

Supplementary Methods

Data cleaning - Human Serum AMD vs Controls

Given their strongly skewed distribution, metabolic abundances were log-transformed to achieve distribution symmetry. Metabolic abundances were normalised by dividing the global median log-abundance of area/batch combination. Principal components plot revealed 3 subjects belonging to the same batch which were global outliers, these were discarded from further analyses (**Figure S11**).

Before imputation we discarded 209 metabolites presenting missingness rates higher than 20% in any of the 4 subgroups. Imputation was performed in the same fashion as in the primate study with 132 metabolites imputed assuming MCAR missingness and 110 with MNAR missingness. Quantile normalisation was performed as described in the primate study. No metabolites and no further samples were discarded after quantile normalization. The final dataset was composed by 348 samples and 763 metabolites. Metabolites were divided into 50 groups highlighting their biological pathways. The list of metabolites and relative groups is available in **Table S3**.

Statistical Analysis - Human Serum AMD vs Controls

Metabolic differential abundance was tested using the same approach as described for the primate study. In model contained in age, sex at birth and processing batch as covariates. No intra-subject correlation was taken into account since only one observation was included per subject for each metabolite. To achieve maximal discovery power, subject-specific weights were calculated using the function *arrayWeights* from the Limma package and were included in the models. We performed two main analyses, one testing for all AMD patients against controls and another one testing each sub disease again controls. Enrichment analysis was performed for both main analyses as previously described and no principal component analysis was performed for this study.

Supplementary Results

Metabolite missingness in primate data

Given the relatively small number of detected metabolites and their strong correlations with each other, in this study, we used a somewhat forgiving missingness threshold to discard a-priority any metabolite (>80%). However, we found that only 44 significant metabolites had an initial missingness rate greater than 20% and only 23 greater than 50% (**Table S3**). We tested whether significance was higher for those metabolites presenting different missingness rates between areas as a quality check. We found that significant metabolites had an average standard deviation of missingness across areas lower than those non-significant (AveMissSD 0.055 vs 0.068) which confirmed that our results were not biased towards metabolites presenting different missingness rates across areas.

Metabolite differential abundance AMD vs Controls

We found a total of 4 metabolites whose abundance was significantly different between AMD cases and controls. All these were depleted and only significant when comparing CNV-AMD to controls and only 1 of these was significant when comparing all AMD cases against controls (**Table S3, Figure S12**).

Pathway differential abundance AMD vs Controls

We then investigated whether metabolites belonging to the same metabolic pathway shared a similar abundance difference across retinal areas. We found two biological pathways to be enriched with metabolites sharing similar signals for disease status. Both were only significant when comparing CNV-AMD to controls and only one of these was significant when comparing all AMD cases against controls (**Table S3, Figure S13**).

Supplementary Figures

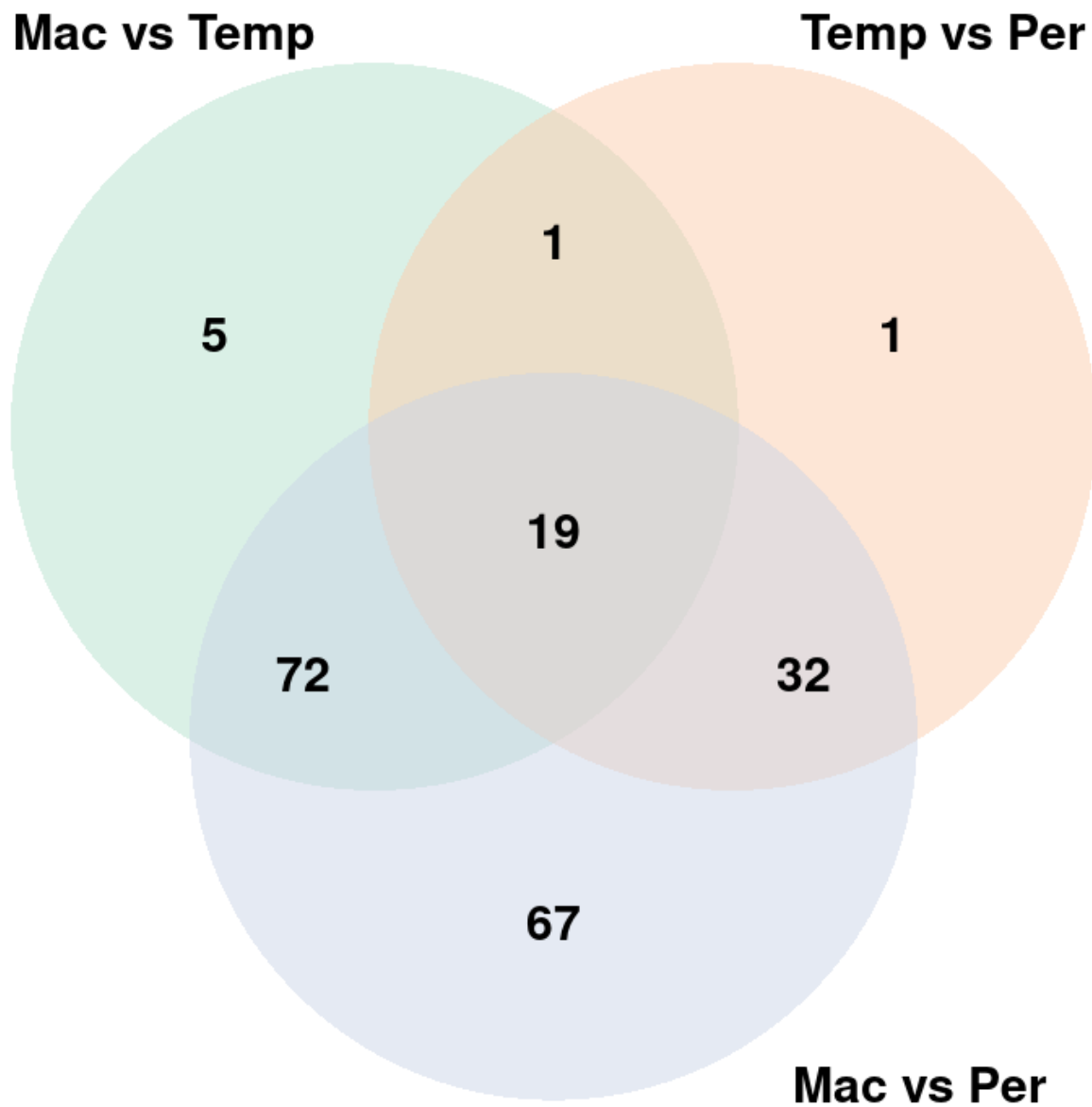


Figure S1: Venn diagram representing frequencies of significant metabolite by the three contrasts *Macula vs Temporal* (Mac vs Temp), *Temporal vs Periphery* (Temp vs Per), and *Macula vs Periphery* (Mac vs Per).



