



## Original Article

# Small Volatile Inhibitors Targeting lncRNA in Radioresistant Non-small Cell Lung Cancer



Arnab Mukherjee and Mukunthan Kuppasamy Selvam\* 

Department of Biotechnology, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India

Received: January 28, 2023 | Revised: March 24, 2023 | Accepted: May 11, 2023 | Published online: June 12, 2023

### Abstract

**Background and objectives:** Radiotherapy is the most crucial nonsurgical therapeutic method in the multidisciplinary care of patients with non-small cell lung cancer (NSCLC). However, radiation resistance continues to be a significant clinical issue, negatively affecting cancer prognosis in patients. The small molecules that target these RNAs offer therapeutic modulation of multiple biological processes. The study aims to identify the genes and a long non-coding RNA (lncRNA) regulating their expression levels, as well as the binding potential of small volatile inhibitors to the lncRNA to overcome radioresistance in NSCLC.

**Methods:** The non-coding RNA microarray dataset of NSCLC cells was analysed to identify the differentially expressed genes regulated by lncRNA, which drives radioresistance. The study comprises three volatile ligands due to their good pharmacokinetic profile to target the identified lncRNA.

**Results:** The analysis revealed the dysregulation of the cell cycle, evasion of apoptosis and cancer immune response. A co-expression analysis with a network pharmacology approach revealed an lncRNA ENST00000605056 regulating three highly ranked hub genes. The molecular interaction studies uncovered their high binding affinity to its binding pocket with a preponderance of non-covalent bond interactions between the ligand and the nucleotides. The molecular dynamics simulations revealed the binding stability of ligands to the lncRNA with a very low deviation compared to the control.

**Conclusion:** This study demonstrates the ability of small molecules to target lncRNA in addressing the global concern of radioresistance among NSCLC patients, thereby facilitating future translational studies.

### Introduction

Lung carcinoma remains one of the most prevalent and is estimated to be a primary cause of cancer mortality globally.<sup>1,2</sup> Lung adenocarcinoma constitutes 80–85% of lung cancer incidences, also termed non-small cell lung cancer (NSCLC).<sup>3</sup> Radiotherapy is the most important nonsurgical strategy for NSCLC patients and is required for approximately 70% of NSCLC patients during treat-

ment.<sup>4</sup> Radiation therapy is effective when surgery is not an alternative in the early stages of NSCLC. However, it often recurs after radiation therapy, and the relapsed NSCLC is commonly re-irradiated in this condition, with poor outcomes; as a result, individuals may acquire radioresistance.<sup>5</sup> It develops in certain patients due to tumour heterogeneity concerning cell origin, aetiology, and molecular mechanisms, which is the most critical question to address since it results in radiotherapy failure.<sup>6</sup> Furthermore, the critical components involved in radioresistance are still unclear.

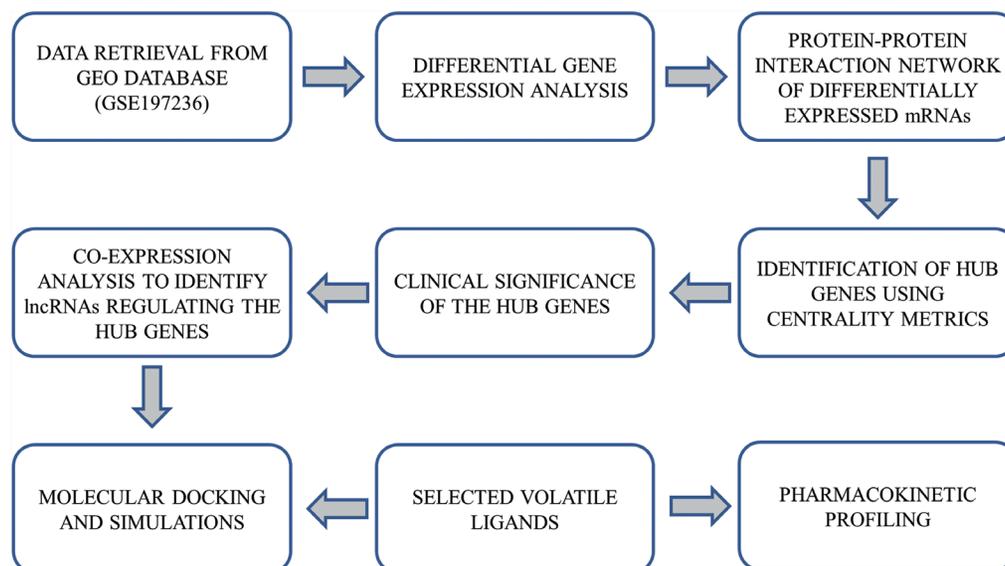
Long non-coding RNA (lncRNA) is a non-coding RNA with more than 200 nucleotides involved in the progression of diseases and contributes to various cellular and biological processes.<sup>7,8</sup> The upregulation of the lncRNAs actively participates in cancer progression by inducing activation of oncogenes and silencing tumour suppressors.<sup>9,10</sup> Recently, several functional lncRNAs have been considered critical factors in the induction of radioresistance responses by modulating the transcriptional or post-transcriptional levels of their target gene expression.<sup>11</sup> lncRNAs offer tremendous potential for developing novel cancer prognoses and therapeutic strategies. The co-expression networks showed important roles in

**Keywords:** Radioresistance; Non-small cell lung cancer; lncRNA; Co-expression; Volatile inhibitors; Docking; Molecular dynamics.

**Abbreviations:** CID, identification numbers; DE, differentially expressed; DEGs, differentially expressed genes; FDR, false discovery rate; GO, gene ontology; HR, hazard ratio; lncRNA, long non-coding RNA; MD, molecular dynamics; NSCLC, non-small cell lung cancer; RMSD, Root Mean Square Deviation.

\***Correspondence to:** Mukunthan KS, Department of Biotechnology, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India. ORCID: <https://orcid.org/0000-0001-9484-2556>. Tel: 08202924322, Email: [mukunthanselvam@gmail.com](mailto:mukunthanselvam@gmail.com); [mukunthan.ks@manipal.edu](mailto:mukunthan.ks@manipal.edu)

**How to cite this article:** Mukherjee A, Kuppasamy Selvam M. Small Volatile Inhibitors Targeting lncRNA in Radioresistant Non-small Cell Lung Cancer. *Gene Expr* 2023;22(2):92–101. doi: 10.14218/GE.2023.00005S.



