

Regulatory polymorphisms of *MSH6*, *MSH2*, *FBXO11*, and *PPP1R21* genes affect survival of patients with immunotherapy-treated lung cancer

Martina Esposito,¹ Sara Noci,² Francesca Minnai,^{1,3} Tania Camboni,¹ Eleonora Mangano,¹ Manuela Gariboldi,² Elisa Frullanti,^{4,5} Claudia Bareggi,⁶ Elena Collovà,⁷ Serena Girelli,⁸ Sheila Piva,⁸ Gabriella Farina,⁸ Arianna Pagliaro,⁹ Luca Toschi,⁹ Luca Sala,¹⁰ Diego Luigi Cortinovis ,^{10,11} Francesca Colombo ¹

To cite: Esposito M, Noci S, Minnai F, *et al.* Regulatory polymorphisms of *MSH6*, *MSH2*, *FBXO11*, and *PPP1R21* genes affect survival of patients with immunotherapy-treated lung cancer. *Journal for ImmunoTherapy of Cancer* 2025;**13**:e011526. doi:10.1136/jitc-2025-011526

► Additional supplemental material is published online only. To view, please visit the journal online (<https://doi.org/10.1136/jitc-2025-011526>).

Accepted 04 August 2025



© Author(s) (or their employer(s)) 2025. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ Group.

For numbered affiliations see end of article.

Correspondence to

Dr Francesca Colombo;
francesca.colombo@cnr.it

ABSTRACT

Background Immune checkpoint inhibitors (ICI) improved survival of patients with non-small cell lung cancer (NSCLC), yet many patients do not respond to treatment. The identification of markers for ICI response remains an unmet clinical need. This study hypothesizes that host genetics influences the response to ICI, contributing to the variability in efficacy among individuals.

Methods We conducted a genome-wide association study (GWAS) in patients with NSCLC on ICI monotherapy with nivolumab, pembrolizumab, or atezolizumab, to identify germline variants associated with objective response rate (ORR), progression-free survival (PFS), and overall survival (OS) at 24 months after the start of ICI therapy. Genomic DNA was genotyped using Axiom Precision Medicine Research Arrays. Raw data were processed with Axiom Analysis Suite, and quality checked with PLINK software. Imputation to the whole genome was done on the Michigan Imputation Server. Association analyses were performed for ORR (logistic regression with PLINK2 software) and survival (Cox proportional hazards model, with GenABEL package in R environment), with appropriate covariates. Variants were annotated for functional significance using SNPnexus and FUMA. Post-GWAS analyses, including colocalization, were performed to explore the function of the identified variants. Their possible role as expression quantitative trait loci was investigated in different databases (GTEx, eQTLGen, TCGA).

Results No genome-wide significant associations were found for ORR or PFS, while a locus on chromosome 2 (lead variant: rs111648355) showed near genome-wide significance (p value = 6.3×10^{-8}) for OS. Patients with minor alleles of these variants exhibited significantly worse OS (HR = 5.1, 95% CI: 2.9 to 9.2). Functional annotation linked these variants to regulatory effects on genes including *MSH2*, *MSH6*, *PPP1R21*, *FBXO11*, and *STON1*. These genes play a role in mismatch repair, endosomal trafficking, or major histocompatibility complex class II regulation, and might influence the response to immunotherapy.

Conclusions This study identifies an association between a genomic locus on chromosome 2 and OS in patients with NSCLC treated with ICI. Although these results need validation in larger cohorts and functional studies to

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Immune checkpoint inhibitors (ICI) have improved survival rates in non-small cell lung cancer (NSCLC), but only 27–46% of patients respond effectively. Existing biomarkers like tumor programmed death-ligand 1 expression, tumor microsatellite instability, and tumor mutation burden have limitations in predicting ICI response, particularly in clinical settings. The role of host genetics in influencing ICI response remains underexplored.

WHAT THIS STUDY ADDS

⇒ This study identifies a genomic locus on chromosome 2 where variants are associated with overall survival in patients with NSCLC treated with ICI. In silico functional analyses link these variants to regulatory effects on genes such as *MSH6*, *MSH2*, *PPP1R21*, and others, highlighting their potential roles in mismatch repair and immune modulation.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ These findings suggest germline genetic variants may serve as predictive biomarkers for ICI response, paving the way for personalized cancer immunotherapy. Further validation in larger cohorts could refine patient stratification and treatment optimization strategies.

elucidate the underlying mechanisms, they highlight the potential of germline variants as predictive biomarkers of response to ICI.

BACKGROUND

In recent years, immune checkpoint inhibitors (ICI), therapeutic monoclonal antibodies targeting immune checkpoint proteins such as cytotoxic T-lymphocyte associated protein 4, programmed cell death 1 (PD-1), and programmed death-ligand 1 (PD-L1), have revolutionized cancer treatment,

including non-small cell lung cancer (NSCLC). Namely, these drugs are nivolumab, pembrolizumab, ipilimumab, atezolizumab, and cemiplimab. Several clinical trials have demonstrated the beneficial effects of these drugs, in terms of overall survival (OS), for patients with NSCLC, as compared with traditional chemotherapy^{1–4}. In the first-line setting, in particular, several studies have demonstrated how ICI in monotherapy or in chemo-free combinations can determine a significant impact in terms of increasing median OS, leading to a long-term 5-year survival rate of over 30% in patients with PD-L1 tumor proportion score $\geq 50\%$.^{1–5}

Despite the benefits of ICI, only approximately 27%–46% of patients with NSCLC effectively respond to these therapies.^{6–8} Consequently, increasing emphasis is being placed on identifying biomarkers that can predict which patients are most likely to benefit from immunotherapy, allowing for better patient stratification. The US Food and Drug Administration has approved three biomarkers to predict responses to ICI: intratumoral PD-L1 expression,⁹ tumor microsatellite instability,¹⁰ and tumor mutation burden.¹¹ However, these markers are not yet fully standardized in clinical practice for patient selection and, in Italy, only tumor PD-L1 expression is considered in clinical practice.

Emerging research on host genetics suggests a potential genetic predisposition that influences individual response to immunotherapy, both in terms of efficacy and toxicity. This may be linked to variations in genes encoding proteins that interact with the immune system and tumors, such as checkpoint inhibitors,^{12–15} human leukocyte antigen class I and II,^{16–17} or those related to immune cell recruitment within tumors, as chemokines and their receptors.^{18–19} Furthermore, genetic variants associated with autoimmune diseases could have an impact on the individual responses to ICI, at least in terms of toxicity.^{20–21} Recently, H Carter's group investigated the role of 525 germline variants, affecting gene expression in the tumor immune microenvironment, on the response to immunotherapy, mainly in patients with melanoma.^{22–23} However, a comprehensive study investigating the association of all the germline variants with the response to ICI in patients with lung cancer, specifically, has yet to be performed.

Here, we carried out a pilot genome-wide association (GWAS) study in 152 Italian patients with NSCLC, treated with ICI in monotherapy, looking for germline variants associated with the objective response rate (ORR), progression-free survival (PFS), and OS at 24 months after the start of immunotherapy.

MATERIALS AND METHODS

Patient enrollment and research ethics

This multicenter study was done in a real-world setting, recruiting patients with NSCLC, treated with ICI in monotherapy, in five hospitals in Milan metropolitan area: Fatebenefratelli Hospital, Milan; Fondazione IRCCS San

Gerardo dei Tintori, Monza (MB); IRCCS Istituto Clinico Humanitas, Rozzano (MI); Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan; Legnano Hospital, Legnano (MI). Patients were retrospectively and prospectively enrolled from 2019, including patients undergoing immunotherapy in the years from 2017 to 2024. Patients were considered eligible for enrollment into the study if they met all the inclusion criteria in online supplemental table 1. Patients provided written informed consent to the collection and use of their biological samples and data for research purposes, according to the European General Data Protection Regulation. The study was conducted in accordance with the Declaration of Helsinki and approved by the ethics committees of the recruiting centers (see Ethical approval details in Declarations section).

Clinical data and biological samples

The following personal and clinical information was collected: age at immunotherapy start, sex, smoking status, ECOG (Eastern Cooperative Oncology Group) performance status, clinical/pathological stage, previous treatments (chemotherapy and/or radiotherapy), NSCLC histotype, result of immunohistochemical testing for PD-L1 in tumor specimens, tumor mutational status of lung cancer driver genes (ie, *KRAS*, *EGFR*, *ALK*, *ROS1*, *RET*), administered ICI, line of treatment (first or further), disease response to ICI treatment evaluated according to the Response Evaluation Criteria in Solid Tumors (RECIST V.1.1²⁴), and live/dead status after 24 months of ICI treatment, or at least until disease progression or end of treatment due to severe immune-related toxicity.

For each patient, a peripheral blood sample was collected at enrollment. Genomic DNA was extracted using the DNeasy Blood and Tissue (Qiagen) or Maxwell RSC Blood DNA (Promega) kits. DNA was fluorimetrically quantified with the Quant-iT PicoGreen dsDNA assay kit (Thermo Fisher Scientific) on a VarioScan LUX multiplate reader (Thermo Fisher Scientific).

Genome-wide genotyping

DNA samples from 166 patients were genome-wide genotyped using Axiom Precision Medicine Research Arrays (PMRA; Thermo Fisher Scientific) on GeneTitan instrument (Thermo Fisher Scientific). Axiom Analysis Suite software (AxAS, Thermo Fisher Scientific) was used to call genotypes of these samples using the “best practice” workflow. Genotype data were subject to quality control (QC) using PLINK V.1.9 software.²⁵ In detail, in the per-sample QC steps, we filtered out samples with a sample call rate $< 98\%$, sex discrepancies, and excessive heterozygosity rate. In detail, the mean heterozygosity rate for an individual was calculated using the formula: $(N(NM) - O(HM)) / N(NM)$, where $N(NM)$ represents the proportion of non-missing genotypes for the individual, and $O(HM)$ denotes the observed proportion of homozygous genotypes for the individual. We removed samples for which heterozygosity rate was found to

differ by three SDs from the population average. We also excluded related patients and duplicates, identified using an identity-by-descent test (as described in²⁶). Per-marker QC filters were genotyping call rate <95%, a minor allele frequency (MAF)<1%, or a Hardy-Weinberg equilibrium (HWE) test $p < 1.0 \times 10^{-10}$. Principal component analysis (PCA) was performed using PLINK2, to identify non-European patients by plotting the principal components of our samples together with those of 2,504 individuals from five different populations (Africans, Americans, East Asians, Europeans, and South-East Asians) selected from the 1,000 Genomes Project.²⁷

The clean dataset was used for imputation to whole genome sequence (for autosomal variants). We used the Minimac4 algorithm on the Michigan Imputation Server, setting GRCh37/hg19 as array build. The HRC r1.1 2016 (GRCh37/hg19) was used as reference panel. Data were phased with Eagle V.2.4.^{28–32} The imputed genotypes were then filtered to exclude variants with an MAF<5% and a low-quality imputation (R^2 info score ≤ 0.7).³³

Statistical analyses

Patient distributions among the recruiting centers and between our patient series and the validation one (from phs002822.v1.p1 dataset³⁴ in dbGap; see below), according to the different personal and clinical variables were compared using the Fisher's exact test (categorical variables) and Kruskal-Wallis test (quantitative variables). Logistic regressions between ORR (defined as the percentage of patients who completely or partially responded to immunotherapy per RECIST V.1.1 criteria) and personal and clinical variables were done using the generalized linear model, glm function of the stats package in R environment. Univariable Cox proportional hazard models³⁵ were tested using the coxph function of the survival package³⁶ in R. We analyzed OS and PFS, at 24 months of follow-up after immunotherapy start, testing the effect of the following variables as possible prognostic factors: age, sex, smoking habit, ECOG performance status, previous chemotherapy and/or radiotherapy, NSCLC histotype, tumor stage, PD-L1 tumor expression, administered ICI, and line of treatment (patients with missing data were excluded). We also evaluated differences among the recruiting centers. The significant variables were then tested in multivariable Cox proportional hazard models (with stepwise model selection, based on Akaike information criterion) to identify independent prognostic factors, for both PFS and OS. The survfit function of the same survival package was used to plot Kaplan-Meier curves and to perform log-rank test. For these analyses, the statistical significance threshold was set at p value<0.05.

