Journal of Translational Medicine

RESEARCH

Open Access

Comprehensive genomic profiling of Chinese lung cancer characterizes germline-somatic mutation interactions influencing cancer risk

Ning Zhou[1†], Yuanyuan Xu[2†], Yumin Huang[3], Guoxiang Ye[4], Liang Luo[5*] and Zuodong Song[6*]

Abstract

Background Germline mutations in numerous genes, particularly tumor suppressor genes, markedly heighten the risk of various cancers, including lung cancer, which is the leading cause of cancer-related deaths worldwide. Despite extensive research on well-known genes like *BRCA1*, *BRCA2*, and mismatch repair genes, many genetic factors contributing to cancer susceptibility remain unidentified.

Methods This study reviewed sequencing data from 4,934 Chinese lung cancer patients. Matched white blood cell samples were sequenced using WES or gene panels to identify germline mutations. Analysis included statistical tests to compare patient demographics, clinical characteristics, and somatic mutation profiles.

Results Among the cohort, 89 patients carried pathogenic/likely pathogenic (P/LP) germline mutations in 20 known cancer susceptibility genes, with *ATM*, *BRCA2*, and *CHEK2* being the most common. *TP53* mutations were linked to early-onset lung cancer, while *ATM* mutations correlated with late-onset and higher PD-L1 expression, suggesting immunotherapy benefits. Germline mutations were more prevalent in younger patients and females. Somatic mutation profiles showed similarities in common mutations but differences in *MTOR* ($p=0.044$) and *MSH6* ($p=0.018$) mutations in P/LP carriers. GO and KEGG analyses indicated distinct biological processes and pathways in patients with P/LP germline mutations. Gene exclusivity analysis revealed correlations and mutual exclusivity between specific germline and somatic mutations. Comparative analysis with the gnomAD database showed a higher prevalence of these mutations in lung cancer patients compared to the general and East Asian populations, suggesting an association with increased lung cancer risk in the Chinese cohort.

Conclusion This study underscores the prevalence of germline mutations in Chinese lung cancer patients, identifying significant associations with clinical characteristics and somatic mutation profiles. The findings highlight the importance of considering germline mutations in lung cancer treatment and the potential benefits of personalized therapy based on genetic susceptibility.

†Ning Zhou and Yuanyuan Xu contributed equally to this work.

*Correspondence:
Liang Luo
28279238@qq.com
Zuodong Song
szd1990@hotmail.com

Full list of author information is available at the end of the article

Keywords Germline mutation, Somatic mutation, Genetic susceptibility, Lung cancer, Cancer risk

Introduction

Germline mutations in over 100 genes, most of which are tumor suppressor genes, significantly increase the risk of various cancers, including breast, ovarian, colorectal cancers, and melanoma [1, 2]. This phenomenon is known as genetic susceptibility. For example, germline mutations in *BRCA1* and *BRCA2* genes lead to susceptibility to breast and ovarian cancers, while mismatch repair (MMR) gene variants are associated with Lynch syndrome-related cancers [3, 4]. Patients with these mutations may exhibit different biological and clinical characteristics and require different treatment approaches. However, these well-known genes account for only a small portion of the genetic burden of cancer, and many genetic alterations that may lead to potential hereditary cancer susceptibilities remain largely unknown.

Lung cancer, the leading cause of cancer-related deaths worldwide, is a multifactorial malignancy driven by environmental exposures, genetic polymorphisms, and somatic and germline mutations [5, 6]. A family history of lung cancer is significantly associated with an increased risk of developing lung cancer in both smokers and non-smokers, suggesting an underlying genetic susceptibility [7–9]. However, well-defined, high-penetrance hereditary lung cancer syndromes are rare. Recent data, mostly from Western populations, indicate that 3.5–8.5% of lung cancer patients harbor pathogenic germline mutations [10, 11]. In addition to *EGFR*, several well-known susceptibility genes such as *ATM*, *BRCA2*, *TP53*, *HER2*, *RET*, *YAP1*, and *CHEK2* have been reported to be associated with lung cancer risk [12–16]. Currently, the standard of care for treating metastatic lung cancer is based on identifying actionable somatic driver gene mutations [17]. There is limited research on the interaction between germline mutations and somatic oncogenic alterations, as well as the overall somatic mutation landscape in the presence of pathogenic germline mutations.

This study aims to assess the prevalence of cancer susceptibility gene mutations in Chinese colorectal cancer patients, identify differences between Chinese and Western patients, and explore the correlation between germline and somatic variations. A comprehensive evaluation of germline mutations in Chinese lung cancer patients could provide evidence to support clinical practice, promote primary prevention, and enhance health benefits for patients and their families.

Materials and methods

Study samples

We retrospectively reviewed sequencing data from 4,934 lung cancer patients who received treatment and genetic testing selection between January 2023 and January 2024. Samples were collected from multiple hospitals across China, predominantly from the eastern and western regions. The patients included had a broad age range of 18 to 94years, with a mean age of 64 years. Gender distribution was balanced, with approximately 58% male and 42% female participants. We aimed to recruit as many lung cancer patients as possible to create a comprehensive and diverse dataset. The primary inclusion criterion was a diagnosis of lung cancer confirmed by institutional records or pathology reports. Exclusion Criteria: (1) Patients with insufficient or poor-quality DNA samples unsuitable for genomic sequencing. (2) Patients with lung cancer who were also diagnosed with other types of cancer. (3) Duplicate or redundant samples identified during quality control processes. The recruited patients provided matched white blood cell (WBC) samples for parallel sequencing to filter germline mutations. Samples were sequenced using WES and capture panels containing 105, 165, or 556 tumor-associated genes (provided by Shanghai Tongshu BioTech Co., Ltd). Among the 4,934 patients, 25 were excluded as no germline variants were detected in the aforementioned tumor-associated genes. Out of the remaining 4,909 patients with germline variants in tumor-associated genes, 89 were identified as carrying likely pathogenic/pathogenic (LP/P) germline variants. The results of this study are not returned to participants or their clinicians for decision-making. Patient sex, age at diagnosis, and pathological characteristics were obtained from medical records, and patients were not selected based on age, sex, or family cancer history. This study was conducted according to the guidelines of the Declaration of Helsinki. It was based on retrospective, de-identified clinical data and received a waiver of patient consent from the Institutional Review Board.

