A Panoply of errors: Polymerase proofreading domain mutations in cancer

Emily Rayner[1,*], Inge van Gool[2,*], Claire Palles[1], Stephen E. Kearsey[3], Tjalling Bosse[2], Ian Tomlinson[1], David N. Church[1,†]

[1] Molecular and Population Genetics Laboratory, The Wellcome Trust Centre for Human Genetics, University of Oxford. Roosevelt Drive, Oxford, OX3 7BN, UK, Tel: +44 (0)1865 287530

[2] Department of Pathology, Leiden University Medical Center, Albinusdreef 2, Postbus 9600, 2300 RC Leiden, The Netherlands. Tel +31 71 526 6625

[3] Department of Zoology, University of Oxford, South Parks Road, Oxford, OX1 3PS, UK

* denotes equal contribution

† denotes corresponding author email: dchurch@well.ox.ac.uk

Conflict of interest

None

Note

Self-archived manuscript version is that accepted following peer review, but prior to copy editing by journal Editor (i.e differs slightly from the final version published by Nature Reviews Cancer).

Abstract

Although it has long been recognised that the exonucleolytic proofreading activity intrinsic to the replicative DNA polymerases (Pols) δ and ε is essential for faithful replication of DNA, evidence that defective polymerase proofreading contributes to human malignancy has been limited. However, recent studies have shown that germline mutations in the proofreading domains of Pols δ and ε predispose to cancer, and that somatic Pol ε proofreading domain mutations occur in multiple sporadic tumours, where they underlie a phenotype of ultramutation and favourable prognosis. In this Review, we summarise current understanding of the mechanisms and consequences of polymerase proofreading domain mutations in human malignancies, and highlight the potential utility of these variants as novel cancer biomarkers.

Introduction

Accurate replication of DNA prior to cell division is a prerequisite for the suppression of mutagenesis and tumour development. The remarkable fidelity of eukaryotic DNA replication – estimated at one incorrect base for every 10^9 to 10^{10} nucleotides replicated [1] – results from a combination of highly accurate base incorporation and exonuclease proofreading by the replicative DNA polymerases (Pols) δ and ε, and post-replication surveillance by the DNA mismatch repair (MMR) apparatus [2,3]. Defects in either polymerase proofreading or MMR increase the mutation rate in yeast and cause tumours in mice [4-8]. While the importance of MMR deficiency (MMR-D) in human cancer has been recognised for more than two decades [9,10], until recently, evidence that defective polymerase proofreading contributes to human malignancy has been scarce [11]. However, over the last three years, studies have shown that germline mutations in the proofreading domains of *POLD1* and *POLE* (which encode the major subunits of Pols δ and ε respectively in humans) predispose to colorectal cancer (CRC) and other malignancies [12]. Somatic *POLE* proofreading domain mutations are found in 7-12% of endometrial cancers (ECs), 1-2% of CRCs, and occasional tumours of the breast, stomach, pancreas and brain, where they define a distinct, ultramutated tumour subgroup [13-19]. *POLE* proofreading domain mutation predicts favourable prognosis in EC[20-24], and may also do so in glioblastomas (GBMs) [14], possibly because the exceptional number of mutations in these cancers causes an enrichment of antigenic neo-peptides, leading to an anti-tumour immune response [25,26].

In this Review, we summarise current understanding of the mechanisms and consequences of polymerase proofreading domain mutations in human malignancies. Although we provide an outline of the organisation and function of replicative DNA polymerases as background, we do not cover this subject in detail as

it has been described comprehensively in several excellent reviews [27-30]. Rather, we seek to highlight the distinctive clinicopathological and molecular characteristics of polymerase proofreading domain mutant tumours, and to focus on the potential utility of these variants as novel cancer biomarkers and targets for therapy.

DNA replication and polymerase proofreading

Pols δ and ε are the principal eukaryotic DNA replicases, and together are responsible for the bulk of DNA replication, following priming by Pol α[31-34]. They are B family polymerases, and unlike Pol α have a 3'-5' exonuclease activity that proofreads the newly synthesised DNA strand [3,35]. Both Pol δ and Pol ε comprise four subunits in humans. The largest of these contains the catalytic and proofreading exonuclease active sites, and is encoded in humans by *POLD1* and *POLE* respectively [36]. The other, smaller subunits also perform several important or essential roles. In the case of Pol δ, they stabilize the holoenzyme complex and stimulate polymerase activity via interactions with the replication processivity factor PCNA [37-39]. The essential second subunit of Pol ε mediates interaction with GINS and may help target Pol ε to the leading strand during initiation of replication [40-42], while the non-essential third and fourth subunits are critical for binding dsDNA and for processive DNA synthesis and processive 3'-5' exonuclease degradation [43].

Studies of mutator polymerases in yeast have suggested a model of the replication fork in which Pol δ replicates the Okazaki fragments of the lagging strand following priming by Pol α, while Pol ε replicates the leading strand [44-46]. This division of labour has been corroborated by analysis of yeast mutants engineered to misincorporate rNTPs (reviewed in[47]), and by biochemical reconstitution experiments[48,49], and is broadly accepted. Indeed, further support for this model – albeit indirect – was

provided by the recently reported crystal structure of *S. cerevisiae* Pol ε, which revealed a domain absent in Pol δ which could explain its enhanced processivity [50]. However, these roles have been questioned by a very recent publication, which suggests that previous results may have been confused by differential MMR, and proposes that Pol δ replicates both leading and lagging strands, with the functions of Pol ε limited to repair synthesis and proofreading of the leading strand [51]. Although the precise contribution of Pol ε to leading strand replication awaits definitive clarification, current data are concordant in indicating that its exonuclease domain preferentially proofreads the leading strand [27,51]. In addition to their role in replication, both polymerases also function in base excision repair (BER) [52,53], nucleotide excision repair (NER) [54,55], mismatch repair [56-58] and double strand break repair [59,60]. Pol ε has also been implicated in cell cycle checkpoint regulation and propagation of chromatin modification states (reviewed in [61]), thus mutations affecting this protein could potentially affect a wider range of cellular activities than just replication fidelity.

The proofreading function of both Pols δ and ε requires several highly conserved exo motifs in their exonuclease domains, within which lie the catalytic site residues (D316 and E318 in Pol δ, and D275 and E277 in Pol ε in humans) essential for exonuclease activity (Figure 1*A*-*D*) [27,28]. Misincorporation of a base into the primer strand results in pausing of the polymerase, and a switch from the catalytic to the exonuclease domain [62]. The incorrect base is then excised and the correct base inserted before DNA synthesis continues [62]. Multiple studies in model organisms have confirmed the essential role of polymerase proofreading in maintenance of genomic stability. Pol δ or ε containing alanine substitutions of the exonuclease active site residues have no exonuclease activity and cells expressing these variants show a ~100-fold increase in mutation rate [63-65]. Mice harbouring the equivalent substitutions show an increased mutation rate and develop tumours [6-8]. Notably, cancers only develop in animals homozygous for proofreading-null Pol δ or ε alleles [6-8], and the tumour spectrum

differs between the two – Pol δ mutant mice develop lymphomas and carcinomas of the skin and lung [7], while Pol ε mutants are prone to intestinal tumours and histiocytic sarcomas [6]. A simple explanation for these phenotypes is that defective proofreading leads to an increased mutation rate, as some misincorporated nucleotides escape subsequent correction by MMR, but the reality may be more complicated. For instance, studies in *S. cerevisiae* indicate that elevation of dNTP levels by cell cycle checkpoint activation is responsible for the mutator phenotype of proofreading defective polymerases [66,67].

