SCIENCE PHOTONS Publishers

Research Article

Innovative Insights in Digital Health

Integrating Serine and Pyrimidine Biosynthesis into the E-TRIPOD Signature for Enhanced Lung Cancer Screening

Raúl Isea*

Fundación IDEA, Hoyo de la Puerta, Baruta, Venezuela.

Received Date: 07 Feb 2026;
Accepted Date: 09 Mar 2026;
Published Date: 12 Mar 2026.

***Correspondence:** Raúl Isea, Fundación IDEA, Hoyo de la Puerta, Baruta, Venezuela.*

Citation: Raúl Isea. Integrating Serine and Pyrimidine Biosynthesis into the E-TRIPOD Signature for Enhanced Lung Cancer Screening. Innov Insights Digit Health. 2026; 2(1): 1-7.

ABSTRACT

Introduction: *Early diagnosis of lung adenocarcinoma is a clinical priority to reduce global mortality, driving the development of various transcriptomic signatures. However, most lack a mechanistic axis connecting gene expression with external diagnostic signals.*

Objective: *To characterize the Extended TRIPOD (E-TRIPOD) model, a five-gene signature (COL11A1, CDC20, PSAT1, LDHA, and UCK2) that integrates structural, proliferative, and metabolic drivers as a biological and theoretical foundation for future e-Nose-based screening.*

Methods: *We analyzed the GSE19188 dataset and performed cross-validation on TCGA real-world data cohorts (n > 500). Kaplan-Meier survival analysis, functional enrichment in g:Profiler, and metabolic mapping via RSEA and KEGG were used to identify potential volatile byproducts.*

Results: *The model showed exceptional prognostic power with a Hazard Ratio of 2.1 ($P=8.8\times10-16$). Unlike standard models, E-TRIPOD reveals a mechanistic axis based on the pyrimidine salvage pathway. Here, the synergy of UCK2 and PSAT1 leads to urea accumulation and its degradation into volatile ammonia (NH_3), as mapped through KEGG pathways.*

Conclusion: *These results position E-TRIPOD as a robust biomarker that turns genomics into an actionable diagnostic tool. This synergy establishes a theoretical framework for the development of biophysically grounded electronic nose technologies, setting a new potential standard for non-invasive screening.*

Keywords: *Lung adenocarcinoma, E-TRIPOD, TRIPOD, Pyrimidine Metabolism, Electronic Nose (e-Nose), Volatile Organic Compounds (VOCs).*

Introduction

Recently was proposed the transcriptomic signature known as TRIPOD [1], designed to identify volatile metabolites for the early detection of lung cancer. However, this model needs to be extended to integrate the complete biosynthetic pathways of lung adenocarcinoma. Therefore, this work presents an expansion called E-TRIPOD (Expanded-TRIPOD). Based on the analysis of 11 independent cohorts and 108 previous genetic signatures, this model reveals that tumor aggressiveness is orchestrated by a highly coordinated carbon and energy supply axis.

It is important to note that PSAT1 acts as the anaplerosis engine, diverting glycolysis precursors to fuel high-demand biosynthetic processes. This extends the scope of the original model toward other critical metabolic routes. Additionally, UCK2 (Uridine-Cytidine Kinase 2) was incorporated as a fundamental pillar of the signature, as it consumes the resources generated by the serine pathway.

In this way, the signature evolves from observing isolated events

to a molecular flow map: captured glucose is processed by PSAT1 to provide carbon equivalents, which UCK2 then transforms into the structural building blocks needed for CDC20 to execute cell division. This synergy gives E-TRIPOD superior robustness, prioritizing biochemical functionality over the simple accumulation of genes. Finally, this expansion allows for the informed design of an electronic nose (e-Nose). Such technology can detect the specific chemical footprint generated by the coordinated activity of PSAT1 and UCK2, even in preclinical stages.

Current State of the Art

The search for RNA-based biomarkers has reached significant milestones, such as the meta-analysis by Ge et al. [2]. However, this extensive work reveals that most signatures have redundant proliferation markers and lack a mechanistic axis connecting gene expression with exhaled metabolic byproducts. This biological fragmentation explains the low clinical implementation of these signatures; they identify genes ("what"), but omit the pathways that generate detectable external signals ("the how").

In parallel, the TRIPOD model was recently introduced [1], selecting a triad of drivers (COL11A1, CDC20, and PSAT1) for their role as biological bottlenecks linked to the production of volatile compounds. Recent studies have also explored the complexity of the tumor microenvironment from different genomic angles: Chen et al. [5] established robust signatures based on immune phenotypes and non-coding RNA (lncRNA) networks, while Liu et al. [6] identified specific biomarkers based on alternative splicing and factors like QKI.

Despite the technical depth of these studies, they focus exclusively on transcriptomic results and lack a simplified model that mechanistically connects genomics with real-time metabolite production.