**Fig. 1.** The flowchart provides an overview of the techniques and approaches used in the study.

cell proliferation in cancer and regulating related genes in cancer.<sup>12</sup>

The rapid progression in personalized and targeted therapies is paving the way for developing next-generation therapeutics. However, only 15% of the proteins are disease-related and druggable.<sup>13</sup> The small compounds targeting the proteins demand a deep binding pocket on their spatial structure. The small molecules are also deemed a potential lead for targeting lncRNA, eventually inhibiting cancer and metastasis.<sup>14</sup> Moreover, these small molecules offer advantages in enhanced pharmacokinetic attributes, cost, and patient acceptability. Therefore, the authors focused on three selected plant-based volatile ligands recognized by the Joint FAO/WHO Expert Committee on Food Additives as flavouring agents. The reported bioactivities and predicted pharmacokinetic properties of the ligands piqued our interest in exploring their potential in modulating the dysregulated biological processes driving radioresistance in NSCLC pathogenesis. The bioactivities of the selected ligands and their corresponding Chemical Identification numbers (CID) were noted. Syringaldehyde (CID:8655) and chalcone demonstrated anti-breast cancer activity by inhibiting cell proliferation and metastasis.<sup>15</sup> Anti-angiogenic and anti-inflammatory properties and engagement in suppressing reactive oxygen species were reported with vanillyl alcohol (CID:62348).<sup>16,17</sup> Currently, antimicrobial drugs are being repurposed for anticancer therapy, such as Ethyl 2-benzylacetoacetate (CID:246929) that is associated with novel antimicrobial and anticancer drug synthesis and related activity.<sup>18</sup>

Advancements in the high throughput screening of lncRNAs are being explored against cancer, and small molecules targeting these oncogenic lncRNAs are expanding rapidly. Therefore, the study focused on identifying and targeting the lncRNA regulating the clinically significant biomarkers in NSCLC pathogenesis. The study demonstrated the binding potential of the volatile ligands to the lncRNA and the ability to exert enhanced pharmacokinetic effects to inhibit the dysregulated cellular and biological processes driving radioresistance in NSCLC cells.

## Methodology

The research focuses on discovering a lncRNA that regulates clini-

cally significant genes linked to radioresistance in NSCLC and exploring the possible interaction between small volatile inhibitors and the identified lncRNA. Figure 1 illustrates the methods employed in this study, while Supplementary Table 1 lists the bioinformatic tools used. The study does not involve any human or animal subjects.

### Microarray data analysis

The bioinformatics web services based on dry lab approaches are crucial for processing and interpreting microarray data aiding in identifying novel targets for drug development and personalized medicine.<sup>19</sup> The Gene Expression Omnibus (GEO) database was used to retrieve a non-coding RNA profiling microarray dataset (GSE197236) of radioresistant A549 cells, which included mRNAs and lncRNAs. Three radioresistant and three sensitive A549 cells were yielded from the dataset. The microarray used a GPL26963: Agilent-085982 Arraystar human lncRNA V5 microarray platform. The differentially expressed (DE) mRNAs and lncRNAs between the Radioresistant and sensitive A549 cells were determined using the GEO2R tool. The  $\log_2FC$  is a measure used to compare the expression levels of a gene in two different conditions, such as a test and a control group. Differentially expressed genes (DEGs) were achieved by using the Benjamini and Hochberg false discovery rate (FDR)  $< 0.05$  to control false positive rates, with the criteria of a  $\log_2FC$  ( $\log_2$ Fold Change)  $> 1$ .<sup>20</sup>

### Identification of hub genes

The DEGs were leveraged to the Search Tool for the Retrieval of Interacting Genes (STRING, <https://string-db.org/>) to assess their interactions with medium confidence  $> 0.4$  and FDR stringency at 5% significance. The constructed network was visualised using Cytoscape (V 3.9.1) to determine key nodes in the network by employing Molecular Complex Detection (MCODE) clustering algorithm (V2.0). The cytoHubba plugin (V 0.1.) was used to identify the top 5 hub genes based on Maximum Clique Centrality (MCC) topological algorithm.<sup>21,22</sup>

### Functional enrichment analysis of hub genes

A gene ontology (GO) functional enrichment analysis for the hub

genes was performed using ShinyGO V0.76.1 (<http://bioinformatics.sdstate.edu/go/>) to uncover dysregulated biological processes in radioresistant NSCLC cells. The GO biological processes were evaluated at 5% FDR stringency and restricted to Homo sapiens.<sup>23</sup>

### Clinical significance

The KM plotter database (<https://kmpplot.com/analysis/>) assessed the association between the hub gene expression cohorts and the overall survival of NSCLC data (n=719). The yielded KM survival plots to assess the number of samples at risk for various time points of 200 months in subgroups with the expression cohorts of individual hub genes with a Hazard Ratio (HR) to compare the risk of an NSCLC progression between two expression cohorts. An HR > 1 indicates that one group has a higher risk of the event than the others.<sup>24</sup>

### mRNA-lncRNA coexpression analysis

A co-expression network of DE lncRNA-mRNA transcripts was constructed to investigate the underlying functions of the DE mRNAs and the interactions between lncRNAs. Pearson correlation is commonly used to assess the strength and direction of the linear relationship between the expression levels of two set genes. The DE mRNA-lncRNA pairs that co-express each other based on average expression and have a Pearson correlation coefficient > 0.99, representing positive correlation, were determined using the “diffcoexp” package of R (V4.1.2).<sup>25</sup> Finally, the lncRNAs having a high correlation to the hub genes were chosen to construct a co-expression network and visualised using Cytoscape.<sup>26,27</sup>

### lncRNA structure prediction and molecular docking

The lncRNA's secondary structure and folding affinity were assessed using the RNAfold web server (<http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>). Further, the three-dimensional structure of the lncRNA was predicted using the RNAComposer (<https://rnacomposer.cs.put.poznan.pl/>).<sup>28</sup> The selected volatile ligands are orally bioavailable and have low molecular weight and drug-likeness with Hydrogen bond acceptors > 2, demonstrating the potential of forming hydrogen bonds with the target was determined using SwissADME (<http://www.swissadme.ch/>). They were subjected to knowledge-based molecular docking with the lncRNA using HADDOCK 2.4 (<https://wenmr.science.uu.nl/haddock2.4/>) by automatically specifying active and passive residues. The minimum relative solvent accessibility was set as the default of 15%. Unlike ab initio docking procedures, which solely take the coordinates of the structures into account, it uses experimental data to guide the molecular docking process. The HADDOCK's docking technique consists of the following steps: randomising orientations preceded by energy minimisation to avert steric clashes, torsion angle dynamics exploiting torsion angles as degrees of freedom, and tweaking the system using an explicit solvent in Cartesian space.<sup>29</sup>

### Molecular dynamics simulation

The molecular dynamics (MD) simulations provide critical information about lncRNA-ligand complexes' flexibility and stability. It depicts the interaction between the lncRNA and the ligand, including binding sites and conformational changes during the bound and unbound states.<sup>30</sup> The MD simulation steps of the lncRNA-ligand complexes were partially adapted from Junaid *et al.*<sup>31</sup> The MD simulation was carried out using GROMACS (2022.1), in contrast to the ligand-free lncRNA, which served as the control. The AMBER force field generated lncRNA and lncR-

NA-ligand complex topologies. The complexes were solvated using the TIP3P water model and were neutralised by incorporating Na<sup>+</sup> and Cl<sup>-</sup> ions. Additionally, the complexes were relaxed using the steepest descent energy minimisation technique. The minimised complexes were subjected to the equilibration step, maintaining the constant volume and pressure using the Nose-Hoover algorithm. The complexes were then subjected to the production step at 300 K and 1 atm pressure for 20 nanoseconds (ns) to assess the system's stability, and the resultant trajectory data were used to demonstrate the Root Mean Square Deviation (RMSD) of the complexes.

## Results

The study identified overexpression of a lncRNA that regulated clinically significant genes associated with radioresistance in NSCLC. The study's findings are explained in the following section and illustrated in [Supplementary Figure 1](#) and [Supplementary Table 2](#).

