

Integrative Analysis for Identification of Therapeutic Targets and Prognostic Signatures in Non-Small Cell Lung Cancer

Bioinformatics and Biology Insights
Volume 16: 1–12
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/11779322221088796



Özgür Cem Erkin¹, Betül Cömerpay¹ and Esra Göv

Department of Bioengineering, Faculty of Engineering, Adana Alparslan Türkeş Science and Technology University, Adana, Turkey.

ABSTRACT: Differential expressions of certain genes during tumorigenesis may serve to identify novel manageable targets in the clinic. In this work with an integrated bioinformatics approach, we analyzed public microarray datasets from Gene Expression Omnibus (GEO) to explore the key differentially expressed genes (DEGs) in non-small cell lung cancer (NSCLC). We identified a total of 984 common DEGs in 252 healthy and 254 NSCLC gene expression samples. The top 10 DEGs as a result of pathway enrichment and protein–protein interaction analysis were further investigated for their prognostic performances. Among these, we identified high expressions of *CDC20*, *AURKA*, *CDK1*, *EZH2*, and *CDKN2A* genes that were associated with significantly poorer overall survival in NSCLC patients. On the contrary, high mRNA expressions of *CBL*, *FYN*, *LRKK2*, and *SOCS2* were associated with a significantly better prognosis. Furthermore, our drug target analysis for these hub genes suggests a potential use of Trichostatin A, Pracinostat, TGX-221, PHA-793887, AG-879, and IMD0354 antineoplastic agents to reverse the expression of these DEGs in NSCLC patients.

KEYWORDS: Non-small cell lung cancer, differential gene expressions, integrative bioinformatics analysis, drug target potential analysis

RECEIVED: December 10, 2021. **ACCEPTED:** February 27, 2022.

TYPE: Original Research

FUNDING: The author(s) received no financial support for the research, authorship, and/or publication of this article.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Esra Göv, Department of Bioengineering, Faculty of Engineering, Adana Alparslan Türkeş Science and Technology University, Balcalı Mah., Çatalan Caddesi No: 201/1, Sarıçam, 01250 Adana, Turkey. Email: egov@atu.edu.tr

Introduction

Lung cancer is one of the deadliest diseases all around the world. The GLOBOCAN estimations by the International Agency for Research on Cancer predict approximately 13 million new cancer cases by the year 2040. Currently, lung cancer is the most commonly diagnosed form of cancer (11.6% of total cases) and the leading cause of cancer deaths (18.4% of total deaths).¹ Histologically, lung cancer can be divided into non-small cell lung cancer (NSCLC), which accounts for approximately 85% of lung cancer cases and small cell lung cancer.² The lack of early-stage symptoms and effective diagnostic markers restrain the treatment success in NSCLC. Thus, most of the patients are diagnosed at an advanced stage, and half of them have distant metastatic disease at initial diagnosis.³ Over the last decade, there have been considerable improvements in chemotherapy, radiation therapy, surgery, and targeted therapy for lung cancer. Especially with the recent advances in molecular biology, significant progress has been made through molecule-targeted therapy in NSCLC. For instance, it has been shown that about 20% of Caucasian and 50% of Asian NSCLC patients had mutations on their EGFR (epidermal growth factor receptor) genes.⁴ However, with the application of EGFR targeting small tyrosine kinase inhibitors (EGFR-TKIs) erlotinib and gefitinib, both response rate and median survival of these patients were found to be improved.⁵ Also, approximately 7% of NSCLC patients bear an activated *ALK* gene, but treatment with ALK inhibitor crizotinib has been shown to improve both response and 6-month

progression-free survival rates in these patients.⁶ Nonetheless, the 5-year survival rates of NSCLC patients remain low with a poor prognosis due to the development of intrinsic or acquired chemoresistance against therapeutic drugs.⁷

Non-small cell lung cancer is a result of the accumulation of several genetic and epigenetic modifications, which could have originated from multiple reasons.⁸ The discovery and characterization of new prognostic or diagnostic markers together with enhanced therapeutic approaches for NSCLC are of top priority for the successful treatment of this disease. Unfortunately, information on the heterogeneous nature of the tumor and the involvement of affecting factors in the process of NSCLC tumor development are far from completely resolved. Therefore, it is highly important to shed light on the molecular mechanisms governing the pathogenesis of NSCLC and to identify effective diagnostic and/or prognostic biomarkers for novel treatment options. High-throughput technologies such as microarrays and integrated bioinformatics methods are used to obtain gene alterations during tumorigenesis and to identify novel prognostic markers in patients with cancer.^{9,10} For instance, in a recent study, Huang and Gao¹¹ have demonstrated that *CDC20*, *CENPF*, *KIP2C*, and *ZWINT* genes were differentially expressed in NSCLC tissues. Similarly, Xiao et al¹² have identified *CCNB1*, *CCNA2*, *CEP55*, *PBK*, and *HMMR* as hub genes and key differentially expressed genes (DEGs) associated with NSCLC by bioinformatics analyses. Interestingly, Wang et al¹³ have shown *CCND1* as the most enriched gene and a potential prognostic biomarker in NSCLC

Table 1. Transcriptome datasets employed in the present study.

| SOURCE—ID | PURPOSE | NO. OF TUMOR SAMPLES | NO. OF CONTROL SAMPLES | REFERENCES |
|---------------|---|----------------------|------------------------|--------------------------------------|
| GEO-GSE18842 | Gene expression analysis of human lung cancer and control samples | 46 | 45 | Sanchez-Palencia et al ¹⁷ |
| GEO-GSE19804 | Genome-wide screening of transcriptional modulation in non-smoking female lung cancer in Taiwan | 60 | 60 | Lu et al ¹⁸ |
| GEO-GSE27262 | Gene expression profiling of non-small cell lung cancer in Taiwan | 25 | 25 | Wei et al ¹⁹ |
| GEO-GSE102287 | Gene and microRNA expression data from African Americans and European Americans with non-small cell lung cancer | 123 | 122 | Mitchell et al ²⁰ |

Abbreviation: GEO, Gene Expression Omnibus.

through a gene set enrichment analysis. More recently, Zhang et al¹⁴ have identified TOP2A, CCNB1, BIRC5, and TTK as well as miR-21-5p and miR-31-5p to be significantly associated with NSCLC prognosis through an integrative analysis of mRNA and miRNA expression profiles. Nevertheless, the genes that are discovered by 1 cohort might be difficult to be identified in other cohorts.¹⁵ For this reason, it is essential to validate genes in several independent studies.

In this study, we sought to identify potential therapeutic targets or prognostic biomarkers among the DEGs associated with NSCLC through an integrated bioinformatics approach. For this purpose, we retrieved 4 different microarray datasets from Gene Expression Omnibus (GEO) database and screened for DEGs between NSCLC tumor and neighboring normal tissues. After gene set enrichment analysis to identify associated biological processes, a protein–protein interaction (PPI) network analysis was performed to elucidate potential key DEGs. We also explored the significance of candidate key DEGs and their correlation to patient prognosis through survival analysis. Finally, potential therapeutic drugs that may target and reverse the expression of these key DEGs were predicted by using the L1000CDS2 signature search engine.

Materials and Methods

Microarray datasets

A comprehensive database search was conducted for identifying appropriate datasets including NSCLC tumor tissue and matched adjacent normal samples from the public GEO database.¹⁶ To avoid microarray platform differences, the datasets originating from Affymetrix microarrays utilizing Human Genome U133 Plus 2.0 chips (Thermo Fisher Scientific, Inc., Waltham, MA, USA) were selected. Four datasets with accession numbers GSE18842, GSE19804, GSE27262, and GSE102287 were identified and downloaded for further analysis (Table 1). GSE19804¹⁸ included 60 pairs of NSCLC tumor and matched adjacent normal lung tissue, while GSE18842¹⁷ comprised 46 NSCLC tumors and 45 paired

controls, GSE27262¹⁹ contained 25 tumors and normal tissue pairs from stage I lung adenocarcinoma, and GSE102287²⁰ contained 66 matched NSCLC tumor and normal tissues. A total of 506 gene expression samples including 252 healthy and 254 NSCLC tissues were evaluated in this study.

Data preprocessing and identification of DEGs

To identify DEGs, raw data in the form of CEL files were downloaded from the GEO database. Affy package of R/Bioconductor platform (version 3.6) was utilized, and the gene expression datasets were normalized by using the Robust Multi-Array (RMA) techniques.²¹ Linear models for microarray data (LIMMA) method²² was used in statistical analysis of each dataset by comparing gene expressions in NSCLC tissues with neighboring healthy tissues to identify DEGs. Differentially expressed genes in each dataset were selected according to computed *P* values <.05. The regulatory pattern of each DEG was determined by fold changes (FCs) of up- and downregulation (FC >2.0 and <0.5, respectively), and at least 50% change was considered significant.

Functional and pathway enrichment analysis

To identify functional annotations significantly associated with the gene products, pathway enrichment analyses were performed using ConsensusPathDB.²³ Genomic, chemical, and systematic functional information of DEG pathways were provided by Kyoto Encyclopedia of Genes and Genomes (KEGG).²⁴ For each of the predefined sets, a *P* value is calculated according to the hypergeometric test, and enrichment results with *P* < .01 were considered statistically significant.

