Integrative Bioinformatics for Identifying Key Genes and Networks in NSCLC: Insights for Targeted Therapy.

NU302 R&D Project

Sem II, AY 2024-25

Under the supervision of

Dr. Utkarsh Raj

By

Banju Shriya (BT22ABT560)

Anushka Kushwaha (BT22GBT225)

Rasuti (BT22GBT287)

Abstract

Our study aimed to identify the significant DEGs in NSCLC and study their gene ontology to discover potential biomarkers to make prognosis more efficient. We started our study with the extraction of samples from the GEO database using GSE17073, GSE27262, GSE102287, and GSE33532. With the help of the R programming language, we identified the significant DEGs and later used the Funrich software for extracting overlapping DEGs. We performed the functional enrichment analysis for the top 15 DEGs that were most significant. The key finding from our research was that we identified a total of 476 overlapping DEGs, out of which we chose the top 15 (ZBTB18, KDM6A, CALD1, IRF6, THBS1, IQGAP1, TAX1BP3, TMEM165, NASP, STX6, TBL1XR1, IRS1, SLF2, PHLPP1, GLTP). Our research concludes that we will be identifying the significant DEGs in NSCLC for which we will be performing the functional enrichment analysis and study the protein-protein interaction, which will reveal the biological processes and pathways associated with the identified genes. This will provide us with deeper insights about the underlying molecular mechanisms.

Keywords
NSCLC (Non-Small Cell Lung Cancer)
DEGs (Differentially Expressed Genes)
GO (Gene Ontology)
PPI (Protein-Protein Interaction)
KEGG (Kyoto Encyclopedia of Genes and Genomes)
GEO (Gene Expression Omnibus)

Table of contents

Contents

Introduction

NSCLC is the most common and deadliest type of lung cancer, comprising 85% of the lung cancer cases worldwide. In comparison with SCLC, it shows slower growth and metastasis. Adenocarcinoma, large cell carcinoma, and squamous cell carcinoma are said to be the three foremost kinds of NSCLC. About 81,000 adult males and 30,000 females will develop lung cancer by 2025, mainly NSCLC. There are various unanswered questions about NSCLC, mainly related to early prognosis, treatment, and its impact on our immune system. Drug resistance and delayed diagnosis are the main causes of NSCLC's poor survival rate. To comprehend tumor growth and drug interactions and provide more tailored therapy, our research attempts to uncover and validate important DEGs. The specific aims and goals of our research is identifying the significant DEGs in NSCLC using R Programming, identifying overlapping DEGs using Funrich Software, performing functional Enrichment analysis using Enrichr and DAVID and study gene annotation using BioGPS. Our research is significant, as it will help us in identifying the potential biomarkers and therapeutic targets and help in early prognosis for NSCLC. Our research has several limitations. Firstly, we were unable to identify the uncommon DEGs that have not been studied earlier. The samples that we have chosen are from different tissue samples, hence the results are not that significant, and we did not perform wet lab experiments for our study, which can hamper our findings.

Literature review

Introduction

Currently, non-small cell lung cancer is about 85% of the total lung cancer cases in the world. Bioinformatics techniques are used to identify biomarkers, molecular pathways, and potential therapeutic targets to overcome EGFR-TKI resistance. In this review, recent research efforts have been grouped critically and thematically to address this challenge. Non-small cell lung cancer (NSCLC) biomarker and differentially expressed gene (DEG) identification.

Identification of Differentially Expressed Genes (DEGs) and Biomarkers in NSCLC

In recent years, many studies have shown the use of a variety of bioinformatics tools for the identification of DEGs and biomarkers that are related to drug resistance and progression of NSCLC. Zhu et al. (2023) used patient-derived xenotransplantation models (PDX) and two datasets - GSE64472 and GSE130160, and identified 1302 DEGs and 10 hub genes that were linked to EGFR-TKI resistance. Only ITGAM, however, was suggested as a possible biomarker, and its low disease-free survival rates underscored the necessity for additional experimental research to confirm the findings. Sultana et al. used single-cell RNA sequencing (scRNA-seq) from the GSE127471 dataset in 2022 and discovered 12 important genes as novel biomarkers, including RPLP1 and TYROBP. Their research focused on integrated miRNA and transcription factor (TF) regulatory network analyses and the importance of tumor heterogeneity in NSCLC. The study's limitations included using only one dataset and lacking wet-lab validation despite their success in finding putative biomarkers. Xiao & Co. (2018) also investigated DEGs with four microarray datasets, finding 25 hub genes and 195 DEGs. Although their integrative approach improved the findings robustness, the emphasis on overlapping DEGs ran the risk of leaving out important genes specific to a given subtype, and there was no differentiation between NSCLC subtypes, which might have limited the findings' clinical relevance. Novel Therapeutic Approaches for Overcoming EGFR-TKI Resistance