The E-TRIPOD model emerges to bridge this research gap by strengthening prognostic potential through the integration of pyrimidine and serine pathways. The objective of this work was to characterize E-TRIPOD as a biological standard for preventive screening. We demonstrate that this integration not only defines mortality risk but also provides the functional basis to conceptualize lung cancer diagnosis into a non-invasive process, establishing a definitive bridge between bioinformatics and clinical implementation via potential e-Nose technology.

Methods

This study builds on the evolution of the TRIPOD model [1], which originally selected a triad of key biological drivers: COL11A1 (extracellular matrix remodeling), CDC20 (cell cycle regulator), and PSAT1 (the engine of tumor anaplerosis and bioenergetics). To strengthen the transition to the E-TRIPOD extended model and ensure its clinical use, the methodology followed four sequential phases: 1) model selection through consensus meta-analysis; 2) translating gene expression into biochemical flows using RSEA and KEGG; 3) validating genomic co-occurrence in real patient cohorts; and 4) evaluating the signature's prognostic potential and statistical significance.

1. Model Selection (E-TRIPOD)

The expansion into E-TRIPOD was finalized by integrating LDHA and UCK2. The goal was to map the pyrimidine salvage pathway, an enzymatic hub essential for tumor survival in harsh environments. This strategic selection was validated against 108 transcriptomic signatures compiled by Ge et al. [2], based on the criteria standards:

- Consensus Validation: CDC20's position as a universal driver of aggression was validated by the fact that it was found to be the most prevalent gene in the literature (found in 28 out of 108 signatures).
- Supply Axis: UCK2 was incorporated. It acts as the nucleotide synthesis node that consumes the carbon resources generated by PSAT1.
- The Metabolic Switch: LDHA was included to monitor the shift in glycolytic flow. While standard literature identifies SLC2A1 (GLUT1) as the fuel entry point, E-TRIPOD focuses on how that flow is processed into biosynthesis and salvage pathways.

2. Translating Gene Expression into Biochemical Flows

Unlike standard genomic analyses that look at genes in isolation, we applied Reaction Set Enrichment Analysis (RSEA) [7]. This method uses the Recon3D human metabolism reconstruction [8] as a computational framework. By integrating E-TRIPOD expression levels into the 13,543 reactions of Recon3D, we quantified actual metabolic flux. This process turned transcriptomic data into an enzymatic activity signature. These flows were then projected onto KEGG pathways (hsa00260 and hsa00240) [9], revealing that reactions R00512 and R01665 act as bottlenecks. These lead to the accumulation of β-Alanine and Urea, which are precursors of detectable ammonia (NH_3).

3. Real Data Validation

We performed a genomic consistency analysis using the cBioPortal for Cancer Genomics to ensure clinical validity and translational relevance [10]. We examined molecular profiles from the TCGA (The Cancer Genome Atlas) cohort's Lung Adenocarcinoma dataset (PanCancer Atlas), which includes over 500 multi-omic patient samples [11].

4. Survival Analysis and Prognostic Potential

The clinical relevance and accuracy of the model were determined using Kaplan-Meier survival analysis [12]. This approach effectively handled filtered observations while estimating survival probability over time. To quantify the impact of the E-TRIPOD signature, we calculated the Hazard Ratio (HR) [13] using a proportional hazards model, along with the P-value (log-rank) to assess the statistical separation of the curves. We focused on the classifier's ability to group patients into high and low risk, especially in early stages. This approach effectively handled filtered observations while estimating survival probability over time.

Together, these four methodological steps ensure that E-TRIPOD is more than just a statistical tool. They provide the evidence need-

ed to develop e-Nose technology, as described later. In this way, the methodology builds a robust bridge between genomic big data and the design of a non-invasive diagnostic system specialized for lung adenocarcinoma.

Results

By comparing the meta-analysis by Ge et al. [2] with the E-TRIPOD model, we identified that CDC20 and UCK2 operate in strict biochemical synergy. The data reveal that the accelerated cell division driven by CDC20 activity is biologically impossible without the massive supply of nucleotides provided by the "salvage pathway". In this context, UCK2 acts as the rate-limiting enzymatic factor, dictating the tumor's replication capacity (details later).

A critical finding of this study was the detection of a phenomenon we called the Anaplerosis Switch. Unlike conventional Warburg theory, which predicts a systematic overexpression of lactate dehydrogenase, our model detected a modulation of LDHA (LogFC = −1.27) in favor of a drastic upregulation of PSAT1 (LogFC = +3.21). This result indicates that the tumor prioritizes diverting glucose precursors toward building biomass and essential amino acids through the serine pathway, rather than simple fermentation into lactate. This convergence between genetic recurrence (the "what") and metabolic flux (the "how") gives E-TRIPOD a mechanistic robustness that justifies its use as a basis for screening via volatile nitrogenous signals.