POLD1 and POLE proofreading domain mutations in human cancer

Until recently, evidence for a role of defective polymerase proofreading in human cancer has been limited [11]. However, in 2012, two exome sequencing analyses of sporadic CRCs revealed a subset of "ultramutated", yet microsatellite stable tumours, with recurrent somatic mutations within the *POLE* exonuclease domain [13,16]. The most common of these involved the replacement of proline by either arginine or histidine at codon 286 (P286R/H), and recurring substitutions were also found at codons 411 (V411L) and 459 (S459F) (Table 1) [13,16]. Shortly after this, two studies detected heterozygous somatic *POLE* proofreading domain mutations, including the P286R and V411L substitutions, in ~7% of sporadic ECs, where they were also associated with ultramutation and microsatellite stability (Table 1) [15,17]. These mutations localised to highly conserved or invariant residues within, or close to, the exo motifs essential for proofreading activity, and were predicted to perturb DNA binding by structural mapping (Figure 1*B*,*D*) [15]. In parallel with these reports, an independent study used linkage analysis and whole genome sequencing to show that families with multiple colorectal tumours, but without known predisposition mutations, carried heterozygous germline mutations in the proofreading domains of

POLD1 (S478N) and *POLE* (L424V) (Table 2)[12]. Interestingly, the *POLD1* mutation was also associated with EC. Similar to the somatic *POLE* variants, these mutations affected highly conserved residues in, or adjacent to, the exo motifs at the DNA binding interface (Figure 1*A*,*C*)[12]. Furthermore, the germline variant alleles have been shown to reduce proofreading activity and a cause a mutator phenotype in yeast [12,65].

Current data suggest that germline *POLE* and *POLD1* mutations occur in 0.5-2% of highly enriched CRC and polyposis sample sets [12,68-70]. The *POLE* L424V mutation appears to be the most common deleterious germline variant, with 21 independent carriers identified to date [69,71,72]. Although predominantly associated with CRC, this variant is now known to also predispose to EC, and may also confer moderately increased risk of many other tumour types (cancers of the breast, stomach and ovary, brain tumours, and duodenal adenomas and cancers; Table 2). The demonstration of *de novo* L424V mutations [71,72] and absence of a detectable haplotype shared among apparently unrelated families [12] suggests this is a mutational hotspot. Other recently reported germline *POLE* mutations include N363K, which maps to the exo II motif active site and is associated with CRC and pancreatic cancer [73], Y458F, which affects the exo III motif active site and predisposes to multiple tumours [74], and W347C, which lies outside the exo motifs, and has been associated with melanoma [75] (Table 2, Figure 1*B*,*D*). The first two of these variants cause strong phenotypes (malignancy in 11/12 and 9/13 cases respectively)[73,74] (Table 2). The last is of uncertain significance, as it appears to display lower penetrance, with six unaffected carriers in the family, and does not appear to confer a strong risk of CRC (<10% cases), despite evidence of pathogenicity in *S. pombe* [75]. Although no additional *POLD1* S478N carriers have been reported since 2012, five other pathogenic germline *POLD1* proofreading mutations have been identified (Table 2, Figure 1*A,C*). One of these, a recurrent

mutation at *POLD1* L474P [68,71], affects the equivalent residue to *POLE* L424V, and another, *POLD1* D316G affects the exo I motif active site [68]. The cancer risk in carriers of *POLD1* germline proofreading mutations appears to be limited to CRC and EC, with no evidence of a predisposition to duodenal or ovarian malignancies. However, caution is required, as the modest number of germline *POLD1* variant carriers makes phenotypic characterisation less certain than it is for *POLE* variant carriers. Interestingly, similar to the *POLE* active site variants, the germline *POLD1* D316G mutation appears to confer a strong phenotype (malignancy in 4/4 carriers)[68].

Somatic *POLE* proofreading domain mutations occur in ~2% of CRCs [13,16] and 7-12% of ECs [15,17,20-23], and have also been detected in ultramutated tumours of the brain, pancreas, ovary, breast and stomach [14,18,76], as well as uterine carcinosarcomas [77]. The most common variants are the P286R and V411L substitutions (Table 1, Figure 1*B*,*D*), while other recurrent substitutions include S297F, A456P and S459F. Most of these occur within, or close to the exo motifs, though unlike the germline mutations, the catalytic sites themselves are seldom affected [13,15]. In contrast to *POLE*, to date only two cases of possibly pathogenic somatic *POLD1* proofreading domain mutations have been identified. The first is a C319Y variant detected in an ultramutated paediatric glioblastoma from a patient with biallelic congenital MMR deficiency (CMMR-D) [78]. This mutation has also been detected in a myeloma with normal mutation burden [79], and its functional effect is uncertain at present. The second is a D316N substitution in a highly mutated sporadic gastric cancer [80]. This tumour displays a high proportion of indels, and as the *POLD1* variant allele fraction is only 0.13, it is unclear to what extent this substitution has contributed to the mutation burden in this tumour. The difference in the frequency of somatic proofreading mutations between *POLD1* and *POLE* is both notable and unexplained (it appears unlikely to be due to bias in the tumour types analysed to date: the exomes of more than 400 lung adenocarcinomas – a common

tumour type in Pol δ proofreading-null mice – have been sequenced at the time of writing). Given the current uncertainty regarding the roles of pols δ and ε at the replication fork [51], it will be intriguing to see whether future studies reveal to what extent, if any, this discrepancy reflects differential contributions of the two polymerases to DNA replication and repair, or other processes.

Functional insights into pathogenicity of proofreading domain mutations

The deleterious consequences of cancer-associated polymerase proofreading domain mutations on exonuclease activity and mutation rate have been confirmed by functional studies. In cell-free assays the somatic *POLE* P286R/H and S459F mutations reduce exonuclease activity to a similar extent as a proofreading null Pol ε construct with substitution of both exo I active site residues, while the somatic V411L and F367S and somatic/germline L424V variants substantially reduce, but do not eliminate exonuclease function [18]. Interestingly, the substitution equivalent to human *POLE* P286R was recently shown to confer an exceptionally strong mutator phenotype in *S. cerevisiae* [81]. This far exceeded that of the proofreading-null allele, suggesting that *POLE* P286R may exert effects beyond proofreading alone [81]. Importantly, in this study the mutation rate was also substantially increased in diploid strains heterozygous for the P286R mutation [81]. These results provide a possible explanation for why mice with proofreading-null Pol ε alleles only develop tumours in the homozygous state [6], yet somatic *POLE* proofreading mutations in human cancers appear to be mostly, if not all, heterozygous changes [15,18]. They may also explain the paucity of exonuclease active site mutations in sporadic cancers. Given these exciting data, the results of similar analyses of other cancer-associated mutations are eagerly awaited.

Clinicopathological characteristics of polymerase proofreading domain-mutant cancers

Polyposis appears to be more severe in carriers of germline *POLE* mutations than those with *POLD1* variants (10-50 vs. <20 colonic polyps per patient), though in both cases the polyps themselves are typically of unremarkable histology. Similarly, with the exception of young age at onset (typically <50 years), CRCs and other malignancies in these patients do not appear to display any distinguishing features, though numbers are limited at present. Tumours with somatic *POLE* proofreading domain mutations also show a strong tendency to occur in young patients, and, where examined, typically exhibit several other notable characteristics, including an association with high tumour grade, a prominent lymphocytic infiltrate, and the presence of multiple giant cells of bizarre morphology [14,25,82].

An association of somatic *POLE* proofreading domain mutations with good prognosis was first suggested by the TCGA EC study [17]. This has since been confirmed in EC by several groups [20-24], and was also proposed for glioblastomas by a recent report [14]. As many patients in these series received post-operative radiotherapy, either alone or in combination with chemotherapy, it is currently not possible to draw definitive conclusions on whether *POLE* mutation is a prognostic or predictive biomarker, or indeed both. However, the lack of recurrences of high-grade *POLE* proofreading domain-mutant ECs in two large clinical trials in which chemotherapy was not used, and in a subset of patients who did not receive radiotherapy in one of these studies [20], suggests that *POLE* mutation may portend good prognosis in the absence of post-operative treatment. While this is reminiscent of the favourable clinical outcome of hypermutated early-stage MMR-D CRCs [83,84], it should be noted

that the impact of mismatch repair deficiency on EC prognosis is unclear [85-87], and the prognostic import of *POLE* mutations in CRC awaits confirmation.