Sequencing analysis and variant annotation

Genomic DNA extracted from formalin-fixed paraffin-embedded (FFPE) tumor samples or blood samples was subjected to NGS using an Illumina NovaSeq 6000 system (Illumina, San Diego, CA, USA) according to the manufacturer's recommendations. Base calling was performed using the Illumina Analysis Pipeline. Low-quality data were removed, and each barcode dataset was separated. Alignment of these sequences to the human reference genome GRCh37/hg19 was performed

using the Burrows-Wheeler Aligner (BWA). Strelka2 was utilized for the detection of single nucleotide variants (SNVs), insertions, and deletions (indels) with default parameters. Detected genomic alterations included SNVs, small indels, copy number variations, and gene fusions. According to the guidelines for sequence interpretation from the American College of Medical Genetics and Genomics (ACMG), the pathogenicity of germline mutations was defined and predicted using a five-tier classification system. As a result, all germline mutations were categorized as pathogenic/likely pathogenic (P/LP+) or non-pathogenic (P/LP-). Tumor mutational burden (TMB) was defined as the number of somatic mutations detected per million bases in the coding region of tumor tissue. The upper quartile TMB value of tumor tissue samples was used as the threshold to distinguish between high and low TMB levels.

Statistical analysis

All statistical analysis was performed using R software. Mann–Whitney test was employed to compare age between groups. The Chi-square test or Fisher's exact test was performed to test frequency between groups. Statistical significance was defined as a two-sided p value of < 0.05.

Results

Germline mutation landscape of Chinese lung cancer patients

In our cohort of 89 patients, we identified 66 pathogenic/likely pathogenic (P/LP) germline variants from 20 known cancer susceptibility genes. Each of the 89 patients carried only one P/LP germline mutation; no patient had two or more P/LP germline mutations. The demographics, clinical, and pathological characteristics of the patients, as well as the prevalence of germline mutations, are shown in Table 1. In this group of Chinese lung cancer patients, the most common mutated gene was *ATM*, found in 17 patients, followed by *BRCA2* in 13 patients, *CHEK2* in 10 patients, *RAD50* in 7 patients, BRIP1 in 6 patients, and so on (Fig. 1A).

Table 1 Clinical characteristics and prevalence of germline mutations

Characteristics	*P*/LP+	*P*/LP-	Prevalence (%)
Number of patients(n)	89	4845	1.804
Age at diagnosis—yrs			$P = 0.597$
Median	67	64	
NA(n)	14	661	
Gender—n			$P = 0.914$
Female	38 (42.7%)	2027 (41.8%)	1.840
Male	51 (57.3%)	2803 (57.9%)	1.787
NA	0 (0)	15 (0.3%)	0
Histologic diagnosis—n			$P = 0.382$
LUAD	42 (47.2%)	2342 (48.3%)	1.762
LUSC	7 (7.9%)	234 (4.8%)	2.905
NOS/Others	40 (44.9%)	2269 (46.8%)	1.732

P/LP, pathogenic/likely pathogenic; NA, not available; NOS, not otherwise specified

TP53 P/LP germline mutations linked to early-onset lung cancer; ATM to late-onset with potential immunotherapy benefit

We then explored whether the clinicopathological characteristics of lung cancer differed in patients with P/LP germline mutations. The results showed that the rates of P/LP germline mutations were similar between lung adenocarcinoma and squamous cell carcinoma patients (Fig. 1B). Interestingly, immunohistochemistry revealed significantly higher PD-L1 expression levels in patients with *ATM* germline mutations (TPS, $P = 0.06$; CPS, $P = 0.02$) (Supplementary Fig. 1), suggesting potential immunotherapy benefits. Additionally, germline mutations appeared more common in female patients across different age groups compared to males (Fig. 1C). Overall, P/LP germline mutations were more prevalent in younger patients and plateaued after age 55. Among our cohort, three patients were identified with *TP53* P/LP germline mutations, with the oldest being 43 years old and the other two under 40, significantly younger than patients without germline mutations (64 years, $p = 0.001$) or those with other germline mutations (67 years, $p < 0.001$). Patients with *BRCA1/2* germline mutations were also younger than those without germline mutations (median age 58 vs. 64 years, $p = 0.147$) or those with other germline mutations (median age 58 vs. 67 years, $p = 0.12$), although these differences were not statistically significant. Conversely, patients with P/LP germline *ATM* ($p < 0.001$), *RAD51D* ($p = 0.001$), and *RNF43* ($p < 0.001$) mutations were older than those without germline mutations (Fig. 1D).

Somatic characteristics of germline P/LP variants carriers

Next, we investigated whether lung cancer patients with P/LP germline mutations exhibited different somatic mutation profiles. The most frequently mutated genes were *EGFR*, *TP53*, and *KRAS* (Fig. 2A). Patients with and without P/LP germline mutations had similar frequencies of these common mutations and tumor mutation burden (TMB) (Fig. 2B). Interestingly, patients with P/LP germline mutations were significantly enriched for *MTOR* ($p = 0.044$) and *MSH6* somatic mutations ($p = 0.018$) (Fig. 2C). Additionally, a multivariable analysis adjusting for sex, age, and histology revealed an odds ratio of 5.081 (0.782–18.827 [95% CI], $p = 0.035$) for *NRAS* mutations (Table 2). Other common key targetable somatic mutations had similar frequencies between patients with and without P/LP germline mutations. These data suggest that while common oncogene

