Mutant POLQ and POLZ/REV3L DNA polymerases may contribute to the favorable survival of patients with tumors with POLE mutations outside the exonuclease domain

Fangjin Huang, Hisashi Tanaka, Beatrice S. Knudsen* and Joanne K. Rutgers

Abstract

Background: Mutations in the exonuclease domain of POLE, a DNA polymerase associated with DNA replication and repair, lead to cancers with ultra-high mutation rates. Most studies focus on intestinal and uterine cancers with POLE mutations. These cancers exhibit a significant immune cell infiltrate and favorable prognosis. We questioned whether loss of function of other DNA polymerases can cooperate to POLE to generate the ultramutator phenotype.

Methods: We used cases and data from 15 cancer types in The Cancer Genome Atlas to investigate mutation frequencies of 14 different DNA polymerases. We tested whether tumor mutation burden, patient outcome (disease-free survival) and immune cell infiltration measured by ESTIMATE can be attributed to mutations in POLQ and POLZ/REV3L.

Results: Thirty six percent of colorectal, stomach and endometrial cancers with POLE mutations carried additional mutations in POLQ (E/Q), POLZ/REV3L (E/Z) or both DNA polymerases (E/Z/Q). The mutation burden in these tumors was significantly greater compared to POLE-only (E) mutant tumors (p < 0.001). In addition, E/Q, E/Z, and E/Q/Z mutant tumors possessed an increased frequency of mutations in the POLE exonuclease domain (p = 0.013). Colorectal, stomach and endometrial E/Q, E/Z, and E/Q/Z mutant tumors within TCGA demonstrated 100% disease-free survival, even if the POLE mutations occurred outside the exonuclease domain (p = 0.003). However, immune scores in these tumors were related to microsatellite instability (MSI) and not POLE mutation status. This suggests that the host immune response may not be the sole mechanism for prolonged disease-free survival of ultramutated tumors in this cohort.

Conclusion: Results in this study demonstrate that mutations in POLQ and REV3L in POLE mutant tumors should undergo further investigation to determine whether POLQ and REV3L mutations contribute to the ultramutator phenotype and favorable outcome of patients with POLE mutant tumors.
Background
The analysis of thousands of cancers by The Cancer Genome Atlas (TCGA) consortium and academic institutions revealed a small group of cancers with mutations in POLE and an ultramutator phenotype [1, 2]. Multiple studies have found significantly improved survival in patients with POLE mutated endometrial cancers [1, 3–6], while the survival benefit was not as profound in POLE mutated colorectal carcinoma [7]. The POLE gene encodes the catalytic subunit of DNA polymerase epsilon, which catalyzes the leading strand synthesis during DNA replication. POLE possesses high-fidelity DNA polymerization, proofreading and 3′–5′ exonuclease activities, which promote accurate DNA synthesis [8]. First identified and reported in 2–6% of colorectal carcinomas [2, 9, 10], POLE mutations were also noted at frequencies of 6–9% amongst uterine corpus endometrial cancers [1, 11] and in gastric adenocarcinoma [12]. Mutations can be found across the entire POLE gene, but those in the POLE exonuclease domain are most prevalent in cancers with ultra-high mutation rates (>100 mut/Mb). These cancers exhibit higher mutation rates than microsatellite instable (MSI) tumors associated with mismatch repair abnormalities. In addition, POLE ultra-mutant cancers also possess a high frequency of cisplatin sensitivity [28–31]. Because studies describe overlapping functions and synergy among the over 14 DNA polymerases [27, 32, 33], we undertook an unbiased approach to identify mutations in polymerases within the TCGA tumor compendium. This analysis revealed a greater frequency of mutations in POLQ and POLZ/REV3L compared to other polymerases of similar gene length. Therefore, we propose that POLQ and POLZ/REV3L may cooperate with POLE in generating ultrahigh mutation rates.

Methods
Data acquisition
The data used in this study are based on the whole exome sequence data sets generated by the TCGA Research Network: http://cancergenome.nih.gov/. The locations and frequencies of somatic mutations, MSI status and clinical stage and follow up information in TCGA Provisional datasets were obtained from cBioportal (http://www.cbioportal.org) [34, 35] up to 06/22/2016 (Supplementary Table 1). Mutation data were obtained by first selecting the “Query” tab from cBioPortal. The cancer type specific data were chosen from the “TCGA PanCancer Atlas Studies”. We entered the genes of interest, POLE, POLQ and REV3L to retrieve mutation data from the genes. When the result page was displayed, we accessed the information underneath the “Mutations” tab. We downloaded the number of mutations in each sample, the annotation of protein amino acid change, and the type of mutation. The survival information for each case (months after diagnosis) was obtained through the “Comparison/Survival” link. Functional domains of the proteins were provided by Pfam database [36]. Data visualization for mutations was performed with MutationMapper in cBioportal.

