



Landscape and clinical impact of metabolic alterations in non-squamous non-small cell lung cancer

Anna V. Ivanina Foureau¹, Wei Sha², David M. Foureau¹, James T. Symanowski², Carol J. Farhangfar¹, Kathryn F. Mileham³

¹Translational Research, Levine Cancer Institute, Atrium Health, Charlotte, NC, USA; ²Cancer Biostatistics, Levine Cancer Institute, Atrium Health, Charlotte, NC, USA; ³Thoracic Medical Oncology, Levine Cancer Institute, Atrium Health, Charlotte, NC, USA

Contributions: (I) Conception and design: AV Ivanina Foureau, W Sha, DM Foureau, CJ Farhangfar, KF Mileham; (II) Administrative support: AV Ivanina Foureau; (III) Provision of study materials or patients: AV Ivanina Foureau, W Sha, DM Foureau; (IV) Collection and assembly of data: AV Ivanina Foureau, DM Foureau; (V) Data analysis and interpretation: AV Ivanina Foureau, W Sha, DM Foureau, JT Symanowski; (VI) Manuscript writing: All authors; (VII) Final approval of manuscripts: All authors.

Correspondence to: Kathryn F. Mileham. Levine Cancer Institute, Atrium Health, 1021 Morehead Medical Drive, Charlotte, NC 28204, USA.

Email: kathryn.mileham@atriumhealth.org.

Background: Metabolomics studies to date have described widespread metabolic reprogramming events during the development of non-squamous non-small cell lung cancer (NSCLC). Extending far beyond the Warburg effect, not only is carbohydrate metabolism affected, but also metabolism of amino acids, cofactors, lipids, and nucleotides.

Methods: We evaluated the clinical impact of metabolic reprogramming. We performed comparative analysis of publicly available data on non-squamous NSCLC, to identify consensus altered metabolic pathways. We investigated whether alterations of metabolic genes controlling those consensus metabolic pathways impacted clinical outcome. Using the clinically annotated lung adenocarcinoma (LUAD) cohort from The Cancer Genome Atlas, we surveyed the distribution and frequency of function-altering mutations in metabolic genes and their impact on overall survival (OS).

Results: We identified 42 metabolic genes of clinical significance, the majority of which (37 of 42) clustered across three metabolic superpathways (carbohydrates, amino acids, and nucleotides) and most functions (40 of 42) were associated with shorter OS. Multivariate analyses showed that dysfunction of carbohydrate metabolism had the most profound impact on OS [hazard ratio (HR) =5.208; 95% confidence interval (CI): 3.272 to 8.291], false discovery rate (FDR)- $P \leq 0.0001$, followed by amino acid metabolism (HR =3.346; 95% CI: 2.129 to 5.258), FDR- $P \leq 0.0001$ and nucleotide metabolism (HR =2.578; 95% CI: 1.598 to 4.159), FDR- $P = 0.0001$. The deleterious effect of metabolic reprogramming on non-squamous NSCLC was observed independently of disease stage and across treatments groups.

Conclusions: By providing a detailed landscape of metabolic alterations in non-squamous NSCLC, our findings offer new insights in the biology of the disease and metabolic adaptation mechanisms of clinical significance.

Keywords: Non-squamous non-small cell lung cancer; cancer metabolites; metabolic gene mutations; post-genomic pipeline

Submitted May 18, 2022. Accepted for publication Nov 06, 2022.

doi: [10.21037/tlcr-22-377](https://doi.org/10.21037/tlcr-22-377)

View this article at: <https://dx.doi.org/10.21037/tlcr-22-377>

Introduction

Cancer-associated metabolic reprogramming, a hallmark of tumorigenesis, refers to the biochemical processes that allow cells to retain energy, balance oxidation-reduction, and synthesize the extra building blocks required for cancer cell development, proliferation, and survival (1). In non-small cell lung cancer (NSCLC), several metabolic alterations have been shown to support cancer progression. Driver mutations such as EGFR, ALK, and KRAS have been found in 30–40% of patients with NSCLC (2,3). Mutations in EGFR can promote metabolic remodeling in NSCLC by increasing glycolysis and altering pyrimidine biosynthesis (4). ALK rearrangements lead to upregulation of glucose metabolism, which is usually associated with more aggressive cancer phenotypes (5). Mutations in KRAS have been associated with upregulation of glucose uptake and the Warburg effect (6). Metabolic reprogramming is also a therapeutic target in non-squamous NSCLC. The clinical utility of antimetabolites such as pemetrexed and fluoropyrimidines is partly attributable to their ability to interfere with the increased metabolic demands of cancer cells for nucleotide biosynthesis and DNA replication (7–9).

Several groups have explored metabolic reprogramming role in non-squamous NSCLC by using mass spectrometry (10–12) and produced comprehensive maps of altered metabolites for different disease stages, demonstrating clear differences between malignant and benign tissues. Wikoff *et al.* (10), analyzed cellular metabolites from tissue samples of patients with stages IA-B lung adenocarcinoma (LUAD) using gas chromatography-time-of flight mass spectrometry (GC-TOF-MS) and Moreno *et al.* (11) analyzed cellular metabolites from tissue samples of patients with stages I-III LUAD using ultra performance liquid chromatography tandem mass spectrometry (UPLC/MS/MS) with gas chromatography mass spectrometry (GC-MS). Results from both studies showed broad metabolic alterations between malignant and non-malignant LUAD tissue, which affected 40% of metabolic functions that were investigated.

However, one key question remains: Do all metabolic alterations associated with LUAD equally impact clinical outcome or are some merely bystander events resulting from rapid cell growth? To answer this question, we applied a systems biology approach to (I) define consensus altered metabolic functions in non-squamous NSCLC, (II) identify the pool of genes directly controlling those metabolic functions using Kyoto Encyclopedia of Genes and Genomes (KEGG), and (III) interrogate The Cancer Genome Atlas (TCGA) to identify the metabolic alterations associated

with overall survival (OS) in LUAD. Among consensus metabolic alterations, we identified 42 metabolic genes of clinical significance. The majority of them clustered within three metabolic superpathways (carbohydrates, amino acids, and nucleotides) and were associated with shorter OS. We present the following article in accordance with the STARD reporting checklist (available at <https://tclr.amegroups.com/article/view/10.21037/tclr-22-377/rc>).

Methods

Ethical statement

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Metabolic data normalization and comparison

Quantitative data from two independent retrospective studies by Wikoff *et al.* (10) and Moreno *et al.* (11), that compared metabolic differences between malignant and benign NSCLC tissues, were compared. The Wikoff study surveyed 39 stages IA-B LUAD tumors and identified 183 altered metabolites; the Moreno study surveyed 33 stages IA-IIIB LUAD tumors and identified 851 altered metabolites (Table S1). To compare metabolic alterations between these two studies, we compiled the reported data, normalized the cellular metabolite nomenclature across studies using the KEGG ENZYME database (13), and calculated (or compiled when such data was available) \log_2 fold changes between malignant and benign NSCLC tissues. Cellular metabolites with \log_2 fold-change ranging from -1.52 to $+3.20$ and a false discovery rate (FDR)-adjusted P value <0.05 in at least one of the studies were considered significantly altered in non-squamous NSCLC.

Metabolic pathways annotation and gene ontology

To group cellular metabolites into metabolic pathways and superpathways, based on the biosynthesis or metabolism of related compounds, we used the KEGG PATHWAY database (RRID: SCR_012773) (14). We used the KEGG ENZYME database to identify each catabolic and anabolic enzyme associated with those metabolic pathways and the genes coding them (13), thereby converting the altered metabolic pathways and superpathways into gene signatures.

Whole exome sequencing (WES) data were used to identify function-altering mutations of metabolic enzyme coding genes. Transcriptomic (RNAseq) data were used to

confirm and quantify metabolic gene expression. WES and RNAseq data from non-squamous NSCLC (LUAD cohort) were obtained from TCGA v21.0 (RRID: SCR_003193) (<https://portal.gdc.cancer.gov/>). Function-altering mutations were defined as non-synonymous mutations with an allele frequency >10%, a SIFT score <1, and a PolyPhen score >0. Included in the list of non-synonymous mutations were mutations modifying upstream or downstream genes, as well as splice region, frameshifts, mutations in introns, and 3' and 5' untranslated regions. Metabolic gene transcript abundance estimates calculated by the Cancer Genome Analysis group at the Broad Institute using RNA-Seq by Expectation Maximization (RSEM) (15) were compiled from GDAC Firehose (gdac_rnaseqv2_genes_RSEM_normalized_Level_3, 2016-02-18). Relative expression of metabolic genes are reported as number of transcripts per million mapped reads.

Clinical annotation of metabolomic and genomic non-squamous NSCLC cohorts

For each metabolomic cohort, we compiled demographic and clinical data from the 39 patients surveyed in the Wikoff study and 33 in the Moreno study (10,11). For the genomic cohort, demographic, clinical, and treatment information were compiled using GDAC Firehose (<http://gdac.broadinstitute.org/>), a TCGA data analysis infrastructure developed by the Broad Institute. The TCGA LUAD cohort contained 585 patients; however, clinical information was missing for 63 patients and thus were removed from the analyses; 522 TCGA LUAD patients were included in the final analyses (Table S1). All patients from both metabolomic cohorts and TCGA were treatment naïve.

Statistical analysis

Descriptive statistical analysis was performed with Graphpad Prism 9. ANOVA and pair-wise comparison analysis was performed using JMP15 software (RRID: SCR_014242) (SAS Institute, Cary, NC, USA). One-way ANOVA was used to test the frequency of function-altering mutations between metabolic pathways, and pair-wise comparison was used to compare differences between each metabolic pathway. Pair-wise comparison was also used to compare differences in metabolic genes expression between mutated *vs.* non-mutated genes. Fisher's exact test was applied to determine whether demographic or clinical

factors affect the frequency of mutations in genes of clinical significance in OS. Differences were considered significant at $P < 0.05$.

Cox proportional hazards model was used to identify genes associated with survival in LUAD patients. The Benjamini-Hochberg procedure was used to control the FDR. Genes associated with survival (raw P value <0.1) were selected and grouped into their respective biochemical pathways.

For each patient, if a mutation was detected in at least one gene on the pathway, the patient was considered to have a mutation in this pathway. Cox proportional hazards model was then used to study the association between mutation on each pathway and survival time. Age, gender, smoking status, tumor stage, and therapy (chemo- or radiation therapy) were adjusted in the model. Hazard ratios and confidence intervals were calculated. Pathways with $FDR-P < 0.05$ were considered statistically significant. To determine whether the effect of pathway mutations on OS varies by cancer stages or treatment type, pathway \times stage interaction and treatment \times stage interaction effect were also included in the initial Cox regression model. Interaction terms with $P > 0.2$ were removed from the final model.

Results

Conserved patterns of metabolic alterations in non-squamous NSCLC across metabolomics studies

We compared quantitative analyses data from two retrospective studies by Wikoff *et al.* and Moreno *et al.* (10,11), comparing metabolic differences between malignant and benign NSCLC tissues. To account for methodological differences, we restricted our analyses to metabolites quantified in both studies. To compare data between these two studies, we normalized nomenclature of metabolites of reported data, compiled (or calculated) \log_2 fold changes and FDR-adjusted P values for each cellular metabolite detected in both studies. Thirty one of 183 tested (16.9%) in the Wikoff study and 49 of 581 tested (8.4%) in the Moreno study (Figure 1, $FDR-P < 0.05$) were detected at significantly different levels between normal tissue and LUAD tissue, while 27 significantly altered metabolites were shared by both studies.

In total, we found that 53 cellular metabolites were detected at different levels between malignant and benign LUAD tissues in at least one of the studies (Figure 1,

Metabolite	Superpathway	Pathway	Lung Adenocarcinoma_Wikoff log ₂ fold change	Lung Adenocarcinoma_Moreno log ₂ fold change	
alanine			0.26*	0.29*	
asparagine		Alanine, aspartate and glutamate metabolism	0.26	0.89*	
aspartate ^{1,2}	0.00		-0.27*		
glutamate ^{1,3,4}	0.49*		0.31*		
citrulline		Arginine biosynthesis	-0.58*	-0.21	
glutamine ^{5,6,7}	0.14		0.14*		
urea ^{5,6,7}	0.00		0.21*		
creatine	Amino acids		0.49*	0.52*	
ornithine ^{1,4}			-0.93*	-0.16	
proline		Arginine and proline metabolism	0.49*	0.38*	
putrescine ⁴			0.14	1.71*	
spermidine ⁴		-0.49*	0.41*		
trans-4-hydroxyproline			0.38*	0.41*	
ascorbate		Glutathione metabolism	0.68	3.20*	
cysteine			0.68*	1.01*	
glycine ⁵			0.14	0.27*	
5-oxoproline			0.14*	0.00	
fructose ⁸		Amino sugar metabolism	0.58*	1.21*	
citrate ⁷		Citrate cycle	-0.68*	-0.68*	
malate			0.38*	0.71*	
fructose	Carbohydrates	Fructose and mannose metabolism	0.26	0.54*	
mannitol			-0.49	0.37*	
sorbitol			0.49	1.16*	
3-phosphoglycerate ^{9,10}		Glycolysis, gluconeogenesis	-1.14*	-1.52*	
glucose ^{9,11}	-1.00*		-0.64*		
ribose		Pentose phosphate cycle	0.00	1.06*	
nicotinamide	Cofactors	Nicotinate and nicotinamide metabolism	0.49*	0.49*	
flavin adenine dinucleotide		Riboflavin metabolism	0.26*	0.39*	
ribose			0.85*	1.06*	
1-monopalmitin	Lipids	Fatty Acid biosynthesis	0.14*	0.32*	
2-monopalmitin			0.26	0.68*	
glycerate ⁹		Glycerolipid metabolism	0.49*	-0.92*	
glycerol			0.26*	0.42*	
glycerol 3-phosphate		0.26	0.31*		
dihydrospingosine	Sphingolipid metabolism		-0.58*	-0.47	
phosphoethanolamine			0.49	0.64*	
serine			0.14	0.23*	
adenine	Nucleotides	Purine metabolism	0.58*	0.34*	
adenosine			0.38*	0.43*	
guanosine			0.00	0.63*	
hypoxanthine			0.14	0.17*	
inosine			0.14	0.64*	
inosine 5'-monophosphate			1.81*	0.10*	
urate			0.49*	0.24*	
xanthine			1.43*	0.59*	
xanthosine			0.14	1.03*	
5,6-dihydrouracil				1.26*	1.64*
beta-alanine				0.14	0.34*
orotate	Pyrimidine metabolism	0.26	0.62*		
thymine		0.26	1.08*		
uracil		0.49*	0.60*		
uridine		0.49*	0.15*		
uridine monophosphate (5' or 3')		0.14	1.06*		

Figure 1 Metabolic alterations in non-squamous NSCLC detected by mass spectrometry across metabolomic studies. Log₂ fold changes in metabolites between malignant and non-malignant non-squamous NSCLC tissues were calculated in a comparative analysis of publicly available data from Wikoff *et al.* (10) and Moreno *et al.* (11). Red indicates metabolite concentrations with log₂ fold changes between -1.52 and 0 in tumor tissues; green indicates metabolite concentrations with log₂ fold changes between 0.1 and +1.82 in tumor tissues. Asterisk (*) indicates values that were significantly altered between malignant and non-malignant tissues (FDR-adjusted P values <0.05). Numbered annotations identify metabolites that belong to the following metabolic pathways, as determined by the KEGG annotation tool: ¹, arginine biosynthesis; ², nicotinate and nicotinamide metabolism; ³, arginine and proline metabolism; ⁴, glutathione metabolism; ⁵, purine metabolism; ⁶, pyrimidine metabolism; ⁷, alanine, aspartate and glutamate metabolism; ⁸, fructose and mannose metabolism; ⁹, pentose metabolism; ¹⁰, glycerolipid metabolism; ¹¹, amino sugar metabolism. NSCLC, non-small cell lung cancer; KEGG, Kyoto Encyclopedia of Genes and Genomes; FDR, false discovery rate.

