scientific reports

OPEN

A genetic variant in telomerase reverse transcriptase (*TERT*) modifies cancer risk in Lynch syndrome patients harbouring pathogenic *MSH2* variants

Mariann Unhjem Wiik[1], Tiffany-Jane Evans[2], Sami Belhadj[3], Katherine A. Bolton[4], Dagmara Dymerska[5], Shantie Jagmohan-Changur[6], Gabriel Capellá[3], Grzegorz Kurzawski[5], Juul T. Wijnen[6], Laura Valle[3], Hans F. A. Vasen[7], Jan Lubinski[5], Rodney J. Scott[2,4,8] & Bente A. Talseth-Palmer[1,2,4✉]

Individuals with Lynch syndrome (LS), have an increased risk of developing cancer. Common genetic variants of *telomerase reverse transcriptase (TERT)* have been associated with a wide range of cancers, including colorectal cancer (CRC) in LS. We combined genotype data from 1881 LS patients, carrying pathogenic variants in *MLH1*, *MSH2* or *MSH6*, for rs2075786 (G>A, intronic variant), 1207 LS patients for rs2736108 (C>T, upstream variant) and 1201 LS patients for rs7705526 (C>A, intronic variant). The risk of cancer was estimated by heterozygous/homozygous odds ratio (OR) with mixed-effects logistic regression to adjust for gene/gender/country of sample origin considering family identity. The AA genotype of SNP rs2075786 is associated with 85% higher odds at developing cancer compared to GG genotype in *MSH2* pathogenic variant carriers ($p = 0.0160$). Kaplan–Meier analysis also shows an association for rs2075786; the AA allele for *MSH2* variant carriers confers risk for earlier diagnosis of LS cancer (log-rank $p = 0.0011$). We report a polymorphism in *TERT* to be a possible modifier of disease risk in *MSH2* pathogenic variant carriers. The rs2075786 SNP in *TERT* is associated with a differential risk of developing cancer for *MSH2* pathogenic variant carriers. Use of this information has the potential to personalise screening protocols for LS patients.

Lynch syndrome (LS) is an autosomal dominantly inherited cancer syndrome that accounts for 2–5% of all colon cancer cases and approximately 2% of all endometrial cancer (EC) cases. LS is characterised by early-onset epithelial cancers in a variety of organs (including but not limited to; CRC, EC, cancer of the kidneys, duodenum, ureter, ovaries, stomach and brain) at a mean age of disease onset that appears to be much lower than that of the general population[1]. LS is a result of pathogenic germline variants in four DNA mismatch repair (MMR) genes (*MLH1*, *MSH2*, *MSH6* and *PMS2*) or by epigenetic inactivation of *MSH2* due to a deletion at the 3' end of *EPCAM*[2–6], but an obvious pre-malignant phenotype/genotype correlation is not evident. Genetic variation is the foundation of diversity observed in the human phenotype, and accounts for the large variety of susceptibilities to common diseases[7]. Differences in disease expression are not only observed between patients who harbour causative germline variants in different MMR genes, but also between patients carrying variants in the same gene and even in patients harbouring the same variant[8], suggesting that other genetic and environmental factors,

[1]Research Unit, Ålesund Hospital, Møre and Romsdal Hospital Trust, Ålesund, Norway. [2]Hunter Medical Research Institute, New Lambton Heights, Australia. [3]Hereditary Cancer Program, Catalan Institute of Oncology, IDIBELL and CIBERONC, Hospitalet de Llobregat, Barcelona, Spain. [4]School of Biomedical Science and Pharmacy, Faculty of Health and Medicine, University of Newcastle, Newcastle, NSW, Australia. [5]Department of Genetics and Pathology, International Hereditary Cancer Center, Pomeranian Medical University, Szczecin, Poland. [6]Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands. [7]Department of Gastroenterology and Hepatology, Leiden University Medical Center, Leiden, the Netherlands. [8]Division of Genetics, NSW Health Pathology, John Hunter Hospital, Newcastle, NSW, Australia. ✉email: bente.talseth-palmer@newcastle.edu.au

are likely to be involved in the disease process. The search for genetic variants that have a modifying effect on disease expression has been ongoing but without definitive results[9–12].

Telomeres are located at the end of chromosomes and have many functions that are critical for genome stability and integrity[13]. Telomere shortening limits the proliferation of normal somatic cells but not cancer cells, which can maintain long telomeres, usually via the enzymatic activity of telomerase[14–16]. Human cancer cells have the capacity of unlimited proliferation potential, associated with the expression of telomerase activity[17]. Telomere shortening has also been shown to result in chromosomal instability[18]. Besides telomere shortening, telomeres may malfunction due to genetic variation in telomere maintenance genes, especially that orchestrated by *telomerase reverse transcriptase (TERT)*[19]. Significant up-regulation of *TERT* is found in Lynch syndrome CRC as well as microsatellite instable sporadic CRC, indicating the influence this gene has on telomere length[20]. The study also shows that mean telomere length systematically shortened in all tumour tissue in LS cancer and sporadic CRC compared to reference systems[20]. Many *TERT* single nucleotide polymorphisms (SNPs) have been associated with a wide range of cancers[21–25], including CRC[26–28] and LS cancers[29]. Studies have described polymorphisms that can increase the transcriptional activity of the *TERT* promoter and thus may promote cancer progression[30,31]. Moreover, genome-wide association studies (GWASs) and case–control studies have demonstrated that polymorphisms in *TERT* are associated with various cancers such as skin cancer, CRC and breast cancer[28,32]. While individual *TERT* SNPs have been associated with different types of cancer and telomere length[22,26,33,34] and a common haplotype with decreased cancer risk substantially reducing *TERT* promoter activity[34].

