Original Title: Plasma metabolomics supports non-fasted sampling for early detection of impaired glucose tolerance in the Nile rat model of type 2 diabetes

New Title: Plasma metabolomics supports non-fasted sampling for metabolic profiling across a spectrum of glucose tolerance in the Nile rat model for type 2 diabetes

## Authors

Benton J. Anderson (1), Anne M. Curtis (2, 3), Annie Jen (4), James A. Thomson (2, 3, 5), Dennis O. Clegg (2, 3), Peng Jiang (6, 7, 8), Joshua J. Coon (1, 4, 5), Katherine A. Overmyer\* (4, 5), Huishi Toh\* (3)

\* Co-corresponding authors

## Affiliations

(1) Department of Chemistry, University of Wisconsin-Madison, Madison, WI 53706

(2) Department of Molecular, Cellular, and Developmental Biology, University of California, Santa Barbara, CA 93106

(3) Neuroscience Research Institute, University of California, Santa Barbara, CA 93106

(4) Department of Biomolecular Chemistry, University of Wisconsin-Madison, Madison, WI 53706

(5) Morgridge Institute for Research, Madison, WI 53706

(6) Department of Biological, Geological and Environmental Sciences, Cleveland State University, Cleveland, OH 44115

(7) Center for Gene Regulation in Health and Disease, Cleveland State University, Cleveland, OH 44115

(8) Center for RNA Science and Therapeutics, School of Medicine, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106

# Abstract

Type 2 diabetes is a challenge in modern healthcare, and animal models are necessary to identify underlying mechanisms, where we can achieve much better environmental control than what is practical in human studies. The Nile rat (*Arvicanthis niloticus*) develops diet-induced diabetes rapidly on a conventional rodent chow diet without genetic or chemical manipulation. Unlike common laboratory models, the outbred Nile rat model is diurnal and has a wide range of overt diabetes onset and diabetes progression patterns in both sexes, better mimicking the heterogeneous diabetic phenotype in humans. Multiple human studies indicate that compared to fasting glucose, post-prandial blood glucose is more sensitive to the initial stages of diabetes, suggesting that we should capture the non-fasted state to study underlying mechanisms of diabetes progression. There is a long-held assumption that *ad libitum* feeding in rodent models leads to increased variance thus masking diabetes-related metabolic changes in the plasma. In this study, we compared the repeatability within triplicate non-fasted or fasted plasma samples and assessed metabolic changes relevant to glucose tolerance in fasted and non-fasted plasma of 8-10 week male Nile rats over a spectrum of glucose tolerance measured at 12 weeks old. We used liquid chromatography-mass spectrometry lipidomics and polar metabolomics to measure relative metabolite abundances in the plasma samples. Compared to fasted metabolites, non-fasted plasma metabolites are not only more strongly associated with glucose tolerance based on unsupervised clustering and elastic net regression model (R2=0.67 non-fasted, R2=0.52 fasted) but also have a lower replicate variance (22.2% RSD non-fasted, 24.9% RSD fasted). Between the two sampling groups, we detected 66 non-fasted metabolites and 32 fasted metabolites that were associated with glucose tolerance using a combined approach with multivariable elastic net and individual metabolite linear models. Further, to test if metabolite replicate variance is affected by age and sex, we measured non-fasted replicate variance in a cohort of mature 30-week male and female Nile rats (18.4% RSD males and 16.9% RSD females). Our results support using non-fasted plasma metabolomics to study glucose tolerance in Nile rats across the progression of diabetes.

# 

# Introduction

Diabetes is an urgent global health challenge with an accelerating incidence rate in recent decades. Currently, 537 million adults are living with diabetes and 541 million adults have impaired glucose tolerance with a high risk of developing type 2 diabetes[1](https://www.zotero.org/google-docs/?tULKVh). To better understand metabolic changes associated with increasing glucose tolerance, we need suitable animal models and experimental methods that can capture these changes.

The Nile rat (*Arvicanthis niloticus*) is a model of type 2 diabetes with key benefits over other rodent models. First, diabetes is rapidly induced in both sexes by conventional laboratory rodent chow that is hypercaloric for the Nile rat compared to its native fiber-rich diet[2,3](https://www.zotero.org/google-docs/?uVcuDI). On conventional rodent chow, the onset of diabetes can range from a month to a year of age, and by 6 months of age, most of the Nile rats would have developed diabetes[4](https://www.zotero.org/google-docs/?CUjYDG). In contrast, common laboratory mice and rats are relatively resistant to diet-alone induced diabetes, where additional chemical or genetic manipulations are used to promote diabetes[5](https://www.zotero.org/google-docs/?zOJGNr). Second, diabetic Nile rats can develop long-term diabetic complications mimicking clinical features of diabetic patients[6–8](https://www.zotero.org/google-docs/?R1TLCB), including diabetic retinopathy[9,10](https://www.zotero.org/google-docs/?J76eWB). Third, the Nile rat model is outbred and displays a wide range of diabetic phenotypes[11](https://www.zotero.org/google-docs/?bKs0Gc), reflecting its underlying genetic diversity. Fourth, the Nile rat, like humans, are active during the day[12](https://www.zotero.org/google-docs/?qgLlo8), unlike common nocturnal rodent models. Additionally, the Nile rat has a reference genome for mechanistic studies[13](https://www.zotero.org/google-docs/?UQiVNX). Overall, the Nile rat is highly suited to study the underlying mechanisms of glucose tolerance in diet-induced diabetes.

When considering experimental methods for studying diabetes, a majority of studies looking for metabolic changes will use blood that has been sampled under fasted state to avoid excess variability from unrestricted eating behavior. However, for the early progression of diabetes, it is known that postprandial hyperglycemia precedes fasted hyperglycemia, and thus is a more sensitive measurement for early diabetes demonstrated in human studies[14–16](https://www.zotero.org/google-docs/?xVpddx) as well as in the Nile rat model[11,17–19](https://www.zotero.org/google-docs/?wko6rl) In addition, there is some evidence that postprandial state might have reduced variability in blood metabolites[20](https://www.zotero.org/google-docs/?eBM3Jz). For rodent models, non-fasted state likely represents a postprandial state given the high frequency of food intake. Yet, to date, no study has validated the use of non-fasted sampling for metabolomics studies in rodent models. Specifically, reproducibility and replicate variability between fasted and non-fasted states have not been sufficiently analyzed, thus this study compares metabolite variance between non-fasted and fasted blood sampling for studying progressive glucose intolerance in Nile rats.

