

RESEARCH

Open Access



Genomic susceptibility to gastric cancer in Northwest Iran: population-based and case–control studies

Homa Akhavan Aghghaleh¹ , Najmeh Ranji^{1*} and Hadi Habibollahi¹

Abstract

Introduction The age-standardized incidence rate for gastric cancer is estimated to be 11.1% worldwide and 39.1% for Ardabil province in northwest Iran. Single nucleotide polymorphisms (SNPs) occur in coding and non-coding regions, contributing to cancer susceptibility. To identify SNPs predisposing individuals to gastric cancer in this region, we compared 263 variants between the Ardabil population and other populations.

Materials and methods Whole exome sequencing was used to determine the distribution of variants in the genomic DNA of 150 volunteers (aged < 35 years) from the general population of Ardabil. We compared allele frequencies with databases such as Iranome, Alfa, GnomAD, and 1000G, and statistically analyzed their correlation with age-standardized incidence rates (ASRs) for gastric cancer in related populations using the Pearson correlation test. Some findings were validated using Sanger-based PCR-Sequencing. We determined the frequency of seventeen variants among 150 individuals with gastric cancer and 150 healthy volunteers (matched for age and sex) as the control group.

Results Nineteen variants, including rs10061133, rs1050631, rs12220909, rs12983273, rs1695, rs2274223, rs2292832, rs2294008, rs2505901, rs2976391, rs33927012, rs3744037, rs3745469, rs4789936, rs4986790, rs4986791, rs6194, rs63750447, and rs6505162, were found to be significantly different between the general population of Ardabil and other populations. Among them, the variants rs1050631, rs12983273, rs1695, rs2274223, rs2292832, rs2505901, rs33927012, rs374569, and rs6505162 showed significant differences between the cases and controls.

Discussion In this study, 17 variants appeared to be involved in the etiology of the high frequency of gastric cancer in the Ardabil population. Some of the observed differences were consistent with previous case–control and meta-analysis reports from various parts of the world. These findings motivate further cohort investigations in this population. Ultimately, identifying prognostic factors can help diagnose individuals predisposed to gastric cancer in this population.

Introduction

Gastric cancer begins with the uncontrolled proliferation of cells in the stomach. Despite a decline in incidence, it remains one of the most common types of cancer worldwide. According to GLOBOCAN 2020, gastric cancer ranks fifth in incidence at 5.6%. However, it is the fourth leading cause of cancer deaths globally. This type of cancer is more common in males, ranking fourth among males. However, it is less significant among common cancers in women. Out of the 1,089,103 new cases of gastric

*Correspondence:

Najmeh Ranji

najmehranji@gmail.com; n_ranji@iaurasht.ac.ir

¹ Department of Biology, Faculty of Sciences, Rasht Branch, Islamic Azad University, P.O. Box 3516-41335, Rasht, Iran

cancer, 75.3% (819,944 cases) occurred in the Asian population. Additionally, the Asian population accounted for 74.8% of the mortality cases. The age-standardized incidence rate (ASR) for gastric cancer was estimated to be 11.1% worldwide. Eastern Asia, with the highest level of incidence rate, had an ASR of 22.4. Males and females in this area had ASRs of 32.5 and 13.2, respectively [1].

Ardabil province is located in the northwest of Iran. Based on an ASR of 49.1 for males and 25.4 for females in this area, compared to the Korean Republic (39.7 for males and 17.6 for females), its population is among the highest prevalence population affected by gastric cancer [2].

Among the subtypes of gastric cancer, the cardia type has been shown to be significantly higher in the Ardabil region compared to some European countries, the USA, Japan, and Korea. The highest incidence of gastric cancer, especially the cardia type, in the Ardabil region provides an exceptional opportunity to investigate its etiology [3].

Through genome-wide association studies (GWAS), more than 430 cancer-associated loci have been identified [4]. Multiple GWAS have been conducted for gastric cancer, leading to the determination of many cancer-associated loci. The SNPs associated with cancer susceptibility are located in coding or non-coding regions of the genome. Synonymous and non-synonymous mutations are two subtypes of coding SNPs [5].

Despite not affecting the amino acid sequence, synonymous variants may change the expression level and function of the gene product through post-transcriptional modification, translation rate stability of mRNA, failure of splicing regulatory proteins to interact with some exons, affecting the kinetics of translation, followed by slowing down the rate of protein synthesis and/or modification, and ultimately altering the pause site, which may result in a different conformation of the protein. Non-coding variants are located in intervening and intergenic sequences, accounting for more than 90% of interindividual variations. Depending on the variant's location, it may be involved in the regulation of gene expression by harboring response elements such as promoters and enhancers, as well as post-transcriptional and translational processing [6].

Additionally, changes in epigenetic modification and chromatin structure caused by non-coding variants can alter the target gene expression level. Variants in the 3' UTR of genes can modulate the interactions between mRNA and related microRNAs, as well as polyadenylation, and can impact translation efficiency and/or mRNA stability [7].

MicroRNAs are small (18–20 nucleotide) non-coding RNAs that regulate mRNA translation by directly binding to their 3'-UTR. This regulation is managed through

cleavage of the mRNA transcript and/or repression of translation, depending on the complementarity between the miRNA and its targeted mRNA. Therefore, SNPs in the sequence and/or 3'-UTR of the mRNA transcript can be valuable markers for predicting cancer susceptibility. SNPs in miRNAs and miRNA binding sites play critical roles as cancer risk biomarkers. For example, the miR-453 binding site in the ESR1 gene, the miR-638 binding site in the BRCA1 gene, and the miR-628-5p binding site in the TGFBR1 gene are important in breast cancer. Variants such as rs3783553 in hepatocellular carcinoma create new binding sites for miR-122 and miR-378 in the IL1A gene. There is also a relationship between SNPs within miRNA binding sites of the CD886, INSR, RPA2, and GTF2H1 genes and colorectal cancer. Furthermore, certain SNPs in the 3'-UTR of the MYCL1 and NBS1 genes are prognostic biomarkers in lung cancer, as well as breast, ovarian, and bladder cancers, while also serving as predictive biomarkers in bladder cancer (for response to radiotherapy), prostate cancer (for response to ADT), and response to Methotrexate and Cisplatin. SNPs in the seed sequence of miRNAs can also affect their processing or binding. Reported variants such as rs2910164 in miR-146a, rs3746444 in miR-499, rs12975333 in miR-125a, rs34059726 in miR-124, and rs11614913 in miR-196-a2 indicate an association between miRNA SNPs and cancer risk [8, 9].

In this study, we compared the frequency of 263 variants in coding and non-coding sequences between populations based on previous case-control, meta-analysis, and GWAS reports. The aim of this research was to identify variants that could be considered as predictive values for susceptibility to gastric cancer in the Ardabil population.

Materials and methods

Study design

All participants were provided signed informed consent declaring to investigate the peripheral blood samples according to the ethical rule of the Iranian Ministry of Health.

One hundred and fifty healthy individuals aged 18–35 years (mean age 27.32 ± 6.32), 150 cases of primary gastric cancer aged 36–72 years (mean age 63.2 ± 11.3), and 150 healthy age- and sex-matched individuals for the case group aged 37–70 years (mean age 60.1 ± 12.4) were included in this study. Among the general population group, 86 individuals (57.3%) were male. This distribution was 60% for cases and 56.7% for controls, respectively. All of the volunteers were born and currently reside in Ardabil province in the northwest of Iran. The inclusion criteria for the general population group were birth and residence in Ardabil, while the exclusion criteria were

having cancer, having first-degree relatives with any type of cancer, and having gastrointestinal disorders or diseases associated with DNA repair system deficiencies. Additionally, having a relative affected by cancer was an exclusion criterion for the control group. The ethical and regulatory issues related to the collection of human specimens for research purposes were approved by the Ethics Committee of Islamic Azad University-Rasht Branch (Approval ID: IR.IAU.RASHT.REC.1398.057).

All participants provided signed informed consent, in accordance with the ethical rules of the Iranian Ministry of Health, to investigate their peripheral blood samples.

Whole exome sequencing

Genomic DNA was isolated from the patient's specimen, EDTA anticoagulated peripheral blood, using a silica-membrane-based DNA purification method recommended by the manufacturer (QIAamp DNA Blood Mini Kit, QIAGEN, Germany). The integrity of the DNA was verified using the Qubit 4 Fluorometer (Thermo Fisher Scientific, USA). A total of 1.0 µg of genomic DNA per sample was used for the DNA sample preparation.

Sequencing libraries were created using the Agilent SureSelect Human All ExonV7 kit (Agilent Technologies, CA, USA) following the manufacturer's instructions. Index codes were added to attribute sequences to each sample. Fragmentation was performed using a hydrodynamic shearing system (Covaris, Massachusetts, USA) to generate fragments of 180–280 bp. Remaining overhangs were converted into blunt ends using exonuclease/polymerase activities, and enzymes were then removed. After adenylation of the 3' ends of the DNA fragments, adapter oligonucleotides were ligated. DNA fragments with adapter molecules on both ends were selectively enriched in a PCR reaction. Captured libraries were further enriched in a PCR reaction to add index tags for hybridization. The products were purified using the AMPure XP system (Beckman Coulter, Beverly, USA) and quantified using the Agilent high sensitivity DNA assay on the Agilent Bioanalyzer 2100 system. The qualified libraries were then sequenced using the NovaSeq 6000 Illumina sequencers.

Data quality control, analysis, and interpretation were performed on a G9 generation HP server using a Unix-based operating system. Variant analysis and classification utilized relevant databases and references [10–18].

Sanger-based PCR-sequencing

Despite using the WES 100X for variant reading, 28 variants were randomly validated using Sanger sequencing. The selection criteria for these variants were based on their location within the gene, including coding sequences, promoters, introns, and non-coding

sequences such as microRNAs. The PCR products were sequenced directly on the automated genetic analyzer (Applied Biosystems 3130xl; USA), followed by blasting on the NCBI website (<http://www.ncbi.nlm.nih.gov/blast>) and comparison with normal sequences. Twenty-six primer pairs were used to amplify the selected variants (see Additional file 1: Table S1). The Primer3 software (<https://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>) and previous studies [19–29] were utilized for primer design and selection. Ultimately, seventeen variants were included in the case–control study.

