ORIGINAL ARTICLE

Gene Expression Profiling of Lung Atypical Carcinoids and Large Cell Neuroendocrine Carcinomas Identifies Three Transcriptomic Subtypes with Specific Genomic Alterations

Michele Simbolo, PhD,[a] Stefano Barbi, PhD,[a] Matteo Fassan, MD,[b] Andrea Mafficini, PhD,[b] Greta Ali, MD,[c] Caterina Vicentini, PhD,[a,b] Nicola Sperandio, DScTech,[a,b] Vincenzo Corbo, PD,[a] Borislav Rusev, MD,[b] Luca Mastracci, MD,[d] Federica Grillo, MD,[d] Sara Pilotto, MD, PhD,[e] Giuseppe Pelosi, MD,[f] Serena Pelliccioni, MD,[c] Rita T. Lawlor, PhD,[b] Giampaolo Tortora, MD, PhD,[e,g,h] Gabriella Fontanini, MD,[c] Marco Volante, MD,[i] Aldo Scarpa, MD, PhD,[a,b,*] Emilio Bria, MD[g,h]

[a]*Department of Diagnostics and Public Health, Section of Anatomical Pathology, University and Hospital Trust of Verona, Verona, Italy*
[b]*ARC-Net Research Centre, University and Hospital Trust of Verona, Verona, Italy*
[c]*Department of Surgical, Medical, Molecular Pathology and Critical Area, University of Pisa, AOU Pisana, Pisa, Italy*
[d]*Department of Surgical and Diagnostic Sciences, University of Genoa and IRCCS S. Martino-IST University Hospital, Genoa, Italy*
[e]*Department of Medicine, Section of Medical Oncology, University and Hospital Trust of Verona, Verona, Italy*
[f]*Department of Oncology and Hemato-Oncology, University of Milan, and Inter-Hospital Pathology Division, IRCCS MultiMedica, Milan, Italy*
[g]*Comprehensive Cancer Center, Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy*
[h]*Sacred Hearth Catholic University, Rome, Italy*
[i]*Department of Oncology, University of Turin at San Luigi Hospital, Orbassano, Turin, Italy*

Received 11 January 2019; revised 25 March 2019; accepted 6 May 2019
Available online - 11 May 2019

ABSTRACT

Introduction: DNA mutational profiling showed that atypical carcinoids (ACs) share alterations with large cell neuroendocrine carcinomas (LCNECs). Transcriptomic studies suggested that LCNECs are composed of two subtypes, one of which shares molecular anomalies with SCLC. The missing piece of information is the transcriptomic relationship between ACs and LCNECs, as a direct comparison is lacking in the literature.

Methods: Transcriptomic and genomic alterations were investigated by next-generation sequencing in a discovery set of 14 ACs and 14 LCNECs and validated on 21 ACs and 18 LCNECs by using custom gene panels and immunohistochemistry for Men1 and Rb1.

Results: A 58-gene signature distinguished three transcriptional clusters. Cluster 1 comprised 20 LCNECs and one AC harboring concurrent inactivation of tumor protein p53 gene (*TP53*) and retinoblastoma 1 gene (*RB1*) in the absence of menin 1 gene (*MEN1*) mutations; all cases lacked Rb1 nuclear immunostaining. Cluster 3 included 20

***Corresponding author.**

Disclosure: Dr. Bria has received honoraria or speakers' fees from MSD, AstraZeneca, Celgene, Pfizer, Helsinn, Eli Lilly, BMS, Novartis, and Roche, as well as research grants from Italian Association for Cancer Research (Associazione Italiana Ricerca sul Cancro, AIRC), the International Association for the Study of Lung Cancer (IASLC), Lega Italiana per la Lotta contro i Tumori (LILT), Cariverona Foundation (Fondazione Cariverona), AstraZeneca, Roche, and Open Innovation outside the submitted work. Dr. Pilotto reports personal fees from AstraZeneca, Eli Lilly, BMS, Boehringer Ingelheim, Roche, MSD and Istituto Gentili outside the submitted work. Dr. Tortora reports grants and personal fees from Celgene, Novartis, Roche, Incyte, and Merck Serono for serving on advisory boards and as a consultant, as well as grants from AIRC and Fondazione Cariverona outside the submitted work. The remaining authors declare no conflict of interest.

Address for correspondence: Aldo Scarpa, MD, PhD, University of Verona, Polyclinic Gian Battista Rossi, Piazzale Ludovico Antonio Scuro 10, Verona 37134, Italy. E-mail: aldo.scarpa@univr.it

ISSN: 1556-0864

https://doi.org/10.1016/j.jtho.2019.05.003

ACs and four LCNECs lacking *RB1* alterations and having frequent *MEN1* (37.5%) and *TP53* mutations (16.7%); menin nuclear immunostaining was lost in 75% of cases. Cluster 2 included 14 ACs and eight LCNECs showing intermediate features: *TP53*, 40.9%; *MEN1*, 22.7%; and *RB1*, 18.2%. Patients in cluster C1 had a shorter cancer-specific survival than did patients in C2 or C3.

Conclusions: ACs and LCNECs comprise three different and clinically relevant molecular diseases, one AC-enriched group in which *MEN1* inactivation plays a major role, one LCNEC-enriched group whose hallmark is *RB1* inactivation, and one mixed group with intermediate molecular features. These data support a progression of malignancy that may be traced by using combined molecular and immunohistochemical analysis.

Keywords: Lung neuroendocrine tumors; Atypical carcinoid; Large cell neuroendocrine carcinoma; Gene expression profiling; Next-generation sequencing; Transcriptomics

Introduction

The current WHO classification divides lung neuroendocrine tumors (LNETs) into four histological variants: typical carcinoid (TC), atypical carcinoid (AC), large cell neuroendocrine carcinoma (LCNEC), and SCLC.[1] From a clinical standpoint, carcinoids (TCs and ACs) are distinguished from carcinomas (LCNEC and SCLC). TCs are low-grade tumors, with patients having a long life expectancy, and ACs are intermediate-grade tumors with varying clinical behavior. Conversely, LCNECs and SCLCs are both high-grade tumors with a dismal prognosis.[1–3]

The insufficient knowledge of LNET biology limits the comprehension of these tumor subtypes, which to date have been considered as having a separate pathogenesis.[4–7] Recent DNA mutational profiles showed that carcinoids and carcinomas share similar gene alterations, but with different prevalence among the subtypes.[8] Alterations in chromatin remodeling genes are found in all four variants, whereas menin 1 gene (*MEN1*) alterations are found mainly in carcinoids, and inactivation of tumor protein p53 gene (*TP53*) and retinoblastoma 1 gene (*RB1*) is significantly enriched in carcinomas.[5,8–12] The fact that the same gene alterations found in carcinomas are identified in low-grade tumors but at a lower prevalence may suggest the existence of progression of malignancy and the development of secondary high-grade neuroendocrine carcinomas from preexisting carcinoids.[8,13]

Gene expression profiles produce a global picture of cellular function, and it has been shown that the transcriptional phenotypes of lung cancers mimic the WHO classification.[6,14] They may also provide additional stratification within histological subtypes (that is, the potential to identify molecular subgroups within tumors showing similar morphological features).[14,15]

Three gene expression profiling studies of LNETs have been recently published.[9,14,16] Asiedu et al. reported that transcriptomic profiles could distinguish between carcinoids (TCs and ACs) and SCLC[16]; in this study LCNECs were not included. Karlsson et al. analyzed an ample series of large cell lung carcinomas, including LNECs, and observed a clear separation of three transcriptional groups: adenocarcinoma, squamous cell carcinoma, and a third neuroendocrine group comprising SCLC and LCNEC.[14] Moreover, the comparison of LCNEC and SCLC showed that LCNEC exhibited two different transcriptional profiles associated with different *TP53* and *RB1* genes alteration patterns, corresponding to a proposed genetic division of LCNEC into SCLC-like and NSCLC-like cancers.[14] A study of 75 LCNECs by George et al. confirmed the existence of two LCNEC subtypes, one (type II) characterized by the concurrent inactivation of *TP53* and *RB1* (42%) and one (type I) with *TP53* and serine/threonine kinase 11 gene (*STK11*)/kelch like ECH associated protein 1 gene (*KEAP1*) alterations (37%), but clearly showed that LCNECs have no transcriptional relationship with adenocarcinomas and squamous cell carcinomas.[9]

The missing piece of information is the transcriptomic relationship between AC and LCNEC, as a direct comparison, is lacking in the literature.