PPI analysis

Human protein interactome containing 43 219 physical interactions between 2294 human proteins was collected from the BioGRID database.²⁵ The physical interactions of the proteins

encoded by the common genes of all datasets were analyzed in Cytoscape²⁶ by the reconstruction of PPI networks. Hub genes were identified according to degree and betweenness scores.

Prognostic performance analysis

Survival analyses were performed by Kaplan–Meier Plotter web tool²⁷ using mRNA data to identify the prognostic performances of each PPI hub gene. The subjects were partitioned into low- and high-risk groups according to their hazard ratio index. The prognostic capabilities of the genes were characterized through Kaplan–Meier plots, and the log-rank P value $<.05$ was considered the cutoff value to describe statistical significance. In addition, prognostic performances of the genes were validated using datasets with available clinical information obtained from CAARRAY, GSE14814, GSE19188, GSE29013, GSE30219, GSE31210, GSE3141, GSE31908, GSE37745, GSE43580, GSE4573, GSE50081, GSE8894, and TCGA in total of 1926 samples.

Drug target potential analysis

L1000CDS2²⁸ is a signature search engine which lists small molecules that are predicted to mimic or reverse the input gene expression profiles utilizing LINCS L1000 data.²⁹ We performed L1000CDS2 analysis using the hub gene list as an input, to assess the potential of agents that may reverse the expression of up- or downregulated hub genes in NSCLC. Resultant small molecules were selected according to 2 main criteria: (1) perturbed human cells as lung cancer cell lines and (2) top 2 overlap scores.

Results

Transcriptomic signatures of NSCLC

We comparatively analyzed 4 distinct transcriptome datasets composed of NSCLC and healthy tissue samples. After preprocessing of all 4 datasets, DEGs were identified between tumor samples and matched normal tissues according to our threshold criteria ($FC > 2.0$ and $FC < 0.5$ with computed $P < .05$). We obtained 3381, 1581, 2488, and 1934 respective DEGs in GSE18842, GSE19804, GSE27262, and GSE102287 datasets. Among these, 1451, 455, 966, and 621 genes were upregulated, while 1951, 1075, 1537, and 1322 genes were downregulated in GSE18842, GSE19804, GSE27262, and GSE102287 datasets, respectively (Figure 1A and Supplementary Table 1). Combination and comparison of all 4 total DEG groups resulted in the identification of 984 common DEGs (Figure 1B), with 246 upregulated (Figure 1C) and 734 downregulated (Figure 1D) genes. It was determined that the number of downregulated genes was higher than the number of upregulated genes in all datasets.

Biological insights of common transcriptomic signatures

The pathway enrichment analysis based on situated common up- and downregulated DEGs indicated significant results for general cancer signaling, and metabolic and immune systems-related pathways. Our results showed that commonly upregulated genes were significantly enriched in 10 KEGG pathways, which included cell cycle, p53 signaling pathway, and extracellular matrix (ECM)–receptor interaction (Figure 2A). On the contrary, the downregulated common DEGs were enriched in 32 terms, including malaria, cell adhesion molecule, and cAMP signaling pathway (Figure 2B). The top 15 downregulated pathways were used in Figure 2B. The rest of the enriched terms can be provided upon request.

NSCLC-specific PPI network

To identify hub genes and to reconstruct a PPI network, we utilized the BioGRID database and analyzed proteins from 984 identified DEGs with Cytoscape. In topological analyses of the NSCLC-specific PPI network, degree (local-based) and betweenness centrality (global-based) metrics³⁰ were used to identify the highly connected proteins, that is, hub proteins (Figure 3A), which might play an important role in cancer pathogenesis. According to their degree and betweenness scores, the top 10 hub genes with strong interactions with others were determined (Figure 3B). We selected Cbl Proto-oncogene (CBL), Enhancer of Zeste 2 Polycomb Repressive Complex 2 Subunit (EZH2), Cyclin Dependent Kinase 1 (CDK1), Cell Division Cycle 20 (CDC20), Cyclin Dependent Kinase Inhibitor 2A (CDKN2A), Aurora Kinase A (AURKA), FYN Proto-oncogene (FYN), ETS Transcription Factor ERG (ERG), Suppressor of Cytokine Signaling 2 (SOCS2), and Leucine Rich Repeat Kinase 2 (LRKK2) as the most significant nodes for further evaluation.

Prognostic hub genes in NSCLC

To evaluate the prognostic performances of the top 10 hub genes, we performed a survival analysis by Kaplan–Meier Plotter. The overall survival times for patients with NSCLC were obtained according to the low and high expressions of each hub gene (Figure 4). Our results showed that high mRNA expression of CDC20 (Hazard ratio [HR], 1.82; confidence interval [CI], 1.6–2.07) as well as AURKA (HR, 1.52; CI, 1.33–1.72), CDK1 (HR, 1.4; CI, 1.23–1.59), EZH2 (HR, 1.31; CI, 1.15–1.48), and CDKN2A (HR, 1.29; CI, 1.13–1.46) were associated with significantly poorer overall survival ($P < .05$) for NSCLC patients. According to our statistical analysis, these genes were among the commonly upregulated genes with significantly higher expression FC levels in NSCLC samples in all datasets (Figure 5). On the contrary,

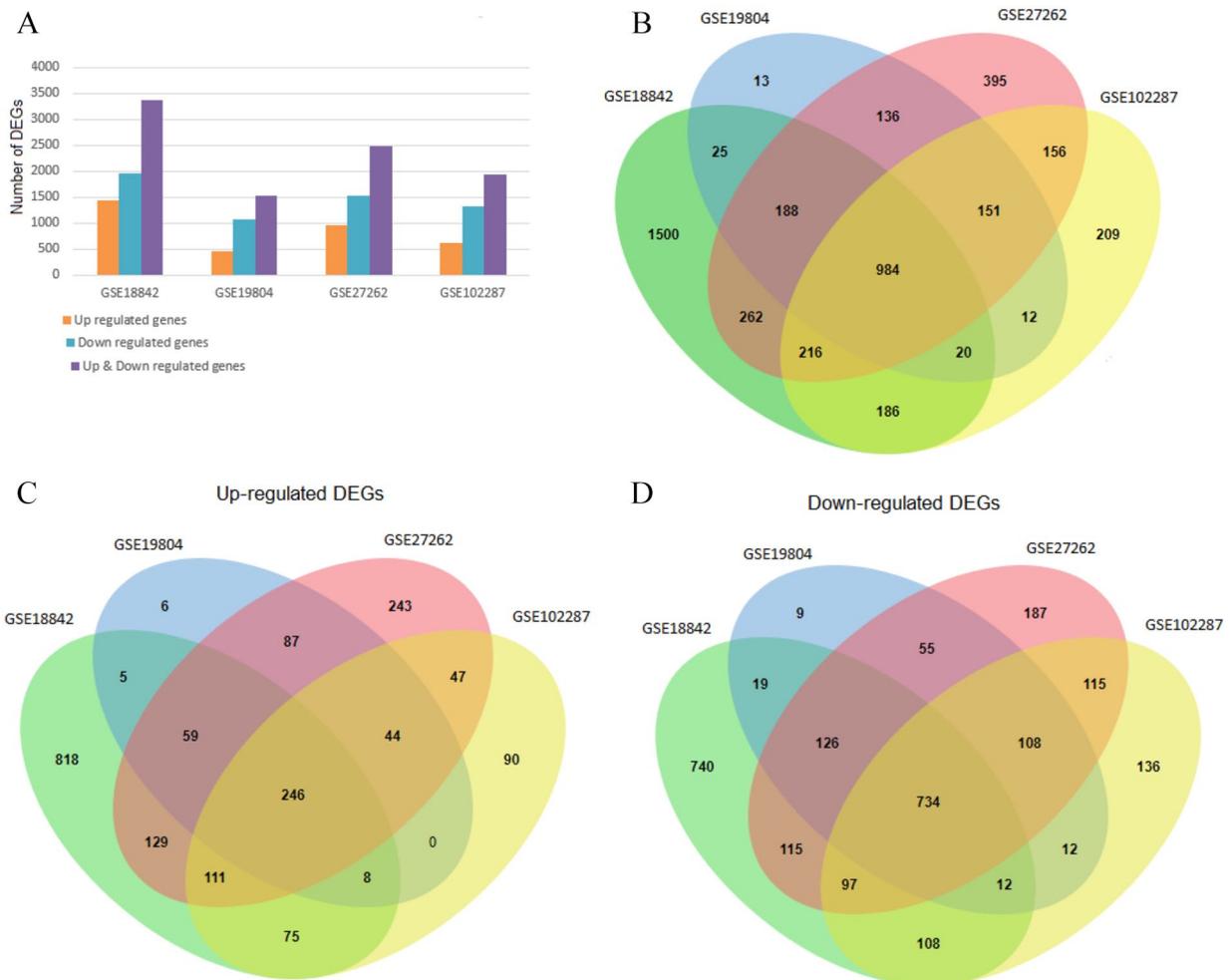


Figure 1. DEGs in non-small cell lung cancer. (A) The distribution of DEGs in non-small cell lung cancer–associated transcriptome datasets. Downregulation and upregulation of DEGs were represented by blue and orange colors, respectively. (B) The Venn diagram represents the comparison of DEGs among all datasets. (C) The Venn diagram represents the comparison of upregulated DEGs among all datasets. (D) The Venn diagram represents the comparison of downregulated DEGs among all datasets. DEGs indicate differentially expressed genes.

