ESMO GOOD SCIENCE BETTER MEDICINE BEST PRACTICE

ESMO OPEN CANCER HORIZONS

ORIGINAL RESEARCH

Integrative molecular analysis of combined small-cell lung carcinomas identifies major subtypes with different therapeutic opportunities

M. Simbolo[1†], G. Centonze[2†], G. Ali[3], G. Garzone[2], S. Taormina[1], G. Sabella[2,4], C. Ciaparrone[1], A. Mafficini[1,5], F. Grillo[6], A. Mangogna[7], M. Volante[8], L. Mastracci[6], G. Fontanini[3], S. Pilotto[9], E. Bria[10], M. Infante[11], C. Capella[12], L. Rolli[13], U. Pastorino[13], M. Milella[9], M. Milione[2*‡] & A. Scarpa[1,5‡]

[1]Section of Pathology, Department of Diagnostics and Public Health, University of Verona, Verona; [2]Pathology Unit 1, Pathology and Laboratory Department, Fondazione IRCCS Istituto Nazionale dei Tumori, Milano; [3]Department of Surgical, Medical, Molecular Pathology and Critical Area, University of Pisa, Pisa; [4]School of Pathology, University of Milan, Milan; [5]ARC-Net Research Centre for Applied Research on Cancer, University of Verona, Verona; [6]Department of Surgical and Diagnostic Sciences (DISC), University of Genova and IRCCS S. Martino-IST University Hospital, Genoa; [7]Institute for Maternal and Child Health, IRCCS Burlo Garofalo, Trieste; [8]Department of Oncology, University of Turin at San Luigi Hospital, Orbassano, Torino; [9]Section of Oncology, Department of Medicine, University of Verona, Verona; [10]Policlinico Universitario Agostino Gemelli IRCCS, Università Cattolica del Sacro Cuore, Rome; [11]Thoracic Surgery, University and Hospital Trust of Verona, Verona; [12]Unit of Pathology, Department of Medicine and Surgery and Research Centre for the Study of Hereditary and Familial tumors, University of Insubria, Varese; [13]Thoracic Surgery Unit, Fondazione IRCCS Istituto Nazionale Tumori, Milan, Italy

Available online XXX

Background: Combined small-cell lung cancer (C-SCLC) is composed of SCLC admixed with a non-small-cell cancer component. They currently receive the same treatment as SCLC. The recent evidence that SCLC may belong to either of two lineages, neuroendocrine (NE) or non-NE, with different vulnerability to specific cell death pathways such as ferroptosis, opens new therapeutic opportunities also for C-SCLC.

Materials and methods: Thirteen C-SCLCs, including five with adenocarcinoma (CoADC), five with large-cell neuroendocrine carcinoma (CoLCNEC) and three with squamous cell carcinoma (CoSQC) components, were assessed for alterations in 409 genes and transcriptomic profiling of 20 815 genes.

Results: All 13 cases harbored *TP53* (12 cases) and/or *RB1* (7 cases) inactivation, which was accompanied by mutated *KRAS* in 4 and *PTEN* in 3 cases. Potentially targetable alterations included two *KRAS* G12C, two *PIK3CA* and one *EGFR* mutations. Comparison of C-SCLC transcriptomes with those of 57 pure histology lung cancers (17 ADCs, 20 SQCs, 11 LCNECs, 9 SCLCs) showed that CoLCNEC and CoADC constituted a standalone group of NE tumors, while CoSQC transcriptional setup was overlapping that of pure SQC. Using transcriptional signatures of NE versus non-NE SCLC as classifier, CoLCNEC was clearly NE while CoSQC was strongly non-NE and CoADC exhibited a heterogeneous phenotype. Similarly, using ferroptosis sensitivity/resistance markers, CoSQC was classified as sensitive (as expected for non-NE), CoLCNEC as resistant (as expected for NE) and CoADC showed a heterogeneous pattern.

Conclusions: These data support routine molecular profiling of C-SCLC to search for targetable driver alterations and to precisely classify them according to therapeutically relevant subgroups (e.g. NE versus non-NE).

Key words: combined small-cell lung carcinoma, small-cell lung cancer, neuroendocrine carcinoma, next-generation sequencing, transcriptomics

INTRODUCTION

Combined small-cell lung carcinoma (C-SCLC) is defined by the World Health Organization (WHO) classification of lung tumors as small-cell carcinoma (SCLC) combined with additional components belonging to any of the histological types of non-small-cell lung carcinoma (NSCLC). These usually include adenocarcinoma (ADC), squamous cell carcinoma (SQC), large-cell neuroendocrine carcinoma (LCNEC) and less commonly spindle cell carcinoma or giant cell carcinoma.[1] No minimum percentage of the additional component is required for a C-SCLC diagnosis with the exception of mixed LCNEC and SCLC, where a minimum of 10% LCNEC component is required given the frequent presence of scattered large cells in surgically resected SCLC.[1]

Patients diagnosed with C-SCLC are currently recommended to receive the same treatment as SCLC in the absence of clear evidence suggesting different strategies.[2,3]

*Correspondence to: Dr Massimo Milione, Pathology Unit 1, Pathology and Laboratory Department, Fondazione IRCCS Istituto Nazionale dei Tumori, Via Venezian 1, 20133, Milan, Italy. Tel: +390223903460

E-mail: massimo.milione@istitutotumori.mi.it (M. Milione).

†Shared first authorship.

‡Shared last authorship.

Combination of chemotherapy and immunotherapy targeting the programmed cell death protein 1/programmed death-ligand 1 checkpoint has been suggested as the best standard for both SCLC and NSCLC without altered genomic drivers.[4]

Reports on the molecular characterization of C-SCLC are still scarce and their biological and clinical features are poorly understood. To date, only three papers have reported molecular data on C-SCLC obtained by next-generation sequencing (NGS). Two studies addressed their genomic profiling using multigene targeted sequencing.[5,6] The first studied 10 C-SCLCs, comprising 5 with an SQC, 3 with an LCNEC and 2 with an ADC component.[5] The second work profiled 12 cases, 6 with an SQC component and 6 with an ADC component.[6] These studies highlighted *TP53* and *RB1* as the most frequently altered genes in C-SCLC and suggested that the different C-SCLC components could derive from the same single pluripotent clone, as they shared driver molecular alterations. The third study was an integrated genomic and transcriptomic analysis of 75 LCNECs including 9 C-SCLCs/LCNECs, which were distributed across the different transcriptomic categories described.[7]

Two mechanisms have been suggested regarding the origin of C-SCLC: heterogeneity and transdifferentiation. Heterogeneity seems to be embedded in SCLC oncogenesis as deduced from genetically engineered mouse models (GEMMs).[8] Indeed, five different models featuring *Rb/p53* knockout, alone or in combination with other genes, all produced combined tumors to a different extent and with a variable admixture of SCLC, LCNEC or NSCLC. Transdifferentiation is the switch of cells from one type to another[9] and has been reported as a mechanism of acquired resistance to therapy in lung ADC.[10,11] Tumors relapsing after treatment displayed SCLC features while retaining the same driver mutations of the initially diagnosed ADC. These data highlight the plasticity of lung cancer histology with respect to its genetic background.

The long-standing notion that SCLC is a single neuroendocrine (NE) entity has been recently challenged. This issue has been revised by Rudin et al. who proposed the existence of two lineages, NE and non-NE.[12] Belonging to either SCLC lineage entails a different vulnerability to specific regulated cell death pathways. Bebber et al. demonstrated that non-NE SCLCs were vulnerable to ferroptosis, while NE SCLCs were resistant to ferroptosis but addicted to the thioredoxin pathway.[13]

These new possibilities related to SCLC heterogeneity are especially implied in C-SCLC, where the tumor lineages and the potential therapeutic approaches could be reflected by the different non-SCLC components. To investigate this scenario, we carried out the integrative analysis of 13 C-SCLCs by investigating the mutational status of 409 cancer-related genes and the transcriptional status of 20 815 genes.