### Differential expression of the genes

The RNA microarray dataset contained six samples with sensitive and Radioresistant A549 cells. 1085 DEGs, of which 589 and 496 DE mRNAs and lncRNAs were identified according to the criteria of log<sub>2</sub>FC > 1 as upregulated and log<sub>2</sub>FC < -1 as downregulated, represented with red and blue dots, respectively, with adjusted *P* value < 0.05. The volcano plots ([Fig. 2a-c](#)) represent the differential gene expression of control vs test, DE mRNAs and lncRNAs.

### PPI network and topological analysis

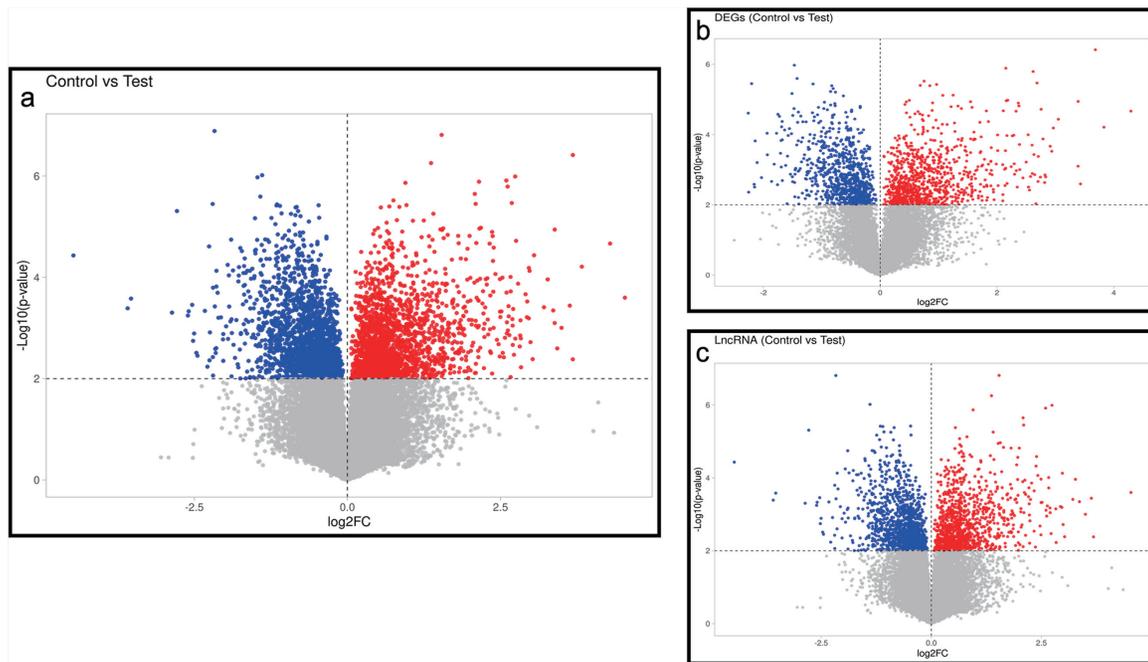
The protein-coding genes were leveraged to the STRING database for elucidating their physical and functional relationship, restricted to “Homo sapiens”. For further visualisation and analysis, the network was exported to Cytoscape. The network with significant DEGs consists of 589 nodes (DE mRNAs) and 1036 edges ([Supplementary Fig. 2](#)) with a clustering coefficient of 0.23, determined through the Analyzer tool of Cytoscape. MCODE determined 19 clusters of densely interconnected nodes from the network. A cluster with 20 nodes and 61 edges and a score of 6.421, cluster 1 was considered significant compared to the other clusters' scores, possessing 12 downregulated and 8 upregulated nodes ([Supplementary Fig. 3a](#)). The cluster uncovered the top 5 hub genes (CKS1B, CDC25C, IRF1, IFNB1, CDT1) using the MCC algorithm ([Supplementary Fig. 3b](#)), which offers better accuracy in assessing the topological relevance of nodes when compared to other centrality algorithms.

### Biological processes and functional enrichment

The GO biological processes were determined for the hub genes at a 5% FDR cut-off, and the percentage of the hub genes enriched GO was based on total fold enrichment. Notably, the 20 GO terms uncovered the dysregulation of cancer immune response, cell cycle, and post-regulation of the cell cycle. The count of genes implicated with the determined biological process driving radioresistance is depicted in a chart ([Fig. 3](#)).

### Survival risk of the hub genes in NSCLC

The KM plotter was employed to determine the overall survival of the hub genes. The survival median and log-rank *P*-value of the KM plot assessed the overall survival between the high and low expression cohorts between the individual hub genes and the overall NSCLC subgroups at 95% confidence. The KM plots explained



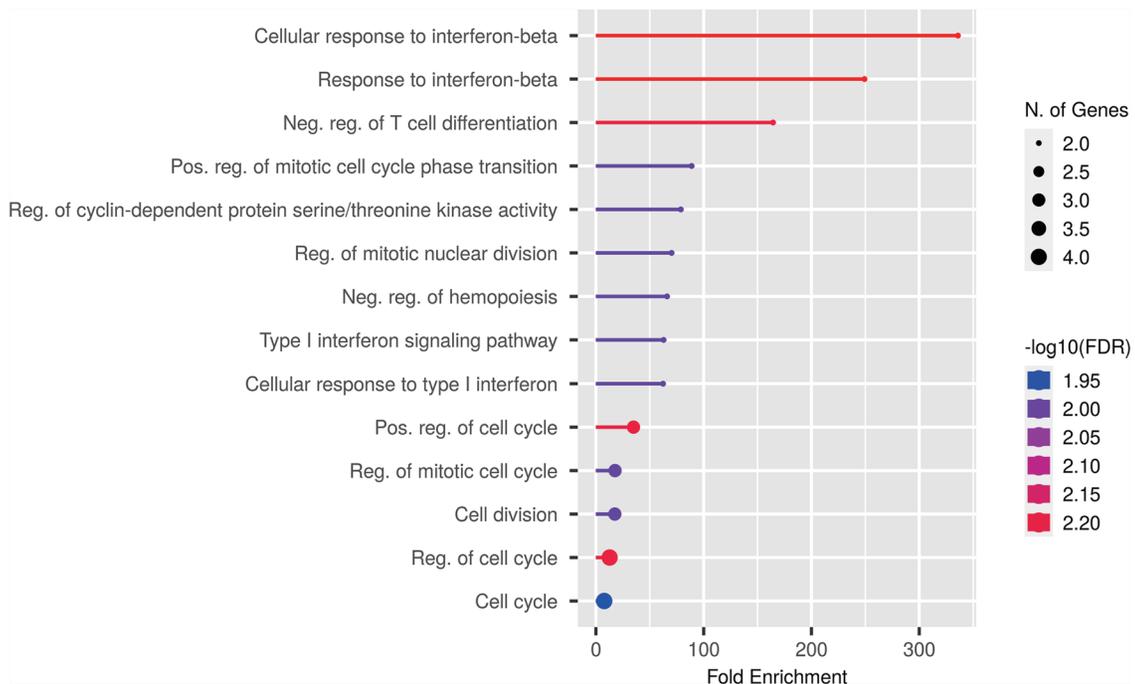
**Fig. 2.** The volcano plot representing the differential gene expression ( $-\log_{10}(\text{p-value})$  vs  $\log_2\text{FC}$ ) with an adjusted P value of  $<0.05$ . (a) Differential gene expression of sensitive (Control) and radioresistant (Test) A549 cells. (b) Significant mRNAs (Control vs Test). (c) Significant lncRNAs (Control vs Test).

that the high expression cohorts drive NSCLC pathogenesis with a Hazard ratio  $>1$  (Supplementary Fig. 4, Supplementary Table 3).