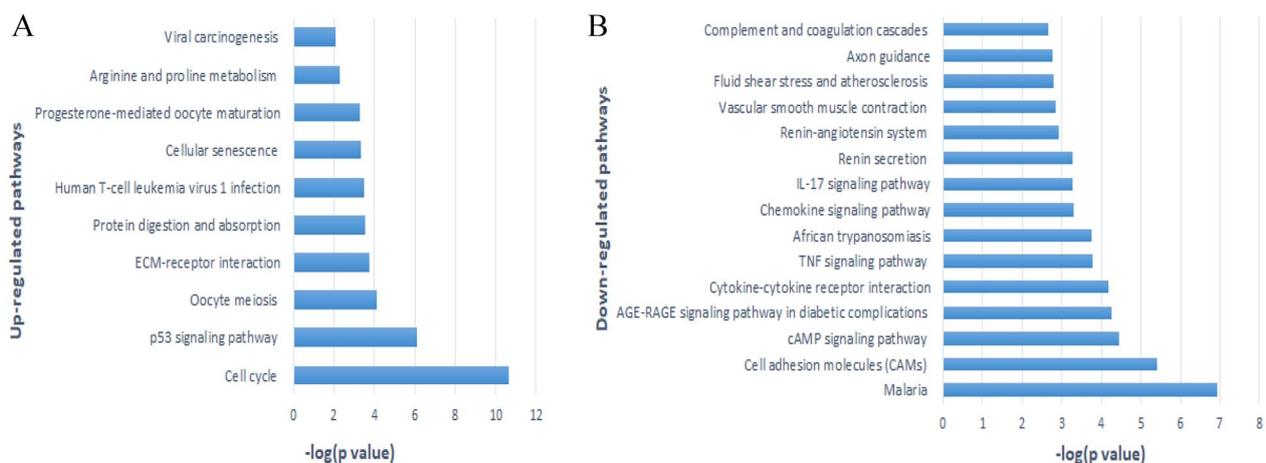


Figure 2. Statistically significant biological pathways of common DEGs: (A) upregulated pathways and (B) downregulated pathways. Up- and downregulated common DEGs list was used to obtain up- and downregulated pathways. In both cases, top 15 enriched pathways were represented. DEGs indicate differentially expressed genes.

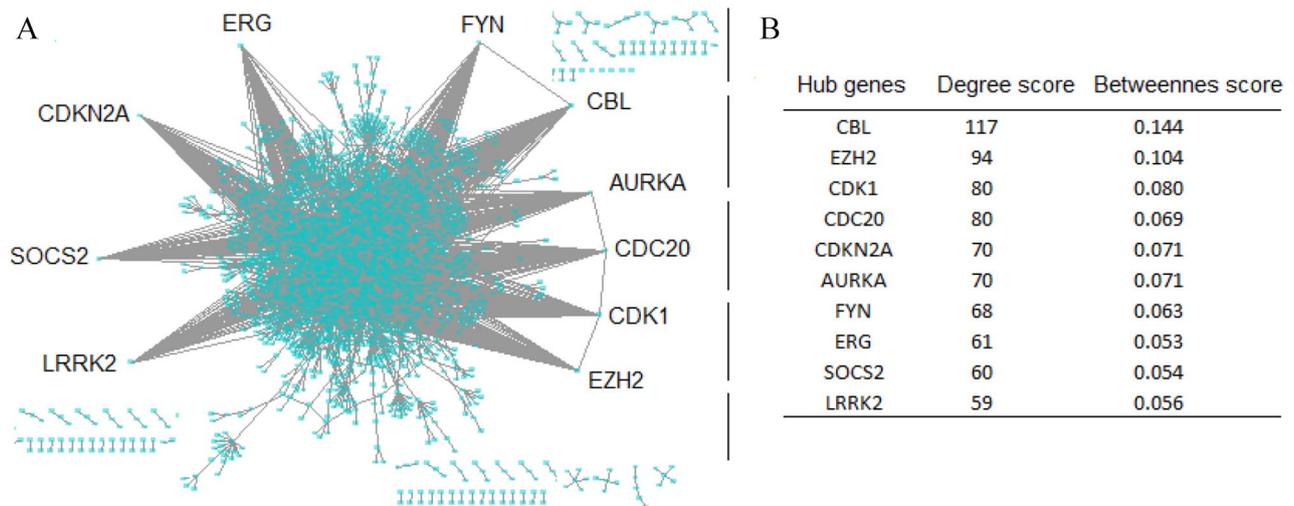


Figure 3. Non-small cell lung cancer-specific protein-protein interaction network and hub proteins. AURKA indicates Aurora Kinase A; CBL, Cbl Proto-oncogene; CDC20, Cell Division Cycle 20; CDK1, Cyclin Dependent Kinase 1; CDKN2A, Cyclin Dependent Kinase Inhibitor 2A; ERG, ETS Transcription Factor ERG; EZH2, Enhancer of Zeste 2 Polycomb Repressive Complex 2 Subunit; FYN, FYN Proto-oncogene; LRRK2, Leucine Rich Repeat Kinase 2; SOCS2, Suppressor of Cytokine Signaling 2.

high mRNA expression of CBL (HR, 0.73; CI, 0.62-0.86) as well as FYN (HR, 0.71; CI, 0.62-0.80), LRRK2 (HR, 0.62; CI, 0.52-0.73), and SOCS2 (HR, 0.62; CI, 0.53-0.74) were associated with significantly better prognosis ($P < .05$). Accordingly, expression levels of these genes were found to have significantly lower FC values in NSCLC tumor samples in all datasets (Figure 5). Therefore, it is logical to predict a poor prognosis associated with these 4 downregulated hub genes. Unfortunately, ERG gene was found to have no statistically significant prognostic potential in a total of 1926 NSCLC samples.

Drug target potential of the hub genes

To assess the potential drugs to therapeutically target identified hub proteins, we have utilized the LINCS L1000 connectivity map data and characteristic direction signature search engine L1000CDS2. The list of hub genes was entered into a web tool to search for substances that can reverse the expression changes in identified DEGs. Thirteen drugs or small molecules were identified, showing potential with an overlap score of >0.4 to reverse expression profiles on upregulated and/or downregulated gene expressions in lung cancer cell lines (Table 2). A considerable portion (6 out of 13) of the drugs was antineoplastic agents, which were suggested for the treatment and management of the cancers. The remaining drug candidates were originally used for purposes other than cancer treatments; however, they possess the potential to reverse the expression of the top 10 DEGs as shown in our analysis.

Discussion

Lung cancer tumorigenesis, progression, and metastasis are very complicated processes involving defects in multiple genes

and cellular pathways.³¹ Multiple interacting DEGs affecting other genes may constitute the core functional network of genes in promoting carcinogenesis.³² For an improved diagnosis and treatment, it is crucial to discover these abnormal genes and understand their roles in the molecular mechanisms of NSCLC. Developments in microarray and high throughput technologies allow us to detect cancer etiology by examining abnormalities at the whole-genome level. These technologies have been widely used to predict the potential therapeutic targets for cancers. On the contrary, a combination of large data on chemical perturbational profiles of human cell lines with differential gene expression analysis may provide not only information on abnormal gene activities but also some important clues on therapeutic options to overcome these defects in NSCLC.

In this study, we identified 246 upregulated, 734 downregulated, and a total of 984 common DEGs between NSCLC and normal tissue samples by using public gene expression profiles of GSE18842, GSE19804, GSE27262, and GSE102287 datasets.