MATERIALS AND METHODS

Cases

A retrospective series (2007-2019) of 13 surgically resected primary C-SCLCs was collected from five Italian institutions (ARC-Net Research Centre-Verona; IRCCS San Martino-Genova; University of Pisa; AUO Orbassano-University of Turin; Fondazione IRCCS Istituto Nazionale Tumori-Milan). In addition, 57 surgically resected primary lung tumors with pure histological features were retrieved for histotype-based comparison of expression profiles, including 17 ADCs, 11 (LCNECs, 9 SCLCs and 20 SQCs. None of the patients had received preoperative therapy. All cases were formalin-fixed and paraffin-embedded (FFPE) for routine histological evaluation. Reclassification according to WHO 2015 criteria[1] and confirmation of histological diagnosis were carried out by consensus meetings of pathologists before inclusion of each case in the study. Tumor stage was according to the seventh edition of the TNM (tumor—node—metastasis) classification of malignant tumors.[14] Eight non-neoplastic lung tissues were used as control samples for transcriptomic analysis.

Immunohistochemistry

Monoclonal antibodies and key staining steps are listed in Supplementary Table S1, available at https://doi.org/10.1016/j.esmoop.2021.100308. NE markers synaptophysin (SYN) and chromogranin A were used to confirm the NE nature and to properly estimate the SCLC percentage in each C-SCLC. Napsin A (NAPSA) and thyroid transcription factor 1 (TTF-1) were utilized to identify ADC components, and p40 to detect squamous components. Tumor proliferative index was evaluated only for the NE components by counting the percentage of Ki67-positive cells in areas of strongest nuclear labeling (hot spots).[15]

Mutational and copy number variation status of 409 cancer genes

The Oncomine Tumor Mutational Load (TML) panel (Thermo Fisher Scientific, Waltham, MA) was used to perform NGS of DNA extracted from FFPE tumor tissues. The assay covers 1.65 Mb including the exons of 409 cancer-related genes. Detailed information on technical procedures and data analysis is reported in Supplementary Methods, available at https://doi.org/10.1016/j.esmoop.2021.100308.

Tumor mutational load and mutational signatures

TML and mutational spectrum were evaluated using the Oncomine TML 5.10 plugin on IonReporter (Thermo Fisher Scientific) as detailed in Supplementary Methods, available at https://doi.org/10.1016/j.esmoop.2021.100308. TML is expressed as the number of mutations per megabase (muts/Mb). The mutational spectrum of individual tumors was obtained considering six major mutation classes: C > T; C > A; C > G; T > A; T > C; T > G.[16]

Fusion gene detection by next-generation sequencing

The FusionPlex Solid Tumor Panel (ArcherDX, Boulder, CO) was used to screen for fusions in 57 genes. Details of the protocol and data analysis are reported in Supplementary

Methods, available at https://doi.org/10.1016/j.esmoop.2021.100308.

Gene expression analysis by next-generation sequencing

The Ampliseq Transcriptome Human Gene Expression Kit (Thermo Fisher Scientific) was used to analyze the expression status of 20 815 human genes. Libraries were prepared using AmpliSeq technology and 1 μg of retro-transcribed RNA for each multiplex PCR amplification. Clonal amplification was carried out using the Ion Chef System (Thermo Fisher Scientific). Sequencing was run on the Ion S5XL (Thermo Fisher Scientific) loaded with Ion 540 Chip. The AmpliSeqRNA plugin was used to generate expression data (counts per transcript) for each sample. Counts were normalized and transformed using the 'DESeq2' package for R.[17] Batch effect was removed using 'LIMMA' package for R.[18] Visualization and clustering were carried out using the 'ComplexHeatmap' package for R.[19]

Differential expression analysis between subtypes was carried out using Deseq2 algorithm. A gene was considered differentially expressed if it showed an adjusted *P* value <0.05.

For clustering analysis, best number of clusters (k) was estimated using the NbClust package of R,[20] which provides 30 indexes to determine the number of clusters in a dataset also offering the best clustering scheme. The consensus matrix for the cluster solution identified was carried out using ConsensusClusterPlus package of R[21] with hierarchical clustering, applying average linkage and a Pearson correlation-based distance, while hybrid hierarchical k-means was used to cluster the samples.[22] Cluster-consensus, the average pairwise item consensus of items in a consensus cluster, is reported below for each solution.

Gene set enrichment analysis

We downloaded c2 and c5 pathways from MSigDB[23,24] and determined the cluster-specific enriched gene sets using the normalized and batch-corrected count matrix. We applied gene set enrichment analysis (GSEA) using GAGE[25] R package between clusters to get pairwise significant up- and down-regulated pathways. While we used an approach based on the single-sample GSEA (ssGSEA) score for determining the biological processes differently enriched between all the clusters. We carried out a z-score normalization of the pathway scores in the clusters.

Statistical analysis

One-way analysis of variance, Kruskal—Wallis test, Fisher's test with Monte Carlo simulation and Fisher's exact test were used as appropriate; correction for multiple comparisons was carried out according to Benjamini—Hochberg. Disease-specific survival was assessed by the Kaplan—Meier method, using the date of surgery as the entry point and the patient's death as the endpoint. Patients died of diseases independent from tumor were censored. The Mantel—Cox log-rank test was applied to assess the strength of association between disease-specific survival and molecular alterations as a single variable. A *P* value <0.05 was considered as significant. All analyses were carried out using MedCalc for Windows version 15.6 (MedCalc Software, Ostend, Belgium) and R v. 3.6.3.[26]

Ethics

Ethics committee approval (ECA) was obtained at the five institutions participating in the study: ARC-Net Research Centre-Verona, ECA no. 2173-prot.26775 (1 June 2012); AUO Orbassano-University of Torino, ECA n. 167/2015-prot. 17975 (21 October 2015); IRCCS San Martino-Genova: ECA n. 027/2016LM (16 March 2016); University of Pisa: ECA n. 1040/16 (31 March 2016); Fondazione IRCCS Istituto Nazionale Tumori-Milan ECA n. INT 171/16 (16 November 2016). The study was conducted in accordance with the ethical principles of the Declaration of Helsinki. Results presented in this article contain no personally identifiable information from the study.

RESULTS

Thirteen C-SCLCs were assessed for alterations in 409 genes and their transcriptomic profiles were compared with those of 57 pure histology lung cancers (Supplementary Figure S1, available at https://doi.org/10.1016/j.esmoop.2021.100308).

Clinico-pathological features

The clinical-pathological features of 13 C-SCLCs patients are summarized in Table 1 and detailed in Supplementary Table S2, available at https://doi.org/10.1016/j.esmoop.2021.100308. The series comprised 3 females and 10 males with a median age of 70 years (range: 23-78 years). All cases were reclassified by consensus meeting of pathologists. The non-SCLC component was ADC in five cases (CoADC), LCNEC in five cases (CoLCNEC) and SQC in three cases (CoSQC). One CoADC (#367) was a recurrence in a patient after 2 years from completion of chemotherapy for an ADC diagnosed in 2008.

The different histological components were highly mixed in all cases, with no obvious boundaries. Thus, the proportion of each component was estimated based on both morphology and immunohistochemical markers (Figure 1).

