



Article

# Signatures of Co-Deregulated Genes and Their Transcriptional Regulators in Lung Cancer

Angeliki Chatziantoniou <sup>1,2</sup> and Apostolos Zaravinos <sup>1,2,\*</sup>

<sup>1</sup> Department of Life Sciences, School of Sciences, European University Cyprus, Nicosia 1516, Cyprus

<sup>2</sup> Cancer Genetics, Genomics and Systems Biology Laboratory, Basic and Translational Cancer Research Center (BTCRC), Nicosia 1516, Cyprus

\* Correspondence: a.zaravinos@euc.ac.cy

**Abstract:** Despite the significant progress made towards comprehending the deregulated signatures in lung cancer, these vary from study to study. We reanalyzed 25 studies from the Gene Expression Omnibus (GEO) to detect and annotate co-deregulated signatures in lung cancer and in single-gene or single-drug perturbation experiments. We aimed to decipher the networks that these co-deregulated genes (co-DEGs) form along with their upstream regulators. Differential expression and upstream regulators were computed using Characteristic Direction and Systems Biology tools, including GEO2Enrichr and X2K. Co-deregulated gene expression profiles were further validated across different molecular and immune subtypes in lung adenocarcinoma (TCGA-LUAD) and lung adenocarcinoma (TCGA-LUSC) datasets, as well as using immunohistochemistry data from the Human Protein Atlas, before being subjected to subsequent GO and KEGG enrichment analysis. The functional alterations of the co-upregulated genes in lung cancer were mostly related to immune response regulating the cell surface signaling pathway, in contrast to the co-downregulated genes, which were related to S-nitrosylation. Networks of hub proteins across the co-DEGs consisted of overlapping TFs (SOX2, MYC, KAT2A) and kinases (MAPK14, CSNK2A1 and CDKs). Furthermore, using Connectivity Map we highlighted putative repurposing drugs, including valproic acid, benticine and astemizole. Similarly, we analyzed the co-DEG signatures in single-gene and single-drug perturbation experiments in lung cancer cell lines. In summary, we identified critical co-DEGs in lung cancer providing an innovative framework for their potential use in developing personalized therapeutic strategies.

**Keywords:** lung cancer; tumor heterogeneity; Characteristic Direction; co-deregulated genes; single-gene perturbation; single-drug perturbation; drug repurposing; GEO2Enrichr; X2K



**Citation:** Chatziantoniou, A.; Zaravinos, A. Signatures of Co-Deregulated Genes and Their Transcriptional Regulators in Lung Cancer. *Int. J. Mol. Sci.* **2022**, *23*, 10933. <https://doi.org/10.3390/ijms231810933>

Academic Editor: Robert Arthur Kratzke

Received: 6 September 2022

Accepted: 13 September 2022

Published: 18 September 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Lung cancer (LC) maintains the highest mortality rate among cancer-related deaths, with almost an equal distribution between females and males [1]. Molecularly diverse subtypes of lung cancer have been investigated and proposed, but the therapeutic outcome in these patients is low, as a result of drug resistance [2]. The gradual accumulation of genetic and epigenetic alterations, as well as environmental factors, destabilizes the DNA, and leads to an abnormal gene expression stemming from the coordinated deregulation of transcription factors (TFs). In addition, mutations in genes encoding protein kinases are critical for the onset of carcinogenesis.

Thus far, most efforts focus on the detection of differentially expressed genes (DEGs) between cancerous and normal lung tissue. Nevertheless, the reported DEGs vary from study to study, depending on biological differences in the profiled samples (e.g., sex, mutation status, subtype, stage, etc.), as well as differences in sample numbers or the computational methodology followed [3]. One intriguing issue not yet contemplated is the analysis of the genes and signaling pathways being simultaneously deregulated across

molecularly heterogeneous subtypes in lung cancer, including lung adenocarcinomas and squamous cell carcinomas. Challenging as it may seem, the analysis of pooled, diverse datasets of lung cancer could reveal the commonly deregulated signaling pathways across these tumor entities. Such information could be very useful for the treatment of what really should be viewed as many different diseases.

Extracting large-scale data from the Gene Expression Omnibus (GEO) offers extensive possibilities of simultaneously managing and analyzing multiple gene signatures. Collecting gene expression profiles of co-deregulated genes (co-DEGs) at the transcriptional level can help us understand the complete activation and overlapping of molecularly diverse oncogenic paths. Abnormal patterns of gene expression can also be used for the detection of disease biomarkers, as well as for therapeutic purposes.

There are many genes being differentially expressed across distinct histological or molecular subtypes in lung cancer [4–11]. Nevertheless, the sensitivity and specificity of these biomarkers is not always necessarily adequate, urging the need to identify an updated panel of genes that can be used as better diagnostic and preventive biomarkers, or even as therapeutic targets. In addition, the co-DEGs across distinct lung tumors vary and, to our knowledge, their role within signaling networks or their transcriptional regulatory mechanisms has been poorly investigated. A recent new multivariate method called the Characteristic Direction (CD), can be used to compute signatures using the orientation of the separating hyperplane from a linear classification scheme to define a direction that characterizes differential expression [12].

The purpose of this study was to use the CD method to identify co-DEGs across various independent datasets in lung cancer. We used GEO datasets to annotate and extract gene expression signatures, and validated our results in the TCGA-LUAD and TCGA-LUSC datasets, as well as in the Human Protein Atlas, aiming to better understand the links between co-deregulated genes, drugs and lung cancer. Importantly, we aimed to identify the “hubs” in the gene networks composed of transcription factors and protein kinases differentially expressed in lung cancer. We also intended to identify drugs (or drug combinations) targeting these hubs. Such drugs could thus be used as new and effective treatment regimes. The lists of co-deregulated gene signatures that we propose can provide further insights into lung carcinogenesis.

## 2. Results

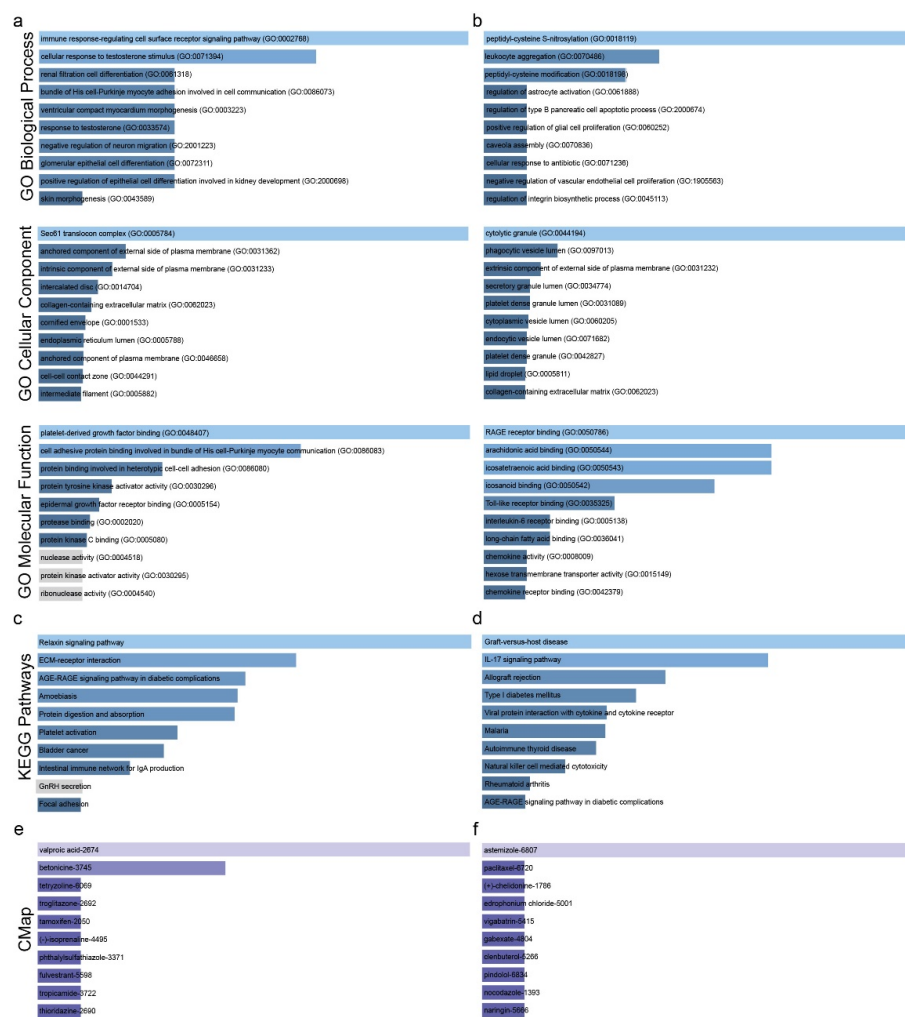
Using a Systems Biology approach, we extracted the co-deregulated signatures (i.e., signatures containing the same deregulated genes present in at least two independent studies) from 25 independent GEO datasets and classified them as: (1) lung cancer vs. healthy tissue; lung cancer (tissue or cell lines) with a single gene (2); or single-drug (3) perturbation (Table S1). We asked whether signature similarity within and across these three categories could recover prior knowledge and discover new connections. To globally assess associations between signatures within each category, we computed the signatures using the CD method, and compared ranked signature associations with prior knowledge. We concluded such lists of co-deregulated genes in each category after stringent filtering and excluding the genes that were deregulated in a single study. These lists contained 20 co-upregulated and 25 co-downregulated genes in LC versus the normal tissue; 333 co-upregulated and 528 co-downregulated genes after single-gene perturbation; and 459 co-upregulated and 439 co-downregulated genes after single-drug perturbation (Table S1). In each category we identified the TFs, PPIs, and kinases accountable for the observed changes in the mRNA expression of these co-DEGs, the drugs that suppress or induce these co-DEGs, respectively, and finally the biological pathways in which the abovementioned genes are involved.

Next, we sought to identify which of these upstream regulators are most probably responsible for the deregulation in the expression of the identified gene lists in lung cancers. To this end, in each category we identified the phosphorylation reactions possibly being carried out by upstream regulatory kinases. We also investigated the drugs that suppress over-expressed genes, or those that induce the expression of under-expressed ones. The

top TFs and protein kinases from each category of co-DEGs were classified based on the highest value of a combined score of the  $p$ -value and the  $z$ -score (Figure S1).

### 2.1. Co-Deregulated Genes in Lung Cancers

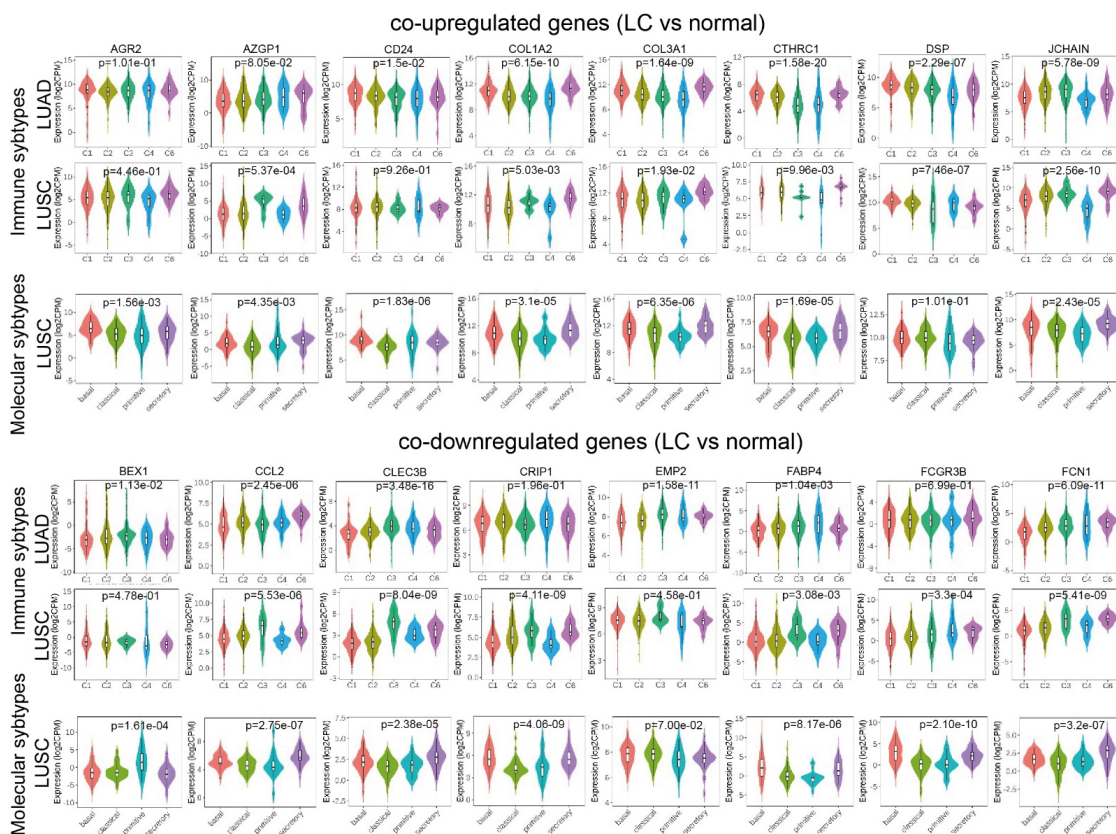
We analyzed two GEO datasets to detect the co-DEGs in lung cancer samples against their adjacent healthy tissue (Table S3). The co-upregulated genes were mainly enriched in the “immune response-regulating cell surface receptor signaling pathway” (GO biological process) and the “Sec61 translocon complex” (GO cellular component), as well as “platelet-derived growth factor binding” and “cell adhesive protein binding involved in bundle of His cell–Purkinje myocyte communication” (GO molecular function) (Figure 1a and Table S3).