To investigate the metabolic differences between the fasted state and non-fasted state, we performed metabolomics using liquid chromatography coupled to mass spectrometry (LC-MS) to measure a broad range of plasma biomolecules[21](https://www.zotero.org/google-docs/?FaVM6a). Further, LC-MS has been used to analyze variance of plasma sampling across metabolites[20,22–25](https://www.zotero.org/google-docs/?qJ9GXg) and to detect plasma biomarkers relevant to diabetes[26](https://www.zotero.org/google-docs/?7Zn9KD),[27](https://www.zotero.org/google-docs/?FXijeN). To assess replicate variance in fasted versus non-fasted and to capture markers of diabetes, we measure non-fasted and fasted plasma samples in triplicate. This cohort consisted of male Nile rats aged 8-10 weeks and varied in levels of glucose tolerance, from non-diabetic to overtly diabetic, at 12 weeks old. We found that metabolites in non-fasted plasma sampling enabled better predictive power of impaired glucose tolerance. Counter to the accepted wisdom that *ad libitum* feeding leads to increased variability, we found that metabolites in non-fasted sampling had lower median replicate variance. To validate improved reproducibility in non-fasted sampling, we used a different age group of Nile rats and included both sexes and assessed metabolite variance in triplicate plasma samples. In this validation cohort, replicate variance of metabolites was similar or lower than that found in the primary cohort. Our data supports the use of non-fasted state for plasma sampling in metabolic studies of Nile rats due to better association to glucose tolerance levels and lower replicate variance independent of age or sex.

In this study, we performed plasma metabolomics to compare the non-fasted versus fasted state in Nile rats. We employed a study design that measured both states within each Nile rat, and investigated how metabolite abundances and variance are affected by glucose tolerance and plasma sampling method. We concluded that plasma metabolomics using non-fasted sampling is valuable for studies of diabetes across a spectrum of glucose tolerance in Nile rats.

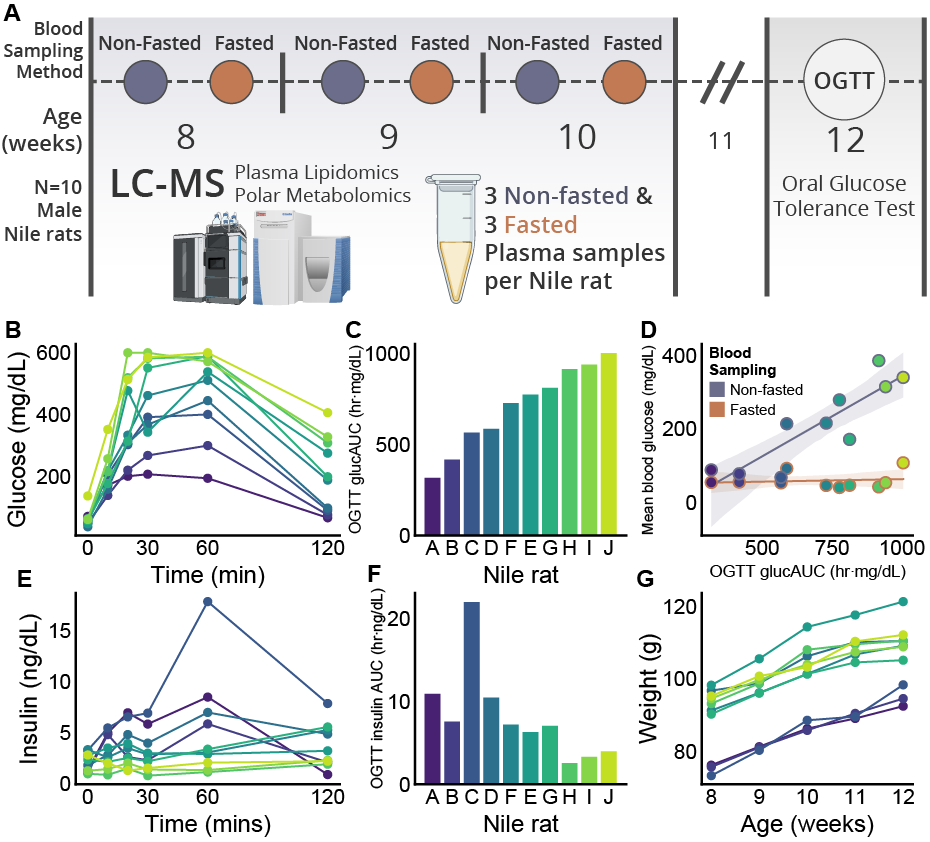
# Results

## Metabolomic profiling of non-fasted and fasted plasma associated with glucose tolerance

Nile rats have been well-described to develop glucose intolerance when consuming conventional rodent chow 5008, in part due to the glycemic load of 5008[11,17–19,28](https://www.zotero.org/google-docs/?kfgb7A). Additionally, males progress toward glucose intolerance more rapidly than females on the same dietary challenge[28](https://www.zotero.org/google-docs/?jlpXFm). Because the Nile rat model is genetically diverse, they display a spectrum of glucose intolerance. To evaluate plasma metabolites that show trends over a spectrum of glucose tolerance, we selected juvenile males fed rodent chow.

Specifically, to evaluate replicate variance between non-fasted and fasted states for plasma sampling, we collected 60 samples using 10 male Nile rats with two sampling states in triplicate. These samples were taken at 8, 9 and 10 weeks of age, and later, an oral glucose tolerance test (OGTT) was performed at 12 weeks (**Figure 1A**). Time of day of sampling for the non-fasted state was found to have no significant effect on the observed blood glucose value (**Supplemental Figure 1A-B**). Sampling the plasma at least two weeks prior to OGTT enabled the animals to recover from weekly blood collections before OGTT. **Figure 1B** shows the glucose excursion across two hours during the OGTT. Nile rats labeled A to J are ordered based on area under the curve of glucose levels during OGTT (OGTT glucAUC) (**Figure 1C**) and show a range of glucose tolerance evenly distributed across these 10 Nile rats. Within the range of glucose tolerances captured in our study cohort, random blood glucose (RBG) exhibited a positive trend with subsequent OGTT glucAUC whereas there was no association to fasted blood glucose (FBG) (**Figure 1D**).