Materials and Methods

Cases

A cohort of 67 surgically resected LNETs was collected from four Italian institutions (Applied Research on Cancer-Network [ARC-Net] Research Centre-Verona, IRCCS San Martino-Genova, University of Pisa, and AUO Orbassano-University of Turin). All cases were reclassified according to the WHO 2015 criteria[1] and included 35 ACs and 32 LCNECs. Neuroendocrine differentiation was assessed by using immunostaining for chromogranin, synaptophysin, and CD56.[1,17,18] The AC diagnostic criteria included a well-differentiated morphology with between two and 10 mitoses per 2 mm^2 of area and/or presence of focal necrosis.[19,20] LCNECs were diagnosed on the basis of non–small cell cytologic features, including large cell size, low nuclear-to-cytoplasmic ratio, prominent nucleoli or vesicular chromatin, a mitotic rate of more than 10 mitoses per 2 mm^2 (average 60–80 mitoses per 2 mm^2), and more extensive necrosis.[18–20] Tumor stage was assigned according to the seventh edition of the TNM classification

of malignant tumors.[21] None of the patients received preoperative therapy. Samples were divided into two groups: a discovery set including 28 samples (14 ACs and 14 LCNECs) and a validation set of 39 samples (21 ACs and 18 LCNECs).

Ethics

Ethics committee approval (ECA) was obtained at the four institutions: ARC-Net Research Centre-Verona (ECA no. 2173-prot.26775 [1 June 2012]), AUO Orbassano-University of Torino (ECA no. 167/2015-prot. 17975 [October 21, 2015]), IRCCS San Martino-Genova (ECA no. 027/2016LM [16 March 2016]), and University of Pisa [ECA no. 1040/16 [March 31, 2016]).

Mutational, CNV, and Expression Analysis by Next-Generation Sequencing

The details of the experimental procedures are described in the Supplementary Methods. Briefly, nucleic acids were obtained from formalin-fixed paraffin-embedded tissues as reported.[22,23] Sequencing was performed on Ion Torrent platform (Thermo Fisher Scientific). Data analysis including variant calling was done by using Torrent Suite Software, version 5.0 (Termo Fisher Scientific). Unfiltered variants are reported in Supplementary Tables 1 and 2. Filtered variants were annotated by using a custom pipeline based on vcflib (https://github.com/ekg/vcflib), SnpSift,[24] and Variant Effect Predictor.[25] Annotated variants were filtered by using only the canonical transcripts. Only missense, nonsense, frameshift, or splice site variants were retained. Germline variants were removed. Alignments were visually verified with the Integrative Genomics Viewer, version 2.3,[26] to confirm the presence of identified mutations and to exclude sequencing artefacts. The mutational profile[27] of each sample was obtained with the MuSiCa software.[28] Copy number variation (CNV) was evaluated by using OncoCNV software, version 6.8.[29] The AmpliSeqRNA plugin was used to analyze expression profiling data. Differential analysis was performed by using the DESeq2[30] package for R; an adjusted p value less than 0.05 was considered significant. For gene set analysis, the GSVA[31] package was used.

Immunohistochemistry

Immunostaining was performed by using the Bond Polymer Refine Detection kit (Leica Biosystems) in a BOND-MAX system (Leica Biosystems) on 4-μm-thick formalin-fixed paraffin-embedded sections with the primary antibodies for menin (clone A300-105A [Bethyl Laboratories], dilution 1:1000) and Rb (clone 4H1 [Cell Signaling Technology], dilution 1:250). Appropriate positive and negative controls were run concurrently.

Statistical Analysis

One-way analysis of variance, the Kruskal-Wallis test, the Fisher test with Monte Carlo simulation, and the Fisher exact test were used as appropriate; correction for multiple comparisons was performed according to Benjamini-Hochberg. The Mantel-Cox test was used to compare survival curves. A p value less than 0.05 was considered as significant. Analyses were performed by using Medcalc for Windows, version.18.11 (MedCalc Software), and R software, version 3.5.3.[32]

Results

Study Design

The cohort of 67 cases was divided into a discovery set and a validation set, consisting of 28 and 39 cases, respectively. The study workflow is depicted in Supplementary Figure 1.

The discovery screen was performed on 14 ACs and 14 LCNECs with use of the Ampliseq Transcriptome Human Gene Expression Kit (ThermoFisher), which investigates the expression of 20,815 human genes, and the Ampliseq Comprehensive Cancer Panel (ThermoFisher) for mutational and CNV analysis of 409 genes.

The validation set comprised 21 ACs and 18 LCNECs and was analyzed by targeted sequencing with the use of two custom panels. The first panel was designed to assess the mRNA expression level of 60 genes, including 58 that were differentially expressed in the discovery set plus *MEN1* and *RB1*. The second panel was devised to evaluate DNA alterations in 16 genes: seven genes for mutational analysis only, three genes for CNV analysis only, and six genes for both mutational and CNV analysis.

The expression profiles clustering analysis and the prevalence of mutations and CNVs were finally computed on the entire cohort of 67 cases.

Clinicopathologic Features

Clinicopathologic data are summarized in Supplementary Table 3 and detailed in Supplementary Table 4. The 67 patients had a mean age of 66.2 years and a median clinical follow-up time of 17 months (range 2–100). Of the 67 cases, 33 (49%) were stage I, 24 (36%) stage II, six (9%) were stage III, and four (6%) were stage IV. ACs and LCNECs differed by patient age, patient sex, tumor size, and Ki67 index, whereas there was no statistically significant difference for smoking status and stage (see Supplementary Table 3).

Gene Expression Analysis and Unsupervised Hierarchical Clustering of Discovery Set

Gene expression analysis was performed on the 28 samples of the discovery set and eight nonneoplastic lung samples. Hierarchical unsupervised clustering analysis