Figure S2: Metabolic distribution of biochemical families between retinal areas. In this image, every boxplot represents the covariate-corrected distribution of one metabolite whose abundance was significantly different for any of the three contrasts. Metabolites are divided into biochemical families which are sorted according to greater family-wise abundance in the macula (top) or in the periphery (bottom).

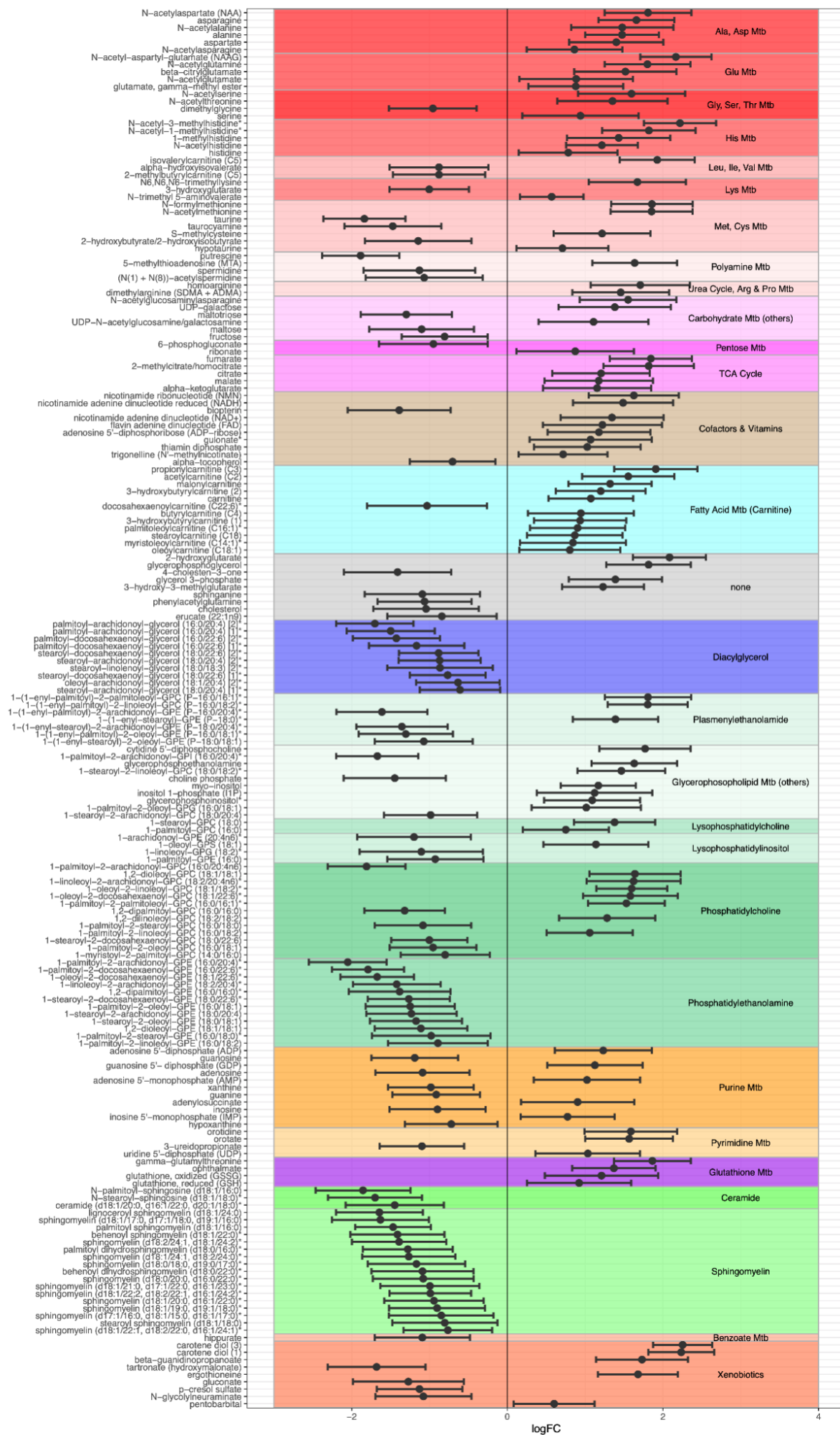


Figure S3: Log-fold changes and 95% confidence interval of metabolites with significant differential abundance between macula and periphery. Positive log-fold changes values in this figure indicate that the metabolite abundance was higher in the macula while negative represent higher abundance in the periphery. Metabolites have been divided and coloured by their respective biological pathways.

