Received: 23 January 2020 | Revised: 26 May 2020 | Accepted: 28 May 2020

DOI: 10.1002/mgg3.1368

ORIGINAL ARTICLE

Molecular Genetics & Genomic Medicine Open Access WILEY

POLE and POLD1 germline exonuclease domain pathogenic variants, a rare event in colorectal cancer from the Middle East

Abdul K. Siraj[1] | **Rong Bu**[1] | **Kaleem Iqbal**[1] | **Sandeep K. Parvathareddy**[1] | **Tariq Masoodi**[1] | **Nabil Siraj**[1] | **Maha Al-Rasheed**[1] | **Yan Kong**[1] | **Saeeda O. Ahmed**[1] | **Khadija A. S. Al-Obaisi**[1] | **Ingrid G. Victoria**[1] | **Maham Arshad**[1] | **Fouad Al-Dayel**[2] | **Alaa Abduljabbar**[3] | **Luai H. Ashari**[3] | **Khawla S. Al-Kuraya**[1]

[1]Human Cancer Genomic Research, Research Center, King Faisal Specialist Hospital and Research Center, iyadh, Saudi Arabia

[2]Department of Pathology and Laboratory Medicine, King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia

[3]Colorectal Section, Department of Surgery, King Faisal Specialist Hospital and Research Center, Riyadh, Saudi Arabia

Correspondence
Khawla S. Al-Kuraya, Human Cancer Genomic Research, Research Center, King Faisal Specialist Hospital and Research Center, MBC#98-16, P.O. Box 3,354, Riyadh 11211, Saudi Arabia.
Email: kkuraya@kfshrc.edu.sa

Abstract

Background: Colorectal cancer (CRC) is a major contributor to morbidity and mortality related to cancer. Only ~5% of all CRCs occur as a result of pathogenic variants in well-defined CRC predisposing genes. The frequency and effect of exonuclease domain pathogenic variants of *POLE* and *POLD1* genes in Middle Eastern CRCs is still unknown.

Methods: Targeted capture sequencing and Sanger sequencing technologies were employed to investigate the germline exonuclease domain pathogenic variants of *POLE* and *POLD1* in Middle Eastern CRCs. Immunohistochemical analysis of POLE and POLD1 was performed to look for associations between protein expression and clinico-pathological characteristics.

Results: Five damaging or possibly damaging variants (0.44%) were detected in 1,135 CRC cases, four in *POLE* gene (0.35%, 4/1,135) and one (0.1%, 1/1,135) in *POLD1* gene. Furthermore, low POLE protein expression was identified in 38.9% (417/1071) cases and a significant association with lymph node involvement ($p = .0184$) and grade 3 tumors ($p = .0139$) was observed. Whereas, low POLD1 expression was observed in 51.9% (555/1069) of cases and was significantly associated with adenocarcinoma histology ($p = .0164$), larger tumor size (T3 and T4 tumors; $p = .0012$), and stage III tumors ($p = .0341$).

Conclusion: *POLE* and *POLD1* exonuclease domain pathogenic variants frequency in CRC cases was very low and these exonuclease domain pathogenic variants might be rare causative events of CRC in the Middle East. *POLE* and *POLD1* can be included in multi-gene panels to screen CRC patients.

KEYWORDS

colorectal cancers, Middle East, *POLD1*, *POLE*, variant

Abdul K. Siraj and Rong Bu contributed equally to the study.

1 | INTRODUCTION

Colorectal cancer (CRC) is one of the main causes of cancer-related morbidity and mortality worldwide (Bray et al., 2018). In Saudi Arabia, CRC ranked first among males and third in females, accounting for 12.2% of all newly diagnosed cancer cases in 2015. Furthermore, the median age at diagnosis is approximately 10 years younger in Saudi Arabia compared to Western countries (Alrawaji, Alshahrani, Alzahrani, Alomran, & Almadouj, 2018; Howlader et al., 2016), and a nearly two-fold increase in age-standardized rate of CRC was observed in Saudi Arabia from 1994 to 2015 whereas the incidence of CRC has stabilized or decreased in the West (Al-Eid, 1994; Alrawaji et al., 2018). The causes of these disparities are still not fully elucidated, but the differences of genetic features in different populations may play an important role.

Although up to one-third of CRCs are attributed to heritable genetic factors, only ~5% of all CRCs are known to be caused by well-defined, inherited germline pathogenic variants in CRC predisposing genes with high penetrance including *MLH1, MSH2, MSH6,* and *PMS2* (MMR genes), *EPCAM* and *MUTYH* (Al-Tassan et al., 2002; Grady, 2003; Jasperson, Tuohy, Neklason, & Burt, 2010; Lichtenstein et al., 2000; Weren et al., 2015). The remaining cases are not fully understood. There is a probability that the pathogenic variants in less penetrant genes could be associated with the well-characterized syndromes. Therefore, a precise understanding of the genetic characteristics of CRC is critical for identifying individuals with high risk to develop CRC, improving cancer prevention and surveillance strategies and developing better diagnostic and therapeutic approaches.

Owing to the advances in sequencing technology, several novel cancer-related genes have recently been discovered, including *POLE* (OMIM: *174,762) and *POLD1* (OMIM: *174,761) genes (Bellido et al., 2016; Calva-Cerqueira et al., 2010; Gala et al., 2014; Houlston et al., 1998; Spier et al., 2015). *POLE* and *POLD1* encode the catalytic subunit of the polymerase enzyme complexes Epsilon (ε) and Delta (δ), respectively, and play a key role in DNA replication and repair (Nick McElhinny, Gordenin, Stith, Burgers, & Kunkel, 2008; Pursell, Isoz, Lundstrom, Johansson, & Kunkel, 2007). Formal estimates of penetrance suggest that pathogenic variants in POLD1 are associated with high penetrance for the disease (Bellido et al., 2016; Buchanan et al., 2018; Palles et al., 2013; Spier et al., 2015). Subsequent screenings revealed that exonuclease domain pathogenic variants were present in 0.1%–0.8% of CRCs with family history of CRC, multiple adenomas or early-onset in Western populations (Chubb et al., 2015; Elsayed et al., 2015; Palles et al., 2013; Valle et al., 2014), whereas relatively higher frequency (2.9%) of pathogenic variants in *POLE* and *POLD1* genes was seen in a CRC cohort of Chinese population (Dong et al., 2019). Furthermore, the exonuclease domain pathogenic variants of these two genes were shown to be associated with hyper-mutation phenotype and significant response to immunotherapy using immune checkpoints inhibitor (Bourdais et al., 2017; Ciardiello et al., 2019; Gong, Wang, Lee, Chu, & Fakih, 2017; Palles et al., 2013). However, the underlying causes of CRC in genetically unexplained cases and their clinico-pathological characteristics in CRCs from the Middle East are still unknown.