Immunohistochemical features

The immunohistochemical data are summarized in Table 1. SCLC and LCNEC components were positive for SYN, while non-SCLC components were negative. NAPSA immunolabeling was positive in ADC components and p40 in the squamous components. As expected, TTF-1 staining was present in SCLC and LCNEC components, but also focally in non-NE components. The proliferation index (Ki67) was always higher in the SCLC component. In the case of CoLCNECs, the large-cell and small-cell components had a similar proliferation index. Immunostaining for P53 and RB1 was always homogeneous and coherent with the mutational results (Figure 2).

Table 1. Clinical-pathological features of 13 combined small-cell carcinoma patients

	All patients *n* (%)	CoADC *n* (%)	CoSQC *n* (%)	CoLCNEC *n* (%)	*P* value*
Total	13 (100)	5 (100)	3 (100)	5 (100)	
Age, years					
<50	3 (23.1)	0 (0.0)	0 (0.0)	3 (60)	
50-69	3 (23.1)	1 (20)	1 (33.3)	1 (20)	
70+	7 (53.8)	4 (80)	2 (66.6)	1 (20)	0.15
Gender					
Male	10 (76.9)	4 (80)	3 (100)	3 (60)	
Female	3 (23.1)	1 (20)	0 (0.0)	2 (40)	0.42
Smoking					
Actual smoker	11 (84.6)	5 (100)	3 (100)	3 (60)	
Former smoker	2 (15.4)	0 (0.0)	0 (0.0)	2 (40)	0.15
Stage					
I	2 (15.4)	0 (0.0)	0 (0.0)	2 (40)	
II	2 (15.4)	0 (0.0)	1 (33.3)	1 (20)	
III	9 (69.2)	5 (100)	2 (66.6)	2 (40)	0.20
Ki-67					
Median [range]	80 [61-90]	85 [70-90]	80 [70-90]	82.5 [61-91]	0.41
Synaptophysin					
Absent	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	
Present	13 (100)	5 (100)	3 (100)	5 (100)	0.73
Chromogranin A					
Absent	2 (15.4)	2 (40)	0 (0.0)	0 (0.0)	
Present	11 (84.6)	3 (60)	3 (100)	5 (100)	0.15
Napsin A					
Absent	8 (61.5)	0 (0.0)	3 (100)	5 (100)	
Present	5 (38.4)	5 (100)	0 (0.0)	0 (0.0)	**0.0015**
p40					
Absent	10 (76.9)	5 (100)	0 (0.0)	5 (100)	
Present	3 (23.1)	0 (0.0)	3 (100)	0 (0.0)	0.0936
p53					
Absent	2 (15.4)	0 (0.0)	1 (33.3)	1 (20)	
Present	11 (84.6)	5 (100)	2 (66.6)	4 (80)	0.42
rb1					
Absent	7 (53.8)	2 (40)	1 (33.3)	4 (80)	
Present	6 (46.2)	3 (60)	2 (66.6)	1 (20)	0.32
TTF-1					
Absent	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	
Present	13 (100)	5 (100)	3 (100)	5 (100)	0.73

CoADC, small-cell lung cancer combined with adenocarcinoma; CoLCNEC, small-cell lung cancer combined with large-cell neuroendocrine carcinoma; CoSQC, small-cell lung cancer combined with squamous cell carcinoma; TTF-1, thyroid transcription factor 1.
Statistically significant *P* value are reported in bold.
* *P* value based on the Fisher Exact test for categorical variables or the Kruskal-Wallis test for continuous variables.

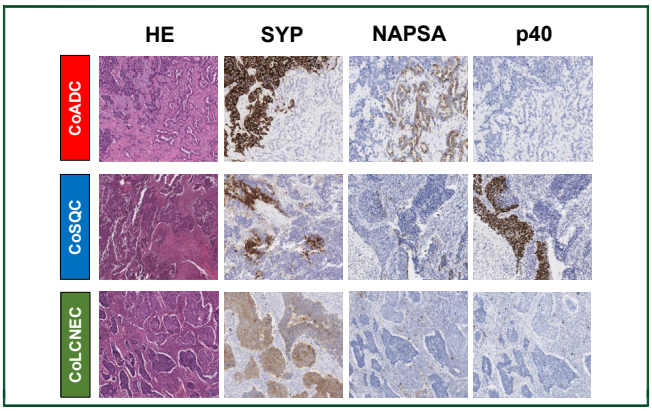

Figure 1. Representative cases of combined-SCLC.
Morphological and immunohistochemical characteristics of the different components in three selected cases of combined small cell neuroendocrine carcinomas. SYP immunostaining identifies both the SCLC and LCNEC components, while NAPSA the adenocarcinoma and p40 the squamous component.
CoLCNEC, small-cell lung cancer combined with large-cell neuroendocrine carcinoma; CoSQC, small-cell lung cancer combined with squamous cell carcinoma; HE, hematoxylin and eosin; NAPSA, napsin A; SYP, synaptophysin.

Mutational status of 409 genes

All cases were analyzed for 409 cancer-related genes, with an average sequencing coverage of 358× (164-865×; Supplementary Table S3, available at https://doi.org/10.1016/j.esmoop.2021.100308). The results are illustrated in Figure 2 and detailed in Supplementary Table S4, available at https://doi.org/10.1016/j.esmoop.2021.100308.

A total of 48 mutations in 25 genes were identified, including 30 missense, 6 nonsense, 6 frameshift, 4 splice site alterations and 2 in-frame deletions. All cases displayed at least one mutation. The most recurrently altered genes were *TP53* (12/13 cases; 92.3%), *RB1* (7/13; 53.8%) and *KRAS* (4/13; 30.8%). Potentially targetable gene alterations included two cases with *KRAS* G12C mutations,[27] two cases with a *PIK3CA* mutation and one case featuring an *EGFR* L858R mutation, amplification of the mutated allele and a subclonal T790M *EGFR* mutation.[28,29] One of the two cases with the *KRAS* G12C mutation was a CoSQC and had a second *KRAS* mutation in codon 13 (G13A) while the tumor with *EGFR* alterations (#367) belonged to the patient previously treated for ADC.

A median TML of 14.5 mutations for Mb (range: 6.8-26.4) was estimated for all C-SCLC (Figure 2 and Supplementary Table S5, available at https://doi.org/10.1016/j.esmoop.2021.100308), which is higher than that of ADC and SQC of the lung,[30] and similar to that observed in SCLC.[31,32] The mutational signatures showed no specific pattern.

Gene and chromosomal copy number alterations

The copy number variation (CNV) status was estimated for all 409 genes from sequencing data. Eight genes showed focal amplification (Figure 2): *CCNE1, EGFR, KIT, MYC, MYCL1, MYCN, PDGFRA, PIM1*. Three genes showed homozygous deletion, including *CDKN2A/B and RB1* in 2/13 cases each (15.4%) and *TP53* gene in 1 case. The status of chromosome arms was also inferred based on the chromosomal position and CNV status of targeted genes. Recurrent events were gain of chromosome arm 8q or part of it including *MYC* (5/13, 38.5%) and loss of heterozygosity (LOH) of chromosome arm 10q (5/13, 38.5%). Cases with a CoADC component exhibited a larger number of alterations but no specific pattern (Supplementary Figure S2, available at https://doi.org/10.1016/j.esmoop.2021.100308).

Fusion genes

No fusion genes were detected by targeted RNA sequencing with the 57-gene panel.