Statistical analysis

Descriptive statistics of the variants were presented as frequencies. The frequency of the variants was calculated using standard methods. Differences in frequencies between groups were compared using the Pearson correlation test. Additionally, scatter plots were drawn for each variant. Genotype and allele frequencies between the cases and controls were analyzed using the Chi-square test. For all analyses, a *P* value of less than 0.05 was considered statistically significant. The IBM SPSS Statistics version 25 was used for the statistical analyses. Three online Chi-square software tools (<https://genecalc.pl/hardy-weinberg-page>, <https://www.had2know.org/academics/hardy-weinberg-equilibrium-calculator-2-alleles.html>, and <https://wpcalc.com/en/equilibrium-hardy-weinberg/>) were used to test for deviation from the Hardy–Weinberg equilibrium (HWE) in the studied groups.

Results

Significant differences of allele distribution among different populations

For the 260 variants selected to be correlated with gastric carcinogenesis through a literature review, allele frequencies were obtained from the whole exome sequencing (WES) results of the general population group (Additional file 2: Table S2).

Comparison of these results with frequencies from genomic databases such as Iranome, Alfa, 1000G, ExAC, and gnomAD, for Iran, Europe, and the world, revealed considerable differences. These variant frequencies were statistically analyzed using the Pearson correlation test and scatter plots, based on their reported age-standardized rates (ASR). Table 1 and Additional file 4: Fig. S1 display the significant differences ($P < 0.05$) and their corresponding scatter plots. The variants that showed significant differences were rs10061133, rs1050631, rs12220909, rs12983273, rs1695, rs2274223, rs2292832, rs2294008, rs2505901, rs2976391, rs33927012, rs3744037, rs3745469, rs4789936, rs4986790, rs4986791, rs6194, rs63750447, and rs6505162.

Table 1 Statistical association between the variant frequencies in Ardabil population comparing Iran, Europe, and World data, based on their reported Age-Standardized Rates (ASR for Ardabil: 39.1, Iran: 16.6, Europe: 13.7, and World: 11.1)

Variant	Ardabil*	Iran	Europe	World	Pearson Cor.**	P Value ₁	Pearson Cor.***	P Value ₂
rs10061133	0.13	0.127	0.097	0.091	0.741	0.259	0.998	0.038
rs10069690	0.127	0.13	0.27	0.26	-0.716	0.284	-0.987	0.101
rs1044165	0.077	0.07	0.14	0.13	-0.637	0.363	-0.970	0.156
rs1047768	0.46	0.52	0.59	0.57	-0.912	0.088	-0.974	0.145
rs1050631	0.27	0.33	0.36	0.35	-0.973	0.027	-0.983	0.118
rs10509577	0.01	0.01	0.07	0.06	-0.693	0.307	-0.968	0.161
rs1051740	0.29	0.24	0.3	0.3	0.020	0.980	-0.996	0.054
rs10519613	0.163	0.14	0.1	0.11	0.878	0.122	0.970	0.156
rs1052133	0.3	0.31	0.22	0.22	0.633	0.367	0.996	0.054
rs1052536	0.35	0.35	0.46	0.41	-0.639	0.307	-0.931	0.238
rs1052555	0.36	0.36	0.33	0.31	0.638	0.362	0.875	0.322
rs1059394	0.49	0.48	0.31	0.34	0.722	0.278	0.968	0.161
rs1061302	0.37	0.4	0.31	0.31	0.473	0.527	0.996	0.054
rs11134527	0.25	0.36	0.26	0.25	-0.209	0.791	-0.581	0.605
rs11170877	0.08	0.08	0.1	0.11	-0.658	0.342	-0.909	0.274
rs1143633	0.28	0.26	0.35	0.35	-0.559	0.441	-0.996	0.054
rs11797	0.33	0.38	0.55	0.58	-0.803	0.197	-0.979	0.132
rs12220909	0.003	0.01	0.0009	0.0007	0.057	0.943	1.0	0.004
rs1229984	0.75	0.81	0.96	0.94	-0.848	0.152	-0.983	0.117
rs12691693	0.64	0.54	0.61	0.59	0.561	0.439	0.875	0.322
rs12983273	0.107	0.136	0.138	0.136	-0.983	0.017	-0.990	0.090
rs12983508	0.1	0.11	0.17	0.17	-0.768	0.232	-0.996	0.054
rs138978736	0	0.0012	0.00075	0.00064	-0.756	0.26	-0.999	0.148
rs1562313	0.18	0.21	0.2	0.2	-0.855	0.145	-0.996	0.054
rs1695	0.297	0.319	0.326	0.327	-0.994	0.006	-0.998	0.035
rs16999593	0.013	0.007	0.0003	0.002	0.918	0.082	0.978	0.132
rs17276588	0.047	0.03	0.029	0.041	0.665	0.335	0.689	0.516
rs1799950	0.1	0.05	0.07	0.065	0.816	0.184	0.974	0.146
rs1800469	0.52	0.58	0.68	0.67	-0.882	0.118	-0.990	0.089
rs1800668	0.37	0.33	0.31	0.28	0.868	0.132	0.909	0.274
rs1801394	0.4	0.45	0.55	0.52	-0.858	0.142	-0.959	0.183
rs1870134	0.02	0.0187	0.00369	0.00496	0.746	0.255	0.999	0.106
rs1884444	0.497	0.54	0.53	0.53	-0.914	0.086	-0.996	0.054
rs2010963	0.58	0.63	0.68	0.68	-0.933	0.067	-0.996	0.054
rs2066702	0	0.004	0.002	0.01	-0.715	0.285	-0.725	0.484
rs2066981	0.41	0.42	0.48	0.47	-0.771	0.229	-0.974	0.146
rs2069763	0.22	0.23	0.33	0.32	-0.750	0.250	-0.984	0.114
rs2070600	0.007	0.02	0.04	0.04	-0.890	0.110	-0.996	0.054
rs2074570	0.01	0.05	0.04	0.056	-0.888	0.112	-0.902	0.284
rs2075570	0.38	0.43	0.5	0.51	-0.894	0.106	-0.986	0.106
rs217727	0.13	0.13	0.19	0.19	-0.694	0.306	-0.996	0.054
rs2228570	0.74	0.75	0.61	0.61	0.656	0.344	0.996	0.054
rs2257082	0.19	0.22	0.27	0.27	-0.882	0.118	-0.996	0.054
rs2274084	0.027	0.02	0.002	0.01	0.806	0.194	0.915	0.264
rs2274223	0.42	0.36	0.32	0.32	0.964	0.036	0.996	0.054
rs2274976	0.047	0.037	0.044	0.046	0.341	0.659	0.689	0.516
rs2292832	0.51	0.64	0.7	0.7	-0.983	0.017	-0.996	0.054
rs2294008	0.4	0.41	0.45	0.45	-0.964	0.323	-0.996	0.254

Table 1 (continued)

Variant	Ardabil*	Iran	Europe	World	Pearson Cor.**	P Value ₁	Pearson Cor.***	P Value ₂
rs2296147	0.34	0.41	0.48	0.45	-0.916	0.084	-0.959	0.184
rs2305158	0.2	0.23	0.23	0.22	-0.866	0.134	-0.866	0.134
rs2505901	0.32	0.45	0.5	0.49	-0.986	0.014	-0.991	0.085
rs2839698	0.39	0.45	0.51	0.45	-0.803	0.197	-0.814	0.395
rs2920297	0.35	0.37	0.46	0.45	-0.796	0.204	-0.984	0.114
rs2976391	0.57	0.51	0.46	0.45	0.944	0.056	1	0.006
rs2976394	0.37	0.41	0.46	0.46	-0.912	0.088	-0.996	0.054
rs3171425	0.5	0.55	0.59	0.57	-0.931	0.069	-0.956	0.189
rs33927012	0.05	0.02	0.01	0.01	0.992	0.008	0.996	0.054
rs3660	0.29	0.37	0.52	0.51	-0.869	0.131	-0.992	0.078
rs3730089	0.12	0.13	0.16	0.17	-0.793	0.207	-0.959	0.183
rs3732183	0.21	0.23	0.27	0.3	-0.832	0.168	-0.972	0.151
rs3738590	0.24	0.24	0.26	0.26	-0.694	0.306	-0.996	0.054
rs3741216	0.02	0.04	0.03	0.04	-0.849	0.151	-0.910	0.272
rs3741219	0.37	0.46	0.5	0.45	-0.886	0.114	-0.883	0.311
rs3744037	0.2	0.166	0.163	0.167	0.976	0.024	0.983	0.116
rs3745469	0.24	0.17	0.1	0.09	0.930	0.070	1.0	0.016
rs3762272	0.08	0.06	0.02	0.03	0.891	0.137	0.998	0.161
rs3764880	0.43	0.34	0.26	0.29	0.932	0.068	0.969	0.159
rs3765524	0.41	0.35	0.29	0.3	0.932	0.068	0.987	0.101
rs3804100	0.14	0.14	0.07	0.07	0.694	0.306	0.996	0.054
rs3810366	0.54	0.52	0.3	0.3	0.737	0.263	0.996	0.054
rs3825569	0.65	0.55	0.63	0.62	0.427	0.573	0.972	0.151
rs4072037	0.42	0.45	0.52	0.54	-0.823	0.177	-0.968	0.161
rs4652	0.39	0.4	0.42	0.43	-0.837	0.163	-0.987	0.101
rs4754	0.37	0.35	0.27	0.28	0.793	0.207	0.985	0.111
rs4711690	0.23	0.23	0.17	0.19	0.658	0.342	0.909	0.274
rs4789936	0.54	0.51	0.51	0.5	0.988	0.012	0.987	0.101
rs4796030	0.5	0.52	0.56	0.56	0.987	0.101	-0.996	0.054
rs4880	0.45	0.47	0.5	0.49	-0.887	0.113	-0.959	0.183
rs4919510	0.17	0.22	0.19	0.2	-0.702	0.298	-0.972	0.151
rs4986790	0.03	0.06	0.06	0.06	-0.985	0.015	-0.996	0.054
rs4986791	0.03	0.06	0.06	0.06	-0.985	0.015	-0.996	0.054
rs5744168	0.03	0.05	0.06	0.05	-0.915	0.085	-0.914	0.266
rs6194	0.003	0.001	0.0001	0.0005	0.972	0.028	0.978	0.135
rs63750447	0.007	0.0006	0	0.0002	0.991	0.009	0.994	0.070
rs6413413	0.0067	0.004	0.006	0.005	0.567	0.433	0.751	0.460
rs6493	0.19	0.18	0.17	0.17	0.933	0.067	0.996	0.054
rs6494	0.2	0.21	0.22	0.2	-0.430	0.570	-0.414	0.728
rs6505162	0.58	0.56	0.46	0.45	0.780	0.220	1.0	0.010
rs6513497	0.1	0.11	0.07	0.06	0.565	0.435	0.987	0.101
rs698	0.24	0.24	0.39	0.37	-0.698	0.302	-0.976	0.140
rs699517	0.48	0.49	0.31	0.34	0.686	0.339	0.998	0.167
rs709816	0.36	0.47	0.35	0.4	-0.309	0.691	-0.417	0.726
rs7165402	0.12	0.16	0.13	0.14	-0.569	0.431	-0.910	0.272
rs731236	0.34	0.35	0.4	0.39	-0.807	0.193	-0.968	0.161
rs762551	0.59	0.6	0.7	0.69	-0.750	0.250	-0.984	0.114
rs763780	0.04	0.06	0.05	0.05	-0.715	0.285	-0.996	0.054
rs8113645	0.61	0.68	0.64	0.68	-0.812	0.188	-0.873	0.324