MSH2 deficiency has been shown to lead to accelerated telomere shortening in normal human cells[35]. If a reduction of *MSH2* expression to 50% is sufficient to increase the rate of telomere shortening in normal differentiated tissues in vivo, then there would be an increased risk of early onset of telomere dysfunction leading to cellular senescence in heterozygous carriers that may affect tissue architecture and cancer progression[35]. With telomere shortening being an early event in CRC carcinogenesis, this makes *MSH2* pathogenic variant carriers especially vulnerable.

In this study we have used genotypes of three SNPs in *TERT,* located on chromosome 5p15.33; rs2736108 (upstream variant) and rs7705526 (intronic variant) both of which have been associated with longer telomeres and breast cancer[22], and rs2075786 (intronic variant) reported to be associated with shorter telomeres and increased cancer risk in LS[29], using data from four different LS cohorts (two of the cohorts previously described for one of the SNPs in[29]). With additional analysis, we aim to determine if these polymorphisms are associated with the age of cancer onset or cancer risk in this susceptible population. If targeted genetic screening is used to identify patients with further increased risk of developing cancer, more personalized screening strategies may be appropriate to reduce the likelihood of LS patients presenting with cancer.

Materials and methods

The study complies with the ethical considerations from Hunter New England Research Ethics Committee (Australia), University of Newcastle Human Research Ethics Committee (Australia), the ethics committees of the Pomeranian Academy of Medicine (Poland), ethics committee of Institut d'Investigació Biomèdica de Bellvitge (Spain), Leiden University Medical Centre (the Netherlands) and Regional Committees for Medical and Health Research Ethics (Norway)—all experiments were performed in accordance with institutional guidelines and regulations. Written, informed consent was obtained from all participants. A parent or guardian provided informed consent for participants under the age of 18 years of age.

Sample cohort. This study consists of 1971 LS patients carrying pathogenic variants (class 4 and 5) in *MLH1*, *MSH2* or *MSH6* (681 Australian, 396 Polish, 240 Spanish and 654 Dutch) from 716 families, representing one of the largest LS cohort published for modifier genes to date. The Spanish and Dutch genotype data for SNP rs2075786 has previously been published[29], while the two additional SNPs and the Australian and Polish dataset has not. The reason for doing a combined analysis instead of using Australian and Polish samples as a validation cohort is the increased statistical power the larger sample size provides. Note that the statistical power for SNP rs 2075786 is larger than for the two other SNPs due to it being analysed in a larger sample cohort. This study also represents an extension of the previous study with more and deeper analyses. In addition, it has accounted for country in the multi-variable analysis of this study.

Genotyping. The Australian and Polish LS patient samples were genotyped for SNPs (major > minor allele according to GnomAD) in *TERT*; rs2075786 (G>A), rs2736108 (C>T) and rs7705526 (C>A) using TaqMan SNP assays (Applied Biosystems) for the Australian and Polish sample cohorts. Thermo-cycling was undertaken according to the TaqMan SNP Genotyping Assay Protocol, involving; 10 min at 95 degrees; 40 cycles of 15 s at 95 degrees; and 1 min at 60 degrees. Raw data was generated using the 7500 standard real-time PCR system (Applied Biosystems). Raw data was analysed using TaqMan Genotyper Software (Life Sciences, Foster City, CA).

Statistical analysis. Statistical analysis was performed using Stata 12.1 (StataCorp LP, TX USA). Pearson's Chi-square test was used to evaluate deviation from the expected Hardy–Weinberg equilibrium (HWE) and genotype frequency differences between sample cohorts (2×3 contingency tables). We applied Bonferroni correction for multiple testing, resulting in a corrected significance threshold of $p = 0.0167$ (0.05 divided by the 3 SNPs tested).

Variation in age of diagnosis between each SNPs genotype was examined using Kaplan–Meier estimator analysis using Wilcoxon's (Breslow), Log-rank and Tarone-Ware tests to examine homogeneity of the Kaplan–Meier