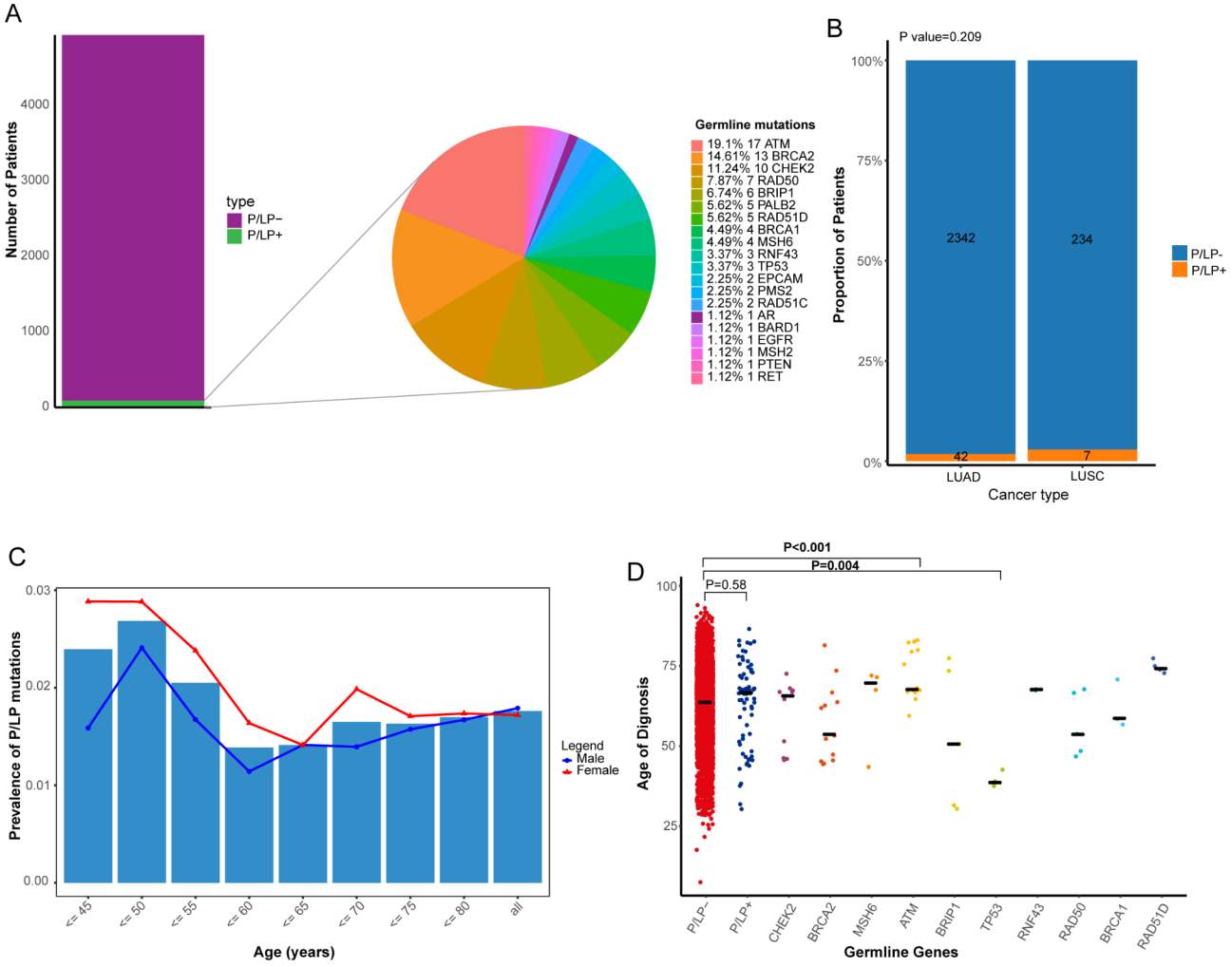

Fig. 1 Distribution of P/LP germline mutations and the age at diagnosis. (**A**) Bar plot indicated the prevalence of P/LP germline mutation (Green). The genes, number of patients, and mutation frequency of each gene are shown in the pie plot. (**B**) Frequency of pathogenic and likely pathogenic germline variants in patients of different ages ($n = 4257$ patients with information on age of onset). (**C**) Bar plot and lines shows the frequency of germline variants in patients under certain age (bar) and frequency in female and male patients (lines). (**D**) The panels show the age of onset for patients without germline mutations ($n = 4182$ patients) and patients with different germline genes ($n = 75$ patients). Horizontal lines indicate median age

mutations are similar in lung cancers with and without P/LP germline mutations, there may be genetic constraints in some patients with cancer susceptibility germline mutations.

To further investigate the differences and mechanisms between the P/LP + and P/LP- groups regarding somatic mutations, we conducted Gene Ontology (GO, standardized vocabulary that describes the functions of genes and gene products in any organism, used to facilitate the annotation and analysis of genomic data) and Kyoto Encyclopaedia of Genes and Genomes (KEGG, which represent networks of molecular interactions and reactions in cells) cluster analyses and compared the results for each group. The KEGG results showed that lung cancer patients with P/LP germline mutations were significantly enriched in "HPV infection," while patients without P/LP germline mutations were significantly enriched in pathways such as "focal adhesion," "proteoglycans in cancer," and "axon guidance" (Supplementary Fig. 2A). The GO results indicated that lung cancer patients with P/LP germline mutations were significantly enriched in biological processes like "reproductive structure development," "reproductive system development," and "epithelial cell proliferation" (Supplementary Fig. 2B). Additionally, we analyzed the differences in clinicopathological features and somatic mutation characteristics between patients with and without germline DNA damage repair (DDR)/MMR mutations (Supplementary Fig. 3). Patients with DDR germline variants also appeared to be associated with early-onset lung cancer, as these patients were significantly younger ($p = 0.007$). Possibly due to the larger proportion of patients with DDR germline variants, the KEGG enrichment results were similar to