**Figure 1.** GO enrichment of the co-upregulated (a) and the co-downregulated (b) genes in lung cancer vs. the normal tissue. KEGG enrichment for the co-upregulated (c) and the co-downregulated (d) genes in lung cancer vs. the normal tissue. Repurposing drugs (CMap) targeting the co-up- (e) and co-downregulated (f) genes in lung cancer. Bar graphs were sorted using a combined score between the Benjamini–Hochberg (BH)-adjusted  $p$ -value and the  $z$ -score of the deviation from the expected rank. Each bar’s length shows the significance of the corresponding term, being relative to the color brightness of each bar.

We further explored differences in the expression of individual genes within the top enriched terms of the co-up- or co-downregulated genes across different immune and molecular subtypes in lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) datasets from the Cancer Genome Atlas (TCGA), respectively. Specifically, the

immune subtypes we explored were: C1 (wound healing); C2 (IFN-gamma dominant); C3 (inflammatory); C4 (lymphocyte depleted); C5 (immunologically quiet); and C6 (TGF-b dominant), as previously described by Thorsson et al. [13]. The molecular subtypes of LUSC contained basal, classical, primitive, and secretory lung tumors. Among the top co-upregulated genes, we examined “anterior gradient 2, protein disulphide isomerase family member” (AGR2), “alpha-2-glycoprotein 1, zinc binding” (AZGP1), “small cell lung carcinoma cluster 4 antigen” (CD24), “collagen, type I, alpha 2” (COL1A2), “collagen, type III, alpha 1” (COL3A1), “collagen triple-helix repeat containing 1” (CTHRC1), “desmoplakin” (DSP), and “joining chain of multimeric IgA and IgM” (JCHAIN). Similarly, among the top co-downregulated genes, we explored “brain-expressed, X-linked 1” (BEX1); “chemokine (C-C motif) ligand 2” (CCL2); “C-type lectin domain family 3, member B” (CLEC3B); “cysteine-rich intestinal protein 1” (CRIP1); “epithelial membrane protein 2 (EMP2)”; “fatty acid-binding protein 4, adipocyte (FABP4)”; “Fc fragment of IgG, low-affinity IIIb, receptor (CD16b)” (FCGR3B) and “ficolin (collagen/fibrinogen domain containing) 1” (FCN1). Interestingly, we found significant differences in the expression of these genes across the different molecular and immune subtypes in lung cancer. For example, AGR2, CD24, COL3A1, CTHRC1, DSP and JCHAIN were upregulated in basal LUSC tumors. Similarly, AZGP1, COL1A2, COL3A1 and CTHRC1 were upregulated in “TGF-b dominant” (C6) LUSC samples (Figure 2).



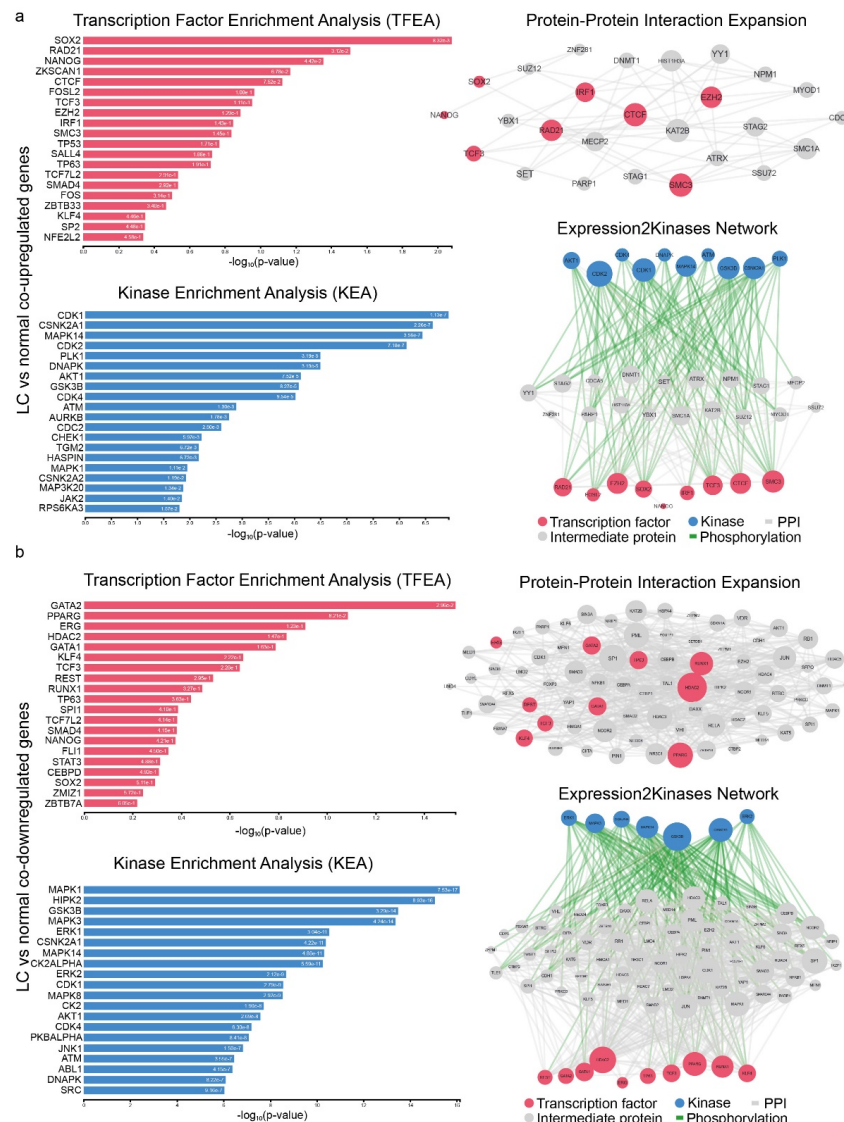
**Figure 2.** Expression of the top co-up- or co-downregulated genes across different immune and molecular subtypes of lung adenocarcinoma (LUAD) or lung squamous cell carcinoma (LUSC). The immune subtypes were: C1 (wound healing); C2 (IFN-gamma dominant); C3 (inflammatory); C4 (lymphocyte depleted); C5 (immunologically quiet) and C6 (TGF-b dominant), as defined in Thorsson et al. [13]. The molecular subtypes of LUSC contained basal, classical, primitive, secretory tumors. LUAD immune subtypes, n=C1,83;C2,147;C3,179;C4,20;C6,28. LUSC immune subtypes, n = C1,275;C2,182;C3,8;C4,7;C6,14. LUSC molecular subtypes, n=basal 42, classical 63, primitive 26, secretory 39.



The co-downregulated genes, on the other hand, were mainly enriched in “peptidyl-cysteine S-nitrosylation”, “leukocyte aggregation” (GO biological process) (Figure 1b and Table S3), “cytolytic granule” (GO cellular component), “RAGE receptor binding”, “arachidonic acid binding”, and “icosatetraenoic acid binding” (GO molecular function).

KEGG enrichment analysis highlighted the involvement of the co-upregulated genes in the pathways: “Relaxin signaling pathway” and “ECM–receptor interaction” (Figure 1c and Table S4). As for the co-downregulated genes, these were overrepresented in “Graft-versus-host disease”, “IL-17 signaling pathway” and “Allograft rejection” (Figure 1d and Table S4).

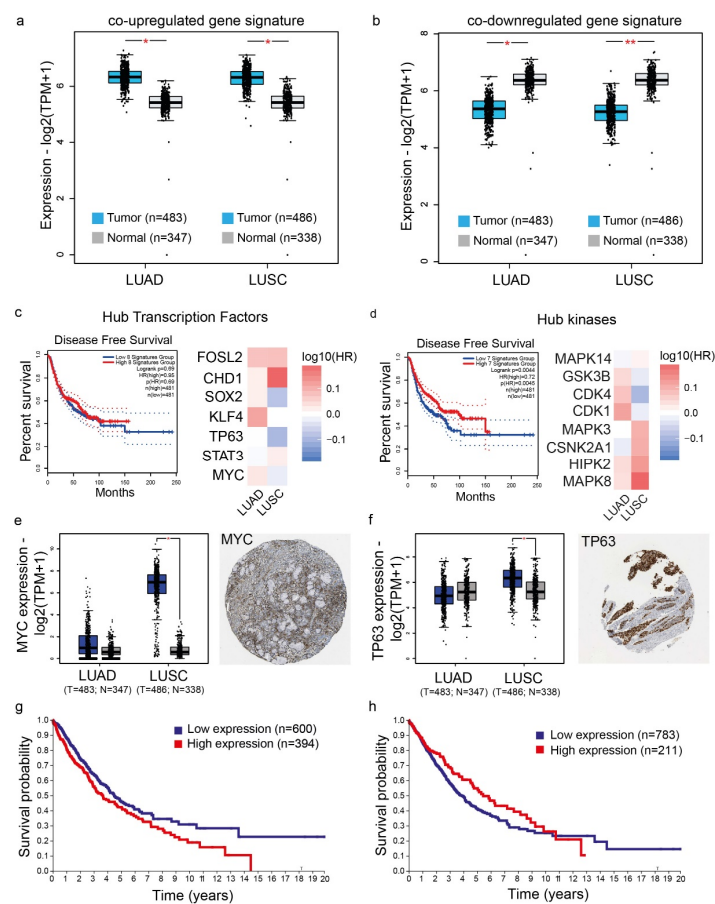
We then constructed the PPI networks containing the critical hub proteins in lung carcinogenesis, as they potentially regulate the expression of the co-DEGs. The responsible upstream regulators for the co-upregulated genes involved CTCF, EZH2, TCF3, SOX2, RAD21, FOSL2, IRF1 and SMC3 (TFs) and CDK1/2, GSK3B, PLK1, AKT1, CDK4, DNAPK, ATM, CSNK2A1 and MAPK14 (kinases) (Figure 3a and Table S5). The co-downregulated genes on the other hand, were significantly associated with the TFs HDAC2, PPARC and RUNX1 and the kinases GSK3B, MAPK14, MAPK3, ERK1/2, CSNK2A1 and CK2ALPHA (Figure 3b and Table S6).



**Figure 3.** Upstream regulatory networks for co-upregulated (a) and co-downregulated (b) gene signatures in lung cancer vs. the normal tissue. The networks depict transcription factors (TFs, red nodes), intermediate

proteins (gray nodes), and kinases (blue nodes). Gray edges indicate PPI interactions and green edges depict kinase-driven phosphorylation events. Node size is relative to expression. Upstream regulatory networks were constructed using X2K.

Apart from selected genes, we also successfully validated the co-deregulated gene signatures in lung cancer (termed “UP genes” and “DOWN genes”, respectively) using the TCGA-LUAD and TCGA-LUSC datasets. Both gene expression signatures confirmed their significant difference in the LUAD and LUSC TCGA datasets, compared to the control tissues. As controls, we used normal lung samples from both the TCGA and GTEx projects (Figure 4a,b).



**Figure 4.** The expression patterns of the co-upregulated (a) and co-downregulated (b) gene signatures were verified in the TCGA-LUAD and TCGA-LUSC datasets, respectively (LUAD, 483 lung adenocarcinomas (T) and 347 normal (N) samples; LUSC, 486 lung squamous cell carcinomas (T) and 338 normal (N) samples). The significantly elevated expressions of the major hub transcription factors MYC, STAT3, TP63 and KLF4 were verified among LUAD and LUSC, respectively. The upregulated levels of the hub kinases CDK1, GSK3B, CSNK2A1 and MAPK14 were also validated in LUAD and LUSC tumors (\*,  $p < 0.001$ ; \*\*,  $p < 0.0001$ ). The Kaplan–Meier curves depict disease-free survival of lung patients with high or low expression in signatures composed of the transcription factors (FOSL2, CHD1, SOX2, KLF4, TP63, STAT3 and MYC) (c) or kinases (MAPK14, GSK3B, CDK4, CDK1, MAPK3, CSNK2A1, HIPK2, MAPK8) (d) acting as main hubs, respectively. High expression of the signature composed of the hub kinases was significantly associated with disease-free survival ( $p < 0.05$ ), but that of the hubs’ TFs was not. MYC and TP63 were significantly upregulated in LUSC (but not LUAD) patients, exhibited moderate protein expression (e,f), and showed a reverse pattern of association with patient survival ( $p \leq 0.05$ , Log-rank) (g,h). The TCGA-LUAD and TCGA-LUSC patient cohorts were analyzed using GEPIA2. IHC data were extracted from the Human Protein Atlas.

