Beyond tobacco: genomic disparities in lung cancer between smokers and never-smokers

Javiera Garrido[1], Yanara Bernal[1], Evelin González[1], Alejandro Blanco[1], Gonzalo Sepúlveda-Hermosilla[1,6], Matías Freire[6], Karen Oróstica[19], Solange Rivas[1], Katherine Marcelain[2,3], Gareth Owen[4,21], Carolina Ibañez[4], Alejandro Corvalan[4], Marcelo Garrido[5], Rodrigo Assar[6,7], Rodrigo Lizana[6], Javier Cáceres-Molina[6], Diego Ampuero[6], Liliana Ramos[6], Paola Pérez[8], Osvaldo Aren[9], Sara Chernilo[10], Cristina Fernández[10], María Loreto Spencer[11], Jacqueline Flores Aguila[12], Giuliano Bernal Dossetto[13], Mónica Ahumada Olea[2,3,14], Germán Rasse[15], Carolina Sánchez[16], Maria Galli de Amorim[17], Thais F. Bartelli[17], Diana Noronha Nunes[17], Emmanuel Dias-Neto[17,20], Helano C. Freitas[17,18] and Ricardo Armisén[1*]

Abstract

Background Tobacco use is one of the main risk factors for Lung Cancer (LC) development. However, about 10–20% of those diagnosed with the disease are never-smokers. For Non-Small Cell Lung Cancer (NSCLC) there are clear differences in both the clinical presentation and the tumor genomic profiles between smokers and never-smokers. For example, the Lung Adenocarcinoma (LUAD) histological subtype in never-smokers is predominately found in young women of European, North American, and Asian descent. While the clinical presentation and tumor genomic profiles of smokers have been widely examined, never-smokers are usually underrepresented, especially those of a Latin American (LA) background. In this work, we characterize, for the first time, the difference in the genomic profiles between smokers and never-smokers LC patients from Chile.

Methods We conduct a comparison by smoking status in the frequencies of genomic alterations (GAs) including somatic mutations and structural variants (fusions) in a total of 10 clinically relevant genes, including the eight most common actionable genes for LC (EGFR, KRAS, ALK, MET, BRAF, RET, ERBB2, and ROS1) and two established driver genes for malignancies other than LC (PIK3CA and MAP2K1). Study participants were grouped as either smokers (current and former, $n = 473$) or never-smokers ($n = 200$) according to self-report tobacco use at enrollment.

Results Our findings indicate a higher overall GA frequency for never-smokers compared to smokers (58 vs. 45.7, p-value < 0.01) with the genes EGFR, KRAS, and PIK3CA displaying the highest prevalence while ERBB2, RET, and ROS1 the lowest. Never-smokers present higher frequencies in seven out of the 10 genes; however, smokers harbor a more complex genomic profile. The clearest differences between groups are seen for EGFR (15.6 vs. 21.5, p-value: < 0.01), PIK3CA (6.8 vs 9.5) and ALK (3.2 vs 7.5) in favor of never-smokers, and KRAS (16.3 vs. 11.5) and MAP2K1 (6.6 vs. 3.5) in favor of smokers. Alterations in these genes are comprised almost exclusively by somatic mutations in EGFR and mainly by fusions in ALK, and only by mutations in PIK3CA, KRAS and MAP2K1.

Conclusions We found clear differences in the genomic landscape by smoking status in LUAD patients from Chile, with potential implications for clinical management in these limited-resource settings.

---

*Correspondence:
Ricardo Armisén
rarmisen@udd.cl
Full list of author information is available at the end of the article

Garrido *et al. BMC Cancer*   (2024) 24:951

Page 2 of 15

Introduction

Lung cancer (LC) is the leading cause of cancer-related death worldwide and the third cause in the Latin American (LA) and the Caribbean populations [1]. About 85% of LCs are Non-Small Cell Lung Cancer (NSCLC), and Lung Adenocarcinoma (LUAD) is the most frequent histological subtype with approximate 40% of cases, followed by Squamous Cell Carcinoma (SqCC, 30%), and large cell carcinoma (15%) [2]. More than 20 agents related to environmental and occupational exposure have been identified as lung carcinogens [3]; however, tobacco use is the leading cause of LC, followed by second-hand smoke. Nevertheless, 10–20% of LC diagnoses are in people who have never smoked [4]. LC in never-smokers (LCINS) has been reported as the 11th most frequent cause of mortality in men, and 8th in women [5]. Second-hand smoke and environmental and occupational exposure partially explain some LC diagnoses in never-smokers [6]. Most of the efforts to understand this condition have been conducted in European, North American, and Asian individuals; however, the LCINS from LA populations are understudied.

Studies on non-LA populations suggest that LCINS has different clinical and genomic characteristics than in smokers [7]. In never-smokers, the diagnosis is predominately of the LUAD subtype and occurs more frequently in Asian and Hispanic young women with advanced disease. Surprisingly, Hispanic never-smokers with LC have shown a poorer survival outcome than non-Hispanic [8]. On the other hand, efforts have been made to understand the tumor profiles in never-smokers, particularly in those of European, North American, and Asian descent. From 63,000 LC patients from 167 different studies, it was shown that never-smokers had a higher prevalence of mutations in *EGFR* and fusion rearrangements in *ALK* while a lower prevalence of mutations on *KRAS*, with some differences between Caucasian and Asian individuals [9]. From a French cohort of 17,664 LC patients, *PIK3CA* and *BRAF* showed a higher mutation frequency in never-smokers compared to smokers [10]. The Belgian FIELTS-2 study found that the frequency of mutations in *ERBB2* and amplifications on *MET* was higher in never-smokers than in smokers [7]. Studies in the Asian population demonstrated that fusion rearrangements in *ROS1* were higher in never-smokers than smokers [11]. In addition, the *RET* proto-oncogene is more prevalent mutated in the LUAD subtype, mostly in the group of young never-smokers [12]. Finally, the frequency of mutation in *MAP2K1* was higher in smokers than in never-smokers [13].

In Chile, recent evidence revealed that although smoking consumption have decreased in the last years, its effect on mortality has not been mirrored, possibly due to the existence of additional risk factors [14]. In addition, the prevalence of LCINS has increased over the years and it is believed that environmental exposure could be related to this trend [15, 16]. On the other hand, tumor genomic profiles of LC patients from Chile and other LA countries are poorly characterized, particularly in never-smokers. A better understanding of the genomic profiles of this underrepresented population could not only provide mechanistic insight into lung carcinogenesis not mediated by tobacco smoke, but also could improve the clinical management of these group of patients with more personalized therapy strategies.

In this study, we perform a comparative-descriptive analysis of the genomic profiles of Chilean LC patients according to smoking status and other variables of clinical relevance (including sex, age, NSCLC subtype, cancer stage, Personal History of Cancer [PHC] and Family History of Cancer [FHC]). We focus our analyses on 10 different genes, including the eight most common actionable genes for LC (*EGFR*, *KRAS*, *ALK*, *MET*, *BRAF*, *RET*, *ERBB2*, and *ROS1*) in addition to two driver genes of high mutation prevalence and established carcinogenic potential (*PIK3CA* and *MAP2K1*). We calculate the prevalence of genomic alterations (GAs) for each gene, including somatic mutations and structural variants (fusions), for the complete set of participants (overall prevalence) and for the group of smokers and never-smokers separately (and for the other relevant clinical variables). Furthermore, we attempt to unravel possible patterns of co-occurrences and exclusions in the genomic landscape of smokers and never-smokers. Finally, a detailed description of the complete set of GAs found in the population is provided.

Materials and methods

Study participants

The population is derived from the study protocol Characterization and Validation of Molecular Diagnostic Technologies for LC Patients from Chile, Brazil, and Peru (registered at clinicaltrials.gov as NCT03220230). The recruitment period was between July 2015 to October 2018 and encompassed 37 health centers from these three countries. A complete description on the study

Garrido *et al. BMC Cancer*   (2024) 24:951

Page 3 of 15

protocol has been provided previously [46, 47]. Of the 5030 recruited participants, genomics profiles and valid clinical information were successfully acquired for 1864 individuals. After the exclusion of patients with missing data in the relevant clinical variables (sex, age, NSCLC subtype, cancer stage, PHC and FHC), a total of 673 Chilean patients were included as part of the study population of this current study. Subjects were grouped as either smokers (current and former) or never-smokers according to self-report tobacco use at enrollment. More details on the final number of study participants are presented in Additional File 1: Fig. S1.

Sequencing quality control and variant classification

From tumor samples, sections with at least 5% tumor tissue were included. We selected up to 8 FFPE sections of 5 μm, the Recover All extraction kit (Thermo Fisher Scientific) was used for the isolation of RNA and DNA. The Oncomine Focus Assay (OFA, Thermo Fisher Scientific) was employed to prepare libraries, and sequenced in the Ion Personal Genome Machine System. OFA is a Next-Generation Sequencing (NGS) panel aimed at discovering Single Nucleotide Variants (SNVs), Indels, Copy Number Variations (CNVs) and gene fusions. The QC metric thresholds were at least 240 median reads per amplicon and 60% of aligned reads for DNA libraries, and 20,000 correctly mapped reads plus three out of five expression control amplicons detected for RNA.

For the alignment and variant calling, strict parameters were defined to call SNVs and Indels: minimum allele frequency of 5% (SNVs) and 7% (Indels), the minimum coverage that admits a variant was 10x (SNVs and Indels). In addition, the minimum coverage of the variant location is 50x, with minimum variant scores in phred-scaled values set at 6 for SNVs and 20 for Indels. Defined parameters of 70% overlap reading alignment with reference and 66% exact matches were used for the alignment of the fusions, and a minimum valid mapped reading of 20× and 15× for fusions and expression controls, respectively. All remaining reference/reference sites, variants with allelic frequency < 5%, and observed alternative alleles < 10 reads were removed from the DNA Variant Call Format (VCF) files. For RNA VCFs, only fusions with more than 20 reads were maintained. Oncomine variants were selected, defined as those located in positions within the predefined hotspots of Oncomine Focus DNA Hotspots v1.4.

Statistical analyses

Overall prevalence of GA per gene was calculated by computing the proportion between the number of patients with at least one GA as the numerator and the total number of study participants as the denominator. Stratified prevalence by smoking status and relevant clinical variables for each gene was calculated by computing the proportion between the number of patients with at least one GA in each group as the numerator and the total number of study participants in each group as the denominator. Statistical significance for difference in proportions between groups was calculated using either the chi-squared test or the fisher exact test for small sample sizes or low expected frequencies. Statistical significance level was set to 5% and 1% and reported appropriately. In addition, we accounted for multiple testing using a Bonferroni correction, controlling the Family Wise Error Rate (FWER) below 1%.

