Explosive mutation accumulation triggered by heterozygous human Pol ϵ proofreading-deficiency is driven by suppression of mismatch repair

Karl P Hodel, Richard de Borja, Erin E Henninger, Brittany B Campbell, Nathan Ungerleider, Nicholas Light, Tong Wu, Kimberly G Lecompte, a Yasemin Goksenin, Bruce A Bunnell, et al.

Explosive mutation accumulation triggered by heterozygous human Pol ε proofreading-deficiency is driven by suppression of mismatch repair

Karl P Hodel[1†], Richard de Borja[2†‡], Erin E Henninger[1†‡], Brittany B Campbell[3,4], Nathan Ungerleider[5], Nicholas Light[2], Tong Wu[5], Kimberly G LeCompte[1§], A Yasemin Goksenin[1#], Bruce A Bunnell[6,7], Uri Tabori[2,3,8], Adam Shlien[2,9,10], Zachary F Pursell[1,11]*

[1]Department of Biochemistry and Molecular Biology, Tulane University School of Medicine, New Orleans, United States; [2]Program in Genetics and Genome Biology, The Hospital for Sick Children, Toronto, Canada; [3]The Arthur and Sonia Labatt Brain Tumour Research Centre, The Hospital for Sick Children, Toronto, Canada; [4]Institute of Medical Science, Faculty of Medicine, University of Toronto, Toronto, Canada; [5]Department of Pathology, Tulane University School of Medicine, New Orleans, United States; [6]Department of Pharmacology, Tulane University School of Medicine, New Orleans, United States; [7]Tulane Center for Stem Cell Research and Regenerative Medicine, Tulane University School of Medicine, New Orleans, United States; [8]Division of Hematology/Oncology, The Hospital for Sick Children, Toronto, Canada; [9]Department of Paediatric Laboratory Medicine, The Hospital for Sick Children, Toronto, Canada; [10]Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Canada; [11]Tulane Cancer Center, Tulane University School of Medicine, New Orleans, United States

*For correspondence:
zpursell@tulane.edu

[†]These authors contributed equally to this work

Present address: [‡]Sorbonne Universités, UPMC Univ Paris 06, CNRS, UMR8226, Laboratoire de Biologie Moléculaire et Cellulaire des Eucaryotes, Institut de Biologie Physico-Chimique, Paris, France; [§]Louisiana State University, Baton Rouge, United States; [#]Department of Pediatrics, University of California, San Francisco, United States

Abstract Tumors defective for DNA polymerase (Pol) ε proofreading have the highest tumor mutation burden identified. A major unanswered question is whether loss of Pol ε proofreading by itself is sufficient to drive this mutagenesis, or whether additional factors are necessary. To address this, we used a combination of next generation sequencing and in vitro biochemistry on human cell lines engineered to have defects in Pol ε proofreading and mismatch repair. Absent mismatch repair, monoallelic Pol ε proofreading deficiency caused a rapid increase in a unique mutation signature, similar to that observed in tumors from patients with biallelic mismatch repair deficiency and heterozygous Pol ε mutations. Restoring mismatch repair was sufficient to suppress the explosive mutation accumulation. These results strongly suggest that concomitant suppression of mismatch repair, a hallmark of colorectal and other aggressive cancers, is a critical force for driving the explosive mutagenesis seen in tumors expressing exonuclease-deficient Pol ε.
DOI: https://doi.org/10.7554/eLife.32692.001

Introduction

Human cancers share common features of genome instability and mutagenesis (*Hanahan and Weinberg, 2011*) that are the sources of the 10^3 to 10^6 somatic mutations observed in the genomes of most types of adult tumors (*Stratton, 2011*; *Wheeler and Wang, 2013*). The total mutation burden in a tumor is the result of multiple mutational pathways operating within the cells at varying rates

eLife digest New cells are made when an existing cell divides in two. Each time a cell divides, it duplicates its DNA so that each new cell inherits a complete copy. Molecular machines called DNA polymerases make these DNA copies. The main DNA polymerases, known as delta and epsilon, can "proofread" the new DNA, which ensures that the genetic information stored in the DNA is correctly copied. Cells also use another system, called mismatch repair, to catch any errors that get missed by the polymerases.

Cancer cells contain many mutations in genes that regulate the growth and production of new cells, which is why cancers grow out of control and produce tumors. Research shows that many cancer cells with high numbers of mutations have lost their proofreading ability. Yet it is not clear if the loss of proofreading is enough to cause cancers, or if other systems, such as mismatch repair, must also be defective.

Hodel, de Borja, Henninger et al. examined human cells grown in the laboratory to understand the importance of proofreading in cancer. It turns out that even the partial loss of polymerase epsilon proofreading could lead to distinctive mutations. Yet, these mutations were repaired by mismatch repair, so they actually are only found in cells when mismatch repair is also defective. This result demonstrates that the lack of proofreading is not enough to cause a large number of mutations. These cancers only happen when other systems are damaged too.

These new findings add to the current understanding of the origins of mutations in cancers and how mutations accumulate over time. It should lead scientists to further investigate the patterns of mutations that happen in the absence of proofreading. It may also enhance our knowledge of proofreading-deficient cancers.

DOI: https://doi.org/10.7554/eLife.32692.002

over time. This can complicate attempts to assign the relative contributions of each pathway to the mutation spectrum of a tumor. One essential tool to our understanding of how mutations accumulate and influence tumor progression is using computational means to extract multiple individual signatures from many tumor genomes (*Alexandrov et al., 2013a*; *Alexandrov and Stratton, 2014*; *Haradhvala et al., 2016*). This is proving to be instrumental in resolving the relative extents to which pathways contribute to the ultimate mutation spectrum in a tumor (*Nik-Zainal et al., 2016*; *Roberts et al., 2013*). Comparing these tumor mutation signatures to those generated in experimental cell lines is another critical tool to understanding the relative rates and causality of mutation acquisition (*Fox et al., 2016*; *Helleday et al., 2014*). Traditionally, these measurements have relied on assays using reporter genes, which necessarily look at a tiny fraction of the genome and may miss global contributions to genome instability. Advances in next generation sequencing now allow for detailed genome-wide analyses of mutation accumulation over defined periods of cellular growth. Since each nucleotide in the genome is subject to the three major determinants of replication fidelity - nucleotide selection, proofreading and mismatch repair (MMR) - during every round of replication, tumors and cells with defects in replication fidelity are uniquely poised to address these issues.

Proofreading defects are now known to occur in a wide variety of tumors, with significant enrichment in colorectal and endometrial tumors (*Cancer Genome Atlas Network, 2012*; *Kandoth et al., 2013*; *Heitzer and Tomlinson, 2014*; *Rayner et al., 2016*). Mutations in DNA polymerase (Pol) ε cluster in the exonuclease proofreading domain and the tumors are clinically characterized by several criteria, including being ultrahypermutated, having a unique mutation spectrum, containing a heterozygous Pol ε mutation with no evidence of loss of heterozygosity (LOH) and being microsatellite stable (MSS) (*Briggs and Tomlinson, 2013*; *Church et al., 2013*; *Palles et al., 2013*; *Zhao et al., 2013*; *Henninger and Pursell, 2014*; *Shinbrot et al., 2014*; *Shlien et al., 2015*; *Barbari and Shcherbakova, 2017*). Whole genome and whole exome analyses of tumors have been the primary means to establish the ultrahypermutated (>100 Mutations per megabase) unique mutational signature that distinguish Pol ε tumors from other cancers (*Alexandrov et al., 2013a*; *Alexandrov and Stratton, 2014*; *Shinbrot et al., 2014*; *Shlien et al., 2015*; *Alexandrov et al., 2013b*; *Campbell et al., 2017*). While there is a rich history of studies on the effects of exonuclease defects on mutagenesis in model

organisms, the extent to which Pol ε proofreading-deficiency by itself drives each of these criteria remains poorly understood.

It is clear from studies in model organisms that complete, biallelic inactivation of Pol ε proofreading activity causes mutagenesis and carcinogenesis in model organisms, where mutation rates have been precisely measured using reporter genes. For example, mutation rates are increased in haploid or diploid yeast strains expressing only proofreading-deficient alleles of Pols ε (*Morrison et al., 1991*; *Morrison and Sugino, 1994*; *Shcherbakova et al., 2003*) or δ (*Morrison et al., 1993*; *Simon et al., 1991*; *Herr et al., 2011a*). These rates are further elevated when combined with defects in mismatch repair, indicating that these errors are made during replication (*Morrison and Sugino, 1994*; *Tran et al., 1999*; *Tran et al., 1997*; *Kennedy et al., 2015*). In mouse models, homozygous inactivation of both copies of either Pol ε or δ exonuclease activity (Pol ε$^{exo-/exo-}$ or Pol δ$^{exo-/exo-}$) causes increased mutation rates and cancer (*Albertson et al., 2009*; *Goldsby et al., 2002*; *Goldsby et al., 2001*). Interestingly, their tumor spectra are different, with gastrointestinal tumors predominant in Pol ε$^{exo-/exo-}$ mice while thymic lymphomas are the major tumor in Pol δ$^{exo-/exo-}$ mice.

However, mice with a heterozygous inactivation of a single Pol ε proofreading allele (the monoallelic Pol ε$^{wt/exo-}$ genotype) fail to develop tumors when mismatch repair is functional (*Albertson et al., 2009*). The equivalent diploid heterozygous Pol ε exonuclease mutant in yeast is also a mutator, but the effect is modest and partially dominant to the wild type allele and lacks the unique mutation spectrum seen in human tumors (*Morrison and Sugino, 1994*; *Shcherbakova et al., 2003*; *Morrison et al., 1993*; *Kane and Shcherbakova, 2014*). These results raise critical questions as to the source of the unique, ultrahypermutant phenotype in human tumors with heterozygous Pol ε exonuclease-deficiency.

Mismatch repair is responsible for the recognition and removal of replication errors and deficiencies in this activity cause genome instabilities that can lead to cancer (*Kunkel and Erie, 2005*; *Li, 2008*; *Jiricny, 2013*; *Modrich, 2006*). Mismatch repair is normally an extremely efficient process, correcting more than 99% of replication errors. However, genome-wide studies have recently shown that MMR efficiencies can vary by over two orders of magnitude and are influenced by a number of factors, including the strand on which the mismatch occurs, the polymerase that made the error, the nature of the mismatch, local sequence context, distance from the origin and replication timing (*Hawk et al., 2005*; *Hombauer et al., 2011*; *Lujan et al., 2014*; *Lujan et al., 2012*; *Supek and Lehner, 2015*). Patients with biallelic mismatch repair disorder (bMMRD) have biallelic germline inactivating mutations in a mismatch repair gene and are completely lacking mismatch repair and develop a number of early-onset tumors in which microsatellite instability (MSI) is readily detectable (*Durno et al., 2017*; *Wimmer et al., 2014*). A subset of these patients acquires a later somatic mutation in a single allele of Pol ε, leading to very aggressive tumor development. Mutation rates from these Pol ε$^{wt/exo-}$ MMR$^{-/-}$ tumors have been estimated on the order of several hundred per genome duplication (*Shlien et al., 2015*). This is consistent with results from model systems as mice with the equivalent genotype (heterozygous Pol ε$^{wt/exo-}$ combined with homozygous MMR$^{-/-}$) develop tumors within 1–2 months (*Treuting et al., 2010*). The equivalent yeast strains are strong mutators as well (*Shcherbakova et al., 2003*; *Morrison et al., 1993*; *Kennedy et al., 2015*).

However, since sporadic POLE tumors are generally microsatellite stable, the role of MMR in Pol ε proofreading-deficiency in the development of these MSS tumors remains a critical unanswered question. Whether MMR and POLE defects together are required for ultramutation, elevated mutation rates or for establishing the unique mutation signature is unknown. Understanding how MMR function or dysfunction affects proofreading-dependent mutagenesis is essential to understanding the mechanisms of mutagenesis during cancer development.

In the current study, we constructed a human cell line model system to address the roles of Pol ε proofreading in driving the clinical characteristics that define Pol ε tumors. Critically, we used a targeted knock-in approach to inactivate one copy of Pol ε 3′–5′ exonuclease activity, since human tumors contain heterozygous, monoallelic Pol ε mutations. Using mutation rates measured at a reporter gene in combination with whole-exome and whole-genome sequencing we found a rapid accumulation of large numbers of Pol ε-specific mutations in mismatch repair-deficient cells. This confirms results suggested by observations in Pol ε mutant bMMRD tumors. We further show that mismatch repair is able to suppress exonuclease-deficient Pol ε-induced mutation rates back to wild type levels using a combination of reporter gene and whole-exome sequencing (WES). These results support the idea that additional unique features beyond a single exonuclease active site inactivation
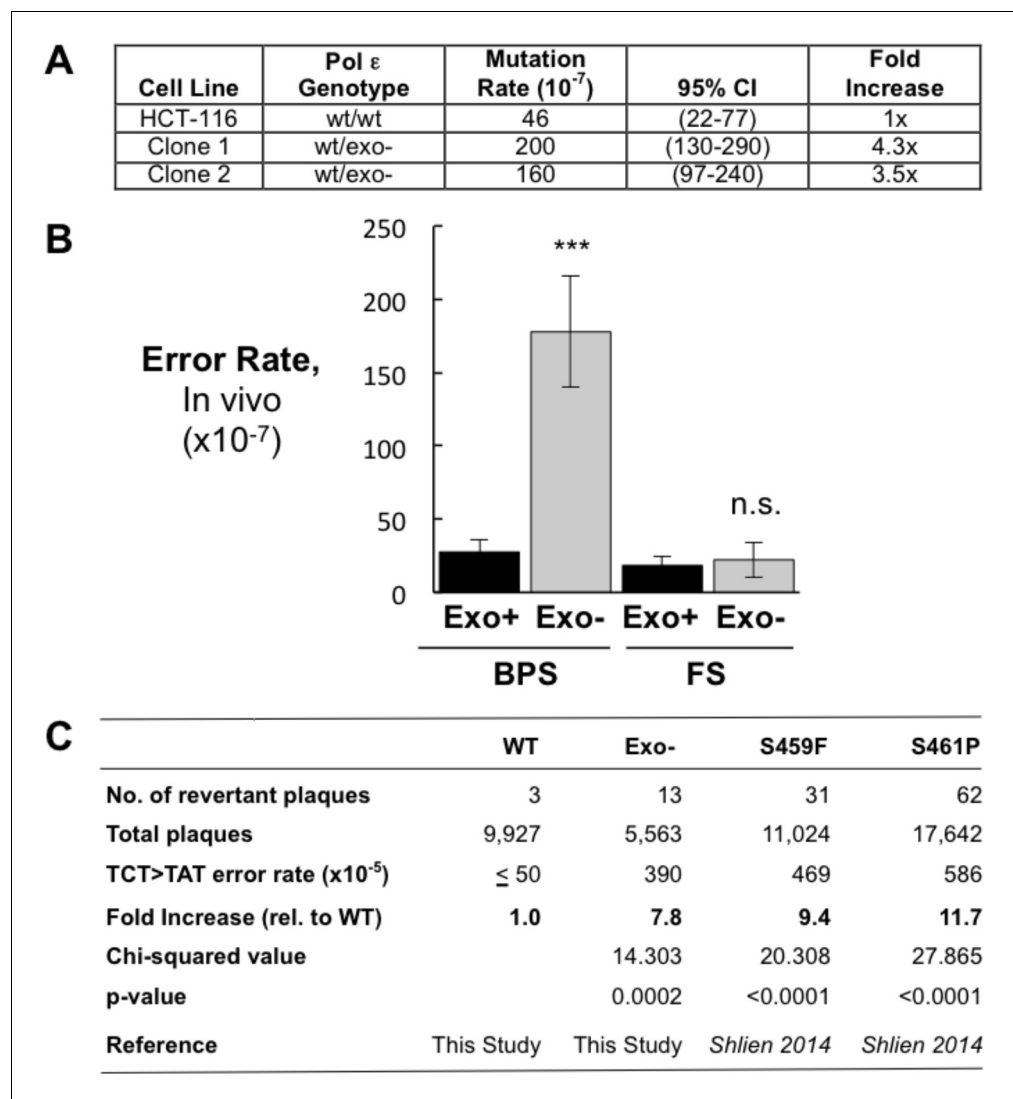
Figure 1. Heterozygous inactivation of Pol ε proofreading causes an increase in specific base pair substitutions. (**A**) Mutation rates were measured using the fluctuation assay at the HPRT1 locus by resistance to 6-thioguanine. Mutation rates and 95% confidence intervals were measured by fluctuation analysis as described in the Methods using the Ma-Sandri-Sarkar Maximum Likelihood Estimator. Twelve independent isolates of both the parental (wt/wt) cell line and two independently derived clones of the heterozygous cell lines (wt/exo-) were used. All cell lines were mismatch repair-deficient. P-values for Clones 1 and 2 (p=0.0017 and p=0.008, respectively) were calculated using an unpaired t-test relative to wt/wt. Mutation rates for Clone 1 and Clone 2 were not significantly different from one another (p=0.4727). (**B**) Error rates for base pair substitutions (BPS) and small insertion/deletion frameshift mutations (FS) were calculated using the mutation rate data from ***Figure 1A***. Exo + BPS Error Rate = 27.6×10^{-7}, SEM = 8.48×10^{-7}, n = 12; Exo- BPS Error Rate = 178×10^{-7}, SEM = 37.8×10^{-7}, n = 8; p=0.0002. Exo + FS Error Rate = 18.4×10^{-7}, SEM = 5.73×10^{-7}, n = 8; Exo- FS Error Rate = 22.2×10^{-7}, SEM = 12.1×10^{-7}, n = 1; p=0.7759. Error rate data shown for Exo- is from Clone 1 (See ***Figure 1A***). The HPRT1 ORF was sequenced from independently derived isolates of 6-TG resistant clones (these included 20 mismatch repair-deficient Pol ε[wt/wt] and 25 mismatch repair-deficient Pol ε[wt/exo-] clones; see Materials and methods). Sequence changes used to calculate error rates are in ***Figure 1—source data 2***. ***p<0.001; n.s., p>0.05. (**C**) Errors rates were calculated using a lacZ reversion substrate that reverts via TCT→TAT transversion. P values were calculated using chi-square tests with Yates correction. Error rates are the averages of two experiments, each conducted with independent DNA and enzyme preparations for each construct tested. ≤indicates the value is a maximal estimate as it is identical to the assay background.
DOI: https://doi.org/10.7554/eLife.32692.003