Novel Therapeutic Approaches for Overcoming EGFR-TKI Resistance

In addition to finding biomarkers, the researchers are also approaching other treatment options. By using molecular docking, in vitro tests, and structure-activity relationship (SAR) analysis, He et al. (2022) mainly concentrated on pyrrole-based EGFR inhibitors. In comparison to the first-generation inhibitors, pyrrole-based compounds showed higher binding affinity to EGFR, indicating intriguing therapeutic potential. We still need clinical trials to evaluate any possible off-target effects and establish their effectiveness.

Li et al. (2018) investigated by a different method, by combining ongoing microwave ablation with EGFR-TKI therapy. Through their research, it was determined that a small sample of 15 people had a longer universal existence and development, free survival. Notwithstanding such true results, the findings' potential to generalize changed into restricted due to the sample size and selection bias. In order to track changing resistance mechanisms, the study also underlined the importance of routine re-biopsies.

Integrated Bioinformatics for Target Identification

Altaf et al. employed integrated bioinformatics analyses for a number of datasets. Ten important DEGs for non-small cell lung cancer were found in 2023; these included ID2, GJA4 and DOCK4. Functional enrichment mutation profiling and protein-protein interactions (PPI) analysis were conducted using a variety of tools, including STRING, DAVID, and ActiveDriverDB. Despite the identification of candidate genes in their study they have recognized the need for additional experimental confirmation and a deeper understanding of the underlying molecular mechanisms.

Critical Synthesis and Research Gaps

There are a number of recurring issues and gaps in these studies. NSCLC tumor heterogeneity makes it more difficult to identify biomarkers and customize treatment. Without adequate in vitro or in vivo experimental validation the majority of studies mainly relied on bioinformatics predictions. Results robustness and applicability are further constrained by single-dataset analyses and small sample sizes. Additionally there is not enough integration of multi-omics (e. g. G. Proteomics transcriptomics and genomics) that may offer a more thorough comprehension of drug resistance mechanisms. Finally in the absence of thorough validation and clinical trials the clinical translation of innovative therapeutic approaches is still unknown.

Research Contribution

In this study key molecular pathways, differentially expressed genes and possible therapeutic targets in non-small cell lung cancer (NSCLC) were identified by applying integrated bioinformatics approaches to publicly available transcriptomic data (GSE33532, GSE17073, GSE19188, GSE4495) . Functional enrichment analysis, the creation of protein-protein interaction networks and hub gene identification were used for computational validation. It is suggested that future research will include more experimental validation.

Conclusion

The reviewed literature highlights the significant progress made in information EGFR-TKI resistance in NSCLC, while also underscoring persistent demanding situations which include tumor heterogeneity, lack of validation and constrained medical translation. By addressing these gaps, future studies, along with the prevailing observation, we can contribute to more effective, personalized therapeutic strategies for NSCLC sufferers.

Methodology

A systematic quantitative bioinformatic analysis was carried out in order to identify and characterize significant differentially expressed genes (DEGs) linked to non-small cell lung cancer (NSCLC). Human NSCLC microarray studies of superior quality were obtained from the NCBI Gene Expression Omnibus (GEO) database. The Affymetrix Human Genome U133 Plus 2. 0 array consistent platform annotation, the availability of raw CEL files and a sufficient sample size (≥20 samples) were the criteria used to select the datasets GSE17073, GSE102287, GSE27262 and GSE33532. Every raw data set was downloaded and processed locally. Utilizing the affyQCReport package initial quality control was carried out in R (v4. 1) to identify outlier arrays and evaluate probe performance in order to guarantee data integrity.