www.nature.com/scientificreports/

A	Samples n	Average age* years	Gender Male (%) / female (%)	Gene *MLH1* (%) / *MSH2* (%) / *MSH6* (%)
rs2075786 (n = 1895)				
LS cancer	830	45 (16–81)	351 (42)/479 (58)	363 (44) / 342 (41)/124 (15)*
CRC	632	44 (16–79)	316 (50)/316 (50)	315 (50) / 250 (40)/66 (10)
Cancer-free	1065	42 (11–89)	451 (42)/614 (58)	472 (44) / 363 (34)/230 (22)
Total	**1895**			
rs2736108 and rs7705526 (n = 1241)				
LS cancer	633	45 (18–84)	266 (42)/367 (58)	292 (46) / 260 (41) / 81 (13)
	503	44 (18–79)	241 (48)/262 (52)	259 (51) / 199 (40) / 45 (9)
Cancer-free	608	40 (12–81)	246 (40)/362 (60)	303 (50) / 205 (34) / 100 (16)
Total	**1241**			
B	**Sample cohort demographics by country**			
	Samples n	Average age* Years (range)	Gender Male (%) / female (%)	Gene *MLH1* (%)/*MSH2* (%)/*MSH6* (%)
LS cancer AU	298	45 (19–81)	173 (58)/125 (42)	106 (36)/136 (46)/56 (18)
LS cancer PL	189	45 (18–71)	121 (64)/68 (36)	101 (54)/74 (39)/14 (7)
LS cancer ES	146	44 (21–79)	73 (50)/72 (50)	85 (58)/50 (34)/11 (8)
LS cancer NL	197	45 (16–78)	112 (57)/85 (43)	71 (36)/82 (42)/43 (22)
Total	**803**			
CRC AU	218	44 (19–75)	113 (52)/105 (48)	92 (42)/99 (45)/27 (13)
CRC PL	150	44 (18–69)	86 (58)/64 (42)	85 (57)/55 (37)/10 (6)
CRC ES	135	44 (21–79)	63 (47)/71 (53)	82 (61)/45 (33)/8 (6)
CRC NL	129	44 (16–72)	54 (42)/75 (58)	56 (43)/51 (40)/21 (17)
Total	**632**			
Cancer-free AU	321	40 (11–81)	196 (61)/124 (39)	129 (40)/112 (35)/80 (25)
Cancer-free PL	193	39 (13–73)	116 (60)/77 (40)	114 (59)/68 (35)/11 (6)
Cancer-free ES	94	41 (12–82)	50 (53)/44 (47)	60 (64)/25 (26)/9 (10)
Cancer-free NL	457	45 (20–89)	252 (55)/205 (45)	169 (37)/158 (35)/130 (28)
Total	**1065**			

Table 1. Demographic data. (A) Displays demographics for the combined LS cohort (rs2075786 and rs2736108 + rs7705526), while (B) Displays demographics for the four countries separately. * Average age of cancer for cancer patients and average age at last follow up for cancer-free patients.

plots. For the Kaplan–Meier analysis, age of diagnosis of LS cancer or CRC is the endpoint for analysis and individuals free from cancer/polyposis were censored at their age at last follow up.

Risk of cancer was estimated for each SNP by genotypic odds ratio (OR) using multilevel mixed-effects logistic regression taking into account family id (because we have both probands and relatives in the cohort), while adjusting for country, gender and gene. Odds ratios, 95% confidence intervals and p-values are presented using forest plots; model coefficients for each model are presented in the supplementary material. Due to significant findings for SNP rs2075786 and to replicate a previous study[29], additional analysis within *MSH2* carriers was conducted (as described above) but modelling LS cancer diagnosis < 45 years of age, versus those diagnosed after 45 years of age or who were cancer-free with no age restriction.

Results

There were 1971 samples across four countries (654 samples from the Netherlands and 240 samples from Spain, both previously described[29], and 681 samples from Australia and 396 samples from Poland) with enough clinical data to warrant inclusion in the current analysis from which 76 samples failed to genotype for all three *TERT* SNPs (sample cohort of 1895).

Genotyping. Samples from the Netherlands and Spain were originally only genotyped for rs2075786[29], while the Australian and Polish were genotyped for this SNP for the purpose of the current study. A further 14 samples failed genotyping yielding an analysis cohort of 1881 for this SNP. The Spanish, Australian and Polish samples were genotyped for the two additional SNPs; making 1241 samples available for rs2736108, an additional 34 samples failed genotyping yielding 1207 for analysis, and for rs7705526, 40 samples failed genotyping yielding 1201 samples. Demographic data from combined and individual sample cohorts can be seen in Table 1.

Table 2 displays genotype frequencies from sample cohorts and as expected there are somewhat significant differences between the cohorts; Australian/Dutch and Polish/Dutch for SNP rs2075786, and Australian/Spanish for SNPs rs2736108 and rs7705526 (all $p > 0.03$). For SNP rs2736108 the Australian genotype frequency is significantly different from the Polish ($p = 0.001$).

Scientific Reports | (2021) 11:11401 | https://doi.org/10.1038/s41598-021-90501-2   **nature** portfolio   3

TERT SNP	Combined sample cohort Total n (%)	Australian sample cohort Total n (%)	Polish sample cohort Total n (%)	Spanish sample cohort Total n (%)	Netherland sample cohort Total n (%)
rs2075786	1881	608	380	239	654
GG	782 (41.6)	228 (37.5)	155 (40.8)	109 (45.6)	290 (44.3)
GA	852 (45.3)	296 (48.6)	162 (42.6)	102 (42.7)	292 (44.7)
AA	247 (13.1)	84 (13.9)	63 (16.6)	28 (11.7)	72 (11.0)
rs2736108	1207	601	381	225	
CC	629 (52.1)	279 (46.4)	222 (58.3)	128 (56.9)	
CT	475 (39.4)	261 (43.4)	133 (34.9)	81 (36.0)	
TT	103 (8.5)	61 (10.2)	26 (6.8)	16 (7.1)	
rs7705526	1201	601	375	225	
CC	513 (42.7)	280 (46.5)	146 (38.9)	87 (38.7)	
CA	549 (45.7)	263 (43.8)	182 (48.6)	104 (46.2)	
AA	139 (11.6)	58 (9.7)	47 (12.5)	34 (15.1)	

Table 2. Genotype frequencies in the combined sample cohort and for each of the four countries included.