Transcriptional profiles of combined-SCLCs

We conducted differential expression analysis by grouping the cases according to their non-small-cell cancer counterpart, to identify the differences between the three types of

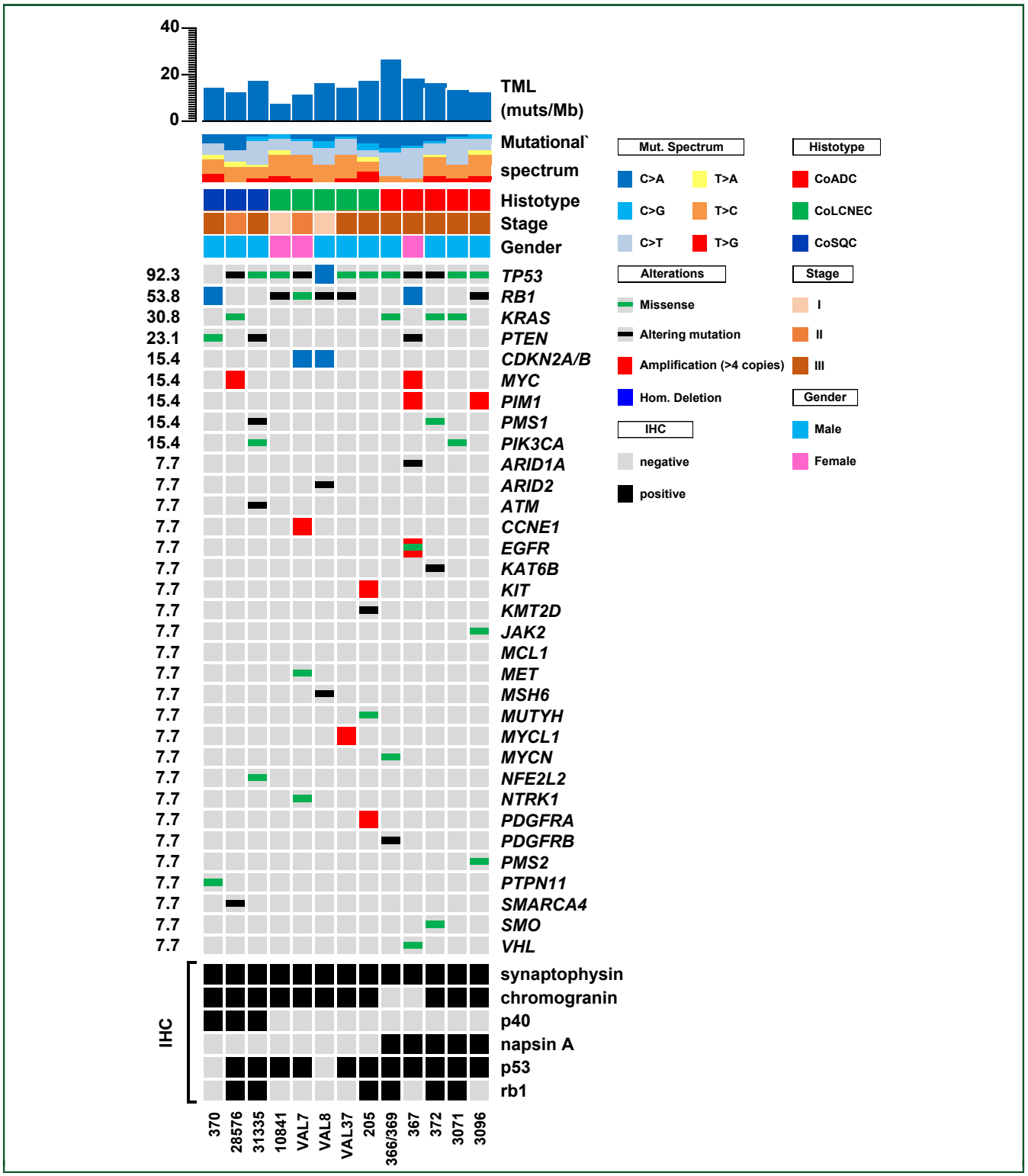

Figure 2. Genomic landscape of 13 combined small cell neuroendocrine carcinomas.
The upper histogram shows tumor mutational load (TML) for each sample, expressed as number of mutations per megabase (muts/Mb). The central matrix shows all altered genes, whose molecular alterations are annotated as illustrated in panel on the right. Genes were ranked according to their alteration frequency. Bottom matrix summarizes immunohistochemistry (IHC) results for the indicated markers.
CoADC, small-cell lung cancer combined with adenocarcinoma; CoLCNEC, small-cell lung cancer combined with large-cell neuroendocrine carcinoma; CoSQC, small-cell lung cancer combined with squamous cell carcinoma; IHC, immunohistochemistry; TML, tumor mutational load.

C-SCLCs under study. We carried out the following comparisons: CoADC versus CoSQC; CoLCNEC versus CoADC; CoLCNEC versus CoSQC. The differential expression analysis identified 630 genes which were able to separate the three groups (Figure 3A and Supplementary Table S6, available at https://doi.org/10.1016/j.esmoop.2021.100308).