Additionally, we measured blood insulin concentration during the OGTT. In humans, patterns of insulin concentration during OGTT can predict incident type 2 diabetes[29](https://www.zotero.org/google-docs/?ZgNRDj). Here, we observed similar patterns where the healthier Nile rats A to D had higher insulin levels at 60 minutes than 120 minutes, compared to Nile rats E to G (**Figure 1E**). Notably, the area under the curve from plasma insulin during OGTT (OGTT insAUC) was exceptionally high in Nile rat C (**Figure 1F**). This hyperinsulinemic response suggests that Nile rat C was at a pre-diabetic or at an early stage of diabetes. Conversely, Nile rats H, I and J were hypoinsulinemic, indicating that these rats were more advanced in the diabetes spectrum. Weekly body weight (**Figure 1G**) and blood glucose (**Supplemental Figure 1C**) were measured from weeks 8 to 12. Though the growth rates were similar, the initial weights taken at 8 weeks segregated the animals into two groups, with A, B and C at lower weights and D through J at higher weights. Based on 8 week RBG, 7 (Nile rats D to J) had non-fasted hyperglycemia (RBG > 200 mg/dL) and two Nile rats had fasted hyperglycemia (FBG > 126 mg/dL in Nile rats D and J).

****

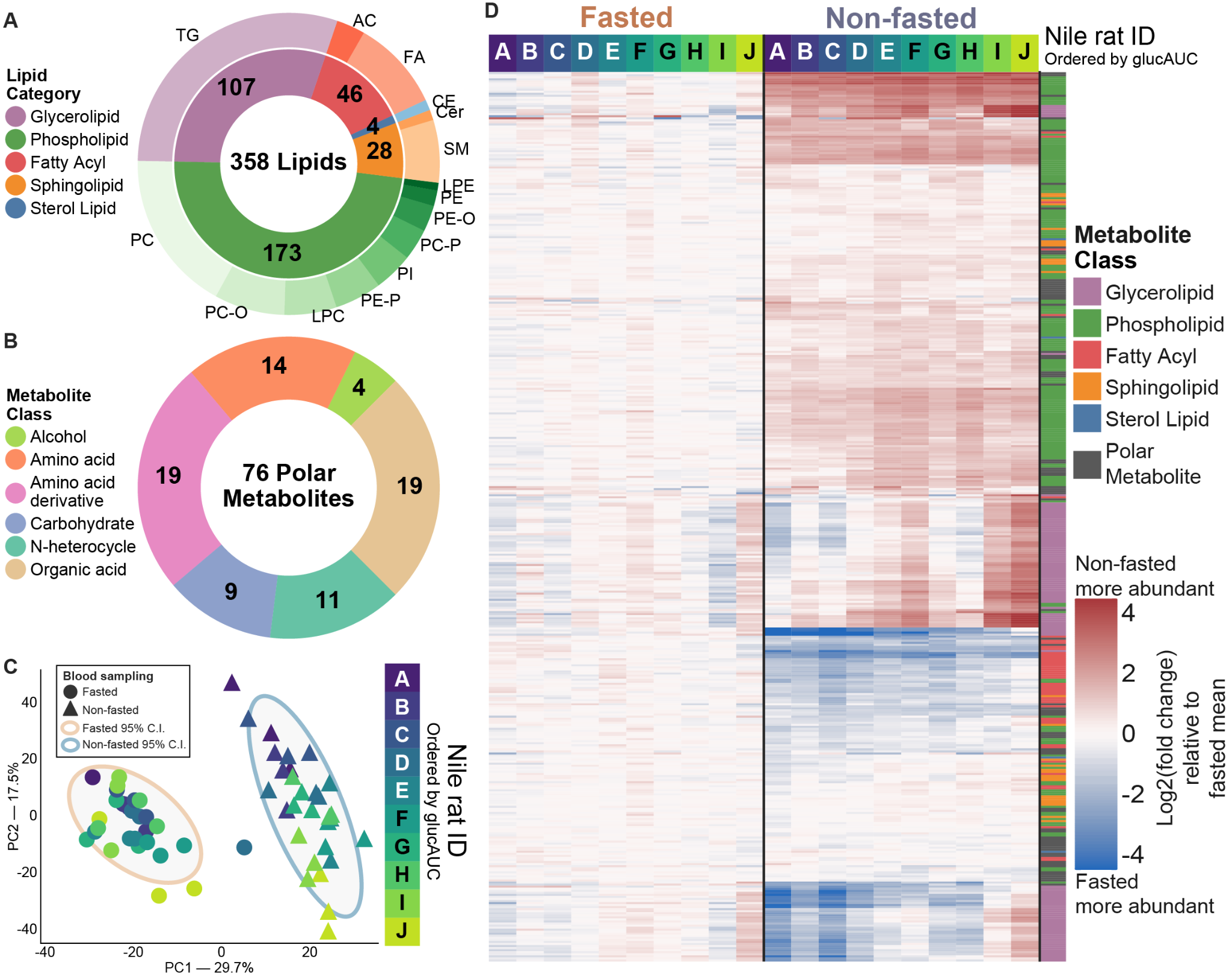
**Figure 1. Metabolomic study design and characterization of glucose tolerance** (**A**) Overview of study design analyzing plasma metabolites and lipids at 8-10 weeks and measuring glucose tolerance by OGTT at 12 weeks. (**B**) OGTT curve with blood glucose levels taken at 0, 10, 20, 30, 60 and 120 minutes after ingesting glucose delivered by gavage. (**C**) Area under the curve calculated from glucose OGTT (OGTT glucAUC). Nile rats ordered by increasing OGTT glucAUC. (**D**) The correlation between OGTT glucAUC at 12 weeks and mean blood glucose levels at 8-10 weeks, measured in fasted versus non-fasted state of the Nile rats. Shaded region depicts the 95% bootstrapped confidence interval. (**E**) OGTT curve with plasma insulin levels taken at 0, 10, 20, 30, 60 and 120 minutes after ingesting glucose. (**F**) Area under the curve calculated from insulin OGTT (insAUC). Nile rats ordered by increasing OGTT glucAUC. (**G**) Growth chart from 8 to 12 weeks based on whole body weight for Nile rats A to J.

