turnitin unesa1

325 Production

Project 10

Document Details

Submission ID
trn:oid:::3618:142100657

Submission Date
Jun 8, 2026, 12:09 PM GMT+7

Download Date
Jun 8, 2026, 12:18 PM GMT+7

File Name
similarity erta_10+Layout+Article+176.pdf

File Size
1.8 MB

20 Pages

9,984 Words

59,034 Characters

*% detected as AI

AI detection includes the possibility of false positives. Although some text in this submission is likely AI generated, scores below the 20% threshold are not surfaced because they have a higher likelihood of false positives.

Caution: Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Journal of Current Studies in SDGs
Homepage : https://journal.i-ros.org/index.php/JOCSIS
Email : jocsis@i-ros.org

e-ISSN 3089-7564
Vol.3, No.4, December 2027

Open-Data Serum Metabolomics of Lung Adenocarcinoma for Equitable Early Detection under SDG Target 3.4

Muhammad Nurrohman Sidiq*, Rudiana Agustini, Nuniek Herdyastuti, Prima Retno Wikandari, Mirwa Adiprahara Anggarani

State University of Surabaya, Surabaya, Indonesia

DOI : https://doi.org/10.63230/jocsis.3.4.325

Sections Info

Article history:
Submitted: March 12, 2026
Final Revised: July 8, 2026
Accepted: July 30, 2026
First Available Online: September 5, 2026
Publication Date: December 27, 2027

Keywords:
Machine Learning;
Metabolomics;
SDG Target 3.4.

ABSTRACT

Objective: *Sustainable Development Goal (SDG) target 3.4 addresses premature non-communicable disease mortality, yet lung cancer leads cancer mortality among Indonesian men, and low-dose computed tomography screening presumes imaging capacity most low-income settings lack. This study asked which serum pathways are recoverable from open lung adenocarcinoma metabolomics and whether their enzymes are altered in tumor tissue.* ***Method:*** *Serum GC-TOF MS data from the ST000385 ADC2 cohort (43 cases, 43 controls, 152 compounds) were reanalysed in MetaboAnalyst 6.0 using PLS-DA with permutation testing, multivariate ROC, and KEGG enrichment. Enzymes were examined in TCGA-LUAD (483 tumor, 347 normal) with GEPIA2 and survival survival analysis, and the cohort ST000386 underwent metadata testing.* ***Results:*** *The one-component model separated the groups ($R^2 = 0.492$, $Q^2 = 0.302$, permutation $p < 5 \times 10^{-4}$) and reached an AUC of 0.882 on five metabolites, which a sensitivity analysis showed to rest on one correlated axis. Twenty metabolites passed a 5% FDR, and arginine biosynthesis was the most enriched pathway (FDR = 5.8×10^{-5}), whereas the TCA cycle and taurine metabolism did not survive correction. LDHA, PKM and CDO1 predicted survival. ARG1, OTC and CPS1 medians fell in tumor while ASS1 and ARG2 rose modestly. ST000386 carried subject-level metadata that does not associate with its own metabolome.* ***Novelty:*** *The dominant pathway is arginine biosynthesis rather than the Warburg axis of earlier readings, and the enzyme layer shows a dysregulated rather than silenced urea cycle.*

INTRODUCTION

Sustainable Development Goal (SDG) target 3.4 commits member states to cutting premature mortality from non-communicable diseases by one third by 2030, and cancer is the component of indicator 3.4.1 making the least progress (Cancela et al., 2023; Frieden et al., 2020; Liu et al., 2021). Lung cancer dominates that shortfall, with 2.48 million new cases and 1.8 million deaths estimated for 2022 (Zhou et al., 2024), and it is the leading cause of cancer death among Indonesian men, while the regional burden concentrates rather than eases (Asmara et al., 2023; Dee et al., 2025). Lung Adenocarcinoma (LUAD) is the commonest subtype of Non-Small Cell Lung Cancer (NSCLC) and is heterogeneous in driver mutation and metabolic phenotype (Herbst et al., 2018; Travis et al., 2015). The gap to the SDG trajectory largely stems from late-stage diagnosis, since stage III or IV presentation drives excess mortality in low- and middle-income settings (Hanafi et al., 2024; Henke et al., 2023). Low-Dose Computed Tomography (LDCT) screening lowers lung cancer mortality by approximately 20% in heavy smokers (National Lung Screening Trial Research Team, 2011) but flags most benign findings as positive (Mazzone et al., 2021) and presumes imaging capacity that cannot be assumed where the burden grows fastest, whereas no blood-based marker yet refines risk well enough to deploy (Seijo et al., 2019).

Metabolomics integrates genetic, transcriptional, and environmental input into one phenotypic readout (Wishart, 2019), and Gas Chromatography–Time-Of-Flight Mass Spectrometry (GC-TOF MS) resolves the primary metabolites that shift under oncometabolic adaptation (Fiehn, 2016). Recent lung cancer panels report area-under-curve values of 0.87 to 0.95 while leaving the biochemistry under-interpreted (Li et al., 2022; Liu et al., 2023; Qi et al., 2021). Fahrmann and colleagues deposited two case–control cohorts, ST000385 and ST000386, in the NIH Common Fund Metabolomics Workbench and built classifiers from them (Fahrmann et al., 2015); current workflows can use the same data for pathway-level interpretation (Pang et al., 2024).

Metabolic reprogramming is a hallmark of cancer (Hanahan & Weinberg, 2011) and the Warburg effect still anchors the field (Liberti & Locasale, 2016; Pavlova & Thompson, 2016), with LDHA elevated in most NSCLC tumors (Xie et al., 2014) and the pyruvate kinase M2 dimer–tetramer switch itself regulated in this tumor type (Anastasiou et al., 2011; Christofk et al., 2008; Liang et al., 2024). Nitrogen metabolism has drawn less attention despite carrying several of the strongest lung-specific dependencies. Oncogenic KRAS reroutes glutamine through a cytosolic NADPH-generating cycle (Berardinis & Cheng, 2010; Son et al., 2013), KEAP1/NRF2-mutant LUAD leans on it hardest (Romero et al., 2017; Lu et al., 2024), glutaminase inhibition is in trial (Harding et al., 2021), and KRAS also silences argininosuccinate synthase 1 (ASS1), the enzyme condensing aspartate with citrulline, which diverts aspartate into pyrimidine synthesis and creates arginine auxotrophy (Gai et al., 2024). A serum readout of that rewiring would involve aspartate, glutamate, glutamine, ornithine, and citrulline rather than lactate alone. Loss of a single enzyme is not the usual pattern, however. Lee and colleagues found that most urea cycle enzymes change expression in many tumors, a hallmark named urea cycle dysregulation, which diverts nitrogen toward pyrimidine synthesis and produces measurable shifts in nitrogen metabolites in patient biofluids (Lee et al., 2018). In lung cancer specifically, KRAS and LKB1 co-mutant tumors express carbamoyl phosphate synthetase 1 and route ammonia nitrogen into pyrimidines (Kim et al., 2017), arginase activity in tumor myeloid cells depletes arginine and suppresses T cells (Miret et al., 2019), and arginine availability governs nitric oxide synthesis (Keshet & Erez, 2018), while ASS1 expression itself stratifies response to arginine deprivation (Barnett et al., 2023). Tumors also remodel the Tricarboxylic Acid (TCA) cycle (Metallo et al., 2012; Sciacovelli & Frezza, 2016). A third axis is sulfur amino acid metabolism, where Cysteine Dioxygenase 1 (CDO1) is silenced by promoter hypermethylation (Brait et al., 2012; Jeschke et al., 2013), shielding NSCLC cells from sulfite toxicity and NADPH depletion (Kang et al., 2019), with plasma CDO1 methylation proposed for early detection (Wang et al., 2020) and taurine itself active in immune evasion and glycolysis (Cao et al., 2024; Sharma et al., 2025).

Public deposits are increasingly reused, and the quality of their metadata determines whether that reuse is meaningful. The two Fahrmann cohorts were designed as a discovery set and a test set, making them an obvious pairing for replication; however, whether the deposited files support that pairing has not been examined. This study therefore re-derives the discriminating serum metabolites and enriched pathways of LUAD from ST000385 without inheriting earlier interpretations, tests whether the catalyzing enzymes are dysregulated and prognostic in TCGA-LUAD, attempts external validation in ST000386, and asks what such a workflow can and cannot contribute where LDCT screening is unavailable.

