**Network-based Identification of Key Proteins and Repositioning of Drugs for Non-Small Cell Lung Cancer**

Oluwatosin Maryam Adeyemo a,e, Zainab Ashimiyu-Abdusalamb,e, Mary Adewunmic,e, Temitope Ayanfunke Ayanod,e, Muhammad Sohaibe, Reem Abdel-Salamf,e

1. Department of Biochemistry, Federal University of Technology, Akure, Nigeria.
2. Department of Biochemistry and Nutrition, Nigeria Institute of Medical Research, Lagos, Nigeria.
3. College of Health and Medicine, University of Tasmania, Hobart, Tasmania
4. Department of Microbiology, Obafemi Awolowo University, Ile-Ife, Nigeria
5. Cancer Research with AI (CARESAi)
6. Department of Computer Engineering, Faculty of Engineering, Cairo University, Cairo, Egypt

[omadeyemo@gmail.com](mailto:omadeyemo@gmail.com) (Corresponding author)

[ziabdusalam@gmail.com](mailto:ziabdusalam@gmail.com)

[mary.adewunmi@utas.edu.au](mailto:mary.adewunmi@utas.edu.au) (Corresponding author)

[ayanotemitope@gmail.com](mailto:ayanotemitope@gmail.com)

[muhammad.sohaib@tju.edu.cn](mailto:muhammad.sohaib@tju.edu.cn)

[reem.abdelsalam13@gmail.com](mailto:reem.abdelsalam13@gmail.com)

**Source of funding**

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

**Conflict of interest**

The authors have no conflict of interest to declare.

**Ethical approval**

There are no ethical or financial issues, conflicts of interest, or animal experiments related to this research.

**Authors’ contributions**

OMA conceived and designed the study, wrote the methodology section of the manuscript. OMA and ZAA conducted the research, collected and analyzed the data. TAA and MS wrote the discussion of the research and MS did initial formatting of the manuscript. RA wrote the conclusion. MA gave a critical revision of the research and wrote the introduction. OMA and MA edited the manuscript and reviewed the final version of the manuscript submitted for publication. All authors provided logistical support and approved the final draft, and are accountable for the manuscript’s content and similarity index.

**ABSTRACT**

NSCLC is a lethal cancer that is highly prevalent and accounts for 85% of cases of lung cancer. Conventional cancer treatments, such as chemotherapy and radiation, frequently exhibit limited efficacy and notable adverse reactions. Therefore, a drug repurposing method is proposed for effective NSCLC treatment. This study aims to evaluate candidate drugs that are effective for NSCLC at the clinical level using systems biology and network analysis approach. Differentially expressed genes of transcriptomics data were identified using the systems biology and network analysis approach. A network of gene co-expression was developed with the aim of detecting two modules of gene co-expression. Subsequently, the Drug-Gene interaction database was employed to pinpoint potential pharmaceutical agents that target crucial genes within two gene co-expression modules associated with non-small cell lung cancer (NSCLC). The construction of a drug-gene interaction network was facilitated with the utilisation of Cytoscape. Finally, the gene set enrichment analysis was done to validate candidate drugs. Unlike previous research on repositioning drugs for NSCLC, which uses a gene co-expression network, this project is the first to research both gene co-expression and co-occurrence networks. And the co-occurrence network also accounts for differentially expressed genes in cancer cells and their adjacent normal cells. Drugs exhibiting elevated gene regulation and gene affinity within the drug-gene interaction network are deemed noteworthy for the efficacious management of non-small cell lung cancer (NSCLC). According to this discourse, NSCLC genes exert a high degree of regulation over medications such as vincristine, fluorouracil, methotrexate, clotrimazole, etoposide, tamoxifen, sorafenib, doxorubicin, and pazopanib. Hence, there is a possibility of repurposing these drugs for the treatment of non-small cell lung cancer.

Key words: non-small cell lung cancer (NSCLC), drug repurposing, network analysis, drug-gene interaction, therapeutics

**Network-based Identification of Key Proteins and Repositioning of Drugs for Non-Small Cell Lung Cancer**

According to GLOBOCAN 2020, lung cancer has the highest mortality rate (Siegel, Miller et al. 2022) and is responsible for 18% of the deaths associated with cancer worldwide (Sung, Ferlay et al. 2021). Nearly 84% of lung cancer cases are non-small cell lung carcinoma (NSCLC) and 15% are small cell lung carcinoma (SCLC) (Zappa and Mousa 2016). NSCLC is categorized into three sub-types: adenocarcinoma, squamous cell carcinoma, and large cell carcinoma. Approximately, 45% of non-small cell lung cancers are adenocarcinomas, 25–30% are squamous cell carcinomas, and 5–10% are large cell carcinomas (Zappa and Mousa 2016). The poor survival rate of lung cancer patients at the metastatic stage is primarily attributable to the late diagnosis of disease at stages III and IV (Crino, Weder et al. 2010). Approximately 92% of patients diagnosed at stage IA1 could live for at least 5 years(Kalinke, Thakrar et al. 2021) in contrast to the percentage of individuals diagnosed at stage IV is 10%, which is comparatively lower. Furthermore, a marginal escalation in the dimensions of the tumour from 1 cm (stage IA1) to greater than 2 cm (stage IA3) may result in a decline in the 5-year survival rates of patients from 92% to 77%. (Kalinke, Thakrar et al. 2021). Surgery, chemotherapy, and radiotherapy for non-small cell lung cancer are not associated with reduced mortality rates (Rajasegaran, How et al. 2023, Wang, Zeng et al. 2023). These methods lack precision and are typically constrained by low drug bioavailability because of high first-pass metabolism. In addition, serious adverse effects are caused by the non-specificity of chemotherapeutics, which negatively affect healthy cells (Dongsar, Dongsar et al. 2023). Individualised treatment is preferred to enhance the survival rate of patients with non-small cell lung cancer (NSCLC). Molecularly targeted therapies have been developed to target specific molecular pathways involved in cancer progression. One such therapy is the epidermal growth factor receptor tyrosine kinase inhibitor (EGFR-TKI), which has shown promise in inhibiting the growth and proliferation of lung tumours that carry EGFR mutations. LC, being a disease with heterogeneous characteristics, exhibits intricate molecular mechanisms that lead to uncontrolled cell proliferation. These mechanisms may arise from aberrant gene expression, promoter methylation, and possibly other factors. mutations in tumour suppressor genes and oncogenes(Singh, Kumar et al. 2022).