Case selection criteria
We searched all cancer types in TCGA for those that possess 2 or more cases with POLE mutations. This yielded 15 tumor types that we named the PANCAN data set in this study. We then determined the frequency of mutations within additional polymerase within PANCAN. We selected the three adenocarcinomas (uterine corpus endometrial adenocarcinoma, colorectal adenocarcinoma, and stomach adenocarcinoma) with the highest frequency of double or triple mutated DNA polymerase status for more detailed analysis.
Studies performed
Kaplan Meier survival plots were generated using clinical follow-up data available within the TCGA database. To determine the global mutational spectrum, we classified 6 types of nucleotide transitions or transversions. The frequency of each mutation type was calculated.

Digital images of all E/Q mutant tumors and representative cases of tumors with neither POLE nor POLQ mutations, and of MSI tumors were assessed by one author (JR) for the amount of tumor immune infiltrate. The combination of tumor infiltrating lymphocytes and the peritumoral lymphoid infiltrate was graded on a scale between 0 and 3. Tumor associated lymphocytes were graded as none (0), minimal (rare) to 1 per high powered field (HPF) (1), 2 to 5 per HPF (2) and > 5 per HPF (3). Peritumoral lymphocytes were assessed at the deepest advancing tumor front, graded at low power: none (0), minimal (1), mild (2), moderate (3) and marked (4). These visual semi-quantitative scores were compared with the immune scores evaluated using ESTIMATE (Estimation of STromal and Immune cells in MA lignant Tumor tissues using Expression data) [37]. ESTIMATE score were obtained from the website of ESTIMATE at MD Anderson: https://bioinformatics.mdanderson.org/estimate/disease.html. The platform type selected was RNA-Seq-V2. Data for colorectal, stomach and endometrial cancer types were downloaded. Next, we separated the colorectal, stomach and endometrial cancer cases into 3 quartiles defined by the highest 25%, intermediate 50%, and lowest 25% of mutational counts or MSI status, and calculated the immune scores for each group.

Statistical data analysis
All statistical analyses were conducted in R v3.1.3. Plots were generated using ggplot2 package in R [38]. Data visualization methods were described previously [39]. The horizontal lines in the boxplots represent the 1st, 2nd and 3rd quartiles and whiskers outside the box show the 1.5 interquartile range. The significance of the differences of data illustrated in the boxplots was determined using the Wilcoxon rank-sum tests. The Chi-square test was performed to test the significance of differences in frequencies of all tables. The significance in the Kaplan–Meier survival plot was calculated using the log rank test. Statistical significance was accepted at p < 0.05.

Results
Of the 33 cancer types in the TCGA database [40], we identified 15 cancer types (PANCAN) with 2 or more cases that possessed mutations in the POLE protein coding region (Fig. 1A and Supplementary Table 1). These 15 cancer types contained 138 cases with POLE mutations anywhere in the exome. 53% of these POLE mutant tumors carried mutations in one or more of the 14 other DNA polymerases. We observed DNA polymerases POLQ and POLZ/REV3L to be most frequently mutated (Fig. 1b). These two polymerases were mutated in 36% of tumors with POLE mutations. In fact, these two polymerases were even more commonly mutated than POLE in our PANCAN cohort (Supplementary Figure 1). Altogether, 14 cases with POLE and POLQ mutations (E/Q), 16 cases with POLE and POLZ/REV3L mutations (E/Z) and 20 cases with POLE, POLQ and POLZ/REV3L (E/Q/Z) mutations were identified in the PANCAN cohort (Fig. 1c). Mutations in the exonuclease domain of POLE are responsible for causing the ultramutator phenotype in colorectal and uterine corpus cancers [1, 2, 11, 41]. In order to determine the contribution of POLQ and REV3L to the ultramutator phenotype, we compared the mutation frequencies of tumors with mutations in only POLE to E/Q, E/Z and E/Q/Z mutant tumors. Mutation frequencies in the cellular genome increased in the following order: no POLE mutations < POLE-only mutations (anywhere with the POLE exome) < E/Q, E/Z, E/Q/Z mutations (Fig. 1d). The median mutation count of E/Q/Z tumors was more than 10-fold higher (p < 0.001) than that of tumors with only POLE mutations. E + Q and E + Z mutant tumors also displayed significantly higher mutation counts compared to E-only mutant tumors, suggesting a contribution of mutationally altered POLQ or POLZ/REV3L to the overall cancer mutation rates. Next, we determined the number of mutations in the exonuclease and polymerase domains of POLE in the 15 cancer types within our PANCAN compendium (Fig. 1e). The percentage of POLE mutations in the exonuclease domain was greater in E/Q, E/Z and E/Q/Z mutant tumors compared to POLE-only mutant tumors (p = 0.013). In contrast, the percentage of mutations in the DNA polymerase domain was similar. Thus, our data confirm the notion that mutations in the exonuclease domain of POLE are responsible for ultra-high mutation rates. In addition, mutations in POLQ and REV3L may further increase the mutation burden. However, why mutations in POLQ and REV3L preferentially increase tumor mutation frequencies remains elusive.