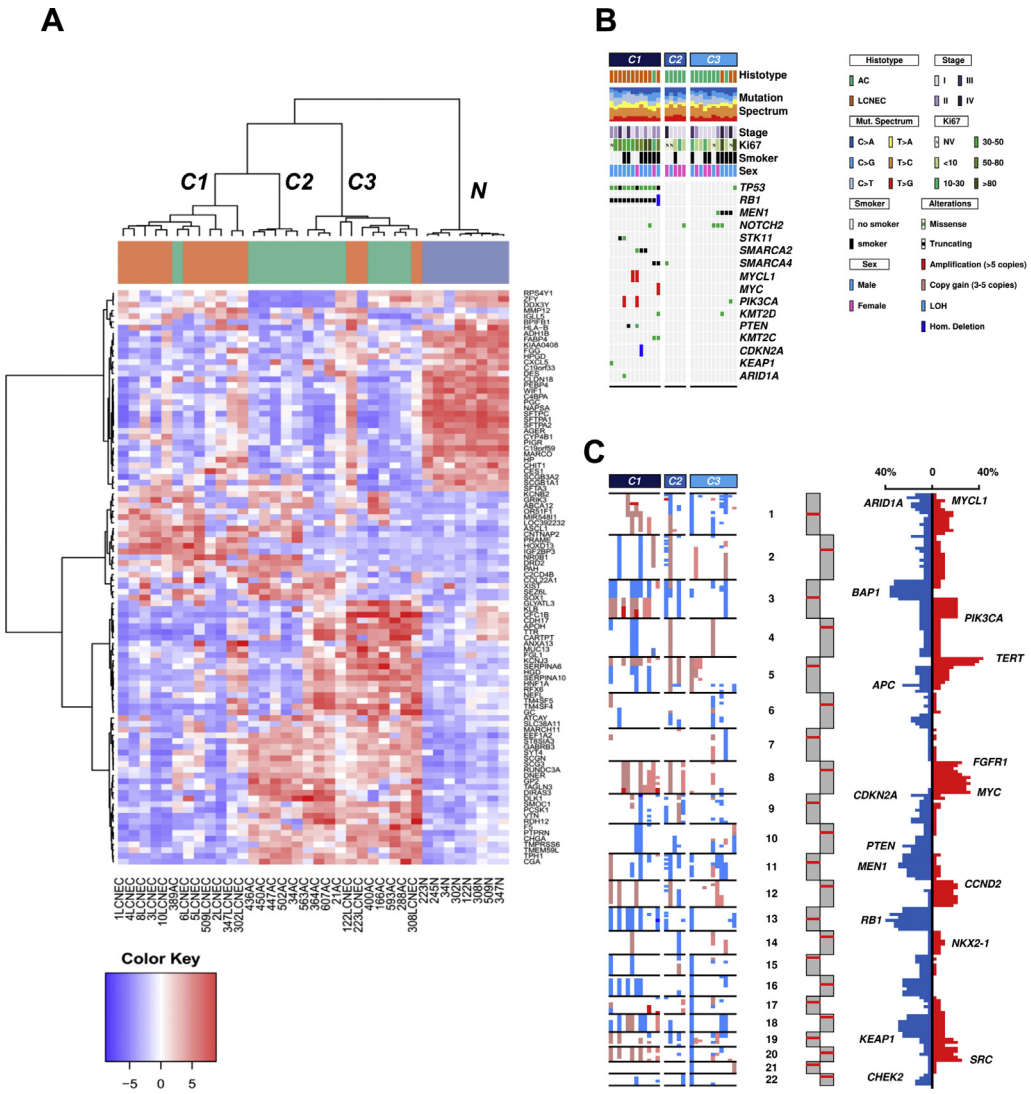

Figure 1. Transcriptome analysis of the discovery set of 28 lung neuroendocrine neoplasms distinguished three molecular clusters of tumors. (*A*) Hierarchical unsupervised clustering of transcriptomes of 14 atypical carcinoids (ACs) (*green*) and 14 large cell neuroendocrine carcinomas (LCNEC) (*orange*) plus eight normal lung (N) (*light blue*) samples with use of the Ward D2 algorithm Tumors were grouped in three separate clusters (C1, C2, and C3) that differ from normal lung samples. Case ID is indicated at the bottom, gene names are indicated on the right. (*B*) Alterations in 16 genes at sequencing analysis; the legend for pathological and molecular alterations is reported in the panel on the right. The mutation spectrum takes into consideration all nonsynonymous variants detected per megabase of exonic sequence, grouped into six classes; stacked bars represent the percentage of each group in each sample. (*C*) Copy number variations within the three clusters (left panel) refer to chromosomes; frequency of copy number variation alterations (*right panel*) where copy gain events are indicated in red and losses in blue. *TP53*, tumor protein p53 gene; *RB1*, retinoblastoma 1 gene; *MEN1*, menin 1 gene; *NOTCH2*, notch 2 gene; *STK11*, serine/threonine kinase 11 gene; *SMARCA2*, SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2 gene; *SMARCA4*, SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4 gene; *MYCL1*, v-myc avian myelocytomatosis viral oncogene lung carcinoma derived homolog gene; *MYC*, v-myc avian myelocytomatosis viral oncogene homolog gene; *PIK3CA*, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha gene; *KMT2D*, lysine methyltransferase 2D gene; *PTEN*, phosphatase and tensin homolog gene; *KMT2C*, lysine methyltransferase 2C gene; *CDKN2A*, cyclin dependent kinase inhibitor 2A gene; *KEAP1*, kelch like ECH associated protein 1 gene; *ARID1A*, AT-rich interaction domain 1A gene; *BAP1*, BRCA1 associated protein 1 gene; *TERT*, telomerase reverse transcriptase; *APC*, APC, WNT signaling pathway regulator; *FGRF1*, gene; *CCND2*, cyclin D2 gene; *NKX2-1*, NK2 homeobox 2 gene; *SRC*, SRC proto-oncogene, non-receptor tyrosine kinase gene; *CHEK2*, checkpoint kinase 2 gene.

using the Ward D2 algorithm identified four clusters (Fig. 1*A*): cluster 1 (C1), which included 12 samples (11 LCNECs and one AC); cluster 2 (C2), which included five samples (all ACs); cluster 3 (C3) , which included 11 samples (eight ACs and three LCNECs); and a fourth group (N), which included all nonneoplastic lung samples.

To verify the robustness of clusters, nonnegative matrix factorization (NMF) based on the expression

profile of the top varying 5000 genes was performed (Supplementary Fig. 2). NMF supported the presence of three LNET classes and one class of normal samples (Supplementary Fig. 2*A*). Comparison of the four NMF classes versus the four hierarchical clustering clusters showed that 32 of 35 cases (91.4%) were consistently assigned to corresponding groups (Supplementary Fig. 2*B*), confirming the reliability of clusters.

To analyze differentially affected pathways, gene set variation analysis was performed[31] aggregating gene expression data according to the "hallmark" gene sets collection from the Molecular Signatures Database (http://software.broadinstitute.org/gsea/msigdb). Gene set variation analysis identified differentially enriched gene sets between tumor clusters (Supplementary Fig. 3). In particular, the gene sets of E2F targets and G2/M checkpoint showed a progressively higher score moving from cluster C3 toward cluster C1 ($p < 0.001$). The E2F targets gene set includes downstream targets of the E2F transcription factors family, which play a major role in G1/S transition.[33] Similarly, genes of the G2M checkpoint set mediate progression through the cell cycle. Thus, coordinated enrichment of both sets is consistent with increased proliferation.

Conversely, the bile acid metabolism gene set, including members involved in peroxisome organization, was increasingly up-regulated moving from C1 to C3 ($p = 0.0217$). A similar trend was also observed in the other gene sets enriched at a false discovery rate of 0.1. Indeed, genes involved in the mitotic spindle and v-myc avian myelocytomatosis viral oncogene homolog gene (*MYC*) targets were enriched in C1, consistent with recurrent *MYC* copy gain in this cluster, whereas gene sets related to bile acid metabolism (fatty acid metabolism, xenobiotic metabolism, and peroxisome) were enriched in C3. Finally, C1 and C2 displayed enrichment in Wnt signalling compared with C3 (Supplementary Table 5 and see also Supplementary Fig. 3).

A set of 58 genes was identified as differentially expressed among the three clusters ($p < 0.05$). The details on differentially expressed genes are reported in Supplementary Table 6 and their distribution in the three LNET clusters in Supplementary Figure 4.