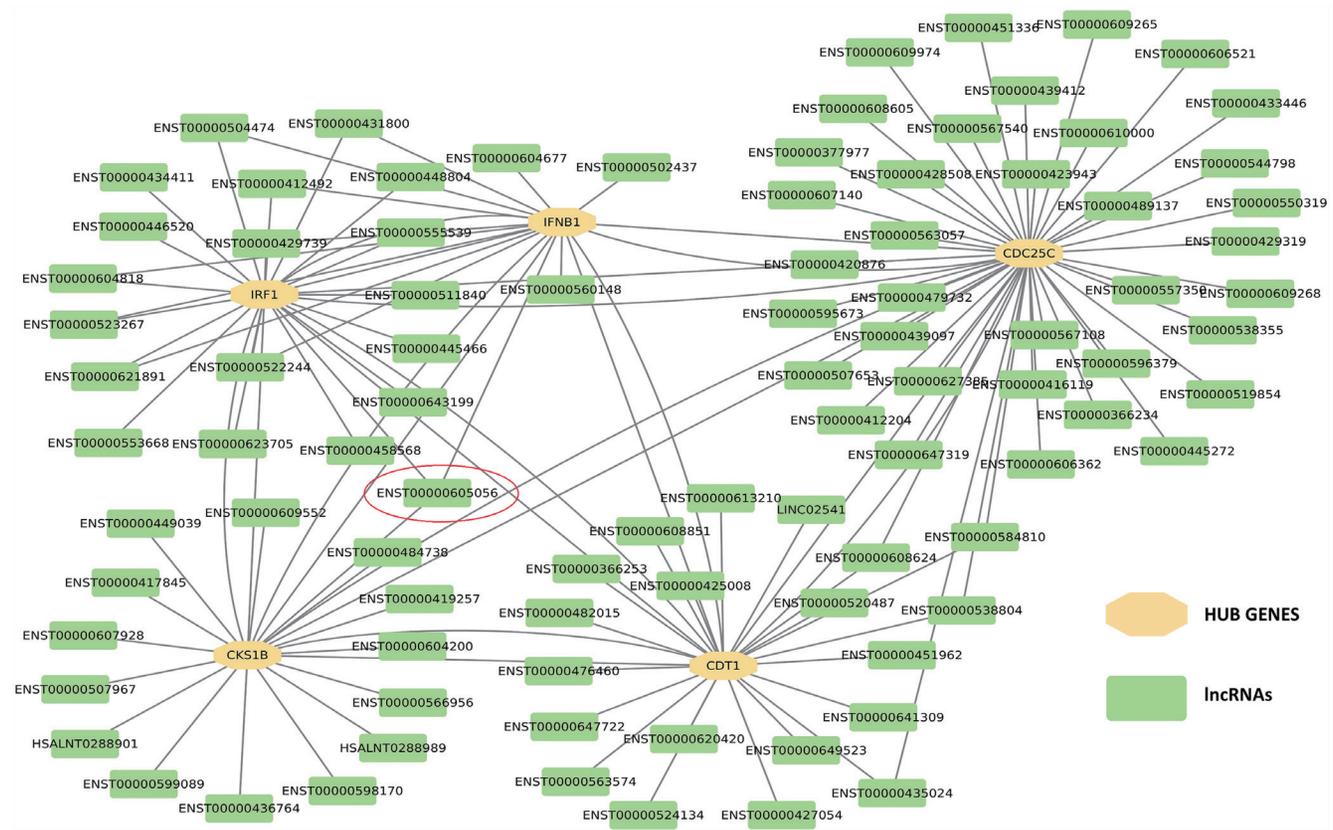
**Hub gene-lncRNA co-expression network**

The co-expression network construction was employed to investigate the functions of lncRNAs. The Pearson correlation coefficient,

a co-expression metric, determined the correlation between the hub genes and lncRNAs based on average expression (Fig. 4). Our investigation revealed that ENST00000605056 lncRNA was upregulated and exhibited the maximum interactions with three highly ranked hub genes (CKS1B, IRF1, and IFNB1), rendering it the most active lncRNA. However, few other lncRNAs (ENST00000608624,



**Fig. 3.** The number of hub genes involved with the determined GO and biological processes at 5% FDR. FDR, false discovery rate; GO, gene ontology.



**Fig. 4.** The mRNA-lncRNA co-expression determined an active lncRNA (ENST00000605056) involved in regulating three highly ranked hub genes. The yellow node represents hub genes, and the green represents the correlated lncRNAs.

ENST00000538804, ENST00000435024, and ENST00000584810 possessed interaction with two hub genes (CDC25C and CDT1). Moreover, 7 lncRNAs (ENST00000504474, ENST00000412492, ENST00000431800, ENST00000448804, ENST00000604818, ENST00000523267, and ENST00000621891) were correlated with IRF1 and IFN1.

**Structure prediction of the lncRNA**

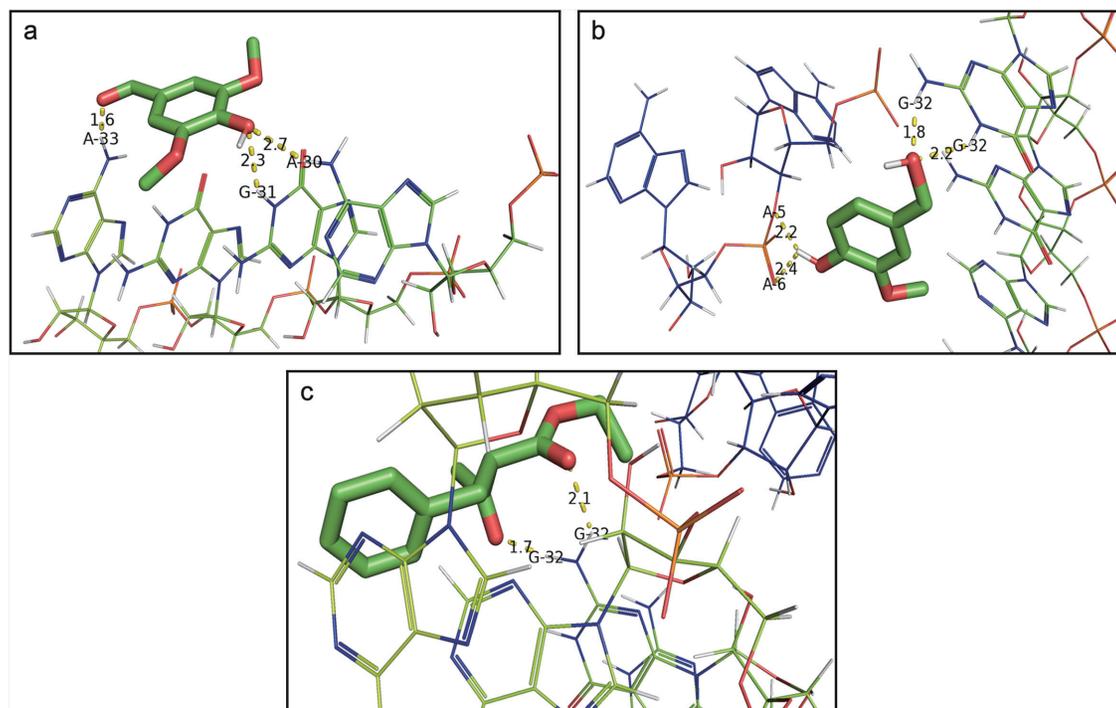
The obtained dot-bracket annotations from RNAfold were used to predict the three-dimensional structure of the ENST00000605056 on the RNA COMPOSER webserver (Supplementary Fig. 5). The secondary structure, folding energy, dot-bracket notations and chromosome coordinates were noted (Table 1).

