Check for updates

A common genetic variation in GZMB may associate with cancer risk in patients with Lynch syndrome

Vince Kornél Grolmusz[1,2,3,4]*, Petra Nagy[1], István Likó[2,4], Henriett Butz[1,2,3,4,5], Tímea Pócza[1], Anikó Bozsik[1,2,4], János Papp[1,2,4], Edit Oláh[1] and Attila Patócs[1,2,3,4]*

[1]Department of Molecular Genetics, National Institute of Oncology, Budapest, Hungary, [2]Hereditary Cancers Research Group, Eötvös Loránd Research Network − Semmelweis University, Budapest, Hungary, [3]Department of Laboratory Medicine, Semmelweis University, Budapest, Hungary, [4]National Tumorbiology Laboratory, National Institute of Oncology, Budapest, Hungary, [5]National Oncology Biobank Center, National Institute of Oncology, Budapest, Hungary

Lynch syndrome (LS), also known as hereditary nonpolyposis colorectal cancer syndrome (HNPCC) is a common genetic predisposition to cancer due to germline mutations in genes affecting DNA mismatch repair. Due to mismatch repair deficiency, developing tumors are characterized by microsatellite instability (MSI-H), high frequency of expressed neoantigens and good clinical response to immune checkpoint inhibitors. Granzyme B (GrB) is the most abundant serine protease in the granules of cytotoxic T-cells and natural killer cells, mediating anti-tumor immunity. However, recent results confirm a diverse range of physiological functions of GrB including that in extracellular matrix remodelling, inflammation and fibrosis. In the present study, our aim was to investigate whether a frequent genetic variation of *GZMB*, the gene encoding GrB, constituted by three missense single nucleotide polymorphisms (rs2236338, rs11539752 and rs8192917) has any association with cancer risk in individuals with LS. *In silico* analysis and genotype calls from whole exome sequencing data in the Hungarian population confirmed that these SNPs are closely linked. Genotyping results of rs8192917 on a cohort of 145 individuals with LS demonstrated an association of the CC genotype with lower cancer risk. *In silico* prediction proposed likely GrB cleavage sites in a high proportion of shared neontigens in MSI-H tumors. Our results propose the CC genotype of rs8192917 as a potential disease-modifying genetic factor in LS.

Introduction

Cancer neoantigens are *de novo* amino acid sequences produced by cancer cells provoking antitumor immune response (1). As multiple layers of evidence support the clinically effective modification of the host's immune system in fighting neoplastic diseases mainly through pharmacologic inhibition of immune checkpoints (2, 3), cutting-edge approaches in oncoimmunology aim to stimulate immune responses by cancer vaccines (4) and *in vitro* modification of effector T cells (5). Although most cancer neoantigens are unique to one's cancer, several cancers, especially those characterized by microsatellite instability (MSI-H) share several recurrent neoantigens (6, 7), providing a rationale for off-the-shelf cancer vaccines not only in the adjuvant/metastatic setting (8), but also as a part of primary prevention in high-risk individuals (9, 10).

Although MSI-H cancers are sporadic tumors in the majority of the cases, Lynch syndrome (LS), a frequent cancer predisposition syndrome with a prevalence of 1:250-500 significantly elevates the risk of developing MSI-H neoplasms (11). LS is caused by germline pathogenic mutations in genes disrupting the optimal function of the DNA mismatch repair machinery (*MLH1*, *MSH2*, *MSH6*, *PMS2*, *EPCAM*), and is inherited in an autosomal dominant manner. Frequent manifestations include colorectal cancer (CRC) and endometrial cancer, although it mildly elevates the risk of a spectrum of malignancies (12). Specific cancer risks vary by the gene concerned; mutations in *MLH1*, *MSH2* and *EPCAM* result in similarly higher risks for CRC (~50% lifetime risk) and endometrial cancer (~50% lifetime risk) (12). Previous studies have demonstrated high frequencies of effective immunity against the shared neoantigens formerly described in MSI-H tumors not only in LS-associated cancer patients but also in healthy individuals living with LS (13, 14). Optimal immunosurveillance acquires a key part in the host's prevention of carcinogenesis (15). In healthy patients with LS, normal colonic mucosa exhibits higher frequencies of a wide range of immune cell populations, while this is reduced in LS-associated cancer (16). In colonic premalignant lesions, a higher amount of lymphocyte –activation gene 3 immune checkpoint expression facilitates immune evasion and carcinogenesis (17). Immune-tumor interactions, and immunoediting in particular, are dynamically changing after malignant transformation, the most important of which is the frequent loss of β_2 microglobulin expression, resulting in impaired HLA class I antigen expression (15, 18–20).

Granzymes are serine proteases, which are the main components of the granules of cytotoxic T-cells and natural killer (NK) cells eliciting perforin-mediated apoptosis of target cells (21). This mechanism is a key effector element in antimicrobial and antitumor immune responses. Granzyme B (GrB) is the most abundant granzyme present in cytotoxic granules, however, recent studies have uncovered several additional molecular mechanisms, in which GrB maintains significant roles (22). In particular, by the expression of GrB in a wide variety of normal epithelial cells and cancer cells, GrB alters extracellular matrix remodeling, epithelial-to-mesenchymal transition and fibrosis (22).

GrB is a 33 kDa protein, which is encoded by the gene *GZMB* in humans. Former genetic studies have confirmed three closely linked common single nucleotide polymorphisms (SNPs) all resulting in

missense mutations (rs8192917 Q48R, rs11539752 P88A and rs2236338 Y245H) (Figure 1A) (26, 27). An initial study proposed that Granzyme B harbouring the minor RAH haplotype (resulting from the minor alleles of these SNPs) is incapable of inducing apoptosis (26), however a follow-up study leveraging several layers of evidence including enzyme activity assays dismissed this possibility (27). Nevertheless, there is still no clear evidence on how these three missense alterations, relatively far from the catalytic triad might affect the function of the enzyme.