The following source data and figure supplements are available for figure 1:

Figure 1 continued on next page

are helping facilitate the massive mutation acquisition seen in microsatellite stable tumors containing mutant Pol ε.

Results

Inactivation of Pol ε proofreading causes a mutator phenotype in human cells

Tumors with mutations in the exonuclease domain of POLE are generally microsatellite stable and show no or low loss of heterozygosity, suggesting that inactivation of exonuclease activity in one allele is sufficient to drive mutagenesis and tumor development, though this has not been directly tested previously. To test whether inactivation of a single allele of Pol ε proofreading was sufficient to cause a mutator phenotype in human cells, we used recombinant adenoassociated virus (rAAV)-mediated gene targeting to engineer a diploid human cell line to express one allele of Pol ε with the D275A/E277A double substitution (*Figure 1—figure supplements 1–2*; *Figure 1—source data 1*). We chose the D275A/E277A mutation because it inactivates exonuclease proofreading in vitro (*Shcherbakova et al., 2003*; *Korona et al., 2011*). The parental cell line, HCT-116, is constitutively mismatch repair-deficient due to an inactivating mutation in Mlh1, thus allowing us to first define the contributions of proofreading deficiency separately to mutagenesis. We then measured the mutation rate at the hypoxanthine-guanine phosphoribosyltransferase (HPRT1) locus using 6-thioguanine (6-TG) resistance and a fluctuation assay. The measurements were repeated in clones derived from independent exonuclease-deficient (exo-) allele integration events. A moderate mutator effect was seen in Pol ε^wt/exo- heterozygotes (*Figure 1A*), indicating the exo- allele was partially dominant over the endogenous exo + allele, similar to what is seen in a mismatch repair-deficient diploid cell line heterozygous for a Pol ε proofreading mutation, *pol2-4/+pms1/pms1* (*Pavlov et al., 2004*). Mutation rates were not measured in cells from the comparable heterozygous Pol ε^wt/exo- mice lacking mismatch repair (*Albertson et al., 2009*).

To begin measuring the effect of inactivating a single Pol ε exonuclease allele on mutation rates in cells, we sequenced the HPRT1 gene from twenty and twenty-five independently derived 6-TG^R (and thus HPRT1 mutant) clones from mismatch repair-deficient Pol ε^wt/wt and Pol ε^wt/exo- cells, respectively (*Figure 1—source data 2*). This allowed comparison to previously measured mutation rates from different groups using the same parental cell line. Mutation rates from the Pol ε^wt/wt cells were similar to the spontaneous mutation rates reported by three previous studies (*Bhattacharyya et al., 1995*; *Glaab and Tindall, 1997*; *Ohzeki et al., 1997*). These results suggest that the baseline rates of mutagenesis are an accurate measure of comparison for the Pol ε^wt/exo- cell lines.

The increase in mutation rate seen in the Pol ε^wt/exo- mismatch repair-deficient cells was primarily due to base pair substitutions (*Figure 1B*). Frameshift error rates did not change, in agreement with previous findings in vitro that Pol ε proofreading primarily strongly corrects base-base mispairs with little effect on frameshift fidelity (*Korona et al., 2011*). However, the number of mutational events scored by this method is insufficient to make statistical claims regarding individual mutations, reinforcing the need for genome sequencing to examine mutations in all possible sequence contexts.

Using an in vitro lacZ reversion substrate that specifically measures TCT→TAT transversions (*Shinbrot et al., 2014*; *Shlien et al., 2015*), the D275A/E277A mutant made these errors at a
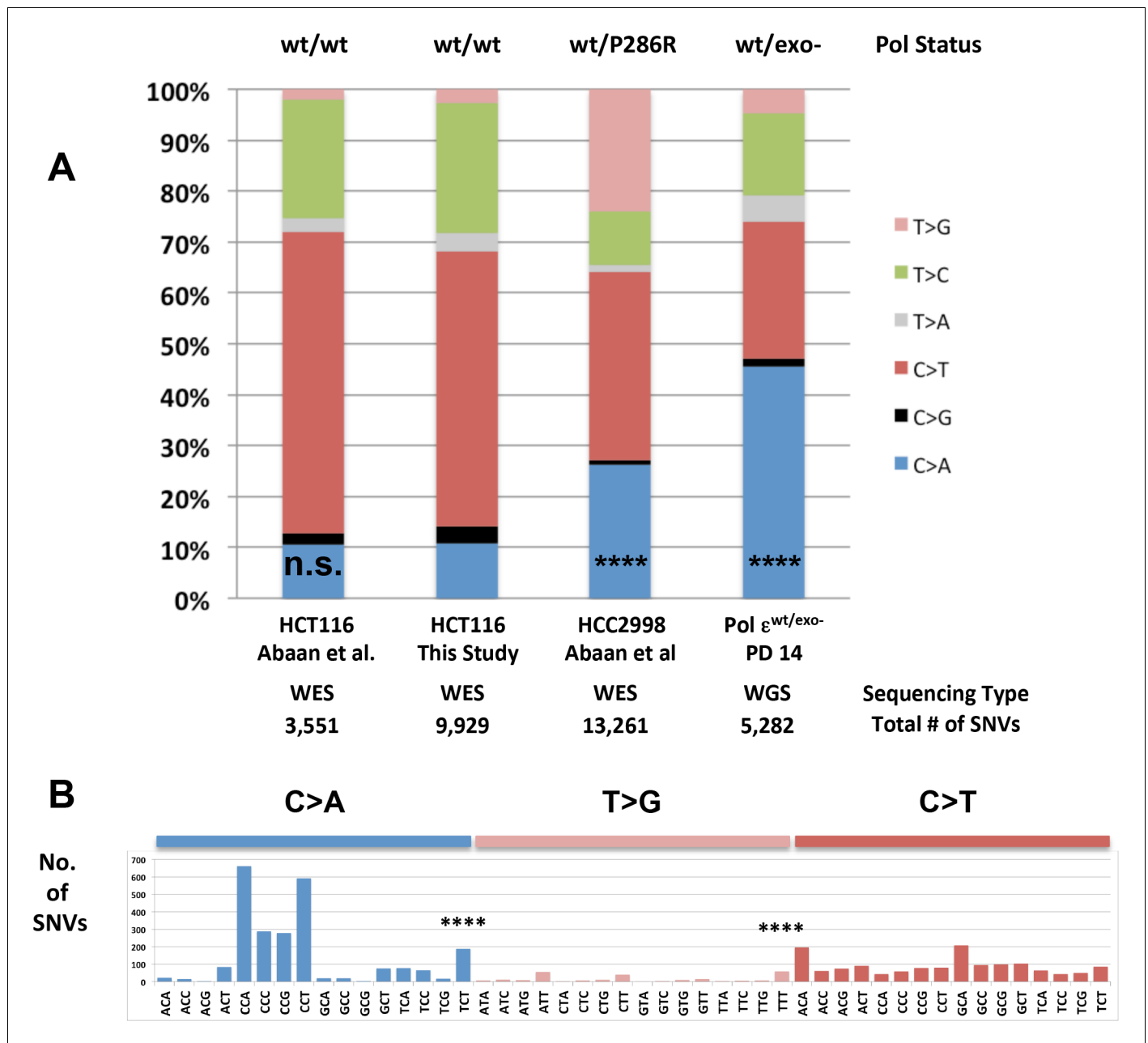
Figure 2. Whole-genome sequencing from defined population doubling Pol ε$^{wt/exo-}$ mismatch repair-deficient cells. (**A**) Whole genome sequencing (2.8 × 10^9 bp, average 30X coverage) was performed on Pol ε$^{wt/exo-}$ cells lacking mismatch repair at two defined population doubling levels, P0 and P14, as described in the Methods. P0 was used as the matched normal cells to define only those mutations arising during the 14 population doublings. The fraction of each type of base pair substitution from the PD 14 Pol ε$^{wt/exo-}$ cells was plotted and compared to the fraction of each type of mutation from HCT116 ((*Abaan et al., 2013*) and this study) and HCC2998 cells (*Abaan et al., 2013*). Chi square tests with Yates correction were used to calculate p values relative to SNVs found in Pol ε$^{wt/wt}$ mismatch repair-deficient cells in this study. Pol ε$^{wt/wt}$ (Abaan et al.) χ^2 = 0.033, p=0.8551; Pol ε$^{wt/P286R}$ (Abaan et al.) χ^2 = 872.341, p<0.0001; Pol ε$^{wt/exo-}$ χ^2 = 2,3680.508, p<0.0001. ****p<0.0001; n.s., not significant. (**B**) The number of each indicated base pair substitution in a specific trinucleotide context was plotted from the PD 14 Pol ε$^{wt/exo-}$ mismatch repair-deficient cells. The base pair substitutions shown (C > A and T > G transversions, left; C > T transitions, right) are those found enriched in POLE tumors. Chi square tests with Yates correction were used to calculate p-values relative to SNVs found in Pol ε$^{wt/wt}$ mismatch repair-deficient cells in this study. C > A TCT χ^2 = 152.772, p<0.0001; T > G TTT χ^2 = 72.254, p<0.0001. ****p<0.0001.

DOI: https://doi.org/10.7554/eLife.32692.008

The following figure supplements are available for figure 2:

Figure supplement 1. Whole genome SNVs identified in Pol ε$^{wt/exo-}$ (PDL = 14) in cells lacking functional mismatch repair identified.

DOI: https://doi.org/10.7554/eLife.32692.009

Figure 2 continued on next page

Figure 2 continued

Figure supplement 2. POLE mutation signature extracted from POLE-mutant cell lines.

DOI: https://doi.org/10.7554/eLife.32692.010

Figure supplement 3. Mutation counts in the indicated trinucleotide context (*top*) were plotted as a proportion of their occurrence (*bottom*) in WGS samples.

DOI: https://doi.org/10.7554/eLife.32692.011

Figure supplement 4. Mutation counts in the indicated trinucleotide context (*top*) were plotted as a proportion of their occurrence (*bottom*) in WES samples.

DOI: https://doi.org/10.7554/eLife.32692.012

Figure supplement 5. Relative contributions of Cosmic Mutation Signatures to individual patient mutation spectra were determined using deconstructSig.

DOI: https://doi.org/10.7554/eLife.32692.013

Figure supplement 6. Mean coverage was greater than 90x for each WES sample.

DOI: https://doi.org/10.7554/eLife.32692.014

Figure supplement 7. Alignment rate to the reference genome exceeded 99% for each WES sample.

DOI: https://doi.org/10.7554/eLife.32692.015

Figure supplement 8. Total reads exceeded 60 million for each WES sample.

DOI: https://doi.org/10.7554/eLife.32692.016

Figure supplement 9. Greater than 90% of the bases in the WES genome exceeded 30x coverage.

DOI: https://doi.org/10.7554/eLife.32692.017

Figure supplement 10. Greater than 85% of the bases in the WGS genome exceeded 20x coverage.

DOI: https://doi.org/10.7554/eLife.32692.018

Figure supplement 11. Average alternate base quality to reference base quality of ~1.0.

DOI: https://doi.org/10.7554/eLife.32692.019

significantly higher rate in vitro than the wild type exonuclease-proficient Pol ε enzyme (*Figure 1C*). We used a construct comprised of the N-terminal 140 kDa of Pol ε, which contains the DNA polymerase and exonuclease domains and has similar fidelity and catalytic activity to the complete four subunit holoenzyme (*Aksenova et al., 2010*; *Ganai et al., 2015*; *Zahurancik et al., 2015*). Importantly, the elevated TCT→TAT error rate we observed with the D275A/E277A mutant was not statistically different from those measured with the S459F and S461P Pol ε cancer mutants previously (*Shinbrot et al., 2014*; *Shlien et al., 2015*), suggesting a common mechanism of mutagenesis for these hotspot mutations.

Mutation rates calculated using reporter genes (μ_L) can be used to extrapolate to genome-wide per base pair mutation rates (μ_{BS}) (*Drake, 1991*; *Lynch, 2010*). The availability of high-throughput DNA sequencing now allows for empirical validation of these calculations in addition to providing insight into the influence of genomic context on mutagenesis. To address this we performed whole-genome sequencing (2.8×10^9 bp at an average depth of 36.1x) on genomic DNA prepared from Pol ε$^{wt/exo-}$ cells. Based on our measured mutation rate for HPRT1 (μ_L) in Pol ε$^{wt/exo-}$ cells lacking mismatch repair (180×10^{-7}), we calculated a μ_{BS} value of 0.23×10^{-7} mutations per base pair per genome duplication.

