

P3.084: Integrated Transcriptomic and Systems Modeling Identifies Novel
UHRF1 SRA-Domain Inhibitors in Lung Squamous Cell Carcinoma

Introduction

Lung squamous cell carcinoma (LUSC) remains a clinical challenge due to its highly heterogeneous genomic landscape and lack of actionable drivers. This project employs a systems-biology pipeline to identify novel epigenetic inhibitors targeting UHRF1, a regulator in tumor suppressor silencing and immune evasion in LUSC, using a multi-phase computational framework:

- Transcriptomic Profiling:** Identifying LUSC-specific gene expression patterns and dysregulated epigenetic pathways.
- Structure-Based Discovery:** Utilizing molecular docking to target the UHRF1 SRA-domain pocket with high-affinity ligands
- Kinetic Systems Modeling:** Simulating the competitive inhibition of UHRF1 in an intracellular environment to predict therapeutic duration and gene protection efficacy.

Background

Why LUSC Needs New Targets?

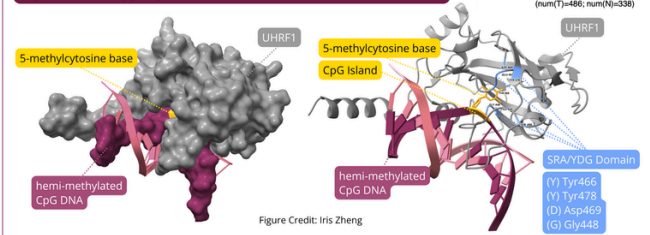
Traditional epigenetic agents (e.g., Azacitidine, Tazemetostat) function as non-selective enzyme traps. By locking enzymes onto DNA, they trigger systemic DNA-damage responses and hematological toxicity, limiting their clinical utility.

Why UHRF1?

(Ubiquitin-like with PHD and RING finger domains 1)

UHRF1 is a modular protein that recruits DNMT1 to maintain DNA methylation patterns by recognizing hemi-methylated CpG sites via its SRA domain (Fig. 1). In LUSC, UHRF1 is frequently dysregulated (Fig. 2), leading to the epigenetic silencing of essential tumor suppressor genes (TSGs). Targeting UHRF1 offers a dual-action path to selectively reactivate these genes while minimizing systemic toxicity and potentially restoring immune visibility.

Fig. 1 Base-Flipping Mechanism in SRA Domain of UHRF1



Research Questions

- Which genes and pathways are differentially expressed in paired LUSC vs. adjacent non-cancerous tissues?
- Can we computationally design selective inhibitors that target the UHRF1 SRA-domain to reactivate tumor suppressor genes and improve therapeutic outcomes?

Predictive Framework & Hypothesis

Phase 1 — Differential Expression & Pathway Analysis

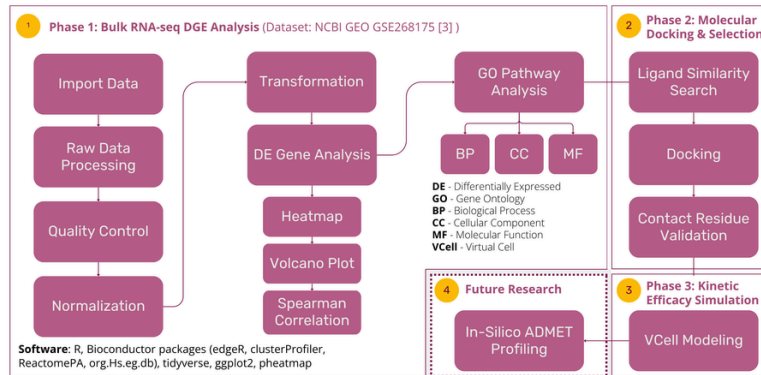
UHRF1 and associated epigenetic regulators will show significant up-regulation in LUSC compared to matched healthy tissue, with pathway enrichment highlighting dysregulated chromatin organization

Phase 2 — In Silico Ligand Discovery & Interaction Mapping
Ligands structurally analogous to AMSA-2 [2] will exhibit high-affinity binding to the UHRF1 SRA-domain, with predicted contact residues overlapping critical DNA-binding sites

Phase 3 — Kinetic Simulation & Efficacy Prediction

Molecules with optimal docking profiles and residue alignment will demonstrate superior gene protection and extended t_{50} (survival-time proxy) within a simulated intracellular environment in a simplified Virtual Cell (VCell) model.