Molecular characteristics of polymerase proofreading domain-mutant tumours

Although the number of tumours from germline *POLD1* and *POLE* mutations analysed to date is relatively modest, most have been found to be microsatellite stable [12], though microsatellite instability has been noted in some cases [72]. Analysis of mutation spectrum and context has been limited to the tumours from *POLE* L424V and *POLD1* S478N carriers, which revealed a phenotype of base substitutions and missense substitutions with relatively few frameshift mutations [12]. As noted earlier, the most striking feature of tumours with somatic *POLE* proofreading domain mutations is their exceptional mutation burden [13,16,18,19]. Similar to the tumours from germline *POLE* variant carriers, these are predominantly missense mutations [13,15,16,18,19], and occur within a unique mutation signature, with a 100-fold increase in C>A transversions in the context TCT and 30-fold increase in C>T transitions in the context of TCG [15,18,19]. This results in a strong bias for particular amino acid changes, with an overrepresentation of serine to tyrosine/leucine, and arginine to isoleucine/glutamine substitutions, and a substantial increase in glutamic acid to stop codon mutations [18]. These are manifest as a distinctive pattern of substitution and truncation mutations in oncogenes and tumour suppressors, including *PIK3CA* (R88Q), *PTEN* (R130Q), and *APC* (R1114X, Q1338X), *MSH6* (E946X, E1322X), *FBXW7* (E369X) and *TP53* (R213X) [12,13,15,16,18,88]

Sporadic *POLE* proofreading domain mutant cancers display few copy number alterations (CNAs) and, similar to tumours in germline variant carriers, a strong tendency to microsatellite stability, despite frequent mutations in MMR genes [13,15-17].

Accumulating evidence suggests that the interaction between defective polymerase proofreading and MMR is complex and may depend on the extent to which the function of each is compromised. For example, while the combination of error-prone polymerases and partial MMR function causes attenuated growth in yeast [89], the combination with complete MMR loss is synthetically lethal in *S. cerevisiae* and mice [6,89,90]. Interestingly, a recent study of patients with CMMR-D who developed glioblastomas showed that these tumours occurred following the acquisition of polymerase-proofreading domain mutations and were ultramutated, yet predominantly microsatellite stable [78]. Most of the germline MMR mutations in these cases involved either *PMS2* or *MSH6*, while sporadic polymerase proofreading domain tumours tend to acquire *MSH6* mutations [13,17,78]. Loss of MSH6 or PMS2 generally causes a milder mutator phenotype, sometimes lacking MSI, than loss of other MMR components such as MSH2 and MLH1 [89,91]. While it is tempting to speculate that the combination of defective proofreading and profound MMR deficiency may result in a mutation rate that exceeds the optimum for tumour fitness [92,93], empirical verification of this is currently lacking, and it must be noted that though rare, tumours with polymerase proofreading domain mutations and MSI have been reported [21,72,78]. It will be of particular interest to determine the timing and clonality of both events in these cancers, and to examine whether these tumours harbour secondary 'antimutator' mutations that permit continued viability, as has been demonstrated in yeast [89,94].

Challenges in determining pathogenicity of POLD1 and POLE proofreading domain mutations

While the number of confirmed germline *POLD1* and *POLE* proofreading domain mutations has grown since their initial report in 2012 (Table 2, also curated in the

Leiden Open Variation Database (LOVD): http://www.lovd.nl/3.0/home), differentiation of pathogenic from non-pathogenic variants in a patient with suggestive clinical features and family history remains challenging. The Exome Aggregation Consortium (ExAC) database currently lists 75 and 56 missense or loss of function variants mapping to the exonuclease domains of *POLE* and *POLD1* respectively, all of which have a population frequency of <1%. Most of these are likely to have no impact on proofreading function. Indeed, of the germline variants reviewed here only *POLD1* D316H is present in ExAC, at a frequency of 0.001% (1/ 98012 alleles). Therefore, filtering variants according to their presence in ExAC and similar databases will be of limited utility, as it will only exclude variants with *relatively* high frequencies, such as those >0.1%. Instead we suggest that the following criteria should ideally be satisfied to prove pathogenicity of a variant: i) segregation with affection status in pedigrees; ii) conservation of the affected residue between human polymerases and those from other species; iii) evidence of functional effect from at least one of the following: analysis of corresponding residue in model organisms; cell-free exonuclease or other polymerase assays [18,95], and; iv) sequencing of tumour DNA demonstrates elevated frequency of base substitutions .

Confirming the pathogenicity of somatic polymerase proofreading domain mutations may also be difficult. Both *POLD1* and *POLE* are large genes and are likely to acquire somatic mutations secondary to other causes of increased mutation burden, such as MMR deficiency. Given the association of somatic *POLE* proofreading domain mutations with prognosis [14,17,20-24], it is important to differentiate these pathogenic variants from passenger mutations of no functional significance, particularly given examples in the recent literature where a functional role has been inferred for *POLE* and *POLD1* variants of uncertain pathogenicity [96]. Based on analysis of TCGA data, Shinbrot and colleagues [18] have proposed criteria to identify bona fide pathogenic somatic *POLE* proofreading mutations. These include: i)

ultramutation (often exceeding 100 mutations/Mb); ii) increased proportion of C>A transversions, exceeding 20% of all substitutions (though it should be noted that the most common substitutions are typically C>T transitions) ; iii) *POLE* mutation at residue that is recurrently mutated in cancer (Figure 1A, Table 1). While most *POLE* proofreading domain-mutant cancers display these characteristics, we advocate a degree of flexibility in their application, as the predominant mutation caused by proofreading-deficient human Pol ε in vitro is T>A transversions [97], and it is possible that novel pathogenic *POLE* proofreading domain mutations could cause a different mutational signature. We also suggest that consideration should also be given to iv) preponderance of missense mutations (i.e relative absence of indels); v) presence of flanking nucleotide bias[15,18,19]; and vi) relative lack of copy number alteration[15,17]. While the co-existence of microsatellite instability does not preclude the presence of a deleterious polymerase proofreading variant, it should prompt careful evaluation of its pathogenicity, ideally using bioinformatic predictors (e.g. MutationTaster, SIFT, Polyphen) and structural mapping using the conserved yeast Pol ε structure. Although some tumours carrying bona fide pathogenic *POLE* proofreading domain mutations will fall outside this classification [21], we believe it represents a reasonable starting point that can be refined and improved as further data are accumulated in the future.

Timing of polymerase proofreading domain mutations in tumourigenesis

The association of germline variants with malignancy [12] suggests that polymerase proofreading domain mutations are able to initiate cancer, and current data, while limited, are consistent with somatic *POLE* mutations occurring as an early event in sporadic tumours. Analysis of variant allele fraction in glioblastomas suggests that *POLE* mutations are present in the earliest persisting clone [14,78], and genes in which driver mutations are known to occur early in CRC and EC frequently display

evidence of the *POLE* mutation signature [13-18]. Interestingly, analysis of tumours from CMMR-D patients suggests that acquisition of somatic pathogenic polymerase proofreading domain mutations is associated with rapid tumour growth [78], consistent with the prediction of mathematical models that a strong mutator phenotype confers a preferential advantage in early tumourigenesis [1,98].

Possible explanations for the good prognosis of polymerase proofreading domain-mutant cancers

As noted previously, in addition to ultramutation and good prognosis, *POLE* proofreading domain-mutant tumours frequently display a prominent lymphocytic infiltrate [14,25,82], similar to that observed in MMR-D CRCs[99]. Recent characterisation of this in EC has shown that it represents a population of $CD8^+$ cytotoxic lymphocytes, and is accompanied by increased expression of cytotoxic lymphocyte differentiation markers (T-bet, Eomes) and effectors (IFNG, perforin, and granzyme B) compared to other ECs [25]. Bioinformatic analysis revealed that the number of 'antigenic mutations' – that is, mutations in expressed genes that encode neo-peptides predicted to bind MHC molecules – is substantially greater in *POLE*-mutant tumours than other ECs, providing a potential explanation for these results [25,26]. Interestingly, *POLE* tumours displayed increased expression of immunosuppressive checkpoints, including PD1, PD-L1, CTLA4, LAG3, TIM-3, and TIGIT [25,26], suggesting that upregulation of these molecules may be required for tumour growth. It will be important to investigate whether these findings are observed in *POLE* proofreading domain-mutant tumours of other tissues, and to investigate whether tumours from carriers of germline polymerase proofreading variants display evidence of enhanced immunogenicity that may impact on their prognosis.