When exploring the biochemical architecture through RSEA analysis, we determined that the pyrimidine metabolism subsystem (RN00240) is the main driver of the model's efficiency. Identifying enzymatic reactions R00512 and R01665, under UCK2 control, confirms a nucleotide salvage route that sustains accelerated DNA replication. This flow not only ensures the tumor's biosynthetic autonomy but also generates the byproducts (urea) that allow for detection using e-Nose technology.

The model's predictive power was validated through Kaplan-Meier survival analysis in the TCGA cohort. The multigene classifier showed a drastic and statistically significant separation between risk groups, reaching a Hazard Ratio (HR) of 2.1 (95% CI: 1.75–2.53; P=8.8×10−16) [Figure 1A]. Patients in the high-expression group had a median survival of just 54.57 months, compared to 124 months in the low-expression group. Robustness analysis (Auto-cutoff plot) confirmed that this statistical significance remains stable despite the biological variability of the samples [Figure 1B].

The connection between the genome and the exhalome is strengthened by observing in KEGG that this metabolic processing leads to the accumulation of urea and β-alanine. Figure 2 shows how the model captures the glycolytic shift toward serine biosynthesis mediated by PSAT1. This serine acts as fuel for pyrimidine metabolism (Figure 3), where UCK2 overexpression activates a salvage pathway that ends in nitrogenous byproducts.

Specifically, the model's functionality is anchored in the pyrimidine salvage route (hsa00240), where UCK2 acts as the flow regulator. As illustrated in Figure 3, this process is not an isolated metabolic sink; instead, it feeds a degradation cascade that converges in the production of β-alanine and urea. The hydrolysis of urea within the lung microenvironment releases volatile ammonia (NH3), turning gene overexpression into a specific chemical noise detectable by the proposed e-Nose system.

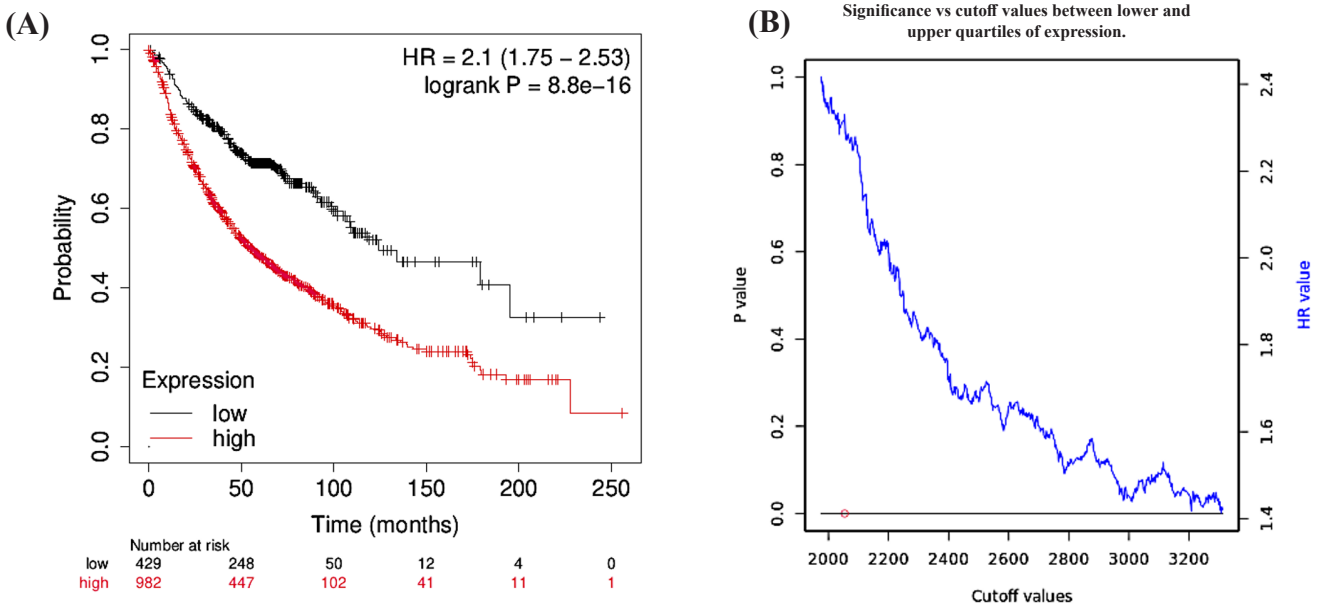

Figure 1. *Clinical validation and robustness of the E-TRIPOD model.*

(A) Kaplan-Meier survival curve showing a clear stratification of patients based on the five-gene signature. An extremely low P-value and a Hazard Ratio (HR) of 2.1 define this model as a high-lethality biomarker.

