

Supplementary data.

Contains:

Figs S1 and S2
Tables S1-S4

Supplementary Figure legends

Figure S1. Chromatograms showing *POLE* and *POLD1* EDMs reported in this study. The upper reads in each panel are the reference sequence generated by the Mutation Surveyor program.

Figure S2. Sequence alignment of human *POLE* exonuclease domain (residues 268-471) and orthologues with exo motifs highlighted and invariant residues indicated by asterisks. Mutated residues detected in this study and the TCGA dataset are indicated.

Figure S1.

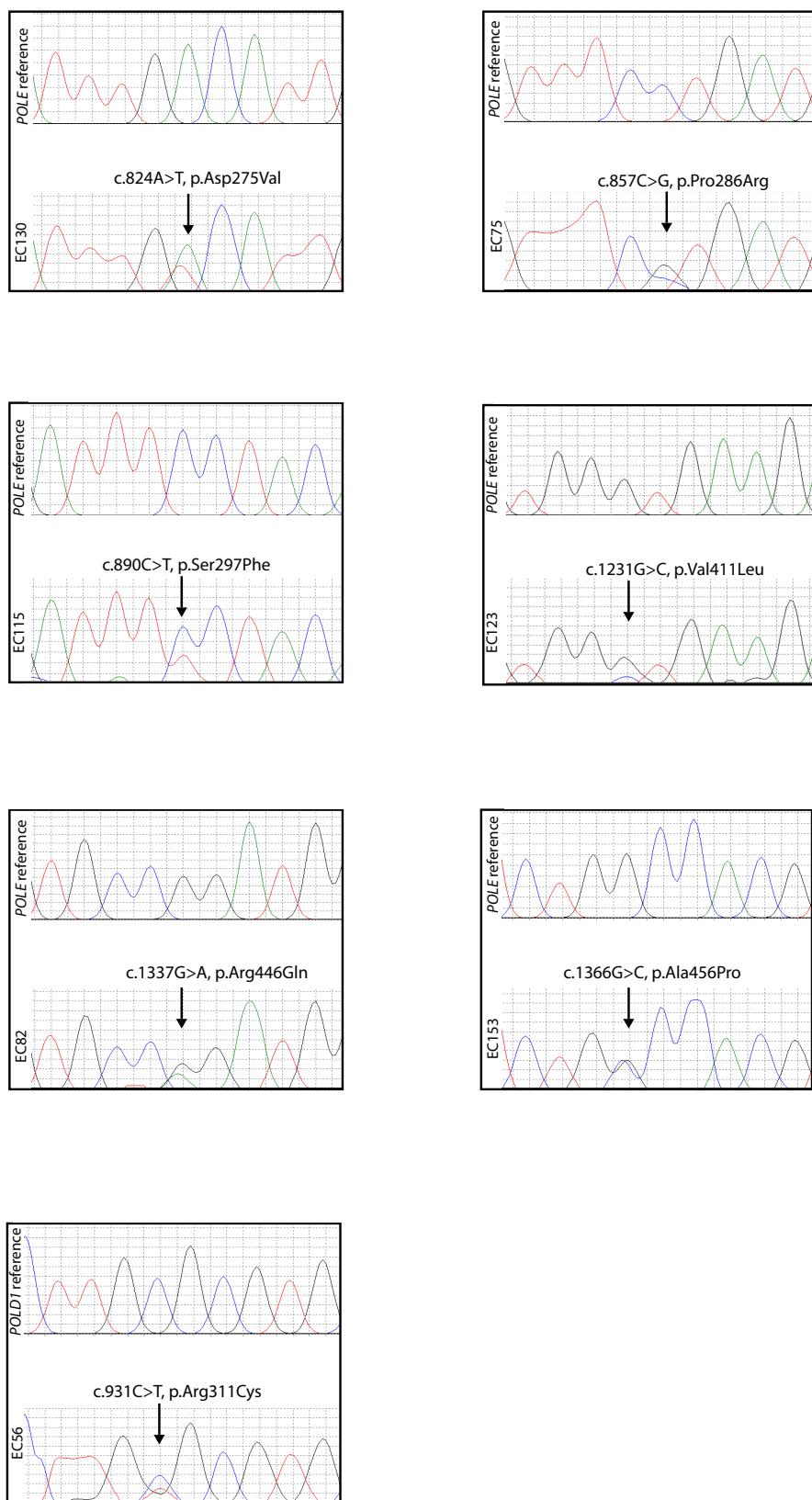


Figure S2

Human	195	LQRGGVITDEEE-TSKKIADQLDNI	VDMREYDVPYHIRLSIDLKIHVAHWYNVRYR-GNAFPVEITRRDDLVERPDPVVL	272				
Mus musculus	195	LQGGSVITDEDE-TSKKIADQLDNI	VDMREYDVPYHIRLSIDLKIHVAHWYNVRFR-GNAFPVEITRRDDLVERPDPVVL	272				
Danio rerio	193	LVGGSVVTEDEGSSKKMTEQLDNL	ILDREYDVPYHVRVSIDLKHVAHWYNVRYR-GSAYPVEIVRRDDLVERPDPVVL	271				
Xenopus	193	LTGNNNMAGEEG-PSKKISDQMENI	VDMREYDVPYHVRVSIDLKHVAHWYNIRYR-GSSSPEITRRDDLVERPDPVVL	270				
Drosophila	193	LAQSSAGSEDAT-LGKRQQDYMDCIVD	IREDHDPVYHVRVSIDLFCGQWYNRCR-SGVELPTTCRDPDILDRPEPVVL	270				
C.Elegans	182	LSKGKGDSDKDQO---	-LNGLNQIVDIREYDVPFHMRSIDEKIFVGLWVDVKGI-GPNRVTIKRKDLPLFHAKPKVL	256				
S. pombe	222	YINVAANGSEKV-----DAKHL	IEDIREYDVPYHVRVSIDKDIRVGKWKVTTQQ-GFIEDT---RK---IAFADPVVM	287				
S.Cerevisiae	206	LNGIIENAFE-----DPLNHVLDIREYDVPYHRSRTLIDL	NIRVGQWYTVSYHeGHQOISLLASR---IERAEP	273				
