Discovery of genetic defects in unexplained colorectal cancer syndromes

Jansen, A.M.L.

Chapter 3

Combined mismatch repair and *POLE/POLD1* defects explain unresolved suspected Lynch Syndrome cancers

AML Jansen[1,2], T. van Wezel[1], BEWM van den Akker[1], M Ventayol Garcia[1], D Ruano[1], CMJ Tops[3], A Wagner[4], TGW Letteboer[5], EB Gómez-García[6], P Devilee[2], JT Wijnen[2,3], FJ Hes[3], H Morreau[1]

EJHG, 2016; 24 (7):1089-92

Departments of [1]Pathology, [2]Human Genetics and [3]Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands.

[4]Department of Clinical Genetics, Erasmus University Medical Center, Rotterdam, The Netherlands

[5]Department of Medical Genetics, University Medical Center Utrecht, Utrecht, The Netherlands

[6]Department of Clinical Genetics, University Hospital Maastricht, Maastricht, The Netherlands

Abstract

Many suspected Lynch Syndrome (sLS) patients who lack mismatch repair (MMR) germline gene variants and *MLH1* or *MSH2* hypermethylation are currently explained by somatic MMR gene variants or, occasionally, by germline *POLE* variants. To further investigate unexplained sLS patients, we analyzed leukocyte- and tumor DNA of a cohort of 62 sLS patients using gene panel sequencing including the *POLE, POLD1* and MMR genes. Forty tumors showed either one, two or more somatic MMR variants predicted to affect function. Nine sLS-tumors showed a likely ultramutated phenotype and were found to carry germline- (n=2) or somatic variants (n=7) in the *POLE/POLD1* exonuclease domain (EDM). Six of these *POLE/POLD1* mutated tumors also carried somatic MMR variants. Our findings suggest that faulty proofreading may result in loss of MMR and thereby in microsatellite instability.

Keywords: POLE; POLD1; suspected Lynch Syndrome; Colorectal Cancer; Mismatch Repair Deficiency

Introduction

Inactivation of the mismatch repair (MMR) genes *MLH1*, *MSH2*, *MSH6* and *PMS2* causes Lynch Syndrome (LS), an autosomal dominant predisposition for colorectal and endometrial cancer.[1] Inactivation of the mismatch repair pathway can also occur sporadically, through somatic *MLH1* methylation[1] or by acquired bi-allelic somatic inactivation (variant affecting function or loss of heterozygosity (LOH)) of the MMR genes.[2,3] Inaccurate DNA repair leads to a high frequency of somatic variants, with loss of MMR leading to 'hypermutated' tumors with 10-100 variants/Mb.[4] LS tumors are characterized by microsatellite instability (MSI) and immunohistochemical loss of expression of MMR proteins.[1] However, germline variants affecting function cannot be detected in up to 59% of patients displaying MSI and/or loss of MMR, referred to as 'suspected LS' (sLS).[5]

Recently, germline and somatic variants in the exonuclease domains (EDM) of DNA polymerase ε (*POLE*) and polymerase δ (*POLD1*) were described.[6-14] These *POLE*/*POLD1* variants affect proofreading function and lead to an ultramutated phenotype with a variant incidence exceeding 100 variants/Mb**.** Germline *POLE*-EDM variants can result in a LS phenotype and microsatellite unstable CRCs.[6,15] The exact role of somatic *POLE*/*POLD1* variants in tumors with high microsatellite instability (MSI-H) remains unclear.

The aim of our study was to identify the underlying genetic cause of disease in a cohort of 64 suspected LS cases - selected on the basis of MSI, loss of MMR, young onset and often a family history for LS - by screening the MMR, *POLE* and *POLD1* genes in both leukocyte and tumor DNA.

Subjects and Methods

This study included 64 patients with Lynch-associated tumors recruited in four academic centers in the Netherlands between 1990 and 2014: Leiden University Medical Center (n=36), Maastricht University Medical Center (n=11), Erasmus Medical Center (n=9) and University Medical Center Utrecht (n=7). Demographic and clinical data, as well as informed consent, were obtained at the time of diagnosis. Patients were selected based on loss of MMR (as indicated by immunohistochemical staining) and/or microsatellite instability (MSI). Unexplained tumors with low microsatellite instability, or tumors with inconclusive IHC results were also included in this study (See Supplementary Table 1 and Supplementary Methods).

Fifty-six (90%) patients fulfilled Bethesda criteria[16], and families of twenty-two (34%) patients also fulfilled Amsterdam II criteria.[17] Patients were previously screened in a diagnostic setting for germline MMR variants. While 58 patients showed no disease causing germline variants, six patients were found to have a germline variant of unknown significance (VUS). Of the total cohort, 75% of patients presented with colorectal cancer (CRC, n=48), 14% with endometrial cancer (EC, n=9), and 11% with another LS-associated tumor (See Supplementary Table 1).

The average age of onset was 52.1 years. Two patients were excluded from the analysis due to poor DNA quality. Of the remaining 62 tumors, tumor- and leukocyte DNA was sequenced for variants in the exonic regions of *MLH1*, *MSH2*, *MSH6*, *PMS2, POLE* and *POLD1* using the Ion PGM™ System (Life Technologies, Carlsbad, CA). Raw data analysis, alignments, and variant calling was carried out using the default parameters in Torrent Suite v4.0 (see Supplementary information for detailed description). Variants were functionally annotated using ANNOVAR.[18]

The full dataset was filtered and prioritized by variant frequency (>10%) and coverage (>50x). Interesting variants under the 10% were manually curated. *In silico* prediction programs were used to predict pathogenicity (see Supplementary Methods). All variants (likely) affecting function, including two variants with a 9% variant frequency, were validated with Sanger sequencing. For all *PMS2* variants, *PMS2* specific primers were created, to validate that the variant is present in *PMS2* and not in a *PMS2* pseudogene.

