transcription elongation function of Rpb9 is involved in transcription-coupled DNA repair in Saccharomyces cerevisiae. *Mol Cell Biol, 26*(24), 9430-9441. doi:10.1128/MCB.01656-06

Li, S., & Smerdon, M. J. (2002). Rpb4 and Rpb9 mediate subpathways of transcription-coupled DNA repair in Saccharomyces cerevisiae. *EMBO J, 21*(21), 5921-5929.

Li, X., & Manley, J. L. (2006). Cotranscriptional processes and their influence on genome stability. *Genes Dev, 20*(14), 1838-1847. doi:10.1101/gad.1438306

Liberfarb, R. M., & Bryson, V. (1970). Isolation, characterization, and genetic analysis of mutator genes in Escherichia coli B and K-12. *J Bacteriol, 104*(1), 363-375.

Liberti, S. E., Andersen, S. D., Wang, J., May, A., Miron, S., Perderiset, M., Keijzers, G., Nielsen, F. C., Charbonnier, J. B., Bohr, V. A., & Rasmussen, L. J. (2011). Bi-directional routing of DNA mismatch repair protein human exonuclease 1 to replication foci and DNA double strand breaks. *DNA Repair (Amst), 10*(1), 73-86. doi:10.1016/j.dnarep.2010.09.023

Liberti, S. E., Larrea, A. A., & Kunkel, T. A. (2013). Exonuclease 1 preferentially repairs mismatches generated by DNA polymerase alpha. *DNA Repair (Amst), 12*(2), 92-96. doi:10.1016/j.dnarep.2012.11.001

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature, 362*(6422), 709-715. doi:10.1038/362709a0

Lippert, M. J., Kim, N., Cho, J. E., Larson, R. P., Schoenly, N. E., O'Shea, S. H., & Jinks-Robertson, S. (2011). Role for topoisomerase 1 in transcription-associated mutagenesis in yeast. *Proc Natl Acad Sci U S A, 108*(2), 698-703. doi:10.1073/pnas.1012363108

Liu, L. F., & Wang, J. C. (1987). Supercoiling of the DNA template during transcription. *Proc Natl Acad Sci U S A, 84*(20), 7024-7027.

Loyola, A., & Almouzni, G. (2004). Histone chaperones, a supporting role in the limelight. *Biochim Biophys Acta, 1677*(1-3), 3-11. doi:10.1016/j.bbaexp.2003.09.012

Lujan, S. A., Clark, A. B., & Kunkel, T. A. (2015). Differences in genome-wide repeat sequence instability conferred by proofreading and mismatch repair defects. *Nucleic Acids Res, 43*(8), 4067-4074. doi:10.1093/nar/gkv271

Lujan, S. A., Williams, J. S., Clausen, A. R., Clark, A. B., & Kunkel, T. A. (2013). Ribonucleotides are signals for mismatch repair of leading-strand replication errors. *Mol Cell, 50*(3), 437-443. doi:10.1016/j.molcel.2013.03.017

Luria, S. E., & Delbruck, M. (1943). Mutations of Bacteria from Virus Sensitivity to Virus Resistance. *Genetics, 28*(6), 491-511.

Lynch, H. T., Lanspa, S., Smyrk, T., Boman, B., Watson, P., & Lynch, J. (1991). Hereditary nonpolyposis colorectal cancer (Lynch syndromes I & II). Genetics, pathology, natural history, and cancer control, Part I. *Cancer Genet Cytogenet, 53*(2), 143-160.

Markowitz, S., Wang, J., Myeroff, L., Parsons, R., Sun, L., Lutterbaugh, J., Fan, R. S., Zborowska, E., Kinzler, K. W., Vogelstein, B., & et al. (1995). Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability. *Science, 268*(5215), 1336-1338.

McCord, J. M., & Fridovich, I. (1969). Superoxide dismutase. An enzymic function for erythrocuprein (hemocuprein). *J Biol Chem, 244*(22), 6049-6055.

McCulloch, S. D., Gu, L., & Li, G. M. (2003). Bi-directional processing of DNA loops by mismatch repair-dependent and -independent pathways in human cells. *J Biol Chem, 278*(6), 3891-3896. doi:10.1074/jbc.M210687200

Mechanic, L. E., Frankel, B. A., & Matson, S. W. (2000). Escherichia coli MutL loads DNA helicase II onto DNA. *J Biol Chem, 275*(49), 38337-38346. doi:10.1074/jbc.M006268200

Medina-Arana, V., Delgado, L., Bravo, A., Martin, J., Fernandez-Peralta, A. M., & Gonzalez-Aguilera, J. J. (2012). Tumor spectrum in lynch syndrome, DNA mismatch repair system and endogenous carcinogens. *J Surg Oncol, 106*(1), 10-16. doi:10.1002/jso.23054

Mellon, I., Rajpal, D. K., Koi, M., Boland, C. R., & Champe, G. N. (1996). Transcription-coupled repair deficiency and mutations in human mismatch repair genes. *Science, 272*(5261), 557-560.

Mellon, I., Spivak, G., & Hanawalt, P. C. (1987). Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. *Cell, 51*(2), 241-249.

Mendillo, M. L., Mazur, D. J., & Kolodner, R. D. (2005). Analysis of the interaction between the Saccharomyces cerevisiae MSH2-MSH6 and MLH1-PMS1 complexes with DNA using a reversible DNA end-blocking system. *J Biol Chem, 280*(23), 22245-22257. doi:10.1074/jbc.M407545200

Mihaylova, V. T., Bindra, R. S., Yuan, J., Campisi, D., Narayanan, L., Jensen, R., Giordano, F., Johnson, R. S., Rockwell, S., & Glazer, P. M. (2003). Decreased expression of the DNA mismatch repair gene Mlh1 under hypoxic stress in mammalian cells. *Mol Cell Biol, 23*(9), 3265-3273.

Moarefi, I., Jeruzalmi, D., Turner, J., O'Donnell, M., & Kuriyan, J. (2000). Crystal structure of the DNA polymerase processivity factor of T4 bacteriophage. *J Mol Biol, 296*(5), 1215-1223. doi:10.1006/jmbi.1999.3511

Modrich, P. (1989). Methyl-directed DNA mismatch correction. *J Biol Chem, 264*(12), 6597-6600.

Modrich, P., & Lahue, R. (1996). Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu Rev Biochem, 65*, 101-133. doi:10.1146/annurev.bi.65.070196.000533

Mullenders, L. H., Vrieling, H., Venema, J., & van Zeeland, A. A. (1991). Hierarchies of DNA repair in mammalian cells: biological consequences. *Mutat Res, 250*(1-2), 223-228.

Ni, T. T., Marsischky, G. T., & Kolodner, R. D. (1999). MSH2 and MSH6 are required for removal of adenine misincorporated opposite 8-oxo-guanine in S. cerevisiae. *Mol Cell, 4*(3), 439-444.

Nick McElhinny, S. A., Kumar, D., Clark, A. B., Watt, D. L., Watts, B. E., Lundstrom, E. B., Johansson, E., Chabes, A., & Kunkel, T. A. (2010).