Despite numerous papers on LC in patients with different classifications (Schwendenwein et al., 2021; Wang et al., 2022; Hu et al., 2021), no collection-based study has examined NSCLC subtype genetics. Most functional or clinical studies focus on one gene and fail to address carcinogenesis for all cancer subtypes. This experiment examines gene co-expression and co-occurrence networks and contrasts earlier studies on NSCLC drug repositioning, which utilised a gene co-expression network. (MotieGhader et al., 2022). In addition, the co-occurrence network considers the differential expression of genes between cancer cells and their adjacent normal cells. This study aims to evaluate candidate drugs that are effective for NSCLC at the clinical level using systems biology and network analysis approach.

1. **Methodology**
   1. **Dataset and preprocessing**

The gene expression data from the NCBI Gene Expression Omnibus (GEO) database was utilised, with reference to the corresponding accession numbers. GSE27262 (Wei et al., 2012, 2014) and GSE21933 (Lo et al., 2012) to compare Non-small Cell Lung Cancer (NSCLC) and normal cell transcriptomes. We used GEO2R to find differentially expressed genes (DEGs) between these two groups (Barrett et al., 2013) with a significance threshold of P-value < 0.05 and |Log2Fold| > 2.0. Venn diagram (<http://bioinformatics.psb.ugent.be/webtools/Venn/>) was used to visualize the intercepting genes in the two datasets. This rigorous selection process allowed us to focus on the most robust and biologically relevant genes for subsequent analyses.

* 1. **Network Construction**

We used publicly accessible human interactome data from the STRING database to uncover potential interactions among differentially expressed genes (DEGs) to better understand Non-small Cell Lung Cancer (NSCLC) processes (Szklarczyk et al., 2021). A comprehensive human protein interactome network was constructed, followed by the construction of a gene co-expression and co-occurrence network using a minimum interaction score of 0.04. The co-expression analysis in STRING uses a method known as FAVA (Functional Associations using Variational Autoencoders) (Szklarczyk et al.,2023). This network was then transferred to the Cytoscape software (Shannon et al., 2003) for further analysis. To identify densely connected regions in the network, we employed the Molecular Complex Detection (MCODE) algorithm (Bader & Hogue, 2003), a graph-theoretic clustering approach made to find areas in a network that are strongly related. Based on the cluster's and the neighborhood's densities, this algorithm locates seed nodes and grows them (Bader & Hogue, 2003). We utilized the MCODE algorithm with the following parameters: Degree threshold = 2, Node Score Threshold = 0.2, K-Core Threshold = 2, and Maxdepth=100. Through the application of the MCODE algorithm, we identified two distinct gene co-expression and co-occurrence modules within the network. These modules represent biologically relevant regions within the network that may have functional significance in NSCLC tumorigenesis. Our use of the MCODE algorithm enabled us to identify these modules in a rigorous and comprehensive manner, providing us with deeper insights into the complex interplay between differentially expressed genes and potential molecular pathways involved in NSCLC.

* 1. **Enrichment Analysis**

To gain a better understanding of the potential biological functions and pathways associated with the largest component network identified through our gene co-expression and co-occurrence analysis, we utilized Metascape (Zhou et al., 2019) for functional enrichment analysis. Metascape is a powerful tool that enables us to explore the biological functions and pathways of our gene sets in a comprehensive and intuitive manner (Tripathi et al., 2015). In our analysis, we focused on six specific terms for enrichment analysis: Gene Ontology Biological Process (GO-BP), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, REACTOME pathways, WikiPathways, Canonical pathways, and CORUM pathway (Tripathi et al., 2015). These terms were selected for their relevance to NSCLC and their potential to illuminate molecular pathways implicated in carcinogenesis. We also used Metascape's Molecular Complex Detection (MCODE) technique to cluster the enriched phrases into larger groupings to discover common biological themes and better understand NSCLC's processes. Clustering these terms revealed NSCLC tumours' most enriched and significant biological processes and pathways.

**Drug-Gene Interaction**

We utilized the Drug Gene Interaction Database (DGIdb) to identify potential drugs that target key genes within the two gene co-expression and co-occurrence modules identified in our analysis (Cotto et al., 2018). The DGIdb is a comprehensive database that integrates multiple sources of information to identify potential drug targets and interactions. Using this database, we were able to identify candidate drugs that target key genes within the two modules, providing us with a list of potential drug candidates for further investigation. Using Cytoscape, a potent software for visualising and analysing complicated networks, we created a drug-gene interaction network to better understand the potential interactions between these medications and the essential genes contained inside the modules (Shannon et al., 2003). The drug-gene interaction network allowed us to gain a more comprehensive understanding of the potential interactions between the candidate drugs and the key genes within the modules, highlighting potential drug targets.