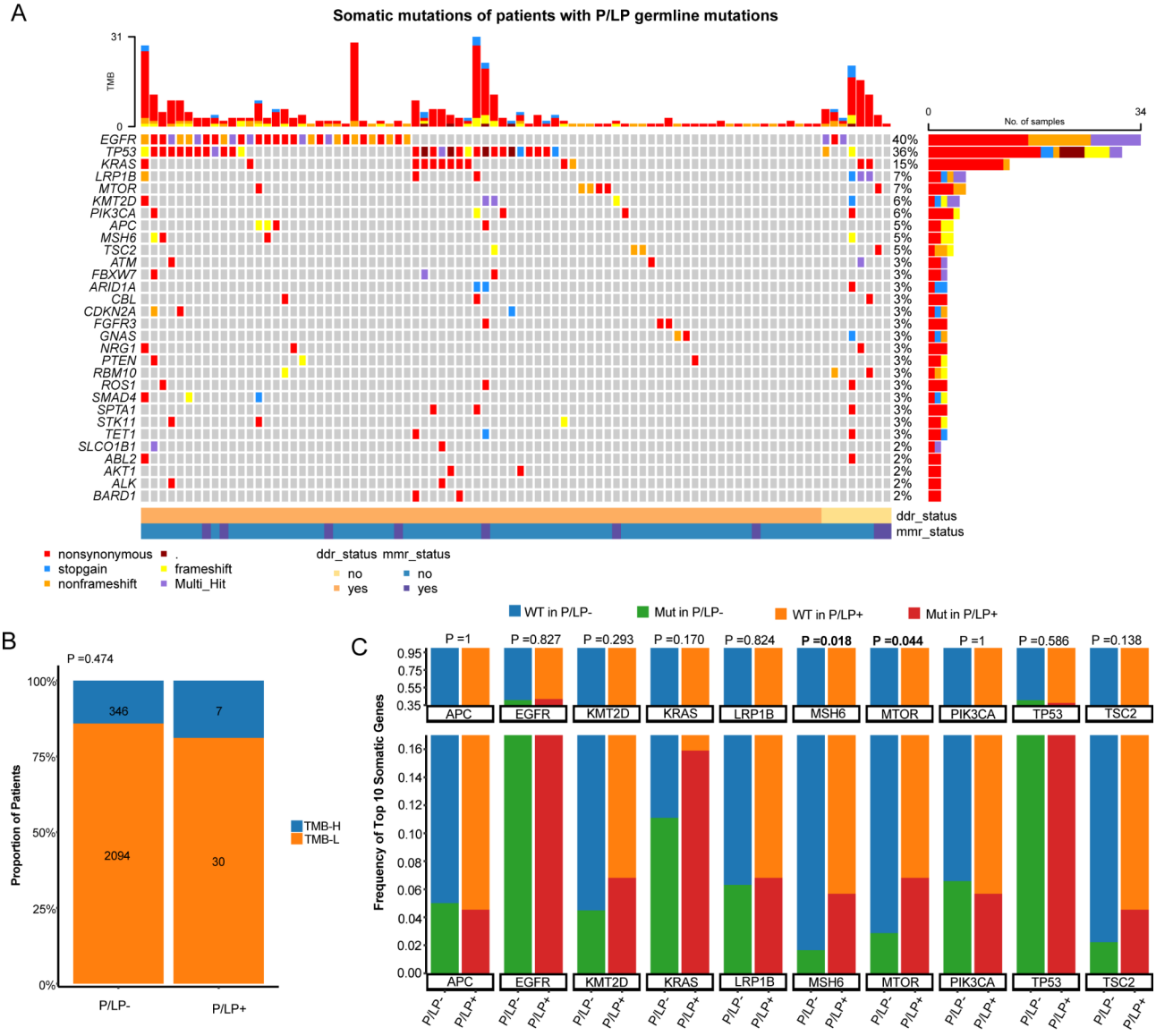

Fig. 2 Common somatic cancer gene mutations in lung cancer patients with and without germline variants. (**A**) Genomic landscape of somatic mutation in patients with P/LP germline mutation. (**B**) The TMB of patients with and without P/LP germline mutations. (**C**) Normalized bar plot illustrates the frequency of top 10 commonly mutated genes. TMB, tumor mutation burden

the overall results. However, KEGG analysis showed no significant differences between the two groups of patients with and without MMR germline variants.

Gene exclusive analysis and potential impact of germline-somatic mutation interactions on lung cancer mutagenesis

Many oncogenes exhibit strong mutual exclusivity or co-occurrence in their mutation patterns, whether germline or somatic. We respectively analyzed the distribution of the top 10 mutated germline and somatic genes among the 89 patients. The results showed mutual exclusivity between germline *ATM* and *BRCA2* variants ($p = 0.057$); for somatic mutations, *KRAS* mutations positively correlated with *TP53* ($p = 0.079$), *RNF43* ($p = 0.014$), and *LRP1B* ($p < 0.001$) mutations but were mutually exclusive with *EGFR* mutations ($p = 0.034$) (Fig. 3A). No patient in our cohort had the same germline mutations among this set of genes (Fig. 3B). We also found some germline-somatic mutation exclusivity or co-occurrence, such as a positive correlation between germline *ATM* mutations and somatic *KRAS* ($p = 0.015$) and *MTOR* ($p = 0.049$) mutations, and mutual exclusivity between germline *RAD50* and somatic *EGFR* mutations ($p = 0.025$).

We further explored whether germline mutations might influence the mutational profile of lung cancer in this cohort. Mutational signature (patterns of somatic alterations in the genome caused by carcinogenic

Table 2 Correlation between germline mutation status and somatic mutations

	Univariate analysis		Multivariate analysis	
	OR (95%CI)	***p***	**OR (95%CI)**	***p***
EGFR mutation				
Age of diagnosis	1.009(1.004–1.014)	**0.001**	1.018 (1.010–1.026)	**< 0.001**
Gender (Female vs. Male)	0.375(0.330–0.426)	**< 0.001**	0.451 (0.378–0.538)	**< 0.001**
Histology (LUSC vs. LUAD)	0.151(0.095–0.230)	**< 0.001**	0.184 (0.115–0.283)	**< 0.001**
Germline mutation (pos vs. neg)	1.112(0.692–1.766)	0.655	1.124 (0.593–2.106)	0.716
KRAS mutation				
Age of diagnosis	1.031(1.022–1.040)	**< 0.001**	1.034 (1.020–1.048)	**< 0.001**
Gender (Female vs. Male)	2.650(2.115–3.349)	**< 0.001**	3.727 (2.705–5.223)	**< 0.001**
Histology (LUSC vs. LUAD)	0.451(0.228–0.802)	**0.012**	0.257 (0.129–0.462)	**< 0.001**
Germline mutation (pos vs. neg)	1.471(0.729-2.700)	0.243	1.464 (0.574–3.263)	0.383
NRAS mutation				
Age of diagnosis	1.032(0.999–1.067)	0.059	1.028 (0.988–1.070)	0.179
Gender (Female vs. Male)	1.627(0.761–3.766)	0.227	2.412 (0.909–7.538)	0.095
Histology (LUSC vs. LUAD)	1.090(0.173–3.811)	0.908	0.694 (0.108–2.537)	0.633
Germline mutation (pos vs. neg)	4.171(0.665–14.276)	0.054	5.081(0.782–18.827)	**0.035**
TP53 mutation				
Age of diagnosis	1.021(1.016–1.026)	**< 0.001**	1.013 (1.005–1.021)	**0.002**
Gender (Female vs. Male)	2.095(1.839–2.388)	**< 0.001**	1.891 (1.576–2.271)	**< 0.001**
Histology (LUSC vs. LUAD)	4.829(3.563–6.620)	**< 0.001**	3.684 (2.692–5.093)	**< 0.001**
Germline mutation (pos vs. neg)	0.843(0.514–1.351)	0.485	0.700 (0.351–1.337)	0.293
BRAF mutation				
Age of diagnosis	1.015(1.001–1.030)	**0.036**	1.004 (0.984–1.025)	0.699
Gender (Female vs. Male)	1.195(0.862–1.671)	0.290	1.277 (0.798–2.065)	0.311
Histology (LUSC vs. LUAD)	0.823(0.316–1.768)	0.652	0.742 (0.280–1.633)	0.499
Germline mutation (pos vs. neg)	0.349(0.020–1.591)	0.297	0.647 (0.036–3.038)	0.669

exposures or aberrant cellular processes) analysis potentially indicated the contributions of certain genes, such as *BRCA1/2* and MMR genes, to tumorigenesis. However, the number of mutations per cancer gene panel sequencing sample was too small for reliable signature analysis. Therefore, we compared SNVs of tumors with germline DDR/MMR gene mutations to those without germline DDR/MMR mutations. We observed that COSMIC mutation signature 3 (associated with BRCA mutations) accounted for 16.9% of SNVs in the DDR group and 4.7% in the MMR group. Similarly, signatures 6 and 15, associated with MMR defects, contributed 3.8% and 12.9% of SNVs in the DDR group, respectively, while signature 6 contributed 21.1% in the MMR group (Supplementary Fig. 4). These data suggest that germline mutations might promote tumorigenesis by inducing specific mutation types. Comprehensive studies at the whole-exome sequencing level are needed to validate these findings.