In addition, we selected specific hub proteins to verify their gene expression, using GEPIA2 [14]. We further compared disease-free survival in patients with high expression in signatures composed of the hub TFs FOSL2, CHD1, SOX2, KLF4, TP63, STAT3 and MYC (Figure 4c) or the kinases MAPK14, GSK3B, CDK4, CDK1, MAPK3, CSNK2A1, HIPK2, and MAPK8 (Figure 4d) against those with low expression and found that, although high expression of the TF hubs' signatures was not associated with survival, the high expression of the signature composed of the eight hub kinases was significantly associated with disease-free survival ( $p = 0.0044$ , Log-rank). Of these, GSK3B, CDK1 and CDK4 were positively correlated with LUAD patients' disease-free survival (DFS), whereas MAPK3, CSNK2A1, HIPK2 and MAPK8 were correlated with LUSC patients' DFS (Figure 4d).

The TFs MYC and TP63 were among the main hubs, and their upregulated levels were mainly verified in LUSC (Figure 4e,f). MYC stained high in 10% of the cases using IHC data from the Human Protein Atlas and its high expression was an unfavorable marker in lung cancer ( $p = 0.0049$ , Log-rank) (Figure 4g). On the other hand, TP63 stained high in 40% of the cases (Figure 4f) and its high expression correlated with better patient survival ( $p = 0.052$ , Log-rank) (Figure 4h). Similar analyses were performed for other hub proteins, including CEBPB, RUNX1, GATA1, etc.

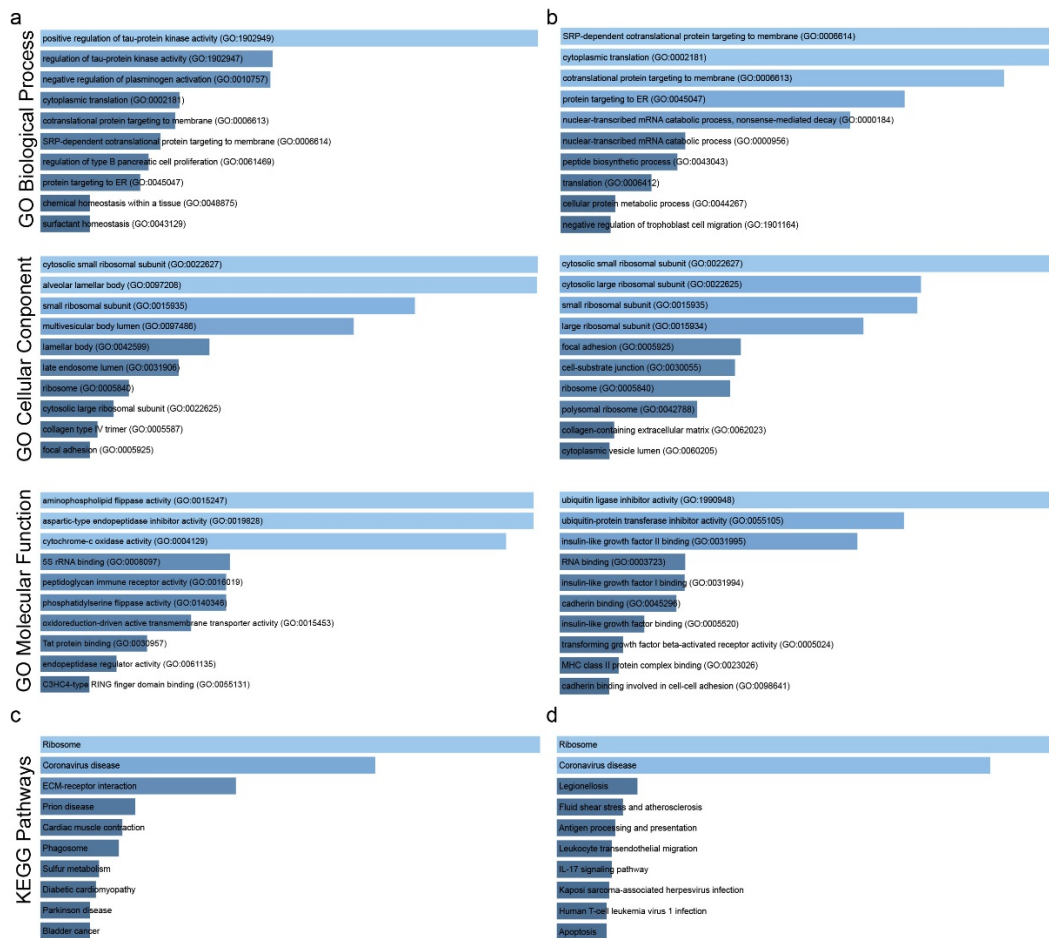
## 2.2. Drug Repurposing against Lung Cancer

Drug repurposing is a strategy for identifying new uses for approved or investigational drugs that are outside the scope of the original medical indication. Here, using Connectivity Map (CMap) and the identified co-DEGs, we found two existing drugs mainly targeting the co-upregulated genes (valproic acid and betonicine) and one drug (astemizole) against the co-downregulated genes (Figure 1e,f and Table S7) [15–19].

## 2.3. Signatures of Co-Deregulated Genes in Single-Gene Perturbation Experiments

We extracted the co-DEG signatures from 11 single-gene perturbation studies that were enriched in biological processes involving “positive regulation of Tau protein kinase activity”, “regulation of Tau protein kinase activity”, “negative regulation of plasminogen activation”, “cytoplasmic translation” and “co-translational protein targeting to membrane” (Figure 5a and Table S8). A remarkable finding is the modulated expression of Tau protein, which normally penetrates to the brain tissue, taking part in the organization of the axial microtubules of nerve cells. This correlated relationship is admittedly anticipated, as similar risk factors such as cell ageing, irregular cell cycle and cell death due to DNA damage turn out to be present in both cancer development and neurological disorders. It has recently been proposed that Tau protein appears in smooth muscle cells of the vascular and airway systems of the lung [15], and there is a positive correlation between Tau protein and the evolution of pulmonary adenocarcinoma [16]. The common cellular pathogenetic events between these two clinically distinct diseases require further study in light of molecular cross-linking given the tremendously high rate of metastatic tendency of lung tumors to the brain [17,18]. Subsequently, these genes co-overexpress in the negative regulation of plasminogen activation, a system that has dissimilar gene expression patterns among lung cancer subtypes [19]. They are also enriched in cellular components such as “cytosolic small ribosomal subunit”, “alveolar lamellar body”, “small ribosomal subunit”, and “multivesicular body lumen”, as well as in operating functions such as “aminophospholipid flippase activity”, “aspartic-type endopeptidase inhibitor activity” and “cytochrome-c oxidase activity”. Aminophospholipid flippases are phospholipid transporters responsible for the formation and maintenance of the cytoplasmic asymmetry of membranes, a unique architecture that contributes to various signaling processes such as apoptosis mediated by phagocytes. A high flippase activity has been found to regulate the low surface exposure of phosphatidylserine residues in human cancer cells [20], an altered pattern of which is responsible for immunological imbalance through silencing key inflammatory signals. In this way, the deregulated action of flippases brings about an immune-escape phenomenon of malignant cells, as well as induces immunosuppressive

phosphatidylserines [21]. Aspartic-type endopeptidase inhibitor activity assists in the hydrolysis of a peptide bond in the side chains of aspartic acid so that a nucleophilic character is acquired. Dysfunctionality arises from an endogenous abnormal feature of carcinogenesis to disorientate transcription factors from functioning properly.



**Figure 5.** Gene Ontology (GO) enrichment analysis of the co-DEGs in single-gene perturbation experiments in lung cancer. **(a)** GO enrichment results of the co-upregulated genes in lung cancer against the normal tissue. **(b)** GO enrichment results of the co-downregulated genes in lung cancer against the normal tissue. KEGG enrichment analysis regarding the co-DEGs in single-gene perturbation experiments in lung cancer. **(c)** KEGG enrichment for the co-upregulated genes and **(d)** the co-downregulated genes. Bar graphs were sorted using a combined score between the Benjamini–Hochberg (BH)-adjusted  $p$ -value and the  $z$ -score of the deviation from the expected rank. Each bar’s length shows the significance of the corresponding term being relative to the color brightness of each bar.

Our KEGG analysis revealed the enrichment of both classes of co-DEGs (i.e., co-up- and co-downregulated genes), mainly in the “Ribosome” and “Coronavirus disease” pathways (Figure 5c,d and Table S9). Ribosomal malfunction reflects the central role that ribosomes play in the control of gene expression. Increased ribosomal biogenesis leads to high levels of protein synthesis as a principal point in maintaining and advancing cancer, without the complete mechanism being known yet.

Interestingly, the second pathway hints towards the molecular association of COVID-19 disease with lung cancer [22]. More specifically, these two separate conditions share a common pathogenetic ground, partially aligned with the lungs being vulnerable to COVID-19. Currently, there is evidence under consideration for developing lung malignancy following SARS-CoV-2 infection.

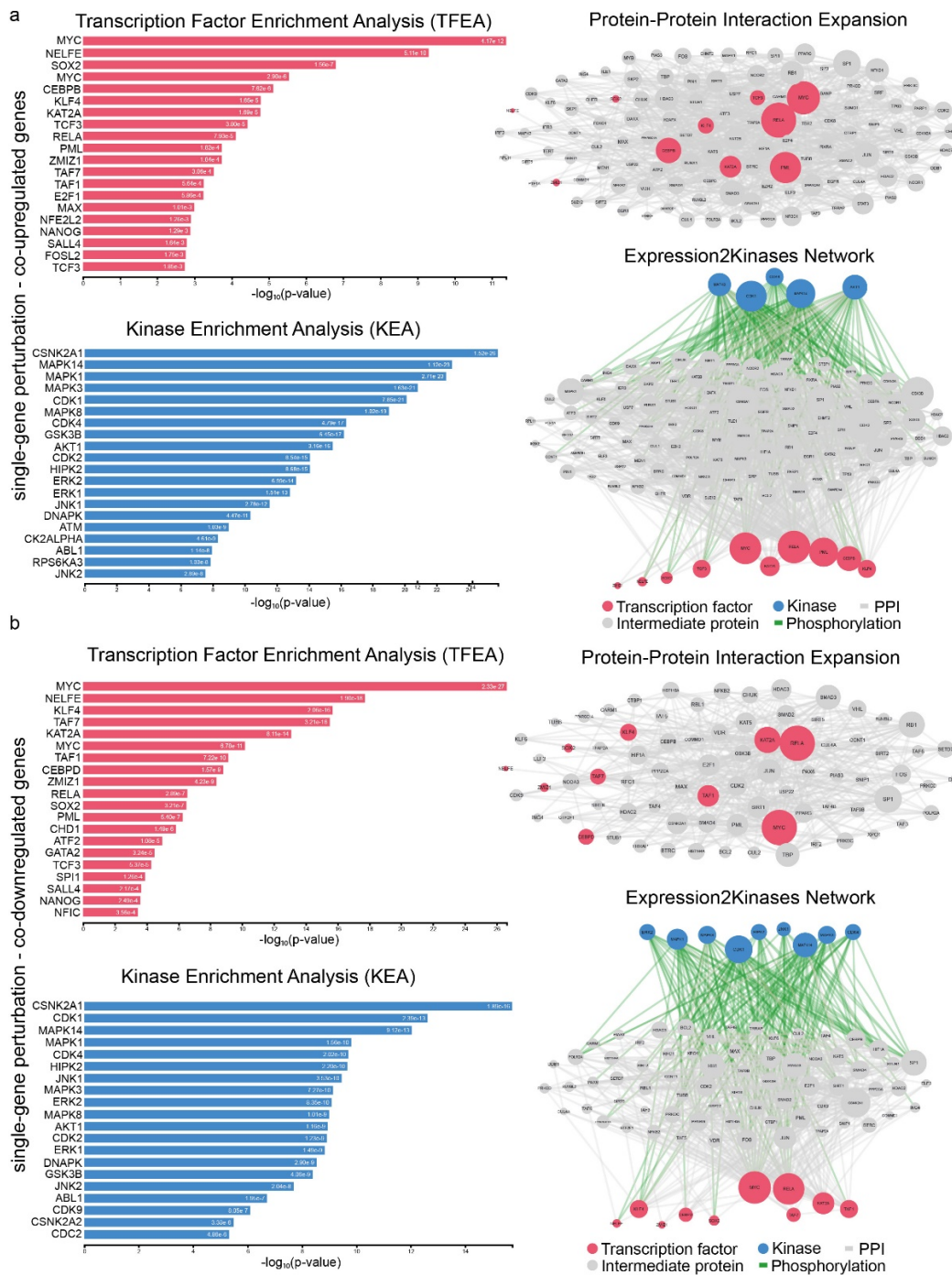


As for the co-downregulated genes, we detected enrichment in “SRP-dependent co-translational protein targeting to membrane”, “cytoplasmic translation”, “co-translational protein targeting to membrane”, “protein targeting to ER” and “nuclear-transcribed mRNA catabolic process, nonsense-mediated decay” (GO biological process); “cytosolic small ribosomal subunit”, “cytosolic large ribosomal subunit”, “small ribosomal subunit” and “large ribosomal subunit” (GO cellular component); and “ubiquitin ligase inhibitor activity”, “ubiquitin–protein transferase inhibitor activity” and “insulin-like growth factor II binding” (GO molecular function) (Figure 5b). Most successive stages of protein synthesis are disrupted by carcinogenetic processes that create a viable tumor microenvironment by controlling the protein interactions at the cellular level. In addition, the deviation from normal protein functionality is reported to be decisive in carcinogenetic events, as there is a dual pathological targeting of translational and post-translational modification processes, as the above-mentioned results reference. Ubiquitination is the physiological process of protein degradation being involved in signal transduction, and results in morbid effects of the solidification of lung tumors. Furthermore, cancer growth signaling moderated by the IGFBP2/IGF pathway has been extensively studied due to its mitogenic nature. Its complex involvement in a plethora of biological procedures results in an uneven motif of overexpression among patients with a diverse clinical picture. In cancer cells, IGFBP2 expression is topologically imported into the cytoplasm and the nucleus is absent of a nuclear localization sequence. Conflicting points of view do not fully explain the oncogenic or tumor suppressive effect of IGFBP2, stating that more *in vitro* functional studies are needed.

