



Common inherited variants of *PDCD1*, *CD274* and *HAVCR2* genes differentially modulate the risk and prognosis of adenocarcinoma and squamous cell carcinoma

Nafeesa Moksud¹ · Marta Wagner¹ · Konrad Pawelczyk² · Irena Porębska³ · Beata Muszczyńska-Bernhard² · Aneta Kowal² · Andrzej Wiśniewski⁴ · Monika Kosacka³ · Julia Kończak⁵ · Paweł Karpiński⁶ · Dominik Frydryk⁵ · Anna Andrzejczak¹ · Lidia Karabon¹ · Piotr Kuśnierczyk⁴ · Monika Jasek¹

Received: 21 December 2022 / Accepted: 25 January 2023
© The Author(s) 2023

Abstract

Background To investigate the association between single nucleotide polymorphisms (SNPs) of *PDCD1*, *CD274*, and *HAVCR2* genes with the risk and outcomes of non-small cell lung cancer (NSCLC) subtypes: squamous cell lung cancer (LUSC) and lung adenocarcinoma (LUAD).

Methods TaqMan SNP genotyping assays or polymerase chain reaction-restriction fragment length polymorphism methods were used to determine genotypes of: *PDCD1*: rs36084323, rs7421861, rs11568821, rs2227981, rs10204525; *CD274*: rs822335, rs10815225, rs17718883, rs2297136, rs4742098, rs4143815; *HAVCR2*: rs10057302, rs1036199. Among 383 NSCLC patients, 112 were diagnosed with LUAD and 116 with LUSC. The control group consisted of 433 unrelated, cancer-free subjects.

Results A CC genotype of rs4143815 and GG genotype of rs4742098 were associated with two times higher risk of developing LUSC (CC vs. GG + GC, OR = 2.31; 95% CI = 1.32, 4.06; $P = 0.003$; GG vs. AA + AG, OR = 2.26; 95% CI = 1.17, 4.36; $P = 0.016$, respectively). Moreover, rs4143815 was an independent predictor of the age at diagnosis of LUAD. The carriers of C allele were diagnosed 4.81 years later (95% CI = 1.47, 8.15; $P = 0.006$) than patients with the GG genotype. The rs10057302 CA genotype was an independent predictor of overall survival in LUSC (adjusted HR = 0.13; 95% CI = 0.02, 0.93; $P = 0.043$). NSCLC carriers of rs11568821 T allele had almost double the risk of death (adjusted HR = 2.05; 95% CI = 1.28, 3.29; $P = 0.003$) compared to carriers of CC genotype.

Conclusions Our results provided additional evidence that SNPs of genes for PD-1, PD-L1 and TIM-3 differentially modulate the risk and prognosis of LUSC and LUAD.

Keywords SNPs · *PDCD1* · *CD274* · *HAVCR2* · NSCLC

Abbreviations

NSCLC Non-small cell lung cancer
LUAD Lung adenocarcinoma
LUSC Lung squamous cell carcinoma

PD-1 Programmed cell death protein 1
PD-L1 Programmed death ligand 1
mAb Monoclonal antibody
ICIs Immune checkpoint inhibitors

✉ Monika Jasek
monika.jasek@hirsfeld.pl

¹ Laboratory of Genetics and Epigenetics of Human Diseases, Hirsfeld Institute of Immunology and Experimental Therapy, Polish Academy of Sciences, Rudolfa Weigla 12, 53-114 Wrocław, Poland

² Lower Silesian Centre of Oncology, Pulmonology and Hematology, Wrocław, Poland

³ Department of Pulmonology and Lung Oncology, Wrocław Medical University, Wrocław, Poland

⁴ Laboratory of Immunogenetics and Tissue Immunology, Hirsfeld Institute of Immunology and Experimental Therapy, Polish Academy of Sciences, Wrocław, Poland

⁵ Department of Genetics and Cell Physiology, University of Wrocław, Wrocław, Poland

⁶ Department of Genetics, Wrocław Medical University, Wrocław, Poland

IRs	Immune checkpoint receptors
TIM-3	T-cell immunoglobulin domain and mucin domain-3
SNPs	Single nucleotide polymorphisms
PCR–RFLP	Polymerase chain reaction–restriction fragment length polymorphism
LD	Linkage disequilibrium
eQTL	Expression quantitative trait loci
GTEX	The Genotype-tissue expression
UALCAN	The University of Alabama at Birmingham Cancer Data Analysis Portal
95% CI	95% Confidence interval
HR	Hazard ratios
KM curve	Kaplan–Meier curve
OS	Overall survival
aHR	Adjusted hazard ratio
mRNA	Messenger RNA
IHC	Immunohistochemistry
NHLF	Normal human lung fibroblast
PFS	Progression-free survival
MAF	Minor allele frequency

Introduction

Non-small cell lung cancer (NSCLC) constitutes a heterogeneous group of cancers with three major histological subtypes: lung adenocarcinoma (LUAD) (40%), lung squamous cell carcinoma (LUSC) (25–30%) and large cell carcinoma (10%). LUAD and LUSC display many distinctive features related to histopathology, genomic and genetic architecture (Relli et al. 2019) and disease course, including tumour differentiation, stage-specific survival, underlying drivers and response to various therapeutic options (Faruki et al. 2017). Relli and colleagues have summarised the main differences between LUAD and LUSC, highlighting that the impact of prognostic determinants was frequently partially masked or even hidden when LUSC and LUAD were analysed together as NSCLC (Relli et al. 2019).

Immune checkpoint receptors such as PD-1 (programmed cell death protein 1, CD279; gene *PDCDI*) along with its ligands play a vital role in regulating the immune response. Under physiological conditions, immune checkpoints are a critical mechanism responsible for maintaining immunological homeostasis and preventing the development of autoimmune diseases (Lucibello et al. 2021; Andrews et al. 2019). In cancer, tumour cells employ immune checkpoint pathways to evade anti-tumour immune responses by the host's immune system (Andrews et al. 2019).