The original data on these datasets were generated to identify SEMA5A,¹⁸ protein arginine methyltransferase 5,³³ and methylosome protein 50¹⁹ as potential biomarkers in NSCLC. On the contrary, GSE18842 was generated to analyze DEGs as a function of tumor type, stage, and differentiation grade in NSCLC. Keratin 15 and plakophilin 1 were identified as potentially good markers to distinguish squamous cell carcinoma, whereas a significant downregulation of desmoglein 3 was observed in early-stage tumor samples.¹⁷ A comparative transcriptome profiling was utilized to show coding and noncoding RNA differences between NSCLC from different human races. In this study, the dataset GSE102287 was generated to show 40 different novel population-specific gene expressions

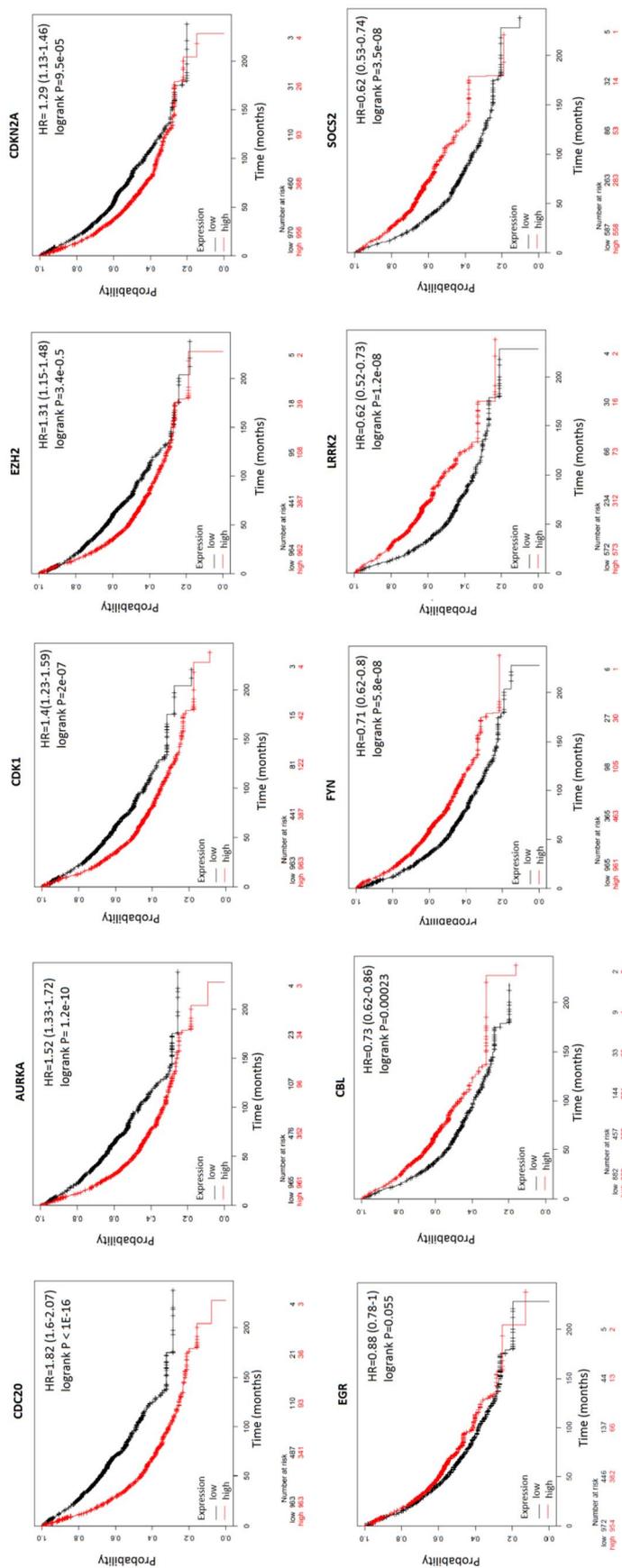


Figure 4. Survival analysis of hub differentially expressed genes. Kaplan-Meier Plot for patients with non-small cell lung cancer obtained from several independent databases. Downregulated expression is represented in a black plus sign, whereas upregulated expression in a red plus sign. AURKA indicates Aurora Kinase A; CBL, Cbl Proto-oncogene; CDC20, Cell Division Cycle 20; CDK1, Cyclin Dependent Kinase 1; ERG, ETS Transcription Factor ERG; FYN, FYN Proto-oncogene.

| Hub gene | GSE18842 | GSE19804 | GSE27262 | GSE102287 |
|----------|----------|----------|----------|-----------|
| AURKA | 7.659 | 2.489 | 4.050 | 3.669 |
| CBL | 0.343 | 0.294 | 0.291 | 0.300 |
| CDC20 | 17.505 | 3.460 | 5.455 | 4.288 |
| CDK1 | 10.797 | 2.395 | 3.918 | 2.933 |
| CDKN2A | 2.808 | 2.672 | 4.583 | 2.308 |
| ERG | 0.236 | 0.281 | 0.222 | 0.286 |
| EZH2 | 5.888 | 2.293 | 4.240 | 2.992 |
| FYN | 0.395 | 0.358 | 0.254 | 0.442 |
| LRRK2 | 0.062 | 0.231 | 0.258 | 0.248 |
| SOCS2 | 0.226 | 0.195 | 0.158 | 0.234 |

█ Up-regulated
█ Down-regulated

Figure 5. Fold changes of hub genes in each dataset. Red and blue grades represent upregulated and downregulated genes, respectively. AURKA indicates Aurora Kinase A; CBL, Cbl Proto-oncogene; CDC20, Cell Division Cycle 20; CDK1, Cyclin Dependent Kinase 1; ERG, ETS Transcription Factor ERG; EZH2, Enhancer of Zeste 2 Polycomb Repressive Complex 2 Subunit; FYN, FYN Proto-oncogene; SOCS2, Suppressor of Cytokine Signaling 2.

Table 2. Potential drugs or small molecules that can reverse the expression change in identified hub proteins through LINCS L1000 connectivity map data.

| RANK | CANDIDATE DRUGS | CELL LINE | DOSE | TIME | OVERLAP SCORE | REVERSIBLE GENE EXPRESSIONS |
|------|---------------------------|-----------|----------|-------|---------------|----------------------------------|
| 1 | Z-Leu3-VS | HCC515 | 10.0 μm | 24.0h | 0.5556 | AURKA, CDC20, CDK1, CDKN2A, EZH2 |
| 2 | TGX-115 | A549 | 160.0 μm | 24.0h | 0.5556 | CDC20, CDK1, CDKN2A, EZH2, SOCS2 |
| 3 | Ingenol 3,20-dibenzoate | HCC515 | 10.0 μm | 24.0h | 0.5556 | AURKA, CDC20, CDK1, EZH2, SOCS2 |
| 4 | Maprotiline hydrochloride | A549 | 10.0 μm | 24.0h | 0.5556 | AURKA, CDC20, CDK1, CDKN2A, EZH2 |
| 5 | S1169 | A549 | 10.0 μm | 24.0h | 0.5556 | AURKA, CDC20, CDK1, EZH2, SOCS2 |
| 6 | RF 01079 | A549 | 10.0 μm | 24.0h | 0.5556 | AURKA, CDC20, CDK1, EZH2, SOCS2 |
| 7 | HY-11001 | A549 | 10.0 μm | 24.0h | 0.5556 | AURKA, CDC20, CDK1, EZH2, SOCS2 |
| 8 | Pracinostat | HCC515 | 3.33 μm | 24.0h | 0.5556 | CDC20, CDK1, CDKN2A, EZH2, SOCS2 |
| 9 | Trichostatin A | HCC515 | 10.0 μm | 24.0h | 0.4444 | AURKA, CDC20, CDK1, CDKN2A |
| 10 | NTNCB hydrochloride | HCC515 | 10.0 μm | 24.0h | 0.4444 | AURKA, CDC20, CDK1, EZH2 |
| 11 | Rottlerin | HCC515 | 10.0 μm | 24.0h | 0.4444 | AURKA, CDC20, CDK1, EZH2 |
| 12 | AG-879 | HCC515 | 10.0 μm | 24.0h | 0.4444 | AURKA, CDC20, CDK1, EZH2 |
| 13 | IMD 0354 | HCC515 | 10.0 μm | 24.0h | 0.4444 | AURKA, CDC20, CDK1, FYN |

Abbreviations: AURKA, Aurora Kinase A; CDC20, Cell Division Cycle 20; CDK1, Cyclin Dependent Kinase 1; CDKN2A, Cyclin Dependent Kinase Inhibitor 2A; DEGs, differentially expressed genes; EZH2, Enhancer of Zeste 2 Polycomb Repressive Complex 2 Subunit; FYN, FYN Proto-oncogene; SOCS2, Suppressor of Cytokine Signaling 2.

such as ARL17A, LRCC37A3, and KANSL1.²⁰ Although the data were supported and validated by several wet-lab experimentations, none of these original works utilized a similar integrated bioinformatics approach. All 4 datasets re-analyzed in a single set for the first time in our study.

In our PPI network analysis, we identified a series of hub genes, including *CDC20*, *AURKA*, *CDK1*, *EZH2*, *CDKN2A*, *CBL*, *FYN*, *LRRK2*, *SOCS2*, and *ERG* as the top 10 significant nodes. Among them, *CDC20* appears to act as a regulatory protein interacting with several other proteins such as mitotic spindle checkpoint protein MAD2L1 and anaphase-promoting complex/cyclosome (APC/C) at multiple points in the cell cycle. It is required for 2 microtubule-dependent

processes, nuclear movement before anaphase and chromosome separation.³⁴ Disorders involving the *CDC20* gene are mosaic variegated aneuploidy syndrome 1, neuronal ceroid lipofuscinosis, and prostate cancer.³⁵ *CDC20* was identified as frequently upregulated in many types of malignancies including NSCLC and suppressed by p53 expression.³⁶ Importantly, *CDC20* was identified among the key genes with prognostic value in NSCLC in several other studies using bioinformatics methods and similar datasets.^{37,38} Therefore, it may represent a potential molecular target and a critical molecular marker for NSCLC progression. All drugs on our candidate list were found to potentially interfere with *CDC20* expression.