Kendall correlation estimates were computed to the obtain matrices of GA co-occurrences and exclusions and the significance level was set at 5%. We used R programming language version 4.3.1 for all statistical analyses.

Results

Study population

A description of the study population by smoking status concerning the main clinical variables is depicted in Table 1. We observe that the never-smoker group

Table 1 Characteristics of the study population by smoking status with respect to main clinical variables ($n = 673$)

	Smoker	Never-Smoker	*P*-value
	($n = 473$)	($n = 200$)	
Sex			
Male	269 (56.9%)	85 (42.5%)	< 0.001
Female	204 (43.1%)	115 (57.5%)	
Age			
55 or less	56 (11.8%)	41 (20.5%)	0.00506
More than 55	417 (88.2%)	159 (79.5%)	
NSCLC Type			
LUAD	413 (87.3%)	148 (74.0%)	< 0.001
SqCC	52 (11.0%)	48 (24.0%)	
Other	8 (1.69%)	4 (2.00%)	
Cancer Stage			
Stage I	25 (5.29%)	8 (4.00%)	0.00187
Stage II	19 (4.02%)	5 (2.50%)	
Stage III	72 (15.2%)	11 (5.50%)	
Stage IV	357 (75.5%)	176 (88.0%)	
PHC			
No	433 (91.5%)	192 (96.0%)	0.0589
Yes	40 (8.46%)	8 (4.00%)	
FHC			
No	346 (73.2%)	171 (85.5%)	< 0.001
Yes	127 (26.8%)	29 (14.5%)	

Counts and percentages are reported, *p*-value for difference in proportions was calculated using the chi-squared test

LUAD Lung Adenocarcinoma, *SqCC* Squamous Cell Carcinoma, *PHC* Personal History of Cancer, *FHC* Family History of Cancer

Garrido *et al. BMC Cancer*   (2024) 24:951

Page 4 of 15

is predominately formed by females (57.5%), and the smoker by males (56.9%). In addition, smokers are characterized by a higher proportion of older participants, LUAD subtype, and self-reported PHC and FHC, while never-smokers by higher frequencies of SqCC subtype, and advanced disease at diagnosis. We found statistically significant results for differences in proportions between groups at 0.1% level for sex, NSCLC type and FHC.

Genomic profile landscape

The co-oncoprint plot by smoking status for the complete set of participants is displayed in Fig. 1. Broadly, we observe that around half of the patients in each group present at least one GA, with most of the tumor samples harboring only one. The highest number of GA in a single tissue sample among smokers is seven ($n=1$), and five among never-smokers ($n=1$). In that same line, samples containing more than one GA are more common in smokers, as co-occurrence events in never-smokers are scarce. The two most altered genes, EGFR and KRAS, display a clear exclusion pattern as tumors with alterations in EGFR do not contain alterations in KRAS, and vice versa. EGFR events co-occur mainly with PIK3CA, MAP2K1, ALK and MET predominately in smokers. Somatic missense mutations (MM) are the most common type of GA and are mainly observed in EGFR, KRAS, PIK3CA and MAP2K1. Other types of somatic mutations such as In−Frame Deletions (InF Del) and In−Frame Insertions (InF Ins) are infrequent, typically present in the EGFR and ERBB2 genes. Structural variants are mostly present in ALK, MET, RET and ROS1. On the other hand, multi-hit events (i.e., two or more GA co-occurring in the same gene) are also infrequent and mainly observed in EGFR and MET. In relation to clinical variables, we can observe a distinguishable pattern for NSCLC subtype, cancer stage, PHC and FHC. Regarding disease subtype, altered samples of smokers seem to have a higher proportion of the LUAD subtype than those of never-smokers; conversely, there are appeared to be a similar proportion of the SqCC subtype in non-altered samples in both smokers and never-smokers. Concerning stage, unaltered samples of never-smokers seem to have
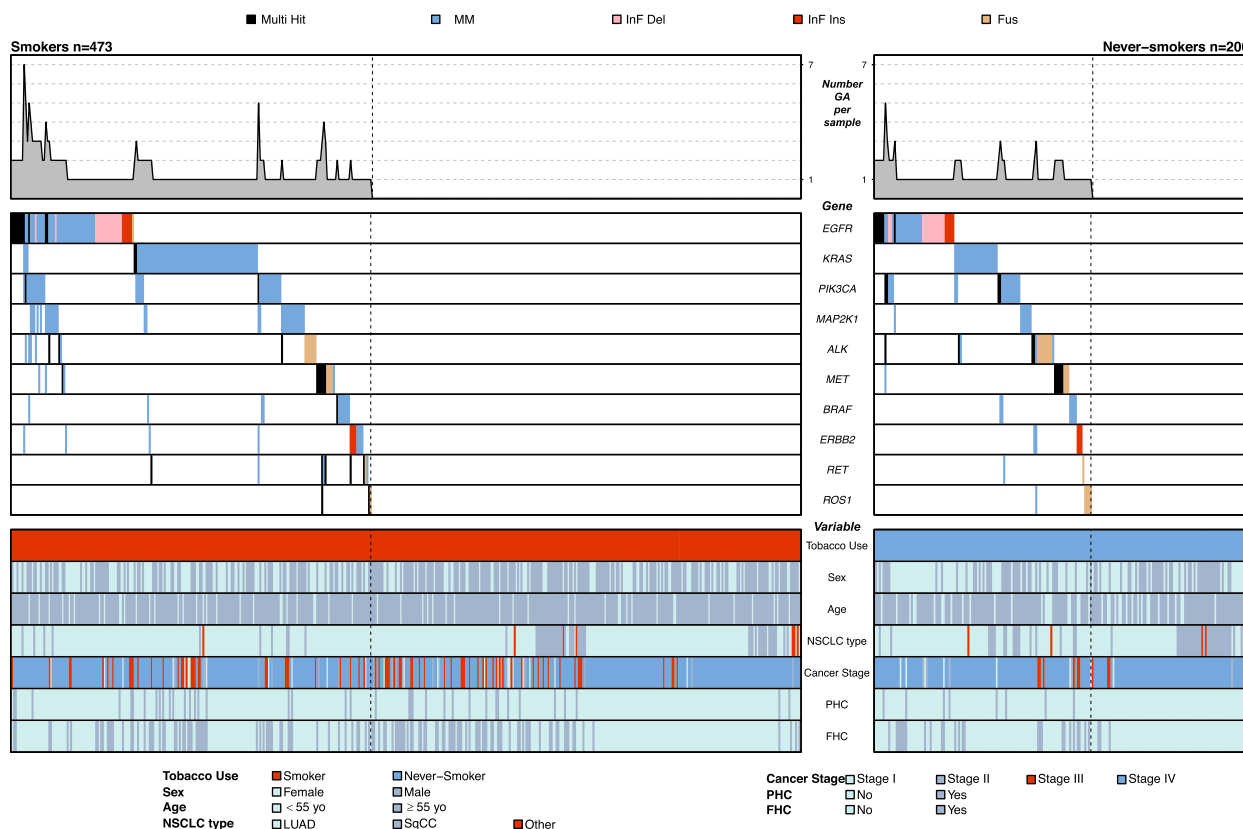
Fig. 1 Genomic landscape of the study population by smoking status. Left panel figures represent smokers ($n=473$) and right panel figures never-smokers ($n=200$). Top panels display the absolute number of GA per tumor sample. Middle panels are the oncoprint plots for each group of participants. Bottom panels indicate the characteristics of the patients for the studied clinical variables. Dashed vertical lines separate the set of samples with at least one GA from the those without GAs. GA: Genomic Alteration, MM: Missense Mutation, InF Del: In-Frame Deletion, InF Ins: In-Frame Insertion, Fus: Fusion, LUAD: Lung Adenocarcinoma, SqCC: Squamous Cell Carcinoma

Garrido *et al. BMC Cancer*   (2024) 24:951

Page 5 of 15

a higher proportion of advanced disease patients. As to PHC and FHC, a negative self-report in both conditions is proportionately more frequent in unaltered samples of never-smokers.

Prevalence of genomic alterations (GAs)

The total number of participants harboring at least one GA in any of the 10 driver genes under study is 332 (49,3%). As depicted in Fig. 2, EGFR, KRAS and PIK3CA are among the most altered genes with an overall prevalence of 17.4%, 14.9% and 7.6%, respectively; ERBB2, RET and ROS1 are the least altered ones with respective overall prevalence of 2.5%, 1.6% and 1.2%. Stratified analyses by smoking status reveal a higher proportion of GA among never-smokers than smokers (n = 116, 58% vs. n = 216, 45.7%, *p*-value < 0.01). Smokers present a higher prevalence for three out of the 10 genes, namely KRAS (16.8% vs 11.5%), MAP2K1 (6.5% vs. 3.5%), and RET (1.9% vs. 1%). Never-smokers present a higher prevalence for the remaining seven genes, with the most noticeable

differences in EGFR (21.5% vs. 15.6%), PIK3CA (6.7% vs 9.5%), and ALK (3.2% vs 7.5%). Statistical significance difference between groups was only reached for EGFR (*p*-value < 0.01).

When estimating the prevalence of somatic mutations and structural variants separately, frequencies remain unchanged for KRAS, PIK3CA, MAP2K1, BRAF and ERRB2 as these genes only display mutations. In contrast, ALK, MET, RET and ROS1 present a prevalence of fusions that is higher than the respective prevalence of mutations; the most notable difference is observed for ALK while MET shows similar prevalence values for the two types of GAs.

