Homologous Recombination Repair Truncations Predict Hypermutation in Microsatellite Stable Colorectal and Endometrial Tumors

Minyi Lee, BS[1], George Eng, MD, PhD[1,2], Stephanie R. Barbari, BS[3], Vikram Deshpande, MBBS[2], Polina V. Shcherbakova, PhD[3] and Manish K. Gala, MD[1,4]

INTRODUCTION: **Somatic mutations in *BRCA1/2* and other homologous recombination repair (HRR) genes have been associated with sensitivity to PARP inhibitors and/or platinum agents in several cancers, whereas hypermutant tumors caused by alterations in *POLE* or mismatch repair genes have demonstrated robust responses to immunotherapy. We investigated the relationship between somatic truncations in HRR genes and hypermutation in colorectal cancer (CRC) and endometrial cancer (EC).**

METHODS: **We analyzed the mutational spectra associated with somatic *BRCA1/2* truncations in multiple genomic cohorts (N = 2,335). From these results, we devised a classifier incorporating HRR genes to predict hypermutator status among microsatellite stable (MSS) tumors. Using additional genomic cohorts (N = 1,439) and functional *in vivo* assays, we tested the classifier to disambiguate *POLE* variants of unknown significance and identify MSS hypermutators without somatic *POLE* exonuclease domain mutations.**

RESULTS: **Hypermutator phenotypes were prevalent among CRCs with somatic *BRCA1/2* truncations (50/62, 80.6%) and ECs with such mutations (44/47, 93.6%). The classifier predicted MSS hypermutators with a cumulative true-positive rate of 100% in CRC and 98.0% in EC and a false-positive rate of 0.07% and 0.63%. Validated by signature analyses of tumor exomes and *in vivo* assays, the classifier accurately reassigned multiple *POLE* variants of unknown significance as pathogenic and identified MSS hypermutant samples without *POLE* exonuclease domain mutations.**

DISCUSSION: **Somatic truncations in HRR can accurately fingerprint MSS hypermutators with or without known pathogenic exonuclease domain mutations in *POLE* and may serve as a low-cost biomarker for immunotherapy decisions in MSS CRC and EC.**

SUPPLEMENTARY MATERIAL accompanies this paper at http://links.lww.com/CTG/A240, http://links.lww.com/CTG/A238, http://links.lww.com/CTG/A239, http://links.lww.com/CTG/A241

Clinical and Translational Gastroenterology 2020;11:e00149. https://doi.org/10.14309/ctg.0000000000000149

INTRODUCTION

Panel-based, somatic mutational profiling of tumors has become routine in the selection of therapies for precision oncology. Several reports have shown that BRCA1 or BRCA2 dysfunction sensitizes cells to PARP inhibition (1,2). Several PARP inhibitors have been approved as monotherapies for *BRCA1/2*-mutated ovarian cancer and *BRCA1/2*-mutated, metastatic triple-negative breast cancer (3–5). In addition, mutations in other homologous recombination repair (HRR) genes have been associated with PARP inhibitor sensitivity (6). Early investigation has begun into the use of PARP inhibitors for other tumor types (7,8). However, the role of these mutations in colorectal cancer (CRC) and endometrial cancer (EC) is unknown.

Molecular profiling has also been used to identify subtypes of cancers responsive to immunotherapy. Microsatellite unstable (MSI-H) CRCs and ECs, and more recently microsatellite stable (MSS) hypermutant tumors (often *POLE*-mutated), have demonstrated robust responses to immunotherapy, primarily attributed to their high neoantigen burden and strong immune infiltrates

[1]Gastrointestinal Unit, Harvard Medical School, Massachusetts General Hospital, Boston, Massachusetts, USA; [2]Department of Pathology, Harvard Medical School, Massachusetts General Hospital, Boston, Massachusetts, USA; [3]Eppley Institute for Research in Cancer and Allied Diseases, Fred & Pamela Buffett Cancer Center, University of Nebraska Medical Center, Omaha, Nebraska, USA; [4]Clinical and Translational Epidemiology Unit, Harvard Medical School, Massachusetts General Hospital, Boston, Massachusetts, USA. **Correspondence:** Manish K. Gala, MD. E-mail: mgala@mgh.harvard.edu.
Minyi Lee, BS and George Eng, MD, PhD share co-first authorship.
Received November 18, 2019; accepted February 13, 2020; published online March 24, 2020

(9–13). Recent survival analyses have demonstrated *POLE*-mutated tumors to be an independent risk factor for identifying individuals who benefited from immune checkpoint inhibitors, with effects on survival similar to those seen with MSI-H tumors (14). Although low-cost methodologies such as immunohistochemistry of mismatch repair proteins and microsatellite testing exist to screen for MSI, comprehensive identification of MSS hypermutators relies on expensive tumor mutational burden assays that require sequencing of >1 million megabases (>400 genes) to achieve reliable results (15,16). Further limiting, these tests are clinically available from only a few commercial laboratories and academic medical centers and lack standardized approaches to calculate and report out high tumor mutational burden samples. Alternatively, sequencing of the exonuclease domain of *POLE* has been embraced as a lower cost alternative by several institutions and clinical trial sponsors. However, this approach limits detection to only those hypermutators with established hotspot pathogenic mutations.

In this study, we investigated the relationship between somatic truncations in *BRCA1/2* and other HRR genes with hypermutant subtypes of CRC and EC and leveraged this unexpected observation to develop an accurate and low-cost methodology to screen for MSS hypermutators using a small number of genes already incorporated by many health care institutions for somatic profiling.