Loss of heterozygosity (LOH) was determined for every heterozygous SNP by comparing the ratio of allele A to allele B in leukocyte and tumor DNA samples. Furthermore, for every heterozygous SNP the allelic imbalance factor (AIF) [19] was calculated and a Fisher exact test was performed to determine whether the difference between normal and tumor is significant. If all heterozygous SNPs of one gene showed loss of heterozygosity with an AIF>2 and Fisher exact p-value <0.05 LOH was called. (Supplementary Table 1).

Results

Six patients with a germline MMR VUS (class 3) were included in this study (Supplementary Table 1). In all cases, the variant is detected with NGS, in leukocyte and tumor DNA. During the course of the study, three of these germline variants were reclassified as class 4 or 5 ((probably) affects function) by the International Society for Gastrointestinal Hereditary Tumors Incorporated (InSiGHT). Three of these patients displayed a second somatic MMR variant predicted to affect function. One tumor displayed LOH (Supplementary Table 1). Six cases with somatic *MLH1* hypermethylation fulfilling Revised Bethesda criteria (three from families fulfilling Amsterdam II criteria) were sequenced for underlying hereditary defects explaining the family history, but no germline variants were found. One of these *MLH1* methylated tumors carried a somatic *MLH1* variant likely to affect function and one displayed *MLH1* LOH (Supplementary Table 1).

One (n=27, 44%) or two (n=13, 21%) somatic aberrations (variant or LOH) in a MMR gene were found in a total of 62 tumors (See Supplementary Table 1). Bi-allelic inactivation was concordant with IHC. Twelve of the thirteen tumors with two somatic aberrations had variants in *MLH1* or *MSH2* and were MSI-H. The thirteenth tumor, sLS-07, showed expression loss of *MSH6* and was MSS. While the majority (81%) of tumors showed less than 10 somatic variants in the genomic region analyzed, ten cases displayed a larger number of somatic variants, ranging from 16 to 375 somatic variants within the sequenced area of 31 kb. Nine out of ten tumors showed a *POLE* or *POLD1* variant which (probably) affects function (Table 1). Of the highly mutated tumors, two carried novel germline heterozygous *POLE/POLD1* variants that are predicted to affect proofreading (Supplementary Table 2). Of these two germline cases, tumor sLS-67, was also found to carry two somatic *MLH1* variants, explaining the tumor phenotype (loss of MLH1 expression and MSI-H). The second tumor, sLS-16, was MSI-L, showed positive MMR expression and had no somatic MMR variants.

Seven of the highly mutated tumors showed somatic *POLE*/*POLD1* variants likely to affect function. Six tumors carried a somatic *POLE/POLD1*-EDM hotspot variant (*POLE*: c.857C>G, c.856C>T, c.1231G>T, c.1366G>C, c.1367C>T or c.1376C>T and *POLD1* c.1433G>A) that has previously been described to impair proofreading [4]. In the seventh tumor (sLS-105) a novel *POLE* c.846_847delinsTT variant was detected. This variant lies close to a known *POLE* hotspot site (*POLE* c.857) and is predicted to be affect function by 2 out of 3 prediction programs (Supplementary Table 2). All *POLE* variants were heterozygous, in agreement with previous research [7]. Four *POLE/POLD1-EDM* mutated tumors displayed additional somatic nonsense *POLE* variants outside the exonuclease domain (See Supplementary Table 3). Only one of these was upstream of the exonuclease domain (sLS-16).

Eight of the nine *POLE/POLD1*-EDM mutated tumors in our study showed microsatellite instability (3 MSI-H and 5 MSI-L). In six of these tumors IHC detected loss of at least one of the MMR proteins and all six tumors displayed somatic variants in the affected MMR gene likely to affect function. well as two somatic *PMS2* aberrations, while IHC showed solitary loss of PMS2 expression. Reanalysis of staining also showed ambiguous MLH1 staining (cytoplasmic enhancement and vague, focal nuclear staining).Tumor sLS-19 with two POLE variants, was found to have two somatic *MLH1* aberrations, as well as two somatic *PMS2* aberrations, while IHC showed solitary loss of PMS2 expression.

Table 1: Overview of patients with a *POLE/POLD1*-EDM mutated tumor

Patient	IHC	MSI	Onset age	Family history	# var	Gene	Variant	Amino acid alteration	%
Somatic *POLE/POLD1*-EDM variant									
sLS-05	None	L	62	FDR	330	*POLE*	c.1367C>T	p.(A456V)	28%
						MLH1	LOH	-	-
						PMS2	LOH	-	-
sLS-07	MSH6	S	39	TDR	37	*POLE*	c.1231G>T	p.(V411L)	44%
						MSH6	c.2735G>A	p.(W912*)	10%
						MSH6	c.2876 G>A	p.(R959H)	14%
sLS-09	MSH6	L	42	FDR	16	*POLE*	c.857C>G	p.(P286R)	38%
						MSH6	c.2539G>T	p.(E847*)	36%
sLS-19	PMS2	H	45	FDR	221	*POLE*	c.1376C>T	p.(S459F)	21%
						POLE	c.856C>T	p.(P286S)	9%
						MLH1	c.199G>A	p.(G67R)	19%
						MLH1	LOH	-	-
						PMS2	c.308C>T	p.(T103I)	11%
						PMS2	LOH	-	-
sLS-24	None	L	34	FDR	115	*POLE*	c.1366G>C	p.(A456P)	28%
sLS-66	MSH6	L	66	SDR	25	*POLD1*	c.1433G>A	p.(S478N)	32%
						MSH6	c.3600_3601del	p.(L1201Hfs*13)	28%
sLS-80	MSH2/ MSH6	H	52	FDR	5	*POLD1*	c.1429G>A	p.(V477M)	26%
						MSH6	c.3961A>G**	p.R1321G	52%
						MSH6	c.3186C>A	p.C1062*	25%
sLS-87	MSH2/ MSH6	H	49	FDR	9	*POLE*	c.1218C>G	p.(N406E)	16%
						MSH6	c.3473_3475del**	p.(C1158del)	53%
						MSH6	c.3311_3312del	p.(F1104Trpfs*3)	21%
sLS-101	PMS2	H	55	FDR	5	*POLD1*	c.1003A>G	p.(I335V)	22%
						PMS2	c.1687C>T	p.R563*	30%
sLS-105	MLH1/ PMS2	H	49	No	184	*POLE*	c.846_847delinsTT	p.(L283F)	13%
						MLH1	c.1614G>A	p.(W538*)	13%
Germline *POLE/POLD1*-EDM variant									
sLS-16	None	L	41	FDR	185	*POLD1*	c.961G>A	p.(G321S)	55%
sLS-67	MLH1/ PMS2	H	53	SDR	91	*POLE*	c.861T>A	p.(D287E)	50%
						MLH1	c.208-1G>A	p.?	14%
						MLH1	c.440_447del	p.(G147Dfs*22)	19%