RESEARCH METHOD

Data sources and metabolomic processing

This is an integrative in silico reanalysis using only MetaboAnalyst 6.0 and GEPIA2. No new human samples were collected, and no identifiable participant data were accessed, so institutional ethical approval was not required. We imported study ST000385 into MetaboAnalyst 6.0 through the Metabolomics Workbench accession interface (Pang et al., 2024). The deposit holds 192 samples across serum, plasma, and pooled quality-control aliquots, annotated for matrix, diagnosis, smoking status, and sex. Retaining serum cases and healthy controls gave 43 against 43 across 152 annotated compounds, with no missing values. This composition identifies the deposit as the ADC2 cohort of Fahrmann et al. (2015) rather than ADC1, and earlier descriptions of ST000385 as ADC1 with 52 cases and 31 controls do not match the file.

The deposited intensities are already total-quantity normalized, since every sample summed to approximately 2,000,000 at a coefficient of variation of 0.0%. Sample normalization was therefore None, transformation $\log_{10}$, and scaling Pareto, with no feature filtering because 152 features fall below the threshold MetaboAnalyst recommends. PLS-DA used 10-fold cross-validation and 2000 label permutations, with features ranked by Variable Importance In Projection (VIP) on component 1. Multivariate exploratory receiver operating characteristic (ROC) analysis used PLS-DA for classification and ranking, with Monte Carlo cross-validation and balanced subsampling over panels of 5 to 100 features. Univariate statistics used Welch t-tests on the same matrix with Benjamini–Hochberg correction across all 152 features, and fold changes as the ratio of group medians on the raw scale. Salicylic acid, lactamide, and the analytical standard. Standard 3 passed the 5% false discovery rate but were excluded from biological interpretation.

Pathway, sensitivity, and metadata analyses

The 39 features with VIP ≥ 1.0 were submitted to the Pathway Analysis module as a compound list, of which 34 mapped to KEGG. Enrichment used the hypergeometric test against the Homo sapiens library, with all library compounds as the reference metabolome, and impact was computed by relative-betweenness centrality, with enrichment claimed only below a 5% false discovery rate. Because aspartate, glutamate, taurine, and lactate are the metabolites most sensitive to hemolysis and delayed serum separation (Yin et al., 2013) and all four ranked among the strongest discriminators, their first principal component was taken as a pre-analytical proxy axis; every feature was regressed on it, and the comparison was repeated on the residuals.

Study ST000386 was imported by the same route as a candidate validation cohort. Its serum stratum holds 49 cases and 31 controls, matching the published description of ADC1. We first tested metadata integrity by asking whether each annotation associates with the metabolome at all. Diagnosis, sex, and smoking status are properties of the subject, whereas the matrix is a property of the aliquot. We ran Welch t-tests for each annotation across all features in both cohorts, compared the p-value distributions to a uniform distribution using the Kolmogorov–Smirnov test, and retested metabolites associated with sex in one cohort at $p < 0.001$ in the other.

Transcriptomic analysis

Enzyme expression was examined in TCGA-LUAD (483 tumors, 347 normal samples) using GEPIA2 (Tang et al., 2019) in Expression DIY, Box Plot mode, with matched TCGA normal and GTEx data checked, log scale on, and $|\log_2 FC| \geq 0.5$ with $p < 0.01$ as the reporting threshold. We assembled the panel in two stages. Eleven enzymes were

selected from the pathways treated as enriched in an earlier reading of this dataset, by listing two to three enzymes catalyzing the reactions in which the matched metabolites participate: GOT1, GOT2, and GLS for transamination and glutaminolysis; IDH1, OGDH, and SDHA for the TCA cycle; LDHA, PKM, and PDHA1 for pyruvate handling; and CDO1 and CSAD for cysteine–taurine biosynthesis. After arginine biosynthesis emerged as the leading pathway in the corrected analysis, we added six urea cycle enzymes and queried them under identical settings: ASS1, ASL, ARG1, ARG2, OTC, and CPS1. We assessed overall survival using median expression cutoffs, Cox proportional-hazards models, and log-rank tests.

RESULTS AND DISCUSSION

Results

Multivariate discrimination

Cross-validation selected a one-component PLS-DA model returning R^2 = 0.492, Q^2 = 0.302, and a cross-validated accuracy of 0.740 (Figure 1B). Permutation testing placed the observed statistic outside the entire null distribution, with no permutation of 2000 reaching the observed value (Figure 1C). The score plot shows partial but systematic separation along component 1 (Figure 1A), and 39 features exceeded the VIP threshold of 1.0 (Figure 1D).

Multivariate ROC analysis returned an AUC of 0.882 (95% CI 0.751–0.968) for a five-metabolite panel and 0.882 for ten, declining monotonically to 0.838 for 100 metabolites (Figure 2A), with predictive accuracy reaching 81.1% at five features (Figure 2B) and aspartate, glutamate, and maltose selected in more than 90% of Monte-Carlo iterations (Figure 2C). These values are substantially higher than the 0.676 to 0.692 reported in an earlier reading of this dataset, which used a cohort description that does not match the deposited file.

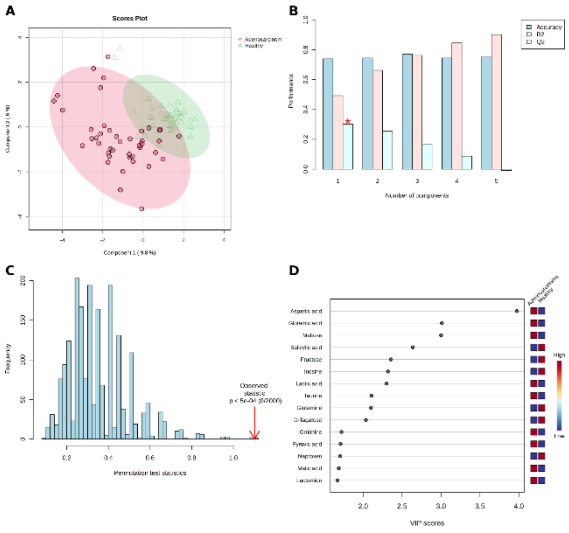

Figure 1. Multivariate model for LUAD discrimination in ST000385, (A) PLS-DA score plot with 95% confidence ellipses, (B) accuracy, R^2 and Q^2 by number of components; the red asterisk marks the selected model, (C) permutation test over 2000 label permutations, (D) VIP scores on component 1

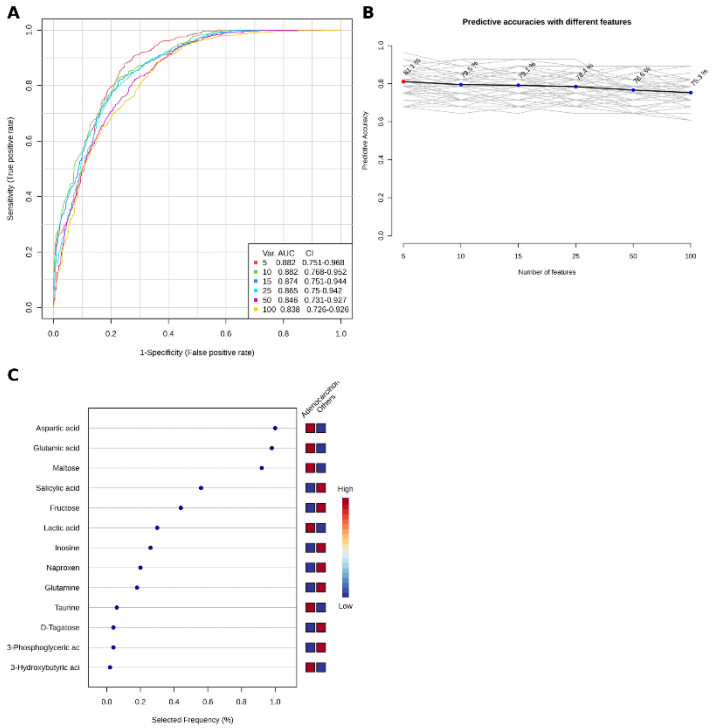

Figure 2. Classification performance, (A) multivariate exploratory ROC curves for panels of 5 to 100 features, (B) cross-validated predictive accuracy by panel size, (C) selection frequency across Monte-Carlo iterations

Differentially abundant metabolites

Twenty features passed a 5% false discovery rate, of which seventeen are endogenous metabolites (Table 1, full list in Supplementary Table S1). Aspartate carried the largest effect at 3.55-fold higher in cases, followed by glutamate and maltose, while glutamine and inosine were the most strongly reduced and ornithine, lactate, and pyruvate were all elevated.