Using the affy and oligo packages the Robust Multiarray Average (RMA) algorithm was used to perform background correction quantile normalization and log2 transformation as subsequent preprocessing steps. The probe-level data were collapsed to distinct gene symbols by averaging probes mapping to the same gene and genes with consistently low variance (bottom 20th percentile) were filtered out to minimize noise. The limma package was used to perform differential expression analysis independently for each dataset fitting linear models to each gene and using empirical Bayes moderation to increase statistical power. It was specified to compare tumor samples to nearby normal tissues using a two-group design matrix. Between 1200 and 2500 DEGs per dataset were found by applying strict criteria which included an absolute log2 fold-change greater than or equal to 1 and an adjusted p-value (Benjamini–Hochberg FDR) less than 0. 05. Next FunRich (v3. 1) was used to import the DEG lists from the four datasets. 19 common DEGs across all studies are highlighted in a Venn diagram created by identifying overlapping genes.

Top 15 genes were chosen for additional examination based on the overlapping DEGs (the average log2 fold-change magnitude). Additional tools like DAVID (v6. 8) were used to cross-check the cellular components, biological process and molecular function annotations and to find the significant pathway associations by using a 0.5 threshold value for Benjamini–Hochberg-corrected p-value. The Enrichr tool was used to analyse the Gene Ontology (GO) terms and KEGG pathways. The BioGPS tool was used to study individual gene annotation for interpretation of the top 15 genes.

Results

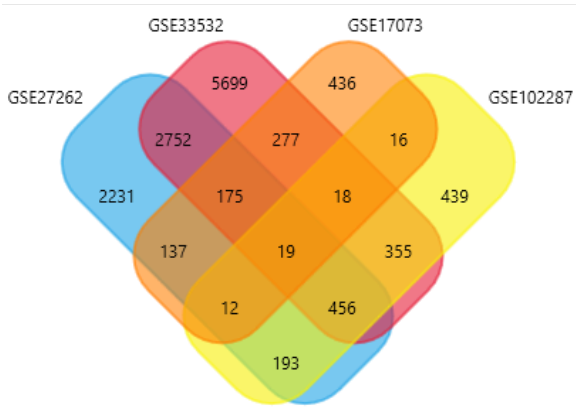


Fig.1 Out of the 19 overlapping DEGs that were identified, we took the top 15 DEGs (ZBTB18, KDM6A, CALD1, IRF6, THBS1, IQGAP1, TAX1BP3, TMEM165, NASP, STX6, TBL1XR1, IRS1, SLF2, PHLPP1, GLTP, TLK2, NAMPT, SEMA3F, EHF).

KEGG 2021 Human Bar Graph | **Table** | Clustergram | Appyter

Hover each row to see the overlapping genes.

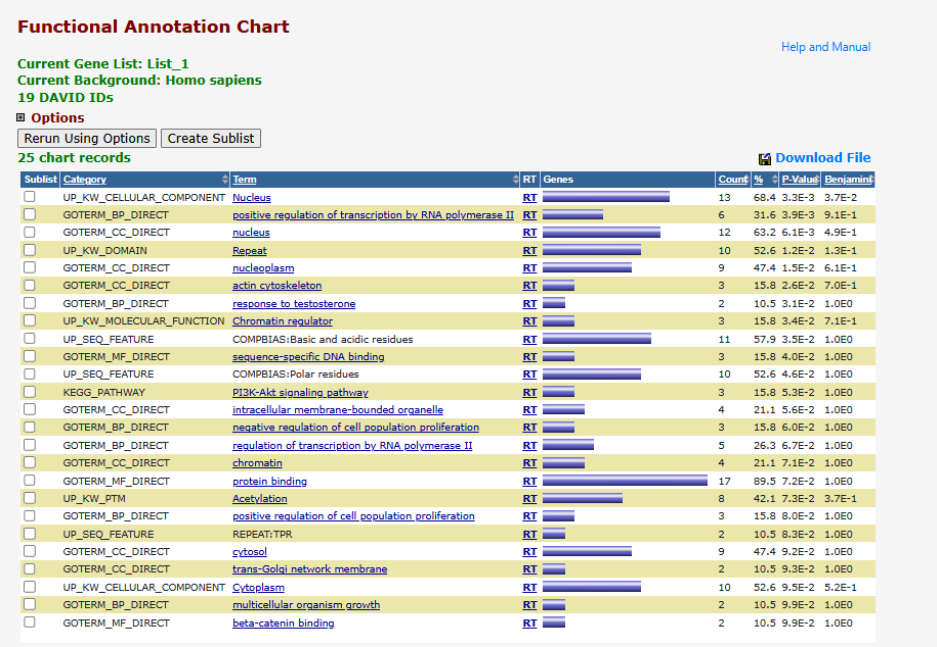
10 entries per page Search:

Index	Name	P-value	Adjusted p-value	Odds Ratio	Combined score
1	PI3K-Akt signaling pathway	0.004316	0.1640	10.49	57.10
2	Proteoglycans in cancer	0.01594	0.1931	11.46	47.44
3	SNARE interactions in vesicular transport	0.03090	0.1931	34.63	120.42
4	Nicotinate and nicotinamide metabolism	0.03275	0.1931	32.59	111.44
5	MicroRNAs in cancer	0.03442	0.1931	7.51	25.32
6	Aldosterone-regulated sodium reabsorption	0.03459	0.1931	30.78	103.55
7	Bladder cancer	0.03826	0.1931	27.70	90.38
8	Type II diabetes mellitus	0.04283	0.1931	24.61	77.54
9	Malaria	0.04647	0.1931	22.60	69.36
10	Regulation of lipolysis in adipocytes	0.05100	0.1931	20.50	61.01

Showing 1 to 10 of 38 entries | Export entries to table Previous Next

Terms marked with an * have an overlap of less than 5

Fig. 2 This table from KEGG pathway enrichment analysis lists the top biological pathways which are related to our gene set. It includes the pathway's name, its statistical significance (p-value and adj p-value), It also explains the strength of the genes in that pathway.



Functional Annotation Chart

Help and Manual

Current Gene List: List_1
Current Background: Homo sapiens
19 DAVID IDs

Options

Rerun Using Options | Create Sublist

25 chart records

Download File

Sublist	Category	Term	RT	Genes	Count	%	P-Value	Benjamini
	UP_KW_CELLULAR_COMPONENT	Nucleus	RT		13	68.4	3.3E-3	3.7E-2
	GOTERM_BP_DIRECT	positive regulation of transcription by RNA polymerase II	RT		6	31.6	3.9E-3	9.1E-1
	GOTERM_CC_DIRECT	nucleus	RT		12	63.2	6.1E-3	4.9E-1
	UP_KW_DOMAIN	Repeat	RT		10	52.6	1.2E-2	1.3E-1
	GOTERM_CC_DIRECT	nucleoplasm	RT		9	47.4	1.5E-2	6.1E-1
	GOTERM_CC_DIRECT	actin cytoskeleton	RT		3	15.8	2.6E-2	7.0E-1
	GOTERM_BP_DIRECT	response to testosterone	RT		2	10.5	3.1E-2	1.0E0
	UP_KW_MOLECULAR_FUNCTION	Chromatin regulator	RT		3	15.8	3.4E-2	7.1E-1
	UP_SEQ_FEATURE	COMPBIAS:Basic and acidic residues	RT		11	57.9	3.5E-2	1.0E0
	GOTERM_MF_DIRECT	sequence-specific DNA binding	RT		3	15.8	4.0E-2	1.0E0
	UP_SEQ_FEATURE	COMPBIAS:Polar residues	RT		10	52.6	4.6E-2	1.0E0
	KEGG_PATHWAY	PI3K-Akt signaling pathway	RT		3	15.8	5.3E-2	1.0E0
	GOTERM_CC_DIRECT	intracellular membrane-bounded organelle	RT		4	21.1	5.6E-2	1.0E0
	GOTERM_BP_DIRECT	negative regulation of cell population proliferation	RT		3	15.8	6.0E-2	1.0E0
	GOTERM_BP_DIRECT	regulation of transcription by RNA polymerase II	RT		5	26.3	6.7E-2	1.0E0
	GOTERM_CC_DIRECT	chromatin	RT		4	21.1	7.1E-2	1.0E0
	GOTERM_MF_DIRECT	protein binding	RT		17	89.5	7.2E-2	1.0E0
	UP_KW_PTM	Acetylation	RT		8	42.1	7.3E-2	3.7E-1
	GOTERM_BP_DIRECT	positive regulation of cell population proliferation	RT		3	15.8	8.0E-2	1.0E0
	UP_SEQ_FEATURE	REPEAT:TPR	RT		2	10.5	8.3E-2	1.0E0
	GOTERM_CC_DIRECT	cytosol	RT		9	47.4	9.2E-2	1.0E0
	GOTERM_CC_DIRECT	trans-Golgi network membrane	RT		2	10.5	9.3E-2	1.0E0
	UP_KW_CELLULAR_COMPONENT	Cytoplasm	RT		10	52.6	9.5E-2	5.2E-1
	GOTERM_BP_DIRECT	multicellular organism growth	RT		2	10.5	9.9E-2	1.0E0
	GOTERM_MF_DIRECT	beta-catenin binding	RT		2	10.5	9.9E-2	1.0E0

Fig.3 Functional Annotation Chart in DAVID. Functional annotation includes gene Ontology(GO), biological processes, molecular function and cellular components categories. Analysis was based on Homo sapiens set in the background. The horizontal bars show the ratio of genes in the input list associated with each term, along with the p-values showing its significance in enrichment.