EDM, exonuclease domain. MSI-status is defined as MSI-H (H), MSI-L (L) or MSS (S). Onset age is the age at which the first LS-associated tumor occurred. All patients presented with colorectal cancer, except patient sLS-87 and sLS-101 which presented with endometrial cancer. #var depicts the number of somatic variants with a frequency >10% identified in the sequenced region of 31 kb. Stop codons are indicated with an asterisk (*). Germline variants of unknown significance (VUS) are indicated with a double asterisk (**) % shows the percentage of variant reads. Family history is defined closest relative with LS-associated tumor, FDR: first-degree relative, SDR: second-degree relative, TDR: third-degree relative or no LS in the family (No).

Reanalysis of staining also showed ambiguous MLH1 staining (cytoplasmic enhancement and vague, focal nuclear staining). In three *POLE/POLD1* mutated tumors with positive MMR expression and MSI-L phenotype, no somatic MMR variants (likely) to affect function were found. However, in one of those three tumors (sLS-05) solitary *MLH1* LOH and *PMS2* LOH without variants was found (See Table 1).

In three non-ultramutated tumors a *POLE/POLD1*-EDM variant was found (sLS-80, sLS-87 and sLS-101, see Table 1). These variants have not been described before, but are predicted to affect function (See Supplementary Table 2). Two variants co-occur with a germline *MSH6* VUS and somatic *MSH6* variant (sLS-80 and sLS-87), while one (sLS-101) co-occurs with a somatic *PMS2* variant. Four additional non-ultramutated tumors showed *POLE* -(sLS-18, sLS-21) or *POLD1* LOH (sLS-12, sLS-49) in all heterozygous SNPs (Supplementary Table 1), without germline or somatic *POLE* variants. None of these tumors displayed an ultramutated phenotype.

Discussion

POLE-EDM variants are reported to be the mutagenic factor driving ultramutation in tumors.[20] The number of variants detected in the sequenced area in the present study implicates an ultramutated phenotype, with >100 variants/Mb in all *POLE/POLD1* mutated tumors in this cohort. Since only a limited region (31 Kb) was sequenced, we can only extrapolate the total number of variants per Mb. In our cohort the *POLE/POLD1* mutated MMR-deficient tumors display two deficient pathways increasing the mutational load. Comparing frequencies of the different variants found in these tumors, it might be concluded that faulty proofreading may be the initiating event in these tumors, possibly resulting in loss of MMR and thereby in microsatellite instability. Interestingly, four tumors show *POLE/POLD1* LOH without germline or somatic *POLE/POLD1* variants. These tumors however do not show the typical ultramutated phenotype, whereas single variants without LOH do show that. This phenomenon of LOH without variants affecting the exonuclease domain has not yet been described. Possible the remaining allele is enough to maintain proofreading. Furthermore, three tumors show somatic *POLE/POLD1* variants, without the ultramutated phenotype. All three variants are missense, but are predicted to affect function (Supplementary Table 2). Since these variants are not found in ultramutated tumors, evidence of pathogenicity is lacking.

In conclusion, targeted next-generation sequencing of 62 sLS cases led to the detection of nine higly mutated tumors with a germline- (n=2) or somatic- (n=7) *POLE/POLD1-EDM* variant. Even though *POLE* germline variants have previously been shown to co-occur with somatic MMR variants[11], in this study we found germline and somatic *POLE/POLD1* variants in a cohort selected for sLS characteristics. Importantly, while current literature mainly addressed *POLE/POLD1* variants in MSS tumors, somatic *POLE/POLD1* variants in sLS patients are likely to be overlooked. Screening of *POLE/POLD1* should be added to the current germline and somatic diagnostic screening for MSI-H and MMR-deficient cases and could resolve the causal defect in these presently unexplained cases.