Another possible contributor to the favourable outcome of polymerase proofreading domain-mutant tumours relates directly to their mutator phenotype. Studies of mutator Pol δ and ε alleles in yeast have demonstrated the existence of an 'error threshold' which, if exceeded, results in reduced viability [89,90,94]. The combination of defective polymerase proofreading and complete loss of MMR is lethal as it exceeds this threshold [89,94]. Interestingly, sequential biopsy of cancers in CMMR-D patients demonstrated that while the acquisition of somatic polymerase proofreading domain mutations was associated with dramatic accumulation of mutations and rapid tumour growth, the total mutation burden subsequently appeared to plateau [78]. The upper limit of 10,000–20,000 exonic mutations suggested by that study is highly concordant with the number observed in sporadic *POLE*-mutant adult cancers [19]. Tumour mutation burden is not simply a reflection of mutation rate, and the relationship between the two in polymerase-mutant tumours in uncertain. Nevertheless, it may be speculated that while the proofreading-deficient mutator phenotype confers a growth advantage in early tumourigenesis, it compromises fitness in later stage cancers. This may be explained by an intriguing recent study, which used mathematical modelling to predict the effect of passenger mutations on (non-hypermutated) tumour fitness [100]. This showed that while strongly deleterious passengers are subject to negative selection, those with individually moderate negative effects are able to fixate, and collectively are predicted to exert a negative impact on tumour growth [100]. As this model was based on tumours with only relatively small numbers of protein-coding mutations (average 10 to 366 per tumour type), it will be interesting to examine the effect of passenger mutations on the fitness of polymerase proofreading domain-mutant tumours.

Clinical management of polymerase proofreading domain-mutant cancers

The PPAP phenotype overlaps with that of Lynch Syndrome and *MUTYH*-associated polyposis, and screening and management algorithms are broadly similar. Valle and colleagues [71] have proposed that screening of the *POLD1* and *POLE* exonuclease domains is indicated in patients with polyposis (10-100 adenomas) and/or early-onset CRC (diagnosis at age <50), who lack germline mutations in MMR genes, *APC* or *MUTYH*. The concomitant presence of extracolonic tumours (Table 2), particularly EC and stomach/duodenal tumours, should increase clinical suspicion of germline *POLD1*/*POLE* mutations. However, at present there is insufficient evidence to recommend screening of patients who lack colonic phenotypes. Similar to patients with Lynch syndrome [101], carriers of germline *POLD1* and *POLE* mutations should be offered colonoscopies at 1-2 year intervals from age 25 and upper GI endoscopy to check for duodenal tumours (particularly in carriers of *POLE* variants). Although EC screening is not of proven benefit, women might pragmatically be offered this from age 40 and clinicians should be aware of the potential increased risk of ovarian, brain and breast cancers in these patients, with consideration given to preventive measures, where available.

The distinctive characteristics of tumours with somatic *POLE* proofreading domain mutations also have potential implications for patient management. Most notably, the association with good prognosis in EC [17,20-24] suggests that *POLE* mutation might identify a group of patients who are less likely to benefit from adjuvant treatment following surgery. This is clinically relevant, as mutations are more common in tumours defined as 'high risk' by conventional criteria [20,22], for which postoperative radiotherapy and chemotherapy are often recommended. However, as noted earlier, the possibility that the favourable outcome of these tumours reflects increased sensitivity to treatment cannot be excluded at present, and further preclinical and clinical studies will be required before *POLE* mutations can be used to guide

management in EC. Similar studies will also be needed to confirm the utility of *POLE* as a biomarker in other cancer types.

The remarkable mutation burden of *POLE*-mutant tumours also raises the possibility that they may be particularly sensitive to specific therapeutic strategies. Perhaps most obvious of these is the use of immune checkpoint inhibitors that target immunosuppressive molecules including PD1 and PD-L1. These agents have recently demonstrated striking activity against highly mutated MMR-deficient CRCs, melanomas and non-small cell lung cancers (NSCLCs) [102-105], where response appears to correlate with increased tumour antigenic mutation burden [96,106] and the presence of a pre-treatment intratumoural $CD8^+$ / $PD\text{-}L1^+$ lymphocyte infiltrate [107]. Both of these are prominent features of *POLE* proofreading domain-mutant tumours [25,26], suggesting that these cancers may be excellent candidates for these drugs. Furthermore, the recent demonstration that immune checkpoint inhibition may be potentiated by radiation [108], suggests that investigation of these combinations against polymerase mutant cancers in preclinical models and clinical trials may be worthwhile. Another potential therapeutic strategy against polymerase proofreading domain-mutant cancers relates to the concept of the error threshold discussed earlier. In theory, agents such as mutagenic nucleosides or inhibitors of DNA repair [109] could be used to increase the mutation rate in these tumours to a level that exceeds this, resulting in lethal mutagenesis and loss of viability [93]. Clearly, in the first instance such a strategy would only be appropriate in patients with incurable disease lacking other treatment options, though in selected cases this may be worthy of exploration [93]. Preclinical studies suggest that a similar effect may result from modification of the dNTP pool [67,110], though differences in nucleotide synthesis and the DNA damage response between yeast and humans mean further work is required before the possible utility of this approach in humans can be predicted.

Given the modest frequency of polymerase proofreading domain mutations overall, any therapeutic study is likely to have to recruit patients with multiple tumour types (a design frequently referred to as a basket trial [111]), or combine proofreading-deficient tumours with other hypermutated cancers, such as those with defective MMR.

Conclusions and future directions

The longstanding postulate that defective polymerase proofreading may contribute to human cancer [112] has been proven only recently. Nevertheless, it is now clear that germline mutations in the exonuclease domains of *POLD1* and *POLE* predispose to polyposis, CRC and other malignancies [12,68,69], and that somatic *POLE* mutations cause ultramutation in sporadic ECs, CRCs and several other cancers [13-19]. In the latter case, the exceptional mutation load in *POLE*-mutant ECs is associated with an enhanced immune response [25,26] and an excellent prognosis [17,20-24]. The possibility that mutation rate in these tumours approaches the maximum compatible with continued viability is an intriguing one, the investigation of which may provide novel insights into the consequences of a mutator phenotype in cancer. It will also be of interest to determine whether the ultramutator phenotype in polymerase proofreading domain-mutant tumours represents an Achilles' heel that can be exploited for therapy, as has recently been suggested [78].

From a clinical perspective, PPAP should be considered and tested for in patients with unexplained polyposis and/or early onset CRC, particularly where family members have EC or other extracolonic cancers suggestive of germline *POLD1* or *POLE* mutations [71]. Meanwhile, somatic polymerase proofreading domain mutations exemplify the challenge of implementing precision medicine. For example, the modest frequency of *POLE* mutations in EC (7-12%) has limited the power of some

studies to evaluate their impact on prognosis [21], a problem that is likely to prove even greater in other tumours where *POLE* mutations are less common. As most novel cancer variants occur at a similarly low frequency (<10%), evaluation of these as prognostic and predictive biomarkers will require large-scale collaborations and coordinated analysis. At a more basic level, the results of studies performed to date pose several fundamental questions. For example, why are somatic mutations largely restricted to *POLE*, when germline variants in both *POLE* and *POLD1* cause cancer? Why are substitutions of the exo motif catalytic sites proportionally more common among germline than somatic proofreading domain mutations? Why does the mutator phenotype of *POLE* P286R exceed that of the exonuclease null allele? Determining the answers to these will be a priority for future studies.

During the last three years, defects in polymerase proofreading have been recognised to drive tumourigenesis in a small, but important fraction of common cancers. Given the rapid progress in the field, we are optimistic that the next three years will see similar advances in our understanding of this novel tumour subgroup, with consequent benefits for patients.

Resources cited:

Leiden Open Variation Database (LOVD): http://www.lovd.nl/3.0/home

MutationTaster: http://www.mutationtaster.org

SIFT: http://sift.jcvi.org

Polyphen: http://genetics.bwh.harvard.edu/pph2/

Funding:

Work in the host laboratories is supported by an Academy of Medical Sciences / Health Foundation Clinician Scientist Fellowship Award (to DC), Cancer Research UK (Programme grant C6199/A10417; to IPMT), the European Research Council (EVOCAN Grant agreement: 340560 to IMPT), the Dutch Cancer Society (Grant ref: UL2012-5719 to TB) and the Medical Research Council (grant ref MR/L016591/1 to SEK) and core funding to the Wellcome Trust Centre for Human Genetics from the Wellcome Trust (Ref: 090532/ Z/09/Z). DC has received funding from the Oxford Cancer Centre.

Acknowledgements:

We would like to thank Jon Grimes for help in generating the Pol δ and ε structures and Niels de Wind for helpful comments on the manuscript.