The aim of this study was to explore the spectrum, frequency, and phenotype of *POLE* and *POLD1* exonuclease domain pathogenic variants in a large CRC cohort from the Middle East. This may increase the knowledge of mechanisms underlying CRC and assist in the implementation of therapeutic and/or preventive strategies in Middle Eastern population.

2 | MATERIALS AND METHODS

2.1 | Editorial policies and ethical considerations

This study was approved by Institutional Review Board. Waiver of consent was provided by Research Advisory Council to obtain archival paraffin tissue blocks under project RAC # 2040 004.

2.2 | Sample selection

One thousand one hundred and thirty-five CRC archival samples of patients diagnosed at KFSHRC between 1990 and 2015 were included in the present study. Clinico-pathological details for the cases were collected from medical records, which are summarized in Table 1.

2.3 | DNA extraction

DNAs were extracted from CRC non-tumor formalin-fixed and paraffin-embedded (FFPE) tissues utilizing Gentra DNA isolation kit (Gentra) according to manufacturer's protocols as elaborated in the previous studies (Abubaker et al., 2008; Khoo et al., 1999).

2.4 | Targeted capture sequencing of germline exonuclease domain variants in *POLE* and *POLD1* genes

Capture sequencing of 376 CRC cases was carried out using a capture panel as previously reported (Siraj et al., 2017). The concentrations were estimated and DNA samples having A260/A280 ratio from 1.8 to 2.0 were used for library

TABLE 1 Clinico-pathological variables for the patient cohort (n = 1,135)

Clinico-pathological parameter	*n* (%)
Age	
Median	55.0
Range(IQR)	46.4–66.1
Gender	
Male	599 (52.8)
Female	536 (47.2)
Histological subtype	
Adenocarcinoma	1,002 (88.3)
Mucinous carcinoma	133 (11.7)
Histological grade	
Well differentiated	105 (9.3)
Moderately differentiated	881 (77.6)
Poorly differentiated	112 (9.9)
Unknown	37 (3.2)
Tumor site	
Left	915 (80.6)
Right	206 (18.1)
Unknown	14 (1.2)
TNM stage	
I	153 (13.5)
II	357 (31.5)
III	452 (39.8)
IV	142 (12.5)
Unknown	31 (2.7)

Abbreviation: IQR, Inter quartile range.

preparation. Random fragmentation of DNA was carried out to prepare the sequencing library, followed by adapter ligation at 5′ and 3′ ends. These fragments were then amplified by PCR and purified after gel electrophoresis. The library was then transferred to a flow cell to generate clusters by attachment to complementary adapters bound to surface of the flow cell. The fragments were amplified into clusters by bridge amplification. The base call data were generated by Illumina's Hiseq control, version 3.3.

The base call data (BCL files) were converted into FASTQ format using Illumina's bcl2fastq (v2.16). The FASTQ reads were mapped to the human reference genome hg19 by burrows wheeler aligner (BWA; Li & Durbin, 2010). The bam files generated by BWA were converted into compressed bam format using Picard-tools. PCR duplicates were marked and realignment was carried out on the bam files to get high quality variants using genome analysis toolkit (GATK) and Picard-tools (McKenna et al., 2010).

The variants were called using GATK. The variants were annotated with different annotation databases like dbSNP, 1000 genomes, Exome Aggregation Consortium (ExAC), Exome Sequencing Project (ESP6500) and ClinVar and an in-house database of around 3,000 healthy exomes using ANNOVAR (Wang, Li, & Hakonarson, 2010).

Our targeted capture panel uses similar sequencing technology as Liang et al., 2016 and Beaubier et al., 2019. These studies have estimated the sensitivity and specificity of mutation detection by targeted capture sequencing to be >98% for all types of alterations with expected MAF of >5% at a sequencing coverage of >2000X. The depth of coverage for our targeted capture sequencing is >1500X and being germline mutations where MAF is close to 50%, the specificity and sensitivity will be more than 98%. (Beaubier et al., 2019; Liang et al., 2016).

2.5 | PCR and sanger sequencing to detect germline exonuclease domain variants in *POLE* and *POLD1* genes

Sequencing of whole coding and splicing regions of *POLE* (GenBank: NG_033840.1) and *POLD1* (GenBank: NG_033800.1) exonuclease domain among 759 samples were carried out using Sanger sequencing technology. Additionally, variants detected by Capture sequencing were subsequently validated by Sanger sequencing. Primer 3 online software was utilized to design the primers (available upon request). PCR and Sanger sequencing were carried out as described previously (Siraj et al., 2019).

2.6 | Assessment of pathogenicity of variants

ACMG/AMP 2015 guideline was utilized first for interpretation of sequence variants (Richards et al., 2015). All the variants of uncertain significance (VUSs) interpreted by ACMG/AMP 2015 were further analyzed using five in silico pathogenicity prediction tools: PolyPhen-2 (Adzhubei et al., 2010), MutationAssessor (Reva, Antipin, & Sander, 2011), SIFT (Vaser, Adusumalli, Leng, Sikic, & Ng, 2016), CADD (Rentzsch, Witten, Cooper, Shendure, & Kircher, 2019) and Mutation Taster (Schwarz, Cooper, Schuelke, & Seelow, 2014). The variants predicted as damaging or possibly damaging by three or more in silico prediction tools were further investigated.