Stratified prevalence analyses by relevant clinical variables of interest reveal a statistically significant difference for EGFR between categories in all variables, favoring a higher proportion of GA for younger female patients with stage I LUAD subtype, and positive PHC and FHC. Additional statistical significance differences were also observed for RET and ROS1 in the NSCLC subtype
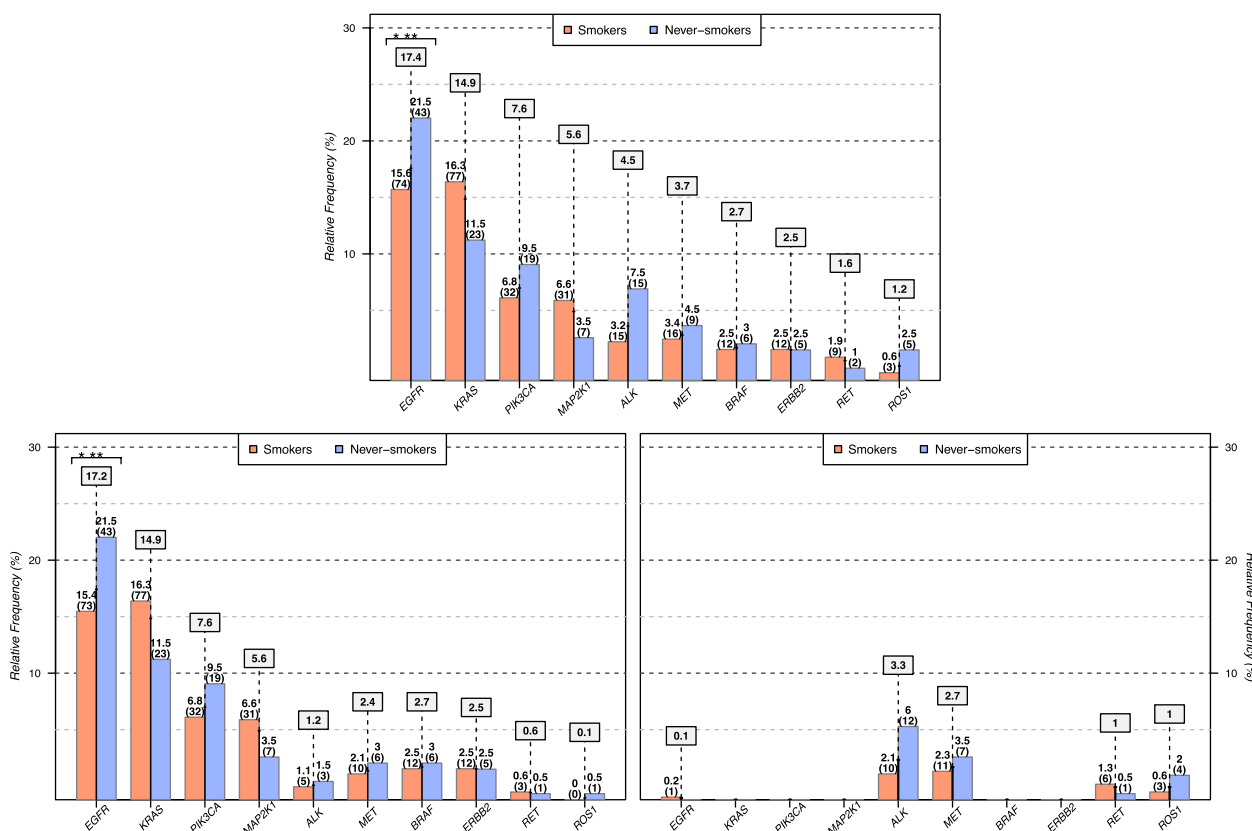
Fig. 2 Prevalence of GAs for the 10 driver genes under study by smoking status. Genes are ordered by decreasing overall prevalence. Absolute and relative frequencies are specified for each group at the top of each bar. Prevalence values for the complete population are shown in grey rectangles. *p*-value for difference in proportions between groups was calculated using either the chi-squared test or the fisher exact test for small expected counts. Top panel displays the overall prevalence for smokers and never-smokers including both somatic mutations and structural variants. Bottom panels display relative frequencies separately for somatic mutations (left) and structural variants (right). (*) represents statistically significant difference. (*: *p*-value < 0.05, **: *p*-value < 0.01, ***: Bonferroni adjusted *p*-value < 0.001). GA: Genomic Alteration

Garrido *et al. BMC Cancer*   (2024) 24:951

Page 6 of 15

(favoring higher frequencies for LUAD subtype); BRAF, ERRB2 and ROS1 in cancer stage (favoring higher frequencies for early-stage disease); and for MAP2K1 in PHC (favoring higher frequencies for a positive history). (Fig. 2 and Additional File 1: Fig. S2).

When treating age as a continuous variable, we found statistically significant findings for overall prevalence for *KRAS* in the age value 63 y/o (prevalence for smokers = 0% [n smokers = 12] vs prevalence for never-smokers = 50% [n never-smokers = 4], *p*-value < 0.05), and *MET* in the age value of 64 y/o (prevalence for smokers = 0% [n smokers = 20] vs prevalence for never-smokers = 100% [n never-smokers = 1], *p*-value < 0.05). When categorizing age by 20th percentiles (20th: 58 y/o, 40th; 65y/o, 60th:70 y/o, and 80th: 75 y/o), we found statistically significant findings for overall prevalence for *KRAS* in the age range ≤ 58 y/o and 70–75 y/o, *MAP2K1* in the age range ≤ 58 y/o, and *ROS1* in the age range 58–65 y/o (*p*-value < 0.05) (Additional File 1: Fig. S3).

Distribution and type of GAs per gene by smoking status
Most patients only harbor one GA in their genomic profiles. More precisely, out of participants with altered

profiles, 73.1% (*n* = 158) of smokers and 76.7% (*n* = 89) of never-smokers present samples with one GA. As depicted in Fig. 3 (top), the prevalence of the genes MAP2K1, ALK, ERBB2, RET and ROS1 are exclusively based on patients with one GA for both smokers and never-smokers. In the case of KRAS and BRAF, one-GA samples are observed solely for never-smokers, while smokers present a comparatively low proportion of samples with more than one GA. On the other hand, PIK3CA and MET display a similar distribution of one- vs. more-than-one GA samples in the two groups, with a higher proportion of the later in the never-smoker group. Notably, a similar distribution between groups is observed for EGFR. Analyses for difference in proportion between smokers and never-smokers did not detect statistically significant findings.

The type of GA per gene by smoking status for samples harboring one GA and more than one GA is depicted in Fig. 3 (bottom left and bottom right, respectively). The prevalence estimates for one-GA samples in the genes KRAS, PIK3CA, MAP2K1, and BRAF are exclusively comprised by somatic mutations of the missense type. Fusions are almost exclusively found in ALK, MET, RET
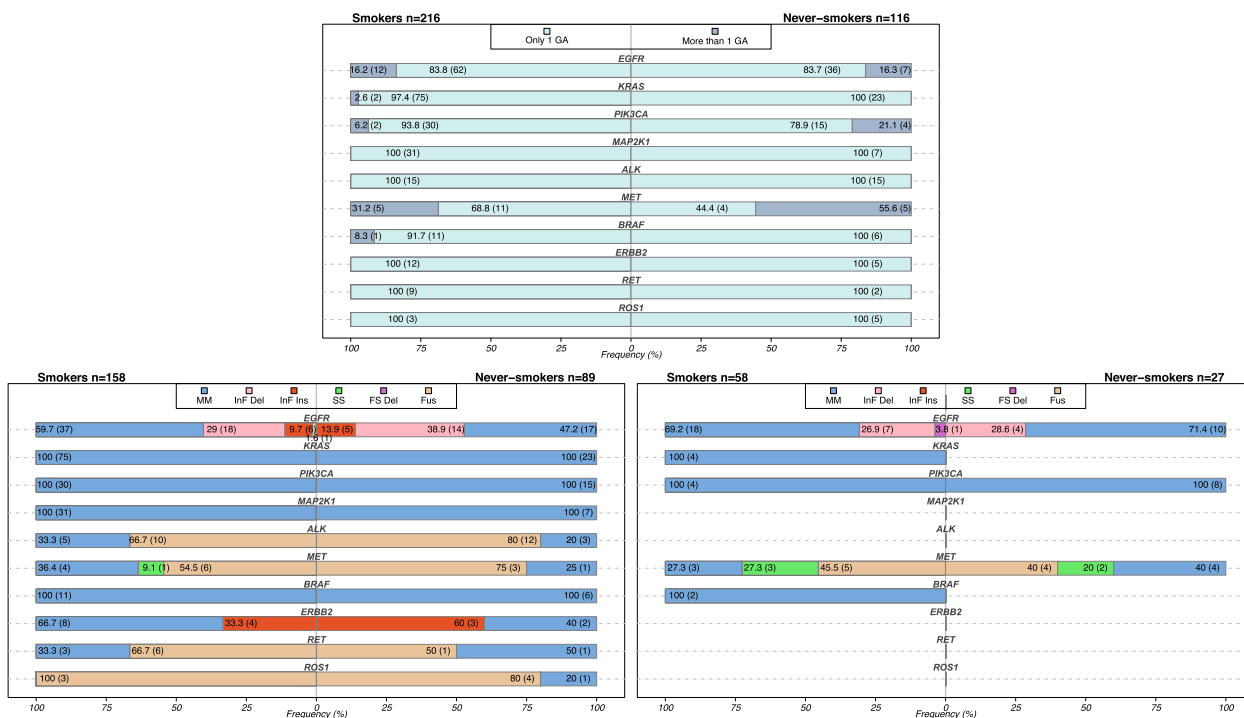
Fig. 3 Distribution and type of GAs per gene. For each plot, smokers are shown in the left and never-smokers in the right. Top panel includes all tumor samples with at least one GA and the distribution of samples with only one or more than one GA for each gene is shown. Bottom left panel includes samples with only one GA and the proportion of the different types of GA for each gene is shown. Bottom right panel includes samples with more than one GA and the proportion of the different types of GA for each gene is shown. In each panel, the number of tumor samples included in the analysis for each group of patients is specified at the top. Absolute and relative frequencies for the distribution and type GA are specified inside bars. GA: Genomic Alteration

and ROS1, and mostly co-occur with MMs. Of those four genes, never-smokers present a slightly higher proportion of fusions in ALK and in MET than smokers. InF Ins events occur only in EGFR and ERRB2 with a higher proportion in never-smokers. InF Del occur only in EGFR in a similar ratio between groups. More infrequent events include splice sites (SS) in MET and fusions in EGFR, both present in smokers only. On the other hand, the prevalence estimates for samples with more than one GA are exclusively based on MM for KRAS, PIK3CA and BRAF. Co-occurrence events with different types of GAs are only seen for EGFR and MET in a similar proportion between smokers and never-smokers. Statistically significant difference in the proportions of GA types between groups did not reveal significant findings.

Correlation patterns between GAs

The matrices of co-occurrences and exclusions of GA grouped by genes and stratified by smoking status are shown in Fig. 4. Overall, we observe that in both smokers and never-smokers most pair-wise correlation coefficients are negative (exclusions) and non-significant. On the contrary, most statistically significant correlations are positive (co-occurrences): 11 out of 12 of the significant coefficients in smokers and three out of four in never-smokers; in the former group these are mainly driven by EGFR, PIK3CA, ALK and MET. The only negative coefficient reaching statistical significance in the two groups is

between EGFR and KRAS (smokers: $R = -0.14$ vs. never-smokers: $R = -0.19$, p-value: < 0.05). On the other hand, the only positive and statistically significant coefficient in the two groups is between PIK3CA and BRAF (smokers: $R = 0.12$ vs. never-smokers: $R = 0.14$, p-value: < 0.05).