To further investigate a potential role of these mutant DNA polymerases in the ultramutator phenotype, we focused on colorectal [38], endometrial (UCEC) and stomach (STAD) cancers. These cancer types contain the highest numbers of tumors with E/Q, E/Z and E/Q/Z amongst the 15 cancer types included in PANCAN (Fig. 1a, Supplementary Table 1). Among these 3 cancer types, we identified 6 cases with E/Q, 12 cases with E/Z and 16 cases with E/Q/Z mutations (Fig. 2a). In these cancers, the mutation burden in POLE, E/Q, E/Z and E/
Fig. 1 Cancer types (PANCAN) with POLE/E/Q/Z mutations in TCGA. 

a. Number of cases with POLE- only, E/Q, E/Z and E/Q/Z mutations in 15 cancer types (cohort referred to as PANCAN) within TCGA. The x-axis shows the actual number of cases with POLE (green), E/Q (orange), E/Z (pink) and E/Q/Z (blue) mutations. The y-axis displays the 15 cancer types: Uterine corpus endometrial carcinoma (UCEC), Stomach adenocarcinoma (STAD), Colon and rectum adenocarcinoma [38], Skin cutaneous melanoma (SKCM), Lymphoid Neoplasm Diffuse Large B-cell Lymphoma (DLBC), Lung adenocarcinoma (LUAD), Breast invasive carcinoma (BRCA), Sarcoma (SARC), Cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC), Pancreatic adenocarcinoma (PAAD), Lung squamous cell carcinoma (LUSC), Head and neck squamous cell carcinoma (HNSC), Bladder urothelial carcinoma (BLCA), Kidney renal clear cell carcinoma (KIRC), and Liver hepatocellular carcinoma (LIHC).

b. Case numbers with mutations in polymerase genes. The number of cases in PANCAN with mutations in the following polymerases is displayed on the Y-axis: DNTT, POLA1, POLB, POLD1, POLG, POLH, POLK, POLL, POLM, POLN, POLQ, REV1, REV3L.

c. Venn diagram displaying the number of cases in PANCAN with mutations in 1, 2 or 3 POL genes.

D. Mutations per Mb (y-axis) of PANCAN cases without POLE mutations (other) or with POLE, E/Q, E/Z and E/Q/Z (x-axis) mutations. Number of cases in each group are listed in parenthesis.

e. Mutation frequencies in POLE exonuclease and polymerase domains as a percentage of total number of mutations in the POLE exome. “Other” refers mutations in the entire Exome outside the exonuclease or polymerase domains. The cases are grouped by their polymerase mutation status on the y-axis, and the number in parenthesis represents the total number of POLE mutations within each group.
Q/Z mutant tumors paralleled the mutation burden in the whole PANCAN cohort (compare Fig. 2b and Fig. 1c, Supplementary Figure 2A-D). E/Q/Z mutant tumors demonstrated, on average, an 8-fold increase in mutation frequencies compared to tumors with only POLE mutations. Amongst POLE mutant tumors, 26 tumors carried mutations outside the POLE exonuclease domain, while 31 tumors carried mutations within the exonuclease domain. The frequency of POLE exonuclease mutations (15/16 cases in Fig. 2c) provides a valid explanation for the difference in mutation rates and potential association of POLE exonuclease domain mutations with mutations in POLQ and POLZ/REV3L, which are the other two most frequently mutated DNA polymerases.

Overall, exonuclease domain mutations were identified in 6/23 cases of POLE only mutant tumors, 4/6 E/Q cases, 6/12 E/Z cases and 15/16 of E/Q/Z cases (Fig. 2c). Stratified by cancer types, POLE exonuclease domain mutations occurred in 7/15 colorectal, 18/26 endometrial and 6/16 stomach tumors, demonstrating cancer type specific frequencies (Supplementary Figure 3A). In contrast to the POLE gene that demonstrates mutational hotspots in the exonuclease domain, mutational hotspots in the POLQ gene are not associated with a functional protein domain (Supplementary Figure 3B, C & D). While REV3L does not reveal mutational hotspots, approximately 50% of mutations lead to truncated protein expression (Supplementary Figure 3E & F). Another characteristic of POLE mutant tumors are C to A and G
to T transversions [2]. We observed the greatest increase of nucleotide transversions in cancers with E/Q/Z mutations (Fig. 2d), consistent with the loss of POLE exonuclease activity in these tumors.

Since mutations in POLE confer increased disease free survival (DFS) in patients with uterine cancer, even in those patients with high-grade tumors [3, 42], we investigated the prognostic role of POLQ and REV3L mutations in POLE mutant tumors. Kaplan-Meier curves were constructed for colorectal, endometrial and stomach cancer cases with follow-up data (Fig. 3a). Using the TCGA annotations of DFS in individual patients, no cancer recurrences were observed in the E/Q, E/Z and E/Q/Z mutant groups. POLE exonuclease domain mutations were observed in 29 cases in the good survival group and 1 case in the poor survival group, consistent with the expected long DFS periods of patients with POLE exonuclease domain positive tumors. In addition, 19 cases with mutations in POLE outside the exonuclease domain were in the good survival group. Of those 7 (37%) had concurrent mutations in POLQ or REV3L or in both polymerases (Fig. 3b). Furthermore, a Kaplan-Meier analysis in the PANCAN cohort revealed improved DFS associated with mutations in POLQ and REV3L. The favorable survival outcome was observed in colorectal, endometrial, and stomach cancers. However, no favorable outcome was observed in diffuse B-cell lymphoma (p = 0.35). These data provide preliminary evidence of cancer-type specific, favorable survival outcomes in tumors with POLE mutations that are located outside the POLE exonuclease domain if concurrent mutations in POLQ, REV3L or in both polymerases are present.