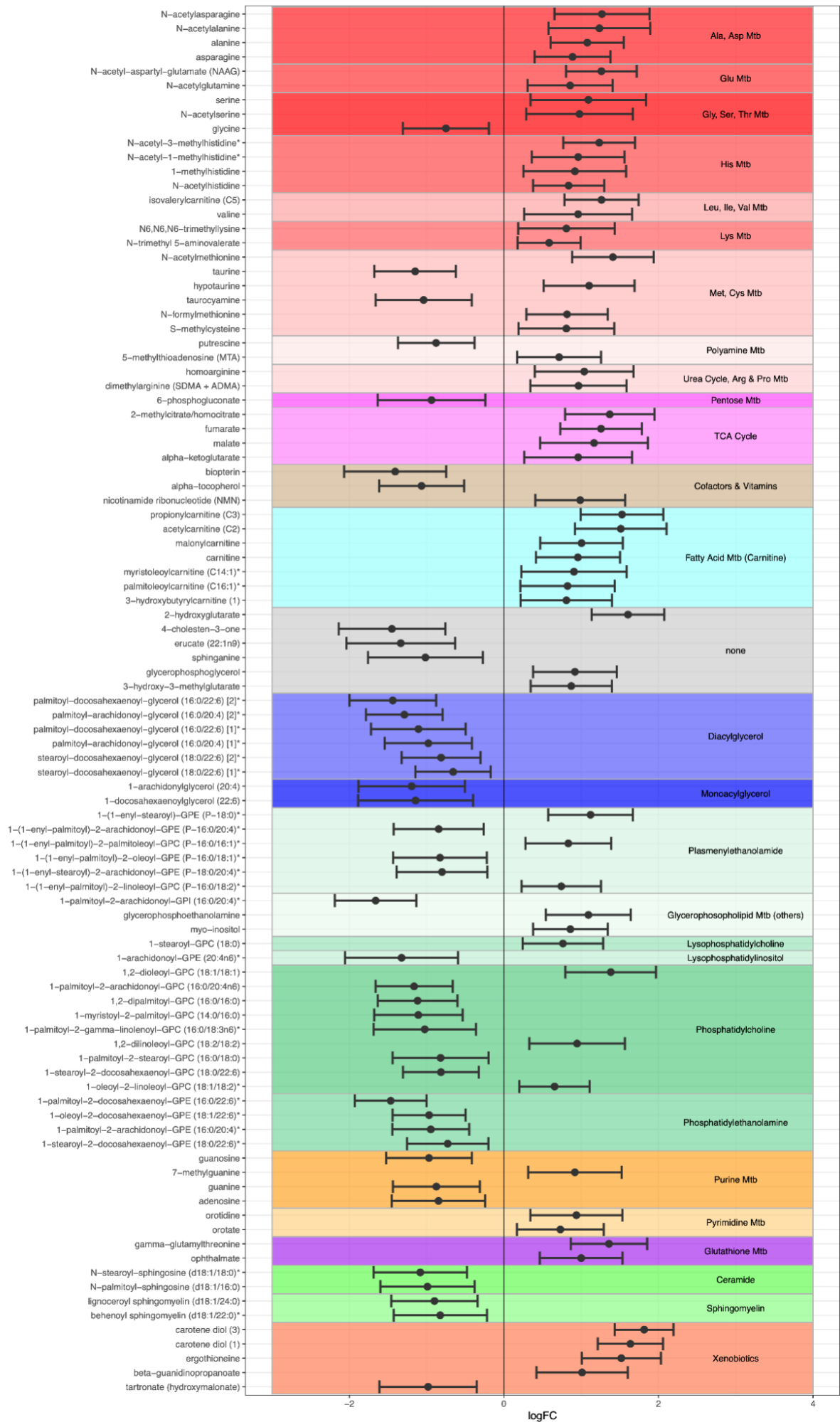


Figure S4: Log-fold changes and 95% confidence interval of metabolites with significant differential abundance between macula and temporal. Positive log-fold changes values in this figure indicate that the metabolite abundance was higher in the macula while negative represent higher abundance in the periphery. Metabolites have been divided and coloured by their respective biological pathways.

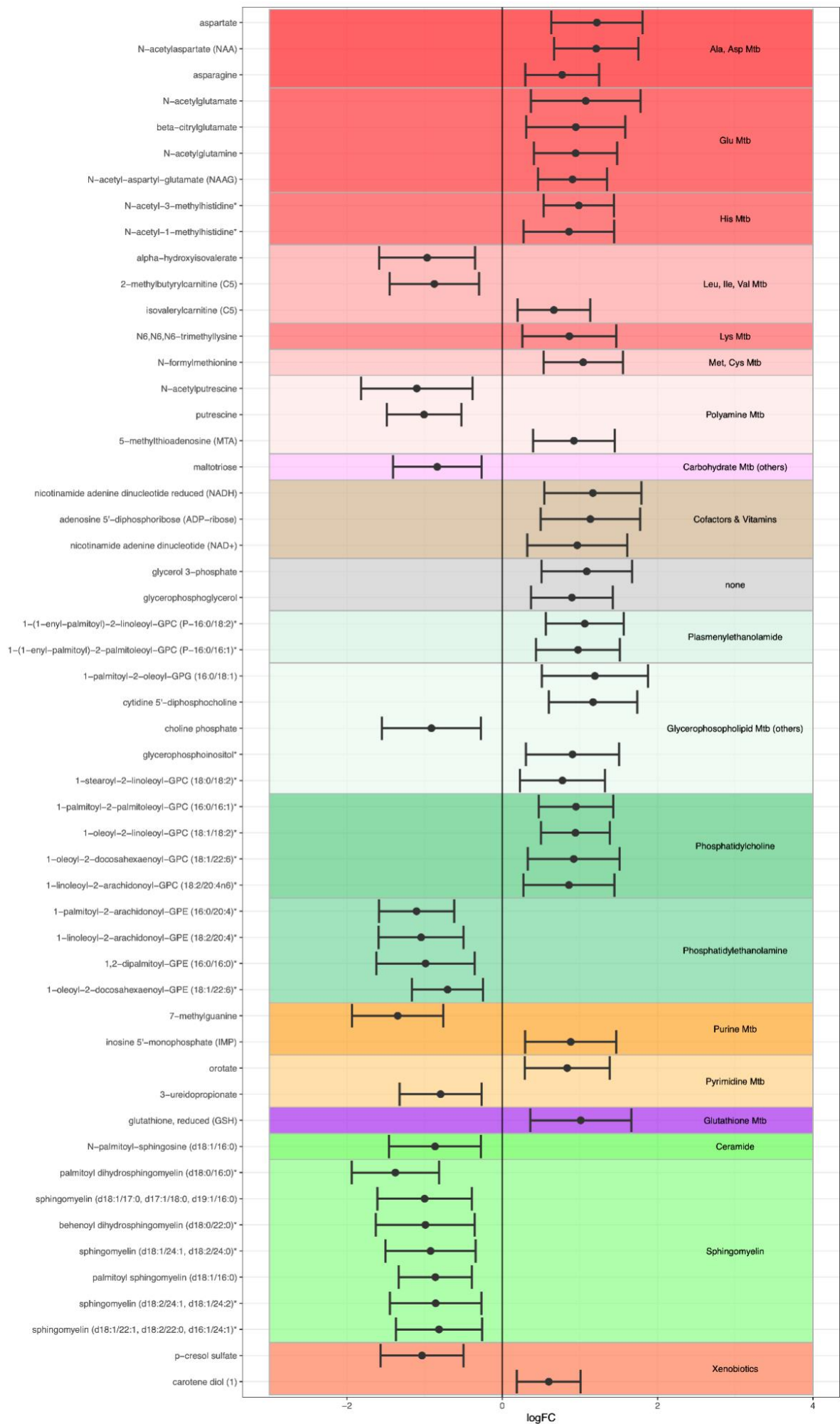


Figure S5: Log-fold changes and 95% confidence interval of metabolites with significant differential abundance between temporal and periphery. Positive log-fold changes values in this figure indicate that the metabolite abundance was higher in the macula while negative represent higher abundance in the periphery. Metabolites have been divided and coloured by their respective biological pathways.

