

SUPPLEMENTARY INFORMATION FOR

Longitudinal Metabolomics in Amyotrophic Lateral Sclerosis Implicates Impaired Lipid Metabolism

Kai Guo, PhD,^{1,2,†} Masha G Savelieff, PhD,^{3,†} Dae-Gyu Jang, PhD,^{1,2} Samuel J Teener,^{1,2} Lili Zhao, PhD,⁴ Junguk Hur, PhD,³ Stephen A Goutman, MD, MS,^{1,2,*} and Eva L Feldman, MD, PhD^{1,2,*}

¹Department of Neurology, University of Michigan, Ann Arbor, MI

²NeuroNetwork for Emerging Therapies, University of Michigan, Ann Arbor, MI

³Department of Biomedical Sciences, University of North Dakota, Grand Forks, ND

⁴Department of Biostatistics, University of Michigan, Ann Arbor, MI

[†]Co-first authors

*Corresponding Authors:

Eva L. Feldman, MD, PhD

109 Zina Pitcher Place

Ann Arbor, MI 48109-5223

Phone: 734-936-8586

Fax: 734-936-5185

e-mail: efeldman@umich.edu

Stephen A Goutman, MD, MS

109 Zina Pitcher Place

Ann Arbor, MI 48109-2200

Phone: 734-936-8586

Fax: 734-936-5185

E-mail: sgoutman@med.umich.edu

Table of contents

Detailed materials and methods	4
<i>Participants and biosamples</i>	4
<i>Metabolomic profiling.....</i>	4
<i>Statistical and data analysis</i>	4
<i>Participant demographics.....</i>	4
<i>Plasma metabolite to ALSFRS-R models.....</i>	5
<i>Plasma metabolomics Model 1: Linear mixed effects model</i>	5
<i>Plasma metabolomics Model 2: Longitudinal interaction model.....</i>	6
<i>Plasma metabolomics Model 3: Progression to event model</i>	7
<i>Plasma metabolomics pathway analysis</i>	7
<i>Plasma metabolomics sex interaction analysis</i>	7
<i>Spinal cord and brain metabolomics analyses</i>	7
<i>Mendelian randomization</i>	8
Supplementary Figures	9
Supplementary Figure S1. Plasma metabolomics linear mixed effects Model 1: Metabolite abundance to ALSFRS-R across all time points	9
Supplementary Figure S2. Plasma metabolomics longitudinal interaction Model 2: Baseline metabolite abundance to predict ALSFRS-R change at 6 months	10
Supplementary Figure S3. Sample and metabolite overlap in the ALS spinal cord and brain cohort	11
Supplementary Figure S4. Enrichment analysis of differential metabolites in onset versus second segment spinal cord tissue in ALS	12
Supplementary Figure S5. Enrichment analysis of Mendelian randomization metabolites	13
Supplementary Tables.....	14
Supplementary Table S1. Plasma metabolites selected by linear mixed effects Model 1.....	14
Supplementary Table S2. Plasma sub-pathways selected by linear mixed effects Model 1 ...	14
Supplementary Table S3. Plasma sub-pathways selected by sex interaction analysis of linear mixed effects Model 1.....	14
Supplementary Table S4. Plasma metabolites selected by longitudinal interaction Model 2 at 6 months	14
Supplementary Table S5. Plasma sub-pathways selected by longitudinal interaction Model 2	14
Supplementary Table S6. Plasma metabolites selected by progression Model 3, PTE20 and PTE50, and enrichment.....	14
Supplementary Table S7. Plasma sub-pathways selected by progression Model 3, PTE20 and PTE50	15
Supplementary Table S8. Metabolites and sub-pathways detected across ALS spinal cord and motor cortex tissue	15

Supplementary Table S9. Metabolites differing in onset versus second spinal cord segment in ALS	15
Supplementary Table S10. Metabolites linked to ALS risk identified by Mendelian randomization.....	15
Supplementary Table S11. Heterogeneity and pleiotropy tests for metabolites from the Mendelian randomization for ALS risk	16
References	17

Detailed materials and methods

Participants and biosamples

Study protocols were previously published.¹⁻³ Briefly, all patients seen at the University of Michigan Pranger ALS Clinic older than 18 years and able to communicate in English were invited to participate and provided oral and written informed consent. ALS patients were diagnosed using the original/revised El Escorial criteria, and all were included meeting the criteria of definite, probable, probable-lab supported, possible (revised criteria),⁴ or suspected (original criteria).⁵ All ALS patients also met the Gold Coast criteria.⁶ For plasma metabolomics, non-fasting plasma samples were obtained by peripheral venipuncture and centrifuged at 2000g for 10 min at 4 °C, aliquoted into cryovials, and stored at -80 °C, following good clinical practice. Samples were collected at baseline (T1) and at approximately 6-month intervals corresponding to routine clinical visits (T2 at 6 months, T3 at 12 months; Figure 1). Brain and spinal cord samples were obtained from patients who agreed to donate tissue following death. The Michigan Brain Bank coordinated sample collection after death, which included informed consent from a legal representative. Brain and spinal cord tissue samples were frozen upon collection. For all participants, demographic and ALS clinical details were abstracted from the medical records, including time from symptom onset and diagnosis, onset segment, and the ALS functional rating score-revised (ALSFRS-R). The University of Michigan institutional review board approved the study and use of samples (HUM00028826).

Metabolomic profiling

Metabolon (Durham, North Carolina) performed untargeted metabolomics profiling of plasma and central nervous tissue samples by ultrahigh-performance liquid chromatography-tandem mass spectroscopy (UPLC-MS/MS).^{7,8} Plasma and central nervous tissue samples were prepared using an automated sample handling system. Several recovery standards were added to each sample for quality control, followed by methanol to extract chemically diverse metabolites, liberate protein-bound metabolites, and precipitate proteins. Samples were shaken vigorously for 2 minutes and centrifuged. Sample extracts were analyzed by reverse-phase UPLC-MS/MS, in positive and negative ion mode, and hydrophilic interaction chromatography UPLC-MS/MS to detect metabolites after sample extraction. Retention time/index, mass-to-charge ratio, and chromatographic data against authenticated standards, identified the metabolites, which were validated by data curation. Metabolomics profiling was performed in two batches (designated cohort 1 and cohort 2 in the prior publication¹). To account for instrument variability, internal standards were added to each sample prior to injection into the mass spectrometers, and the median relative standard deviations were calculated for each standard. The median relative standard deviation for all endogenous metabolites (*i.e.*, non-instrument standards) present in a pool of well-characterized human plasma technical replicates was calculated to determine process variability.