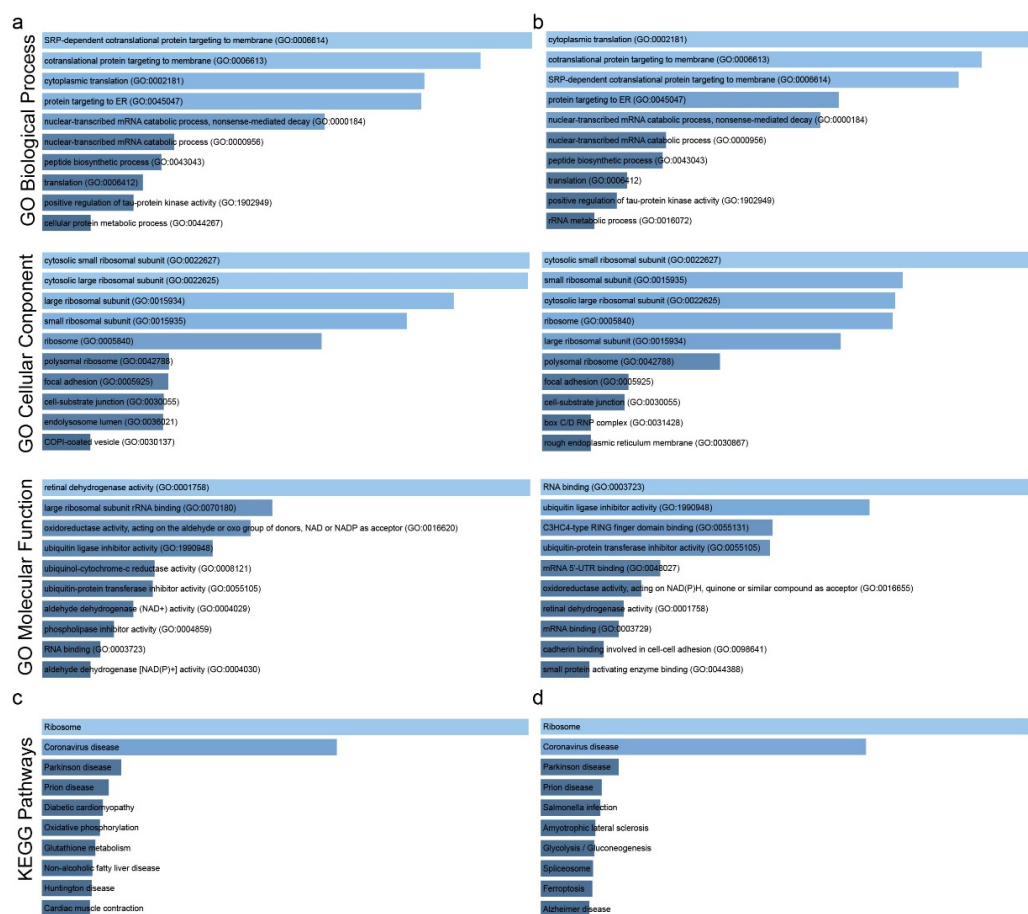
The TFs RELA and MYC, along with the kinases CDK1 and MAPK14, were significant regulatory mediators of both the co-up- and co-downregulated genes following single-gene perturbation. In addition, PML, MYC, RELA, KAT2A, CEBPB, KLF4, TCF3, SOX2, NELFE and ZMIZ1 (TFs), as well as MAPK8/14, CDK1/4 and AKT1 (kinases), were central hub proteins in the co-upregulatory network. Similarly, MYC, RELA, KAT2A, TAF1/7, CEBPD, KLF4, SOX2, NELFE, and ZMIZ1 (TFs), and MAPK1/3/8/14, ERK2, CDK1/4, HIPK2, and JNK1 (kinases), were central hubs in the co-downregulatory network, respectively (Figure 6a,b and Tables S10 and S11).

#### 2.4. Signatures of Co-Deregulated Genes in Single-Drug Perturbation Experiments

We reanalyzed 12 independent GEO datasets to identify the co-DEGs in samples undergoing a single-drug perturbation. The co-upregulated genes were enriched in “SRP-dependent co-translational protein targeting to membrane”, “co-translational protein targeting to membrane”, “cytoplasmic translation”, “protein targeting to ER” and “nuclear-transcribed mRNA catabolic process, nonsense-mediated decay” (GO biological process), “cytosolic small ribosomal subunit”, “cytosolic large ribosomal subunit”, “large ribosomal subunit”, “small ribosomal subunit” and “ribosome” (GO cellular component), as well as in “retinal dehydrogenase activity”, “large ribosomal subunit rRNA binding”, “oxidoreductase activity, acting on the aldehyde or oxo group of donors, NAD or NADP as acceptor”, “ubiquitin ligase inhibitor activity” and “ubiquinol-cytochrome-c reductase activity” (GO molecular function) (Figure 7a and Table S12). Alternated patterns of protein synthesis are observed in controlling their production machinery, as well as the roles the proteins perform. Regarding the molecular processes, the co-upregulated genes were related to biosynthetic biochemical events, and took part in the transmission of intracellular signals. Indeed, in certain cases of non-small cell carcinoma, the expression of retinal dehydrogenase is particularly high and plays a role in the synthesis of increased ATP being used by cancer lung cells [23]. Ubiquinol-cytochrome-c protein reductase is actively implicated in the electron transport chain, but even though it has been suggested as a possible diagnostic biomarker in pulmonary adenocarcinomas [24], its usefulness remains under investigation.



**Figure 6.** Upstream regulatory networks for co-upregulated (a) and co-downregulated (b) gene signatures in single-gene perturbation experiments in lung cancer. The networks depict transcription factors (TFs, red nodes), intermediate proteins (gray nodes), and kinases (blue nodes). Gray edges indicate PPI interactions and green edges depict kinase-driven phosphorylation events. Node size is relative to expression. Upstream regulatory networks were constructed using X2K.

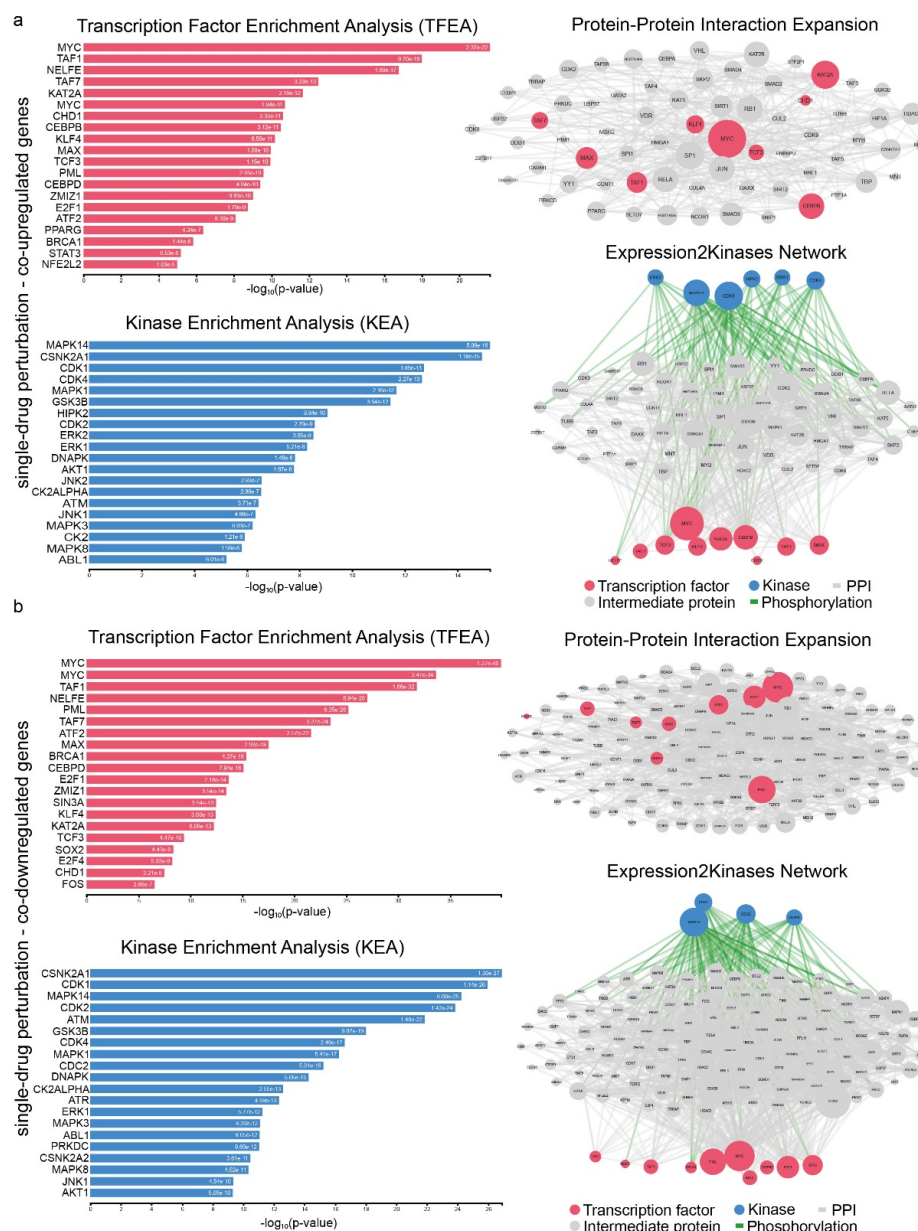


**Figure 7.** GO enrichment of the co-upregulated (a) and the co-downregulated (b) genes in drug perturbation experiments in lung cancer. KEGG enrichment for the co-upregulated (c) and the co-downregulated (d) genes in single-gene perturbation experiments in lung cancer. Bar graphs were sorted using a combined score between the Benjamini–Hochberg (BH)-adjusted  $p$ -value and the z-score of the deviation from the expected rank. Each bar’s length shows the significance of the corresponding term relative to the color brightness of each bar.

In addition, the co-downregulated genes were enriched in “cytoplasmic translation”, “SRP-dependent co-translational protein targeting to membrane”, “protein targeting to ER”, “nuclear-transcribed mRNA catabolic process, nonsense-mediated decay” (GO biological process); “cytosolic small ribosomal subunit”, “small ribosomal subunit”, “Cytosolic large ribosomal subunit”, “ribosome” and “large ribosomal subunit” (GO Cellular component); and in “RNA binding”, “ubiquitin ligase inhibitor activity”, “C3HC4-type RING finger domain binding”, “ubiquitin–protein transferase inhibitor activity” and “mRNA 5'-UTR binding” (GO molecular function) (Figure 7b). The “C3HC4-type zinc finger” region of the RING protein complex is part of several proteins involved in a variety of cell growth and differentiation functions, justifying its transcriptional regulatory application in lung oncogenesis. This factor has been reported to be important in facilitating cell proliferation in lung tissue, diverting ubiquitination pathways in combination with metastatic penetration [25]. The 5'-UTR mRNA binding process reduces the expression of the five untranslated regions of mature mRNA which are responsible for the gene expression post-transcriptional activities necessary for cellular homeostasis. The non-expression of these regulatory elements provokes the consequent deregulation of the normal expression pattern, enhancing pathogenetic effects in lung cancer.

Furthermore, similar to the single-gene perturbation experiments, KEGG analysis for both co-DEGs in this category, revealed enrichment in “Ribosome” and “Coronavirus disease” (Figure 7c,d and Table S13).