Statistical analysis—risk of LS cancer. Both simple logistic regression and Kaplan–Meier analysis show that the patient cohort is consistent with published literature; carriers of *MSH6* pathogenic variants have a decreased risk of LS cancer compared to *MLH1* pathogenic variant carriers (OR = 0.68, CI = 0.50–0.92, p = 0.014), and later age of onset compared *MLH1* and *MSH2* variants carriers, see supplementary Fig. S1. None of the three *TERT* SNPs demonstrated independent associations with risk of LS cancer or age of diagnosis (i.e. when not accounting for LS mutant gene); observed genotype frequencies, crude genotypic odds ratios and odds ratios are presented in supplementary table 1. There were no deviations from HWE (rs2075786 p = 0.53, rs2736108 p = 0.32, rs7705526 p = 0.67).

Results for the mixed-effects logistic regression investigating the interactions of each SNP with the respective gene, adjusting for confounders, are presented in Fig. 1; depicted are the corresponding odds ratios by each level of gene and genotype with the reference group (*MLH1* homozygous major allele) set at unity. There was weak evidence that effects for rs2075786 and rs2736108 genotypes were different by gene (p for interaction = 0.07 and 0.05, respectively). Model coefficients, confidence intervals and p-values are presented in supplementary tables S2-S4.

For SNP rs2075786 in Fig. 1, it was apparent that within *MLH1* and *MSH6* variant carriers, the genotype risk patterns were similar; the GG genotype confers the greatest risk but it was not statistically significantly different from the other genotypes (95% confidence intervals overlap substantially). Whereas for *MSH2* variant carriers, harbouring the heterozygous genotype was associated with greater risk, and those homozygous for A had the greatest risk of cancer across the cohort.

Within *MSH2* variant carriers, the AA genotype of rs2075786 is associated with 85% higher odds of developing cancer compared to *MSH2* carriers with the GG genotype (Fig. 2, ORs within *MSH2* pathogenic variant carriers only, estimated from the same model as Fig. 1).

A previous study of LS patients with *MSH2* variants, linked SNP rs2075786 to an increased risk of cancer diagnosis younger than 45 years age using logistic regression[34]. When we analysed this outcome we also observed an association (see supplementary Fig. S2); the heterozygous genotype (GA) is linked to the greatest risk of cancer (OR 1.79, 95% CI 1.2 to 2.7, p = 0.005; Model coefficients, confidence intervals and p- values are presented in supplementary table S5). Our Kaplan–Meier analysis also demonstrated that *MSH2* variant carriers who also carry the rs2075786 A allele have earlier onset of cancer (Kaplan–Meier analysis log-rank p = 0.0011, Wilcoxon p = 0.0006 and Tarone-Ware p = 0.0007 see Fig. 3), which was not observed for *MLH1* and *MSH6* variant carriers (log-rank p = 0.3524 and p = 0.3763, respectively).

For rs2736108, the risk patterns observed for *MLH1* variant carriers was similar to *MSH2* variant carriers whereas, the pattern for *MSH6* variant carriers was different (see Fig. 1). Within *MSH6* variant carriers, the C allele confers greater risk than the T allele (Fig. 4 presents the odds ratios for the genotypes). The odds of LS cancer were 48% lower for the CT genotype vs the CC genotype (OR 0.52 95%CI 0.29 to 098) and the odds for the TT genotype were lower again however the confidence interval overlapped one due to the small number of observations with this covariate pattern.

There was a lack of evidence that the risk pattern for genotypes of rs7705526 differed by gene (p for interaction = 0.73).

Discussion

Here we present results from a large sample cohort of 1881 LS patients that statistically show that a polymorphism in *TERT* influences disease risk in LS patients. SNP rs2075786 in *TERT* modifies cancer risk in LS patients with mutations in *MSH2*, the variant AA genotype is associated with 85% higher odds of cancer compared to the wildtype GG genotype.