* 1. **Candidate Drug Validation**

We used the STITCH database and gene set enrichment analysis (GSEA) to further validate the probable candidate medications discovered through our investigation. With the help of the sophisticated computational technique GSEA, we can identify whether a group of genes exhibits statistically significant variations between two biological states (Subramanian et al., 2005). In this case, we utilized GSEA to determine whether the identified candidate drugs were enriched in NSCLC-related gene sets (Subramanian et al., 2005), providing us with insights into their potential efficacy for NSCLC therapy. In addition to GSEA, we also utilized the STITCH database, a comprehensive resource that integrates various sources of data to identify potential drug targets and interactions (Szklarczyk et al., 2019). By using the STITCH database, we were able to evaluate the identified candidate drugs and determine whether they had previously been shown to be associated with NSCLC or related biological pathways.

1. **Results**
   1. **Differentially expressed Genes analysis**

According to gene expression data from two independent studies, GSE27262 and GSE21933, a total of 50 and 42 samples were analyzed, respectively. Utilizing the GEO2R tool, it was found that 458 genes exhibited differential expression in GSE27262, while GSE21933 displayed differential expression in 797 genes. Among these genes, a total of 400 were upregulated, indicating increased expression levels, while 675 genes were downregulated, indicating decreased expression levels (Figure 1). The relevant genes were used for subsequent analysis.

**Figure 1a.** Venn diagram of upregulated intercepted genes.  **Figure 1b.** Venn diagram of downregulated intercepted genes.

* 1. **Module analysis & Gene co-expression Network Construction**

In order to explore the relationship and interactions between differentially expressed genes in both GSE27262 and GSE21933 datasets, a gene co-expression and co-occurrence network was meticulously constructed utilizing the STRING database. Prior to network construction, any disconnected genes were systematically removed from the analysis, ensuring a focused and interconnected network. As a result, the resulting network encompassed a substantial number of nodes, totaling 1024, and demonstrated numerous edges, amounting to 8652 in total.

To delve deeper into the NSCLC human interactome data, the gene co-expression network was further analyzed using the STRING database. Employing this approach, two distinct gene modules (Figure 2) were successfully identified within the human interactome network, employing the MCODE algorithm. These modules effectively represented clusters of genes that displayed significant co-expression patterns and shared functional relationships. Notably, the purple module emerged as the smaller module, housing 28 genes that exhibited intricate interconnections and concerted expression behavior. On the other hand, the largest module, known as the blue module, comprised an impressive collection of 92 genes, showcasing a robust network of interdependencies and interplay.

Module A

Module B

**Figure 2.**  Module (A and B) of the interaction network of connected genes drawn using Cytoscape (Shannon et al., 2003) v. 3.9.1. The circle nodes represent genes while the line edges show their interactions.

* 1. **Enrichment Analysis of Gene Module**

The biological function of the genes in the two modules was carried out using a hypergeometric test and Benjamini-Hochberg statistical correction algorithm in Metascape. It was done using the following ontology sources: KEGG Pathway (Kanehisa et al., 2017), GO Biological Processes (Gene Ontology Consortium, 2004), Reactome Gene Sets (Gillespie et al., 2022), CORUM (Tsitsiridis et al., 2023), PANTHER Pathway ((Mi & Thomas, 2009), and WikiPathways (Martens et al., 2021). The result shows that the terms were primarily involved in mitotic cell cycle, kinetochore organization and DNA metabolic processes (Figure 3). They are also involved in degradation of collagen and cell cycle proteins.

**Figure 3.** Bar graph of p-value-colored enriched terms from gene lists obtained from Metascape (Zhou et al., 2019).

In addition, Clustering Analysis was done using the MCODE algorithm in Metascape resulting in eight modules (Figure 4), for molecular detection and to identify similar biological themes and relevant biological pathways associated with NSCLC. As detailed in Table 1, the genes are associated with the cell cycle and involved in processes such as mitosis, collogen formation, DNA replication and repair.

**Figure 4.** Module detection using MCODE algorithm in Metascape (Zhou et al., 2019). Circles represent protein nodes. Nodes in each subgraph are colored differently for different modules.

* 1. **Drug-Gene Interaction and Validation**

A drug-gene interaction network was constructed using the genes co-expressed in the two important gene modules identified and drugs in the Drug Gene Interaction Database (DGIdb) (Cotto et al., 2018) to detect potential drugs for NSCLC treatment. The interaction network in was reconstructed using Cytoscape (Shannon et al., 2003) v. 3.9.1. The network is shown in Figure 5. 8 drugs show high (above 3) interactions with genes while 10 drugs have 3 interactions with genes (Table 2).

**Figure 5.** Drug-gene interaction network drawn with Cytoscape (Shannon et al., 2003) v. 3.9.1. The red nodes indicate the gene while the blue nodes are drugs that interact with those genes. The size of the blue nodes (drug) indicates how many genes they interact with; he larger the size, the more genes they interact with.

STITCH v5.0 (Kuhn et al., 2007), a drug-protein interaction database, was used to validate the results from the drug-gene interaction network using a medium confidence cutoff of 0.4. The STITCH result is shown in Figure 6. Pathway enrichment analysis from STITCH shows that Sorafenib, Pazopanib and Methotrexate interact with proteins in the peptidyl-tyrosine phosphorylation pathway, vascular endothelial growth factor signaling pathway and uracil metabolic process and they are involved in transmembrane receptor protein tyrosine kinase activity.