Genome instability due to ribonucleotide incorporation into DNA. *Nat Chem Biol, 6*(10), 774-781. doi:10.1038/nchembio.424
Nielsen, F. C., Jager, A. C., Lutzen, A., Bundgaard, J. R., & Rasmussen, L. J. (2004). Characterization of human exonuclease 1 in complex with mismatch repair proteins, subcellular localization and association with PCNA. *Oncogene, 23*(7), 1457-1468. doi:10.1038/sj.onc.1207265
Noll, D. M., Mason, T. M., & Miller, P. S. (2006). Formation and repair of interstrand cross-links in DNA. *Chem Rev, 106*(2), 277-301. doi:10.1021/cr040478b
Novick, P., Osmond, B. C., & Botstein, D. (1989). Suppressors of yeast actin mutations. *Genetics, 121*(4), 659-674.
Olinski, R., Gackowski, D., & Cooke, M. S. (2018). Endogenously generated DNA nucleobase modifications source, and significance as possible biomarkers of malignant transformation risk, and role in anticancer therapy. *Biochim Biophys Acta Rev Cancer, 1869*(1), 29-41. doi:10.1016/j.bbcan.2017.11.002
Owiti, N., Lopez, C., Singh, S., Stephenson, A., & Kim, N. (2017). Def1 and Dst1 play distinct roles in repair of AP lesions in highly transcribed genomic regions. *DNA Repair (Amst), 55*, 31-39. doi:10.1016/j.dnarep.2017.05.003
Paques, F., & Haber, J. E. (1997). Two pathways for removal of nonhomologous DNA ends during double-strand break repair in Saccharomyces cerevisiae. *Mol Cell Biol, 17*(11), 6765-6771.
Paulsen, R. D., & Cimprich, K. A. (2007). The ATR pathway: fine-tuning the fork. *DNA Repair (Amst), 6*(7), 953-966. doi:10.1016/j.dnarep.2007.02.015
Pavlov, Y. I., Mian, I. M., & Kunkel, T. A. (2003). Evidence for preferential mismatch repair of lagging strand DNA replication errors in yeast. *Curr Biol, 13*(9), 744-748.
Peltomaki, P. (2003). Role of DNA mismatch repair defects in the pathogenesis of human cancer. *J Clin Oncol, 21*(6), 1174-1179. doi:10.1200/JCO.2003.04.060
Phaniendra, A., Jestadi, D. B., & Periyasamy, L. (2015). Free radicals: properties, sources, targets, and their implication in various diseases. *Indian J Clin Biochem, 30*(1), 11-26. doi:10.1007/s12291-014-0446-0
Pluciennik, A., Dzantiev, L., Iyer, R. R., Constantin, N., Kadyrov, F. A., & Modrich, P. (2010). PCNA function in the activation and strand direction of MutLalpha endonuclease in mismatch repair. *Proc Natl Acad Sci U S A, 107*(37), 16066-16071. doi:10.1073/pnas.1010662107
Prado, F., Piruat, J. I., & Aguilera, A. (1997). Recombination between DNA repeats in yeast hpr1delta cells is linked to transcription elongation. *EMBO J, 16*(10), 2826-2835. doi:10.1093/emboj/16.10.2826
Prakash, S., & Prakash, L. (2000). Nucleotide excision repair in yeast. *Mutat Res, 451*(1-2), 13-24.
Qiu, J., Qian, Y., Frank, P., Wintersberger, U., & Shen, B. (1999). Saccharomyces cerevisiae RNase H(35) functions in RNA primer removal during lagging-strand DNA synthesis, most efficiently in cooperation with Rad27 nuclease. *Mol Cell Biol, 19*(12), 8361-8371.

Ranish, J. A., Hahn, S., Lu, Y., Yi, E. C., Li, X. J., Eng, J., & Aebersold, R. (2004). Identification of TFB5, a new component of general transcription and DNA repair factor IIH. *Nat Genet, 36*(7), 707-713. doi:10.1038/ng1385

Rasmussen, L. J., Heinen, C. D., Royer-Pokora, B., Drost, M., Tavtigian, S., Hofstra, R. M., & de Wind, N. (2012). Pathological assessment of mismatch repair gene variants in Lynch syndrome: past, present, and future. *Hum Mutat, 33*(12), 1617-1625. doi:10.1002/humu.22168

Reagan, M. S., & Friedberg, E. C. (1997). Recovery of RNA polymerase II synthesis following DNA damage in mutants of Saccharomyces cerevisiae defective in nucleotide excision repair. *Nucleic Acids Res, 25*(21), 4257-4263.

Reha-Krantz, L. J. (2010). DNA polymerase proofreading: Multiple roles maintain genome stability. *Biochim Biophys Acta, 1804*(5), 1049-1063. doi:10.1016/j.bbapap.2009.06.012

Repmann, S., Olivera-Harris, M., & Jiricny, J. (2015). Influence of oxidized purine processing on strand directionality of mismatch repair. *J Biol Chem, 290*(16), 9986-9999. doi:10.1074/jbc.M114.629907

Roberts, S. A., & Gordenin, D. A. (2014). Hypermutation in human cancer genomes: footprints and mechanisms. *Nat Rev Cancer, 14*(12), 786-800. doi:10.1038/nrc3816

Rodriguez, K., Talamantez, J., Huang, W., Reed, S. H., Wang, Z., Chen, L., Feaver, W. J., Friedberg, E. C., & Tomkinson, A. E. (1998). Affinity purification and partial characterization of a yeast multiprotein complex for nucleotide excision repair using histidine-tagged Rad14 protein. *J Biol Chem, 273*(51), 34180-34189.

Rustgi, A. K. (2007). The genetics of hereditary colon cancer. *Genes Dev, 21*(20), 2525-2538. doi:10.1101/gad.1593107

Sancar, A. (1996). DNA excision repair. *Annu Rev Biochem, 65*, 43-81. doi:10.1146/annurev.bi.65.070196.000355

Schiestl, R. H., & Prakash, S. (1988). RAD1, an excision repair gene of Saccharomyces cerevisiae, is also involved in recombination. *Mol Cell Biol, 8*(9), 3619-3626.

Schiestl, R. H., & Prakash, S. (1990). RAD10, an excision repair gene of Saccharomyces cerevisiae, is involved in the RAD1 pathway of mitotic recombination. *Mol Cell Biol, 10*(6), 2485-2491.

Schmidt, K. H., Derry, K. L., & Kolodner, R. D. (2002). Saccharomyces cerevisiae RRM3, a 5' to 3' DNA helicase, physically interacts with proliferating cell nuclear antigen. *J Biol Chem, 277*(47), 45331-45337. doi:10.1074/jbc.M207263200

Schmidt, K. H., Wu, J., & Kolodner, R. D. (2006). Control of translocations between highly diverged genes by Sgs1, the Saccharomyces cerevisiae homolog of the Bloom's syndrome protein. *Mol Cell Biol, 26*(14), 5406-5420. doi:10.1128/MCB.00161-06

Schmidt, T. T., Reyes, G., Gries, K., Ceylan, C. U., Sharma, S., Meurer, M., Knop, M., Chabes, A., & Hombauer, H. (2017). Alterations in cellular metabolism triggered by URA7 or GLN3 inactivation cause imbalanced dNTP pools and increased mutagenesis. *Proc Natl Acad Sci U S A, 114*(22), E4442-E4451. doi:10.1073/pnas.1618714114

Schopf, B., Bregenhorn, S., Quivy, J. P., Kadyrov, F. A., Almouzni, G., & Jiricny, J. (2012). Interplay between mismatch repair and chromatin assembly. *Proc Natl Acad Sci U S A, 109*(6), 1895-1900. doi:10.1073/pnas.1106696109

Shell, S. S., Putnam, C. D., & Kolodner, R. D. (2007). The N terminus of Saccharomyces cerevisiae Msh6 is an unstructured tether to PCNA. *Mol Cell, 26*(4), 565-578. doi:10.1016/j.molcel.2007.04.024

Shen, Y., Koh, K. D., Weiss, B., & Storici, F. (2011). Mispaired rNMPs in DNA are mutagenic and are targets of mismatch repair and RNases H. *Nat Struct Mol Biol, 19*(1), 98-104. doi:10.1038/nsmb.2176

Shia, J. (2008). Immunohistochemistry versus microsatellite instability testing for screening colorectal cancer patients at risk for hereditary nonpolyposis colorectal cancer syndrome. Part I. The utility of immunohistochemistry. *J Mol Diagn, 10*(4), 293-300. doi:10.2353/jmoldx.2008.080031

Shimodaira, H., Filosi, N., Shibata, H., Suzuki, T., Radice, P., Kanamaru, R., Friend, S. H., Kolodner, R. D., & Ishioka, C. (1998). Functional analysis of human MLH1 mutations in Saccharomyces cerevisiae. *Nat Genet, 19*(4), 384-389. doi:10.1038/1277

Shukla, P., Solanki, A., Ghosh, K., & Vundinti, B. R. (2013). DNA interstrand cross-link repair: understanding role of Fanconi anemia pathway and therapeutic implications. *Eur J Haematol, 91*(5), 381-393. doi:10.1111/ejh.12169

Smith, K. C. (1992). Spontaneous mutagenesis: experimental, genetic and other factors. *Mutat Res, 277*(2), 139-162.

