

## RESEARCH ARTICLE

# In-depth proteomic analysis reveals unique subtype-specific signatures in human small-cell lung cancer

Beáta Szeitz<sup>1</sup>  | Zsolt Megyesfalvi<sup>2,3,4</sup> | Nicole Woldmar<sup>5,6</sup> | Zsuzsanna Valkó<sup>2,3</sup> | Anna Schwendenwein<sup>3</sup> | Nándor Bárány<sup>2,3,7</sup> | Sándor Paku<sup>7</sup> | Viktória László<sup>2,3</sup> | Helga Kiss<sup>4,8</sup> | Edina Bugyik<sup>2,7</sup> | Christian Lang<sup>3</sup> | Attila Marcell Szász<sup>2,9</sup> | Luciana Pizzatti<sup>6</sup> | Krisztina Bogos<sup>2</sup> | Mir Alireza Hoda<sup>3</sup> | Konrad Hoetzenecker<sup>3</sup> | György Marko-Varga<sup>5</sup> | Peter Horvatovich<sup>10</sup> | Balázs Döme<sup>2,3,4,11</sup>  | Karin Schelch<sup>12</sup> | Melinda Rezeli<sup>5</sup> 

<sup>1</sup>Division of Oncology, Department of Internal Medicine and Oncology, Semmelweis University, Budapest, Hungary

<sup>2</sup>National Korányi Institute of Pulmonology, Budapest, Hungary

<sup>3</sup>Department of Thoracic Surgery, Medical University of Vienna, Vienna, Austria

<sup>4</sup>Department of Thoracic Surgery, National Institute of Oncology, Semmelweis University, Budapest, Hungary

<sup>5</sup>Division of Clinical Protein Science, & Imaging, Department of Clinical Sciences (Lund) and Department of Biomedical Engineering, Lund University, Lund, Sweden

<sup>6</sup>Laboratory of Molecular Biology and Proteomics of Blood/LADETEC, Institute of Chemistry, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

<sup>7</sup>First Department of Pathology and Experimental Cancer Research, Semmelweis University, Budapest, Hungary

<sup>8</sup>University of Pécs, Pécs, Hungary

<sup>9</sup>Department of Bioinformatics, Semmelweis University, Budapest, Hungary

<sup>10</sup>Department of Analytical Biochemistry, Groningen Research Institute of Pharmacy, University of Groningen, Groningen, The Netherlands

<sup>11</sup>Department of Translational Medicine, Lund University, Lund, Sweden

<sup>12</sup>Center for Cancer Research, Medical University of Vienna, Vienna, Austria

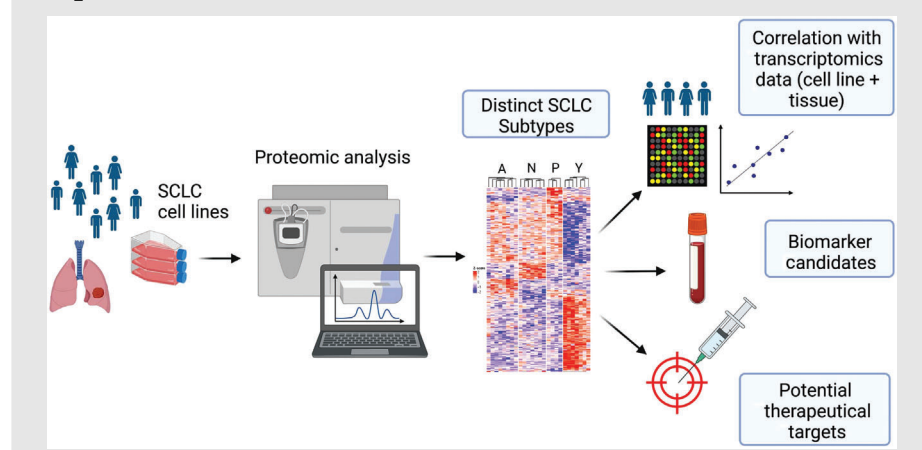
**Correspondence**

Balázs Döme, Department of Thoracic Surgery, Medical University of Vienna, Waehringer Guertel 18-20, 1090 Vienna, Austria.

Email: [balazs.dome@meduniwien.ac.at](mailto:balazs.dome@meduniwien.ac.at)

Melinda Rezeli, Division of Clinical Protein Science & Imaging, Department of Clinical Sciences (Lund) and Department of Biomedical Engineering, Lund University, Lund, Sweden.

Email: [melinda.rezeli@bme.lth.se](mailto:melinda.rezeli@bme.lth.se)

**Graphical Abstract**

1. The first mass spectrometry-based proteomic analysis on small-cell lung cancer (SCLC) cell lines, reporting unique proteomic profiles of its proposed molecular subtypes.
2. SCLC subgroups detected by proteomics are consistent with mRNA-based subtypes.
3. Unique proteomic profiles of SCLC subtypes highlight potential subtype-specific therapeutic vulnerabilities and diagnostic biomarkers.

## RESEARCH ARTICLE

# In-depth proteomic analysis reveals unique subtype-specific signatures in human small-cell lung cancer

Beáta Szeitz<sup>1</sup>  | Zsolt Megyesfalvi<sup>2,3,4</sup> | Nicole Woldmar<sup>5,6</sup> | Zsuzsanna Valkó<sup>2,3</sup> | Anna Schwendenwein<sup>3</sup> | Nándor Bárányi<sup>2,3,7</sup> | Sándor Paku<sup>7</sup> | Viktória László<sup>2,3</sup> | Helga Kiss<sup>4,8</sup> | Edina Bugyik<sup>2,7</sup> | Christian Lang<sup>3</sup> | Attila Marcell Szász<sup>2,9</sup> | Luciana Pizzatti<sup>6</sup> | Krisztina Bogos<sup>2</sup> | Mir Alireza Hoda<sup>3</sup> | Konrad Hoetzenecker<sup>3</sup> | György Marko-Varga<sup>5</sup> | Peter Horvatovich<sup>10</sup> | Balázs Döme<sup>2,3,4,11</sup>  | Karin Schelch<sup>12</sup> | Melinda Rezeli<sup>5</sup> 

<sup>1</sup>Division of Oncology, Department of Internal Medicine and Oncology, Semmelweis University, Budapest, Hungary

<sup>2</sup>National Korányi Institute of Pulmonology, Budapest, Hungary

<sup>3</sup>Department of Thoracic Surgery, Medical University of Vienna, Vienna, Austria

<sup>4</sup>Department of Thoracic Surgery, National Institute of Oncology, Semmelweis University, Budapest, Hungary

<sup>5</sup>Division of Clinical Protein Science, & Imaging, Department of Clinical Sciences (Lund) and Department of Biomedical Engineering, Lund University, Lund, Sweden

<sup>6</sup>Laboratory of Molecular Biology and Proteomics of Blood/LADETEC, Institute of Chemistry, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

<sup>7</sup>First Department of Pathology and Experimental Cancer Research, Semmelweis University, Budapest, Hungary

<sup>8</sup>University of Pécs, Pécs, Hungary

<sup>9</sup>Department of Bioinformatics, Semmelweis University, Budapest, Hungary

<sup>10</sup>Department of Analytical Biochemistry, Groningen Research Institute of Pharmacy, University of Groningen, Groningen, The Netherlands

<sup>11</sup>Department of Translational Medicine, Lund University, Lund, Sweden

<sup>12</sup>Center for Cancer Research, Medical University of Vienna, Vienna, Austria

## Correspondence

Balázs Döme, Department of Thoracic Surgery, Medical University of Vienna, Waehringer Guertel 18-20, 1090 Vienna, Austria.

Email: [balazs.dome@meduniwien.ac.at](mailto:balazs.dome@meduniwien.ac.at)

Melinda Rezeli, Division of Clinical Protein Science & Imaging, Department of Clinical Sciences (Lund) and Department of Biomedical Engineering, Lund University, Lund, Sweden.

Email: [melinda.rezeli@bme.lth.se](mailto:melinda.rezeli@bme.lth.se)

## Abstract

**Background:** Small-cell lung cancer (SCLC) molecular subtypes have been primarily characterized based on the expression pattern of the following key transcription regulators: *ASCL1* (SCLC-A), *NEUROD1* (SCLC-N), *POU2F3* (SCLC-P) and *YAPI* (SCLC-Y). Here, we investigated the proteomic landscape of these molecular subsets with the aim to identify novel subtype-specific proteins of diagnostic and therapeutic relevance.

**Methods:** Pellets and cell media of 26 human SCLC cell lines were subjected to label-free shotgun proteomics for large-scale protein identification

Beáta Szeitz and Zsolt Megyesfalvi contributed equally.

Balázs Döme, Karin Schelch, and Melinda Rezeli share the last authorship.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Clinical and Translational Medicine* published by John Wiley & Sons Australia, Ltd on behalf of Shanghai Institute of Clinical Bioinformatics.

and quantitation, followed by in-depth bioinformatic analyses. Proteomic data were correlated with the cell lines' phenotypic characteristics and with public transcriptomic data of SCLC cell lines and tissues.

**Results:** Our quantitative proteomic data highlighted that four molecular subtypes are clearly distinguishable at the protein level. The cell lines exhibited diverse neuroendocrine and epithelial–mesenchymal characteristics that varied by subtype. A total of 367 proteins were identified in the cell pellet and 34 in the culture media that showed significant up- or downregulation in one subtype, including known druggable proteins and potential blood-based markers. Pathway enrichment analysis and parallel investigation of transcriptomics from SCLC cell lines outlined unique signatures for each subtype, such as upregulated oxidative phosphorylation in SCLC-A, DNA replication in SCLC-N, neurotrophin signalling in SCLC-P and epithelial–mesenchymal transition in SCLC-Y. Importantly, we identified the *YAPI*-driven subtype as the most distinct SCLC subgroup. Using sparse partial least squares discriminant analysis, we identified proteins that clearly distinguish four SCLC subtypes based on their expression pattern, including potential diagnostic markers for SCLC-Y (e.g. GPX8, PKD2 and UFO).

**Conclusions:** We report for the first time, the protein expression differences among SCLC subtypes. By shedding light on potential subtype-specific therapeutic vulnerabilities and diagnostic biomarkers, our results may contribute to a better understanding of SCLC biology and the development of novel therapies.

#### KEYWORDS

diagnostic biomarkers, molecular targets, proteomics, secretome, small-cell lung cancer, subtype, transcriptomics

## 1 | INTRODUCTION

Small-cell lung cancer (SCLC) represents about 13%–15% of all lung cancers and with a 5-year survival rate of less than 7%, it remains one of the most lethal forms of malignant diseases.<sup>1–3</sup> It has a very aggressive course and is characterized by extensive chromosomal rearrangements, high mutation burden and almost universal inactivation of the tumour suppressor genes *TP53* and *RBI*.<sup>1,4</sup> Therefore, the vast majority of SCLC patients are diagnosed with extensive-stage disease when surgery is not feasible, and the treatment options are mostly limited to cytotoxic chemotherapy and radiation.<sup>4</sup> Importantly, targeted therapies for these patients have so far failed, and the success of immunotherapy in non-SCLC has not been reflected in SCLC.<sup>3,5</sup>

Although SCLC has been formerly considered a homogeneous disease with a single morphological type, recent advances in SCLC research have led to the development of subtype-specific classifications primarily based

on neuroendocrine (NE) features and unique molecular profiles.<sup>3,6,7</sup> Accordingly, SCLC can be classified into NE-high and NE-low subtypes based on the expression pattern of key NE markers (i.e. *SYP*, *CHGA*, *NCAM1/CD56* and *GRP*).<sup>1,8</sup> To further elucidate the clinicopathological relevance of NE subtypes, Zhang et al. developed a numeric score to evaluate the degree of NE differentiation in SCLC.<sup>7</sup> By using the top 50 genes most strongly associated with NE differentiation, they successfully separated NE high and low subsets of SCLC tumours and cell lines and showed that these subtypes had widely different growth properties and morphology.<sup>7</sup> Of note, some SCLCs lack NE differentiation and are termed non-NE tumours.<sup>6</sup> Importantly, the NE differentiation also serves as a framework for the recently emerged molecular subtypes.<sup>6</sup> These newly proposed molecular subgroups (SCLC-A, SCLC-N, SCLC-P and SCLC-Y) have been defined based on the expression pattern of transcription factors *ASCL1*, *NEUROD1*, *POU2F3* and *YAPI*, respectively.<sup>4,6</sup> SCLC-P and SCLC-Y show a non-NE phenotype, whereas

SCLC-A and SCLC-N are classified as NE tumours (NE high and low, respectively).<sup>4,6</sup> Moreover, our group also found recently that high POU2F3 expression is associated with improved survival outcomes, whereas elevated ASCL1 expression is an independent negative prognosticator in surgically treated SCLC patients,<sup>9</sup> highlighting the prognostic relevance of SCLC subgroups. So far, the vast majority of SCLC profiling studies have been conducted by analysing transcriptomic data and gene expression profiles of clinical samples and preclinical models, such as patient-derived and circulating tumour cell-derived SCLC xenografts.<sup>6,8,10,11</sup> Although recent immunohistochemistry (IHC) analyses confirmed that subtype-specific markers are indeed detectable in human tumour tissue samples, these studies failed to distinguish a unique *YAP1*-driven subtype.<sup>12</sup> Of note, based on tumour expression data analysis by non-negative matrix factorization, Gay et al. proposed a unique inflamed (SCLC-I) subtype characterized by an inflamed gene signature as an alternative to the *YAP1*-defined subtype.<sup>13</sup> Additionally, not all previously expected RNA-based correlation patterns between subtype markers and NE features could have been validated by IHC either.<sup>12</sup> Therefore, subtype classification based on transcriptional profiling might not be exclusively representative concerning protein expression.

Mass spectrometry (MS)-based proteomics enables a large-scale analysis of complex biological systems, such as cells, tissues or blood plasma. Parallel detection and quantitation of thousands of proteins, including those with lower abundance, is feasible with modern high-resolution mass spectrometers and advanced sample preparation workflows, thus achieving a better understanding of molecular interactions and signalling pathways in cancer.<sup>14,15</sup> The scarcity of available material has so far hindered a comprehensive proteomic examination of SCLC tissues, whereas the proteome of non-SCLC tissues were thoroughly mapped previously.<sup>16,17</sup> To date, proteomic studies considered SCLC as a single entity and compared SCLC samples with normal bronchial epithelial tissues,<sup>18</sup> non-SCLC cell lines<sup>19</sup> or carcinoid tumour tissues.<sup>20</sup> Of note, these studies reported relatively low number of identifiable proteins, ranging from 193 to 1991. A proteomic study of 949 cancer cell lines from 28 tissue types, including 57 SCLC cell lines, was published only recently, in which a total of 8498 proteins were quantified.<sup>21</sup>

In the current work, we conducted a comparative proteomic analysis of SCLC subtypes using cell lines categorized based on the relative expression of four key transcription factors. Our label-free shotgun proteomic approach assessing both the cell pellet (CP) and culture media (CM) resulted in the confident quantification of

nearly 9000 proteins. This comprehensive proteomic evaluation of SCLC cell lines integrated with knowledge from existing transcriptomic datasets can give a clearer definition of SCLC subtypes and might provide insights into their specific features that govern therapy response.

## 2 | MATERIALS AND METHODS

### 2.1 | Cell culture

Human SCLC cell lines were maintained in RPMI-1640 with 10% fetal calf serum (Sigma Chemical Co.), 100 U/ml penicillin and 10 mg/ml streptomycin (Sigma Chemical Co.) at 37°C in a humidified incubator with 5% CO<sub>2</sub>. SCLC cell lines were either purchased from ATCC or kindly provided by our collaborators. All cell lines were regularly checked for mycoplasma contamination using the luminescence-based MycoAlert mycoplasma detection kit (Lonza) with supernatant from cells cultured for >3 days and used within 10 passages after authentication. The cell lines and their key characteristics are listed in Table 1.