Statistical and data analysis

Participant demographics

Participant demographics were summarized by median (25th – 75th percentile) for continuous variables and counts (percentages) for categorical variables. Significant changes over time were assessed by a linear mixed model.

Plasma metabolite to ALSFRS-R models

A dataset of longitudinal metabolomics and ALSFRS-R measures was leveraged to investigate associations between potential metabolite biomarkers and disease progression. This dataset contained a total of 692 metabolites measured from 548 samples from 317 participants. Three models were constructed: Model 1 (linear mixed effects model), Model 2 (longitudinal interaction model), and Model 3 (progression model).

Plasma metabolomics Model 1: Linear mixed effects model

This model employed all metabolite abundances measured at all time points, *i.e.*, from baseline up to 12 months, and their corresponding ALSFRS-R scores at the time of sample collection to understand the association of metabolites to ALSFRS-R by participant. Model 1 is a linear mixed model accounting for group-level variability with a “participant” random intercept, adjusted for age, sex, BMI at study entry, and onset segment.

For participant “i” in group “j”, the model is given by:

$$ALSFRS_R_{i,j} = \beta_0 + \beta_1 age_{i,j} + \beta_2 sex_{i,j} + \beta_3 BMI_{i,j} + \beta_4 onset_segment_{i,j} + \beta_5 metabolite_{i,j} + \beta_6 participant_j + e_{i,j}$$

- $ALSFRS_R_{i,j}$: The response variable (outcome), representing the ALSFRS-R score for the i-th observation in the j-th participant.
- β_0 : The intercept term, representing the baseline level of ALSFRS-R when all predictors are zero.
- $\beta_1 age_{i,j}$: The fixed effect of **age** for the i-th observation of the j-th participant, where β_1 is the coefficient indicating the effect of age on ALSFRS-R.
- $\beta_2 sex_{i,j}$: The fixed effect of **sex** for each observation, where β_2 is the coefficient indicating the effect of sex on ALSFRS-R.
- $\beta_3 BMI_{i,j}$: The fixed effect of **Body Mass Index (BMI)**, where β_3 is the coefficient representing the effect of BMI on ALSFRS-R.
- $\beta_4 onset_segment_{i,j}$: The fixed effect of **onset segment** (the part of the body where ALS symptoms began), where β_4 measures this effect on ALSFRS-R.
- $\beta_5 metabolite_{i,j}$: The fixed effect of a specific **metabolite** level, where β_5 indicates the effect of the metabolite on ALSFRS-R.
- $\beta_6 participant_j$: A random effect associated with each **participant j**, accounting for the fact that there may be individual differences among participants. This term allows each participant to have a unique baseline, capturing unobserved heterogeneity.
- $e_{i,j}$: The residual error term, representing unexplained variance in $ALSFRS_R_{i,j}$ for each observation.

Explanation of Indices i and j:

- **i**: Refers to each observation within a participant. In longitudinal studies, participants have multiple measurements over time, so i indexes represent these repeated measures.
- **j**: Refers to each unique participant. Random effects are modeled at the participant level, allowing individual variations.

Plasma metabolomics Model 2: Longitudinal interaction model

This model employed baseline metabolite abundance to predict longitudinal change in ALSFRS-R, from baseline to 6 months, to examine disease progression, assuming a constant rate over time. Model 2 is an interaction linear mixed model accounting for group-level variability with a “participant” random intercept, adjusted for age, sex, BMI at study entry, and onset segment. Model 2 builds on Model 1 by including time and an interaction term between “Metabolite” and “Time”, which allows assessment of how the “Metabolite” affects “ALSFRS-R” changes over time.

For participant “i” in group “j”, the model is given by:

$$ALSFRS_R_{i,j} = \beta_0 + \beta_1 age_{i,j} + \beta_2 sex_{i,j} + \beta_3 BMI_{i,j} + \beta_4 onset_segment_{i,j} + \beta_5 metabolite_{i,j} + \beta_6 time_{i,j} + \beta_7 metabolite_{i,j} * time_{i,j} + \beta_8 participant_j + e_{i,j}$$

- $ALSFRS_R_{i,j}$: The response variable (outcome), representing the ALSFRS-R score for the i-th observation in the j-th participant.
- β_0 : The intercept term, representing the baseline level of ALSFRS-R when all predictors are zero.
- $\beta_1 age_{i,j}$: The fixed effect of **age** for the i-th observation of the j-th participant, where β_1 is the coefficient indicating the effect of age on ALSFRS-R.
- $\beta_2 sex_{i,j}$: The fixed effect of **sex** for each observation, where β_2 is the coefficient indicating the effect of sex on ALSFRS-R.
- $\beta_3 BMI_{i,j}$: The fixed effect of **Body Mass Index (BMI)**, where β_3 is the coefficient representing the effect of BMI on ALSFRS-R.
- $\beta_4 onset_segment_{i,j}$: The fixed effect of **onset segment** (the part of the body where ALS symptoms began), where β_4 measures this effect on ALSFRS-R.
- $\beta_5 metabolite_{i,j}$: The fixed effect of a specific **metabolite** level, where β_5 indicates the effect of the metabolite on ALSFRS-R.
- $\beta_6 time_{i,j}$: The fixed effect of **time** (likely representing time since diagnosis or another time measure), where β_6 is the coefficient showing how ALSFRS-R changes over time.
- $\beta_7 metabolite_{i,j} * time_{i,j}$: The interaction term between **metabolite** and **time**, where β_7 represents how the impact of the metabolite on ALSFRS-R varies over time.
- $\beta_8 participant_j$: A random effect associated with each **participant j**, accounting for the fact that there may be individual differences among participants. This term allows each participant to have a unique baseline, capturing unobserved heterogeneity.
- $e_{i,j}$: The residual error term, representing unexplained variance in $ALSFRS_R_{i,j}$ for each observation.