In the present study, we aimed to analyze the association between rs8192917, a tagging SNP of the RAH haplotype and cancer risk in high-risk individuals with LS. Additionally, based on our *in silico* analysis we anticipate that a significant portion of shared neoantigens in MSI-H cancers can be cleaved by GrB. This study strengthens the immune-related pathogenetic contribution to LS-associated tumorigenesis and invites further investigations in independent LS cohorts to validate the observed disease-modifying nature of rs8192917.

Materials and methods

Subjects, DNA isolation and genotyping

145 individuals harboring germline pathogenic mutations of *MLH1*, *MSH2* or terminal deletions of *EPCAM* were consecutively enrolled between 1994 and 2021. Baseline characteristics of study subjects are presented in Table 1. Following genetic counseling and written informed consent, DNA was isolated from peripheral blood using the Gentra Puregene Blood Kit (Qiagen, Cat No.: 158389). Mutation analysis was performed using methods described earlier (28, 29) including Sanger sequencing and multiplex ligation-dependent probe amplification (MLPA). The study was approved by the Scientific and Research Ethics Committee of the Medical Research Council of Hungary (ETT-TUKEB 53720-7/2019/EÜIG).

For the genotyping of rs8192917, a predesigned TaqMan Allelic Discrimination assay was used (assay ID: C:_2815152_20). Quantitative real-time PCR was performed on an Applied Biosystems 7500 Fast PCR instrument according to the manufacturer's instructions.

Linkage disequilibrium analysis of rs2236338, rs11539752 and rs8192917

To investigate if a tagging SNP from rs2236338, rs11539752 and rs8192917 can be selected we analyzed the linkage disequilibrium of these SNPs. Firstly, LDMatrix online tool (LDLink, version 5.3.3) was used to assess the linkage disequilibrium of these SNPs in various populations (30).

Additionally, to investigate linkage disequilibrium in the same population from which our LS cohort was selected, we leveraged whole exome sequencing (WES) data from our center obtained between 2017-2022. Patients included in this cohort were advised to have their germline DNA subjected to clinical WES based on their individual and family cancer history and the lack of germline pathogenic mutation detection in prior targeted genetic testing. As patients throughout the entire country were included in this database following the recommendations of national centers of
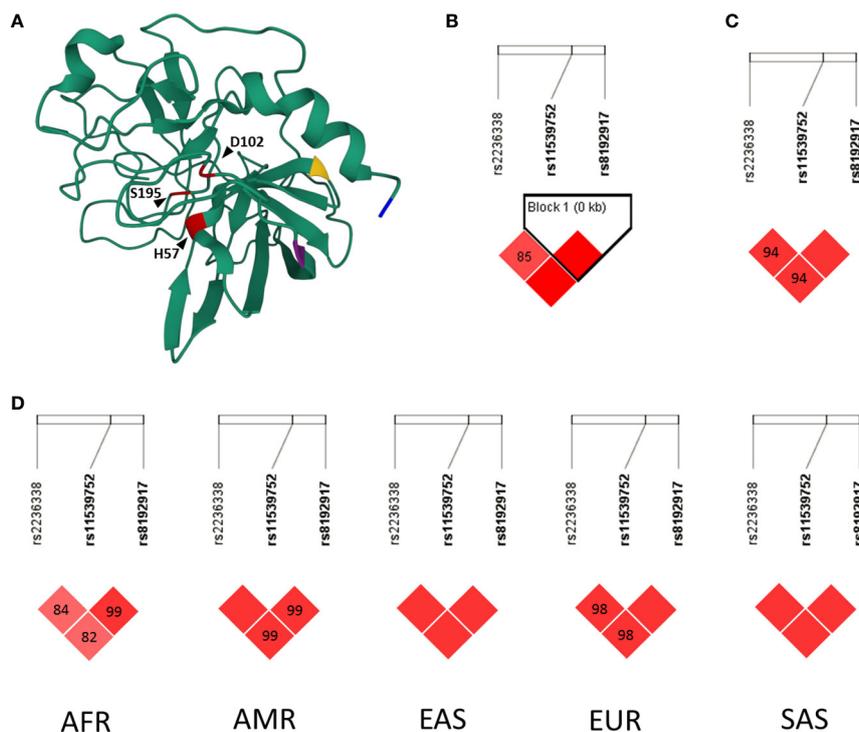
FIGURE 1

Structural representation and linkage disequilibrium analyses of rs2236338, rs11539752 and rs8192917 of GZMB Panel **(A)**: 3D visualization of the crystallographic structure of Granzyme **(B)** Members of the catalytic triad are highlighted in red (H57, D102, S195). Amino acids altered by rs8192917, rs11539752 and rs2236338 are highlighted in yellow (Q48R), purple (P88A) and blue (Y245H), respectively. 3D protein structure (1FQ3) was deposited to Protein Data Bank (rcsb.org) by Estebanez-Perpina et al. (23) and was visualized using Mol* Viewer (24). Panels **(B−D)**: Linkage disequilibrium (LD) analyses of SNPs rs2236338, rs11539752 and rs8192917. Haploview 4.2 was used for LD analysis of the Hungarian WES cohort (Panel **B**). LDMatrix online tool of the LDLink package was used to analyze LD in the CEU (Utah residents from North and West Europe) population (Panel **C**) and in additional populations (Panel **D**). The D' values of each SNP pair are indicated within the squares in the matrices. Solid red squares indicate D' value of 100%. D' values characterize the extent two which two alleles are nonrandomly associated (25). A D' value of 100% between two SNPs corresponds to fully dependent inheritance, while a D' value of 0% signals that the inheritance of the two SNPs are statistically independent from each other. AF, African; AMR, Ad Mixed American; EAS, East Asian; EUR, European; SAS, South Asian.

excellence, this database can be considered as representative of the current Hungarian population.