**Figure 6. Drug-protein interaction network of candidate drugs targeting the key proteins resulted from STITCH (Kuhn et al., 2007) v5.0 a. Drug-protein interaction network in drugs with more than three nodes in the drug-gene interaction database. b. Drug-protein interaction network of candidate drugs with three nodes in the drug-gene interaction database**

1. **Discussion:**

The goal of this study was aimed at discovering potentially useful drugs for the treatment of non-small cell lung cancer (NSCLC) by using a systems biology and network analysis technique to identify important genes and proteins involved in NSCLC and to build a drug-gene interaction network for potential drug repurposing. Due to the ineffectiveness and considerable side effects of conventional cancer treatments, drug repurposing has emerged as a possible method for developing new medicines for NSCLC. The study used transcriptomics data from two GEO datasets to find differentially expressed genes and build a gene co-expression network in order to find new medication candidates. As a result, two gene co-expression modules were discovered. Candidate medications that target crucial genes in the two NSCLC modules were found using the drug-gene interaction database. The creation of a drug-gene interaction network was made easier by Cytoscape, and candidate medications were validated using gene set enrichment analysis.

Enrichment analysis is a crucial tool for identifying significant biological processes and pathways in cancer research (Liu et al., 2022). In this study, the enrichment analysis revealed that several processes, including cell division, the mitotic cell cycle, collagen formation, and DNA repair, are important in cancer development and progression. Dysregulation of cell division and the mitotic cell cycle leads to uncontrolled proliferation, which is a hallmark of cancer development (Matthews et al., 2021). The enrichment analysis showed that genes involved in the mitotic cell cycle are significantly enriched in cancer cells, highlighting their critical role in cancer development. The dysregulation of collagen formation and degradation, which plays a crucial role in the extracellular matrix that provides structural support to cells, was also found to be critical in cancer progression and metastasis (Martins Cavaco et al., 2020). Additionally, DNA repair was identified as a critical process in preventing the accumulation of mutations that can lead to cancer development (Li et al., 2021). The significant enrichment of genes involved in DNA metabolic processes in cancer cells indicated that dysregulation of DNA repair is critical in cancer development. Therefore, targeting these processes through drug intervention could offer potential therapeutic benefits for NSCLC patients.

The drug-gene interaction network analysis in this study identified seven modal genes (TOP2A, TYMS, BIRC5, GMNN, CCNA2, AURKA, and AURKB) that have potential roles in NSCLC treatment. These genes are known to be involved in cell division, DNA replication, and DNA repair processes, which are essential for cancer cell growth and proliferation. TOP2A is an enzyme involved in DNA replication and transcription, and its overexpression has been associated with poor prognosis in NSCLC patients (Kou et al., 2020; Li et al., 2021). Also, several TOP2A inhibitors, such as etoposide and doxorubicin, have shown promising results in preclinical studies for cancer treatment (Matias-Barrios and Dong, 2023). TYMS, a key enzyme involved in DNA synthesis, has been shown to be upregulated in NSCLC and is a potential target for chemotherapy drugs like 5-fluorouracil (Kotoula et al., 2012).

BIRC5 (survivin) is a protein involved in cell division and is associated with drug resistance and a poor prognosis (Frazzi et al., 2020). GMNN encodes geminin, a protein that plays a critical role in DNA replication and cell cycle regulation. GMNN overexpression has been observed in breast cancer, and it is associated with a poor prognosis (Hernández-Pérez et al., 2017). CCNA2 encodes cyclin A2, a protein that plays a critical role in cell cycle regulation. In a study by Qian et al. (2015), it was found that the overexpression of the CCNB2 protein is associated with clinical progression and a poor prognosis in NSCLC. AURKA and AURKB encode aurora kinases A and B, respectively, which play critical roles in mitosis and cell division. Overexpression of AURKA and AURKB has been observed in various cancers, including NSCLC, and is associated with poor prognosis and drug resistance (Yu et al., 2018). In summary, all the modal genes in these studies have potential roles in NSCLC treatment and have been previously reported as potential targets for NSCLC treatment.

The gene co-expression and co-occurrence network approach used in this study has a unique feature that sets it apart from previous studies. The accuracy of the research results improves since it takes into consideration the variable gene expression patterns between cancer cells and the adjacent normal cells.. The findings of this study support prior studies that identified several candidate drugs that could be repurposed for NSCLC treatment, such as vincristine, fluorouracil, methotrexate, clotrimazole, etoposide, tamoxifen, sorafenib, doxorubicin, and pazopanib, based on their strong gene regulation.

Vincristine, a vinca alkaloid (Dhyani et al., 2022) and a conventional chemotherapeutic medication, has been studied for its synergistic effect in the treatment of NSCLC in multiple studies. (Sampson et al., 2022; Ghosh et al., 2020; Ghosh et al., 2021; Kumar et al., 2018). This alkaloid (vincristine drug) acts as an anti-microtubule agent that blocks mitosis by blocking cells in the metaphase (Martino et al., 2018).

Fluorouracil is a chemotherapy drug that targets TYMS and has shown efficacy in treating NSCLC. 5-fluorouracil has been investigated for its anticancer properties, including against NSCLC (Dhumad et al., 2021; Arias, 2008; Wang et al., 2015). Methotrexate is another chemotherapy drug that targets folate metabolism and has been investigated for the synergistic chemotherapy of NSCLC (Subaiea et al., 2023; Yi et al., 2012). Clotrimazole, which is commonly used as an antifungal medication, has shown potential as an antitumor drug in reducing the size and growth of neoplasms in previous studies (Kadavakollu et al., 2014). In a study carried out by Sebastian et al., there was an increase towards prolonged life with concomitant clotrimazole and ICI treatment for advanced NSCLC (Sebastian et al., 2021). Etoposide is a topoisomerase II inhibitor that has been used in combination with other chemotherapy drugs for the treatment of NSCLC (Ferguson et al., 2015). Tamoxifen is a selective estrogen receptor modulator that has been investigated for its potential in treating breast cancer (Peng et al., 2019) and is still under investigation for treating NSCLC. Sorafenib is a multi-kinase inhibitor that targets several signalling pathways involved in cancer development and progression (Adnane et al., 2006). Doxorubicin is an anthracycline antibiotic and one of the most used chemotherapy drugs (Tacar et al., 2013). Pazopanib is a tyrosine kinase inhibitor that targets angiogenesis and has been investigated as a potential treatment for cancer (Tullemans et al., 2018). These drugs hold promise as potential treatments for NSCLC and warrant further investigation.