Compared to microsatellite stable tumors (MSS), microsatellite instability (MSI) in colorectal cancer confers a better prognosis [43]. To determine whether

![Fig. 3](image-url)
the favorable outcome of E/Q, E/Z and E/Q/Z mutant cancers can be explained by MSI or TMN stage, we examined the relationship between MSI status, tumor stage and polymerase mutations in colorectal, endometrial and stomach cancers (Fig. 3c and Supplementary Figure 6 and Supplementary Table 2). Despite improved DFS rates, the full range of tumor stages was observed amongst E/Q, E/Z and E/Q/Z tumors (p = 0.42) (Supplementary Table 2A). Comparing the POLE-only and E/Q/Z mutant cancers did not reveal a significant difference in tumor stage, but differed in the frequency of MSI cases (p < 0.001). POLE mutant tumors were more frequent in the MSI group (29/163, 17.8%) than in the MSS group (30/595, 5%). In addition, the frequency of MSI cases in POLE mutant tumors differed between the 3 cancer types (p < 0.001) (Supplementary Table 2B & C). Although MSI is enriched in samples with high mutation levels (Supplementary Table 2D), as expected, 10-fold higher mutation counts (P < 0.001) were observed in cancers with E/Q/Z mutations compared to MSI without E/Q/Z mutations (Supplementary Figure 5). These results suggest that mutations in E/Q, E/Z and E/Q/Z confer a better prognosis independent of MSI status and TMN stage in colorectal, endometrial and stomach adenocarcinomas.

We next examined the amount of the cancer-associated immune infiltrate. The immune score obtained through ESTIMATE [37] corresponded to the categorical score of the immune infiltrate derived from digital H&E images (Supplementary Figure 6). Therefore, we used the ESTIMATE immune scores for further analysis of colorectal, endometrial and stomach cancers. As shown in Fig. 4a, a significant difference was observed in the median immune scores between groups with low, intermediate and high mutation burden, grouped based on mutation burden and not on E, Q, Z mutant status (see Methods) and, as expected, the median immune scores increased with total mutation levels. Surprisingly, the immune scores in E/Q/Z mutant tumors did not differ significantly from tumors with a low level of mutations (Fig. 4b). As expected, MSI tumors possessed higher immune scores than MSS tumors (p < 0.001) [44] (Fig. 4c). Finally, immune scores of MSI and E/Q/Z mutation tumors were similar to those in the MSI group and higher than MSS and E/Q/Z mutation tumors (Fig. 4d). Within the group of tumors with POLE exonuclease domain mutations, MSS tumors possessed lower immune scores than MSI tumors, but the difference was not significant (p = 0.29) (Supplementary Figure 7). Together, results in this TCGA cohort demonstrate that the immune response is driven by MSI rather than POLE exonuclease domain mutations.

Discussion
An analysis of 14 DNA polymerases in tumors with mutations in the POLE revealed additional mutations in specific polymerases, most commonly in POLQ and POLZ/REV3L. Among the 15 cancer types, colorectal, uterine and stomach cancers were most frequently afflicted by these mutations. Cancers with mutations in POLE and POLQ (E/Q), POLE and POLZ/REV3L (E/Z) and in all 3 polymerases (E/Q/Z) were associated with the highest mutation burden and an excellent prognosis independent of MSI status and tumor stage. Mutations in the exonuclease domain were observed in 94% (15/16) of E/Q/Z mutant tumors, but only in 26% of POLE-only mutant tumors or in 55% of E/Q + E/Z tumors. However, despite harboring 10-fold more mutations than MSI tumors and 8-fold more mutations than the mutation frequencies associated with POLE-only mutant tumors, E/Q/Z mutant tumors did not display significantly more inflammation.

The main result from that analysis is that patients with colorectal, stomach and endometrial cancers bearing E/Q, E/Z and E/Q/Z mutations have 100% disease free survival (DFS) at a median follow up time of 33 months. In contrast, patients with tumors bearing mutations in POLE only, most of which outside the POLE exonuclease domain, had a DFS of 76% at follow up of 18.4 months. The favorable DFS in E/Q, E/Z and E/Q/Z mutant tumors occurred even in tumors with mutations in POLE that are located outside the POLE exonuclease domain. The contribution of mutations in POLE to TMB, C>A substitutions and cancer type associations are described in Table 1 of Raynor et al, 2016 using a larger resource of cases and should be used to interpret the mutations in the current study, listed in Supplementary Figure 3.