The current study was undertaken as cancer-affected LS patients with the AA genotype of SNP rs2075786 had been shown to have shorter telomeres than those with GG genotype and the A allele was marginally associated with LS cancer in patients < 45 years of age[29]. In silico analysis predicted the A allele of SNP rs2075786
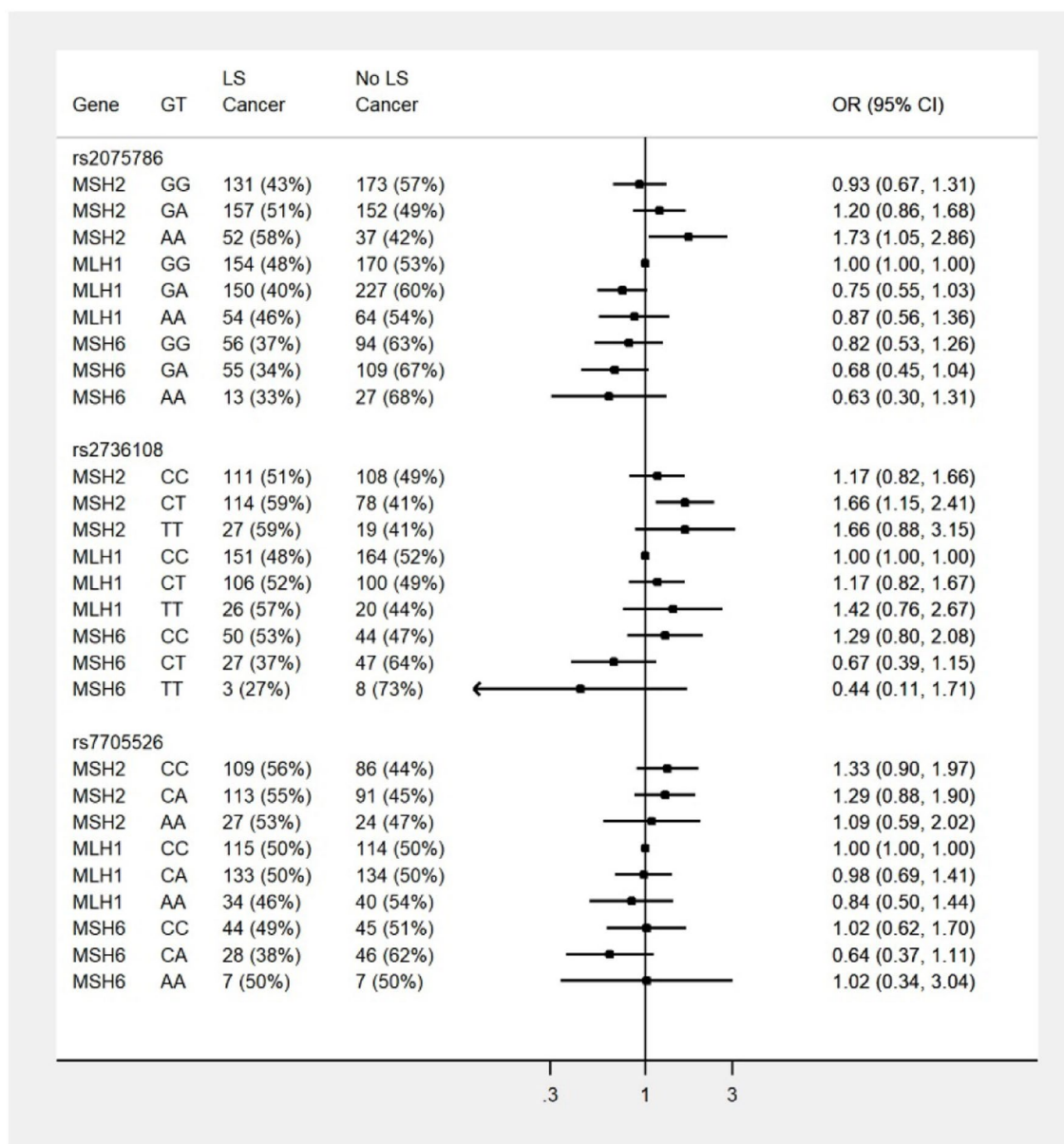
Figure 1. This forest plot displays across cohort, odds ratios for risk of LS cancer by gene and genotype for the three *TERT* SNPs in the current study, the reference group is *MLH1* major genotype. The reference group for each model, is the gene/genotype group set to unity. All other ORs are relative to this reference group.

eliminates a retinoid binding site, causing natural retinoids not to efficiently limit *TERT* expression, culminating in accelerated tumour growth[29]. This finding is consistent with another study that revealed leukocyte telomeres of patients with LS cancer were shorter than those of controls and unaffected LS patients[36], suggesting that shortened telomeres are a result of the disease or an additional risk factor for LS patients. A second LS study reported no evidence of association between *TERT* SNPs and risk of CRC, overall or when stratified by gender and MMR gene after adjustment for multiple testing and censored by age 45 years[37], but differently to the current study they only considered CRC risk not including all LS associated cancers. Another advantage of the current study is the large sample size and the ability to detect smaller effect sizes.