2.7 | Tissue microarray construction & Immunohistochemistry

Tissue microarray (TMA) was utilized for analyses of samples. Construction of TMA was performed as described previously (Siraj et al., 2007). Two cores from each case of CRC were arrayed.

TABLE 2 Characteristics of five damaging or possibly damaging variants identified in our cohort

Gene	Variant	ACMG	PolyPhen-2	Mutation assessor	SIFT	CADD	Mutation taster	Onset age	Family history	Conserved	Frequency in ExAC
POLE[a]	c.1370C>T:p.(Thr457Met)	Uncertain significance	Probably damaging	Low	Tolerated	25.1	Disease causing	52	NA	4/6	ABSENT
POLE	c.940T>G:p.(Ser314Ala)	Uncertain significance	Benign	Medium	Tolerated	22.1	Disease causing	47	Positive	6/6	0.00008
POLE	c.1024C>T:p.(His342Tyr)	Uncertain significance	Benign	Low	Damaging	22.3	Disease causing	47	NA	3/6	Absent
POLE	c.1184G>A:p.(Gly395Glu)	Uncertain significance	Possibly damaging	Low	Tolerated	28	Disease causing	50	NA	5/6	0.00009
POLD1[b]	c.932G>A:p.(Arg311His)	Uncertain significance	Probably damaging	High	Damaging	24.7	Disease causing	90	NA	7/7	Absent

[a]GenBank accession number NG_033840.1
[b]GenBank accession number NG_033800.1

Immunohistochemistry (IHC) staining was performed using a standard protocol as described previously (Siraj et al., 2019). Staining of slides with omission of the primary antibody served as negative controls. Simultaneous staining of fresh cut slides was performed to reduce the effect of slide aging and to optimize the staining.

For evaluation of IHC staining, intensity (I) and percentage (P) of positively stained tumor cells was assessed for each case as follows: intensity score was assigned ranging from 0 to 3 (I0-negative; I1-weak; I2-moderate; I3-strong) and percentage of positively stained tumor cells for each intensity score was assigned ranging from 0%–100% (P0–P3). The product of the scores obtained for each intensity and the corresponding percentage of tumor cells stained were added to obtain the final H score (=I1XP1 + I2XP2 + I3XP3), which ranged from 0 to 300. Using X-tile version 3.6.1 (Camp, Dolled-Filhart, & Rimm, 2004), H score of 70 and 120 was defined as the cutoff for POLE and POLD1, respectively. The cases having score below the cutoff were considered to have low expression and those having score above the cutoff were classified as over expression.

Mismatch repair protein staining and evaluation was done as described previously (Siraj et al., 2015).

2.8 | Statistical analysis

The associations between protein expression and clinicopathological variables were performed using contingency table analysis and Chi-square tests. Kaplan–Meier method was used to generate overall survival curves and Mantel–Cox log-rank test was used to assess significance. Two-sided tests were used for statistical analyses with a limit of significance defined as $p < .05$. The JMP11.0 (SAS Institute, Inc.) software was used for data analyses.

3 | RESULTS

3.1 | Sample characteristics

A total of 1,135 CRC cases were included in the study. The median age of the study cohort was 55 years. The majority of the cases was of adenocarcinoma histologic subtype (88.3%), moderately differentiated (77.6%), and left-sided (80.6%) tumors (Table 1).

3.2 | Identification of germline exonuclease domain variants in *POLE* and *POLD1* genes

Out of 376 CRC cases sequenced utilizing capture sequencing, three variants, two in *POLE* and one in *POLD1*, were

detected and interpreted as of uncertain significance utilizing ACMG/AMP 2015 guideline. Furthermore, five in silico prediction tools were utilized for further analysis. Two variants (c.940T>G: p.Ser314Ala and c.1370C>T: p.Thr457Met) in *POLE* were assessed as damaging or possibly damaging, accounting for 0.5% (2/376) of all samples analyzed by Capture sequencing. To increase our cohort, we included another 759 CRC cases sequenced by Sanger sequencing technology. In this cohort, six variants were interpreted as of uncertain significance by ACMG/AMP 2015 guideline, three in *POLE* and three in *POLD1*. Further analysis using pathogenicity prediction tools revealed that three (0.4%, 3/759) were damaging or possibly damaging, two (c.1024C>T: p.His342Tyr and c.1184G>A: p.Gly395Glu) in *POLE* and one (c.932G>G: p.Arg311His) in *POLD1*. Altogether, five variants (0.44%) were classified by ACMG guidelines as variants of uncertain significance, in silico prediction tools in the 1,135 CRC cases identified four variants in *POLE* (0.35%, 4/1,135) and one variant in *POLD1* (0.1%, 1/1,135; Table 2) to warrant further investigation.

Two heterozygous POLE variants p.(Ser314Ala) and p.(Thr457Met) were detected by capture sequencing. The *POLE* heterozygous variant p.(Ser314Ala) was identified in an early-onset CRC patient (diagnosed at 47 years of age) with large tubulovillous adenoma. The patient carrying this variant has strong family history of CRC and Familial adenomatous polyposis (FAP; Figure 1). Furthermore, this variant is fully conserved and is also identified in population database ExAC at a very low frequency of 0.00008. p.(Thr457Met) in POLE was detected in CRC patient diagnosed at 52 years. This variant was also highly conserved and absent in ExAC. The patients carrying these two variants are not known to carry any germline pathogenic variants in other CRC-predisposing genes. The variants of p.(His342Tyr) and p.(Gly395Glu) in *POLE* were detected in CRC patients diagnosed at 47 and 50 years, respectively. p.(His342Tyr) variant was highly conserved and absent in ExAC whereas p.(Gly395Glu) was observed at a low frequency of 0.00009 in ExAC (Table 2). IHC expression of MMR genes in the cases having these two variants was normal therefore further sequencing analysis was not performed.