References:

1. Loeb, L. A. Mutator Phenotype May Be Required for Multistage Carcinogenesis. *Cancer Res.* **51**, 3075-3079 (1991).
2. Kunkel, T. A. DNA Replication Fidelity. *J. Biol. Chem.* **279**, 16895-16898 (2004).
3. Loeb, L. A. & Monnat, R. J. DNA polymerases and human disease. *Nat. Rev. Gen.* **9**, 594-604 (2008).
4. Morrison, A., Johnson, A. L., Johnston, L. H. & Sugino, A. Pathway correcting DNA replication errors in Saccharomyces cerevisiae. *EMBO J.* **12**, 1467-1473 (1993).
5. Edelmann, W. et al. Mutation in the mismatch repair gene Msh6 causes cancer susceptibility. *Cell* **91**, 467-477 (1997).
6. Albertson, T. M. et al. DNA polymerase and proofreading suppress discrete mutator and cancer phenotypes in mice. *Proc. Natl. Acad. Sci. USA* **106**, 17101-17104 (2009).
7. Goldsby, R. E. et al. High incidence of epithelial cancers in mice deficient for DNA polymerase delta proofreading. *Proc. Natl. Acad. Sci. USA* **99**, 15560-15565 (2002).
8. Goldsby, R. E. et al. Defective DNA polymerase-delta proofreading causes cancer susceptibility in mice. *Nat. Med.* **7**, 638-639 (2001).
9. Fishel, R. et al. The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell* **75**, 1027-1038 (1993).
10. Aaltonen, L. A. et al. Clues to the pathogenesis of familial colorectal cancer. *Science* **260**, 812-816 (1993).
11. Yoshida, R. et al. Concurrent genetic alterations in DNA polymerase proofreading and mismatch repair in human colorectal cancer. *Eur. J. Hum. Gen.* **19**, 320-325 (2011).
12. Palles, C. et al. Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. *Nat. Gen.* **45**, 136-144 (2012).
13. The Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330-337 (2012).
14. Erson-Omay, E. Z. et al. Somatic POLE mutations cause an ultramutated giant cell high-grade glioma subtype with better prognosis. *Neuro-Oncol.* 1-9 (Advance online publication March 2015).

15. Church, D. N. et al. DNA polymerase epsilon and delta exonuclease domain mutations in endometrial cancer. *Hum. Mol. Genet.* **22**, 2820-2828 (2013).
16. Seshagiri, S. et al. Recurrent R-spondin fusions in colon cancer. *Nature* **488**, 660-664 (2012).
17. The Cancer Genome Atlas Network. Integrated genomic characterization of endometrial carcinoma. *Nature* 497, 67-73 (2013).
18. Shinbrot, E. et al. Exonuclease mutations in DNA polymerase epsilon reveal replication strand specific mutation patterns and human origins of replication. *Genome Res.* **24**, 1740-1750 (2014).
19. Alexandrov, L. B. et al. Signatures of mutational processes in human cancer. *Nature* **500**, 415-421 (2013).
20. Church, D. N. et al. Prognostic Significance of POLE Proofreading Mutations in Endometrial Cancer. *J. Nat. Can. Inst.* **107**, 1-8 (2015).
21. Billingsley, C. C. et al. Polymerase ε (POLE) mutations in endometrial cancer: Clinical outcomes and implications for Lynch syndrome testing. *Cancer* **121**, 386-94 (2014).
22. Stelloo, E. et al. Refining prognosis and identifying targetable pathways for high-risk endometrial cancer; a TransPORTEC initiative. *Mod. Pathol.* **28**, 836-844 (2015).
23. Talhouk, A. et al. A clinically applicable molecular-based classification for endometrial cancers. *Br. J. Cancer* **113**, 299-310 (2015).
24. Meng, B. et al. POLE exonuclease domain mutation predicts long progression-free survival in grade 3 endometrioid carcinoma of the endometrium. *Gynecol. Oncol.* **134**, 15-19 (2014).
25. van Gool, I. C. et al. POLE proofreading mutations elicit an anti-tumor immune response in endometrial cancer. *Clin. Cancer Res.* **14**, 3347-3355 (2015).
26. Howitt, B. E. et al. Association of POLE-mutated and MSI endometrial cancers with an elevated number of tumor-infiltrating and peritumoral lymphocytes and higher expression of PD-L1. *JAMA Oncol.* **33** (Advance Online Publication, July 2015).
27. Henninger, E. E. & Pursell, Z. F. DNA polymerase ε and its roles in genome stability. *IUBMB Life* **66**, 339-351 (2014).
28. Prindle, M. J. & Loeb, L. A. DNA polymerase delta in DNA replication and genome maintenance. *Env. Mol. Mutagen.* **53**, 666-682 (2012).
29. Pavlov, Y. I. & Shcherbakova, P. V. DNA polymerases at the eukaryotic fork—20 years later. *Mut. Res.* **685**, 45-53 (2010).

30. Johansson, E. & Dixon, N. Replicative DNA polymerases. *Cold Spring Harb. Perspect. Biol.* **5**, a012799 (2013).
31. Morrison, A., Araki, H., Clark, A. B., Hamatake, R. K. & Sugino, A. A third essential DNA polymerase in S. cerevisiae. *Cell* **62**, 1143-1151 (1990).
32. Budd, M. & Campbell, J. L. Temperature-sensitive mutations in the yeast DNA polymerase I gene. *Proc. Natl. Acad. Sci. USA* **84**, 2838-2842 (1987).
33. Budd, M. E. & Campbell, J. L. DNA polymerases delta and epsilon are required for chromosomal replication in Saccharomyces cerevisiae. *Mol. Cell Biol.* **13**, 496-505 (1993).
34. Francesconi, S., Park, H. & Wang, T. S. Fission yeast with DNA polymerase delta temperature-sensitive alleles exhibits cell division cycle phenotype. *Nucleic Acids Res.* **21**, 3821-3828 (1993).
35. Bernad, A., Blanco, L., Lázaro, J., Martín, G. & Salas, M. A conserved 3′→5′ exonuclease active site in prokaryotic and eukaryotic DNA polymerases. *Cell* **59**, 219-228 (1989).
36. Shevelev, I. V. & Hübscher, U. The 3' 5' exonucleases. *Nat. Rev. Mol. Cell Biol.* **3**, 364-376 (2002).
37. Baranovskiy, A. G. et al. X-ray structure of the complex of regulatory subunits of human DNA polymerase delta. *Cell Cycle* **7**, 3026-3036 (2008).
38. Podust, V. N., Chang, L. S., Ott, R., Dianov, G. L. & Fanning, E. Reconstitution of human DNA polymerase delta using recombinant baculoviruses: the p12 subunit potentiates DNA polymerizing activity of the four-subunit enzyme. *J. Biol. Chem.* **277**, 3894-3901 (2002).
39. Zhou, Y., Meng, X., Zhang, S., Lee, E. Y. & Lee, M. Y. Characterization of human DNA polymerase delta and its subassemblies reconstituted by expression in the MultiBac system. *PLoS One* **7**, e319156 (2012).
40. Garbacz, M. et al. Fidelity consequences of the impaired interaction between DNA polymerase epsilon and the GINS complex. *DNA Repair* **29**, 23-35 (2015).
41. Langston, L. D. et al. CMG helicase and DNA polymerase epsilon form a functional 15-subunit holoenzyme for eukaryotic leading-strand DNA replication. *Proc. Natl. Acad. Sci. USA* **111**, 15390-15395 (2014).
42. Sengupta, S., van Deursen, F., de Piccoli, G. & Labib, K. Dpb2 integrates the leading-strand DNA polymerase into the eukaryotic replisome. *Curr. Biol.* **23**, 543-552 (2013).