Table 1. The ten most significant endogenous serum metabolites differing between LUAD cases and healthy controls in ST000385

Metabolite	VIP	Fold change	$\log_2$FC	FDR	Direction
Aspartic acid	3.98	3.55	1.83	1.4×10^{-8}	Higher in LUAD
Glutamic acid	3.01	2.30	1.20	1.8×10^{-7}	Higher in LUAD
Maltose	3.00	1.91	0.93	9.0×10^{-6}	Higher in LUAD
Taurine	2.11	1.55	0.63	2.0×10^{-4}	Higher in LUAD
Lactic acid	2.30	1.69	0.76	2.5×10^{-4}	Higher in LUAD
Glutamine	2.10	0.58	−0.79	6.0×10^{-4}	Lower in LUAD
Ornithine	1.72	1.50	0.59	1.5×10^{-3}	Higher in LUAD
Inosine	2.32	0.41	−1.27	2.9×10^{-3}	Lower in LUAD
D-Tagatose	2.04	0.65	−0.61	0.010	Lower in LUAD

Metabolite	VIP	Fold change	$\log_2$FC	FDR	Direction
Hippuric acid	1.54	1.20	0.27	0.012	Higher in LUAD

Table 1 shows the ten most significant endogenous serum metabolites differing between LUAD cases and healthy controls in ST000385. All seventeen are listed in Supplementary Table S1. Fold change is the ratio of group medians on the raw intensity scale. VIP is the component 1 score from the PLS-DA model. FDR is Benjamini–Hochberg adjusted across all 152 features.

Pathway enrichment

Three pathways survived correction (Table 2 and Figure 3). Arginine biosynthesis was the most strongly enriched, with five of its fourteen compounds matched, a false discovery rate of 5.8×10^{-5}, and a topology impact of 0.426. The matched compounds were aspartate, glutamate, glutamine, ornithine, and citrulline. Glyoxylate and dicarboxylate metabolism followed, although its impact of 0.087 places the matched compounds at the periphery of that network, and alanine, aspartate and glutamate metabolism was third with the highest impact in the analysis at 0.534.

Nitrogen metabolism and pyruvate metabolism fell just outside the threshold. Two pathways emphasised in earlier readings did not survive correction, since the tricarboxylic acid cycle returned 0.216 on two of twenty matched compounds and taurine and hypotaurine metabolism returned 0.400 on a single matched compound. Both results are partly a matter of platform coverage, because 152 compounds cannot populate those maps densely, but a single hit cannot support a pathway-level claim either way.

Table 2. KEGG pathway enrichment for the 39 metabolites with VIP ≥ 1.0

KEGG pathway	Hits	Raw p	FDR	Impact
Arginine biosynthesis	5/14	7.3×10^{-7}	5.8×10^{-5}	0.426
Glyoxylate and dicarboxylate metabolism	6/32	3.2×10^{-6}	1.3×10^{-4}	0.087
Alanine, aspartate and glutamate metabolism	4/28	5.4×10^{-4}	0.014	0.534
Nitrogen metabolism	2/6	2.9×10^{-3}	0.058	0.000
Pyruvate metabolism	3/23	3.9×10^{-3}	0.062	0.220
Glycine, serine and threonine metabolism	3/33	0.011	0.145	0.239
Citrate cycle (TCA cycle)	2/20	0.032	0.216	0.090
Taurine and hypotaurine metabolism	1/8	0.110	0.400	0.429

Table 2 shows the KEGG pathway enrichment for the 39 metabolites with VIP ≥ 1.0. Hits are matched compounds across pathway. Impact is relative betweenness centrality. The first three pathways survive a 5% false discovery rate.

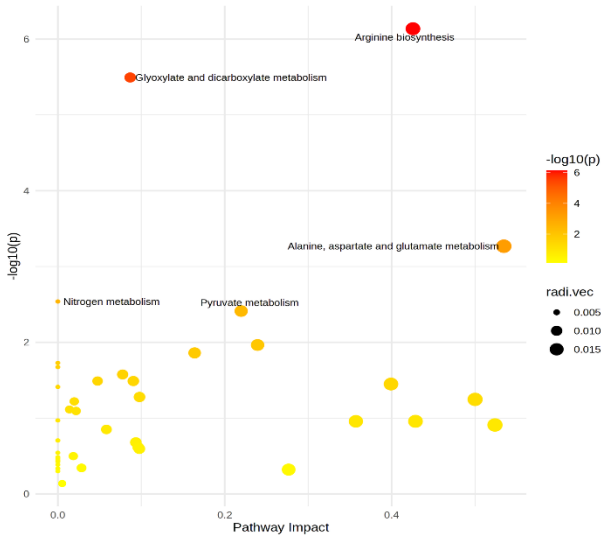

Figure 3. KEGG pathway enrichment

From Figure 3, explain the pathway impact on the horizontal axis against $-\log_{10}(p)$ on the vertical axis. Point size scales with pathway size, and colour scales with significance.

Sensitivity to pre-analytical structure

The four hemolysis-sensitive metabolites form a tight block on the log scale, with Pearson correlations of 0.80 between aspartate and glutamate, 0.86 between aspartate and taurine, and 0.60 between aspartate and lactate. Their first principal component captures 73.9% of their joint variance and separates cases from controls on its own (AUC of 0.852 and $p = 1.5 \times 10^{-10}$). That axis does not track analytical run date (ANOVA $p = 0.204$) and separates the groups consistently within each of the four run days, so instrument drift does not explain it. After regressing every feature on that axis, no metabolite retained a false discovery rate below 0.05, and only three of 152 reached a raw p below 0.05, where 7.6 would be expected. The case–control difference in this dataset is therefore essentially one-dimensional. This does not by itself distinguish a biological axis from a pre-analytical one, since adjusting for an axis that is itself the signal will always remove the signal, and this point is developed in the Limitations.

External validation attempt

Replication in ST000386 was not attainable. Its serum stratum holds 49 cases and 31 controls, matching the published description of ADC1, but none of its three subject-level annotations associate with the metabolome. Diagnosis returned seven features at $p < 0.05$ where 9.1 were expected, sex returned three and smoking status nine, and no p-value distribution departs from uniformity in the direction of signal. The matrix annotation, a property of the aliquot rather than of the subject, behaves entirely differently in the same file and returns 85 features at a Kolmogorov–Smirnov p-value of 4.6×10^{-33}, whereas in ST000385 all four annotations behave as expected and sex alone returns 33 features. Five metabolites associated with sex in ST000385 at $p < 0.001$ are present in ST000386, and none reaches $p < 0.05$ there, with the correlation of sex effect sizes between cohorts at −0.08 in serum. Sex differences in the serum metabolome are established biology and independent of disease, so a correctly aligned annotation at $n = 80$ would recover them, and randomized labels destroy signal rather than create it. Consistent with this, aspartate, reported by Fahrmann and colleagues as their best individual serum classifier from ADC1, reaches only $p = 0.026$ at a false discovery rate of 0.98 in the deposited ADC1 file.

The cohort was therefore withdrawn; all results reported here are based on ST000385 alone, and the full test is provided in Supplementary Table S4.

Urea cycle enzymes in TCGA-LUAD

The six urea cycle enzymes were not altered in one direction (Figure 4). ARG1 was the only enzyme in the group to meet the reporting threshold, and it was strongly reduced in tumor, with a median close to zero against a normal median near 2.1 on the $\log_2$(TPM+1) scale. OTC and CPS1 medians were also lower in tumor, whereas ASS1 and ARG2 were modestly higher and ASL was unchanged. CPS1 showed the most distinctive distribution, as its tumor median fell below the normal median. In contrast, a substantial minority of tumors expressed it far above any normal sample, reaching values above 10 on the same scale. ASS1 was therefore not silenced at the level of bulk tumor messenger RNA.