WES data of 276 independent individuals were included in this anonymized Hungarian germline sequencing database. WES was performed using standard methods. Sequencing data analysis was performed following the Genome Analysis Toolkit (GATK) best practices guidelines. Paired-end sequencing data were obtained in FASTQ file format and reads were trimmed using Cutadapt to remove adapters and bases where the PHRED quality value was less than 20. The trimmed reads were aligned to Genome Reference Consortium Human Build hg19 using Burrows-Wheeler Aligner (bwa-mem2-2.0) (31). Picard tools were used to sort, marking duplicates and index reads. Base Quality Score Recalibration (BQSR) was performed using (GATK) (32). Variant discovery was performed in two steps: single-sample variant calling was performed using HaplotypeCaller in GATK; this was followed by GenotypeGVCFs to combine variants from single-sample gVCFs to the multi-sample VCF. Variant annotations were executed using Funcotator. Genotype calls of rs2236338, rs11539752 and rs8192917 were performed as earlier reported (33). Briefly, allele ratios of 0-0.1 and 0.90-1 were considered homozygous, 0.3-0.7 were considered heterozygous, while allele ratios of 0.1-0.3 and 0.7-0.9 were uncalled. The minimum sequencing depth for homozygous

calls was 5 reads/allele, and for heterozygous calls was 10 reads/allele. Genotype calls were further analyzed using Haploview 4.2 software (34).

In silico analysis of Granzyme B cleavage sites in shared frameshift neoantigens

The ability of Granzyme B to cleave shared neoantigens present in MSI-H tumors was analyzed by the PROSPERous online tool (35). Briefly, shared neoantigens which are developed as frameshift peptides due to repeated mismatch repair deficiency provoking an immune response in independent patients with MSI-H tumors were selected based on the prior study of Ruodko et al. (7). The amino acid sequences of the selected neoantigens were subjected to the PROSPERous prediction algorithm using the P4-P2' cleavage site and logistic regression options.

Statistical analysis

Statistical analysis was performed using GraphPad Prism 9.0 software. For analyzing the correlation between cancer occurrence

TABLE 1   Baseline characteristics of individuals with LS involved in this study.

n	145
Sex, female (%)	70 (48.3%)
age (years)	42.9 ± 15.2
LS-associated tumor occurrence (%)	110 (75.9%)
age at LS-associated tumor occurrence (years)	39.5 ± 9.4
CRC (% of total)	98 (67.6%)
age at CRC occurrence (years)	40.1 ± 10.9
endometrial cancer (% of women)	28 (40%)
age at endometrial cancer occurrence (years)	47.0 ± 6.5
ovarian cancer (% of women)	1 (1.4%)
age at ovarian cancer occurrence (years)	37.9
cancer of the small intestine (% of total)	3 (2.1%)
age at small intestine cancer occurrence (years)	55.0 ± 8.1
pancreatic cancer (% of total)	2 (1.4%)
age at pancreatic cancer occurrence (years)	45.2 ± 1.9
gastric cancer (% of total)	4 (2.8%)
age at gastric cancer occurrence (years)	50.3 ± 14.5
cancer of the bile ducts (% of total)	1 (0.7%)
age at bile duct cancer occurrence (years)	61.1
cancer of the urinary tract (% of total)	3 (2.1%)
age at urinary cancer occurrence (years)	58.2 ± 8.5
keratoacanthomas of the skin (% of total)	7 (4.8%)
age at skin keratoacanthoma occurrence (years)	45.4 ± 20.0
brain tumor (% of total)	0 (0%)

If multiple cancers in the same organ occurred in the same patient, the age at the first occurrence was included. Data are presented as n (% of total/female subjects) or mean ± standard deviation.

and SNP status log-rank test, χ^2 and Fisher's exact tests were used. P-values < 0.05 were considered statistically significant. *Post-hoc* power calculations were performed on significant differences observed by Fisher's exact test to assess the role of sample size. Statistical power > 80% was considered to be strong statistical power.

Results

SNPs rs2236338, rs11539752 and rs8192917 are in linkage disequilibrium in the investigated Central European population

We performed a linkage disequilibrium analysis of SNPs rs2236338, rs11539752 and rs8192917. First, we analyzed an anonymized exome sequencing database of Hungarian patients

with various medical histories. This cohort includes patients from the whole country and therefore is representative of the current Hungarian population. Within this cohort, the three SNPs were found to be in linkage disequilibrium (Figure 1B). Additional *in silico* analysis in the LDmatrix database verified the universal nature of this association in various populations (Figures 1C, D).

Since the results indicated an extremely strong linkage disequilibrium between the analyzed three SNPs, we selected rs8192917 as the tagging SNP for this haplotype for further analysis based on the fact that extensive previous research initiatives analyzed its' association with various conditions and diseases (36–42).

The minor allele of rs8192917 correlates with delayed cancer occurrence in patients with LS

Our LS cohort included 63 (43.4%), 61 (42.1%) and 21 (14.5%) individuals with germline pathogenic variants in *MLH1*, *MSH2* and *EPCAM* genes, respectively. This distribution correlates well with data from the Prospective Lynch syndrome Database (PLSD), where data from 2607 (51.1%) *MLH1* and 2495 (48.9%) *MSH2* (*EPCAM* carriers were included in the *MSH2* cohort) carriers were analyzed (12).

As expected, the most frequent manifestation of LS was CRC followed by endometrial cancer in women, with more than 75% of the cohort exhibiting any manifestation of LS (Table 1). This LS-associated cancer risk corresponds well with data from the PLSD (70-80%), while our cohort contains a larger percentage of CRC cases (67% vs. 50%) possibly due to referral bias (12).

Genotyping of rs8192917 in our LS cohort revealed 85 individuals with TT, 49 individuals with CT and 11 individuals with CC genotype resulting in a minor allele frequency (MAF) of 0.24. This MAF is similar to the MAF of 0.23 observed in the European (non-Finnish) population of the gnomAD database (version 2.1.1) (43).