Table 1 (continued)

Variant	Ardabil*	Iran	Europe	World	Pearson Cor.**	P Value ₁	Pearson Cor.***	P Value ₂
rs8177812	0.22	0.17	0.13	0.14	0.948	0.052	0.983	0.118
rs897986	0.69	0.68	0.64	0.61	0.711	0.289	0.888	0.304
rs910924	0.13	0.17	0.27	0.25	-0.867	0.156	-0.999	0.146
rs937283	0.25	0.27	0.41	0.38	-0.743	0.257	-0.966	0.166

Bold values indicate a statistically significant difference in allele frequency between Ardabil and other populations

* Our results

** Comparing our results with Iran, Europe, and World

*** Comparing our results with Europe and World

To validate the accuracy of the detected frequencies, 26 randomly selected variants were confirmed using Sanger-based PCR-Sequencing (Additional file 5: Fig. S2).

Significant differences of allele distribution among case and control groups

Based on their significant differences with other populations and the frequency of the variant in our population (Table 1), seventeen variants were selected for evaluation in terms of frequency difference between two case and control groups. The genotype frequencies of the variants rs12220909, rs2292832, rs2505901, rs33927012, rs4789936, rs6194, and rs63750447 did not deviate from HWE, while the others did deviate from HWE. Among the 17 selected variants, the variants rs1050631, rs12983273, rs1695, rs2274223, rs2292832, rs2505901, rs33927012, rs374569, and rs6505162 showed significant differences (Table 2).

The CT (36.7%), TT (0.047%), and CT+TT (41.3%) genotypes of the rs1050631 C>T polymorphism (deviated from HWE: $\chi^2=2.65$; p Value=0.104) were less frequently observed in the gastric cancer affected cases than in the controls (46.7%, 14%, and 52.7%, respectively). Comparisons of the major allele (C) frequencies and genotypes (CC) indicated that they were significantly associated with gastric cancer using a dominant model (CC vs. CT+TT: $P<0.0001$).

Comparisons of the minor allele (T) frequencies and genotypes of the variant rs12983273 (deviated from HWE: $\chi^2=2.03$; P value=0.154) indicated that they had a significant association using recessive models: allele T (7.7% in cases and 13.7% in controls) vs. C, $P=0.02$; genotype CT+TT (14.7% and 24.7% for the cases and controls, respectively), $P=0.031$.

The AG+GG (62.7%) genotype of the rs1695 A>G polymorphism (deviated from HWE: $\chi^2=0.498$; P value=0.48) was more frequently observed in the gastric cancer affected cases than in the controls (50.7%). Additionally, the higher frequency of the allele G among the cases (37.3% vs. 29.7% for the controls) indicated that

Table 2 Consistency with Hardy Weinberg’s law of the selected variants for the general population

Variant	HWE	
	χ^2	P value
rs10061133	3.63	0.057
rs1050631	2.65	0.104
rs12220909	0.002	0.962
rs12983273	2.03	0.154
rs1695	0.498	0.48
rs2274223	0.49	0.48
rs2292832	13.42	0.0002
rs2505901	51.59	0
rs33927012	16.94	0
rs3744037	3.22	0.073
rs3745469	5.74	0.057
rs4789936	7.79	0.005
rs4986790	0.47	0.5
rs4986791	0.4	0.53
rs6194	0.002	0.96
rs63750447	0.005	0.94
rs6505162	2.7	0.1

they were significantly associated with gastric cancer using a recessive model: allele (G vs. A), $P=0.047$; genotype, $P=0.036$; for the genotype (AG+GG vs. AA).

The genotype GG of the variant rs2274223 A>G (deviated from HWE: $\chi^2=0.49$; P value=0.48) showed to be more frequent among the cases (24% vs. 16% for the controls). This difference indicated that it was significantly associated with gastric cancer using a recessive model: genotype GG vs. AA, $P=0.035$.

The allele C (35.3%), and TC (38.7%) and CC (18%) genotypes of the rs2292832 T>C polymorphism (did not deviate from HWE: $\chi^2=13.42$; P value=0.0002) were less frequently observed in the cases than in the healthy controls (64.7%, 43.3%, and 26.7%, respectively). Comparisons of the major allele (T) frequencies and genotypes (TT) indicated that they were significantly associated

with GDM using both dominant and recessive models (for the recessive models: allele (C vs. T), $P=0.001$, and genotype (TC+CC vs. TT), $P=0.01$; for the dominant models (CC vs. TT), $P=0.01$, and (CT vs. TT), $P=0.046$).

The rs2505901 T>C did not deviate from HWE ($\chi^2=51.59$; P value=0). However, comparisons of the minor allele (C) frequencies (28.3% and 36.3% for cases and controls, respectively) and genotypes (TC+CC;

Table 3 Significant differences between cases and controls for some variants

Variant			Cases	Controls	P value	Odds ratio (95% CI)
rs1050631	Genotype (No.)	CC (reference)	88	71		
		CT	55	70	< 0.0001	3.37 (2.16–5.27)
		TT	7	9	0.019	3.4 (1.22–9.49)
		CT+TT	62	79	< 0.0001	3.374 (2.19–5.19)
rs12983273	Genotype (No.)	CC (reference)	128	113		
		CT+TT	22	37	0.031	1.9 (1.1–3.4)
	Allele (Freq.)	C (reference)	0.923	0.863		
		T	0.077	0.137	0.02	1.91 (1.1–3.3)
rs1695	Genotype (No.)	AA (reference)	56	74		
		AG+GG	94	76	0.036	0.612 (0.39–0.97)
	Allele (Freq.)	A (reference)	0.627	0.703		
		G	0.373	0.297	0.047	0.708 (0.5–0.99)
rs2274223	Genotype (No.)	AA (reference)	34	47		
		GG	36	24	0.035	0.482 (0.24–0.95)
rs2292832	Genotype (No.)	TT (reference)	68	45		
		TC	58	65	0.046	1.693 (1.01–2.84)
		CC	27	40	0.01	2.239 (1.21–4.15)
		TC+CC	85	105	0.01	1.867 (1.16–3)
	Allele (Freq.)	T (reference)	0.647	0.517		
		C	0.353	0.483	0.001	1.712 (1.23–2.38)
rs2505901	Genotype (No.)	TT (reference)	88	71		
		TC+CC	62	79	0.0497	1.58 (1.0005–2.5)
	Allele (Freq.)	T (reference)	0.717	0.637		
		C	0.283	0.363	0.0366	1.44 (1.02–2.04)
rs33927012	Allele (Freq.)	T (reference)	0.933	0.97		
		C	0.067	0.03	0.0412	0.43 (0.19–0.97)
rs3745469	Genotype (No.)	GG (reference)	80	97		
		GA+AA	70	53	0.046	0.62 (0.39–0.99)
	Allele (Freq.)	G (reference)	0.72	0.79		
		A	0.28	0.21	0.047	0.68 (0.47–0.99)
rs6505162	Genotype (No.)	AA (reference)	19	34		
		AC	73	60	0.02	0.46 (0.24–0.89)
		AC+CC	131	116	0.025	0.5 (0.27–0.91)

41.3% and 52.7% for cases and controls, respectively) indicated that they were significantly associated with the inverse risk of gastric cancer using a recessive model (T vs. C: $P=0.0366$; TC+CC vs. TT: $P=0.0497$).

Despite not deviating from HWE ($\chi^2=16.94$; P value=0) and not determining any significant difference for the distribution of the genotypes among the cases and controls for the variant rs33927012 T>C, the higher frequency of the allele C (6.7% for the cases vs. 3% for the controls) led to a significant difference ($P=0.0412$).

The GA+AA (46.7%) genotype of the rs3745469 G>A polymorphism (deviated from HWE: $\chi^2=5.74$; P value=0.057) was more frequently observed in the gastric cancer affected cases than in the controls (35.3%). Additionally, the allele A was more frequent among the cases (28% vs. 21% for the controls). These findings indicated that they were significantly associated with gastric cancer using a recessive model: allele (A vs. G), $P=0.047$; genotype (GA+AA vs. GG), $P=0.046$.

When the frequency of rs6505162 AA genotype was used as a reference, individuals carrying the AC and AC+CC genotypes had an increased risk of gastric cancer (crude OR0.46, 95% CI 0.24–0.89 for AC vs. AA, $P=0.02$; crude OR0.5, 95% CI 0.27–0.91 for AC+CC vs. AA, $P=0.02$). This variant deviated from HWE ($\chi^2=2.7$; P value=0.1). The mentioned points are shown in detail in Tables 2, and 3, and Additional file 3: Table S3.

