Chapter 4

Polymerase mutations in mammalian systems and in combination with other mutations

Overarching hypothesis

DNA polymerase mutations exert their mutagenic function in a manner separate from exonuclease deficiency and mutation accumulation increases identfied in *Saccharomyces cerevisiae* can be confirmed in mammalian systems.

Aims:

- To determine whether polymerase mutations and mismatch repair deficiency act synergistically when present in the same cell.
- To examine whether the mutagenesis observed in mutant strains is independent of Polε exonuclease activity.
- To determine the mutagenesis observed in mutant strains may be due to increased involvement of a translesion polymerase in DNA replication.
- To investigate candidate polymerase mutations identified as mutators in *Saccharomyces cerevisiae* in mammalian systems.

4.1 Introduction

In the previous chapter, candidate polymerase mutations were assessed for their mutagenic potential. The exonuclease domain mutations *pol2-P301R, pol2-S312F, pol2-L439V, pol2-M459K* or *pol3-S483N* were all shown to significantly increase the number of acquired mutations when in heterozygosis. Strikingly, the *pol2*-4 exonuclease deficient allele does not confer a comparably strong increase in mutation numbers. To get a better understanding of how these candidate polymerase mutations lead to more mutagenesis, in this chapter a series of double mutant experiments is described.

Additionally, while the budding yeast *Saccharomyces cerevisiae* was chosen to investigate the genome-wide effects of these mutations in a cost and time effective manner in an organism that has been shown to yield insights into DNA replication and repair that shows remarkable conservation, it remains to validate these in mammalian systems. Thus, this chapter also describes the introduction of the *pol2-L439V* and *pol3-S483N* mutations into a mouse model. These two mutations were identified in the germline of families with a predisposition to colorectal cancer and the whole organism aspect of the mouse model may give insights into the tumorigenesis, which a single-celled organism cannot provide. The *pol2-P301R* mutation, resulting in similarly striking mutation increases as *pol3-S483N* and one of the most common DNA polymerase alterations identified in human cancers, is introduced into a human cell line to show that the insights gained in yeast can be translated into a human system, while acuired at a fraction of the cost.

4.2 Synthetic lethality with mismatch repair deficiency

Given the differences between the mutational patterns observed in polymerase mutant cells and mismatch repair deficient cells and the fact that both are promoting tumor progression and predispose to colorectal cancer, the question is how the genome is affected when both fidelity systems are impaired simultaneously.

To obtain double mutants, a MAT α strain deleted for *MSH2 (msh2*Δ) was created (see Chapter 6.6) and mated to haploid MATa cells: *pol2-P301R, pol2-S312F, pol2-A480V* and *pol2-M459K* (see Chapter 6.6). The strains now heterozygous diploid for *msh2*Δ and the polymerase mutation were kept on sporulation medium to undergo meiosis. Tetrads were issolated and dissected. After two days of growth tetrads were replicated onto selection plates selecting for the polymerase mutation (ura- plates), the *msh2*Δ mutation (G418 plates) and plates selecting for either mating type. This allowed confirmation that tetrads were haploid