### 2.2 | RNA isolation and qPCR

Cells were incubated in T25 flasks until 70% confluency. Total RNA was isolated with TRIzol reagent and reverse transcribed with MMLV reverse transcriptase (Thermo Fisher Scientific) according to the manufacturer's protocol. Gene expression was analysed by quantitative polymerase chain reaction (qPCR) using TaqMan gene expression assays on the Applied Biosystems 7500 Fast Real-Time PCR System (Assay IDs: *ASCL1*: Hs00269932\_m1, *NEUROD1*: Hs00159598\_m1, *POU2F3*: Hs00205009\_m1, *YAP1*: Hs00371735\_m1, *GAPDH*: Hs02786624\_g1; Thermo Fisher Scientific). Gene expression was calculated as 2<sup>-ΔCt</sup> using glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) as reference gene.<sup>22</sup>

### 2.3 | Sample processing for proteomics

CPs and CM from 26 cell lines were processed and subjected to an MS-based proteomic analysis. In brief, the CPs were solubilized with protein extraction buffer (25 mM dithiothreitol [DTT], 10% sodium dodecyl sulphate (SDS), 100 mM triethylammonium bicarbonate [TEAB], pH 8), incubating for 5 min at 95°C with 500 rpm shaking. The volume of the buffer was adjusted to the number of cells in each sample, that is 250 μl of protein extraction buffer

TABLE 1 Cell lines included in the study, and their general characteristics

Cell line ID	Other cell line ID	Subtype	Cell line origin	Chemotherapy	Culture type
DMS153	CRL-2064	SCLC-A	Metastatic	Post-chemo	Semi-adherent
DMS53	CRL-2062	SCLC-A	Lung	Chemo-naïve	Adherent
H146	HTB-173	SCLC-A	Metastatic	Chemo-naïve	Suspension
H1688	CCL-257	SCLC-A	Metastatic	Chemo-naïve	Adherent
H1882	CRL-5903	SCLC-A	Metastatic	N/A	Adherent
H209	HTB-172	SCLC-A	Metastatic	Chemo-naïve	Suspension
H378	CRL-5808	SCLC-A	Lung	Post-chemo	Suspension
SHP77	CRL-2195	SCLC-A	Lung	N/A	Adherent
GLC4	N/A	SCLC-N	Pleural eff.	Chemo-naïve	Suspension
H1694	CRL-5888	SCLC-N	Lung	N/A	Semi-adherent
H2171	CRL-5929	SCLC-N	Pleural eff.	Post-chemo	Suspension
H446	HTB-171	SCLC-N	Pleural eff.	N/A	Adherent
H524	CRL-5831	SCLC-N	Metastatic	Post-chemo	Suspension
H82	HTB-175	SCLC-N	Metastatic	N/A	Semi-adherent
N417	CRL-5809	SCLC-N	Lung	N/A	Suspension
COR-L311	N/A	SCLC-P	Lung	Post-chemo	Suspension
H1048	CRL-5853	SCLC-P	Pleural eff.	N/A	Adherent
H211	CRL-5824	SCLC-P	Lung	Post-chemo	Suspension
H526	CRL-5811	SCLC-P	Metastatic	Chemo-naïve	Suspension
CRL-2066	DMS 114	SCLC-Y	Lung	Chemo-naïve	Adherent
CRL-2177	SW1271	SCLC-Y	Lung	N/A	Adherent
H1341	CRL-5864	SCLC-Y	Metastatic	N/A	Adherent
H196	CRL-5823	SCLC-Y	Pleural eff.	Post-chemo	Adherent
H372	N/A	SCLC-Y	Metastatic	N/A	Adherent
H841	CRL-5845	SCLC-Y	Lung	Post-chemo	Adherent
HLHE	N/A	SCLC-Y	Metastatic	N/A	Adherent

Note: Pleural eff., pleural effusion.

Abbreviations: N/A, not available; SCLC, small-cell lung cancer.

was added to samples containing 5 M cells. Proteins were extracted via a 20 min sonication at 4°C (Bioruptor Plus, Diagenode) with 40 cycles (15 s on/15 s off), followed by a brief centrifugation at 20000 × g at 18°C, discarding the cell debris. Protein determination was performed using a Pierce 660 nm Protein Assay kit (Thermo Scientific), following the manufacturer's instructions.

Additionally, the filtered CM samples were concentrated using spin concentrators (5K 4 ml, Agilent Technologies) to ~100 µl. Protein determination was performed with NanoDrop (DeNovix DS-11 FX +). Prior to digestion, SDS was added to a final concentration of 3%, and 100 mM TEAB was added to increase the pH. Reduction was performed with 10 mM DTT and a 1 h incubation at 37°C.

Protein digestion for both the CP and media (100 µg of proteins per sample) was accomplished using the S-Trap technology (ProtiFi) with few modifications, as previously

described by our group.<sup>23</sup> In brief, the samples were alkylated with 50 mM iodoacetamide and acidified with 1.2% phosphoric acid (final concentration). Then, S-Trap binding buffer was added (90% methanol, 100 mM TEAB) to 7× the final sample volume, and the samples were transferred to the S-Trap 96-well digestion plate. Captured proteins were washed four times with 200 µl S-Trap binding buffer and brief centrifugations (2 min at 1000 × g). Digestion buffer (50 mM TEAB) containing LysC at 1:50 enzyme-to-protein ratio was added on top of the filters, incubating for 2 h at 37°C. Next, digestion buffer containing trypsin at 1:50 enzyme-to-protein ratio was added to the samples, incubating overnight at 37°C. On the following day, the peptides were eluted in three steps, first with 80 µl of digestion buffer, then with 80 µl of 0.2% formic acid, and finally with 80 µl of 50% acetonitrile (ACN) containing 0.2% formic acid. All peptides were dried down

in a vacuum concentrator. Peptide determination was performed for all samples using Pierce Quantitative Colorimetric Peptide Assay kit (Thermo Scientific), following the manufacturer's instructions.

## 2.4 | Nano LC–MS/MS analysis

The nLC–MS/MS analysis was performed on a Q Exactive HF-X mass spectrometer coupled to a Dionex UltiMate 3000 RSLCnano UPLC system (Thermo Scientific), with an EASY-Spray ion source. Peptides from CPs and media were injected in triplicates (1.5 and 1  $\mu\text{g}$ , respectively), utilizing two MS methods detailed later. All samples were loaded onto an Acclaim PepMap 100 C18 (75  $\mu\text{m}$   $\times$  2 cm, 3  $\mu\text{m}$ , 100  $\text{\AA}$ , nanoViper) trap column and separated on an Acclaim PepMap RSLC C18 column (75  $\mu\text{m}$   $\times$  50 cm, 2  $\mu\text{m}$ , 100  $\text{\AA}$ ) (Thermo Scientific) using a flow rate of 300 nl/min, a column temperature of 60°C. A 145-min gradient was applied for separation, using solvents A (0.1% formic acid) and B (0.1% formic acid in 80% ACN), increasing solvent B from 2% to 25% in 115 min, then to 32% in the next 10 min, and to 45% in 7 min. Finally, the gradient increased to 90% solvent B in 8 min, continuing for another 5 min.

Regarding the MS approach, peptides from the CP were analysed with one data-dependent acquisition (DDA) and two data-independent acquisitions (DIA), whereas peptides from the CM were analysed with two DDA and one DIA runs.

The top 20 DDA method was applied with full MS1 scans at  $m/z$  375–1500, resolution of 120000 (at 200  $m/z$ ), target AGC value of  $3 \times 10^6$  and maximum injection time of 100 ms. Fragmentation was done with an NCE of 28, and the isolation window was set to 1.2  $m/z$ . MS2 scans were acquired with a resolution of 15000 (at 200  $m/z$ ), target AGC value of  $1 \times 10^5$ , maximum injection time of 50 ms, ion selection threshold of  $8 \times 10^3$  and dynamic exclusion of 40 s.

For the DIA analysis, a complete acquisition cycle consisted of 3 MS1 full scans, each followed by 18 MS2 DIA scans with variable isolation windows. MS1 full scans were acquired at  $m/z$  375–1455, with a resolution of 120000 (at 200  $m/z$ ), target AGC value of  $3 \times 10^6$  and maximum injection time of 50 ms. The MS2 scans were acquired with a resolution of 30000 (at 200  $m/z$ ), fragmentation with NCE of 28, target AGC value of  $1 \times 10^6$ , automatic maximum injection time, fixed first mass of 200  $m/z$  and the variable isolation windows were 13.0, 16.0, 26.0 and 61.0  $m/z$  (with 27, 13, 8 and 6 loop counts, respectively).

See Table S1 for details on biological and technical repeats. Samples were run in a randomized order, in two distinct batches.

## 2.5 | Proteomic data processing

### 2.5.1 | Database search

Database search was done on Proteome Discoverer v2.4 using SEQUEST HT search engine combined with spectral library search, using the UniProtKB human database (accessed on 15 January 2019) and Proteome Tools spectral libraries. Dynamic modifications included oxidation of methionine and N-terminal acetylation, whereas carbamidomethylation of cysteine was implemented as a static modification. Precursor tolerance was set to 10 ppm and fragment mass tolerance was set to 0.02 Da. Additionally, a maximum of two missed cleavages were allowed, and a false discovery rate (FDR) of 1% was set both on peptide and protein levels. For protein quantitation, the top three average methods were used (i.e. it was calculated based on the average of top three peptide abundances unique to the protein group).

### 2.5.2 | Post-processing

The data-processing steps were performed separately for the CP and CM samples. Both the processing and the statistical tests were performed in R v.4.2.0<sup>24</sup> unless specified otherwise. See [https://github.com/bszeit/SCLC\\_proteomics](https://github.com/bszeit/SCLC_proteomics) for details on all used packages.

The raw protein intensities were  $\log_2$ -transformed, and individual measurements were median normalized by centring their intensities around the global median. Measurements from the same batch were then checked for outliers. After removal of low-quality measurements in the CM data (see Table S1 for details on removed samples), the median intensity of proteins in repeated measurements from the same MS vial and in the same batch was used for further analysis. Proteins with no quantitative values in any samples were removed, leading to the 9570 and 6425 proteins in the CP and CM, respectively. A strong batch effect that was independent from other factors was observed in the data (Figure S1a); therefore, differences in protein expression caused by batch effect were removed using univariate linear regression. Proteins that had no quantitative values in one of the batches were excluded from the batch effect correction. Consequently, the variance explained by the subtype increased (Figure S1b), and the technical and biological replicates from the same cell line showed high similarity according to unsupervised hierarchical clustering results (Figure S1c). All replicates from the same cell line were, therefore, averaged by taking their arithmetic mean.

For statistical analyses, the expression table was filtered for proteins with min. 80% valid values across samples, resulting in 8405 and 5408 proteins in the CP and CM, respectively. Missing value imputation was performed in Perseus v.1.6<sup>25</sup> using imputation based on normal distribution (width = 0.3, down shift = 1.8). The protein intensity histogram per sample, also highlighting the frequency distribution of imputed values, is shown on Figure S1d for the samples with the highest missing value content. Additionally, potential ‘on/off proteins’ were also sought out during the filtering step. The ‘on/off proteins’ were defined as proteins that were present in min. 85% of the samples in one subtype and simultaneously, in max. 15% of the samples in the other subtypes (‘on’), or in max. 15% of the samples in one subtype and simultaneously, in min. 85% of the samples in the other subtypes (‘off’). These proteins, together with the filtered and imputed expression table elements, were also examined as part of the analysis of subtype-specific proteins.

### 2.5.3 | Annotation of secreted, cell-surface, plasma and ‘druggable’ proteins

To map secreted proteins from the list of identified proteins, three secretome databases were utilized: Human Protein Atlas (The Human Protein Atlas v20.1 and Ensembl v92.38, retrieved on 11 April 2021),<sup>26</sup> SPROME (accessed on 7 April 2021)<sup>27</sup> and MetazSecKB (accessed on 7 April 2021).<sup>28</sup> These databases were created using curated experimental evidence and a number of prediction tools that do not completely overlap. All proteins mentioned in at least two databases were considered secreted proteins. In sum, 422 and 636 secreted proteins were detected in the CP and CM, respectively, out of which 295 and 514 were also quantified in min. 80% of the samples.

Additionally, information from the Human Protein Atlas (The Human Protein Atlas v21.1, <https://www.proteinatlas.org/humanproteome/blood+protein>, accessed on 23 July 2022) was used to annotate proteins detectable in human blood plasma by immunoassay, MS or proximity extension assays. This was supplemented with a list of proteins actively secreted into the blood ([https://www.proteinatlas.org/search/sa\\_location%3Asecreted+to+blood](https://www.proteinatlas.org/search/sa_location%3Asecreted+to+blood)). In total, 3076 and 3126 plasma proteins were identified in the CP and CM, respectively, of which 2763 and 2823 were quantified in at least 80% of the samples. Among the proteins actively secreted in the blood, 229 and 334 were found in the CP and CM, respectively, and 172 and 278 proteins were quantified in min. 80% of the samples.

Two databases were downloaded to retrieve a list of cell-surface proteins, namely The Cancer Surfaceome Atlas<sup>29</sup>

(<http://fcgportal.org/TCSA/Download.php>, accessed on 28 July 2022) and the in silico human surfaceome by Bausch-Fluck et al.<sup>30</sup> (<http://wlab.ethz.ch/surfaceome/>, accessed on 28 July 2022). In the CP and CM, we found a total of 682 and 549 cell-surface proteins that are included in both databases, of which 477 and 416 were quantified in at least 80% of the samples.

To annotate ‘druggable’ proteins with subtype specific profiles, we used The Druggable Proteome database available at the Human Protein Atlas website (<https://www.proteinatlas.org/humanproteome/tissue/druggable>, accessed on 22 October 2021), which contains the protein targets of the current Food and Drug Administration (FDA)-approved drugs, including enzymes, transporters, ion channels, and receptors. Functional annotations were collected from the UniProt<sup>31</sup> website (<https://www.uniprot.org/>, Release 2022\_03) for selected proteins, and the list of FDA-approved drugs that directly interact with those proteins as part of their mechanism of action was gathered from the DrugBank database ([www.drugbank.ca](http://www.drugbank.ca)).<sup>32</sup>

## 2.6 | External data retrieval and processing

### 2.6.1 | Transcriptomic data of SCLC tissue samples

The mRNA data of SCLC tissue samples was accessed from cBioPortal<sup>33,34</sup> on 24 May 2021, downloading George et al.’s work.<sup>8</sup> Only samples that were also categorized by Rudin et al.<sup>6</sup> were further analysed (52 samples out of 81). This resulted in 38 SCLC-A, 5 SCLC-N, 7 SCLC-P and 2 SCLC-Y samples. These 52 samples’ Fragments per kilobase million (FPKM) values were utilized to perform the single-sample gene set enrichment analysis (ssGSEA), as well as the log-transformed mRNA Z-scores (where Z-score was calculated compared to the expression distribution of all samples), were used to assess the expression profile differences between selected transcripts across subtypes (SCLC-A/N/P/Y).