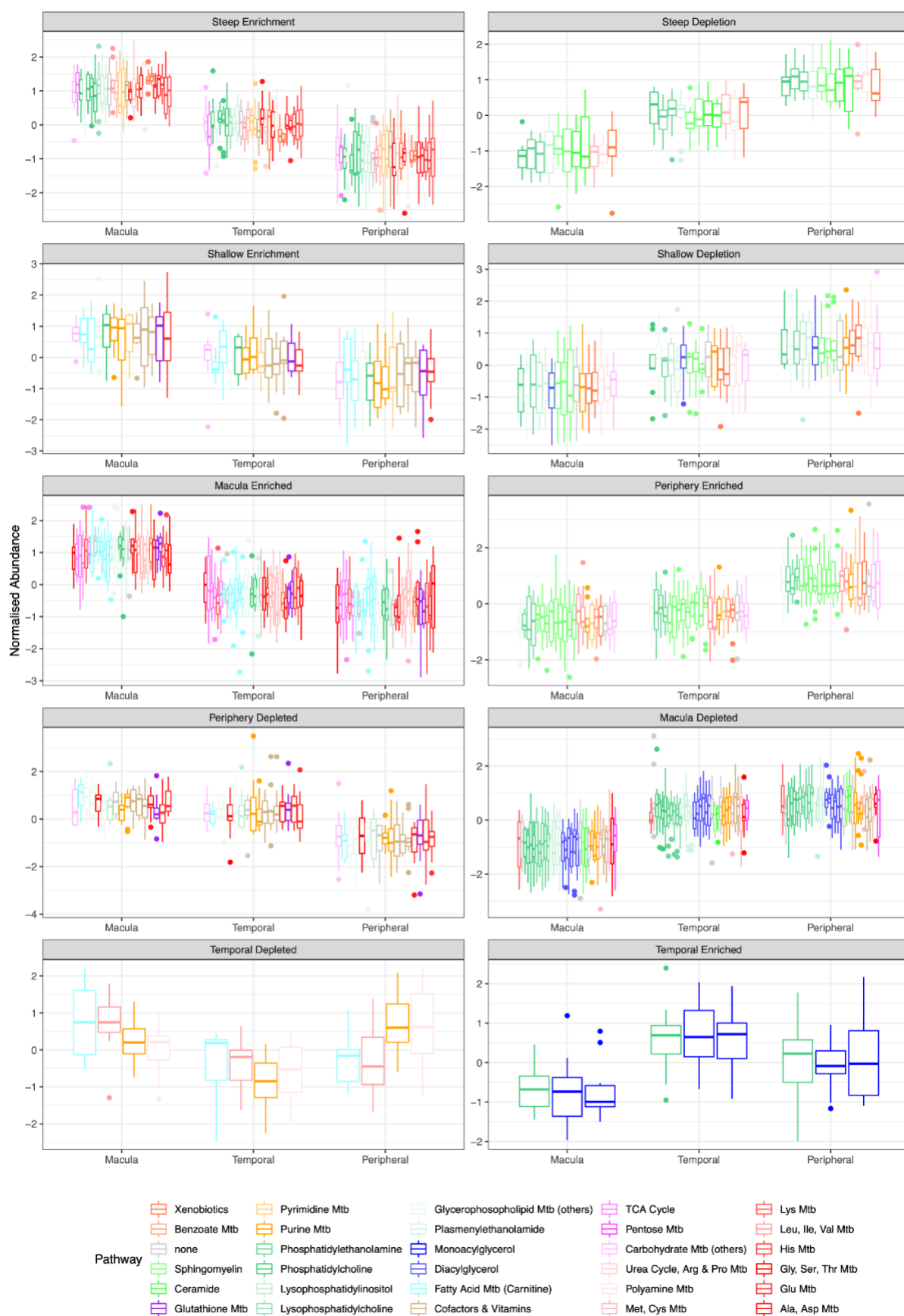


Figure S6: Metabolic distribution between clusters. In this image, every boxplot represents the covariate-corrected distribution of one metabolite whose abundance was significantly different for any of the three contrasts. Metabolites are divided into clusters

according to their combination of significant results and respective log-fold change direction, further defined in Table 1 (summary statistics provided in **Table S5**). Boxplots have been coloured by the biological pathway of each metabolite. Increasing here is used to represent an increment of metabolic abundance from the macula to the periphery while decreasing represents the opposite.

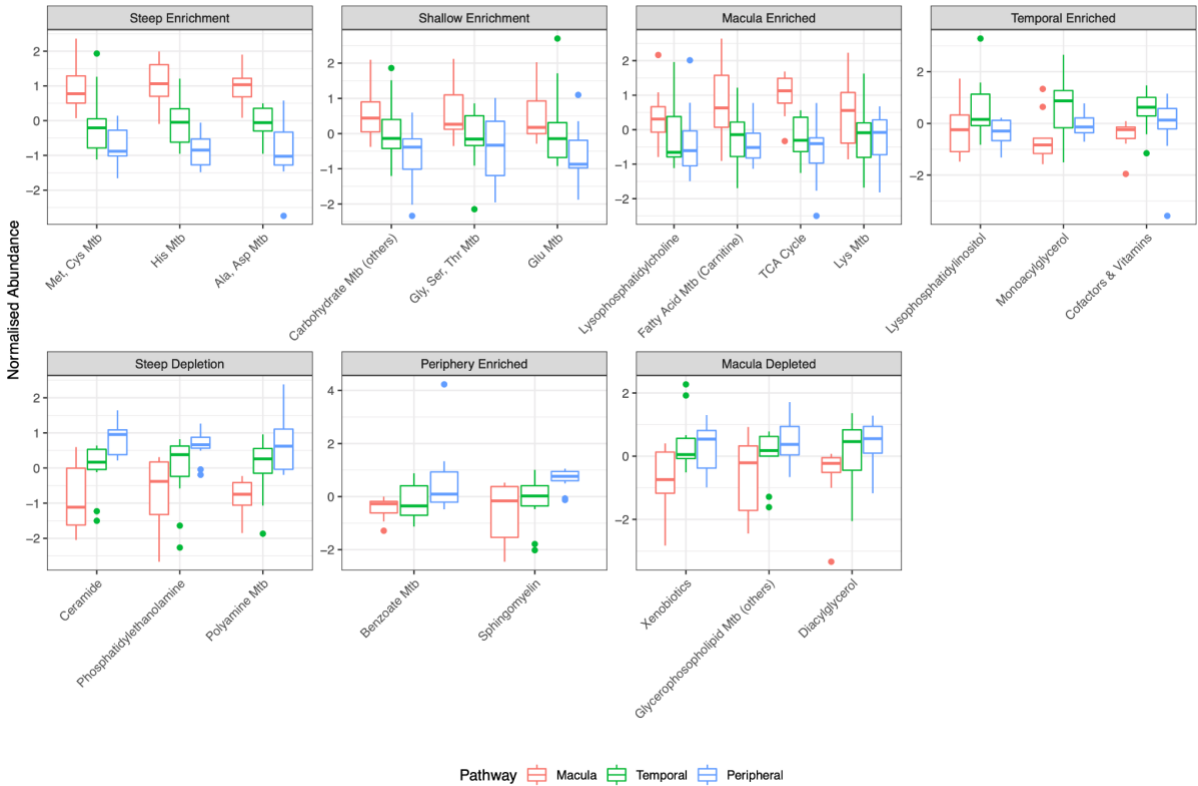


Figure S7: Pathway first PC distribution between retinal areas and divided by cluster. In this image, every boxplot represents the covariate-corrected distribution of the first PC for one pathway where this was significantly different for any of the three contrasts. Pathways are divided according to their combination of significant results and respective log-fold change direction (**Table S5**).