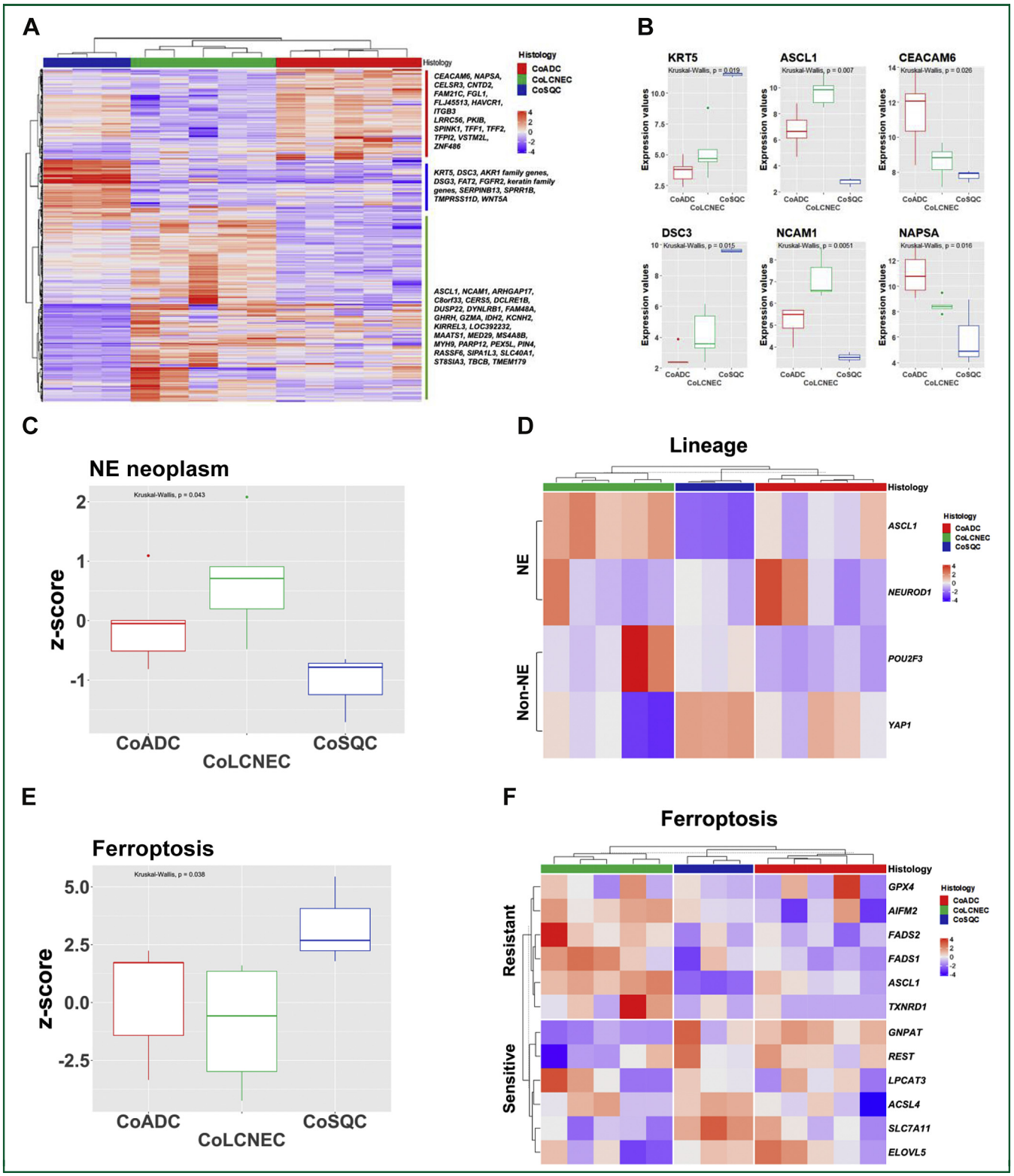

Figure 3. Genes differently expressed between the three subtypes of combined small cell neuroendocrine carcinomas.
(A) Heatmap resulting from 630 differentially expressed genes, which were able to separate the three histological subtypes of combined cancers. The most representative genes for each subtype are listed on the right. (B) Box and whisker plots displaying the expression values for the combined LCNEC (CoLCNEC) markers *ASCL1* and *NCAM1*, combined ADC (CoADC) markers *CEACAM6* and *NAPSA* and combined SQC markers *KRT5* and *DSC3* (box plots: median and interquartile range, whiskers: min—max values). (C) Box and whisker plots displaying the normalized enrichment z-score for the neuroendocrine neoplasm signature (HP_NEUROENDOCRINE_NEOPLASM). ssGSEA was used to obtain the enrichment score. (D) Heatmap of differentially expressed genes identified as discriminating between the Neuroendocrine (NE) and non-Neuroendocrine (non-NE) phenotype in current literature. The analysis showed CoLCNECs as Neuroendocrine-like, CoSQCs as non-Neuroendocrine-like, and CoADCs as intermediate phenotypes. (E) Box and whisker plots displaying the normalized enrichment z-score for the ferroptosis signature (WP_FERROPTOSIS). ssGSEA was used to obtain the enrichment score. (F) Heatmap of differentially expressed genes related to ferroptosis sensitivity (Sensitive) and resistance (Resistant). The analysis showed that CoLCNECs have high expression levels of genes marking sensitivity to ferroptosis.
CoADC, small-cell lung cancer combined with adenocarcinoma; CoLCNEC, small-cell lung cancer combined with large-cell neuroendocrine carcinoma; CoSQC, small-cell lung cancer combined with squamous cell carcinoma; NE, neuroendocrine.

As expected, the main discriminating overexpressed genes were those typical of the non-small-cell component. Among these, six genes were sufficient to distinguish the three groups: *KRT5* and *DSC3* overexpressed in CoSQC, *ASCL1* and *NCAM1* in CoLCNEC, *CEACAM6* and *NAPSA* in CoADC (Figure 3B). Other distinctive markers were identified using the following three criteria: (i) the marker had to be differentially expressed in all comparisons involving the specific subtype; (ii) the marker had to have the highest mean expression among the genes identified in the same comparison; (iii) the marker had to have the lowest adjusted *P* value identified in the same comparison. According to these criteria, distinctive markers for CoADCs included *CNTD2*, *FGL1* and *SPINK1*, for CoSQC *FGFR2*, Keratins and *WNT5A* genes, while for CoLCNEC *IDH2*, *MED29* and *SLC40A1* genes.

We further investigated the NE profile of the three CoSCLC groups by verifying gene set enrichment score for NE neoplasm signature available on MSigDB database. As illustrated in Figure 3C, CoLCNECs exhibited the most NE-like profile while CoSQCs exhibited an opposite profile. CoADCs, on the other hand, showed an intermediate profile.

Furthermore, using the recently proposed molecular classification of SCLC by Rudin et al.,[12] CoSQC could be classified as 'non-neuroendocrine (non-NE) SCLC-Y', based on high levels of *YAP1.* CoLCNECs, could be classified as 'NE SCLC-A', due to high expression of *ASCL1*; interestingly, two CoLCNEC samples co-expressed *POU2F3* and one co-expressed *NEUROD1*. Finally, CoADCs showed a very heterogeneous molecular profile: one sample exhibited high levels of *ASCL1,* defining it as 'NE SCLC-A', two samples overexpressed *NEUROD1*, defining them as 'NE SCLC-N', while the two remaining samples expressed high levels of *YAP1,* defining them as 'non-NE SCLC-Y' (Figure 3D).

A close link between the non-NE profile and ferroptosis has recently been described in SCLC.[13] We applied ssGSEA method to verify this association in our cohort, using a ferroptosis signature filed in the MSigDB database. As expected, the most non-NE C-SCLCs, CoSQC, highly expressed a ferroptosis signature (Figure 3E).

We thus evaluated the expression levels of the main sensitivity and resistance markers described for ferroptosis,[13,33,34] selecting genes with a high differential expression (fold change >2 in absolute value) among our three C-SCLC groups (Figure 3F). High levels of *ASCL1* and *AIFM2* and low levels of *REST* defined a ferroptosis-resistant profile in CoLCNEC, whereas a ferroptosis-sensitive profile was observed in CoSQC and CoADC, with high levels of *ACSL4*, *ELOVL5* and *SLC7A11*.

Comparison of combined-SCLC profiles with histologically pure lung cancers

To understand the relationship of C-SCLCs with lung cancers having a single component, we compared expression data of the 13 C-SCLCs with those of 57 single-component lung cancers (17 ADCs, 20 SQCs, 11 LCNECs, 9 SCLCs) and 8 non-neoplastic lung samples. For clustering purposes, we selected a set of genes whose expression levels were both high on average and highly variable. The 0.75 quantile of mean normalized expression was used as a cut-off value for highly expressed genes, while the 0.9 quantile of standard deviation was used as a threshold to identify genes with highly variable expression. This filtering yielded a final set of 1827 genes (Supplementary Figure S3A, available at https://doi.org/10.1016/j.esmoop.2021.100308), which were used for clustering analysis.

We first used the NbClust package to estimate the best number of clusters, which resulted to be 6 ($k = 6$; Supplementary Figure S3B, available at https://doi.org/10.1016/j.esmoop.2021.100308). Secondly, we applied the hybrid hierarchical k-means approach to the 1827 genes setting the number of clusters to 6, to perform principal component analysis (PCA) (Supplementary Figure S3C, available at https://doi.org/10.1016/j.esmoop.2021.100308) and construct a dendrogram showing the relationships between samples (Supplementary Figure S3D, available at https://doi.org/10.1016/j.esmoop.2021.100308). To verify the resulting associations between samples, unsupervised consensus clustering was carried out using ConsensusClusterPlus. The resulting consensus matrix confirmed the associations obtained by PCA and dendrogram (Supplementary Figure S3E, available at https://doi.org/10.1016/j.esmoop.2021.100308).

The six clusters obtained were named, based on the most represented histological type, 'N' (normal), 'ADC' (adenocarcinoma), 'SQC' (squamous), 'SCLC' (small cell), 'LCNEC' (large cell) and 'C-SCLC' (combined) (Figure 4A). Indeed, cluster 'N' included all eight non-neoplastic lung samples; cluster 'ADC' included 18 samples (17 ADCs and 1 CoADC); cluster 'SQC' included 23 samples (20 SQCs and all 3 CoSQCs); cluster 'SCLC' included 11 samples (9 SCLCs and 2 LCNECs); cluster 'LCNEC' included 10 samples (9 LCNECs and 1 CoLCNEC); cluster 'Combined-SCLC' included 8 samples (4 CoADCs and 4 CoLCNECs).