Song, L., Yuan, F., & Zhang, Y. (2010). Does a helicase activity help mismatch repair in eukaryotes? *IUBMB Life, 62*(7), 548-553. doi:10.1002/iub.349

Soulas-Sprauel, P., Rivera-Munoz, P., Malivert, L., Le Guyader, G., Abramowski, V., Revy, P., & de Villartay, J. P. (2007). V(D)J and immunoglobulin class switch recombinations: a paradigm to study the regulation of DNA end-joining. *Oncogene, 26*(56), 7780-7791. doi:10.1038/sj.onc.1210875

Sparks, J. L., Chon, H., Cerritelli, S. M., Kunkel, T. A., Johansson, E., Crouch, R. J., & Burgers, P. M. (2012). RNase H2-initiated ribonucleotide excision repair. *Mol Cell, 47*(6), 980-986. doi:10.1016/j.molcel.2012.06.035

Stith, C. M., Sterling, J., Resnick, M. A., Gordenin, D. A., & Burgers, P. M. (2008). Flexibility of eukaryotic Okazaki fragment maturation through regulated strand displacement synthesis. *J Biol Chem, 283*(49), 34129-34140. doi:10.1074/jbc.M806668200

Stundon, J. L., & Zakian, V. A. (2015). Identification of Saccharomyces cerevisiae Genes Whose Deletion Causes Synthetic Effects in Cells with Reduced Levels of the Nuclear Pif1 DNA Helicase. *G3 (Bethesda), 5*(12), 2913-2918. doi:10.1534/g3.115.021139

Su, S. S., Lahue, R. S., Au, K. G., & Modrich, P. (1988). Mispair specificity of methyl-directed DNA mismatch correction in vitro. *J Biol Chem, 263*(14), 6829-6835.

Su, X. A., & Freudenreich, C. H. (2017). Cytosine deamination and base excision repair cause R-loop-induced CAG repeat fragility and instability

in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A, 114*(40), E8392-E8401. doi:10.1073/pnas.1711283114

Sugawara, N., Paques, F., Colaiacovo, M., & Haber, J. E. (1997). Role of Saccharomyces cerevisiae Msh2 and Msh3 repair proteins in double-strand break-induced recombination. *Proc Natl Acad Sci U S A, 94*(17), 9214-9219.

Svejstrup, J. Q., Wang, Z., Feaver, W. J., Wu, X., Bushnell, D. A., Donahue, T. F., Friedberg, E. C., & Kornberg, R. D. (1995). Different forms of TFIIH for transcription and DNA repair: holo-TFIIH and a nucleotide excision repairosome. *Cell, 80*(1), 21-28.

Svilar, D., Goellner, E. M., Almeida, K. H., & Sobol, R. W. (2011). Base excision repair and lesion-dependent subpathways for repair of oxidative DNA damage. *Antioxid Redox Signal, 14*(12), 2491-2507. doi:10.1089/ars.2010.3466

Sweder, K. S., Verhage, R. A., Crowley, D. J., Crouse, G. F., Brouwer, J., & Hanawalt, P. C. (1996). Mismatch repair mutants in yeast are not defective in transcription-coupled DNA repair of UV-induced DNA damage. *Genetics, 143*(3), 1127-1135.

Szankasi, P., & Smith, G. R. (1992). A DNA exonuclease induced during meiosis of Schizosaccharomyces pombe. *J Biol Chem, 267*(5), 3014-3023.

Szankasi, P., & Smith, G. R. (1995). A role for exonuclease I from S. pombe in mutation avoidance and mismatch correction. *Science, 267*(5201), 1166-1169.

Tham, K. C., Hermans, N., Winterwerp, H. H., Cox, M. M., Wyman, C., Kanaar, R., & Lebbink, J. H. (2013). Mismatch repair inhibits homeologous recombination via coordinated directional unwinding of trapped DNA structures. *Mol Cell, 51*(3), 326-337. doi:10.1016/j.molcel.2013.07.008

Thompson, E., Meldrum, C. J., Crooks, R., McPhillips, M., Thomas, L., Spigelman, A. D., & Scott, R. J. (2004). Hereditary non-polyposis colorectal cancer and the role of hPMS2 and hEXO1 mutations. *Clin Genet, 65*(3), 215-225.

Toft, N. J., Winton, D. J., Kelly, J., Howard, L. A., Dekker, M., te Riele, H., Arends, M. J., Wyllie, A. H., Margison, G. P., & Clarke, A. R. (1999). Msh2 status modulates both apoptosis and mutation frequency in the murine small intestine. *Proc Natl Acad Sci U S A, 96*(7), 3911-3915.

Tran, P. T., Erdeniz, N., Symington, L. S., & Liskay, R. M. (2004). EXO1-A multi-tasking eukaryotic nuclease. *DNA Repair (Amst), 3*(12), 1549-1559. doi:10.1016/j.dnarep.2004.05.015

Tran, P. T., Simon, J. A., & Liskay, R. M. (2001). Interactions of Exo1p with components of MutLalpha in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A, 98*(17), 9760-9765. doi:10.1073/pnas.161175998

Umar, A., Boland, C. R., Terdiman, J. P., Syngal, S., de la Chapelle, A., Ruschoff, J., Fishel, R., Lindor, N. M., Burgart, L. J., Hamelin, R., Hamilton, S. R., Hiatt, R. A., Jass, J., Lindblom, A., Lynch, H. T., Peltomaki, P., Ramsey, S. D., Rodriguez-Bigas, M. A., Vasen, H. F., Hawk, E. T., Barrett, J. C., Freedman, A. N., & Srivastava, S. (2004). Revised Bethesda Guidelines for hereditary nonpolyposis colorectal

cancer (Lynch syndrome) and microsatellite instability. *J Natl Cancer Inst, 96*(4), 261-268.

van Gool, A. J., Verhage, R., Swagemakers, S. M., van de Putte, P., Brouwer, J., Troelstra, C., Bootsma, D., & Hoeijmakers, J. H. (1994). RAD26, the functional S. cerevisiae homolog of the Cockayne syndrome B gene ERCC6. *EMBO J, 13*(22), 5361-5369.

Vasen, H. F., Watson, P., Mecklin, J. P., & Lynch, H. T. (1999). New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC. *Gastroenterology, 116*(6), 1453-1456.

Verhage, R. A., Van de Putte, P., & Brouwer, J. (1996). Repair of rDNA in Saccharomyces cerevisiae: RAD4-independent strand-specific nucleotide excision repair of RNA polymerase I transcribed genes. *Nucleic Acids Res, 24*(6), 1020-1025.

Viswanathan, M., & Lovett, S. T. (1998). Single-strand DNA-specific exonucleases in Escherichia coli. Roles in repair and mutation avoidance. *Genetics, 149*(1), 7-16.

Vo, A. T., Zhu, F., Wu, X., Yuan, F., Gao, Y., Gu, L., Li, G. M., Lee, T. H., & Her, C. (2005). hMRE11 deficiency leads to microsatellite instability and defective DNA mismatch repair. *EMBO Rep, 6*(5), 438-444. doi:10.1038/sj.embor.7400392

Volker, M., Mone, M. J., Karmakar, P., van Hoffen, A., Schul, W., Vermeulen, W., Hoeijmakers, J. H., van Driel, R., van Zeeland, A. A., & Mullenders, L. H. (2001). Sequential assembly of the nucleotide excision repair factors in vivo. *Mol Cell, 8*(1), 213-224.