We investigated the association between polymorphisms (tested in an additive model) and ORR, PFS and OS. Association with ORR was tested in a genome-wide logistic regression, using PLINK software (with no covariates, since they were not statistically significant). Two genome-wide survival analyses for PFS and OS were

carried out using the GenABEL package in R V.4.1.1³⁷ to test multivariable Cox proportional hazard models. For PFS, age, sex, administered ICI, and the recruiting center were used as covariates; the same variables as above, except for the center, and with the addition of the ECOG performance status, were included in the Cox model for OS analysis. The selected covariates were those passing Akaike stepwise model selection. A p value< 5.0×10^{-8} was conventionally set as the genome-wide statistical significance threshold. A suggestive threshold was considered at $p < 1.0 \times 10^{-5}$. Manhattan and QQ plots were obtained using the library qqman in R.

In silico functional analyses

Identified variants were annotated using SNPnexus tool³⁸ (accessed on September 16, 2024). We functionally annotated and prioritized the results obtained in the genome-wide OS analysis using the FUMA platform,³⁹ with our summary statistics as input data (accessed on June 30, 2025). With the SNP2GENE function, we performed both positional mapping (with a maximum distance of 250 kb) and expression quantitative trait locus (eQTL) mapping (cis-eQTL within 10 kb; using GTEx V.8 for all tissues and the other default databases). Default settings were used for both SNP2GENE and GENE2FUNC, except for maximum p value for lead single nucleotide polymorphisms (SNPs) set at the suggestive significance threshold (p value< 1.0×10^{-5}), and for MAGMA analysis. In addition, variants were searched in QTLbase V.2.0 database (downloaded on November 13, 2024).

The coloc package,⁴⁰ in R, was used for colocalization analyses and plots, with default priors and incorporating our GWAS summary statistics alongside data from the eQTLGen dataset. This analysis evaluated four hypotheses in addition to the null hypothesis (H0), which states there is no association with either trait. The other hypotheses included: association with gene expression only (H1), or with OS only (H2), association with both traits through independent causal variants (H3), and a shared causal variant affecting both traits (colocalization, H4). A posterior probability ≥ 0.8 for H4 indicated strong evidence of colocalization, while values between 0.5 and 0.8 suggested moderate colocalization.⁴¹

Genes reported as being regulated by the most significant variants in our OS GWAS were further investigated with the Kaplan-Meier Plotter online tool (accessed on November 20, 2024). We looked for associations between survival and the expression levels of these genes according to published gene expression data from lung tumor tissue (both adenocarcinoma and squamous cell carcinoma).⁴² We selected the option of using only the best probe set and a follow-up threshold of 24 months. Multivariable Cox regression with sex, stage, smoking status, and histology as covariates was performed. The other settings were the default ones. A two-sided p value<0.05 was set as the significance threshold for this analysis.

eQTL detection in lung tumor tissue

To test whether the significant eQTLs identified in databases were also present in lung tumor tissue, we used The Cancer Genome Atlas (TCGA) RNA-sequencing data (phs000178.v11.p8) from lung adenocarcinoma (LUAD) and squamous cell lung carcinoma (LUSC) tissues and germline genotyping data from the same patients.^{43 44} Genotyping and gene expression data were downloaded from the National Cancer Institute Genomic Data Commons portal,⁴⁵ accessed on October 25, 2024 (data access request number 131208). In the downloaded germline variants files, for each SNP was reported the genotype and a confidence score. Genotype calls with a score larger than 0.1 were set as missing and the data was processed to be used in PLINK V.2. Samples with call rates <95% were discarded, as well as duplicates and samples with excess of heterozygosity and sex inconsistencies. Variants with call rate <95%, MAF<0.01, and HWE test p value<1×10⁻¹⁰ were filtered out. PCA was carried out to identify and remove non-European patients. Imputation to whole sequence for chromosome 2 variants only was carried out, as described above for the whole genome, using the Michigan imputation server (V.2.0.6). From the downloaded RNA sequencing data, unstranded counts were normalized using Variance Stabilizing Transformation function of DESeq2 R package.⁴⁶ Expressed transcripts were annotated using GENCODE Release 36 (GRCh38.p13). Linear regressions, between the expression levels of candidate genes on chromosome 2 (ie, coding genes in a region 1Mb spanning the investigated SNPs; online supplemental table 2) and the genotypes (coded as numerical variables from 0 to 2 indicating an increasing number of minor alleles) of the top-10 SNPs identified as associated with OS, were carried out using Matrix eQTL package in R,⁴⁷ with sex, age, stage, and analysis batch (LUAD or LUSC) as covariates. A two-sided p value<0.05 was set as the significance threshold for this analysis.

Validation series, genotype imputation and association with OS

We obtained access to the exome sequencing data and clinical information (including OS data) from normal samples of 182 patients with NSCLC treated with ICI from phs002822.v1.p1 dataset³⁴ in dbGap. SRA files, obtained using SRA toolkit V.2.11.3, were converted into bam files and germline variants on chromosome 2 were called using DeepVariant V.1.6.1 (b37 genomic release, as reference) with the nf-core/sarek V.3.5.1 workflow.⁴⁸ The merged vcf file with passed variants was processed with Conform-gt (Apache License, V.2.0) and then used as input for the Michigan Imputation Server (reference panel: 1000 Genome Project, phasing: Eagle) for SNP imputation. Variants with MAF<5% and imputation quality score R²<0.3 were filtered out. Linkage disequilibrium (LD) between imputed variants and our top significant SNPs was calculated using the LDmatrix tool of LDlink.⁴⁹ Variants in LD (D'>0.8) with our top significant SNP were tested for association with OS in Cox proportional hazard model, with histotype, line of ICI

treatment, and recruiting center, as covariates in R with survival package (covariates were selected following the same method described above for the discovery series). Kaplan-Meier curves were drawn to compare individuals carrying at least one minor allele with patients homozygous for the major allele and log-rank p values were calculated, as above in R environment with survival package.

RESULTS

In this study, we genotyped genomic DNA samples from 166 patients with NSCLC, treated with ICI in monotherapy. Two samples did not pass the QC filters of the AxAS software. PLINK per-sample QC steps excluded two samples with a genotyping call rate <98%, seven samples with inconsistencies between recorded sex and sex imputed by the genotypes, and three samples for heterozygosity issues (online supplemental figure 1a). Thus, we analyzed the genotyping data from 152 patients.

Exploratory analyses of the effects of personal and clinical variables on ORR, PFS and OS.

Clinical characteristics of the patients are reported in table 1. The median age at enrollment was 72 years and two-thirds of patients were men. Most patients (85%) were former or current smokers and, as expected, of advanced stages (~90%), but with good performance status (ECOG=0 or 1; 93%). More than two-thirds of patients were diagnosed with LUAD (68%) and ~64% of patients received previous chemotherapy and/or radiotherapy. From a molecular point of view, 64% of patients were strong PD-L1 expressors. About 20% of the patients carried a mutation in the *KRAS* gene, whereas *EGFR* mutations and *ROS1* rearrangements were less frequent. However, the mutational status was not available for ~18% of the patients. Patients were unevenly distributed among the recruiting centers, according to their performance status (Fisher's exact test p value=0.0052), PD-L1 expression in their tumors (Fisher's exact test p value=0.011), and type and line of administered ICI (Fisher's exact test p values<0.001 and =0.0012, respectively; table 1). Patients enrolled in the study received ICI in monotherapy; in detail, pembrolizumab was administered to more than half of patients (mostly as first-line treatment), 27% received nivolumab in second line, as per Agenzia Italiana del Farmaco (AIFA) indications, and the remaining were treated with atezolizumab (mostly as second-line treatment; Table 1). Overall, the ORR to immunotherapy was 20%, similar to that reported in previous studies.⁵⁰ The median follow-up time of our patient series was 21.6 months (range 0.26–84) and, at 24 months of follow-up, about 54% of patients had disease progression and 43% of patients had died.

(AIFA) indications, and the remaining were treated with atezolizumab (mostly as second-line treatment; table 1). Overall, the ORR to immunotherapy was 20%, similar to that reported in previous studies. The median follow-up time of our patient series was 21.6 months

Table 1 Characteristics of the 152 patients included in the analyses, grouped by recruiting center. *P*-values in bold are statistically significant.