METHODS

Study populations

We performed analyses on 2,335 published samples from the CRC and uterine corpus EC (UCEC) subsets of The Cancer Genome Atlas (TCGA) PanCancer study, the MSK-IMPACT cohort (Memorial Sloan Kettering Cancer Center [MSKCC]) of EC and metastatic CRC (17–19). All clinical characteristics and annotated somatic mutation data were taken from the cBioPortal (20,21). MSI status was unavailable for 11 samples from the TCGA samples. We determined MSI status for these samples using somatic signature analyses from Mutect2 variant call files derived from whole-exome sequencing. Samples from MSK-IMPACT were deemed MSI-H if either established by immunohistochemistry or MSIsensor. Annotated somatic mutations of CRCs and ECs analyzed by DFCI-OncoPanel-3 (Dana-Farber Cancer Institute) and ECs analyzed by MSK-IMPACT 410 and MSK-IMPACT 468 panels ($N = 1{,}439$) were also obtained from the American Association for Cancer Research (AACR) Project Genomics Evidence Neoplasia Information Exchange (GENIE) v5.0 deposited on the cBioPortal (22). MSI status data were unavailable for these AACR Project GENIE samples, conservatively inferred as MSI-H by the presence of *BRAF* V600E mutations, pathogenic mutations in mismatch repair genes with variant allele frequencies approaching 0.5 suggesting germline susceptibility to Lynch syndrome, or the presence of frameshift mutations in *ACVR2A*, *TGFBR2*, or *RNF43* (23). In addition to annotated somatic mutations, nonsynonymous mutation counts were analyzed in context of the sequencing panel used.

Development and validation of a microsatellite stable hypermutator classifier

We devised a classifier that would return positive if 1 or more of 3 criteria were fulfilled:

1. Tumor contained a known pathogenic exonuclease domain mutation in Polε (amino acids 268–471). The set of known pathogenic exonuclease domain mutations was defined as those annotated as pathogenic in ClinVar or previously demonstrated by functional *in vivo* validation studies (24,25).
2. The presence of a variant of unknown significance (VUS) in the exonuclease domain of Polε AND either a somatic truncating mutation (nonsense, splice site, or frameshift) in any one of 13 HRR genes (*BRCA1, BRCA2, PTEN, ATM, ATR, PALB2, MRE11, BARD1, BRIP1, RAD50, RAD51B, RAD51C*, and *RAD51D*) or pathogenic missense mutation in *PTEN*.
3. Somatic truncations in either *BRCA1, BRCA2*, or *PTEN* or a pathogenic missense mutation in *PTEN* AND

Somatic truncation in any of the 13 HRR genes or pathogenic mutation in *PTEN* as long this mutation was not applied to satisfy in the antecedent clause.

Pathogenic missense mutations in *PTEN* were defined as those annotated by OncKB as shown on cBioPortal and incorporated into the criteria, given their previously demonstrated association with *POLE*-mutated cancers (26–28). The CRC subset of the TCGA PanCancer analysis was used as the discovery cohort. The UCEC subset of the TCGA PanCancer analysis and MSKCC Metastatic CRC Cohort were used as validation cohorts. Somatic signature analysis of whole-exome data from TCGA tumors was performed to ascertain true positives of Polε subtypes (see the Statistical Analysis section). In the MSKCC cohorts, samples were already annotated with Polε subtypes in the metadata through similar somatic signature analyses previously performed. The use of *RAD51B* was omitted from the classifier in analysis of tumors from the DFCI-OncoPanel-3 cohort, given its exclusion in the sequencing panel. Sequence data from all other 13 genes in the classifier were available for the other data sets analyzed from AACR GENIE.

Saccharomyces cerevisiae strains and mutation rate measurements

Mutations analogous to those found in tumor specimens were made in the *POL2* gene encoding the catalytic subunit of *S. cerevisiae* Polε. The mutations were created in the *URA3*-based integrative plasmid YIpDK1 containing the wild-type *POL2* fragment (29) by site-directed mutagenesis. To construct haploid *pol2* mutants, the PSD93 diploid (*MAT*a/*MATα ade5-1/ade5-1 lys2:InsE$_{A14}$/lys2:InsE$_{A14}$ trp1-289/trp1-289 his7-2/his7-2 leu2-3,112/leu2-3,112 ura3-52/ura3-52* (30)) was transformed with the linearized YIpDK1-pol2-x plasmid to target integration of the plasmid into the chromosomal *POL2*. In the resulting diploids, one *POL2* locus is intact, and the other contains the *URA3* marker between a full-length *pol2-x* allele and a truncated copy of *POL2* without the mutation. These diploids were then sporulated, and Ura$^+$ haploids containing the *pol2-x* allele were obtained by tetrad dissection. The *URA3* marker was lost through selection on media containing 5-fluoroorotic acid, and clones that retained the full-length *pol2-x* allele were identified by DNA sequencing. The rate of spontaneous mutation was measured by fluctuation analysis as described previously (25).

Statistical Analysis

Somatic mutational signatures were calculated from whole-exome sequencing data processed by the Mutect2 somatic variant caller (31). Mutect2-processed VCF files of TCGA samples were downloaded from the National Cancer Institute Genomic Data Commons and analyzed through a non-negative matrix factorization algorithm to decompose mutational spectra. The DeconstructSigs package in R version 3.3 was used to compare components in the context of the 30 reference signatures identified from the Wellcome Trust Sanger Institute Mutational Signature Framework (32,33). Wellcome Trust Sanger Institute signature 10 has been previously demonstrated to characterize Polε hypermutation (strand bias for C>A mutations at TpCpT context and T>G mutations at TpTpT context), whereas signature 14 describes hypermutation of unknown etiology in a small subset of ECs. Signatures 6, 15, 20, and 26 have been associated with mismatch repair deficiency. Fraction of somatic alterations accounted by each signature was calculated for each tumor sample, and subtype was ascribed to the largest representative signature. Signature 10 analyses for MSKCC samples (Polε subtype) have been precalculated by similar methodologies and have been directly deposited into cBioPortal metadata (34).

The performance of the classifier was analyzed in the context of the MSS tumors present in each cohort. The statistical significance of differences of mutation rates in yeast was assessed using the Mann-Whitney-Wilcoxon nonparametric tests.