**lncRNA-ligand docking**

The potential binding mode of interaction between lncRNA and the volatile ligands was studied employing knowledge-based RNA-Ligand docking using the HADDOCK webserver. The docking findings resulted in the generation of 5 clusters, each including a wide range of structures that depicted 95.5% of water-refined models. A weighted average of the van der Waals intermolecular energy, electrostatic intermolecular energy, desolvation energy, restraints violation energy, and buried surface area is used to calculate the HADDOCK score. The lower the HADDOCK scores, the stronger the complex’s binding affinity. The number of standard deviations from the mean at which a specific cluster is positioned was represented by a Z-score. Negative Z-scores are thought to

**Table 1.** The predicted secondary structure of the identified lncRNA with dot-bracket notations and its chromosome number and coordinates

Identified lncRNA	2° Structure of lncRNA and MFE	Dot bracket	Chromosome number and coordinates
ENST00000605056/AC104695.2		.....(((((((.....)))))).....	Chr2:28425945-28426719[+]



**Fig. 5.** Molecular interactions of the best-docked poses of lncRNA and the selected volatile ligands (a) 8655 (b) 62348 (c) 246929.

illustrate a strong HADDOCK cluster better. Based on the Z score and HADDOCK score, the best cluster for each complex was considered, along with other intermolecular energies and restraints (Supplementary Table 4). The best cluster possessed the top 4 poses for each of the complexes of our study and was visualised with PyMOL2 (V2.5.2), and non-covalent interactions at ( $<3.5 \text{ \AA}$ ) were considered. The best pose of the 8655-lncRNA complex possessed three hydrogen bonds between the nucleotides (A33, G31, and A30) and the ligand, possessing a HADDOCK score of  $-21.1 \pm 1.7$ . Similarly, the 62348-lncRNA complex was involved in three hydrogen bonds at G32, A5, and A6 with a score of  $-17.0 \pm 1.7$ . However, the 246929-lncRNA complex possessed two H bonds at G32 with the highest score of  $-22.6 \pm 2.4$ . The H bond distance between the ligands and the nucleotides was acceptable. Moreover, our docking analysis uncovered the Van der Waals interactions between the ligands and the lncRNA (Fig. 5, Table 2).

#### Binding stability of the complexes

The MD simulation assessed the binding stability of the best-docked lncRNA-ligand complexes. The ligand-free lncRNA served as the control for the simulation. The RMSD calculations were used to calculate the complex's deviation during the

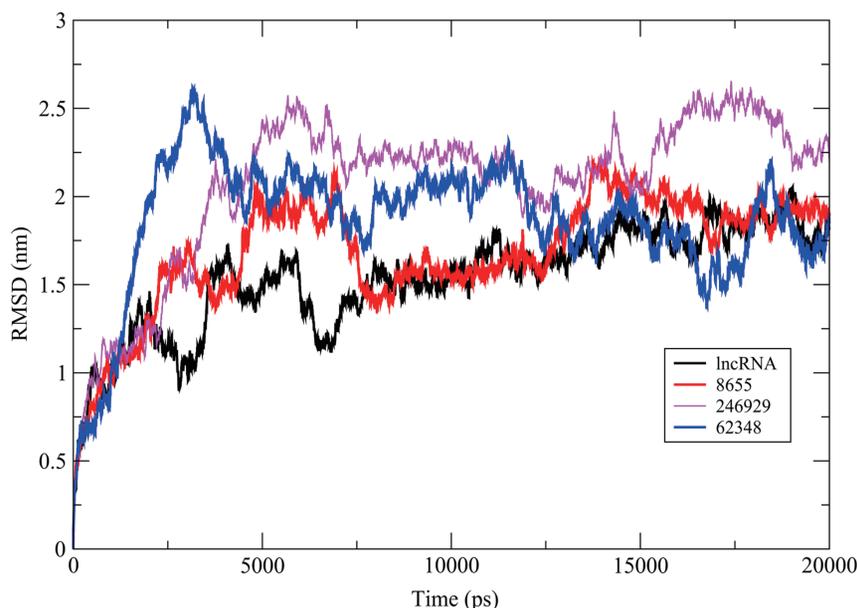
20 ns trajectory period with a snapshot of the atomic coordinates every 1 picosecond (ps). In contrast to the control, the average RMSD value of 1.63 nm, 1.7 nm for 8655-lncRNA, 1.86 nm for 62348-lncRNA, and 2.08 nm for 246929-lncRNA was determined, respectively (Fig. 6).

#### Discussion

The study involved identifying and targeting a lncRNA, dysregulating the multiple biological processes driving the radioresistance in NSCLC cells. Since nucleic acid-based techniques often entail big, generally highly charged molecules and pose delivery challenges, small-molecule therapeutics targeting RNA are frequently favoured.<sup>32</sup> The authors investigated the RNA microarray dataset and estimated gene expression based on  $\log_2FC$  among radioresistant and sensitive A549 cells. The experimental investigators of the dataset claim 529 DEGs, enriched to impaired immune response and TNF signalling. However, our statistical analysis revealed the presence of 589 DE mRNAs and 496 DE lncRNAs. The DE mRNAs were further leveraged for the downstream analysis, involving the physical and functional relationship through a static network-based approach.<sup>33</sup> A noteworthy cluster was identified

**Table 2.** The HADDOCK score and non-covalent bond interactions ( $<3.5 \text{ \AA}$ ) of the best-docked complexes of the lncRNA and the volatile ligands

Complexes	HADDOCK Score	H-bonds and distance ( $\text{\AA}$ ) between lncRNA-ligand complex	Nucleotides participating in Van der Waals interactions
lncRNA-8655	$-21.1 \pm 1.7$	A33 (H62)-O4 (1.6); G31 (H1)-O3 (2.3); A30 (H62)-O3 (2.7)	A30
lncRNA-62348	$-17.0 \pm 1.7$	G32(H1)-O3 (1.8); G32(H22)-O3 (2.2); A5(O3)-O2HAA (2.2); A6(O1)-O2HAA (2.4)	A34 and A33
lncRNA-246929	$-22.6 \pm 2.4$	G32(H22)-O2 (1.7); G32(H21)-O3 (2.1)	A33



**Fig. 6.** The binding stability of the complexes and free lncRNA as control was assessed during a 20 ns molecular dynamics simulation. RMSD, Root Mean Square Deviation.

based on densely interconnected regions within the network by important genes, their relationships, and their significance in regulating biological processes. The  $\log_2$ FC-based retrieved hub genes explain the dysregulation of the cell cycle and evasion of cancer immune responses.