Supplementary information accompanies this paper on European Journal of Human Genetics website (*http://www.nature.com/ejhg*)

References

1. Colas C, Coulet F, Svrcek M, et al. Lynch or not Lynch? Is that always a question? Adv Cancer Res 2012;113:121-66.
2. Mensenkamp AR, Vogelaar IP, van Zelst-Stams WA, et al. Somatic mutations in MLH1 and MSH2 are a frequent cause of mismatch-repair deficiency in Lynch syndrome-like tumors. Gastroenterology 2014;146:643-6 e8.
3. Geurts-Giele WR, Leenen CH, Dubbink HJ, et al. Somatic aberrations of mismatch repair genes as a cause of microsatellite-unstable cancers. J Pathol 2014;234:548-59.
4. 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.
5. Buchanan DD, Rosty C, Clendenning M, Spurdle AB, Win AK. Clinical problems of colorectal cancer and endometrial cancer cases with unknown cause of tumor mismatch repair deficiency (suspected Lynch syndrome). Appl Clin Genet 2014;7:183-93.
6. Heitzer E, Tomlinson I. Replicative DNA polymerase mutations in cancer. Curr Opin Genet Dev 2014;24:107-13.
7. Palles C, Cazier JB, Howarth KM, et al. Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. Nat Genet 2013;45:136-44.
8. Valle L, Hernandez-Illan E, Bellido F, et al. New insights into POLE and POLD1 germline mutations in familial colorectal cancer and polyposis. Hum Mol Genet 2014;23:3506-12.
9. Briggs S, Tomlinson I. Germline and somatic polymerase epsilon and delta mutations define a new class of hypermutated colorectal and endometrial cancers. J Pathol 2013;230:148-53.
10. Spier I, Holzapfel S, Altmuller J, et al. Frequency and phenotypic spectrum of germline mutations in POLE and seven other polymerase genes in 266 patients with colorectal adenomas and carcinomas. Int J Cancer 2015;137:320-31.
11. Haraldsdottir S, Hampel H, Tomsic J, et al. Colon and endometrial cancers with mismatch repair deficiency can arise from somatic, rather than germline, mutations. Gastroenterology 2014;147:1308-16 e1.
12. Yoshida R, Miyashita K, Inoue M, et al. Concurrent genetic alterations in DNA polymerase proofreading and mismatch repair in human colorectal cancer. Eur J Hum Genet 2011;19:320-5.
13. 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.
14. Billingsley CC, Cohn DE, Mutch DG, Stephens JA, Suarez AA, Goodfellow PJ. Polymerase varepsilon (POLE) mutations in endometrial cancer: clinical outcomes and implications for Lynch syndrome testing. Cancer 2015;121:386-94.
15. Elsayed FA, Kets CM, Ruano D, et al. Germline variants in POLE are associated with early onset mismatch repair deficient colorectal cancer. European journal of human genetics : EJHG 2015 Aug; 23(8):1080-4
16. Umar A, Boland CR, Terdiman JP, et al. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. J Natl Cancer Inst 2004;96:261-8.
17. Park J-G, Vasen HF, Park Y, et al. Suspected HNPCC and Amsterdam criteria II: evaluation of mutation detection rate, an international collaborative study. International Journal of Colorectal Disease 2001;17:109-14.
18. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010;38:e164.
19. Devilee P, Cleton-Jansen AM, Cornelisse CJ. Ever since Knudson. Trends Genet 2001;17:569-73.
20. Helleday T, Eshtad S, Nik-Zainal S. Mechanisms underlying mutational signatures in human cancers. Nat Rev Genet 2014;15:585-98.

Supplementary Methods

Study subjects
This retrospective cohort was collected between 1997 and 2014. Leukocyte DNA isolated from peripheral blood was available for all patients. Tumor DNA was isolated from formalin-fixed paraffin-embedded tumor tissue (FFPE) at time of diagnosis. Two patients were excluded from the analysis due to poor quality DNA. This study was approved by the local medical ethical committee of the LUMC (P01-019E).

Immunohistochemical analysis (IHC) and microsatellite instability testing was performed previously at request of board certified Clinical Genetics medical specialists. Routine testing of all four MMR proteins became available only around 2004. Therefore not of all tumors in our cohort complete MMR immunohistochemical results are available. At time of the current study leukocyte- and tumor DNA was retrieved from our archives. In a minority of cases FFPE-blocks were still available for retrospective testing. In 38 patients (61%) immunohistochemistry data was complete. In 11 tumors PMS2 immunohistochemistry was not performed. The remaining 13 tumors had one of more inconclusive immunohistochemical results. Patient sLS-05 and sLS-24 were initially included due to negative MSH6 staining, but reanalysis during the study showed no MSH6 expression loss. Patient sLS-11 was included due to a reported MSI-H status, but reanalysis showed no instability.

Targeted next-generation sequencing
Tumor and leukocyte DNA was sequenced for variants in *MLH1*, *MSH2*, *MSH6*, *PMS2, POLE* and *POLD1* using the Ion PGM™ System (Life Technologies). Ion AmpliSeq™ Custom Panels were designed with the Ion AmpliSeq™ Designer tool. The complete sequencing panel consisted of 307 amplicons (31094 bp), covering 98%, 90%, 98%, 75%, 95%, and 78% of the coding regions of *MLH1, MSH2, MSH6, PMS2, POLE, and POLD1*, respectively. Libraries were prepared with Ion AmpliSeq™ Library Kit 2.0 according to the manufacturer's protocol. For template preparation the Ion OneTouch™ 2 System and the Ion Chef™ System were used.

Data analysis
Raw data analysis, alignments, and variant calling was carried out using the default parameters in Torrent Suite v4.0. The Variant Caller Parameter Setting was set on 'Somatic – PGM – Low Stringency'. Variants were functionally annotated using ANNOVAR.[1] The following Genbank reference sequences were used: NM_000249.3 for *MLH1*, NM_000251.2 for *MSH2*, NM_000179.2 for *MSH6*, NM_000535.5 for *PMS2*, NM_006231.2 for *POLE* and NM_001256849.1 for *POLD1*. Recommendations of the Human Genome Variation Society (HGVS) to use the terms "variant" and "likely to affect function" instead of "mutation" and "pathogenic" were followed (*http://www.hgvs.org/mutnomen/recs.html*). Classification of the functional effects of the variants was done according to the five-tiered InSiGHT scheme.[2]