RESULTS

Mutation profiles associated with somatic *BRCA1/2* truncations

The clinical characteristics of the 4 genomic cohorts initially analyzed are described in Table 1. In the TCGA CRC PanCancer and the MSKCC metastatic CRC cohorts, 3.9% (21/533) and 3.7% (41/1,099) of samples harbored somatic truncating mutations (nonsense, splice site, or frameshift) in *BRCA1/2*, respectively (Figure 1a). In the TCGA UCEC PanCancer and MSKCC EC cohorts, 8.2% (42/515) and 2.7% (5/188) of samples had *BRCA1/2* truncations, respectively. Mutational signature analysis of *BRCA1/2*-truncated tumors from the TCGA cohorts demonstrated that 85.7% of such CRCs showed hypermutant signatures (38.1% Polε subtype, 47.6% MSI-H), and 95.3% of such ECs (54.8% Polε subtype, 40.5% MSI-H). Among the published MSKCC cohorts, 78% of the *BRCA1/2*-mutated CRCs demonstrated either known pathogenic Polε exonuclease mutations (14.6%) or MSI (63.4%). All 5 *BRCA1/2*-mutated tumors in the EC cohort exhibited MSI-H phenotypes. Among MSI-H tumors, frameshift mutations were more common than nonsense and splice site mutations.

Exome-wide survey of the other somatic mutations present in TCGA *BRCA1/2*-mutated CRC and EC demonstrated co-occurring truncations in additional HRR genes and frequent, pathogenic missense variants in *PTEN* (Figure 1b).

Fourteen-gene classifier to identify microsatellite stable hypermutators

Given the hypermutant profiles exhibited in *BRCA1/2*-truncated tumors and frequent co-occuring HRR gene mutations, we subsequently devised a classifier to identify MSS hypermutators using *POLE* and 13 HRR genes routinely tested in germline cancer predisposition and/or PARP inhibitor sensitivity assays. We used MSS tumors from the TCGA CRC PanCancer cohort as our discovery data set. The MSKCC metastatic CRC cohort and the TCGA UCEC PanCancer cohort were used as validation data sets. The endometrial MSKCC data set was excluded, given the paucity of MSS hypermutators.

The classifier demonstrated a true-positive rate (TPR) of 100% and a false-positive rate (FPR) of 0% in the discovery data set. The TPR/FPR for each HRR gene in the discovery cohort can be seen in Figure 2, and the performance of each criterion in the classifier in Table 2. Notably, criterion 3, which only requires 2 qualifying mutations in HRR irrespective of Polε status, demonstrated a TPR of 75% and a FPR of 0%.

The performance of the classifier among discovery and validation cohorts is presented in Table 3. In the MSKCC metastatic CRC validation data set, the TPR of the classifier was 100%

Table 1. Characteristics of the genomic cohorts analyzed

	Colorectal cancer		Endometrial cancer	
Characteristic	**TCGA PanCancer (n = 533)**	**MSKCC (n = 1,099)**	**TCGA PanCancer (n = 515)**	**MSKCC (n = 188)**
Sex, no. (%)				
Male	277 (52.0)	597 (54.3)	0 (0.0)	0 (0.0)
Female	254 (47.6)	502 (45.7)	515 (100.0)	188 (100.0)
Missing	2 (0.4)	0 (0.0)	0 (0.0)	0 (0.0)
MSI status (%)				
MSI-H	68 (12.8)	144 (13.1)	150 (29.1)	30 (16.0)
MSS	465 (87.2)	955 (86.9)	365 (70.9)	158 (84.0)
Stage, no. (%)				
I	95 (17.8)	40 (3.6)	319 (61.9)[a]	N/A
II	200 (37.5)	128 (11.6)	50 (9.7)[a]	N/A
III	151 (28.3)	267 (24.3)	119 (23.1)[a]	N/A
IV	73 (13.7)	664 (60.4)	27 (5.2)[a]	N/A
Missing	14 (2.6)	0 (0.0)	0 (0.0)[a]	N/A

MSI, microsatellite instability; MSI-H, microsatellite unstable; MSKCC, Memorial Sloan Kettering Cancer Center; MSS, microsatellite stable; TCGA, The Cancer Genome Atlas.
[a]Clinical staging was provided for the Uterine Corpus Endometrial Carcinoma subset of the TCGA PanCancer analysis.