As a whole, the current study expands the spectrum of POLE mutant tumors with an excellent prognosis. The favorable prognosis included patients with high tumor stage, which echoes prior studies demonstrating a favorable outcome of uterine tumors with POLE exonuclease mutations despite adverse standard clinicopathologic indicators including high grade, high stage, and lymphovascular invasion [3, 45]. While the high mutation frequencies may cause an early growth advantage [46], as tumors evolve they may succumb to high mutation burden as new mutations can no longer be tolerated and cause tumor cell death [47, 48] or increased sensitivity to therapeutic agents.

The prevailing hypothesis for the favorable prognosis of cancers displaying the hypermutator phenotype is the increased attack by the immune system. Evidence in support of this theory is the observation that tumor infiltrating [49] and peritumoral lymphocytes are increased and that cytotoxic activities in CD8+ and CD4+ lymphocyte populations are heightened in POLE...
mutated endometrial cancers [50–53], similar to hypermutated MSI tumors [54]. This observation has led to the hypothesis that immune checkpoint inhibitors may be efficacious in POLE ultramutated tumors [51]. Our results question a direct relationship between mutation burden, tumor immune response and PD-L1 expression, also raised in a larger study across 5722 cases from 21 cancer types in TCGA [55]. While we observed a concordance between the computational and histological assessments of the immune infiltrate, the immune score in tumors with E/Q/Z mutations depended on MSI status. This result suggests that the immune infiltrate attributable to mutations in E/Q/Z mutant tumors may be less, or that its composition may involve immune cells other than lymphocytes. Lesser CD8+ and gamma-interferon gene expression signatures have also been observed in gastrointestinal tumors with a large single nucleotide variant (SNV) burden that was attributed largely to POLE exonuclease mutations [49]. Perhaps E/Q/Z mutations occur at a later point in tumor evolution [56] when immunosuppressive factors already dominate. We also cannot rule out the possibility of increased numbers of cytotoxic lymphocytes intermixed with E/Q/Z mutant tumor cells, because computational methods and inspection of H&E images are not sensitive enough to detect small differences in tumor infiltrating lymphocytes (TILs) that may have large anti-tumoral effects.

A significant limitation of the study lies in the relatively small number of tumors. This limitation cautions the generalization of results and seemingly novel insights.

**Fig. 4** ESTIMATE immune scores by mutation frequency quartiles, E/Q/Z mutation groups, and MSI in colorectal [38], endometrial (UCEC) and stomach (STAD) cancers. a ESTIMATE immune scores in cancers within high, intermediate and low overall mutation quartiles. b Immune scores of samples with E/Q, E/Z and E/Q/Z mutations compared to the low mutation quartile from panel A. c Immune scores in groups of cancers separated by MSI status. d Immune scores in MSI and MSS E/Q/Z cases compared to all other MSI cases. For each panel the number of cases within each group is included in parentheses on the x-axis.
into the hypermutator phenotype. Studies by other groups attributed hypermutator phenotype to specific mutations primarily within the POLE exonuclease domain. Given the proofreading function of the exonuclease domain, it makes sense that mutations outside the domain have a lesser effect on TMB. In agreement with this concept, our study reveals that (1) compared to tumors harboring only a mutation in POLE (POLE single-mutant tumors, 6 of 23 cases), POLE exonuclease domain mutations are more common in tumors with both double (E/Z and E/Q) and triple (E/Q/Z) DNA polymerase mutations (25 of 34 cases) and (2) double and triple mutant tumors have higher mutation counts than POLE single mutant tumors. While it cannot be fully excluded that POLQ and POLZ/REV3L mutations are bystander events in POLE mutant tumors, we observe a higher TMB in cases with mutations in all three polymerases (Supplementary Figure 2A and B). Mechanistically, POLQ and POLZ are thought to function in different repair processes: POLQ in alternative (microhomology-mediated) non-homologous DNA repair pathway and POLZ in translesion DNA synthesis. How these DNA repair processes cooperate with the replicative DNA polymerase, POLE, to prevent genome instability remains unknown. This will be an important subject for further understanding of the mechanism underlying the hypermutator phenotype.

Conclusions
If validated in additional cohorts, our findings may have important clinical implications. They build upon and expand the previously well documented good prognostic impact of POLE exonuclease mutations in uterine cancer, that have generated intense interest in part due to the paradox of a favorable prognosis in tumors with pathologic indicators of poor prognosis. While in this study, prolonged DFS is observed in colorectal, endometrial and stomach cancers with E/Q/Z mutations, this is not the case in other non-carcinoma cancer types within TCGA. Thus, we find that the positive outcome prediction is cancer type specific. Altogether, results from this study provide a rationale for including POLQ and/or POLZ/REV3L mutations in clinical outcome studies of tumors with POLE mutations. However, future validation is required to confirm the concept that is revealed in the current study.

Supplementary information
Supplementary information accompanies this paper at https://doi.org/10.1186/s12881-020-01089-9.