Comparison to population database

To illustrate the potential association of these P/LP germline mutations with lung cancer in this cohort, we searched the gnomAD database for the prevalence of all identified germline mutations in the general and East Asian populations (Table 3). We found 24 P/LP germline mutations among the same 20 cancer susceptibility genes in the gnomAD, significantly lower than in our lung cancer patient group. By comparing the frequency of germline mutations identified in this study with the variant prevalence in the general and East Asian populations, we calculated the odds ratios (OR) for these germline mutations. Nineteen of the 20 genes had higher P/LP germline mutation allele frequencies in lung cancer patients compared to the general population (with five genes significantly higher), and 18 of the 19 genes had higher frequencies compared to the East Asian population, indicating an enrichment of these germline mutations in lung cancer patients. These data suggest that the P/LP germline mutations identified in this study may be associated with an increased risk of lung cancer development in the Chinese population. We also compared the prevalence of germline mutations in Chinese lung cancer patients with that in Western lung cancer patients by comparing our results with germline mutation data from TCGA lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) cohorts. Overall, the current Chinese lung cancer cohort had a significantly lower germline mutation rate than the TCGA cohort (Supplementary Table 1), with similar results for lung adenocarcinoma, while no significant

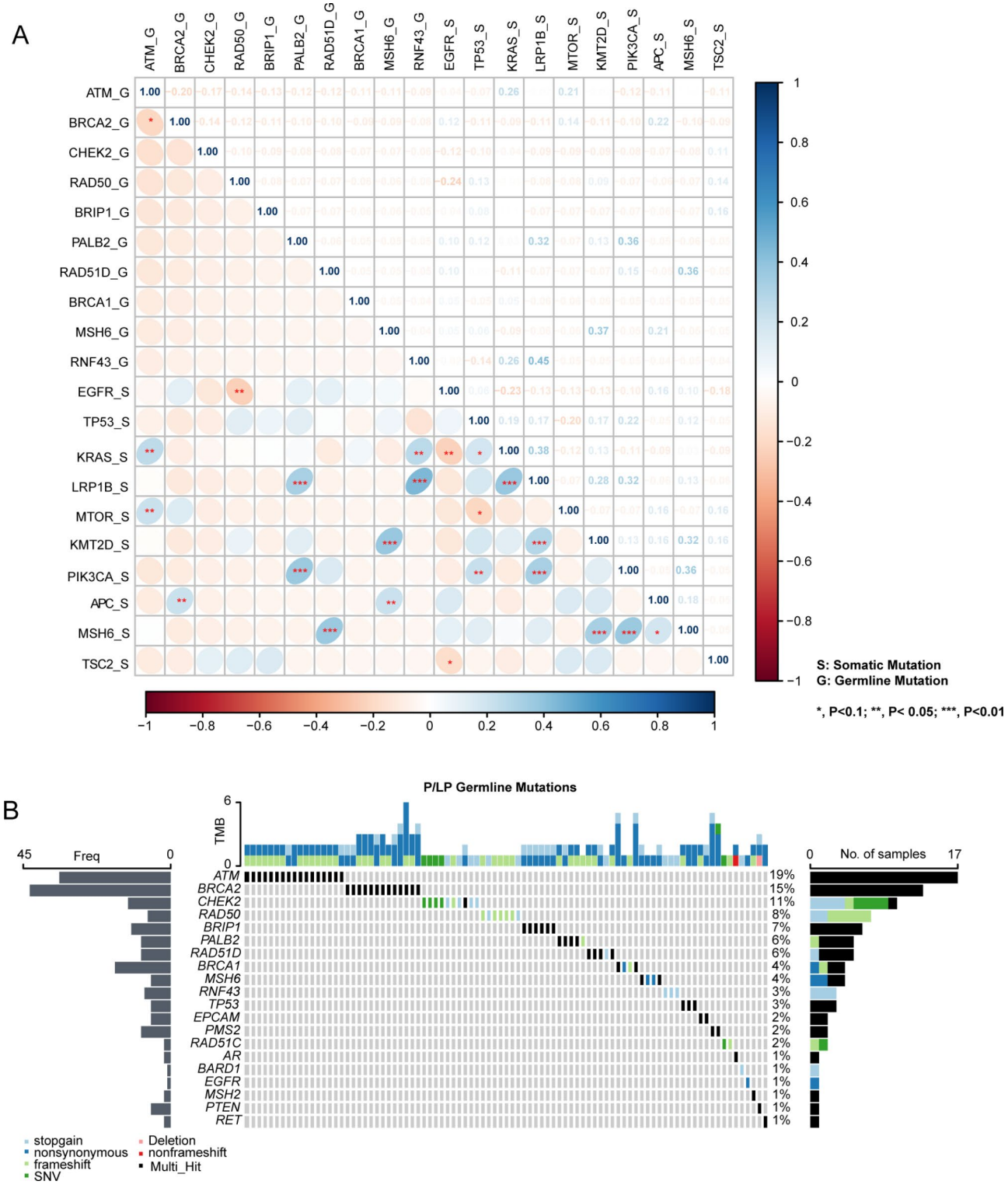

Fig. 3 Mutual exclusive analysis of the selected genes. (**A**) Mutual exclusive analysis of the top 10 mutated germline and somatic genes. The upper right number represents the correlation coefficient (positive numbers represent positive correlation, negative numbers represent negative correlation); asterisks represent significance (*, $P < 0.1$; **, $P < 0.05$; ***, $P < 0.01$) (**B**) Repertoire of P/LP germline genetic alterations of lung cancer in the present cohort