Table 1 Genes regulating metabolic pathways altered in non-squamous non-small cell lung cancer

Metabolic superpathway	Metabolic pathway	Number of enzymes	Number of enzyme-encoding genes
Amino acids	Alanine, aspartate & glutamate metabolism	35	37
	Glutathione metabolism	31	57
	Arginine biosynthesis	13	22
	Arginine and proline metabolism	43	51
Carbohydrates	Amino sugar	44	48
	Citrate cycle	20	30
	Fructose and mannose metabolism	23	33
	Glycolysis, gluconeogenesis	43	67
	Pentose metabolism	25	30
Cofactors	Nicotinate and nicotinamide metabolism	24	35
	Riboflavin metabolism	9	8
Lipids	Fatty acid biosynthesis	15	18
	Glycerolipid metabolism	33	61
	Sphingolipid metabolism	32	49
Nucleotides	Purine metabolism	66	133
	Pyrimidine metabolism	34	57

FDR- $P < 0.05$). Across both data sets, 43 of those cellular metabolites (81.2%) were at higher concentrations in LUAD tumor tissue than in normal tissue, and 4 (7.5%) were at lower concentrations; 6 (11.3%) had discrepant results between studies. The magnitude of metabolic alterations was also very similar between the two studies (median \log_2 fold-change: 0.26, range -1.14 to $+1.81$ vs. 0.46, range -1.52 to $+3.20$; $P = 0.912$). Using the KEGG PATHWAY database, we categorized the 53 metabolites associated with LUAD into their respective 16 metabolic pathways, which clustered into 5 superpathways: amino acid, carbohydrate, cofactor, lipid, and nucleotide metabolism (Figure S1).

Evenly distributed cellular metabolism alterations in non-squamous NSCLC across metabolic superpathways

Given the similarities in metabolic alterations between the two LUAD cohorts, we hypothesized that gene mutations in those 16 metabolic pathways may be responsible for the alterations in the corresponding metabolites. Using KEGG ENZYME database, we catalogued 736 genes coding for the 490 enzymes that control the 16 metabolic pathways

(Table 1). One hundred seventy six of 736 genes (23.9%) coded for enzymes directly controlling anabolism or catabolism of the 53 cellular metabolites altered in non-squamous NSCLC, the remaining 560 genes more broadly controlling the 16 metabolic pathways of interest. The number of genes that control each metabolic pathway varied between 8 for riboflavin metabolism and 133 for purine metabolism, with a median of 42.5 genes per pathway (Figure S1).

We analyzed the WES data of the 522 patients represented in TCGA with LUAD for function-altering mutations in those 736 genes controlling cellular metabolism, and we identified 4,608 unique variants. The incidence of metabolic gene mutations was low ($6.29\% \pm 0.25\%$), and the frequency of function-altering mutations was evenly distributed across superpathways (Figure 2A). However, pair-wise comparisons of pathway mutation frequencies did show some significant differences. The glutathione metabolism pathway (GSH), part of the amino acids superpathway, had a lowest mutation frequency than other pathways of amino acids superpathway (0.74% vs. 1.66% or 1.83% , $P = 0.0034$ or $P = 0.0302$ for alanine, aspartate and glutamate metabolism or arginine and proline metabolism, respectively). Mutation frequency

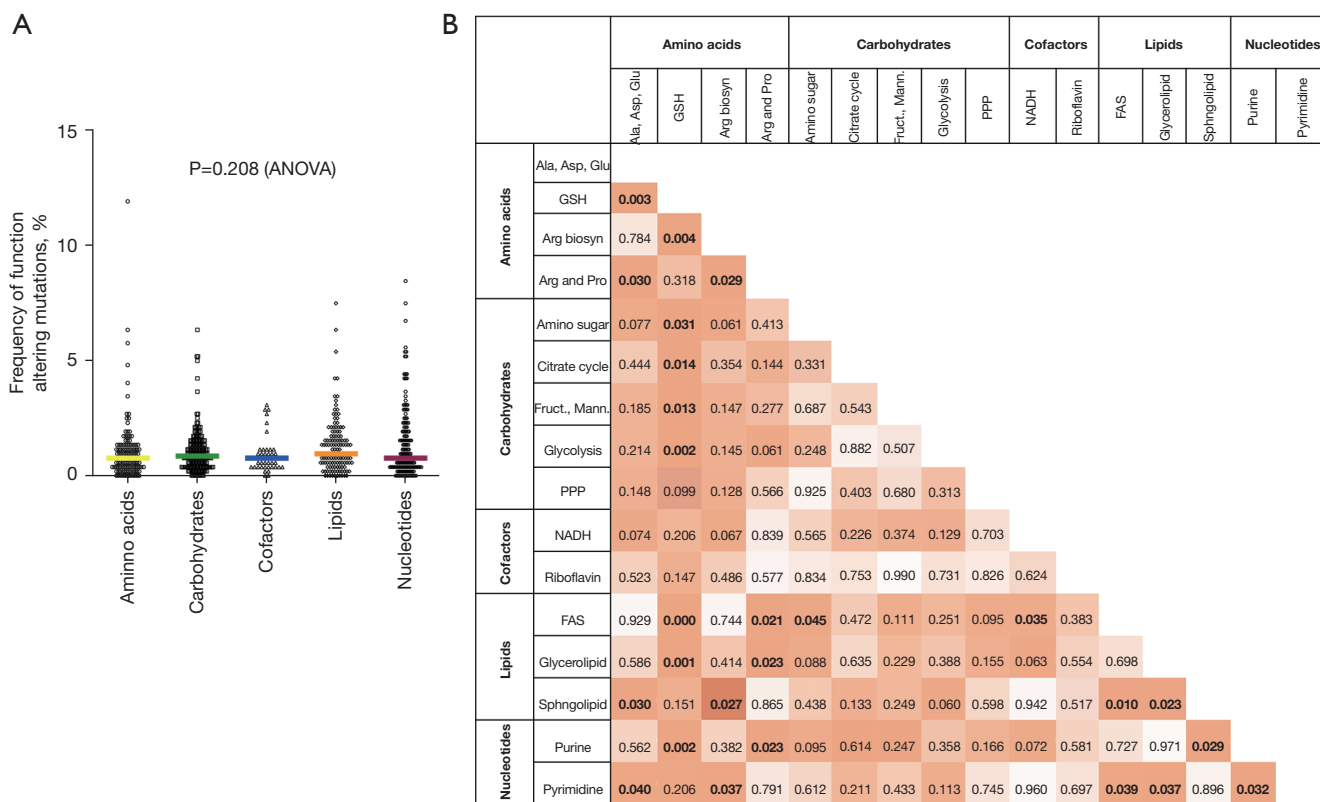


Figure 2 Frequency of metabolic gene mutations in non-squamous non-small cell lung cancer. (A) Frequency of metabolic gene mutations in lung adenocarcinoma TCGA cohort. Each point represents the frequency (%) of function-altering mutations in individual metabolic genes classified by metabolic superpathway. (B) Heatmap of pair-wise comparisons of the mutation frequency of genes among the metabolic pathways to which they belong. Values in bold are significantly different ($P < 0.05$). ANOVA, analysis of variance; Ala, Asp, Glu: alanine, aspartate and glutamate metabolism; GSH, glutathione biosynthesis; Arg biosyn, arginine biosynthesis; Arg and Pro, arginine and proline metabolism; Amino sugar, amino sugar metabolism; Citrate cycle, citric acid cycle; Fruct., Mann., fructose and mannose metabolism; glycolysis, glycolysis and gluconeogenesis; PPP, pentose metabolism; NADH, nicotinate and nicotinamide metabolism; Riboflavin, riboflavin metabolism; FAS, fatty acid biosynthesis; Glycerolipid, glycerolipid metabolism; Sphngolipid, sphngolipid metabolism; Purine, purine metabolism; Pyrimidine, pyrimidine metabolism; TCGA, The Cancer Genome Atlas.

of GSH was significantly lower than metabolic pathways of carbohydrates (0.74% vs. 1.05%; 1.30%; 1.13% and 1.26%; $P=0.0305$, $P=0.0143$, $P=0.0131$, $P=0.0015$ for amino sugar metabolism; citrate cycle, fructose and mannose metabolism or glycolysis, respectively), lower than fatty acid synthesis and glycerolipid metabolism (0.74% vs. 1.61% and 1.46%, $P=0.0003$ and $P=0.0015$ for fatty acid synthesis and glycerolipid metabolism, respectively). The frequency of mutations in GSH was also lower than the purine metabolism pathway (0.74% vs. 1.45%, $P=0.0015$) (Figure 2B). Only one gene was found mutated in more than 10% of patients—the gene encoding *CPS1*, part of the alanine, aspartate and glutamate metabolic pathway (amino

acids superpathway). *CPS1* is a mitochondrial enzyme catalyzing synthesis of carbamoyl phosphate from ammonia and bicarbonate.

Metabolic gene mutations associated with OS in non-squamous NSCLC

We estimated the median follow up time to be 23.93 months for the 522 patients using the reverse Kaplan-Meier approach (16) (Table S1). Using clinical outcome data for these patients available through TCGA, we performed univariable analyses to determine whether any function-altering mutations in the 736 metabolic genes were

associated with OS from TCGA LUAD cohort. Each gene was analyzed using a separate Cox regression model. Mutations in 42 genes were associated with OS (raw $P < 0.1$) (Figure S2); 40 were associated with shorter OS (average HR =9.46, median HR =3.74, min–max 1.80–141.83), and two genes were associated with longer OS (PDE10A: HR =0.37, $P = 0.0866$; GAD2: HR =0.30, $P = 0.0901$). The mutation frequency for genes associated with OS was not significantly different from that of other metabolic genes ($P = 0.71$; Figure S3).

We tested whether the 42 genes of clinical significance were expressed among tumors containing function altering mutations. All but one gene (GAD2) was expressed in non-squamous NSCLC tumors. Three metabolic genes (7.1%) were differentially expressed, each of them had higher transcript abundance estimate (i.e., transcripts per million estimated reads) in tumors containing function altering mutations compared with tumors with non-mutated metabolic genes: GBA (2,440±289 vs. 1,849±34); DGKG (150±25 vs. 97.1±4.1); GSS (1,653±204 vs. 1,215±20), all $P < 0.05$ (Table S2).

To validate our data in a biological context, we further investigated whether the 42 clinically significant metabolic genes identified using our genomic TCGA cohort (Figure 3), were associated with metabolic alterations detected by mass spectrometry (MS) in the metabolomic studies (Figure 1). Using the KEGG ENZYME database we found that 34 of the 42 metabolic genes of clinical significance (81%) controlled non-squamous NSCLC metabolic alteration that was identified by MS (Figure 3, Figure S4). While 23 metabolic genes (54.8%) were associated with metabolic alterations detected in both metabolomic studies (10,11), there were eleven genes (26.2%) with metabolic alterations detected only in the Moreno study (11). The remaining 8 metabolic genes (19%), coding for 6 metabolic enzymes, did not match metabolomic studies findings (Figure 3). Lastly, among the 53 non-squamous NSCLC metabolic alterations detected by MS, 11 showed no association with metabolic genes of clinical significance (Figure 1, Figure S4).

Ninety-three percent of the clinically significant metabolic genes in LUAD (38 of 42) were concentrated across three superpathways: amino acids, carbohydrates and nucleotides. Within the amino acids superpathway, non-silent mutations of genes controlling glutathione catabolism or proline levels (Figure S4B,S4D) were associated with shorter OS. Similarly, alterations to glucose, fructose and mannose metabolism (carbohydrates superpathway, Figure S4G-S4H) and uric acid biosynthesis (nucleotides

superpathway, Figure S4O) were also associated with shorter survival. Two metabolic alterations were associated with longer OS: glutamate/glutamine biosynthesis (glutamate decarboxylase 2, GAD2; Figure S4A) and cAMP/cGMP hydrolysis (phosphodiesterase 10A, PDE10A, Figure S4O). The fourteen metabolites that were not associated with clinical significance were evenly distributed between the five superpathways (Figure S4) (11).