The introduction of immunotherapy into clinical practice represents a significant development in cancer treatment, with the PD-1/PD-L1 (programmed death ligand 1; CD274; gene *CD274*) pathway being a primary target of this strategy in

several cancers, including NSCLC. It has been shown that treatment with a single monoclonal antibody (mAb) directed against PD-1 or PD-L1 allowed a durable clinical response in about 18% of patients with advanced NSCLC (Datar et al. 2019). Despite the significant contribution of immune checkpoint inhibitors (ICIs) towards improving the outcomes of some patients, many others are resistant or develop adaptive resistance to immunotherapy or suffer disease progression despite treatment (Andrews et al. 2019; Curigliano et al. 2021). As a result, there is an urgent need for novel therapies or combinational strategies employing the blockade of other immune checkpoint receptors (IRs), such as TIM-3 (T-cell immunoglobulin domain and mucin domain-3, CD366, gene *HAVCR2*).

TIM-3 is frequently co-expressed with PD-1 (Curigliano et al. 2021), and its upregulation in NSCLC is considered a mechanism of adaptive resistance to PD-1 blockade (Andrews et al. 2019). Preclinical studies on animal models demonstrated that therapy involving simultaneous blockade of TIM-3 and PD-1 pathways improved anti-tumour response and effectively suppressed tumour growth, compared to when only one pathway was targeted (Andrews et al. 2019; Curigliano et al. 2021). These observations led to many clinical trials investigating the effectiveness of combining PD-1 and TIM-3 blockade in solid tumours (Acharya et al. 2020).

Genetic variants of genes for PD-1, PD-L1 and TIM-3 molecules have been investigated in several cancers to determine their potential association with disease risk and outcomes. Their association with response to immunotherapy and other types of treatment has also been studied, particularly for polymorphisms located in *PDCDI* and *CD274* genes. However, most of these studies were conducted mainly in East Asian cohorts. Since genotype frequencies and allele distribution of these variants differ significantly in various populations, more studies are needed to establish population-specific and “general” variants associated with risk, outcomes and response to treatment (Wagner et al. 2021). Moreover, only a few analyses considered NSCLC by histology.

Therefore, this study aimed to investigate the association between widely studied variants of *PDCDI*, *CD274* and *HAVCR2* genes with NSCLC risk and outcomes as well as by LUSC and LUAD subtypes separately. We also examined the contribution of haplotypes to the development and the course of NSCLC and its subtypes.

Materials and methods

NSCLC patients and control group

383 patients with a confirmed diagnosis of NSCLC were enrolled by the Department of Pulmonology and Lung

Cancer, Wrocław Medical University and the Thoracic Surgery Centre, Lower Silesian Centre of Lung Disease, Wrocław, from 2005 to 2017. Table 1 presents baseline characteristics of NSCLC patients and the control group. Histological subtypes of 128 patients were unable to be determined as the diagnostic criteria of small biopsies and non-resection specimens were not established until 2011. The control group consisted of 433 unrelated cancer-free volunteers and blood donors. Written informed consent was obtained from all participants. The study was conducted following the Declaration of Helsinki and was approved by the Bioethics Committee of the Wrocław Medical University (KB-586/2020 September 28, 2020; Addendum; October 13, 2020).

Selection of single nucleotide polymorphisms (SNPs) and genotyping

We selected widely investigated genetic variants of *PDCD1* (rs36084323 C > T, rs7421861 A > G, rs11568821 C > T, rs2227981 A > G, rs10204525 C > T), *CD274* (rs822335 T > C, rs10815225 G > C, rs17718883 C > G, rs2297136 G > A, rs4742098 A > G, rs4143815 G > C), and *HAVCR2* ((rs10057302 C > A as a tagging SNP for rs10053538 C > A, LD $r^2 = 1.0$) and rs1036199 C > A) in different types of cancer. A detailed description of these variants can be found in our review article, where we summarised the association between inherited variants in these genes and their functional relevance (Wagner et al. 2021). Allelic discrimination using TaqMan SNP Genotyping Assays Applied Biosystems™ and TaqMan™ Genotyping Master Mix Applied Biosystems™ (Cat. No.: 4351376, Cat. No.: 4381656, respectively, Thermo Fisher Scientific, Waltham, MA, USA) were used for genotyping all the examined SNPs according to the manufacturer's protocol, except for rs2297136 and rs17718883, which were genotyped using polymerase chain reaction–restriction fragment length polymorphism (PCR–RFLP). The PCR products were digested with *BsuRI* (*HaeIII*) and *HinfI* restriction enzymes, respectively (Cat. No: ER0151 and ER0801, respectively, Thermo Fisher Scientific, Waltham, MA, USA). Supplementary Table S1 contains a detailed list of genotyping assays used in this study. Primer sequences, annealing temperatures, restriction enzymes and restriction fragments length are listed in Supplementary Table S2. TaqMan SNP Genotyping Assays were run on ViiA 7 Real-Time PCR system and analysed with Quant Studio Real-Time PCR Software 2016 v.1.3 (Applied Biosystems, Thermo Fisher Scientific, Waltham, MA, USA). PCRs were carried out on T100™ Thermal

Table 1 Baseline characteristics of NSCLC patients and control group

Variables	Patients		Control group	
	Number	%	Number	%
Overall	383	100	433	100
Sex				
Male	276	72.06	313	72.29
Female	107	27.94	119	27.48
Missing value	0	0	1	0.23
Median age (range) ^a				
Male	64 (45–87)		32 (18–75)	
Female	62 (36–87)		33 (19–63)	
All	64 (36–87)		32 (19–75)	
Smoking status				
Ever smoker	318	83.03	98	22.63
Non-smoker	51	13.32	55	12.70
Missing value	14	3.65	280	64.67
Histological type				
LUAD	112	29.24		
LUSC	116	30.29		
LCC	21	5.49		
NEC	3	0.78		
ASC	3	0.78		
NOS	128	33.42		
Tumour stage				
I	86	22.45		
II	49	12.80		
III	118	30.81		
IV	123	32.11		
Missing value	7	1.83		
Therapy				
Surgery				
Yes	181	47.26		
No	200	52.22		
Missing value	2	0.52		
Chemotherapy				
Yes	250	65.27		
No	121	31.60		
Missing value	12	3.13		
Survival status				
Alive ^b	57	14.88		
Deceased	266	69.45		
Missing value	60	15.67		

LUAD adenocarcinoma; LUSC squamous cell carcinoma; LCC Large cell carcinoma; ASC adenosquamous cell carcinoma; NEC neuroendocrine carcinoma; NOS not otherwise specified

^aPatients' age at diagnosis

^bStatus on 31.12.2021

Cycler (Bio-Rad, Hercules, CA, USA). Genotyping accuracy for all SNPs were verified by direct sequencing of a few samples representing homozygotes of two types and heterozygotes for each investigated SNP. These samples were used as reference samples in the subsequent genotyping experiments.