Variant filtering and validation
The full dataset was filtered by variant frequency (>10%) and coverage (>50x). Variants with a minor allele frequency (MAF) >0.05, as reported in the exome-database (*http://www.ncbi.nlm.nih.gov/projects/SNP/*), or a MAF >0.01, as reported in the Genome of the Netherlands (*http://nlgenome.nl*), were excluded from further analysis. In addition, coding variants

were filtered further based on predictions by *in silico* missense prediction software Align GVGD[3], SIFT (*http://sift.jcvi.org/*), PolyPhen-2 (*http://genetics.bwh.harvard.edu/pph2/*), MutationTaster (*http://www.mutationtaster.org/*) and MAPP (*http://mendel.stanford.edu/SidowLab/downloads/MAPP/index.html*). The Leiden Open Variation Database (LOVD) was consulted to find variants previously described and classified (*http://www.lovd.nl/3.0/home*). All predicted to affect function were visually inspected with the Integrative Genomics Viewer (*https://www.broadinstitute.org/igv/home*).Variants predicted to affect function were validated with Sanger sequencing. Germline variants of this study are submitted to the Leiden Open (source) Variation Database (LOVD) available at *http://www.lovd.nl/3.0/home* and somatic variants are submitted to the COSMIC database (*http://cancer.sanger.ac.uk/cosmic*) .

References

1. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic acids research. 2010;38(16):e164.
2. Thompson BA, Spurdle AB, Plazzer JP, Greenblatt MS, Akagi K, Al-Mulla F, et al. Application of a 5-tiered scheme for standardized classification of 2,360 unique mismatch repair gene variants in the InSiGHT locus-specific database. Nature genetics. 2014;46(2):107-15.
3. Tavtigian SV, Deffenbaugh AM, Yin L, Judkins T, Scholl T, Samollow PB, et al. Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. Journal of medical genetics. 2006;43(4):295-305.

Supplementary Table 1: Overview of all variants that are prediced to affect function

Patient	Gender	IHC MLH1	IHC MSH2	IHC MSH6	IHC PMS2	Tumor	Age of onset	MSI	Family History	# var	POLE/POLD1 variant	%	MMR variants	%
sLS-01	Male	-	+	+	-	CRC	54	Unknown	FDR	4	-	-	MLH1 c.1985_1988del	34%
sLS-03	Male	-	+	+	-	CRC	52	High	Unknown	5	-	-	MLH1 c.1252_1253delGA MLH1 LOH	64% -
sLS-05	Female	+	+	+	+	CRC	62	Low	FDR	330	POLE c.1367C>T	28%	MLH1 LOH PMS2 LOH	- -
sLS-06	Male	INC	-	NT	NT	CRC	47	High	FDR	8	-	-	MSH2 c.2131 C>T MSH2 c.1600_1601delCG	20% 19%
sLS-07	Male	INC	INC	-	NT	CRC	39	Stable	TDR	37	POLE c.1231 G>T	44%	MSH6 c.2876G>A MSH6 LOH	14% -
sLS-09	Male	+	+	-	NT	CRC	42	Low	FDR	16	POLE c.857C>G	38%	MSH6 c.2539G>T MSH2 c.1166G>A	36% 38%
sLS-11	Female	INC	INC	NT	NT	Ad	44	Stable	FDR	2	-	-	-	-
sLS-12	Male	+	+	NT	NT	CRC	30	High	FDR	0	POLD1 LOH	-	MLH1 LOH	-
sLS-14	Female	+	+	NT	NT	CRC	62	High	FDR	3	-	-	-	-
sLS-16	Male	+	+	NT	NT	CRC	41	Low	FDR	185	POLD1 c.961G>A*	55%	-	-
sLS-17	Female	+	+	-	NT	CRC	39	Stable	FDR	0	-	-	-	-
sLS-18	Female	+	+	NT	NT	CRC, Oes	62, 71	High	FDR	1	POLE LOH	-	MLH1 LOH	-
sLS-19	Male	AMB	+	+	-	CRC	45	High	FDR	221	POLE c.1376C>T POLE c.856C>T	21% 9%	MLH1 c.199G>A MLH1 LOH PMS2 c.308C>T PMS2 LOH	19% - 11% -
sLS-20	Female	+	INC	INC	NT	CRC	55	Low	FDR	1	-	-	-	-
sLS-21	Female	+	+	-	+	Ad	61	High	FDR	7	POLE LOH	-	MSH6 c.1134_1135delAA	31%
sLS-24	Female	+	+	+	+	CRC	34	Low	FDR	115	POLE c.1366G>C	28%	-	-
sLS-26	Female	+	-	-	NT	EC, OC	57, 57	Low	SDR	0	-	-	-	-

Supplementary Table 1: Overview of all variants that are prediced to affect function; continued