Despite identifying potential drug candidates for NSCLC treatment, further investigation is necessary to evaluate their efficacy for specific cancer conditions. In-silico studies could provide additional evidence to support the potential repurposing of these drugs. The findings of this study highlight the importance of gene co-expression and network analysis in drug development for existing diseases and the advantages of drug repurposing as a quicker and more cost-effective approach with potential improved safety profiles. The identified drugs in this study hold promise as effective treatments for NSCLC and warrant further investigation.

1. **Conclusion:**

In this research paper, we proposed a systems biology and network analysis approach to identify potential drugs for the treatment of non-small cell lung cancer (NSCLC). We addressed the limitations of current treatments, such as drug resistance, toxicity, and low survival rates, by utilizing transcriptomics data and constructing gene co-expression and co-occurrence networks. Through our analysis, we identified differentially expressed genes in NSCLC and identified two gene co-expression modules. By leveraging the Drug-Gene interaction database, we identified candidate drugs that target essential genes within these modules. Additionally, we constructed a drug-gene interaction network and validated the candidate drugs using gene set enrichment analysis.

Unlike previous research that solely relied on gene co-expression networks, our approach considered both cancer cells and adjacent normal cells, thereby potentially reducing the side effects of treatment. The candidate drugs we identified, including topoisomerase inhibitors and proteasome inhibitors, have demonstrated effectiveness in preclinical and clinical studies for NSCLC and other cancers. Nevertheless, our study has several limitations. It relies on the assumption that the Drug-Gene interaction database is comprehensive and accurate, which may not always be the case. Additionally, transcriptomics data may not fully capture proteomics and metabolomics changes in NSCLC. Furthermore, being a computational analysis, our findings require further experimental validation.

**REFERENCES**

Adnane, L., Trail, P. A., Taylor, I., & Wilhelm, S. M. (2006). Sorafenib (BAY 43-9006, Nexavar), a dual-action inhibitor that targets RAF/MEK/ERK pathway in tumor cells and tyrosine kinases VEGFR/PDGFR in tumor vasculature. Methods Enzymol, 407, 597-612. doi: 10.1016/S0076-6879(05)07047-3. PMID: 16757355.

Arias JL. Novel strategies to improve the anticancer action of 5-fluorouracil by using drug delivery systems. Molecules. 2008 Oct 1;13(10):2340.

Bader, G. D., & Hogue, C. W. (2003). An automated method for finding molecular complexes in large protein interaction networks. BMC bioinformatics, 4(1), 2. <https://doi.org/10.1186/1471-2105-4-2>

Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., Marshall, K. A., Phillippy, K. H., Sherman, P. M., Holko, M., Yefanov, A., Lee, H., Zhang, N., Robertson, C. L., Serova, N., Davis, S., Soboleva, A. (2013). NCBI GEO: archive for functional genomics data sets--update. Nucleic acids research, 41(D1), D991-D995. <https://doi.org/10.1093/nar/gks1193>

Cotto, K. C., Wagner, A. H., Feng, Y. Y., Kiwala, S., Coffman, A. C., Spies, G., ... & Griffith, O. L. (2018). DGIdb 3.0: a redesign and expansion of the drug–gene interaction database. Nucleic acids research, 46(D1), D1068-D1073.

Crino, L., Weder, W., Van Meerbeeck, J., & Felip, E. S. M. O. (2010). Early stage and locally advanced (non-metastatic) non-small-cell lung cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of oncology*, *21*, v103-v115.

Dhumad AM, Majeed HJ, Zandi H, Harismah K. FeC19 cage vehicle for fluorouracil anticancer drug delivery: DFT approach. Journal of Molecular Liquids. 2021 Jul 1;333:115905.

Dhyani, P., Quispe, C., Sharma, E. *et al.* Anticancer potential of alkaloids: a key emphasis to colchicine, vinblastine, vincristine, vindesine, vinorelbine and vincamine. *Cancer Cell Int* 22, 206 (2022). <https://doi.org/10.1186/s12935-022-02624-9>

Dongsar, T. T., Dongsar, T. S., Gupta, N., Almalki, W. H., Sahebkar, A., & Kesharwani, P. (2023). Emerging potential of 5-Fluorouracil-loaded chitosan nanoparticles in cancer therapy. *Journal of Drug Delivery Science and Technology*, 104371.

Ferguson, D. M., Jacobson, B. A., Jay-Dixon, J., Patel, M. R., Kratzke, R. A., & Raza, A. (2015). Targeting Topoisomerase II Activity in NSCLC with 9-Aminoacridine Derivatives. Anticancer Research, 35(10), 5211-5217.

Frazzi, R. (2021). BIRC3 and BIRC5: multifaceted inhibitors in cancer. Cell Bioscience, 11(1), 8. <https://doi.org/10.1186/s13578-020-00521-0>

Gene Ontology Consortium (2004) ‘The Gene Ontology (GO) database and informatics resource’, *Nucleic Acids Research*, 32(90001). doi:10.1093/nar/gkh036.