LS individuals harboring the CC genotype were less likely to develop CRC, the main manifestation of the syndrome (Figures 2A, B, Table 2A). Moreover, individuals homozygous for the minor allele 'C' were less likely to develop LS-associated tumor manifestations (Figures 2C, D, Table 2B). Since LS-associated tumor risks differ in women vs. men mainly due to the risk of endometrial cancer, we performed subgroup analyses based on sex as well (Supplementary Figure 1, Tables 2C, D) and have found that rs8192917-dependent differences are maintained in the female cohort. Statistical power calculations revealed the associations found in the total cohort to be mild, however statistical power was found to be strong when analyzing LS-associated tumor occurrence restricted to women (Table 2). Additional calculations restricted to LS-associated tumor manifestations without CRC or restricted to endometrial cancer occurrence in women have found no associations, probably due to the low sample sizes in these subgroups (Supplementary Table 1).
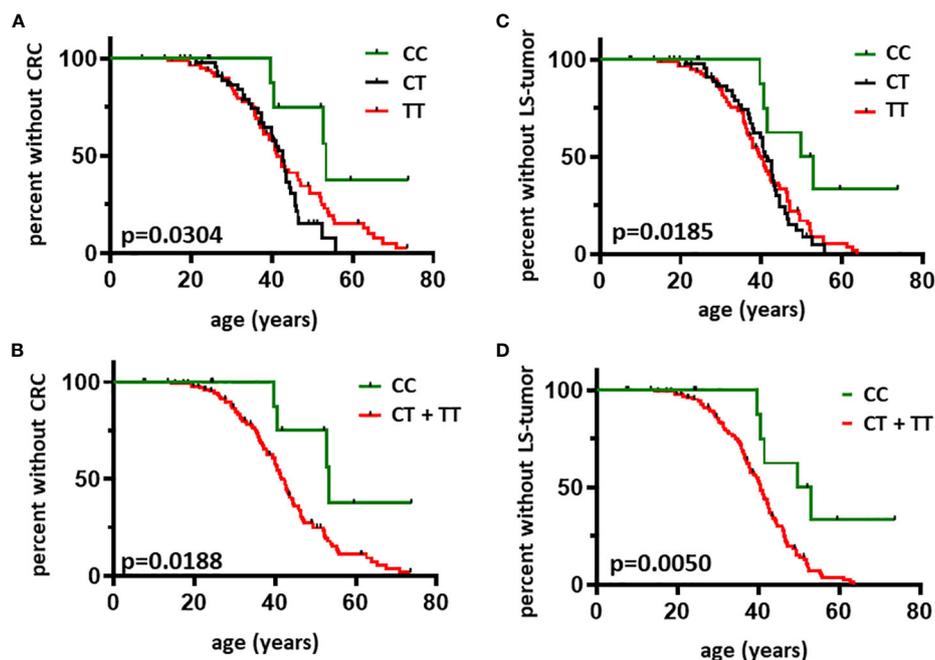
FIGURE 2
Age-related occurrence of CRC and LS-associated tumors in individuals with LS according to rs8192917 genotype Kaplan-Meier curves were plotted to visualize the first CRC diagnosis based on CC vs. CT vs. TT genotypes (Panel **A**) and CC vs. CT + TT genotypes (Panel **B**). Kaplan-Meier curves were plotted to visualize the first LS-associated tumor diagnosis based on CC vs. CT vs. TT genotypes (Panel **C**) and CC vs. CT + TT genotypes (Panel **D**). LS-associated tumors included malignant tumors of the gastrointestinal tract (CRC, gastric cancer, pancreatic cancer, cancer of the small intestine and of the bile ducts), endometrial cancer, ovarian cancer, malignant tumors of the urinary tract and keratoacanthomas of the skin. Curves were compared by log-rank test. n = 145. P-values < 0.05 were considered statistically significant.

Granzyme B is predicted to cleave multiple shared neoantigens in MSI-H cancers

Based on the study of Roudko et al, 46 neoantigens shared between MSI-H tumors were included in the *in silico* analysis (7) to assess probable cleavage by Granzyme B (35). Upon the scores of each cleavage sites, putative loci were classified as likely (score>5.0, n = 27) and unlikely (score<1.0, n = 1575) cleavage sites (Figure 3A). Applying this threshold, 19 out of 46 (41.3%) shared neoantigens harboured likely cleavage sites (Figure 3B, Supplementary Table 2).

Discussion

Identifying disease-modifying genetic factors has tremendous potential in personalizing screening protocols and treatment options, especially in individuals living with high risk for cancer development. Primary studies in individuals living with hereditary predisposition to breast and ovarian cancer found multiple SNPs which may alter cancer risk (44, 45). Based on these results recent studies started to investigate the predictive ability of polygenic risk scores for cancer risk prediction in women living with germline mutations in *BRCA1* and *BRCA2* (46, 47). In LS, earlier studies have found some SNPs which can affect personalized risk of cancer (48–51). In particular, two studies confirmed the association of genomic regions 8q23.3 and 11q23.1 with CRC risk in individuals with LS

(48, 51). A recent study found an association between rs2075786, a SNP of the gene encoding telomerase reverse transcriptase (*TERT*) and cancer risk in individuals with germline *MSH2* mutation in a large international cohort (50). Nevertheless, studies leveraging these results to construct polygenic risk scores in this cohort are still scarce. However, a pioneering study in individuals with germline *PMS2* mutations provided a rationale to investigate this matter more thoroughly (49). In anticipation of these studies we aimed to investigate if immunogenetic factors, specifically genetic variants of GrB might associate with cancer risk in LS.

Rs2236338, rs1159752 and rs8192917 are three functionally active SNPs resulting in amino acid change. Jeong et al. previously showed that these three SNPs constitute a haploblock in the Korean population (36). Similarly to this result, by the analysis of the Central European population in the LDMatrix online tool and also in our Hungarian cohort of 276 independent patients, where WES was performed, we found that these three SNPs are closely linked together. Based on these results we selected rs8192917 as a tagging SNP of this haplotype for further genotyping in our LS cohort.

The CC genotype of rs8192917 was associated with decreased risk for LS-associated tumors in our cohort. This association was significant when the analysis was restricted to females, while in males a tendency toward this association was verified. The association of the CC genotype with decreased cancer occurrence was also validated when the analysis was restricted to CRC, but not to endometrial cancer and other, less frequent LS manifestations, possibly due to the insufficient sample size in this regard. Earlier, the

TABLE 2 Analysis of frequencies of CRC and LS-associated tumors in LS individuals based on their rs8192917 genotype.