When analyzing the co-occurrence and exclusion patterns at the level of individual GA, we observe a much sparser matrix as most coefficients are weak and non-significant, particularly for smokers. In both groups, a minority of pair-wise correlations reach statistically significant coefficients of one (34 out of 110 [25.9%] in smokers and 12 out of 74 [15.4%] in never-smokers). Among those, the following pair-wise correlations can be highlighted: p.A1200V (ALK)/ p.R841K (EGFR), p.L1204F (ALK)/ p.L798F (EGFR), p.G776D (ERBB2) / p.G128S (MAP2K1) in smokers, and p.D1045N (PIK3CA)/ p.G863S (EGFR) in never-smokers (Fig. 5 and Additional File 1: Fig. S4). The complete set of co-occurrences in each group of patients is listed in Additional File 2: Tables S1-S2.

Characterization of individual GA

The total number of unique GAs identified in our study population is 152, of which 78 (51.3%) are solely found in smokers, 42 (27.4%) solely in never-smokers and 32 (21.1%) are common between the two groups. EGFR and PI3KCA are the genes with the highest numbers of unique GAs ($n = 48$ and $n = 33$, respectively), followed
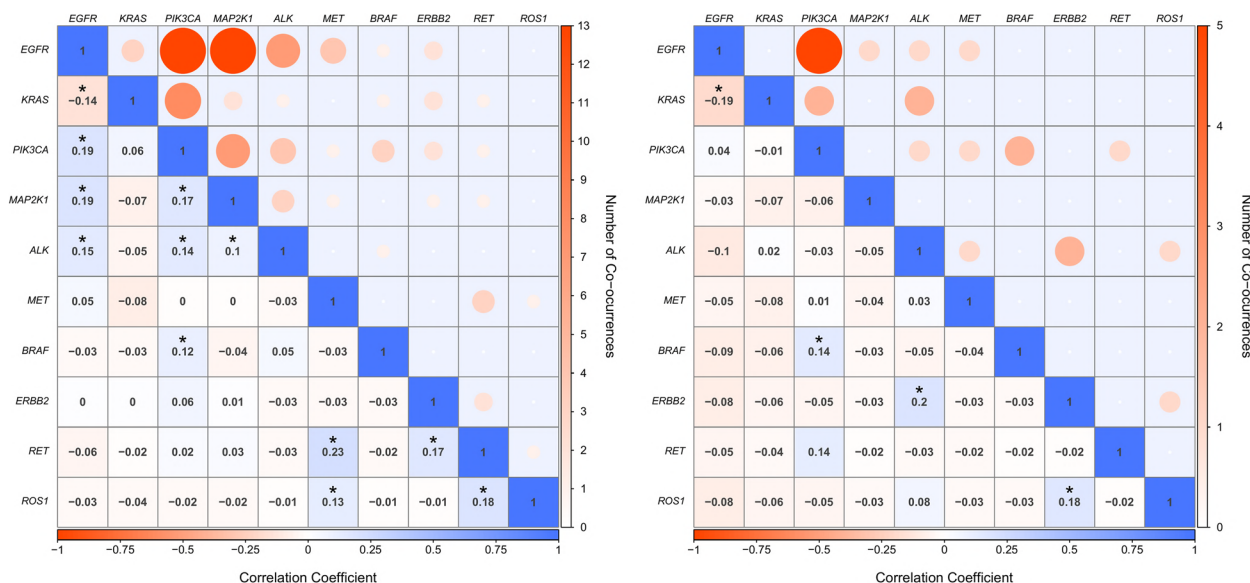
Fig. 4 Matrices of co-occurrences and exclusions of GA grouped by genes. The group of smokers are shown in the left and never-smokers in the right. Lower triangles of the matrices represent Kendall correlation coefficients, with negative correlation (exclusion) coloured in red and positive correlation (co-occurrence) in blue. (*) represents statistically significant correlations at a 5% level. Upper triangles of the matrices represent the actual number of co-occurrences, with bigger and redder circles representing higher absolute number of co-occurrences. GA: Genomic Alteration

Garrido *et al. BMC Cancer*   (2024) 24:951

Page 8 of 15
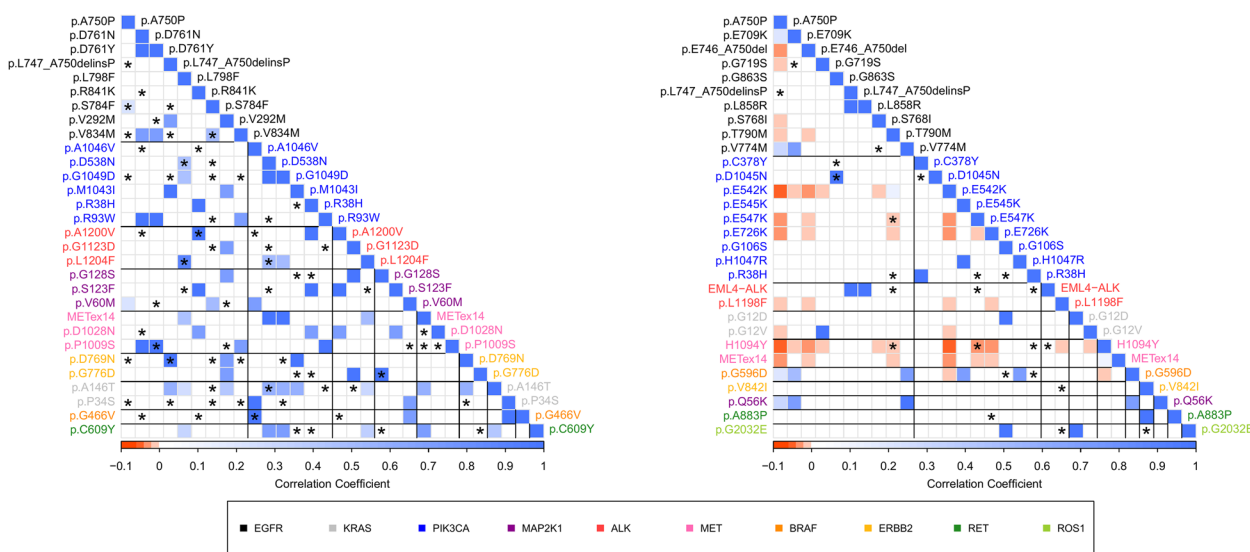
Fig. 5 Matrices of co-occurrences and exclusions of the top 30 GA with the highest coefficients. The group of smokers are shown in the left and never-smokers in the right. Matrices display the first 30 GAs with the highest sum of the absolute values of all pair-wise coefficients. GAs are coloured and ordered by the genes to which they belong. Negative correlations (exclusions) are coloured in red and positive correlation (co-occurrences) in blue. (*) represents statistically significant correlations at a 5% level. GA: Genomic Alteration

by KRAS and BRAF ($n=12$), MET and ERBB2 ($n=9$), MAP2K1 and ALK ($n=8$), RET ($n=7$) and ROS1 ($n=6$). In the case of EGFR, most of these GA are located in exons 19 ($n=19$, 36.9%) and 20 ($n=13$, 27.1%), where the tyrosine-kinase domain is located; for PIK3CA in exons 10 ($n=8$, 24.2%), 2 and 21 ($n=7$, 21.2%), and in the PKc MEK1 domain. Details on the unique set of GAs for these two genes are provided in Figs. 6 and 7 (for the remaining genes see Additional File 1: Fig. S4 and Additional File 2: Tables S3-S12). Common actionable mutations in EGFR including deletions in exon 19 and point mutation p.L858R reach overall relative frequencies of 6.5% and 5%, respectively. For other common actionable GAs these values are: 5.6% and 2.8% for KRAS p.G12C and p.G12D, respectively; 2.7% for METex14; and 0.9% for BRAF p.V600E. For smokers, individual GAs with the highest frequencies are p.G12C with 6.6%, and MAP2K1 p.Q56K with 5.3%; for never-smokers, p.L858R and p.G12C with 7% and 3.5%, respectively. When calculating differences in relative frequencies at the level of individual GA for each gene by smoking status, our analyses did not find statistically significant results (see also Additional File 2: Tables S3-S12). In relation to the common GAs between groups, 30 out of 32 are somatic mutations and only two are structural variants. These common alterations are present in 264 participants (smokers: $n=173$, never-smokers: $n=91$). Most of the shared mutations are found in EGFR ($n=10$), KRAS ($n=6$), PIK3CA ($n=5$) and MET ($n=5$), while shared fusions are found in ALK and MET ($n=1$). Analyses for differences in proportion between groups did not find statistically significant results for any of the common GAs (Table 2).