43. Aksenova, A. et al. Mismatch Repair–Independent Increase in Spontaneous Mutagenesis in Yeast Lacking Non-Essential Subunits of DNA Polymerase ε. *PLoS Gen.* **6**, e1001209 (2010).
44. Pursell, Z. F., Isoz, I., Lundström, E.-B., Johansson, E. & Kunkel, T. A. Yeast DNA polymerase epsilon participates in leading-strand DNA replication. *Science* **317**, 127-130 (2007).
45. McElhinny, S. A., Gordenin, D. A., Stith, C. M., Burgers, P. M. J. & Kunkel, T. A. Division of labor at the eukaryotic replication fork. *Mol. Cell* **30**, 137-144 (2008).
46. Miyabe, I., Kunkel, T. A. & Carr, A. M. The Major Roles of DNA Polymerases Epsilon and Delta at the Eukaryotic Replication Fork Are Evolutionarily Conserved. *PLoS Gen.* **7**, e1002407 (2011).
47. Vazquez, E. & Antequera, F. Replication dynamics in fission and budding yeasts through DNA polymerase tracking. *Bioessays* (Advance online publication August 2015).
48. Georgescu, R. E. et al. Reconstitution of a eukaryotic replisome reveals suppression mechanisms that define leading/lagging strand operation. *Elife* **4**, e04988 (2015).
49. Georgescu, R. E. et al. Mechanism of asymmetric polymerase assembly at the eukaryotic replication fork. *Nature Struct. Mol. Biol.* **21**, 664-670 (2014).
50. Hogg, M. et al. Structural basis for processive DNA synthesis by yeast DNA polymerase epsilon. *Nature Struct. Mol. Biol.* **21**, 49-55 (2014).
51. Johnson, Robert E., Klassen, R., Prakash, L. & Prakash, S. A Major Role of DNA Polymerase δ in Replication of Both the Leading and Lagging DNA Strands. *Mol. Cell* **59**, 163-175 (2015).
52. Blank, A., Kim B Fau - Loeb, L. A. & Loeb, L. A. DNA polymerase delta is required for base excision repair of DNA methylation damage in Saccharomyces cerevisiae *Proc. Natl. Acad. Sci. USA* **91**, 9047-9051 (1994).
53. Stucki, M. et al. Mammalian base excision repair by DNA polymerases delta and epsilon. *Oncogene* **17**, 835-843 (1998).
54. Nishida, C., Reinhard P Fau - Linn, S. & Linn, S. DNA repair synthesis in human fibroblasts requires DNA polymerase delta. *J. Biol. Chem.* **265**, 501-510 (1988).
55. Lehmann, A. R. DNA polymerases and repair synthesis in NER in human cells. *DNA Repair* **10**, 730-733 (2011).
56. Tran, H. T., Gordenin, D. A. & Resnick, M. A. 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. *Mol. Cell Biol.* **19**, 2000-2007 (1999).

57. Bowen, N. et al. Reconstitution of long and short patch mismatch repair reactions using Saccharomyces cerevisiae proteins. *Proc. Natl. Acad. Sci. USA* **110**, 18472-18477 (2013).
58. Zhang, Y. et al. Reconstitution of 5'-directed human mismatch repair in a purified system. *Cell* **122**, 693-705, (2005).
59. Costantino, L. et al. Break-induced replication repair of damaged forks induces genomic duplications in human cells. *Science* **343**, 88-91 (2014).
60. Lydeard, J. R., Jain, S., Yamaguchi, M., Haber, J.E. Break-induced replication and telomerase-independent telomere maintenance require Pol32. *Nature* **448**, 820-823 (2007).
61. Pursell, Z. F. & Kunkel, T. A. DNA polymerase epsilon: a polymerase of unusual size (and complexity). *Prog. Nucleic Acid Res. Mol. Biol.* **82**, 101-145 (2008).
62. Ganai, R. a., Bylund, G. O. & Johansson, E. Switching between polymerase and exonuclease sites in DNA polymerase ε. *Nucleic Acids Res*. **4**, 1-11 (2014).
63. Morrison, A., Johnson, A.L., Johnston, L.H., Sugino, A. Pathway correcting DNA replication errors in Saccharomyces cerevisiae. *EMBO J.* **12**, 1467-1473 (1993).
64. Simon, M., Giot, L. & Faye, G. The 3' to 5' exonuclease activity located in the DNA polymerase delta subunit of Saccharomyces cerevisiae is required for accurate replication. *EMBO J.* **10**, 2165-2170 (1991).
65. Murphy, K., Darmawan, H., Schultz, A., da Silva, E. F. & Reha-Krantz, L. J. A method to select for mutator DNA polymerase Deltas in Saccharomyces cerevisiae. *Genome* **49**, 403-410 (2006).
66. Datta, A. et al. Checkpoint-dependent activation of mutagenic repair in Saccharomyces cerevisiae pol3-01 mutants. *Mol. Cell* **6**, 593-603 (2000).
67. Williams, L. N. et al. dNTP pool levels modulate mutator phenotypes of error-prone DNA polymerase ε variants. *Proc. Natl. Acad. Sci. USA* **112**, E2457-E2466 (Advance online publication March 2015).
68. Bellido, F. et al. POLE and POLD1 mutations in 529 kindred with familial colorectal cancer and/or polyposis: review of reported cases and recommendations for genetic testing and surveillance. *Genet. Med.* 1-8 (Advance online publication July 2015).

69. Spier, I. et al. Frequency and phenotypic spectrum of germline mutations in POLE and seven other polymerase genes in 266 patients with colorectal adenomas and carcinomas. *Int. J. Can.* **137**, 320-31 (2015).
70. Chubb, D. et al. Genetic diagnosis of high-penetrance susceptibility for colorectal cancer (CRC) is achievable for a high proportion of familial CRC by exome sequencing. *J. Clin. Oncol.* **33**, 426-432 (2015).
71. Valle, L. et al. New insights into POLE and POLD1 germline mutations in familial colorectal cancer and polyposis. *Hum. Mol. Gen.* **23**, 3506-3512 (2014).
72. Elsayed, F. a. et al. Germline variants in POLE are associated with early onset mismatch repair deficient colorectal cancer. *Eur. J. Hum. Gen.* **23**, 1080-1084 (2015).
73. Rohlin, A. et al. A mutation in POLE predisposing to a multi-tumour phenotype. *Int. J. Oncol.* **45**, 77-81 (2014).
74. Hansen, M. F. et al. A novel POLE mutation associated with cancers of colon, pancreas, ovaries and small intestine. *Fam. Cancer* **14**, 437-48 (2015).
75. Aoude, L. G. et al. POLE mutations in families predisposed to cutaneous melanoma. *Fam. Cancer* (Advance online publication August 2015).
76. Zou, Y. et al. Frequent POLE1 p.S297F mutation in Chinese patients with ovarian endometrioid carcinoma. *Mut. Res.* **761**, 49-52 (2014).
77. Jones, S. et al. Genomic analyses of gynaecologic carcinosarcomas reveal frequent mutations in chromatin remodelling genes. *Nat. Commun.* **5**, 5006 (2014).
78. Shlien, A. et al. Combined hereditary and somatic mutations of replication error repair genes result in rapid onset of ultra-hypermutated cancers. *Nat. Gen.* **47**, 257-62 (2015).
79. The Cancer Genome Atlas Network. Widespread genetic heterogeneity in multiple myeloma: Implications for targeted therapy. *Cancer Cell* **25**, 91-101 (2014).
80. The Cancer Genome Atlas Network. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* **513**, 202-209 (2014).
81. Kane, D. P. & Shcherbakova, P. V. A common cancer-associated DNA polymerase ε mutation causes an exceptionally strong mutator phenotype, indicating fidelity defects distinct from loss of proofreading. *Cancer Res.* **74**, 1895-901 (2014).