A total cohort (n = 145)				
genotype	CRC diagnosis	no CRC diagnosis	p-value	*post-hoc* power
CC	4 (36.4%)	7 (63.6%)		
CT	35 (71.4%)	14 (28.6%)		
TT	59 (69.4%)	26 (30.6%)	0.0688	
CC	4 (36.4%)	7 (63.6%)		
CT + TT	94 (70.1%)	40 (29.9%)	**0.0390**	62.8%
B total cohort (n = 145)				
genotype	LS-tumor diagnosis	no LS-tumor diagnosis	p-value	*post-hoc* power
CC	5 (45.5%)	6 (54.5%)		
CT	38 (77.6%)	11 (22.4%)		
TT	67 (78.8%)	18 (21.2%)	**0.0489**	
CC	5 (45.5%)	6 (54.5%)		
CT + TT	105 (78.4%)	29 (21.6%)	**0.0240**	66.6%
C male cohort (n = 75)				
genotype	LS-tumor diagnosis	no LS-tumor diagnosis	p-value	*post-hoc* power
CC	3 (75.0%)	1 (25.0%)		
CT	21 (70.0%)	9 (30.0%)		
TT	30 (73.2%)	11 (26.8%)	0.9487	
CC	3 (75.0%)	1 (25.0%)		
CT + TT	51 (71.8%)	20 (28.2%)	>0.9999	
D female cohort (n = 70)				
genotype	LS-tumor diagnosis	no LS-tumor diagnosis	p-value	*post-hoc* power
CC	2 (28.6%)	5 (71.4%)		
CT	17 (89.5%)	2 (10.5%)		
TT	37 (84.1%)	7 (15.9%)	**0.0014**	
CC	2 (28.6%)	5 (71.4%)		
CT + TT	54 (85.7%)	9 (14.3%)	**0.0027**	92.9%

Panel **(A)** demonstrates individuals with and without CRC diagnosis (n = 145), while Panel **(B–D)** demonstrates individuals with LS-associated tumor diagnosis in the total cohort (Panel **B**, n = 145), in males (Panel **C**, n = 75) and females (Panel **D**, n = 70). LS-associated tumors included malignant tumors of the gastrointestinal tract (CRC, gastric cancer, pancreatic cancer, cancer of the small intestine and of the bile ducts), endometrial cancer, ovarian cancer, malignant tumors of the urinary tract and keratoacanthomas of the skin. 3X2 contingency tables (CC vs. CT vs. TT genotypes) were analyzed by χ^2 tests, while 2X2 contingency tables (CC vs. CT + TT genotypes) were analyzed by Fisher's exact tests. P-values < 0.05 were considered statistically significant. In the case of Fisher's exact tests, *post-hoc* power calculation has additionally been performed. Statistical power > 80% was considered strong.

'C' allele of rs8192917 has been associated with vitiligo, an autoimmune dermatologic condition in three independent cohorts (36–38). Moreover, rs8192917 has also been associated with subacute sclerosing panencephalitis (39) and postoperative keloids (40) in single cohorts. Additionally, this SNP has been associated with transplantation outcomes after HLA-matched unrelated bone marrow transplantation (41). On the functional level, a study has suggested that rs8192917 associates with natural killer cell cytotoxicity (42), however further functional validation is needed to determine the specific nature of this finding. Nonetheless, another *GZMB* SNP, independent of rs8192917 has been associated with joint destruction in rheumatoid arthritis (52), strengthening

the role of genetic variants of *GZMB* in immune-related pathogenic mechanisms.

Although these previous lines of evidence suggest a wide-raging disease-modifying effect of GZMB genetic variants, mechanistic insights are still lacking. Sun et al. have clearly demonstrated that the minor RAH haplotype retains the pro-apoptotic activity of GrB (27), however, GrB has multiple roles in extracellular matrix remodeling, epithelial-to-mesenchymal transition, inflammation and fibrosis (22). QPY and RAH variants may have different substrate-specificities, which might directly affect some of these mechanisms. Regarding autoimmune pathomechanisms as in the case of generalized vitiligo, it has already been suggested that altered
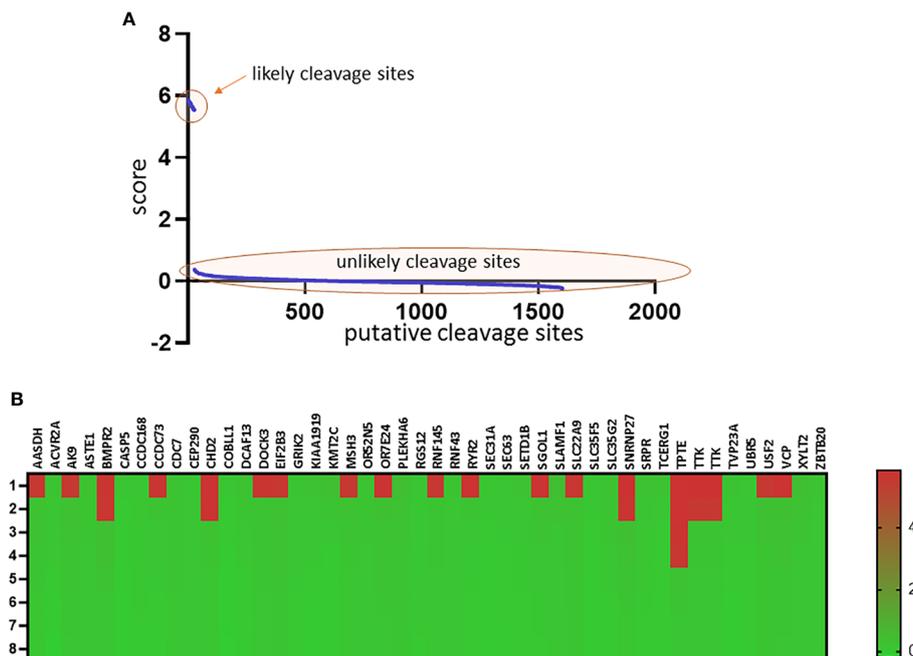
FIGURE 3