The TFs MYC, KAT2A and CEBPB seem to be the most significant hubs among the co-upregulated genes, while MYC, PML, ATF2 and E2F1 seem to be the most significant hubs among the co-downregulated genes in single-drug perturbation experiments (Figure 8a and Table S14). Additionally, MAPK14, CDK1, ERK1/2, HIPK2 and CDK4 were the main hub kinases regulating the expression of the co-upregulated genes, whereas MAPK14, CDC2, DNAPK and CDK4 were the main hubs across the co-downregulated genes in this group (Figure 8b and Table S15).



**Figure 8.** Upstream regulatory networks for co-upregulated (a) and co-downregulated (b) gene signatures in single-drug perturbation experiments in lung cancer. The networks depict transcription factors (TFs, red nodes), intermediate proteins (gray nodes), and kinases (blue nodes). Gray edges indicate PPI interactions and green edges depict kinase-driven phosphorylation events. Node size is relative to expression.



### 3. Discussion

In the present work, we used a Systems Biology approach to detect the co-deregulated genes across different lung cancer studies, orchestrated by an extended network of TFs and kinases. We investigated these co-deregulated gene signatures in-depth and classified them into three different categories, focusing on their upstream regulators, networks, hub proteins and interactions. The new signatures of co-DEGs recovered prior knowledge, but also discovered new connections, adding significant information to the pathobiological basis of lung carcinogenesis.

Assembling the co-DEG signatures in lung cancer and exploring their pathways, we found that they are primarily enriched in immunological destabilization, a major hit in the lung. Immune deregulation is a determining point in the development of tumorigenesis in the lungs, alongside presenting a deteriorating clinical status among patients. Cancerous cells attempt to reverse this protective function by controlling primary cellular signals to deviate from the well-orchestrated immune system. In particular, the perception of the extracellular and intercellular microenvironment is of vital significance for a coordinated immunological response. A similar bioinformatics analysis highlighted the enrichment of the down-regulated genes in angiogenesis in calcium ion binding and cell adhesion [26], whereas the up-regulated genes were significantly enriched in the extracellular matrix disassembly, collagen catabolic process, chemokine-mediated signaling pathway, and endopeptidase inhibitor activity. Our study demonstrates the existence of a strong correlation between the disruption of canonical immunological procedures in co-upregulated and down-regulated genes, respectively. Similar to our results, Yu et al. found enriched critical factors in the regulation of immune responses, inducing tumor growth and metastasis [26]. In a broader context, both studies address the importance of deregulated inflammatory responses in lung cancer, and any differences should be attributed to the different sample numbers analyzed in each GEO dataset or the methodological plan followed.

The co-upregulated genes negatively affect cancerous transformation, as their presence at key topological sites of protein synthesis and transcription has a substantial impact on the expression of genetic information. The enriched terms referring to the molecular function also showed a strong association with lung cancer. In addition, platelet-derived growth factor (PDGF) signaling contributes to a wide variety of developmental procedures, and multiple abnormalities have been documented to occur during lung tumorigenesis. Interestingly, “Relaxin signaling pathway” and “ECM–receptor interaction” were found to be enriched in KEGG pathway analysis. Relaxin is a peptide hormone that acts both in an autocrine and paracrine manner by stimulating the nitric oxide (NO) guanosine pathway after binding to its RXFP1 receptor [27]. Relaxin is involved in the processes of lung structure and remodeling [28], but it seems that its upregulation is also associated with tumor invasion [29]. Furthermore, enhanced RXFP1 activation mediates anti-apoptotic and angiogenic events [30] through ECM degradation [31], providing several implications in the development of chemoresistance [29]. In addition, the metastatic tendency of lung cancer is of high concern owing to the cellular structural changes taking place within the ECM architecture, as the process of malignant establishment continues to occur.

The expression of co-downregulated genes is directly related to dysregulated immunological pathways, ensuring that immune tolerance to malignant cells is upheld and no antioncogenic action occurs. At a biochemical level, the addition of NO to cysteine residues (S-nitrosylation) has been demonstrated to amplify tumor progression and provide no effective response to anticancer medication [32]. Tumor-infiltrating leukocytes are known to have a major, yet antagonistic, role in cancer surveillance and appear to disrupt epithelial tissues during inflammation. More specifically, angiogenic events happen during the onset of tumorigenesis, as the infiltrated leukocytes produce pro-angiogenic agents that maintain a favorable tumor microenvironment. Likewise, cell adhesion molecules are the main membrane glycoproteins that regulate immune cell migration and activation, through which malignant cells can spread to secondary tissues. In addition, it was recently proposed that leukocyte aggregates are associated with high thrombosis risk in lung cancer

patients [33], corresponding to the fact that leukocytes accumulate in the tumor microenvironment in order to facilitate their progression and survival. In this context, it is assumed that a high leukocyte aggregation triggers a metastatic profile of lung cancer to distant sites, and is correlated with fatal cardiovascular incidents. Indeed, in a recent observational study, cancer patients were shown to have a high mortality risk due to cardiovascular disease [34]. Thus, the co-downregulated genes were mainly enriched in “cytolytic granule” (GO cellular component) and took part in molecular processes such as “RAGE receptor binding”, “arachidonic acid binding”, “icosatetraenoic acid binding”, “eicosanoid binding” and “Toll-like receptor binding”.

The co-downregulated genes were enriched in “RAGE receptor functions of glycosylation end products”, which are mainly presented in diabetic pathogenesis, but the RAGE receptor is associated with lung carcinogenesis through the induction of multiple intracellular oxidative stress pathways [35]. Nevertheless, RAGE receptors appear to exhibit multiligand characteristics by activating responses in inflammatory conditions, and they are involved in the maturation process of dendritic cells, T cell proliferation, and polarization to CD4+ cells. At the same time, there is a functional relationship with the TLR receptor family through the initiation of innate immune responses because of the joint interaction ligands. TLRs have regulatory features to mediate cancer cell growth, secondary tissue penetration, angiogenesis, and cancerous conversion in different molecular mechanisms.

In addition, the metabolic pathway of arachidonic acid is linked to the biological basis of lung cancer due to the involvement of biosynthetic cyclooxygenases (COX-2) and phospholipase A2 (PLA2) in the MAPK/ERK and EGFR pathways [36]. Regarding the binding function of eicosapentaenoic acid, omega-3 fatty acid supplements are well-known to have anti-inflammatory properties against cancer cachexia syndrome [37], playing a beneficial role in the inhibition of cell proliferation and the reversal of the arachidonic acid metabolism [38]. This correlates with previous findings, reinforcing the need for the co-prescription of supplements in lung cancer patients. Finally, similar effects are revealed in eicosanoid acid-binding processes that act as lipid mediators of arachidonic acid, the production of which is higher both in tumor cells and their surrounding microenvironment [39,40], hinting towards a new treatment approach that targets the biochemical cycle of lung cancer cells.

Diabetic cardiomyopathy may be associated with lung carcinogenesis via the activation of the AGE–RAGE signaling pathway. The disease also associates with the intracellular stimulation of oxidative stress, whereas the RAGE receptor modulates immune responses in the lungs [41]. Pathogenic Salmonella infection is not correlated with tumorigenesis in the lung, but the organ remains vulnerable to infections. This finding warrants further investigation in the lung microbiome to better understand the properties of its drug resistance.

ECM is a highly complex structure of biologically active macromolecules that control key cellular functions. The interaction between ECM and tumor cells underlines the tendency of the latter to infiltrate the tumor microenvironment. The dynamic and organizational structure of ECM is considered to be the basis of cellular behavior. Tumor-associated macrophages (TAMs) and neutrophils also play a key role in the onset of carcinogenesis, performing phagocytic processes (“phagosome”) through lysosomal digestion. In the majority of different cancer types, antiphagocytic signals are expressed to escape immune surveillance, and thus malignant proliferation is not eliminated.

The terms that we found to correspond to the KEGG pathway analysis for the co-downregulated genes refer to certain immunological procedures taking place during histocompatibility and the development of an acute inflammation response. After all, lung is a target organ of the graft-versus-host disease response due to the consecutive emergence of cytotoxic T-lymphocytes and NK cells influenced by the cytokines IL-1/TNF $\alpha$  [42]. The interleukin 17 (IL-17) signaling pathway may be enriched because of its participation in tumor-associated inflammation. In fact, IL-17 is a pro-inflammatory cytokine [43] that contributes to the metastasis of lung cancer cells [44].

PPI networks contribute to a further understanding of the transcriptional machinery and phosphorylation reactions driving the pathobiological events of lung carcinogenesis. Here, we sought to investigate the interactions among the co-DEGs in lung cancer and certain hub proteins as critical mediators in the signaling pathways. We highlight EZH2, CTCF, RAD21, HDAC2, RUNX1 and PPARG as the principal hubs closely connected to this disease. Of these, we emphasize on the enhancer of Zeste Homologue 2 (EZH2), the CCCTC-binding factor (CTCF), and the RAD21 Cohesin Complex Component (RAD21), across the main co-upregulated TFs.

EZH2 possesses a crucial role in chromosomal remodeling, as well as in the regulation of the silencing of several tumor suppressor genes and those involved in immune cell development [45]. The overexpression of EZH2 allows cancer cells to divide uncontrollably and plays an important role in the acquired chemoresistance of cancer cells [46].

CTCF has multiple functional roles in chromosomal interactions, leading to gene expression, while its rs60507107 variant correlates with an increased risk of lung cancer [47]. As a chromatin architecture mediator, it epigenetically regulates transcription and CTCF-binding alterations can be considered as epigenomic signatures of cancer development. A cancerous transition mechanism in lung fibroblasts relies upon the deregulated expression of Rb2/p130, which is controlled by CTCF, and has been shown to promote the progression and recurrence of lung cancer after treatment. Thus, CTCF upregulation promotes carcinogenic effects, as it ultimately organizes the genome structure and can alter gene expression.

RAD21 normally participates in sister chromatid cohesion and separation when needed during transcription, DNA replication, or DNA damage repair mechanisms. RAD21 knock down shows resistance to DNA-damaging chemotherapeutic drugs *in vitro* [48], and its increased expression is evident in poorly differentiated lung cancers due to its contribution to the regulation of the cell cycle [49].

On the other hand, among the co-downregulated transcription factors in lung cancer, we focused on RUNX1, histone deacetylase 2 (HDAC2), and peroxisome proliferator-activated receptor gamma (PPARG).

RUNX1 participates in different hallmarks of cancer, such as the developmental differentiation of multiple human cell lines. Its downregulation was recently linked with poor lung cancer patient survival [50], while its altered methylation pattern is used as a biomarker in non-small cell lung cancer (NSCLC) [51]. Furthermore, RUNX1 is known to inhibit the transcription of YAP [52] in breast cancer, a molecular interplay that leads to immunosuppressive events during lung tumorigenesis [53].

HDAC2 functions as a part of large multiprotein complexes that repress transcription through the deacetylation of lysine residues on the N-terminal part of histones. In NSCLC, HDAC2 knock down correlates with a low expression of fibronectin (FN) [54] and enhances the metastatic potential of lung cancer cells [55]. The functional alteration and imbalance in epigenetic modulation affects major pleiotropic cellular events and growing interest has been raised towards clinical utility.

A decreased expression of PPARG is a prognostic marker for NSCLC. PPARs regulate cancer-relevant processes, such as cell differentiation, proliferation and apoptosis [56,57]. In colorectal cancer, emerging evidence has shown that PPARG signaling is downregulated due to the regulatory actions of EZH2 and HDAC1 [58]. In fact, our results come in accordance with this statement and we, therefore, consider that the axis between PPARs, EZH2 and HDACs is a novel transcriptional interplay underlying lung carcinogenesis, and thus needs further investigation.

Moreover, we detected specific repurposing drugs to potentially extend the therapeutic opportunities in lung cancer. A beneficial effect in lung cancer treatment could result from the pharmacological regulation in the expression of co-DEGs. Here, we paid specific attention to valproic acid (VPA), betonicine, and astemizole.

Valproic acid is a histone deacetylase (HDAC) inhibitor used in the treatment of epilepsy because of its mood-stabilization properties. As such, VPA can regulate the expres-

sion of genes in small cell lung cancer (SCLC), as well as balance apoptotic modulators and suppress cell growth via the activation of the Notch 1 signaling pathway [59]. VPA was also recently reported to effectively inhibit the proliferation of lung cancer cells *in vivo*, when combined with arsenic trioxide, via the induction of apoptotic signals [60]. In addition, VPA was shown to improve a second-line regimen of small cell lung carcinoma in preclinical models, opening new prospects for improved therapies [61]. The anticancer activity of VPA was also shown in non-small cell lung cancer (NSCLC) cell lines when combined with a new cyclin-dependent kinase (CDK) inhibitor (P276-00) [62]. Therefore, our findings corroborate previous reports that VPA is a promising new molecularly targeted therapeutic approach for the treatment of lung cancer.