82. Hussein, Y. R. et al. Clinicopathological analysis of endometrial carcinomas harboring somatic POLE exonuclease domain mutations. *Mod. Pathol.* **28**, 505-14 (2014).
83. Bertagnolli, M. M. et al. Microsatellite instability and loss of heterozygosity at chromosomal location 18q: prospective evaluation of biomarkers for stages II and III colon cancer--a study of CALGB 9581 and 89803. *J. Clin. Oncol.* **29**, 3153-3162 (2011).
84. Hutchins, G. et al. Value of mismatch repair, KRAS, and BRAF mutations in predicting recurrence and benefits from chemotherapy in colorectal cancer. *J. Clin. Oncol.* **29**, 1261-1270 (2011).
85. Nelson, G. S. et al. MMR deficiency is common in high-grade endometrioid carcinomas and is associated with an unfavorable outcome. *Gynecol. Oncol.* **131**, 309-314 (2013).
86. Stelloo, E. et al. Refining prognosis and identifying targetable pathways for high-risk endometrial cancer; a TransPORTEC initiative. *Mod. Pathol.* **28**, 836-844 (2015).
87. Diaz-Padilla, I. et al. Mismatch repair status and clinical outcome in endometrial cancer: A systematic review and meta-analysis. *Crit. Rev. Oncol. Hematol.* **88**, 154-167 (2013).
88. Heitzer, E. & Tomlinson, I. Replicative DNA polymerase mutations in cancer. *Curr. Opin. Genet. Dev.* **24**, 107-113 (2014).
89. Williams, L. N., Herr, A. J. & Preston, B. D. Emergence of DNA polymerase ε antimutators that escape error-induced extinction in yeast. *Genetics* **193**, 751-770 (2013).
90. Herr, A. J., Kennedy, S. R., Knowels, G. M., Schultz, E. M. & Preston, B. D. DNA replication error-induced extinction of diploid yeast. *Genetics* **196**, 677-691 (2014).
91. Heyer, J., Yang, K., Lipkin, M., Edelmann, W. & Kucherlapati, R. Mouse models for colorectal cancer. *Oncogene* **18**, 5325-5333 (1999).
92. Sole, R. V. & Deisboeck, T. S. An error catastrophe in cancer? *J. Theoret. Biol.* **228**, 47-54 (2004).
93. Loeb, L. A. Human cancers express mutator phenotypes: origin, consequences and targeting. *Nat. Rev. Cancer* **11**, 450-457 (2011).
94. Herr, A. J. et al. Mutator Suppression and Escape from Replication Error-Induced Extinction in Yeast. *Plos Genet.* **7**, e1002282 (2011).

95. Ghodgaonkar, M. M. et al. Phenotypic characterization of missense polymerase-δ mutations using an inducible protein-replacement system. *Nat. Commun*. **5**, 4990-4990 (2014).
96. Rizvi, N. A. et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* **348**, 124-128 (2015).
97. Korona, D. A., LeCompte, K. G. & Pursell, Z. F. The high fidelity and unique error signature of human DNA polymerase. *Nucleic Acids Res.* **39**, 1763-1773 (2011).
98. Fox, E. J., Prindle, M. J. & Loeb, L. A. Do mutator mutations fuel tumorigenesis? *Cancer Met. Rev*. **32**, 353-361 (2013).
99. Llosa, N. J. et al. The vigorous immune microenvironment of microsatellite instable colon cancer is balanced by multiple counter-inhibitory checkpoints. *Cancer Discov.* **5**, 43-51 (2015).
100. McFarland, C. D., Mirny, L. A. & Korolev, K. S. Tug-of-war between driver and passenger mutations in cancer and other adaptive processes. *Proc. Natl. Acad. Sci. USA* **111**, 15138-15143 (2014).
101. Vasen, H. F. et al. Revised guidelines for the clinical management of Lynch syndrome (HNPCC): recommendations by a group of European experts. *Gut* **62**, 812-823 (2013).
102. Hamid, O. et al. Safety and tumor responses with lambrolizumab (anti-PD-1) in melanoma. *N. Engl. J. Med*. **369**, 134-144 (2013).
103. Brahmer, J. R. et al. Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. *N. Engl. J. Med*. **366**, 2455-2465 (2012).
104. Topalian, S. L. et al. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N. Engl. J. Med*. **366**, 2443-2454 (2012).
105. Le, D. T. et al. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N. Engl. J. Med*. **372**, 2509-2520 (2015).
106. Snyder, A. et al. Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N. Engl. J. Med*. **371**, 2189-2199 (2014).
107. Tumeh, P. C. et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* **515**, 568-571 (2014).
108. Twyman-Saint Victor, C. et al. Radiation and dual checkpoint blockade activate non-redundant immune mechanisms in cancer. *Nature* **520**, 373-377 (2015).
109. Jin, Y. H. et al. Cadmium is a mutagen that acts by inhibiting mismatch repair. *Nat. Genet*. **34**, 326-329 (2003).

110. Mertz, T. M., Sharma, S., Chabes, A. & Shcherbakova, P. V. Colon cancer-associated mutator DNA polymerase δ variant causes expansion of dNTP pools increasing its own infidelity. *Proc. Natl. Acad. Sci. USA* **112**, E2467-76 (Online publication March 2015).
111. Hyman, D. M. et al. Vemurafenib in Multiple Nonmelanoma Cancers with BRAF V600 Mutations. *N. Engl. J. Med.* **373**, 726-736 (2015).
112. Loeb, L. A., Bielas, J. H. & Beckman, R. A. Cancers exhibit a mutator phenotype: clinical implications. *Cancer Res*. **68**, 3551-3557 (2008).
113. Abdus Sattar, A. K., Lin, T. C., Jones, C. & Konigsberg, W. H. Functional consequences and exonuclease kinetic parameters of point mutations in bacteriophage T4 DNA polymerase. *Biochemistry* **35**, 16621-16629 (1996).
114. Hoang, L. N. M. D. et al. Polymerase Epsilon Exonuclease Domain Mutations in Ovarian Endometrioid Carcinoma. *Int. J. Gynecol. Cancer* **25**, 1187-93 (2015).
115. Cerami, E. et al. The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discov*. **2**, 401-404 (2014).
116. Stenzinger, A. et al. Mutations in POLE and survival of colorectal cancer patients - link to disease stage and treatment. *Cancer Med*. **3**:1527-38 (2014).
117. Swan, M. K., Johnson, R. E., Prakash, L., Prakash, S. & Aggarwal, A. K. Structural basis of high-fidelity DNA synthesis by yeast DNA polymerase δ. *Nature Struct. Mol. Biol*. **16**, 979-986 (2009).
118. Wang, J., Yu, P., Lin, T. C., Konigsberg, W. H. & Steitz, T. A. Crystal Structures of an NH2-Terminal Fragment of T4 DNA Polymerase and Its Complexes with Single-Stranded DNA and with Divalent Metal Ions †. *Biochemistry* **35**, 8110-8119 (1996).

Figure 1. Frequency and location of germline and somatic *POLD1* and *POLE* proofreading exonuclease domain mutations in cancers

Representation of exonuclease domains of *POLD1* (*A*) and *POLE* (*B*), showing conserved exo motifs (I-V), position and frequency of germline and somatic mutations. Exo I active site residues are highlighted. Position of germline (blue) and somatic (red) proofreading domain mutations mapped to *S. cerevisiae* Pol δ (PDB 3IAY[117]) (*C*) and Pol ε (PBD 4M8O[50]) (*D*) exonuclease domain structures. ssDNA from the aligned T4 polymerase complex (PDB 1NOY[118]) is shown in yellow. The exo I motif active site residues are highlighted in magenta (with exception of mutated *POLD1* active site residue D316). Note that Pol ε residues 424 and 436 are the site of both germline and somatic mutations.

Table 1. Pathogenic somatic *POLD1* and *POLE* proofreading exonuclease domain mutations in sporadic cancers

Gene / nucleotide substitution	Protein change	Evidence supporting pathogenicity				Molecular characteristics†			Tumour type
		Structural mapping	PhyloP score*	Yeast fluctuation assay	Biochemical evidence for functional effect‡	Median no. mutations / exome (range)	% MSS	% tumours with ≥ 20% C>A substitutions	
POLD1									
c.956G>A	p.C319Y	Within exo I motif	2.38	Yes[65]	NR	7346 (44-14654)	100	N/A	GBM[78], MM[79]
POLE									
c.857C>G/A/T	p.P286R/H/L	Flanking Exo I motif	2.57	Yes[81]	Yes[18]	5147 (738-16248)	100	93	CRC[13,16,81], EC[15,17,20,21,24], GBM[14], EOC[114] BrC[115]
c.890C>T/A	p.S297F/Y	Flanking Exo I motif	2.67	NR	NR	5419 (4918-15545)	50	66	EC[15,17,20,21,24], GBM[14,78], EOC[76], CC[115]
c.1100T>C	p.F367S	Exo II active site	2.19	NR	Yes[18,113]	2934	100	100	CRC[11,13]

c.1231G>C/T	p.V411L	Flanking Exo IV Motif	2.66	NR	Yes[18]	6294 (955-14074)	100	88	CRC[13], EC[15,17,20,21,24], GBM[14], EOC[114], GC[80]
c.1270C>G/A	p.L424V/I	Exo IV active site	2.66	Yes[65]	Yes[18,113]	163 (85-6724)	50	100	EC[17], BrC[115]
c.1307C>G	p.P436R	Exo V Motif	3.53*	NR	Yes[113]	6131	100	100	CRC[13], EC[21]
c.1331T>A	p.M444K	Flanking exo V motif	2.15	NR	NR	1204	100	100	EC[17]
c.1366G>C	p.A456P	Within Exo III motif	2.61	NR	NR	5968	100	100	CRC[18,116], EC[15,17,21]
c.1376C>T	p.S459F	Within Exo III motif	2.52	NR	Yes[18,113]	4780 (1868-9907)	100	75	CRC[13,18,116] GBM[14] AA[78]

MSS – microsatellite stable

AA – anaplastic astrocytoma; BrC, Breast cancer; CC – Squamous cell cervical carcinoma; CRC – Colorectal carcinoma; EC – Endometrial carcinoma; EOC – Endometrioid ovarian carcinoma; GC, Gastric cancer; GBM – Glioblastoma multiforme; MM – Multiple myeloma; PC – Pancreatic adenocarcinoma.