The *POLD1* variant of p.(Arg311His) was identified by Sanger sequencing in a patient with late onset of CRC (diagnosed at 90 years of age). However, this variant is completely conserved among species and is predicted as damaging or possibly damaging variant by all five prediction tools. Additionally, the p.(Arg311His) variant is absent in ExAC (Table 2). This case was also enrolled in another ongoing study in which Sanger sequencing was employed to screen the germline and somatic MMR genes variants. However, the patient does not carry any germline or somatic pathogenic variants in MMR genes (data not shown). Interestingly, the tumor from this patient is deficient MMR (dMMR) as assessed by IHC.

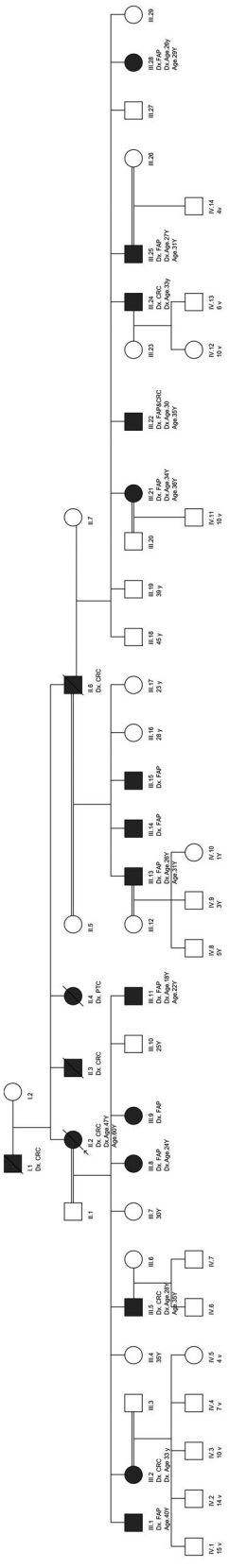

FIGURE 1 Pedigree of the patient having *POLE* c.940T>G: p.(Ser314Ala) variant. The indexed patient is indicated by arrow. Circles indicate female; square, male; slash forward, individual deceased; dark square or circle, affected by cancer; CRC indicates colorectal cancer; FAP indicates Familial adenomatous polyposis; PTC indicates Papillary Thyroid cancer; Dx. Age indicates Age at diagnosis

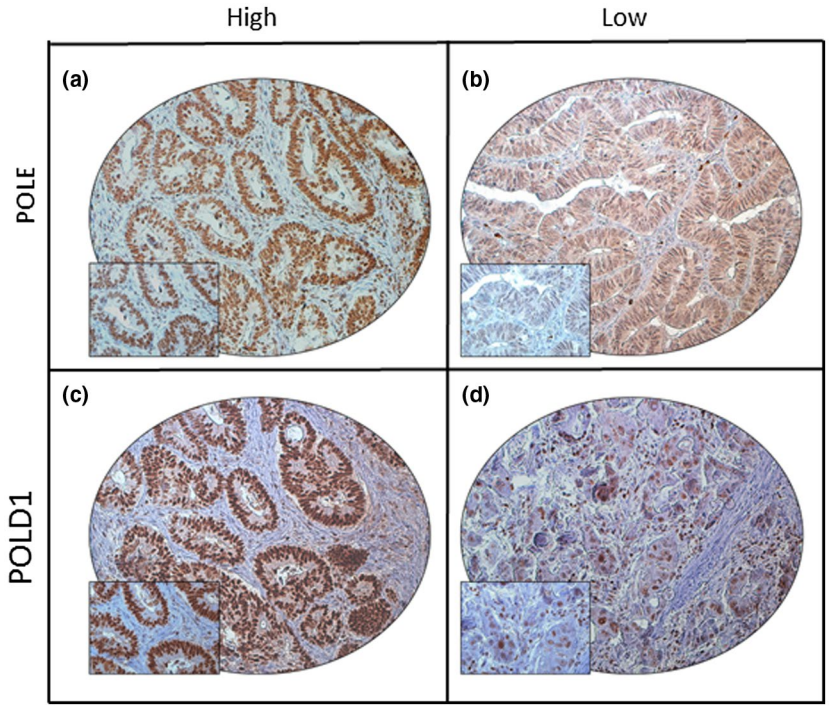

FIGURE 2 Tissue microarray based immunohistochemistry analysis of *POLE* and *POLD1* in CRC patients. CRC TMA cores showing overexpression of POLE (a) and POLD1 (c). In contrast, another set of TMA cores showing reduced expression of POLE (B) and POLD1 (e). 20 X/0.70 objective on an Olympus BX 51 microscope. (Olympus America Inc.) with the inset showing a 40 × 0.85 aperture magnified view of the same TMA spot

We further sought to explore the possible associations between *POLE*/*POLD1* damaging or possibly damaging variants and clinico-pathological parameters. However, no significant association was found between *POLE*/*POLD1* disease–causing variants and clinico-pathological characteristics such as age, gender, grade, stage, histologic subtype, tumor site, and MSI status.

3.3 | Protein expression of POLE and POLD1 genes in CRC and their clinico-pathological associations

We further examined the expression of POLE and POLD1 proteins by IHC in 1,135 CRC cases in TMA format. POLE expression was interpretable in 1,071 cases. Low expression of POLE was detected in 38.9% (417/1071) of cases (Figure 2a,b) and revealed a significant association with lymph node involvement ($p = .0184$) and grade 3 tumors ($p = .0139$; Table 3). POLD1 expression was interpretable in 1,069 cases. Low expression of POLD1 was observed in 51.9% (555/1069) cases (Figure 2c,d) and showed significant association with adenocarcinoma histology ($p = .0164$), larger tumor size (T3 and T4 tumors; $p = .0012$), and stage III tumors ($p = .0341$; Table 4). However, there was no association between protein expression and these five *POLE*/*POLD1* exonuclease domain variants.