(B) Auto-cutoff optimization plot. These results show that the prognostic value of E-TRIPOD is an intrinsic tumor trait, not a data selection artifact, as shown by the stable, low P-values across cutoffs.

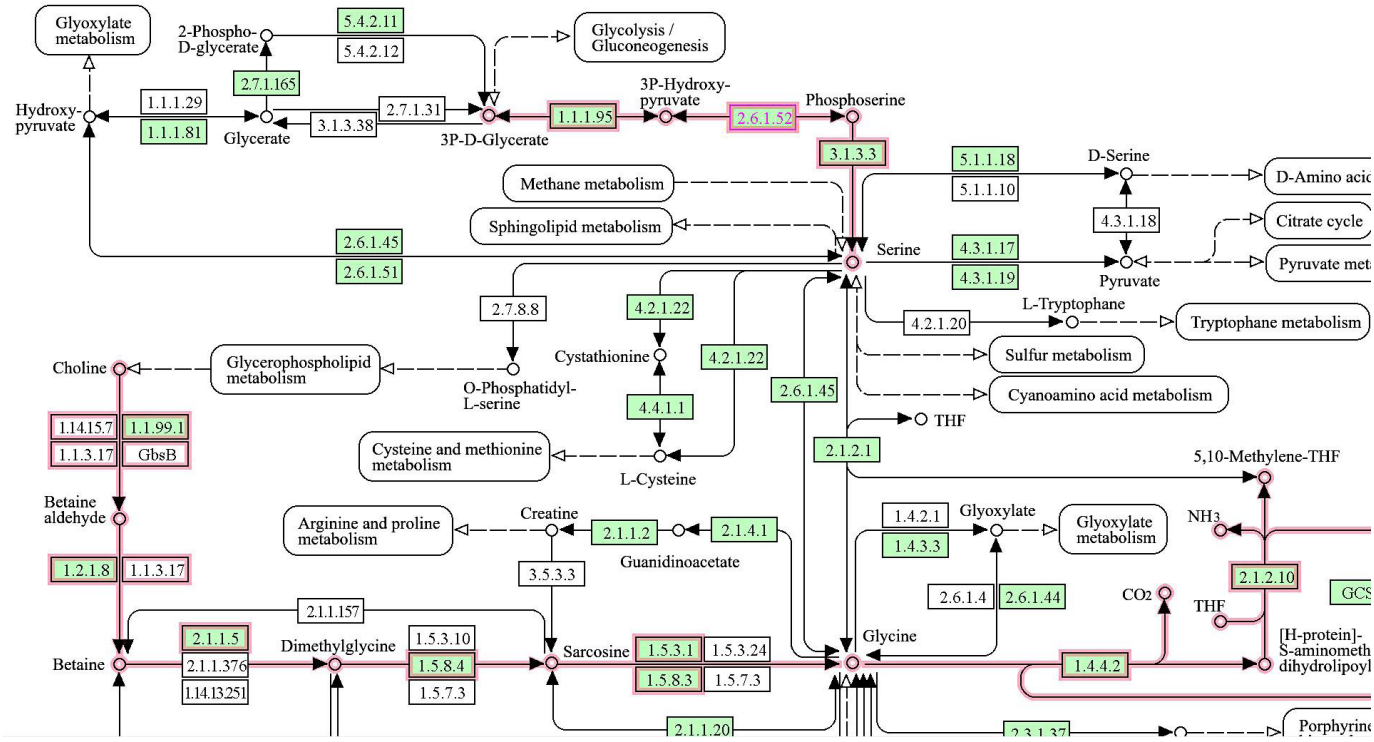

Figure 2.. *Reprogramming of the serine pathway (KEGG hsa00260). Detail of the PSAT1 enzymatic node (highlighted in fuchsia) connecting glycolysis with the production of essential amino acids for tumor growth.*