Betonicine is a pyrrolidine alkaloid isolated from *Achillea millefolium* that inhibits bacterial signaling molecules being used as an adjuvant to anti-infective treatments. It targets several enzymes involved in drug metabolism (mainly CYPs). Some types of betonicine's inhibitory activity are against  $\alpha$  and  $\beta$  glucosidases, as well as  $\beta$ -mannosidase and eukariotic DNA polymerases. The toxic effect of pyrrolidine alkaloids is generally manifested in the liver [63,64]; nevertheless, betonicine's anti-tumor effects in lung cancer are widely unknown.

As for the repurposing drugs that seem to induce the co-downregulated genes in lung cancer, here we highlight astemizole. This drug has been commonly used in the treatment of allergies as an antihistamine, but it was reported to cause arrhythmias when administered in high doses. Astemizole targets several proteins involved in tumor cell proliferation [65], and its combined pharmacological action with gefitinib was recently shown to have promising results in lung cancer [66]. Our findings corroborate these reports and strongly suggest that astemizole could be used as a repurposing therapeutic drug for lung cancer patients. Together, these findings identify the necessity of integrating repurposed drugs into therapeutic efforts to treat lung cancer.

All in all, here, we analyzed pooled data sets across heterogeneous subtypes of lung cancer and identified the top co-DEGs and their hubs along with targeted therapeutic drugs. Our results show that all of the three signature categories recover prior knowledge associations between genes, drugs, and diseases.

## 4. Materials and Methods

### 4.1. GEO Data Extraction and Filtering

We first assorted 24 studies from the Gene Expression Omnibus (GEO) to extract differentially expressed genes using GEO2Enrichr [67,68]. The studies were classified into three categories: (1) those assessing gene expression profiles in LC vs. healthy samples (2 GEO studies), (2) those subjected to single-gene perturbation (10 GEO studies), and (3) those observing any therapeutic effect through single-drug perturbation (12 GEO studies) in lung cancer. The selection process was strictly focused on gene expression studies containing tissue samples or cell lines of human or mouse origin. The standard naming of genes, diseases, and drugs was provided as an autocomplete option in the submission forms created from HGNC [68], Disease Ontology [69], and DrugBank [70], respectively.

The corresponding mined data sets were as follows: GDS4794, consisting of 23 lung carcinoma tissue samples and 2 normal tissue samples; GDS3837, containing 60 samples of healthy tissue and 60 samples of non-small cell lung cancer from women; GDS5418, containing 4 normal A456 cell lines and 4 SRC-/- A456 cell lines; GDS5391, consisting of 4 NCI-H1299 and NCI cell lines samples (2 for each line) of pulmonary adenocarcinoma not expressing the protein kinase of tyrosine PTK7 and 4 normal samples (control); GDS2489, consisting of 18 normal samples of lung airway epithelial cells and 26 samples of the same cell line from smokers; GDS3510, composed of CL1-5-derived cells overexpressing Claudin-1 (CLDN1) derived from pulmonary adenocarcinoma tissue; GDS3029, having 27 small cell lung cancer cell samples resistant to Bcl-2 antagonist ABT-737 and 7 normal samples (control); GDS3826, containing 5 transgenic SP-C/c-raf lung samples and 10 healthy samples; GDS5206, consisting of 9 samples of A549 epithelial cell line from



pulmonary adenocarcinoma and 9 healthy controls; GDS4847, containing 40 lung samples from B6C3F1 mice that were given chloprene and 10 normal tissue samples; GDS5067, consisting of 5 control samples and 9 A549 cell line samples treated with oligopiperazines (OOPs) BB2-125, BB2-162 and BB2-282; GDS3101, composed of A459-derived cis-platinum-resistant cells, 3 samples that were treated with cis-platinum, and 3 more samples that were not; GDS5496, consisting of 2 NCI-H441 lung-derived cancer cells expressing hsa-miR-365-2 and 2 normal samples; GDS3825, analyzing 5 transgenic SP-C/c-raf cancerous lung cells, 5 non-transgenic SP-C/c-raf samples and 5 completely normal samples; GDS5201, containing 4 tumor samples resulting from activation of the Wnt/beta-catenin signaling pathway in combination with the expression of KRAS protein and 2 samples of normal lung tissues; GDS2966, consisting of A549 lung cancer cells, of which 2 samples were exposed to resveratrol and 2 more remained intact; GDS5648, consisting of 5 K-RAS-driven mouse tissue and 5 controls; GDS1649, containing 15 samples of mouse lung cancer and 29 samples that were subjected to iodine urethane; GDS3321, consisting of 4 mouse lung tumors which had undergone transgenic intervention in alveolar epithelial cells to overexpress c-Myc proto-oncogene and 4 control samples; GDS2958, using 2 samples of HCC827-derived cells with inactive PTEN inhibitor and 2 normal samples; GDS2298, containing 7 samples of NSCLC-derived cell lines sensitive to Gefitinib and 11 control samples; GDS4840, containing 3 samples of CCR5-/-mutant mouse-derived lung tissue and 3 samples from normal tissues; GDS2604, consisting of 14 samples from epithelial and mesothelial lung cancer cell lines and 13 control samples; GDS5247, using 3 samples of H460 parent-derived lung cell line and 3 samples of the same cell line with cis-platin resistance; GDS2499, containing 3 cell samples were cultured in mannitol, 3 samples cultured in actinomycin D, and 6 maintained in saffron PCI-2050 (Table S2).

We re-processed the extracted expression signatures to filter their quality and check data integrity, as previously explained in detail [3,66]. We also quantified batch effects with variance component analysis [71] and corrected them with surrogate variable analysis (SVA) [68].

#### 4.2. Differentially Expressed Genes and Co-DEGs

Genes presenting a differential expression profile (cutoff of 500 genes) were estimated based on the Characteristic Direction (CD) algorithm [12]. The GEO2Enrichr extracted gene sets of the over- and under-expressed genes were manually sorted based on the CD metric. In the event of single-gene perturbation (i.e., knock out, knock down, knock in, RNAi, or overexpression) experiments, the differential expression signatures were considered against their normal (wt) alleles. Regarding single-drug experiments, DEGs were calculated against the non-treated cells or tissues. Differentially expressed genes that were found only in one study were excluded from further analysis. For each group, the DEGs that were identified between at least two independent studies were termed as “co-upregulated” or “co-downregulated”, respectively (co-DEGs).

#### 4.3. Upstream Regulators of co-DEGs and Protein–Protein Interaction Networks

Expression2Kinases (X2K) was utilized to investigate the upstream regulatory networks from the co-DEGs signatures within each group. Transcription factors (TFs), intermediate proteins and protein kinases participating in the regulation of transcriptional and expression processes of the inputted co-DEGs were produced as previously described [72]. X2K detects the expected elements within each category using an integrated promoter analysis of ChIP-X to construct a complete network between TFs and known protein interactions [73]. Kinase enrichment analysis was used to locate the upstream protein kinases, which are critical for carrying out phosphorylation reactions that regulate the expression of the identified co-DEGs. We then used Enrichr [74] to create protein–protein interaction networks (PPI) for each group of co-DEGs, involving the implicated TFs, kinases and their intermediated proteins, indicating the nodes (genes) and edges (lines) in each network. The

interacting proteins along with the phosphorylation paths involved were also included for the visualization purposes of each PPI network.

#### 4.4. GO and KEGG Enrichment Analysis

We performed Gene Ontology (GO) enrichment analysis to reveal the overrepresented GO terms, focusing on the molecular-level activities performed by the top co-DEGs, the locations relative to the cellular structures in which their gene products perform their functions, and the biological processes accomplished by multiple molecular activities using Enrichr [74].

In addition, we used the Kyoto Encyclopedia of Genes and Genomes (KEGG) [75–77] to assess the pathway information of the top co-DEGs in each group. Results were sorted using a combined score, defined as  $c = \log(p) \cdot z$ , i.e., the log of the BH-adjusted  $p$ -value from the Fischer's exact test, multiplied by the  $z$ -score of the deviation from the expected rank. An adjusted  $p$ -value (adj- $p$ ) of 0.05 was used as the threshold of statistical significance.

#### 4.5. Connectivity Map Analysis

Connectivity Map (CMap, <https://clue.io/cmap>, accessed on 21 September 2021) is a large-scale database including shareable data on pharmaceutical agents [78] and substances with a previously defined pharmacological action. These are statistically correlated with provided gene expression signatures, and could potentially induce or reverse LC based on this. We used CMap to identify repurposing drugs for lung cancer based on a connectivity score ranging from  $-1$  to  $+1$ . The given range indicates that values closer to  $+1$  have a positive connectivity, as drugs enhance the carcinogenic effects in the lung. In contrast, values closer to  $-1$  have a reversible gene–drug relationship, suppressing the growth of lung cancer cells. We submitted the lists of co-DEGs (over- and under-expressed genes related to lung cancer), and found a group of substances that act reversibly on the studied expression patterns of the co-DEGs. A hypergeometric probability test was used to associate drugs with disease.

#### 4.6. Validation of the Co-Deregulated Gene Signatures and Hub Genes in the TCGA and the Human Protein Atlas

We validated the calculated expression signatures in lung cancer after extracting the read counts of RNA-seq data from the Cancer Genome Atlas of lung squamous cell carcinoma (TCGA-LUSC) (486 tumor samples and 338 controls) and the lung adenocarcinoma (LUAD) (483 tumors and 347 controls) datasets using the Genomic Data Commons data portal (<https://portal.gdc.cancer.gov/>, accessed on 23 January 2022). We then normalized the read counts to  $\log_2$  (TPM + 1) values, as previously described [79]. To increase the sample number of the controls, the TCGA normal data were matched with normal lung samples from the Genotype-Tissue Expression (GTEx) project (<https://gtexportal.org/home/>, accessed on 23 January 2022). The expression levels of the “UP genes” signature, composed of 54 genes, and of the “DOWN genes” signature, composed of 48 genes, were explored with Limma [80] using  $\log_2FC = 1$  and  $q$ -value = 0.01 as thresholds of significance.

In addition, we explored the expression of the top deregulated genes in lung cancer across different molecular subtypes in LUSC (basal, classical, primitive, secretory), as well as different immune subtypes in LUSC and LUAD (C1, wound healing; C2, IFN-gamma dominant; C3, inflammatory; C4, lymphocyte depleted; C5, immunologically quiet; C6, TGF- $\beta$  dominant) [13].

To further investigate the hub genes and associate their expression with patient survival, we used the Gene Expression Profiling Interactive Analysis 2 (GEPIA2) [14] and immunohistochemistry (IHC) data from the Human Protein Atlas (HPA) [81]. Regarding the IHC protocol that was followed, in brief, FFPE sections (4  $\mu$ m) were heated at 50 °C overnight. Then, they were deparaffinized in xylene and rehydrated in graded ethanol to distilled water. During hydration, a 5 min blocking for endogenous peroxidase was completed in 0.3% H<sub>2</sub>O<sub>2</sub> in 95% ethanol. Prior to immunostaining, the sections were im-

mersed in 10mM citrate buffer (pH 6.0), rinsed in Tris-buffered saline (TBS) and subjected to heat-induced epitope retrieval (HIER) using a pressure boiler. Sections were then incubated overnight at 4 °C with mouse monoclonal antibodies (mAbs) against MYC (1:20, Atlas Antibodies Cat#HPA055893, RRID: AB\_2682960), TP63 (1:75, Atlas Antibodies Cat#HPA006288, RRID: AB\_1080334). The UltraVision LP HRP polymer<sup>®</sup>, Ultra V Block and DAB quanto substrate system<sup>®</sup> (Thermo scientific, La Jolla, CA, USA) were used for detection. Finally, slides were rinsed in tap water, counterstained with hematoxylin, dehydrated in grade ethanol and cover-slipped. Slides were then assessed independently by two observers.

#### 4.7. Patient Survival Analysis

Disease-free survival analysis and corresponding maps for the patients in the TCGA-LUSC and TCGA-LUAD datasets were further constructed using two gene signatures: one containing the hub TFs (MYC, STAT3, TP63, KLF4, SOX2, CHD1 and FOSL2) and another containing the hub kinases (JNK1, HIPK2, CSNK2A1, MAPK3, CDK1, CDK4, GSK3B, ERK1 and MAPK14). For the Kaplan–Meier curves we used hazard ratio (HR), which was calculated based on the Cox PH model and a 95% confidence interval (CI). For the survival maps, we used an FDR-adjusted *p*-value = 0.05 as the threshold of significance and median value cutoff. All experiments were performed in accordance with the TCGA relevant guidelines and regulations.

## 5. Conclusions

Overall, the study of commonly deregulated pathways and molecular interactions reveals the dynamic processes taking place during lung carcinogenesis.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms231810933/s1>.

**Author Contributions:** Conceptualization, A.Z.; methodology, A.C. and A.Z.; software, A.Z. and A.C.; validation, A.Z.; formal analysis, A.Z. and A.C.; investigation, A.Z. and A.C.; resources, A.Z. and A.C.; writing—original draft preparation, A.C. and A.Z.; writing—review and editing, A.Z.; visualization, A.Z.; supervision, A.Z.; project administration, A.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The figshare repository was used to store the supporting data of our findings (doi:10.6084/m9.figshare.16736896).