## Unsupervised clustering revealed better association of metabolite abundance to glucose tolerance in non-fasted plasma

To characterize the plasma biomolecules in Nile rats under fasted and non-fasted conditions, we performed discovery metabolomics and lipidomics by LC-MS/MS, and calculated relative quantification by integrating chromatographic peak area. We annotated 358 lipids across 5 lipid categories[26,30](https://www.zotero.org/google-docs/?Idlbgn), including glycerolipids, phospholipids, sphingolipids, fatty acyls, and sterol lipids; 556 lipid chromatographic features remain unannotated but were included in some of the downstream analyses (**Figure 2A**). Of the annotated lipids, 200 were identified at species level and 158 were at molecular species level[31](https://www.zotero.org/google-docs/?rTKLgV). Among polar metabolites, we annotated 76 compounds from 6 compound classes, including organic alcohols, amino acids, amino acid derivatives, nitrogen heterocycles, carbohydrates and organic acids. (**Figure 2B**). 419 polar metabolite features remained unannotated.

Next we performed principal component analysis (PCA) using all metabolite features in our 60 plasma samples. PCA reveals two clusters separated on the first principal component by fasted or non-fasted sampling conditions (**Figure 2C**). Within the non-fasted cluster, the samples appear to be ordered by OGTT glucAUC along the second principal component, whereas a similar ordering is absent in the fasted cluster. This suggests that non-fasted metabolomic changes are associated with glucose tolerance.

To further explore high-level trends in plasma metabolites, we constructed a heatmap ordered by Nile rat OGTT glucAUC on the columns, with hierarchical clustering on the rows of annotated metabolites (**Figure 2D**). From here on, we refer to plasma samples collected in the non-fasted or fasted state as ‘non-fasted samples’ or ‘fasted samples’, respectively. Overall, non-fasted samples display greater log2 fold changes relative to mean metabolite abundance in fasted samples. In general, the lipids appear to have a larger dynamic range than the polar metabolites. Glycerolipids show the most apparent trends in association to OGTT glucAUC ranking.

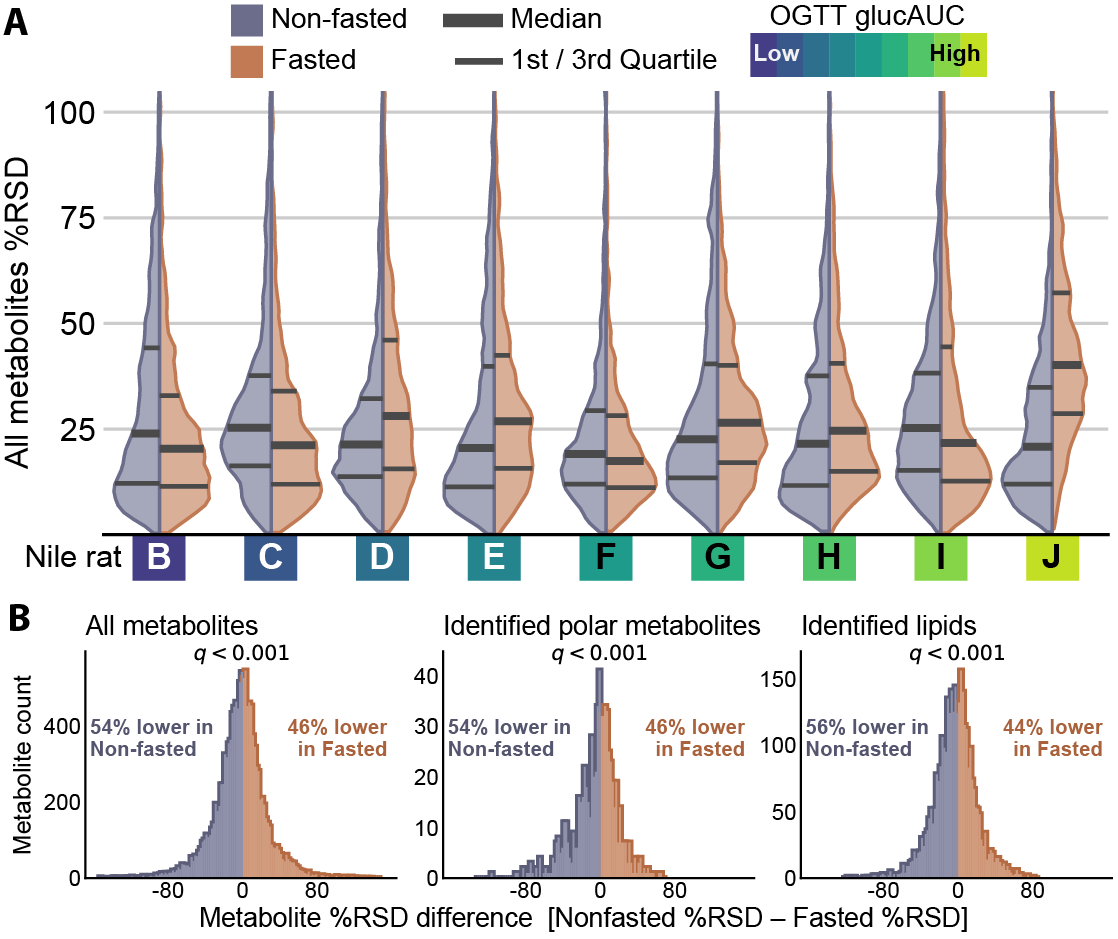


**Figure 2. Plasma lipids and polar metabolites separate non-fasted and fasted sampling methods in unsupervised clustering.**

**(A)** A summary of counts of 358 identified lipids, divided into 5 lipid categories on the inner circle, and lipid class on the outer circle. **(B)** A summary of counts of 76 total identified polar metabolites from 6 classes. **(C)** Principal component analysis of all non-fasted and fasted plasma samples using all features from both HILIC and reversed phase LC-MS methods. Principal components 95% confidence intervals (CI) are displayed around non-fasted and fasted points. Points are colored according to Nile rat ID, with color gradation based on OGTT glucose AUC order. **(D)** Heatmap of all identified lipids and polar metabolites, separated by fasted (left side) and non-fasted (right). Each column is the mean of triplicate Nile rat samples. Rows are ordered based on clustering (see methods) using non-fasted plasma samples. Lipid or polar metabolite indicator are given in colorbar. Metabolite abundances are given as log2 fold change of the difference from mean of fasted samples.