Methods & Materials



Software: R, Bioconductor packages (edgeR, clusterProfiler, ReactomePA, org.Hs.eg.db), tidyverse, ggplot2, pheatmap

Fig. 4 Heatmap of Top-30 Up-regulated Genes in LUSC Cancer Tissue

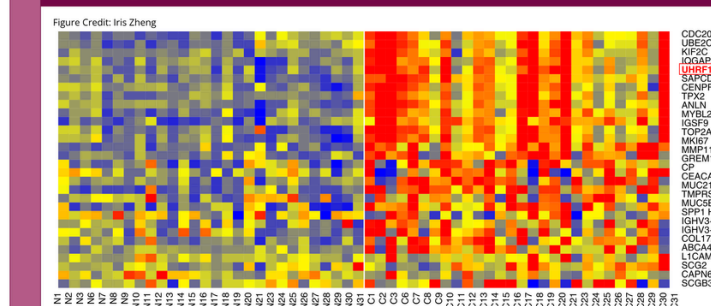


Fig. 4 Hierarchical clustering (n=28) confirms distinct expression profiles between Normal (N) and Cancerous (C) tissue. UHRF1 (ranked 22nd; log₂FC = 3.79, FDR = 1.76E-19) consistently shows significant upregulation, corroborating its role as a key epigenetic driver in LUSC progression.

Fig. 5 Volcano Plot of DEGs

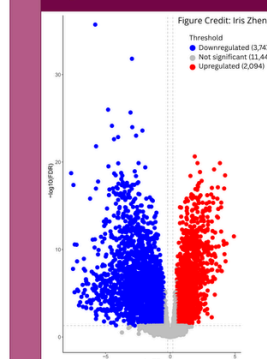
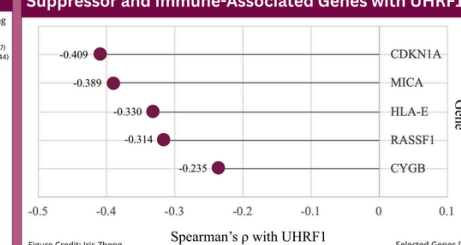


Fig. 5 Volcano plot highlighting 5,841 significant DEGs (FDR < 0.05 & |log₂FC| > 0.5). Red and blue clusters denote upregulated and downregulated genes, respectively, driving LUSC pathology.

Fig. 6 Spearman Correlation of Selected Tumor Suppressor and Immune-Associated Genes with UHRF1



Gene	Function	P
CDKN1A	cell-cycle arrest	-0.409
MICA	stress-induced immune ligand	-0.389
HLA-E	immune recognition modulator	-0.330
RASSF1	TSG; apoptosis	-0.314
CYGB	redox & oxidation stress protection	-0.235

Fig. 7 GO BP Enrichment of Upregulated Genes

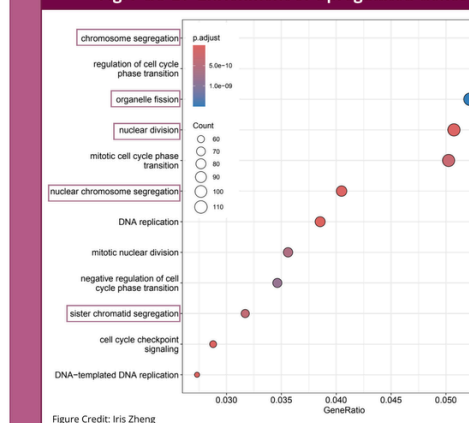


Fig. 7 Top-ranked biological processes associated with UHRF1-driven cancer progression.

Enrichment highlights key pathways, including chromosome segregation and nuclear division, confirming a mechanistic link between UHRF1 overexpression and the dysregulated epigenetic landscape of LUSC.