Additional file 1: Supplementary Figure 1. Number of cases POLE mutations (n = 138) and mutations in the exomes of DNA polymerase genes in PANCAN. Supplementary Figure 2. Mutation counts for colorectal (CORE), stomach (STAD) and endometrial (UCEC) cancers by specific polymerase mutated groups in TCGA data sets. Supplementary Figure 3. Locations of mutations in POLE, POLQ and REV3L exomes in individual colorectal (CORE), stomach (STAD) and uterine cancers (UCEC). Supplementary Figure 4. Relationships between mutation spectrum and mutation counts, POLE, POLQ and REV3L exome mutations, MSI and tumor stage of individual cases. Supplementary Figure 5. Mutation rates per Mb (y-axis) of CORE, STAD and UCEC cases with MSI and E/Q, E/Z and E/Q/Z (x-axis) mutations. Supplementary Figure 6. Relationship between pathology inflammation score and ESTIMATE immune scores for CORE, STAD, and UCEC. Supplementary Figure 7. ESTIMATE immune scores in colorectal (CORE), endometrial (UCEC) and stomach (STAD) cancers. Supplementary Table 1. Number of cases with POLE, POLQ, Z/REV3L, or multiple exome mutations in the PANCAN cohort. Supplementary Table 2. Contingency tables showing number of cases of colorectal, endometrial and stomach cancers in each category.

Abbreviations
POLE: Polymerase epsilon; POLQ: Polymerase theta; POLZ/REV3L: REV3 like, DNA directed polymerase zeta (POLZ) catalytic subunit; TCGA: The Cancer Genome Atlas; MSI: Micro Satellite Instability; MSS: Micro Satellite Stability; TMB: Tumor Mutation Burden; TNM: Tumor, lymph Node, Metastasis; ESTIMA TE: Estimation of STromal and Immune cells in MAIgnant Tumor tissues using Expression data

Acknowledgements
Not applicable.

Authors’ contributions
All authors have read and approved the content of this manuscript. FH: study design and data analysis, HT: data interpretation, JR: data analysis and manuscript writing, BSK: data interpretation and manuscript writing.

Funding
The Prostate Cancer Foundation funded salaries of FH and BSK for computational analysis, Steven Spielberg Team Science Award and RO1CA131255 (BSK) funded salaries for data analysis and paper writing of FH and BSK. We also acknowledge the institutional support of salaries through the NIH G20 RR030860 to the Cedars-Sinai Biobank and Translational Research Core for salaries of BSK and FH. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does any mention of trade names, commercial products, or organizations imply any endorsement by the U.S. Government.

Availability of data and materials
Sequencing data can be obtained from National Cancer Institute GDC data portal and are published in [57]. Raw genomic and clinical data can be found at the NCI Genomic Data Commons https://portal.gdc.cancer.gov/legacy-archive/ The MC3 mutation annotation file can be accessed at https://gdc.cancer.gov/about-data/publications/mc3-2017 and the processed data files can be viewed at https://gdc.cancer.gov/about-data/publications/pancanatlas

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors do not declare any competing interests.

Author details
1Department of Biomedical Sciences, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA. 2Samuel Oschin Cancer Research Institute (SOCCI), Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA. 3Surgery, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA. 4Pathology and Laboratory Medicine, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA. 5Department of Pathology, University of Utah, Salt Lake City, UT 84112, USA.
References

1. Kandath C, Schultz N, Chenniack AD, Albani R, Liu Y, Shen H, Robertson AG, Pastorin I, Shen R, Benz CC, et al. Integrated genomic characterization of endometrial carcinoma. Nature. 2013;497(7447):67–73.

2. Cancer Genome Atlas N. Comprehensive molecular characterization of human colon and rectal cancer. Nature. 2012;487(740):330–7.

3. Church DN, Stellio E, Nout RA, Valtcheva N, Depreeuw J, ter Haar N, Noske A, Arnott F, Tomlinson IP, Wild PJ, et al. Prognostic significance of POLE proofreading mutations in endometrial cancer. J Natl Cancer Inst. 2015;107(1):402.

4. Bilingsley CC, Cohn DE, Mutch DG, Stephens JA, Suarez AA, Goodfellow PJ. Polymerase varepsilon (POLE) mutations in endometrial cancer: clinical outcomes and implications for Lynch syndrome testing. Cancer. 2015;121(3):386–94.

5. Meng B, Hoang LN, McIntyre JB, Duggan MA, Nelson GS, Lee CH, Kobel M. POLE exornuclease domain mutation predicts long progression-free survival in grade 3 endometrial carcinoma of the endometrium. Gynecol Oncol. 2014;134(1):15–9.

6. Imboden S, Nastic D, Ghaderi M, Rydberg F, Rau TT, Mueller MD, Epstein E, Carlson JW. Phenotype of POLE-mutated endometrial cancer. PLoS One. 2019;14(9):e241481.

7. Rayner E, van Gool IC, Palles C, Kearsey SE, Bosse T, Tomlinson I, Church DN. A panoply of errors: polymerase proofreading domain mutations in cancer. Nat Rev Cancer. 2016;16(2):71–81.