4 | DISCUSSION

Although a wealth of knowledge has been acquired, the picture of germline genetic alterations causing CRC is still incomplete. In this study, we examined the germline exonuclease domain variants of *POLE* and *POLD1* genes to identify predisposing variants in a large CRC cohort from the Middle East. We detected four heterozygous damaging or possibly damaging variants in *POLE* and one heterozygous damaging or possibly damaging variant in *POLD1*, accounting for 0.4% of all CRC patients. The frequency of damaging or possibly damaging variants in Middle Eastern CRC was similar to that in Western populations (Chubb et al., 2015; Elsayed et al., 2015; Palles et al., 2013; Valle et al., 2014). However, there is insufficient data to assess the functional impact of these genetic variants on the protein; therefore, these variants are classified as a Variant of Uncertain Significance by ACMG. In our cohort, we did not identify any damaging or possibly damaging variants which were frequently detected previously in Western populations, indicating the unique spectrum of germline exonuclease domain variants of *POLE* and *POLD1* in our population.

Our data provide some evidence that POLE p.Ser314Ala is associated with increased risk for CRC. The patient carrying the variant of *POLE* p.(Ser314Ala) had large adenoma and early onset phenotype with strong family history of CRC and FAP, similar to that of previously reported pathogenic *POLE* variant p.(Leu424Val; Palles et al., 2013) and resembling Lynch syndrome. Furthermore, absence of other germline pathogenic variants in CRC predisposing genes including *APC*, *MUTYH,* and MMR genes indicates that this variant might be the causative variant of CRC. Moreover this variant is fully conserved between six species and observed in population database at a much lower frequency. These data might indicate the pathogenic effect of this variant. Interestingly one sister of the patient carrying this variant developed papillary thyroid carcinoma (PTC), therefore further

TABLE 3 Correlation of POLE IHC expression with clinico-pathological parameters in colorectal carcinoma

	Total		Low		High		*p* value
	N	%	*N*	%	*N*	%	
Total number of cases	1,071		417	38.9	654	61.1	
Age							
≤50 years	367	34.4	138	37.6	229	62.4	.5421
>50 years	701	65.6	277	39.5	424	60.5	
Sex							
Male	564	52.7	221	39.2	343	60.8	.8602
Female	507	47.3	196	38.7	311	61.3	
Tumor site							
Left colon	857	80.9	329	38.4	528	61.6	.4037
Right colon	202	19.1	84	41.6	118	58.4	
histological type							
Adenocarcinoma	948	88.8	363	38.3	585	61.7	.2883
Mucinous carcinoma	120	11.2	52	43.3	68	56.7	
pT							
T1/T2	198	19.2	71	35.9	127	64.1	.3495
T3/T4	834	80.8	329	39.5	505	60.5	
pN							
N0	508	49.7	178	35.0	330	65.0	.0184
N1/N2	514	50.3	217	42.2	297	57.8	
pM							
M0	916	87.7	357	39.0	559	61.0	.8802
M1	128	12.3	49	38.3	79	61.7	
Tumor stage							
I	143	13.8	50	35.0	93	65.0	.2041
II	345	33.1	127	36.8	218	63.2	
III	425	40.8	183	43.1	242	56.9	
IV	128	12.3	49	38.3	79	61.7	
Differentiation							
Well	99	9.5	34	34.3	65	65.7	.0139
Moderate	842	80.4	319	37.9	523	62.1	
Poor	106	10.1	55	51.9	51	48.1	
MSH2 IHC							
Positive	942	90.1	361	38.3	581	61.7	.2916
Negative	103	9.9	45	43.7	58	56.3	
POLE mutation							
Yes	3	0.3	2	66.7	1	33.3	.3315
No	1,068	99.7	415	38.9	653	61.1	
Survival							
OS 5 years				71.9		72.8	.5934

analysis is required to investigate whether *POLE* p.(Ser-314Ala) variant is driver of PTC. The other three *POLE* variants p.(His342Tyr), p.(Gly395Glu), and p.(Thr457Met) were also found in early onset CRC patients and absent or observed at low frequency in ExAC. All the variants were partially conserved and predicted as pathogenic by three in silico prediction tools. Interestingly, the variant p.(His342Tyr) was previously reported to be present in a hyper-mutant

TABLE 4 Correlation of POLD1 IHC expression with clinico-pathological parameters in colorectal carcinoma

	Total		Low		High		p value
	N	%	*N*	%	*N*	%	
Total number of cases	1,069		555	51.9	514	48.1	
Age							
≤50 years	366	34.3	187	51.1	179	48.9	.6786
>50 years	700	65.7	367	52.4	333	47.6	
Sex							
Male	562	52.6	290	51.6	272	48.4	.8275
Female	507	47.4	265	52.3	242	47.7	
Tumor site							
Left colon	856	81.0	453	52.9	403	47.1	.1162
Right colon	201	19.0	94	46.8	107	53.2	
Histological type							
Adenocarcinoma	946	88.7	504	53.3	442	46.7	.0164
Mucinous carcinoma	120	11.3	50	41.7	70	58.3	
pT							
T1/T2	198	19.2	82	41.4	116	58.6	.0012
T3/T4	832	80.8	451	54.2	381	45.8	
pN							
N0	507	49.7	246	48.5	261	51.5	.0802
N1/N2	513	50.3	277	54.0	236	46.0	
pM							
M0	914	87.7	475	52.0	439	48.0	.6763
M1	128	12.3	64	50.0	64	50.0	
Tumor stage							
I	143	13.8	59	41.3	84	58.7	.0341
II	344	33.1	182	52.9	162	47.1	
III	424	40.8	234	55.2	190	44.8	
IV	128	12.3	64	50.0	64	50.0	
Differentiation							
Well	99	9.5	57	57.6	42	42.4	.2026
Moderate	840	80.4	440	52.4	400	47.6	
Poor	106	10.1	48	45.3	58	54.7	
MSH2 IHC							
Positive	941	90.2	487	51.7	454	48.3	.8196
Negative	102	9.8	54	52.9	48	47.1	
POLD1 mutation							
Yes	1	0.1	1	100.0	0	0.0	.2521
No	1,068	99.9	554	51.9	514	48.1	
Survival							
OS 5 years				71.9		73.0	.3640

tumor (Campbell et al., 2017). This hyper-mutation phenotype indicates that p.(His342Tyr) might have strong mutagenic effect. However, due to conservative culture in the Middle East, family history information from the carriers of p.(His342Tyr) and p.(Thr457Met) variants were not available (Bou Khalil, 2013; Eldeek et al., 2014).