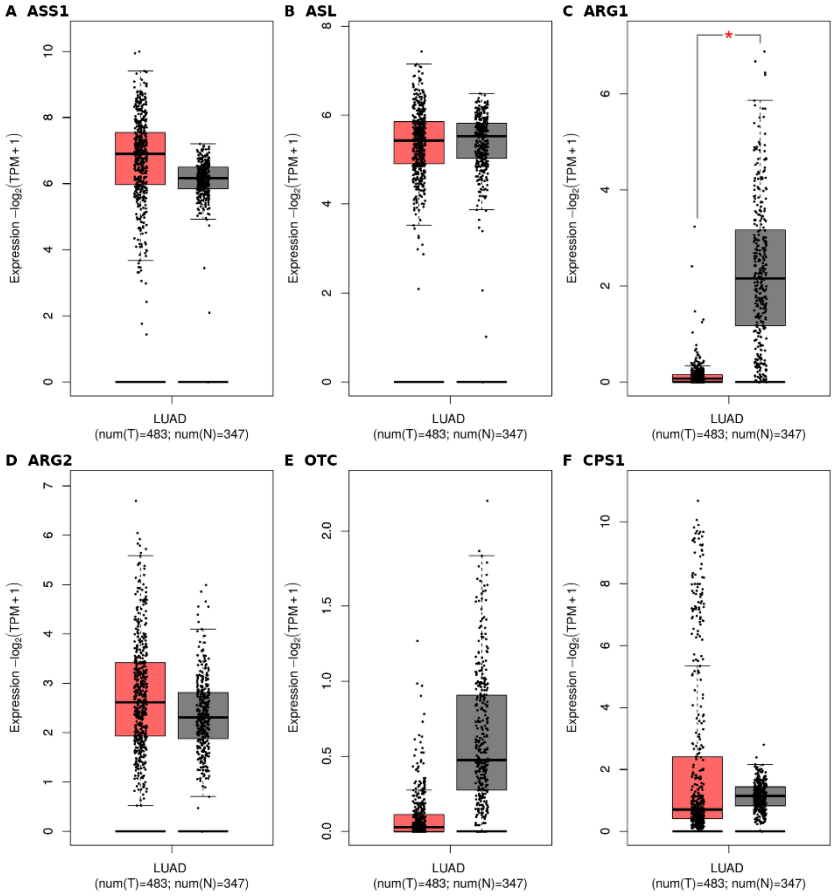

Figure 4. Expression of six urea cycle enzymes in TCGA-LUAD (483 tumor in red, 347 normal in grey), from GEPIA2 with matched TCGA and GTEx normal tissue

Figure 4 shows the expression of six urea cycle enzymes in TCGA-LUAD (483 tumor in red, 347 normal in grey), from GEPIA2 with matched TCGA and GTEx normal tissue. The red asterisk in panel C marks the only comparison meeting $|\log_2 FC| \geq 1$ with $p < 0.01$ at the GEPIA2 display cutoff.

The first enzyme panel and overall survival

In the eleven-enzyme panel assembled from the earlier pathway ranking, only LDHA, CDO1 and CSAD met the GEPIA2 display threshold, LDHA higher in tumor and the two cysteine–taurine enzymes lower, with CDO1 falling from a normal median near 3.5 to a tumor median near 0.9 (Supplementary Figure S2 and Supplementary Table S3). GOT1,

GOT2, PKM and IDH1 were higher in tumor without reaching that threshold, whereas GLS was lower in tumor, and OGDH, SDHA and PDHA1 differed little between groups. Three enzymes were prognostic for overall survival (Table 3 and Figure 5): LDHA at a hazard ratio of 1.9 with log-rank p = 3×10^{-5}, PKM at 1.4 with p = 0.017, and CDO1 at 0.69 with p = 0.017, where a ratio below one means that patients with high expression survived longer.

Table 3. Overall survival analysis in TCGA-LUAD (GEPIA2, median cutoff, 239 patients per group, except where noted)

Gene	Metabolic role	HR (high)	Log-rank p	Interpretation
LDHA	Pyruvate to lactate	1.9	3×10^{-5}	Poor prognosis
PKM	Glycolytic PKM2 isoform	1.4	0.017	Poor prognosis
CDO1	Cysteine to taurine	0.69	0.017	Protective
IDH1	α-Ketoglutarate metabolism	1.3	0.13	Not significant
CSAD	Hypotaurine to taurine	0.98	0.86	Not significant

From Table 3, explain the overall survival analysis in TCGA-LUAD (GEPIA2, median cutoff, 239 patients per group, except where noted). Hazard ratios are for the high-expression group.

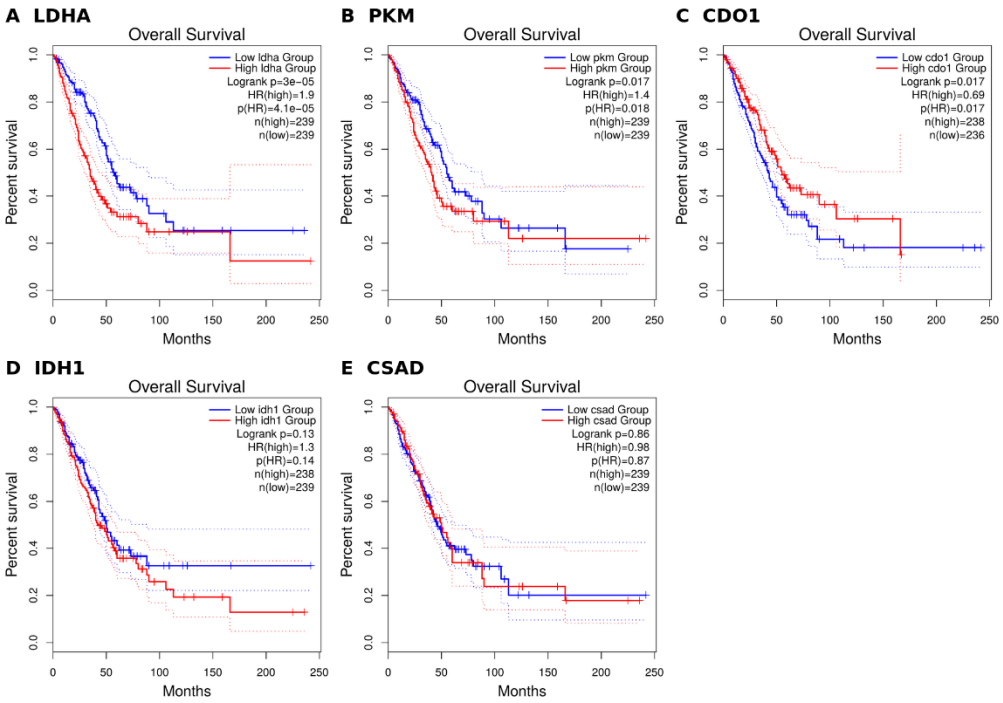

Figure 5. Kaplan–Meier overall survival curves in TCGA-LUAD by median expression, for the five enzymes taken to survival analysis

Figure 5 shows Kaplan–Meier overall survival curves in TCGA-LUAD by median expression for the five enzymes used in the survival analysis; blue denotes the low-expression group and red denotes the high-expression group, with dotted lines showing 95% confidence bands. Hazard ratios and log-rank p-values are printed within each panel.