Waris, G., & Ahsan, H. (2006). Reactive oxygen species: role in the development of cancer and various chronic conditions. *J Carcinog, 5*, 14. doi:10.1186/1477-3163-5-14

Wei, K., Clark, A. B., Wong, E., Kane, M. F., Mazur, D. J., Parris, T., Kolas, N. K., Russell, R., Hou, H., Jr., Kneitz, B., Yang, G., Kunkel, T. A., Kolodner, R. D., Cohen, P. E., & Edelmann, W. (2003). Inactivation of Exonuclease 1 in mice results in DNA mismatch repair defects, increased cancer susceptibility, and male and female sterility. *Genes Dev, 17*(5), 603-614. doi:10.1101/gad.1060603

Wen, Q., Scorah, J., Phear, G., Rodgers, G., Rodgers, S., & Meuth, M. (2008). A mutant allele of MRE11 found in mismatch repair-deficient tumor cells suppresses the cellular response to DNA replication fork stress in a dominant negative manner. *Mol Biol Cell, 19*(4), 1693-1705. doi:10.1091/mbc.E07-09-0975

Wilson, M. A., Kwon, Y., Xu, Y., Chung, W. H., Chi, P., Niu, H., Mayle, R., Chen, X., Malkova, A., Sung, P., & Ira, G. (2013). Pif1 helicase and Poldelta promote recombination-coupled DNA synthesis via bubble migration. *Nature, 502*(7471), 393-396. doi:10.1038/nature12585

Woudstra, E. C., Gilbert, C., Fellows, J., Jansen, L., Brouwer, J., Erdjument-Bromage, H., Tempst, P., & Svejstrup, J. Q. (2002). A Rad26-Def1 complex coordinates repair and RNA pol II proteolysis in response to DNA damage. *Nature, 415*(6874), 929-933. doi:10.1038/415929a

Wu, Q., Christensen, L. A., Legerski, R. J., & Vasquez, K. M. (2005). Mismatch repair participates in error-free processing of DNA interstrand crosslinks

in human cells. *EMBO Rep, 6*(6), 551-557. doi:10.1038/sj.embor.7400418

Zhang, N., Lu, X., Zhang, X., Peterson, C. A., & Legerski, R. J. (2002). hMutSbeta is required for the recognition and uncoupling of psoralen interstrand cross-links in vitro. *Mol Cell Biol, 22*(7), 2388-2397.

Zhang, Y., Yuan, F., Presnell, S. R., Tian, K., Gao, Y., Tomkinson, A. E., Gu, L., & Li, G. M. (2005). Reconstitution of 5'-directed human mismatch repair in a purified system. *Cell, 122*(5), 693-705. doi:10.1016/j.cell.2005.06.027

Zou, L., & Elledge, S. J. (2003). Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science, 300*(5625), 1542-1548. doi:10.1126/science.1083430

9 Appendix

Table 9.1 Oligos used in this study

Oligo number in Hoffmann collection	Oligo name	Sequence	Used for
O0073	K2	TTCAGAAACAACTCTGGCGCA	Reverse primer 592 bp into the ORF of *KANMX6* cassette. to verify Integration of *KANMX6*
O0074	K3	CATCCTATGGAACTGCCTCGG	Forward primer ~400bp from the end of the ORF of *KANMX6* cassette. to verify Integration of *KANMX6*
O0151	N2	GATTCGTCGTCCGATTCGTC	Reverse primer 640 bp into the ORF of *NATMX4* cassette to verify Integration of *NATMX4* cassette
O0152	N3	AGGTCACCAACGTCAACGCA	Forward primer ~400bp from the end of the ORF of *NATMX4* cassette. to verify Integration of *NATMX4*
O0628	PMS1_MX4.F	CGAAAAGAAAAGACGCGTCTCTCTTAATAATCATTATGCGATAAAcgtacgctgcaggtcgac	forward primer anneals upstream of *PMS1* start codon, used for amplification of cassette to replace *PMS1*
O0629	PMS1_MX4.R	ATAATGTATTTGTTAATTATATAATGAATGAATATCAAAGCTAGAatcgatgaattcgagctcg	Reverse primer anneals downstream of *PMS1,* used for amplification of cassette to replace *PMS1*
O0349	PMS1_A	GCAGTTTCCATCAGCTATTTATGTT	Forward primer anneals ~400bp upstream of *PMS1.* to check integration or for amplifying *PMS1*
O0350	PMS1_B	GCAGTTGTTAAGTCGGTGATAACTT	Reverse primer anneals ~400bp downstream of start site for *PMS1.* to check integration
O0351	PMS1_C	CTTTTTGATTTGGGTGATTTTAATG	Forward primer anneals ~400bp upstream of end of *PMS1 ORF* to check integration
O0352	PMS1_D	GCTTAATAAAATCGATAGACGTGGA	Reverse primer anneals ~400bp downstream of *PMS1.* to check integration or for amplifying *PMS1*
O0626	MSH2_MX4.F	CCCATCATCTCGGTTTGAGGAACAGACGCCTTTTCATAGTTTTGGcgtacgctgcaggtcgac	forward primer anneals upstream of *MSH2* start codon, used for amplification of cassette to replace *MSH2*
O0627	MSH2_MX4.R	TATGAAGTAATCTATTGTGCTGAGTGGTGATAGTGCACCCGATCAatcgatgaattcgagctcg	Reverse primer anneals downstream of *MSH2* used for amplification of cassette to replace *MSH2*
O0353	MSH2_A	CGTATAAACAAAGCCAAAGACAAGT	Forward primer anneals ~400bp upstream of *MSH2* to check integration or for amplifying *MSH2*
O0354	MSH2_B	CCCAATTGAATCAAGAAACTCTCTA	Reverse primer anneals ~400bp downstream of start site for MSH2. to check integration

Table 9.1 continued Oligos used in this study			
O0355	MSH2_C	TGAATTGACAGAATTGTCTGAAAAA	Forward primer anneals ~400bp upstream of end of *MSH2* ORF to check integration
O0356	MSH2_D	ACATCTCTTGTTTATCCCATCCATA	Reverse primer anneals ~400bp downstream of *MSH2* to check integration or for amplifying *MSH2*
Oligo number in Hoffmann collection	**Oligo name**	**Sequence**	**Used for**
O0624	MLH1_MX4.F	GATAGTAAATGGAAGGTAAAAATAACATAGACCTATCAATAAG CAcgtacgctgcaggtcgac	forward primer anneals upstream of *MLH1* start codon, used for amplification of cassette to replace *MLH1*
O0625	MLH1_MX4.R	GAAATAAACAAAAAACTTTGGTATTACAGCCAAAACGTTTTAA AGatcgatgaattcgagctcg	Reverse primer anneals downstream of *MLH1* used for amplification of cassette to replace *MLH1*
O0585	MLH1_A	TATAGTTGAGGAGTTCTCGAAGACG	Forward primer anneals ~400bp upstream of *MLH1* to check integration
O0358	MLH1_B	GTATTCATCATTATGGGACCTCAAG	Reverse primer anneals ~400bp downstream of start site for *MLH1* to check integration
O0359	MLH1_C	CGCAAACTTTGGTAAGATAAACCTA	Forward primer anneals ~400bp upstream of end of *MLH1* ORF to check integration
O0586	MLH1_D	TTTCCTAACTGCAACCATATTTCAT	Reverse primer anneals ~400bp downstream of *MLH1* to check integration or for amplifying *MLH1*
O0789	TRM2_A	ATCAAGACGAAACAATTAGAACCTG	Forward primer anneals ~400bp upstream of *TRM2*.. to check integration or for amplifying *TRM2*
O0790	TRM2_D	TTTTCCATTTTGGTATGATGTTTCT	Reverse primer anneals ~400bp downstream of *TRM2*. to check integration or for amplifying *TRM2*
O0791	TRM2_F	ATGCTCCCAGAGAGCCTACA	forward primer anneals upstream of *TRM2* start codon, used for amplification of cassette to replace *TRM2*
O0792	TRM2_R	TTGAACAGATGCGTGGTCAT	Reverse primer anneals downstream of *TRM2* used for amplification of cassette to replace *TRM2*
O1358	TRM2_MXF	AATCTGTCATTTTATTTTAGAGGAATAGTTTAGGACAAAGTCAT TCGTACGCTGCAGGTC	forward primer anneals upstream of *TRM2* start codon, used for amplification of cassette to replace *TRM2*
O1359	TRM2_MXR	TACTCTAGAAAGATATACATAGTGATAGATATTTTATATGTGCA AATCGATGAATTCGAG	Reverse primer anneals downstream of *TRM2* used for amplification of cassette to replace *TRM2*
O0793	TEL1_A	CACATGATATTATGAGCGTGATAGG	Forward primer anneals ~400bp upstream of *TEL1.* to check integration or for amplifying *TEL1.*