Retinal Metabolic Enrichment

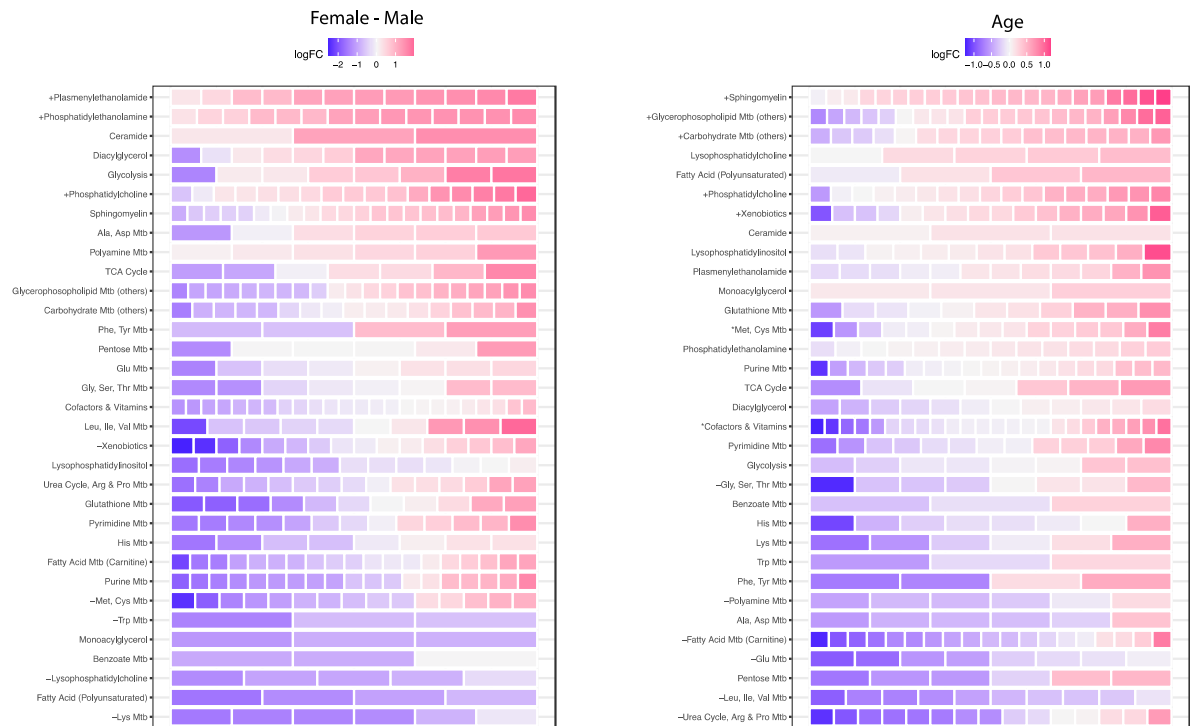