The second hub gene *AURKA* is another cell cycle-related protein which is a serine/threonine kinase that takes part in the regulation of cell cycle progression. During mitosis, it associates with centrosomes and spindle microtubules, and plays an essential role in various mitotic events such as the establishment of mitotic spindle, centrosome duplication and separation in addition to chromosomal alignment, spindle assembly checkpoint, and cytokinesis. In the checkpoint response pathways that are critical for oncogenic transformation of cells, it acts as a key regulatory component of p53/TP53 pathway by phosphorylating and stabilizing p53 itself. Interestingly, *AURKA* is required for the initial activation of CDK1 at centrosomes.³⁹ Aberrant expression of *AURKA* is associated with several cancer types, including colorectal cancer, laryngeal squamous cell carcinoma, atypical teratoid rhabdoid tumors, as well as tetraploidy syndrome.⁴⁰ In a recent study using gene expression meta-analysis integrated with neural network algorithms, *AURKA* was identified as the most obvious class of hub genes associated with NSCLC.⁴¹ Furthermore, the amplification or activation of *AURKA*-induced impairment of the LKB1/AMPK axis was found to contribute to NSCLC initiation and progression, and suggests *AURKA* as a potential therapeutic target.⁴² In a few bioinformatics studies using very similar integrative approach and some of the datasets used in this study, *AURKA* was suggested as one of the hub genes in NSCLC.^{43,44} CDK1 plays a key role in the control of the eukaryotic cell cycle by modulating the centrosome cycle as well as the mitotic onset. Besides the promotion of G2-M and G1S transitions, CDK1 regulates G1 progress via association with multiple interphase cyclins and it is required for entry into both S phase and mitosis in higher cells. In addition, as a master modulator in cell cycle onset and progression, it phosphorylates and activates numerous proteins, including CDC20 and EZH2. Loss of CDK1 activity or the aberrant expression of CDK1 is involved in G2 phase arrest in many tumor types. Therefore, an interest has been developed to search for potent CDK1 inhibitors and it comprises an attractive target in oncology.⁴⁵ Our results together with some recent other bioinformatics studies utilizing GSE18842 dataset⁴⁶ support the critical role of CDK1 as an important oncological target in NSCLC as well. Among the other identified hub genes, EZH2 is a polycomb group protein and the core catalytic subunit of a protein complex that participates in transcriptional repression of the affected target genes by methylated lysines on histone H3 through multiple modes of action.⁴⁷ In tumor progression, epigenetic malfunctions display important roles in addition to genetic factors. An abnormal expression of EZH2 induces permanent silencing on some critical tumor suppressor genes, represses their transcription, and therefore contributes indirectly to tumor proliferation, invasion, and metastasis. Accordingly, it is highly expressed in several types of tumors such as breast cancer, prostate cancer, ovarian cancer, lymphoma, as well as lung cancer.⁴⁸ In NSCLC, overexpression of

EZH2 mRNA was found to be a negative prognostic indicator and it was suggested as a biomarker to predict the response to histone deacetylase (HDAC) inhibitors.⁴⁹ On the contrary, EZH2 itself is an actionable molecular target and there are several ongoing clinical trials of EZH2 in hematological malignancies.⁵⁰ In a recent study, benzomorpholine derivatives were reported as novel EZH2 inhibitors for anti-NSCLC activity.⁵¹ Perhaps EZH2 can be targeted molecularly with potent inhibitors in appropriate NSCLC patients for therapeutic purposes in the future. The next top hub gene found in our study, *CDKN2A* or *p16INK4a*, is a powerful inhibitor of CDK4/6 and capable of inducing cell cycle arrest in G1 and G2 phases. It acts as a tumor suppressor and its loss is a significant event in several cancer types. In the majority of the cases, *CDKN2A* is inactivated by homozygous deletions or through hypermethylation of the promoter region of the gene. However, in accordance with our data, a retained overexpression of *CDKN2A* may result in metastatic and invasive tumor phenotypes and is associated with poor prognosis.⁵²

CBL is a proto-oncogene associated with diseases such as juvenile myelomonocytic leukemia and other hematological cancers.⁵³ It acts as an E3 ubiquitin-protein ligase that can function as a negative regulator of many signaling pathways triggered by cell surface receptors such as FGFR1, FGFR2, PDGFRA, PDGFRB, and EGFR.⁵⁴ Although the role of CBL family proteins in NSCLC is largely unknown, in a few recent studies their upregulation were shown to inhibit especially mutated EGFR expression by mediating proteasome degradation of the protein.⁵⁵ Therefore, it may be speculated that the positive prognostic value of CBL in our study can be attributed to its overexpression and anticancer effects in patients with mutant EGFR expression.

Non-receptor tyrosine-protein kinase FYN is an oncogene that plays a role in several biological processes such as cell motility, cytoskeletal remodeling, cell adhesion, integrin-mediated signaling, and cell growth and survival. FYN is overexpressed in various cancers, including head and neck cancer, melanoma, squamous cell carcinoma, and prostate cancer.⁵⁶ High FYN expression was associated with epithelial-mesenchymal transition and has various roles in cancer-metastasis-related pathways.⁵⁷ On the contrary, the prognostic value of FYN was shown in a recent work aiming to predict metastasis in NSCLC.⁵⁸ Consistent with our results, they have shown a high FYN expression associated with an improved survival time. Although the reason is unknown, this contradiction with previous findings needs to be clarified by further studies in the future. One of the hub genes with a positive prognostic value in our PPI network analysis was *LRKK2*, which is a member of the leucine-rich repeat kinase family. *LRKK2* is a serine/threonine kinase that modulates a large number of proteins involved in multiple processes like vesicle trafficking, autophagy, and neuronal plasticity.⁵⁹ Abnormalities in this gene are largely associated with neurodegenerative disorders such as Parkinson's disease.⁶⁰ In a

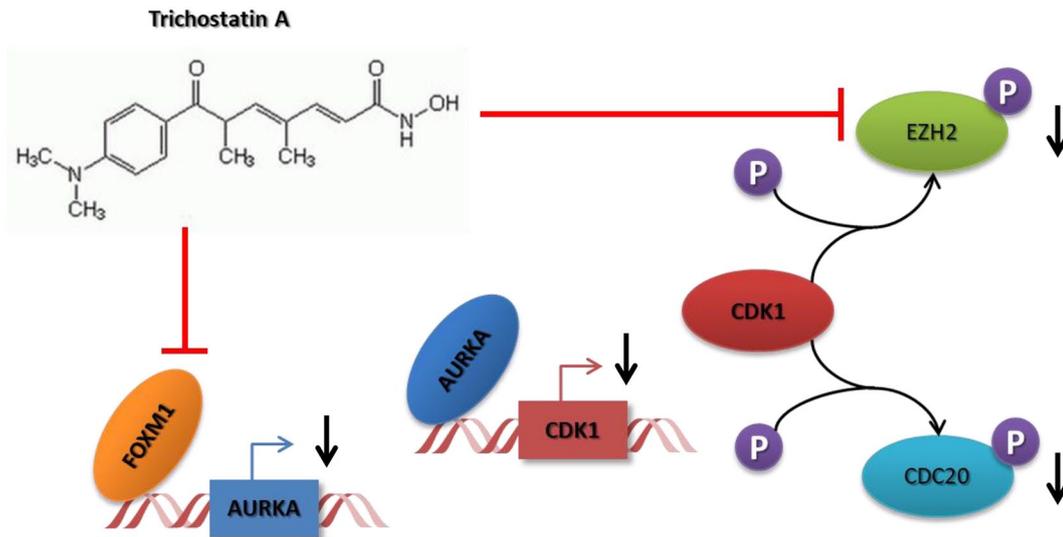


Figure 6. Proposed mechanism of action for reversing effects of Trichostatin A on hub gene expressions. Trichostatin A can interfere with AURKA expression through inhibiting FOXM1 activator. Reduced AURKA expression may interfere with CDK1 activation and thus expression of its downstream targets. Blocking EZH2 expression as well may help to decrease tumorigenic properties of NSCLC cells. Black arrow: downregulation. AURKA indicates Aurora Kinase A; CDK1, Cyclin Dependent Kinase 1; EZH2, Enhancer of Zeste 2 Polycomb Repressive Complex 2 Subunit; NSCLC, non-small cell lung cancer.

recent study, LRRK2 was shown to serve as a scaffold during activation of the WNT/ β -catenin pathway, which is activated in different types of human cancers.⁶¹ LRRK2 overexpression was shown to have important antitumor activities like suppression of proliferative, migrative, and invasive properties of tumor cells, and induction of apoptosis together with the arrest of the cell cycle.⁶² LRRK2 was shown as a hub gene in colon cancer by a bioinformatics approach⁶³ and was found among frequently mutated genes in lung squamous carcinoma.⁶⁴ SOCS2 gene is a member of the suppressor of cytokine signaling family proteins and it takes part in the negative control of transduction in this signaling cascade through the Janus kinase (JAK)- signal transducer and activator of transcription (STAT) pathway and in part by ubiquitination.⁶⁵ In general, SOCS2 seems to act largely as a positive prognostic factor in different human cancers including hepatocellular carcinoma, breast cancer, and colorectal cancer, with an exception in acute myeloid leukemia (AML) in a large cohort.⁶⁶ Inhibition of SOCS2 with different miRNAs in lung cancer results in cell proliferation and metastasis.⁶⁷ In addition, SOCS2 was identified as one of the downregulated genes in lung tumor tissue in a cDNA array-based study.⁶⁸ Our survival analysis results have shown a longer patient survival time with high mRNA expression of SOCS2; therefore, it may be considered as a positive prognostic marker of NSCLC.