Explanation of Indices i and j”

- **i**: Refers to each observation within a participant. In longitudinal studies, participants have multiple measurements over time, so iii indexes these repeated measures.
- **j**: Refers to each unique participant. Random effects are modeled at the participant level, allowing individual variations.

Missing ALSFRS-R scores were imputed with the average values between the data points preceding and following the time point of the missing value. BH-adjusted p-values<0.05 were

computed for Model 2 (**Supplementary Tables S4, S5**) but results were reported using a nominal p-value<0.05 due to the more exploratory nature of this analysis, with lower power to detect due to the added variable of time versus Model 1.

Plasma metabolomics Model 3: Progression to event model

This is a Cox regression model, which correlated baseline metabolite abundance to progression to an event (PTE) based on longitudinal ALSFRS-R scores. The PTE event for each participant was defined as either a 20% (PTE20) or 50% (PTE50) decline in ALSFRS-R relative to the diagnostic ALSFRS-R score. PTE20 was defined as the first longitudinal ALSFRS-R score that was $\leq 0.8 \times$ initial ALSFRS-R or death. PTE50 was defined as the first longitudinal ALSFRS-R score that was $\leq 0.5 \times$ initial ALSFRS-R or death. The PTE thresholds were selected based on a survey of ALS clinicians and researchers, where 90% of respondents considered a 20% drop in ALSFRS-R as “somewhat clinically significant” and a 50% drop as “very clinically meaningful”.⁹ PTE20 and PTE50 are defined regardless of subsequent decline or death status. Participants who did not experience PTE20 or PTE50 or were lost to follow-up were censored. The model included 316 ALS participants and was adjusted for age at initial sample, sex, BMI at entry, onset segment (bulbar versus non-bulbar, i.e., cervical, lumbar), diagnostic ALSFRS-R, time between symptom onset and diagnosis, and definite versus non-definite (i.e., probable, probable-laboratory supported, possible, suspected) diagnostic El Escorial criteria. BH-adjusted p-values<0.05 were computed for Model 3 (**Supplementary Tables S6-S7**), but results were reported using a nominal p-value<0.05 due to the more exploratory nature of this analysis, with lower power to detect due to the right-censored data.

Plasma metabolomics pathway analysis

Gene set variation analysis (GSVA) was performed to assess the pathway activity in each participant sample using the R GSVA package.¹⁰ GSVA is an unsupervised gene set enrichment method that estimates the variation in a pathway’s activity within a sample population. The GSVA pathway score, the normalized enrichment statistic, was calculated based on the abundance of all metabolites within each pathway. The three models (mixed effects, longitudinal interaction, progression to event) were repeated using GSVA instead of individual metabolites to determine the significant sub-pathways associated with changes in ALSFRS-R.

Plasma metabolomics sex interaction analysis

To examine the effect of sex on the correlation of metabolites to ALSFRS-R, a sex interaction analysis was performed on Model 1 (mixed effects) and Model 2 (longitudinal interaction). The interaction effects between sex and pathway were added as covariates into the models. Moreover, the models were adjusted for age, BMI, and onset segment. A nominal p-value<0.05 was considered significant in all models.

Spinal cord and brain metabolomics analyses

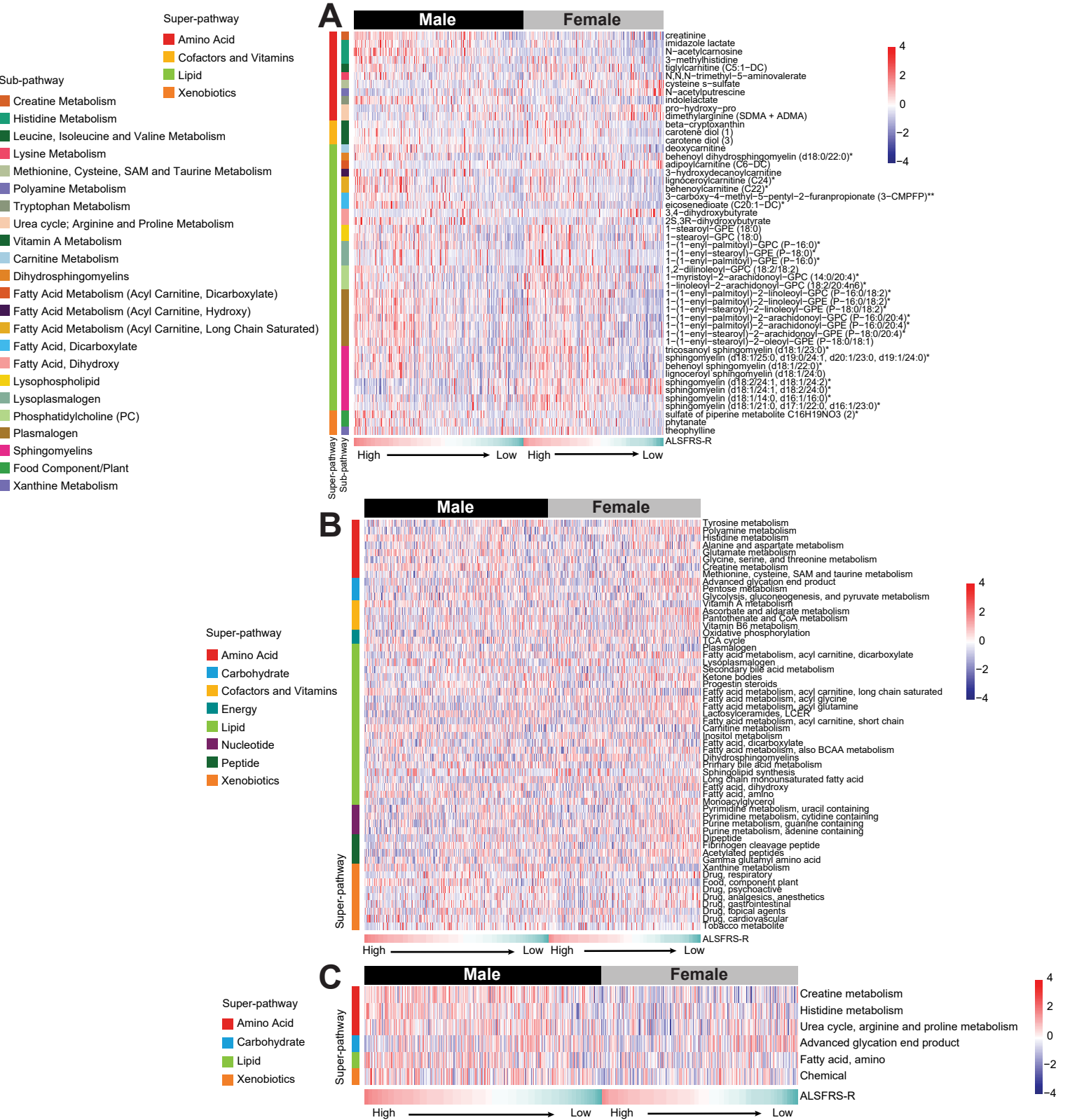
Metabolites with over 20% missingness across samples were removed from downstream analysis. The data were log-transformed, and missing values were imputed using the minimum observed value for each metabolite. The analysis first examined unique and overlapping metabolites detected across all samples since they spanned diverse sources, including lumbar