In silico analysis

We applied several publicly available tools to indicate possible links between investigated germline variants in studied genes and the impact on development and course of the disease. dbSNP was the reference source for providing the SNP ID number, alleles and frequencies. RegulomeDB 2.0.3 (Boyle et al. 2012) and 3DSNP 2.0 tools (Quan et al. 2022) were used to examine the putative regulatory potential of SNPs. HaploRegv4.1 was applied for the functional annotation of SNPs as well as for analysis of Linkage Disequilibrium (LD) (Ward and Kellis 2012). Expression quantitative trait loci (eQTL) analysis was conducted using the Genotype-Tissue Expression (GTEx) portal V8 release. The University of Alabama at Birmingham Cancer (UALCAN) data analysis portal was used for the examination of hsa-mir-570 expression levels in LUAD and LUSC (Chandrashekar et al. 2017, 2022).

Statistical analysis

The difference between genotype distribution in patients and the control group was analysed with a chi-squared test. The odds ratio and 95% confidence interval (95% CI) were applied as a measure of effect size. All analyses used the control group's most frequent homozygote and allele as the reference category. Association between examined genetic variants and age at diagnosis was tested by linear regression model, adjusted for smoking and gender. Chi-squared or Fisher's exact test was used to test the association between SNPs and clinical stage, lymph and distant metastases considering patients' smoking status. Multivariable-adjusted hazard ratios (HRs) and 95% CI for all-cause survival were estimated using Cox proportional hazards models with adjustment for smoking and gender, using age at diagnosis as the time metric. Follow-up started at the day of NSCLC diagnosis and ended at the day of death or the end of follow-up (31/12/2021). Kaplan–Meier (KM) survival curves were estimated for each SNP between subgroups of surgical intervention (patients who underwent surgery vs. patients who did not undergo surgery). Haplotypes and their frequency were estimated using the likelihood-based expectation–maximization algorithm EM algorithm for *PDCD1* and *CD274* genes; $P < 0.05$ was considered statistically significant.

Statistical analysis was conducted using R, version 4.1.3 (R Project for Statistical Computing) and STATA 14 (College Station, TX: StataCorp LLC).

Results

PDCD1, *CD274*, and *HAVCR2* genetic variants and susceptibility to NSCLC

Genotypes of selected SNPs were determined in 383 NSCLC and 433 unrelated cancer-free subjects. The detected allele frequencies for all SNPs were similar to the European population from the 1000 Genomes Project, as listed in dbSNP (Supplementary Table S3). There were no differences in genotype distribution between NSCLC patients and control subjects for any of our examined SNPs (Supplementary Tables S4–S6). As LUSC and LUAD are suspected of having different genetic backgrounds, we investigated the association between our selected SNPs and the risk of developing either subtype (Table 2, Supplementary Tables S7–S9 and S10–S12, respectively). We noticed a higher frequency of rs4143815 CC homozygotes among LUSC patients (19.0% vs. 9.2%). Rs4143815 CC genotype was associated with two times higher LUSC risk (CC vs. GG + GC, OR = 2.31; 95% CI = 1.32, 4.06; $P = 0.003$). Also, rs4742098 GG homozygotes were more frequent in LUSC patients (12.9% vs. 6.2%) and had two times higher LUSC development risk (GG vs. AA + GG, OR = 2.26; 95% CI = 1.17, 4.36; $P = 0.016$) (Table 2). In addition, we noticed an interesting, however, not statistically significant, association for rs11568821 C > T of *PDCD1*. Subjects carrying TT homozygotes might had more than three times higher risk of developing this NSCLC subtype compared to subjects with CC or CT genotype (TT vs. CC + CT, OR = 3.12; 95% CI = 0.88, 11.02; $P = 0.084$). We did not notice any significant results for any examined SNPs for LUAD.

PDCD1, *CD274*, and *HAVCR2* genetic variants and NSCLC prognosis

Next, we investigated the prognostic outcomes of selected SNPs. First, we checked if any of selected SNPs were associated with age at diagnosis in NSCLC as a whole group and by LUSC and LUAD separately. Rs4143815 of *CD274* appeared to be an independent predictor of the age at diagnosis of LUAD. LUAD patients carrying C allele (GC + CC) were diagnosed 4.81 years later than patients with GG genotype (parameter = 4.81; 95% CI = 1.47, 8.15; $P = 0.006$) (Table 3).

Table 2 Association of *PDCD1* and *CD274* variants and the risk of LUSC

Gene variant		LUSC patients <i>n</i> = 116		Controls <i>n</i> = 433		OR	95% CI	Patients vs. controls <i>P</i> value
<i>PDCD1</i> (-strand)		<i>n</i>	%	<i>n</i>	%			
rs11568821 C>T	CC	95	81.9	362	83.6	1	Ref	$X^2_{df=2} = 2.9886$
c.627+189 intron 4	CT	17	14.7	66	15.2	1.00	0.56–1.77	$P = 0.2244$
PD-1.3	TT	4	3.4	5	1.2	3.11	0.88–11.02	
HWE		$P_{HWE} = 0.0104$		$P_{HWE} = 0.3175$				
	CC+CT	112	96.6	428	98.8	1	Ref	$X^2_{df=1} = 2.9792$
	TT	4	3.4	5	1.2	3.12	0.88–11.02	$P = 0.0843$
<i>CD274</i>								
rs4143815 G>C	GG	50	43.1	206	47.6	1	Ref	$X^2_{df=2} = 8.6590$
c.*395 exon 7 (3'UTR)	GC	44	37.9	187	43.2	0.97	0.62–1.52	$P = 0.0132$
	CC	22	19.0	40	9.2	2.27	1.25–4.14	
HWE		$P_{HWE} = 0.0362$		$P_{HWE} = 0.7938$				
	GG+GC	94	81.0	393	90.8	1	Ref	$X^2_{df=1} = 8.6262$
	CC	22	19.0	40	9.2	2.31	1.32–4.06	$P = 0.0033$
rs4742098 A>G	AA	61	52.6	230	53.1	1	Ref	$X^2_{df=2} = 6.2495$
c.*2635A exon 7 (3'UTR)	AG	40	34.5	176	40.7	0.86	0.55–1.34	$P = 0.0439$
	GG	15	12.9	27	6.2	2.11	1.07–4.18	
HWE		$P_{HWE} = 0.0504$		$P_{HWE} = 0.3827$				
	AA+AG	101	87.1	406	93.8	1	Ref	$X^2_{df=1} = 5.7947$
	GG	15	12.9	27	6.2	2.26	1.17–4.36	$P = 0.0161$