Overall, *CDC20*, *AURKA*, *CDK1*, *EZH2*, and *LRKK2* were identified as important hub genes in other bioinformatics studies using various integrated analysis methods by utilizing some of the NSCLC datasets in our study. However, in the present study, we suggest *CDKN2A*, *CBL*, *FYN*, and *SOCS2* as novel important biomarker candidates in NSCLC.

Our drug target analysis considering the identified top 10 DEGs has revealed 13 drugs or small molecules demonstrating

a potential to reverse their aberrant expressions. Among these drugs, we identified Pracinostat, Trichostatin A (a histone deacetylase-HDAC inhibitor), S1169 (TGX-221-a PI3K inhibitor), HY-11001 (PHA-793887-a potent CDK inhibitor), AG-879, and IMD 0354 with antineoplastic effects. On the contrary, Z-Leu3-VS (a proteasome inhibitor), TGX-115 (an antimalarial agent), Ingenol 3,20-dibenzoate (an antiviral agent), Maprotiline hydrochloride (a selective noradrenaline reuptake inhibitor), NTNCB hydrochloride (a neuropeptide Y receptor Y5 antagonist), Rottlerin (antiplasmodial activity), and RF01079 with no functional information were also identified as potential drug candidates to reverse the expression of top 10 DEGs in our analysis. Pracinostat is a small-molecule HDAC inhibitor that may take part in chromatin remodeling through an accumulation of histone hyperacetylation. Combination therapy with pracinostat was suggested for elderly patients of AML in a phase II clinical trial.⁶⁹ However, we could not find any literature regarding the effects of pracinostat in NSCLC. Trichostatin A is another HDAC inhibitor that is identified in our study and is long known as an antineoplastic agent against NSCLC.⁷⁰ Although there are many reported studies on the antitumoral effects of Trichostatin A in NSCLC, perhaps the most critical impact was about the reversal of chemoresistance resulting from high IGF2BP2 expression.⁷¹ Trichostatin A was found to inhibit both EZH2⁷² and AURKA activator FOXM1⁷³ which had an FC value above 2 and thus upregulated in all datasets we analyzed (Supplementary Table 1). As AURKA-induced CDK1 expression may trigger both CDC20 and EZH2 gene expressions, we propose an action mechanism in which Trichostatin A could inhibit all these hub genes through its negative actions on FOXM1 gene expression (Figure 6).

AG-879 or Tyrphostin AG-879 is a specific TKI for ErbB2 (HER2). Although its antineoplastic effects are long known,⁷⁴ the utility of AG-879 is very limited in recent cancer studies. IMD0354 is an IKK-2 inhibitor V with therapeutic effects on inflammation and insulin resistance. The antineoplastic effect of IMD 0354 was reported on breast cancer stem cells to prevent chemoresistance in a murine model.⁷⁵ As of now, its antitumoral effects and their molecular mechanisms in NSCLC largely remain to be elucidated. In a recent study, it was shown that IMD0354 could effectively suppress cancer cell proliferation, invasion, and migration through inhibition of elevated expressions of transmembrane serine protease 4 (TMPRSS4) in NSCLC cells.⁷⁶ Therefore, IMD0345 has considerable potential for further development as a novel anticancer agent for NSCLC treatment. S1169 or TGX-221 is a potent PI3K isoform-specific inhibitor that was derived from a natural compound Quercetin.⁷⁷ The derivatives of TGX-221 have been utilized especially against prostate cancer cells as well as in xenograft models.⁷⁸ Antitumoral effects of TGX-221 in NSCLC are somewhat contradictory and the number of studies is limited. It was suggested that PI3K isoforms may functionally compensate for one another thus limiting the efficacy of single-agent treatment.⁷⁹ HY-11001 or PHA-793887 is a potent, ATP-competitive CDK inhibitor acting on CDK1, CDK2, CDK4, and CDK9, which has been used in a phase I clinical trial studying the treatment of advanced or metastatic solid tumors.⁸⁰ However, this study was not continued due to severe hepatic toxicity. There was no evidence for in vitro or in vivo antitumoral effects of PHA-793887 in NSCLC. Although there is a line of evidence for drug repurposing for other anti-malarial drugs such as hydroxychloroquine and artemisinin,⁸¹ there was no indication of TGX-115 usage in NSCLC treatment. We could not find any cancer-associated literature for remaining small molecules Z-Leu3-VS, TGX-115, Ingenol 3,20-dibenzoate, Maprotiline hydrochloride, NTNCB hydrochloride, or Rottlerin. On the contrary, RF01079 was found to be inactive in most of the tested bioassays according to PubChem. Overall, we identified Trichostatin A, IMD0354, and TGX-221 as potent drug repurposing candidates for NSCLC treatment in our study.

This study integrates gene expression profiles and PPI networks to identify prognostic genes and candidate drug molecules. This analysis, however, is limited in that it does not include the expression of these genes in the different NSCLC subtypes. Here, we showed that high expressions of *CDC20*, *AURKA*, *CDK1*, *EZH2*, and *CDKN2A* genes were associated with significantly poorer overall survival, whereas upregulations of *CBL*, *FYN*, *LRKK2*, and *SOCS2* were associated with a significantly better prognosis. Our drug target analysis of hub genes suggests a potential use of Trichostatin A, Pracinostat, TGX-221, PHA-793887, AG-879, and IMD0354 antineoplastic agents to reverse the expression of hub genes in NSCLC patients. While our work showed the benefit and usefulness of gene expression data analysis in bringing out the DEGs and

hub genes that can be potential prognostic and treatment targets of NSCLC, the need for improved analysis and prospective clinical studies is still imperative. Finally, this study can contribute to the overall understanding of the underlying molecular mechanisms of NSCLC and serve as a guide to subsequent experimental studies.

Author Contributions

ÖCE, EG study conception and design. ÖCE data collection. ÖCE, BC analysis and interpretation of results. ÖCE, BC and EG draft manuscript preparation. All authors reviewed the results and approved the final version of the manuscript.

ORCID iDs

Özgür Cem Erkin  <https://orcid.org/0000-0001-9818-3447>
Betül Cömertpay  <https://orcid.org/0000-0002-3515-1461>

Supplemental Material

Supplemental material for this article is available online.

REFERENCES

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68:394-424. doi:10.3322/caac.21492.
- Herbst RS, Heymach JV, Lippman SM. Lung cancer. *N Engl J Med*. 2008;359:1367-1380. doi:10.1056/NEJMra0802714.
- Sibille A, Paulus A, Martin M, et al. Management of non-small cell lung cancer. *Rev Med Liege*. 2015;70:432-441.
- Oxnard GR, Lo PC, Nishino M, et al. Natural history and molecular characteristics of lung cancers harboring EGFR exon 20 insertions. *J Thorac Oncol*. 2013;8:179-184. doi:10.1097/JTO.0b013e3182779d18.
- Janku F, Stewart DJ, Kurzrock R. Targeted therapy in non-small-cell lung cancer—is it becoming a reality? *Nat Rev Clin Oncol*. 2010;7:401-414. doi:10.1038/nrclinonc.2010.64.
- Gerber DE, Minna JD. ALK inhibition for non-small cell lung cancer: from discovery to therapy in record time. *Cancer Cell*. 2010;18:548-551. doi:10.1016/j.ccr.2010.11.033.
- Miller KD, Siegel RL, Lin CC, et al. Cancer treatment and survivorship statistics, 2016. *CA Cancer J Clin*. 2016;66:271-289. doi:10.3322/caac.21349.
- Wistuba I. Genetics of preneoplasia: lessons from lung cancer. *Curr Mol Med*. 2007;7:3-14. doi:10.2174/156652407779940468.
- Xu Z, Zhou Y, Cao Y, Dinh TL, Wan J, Zhao M. Identification of candidate biomarkers and analysis of prognostic values in ovarian cancer by integrated bioinformatics analysis. *Med Oncol*. 2016;33:130. doi:10.1007/s12032-016-0840-y.
- Kori M, Gov E, Arga KY. Novel genomic biomarker candidates for cervical cancer as identified by differential co-expression network analysis. *OMICS*. 2019;23:261-273. doi:10.1089/omi.2019.0025.
- Huang R, Gao L. Identification of potential diagnostic and prognostic biomarkers in non-small cell lung cancer based on microarray data. *Oncol Lett*. 2018;15:6436-6442. doi:10.3892/ol.2018.8153.
- Xiao Y, Feng M, Ran H, Han X, Li X. Identification of key differentially expressed genes associated with non-small cell lung cancer by bioinformatics analyses. *Mol Med Rep*. 2018;17:6379-6386. doi:10.3892/mmr.2018.8726.
- Wang J, Wang N, Li ZJ, et al. Identification of key genes and construction of microRNA-mRNA regulatory networks in non-small cell lung cancer. *Cancer Genet*. 2018;228-229:47-54. doi:10.1016/j.cancergen.2018.08.003.
- Zhang J, Li D, Zhang Y, et al. Integrative analysis of mRNA and miRNA expression profiles reveals seven potential diagnostic biomarkers for non-small cell lung cancer. *Oncol Rep*. 2020;43:99-112. doi:10.3892/or.2019.7407.
- Ioannidis JP, Allison DB, Ball CA, et al. Repeatability of published microarray gene expression analyses. *Nat Genet*. 2009;41:149-155. doi:10.1038/ng.295.
- Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets-update. *Nucleic Acids Res*. 2013;41:D991-D995. doi:10.1093/nar/gks1193.