spinal cord, cervical spinal cord, and brain motor cortex. Sub-pathways and super-pathways were then summarized for the unique and overlapping metabolites. Next, an analysis of differential metabolites in the onset segment versus a second segment in spinal cord was used as a surrogate of disease progression. For participants with lumbar onset, second segment of disease progression was cervical spinal cord. For participants with cervical onset, the second segment of disease progression was lumbar spinal cord. A pairwise Wilcoxon test between onset versus second segment identified metabolites that significantly differed with disease progression based on a nominal p-value<0.05.

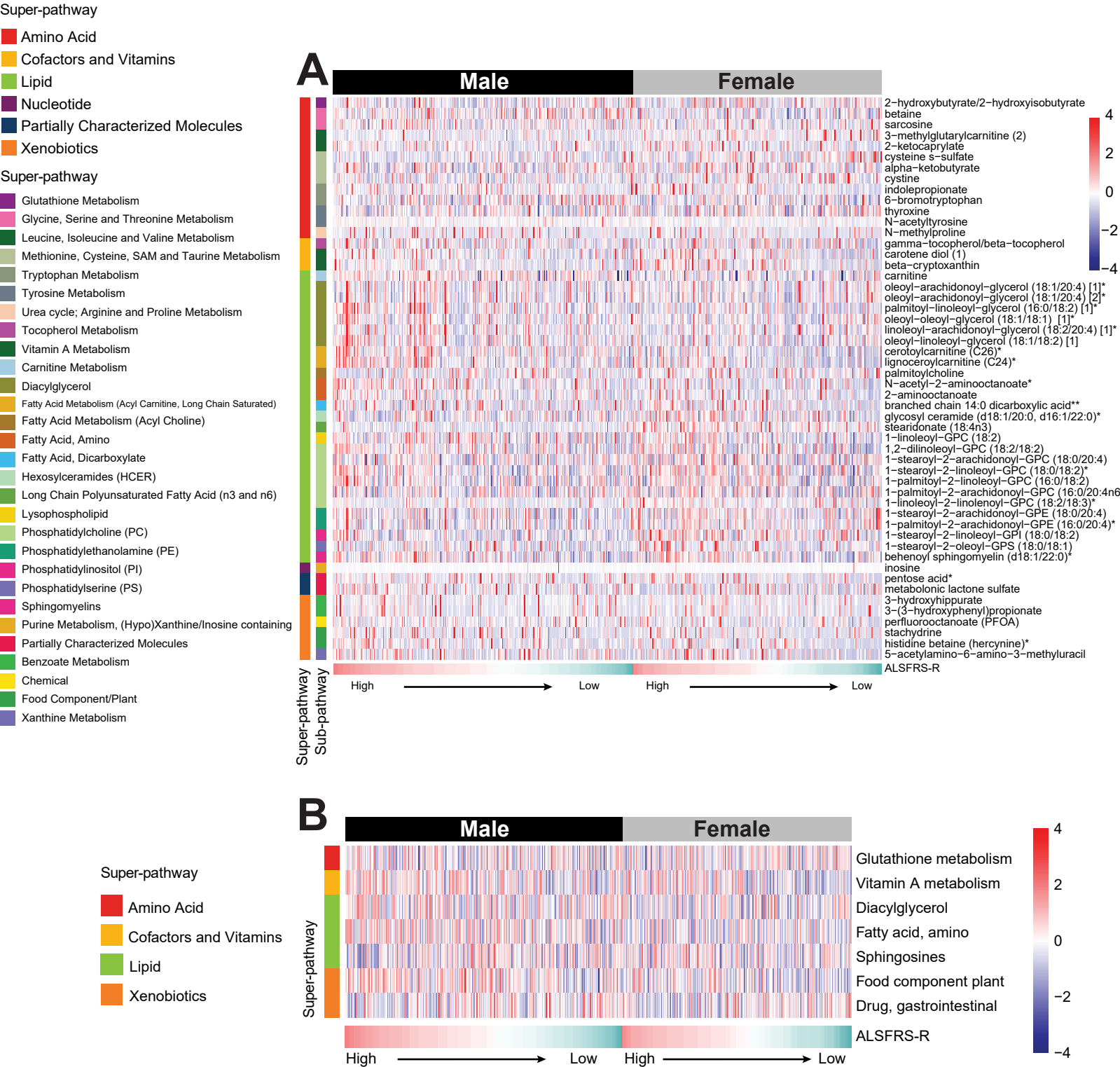
Mendelian randomization

Mendelian randomization (MR) was performed using single-nucleotide polymorphisms (SNPs) from a published genome wide association study of ALS patients (n=27,205) and controls (n=110,881) of European ancestry, accession number GCST90027164.¹¹ For the exposure dataset, a genome wide association study of SNPs from 8,299 participants of European ancestry related to 1,702 plasma metabolites was employed, accession number GCST90199621-90201020.¹² Datasets of European ancestry were selected since participants in this study were 97% White. SNPs in the GCST90027164 dataset were annotated to GrCh37 version of the human genome while SNPs to the GCST90199621-90201020 dataset were annotated to GrCh38; therefore, to conduct the MR analysis, SNPs from the GCST90199621-90201020 dataset were converted from GrCh38 to GrCh37 human genome annotation. For the MR analysis, SNPs were filtered to the criteria: p-value<5x10⁻⁵, linkage disequilibrium (LD) R²<0.01, distance 1 Mb, F-statistic>10. A p-value cutoff of 5x10⁻⁵ was selected due to the large number of metabolite SNPs and ALS SNPs, which is aligned with prior MR analyses in ALS.^{13,14} Significance was determined according to the criteria: inverse variance weighted (IVW) p-value<0.05, pleiotropy test p-value>0.05, and heterogeneity test p-value>0.05. The MR identified 114 metabolites with IVW p-value<0.05, of which 98 metabolites had a known name.

Supplementary Figure S1. Plasma metabolomics linear mixed effects Model 1: Metabolite abundance to ALSFRS-R across all time points. (A) Heatmap visualization of the relative abundance in the top 50 metabolites out of 80 (**Supplementary Table S1**) with adjusted p-value (Padj)<0.05 and |beta|>0.2 to ALSFRS-R. **(B)** Heatmap visualization of the relative abundance in enriched sub-pathways with Padj<0.05 to