Discussion

Nitrogen handling rather than glycolysis

The pathway result reorder the biological interpretation of this dataset. Arginine biosynthesis is the strongest enrichment by four orders of magnitude in raw p, driven by five compounds each independently altered or close to it. ASS1 consumes aspartate and citrulline, so any change in urea cycle flux would be expected to move exactly this set of metabolites, and elevated serum aspartate with elevated ornithine and reduced glutamine fits a rewired nitrogen-disposal axis. However, serum abundance cannot resolve which tissue the change originates in. Higher GOT1 and GOT2 in tumor are consistent with increased aspartate transamination, which is the step the KRAS, glutamine, and aspartate axis predicts (Lu et al., 2024; Romero et al., 2017; Son et al., 2013). GLS, however, was lower in tumor rather than higher, so the glutaminolysis arm of that axis is not supported at the transcript level here, and the serum fall in glutamine cannot be attributed to tumor glutaminase on this evidence. This differs from earlier readings of the same deposit, which placed Warburg metabolism first. Lactate and pyruvate are elevated, and both LDHA and PKM are elevated and prognostic. However, only LDHA met the expression threshold, so the glycolytic component is real but modest, and pyruvate metabolism did not survive correction, and the TCA cycle returned a false discovery rate of 0.216 on two matched compounds. The honest ordering places nitrogen handling first, glycolysis second as a single-metabolite observation, and the TCA cycle nowhere.

A dysregulated urea cycle rather than a silenced enzyme

The enzyme result refines the mechanism rather than confirming it. If serum aspartate rose because KRAS silenced ASS1, ASS1 messenger RNA should be lower in the tumor, but it is not. The six enzymes show a cycle pulled apart, with ARG1, OTC, and CPS1 medians down while ASS1 and ARG2 rise modestly. That is the pattern Lee and colleagues termed urea cycle dysregulation, in which most enzymes change expression, nitrogen is diverted toward pyrimidine synthesis, and nitrogen metabolites shift measurably in patient biofluids (Lee et al., 2018). The account predicts a serum signature dominated by aspartate, glutamate, glutamine, ornithine, and citrulline, and it fits these data better than a single silencing event. The bimodal CPS1 distribution adds a second reading, since a minority of tumors express CPS1 far above any normal sample, which is the behavior reported for KRAS and LKB1 co-mutant lung tumors that route ammonia nitrogen into pyrimidines (Kim et al., 2017). The ASS1 silencing described by Gai and colleagues is specific to KRAS-mutant cells (Gai et al., 2024). In contrast, TCGA-LUAD mixes genotypes, so a restricted effect could be masked, and mutation-stratified analysis would be needed to test this.

ARG1 needs the most caution despite being the only enzyme meeting the threshold. Arginase 1 in lung tissue is expressed mainly by granulocytic myeloid cells rather than by epithelium, and its activity in lung tumors has been reported as elevated at the protein level in exactly those cells (Miret et al., 2019). A fall in bulk messenger RNA against GTEx normal lung therefore reflects cell composition at least as much as tumor-cell biology, and the same caution applies to any bulk comparison of a lineage-restricted gene. What can be said is that arginine availability is under regulation from several directions at once, through synthesis, catabolism, and nitric oxide production (Keshet & Erez, 2018), and that ASS1 expression already stratifies patients for arginine deprivation therapy in another tumor type (Barnett et al., 2023).

Taurine

Taurine was elevated 1.55-fold, and CDO1 and CSAD were the most strongly downregulated enzymes in the eleven-enzyme panel, with CDO1 also protective for survival, which is consistent with the tumor-suppressor behavior of CDO1 and with the account in which silencing shields NSCLC cells from sulfite toxicity and NADPH depletion (Brait et al., 2012; Jeschke et al., 2013; Kang et al., 2019). Pathway-level enrichment cannot be claimed, since taurine was the only compound of the eight in that KEGG map present on this platform, and recent findings complicate a simple reading in any case (Cao et al., 2024; Sharma et al., 2025).

Open-data reuse and implications for SDG target 3.4

The metadata defect documented here is not peripheral. Any group attempting the obvious biological reuse- validation of an ADC2 finding in ADC1- would obtain a null result and could reasonably report it as a failure to replicate. That would be wrong, and wrong in a direction that discourages further reuse of a valuable deposit. The test is cheap, since running the disease comparison alongside a biological control annotation costs one additional analysis and settles whether a null result is informative. The point generalizes to machine-learning practice on public repositories. Where a deposit is used to benchmark algorithms, shuffled subject labels are harmless, because only the relative ranking of methods is at stake and every method is handicapped equally, whereas biological reuse has no such protection. Reporting a biological control annotation alongside the disease label costs nothing. It would flag deposits that cannot support downstream biological claims, and this check is recommended as routine before using any public deposit as a validation cohort.

Two implications follow for the SDG 3.4 agenda. Where LDCT screening is unavailable, the binding constraint on earlier diagnosis is infrastructure rather than analytic sophistication (Asmara et al., 2023; Hanafi et al., 2024), and a blood-drawn assay shifts that constraint to laboratory capacity, which is more widely distributed. An AUC of 0.882 on five metabolites falls within the range reported by dedicated cohorts (Li et al., 2022; Qi et al., 2021), but it should be read as an upper bound. The classification layer here is a supervised machine-learning pipeline, and PLS-DA on high-dimensional data with few samples is prone to over-optimistic estimates under internal cross-validation, with the size of that bias set by the resampling scheme (Rodríguez-Pérez et al., 2018). Performance also declined from five to one hundred features, which marks features that add variance without adding information, and the sensitivity analysis found a single correlated axis carrying almost the whole separation. A more expressive learner such as a random forest or a neural network (Galal et al., 2022) would therefore change the reported number without changing what it means, since the limiting factor is sample provenance rather than model capacity, and the panel is not a candidate assay. This analysis contributes a prioritized target set with a specific mechanistic hypothesis about ASS1. The second implication is methodological. Every dataset and tool used here is free, and the analysis required no sample collection, mass-spectrometry facilities, or licensed software. Hence, this class of reanalysis converts existing global data into locally directed hypotheses at minimal cost and can be taught. The ST000386 finding shows the other half of that argument, since groups working from public deposits must test what they inherit rather than assume it (Cancela et al., 2023; Dee et al., 2025; Liu et al., 2021).

The strongest limitation concerns pre-analytical variation. The four metabolites dominating the signal are the most sensitive to hemolysis and to delay between blood draw and separation of serum from cells (Yin et al., 2013); they define a single axis carrying 73.9% of their joint variance, and that axis alone reaches an AUC of 0.852. A 3.55-

fold difference in serum aspartate exceeds what whole-body tumor metabolism would be expected to produce, and cases and controls in a hospital-recruited study are frequently sampled in different clinical settings. The deposit records no collection time, processing delay, or hemolysis index, so the explanation cannot be confirmed or excluded. This does not invalidate the results, and the original authors reported aspartate independently, but any prospective evaluation must standardize sample handling and record a hemolysis index.

Six further limitations apply. External validation could not be completed, so all metabolomic results rest on 43 cases and 43 controls with internal cross-validation only. The pathway analysis used the VIP ≥ 1.0 list rather than the false-discovery-rate list (MetaboAnalyst default), a more permissive input, and five of 39 names did not map to KEGG. Platform coverage constrains which pathways can be detected, so absence of enrichment is not evidence of absence of involvement. The transcriptomic layer relies on GEPIA2 summaries of bulk tissue, which cannot separate tumor cells from stroma and immune infiltrate. It cannot be stratified by KRAS, LKB1, EGFR, or TP53 status, which matters directly because both mechanisms invoked here are genotype-restricted. The metabolomic and transcriptomic layers derive from different individuals, so the enzyme findings corroborate the pathway inference without confirming it within the same patients, and all correlation and survival findings are associative. Residual confounding from medication, fasting state, and comorbidity could not be addressed, and salicylic acid and naproxen among the discriminating features show that medication differences between the groups are real.

CONCLUSION

Fundamental Finding: Reanalysis of the ST000385 serum cohort places arginine biosynthesis and nitrogen handling, not the Warburg axis, at the centre of the LUAD serum metabolomic signature, and five compounds of that pathway are altered with an enrichment that survives correction by four orders of magnitude; the urea cycle enzymes in TCGA-LUAD are dysregulated rather than switched off, since ARG1, OTC and CPS1 medians fall while ASS1 and ARG2 rise modestly and CPS1 is expressed far above normal in a minority of tumors. **Implication:** Glycolytic change stands only as a single-metabolite observation supported by elevated and prognostic LDHA and PKM, while claims about the tricarboxylic acid cycle and taurine biosynthesis are not supported at pathway level, so an open deposit can redirect a mechanistic reading for groups without cohort access, which is where the wider value for SDG target 3.4 lies. **Limitation:** The five-metabolite panel reached an AUC of 0.882 under internal cross-validation, but the separation is one-dimensional and dominated by metabolites sensitive to sample handling, so the figure is an upper bound rather than a candidate assay and no additional model capacity would change it; external validation could not be completed because the subject-level metadata of the companion deposit ST000386 does not associate with its own metabolome, a defect documented here for the first time. **Future Research:** Mutation-stratified testing of the urea cycle in LUAD, in cohorts with verified metadata and controlled pre-analytical handling, is the priority, alongside a metadata-integrity check as a routine first step before any public deposit is reused.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge Universitas Negeri Surabaya for providing financial support for this research.