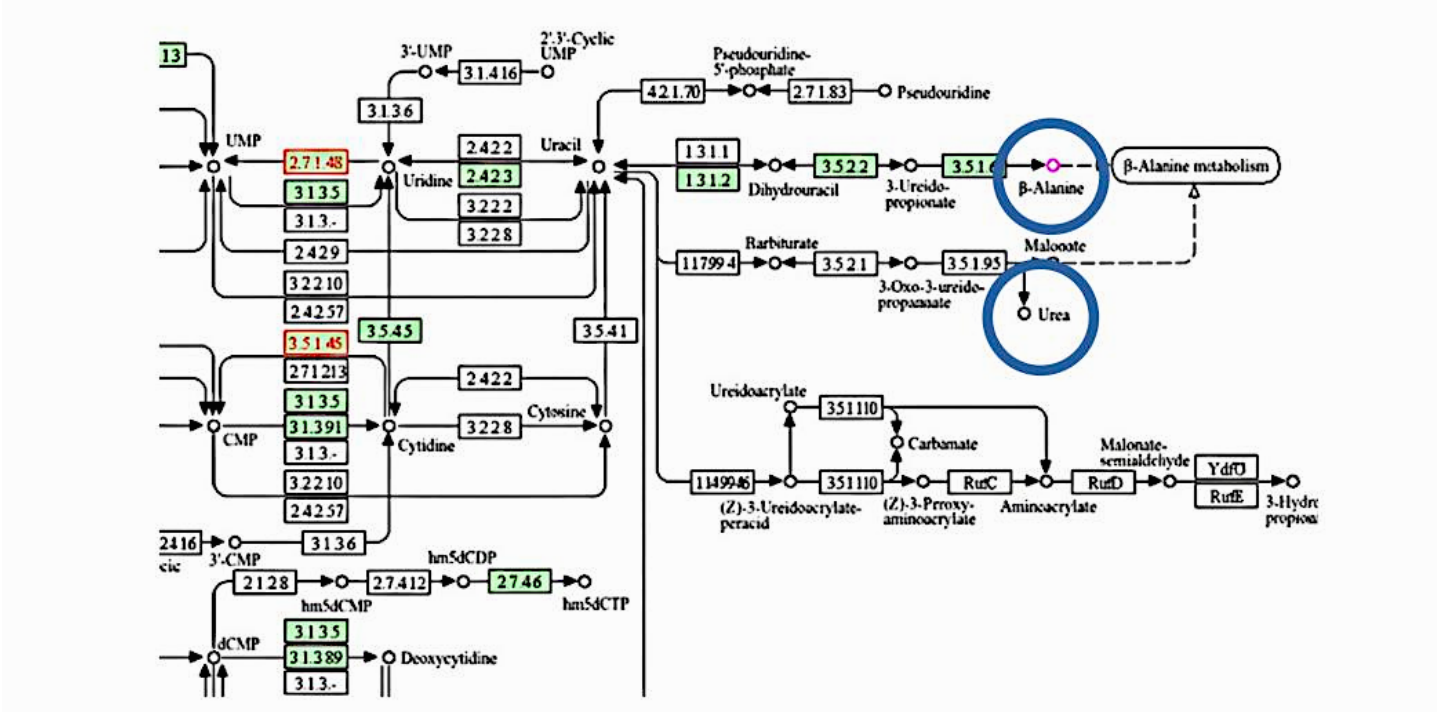

Figure 3. *Mechanism of nitrogenous byproduct generation through the pyrimidine salvage pathway. The strategic position of UCK2 (pink boxes) is highlighted, showing its role in the metabolic flow that leads to the formation of β-alanine and urea (blue circles), which are precursors of the gaseous signal (NH3).*

We used the Lung Adenocarcinoma cohort (TCGA, PanCancer Atlas, n>500) for cross-validation in cBioPortal (Figure 4) in order to validate these results in actual biological samples. The results revealed a significant trend toward genetic co-occurrence ($p<0.05$) among the five components of the signature. Scatter plots for key drivers such as PSAT1 (Figure 4A) and COL11A1 (Figure 4B) show a consistent distribution of alterations that support the model's robustness. Furthermore, OncoPrint analysis confirmed that gain in copy number and increased mRNA are directly linked to lower overall survival, keeping the PSAT1-UCK2 axis as a high-activity node across various patient subgroups. This massive clinical evidence reinforces our results, turning bioinformatics observations into a robust biomarker validated in the genomic data real.

Finally, to confirm the biological cohesion of the model, we conducted a functional enrichment analysis using g:Profiler [Figure 5]. Uridine kinase activity (Padj = 1.99×10−1), serine metabolism (Padj = 9.80×10−1), and the regulation of the anaphase-promoting complex (Padj = 4.48×10−1) were highlighted in the results as exhibiting a significant convergence in processes essential for tumor progression. E-TRIPOD functions as a coordinated functional unit, where nucleotide and amino acid production synchronize with the cell cycle to sustain adenocarcinoma aggressiveness, as evidenced by the agreement between Reactome and Gene Ontology (GO) terminology.

Discussion

The development of the E-TRIPOD model marks a major shift from previous transcriptomic studies [1,2]. While most signatures documented by Ge et al. [2] focus on redundant proliferation genes, our model integrates pyrimidine metabolism as a core predictor of aggressiveness. The metabolic condition of the tumor in its early phases is captured by this structure. This capacity was confirmed by cBioPortal data, which indicate that the development of adenocarcinoma is inherently connected to the generation of exhalable nitrogenous compounds.