Because the parental HCT-116 cell line already carries a significant number of single nucleotide variants (SNVs) relative to the human reference sequence ([*Abaan et al., 2013*] and see Discussion), we needed a way of measuring de novo mutations resulting from Pol ε-dependent replication errors. To do this we first performed whole genome sequencing (WGS) on genomic DNA prepared from mismatch repair-deficient Pol ε$^{wt/exo-}$ cells, which we then used as a matched normal control. We termed this mutation spectrum P0. We then passaged these cells through a calculated 13.9 population doublings and then performed WGS again on the passaged population, which we termed P14. Mutations unique to P14 arose during the defined number of population doublings. The P0 and P14 samples contained 140.3 and 141.4 Mut/Mb, respectively. Given the calculated μ_{BS} and the 2.8×10^9 bp sequenced, we predicted the accumulation of 906 novel genome-wide mutations after 14 population doublings. Whole-genome sequencing revealed 5,282 SNVs unique to the P14 population, 5.8-fold higher than that predicted from the μ_L at HPRT1. Mutations observed in HPRT1 in this cell line may thus slightly underrepresent those found genome-wide. This difference is consistent with what is seen in microbes, where reporter gene mutation rates are consistently 6–8-fold lower than concurrently measured whole-genome mutation rates, likely due to phenotypic lag, strong
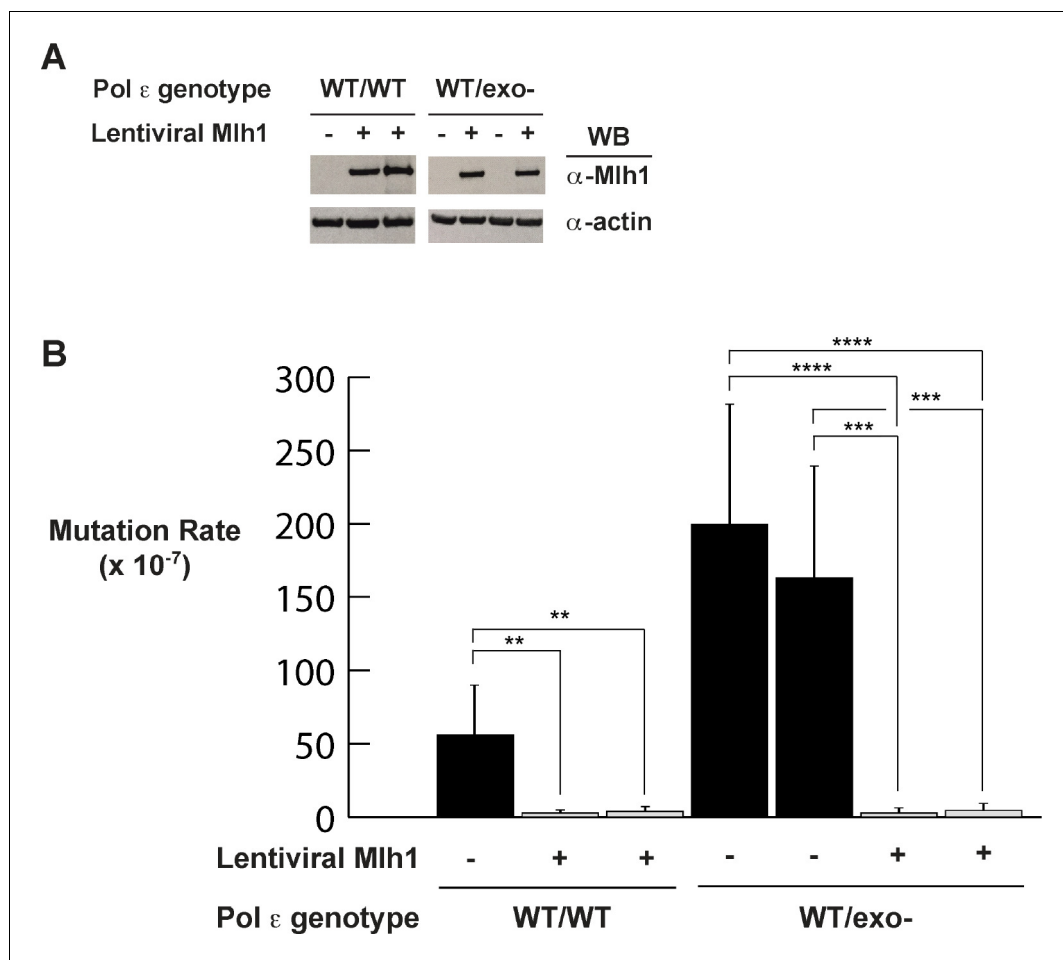
Figure 3. Mismatch repair suppresses exonuclease-deficient Pol ε-induced mutation rate increase. (**A**) Lentivirus encoding human Mlh1 was generated and used to infect parental cells with wild type Pol ε and cells heterozygous for Pol ε exonuclease deficiency. Cell lysates were probed by Western blot using antibodies against Mlh1 and β-actin. (**B**) Mutation rates were measured by fluctuation analysis as described in the Methods using the Ma-Sandri-Sarkar Maximum Likelihood Estimator. Twelve independent isolates from each of two parental (wt/wt) and two heterozygous cell lines (wt/exo-) expressing Mlh1 were used. 95% confidence intervals are shown. Pol $\varepsilon^{wt/wt}$ Mlh1+ Clone 1 Mutation Rate = 1.7×10^{-7}, SEM = 0.72×10^{-7}, p=0. 0046. Pol $\varepsilon^{wt/wt}$ Mlh1+ Clone 2 Mutation Rate = 2.5×10^{-7}, SEM = 1.1×10^{-7}, p=0.0053. Pol $\varepsilon^{wt/exo-}$ Mlh1+ Clone 1 Mutation Rate = 2.3×10^{-7}, SEM = 0.81×10^{-7}, p<0.0001 (vs. Pol $\varepsilon^{wt/exo-}$ Mlh1- Clone 1) and p=0.0003 (vs. Pol $\varepsilon^{wt/exo-}$ Mlh1- Clone 2). Pol $\varepsilon^{wt/exo-}$ Mlh1+ Clone 2 Mutation Rate = 3×10^{-7}, SEM = 1.3×10^{-7}, p<0.0001 (vs. Pol $\varepsilon^{wt/exo-}$ Mlh1- Clone 1) and p=0.0003 (vs. Pol $\varepsilon^{wt/exo-}$ Mlh1- Clone 2). Mutation Rates for Pol $\varepsilon^{wt/exo-}$ Mlh1+ Clone 1 and Clone 2 were not significantly different (p=0.6485). Mutation rates from cells lacking mismatch repair (from **Figure 1A**) are shown for comparison.
DOI: https://doi.org/10.7554/eLife.32692.020

selective pressure and transcription in the reporter gene (**Drake, 2012**; **Jee et al., 2016**; **Lee et al., 2012**).

C→A transversions exceeding 20% of all base pair substitutions is a primary characteristic of mutation spectra from tumors containing Pol ε exonuclease domain mutations (**Rayner et al., 2016**; **Shinbrot et al., 2014**). C→A transversions were increased significantly in the Pol $\varepsilon^{wt/exo-}$ cells as compared to the control Pol $\varepsilon^{wt/wt}$ spectrum, accounting for 46% of all base pair substitutions (**Figure 2A**, χ^2 = 11.874, p<0.0001). These were not cell line artifacts, as whole exome sequencing from HCT-116 cells from two independent studies ([**Abaan et al., 2013**] and this study) showed roughly 10% C→A transversions (**Figure 2A**, p>0.5). HCC2998 cells, which harbor the Pol $\varepsilon^{wt/P286R}$ mutation, also showed a significant increase in C→A transversions relative to Pol $\varepsilon^{wt/wt}$ cells (**Figure 2A**, p<0.0001).

Two sequence context mutational hotspots were observed that are consistent with Pol ε exonuclease domain mutant spectra: C→A transversions in TCT context and T→G transversions in TTT context and, to a lesser extent, ATT and GTT contexts (*Figure 2B*). These hotspots are seen in Pol ε tumors from patients with bMMRD (*Shlien et al., 2015*), colorectal and endometrial cancer (*Alexandrov and Stratton, 2014*; *Cancer Genome Atlas Network, 2012*; *Kandoth et al., 2013*; *Shinbrot et al., 2014*), as well as in the Pol ε-P286R HCC2998 cells (*Figure 2—figure supplement 1*, data extracted from [*Abaan et al., 2013*]). These are not mutational hotspots in HCT-116 cells, which contain wild type Pol ε (*Figure 2—figure supplement 1*). The largest number of mutations that arose during the 14 doublings were C→A transversions in triplet contexts containing adjacent cytosines: CCA, CCT, CCC and CCG. Triplet nucleotide occurrences can vary in the regions captured by WGS and WES. In order to address this we reanalyzed each sample relative to the number of times each trinucleotide is found in the relevant sample and found the hotspot patterns are all retained (*Figure 2—figure supplements 3–4*). The increase in C→A mutations in the CCT context was also seen in Pol ε exonuclease domain (EDM) tumors from bMMRD patients (*Shlien et al., 2015*), suggesting a link between Pol ε replication errors left uncorrected by mismatch repair. C→A mutations in CCA, CCC and CCG contexts are slightly elevated in Mutation Signature 20, which has been associated with loss of mismatch repair (*Alexandrov et al., 2013b*). These transversions were seen in the HCT116 cell line with wild type Pol ε (*Figure 2—figure supplement 1*), though to a lesser extent. The lack of C→T transitions in TCG contexts is significantly different from colorectal and endometrial Pol ε tumors, but consistent with their absence from bMMRD tumors with Pol ε EDM mutations (*Cancer Genome Atlas Network, 2012*; *Kandoth et al., 2013*; *Shinbrot et al., 2014*).

Expression of MMR suppresses Pol ε^wt/exo- mutagenesis

While it is clear that Pol ε-dependent mutagenesis in the absence of functional MMR accounts for the ultramutated phenotype in bMMRD tumors with Pol ε mutations, the role of MMR in Pol ε somatic tumors is less clear. In order to measure the effects of MMR on Pol ε exonuclease-dependent replication errors, we wanted to measure error rates in both the presence and absence of MMR. Previous studies have restored MMR by stably adding Mlh1-expressing chromosome 3 to cells (*Glaab and Tindall, 1997*). We made Mlh1-encoding lentivirus and used this to infect Mlh1-deficient HCT-116 cells containing wild type and mutant Pol ε (*Figure 3A*). Lentiviral Mlh1 expression reduced mutation rates at the HPRT1 locus by 14- to 20-fold in the wild type polymerase background (*Figure 3B*), similar to the 12-fold reduction reported when the Mlh1-encoding chromosome 3 was added back to HCT-116 cells ([*Glaab and Tindall, 1997*; *Tindall et al., 1998*]; 73×10^{-7} and 5.9×10^{-7}; 12.4-fold reduction), indicating that the expressed Mlh1 is functional.

Mlh1 expression in Pol ε^wt/exo- cells caused an over 50-fold decrease in the mutation rate (to 2.3 and 3.0×10^{-7}, *Figure 3B*), making them indistinguishable from those measured in Pol ε^wt/wt cells with Mlh1 expressed (*Figure 3B*). This result also suggests that Msh3 is unlikely to play a significant role in correcting the exonuclease-deficient Pol ε errors since HCT-116 cells are deficient in this factor and it was not added back in these experiments (*Papadopoulos et al., 1994*).

When fluctuation assay mutation rates are very low due to a significant number of independent isolates giving rise to zero HPRT1-mutant colonies, as was the case here, an alternative method to measure mutation rates can be used. We chose to periodically measure HPRT1 mutant frequencies at increasing population doubling level (PDL), where the slope of the plotted line is equal to the mutation rate (*Glaab and Tindall, 1997*). We measured HPRT1 mutant frequencies at several population doublings from PDL = 0 to PDL = 70 or 71 in Pol ε^wt/wt and Pol ε^wt/exo- cells expressing Mlh1, respectively (*Figure 4A*). At each PDL we scored between 1 and 19 6-TG-resistant colonies. However, when we sequenced the HPRT1 ORF from all 6-TG-resistant colonies we saw many instances of repeat mutations in a collection from a single PDL, indicative of a single mutational event that expanded throughout the population. Plotting mutant frequency values calculated for the indicated PDL using only the unique HPRT1 mutations (*Figure 4—source data 1*) returned a line with slope of ~1, suggesting that the mutation rates were at or near the level of detection of this assay. The Pol ε^wt/exo- mutant frequencies were consistently higher than those from the Pol ε^wt/wt cells, but this difference was not statistically significant (*Figure 4A*).

To determine if this phenomenon held throughout the genome, we carried out whole-exome sequencing to an average depth of 100x on the early (PDL = 0) and late (PDL = 70) samples from
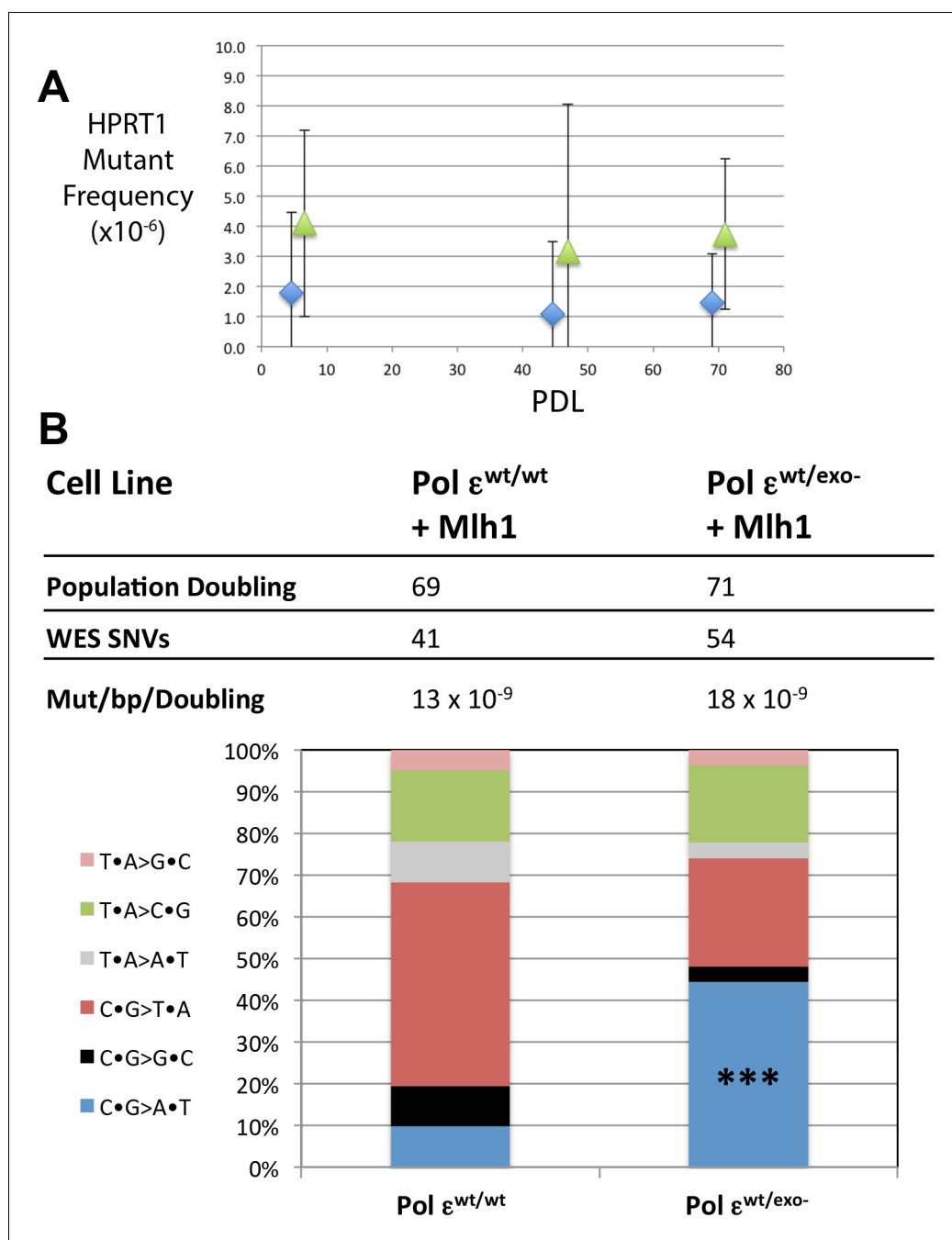
Figure 4. Mismatch repair suppresses the majority of exonuclease-deficient Pol ε mutation specificity. (**A**) Cells were continuously passaged and PDL was calculated using the following equation: PDL = [ln(N_t)-ln(N_0*PE)]/ln2. N_t = Number of viable cells counted after passage; N_0 = Number of cells seeded prior to passage; PE = plating efficiency. Mutant frequencies were measured for each mismatch repair proficient strain at the indicated PDL (diamonds, Pol ε$^{wt/wt}$; triangles, Pol ε$^{wt/exo-}$). Ten plates for each cell lines were seeded with 2×10^5 cells at each PDL into media containing 6-TG and grown for 12–14 days. Each 6^{TG}-resistant clone was isolated, expanded and the HPRT1 ORF was sequenced. Mutant frequencies were calculated based on the number of unique HPRT1 mutations at each PDL. Pol ε$^{wt/wt}$ PDL6.4 MF = 1.8×10^{-6}, SEM = 2.7×10^{-6}, n = 4; Pol ε$^{wt/exo-}$ PDL6.6 MF = 4.1×10^{-6}, SEM = 3.1×10^{-6}, n = 3, p=0.6003. Pol ε$^{wt/wt}$ PDL44.6 MF = 1.1×10^{-6}, SEM = 2.3×10^{-6}, n = 2; Pol ε$^{wt/exo-}$ PDL47.9 MF = 3.2×10^{-6}, SEM = 4.7×10^{-6}, n = 8, p=0.9066. Pol ε$^{wt/wt}$ PDL69 MF = 1.5×10^{-6}, SEM = 1.6×10^{-6}, n = 5; Pol ε$^{wt/exo-}$ PDL71 MF = 3.7×10^{-6}, SEM = 2.6×10^{-6}, n = 5, p=0.4917. (**B**) Whole exome sequencing (30×10^6 bp, average 101x coverage) was performed on the indicated cell line at two defined population doubling levels, P0 and P69 or P71, as described in the Methods. P0 for each cell line was used as the matched