Although the *POLD1* variant of c.932G>A: p.(Arg311His) was found in a late onset patient, the amino acid Arg is fully conserved between all seven species and is predicted as damaging or possibly damaging by all prediction tools. Furthermore, this variant is absent in ExAC. All these data suggested that this variant may cause loss of function of *POLD1*. In addition, the tumor from this patient is dMMR by IHC assessment. Additional, germline and somatic pathogenic variants of MMR genes by Sanger sequencing were not detected in this patient (data not shown). This result is consistent with previous finding that MSI–H tumor might possess exonuclease domain pathogenic variants in *POLD1* (Haradhvala et al., 2018).

In this study, we did not observe association between germline damaging or possibly damaging variants of *POLE* and *POLD1* genes and their protein expression. This result might be partially explained by the fact that *POLE* and *POLD1* somatic exonuclease domain variants were not investigated in current study. Interestingly, a previous study reported that compared to the exonuclease domain region more truncated variants were detected in the region outside of this domain, and more than 30% of *POLE* and *POLD1* truncated variants did not result in high tumor mutation burden (TMB) in the tumors (Campbell et al., 2017). In addition, another study also reported that POLE IHC has no predictive value for effect of mutation (Elsayed et al., 2015). All these results implied that IHC analysis of these two genes might not be suitable to select the patients for immune checkpoint inhibitor therapy.

To the best of our knowledge, we are the first to report the frequency and clinico-pathological associations of protein expression in POLE and POLD1 genes in CRCs. POLE and POLD1 low expression was associated with adverse prognostic factors. However, no significant survival outcomes were noted with both biomarkers. Further studies are warranted to validate our findings.

In conclusion, we showed a low frequency of germline damaging or possibly damaging variants in exonuclease domain of *POLE* and *POLD1* in CRC patients from the Middle East. We recommend that screening of *POLE* and *POLD1* exonuclease region should be applied for identification of individuals that are at an increased risk of developing CRC and patients who might be suitable for immunotherapy. Compared to single gene testing next generation sequencing technology provides critical benefits by reduction in cost and time therefore, it is feasible to screen multiple cancer predisposing genes in CRC patients using multi-gene panels including *POLE* and *POLD1*.

ACKNOWLEDGMENTS

The study is not funded by any local or foreign grant.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

AUTHOR CONTRIBUTION

Abdul Khalid Siraj contributed to formal analysis (lead); methodology (lead); writing—original draft (equal); conceptualization (supporting). Rong Bu contributed to methodology (equal); formal analysis (equal); writing—review and editing (equal). Kaleem Iqbal contributed to writing—review and editing (supporting); formal analysis (supporting), data curation (equal). Sandeep Kumar Parvathareddy contributed to methodology (equal); writing—review and editing (equal); data curation (equal). Tariq Masoodi contributed to software (lead); formal analysis (equal). Nabil Siraj, Maha Al-Rasheed, and Yan Kong contributed to data curation (equal); investigation (supporting). Saeeda O. Ahmed, Khadija A.S. Al-Obaisi, Ingrid G. Victoria, and Maham Arshad contributed to data curation (supporting), investigation (supporting). Fouad Al-Dayel, Alaa Abduljabbar, and Luai H. Ashari contributed to resources (equal); data curation (supporting). Khawla S. Al-Kuraya contributed to supervision (lead); conceptualization (lead); methodology (lead); writing—original draft (lead).

DATA AVAILABILITY STATEMENT

The raw data of the findings are available from the corresponding author upon request.

ORCID

Khawla S. Al-Kuraya https://orcid.org/0000-0002-4126-3419

REFERENCES

Abubaker, J., Jehan, Z., Bavi, P., Sultana, M., Al-Harbi, S., Ibrahim, M., … Al-Kuraya, K. S. (2008). Clinicopathological analysis of papillary thyroid cancer with PIK3CA alterations in a Middle Eastern population. *Journal of Clinical Endocrinology and Metabolism*, *93*(2), 611–618. https://doi.org/10.1210/jc.2007-1717

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., … Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nature Methods*, *7*(4), 248–249. https://doi.org/10.1038/nmeth0410-248

Al-Eid, H. S. (1994). Colo-rectal cancer (c 18-c 20). in, bazarbashi S. *Saudi Cancer Registry*, *1996*, 28–29.

Alrawaji, A., Alshahrani, Z., Alzahrani, W., & Alomran, F., & Almadouj, A. (2018). Registry SC. *Cancer incidence report Saudi Arabia 2014* (p. 60). Riyadh: Saudi Cancer Registry, Ministry of Health.

Al-Tassan, N., Chmiel, N. H., Maynard, J., Fleming, N., Livingston, A. L., Williams, G. T., … Cheadle, J. P. (2002). Inherited variants of MYH associated with somatic G:C–>T: A mutations in colorectal tumors. *Nature Genetics*, *30*(2), 227–232. https://doi.org/10.1038/ng828

Beaubier, N., Tell, R., Lau, D., Parsons, J. R., Bush, S., Perera, J., … White, K. P. (2019). Clinical validation of the tempus xT next-generation targeted oncology sequencing assay. *Oncotarget*, *10*(24), 2384–2396. https://doi.org/10.18632/oncotarget.26797

Bellido, F., Pineda, M., Aiza, G., Valdés-Mas, R., Navarro, M., Puente, D. A., … Valle, L. (2016). 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. *Genetics in Medicine*, *18*(4), 325–332. https://doi.org/10.1038/gim.2015.75

Bou Khalil, R. (2013). Attitudes, beliefs and perceptions regarding truth disclosure of cancer-related information in the Middle East: A review. *Palliative and Supportive Care*, *11*(1), 69–78. https://doi.org/10.1017/S1478951512000107