Discussion

Our main findings reveal that the genes with the highest overall genomic prevalence are *EGFR*, *KRAS* and *PIK3CA* while *ERBB2*, *RET* and *ROS1* present the lowest. Compared to smokers, never-smokers harbor at least a single alteration in their tumor samples more frequently (58 vs. 45.7, p-value < 0.01), with higher genomic prevalence in seven out the 10 genes under study. The clearest differences in favor of never-smokers are observed for *EGFR* (15.6 vs. 21.5, p-value: < 0.01), *PIK3CA* (6.8 vs 9.5) and *ALK* (3.2 vs 7.5). The clearest differences in favor of smokers are seen for *KRAS* (16.3 vs. 11.5) and *MAP2K1* (6.6 vs. 3.5). Despite the lower prevalence, the group of smokers harbor a more complex genomic profile as i) there is a higher proportion of samples with more than one alteration (26.9 vs. 23.3), ii) co-occurrence events are more common (at the level individual GAs and grouped by genes), and iii) isolated events such as fusions and frame shift deletions in *EGFR*, and SSs in *MET* are present in smokers only. On the other hand, our analyses also reveal common features between smokers and never-smokers: i) the distribution and type of alterations across genes is similar, and ii) a clear exclusion pattern between *EGFR* and *KRAS* events is present in the genomic profiles of the two groups. With respect to the relevant clinical variables, we can draw the following conclusions: i) never-smokers are more likely to be younger women
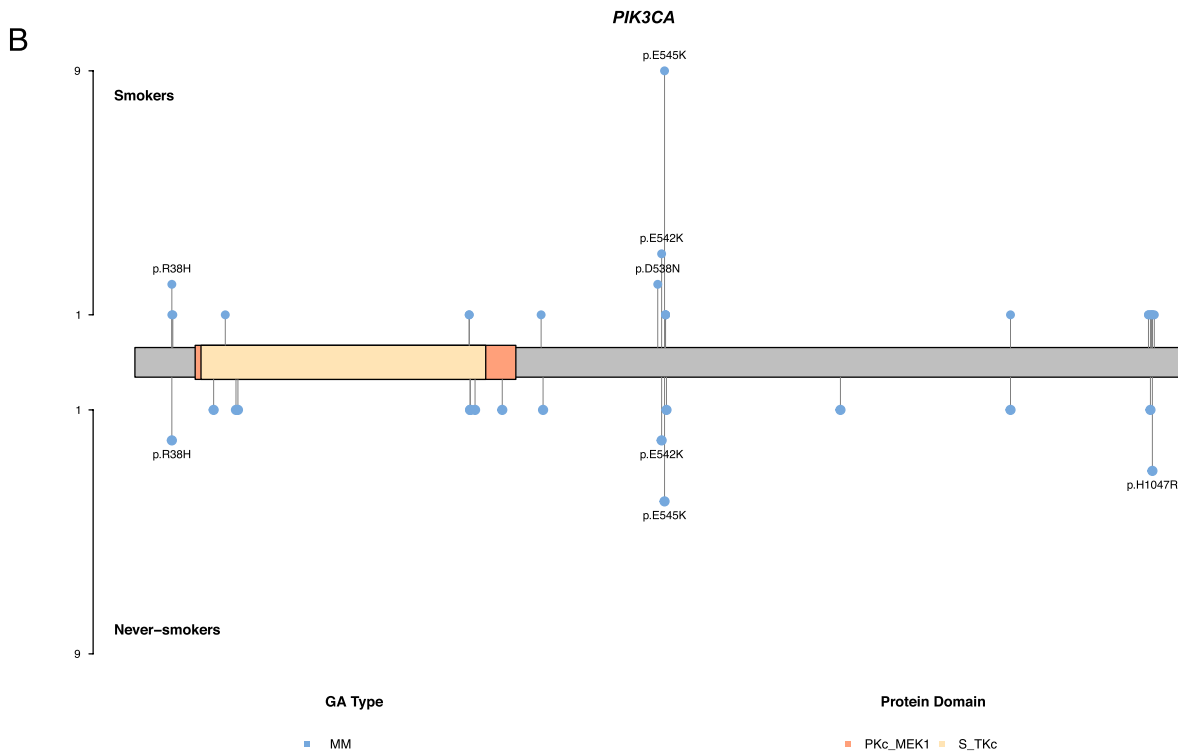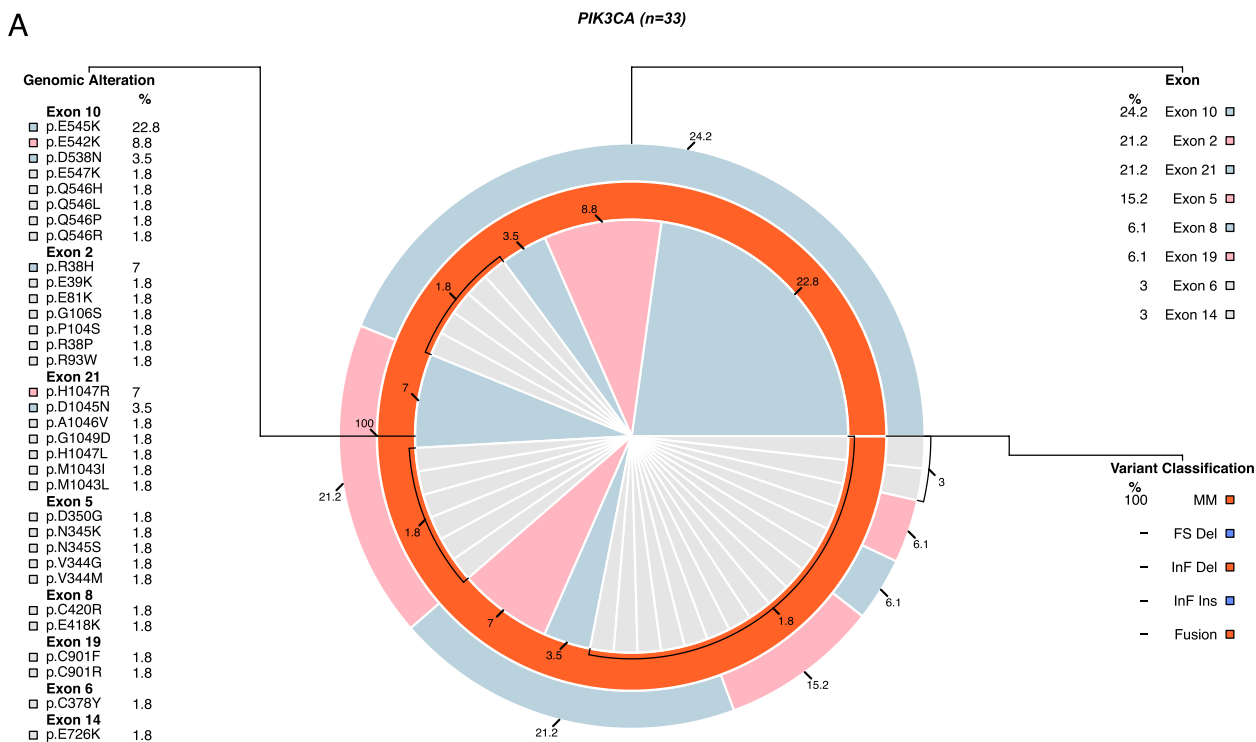
Garrido *et al. BMC Cancer*   (2024) 24:951

Page 9 of 15
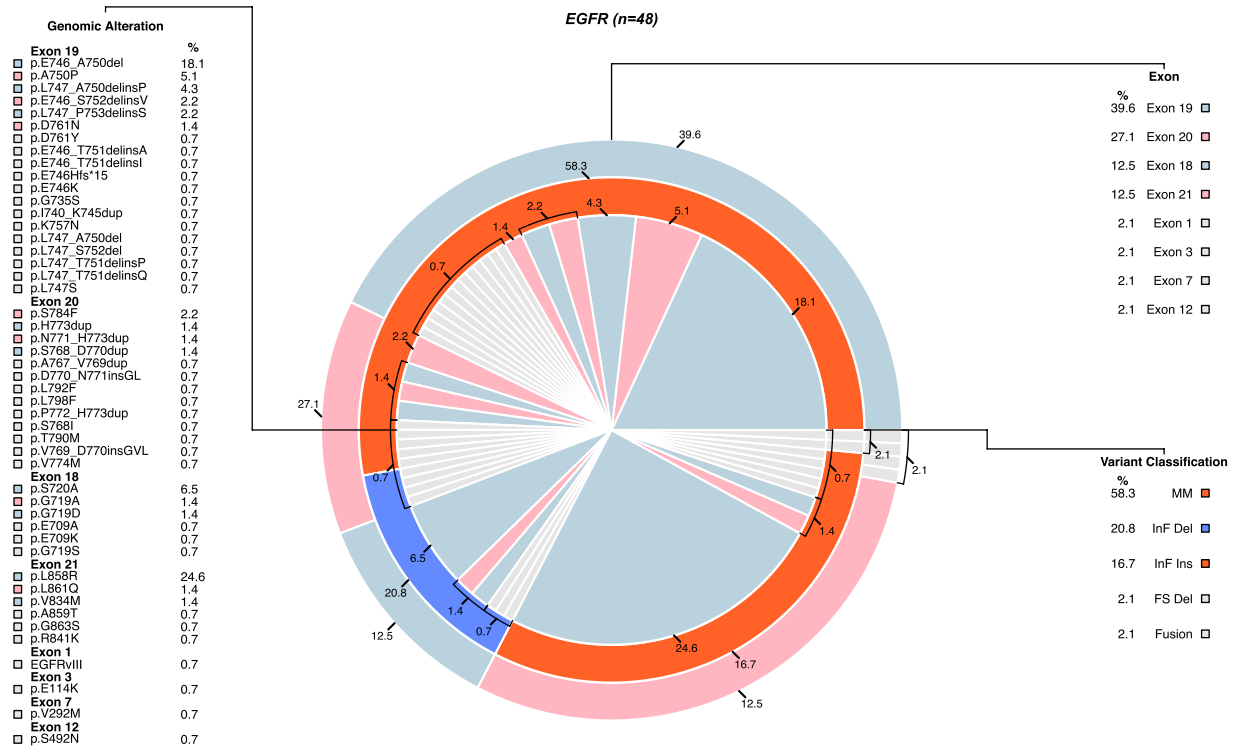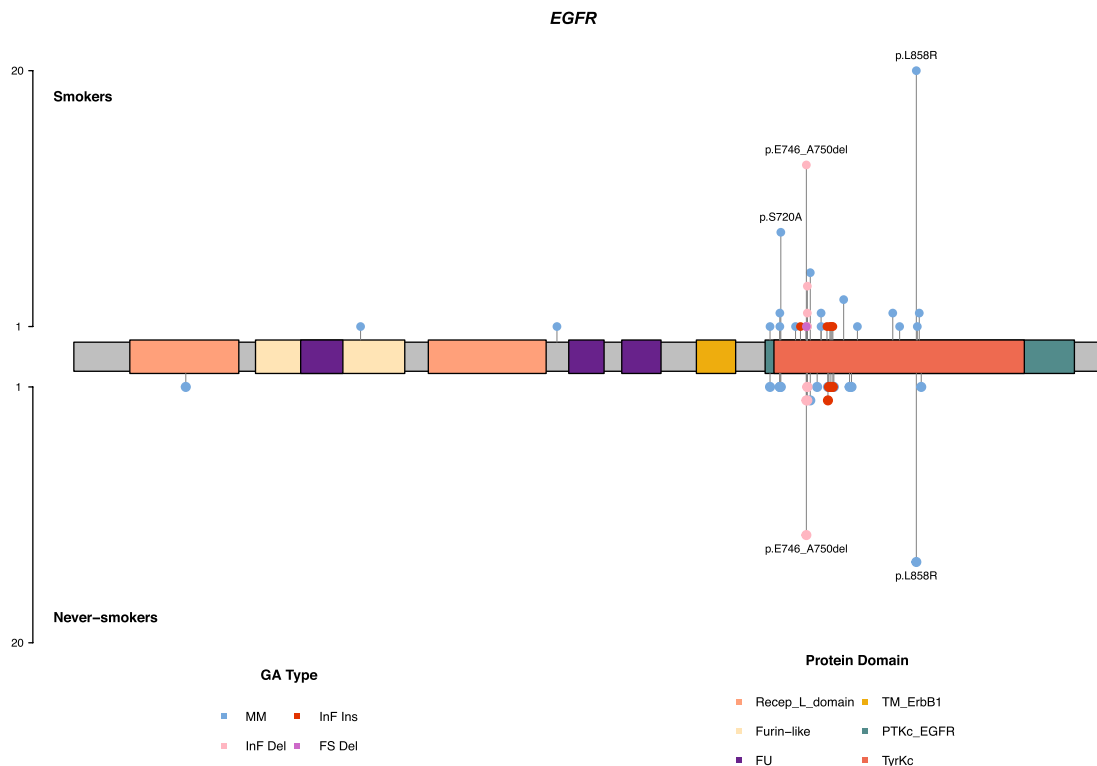
Fig. 6 Characterization of individual GA identified in the gene EGFR (*n* = 48). Panel A: pie chart displaying the relationship between the frequency of individual GAs (inner circle), type of GA (middle circle), and exon number to which individual GAs belong (outer circle). For all circles, grey color represents GA present only once (*n* = 1). Panel B: lolliplot specifying the location and counts of individual GAs for the group of smokers (top) and never-smokers (bottom). For more details see Additional File 2: Table S3