Functional Annotation Table

Help and Manual

Current Gene List: List_1
Current Background: Homo sapiens
19 DAVID IDs

19 record(s) Download File

EHF	ETS homologous factor(EHF)	Related Genes	Homo sapiens
GOTERM_BP_DIRECT	regulation of transcription by RNA polymerase II, cell differentiation, epithelial cell differentiation, positive regulation of transcription by RNA polymerase II,		
GOTERM_CC_DIRECT	chromatin, nucleus, nucleoplasm, Golgi apparatus,		
GOTERM_MF_DIRECT	RNA polymerase II cis-regulatory region sequence-specific DNA binding, DNA-binding transcription factor activity, RNA polymerase II-specific, DNA-binding transcription activator activity, RNA polymerase II-specific, DNA binding, protein binding, sequence-specific DNA binding,		
INTERPRO	Ets_dom, Pointed_dom, SAM/pointed_sf, EHF_SAM_Pointed_dom, WH-like_DNA-bd_sf, WH_DNA-bd_sf, ETS_fam,		
SMART	SAM_PNT, ETS,		
UP_KW_BIOLOGICAL_PROCESS	Transcription, Transcription regulation,		
UP_KW_CELLULAR_COMPONENT	Nucleus,		
UP_KW_MOLECULAR_FUNCTION	DNA-binding,		
UP_SEQ_FEATURE	COMPBIAS:Basic and acidic residues, DNA_BIND:ETS, DOMAIN:PNT, REGION:Disordered,		
IQGAP1	**IQ motif containing GTPase activating protein 1(IQGAP1)**	**Related Genes**	**Homo sapiens**
COG_ONTOLOGY	Cell division and chromosome partitioning / Signal transduction mechanisms,		
GOTERM_BP_DIRECT	MAPK cascade, regulation of cytokine production, signal transduction, epidermal growth factor receptor signaling pathway, regulation of mitotic cell cycle, fibroblast growth factor receptor signaling pathway, fibroblast migration, cell migration, negative regulation of dephosphorylation, cellular response to platelet-derived growth factor stimulus, positive regulation of MAPK cascade, positive regulation of protein kinase activity, platelet-derived growth factor receptor signaling pathway, caveola assembly, cellular response to calcium ion, cellular response to epidermal growth factor stimulus, podocyte development, mitotic actomyosin contractile ring assembly actin filament organization, neuron projection extension,		
GOTERM_CC_DIRECT	ruffle, nucleus, cytoplasm, cytosol, microtubule, actin filament, plasma membrane, focal adhesion, cell cortex, cytoplasmic side of plasma membrane, actin cytoskeleton, microtubule cytoskeleton, basolateral plasma membrane, apical plasma membrane, lateral plasma membrane, axon, growth cone, midbody, secretory granule membrane, cortical actin cytoskeleton, slit diaphragm, cytoplasmic ribonucleoprotein granule, neuron projection, extracellular exosome, plasma membrane bounded cell projection, ribonucleoprotein complex,		
GOTERM_MF_DIRECT	MAP-kinase scaffold activity, GTPase inhibitor activity, GTPase activator activity, calcium ion binding, protein binding, calmodulin binding, phosphatidylinositol-3,4,5-trisphosphate binding, protein kinase binding, protein phosphatase binding, protein domain specific binding, small GTPase binding, protein serine/threonine kinase activator activity, S100 protein binding, cadherin binding, actin filament binding, molecular adaptor activity,		
INTERPRO	IQ_motif_EF-hand-BS, RasGAP_C, WW_dom, CH_dom, RasGAP_dom, Rho_GTPase_activation_prot, RasGAP_CS, P-loop_NTPase, CH_dom_sf,		
KEGG_PATHWAY	Adherens junction, Regulation of actin cytoskeleton, Proteoglycans in cancer,		
SMART	IQ, CH, RasGAP, WW,		
UP_KW_BIOLOGICAL_PROCESS	Host-virus interaction,		
UP_KW_CELLULAR_COMPONENT	Membrane, Nucleus, Cytoplasm, Cell membrane,		
UP_KW_DOMAIN	Coiled coil, Repeat, Signal,		
UP_KW_MOLECULAR_FUNCTION	Calmodulin-binding,		
UP_KW_PTM	Acetylation, Phosphoprotein,		
UP_SEQ_FEATURE	COMPBIAS:Basic and acidic residues, DOMAIN:Calponin-homology (CH), DOMAIN:IQ 1, DOMAIN:IQ 2, DOMAIN:IQ 3, DOMAIN:IQ 4, DOMAIN:Ras-GAP, DOMAIN:WW, MUTAGEN:S->A: Abolishes neurite outgrowth promoting activity; when associated with A-1441., MUTAGEN:S->A: Abolishes neurite outgrowth promoting activity; when associated with A-1443., MUTAGEN:S->D: Strongly enhances neurite outgrowth promoting activity; when associated with A-1441., MUTAGEN:S->E: Strongly enhances neurite outgrowth promoting activity; when associated with A-1443., REGION:C1, REGION:C2, REGION:Disordered,		