		* * * * *	* * * * *	*				
		D275	P286	S297				
		▼	▼	▼				
Human	273	AFDIETTKLPKF DAETDQIMMISY	MIDQGQYLITNREIVSEDIEDFEFTPKPEYE	GPGFCVNEPDEAHЛИQRWFЕHVQ	352			
Mus musculus	273	AFDIETTKLPKF DAETDQIMMISY	MIDQGQYLITNREIVSEDIEDFEFTPKPEYE	GPGFCVNEPDEAHЛИQRWFЕHQ	352			
Danio rerio	272	AFDIETTKLPKF DAETDQIMMISY	MIDQGQYLITNREIVSEDIEDFEFTPKPEYE	GPGFTI	351			
Xenopus	271	AFDIETTKLPKF DAETDQIMMISY	MIDQGQYLITNREIVSEDIEDFEFTPKPEYE	GPGFCI	350			
Drosophila	271	AFDIETTKLPKF DAQTDQVMMISY	MIDQGQYLITNREI	ISSNVDDPEYTPKPEFEGNFIVVNENEMOLLQRFFDHIM	350			
C.Elegans	257	AFDIETTKLPKF PDASDQIMMISY	MIDQGQYLITNREI	REVSADINAPEYTPKPEYEGPFTI	336			
S. pombe	288	AFDIETTKLPKF PDASDQIMMISY	MIDQGQYLITNREI	EDFEYTPKPEYEGPFTI	367			
S.Cerevisiae	274	AFDIETTKLPKF PDSSFDKIMMISY	MIDQGQYLITNREI	ISQNIEDFHYPREEFEGPFII	353			
		*****	*****	*****	*****			
		Exo I						
		D411	L424	A428				
Human	353	ETKPTIMVTYNGDFDWPV EARAAVHGLSMQEIGFKQDSQGEYKAPQCIC	HMDCLRWVKRDSYLPVGSHNLKAAAKAKL	432				
Mus musculus	353	ETKPTIMVTYNGDFDWPV EARAAIHLGMSMQEIGFKQDSQGEYKAPQCIC	HMDCLRWVKRDSYLPVGSHNLKAAAKAKL	432				
Danio rerio	352	ETKPNIFVITYNGDFDWPV EARAAQLGLSMRHEIGFKDNOGEYKASQAI	HMDCLRWVKRDSYLPVGSHNLKAAAKAKL	431				
Xenopus	351	ETKPNIIVTYNGDFDWPV ETRAAVHGMMSLQEIGFKQDKNOGEYKSPPC	IHMDCCLRWVKRDSYLPVGSHNLKAAAKAKL	430				
Drosophila	351	EVRPHIIVTYNGDFDWPV	FETRAAVYDLMQEIIGFSKLRDQNYLSRPATI	HMDCLCWVKRDSYLPVGSGQGLKAVAKAKL	430			