Garrido *et al. BMC Cancer*   (2024) 24:951

Page 10 of 15

Fig. 7 Characterization of individual GA identified in the gene PIK3CA (*n* = 33). Panel A: pie chart displaying the distribution and relationship between the frequency of individual GAs (inner circle), type of GA (middle circle), and exon number to which individual GAs belong (outer circle). For all circles, grey color represents GA present only once (*n* = 1). Panel B: lolliplot specifying the location and counts of individual GAs for the group of smokers (top) and never-smokers (bottom). For more details see Additional File 2: Table S4

Garrido *et al. BMC Cancer*   (2024) 24:951

Page 11 of 15

Table 2 Detail of the common GA between smokers and never-smokers (*n* = 32)

GA Name/ Gene	Total (n)	Smoker (n)	Never-smoker (n)	GA Type
EGFR (*n* = 10)				
1 p.L858R	34	20	14	MM
2 p.E746_A750del	25	13	12	InF Del
3 p.S720A	9	8	1	MM
4 p.A750P	7	5	2	MM
5 p.L747_A750delinsP	6	4	2	InF Del
6 p.E746_S752delinsV	3	1	2	InF Del
7 p.L747_P753delinsS	3	2	1	InF Del
8 p.G719A	2	1	1	MM
9 p.H773dup	2	1	1	InF Ins
10 p.N771_H773dup	2	1	1	InF Ins
KRAS (*n* = 6)				
1 p.G12C	38	31	7	MM
2 p.G12V	20	14	6	MM
3 p.G12D	19	14	5	MM
4 p.G12A	9	7	2	MM
5 p.G12S	3	1	2	MM
6 p.G13D	2	1	1	MM
PIK3CA (*n* = 5)				
1 p.E545K	13	9	4	MM
2 p.E542K	5	3	2	MM
3 p.H1047R	4	1	3	MM
4 p.R38H	4	2	2	MM
5 p.D1045N	2	1	1	MM
MET (*n* = 5)				
1 X1010_splice	4	3	1	SS
2 D1010H	3	2	1	MM
3 H1094Y	3	1	2	MM
4 p.P1009S	3	2	1	MM
5 METex14	18	11	7	Fus
ALK (*n* = 2)				
1 p.L1204F	2	1	1	MM
2 EML4-ALK	22	10	12	Fus
BRAF (*n* = 2)				
1 p.V600E	6	5	1	MM
2 p.G466V	2	1	1	MM
MAP2K1 (*n* = 1)				
1 p.Q56K	30	25	5	MM
ERBB2 (*n* = 1)				
1 p.Y772_A775dup	6	4	2	InF Ins

MM Missense Mutation, *InF Del* In-Frame Deletion, *InF Ins* In-Frame Insertion, *SS* Splite Site, *Fus* Fusion

with SqCC subtype and advanced disease at diagnosis, and a negative history of cancer, ii) never-smokers without alterations in their profiles are more likely to have advanced disease and a negative history of cancer, and

iii) higher prevalence of alterations in *EGFR* are found in never-smoker young women with early-stage disease and LUAD subtype, and a positive PHC and FHC.

Ancestry-mediated variations in the prevalence of alterations of LC from LA are well-documented, particularly for genes with approved targeted therapy in the region [17–20]. Compared to Peru, Mexico and Ecuador, lower prevalence estimates of *EGFR* alterations are reported for Chile, with values similar to those found in European and White patients [21, 22]. Differences based on raced are also found for individual alterations. Common mutations such as deletions of exon 19 and p.L858R point mutation of exon 21 account for 80% to 90% of all *EGFR* mutations [20]. In our study, however, these represent around 50% of all alterations of this gene, with individual mutation prevalence of 6.4 and 5% respectively, which is considerably lower to has been reported in other populations [20–22]. Less common mutations like exon 20 insertions and points mutations p.G719X in exon 18, p.L861X in exon 21, and S768I in exon 20 were all found to have a prevalence close to 1% in the present study, which is consistent with other reports [20, 23]. Interestingly, multiple mutations in *EGFR* were found in < 3% of the patients in our study, which is higher than frequencies reported elsewhere [20, 24]. Compound *EGFR* mutations have been shown to be less responsive to therapy targeting this gene than single mutations [25]. In regard to clinical variables, our finding of higher proportion of *EGFR*-mutated samples for never-smoker women with LUAD and early-stage disease confirms what has been reported in recent metanalysis of LA patients [20]; age, however, appears to be a discordant factor as our results point to younger women. With respect to *KRAS*, the estimates of 14.9% is within the range found for other LA countries [20]; nevertheless, it is lower than values reported for Whites, Blacks, the GENIE database, and higher than values for Asians [26]. The p.G12C mutation was the most common at 5.6%, making it less common than prevalence estimates of other LA countries (7%); likewise for p.G12D (2.8 vs. 4%) [20]. Studies have shown that mutation p.G12C is more predominant in former/current smokers while p.G12D in never smokers. However, our data shows both of these alterations are more frequent in smokers. Race-driven variability for other actionable genes like *ALK*, *ROS1*, *BRAF* is less clear, given the low prevalence estimates. Nevertheless, our values are similar to those reported elsewhere: between 2.8 and 5% for *ALK* fusions [20, 27, 28], 1.9 and 2.2 for *ROS1* rearrangements [11, 20], and 2 and 6% for *BRAF* mutations [20, 29]. In our study, the most common *BRAF* mutation, V600E, accounts for 31.5% of alterations in this gene, which is lower than what has been found in other LA populations (50% and 68%) [30, 31]. For *MET*, while

Garrido *et al. BMC Cancer*   (2024) 24:951

Page 12 of 15

a previous study did not identify alterations in this gene for Chilean patients [32], our data shows a prevalence of 4.5%, with METex14 accounting for 50% of all *MET* alterations. No previous data for Chilean patients were found for *ERBB2* and *RET*. For *ERRB2*, our estimate of 2.5 is lower than what has been found in other LA populations with values ranging from 4.9 to 11% [29, 31, 33, 34]. For *RET*, our estimate of 1.6% is similar to those reported elsewhere [20, 29, 35].

Contrasting results were observed for non-actionable genes. While our *PIK3CA* estimate is within the range of values reported in LA, estimates for *MAP2K1* are notably higher than those reported in the literature (5.6 vs < 1%), particularly for smokers [36, 37]. Alterations in *MAP2K1* have been found to be more frequent in patients of African descent [37]. The present study also shows a potential enrichment of *MAP2K1*-mutant samples in this group of Chilean patients. Point mutation p.Q56X is significantly more common that p.K57X (78.9 vs 2.6), which is contrary to frequencies reported by other authors [36, 37]. These findings constitute valuable discoveries that warrant further investigation, specially giving the promising results that targeted therapies for this gene have shown in LC and other solid tumors.

The results from the present study could have important implications for the management of Chilean LC patients. Current international guidelines recommend molecular testing of *EGFR*, *ALK*, and *ROS1* for all patients with advanced-stage LC with an adenocarcinoma component, and *ERBB2, MET, BRAF, KRAS*, and *RET* in laboratories performing NGS [38]. However, access to standard of care molecular diagnosis for LC in Chile is limited. The most widely available techniques are qPCR and immunohistochemistry-based assays for the assessment of established actionable alterations in *EGFR*, *ALK*, and *ROS1*. Target or comprehensive NGS based assays are available in a small number of private hospitals in the capital city, and these lack insurance coverage or reimbursements [32, 39–41]. On the other hand, targeted therapy currently approved by local regulatory authorities for metastatic NSCLC include *EGFR* tyrosine kinase inhibitors (TKIs), and *ALK*, *ROS1* and *BRAF* inhibitors [39]. Similar to molecular testing, access to these drugs pose a significant problem given their high cost, as in the vast majority of cases they are not reimbursed and must be paid directly by patients. In addition, many FDA-approved drugs are yet to be registered in the country, particularly for *KRAS*, *MET*, *RET* and *ERRB2*, as well as some second- and third generation *EGFR*-TKIs. In this scenario, clinical trial involvement is a viable alternative to receive newer and more effective therapies. The associations found in our study of smoking status and clinical variables with actionable alterations may guide a risk-based selection of patients for access to molecular testing and targeted therapies in these unfavorable settings where financial considerations impose a major constraint [42, 43]. In particular, the more complex genomic profile of smokers also makes molecular testing in this group of patients more relevant, as a more careful consideration of the therapy to use is needed. Furthermore, given that our study indicates that i) estimates of established actionable alterations including those in *EGFR* and *BRAF* are lower than those reported in other populations, and ii) estimates for *ALK* fusions and *ROS1* rearrangements are equally low to those described elsewhere, approval and testing of drugs targeting other genes is crucial. Based on our findings, drugs targeting genes such as *KRAS*, *MET* and *MAP2K1* should be prioritized.

This study characterized for the first time the differences in the genomic profiles between smokers and never-smokers LC patients from Chile; as such constitutes a valuable effort to close the gap in the understanding of underrepresented populations. Nonetheless, it is not without limitations. First, small sample sizes and class imbalances may have hindered the possibility of detecting statistically significant findings. Second, our descriptive analyses did not correct for known and unknown confounding factors; therefore, whether the observed group differences can be explained by factors other than tobacco consumption is yet to be determined. Third, categorization between smokers and never-smokers was made according to self-report at enrollment. More formally, never-smokers are defined as people who have smoked less than 100 cigarettes in their lifetime, in contrast to ever smokers (current and former) who are people who have smoked more than 100 cigarettes over their life [44]**.** Thus, a more robust categorization would have included this aspect into account or employed metrics such the Comprehensive Smoking Index**,** which aggregates duration, intensity and time since cessation [45]. Fourth, the study utilized FFPE with at least 5% tumor content, possibly limiting the detection of variants as most assays require closer to 20–30%. Finally, the OFA is targeted NGS panel designed to detect the presence of specific and restricted number of alterations (> 1,000), which may result in an underestimation of the genomic prevalence when compared to more comprehensive untargeted assays. Also, the determination of important somatic genomic features such as tumor mutational burden and mutational signatures are not possible because of the small size of the OFA DNA targeted regions.