Figure 4 continued on next page

Figure 4 continued

normal cells to define only those mutations arising during the 70 or 71 population doublings. The fraction of each type of base pair substitution found unique to PDL 69 (for Pol ε$^{wt/wt}$) or PDL 71 (Pol ε$^{wt/exo-}$) was plotted and compared. Fisher's exact tests were used to calculate p values. p=0.0002 (***p<0.001).

DOI: https://doi.org/10.7554/eLife.32692.021

The following source data and figure supplement are available for figure 4:

Source data 1. HPRT1 mutations sequenced from mismatch repair-proficient cells.

DOI: https://doi.org/10.7554/eLife.32692.023

Source data 2. Pol ε mutation spectra calculation of cosine similarity to cancer mutation spectra.

DOI: https://doi.org/10.7554/eLife.32692.024

Figure supplement 1. Whole exome SNVs identified in Pol ε$^{wt/exo-}$ (PDL = 71) and Pol ε$^{wt/wt}$ (PDL = 70) cells expressing Mlh1 were analyzed for their triplet nucleotide sequence context.

DOI: https://doi.org/10.7554/eLife.32692.022

---

both Pol ε$^{wt/wt}$ and Pol ε$^{wt/exo-}$ mismatch repair-proficient cell lines (*Figure 4B*). Using the PDL = 0 samples as matched normal controls, we measured similar low mutation rates in Pol ε$^{wt/wt}$ and Pol ε$^{wt/exo-}$ cells (13 × 10^{-9} Mut/bp/doubling and 18 × 10^{-9} Mut/bp/doubling, respectively). The total numbers of all mutations acquired were essentially no different than with wild type Pol ε. Interestingly, there was a statistically significant increase in C→A transversions (p=0.0002) between the mismatch repair-proficient Pol ε$^{wt/exo-}$ cells and the mismatch repair-proficient Pol ε$^{wt/wt}$ cells, while no statistically significant difference was found in any other class of base pair substitution (p>0.2 for each of the six classes, Fisher's Exact Test). Further, all triplet context mutations were observed in insufficient numbers to evaluate statistically. C→A mutations were, however, observed in all triplet contexts seen as hotspots in the MMR-deficient cells (CCA, CCC, CCG, CCT and TCT, *Figure 4—figure supplement 1*). Mutation signature 10, the unique Pol ε mutation signature, was extracted from Pol ε exonuclease-deficient mutation spectra from cells with and without mismatch repair (*Figure 2—figure supplement 2* and *Figure 4—source data 2*). The relative contribution of signature 10 in Pol ε exo-deficient cells is closer to that seen in bMMRD patients (*Figure 2—figure supplement 5*), most likely due to the relative absence of C→T transitions in TCG context. These results indicate that the majority of replication errors made by the Pol ε-D275A/E277A mutant are in fact corrected by mismatch repair.

Discussion

In the current study we examined the relative contributions of two essential determinants of replication fidelity, proofreading and mismatch repair, on mutagenesis in human cells. We used a combination of gene editing, reporter gene studies and next generation sequencing to measure mutation rates and specificities in human cells engineered to model proofreading-deficient tumors with and without mismatch repair. This is the equivalent to what occurs in human tumors with mutations in the Pol ε exonuclease domain and genomic mutation frequencies exceeding 100 mutations per Mb (*Cancer Genome Atlas Network, 2012*; *Rayner et al., 2016*; *Shinbrot et al., 2014*; *Shlien et al., 2015*). We show that large and rapid mutation accumulation occurs when Pol ε exonuclease domain mutations occur along with inactivation of mismatch repair. Most of these are specific transversion mutations known to be hotspots of exonuclease-deficient Pol ε mutagenesis. We further show that this large increase in mutation rate is largely suppressed by functional mismatch repair. Taken together, these results suggest that the mechanism of replication error mutagenesis in sporadic tumors with heterozygous Pol ε mutations likely requires an additional feature, several of which are described below, including suppression of MMR and alternative effects on Pol ε activity.

We used rAAV-mediated gene targeting to replace two exonuclease active site residues, D275 and E277, with alanines on a single POLE allele. The single allelic inactivation was chosen to model the case in tumors with heterozygous Pol ε mutations. This double amino acid substitution has been shown to inactivate exonuclease proofreading in vitro and cause increased reporter gene mutation rates in yeast and mammalian cells (*Morrison et al., 1991*; *Morrison and Sugino, 1994*; *Tran et al., 1999*; *Albertson et al., 2009*; *Korona et al., 2011*; *Shcherbakova and Pavlov, 1996*; *Agbor et al., 2013*). Next generation sequencing on these cells in the presence or absence of mismatch repair over defined

numbers of population doublings allowed us to compare genome-wide mutation rates and spectra to the mutation spectra from patient tumors.

Unbiased whole-genome sequencing confirmed the rapid accumulation of Pol ε-specific mutations seen in POLE tumors lacking functional mismatch repair (*Shlien et al., 2015*). The total number of measured SNVs suggests a mutation rate of 380 mutations per population doubling, similar to the 608 mutations per cell cycle calculated for a mismatch repair-deficient brain tumor harboring a Pol ε exonuclease domain mutation. Our cellular mutation rate values possibly underestimate the true Pol ε exonuclease-deficient mutation rate for several reasons. Our data were generated from a cancer cell line with a large number of pre-existing mutations (*Abaan et al., 2013*), as well as additional mutations that have assuredly arisen during passaging in the laboratory. These could conceivably include suppressor mutations functioning to restrain elevated mutation rates (*Morrison and Sugino, 1994*; *Herr et al., 2011a*; *Williams et al., 2013*). Importantly, no additional mutations in POLE were sequenced, suggesting that viability of this cell line is not due to an acquired mutation elsewhere in POLE acting to suppress the mutation rate, as occurs frequently in yeast (*Herr et al., 2011a*; *Williams et al., 2013*; *Herr et al., 2011b*; *Dennis et al., 2017*). While we cannot formally exclude the possibility that a de novo mutation in another gene acted to suppress the mutation rate in trans, no obvious candidates were identified.

An additional reason that our mutation rates may underestimate the true mutation rate is that mutations that arise in the last several rounds of replication and those that fall below 5% allele frequency would not meet the threshold for scoring as a true SNV. The genome data was generated using an instrument with high accuracy (<1% error rate) and variants were called using an established algorithm, however there are indeed a small number of areas in the genome that are inaccessible – either due to gaps in the reference assembly, or excessive numbers of repeats that prevent proper alignment. Experiments using single-cell sequencing could address these issues, ideally by selecting single cells, expanding subclones and then measuring mutations at higher stringency values than used here. These rates are also similar to the per base pair mutation rates in haploid yeast with complete Pol ε exonuclease deficiency and disrupted MMR (*Kennedy et al., 2015*). This similarity is striking considering our measurements were made in a heterozygous diploid human cell line. A key finding from the yeast study was that individual cell mutation rates could vary by an order of magnitude. We are currently unable to measure mutation rates in individual cells, but this remains a critical issue to address in future studies.

The unique mutation spectrum seen in POLE tumors was recapitulated in our gene-targeted cell lines, with one notable exception. In tumors, many C→A transversions occur in a highly specific triplet sequence context, TCT, which we also see in the cell lines, though not to the same proportion as in the tumor genomes. Interestingly, this particular mutation is also enriched in yeast with the P286R equivalent allele (*Barbari and Shcherbakova, 2017*). We also observe increased T→G transversions in TTT (and to a lesser extent ATT and CTT) context, similar to Pol ε tumors. Because of the limited number of mutational target sites we cannot at this time draw conclusions as to Pol ε strand usage during replication. Experiments designed to assess strand bias in these errors are currently underway. What is notable, however, is the lack of TCG→TTG transitions in our dataset. This is the second most frequent Pol ε-specific mutation in the TCGA database. TCG→TTG transitions were also not found elevated in the Pol ε bMMRD brain tumor mutation spectrum. This difference may reflect interesting, but as-yet undefined tissue differences.

Another possible explanation for these differences is that the Pol ε mutants found in tumors are somehow intrinsically different biochemically from the double alanine substitution mutant used in the current study. Depending on the reporter gene used, the monoallelic Pol ε-P286R mutant is a 2.3- to 12-fold stronger mutator than the pol2-4 mutant (equivalent to the human Pol ε D275A/E277A studied here) when measured in a diploid yeast strain (*Kane and Shcherbakova, 2014*). However, a number of direct biochemical comparisons of activity and fidelity (*Figure 1C*, (*Shinbrot et al., 2014*; *Shlien et al., 2015*) and unpublished observations) between several cancer mutant constructs and the D275A/E277A construct have not yet shown any significant differences that could account for this. Certain DNA Pol mutants, including some found in human tumors, can cause increased mutagenesis by inducing expansions of normal dNTP pools in yeast and human cells (*Dennis et al., 2017*; *Mertz et al., 2015*; *Williams et al., 2015*). Interestingly, the pol2-4 allele has no effect on dNTP pools in yeast, suggesting a possible explanation for possible allelic differences with functional MMR.
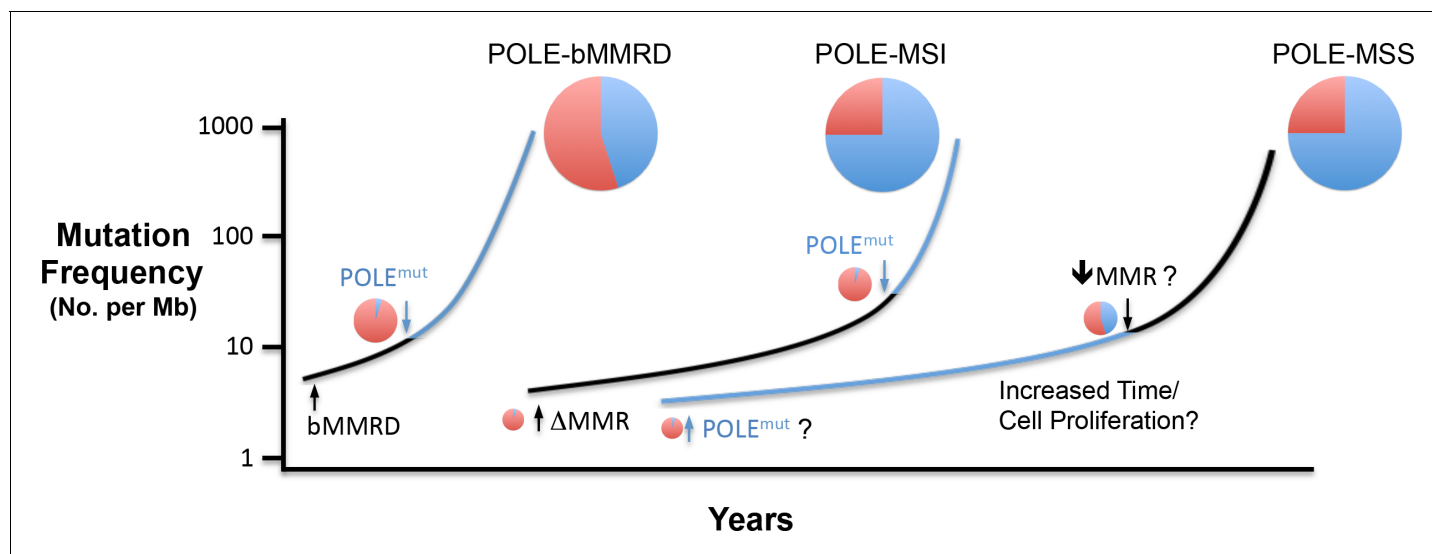
Figure 5. Model for Pol ε-dependent tumor mutation signature development. Rapid, massive mutation accumulation and Pol ε mutation signature acquisition (blue circles) depends on both Pol ε exonuclease domain mutation and compromised mismatch repair function. In somatic tumors, the partial MSI phenotype seen in a subset of POLE patients is likely the result of mismatch repair loss preceding Pol ε mutation (black line), leading to an accumulation of Pol ε-independent mutations (red circles). Mutations in bMMRD patients develop with similar mutation patterns, but accelerated timing due to germline loss of mismatch repair. When the Pol ε mutation occurs first during somatic tumor development, the mutation signature likely requires an additional characteristic for the explosive mutation acquisition to occur (blue line). Possibilities include subsequent suppression of mismatch repair (↓MMR?), unique biochemical properties (POLE^{mut}?) or increased time and or cellular proliferation.

DOI: https://doi.org/10.7554/eLife.32692.025

The following figure supplement is available for figure 5:

Figure supplement 1. Oncoprints were made using cBioPortal for colorectal (n = 8) and endometrial (n = 18) tumors from the TCGA studies containing Pol ε exonuclease domain mutations.