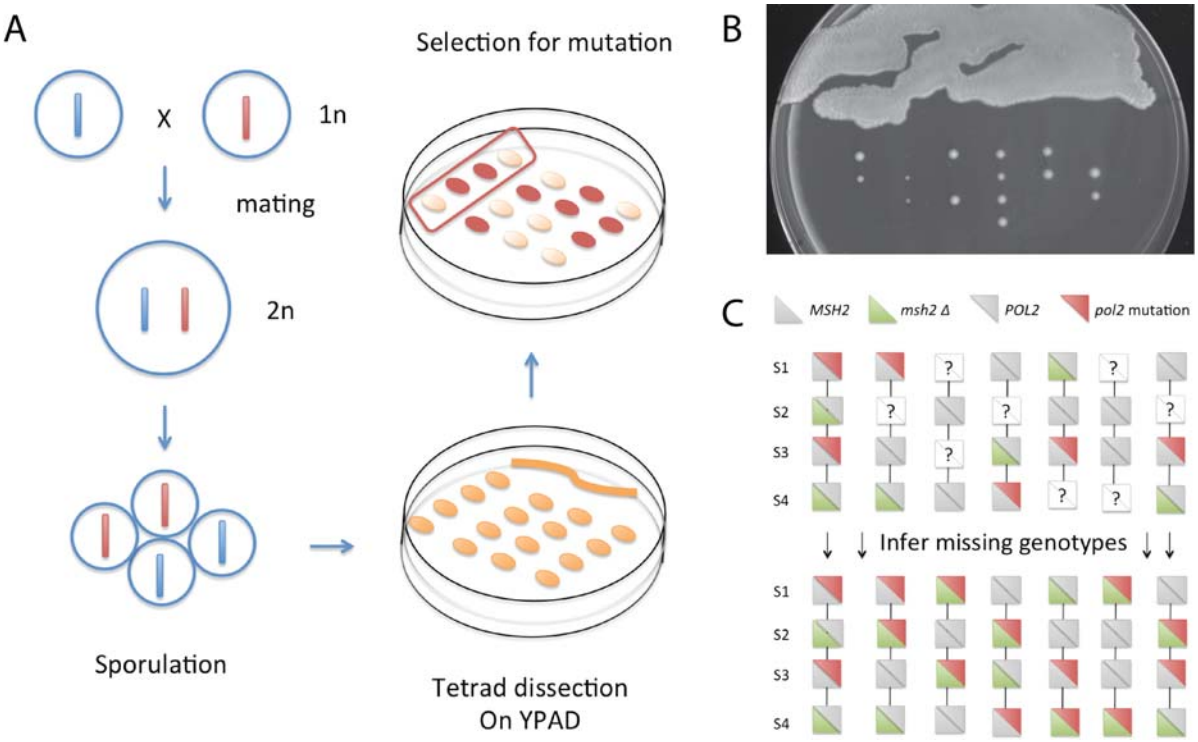

Figure 4.1: Tetrad dissection to generate double mutants and detect synthetic lethality **A|** For tetrad dissection, two haploid single mutant strains of opposite mating types are mated and the resulting diploid forced to undergo meiosis resulting in a tetrad. Alternatively, an already existent diploid can be used. Replication followed by segregation results in each allele (here depicted as a "blue" wild-type and a "red" mutant allele) being present in two out of the four tetrads. This is mutually exclusive (for instance a tetrad spore will under regular circumstances contain either the "red" or the "blue" allele, not both or neither as that would result in aneuploidy). Using a micromanipulator tetrads are dissected on rich medium plates (YPAD) and allowed to germinate and expand to colony size. Testing growth in selective conditions, can reveal mating type or, in cases where a mutant allele is marked by for instance an antibiotic resistance, which spores contain the mutant allele. Double mutants can be generate analogously, by mating two single mutant haploid strains. If both mutatant alleles are unlinked (on different chromosomes), they should be randomly assorted during meiosis, resulting in wild-type, single mutant and double mutant cells, which can be identified by tetrad analysis. **B|** An example, of a tetrad dissection of a mating of *msh2*Δ and *pol2-P301R*. When not all four spores result in a colony, microscopy can be used to confirm the spores germinated, but ceased to divide after a few division cycles. Should all missing colonies have been double mutants, synthetic lethality is a likely conclusion. **C|** If not all spores reached full colony size (denoted by "?"), then their genotype might be inferred by using selection to determine the genotype of the remaining colonies. Since every allele (wild-type and mutant) should occur twice, the genotype of the failed colonies can be determined. In the case of synthetic lethality, double mutants will not be among the full-sized colonies and most if not all failed colonies have an inferred double mutant genotype.

(either MATa or MATα) and whether they carried one or both mutations. Among all strains I sporulated, I observed spores that germinated, but didn't grow to form a colony. Across a tetrad, if the two mutations are not linked, each mutation should be observed twice, randomly assorted (Fig. 4.1). As Fig. 4.1 illustrates that means if three meiotic products grew to a colony and one germinated and underwent extinction, the genotype of the extinct cells can be inferred. Should these cells be double mutants and should no living double mutants be recovered, a case for synthetic lethality can be made. Thus, the number of double mutants we should recover was determined from genotype information and the number of double mutants actually recovered was obtained (Table 4.1). As can be seen, *pol2-P301R, pol2-S312F* and *pol2-M459K* are likely synthetic lethal with *msh2*Δ in a haploid background. *pol2-A480V msh2*Δ double mutants can be viable, but grow visibly slower on rich medium plates.

Since of all four polymerase mutants, the one causing the lowest increase in mutation number (Fig. 3.1) is the only one to allow for a double mutant with *msh2*Δ in a haploid context, the data is in line with the notion of a threshold for lethal mutagenesis. This postulates that a certain elevation of mutation rate will overwhelm the population and cause extinction. If this is the explanation, then predictions about the viablity of other polymerase mutants in combination with mismatch repair deficiency can be made. For instance *pol2-Q468R msh2*Δ double mutant should be viable, a *pol3-S483N msh2*Δ double mutant should not be and the viability of *pol2-04 msh2*Δ and *pol2-L439V msh2*Δ double mutants is uncertain. This will be the next experiments carried out to further underpin the relationship between polymerase mutants and mismatch repair deficiency.