17. Sanchez-Palencia A, Gomez-Morales M, Gomez-Capilla JA, et al. Gene expression profiling reveals novel biomarkers in non-small cell lung cancer. *Int J Cancer*. 2011;129:355-364. doi:10.1002/ijc.25704.
18. Lu TP, Tsai MH, Lee JM, et al. Identification of a novel biomarker, SEMA5A, for non-small cell lung carcinoma in nonsmoking women. *Cancer Epidemiol Biomarkers Prev*. 2010;19:2590-2597. doi:10.1158/1055-9965.EPI-10-0332.
19. Wei TY, Hsia JY, Chiu SC, et al. Methylosome protein 50 promotes androgen- and estrogen-independent tumorigenesis. *Cell Signal*. 2014;26:2940-2950. doi:10.1016/j.cellsig.2014.09.014.
20. Mitchell KA, Zingone A, Toulabi L, Boeckelman J, Ryan BM. Comparative transcriptome profiling reveals coding and noncoding RNA differences in NSCLC from African Americans and European Americans. *Clin Cancer Res*. 2017;23:7412-7425. doi:10.1158/1078-0432.CCR.
21. Irizarry RA, Hobbs B, Collin F, et al. Exploration, normalization, and summaries of high-density oligonucleotide array probe level data. *Biostatistics*. 2003;4:249-264. doi:10.1093/biostatistics/4.2.249.
22. Smyth GK. limma: linear models for microarray data. In: Gentleman R, Carey VJ, Huber W, Irizarry RA, Dudoit S (eds) *Bioinformatics and Computational Biology Solutions Using R and Bioconductor. Statistics for Biology and Health*. New York: Springer; 2005. doi:10.1007/0-387-29362-0_23.
23. Kamburov A, Stelzl U, Lehrach H, Herwig R. The ConsensusPathDB interaction database: 2013 update. *Nucleic Acids Res*. 2013;41:D793-D800. doi:10.1093/nar/gks1055.
24. Kanehisa M, Sato Y, Furumichi M, Morishima K, Tanabe M. New approach for understanding genome variations in KEGG. *Nucleic Acids Res*. 2019;47:D590-D595. doi:10.1093/nar/gky962.
25. Oughtred R, Stark C, Breitkreutz BJ, et al. The BioGRID interaction database: 2019 update. *Nucleic Acids Res*. 2019;47:D529-D541. doi:10.1093/nar/gky1079.
26. Cline MS, Smoot M, Cerami E, et al. Integration of biological networks and gene expression data using Cytoscape. *Nat Protoc*. 2007;2:2366-2382. doi:10.1038/nprot.2007.324.
27. Györfy B, Lanczky A, Eklund AC, et al. An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1809 patients. *Breast Cancer Res Treat*. 2010;123:725-731. doi:10.1007/s10549-009-0674-9.
28. Duan Q, Reid SP, Clark NR, et al. L1000CDS2: LINCS L1000 characteristic direction signatures search engine. *NPJ Syst Biol Appl*. 2016;2:16015. doi:10.1038/npsba.2016.15.
29. Keenan AB, Jenkins SL, Jagodnik KM, et al. The library of integrated network-based cellular signatures NIH program: system-level cataloging of human cells response to perturbations. *Cell Syst*. 2018;6:13-24. doi:10.1016/j.cels.2017.11.001.
30. Comertpay B, Gov E. Identification of key biomolecules in rheumatoid arthritis through the reconstruction of comprehensive disease-specific biological networks. *Autoimmunity*. 2020;53:156-166. doi:10.1080/08916934.2020.1722107.
31. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. *Nat Med*. 2004;10:789-799. doi:10.1038/nm1087.
32. Saito M, Shiraiishi K, Kunitoh H, Takenoshita S, Yokota J, Kohno T. Gene aberrations for precision medicine against lung adenocarcinoma. *Cancer Sci*. 2016;107:713-720. doi:10.1111/cas.12941.
33. Wei TYW, Juan CC, Hsia JY, et al. Protein arginine methyltransferase 5 is a potential oncoprotein that upregulates G1 cyclins/cyclin-dependent kinases and the phosphoinositide 3-kinase/AKT signaling cascade. *Cancer Sci*. 2012;103:1640-1650. doi:10.1111/j.1349-7006.2012.02367.x.
34. Fang G, Yu H, Kirschner MW. The checkpoint protein MAD2 and the mitotic regulator CDC20 form a ternary complex with the anaphase-promoting complex to control anaphase initiation. *Genes Dev*. 1998;12:1871-1883. doi:10.1101/gad.12.12.1871.
35. Zhou Z, He M, Shah AA, Wan Y. Insights into APC/C: from cellular function to diseases and therapeutics. *Cell Div*. 2016;11:9. doi:10.1186/s13008-016-0021-6.
36. Kidokoro T, Tanikawa C, Furukawa Y, Katagiri T, Nakamura Y, Matsuda K. CDC20, a potential cancer therapeutic target, is negatively regulated by p53. *Oncogene*. 2008;27:1562-1571. doi:10.1038/sj.onc.1210799.
37. Wang L, Qu J, Liang Y, et al. Identification and validation of key genes with prognostic value in non-small-cell lung cancer via integrated bioinformatics analysis. *Thorac Cancer*. 2020;11:851-866. doi:10.1111/1759-7714.13298.
38. Zeng Y, Li N, Chen R, et al. Screening of hub genes associated with prognosis in non-small cell lung cancer by integrated bioinformatics analysis. *Transl Cancer Res*. 2020;9:7149-7164. doi:10.21037/tcr-20-1073.
39. Carvalhal S, Ribeiro SA, Arocena M, et al. The nucleoporin ALADIN regulates Aurora A localization to ensure robust mitotic spindle formation. *Mol Biol Cell*. 2015;26:3424-3438. doi:10.1091/mbc.E15-02.
40. Nikonova AS, Astsaturov I, Serebriiskii IG, Dunbrack RL, Golemis EA. Aurora A kinase (AURKA) in normal and pathological cell growth. *Cell Mol Life Sci*. 2013;70:661-687. doi:10.1007/s00018-012.
41. Selvaraj G, Kaliamurthi S, Kaushik AC, et al. Identification of target gene and prognostic evaluation for lung adenocarcinoma using gene expression meta-analysis, network analysis and neural network algorithms. *J Biomed Inform*. 2018;86:120-134. doi:10.1016/j.jbi.2018.09.004.
42. Zheng X, Chi J, Zhi J, et al. Aurora-A-mediated phosphorylation of LKB1 compromises LKB1/AMPK signaling axis to facilitate NSCLC growth and migration. *Oncogene*. 2018;37:502-511. doi:10.1038/onc.2017.354.
43. Li ZH, Sang MX, Tian ZQ, et al. Identification of key biomarkers and potential molecular mechanisms in lung cancer by bioinformatics analysis. *Oncol Lett*. 2019;18:4429-4440. doi:10.3892/ol.2019.10796.
44. Ye X, Gao Q, Wu J, Zhou L, Tao M. Identification of significant genes in non-small cell lung cancer by bioinformatics analyses. *Transl Cancer Res*. 2020;9:4330-4340. doi:10.21037/tcr-19-2596.
45. Wang Q, Su L, Liu N, Zhang L, Xu W, Fang H. Cyclin dependent kinase 1 inhibitors: a review of recent progress. *Curr Med Chem*. 2011;18:2025-2043. doi:10.2174/092986711795590110.
46. Song JY, Tan J, Gao XH, Wang LX. Integrated analysis reveals key genes with prognostic value in lung adenocarcinoma. *Cancer Manag Res*. 2018;10:6097-6108.
47. Wan Z, Jiang H, Li L, Zhu S, Hou J, Yu Y. Carcinogenic roles and therapeutic effects of EZH2 in gynecological cancers. *Bioorg Med Chem*. 2020;28:115379. doi:10.1016/j.bmc.2020.115379.
48. Hubaux R, Thu KL, Coe BP, MacAulay C, Lam S, Lam WL. EZH2 promotes E2F-driven SCLC tumorigenesis through modulation of apoptosis and cell-cycle regulation. *J Thorac Oncol*. 2013;8:1102-1106. doi:10.1097/JTO.0b013e318298762f.
49. Shi B, Behrens C, Vaghani V, et al. Oncogenic enhancer of zeste homolog 2 is an actionable target in patients with non-small cell lung cancer. *Cancer Med*. 2019;8:6383-6392. doi:10.1002/cam4.1855.
50. Tremblay-LeMay R, Rastgoo N, Pourabdollah M, Chang H. EZH2 as a therapeutic target for multiple myeloma and other haematological malignancies. *Biomark Res*. 2018;6:34. doi:10.1186/s40364-018-0148-5.
51. Feng Q, He H, Gao T, et al. Synthesis and biological evaluation of benzomorpholine derivatives as novel EZH2 inhibitors for anti-non-small cell lung cancer activity. *Mol Divers*. 2019;23:681-696. doi:10.1007/s11030-018-9903-7.
52. Fontana R, Ranieri M, La Mantia G, Vivo M. Dual role of the alternative reading frame ARF protein in cancer. *Biomolecules*. 2019;9:87. doi:10.3390/biom9030087.
53. Naramura M, Nadeau S, Mohapatra B, et al. Mutant Cbl proteins as oncogenic drivers in myeloproliferative disorders. *Oncotarget*. 2011;2:245-250. doi:10.18632/oncotarget.233.
54. Schmidt MH, Dikic I. The Cbl interactome and its functions. *Nat Rev Mol Cell Biol*. 2005;6:907-918. doi:10.1038/nrm1762.
55. Wei TT, Lin YC, Lin PH, et al. Induction of c-Cbl contributes to anti-cancer effects of HDAC inhibitor in lung cancer. *Oncotarget*. 2015;6:12481-12492. doi:10.18632/oncotarget.3489.
56. Saito YD, Jensen AR, Salgia R, Posadas EM. Fyn: a novel molecular target in prostate cancer. *Cancer*. 2010;116:1629-1637. doi:10.1002/cncr.24879.
57. Lewin B, Siu A, Baker C, et al. Expression of Fyn kinase modulates EMT in oral cancer cells. *Anticancer Res*. 2010;30:2591-2596.
58. Sun R, Meng X, Wang W, et al. Five genes may predict metastasis in non-small cell lung cancer using bioinformatics analysis. *Oncol Lett*. 2019;18:1723-1732. doi:10.3892/ol.2019.10498.
59. Gómez-Suaga P, Luzón-Toro B, Churamani D, et al. Leucine-rich repeat kinase 2 regulates autophagy through a calcium-dependent pathway involving NAADP. *Hum Mol Genet*. 2012;21:511-525. doi:10.1093/hmg/ddr481.
60. Paisán-Ruiz C, Nath P, Washecka N, Gibbs JR, Singleton AB. Comprehensive analysis of LRRK2 in publicly available Parkinson's disease cases and neurologically normal controls. *Hum Mutat*. 2008;29:485-490. doi:10.1002/humu.20668.
61. Salašová A, Yokota C, Potěšil D, Zdráhal Z, Bryja V, Arenas E. A proteomic analysis of LRRK2 binding partners reveals interactions with multiple signaling components of the WNT/PCP pathway. *Mol Neurodegener*. 2017;12:54. doi:10.1186/s13024-017.
62. Jiang ZC, Chen XJ, Zhou Q, Gong XH, Chen X, Wu WJ. Downregulated LRRK2 gene expression inhibits proliferation and migration while promoting the apoptosis of thyroid cancer cells by inhibiting activation of the JNK signaling pathway. *Int J Oncol*. 2019;55:21-34. doi:10.3892/ijo.2019.4816.
63. Peng Q, Lin K, Chang T, Zou L. Identification of genomic expression differences between right-sided and left-sided colon cancer based on bioinformatics analysis. *Oncol Targets Ther*. 2018;11:609-618. doi:10.2147/OTT.S154207.
64. Meng F, Zhang L, Ren Y, Ma Q. The genomic alterations of lung adenocarcinoma and lung squamous cell carcinoma can explain the differences of their overall survival rates. *J Cell Physiol*. 2019;234:10918-10925. doi:10.1002/jcp.27917.
65. Letellier E, Haan S. SOCS2: physiological and pathological functions. *Front Biosci (Elite Ed)*. 2016;8:189-204. doi:10.2741/E760.
66. Nguyen CH, Glüxam T, Schlerka A, et al. SOCS2 is part of a highly prognostic 4-gene signature in AML and promotes disease aggressiveness. *Sci Rep*. 2019;9:9139. doi:10.1038/s41598-019.