DOI: https://doi.org/10.7554/eLife.32692.026

In heterozygous Pol ε$^{wt/exo-}$ cells with functional mismatch repair, mutation rates were suppressed to the levels seen in cells with wild type Pol ε. These rates would be insufficient to give rise to ultra-hypermutated tumors in a matter of months. In addition, there is no explosive accumulation of triplet context-specific mutations in the MMR-proficient Pol ε$^{wt/exo-}$ cells that is seen in these tumors.

Given that the HCT-116 cells used in these studies are mutators themselves, it is possible that pre-existing deficiencies in other DNA repair or replication proteins could contribute to the observed mutagenesis. While direct contribution is unlikely given the absence of POLE mutation spectrum in the wild type Pol ε cells, cooperation with exonuclease-deficient Pol ε remains a formal possibility. To address this we used gene ontology to identify 58 DNA repair and replication proteins mutated in our HCT-116 cells, including 38 non-synonymous and 20 indel mutations. While several interesting candidates with known links to mutagenesis were identified, all have been shown by other groups to be expressed in this cell line and each, when tested, is functional (e.g. ATM, SETD2, Pol η, Pol ζ [*Bhat et al., 2013*; *Hahn et al., 2011*; *Nicolay et al., 2012*; *Zhou et al., 2013*; *Zhu et al., 2009*]). No mutation that arose during the population doubling experiments showed any obvious link to mutagenesis.

Our results support a model in which simple heterozygous loss of two Pol ε exonuclease metal chelating residues on a single allele of POLE is insufficient to drive Pol ε ultramutational specificity. Additional factors are likely required to help drive the ultramutated phenotype observed in POLE tumors, including suppression of mismatch repair, discussed below. In bMMRD, the complete lack of mismatch repair prior to Pol ε mutation leads to the moderate accumulation of Pol ε-independent replication errors (*Figure 5*). Mutation rates then increase dramatically upon loss of proofreading in one allele, with the Pol ε error signature representing a smaller fraction of the total errors, which is seen in these tumors (*Shlien et al., 2015*). Our results suggest that Pol ε mutations in somatic tumors can occur first and early, but later suppression of MMR would then accelerate overall mutation rates

to that seen in the ultramutated tumors, while the signature mutation proportion remains high (*Cancer Genome Atlas Network, 2012*; *Kandoth et al., 2013*).

Analysis of the mutational status of all mismatch repair genes in Pol ε tumors sequenced by TCGA supports the model of mismatch repair loss dramatically accelerating the acquisition of Pol ε-specific mutations. 85% (22/26) of the TCGA Pol ε tumors also have a mutation in at least one mismatch repair gene, most of which (18/22) harbor at least one nonsense mutation, which are more likely to be inactivating mutations (*Figure 5—figure supplement 1*). This predicts that at least some tumors would show evidence of MSI. In the original TCGA studies, several POLE tumors were actually first classified as MSI (three as MSI-H; five as MSI-L) (*Cancer Genome Atlas Network, 2012*; *Kandoth et al., 2013*). Analysis of sequencing reads from 46 homonucleotide runs in the POLE endometrial tumors showed no evidence of instability, so the POLE tumors were then reclassified as MSS (*Shinbrot et al., 2014*). However, the initial TCGA studies used both homo- and di-nucleotide loci to score MSI, raising the possibility that a subset of POLE tumors have a microsatellite instability defect at repeats more complex than homonucleotides. Indeed, the repeat unit size, the number of repeats and the repeat sequence composition are known to have very strong influences on the variability of microsatellite mutagenesis (*Shah et al., 2010*). Curiously, however, most (15/18) of the MMR gene nonsense mutations are the result of TCT→TAT transversions, raising the possibility that Pol ε mutation occurs first and possibly even promotes subsequent mutational inactivation of MMR.

Of all the Pol ε mutant colorectal and endometrial tumors sequenced in the TCGA studies, 15% (4/26) lacked a mutation in any mismatch repair gene and also showed no evidence of MLH1 promoter hypermethylation, demonstrating that the ultramutated phenotype can arise when mismatch repair is intact at both the genetic and epigenetic level. An alternative possibility is that mismatch repair activity is suppressed at some point during POLE tumor development. In this scenario, mutations introduced by the mutant Pol ε could accumulate slowly even in the presence of genotypically and epigenetically wild type mismatch repair. A number of conditions have been shown to transiently and reversibly lower mismatch repair protein levels and inhibit mismatch repair activity, including hypoxia, oxidative damage, inflammation, reduced pH, exposure to adriamycin or cadmium and treatment with mutagenic dNTP analogs (*Banerjee and Flores-Rozas, 2005*; *Francia et al., 2005*; *Larson and Drummond, 2001*; *Mihaylova et al., 2003*; *Chang et al., 2002*; *Hile et al., 2013*; *Iwaizumi et al., 2013*; *Lu et al., 2014*; *Negishi et al., 2002*). The variable nature and duration of such a suppression event would be expected to result in a complex effect on microsatellite instability. Perhaps even more intriguingly, transient mismatch repair suppression has been seen in the context of proofreading-deficiency in E. coli (*Fijalkowska and Schaaper, 1996*; *Schaaper and Radman, 1989*). While replication errors made by the proofreading-deficient allele tested here were clearly insufficient to suppress MMR, it is possible that the nature and rate of errors made by cancer-associated alleles might be sufficient to saturate and overwhelm MMR pathways.

Our results support the idea that loss of a single Pol ε proofreading allele is sufficient to drive a subset of the observed clinical characteristics of Pol ε tumors, provided mismatch repair is compromised in some way. These observations further support the idea that in the presence of fully functional MMR the appearance of the ultrahypermutated mutation signature may be more directly related to some as yet uncharacterized additional defect in the mutant polymerase (*Barbari and Shcherbakova, 2017*). These ideas are not mutually exclusive of one another.

Given the recent success of immune checkpoint therapies in treating tumors with high mutation burden (*Shlien et al., 2015*; *Bouffet et al., 2016*; *Hodi et al., 2010*; *Le et al., 2015*; *Santin et al., 2016*), it is of great interest to understand the mechanisms that result in ultrahypermutated tumors harboring DNA polymerase mutations.

Materials and methods

Key resources table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
cell line (Homo sapiens, Male)	HCT116 cells	Other	RRID:CVCL_0291	Prescott Deininger at Tulane Univeristy LCRC

Continued on next page

Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
cell line (H. sapiens, Male)	HCT116 + Mlh1	This paper	NA	HCT116 cells stably expressing human Mlh1-ORF via lentivirus-mediated integration
cell line (H. sapiens, Male)	Exo-; Exonuclease-deficient HCT116 Cells	This paper	NA	HCT116 cells infected with rAAV containing D275A and E277A POLE mutations
cell line (H. sapiens, Male)	Exo-; Exonuclease-deficient HCT116 Cells + Mlh1	This paper	NA	HCT116 cells stably expressing human Mlh1-ORF via lentivirus-mediated integration and infected with rAAV containing D275A and E277A POLE mutations
recombinant DNA reagent	ExoI-targeting rAAV vector	This paper	NA	Homology arms/SEPT Cassette/Exo- mutations
recombinant DNA reagent	pCMV-XL5-Mlh1	Other	NA	Victoria Belancio at Tulane Univeristy LCRC
antibody	Mlh1 Antibody	Pharmingen	G168-728; RRID: AB_395227	Rabbit monoclonal; (1:100) in Milk (1%) TBST (1X) x 1 hr at RT
chemical compound, drug	6-Thioguanine; 6-TG	Sigma-Aldrich	A4882	Used at 5 ug/mL final concentration
chemical compound, drug	Hypoxanthine-Aminopterin-Thymidine; HAT	Thermo Fisher Scientific	21060017	Used at 1X final concentration
chemical compound, drug	Geneticin; G418	Thermo Fisher Scientific	10131027	Used at 400 ug/mL final concentration
other	Ad-CMV-Cre	Vector Biolabs	1045	Adenovirus expressing Cre recombinase for excision of SEPT cassette from ExoI-targeted cell lines
software, algorithm	BWA-MEM v0.7.8	PMID: 19451168	NA	Used to align reads to human reference
software, algorithm	Picard v1.108	Broad Institute; https://broadinstitute.github.io/picard/.	NA	Identify duplicate reads
software, algorithm	The Genome Analysis Toolkit (GATK) v2.8.1	PMCID: PMC2928508	NA	locally realign reads to known indels and recalibrate base quality scores
software, algorithm	MuTect v1.1.4	PMCID: PMC3833702	NA	Identiy somatic point mutations between the tumour and matched normal
other	WES/WGS raw sequencing data	This paper	NCBI GEO Accession: PRJNA327240	Raw FASTQ files for WES and WGS performed in this study

Materials

Trypsin-EDTA was from Life Technologies and Geneticin was from Invitrogen. Antibodies against Mlh1 (mouse α-human Mlh1, G168-728) and β-actin (mouse α-human beta-actin, A1978) were from Pharmingen and Sigma, respectively.

Cell culture

The human colorectal cancer cell line HCT-116 (a kind gift from Dr. Prescott Deininger) was grown in HyClone MEM/EBSS (Thermo Scientific) supplemented with 10% fetal bovine serum (Atlanta Biologicals), 1% sodium pyruvate (Invitrogen) and 1% MEM-NEAA (Invitrogen). The HCT-116 cells used in this study were validated via analysis of genome-wide mutation signature, microsatellite instability and biomarker. HCT-116 cells lack Mlh1 resulting in a well-characterized MSI phenotype (*Lynch et al., 1993*; *Parsons et al., 1993*; *Boland and Goel, 2010*). They further have a unique mutational spectrum that can be evaluated via next-generation sequencing (*Abaan et al., 2013*). Western blot analyses (*Figure 3A*) showed a lack of Mlh1 protein. The mutation spectrum from our whole-exome sequencing of HCT-116 cells (*Figure 2A* and *Figure 2—figure supplement 1*) is identical with that reported by Abaan

(*Abaan et al., 2013*). Lastly, we performed microsatellite stability analysis in our HCT116 cells at five mononucleotide homopolymeric run loci (NR27, NR21, NR24, BAT25, BAT26) using capillary electrophoresis, which showed instability at these loci providing a phenotypic readout consistent with the lack of Mlh1 expression in our cells (data not shown). The HCT-116 cell line is also not in the 488 commonly misidentified cell lines from the most recent ICLAS database (Version 8.0) and tested negative for mycoplasma.

Generation of targeting constructs

In order to target the proofreading inactivating mutations to the POLE locus in vivo, we used rAAV with a synthetic exon promoter trap (*Rago et al., 2007*). A 1045 bp fragment containing POLE exons 7 and 8 along with intron 7 (termed HA1) was PCR amplified from HCT-116 genomic DNA using primers designed to add unique NotI and SacI sites to the 5′ and 3′ ends, respectively. A 1057 bp fragment containing exons 9, 10 and 11 along with introns 9 and 10 (termed HA2) was PCR amplified from HCT-116 genomic DNA using primers designed to add unique EcoRI and NotI sites to the 5′ and 3′ ends, respectively. Both HA1 and HA2 were first cloned into pCR-TOPO and sequence verified. The catalytic exonuclease DIE residues located in HA2 (exon 9) were changed to AIA using site-directed mutagenesis and sequence verified. The Pol ε rAAV shuttle vector was assembled by four-way ligation using the restriction enzyme-digested gene-specific HA1 and HA2 fragments, along with the SEPT/loxP cassette digested with NotI-EcoRI and the ITR-containing pAAV shuttle vector digested with NotI (SEPT/loxP cassette and pAAV shuttle vectors were kind gifts of Dr. Fred Bunz, Johns Hopkins University). The Exo-targeting vector was used to package high-titer (1.6×10^6 PFU/ml) recombinant adeno-associated virus into AAV2 serotype capsids.

Gene targeting and isolation of recombinant cell lines

Cells were grown in 100 mm dishes and infected with rAAV when ~75–80% confluent. At the time of infection, cells were washed with 1x Hanks buffered saline solution (Invitrogen) before adding 3 ml of media containing 75 µl of a 1:250 dilution of rAAV lysate. 3 hr after infection an additional 6 ml of media was added to plates and allowed to incubate at 37°C for 48 hr. After 48 hr, media was changed and Geneticin was added to a final concentration of 400 µg/ml. Plates were then incubated under selection for an additional 14 days. At the end of the selection period, colonies from plates were isolated using glass cloning rings and 0.05% trypsin (Invitrogen) was used to transfer colonies to 6-well plates for subsequent expansion. Genomic DNA was extracted from expanded clones using DNeasy Blood and Tissue kit (QIAgen) according to the manufacturer's protocol and eluted in 100 µl of elution buffer. Locus-specific integration was assessed by PCR using a primer that annealed outside the homology region and another that annealed within the *neo* cassette.

Cre-mediated excision

To remove the SEPT cassette from correctly targeted clones, cells were infected in a 25 cm² flask with adenovirus that expresses the Cre recombinase (1.0×10^6 PFU/ml, Vector Biolabs, Philadelphia, PA). Cells were plated at a limiting dilution in nonselective medium 24 hr after infection. 12 days after infection, single cell colonies were plated in duplicate and geneticin was added to one set of wells at a final concentration of 400 µg/ml to test for sensitivity. During this time, genomic DNA was extracted as previously described and screened using primers that annealed across both homology arms. PCR products were digested with SacI to distinguish between the wild type and recombinant locus.

Southern blot analysis

Genomic DNA was harvested from the knock-in cell lines using the DNeasy Blood and Tissue Kit (Qiagen), and double digested with SacI and SalI. Hoechst fluorimetry was used to determine the concentration of DNA samples for accurate loading of samples. 4 µg of each sample was run on a 0.8% agarose gel in TBE. DNA was transferred to Hybond N + membrane (Amersham), blotted with a probe to HA2 at 65°C overnight, and washed at 65°C. To make the probe, a 300 bp sequence was amplified from the HA2-pCR-TOPO clone using the primers: 5′-GCATCTGCCCCACTGTTAGT-3′ and 5′-CTCCCTGTTGGTGATGAGGT-3′. The PCR product was labeled using the Prime-It II Random Primer Labeling Kit (Agilent) and α-^{32}P-dCTP (Perkin Elmer). Membrane was blocked in Denhardt's

pre-hybridization buffer [6x SSC, 0.5% SDS, 0.1% Ficoll 70, 0.1% Ficoll 400, 0.2% PVP, and 0.2%] at 65°C for 1 hr. The probe was added to hybridization buffer [6x SSC, 0.5% SDS, and 10% Dextran Sulfate] and incubated overnight at 65°C. To wash off excess probe, the blot was washed for 2×15 min washes in wash 1 [10x SSC, 0.5% SDS], 2×15 min washes in wash 2 [1x SSC, 1% SDS], and 2×30 min washes in wash 3 [0.1x SSC, 1% SDS]. The gel was exposed to a PhosphorImage screen and scanned on a Typhoon Imager.

Purification of human Pol ε

An expression vector encoding residues 1–1189 of the catalytic subunit of human Pol ε containing the D275A/E277A substitution was prepared as described (Korona et al., 2011). Briefly, the human Pol ε was coexpressed in autoinduction medium with pRK603, which allows coexpression of TEV protease, at 25°C until the culture was saturated. Peak fractions from the HisTrap column were pooled, dialyzed into 50 mM HEPES, pH 7.5, 1 mM DTT, 5% glycerol and bound to SP sepharose. Bound protein was eluted with a 0–1 M with NaCl gradient. Peak fractions were pooled, dialyzed into 50 mM Tris, pH 7.5, 1 mM DTT, 5% glycerol, 100 mM NaCl and bound to Q Sepharose. Bound protein was eluted with a 100 mM–M M NaCl gradient. Peak fractions were pooled, concentrated and passed through a pre-equilibrated Superdex200 size exclusion column. Fractions containing the purified 140 kDa protein were pooled, dialyzed into 50 mM Tris, pH 8.0, 1 mM DTT, 5% glycerol and aliquots were frozen and stored at −80°C.