Predicted cleavage sites of Granzyme B in shared neoantigens in MSI-H tumors Representation of each predicted cleavage site score in decreasing order. Sites with the highest probability are located on the left side, sites with the lowest probability are located on the right side (Panel **A**). Heatmap representing the top 8 cleavage sites of each shared neoantigen and their corresponding scores (Panel **B**). On the top of the heatmap gene IDs from which each neoantigen originates are presented. Detailed results of each cleavage site prediction are presented in Supplementary Table 2.

cleavage of autoantigens might explain the previously confirmed association of rs8192917 with this disease (38). Transposing this hypothesis to MSI-H carcinogenesis in LS, our *in silico* analysis revealed that a large proportion of shared neoantigens in MSI-H cancers encompass likely GrB cleavage sites. Differential cleavage of these frameshift peptides serving as neoantigens might directly affect optimal immunosurveillance, a key feature in LS-related tumorigenesis (15). Moreover, performing in-depth analyses regarding autoimmune conditions and cancer risk in individuals with LS might further shed light on the immunological pathomechanism of these diseases which can possibly be affected by GrB and other mediators.

It is important to note the limitations of our study. Deriving from its' monocentric nature, we cannot infer the observed association to other populations. Although our cohort is smaller than those of large international consortia, such as the PLSD (12), it is still the largest Hungarian Lynch syndrome cohort ever studied and is comparable to a Swedish national study, where the country's population is also comparable to Hungary (53). By disregarding individuals with *MSH6* and *PMS2* mutations, where cancer risk is significantly lower, and selecting only individuals with *MLH1*, *MSH2* or *EPCAM* mutations, where lifetime colorectal cancer risk is approximately 50%, we were able to perform our comparative study in a relatively homogeneous population.

In conclusion, we found that rs8192917 SNP of the gene encoding GrB correlates with cancer risk in our LS cohort. Following validation of this finding in independent cohorts, this SNP can be included in personalized risk stratification and screening recommendations in affected individuals. Further research avenues might also include the functional assessment of the QPY and RAH variants of GrB to investigate possible differences in substrate-specificities that might explain the observed protective effect.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: European Variation Archive, Project ID: PRJEB55654, Analysis ID: ERZ13336597.

Ethics statement

The studies involving human participants were reviewed and approved by Scientific and Research Ethics Committee of the Medical Research Council of Hungary. Written informed consent

to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

Hypothesis and objective: VG and AP. Genetic counseling: VG, HB, AP, and EO. Molecular genetic diagnosis: AB, TP, JP, HB, AP, and EO. Genotyping: PN and VG. Bioinformatics analysis of WES: IL. In silico analysis: VG. Manuscript preparation: VG, IL, AB, JP, HB, EO, and AP. All authors contributed to the article and approved the submitted version.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2023.1005066/full#supplementary-material

References

1. Schumacher TN, Schreiber RD. Neoantigens in cancer immunotherapy. *Science* (2015) 348(6230):69–74. doi: 10.1126/science.aaa4971

2. Andre T, Shiu KK, Kim TW, Jensen BV, Jensen LH, Punt C, et al. Pembrolizumab in microsatellite-Instability-High advanced colorectal cancer. *New Engl J Med* (2020) 383(23):2207–18. doi: 10.1056/NEJMoa2017699

3. Gandhi L, Rodriguez-Abreu D, Gadgeel S, Esteban E, Felip E, De Angelis F, et al. Pembrolizumab plus chemotherapy in metastatic non-Small-Cell lung cancer. *New Engl J Med* (2018) 378(22):2078–92. doi: 10.1056/NEJMoa1801005

4. Blass E, Ott PA. Advances in the development of personalized neoantigen-based therapeutic cancer vaccines. *Nat Rev Clin Oncol* (2021) 18(4):215–29. doi: 10.1038/s41571-020-00460-2

5. Brown CE, Alizadeh D, Starr R, Weng L, Wagner JR, Naranjo A, et al. Regression of glioblastoma after chimeric antigen receptor T-cell therapy. *New Engl J Med* (2016) 375(26):2561–9. doi: 10.1056/NEJMoa1610497

6. Ballhausen A, Przybilla MJ, Jendrusch M, Haupt S, Pfaffendorf E, Seidler F, et al. The shared frameshift mutation landscape of microsatellite-unstable cancers suggests immunoediting during tumor evolution. *Nat Commun* (2020) 11(1):4740. doi: 10.1038/s41467-020-18514-5

7. Roudko V, Bozkus CC, Orfanelli T, McClain CB, Carr C, O'Donnell T, et al. Shared immunogenic poly-epitope frameshift mutations in microsatellite unstable tumors. *Cell* (2020) 183(6):1634–49 e17. doi: 10.1016/j.cell.2020.11.004

8. Kloor M, Reuschenbach M, Pauligk C, Karbach J, Rafiyan MR, Al-Batran SE, et al. A frameshift peptide neoantigen-based vaccine for mismatch repair-deficient cancers: A phase I/Iia clinical trial. *Clin Cancer* (2020) 26(17):4503–10. doi: 10.1158/1078-0432.CCR-19-3517

9. Enokida T, Moreira A, Bhardwaj N. Vaccines for immunoprevention of cancer. *J Clin Invest* (2021) 131(9). doi: 10.1172/JCI146956

10. Roudko V, Cimen Bozkus C, Greenbaum B, Lucas A, Samstein R, Bhardwaj N. Lynch syndrome and msi-h cancers: From mechanisms to "Off-the-Shelf" cancer vaccines. *Front Immunol* (2021) 12:757804. doi: 10.3389/fimmu.2021.757804

11. Win AK, Jenkins MA, Dowty JG, Antoniou AC, Lee A, Giles GG, et al. Prevalence and penetrance of major genes and polygenes for colorectal cancer. *Cancer epidemiol Biomarkers Prev* (2017) 26(3):404–12. doi: 10.1158/1055-9965.EPI-16-0693