C.Elegans	337	QVRPNIVVYTYNGDFDWPV	EARAKIRGFNMREIGFSKDSADEYK5RNCHMDAFRWVKRDSYLPVGSGQGLKAVITKAKL	416				
S. pombe	368	DVRPTIVSTFNGDFDWPV	FIHNRSKIHLGDMFDEIGFAPDAE	GEYKSSYCSHMDCFRWVKRDSYLPQGSQGLKAVTOSKL	447			
S.Cerevisiae	354	SAKPSIVITYNGDFDWPV DARAFAHGLNLTEETGFRDAE	YKSSYCSHMDAFRWVKRDSYLPQGSQGLKAVTVSKL	433				
		*****	*****	*****	*****			
		Exo I						
		M444 R446 Q452 A456 A465						
		Exo IV						
Human	433	GYDPVELDPE DMCRMATEQ	PQTLATYSVSDAVATYYLYMKVYHPPFIFALCTIIPMEPDEVLRKGSGTLCE	ALLMVQAFHA	512			
Mus musculus	433	GYDPVELDPE DMCRMATEQ	PQTLATYSVSDAVATYYLYMKVYHPPFIFALCTIIPMEPDEVLRKGSGTLCE	ALLMVQAFHA	512			
Danio rerio	432	GYDPVELDPE EMCRMATEEP	PQTLATYSVSDAVATYYLYMKVYHPPFIFALCTIIPMEPDEVLRKGSGTLCE	ALLMVQAYHV	511			
Xenopus	431	GYDPVELDPE EMCRMATEEP	PQTLATYSVSDAVATYYLYMKVYHPPFIFALCTIIPMEPDEVLRKGSGTLCE	ALLMVQAYHA	510			
Drosophila	431	RYDPVELDPE DMCRMATEQ	PQVLANYSVSDAVATYYLYMKVYHPPFIFALNTIIPMEPDEVLRKGSGTLCE	TLLMVEAYHA	510			
C.Elegans	417	RYDPVELDPE ELMCKMAREQ	PQQLANYSVSDAVATYYLYMKVYHQPIFALCTIIPLGAD	DLVRKGSGTLCE	ALLMVRAEHN	496		
S. pombe	448	GYNPIELDPE MLTPYAFEK	PQHLSEYSVSDAVATYYLYMKVYHPPFIFSLCTIIPLN	PLNPDET	LRKGSGTLCE	CEMLLMVQAYH	527	
S.Cerevisiae	434	GYNPIELDPE MLTPYASEKP	QVLAQYSVSDAVATYYLYMKVYHPPFIFSLCN	IIPLNPDE	VLRKGTGTLC	CE	TLLLTVEACTK	513
		*****	*****	*****	*****			
		Exo V						
		Exo III						