Table 3 Pathogenic/likely pathogenic germline mutations identified in this study

Genes	Variants	TS-AC	TS-AN	TS-AF	General population*			*p* value	East Asian*			*p* value
					AF	OR	95%CI		AF	OR	95%CI	
AR	p.P392R	1	8263	0.0001	1.44E-05	8.41	0.142-161.834	0.1590	/	NA	NA	NA
ATM	p.Tyr1252Ter	1	8263	0.0001	NA	NA	NA	NA	/	NA	NA	NA
ATM	p.Arg2849Ter	1	8263	0.0001	4.07E-06	29.75	0.379-2285.140	0.0640	0	Inf	0.054-Inf	0.3239
ATM	p.F2958fs	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
ATM	p.Glu1267fs	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
ATM	p.K468fs	5	8259	0.0006	4.06E-05	14.90	3.996–47.881	**0.0001**	2.32E-04	2.61	0.562–13.158	0.1603
ATM	p.L2006fs	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
ATM	p.Leu1444fs	3	8261	0.0004	NA	NA	NA	NA	NA	NA	NA	NA
ATM	p.Phe1097fs	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
ATM	p.Pro292Leu	1	8263	0.0001	8.21E-06	14.74	0.249-282.599	0.0952	5.82E-05	2.08	0.027-163.132	0.5439
ATM	p.S2812fs	2	8262	0.0002	1.22E-05	19.82	1.655-173.151	**0.0099**	1.74E-04	1.39	0.116–12.117	0.6621
BARD1	p.R150*	1	8263	0.0001	8.15E-06	14.82	0.251-284.748	0.0945	0	Inf	0.053-Inf	0.3240
BRCA1	p.E879Ter	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
BRCA1	p.Gln858Ter	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
BRCA1	p.N1355fs	1	8263	0.0001	8.14E-06	14.87	0.251–284.930	0.0945	0	Inf	0.053-Inf	0.3241
BRCA1	p.T1691K	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
BRCA2	p.Arg2494Ter	2	8262	0.0002	3.25E-05	7.45	0.770-37.328	**0.0399**	1.74E-04	1.39	0.116–12.146	0.6616
BRCA2	p.E3330*	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
BRCA2	p.A2032fs	2	8262	0.0002	NA	NA	NA	NA	NA	NA	NA	NA
BRCA2	p.Glu1548Ter	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
BRCA2	p.Glu2123fs	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
BRCA2	p.K936fs	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
BRCA2	p.Phe1460fs	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
BRCA2	p.Ser2120Ter	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
BRCA2	p.T17fs(E2)	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
BRCA2	p.Thr219fs	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
BRCA2	p.Tyr1894Ter	1	8263	0.0001	4.08E-06	29.67	0.378-2279.549	0.0641	0	Inf	0.053-Inf	0.3240
BRIP1	p.Q208*	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
BRIP1	p.Arg356Ter	2	8262	0.0002	8.13E-06	29.77	2.158-407.384	**0.0061**	5.80E-05	4.17	0.217-245.923	0.2468
BRIP1	p.Arg798Ter	1	8263	0.0001	1.73E-04	0.70	0.017–4.107	1.0000	0	Inf	0.052-Inf	0.3322
BRIP1	p.Tyr748Ter	2	8262	0.0002	NA	NA	NA	NA	NA	NA	NA	NA
CHEK2	p.Gln487Ter	3	8261	0.0004	1.75E-05	20.81	3.047–122.700	**0.0013**	2.36E-04	1.54	0.225–9.082	0.6904
CHEK2	c.909-1G > A	4	8260	0.0005	NA	NA	NA	NA	NA	NA	NA	NA
CHEK2	p.Y139*	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
CHEK2	p.Gln439Ter	1	8263	0.0001	4.07E-06	29.74	0.378-2284.396	0.0640	0	Inf	0.054-Inf	0.3239
CHEK2	p.Lys373fs	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
EGFR	p.Thr790Met	1	8263	0.0001	2.84E-05	4.26	0.094–33.148	0.2321	0	Inf	0.054-Inf	0.3240
EPCAM	p.Tyr251Ter	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA

Table 3 (continued)

Genes	Variants	TS-AC	TS-AN	TS-AF	General population*			*p* value	East Asian*			*p* value
					AF	OR	95%CI		AF	OR	95%CI	
EPCAM	p.Lys106fs	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
MSH2	p.P476fs	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
MSH6	p.Glu207Ter	1	8263	0.0001	2.84E-05	4.26	0.094–33.148	0.2321	0	Inf	0.054-Inf	0.3240
MSH6	p.His367Arg	2	8262	0.0002	NA	NA	NA	NA	NA	NA	NA	NA
MSH6	p.Glu877Ter	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
PALB2	p.L531fs	2	8262	0.0002	1.83E-04	1.32	0.155–5.072	0.6657	0	Inf	0.392-Inf	0.1049
PALB2	p.K148fs	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
PALB2	p.Gln251Ter	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
PALB2	p.M723fs	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
PMS2	c.2275 + 1G > T	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
PMS2	p.Lys580Ter	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
PTEN	p.Tyr88Ter	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
RAD50	p.Gln755Ter	1	8263	0.0001	3.26E-05	3.72	0.083–27.720	0.2576	2.32E-04	0.52	0.011–5.271	1.0000
RAD50	p.Leu719fs	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
RAD50	p.R1185*	1	8263	0.0001	2.44E-05	4.97	0.107–40.937	0.2064	5.80E-05	2.09	0.027-163.643	0.5429
RAD50	p.Ile1252fs	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
RAD50	p.K722fs	2	8262	0.0002	1.63E-05	14.82	1.340-103.540	**0.0146**	2.32E-04	1.04	0.094–7.279	1.0000
RAD50	p.K994fs	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
RAD51C	c.1025_1026 + 4del	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
RAD51C	c.905–2 A > C	1	8263	0.0001	8.13E-06	14.89	0.252-285.452	0.0943	1.16E-04	1.04	0.018–20.056	1.0000
RAD51D	p.Q151*	1	8263	0.0001	4.06E-06	29.79	0.379-2288.280	0.0639	0	Inf	0.054-Inf	0.3239
RAD51D	p.R232*	1	8263	0.0001	1.66E-05	7.28	0.147–73.603	0.1553	0	Inf	0.053-Inf	0.3275
RAD51D	p.Lys91fs	3	8261	0.0004	NA	NA	NA	NA	NA	NA	NA	NA
RET	p.Val804Met	1	8263	0.0001	1.27E-04	0.96	0.023–5.762	1.0000	6.04E-05	2.00	0.026-157.162	0.5549
RNF43	p.Arg132Ter	3	8261	0.0004	NA	NA	NA	NA	NA	NA	NA	NA
TP53	p.Arg158His	1	8263	0.0001	4.06E-06	29.77	0.379-2287.155	0.0639	0	Inf	0.054-Inf	0.3239
TP53	p.C242F	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA
TP53	p.Gly108fs	1	8263	0.0001	NA	NA	NA	NA	NA	NA	NA	NA

difference was observed in the lung squamous cell carcinoma cohort.

Discussion

Our study provides a comprehensive analysis of the germline mutation landscape in Chinese lung cancer patients, revealing significant insights into genetic susceptibility and the interplay between germline and somatic mutations.

Among the 4,934 patients assessed, 89 were found to carry P/LP germline variants across 20 known cancer susceptibility genes. This prevalence underscores the importance of considering genetic predispositions in lung cancer, a multifactorial disease driven by both environmental and genetic factors. Comparing our cohort to the gnomAD database, we found that the prevalence of P/LP germline mutations was significantly higher in lung cancer patients than in the general and East Asian populations. This enrichment underscores the potential role of these mutations in increasing lung cancer risk. Furthermore, our comparison with Western lung cancer cohorts from TCGA revealed a lower overall germline mutation rate in the Chinese cohort, particularly for lung adenocarcinoma, while no significant difference was observed for lung squamous cell carcinoma. This suggests potential ethnic and regional variations in genetic susceptibility to lung cancer. However, despite the lower overall prevalence, the spectrum of mutated genes, including *ATM*, *BRCA2*, and *CHEK2*, aligning with findings from Western populations [18], highlights shared genetic risk factors across different populations.

Our data reveal distinct patterns in the age and sex distribution of patients with germline mutations. Notably, P/LP germline mutations are more prevalent in younger patients, particularly those with *TP53* and *BRCA2* mutations associated with early-onset lung cancer [19]. Although there were only three patients with *TP53* germline variants, their average age was just 40 years. This finding aligns with previous research indicating that *TP53* germline mutations are linked to early-onset cancers in Li-Fraumeni syndrome patients [20, 21]. If validated, these data support the necessity for earlier lung cancer screening in similar high-risk populations. Conversely, mutations in genes like *ATM*, *RAD51D*, and *RNF43* are more common in older patients, suggesting different pathways of carcinogenesis influenced by age-related factors. Additionally, patients with *ATM* germline mutations exhibited significantly higher PD-L1 expression levels, suggesting potential benefits from immunotherapy for this subgroup. This aligns with emerging evidence that certain germline mutations may influence the tumor microenvironment and response to treatment [22, 23].