A 58-Gene Signature Identifies Three Expression Profiling Clusters with Distinct Clinicopathologic Features

An RNA targeted sequencing custom panel, designed by using the 58-gene signature identified in the discovery set, was used to analyze the entire series of 67 cases (35 ACs and 32 LCNECs) comprising the 28 of the discovery set and the 39 of the validation set. Additionally, *MEN1* and *RB1* transcripts were included in the custom panel because of their known involvement in AC and LCNEC, respectively,[9,14,34] to correlate their expression levels with the mutational status.

Hierarchical clustering using the 58 genes and Ward D2 algorithm categorized the cases in three clusters (Fig. 2), which were consistent with those obtained by the analysis of 20,815 genes in the discovery set. Clinicopathologic features of the 67 cases are compared across the three clusters in Table 1. Cluster 1 contained 20 LCNECs and one AC. This cluster showed a higher Ki67 index (mean 66% [$p < 0.0001$]) and shorter cancer-specific survival ($p = 0.26$). C3 included 20 ACs and four LCNECs characterized by a lower Ki67 index (mean 21%) and did not reach the median cancer-specific survival (Supplementary Fig. 5). C2 included 14 ACs and eight LCNECs showing intermediate features between those in C1 and C3, with a mean Ki67 index of 36% and a median cancer-specific survival of 47 months. Remarkably, this intermediate C2 cluster was composed of two subclusters (see Fig. 2). C2a included eight LCNECs and three ACs, the former characterized by *TP53* mutations in all eight cases associated with a heterozygous *RB1* mutations in four of them, whereas one of the three ACs had a *MEN1* mutation. C2b included 11 ACs, four of which harbored a *MEN1* mutation and one had a *TP53* mutation. The Ki67 index in C2a (mean 60.0, median 60.0, range 10%–80%) was higher than in C2b (mean 12.0; median 7.0; range 3-35%).

MEN1 and *RB1* expression levels were differentially distributed among the clusters, with significant underexpression of *RB1* in all C1 cases and significant underexpression of *MEN1* in most of the C3 samples and one-third of the C2 samples (Fig. 3*A*).

Discovery Screen of Mutations and CNVs of 14 ACs and 14 LCNECs

Mutational analysis was performed on the discovery set for the coding sequence of 409 genes. Sequencing achieved an average coverage of 698× (198×–1657×) in tumor and 386× (26×–981×) in normal samples (Supplementary Table 7).

Mutations were identified in 22 of the 28 cases. All 12 cases in C1 harbored mutations, whereas two of five cases in C2 and four of 11 in C3 had no mutations. A total of 79 mutations were identified in 36 genes: 56 missense, 10 nonsense, eight frameshift, and five splice site (Supplementary Table 8). *TP53* was the most frequently mutated gene (13 of 28 [46.4%]), followed by *RB1* (11 of 28 [39.3%]), notch 2 gene (*NOTCH2*) (five of 28 [17.9%]), and *MEN1* (four of 28 [14.3%]) (Fig. 1*B*). The mutational spectrum was prevalently characterized by T>C and C>T transitions, with different relative contributions in individual tumors.

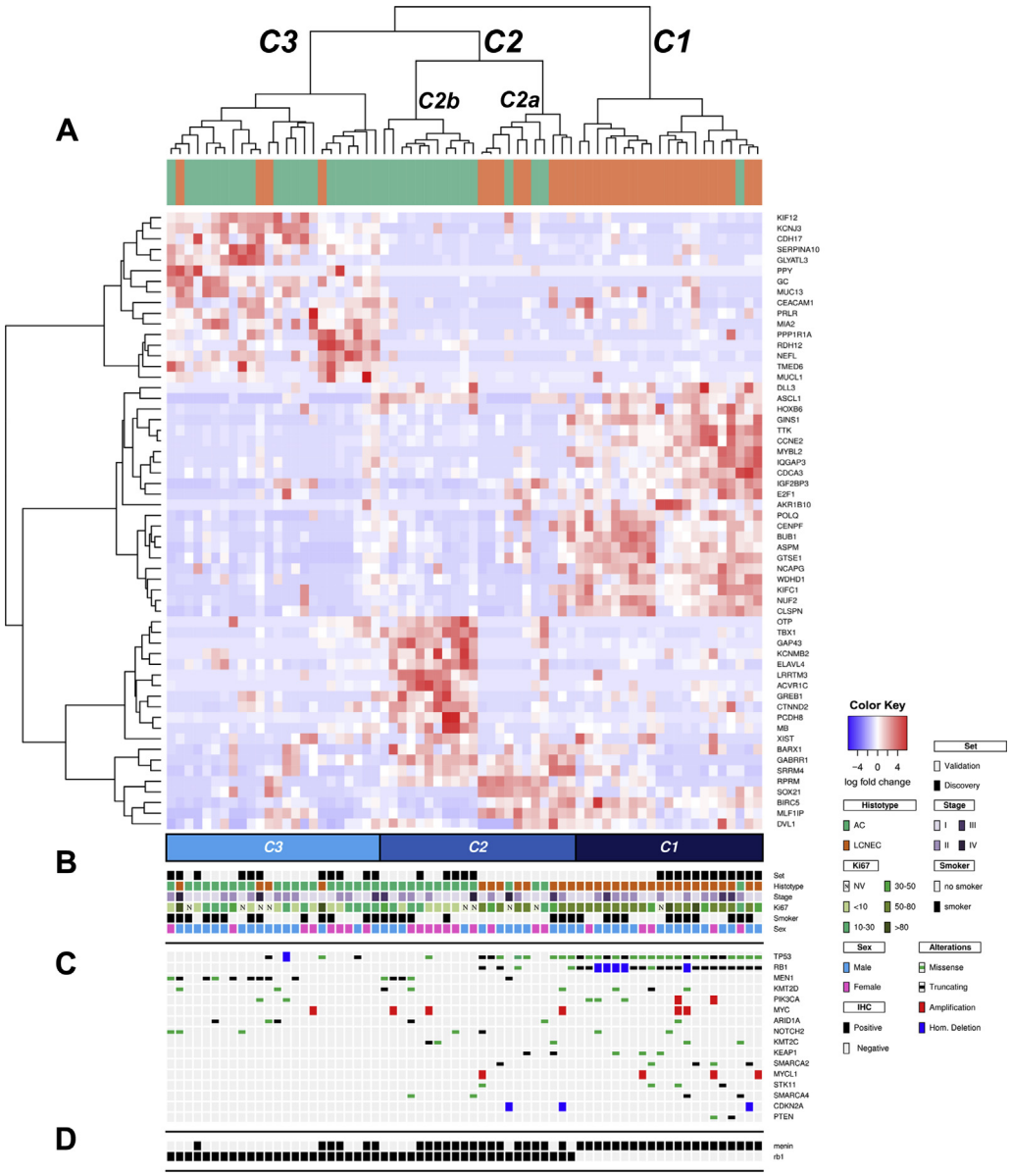

Figure 2. Validation analysis on 67 lung neuroendocrine neoplasms confirms the existence of three clusters of tumors. (*A*) Hierarchical clustering of 35 atypical carcinoids (ACs) and 32 large cell neuroendocrine carcinomas (LCNECs) with use of an RNA custom panel of 58 genes confirmed the presence of the three different molecular subgroups identified by whole transcriptome analysis (see Fig. 1 and Supplementary Fig. 4) and suggests further splitting of cluster 2. (*B*) Clinicopathologic features of the 67 samples; the legend for clinical pathological and molecular alterations is reported in the panel on the right. (*C*) The 16 genes that were altered at sequencing analysis; the legend for alteration type is reported in the panel on the right. (*D*) Immunohistochemical analysis data of menin and Rb.