Characteristic	Recruiting center					<i>P</i> value*
	Total (n=152)	FBF (n=28)	HUM (n=44)	OLE (n=15)	POL (n=22)	SGM (n=43)
Age at enrollment (range)	72 (45–91)	69 (45–85)	68 (52–86)	69 (56–80)	74 (53–91)	73 (54–84)
Sex, n (%)	100 (65.79)	13 (46.43)	31 (70.45)	4 (26.67)	13 (59.09)	32 (74.42)
Male						
Female	52 (34.21)	15 (53.57)	13 (29.55)	11 (73.33)	9 (40.91)	11 (25.58)
Smoking status, n (%)	50 (32.90)	7 (25.00)	14 (31.82)	6 (40.00)	7 (31.82)	16 (37.21)
Current smoker						
Former smoker	81 (53.29)	11 (39.29)	25 (56.82)	9 (60.00)	12 (54.55)	24 (55.81)
Never smoker	20 (13.15)	10 (35.71)	5 (11.36)	0 (0.00)	2 (9.09)	3 (6.98)
Not available	1 (0.66)	0 (0.00)	0 (0.00)	0 (0.00)	1 (4.55)	0 (0.00)
Stage, n (%)	9 (5.92)	1 (3.57)	5 (11.36)	0 (0.00)	2 (9.09)	1 (2.33)
I						
II	4 (2.63)	0 (0.00)	1 (2.27)	2 (13.33)	1 (4.55)	0 (0.00)
III	24 (15.79)	4 (14.29)	10 (22.73)	0 (0.00)	2 (9.09)	8 (18.6)
IV	113 (74.34)	23 (82.14)	28 (63.64)	13 (86.67)	15 (68.18)	34 (79.07)
Not available	2 (1.32)	0 (0.00)	0 (0.00)	0 (0.00)	2 (9.09)	0 (0.00)
Histotype, n (%)	104 (68.42)	24 (85.71)	25 (56.82)	9 (60.00)	17 (77.27)	29 (67.44)
Adenocarcinoma						
Squamous carcinoma	39 (25.66)	3 (10.71)	12 (27.27)	6 (40.00)	5 (22.73)	13 (30.23)
Large cell carcinoma	2 (1.32)	1 (3.57)	1 (2.27)	0 (0.00)	0 (0.00)	0 (0.00)
Other NSCLC	7 (4.61)	0 (0.00)	6 (13.64)	0 (0.00)	0 (0.00)	1 (2.33)
Performance status (ECOG) at baseline, n (%)	88 (57.89)	19 (67.86)	30 (68.18)	13 (86.67)	8 (36.36)	18 (41.86)
0						
1	54 (35.53)	7 (25)	14 (31.82)	2 (13.33)	11 (50)	20 (46.51)
2	8 (5.26)	1 (3.57)	0 (0.00)	0 (0.00)	3 (13.64)	4 (9.3)
Not available	2 (1.32)	1 (3.57)	0 (0.00)	0 (0.00)	0 (0.00)	1 (2.33)
Previous therapy, n (%)	42 (27.63)	12 (42.86)	0 (0.00)	4 (26.67)	11 (50.00)	15 (34.88)
Chemotherapy						
Radiotherapy	13 (8.55)	1 (3.57)	12 (27.27)	0 (0.00)	0 (0.00)	0 (0.00)
Chemo/radiotherapy	41 (26.97)	8 (28.57)	12 (27.27)	6 (40.00)	3 (13.64)	12 (27.91)
None	53 (34.87)	7 (25.00)	18 (40.91)	5 (33.33)	7 (31.81)	16 (37.21)
Not available	3 (1.97)	0 (0.00)	2 (4.55)	0 (0.00)	1 (4.55)	0 (0.00)
PD-L1 expression in the tumor, n (%)	36 (23.68)	5 (17.86)	6 (13.64)	4 (26.67)	7 (31.82)	14 (32.56)
<1%						
1%–49%	16 (10.53)	4 (14.29)	1 (2.27)	1 (6.67)	2 (9.09)	8 (18.6)
≥50%	82 (53.95)	10 (35.71)	37 (84.09)	7 (46.67)	9 (40.91)	19 (44.19)
Not available	18 (11.84)	9 (32.14)	0 (0.00)	3 (20.00)	4 (18.18)	2 (4.65)

Continued

Table 1 Continued

Characteristic	Recruiting center						P value*	
	Total (n=152)	FBF (n=28)	HUM (n=44)	OLE (n=15)	POL (n=22)	SGM (n=43)		
Mutational status, n (%)	KRAS	30 (19.74)	4 (14.29)	10 (22.73)	1 (6.67)	3 (13.64)	12 (27.91)	n.a.
	EGFR	5 (3.29)	2 (7.14)	1 (2.27)	0 (0.00)	0 (0.00)	2 (4.65)	
	ROS1	7 (4.61)	0 (0.00)	0 (0.00)	0 (0.00)	1 (4.55)	6 (13.95)	
Administered ICI, n (%)	Not available	27 (17.76)	3 (10.71)	8 (18.18)	6 (40.00)	3 (13.64)	7 (16.28)	
	Atezolizumab	31 (20.40)	5 (17.86)	7 (15.91)	2 (13.33)	6 (27.27)	11 (25.58)	<0.001
	Nivolumab	42 (27.63)	12 (42.86)	0 (0.00)	7 (46.67)	7 (31.82)	16 (37.21)	
	Pembrolizumab	79 (51.97)	11 (39.29)	37 (84.09)	6 (40.00)	9 (40.91)	16 (37.21)	
	First	66 (43.42)	20 (71.43)	12 (27.27)	10 (66.67)	14 (63.64)	27 (62.79)	0.0012
Line of ICI administration, n (%)	Subsequent†	83 (54.61)	8 (28.57)	30 (68.18)	5 (33.33)	7 (31.81)	16 (37.21)	
	Not available	3 (1.97)	0 (0.00)	2 (4.55)	0 (0.00)	1 (4.55)	0 (0.00)	
	Response‡	31 (20.39)	5 (17.86)	14 (31.82)	3 (20.00)	4 (18.18)	5 (11.63)	0.069
Response to immunotherapy, n (%)	Stable disease	32 (21.05)	4 (14.29)	8 (18.18)	7 (46.67)	5 (22.73)	8 (18.60)	
	Disease progression	89 (58.55)	19 (67.86)	22 (50)	5 (33.33)	13 (59.09)	30 (69.77)	
Alive/dead status§, n (%)	Alive	86 (56.58)	16 (57.14)	25 (56.82)	15 (100)	13 (59.09)	17 (39.53)	<0.001
	Dead	66 (43.42)	12 (42.86)	19 (43.18)	0 (0.00)	9 (40.91)	26 (60.47)	

n.a., not analyzed due to the high number of missing values.

*Fisher's exact test (or Kruskal-Wallis for age, which is a continuous variable) p value.

†All patients treated with nivolumab (n=42), as per AIFA indications, 71% of atezolizumab-treated patients (n=24), and 21.5% of patients who received pembrolizumab (n=17).

‡Complete or partial response.

§At 24 months of follow-up, after immunotherapy start.

SGM, Fondazione IRCCS San Gerardo dei Tintori; ECOG, Eastern Cooperative Oncology Group; FBF, Fatebenefratelli-Sacco Hospital; HUM, IRCCS Istituto Clinico Humanitas; ICI, immune checkpoint inhibitor; NSCLC, non-small cell lung cancer; OLE, Legnano Hospital; PD-L1, programmed death ligand-1; POL, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico.

(range 0.26–84) and, at 24 months of follow-up, about 54% of patients had disease progression and 43% of patients had died.

We performed exploratory analyses (in univariable and multivariable models) to identify clinical variables associated with ORR, PFS, and OS, which might be confounding factors in the genetic analyses. Although some ORR univariable analyses gave significant results, the multivariable analysis did not show any significant association between clinical variables and the response to ICI (p value >0.05 ; online supplemental table 3). Therefore, the GWAS with ORR phenotype was not corrected for any covariate. Instead, the multivariable Cox model, with the variables significantly associated with PFS in the univariable analyses (ie, age, performance status, PD-L1 expression, ICI drug, and the recruiting center), indicated that the administered ICI and the recruiting center were significant independent confounding factors for PFS (table 2). Similarly, among the variables significantly associated with OS in the univariable analyses (ie, performance status, PD-L1 expression, and ICI), the administered ICI and the performance status were significant independent confounding factors (table 2). Therefore, we corrected the genetic analyses for these confounders, by adding these covariates to the genome-wide Cox models with the genotypes. Kaplan-Meier curves are shown in figure 1.

GWAS identifies variants on chromosome 2 associated with OS of patients with ICI-treated lung cancer

For each patient we genotyped 855,568 variants. In the per-marker QC steps, 420 markers with missing rates $>5\%$ were removed, 439,253 polymorphisms with $MAF < 1\%$ were filtered out, and 208 variants were excluded because in Hardy-Weinberg disequilibrium (p value $< 1.0 \times 10^{-10}$). The clean dataset, including 152 patients and 415,687 SNPs, was used for imputation to whole genome sequence. Imputation quality and MAF filters left 5,313,198 polymorphisms that were used in the genome-wide analyses (online supplemental figure 1b).

We first performed the GWAS for the ORR phenotype. No genome-wide statistically significant variants were identified. Two SNPs, on chromosomes 7 and 20, passed the suggestive threshold (p value $< 1.0 \times 10^{-5}$, (online supplemental figure 2a), but they were isolated variants, and no peaks of associations were distinguishable.

Then, we analyzed the association between polymorphisms and PFS. Although this analysis did not identify genome-wide statistically significant variants, 44 SNPs passed the suggestive threshold, set at p value $< 1.0 \times 10^{-5}$ (online supplemental table 4 and figure 2b). Among these, there were 19 variants mapping in an intergenic region, spanning 10kb, on chromosome 8, ahead of *MCPH1* gene, previously proposed as a putative tumor suppressor gene in lung cancer.⁵¹ The minor alleles of these variants were associated with a higher risk of disease progression ($HR > 1$), suggesting a possible role as negative prognostic factors.

Interestingly, the OS genome-wide analysis identified 10 nearly significant variants (p value $= 6.3 \times 10^{-8}$) on chromosome 2, in complete LD with each other (figure 2). Patients heterozygous or homozygous for the minor alleles of these SNPs had a higher risk of death than patients homozygous for the major alleles ($HR = 5.1$, 95% CI: 2.9 to 9.2). A Kaplan-Meier curve was plotted, showing the survival probability curves of patients, grouped according to the genotype at rs111648355. As in a dominant model, heterozygotes and patients homozygous for the minor allele were grouped and compared with patients homozygous for the major allele (online supplemental figure 3). In the same region on chromosome 2, spanning from position 48,163,611 to 48,755,968bp, further 274 variants, in high LD with the top 10 ($R^2 \geq 0.6$, $D' \geq 0.8$), were found to be associated with OS, at p value $< 1.0 \times 10^{-5}$ (online supplemental table 5). All these polymorphisms had $HR > 1$, indicating that they were all negative prognostic markers.

Overall, 406 variants were associated with OS at p value $< 1.0 \times 10^{-5}$ (figure 2 and online supplemental table 5). Among them, 53 SNPs mapped on chromosome 4, in the region spanning from 70,045,614 to 70,356,404bp, and other 20 were on the long arm of chromosome 2 (from position 210,393,521 to 210,450,687bp). Other suggestive association signals were on chromosomes 3, 5, 7, 11, 12, 13, 14, 16, 17, 18, 20, and 22. Except for the three variants on chromosome 17, the minor alleles of all the other SNPs were associated with a higher risk of death ($HR > 1$).

Then, we wanted to understand whether the identified locus was specifically associated with OS after ICI treatment. To do this, we took advantage of our previous results of a survival GWAS in a large series of patients with LUAD who were not treated with ICI (and not overlapping with the series herein investigated).⁵² We compared the summary statistics of 207 out of 406 SNPs associated with OS in patients treated with ICI obtained in the two studies (the summary statistics of the remaining 199 variants were not available in the previous study). Only rs59650998 on chromosome 22 was nominally significant in the previous analysis (p value $= 0.02$). Looking at the top 10 SNPs (nearly significant at genome-wide level), only 5 were analyzed also in the previous analysis, and they had p values ranging from 0.5 to 0.7, thus confirming that the SNPs identified in this study are specifically associated with OS in the ICI context. Online supplemental table 6 shows the p values, obtained in our previous genome-wide OS analysis, of the 207 SNPs associated with OS after ICI treatment; p values obtained in this new study are also reported for comparison.