Independent association of metabolic alterations in non-squamous NSCLC with shorter OS

To measure the relative association of each metabolic superpathway with OS, we grouped the 40 metabolic genes associated with shorter survival (Figure S2) as follows: amino acids (14 genes), carbohydrates (15 genes), cofactors (two genes), lipids (6 genes) and nucleotides (7 genes). Four genes were shared across more than one superpathway (Figure 3).

We first tested the effect of gender (male vs. female), smoking status (non-smoking vs. smoking) and disease stage (early vs. late stage) on the frequency of mutated metabolic genes in the genomic TCGA LUAD cohort, grouping clinically significant genes by superpathways. We found that smoking status significantly impacted mutation frequency of clinically relevant genes controlling amino acid and nucleotide metabolic superpathway (Table S3). In both superpathways, smokers had almost double the mutation frequency of non-smokers: [amino acid 19.2% vs. 10.8% ($P = 0.027$)] and [nucleotide 20.9% vs. 10.0% ($P = 0.004$)]. We then tested the impact of the aforementioned demographic and clinical variables on clinical outcome and found that patients presenting with later stage disease had a worse overall clinical outcome ($P < 0.0001$). No differences were observed among the other variables. On the basis of these results, we adjusted our model for gender, age, smoking status, disease stage, and therapy modality (chemotherapy and radiation therapy). Interaction terms were not considered in this analysis. After controlling for confounding factors, we found that carbohydrate metabolism had the biggest effect on OS (HR =5.21; 95% CI: 3.27 to 8.29; FDR- $P < 0.0001$) followed by lipids (HR =4.05; 95% CI: 2.55 to 6.69; FDR- $P < 0.0001$), cofactors (HR =4.00; 95% CI: 1.90 to 8.45; FDR- $P = 0.0003$), amino acids (HR =3.35; 95% CI: 2.13 to 5.26; FDR- $P < 0.0001$), and nucleotides (HR =2.58; 95% CI: 1.60 to 4.16; FDR- $P = 0.0001$; Table 2).

Compared with the two metabolomic cohorts, the genomic TCGA LUAD cohort is comprised of a

Metabolite	Metabolic enzyme of clinical significance		Superpathway	Pathway
	Direct metabolizing enzyme (1)	Indirect metabolizing enzyme (2)		
Glutamate*	GAD2 (4.1.1.15) GLS2 (3.5.1.2) LAP3 (3.4.11.1)			Alanine, aspartate and glutamate meta
Cysteine		GGT5 (2.3.2.2) GGT5 (3.4.19.13)		
5-oxoproline		GGT5 (2.3.2.2)		
Glycine	LAP3 (3.4.11.1) GSS (6.3.2.3)	GGT5 (3.4.19.13) GGT5 (2.3.2.2)		
Glutamate*	GGT5 (2.3.2.2)	MGST3 (2.5.1.18) GSTA1 (2.5.1.18) GSTM5 (2.5.1.18)		Glutathione metabolism
Glutathione (GSH)	MGST3 (2.5.1.18) GSTA1 (2.5.1.18) GSTM5 (2.5.1.18) GGT5 (3.4.19.13) GGT5 (2.3.2.2)		Amino acids	
Glutamine	GLS2 (3.5.1.2)	LAP3 (3.4.11.1)		Arginine biosynthesis
Proline	PYCR2 (1.5.1.2) P4HA2 (1.14.11.2) LAP3 (3.4.11.5)			
Putrescine	AOC1 (1.4.3.22)	ALDH1B1 (1.2.1.3)		Arginine and proline metabolism
Spermidine	SMS (2.5.1.22)			
Trans-4-hydroxyproline	P4HA2 (1.14.11.2) PYCR2 (1.5.1.2)			
Canosine synthesis	CARNST (6.3.2.11)			
UDP-GlcA	UGDH (1.1.1.22)			
N-acetyl-D-mannosamine	GNE (3.2.1.183)	UAPL1 (2.7.7.23)		Amino sugar metabolism
N-acetylneuraminate	CMAS (2.7.7.43)	CYB5RL (1.6.2.2)		
Citrate	CS (2.3.3.1)			
Malate		CS (2.3.3.1)		Citrate cycle
2-oxoglutarate to succinyl-CoA conversion	DLST (2.3.1.61)			
Fructose		AKR1B1 (1.1.1.21)		
Mannitol		AKR1B1 (1.1.1.21)		
Sorbitol	AKR1B1 (1.1.1.21)		Carbohydrates	Fructose and mannose metabolis
Mannose-6P	PMM1 (5.4.2.8)	GMPPA (2.7.7.13)		
3-phosphoglycerate	BPGM (5.4.2.11)			
Glucose		BPGM (5.4.2.4) PGM2 (5.4.2.2)		Glycolysis, gluconeogenesis
Acetyl-CoA	ACSS2 (6.2.1.1)	ALDH3A1 (1.2.1.5)		
L-lactate	LDHAL6A (1.1.1.27)			
Ribose	RBKS (2.7.1.15) PGM2 (5.4.2.7)			Pentose phosphate cycle
NAD+ to nicotinamide mononucleotide conversion	NMNAT3 (2.7.7.1)		Cofactors	Nicotinate and nicotinamide metabolism
Flavin adenine dinucleotide	ENPP3 (3.6.1.9)			Riboflavin metabolism
Glycerate	ALDH1B1 (1.1.1.3)			
Glycerol	AKR1B1 (1.1.1.21)			
Glycerol 3 phosphate		ALDH1B1 (1.1.1.3) AKR1B1 (1.1.1.21)	Lipids	Glycerolipid metabolism
Ethanolamine	DGKG (2.7.1.107) LPL (3.1.1.34)			
Ceramide formation	ENPP7 (3.1.4.12) GBA (3.2.1.45)			Sphingolipid metabolism
Adenine	ADA (3.5.4.4)			
Adenosine	ADA (3.5.4.4)			
Guanosine		ENPP3 (3.6.1.9)		
2'-deoxyguanosine	DGUOK (2.7.1.113)			
Hypoxanthine		ADA (3.5.4.4)		
Inosine	ADA (3.5.4.4)			Purine metabolism
Inosine 5'-monophosphate		ADA (3.5.4.4)		
Cyclic nucleotides concentration regulation	ADCY2 (4.6.1.1) GUCY2C (4.6.1.2) PDE10 (3.1.4.35)			
CAMP concentration regulation	PDE10 (3.1.4.17) ADCY2 (4.6.1.1)			
Thymine		TK1 (2.7.1.21)		Pyrimidine metabolism
Uridine monophosphate (5' or 3')	ENPP3 (3.6.1.9)			

Figure 3 Association between altered metabolites identified by mass-spec and metabolic genes of clinical significance in non-squamous non-small cell lung cancer. Metabolites in dark blue are those found to be significantly altered by both gas chromatography-time-of flight mass spectrometry (GC-TOF-MS) and liquid chromatography tandem mass spectrometry coupled with gas chromatography mass spectrometry (UPLC/MS/MS coupled with GC-MS); metabolites in light blue are those to be found significantly altered only by UPLC/MS/MS coupled with GC-MS, metabolic functions colored in white were not identified by either technique. Orange indicates non-synonymous alterations in metabolic genes associated with shorter OS, and green indicates non-synonymous alterations in metabolic genes associated with longer OS. Asterisk (*) indicates metabolite, being present in several metabolic pathways. 1, enzymes, directly catabolize or anabolize metabolite; 2, enzymes, catabolize or anabolize of an immediate metabolic precursor or byproduct. UDP-GlcA, uridine 5'-diphosphoglucuronic acid; GC-MS, gas chromatography mass spectrometry; OS, overall survival.

Table 2 The relative association of individual metabolic alterations with overall survival of patients from lung adenocarcinoma cohort

Metabolic superpathway	HR	95% CI	P value	FDR-P value
Amino acids	3.346	2.129–5.258	<0.0001	<0.0001
Carbohydrates	5.208	3.272–8.291	<0.0001	<0.0001
Cofactors	4.003	1.896–8.449	0.0003	0.0003
Lipids	4.054	2.547–6.690	<0.0001	<0.0001
Nucleotides	2.578	1.598–4.159	0.0001	0.0001

Forty genes that were associated with lower OS were grouped into five metabolic superpathways. The HR +/- 95% CI, P value and FDR corrected P value (FDR-P value) for each clinically relevant metabolic superpathway in lung adenocarcinoma cohort from the TCGA database. HR, hazard ratio; CI, confidence interval; FDR, false discovery rate; OS, overall survival; TCGA, The Cancer Genome Atlas.

heterogeneous NSCLC patient population (Table S1). Most patients (77%) had early-stage disease (I–II). Fifty-four percent received chemotherapy only and 31% patients received radiation therapy only. For both treatment groups, patients with early-stage disease had more favorable clinical outcomes than those with later stage disease. For patients treated with chemotherapy only, the median survival time for early stage disease was 103.9 months, and for later stage disease was 41.5 months. For patients treated with radiation therapy only, the median survival time for early stage disease was 76.0 months, and for later stage disease was 15.0 months.

To examine whether metabolic superpathway dysfunction may have differentially impacted clinical outcomes by disease stage, we added the pathway \times stage interaction effect to the multivariable Cox regression model. We conducted three separate analyses by patients groups: all patients, patients who received chemotherapy only, and patients who received radiation therapy only). The pathway \times stage interaction effect was not significant for any of the superpathways in these analyses; P values ranged from 0.23 to 0.99. Similarly, pathway \times treatment (chemotherapy or radiation therapy) interaction effect was also not significant, with P values ranging from 0.61 to 0.92. Confidence intervals for hazard ratios overlapped between early stages (I–II) and late stages (III–IV), and between different treatment types (Figure S5). Taken together, dysfunction of individual metabolic superpathways (amino acids, carbohydrates, nucleotides, lipids or cofactors) was associated with OS independently of disease stage and treatment modality.

Discussion

In this study, we sought to assess whether metabolic alterations in non-squamous NSCLC are a byproduct of accelerating cell growth or whether they may be a driving

factor of tumor biology. Building on consensus findings from MS-based metabolomics studies, we developed a post-genomics research strategy to assess the clinical impact of metabolic alterations in non-squamous NSCLC. We specifically focused on five metabolic superpathways (amino acids, carbohydrates, cofactors, lipids, and nucleotides), reproducibly found altered in LUAD across metabolomic studies (10,11). Using KEGG PATHWAY and KEGG ENZYME databases to identify metabolic genes controlling those metabolic superpathways, we queried the TCGA LUAD genomic cohort for function-altering mutations affecting those genes (KEGG-TCGA analysis), as a surrogate marker for metabolic alterations. While this approach allowed us to probe the association between clinical outcomes and specific metabolic functions in large and clinically annotated cohorts, it is not without limitations. It only provides an indirect assessment of metabolic functions in lung cancer and is restricted to metabolic genes presenting with function-altering mutations. As a result, the overall frequency of metabolic gene mutations may introduce a selection bias and diminished statistical power. Nonetheless, the metabolic genes we identified as clinically relevant did not present with a higher frequency of mutations or mutation severity score (SIFT and PolyPhen) in studied cohort.

The TCGA LUAD genomic cohort data has several important features that we were able to leverage in our study. The data represents a large sample size, extensive clinical annotation with almost 2 years of clinical follow-up and includes comprehensive genomic data. While it provided us with a valuable hypothesis-generating dataset, we also acknowledge that it bears some limitations as it lacks extensive proteomic and metabolomic data and a matching validation cohort of comparable size, annotations, and length of clinical follow-up. In the context of metabolic

studies, these limitations are compounded by the fact that a complex network of endogenous and exogenous factors regulates cellular metabolomic activity (17). We demonstrated that 41 of 42 metabolic genes of clinical significance were indeed expressed in non-squamous NSCLC tumors. However, our genomic data do not account for post-translational modification and metabolic activity regulation, for example, protein folding or stability, protein phosphorylation/acetylation/ubiquitination, substrate and cofactor availability, enzymatic activity overlap between metabolic enzyme isoforms or interplay between anabolic and catabolic enzymes. To overcome such limitations, we employed a biological validation approach in which we paired genomic data from the TCGA LUAD cohort with metabolomic data derived from two independent non-squamous NSCLC cohorts (8,9). We showed that 81% of clinically relevant genes (34 of 42) controlled anabolism or catabolism of altered metabolites in non-squamous NSCLC (converging of KEGG-TCGA and MS finding). In the context of our analytical input, where only 176 out of 736 (23.9%) metabolic genes tested directly controlled non-squamous NSCLC metabolic alterations, such high level of overlap (81%) between genomic and metabolomic findings suggests a non-stochastic distribution of metabolic genes of clinical significance. Eleven non-squamous NSCLC metabolic alterations were unrelated to metabolic functions of clinical significance (MS findings only) and could be a byproduct of accelerating cell growth.

We found that clusters of genes controlling key metabolic functions of the carbohydrates metabolic superpathway had the most profound impact on OS. For instance, alterations in genes controlling glucose and fructose catabolism, two carbohydrates and key components of the Warburg effect, were associated with shorter OS. Metabolomic studies and clinical PET-CT examinations have previously shown that lung tumors exhibit hyperactive glucose metabolism (18,19). Lung tumors are dependent on glucose metabolism, and increased expression of glycolytic enzymes correlates with poor prognosis (20). In an unselected cohort of 342 patients with newly diagnosed NSCLC, patients with high levels of glucose in blood had shorter OS than those with normal glucose levels, independently of other prognostic factors (21). However, while hyperglycemia decreases the antiproliferative effect of chemotherapy in preclinical models, glycemic control interventions have shown inconsistent results in clinical trials (20). Beyond the Warburg effect, we found that alterations in mannose metabolism, a key

carbohydrate in glycosylation (22), were also associated with shorter OS. Consistent with our results, high mannose-type glycans are prevalent in LUAD tissues (23) and are associated with disease progression (24).