4.3 Epistatic relationship of mutations with exonuclease deficiency

To gain further understanding of why some candidate *pol2* and *pol3* mutant strains acquire more mutations than the *pol2*-4 and *pol3-01* exonuclease deficienct strains, the relationship between *pol2* candidate mutations and the exonuclease function was further explored. *Pol2* candidate mutations (*pol2-P301R, pol2-S312F, pol2-A480V, pol2-M459K)* were combined with the mutations in the strain (*pol2-D290A,E292A*) in the same gene, to construct MATa haploid yeasts trains each carrying three point mutations in the *POL2* gene. Wild-type, single mutant and combined strains were propagated using single colony bottlenecks for 13 passages (1.5 months) in 18 parallel lines each. Strains were sequenced and analysed for acquired SNVs and INDELs as in Chapter 2.4.2 and 3.2. While there is no apparent difference be-

polymerase mutant	tetrads dissected	double mutants expected	double mutants recovered
pol2-A480V	4	4	3*
pol2-M459K	8	10	0
pol2-P301R	10	14	0
pol2-S312F	15	14	0

Table 4.1: Synthetic lethality of polymerase mutants and mismatch repair deficiency
Diploid strains heterozygous for both *msh2*Δ and a polymerase mutant (see first column) were sporulated and dissected after meiosis allowing the recovery of all four meiotic products. The number of tetrads dissected is the number of tetrads where all four meiotic products germinated. Replica-plating on selective medium plates allowed the identification of double mutants. The number of "double mutants expected" is the number of double mutants we expect to obtain from all tetrads considering the genotypes in the tetrad, the number of "double mutants recovered" is the number of double mutants that actually made it from germination to colony. A left-tailed 2X2 Fisher's Exact Test was performed to determine the significance of the negative association between the polymerase mutations and mismatch repair deficiency.
* Three double mutants were recovered. However, they were smaller than all single mutant and wild-type strains.

tween the number of SNVs acquired by *pol2-A480V* samples and the number acquired by *pol2-D290A,E292A,A480V* samples, a significant difference between the number of SNVs acquired by *pol2-P301R, pol2-S312F, pol2-M459K* samples and their *pol2-D290A,E292A* combined counterparts is observed (Fig. 4.2). In each of those cases the number of acquired mutations is much less when the *pol2* candidate point mutation is combined with the other two point mutations, however, the numbers still exceed those acquired by a *pol2*-4 strain, suggesting that the observed mutagenesis is at least in part dependent on the exonuclease function of the protein.

4.4 Observed mutagenesis in *pol2-P301R* strains is not due to increased participation of Polζ in DNA replication

To investigate the high number of mutations in polymerase mutant strains (*pol2-P301R, pol2-S312F, pol2-L439V, pol2-M459K* and *pol3-S483N*) compared to the *pol2-4* and *pol3-01* exonuclease deficient mutant strains, the possibility of the involvement of another polymerase was explored. Recent reports indicate that in cases of replisome instability polymerase ζ, unique in its ability to extend primers with a terminal mismatch, can take over more of the replication burden and account for some of the mutagenesis[883, 884]. It was described that cells carrying the *dpb2-100* mutant allele have decreased interaction with the CMG (Cdc45-