Table S1. Mutations in EC driver genes in non-EDM and *POLE*-EDM cancers

Tumour ID	MSI	EDM status	Somatic mutation	Protein alteration
EC8	MSS	Non-mutant	<i>PTEN</i> c.518G>A	PTEN p.Arg173His
EC25	MSS	Non-mutant	<i>PIK3CA</i> c.3127A>G <i>PIK3CA</i> c.263G>A	PK3CA p.Met1043Val PK3CA p.Arg88Gln
EC34	MSS	Non-mutant	<i>FBXW7</i> c.1385C>T¶	FBXW7 p.Ser462Phe¶
EC145	MSS	Non-mutant	<i>PIK3CA</i> c.278G>A <i>PTEN</i> c.389G>A	PK3CA p.Arg93Gln PTEN p.Arg130Gln
EC71	MSS	<i>POLE</i> Pro286Arg	<i>FBXW7</i> c.751G>T <i>PTEN</i> c.766G>T	FBXW7 p.Glu251* PTEN p.E256*
EC74	MSS	<i>POLE</i> Pro286Arg	<i>PIK3CA</i> c.3141T>G <i>PIK3CA</i> c.1070G>A	PK3CA p.His1047Gln PK3CA p.Arg357Gln
EC167	MSS	<i>POLE</i> Pro286Arg	<i>FBXW7</i> c.1514G>T¶	FBXW7 p.Arg505Leu¶
EC173	MSS	<i>POLE</i> Ser297Phe	<i>PIK3CA</i> c.263G>A <i>PIK3CA</i> c.3062A>G <i>PTEN</i> c.388C>G	PK3CA p.Arg88Gln PK3CA p.Tyr1021Cys PTEN p.Arg130Gly

¶ Indicates variant validated by Sanger sequencing

Table S2. Site and predicted consequence of *POLE* and *POLD1* exonuclease domain mutations not detected in our panel but present in the TCGA EC set

Gene	Nucleotide change	Amino acid change	Number of tumours	Site	SIFT score	PolyPhen-2 score	PhyloCons Score	MutationTaster score
<i>POLE</i>	c.1270C>G	p.Leu424Val	2	Exo IV motif	0.01	0.993	1.00	0.87
	c.1282G>A	p.Ala428Thr	1	Exo IV motif	0.79	0.041 (Benign)	1.282	1.58
	c.1331T>A	p.Met444Lys	1	Between Exo III and V motifs	0.00	1.000	1.00	2.59
	c.1358A>G	p.Gln453Arg	1	Exo III motif	0.06	0.025 (benign)	1.00	1.17
	c.1394C>T	p.Ala465Val	1	Exo III motif	0.00	1.000	1.00	1.75
<i>POLD1</i>	c.1174G>A	p.Val392Met	1	Flanking Exo I motif	0.000	0.946	0.999	Polymorphism 0.57

Table S3. Frequency and type of base substitutions according to tumour POLE/POLD mutation status in TCGA dataset