Discussion

Somatic mutations in oncogenes and tumor suppressor genes, in collaboration with environmental exposures such as tobacco and carcinogenic chemicals, can trigger cancer. While hereditary cancers account for only 5–10% of all cases and are associated with germline alterations in oncogenes and tumor suppressor genes, non-hereditary cancer can be predisposed by single-nucleotide polymorphisms (SNPs). These genetic variants, known as “drivers,” primarily cause uncontrolled cellular proliferation and occur in “cancer driver genes” [30]. A recent study identified about 570 cancer driver genes that mediate molecular regulatory networks and changes in the tumor microenvironment. Mutations in these genes, as well as abnormal expression levels, can lead to uncontrolled tumor cell proliferation, invasion, and drug resistance [31].

Some case–control studies have shown a correlation between SNPs and susceptibility to gastric cancer in the Ardabil population. Among these, the relationship between IL-1 β -511 and MTHFR C677T polymorphisms and gastric cancer is notable [32, 33]. However, some SNPs previously reported as gastric cancer-related variants were not associated with this population, including the polymorphism P53 Arg72Pro [34].

Previous studies have suggested an important role of miR-149 in carcinogenesis. This microRNA acts as a tumor suppressor gene in some cancers and an oncogene in others. Upregulation of miR-149 has been reported in AML, prostate cancer, glioblastoma, and melanoma, while it is downregulated in gastric, hepatocellular, renal cell, lung, colorectal, breast, and thyroid cancers, as well as neuroblastoma. Upregulation of miR-149 in tumor cells and gastric cancer cell lines inhibits cell proliferation and induces G0/G1 cell cycle arrest [35].

In our general population, the frequency of the C allele of rs2292832 (a T to C nucleotide change in miR-149) was found to be the lowest (0.51 vs. 0.64 for Iran, 0.7 for Europe, and 0.7 for the world), indicating a significant difference ($P=0.017$). Consistent with this finding, a lower frequency of the C allele ($P=0.001$) and the genotypes TC ($P=0.046$), CC ($P=0.01$), and TC+CC ($P=0.01$) were observed compared to the controls. Based on the reported relationship between a reduced risk of gastric cancer and the rs2292832 C allele in two case–control studies on Chinese and Korean populations, as well as a meta-analysis [36–38], the lower frequency of the C allele in Ardabil could be associated with a higher risk of gastric cancer. However, the absence of deviation from Hardy–Weinberg equilibrium and the lack of association in other studies [39–41] necessitate further research, including cohort investigations, in our population.

The variant rs12983273 C>T is located in the MIR371B and MIR373 genes. It has been reported that hsa-miR-373 is downregulated in the gastric cancer recurrent group [42]. PRDM4, which has been identified as a risk biomarker for gastric cancer patients, is regulated by hsa-miR-373 [43]. In addition to the upregulation of miR373 in gastric adenocarcinoma tissue and gastric carcinoma cell lines, the inhibition of migration and invasion in some gastric cancer cell lines has been observed [44].

There were no reported associations for rs12983273 and female neoplasm, breast cancer, ESCC, Oral Premalignant Lesions (OPL), and esophageal cancer risk [45]. Compared to Iranian, European, and world populations, the T allele has a lower frequency in Ardabil (0.107 vs. 0.136). This difference tends to be significant ($P=0.017$). Consistent with this finding, the frequency of allele T and genotype CT+TT in cases affected with gastric cancer were lower than in controls (P values of 0.02 and 0.031, respectively). Cohort studies are needed to validate the impact of rs12983273 in conferring gastric cancer predisposition in our population.

The variant rs1695 is located on exon 5 of the GSTP1 gene. The substitution of an A nucleotide with a G leads to a missense mutation (Leu>Val). This polymorphism could reduce the detoxification activity of the enzyme.

Furthermore, different genotypes of rs1695 have important interactions with environmental factors, including *Helicobacter pylori* infection, smoking, and alcohol consumption in cases with gastric cancer [46]. While some reports show no association with gastric cancer [47, 48], Val/Val (vs. Ile/Ile) and the G allele have been found to be related to gastric cancer. Additionally, the GA+GG genotype was associated with a larger tumor size [49]. These results do not align with our finding in the general population, which has a lower frequency of the G allele. However, our case–control study revealed that the frequency of the G allele and genotype AG+GG were higher among the cases compared to the controls (*P* values of 0.047 and 0.036, respectively). A meta-analysis found no association between rs1695 and gastric cancer in the Caucasian population (the Ardabil population is part of Caucasians), in contrast to Asians [50]. This report supports our finding in the general population. The contradiction between our population-based and case–control studies can be resolved by conducting a cohort study.

The Phospholipase C epsilon1 gene (PLCE1) is involved in cell growth and differentiation, as well as gene expression. The phospholipase C encoded by PLCE1 induces GTPases and hydrolyzes 1,4,5-phosphatidylinositol to inositol 1,4,5-triphosphate and 4,5-diacylglycerol [51]. Variant rs2274223 is a missense variant (A>G; His>Arg) in the coding region of the PLCE1 gene.

There is a 30% difference in the frequency of the G allele between the Ardabil general population and the world population (0.42 vs. 0.32), which is statistically significant (*P*=0.036). This difference, along with the consistent relationship between rs2274223 and gastric cancer reported previously, could explain the important role of this variant in conferring susceptibility to gastric cancer in Ardabil [52–57]. Additionally, the GG genotype was found to be more frequent in the cases compared to the controls (*P*=0.035).

There were no reported associations for rs12983273 and female neoplasm, breast cancer, ESCC, Oral Premalignant Lesions (OPL), and esophageal cancer risk [45]. Compared to Iranian, European, and world populations, the T allele has a lower frequency in Ardabil (0.107 vs. 0.136). This difference tends to be significant (*P*=0.017). Consistent with this finding, the frequency of allele T and genotype CT+TT in cases affected with gastric cancer were lower than in controls (*P* values of 0.02 and 0.031, respectively). Cohort studies are needed to validate the impact of rs12983273 in conferring gastric cancer predisposition in our population.

The variant rs1695 is located on exon 5 of the GSTP1 gene. The substitution of an A nucleotide with a G leads to a missense mutation (Leu>Val). This polymorphism could reduce the detoxification activity of the enzyme.

Furthermore, different genotypes of rs1695 have important interactions with environmental factors, including *Helicobacter pylori* infection, smoking, and alcohol consumption in cases with gastric cancer [46]. While some reports show no association with gastric cancer [47, 48], Val/Val (vs. Ile/Ile) and the G allele have been found to be related to gastric cancer. Additionally, the GA+GG genotype was associated with a larger tumor size [49]. These results do not align with our finding in the general population, which has a lower frequency of the G allele. However, our case–control study revealed that the frequency of the G allele and genotype AG+GG were higher among the cases compared to the controls (*P* values of 0.047 and 0.036, respectively). A meta-analysis found no association between rs1695 and gastric cancer in the Caucasian population (the Ardabil population is part of Caucasians), in contrast to Asians [50]. This report supports our finding in the general population. The contradiction between our population-based and case–control studies can be resolved by conducting a cohort study.

The Phospholipase C epsilon1 gene (PLCE1) is involved in cell growth and differentiation, as well as gene expression. The phospholipase C encoded by PLCE1 induces GTPases and hydrolyzes 1,4,5-phosphatidylinositol to inositol 1,4,5-triphosphate and 4,5-diacylglycerol [51]. Variant rs2274223 is a missense variant (A>G; His>Arg) in the coding region of the PLCE1 gene.

There is a 30% difference in the frequency of the G allele between the Ardabil general population and the world population (0.42 vs. 0.32), which is statistically significant (*P*=0.036). This difference, along with the consistent relationship between rs2274223 and gastric cancer reported previously, could explain the important role of this variant in conferring susceptibility to gastric cancer in Ardabil [52–57]. Additionally, the GG genotype was found to be more frequent in the cases compared to the controls (*P*=0.035).

The variant rs6194 is a synonymous variant (p.His=) in the NR3C1 gene. Although synonymous variants are generally considered to have no pathogenic effect, some reports suggest that they may play a role in splicing, RNA structure, and translation rate in carcinogenesis processes [6]. Investigating the role of rs6194 in conferring cancer susceptibility is therefore important, as it has been found to create premature stop codons and yield a shorter mRNA via exon skipping [69]. However, despite the high frequency of the T allele in the general population of Ardabil (Ardabil: 0.003; World: 0.0005; *P*=0.028), the very low frequency of this variant in the population of Ardabil means that it lacks diagnostic and prognostic validity. Our case–control study results confirm this issue.

The variant rs63750447 is a missense variant (A>T; Val>Asp) in the mismatch repair gene hMLH1. In Silico analysis, including FATHM, FATHM-MKL, MutationTaster, MutationAssessor, SIFT, Polyphen2 HVAR, PROVEAN, BayesDel_addAF, DANN, DEOGEN2, EIGEN, LIST-S2, and MetaLR, predicts this variant to be deleterious. We found a significant difference in the frequency of the T allele in our general population ($P=0.009$). A meta-analysis and a study in a Chinese population have reported associations between rs63750447 and increased risk of colorectal, endometrial, and gastric cancer [70, 71]. However, another study in a Chinese population did not find any association [72]. Therefore, this variant could be considered a candidate for increasing gastric cancer frequency in our population. The inconsistency between our case-control study results and the general population findings may be due to not deviating from the Hardy-Weinberg equilibrium.

The variant rs1047768 is a silent variant (p.His=) in the ERCC5 gene, which replaces a T nucleotide with a C. ERCC5 is an essential element of the Nucleotide Excision Repair (NER) pathway and encodes an endonuclease that catalyzes 3' incisions [73, 74]. Previous studies have suggested that rs1047768 plays an important role in carcinogenesis, abnormal cell proliferation and differentiation, and increased cancer susceptibility. This variant has also been reported to impact sensitivity to chemotherapy [75]. While some reports have found no association between rs1047768 and gastric cancer [76-79], our results are consistent with a study in a Chinese population, which found a decreased risk for gastric cancer with the TC genotype compared to TT [80]. Therefore, the lower frequency of the C allele in the Ardabil population (0.46 vs. 0.57 for the World) could confer susceptibility to gastric cancer.