8. Briggs S, Tomlinson I. Germline and somatic polymerase epsilon and delta mutations define a new class of hypermutated colorectal and endometrial cancers. J Pathol. 2013;230(2):148–53.

9. Church JM. Polymerase proofreading-associated polymorphs: a new, dominantly inherited syndrome of hereditary colorectal cancer predisposition. Dis Colon Rectum. 2014;57(3):386–7.

10. Jesinghaus M, Pfarr N, Endris V, Kloor M, Volckmar AL, Brandt R, Herpel E, Byrnes G, Tifft A, Brunnauer M, et al. Clinical profile of POLE-mutated colorectal cancer. Cancer. 2017;123(2):188–96.

11. Millar AJ, Wakelin D, Smith S, Cooper CS, et al. The cBio cancer genomics portal: an integrated platform for exploring cancer genomics data in context. Nucleic Acids Res. 2017;45(Database issue):D804–11.

12. Proctor AK, Currin J, Firth H, Tischendorf L, Paul S, birth Orr S, et al. Pathogenic POLE mutations in colorectal cancer are enriched across all European populations. Cancer. 2017;123(1):261–73.

13. Fawzy A, Novara G, Bass AJ, et al. The cBio cancer genomics portal: an integrated platform for exploring cancer genomics data in context. Nucleic Acids Res. 2017;45(Database issue):D804–11.

14. Liu J, Petricoin EF, Garber AJ, et al. DNA polymorphism in colorectal and endometrial cancers. J Pathol. 2013;230(2):148–53.

15. Scholz SW, Richardson AL, Gao J, et al. Comprehensive molecular characterization of gastric adenocarcinoma. Cancer Cell. 2017;84:290–302.

16. Li D, Ruan J,顾 perks J, et al. The cBio cancer genomics portal: an integrated platform for exploring cancer genomics data in context. Nucleic Acids Res. 2017;45(Database issue):D804–11.

17. Zhang Z, Chen H, Zhao X, Cao J, Tong J, Liu W, Wang S, Shao Q, Liu D. REV3L-2UTR 460 T>C polymorphism in microRNA target sites contributes to lung cancer susceptibility. Oncogene. 2013;32(2):242–50.

18. Dubarry M, Lawless C, Banks AP, Cockell S, Lydall D. Genetic Networks Companion: a tool for modeling genome-scale genetic networks. Nucleic Acids Res. 2012;40:751–6.

19. Loeb LA, Matzinger P. DNA polymerses and human disease. Nat Rev Genet. 2008;9(8):594–604.

20. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer S, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012;2(5):401–4.

21. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer S, Sun Y, Jacobsen A, Sinha R, Larsson E, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013;6(296):pl1.

22. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, Sabedot TS, Canalis R, Pina L, et al. TCGAbiolinks: an R/bioconductor companion for integrative analysis of TCGA data. Nucleic Acids Res. 2016;44(12):D279–85.

23. Yoshikara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, Trevino V, Shen H, Laird PW, Levine DA, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013;4:2612.

24. Core R, Team: R: a language and environment for statistical computing. In Vienna, Austria: R Foundation for statistical. Computing. 2015.

25. Killock D. Targeted therapies: DNA polymerase theta—a new target for breast cancer. Breast Cancer Res Treat. 2011;129(1):235–45.

26. Wang W, Sheng W, Yu C, Cao J, Zhou J, Wu J, Zhang H, Zhang S. REV3L modulates cisplatin sensitivity of non-small cell lung cancer H1299 cells. Oncol Rep. 2015;34(3):1460–8.

27. Zhang S, Chen H, Zhao X, Cao J, Tong J, Liu W, Wu W, Shao Q, Liu D. REV3L-2UTR 460 T>C polymorphism in microRNA target sites contributes to lung cancer susceptibility. Oncogene. 2013;32(2):242–50.

28. Munger P, Eberhardt RY, Mistry J, Mitchell AL, Potter SC, Wiemels JL. DNA polymorphisms in human colon and rectal cancer. Nature. 2012;487(7407):330–3.

29. Genet. 2008;9(8):594–604.

30. Mano N. DNA polymerses and human disease. Nat Rev Genet. 2013;9(8):408–10.

31. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, Sabedot TS, Canalis R, Pina L, et al. TCGAbiolinks: an R/bioconductor companion for integrative analysis of TCGA data. Nucleic Acids Res. 2016;44(12):D279–85.

32. Yoshikara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, Trevino V, Shen H, Laird PW, Levine DA, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013;4:2612.

33. Core R, Team: R: a language and environment for statistical computing. In Vienna, Austria: R Foundation for statistical. Computing. 2015.

34. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer S, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012;2(5):401–4.

35. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer S, Sun Y, Jacobsen A, Sinha R, Larsson E, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013;6(296):pl1.

36. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Marshall AP, Potter SC, Punta M, Quevreux M, Sangrador-Vegas A, et al. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. 2016;44(D1):D279–85.

37. Yoshikara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, Trevino V, Shen H, Laird PW, Levine DA, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013;4:2612.