ALSFRS-R (**Supplementary Table S2**). **(C)** Heatmap visualization of relative abundance in enriched sub-pathways with a sex interaction with p-value<0.05 to ALSFRS-R (**Supplementary Table S3**). Heatmaps grouped by male (black) and female (grey) and sorted by ALSFRS-R, from high (red; better functional status) to low score (turquoise; poorer functional status). Relative abundance scaled from higher (red) to lower (blue) according to the legend. Single asterisk * indicates a compound that has not been confirmed based on a standard, but that the analytical platform is confident in its identity. Double asterisk ** indicates a compound for which a standard is not available, but that the analytical platform is reasonably confident in its identity, or the information provided. ADMA, asymmetric dimethylarginine; GPC, glycerophosphocholine; GPE, glycerophosphoethanolamine; SDMA symmetric dimethylarginine. Legends ordered according to metabolite/sub-pathway order in heatmaps.

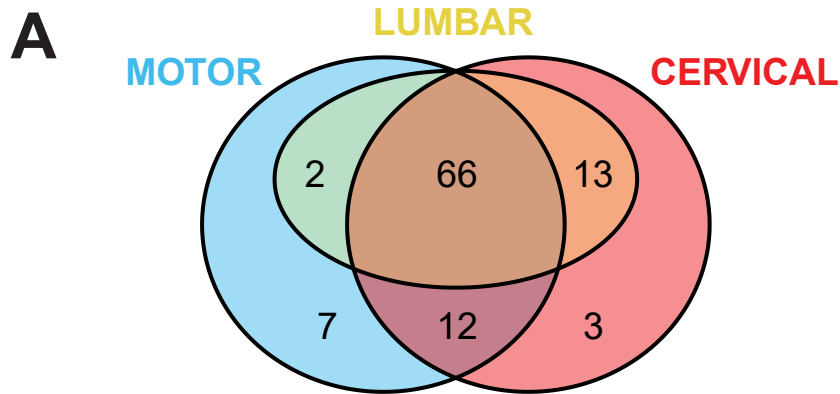


Supplementary Figure S2. Plasma metabolomics longitudinal interaction Model 2: Baseline metabolite abundance to predict ALSFRS-R change at 6 months. (A) Heatmap visualization of the relative abundance of the 52 baseline metabolites (**Supplementary Table S4**) with nominal p-value<0.05 to change in ALSFRS-R. **(B)** Heatmap visualization of the relative abundance of enriched sub-pathways with p-value<0.05 to change in ALSFRS-R (**Supplementary Table S5**). Heatmaps grouped by male (black) and female (grey) and sorted by ALSFRS-R, from high (red; better functional status) to low score (turquoise; poorer functional status). Relative abundance scaled from higher (red) to lower (blue) according to the legend. Single asterisk * indicates a compound that has not been confirmed based on a standard, but that the analytical platform is confident in its identity. GPC, glycerophosphocholine; GPE, glycerophosphoethanolamine. Legends ordered according to metabolite/sub-pathway order in heatmaps.



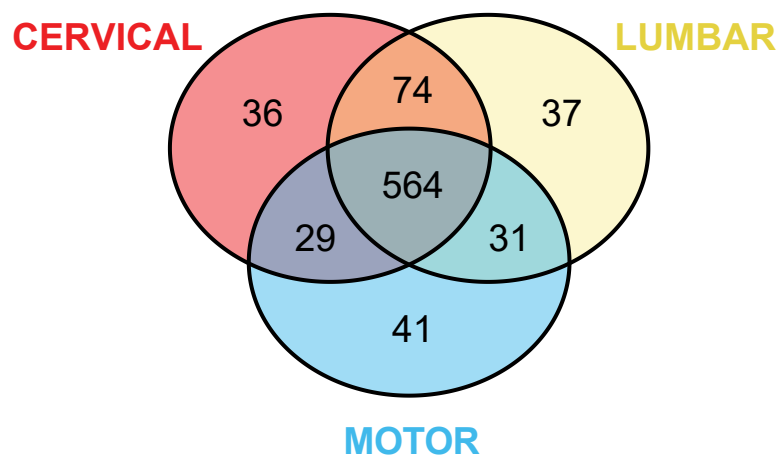
Supplementary Figure S3. Sample and metabolite overlap in the ALS spinal cord and brain cohort

(A) Venn diagram showing overlap of samples provided by ALS participants in the brain and spinal cord cohort.
 (B) Top: Table of total metabolites, metabolites with <20% missingness, and annotated metabolites detected across all tissue samples, cervical spinal cord, lumbar spinal cord, and motor cortex. Bottom: Venn diagram showing overlap of metabolites detected across all tissue samples.