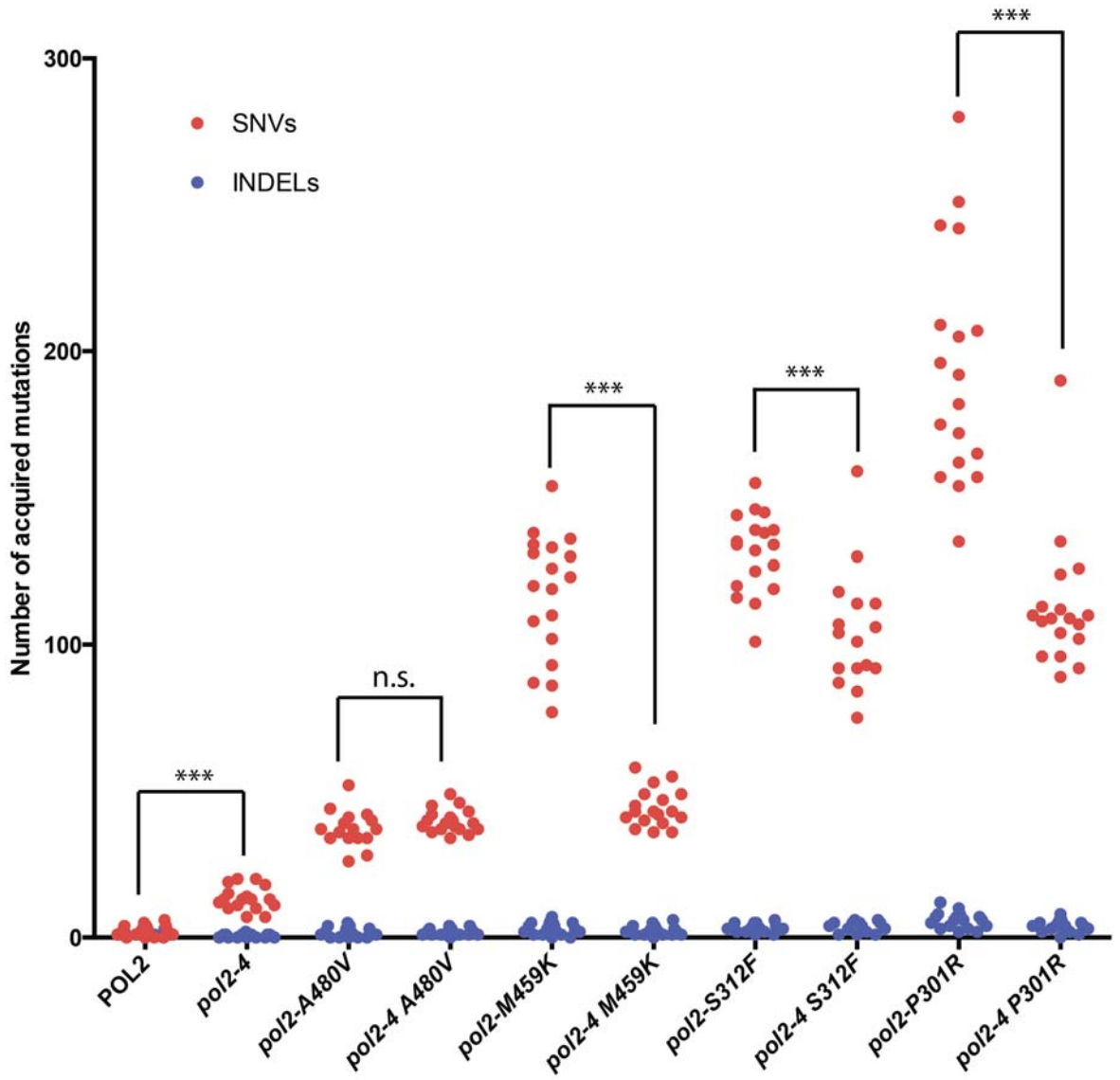

Figure 4.2: The mutagenesis observed in strong mutator strains is partially rescued by mutating critical residues in the exonuclease domain active site
Combinations of candidate *pol2* mutations and the *pol2-D290A,E292A* mutations in the same gene were achieved by site directed mutagenesis of the plasmids used to construct the strains. When haploid strains expressing Polε with all three missense mutations were obtained, they were propagated alongside haploid single mutants for 13 passages using single-colony bottlenecks. Samples were sequenced before and after propagation and acquired single nucleotide variants (SNVs) and INDELs were determined as detailed in Chapter 2. The number of parallel lines are as follows: n=18 (POL2, *pol2-S312F*, *pol2-P301R*, *pol2-4 A480V*, *pol2-4 M459K*, *pol2-4 P301R*), n=17 (*pol2-A480V*, *pol2-4*) and n=16 (*pol2-4 S312F*). Significance of the difference in number of SNVs was determined by two-tailed, unpaired t-Tests. n.s. $P > 0.05$; *** $P \leq 0.001$

MCM-GINS) complex, which is thought to bring the polymerase ε into the replisome. Polymerase ζ, recently found to also interact with polymerase δ subunits Pol31 and Pol32[885, 886], has been proposed to be the fourth polymerase in the replisome and is expected to take over replication when the main replicase has a problem with primer extension. Deletion of the catalytic subunit of Polζ, *REV3*, substantially decreases the mutator phenotype found in *dpb2-100* cells, but has little effect on the mutator phenotype observed in *pol2-4* cells[883].

To investigate whether this could partly explain the apparent divergence between *pol2-4* mutant strains and the other polymerase mutations, whose presence results in higher mutation numbers than exonuclease deficiency, a *rev3*Δ strain was constructed and double mutants of polymerase mutations with *rev3*Δ were generated by mating and tetrad dissection (see Chapter 4.2). The resulting haploid double mutant strains, the single mutant precursor strains and a wild-type control were propagated in 18 parallel lines using single-colony bottlenecks for 13 passages (1.5 months) on non-selective rich medium. Starting and final colonies were sent for whole-genome sequencing and analysed for single nucleotide variants (SNVs) and small insertions/deletions (INDELs) as before (see Chapter 3.2).

As expected, the muttaion numbers observed in the *rev3*Δ strains is lower than those in wild-type strains with a mean number of mutations of 1.89 compared to 3.55 in the wild-type (Fig. 4.3). However, unlike the work with *dpb2-100* cells[883], where a double mutant with *rev3*Δ decreased mutation numbers, leading them to conclude that Polζ is responsible for a substantial part of the mutagenesis in *dpb2-100* cells, here no such drop in mutation numbers can be observed. While there is only small increases in mutation number when one compares *pol2-A480V* strains with the corresponding *pol2-A480V rev3*Δ double mutant (a mean of 38.6 versus a mean of 49.1 mutations) and the *pol2-S312F* with the corresponding *pol2-S312F rev3*Δ cells (a mean of 134.6 versus a mean of 161.6 mutations), the mutation increase oberserved in *pol2-P301R* strains upon combination with a deletion of *REV3* is striking: *pol2-*P301R cells acquire a mean of 116.8 mutations, while *pol2-P301R rev3*Δ cells acquire a mean of 347.1 mutations in the same time frame.