### 2.6.2 | Genomic and transcriptomic data of SCLC cell lines

Mutation and RNA-Seq data (gene count data, normalized using RNA-Seq by Expectation-Maximization, i.e. RSEM method) of Cancer Cell Line Encyclopedia (CCLE) cancer cell lines were accessed on 28 June 2022 and 23 August 2021, respectively, from the Cancer Dependency Map website, <https://depmap.org/portal/>.<sup>35</sup> Compared to



our set of cell lines, only five cell lines are not included in the mutation table (GLC4, HLHE, H1882, N417, H372), and two additional cell lines are not present in the RNA sequencing table (H378 and H1688). The RNA-Seq data contains the measurements from 50 SCLC cell lines, which were also categorized into subtypes by Rudin et al.<sup>6</sup> (26 SCLC-A, 12 SCLC-N, 4 SCLC-P and 8 SCLC-Y). The transcriptomic data was further processed via limma R package v.3.46.0,<sup>36</sup> following the standard RNA-Seq data-processing steps. This included normalization factor calculation with default settings, filter of low-expressed genes (only genes with minimum one expression value equal to or higher than 50 should be kept; fulfilled by 9237 genes) and voom transformation with default settings where model matrix included the subtype assignment. This was followed by differential expression analysis described in detail in Section 2.7. Log<sub>2</sub> CPM counts were calculated using the 'cpm' function from edgeR v.3.32.1<sup>37–39</sup> in which prior count was set to 3 (i.e. the average count to be added to each observation to avoid zeros in the dataset during log<sub>2</sub>-transformation). This expression matrix was used to display the differential expression of selected transcripts across subtypes.

### 2.6.3 | Drug sensitivity data of SCLC cell lines

CancerRxGene drug sensitivity data<sup>40</sup> (Release 8.3, available from June 2020) was downloaded from the FTP Server of the Wellcome Sanger Institute ([ftp://ftp.sanger.ac.uk/pub/project/cancerrxgene/releases/current\\_release/](ftp://ftp.sanger.ac.uk/pub/project/cancerrxgene/releases/current_release/)) on 18 July 2022. Both the Genomics of Drug Sensitivity in Cancer 1 (GDSC1) and GDSC2 datasets were investigated. In total, the subtype classification of 38 SCLC cell lines is known,<sup>6</sup> of which 20 were also measured by proteomics in our study (the missing cell lines: DMS153, GLC4, H1882, H372, HLHE, N417). The logarithmic IC<sub>50</sub> values of these cell lines for selected drugs were then compared across the subtypes or were correlated with protein abundance.

## 2.7 | Statistics and bioinformatics

### 2.7.1 | Differential expression analyses

Differential expression analyses for the proteomic data were performed via ANOVA, followed by multiple testing correction (Benjamini–Hochberg, BH) of ANOVA *p*-values and Tukey's honestly significant difference (HSD) post hoc tests. Comparisons were done between culture types (adherent, suspension) and subtypes (SCLC-A, -N, -P, -Y). A protein was considered significant for a comparison if both the ANOVA FDR and the corresponding

pairwise Tukey's HSD test *p*-value were less than 0.05. Subtype-specific proteins were selected based on the following criteria: ANOVA FDR < 0.05, as well as Tukey's HSD test *p* < 0.05 for all comparisons of the subtype of interest against the other subtypes. Differential expression analysis for the CCLE RNA-Seq data was performed via limma by fitting a linear model for the six subtype comparisons, followed by Empirical Bayes smoothing of standard errors. Multiple testing correction with the BH method was applied and FDR < 0.05 was considered significant. For the transcriptomic dataset by George et al., pairwise Wilcoxon tests were performed on the Z-scored values to check gene expression differences between the subtypes. Significance level was set to nominal *p* < 0.05.

Overrepresentation analyses (ORAs) were conducted using clusterProfiler v.3.18.1<sup>41,42</sup> and ReactomePA v.1.34.0.<sup>43</sup> The default list of human genes was used as background.

### 2.7.2 | Neuroendocrine and epithelial–mesenchymal transition scores

Two previously published gene sets were used to introduce a scoring system for the NE and epithelial–mesenchymal transition (EMT) characteristics of cell lines. For NE characteristics, a 50-gene NE signature built for SCLC cell lines described by Zhang et al.<sup>7</sup> was examined. The number of NE (i.e. markers involved in NE differentiation) and non-NE markers present in our CP data was 19 and 17, respectively (Figure S3a). NE score was calculated the following way for each cell line:

NE score = (Mean Z-score of NE markers) – (Mean Z-score of non-NE markers).

Similarly, a gene signature with 22 epithelial and 15 mesenchymal genes published by Kohn et al.<sup>44</sup> was examined, from which 12 and 10 corresponding proteins were quantified in our CP data (Figure S3b). EMT score was calculated the following way for each cell line:

EMT score = (Mean Z-score of mesenchymal markers) – (Mean Z-score of epithelial markers).

### 2.7.3 | Consensus clustering

Samples were grouped in an unsupervised manner using the consensus clustering algorithm, implemented in the ConsensusClusterPlus R package v.1.54.0.<sup>45</sup> The basis of clustering was the samples' global protein expression profile (filtered and imputed) in the CP dataset, which was further restricted to only those proteins that have an SD above 1.25. This list of proteins included protein products of *POU2F3* and *YAPI*, both showing large SD (SD = 1.40

and 2.05, respectively), whereas *NEUROD1* did not show sufficiently large variance ( $SD = 0.50$ ), and *ASCL1* was not quantified in enough samples to be included in the list. The resulting proteome data (including the expression of 1157 proteins without Z-score normalization) were resampled 1000 times via the bootstrap method with a probability of 0.8 for selecting any item (i.e. sample) and any feature (i.e. protein). The bootstrap sample datasets were clustered using the partitioning around medoids method with the Pearson distance and complete linkage, exploring the range of two to seven clusters. The visual inspection of consensus matrices, as well as the silhouette plots for  $K = 4$  and  $K = 5$  revealed that a cleaner separation of the clusters can be achieved with  $K = 4$ ; moreover, the average silhouette width was higher for  $K = 4$ . Based on these observations, and considering the relatively low number of samples, consensus clustering with  $K = 4$  was used as sample grouping. Graphical outputs from the ConsensusClusterPlus R package and silhouette plots for  $K = 4$  and  $K = 5$  are shown in Figure S4a–c. Silhouette information was computed using the ‘silhouette’ function from R package ‘cluster’ v.2.1.1.

## 2.7.4 | Gene set enrichment analysis

Pre-ranked GSEA (pGSEA) was performed via the ‘GSEA’ function of the clusterProfiler R package for all six subtype comparisons. As input, the hallmark,<sup>46</sup> KEGG,<sup>47</sup> Reactome,<sup>48</sup> Gene Ontology biological process<sup>49,50</sup> and oncogenic curated gene sets were used, downloaded from the Molecular Signatures Database (MSigDB) v.7.4,<sup>51,52</sup> as well as a ranked list of genes/proteins (based on the multiplication of the FC and  $-\log_{10}$  ( $p$ -value of pairwise test), calculated for all proteins/genes used in the statistical analyses). The pairwise test was the Tukey’s HSD post hoc test for proteomics, whereas for CCLE RNA-Seq, the  $p$ -values were derived from the moderated  $t$ -test statistic. The BH method was used for  $p$ -value adjustment in pGSEA. To obtain the list of subtype-characteristic gene sets, the following filter was applied on the pGSEA results:  $p < 0.01$  in all comparisons of the subtype of interest against the other subtypes with an unequivocal sign (either positive or negative) normalized enrichment score (NES). Gene sets with a  $p < 0.01$  for at least one comparison are listed in Table S4. The relationship between proteomic and transcriptomic results was examined on these subtype-specific gene sets as follows: if a gene set in one dataset (either proteomic or transcriptomic) was found to be characteristic for a subtype, and the  $p$ -value of the same gene set was less than 0.1 in relevant comparisons of the other dataset, as well as its NES values showed the same direction, the gene set was declared significant in both proteomics and transcriptomics. In contrast, if the gene set in the other dataset did

not meet the previous criteria, the gene set was declared significant in only one dataset. A visual explanation for the filtering steps can be seen in Figure S5 and a summary of the obtained results in Table S4.

ssGSEA<sup>53</sup> was performed on the transcriptomic dataset by George et al., using only transcripts with sum FPKM higher than 50, and only one transcript for each gene was retained (the transcript with the highest sum FPKM value). This resulted in an expression matrix with 13542 transcripts. It should be noted that different transcripts from the same gene may have different biological functions; however, we cannot preserve such depth of information in this analysis. The NES was calculated for subtype-specific gene sets as defined earlier (the list can be accessed in Table S4), with the script available on <https://github.com/broadinstitute/ssGSEA2.0> (accessed on 05 February 2022). The parameters were set as follows: sample.norm.type = ‘rank’, weight = 0.75, statistic = ‘area.under.RES’, output.score.type = ‘NES’, nperm = 1000, min.overlap = 5, and correl.type = ‘z.score’. The gene sets were then filtered for the most representative gene sets (Table S4), and ssGSEA results for these are depicted in Figure S6. For visualization purposes, NES values of each gene set were min.–max. scaled across the samples.

## 2.7.5 | Sparse partial least squares discriminant analysis

The sparse partial least squares discriminant analysis (sPLS-DA) was performed via the mixOmics R package v.6.14.0,<sup>54</sup> following the guidelines provided on its website (<http://mixomics.org/case-studies/splsda-srbct/>, accessed on 27 July 2021). The full CP and CM data (filtered and imputed) were analysed separately. First, a PLS-DA model was fitted with 10 components to evaluate the performance and to select the optimal number of components for the final model. The performance plot showed that five and four components with centroid distance measure are sufficient for good performance (0.169 and 0.476 balanced error rate) in the CP and CM data, respectively (Figure S7a). Then sPLS-DA model was tuned, estimating the classification performance with respect to the number of selected variables in the model. Maximum number of components was set to five and four in CP and CM, respectively, threefold cross-validation was used which was repeated 50 times, and centroid distance measure was selected. As a result, the settings of the final model were the following in CP: (1) optimal number of components is 3, (2) the number of proteins to select on each component are 45, 30 and 7. In the case of CM, the settings of the final model: (1) optimal number of components is 2; (2) the number of proteins to

select on each component are 8 and 15. The sample plots of the first three components in CP (Figure S7b) showed that SCLC-Y is well separated on the first component, SCLC-A and -N are distinguishable on the second component and adding the third component further discriminates SCLC-P from the rest. Similarly, the first two components were visualized for CM, displaying that SCLC-Y is best separated on the first component, whereas SCLC-A/P can be distinguished from SCLC-N on the second component. The classification performance of these final sPLS-DA models were 0.113 and 0.484 with centroid distance for CP and CM, respectively (Figure S7c). Additionally, selected proteins in each component with their loading weight and stability (i.e. ratio of how many times it was selected across the cross-validation runs) were extracted. The selected proteins are shown in Supplementary Figure S8. To select the transcripts showing the best match with proteomic results, transcripts were ranked based on the multiplication of the fold change (FC) and  $-\log_{10}$  (*p*-value of pairwise test).

### 2.7.6 | Visualizations

Individual clustering settings for heat map visualizations were hierarchical clustering, Euclidean distance and complete linkage, unless specified otherwise in the figure legends. Heat maps were plotted using ComplexHeatmap R package v.2.6.2.<sup>55</sup> Principal component analyses (PCAs) and principal variance component analyses (PVCAs) were done on the Z-scored protein expression tables. PCA biplots were drawn using the ggbiplot R package v.0.55 (<https://github.com/vqv/ggbiplot>). Further visualizations were done using R v.4.2.0 and GraphPad Prism v.8 for Windows, GraphPad Software, San Diego, California USA, [www.graphpad.com](http://www.graphpad.com).

## 3 | RESULTS

### 3.1 | Molecular heterogeneity of SCLC cell lines detected via proteomics

We characterized 26 cell lines derived from primary or metastatic human SCLC lesions (Table 1) using label-free proteomic analysis. A total of 10161 proteins were identified and quantified (9570 and 6425 proteins in CP and CM, respectively), and the majority of these proteins were quantified in minimum 80% of the samples (8405 and 5408 proteins in CP and CM, respectively). Altogether, we annotated 699 secreted proteins, 800 cell-surface proteins, 3440 proteins detectable in human blood plasma, of which 367 are actively secreted into the blood, and 289 'druggable'

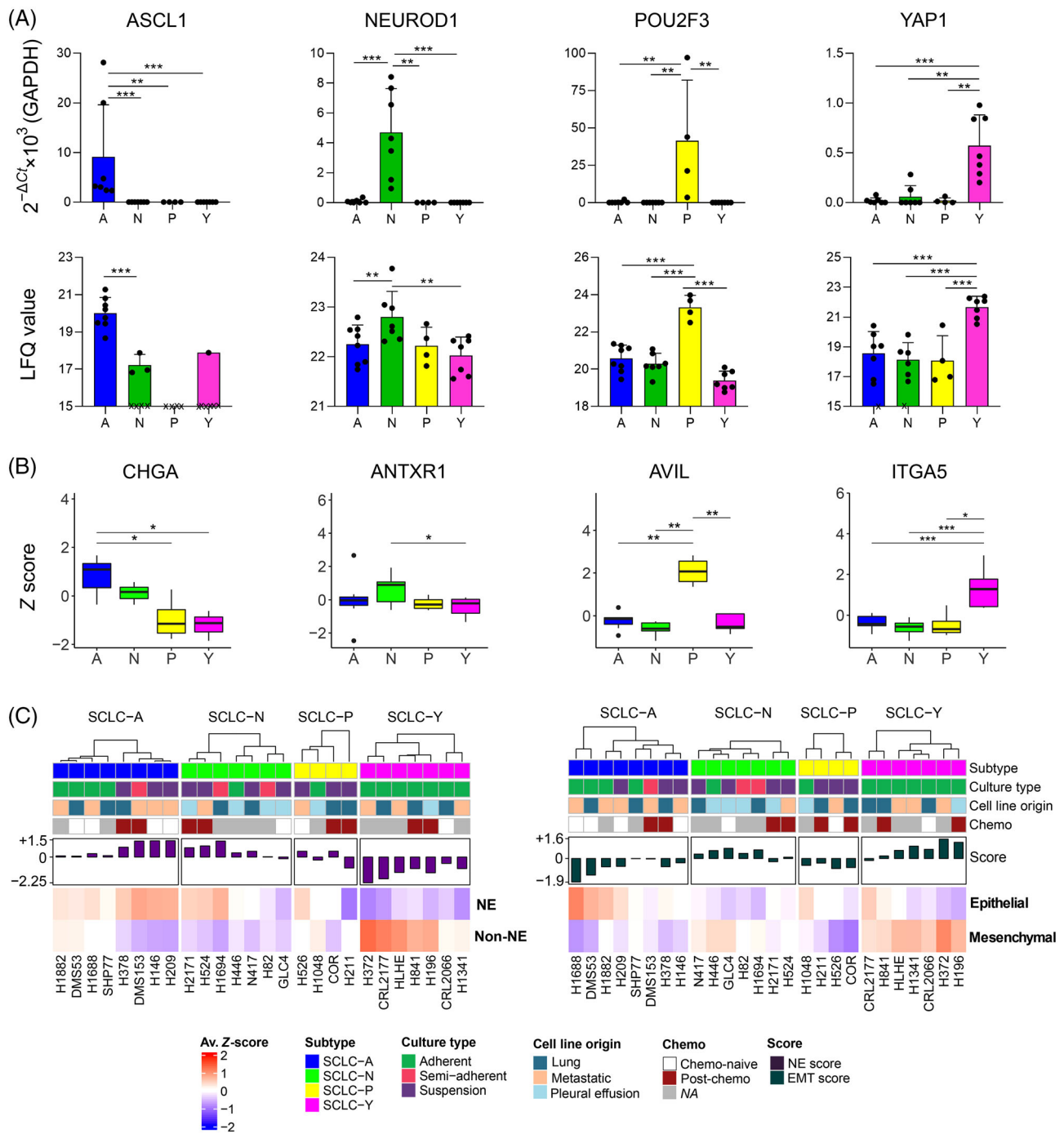
proteins in our cell lines. The CM consisted of relatively more secreted and plasma proteins, highlighting its added value in the search for potential blood-based biomarkers.

First, according to their *ASCL1*, *NEUROD1*, *POU2F3* and *YAPI* mRNA expression patterns – we grouped the cell lines into one of the four respective subgroups: SCLC-A, SCLC-N, SCLC-P and SCLC-Y (eight, seven, four and seven cell lines, respectively) (Figure 1A, upper panel). These transcription factors also showed increased protein levels in their respective subtype (Figure 1A, lower panel).