In silico functional annotation suggests a possible regulatory role for OS-associated variants

We annotated the results obtained from the genome-wide OS analysis using FUMA. Looking at the most significant locus on chromosome 2, whose regional plot is reported in figure 3, two independent SNPs were identified

Table 2 Factors associated with progression-free and overall survival (PFS and OS), according to univariable and multivariable Cox regression models. *P*-values in bold are statistically significant.

	PFS			OS		
	Univariable		Multivariable*	Univariable		Multivariable*
	HR (95% CI)	P value		HR (95% CI)	P value	
Age	1.03 (1.00 to 1.06)	0.027	1.02 (0.995 to 1.06)	1.02 (0.991 to 1.05)	0.11	1.02 (0.991 to 1.05)
Sex female	1.00			1.00		
Male	1.49 (0.922 to 2.41)	0.10		1.55 (0.902 to 2.67)	0.11	
Smoking status never	1.00			1.00		
Former	0.906 (0.482 to 1.70)	0.76		1.20 (0.534 to 2.68)	0.66	
Current	0.561 (0.274 to 1.15)	0.11		0.991 (0.419 to 2.35)	0.98	
Stage ≤III	1.00			1.00		
IV	1.03 (0.613 to 1.71)	0.97		0.926 (0.526 to 1.63)	0.79	
Histotype adenocarcinoma	1.00			1.00		
Other	1.23 (0.784 to 1.94)	0.36		1.42 (0.861 to 2.33)	0.17	
ECOG performance status 0	1.00			1.00		1.00
1 or 2	2.02 (1.30 to 3.12)	0.0017		2.19 (1.34 to 3.59)	0.0017	1.69 (1.00 to 2.86)
Previous therapy no	1.00			1.00		
Yes	1.29 (0.813 to 2.05)	0.28		1.25 (0.739 to 2.12)	0.40	
Line of ICI administration first	1.00			1.00		
Subsequent	1.43 (0.918 to 2.24)	0.11		1.33 (0.807 to 2.20)	0.26	
PD-L1 expression in the tumor <1%	1.00			1.00		
1–49%	0.754 (0.369 to 1.54)	0.44		0.828 (0.379 to 1.81)	0.64	
≥50%	0.500 (0.300 to 0.834)	0.0080		0.542 (0.311 to 0.944)	0.031	
Recruiting center FBF	1.00		1.00	1.00		
HUM	0.810 (0.434 to 1.51)	0.51	1.00 (0.479 to 2.10)	0.99	1.11 (0.538 to 2.28)	0.78
OLE	0.206 (0.0603 to 0.700)	0.011	0.102 (0.0225 to 0.465)	0.0032	1.41×10 ⁻⁸ †	1.0
POL	0.598 (0.276 to 1.30)	0.19	0.342 (0.139 to 0.840)	0.019	0.887 (0.374 to 2.11)	0.79
SGM	1.37 (0.760 to 2.46)	0.3	0.937 (0.480 to 1.83)	0.85	1.81 (0.909 to 3.59)	0.091
ICI pembrolizumab	1.00		1.00	1.00		1.00
Nivolumab	1.54 (0.919 to 2.59)	0.10	2.80 (1.41 to 5.53)	0.0031	1.59 (0.874 to 2.89)	0.13
Atezolizumab	2.67 (1.57 to 4.54)	<0.001	3.05 (1.61 to 5.80)	<0.001	3.45 (1.93 to 6.15)	<0.001

Continued

Table 2 Continued

	PFS		OS	
	Univariable		Univariable	
	HR (95% CI)	P value	HR (95% CI)	P value
			Multivariable*	Multivariable*
			HR (95% CI)	HR (95% CI)
			P value	P value

*Including variables that were statistically significant in the univariable model; only the results of the model selected with the stepwise procedure (based on Akaike information criterion, AIC) are shown.

†Group with zero events.

SGM, Fondazione IRCCS San Gerardo dei Tintori; ECOG, Eastern Cooperative Oncology Group; FBF, Fatebenefratelli-Sacco Hospital; HUM, IRCCS Istituto Clinico Humanitas; ICI, immune checkpoint inhibitor; OLE, Legnano Hospital; PD-L1, programmed death ligand-1; POL, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico.

(rs111648355 and rs7819296) and the SNP rs111648355 was prioritized as lead variant. Positional mapping of the SNPs was performed: *FBXO11* is the coding gene closest to the top-significant variants and contains most of the SNPs of the locus passing the suggestive threshold of significance.

Among the SNPs in the identified genomic locus at p value $< 1.0 \times 10^{-5}$, there was an exonic variant of *PPP1R21* gene, rs76587580 (NM_001135629.3:c.261C>G); it is a synonymous variant (NP_001129191.1:p.Gly87=) with a Combined Annotation Dependent Depletion (CADD) score=12.29, and reported as benign in ClinVar. All the other polymorphisms, including the 10 top-significant ones, mapped in non-coding regions of the genome (mainly in introns and intergenic regions). In addition, rs10495950 had the highest RegulomeDB ranking score (=1b), meaning that it is likely to affect transcription factor binding, gene expression, or chromatin accessibility. Other three variants (rs7583719, rs11683689, and rs11692537) are likely to be functional regulatory variants, having a RegulomeDB ranking score=1f.

All the 284 variants in the region were reported to be eQTLs of at least 1 of the 11 genes mapping in this same region (*STPG4*, alias *C2orf61*, *CALM2*, *FOXN2*, *FSHR*, *GTF2AIL*, *LHCGR*, *MSH2*, *MSH6*, *PPP1R21*, *STON1*, and *STON1-GTF2AIL*) in several tissues (false discovery rate (FDR) < 0.05 ; online supplemental table 7). Only three eQTLs were reported in lung tissue: rs11693533, rs77902227, and rs6747738 associated with *GTF2AIL* expression. The most numerous eQTLs were those of *PPP1R21*, followed by *GTF2AIL*, *MSH6*, *STON1*, and *STON1-GTF2AIL*, while the most significant ones (p value $< 5.0 \times 10^{-8}$) were those of *MSH6*, followed by *PPP1R21* and *STON1*. Our 10 top-significant variants associated with OS were most significantly associated with *MSH6* expression in blood tissue. Mainly, the minor alleles of these SNPs were significantly associated with lower gene expression levels.

In addition, we searched our OS-associated variants in the QTLbase dataset, where data about different types of QTLs are recorded, such as, for example, QTLs of methylation (mQTLs), splicing, and chromatin accessibility, just to cite some of them. Moreover, QTLbase includes eQTLs data not considered by FUMA. This is particularly relevant for the *FBXO11* gene for which FUMA did not report eQTLs. We observed that our top-significant variants were already reported as eQTLs also of the *FBXO11* gene, in blood (online supplemental table 8). In this same tissue, the 10 SNPs most significantly associated with OS were also reported as mQTLs of *MSH6*, *FBXO11*, and *STON1-GTF2AIL*.

We then used the eQTLGen data to perform colocalization analyses and test whether the genetic factors associated with OS in our patients with ICI-treated cancer were the same as those regulating the expression of *MSH6*, *MSH2*, *PPP1R21* and *STON1* genes (ie, the four eQTL target genes in blood). The highest posterior probabilities (PP) for shared variants regulating both expression

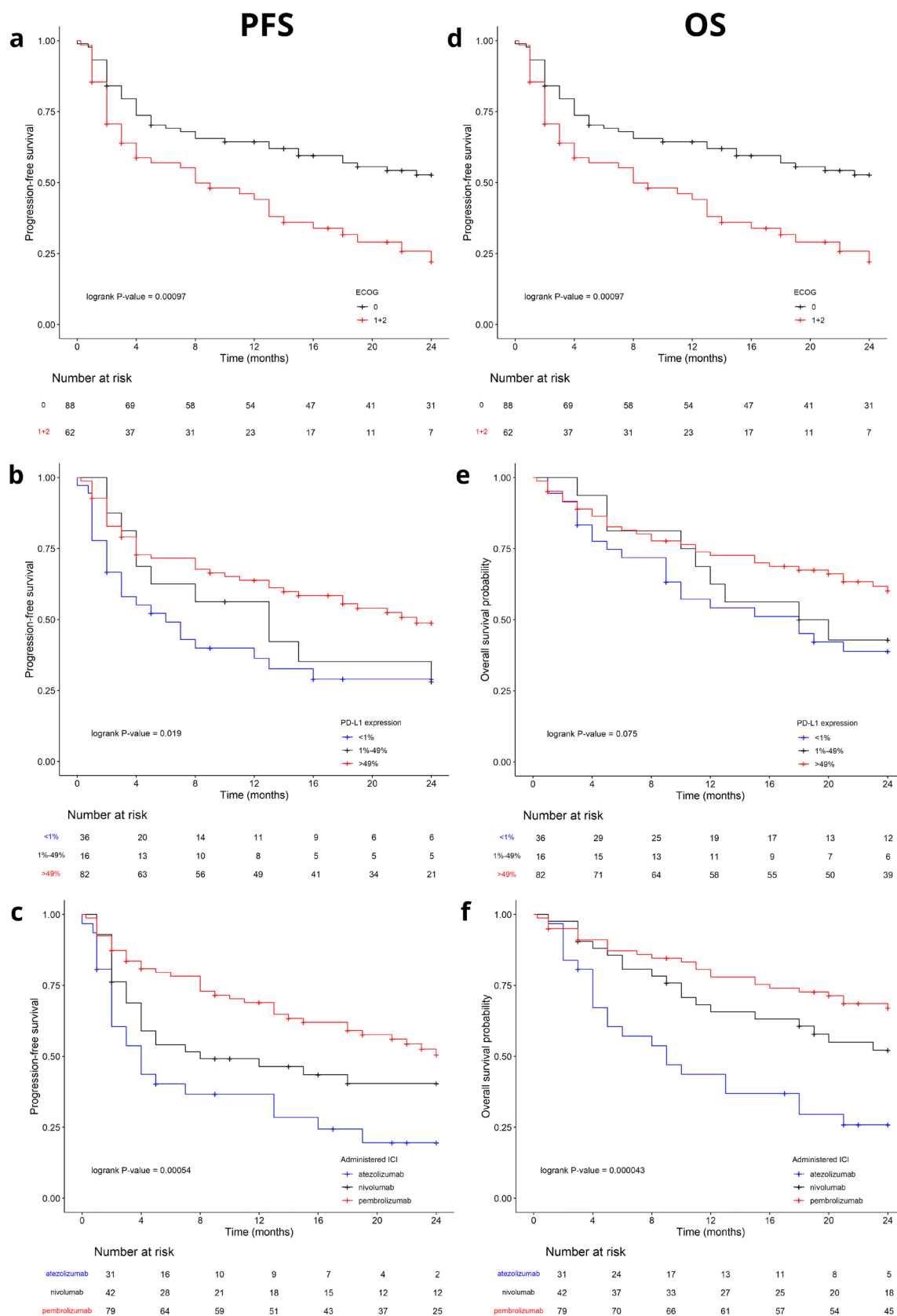


Figure 1 Variables associated with PFS and OS in univariable Cox models. Kaplan-Meier (KM) survival curves (at 24 months of follow-up) of patients with lung adenocarcinoma treated with ICI, according to their ECOG performance status (a, d), programmed death-ligand 1 expression level in their tumors (b, e), and taken ICI (c, f). PFS and OS KM are reported in the left and right panels, respectively. Crosses denote censored samples. Below each plot are indicated the numbers of patients at risk in each group. Log-rank p values are shown. ECOG, Eastern Cooperative Oncology Group; OS, overall survival; PFS, progression-free survival.