Ghosh S, Lalani R, Maiti K, Banerjee S, Bhatt H, Bobde YS, Patel V, Biswas S, Bhowmick S, Misra A. Synergistic co-loading of vincristine improved chemotherapeutic potential of pegylated liposomal doxorubicin against triple negative breast cancer and non-small cell lung cancer. Nanomedicine: Nanotechnology, Biology and Medicine. 2021 Jan 1;31:102320.

Ghosh S, Lalani R, Maiti K, Banerjee S, Patel V, Bhowmick S, Misra A. Optimization and efficacy study of synergistic vincristine coloaded liposomal doxorubicin against breast and lung cancer. Nanomedicine. 2020 Sep;15(26):2585-607.

Gillespie, M., Jassal, B., Stephan, R., Milacic, M., Rothfels, K., Senff-Ribeiro, A., ... & D’Eustachio, P. (2022). The reactome pathway knowledgebase 2022. *Nucleic acids research*, *50*(D1), D687-D692.

Hernández-Pérez, S., Cabrera, E., Salido, E., Lim, M., Reid, L., Lakhani, S. R., Khanna, K. K., Saunus, J. M., & Freire, R. (2017). DUB3 and USP7 de-ubiquitinating enzymes control replication inhibitor Geminin: molecular characterization and associations with breast cancer. Oncogene, 36(33), 4802-4809. doi: 10.1038/onc.2017.21. Erratum in: Oncogene. 2017 Jun 26;: Erratum in: Oncogene. 2019 Jun;38(24):4886. PMID: 28288134.

Hu, H., Piotrowska, Z., Hare, P. J., Chen, H., Mulvey, H. E., Mayfield, A., ... & Engelman, J. A. (2021). Three subtypes of lung cancer fibroblasts define distinct therapeutic paradigms. *Cancer Cell*, *39*(11), 1531-1547.

Kadavakollu S, Stailey C, Kunapareddy CS, White S. Clotrimazole as a cancer drug: a short review. Medicinal chemistry. 2014;4(11):722.

Kalinke, L., Thakrar, R., & Janes, S. M. (2021). The promises and challenges of early non‐small cell lung cancer detection: patient perceptions, low‐dose CT screening, bronchoscopy and biomarkers. *Molecular Oncology*, *15*(10), 2544-2564.

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., & Morishima, K. (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic acids research*, *45*(D1), D353-D361.

Kotoula, V., Krikelis, D., Karavasilis, V., & et al. (2012). Expression of DNA repair and replication genes in non-small cell lung cancer (NSCLC): a role for thymidylate synthetase (TYMS). BMC Cancer, 12, 342. <https://doi.org/10.1186/1471-2407-12-342>

Kou, F., Sun, H., Wu, L., Li, B., Zhang, B., Wang, X., & Yang, L. (2020). TOP2A promotes lung adenocarcinoma cells' malignant progression and predicts poor prognosis in lung adenocarcinoma. *Journal of Cancer*, *11*(9), 2496.

Kuhn, M., von Mering, C., Campillos, M., Jensen, L. J., & Bork, P. (2007). STITCH: interaction networks of chemicals and proteins. *Nucleic acids research*, *36*(suppl\_1), D684-D688.

Kumar N, Salar RK, Prasad M, Ranjan K. Synthesis, characterization and anticancer activity of vincristine loaded folic acid-chitosan conjugated nanoparticles on NCI-H460 non-small cell lung cancer cell line. Egyptian Journal of Basic and Applied Sciences. 2018 Mar 1;5(1):87-99.

Li, L. Y., Guan, Y. D., Chen, X. S., Yang, J. M., & Cheng, Y. (2021). DNA Repair Pathways in Cancer Therapy and Resistance. Frontiers in Pharmacology, 11, 629266. https://doi.org/10.3389/fphar.2020.629266

Liu, H., Yuan, M., Mitra, R., Wang, J., Fang, X., Du, Y., ... & Liu, J. S. (2022). CTpathway: a CrossTalk-based pathway enrichment analysis method for cancer research. Genome Medicine, 14(1), 118. https://doi.org/10.1186/s13073-022-01119-6

Lo, F. Y., Chang, J. W., Chang, I., Chen, Y. J., Hsu, H. S., Huang, S. F. K., ... & Wang, Y. C. (2012). The database of chromosome imbalance regions and genes resided in lung cancer from Asian and Caucasian identified by array-comparative genomic hybridization. *BMC cancer*, *12*(1), 1-13.

Martens, M., Ammar, A., Riutta, A., Waagmeester, A., Slenter, D. N., Hanspers, K., ... & Kutmon, M. (2021). WikiPathways: connecting communities. *Nucleic acids research*, *49*(D1), D613-D621.

Martino E, Casamassima G, Castiglione S, Cellupica E, Pantalone S, Papagni F, Rui M, Siciliano AM, Collina S. Vinca alkaloids and analogues as anti-cancer agents: looking back, peering ahead. Bioorg Med Chem Lett. 2018;28(17):2816–26.

Martins Cavaco, A. C., Dâmaso, S., Casimiro, S., & Costa, L. (2020). Collagen biology making inroads into prognosis and treatment of cancer progression and metastasis. Cancer and Metastasis Reviews, 39(3), 603–623. https://doi.org/10.1007/s10555-020-09888-5

Matias-Barrios, V. M., & Dong, X. (2023). The Implication of Topoisomerase II Inhibitors in Synthetic Lethality for Cancer Therapy. Pharmaceuticals (Basel, Switzerland), 16(1), 94. https://doi.org/10.3390/ph16010094

Matthews, H. K., Bertoli, C., & de Bruin, R. A. M. (2021). Cell cycle control in cancer. Nature Reviews Molecular Cell Biology. <https://doi.org/10.1038/s41580-021-00404-3>

Mi, H., & Thomas, P. (2009). PANTHER pathway: an ontology-based pathway database coupled with data analysis tools. *Protein networks and pathway analysis*, 123-140.