## Non-fasted Nile rat plasma yields lower replicate variance across metabolites

A major concern of non-fasted sampling is the excess variability driven by *ad libitum* feeding and varying degrees of postprandial state. To assess the plasma metabolite variability within replicate samples between fasted and non-fasted sampling states, we calculated percent relative standard deviation[32](https://www.zotero.org/google-docs/?tQIaGj) (%RSD) across an individual’s triplicate 8 to 10 week plasma samples. The distribution %RSDs for all metabolites is shown for each Nile rat in **Figure 3A**, grouped by sampling method. We excluded Nile rat A which had 2 out of 3 replicate fasted samples. Of the remaining 9 animals, 5 had lower median metabolite %RSD in non-fasted replicates. The median %RSD across all triplicate metabolite measurements is smaller in non-fasted samples (22.2%) compared to fasted samples (24.9%). At an individual metabolite level, we calculated the percentage point difference between non-fasted and fasted %RSD for each metabolite per Nile rat and show the distribution of these paired differences in **Figure 3B**. A larger number of %RSD differences were lower in non-fasted replicates for all metabolites (54%), identified polar metabolites (54%) and identified lipids (56%). All 3 groups show significant difference from 0 percentage point difference (*q < 0.001,* see methods). Similar analysis for other groups based on metabolite class, lipid class and lipid category are given in **Supplemental Figure 2**. 8 out of 15 lipid classes (LysoPC, PC, SM, Plasmenyl-PC, Plasmanyl-PC, Plasmenyl-PE, Plasmanyl-PE, and TG) have significantly lower (*q < 0.05*) %RSDs in non-fasted replicates. Among polar metabolite groups, carbohydrates, organic acids and amino acids yield significantly lower %RSDs in non-fasted replicates. Additionally, across all groupings metabolites tested, none showed a significantly lower %RSD in fasted replicates, affirming that non-fasted sampling results in lower replicate variance. These results are supported by a similar analysis using 1H-NMR metabolomics on postprandial versus fasted human plasma sampling[20](https://www.zotero.org/google-docs/?HFE6dI).

****

**Figure 3. Non-fasted plasma metabolites show similar variance within Nile rats and lower variance within each metabolite compared to fasted sampling** (**A**) Comparison of %RSD distributions for non-fasted and fasted within each Nile rat across triplicate plasma sampling. 5 out of 9 Nile rats show lower median %RSD in non-fasted sampling. Overall median %RSD in non-fasted is 22.2%, compared to overall median %RSD of 24.9% in fasted. Nile rats are ordered based on OGTT glucAUC, with more glucose intolerant to the right. (**B**) Calculating %RSD in each sampling method and then subtracting non-fasted %RSD from fasted %RSD for each metabolite in all Nile rats yields percentage point differences. Distribution of percentage point differences are shown, where the blue portion on left are metabolites whose %RSD difference is less than 0%, i.e. it has a lower %RSD in non-fasted. Across all metabolites, 54% of matched measurements have lower non-fasted %RSD. Across identified polar metabolites, 54% are lower in non-fasted and in identified lipids, 56% are lower in non-fasted. Significance testing using Wilcoxon signed rank test reveals that the median for each distribution significantly differs from 0% (*q<0.001* for all three distributions)