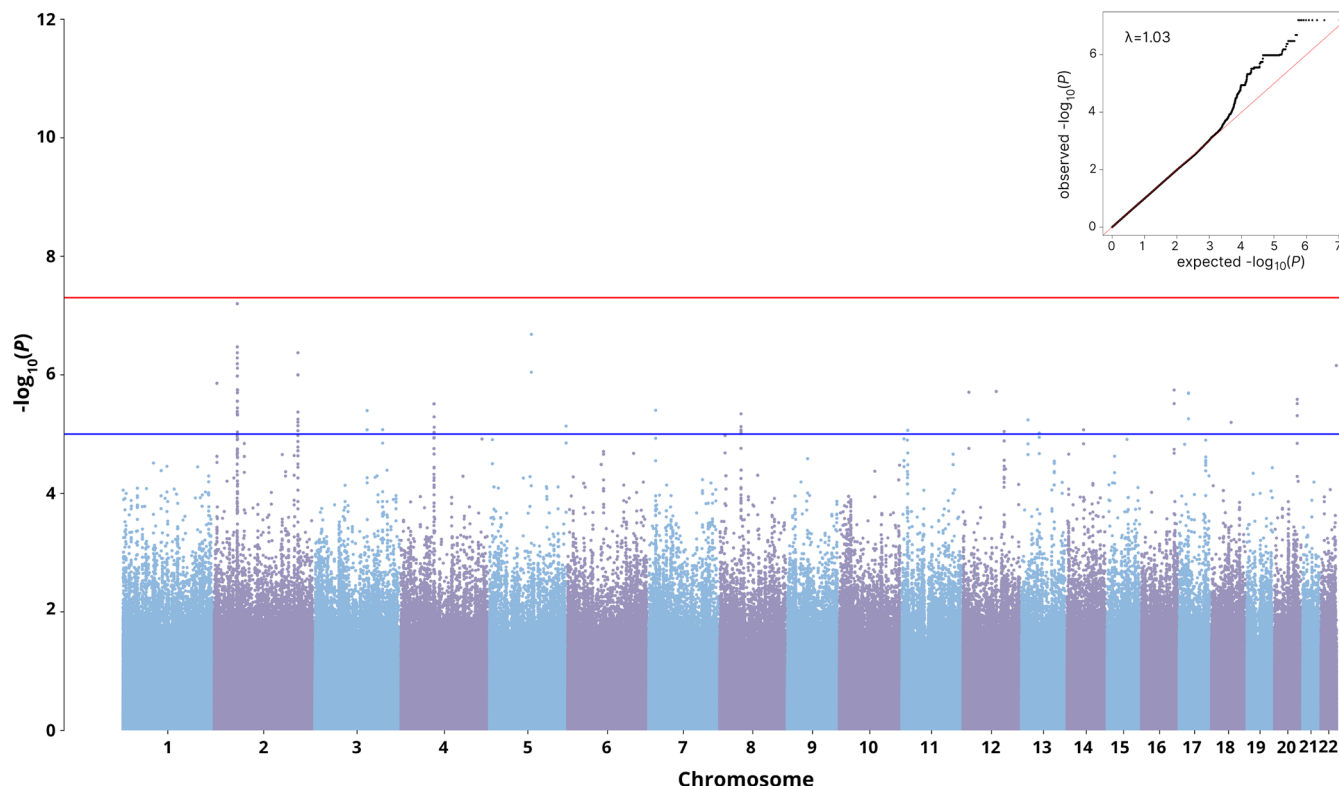


Figure 2 A nearly significant locus on chromosome 2 is associated with overall survival of patients with immune checkpoint inhibitor-treated lung adenocarcinoma. Manhattan plot of the results of the genome-wide association for overall survival, at 24 months of follow-up, with age, sex, drug, and ECOG as covariates ($n=150$, due to missing ECOG data for two patients). Imputed variants are plotted along the x-axis according to their genomic position (GChr37, hg19 release). On the y-axis is reported the level of association ($-\log_{10}(p \text{ values})$) with survival probability of each single nucleotide polymorphism. The horizontal red and blue lines represent the threshold of genome-wide significance ($p < 5.0 \times 10^{-8}$) and a suggestive threshold at $p < 1.0 \times 10^{-5}$, respectively. In the upper-right corner the QQ plot of observed and expected p values is shown. Genomic inflation factor (λ) is reported. ECOG, Eastern Cooperative Oncology Group.

and OS (H4) were found for *PPP1R21* ($PP_{H4}=66.8\%$) and *MSH2* ($PP_{H4}=66.2\%$), followed by *MSH6*, with $PP_{H4}=46.2\%$; finally, the PP_{H4} for *STON1* was the lowest (23.3%). The results of colocalization analyses are reported in online supplemental table 9, while figure 4a–d shows the colocalization plots. We carried out the same analysis for the *FBXO11* gene, using the expression data of this gene in blood, reported in the dataset described in.⁵³ Low evidence for shared variants affecting both OS and *FBXO11* expression was observed ($PP_{H4}=33.2\%$; online supplemental table 9 and figure 4e).

Then, we used TCGA expression and genotyping data from patients with lung cancer ($n=709$, including 358 LUAD and 351 LUSC; online supplemental figure 4 shows the quality check steps followed to filter the downloaded data) to look for the same eQTLs reported in blood. In online supplemental table 10, the characteristics of this patient series are summarized. Of note, these patients are quite different from ours since most of them had tumors at lower stages than those in our study. Nevertheless, the eQTL analysis in lung tumor tissue, investigating the 10 top-significantly OS-associated SNPs and 17 coding

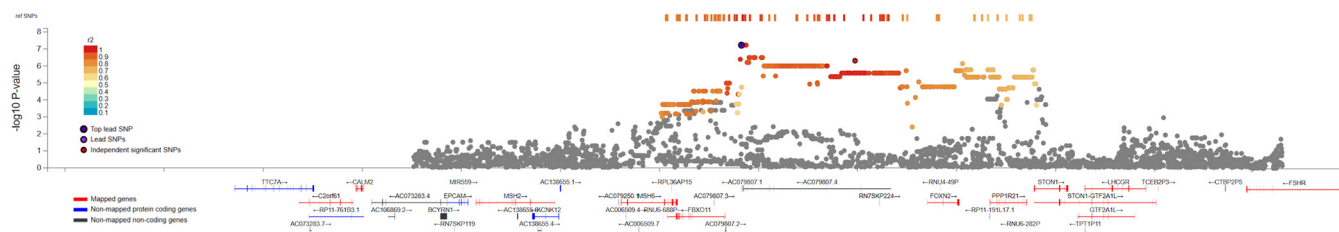


Figure 3 Regional plot of the locus on chromosome 2 identified in the genome-wide association. Plot spans the region from 47.15 Mbp to 49.25 Mbp (GRCh37/hg19), containing the mapped genes and the analyzed imputed variants. SNPs are plotted on the x-axis according to their position on chromosome 2. Dot color represents the level of linkage disequilibrium, expressed as R^2 between each SNP and the lead variant (rs111648355, purple dot). SNP, single nucleotide polymorphism.