MotieGhader, H., Tabrizi-Nezhadi, P., Deldar Abad Paskeh, M., Baradaran, B., Mokhtarzadeh, A., Hashemi, M., ... & Masoudi-Nejad, A. (2022). Drug repositioning in non-small cell lung cancer (NSCLC) using gene co-expression and drug–gene interaction networks analysis. *Scientific Reports*, *12*(1), 9417.

Peng, J., Sengupta, S., & Jordan, V. C. (2009). Potential of selective estrogen receptor modulators as treatments and preventives of breast cancer. Anti-cancer agents in medicinal chemistry, 9(5), 481–499. https://doi.org/10.2174/187152009788451833

Qian, X., Song, X., He, Y., Yang, Z., Sun, T., Wang, J., Zhu, G., Xing, W., & You, C. (2015). CCNB2 overexpression is a poor prognostic biomarker in Chinese NSCLC patients. Biomed Pharmacother, 74, 222-227. doi: 10.1016/j.biopha.2015.08.004. PMID: 26349989.

Rajasegaran, T., How, C. W., Saud, A., Ali, A., & Lim, J. C. W. (2023). Targeting Inflammation in Non-Small Cell Lung Cancer through Drug Repurposing. *Pharmaceuticals*, *16*(3), 451.

Sampson J, Ju HM, Song JY, Fry AM, Bayliss R, Choi J. A Polytherapy Strategy Using Vincristine and ALK Inhibitors to Sensitise EML4-ALK-Positive NSCLC. Cancers. 2022 Feb 2;14(3):779.

Schwendenwein, A., Megyesfalvi, Z., Barany, N., Valko, Z., Bugyik, E., Lang, C., ... & Dome, B. (2021). Molecular profiles of small cell lung cancer subtypes: therapeutic implications. *Molecular Therapy-Oncolytics*, *20*, 470-483.

Sebastian N, Stokes WA, Behera M, Jiang R, Gutman D, Giuste F, Burns A, Ramalingam S, Sukhatme V, Lowe MC, Ramalingam SS. Association of azole antifungals with survival in patients with non-small cell lung cancer receiving immunotherapy.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., ... & Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome research, 13(11), 2498-2504.

Siegel, R. L., Miller, K. D., Fuchs, H. E., & Jemal, A. (2022). Cancer statistics, 2022. *CA: a cancer journal for clinicians*, *72*(1), 7–33. <https://doi.org/10.3322/caac.21708>

Singh, M., Kumar, V., Sehrawat, N., Yadav, M., Chaudhary, M., Upadhyay, S. K., ... & Sharma, A. K. (2022, August). Current paradigms in epigenetic anticancer therapeutics and future challenges. In *Seminars in Cancer Biology* (Vol. 83, pp. 422-440). Academic Press.

Subaiea G, Rizvi SM, Yadav HK, Al Hagbani T, Abdallah MH, Khafagy ES, Gangadharappa HV, Hussain T, Abu Lila AS. Ganetespib with Methotrexate Acts Synergistically to Impede NF-κB/p65 Signaling in Human Lung Cancer A549 Cells. Pharmaceuticals. 2023 Feb 2;16(2):230.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., ... & Mesirov, J. P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences, 102(43), 15545-15550.

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, *71*(3), 209-249.

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., ... & Jensen, L. J. (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.

Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., ... & von Mering, C. (2023). The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Research*, *51*(D1), D638-D646.

Tacar, O., Sriamornsak, P., & Dass, C. R. (2013). Doxorubicin: an update on anticancer molecular action, toxicity and novel drug delivery systems. Journal of Pharmacy and Pharmacology, 65, 157-170. <https://doi.org/10.1111/j.2042-7158.2012.01567.x>

Tsitsiridis, G., Steinkamp, R., Giurgiu, M., Brauner, B., Fobo, G., Frishman, G., ... & Ruepp, A. (2023). CORUM: the comprehensive resource of mammalian protein complexes–2022. *Nucleic Acids Research*, *51*(D1), D539-D545.

Tripathi, S., Pohl, M. O., Zhou, Y., Rodriguez-Frandsen, A., Wang, G., Stein, D. A., ... & García-Sastre, A. (2015). Meta-and orthogonal integration of influenza “OMICs” data defines a role for UBR4 in virus budding. Cell host & microbe, 18(6), 723-735.

Tullemans BME, Nagy M, Sabrkhany S, Griffioen AW, Oude Egbrink MGA, Aarts M, Heemskerk JWM, Kuijpers MJE. (2018). Tyrosine Kinase Inhibitor Pazopanib Inhibits Platelet Procoagulant Activity in Renal Cell Carcinoma Patients. Front Cardiovasc Med, 5, 142. doi: 10.3389/fcvm.2018.00142. PMID: 30460241; PMCID: PMC6232667.

Wang H, Yang T, Wu X. 5-Fluorouracil preferentially sensitizes mutant KRAS non-small cell lung carcinoma cells to TRAIL-induced apoptosis. Molecular oncology. 2015 Nov 1;9(9):1815-24.

Wang, W. Z., Shulman, A., Amann, J. M., Carbone, D. P., & Tsichlis, P. N. (2022, April). Small Cell Lung Cancer: Subtypes and Therapeutic Implications. In *Seminars in Cancer Biology*. Academic Press.

Wei, T. Y. W., Hsia, J. Y., Chiu, S. C., Su, L. J., Juan, C. C., Lee, Y. C. G., ... & Yu, C. T. R. (2014). Methylosome protein 50 promotes androgen-and estrogen-independent tumorigenesis. *Cellular signalling*, *26*(12), 2940-2950.