In general, we identified several members of the Myc proto-oncogene (MYC) family universally expressed across our samples, such as MYC and protein L-Myc (MYCL),<sup>6</sup> or the transcription activator nuclear factor 1 B-type (Figure S2a).<sup>56</sup> The protein products of *TP53* and *RBI* genes, which are well known to be genetically altered in SCLC,<sup>57–59</sup> were quantified in 100.0% and 88.5% of the cell lines, respectively, independent of the previously described *TP53* and *RBI* mutational status of the cell lines<sup>35</sup> (Figure S2b). Several well-accepted subtype markers<sup>3</sup> showed the expected protein expression profile across the subtypes (Figure S2c), such as chromogranin-A (SCLC-A marker), anthrax toxin receptor 1 (SCLC-N marker), advillin (SCLC-P marker) and multiple integrins (SCLC-Y markers) (Figure 1B).

The cell lines were further characterized by their NE and EMT features. NE scores built from 19 NE and 17 non-NE markers,<sup>7</sup> and EMT scores based on 12 epithelial and 10 mesenchymal markers,<sup>44</sup> were calculated for each cell line (Figure S3a,b). The mean protein abundances of NE, non-NE, epithelial and mesenchymal markers in each sample are shown in Figure 1D, whereas the mean NE and EMT scores across the subtypes are depicted in Figure S3c. As expected, most SCLC-A cell lines expressed NE and epithelial markers more strongly than non-NE and mesenchymal markers ( $M_{NE\ score} = 0.71$ ,  $M_{EMT\ score} = -0.70$ ). SCLC-N was found to be a rather NE subtype with mixed epithelial–mesenchymal characteristics ( $M_{NE\ score} = 0.56$  and  $M_{EMT\ score} = 0.39$ ). SCLC-P exhibited moderate non-NE characteristics in our dataset (i.e. lower than SCLC-A and -N, but higher than SCLC-Y); however, a high expression of epithelial markers was detected ( $M_{NE\ score} = -0.05$  and  $M_{EMT\ score} = -0.61$ ) in this subtype. In contrast, SCLC-Y cell lines exhibited prominent non-NE and mesenchymal traits ( $M_{NE\ score} = -1.35$  and  $M_{EMT\ score} = 0.75$ ). In line with these findings, delta-like protein 3 (DLL3) protein, an inhibitory Notch pathway ligand,<sup>60</sup> was expressed in decreasing amounts from SCLC-A to SCLC-Y, indicating the Notch pathway's gradual activation (Figure S2c).

Comparing NE and EMT scores between cell lines with different properties, such as culture type, cell line origin and treatment with chemotherapy, we found that adherent cell lines harbour significantly lower NE scores than the non-adherent ones (Figure S3d–f). In addition, a



**FIGURE 1** Proteomic analysis of small-cell lung cancer (SCLC) cell lines highlights molecular heterogeneity: (A) The mRNA expression of key genes *ASCL1*, *NEUROD1*, *POU2F3* and *YAPI* to determine the molecular subtypes (top). Data is shown as mean  $\pm$  SD of the  $2^{-\Delta Ct} \times 1000$  value, normalized to glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*). Each dot represents one cell line and is the mean of two biological replicates performed in triplicates. The significance of Mann–Whitney *U* tests is indicated above the boxplots ( $*p < 0.05$ ;  $**p < 0.01$ ;  $***p < 0.001$ ). Bottom panel shows label-free quantitation (LFQ) values derived from the proteomic analysis, shown as mean  $\pm$  SD of each cell line, for each defined subtype. Missing values are indicated by an x. The significance of independent *t*-tests is indicated above the boxplots ( $*p < 0.05$ ;  $**p < 0.01$ ;  $***p < 0.001$ ); (B) protein expression profile of four well-known subtype markers (from left to right: *CHGA*, *ANTXR1*, *AVIL* and *ITGA5*, markers for SCLC-A, -N, -P and -Y, respectively). The significance of independent *t*-tests is indicated above the boxplots ( $*p < 0.05$ ;  $**p < 0.01$ ;  $***p < 0.001$ ); (C) mean Z-score values for neuroendocrine (NE) and non-NE markers in each cell line (left), and mean Z-score values for epithelial and mesenchymal markers (right)

significant negative correlation was observed between NE and EMT scores (Figure S3g).

### 3.2 | Manifestation of heterogeneous in vitro growth characteristics in the proteome

Although maintained in the same in vitro conditions, the cell lines showed significantly different growth characteristics. Specifically, out of the 26 cell lines, 10 (38.5%) grew in suspension, 3 (11.5%) in a semi-adherent form and the other 13 (50.0%) grew adherent on the plastic (Figure 2A). Adherent and non-adherent cell lines showed clearly distinct protein expression profiles (Figure 2B,C). In total, 270 and 148 proteins were significantly (BH adjusted  $p < 0.05$ ) downregulated in suspension cell lines compared to adherent cell lines in the CP and CM, respectively, whereas 244 and 244 proteins were significantly upregulated in suspension cell lines in the CP and CM, respectively (Figure 2B,C, left). An ORA of differentially expressed proteins, separately for CP and CM but combining up- and downregulated proteins, showed that KEGG pathways such as protein processing in endoplasmic reticulum, lysosome and glycosaminoglycan degradation were significantly ( $p < 0.05$ ) enriched both in CP and CM, as well as other pathways such as endocytosis in CP and gap junction in CM were overrepresented, thereby supporting the phenotypic cell line differences on protein level (Figure 2B,C, right and Table S2).

### 3.3 | Proteome-based SCLC subgroups match with mRNA-based subtyping

To investigate whether proteomic subtyping correlates with the mRNA-based classification, unsupervised consensus clustering of the CP samples was performed based on the most variable proteins (Figure S4a–c). The analysis revealed four clusters in the proteomic data, which agrees with the mRNA-based subtyping (Figure 3A). A discrepancy was detected in only one cell line (H1882), which was classified in the SCLC-A subgroup according to the qPCR data and in the SCLC-P subset based on the proteomic results. Of note, this cell line also displayed a higher *POU2F3* mRNA expression than other SCLC-A cell lines (Figure 3A). Additionally, two adherent SCLC-A samples (H1688, DMS53) were rather separated from their group members (Figure S4c). Notably, SCLC-Y samples exhibited the most distinct protein expression profile. The previous observations are also well reflected on a PCA plot of the CP samples (Figure 3B). The CM samples, on the other hand, exhibit a rather heterogeneous expression profile based on

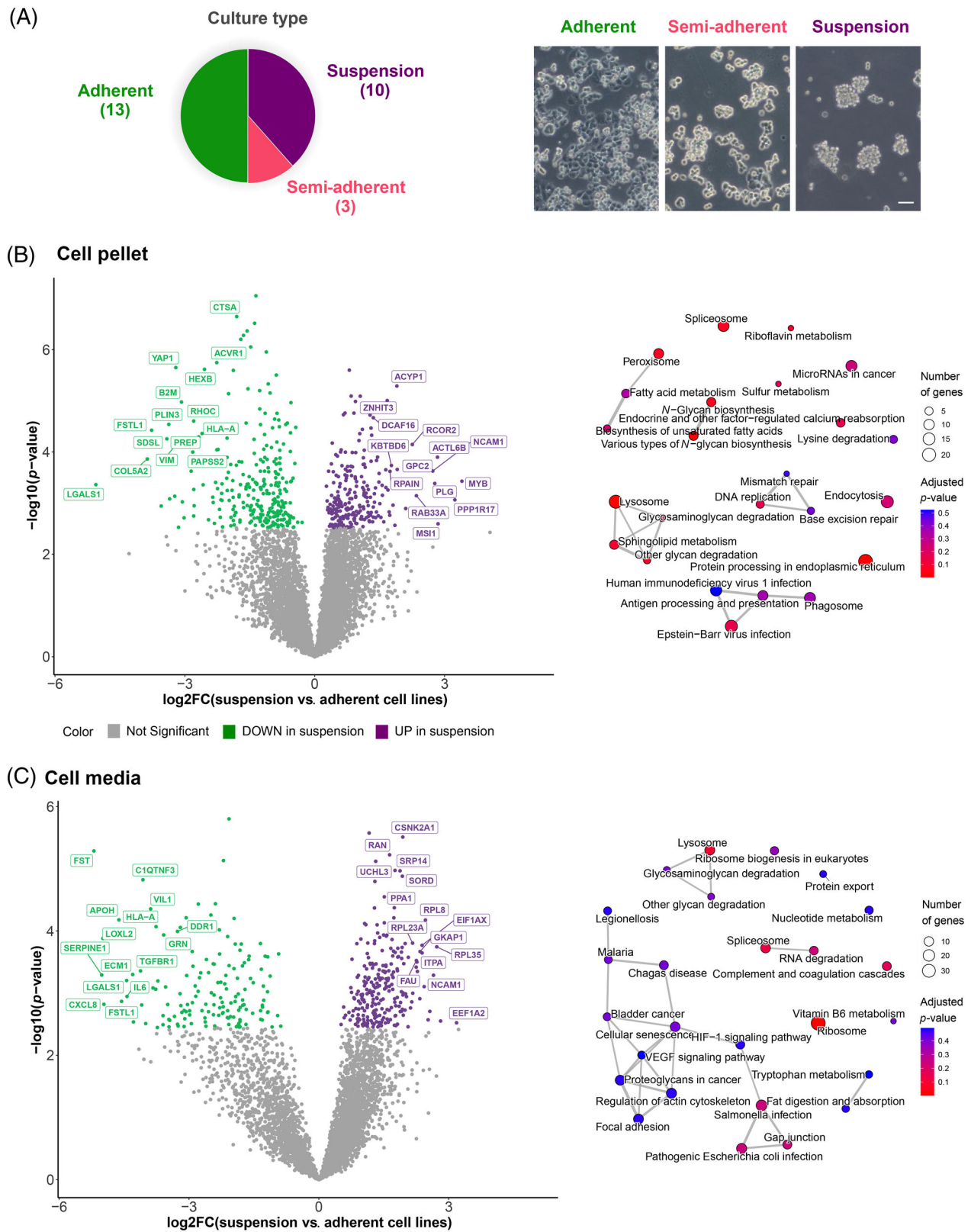
their most variable proteins, and the PCA plot showed no apparent separation according to the mRNA-based subtype classification (Figure 3C). Correspondingly, PVCA revealed that the molecular subtypes were less prominent contributors to protein expression variability in CM compared to CP (Figure 3D). Furthermore, culture type was identified as an important contributor to the proteomic profiles of SCLC cell lines, which was more pronounced in CM.

### 3.4 | Multi-omic portraits of SCLC subtypes outline potential subtype-specific vulnerabilities

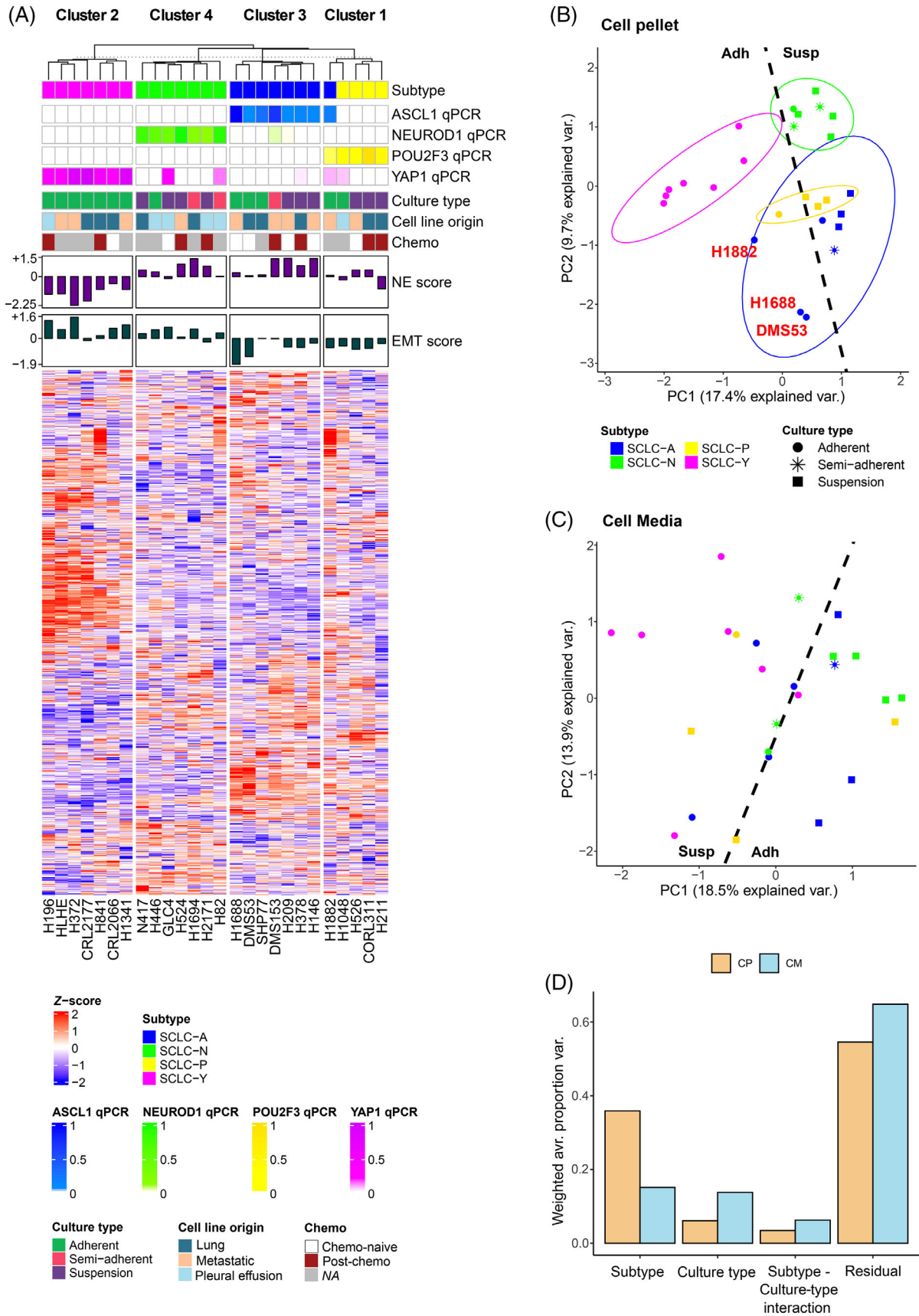
According to the high concordance between mRNA- and proteome-based subtypes, we used the mRNA-based classification system in differential expression analyses between subtypes. We found 367 and 34 subtype-specific proteins (ANOVA BH adjusted  $p < 0.05$ , Tukey's HSD post hoc  $p < 0.05$ ), the levels of which differ in a given subtype compared to all the three other subtypes in the CP and CM data, respectively (Table S3). This also included four proteins with on/off characteristics in the CP data, namely achaete-scute homolog 1 (ASCL1; 'on' in SCLC-A), regulator of G-protein signalling 22 (RGS22; 'on' in SCLC-P), neurexophilin-4 and puratrophin-1 (NXPH4 and PKHG4; 'off' in SCLC-Y). All subtype-specific proteins (including those from CP and CM) were then subjected to the pathway analysis (Table S3).

In order to more systematically investigate the underlying biology of SCLC subtypes, in addition to identifying subtype-specific proteins, we also performed a pathway-based comparison of subgroups using pGSEAs. We looked for significant pathways that were concordantly activated or suppressed in a certain subtype compared to the other three subtypes. To do this, all pairwise subtype comparisons were performed, using the full list of quantified proteins ( $n = 8405$ ) from the CP data. Furthermore, we performed the same analysis on the RNA-Seq data of 50 SCLC cell lines from CCLE<sup>35</sup> ( $n = 9237$  genes) and finally assessed the relationship between proteomic and transcriptomic results (Table S4). These steps are described in detail in Section 2 and Figure S5.