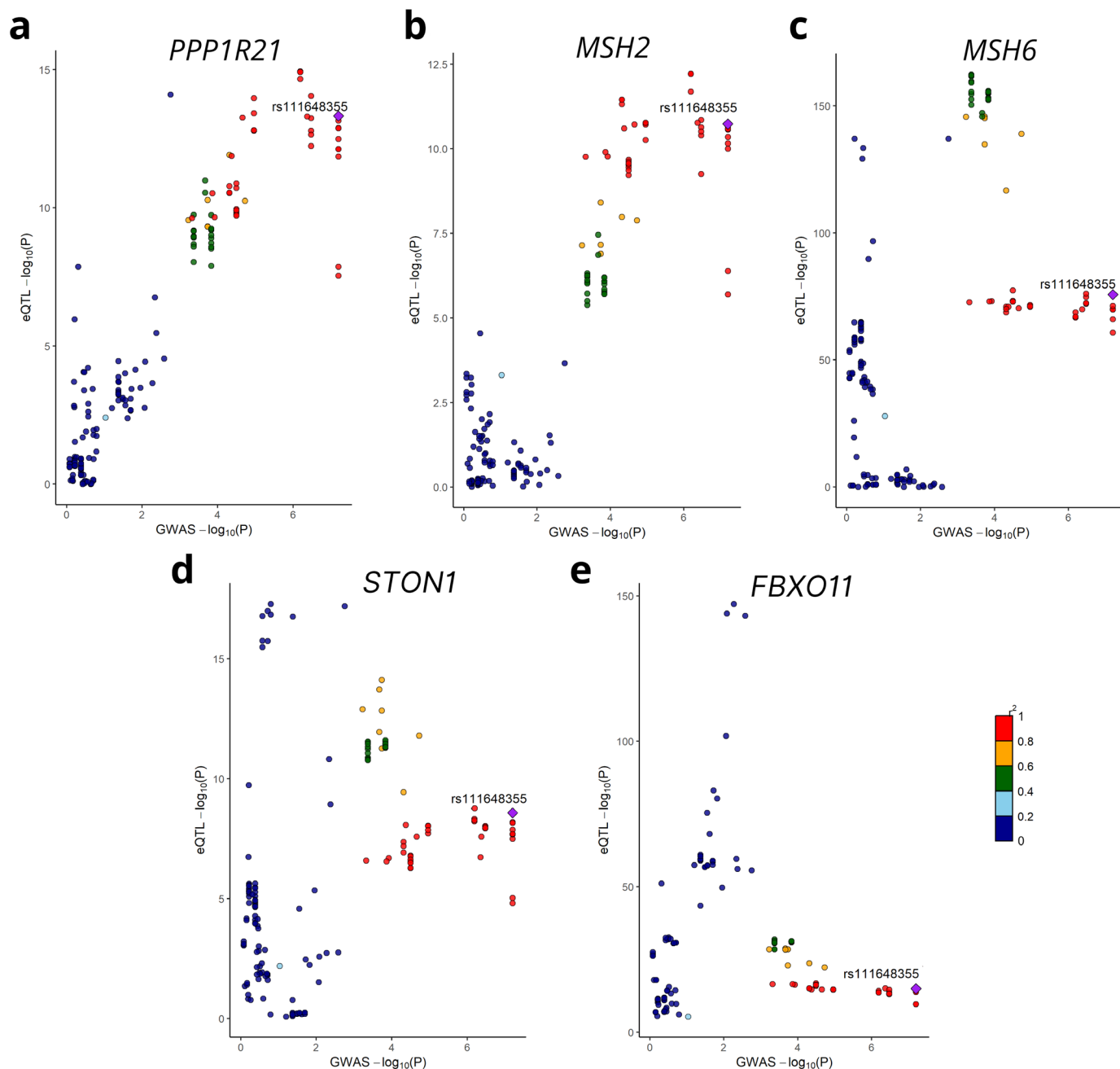


Figure 4 The evidence for shared variants between OS and gene expression is moderate. Colocalization plots for the (a) *PPP1R21*, (b) *MSH2*, (c) *MSH6*, and (d) *STON1* genes using the whole-blood tissue expression data from the eQTLGen database (numbers of SNPs included in the analyses were 190); (e) colocalization plot for *FBXO11*, with expression data from, including 180 SNPs. Dots represent variants along the portion of chromosome 2 spanning from 48.0 to 48.2 Mb and are colored based on linkage disequilibrium (R^2) with the lead variant of our genome-wide association for OS, rs111648355 (purple dot). Dot coordinates are defined, on the x-axis, by the p values ($-\log_{10}$ -transformed) of association with OS, whereas, on the y-axis, those with the gene expression levels. OS, overall survival; SNP, single nucleotide polymorphism.

genes in the same region of chromosome 2, found associations between the genotypes of these variants and the expression levels of *STON1-GTF2A1L*, *MCFD2*, *FOXN2*, *PPP1R21*, *GTF2A1L*, and *LHCGR* (p value < 0.05; online supplemental table 11). A negative correlation between the number of minor alleles and the expression levels of *STON1-GTF2A1L*, *GTF2A1L*, and *LHCGR* was observed, whereas for the remaining genes, there was a positive correlation.

Finally, taking advantage of public gene expression data in lung tumor tissue⁴² we looked for associations between OS (at 24 months of follow-up) and the expression levels of genes, in the locus on chromosome 2, for which a possible regulatory role of the identified variants was described. Although the probability of survival of patients expressing high or low levels of *MSH2*, *MSH6*, *MCFD2*, *GTF2A1L*, *STON1-GTF2A1L*, and *LHCGR* in lung tumor tissue was not significantly different, low levels of *FBXO11*,

PPP1R21, *FOXN2*, and *STON1* were associated with a poor prognosis (online supplemental figure 5). The associations with OS were significant with both the log-rank test and multivariable Cox proportional hazard model, with sex, stage, smoking and histotype as covariates.

Validation and comparison with previously reported associations

To validate our findings, we used an independent series of 182 patients with NSCLC treated with ICI, with publicly available data from exome sequencing of normal samples (blood or lung tissue).³⁴ From these exome data, we called 516,386 germline variants, which were used to impute genomic variants on chromosome 2 ($n=3,788,689$). Imputation quality filtering retained 39,813 variants and only 19,293 of them passed the MAF filter. Unfortunately, none of the SNPs significantly associated with OS in our GWAS study at $p\text{ value}<1.0\times10^{-5}$, were imputed with sufficient quality. This was quite expected as they were mainly intergenic variants, far from exonic regions. Therefore, we searched for imputed variants in LD with our top significant SNPs in a region spanning from 47.15 Mbp to 49.25 Mbp. This region contained 18 SNPs and only four of them were in LD ($D'>0.8$) with our 10 OS-associated variants (online supplemental table 12). Therefore, we tested the association with OS, at 24 months of follow-up, of these four polymorphisms in a subset of 142 patients with complete OS information, of Caucasian origin (considering the ethnicity reported in the dataset), and who received monotherapy ICI (online supplemental table 13). This subset was selected for homogeneity with our series. Nonetheless, the two series were significantly different, mainly in terms of age, sex, and administered ICI (Kruskal-Wallis or Fisher's exact test $p\text{ values}<0.001$; online supplemental table 13). In addition, they significantly differ for the median follow-up time (10.5 months in the validation series vs 21 in our series; Kruskal-Wallis test $p\text{ value}<0.001$). Having at least one minor allele of rs2072447 or rs3136316, two *MSH6* intronic variants, was a positive prognostic factor (log-rank $p\text{ values}=0.014$ and 0.15 , respectively; Kaplan-Meier curves are shown in online supplemental figure 6). Also, multivariable Cox models with clinical variables resulted significantly associated with OS (ie, histotype, line of treatment, and recruiting center, (online supplemental table 14) as covariates, showed the association of the minor allele of rs2072447 and rs3136316 with better prognosis (rs2072447: HR=0.53, 95% CI: 0.3 to 0.91, $p\text{ value}=0.02$; rs3136316: HR=0.55, 95% CI: 0.31 to 0.97, $p\text{ value}=0.04$).

We searched the literature for previously reported associations between germline variants and the response to ICI treatment (in terms of OS or PFS) in patients with lung cancer. Several candidate gene studies, mainly focused on variants in *CD274* (PD-L1), but also in *ERCC1*, *ALDH2*, *CD47*, *TLR4*, *LEP*, and *ABO* genes, reported some significant associations (at $p\text{ value}<0.05$).^{12 14 19 54–63} We compared the reported $p\text{ values}$ and HR with the results from our OS GWAS. Online supplemental table 15 summarizes the results of these comparisons. Briefly,

5 of 15 previously reported variants had a nominal $p\text{ value}<0.05$ also in our genome-wide analysis: 4 were located in PD-L1 variants and 1 is in *ABO* gene. However, only for rs7041009 in PD-L1, the direction of the effect of the minor allele (A) was consistent across both studies, being associated with a poor prognosis.⁵⁶

Furthermore, we compared the statistics of associations with the response to immunotherapy (but not in terms of OS) reported in a study including 279 patients with different types of cancer, that investigated 525 SNPs acting as tumor immune microenvironment eQTLs.²² The authors reported 32 variants associated with ICI response (at $p\text{ value}<0.05$) and, among them, only 1 (rs2070746) had nominally significant $p\text{ value}$ ($=0.031$) in our study, with the T allele associated with a better OS (HR=0.64) in our study, and with a better response in the study by Pagadala *et al* (Z score=2.18). In a subsequent article,²³ the same authors re-analyzed their data and reported an extended list of 229 variants associated with ICI response. 10 of them had $p\text{ value}<0.05$ also in our genome-wide OS analysis and some displayed concordant effects (online supplemental table 16).

Finally, we checked whether associations reported by other groups,^{64 65} investigating the genetics of antitumor immunity in general and broadly across different tumor types, showed nominal significance in our study. The two SNPs (rs3366 and rs4819959), associated with the amount of infiltrating follicular helper T cells and with the T helper 17 cell signature in the study by Shahamatdar *et al*, were not significantly associated with the OS of our ICI-treated patients ($p\text{ value}=0.85$ and $=0.081$, respectively). Among the 598 SNPs, identified by Sayaman *et al* as genome-wide significantly associated with 10 immunological traits in 30 different solid tumors, 10 had a $p\text{ value}<0.05$ also in our study (online supplemental table 16).

DISCUSSION

In this study, we investigated germline genetic factors associated with the response to immunotherapy, in terms of ORR, PFS, and OS at 24 months of follow-up, of patients with lung cancer treated with ICIs, adjusting for possible confounders such as age, sex, the recruiting center, and the administered drug. As expected from a small sample size, we did not find significant associations for ORR and PFS. Nevertheless, we identified a nearly genome-wide significant locus on chromosome 2 associated with the OS of patients with lung cancer, 24 months after the beginning of ICI treatment. This association is specific to patients with ICI-treated lung cancer, as the variants herein identified were not reported as prognostic factors for patients with lung cancer in a previous GWAS.⁵² Of note, the identified germline variants had already been reported as regulatory polymorphisms of the expression, in blood tissue, of the genes mapping in the locus, including *MSH6*, *MSH2*, *FBXO11*, *PPP1R21*, and *STON1*.

The minor alleles of the identified variants were associated with low messenger RNA (mRNA) levels of mutS

homolog 6 (*MSH6*) and high levels of mutS homolog 2 (*MSH2*) genes, two key players in the DNA mismatch repair system. The proteins encoded by *MSH6* and *MSH2* heterodimerize to form a complex that binds and dissociates mismatches. When these genes are mutated and their function is reduced, mismatches are not properly repaired and there is a hypermutability condition called microsatellite instability (MSI), which predisposes to several types of tumors,⁶⁶ but not lung cancer. Clinical studies of the efficacy of pembrolizumab including patients harboring cancers with MSI have reported good responses to immunotherapy (as reviewed in⁶⁷). Therefore, the US Food and Drug Administration has approved MSI as a biomarker for pembrolizumab indication regardless of the primary tumor site.^{10 68} The available eQTL data indicated that the minor alleles of our top-significant variants, associated with negative response to ICI in terms of OS, were associated with reduced *MSH6* expression and with higher levels of *MSH2* in blood. Therefore, we could not rule out the possibility that patients with these germline variants have a more efficient mismatch repair system that might predispose to a poor response to ICI.

We investigated whether the same variants regulated the expression levels of *MSH2* and *MSH6* in lung tumor tissue (from the TCGA dataset), but we did not find any significant associations. However, patients from the TCGA dataset were not treated with immunotherapy and prevalently had stage I tumors; in addition, data were only available for 709 patients, which is a limited number as compared with that of eQTLGen, which reported the above-mentioned eQTLs. Thus, a better and larger dataset should be needed. Further functional investigations should also be carried out to understand the mechanism by which the identified variants negatively affect OS.

The identified polymorphisms were also reported to regulate the expression of *STON1*; in particular, there was an inverse correlation between the number of minor alleles of the identified SNPs and the expression level of *STON1* in blood. To the best of our knowledge, no information is available on the role of this gene in lung cancer, but a study in renal clear cell carcinoma⁶⁹ showed that patients with low *STON1* expression had a worse prognosis than those with high levels of this gene and a greater tumor infiltration of immune cells. This latter observation led those authors to hypothesize that low *STON1* expression might be a marker of good response to immunotherapy. Although that hypothesis seems to contrast with our findings, it would be necessary to investigate the negative correlation between *STON1* expression in lung cancer tissue of patients treated with ICIs and the alleles associated with poor OS. So far, we know that low levels of *STON1* in lung tumor tissue were negative prognostic factors. Our eQTL analysis in lung tumor tissue from TCGA patients did not identify a statistically significant association between OS-associated SNPs and *STON1* levels but, as mentioned above, a more appropriate dataset of patients with lung cancer treated with ICI and with genotyping and expression data would be needed to better

investigate the regulatory role of the identified polymorphisms in tumor tissue.