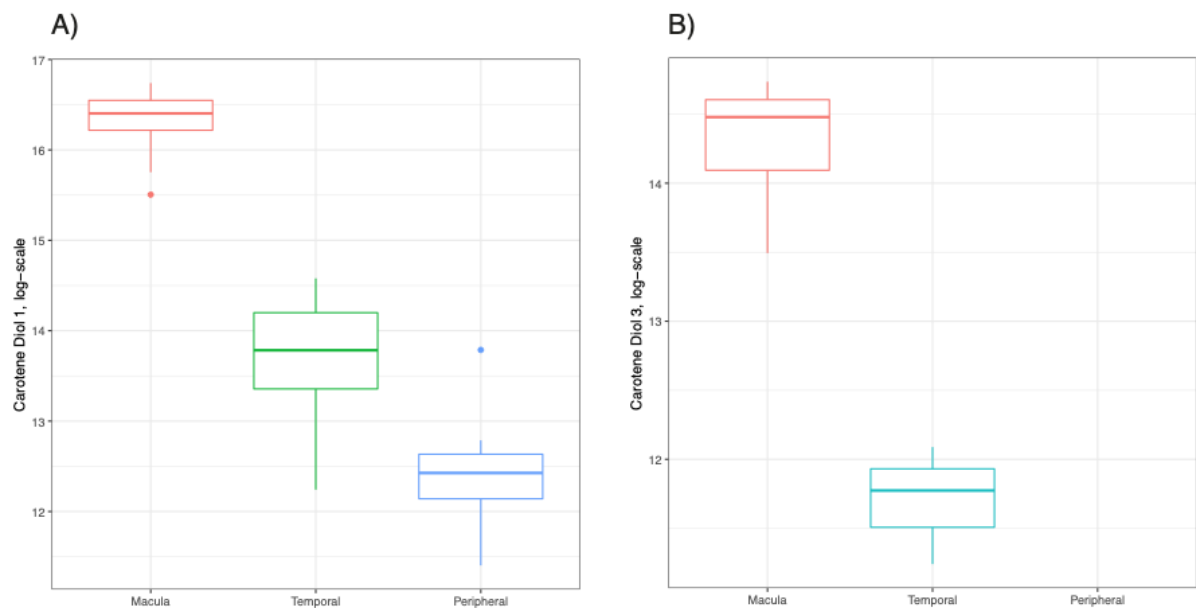
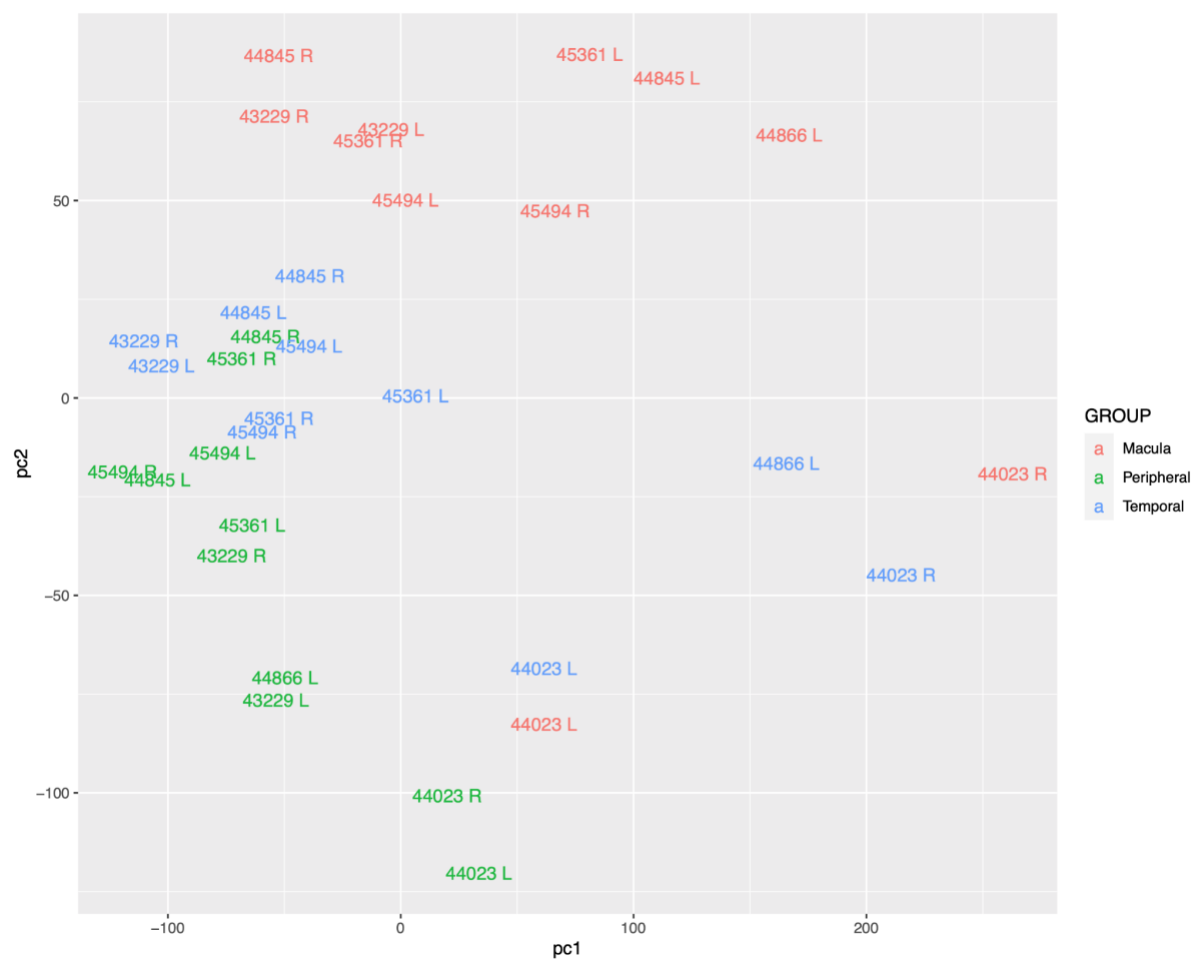