The variant rs1050631 is a C to T nucleotide change in exon 6 of the SLC39A4 gene, resulting in a synonymous mutation. Although it is expected to have no impact on protein function, this synonymous mutation may be involved in various pathogenic mechanisms, such as disease susceptibility, treatment outcomes, protein levels, structure and function, RNA processing, post-transcriptional regulation, translation initiation, early and late translation elongation, and co-translational folding [6, 69]. Our current findings reveal a lower allelic frequency of the T nucleotide in the general population of the Ardabil province (0.27 vs. World: 0.35; $P=0.027$). Additionally, the frequencies of CT, TT, and CT+TT genotypes were lower among affected cases compared to healthy controls (P value < 0.0001, 0.019, and < 0.0001, respectively). These results are consistent with a Chinese population, where the CT+TT genotype was associated with a higher risk of recurrence and death in patients with

gastroadenocarcinoma [81]. However, another study did not find any association [82]. Therefore, rs1050631 may be associated with the high rate of gastric cancer in the Ardabil population.

DNA methylation, which is regulated by DNA methyltransferases (DNMTs), plays a crucial role in transcription regulation. Abnormal expression of DNMT1, the first discovered DNMT, has been linked to various types of cancer. In gastric cancer, overexpression of DNMT1 is associated with abnormal differentiation, higher tumor stage, and increased mortality risk [83, 84]. Knockdown of DNMT1 leads to cell cycle arrest, increased apoptosis, decreased invasion, and improved response to chemotherapy. Patients with low DNMT1 expression levels have shown better histopathological and clinical responses. DNMT1 expression is higher in gastric cancer compared to normal, para-cancerous, and dysplasia tissues [85].

The variant rs16999593 is a substitution of His amino acid with Arg (A>G) in exon 4 of the DNMT1 gene. Although no association with gastric cancer has been found in some populations, such as Iranian and Chinese [86, 87], our findings are consistent with a higher frequency of the G allele in the Ardabil population compared to the global population (0.013 vs. 0.002). Strong associations with gastric cancer have been reported in three different meta-analyses, with a 45% increased risk in G allele carriers [76, 88-90].

The variant rs1800469 is located in the promoter region of the TGFB gene. The TGF β 1 polypeptide can have a dual role in gastric carcinogenesis, either inhibiting immunosurveillance and promoting EMT and metastasis, or suppressing tumor growth by inhibiting cell cycle progression [91]. Several studies conducted on Chinese, American, and Polish populations did not find any significant association between this variant and gastric cancer [100]. Additionally, there was no relationship between rs1800469 and pancreatic cancer in the Iranian population [92].

The decreased risk of gastric cancer in carriers of the C allele (genotypes CC and CT compared with TT) in a Chinese population, as well as the significant association found in certain subtypes of gastric cancer including intestinal type, poorly differentiated, and stage TNM, suggests a relationship between rs1800469 and a decrease in gastric cancer risk [93]. These findings indicate that the lower frequency of the C allele in the Ardabil population (0.52 vs. 0.67 worldwide) could be related to a higher frequency of gastric cancer in this geographical location.

The GG genotype of the miR-449 variant rs10061133 showed a predictive effect on esophageal squamous cell carcinoma (ESCC) in a Chinese population [94]. However, no association was found for lung cancer [95]. In our study, we found that Allele G was more frequent in

our general population compared to the worldwide frequency (0.13 vs. 0.091; $P=0.038$). Given the reported correlations with gastric cancer, including downregulation of miR-449, inhibition of cell proliferation, induction of senescence and apoptosis, and the upregulation of CDK6, as well as its high frequency in the Korean population (0.25) with a gastric cancer age-standardized rate (ASR) of 39.6, investigating the role of rs10061133 in gastric carcinogenesis seems to be useful.

It was observed that SPHK1 expression is elevated in gastric cancer tissues and cells. Additionally, SPHK1 knockdown tends to suppress cell proliferation, migration, and invasion of gastric cancer cell lines, as well as block the cell cycle and induce apoptosis [97]. In the Chinese population, higher expression levels of SPHK1 are associated with shorter overall survival time in gastric cancer patients [98]. The variant rs3744037 is a stop-gained variant that leads to premature termination of translation. Although it appears to reduce SPHK1 expression, based on five benign predictions (by in silico analysis including BayesDel_addAF, DANN, EIGEN, FATHMM-MKL, and MutationTaster), and the fact that it is not in a highly conserved position, it may not have a mitigating effect. The higher frequency of this variant in the Ardabil general population (0.2 vs. 0.167 worldwide; $P=0.024$) makes it an attractive subject for investigation. However, the findings of our case-control study did not confirm this conclusion.

A significant association was found for the apoptosis-related gene, BCL2L12, with early-stage (I/II) gastric tumors and the intestinal histotype, as well as disease-free and overall survival of patients [99]. The variant rs3745469 is located in the non-coding region of the BCL2L12 gene and has been found to be more frequent in the Ardabil population (0.24 vs. worldwide: 0.09; $P=0.016$). Our case-control study helped us reveal its role in gastric carcinogenesis. The frequency of allele A and the genotype GA+AA were determined to be higher in our affected cases compared to controls (P value, 0.047, and 0.046, respectively).

In a Chinese population, allele C of rs12220909 has been associated with a decreased risk of esophageal squamous cell carcinoma [94]. Similarly, the frequency of allele C in patients affected by non-small cell lung cancer (NSCLC) was significantly lower than in the Chinese normal population (13.3% in cases vs. 16.9% in controls; $P=0.001$) [95]. Therefore, the low frequency of this variant in the Ardabil general population (0.003 compared to Iran: 0.01, 1000G: 0.0381, and gnomAD: 0.01) could be considered a risk factor for gastric cancer.

Bioinformatics analysis, including RegulomeDB, HaploReg, and GTEx database, found that rs4789936 has an effect on transcription factor binding, motif changes,

DNase footprint, and DNase peaks of the TIMP2 gene. Additionally, it was shown that carriers of the TT genotype have a higher expression level compared to CC carriers, and the risk of this variant is associated with an increased expression level of the gene [100]. The higher frequency of the T allele in the Ardabil population (0.54 vs. 0.5 worldwide; $P=0.012$) could be related to the observed incidence rate of gastric cancer. However, the results of the case-control study showed contradictory results. This contradiction may be due to not deviating from the Hardy-Weinberg equilibrium. Therefore, designing a cohort study could be useful.

TLR4 is highly expressed in monocytes, lymphocytes, and splenocytes. Its product plays a critical role in the inflammatory-related immune response to *H. Pylori* infection in the gastric carcinogenesis process. A meta-analysis detected a significantly increased gastric cancer risk in Caucasian populations for rs4986790 and rs4986791 [100]. Despite our findings in the general population (0.03 in Ardabil vs. 0.06 in Iran and worldwide), the case-control study did not reveal any significant difference.

Three unrelated meta-analyses revealed a significant association between the miR-423 rs6505162 A>C variant and overall cancer risk, including gastric cancer [101–103]. The higher frequency of allele C in our population (0.58 vs. 0.45 worldwide; $P=0.010$), as well as the higher frequencies of the genotypes AC ($P=0.02$) and AC+CC ($P=0.025$), suggest that this variant may be a risk factor for gastric cancer in this region.

PSCA (Prostate Stem Cell Antigen), a member of the LY-6 family of surface proteins, is abundantly expressed in the normal esophagus and stomach but undetectable in esophageal or gastric tumors. PSCA is expressed in differentiating gastric epithelial cells, and in vitro studies have detected its cell proliferation inhibition activity and silencing in gastric cancer [104].

Various SNPs in the PSCA gene have been associated with gastric cancer in Chinese, Korean, Japanese, Tibetan, and Caucasian populations [105–110]. The variant rs2976391 C>A is an intronic variant of the PSCA gene. In a meta-analysis with 81,961 cases affected by gastric cancer and 442,932 healthy controls, a strong association was found for this variant [111]. The frequency of allele A in our population was found to be greater than in other populations (0.57 vs. 0.41 in Iran, 0.4 in gnomAD, and 0.405 in 1000G). This suggests that this variant may play a role in conferring gastric cancer susceptibility in Ardabil province.

The PLCE1 (Phospholipase C epsilon 1) gene is involved in the regulation of cell growth, differentiation, and oncogenesis. The substitution of a C nucleotide with a T at residue 4406 of the PLCE1 gene results

in the substitution of threonine amino acid with isoleucine, known as rs3765524. This variant has been reported to be correlated with an increased risk of gastric cancer in some populations, including Chinese and Kashmir Valley [52, 53]. Despite finding no association in the north of Iran [62], the higher frequency of allele T in Ardabil (0.41 vs. 0.35 in Iran and 0.31 in gnomAD) may be related to the high incidence of gastric cancer.

The frequency of allele C in the DNA repair gene ERCC2 rs3810366 is higher in the Ardabil population compared to other populations (Ardabil: 0.54; Iran: 0.52; 1000G: 0.42). This finding is consistent with studies conducted on southern Chinese and Taiwanese populations (112).

The differences in observed frequencies, whether significant or not, motivate researchers to conduct cohort studies in the Ardabil population. By doing so, we aim to identify prognostic factors that can help diagnose individuals predisposed to gastric cancer in this population.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s43042-024-00474-w>.

Additional file 1. Table S1. The primers used for validation of some of the Whole Exome Sequencing results.

Additional file 2. Table S2. Frequency of the variants associated with gastric cancer susceptibility in Ardabil (general), Azeri, Iranian, and world populations.

Additional file 3. Table S3. Statistical analysis of the selected variants among the cases and controls.

Additional file 4. Fig. S1. Scatter plots for the selected variants considering the allele frequency among different populations based on their gastric cancer ASR.

Additional file 5. Fig. S2. Some of the validated results by Sanger-based PCR-Sequencing.

Acknowledgements

We are grateful to the Department of Genomics, Gastrointestinal Diseases Research Center at Ardabil University of Medical Sciences, for their kind cooperation in the sampling.

Author contributions

HA, NR, and HH contributed to the study design. HA collected the samples, contributed to the laboratory works, and statistical analyses. HA and NR wrote and revised the manuscript. All authors read and approved the final manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The participants provided their written informed consent to be included in the study. The Ethical and regulatory issues related to human specimen

collection for research purposes have been approved by the Ethics Committee of Islamic Azad University-Rasht Branch (Approval ID: IRI.AU.RASHT.REC.1398.057).

Consent for publication

All participants were provided signed informed consent declaring to investigate the peripheral blood samples.

Competing interests

The authors declare that they have no competing interests.