Patient	Gender	IHC MLH1	IHC MSH2	IHC MSH6	IHC PMS2	Tumor	Age of onset	MSI	Family History	# var	POLE/POLD1 variant	%	MMR variants	%
sLS-29	Female	+	-	-	NT	CRC	49	High	FDR	2	-	-	MSH2 c.1786_1788del* MSH2 LOH MSH6 LOH	46% - -
sLS-36	Female	+	+	+	NT	Ad	43	High	No	6	-	-	-	-
sLS-39	Female	+	+	+	NT	OC	68	High	FDR	1	-	-	-	-
sLS-40	Male	+	+	+	INC	CRC	52	High	TDR	3	-	-	-	-
sLS-42	Male	+	INC	INC	INC	CRC	69	High	FDR	47	-	-	PMS2 LOH	-
sLS-43	Male	INC	-	-	NT	CRC	74	High	FDR	1	-	-	-	-
sLS-44	Female	+	-	-	NT	CRC	41	High	FDR	13	-	-	-	-
sLS-45	Female	+	+	-	NT	CRC	47	High	No	2	-	-	MSH2 c.1638_1640del	30%
sLS-46	Female	+	+	+	NT	CRC, EC	76, 79	High	No	2	-	-	-	-
sLS-49	Male	-	+	+	-	CRC	44	High	SDR	4	POLD1 LOH	-	MLH1 c.885-2A>G MLH1 LOH	57% -
sLS-51	Female	-	+	+	-	CRC	69	High	FDR	4	-	-	-	-
sLS-52	Female	-	+	+	-	EC	67	High	FDR	2	-	-	-	-
sLS-53	Male	-	+	+	-	CRC	81	High	FDR	3	-	-	MLH1 c.769A>C	20%
sLS-54	Female	-	+	INC	-	EC	64	High	SDR	2	-	-	-	-
sLS-55	Female	+	-	-	+	OC, EC	48, 48	High	FDR	5	-	-	MSH2 LOH MSH6 c.3974_3976del	- 24%
sLS-56	Male	+	-	-	+	CRC 2x	64 2x	High	Unknown	5	-	-	MSH2 c.1710T>A	19%
sLS-58	Male	-	+	+	-	CRC	39	Low	FDR	6	-	-	MLH1 c.790+1 G>A	28%
sLS-60	Female	-	+	+	NT	OC	44	High	No	1	-	-	MLH1 LOH	-
sLS-61	Female	-	+	+	-	CRC 2x	49, 51	High	FDR	1	-	-	-	-
sLS-62	Male	+	-	-	NT	CRC	35	Low	FDR	2	-	-	-	-

Supplementary Table 1: continued

Patient	Gender	IHC MLH1	IHC MSH2	IHC MSH6	IHC PMS2	Tumor	Age of onset	MSI	Family History	# var	POLE/POLD1 variant	%	MMR variants	%
sLS-64	Male	-	+	+	-	CRC	48	High	FDR	1	-	-	MLH1 c.2059C>T	28%
sLS-66	Male	+	+	-	+	CRC	66	Low	TDR	25	POLD1 c.1433 G>A	32%	MSH6 c.3600_3601delAC	28%
sLS-67	Male	-	+	+	-	CRC 2x	57, 60	High	SDR	91	POLE c.861 T>A*	51%	MLH1 c.208-1 G>A MLH1 c.440_447del	14% 19%
sLS-71	Male	+	-	-	+	CRC, Ur	52, 52	High	FDR	1	-	-	MSH2 c.493T>G*	35%
sLS-72	Female	+	-	-	+	EC, CRC	68, 73	High	Unknown	1	-	-	MSH2 c.1576dupA	29%
sLS-77	Female	+	-	-	+	CRC, Pol	45, 45	High	No	5	-	-	MSH2 c.2470C>T	38%
sLS-80	Male	+	-	-	+	CRC	52	High	FDR	5	POLD1 c.1429G>A	-	MSH6 c.3961A>G** MSH6 c.3186C>A	52% 25%
sLS-82	Female	-	+	+	-	CRC	69	High	FDR	6	-	-	-	-
sLS-84	Female	-	+	+	-	EC	46	High	Unknown	1	-	-	MLH1 c.1989+1 G>A MLH1 c.2224 C>T	11% 9%
sLS-85	Male	+	+	-	+	Ki, CRC	41, 44	High	SDR	4	-	-	MSH6 c.3139T>G	30%
sLS-87	Female	+	-	-	+	EC	49	High	FDR	9	POLE c.1218C>G	-	MSH6 c.3473_3475del** MSH6 c.3311_3312delTT	53% 21%
sLS-89	Male	-	+	+	-	CRC 3x	48, 51, 61	High	TDR	7	-	-	MLH1 c.2041G>A* MLH1 c.677 G>A	50% 28%
sLS-90	Male	+	-	-	+	CRC	46	High	FDR	2	-	-	MSH2 c.490G>T** MSH2 c.1511-1G>T	40% 18%
sLS-91	Male	-	+	+	-	CRC	49	Unknown	No	3	-	-	MLH1 c.350 C>T MLH1 LOH	16% -
sLS-92	Female	-	+	INC	-	CRC	45	High	FDR	3	-	-	MLH1 c.790+1 G>A MLH1 LOH	78% -

Supplementary Table 1: Overview of all variants that are prediced to affect function;continued

Patient	Gender	IHC MLH1	IHC MSH2	IHC MSH6	IHC PMS2	Tumor	Age of onset	MSI	Family History	# vari	POLE/POLD1 variant	%	MMR variants	%
sLS-101	Female	+	+	+	-	EC	55	High	FDR	5	POLD1 c.1003A>G	-	PMS2 c.1687C>T	30%
sLS-104	Male	+	-	-	+	SB	47	High	No	1	-	-	MSH2 c.271delG MSH2 LOH	92% -
sLS-105	Male	-	+	+	-	CRC	49	High	No	184	POLE c.847_846delinsTT	13%	MLH1 c.1614G>A	13%
sLS-109	Female	-	+	+	-	CRC	49	High	No	2	-	-	MLH1 c.1192C>T MLH1 c.207+1 G>A	34% 19%
sLS-111	Female	+	-	-	+	EC	58	High	FDR	8	-	-	MSH2 c.687delA MSH2 c.773 T>A	44% 41%
sLS-113	Male	-	+	+	-	CRC	46	High	SDR	2	-	-	MLH1 c.982 C>T	16%
sLS-115	Female	-	+	+	-	CRC	38	High	Unknown	3	-	-	MLH1 c.1852_1854del MLH1 LOH	54% -
sLS-118	Female	-	+	+	-	EC	55	High	No	8	-	-	MLH1 c.1975C>T MLH1 c.2104-1G>C	33% 36%
sLS-120	Male	-	+	+	-	CRC	57	Stable	FDR	3	-	-	-	-
sLS-124	Male	+	+	-	+	CRC	50	High	FDR	4	-	-	MSH6 c.3484G>C**	48%