The most frequent trinucleotide substitution was C[T>C]G, followed by A[T>C]G and A[C>T]G. In particular, all three clusters showed C[T>C]G as the most frequent substitution. The A[T>C]G substitution was the second most frequent in C1 and C2 but not in C3, in which the second most frequent substitution was A[C>T]G. However, no specific substitution was predominant in any cluster.

The CNV status was estimated for all 409 genes by using sequencing data (Supplementary Table 9). The most frequently altered were 20 genes (Supplementary Table 10), including gains in succinate dehydrogenase complex flavoprotein subunit A gene (*SDHA*), RPTOR independent companion of MTOR complex 2 gene (*RICTOR*), telomerase reverse transcriptase gene (*TERT*) (12 of 28 each [42.9%]), and *MYC* (11 of 28 [39.3%]) and losses in BRCA1 associated protein 1 gene (*BAP1*) (12 of 28 [42.9%]), *RB1* (10 of 28 [35.7%]), and *MEN1* (eight of 28 [28.6%]). On the basis of the chromosomal position of each gene, the status of chromosome arms was inferred

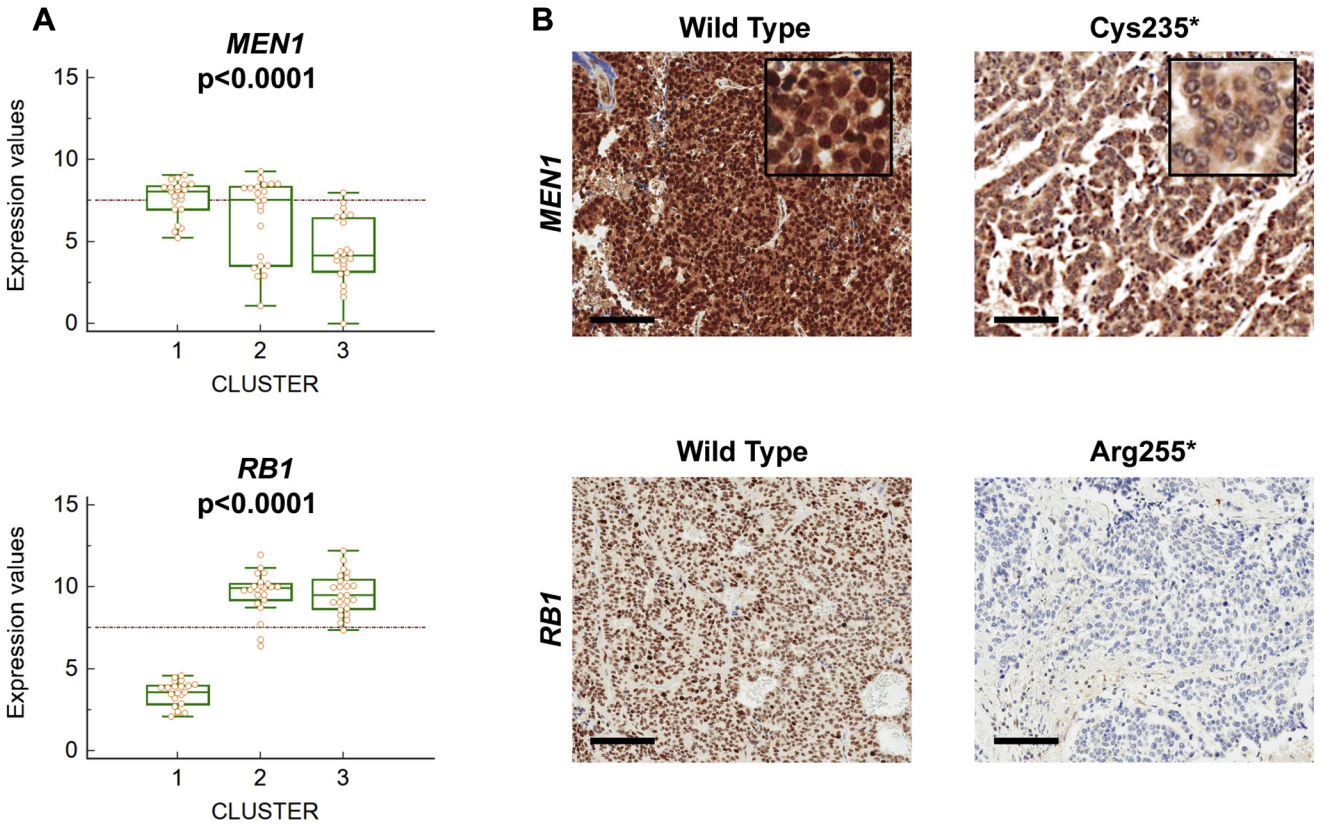

Figure 3. mRNA and immunohistochemical expression analysis of *MEN1* and *RB1*. (*A*) Expression levels (natural logarithm values on *y* axes) of menin 1 gene (*MEN1*) and retinoblastoma 1 gene (*RB1*) genes in the three molecular clusters (on *x* axes) identified by gene expression profiling (see Fig. 2); the dashed red line represents the average gene expression across the entire cohort (reference line). (*B*) Representative images of positive and negative nuclear immunostaining for menin and Rb proteins in two cases harboring a truncating mutation of the corresponding gene, namely Cys235* truncating mutation in *MEN1* gene and Arg255* truncating mutation in *RB1* gene. (scale bars = 100 μm; original magnifications, $\times$10 and $\times$40 [inset]).

(Fig. 1*C* and Supplementary Table 11). Such analysis showed major alterations, namely, losses of chromosome arm 3p (14 of 28 [50.0%]) and whole chromosomes 18 (11 of 28 [39.3%]) and 11 (nine of 28 [32.1%]) and gains in chromosome arms 8q (12 of 28 [42.9%]) and 5p (11 of 28 [39.3%]). Chromosome alterations according to tumor type and expression profile clusters are illustrated in Supplementary Figure 6.

Validation of Mutations and CNVs in 21 ACs and 18 LCNECs by Targeted Sequencing of 16 Genes

The validation set was analyzed by targeted sequencing using a DNA custom panel of 16 genes altered in both the present study and other studies investigating LNETs,[5,9,16,35] investigating the mutational status of seven genes, CNV of three genes, and both alterations of six genes (Supplementary Table 12). Sequencing achieved an average coverage of 1649$\times$ (329$\times$–3636$\times$) in tumor and 468$\times$ (165$\times$–1643$\times$) in normal samples (see Supplementary Table 7). A total of 67 mutations were identified: 38 missense mutations, 13 nonsense mutations, 11 frameshift mutations, three in-frame deletions, and two splice site mutations (see Supplementary Table 8). *TP53* was the most frequently mutated gene (20 of 39 [51.3%]), followed by *MEN1* (10 of 39 [25.6%]) and *RB1* (nine of 39 [23.1%]).

Analysis of CNV status for the nine selected genes identified gains in *MYC* (five of 39 [12.8%]) and losses in *RB1* (four of 39 [10.3%]) as the most frequent alterations. The comparison between discovery and validation set is reported in Supplementary Table 13.