**Acknowledgments:** We would like to acknowledge the GEO studies from which data were extracted.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [[CrossRef](#)]
2. Liu, W.-J.; Du, Y.; Wen, R.; Yang, M.; Xu, J. Drug resistance to targeted therapeutic strategies in non-small cell lung cancer. *Pharmacol. Ther.* **2019**, *206*, 107438. [[CrossRef](#)] [[PubMed](#)]
3. Mastrogamvraki, N.; Zaravinos, A. Signatures of co-deregulated genes and their transcriptional regulators in colorectal cancer. *NPJ Syst. Biol. Appl.* **2020**, *6*, 23. [[CrossRef](#)] [[PubMed](#)]
4. Wen, S.; Peng, W.; Chen, Y.; Du, X.; Xia, J.; Shen, B.; Zhou, G. Four differentially expressed genes can predict prognosis and microenvironment immune infiltration in lung cancer: A study based on data from the GEO. *BMC Cancer* **2022**, *22*, 1–12. [[CrossRef](#)] [[PubMed](#)]
5. Zengin, T.; Önal-Süzek, T. Analysis of genomic and transcriptomic variations as prognostic signature for lung adenocarcinoma. *BMC Bioinform.* **2020**, *21* (Suppl. S14), 368. [[CrossRef](#)] [[PubMed](#)]
6. Gao, M.; Kong, W.; Huang, Z.; Xie, Z. Identification of key genes related to lung squamous cell carcinoma using bioinformatics analysis. *Int. J. Mol. Sci.* **2020**, *21*, 2994. [[CrossRef](#)]

7. Zhang, Y.; Liu, X.; Liu, L.; Li, J.; Hu, Q.; Sun, R. Expression and Prognostic Significance of m6A-Related Genes in Lung Adenocarcinoma. *Med Sci. Monit.* **2020**, *26*, e919644. [[CrossRef](#)]
8. Long, T.; Liu, Z.; Zhou, X.; Yu, S.; Tian, H.; Bao, Y. Identification of differentially expressed genes and enriched pathways in lung cancer using bioinformatics analysis. *Mol. Med. Rep.* **2019**, *19*, 2029–2040. [[CrossRef](#)]
9. Li, R.; Liu, Y.; Wang, T.; Tang, J.; Xie, L.; Yao, Z.; Li, K.; Liao, Y.; Zhou, L.; Geng, Z.; et al. The characteristics of lung cancer in Xuanwei County: A review of differentially expressed genes and noncoding RNAs on cell proliferation and migration. *Biomed. Pharmacother.* **2019**, *119*, 109312. [[CrossRef](#)]
10. Yu, H.; Pang, Z.; Li, G.; Gu, T. Bioinformatics analysis of differentially expressed miRNAs in non-small cell lung cancer. *J. Clin. Lab. Anal.* **2021**, *35*, e23588. [[CrossRef](#)]
11. Tian, S. Identification of monotonically differentially expressed genes for non-small cell lung cancer. *BMC Bioinform.* **2019**, *20*, 177. [[CrossRef](#)] [[PubMed](#)]
12. Clark, N.R.; Hu, K.S.; Feldmann, A.S.; Kou, Y.; Chen, E.Y.; Duan, Q.; Ma'Ayan, A. The characteristic direction: A geometrical approach to identify differentially expressed genes. *BMC Bioinform.* **2014**, *15*, 79. [[CrossRef](#)]
13. Thorsson, V.; Gibbs, D.L.; Brown, S.D.; Wolf, D.; Bortone, D.S.; Ou Yang, T.-H.; Porta-Pardo, E.; Gao, G.F.; Plaisier, C.L.; Eddy, J.A.; et al. The Immune Landscape of Cancer. *Immunity* **2018**, *48*, 812–830.e14. [[CrossRef](#)] [[PubMed](#)]
14. Tang, Z.; Kang, B.; Li, C.; Chen, T.; Zhang, Z. GEPIA2: An enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res.* **2019**, *47*, W556–W560. [[CrossRef](#)] [[PubMed](#)]
15. Shults, N.V.; Seeherman, S.; Sariipek, N.E.; Rybka, V.; Marocci, L.; Gychka, S.G.; Ibrahim, Y.F.; Suzuki, Y.J. Tau Protein in Lung Smooth Muscle Cells. *J. Respir.* **2020**, *1*, 30–39. [[CrossRef](#)]
16. Papin, S.; Paganetti, P. Emerging Evidences for an Implication of the Neurodegeneration-Associated Protein TAU in Cancer. *Brain Sci.* **2020**, *10*, 862. [[CrossRef](#)] [[PubMed](#)]
17. Gavrilovic, I.; Posner, J.B. Brain metastases: Epidemiology and pathophysiology. *J. Neuro-Oncol.* **2005**, *75*, 5–14. [[CrossRef](#)]
18. Wang, G.; Xu, J.; Qi, Y.; Xiu, J.; Li, R.; Han, M. Distribution of Brain Metastasis from Lung Cancer. *Cancer Manag. Res.* **2019**, *11*, 9331–9338. [[CrossRef](#)]
19. Pappot, H.; Br unner, N. The plasminogen activation system and its role in lung cancer. A review. *Lung Cancer* **1995**, *12*, 1–12. [[CrossRef](#)]
20. Vallabhapurapu, S.D.; Blanco, V.M.; Sulaiman, M.K.; Vallabhapurapu, S.L.; Chu, Z.; Franco, R.S.; Qi, X. Variation in human cancer cell external phosphatidylserine is regulated by flippase activity and intracellular calcium. *Oncotarget* **2015**, *6*, 34375–34388. [[CrossRef](#)]
21. Birge, R.B.; Boeltz, S.; Kumar, S.; Carlson, J.; Wanderley, J.; Calianese, D.; Barcinski, M.; Brekken, R.A.; Huang, X.; Hutchins, J.T.; et al. Phosphatidylserine is a global immunosuppressive signal in efferocytosis, infectious disease, and cancer. *Cell Death Differ.* **2016**, *23*, 962–978. [[CrossRef](#)] [[PubMed](#)]
22. Soon Nan, K.; Karuppanan, K.; Kumar, S.; Alam, S. Identification of common key genes and pathways between COVID-19 and lung cancer by using protein-protein interaction network analysis. *bioRxiv* **2021**. [[CrossRef](#)]
23. Kang, J.H.; Lee, S.-H.; Hong, D.; Lee, J.-S.; Ahn, H.-S.; Ahn, J.-H.; Seong, T.W.; Lee, C.-H.; Jang, H.; Hong, K.M.; et al. Aldehyde dehydrogenase is used by cancer cells for energy metabolism. *Exp. Mol. Med.* **2016**, *48*, e272. [[CrossRef](#)] [[PubMed](#)]
24. Gao, F.; Liu, Q.; Li, G.; Dong, F.; Qiu, M.; Lv, X.; Zhang, S.; Guo, Z. Identification of ubiquinol cytochrome c reductase hinge (UQCRH) as a potential diagnostic biomarker for lung adenocarcinoma. *Open Biol.* **2016**, *6*, 150256. [[CrossRef](#)] [[PubMed](#)]
25. Cassandri, M.; Smirnov, A.; Novelli, F.; Pitolli, C.; Agostini, M.; Malewicz, M.; Melino, G.; Raschella, G. Zinc-finger proteins in health and disease. *Cell Death Discov.* **2017**, *3*, 17071. [[CrossRef](#)]
26. Yu, Y.; Tian, X. Analysis of genes associated with prognosis of lung adenocarcinoma based on GEO and TCGA databases. *Medicine* **2020**, *99*, e20183. [[CrossRef](#)] [[PubMed](#)]
27. Fallowfield, J.A.; Hayden, A.L.; Snowdon, V.K.; Aucott, R.L.; Stutchfield, B.M.; Mole, D.J.; Pellicoro, A.; Gordon-Walker, T.T.; Henke, A.; Schrader, J.; et al. Relaxin modulates human and rat hepatic myofibroblast function and ameliorates portal hypertension in vivo. *Hepatology* **2014**, *59*, 1492–1504. [[CrossRef](#)]
28. Wyatt, T.; Sisson, J.; Forg t, M.; Bennett, R.; Hamel, F.; Spurzem, J. Relaxin Stimulates Bronchial Epithelial Cell PKA Activation, Migration, and Ciliary Beating. *Exp. Biol. Med.* **2002**, *227*, 1047–1053. [[CrossRef](#)]
29. Rizvi, S.; Gores, G.J. The Two Faces of Relaxin in Cancer: Antitumor or Protumor? *Hepatology* **2020**, *71*, 1117–1119. [[CrossRef](#)]
30. Thanasupawat, T.; Glogowska, A.; Nivedita-Krishnan, S.; Wilson, B.; Klonisch, T.; Hombach-Klonisch, S. Emerging roles for the relaxin/RXFP1 system in cancer therapy. *Mol. Cell. Endocrinol.* **2019**, *487*, 85–93. [[CrossRef](#)]
31. Samuel, C.S.; Lekgabe, E.D.; Mookerjee, I. The Effects of Relaxin on Extracellular Matrix Remodeling in Health and Fibrotic Disease. *Adv. Exp. Med. Biol.* **2007**, *612*, 88–103. [[CrossRef](#)] [[PubMed](#)]
32. Ben-Lulu, S.; Ziv, T.; Weisman-Shomer, P.; Benhar, M. Nitrosothiol-trapping-based proteomic analysis of s-nitrosylation in human lung carcinoma cells. *PLoS ONE* **2017**, *12*, e0169862. [[CrossRef](#)] [[PubMed](#)]
33. Meikle, C.K.; Meisler, A.J.; Bird, C.M.; Jeffries, J.A.; Azeem, N.; Garg, P.; Crawford, E.L.; Kelly, C.A.; Gao, T.Z.; Wuescher, L.M.; et al. Platelet-T cell aggregates in lung cancer patients: Implications for thrombosis. *PLoS ONE* **2020**, *15*, e0236966. [[CrossRef](#)] [[PubMed](#)]
34. Herrmann, J. From trends to transformation: Where cardio-oncology is to make a difference. *Eur. Hear. J.* **2019**, *40*, 3898–3900. [[CrossRef](#)] [[PubMed](#)]