Fig.4 Functional Annotation Table in DAVID.This was generated using DAVID for the input gene list (Homo sapiens background). The table shows enriched Gene Ontology terms, protein domains. KEGG pathways, sequence features and gene annotations with associated genes.

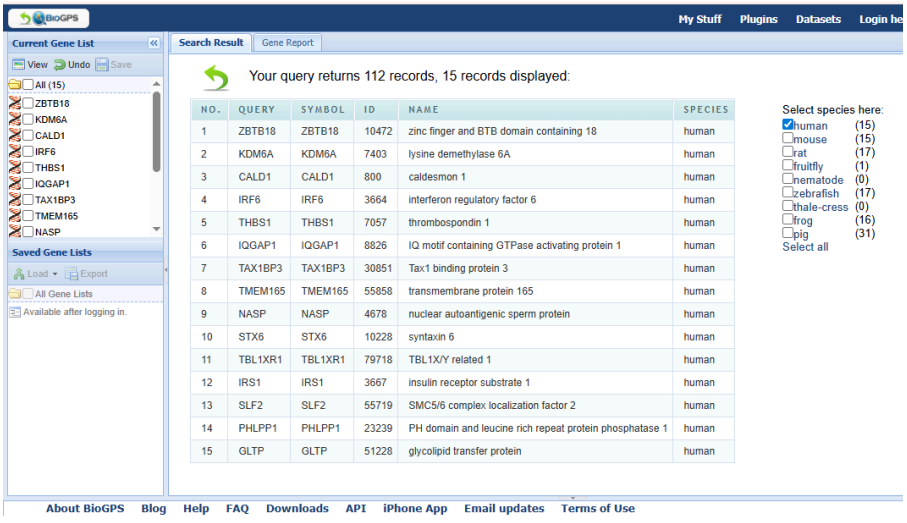


Fig.5 BioGPS gene annotation results. This summarizes information about the 15 genes from our query - its unique ID, query name, full gene name and the official gene symbol - thus describing its function, species, etc.

Discussion

We integrated four NSCLC microarray datasets i.e. GSE17073, GSE27262, GSE102287, GSE33532 and identified 19 consistent dysregulated genes, of which we selected the following top 15 genes: ZBTB18, KDM6A, CALD1, IRF6, THBS1, IQGAP1, TAX1BP3, TMEM165, NASP, STX6, TBL1XR1, IRS1, SLF2, PHLPP1 and GLTP, along with recurrent candidates TLK2, NAMPT, SEMA3F, and EHF (Fig1).

The enrichment analysis highlighted the role of these genes in cytokine signalling, cell adhesion and ECM modelling and also in the epigenetic modulation. THBS1 and IRS1 were clinically relevant genes which were linked to tumor progression. ZBTB18 and SLF2 were less studied but they could be novel biomarkers (Fig.3, Fig.4).

These DEGs show promising results for blood or tissue related diagnostics and targeted therapies (Fig.5). By inhibiting THBS1/SEMA3F it could result in blocking angiogenesis, targeting IQGAP1/IRS1 signaling and exploiting NAMPT's metabolic role.