Prevalence of Gene Mutations and Copy Number Alterations in the Three Different Molecular Clusters

Considering the whole series of 67 cases (35 ACs and 32 LCNECs), the three expression profile clusters showed differences in mutation and CNV frequency for specific genes (Fig. 2*C*). Mutations were present in 56 of 67 cases (see Supplementary Table 8). All 21 patients in C1 harbored mutations, whereas three of 22 in C2 and eight of 24 in C3 had no mutations. Different distributions were identified for *TP53* and *RB1* alterations (each $p < 0.0001$), SWI/SNF related, matrix associated, actin

dependent regulator of chromatin, subfamily a, member 2 gene (*SMARCA2*) ($p = 0.043$) and *MEN1* ($p = 0.008$) mutations, and phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha gene (*PIK3CA*) ($p = 0.035$) alterations (Table 2). C1 was characterized by concurrent *TP53* and *RB1* alterations (21 of 21; 100%); also peculiar to this cluster were alterations in *PIK3CA* (five of 21 [23.8%], $p = 0.038$) and *SMARCA2* (four of 21 [19.0%], $p = 0.047$). C2 showed *TP53* alterations as the most frequent event (nine of 22 [40.9%]), followed by *MEN1* (five of 22 [22.7%]) and *RB1* (four of 22 [18.2%]). C3 had *MEN1* mutation as the most frequent event (nine of 24 [37.5%]), followed by *TP53* alterations (four of 24 [16.7%]), whereas no *RB1* anomaly was displayed.

Expression Levels and Immunohistochemistry of Menin and Rb

Menin and Rb mRNA levels were assessed in all 67 samples. For each sample, normalized log-transformed next-generation sequencing counts were compared between mRNA expression clusters by Kruskal-Wallis test. *RB1* and *MEN1* mRNAs showed differential expression between clusters (each $p < 0.0001$) (see Fig. 3*A*). In particular, *MEN1* mRNA level was very low in 24 samples, of which 17 were in C3 (17 of 24; 70.8%) and seven were in C2 (seven of 22; 31.8%), whereas *RB1* mRNA level was very low in all C1 samples (see Fig. 3*A*). All 14 samples harboring *MEN1* mutations showed very low *MEN1* mRNA.

All cases were immunostained for menin and Rb, and nuclear negativity was interpreted as abnormal (representative cases in Fig. 3*B*). Both *MEN1* and *RB1* displayed strong correlation between mRNA levels and protein immunolabeling ($p < 0.00001$ [Supplementary Fig. 7]).

Lack of menin nuclear immunostaining was detected in 24 of 67 cases (35.8%), including the seven samples in C2 and the 17 in C3 that had low mRNA levels, whereas all cases in C1 had positive immunostaining. A direct correlation between presence of mutations, low mRNA level, and loss of protein nuclear immunostaining was observed ($p < 0.0001$). In detail, all 14 cases with *MEN1* mutation, including 11 ACs and three LCNECs, had low mRNA levels and negative nuclear immunostaining; of these 14 cases, the three LCNECs and six ACs were in C3 whereas five ACs were in C2. Interestingly, 10 cases were wild-type for *MEN1* at sequencing; nine ACs belonging to C3 and one LCNEC belonging to C2, showed loss of nuclear menin, suggesting the existence of additional mechanisms of *MEN1* inactivation.[34] All cases in C1 were *MEN1* wild-type at sequencing and had positive menin immunostaining.[36]

Lack of Rb nuclear immunostaining was found in all 21 samples of C1 (20 LCNECs and one AC), with concomitant low mRNA levels and biallelic inactivation of *RB1* owing to homozygous deletion in five cases or heterozygous mutation and loss of the wild-type allele in 16 cases (Supplementary Table 14 and see also Fig. 2). Conversely, all samples in C2 and C3 had positive nuclear immunostaining, including four cases in C2 with *RB1* heterozygous mutations and retention of the wild-type allele.

Discussion

It has been suggested that high-grade neuroendocrine carcinomas may represent a progression of malignancy of preexisting carcinoids.[13] Indeed, recent genomic and transcriptomic data indicate that ACs are hybrid tumors sharing genomic features with both low-grade (TCs) and high-grade (LCNECs and SCLCs) neuroendocrine neoplasms,[8] and that LCNECs may be subdivided in at least two molecular subgroups, one of which shows molecular similarities with SCLCs.[7,9,14] This prompted us to perform a direct comparison of molecular alterations of ACs and LCNECs, which is lacking in literature.

The comparative transcriptomic analysis of ACs and LCNECs reported herein discriminated three transcriptional clusters, defined as C1, C2, and C3, which also showed specific genomic patterns (Fig. 4).

C3 was an AC-enriched cluster that included 20 ACs (83.3%) and four LCNECs (16.7%). *MEN1* mutations were the most frequent events (nine of 24 [37.5%]), followed by *TP53* mutations (16.7%). No case had *RB1* alterations. Interestingly, three of the four LCNECs in this cluster harbored *MEN1* mutations, which are relatively rare in LCNECs, in which they account for 4% of cases.[8,35] That *MEN1* alterations may represent a major event in this cluster is suggested by loss of menin nuclear immunostaining in most cases (18 of 24 [75.0%]), including nine samples harboring a *MEN1* mutation and nine that were determined to be wild-type, suggesting the existence of additional inactivation mechanisms.[34] Indeed, the immunohistochemical findings in LNETs of the present series parallel those in pancreatic neuroendocrine neoplasms, in which of 80% of cases lacking menin nuclear immunostaining, only 30% revealed *MEN1* mutations by sequencing analysis,[36] and subsequent whole-genome analysis showed gross chromosome 11 alterations in many cases.[37]

C1 was an LCNEC-enriched cluster consisting of 20 LCNECs and one AC. All cases in this cluster had concurrent inactivation of *TP53* and *RB1* genes (100%) and lacked *MEN1* mutations. All samples showed low *RB1* mRNA and loss of nuclear immunostaining for Rb protein. Other frequent alterations in C1 were found in

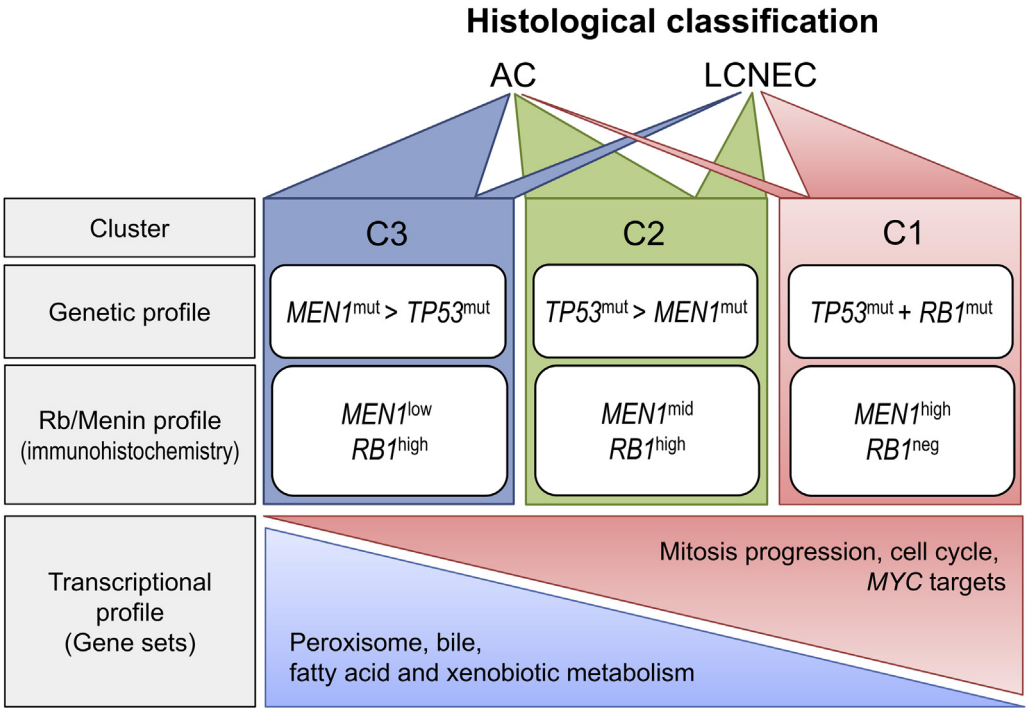

Figure 4. Outline of main differences between the three clusters of the lung neuroendocrine tumors. Cluster 3 is mainly composed of atypical carcinoids (ACs), is Rb proficient, features frequent menin 1 gene (*MEN1*) and rare retinoblastoma 1 gene (*TP53*) mutations, and displays high levels of oxidative metabolism. Cluster 1 is composed almost exclusively of large cell neuroendocrine carcinomas (LCNECs), has Rb loss of expression, always features *TP53* and *RB1* inactivation, and displays high levels of cell cycle deregulation. Cluster 2 has an intermediate transcriptional profile compared with C1 and C3 and is composed of both ACs (one-third of which bear a *MEN1* mutation) and LCNECs (featuring *TP53* mutation but retaining Rb expression). profile: neg, negative staining in all cases; low, negative staining in most cases; mid, similar proportion of positive and negative cases; high, positive staining in most cases; mut, mutation; *MYC*, v-myc avian myelocytomatosis viral oncogene homolog gene.