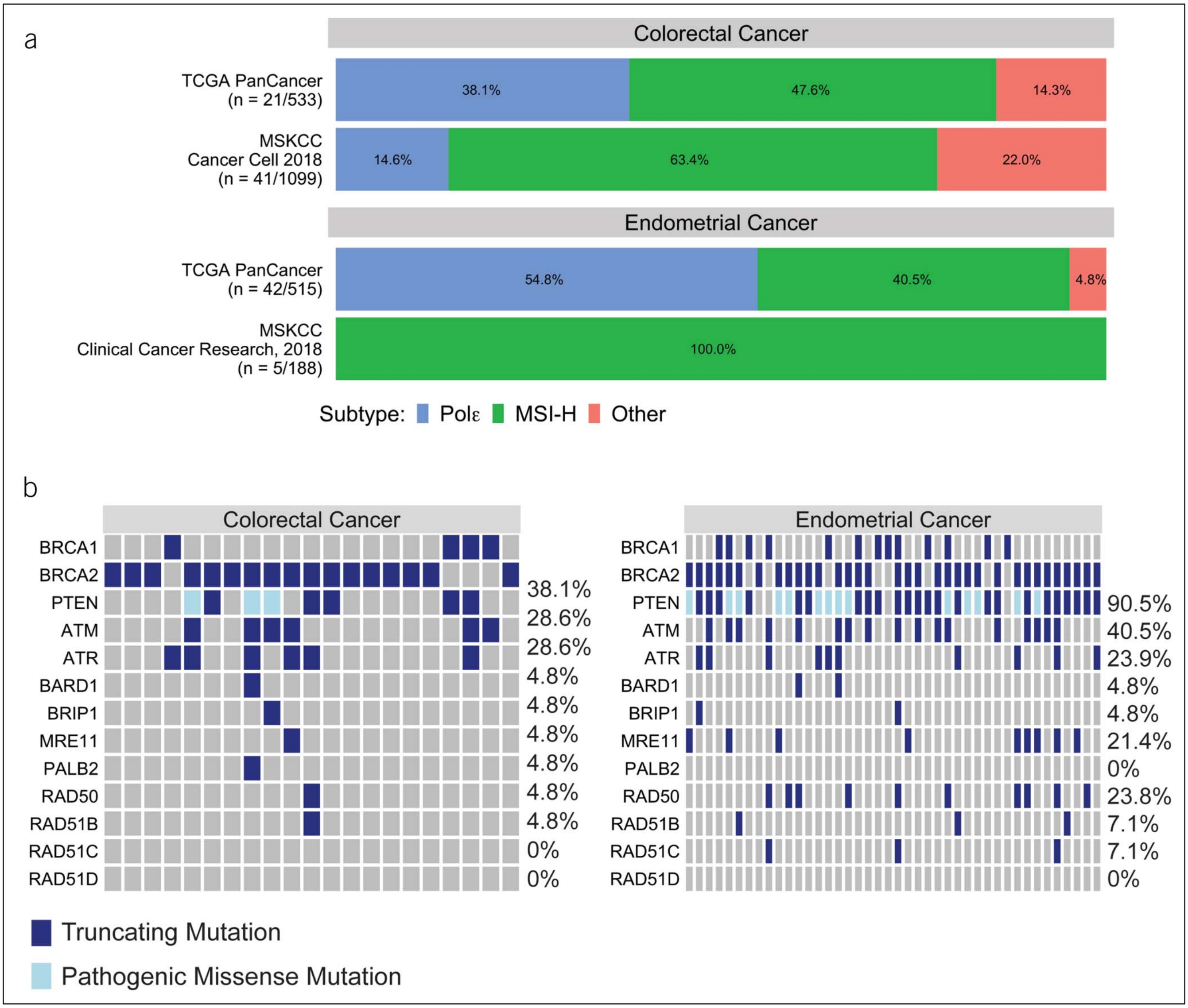

Figure 1. Mutational profiles associated with *BRCA1/2* truncations in colorectal and endometrial cancers. (**a**) Hypermutator signatures associated with *BRCA1/2* truncations in colorectal and endometrial cancers. (**b**) Co-occurring mutations in homologous recombination repair genes present in *BRCA1/2*-mutated colorectal cancers and endometrial cancers from the TCGA PanCancer analysis. MSI-H, microsatellite unstable; MSKCC, Memorial Sloan Kettering Cancer Center; TCGA, The Cancer Genome Atlas.

(7/7), and the FPR was 0.11% (1/948). The positive predictive value was 87.5%, and the negative predictive value was 100%. In the TCGA UCEC PanCancer data set, the TPR was 98.0% (48/49), and the FPR was 0.63%. In this cohort, the positive predictive value was 96.0%, and the negative predictive value was 99.7%. All false positives (3) in the validation cohorts had no pathogenic exonuclease domain mutation in *POLE*, but demonstrated concurrent pathogenic mutations in *PTEN* and truncations in *ATM*. The performance of each criterion among the validation data sets is presented in Table 4.

Disambiguation of *POLE* variants of unknown significance and identification of hypermutators without exonuclease mutations

The classifier was able to correctly identify 1 hypermutant CRC with a Polε mutational signature in the discovery cohort without somatic alteration in the exonuclease domains of *POLE* or *POLD1*. The tumor harbored truncations in multiple HRR genes: *ATM, ATR, BRCA2,* and *MRE11A*. Furthermore, all tumors with a *POLE* VUS in the exonuclease domain and a qualifying mutation in 1 of 13 HRR genes were accurately classified as having a Polε mutational signature (see Table, Supplementary Digital Content 1, http://links.lww.com/CTG/A238). Given the performance of the classifier in the discovery and validation cohorts, we next hypothesized that the classifier could serve to disambiguate VUS as pathogenic and identify additional hypermutators without known exonuclease mutations. To confirm this hypothesis, we analyzed additional clinical cohorts (N = 1,439) from the AACR Project GENIE version 5.0.

Given the lack of MSI status information in AACR Project GENIE, we selected all tumors with nonsynonymous mutation counts that were equal to or higher than the tumor with the lowest mutation count harboring a known pathogenic Polε mutation (Figure 3a and see Table, Supplementary Digital

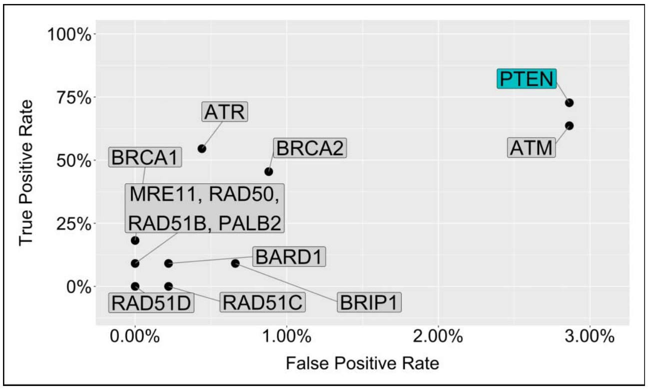

Figure 2. Performance of each individual homologous recombination repair gene for identifying microsatellite stable hypermutators in the discovery cohort. Any pathogenic mutation in *PTEN* was considered and highlighted in blue, whereas truncating mutations (nonsense, splice site, and frameshift) were only considered for the other homologous recombination repair genes. The discovery cohort was composed of microsatellite stable colorectal cancers from the TCGA PanCancer data set. TCGA, The Cancer Genome Atlas.