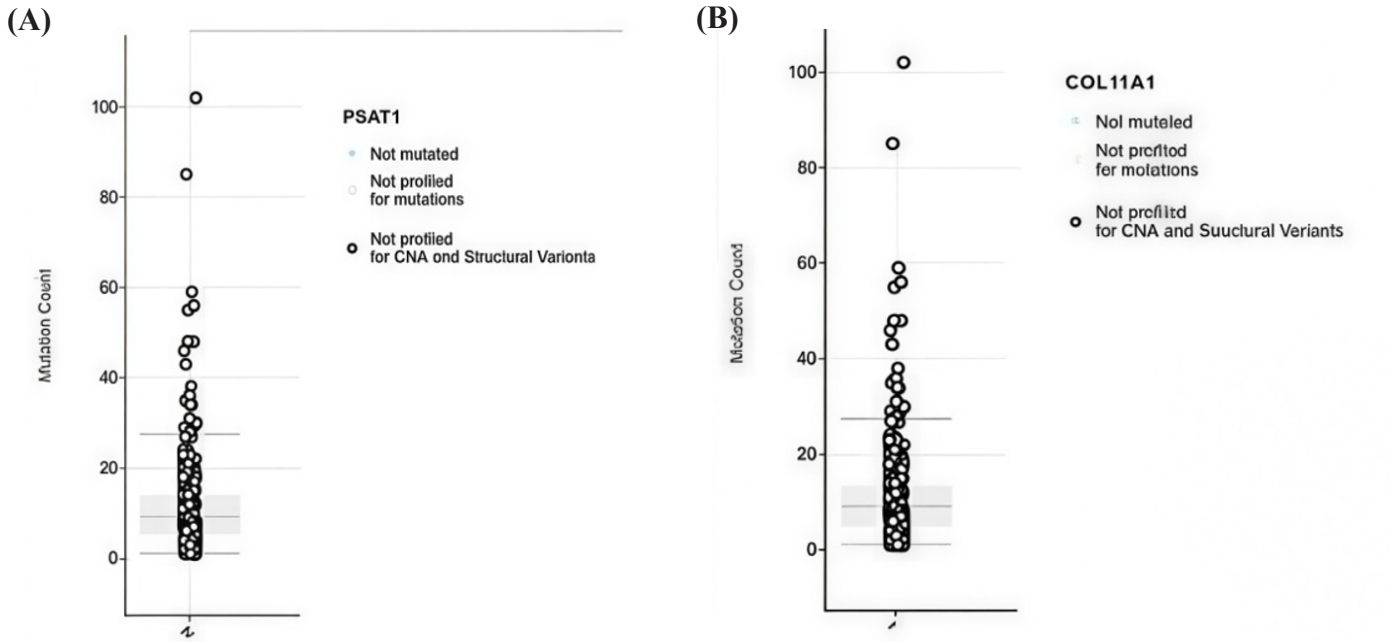

Figure 4. *Genomic alterations and co-occurrence profile in the TCGA cohort. Mutation load and copy number variations are shown for two core pillars of the model: (A) PSAT1 and (B) COL11A1. Data distribution in real-world patients (n>500) proves that E-TRIPOD alterations are not isolated events. Instead, they represent coordinated genomic states consistently found in clinical populations. This validates the model's potential as a target for exhalome-based screening.*

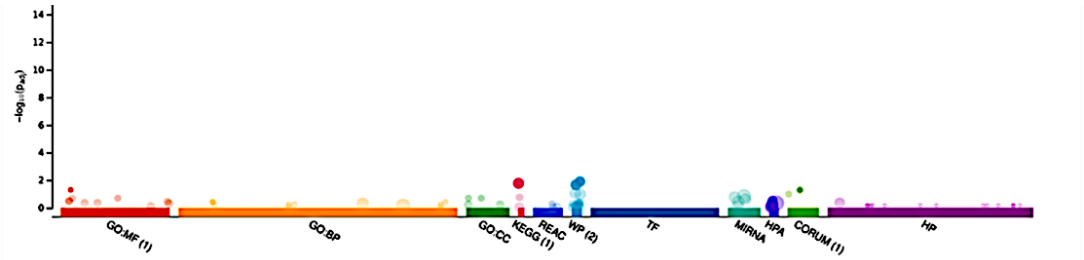

Figure 5. *Functional and ontological enrichment analysis of the E-TRIPOD model (g:Profiler).*

Analyzing the Hallmarks of Cancer [14] revealed that LDHA and PSAT1 are significantly enriched in the Reprogramming of Energy Metabolism, with an Odds Ratio of 13.79 (P=0.022, Figure 6). This finding is crucial; it confirms that E-TRIPOD does not just identify isolated genes. Instead, it captures the essential metabolic vulnerabilities that define lung cancer aggressiveness. Unlike the approaches of Chen et al. [5] and Liu et al. [6], which dive into the structural complexity of the microenvironment and splicing, E-TRIPOD offers a simplified, functional path. By focusing on the PSAT1-UCK2 axis, we successfully mapped the nitrogen flow that ends in urea production and its eventual degradation into ammonia (NH3).