Alterations to three additional gene clusters that control glutathione catabolism, proline synthesis (amino acids superpathway), and uric acid anabolism (nucleotides superpathway) were also associated with worse OS. The frequency of metabolic genes mutations controlling both amino acid and nucleotide superpathways was higher among smokers than non-smokers. Smoking is a known risk factor in NSCLC and is associated with poor clinical outcomes (25). In our model that was adjusted for variables such as smoking status, alterations in amino acid and nucleotide superpathways were independently associated with clinical outcome. In a hydrophilic environment, both glutathione and uric acid are antioxidants, acting as scavengers of free radicals produced by reactive oxygen species. Higher glutathione levels have been associated with resistance to both chemotherapy and radiation therapy (26,27). However, conflicting data in the literature supports the role of uric acid in cancer. Results of a study involving 354,110 participants showed that serum uric acid was cancer-protective and reduced risk cancer mortality (28), while data from another study involving 83,693 patients showed that high levels of serum uric acid were associated with a higher risk of all cancer mortality (29). In our study, alterations to genes controlling uric acid precursors anabolism was associated with shorter OS, suggesting that uric acid may indeed be harmful in non-squamous NSCLC. Similarly, our data showed that alterations to the biosynthesis of proline (i.e., *PYCR2*, *P4HA2*, and *LAP3*), a non-essential amino acid involved in cancer cell redox homeostasis (30,31) was associated with shorter OS. Overexpression of *PYCR2* and *P4HA2* has been described in many malignancies, including lung cancer (32,33). Conflicting with our findings, preclinical data attributed a pro-tumorigenic role to both genes in epithelial cancers (34,35).

While most of the metabolic genes that we identified as clinically relevant were associated with shorter OS, alterations to two genes were associated with improved OS: *GAD2* and *PDE10*. Both metabolic genes are typically expressed in the brain and central nervous system but their ectopic expression in tumor types such as lung tumors has been described (36,37). Only *PDE10* ectopic expression in non-squamous NSCLC tumors was confirmed in our study, *GAD2* transcripts were not detected. *GAD* exists as two isoforms, *GAD1* and *GAD2*, that shares enormous

sequence homology at the protein level, but have different affinity for the cofactor pyridoxal 5'phosphate and distinct intracellular localization. Evidence suggest the presence of alternative splicing of GAD1 and a functional truncated version of GAD2 (38). From analyses of the LUAD TCGA data, we observed that GAD1 mRNA expression was a 5-fold increase in GAD1 mRNA expression within GAD2 mutated tumor compared with non-mutated (313.2 ± 47.4 vs. 63.6 ± 5 , $P < 0.0001$). GAD1/GAD2 catalyzes decarboxylation of glutamate to γ -aminobutyric acid (GABA), and PDE10A regulates various neurotransmitter receptors, including dopamine D1 and D2 (39,40). While emerging data suggest that cancer cells take advantage of the neurotransmitters-initiated signaling pathway to activate proliferation and tumor progression (41), both GABA and dopamine mostly activate cancer growth (42,43). In addition to regulating transmitter receptors, PDE10 also hydrolyzes cGMP, a strong antineoplastic agent across several cancer types (44). While our data cannot resolve those contradictory roles for GAD1-2/PDE10 in tumor biology, it appeared that function-altering mutations of those genes were associated with improved clinical outcomes for the cohort of patients that we studied. As such, synthetic inhibitors of those two metabolic enzymes, some of which are currently in development (45,46), may warrant further investigation for the treatment of non-squamous NSCLC.

This study presents a novel post-genomic research strategy for identifying metabolic alterations associated with clinical outcomes in non-squamous NSCLC. We identified a strong consensus between cancer metabolism data obtained by MS methods and clinically annotated genomics data. By providing a detailed map of the landscape of metabolic alterations in non-squamous NSCLC, our findings may offer valuable insights into the biology of the disease and enable the identification of new therapeutic targets to overcome resistance to chemotherapy and radiation therapy. Additional clinical and functional studies will be warranted to validate those finding and determine the actionability of those metabolic alterations.

Acknowledgments

Scientific writing and editing support were provided by Hazel F O'Connor, PhD, on behalf of Levine Cancer Institute. The results shown here are in part based upon data generated by the TCGA Research Network: <https://www.cancer.gov/tcga>.

Funding: This work was supported by the North Carolina

Lung Cancer Initiative Research Fellow Grant to AV Ivanina Foureau.

Footnote

Reporting Checklist: The authors have completed the STARD reporting checklist. Available at <https://tcr.amegroups.com/article/view/10.21037/tcr-22-377/rc>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://tcr.amegroups.com/article/view/10.21037/tcr-22-377/coif>). AVIF reports that this work was supported by North Carolina Lung Cancer Initiative Research Fellow grant. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Pavlova NN, Zhu J, Thompson CB. The hallmarks of cancer metabolism: Still emerging. *Cell Metab* 2022;34:355-77.
2. Zito Marino F, Ronchi A, Accardo M, et al. Concomitant ALK/KRAS and ALK/EGFR mutations in non small cell lung cancer: different profile of response to target therapies. *Transl Cancer Res* 2017;6:S457-S460.
3. Wang M, Herbst RS, Boshoff C. Toward personalized treatment approaches for non-small-cell lung cancer. *Nat Med* 2021;27:1345-56.
4. Jin N, Bi A, Lan X, et al. Identification of metabolic vulnerabilities of receptor tyrosine kinases-driven cancer. *Nat Commun* 2019;10:2701.

5. Majem B, Nadal E, Muñoz-Pinedo C. Exploiting metabolic vulnerabilities of Non small cell lung carcinoma. *Semin Cell Dev Biol* 2020;98:54-62.
6. Mendes C, Serpa J. Metabolic Remodelling: An Accomplice for New Therapeutic Strategies to Fight Lung Cancer. *Antioxidants (Basel)* 2019;8:603.
7. Luengo A, Gui DY, Vander Heiden MG. Targeting Metabolism for Cancer Therapy. *Cell Chem Biol* 2017;24:1161-80.
8. Li X, Song H, Kong F, et al. Pemetrexed exerts anticancer effects by inducing G(0)/G(1)-phase cell cycle arrest and activating the NOXA/Mcl-1 axis in human esophageal squamous cell carcinoma cells. *Oncol Lett* 2019;17:1851-8.
9. Wu HL, Gong Y, Ji P, et al. Targeting nucleotide metabolism: a promising approach to enhance cancer immunotherapy. *J Hematol Oncol* 2022;15:45.
10. Wikoff WR, Grapov D, Fahrman JF, et al. Metabolomic markers of altered nucleotide metabolism in early stage adenocarcinoma. *Cancer Prev Res (Phila)* 2015;8:410-8.
11. Moreno P, Jiménez-Jiménez C, Garrido-Rodríguez M, et al. Metabolomic profiling of human lung tumor tissues - nucleotide metabolism as a candidate for therapeutic interventions and biomarkers. *Mol Oncol* 2018;12:1778-96.
12. Rocha CM, Barros AS, Goodfellow BJ, et al. NMR metabolomics of human lung tumours reveals distinct metabolic signatures for adenocarcinoma and squamous cell carcinoma. *Carcinogenesis* 2015;36:68-75.
13. Kanehisa M. Enzyme Annotation and Metabolic Reconstruction Using KEGG. *Methods Mol Biol* 2017;1611:135-45.
14. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 2000;28:27-30.
15. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 2011;12:323.
16. Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. *Control Clin Trials* 1996;17:343-6.
17. Zapalska-Sozoniuk M, Chrobak L, Kowalczyk K, et al. Is it useful to use several "omics" for obtaining valuable results? *Mol Biol Rep* 2019;46:3597-606.
18. Feinberg T, Alkoby-Meshulam L, Herbig J, et al. Cancerous glucose metabolism in lung cancer-evidence from exhaled breath analysis. *J Breath Res* 2016;10:026012.
19. Tan H, Ma M, Huang J, et al. Preoperative Body Composition Combined with Tumor Metabolism Analysis by PET/CT Is Associated with Disease-Free Survival in Patients with NSCLC. *Contrast Media Mol Imaging* 2022;2022:7429319.
20. Gerards MC, van der Velden DL, Baars JW, et al. Impact of hyperglycemia on the efficacy of chemotherapy-A systematic review of preclinical studies. *Crit Rev Oncol Hematol* 2017;113:235-41.
21. Luo J, Chen YJ, Chang LJ. Fasting blood glucose level and prognosis in non-small cell lung cancer (NSCLC) patients. *Lung Cancer* 2012;76:242-7.
22. Nan F, Sun Y, Liang H, et al. Mannose: A Sweet Option in the Treatment of Cancer and Inflammation. *Front Pharmacol* 2022;13:877543.
23. Ruhaak LR, Taylor SL, Stroble C, et al. Differential N-Glycosylation Patterns in Lung Adenocarcinoma Tissue. *J Proteome Res* 2015;14:4538-49.
24. Lattová E, Skříčková J, Hausnerová J, et al. N-Glycan profiling of lung adenocarcinoma in patients at different stages of disease. *Mod Pathol* 2020;33:1146-56.
25. Bade BC, Dela Cruz CS. Lung Cancer 2020: Epidemiology, Etiology, and Prevention. *Clin Chest Med* 2020;41:1-24.
26. Estrela JM, Ortega A, Obrador E. Glutathione in cancer biology and therapy. *Crit Rev Clin Lab Sci* 2006;43:143-81.
27. Kennedy L, Sandhu JK, Harper ME, et al. Role of Glutathione in Cancer: From Mechanisms to Therapies. *Biomolecules* 2020.
28. Kuo CF, See LC, Yu KH, et al. Significance of serum uric acid levels on the risk of all-cause and cardiovascular mortality. *Rheumatology (Oxford)* 2013;52:127-34.
29. Strasak AM, Rapp K, Hilbe W, et al. Serum uric acid and risk of cancer mortality in a large prospective male cohort. *Cancer Causes Control* 2007;18:1021-9.
30. Hollinshead KER, Munford H, Eales KL, et al. Oncogenic IDH1 Mutations Promote Enhanced Proline Synthesis through PYCR1 to Support the Maintenance of Mitochondrial Redox Homeostasis. *Cell Rep* 2018;22:3107-14.
31. Schwörer S, Berisa M, Violante S, et al. Proline biosynthesis is a vent for TGFβ-induced mitochondrial redox stress. *EMBO J* 2020;39:e103334.
32. Nilsson R, Jain M, Madhusudhan N, et al. Metabolic enzyme expression highlights a key role for MTHFD2 and the mitochondrial folate pathway in cancer. *Nat Commun* 2014;5:3128.
33. Hu D, Ansari D, Zhou Q, et al. Low P4HA2 and high PRTN3 expression predicts poor survival in patients with pancreatic cancer. *Scand J Gastroenterol* 2019;54:246-51.
34. Ou R, Zhang X, Cai J, et al. Downregulation of pyrroline-

- 5-carboxylate reductase-2 induces the autophagy of melanoma cells via AMPK/mTOR pathway. *Tumour Biol* 2016;37:6485-91.
35. Ding Z, Ericksen RE, Escande-Beillard N, et al. Metabolic pathway analyses identify proline biosynthesis pathway as a promoter of liver tumorigenesis. *J Hepatol* 2020;72:725-35.
36. Dade M, Berzero G, Izquierdo C, et al. Neurological Syndromes Associated with Anti-GAD Antibodies. *Int J Mol Sci* 2020;21:3701.
37. Zhu B, Lindsey A, Li N, et al. Phosphodiesterase 10A is overexpressed in lung tumor cells and inhibitors selectively suppress growth by blocking β -catenin and MAPK signaling. *Oncotarget* 2017;8:69264-80.
38. Trifonov S, Yamashita Y, Kase M, et al. Glutamic acid decarboxylase 1 alternative splicing isoforms: characterization, expression and quantification in the mouse brain. *BMC Neurosci* 2014;15:114.
39. Al-Wadei HA, Al-Wadei MH, Schuller HM. Cooperative regulation of non-small cell lung carcinoma by nicotinic and beta-adrenergic receptors: a novel target for intervention. *PLoS One* 2012;7:e29915.
40. Niccolini F, Haider S, Reis Marques T, et al. Altered PDE10A expression detectable early before symptomatic onset in Huntington's disease. *Brain* 2015;138:3016-29.
41. Jiang SH, Hu LP, Wang X, et al. Neurotransmitters: emerging targets in cancer. *Oncogene* 2020;39:503-15.
42. Liang Y, Li H, Gan Y, et al. Shedding Light on the Role of Neurotransmitters in the Microenvironment of Pancreatic Cancer. *Front Cell Dev Biol* 2021;9:688953.
43. Qian X, Zhang D, Cao Z, et al. Dopamine Pathway Mediated by DRD5 Facilitates Tumor Growth via Enhancing Warburg Effect in Esophageal Cancer. *Front Oncol* 2021;11:655861.
44. Tuttle TR, Mierzwa ML, Wells SI, et al. The cyclic GMP/protein kinase G pathway as a therapeutic target in head and neck squamous cell carcinoma. *Cancer Lett* 2016;370:279-85.
45. Choi YK, Park KG. Targeting Glutamine Metabolism for Cancer Treatment. *Biomol Ther (Seoul)* 2018;26:19-28.
46. Amin HS, Parikh PK, Ghate MD. Medicinal chemistry strategies for the development of phosphodiesterase 10A (PDE10A) inhibitors - An update of recent progress. *Eur J Med Chem* 2021;214:113155.