Content 2, http://links.lww.com/CTG/A239). Tumors harboring mutations associated with MSI or demonstrating overlap with MSKCC samples already analyzed were removed (23). Application of the classifier resulted as positive for all MSS hypermutant samples (34/34), including 5 with Polε VUS in the exonuclease domain and 2 without somatic Polε exonuclease mutations. Thus, the classifier increased the number of identified MSS hypermutant tumors by 26% over a strategy of using known pathogenic *POLE* mutations.

The most frequent VUS associated with a positive classification in all cohorts was the A456P mutation. Previous studies have inferred that this mutation might be pathogenic due to its location in the exonuclease domain and occurrence in Polε hypermutated tumors (28,35–37). However, no functional *in vitro* or *in vivo* assays have been performed to confirm its ability to impair DNA proofreading or elevate the mutation rate.

Table 2. Performance of each criterion from the classifier in the discovery cohort

Criterion	TPR	FPR	PPV	NPV
1	10/12	0/453	100%	99.6%
2	1/12	0/453	100%	97.6%
3	8/12	0/453	100%	99.1%
1 + 2	11/12	0/453	100%	99.8%
1 + 3	11/12	0/453	100%	99.8%
1 + 2 + 3	12/12	0/453	100%	100%

Criterion 1 consists of known pathogenic exonuclease domain mutations in *POLE*. Criterion 2 consists of variants of unknown significance in the exonuclease domain with a qualifying mutation in a homologous recombination repair gene. Criterion 3 consists of 2 qualifying mutations in homologous recombination repair genes. The discovery cohort comprised microsatellite stable colorectal cancers from the TCGA PanCancer analysis.
FPR, false-positive rate; NPV, negative predictive value; PPV, positive predictive value; TCGA, The Cancer Genome Atlas; TPR, true-positive rate.

Accordingly, ClinVar has designated the variant as of unknown significance. We performed mutational signature analyses of exomes from all Polε A456P-mutated cancers in TCGA, irrespective of tumor type. All tumors (4/4) bore signatures consistent with hypermutation (Figure 3b). For direct *in vivo* confirmation, we modeled this mutation in yeast and determined its effect on the mutation rate. The amino acid sequence around alanine 456 is highly conserved between human and yeast Polε with the exception of 3 residues that include alanine 456 itself (serine in yeast), threonine 454, and threonine 457 (see Figure, Supplementary Digital Content 3, http://links.lww.com/CTG/A240). We first constructed a yeast strain in which the entire 454–457 amino acid segment (HLSE) was replaced with the corresponding human sequence (TLAT) to mimic the wild-type human Polε. We then introduced a mutation to replace the alanine in the TLAT sequence with a proline to mimic A456P. The mutation rate was measured using 3 different reporter assays. The *CAN1* forward mutation reporter detects a variety of single-base substitutions, frameshifts, and larger rearrangements that inactivate the gene. The *his7-2* reversion reporter scores predominantly +1 frameshifts. The *lys2-InsE*$_{A14}$ reporter allele scores frameshift mutations in a long homonucleotide run, thus providing a readout for MSI. The HLSE-to-TLAT substitution did not affect mutagenesis in any of the assays. The A456P mimic increased the rate of *CAN1* mutation and *his7-2* reversion 3.7-fold and 4.6-fold, respectively, compared with wild type and did not affect instability at the *lys2-InsE*$_{A14}$ locus (Figure 3C and see Table, Supplementary Digital Content 4, http://links.lww.com/CTG/A241), consistent with the hypermutator MSS phenotype of A456P tumors. To further validate our classifier's ability to resolve VUS, we selected an additional VUS (M295R) from a TCGA UCEC sample successfully identified as hypermutant by our classifier. Modeling of the M295R mutation in yeast demonstrated a 16- and 19-fold increase in the mutation rate over wild type at the *CAN1* and *his7-2* loci, respectively, with a minimal effect on the instability of the *lys2- InsE*$_{A14}$ homonucleotide run.

DISCUSSION

In this study, we demonstrate an unexpected finding that somatic truncations in *BRCA1/2* and other HRR genes are highly specific for hypermutator phenotypes in CRC and EC. Although previous studies have observed small increases in tumor mutation burden for other solid tumor types with somatic mutations in HRR genes, no previous association with hypermutation has been reported (38). We leveraged this novel association to develop and validate a 14-gene classifier for MSS hypermutators through the analysis of publicly available cohorts. Furthermore, we used the sensitivity and specificity of the classifier to formally disambiguate VUS in *POLE* and identify MSS hypermutators without any detectable exonuclease domain mutations. Through analysis of additional genomic cohorts, our classifier identified an additional 26% more MSS hypermutant cancers over the existing strategy of testing for known pathogenic *POLE* mutations.

It should be noted that the MSS hypermutant tumors identified without exonuclease domain mutations still demonstrated *POLE*-mutated signatures; these tumors harbored strand bias for C>A mutations at TpCpT context and T>G mutations at TpTpT context. The lack of missense exonuclease domain

COLON

Table 3. Performance of a 14-gene classifier to identify microsatellite stable hypermutators

Cohort	TPR	FPR	PPV	NPV
TCGA PanCancer colorectal cancer	100% (12/12)	0% (0/453)	100%	100%
MSKCC metastatic colorectal cancer	100% (7/7)	0.11% (1/948)	87.5%	100%
TCGA PanCancer uterine corpus endometrial carcinoma	98.0% (48/49)	0.63% (2/316)	96.0%	99.7%

Only microsatellite stable tumors were analyzed in each cohort. The colorectal subset of the TCGA PanCancer analysis was used as discovery. The other 2 cohorts were used as validation.
FPR, false-positive rate; MSKCC, Memorial Sloan Kettering Cancer Center; NPV, negative predictive value; PPV, positive predictive value; TCGA, The Cancer Genome Atlas; TPR, true-positive rate.

mutations in these tumors may suggest the existence of alternative genetic mechanisms of how the proofreading ability of *POLE* may be disrupted. For example, current gene sequencing strategies often fail to detect mutations in deep intronic segments that may cause aberrant splicing events of the exonuclease domain. Alternatively, pathogenic missense mutations may be miscalled, given that no tumor variant calling algorithm demonstrates perfect accuracy (39).