TCT→TAT in vitro error rate

We previously reported that the lacZ forward mutation assay template lacks sites at which TCT→TAT transversions are phenotypically detectable (Shlien et al., 2015). To overcome this limitation we previously made a reversion substrate that reports only this mutation by using site-directed mutagenesis to change A_{-11} to C_{-11}. Double-stranded M13mp2 DNA containing the $TC_{-11}T$ sequence was used as a substrate in reactions containing 0.15 nM DNA, 50 mM Tris-Cl, pH 7.4, 8 mM MgCl2, 2 mM DTT, 100 μg/ml BSA, 10% glycerol, 250 μM dNTPs and 1.5 nM Pol ε at 37°C. Completely filled product was transfected into Escherichia coli cells, which were used to determine the frequency of dark blue revertant plaques that occurred as a result of TCT→TAT transversions arising during DNA synthesis. In this assay, accurate DNA synthesis yields colorless plaques. Error rates were calculated according to the following equation: error rate (per nucleotide synthesized) = ((number of mutants of a particular class) × (mutant frequency)) / ((number of mutations sequenced) × (0.6) × (number of detectable sites)).

Mlh1 lentivirus construction

Mlh1 ORF was PCR amplified using the pCMV-XL5-Mlh1 vector (kindly provided by Victoria Belancio, Tulane University), forward and reverse primers (fwd 5′-TCGACTCGAGTCCACCATGTCGTTCG TGGCAGG-3′; rev 5′-TCGAGGATCCGTTACTTAACACCTCTCAAAGAC-3′) and Q5 DNA polymerase (NEB). After gel purification, dA was added to the 3′ ends with Taq and the Mlh1 ORF was cloned into pLenti6.3/V5-TOPO (Invitrogen). Mlh1 was found to have a common I219V SNP that does not affect Mlh1 function (Plotz et al., 2008). Mlh1 Lentiviral particles were made using the Vira-Power Lentiral Expression System (Invitrogen). Briefly, 293FT cells were transfected with pLenti6.3/V5-TOPO-Mlh1 and a mixture of plasmids encoding lentiviral packaging factors. Viral supernatant was harvested 48 hr after transfection, filter sterilized and stored in aliquots at −80°C. After titering, HCT-116 cells were transduced with Mlh1 lentivirus at MOI of 1.0. Cells were selected for 1 week in 10 μg/ml blasticidin. Blasticidin-resistant clones were identified and cells were harvested, lysed and probed by Western blot (mouse α-human Mlh1, G168-728, Pharmingen) to confirm Mlh1 expression.

Mutation rate and mutant frequency measurements

Prior to mutation rate measurements, preexisting HPRT1 mutants were eliminated from cell populations by incubating cells in HAT medium (1x Hypoxanthine-Aminopterin-Thymidine) for five passages. For each cell line analyzed, 500 cells were seeded and grown to confluence in 12 wells across two 6-well plates. Cells from one well were harvested and counted to estimate cell number in the remaining 11 wells. For mutation rate measurement, 500 cells from each of the remaining eleven wells were seeded per dish in 3×100 mm dishes in media lacking 6-TG to be used to measure

plating efficiency. At the same time, 5×10^5 cells from each of the remaining eleven wells were plated in 5×100 mm dishes in media containing 6-TG. After 7 days, colonies on the plating efficiency wells were stained with crystal violet and counted. After 12–14 days, the 6-TG resistant colonies were also stained with crystal violet and counted. Mutation rate was calculated using the Ma-Sandri-Sarkar Maximum Likelihood Estimator (MSS-MLE) method (*Rosche and Foster, 2000*).

For mutant frequency measurement, 500 cells per clone were seeded in duplicate in 6-well plates in media lacking 6-TG and allowed to grow for 5–7 days to determine plating efficiency. The remaining wells were seeded with 5×10^4 cells in media containing 6-TG and allowed to grow for 12–14 days. After the indicated time, colonies were stained with crystal violet and counted. Mutant frequency was calculated by the following equation: (# 6-TG resistant colonies) / ([(# colonies scored$_{PE}$)/ (# cells seeded$_{PE}$)] x (# cells seeded$_{6-TG}$)). PE refers to plating efficiency. Colonies were defined as $\geq$50 cells.

HCT116 and HCT116 + Mlh1 cells were seeded into T75 flasks and grown at 37°C/5% CO_2 until 80% confluency was reached. Cells were counted using the Countess Automated Cell Counter (Invitrogen) and 1×10^6 cells were seeded into new T75 flasks and incubated until 80% confluency was reached. The above protocol was repeated at regular intervals (3–4 days) and population doubling (PDL) numbers calculated using the following equation: PDL = [ln(N_t)-ln(N_0*PE)]/ln2. N_t = Number of viable cells counted after passage; N_0 = Number of cells seeded prior to passage; PE = plating efficiency. At PDL ~ 6, 44 and 69 cells were trypsinized and counted. For mutant frequency measurement, 300 cells were seeded into each of 3×100 mm dishes in media lacking 6-TG to be used to measure plating efficiency. Concurrently, 2×10^5 cells were seeded into each of 10×100 mm dishes in media supplemented with 6-TG to a final concentration of 5 µg/mL. After 7 days, colonies on the plating efficiency dishes were stained with crystal violet and counted. After 12–14 days, 6-TG resistant colonies were isolated using glass cloning rings and 0.05% trypsin and transferred into 24-well plates for expansion and RNA isolation. Additionally, at the above PDLs an aliquot of cells were harvested, lysed and probed by Western blot (mouse α-human Mlh1, G168-15, Abcam) to confirm maintenance of Mlh1 expression.

Genomic per base pair mutation rates (μ_{BS}) were calculated using the method of Drake (*Drake, 1991*) with modifications as applied in Lynch (*Lynch, 2010*). The equation used was: $\mu_{BS} = (\mu_L \bullet f_T \bullet f_{BS}) / (L \bullet f_L \bullet [x (n_m + n_n)/n_n])$, where μ_L is the measured mutation rate at the HPRT1 reporter gene, f_T is the fraction of mutants found after sequencing, f_{BS} is the fraction of mutations due to base pair substitutions, L is the length (in nt) of the reporter gene, f_L is the fraction of HPRT1 that gives rise to detectable mutations, x is the fraction of mutations that would give rise to chain terminator mutations, n_m is the observed number of missense mutations and n_n is the observed number of nonsense mutations. We used 126 HPRT1 mutations from three independent studies (*Bhattacharyya et al., 1995*; *Glaab and Tindall, 1997*; *Ohzeki et al., 1997*) to calculate μ_{BS}. The values used were: $f_T = 1.0$, $f_{BS} = 79/126 = 0.627$; $L = 627$ nt; $f_L = 1$; $x = 3/64 = 0.047$; $n_m = 74$; $n_n = 5$. The μ_L value for Pol ε mutant cell lines was determined empirically using fluctuation analysis.

HPRT1 sequencing

Total RNA was isolated using the Qiagen RNeasy kit (Qiagen) according to the manufacturer's protocol. RT-PCR was performed with SuperScript III Reverse Transcriptase (Invitrogen) according to the manufacturer's protocol using 1 µg of RNA as a template. Primer-specific cDNA was amplified for 32 cycles at an annealing temperature of 60°C using the following HPRT1 primers: 123(fwd) CTTCC TCCTCCTGAGCAGTC and 1041 (rev) GCCCAAAGGGAACTGATAGTC. From the HPRT1 sequencing of 6-TG resistant colonies, one clone was found to have exon 2 completely deleted. Exon deletions in HPRT1 have been shown to be caused by splice site mutations (*Bhattacharyya et al., 1995*). We therefore amplified exon 2 and its flanking region from genomic DNA prepared from the appropriate clone using the following primers: Forward: TTGTTTTCTTACATAATTCATTATCATACC; Reverse: TTACTTTGTTCTGGTCCCTACAGAG.

Whole genome and exome sequencing

Next generation sequencing was performed as per the published protocols. Whole genome sequencing (WGS) was performed on an Illumina HiSeq Xten instrument with libraries prepared using the manufacturer's TruSeq Nano DNA Library Prep kit and sequenced to a depth of 36.1x. For

exome sequencing, DNA was enriched using Agilent SureSelect Human Exome Library Preparation V5 kit, then sequenced to a depth of 101.38x (96.61x-108.19x).

Substitution detection from next generation sequenced data

All samples were processed from raw reads (FASTQ files) from paired end libraries. The reads were aligned to the human reference (GRCh37 with decoy sequences) using BWA-MEM v0.7.8 (*Li and Durbin, 2009*). Duplicate reads were identified and marked using Picard v1.108 (https://broadinstitute.github.io/picard/). The Genome Analysis Toolkit (GATK) v2.8.1 (*McKenna et al., 2010*) was used to locally realign reads to known indels and recalibrate base quality scores. Quality metrics were generated from the final BAM files to ensure high quality alignment. This includes:

- average coverage >90 x in whole exome data (*Figure 2—figure supplement 6* Mean_Coverage_Per_Sample.pdf)
- alignment rate to the reference genome >99% across whole exome data (*Figure 2—figure supplement 7* Proportion_of_properly_paired_reads.pdf) with >60M reads per sample (*Figure 2—figure supplement 8* Total_Reads_Exome.pdf)
- >90% of bases in the genome at >20 x coverage and >90% of bases in the exome at >30 x coverage (*Figure 2—figure supplement 9*)

Limitations in the genome due to low-complexity regions and incomplete areas in the genome (*Li, 2014*) prevent proper alignment resulting in sources of error.

Somatic point mutations between the tumour and matched normal were identified using MuTect v1.1.4 (*Cibulskis et al., 2013*). In addition, we used MuTect v1.1.4 in single sample mode to detect all mutations in each sample. All mutations were annotated using ANNOVAR v20130823 (*Wang et al., 2010*). Subsequent filtering was performed to reduce potential false positives and allow only high confidence mutations in the dataset using a custom R package (ShlienLab.Core.SNV v0.09). Mutations were retained if they met the following criteria:

- not identified in common mutation databases including: dbSNP (138), 1000 genomes (1000g2012feb), complete genomics (CG69), Exome sequencing project (ESP 6500si)
- for exome data, must have at least 20x normal and 30x tumour
- for WGS data, must have at least 10x normal and 10x tumour (*Figure 2—figure supplement 10*)

To investigate the quality of somatic mutations, we also identified key metrics including:

- Average alternate base quality to reference base quality of ~1.0 (*Figure 2—figure supplement 11*, mean_ratio_tumour_alt_ref_base_quality.pdf)

Data access

DNA sequencing data from this study have been submitted to the NCBI Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/) under accession number PRJNA327240.

Acknowledgements

Additional information

Author contributions

Karl P Hodel, Conceptualization, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing—original draft, Writing—review and editing; Richard de Borja, Data curation, Software, Formal analysis, Investigation, Visualization, Methodology, Writing—review and editing; Erin E Henninger, Conceptualization, Formal analysis, Investigation, Visualization, Methodology, Writing—review and editing; Brittany B Campbell, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing—review and editing; Nathan Ungerleider, Resources, Software, Formal analysis, Investigation, Visualization, Methodology; Nicholas Light, Data curation, Software, Formal analysis, Investigation, Methodology, Writing—review and editing; Tong Wu, Kimberly G LeCompte, Bruce A Bunnell, Investigation, Methodology, Writing—review and editing; A Yasemin Goksenin, Investigation, Methodology; Uri Tabori, Conceptualization, Formal analysis, Supervision, Funding acquisition, Validation, Methodology, Writing—review and editing; Adam Shlien, Conceptualization, Resources, Data curation, Software, Formal analysis, Supervision, Funding acquisition, Validation, Investigation, Visualization, Methodology, Writing—review and editing; Zachary F Pursell, Conceptualization, Resources, Formal analysis, Supervision, Funding acquisition, Validation, Investigation, Visualization, Methodology, Writing—original draft, Project administration, Writing—review and editing

Author ORCIDs

Zachary F Pursell [iD] http://orcid.org/0000-0001-5871-7192

Decision letter and Author response

Decision letter https://doi.org/10.7554/eLife.32692.033
Author response https://doi.org/10.7554/eLife.32692.034

Additional files

Supplementary files

• Transparent reporting form
DOI: https://doi.org/10.7554/eLife.32692.027

Major datasets

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database, license, and accessibility information
Hodel KP	2016	Homo sapiens Raw sequence reads - BioProject	https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA327240	Publicly available at NCBI BioProject (Accession no. PRJNA327240)

The following previously published dataset was used:

Author(s)	Year	Dataset title	Dataset URL	Database, license, and accessibility information
Abaan OD	2013	NCI-60 mutation dataset	https://cancer.sanger.ac.uk/cosmic/download	Publicly available in the Catalogue Of Somatic Mutations In Cancer (COSMIC) database at the Wellcome Sanger Institute (file labelled: 'CosmicMutantExport.tsv.gz')

References

Abaan OD, Polley EC, Davis SR, Zhu YJ, Bilke S, Walker RL, Pineda M, Gindin Y, Jiang Y, Reinhold WC, Holbeck SL, Simon RM, Doroshow JH, Pommier Y, Meltzer PS. 2013. The exomes of the NCI-60 panel: a genomic resource for cancer biology and systems pharmacology. *Cancer Research* **73**:4372–4382. DOI: https://doi.org/10.1158/0008-5472.CAN-12-3342, PMID: 23856246

Agbor AA, Göksenin AY, LeCompte KG, Hans SH, Pursell ZF. 2013. Human Pol ε-dependent replication errors and the influence of mismatch repair on their correction. *DNA Repair* **12**:954–963. DOI: https://doi.org/10.1016/j.dnarep.2013.08.012, PMID: 24051051

Aksenova A, Volkov K, Maceluch J, Pursell ZF, Rogozin IB, Kunkel TA, Pavlov YI, Johansson E. 2010. Mismatch repair-independent increase in spontaneous mutagenesis in yeast lacking non-essential subunits of DNA polymerase ε. *PLoS Genetics* **6**:e1001209. DOI: https://doi.org/10.1371/journal.pgen.1001209, PMID: 21124948

Albertson TM, Ogawa M, Bugni JM, Hays LE, Chen Y, Wang Y, Treuting PM, Heddle JA, Goldsby RE, Preston BD. 2009. DNA polymerase epsilon and delta proofreading suppress discrete mutator and cancer phenotypes in mice. *PNAS* **106**:17101–17104. DOI: https://doi.org/10.1073/pnas.0907147106, PMID: 19805137

Alexandrov LB, Nik-Zainal S, Wedge DC, Campbell PJ, Stratton MR. 2013a. Deciphering signatures of mutational processes operative in human cancer. *Cell Reports* **3**:246–259. DOI: https://doi.org/10.1016/j.celrep.2012.12.008, PMID: 23318258

Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Børresen-Dale AL, Boyault S, Burkhardt B, Butler AP, Caldas C, Davies HR, Desmedt C, Eils R, Eyjörd JE, Foekens JA, Greaves M, et al. 2013b. Signatures of mutational processes in human cancer. *Nature* **500**:415–421. DOI: https://doi.org/10.1038/nature12477, PMID: 23945592