Figure S8: Relative abundance of N metabolites (filled rectangles, or 'bricks') grouped by biochemical family/pathway (y axis) and ordered by greater family-wise abundance in females (left) and older (right). The number of bricks indicates the number of metabolites in the pathway available for this study. The colour of each brick represents the log-fold change a particular comparison (female vs males on the right and older vs younger left). Positive log-fold change (reds) indicates that the metabolite is more abundant while negative values (blues) indicate that the metabolite is depleted. + denotes family is significantly enriched (mainly red); - denotes family is significantly depleted (mainly blue); * denotes family is more differentially abundant in both directions than would be expected by chance ('mixed directionality'/signal mixture effect).





C)

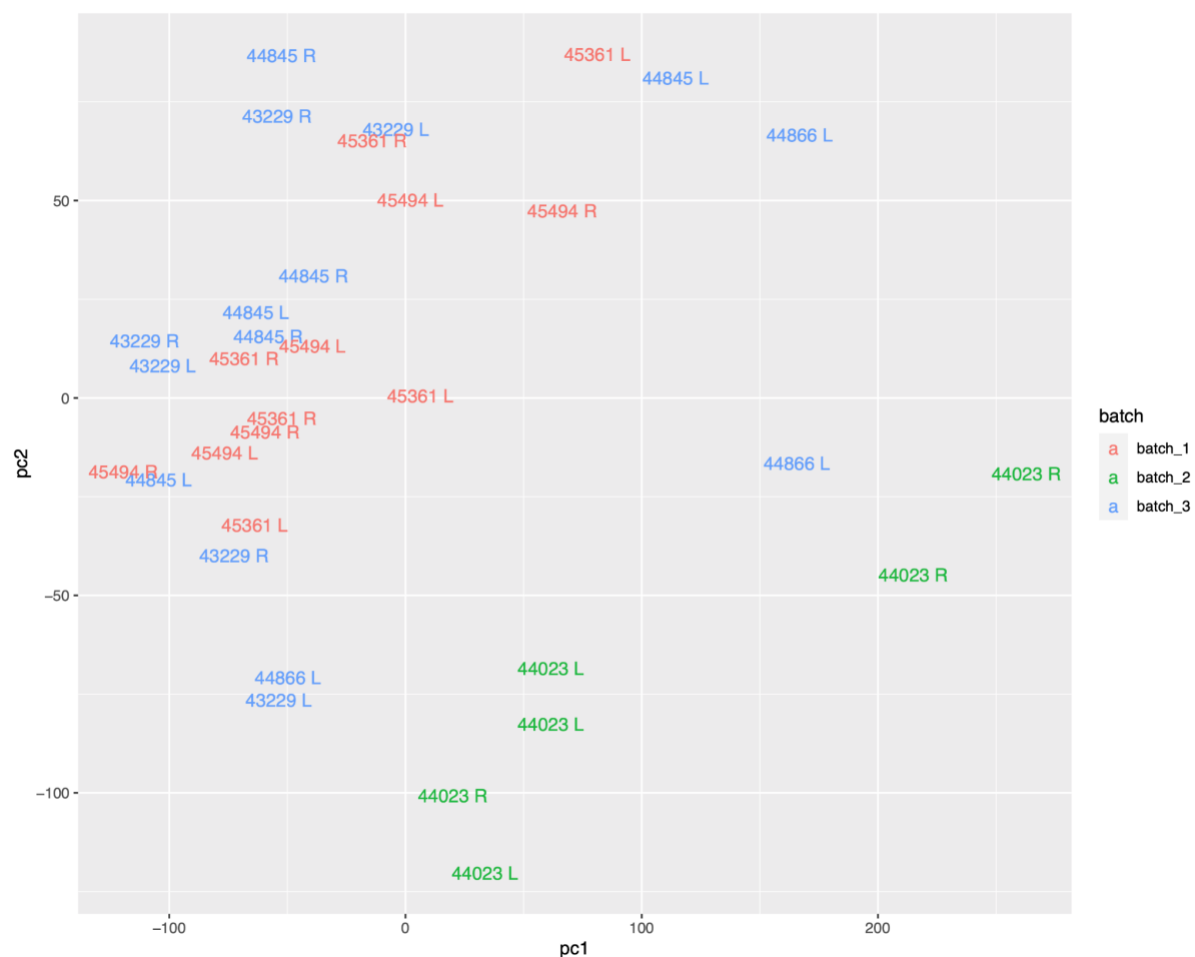


Figure S10: First two principal components plot showing the degrees of separation between biological samples in A) non-normalised data B) quantile normalised data. Animal ID number and eye (R=right eye, L=Left eye) is presented as text in the plot. C) First two principal components plot showing the degrees of separation between batches in the quantile normalised data.

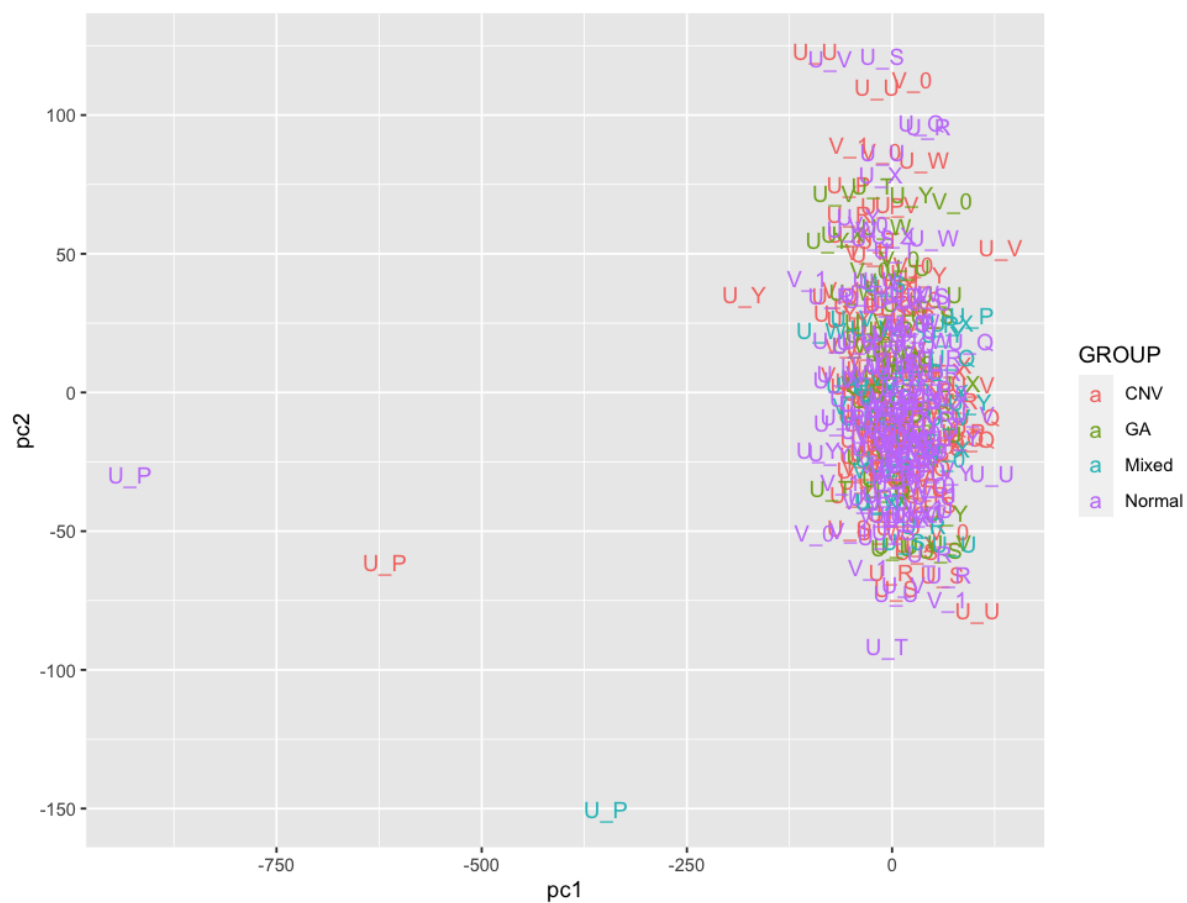


Figure S11: First two principal components plot on AMD metabolic data identifying 3 outliers. This plot every sample was colored according to the disease status and the string represented the sample preparation sub-batch.