AUTHOR CONTRIBUTIONS

Muhammad Nurrohman Sidiq: conceptualization, methodology, software, formal analysis, data curation, visualization, and writing of the original draft. **Rudiana Agustini**: conceptualization, validation, supervision, and writing through review and editing. **Nuniek Herdyastuti**: methodology, validation, and writing through review and editing. **Prima Retno Wikandari**: investigation, resources, and writing through review and editing. **Mirwa Adiprahara Anggarani**: project administration, data curation, and writing through review and editing. All listed authors have reviewed and approved the final version of this submission.

CONFLICT OF INTEREST STATEMENT

The authors confirm that there are no conflicts of interest, either financial or personal, that may have influenced the content or outcome of this study.

ETHICAL COMPLIANCE STATEMENT

This manuscript complies with research and publication ethics. The authors affirm that the work is original, conducted with academic integrity, and free from any unethical practices, including plagiarism.

STATEMENT ON THE USE OF AI OR DIGITAL TOOLS IN WRITING

The authors acknowledge the use of digital tools, including AI-based technologies, as support in the research and writing stages of this article. MetaboAnalyst 6.0 was used for multivariate modelling, receiver operating characteristic analysis, and pathway enrichment; GEPIA2 was used for the transcriptomic and survival queries in TCGA-LUAD; Python with pandas, NumPy, and SciPy was used for the univariate statistics, the pre-analytical sensitivity analysis, and the metadata integrity test, and the script is reproduced in the Supplementary Material. Quillbot and Claude was used solely for language editing and proofreading. No text, figure, or numerical result was accepted without verification against the primary data and the cited sources. All outputs generated with digital assistance were critically evaluated and revised to ensure academic rigour and ethical standards were upheld. The final responsibility for the manuscript rests entirely with the authors.

REFERENCES

Anastasiou, D., Poulogiannis, G., Asara, J. M., Boxer, M. B., Jiang, J.-K., Shen, M., Bellinger, G., Sasaki, A. T., Locasale, J. W., Auld, D. S., Thomas, C. J., Vander Heiden, M. G., & Cantley, L. C. (2011). Inhibition of pyruvate kinase M2 by reactive oxygen species contributes to cellular antioxidant responses. *Science, 334*(6060), 1278–1283. https://doi.org/10.1126/science.1211485

Asmara, O. D., Tenda, E. D., Singh, G., Pitoyo, C. W., Rumende, C. M., Rajabto, W., Ananda, N. R., Trisnawati, I., Budiyono, E., Thahadian, H. F., Boerma, E. C., Faisal, A., Hutagaol, D., Soeharto, W., Radityamurti, F., Marfiani, E., Romadhon, P. Z., Kholis, F. N., Suryadinata, H., Van Geffen, W. H. (2023). Lung cancer in Indonesia. *Journal of Thoracic Oncology, 18*(9), 1134–1145. https://doi.org/10.1016/j.jtho.2023.06.010

Barnett, S. E., Kenyani, J., Tripari, M., Butt, Z., Grosman, R., Querques, F., Shaw, L., Silva, L. C., Goate, Z., Marciniak, S. J., Rassl, D. M., Jackson, R., Lian, L.-Y., Szlosarek, P. W., Sacco, J. J., & Coulson, J. M. (2023). BAP1 loss is associated with higher ASS1 expression in epithelioid mesothelioma: Implications for therapeutic stratification.

Molecular Cancer Research, 21(5), 411–427. https://doi.org/10.1158/1541-7786.MCR-22-0635

Berardinis, R. J. D., & Cheng, T. (2010). Q's next: The diverse functions of glutamine in metabolism, cell biology and cancer. *Oncogene, 29*(3), 313–324. https://doi.org/10.1038/onc.2009.358

Brait, M., Ling, S., Nagpal, J. K., Chang, X., Park, H. L., Lee, J., Okamura, J., Yamashita, K., Sidransky, D., & Kim, M. S. (2012). Cysteine dioxygenase 1 is a tumor suppressor gene silenced by promoter methylation in multiple human cancers. *PLOS ONE, 7*(9), 1-19. https://doi.org/10.1371/journal.pone.0044951

Cancela, M. D. C., de Souza, D. L. B., Martins, L. F. L., Borges, L., Schilithz, A. O., Hanly, P., Sharp, L., Pearce, A., & Soerjomataram, I. (2023). Can the sustainable development goals for cancer be met in Brazil? A population-based study. *Frontiers in Oncology, 12*(1), 1-12. https://doi.org/10.3389/fonc.2022.1060608

Cao, T., Zhang, W., Wang, Q., Wang, C., Ma, W., Zhang, C., Ge, M., Tian, M., Yu, J., Jiao, A., Wang, L., Liu, M., Wang, P., Guo, Z., Zhou, Y., Chen, S., Yin, W., Yi, J., Guo, H., Zhao, X. (2024). Cancer SLC6A6-mediated taurine uptake transactivates immune checkpoint genes and induces exhaustion in CD8+ T cells. *Cell, 187*(9), 2288–2304. https://doi.org/10.1016/j.cell.2024.03.011

Christofk, H. R., Vander Heiden, M. G., Wu, N., Asara, J. M., & Cantley, L. C. (2008). Pyruvate kinase M2 is a phosphotyrosine-binding protein. *Nature, 452*(7184), 181–186. https://doi.org/10.1038/nature06667

Dee, E. C., Laversanne, M., Bhoo-Pathy, N., Ho, F. D. V., Feliciano, E. J. G., Eala, M. A. B., Ting, F. I. L., Ginsburg, O., Moraes, F. Y., Gyawali, B., Gomez, S. L., Ng, K., Wu, J. F., Jain, U., Jain, B., Columbres, R. C., Matsuda, T., Sangrajrang, S., Sinuraya, E. S., Bray, F. (2025). Cancer incidence and mortality estimates in 2022 in Southeast Asia: A comparative analysis. *The Lancet Oncology, 26*(4), 516–528. https://doi.org/10.1016/S1470-2045(25)00017-8

Fahrmann, J. F., Kim, K., DeFelice, B. C., Taylor, S. L., Gandara, D. R., Yoneda, K. Y., Cooke, D. T., Fiehn, O., Kelly, K., & Miyamoto, S. (2015). Investigation of metabolomic blood biomarkers for detection of adenocarcinoma lung cancer. *Cancer Epidemiology, Biomarkers & Prevention, 24*(11), 1716–1723. https://doi.org/10.1158/1055-9965.EPI-15-0427

Fiehn, O. (2016). Metabolomics by gas chromatography–mass spectrometry: Combined targeted and untargeted profiling. *Current Protocols in Molecular Biology, 114*(1), 1-43. https://doi.org/10.1002/0471142727.mb3004s114

Frieden, T. R., Cobb, L. K., Leidig, R. C., Mehta, S., & Kass, D. (2020). Reducing premature mortality from cardiovascular and other non-communicable diseases by one third: Achieving sustainable development goal indicator 3.4.1. *Global Heart, 15*(1), 1-10. https://doi.org/10.5334/gh.531

Gai, X., Liu, Y., Lan, X., Chen, L., Yuan, T., Xu, J., Li, Y., Zheng, Y., Yan, Y., Yang, L., Fu, Y., Tang, S., Cao, S., Dai, X., Zhu, H., Geng, M., Ding, J., Pu, C., & Huang, M. (2024). Oncogenic KRAS induces arginine auxotrophy and confers a therapeutic vulnerability to SLC7A1 inhibition in non-small cell lung cancer. *Cancer Research, 84*(12), 1963–1977. https://doi.org/10.1158/0008-5472.CAN-23-2095

Galal, A., Talal, M., & Moustafa, A. (2022). Applications of machine learning in metabolomics: Disease modeling and classification. *Frontiers in Genetics, 13*, 1017340. https://doi.org/10.3389/fgene.2022.1017340

Hanafi, A. R., Hanif, M. A., Pangaribuan, M. T. G., Ariawan, W. P., Sutandyo, N., Kurniawati, S. A., Setiawan, L., Cahyanti, D., Rayhani, F., & Imelda, P. (2024).