This mechanistic foundation is what bridges the gap between bioinformatics and clinical use. e-Nose technology finds its biological basis in E-TRIPOD: the ammonia detected in breath is not random noise. It is the physical messenger of a synchronized tumor machinery. By showing that the co-expression of these five genes forms a functional unit in massive cohorts (n>500), we set a new standard for preventive biomarkers. E-TRIPOD allows lung cancer detection to evolve from late-stage X-ray observations toward a digital reading that is immediate, robust, and, above all, biologically grounded.

Design of the Electronic Nose (e-Nose)

The E-TRIPOD model architecture strengthens diagnostic capacity by linking serine biosynthesis with the pyrimidine salvage pathway. Functional analysis in g:Profiler showed a synchronized machinery where metabolism directly fuels cell division and tumor microenvironment remodeling. Validation via Recon3D allowed us to translate the expression of these genes into core reactions of the RN00240 subsystem (P=1.58×10−16), confirming that the tumor uses this metabolic flow for efficient biomass construction.

By tracking this flow, we identified β-Alanine (C00099) and Urea (C00086) as the terminal byproducts of the PSAT1-UCK2 axis. This finding provides the theoretical pillar for e-Nose engineering: urea hydrolysis releases Ammonia (NH3), which becomes the traceable smellprint in the patient's breath. This biological base closes the mechanistic gap that limited pioneering studies [15,16], which lacked a causal link between the genome and the exhalome. While current engineering focuses on improving deep learning algorithms [17] or devices like the Aeonose [18], the E-TRIPOD model provides the biological justification needed to optimize detection sensitivity in future clinical applications.

The proposed technical implementation envisions a metal oxide sensor (MOS) array. These are specifically calibrated to detect ammonia and nitrogenous derivatives. By processing the exhaled profile through artificial intelligence, the device transforms transcriptome complexity into an immediate digital reading. This synergy creates a scalable platform for mass preventive screening. It turns lung cancer detection into a process that is non-invasive, low-cost, and, most importantly, biologically grounded.

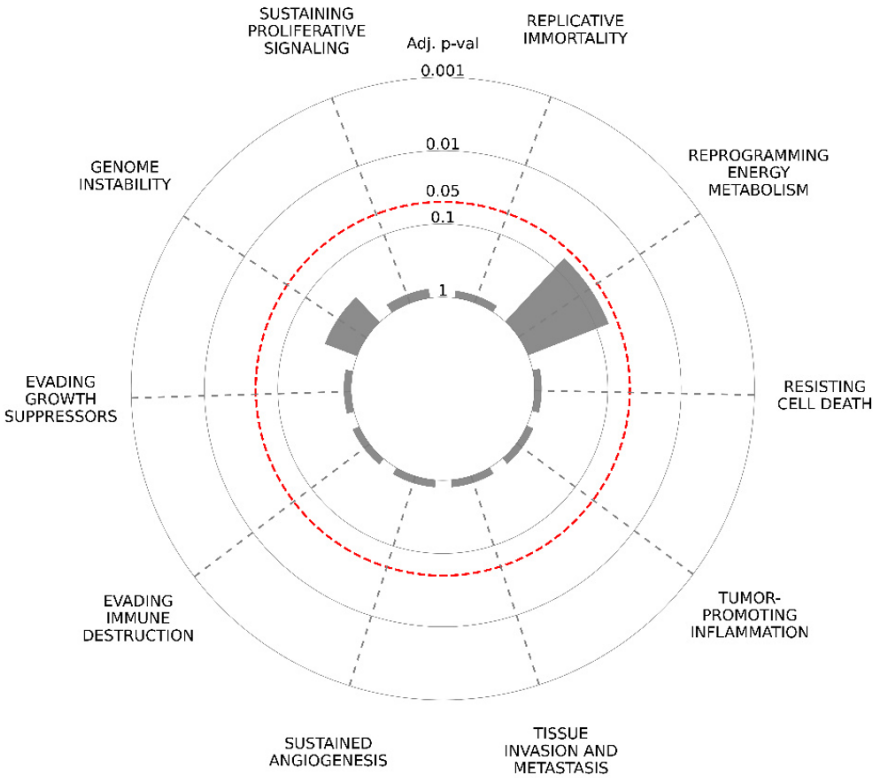

Figure 6. *Mapping the E-TRIPOD model onto the Hallmarks of Cancer. The plot shows how the signature functionally converges toward the Reprogramming of Energy Metabolism. This validates the PSAT1-LDHA axis as the main engine driving the tumor phenotype.*

Conclusions

This research characterizes the E-TRIPOD model as a high-resolution transcriptomic signature for lung adenocarcinoma. By integrating structural, proliferative, and metabolic drivers with UCK2 enzymatic activity, we showed that tumor aggressiveness is the result of a coherent molecular orchestration, detectable even in its earliest stages. The clinical robustness of this model offers a powerful yet simplified stratification tool.