Significantly overrepresented ( $p < 0.05$ ) KEGG processes in SCLC-A according to subtype-specific proteins ( $n = 33$ ) include oxidative phosphorylation (OXPHOS), as well as phenylalanine metabolism and leukocyte transendothelial migration (Figure 4A). Consistent with this, pGSEA results also showed upregulation of OXPHOS and respiratory chain elements based on proteomic data. Positive regulation of neural precursor cell proliferation was supported by both datasets, whereas transcriptomics



**FIGURE 2** In vitro growth characteristics mirrored in the proteome: (A) Pie chart and representative images (taken on a Zeiss Axiovert 40 C microscope) of the different culture types of small-cell lung cancer (SCLC) cell lines ( $n = 26$ ). Scale bar = 100  $\mu\text{m}$ ; (B) volcano plot depicting the results of differential expression analysis results between suspension and adherent cell lines (left) with the corresponding enrichment map of the overrepresented KEGG pathways (right) in the cell pellet (CP) data; (C) volcano plot depicting the results of differential expression analysis between suspension and adherent cell lines (left) with corresponding enrichment map of the overrepresented KEGG pathways (right) in the culture media (CM) data



**FIGURE 3** The mRNA-based classification of small-cell lung cancer (SCLC) subtypes correlates with proteomic data: (A) Heat map of consensus clustering results using proteins from the cell pellet (CP) data showing high variation (>1.25 SD). Samples are sorted according to their cluster assignments and their representative protein expression profiles are shown; (B) principal component analyses (PCA) plot of the

showed the activation of the subpallium development gene set in SCLC-A (Figure 4B).

Subtype-specific proteins in SCLC-N ( $n = 54$ ) contributed to the significant overrepresentation of KEGG pathways, such as cell cycle, phagosome, riboflavin metabolism and peroxisome proliferator-activated receptor signalling pathway (Figure 4C). SCLC-N could be further characterized by the suppression of epidermis development processes based on pGSEA. Proteomic data also outlined the downregulation of immune response, cytokine signalling, cell adhesion and cytoskeleton organization, as well as upregulation of transcription and DNA replication (Figure 4D).

Considering SCLC-P specific proteins ( $n = 32$ ), three significantly enriched KEGG pathways were detected, namely phospholipase D signalling, lysosome as well as other glycan degradation (Figure 4E). Furthermore, pGSEA showed activation of the neurotrophin signalling pathway and the lamellipodium organization gene set in SCLC-P (Figure 4F).

Regarding SCLC-Y, multiple KEGG pathways were significantly overrepresented in the context of subtype-specific proteins ( $n = 271$ ), such as extracellular matrix (ECM)-receptor interaction, focal adhesion, spliceosome, peroxisome or O-glycan biosynthesis (Figure 5A). Similarly, a bulky list of processes showed upregulation in SCLC-Y compared to other subtypes according to pGSEA, such as ECM organization, cytokine-mediated signalling, interleukin signalling, inflammatory response, EMT, response to growth factors, cell-substrate adhesion and mitogen-activated protein kinase (MAPK) cascade. Transcriptomic data showed an activation of apoptotic pathways and the Janus kinase-signal transducer and activator of transcription signalling, whereas proteomics revealed the upregulation of signalling by Rho-GTPases, as well as activation of transmembrane transporter disorder-related processes. Moreover, DNA repair, protein acetylation and chromatin modification were found downregulated in this subtype, as detected by the proteomic data (Figure 5B).

In order to verify the subtype-specificity observed in our cell lines in SCLC tissues as well, we investigated the behaviour of representative gene sets ( $n = 33$ ) of subtype-specific processes outlined earlier in the SCLC tissue transcriptomic dataset published by George et al.<sup>8</sup> ssGSEA

identified 22 gene sets for which subtype-specificity was confirmed to some extent in the tissue data (Figures 5C and S6). Of these, OXPHOS activation in SCLC-A, upregulated DNA replication and downregulated immune response in SCLC-N, as well as more active EMT and suppressed DNA repair in SCLC-Y could be highlighted. Neither of the subtype-specific processes in SCLC-P could be verified by tissue transcriptomics.

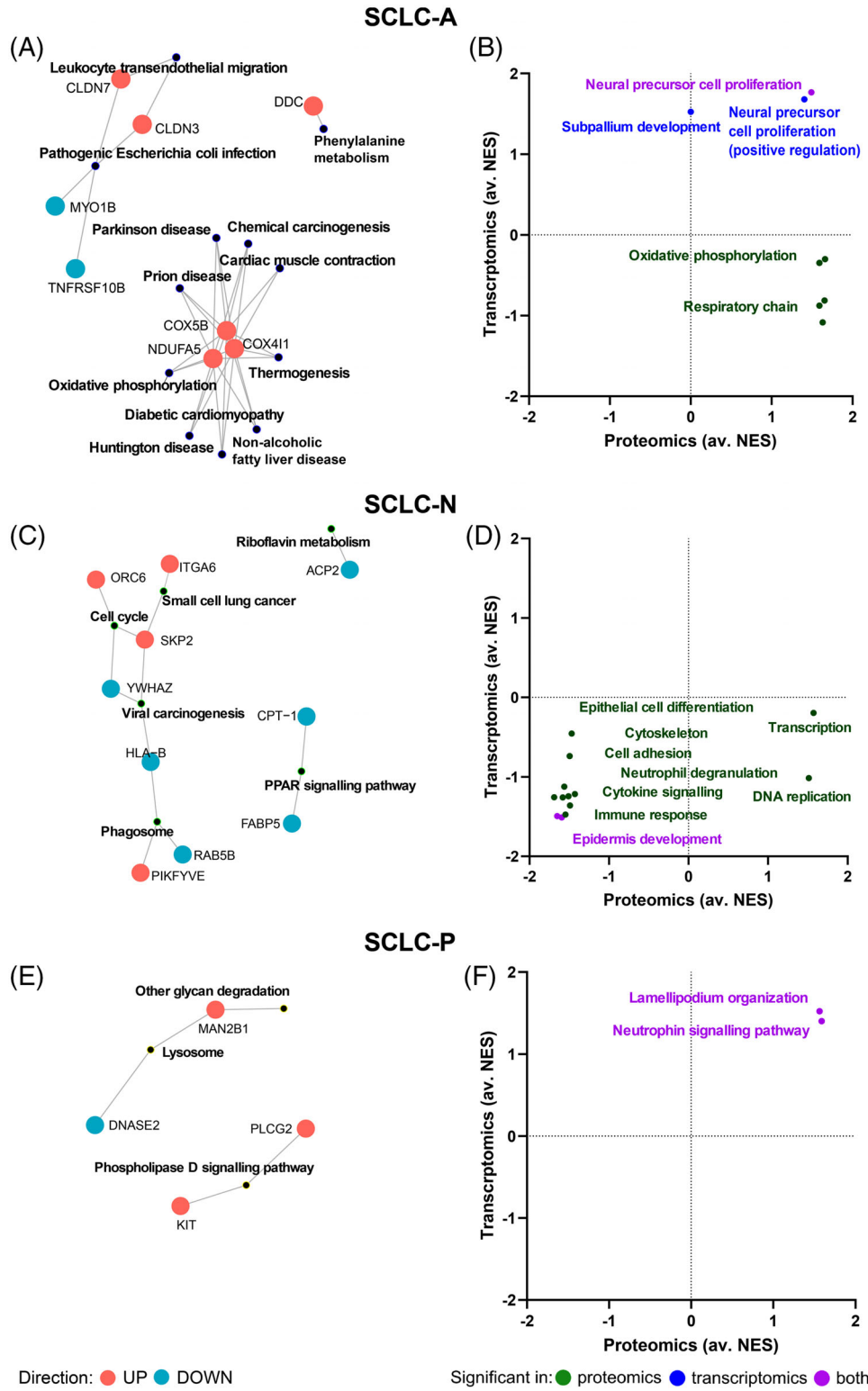
### 3.5 | Proteomic analysis identifies potential diagnostic markers and druggable targets for SCLC subtypes

In addition to differential expression analysis, sPLS-DA (Figure S7) was performed separately for CP and CM to identify the proteins that are most suitable for subtype classifications based on their expression patterns (i.e. potential IHC- or blood-based markers). The analysis resulted in 104 proteins (82 and 23 in the CP and CM datasets, respectively; one protein was detected in both datasets) showing clearly distinct profiles between at least two subtypes (Figure S8). The proteins selected by sPLS-DA could be separated into the following expression pattern categories: upregulated in SCLC-A versus -N, upregulated in SCLC-N versus SCLC-A, upregulated in SCLC-P (vs. other subtypes), upregulated in SCLC-Y and downregulated in SCLC-Y (Figure 6A). Of note, 35 and 17 out of the 104 proteins showed either clear up- or downregulation in the SCLC-Y subtype, respectively. For several markers, the expression pattern was matching with transcriptomic data from SCLC tissues (top three is shown in Figure 6B). Notably, eight of such proteins were found to be overexpressed in SCLC-Y, including probable glutathione peroxidase 8 (*GPX8*), pyruvate dehydrogenase kinase isoform 2 (*PDK2*) and tyrosine-protein kinase receptor UFO (*AXL*) (Figures 6B and S9a).

Among the proteins confirmed on tissue transcriptome level, six are also detectable in human blood plasma by MS, namely tyrosine-protein kinase receptor UFO (*AXL*, upregulated in SCLC-Y), basal cell adhesion molecule (*BCAM*, upregulated in SCLC-A vs. -N), GTP-binding protein 1 (*GTPBP1*, upregulated in SCLC-N vs. -A), keratin-18 (*KRT18*, upregulated in SCLC-A vs. -N), noelin (*OLFML1*, upregulated in SCLC-N vs. -A) and phospholipase C-gamma-2 (*PLCG2*, upregulated in SCLC-P). The most

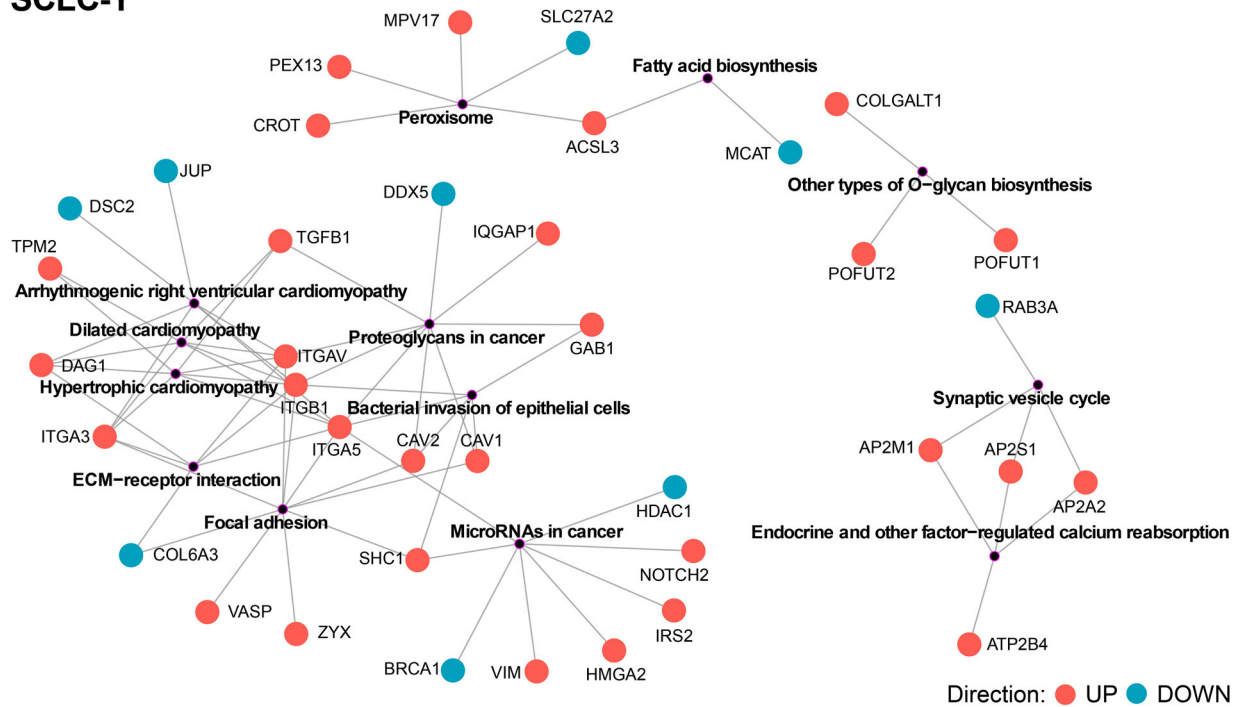
CP data using the most variable ( $>1.25$  SD) proteins. Cell lines are coloured according to the subtype and shape corresponds to culture type. Normal data ellipses are also drawn for each subtype (probability = 68%); (C) PCA plot of the culture media (CM) data using only the most variable ( $>1.25$  SD) proteins. Cell lines are coloured according to the subtype and shape corresponds to culture type; (D) principal variance component analysis (PVCA) reveals the contribution of subtype and culture type to the proteomic profile differences in CP and CM. The residual variation (noted as 'residual') represents the remaining biological and technical variance in the dataset which could not be attributed to the abovementioned factors.



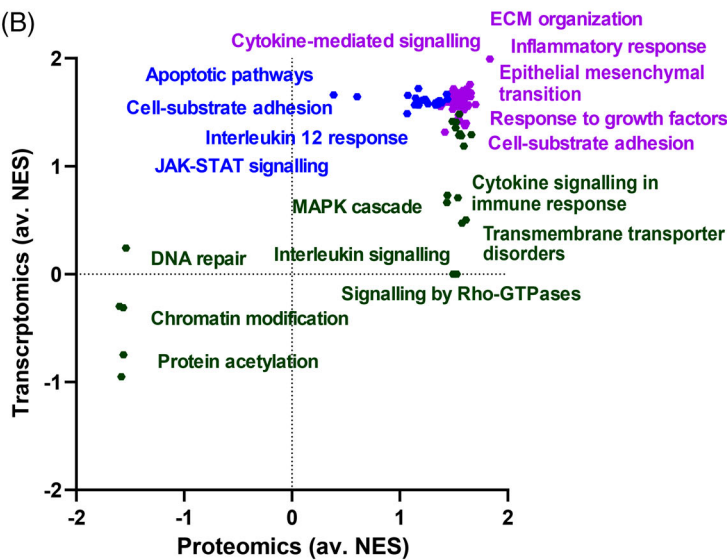


**FIGURE 4** Subtype-specific biological processes in small-cell lung cancer (SCLC)-A/N/P. Significantly overrepresented KEGG pathways ( $p < 0.05$ ) derived from the list of subtype-specific proteins, as well as the members of these pathways (red and blue colours mean up- and downregulated in the given subtype compared to the other subtypes, respectively) are shown on the left side of the panel. The characteristic gene sets for each subtype, as determined by pre-ranked gene set enrichment analysis (pGSEA), are shown on the right side of the panel. The x axis indicates the average normalized enrichment score (av. NES) in proteomics for comparisons SCLC-A versus -N/P/Y, SCLC-N versus -A/P/Y, or SCLC-P versus -A/N/Y, whereas y axis indicates the av. NES in transcriptomics. Dots refer to individual gene sets, which are summarized by keywords. Gene set activation or suppression supported by both omic data is shown in purple, whereas green and blue means gene sets supported only by proteomics or transcriptomics, respectively. (A and B) SCLC-A, (C and D) SCLC-N and (E and F) SCLC-P