Table 9.1 continued Oligos used in this study			
O0794	TEL1_D	ATCTACGTCGATTTCTTTCATTTTG	Reverse primer anneals ~400bp downstream of *TEL1.* to check integration or for amplifying *TEL1.*
O0805	RTT107_A	ACTTAACCACAGAATGTTCTTCGAC	Forward primer anneals ~400bp upstream of *RTT107.* to check integration or for amplifying *RTT107*
O0806	RTT107_D	TTGTAGAAAAATTAAAAGGTTTGCG	Reverse primer anneals ~400bp downstream of *RTT107*. to check integration or for amplifying *RTT107*
O0809	RRM3_A	TATCTTCCCTTACCGGATTTATTTC	Forward primer anneals ~400bp upstream of *RRM3*. to check integration or for amplifying *RRM3*
O0810	RRM3_D	GGCTAGATCTCCTTTTTCAGTTTCT	Reverse primer anneals ~400bp downstream of *RRM3.* to check integration or for amplifying *RRM3*
O1356	RRM3_MXF	GAACAAGCTCAAAAGTCGAGAGATTTGTTCTTATAAGACATCC CGCGTACGCTGCAGGTC	forward primer anneals upstream of *RRM3* start codon, used for amplification of cassette to replace *RRM3*
O1357	RRM3_MXR	CTTGCAACGAATAAATGCATATACTCTAGTTGAAGTTTTCTTTT CATCGATGAATTCGAG	Reverse primer anneals downstream of *RRM3* used for amplification of cassette to replace *RRM3*
O0813	RNR4_A	TTCAATGTTCCTAAAGTTTCATTCC	Forward primer anneals ~400bp upstream of *RNR4*. to check integration or for amplifying *RNR4*
O0814	RNR4_D	TTACCTCATCCATTGTTGCTATGTA	Reverse primer anneals ~400bp downstream of *RNR4* to check integration or for amplifying *RNR4*
O2041	TFB5_A	CTGTAGTCATTAATGTCTAATAGTAGG	Forward primer anneals ~400bp upstream of *MSH2* to check integration or for amplifying *MSH2*
O2042	TFB5_B	AGATGGGTGTCATCCAACTCCTCTA	Reverse primer anneals ~400bp downstream of start site for TFB5 to check integration
O2043	TFB5_C	CATCTTTTAGTAAATCCCTCGAAAG	Forward primer anneals ~400bp upstream of end of *TFB5* orf to check integration
O2044	TFB5_D	AGAAAAAATTAAACGGGATTTGGCT	Reverse primer anneals ~400bp downstream of *MSH2* to check integration or for amplifying *MSH2*
O2047	THO2_A	ATTTACATGTTCTGAATGAGAAGGC	Forward primer anneals ~400bp upstream of *THO2* to check integration or for amplifying *THO2*
O2048	THO2_B	GGGAGAATTTTCATCGTTTTTATTT	Reverse primer anneals ~400bp downstream of start site for THO2 to check integration
O2049	THO2_C	CGAATAAAAGATTCAAGAAGGATGA	Forward primer anneals ~400bp upstream of end of *THO2* orf to check integration
O2050	THO2_D	CGATAAAAGAAAGAACGGTTTGTTA	Reverse primer anneals ~400bp downstream of *THO2* to check integration or for amplifying *THO2*

Table 9.1 continued Oligos used in this study			
O2053	DEF1_A	TTACCTTTTTAGGATATCCGTCACC	Forward primer anneals ~400bp upstream of *DEF1* to check integration or for amplifying *DEF1*
O2054	DEF1_B	CCCTTGTGCTTCTAATTCTTCATTA	Reverse primer anneals ~400bp downstream of start site for *DEF1* to check integration
Oligo number in Hoffmann collection	**Oligo name**	**Sequence**	**Used for**
O2055	DEF1_C	GCTGGTCAATATCCATACCAGTTAC	Forward primer anneals ~400bp upstream of end of *DEF1* orf to check integration
O2056	DEF1_D	CTCCCACGTCTATACCAATGTACTC	Reverse primer anneals ~400bp downstream of *DEF1* to check integration or for amplifying *DEF1*
O2059	SAC3_A	AGAATTCGTGATTCAATGTTCATTAT	Forward primer anneals ~400bp upstream of *SAC3* to check integration or for amplifying *SAC3*
O2060	SAC3_B	CTTTATCCACTAGCCCTTTTTCTTC	Reverse primer anneals ~400bp downstream of start site for *SAC3* to check integration
O2061	SAC3_C	ATACAGTCCCAACAATTAATCGAAA	Forward primer anneals ~400bp upstream of end of *SAC3* orf to check integration
O2062	SAC3_D	AGCAGTTTATAACTTTTTGGCCTCT	Reverse primer anneals ~400bp downstream of *SAC3* to check integration or for amplifying *SAC3*
O2174	Mlh1_E31_Check	TGTAGCATTCGCATCGATGGAATTCTCCATCATTg	to check E31A substitution

Table 9.2 Yeast strains used in this study

strain number	strain background	genotype	mating type
Y937	S288C	his3Δ1 leu2Δ0 met15Δ0 ura3Δ0	Mat a
Y938	S288C	his3Δ1 leu2Δ0 met15Δ0 ura3Δ0	MATα
Y5785	S288C	hpr1::KANMX4, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y5786	S288C	rad26::KANMX4, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y5787	S288C	rpb9::KANMX4, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y5788	S288C	rad10::KANMX4, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y5789	S288C	pol1-L868M.natNT2, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y5790	S288C	pol2-04.natNT2, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y5791	S288C	pol2-M644G.natNT2, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y5792	S288C	pol3-L612M.natNT2, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y5793	S288C	mlh1::HphMX4, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y5794	S288C	pol3-L612M.natNT2, mlh1::HphMX4,, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y5795	S288C	pol2-04.natNT2, mlh1::HphMX4,, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y5796	S288C	pol2-04.natNT2, mlh1::HphMX4,, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y5797	S288C	pol1-L868M.natNT2, mlh1::HphMX4, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y5798	S288C	pol1-L868M.natNT2 ,mlh1::HphMX4,, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y5799	S288C	pol3-L612M.natNT2, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y5800	S288C	pol1-L868M.natNT2,, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y5719	S288C	tfb5Δ:KanMX4, mlh1Δ:NATMX4, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y5720	S288C	sac3Δ:KanMX4,mlh1Δ:NATMX4, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y4845	S288C	mlh1::natMX4,, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, cyhR	Mat a
Y714	S288c derived	his3, leu2-112, lys2-BglII CEN5-lacO-LEU2, pCYC1-LacI-GFP-HIS3, ilv1 (S.cerevisiae chromosome V)	MATα
Y5893	S288c derived	msh2::natMX4, cyhR, his3, leu2-112, lys2-BglII CEN5-lacO-LEU2, pCYC1-LacI-GFP-HIS3, ilv1 (S.cerevisiae chromosome V)	MATα