Detailed analysis of KEGG pathways identified that the metabolic flow governed by E-TRIPOD, specifically within the pyrimidine metabolism axis, ends in urea accumulation and its degradation into ammonia (NH3). This finding provides a scientific rationale for the development of biologically grounded electronic noses, moving away from black box empirical models. It ensures that gas detection could be a true reflection of tumor enzymatic activity rather than an empirical process.

This approach moves past the limitations of previous studies that failed to create a non-invasive diagnostic output. By integrating serine and pyrimidine pathways, we bridge the gap between bioinformatics and potential clinical practice through volatile metabolite detection. Consequently, the E-TRIPOD model establishes the theoretical and technical feasibility of identifying lung adenocarcinoma in early stages through a simple breath test.

Study Limitations

Although the E-TRIPOD model demonstrates high statistical robustness and a clear mechanistic link between transcriptomics and metabolic flux, this study has limitations that must be acknowledged. First, it is a computational and hypothesis-generating framework based on retrospective data from the TCGA and GEO cohorts. While the biochemical pathways leading to ammonia (NH3) production are well-supported by Recon3D and KEGG mapping, these findings have not yet been validated in a prospective clinical trial. Second, the proposed application for e-Nose technology remains theoretical; further experimental research is required to correlate physical sensor readings with the intracellular activity of the PSAT1-UCK2 axis in real-time clinical settings.

Conflicts of Interest

The authors declare no conflict of interest and received no specific funding for this work.

References

1. Isea R. A breath-based transcriptomic signature for lung cancer screening. J Clin Pract Med Res. 2026; 2: 41-46.
2. Ge X, Xu H, Weng S, Zhang Y, Liu L, et al. Systematic analysis of transcriptome signature for improving outcomes in lung adenocarcinoma. J Cancer Res Clin Oncol. 2023; 149: 8951-8968.
3. Ahuja S, Zaheer S. Molecular Mediators of Metabolic Reprogramming in Cancer: Mechanisms, Regulatory Networks, and Therapeutic Strategies. Immunology. 2026; 177: 1-43.
4. Wu Y, Jamal M, Xie T, Sun J, Song T, et al. Uridine-cytidine kinase 2 (UCK2): A potential diagnostic and prognostic biomarker for lung cancer. Cancer Sci. 2019; 110: 2734-2747.
5. Chen Q, Ma J, Wang X, Zhu X. Identification of prognostic candidate signatures by systematically revealing transcriptome characteristics in lung adenocarcinoma with differing tumor microenvironment immune phenotypes. Aging (Albany NY). 2022; 14: 4786-4818.
6. Liu Y, Zhang C, Zheng Y, Jin S, Niu J, et al. Transcriptomic analysis reveals lung cancer and subtype-specific alternative splicing biomarkers regulated by RNA-binding proteins. Mol Ther Nucleic Acids. 2025; 36: 102681.
7. Yarıcı M, Cantürk F, Dursun S, Aydın HN, Karabekmez ME. RSEA: A Web Server for Pathway Enrichment Analysis of Metabolic Reaction Sets. Biotechnol Bioeng. 2025; 122: 2251-2258.
8. Brunk E, Sahoo S, Zielinski DC, Altunkaya A, Dräger A, et al. Recon3D enables a three-dimensional view of gene variation in human metabolism. Nat Biotechnol. 2018; 36: 272-281.
9. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000; 28: 27-30.
10. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012; 2: 401-404.
11. Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol (Pozn). 2015; 19: A68-A77.
12. Rich JT, Neely JG, Paniello RC, Voelker CC, Nussenbaum B, et al. A practical guide to understanding Kaplan-Meier curves. Otolaryngol Head Neck Surg. 2010; 143: 331-336.
13. Trkulja V, Hrabač P. Hazard ratio. Croat Med J. 2019; 60: 293-296.
14. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000; 100: 57-70.
15. Tran VH, Chan HP, Thurston MC, Jackson P, Lewis CR, et al. Breath analysis of lung cancer patients using an electronic nose detection system. IEEE Sens J. 2010; 10: 1514-1518.
16. D'Amico A, Pennazza G, Santonico M, Martinelli E, Roscioni C, et al. An investigation on electronic nose diagnosis of lung cancer. Lung Cancer. 2010; 68: 170-176.
17. Lee MR, Kao MH, Hsieh YC, Sun M, Tang KT, et al. Cross-site validation of lung cancer diagnosis by electronic nose with deep learning: a multicenter prospective study. Respir Res. 2024; 25: 203.
18. van de Goor R, van Hooren M, Dingemans AM, Kremer B, Kross K. Training and validating a portable electronic nose for lung cancer screening. J Thorac Oncol. 2018; 13: 676-681.