This study highlights the continued importance of interrogating tissue specificity for pharmacogenomic associations. The promulgation of molecular basket trials, where the effect of 1 drug on a single mutation is evaluated in a variety of tumor types at the same time, has changed the paradigm of clinical trials in precision oncology. Our work reinforces the additional importance of global assessments of tumor mutations in therapy decisions. Analyses of single gene mutations in the broader context of exome-wide signatures enabled us to postulate a novel pharmacogenomic association of HRR genes with immunotherapy. Moreover, this analysis could only be performed thanks to the commitment of academic institutions and nonprofit organizations to share genomic data sets to the wider scientific community.

Immune checkpoint blockade has advanced the treatment of CRC and EC with high neoantigen loads, particularly those tumors with MSI (9,10). Accounting for ~2% of CRCs and 7% of ECs, MSS hypermutant tumors often demonstrate neoantigen burdens that exceed those of MSI-H tumors and have also demonstrated robust responses to immune checkpoint blockade and survival advantages (11–14,40–43). Despite these observations, identification of MSS hypermutant cancers for clinical trial recruitment and palliative immunotherapy remains challenging. The absence of a low-cost, comprehensive screening test and the lower prevalence of MSS hypermutant tumors than MSI-H tumors contribute to their underrecognition and undertreatment in the clinical environment. To address this issue, we intentionally selected genes in our classifier that are routinely tested in panels performed by many medical centers. Furthermore, the nature of the criteria also allows for medical providers to scale down the number of HRR genes sequenced (as long as a minimum *PTEN*, *BRCA1*, and *BRCA2* are incorporated) for an acceptable trade-off in sensitivity, yet without compromising specificity.

Although we analyzed multiple cohorts that have been previously deemed suitable for US Food and Drug Administration applications for accompanying diagnostics, future prospective trials and cost-effective analyses will be required to further demonstrate the clinical utility of our classifier as a comprehensive, low-cost alternative to tumor mutation burden testing for immunotherapy in these cancer types.

Table 4. Performance of each criterion from the classifier in the validation cohort

	Criterion	TPR	FPR	PPV	NPV
MSKCC metastatic colorectal cancer	1	7/7	0/948	100%	100%
	2	0/7	0/948	0%	99.3%
	3	6/7	1/948	85.7%	99.9%
	1 + 2	7/7	0/948	100%	100%
	1 + 3	7/7	1/948	87.5%	100%
	1 + 2 + 3	7/7	1/948	87.5%	100%
TCGA PanCancer uterine corpus endometrial carcinoma	1	44/49	0/316	100%	98.4%
	2	4/49	0/316	100%	87.5%
	3	38/49	2/316	95%	96.6%
	1 + 2	48/49	0/316	100%	99.7%
	1 + 3	45/49	2/316	100%	98.7%
	1 + 2 + 3	48/49	2/316	100%	99.7%

Criterion 1 consists of known pathogenic exonuclease domain mutations in *POLE*. Criterion 2 consists of variants of unknown significance in the exonuclease domain with a qualifying mutation in a homologous recombination repair gene. Criterion 3 consists of 2 qualifying mutations in homologous recombination repair genes. Microsatellite stable cancers were only analyzed from each cohort.
FPR, false-positive rate; MSKCC, Memorial Sloan Kettering Cancer Center; NPV, negative predictive value; PPV, positive predictive value; TCGA, The Cancer Genome Atlas; TPR, true-positive rate.