Supplementary Table 1: continued

Patient	MLH1 Meth.	Conclusion	ACII or BC	LOH MLH1	LOH MSH2	LOH MSH6	LOH PMS2	LOH POLE	LOH POLD1
sLS-01	No	1 somatic MMR variant	No	-	NP	NP	NP	NP	NP
sLS-03	No	2 somatic MMR variants	BC	1/1	NP	NP	-	NP	NP
sLS-05	NA	Somatic POLE variant	ACII	1/1	-	-	1/1	-	-
sLS-06	NA	2 somatic MMR variants	ACII	-	NP	-	-	NP	NP
sLS-07	NA	Somatic POLE variant	BC	-	NP	3/3	-	-	-
sLS-09	NA	Somatic POLE variant	BC	-	-	-	-	-	-
sLS-11	No	No variant	No	-		-	-	-	-
sLS-12	NA	1 somatic MMR variant	ACII	1/1	-	-	-	-	5/5
sLS-14	NA	No variant	ACII	NP	NP	-	NP	-	-
sLS-16	NA	Germline POLD1 variant	ACII	NP	NP	NP	NP	NP	NP
sLS-17	NA	No variant	BC	-	-	-	-	-	NP
sLS-18	NA	1 somatic MMR variant	BC	1/1	-	-	-	14/14	-
sLS-19	No	Somatic POLE variant	ACII	2/2	-	-	2/2	-	-
sLS-20	NA	No variant	ACII	-	-	-	-	NP	-
sLS-21	No	1 somatic MMR variant	ACII	NP	NP	NP	-	1/1	NP
sLS-24	NA	Somatic POLE variant	ACII	-	NP	-	NP	-	NP
sLS-26	NA	No variant	BC	NP	-	-	-	-	-
sLS-29	NA	Germline MMR variant with LOH	BC	NP	2/2	1/1	-	-	NP
sLS-36	NA	No variant	No	-	NP	-	-	-	-
sLS-39	NA	No variant	No	NP	NP	-	-	NP	-
sLS-40	NA	No variant	BC	-	NP	NP	-	-	NP
sLS-42	No	1 somatic MMR variant	ACII	-	-	-	4/4	NP	-
sLS-43	NA	No variant	BC	-	-	-	-	-	NP

Supplementary Table 1: Overview of all variants that are prediced to affect function; continued

Patient	MLH1 Meth.	Conclusion	ACII or BC	LOH MLH1	LOH MSH2	LOH MSH6	LOH PMS2	LOH POLE	LOH POLD1
sLS-44	NA	No variant	ACII	NP	NP	NP	NP	NP	NP
sLS-45	NA	1 somatic MMR variant	ACII	-	-	-	NP	-	-
sLS-46	NA	No variant	BC	NP	NP	-	NP	-	-
sLS-49	No	2 somatic MMR variants	BC	1/1	-	-	-	-	5/5
sLS-51	Yes	Sporadic - MLH1 Methylation	ACII	NP	-	NP	-	NP	NP
sLS-52	Yes	Sporadic - MLH1 Methylation	BC	-	-	-	-	-	NP
sLS-53	Yes	Sporadic - MLH1 Methylation	ACII	-	NP	NP	-	-	-
sLS-54	Yes	Sporadic - MLH1 Methylation	ACII	-	NP	-	-	NP	NP
sLS-55	NA	1 somatic MMR variant	BC	-	1/1	NP	-	NP	-
sLS-56	NA	1 somatic MMR variant	BC	-	-	-	-	-	NP
sLS-58	No	1 somatic MMR variant	ACII	NP	-	NP	NP	-	NP
sLS-60	Yes	Sporadic - MLH1 Methylation	BC	2/2	NP	-	-	NP	-
sLS-61	Yes	Sporadic - MLH1 Methylation	BC	NP	NP	-	-	-	NP
sLS-62	NA	No variant	BC	-	-	-	-	-	-
sLS-64	No	1 somatic MMR variant	BC	-	NP	-	-	-	-
sLS-66	NA	Somatic POLD1 variant	No	NP	-	-	-		-
sLS-67	No	Germline POLE variant	BC	-	-	-	NP	-	NP
sLS-71	NA	Germline MMR variant	ACII	-	NP	-	-	-	-
sLS-72	NA	1 somatic MMR variant	BC	NP	NP	NP	NP	NP	NP
sLS-77	NA	1 somatic MMR variant	BC	-	-	-	-	-	-
sLS-80	NA	Somatic POLD1 variant, germline VUS and 1 somatic MMR variant	ACII	-	-	-	-	-	-
sLS-82	No	No variant	BC	-	NP	-	-	-	NP
sLS-84	No	2 somatic MMR variants	BC	-	NP	-	-	-	-

Supplementary Table 1: continued

Patient	MLH1 Meth.	Conclusion	ACII or BC	LOH MLH1	LOH MSH2	LOH MSH6	LOH PMS2	LOH POLE	LOH POLD1
sLS-85	NA	1 somatic MMR variant	BC	NP	-	-	-	-	NP
sLS-87	NA	Somatic POLE variant, germline VUS and 1 somatic MMR variant	ACII	-	-	-	-	-	-
sLS-89	No	Germline MMR variant	BC	-	NP	-	-	-	-
sLS-90	NA	Germline VUS and 1 somatic MMR variant	ACII	-	-	-	-	-	-
sLS-91	No	1 somatic MMR variant	BC	1/1	NP	NP	-	-	-
sLS-92	No	2 somatic MMR variants	ACII	1/1	-	NP	-	-	NP
sLS-101	No	Somatic POLE variant	BC	NP	-	-	NP	-	-
sLS-104	NA	2 somatic MMR variants	No	NP	1/1	NP	-	-	-
sLS-105	No	Somatic POLE variant	BC	NP	NP	NP	NP	NP	NP
sLS-109	No	2 somatic MMR variants	BC	-	-	-	-	-	-
sLS-111	NA	2 somatic MMR variants	ACII	NP	-	-	NP	-	-
sLS-113	No	1 somatic MMR variant	ACII	-	NP	-	NP	-	NP
sLS-115	No	2 somatic MMR variants	BC	NP	-	-	-	-	NP
sLS-118	No	2 somatic MMR variants	BC	-	-	-	-	-	-
sLS-120	No	No variant	BC	-	-	-	-	-	-
sLS-124	NA	Germline VUS	BC	NP	NP	-	-	-	-