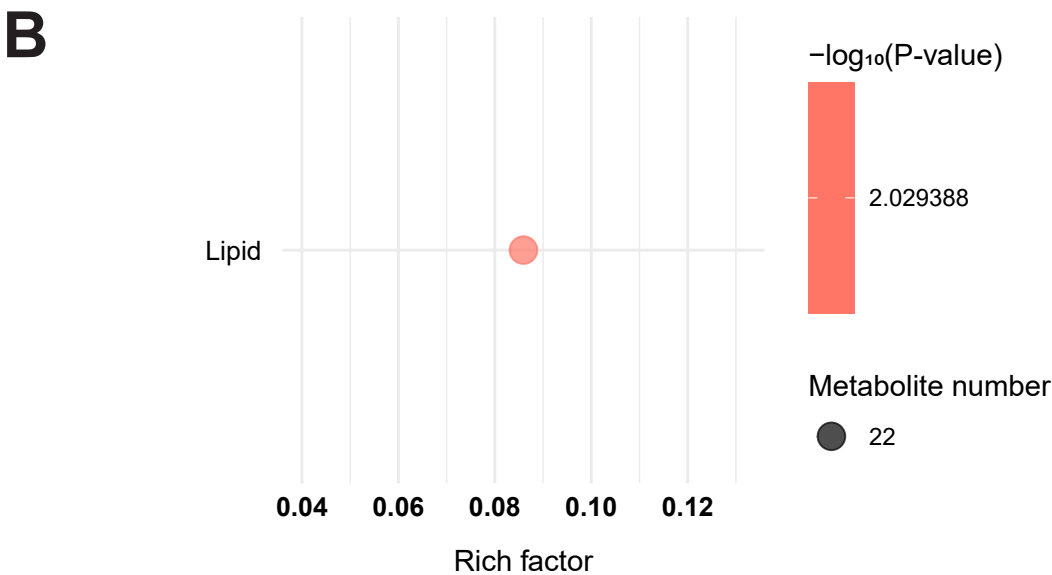
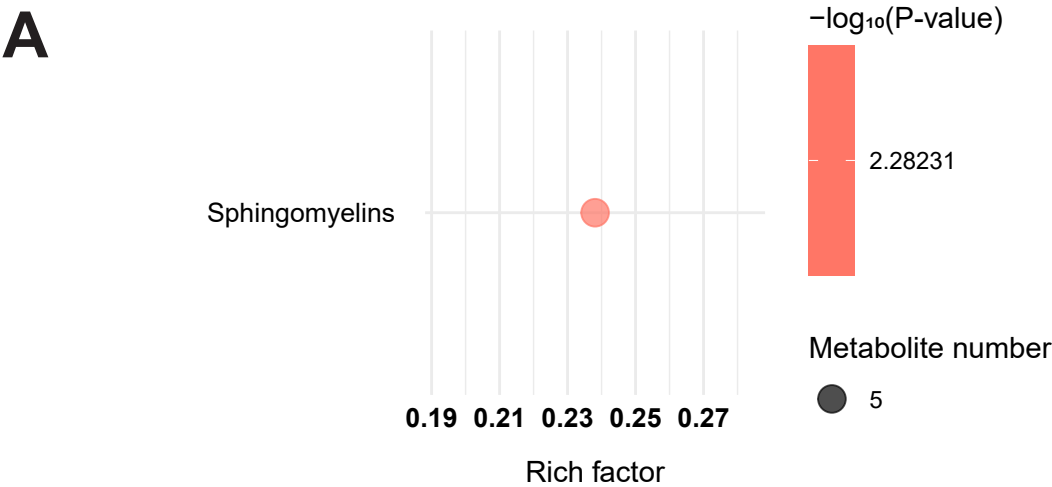
B

	Cervical	Lumbar	Motor
Total metabolites	946	1041	866
<20% missingness	734	764	701
Annotated metabolites	703	706	665



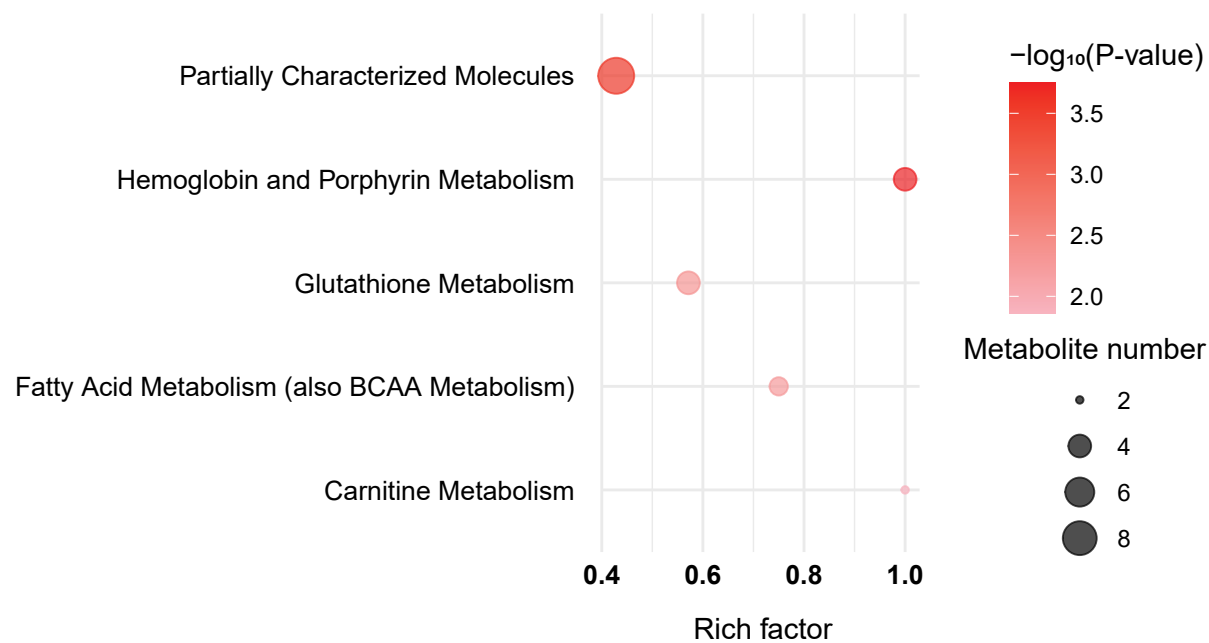
Supplementary Figure S4. Enrichment analysis of differential metabolites in onset versus second segment spinal cord tissue in ALS