Of note, the C-SCLC cluster included 4/5 CoLNECs and 4/5 CoADCs, all with a small-cell component variable between 30% and 60%. One CoADC with a small-cell component of 10% clustered with pure ADCs and one CoLCNEC with a small-cell component of 50% clustered with pure LCNECs. On the contrary, all three CoSQCs clustered with pure SQC independently from the proportion of small-cell component that was 15%, 45% and 60%.

Survival analysis

Follow-up was available for all cases. The median follow-up for all C-SCLC was 5 months and the mean was 25.6 months (range: 1-120 months). Six (46.1%) subjects died of disease. No difference was observed in survival based on non-SCLC component, smoking status or gender. No molecular alteration showed prognostic significance. Supplementary Figure S4, available at https://doi.org/10.1016/j.esmoop.2021.100308, shows disease-specific survival curves,

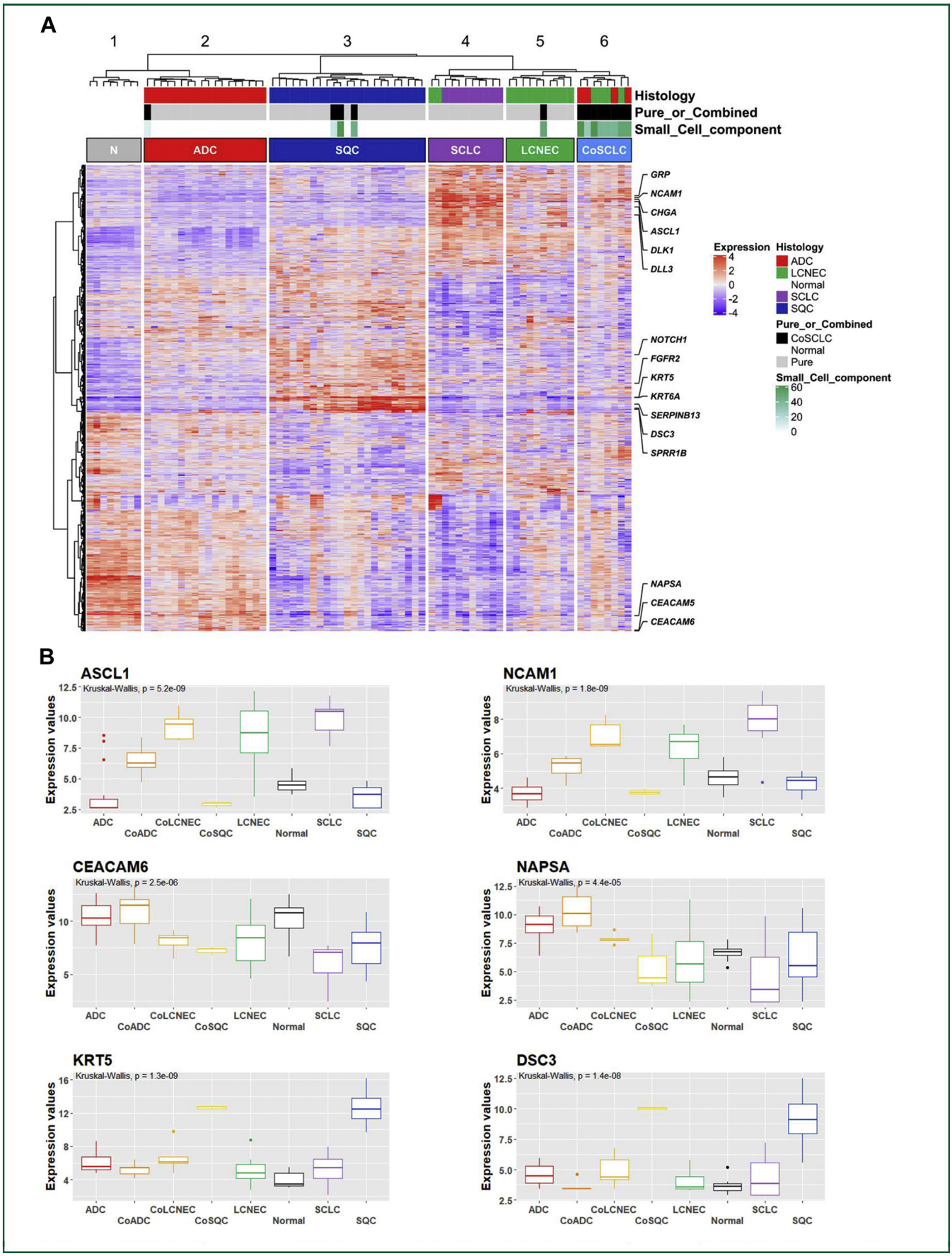

Figure 4. Clustering of 13 combined-SCLC and 57 lung cancers with pure histology.
1827 genes with high and highly variable expression were selected as input for the analysis and the number of clusters was set to 6. (A) In the heatmap, tumor samples are arranged in columns, and annotated for the histological subtype (Histology), for their pure or combined histology status (Pure_or_Combined) and for the proportion

demonstrating a distinct survival profile between C-SCLCs (as a group) and their 'pure' counterparts ($P = 0.0006$).

DISCUSSION

The present study included 13 surgically resected C-SCLCs comprising 5 CoLCNECs, 5 CoADCs and 3 CoSQCs with an SCLC component ranging from 30% to 60% in 11 cases, 10% in 1 CoADC and 15% in 1 CoSQC.

The genetic landscape of C-SCLC overlapped that of SCLC for the presence of *TP53* and/or *RB1* alterations, which were associated with other genetic lesions that correlated with the non-SCLC component. Potentially targetable gene alterations including *KRAS* G12C mutations,[27] *PIK3CA* mutations and an amplified *EGFR* mutation[28,29] were mainly represented in CoADCs and in part in CoSQCs; these findings call for routine testing of CoADC and CoSQC for potentially targetable mutations.

Genetic alterations identified also provide information about origin of C-SCLC. Recent studies proposed two mechanisms: molecular heterogeneity and transdifferentiation. Heterogeneity is supported by GEMMs.[8] Transdifferentiation is sustained by the switch of ADC into SCLC as a mechanism of acquired resistance to therapy.[10,11] Both these mechanisms are supported by selected cases in our cohort. In fact, two CoSQC cases showed mutations in *PTEN* and alteration of *TP53*/*RB1*, supporting C-SCLC as a result of tumor heterogeneity arising from a genetic background similar to that of *Rb/p53/pten* triple knockout models.[8] A third case (#367) is instead compatible with transdifferentiation, as it was a recurrence in a patient with a previous ADC treated with chemotherapy and presenting an *EGFR* gene affected by amplification and two mutations (L858R and T790M), which are typical features of lung ADC.

Comparison of the three subtypes by differential expression analysis highlighted that CoLCNECs were very well separated from the other two subtypes by the NE genes *ASCL1* and *NCAM1*. Overexpression of *NAPSA* and *CEACAM6* was the best marker distinguishing CoADC from the other C-SCLCs. Other distinctive markers of CoADC included *SPINK1*, *CNTD2* and *FGL1*, which have been associated with increase in cell growth, migration and metastasis in lung ADC and other tumor types.[35-37] Finally, the CoSQC subtype was distinguished by overexpression of *DSC3*,[38,39] Keratin family genes, *FGFR2* and *WNT5A*; the latter two may have therapeutic implications as FGFR and WNT inhibitors are currently in clinical development.[40-42]

Comparison of C-SCLC with pure histology lung cancers showed that CoSQC transcriptional setup was overlapping that of pure SQC regardless of the proportion of small-cell component that was 60%, 45% and 15%, respectively, while CoLCNEC and CoADC constituted a standalone group of NE tumors. These results contrast with the previous hypothesis of Murase et al., who suggested a closer similarity of CoSQCs to SCLCs than to CoADCs, based on their analysis of six C-SCLC cases by immunohistochemistry, *TP53* sequence and LOH analysis for chromosome 3p.[43] Moreover, it must be noted that the C-SCLCs analyzed by Murase and colleagues were all CoADCs, three of which also presented a squamous component. By contrast, an adeno-like component seems to have a weaker impact on the overall transcriptomic profile but may retain clinically relevant targetable drivers.