Alexandrov LB, Stratton MR. 2014. Mutational signatures: the patterns of somatic mutations hidden in cancer genomes. *Current Opinion in Genetics & Development* **24**:52–60. DOI: https://doi.org/10.1016/j.gde.2013.11.014, PMID: 24657537

Banerjee S, Flores-Rozas H. 2005. Cadmium inhibits mismatch repair by blocking the ATPase activity of the MSH2-MSH6 complex. *Nucleic Acids Research* **33**:1410–1419. DOI: https://doi.org/10.1093/nar/gki291, PMID: 15746000

Barbari SR, Shcherbakova PV. 2017. Replicative DNA polymerase defects in human cancers: Consequences, mechanisms, and implications for therapy. *DNA Repair* **56**:16–25. DOI: https://doi.org/10.1016/j.dnarep.2017.06.003, PMID: 28687338

Bhat A, Andersen PL, Qin Z, Xiao W. 2013. Rev3, the catalytic subunit of Polζ, is required for maintaining fragile site stability in human cells. *Nucleic Acids Research* **41**:2328–2339. DOI: https://doi.org/10.1093/nar/gks1442, PMID: 23303771

Bhattacharyya NP, Ganesh A, Phear G, Richards B, Skandalis A, Meuth M. 1995. Molecular analysis of mutations in mutator colorectal carcinoma cell lines. *Human Molecular Genetics* **4**:2057–2064. DOI: https://doi.org/10.1093/hmg/4.11.2057, PMID: 8589681

Boland CR, Goel A. 2010. Microsatellite instability in colorectal cancer. *Gastroenterology* **138**:2073–2087. DOI: https://doi.org/10.1053/j.gastro.2009.12.064, PMID: 20420947

Bouffet E, Larouche V, Campbell BB, Merico D, de Borja R, Aronson M, Durno C, Krueger J, Cabric V, Ramaswamy V, Zhukova N, Mason G, Farah R, Afzal S, Yalon M, Rechavi G, Magimairajan V, Walsh MF, Constantini S, Dvir R, et al. 2016. Immune checkpoint inhibition for hypermutant glioblastoma multiforme resulting from germline biallelic mismatch repair deficiency. *Journal of Clinical Oncology* **34**:2206–2211. DOI: https://doi.org/10.1200/JCO.2016.66.6552, PMID: 27001570

Briggs S, Tomlinson I. 2013. Germline and somatic polymerase ε and δ mutations define a new class of hypermutated colorectal and endometrial cancers. *The Journal of Pathology* **230**:148–153. DOI: https://doi.org/10.1002/path.4185, PMID: 23447401

Campbell BB, Light N, Fabrizio D, Zatzman M, Fuligni F, de Borja R, Davidson S, Edwards M, Elvin JA, Hodel KP, Zahurancik WJ, Suo Z, Lipman T, Wimmer K, Kratz CP, Bowers DC, Laetsch TW, Dunn GP, Johanns TM, Grimmer MR, et al. 2017. Comprehensive analysis of hypermutation in human cancer. *Cell* **171**:1042–1056. DOI: https://doi.org/10.1016/j.cell.2017.09.048, PMID: 29056344

Cancer Genome Atlas Network. 2012. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**:330–337. DOI: https://doi.org/10.1038/nature11252, PMID: 22810696

Chang CL, Marra G, Chauhan DP, Ha HT, Chang DK, Ricciardiello L, Randolph A, Carethers JM, Boland CR. 2002. Oxidative stress inactivates the human DNA mismatch repair system. *American Journal of Physiology-Cell Physiology* **283**:C148–C154. DOI: https://doi.org/10.1152/ajpcell.00422.2001, PMID: 12055083

Church DN, Briggs SE, Palles C, Domingo E, Kearsey SJ, Grimes JM, Gorman M, Martin L, Howarth KM, Hodgson SV, Kaur K, Taylor J, Tomlinson IP, NSECG Collaborators. 2013. DNA polymerase ε and δ exonuclease domain mutations in endometrial cancer. *Human Molecular Genetics* **22**:2820–2828. DOI: https://doi.org/10.1093/hmg/ddt131, PMID: 23528559

Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, Gabriel S, Meyerson M, Lander ES, Getz G. 2013. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nature Biotechnology* **31**:213–219. DOI: https://doi.org/10.1038/nbt.2514, PMID: 23396013

Dennis DG, McKay-Fleisch J, Eitzen K, Dowsett I, Kennedy SR, Herr AJ. 2017. Normally lethal amino acid substitutions suppress an ultramutator DNA Polymerase δ variant. *Scientific Reports* **7**:46535. DOI: https://doi.org/10.1038/srep46535, PMID: 28417960

Drake JW. 1991. A constant rate of spontaneous mutation in DNA-based microbes. *PNAS* **88**:7160–7164. DOI: https://doi.org/10.1073/pnas.88.16.7160, PMID: 1831267

Drake JW. 2012. Contrasting mutation rates from specific-locus and long-term mutation-accumulation procedures. *G3: Genes|Genomes|Genetics* **2**:483–485. DOI: https://doi.org/10.1534/g3.111.001842, PMID: 22540039

Durno C, Boland CR, Cohen S, Dominitz JA, Giardiello FM, Johnson DA, Kaltenbach T, Levin TR, Lieberman D, Robertson DJ, Rex DK. 2017. Recommendations on surveillance and management of biallelic mismatch repair deficiency (BMMRD) syndrome: a consensus statement by the US multi-society task force on colorectal cancer. *Gastroenterology* **152**:1605–1614. DOI: https://doi.org/10.1053/j.gastro.2017.02.011, PMID: 28363489

Fijalkowska IJ, Schaaper RM. 1996. Mutants in the Exo I motif of Escherichia coli dnaQ: defective proofreading and inviability due to error catastrophe. *PNAS* **93**:2856–2861. DOI: https://doi.org/10.1073/pnas.93.7.2856, PMID: 8610131

Fox EJ, Salk JJ, Loeb LA. 2016. Exploring the implications of distinct mutational signatures and mutation rates in aging and cancer. *Genome Medicine* **8**:30. DOI: https://doi.org/10.1186/s13073-016-0286-z, PMID: 26987311

Francia G, Green SK, Bocci G, Man S, Emmenegger U, Ebos JM, Weinerman A, Shaked Y, Kerbel RS. 2005. Down-regulation of DNA mismatch repair proteins in human and murine tumor spheroids: implications for multicellular resistance to alkylating agents. *Molecular Cancer Therapeutics* **4**:1484–1494. DOI: https://doi.org/10.1158/1535-7163.MCT-04-0214, PMID: 16227397

Ganai RA, Osterman P, Johansson E. 2015. Yeast DNA polymerase ε catalytic core and holoenzyme have comparable catalytic rates. *Journal of Biological Chemistry* **290**:3825–3835. DOI: https://doi.org/10.1074/jbc.M114.615278, PMID: 25538242

Glaab WE, Tindall KR. 1997. Mutation rate at the hprt locus in human cancer cell lines with specific mismatch repair-gene defects. *Carcinogenesis* **18**:1–8. DOI: https://doi.org/10.1093/carcin/18.1.1, PMID: 9054582

Goldsby RE, Lawrence NA, Hays LE, Olmsted EA, Chen X, Singh M, Preston BD. 2001. Defective DNA polymerase-delta proofreading causes cancer susceptibility in mice. *Nature Medicine* **7**:638–639. DOI: https://doi.org/10.1038/88963, PMID: 11385474

Goldsby RE, Hays LE, Chen X, Olmsted EA, Slayton WB, Spangrude GJ, Preston BD. 2002. High incidence of epithelial cancers in mice deficient for DNA polymerase delta proofreading. *PNAS* **99**:15560–15565. DOI: https://doi.org/10.1073/pnas.232340999, PMID: 12429860

Hahn MA, Wu X, Li AX, Hahn T, Pfeifer GP. 2011. Relationship between gene body DNA methylation and intragenic H3K9me3 and H3K36me3 chromatin marks. *PLoS One* **6**:e18844. DOI: https://doi.org/10.1371/journal.pone.0018844, PMID: 21526191

Hanahan D, Weinberg RA. 2011. Hallmarks of cancer: the next generation. *Cell* **144**:646–674. DOI: https://doi.org/10.1016/j.cell.2011.02.013, PMID: 21376230

Haradhvala NJ, Polak P, Stojanov P, Covington KR, Shinbrot E, Hess JM, Rheinbay E, Kim J, Maruvka YE, Braunstein LZ, Kamburov A, Hanawalt PC, Wheeler DA, Koren A, Lawrence MS, Getz G. 2016. Mutational strand asymmetries in cancer genomes reveal mechanisms of DNA damage and repair. *Cell* **164**:538–549. DOI: https://doi.org/10.1016/j.cell.2015.12.050, PMID: 26806129

Hawk JD, Stefanovic L, Boyer JC, Petes TD, Farber RA. 2005. Variation in efficiency of DNA mismatch repair at different sites in the yeast genome. *PNAS* **102**:8639–8643. DOI: https://doi.org/10.1073/pnas.0503415102, PMID: 15932942

Heitzer E, Tomlinson I. 2014. Replicative DNA polymerase mutations in cancer. *Current Opinion in Genetics & Development* **24**:107–113. DOI: https://doi.org/10.1016/j.gde.2013.12.005, PMID: 24583393

Helleday T, Eshtad S, Nik-Zainal S. 2014. Mechanisms underlying mutational signatures in human cancers. *Nature Reviews Genetics* **15**:585–598. DOI: https://doi.org/10.1038/nrg3729, PMID: 24981601

Henninger EE, Pursell ZF. 2014. DNA polymerase ε and its roles in genome stability. *IUBMB Life* **66**:339–351. DOI: https://doi.org/10.1002/iub.1276, PMID: 24861832

Herr AJ, Ogawa M, Lawrence NA, Williams LN, Eggington JM, Singh M, Smith RA, Preston BD. 2011a. Mutator suppression and escape from replication error-induced extinction in yeast. *PLoS Genetics* **7**:e1002282. DOI: https://doi.org/10.1371/journal.pgen.1002282, PMID: 22022273

Herr AJ, Williams LN, Preston BD. 2011b. Antimutator variants of DNA polymerases. *Critical Reviews in Biochemistry and Molecular Biology* **46**:548–570. DOI: https://doi.org/10.3109/10409238.2011.620941, PMID: 21977975

Hile SE, Shabashev S, Eckert KA. 2013. Tumor-specific microsatellite instability: do distinct mechanisms underlie the MSI-L and EMAST phenotypes? *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **743-744**:67–77. DOI: https://doi.org/10.1016/j.mrfmmm.2012.11.003, PMID: 23206442

Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, Haanen JB, Gonzalez R, Robert C, Schadendorf D, Hassel JC, Akerley W, van den Eertwegh AJ, Lutzky J, Lorigan P, Vaubel JM, Linette GP, Hogg D, Ottensmeier CH, Lebbé C, Peschel C, et al. 2010. Improved survival with ipilimumab in patients with metastatic melanoma. *New England Journal of Medicine* **363**:711–723. DOI: https://doi.org/10.1056/NEJMoa1003466, PMID: 20525 992

Hombauer H, Campbell CS, Smith CE, Desai A, Kolodner RD. 2011. Visualization of eukaryotic DNA mismatch repair reveals distinct recognition and repair intermediates. *Cell* **147**:1040–1053. DOI: https://doi.org/10.1016/j.cell.2011.10.025, PMID: 22118461

Iwaizumi M, Tseng-Rogenski S, Carethers JM. 2013. Acidic tumor microenvironment downregulates hMLH1 but does not diminish 5-fluorouracil chemosensitivity. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **747-748**:19–27. DOI: https://doi.org/10.1016/j.mrfmmm.2013.04.006, PMID: 23643670

Jee J, Rasouly A, Shamovsky I, Akivis Y, Steinman SR, Mishra B, Nudler E. 2016. Rates and mechanisms of bacterial mutagenesis from maximum-depth sequencing. *Nature* **534**:693–696. DOI: https://doi.org/10.1038/nature18313, PMID: 27338792

Jiricny J. 2013. Postreplicative mismatch repair. *Cold Spring Harbor Perspectives in Biology* **5**:a012633. DOI: https://doi.org/10.1101/cshperspect.a012633, PMID: 23545421

Kandoth C, Schultz N, Cherniack AD, Akbani R, Liu Y, Shen H, Robertson AG, Pashtan I, Shen R, Benz CC, Yau C, Laird PW, Ding L, Zhang W, Mills GB, Kucherlapati R, Mardis ER, Levine DA, Cancer Genome Atlas Research Network. 2013. Integrated genomic characterization of endometrial carcinoma. *Nature* **497**:67–73. DOI: https://doi.org/10.1038/nature12113, PMID: 23636398

Kane DP, Shcherbakova PV. 2014. A common cancer-associated DNA polymerase ε mutation causes an exceptionally strong mutator phenotype, indicating fidelity defects distinct from loss of proofreading. *Cancer Research* **74**:1895–1901. DOI: https://doi.org/10.1158/0008-5472.CAN-13-2892, PMID: 24525744

Kennedy SR, Schultz EM, Chappell TM, Kohrn B, Knowels GM, Herr AJ. 2015. Volatility of Mutator Phenotypes at Single Cell Resolution. *PLoS Genetics* **11**:e1005151. DOI: https://doi.org/10.1371/journal.pgen.1005151, PMID: 25868109

Korona DA, Lecompte KG, Pursell ZF. 2011. The high fidelity and unique error signature of human DNA polymerase epsilon. *Nucleic Acids Research* **39**:1763–1773. DOI: https://doi.org/10.1093/nar/gkq1034, PMID: 21036870

Kunkel TA, Erie DA. 2005. DNA mismatch repair. *Annual Review of Biochemistry* **74**:681–710. DOI: https://doi.org/10.1146/annurev.biochem.74.082803.133243, PMID: 15952900

Larson ED, Drummond JT. 2001. Human mismatch repair and G*T mismatch binding by hMutSalpha in vitro is inhibited by adriamycin, actinomycin D, and nogalamycin. *Journal of Biological Chemistry* **276**:9775–9783. DOI: https://doi.org/10.1074/jbc.M006390200, PMID: 11134041

Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, Skora AD, Luber BS, Azad NS, Laheru D, Biedrzycki B, Donehower RC, Zaheer A, Fisher GA, Crocenzi TS, Lee JJ, Duffy SM, Goldberg RM, de la Chapelle A, Koshiji M, et al. 2015. PD-1 blockade in tumors with mismatch-repair deficiency. *New England Journal of Medicine* **372**:2509–2520. DOI: https://doi.org/10.1056/NEJMoa1500596, PMID: 26028255

Lee H, Popodi E, Tang H, Foster PL. 2012. Rate and molecular spectrum of spontaneous mutations in the bacterium Escherichia coli as determined by whole-genome sequencing. *PNAS* **109**:E2774–E2783. DOI: https://doi.org/10.1073/pnas.1210309109, PMID: 22991466

Li GM. 2008. Mechanisms and functions of DNA mismatch repair. *Cell Research* **18**:85–98. DOI: https://doi.org/10.1038/cr.2007.115, PMID: 18157157

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**:1754–1760. DOI: https://doi.org/10.1093/bioinformatics/btp324, PMID: 19451168

Li H. 2014. Toward better understanding of artifacts in variant calling from high-coverage samples. *Bioinformatics* **30**:2843–2851. DOI: https://doi.org/10.1093/bioinformatics/btu356, PMID: 24974202

Lu Y, Wajapeyee N, Turker MS, Glazer PM. 2014. Silencing of the DNA mismatch repair gene MLH1 induced by hypoxic stress in a pathway dependent on the histone demethylase LSD1. *Cell Reports* **8**:501–513. DOI: https://doi.org/10.1016/j.celrep.2014.06.035, PMID: 25043185