Limitations of our study include heterogeneity of our datasets, Funrich thresholds and the absence of experimental validation for our results. To improve the precision for our results in the future, we will perform qRT-PCR, western blot and IHC and investigate the mechanisms for 3D models.

Conclusion

In our study, we have identified DEGs by conducting comprehensive computational analysis for multiple GSE datasets. By using R programming, FunRich, Enrichr and DAVID we identified overlapping DEGs successfully and revealed their biological processes, molecular functions, cellular components and pathways.

The findings from our study show how bioinformatics techniques may be used to extract valuable biological insights from the publically accessible data. Our study lays the groundwork for further future experimental validation and therapeutic exploration by identifying important genes and pathways.

Overall our study was focused on integrating multiple datasets and various analytical platforms for reliable gene discovery and functional annotation which will lead to advancing our knowledge in genomics and systems biology.

References

Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Čech, M., & Goecks, J. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Research, 46*(W1), W537–W544. https://doi.org/10.1093/nar/gky379

Altaf, R., Ilyas, U., Ma, A., & Shi, M. (2023). Identification and validation of differentially expressed genes for targeted therapy in NSCLC using integrated bioinformatics analysis. *Frontiers in Oncology, 13*, 1206768. https://doi.org/10.3389/fonc.2023.1206768

Identification and validation of differentially expressed genes for targeted therapy in NSCLC using integrated bioinformatics analysis. (2023, May 31). *Frontiers in Oncology*. Retrieved February 21, 2025, from https://www.frontiersin.org/journals/oncology/articles/10.3389/fonc.2023.1206768/full

Identification of biomarkers, pathways, and therapeutic targets for EGFR-TKI resistance in NSCLC. (2023, October 10). *PubMed NCBI*. https://pubmed.ncbi.nlm.nih.gov/37816585/

Identification of key differentially expressed genes associated with non‐small cell lung cancer by bioinformatics analyses. (2018, May). *PubMed NCBI*. https://pubmed.ncbi.nlm.nih.gov/29532892/

Identification of key differentially expressed genes associated with non-small cell lung cancer by bioinformatics analyses. (2018, March 9). *PubMed Central (PMC)*. https://pmc.ncbi.nlm.nih.gov/articles/PMC5928621/

Non-small cell lung cancer. (2025, January 16). *Cleveland Clinic*. Retrieved February 21, 2025, from https://my.clevelandclinic.org/health/diseases/6203-non-small-cell-lung-cancer#outlook-prognosis

Non-small cell lung cancer treatment (PDQ®)–Patient version. (2025, February 12). *National Cancer Institute*. Retrieved February 21, 2025, from https://www.cancer.gov/types/lung/hp/non-small-cell-lung-treatment-pdq#:~:text=NSCLC%20is%20any%20type%20of,occur%20in%20unusual%20histological%20variants

Pathan, M., Keerthikumar, S., Ang, C. S., Gangoda, L., Quek, C. Y., Williamson, N. A., ... & Mathivanan, S. (2015). FunRich: An open access standalone functional enrichment and interaction network analysis tool. *Proteomics, 15*(15), 2597–2601. https://doi.org/10.1002/pmic.201400515

Pyrrole-based EGFR inhibitors for the treatment of NSCLC: Binding modes and SARs investigations. (2023, January). *PubMed NCBI*. https://pubmed.ncbi.nlm.nih.gov/36394145/

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., & Morris, J. H. (2019). STRING v11: Protein–protein association networks with increased coverage, supporting

functional discovery in genome-wide experimental datasets. *Nucleic Acids Research, 47*(D1), D607–D613. https://doi.org/10.1093/nar/gky1131

Uniqueness of lung cancer in Southeast Asia. (2024, August). *The Lancet Regional Health – Southeast Asia*. Retrieved February 21, 2025, from https://www.thelancet.com/journals/lansea/article/PIIS2772-3682(24)00080-5/fulltext

What steps can you take to lower your risk of lung cancer? (2021, March 12). *Healthline*. Retrieved February 21, 2025, from https://www.healthline.com/health/lung-cancer/lung-cancer-prevention

Elderly patients with advanced NSCLC: The value of geriatric evaluation and the feasibility of CGA alternatives in predicting chemotherapy toxicity. (2019, February). *Translational Lung Cancer Research*, *8*(1), 44–57. Retrieved February 21, 2025, from https://www.sciencedirect.com/science/article/pii/S2531043718301132