It has been shown that cell lines with variants in MMR genes show telomere instability, with highest mutation frequency in *MSH2* deficient cells[35]. Reduction in *MSH2* expression leads to accelerated telomere shortening in normal cells[35] and *MSH2* deficient cells have been shown to have minor telomere capping effects[38]. *MSH2* is associated with the *TERT* promoter and regulates promoter activity, i.e. knockdown of *MSH2* results in a significant reduction of telomerase activity in human oral squamous cell carcinoma cells[39]. *MSH2* variants lead to accelerated telomere shortening in normal cells (an early event in CRC carcinogenesis) and the A allele of SNP rs2075786 is predicted to cause early telomerase activation (carriers of the AA genotype have shorter telomeres[29]). Individually they might just have subtle inhibitory effect on *TERT* but together they may increase LS patients' risk of cancer development. This can explain why we observe the significantly increased risk of cancer in LS
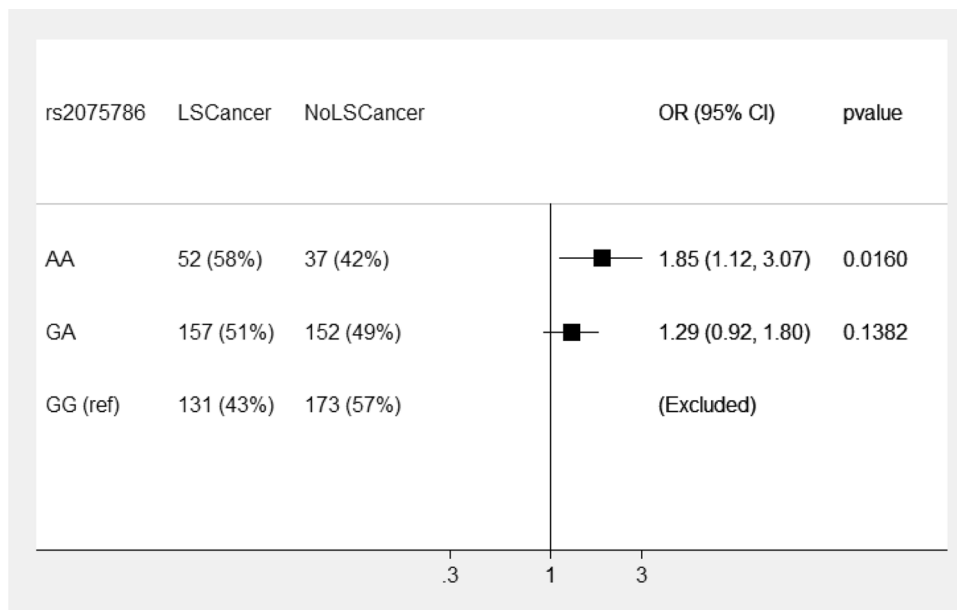
Figure 2. Displays rs2075786 odds ratios for risk of LS cancer in *MSH2* pathogenic variant carriers. The reference is genotype GG (wildtype).
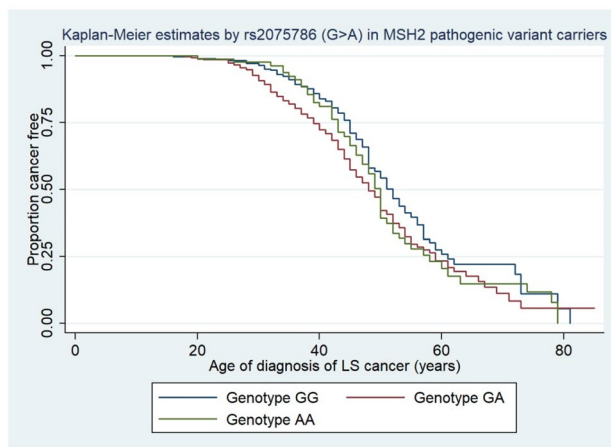

Figure 3. Kaplan–Meier estimated by rs2075786 genotypes in *MSH2* pathogenic variant carriers. The graph shows the effect the genotypes has on age of diagnosis of Lynch Syndrome (LS) cancer in LS patients. A significant difference in the age of diagnosis of LS cancer can be seen between the three genotypes (Log-rank p = 0.0011, Wilcoxon p = 0.0006 and Tarone-Ware p = 0.0007). LS patients over the age of 50 years and carriers of the A allele (GA and AA genotype) will develop LS cancer earlier than LS patients carrying the GG genotype will.

patients compared to LS cancer-free patients and makes SNP rs2075786 a plausible modifier for disease risk in *MSH2* pathogenic variant carriers.

Possible biases in the current study include confounding factors such as lifestyle, smoking and other environmental factors influencing the reported results, however since there was no specific selection for patients these variables are likely to be equally distributed across the patient cohort. Studies on modifier genes in LS are difficult due to all the variables affecting cancer risk, and controversial results have rather been the rule than the exception, but with increased sample sizes we are now hoping to avoid this. Ascertainment bias related to sampling and selection bias (where some members are less likely to be included than others) should not be a problem since there is good representation of both cancer affected and unaffected MMR variant carriers. Our results could aid in explaining the controversial evidence for anticipation in LS[40,41], even though we have not looked into this in the current study, as increased rate of telomere shortening in *MSH2* deficient cells provides a mechanism that may contribute to genetic anticipation in some LS families[35]. We were unable to control for
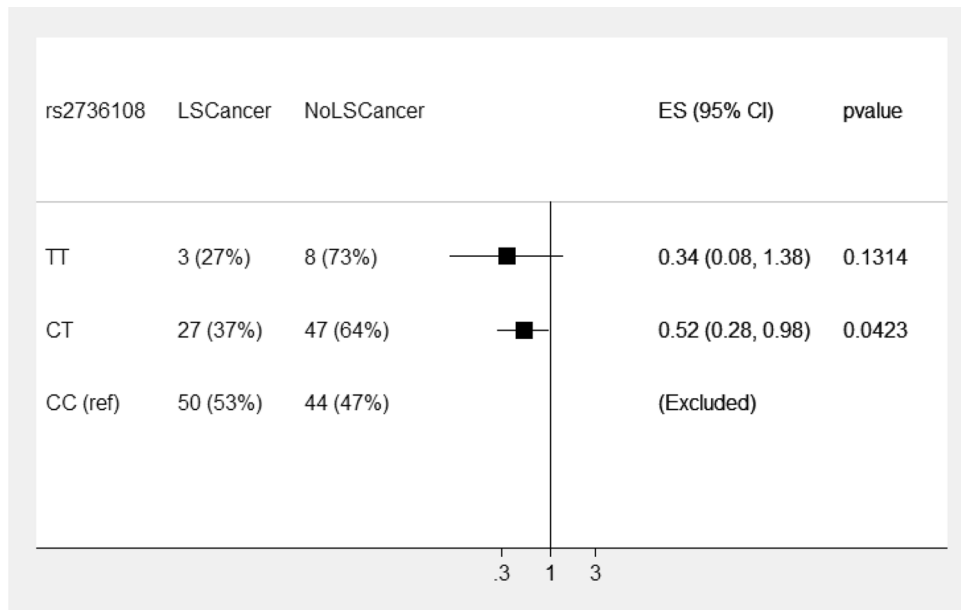
Figure 4. Displays rs2736108 odds ratios for risk of LS cancer in *MSH6* pathogenic variant carriers. The reference is genotype CC.

differences in individual causal germline variants (i.e. frameshift, splice site, etc.) in individual genes as this information was not available for all patients.