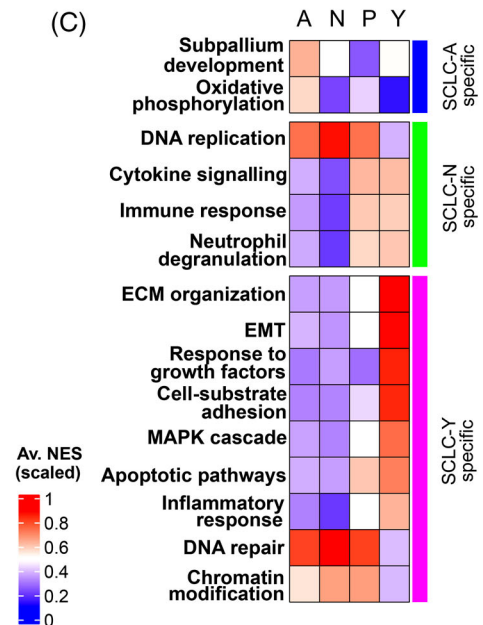
## (A) SCLC-Y



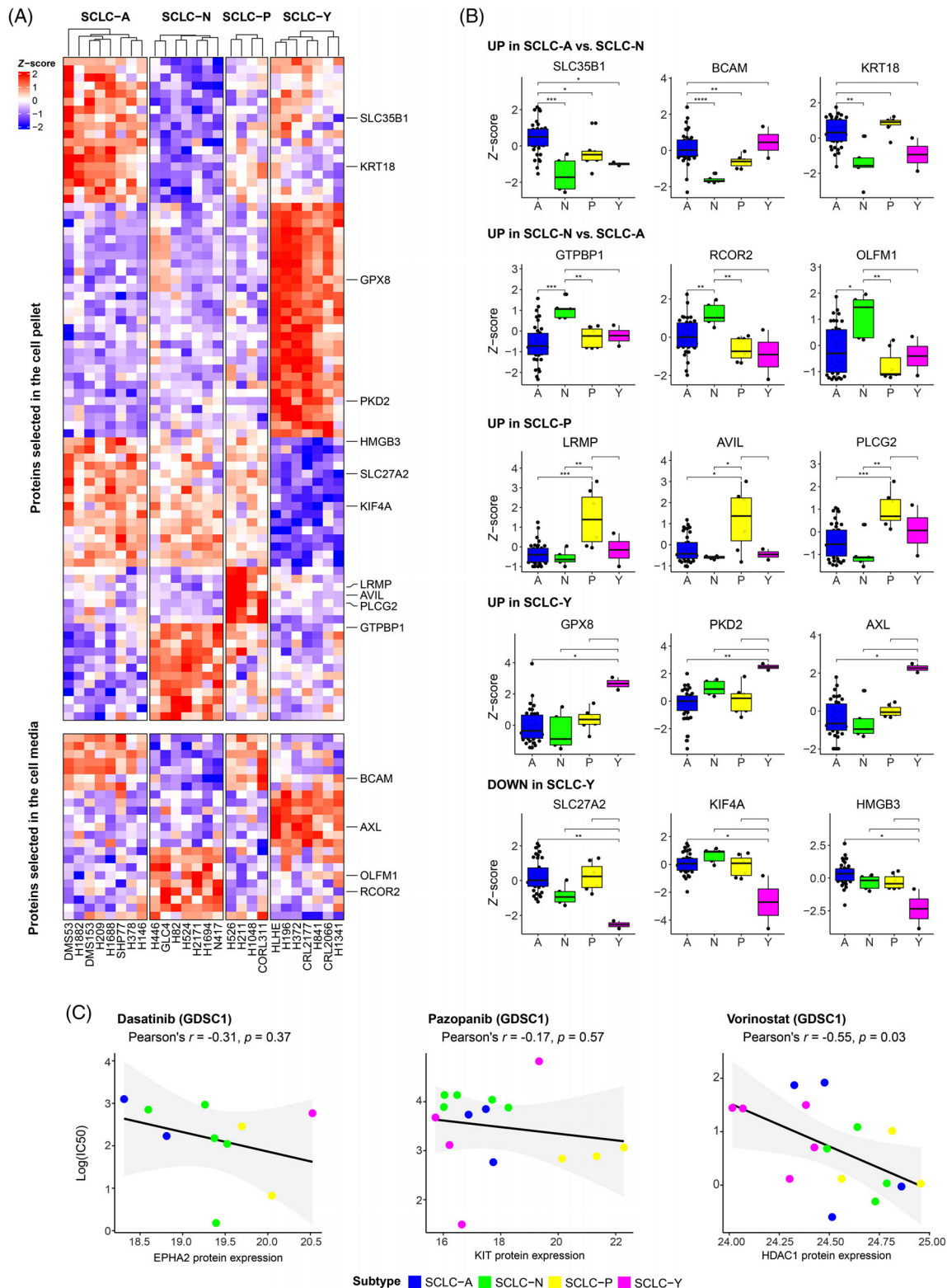
## (B)



## (C)



**FIGURE 5** Subtype-specific biological processes in small-cell lung cancer (SCLC)-Y and processes verified using tissue transcriptomics: (A) Significantly overrepresented KEGG pathways ( $p < 0.05$ ) derived from the list SCLC-Y specific proteins, as well as the members of these pathways (red and blue means up- and downregulated in SCLC-Y, respectively); (B) the characteristic gene sets for SCLC-Y determined by pre-ranked gene set enrichment analysis (pGSEA). The x axis indicates the average normalized enrichment score (NES) in proteomics for comparisons SCLC-Y versus -A/N/P, whereas the y axis indicates the average NES in transcriptomics. Dots show the individual gene sets, which are then summarized by keywords. Colouring is based on whether the gene set activation or suppression was detected by both proteomics and transcriptomics (purple), or only by proteomics (green) or transcriptomics (blue). (C) The average value of NESs per subtype for SCLC tissue samples, derived from single-sample gene set enrichment analysis (ssGSEA) for some representative gene sets where the subtype specificity was also supported by the tissue data.



**FIGURE 6** Proteins with diagnostic and therapeutic relevance in small-cell lung cancer (SCLC) subtypes: (A) The proteins selected by partial least squares discriminant analysis (SPLS-DA), which are suitable for separating the subtypes based on their expression profile. Results from cell pellet (CP) and culture media (CM) are displayed on the top and bottom heat map, respectively. Proteins with the best matching expression profiles between cell line proteomics and tissue transcriptomics are highlighted; (B) gene expression differences of the best matching transcripts showing subtype-specific expression profile. The significance of Wilcoxon tests is indicated above the boxplots ( $*p < 0.05$ ;  $**p < 0.01$ ;  $***p < 0.001$ ); (C) the  $\ln(\text{IC}_{50})$  values of the cell lines for drugs selected from the Genomics of Drug Sensitivity in Cancer 1 (GDSC1) database (from left to right: dasatinib, pazopanib, and vorinostat), as a function of the measured protein expression. The results of the Pearson correlation analysis are indicated above the scatter plots. Dots are coloured according to the cell line's subtype assignment.

TABLE 2 Potentially targetable subtype-specific proteins

Protein name (gene name)	Dataset	Specificity	Annotation	FDA-approved drugs with pharmacological action
Aromatic-L-amino-acid decarboxylase ( <i>DDC</i> )	CP	↑ in A	Enzyme that catalyses dopamine and serotonin synthesis	Benserazide, carbidopa, methyldopa
Ephrin type-A receptor 2 ( <i>EPHA2</i> )	CP	↑ in Y	Receptor tyrosine kinase involved in contact-dependent bidirectional signalling with neighbouring cells	Dasatinib, regorafenib
Histone deacetylase 1 ( <i>HDAC1</i> )	CP	↑ in A/N/P	Histone deacetylase with regulatory function in transcriptional processes	Romidepsin, vorinostat
Integrin alpha-V ( <i>ITGAV</i> )	CP	↑ in Y	Integrin, receptor for a wide array of proteins. CD marker	Antithymocyte immunoglobulin, levothyroxine
Integrin beta-1 ( <i>ITGB1</i> )	CP	↑ in Y	Integrin, receptor for a wide array of proteins. CD marker	Antithymocyte immunoglobulin
Mast/stem cell growth factor receptor kit ( <i>KIT</i> )	CP and CM	↑ in P	Receptor tyrosine kinase, acts as cell-surface receptor for the cytokine KITLG/SCF. CD marker	Ancestim, imatinib, lenvatinib, pazopanib, regorafenib, ripretinib, sorafenib, sunitinib, tivozanib

Abbreviations: ↑, higher expression; A, SCLC-A; CD, cluster of differentiation; CM, culture media; CP, cell pellet; FDA: Food and Drug Administration; N, SCLC-N; P, SCLC-P; SCLC, small cell lung cancer; Y, SCLC-Y.

promising blood-based biomarker from this list is the protein UFO, which has been previously detected via immunoassay as well.

Finally, we investigated whether among the proteins detected by differential expression analysis or sPLS-DA (a total of 418 unique proteins), we could find ‘druggable’ proteins<sup>61</sup> (i.e. drug targets approved by FDA). We identified six of such proteins, listed in Table 2, as potential targets for subtype-specific therapies: aromatic-L-amino-acid decarboxylase (*DDC*, overexpressed in SCLC-A), ephrin type-A receptor 2 (*EPHA2*), integrin alpha-V and beta-1 (*ITGAV*, *ITGB1*, overexpressed in SCLC-Y), histone deacetylase 1 (overexpressed in SCLC-A/N/P vs. SCLC-Y), and mast/stem cell growth factor receptor Kit (*KIT*, overexpressed in SCLC-P). Multiple drugs that directly interact with the aforementioned proteins as part of their mechanism of action could be delineated (Table 2). Moreover, SCLC cell lines have already been tested against seven of these drugs (CancerRxGene database<sup>40</sup>), namely dasatinib (targeting *EPHA2*), vorinostat (targeting *HDAC1*), imatinib, pazopanib, sorafenib, sunitinib and tivozanib (all targeting *KIT*). We therefore investigated whether the subtypes showed differences in sensitivity to these drugs, either in GDSC1 or in GDSC2 datasets (Figure S9b). Generally, we found that lower *EPHA2*, *KIT* and *HDAC1* protein abundance in our cell lines were indicative of increased resistance to the drugs targeting these proteins (Figure 6C) according to the GDSC1 dataset. This trend was not verified by the GDSC2 dataset for dasatinib; however, median

IC50 values of some *KIT*-targeting drugs (pazopanib, sunitinib, tivozanib) were lowest in SCLC-P cell lines, as well as SCLC-Y subtype’s higher resistance to vorinostat could be validated (Figure S9b).

## 4 | DISCUSSION

Although clinically SCLC is still regarded as a homogeneous tumour type with NE characteristics, it was described decades ago that besides the ‘classical’ form of SCLC with suspension growth type and NE phenotype, a ‘variant’ form (NE-low type) of SCLC also exists, mainly forming adherent cell cultures.<sup>62,63</sup> In addition, tumours lacking NE differentiation (non-NE) have been also described recently.<sup>6</sup> In our SCLC cell line panel, 50.0% of the samples grew adherently, whereas the other half semi-adherently or in suspension. A significant positive relationship between the adherent phenotype and the non-NE features was observed. Importantly, the proteomic landscape of the examined cell lines also mirrored these phenotypic differences, and proteins associated with several pathways involved in cell adhesion were identified as differentially expressed between adherently or in-suspension-growing cell lines, including glycosaminoglycan degradation and endocytosis or gap junction.

Recent profiling studies highlighted that the NE and non-NE classes of SCLC can be further divided into additional subtypes, namely SCLC-A and -N (NE-high and

NE-low, respectively) and SCLC-P and -Y (non-NE types).<sup>6</sup> Our cell lines were also categorized accordingly using qPCR, resulting in eight, seven, four and seven *ASCL1*+, *NEUROD1*+, *POU2F3*+ and *YAPI*+ cell lines, respectively. In the global proteomic data, these transcription factors showed higher abundances in their respective subtype, with only the elevated level of neurogenic differentiation factor 1 (*NEUROD1*) in SCLC-N cell lines being not significantly different in all three comparisons.

According to the protein expression of NE and non-NE markers described by Zhang et al.,<sup>7</sup> as expected, SCLC-A and -N cell lines showed NE features, whereas those in the SCLC-P and -Y groups were associated with rather non-NE protein profiles. This is in line with the results of a recent IHC-based study on human tissue specimens, in which the authors also found that the likelihood of *POU2F3* expression in SCLC is quantitatively linked with the level of NE marker expression and SCLC-P tumours are characterized by the near-complete absence of NE differentiation.<sup>64</sup> Therefore, including *POU2F3* as a potential additional diagnostic marker might represent an appealing approach for the diagnosis of SCLC tumours that lack or exhibit minimal level of standard NE markers.<sup>64</sup> Meanwhile, in another study, the same group also showed that the expression of conventional markers linked with NE differentiation is substantially higher in *ASCL1*- and *NEUROD1*-defined tumours (vs. *ASCL1*/*NEUROD1*-double-negative tumours), which as well corresponds with our findings.<sup>12</sup> Regarding epithelial and mesenchymal traits, a study on SCLC cell lines<sup>65</sup> outlined that SCLC-P and -Y carry epithelial and mesenchymal attributes, respectively, whereas SCLC-A and -N subtypes have mixed characteristics. Our proteomic study, based on the epithelial and mesenchymal markers established by Kohn et al.,<sup>44</sup> confirmed these observations. Interestingly, the EMT score was not influenced by the cell lines' site of origin. Our data also showed that strong NE or EMT characteristics tend to be mutually exclusive.

Importantly, we could categorize our SCLC cell lines into four subgroups based on their comprehensive cellular proteomic profiles, which agreed with the pre-defined mRNA-based subtypes (only one cell line was misclassified in proteomics). Of note, the overall high variability of cell media proteins compared to cellular proteins probably interfered with the detection of subtype-specific signatures in this dataset ( $M_{SD}$  in CP = 0.85,  $M_{SD}$  in CM = 1.20). Another driver of the proteomic differences was the culture type, affecting both the CP and CM.

To describe the unique traits of SCLC subtypes, we examined the differentially expressed proteins between subtypes and performed a multi-omic pathway-level analysis using the expression differences of all quantified proteins (from our dataset) and transcripts (from CCLE

transcriptomic data of SCLC cell lines) across the subtypes. Accordingly, we uncovered a list of potential subtype-specific therapeutic vulnerabilities.