Received: 28 November 2023 Accepted: 11 January 2024

Published online: 27 March 2024

References

- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F (2021) Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J Clin* 71(3):209–249
- Babaei M, Jaafarzadeh H, Sadjadi AR, Samadi F, Yazdanbod A, Fallah M, Aghlmandi S, Ramezani R, Haukka J, Hekmat SH, Didevar R (2009) Cancer incidence and mortality in Ardabil: Report of an ongoing population-based cancer registry in Iran, 2004–2006. *Iran J Public Health* 38(4):35–45
- Derakhshan MH, Yazdanbod A, Sadjadi AR, Shokoohi B, McColl KE, Malekzadeh R (2004) High incidence of adenocarcinoma arising from the right side of the gastric cardia in NW Iran. *Gut* 53(9):1262–1266
- Yang W, Zhang T, Song X, Dong G, Xu L, Jiang F (2022) SNP-target genes interaction perturbing the cancer risk in the post-GWAS. *Cancers* 14(22):5636
- Yan C, Zhu M, Ding Y, Yang M, Wang M, Li G, Ren C, Huang T, Yang W, He B, Wang M (2020) Meta-analysis of genome-wide association studies and functional assays decipher susceptibility genes for gastric cancer in Chinese populations. *Gut* 69(4):641–651
- Sauna ZE, Kimchi-Sarfaty C (2011) Understanding the contribution of synonymous mutations to human disease. *Nat Rev Genet* 12(10):683–691
- Malhotra P, Read GH, Weidhaas JB (2019) Breast cancer and miR-SNPs: The importance of miR germ-line genetics. *Non-coding RNA* 5(1):27
- Preskill C, Weidhaas JB. SNPs in microRNA binding sites as prognostic and predictive cancer biomarkers. *Critical Reviews™ in Oncogenesis*. 2013;18(4).
- Chhichholiya Y, Suryan AK, Suman P, Munshi A, Singh S. SNPs in miRNAs and Target sequences: role in cancer and diabetes. *Front Genet*. 2021;12.
- The 1000 Genomes Project Consortium (2012) An integrated map of genetic variation from 1092 human genomes. *Nature* 491:56–65
- Standards and guidelines for the interpretation of sequence variants. *Genet Med*. 2015;17(5):405–23.
- Ambry Genetics Variant Classification Scheme. <http://www.ambrygen.com/variant-classification>.
- Database of Single Nucleotide Polymorphisms (dbSNP) [Internet]. Bethesda (MD): National Center for Biotechnology Information, National Library of Medicine (dbSNP Build ID:135) Available from: www.ncbi.nlm.nih.gov/SNP. Accessed Jan 2012.
- Exome Variant Server, NHLBI Exome Sequencing Project (ESP) [Internet], Seattle WA. Available from: evs.gs.washington.edu/EVS.
- Online Mendelian Inheritance in Man, OMIM®. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD), Copyright® 1966–2012. World Wide Web URL: <http://omim.org>.
- Exome Aggregation Consortium (ExAC) [Internet], Cambridge, MA. Available from: <http://exac.broadinstitute.org>.
- Genome Aggregation Database (gnomAD) [Internet], Cambridge, MA. Available from: <http://gnomad.broadinstitute.org>.
- Kalia SS, Adelman K, Bale SJ, Chung WK, Eng C, Evans JP, Herman GE, Hufnagel SB, Klein TE, Korf BR, McKelvey KD (2017) Recommendations for reporting of secondary findings in clinical exome and genome

- sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. *Genet Med* 19(2):249–255
19. Saleh-Gohari N, Mohammadi-Anaie M, Kalantari-Khandani B (2012) BRCA1 gene mutations in breast cancer patients from Kerman Province, Iran. *Iran J Cancer Prevent* 5(4):210
 20. Pourmoshir N, Motalleb G, Vallian S (2020) hsa-miR-423 rs6505162 is associated with the increased risk of breast cancer in isfahan central province of Iran. *Cell J (Yakhteh)* 22(Suppl 1):110
 21. Skrypnik D, Mostowska A, Jagodziński PP, Bogdański P (2020) Association of rs699947 (– 2578 C/A) and rs2010963 (– 634 G/C) single nucleotide polymorphisms of the VEGF Gene, VEGF-A and leptin serum level, and cardiovascular risk in patients with excess body mass: a case-control study. *J Clin Med* 9(2):469
 22. Malik MA, Umar M, Gupta U, Zargar SA, Mittal B (2014) Phospholipase C epsilon 1 (PLCE1 rs2274223A> G, rs3765524C> T and rs7922612C> T) polymorphisms and esophageal cancer risk in the Kashmir Valley. *Asian Pac J Cancer Prev* 15(10):4319–4323
 23. Li X, Li K, Wu Z (2015) Association of four common SNPs in microRNA polymorphisms with the risk of hepatocellular carcinoma. *Int J Clin Exp Pathol* 8(8):9560
 24. Alikhani R, Taravati A, Hashemi-Soteh MB (2020) Association of MUC1 5 640G> A and PSCA 5057C> T polymorphisms with the risk of gastric cancer in Northern Iran. *BMC Med Genet* 21(1):1–8
 25. Chen Y, Zhou F, Shen D, He X, Zhang Y, Xu L, Xia Q, Liang S, Song Y, Zuo Y (2016) ERCC5 single nucleotide polymorphism (rs2296147) predicts the risk of acute radiation pneumonitis in lung cancer patients undergoing radiotherapy. *Int J Clin Exp Pathol* 9(11):11868–11875
 26. Yamanaka M, Shiga K, Fujiwara S, Mizuguchi Y, Yasuda S, Ishizawa K, Saiki Y, Higashi K, Ogawa T, Kimura N, Horii A (2018) A novel SDHB IVS2-2A> C mutation is responsible for hereditary pheochromocytoma/paraganglioma syndrome. *Tohoku J Exp Med* 245(2):99–105
 27. Liao X, Han C, Qin W, Liu X, Yu L, Zhu G, Yu T, Lu S, Su H, Liu Z, Chen Z (2017) PLCE1 polymorphisms and expression combined with serum AFP level predicts survival of HBV-related hepatocellular carcinoma patients after hepatectomy. *Oncotarget* 8(17):29202
 28. Tian K, Wang L, Wang K, Ma J, Li D, Yang Y, Jia G, Wu Z, Zhang L, Zhang J (2018) Analysis of variants at LGALS3 single nucleotide polymorphism loci in skull base chordoma. *Oncol Lett* 16(1):1312–1320
 29. Du H, Zhao J, Su Z, Liu Y, Yang Y (2018) Sequencing the exons of human glucocorticoid receptor (NR3C1) gene in Han Chinese with high-altitude pulmonary edema. *J Physiol Anthropol* 37(1):1–5
 30. Martínez-Jiménez F, Muiños F, Sentís I, Deu-Pons J, Reyes-Salazar I, Arnedo-Pac C, Mularoni L, Pich O, Bonet J, Kranas H, Gonzalez-Perez A (2020) A compendium of mutational cancer driver genes. *Nat Rev Cancer* 20(10):555–572
 31. Bossi D, Cicalese A, Dellino GI, Luzzi L, Riva L, D'Alesio C, Diaferia GR, Carugo A, Cavallaro E, Piccioni R, Barberis M (2016) In vivo genetic screens of patient-derived tumors revealed unexpected frailty of the transformed phenotype identification of novel epigenetic essential genes. *Cancer Discov* 6(6):650–663
 32. Hosseini-Asl SS, Pourfarzi F, Barzegar A, Mazani M, Farahmand N, Niasti E, Yazdanbod A, Didevar R, Akhavan H, Malekzadeh R (2013) Decrease in gastric cancer susceptibility by MTHFR C677T polymorphism in Ardabil Province. *Iran Turk J Gastroenterol* 24(2):117–121
 33. Seyedhashemi E, Niasti E, Farahmand N, Mazani M, Yazdanbod A, Amani F, Akhavan H, Hossini-Asl S (2020) The Relationship between IL-1β-511 Polymorphism and Gastric Cancer in Ardabil Province, Iran. *J Ardabil Univ Med Sci* 19(4):410–417
 34. Hosseini-Asl S, Mazani M, Barzegar A, Niasti E, Farahmand N, Akhavan H (2014) The polymorphism P53 Arg72Pro could not confer the susceptibility of gaining gastric cancer in Ardabil Province, Iran. *Govaresh* 19(4):223–230
 35. He Y, Yu D, Zhu L, Zhong S, Zhao J, Tang J (2018) miR-149 in human cancer: a systemic review. *J Cancer* 9(2):375
 36. Zhang L, Liu Q, Wang F (2018) Association between miR-149 gene rs2292832 polymorphism and risk of gastric cancer. *Arch Med Res* 49(4):270–277
 37. Xu Q, Liu JW, Yuan Y (2015) Comprehensive assessment of the association between miRNA polymorphisms and gastric cancer risk. *Mutation Res/Rev Mutation Res* 763:148–160
 38. Ahn DH, Rah H, Choi YK, Jeon YJ, Min KT, Kwack K, Hong SP, Hwang SG, Kim NK (2013) Association of the miR-146aC> G, miR-149T> C, miR-196a2T> C, and miR-499A> G polymorphisms with gastric cancer risk and survival in the Korean population. *Mol Carcinog* 52(51):39–51
 39. Jiang J, Jia ZF, Cao DH, Wu YH, Sun ZW, Cao XY (2016) Association of the miR-146a rs2910164 polymorphism with gastric cancer susceptibility and prognosis. *Future Oncol* 12(19):2215–2226
 40. Pu JY, Dong W, Zhang L, Liang WB, Yang Y, Lv ML (2014) No association between single nucleotide polymorphisms in pre-miRNAs and the risk of gastric cancer in Chinese population. *Iran J Basic Med Sci* 17(2):128
 41. Dikeakos P, Theodoropoulos G, Rizos S, Tzanakis N, Zografos G, Gazouli M (2014) Association of the miR-146aC> G, miR-149T> C, and miR-196a2T> C polymorphisms with gastric cancer risk and survival in the Greek population. *Mol Biol Rep* 41(2):1075–1080
 42. Zhang X, Yan Z, Zhang J, Gong L, Li W, Cui J, Liu Y, Gao Z, Li J, Shen L, Lu Y (2011) Combination of hsa-miR-375 and hsa-miR-142-5p as a predictor for recurrence risk in gastric cancer patients following surgical resection. *Ann Oncol* 22(10):2257–2266
 43. Yan Z, Xiong Y, Xu W, Li M, Cheng Y, Chen F, Ding S, Xu H, Zheng G (2012) Identification of recurrence-related genes by integrating microRNA and gene expression profiling of gastric cancer. *Int J Oncol* 41(6):2166–2174
 44. Shi Y, Shi H, Zhang B, Yan Y, Han X, Jiang W, Qian H, Xu W (2018) miR-373 suppresses gastric cancer metastasis by downregulating vimentin. *Mol Med Rep* 17(3):4027–4034
 45. Bastami M, Choupani J, Saadatian Z, ZununiVahed S, Ouladsahebmadarek E, Mansoori Y, Daraei A, SamadiKafil H, Yousefi B, Mahdipour M, Masotti A (2019) Evidences from a systematic review and meta-analysis unveil the role of MiRNA polymorphisms in the predisposition to female neoplasms. *Int J Mol Sci* 20(20):5088
 46. Zhang Y, Sun LP, Xing CZ, Xu Q, He CY, Li P, Gong YH, Liu YP, Yuan Y. Interaction between GSTP1 Val allele and H. pylori infection, smoking and alcohol consumption and risk of gastric cancer among the Chinese population. 2012:e47178.
 47. García-González MA, Quintero E, Bujanda L, Nicolás D, Benito R, Strunk M, Santolaria S, Sopena F, Badía M, Hijona E, Pérez-Aísa MA (2012) Relevance of GSTM1, GSTT1, and GSTP1 gene polymorphisms to gastric cancer susceptibility and phenotype. *Mutagenesis* 27(6):771–777
 48. Kim J, Kim H, Lee J, Choi JJ, Kim YI, Kim J (2021) Antioxidant-rich diet, GSTP1 rs1871042 polymorphism, and gastric cancer risk in a hospital-based case-control study. *Front Oncol* 10:3240
 49. Xu Z, Zhu H, Luk JM, Wu D, Gu D, Gong W, Tan Y, Zhou J, Tang J, Zhang Z, Wang M (2012) Clinical significance of SOD2 and GSTP1 gene polymorphisms in Chinese patients with gastric cancer. *Cancer* 118(22):5489–5496
 50. Mocellin S, Verdi D, Pooley KA, Nitti D (2015) Genetic variation and gastric cancer risk: a field synopsis and meta-analysis. *Gut* 64(8):1209–1219
 51. Cui XB, Peng H, Li S, Li TT, Liu CX, Zhang SM, Jin TT, Hu JM, Jiang JF, Liang WH, Li N (2014) Prognostic value of PLCE1 expression in upper gastrointestinal cancer: a systematic review and meta-analysis. *Asian Pac J Cancer Prev* 15(22):9661–9666
 52. Liang P, Zhang W, Wang W, Dai P, Wang Q, Yan W, Wang W, Lei X, Cui D, Yan Z. PLCE1 polymorphisms and risk of esophageal and gastric cancer in a northwestern Chinese population. *BioMed Res Int*. 2019;2019.
 53. Malik MA, Srivastava P, Zargar SA, Mittal B (2014) Phospholipase C epsilon 1 (PLCE1) haplotypes are associated with increased risk of gastric cancer in Kashmir Valley. *Saudi J Gastroenterol: Off J Saudi Gastroenterol Assoc* 20(6):371
 54. Zhang H, Jin G, Li H, Ren C, Ding Y, Zhang Q, Deng B, Wang J, Hu Z, Xu Y, Shen H (2011) Genetic variants at 1q22 and 10q23 reproducibly associated with gastric cancer susceptibility in a Chinese population. *Carcinogenesis* 32(6):848–852
 55. Cai M, Dai S, Chen W, Xia C, Lu L, Dai S, Qi J, Wang M, Wang M, Zhou L, Lei F (2017) Environmental factors, seven GWAS-identified susceptibility loci, and risk of gastric cancer and its precursors in a Chinese population. *Cancer Med* 6(3):708–720
 56. Wang M, Zhang R, He J, Qiu L, Li J, Wang Y, Sun M, Yang Y, Wang J, Yang J, Qian J (2012) Potentially functional variants of PLCE1 identified by GWASs contribute to gastric adenocarcinoma susceptibility in an eastern Chinese population. *PLoS ONE* 7(3):e31932