Ligand Similarity Search

Reference: Known SRA-Domain inhibitor anthraben (AMSA-2) [2]

Fig. 8 AMSA-2 Analog Molecules & Similarity Score		
Anthraben-1,10-diol (0.776)	1,2-Anthracediol (0.702)	1,2-Dihydroxyphenanthrene
Benzo[k]tetraphene-3,4-diol (0.757)	1,2-Dihydroxynaphthalene (0.680)	Chrysene-1,2-diol (0.655)
Benzo[a]anthracene-7-ol (0.737)	7-Methylbenzo[a]anthracene-3,4-diol (0.686)	1,2-Chrysenediol, 5-methyl- (0.655)
Benzo[a]anthracene-3,4-diol (0.722)	Anthracene (0.678)	1,2,4-Naphthalenetriol (0.651)

Fig. 8 A library of 12 small-molecule candidates was selected based on an Arthor similarity threshold of >0.65 to the AMSA-2 scaffold. Scores represent the chemical structural fidelity to the parent compound while exploring chemical space for optimized SRA-domain binding.

Figure Credit: Modified from Pub Chem [4]

Key SRA Domain Residues: Y466, Y478, D469, G448

Fig. 9 Binding Affinity and Residue Interaction Mapping for Lead Candidates across Selected UHRF1 Structural Models

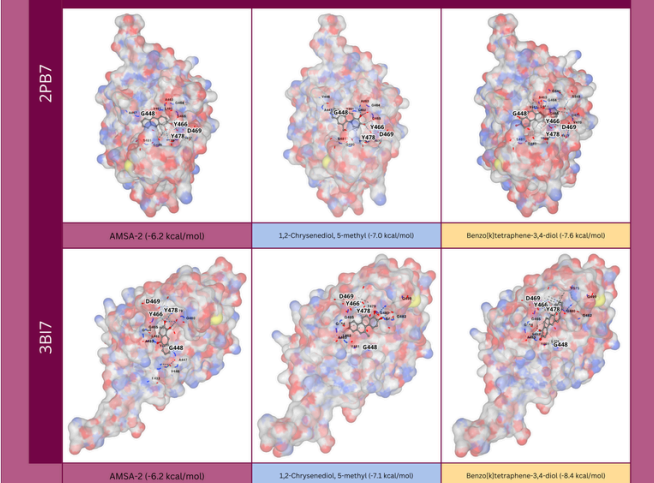


Fig. 9 Binding poses for the reference (AMSA-2) and top-two leads (Benzo[k]tetraphene-3,4-diol and 5-methyl-1,2-chrysenediol) docked into the SRA domain Cavity 1. Models were generated via CB-Dock 2 [5] using two distinct UHRF1 crystal structures (PDB: 2PB7, 3BI7). Candidates were prioritized based on binding affinity (ΔG_{bind}) and the density of stabilizing contacts within the SRA active site.

Fig. 10 Virtual Cell Modeling of Lead Candidates

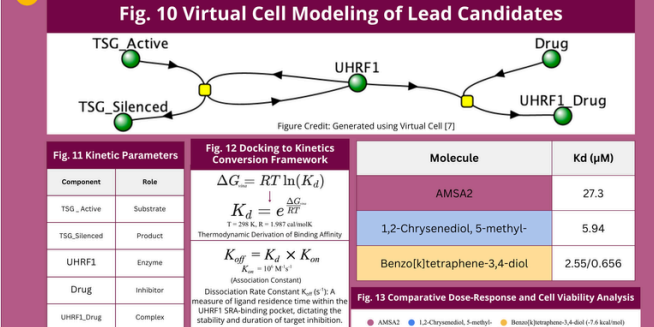
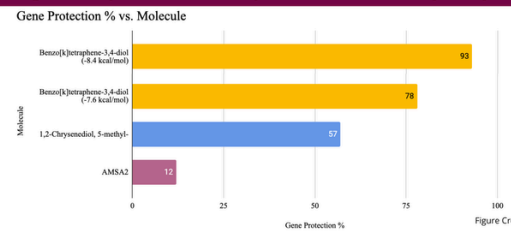


Fig. 13 Survival time proxy (t_{50}) for the lead candidate (Benzo[k]tetraphene-3,4-diol) compared against 1,2-Chrysenediol and the reference compound (AMSA-2). Data demonstrates a dose-dependent reduction in cell viability, with Benzo[k]tetraphene-3,4-diol exhibiting the most potent inhibitory profile.