Bourdais, R., Rousseau, B., Pujals, A., Boussion, H., Joly, C., Guillemin, A., … Tournigand, C. (2017). Polymerase proofreading domain mutations: New opportunities for immunotherapy in hypermutated colorectal cancer beyond MMR deficiency. *Critical Reviews in Oncology Hematology*, *113*, 242–248. https://doi.org/10.1016/j.critrevonc.2017.03.027

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, *68*(6), 394–424. https://doi.org/10.3322/caac.21492

Buchanan, D. D., Stewart, J. R., Clendenning, M., Rosty, C., Mahmood, K., Pope, B. J., … Win, A. K. (2018). Correction: Risk of colorectal cancer for carriers of a germ-line mutation in POLE or POLD1. *Genetics in Medicine*, *20*(10), 1299. https://doi.org/10.1038/gim.2017.265

Calva-Cerqueira, D., Dahdaleh, F. S., Woodfield, G., Chinnathambi, S., Nagy, P. L., Larsen-Haidle, J., … Howe, J. R. (2010). Discovery of the BMPR1A promoter and germline mutations that cause juvenile polyposis. *Human Molecular Genetics*, *19*(23), 4654–4662. https://doi.org/10.1093/hmg/ddq396

Camp, R. L., Dolled-Filhart, M., & Rimm, D. L. (2004). X-tile: A new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clinical Cancer Research*, *10*(21), 7252–7259. https://doi.org/10.1158/1078-0432.CCR-04-0713

Campbell, B. B., Light, N., Fabrizio, D., Zatzman, M., Fuligni, F., de Borja, R., … Shlien, A. (2017). Comprehensive analysis of hypermutation in human cancer. *Cell*, *171*(5), 1042–1056.e10. https://doi.org/10.1016/j.cell.2017.09.048

Chubb, D., Broderick, P., Frampton, M., Kinnersley, B., Sherborne, A., Penegar, S., … Houlston, R. S. (2015). Genetic diagnosis of high-penetrance susceptibility for colorectal cancer (CRC) is achievable for a high proportion of familial CRC by exome sequencing. *Journal of Clinical Oncology*, *33*(5), 426–432. https://doi.org/10.1200/JCO.2014.56.5689

Ciardiello, D., Vitiello, P. P., Cardone, C., Martini, G., Troiani, T., Martinelli, E., & Ciardiello, F. (2019). Immunotherapy of colorectal cancer: Challenges for therapeutic efficacy. *Cancer Treatment Reviews*, *76*, 22–32. https://doi.org/10.1016/j.ctrv.2019.04.003

Dong, H., Bai, Y., Cao, X., Wang, Y., Shi, L., Li, F., Sun, S. (2019). Comprehensive analysis of POLE and POLD1 mutation in 9322 Chinese cancer patients. *Cancer Research*, *79*(13), 5137–5137. https://doi.org/10.1158/1538-7445.AM2019-5137

Eldeek, B., Alahmadi, J., Al-Attas, M., Sait, K., Anfinan, N., Aljahdali, E., … Sait, H. (2014). Knowledge, perception, and attitudes about cancer and its treatment among healthy relatives of cancer patients: Single Institution hospital-based study in Saudi Arabia. *Journal of Cancer Education*, *29*(4), 772–780. https://doi.org/10.1007/s13187-014-0653-7

Elsayed, F. A., Kets, C. M., Ruano, D., van den Akker, B., Mensenkamp, A. R., Schrumpf, M., … van Wezel, T. (2015). Germline variants in POLE are associated with early onset mismatch repair deficient colorectal cancer. *European Journal of Human Genetics*, *23*(8), 1080–1084. https://doi.org/10.1038/ejhg.2014.242

Gala, M. K., Mizukami, Y., Le, L. P., Moriichi, K., Austin, T., Yamamoto, M., … Chung, D. C. (2014). Germline mutations in oncogene-induced senescence pathways are associated with multiple sessile serrated adenomas. *Gastroenterology*, *146*(2), 520–529. https://doi.org/10.1053/j.gastro.2013.10.045

Gong, J., Wang, C., Lee, P. P., Chu, P., & Fakih, M. (2017). Response to PD-1 blockade in microsatellite stable metastatic colorectal cancer harboring a POLE mutation. *Journal of the National Comprehensive Cancer Network*, *15*(2), 142–147.

Grady, W. M. (2003). Genetic testing for high-risk colon cancer patients. *Gastroenterology*, *124*(6), 1574–1594. https://doi.org/10.1016/s0016-5085(03)00376-7

Haradhvala, N. J., Kim, J., Maruvka, Y. E., Polak, P., Rosebrock, D., Livitz, D., … Getz, G. (2018). Distinct mutational signatures characterize concurrent loss of polymerase proofreading and mismatch repair. *Nature Communications*, *9*(1), 1746. https://doi.org/10.1038/s41467-018-04002-4

Houlston, R., Bevan, S., Williams, A., Young, J., Dunlop, M., Rozen, P., … Tomlinson, I. (1998). Mutations in DPC4 (SMAD4) cause juvenile polyposis syndrome, but only account for a minority of cases. *Human Molecular Genetics*, *7*(12), 1907–1912. https://doi.org/10.1093/hmg/7.12.1907

Howlader, N., Noone, A., Krapcho, M., Miller, D., Bishop, K., Altekruse, S., … Tatalovich, Z. (2016). *SEER cancer statistics review, 1975–2013*. Bethesda, MD: National Cancer Institute. https://seer.cancer.gov/archive/csr/1975_2013/.