To get an understanding of the process at work here, the mutation patterns of the SNVs acquired in *pol2-P301R* cells and *pol2-P301R rev3*Δ cells were compared (Fig. 4.4). Even though the number of mutations acquired by the *pol2-P301R rev3*Δ cells is more than double that of the *pol2-P301R* cells, there are no striking differences between the two patterns.

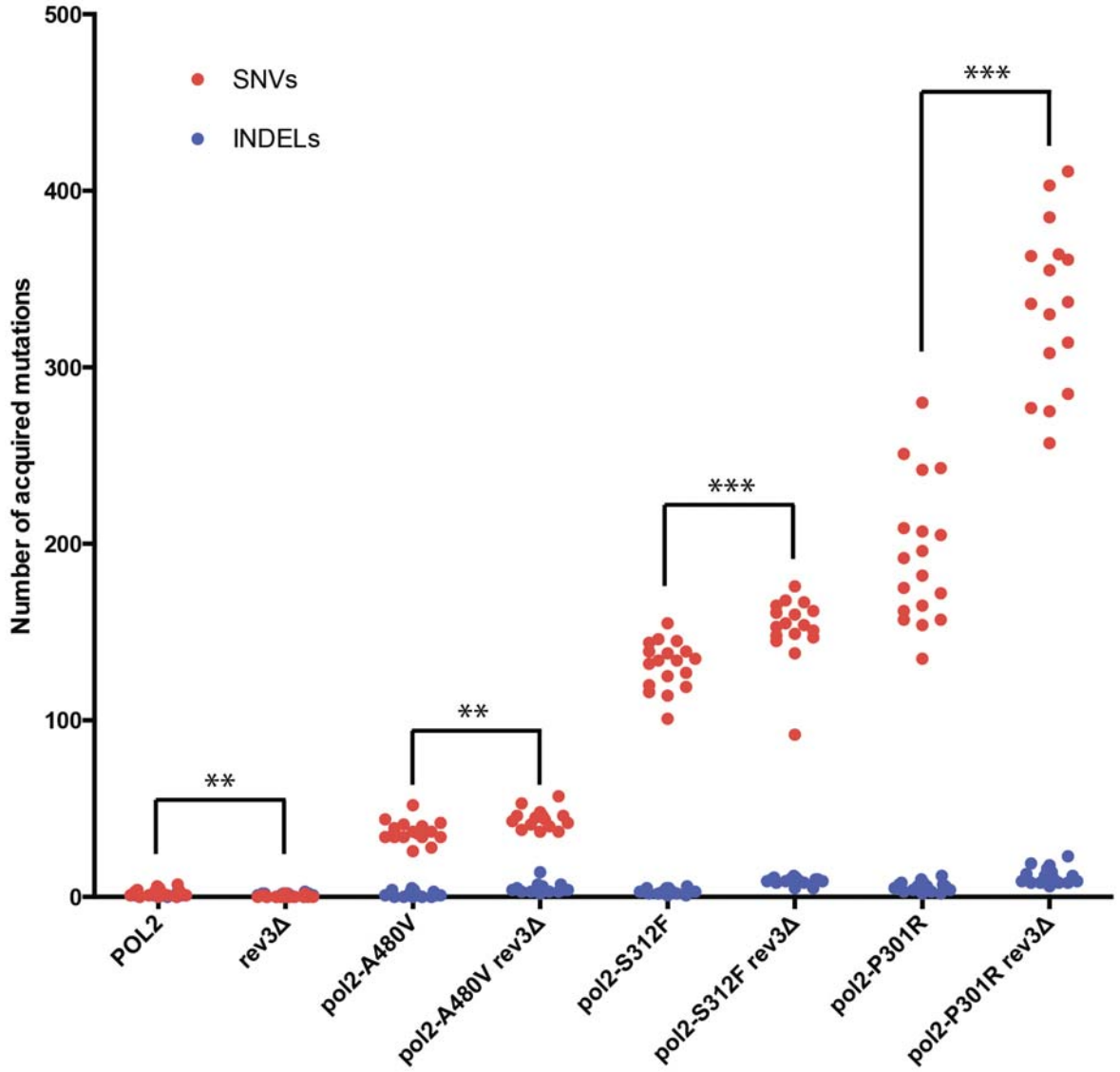

Figure 4.3: Synergystic effects on mutation number between *rev3*Δ and *pol2-P301R*
Polymerase mutation and *rev3*Δ haploid double mutants were obtained by crossing and propagated alongside haploid single mutants for 13 passages using single colony bottlenecks. Samples were sequenced before and after propagation and acquired single nucleotide variants (SNVs) and INDELs were determined as detailed in Chapter 2. The number of parallel lines are as follows: n=18 (POL2, *rev3*Δ, *pol2-S312F*, *pol2-P301R*), n=17 (*pol2-A480V*, *pol2-S312F rev3*Δ), n=16 (*pol2-P301R rev3*Δ) and n=15 (*pol2-A480V rev3*Δ). Significance of the difference in number of SNVs was determined by two-tailed, unpaired t-Tests. ** $P \leq 0.01$; *** $P \leq 0.001$