Statistically significant ($P < 0.05$)

SNP alleles are reported in the Forward orientation in agreement with the dbSNP database. The most frequent homozygote in the control group was used as the reference genotype

OR odds ratio; CI confidence interval; ref. reference genotype; HWE test for Hardy–Weinberg equilibrium

We then examined the association between investigated SNPs and: (1) cancer clinical stage, (2) overall survival (OS), (3) OS and surgery status, (4) lymph nodes metastasis, and (5) distal metastasis. These analyses were performed for NSCLC, LUSC and LUAD, adjusting for smoking and gender where applicable. We found that CA genotype at *HAVCR2* rs10057302 might be an independent predictor of OS in LUSC. Patients with this genotype had about 87% lower chance of death at any time than patients with CC genotype (aHR = 0.13; 95% CI = 0.02, 0.93; $P = 0.043$) (Table 3). Among NSCLC patients who underwent surgery, rs11568821 T allele carriers (CT + TT) had almost double the risk of all-cause mortality (aHR = 2.05; 95% CI = 1.28, 3.29; $P = 0.003$) compared to patients carrying CC genotype (Fig. 1, Table 3).

Moreover, distal metastasis was present in a lower percentage of NSCLC patients who were smokers and *HAVCR2* rs10057302 carriers of CA genotype than patients with CC genotype irrespective of smoking status (CA smokers 16.67% vs. CC Non-smokers 36.36%, CC smokers 32.53%; $P = 0.009$) (Table 3). For *HAVCR2*

rs1036199, we observed a higher percentage of patients with distal metastasis among non-smoking C allele carriers (AA non-smoker 19.35% vs. CA smokers 58%) and smokers with CC genotype (AA smokers 29.03%, AC smokers 29.03% vs. CC smokers 53.33%) (Table 3).

Predictive values of *PDCD1* and *CD274* haplotypes in NSCLC

Supplementary Tables S13 and S14 show the distribution of the most frequent haplotypes in the Polish population for *PDCD1* and *CD274*. The specific combination of adjacent SNPs could be necessary for revealing a particular phenotype; therefore, haplotype analysis may allow the determination of such distinctive SNP composition (Ek 2007). Hence, we also examined the association between haplotypes and NSCLC risk and its subtypes, as well as clinicopathological data, such as age at diagnosis and OS. None of the estimated haplotypes was associated with NSCLC risk. However, the *CD274* haplotype rs822335*C-rs10815225*G-rs17718883*C-rs2297136*A-rs4143815*G-rs4742098*A

Table 3 Prognostic significance of *CD274*, *PDCD1*, and *HAVCR2* variants

SNV gene	LUAD	<i>n</i>	Parameter [year] ^a		95% CI	<i>P</i> ^a
rs4143815 G > C <i>CD274</i>	GG	47	Ref.			
	GC	49	5.54		2.04	9.04
	CC	12	1.80		-3.77	7.37
	GG	47	Ref.			
	GC+CC	61	4.81		1.47	8.15
rs10057302 C > A <i>HAVCR2</i>	LUSC	<i>n</i>	Deaths	HR	95% CI	<i>P</i> ^b
	CC	83	64	1	Ref.	
	CA	5	1	0.13	0.02	0.93
rs11568821 C > T <i>PDCD1</i>	NSCLC Operated patients	<i>n</i>	Deaths	HR	95% CI	<i>P</i> ^b
	CC	134	85	1	Ref.	
	CT	27	22	1.97	1.21	3.22
	TT	2	2	3.43	0.82	14.37
	CC	134	85	1	Ref.	3.29
	CT+TT	29	24	2.05	1.28	
rs10057302 C > A <i>HAVCR2</i>	NSCLC	<i>n</i>	M0	M1	Mx	<i>P</i> ^c
	CC _{NON-SMOKER}	44	22 (50.00%)	16 (36.36%)	6 (13.64%)	0.009
	CC _{SMOKER}	292	187 (64.04%)	95 (32.53%)	10 (3.42%)	
	CA _{SMOKER}	24	17 (70.83%)	4 (16.67%)	3 (12.50%)	
rs1036199 C > A <i>HAVCR2</i>	NSCLC	<i>n</i>	M0	M1	Mx	<i>P</i> ^c
	AC _{NON-SMOKER}	17	6 (35.29%)	10 (58.82%)	1 (5.88%)	0.002
	AA _{NON-SMOKER}	31	19 (61.29%)	6 (19.35%)	6 (19.35%)	
	CC _{SMOKER}	15	7 (46.67%)	8 (53.33%)	0 (0.00%)	
	AC _{SMOKER}	93	64 (68.82%)	27 (29.03%)	2 (2.15%)	
	AA _{SMOKER}	207	132 (63.77%)	64 (30.92%)	11 (5.31%)	

HR hazard ratio; CI confidence intervals; M0 cancer has not spread to other parts of the body; M1 cancer has spread to other parts of the body; Mx metastasis cannot be measured

P^a linear regressions adjusted for smoking and gender; statistically significant (*P* < 0.05)

P^b Cox proportional hazard models adjusted for smoking and gender; statistically significant (*P* < 0.05)

P^c chi-squared or Fisher's exact test; statistically significant (*P* < 0.05)