Radiation resistance is associated with immune escape, involving various established mechanisms for evasion of cancer immune responses, such as cytokine release, silencing T cell activation, and other alterations.<sup>6</sup> In our study, IFNB1, which involved autophagy and immunomodulatory function, was downregulated.<sup>34</sup> The inherent radiosensitivity of cell subpopulations prevalent in low- and high-radiosensitive subsets has been altered. This distinction relies on hypoxia levels, DNA repair ability, the number of dividing and apoptotic cells, and cell cycle stages. Cell cycle regulation plays a significant part in this process.<sup>35</sup> CKS1B, and CDC25C are the highly ranked hub genes identified in our study that are crucial in impaired cell cycle regulation and drug resistance, reflecting our results on their expression levels in radioresistance.<sup>36,37</sup> Also, overexpression of CDT1 triggers aberrant cell replication activates DNA damage checkpoints and promotes numerous malignancies.<sup>38</sup> Furthermore, IRF1, a tumour suppressor gene, was downregulated, indicating apoptosis evasion and possibly negatively regulated by the Wnt/ $\beta$ -catenin signalling pathway.<sup>39</sup>

It is critical to select biomarkers that may predict therapeutic response and clinical outcomes using Kaplan-Meier plots.<sup>40</sup> The survival risk of expression cohorts in the NSCLC patients was explained for the hub genes for 200 months. The lower median of high expression cohorts of the hub genes (CKS1B, CDC25C, and CDT1) revealed their role in impaired cell cycle regulation, increasing the survival risk in NSCLC patients with a hazard ratio greater than 1, which indicates greater risk.<sup>41</sup> The tumour suppressor IRF1 possessed a lower median for the high-expression cohort. However, the study by Xu *et al.*<sup>42</sup> revealed that overexpression of IRF1 hindered the development and enhanced radiosensitivity in cancer cells, probably by modulating interferon-induced proteins, which is parallel to the finding.

This study uncovered a co-expression network to acquire a deeper insight into the possible relationships and related processes of the co-expressed lncRNAs and hub genes. The Pearson correlation coefficient was used to determine the co-expression between the hub genes and DE lncRNAs. The Pearson correlation yields an absolute number between 0 and 1. The correlation's absolute value, closer to 1, was claimed to be significantly co-expressed by the pair of genes (hub gene-lncRNA).<sup>43</sup> The co-expression network uncovered a novel lncRNA (ENST00000605056) regulating three highly ranked hub genes, and the authors claim it to be active and involved in driving radioresistance.<sup>12</sup>

The lncRNA structure was predicted using the most frequently employed method of thermodynamic models. RNAfold leveraged the well-known Zuker algorithm to efficiently predict an optimal secondary structure with the minimum free energy.<sup>44</sup> The dot-bracket notations encoded the topology of the lncRNA secondary structure and aided in exploiting its tertiary structure.<sup>28</sup> The predicted three-dimensional structure of the lncRNA opened an avenue to study its molecular interaction with the selected volatile ligands.

The plant volatiles possess low molecular weight and high vapour pressure<sup>45</sup> and diverse scaffolds and are deemed potential candidates for lead development.<sup>46,47</sup> However, poor pharmacokinetics is frequently a stumbling block during drug development.<sup>48</sup> Therefore, the pharmacokinetic properties of the ligands were determined (Supplementary Fig. 6, Supplementary Table 5). The selected ligands satisfied the Rule of Five, indicating favourable drug-like properties and are bioavailable.<sup>49</sup> The molecular interaction of the ligands with the lncRNA revealed the preponderance of H bonds, Van der Waals, and pi-stacking, promoting high binding affinity and stability.<sup>31</sup> Moreover, the HADDOCK cluster revealed an active site at 5–40 nucleotides, possessing the best poses of the complex (ligand-lncRNA) with high binding affinity.<sup>14</sup>

The molecular interaction analysis of the complexes was further extended to validate the stability of the complexes and facilitate enhanced pharmacokinetic effects to inhibit the dysregulated lncR-

NA. The MD simulation assessed the stability of the complexes to understand the structural and conformational changes. The conformations and stability were assessed by comparing the distance between the ligand-bound nucleotide atoms and the lncRNA's unbound states. Therefore, we employed RMSD calculations to determine the distance between nucleotide atoms. An MD simulation study of the miRNA-ligand complex by Junaid *et al.*<sup>31</sup> revealed the average RMSD > 2 for a trajectory period of 10 ns. However, our findings indicated that the complexes, in contrast to the control, acquired conformational stability in the 20 ns trajectory period. The RMSD of the ligand 8655 and 62346 was <2. The ligand 8655 possessed a similar trajectory of structural deviation compared to the control, demonstrating it to be the potential lead for targeting the dysregulated lncRNA, inducing radioresistance in NSCLC patients.

### Conclusion

The study aimed to identify and target the lncRNA in inducing radiation resistance in NSCLC cells using selected small volatile ligands with favourable pharmacokinetic attributes. The co-expression network revealed the ENST00000605056 lncRNA to regulate clinically significant genes CKS1B, IRF1 and IFNB1 involved in NSCLC pathogenesis. The molecular interaction study demonstrated a strong binding association of the selected volatile ligands with the lncRNA with a preponderance of non-covalent bonds, promoting binding stability. Moreover, the MD simulation determined the dynamic behaviour and stability of the complexes, and the lead 8655 possessed lower RMSD than the other selected ligands, indicating it to be the potential lead in targeting dysregulated lncRNA. This study filled the gap in understanding the drugging of a small molecule with lncRNAs triggering radioresistance with an impaired cell cycle regulation, immune response and apoptosis evasion using an *in-silico* approach. Furthermore, upon experimental confirmation, this strategy would assist in developing small molecules to target specific lncRNAs in the NSCLC aetiology with improved pharmacokinetic attributes.

### Supporting information

Supplementary material for this article is available at <https://doi.org/10.14218/GE.2023.00005S>.

**Supplementary Table 1.** The bioinformatics tools employed in this study and their principles and applications are listed.

**Supplementary Table 2.** The significant results obtained in the study are summarized.

**Supplementary Table 3.** The median of the expression cohorts of the hub genes determines the overall survival risk for NSCLC patients.

**Supplementary Table 4.** The best clusters obtained from the complexes obtained from the HADDOCK with associated intermolecular interaction energies.

**Supplementary Table 5.** The drug-likeness of the selected volatile ligands based on the rule of five assessed using SwissADME.

**Supplementary Fig. 1.** The outcome of this study is presented in a flowchart.

**Supplementary Fig. 2.** The interaction of the DE mRNAs dataset is represented using a PPI network.

**Supplementary Fig. 3.** (a) The top-scored cluster (11.273) was identified from the PPI network using the MCODE plugin of Cytoscape. (b) The top 5 ranked hub genes were determined using the maximum clique centrality (MCC) technique.

**Supplementary Fig. 4.** The Kaplan Meier plots of the hub genes represent their clinical significance based on the median survival of NSCLC patients (n = 719) for 200 months.

**Supplementary Fig. 5.** The tertiary structure of the ENST00000605056 lncRNA was predicted using RNA COMPOSER.

**Supplementary Fig. 6.** The radar plot represents the predicted pharmacokinetic properties of the selected volatile ligands.

### Acknowledgments

The authors thank the organizers and the management of CompBio22 for providing the Jetstream2 HPC facility to carry out this study.

### Funding

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

### Conflict of interest

The authors have no conflict of interests to declare.

### Author contributions

AM: Conceptualization; Methodology; Software; Data curation; Formal analysis; Investigation; Validation; Visualization; Writing-original draft; Writing-review and editing. MKS: Conceptualization; Methodology; Supervision; Writing-review and editing.