Our analysis of somatic mutation profiles in patients with germline mutations revealed several intriguing findings. Previous lung cancer studies have reported germline mutations primarily in *EGFR*, mainly due to the close association between TKI use and *EGFR* mutations [12, 24]. However, *EGFR* mutations are not typically associated with hereditary cancer, and population studies have shown that *EGFR* germline mutations are uncommon in lung cancer, despite reports of *EGFR* germline mutations at multiple sites [24]. In our study, we did not detect any pathogenic *EGFR* germline variants, although *EGFR* remained the most frequently somatic mutated gene. We also found that *EGFR* somatic mutations were mutually exclusive with *KRAS* mutations, consistent with previous reports [25, 26]. Additionally, *EGFR* somatic mutations were found to be mutually exclusive with *RAD50* germline variants.

While common somatic mutations such as *EGFR*, *TP53*, and *KRAS* had similar distributions between patients with and without P/LP germline variants, we observed a significant enrichment of *MTOR* and *MSH6* somatic mutations in the germline mutation group. This suggests that potential interactions between germline and somatic alterations may influence lung cancer development and progression. The mutual exclusivity and co-occurrence patterns observed between certain germline and somatic mutations further highlight the complex interplay between germline and somatic genetic changes. For example, the positive correlation between germline *ATM* mutations and somatic *KRAS* and *MTOR* mutations may indicate specific pathways that are particularly susceptible to disruption in the presence of these germline alterations.

Mutation signature analysis further supports the impact of germline mutations on the lung cancer mutational landscape. For instance, COSMIC mutation signature 3, associated with *BRCA* mutations, was more prevalent in tumors with germline *BRCA1/2* mutations, while signatures associated with MMR defects were more common in tumors with germline MMR mutations. These findings suggest that specific germline mutations may drive tumorigenesis through distinct mutational processes. Although our data are limited by the small number of mutations in panel sequencing, these association-based studies have already become the basis for guideline development and are valuable in determining strategies for screening and preventing certain cancers. Additionally, the significant enrichment of pathways related to reproductive system development and epithelial cell proliferation in the germline mutation group indicates that these mutations may lead to lung cancer through specific biological mechanisms. The enrichment of the HPV infection pathway in this group also highlights the potential role of viral factors

in modulating cancer risk in genetically susceptible individuals.

Differences between Chinese and Westerners may be influenced not only by genetic factors, but also by environmental and lifestyle factors. Higher rates of *EGFR* mutations in Asian lung cancer patients compared to Western populations may be partially attributed to environmental and lifestyle differences, such as lower smoking prevalence among Asian women and higher exposure to indoor pollutants [27]. Research indicates that air pollution significantly contributes to lung cancer incidence in East Asia, potentially amplifying mutation rates in individuals with certain genetic susceptibilities [28]. Nonsmokers with lung cancer—more common in East Asia—tend to show better responses to targeted therapies like EGFR-TKIs [29]. These factors underscore the need for further research to disentangle the effects of genetics, environment, and lifestyle on lung cancer risk and treatment outcomes.

Despite the robust findings, our study has other limitations. As a retrospective real-world data mining study, many patients' clinical information is missing, such as smoking history and environmental exposures, which are well-established contributors to lung cancer etiology and mutation rates. These limitations prevent us from exploring potential interactions between germline mutations and these external factors, which could provide a more nuanced understanding of lung cancer risk. This study maximized inclusivity, however, certain underrepresented regions or populations may introduce bias in the genetic analysis.

Our study highlights germline mutation profiles in Chinese lung cancer patients, providing a basis for understanding genetic predispositions. However, translating these findings into clinical applications was beyond the scope of this work. A key limitation is the lack of clinical or treatment-related data, such as therapy regimens or response outcomes. For instance, while mutations in genes such as *ATM* or *BRCA2* may suggest sensitivity to DNA-damaging agents or immunotherapies, these associations require validation in future studies integrating clinical and functional data. Future research should incorporate prospective cohorts with detailed therapeutic outcomes analyses to explore whether germline mutations can guide personalized treatments.

Finally, in this study, most patients were tested using custom panels. Expanding the scope to include whole genome or whole exome sequencing (WES) could provide deeper insights into the genetic landscape and uncover new susceptibility genes. However, the gene panels were carefully designed to include genes with strong evidence of association with lung cancer based on prior studies and clinical relevance, which prioritize genes known to harbor mutations associated with lung cancer susceptibility (*EGFR*, *TP53*, *KRAS*) and other germline mutations with actionable or predictive significance. The targeted panels provide high-depth sequencing, which ensures superior sensitivity for detecting low-frequency variants compared to WES. The high coverage enhances the accuracy of variant detection, particularly for rare germline mutations, which is critical in the context of our study. Given the large sample size, targeted gene panels offer a more practical and cost-efficient solution than WES. WES would significantly increase the computational and financial resources required for analysis, which could limit the study's feasibility within a large, diverse cohort. Our primary aim was to identify germline mutations in genes known to be associated with lung cancer and correlate them with cancer risk. A targeted approach aligns directly with this objective, allowing us to focus on mutations with established clinical or biological significance. While WES captures a broader range of genetic variations, it can lead to increased noise and challenges in distinguishing variants of uncertain significance (VUS) from pathogenic mutations. The targeted panel's focused approach minimizes this risk and ensures a clearer interpretation of the findings.

All above prevents us from exploring some crucial questions, such as the impact of P/LP germline mutations on mutation characteristics, treatment response, and prognosis. Future research should aim to validate these findings through prospective studies and functional analyses to elucidate the exact mechanisms by which germline mutations promote lung cancer development.

In summary, our study provides a comprehensive overview of the germline mutation landscape in Chinese lung cancer patients, highlighting significant associations with somatic mutations and potential clinical implications. These findings emphasize the importance of considering germline mutations in lung cancer risk assessment and treatment strategies, particularly in diverse populations. Further studies are warranted to validate these results and explore the underlying mechanisms driving these genetic interactions in lung cancer.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12967-025-06096-z.

Supplementary Material 1

Acknowledgements

We thank Shanghai Tongshu Biotechnology Co., Ltd. for technical support.

Author contributions
Zuodong Song and Liang Luo conceived and designed the study. Ning Zhou, Yuanyuan Xu, Yumin Huang and Guoxiang Ye collected data. Ning Zhou and Yuanyuan Xu analyzed, interpreted data and drafted the manuscript. Zuodong Song and Liang Luo revised the manuscript for important intellectual content. All authors participated in manuscript writing and approved the final version of the manuscript.

Funding
This study was supported by the Health Commission of Mianyang City (No: 202055).