^aDifference in the age at diagnosis

(reference haplotype: rs822335*C-rs10815225*G-rs17718883*C-rs2297136*G-rs4143815*G-rs4742098*A) might decrease the age at diagnosis of NSCLC on average by 1 year and 5 months after adjustment for smoking (Parameter = -1.51; 95% CI = -2.73, -0.28; *P* = 0.016) (Supplementary Table S15). The haplotype associated with earlier age at diagnosis differs from the reference haplotype only by A allele of rs2297136. This SNP is located within the binding site for hsa-mir-296-5p, with A allele affecting binding between hsa-mir-296-5p and *CD274* mRNA. Accordingly, in the reporter assay, G allele was associated with lower luciferase activity in relation to A allele (Du et al. 2017). In gastric cancer, AA and AG genotypes of rs2297136 were significantly more frequent among patients with positive PD-L1 expression assayed by immunohistochemistry (IHC) (Wu et al. 2019). Of note, we observed a better prognosis in terms of survival for operated NSCLC patients having G allele,

but after adjustment for smoking and gender, this correlation lost statistical significance (Supplementary Figure 1). By contrast, the haplotype rs822335*T-rs10815225*G-rs17718883*C-rs2297136*A-rs4143815*G-rs4742098*A increased the age of diagnosis of NSCLC by approximately 2 years (Parameter = 1.98; 95% CI = 0.40, 3.56; *P* = 0.014) (Supplementary Table S15). This haplotype had T allele at rs822335 in addition to allele A at rs2297136. The impact of rs822335 on PD-L1 expression was investigated in tissue samples of NSCLC patients by Krawczyk et al. (2019) and Grenada et al. (2021). These authors detected a higher percentage of tumour cells presenting PD-L1 assessed by IHC in samples from patients with CC genotype compared to CT or TT genotypes (Krawczyk et al. 2019; Grenada et al. 2021). We carried out in silico analysis with RegulomeDB and 3DSNP 2.0 databases which showed that rs822335 was annotated with the enhancer state in the A549 cell line and

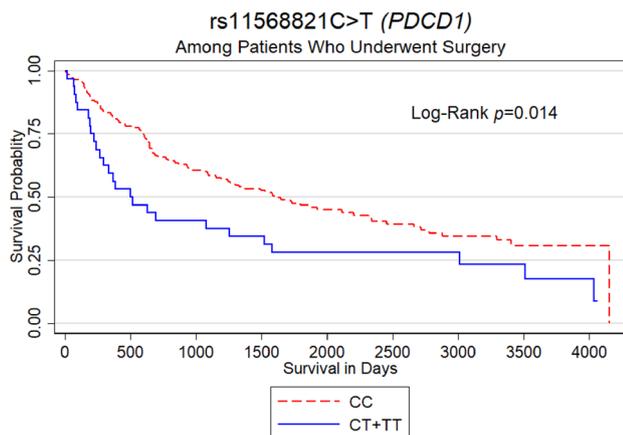


Fig. 1 The rs11568821 germline variant of the *PDCD1* gene predicts the overall survival of NSCLC patients who have undergone surgery. The carriers of allele T (CT+TT) had almost double the risk of all-cause mortality (aHR=2.052; 95% CI=1.281, 3.287; $P=0.003$). Cox proportional hazard models were adjusted for smoking and gender

Lung Fibroblast Primary Cells (NHLF), however, any potential binding site for transcription factor related to rs822335 was revealed.

Two haplotypes of *CD274* decreased the age at diagnosis of LUSC. The abovementioned haplotype rs822335*C-rs10815225*G-rs17718883*C-rs2297136*A-rs4143815*G-rs4742098*A lowered the age at diagnosis by about 3 years (Parameter = -3.26; 95% CI = -5.45, -1.07; $P=0.004$) and haplotype rs822335*T-rs10815225*G-rs17718883*C-rs2297136*A-rs4143815*C-rs4742098*G by nearly 2 years (Parameter = -1.83; 95% CI = -3.44, -0.22; $P=0.026$) (Supplementary Table S16). It is worth noting that the last haplotype had the opposite effect in LUAD, where it was associated with higher age at diagnosis by about 6 years (Parameter = 6.13; 95% CI = 3.42, 8.84; $P<0.001$) (Supplementary Table S17).

When analysing the association between haplotypes and OS, we noticed that rs36084323*C-rs7421861*G-rs11568821*T-rs2227981*G-rs10204525*C haplotype of *PDCD1* after adjustment for smoking increased the probability of death among NSCLC patients who underwent surgery to about 70% (HR = 1.70; 95% CI = 1.13, 2.57; $P=0.011$) in relation to the reference haplotype: rs36084323*C-rs7421861*A-rs11568821*C-rs2227981*A-rs10204525*C (Supplementary Table S18). The haplotype associated with a worse prognosis contained T allele of rs11568821, which in our study was associated with worse survival of patients undergoing surgery, and the G allele of rs7421861, associated with a higher overall cancer risk (Hashemi et al. 2019; Zhang et al. 2021). These haplotypes also differ regarding rs2227981 (also known as PD-1.5 or +7785). The association between A allele and a higher percentage of

PD-1⁺CD4⁺ T cells was detected in Japanese patients with autoimmune type I diabetes (Fujisawa et al. 2015). The data relating to association of rs2227981 with NSCLC are conflicting: Yin and colleagues (2014) observed that A allele was associated with lower NSCLC risk, but no such association was detected by Ma and colleagues (2015). Similarly, an Iranian study did not find associations between rs2227981 and NSCLC (Piredelkhosh et al. 2018). Further studies are needed to confirm the observed correlations between particular haplotypes and clinical outcomes. It is also essential to carry out functional studies to reveal the biological implication of analysed genetic variants as a single locus and variants interacting in a particular combination.

Discussion

Our study aimed to investigate whether SNPs present in genes for PD-1, PD-L1 and TIM-3 molecules could predispose to NSCLC development, progression, and affect patient outcomes. Being aware of possible etiological differences between the LUAD and LUSC, we also examined if selected variants were associated with either subtype.