The expression of non-NE markers in CoADC and CoSQC led us to investigate the NE lineage of each subtype using a well-established molecular signature. The GSEA analysis in fact revealed that CoLCNECs were clearly NE, CoSQCs were strongly non-NE and CoADC exhibited a heterogeneous phenotype. The dichotomy NE versus non-NE has been reported in pure SCLC. Rudin et al.[12] recently proposed a new taxonomy for SCLC including four subtypes. Two of these are defined as NE, comprising SCLC-A featuring high expression of *ASCL1* and SCLC-N showing high expression of *NEUROD1*. The other two are non-NE, SCLC-P and SCLC-Y, defined by the high expression levels of *POU2F3* and *YAP1*, respectively. Our data suggest that CoLCNECs belong to 'NE SCLC-A', while CoSQCs fit in the 'non-NE SCLC-Y' group.

Although current European[2] and North American[3] guidelines indicate the same therapeutic strategy for cancers containing any SCLC component, our findings that a squamous cell component may be molecularly dominant on the SCLC component and that typical ADC drivers (*EGFR, KRAS, PIK3CA*) are found in a proportion of CoADC cases challenge this concept. Indeed, SQC-directed treatments might be more appropriate for CoSQC, while combined targeting of molecular drivers and standard chemotherapy might be of benefit in CoADC. Moreover, recent data produced on both SCLC human cell lines and mouse models suggest an increase in tumor sensitivity to ferroptosis in non-NE SCLC, while NE SCLC appears to be addicted to a TRX-mediated anti-oxidation pathway.[13] Ferroptosis is a recently discovered type of regulated cell death, usually accompanied by a large amount of iron accumulation and lipid peroxidation. Its induction has been recently explored as an alternative mean of reaching cancer cells' death.[44] Markers of sensitivity to ferroptosis include enhanced levels of *ACSL4*, *GNPAT*, *LPCAT3* and *SLC7A11*, which were indeed enriched in our CoSQC samples. Depletion or inhibition of SLC7A11 through the small molecule erastin was shown to induce ferroptotic death of lung ADC cells, representing a novel possible therapeutic

of small cell component (small_cell_component). Six main clusters are evident: normal lung (N), pure adenocarcinoma (ADC), pure squamous cell carcinoma (SQC), pure small cell lung cancer (SCLC), pure large cell neuroendocrine carcinomas (LCNEC) and combined-small cell lung cancer (Co-SCLC). (B) Expression values for the combined LCNEC (CoLCNEC) markers *ASCL1* and *NCAM1*, combined ADC (CoADC) markers *CEACAM6* and *NAPSA* (Napsin-A) and combined SQC markers *KRT5* and *DSC3* and comparison to other lung tumors (box plots: median and interquartile range, whiskers: min—max values).

ADC, adenocarcinoma; CoADC, small-cell lung cancer combined with adenocarcinoma; CoLCNEC, small-cell lung cancer combined with large-cell neuroendocrine carcinoma; CoSQC, small-cell lung cancer combined with squamous cell carcinoma; LCNEC, large-cell neuroendocrine carcinoma; NAPSA, napsin A; SCLC, small-cell lung carcinoma; SQC, squamous cell carcinoma.

option in these cancers.[45] Interestingly, enrichment in non-NE phenotype and sensitization to ferroptosis have been observed in SCLC upon immunotherapy or radiotherapy,[46,47] suggesting the combined induction of ferroptosis and immune checkpoint blockade as a possible therapeutic strategy, with further improvement by inhibition of the thioredoxin pathway also suggested.[13] Furthermore, switching between non-NE and NE phenotypes has also been described in preclinical SCLC models upon selective inhibition of specific redox pathways,[13] further suggesting a high degree of plasticity in SCLC subtypes. Therefore, a more precise, molecular classification of C-SCLC (and SCLC itself) might be important for inclusion in specific clinical trials.

The main limitation of this study is the small number of cases and the impossibility to analyze the different neoplastic components separately, due to their admixture. However, this limitation is also reflected in the other three studies published on C-SCLC.[5-7] It is difficult to quantify the prevalence of C-SCLC, as almost all SCLCs are not resected, and their diagnosis is made on small biopsies or cytological material. For these reasons, several cases of C-SCLC may never be diagnosed, leading to a potential underestimation of their incidence.

In conclusion, our study is the first focused on the genomic and transcriptomic profiles of C-SCLC that highlighted their features, offering new insights for their classification in the context of lung tumors. Our data suggest that CoSQCs have an expression profile that places them within the category of squamous cell lung tumors, defining them as SQCs with SCLC features. Conversely, all CoADCs with an SCLC component >10% cluster together with CoLCNECs within the NE-type neoplasms. These data support routine molecular profiling of C-SCLC to search for targetable driver alterations (such as EGFR, KRAS and PIK3CA) and precisely classify them according to therapeutically relevant subgroups (e.g. NE versus non-NE), with the aim of achieving a truly precise treatment approach even in the context of these still relatively unknown subtypes of lung cancer.

ACKNOWLEDGEMENTS

This work is dedicated to the memory of Laura Salvaterra, a courageous woman who battled against cancer. This is an invitation to fight cancer every day in her name, even after she has left us.

FUNDING

This work was supported by the Italian Ministry of Health (ERP-2017-23671129 'PMTR-pNET') to MM by Fondazione IRCCS Istituto Nazionale Tumori Milano 5x 1000 Funds—grant 'Integrative molecular analysis of pure and combined lung large-cell neuroendocrine carcinoma (LCNEC)' to MM (no grant number); Associazione Italiana per la Ricerca sul Cancro [grant number AIRC 5x 1000 n.12182] to AS; and EB is supported by Università Cattolica del Sacro Cuore (UCSC-project D1-2020) and AIRC IG [grant number n. 20583]. GC is supported by an FIRC-AIRC fellowship for Italy [grant number 25422].

DISCLOSURE

The authors declare the following conflicts of interest. MM reports personal fees from Ipsen and Novartis. SP reports personal fees from AstraZeneca, Eli-Lilly, AMGEN, BMS, Boehringer Ingelheim, Merck & Co. and Roche; and grants from AstraZeneca and BMS outside the submitted work. EB reports personal fees, non-financial support and other from MSD, AstraZeneca, Pfizer, Eli-Lilly, BMS, Novartis and Roche; and grants from AstraZeneca and Roche, outside the submitted work. MM reports personal fees from Pfizer, MSD, AstraZeneca, EUSA Pharma, Boehringer Ingelheim and Ipsen; and grants from Roche and BMS, outside the submitted work. AS reports personal fees from AstraZeneca, AMGEN, Incyte Biosciences, Tesaro-GSK and MSD, outside the submitted work. The remaining authors have declared no conflicts of interest.