## 

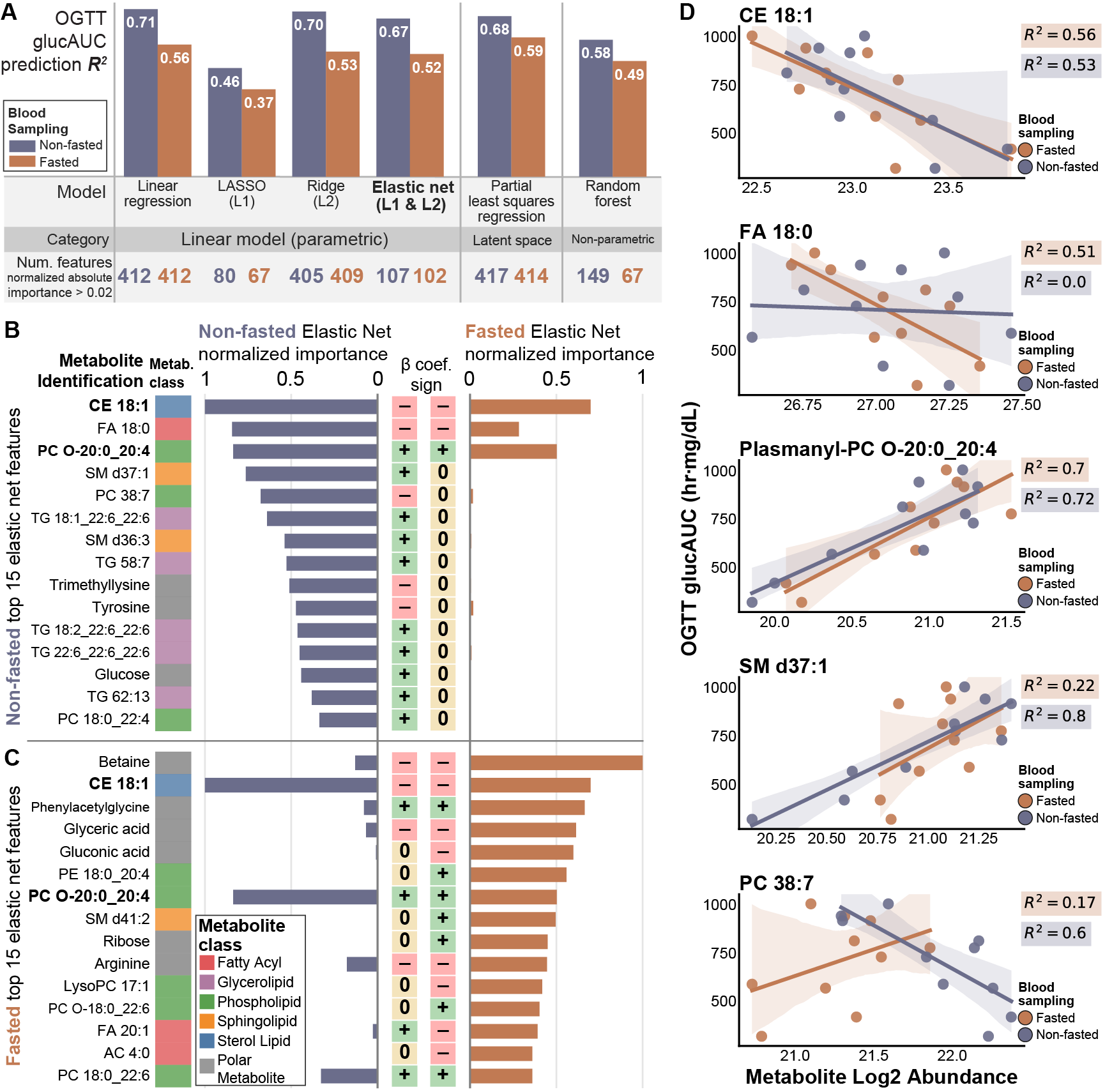
## 

## 

## Non-fasted is superior to fasted Nile rat plasma sampling for predicting OGTT glucAUC in young males

Earlier, we suggested that non-fasted plasma sampling shows stronger associations to OGTT glucAUC compared to fasted sampling based on unsupervised modeling with PCA (Figure 2C). To test this hypothesis, we trained regression models to learn potential metabolite associations to glucose tolerance (**Figure 4A**). Linear regression, least absolute shrinkage and selection operator (LASSO), ridge, elastic net, partial least squares regression (PLSr) and random forest machine learning models were trained to predict 12 week OGTT glucAUC using all annotated lipids and polar plasma metabolites sampled at age 8-10 weeks. We trained competing models using non-fasted versus fasted plasma samples. Model performance was assessed using the median coefficient of determination (R2). Overall, the models trained on non-fasted data yielded a higher median R2 over the same model trained on fasted data. While linear regression was the most performant (R2=0.71 non-fasted, R2=0.56 fasted), biological interpretation of its learned parameters is complicated by the large number of metabolite features retained in the model. The number of features can be minimized by methods such as regularization in linear modeling using LASSO, ridge or elastic net[33](https://www.zotero.org/google-docs/?mnvrz3), bootstrapping in random forests[34](https://www.zotero.org/google-docs/?FU1JZN), or transformation into lower dimensional latent spaces in PLSr[35](https://www.zotero.org/google-docs/?ru4AXz). Of these five model types, elastic net achieved both high performance (R2=0.67 non-fasted, R2=0.52 fasted) and significant coefficient shrinkage (107 and 102 features with normalized absolute importance > 0.02 in non-fasted and fasted, respectively). Compared to other methods, LASSO (R2=0.46 non-fasted, R2=0.37 fasted) and random forest (R2=0.58 non-fasted, R2=0.49 fasted) had lower performance. Ridge and PLSr achieved slightly higher R2 than elastic net, but failed to shrink the number of important metabolite features compared to elastic net. Therefore, elastic net was selected as the optimal model.

The top 15 most important metabolites for predicting OGTT glucAUC in non-fasted and fasted elastic net models are shown in **Figure 4B-C**. There is sparse overlap between the top 15 metabolites in the non-fasted and fasted sampling, except for CE 18:1 and Plasmanyl-PC O-20:0\_20:4. Crucially, a high importance in a multivariate model like elastic net does not ensure that the metabolite predicts OGTT glucAUC well in a univariate model. To demonstrate the performance of univariate prediction, **Figure 4D** gives the linear regression results of predicting OGTT glucAUC from the top five elastic net non-fasted features. CE 18:1 and plasmanyl-PC O-20:0\_20:4 are both in the top 15 fasted and non-fasted elastic net metabolites and achieve approximately similar R2 in both fasted and fed models. In contrast, despite being the 2nd ranked metabolite in non-fasted, FA 18:0 achieves an R2 of 0.0 in predicting OGTT glucAUC. SM d37:1 achieves R2 of 0.8 in non-fasted and 0.22 in fasted. Superior predictive performance by the SM d37:1 model compared to the full elastic net model is due to no cross validation. Finally, PC 38:7 returns better R2 in non-fasted (0.6) and displays a positive correlation, whereas in fasted, it shows negative to no correlation. In summary, some metabolites are useful in a multivariate model by combining their information with other metabolites to boost OGTT glucAUC prediction performance.



**Figure 4 – Linear regression modeling trained on non-fasted plasma samples achieves superior performance in predicting OGTT glucose AUC.** (**A**) Median R2 was compared for 6 different machine learning model architectures trained on non-fasted and fasted plasma metabolite abundances. Across all models, non-fasted data provided higher median R2 values. Linear regression returned the highest R2, but regularized linear models (LASSO, ridge and elastic net) and other models (partial least squares regression and random forest) were trained to perform feature shrinkage or reduction of feature space dimensionality. Elastic net was most performant for biological interpretation due to its nearly equivalent R2 to linear regression and significant coefficient shrinkage. The 6 models were categorized based on the underlying mechanism of prediction, divided between parametric, latent space and non-parametric methods. (**B**) The importance values of the top 15 metabolites in non-fasted and (**C**) top 15 metabolites in fasted elastic net modeling presented along with importance of molecule in the other sampling method. CE 18:1 and PC O-20:0\_20:4 are bolded due to presence in both top 15 lists. (**D**) The top 5 most important metabolites from non-fasted elastic net modeling were individually regressed to OGTT glucAUC. Dots represent the mean value of each Nile rat’s triplicate metabolite abundance; shaded regions are the 95% bootstrapped confidence interval. Abbreviations: LASSO: least absolute shrinkage and selection operator;