All variants described are nonsense, frameshift or missense variants that are likely to affect function in the MMR genes and POLE/POLD1 exonuclease domain. Patients presented with colorectal- (CRC), endometrial- (EC), ovarian- (OC), kidney- (Ki), sebaceous gland- (SB), urinary tract- (Ur), esophageal cancer (Oes) or adenoma with high grade dysplasia (Ad). Staining was defined as positive (+), negative (-), inconclusive (INC) or ambiguous (AMB). Family history is defined as first-degree relative with LS-associated tumors (FDR), second-degree relative with LS-associated tumors (SDR), third-degree relative with LS-associated tumor (TDR) or no Lynch-associated tumors in the family (No). Germline variants are indicated with an asterisk (*). Germline variants of unknown significance (VUS, class 3) are depicted with double asterisks (**). % shows the percentage of variant reads. MLH1 hypermethylation status is defined as yes, no or not applicable (NA). Patients fulfilled revised bethesda criteria (BC), Amsterdam II criteria (ACII) or neither (No). LOH was called when all heterozygous SNPs showed significant imbalance. LOH classified as no LOH (-), LOH, with the number of informative SNPs (#/#) or LOH analysis was not possible (NP).

Supplementary Table 2: Missense prediction of *POLE/POLD1* variants

Missense prediction									
***POLE/POLD1-EDM* variants found in hypermutated tumors**									
Patient	**Gene**	**Variant**	**Amino acid alteration**	**Observed**	**Align GVGD**	**SIFT**	**Mutation Taster**	**Previously described**	**Rs id**
sLS-05	*POLE*	c.1367C>T	p.(A456V)	Somatic	C0	Deleterious	Disease causing	Yes	-
sLS-07	*POLE*	c.1231G>T	p.(V411L)	Somatic	C0	Deleterious	Disease causing	Yes	-
sLS-09	*POLE*	c.857C>G	p.(P286R)	Somatic	C0	Deleterious	Disease causing	Yes	-
sLS-16	*POLD1*	c.961G>A	p.(G321S)	Germline	C55	Deleterious	Disease causing	No	rs41554817
sLS-19	*POLE*	c.856C>T	p.(P286S)	Somatic	C0	Deleterious	Disease causing	Yes	-
sLS-19	*POLE*	c.1376C>T	p.(S459F)	Somatic	C0	Deleterious	Disease causing	Yes	-
sLS-24	*POLE*	c.1366G>C	p.(A456P)	Somatic	C0	Deleterious	Disease causing	No	-
sLS-66	*POLD1*	c.1433G>A	p.(S478N)	Somatic	C45	Deleterious	Disease causing	Yes	rs397514632
sLS-67	*POLE*	c.861T>A	p.(D287E)	Germline	C0	Deleterious	Disease causing	No	rs139075637
sLS-110	*POLE*	c.846_847delinsTT	p.(L283F)	Somatic	C0	Deleterious	Disease causing	No	-
***POLE/POLD1-EDM* variants found in non-hypermutated tumors**									
sLS-80	*POLD1*	c.1429G>A	p.(V477M)	Somatic	C15	Deleterious	Disease causing	No	-
sLS-87	*POLE*	c.1218C>G	p.(N406E)	Somatic	C0	Deleterious	Disease causing	No	-
sLS-101	*POLD1*	c.1003A>G	p.(I335V)	Somatic	C25	Deleterious	Disease causing	No	-

Overview of missense prediction of POLE/POLD1 exonuclease (EDM) variants. Missense prediction is based on 3 prediction programs: Align GVGD [C0 - C65, with high values indicating likely deleterious variants], SIFT [Tolerated/Deleterious], MutationTaster [Polymorphism/Disease Causing].

Supplementary Table 3: Additional nonsense variants in POLE outside the EDM-domain

	POLE/POLD1-EDM variants			Somatic MMR variants			Non-EDM *POLE/POLD1* variants			
Patient ID	**Gene**	**Variant**	**%**	**Gene**	**Variant**	**%**	**Gene**	**Variant**	**%**	**Location**
sLS-05	*POLE*	c.1367C>T	28%	*MLH1*	LOH	-	*POLE*	c.4933C>T	23%	A
				PMS2	LOH	-				
sLS-16	*POLD1*	c.961G>A*	55%	*MLH1*	c.208-1 G>A	14%	*POLE*	c.448C>T	24%	B
				MLH1	c.440_447del	19%				
sLS-19	*POLE*	c.1376C>T	21%	*MLH1*	c.199G>A	19%	*POLE*	c.6181C>T	23%	A
				MLH1	LOH	-				
	POLE	c.856C>T	9%	*PMS2*	c.308C>T	11%	*POLE*	c.5287C>T	14%	A
				PMS2	LOH	-				
sLS-105	*POLE*	c.847_846delinsTT	13%	*MLH1*	c.1614G>A	13%	*POLE*	c.2059C>T	17%	A
							POLE	c.2065C>T	17%	A

Patients with additional POLE nonsense variants outside the exonuclease (EDM) domain. % shows variant frequency. Asterisk (*) indicates germline variants. A/B shows whether nonsense variants were before (B) or after (A) the exonuclease domain.