Due to the well-known regulatory role of *ASCL1* in neural differentiation,<sup>66</sup> neural precursor cell proliferation and subpallium development-related proteins showed a concordant upregulation in SCLC-A subtype. Interestingly, proteomic data clearly indicated that the activation of OXPHOS and respiratory chain elements are strongly specific for SCLC-A. It has been recently described that cell lines not expressing *MYC* (characteristic for SCLC-A<sup>6,8</sup>) relies more on oxidative metabolism<sup>67</sup> suggesting that SCLC-A tumours might be susceptible to OXPHOS inhibitors.<sup>68</sup>

SCLC-N cell lines, which predominantly formed suspension cultures in our study, showed a concordant downregulation of cell adhesion pathways. Harvesting the protein-level data also highlighted the increased activity of DNA replication and transcription and the depletion of cytokine-mediated signalling in this subtype. Interestingly, a similar trend was described in a study where gene expression differences between SCLC and normal tissue were addressed.<sup>69</sup>

In the epithelial-like SCLC-P cell lines, upregulation of the lamellipodium organization pathway was observed, which constitutes a crucial step in EMT leading to increased cell motility and invasive capacities.<sup>70</sup> In addition, upregulation of neurotrophin signalling foreshadows that poly(adenosine diphosphate-ribose) polymerase (PARP) inhibitors can be effective therapeutic agents for SCLC-P as previously proposed,<sup>13</sup> as prolonged PARP activation was found to contribute to neurotrophin-induced neuronal death.<sup>71</sup> Moreover, our data suggests that direct targeting of neurotrophin signalling might be an appropriate treatment option in *POU2F3*-driven SCLC.<sup>72</sup>

SCLC-Y cell lines formed a clearly defined subgroup in our SCLC samples. A unique finding in the proteomic data was the downregulation of protein acetylation, chromatin modification (driven by the decreased levels of several histone acetyltransferases) and DNA double-strand break repair pathways in this subtype. Moreover, we also detected overexpression of the MAPK-cascade and Rho-GTPase signalling members in this subtype. Rho is crucial for YAP/TAZ activity<sup>73</sup>; therefore, the upregulation of the latter pathways can be expected in *YAPI*+ cell lines. Recently, Caeser et al. reported that in SCLC-A, the activation of the MAPK pathway rather has inhibitory functions, compared to SCLC-N/P cell lines. However, no cell lines from the SCLC-Y subtype were tested by these authors.<sup>74</sup> In line with the adherent nature of the *YAPI*-driven cell lines, we also observed that focal adhesion, ECM organization and cell-substrate protein pathways were uniformly upregulated in the SCLC-Y subgroup, just

as peroxisome and endocytosis-related proteins. Tlemsani et al. described that SCLC-Y cell lines demonstrate high presenting and native immune predisposition.<sup>65</sup> Additionally, these cells also have the highest antigen-presenting machinery scores, thus anticipating that the SCLC-Y subtype might be sensitive to immune-checkpoint inhibitors.<sup>65</sup> In line with this, we also identified a distinctive upregulation of cytokine-mediated signalling and inflammatory response in this subtype.

Of note, controversies around SCLC-Y still exists, because comprehensive immunohistochemical and histopathologic analyses of SCLC subtypes in patient samples failed to identify a distinct *YAPI*-driven subtype.<sup>12</sup> In contrast, our preclinical proteomic study clearly identifies a unique SCLC-Y subtype among the examined cell lines. Accordingly, a significant proportion of the 104 proteins with diagnostic relevance was linked to SCLC-Y. The most promising biomarkers for SCLC-Y include *GPX8*, *PKD2* and *UFO*, from which *UFO* is also potentially detectable in the human blood plasma. This is of clinical importance because the lack of appropriate tissue samples highlights the diagnostic relevance of blood-based biomarkers. Indeed, so far only a single study attempted to define a panel of subtype-specific blood-based biomarkers.<sup>75</sup> In their study, however, the authors examined the diagnostic relevance of circulating cell-free DNA, not proteins.<sup>75</sup> Regarding the potential SCLC-Y markers identified in our study, gene expression of *GPX8* was previously correlated with *YAPI* expression in SCLC cell lines.<sup>76</sup> Little is known about the role of protein kinase D (*PKD*) in SCLC, but a previous study has shown that *PKD* is activated in cell lines H69, H345 and H510 via a protein kinase C-dependent pathway activation.<sup>77</sup> The *AXL* gene (encoding the *UFO* protein) is associated with the mesenchymal phenotype and a potential target for overcoming resistance to epidermal growth factor (*EGF*) receptor inhibitors.<sup>78</sup>

Our data outlined six proteins that show remarkable differences in abundance across the subtypes and are also FDA-approved drug targets, namely *DDC* (overexpressed in SCLC-A), *EPHA2*, *ITAV* and *ITBI* (corresponding to genes *EPHA2*, *ITGAV* and *ITGB1*, overexpressed in SCLC-Y), *HDAC1* (downregulated in SCLC-Y) and *KIT* (upregulated in SCLC-P). Cells with high levels of *ASCL1* also showed stronger expression of *DDC*,<sup>79</sup> whereas *EPHA2*, a non-NE marker,<sup>7</sup> was previously described as upregulated in SCLC-Y.<sup>65</sup> Overexpression of integrins in SCLC-Y contributes to chemotherapy resistance through the suppression of chemotherapy-induced apoptosis.<sup>80</sup> In line with this, our group previously reported a positive correlation between *YAPI* protein abundance and resistance to chemotherapeutic agents in SCLC cell lines.<sup>9</sup> An *HDAC* inhibitor resistance in SCLC-Y, foreshadowed by *HDAC1*

downregulation in our study, was also reported recently.<sup>81</sup> *KIT* protein is a known SCLC-P marker.<sup>82</sup> Importantly, our hypothesis that subtype-specific overexpression of proteins indicates sensitivity to certain drugs was supported by data from CancerRxGene.<sup>40</sup> In particular, we confirmed that *KIT*-targeting drugs, such as pazopanib, are potentially adequate to target *POU2F3*-driven SCLCs, whereas *YAPI*-driven tumours are more resistant to vorinostat (targeting *HDAC1*).

Among others, the main advantage of cell line-based studies is that one can examine pure populations of homogeneous tumour cells without admixed stromal or inflammatory cells. This is of scientific importance in drug sensitivity assays and subtyping studies such as the current one.<sup>83,84</sup> Importantly, our proteomic data from cell lines showed a notable overlap with transcriptomic data from SCLC tissue samples in terms of subtype-specific pathways, suggesting that *in vitro* investigations of SCLC cell lines are useful in characterizing SCLC tissue subtypes. A comprehensive multi-omic investigation of SCLC tissue and cell line data by Gay et al. demonstrated the same, namely that SCLCs can be subcategorized both in the presence and absence of tumour microenvironment.<sup>13</sup>

Cell lines, however undoubtedly not ideal models for profiling studies, offer many advantages that complement the use of tumour tissues and animal models for the examination of SCLC.<sup>83</sup> Nevertheless, some study limitations should be underscored. Although we confirmed several previously described subtype markers, and our results are largely consistent with previous transcriptomic data for subtype-specific pathways, the total number of cell lines included was relatively small. Additionally, given that all *YAPI*-driven cell lines had adherent culture type, no subsequent analyses were feasible to investigate subtype and culture type separately. Therefore, the influence of culture type on the distinct characteristics of SCLC-Y is still unclear. It is also important to note that unique up- or downregulation of pathways in a subtype might not translate to dependency or independency from those pathways, as interconnected biological processes and numerous regulatory factors including feedback and feed-forward loops interfere with such processes. The lack of appropriate treatment-related data on some cell lines represents another potential study limitation. This is partly because the majority of the examined cell lines were established in the past century, and therefore, the clinicopathological data concerning the tumour they were derived from could not be retrieved even after an in-depth literature search. Notably, however, neither the NE/mesenchymal characteristics nor the protein expression profile differed significantly according to the presence or absence of chemotherapy (data not shown).

Taken together, our results should primarily be considered hypothesis-generating for future studies, and all findings should be examined in light of the abovementioned limitations. Accordingly, proteomic analyses of larger SCLC cohorts, preferably including fresh tissue samples of patients with homogeneous treatment histories, are needed to validate our findings.

## 5 | CONCLUSIONS

To conclude, this is the first large-scale proteomic study in human SCLC cell lines concerning the newly defined molecular subtypes. We report that SCLC cell lines can be divided into four distinct subtypes by MS-based proteomics, which accords well with the qPCR-based classification. Importantly, a distinct *YAP1*-driven subtype with specific proteomic features could be also distinguished. Comprehensive proteomic profiling of these subtypes has uncovered a list of candidate subtype-specific therapeutic vulnerabilities for this once enigmatic cancer. Additionally, we also identified several potential IHC- and blood-based biomarkers that might facilitate subtype diagnosis in the future. Altogether, our results might pave the way for future subtype-specific personalized therapeutic approaches in SCLC.

## CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

## FUNDING INFORMATION

BS is a recipient of the Semmelweis 250+ Excellence PhD Scholarship (EFOP-3.6.3-VEKOP-16-2017-00009). BD, AMS and ZM acknowledge funding from the Hungarian National Research, Development and Innovation Office (KH130356 and KKP126790 to BD; KNN121510 to AMS, 2020-1.1.6-JÖVŐ and TKP2021-EGA-33 to BD and ZM). BD was also supported by the Austrian Science Fund (FWF I3522, FWF I3977 and I4677). ZM was supported by the UNKP-20-3 and UNKP-21-3 New National Excellence Program of the Ministry for Innovation and Technology of Hungary, and by the Hungarian Respiratory Society (MPA #2020), PH was supported by the Netherlands X-omics Initiative (NWO, project 184.034.019). VL is a recipient of the Bolyai Research Scholarship of the Hungarian Academy of Sciences and the UNKP-19-4 New National Excellence Program of the Ministry for Innovation and Technology. KS was supported by the Austrian Science Fund (FWF No. T 1062-B33) and the City of Vienna Fund for Innovative Interdisciplinary Cancer Research. MR acknowledges funding from the Mrs. Berta Kampradt's Cancer Foundation (FBKS-2020-22-(291)).

## DATA AVAILABILITY STATEMENT

The mass spectrometry proteomics data has been deposited to the ProteomeXchange Consortium via the PRIDE<sup>85</sup> partner repository with the dataset identifiers PXD029805 and 10.6019/PXD029805 (cell pellet data), PXD029821 and 10.6019/PXD029821 (culture media data). The scripts for the proteomic data analyses can be obtained at: [https://github.com/bszeitz/SCLC\\_proteomics](https://github.com/bszeitz/SCLC_proteomics).

## ORCID

Beáta Szeitz  <https://orcid.org/0000-0001-6414-0537>

Balázs Döme  <https://orcid.org/0000-0001-8799-8624>

Melinda Rezeli  <https://orcid.org/0000-0003-4373-5616>

## REFERENCES

- Gazdar AF, Bunn PA, Minna JD. Small-cell lung cancer: what we know, what we need to know and the path forward. *Nat Rev Cancer*. 2017;17(12):725-737. doi: [10.1038/nrc.2017.87](https://doi.org/10.1038/nrc.2017.87)
- Kahnert K, Kauffmann-Guerrero D, Huber RM. SCLC-state of the art and what does the future have in store? *Clin Lung Cancer*. 2016;17(5):325-333. doi: [10.1016/j.clcc.2016.05.014](https://doi.org/10.1016/j.clcc.2016.05.014)
- Schwendenwein A, Megyesfalvi Z, Barany N, et al. Molecular profiles of small cell lung cancer subtypes: therapeutic implications. *Mol Ther – Oncolytics*. 2021;20:470-483. doi: [10.1016/j.omto.2021.02.004](https://doi.org/10.1016/j.omto.2021.02.004)
- Rudin CM, Brambilla E, Faivre-Finn C, Sage J. Small-cell lung cancer. *Nat Rev Dis Primers*. 2021;7(1):3. doi: [10.1038/s41572-020-00235-0](https://doi.org/10.1038/s41572-020-00235-0)
- Taniguchi H, Sen T, Rudin CM. Targeted therapies and biomarkers in small cell lung cancer. *Front Oncol*. 2020;10:741. doi: [10.3389/fonc.2020.00741](https://doi.org/10.3389/fonc.2020.00741)
- 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(5):289-297. doi: [10.1038/s41568-019-0133-9](https://doi.org/10.1038/s41568-019-0133-9)
- Zhang W, Girard L, Zhang Y-A, et al. Small cell lung cancer tumors and preclinical models display heterogeneity of neuroendocrine phenotypes. *Transl Lung Cancer Res*. 2018;7:32-49. doi: [10.21037/tlcr.2018.02.02](https://doi.org/10.21037/tlcr.2018.02.02)
- George J, Lim JS, Jang SJ, et al. Comprehensive genomic profiles of small cell lung cancer. *Nature*. 2015;524(7563):47-53. doi: [10.1038/nature14664](https://doi.org/10.1038/nature14664)
- Megyesfalvi Z, Barany N, Lantos A, et al. Expression patterns and prognostic relevance of subtype-specific transcription factors in surgically resected small-cell lung cancer: an international multicenter study. *J Pathol*. 2022;257(5):674-686. doi: [10.1002/path.5922](https://doi.org/10.1002/path.5922)
- Drapkin BJ, Rudin CM. Advances in small-cell lung cancer (SCLC) translational research. *Cold Spring Harb Perspect Med*. 2021;11(4):a038240. doi: [10.1101/cshperspect.a038240](https://doi.org/10.1101/cshperspect.a038240)
- Caeser R, Egger JV, Chavan S, et al. Genomic and transcriptomic analysis of a library of small cell lung cancer patient-derived xenografts. *Nat Commun*. 2022;13(1):2144. doi: [10.1038/s41467-022-29794-4](https://doi.org/10.1038/s41467-022-29794-4)
- Baine MK, Hsieh M-S, Lai WV, et al. SCLC Subtypes Defined by ASCL1, NEUROD1, POU2F3, and YAP1: a comprehensive immunohistochemical and histopathologic characterization. *J*