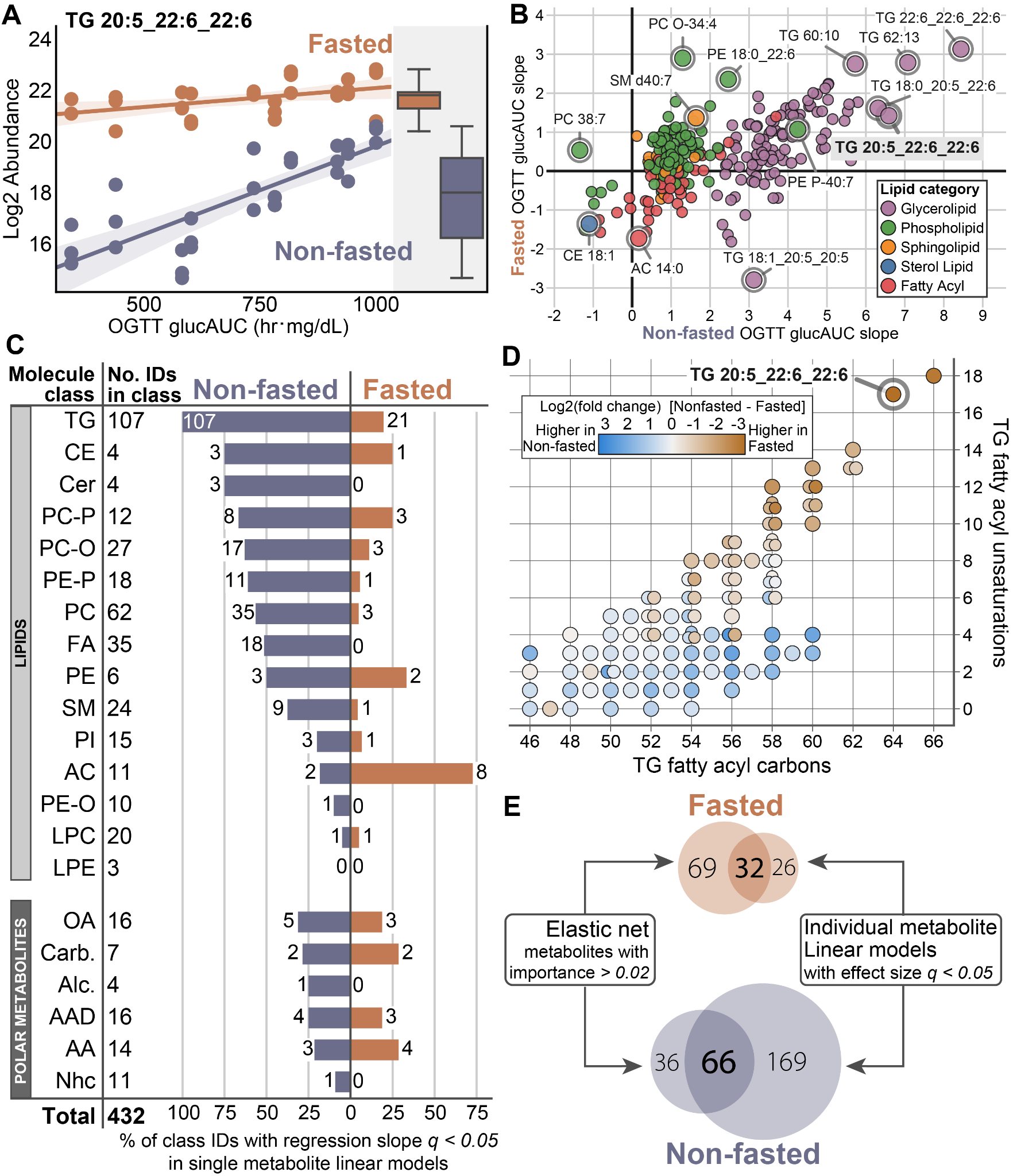
## Non-fasted plasma samples had more metabolites with strong associations to glucose tolerance

Previously, we discovered metabolites that best predicted glucose tolerance in a multivariate model setting (**Figure 4**). Next, we determined which metabolites had individual associations to OGTT glucAUC. Using linear models at the individual metabolite level (see **Methods**) we calculated the effect size of OGTT glucAUC and the interaction between sampling and OGTT glucAUC. An example of the analysis is highlighted in **Figure 5A**, where the abundance of triacylglycerol (TG) 20:5\_22:6\_22:6 significantly increases with OGTT glucAUC in non-fasted (*q<0.0001*), but is not significant in fasted sampling (*q=0.058*). The OGTT glucAUC effect size is greater in non-fasted than fasted sampling (*q=1.8\*10-22*). The mean abundance of TG 20:5\_22:6\_22:6 is significantly greater in fasted samples compared to non-fasted samples (*q<0.0001*) (see also **Supplemental Figure 3A-B**). Results for all metabolites are given in **Supplemental Table 5**.

To explore the OGTT glucAUC effect size of non-fasted and fasted for all metabolites, we plotted regression slopes for all annotated lipids (**Figure 5B**) and polar metabolites (**Supplemental Figure 3C**). The abundance of all TGs were positively associated with OGTT glucAUC in non-fasted sampling, while TGs in fasted sampling had both positive and negative associations to OGTT glucAUC. For all TGs, non-fasted sampling had a steeper regression slope than fasted sampling, indicating stronger associations to glucose tolerance. The TGs most positively associated were TG 22:6\_22:6\_22:6 (66:18), TG 62:13, TG 18:0\_20:5\_22:6 (60:11), TG 60:10, and TG 20:5\_22:6\_22:6 (66:17), all of which contain polyunsaturated fatty acyls (PUFA) such as docosahexaenoic acid (DHA, 22:6n-3). Among lipids that were positively associated with OGTT glucAUC in both sampling methods were PC O-34:4, PE 18:0\_22:6, PE 18:0\_20:4, SM d40:7 and PE P-40:7. In contrast, CE 18:1 was negatively associated with OGTT glucAUC in both fasted and non-fasted sampling conditions. Overall, there were more significant metabolite associations to OGTT glucAUC in non-fasted sampling compared to fasted sampling (**Figure 5C**). Across metabolite classes, only acylcarnitines (AC) and amino acids (AA) had a greater number of metabolites that were significant to OGTT glucAUC in fasted sampling.