The OS-associated variants were also reported as eQTLs of *FBXO11* gene, with minor alleles, here identified as negative prognostic factors, inversely correlated with the expression of this gene. A previous study showed that FBXO11 binds to class II major histocompatibility complex (MHC-II) transactivator (*CIITA*), mediating its ubiquitination and degradation, and thus contributing to the negative regulation of MHC-II in normal cells and breast cancer.⁷⁰ However, the authors did not find a significant association between *FBXO11* expression and survival of patients with breast cancer. Here, using already reported expression data in lung tumor tissue, we observed that low *FBXO11* expression was associated with poor prognosis. Unfortunately, the analysis with TCGA data did not find significant eQTLs in lung tumor for this gene. Further studies are needed to confirm the hypothesis that the mechanism underlying the associations observed herein between germline variants and poor OS in ICI-treated patients is exerted, at least in part, through their regulatory effect on *FBXO11* gene expression.

Finally, we cannot exclude a role of the *PPP1R21* gene, for which the strongest colocalization signal was measured. Very little is known about the function of its encoded protein. It has been implicated in mRNA transport across early endosomes, as part of the five-subunit endosomal Rab5 and RNA/ribosome intermediary complex.⁷¹ This has been observed in neurons, but no studies have been reported in other tissues. It has been shown that endosomal trafficking of plasma membrane PD-L1 in tumor cells is mediated by Rab5.⁷² Therefore, further studies are needed to elucidate the role of the *PPP1R21* gene and its regulatory variants in the response to ICI in patients with lung cancer.

To the best of our knowledge, no GWAS has been performed thus far that specifically looks for germline variants modulating the response to ICI in patients with lung cancer. Our pilot GWAS represents the first attempt in this direction, though it is limited by its small sample size and lack of a validation series. Unfortunately, no publicly available dataset currently includes genotyping data from patients with NSCLC treated with ICI that can be used for validation. In an attempt to confirm our results, we used exome sequencing data from normal samples of patients with NSCLC treated with ICI (accessible from dbGap).³⁴ As expected, imputing intergenic variants from this type of data was unsuccessful. Then, we tested the association of four imputed polymorphisms mapping in the region identified in our GWAS and in LD with our top significant variants with OS at 24 months after ICI treatment. Although two of these SNPs were significantly associated with OS, at a nominal p value < 0.05 in the validation series, these results are not a full validation. Thus, further studies with large and homogeneous patient series are needed.

Previous studies in patients with lung cancer undergoing immunotherapy have investigated candidate genes, typically focusing on a small number of variants in genes

involved in immune regulation or checkpoint pathways. We compared the reported associations with our results, which confirmed a possible prognostic role of rs7041009 in PD-L1 and of rs2070746 in *FPRI*. However, given the low significance and limited sample size, validation in other patient series is necessary.

In conclusion, we are aware that the results presented herein need further validation in larger cohorts of patients with lung cancer treated with ICI as monotherapy. Indeed, the most significant locus identified by our analysis did not reach the genome-wide significance, and the moderate evidence of colocalization between the leading variants and gene expression is also likely due to the low statistical significance of our OS GWAS results. Indeed, the posterior probabilities of the hypotheses of an involvement of the analyzed variants only in the OS phenotype were very low. Also, functional studies will be helpful in understanding the biological mechanism underlying the observed association with OS. However, our findings of a nearly genome-wide significant peak of association with OS in patients treated with ICI are very promising and should prompt the scientific community to further study the pharmacogenomics of cancer immunotherapy, also in patient series of different ancestries, and its potential for personalized medicine.

Author affiliations

¹Institute for Biomedical Technologies, National Research Council, Segrate, MI, Italy

²Department of Experimental Oncology, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy

³Department of Medical Biotechnology and Translational Medicine, Università degli Studi di Milano, Milan, Italy

⁴Med Biotech Hub and Competence Center, Department of Medical Biotechnologies, University of Siena, Siena, Italy

⁵Cancer Genomics and Systems Biology Lab, Department of Medical Biotechnologies, University of Siena, Siena, Italy

⁶SC Medical Oncology, Fondazione IRCCS Ca' Granda Ospedale Maggiore di Milano Policlinico, Milano, Lombardia, Italy

⁷Medical Oncology Unit, ASST Ovest Milanese, Legnano, Lombardy, Italy

⁸Medical Oncology Unit, Fatebenefratelli Hospital, ASST Fatebenefratelli Sacco, Milan, Italy

⁹Medical Oncology & Hematology Unit, IRCCS Humanitas Research Hospital, Rozzano, MI, Italy

¹⁰SC Medical Oncology, Fondazione IRCCS San Gerardo dei Tintori, Monza, Italy

¹¹Medical Department, Università degli Studi di Milano-Bicocca, Milano, Italy

Acknowledgements The results shown here are in part based upon data generated by the TCGA Research Network: <https://www.cancer.gov/tcga>. The authors acknowledge Elisa Sala for technical support and data management activities, and Loredana Ansalone for administrative support.

Contributors Conceptualization: ME and FC. Formal analysis: FM, ME, and FC. Resources: CB, EC, SP, GF, AP, LT, and DLC. Data curation: SG, LS, AP, SN, EM, ME, and FC. Investigation: FC, SN, and TC. Writing—original draft preparation: FC and ME. Writing—review and editing: ME, FM, EF, MG, DLC, and FC. Funding acquisition: FC. All authors have read and agreed to the published version of the manuscript. Guarantor: FC.

Funding This research was partially funded by Fondazione Martalive Onlus, Brughiero (MI), Italy (FC) and the European Union – Next Generation EU (Missione 4 Componente 1) with Italian Ministry of University and Research (MUR) for “Fondo per il Programma Nazionale di Ricerca e Progetti di Rilevante Interesse Nazionale (PRIN)”- PRIN2022 Life Sciences_L2 Integrative Biology: from Genes and Genomes to Systems – Project ID 2022ARXHR2, CUP B53D23007910006 (FC).

Competing interests No, there are no competing interests.

Patient consent for publication Not applicable.

Ethics approval This study was approved by the ethics committees of the recruiting centers (Comitato Etico Milano Area 1, protocol n. 43745/2018; Comitato Etico Brianza, protocol n. 2870/2019; Comitato Etico Indipendente IRCCS Istituto Clinico Humanitas, protocol n. 95/19; Comitato Etico Milano Area 2, protocol n. 297_2019; Comitato Etico Milano Area 3, protocol n. 538-102019). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. Data are available upon reasonable request. The raw genotyping data presented in this study are available on request from the corresponding author due to privacy restrictions. The GWAS summary statistics are publicly available upon publication in the GWAS catalog (GCST90624815). Data from Ravi *et al* were accessed under dbGap accession phs002822.v1.p1.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Diego Luigi Cortinovis <http://orcid.org/0000-0001-7611-7369>

Francesca Colombo <http://orcid.org/0000-0003-2015-4317>

REFERENCES

- 1 Brahmer J, Reckamp KL, Baas P, *et al*. Nivolumab versus Docetaxel in Advanced Squamous-Cell Non-Small-Cell Lung Cancer. *N Engl J Med* 2015;373:123–35.
- 2 Garon EB, Hellmann MD, Rizvi NA, *et al*. Five-Year Overall Survival for Patients With Advanced Non-Small-Cell Lung Cancer Treated With Pembrolizumab: Results From the Phase I KEYNOTE-001 Study. *J Clin Oncol* 2019;37:2518–27.
- 3 Sezer A, Kilickap S, Gümüş M, *et al*. Cemiplimab monotherapy for first-line treatment of advanced non-small-cell lung cancer with PD-L1 of at least 50%: a multicentre, open-label, global, phase 3, randomised, controlled trial. *The Lancet* 2021;397:592–604.
- 4 Herbst RS, Giaccone G, de Marinis F, *et al*. Atezolizumab for First-Line Treatment of PD-L1-Selected Patients with NSCLC. *N Engl J Med* 2020;383:1328–39.
- 5 Reck M, Ciuleanu T-E, Lee J-S, *et al*. Systemic and Intracranial Outcomes With First-Line Nivolumab Plus Ipilimumab in Patients With Metastatic NSCLC and Baseline Brain Metastases From CheckMate 227 Part 1. *J Thorac Oncol* 2023;18:1055–69.
- 6 Reck M, Rodríguez-Abreu D, Robinson AG, *et al*. Five-Year Outcomes With Pembrolizumab Versus Chemotherapy for Metastatic Non-Small-Cell Lung Cancer With PD-L1 Tumor Proportion Score ≥ 50. *J Clin Oncol* 2021;39:2339–49.
- 7 Herbst RS, Baas P, Kim D-W, *et al*. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. *The Lancet* 2016;387:1540–50.
- 8 de Castro G Jr, Kudaba I, Wu Y-L, *et al*. Five-Year Outcomes With Pembrolizumab Versus Chemotherapy as First-Line Therapy in Patients With Non-Small-Cell Lung Cancer and Programmed Death Ligand-1 Tumor Proportion Score ≥ 1% in the KEYNOTE-042 Study. *J Clin Oncol* 2023;41:1986–91.
- 9 Patel SP, Kurzrock R. PD-L1 Expression as a Predictive Biomarker in Cancer Immunotherapy. *Mol Cancer Ther* 2015;14:847–56.
- 10 Marcus L, Lemery SJ, Keegan P, *et al*. FDA Approval Summary: Pembrolizumab for the Treatment of Microsatellite Instability-High Solid Tumors. *Clin Cancer Res* 2019;25:3753–8.