Wei, T. Y. W., Juan, C. C., Hisa, J. Y., Su, L. J., Lee, Y. C. G., Chou, H. Y., ... & Yu, C. T. R. (2012). Protein arginine methyltransferase 5 is a potential oncoprotein that upregulates G 1 cyclins/cyclin‐dependent kinases and the phosphoinositide 3‐kinase/AKT signaling cascade. *Cancer science*, *103*(9), 1640-1650.

Yi H, Cho HJ, Cho SM, Jo K, Park JA, Lee SH, Chang BJ, Kim JS, Shin HC. Effect of 5-FU and MTX on the expression of drug-resistance related cancer stem cell markers in non-small cell lung cancer cells. The Korean Journal of Physiology & Pharmacology. 2012 Feb 1;16(1):11-6.

Yu, J., Zhou, J., Xu, F., Bai, W., & Zhang, W. (2018). High expression of Aurora-B is correlated with poor prognosis and drug resistance in non-small cell lung cancer. The International Journal of Biological Markers, 33(2), 215–221. doi:10.1177/1724600817753098

Zappa, C. and S. A. Mousa (2016). "Non-small cell lung cancer: current treatment and future advances." Translational lung cancer research 5(3): 288.

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., Benner, C., & Chanda, S. K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature communications*, *10*(1), 1523. https://doi.org/10.1038/s41467-019-09234-6

**Table 1.** Clustering of functional enrichment analysis using Metascape (Zhou et al., 2019).

|  |  |  |  |
| --- | --- | --- | --- |
| Functional Component | Term ID | Biological Function | Log10(p-Value) |
| MCODE\_1 | GO:0051301 | Cell division | -24.8 |
| GO:1903047 | mitotic cell cycle process | -24.7 |
| GO:0000278 | mitotic cell cycle | -23.5 |
| MCODE\_2 | R-HSA-2500257 | Resolution of Sister Chromatid Cohesion | -26.8 |
| R-HSA-68877 | Mitotic Prometaphase | -24.2 |
| R-HSA-68882 | Mitotic Anaphase | -23.5 |
| MCODE\_3 | R-HSA-2500257 | Resolution of Sister Chromatid Cohesion | -22.9 |
| R-HSA-141444 | Amplification of signal from unattached kinetochores via a MAD2 inhibitory signal | -20.9 |
| R-HSA-141424 | Amplification of signal from the kinetochores | -20.9 |
| MCODE\_4 | GO:1903047 | mitotic cell cycle process | -9.8 |
| R-HSA-69278 | Cell Cycle, Mitotic | -9.6 |
| GO:0000278 | mitotic cell cycle | -9.3 |
| MCODE\_5 | R-HSA-1650814 | Collagen biosynthesis and modifying enzymes | -16.0 |
| R-HSA-1474290 | Collagen formation | -15.2 |
| M3005 | NABA COLLAGENS | -13.5 |
| MCODE\_6 | WP4946 | DNA repair pathways, full network | -11.3 |
| R-HSA-5693616 | Presynaptic phase of homologous DNA pairing and strand exchange | -10.4 |
| R-HSA-5693579 | Homologous DNA Pairing and Strand Exchange | -10.3 |
| MCODE\_7 | WP129 | Matrix metalloproteinases | -12.1 |
| R-HSA-1592389 | Activation of Matrix Metalloproteinases | -11.9 |
| GO:0030574 | collagen catabolic process | -11.5 |
| MCODE\_8 | R-HSA-176974 | Unwinding of DNA | -10.3 |
| GO:0006268 | DNA unwinding involved in DNA replication | -9.5 |
| R-HSA-69190 | DNA strand elongation | -9.0 |

**Table 2.** Potential drugs with three or more target genes from drug-gene interaction network (Cotto et al., 2018)

|  |  |  |
| --- | --- | --- |
| Drug | Number of interacting nodes (genes) | Target gene |
| Fluorouracil | 8 | TOP2A, AURKA, SELE, TYMS, HMMR, BIRC5, CHEK1, EXO1 |
| Methotrexate | 5 | IL2, TYMS, BIRC5, BMP7, HMMR |
| Etoposide | 5 | CHEK1, TYMS, GMNN, TOP2A, AURKB |
| Vincristine | 5 | TYMS, GMNN, BIRC5, BMPT, TOP2A |
| Tamoxifen | 4 | TYMS, CCNA2, AURKA, AURKB |
| Sorafenib | 4 | AURKA, GMNN, PLK4, AURKB |
| Doxorubicin | 4 | EZH2, HMMR, BIRC5, BMPT |
| Collagenase Clostridium Histolyticum | 4 | COL11A1, COL4A3, COL5A2, MMP1 |
| Ocriplasmin | 3 | COL11A1, COL4A3, COL5A2 |
| Doxycycline | 3 | MMP13, MMP7, MMP1 |
| Doxycycline calcium | 3 | MMP13, MMP7, MMP1 |
| Cisplastin | 3 | CHEK1, TYMS, AURKA |
| Daunorubicin | 3 | TYMS, TOP2A, BMP7 |
| Paclitaxel | 3 | GMNN, AURKA, EZH2 |
| Vorinostat | 3 | GMNN, EZH2, BIRC5 |
| Pazopanib | 3 | NEK2, AURKA, CCNA2 |
| Clotrimazole | 3 | GMNN, CXCR2, CDK1 |
| Capecitabin | 3 | TYMS, SELE, EXO1 |
| Cytarabine | 3 | RRM2, BMP7, TYMS |