38. Caremi E, Gao J, Dogrusoz U, Gross BE, Sumer S, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012;2(5):401–4.

39. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer S, Sun Y, Jacobsen A, Sinha R, Larsson E, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013;6(296):pl1.
42. McConechy MK, Talhouk A, Leung S, Chiu D, Yang W, Senz J, Reha-Krantz LJ, Lee CH, Huntsman DG, Gilks CB, et al. Endometrial carcinomas with POLE exonuclease domain mutations have a favorable prognosis. Clin Cancer Res. 2016;22(12):2865–73.
43. Popat S, Hubner R, Houlston RS. Systematic review of microsatellite instability and colorectal cancer prognosis. J Clin Oncol. 2005;23(3):509–18.
44. Garg K, Soslow RA. Lynch syndrome (hereditary non-polyposis colorectal cancer) and endometrial carcinoma. J Clin Pathol. 2009;62(8):679–84.
45. Billingsley CC, Cohn DE, Mutch DG, Hade EM, Goodfellow PJ. Prognostic significance of POLE exonuclease domain mutations in high-grade endometrioid endometrial Cancer on survival and recurrence: a subanalysis. Int J Gynecol Cancer. 2016;26(5):933–8.
46. Fox EJ, Prindle MJ, Loeb LA. Do mutator mutations fuel tumorigenesis? Cancer Metastasis Rev. 2013;32(3–4):353–61.
47. Bellone S, Centritto F, Black J, Schwab C, English D, Cocco E, Lopez S, Bonazzoli E, Predolfini F, Ferrari F, et al. Polymerase epsilon (POLE) ultra-mutated tumors induce robust tumor-specific CD4+ T cell responses in endometrial cancer patients. Gynecol Oncol. 2015;138(1):1–7.
48. Shlien A, Campbell BB, de Borja R, Alexandrov LB, Merico D, Wedge D, Van Loo P, Tarpey PS, Coupland P, Behjati S, et al. Combined hereditary and somatic mutations of replication error repair genes result in rapid onset of ultra-hypermutated cancers. Nat Genet. 2015;47(3):257–62.
49. Liu Y, Sethi NS, Hinoue T, Schneider BG, Cheniack AD, Sanchez-Vega F, Seoane JA, Fanishidfar F, Bivbly R, Islam M, et al. Comparative molecular analysis of gastrointestinal adenocarcinomas. Cancer Cell. 2018;33(4):721–35 e728.
50. van Gool IC, Eggink FA, Freeman-Mills L, Stelloo E, Marchi E, de Bruyn M, Palles C, Nout RA, de Kroon CD, Osse EM, et al. POLE proofreading mutations elicit an antitumor immune response in endometrial Cancer. Clin Cancer Res. 2015;21(14):3347–55.
51. Howitt BE, Shukla SA, Sholl LM, Ritterhouse LL, Watkins JC, Rodig S, Stover E, Strickland KC, D’Andrea AD, Wu CJ, et al. Association of Polymerase ε-mutated and microsatellite-instable endometrial cancers with Neoantigen load, number of tumor-infiltrating lymphocytes, and expression of PD-1 and PD-L1. JAMA Oncol. 2015;1(9):1319–23.
52. Hussein YR, Weigelt B, Levine DA, Schoolmeester JK, Dao LN, Balzer BL, Liles G, Karlan B, Kobel M, Lee CH, et al. Clinicopathological analysis of endometrial carcinomas harboring somatic POLE exonuclease domain mutations. Mod Pathol. 2015;28(4):505–14.
53. Bakhsh S, Kinloch M, Hoang LN, Soslow RA, Kobel M, Lee CH, McAlpine JN, McConechy MK, Gilks CB. Histopathological features of endometrial carcinomas associated with POLE mutations: implications for decisions about adjuvant therapy. Histopathology. 2016;68(9):916–24.
54. Nosho K, Baba Y, Tanaka N, Shima K, Hayashi M, Meyerhardt JA, Giovannucci E, Dunoff G, Fuchs CS, Ogino S. Tumour-infiltrating T-cell subsets, molecular changes in colorectal cancer, and prognosis: cohort study and literature review. J Pathol. 2010;222(4):350–66.
55. Budczies J, Seidel A, Christopoulos P, Endris V, Kloor M, Gyoryffy B, Seliger B, Schirmacher P, Stendanger A, Denkert C. Integrated analysis of the immunological and genetic status in and across cancer types: impact of mutational signatures beyond tumor mutational burden. Oncoimmunology. 2018;7(12):e1526613.
56. Smid M, Rodriguez-Gonzalez FG, Siewerts AM, Salgado R, Prager-Van der Smissen WJ, Mubt-Daane MV, van Galen A, Nik-Zainal S, Staal J, Brinkman AB, et al. Breast cancer genome and transcriptome integration implicates specific mutational signatures with immune cell infiltration. Nat Commun. 2016;7:12190.
57. Hoadley KA, Yau C, Hinoue T, Wolf DM, Lazar AJ, Drill E, Shen R, Taylor AM, Cheniack AD, Thorsson V, et al. Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 types of Cancer. Cell. 2018;173(2):291–304 e296.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.