12. Dominguez-Valentin M, Sampson JR, Seppala TT, Ten Broeke SW, Plazzer JP, Nakken S, et al. Cancer risks by gene, age, and gender in 6350 carriers of pathogenic mismatch repair variants: Findings from the prospective lynch syndrome database. *Genet Med* (2020) 22(1):15–25. doi: 10.1038/s41436-019-0596-9

13. Schwitalle Y, Kloor M, Eiermann S, Linnebacher M, Kienle P, Knaebel HP, et al. Immune response against frameshift-induced neopeptides in hnpcc patients and healthy hnpcc mutation carriers. *Gastroenterology* (2008) 134(4):988–97. doi: 10.1053/j.gastro.2008.01.015

14. Reuschenbach M, Kloor M, Morak M, Wentzensen N, Germann A, Garbe Y, et al. Serum antibodies against frameshift peptides in microsatellite unstable colorectal cancer patients with lynch syndrome. *Fam Cancer* (2010) 9(2):173–9. doi: 10.1007/s10689-009-9307-z

15. Pastor DM, Schlom J. Immunology of lynch syndrome. *Curr Oncol Rep* (2021) 23(8):96. doi: 10.1007/s11912-021-01085-z

16. Bohaumilitzky L, Kluck K, Huneburg R, Gallon R, Nattermann J, Kirchner M, et al. The different immune profiles of normal colonic mucosa in cancer-free lynch syndrome carriers and lynch syndrome colorectal cancer patients. *Gastroenterology* (2022) 162(3):907–19 e10. doi: 10.1053/j.gastro.2021.11.029

17. Chang K, Taggart MW, Reyes-Uribe L, Borras E, Riquelme E, Barnett RM, et al. Immune profiling of premalignant lesions in patients with lynch syndrome. *JAMA Oncol* (2018) 4(8):1085–92. doi: 10.1001/jamaoncol.2018.1482

18. Walkowska J, Kallemose T, Jonsson G, Jonsson M, Andersen O, Andersen MH, et al. Immunoprofiles of colorectal cancer from lynch syndrome. *Oncoimmunology* (2019) 8(1):e1515612. doi: 10.1080/2162402X.2018.1515612

19. Kloor M, Michel S, Buckowitz B, Ruschoff J, Buttner R, Holinski-Feder E, et al. Beta2-microglobulin mutations in microsatellite unstable colorectal tumors. *Int J Cancer J Int du Cancer* (2007) 121(2):454–8. doi: 10.1002/ijc.22691

20. Dierssen JW, de Miranda NF, Ferrone S, van Puijenbroek M, Cornelisse CJ, Fleuren GJ, et al. Hnpcc versus sporadic microsatellite-unstable colon cancers follow different routes toward loss of hla class I expression. *BMC Cancer* (2007) 7:33. doi: 10.1186/1471-2407-7-33

21. Voskoboinik I, Whisstock JC, Trapani JA. Perforin and granzymes: Function, dysfunction and human pathology. *Nat Rev Immunol* (2015) 15(6):388–400. doi: 10.1038/nri3839

22. Velotti F, Barchetta I, Cimini FA, Cavallo MG. Granzyme b in inflammatory diseases: Apoptosis, inflammation, extracellular matrix remodeling, epithelial-to-Mesenchymal transition and fibrosis. *Front Immunol* (2020) 11:587581. doi: 10.3389/fimmu.2020.587581

23. Estebanez-Perpina E, Fuentes-Prior P, Belorgey D, Braun M, Kiefersauer R, Maskos K, et al. Crystal structure of the caspase activator human granzyme b, a proteinase highly specific for an asp-P1 residue. *Biol Chem* (2000) 381(12):1203–14. doi: 10.1515/BC.2000.148

24. Sehnal D, Bittrich S, Deshpande M, Svobodova R, Berka K, Bazgier V, et al. Mol* viewer: Modern web app for 3d visualization and analysis of Large biomolecular structures. *Nucleic Acids Res* (2021) 49(W1):W431–W7. doi: 10.1093/nar/gkab314

25. Slatkin M. Linkage disequilibrium–understanding the evolutionary past and mapping the medical future. *Nat Rev Genet* (2008) 9(6):477–85. doi: 10.1038/nrg2361

26. McIlroy D, Cartron PF, Tuffery P, Dudoit Y, Samri A, Autran B, et al. A triple-mutated allele of granzyme b incapable of inducing apoptosis. *Proc Natl Acad Sci United States America* (2003) 100(5):2562–7. doi: 10.1073/pnas.0437935100

27. Sun J, Bird CH, Thia KY, Matthews AY, Trapani JA, Bird PI. Granzyme b encoded by the commonly occurring human rah allele retains pro-apoptotic activity. *J Biol Chem* (2004) 279(17):16907–11. doi: 10.1074/jbc.M400563200

28. Papp J, Kovacs ME, Olah E. Germline Mlh1 and Msh2 mutational spectrum including frequent Large genomic aberrations in Hungarian hereditary non-polyposis colorectal cancer families: Implications for genetic testing. *World J Gastroenterol WJG* (2007) 13(19):2727–32. doi: 10.3748/wjg.v13.i19.2727

29. Kovacs ME, Papp J, Szentirmay Z, Otto S, Olah E. Deletions removing the last exon of Tacstd1 constitute a distinct class of mutations predisposing to lynch syndrome. *Hum Mutat* (2009) 30(2):197–203. doi: 10.1002/humu.20942

30. Machiela MJ, Chanock SJ. Ldlink: A web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* (2015) 31(21):3555–7. doi: 10.1093/bioinformatics/btv402

31. Li H, Durbin R. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* (2009) 25(14):1754–60. doi: 10.1093/bioinformatics/btp324

32. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* (2011) 43(5):491–8. doi: 10.1038/ng.806

33. Sarkadi B, Liko I, Nyiro G, Igaz P, Butz H, Patocs A. Analytical performance of ngs-based molecular genetic tests used in the diagnostic workflow of Pheochromocytoma/Paraganglioma. *Cancers (Basel)* (2021) 13(16). doi: 10.3390/cancers13164219

34. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: Analysis and visualization of ld and haplotype maps. *Bioinformatics* (2005) 21(2):263–5. doi: 10.1093/bioinformatics/bth457