Conclusions

High-quality local genomic data is essential to ease the transition to a more widespread use of molecular testing and targeted therapy approaches for the

Garrido *et al. BMC Cancer*   (2024) 24:951

Page 13 of 15

management of LC patients in LA countries. Here, we provided a thorough characterization of the genomic landscape of Chilean LC patients by smoking status. We found important differences in the prevalence of alterations compared to other countries of the region as well as by tobacco use, which are mainly driven by genes *EGFR*, *KRAS*, *MET*, and *MAP2K1*. Smokers appear to face a more challenging prospect as they are less likely to have actionable mutations and more likely to harbor a complex genomic profile. It is our expectation these findings offer guidance to clinicians and regulatory agencies for management of LC patients in these limited-resource settings where financial constraints are a major hurdle.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12885-024-12737-1.

Supplementary Material 1.

Supplementary Material 2.

Availability of data and materials

The data supporting the findings of this study are available within the article and its supplementary materials.

Declarations

Ethics approval and consent to participate

Institutional Review Board Statement: The study was performed in accordance with the Declaration of Helsinki. The protocol was registered with the identifier NCT03220230 (https://clinicaltrials.gov, last accessed May 27, 2021) and approved by local ethics committees for each recruiting hospital: Comite Etico Cientifico Av. Salvador 364, Servicio de Salud Metropolitano Oriente Providencia, SANTIAGO, RM 7500922 CHILE, Comite de Etica Hospital Clinico Universidad de Chile Avenida Santos Dumont 999 Santiago, RM 8380456 CHILE, Comite Etico Cientifico. Servicio de Salud Concepcion San Martin 1436 Concepcion, REGION DEL BIO BIO 4070038 CHILE. Informed Consent Statement: All patients provided informed consent for access to clinical, demographic, and pathology information as well as to available FFPE tumor tissue. Strategies to protect their identity and privacy included anonymization procedures and a unique eight-digit identifier. Patients received no treatment as part of this study.

Consent for publication

Not applicable.

Competing interests

GS, AB, RA, RAC, MF, LR, DA, RL, JC and PP were Pfizer Chile employees. HF, EDN, DNN, GPB, MGA, CF, TFB, JF, MA, SC, OA, MLS, GR, CS, KM and SR received a grant and non-financial support for to perform this work for CEMP Pfizer Chile. Outside this work, HF discloses personal fees and non- financial support from Pfizer and BMS and non-financial support from AstraZeneca and Roche. RA declares honoraria for conferences, advisory boards, and educational activities from Roche, grants, and support for scientific research from Illumina, Pfizer, Roche & Thermo Fisher Scientific, and honoraria for conferences from Thermo Fisher Scientific, Janssen & Tecnofarma. The other authors declare that they have no competing interests. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Author details

[1]Centro Genética y Genómica, Instituto de Ciencias E Innovación en Medicina, Facultad de Medicina Clínica Alemana, Universidad del Desarrollo, Santiago, Chile. [2]Departamento de Oncología Básico Clínica, Facultad de Medicina, Universidad de Chile, Santiago, Chile. [3]Centro Para La Prevención y el Control del Cáncer, Universidad de Chile, Santiago, Chile. [4]Departamento de Hematología y Oncología and Advanced Center for Chronic Diseases, Pontificia Universidad Católica de Chile, Santiago, Chile. [5]Centro de Oncología de Precisión, Universidad Mayor, Santiago, Chile. [6]CORFO Center of Excellence in Precision Medicine Pfizer, Santiago, Chile. [7]Instituto de Ciencias Biomédicas, Facultad de Medicina, Universidad de Chile, Santiago, Chile. [8]NIDCR, National Institute of Health, Bethesda, USA. [9]Centro de Investigación Clínica Bradford Hill, Santiago, Chile. [10]Instituto Nacional del Tórax, Santiago, Chile. [11]Departamento de Patología, Hospital Clínico Regional de Concepción Dr. Guillermo Grant Benavente Chile, Concepcion, Chile. [12]Departamento de Salud Pública, Facultad de Medicina, Universidad Católica del Norte, La Serena, Chile. [13]Departamento de Ciencias Biomédicas, Facultad de Medicina, Universidad Católica Del Norte, La Serena, Chile. [14]Departamento de Medicina Interna, Servicio de Oncología, Hospital Clínico de La Universidad de Chile, Santiago, Chile. [15]Hospital de Puerto Montt,

Garrido *et al. BMC Cancer*   (2024) 24:951

Page 14 of 15

Puerto Montt, Chile. [16]Centro de Genómica y Bioinformática, Universidad Mayor, Santiago, Chile. [17]Laboratory of Medical Genomics, A. C. Camargo Cancer Center, Sao Paulo, Brazil. [18]Department of Clinical Oncology, A.C. Camargo Cancer Center, Sao Paulo, Brazil. [19]Instituto de Investigación Interdisciplinario, Vicerrectoría Académica, Universidad de Talca, Talca, Chile. [20]Department of Radiation Oncology, Cancer Institute of New Jersey, Rutgers University, Newark, NJ, USA. [21]Millennium Institute for Immunology and Immunotherapy, Santiago, Chile.

References

1. Piñeros M, Laversanne M, Barrios E, de Camargo Cancela M, de Vries E, Pardo C, Bray F. An updated profile of the cancer burden, patterns and trends in Latin America and the caribbean. Lancet Reg Health Am. 2022;13:1–14. https://doi.org/10.1016/j.lana.2022.100294.
2. Wakelee HA, Chang ET, Gomez SL, Keegan THM, Feskanich D, Clarke CA, Holmberg L, Yong LC, Kolonel LN, Gould MK, et al. Lung cancer incidence in never-smokers. J Clin Oncol. 2007;25:472–8. https://doi.org/10.1200/JCO.2006.07.2983.
3. Field RW, Withers BL. Occupational and environmental causes of lung cancer. Clin Chest Med. 2012;33:681–703. https://doi.org/10.1016/j.ccm.2012.07.001.
4. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. CA Cancer J Clin. 2020;70:7–30. https://doi.org/10.3322/caac.21590.
5. Pinheiro PS, Callahan KE, Medina HN, Koru-Sengul T, Kobetz EN, Gomez SL, de Lima Lopes G. Lung cancer in never smokers: distinct population-based patterns by age, sex, and race/ethnicity. Lung Cancer. 2022;174:50–6. https://doi.org/10.1016/j.lungcan.2022.10.009.
6. Clément-Duchêne C, Vignaud JM, Stoufflet A, Bertrand O, Gislard A, Thiberville L, Grosdidier G, Martinet Y, Benichou J, Hainaut P, et al. Characteristics of never smoker lung cancer including environmental and occupational risk factors. Lung Cancer. 2010;67:144–50. https://doi.org/10.1016/j.lungcan.2009.04.005.
7. Boeckx B, Shahi RB, Smeets D, De Brakeleer S, Decoster L, Van Brussel T, Galdermans D, Vercauter P, Decoster L, Alexander P, et al. The genomic landscape of nonsmall cell lung carcinoma in never smokers. Int J Cancer. 2020;146:3207–18. https://doi.org/10.1002/ijc.32797.
8. Clément-Duchêne C, Stock S, Xu X, Chang ET, Gomez SL, West DW, Wakelee HA, Gould MK. Survival among never-smokers with lung cancer in the cancer care outcomes research and surveillance study. Ann Am Thorac Soc. 2016;13:58–66. https://doi.org/10.1513/AnnalsATS.201504-241OC.
9. Chapman AM, Sun KY, Ruestow P, Cowan DM, Madl AK. Lung cancer mutation profile of EGFR, ALK, and KRAS: meta-analysis and comparison of never and ever smokers. Lung Cancer. 2016;102:122–34. https://doi.org/10.1016/j.lungcan.2016.10.010.
10. Barlesi F, Mazières J, Merlio J, Debieuvre D, Léna H, Ouafik H, Besse B, Rouquette I, Westeel V. Routine molecular profiling of cancer : results of a one-year nationwide program of the french cooperative thoracic intergroup (IFCT) for advanced non-small cell lung cancer (NSCLC) patients. Lancet. 2016;287:1415–26. https://doi.org/10.1016/S0140-6736(16)00004-0.
11. Zhu Q, Zhan P, Zhang X, Lv T, Song Y. Clinicopathologic characteristics of patients with ROS1 fusion gene in non-small cell lung cancer: a meta-analysis. Transl Lung Cancer Res. 2015;4:300–9. https://doi.org/10.3978/j.issn.2218-6751.2015.05.01.
12. Okudela K, Matsumura M, Arai H, Woo T. The nonsmokers' and smokers' pathways in lung adenocarcinoma: histological progression and molecular bases. Cancer Sci. 2021;112:3411–8. https://doi.org/10.1111/cas.15031.
13. Arcila ME, Drilon A, Sylvester BE, Lovly CM, Borsu L, Reva B, Kris MG, Solit DB, Ladanyi M. MAP2K1 (MEK1) Mutations define a distinct subset of lung adenocarcinoma associated with smoking. Clin Cancer Res. 2015;21:1935–43. https://doi.org/10.1158/1078-0432.CCR-14-2124.
14. Rojas Hitschfeld HA. Análisis epidemiológico: mortalidad por cáncer de pulmón e indicadores de consumo de cigarrillo en Chile. Revista Instituto Salud Pública Chile. 2021;5:21–7. https://doi.org/10.34052/rispch.v5i2.145.
15. Martinez VD, Vucic EA, Lam S, Lam WL. Arsenic and lung cancer in never-smokers: lessons from Chile. Am J Respir Crit Care Med. 2012;185:1131–2. https://doi.org/10.1164/ajrccm.185.10.1131.
16. Burotto M, Aren O, Renner A, Samtani S, de la Jara JJ. Lung cancer in Chile. J Thorac Oncol. 2019;14:1504–9. https://doi.org/10.1016/j.jtho.2019.02.024.
17. Gimbrone NT, Sarcar B, Gordian ER, Rivera JI, Lopez C, Yoder SJ, Teer JK, Welsh EA, Chiappori AA, Schabath MB, et al. Brief report: somatic mutations and ancestry markers in hispanic lung cancer patients. J Thoracic Oncol. 2017;12:1851–6. https://doi.org/10.1016/J.JTHO.2017.08.019.
18. Carrot-Zhang J, Soca-Chafre G, Patterson N, Thorner AR, Nag A, Watson J, Genovese G, Rodriguez J, Gelbard MK, Corrales-Rodriguez L, et al. Genetic ancestry contributes to somatic mutations in lung cancers from admixed Latin American populations. Cancer Discov. 2021;11:591–8. https://doi.org/10.1158/2159-8290.CD-20-1165/333527/AM/GENETIC-ANCESTRY-CONTRIBUTES-TO-SOMATIC-MUTATIONS.
19. Cardona AF, Arrieta-Mercado O, Ruíz-Patiño A, Rojas L, Rolfo C. Características de los pacientes con cancer de pulmon de celula no pequeña en el instituto nacional de cancerología de Colombia. Rev Colomb Cancerol. 2021;25:226–31. https://doi.org/10.35509/01239015.815.
20. Parra-Medina R, Castañeda-González JP, Montoya L, Gómez-Gómez MP, Clavijo Cabezas D, Plazas Vargas M. Prevalence of oncogenic driver mutations in Hispanics/Latin patients with lung cancer. Systematic review meta-analysis. Lung Cancer (Amsterdam, Netherlands). 2023;185:107378. https://doi.org/10.1016/J.LUNGCAN.2023.107378.
21. Zhang YL, Yuan JQ, Wang KF, Fu XH, Han XR, Threapleton D, Yang ZY, Mao C, Tang JL. The Prevalence of EGFR mutation in patients with non-small cell lung cancer: a systematic review and meta-analysis. Oncotarget. 2016;7:78985–93. https://doi.org/10.18632/oncotarget.12587.
22. Melosky B, Kambartel K, Häntschel M, Bennetts M, Nickens DJ, Brinkmann J, Kayser A, Moran M, Cappuzzo F. Worldwide prevalence of epidermal growth factor receptor mutations in non-small cell lung cancer: a meta-analysis. Mol Diagn Ther. 2022;26:7. https://doi.org/10.1007/S40291-021-00563-1.
23. John T, Taylor A, Wang H, Eichinger C, Freeman C, Ahn MJ. Uncommon EGFR Mutations in non-small-cell lung cancer: a systematic literature review of prevalence and clinical outcomes. Cancer Epidemiol. 2022;76:102080. https://doi.org/10.1016/J.CANEP.2021.102080.
24. Castañeda-González JP, Chaves JJ, Parra-Medina R. Multiple Mutations in the EGFR gene in lung cancer: a systematic review. Transl Lung Cancer Res. 2022;11:2148–63. https://doi.org/10.21037/TLCR-22-235/COIF.
25. Kim EY, Cho EN, Park HS, Hong JY, Lim S, Youn JP, Hwang SY, Chang YS. Compound EGFR mutation is frequently detected with co-mutations of actionable genes and associated with poor clinical outcome in lung adenocarcinoma. Cancer Biol Ther. 2016;17:237–45. https://doi.org/10.1080/15384047.2016.1139235.
26. Shi H, Seegobin K, Heng F, Zhou K, Chen R, Qin H, Manochakian R, Zhao Y, Lou Y. Genomic landscape of lung adenocarcinomas in different races. Front Oncol. 2022;12:946625. https://doi.org/10.3389/FONC.2022.946625/FULL.
27. Fan L, Feng Y, Wan H, Shi G, Niu W. Clinicopathological and demographical characteristics of non-small cell lung cancer patients with ALK rearrangements: a systematic review and meta-analysis. PLoS ONE. 2014;9:100866. https://doi.org/10.1371/JOURNAL.PONE.0100866.
28. Lin HM, Wu Y, Yin Y, Niu H, Curran EA, Lovly CM, Humphries MJ. Real-world ALK testing trends in patients with advanced non–small-cell lung cancer in the United States. Clin Lung Cancer. 2023;24:e39–49. https://doi.org/10.1016/j.cllc.2022.09.010.
29. Velazquez A, Olazagasti C, Duma N. P59.09 Landscape of targetable genomic alterations in Hispanic/Latinx patients with non-small cell lung cancers. J Thorac Oncol. 2021;16:S1150–1. https://doi.org/10.1016/j.jtho.2021.08.598.
30. Andreis TF, Correa BS, Vianna FS, Leistner-Segal F, Siebert M, Leistner-Segal S, Hahn EC, Ulbrich JM, RRivero LF, De Oliveira FH, et al. Analysis of predictive biomarkers in patients with lung adenocarcinoma from Southern Brazil reveals a distinct profile from other regions of the country. J Global Oncol. 2019;5:1–9. https://doi.org/10.1200/JGO.19.00174.
31. Mascarenhas E, Gelatti AC, Araújo LH, Baldotto C, Mathias C, Zukin M, Werutsky G, Pacheco P, Gomes R, de Castro G, et al. Comprehensive genomic profiling of brazilian non-small cell lung cancer patients (GBOT