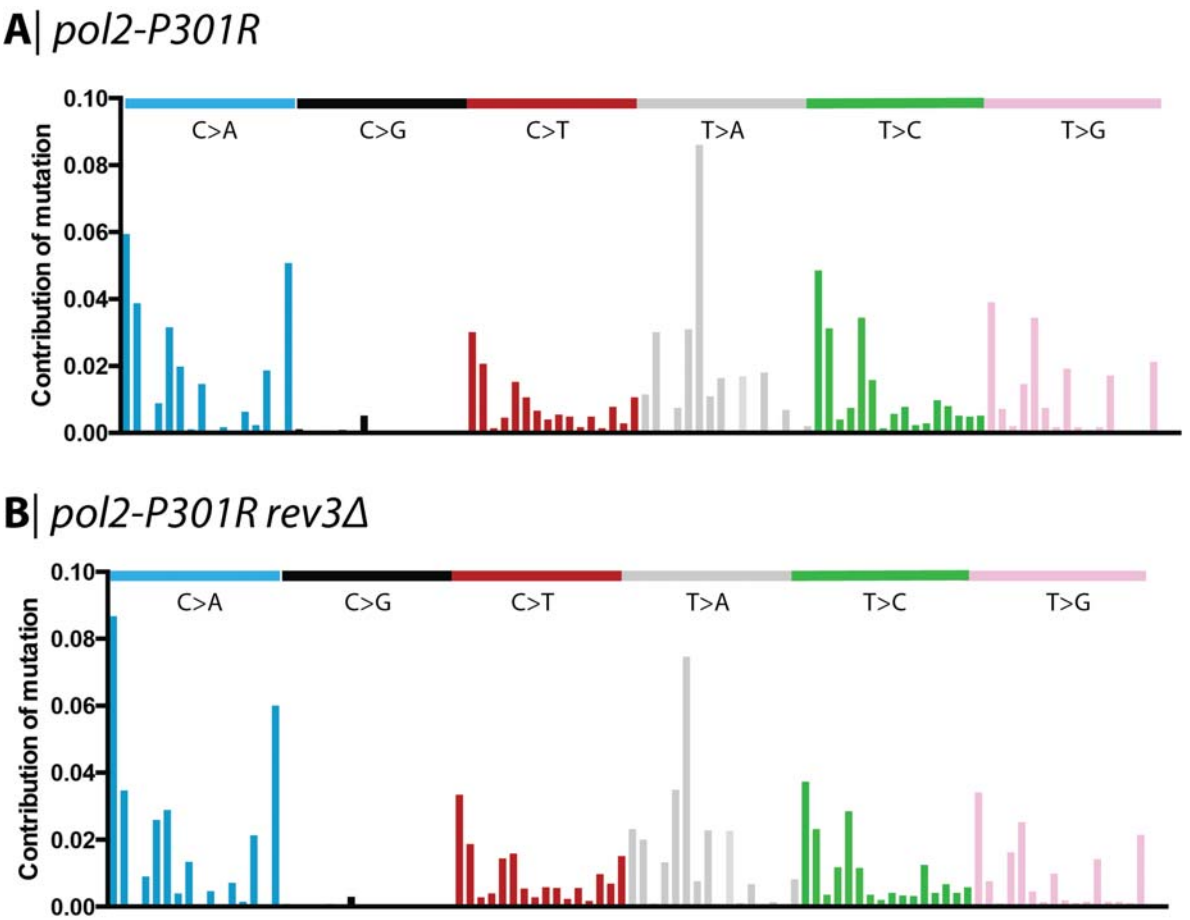

Figure 4.4: Mutational patterns observed in *pol2-P301R* cells and *pol2-P301R rev3*Δ cells are highly similar

All single nucleotide variants identified in the *pol2-P301R* cells and *pol2-P301R rev3*Δ cells displayed in Fig. 4.3 are displayed in their trinucleotide context. Relative contribution to the total of single nucelotide varaints is given. No adjustment for *S. cerevisiae* genome-wide trinucelotide frequencies is made.

4.5 Examining polymerase mutations in other organisms

The work in budding yeast *S. cerevisiae* has identified *pol2-P301R, pol2-S312F, pol2-L439V, pol2-M459K* and *pol3-S483N* as cancer mutations that likely increase the mutation rate, thus promoting tumour progression by increasing the probability of acquiring further cancer promoting mutations. To bring this evaluation of the candidate polymerase mutation full circle, mutations identfied as likely tumor promoting in yeast cells will be introduced into mammalian systems, to show whether the assertions made in budding yeast hold and observing effects in multicellular organisms.