Cite this article as: Ivanina Foureau AV, Sha W, Foureau DM, Symanowski JT, Farhangfar CJ, Mileham KF. Landscape and clinical impact of metabolic alterations in non-squamous non-small cell lung cancer. *Transl Lung Cancer Res* 2022;11(12):2464-2476. doi: 10.21037/tlcr-22-377

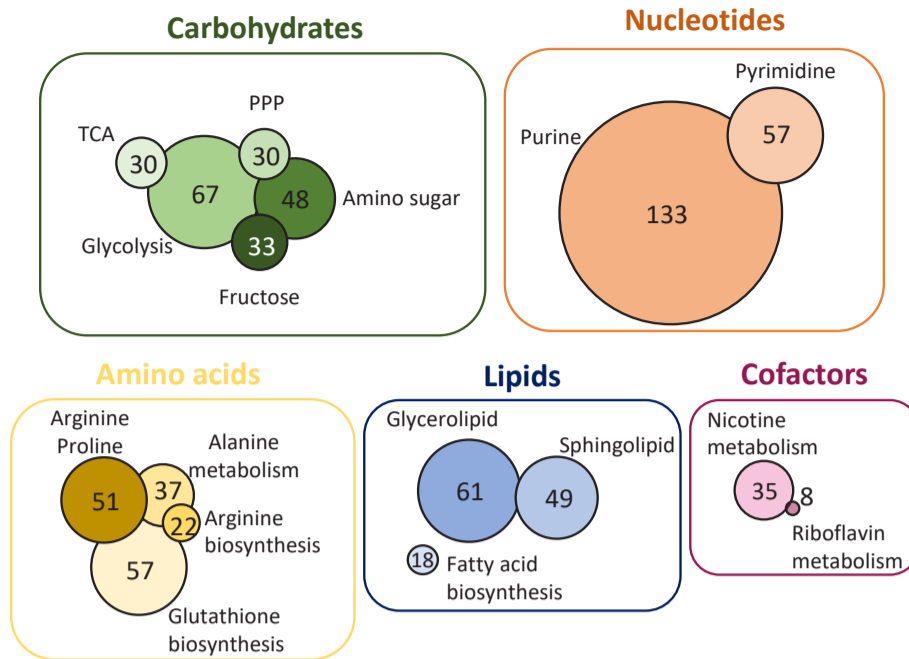


Figure S1 Diagram of 15 metabolic pathways, being altered in lung adenocarcinoma patients. Alanine metabolism- alanine, aspartate and glutamate metabolism; Arginine Proline, arginine and proline metabolism; Amino sugar, amino sugar metabolism; TCA, citric acid cycle; Fructose, fructose and mannose metabolism; Glycolysis, glycolysis and gluconeogenesis; PPP, pentose metabolism; Nicotine metabolism, nicotine and nicotinamide metabolism; Glycerolipid, glycerolipid metabolism; Sphingolipid, sphingolipid metabolism; Purine, purine metabolism; Pyrimidine, pyrimidine metabolism.

Table S1 Lung adenocarcinoma cohorts demographic and clinical information

Characteristic	Lung adenocarcinoma Wikoff (n=39)	Lung adenocarcinoma Moreno (n=33)	Lung adenocarcinoma TCGA (n=522)
Age, y			
Gender	72.3±8.78	62.1±9.73	65.8±0.45
Male	15 (38%)	24 (73%)	242 (46%)
Female	24 (62%)	9 (23%)	280 (54%)
Race			
White			393 (75%)
Black or African American			53 (10%)
Other			9 (2%)
NA	39 (100%)	33 (100%)	67 (13%)
Smoking status			
Smoking	39 (100%)		402 (77%)
Non-smoking			120 (23%)
NA		33 (100%)	
AJCC stage, n (%)			
I	39 (100%)	17 (52%)	279 (53%)
II		10 (30%)	124 (24%)
III		6 (18%)	85 (16%)
IV			26 (5%)
NA			8 (2%)
Treatment			
Radiotherapy only			160 (31%)
Chemotherapy only			284 (54%)
Chemo-radiation			78 (15%)
NA	39 (100%)	33 (100%)	
Length of clinical follow-up from diagnosis, mo			23.93

Gene	DF	Estimate	StdErr	ChiSq	HazardRatio	raw_p	pfdr_p
UAP1L1	1	4.9546	0.84573	34.3206	141.826	<.0001	<.0001
GBA	1	2.4793	0.52804	22.0461	11.933	<.0001	0.000735
GGT5	1	2.27617	0.59609	14.5808	9.739	0.0001	0.024706
LDHAL6A	1	2.43014	0.72568	11.2144	11.36	0.0008	0.094259
NMNAT3	1	3.45468	1.03597	11.1205	31.648	0.0009	0.094259
LAP3	1	1.81057	0.58831	9.4714	6.114	0.0021	0.146998
DLST	1	1.817	0.59158	9.4336	6.153	0.0021	0.146998
ACSS2	1	3.00126	1.0334	8.4347	20.111	0.0037	0.22578
CMAS	1	1.473	0.51464	8.1921	4.362	0.0042	0.232244
GMPPA	1	1.41237	0.51214	7.6054	4.106	0.0058	0.292024
MGST3	1	2.75469	1.02562	7.2139	15.716	0.0072	0.332766
RBKS	1	2.54643	1.01672	6.2728	12.762	0.0123	0.422517
P4HA2	1	1.28102	0.51174	6.2664	3.6	0.0123	0.422517
DGKG	1	1.14326	0.45821	6.2252	3.137	0.0126	0.422517
GNE	1	1.77386	0.71641	6.1308	5.894	0.0133	0.422517
ENPP3	1	0.90288	0.36658	6.0664	2.467	0.0138	0.422517
PMM1	1	1.70392	0.71761	5.638	5.495	0.0176	0.510612
AOC1	1	1.07126	0.46051	5.4114	2.919	0.02	0.552143
TK1	1	1.62524	0.71614	5.1503	5.08	0.0232	0.610941
GUCY2C	1	1.15127	0.51265	5.0434	3.162	0.0247	0.611207
ALDH1B1	1	1.13884	0.51352	4.9183	3.123	0.0266	0.611207
SMS	1	1.55249	0.71674	4.6917	4.723	0.0303	0.643458
PGM2	1	1.06793	0.51106	4.3666	2.909	0.0366	0.697599
GLS2	1	1.47655	0.71879	4.2198	4.378	0.04	0.732086
CS	1	2.06318	1.01018	4.1713	7.871	0.0411	0.732086
GSTA1	1	1.15091	0.58691	3.8453	3.161	0.0499	0.843039
PYCR2	1	1.14325	0.58629	3.8024	3.137	0.0512	0.843039
GSS	1	1.13911	0.58776	3.756	3.124	0.0526	0.843039
UGDH	1	0.98409	0.5105	3.7161	2.675	0.0539	0.843039
ADCY2	1	0.58736	0.30607	3.6827	1.799	0.055	0.843039
CARNS1	1	1.32751	0.71584	3.439	3.772	0.0637	0.917396
LPL	1	1.08953	0.58816	3.4315	2.973	0.064	0.917396
GSTM5	1	0.7746	0.4211	3.3836	2.17	0.0658	0.917396
CYB5RL	1	1.30852	0.7151	3.3483	3.701	0.0673	0.917396
ADA	1	1.85034	1.0144	3.3273	6.362	0.0681	0.917396
DGUOK	1	1.29131	0.71565	3.2558	3.638	0.0712	0.919178
ALDH3A1	1	1.05951	0.58808	3.2459	2.885	0.0716	0.919178
ENPP7	1	0.90736	0.51039	3.1604	2.478	0.0754	0.946461
PDE10A	1	-1.00389	0.58589	2.9359	0.366	0.0866	0.99805
AKR1B1	1	1.22208	0.71865	2.8918	3.394	0.089	0.99805
GAD2	1	-1.20873	0.71314	2.8728	0.299	0.0901	0.99805
BPGM	1	0.99282	0.58619	2.8686	2.699	0.0903	0.99805

Figure S2 Metabolic genes of clinical significance in non-squamous non-small cell lung cancer. Orange indicates hazard ratio of clinically relevant genes associated with shorter OS; green indicates Hazard Ratio of clinically relevant genes associated with longer OS. DF, degree of freedom; Std. Err, standard error; Chi Sq, Chi squared test; HR, hazard ratio; raw_p, raw P value; pfdr, p-FDR corrected P value; OS, overall survival.

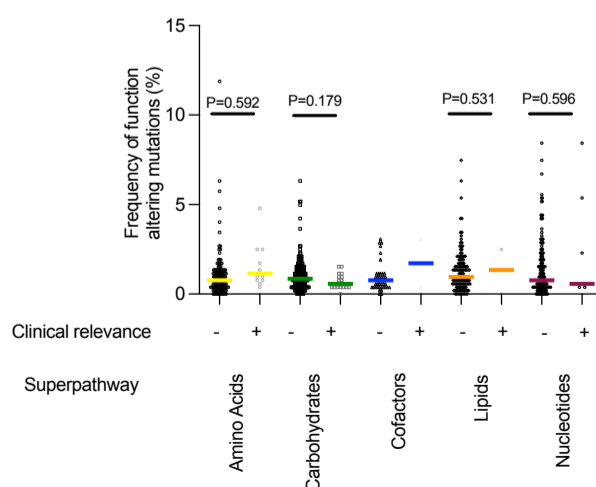


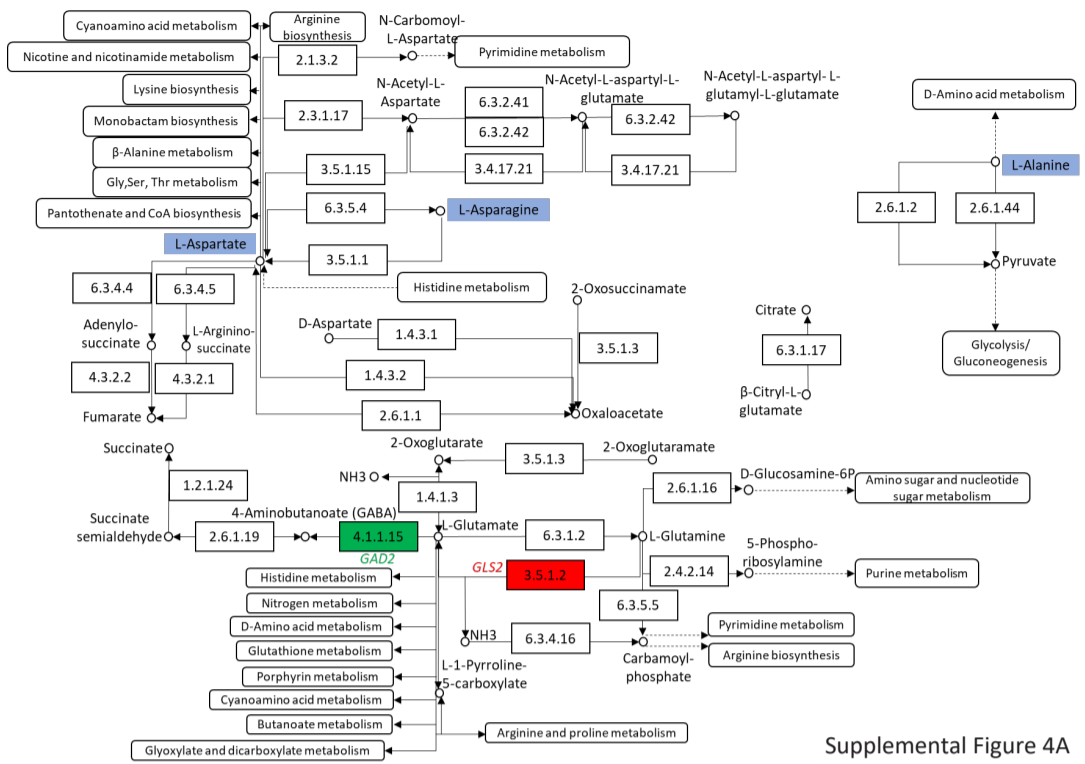
Figure S3 Incidence of function altering mutations of non-clinically relevant and clinically relevant genes controlling cellular metabolism in non-small cell lung cancer. Data presented as frequency of function altering mutations (%) per superpathway. Each point represents frequency of mutation in individual gene. Frequency of mutations in 522 patients from lung adenocarcinoma cohort.

Table S2 RNA expression of metabolic genes of clinical significance in non-squamous non-small cell lung cancer

Gene	Non-mutated	Mutated	P value
<i>UAP1L1</i>	354.1±12.1	168.2±194.7	0.3412
<i>GBA</i>	1849.5±34.0	2440.3±289.7	0.0433*
<i>GGT5</i>	886.3±31.4	640.3±355.4	0.4910
<i>LDHAL6A</i>	2.8±0.7	0.9±11.6	0.8695
<i>NMNAT3</i>	65.3±2.0	42.0±31.2	0.4564
<i>LAP3</i>	2663.4±53.2	2655.3±537.4	0.9880
<i>DLST</i>	2200.4±32.8	1834.1±331.8	0.2724
<i>ACSS2</i>	1589.4±42.6	902.6±683.2	0.3161
<i>CMAS</i>	850.1±33.9	1021.2±312.9	0.5870
<i>GMPPA</i>	1000.7±17.1	958.7±136.2	0.7599
<i>MGST3</i>	1652.1±34.9	1928.8±559.3	0.6218
<i>RBKS</i>	125.4±3.4	70.5±55.0	0.3201
<i>P4HA2</i>	1699.8±43.7	1556.9±403.1	0.7246
<i>DGKG</i>	97.1±4.1	150.3±25.5	0.0401*
<i>GNE</i>	616.9±16.9	668.4±271.5	0.8498
<i>ENPP3</i>	206.9±15.1	114.4±84.6	0.2827
<i>PMM1</i>	956.7±33.3	631.4±376.8	0.3902
<i>AOC1</i>	900.6±71.0	792.9±441.8	0.8098
<i>TK1</i>	1243.7±44.0	1588.7±705.7	0.6258
<i>GUCY2C</i>	38.3±10.1	25.9±65.4	0.8511
<i>ALDH1B1</i>	880.5±24.2	928.3±181.4	0.7939
<i>SMS</i>	1601.3±37.1	1474.1±484.8	0.7938
<i>PGM2</i>	566.4±14.7	669.1±117.0	0.3840
<i>GLS2</i>	160.3±7.8	194.8±78.3	0.6611
<i>CS</i>	3828.5±56.5	3737.1±906.1	0.9199
<i>GSTA1</i>	788.5±145.8	17.2±1344.1	0.5686
<i>PYCR2</i>	1847.3±29.2	2326.4±249.3	0.0569
<i>GSS</i>	1215.5±20.2	1653.6±204.7	0.0336*
<i>UGDH</i>	3356.8±210.1	4794.2±2124.3	0.5010
<i>ADCY2</i>	27.2±4.3	21.6±14.1	0.7011
<i>CARNS1</i>	77.3±6.1	26.5±69.3	0.4649
<i>LPL</i>	1500.2±98.5	389.5±839.7	0.1896
<i>GSTM5</i>	80.5±3.8	57.4±23.8	0.3386
<i>CYB5RL</i>	70.6±1.7	75.2±28.0	0.8686
<i>ADA</i>	178.5±6.2	82.3±99.0	0.3325
<i>DGUOK</i>	845.8±15.5	789.8±202.2	0.7827
<i>ALDH3A1</i>	1762.6±320.7	475.4±2735.0	0.6404
<i>ENPP7</i>	0.7±0.1	1.0±0.40	0.3860
<i>PDE10A</i>	160.3±13.7	158.3±57.3	0.9723
<i>AKR1B1</i>	2062.0±111.9	2530.8±1131.5	0.6803
<i>GAD2</i>	0.2±0.1	0.00E+00	0.7037
<i>BPGM</i>	649.3±18.2	469.9±205.4	0.3847