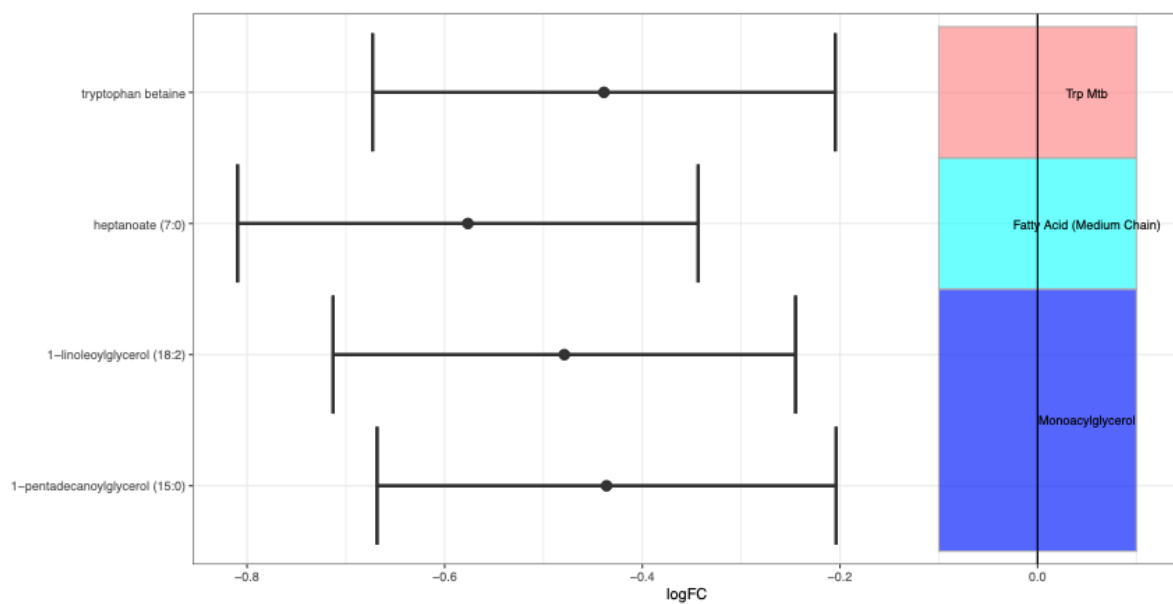


Figure S12: Log-fold changes and 95% confidence interval of metabolites with significant ($FDR < 0.05$) differential abundance between CNV-AMD and healthy controls. Negative log-fold changes values in this figure indicate that the metabolite abundance was lower in CNV-AMD samples compared to controls. Metabolites have been divided and coloured by their respective biological pathways.

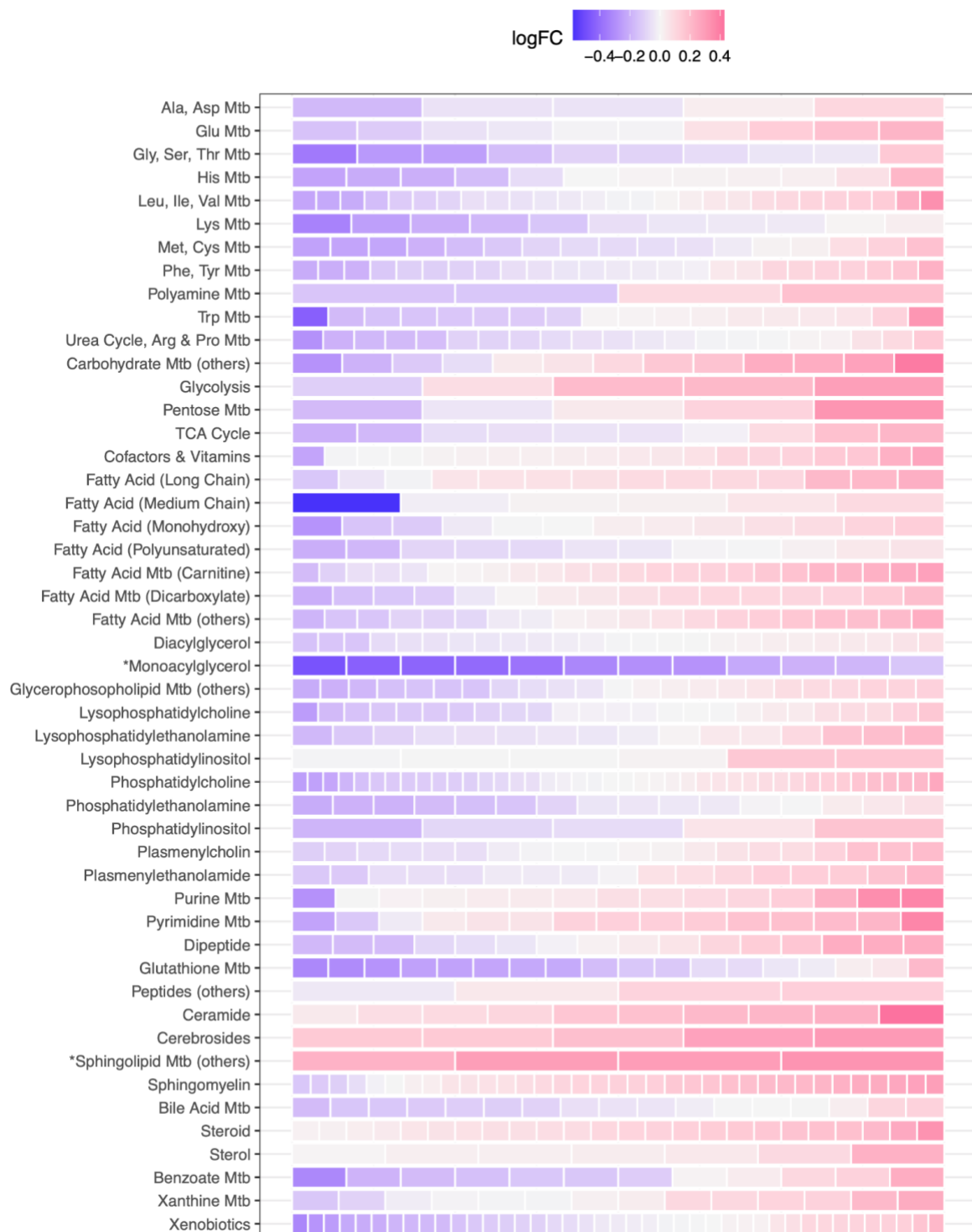


Figure S13: Pathways enrichment results. In this figure, each row represents a biological pathway. Metabolites are presented as "bricks" in each row. The colour of each metabolite represents the log-fold changes of that metabolite between CNV-AMD and healthy controls. High log-fold change indicates that the metabolite is more abundant in the CNV-AMD while negative values indicated that the metabolite is depleted. The symbol '**' next to the pathway name indicates that that pathway was significantly enriched with either abundant or depleted metabolites in CNV-AMD compared to healthy controls.

Supplementary Table Legends

Table S1: Table of differential abundance results for the primate analysis. This table contains most of the results for each metabolite presented in this publication. The results are divided into different categories: Differential Abundance By Metabolite, Pathway Enrichment Analysis, Missingness and Imputation, Additional Metabolites Info, and Pathway First PC Analysis results. (Note that a legend for each column presented in **Table S1** is provided as well as an enrichment analysis only version).

Table S2: Table of differential abundance results for age and sex effects.

Table S3: Table of differential abundance results for the human serum AMD analysis. The results are divided into different categories: Differential Abundance By Metabolite, Pathway Enrichment Analysis, Missingness and Imputation, Additional Metabolites Info. (Note that a legend for each column presented in **Table S3** is provided).

Table S4: Table presenting the list of discarded metabolites. For each metabolite, a reason is presented for discarding it.

Table S5: Retained metabolites used in this study. This list contains all metabolites as well their classification in metabolic pathways as super-pathways.

Table S6: Table presenting metabolic clustering decision rules. For each cluster, the combination of significant effects for each contrast and their direction is presented.