Table 9.2 continued Yeast strains used in this study			
Y5894	S288c derived	pms1::natMX4, cyhR, his3, leu2-112, lys2-BglII CEN5-lacO-LEU2, pCYC1-LacI-GFP-HIS3 ilv1 (S.cerevisiae chromosome V)	MATα
Y5895	S288c derived	mlh1::natMX4, cyhR,his3, leu2-112, lys2-BglII CEN5-lacO-LEU2, pCYC1-LacI-GFP-HIS3, ilv1 (S.cerevisiae chromosome V)	MATα
Y2195	S288C	rnr4::kanMX4, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y2197	S288C	ogg1::kanMX4,his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y2198	S288C	elg1::kanMX4,his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y2200	S288C	exo1::kanMX4,his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y2201	S288C	mgs1::kanMX4,his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y2202	S288C	mms2::kanMX4,his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y2203	S288C	psy3::kanMX4,his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y2204	S288C	rad27::kanMX4,his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y2205	S288C	rad34::kanMX4,his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y2206	S288C	rad5::kanMX4,his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y2207	S288C	rnh201::kanMX4,his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y2208	S288C	trm2::kanMX4,his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y2209	S288C	rrm3::kanMX4,his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y2210	S288C	rtt107::kanMX4,his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y2211	S288C	shu1::kanMX4,his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y2212	S288C	slx4::kanMX4,his3Δ1, leu2Δ0, met15Δ0, ura3Δ0	Mat a
Y5716	W303	Mlh1::NatMX4 delete, RAD5 corrected, his3-11, ade1::ADE2, lys14A, ura3Δ0	Mat α
Y5717	W303	tfb5:KanMX4,, RAD5 corrected, his3-11, ade1::ADE2, lys14A, ura3Δ0	Mat a
Y5718	W303	sac3:KanMX4,, , RAD5 corrected,, his3-11, ade1::ADE2, lys14A, ura3Δ0	Mat a
Y5719	W303	tfb5:KanMX4, mlh1:NATMX4, RAD5 corrected, his3-11, ade1::ADE2, lys14A, ura3Δ0	Mat a
Y5720	W303	sac3:KanMX4,mlh1:NATMX4, RAD5 corrected,, his3-11, ade1::ADE2, lys14A, ura3Δ0	Mat a

Table 9.3 Biological process associated with increased mutation rate from figure 3.7

	Count	list %	genome%	description according to AmiGO2 site
DNA repair	16	12,5	3,437106352	The process of restoring DNA after damage. Genomes are subject to damage by chemical and physical agents in the environment (e.g. UV and ionizing radiations, chemical mutagens, fungal and bacterial toxins, etc.) and by free radicals or alkylating agents endogenously generated in metabolism. DNA is also damaged because of errors during its replication. A variety of different DNA repair pathways have been reported that include direct reversal, base excision repair, nucleotide excision repair, photoreactivation, bypass, double-strand break repair pathway, and mismatch repair pathway. Source: PMID:11563486
cellular response to DNA damage stimulus	15	11,71875	3,707036171	Any process that results in a change in state or activity of a cell (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a stimulus indicating damage to its DNA from environmental insults or errors during metabolism. *Source:* GOC:go_curators
actin filament organization	5	3,90625	0,665826885	A process that is carried out at the cellular level which results in the assembly, arrangement of constituent parts, or disassembly of cytoskeletal structures comprising actin filaments. Includes processes that control the spatial distribution of actin filaments, such as organizing filaments into meshworks, bundles, or other structures, as by cross-linking. Source: GOC:mah
protein localization to pre-autophagosomal structure	3	2,34375	0,161957891	Any process in which a protein is transported to, or maintained at, the phagophore assembly site (PAS). Source: GOC:rb
mismatch repair	4	3,125	0,44988303	A system for the correction of errors in which an incorrect base, which cannot form hydrogen bonds with the corresponding base in the parent strand, is incorporated into the daughter strand. The mismatch repair system promotes genomic fidelity by repairing base-base mismatches, insertion-deletion loops and heterologies generated during DNA replication and recombination. Source: ISBN:0198506732, PMID:11687886
cytoplasm to vacuole targeting pathway	5	3,90625	0,88177074	A cytoplasm to vacuole targeting pathway that uses machinery common with autophagy. The Cvt vesicle is formed when the receptor protein, Atg19, binds to the complexes of the target protein (aminopeptidase or alpha-mannosidase homododecamers), forming the Cvt complex. Atg11 binds to Atg9 and transports the Cvt complex to the pre-autophagosome (PAS). The phagophore membrane expands around the Cvt complex (excluding bulk cytoplasm) forming the Cvt vesicle. This pathway is mostly observed in yeast. Source: PMID:15659643,PMID:12865942
fungal-type cell wall organization	8	6,25	2,51934497	A process that is carried out at the cellular level which results in the assembly, arrangement of constituent parts, or disassembly of the fungal-type cell wall. Source: GOC:dph, GOC:jl, GOC:mtg_sensu, GOC:mah
regulation of DNA double-strand break processing	2	1,5625	0,035990642	Any process that modulates the frequency, rate or extent of DNA double-strand break processing. Source: GOC:TermGenie, GO_REF:0000058, PMID:25203555
double-strand break repair	4	3,125	0,665826885	The repair of double-strand breaks in DNA via homologous and nonhomologous mechanisms to reform a continuous DNA helix. Source: GOC:elh
telomere maintenance	4	3,125	0,701817527	Any process that contributes to the maintenance of proper telomeric length and structure by affecting and monitoring the activity of telomeric proteins, the length of telomeric DNA and the replication and repair of the DNA. These processes includes those that shorten, lengthen, replicate and repair the telomeric DNA sequences. Source: GOC:BHF_telomere, PMID:11092831, GOC:BHF, GOC:elh, GOC:rl
metabolic process	12	9,375	5,110671225	The chemical reactions and pathways, including anabolism and catabolism, by which living organisms transform chemical substances. Metabolic processes typically transform small molecules, but also include macromolecular processes such as DNA repair and replication, and protein synthesis and degradation. Source: ISBN:0198547684, GOC:go_curators
regulation of DNA damage checkpoint	2	1,5625	0,053985964	Any process that modulates the frequency, rate or extent of a DNA damage checkpoint. Source: GOC:obol
chronological cell aging	3	2,34375	0,377901746	The process associated with progression of the cell from its inception to the end of its lifespan that occurs when the cell is in a non-dividing, or quiescent, state. Source: GOC:jh, PMID:12044934
cellular glucan metabolic process	2	1,5625	0,071981285	The chemical reactions and pathways involving glucans, polysaccharides consisting only of glucose residues, occurring at the level of an individual cell. Source: ISBN:0198547684
glucan catabolic process	2	1,5625	0,071981285	The chemical reactions and pathways resulting in the breakdown of glucans, polysaccharides consisting only of glucose residues. Source: GOC:go_curators
arginine transmembrane transport	2	1,5625	0,071981285	The directed movement of arginine across a membrane. Source: GOC:TermGenie, GO_REF:0000069, PMID:18357653
cell cycle arrest	2	1,5625	0,071981285	A regulatory process that halts progression through the cell cycle during one of the normal phases (G1, S, G2, M). Source: GOC:dph, GOC:tb, GOC:mah
chitin catabolic process	2	1,5625	0,071981285	The chemical reactions and pathways resulting in the breakdown of chitin, a linear polysaccharide consisting of beta-(1->4)-linked N-acetyl-D-glucosamine residues. Source: ISBN:0198506732, GOC:jl

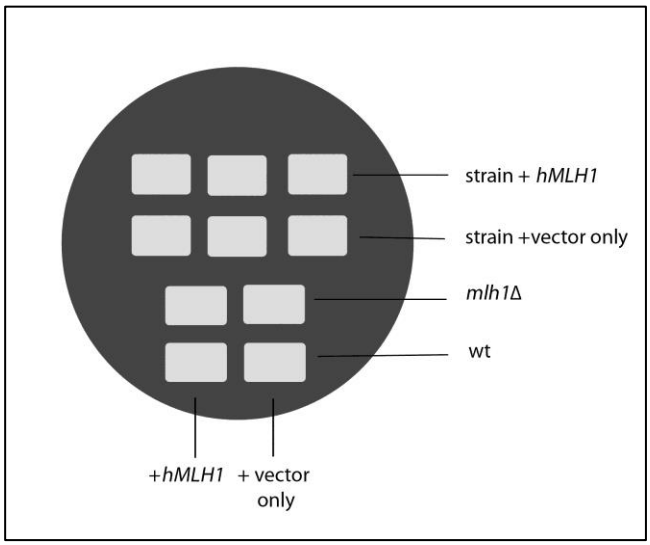
Table 9.4 processess enriched in genes causing lower mutation rate.