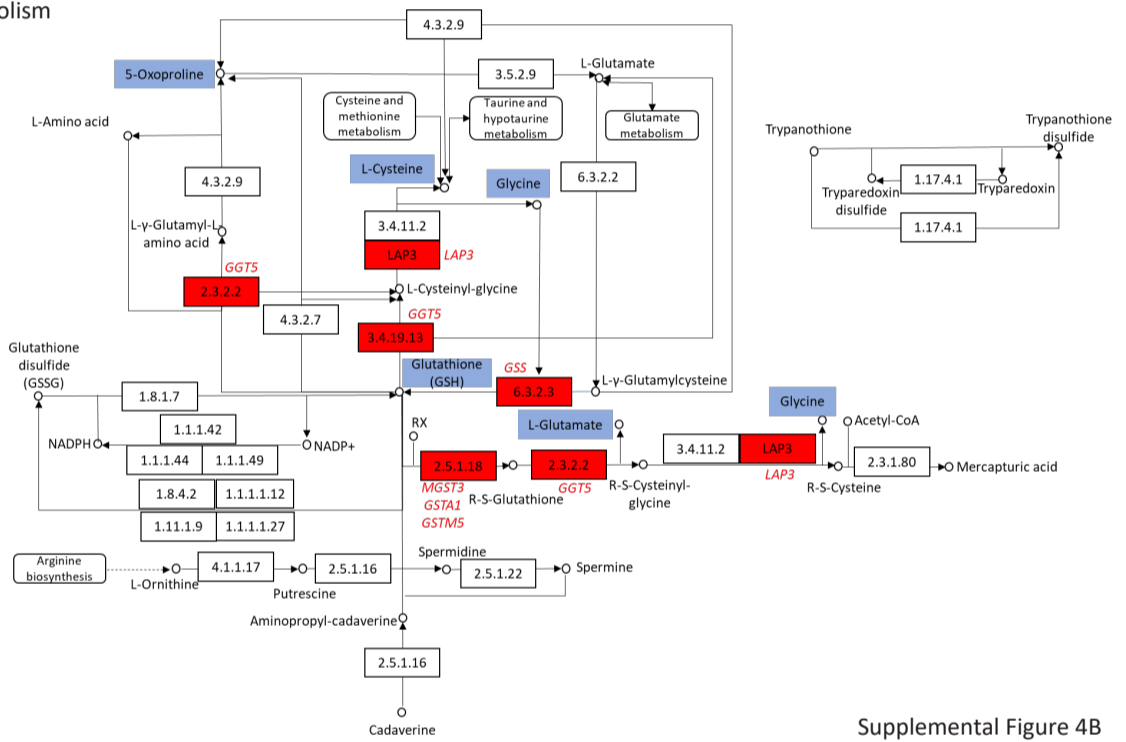
Data presented as mean ± standard error. Putative differences between groups were tested using pair-wise comparison. Significant P values, i.e., P<0.05 between non-mutated vs. mutated metabolic genes of clinical significance are indicated by *.