SMARCA2 (19.0%), *STK11, KEAP1,* and v-myc avian myelocytomatosis viral oncogene lung carcinoma derived homolog gene (*MYCL1*) (each 14.3%) genes.

C2 showed intermediate features between C1 and C3. This group included 14 ACs (64%) and eight LCNECs (36%), in which mutations in *TP53* was the most frequent event (41%), followed by mutations in *MEN1* (23%) and *RB1* (18%) genes. Menin nuclear immunostaining was lost in six samples, comprising the five harboring *MEN1* mutations and one *MEN1* wild-type. Interestingly, nuclear immunostaining of Rb was retained in all cases in this cluster, including the four that harbored a mutated allele but retained the normal allele, supporting the central role of the biallelic inactivation of *RB1* in the transcriptomic shift from C2 to C1.

A comparison of our clusters with the subsets of Rekhtman et al.,[35] and George et al.,[9] both comprising only LCNECs, has identified some interesting overlap. Of note, Rekhtman et al. defined those LCNECs harboring *MEN1* mutations as a third minor subset of "carcinoid-like" LCNECs, and these cases were part of our AC-predominant C3.[35] Our C1 LCNEC-enriched cluster with concurrent inactivation of *TP53* and *RB1* genes and lacking *MEN1* mutations coincide with both the SCLC-like LCNEC subtype of Rekhtman et al.[35] and the type II LCNECs of George et al.[9] Interestingly, other frequent alterations found in our C1 were *SMARCA2*, *STK11, KEAP1,* and *MYCL1*. As these genes are frequently altered in NSCLCs, Rekhtman et al. classified cases harboring these alterations and lacking concurrent *RB1/TP53* inactivation as NSCLC-like LCNEC.[35] Similarly, George et al. suggested that these alterations are typical of their type II LCNECs, but clearly reported that they had no transcriptional relationship with NSCLCs.[9] Furthermore, inactivation of *TP53* with retained Rb immunostaining notable to our C2, are characteristics of the NSCLC-like subtype of Rekhtman et al. and the type I of George et al.[9,35] Notably, our finding that these alterations may occur in association with concurrent *RB1/TP53* inactivation is not surprising, as several cases in the series of both Rekhtman et al. and George et al. displayed the same phenomenon.[9,35]

The identification of three clusters of ACs and LCNECs, two of which were enriched for either AC or LCNEC and the third with intermediate features, suggests the existence of a progression of malignancy for a

proportion of ACs into LCNECs. This is further supported by the fact that the intermediate C2 cluster was composed of two subclusters. The first (C2a) was closer to C1 because of both a similarly high Ki67 index (mean 60%) and the presence of LCNECs with *TP53* mutations associated with heterozygous RB1 alterations. The second (C2b), which is closer to C3, is composed entirely of ACs with a mean Ki67 index of 12% and enriched for *MEN1* mutations. The ACs in C2, especially those sharing the C2a subcluster with LCNECs, might overlap with the recently proposed "supracarcinoids," which were identified by Alcala et al. through supervised machine learning of genomic and transcriptomic data.[38] Indeed, supracarcinoids were defined as a subgroup of LNETs with a clear carcinoid histopathologic pattern but with molecular characteristics similar to those of LCNECs.[38] This further supports the hypothesis that carcinoids may evolve into carcinomas by accumulation of genetic anomalies. A recent publication explicitly suggested the existence of an evolution from AC to LCNEC on the basis of clustering of mutations and CNVs.[13] This might be especially true for a fraction of ACs that do not display *MEN1* loss but show *TP53* alterations.

The three clusters described here also differed from a clinicopathologic point of view, including Ki67 proliferation index and cancer-specific survival. C1 had a mean Ki67 index of 66% versus 37% and 21% for C2 and C3, respectively. Follow-up data were available for 56 patients. Patients in C1 had the shortest survival (median 19 months), whereas patients in C2 had a median survival of 47 months and patients in C3 did not reach the median survival during follow-up.

In conclusion, our study shows that ACs and LCNECs comprise three different molecular diseases of potential clinical relevance, one AC-enriched group in which *MEN1* inactivation plays a major role, one LCNEC-enriched group whose hallmark is *RB1* inactivation, and one group with intermediate features. Indeed, it has been reported that carcinoids harboring *MEN1* mutations, loss of heterozygosity, and low mRNA levels had shorter overall survival,[34] whereas the independent poor prognostic role of *RB1* inactivation is common to AC, LCNEC, and SCLC.[8] Molecular profiling with a combined immunohistochemical and mutational analysis using routinely available paraffin-embedded tissues may complement histological examination to provide better diagnostic definition and prognostic stratification of LNETs that would be helpful for their clinical management.

Acknowledgments

This study was funded by the following organizations: Italian Cancer Genome Project, Ministry of University and Research (grant FIRB RBAP10AHJB); Italian Association for Cancer Research (AIRC grants 5x1000 12182 to Dr. Scarpa and Dr. Tortora, IG 19238 to Dr. Volante, and IG 20583 to Dr. Bria); Cariverona Foundation (grant 2015.0872 to Dr. Tortora and Project Antonio Schiavi funding to Dr. Lawlor); and International Association for the Study of Lung Cancer (IASLC) to Dr. Pilotto. The funding agencies had no role in the collection, analysis, and interpretation of data or in the writing of the article. Dr. Simbolo, Dr. Bria, and Dr. Scarpa conceived the study. Dr. Simbolo designed the study and validation experiments. Dr. Vicentini supervised the validation experiments. Dr. Lawlor coordinated patients and sample data management and supervised ethical protocols. Dr. Ali, Dr. Pilotto, Dr. Mastracci, Dr. Grillo, Dr. Fontanini, Dr. Volante, and Dr. Bria collected materials and clinical data. Dr. Ali, Dr. Fassan, Dr. Fontanini, Dr. Volante, Dr. Pelosi, Dr. Scarpa analyzed histopathologic data. Dr. Rusev and Dr. Fassan microdissected samples. Ms. Sperandio and Dr. Pelliccioni extracted and qualified nucleic acids. Dr. Simbolo, Dr. Mafficini, and Dr. Barbi carried out sequencing and performed bioinformatic analysis. Dr. Ali, Dr. Mastracci, and Dr. Vicentini performed the immunohistochemistry analysis. Dr. Simbolo, Dr. Barbi, Dr. Mafficini, Dr. Fassan, Dr. Corbo, Dr. Tortora, Dr. Volante, Dr. Pelosi drafted the article. Dr. Scarpa and Dr. Bria revised and finalized the article. All authors approved the submitted version.