Category	Proportion of total base substitutions, % (95 CI)						Number of substitutions (range)
	C>T G>A	A>G T>C	C>G G>C	G>T C>A	A>C T>G	A>T T>A	
No <i>POLE/POLD1</i> mutation	59.1 (57.6-59.1)	11.8 (10.9-11.8)	8.7 (7.6-8.7)	13.5 (12.9-13.5)	3.5 (3.1-3.5)	3.4 (3-3.4)	173.9 (140-208)
<i>POLE</i> EDM							
p.Pro286Arg	42.9 (38.7-47)	7.6 (6.6-8.6)	0.3 (0.2-0.3)	37.4 (34.2-40.5)	11.1 (9.3-12.9)	0.8 (0.6-0.9)	7355.1 (4887-9824)
p.Ser297Phe	56.4	6.4	0.1	32.4	3.5	1.2	12699
p.Val411Leu	55.1 (49.3-61)	10.8 (9.3-12.3)	0.5 (0.4-0.7)	23.3 (21-25.6)	9.5 (6.1-12.9)	0.7 (0.6-0.8)	7743 (2474-13012)
p.Leu424Val	58.1	9.9	0.8	26.5	3.2	1.4	4749
p.Ala428Thr	59.3	3.7	11.1	11.1	11.1	3.7	27
p.Met444Lys	38.6	4.8	0.3	44.9	10.1	1.2	1523
p.Gln453Arg	65.1	20.0	0.9	10.0	1.5	2.5	976
p.Ala456Pro	60.1	8.3	0.3	25.1	5.0	1.3	8385
p.Ala465Val	85.0	5.3	0.4	8.3	0.5	0.6	5592
<i>POLE</i> non-ED mutated	58.49 (40.1-76.9)	22.47 (4.8-40.2)	2.4 (0.6-4.2)	9.8 (5.8-13.8)	3.0 (1.4-4.5)	3.9 (2.0-5.9)	636 (205.0-1067.0)
<i>POLD1</i> EDM							
p.Val392Met	81.3	9.8	1.3	5.7	0.9	1.1	1637
<i>POLD1</i> non-ED mutated	63.7 (59.0-68.4)	15.1 (8.6-21.6)	3.8 (0.3-7.2)	11.9 (8.7-15.2)	2.1 (0.7-3.4)	3.5 (2.3-4.7)	606 (72.5-1139.5)

Table S4. Pattern of mutation of EC driver alterations in EDM tumours in TCGA dataset