Alanine, aspartate and glutamate metabolism



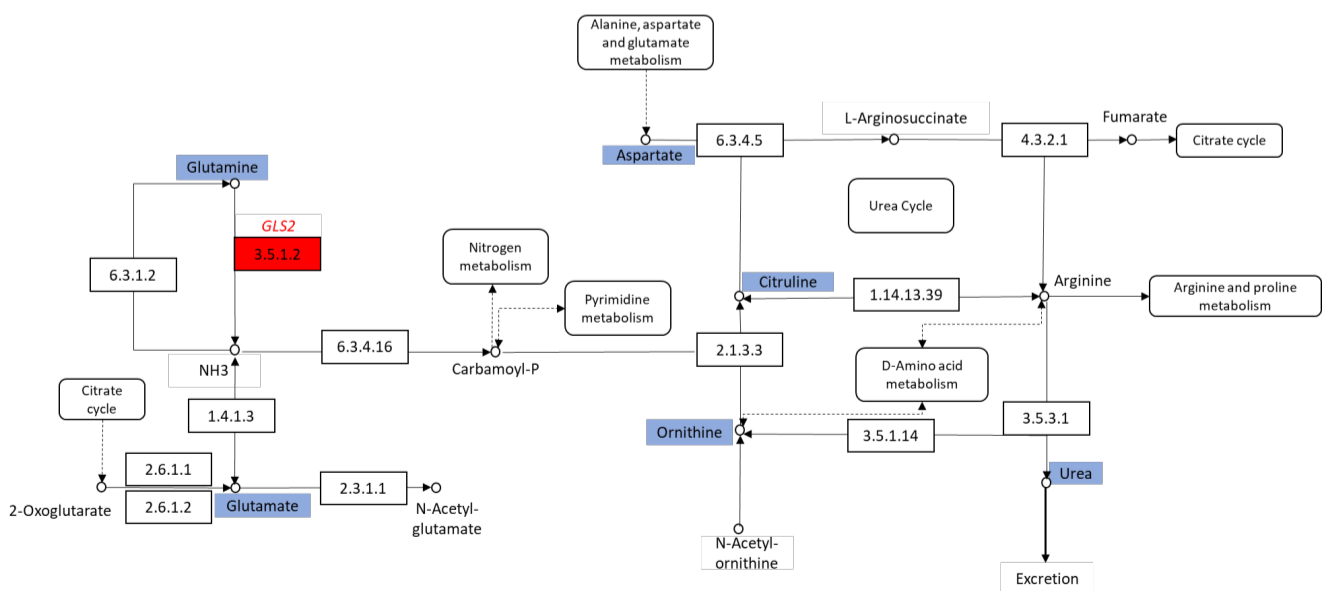
Supplemental Figure 4A

Glutathione metabolism



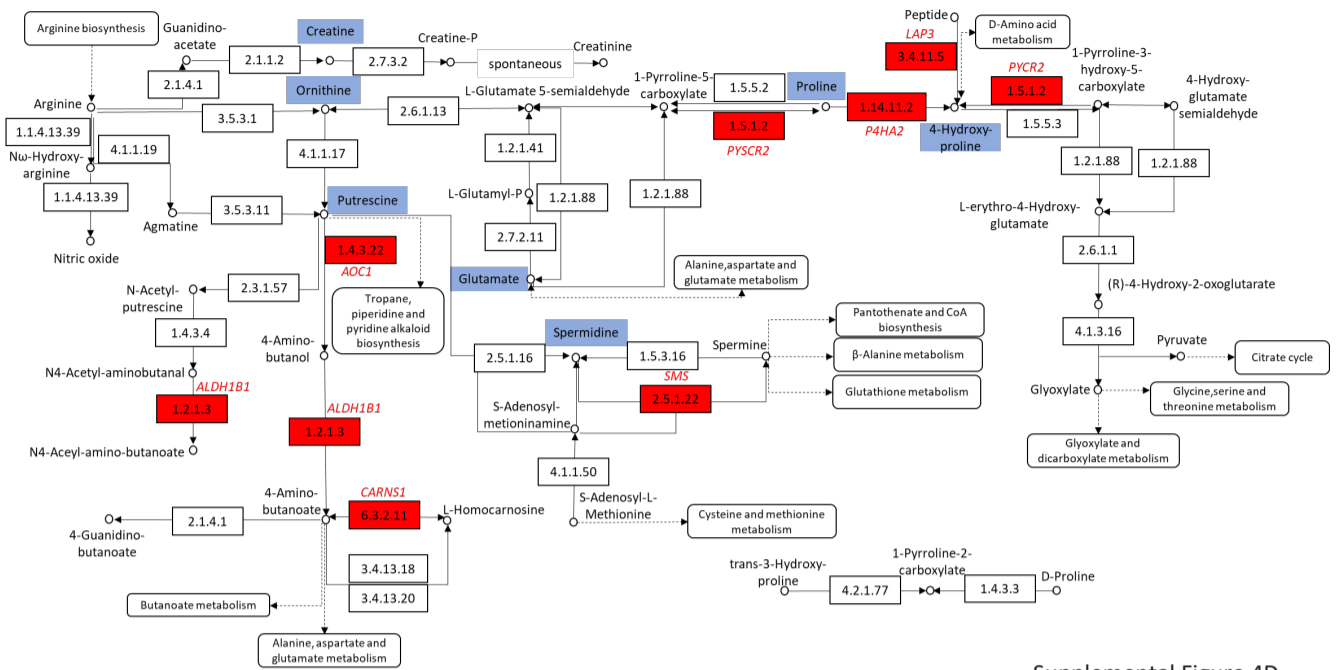
Supplemental Figure 4B

Arginine metabolism



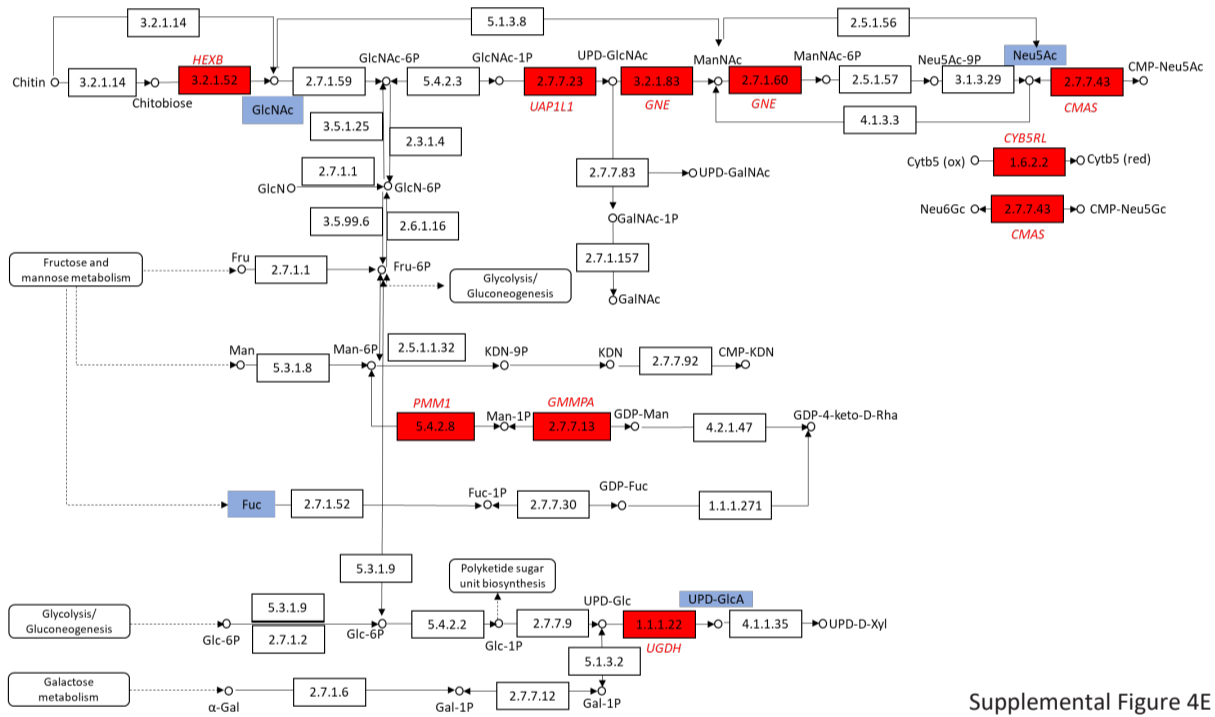
Supplemental Figure 4C

Arginine and proline metabolism



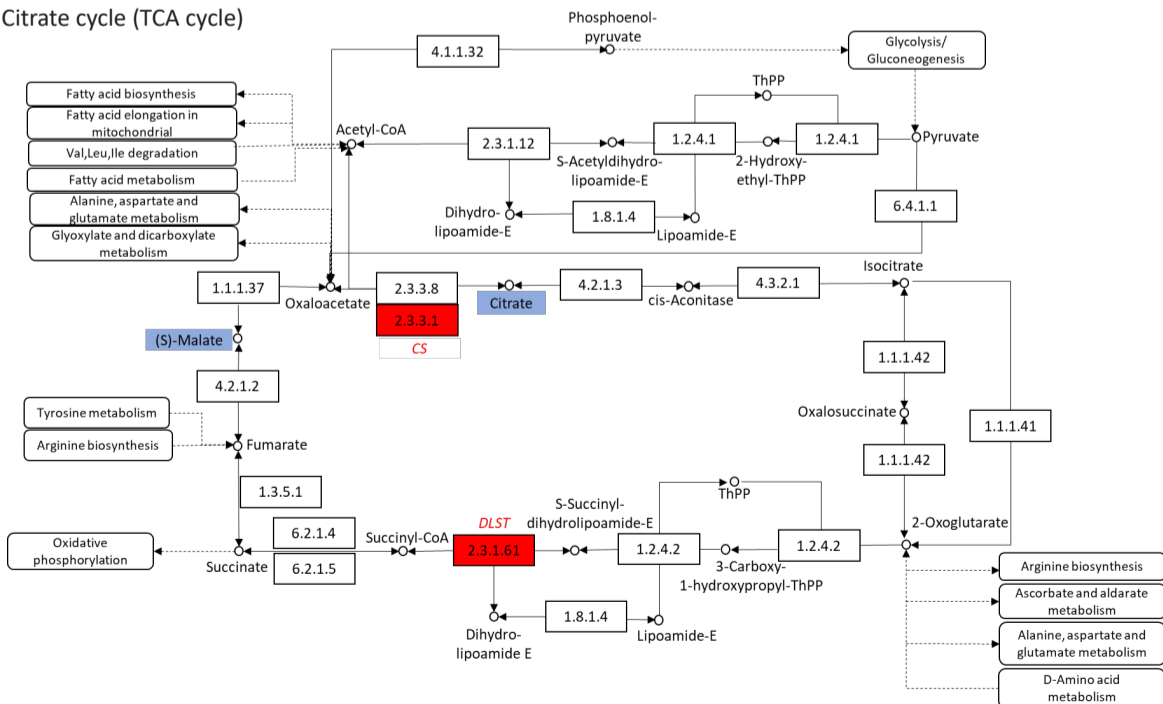
Supplemental Figure 4D

Amino sugar and nucleotide sugar metabolism



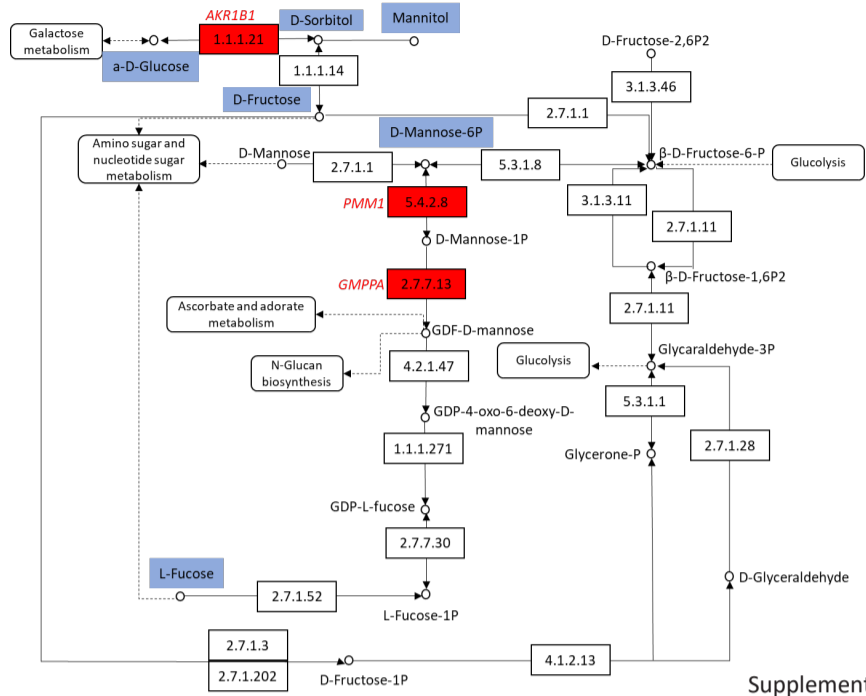
Supplemental Figure 4E

Citrate cycle (TCA cycle)