Genomic features of lung cancer patients in Indonesia's national cancer center. *BMC Pulmonary Medicine, 24*(1), 1-11. https://doi.org/10.1186/s12890-024-02851-y

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell, 144*(5), 646–674. https://doi.org/10.1016/j.cell.2011.02.013

Harding, J. J., Telli, M., Munster, P., Voss, M. H., Infante, J. R., DeMichele, A., Dunphy, M., Le, M. H., Molineaux, C., Orford, K., Parlati, F., Whiting, S. H., Bennett, M. K., Tannir, N. M., & Meric-Bernstam, F. (2021). A phase I dose-escalation and expansion study of telaglenastat in patients with advanced or metastatic solid tumors. *Clinical Cancer Research, 27*(18), 4994–5003. https://doi.org/10.1158/1078-0432.CCR-21-1204

Henke, O., Qader, A. Q., Malle, G. L., Kuiate, J. R., Hennig, L., Demeke, T., Stroetmann, C., Henke, A. A., Alaric, T. T., Rushanyan, M., Enssle, C., & Bussmann, H. (2023). International cooperation to fight cancer's late-stage presentation in low- and middle-income countries. *Clinical & Experimental Metastasis, 40*(1), 1–3. https://doi.org/10.1007/s10585-022-10196-1

Herbst, R. S., Morgensztern, D., & Boshoff, C. (2018). The biology and management of non-small cell lung cancer. *Nature, 553*(7689), 446–454. https://doi.org/10.1038/nature25183

Jeschke, J., O'Hagan, H. M., Zhang, W., Vatapalli, R., Calmon, M. F., Danilova, L., Nelkenbrecher, C., Van Neste, L., Bijsmans, I. T. G. W., Van Engeland, M., Gabrielson, E., Schuebel, K. E., Winterpacht, A., Baylin, S. B., Herman, J. G., & Ahuja, N. (2013). Frequent inactivation of cysteine dioxygenase type 1 contributes to survival of breast cancer cells and resistance to anthracyclines. *Clinical Cancer Research, 19*(12), 3201–3211. https://doi.org/10.1158/1078-0432.CCR-12-3751

Kang, Y. P., Torrente, L., Falzone, A., Elkins, C. M., Liu, M., Asara, J. M., Dibble, C. C., & DeNicola, G. M. (2019). Cysteine dioxygenase 1 is a metabolic liability for non-small cell lung cancer. *eLife, 8*(5), 1-35. https://doi.org/10.7554/eLife.45572

Keshet, R., & Erez, A. (2018). Arginine and the metabolic regulation of nitric oxide synthesis in cancer. *Disease Models & Mechanisms, 11*(8), 1-11. https://doi.org/10.1242/dmm.033332

Kim, J., Hu, Z., Cai, L., Li, K., Choi, E., Faubert, B., Bezwada, D., Rodriguez-Canales, J., Villalobos, P., Lin, Y.-F., Ni, M., Huffman, K. E., Girard, L., Byers, L. A., Unsal-Kacmaz, K., Peña, C. G., Heymach, J. V., Wauters, E., Vansteenkiste, J., DeBerardinis, R. J. (2017). CPS1 maintains pyrimidine pools and DNA synthesis in KRAS/LKB1-mutant lung cancer cells. *Nature, 546*(7656), 168–172. https://doi.org/10.1038/nature22359

Lee, J. S., Adler, L., Karathia, H., Carmel, N., Rabinovich, S., Auslander, N., Keshet, R., Stettner, N., Silberman, A., Agemy, L., Helbling, D., Eilam, R., Sun, Q., Brandis, A., Malitsky, S., Itkin, M., Weiss, H., Pinto, S., Kalaora, S., … Erez, A. (2018). Urea cycle dysregulation generates clinically relevant genomic and biochemical signatures. *Cell, 174*(6), 1559–1570. https://doi.org/10.1016/j.cell.2018.07.019

Li, J., Liu, K., Ji, Z., Wang, Y., Yin, T., Long, T., Shen, Y., & Cheng, L. (2023). Serum untargeted metabolomics reveal metabolic alteration of non-small cell lung cancer and refine disease detection. *Cancer Science, 114*(2), 680–689. https://doi.org/10.1111/cas.15629

Liang, L.-J., Yang, F.-Y., Wang, D., Zhang, Y.-F., Yu, H., Wang, Z., Sun, B.-B., Liu, Y.-T., Wang, G.-Z., & Zhou, G.-B. (2024). CIP2A induces PKM2 tetramer formation and oxidative phosphorylation in non-small cell lung cancer. *Cell Discovery, 10*(1), 1-22. https://doi.org/10.1038/s41421-023-00633-0

Liberti, M. V., & Locasale, J. W. (2016). The Warburg effect: How does it benefit cancer cells? *Trends in Biochemical Sciences, 41*(3), 211–218. https://doi.org/10.1016/j.tibs.2015.12.001

Liu, L., Li, Y., Song, J., Chen, Q., Li, S., Mu, H., Na, J., Zhang, R., Yu, L., Sun, W., & Pan, G. (2021). Current status of premature mortality from four non-communicable diseases and progress towards the Sustainable Development Goal target 3.4: A population-based study in northeast China, 2004–2017. *BMC Public Health, 21*(1), 1608. https://doi.org/10.1186/s12889-021-11611-0

Liu, Z., Wang, L., Gao, S., Xue, Q., Tan, F., Li, Z., & Gao, Y. (2023). Plasma metabolomics study in screening and differential diagnosis of multiple primary lung cancer. *International Journal of Surgery, 109*(3), 297–312. https://doi.org/10.1097/JS9.0000000000000006

Lu, W., Cui, J., Wang, W., Hu, Q., Xue, Y., Liu, X., Gong, T., Lu, Y., Ma, H., Yang, X., Feng, B., Wang, Q., Zhang, N., Xu, Y., Liu, M., Nussinov, R., Cheng, F., Ji, H., & Huang, J. (2024). PPIA dictates NRF2 stability to promote lung cancer progression. *Nature Communications, 15*(1), 4703. https://doi.org/10.1038/s41467-024-48364-4

Mazzone, P. J., Silvestri, G. A., Souter, L. H., Caverly, T. J., Kanne, J. P., Katki, H. A., Wiener, R. S., & Detterbeck, F. C. (2021). Screening for lung cancer: CHEST guideline and expert panel report. *Chest, 160*(5), 427–494. https://doi.org/10.1016/j.chest.2021.06.063

Metallo, C. M., Gameiro, P. A., Bell, E. L., Mattaini, K. R., Yang, J., Hiller, K., Jewell, C. M., Johnson, Z. R., Irvine, D. J., Guarente, L., Kelleher, J. K., Vander Heiden, M. G., Iliopoulos, O., & Stephanopoulos, G. (2012). Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia. *Nature, 481*(7381), 380–384. https://doi.org/10.1038/nature10602

Miret, J. J., Kirschmeier, P., Koyama, S., Zhu, M., Li, Y. Y., Naito, Y., Wu, M., Malladi, V. S., Huang, W., Walker, W., Palakurthi, S., Dranoff, G., Hammerman, P. S., Pecot, C. V., Wong, K.-K., & Akbay, E. A. (2019). Suppression of myeloid cell arginase activity leads to therapeutic response in a NSCLC mouse model by activating anti-tumor immunity. *Journal for ImmunoTherapy of Cancer, 7*(1), 1-32. https://doi.org/10.1186/s40425-019-0504-5

National Lung Screening Trial Research Team. (2011). Reduced lung-cancer mortality with low-dose computed tomographic screening. *New England Journal of Medicine, 365*(5), 395–409. https://doi.org/10.1056/NEJMoa1102873