As most published reports investigated the association between genetic variants of immune checkpoints without distinguishing between LUAD and LUSC, we have also used this approach; however, we did not observe any significant association for NSCLC as a whole group. In our study, CC genotype of rs4143815 appeared to be associated exclusively with a higher LUSC risk. This variant is in the 3'UTR of *CD274* within the seed region of the hsa-mir-570 binding site, with rs4143815 allele C predicted to affect the interaction between hsa-mir-570 and *CD274* mRNA (Wang et al. 2013). Functional relevance of this allele was confirmed in luciferase assays, where C allele was associated with higher luciferase activity in gastric cancer (Wang et al. 2013) and NSCLC (Lee et al. 2017). Moreover, Yeo and colleagues found an association between elevated expression of PD-L1 protein evaluated by IHC and CC genotype of rs4143815 in NSCLC samples (Yeo et al. 2017). It is worth mentioning that increased levels of hsa-mir-570 were detected in NSCLC tumour tissue in relation to adjacent non-malignant tissue (Tong et al. 2015). Since we observed the different impact of rs4143815 on LUSC and LUAD, we checked if hsa-mir-570 is differently expressed in the microenvironment of these two subtypes of NSCLC. As presented in Supplementary Fig. S2, hsa-mir-570 expression was significantly higher in LUSC (Chandrashekar et al. 2017, 2022). Based on available data, it can be assumed with a high probability that NSCLC patients with rs4143815 CC genotype will have elevated PD-L1 expression in tumour tissue, indicating better responses to mAbs blocking the interaction between PD-1 and PD-L1 (Nomizo et al. 2017). This

relation has been already described in literature: Nomizo and colleagues investigated the relationship between rs4143815 genotypes and response to nivolumab (anti-PD-1 mAbs) and found rs4143815 allele C carriers (GC and CC genotypes) treated with nivolumab showed significantly better clinical response than patients with GG genotype, with the highest objective response rate observed for rs4143815 CC genotype as well as the longest progression-free survival (PFS) (Nomizo et al. 2017). In line with this observation, a significantly shorter PFS associated with rs4143815 GG in NSCLC patients treated with nivolumab was observed in a study by Del Re and colleagues (Re et al. 2021). So far, the largest study investigating the response to ICIs in relation to rs4143815 was performed by Minari and colleagues on Italian patients with metastatic NSCLC (Minari et al. 2022). This group did not find a significant correlation between rs4143815 genotypes and clinical outcomes such as PFS or OS. Nevertheless, they reported a significant association between allele C of rs4143815 (GC and CC genotypes) and long-lasting responses to ICIs treatment (Minari et al. 2022).

To the best of our knowledge, no reports have shown an association between rs4143815 variant and NSCLC risk. In work by Du and colleagues, genotypes of this SNP were not associated with risk or analysed clinical covariates (Du et al. 2017). In terms of clinical outcomes, Lee and colleagues reported the association between rs4143815 genotype GG and significantly worse OS in NSCLC under the recessive model. Analysis by histology revealed that this association appeared to be significant only in LUAD (Lee et al. 2017). Yeo et al. obtained a similar result for rs4143815 GG genotype. These authors observed that this genotype was related to a shorter disease-free survival and OS among LUAD patients; however, this association was not significant (Yeo et al. 2017). In our study, rs4143815 genotype GG was not associated with OS either for NSCLC or after stratification by histology; however, this genotype correlated significantly with the earlier onset of LUAD in relation to C allele carriers at this site.

In our study, rs4742098, another 3'UTR variant of *CD274*, was also associated with LUSC, with GG genotype significantly increasing the risk of developing this subtype. This variant is located within the seed region of hsa-mir-138, and its functional relevance was confirmed by luciferase assay where allele G, predicted to impact the binding of hsa-mir-138 to 3'UTR of *CD274* mRNA, was associated with higher luciferase activity in contrast to decreased activity demonstrated for allele A (Du et al. 2017). Hence, it may be assumed that the substitution of A to G in rs4742098 will be responsible for the higher expression of PD-L1. Accordingly, Du and colleagues observed that rs4742098 AG genotype was associated with a higher NSCLC risk in relation to AA genotype; surprisingly GG genotype did not show this effect (Du et al. 2017).

It is established that surgery is the best cure for early-stage NSCLC; however, a significant proportion of operated patients die from cancer relapse. Considerable variability in survival exists even among NSCLC patients within the same stage, indicating that host genetic factors may determine prognosis (Shin et al. 2018).

Among NSCLC patients who underwent surgery in our cohort, rs11568821 T allele carriers had a significantly higher probability of death compared to patients with CC genotype. Two more studies investigated rs11568821 in NSCLC, one in a Chinese (Ma et al. 2015), another in an Iranian population (Piredelkhosh et al. 2018). The association between rs11568821 genotypes and the risk or outcomes of NSCLC was not found in either study, similar to our results. These reports did not stratify analyses by histology. In a retrospective study, melanoma patients carrying CC homozygotes in rs11568821 had better responses to nivolumab or pembrolizumab than patients carrying T allele (CT and TT genotypes), showing C allele significantly associated with longer median PFS (Parakh et al. 2021). It is worth mentioning that rs11568821, according to RegulomeDB, has been annotated with chromatin states related to the strong transcription in different subpopulations of T cells and lung tissue. It can also be considered a *cis*-eQTL variant for the *PDCDI* gene in lung tissue according to GTEx, with T allele associated with higher expression of *PDCDI* (Supplementary Figure 3). Therefore, it can be speculated that the higher mortality rate among operated rs11568821 T allele carriers in our study may be due to insufficient response from the immune system against the remaining tumour cells resulting from higher PD-1 expression, and thus a weaker response from T lymphocytes. Due to the low frequency of T allele in our and other populations, a lack of sufficient data about the impact of rs11568821 on PD-1 expression/activity and only two case-control studies related to this variant in NSCLC, our results must be considered with due caution. Further investigation needs to be conducted to evaluate the biological relevance of *PDCDI* rs11568821, and its association with NSCLC histology, risk and/or outcomes must be validated in larger groups of control subjects and patients.

We investigated two variants of the *HAVCR2* gene, rs10057302 and rs1036199. We found that patients carrying CA genotype of rs10057302 (our tag SNP for rs10053538 C > A) had better OS than patients with CC genotype. We observed this correlation only for patients with LUSC. In addition, we noticed that NSCLC patients who smoked and were rs10057302 CA genotype carriers were less likely to develop distal metastasis compared to those with CC genotype, which may explain the better prognosis observed in our patients with LUSC. To our knowledge, apart from our study, only Bai and colleagues investigated rs10053538 in NSCLC (Bai et al. 2013) but did not find any associations. However, the authors did not stratify patients by histology.

Bai et al. also examined rs1036199 and found that CA genotype was associated with a higher NSCLC risk and shorter survival compared to AA genotype (Bai et al. 2013). We did not find an association between rs1036199 genotypes and NSCLC or its subtypes. Nonetheless, we noticed a higher percentage of patients with distal metastasis among non-smokers with CA genotype and smokers with CC genotype. The association of CA genotype with distal metastasis was also observed in breast cancer and renal cell carcinoma (Cheng et al. 2017; Cai et al. 2012).