Term	Count	list %	genome %	description according to AmiGO2 site
ubiquitin-dependent protein catabolic process	8	1,81	0,27	The chemical reactions and pathways resulting in the breakdown of a protein or peptide covalently tagged with ubiquitin, via the multivesicular body (MVB) sorting pathway; ubiquitin-tagged proteins are sorted into MVBs, and delivered to a lysosome/vacuole for degradation. Source:PMID:11511343, GOC:jl
arginine biosynthetic process	7	1,59	0,20	The chemical reactions and pathways resulting in the formation of arginine, 2-amino-5-(carbamimidamido)pentanoic acid. *Source:* CHEBI:29016, ISBN:0198506732
ATP export	8	1,81	0,31	The directed movement of ATP out of a cell or organelle. Source: GO_REF:0000074, GOC:TermGenie, PMID:24286344
late endosome to vacuole transport	8	1,81	0,34	The directed movement of substances from late endosomes to the vacuole. In yeast, after transport to the prevacuolar compartment, endocytic content is delivered to the late endosome and on to the vacuole. This pathway is analogous to endosome to lysosome transport. Source:PMID:11872141
protein targeting to vacuole	13	2,95	0,97	The process of directing proteins towards the vacuole, usually using signals contained within the protein. Source: GOC:curators
intralumenal vesicle formation	5	1,13	0,13	The invagination of the endosome membrane and resulting formation of a vesicle within the lumen of the endosome. Source: PMID:19234443, GOC:jp
DNA-templated transcription, termination	6	1,36	0,23	The cellular process that completes DNA-templated transcription; the formation of phosphodiester bonds ceases, the RNA-DNA hybrid dissociates, and RNA polymerase releases the DNA. Source: PMID:15020047, PMID:18280161, GOC:txnOH, ISBN:0716720094
positive regulation of transcription elongation from RNA polymerase I promoter	5	1,13	0,16	Any process that activates or increases the frequency, rate or extent of transcription elongation from RNA polymerase I promoter. Source: PMID:20299458
regulation of translation	11	2,49	0,88	Any process that modulates the frequency, rate or extent of the chemical reactions and pathways resulting in the formation of proteins by the translation of mRNA or circRNA. Source: GOC:isa_complete
fatty acid biosynthetic process	7	1,59	0,41	Any process that modulates the frequency, rate or extent of cellular DNA-templated transcription. Source: GOC:go_curators, GOC:txnOH
regulation of transcription, DNA-templated	56	12,70	9,16	Any process that modulates the frequency, rate or extent of cellular DNA-templated transcription. Source: GOC:go_curators, GOC:txnOH
protein maturation	4	0,91	0,13	Any process leading to the attainment of the full functional capacity of a protein. Source: GOC:ai
regulation of transcription-coupled nucleotide-excision repair	4	0,91	0,13	Any process that modulates the frequency, rate, or extent of the nucleotide-excision repair process that carries out preferential repair of DNA lesions on the actively transcribed strand of the DNA duplex. In addition, the transcription-coupled nucleotide-excision repair pathway is required for the recognition and repair of a small subset of lesions that are not recognized by the global genome nucleotide excision repair pathway. Source: GOC:tb
transcription, DNA-templated	56	12,70	9,48	The cellular synthesis of RNA on a template of DNA. Source: GOC:jl, GOC:txnOH
regulation of histone H3-K4 methylation	3	0,68	0,05	Any process that modulates the frequency, rate or extent of the covalent addition of a methyl group to the lysine at position 4 of histone H3. Source: GOC:ai
cellular amino acid biosynthetic process	15	3,40	1,75	The chemical reactions and pathways resulting in the formation of amino acids, organic acids containing one or more amino substituents. Source: ISBN:0198506732
protein transport	43	9,75	7,04	The directed movement of proteins into, out of or within a cell, or between cells, by means of some agent such as a transporter or pore. Source: GOC:ai
mitochondrion inheritance	7	1,59	0,54	The distribution of mitochondria, including the mitochondrial genome, into daughter cells after mitosis or meiosis, mediated by interactions between mitochondria and the cytoskeleton. Source: GOC:mcc, PMID:11389764, PMID:10873824
meiotic nuclear division	8	1,81	0,70	Progression through the phases of the meiotic cell cycle, in which canonically a cell replicates to produce four offspring with half the chromosomal content of the progenitor cell via two nuclear divisions. Source: GOC:ai

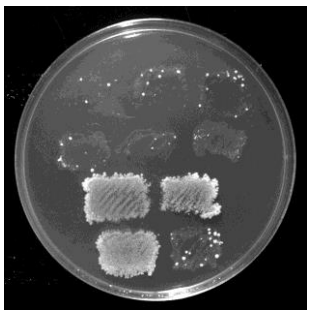
snoRNA transcription from an RNA polymerase II promoter	3	0,68	0,07	the synthesis of small nucleolar RNA (snoRNA) from a DNA template by RNA polymerase II, originating at an RNA polymerase II promoter. Source: GOC:txnOH
positive regulation of phosphorylation of RNA polymerase II C-terminal domain serine 2 residues	3	0,68	0,07	Any process that activates or increases the frequency, rate or extent of phosphorylation of RNA polymerase II C-terminal domain serine 2 residues. Source: PMID:15149594
positive regulation of DNA-templated transcription, elongation	3	0,68	0,07	Any process that activates or increases the frequency, rate or extent of transcription elongation, the extension of an RNA molecule after transcription initiation and promoter clearance by the addition of ribonucleotides catalyzed by a DNA-dependent RNA polymerase. Source:GOC:mah, GOC:txnOH
double-strand break repair via homologous recombination	6	1,36	0,43	The error-free repair of a double-strand break in DNA in which the broken DNA molecule is repaired using homologous sequences. A strand in the broken DNA searches for a homologous region in an intact chromosome to serve as the template for DNA synthesis. The restoration of two intact DNA molecules results in the exchange, reciprocal or nonreciprocal, of genetic material between the intact DNA molecule and the broken DNA molecule. Source: GOC:elh, PMID:10357855
positive regulation of transcription from RNA polymerase I promoter	5	1,13	0,31	Any process that activates or increases the frequency, rate or extent of transcription mediated by RNA polymerase I. Source: GOC:go_curators, GOC:txnOH
translation	35	7,94	5,72	The cellular metabolic process in which a protein is formed, using the sequence of a mature mRNA or circRNA molecule to specify the sequence of amino acids in a polypeptide chain. Translation is mediated by the ribosome, and begins with the formation of a ternary complex between aminoacylated initiator methionine tRNA, GTP, and initiation factor 2, which subsequently associates with the small subunit of the ribosome and an mRNA or circRNA. Translation ends with the release of a polypeptide chain from the ribosome. Source: GOC:go_curators
vacuolar transport	6	1,36	0,45	The directed movement of substances into, out of or within a vacuole. Source: GOC:ai
proteasome assembly	5	1,13	0,32	The aggregation, arrangement and bonding together of a mature, active proteasome complex. Source: PMID:10872471, GOC:go_curators
regulation of histone H2B conserved C-terminal lysine ubiquitination	3	0,68	0,09	Any process that modulates the frequency, rate or extent of histone H2B conserved C-terminal lysine ubiquitination. Source: PMID:17576814
histone modification	3	0,68	0,09	The covalent alteration of one or more amino acid residues within a histone protein. Source: GOC:krc
transcription elongation from RNA polymerase II promoter	9	2,04	0,97	The extension of an RNA molecule after transcription initiation and promoter clearance at an RNA polymerase II promoter by the addition of ribonucleotides catalyzed by RNA polymerase II. Source: GOC:mah, GOC:txnOH
positive regulation of transcription from RNA polymerase II promoter in response to heat stress	4	0,91	0,22	Any process that increases the frequency, rate or extent of transcription from an RNA polymerase II promoter as a result of a heat stimulus, a temperature stimulus above the optimal temperature for that organism. Source: GOC:dph
mRNA export from nucleus	7	1,59	0,67	The directed movement of mRNA from the nucleus to the cytoplasm. Source: GOC:ma
endocytosis	14	3,17	1,89	A vesicle-mediated transport process in which cells take up external materials or membrane constituents by the invagination of a small region of the plasma membrane to form a new membrane-bounded vesicle. Source: ISBN:0716731363, ISBN:0198506732, GOC:mah
intracellular protein transport	13	2,95	1,73	The directed movement of proteins in a cell, including the movement of proteins between specific compartments or structures within a cell, such as organelles of a eukaryotic cell. Source: GOC:mah
histone deubiquitination	3	0,68	0,11	The modification of histones by removal of ubiquitin groups. Source: GOC:ai
cellular sphingolipid homeostasis	3	0,68	0,11	Any biological process involved in the maintenance of an internal steady state of sphingolipids at the level of the cell. Source: GOC:dph, GOC:tb, GOC:ascb_2009
urea cycle	3	0,68	0,11	The sequence of reactions by which arginine is synthesized from ornithine, then cleaved to yield urea and regenerate ornithine. The overall reaction equation is NH3 + CO2 + aspartate + 3 ATP + 2 H2O = urea + fumarate + 2 ADP + 2 phosphate + AMP + diphosphate. Source:ISBN:0198506732, GOC:vw, GOC:pde
meiotic DNA double-strand break formation	4	0,91	0,23	The cell cycle process in which double-strand breaks are generated at defined hotspots throughout the genome during meiosis I. This results in the initiation of meiotic recombination. Source: PMID:11529427, GOC:jl, GOC:elh
mRNA 3'-end processing	5	1,13	0,38	Any process involved in forming the mature 3' end of an mRNA molecule. Source: GOC:mah