Lujan SA, Williams JS, Pursell ZF, Abdulovic-Cui AA, Clark AB, Nick McElhinny SA, Kunkel TA. 2012. Mismatch repair balances leading and lagging strand DNA replication fidelity. *PLoS Genetics* **8**:e1003016. DOI: https://doi.org/10.1371/journal.pgen.1003016, PMID: 23071460

Lujan SA, Clausen AR, Clark AB, MacAlpine HK, MacAlpine DM, Malc EP, Mieczkowski PA, Burkholder AB, Fargo DC, Gordenin DA, Kunkel TA. 2014. Heterogeneous polymerase fidelity and mismatch repair bias genome variation and composition. *Genome Research* **24**:1751–1764. DOI: https://doi.org/10.1101/gr.178335.114, PMID: 25217194

Lynch HT, Smyrk TC, Watson P, Lanspa SJ, Lynch JF, Lynch PM, Cavalieri RJ, Boland CR. 1993. Genetics, natural history, tumor spectrum, and pathology of hereditary nonpolyposis colorectal cancer: an updated review. *Gastroenterology* **104**:1535–1549. DOI: https://doi.org/10.1016/0016-5085(93)90368-M, PMID: 8482467

Lynch M. 2010. Rate, molecular spectrum, and consequences of human mutation. *PNAS* **107**:961–968. DOI: https://doi.org/10.1073/pnas.0912629107, PMID: 20080596

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation

DNA sequencing data. *Genome Research* **20**:1297–1303. DOI: https://doi.org/10.1101/gr.107524.110, PMID: 20644199

Mertz TM, Sharma S, Chabes A, Shcherbakova PV. 2015. Colon cancer-associated mutator DNA polymerase δ variant causes expansion of dNTP pools increasing its own infidelity. *PNAS* **112**:E2467–E2476. DOI: https://doi.org/10.1073/pnas.1422934112, PMID: 25827231

Mihaylova VT, Bindra RS, Yuan J, Campisi D, Narayanan L, Jensen R, Giordano F, Johnson RS, Rockwell S, Glazer PM. 2003. Decreased expression of the DNA mismatch repair gene Mlh1 under hypoxic stress in mammalian cells. *Molecular and Cellular Biology* **23**:3265–3273. DOI: https://doi.org/10.1128/MCB.23.9.3265-3273.2003, PMID: 12697826

Modrich P. 2006. Mechanisms in eukaryotic mismatch repair. *Journal of Biological Chemistry* **281**:30305–30309. DOI: https://doi.org/10.1074/jbc.R600022200, PMID: 16905530

Morrison A, Bell JB, Kunkel TA, Sugino A. 1991. Eukaryotic DNA polymerase amino acid sequence required for 3'——5' exonuclease activity. *PNAS* **88**:9473–9477. DOI: https://doi.org/10.1073/pnas.88.21.9473, PMID: 1658784

Morrison A, Johnson AL, Johnston LH, Sugino A. 1993. Pathway correcting DNA replication errors in Saccharomyces cerevisiae. *The EMBO journal* **12**:1467–1473. PMID: 8385605

Morrison A, Sugino A. 1994. The 3'–>5' exonucleases of both DNA polymerases delta and epsilon participate in correcting errors of DNA replication in Saccharomyces cerevisiae. *MGG Molecular & General Genetics* **242**:289–296. DOI: https://doi.org/10.1007/BF00280418, PMID: 8107676

Negishi K, Loakes D, Schaaper RM. 2002. Saturation of DNA mismatch repair and error catastrophe by a base analogue in Escherichia coli. *Genetics* **161**:1363–1371. PMID: 12196386

Nicolay NH, Carter R, Hatch SB, Schultz N, Prevo R, McKenna WG, Helleday T, Sharma RA. 2012. Homologous recombination mediates S-phase-dependent radioresistance in cells deficient in DNA polymerase eta. *Carcinogenesis* **33**:2026–2034. DOI: https://doi.org/10.1093/carcin/bgs239, PMID: 22822095

Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, Martincorena I, Alexandrov LB, Martin S, Wedge DC, Van Loo P, Ju YS, Smid M, Brinkman AB, Morganella S, Aure MR, Lingjærde OC, Langerød A, Ringnér M, Ahn SM, et al. 2016. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* **534**:47–54. DOI: https://doi.org/10.1038/nature17676, PMID: 27135926

Ohzeki S, Tachibana A, Tatsumi K, Kato T. 1997. Spectra of spontaneous mutations at the hprt locus in colorectal carcinoma cell lines defective in mismatch repair. *Carcinogenesis* **18**:1127–1133. DOI: https://doi.org/10.1093/carcin/18.6.1127, PMID: 9214593

Palles C, Cazier JB, Howarth KM, Domingo E, Jones AM, Broderick P, Kemp Z, Spain SL, Guarino E, Guarino Almeida E, Salguero I, Sherborne A, Chubb D, Carvajal-Carmona LG, Ma Y, Kaur K, Dobbins S, Barclay E, Gorman M, Martin L, et al. 2013. Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. *Nature Genetics* **45**:136–144. DOI: https://doi.org/10.1038/ng.2503, PMID: 23263490

Papadopoulos N, Nicolaides NC, Wei YF, Ruben SM, Carter KC, Rosen CA, Haseltine WA, Fleischmann RD, Fraser CM, Adams MD. 1994. Mutation of a mutL homolog in hereditary colon cancer. *Science* **263**:1625–1629. DOI: https://doi.org/10.1126/science.8128251, PMID: 8128251

Parsons R, Li GM, Longley MJ, Fang WH, Papadopoulos N, Jen J, de la Chapelle A, Kinzler KW, Vogelstein B, Modrich P. 1993. Hypermutability and mismatch repair deficiency in RER+ tumor cells. *Cell* **75**:1227–1236. DOI: https://doi.org/10.1016/0092-8674(93)90331-J, PMID: 8261516

Pavlov YI, Maki S, Maki H, Kunkel TA. 2004. Evidence for interplay among yeast replicative DNA polymerases alpha, delta and epsilon from studies of exonuclease and polymerase active site mutations. *BMC Biology* **2**:11. DOI: https://doi.org/10.1186/1741-7007-2-11, PMID: 15163346

Plotz G, Raedle J, Spina A, Welsch C, Stallmach A, Zeuzem S, Schmidt C. 2008. Evaluation of the MLH1 I219V alteration in DNA mismatch repair activity and ulcerative colitis. *Inflammatory Bowel Diseases* **14**:605–611. DOI: https://doi.org/10.1002/ibd.20358, PMID: 18200512

Rago C, Vogelstein B, Bunz F. 2007. Genetic knockouts and knockins in human somatic cells. *Nature Protocols* **2**:2734–2746. DOI: https://doi.org/10.1038/nprot.2007.408, PMID: 18007609

Rayner E, van Gool IC, Palles C, Kearsey SE, Bosse T, Tomlinson I, Church DN. 2016. A panoply of errors: polymerase proofreading domain mutations in cancer. *Nature Reviews Cancer* **16**:71–81. DOI: https://doi.org/10.1038/nrc.2015.12, PMID: 26822575

Roberts SA, Lawrence MS, Klimczak LJ, Grimm SA, Fargo D, Stojanov P, Kiezun A, Kryukov GV, Carter SL, Saksena G, Harris S, Shah RR, Resnick MA, Getz G, Gordenin DA. 2013. An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. *Nature Genetics* **45**:970–976. DOI: https://doi.org/10.1038/ng.2702, PMID: 23852170

Rosche WA, Foster PL. 2000. Determining mutation rates in bacterial populations. *Methods* **20**:4–17. DOI: https://doi.org/10.1006/meth.1999.0901, PMID: 10610800

Santin AD, Bellone S, Buza N, Choi J, Schwartz PE, Schlessinger J, Lifton RP. 2016. Regression of Chemotherapy-Resistant Polymerase ε (POLE) Ultra-Mutated and MSH6 Hyper-Mutated Endometrial Tumors with Nivolumab. *Clinical Cancer Research* **22**:5682–5687. DOI: https://doi.org/10.1158/1078-0432.CCR-16-1031, PMID: 27486176

Schaaper RM, Radman M. 1989. The extreme mutator effect of Escherichia coli mutD5 results from saturation of mismatch repair by excessive DNA replication errors. *The EMBO journal* **8**:3511–3516. PMID: 2555167

Shah SN, Hile SE, Eckert KA. 2010. Defective mismatch repair, microsatellite mutation bias, and variability in clinical cancer phenotypes. *Cancer Research* **70**:431–435. DOI: https://doi.org/10.1158/0008-5472.CAN-09-3049, PMID: 20068152

Shcherbakova PV, Pavlov YI. 1996. 3′–>5′ exonucleases of DNA polymerases epsilon and delta correct base analog induced DNA replication errors on opposite DNA strands in Saccharomyces cerevisiae. *Genetics* **142**:717–726. PMID: 8849882

Shcherbakova PV, Pavlov YI, Chilkova O, Rogozin IB, Johansson E, Kunkel TA. 2003. Unique error signature of the four-subunit yeast DNA polymerase epsilon. *Journal of Biological Chemistry* **278**:43770–43780. DOI: https://doi.org/10.1074/jbc.M306893200, PMID: 12882968

Shinbrot E, Henninger EE, Weinhold N, Covington KR, Göksenin AY, Schultz N, Chao H, Doddapaneni H, Muzny DM, Gibbs RA, Sander C, Pursell ZF, Wheeler DA. 2014. Exonuclease mutations in DNA polymerase epsilon reveal replication strand specific mutation patterns and human origins of replication. *Genome Research* **24**:1740–1750. DOI: https://doi.org/10.1101/gr.174789.114, PMID: 25228659

Shlien A, Campbell BB, de Borja R, Alexandrov LB, Merico D, Wedge D, Van Loo P, Tarpey PS, Coupland P, Behjati S, Pollett A, Lipman T, Heidari A, Deshmukh S, Avitzur N, Meier B, Gerstung M, Hong Y, Merino DM, Ramakrishna M, et al. 2015. Combined hereditary and somatic mutations of replication error repair genes result in rapid onset of ultra-hypermutated cancers. *Nature Genetics* **47**:257–262. DOI: https://doi.org/10.1038/ng.3202, PMID: 25642631

Simon M, Giot L, Faye G. 1991. The 3′ to 5′ exonuclease activity located in the DNA polymerase delta subunit of Saccharomyces cerevisiae is required for accurate replication. *The EMBO journal* **10**:2165–2170. PMID: 1648480

Stratton MR. 2011. Exploring the genomes of cancer cells: progress and promise. *Science* **331**:1553–1558. DOI: https://doi.org/10.1126/science.1204040, PMID: 21436442

Supek F, Lehner B. 2015. Differential DNA mismatch repair underlies mutation rate variation across the human genome. *Nature* **521**:81–84. DOI: https://doi.org/10.1038/nature14173, PMID: 25707793

Tindall KR, Glaab WE, Umar A, Risinger JI, Koi M, Barrett JC, Kunkel TA. 1998. Complementation of mismatch repair gene defects by chromosome transfer. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **402**:15–22. DOI: https://doi.org/10.1016/S0027-5107(97)00277-7, PMID: 9675233

Tran HT, Keen JD, Kricker M, Resnick MA, Gordenin DA. 1997. Hypermutability of homonucleotide runs in mismatch repair and DNA polymerase proofreading yeast mutants. *Molecular and Cellular Biology* **17**:2859–2865. DOI: https://doi.org/10.1128/MCB.17.5.2859, PMID: 9111358

Tran HT, Gordenin DA, Resnick MA. 1999. The 3′–>5′ exonucleases of DNA polymerases delta and epsilon and the 5′–>3′ exonuclease Exo1 have major roles in postreplication mutation avoidance in Saccharomyces cerevisiae. *Molecular and Cellular Biology* **19**:2000–2007. DOI: https://doi.org/10.1128/MCB.19.3.2000, PMID: 10022887

Treuting PM, Albertson TM, Preston BD. 2010. Case series: acute tumor lysis syndrome in mutator mice with disseminated lymphoblastic lymphoma. *Toxicologic Pathology* **38**:476–485. DOI: https://doi.org/10.1177/0192623310362249, PMID: 20190201

Wang K, Li M, Hakonarson H. 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research* **38**:e164. DOI: https://doi.org/10.1093/nar/gkq603, PMID: 20601685

Wheeler DA, Wang L. 2013. From human genome to cancer genome: the first decade. *Genome Research* **23**:1054–1062. DOI: https://doi.org/10.1101/gr.157602.113, PMID: 23817046

Williams LN, Herr AJ, Preston BD. 2013. Emergence of DNA polymerase ε antimutators that escape error-induced extinction in yeast. *Genetics* **193**:751–770. DOI: https://doi.org/10.1534/genetics.112.146910, PMID: 23307893

Williams LN, Marjavaara L, Knowels GM, Schultz EM, Fox EJ, Chabes A, Herr AJ. 2015. dNTP pool levels modulate mutator phenotypes of error-prone DNA polymerase ε variants. *PNAS* **112**:E2457–E2466. DOI: https://doi.org/10.1073/pnas.1422948112, PMID: 25827226

Wimmer K, Kratz CP, Vasen HF, Caron O, Colas C, Entz-Werle N, Gerdes AM, Goldberg Y, Ilencikova D, Muleris M, Duval A, Lavoine N, Ruiz-Ponte C, Slavc I, Burkhardt B, Brugieres L, EU-Consortium Care for CMMRD (C4CMMRD). 2014. Diagnostic criteria for constitutional mismatch repair deficiency syndrome: suggestions of the European consortium 'care for CMMRD' (C4CMMRD). *Journal of Medical Genetics* **51**:355–365. DOI: https://doi.org/10.1136/jmedgenet-2014-102284, PMID: 24737826

Zahurancik WJ, Baranovskiy AG, Tahirov TH, Suo Z. 2015. Comparison of the kinetic parameters of the truncated catalytic subunit and holoenzyme of human DNA polymerase ε. *DNA Repair* **29**:16–22. DOI: https://doi.org/10.1016/j.dnarep.2015.01.008, PMID: 25684708

Zhao S, Choi M, Overton JD, Bellone S, Roque DM, Cocco E, Guzzo F, English DP, Varughese J, Gasparrini S, Bortolomai I, Buza N, Hui P, Abu-Khalaf M, Ravaggi A, Bignotti E, Bandiera E, Romani C, Todeschini P, Tassi R, et al. 2013. Landscape of somatic single-nucleotide and copy-number mutations in uterine serous carcinoma. *PNAS* **110**:2916–2921. DOI: https://doi.org/10.1073/pnas.1222577110, PMID: 23359684

Zhou W, Chen YW, Liu X, Chu P, Loria S, Wang Y, Yen Y, Chou KM. 2013. Expression of DNA translesion synthesis polymerase η in head and neck squamous cell cancer predicts resistance to gemcitabine and cisplatin-based chemotherapy. *PLoS One* **8**:e83978. DOI: https://doi.org/10.1371/journal.pone.0083978, PMID: 24376779

Zhu H, Miao ZH, Huang M, Feng JM, Zhang ZX, Lu JJ, Cai YJ, Tong LJ, Xu YF, Qian XH, Ding J. 2009. Naphthalimides induce G(2) arrest through the ATM-activated Chk2-executed pathway in HCT116 cells. *Neoplasia* **11**:1226–1234. DOI: https://doi.org/10.1593/neo.09986, PMID: 19881958