- Thorac Oncol.* 2020;15(12):1823-1835. doi: [10.1016/j.jtho.2020.09.009](https://doi.org/10.1016/j.jtho.2020.09.009)
13. Gay CM, Stewart CA, Park EM, et al. Patterns of transcription factor programs and immune pathway activation define four major subtypes of SCLC with distinct therapeutic vulnerabilities. *Cancer Cell.* 2021;39(3):346-360. doi: [10.1016/j.ccell.2020.12.014](https://doi.org/10.1016/j.ccell.2020.12.014)
  14. Macklin A, Khan S, Kislinger T. Recent advances in mass spectrometry based clinical proteomics: applications to cancer research. *Clin Proteomics.* 2020;17(1):17. doi: [10.1186/s12014-020-09283-w](https://doi.org/10.1186/s12014-020-09283-w)
  15. Kelemen O, Pla I, Sanchez A, et al. Proteomic analysis enables distinction of early- versus advanced-stage lung adenocarcinomas. *Clin Transl Med.* 2020;10(2):e106. doi: [10.1002/ctm2.106](https://doi.org/10.1002/ctm2.106)
  16. Gillette MA, Satpathy S, Cao S, et al. Proteogenomic characterization reveals therapeutic vulnerabilities in lung adenocarcinoma. *Cell.* 2020;182(1):200-225. doi: [10.1016/j.cell.2020.06.013](https://doi.org/10.1016/j.cell.2020.06.013)
  17. Satpathy S, Krug K, Jean Beltran PM, et al. A proteogenomic portrait of lung squamous cell carcinoma. *Cell.* 2021;184(16):4348-4371.e40. doi: [10.1016/j.cell.2021.07.016](https://doi.org/10.1016/j.cell.2021.07.016)
  18. Jeong H-C, Kim G-I, Cho S-H, et al. Proteomic analysis of human small cell lung cancer tissues: up-regulation of coactosin-like protein-1. *J Proteome Res.* 2011;10(1):269-276. doi: [10.1021/pr100714b](https://doi.org/10.1021/pr100714b)
  19. Byers LA, Wang J, Nilsson MB, et al. Proteomic profiling identifies dysregulated pathways in small cell lung cancer and novel therapeutic targets including PARP1. *Cancer Discov.* 2012;2(9):798-811. doi: [10.1158/2159-8290.CD-12-0112](https://doi.org/10.1158/2159-8290.CD-12-0112)
  20. Fujii K, Miyata Y, Takahashi I, et al. Differential proteomic analysis between small cell lung carcinoma (SCLC) and pulmonary carcinoid tumors reveals molecular signatures for malignancy in lung cancer. *Proteomics Clin Appl.* 2018;12(6):e1800015. doi: [10.1002/prca.201800015](https://doi.org/10.1002/prca.201800015)
  21. Gonçalves E, Poulos RC, Cai Z, et al. Pan-cancer proteomic map of 949 human cell lines. *Cancer Cell.* 2022;40(8):835-849. doi: [10.1016/j.ccell.2022.06.010](https://doi.org/10.1016/j.ccell.2022.06.010)
  22. Chou T-C, Talalay P. Quantitative analysis of dose-effect relationships: the combined effects of multiple drugs or enzyme inhibitors. *Adv Enzyme Regul.* 1984;22:27-55. doi: [10.1016/0065-2571\(84\)90007-4](https://doi.org/10.1016/0065-2571(84)90007-4)
  23. Kuras M, Woldmar N, Kim Y, et al. Proteomic workflows for high-quality quantitative proteome and post-translational modification analysis of clinically relevant samples from formalin-fixed paraffin-embedded archives. *J Proteome Res.* 2021;20(1):1027-1039. doi: [10.1021/acs.jproteome.0c00850](https://doi.org/10.1021/acs.jproteome.0c00850)
  24. R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing; 2022. <https://www.R-project.org/> (accessed on 4 May 2022)
  25. Tyanova S, Temu T, Sinitcyn P, et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods.* 2016;13:731-740. doi: [10.1038/nmeth.3901](https://doi.org/10.1038/nmeth.3901)
  26. Uhlén M, Karlsson MJ, Hober A, et al. The human secretome. *Sci Signal.* 2019;12(609):eaaz0274. doi: [10.1126/scisignal.aaz0274](https://doi.org/10.1126/scisignal.aaz0274)
  27. Chen G, Chen J, Liu H, et al. Comprehensive identification and characterization of human secretome based on integrative proteomic and transcriptomic data. *Front Cell Dev Biol.* 2019;7:299. doi: [10.3389/fcell.2019.00299](https://doi.org/10.3389/fcell.2019.00299)
  28. Meinken J, Walker G, Cooper CR, Min XJ. MetazSecKB: the human and animal secretome and subcellular proteome knowledgebase. *Database (Oxford).* 2015;2015:bav077. doi: [10.1093/database/bav077](https://doi.org/10.1093/database/bav077)
  29. Hu Z, Yuan J, Long M, et al. The Cancer Surfaceome Atlas integrates genomic, functional and drug response data to identify actionable targets. *Nat Cancer.* 2021;2(12):1406-1422. doi: [10.1038/s43018-021-00282-w](https://doi.org/10.1038/s43018-021-00282-w)
  30. Bausch-Fluck D, Goldmann U, Müller S, et al. The in silico human surfaceome. *Proc Natl Acad Sci USA.* 2018;115(46):E10988-E10997. doi: [10.1073/pnas.1808790115](https://doi.org/10.1073/pnas.1808790115)
  31. The UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.* 2021;49(D1):D480-D489. doi: [10.1093/nar/gkaa1100](https://doi.org/10.1093/nar/gkaa1100)
  32. Wishart DS, Feunang YD, Guo AC, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* 2018;46(D1):D1074-D1082. doi: [10.1093/nar/gkx1037](https://doi.org/10.1093/nar/gkx1037)
  33. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* 2012;2(5):401-404. doi: [10.1158/2159-8290.CD-12-0095](https://doi.org/10.1158/2159-8290.CD-12-0095)
  34. Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal.* 2013;6(269):pl1. doi: [10.1126/scisignal.2004088](https://doi.org/10.1126/scisignal.2004088)
  35. Ghandi M, Huang FW, Jané-Valbuena J, et al. Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature.* 2019;569(7757):503-508. doi: [10.1038/s41586-019-1186-3](https://doi.org/10.1038/s41586-019-1186-3)
  36. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43(7):e47-e47. doi: [10.1093/nar/gkv007](https://doi.org/10.1093/nar/gkv007)
  37. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 2010;26(1):139-140. doi: [10.1093/bioinformatics/btp616](https://doi.org/10.1093/bioinformatics/btp616)
  38. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* 2012;40(10):4288-4297. doi: [10.1093/nar/gks042](https://doi.org/10.1093/nar/gks042)
  39. Chen Y, Lun ATL, Smyth GK. From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline [version 2; peer review: 5 approved]. *F1000Research.* 2016;5:1438. doi: [10.12688/F1000RESEARCH.8987.2](https://doi.org/10.12688/F1000RESEARCH.8987.2)
  40. Yang W, Soares J, Greninger P, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res.* 2013;41(D1):D955-D961. doi: [10.1093/nar/gks1111](https://doi.org/10.1093/nar/gks1111)
  41. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS: J Integr Biol.* 2012;16(5):284-287. doi: [10.1089/omi.2011.0118](https://doi.org/10.1089/omi.2011.0118)
  42. Wu T, Hu E, Xu S, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation.* 2021;2(3):100141. doi: [10.1016/j.xinn.2021.100141](https://doi.org/10.1016/j.xinn.2021.100141)
  43. Yu G, He Q-Y. ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. *Mol Biosyst.* 2016;12(2):477-479. doi: [10.1039/C5MB00663E](https://doi.org/10.1039/C5MB00663E)
  44. Kohn KW, Zeeberg BM, Reinhold WC, Pommier Y. Gene expression correlations in human cancer cell lines define



- molecular interaction networks for epithelial phenotype. *PLoS One*. 2014;9(6):e99269. doi: [10.1371/journal.pone.0099269](https://doi.org/10.1371/journal.pone.0099269)
45. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics*. 2010;26(12):1572-1573. doi: [10.1093/bioinformatics/btq170](https://doi.org/10.1093/bioinformatics/btq170)
  46. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015;1(6):417-425. doi: [10.1016/j.cels.2015.12.004](https://doi.org/10.1016/j.cels.2015.12.004)
  47. Kanehisa M, Goto SKEGG. Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27-30. doi: [10.1093/nar/28.1.27](https://doi.org/10.1093/nar/28.1.27)
  48. Jassal B, Matthews L, Viteri G, et al. The reactome pathway knowledgebase. *Nucleic Acids Res*. 2020;48(D1):D498-D503. doi: [10.1093/nar/gkz1031](https://doi.org/10.1093/nar/gkz1031)
  49. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. 2000;25(1):25-29. doi: [10.1038/75556](https://doi.org/10.1038/75556)
  50. Gene Ontology Consortium. The gene ontology resource: enriching a GOld mine. *Nucleic Acids Res*. 2021;49(D1):D325-D334. doi: [10.1093/nar/gkaa1113](https://doi.org/10.1093/nar/gkaa1113)
  51. 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 USA*. 2005;102(43):15545-15550. doi: [10.1073/pnas.0506580102](https://doi.org/10.1073/pnas.0506580102)
  52. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. *Bioinformatics*. 2011;27(12):1739-1740. doi: [10.1093/bioinformatics/btr260](https://doi.org/10.1093/bioinformatics/btr260)
  53. Barbie DA, Tamayo P, Boehm JS, et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature*. 2009;462(7269):108-112. doi: [10.1038/nature08460](https://doi.org/10.1038/nature08460)
  54. Rohart F, Gautier B, Singh A, Lê Cao K-A. mixOmics: an R package for 'omics feature selection and multiple data integration. *PLoS Comput Biol*. 2017;13(11):e1005752. doi: [10.1371/journal.pcbi.1005752](https://doi.org/10.1371/journal.pcbi.1005752)
  55. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*. 2016;32(18):2847-2849. doi: [10.1093/bioinformatics/btw313](https://doi.org/10.1093/bioinformatics/btw313)
  56. Denny SK, Yang D, Chuang C-H, et al. Nfib promotes metastasis through a widespread increase in chromatin accessibility. *Cell*. 2016;166(2):328-342. doi: [10.1016/j.cell.2016.05.052](https://doi.org/10.1016/j.cell.2016.05.052)
  57. Mori N, Yokota J, Akiyama T, et al. Variable mutations of the RB gene in small-cell lung carcinoma. *Oncogene*. 1990;5(11):1713-1717.
  58. Wistuba II, Gazdar AF, Minna JD. Molecular genetics of small cell lung carcinoma. *Semin Oncol*. 2001;28(2)(suppl 4):3-13.
  59. Arriola E, Cañadas I, Arumí M, Rojo F, Rovira A, Albanell J. Genetic changes in small cell lung carcinoma. *Clin Transl Oncol*. 2008;10(4):189-197. doi: [10.1007/s12094-008-0181-1](https://doi.org/10.1007/s12094-008-0181-1)
  60. Sabari JK, Lok BH, Laird JH, Poirier JT, Rudin CM. Unraveling the biology of SCLC: implications for therapy. *Nat Rev Clin Oncol*. 2017;14(9):549-561. doi: [10.1038/nrclinonc.2017.71](https://doi.org/10.1038/nrclinonc.2017.71)
  61. Uhlén M, Fagerberg L, Hallström BM, et al. Tissue-based map of the human proteome. *Science (80-)*. 2015;347(6220):1260419. doi: [10.1126/science.1260419](https://doi.org/10.1126/science.1260419)
  62. Carney DN, Gazdar AF, Bepler G, et al. Establishment and identification of small cell lung cancer cell lines having classic and variant features. *Cancer Res*. 1985;45(6):2913-2923.
  63. Gazdar AF, Carney DN, Nau MM, Minna JD. Characterization of variant subclasses of cell lines derived from small cell lung cancer having distinctive biochemical, morphological, and growth properties. *Cancer Res*. 1985;45(6):2924-2930.
  64. Baine MK, Febres-Aldana CA, Chang JC, et al. POU2F3 in SCLC: clinicopathologic and genomic analysis with a focus on its diagnostic utility in neuroendocrine-low SCLC. *J Thorac Oncol*. 2022;17:1109-1121. doi: [10.1016/j.jtho.2022.06.004](https://doi.org/10.1016/j.jtho.2022.06.004)
  65. Tlemsani C, Pongor L, Elloumi F, et al. SCLC-CellMiner: a resource for small cell lung cancer cell line genomics and pharmacology based on genomic signatures. *Cell Rep*. 2020;33(3):108296. doi: [10.1016/j.celrep.2020.108296](https://doi.org/10.1016/j.celrep.2020.108296)
  66. Guillemot F, Hassan BA. Beyond proneural: emerging functions and regulations of proneural proteins. *Curr Opin Neurobiol*. 2017;42:93-101. doi: [10.1016/j.conb.2016.11.011](https://doi.org/10.1016/j.conb.2016.11.011)
  67. Cargill KR, Stewart CA, Park EM, et al. Targeting MYC-enhanced glycolysis for the treatment of small cell lung cancer. *Cancer Metab*. 2021;9(1):33. doi: [10.1186/s40170-021-00270-9](https://doi.org/10.1186/s40170-021-00270-9)
  68. Sica V, Bravo-San Pedro JM, Stoll G, Kroemer G. Oxidative phosphorylation as a potential therapeutic target for cancer therapy. *Int J Cancer*. 2020;146(1):10-17. doi: [10.1002/ijc.32616](https://doi.org/10.1002/ijc.32616)
  69. Liao Y, Yin G, Wang X, Zhong P, Fan X, Huang C. Identification of candidate genes associated with the pathogenesis of small cell lung cancer via integrated bioinformatics analysis. *Oncol Lett*. 2019;18(4):3723-3733. doi: [10.3892/ol.2019.10685](https://doi.org/10.3892/ol.2019.10685)
  70. Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol*. 2014;15(3):178-196. doi: [10.1038/nrm3758](https://doi.org/10.1038/nrm3758)
  71. Hwang J-J, Choi S-Y, Koh J-Y. The role of NADPH oxidase, neuronal nitric oxide synthase and poly(ADP ribose) polymerase in oxidative neuronal death induced in cortical cultures by brain-derived neurotrophic factor and neurotrophin-4/5. *J Neurochem*. 2002;82(4):894-902. doi: [10.1046/j.1471-4159.2002.01040.x](https://doi.org/10.1046/j.1471-4159.2002.01040.x)
  72. Griffin N, Faulkner S, Jobling P, Hondermarck H. Targeting neurotrophin signaling in cancer: the renaissance. *Pharmacol Res*. 2018;135:12-17. doi: [10.1016/j.phrs.2018.07.019](https://doi.org/10.1016/j.phrs.2018.07.019)
  73. Dupont S, Morsut L, Aragona M, et al. Role of YAP/TAZ in mechanotransduction. *Nature*. 2011;474(7350):179-183. doi: [10.1038/nature10137](https://doi.org/10.1038/nature10137)
  74. Caesar R, Hulton C, Costa E, et al. MAPK pathway activation selectively inhibits ASCL1-driven small cell lung cancer. *iScience*. 2021;24(11):103224. doi: [10.1016/j.isci.2021.103224](https://doi.org/10.1016/j.isci.2021.103224)
  75. Chemi F, Pearce SP, Clipson A, et al. cfDNA methylome profiling for detection and subtyping of small cell lung cancer. *Nat Cancer*. 2022. doi: [10.1038/s43018-022-00415-9](https://doi.org/10.1038/s43018-022-00415-9)
  76. Horie M, Saito A, Ohshima M, Suzuki HI, Nagase T. YAP and TAZ modulate cell phenotype in a subset of small cell lung cancer. *Cancer Sci*. 2016;107(12):1755-1766. doi: [10.1111/cas.13078](https://doi.org/10.1111/cas.13078)
  77. Paolucci L, Rozengurt E. Protein kinase D in small cell lung cancer cells: rapid activation through protein kinase C. *Cancer Res*. 1999;59(3):572-577.
  78. Byers LA, Diao L, Wang J, et al. An epithelial-mesenchymal transition gene signature predicts resistance to EGFR and PI3K inhibitors and identifies Axl as a therapeutic target for overcoming EGFR inhibitor resistance. *Clin Cancer Res*. 2013;19(1):279-290. doi: [10.1158/1078-0432.CCR-12-1558](https://doi.org/10.1158/1078-0432.CCR-12-1558)

79. Borromeo MD, Savage TK, Kollipara RK, et al. ASCL1 and NEUROD1 reveal heterogeneity in pulmonary neuroendocrine tumors and regulate distinct genetic programs. *Cell Rep*. 2016;16(5):1259-1272. doi: [10.1016/j.celrep.2016.06.081](https://doi.org/10.1016/j.celrep.2016.06.081)
80. Sethi T, Rintoul RC, Moore SM, et al. Extracellular matrix proteins protect small cell lung cancer cells against apoptosis: a mechanism for small cell lung cancer growth and drug resistance in vivo. *Nat Med*. 1999;5(6):662-668. doi: [10.1038/9511](https://doi.org/10.1038/9511)
81. Zhao P, Sun X, Li H, et al. c-Myc targets HDAC3 to suppress NKG2DL expression and innate immune response in N-Type SCLC through histone deacetylation. *Cancers (Basel)*. 2022;14(3). doi: [10.3390/cancers14030457](https://doi.org/10.3390/cancers14030457)
82. Chan JM, Quintanal-Villalonga Á, Gao VR, et al. Signatures of plasticity, metastasis, and immunosuppression in an atlas of human small cell lung cancer. *Cancer Cell*. 2021;39(11):1479-1496. doi: [10.1016/j.ccell.2021.09.008](https://doi.org/10.1016/j.ccell.2021.09.008)
83. Gazdar AF, Gao B, Minna JD. Lung cancer cell lines: useless artifacts or invaluable tools for medical science? *Lung Cancer*. 2010;68(3):309-318. doi: [10.1016/j.lungcan.2009.12.005](https://doi.org/10.1016/j.lungcan.2009.12.005)
84. van Staveren WCG, Solís DYW, Hébrant A, Detours V, Dumont JE, Maenhaut C. Human cancer cell lines: experimental models for cancer cells in situ? For cancer stem cells? *Biochim Biophys Acta*. 2009;1795(2):92-103. doi: [10.1016/j.bbcan.2008.12.004](https://doi.org/10.1016/j.bbcan.2008.12.004)
85. Perez-Riverol Y, Csordas A, Bai J, et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res*. 2019;47(D1):D442-D450. doi: [10.1093/nar/gky1106](https://doi.org/10.1093/nar/gky1106)

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Szeitz B, Megyesfalvi Z, Woldmar N, et al. In-depth proteomic analysis reveals unique subtype-specific signatures in human small-cell lung cancer. *Clin Transl Med*. 2022;12:e1060. <https://doi.org/10.1002/ctm2.1060>