To the best of our knowledge, our study is the first one comprehensively analysing the impact of frequently investigated common inherited variants of genes for PD-1, PD-L1 and TIM-3 in European NSCLC patients.

The main limitation of our study is a relatively low number of patients in the LUSC and LUAD groups, which might impact the significance level obtained for examined SNPs, especially those with low minor allele frequencies (MAFs). Moreover, since our control group largely consisted of blood donors, there was a significant difference between the median age of patients and control subjects. Further research is therefore required. Nonetheless, our study provides strong evidence that germline variants in genes for immune checkpoint molecules may impact NSCLC risk and outcomes and identifying these variants may be helpful in clinical practice.

In summary, our results are in line with the postulate that LUAD and LUSC should be analysed as two distinct biological subtypes (Relli et al. 2019; Faruki et al. 2017). This distinction may be essential when planning the management of patients with NSCLC. Taking this difference into consideration, further research may fine-tune findings which can improve risk and prognosis prediction and even guide new treatment modalities.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00432-023-04602-8>.

Acknowledgements This work was supported by the National Science Centre, Poland, grant number 2019/33/B/NZ5/03029 to MJ. The Genotype-Tissue Expression (GTEx) Project was supported by the Common Found of the Office of the Director of the National Institutes of Health and by the National Cancer Institute (NCI), the National Human Genome Research Institute (NHGRI), the National Heart, Lung, and Blood Institute (NHLBI), the National Institute on Drug Abuse (NIDA), the National Institute of Mental Health (NIMH), and the National Institute of Neurological Disorders and Stroke (NINDS). The data used for the analyses described in this manuscript were obtained from the GTEx Portal on 10/09/22.

Author contributions NM—investigation, data curation, formal analysis, writing-original draft; MW—methodology, investigation, formal analysis, data curation, validation, supervision, writing review and editing; KP—resources, data curation, writing review and editing; IP—resources, data curation, writing review and editing; BM-B—resources, data curation, writing review and editing; AK—resources, data curation, writing review and editing; AW—resources, data curation, writing review and editing; MK—resources, writing review and editing; JK—investigation, writing review and editing, PK—formal analysis,

writing review and editing; DF—investigation, writing review and editing; AA—investigation, writing review and editing; LK—resources, writing review and editing; PK—resources, writing review and editing; MJ—conceptualization, resources, funding acquisition, data curation, methodology, supervision, investigation, validation, formal analysis, writing-original draft, project administration. All authors reviewed the manuscript.

Funding This work was supported by the National Science Centre, Poland, Grant number 2019/33/B/NZ5/03029 to MJ.

Data availability The datasets generated and analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare no potential conflicts of interest.

Ethics approval This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Bioethics Committee of the Wroclaw Medical University (KB-586/2020 September 28, 2020; Addendum; October 13, 2020).

Consent to participate and to publish Written informed consent was obtained from all individual participants included in the study, and the authors affirm that human research participants provided informed consent for publication of their anonymised, aggregated data.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Acharya N, Sabatos-Peyton C, Anderson AC (2020) Tim-3 finds its place in the cancer immunotherapy landscape. *J Immunother Cancer* 8:e000911. <https://doi.org/10.1136/jitc-2020-000911>
- Andrews LP, Yano H, Vignali DAA (2019) Inhibitory receptors and ligands beyond PD-1, PD-L1 and CTLA-4: breakthroughs or backups. *Nat Immunol* 20:1425–1434. <https://doi.org/10.1038/s41590-019-0512-0>
- Bai J, Li X, Tong D et al (2013) T-cell immunoglobulin- and mucin-domain-containing molecule 3 gene polymorphisms and prognosis of non-small-cell lung cancer. *Tumor Biol* 34:805–809. <https://doi.org/10.1007/s13277-012-0610-1>
- Boyle AP, Hong EL, Hariharan M et al (2012) Annotation of functional variation in personal genomes using regulome DB. *Genome Res* 22:1790–1797. <https://doi.org/10.1101/gr.137323.112>
- Cai C, Wang L, Wu Z et al (2012) T-cell immunoglobulin- and mucin-domain-containing molecule 3 gene polymorphisms and renal