Data availability
To ensure transparency and reproducibility of our research, we will upload all relevant data to a public repository with controlled access prior to the article's pre-acceptance (or final acceptance).

Declarations

Ethics approval and consent to participate
This study was conducted in accordance with the guidelines of the Helsinki Declaration. It was based on retrospective, unidentifiable clinical data and obtained institutional review board exemption for patient consent.

Consent for publication
Not applicable.

Competing interests
The authors have no conflicts of interest to declare.

Author details
[1]Department of Pathology, Mianyang 404 Hospital, Mianyang, Sichuan Province, China
[2]Department of Oncology Surgery, Shanghai Chest Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, China
[3]Respiratory department of Yangzhou University Affiliated Hospital, Yangzhou, Jiangsu Province, China
[4]Department of Oncology, Yangzhou Friendliness Hospital, Yangzhou, Jiangsu Province, China
[5]Departments of General Surgery, Mianyang 404 Hospital, Mianyang, Sichuan Province, China
[6]Department of Oncology, Shanghai Lung Cancer Center, Shanghai Chest Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

Received: 4 September 2024 / Accepted: 8 January 2025
Published online: 18 February 2025

References

1. Rahman N. Realizing the promise of cancer predisposition genes. Nature. 2014;505(7483):302–8.
2. Oh M, McBride A, Yun S, et al. BRCA1 and BRCA2 gene mutations and colorectal Cancer risk: systematic review and Meta-analysis. J Natl Cancer Inst. 2018;110(11):1178–89.
3. Kuchenbaecker KB, Hopper JL, Barnes DR, et al. Risks of breast, ovarian, and contralateral breast Cancer for BRCA1 and BRCA2 mutation carriers. JAMA. 2017;317(23):2402–16.
4. de la Chapelle A. Microsatellite instability. N Engl J Med. 2003;349(3):209–10.
5. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424.
6. Alberg AJ, Brock MV, Ford JG et al. Epidemiology of lung cancer: diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. Chest 2013; 143 (5 Suppl): e1S-e29S.
7. Schwartz AG, Yang P, Swanson GM. Familial risk of lung cancer among nonsmokers and their relatives. Am J Epidemiol. 1996;144(6):554–62.
8. Brownson RC, Alavanja MC, Caporaso N, et al. Family history of cancer and risk of lung cancer in lifetime non-smokers and long-term ex-smokers. Int J Epidemiol. 1997;26(2):256–63.
9. Rooney A. Family history reveals lung-cancer risk. Lancet Oncol. 2003;4(5):267.
10. Lu C, Xie M, Wendl MC, et al. Patterns and functional implications of rare germline variants across 12 cancer types. Nat Commun. 2015;6:10086.
11. Huang KL, Mashl RJ, Wu Y, et al. Pathogenic germline variants in 10,389 adult cancers. Cell. 2018;173(2):355–70. e314.
12. Lu S, Yu Y, Li Z, et al. EGFR and ERBB2 germline mutations in Chinese Lung Cancer patients and their roles in genetic susceptibility to Cancer. J Thorac Oncology: Official Publication Int Association Study Lung Cancer. 2019;14(4):732–6.
13. Xiong D, Wang Y, Kupert E, et al. A recurrent mutation in PARK2 is associated with familial lung cancer. Am J Hum Genet. 2015;96(2):301–8.
14. Chen HY, Yu SL, Ho BC, et al. R331W missense mutation of Oncogene YAP1 is a germline risk allele for lung Adenocarcinoma With Medical Actionability. J Clin Oncol. 2015;33(20):2303–10.
15. Wang Y, McKay JD, Rafnar T, et al. Rare variants of large effect in BRCA2 and CHEK2 affect risk of lung cancer. Nat Genet. 2014;46(7):736–41.
16. Esai Selvan M, Klein RJ, Gumus ZH, Rare. Pathogenic germline variants in Fanconi Anemia genes increase risk for squamous Lung Cancer. Clin cancer Research: Official J Am Association Cancer Res. 2019;25(5):1517–25.
17. Hirsch FR, Suda K, Wiens J, et al. New and emerging targeted treatments in advanced non-small-cell lung cancer. Lancet. 2016;388(10048):1012–24.
18. Mukherjee S, Bandlamudi C, Hellmann MD, et al. Germline pathogenic variants Impact Clinicopathology of Advanced Lung Cancer. Cancer Epidemiol Biomarkers Prevention: Publication Am Association Cancer Res Cosponsored Am Soc Prev Oncol. 2022;31(7):1450–9.
19. Peng W, Li B, Li J, et al. Clinical and genomic features of Chinese lung cancer patients with germline mutations. Nat Commun. 2022;13(1):1268.
20. Park KJ, Choi HJ, Suh SP, et al. Germline TP53 mutation and clinical characteristics of Korean patients with Li-Fraumeni Syndrome. Annals Lab Med. 2016;36(5):463–8.
21. Amadou A, Achatz MIW, Hainaut P. Revisiting tumor patterns and penetrance in germline TP53 mutation carriers: temporal phases of Li-Fraumeni syndrome. Curr Opin Oncol. 2018;30(1):23–9.
22. Trigos AS, Pasam A, Banks P et al. Tumor immune microenvironment of primary prostate cancer with and without germline mutations in homologous recombination repair genes. J Immunother Cancer 2022; 10 (6).
23. Das A, Tabori U, Sambira Nahum LC, et al. Efficacy of Nivolumab in Pediatric cancers with High Mutation Burden and Mismatch Repair Deficiency. Clin cancer Research: Official J Am Association Cancer Res. 2023;29(23):4770–83.
24. Shukuya T, Patel S, Shane-Carson K, et al. Lung Cancer patients with germline mutations detected by Next-Generation sequencing and/or Liquid Biopsy. J Thorac Oncology: Official Publication Int Association Study Lung Cancer. 2018;13(2):e17–9.
25. Gazdar A, Robinson L, Oliver D, et al. Hereditary lung cancer syndrome targets never smokers with germline EGFR gene T790M mutations. J Thorac Oncology: Official Publication Int Association Study Lung Cancer. 2014;9(4):456–63.
26. Lan Y, Liu W, Zhang W, et al. Transcriptomic heterogeneity of driver gene mutations reveals novel mutual exclusivity and improves exploration of functional associations. Cancer Med. 2021;10(14):4977–93.
27. Liam CK, Pang YK, Poh ME. EGFR mutations in Asian patients with advanced lung adenocarcinoma. J Thorac Oncol. 2014;9(9):e70–71.
28. Berg CD, Schiller JH, Boffetta P, et al. Air Pollution and Lung Cancer: a review by International Association for the Study of Lung Cancer Early Detection and Screening Committee. J Thorac Oncol. 2023;18(10):1277–89.
29. Morgensztern D, Politi K, Herbst RS. EGFR mutations in Non-small-cell Lung Cancer: find, divide, and conquer. JAMA Oncol. 2015;1(2):146–8.

Publisher's note