57. Sun H, Wu X, Wu F, Li Y, Yu Z, Chen X, Chen Y, Yang W (2015) Associations of genetic variants in the PSCA, MUC1 and PLCE1 genes with stomach cancer susceptibility in a Chinese population. *PLoS ONE* 10(2):e0117576
58. Wadhwa R, Song S, Lee JS, Yao Y, Wei Q, Ajani JA (2013) Gastric cancer—molecular and clinical dimensions. *Nat Rev Clin Oncol* 10(11):643–655
59. Xie Z, Wang B, Chai Y, Chen J (2020) Estimation of associations between 10 common gene polymorphisms and gastric cancer: evidence from a meta-analysis. *J Clin Pathol* 73(6):318–321
60. Yuan J, Li Y, Tian T, Li N, Zhu Y, Zou J, Gao J, Shen L (2016) Risk prediction for early-onset gastric carcinoma: a case-control study of polygenic gastric cancer in Han Chinese with hereditary background. *Oncotarget* 7(23):33608
61. Li M, Huang L, Qiu H, Fu Q, Li W, Yu Q, Sun L, Zhang L, Hu G, Hu J, Yuan X (2013) Helicobacter pylori infection synergizes with three inflammation-related genetic variants in the GWASs to increase risk of gastric cancer in a Chinese population. *PLoS ONE* 8(9):e74976
62. Shekarriz R, Faghani S, Tafazolli A, Hashemi-Soteh MB (2019) The Correlation between Phospholipase C Epsilon (PLCE1) gene polymorphisms and risk of gastric Adenocarcinoma in Iranian population. *Int J Hematol-Oncol Stem Cell Res* 13(3):108
63. Hwang J, Min BH, Jang J, Kang SY, Bae H, Jang SS, Kim JI, Kim KM (2018) MicroRNA expression profiles in gastric carcinogenesis. *Sci Rep* 8(1):1–8
64. Torruella-Loran I, Ramirez Vina MK, Zapata-Conterras D, Muñoz X, Garcia-Ramallo E, Bonet C, Gonzalez CA, Sala N, Espinosa-Parrilla Y, Epic Gastric Cancer Working Group. rs12416605: C>T in MIR938 associates with gastric cancer through affecting the regulation of the CXCL12 chemokine gene. *Mol Genet Genom Med*. 2019;7(8):e832.
65. Wu Y, Jia Z, Cao D, Wang C, Wu X, You L, Wen S, Pan Y, Cao X, Jiang J. Predictive value of miR-219-1, miR-938, miR-34b/c, and miR-218 polymorphisms for gastric cancer susceptibility and prognosis. *Disease Mark*. 2017;2017.
66. Arisawa T, Tahara T, Shiroeda H, Matsue Y, Minato T, Nomura T, Yamada H, Hayashi R, Saito T, Matsunaga K, Fukuyama T (2012) Genetic polymorphisms of IL17A and pri-microRNA-938, targeting IL17A 3'-UTR, influence susceptibility to gastric cancer. *Hum Immunol* 73(7):747–752
67. Karageorgos I, Mizzi C, Giannopoulou E, Pavlidis C, Peters BA, Zagoriti Z, Stenson PD, Mitropoulos K, Borg J, Kalofonos HP, Drmanac R (2015) Identification of cancer predisposition variants in apparently healthy individuals using a next-generation sequencing-based family genomics approach. *Hum Genom* 9(1):1
68. Ricketts CJ, Forman JR, Rattenberry E, Bradshaw N, Lalloo F, Izatt L, Cole TR, Armstrong R, Kumar VA, Morrison PJ, Atkinson AB (2010) Tumor risks and genotype–phenotype–proteotype analysis in 358 patients with germline mutations in SDHB and SDHD. *Hum Mutat* 31(1):41–51
69. Chen R, Davydov EV, Sirota M, Butte AJ (2010) Non-synonymous and synonymous coding snp show similar likelihood and effect size of human disease association. *PLoS ONE* 5(10):e13574
70. Wu J, Wang D, Song L, Li S, Ding J, Chen S, Li J, Ma G, Zhang X (2011) A new familial gastric cancer-related gene polymorphism: T1151A in the mismatch repair gene hMLH1. *Mol Biol Rep* 38(5):3181–3187
71. Li S, Zheng Y, Tian T, Wang M, Liu X, Liu K, Zhai Y, Dai C, Deng Y, Li S, Dai Z (2017) Pooling-analysis on hMLH1 polymorphisms and cancer risk: evidence based on 31,484 cancer cases and 45,494 cancer-free controls. *Oncotarget* 8(54):93063
72. Jaafar L, Flores-Rozas H (2009) Elucidating the role of human mismatch repair factor hMLH3. *Cancer Biol Ther* 8(14):1421–1423
73. González-Corrochano R, Ruiz FM, Taylor NM, Huecas S, Drakulic S, Spínola-Amilibia M, Fernández-Tornero C (2020) The crystal structure of human XPG, the xeroderma pigmentosum group G endonuclease, provides insight into nucleotide excision DNA repair. *Nucleic Acids Res* 48(17):9943–9958
74. Tsutakawa SE, Sarker AH, Ng C, Arvai AS, Shin DS, Shih B, Jiang S, Thwin AC, Tsai MS, Willcox A, Her MZ (2020) Human XPG nuclease structure, assembly, and activities with insights for neurodegeneration and cancer from pathogenic mutations. *Proc Natl Acad Sci* 117(25):14127–14138
75. Sun XH, Hou WG, Zhao HX, Zhao YL, Ma C, Liu Y (2013) Single nucleotide polymorphisms in the NER pathway and clinical outcome of patients with bone malignant tumor. *Asian Pac J Cancer Prev* 14(3):2049–2052
76. Tian J, Liu G, Zuo C, Liu C, He W, Chen H (2019) Genetic polymorphisms and gastric cancer risk: a comprehensive review synopsis from meta-analysis and genome-wide association studies. *Cancer Biol Med* 16(2):361
77. Hua RX, Zhuo ZJ, Zhu J, Jiang DH, Xue WQ, Zhang SD, Zhang JB, Li XZ, Zhang PF, Jia WH, Shen GP (2016) Association between genetic variants in the XPG gene and gastric cancer risk in a Southern Chinese population. *Aging (Albany NY)* 8(12):3311
78. Liu J, Deng N, Xu Q, Sun L, Tu H, Wang Z, Xing C, Yuan Y (2016) Polymorphisms of multiple genes involved in NER pathway predict prognosis of gastric cancer. *Oncotarget* 7(30):48130
79. Carrera-Lasfuentes P, Lanás A, Bujanda L, Strunk M, Quintero E, Santolaria S, Benito R, Sopena F, Piazuolo E, Thomson C, Pérez-Aisa A (2017) Relevance of DNA repair gene polymorphisms to gastric cancer risk and phenotype. *Oncotarget* 8(22):35848
80. Hussain SK, Mu LN, Cai L, Chang SC, Park SL, Oh SS, Wang Y, Goldstein BY, Ding BG, Jiang Q, Rao J (2009) Genetic variation in immune regulation and DNA repair pathways and stomach cancer in China. *Cancer Epidemiol Prevent Biomark* 18(8):2304–2309
81. Gao J, Ren W, Xiao C, Wang L, Huang Q, Zhang Z, Dang Y, Weng P, Wang H, Fang X, Zhuang M (2019) Involvement of SLC39A6 in gastric adenocarcinoma and correlation of the SLC39A6 polymorphism rs1050631 with clinical outcomes after resection. *BMC Cancer* 19(1):1–5
82. Wang N, Qiao Q, Bao G, Wu T, Li Y, Li J, Lu J, He X (2017) Genetic polymorphisms are associated with the risk of gastric and colorectal cancers in a Han Chinese population. *Oncotarget* 8(17):28805
83. Mutze K, Langer R, Schumacher F, Becker K, Ott K, Novotny A, Hapfelmeier A, Höfler H, Keller G (2011) DNA methyltransferase 1 as a predictive biomarker and potential therapeutic target for chemotherapy in gastric cancer. *Eur J Cancer* 47(12):1817–1825
84. Yang J, Wei X, Wu Q, Xu Z, Gu D, Jin Y, Shen Y, Huang H, Fan H, Chen J (2011) Clinical significance of the expression of DNA methyltransferase proteins in gastric cancer. *Mol Med Rep* 4(6):1139–1143
85. Ma T, Li H, Sun M, Yuan Y, Sun LP (2017) DNMT1 overexpression predicting gastric carcinogenesis, subsequent progression and prognosis: a meta and bioinformatic analysis. *Oncotarget* 8(56):96396
86. Khatami F, Noorinayer B, Ghiasi S, Mohebi R, Hashemi M, Zali MR (2009) Lack of effects of single nucleotide polymorphisms of the DNA methyltransferase 1 gene on gastric cancer in Iranian patients: a case control study. *Asian Pac J Cancer Prev* 10(6):1177–1182
87. Jia Z, Wu X, Cao D, Wang C, You L, Jin M, Wen S, Cao X, Jiang J. Polymorphisms of the DNA methyltransferase 1 gene predict survival of gastric cancer patients receiving tumorectomy. *Disease Mark*. 2016;2016.
88. Yang XX, He XQ, Li FX, Wu YS, Gao Y, Li M (2012) Risk-association of DNA methyltransferases polymorphisms with gastric cancer in the Southern Chinese population. *Int J Mol Sci* 13(7):8364–8378
89. Li H, Li W, Liu S, Zong S, Wang W, Ren J, Li Q, Hou F, Shi Q (2016) DNMT1, DNMT3A and DNMT3B polymorphisms associated with gastric cancer risk: a systematic review and meta-analysis. *EbioMedicine* 13:125–131
90. Neves M, Ribeiro J, Medeiros R, Sousa H (2016) Genetic polymorphism in DNMTs and gastric cancer: a systematic review and meta-analysis. *Porto Biomed J* 1(5):164–172
91. Zhao B, Xu S, Dong X, Lu C, Springer TA (2018) Prodomain–growth factor swapping in the structure of pro-TGF-β1. *J Biol Chem* 293(5):1579–1589
92. Farahbakhsh FB, Mojarad EN, Azimzadeh P, Goudarzi F, Alizadeh AH, Haghazali M, Lamoki RM, Aghdaei HA (2017) TGF-β1 polymorphisms-509 C> T and+ 915 G> C and risk of pancreatic cancer. *Gastroenterol Hepatol Bed Bench* 10(1):14
93. Xu L, Zeng Z, Chen B, Wu X, Yu J, Xue L, Tian L, Wang Y, Chen M, Sung JJ, Hu P (2011) Association between the TGFβ1-509C/T and TGFβR2-875A/G polymorphisms and gastric cancer: a case-control study. *Oncol Lett* 2(2):371–377
94. Zhang P, Wang J, Lu T, Wang X, Zheng Y, Guo S, Yang Y, Wang M, Kolluri VK, Qiu L, Shen F (2015) miR-449b rs10061133 and miR-4293 rs12220909 polymorphisms are associated with decreased esophageal squamous cell carcinoma in a Chinese population. *Tumor Biol* 36(11):8789–8795
95. Fan L, Chen L, Ni X, Guo S, Zhou Y, Wang C, Zheng Y, Shen F, Kolluri VK, Mukhtali M, Zhao Z (2017) Genetic variant of miR-4293 rs12220909 is