67. Tong JL, Wang LL, Ling XF, Wang MX, Cao W, Liu YY. MiR-875 can regulate the proliferation and apoptosis of non-small cell lung cancer cells via targeting SOCS2. *Eur Rev Med Pharmacol Sci.* 2019;23:5235-5241. doi:10.26355/eurrev_201906_1818.
68. Wikman H, Kettunen E, Seppänen JK, et al. Identification of differentially expressed genes in pulmonary adenocarcinoma by using cDNA array. *Oncogene.* 2002;21:5804-5813. doi:10.1038/sj.onc.1205726.
69. Manero GG, Atallah E, Khaled SK, et al. A phase 2 study of pracinostat and azacitidine in elderly patients with acute myeloid leukemia (AML) not eligible for induction chemotherapy: response and long-term survival benefit. *Blood.* 2016;128:100. doi:10.1182/blood.V128.22.100.100.
70. Mukhopadhyay NK, Weisberg E, Gilchrist D, Bueno R, Sugarbaker DJ, Jaklitsch MT. Effectiveness of trichostatin A as a potential candidate for anticancer therapy in non-small-cell lung cancer. *Ann Thorac Surg.* 2006;81:1034-1042. doi:10.1016/j.athoracsur.2005.06.059.
71. Tang D, Yao R, Zhao D, et al. Trichostatin A reverses the chemoresistance of lung cancer with high IGFBP2 expression through enhancing autophagy. *Sci Rep.* 2018;8:3917. doi:10.1038/s41598-018.
72. Sun F, Chan E, Wu Z, Yang X, Marquez VE, Yu Q. Combinatorial pharmacologic approaches target EZH2-mediated gene repression in breast cancer cells. *Mol Cancer Ther.* 2009;8:3191-3202. doi:10.1158/1535-7163.MCT.
73. Li L, Fan B, Zhang LH, et al. Trichostatin A potentiates TRAIL-induced anti-tumor effects via inhibition of ERK/FOXM1 pathway in gastric cancer. *Tumour Biol.* 2016;37:10269-10278. doi:10.1007/s13277-016-4816-5.
74. Larsson LI. Novel actions of tyrphostin AG 879: inhibition of RAF-1 and HER-2 expression combined with strong antitumoural effects on breast cancer cells. *Cell Mol Life Sci.* 2004;61:2624-2631. doi:10.1007/s00018-004.
75. Gomez-Cabrero A, Wrasidlo W, Reisfeld RA. IMD-0354 targets breast cancer stem cells: a novel approach for an adjuvant to chemotherapy to prevent multidrug resistance in a murine model. *PLoS ONE.* 2013;8:e73607. doi:10.1371/journal.pone.0073607.
76. Kim S, Ko D, Lee Y, et al. Anti-cancer activity of the novel 2-hydroxydiaryl-amide derivatives IMD-0354 and KRT1853 through suppression of cancer cell invasion, proliferation, and survival mediated by Tmprss4. *Sci Rep.* 2019;9:10003. doi:10.1038/s41598-019.
77. Jackson SP, Schoenwaelder SM, Goncalves I, et al. PI 3-kinase p110beta: a new target for antithrombotic therapy. *Nat Med.* 2005;11:507-514. doi:10.1038/nm1232.
78. Li B, Sun A, Jiang W, Thrasher JB, Terranova P. PI-3 kinase p110β: a therapeutic target in advanced prostate cancers. *Am J Clin Exp Urol.* 2014;2:188-198.
79. Stamatkin C, Ratermann KL, Overley CW, Black EP. Inhibition of class IA PI3K enzymes in non-small cell lung cancer cells uncovers functional compensation among isoforms. *Cancer Biol Ther.* 2015;16:1341-1352. doi:10.1080/15384047.2015.1070986.
80. Massard C, Soria JC, Anthony DA, et al. A first in man, phase I dose-escalation study of PHA-793887, an inhibitor of multiple cyclin-dependent kinases (CDK2, 1 and 4) reveals unexpected hepatotoxicity in patients with solid tumours. *Cell Cycle.* 2011;10:963-970. doi:10.4161/cc.10.6.15075.
81. Jain AS, Prasad A, Pradeep S, et al. Everything old is new again: drug repurposing approach for non-small cell lung cancer targeting MAPK signaling pathway. *Front Oncol.* 2021;11:741326. doi:10.3389/fonc.2021.741326.