Jasperson, K. W., Tuohy, T. M., Neklason, D. W., & Burt, R. W. (2010). Hereditary and familial colon cancer. *Gastroenterology*, *138*(6), 2044–2058. https://doi.org/10.1053/j.gastro.2010.01.054

Khoo, U.-S., Ozcelik, H., Cheung, A. N. Y., Chow, L. W. C., Ngan, H. Y. S., Done, S. J., … Ho, F. C. S. (1999). Somatic mutations in the BRCA1 gene in Chinese sporadic breast and ovarian cancer. *Oncogene*, *18*(32), 4643–4646. https://doi.org/10.1038/sj.onc.1202847

Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, *26*(5), 589–595. https://doi.org/10.1093/bioinformatics/btp698

Liang, J., She, Y., Zhu, J., Wei, L., Zhang, L., Gao, L., … Li, P. (2016). Development and validation of an ultra-high sensitive next-generation sequencing assay for molecular diagnosis of clinical oncology. *International Journal of Oncology*, *49*(5), 2088–2104. https://doi.org/10.3892/ijo.2016.3707

Lichtenstein, P., Holm, N. V., Verkasalo, P. K., Iliadou, A., Kaprio, J., Koskenvuo, M., … Hemminki, K. (2000). Environmental and heritable factors in the causation of cancer–analyses of cohorts of twins from Sweden, Denmark, and Finland. *New England Journal of Medicine*, *343*(2), 78–85. https://doi.org/10.1056/NEJM200007133430201

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., … DePristo, M. A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303. https://doi.org/10.1101/gr.107524.110

Nick McElhinny, S. A., Gordenin, D. A., Stith, C. M., Burgers, P. M., & Kunkel, T. A. (2008). Division of labor at the eukaryotic replication fork. *Molecular Cell*, *30*(2), 137–144. https://doi.org/10.1016/j.molcel.2008.02.022

Palles, C., Cazier, J.-B., Howarth, K. M., Domingo, E., Jones, A. M., Broderick, P., … Tomlinson, I. (2013). Germline mutations affecting

the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. *Nature Genetics*, *45*(2), 136–144. https://doi.org/10.1038/ng.2503

Pursell, Z. F., Isoz, I., Lundstrom, E. B., Johansson, E., & Kunkel, T. A. (2007). Yeast DNA polymerase epsilon participates in leading-strand DNA replication. *Science*, *317*(5834), 127–130. https://doi.org/10.1126/science.1144067

Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J., & Kircher, M. (2019). CADD: Predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Research*, *47*(D1), D886–D894. https://doi.org/10.1093/nar/gky1016

Reva, B., Antipin, Y., & Sander, C. (2011). Predicting the functional impact of protein mutations: Application to cancer genomics. *Nucleic Acids Research*, *39*(17), e118. https://doi.org/10.1093/nar/gkr407

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., … Rehm, H. L. (2015). Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine*, *17*(5), 405–423. https://doi.org/10.1038/gim.2015.30

Schwarz, J. M., Cooper, D. N., Schuelke, M., & Seelow, D. (2014). MutationTaster2: Mutation prediction for the deep-sequencing age. *Nature Methods*, *11*(4), 361–362. https://doi.org/10.1038/nmeth.2890

Siraj, A. K., Bavi, P., Abubaker, J., Jehan, Z., Sultana, M., Al-Dayel, F., … Al-Kuraya, K. S. (2007). Genome-wide expression analysis of Middle Eastern papillary thyroid cancer reveals c-MET as a novel target for cancer therapy. *The Journal of Pathology*, *213*(2), 190–199. https://doi.org/10.1002/path.2215

Siraj, A. K., Masoodi, T., Bu, R., Parvathareddy, S. K., Al-Badawi, I. A., Al-Sanea, N., … Al-Kuraya, K. S. (2017). Expanding the spectrum of germline variants in cancer. *Human Genetics*, *136*(11–12), 1431–1444. https://doi.org/10.1007/s00439-017-1845-0

Siraj, A. K., Parvathareddy, S. K., Bu, R., Iqbal, K., Siraj, S., Masoodi, T., … Al-Kuraya, K. S. (2019). Germline POLE and POLD1 proofreading domain mutations in endometrial carcinoma from Middle Eastern region. *Cancer Cell International*, *19*(1), 334. https://doi.org/10.1186/s12935-019-1058-9

Siraj, A. K., Prabhakaran, S., Bavi, P., Bu, R., Beg, S., Hazmi, M. A., … Al-Kuraya, K. S. (2015). Prevalence of Lynch syndrome in a Middle Eastern population with colorectal cancer. *Cancer*, *121*(11), 1762–1771. https://doi.org/10.1002/cncr.29288

Spier, I., Holzapfel, S., Altmuller, J., Zhao, B., Horpaopan, S., Vogt, S., … Aretz, S. (2015). Frequency and phenotypic spectrum of germline mutations in POLE and seven other polymerase genes in 266 patients with colorectal adenomas and carcinomas. *International Journal of Cancer*, *137*(2), 320–331. https://doi.org/10.1002/ijc.29396

Valle, L., Hernández-Illán, E., Bellido, F., Aiza, G., Castillejo, A., Castillejo, M.-I., … Blanco, I. (2014). New insights into POLE and POLD1 germline mutations in familial colorectal cancer and polyposis. *Human Molecular Genetics*, *23*(13), 3506–3512. https://doi.org/10.1093/hmg/ddu058

Vaser, R., Adusumalli, S., Leng, S. N., Sikic, M., & Ng, P. C. (2016). SIFT missense predictions for genomes. *Nature Protocols*, *11*(1), 1–9. https://doi.org/10.1038/nprot.2015.123

Wang, K., Li, M., & Hakonarson, H. (2010). ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research*, *38*(16), e164. https://doi.org/10.1093/nar/gkq603

Weren, R. D. A., Ligtenberg, M. J. L., Kets, C. M., de Voer, R. M., Verwiel, E. T. P., Spruijt, L., … Hoogerbrugge, N. (2015). A germline homozygous mutation in the base-excision repair gene NTHL1 causes adenomatous polyposis and colorectal cancer. *Nature Genetics*, *47*(6), 668–671. https://doi.org/10.1038/ng.3287

How to cite this article: Siraj AK, Bu R, Iqbal K, et al. POLE and POLD1 germline exonuclease domain pathogenic variants, a rare event in colorectal cancer from the Middle East. *Mol Genet Genomic Med*. 2020;8:e1368. https://doi.org/10.1002/mgg3.1368