In conclusion, we present a polymorphism in *TERT* to be a possible modifier of disease risk in *MSH2* pathogenic variant carriers. The rs2075786 SNP in *TERT* is associated with a differential risk of developing cancer for *MSH2* pathogenic variant carriers. By including this SNP in future risk algorithms, it should be possible to tailor surveillance options for individual patients. Use of this information has the potential to personalize screening protocols for LS patients.

References

1. Sijmons, R. H. & Hofstra, R. M. Review: Clinical aspects of hereditary DNA mismatch repair gene mutations. *DNA Repair (Amst).* **38**, 155–162. https://doi.org/10.1016/j.dnarep.2015.11.018 (2016).
2. Miyaki, M. *et al.* Germline mutation of MSH6 as the cause of hereditary nonpolyposis colorectal cancer. *Nat. Genet.* **17**, 271–272 (1997).
3. Ligtenberg, M. J. *et al.* Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3' exons of TACSTD1. *Nat. Genet.* **41**, 112–117. https://doi.org/10.1038/ng.283 (2009).
4. Fishel, R. *et al.* The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell* **75**, 1027–1038 (1993).
5. Bronner, C. E. *et al.* Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. *Nature* **368**, 258–261 (1994).
6. Nicolaides, N. C. *et al.* Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. *Nature* **371**, 75–80 (1994).
7. Talseth-Palmer, B. & Scott, R. Genetic variation and its role in malignancy. *Int. J. Biomed. Sci.* **7**, 158–171 (2011).
8. Church, J. & Kravochuck, S. Variation in lynch syndrome. *Dis. Colon. Rectum.* **58**, e77. https://doi.org/10.1097/DCR.0000000000000355 (2015).
9. Talseth-Palmer, B. A. *et al.* Combined analysis of three Lynch syndrome cohorts confirms the modifying effects of 8q233 and 11q231 in MLH1 mutation carriers. *Int. J. Cancer.* **132**, 1556–1564. https://doi.org/10.1002/ijc.27843 (2013).
10. Houlle, S. *et al.* Evaluation of Lynch syndrome modifier genes in 748 MMR mutation carriers. *Eur. J. Hum. Genet.* **2**, 2 (2011).
11. Talseth, B. A. *et al.* Age of diagnosis of colorectal cancer in HNPCC patients is more complex than that predicted by R72P polymorphism in TP53. *Int. J. Cancer.* **118**, 2479–2484 (2006).
12. Jones, J. S. *et al.* p53 polymorphism and age of onset of hereditary nonpolyposis colorectal cancer in a Caucasian population. *Clin. Cancer Res.* **10**, 5845–5849 (2004).
13. Kurenova, E. V. & Mason, J. M. Telomere functions. A review. *Biochemistry (Mosc).* **62**, 1242–1253 (1997).
14. Bertorelle, R., Rampazzo, E., Pucciarelli, S., Nitti, D. & De Rossi, A. Telomeres, telomerase and colorectal cancer. *World J. Gastroenterol.* **20**, 1940–1950. https://doi.org/10.3748/wjg.v20.i8.1940 (2014).
15. Stampfer, M. R. & Yaswen, P. Human epithelial cell immortalization as a step in carcinogenesis. *Cancer Lett.* **194**, 199–208 (2003).
16. Counter, C. M. *et al.* Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. *EMBO J.* **11**, 1921–1929 (1992).
17. Ducrest, A. L., Szutorisz, H., Lingner, J. & Nabholz, M. Regulation of the human telomerase reverse transcriptase gene. *Oncogene* **21**, 541–552. https://doi.org/10.1038/sj.onc.1205081 (2002).
18. Mathieu, N., Pirzio, L., Freulet-Marriere, M. A., Desmaze, C. & Sabatier, L. Telomeres and chromosomal instability. *Cell Mol. Life Sci.* **61**, 641–656. https://doi.org/10.1007/s00018-003-3296-0 (2004).