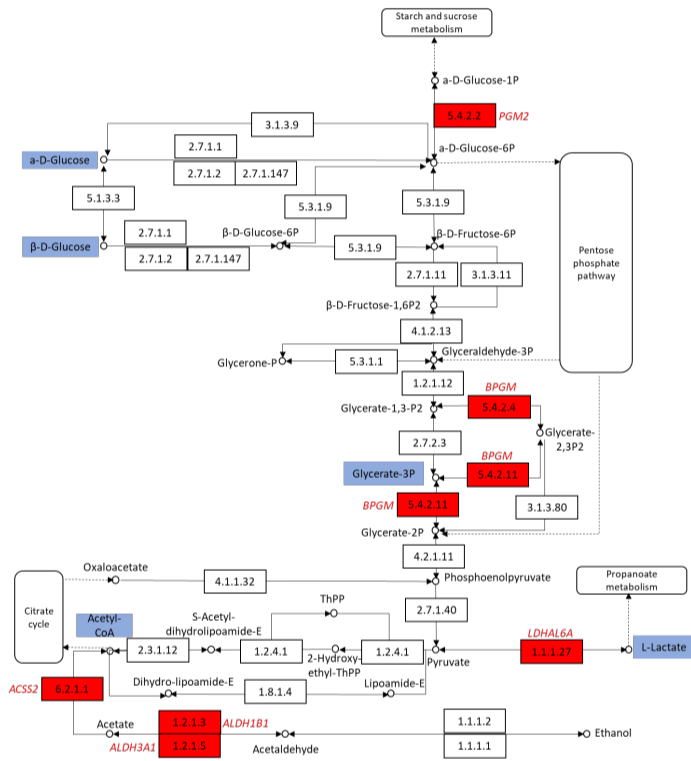
Supplemental Figure 4F

Fructose and mannose metabolism



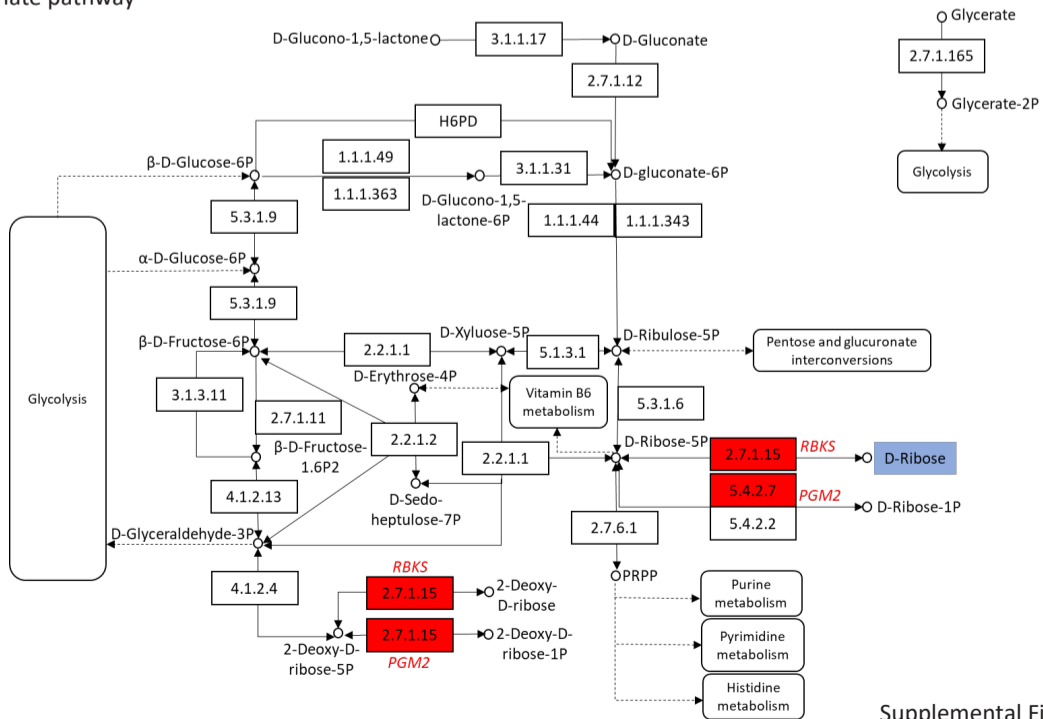
Supplemental Figure 4G

Glycolysis/gluconeogenesis



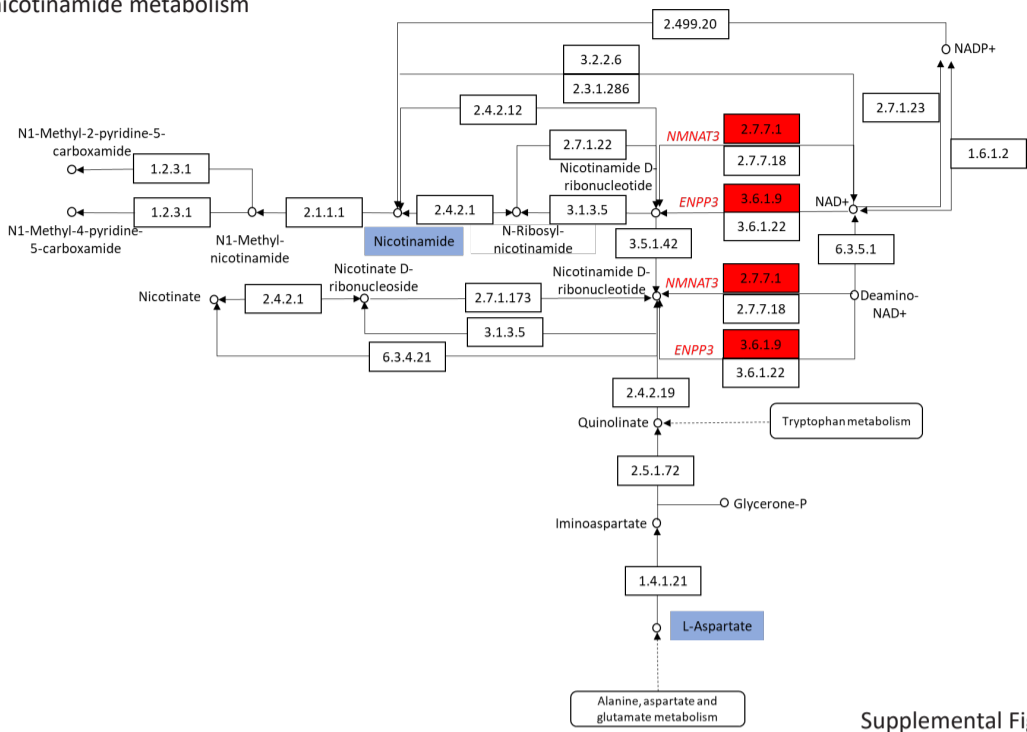
Supplemental Figure 4H

Pentose phosphate pathway



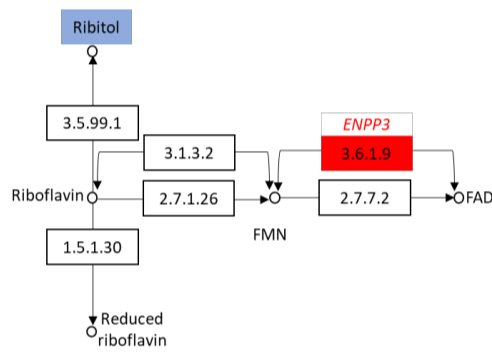
Supplemental Figure 4I

Nicotinate and nicotinamide metabolism



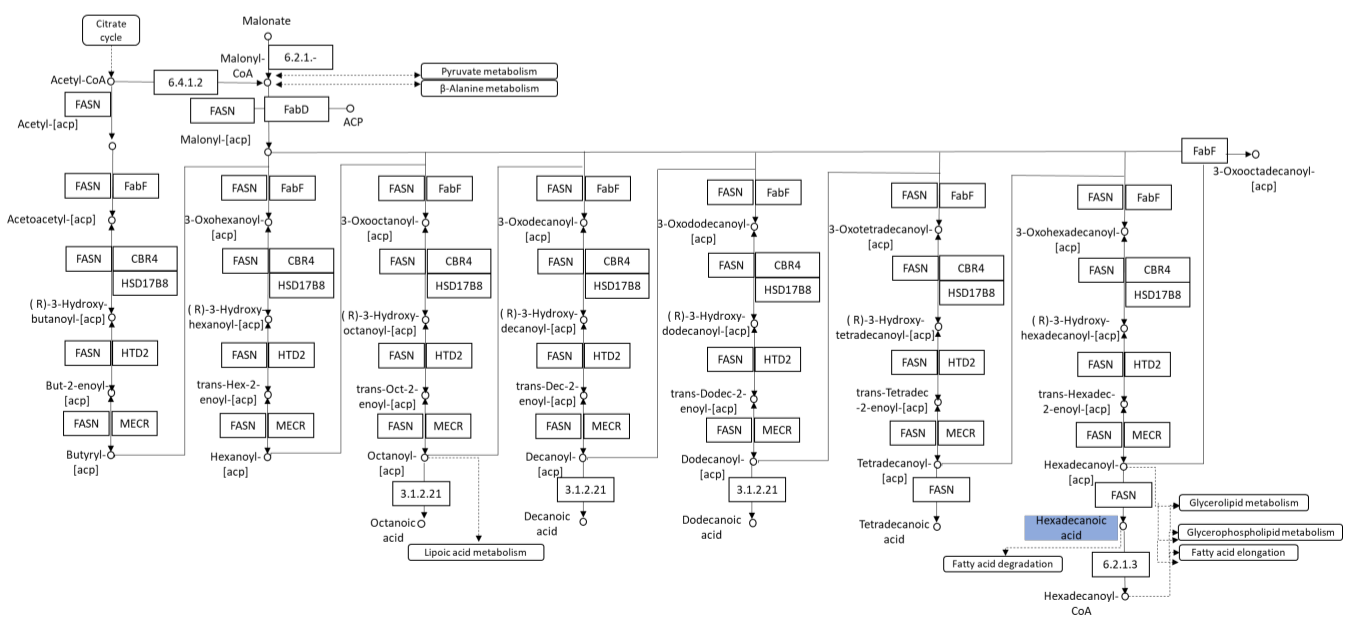
Supplemental Figure 4J

Riboflavin metabolism



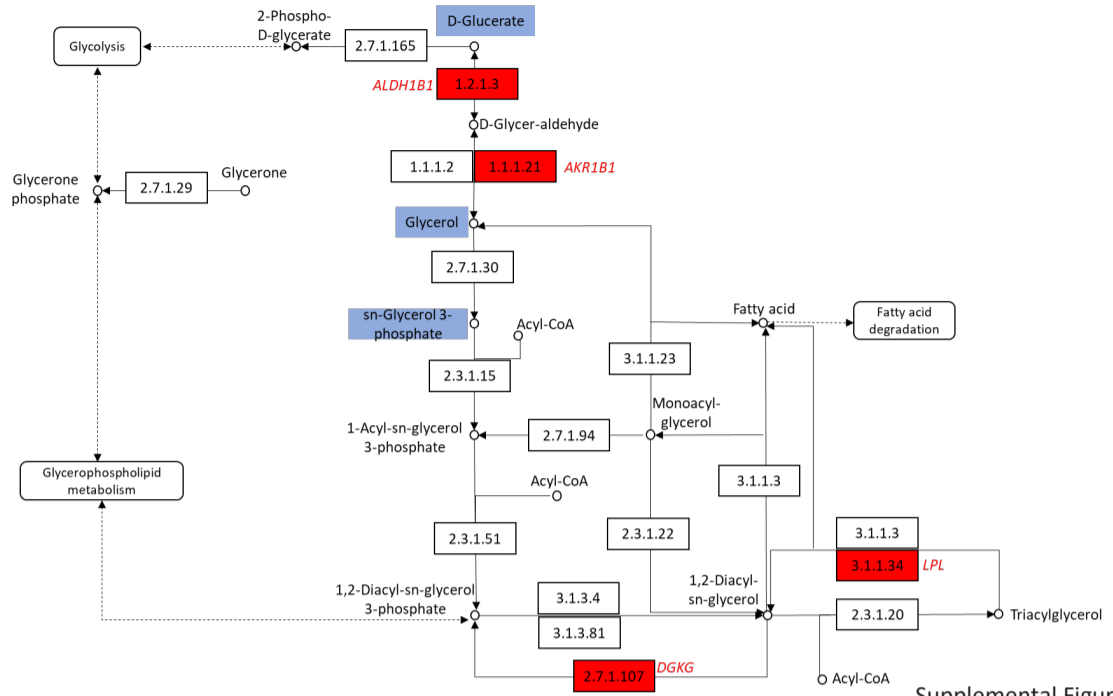
Supplemental Figure 4K

Fatty acid biosynthesis



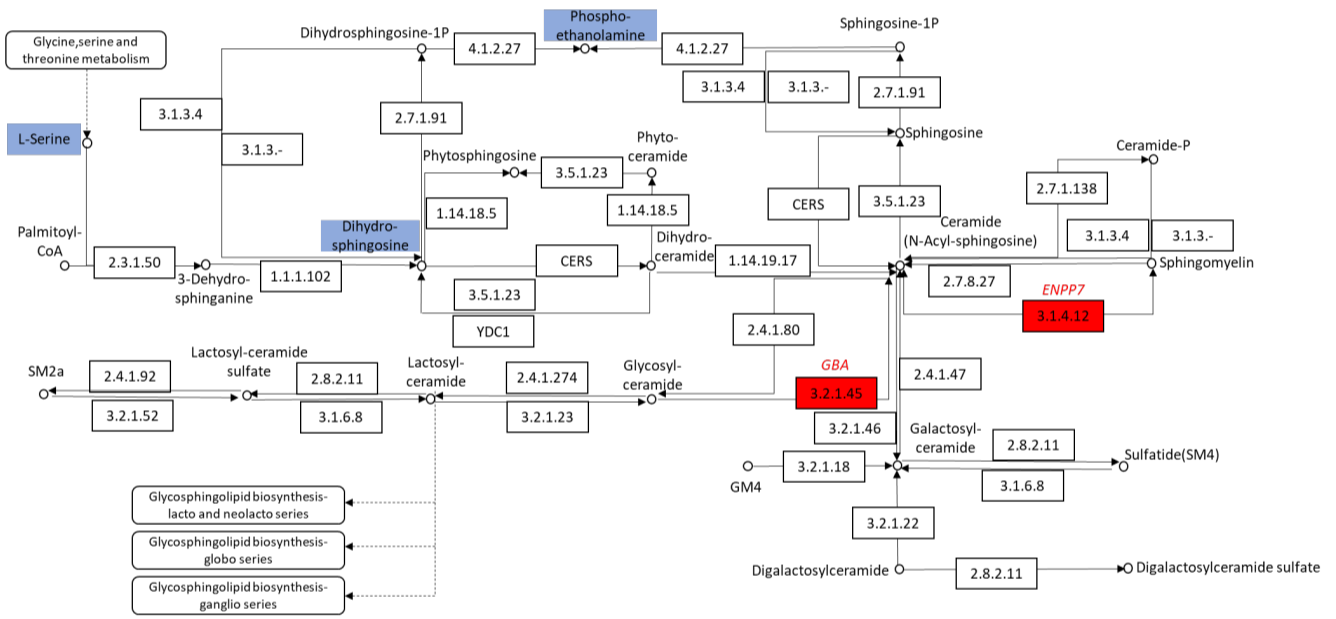
Supplemental Figure 4L

Glycerolipid metabolism



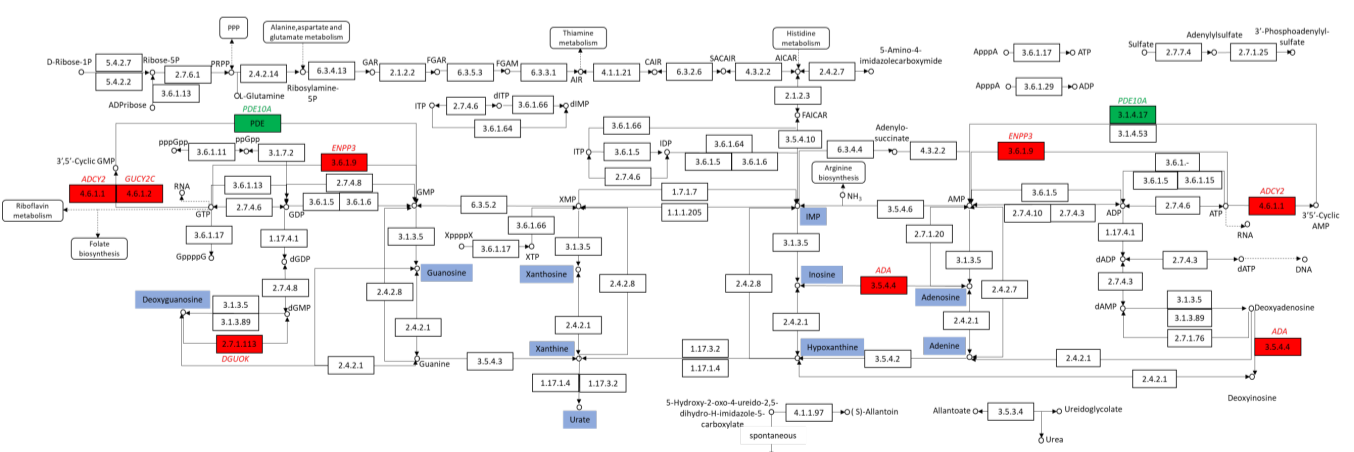
Supplemental Figure 4M

Sphingolipid metabolism



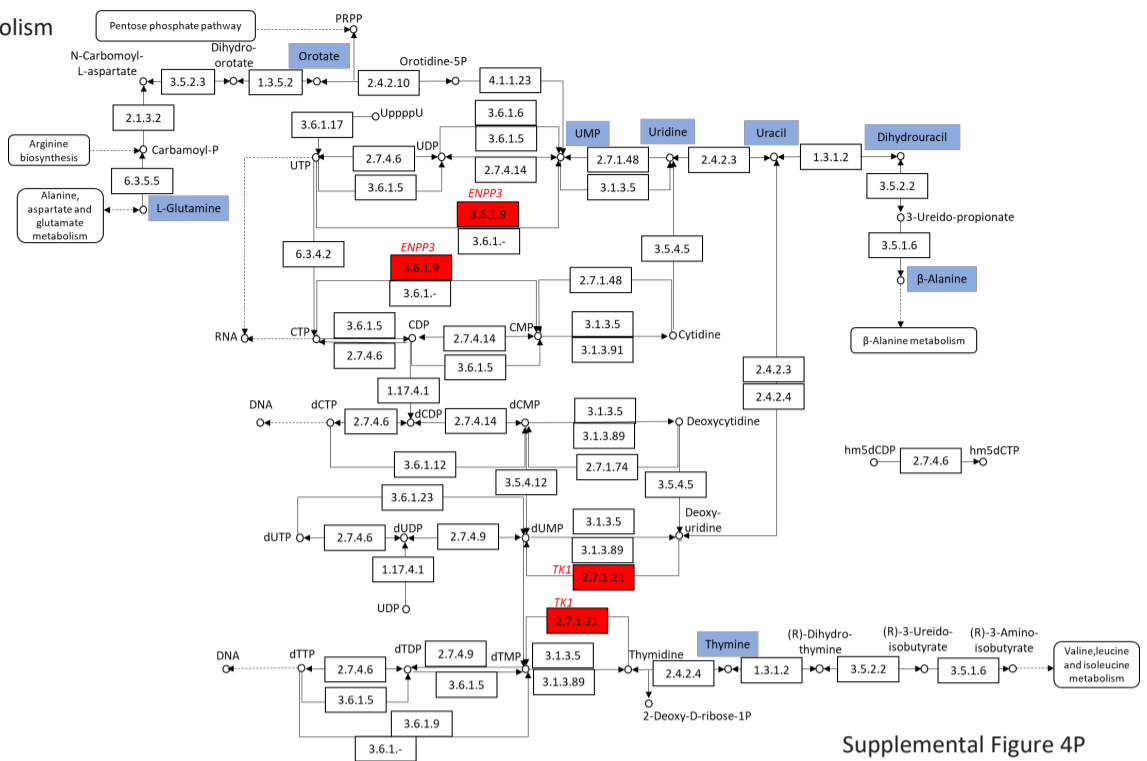
Supplemental Figure 4N

Purine metabolism



Supplemental Figure 4O

Pyrimidine metabolism



Supplemental Figure 4P

Figure S4 Metabolic maps with altered metabolites and clinically relevant genes in lung adenocarcinoma. (A) Alanine, aspartate and glutamate metabolism; (B) Glutathione metabolism; (C) Arginine biosynthesis; (D) Arginine and proline metabolism; (E) Amino sugar and nucleotide sugar metabolism; (F) Citrate cycle; (G) Fructose and mannose metabolism; (H) Glycolysis/gluconeogenesis; (I) Pentose phosphate pathway; (J) Nicotinate and nicotinamide metabolism; (K) Riboflavin metabolism; (L) Fatty acid biosynthesis; (M) Glycerolipid metabolism; (N) Sphingolipid metabolism; (O) Purine metabolism; (P) Pyrimidine metabolism. Metabolites, being altered in lung adenocarcinoma are colored in blue; genes, function altering mutations in which associated with worst prognosis colored in red; and genes, function altering mutations in which associated with good prognosis colored in green.

Table S3 Effect of clinical and demographic factors on frequency of mutations in metabolic superpathway of non-squamous non-small cell lung cancer

Factors	Metabolic superpathways				
	Amino acids	Carbohydrates	Cofactors	Lipids	Nucleotides
Gender (M vs. F, P value)	16.9% vs. 17.5%, P=0.866	12.0% vs. 10.4%, P=0.556	3.7% vs. 3.2%, P=0.753	9.5% vs. 7.5%, P=0.412	21.5% vs. 15.7%, P=0.090
Smoking status (non-smoking vs. smoking, P value)	10.8% vs. 19.2% P=0.027*	8.3% vs. 11.9%, P=0.273	2.5% vs. 3.7%, P=0.502	7.5% vs. 8.7%, P=0.673	10.0% vs. 20.9%, P=0.004*
AJCC stage (stages I + II vs. stages III + IV, P value)	16.6% vs. 19.8% P=0.437	10.2% vs. 14.4%, P=0.221	3.0% vs. 3.6%, P=0.741	7.9% vs. 10.9%, P=0.351	19.4% vs. 11.7%, P=0.052

Frequency of mutations is expressed in percentile value (%) between demographic and clinical groups. Putative differences between groups were tested using Fisher's exact test. Significant P values, i.e., P<0.05 between non-smoker vs. smokers are indicated by *.

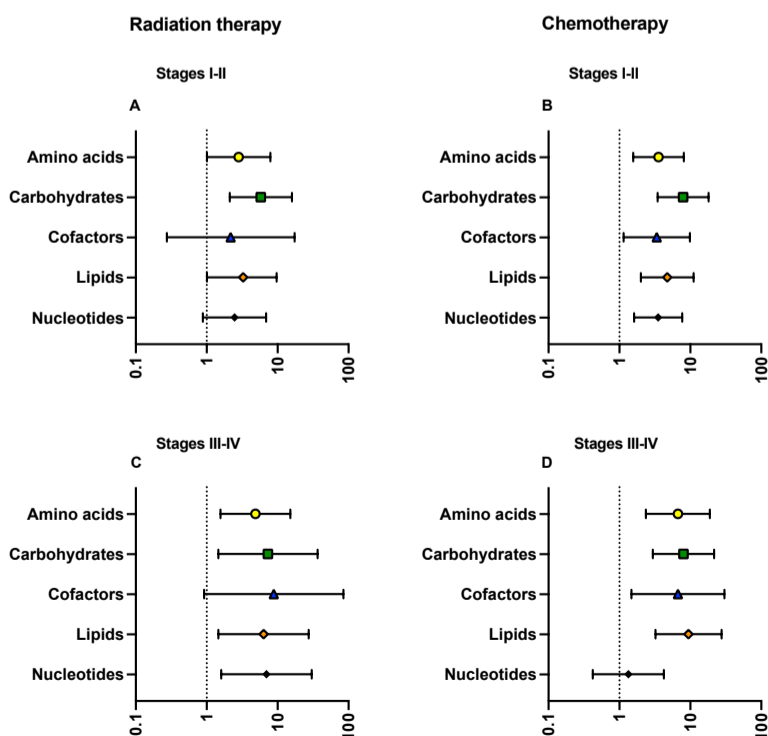


Figure S5 The relative impact of individual metabolic alterations on overall survival of LUAD cohort. Forest plots show the hazard ratio +/- 95% confidence intervals for each clinically relevant metabolic pathway in LUAD from the TCGA database. (A,C) LUAD patients received radiation therapy only. (B,D) LUAD patients received chemotherapy only. (A,B) Patients with early stages of disease (stages I and II are combined), (C,D) Patients with late stages of disease (stages III and IV are combined). LUAD, lung adenocarcinoma.