### Data sharing

The dataset is publicly available at <https://www.ncbi.nlm.nih.gov/geo/> (GSE197236).

### References

- [1] Thandra KC, Barsouk A, Saginala K, Aluru JS, Barsouk A. Epidemiology of lung cancer. *Contemp Oncol* 2021;25(1):45. doi:10.5114/WO.2021.103829, PMID:33911981.
- [2] Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics. *CA Cancer J Clin* 2022;72(1):7–33. doi:10.3322/CAAC.21708, PMID:35020204.
- [3] Chalela R, Curull V, Enríquez C, Pijuan L, Bellosillo B, Gea J. Lung adenocarcinoma: from molecular basis to genome-guided therapy and immunotherapy. *J Thorac Dis* 2017;9(7):2142. doi:10.21037/JTD.2017.06.20, PMID:28840016.
- [4] Zhang H, Zhou F, Wang Y, *et al.* Eliminating Radiation Resistance of Non-Small Cell Lung Cancer by Dihydroartemisinin Through Abrogating Immunity Escaping and Promoting Radiation Sensitivity by Inhibiting PD-L1 Expression. *Front Oncol* 2020;10:1–14. doi:10.3389/fonc.2020.595466, PMID:33194761.
- [5] Leiser D, Samanta S, Eley J, *et al.* Role of caveolin-1 as a biomarker for radiation resistance and tumor aggression in lung cancer. *PLoS One*

- 2021;16:1–20. doi:10.1371/journal.pone.0258951, PMID:34762666.
- [6] Xu LM, Yu H, Yuan YJ, *et al.* Overcoming of Radioresistance in Non-small Cell Lung Cancer by microRNA-320a Through HIF1 $\alpha$ -Suppression Mediated Methylation of PTEN. *Front Cell Dev Biol* 2020;8:912. doi:10.3389/FCCELL.2020.553733, PMID:33304897.
  - [7] Wu W, Zhang S, He J. The Mechanism of Long Non-coding RNA in Cancer Radioresistance/Radiosensitivity: A Systematic Review. *Front Pharmacol* 2022;13:1–12. doi:10.3389/fphar.2022.879704, PMID:35600868.
  - [8] Behzadi P, Issakhanian L, Behzadi P, Issakhanian L. Introductory Chapter: Gene Regulation, an RNA Network-Dependent Architecture. In: Behzadi P (ed). *Gene Regul. Rijeka: IntechOpen; 2019.* doi:10.5772/INTECHOPEN.84535.
  - [9] Fang K, Caixia H, Xiufen Z, Zijian G, Li L. Screening of a Novel Up-regulated lncRNA, A2M-AS1, That Promotes Invasion and Migration and Signifies Poor Prognosis in Breast Cancer. *Biomed Res Int* 2020;20:9747826. doi:10.1155/2020/9747826, PMID:32352014.
  - [10] Xu S, Kong D, Chen Q, Ping Y, Pang D. Oncogenic long noncoding RNA landscape in breast cancer. *Mol Cancer* 2017;16(1):1–15. doi:10.1186/S12943-017-0696-6, PMID:28738804.
  - [11] Lin P, Xing W, Ren Q, Wang Q, Yan J, Mao G. LncRNAs as Theragnostic Biomarkers for Predicting Radioresistance in Cancer: A Systematic Review and Meta-Analysis. *Front Oncol* 2022;12:767750. doi:10.3389/FONC.2022.767750, PMID:35692742.
  - [12] Jiang L, Hong L, Yang W, Zhao Y, Tan A, Li Y. Co-expression network analysis of the lncRNAs and mRNAs associated with cervical cancer progression. *Arch Med Sci* 2019;15(3):754. doi:10.5114/AOMS.2019.84740, PMID:31110543.
  - [13] Pathmanathan S, Grozavu I, Lyakisheva A, Stagliar I. Drugging the undruggable proteins in cancer: A systems biology approach. *Curr Opin Chem Biol* 2022;66:102079. doi:10.1016/J.CBPA.2021.07.004, PMID:34426091.
  - [14] Ren Y, Wang YF, Zhang J, *et al.* Targeted design and identification of AC1NOD4Q to block activity of HOTAIR by abrogating the scaffold interaction with EZH2. *Clin Epigenetics* 2019;11(1):1–16. doi:10.1186/S13148-019-0624-2, PMID:30764859.
  - [15] Jin XX, Mei YN, Shen Z, Zhu JF, Xing SH, Yang HM, *et al.* A chalcone-syringaldehyde hybrid inhibits triple-negative breast cancer cell proliferation and migration by inhibiting CKAP2-mediated FAK and STAT3 phosphorylation. *Phytomedicine* 2022;101:154087. doi:10.1016/J.PHYMED.2022.154087, PMID:35429924.
  - [16] Jung HJ, Song YS, Lim CJ, Park EH. Anti-angiogenic, anti-inflammatory and anti-nociceptive activities of vanillyl alcohol. *Arch Pharmacol Res* 2008 31(10) 2008;31(10):1275–1279. doi:10.1007/S12272-001-2106-1, PMID:18958417.
  - [17] Kim IS, Choi DK, Jung HJ. Neuroprotective effects of vanillyl alcohol in *Gastrodia elata* Blume through suppression of oxidative stress and anti-apoptotic activity in toxin-induced dopaminergic MN9D cells. *Molecules* 2011;16(7):5349–5361. doi:10.3390/MOLECULES16075349, PMID:21705974.
  - [18] Cihan-Üstündağ G, Çapan G. Synthesis and evaluation of functionalized indoles as antimycobacterial and anticancer agents. *Mol Divers* 2012 16(3) 2012;16(3):525–539. doi:10.1007/S11030-012-9385-Y, PMID:22893206.
  - [19] Behzadi P, Ranjbar R. DNA microarray technology and bioinformatic web services. *Acta Microbiol Immunol Hung* 2019;66(1):19–30. doi:10.1556/O30.65.2018.028, PMID:30010394.
  - [20] Han BA, Yang XP, Hosseini DK, *et al.* Identification of candidate aberrantly methylated and differentially expressed genes in Esophageal squamous cell carcinoma. *Sci Reports* 2020 10(1) 2020;10(1):1–13. doi:10.1038/s41598-020-66847-4, PMID:32546690.
  - [21] Cai Y, Mei J, Xiao Z, *et al.* Identification of five hub genes as monitoring biomarkers for breast cancer metastasis in silico. *Hereditas* 2019;156(1):20. doi:10.1186/S41065-019-0096-6, PMID:31285741.
  - [22] Li T, Gao X, Han L, Yu J, Li H. Identification of hub genes with prognostic values in gastric cancer by bioinformatics analysis. *World J Surg Oncol* 2018;16(1):1–12. doi:10.1186/S12957-018-1409-3, PMID:29921304.
  - [23] Sekaran TSG, Kedilaya VR, Kumari SN, Shetty P, Gollapalli P. Exploring the differentially expressed genes in human lymphocytes upon response to ionizing radiation: a network biology approach. *Radiat Oncol J* 2021;39(1):48. doi:10.3857/ROJ.2021.00045, PMID:33794574.
  - [24] Liu K, Kang M, Li J, Qin W, Wang R. Prognostic value of the mRNA expression of members of the HSP90 family in non-small cell lung cancer. *Exp Ther Med* 2019;17(4):2657. doi:10.3892/ETM.2019.7228, PMID:309330968.
  - [25] Wei W, Amberkar S, Hide W. Differential co-expression analysis. R package version 1.18.0; 2022. doi: 10.18129/B9.bioc.diffcoex.
  - [26] Akbari F, Peymani M, Salehzadeh A, Ghaedi K. Integrative in silico and in vitro transcriptomics analysis revealed new lncRNAs related to intrinsic apoptotic genes in colorectal cancer. *Cancer Cell Int* 2020;20(1):1–16. doi:10.1186/S12935-020-01633-W, PMID:33292233.
  - [27] Chao Y, Jin J, Wang L, Jin X, Yang L, Zhang B. Transcriptome Analysis of lncRNA–mRNA Interactions in Chronic Atrophic Gastritis. *Front Genet* 2021;11:1704. doi:10.3389/FGENE.2020.612951, PMID:33505433.
  - [28] Biesiada M, Purzycka KJ, Szachniuk M, Blazewicz J, Adamiak RW. Automated RNA 3D Structure Prediction using RNAComposer. *Methods Mol Biol* 2016;1490:199–215. doi:10.1007/978-1-4939-6433-8\_13, PMID:27665601.
  - [29] Parate S, Rampogu S, Lee G, Hong JC, Lee KW. Exploring the Binding Interaction of Raf Kinase Inhibitory Protein With the N-Terminal of C-Raf Through Molecular Docking and Molecular Dynamics Simulation. *Front Mol Biosci* 2021;8:496. doi:10.3389/FMOLB.2021.655035, PMID:34124147.
  - [30] Manigrasso J, Marcia M, De Vivo M. Computer-aided design of RNA-targeted small molecules: A growing need in drug discovery. *Chem* 2021;7(11):2965–2988. doi:10.1016/J.CHEMPR.2021.05.021.
  - [31] Junaid M, Dash R, Islam N, *et al.* Molecular Simulation Studies of 3,3'-Diindolylmethane as a Potent MicroRNA-21 Antagonist. *J Pharm Bioallied Sci* 2017;9(4):259. doi:10.4103/JPBS.JPBS\_266\_16, PMID:29456377.
  - [32] Warner KD, Hajdin CE, Weeks KM. Principles for targeting RNA with drug-like small molecules. *Nat Rev Drug Discov* 2018 17(8):547–558. doi:10.1038/nrd.2018.93, PMID:29977051.
  - [33] Fionda V. Networks in Biology. *Encycl Bioinforma Comput Biol ABC Bioinforma* 2019;1-3:915–921. doi:10.1016/B978-0-12-809633-8.20420-2.
  - [34] Ambjørn M, Ejlerskov P, Liu Y, Lees M, Jättelä M, Issazadeh-Navikas S. IFN $\beta$ /interferon- $\beta$ -induced autophagy in MCF-7 breast cancer cells counteracts its proapoptotic function. *Autophagy* 2013;9(3):287. doi:10.4161/AUTO.22831, PMID:23221969.
  - [35] Cao R, Ding Q, Li P, *et al.* SHP1-mediated cell cycle redistribution inhibits radiosensitivity of non-small cell lung cancer. *Radiat Oncol* 2013;8(1):1–9. doi:10.1186/1748-717X-8-178, PMID:23842094.
  - [36] Shi W, Huang Q, Xie J, Wang H, Yu X, Zhou Y. CKS1B as Drug Resistance-Inducing Gene—A Potential Target to Improve Cancer Therapy. *Front Oncol* 2020;10:582451. doi:10.3389/FONC.2020.582451, PMID:33102238.
  - [37] Liu K, Zheng M, Lu R, *et al.* The role of CDC25C in cell cycle regulation and clinical cancer therapy: a systematic review. *Cancer Cell Int* 2020 20(1) 2020;20(1):1–16. doi:10.1186/S12935-020-01304-W, PMID:32518522.
  - [38] Cai C, Zhang Y, Hu X, *et al.* CDT1 Is a Novel Prognostic and Predictive Biomarkers for Hepatocellular Carcinoma. *Front Oncol* 2021;11:3803. doi:10.3389/FONC.2021.721644, PMID:34631549.
  - [39] Ohsugi T, Yamaguchi K, Zhu C, *et al.* Anti-apoptotic effect by the suppression of IRF1 as a downstream of Wnt/ $\beta$ -catenin signaling in colorectal cancer cells. *Oncogene* 2019 38(32) 2019;38(32):6051–6064. doi:10.1038/s41388-019-0856-9, PMID:31292489.
  - [40] Zhu C, Menyhart O, Gyorffy B, He X. The prognostic association of SPAG5 gene expression in breast cancer patients with systematic therapy. *BMC Cancer* 2019;19(1):1–12. doi:10.1186/S12885-019-6260-6, PMID:31690268.
  - [41] Rich JT, Neely JG, Paniello RC, Voelker CCJ, Nussenbaum B, Wang EW. A PRACTICAL GUIDE TO UNDERSTANDING KAPLAN-MEIER CURVES. *Otolaryngol Head Neck Surg* 2010;143(3):331. doi:10.1016/J.OTOHNS.2010.05.007, PMID:20723767.
  - [42] Xu X, Wu Y, Yi K, Hu Y, Ding W, Xing C. IRF1 regulates the progression of colorectal cancer via interferon-induced proteins. *Int J Mol Med* 2021;47(6):104. doi:10.3892/IJMM.2021.4937, PMID:33907823.
  - [43] Furlotte NA, Kang HM, Ye C, Eskin E. Mixed-model coexpression: calculating gene coexpression while accounting for expression heterogeneity. *Bioinformatics* 2011;27(13):i288. doi:10.1093/BIOINFOR-