19. Collins, K. & Mitchell, J. R. Telomerase in the human organism. *Oncogene* **21**, 564–579. https://doi.org/10.1038/sj.onc.1205083 (2002).
20. Nersisyan, L. H., Loeffler-Wirth, H., Galle, J. L. & Arakelkyan, A. B. Telomere length maintenance and its transcriptional regulation in lynch syndrome and sporadic colorectal carcinoma. *Front. Oncol.* https://doi.org/10.3389/fonc.2019.01172 (2019).
21. HosgoodCawthon, H. D. R., He, X., Chanock, S. & Lan, Q. Genetic variation in telomere maintenance genes, telomere length, and lung cancer susceptibility. *Lung Cancer* **66**, 157–161. https://doi.org/10.1016/j.lungcan.2009.02.005 (2009).
22. Bojesen, S. E. *et al.* Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer. *Nat. Genet.* **45**, 371–384. https://doi.org/10.1038/ng.2566 (2013).
23. Kote-Jarai, Z. *et al.* Fine-mapping identifies multiple prostate cancer risk loci at 5p15, one of which associates with TERT expression. *Hum. Mol. Genet.* **22**, 2520–2528. https://doi.org/10.1093/hmg/ddt086 (2013).
24. Carvajal-Carmona, L. G. *et al.* Candidate locus analysis of the TERT-CLPTM1L cancer risk region on chromosome 5p15 identifies multiple independent variants associated with endometrial cancer risk. *Hum. Genet.* **134**, 231–245. https://doi.org/10.1007/s00439-014-1515-4 (2015).
25. Grotenhuis, A. J. *et al.* Independent replication of published germline polymorphisms associated with urinary bladder cancer prognosis and treatment response. *Bladder Cancer.* **2**, 77–89. https://doi.org/10.3233/BLC-150027 (2016).
26. Kinnersley, B. *et al.* The TERT variant rs2736100 is associated with colorectal cancer risk. *Br. J. Cancer.* **107**, 1001–1008. https://doi.org/10.1038/bjc.2012.329 (2012).
27. Peters, U. *et al.* Meta-analysis of new genome-wide association studies of colorectal cancer risk. *Hum. Genet.* **131**, 217–234. https://doi.org/10.1007/s00439-011-1055-0 (2012).
28. Hofer, P. *et al.* Association of genetic variants of human telomerase with colorectal polyps and colorectal cancer risk. *Mol. Carcinog.* **51**(Suppl 1), E176–E182. https://doi.org/10.1002/mc.21911 (2012).
29. Bellido, F. *et al.* Genetic variant in the telomerase gene modifies cancer risk in Lynch syndrome. *Eur. J. Hum. Genet.* **21**, 511–516. https://doi.org/10.1038/ejhg.2012.204 (2013).
30. Huang, F. W. *et al.* Highly recurrent TERT promoter mutations in human melanoma. *Science* **339**, 957–959. https://doi.org/10.1126/science.1229259 (2013).
31. Horn, S. *et al.* TERT promoter mutations in familial and sporadic melanoma. *Science* **339**, 959–961. https://doi.org/10.1126/science.1230062 (2013).
32. Shen, J. *et al.* Genetic polymorphisms in telomere pathway genes, telomere length, and breast cancer survival. *Breast Cancer Res. Treat.* **134**, 393–400. https://doi.org/10.1007/s10549-012-2058-9 (2012).
33. Do, S. K. *et al.* Replication of the results of genome-wide and candidate gene association studies on telomere length in a Korean population. *Korean J. Intern. Med.* **30**, 719–726. https://doi.org/10.3904/kjim.2015.30.5.719 (2015).
34. Beesley, J. *et al.* Functional polymorphisms in the TERT promoter are associated with risk of serous epithelial ovarian and breast cancers. *PLoS ONE* **6**, e24987. https://doi.org/10.1371/journal.pone.0024987 (2011).
35. Mendez-Bermudez, A. & Royle, N. J. Deficiency in DNA mismatch repair increases the rate of telomere shortening in normal human cells. *Hum. Mutat.* **32**, 939–946. https://doi.org/10.1002/humu.21522 (2011).
36. Segui, N. *et al.* Telomere length and genetic anticipation in Lynch syndrome. *PLoS ONE* **8**, e61286. https://doi.org/10.1371/journal.pone.0061286 (2013).
37. Win, A. K. *et al.* Genetic variants within the hTERT gene and the risk of colorectal cancer in Lynch syndrome. *Genes Cancer.* **6**, 445–451. https://doi.org/10.18632/genesandcancer.85 (2015).
38. Campbell, M. R., Wang, Y., Andrew, S. E. & Liu, Y. Msh2 deficiency leads to chromosomal abnormalities, centrosome amplification, and telomere capping defect. *Oncogene* **25**, 2531–2536. https://doi.org/10.1038/sj.onc.1209277 (2006).
39. Kang, X., Chen, W., Kim, R. H., Kang, M. K. & Park, N. H. Regulation of the hTERT promoter activity by MSH2, the hnRNPs K and D, and GRHL2 in human oral squamous cell carcinoma cells. *Oncogene* **28**, 565–574. https://doi.org/10.1038/onc.2008.404 (2009).
40. Stella, A. *et al.* Germline novel MSH2 deletions and a founder MSH2 deletion associated with anticipation effects in HNPCC. *Clin. Genet.* **71**, 130–139. https://doi.org/10.1111/j.1399-0004.2007.00745.x (2007).
41. Nilbert, M., Timshel, S., Bernstein, I. & Larsen, K. Role for genetic anticipation in Lynch syndrome. *J. Clin. Oncol.* **27**, 360–364. https://doi.org/10.1200/JCO.2008.16.1281 (2009).

Acknowledgements

Author contributions

S.B. and K.A.B. carried out the molecular genetic studies, T.J.E. provided statistical analysis, and guidance. D.D., S.J.C., G.C., G.K., J.T.W., L.V., H.F.A.V., J.L. and R.J.S. provided samples, input into study design and/or helped in the interpretation of data. MUW participated in the interpretation of data and drafting the manuscript. B.T.P. participated in the design of the study and coordination, performed statistical analysis, interpretation of data and drafted the manuscript. All authors read and approved the final manuscript.

Funding

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-90501-2.

Correspondence and requests for materials should be addressed to B.A.T.-P.

Reprints and permissions information is available at www.nature.com/reprints.