Dot plot of enrichment analysis by **(A)** sub-pathway and **(B)** super-pathway of differential metabolites in onset versus second segment spinal cord tissue in ALS. Dot color represents least (light pink) to most (red) significant enriched sub-pathways; dot size represents number of metabolite or sub-pathway members in sub- and super-pathways, respectively.



Supplementary Figure S5. Enrichment analysis of Mendelian randomization metabolites

Dot plot of enrichment analysis of MR metabolites (Supplementary Table S9). Dot color represents least (light pink) to most (red) significant enriched sub-pathways; dot size represents number of metabolite members in sub-pathways.



Supplementary Tables

Supplementary Table S1. Plasma metabolites selected by linear mixed effects Model 1

Sheet 1: Metabolite abundances that correlated significantly to ALSFRS-R scores across all time points by linear mixed effects Model 1. **Sheet 2:** Metabolite abundances that correlated significantly with $P_{adj} < 0.05$ and $|\beta| > 0.2$ to ALSFRS-R scores across all time points by linear mixed effects Model 1. P_{adj} , Benjamini-Hochberg adjusted p-value.

Single asterisk * indicates a compound that has not been confirmed based on a standard, but that the analytical platform is confident in its identity. Double asterisk ** indicates a compound for which a standard is not available, but that the analytical platform is reasonably confident in its identity, or the information provided.

Supplementary Table S2. Plasma sub-pathways selected by linear mixed effects Model 1

Sub-pathways selected by gene set variation analysis that correlated significantly to ALSFRS-R scores across all time points by linear mixed effects Model 1. P_{adj} , Benjamini-Hochberg adjusted p-value.

Supplementary Table S3. Plasma sub-pathways selected by sex interaction analysis of linear mixed effects Model 1

Sub-pathways selected by gene set variation analysis that correlated significantly to ALSFRS-R scores across all time points in the sex interaction analysis by linear mixed effects Model 1. P_{adj} , Benjamini-Hochberg adjusted p-value.

Supplementary Table S4. Plasma metabolites selected by longitudinal interaction Model 2 at 6 months

Baseline metabolite abundances that correlated to change in ALSFRS-R at 6 months in the longitudinal interaction Model 2. P_{adj} , Benjamini-Hochberg adjusted p-value.

Single asterisk * indicates a compound that has not been confirmed based on a standard, but that the analytical platform is confident in its identity. Double asterisk ** indicates a compound for which a standard is not available, but that the analytical platform is reasonably confident in its identity, or the information provided.

Supplementary Table S5. Plasma sub-pathways selected by longitudinal interaction Model 2

Baseline sub-pathway abundances that correlated to change in ALSFRS-R at 6 months in the longitudinal interaction Model 2. P_{adj} , Benjamini-Hochberg adjusted p-value.

Supplementary Table S6. Plasma metabolites selected by progression Model 3, PTE20 and PTE50, and enrichment

Sheet 1: PTE20, defined as 20% decline relative to diagnostic ALSFRS-R. **Sheet 2:** PTE50, defined as 50% decline relative to diagnostic ALSFRS-R. **Sheet 3:** Pathways enrichment analysis of PTE20 metabolites. **Sheet 4:** Pathways enrichment analysis of PTE50 metabolites. The model was adjusted for age at initial sample, sex, BMI at entry, onset segment (bulbar vs

non-bulbar), diagnostic ALSFRS-R, time between symptom onset and diagnosis, and definite vs non-definite diagnostic EI-Escorial. HR (CI), hazard ratio (confidence interval); Padj, Benjamini-Hochberg adjusted p-value. In the enrichment analysis, “annotated” is the number of member metabolites in the sub-pathway, “significant” is the number of metabolites selected by PTE20 or PTE50.

Single asterisk * indicates a compound that has not been confirmed based on a standard, but that the analytical platform is confident in its identity. Double asterisk ** indicates a compound for which a standard is not available, but that the analytical platform is reasonably confident in its identity, or the information provided.

Supplementary Table S7. Plasma sub-pathways selected by progression Model 3, PTE20 and PTE50

Sheet 1: PTE20, defined as 20% decline relative to diagnostic ALSFRS-R. **Sheet 2:** PTE50, defined as 50% decline relative to diagnostic ALSFRS-R. The model was adjusted for age at initial sample, sex, BMI at entry, onset segment (bulbar vs non-bulbar), diagnostic ALSFRS-R, time between symptom onset and diagnosis, and definite vs non-definite diagnostic EI-Escorial. HR (CI), hazard ratio (confidence interval); Beta, estimated beta value; lcl_beta, lower limit of 95% confidence interval of beta; ucl_beta, upper limit of 95% confidence interval of beta; Padj, Benjamini-Hochberg adjusted p-value.

Supplementary Table S8. Metabolites and sub-pathways detected across ALS spinal cord and motor cortex tissue

The 564 metabolites detected in all three tissues, cervical spinal cord, lumbar spinal cord, and motor cortex (**sheet 1**). Metabolites shared by cervical and lumbar spinal cord (74 metabolites, **sheet 2**), cervical spinal cord and motor cortex (29 metabolites, **sheet 3**), and lumbar spinal cord and motor cortex (31 metabolites, **sheet 4**). Metabolites unique to cervical spinal cord (36 metabolites, **sheet 5**), lumbar spinal cord (37 metabolites, **sheet 6**), and motor cortex (41 metabolites, **sheet 7**) tissue.