35. Song J, Li F, Leier A, Marquez-Lago TT, Akutsu T, Haffari G, et al. Prosperous: High-throughput prediction of substrate cleavage sites for 90 proteases with improved accuracy. *Bioinformatics* (2018) 34(4):684–7. doi: 10.1093/bioinformatics/btx670

36. Jeong KH, Kim SK, Seo JK, Shin MK, Lee MH. Association of gzmb polymorphisms and susceptibility to non-segmental vitiligo in a Korean population. *Sci Rep* (2021) 11(1):397. doi: 10.1038/s41598-020-79705-0

37. Xu M, Liu Y, Liu Y, Li X, Chen G, Dong W, et al. Genetic polymorphisms of gzmb and vitiligo: A genetic association study based on Chinese han population. *Sci Rep* (2018) 8(1):13001. doi: 10.1038/s41598-018-31233-8

38. Ferrara TM, Jin Y, Gowan K, Fain PR, Spritz RA. Risk of generalized vitiligo is associated with the common 55r-94a-247h variant haplotype of gzmb (Encoding granzyme b). *J Invest Dermatol* (2013) 133(6):1677–9. doi: 10.1038/jid.2013.5

39. Yentur SP, Aydin HN, Gurses C, Demirbilek V, Kuru U, Uysal S, et al. Granzyme b gene polymorphism associated with subacute sclerosing panencephalitis. *Neuropediatrics* (2014) 45(5):309–13. doi: 10.1055/s-0034-1378129

40. Wen X, Du H, Hao X, Wang J, Guo Y. Expression of the gzmb gene polymorphism, snp Rs8192917, in 990 han Chinese patients with postoperative keloids. *Med Sci Monit* (2022) 28:e936963. doi: 10.12659/MSM.936963

41. Espinoza LJ, Takami A, Nakata K, Yamada K, Onizuka M, Kawase T, et al. Genetic variants of human granzyme b predict transplant outcomes after hla matched unrelated bone marrow transplantation for myeloid malignancies. *PloS One* (2011) 6(8):e23827. doi: 10.1371/journal.pone.0023827

42. Oboshi W, Watanabe T, Hayashi K, Nakamura T, Yukimasa N. Qpy/Rah haplotypes of the gzmb gene are associated with natural killer cell cytotoxicity. *Immunogenetics* (2018) 70(1):29–36. doi: 10.1007/s00251-017-1014-6

43. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alfoldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* (2020) 581(7809):434–43. doi: 10.1038/s41586-020-2308-7

44. Antoniou AC, Spurdle AB, Sinilnikova OM, Healey S, Pooley KA, Schmutzler RK, et al. Common breast cancer-predisposition alleles are associated with breast cancer risk in Brca1 and Brca2 mutation carriers. *Am J Hum Genet* (2008) 82(4):937–48. doi: 10.1016/j.ajhg.2008.02.008

45. Mulligan AM, Couch FJ, Barrowdale D, Domchek SM, Eccles D, Nevanlinna H, et al. Common breast cancer susceptibility alleles are associated with tumour subtypes in Brca1 and Brca2 mutation carriers: Results from the consortium of investigators of modifiers of Brca1/2. *Breast Cancer Res* (2011) 13(6):R110. doi: 10.1186/bcr3052

46. Barnes DR, Rookus MA, McGuffog L, Leslie G, Mooij TM, Dennis J, et al. Polygenic risk scores and breast and epithelial ovarian cancer risks for carriers of Brca1 and Brca2 pathogenic variants. *Genet Med* (2020) 22(10):1653–66. doi: 10.1038/s41436-020-0862-x

47. Lakeman IMM, van den Broek AJ, Vos JAM, Barnes DR, Adlard J, Andrulis IL, et al. The predictive ability of the 313 variant-based polygenic risk score for contralateral breast cancer risk prediction in women of European ancestry with a heterozygous Brca1 or Brca2 pathogenic variant. *Genet Med* (2021) 23(9):1726–37. doi: 10.1038/s41436-021-01198-7

48. Talseth-Palmer BA, Wijnen JT, Brenne IS, Jagmohan-Changur S, Barker D, Ashton KA, et al. Combined analysis of three lynch syndrome cohorts confirms the modifying effects of 8q23.3 and 11q23.1 in Mlh1 mutation carriers. *Int J Cancer J Int du Cancer* (2013) 132(7):1556–64. doi: 10.1002/ijc.27843

49. Ten Broeke SW, Elsayed FA, Pagan L, Olderode-Berends MJW, Garcia EG, Gille HJP, et al. Snp association study in Pms2-associated lynch syndrome. *Fam Cancer* (2018) 17(4):507–15. doi: 10.1007/s10689-017-0061-3

50. Wiik MU, Evans TJ, Belhadj S, Bolton KA, Dymerska D, Jagmohan-Changur S, et al. A genetic variant in telomerase reverse transcriptase (Tert) modifies cancer risk in lynch syndrome patients harbouring pathogenic Msh2 variants. *Sci Rep* (2021) 11(1):11401. doi: 10.1038/s41598-021-90501-2

51. Wijnen JT, Brohet RM, van Eijk R, Jagmohan-Changur S, Middeldorp A, Tops CM, et al. Chromosome 8q23.3 and 11q23.1 variants modify colorectal cancer risk in lynch syndrome. *Gastroenterology* (2009) 136(1):131–7. doi: 10.1053/j.gastro.2008.09.033

52. Knevel R, Krabben A, Wilson AG, Brouwer E, Leijsma MK, Lindqvist E, et al. A genetic variant in granzyme b is associated with progression of joint destruction in rheumatoid arthritis. *Arthritis Rheum* (2013) 65(3):582–9. doi: 10.1002/art.37808

53. Bernstedt SW, Bjork J, Fritzell K, Spigelman AD, Bjorck E, Backman AS. Room for improvement: One third of lynch syndrome patients presenting for genetic testing in a highly specialised centre in Stockholm already have cancer. *Hered Cancer Clin Pract* (2021) 19(1):18. doi: 10.1186/s13053-021-00171-4