4.5.1 The *Pole* and *Pold1* mutations in mouse models

In a collaboration with Ian Tomlinson's group in Oxford, the Adams group started to construct germline *Pole* and *Pold1* mutations in mice[768]. The Tomlinson lab provided me with two constructs: *Pold1*-S476N and *Pole*-L424V. The former is a conditional knock-in of the S476N mutation into the endogenous *Pold1* gene by a loxP mediated introduction of a mutated exon 12 (Fig. 4.5-B). The latter introduces the L424V mutation similarly by a conditional knock-in of a mutated exon 13 of *Pole* (Fig. 4.5-A). The constructs were designed by Ian Tomlinson and use inverted loxP sites. Inversion of these sites results in an expression of a fluorescent marker and a switch from the unmodified to the mutated exon. Where regions of homology exist, the sequence of the exons were optimised by using synonymous codons (for instance exon 12-14), to decrease the likelihood of secondary DNA structures.

The constructs were introduced into JM8.F6 mouse ES cells (C57BL/6N strain) by Graham Duddy from the Sanger Institute ES Cell Mutagenesis Team. I checked 48 Neo-resistant clones each for proper integration of the constructs by using PCR across the homology arms. Five clones for the *Pold1* construct and three clones for the *Pole* construct were identfied. One of the *Pold1*-targeted clones was excluded for trisomy of chromosome 8. The clones for each construct were micro-injected by the Sanger mouse facility and 40-50 mouse embryos each were transferred for gestation. For both constructs chimeras were obtained and mated to generate non-chimeric progeny. The F1 progeny was genotyped by James Hewinson (Adams group). The mice carrying the conditional *Pole*-L424V mutation have been further crossed to Flp-deleter mice to make them conditional. These mice have been sent to Oxford for phenotyping and further experiments and their sperm cryopreserved should the line need reviving. The mice carrying the conditional *Pold1*-S476N strain were successfully generated after a second round of microinjection and are currently being crossed to Flp-deleter mice. These mice should give further indications about the nature of tumors that arise due to these

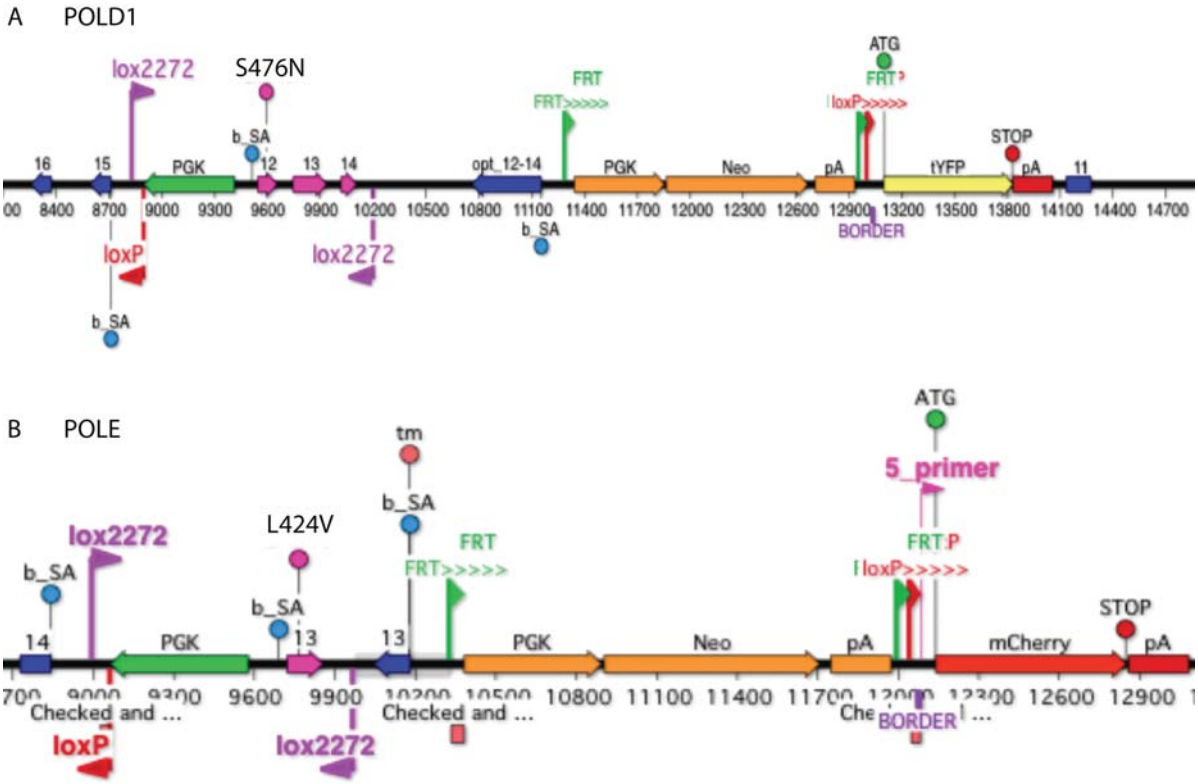

Figure 4.5: Constructs used for conditional knock-in mutations in mice
Constructs used to generate knock-in conditional mutants for **A|** *Pold1* S476N (mouse equivalent of human *POLE* p.L424V) and **B|** *Pole* L424V (mouse equivalent of *POLD1* S478N) in mouse mebryonic stem cells. Inversion of loxP sites results in the expression of a fluorescent marker and the switch of the unmodified to a mutated exon. In the case of *Pold1* exon 12-14 the exons were optimised by synonymous mutations to decrease homology mediated secondary structures. FRT sites are also included alowing the excision of the PGK Neo cassette used for clone selection. The construct was designed by Ian Tomlinson (Oxford).

two mutations: their prognosis, their organ tropism and any possible drug sensitivities.