Single asterisk * indicates a compound that has not been confirmed based on a standard, but that the analytical platform is confident in its identity. Double asterisk ** indicates a compound for which a standard is not available, but that the analytical platform is reasonably confident in its identity, or the information provided.

Supplementary Table S9. Metabolites differing in onset versus second spinal cord segment in ALS

Metabolite abundances that correlated with onset versus second spinal cord segment in as a surrogate of ALS progression. Padj, Benjamini-Hochberg adjusted p-value.

Single asterisk * indicates a compound that has not been confirmed based on a standard, but that the analytical platform is confident in its identity. Double asterisk ** indicates a compound for which a standard is not available, but that the analytical platform is reasonably confident in its identity, or the information provided.

Supplementary Table S10. Metabolites linked to ALS risk identified by Mendelian randomization

Sheet 1: MR metabolites. CEHC, carboxyethyl hydroxymethoxy; GPC, glycerophosphocholine; GPE, glycerophosphoethanolamine. **Sheet 2:** Pathways enrichment analysis of MR metabolites. In the enrichment analysis, “annotated” is the number of member metabolites in the sub-pathway, “significant” is the number of metabolites selected by MR analysis. **Sheet 3:** Overlap of MR metabolites with nominally significant metabolites selected by Models 1, 2, 3, and the model of onset versus second segment.

N SNPs, number of single-nucleotide polymorphisms; Std error, standard error; inverse variance weighted p-value, inverse variance weighted p-value; Lo_CI, lower bound of confidence interval; Up_CI, upper bound of confidence interval; OR, odds ratio; OR_L_CI95, lower bound of the 95% confidence interval for the odds ratio; OR_U_CI95, upper bound of the 95% confidence interval for the odds ratio; Padj, Benjamini-Hochberg adjusted p-value.

Single asterisk * indicates a compound that has not been confirmed based on a standard, but that the analytical platform is confident in its identity. Double asterisk ** indicates a compound for which a standard is not available, but that the analytical platform is reasonably confident in its identity, or the information provided.

Supplementary Table S11. Heterogeneity and pleiotropy tests for metabolites from the Mendelian randomization for ALS risk

Sheet 1: Heterogeneity test using Cochran’s Q Statistic. **Sheet 2:** Pleiotropy test using MR-Egger Intercept.

References

1. Goutman SA, Guo K, Savelieff MG, et al. Metabolomics identifies shared lipid pathways in independent amyotrophic lateral sclerosis cohorts. *Brain : a journal of neurology*. Jan 28 2022;doi:10.1093/brain/awac025
2. Goutman SA, Boss J, Guo K, et al. Untargeted metabolomics yields insight into ALS disease mechanisms. *J Neurol Neurosurg Psychiatry*. Dec 2020;91(12):1329-1338. doi:10.1136/jnnp-2020-323611
3. Su FC, Goutman SA, Chernyak S, et al. Association of Environmental Toxins With Amyotrophic Lateral Sclerosis. *JAMA Neurol*. Jul 1 2016;73(7):803-11. doi:10.1001/jamaneurol.2016.0594
4. Brooks BR, Miller RG, Swash M, Munsat TL. El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotroph Lateral Scler Other Motor Neuron Disord*. Dec 2000;1(5):293-9.
5. Brooks BR. El Escorial World Federation of Neurology criteria for the diagnosis of amyotrophic lateral sclerosis. Subcommittee on Motor Neuron Diseases/Amyotrophic Lateral Sclerosis of the World Federation of Neurology Research Group on Neuromuscular Diseases and the El Escorial "Clinical limits of amyotrophic lateral sclerosis" workshop contributors. *J Neurol Sci*. Jul 1994;124 Suppl:96-107.
6. Shefner JM, Al-Chalabi A, Baker MR, et al. A proposal for new diagnostic criteria for ALS. *Clin Neurophysiol*. Aug 2020;131(8):1975-1978. doi:10.1016/j.clinph.2020.04.005
7. Dehaven CD, Evans AM, Dai H, Lawton KA. Organization of GC/MS and LC/MS metabolomics data into chemical libraries. *J Cheminform*. Oct 18 2010;2(1):9. doi:10.1186/1758-2946-2-9
8. Evans A, Bridgewater B, Liu Q, et al. High Resolution Mass Spectrometry Improves Data Quantity and Quality as Compared to Unit Mass Resolution Mass Spectrometry in High-Throughput Profiling Metabolomics. *Metabolomics*. 2014;4:132.
9. Castrillo-Viguera C, Grasso DL, Simpson E, Shefner J, Cudkowicz ME. Clinical significance in the change of decline in ALSFRS-R. *Amyotroph Lateral Scler*. 2010;11(1-2):178-80. doi:10.3109/17482960903093710
10. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics*. Jan 16 2013;14:7. doi:10.1186/1471-2105-14-7
11. van Rheenen W, van der Spek RAA, Bakker MK, et al. Common and rare variant association analyses in amyotrophic lateral sclerosis identify 15 risk loci with distinct genetic architectures and neuron-specific biology. *Nat Genet*. Dec 2021;53(12):1636-1648. doi:10.1038/s41588-021-00973-1
12. Chen Y, Lu T, Pettersson-Kymmer U, et al. Genomic atlas of the plasma metabolome prioritizes metabolites implicated in human diseases. *Nat Genet*. Jan 2023;55(1):44-53. doi:10.1038/s41588-022-01270-1
13. Julian TH, Glasgow N, Barry ADF, et al. Physical exercise is a risk factor for amyotrophic lateral sclerosis: Convergent evidence from Mendelian randomisation, transcriptomics and risk genotypes. *EBioMedicine*. May 26 2021;68:103397. doi:10.1016/j.ebiom.2021.103397
14. Julian TH, Boddy S, Islam M, et al. A review of Mendelian randomization in amyotrophic lateral sclerosis. *Brain*. Apr 29 2022;145(3):832-842. doi:10.1093/brain/awab420