Tumour ID	Gene	Protein alteration	Hypermutated	MLH1 methylation	Somatic mutation	Protein alteration
A05Z	<i>POLE</i>	p.Pro286Arg	+	-	<i>PTEN</i> c.389G>A, <i>PTEN</i> c.601G>T	PTEN p.Arg130Gln PTEN p.Glu201*
A0J0	<i>POLE</i>	p.Pro286Arg	+	-	<i>CTNNB1</i> c.110C>G <i>PTEN</i> c.389G>A	<i>CTNNB1</i> p.Ser37Cys PTEN p.Arg130Gln
A0JY	<i>POLE</i>	p.Pro286Arg	+	-	<i>CTNNB1</i> c.94G>T <i>PTEN</i> c.19G>T, c.140delG:	<i>CTNNB1</i> p.Asp32Tyr PTEN p.Glu7*, p.Arg47fs
A11N	<i>POLE</i>	p.Pro286Arg	+	-	<i>FBXW7</i> c.1394G>A <i>KRAS</i> c.182A>T <i>PTEN</i> c.19G>T:	<i>FBXW7</i> p.Arg465His <i>KRAS</i> p.Gln61Leu PTEN p.Glu7*
A0UF	<i>POLE</i>	p.Pro286Arg	+	-	<i>PTEN</i> c.389G>A <i>TP53</i> c.637C>T:	PTEN p.Arg130Gln TP53 p.Arg213*
A0UV	<i>POLE</i>	p.Pro286Arg	+	-	<i>PIK3CA</i> c.263G>A <i>PTEN</i> c.389G>A, c.720C>G <i>TP53</i> c.637C>T	<i>PK3CA</i> p.Arg88Gln PTEN p.Arg130Gln, p.Tyr240* TP53 p.Arg213*
A16X	<i>POLE</i>	p.Pro286Arg	+	-	<i>PTEN</i> c.389G>A, c.601G>T	PTEN p.Arg130Gln, p.Glu201*
A17Q	<i>POLE</i>	p.Pro286Arg	+	-	<i>FBXW7</i> c.711G>A <i>PIK3CA</i> c.353G>A <i>PTEN</i> c.1008C>A	<i>FBXW7</i> p.Trp237* <i>PIK3CA</i> p.Gly118Asp PTEN p.Tyr336*
A059	<i>POLE</i>	p.Ser297Phe	+	-	<i>TP53</i> c.718A>G	<i>TP53</i> p.Ser240Gly
A0GP	<i>POLE</i>	p.Val411Leu	+	-	<i>FBXW7</i> c.1393C>T <i>KRAS</i> c.35G>T <i>PIK3CA</i> c.263G>A, c.353G>A <i>PTEN</i> c.G389A, c.697C>T	<i>FBXW7</i> p.Arg465Cys <i>KRAS</i> p.Gly12Val <i>PK3CA</i> p.Arg88Gln, p.Gly118Asp PTEN p.Arg130Gln, p.Arg233*
A056	<i>POLE</i>	p.Val411Leu	+	-	<i>FBXW7</i> c.1105G>T <i>KRAS</i> c.38G>T <i>PIK3CA</i> c.263G>A, c.3062A>G <i>PTEN</i> c.389G>A	<i>FBXW7</i> p.Glu369* <i>KRAS</i> p.Gly13Val <i>PK3CA</i> p.Arg88Gln, p.Tyr1021Cys PTEN p.Arg130Gln
A0LM	<i>POLE</i>	p.Val411Leu	+	-	<i>FBXW7</i> c.1513C>G	<i>FBXW7</i> p.Arg505Gly
A11E	<i>POLE</i>	p.Val411Leu	+	-	<i>KRAS</i> c.35G>T <i>PTEN</i> c.895G>T <i>TP53</i> c.1024C>T	<i>KRAS</i> p.Gly12Val PTEN p.Glu299* TP53 p.Arg342*
A16Y	<i>POLE</i>	p.Val411Leu	+	-	<i>FBXW7</i> c.1972C>T <i>PTEN</i> c.259C>T, c.389G>A <i>TP53</i> c.638G>A	<i>FBXW7</i> p.Arg658* PTEN p.Gln87*, p.Arg130Gln TP53 p.Arg213Gln
A051	<i>POLE</i>	p.Leu424Val	+	+	Nil found	Nil found
A0VX	<i>POLE</i>	p.Leu424Val	-	-	<i>CTNNB1</i> c.94G>T <i>PIK3CA</i> c.1637A>G, c.263G>A	<i>CTNNB1</i> p.Asp32Tyr <i>PK3CA</i> p.Gln546Arg, p.Arg88Gln
A1DQ	<i>POLE</i>	p.Ala428Lys	-	-	<i>PIK3CA</i> c.317G>T <i>TP53</i> c.524G>A	<i>PK3CA</i> p.Gly106Val TP53 p.Arg175His
A0TC	<i>POLE</i>	p.Met444Lys	+	-	<i>KRAS</i> c.35G>T <i>PIK3CA</i> c.323G>A <i>PTEN</i> c.388C>G, c.437T>A	<i>KRAS</i> p.Gly12Val <i>PK3CA</i> p.Arg108His PTEN p.Arg130Gly, p.Leu146*
A11H	<i>POLE</i>	p.Gln453Arg	+/-	-	<i>PTEN</i> c.526_528del, c.606_607del	PTEN p.176_176del, p.202_203del
A103	<i>POLE</i>	p.Ala456Pro	+	+	<i>FBXW7</i> c.1972C>T, c.2065C>T <i>PTEN</i> c.19G>T	<i>FBXW7</i> p.Arg658*X, p.Arg689Trp PTEN p.Glu7*
A0J1	<i>POLE</i>	p.Ala465Val	+	+	<i>PTEN</i> c.389G>A	PTEN p.Arg130Gln
A0VP	<i>POLD1</i>	p.Val392Met	+	-	<i>FBXW7</i> c.40C>T <i>PIK3CA</i> c.G263G>A, c.1625A>C <i>PTEN</i> c.518G>A, c.697C>T	<i>FBXW7</i> p.Arg14* <i>PIK3CA</i> p.Arg88Gln, p.Glu542Ala PTEN p.Arg173His, p.Arg233*