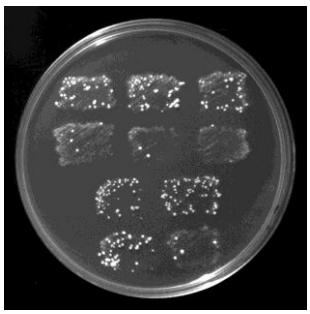
actin cytoskeleton organization	6	1,36	0,54	A process that is carried out at the cellular level which results in the assembly, arrangement of constituent parts, or disassembly of cytoskeletal structures comprising actin filaments and their associated proteins. Source: GOC:dph, GOC:jl, GOC:mah
ubiquitin-dependent protein catabolic process	10	2,27	1,22	The chemical reactions and pathways resulting in the breakdown of a protein or peptide by hydrolysis of its peptide bonds, initiated by the covalent attachment of a ubiquitin group, or multiple ubiquitin groups, to the protein. Source: GOC:go_curators
mitochondrial translation	15	3,40	2,16	The chemical reactions and pathways resulting in the formation of a protein in a mitochondrion. This is a ribosome-mediated process in which the information in messenger RNA (mRNA) is used to specify the sequence of amino acids in the protein; the mitochondrion has its own ribosomes and transfer RNAs, and uses a genetic code that differs from the nuclear code. Source: GOC:go_curators
lipid metabolic process	18	4,08	2,74	The chemical reactions and pathways involving lipids, compounds soluble in an organic solvent but not, or sparingly, in an aqueous solvent. Includes fatty acids; neutral fats, other fatty-acid esters, and soaps; long-chain (fatty) alcohols and waxes; sphingoids and other long-chain bases; glycolipids, phospholipids and sphingolipids; and carotenes, polyprenols, sterols, terpenes and other isoprenoids. *Source:* GOC:ma

Figure 9.1: Qualitative rrepresentation of CAN1 mutation rate assay for genes associated with replication and repair.
For each gene of interest deletion, three independent transformants with *hMLH1* or vector only were patched on canavanine containing media and cells were compared for mutation rate with WT



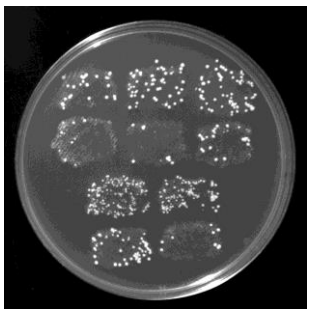
*rnr4*Δ (after 5 days of incubation)

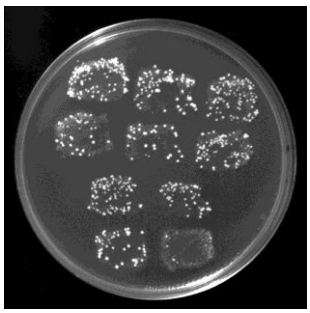




*slx4*Δ

*rad34*Δ

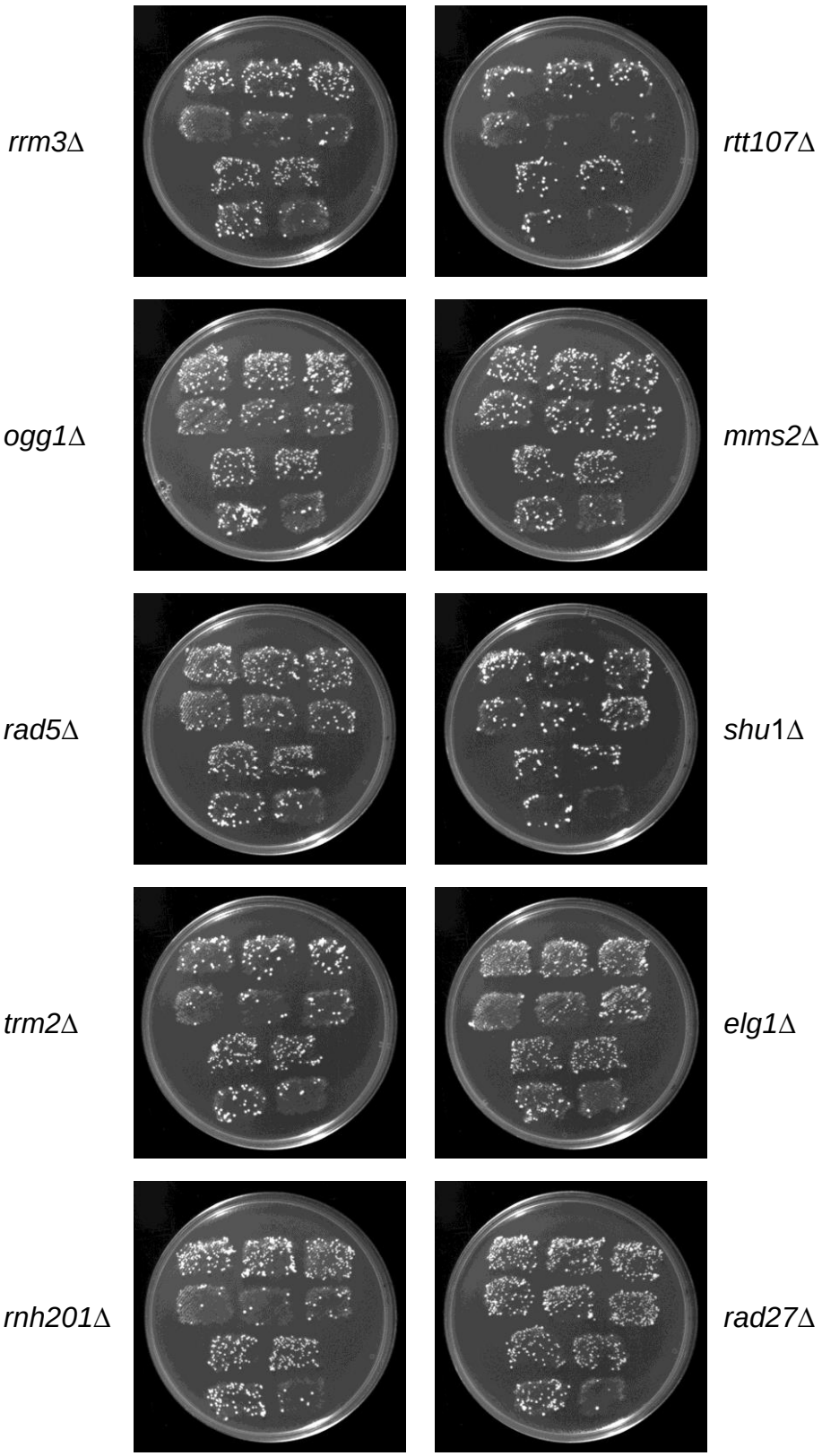




*psy3*Δ

*mgs1*Δ

*tel1*Δ


rrm3Δ
rtt107Δ
ogg1Δ
mms2Δ
rad5Δ
shu1Δ
trm2Δ
elg1Δ
rnh201Δ
rad27Δ