35. Kay, A.M.; Simpson, C.L.; Stewart, J.A. The Role of AGE/RAGE Signaling in Diabetes-Mediated Vascular Calcification. *J. Diabetes Res.* **2016**, *2016*, 1–8. [[CrossRef](#)]
36. Krysan, K.; Reckamp, K.L.; Dalwadi, H.; Sharma, S.; Rozenfurt, E.; Dohadwala, M.; Dubinett, S.M. Prostaglandin E2 Activates Mitogen-Activated Protein Kinase/Erk Pathway Signaling and Cell Proliferation in Non-Small Cell Lung Cancer Cells in an Epidermal Growth Factor Receptor-Independent Manner. *Cancer Res.* **2005**, *65*, 6275–6281. [[CrossRef](#)]
37. Gorjao, R.; dos Santos, C.M.M.; Serdan, T.D.A.; Diniz, V.L.S.; Alba-Loureiro, T.C.; Cury-Boaventura, M.F.; Hatanaka, E.; Levada-Pires, A.C.; Sato, F.T.; Pithon-Curi, T.C.; et al. New insights on the regulation of cancer cachexia by N-3 polyunsaturated fatty acids. *Pharmacol. Ther.* **2018**, *196*, 117–134. [[CrossRef](#)]
38. Yang, P.; Cartwright, C.; Chan, D.; Ding, J.; Felix, E.; Pan, Y.; Pang, J.; Rhea, P.; Block, K.; Fischer, S.M.; et al. Anticancer activity of fish oils against human lung cancer is associated with changes in formation of PGE<sub>2</sub> and PGE<sub>3</sub> and alteration of Akt phosphorylation. *Mol. Carcinog.* **2013**, *53*, 566–577. [[CrossRef](#)]
39. Poczobutt, J.M.; Gijon, M.; Amin, J.; Hanson, D.; Li, H.; Walker, D.; Weiser-Evans, M.; Lu, X.; Murphy, R.C.; Nemenoff, R.A. Eicosanoid profiling in an orthotopic model of lung cancer progression by mass spectrometry demonstrates selective production of leukotrienes by inflammatory cells of the microenvironment. *PLoS ONE* **2013**, *8*, e79633. [[CrossRef](#)]
40. Gomes, R.N.; da Costa, S.F.; Colquhoun, A. Eicosanoids and cancer. *Clinics* **2018**, *73*, e530s. [[CrossRef](#)]
41. Oczypok, E.A.; Perkins, T.N.; Oury, T.D. All the “RAGE” in lung disease: The receptor for advanced glycation endproducts (RAGE) is a major mediator of pulmonary inflammatory responses. *Paediatr. Respir. Rev.* **2017**, *23*, 40–49. [[CrossRef](#)] [[PubMed](#)]
42. Liu, Q.-F.; Luo, X.-D.; Ning, J.; Xu, D.; Fan, Z.-P.; Sun, J.; Zhang, Y.; Xu, B.; Wei, Y.-Q. Association between acute graft versus host disease and lung injury after allogeneic haematopoietic stem cell transplantation. *Hematology* **2009**, *14*, 63–72. [[CrossRef](#)] [[PubMed](#)]
43. Han, X.; Yang, Q.; Lin, L.; Xu, C.; Zheng, C.; Chen, X.; Han, Y.; Li, M.; Cao, W.; Cao, K.; et al. Interleukin-17 enhances immunosuppression by mesenchymal stem cells. *Cell Death Differ.* **2014**, *21*, 1758–1768. [[CrossRef](#)]
44. Wu, F.; Xu, J.; Huang, Q.; Han, J.; Duan, L.; Fan, J.; Lv, Z.; Guo, M.; Hu, G.; Chen, L.; et al. The Role of Interleukin-17 in Lung Cancer. *Mediat. Inflamm.* **2016**, *2016*, 8494079. [[CrossRef](#)] [[PubMed](#)]
45. Huang, J.; Gou, H.; Yao, J.; Yi, K.; Jin, Z.; Matsuoka, M.; Zhao, T. The noncanonical role of EZH2 in cancer. *Cancer Sci.* **2021**, *112*, 1376–1382. [[CrossRef](#)] [[PubMed](#)]
46. Gardner, E.E.; Lok, B.H.; Schneeberger, V.E.; Desmeules, P.; Miles, L.A.; Arnold, P.K.; Ni, A.; Khodos, I.; de Stanchina, E.; Nguyen, T.; et al. Chemosensitive Relapse in Small Cell Lung Cancer Proceeds through an EZH2-SLFN11 Axis. *Cancer Cell* **2017**, *31*, 286–299. [[CrossRef](#)]
47. Kim, S.; Yu, N.-K.; Kaang, B.-K. CTCF as a multifunctional protein in genome regulation and gene expression. *Exp. Mol. Med.* **2015**, *47*, e166. [[CrossRef](#)]
48. Rhodes, D.R.; Yu, J.; Shanker, K.; Deshpande, N.; Varambally, R.; Ghosh, D.; Barrette, T.; Pandey, A.; Chinnaiyan, A.M. Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neo-plastic transformation and progression. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 9309–9314. [[CrossRef](#)]
49. Atienza, J.M.; Roth, R.B.; Rosette, C.; Smylie, K.J.; Kammerer, S.; Rehbock, J.; Ekblom, J.; Denissenko, M.F. Suppression of *RAD21* gene expression decreases cell growth and enhances cytotoxicity of etoposide and bleomycin in human breast cancer cells. *Mol. Cancer Ther.* **2005**, *4*, 361–368. [[CrossRef](#)]
50. Ramsey, J.; Butnor, K.; Peng, Z.; LeClair, T.; van der Velden, J.; Stein, G.; Lian, J.; Kinsey, C.M. Loss of RUNX1 is associated with aggressive lung adenocarcinomas. *J. Cell. Physiol.* **2017**, *233*, 3487–3497. [[CrossRef](#)]
51. Kim, Y.; Bin Lee, B.; Kim, D.; Um, S.; Cho, E.Y.; Han, J.; Shim, Y.M.; Kim, D.-H. Clinicopathological Significance of RUNX1 in Non-Small Cell Lung Cancer. *J. Clin. Med.* **2020**, *9*, 1694. [[CrossRef](#)] [[PubMed](#)]
52. Kulkarni, M.; Tan, T.Z.; Sulaiman, N.B.S.; Lamar, J.M.; Bansal, P.; Cui, J.; Qiao, Y.; Ito, Y. RUNX1 and RUNX3 protect against YAP-mediated EMT, stemness and shorter survival outcomes in breast cancer. *Oncotarget* **2018**, *9*, 14175–14192. [[CrossRef](#)]
53. Meng, G.; Wei, J.; Wang, Y.; Qu, D.; Zhang, J. MiR-21 regulates immunosuppression mediated by myeloid-derived suppressor cells by impairing RUNX1-YAP interaction in lung cancer. *Cancer Cell Int.* **2020**, *20*, 495. [[CrossRef](#)] [[PubMed](#)]
54. Li, L.; Mei, T.; Zeng, Y. HDAC2 promotes the migration and invasion of non-small cell lung cancer cells via upregulation of fibronectin. *Biomed. Pharmacother.* **2016**, *84*, 284–290. [[CrossRef](#)] [[PubMed](#)]
55. Werbajh, S.E.; Urtreger, A.J.; Puricelli, L.I.; De Lustig, E.S.; Joffé, E.B.D.K.; Kornblihtt, A.R. Downregulation of fibronectin transcription in highly metastatic adenocarcinoma cells. *FEBS Lett.* **1998**, *440*, 277–281. [[CrossRef](#)]
56. Michalik, L.; Wahli, W. PPARs Mediate Lipid Signaling in Inflammation and Cancer. *PPAR Res.* **2008**, *2008*, 134059. [[CrossRef](#)]
57. Theocharis, S.; Kanelli, H.; Politi, E.; Margeli, A.; Karkandaris, C.; Philippides, T.; Koutselinis, A. Expression of peroxisome proliferator activated receptor-gamma in non-small cell lung carcinoma: Correlation with histological type and grade. *Lung Cancer* **2002**, *36*, 249–255. [[CrossRef](#)]
58. Sabatino, L.; Fucci, A.; Pancione, M.; Colantuoni, V. PPARγ Epigenetic Dereglulation and Its Role in Colorectal Tumorigenesis. *PPAR Res.* **2012**, *2012*, 687492. [[CrossRef](#)]
59. Platta, C.S.; Greenblatt, D.Y.; Kunnimalaiyaan, M.; Chen, H. Valproic Acid Induces Notch1 Signaling in Small Cell Lung Cancer Cells. *J. Surg. Res.* **2008**, *148*, 31–37. [[CrossRef](#)]
60. Park, H.K.; Han, B.R.; Park, W.H. Combination of arsenic trioxide and valproic acid efficiently inhibits growth of lung cancer cells via G2/M-phase arrest and apoptotic cell death. *Int. J. Mol. Sci.* **2020**, *21*, 2649. [[CrossRef](#)]

61. Hubaux, R.; Vandermeers, F.; Cosse, J.-P.; Crisanti, C.; Kapoor, V.; Albelda, S.M.; Mascaux, C.; Delvenne, P.; Hubert, P.; Willems, L. Valproic acid improves second-line regimen of small cell lung carcinoma in preclinical models. *ERJ Open Res.* **2015**, *1*, 00028–2015. [[CrossRef](#)] [[PubMed](#)]
62. Shirsath, N.; Rathos, M.; Chaudhari, U.; Sivaramakrishnan, H.; Joshi, K. Potentiation of anticancer effect of valproic acid, an antiepileptic agent with histone deacetylase inhibitory activity, by the cyclin-dependent kinase inhibitor P276-00 in human non-small-cell lung cancer cell lines. *Lung Cancer* **2013**, *82*, 214–221. [[CrossRef](#)] [[PubMed](#)]
63. Rakba, N.; Melhaoui, A.; Loyer, P.; Delcros, J.G.; Morel, I.; Lescoat, G. Bgugaine, a pyrrolidine alkaloid from *Arisarum vulgare*, is a strong hepatotoxin in rat and human liver cell cultures. *Toxicol. Lett.* **1999**, *104*, 239–248. [[CrossRef](#)]
64. Rakba, N.; Melhaoui, A.; Rissel, M.; Morel, I.; Loyer, P.; Lescoat, G. Irniine, a pyrrolidine alkaloid, isolated from *Arisarum vulgare* can induce apoptosis and/or necrosis in rat hepatocyte cultures. *Toxicol.* **2000**, *38*, 1389–1402. [[CrossRef](#)]
65. Garcia-Quiroz, J.; Camacho, J. Astemizole: An old anti-histamine as a new promising anti-cancer drug. *Anti-Cancer Agents Med. Chem.* **2012**, *11*, 307–314. [[CrossRef](#)]
66. Gray, K.A.; Daugherty, L.C.; Gordon, S.M.; Seal, R.L.; Wright, M.W.; Bruford, E.A. Genenames.org: The HGNC resources in 2013. *Nucleic Acids Res.* **2012**, *41*, D545–D552. [[CrossRef](#)]
67. Gundersen, G.W.; Jones, M.R.; Rouillard, A.; Kou, Y.; Monteiro, C.D.; Feldmann, A.S.; Hu, K.S.; Ma'Ayan, A. GEO2Enrichr: Browser extension and server app to extract gene sets from GEO and analyze them for biological functions. *Bioinformatics* **2015**, *31*, 3060–3062. [[CrossRef](#)]
68. Wang, Z.; Monteiro, C.D.; Jagodnik, K.M.; Fernandez, N.F.; Gundersen, G.W.; Rouillard, A.D.; Jenkins, S.L.; Feldmann, A.S.; Hu, K.S.; McDermott, M.G.; et al. Extraction and analysis of signatures from the Gene Expression Omnibus by the crowd. *Nat. Commun.* **2016**, *7*, 12846. [[CrossRef](#)]
69. Kibbe, W.A.; Arze, C.; Felix, V.; Mitraka, E.; Bolton, E.; Fu, G.; Mungall, C.J.; Binder, J.X.; Malone, J.; Vasant, D.; et al. Disease Ontology 2015 update: An expanded and updated database of Human diseases for linking bio-medical knowledge through disease data. *Nucleic Acids Res.* **2015**, *43*, D1071–D1078. [[CrossRef](#)]
70. Law, V.; Knox, C.; Djoumbou, Y.; Jewison, T.; Guo, A.C.; Liu, Y.; Maciejewski, A.; Arndt, D.; Wilson, M.; Neveu, V.; et al. DrugBank 4.0: Shedding new light on drug metabolism. *Nucleic Acids Res.* **2013**, *42*, D1091–D1097. [[CrossRef](#)]
71. Boedigheimer, M.J.; Wolfinger, R.D.; Bass, M.B.; Bushel, P.R.; Chou, J.W.; Cooper, M.; Corton, J.C.; Fostel, J.; Hester, S.; Lee, J.S.; et al. Sources of variation in baseline gene expression levels from toxicogenomics study control animals across multiple laboratories. *BMC Genom.* **2008**, *9*, 285. [[CrossRef](#)] [[PubMed](#)]
72. Leek, J.T.; Storey, J.D. Capturing Heterogeneity in Gene Expression Studies by Surrogate Variable Analysis. *PLoS Genet.* **2007**, *3*, 1724–1735. [[CrossRef](#)] [[PubMed](#)]
73. Chen, E.Y.; Xu, H.; Gordonov, S.; Lim, M.P.; Perkins, M.H.; Ma'Ayan, A. Expression2Kinases: mRNA profiling linked to multiple upstream regulatory layers. *Bioinformatics* **2011**, *28*, 105–111. [[CrossRef](#)]
74. Kuleshov, M.V.; Jones, M.R.; Rouillard, A.D.; Fernandez, N.F.; Duan, Q.; Wang, Z.; Koplev, S.; Jenkins, S.L.; Jagodnik, K.M.; Lachmann, A.; et al. Enrichr: A comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **2016**, *44*, W90–W97. [[CrossRef](#)]
75. Kanehisa, M.; Furumichi, M.; Sato, Y.; Ishiguro-Watanabe, M.; Tanabe, M. KEGG: Integrating viruses and cellular organisms. *Nucleic Acids Res.* **2021**, *49*, D545–D551. [[CrossRef](#)] [[PubMed](#)]
76. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **2019**, *28*, 1947–1951. [[CrossRef](#)] [[PubMed](#)]
77. Ogata, H.; Goto, S.; Sato, K.; Fujibuchi, W.; Bono, H.; Kanehisa, M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **2000**, *28*, 29–34. [[CrossRef](#)]
78. Musa, A.; Ghoraie, L.S.; Zhang, S.-D.; Galzko, G.; Yli-Harja, O.; Dehmer, M.; Haibe-Kains, B.; Emmert-Streib, F. A review of connectivity map and computational approaches in pharmacogenomics. *Briefings Bioinform.* **2017**, *19*, 506–523. [[CrossRef](#)]
79. Zaravinos, A.; Roufas, C.; Nagara, M.; de Lucas Moreno, B.; Oblovatskaya, M.; Efstathiades, C.; Ayiomamitis, G.D. Cytolytic activity correlates with the mutational burden and deregulated expression of immune checkpoints in colorectal cancer. *J. Exp. Clin. Cancer Res.* **2019**, *38*, 364. [[CrossRef](#)]
80. Ritchie, M.E.; Belinda, P.; Wu, D.; Hu, Y.; Law, C.W.; Shi, W.; Smyth, G.K. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **2015**, *43*, e47. [[CrossRef](#)]
81. Uhlén, M.; Fagerberg, L.; Hallström, B.M.; Lindskog, C.; Oksvold, P.; Mardinoglu, A.; Sivertsson, Å.; Kampf, C.; Sjöstedt, E.; Asplund, A.; et al. Proteomics. Tissue-Based Map of the Human Proteome. *Science* **2015**, *347*, 1260419. [[CrossRef](#)] [[PubMed](#)]