0118/LACOG0418). Thoracic Cancer. 2021;12:580–7. https://doi.org/10.1111/1759-7714.13777.

32. Martin C, Cuello M, Barajas O, Recondo G, Aruachan S, Perroud H, Sena S, Bonilla C, Orlandi F, Berutti S, et al. Real-world evaluation of molecular testing and treatment patterns for EGFR mutations in non-small cell lung cancer in Latin America. Mol Clin Oncol. 2022;16:6. https://doi.org/10.3892/MCO.2021.2439.

33. Hernández-Pedro N, Soca-Chafre G, Alaez-Versón C, Carrillo-Sánchez K, Avilés-Salas A, Vergara E, Arrieta O. Mutational profile by targeted next generation sequencing of non-small cell lung cancer in the Mexican population. Salud Publica Mex. 2019;61:308–17. https://doi.org/10.21149/10113.

34. Zheng R, Yin Z, Alhatem A, Lyle D, You B, Jiang AS, Liu D, Jobbagy Z, Wang Q, Aisner S, et al. Epidemiologic features of NSCLC gene alterations in Hispanic patients from Puerto Rico. Cancers. 2020;12:1–12. https://doi.org/10.3390/CANCERS12123492.

35. Raez LE, Santos ES, Mourafetis J, Kim AY, Hunis B, Sareli C, Castillero L, Velis E. P1.13 (also presented as PD2.01): lung cancer chromosomal aberrations and gene expression profiles of hispanics living in the US or Latin America are similar. J Thor Oncol. 2016;11:S187–8. https://doi.org/10.1016/j.jtho.2016.08.035.

36. Scheffler M, Holzem A, Kron A, Nogova L, Ihle MA, von Levetzow C, Fassunke J, Wömpner C, Bitter E, Koleczko S, et al. Co-occurrence of targetable mutations in Non-Small Cell Lung Cancer (NSCLC) patients harboring MAP2K1 mutations. Lung Cancer. 2020;144:40–8. https://doi.org/10.1016/j.lungcan.2020.04.020.

37. Cheng ML, Lee JK, Kumar R, Klein H, Raskina K, Schrock AB, Michael KS, Mazor T, Cerami E, Oxnard GR, et al. Response to MEK inhibitor therapy in MAP2K1 (MEK1) K57N non–small-cell lung cancer and genomic landscape of MAP2K1 mutations in non–small-cell lung cancer. JCO Precis Oncol. 2022;6:e2200382. https://doi.org/10.1200/PO.22.00382.

38. Smeltzer MP, Wynes MW, Lantuejoul S, Soo R, Ramalingam SS, Varella-Garcia M, Meadows Taylor M, Richeimer K, Wood K, Howell KE, et al. The international association for the study of lung cancer global survey on molecular testing in lung cancer. J Thorac Oncol. 2020;15:1434–48. https://doi.org/10.1016/j.jtho.2020.05.002.

39. Rivas S, Armisén R. Non-small cell lung cancer in the era of precision medicine. Rev Méd Clín Las Condes. 2022;33:25–35. https://doi.org/10.1016/j.rmclc.2022.01.001.

40. Ríos JA, Marcelain K, Plaza-Parrochia F, Selman C, Bustamante E, Godoy JA, Labbé TP, García-Bloj B. Planificación sanitaria a partir de los datos ómicos: ¿es posible esta idea para Chile? [The contribution of omic sciences for the management of cancer in Chile]. Rev Med Chil. 2021;149:1657–63. https://doi.org/10.4067/S0034-98872021001101657.

41. Alvarez-Gomez RM, De la Fuente-Hernandez MA, Herrera-Montalvo L, Hidalgo-Miranda A. Challenges of diagnostic genomics in Latin America. Curr Opin Genet Dev. 2021;66:101–9. https://doi.org/10.1016/J.GDE.2020.12.010.

42. Sacher AG, Dahlberg SE, Heng J, Mach S, Jänne PA, Oxnard GR. Association between younger age and targetable genomic alterations and prognosis in non-small-cell lung cancer. JAMA Oncol. 2016;2:313–20. https://doi.org/10.1001/JAMAONCOL.2015.4482.

43. Liu X, Jiang G, Sun X, Su G, Zhang X, Shen D, Yan N. Relationship between driver gene mutations and clinical pathological characteristics in older lung adenocarcinoma. Front Oncol. 2023;13:1275575. https://doi.org/10.3389/FONC.2023.1275575.

44. Couraud S, Zalcman G, Milleron B, Morin F, Souquet PJ. Lung Cancer in never smokers – a review. Eur J Cancer. 2012;48:1299–311. https://doi.org/10.1016/J.EJCA.2012.03.007.

45. Leffondré K, Abrahamowicz M, Xiao Y, Siemiatycki J. Modelling smoking history using a comprehensive smoking index: application to lung cancer. Stat Med. 2006;25:4132–46. https://doi.org/10.1002/sim.2680.

46. Sepulveda, G.; Blanco, A.; Freire, M.; Lizana, R.; Caceres-Molina, J.; Ampuero, D.; Perez, P.; Ramos, L.; Aren, O.; Chernilo, S.; et al. Molecular Characterization of Non-Small Cell Lung Cancer Tumors in Latin American Patients from Brazil, Chile and Peru Uncovers Novel Potentially Driver Mutations. medRxiv 2020, 2020.09.11.20171025, https://doi.org/10.1101/2020.09.11.20171025.

47. Sepúlveda-Hermosilla G, Freire M, Blanco A, Cáceres J, Lizana R, Ramos L, Assar Cuevas R, Ampuero D, Aren O, Chernilo S, et al. Concordance analysis of ALK gene fusion detection methods in patients with non-small-cell lung cancer from Chile, Brazil, and Peru. J Mol Diagn. 2021;23:1127–37. https://doi.org/10.1016/j.jmoldx.2021.05.018.

Publisher's Note