- MATICS/BTR221, PMID:21685083.
- [44] Sato K, Akiyama M, Sakakibara Y. RNA secondary structure prediction using deep learning with thermodynamic integration. *Nat Commun* 2021;12(1):1–9. doi:10.1038/s41467-021-21194-4, PMID:33574226.
- [45] Duan X, Li T, Han X, *et al*. The antitumor effect of arsenic trioxide on hepatocellular carcinoma is enhanced by andrographolide. *Oncotarget* 2017;8(53):90905–90915. doi:10.18632/ONCOTARGET.18677, PMID:29207612.
- [46] Dhifi W, Bellili S, Jazi S, Bahloul N, Mnif W. Essential Oils' Chemical Characterization and Investigation of Some Biological Activities: A Critical Review. *Medicines* 2016;3(4):25. doi:10.3390/MEDICINES3040025, PMID:28930135.
- [47] Ríos JL. Essential Oils: What They Are and How the Terms Are Used and Defined. In: Preedy VR (ed). *Essent Oils Food Preserv Flavor Saf*. San Diego: Academic Press; 2016:3–10. doi:10.1016/B978-0-12-416641-7.00001-8.
- [48] Dong J, Wang NN, Yao ZJ, *et al*. ADMETlab: a platform for systematic ADMET evaluation based on a comprehensively collected ADMET database. *J Cheminformatics* 2018;10(1):1–11. doi:10.1186/S13321-018-0283-X, PMID:29943074.
- [49] Abdulrahman HL, Uzairu A, Uba S. Computational pharmacokinetic analysis on some newly designed 2-anilinopyrimidine derivative compounds as anti-triple-negative breast cancer drug compounds. *Bull Natl Res Cent* 2020;44(1):1–8. doi:10.1186/S42269-020-00321-Z.