- associated with susceptibility to non-small cell lung cancer in a Chinese Han population. *PLoS ONE* 12(4):e0175666
96. Kheir TB, Futoma-Kazmierczak E, Jacobsen A, Krogh A, Bardram L, Hother C, Grønbaek K, Federspiel B, Lund AH, Friis-Hansen L (2011) miR-449 inhibits cell proliferation and is down-regulated in gastric cancer. *Mol Cancer* 10(1):1–2
 97. Wang Z, Qu H, Gong W, Liu A (2018) Up-regulation and tumor-promoting role of SPHK1 were attenuated by miR-330-3p in gastric cancer. *IUBMB Life* 70(11):1164–1176
 98. Li W, Yu CP, Xia JT, Zhang L, Weng GX, Zheng HQ, Kong QL, Hu LJ, Zeng MS, Zeng YX, Li M (2009) Sphingosine kinase 1 is associated with gastric cancer progression and poor survival of patients. *Clin Cancer Res* 15(4):1393–1399
 99. Florou D, Papadopoulos IN, Scorilas A (2010) Molecular analysis and prognostic impact of the novel apoptotic gene BCL2L12 in gastric cancer. *Biochem Biophys Res Commun* 391(1):214–218
 100. Wang K, Wang G, Huang S, Luo A, Jing X, Li G, Zhou Y, Zhao X (2019) Association between TIMP-2 gene polymorphism and breast cancer in Han Chinese women. *BMC Cancer* 19(1):1–9
 101. Pipan V, Zorc M, Kunej T (2015) MicroRNA polymorphisms in cancer: a literature analysis. *Cancers* 7(3):1806–1814
 102. Hu Y, Yu CY, Wang JL, Guan J, Chen HY, Fang JY (2014) MicroRNA sequence polymorphisms and the risk of different types of cancer. *Sci Rep* 4(1):1–8
 103. Chen R, Zheng Y, Zhuo L, Wang S (2017) The association between miR-423 rs6505162 polymorphism and cancer susceptibility: a systematic review and meta-analysis. *Oncotarget* 8(25):40204
 104. Saeki N, Gu J, Yoshida T, Wu X (2010) Prostate stem cell antigen: a Jekyll and Hyde molecule? *Clin Cancer Res* 16(14):3533–3538
 105. Sakamoto H, Yoshimura K, Saeki N, Katai H, Shimoda T, Matsuno Y, Saito D, Sugimura H, Tanioka F, Kato S, Matsukura N. Genetic variation in PSCA is associated with susceptibility to diffuse-type gastric cancer. *Nature genetics*. 2008;40(6).
 106. Matsuo K, Tajima K, Suzuki T, Kawase T, Watanabe M, Shitara K, Misawa K, Ito S, Sawaki A, Muro K, Nakamura T (2009) Association of prostate stem cell antigen gene polymorphisms with the risk of stomach cancer in Japanese. *Int J Cancer* 125(8):1961–1964
 107. Lu Y, Chen J, Ding Y, Jin G, Wu J, Huang H, Deng B, Hua Z, Zhou Y, Shu Y, Liu P (2010) Genetic variation of PSCA gene is associated with the risk of both diffuse-and intestinal-type gastric cancer in a Chinese population. *Int J Cancer* 127(9):2183–2189
 108. Ou J, Li K, Ren H, Bai H, Zeng D, Zhang C (2010) Association and haplotype analysis of prostate stem cell antigen with gastric cancer in Tibetans. *DNA Cell Biol* 29(6):319–323
 109. Song HR, Kim HN, Piao JM, Kweon SS, Choi JS, Bae WK, Chung JJ, Park YK, Kim SH, Choi YD, Shin MH (2011) Association of a common genetic variant in prostate stem-cell antigen with gastric cancer susceptibility in a Korean population. *Mol Carcinog* 50(11):871–875
 110. Sala N, Muñoz X, Travier N, Agudo A, Duell EJ, Moreno V, Overvad K, Tjonneland A, Boutron-Ruault MC, Clavel-Chapelon F, Canzian F (2012) Prostate stem-cell antigen gene is associated with diffuse and intestinal gastric cancer in Caucasians: results from the EPIC-EURGAST study. *Int J Cancer* 130(10):2417–2427
 111. Cui H, Tang M, Zhang M, Liu S, Chen S, Zeng Z, Shen Z, Song B, Lu J, Jia H, Gu D (2019) Variants in the PSCA gene associated with risk of cancer and nonneoplastic diseases: systematic research synopsis, meta-analysis and epidemiological evidence. *Carcinogenesis* 40(1):70–83
 112. He J, Zhuo ZJ, Zhang A, Zhu J, Hua RX, Xue WQ, Zhang SD, Zhang JB, Li XZ, Jia WH (2018) Genetic variants in the nucleotide excision repair pathway genes and gastric cancer susceptibility in a southern Chinese population. *Cancer Manag Res* 10:765

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.