- cell carcinoma. *DNA Cell Biol* 31:1285–1289. <https://doi.org/10.1089/dna.2012.1625>
- Chandrashekar DS, Bashel B, Balasubramanya SAH et al (2017) UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. *Neoplasia* 19:649–658. <https://doi.org/10.1016/j.neo.2017.05.002>
- Chandrashekar DS, Karthikeyan SK, Korla PK et al (2022) UALCAN: an update to the integrated cancer data analysis platform. *Neoplasia* 25:18–27. <https://doi.org/10.1016/j.neo.2022.01.001>
- Cheng S, Ju Y, Han F et al (2017) T cell Immunoglobulin- and mucin-domain-containing molecule 3 gene polymorphisms and susceptibility to invasive breast cancer. *Ann Clin Lab Sci* 47:668–675
- Curigliano G, Gelderblom H, Mach N et al (2021) Phase I/IIb clinical trial of sabatolimab, an anti-TIM-3 antibody, alone and in combination with spartalizumab, an anti-PD-1 antibody, in advanced solid tumors. *Clin Cancer Res off J Am Assoc Cancer Res* 27:3620–3629. <https://doi.org/10.1158/1078-0432.CCR-20-4746>
- Datar I, Sanmamed MF, Wang J et al (2019) Expression analysis and significance of PD-1, LAG-3, and TIM-3 in human non-small cell lung cancer using spatially resolved and multiparametric single-cell analysis. *Clin Cancer Res off J Am Assoc Cancer Res* 25:4663–4673. <https://doi.org/10.1158/1078-0432.CCR-18-4142>
- Del Re M, Cucchiara F, Rofi E et al (2021) A multiparametric approach to improve the prediction of response to immunotherapy in patients with metastatic NSCLC. *Cancer Immunol Immunother* 70:1667–1678. <https://doi.org/10.1007/s00262-020-02810-6>
- Du W, Zhu J, Chen Y et al (2017) Variant SNPs at the microRNA complementary site in the B7–H1 3′-untranslated region increase the risk of non-small cell lung cancer. *Mol Med Rep* 16:2682–2690
- Ek S (2007) Haplotype thinking in lung disease. *Proc Am Thorac Soc* 4:4–8. <https://doi.org/10.1513/pats.200607-145JG>
- Faruki H, Mayhew GM, Serody JS et al (2017) Lung adenocarcinoma and squamous cell carcinoma gene expression subtypes demonstrate significant differences in tumor immune landscape. *J Thorac Oncol off Publ Int Assoc Study Lung Cancer* 12:943–953. <https://doi.org/10.1016/j.jtho.2017.03.010>
- Fujisawa R, Haseda F, Tsutsumi C et al (2015) Low programmed cell death-1 (PD-1) expression in peripheral CD4(+) T cells in Japanese patients with autoimmune type 1 diabetes. *Clin Exp Immunol* 180:452–457. <https://doi.org/10.1111/cei.12603>
- Grenda A, Krawczyk P, Kucharczyk T et al (2021) Impact of copy number variant and single nucleotide polymorphism of the programmed death-ligand 1 gene, programmed death-ligand 1 protein expression and therapy regimens on overall survival in a large group of Caucasian patients with non-small cell lung carcinoma. *Oncol Lett* 21:449. <https://doi.org/10.3892/ol.2021.12710>
- Hashemi M, Karami S, Sarabandi S et al (2019) Association between PD-1 and PD-L1 polymorphisms and the risk of cancer: a meta-analysis of case–control studies. *Cancers* 11:1150
- Krawczyk P, Grenda A, Wojas-Krawczyk K et al (2019) PD-L1 gene copy number and promoter polymorphisms regulate PD-L1 expression in tumor cells of non-small cell lung cancer patients. *Cancer Genet* 237:10–18. <https://doi.org/10.1016/j.cancergen.2019.06.001>
- Lee SY, Jung DK, Choi JE et al (2017) Functional polymorphisms in PD-L1 gene are associated with the prognosis of patients with early stage non-small cell lung cancer. *Gene* 599:28–35. <https://doi.org/10.1016/j.gene.2016.11.007>
- Lucibello G, Mograbi B, Milano G et al (2021) PD-L1 regulation revisited: impact on immunotherapeutic strategies. *Trends Mol Med* 27:868–881. <https://doi.org/10.1016/j.molmed.2021.06.005>
- Ma Y, Liu X, Zhu J et al (2015) Polymorphisms of co-inhibitory molecules (CTLA-4/PD-1/PD-L1) and the risk of non-small cell lung cancer in a Chinese population. *Int J Clin Exp Med* 8:16585–16591
- Minari R, Bonatti F, Mazzaschi G et al (2022) PD-L1 SNPs as biomarkers to define benefit in patients with advanced NSCLC treated with immune checkpoint inhibitors. *Tumori* 108:47–55. <https://doi.org/10.1177/03008916211014954>
- Nomizo T, Ozasa H, Tsuji T et al (2017) Clinical impact of single nucleotide polymorphism in PD-L1 on response to nivolumab for advanced non-small-cell lung cancer patients. *Sci Rep* 7:45124
- Parakh S, Musafar A, Paessler S et al (2021) PDCD1 polymorphisms may predict response to anti-PD-1 blockade in patients with metastatic melanoma. *Front Immunol*. <https://doi.org/10.3389/fimmu.2021.672521>
- Piredelkhosh Z, Kazemi T, Haghshenas MR et al (2018) Investigation of programmed cell death-1 (PD-1) gene variations at positions PD1.3 and PD1.5 in Iranian patients with non-small cell lung cancer. *Middle East J Cancer* 9:13–17. <https://doi.org/10.30476/mejc.2018.42095>
- Quan C, Ping J, Lu H et al (2022) 3DSNP 2.0: update and expansion of the noncoding genomic variant annotation database. *Nucleic Acids Res* 50:D950–D955. <https://doi.org/10.1093/nar/gkab1008>
- Relli V, Trerotola M, Guerra E, Alberti S (2019) Abandoning the notion of non-small cell lung cancer. *Trends Mol Med* 25:585–594. <https://doi.org/10.1016/j.molmed.2019.04.012>
- Shin KM, Hong MJ, Lee SY et al (2018) Regulatory variants in cancer-related pathway genes predict survival of patients with surgically resected non-small cell lung cancer. *Gene* 646:56–63. <https://doi.org/10.1016/j.gene.2017.12.055>
- Tong XD, Liu TQ, Wang GB et al (2015) MicroRNA-570 promotes lung carcinoma proliferation through targeting tumor suppressor KLF9. *Int J Clin Exp Pathol* 8(3):2829–2834
- Wagner M, Jasek M, Karabon L (2021) Immune checkpoint molecules—herited variations as markers for cancer risk. *Front Immunol* 11:606721
- Wang W, Li F, Mao Y et al (2013) A miR-570 binding site polymorphism in the B7–H1 gene is associated with the risk of gastric adenocarcinoma. *Hum Genet* 132:641–648. <https://doi.org/10.1007/s00439-013-1275-6>
- Ward LD, Kellis M (2012) HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* 40:D930–934. <https://doi.org/10.1093/nar/gkr917>
- Wu Y, Zhao T, Jia Z et al (2019) Polymorphism of the programmed death-ligand 1 gene is associated with its protein expression and prognosis in gastric cancer. *J Gastroenterol Hepatol* 34:1201–1207. <https://doi.org/10.1111/jgh.14520>
- Yeo M-K, Choi S-Y, Seong I-O et al (2017) Association of PD-L1 expression and PD-L1 gene polymorphism with poor prognosis in lung adenocarcinoma and squamous cell carcinoma. *Hum Pathol* 68:103–111
- Yin L, Guo H, Zhao L, Wang J (2014) The programmed death-1 gene polymorphism (PD-1.5 C/T) is associated with non-small cell lung cancer risk in a Chinese Han population. *Int J Clin Exp Med* 7:5832–5836
- Zhang W, Song Y, Zhang X (2021) Relationship of programmed death-1 (PD-1) and programmed death ligand-1 (PD-L1) polymorphisms with overall cancer susceptibility: an updated meta-analysis of 28 studies with 60,612 subjects. *Med Sci Monit Int Med J Exp Clin Res* 27:e9321461–e93214617. <https://doi.org/10.12659/MSM.932146>

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