4.5.2 Human *POLE* P286R mutant cell lines

Additional to the mice engineered to carry the two identified germline mutations from [768], I decided to also generate the *POLE* Pro286Arg mutant in a human cell line to confirm that mutation rate increases detected in *S. cerevisiae* can also be detected in human cell lines and that other characteristics like mutational patterns are conserved.

To that end, recent advances in gene editing techniques were chosen. To make the point mutation in human cells a CRISPR-Cas9^{D10A} nickase-based system developed in the Jackson group was used to construct the plasmid for transfection[887]. This involves designing guide RNAs (gRNAs) that will target the mutated CAS9 enzyme to the *POLE* gene and introduce single-stranded breaks either side of the residue to be mutated. Furthermore, a 200bp long oligonucleotide (ssODN) is supplied that carries the genomic sequence around that locus with the designed mutation as well as additional mutations to prevent re-nicking from the Cas9^{D10A} after recombination has taken place. gRNAs and the ssODN have been designed and cloned and are awaiting transfection into human cell lines (see Chapter 6.4 for sequences). The human cell line to use is currently being chosen. After genotyping a *POLE* P286R mutated human cell line will be used to assess the effects of this mutation on genomic integrity.

4.6 Summary

In this chapter, a selection of candidate pol2 mutations, shown in Chapter 3 to lead to varying increases in mutation number, were assessed for their behaviour in combination with other mutations. Inactivating mismatch repair in these haploid strains lead to either minature colonies, in case of the weakest mutator *pol2-A480V*, or to synthetic lethality in case of stronger mutators. This is concordant with an hypothesis of a lethal mutation rate leading to extinction. Experiments combining intermediate mutators with mismatch repair deficiency or repeating these experiments in diploid backgrounds are logical next steps.

Since, intriguingly, most mutator *pol2* mutations described in this work lead to mutation accumulation in excess of what results due to the mutations in the *pol2-4* background (mutations of two critical residues in the ExoI motif to alanine), which are reported to abrogate the catalytic activity of the exonuclease domain. Combination of these mutations to alanine with strong mutator candidate *pol2* mutations results in noticeable decreases of mutation accumulation, but mutation numbers are still larger than those accumulated by *pol2-4* strains.

To investigate whether a bulk of the mutagenesis could be due to the increased involvement of a translesion polymerase like Polζ, double mutants were generated and increases in mutation accumulation observed for all double mutants, though deletion of the catalytic subunit of Polζ, itself, leads to a reduction in accumulated mutations. To determine a possible source of the additional mutagenesis mutational patterns were plotted and no difference between the pattern of mutations in the polymarase ε mutants versus the mutations in the Polε/Polζ double mutant was observed. This suggests that, if anything, Polζ limits the mutagenesis due to the *pol2* mutations.

Having identfied a subset of DNA polymerase mutations that lead to an increased accumulation of mutations in budding yeast, it is crucial to demonstrate that these findings are relevant in a mammalian system. To that end, in collaboration with the Sanger Mouse Facility and Ian Tomlinson's group in Oxford, mice with conditional knock-in mutations for *Pole*-L424V and *Pold1*-S476N were created. To take this work back into a human system, we are also in the process of creating a *POLE* P286R mutated human cell line.

Evaluating hypotheses

Aims:

- Polymerase mutations and mismatch repair deficiency act synergistically when present in the same cell.

  When combining those mutations in haploid cells, most lead to synthetic lethality. The only surviving colonies we obtained - pol2-A480V msh2Δ - will need to be assessed for how both mutator phenotypes interact.

- The mutagenesis observed in mutant strains is independent of Polεexonuclease activity.

  Combining polymerase mutations with the exonuclease-deactivating missense mutations in the same gene, leads to a reduction in mutagenesis for strong mutator polymerases suggesting that they are not independent of the exonuclease catalytic site.

- The mutagenesis observed in mutant strains may be due to increased involvement of a translesion polymerase in DNA replication.

  Contrary to what one would expect if this hypothesis were true, cells deficient for the translesion polymerase Polζ display more mutagenesis, not less. Thus, Polζ is not the source of the mutagenesis observed, but other polymerases have not been ruled out.

- Candidate polymerase mutations identified as mutators in *Saccharomyces cerevisiae* will be mutators in mammalian systems.

  Mice carrying conditional knock-in mutations for the reported human germline varaints, POLE L424V and POLD1 S478N, were generated. A POLE P286R mutated human cell line is also currently being made.