REFERENCES

1. Travis WD, Brambilla E, Burke A, et al. *WHO Classification of Tumours. Thoracic Tumours*. Lyon: International Agency for Research on Cancer; 2021.
2. Dingemans AC, Fruh M, Ardizzoni A, et al. Small-cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2021;32:839-853.
3. National Comprehensive Cancer Network. Small Cell Lung Cancer (version 3.2021). Available at https://www.nccn.org/professionals/physician_gls/pdf/sclc.pdf. Accessed September 4, 2021.
4. Houssaini MS, Damou M, Ismaili N. Advances in the management of non-small cell lung cancer (NSCLC): a new practice changing data from asco 2020 annual meeting. *Cancer Treat Res Commun*. 2020;25:100239.
5. Zhao X, McCutcheon JN, Kallakury B, et al. Combined small cell carcinoma of the lung: is it a single entity? *J Thorac Oncol*. 2018;13:237-245.
6. Zhang J, Zhang L, Luo J, et al. Comprehensive genomic profiling of combined small cell lung cancer. *Transl Lung Cancer Res*. 2021;10:636-650.
7. George J, Walter V, Peifer M, et al. Integrative genomic profiling of large-cell neuroendocrine carcinomas reveals distinct subtypes of high-grade neuroendocrine lung tumors. *Nat Commun*. 2018;9:1048.
8. Gazdar AF, Savage TK, Johnson JE, et al. The comparative pathology of genetically engineered mouse models for neuroendocrine carcinomas of the lung. *J Thorac Oncol*. 2015;10:553-564.
9. Meder L, Buttner R, Odenthal M. Notch signaling triggers the tumor heterogeneity of small cell lung cancer. *J Thorac Dis*. 2017;9:4884-4888.
10. Sequist LV, Waltman BA, Dias-Santagata D, et al. Genotypic and histological evolution of lung cancers acquiring resistance to EGFR inhibitors. *Sci Transl Med*. 2011;3:75ra26.
11. Niederst MJ, Sequist LV, Poirier JT, et al. RB loss in resistant EGFR mutant lung adenocarcinomas that transform to small-cell lung cancer. *Nat Commun*. 2015;6:6377.
12. Rudin CM, Poirier JT, Byers LA, et al. Molecular subtypes of small cell lung cancer: a synthesis of human and mouse model data. *Nat Rev Cancer*. 2019;19:289-297.
13. Bebber CM, Thomas ES, Stroh J, et al. Ferroptosis response segregates small cell lung cancer (SCLC) neuroendocrine subtypes. *Nat Commun*. 2021;12:2048.
14. Edge SB, Byrd DR, Compton CC, et al. *AJCC Cancer Staging Manual*. 7th ed. New York: Springer; 2009.
15. Milione M, Maisonneuve P, Pellegrinelli A, et al. Ki67 proliferative index of the neuroendocrine component drives MANEC prognosis. *Endocr Relat Cancer*. 2018;25:583-593.
16. Alexandrov LB, Kim J, Haradhvala NJ, et al. The repertoire of mutational signatures in human cancer. *Nature*. 2020;578:94-101.

17. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550.
18. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43:e47.
19. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*. 2016;32:2847-2849.
20. Charrad M, Ghazzali N, Boiteau V, Niknafs A. NbClust: an R package for determining the relevant number of clusters in a data set. *J Stat Softw*. 2014;61:1-36.
21. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics*. 2010;26:1572-1573.
22. Bernard C, Harrison R, Yi P, et al. *Novel Hybrid Hierarchical-K-means Clustering Method (H-K-means) for Microarray Analysis. 2005 IEEE Computational Systems Bioinformatics Conference—Workshops (CSBW'05)*. IEEE Computer Society; 2005:105-108. https://doi.org/10.1109/CSBW.2005.98.
23. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102:15545-15550.
24. Liberzon A, Subramanian A, Pinchback R, et al. Molecular signatures database (MSigDB) 3.0. *Bioinformatics*. 2011;27:1739-1740.
25. Luo W, Friedman MS, Shedden K, et al. GAGE: generally applicable gene set enrichment for pathway analysis. *BMC Bioinformatics*. 2009;10:161.
26. *R: a language and environment for statistical computing [computer program]. Vienna, Austria*. 2015.
27. Hong DS, Fakih MG, Strickler JH, et al. KRAS(G12C) Inhibition with sotorasib in advanced solid tumors. *N Engl J Med*. 2020;383:1207-1217.
28. Ramalingam SS, Vansteenkiste J, Planchard D, et al. Overall survival with osimertinib in untreated, EGFR-mutated advanced NSCLC. *N Engl J Med*. 2020;382:41-50.
29. Mok TS, Wu YL, Ahn MJ, et al. Osimertinib or platinum-pemetrexed in EGFR T790M-positive lung cancer. *N Engl J Med*. 2017;376:629-640.
30. Lawrence MS, Stojanov P, Polak P, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature*. 2013;499:214-218.
31. George J, Lim JS, Jang SJ, et al. Comprehensive genomic profiles of small cell lung cancer. *Nature*. 2015;524:47-53.
32. Rudin CM, Durinck S, Stawiski EW, et al. Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer. *Nat Genet*. 2012;44:1111-1116.
33. Lee JY, Nam M, Son HY, et al. Polyunsaturated fatty acid biosynthesis pathway determines ferroptosis sensitivity in gastric cancer. *Proc Natl Acad Sci U S A*. 2020;117:32433-32442.
34. Tang D, Chen X, Kang R, et al. Ferroptosis: molecular mechanisms and health implications. *Cell Res*. 2021;31:107-125.
35. Xu L, Lu C, Huang Y, et al. SPINK1 promotes cell growth and metastasis of lung adenocarcinoma and acts as a novel prognostic biomarker. *BMB Rep*. 2018;51:648-653.
36. Sanchez-Botet A, Gasa L, Quandt E, et al. The atypical cyclin CNTD2 promotes colon cancer cell proliferation and migration. *Sci Rep*. 2018;8:11797.
37. Zhang W, Girard L, Zhang YA, et al. Small cell lung cancer tumors and preclinical models display heterogeneity of neuroendocrine phenotypes. *Transl Lung Cancer Res*. 2018;7:32-49.
38. Monica V, Ceppi P, Righi L, et al. Desmocollin-3: a new marker of squamous differentiation in undifferentiated large-cell carcinoma of the lung. *Mod Pathol*. 2009;22:709-717.
39. Righi L, Graziano P, Fornari A, et al. Immunohistochemical subtyping of nonsmall cell lung cancer not otherwise specified in fine-needle aspiration cytology: a retrospective study of 103 cases with surgical correlation. *Cancer*. 2011;117:3416-3423.
40. Desai A, Adjei AA. FGFR signaling as a target for lung cancer therapy. *J Thorac Oncol*. 2016;11:9-20.
41. Laeremans H, Hackeng TM, van Zandvoort MA, et al. Blocking of frizzled signaling with a homologous peptide fragment of wnt3a/wnt5a reduces infarct expansion and prevents the development of heart failure after myocardial infarction. *Circulation*. 2011;124:1626-1635.
42. Uitterdijk A, Hermans KC, de Wijs-Meijler DP, et al. UM206, a selective Frizzled antagonist, attenuates adverse remodeling after myocardial infarction in swine. *Lab Invest*. 2016;96:168-176.
43. Murase T, Takino H, Shimizu S, et al. Clonality analysis of different histological components in combined small cell and non-small cell carcinoma of the lung. *Hum Pathol*. 2003;34:1178-1184.
44. Li J, Cao F, Yin HL, et al. Ferroptosis: past, present and future. *Cell Death Dis*. 2020;11:88.
45. Hu K, Li K, Lv J, et al. Suppression of the SLC7A11/glutathione axis causes synthetic lethality in KRAS-mutant lung adenocarcinoma. *J Clin Invest*. 2020;130:1752-1766.
46. Wagner AH, Devarakonda S, Skidmore ZL, et al. Recurrent WNT pathway alterations are frequent in relapsed small cell lung cancer. *Nat Commun*. 2018;9:3787.
47. Lang X, Green MD, Wang W, et al. Radiotherapy and immunotherapy promote tumoral lipid oxidation and ferroptosis via synergistic repression of SLC7A11. *Cancer Discov*. 2019;9:1673-1685.