Supplementary Data

Note: To access the supplementary material accompanying this article, visit the online version of the *Journal of Thoracic Oncology* at www.jto.org and at https://doi.org/10.1016/j.jtho.2019.05.003.

References

1. Travis WD, Brambilla E, Burke AP, Marx Z, Nicholson AG, eds. *WHO Classification of Tumours of the Lung, Pleura, Thymus and Heart*. 4th ed. Lyon, France: IARC; 2015.
2. Asamura H, Kameya T, Matsuno Y, et al. Neuroendocrine neoplasms of the lung: a prognostic spectrum. *J Clin Oncol*. 2006;24:70–76.
3. Beasley MB, Thunnissen FB, Brambilla E, et al. Pulmonary atypical carcinoid: predictors of survival in 106 cases. *Hum Pathol*. 2000;31:1255–1265.
4. Swarts DR, Ramaekers FC, Speel EJ. Molecular and cellular biology of neuroendocrine lung tumors: evidence for separate biological entities. *Biochim Biophys Acta*. 2012;1826:255–271.
5. Fernandez-Cuesta L, Peifer M, Lu X, et al. Frequent mutations in chromatin-remodelling genes in pulmonary carcinoids. *Nat Commun*. 2014;5:3518.
6. Clinical Lung Cancer Genome Project (CLCGP), Network Genomic Medicine (NGM). A genomics-based classification of human lung tumors. *Sci Transl Med*. 2013;5:209ra153.
7. Derks JL, Leblay N, Lantuejoul S, Dingemans AC, Speel EM, Fernandex-Cuesta L. New insights into the

molecular characteristics of pulmonary carcinoids and large cell neuroendocrine carcinomas, and the impact on their clinical management. *J Thorac Oncol*. 2018;13:752–766.
8. Simbolo M, Mafficini A, Sikora KO, et al. Lung neuroendocrine tumours: deep sequencing of the four World Health Organization histotypes reveals chromatin-remodelling genes as major players and a prognostic role for TERT, RB1, MEN1 and KMT2D. *J Pathol*. 2017;241:488–500.
9. 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.
10. Umemura S, Mimaki S, Makinoshima H, et al. Therapeutic priority of the PI3K/AKT/mTOR pathway in small cell lung cancers as revealed by a comprehensive genomic analysis. *J Thorac Oncol*. 2014;9:1324–1331.
11. Vollbrecht C, Werner R, Walter RF, et al. Mutational analysis of pulmonary tumours with neuroendocrine features using targeted massive parallel sequencing: a comparison of a neglected tumour group. *Br J Cancer*. 2015;113:1704–1711.
12. Govindan R, Ding L, Griffith M, et al. Genomic landscape of non-small cell lung cancer in smokers and never-smokers. *Cell*. 2012;150:1121–1134.
13. Pelosi G, Bianchi F, Dama E, et al. Most high-grade neuroendocrine tumours of the lung are likely to secondarily develop from pre-existing carcinoids: innovative findings skipping the current pathogenesis paradigm. *Virchows Arch*. 2018;472:567–577.
14. Karlsson A, Brunnstrom H, Micke P, et al. Gene expression profiling of large cell lung cancer links transcriptional phenotypes to the new histological WHO 2015 classification. *J Thorac Oncol*. 2017;12:1257–1267.
15. Bhattacharjee A, Richards WG, Staunton J, et al. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc Natl Acad Sci U S A*. 2001;98:13790–13795.
16. Asiedu MK, Thomas CF Jr, Dong J, et al. Pathways impacted by genomic alterations in pulmonary carcinoid tumors. *Clin Cancer Res*. 2018;24:1691–1704.
17. Travis WD. Advances in neuroendocrine lung tumors. *Ann Oncol*. 2010;21(suppl 7):vii65–vii71.
18. Travis WD, Brambilla E, Nicholson AG, et al. The 2015 World Health Organization classification of lung tumors: impact of genetic, clinical and radiologic advances since the 2004 classification. *J Thorac Oncol*. 2015;10:1243–1260.
19. Travis WD, Rush W, Flieder DB, et al. Survival analysis of 200 pulmonary neuroendocrine tumors with clarification of criteria for atypical carcinoid and its separation from typical carcinoid. *Am J Surg Pathol*. 1998;22:934–944.
20. Rekhtman N. Neuroendocrine tumors of the lung: an update. *Arch Pathol Lab Med*. 2010;134:1628–1638.
21. Vallieres E, Shepherd FA, Crowley J, et al. The IASLC Lung Cancer Staging Project: proposals regarding the relevance of TNM in the pathologic staging of small cell lung cancer in the forthcoming (seventh) edition of the TNM classification for lung cancer. *J Thorac Oncol*. 2009;4:1049–1059.
22. Simbolo M, Gottardi M, Corbo V, et al. DNA qualification workflow for next generation sequencing of histopathological samples. *PloS One*. 2013;8:e62692.
23. Zamo A, Bertolaso A, van Raaij AW, et al. Application of microfluidic technology to the BIOMED-2 protocol for detection of B-cell clonality. *J Mol Diagn*. 2012;14:30–37.
24. Cingolani P, Patel VM, Coon M, et al. Using Drosophila melanogaster as a model for genotoxic chemical mutational studies with a new program, SnpSift. *Front Genet*. 2012;3:35.
25. McLaren W, Pritchard B, Rios D, et al. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics*. 2010;26:2069–2070.
26. Robinson JT, Thorvaldsdottir H, Winckler W, Zehir A, Mesirov JP. Integrative genomics viewer. *Nat Biotechnol*. 2011;29:24–26.
27. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. *Nature*. 2013;500:415–421.
28. Diaz-Gay M, Vila-Casadesus M, Franch-Exposito S, Hernández-Ilián E, Lozano JJ, Castellvi-Bel S. Mutational Signatures in Cancer (MuSiCa): a web application to implement mutational signatures analysis in cancer samples. *BMC Bioinformatics*. 2018;19:224.
29. Boeva V, Popova T, Lienard M, et al. Multi-factor data normalization enables the detection of copy number aberrations in amplicon sequencing data. *Bioinformatics*. 2014;30:3443–3450.
30. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550.
31. Hanzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics*. 2013;14:7.
32. R: a language and environment for statistical computing [computer program]. Vienna, Austria: 2015.
33. Chen HZ, Tsai SY, Leone G. Emerging roles of E2Fs in cancer: an exit from cell cycle control. *Nat Rev Cancer*. 2009;9:785–797.
34. Swarts DR, Scarpa A, Corbo V, et al. MEN1 gene mutation and reduced expression are associated with poor prognosis in pulmonary carcinoids. *J Clin Endocrinol Metab*. 2014;99:E374–E378.
35. Rekhtman N, Pietanza MC, Hellmann M, et al. Next-generation sequencing of pulmonary large cell neuroendocrine carcinoma reveals small cell carcinoma-like and non-small cell carcinoma-like subsets. *Clin Cancer Res*. 2016;22:3618–3629.
36. Corbo V, Dalai I, Scardoni M, et al. MEN1 in pancreatic endocrine tumors: analysis of gene and protein status in 169 sporadic neoplasms reveals alterations in the vast majority of cases. *Endocr Relat Cancer*. 2010;17:771–783.
37. Scarpa A, Chang DK, Nones K, et al. Whole-genome landscape of pancreatic neuroendocrine tumours. *Nature*. 2017;543:65–71.
38. Alcala N, Leblay N, Gabriel A, et al. Integrative and comparative genomic analyses identify clinically relevant groups of pulmonary carcinoids and unveil the existence of supra-carcinoids. Paper presented at: 16th Annual European Neuroendocrin Tumor Society Conference. March 6–8, 2019; Barcelona, Spain.