* PhyloP (phylogenetic conservation) scores were calculated per nucleotide using alignment of 46 vertebrates dbNSFPv23. Where a variant mapped to the third position of a codon the average PhyloP score for the codon is displayed (denoted by *).

† Data from exome sequencing studies.

‡ Data from studies of B family polymerases

NR – not reported

Table 2. Pathogenic germline *POLD1* and *POLE* proofreading exonuclease domain mutations

		Evidence supporting pathogenicity of variant								% of carriers with									
Gene / nucleotide change	Protein change	Structural mapping	PhyloP score*	Yeast fluctuation assays	Biochemical evidence for functional effect	Segregates with affection status	Number of carriers†	Number of unrelated carriers†	Mean age at diagnosis (range) §	CRC	EC	BrC	DuC	OC	GBM	Duodenal :adenomas / polyps	Colonic adenomas / polyps	Other cancers reported in carriers	References
POLD1																			
c.947A>G	p.D316G	Exo I motif active site	1.93	Yes[64]	Yes[113]	Yes	2	1	51 (44-57)	50	100	50	0	0	0	0	50	0	[68]
c.946G>C	p.D316H	Exo I motif active site	1.16	Yes[64]	Yes[113]	Yes	2	1	61 (58-64)	50	0	50	0	0	0	0	100	Mesothelioma	[68]
c.981C>G	p.P327L	Flanking Exo I motif	2.16	Yes[81 ǁ]	Yes[18 ǁ]	NR	1	1	70	0	0	0	0	0	0	0	100	0	[12]
c.1225C>T	p.R409W	Flanking Exo II motif	2.26	NR	NR	NR	1	1	32	100	0	0	0	0	0	0	100	0	[68]
c.1421T>C	p.L474P	Exo IV motif	1.92	Yes[65]	Yes[18 ǁ]	Yes	6	2	40 (23-52)	67	33	0	0	0	0	0	17	0	[68,71]
c.1433G>A	p.S478N	Exo IV motif	1.19	Yes[12]	NR	Yes	11	3	35 (26-52)	45	36	0	0	0	0	0	91	Astrocytoma	[12]

POLE																			
c.1089C>A	p.N363K	Exo II motif active site	4.9	NR	NR	Yes	12	1	41 (28-56)	75	17	0	0	25	0	NR	83	Pancreas	[73]
C.1103A>T	D368V	Exo II motif active site	5.13	NR	Yes[113]	NR	1	1	47	100	0	0	0	0	0	0	0	0	[70]
c.1270C>G	p.L424V	Exo IV motif active site	2.66	Yes[65]	Yes[18]	Yes. 2 de novo carriers	48	21	39 (16-64)	61	2	2	2	2	4	19	92	Oligo-dendroglioma, neuro-endocrine carcinoma	[12,69,70,71,72]
c.1306C>T	p.P436S	Within Exo V motif	3.53	NR	NR	de novo	1	1	31	100	0	0	0	0	0	100	100	0	[69]
c.1373A>T	p.Y458F	Exo III motif active site	4.97	NR	Yes[113]	Yes	13	2	48 (38-63)	62	0	0	15	8	0	NR	62	Pancreas	[74]
c.1041G>T	p.W347C	Outside Exo motifs	2.71	Yes[75]	NR	No Δ¶	11	1	49 (14-70)	0	0	0	0	0	0	NR	NR	Prostate, cutaneous and uveal melanoma	[75]

CRC, Colorectal cancer; EC, Endometrial cancer; GBM, Glioblastoma; OC, Ovarian cancer ; BrC, Breast cancer, DuC, duodenal carcinoma

* PhyloP (phylogenetic conservation) scores were calculated per nucleotide using alignment of 46 vertebrates dbNSFPv23. Where a variant mapped to the third position of a codon the average phyloP score for the codon is displayed (denoted by *)

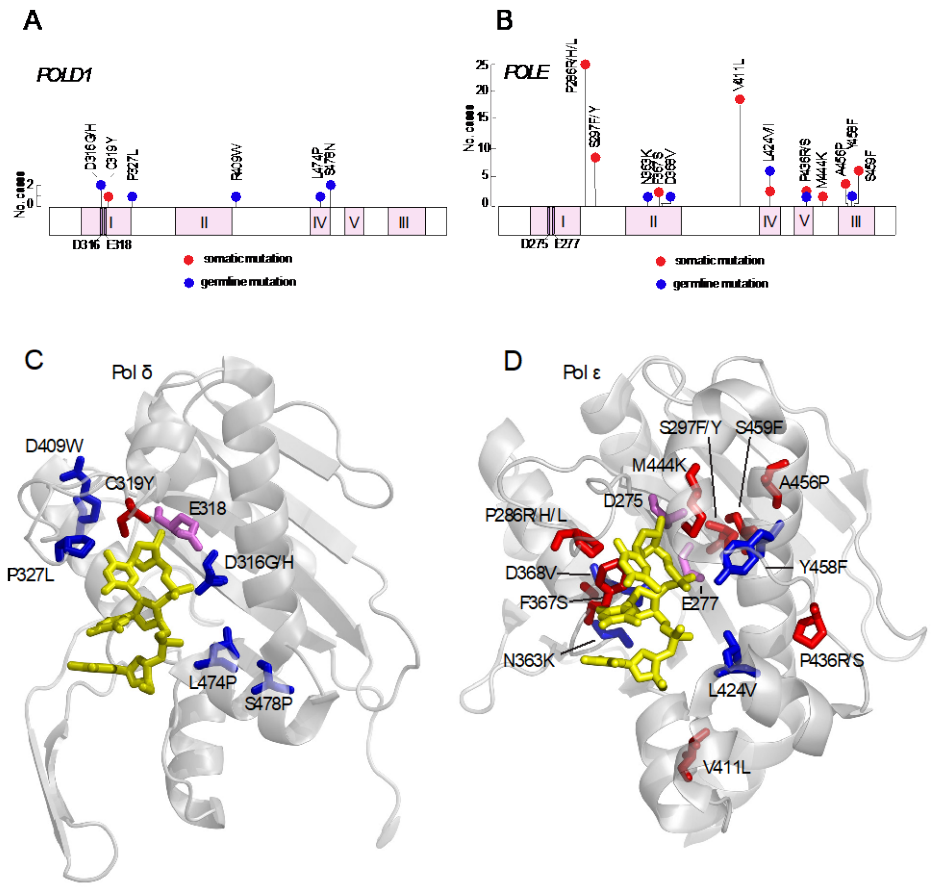
† reported as of Aug 2015

‡ Data from studies of B family polymerases

§ Age at diagnosis refers to cancer or adenoma diagnosis, whichever was earliest

ǁ Functional studies of corresponding residue in Pol ε

¶ Δ 6 mutation carriers were unaffected.


A
POLD1
D316G/H
C319Y
P327L
R409W
L474P
S478N
I
II
IV
V
III
D316 E318
somatic mutation
germline mutation
B
POLE
P286R/H/L
S297F/Y
N363K
F367S
D368V
V411L
L424V/I
P436R/S
M444K
A456P
S459F
D275 E277
somatic mutation
germline mutation
C
Pol δ
D409W
C319Y
E318
P327L
D316G/H
L474P
S478P
D
Pol ε
S297F/Y
S459F
M444K
A456P
D275
P286R/H/L
D368V
Y458F
F367S
E277
N363K
P436R/S
L424V
V411L