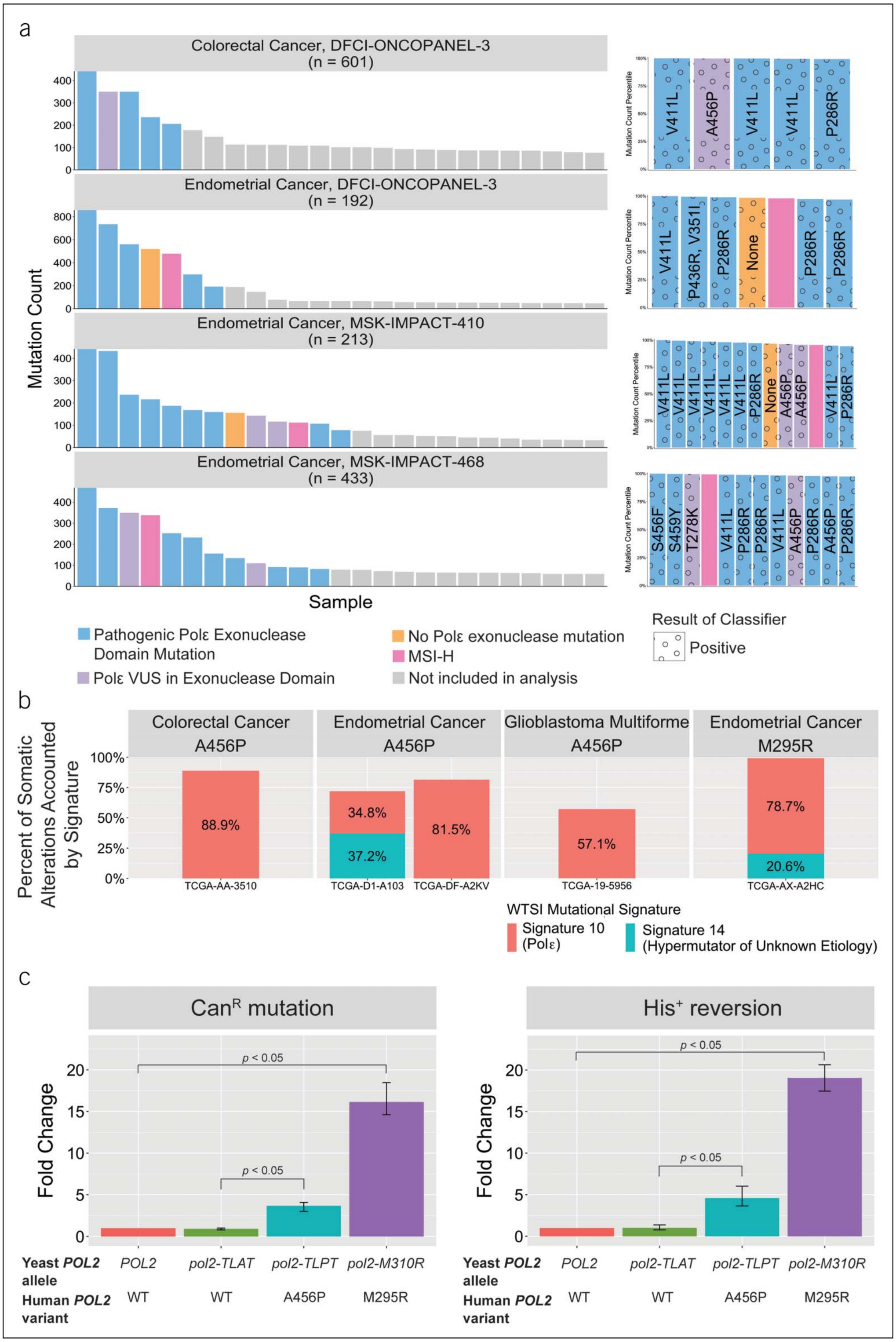

Figure 3. Identification of hypermutators without Polε exonuclease domain mutations and disambiguation of variants of unknown significance. (**a**) Application of the classifier to colorectal and endometrial cancer cohorts from the AACR GENIE project. All tumors with mutation counts ≥ tumors with a known pathogenic exonuclease domain mutation were analyzed. The classifier correctly identified all hypermutators without features of microsatellite instability including those with variants of unknown significance and those without detectable exonuclease mutations. (**b**) Somatic signature analysis of exomes from microsatellite stable cancers with Polε A456P and M295R demonstrates hypermutation in all tumors where such mutations are found. (**c**) Fold increase in various measures of the mutation rate by *in vivo* modeling of Polε A456P and M295R mutations in yeast. AACR, American Association for Cancer Research; GENIE, Genomics Evidence Neoplasia Information Exchange.

CONFLICTS OF INTEREST

Guarantor of the article: Manish K. Gala, MD.
Specific author contributions: M.L, G.E., and S.R.B. designed and performed experiments and wrote the paper. V.D. helped analyze the data. P.V.S. supervised experiments, analyzed data, and helped write the manuscript. M.K.G. conceived the study, helped analyze the data, supervised research, and wrote the manuscript.
Financial support: Funded by the American College of Gastroenterology (M.K.G.) and the National Institutes of Health (T32CA009476 to S.R.B., ES015869 to P.V.S., and DK103119 to M.K.G.)
Potential competing interests: M.K.G. is a cofounder and has equity in New Amsterdam Genomics, Inc. No funding, support, or input from New Amsterdam Genomics was received for this study. The other authors have no conflicts to report.

ACKNOWLEDGEMENTS

We would like to thank Dr. Andrew T. Chan for his helpful comments and Liz Moore for her technical assistance.

Study Highlights

WHAT IS KNOWN

- ✓ *POLE*-mutant tumors account for 1% of CRCs and 7% of ECs.
- ✓ Treatment of *POLE*-mutant, MSS CRC and EC with immunotherapy demonstrates survival advantages similar to those observed with microsatellite unstable tumors.
- ✓ Given the cost-prohibitive nature of tumor mutation burden assays, most clinical providers rely on sequencing the exonuclease domain of *POLE* to identify MSS hypermutant cancers.

WHAT IS NEW HERE

- ✓ Somatic truncations in HRR genes can be used to disambiguate VUS in the exonuclease domain of *POLE.*
- ✓ Somatic truncations in HRR genes can be used to identify MSS hypermutators without *POLE* exonuclease domain mutations.
- ✓ Incorporation of a limited number of HRR genes can increase identification of MSS hypermutant cancers by up to 26% over a strategy of sequencing *POLE* alone.

TRANSLATIONAL IMPACT

- ✓ Sequencing of *POLE* and HRR genes already incorporated into the somatic profiling panels of most cancer centers is a low-cost alternative to tumor mutation burden assays to identify MSS hypermutant colorectal and ECs.