- 11 Fang W, Ma Y, Yin JC, *et al.* Comprehensive Genomic Profiling Identifies Novel Genetic Predictors of Response to Anti-PD-(L)1 Therapies in Non-Small Cell Lung Cancer. *Clin Cancer Res* 2019;25:5015–26.
- 12 Nomizo T, Ozasa H, Tsuji T, *et al.* Clinical Impact of Single Nucleotide Polymorphism in PD-L1 on Response to Nivolumab for Advanced Non-Small-Cell Lung Cancer Patients. *Sci Rep* 2017;7:45124.
- 13 Queirolo P, Dozin B, Morabito A, *et al.* Association of CTLA-4 Gene Variants with Response to Therapy and Long-term Survival in Metastatic Melanoma Patients Treated with Ipilimumab: An Italian Melanoma Intergroup Study. *Front Immunol* 2017;8:386.
- 14 Polcaro G, Liguori L, Manzo V, *et al.* rs822336 binding to C/EBP β and NFIC modulates induction of PD-L1 expression and predicts anti-PD-1/PD-L1 therapy in advanced NSCLC. *Mol Cancer* 2024;23:63.
- 15 Breunis WB, Tarazona-Santos E, Chen R, *et al.* Influence of cytotoxic T lymphocyte-associated antigen 4 (CTLA4) common polymorphisms on outcome in treatment of melanoma patients with CTLA-4 blockade. *J Immunother* 2008;31:586–90.
- 16 Chowell D, Krishna C, Pierini F, *et al.* Evolutionary divergence of HLA class I genotype impacts efficacy of cancer immunotherapy. *Nat Med* 2019;25:1715–20.
- 17 Naranbhai V, Viard M, Dean M, *et al.* HLA-A*03 and response to immune checkpoint blockade in cancer: an epidemiological biomarker study. *Lancet Oncol* 2022;23:172–84.
- 18 Gallerano D, Ciminati S, Grimaldi A, *et al.* Genetically driven. *Int J Cancer* 2020;147:2597–610.
- 19 Ogimoto T, Ozasa H, Yoshida H, *et al.* CD47 polymorphism for predicting nivolumab benefit in patients with advanced non-small-cell lung cancer. *Oncol Lett* 2023;26:364.
- 20 Middha P, Thummalapalli R, Betti MJ, *et al.* Polygenic risk score for ulcerative colitis predicts immune checkpoint inhibitor-mediated colitis. *Nat Commun* 2024;15:2568.
- 21 Groha S, Alaiwi SA, Xu W, *et al.* Germline variants associated with toxicity to immune checkpoint blockade. *Nat Med* 2022;28:2584–91.
- 22 Pagadala M, Sears TJ, Wu VH, *et al.* Germline modifiers of the tumor immune microenvironment implicate drivers of cancer risk and immunotherapy response. *Nat Commun* 2023;14:2744.
- 23 Sears TJ, Pagadala MS, Castro A, *et al.* Integrated Germline and Somatic Features Reveal Divergent Immune Pathways Driving Response to Immune Checkpoint Blockade. *Cancer Immunol Res* 2024;12:1780–95.
- 24 Eisenhauer EA, Therasse P, Bogaerts J, *et al.* New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 2009;45:228–47.
- 25 Purcell S, Neale B, Todd-Brown K, *et al.* PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics* 2007;81:559–75.
- 26 Anderson CA, Pettersson FH, Clarke GM, *et al.* Data quality control in genetic case-control association studies. *Nat Protoc* 2010;5:1564–73.
- 27 Delaneau O, Marchini J, McVean GA, *et al.* Integrating sequence and array data to create an improved 1000 Genomes Project haplotype reference panel. *Nat Commun* 2014;5:3934.
- 28 Fuchsberger C, Abecasis GR, Hinds DA. minimac2: faster genotype imputation. *Bioinformatics* 2015;31:782–4.
- 29 Loh P-R, Danecek P, Palamara PF, *et al.* Reference-based phasing using the Haplotype Reference Consortium panel. *Nat Genet* 2016;48:11:1443–8.
- 30 Das S, Forer L, Schönerr S, *et al.* Next-generation genotype imputation service and methods. *Nat Genet* 2016;48:1284–7.
- 31 McCarthy S, Das S, Kretzschmar W, *et al.* A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet* 2016;48:1279–83.
- 32 Luo Y, Kanai M, Choi W, *et al.* A high-resolution HLA reference panel capturing global population diversity enables multi-ancestry fine-mapping in HIV host response. *Nat Genet* 2021;53:1504–16.
- 33 Verlouw JAM, Clemens E, de Vries JH, *et al.* A comparison of genotyping arrays. *Eur J Hum Genet* 2021;29:1611–24.
- 34 Ravi A, Hellmann MD, Arniella MB, *et al.* Genomic and transcriptomic analysis of checkpoint blockade response in advanced non-small cell lung cancer. *Nat Genet* 2023;55:807–19.
- 35 Clark TG, Bradburn MJ, Love SB, *et al.* Survival analysis part I: basic concepts and first analyses. *Br J Cancer* 2003;89:232–8.
- 36 Therneau TM, Grambsch PM. Modeling Survival Data: Extending the Cox Model. 2000.
- 37 Aulchenko YS, Ripke S, Isaacs A, *et al.* GenABEL: an R library for genome-wide association analysis. *Bioinformatics* 2007;23:1294–6.
- 38 Oscanoa J, Sivapalan L, Gadaleta E, *et al.* SNPexus: a web server for functional annotation of human genome sequence variation (2020 update). *Nucleic Acids Res* 2020;48:W185–92.
- 39 Watanabe K, Taskesen E, van Bochoven A, *et al.* Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* 2017;8:1826.
- 40 Giambartolomei C, Vukcevic D, Schadt EE, *et al.* Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet* 2014;10:e1004383.
- 41 Chen B, Wang L, Pu S, *et al.* Unveiling potential drug targets for hyperparathyroidism through genetic insights via Mendelian randomization and colocalization analyses. *Sci Rep* 2024;14:6435.
- 42 Györfy B. Transcriptome-level discovery of survival-associated biomarkers and therapy targets in non-small-cell lung cancer. *Br J Pharmacol* 2024;181:362–74.
- 43 The Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. *Nature New Biol* 2012;489:519–25.
- 44 Campbell JD, Alexandrov A, Kim J, *et al.* Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. *Nat Genet* 2016;48:607–16.
- 45 Grossman RL, Heath AP, Ferretti V, *et al.* Toward a Shared Vision for Cancer Genomic Data. *N Engl J Med* 2016;375:1109–12.
- 46 Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:550.
- 47 Shabalin AA. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics* 2012;28:1353–8.
- 48 Garcia M, Juhos S, Larsson M, *et al.* n.d. Sarek: A portable workflow for whole-genome sequencing analysis of germline and somatic variants. *F1000Res* 2019;8:63.
- 49 Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* 2015;31:3555–7.
- 50 Aguiar PN Jr, De Mello RA, Barreto CMN, *et al.* Immune checkpoint inhibitors for advanced non-small cell lung cancer: emerging sequencing for new treatment targets. *ESMO Open* 2017;2:e000200.
- 51 Wu X, Liu W, Liu X, *et al.* Overexpression of MCPH1 inhibits the migration and invasion of lung cancer cells. *Oncotargets Ther* 2018;11:3111–7.
- 52 Minnai F, Noci S, Esposito M, *et al.* Germline Polymorphisms Associated with Overall Survival in Lung Adenocarcinoma: Genome-Wide Analysis. *Cancers (Basel)* 2024;16:3264.
- 53 Jansen R, Hottenga J-J, Nivard MG, *et al.* Conditional eQTL analysis reveals allelic heterogeneity of gene expression. *Hum Mol Genet* 2017;26:1444–51.
- 54 Aiello MM, Solinas C, Santoni M, *et al.* Excision Repair Cross Complementation Group 1 Single Nucleotide Polymorphisms and Nivolumab in Advanced Non-Small Cell Lung Cancer. *Front Oncol* 2020;10:1167.
- 55 Del Re M, Cucchiara F, Rofi E, *et al.* A multiparametric approach to improve the prediction of response to immunotherapy in patients with metastatic NSCLC. *Cancer Immunol Immunother* 2021;70:1667–78.
- 56 Machado-Rugolo J, Gutierrez Prieto T, Fabro AT, *et al.* Relevance of PD-L1 Non-Coding Polymorphisms on the Prognosis of a Genetically Admixed NSCLC Cohort. *Pharmacogenomics Pers Med* 2021;14:239–52.
- 57 Matsumoto A, Nakashima C, Kimura S, *et al.* ALDH2 polymorphism rs671 is a predictor of PD-1/PD-L1 inhibitor efficacy against thoracic malignancies. *BMC Cancer* 2021;21:584.
- 58 Yoshida H, Nomizo T, Ozasa H, *et al.* PD-L1 polymorphisms predict survival outcomes in advanced non-small-cell lung cancer patients treated with PD-1 blockade. *Eur J Cancer* 2021;144:317–25.
- 59 Minari R, Bonatti F, Mazzaschi G, *et al.* PD-L1 SNPs as biomarkers to define benefit in patients with advanced NSCLC treated with immune checkpoint inhibitors. *Tumori* 2022;108:47–55.
- 60 Xin Z, You L, Li J, *et al.* Immunogenetic polymorphisms predict therapeutic efficacy and survival outcomes in tumor patients receiving PD-1/PD-L1 blockade. *Int Immunopharmacol* 2023;121:110469.
- 61 Frąk M, Grenda A, Krawczyk P, *et al.* The influence of nutritional status, lipid profile, leptin concentration and polymorphism of genes encoding leptin and neuropeptide Y on the effectiveness of immunotherapy in advanced NSCLC patients. *BMC Cancer* 2024;24:937.
- 62 Xiao C-L, Liu W-H, Luo Z-Y, *et al.* Blood Group Antigen A Carriers Exhibit an Extended Progression-Free Survival with no more Immune-Related Adverse Events. *Clin Pharmacol Ther* 2024;115:545–55.
- 63 Suminaga K, Nomizo T, Yoshida H, *et al.* The impact of PD-L1 polymorphisms on the efficacy of immune checkpoint inhibitors

- depends on the tumor proportion score: a retrospective study. *J Cancer Res Clin Oncol* 2025;151:61.
- 64 Shahamatdar S, He MX, Reyna MA, *et al.* Germline Features Associated with Immune Infiltration in Solid Tumors. *Cell Rep* 2020;30:2900–8.
- 65 Sayaman RW, Saad M, Thorsson V, *et al.* Germline genetic contribution to the immune landscape of cancer. *Immunity* 2021;54:367–86.
- 66 Yamamoto H, Watanabe Y, Arai H, *et al.* Microsatellite instability: A 2024 update. *Cancer Sci* 2024;115:1738–48.
- 67 Zhao P, Li L, Jiang X, *et al.* Mismatch repair deficiency/microsatellite instability-high as a predictor for anti-PD-1/PD-L1 immunotherapy efficacy. *J Hematol Oncol* 2019;12:54.
- 68 Lemery S, Keegan P, Pazdur R. First FDA Approval Agnostic of Cancer Site - When a Biomarker Defines the Indication. *N Engl J Med* 2017;377:1409–12.
- 69 Zheng A, Bai J, Ha Y, *et al.* Integrated analysis of the relation to tumor immune microenvironment and predicted value of Stonin1 gene for immune checkpoint blockage and targeted treatment in kidney renal clear cell carcinoma. *BMC Cancer* 2023;23:135.
- 70 Kasuga Y, Ouda R, Watanabe M, *et al.* FBXO11 constitutes a major negative regulator of MHC class II through ubiquitin-dependent proteasomal degradation of CIITA. *Proc Natl Acad Sci USA* 2023;120.
- 71 Schuhmacher JS, Tom Dieck S, Christoforidis S, *et al.* The Rab5 effector FERRY links early endosomes with mRNA localization. *Mol Cell* 2023;83:1839–55.
- 72 Ye Z, Xiong Y, Peng W, *et al.* Manipulation of PD-L1 Endosomal Trafficking Promotes Anticancer Immunity. *Adv Sci (Weinh)* 2023;10:e2206411.