Results

Fig. 14 Comparative Dose-Response and Cell Viability Analysis



Molecule	Final TSG_Silenced Value (100 sec)
Control	5.0
AMSA2	4.4
1,2-Chrysenediol, 5-methyl-	2.2
Benzo[k]tetraphene-3,4-diol (-7.6 kcal/mol)	1.1
Benzo[k]tetraphene-3,4-diol (-8.4 kcal/mol)	0.33

Fig. 14 Quantification of Tumor Suppressor Gene (TSG) restoration. Percentage of TSG expression recovered in the presence of inhibitor candidates demonstrates that Benzo[k]tetraphene-3,4-diol provides near-total gene rescue at 100 μ M concentration, significantly outperforming current clinical alternatives.

Figure Credit: Iris Zheng

Conclusion

This study bridges human LUSC transcriptomics with structure-based inhibitor screening, identifying UHRF1 as a priority epigenetic target. By validating lead candidates against multiple crystal structures (2PB7, 3BI7), we identified Benzo[k]tetraphene-3,4-diol as a high-affinity SRA-domain antagonist. Our virtual cell model confirms that this molecule achieves superior TSG rescue and cellular viability compared to clinical references, establishing a robust, mechanism-based lead for future in vitro validation.

Future Research/In-Silico ADMET Profiling

Safety & Bioavailability: Using ADMET-AI [7], lead candidate (Benzo[k]tetraphene-3,4-diol) demonstrated favorable oral bioavailability (high Human Intestinal Absorption) and a low predicted clinical toxicity profile.

Fig. 14 ADMET Profiling Credit: Generated using ADMET-AI [7] (Machine learning platform for Absorption, Distribution, Metabolism, Excretion, and Toxicity prediction)

Metabolic Constraints: Secondary candidates (5-methyl-1,2-chrysenediol) were identified as substrates for UGT2A1-mediated glucuronidation (Abdo et al., 2026), indicating a higher risks of rapid hepatic clearance and toxic metabolite formation.

Future Scope: Transition from initial screening to comprehensive safety profiling. This includes *in vitro* hepatic clearance assays and dynamic toxicological modeling to ensure therapeutic candidates exhibit a safe, non-toxic profile prior to clinical consideration.

Limitations & Model Constraints

- Static Structural Limitations:** Crystal structures (PDB: 2PB7, 3BI7) provide static snapshots that may not fully capture the conformational plasticity of the UHRF1 SRA domain.
- Docking Approximations:** Vina scores provide relative binding energies, which may be sensitive to protein protonation states and solvent effects.
- Kinetic Simplification:** VCell parameters are foundational estimates; while suitable for ranking lead candidates, future *in vitro* assays are required to confirm precise kinetic constants (K_d , K_{off}).

Selected References

[1] CEP42 (Gene Expression Interactive Analysis) <http://gepi2.cancer-pku.cn/#index> [2] Cicco S, Mazzoleni V, Javed A, Eller S, Ruff M, Mouzli M, Mori M, Mely Y. Inhibitors of UHRF1 base flipping activity showing cytotoxicity against cancer cells. *Bioorg Chem*. 2023 Aug;137:106616. doi: 10.1016/j.bioorg.2023.106616. Epub 2023 May 22. PMID: 37247964. [3] NCBI GEO: GSE268175. <https://www.ncbi.nlm.nih.gov/geo/> [4] Pub Chem <https://pubchem.ncbi.nlm.nih.gov/> [5] CB-Dock 2 <http://183.56.231.194:8001/cb-dock2/index.php> [6] Chen, Yu, et al. "Effects of MICA Expression on the Progression of Advanced Non-Small Cell Lung Cancer and the Efficacy of CR Therapy." *PLoS ONE*, vol. 8, no. 7, 23 July 2013. pp. e69044-e69044. <https://doi.org/10.1371/journal.pone.0069044>. Accessed 28 Aug. 2023. [7] Virtual Cell: <https://vcell.org/vcell-models> [8] ADMET-AI: <https://admet.ai/geostonebio.com/> Additional references available upon request