REFERENCES

1. Farmer H, McCabe N, Lord CJ, et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. Nature 2005;434:917–21.
2. Bryant HE, Schultz N, Thomas HD, et al. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. Nature 2005;434:913–7.
3. Litton JK, Rugo HS, Ettl J, et al. Talazoparib in patients with advanced breast cancer and a germline BRCA mutation. N Engl J Med 2018;379:753–63.
4. Robson M, Im SA, Senkus E, et al. Olaparib for metastatic breast cancer in patients with a germline BRCA mutation. N Engl J Med 2017;377:523–33.
5. Moore K, Colombo N, Scambia G, et al. Maintenance olaparib in patients with newly diagnosed advanced ovarian cancer. N Engl J Med 2018;379:2495–505.
6. Lord CJ, Ashworth A. PARP inhibitors: Synthetic lethality in the clinic. Science 2017;355:1152–8.
7. Ghiringhelli F, Richard C, Chevrier S, et al. Efficiency of olaparib in colorectal cancer patients with an alteration of the homologous repair protein. World J Gastroenterol 2016;22:10680–6.
8. Bian X, Gao J, Luo F, et al. PTEN deficiency sensitizes endometrioid endometrial cancer to compound PARP-PI3K inhibition but not PARP inhibition as monotherapy. Oncogene 2018;37:341–51.
9. Le DT, Uram JN, Wang H, et al. PD-1 blockade in tumors with mismatch-repair deficiency. N Engl J Med 2015;372:2509–20.
10. Le DT, Durham JN, Smith KN, et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. Science 2017;357:409–13.
11. Veneris JT, Lee EK, Goebel EA, et al. Diagnosis and management of a recurrent polymerase-epsilon (POLE)-mutated endometrial cancer. Gynecol Oncol 2019;153:471–8.
12. Silberman R, Steiner DF, Lo AA, et al. Complete and prolonged response to immune checkpoint blockade in POLE-mutated colorectal cancer. JCO Precision Oncol 2019:1–5.
13. Mehnert JM, Panda A, Zhong H, et al. Immune activation and response to pembrolizumab in POLE-mutant endometrial cancer. J Clin Invest 2016;126:2334–40.
14. Wang F, Zhao Q, Wang YN, et al. Evaluation of POLE and POLD1 mutations as biomarkers for immunotherapy outcomes across multiple cancer types. JAMA Oncol 2019;5:1504–6.
15. Gong J, Pan K, Fakih M, et al. Value-based genomics. Oncotarget 2018;9:15792–815.
16. Melendez B, Van Campenhout C, Rorive S, et al. Methods of measurement for tumor mutational burden in tumor tissue. Transl Lung Cancer Res 2018;7:661–7.
17. Cancer Genome Atlas Research N, Weinstein JN, Collisson EA, et al. The cancer genome Atlas pan-cancer analysis project. Nat Genet 2013;45:1113–20.
18. Yaeger R, Chatila WK, Lipsyc MD, et al. Clinical sequencing defines the genomic landscape of metastatic colorectal cancer. Cancer Cell 2018;33:125–36.e3.
19. Soumerai TE, Donoghue MTA, Bandlamudi C, et al. Clinical utility of prospective molecular characterization in advanced endometrial cancer. Clin Cancer Res 2018;24:5939–47.
20. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012;2:401–4.
21. Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal 2013;6:pl1.
22. Consortium APG. AACR project GENIE: Powering precision medicine through an international consortium. Cancer Discov 2017;7:818–31.
23. Cortes-Ciriano I, Lee S, Park WY, et al. A molecular portrait of microsatellite instability across multiple cancers. Nat Commun 2017;8:15180.
24. Landrum MJ, Lee JM, Benson M, et al. ClinVar: Improving access to variant interpretations and supporting evidence. Nucleic Acids Res 2018;46:D1062–7.
25. Barbari SR, Kane DP, Moore EA, et al. Functional analysis of cancer-associated DNA polymerase epsilon variants in Saccharomyces cerevisiae. G3 (Bethesda) 2018;8:1019–29.
26. Hussein YR, Weigelt B, Levine DA, et al. Clinicopathological analysis of endometrial carcinomas harboring somatic POLE exonuclease domain mutations. Mod Pathol 2015;28:505–14.
27. Chakravarty D, Gao J, Phillips SM, et al. OncoKB: A precision oncology knowledge base. JCO Precis Oncol 2017;2017.
28. Shinbrot E, Henninger EE, Weinhold N, et al. Exonuclease mutations in DNA polymerase epsilon reveal replication strand specific mutation patterns and human origins of replication. Genome Res 2014;24:1740–50.
29. Kane DP, Shcherbakova PV. A common cancer-associated DNA polymerase epsilon mutation causes an exceptionally strong mutator phenotype, indicating fidelity defects distinct from loss of proofreading. Cancer Res 2014;74:1895–901.
30. Daee DL, Mertz TM, Shcherbakova PV. A cancer-associated DNA polymerase delta variant modeled in yeast causes a catastrophic increase in genomic instability. Proc Natl Acad Sci USA 2010;107:157–62.
31. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol 2013;31:213–9.

32. Rosenthal R, McGranahan N, Herrero J, et al. DeconstructSigs: Delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. Genome Biol 2016;17:31.
33. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. Nature 2013;500:415–21.
34. Zehir A, Benayed R, Shah RH, et al. Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. Nat Med 2017;23:703–13.
35. Rayner E, van Gool IC, Palles C, et al. A panoply of errors: Polymerase proofreading domain mutations in cancer. Nat Rev Cancer 2016;16:71–81.
36. Church DN, Briggs SE, Palles C, et al. DNA polymerase epsilon and delta exonuclease domain mutations in endometrial cancer. Hum Mol Genet 2013;22:2820–8.
37. Campbell BB, Light N, Fabrizio D, et al. Comprehensive analysis of hypermutation in human cancer. Cell 2017;171:1042–56 e10.
38. Strickland KC, Howitt BE, Shukla SA, et al. Association and prognostic significance of BRCA1/2-mutation status with neoantigen load, number of tumor-infiltrating lymphocytes and expression of PD-1/PD-L1 in high grade serous ovarian cancer. Oncotarget 2016;7:13587–98.
39. Bian X, Zhu B, Wang M, et al. Comparing the performance of selected variant callers using synthetic data and genome segmentation. BMC Bioinformatics 2018;19:429.
40. Cancer Genome Atlas Research Network, Kandoth C, Schultz N, Cherniack AD, et al. Integrated genomic characterization of endometrial carcinoma. Nature 2013;497:67–73.
41. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. Nature 2012;487:330–7.
42. Gong J, Wang C, Lee PP, et al. Response to PD-1 blockade in microsatellite stable metastatic colorectal cancer harboring a POLE mutation. J Natl Compr Canc Netw 2017;15:142–7.
43. Samstein RM, Lee CH, Shoushtari AN, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. Nat Genet 2019;51:202–6.