Pang, Z., Lu, Y., Zhou, G., Hui, F., Xu, L., Viau, C., Spigelman, A. F., MacDonald, P. E., Wishart, D. S., Li, S., & Xia, J. (2024). MetaboAnalyst 6.0: Towards a unified platform for metabolomics data processing, analysis and interpretation. *Nucleic Acids Research, 52*(1), 398–406. https://doi.org/10.1093/nar/gkae253

Pavlova, N. N., & Thompson, C. B. (2016). The emerging hallmarks of cancer metabolism. *Cell Metabolism, 23*(1), 27–47. https://doi.org/10.1016/j.cmet.2015.12.006

Qi, S.-A., Wu, Q., Chen, Z., Zhang, W., Zhou, Y., Mao, K., Li, J., Li, Y., Chen, J., Huang, Y., & Huang, Y. (2021). High-resolution metabolomic biomarkers for lung cancer diagnosis and prognosis. *Scientific Reports, 11*(1), 11805. https://doi.org/10.1038/s41598-021-91276-2

Rodríguez-Pérez, R., Fernández, L., & Marco, S. (2018). Overoptimism in cross-validation when using partial least squares-discriminant analysis for omics data: A systematic study. *Analytical and Bioanalytical Chemistry, 410*(23), 5981–5992. https://doi.org/10.1007/s00216-018-1217-1

Romero, R., Sayin, V. I., Davidson, S. M., Bauer, M. R., Singh, S. X., LeBoeuf, S. E., Karakousi, T. R., Ellis, D. C., Bhutkar, A., Sánchez-Rivera, F. J., Subbaraj, L.,

Martinez, B., Bronson, R. T., Prigge, J. R., Schmidt, E. E., Thomas, C. J., Goparaju, C., Davies, A., Dolgalev, I., … Papagiannakopoulos, T. (2017). Keap1 loss promotes Kras-driven lung cancer and results in dependence on glutaminolysis. *Nature Medicine, 23*(11), 1362–1368. https://doi.org/10.1038/nm.4407

Sciacovelli, M., & Frezza, C. (2016). Oncometabolites: Unconventional triggers of oncogenic signalling cascades. *Free Radical Biology and Medicine, 100*, 175–181. https://doi.org/10.1016/j.freeradbiomed.2016.04.025

Seijo, L. M., Peled, N., Ajona, D., Boeri, M., Field, J. K., Sozzi, G., Pio, R., Zulueta, J. J., Spira, A., Massion, P. P., Mazzone, P. J., & Montuenga, L. M. (2019). Biomarkers in lung cancer screening: Achievements, promises, and challenges. *Journal of Thoracic Oncology, 14*(3), 343–357. https://doi.org/10.1016/j.jtho.2018.11.023

Sharma, S., Rodems, B. J., Baker, C. D., Kaszuba, C. M., Franco, E. I., Smith, B. R., Ito, T., Swovick, K., Welle, K., Zhang, Y., Rock, P., Chaves, F. A., Ghaemmaghami, S., Calvi, L. M., Ganguly, A., Burack, W. R., Becker, M. W., Liesveld, J. L., Brookes, P. S., Bajaj, J. (2025). Taurine from tumour niche drives glycolysis to promote leukaemogenesis. *Nature, 644*(8075), 263–272. https://doi.org/10.1038/s41586-025-09018-7

Son, J., Lyssiotis, C. A., Ying, H., Wang, X., Hua, S., Ligorio, M., Perera, R. M., Ferrone, C. R., Mullarky, E., Shyh-Chang, N., Kang, Y., Fleming, J. B., Bardeesy, N., Asara, J. M., Haigis, M. C., DePinho, R. A., Cantley, L. C., & Kimmelman, A. C. (2013). Glutamine supports pancreatic cancer growth through a KRAS-regulated metabolic pathway. *Nature, 496*(7443), 101–105. https://doi.org/10.1038/nature12040

Tang, Z., Kang, B., Li, C., Chen, T., & Zhang, Z. (2019). GEPIA2: An enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Research, 47*(1), 556–560. https://doi.org/10.1093/nar/gkz430

Travis, W. D., Brambilla, E., Nicholson, A. G., Yatabe, Y., Austin, J. H. M., Beasley, M. B., Chirieac, L. R., Dacic, S., Duhig, E., Flieder, D. B., Geisinger, K., Hirsch, F. R., Ishikawa, Y., Kerr, K. M., Noguchi, M., Pelosi, G., Powell, C. A., Tsao, M. S., & Wistuba, I. (2015). The 2015 World Health Organization classification of lung tumors: Impact of genetic, clinical and radiologic advances since the 2004 classification. *Journal of Thoracic Oncology, 10*(9), 1243–1260. https://doi.org/10.1097/JTO.0000000000000630

Wang, P., Zhao, H., Shi, R., Liu, X., Liu, J., Ren, F., Zhao, Q., Zhang, H., Li, Y., Liu, H., & Chen, J. (2020). The role of plasma CDO1 methylation in the early diagnosis of lung cancer. *Zhongguo Fei Ai Za Zhi [Chinese Journal of Lung Cancer], 23*(5), 314–320. https://doi.org/10.3779/j.issn.1009-3419.2020.102.20

Wishart, D. S. (2019). Metabolomics for investigating physiological and pathophysiological processes. *Physiological Reviews, 99*(4), 1819–1875. https://doi.org/10.1152/physrev.00035.2018

Xie, H., Hanai, J.-I., Ren, J.-G., Kats, L., Burgess, K., Bhargava, P., Signoretti, S., Billiard, J., Duffy, K. J., Grant, A., Wang, X., Lorkiewicz, P. K., Schatzman, S., Bousamra, M., Lane, A. N., Higashi, R. M., Fan, T. W. M., Pandolfi, P. P., Sukhatme, V. P., & Seth, P. (2014). Targeting lactate dehydrogenase-A inhibits tumorigenesis and tumor progression in mouse models of lung cancer and impacts tumor-initiating cells. *Cell Metabolism, 19*(5), 795–809. https://doi.org/10.1016/j.cmet.2014.03.003

Yin, P., Peter, A., Franken, H., Zhao, X., Neukamm, S. S., Rosenbaum, L., Lucio, M., Zell, A., Häring, H.-U., Xu, G., & Lehmann, R. (2013). Preanalytical aspects and sample quality assessment in metabolomics studies of human blood. *Clinical Chemistry, 59*(5), 833–845. https://doi.org/10.1373/clinchem.2012.199257

Zhou, J., Xu, Y., Liu, J., Feng, L., Yu, J., & Chen, D. (2024). Global burden of lung cancer in 2022 and projections to 2050: Incidence and mortality estimates from

GLOBOCAN. *Cancer Epidemiology, 93*, 102693. https://doi.org/10.1016/j.canep.2024.102693

*** Muhammad Nurrohman Sidiq, S.Si., M.Sc., Ph.D. (Corresponding Author)**
Department of Chemistry, Faculty of Mathematics and Natural Sciences,
State University of Surabaya
Jl. Ketintang, Surabaya, East Java 60231, Indonesia
Email: muhammadsidiq@unesa.ac.id

Prof. Dr. Rudiana Agustini, M.Pd.
Department of Chemistry, Faculty of Mathematics and Natural Sciences,
State University of Surabaya
Jl. Ketintang, Surabaya, East Java 60231, Indonesia
Email: rudianaagustini@unesa.ac.id

Prof. Dr. Nuniek Herdyastuti, M.Si.
Department of Chemistry, Faculty of Mathematics and Natural Sciences,
State University of Surabaya
Jl. Ketintang, Surabaya, East Java 60231, Indonesia
Email: nuniekherdyastuti@unesa.ac.id

Prof. Dr. Prima Retno Wikandari, M.Si.
Department of Chemistry, Faculty of Mathematics and Natural Sciences,
State University of Surabaya
Jl. Ketintang, Surabaya, East Java 60231, Indonesia
Email: primaretno@unesa.ac.id

Mirwa Adiprahara Anggarani, S.Si., M.Si.
Department of Chemistry, Faculty of Mathematics and Natural Sciences,
State University of Surabaya
Jl. Ketintang, Surabaya, East Java 60231, Indonesia
Email: mirwaanggarani@unesa.ac.id