Given the strong associations between numerous TGs and OGTT glucAUC, we explored TGs further by plotting TGs separated by fatty acyl carbon count and number of unsaturations, with dots colored by log2 fold change between non-fasted and fasted sampling (**Figure 5D**). TGs with higher carbon counts and number of unsaturations tended to be more abundant in fasted samples, whereas saturated, monounsaturated and TGs with 3 to 4 unsaturations tended to be more abundant in non-fasted samples. A similar plot of TGs is presented with dots colored by difference in OGTT glucAUC regression slope between non-fasted and fasted samples (**Supplemental Figure 3D**). A greater difference in fasted and fed slopes indicates a larger interaction effect between sampling and glucose tolerance.

To filter our data to the metabolites that are most associated with glucose tolerance, we integrated both our multivariate model and individual metabolite analyses (**Figure 5E**). After filtering, we found 66 metabolites associated to OGTT glucAUC in non-fasted sampling, versus 32 metabolites in fasted sampling. Next, we compared these 66 non-fasted sampled metabolites in Nile rats to a list of metabolites from a meta-analysis of incident type 2 diabetes in humans[36](https://www.zotero.org/google-docs/?BSngCg). We found 2 polar metabolites and 5 lipids that are predictive of diabetes in both our Nile rat cohort and in humans: isoleucine, betaine, PC 18:0\_20:3 (38:3), SM d39:1, TG 16:0\_16:0\_16:0 (48:0), TG 16:0\_16:0\_18:0 (50:0) and TG 56:6 (**Supplemental Table 6**). Of the unmatched 59 Nile rat metabolites, 20 were present in the human meta-analysis, but not found to have a significant relative risk for type 2 diabetes, such as glucose (*q=0.09*), and the remaining 39 metabolites were not listed in the human meta-analysis, including many of the polyunsaturated TGs that featured prominently in **Figure 5B**, such as TG 22:6\_22:6\_22:6. Future studies are needed to determine the relevance of these unmatched metabolites in human diabetes.



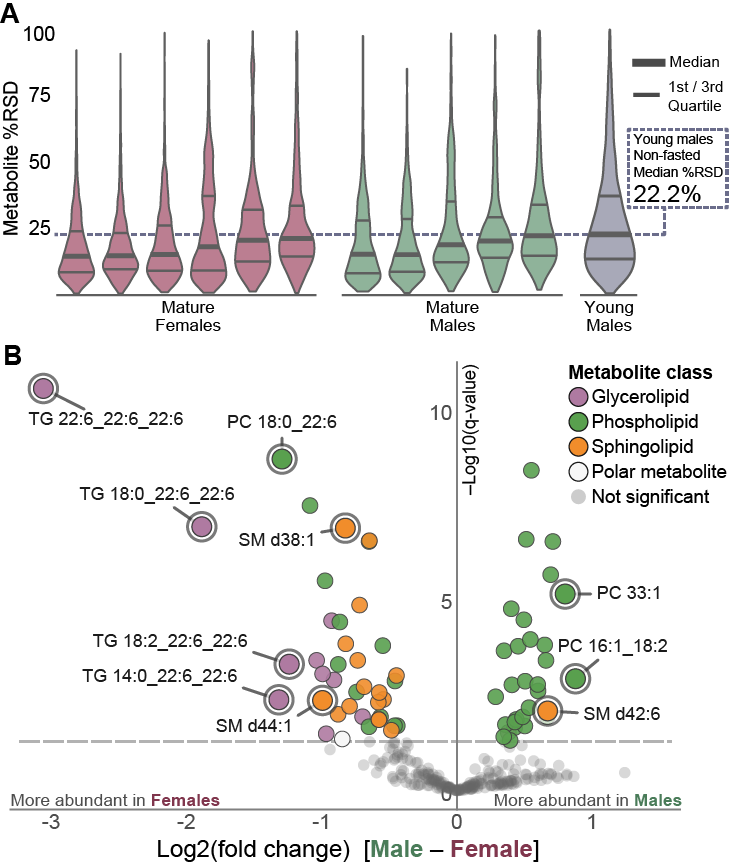
**Figure 5 - Individual metabolite linear modeling is a method for determining plasma metabolite importance in predicting glucose tolerances in Nile rats.** (**A**) Linear regression modeling of log2 abundance versus OGTT glucAUC for TG 20:5\_22:6\_22:6 within each blood sampling method shows higher mean abundance in fasted and a steeper regression slope in non-fasted. (**B**) Plot of regression effect size of OGTT glucAUC versus log2 abundance for all identified lipids in fasted versus non-fasted. Lipids with the steepest slope are dominated by TGs containing PUFAs in non-fasted sampling. (**C**) Linear regression modeling of log2 abundance versus OGTT glucAUC provides regression effect sizes and associated q-values. The number of metabolites in each molecule class with q < 0.05 are shown for both non-fasted and fasted. Across all metabolite classes except AC and AA, non-fasted has more metabolites with significant OGTT glucAUC regression slope than fasted. (**D**) TGs are separated based on fatty acyl carbon count and fatty acyl unsaturation count, with points colored by the mean difference in log2 abundance between non-fasted and fasted. TGs with both high fatty acyl carbons (>57) and high unsaturations (>9) are more abundant in fasted plasma samples. TGs with lower unsaturation counts tend to be more abundant in non-fasted. TGs with identical fatty acyl carbon and unsaturation counts are shown as overlapping dots. (**E**) Filtering metabolites for importance in association with OGTT glucAUC in non-fasted and fasted elastic net models (from Figure 4) and individual metabolite models yields 66 metabolites in non-fasted sampling and 32 metabolites in fasted sampling.

Abbreviations: TG triacylglycerol; CE cholesteryl ester; Cer ceramide; PC phosphatidylcholine; PE phosphatidylethanolamine; -P plasmenyl; -O plasmanyl; FA fatty acyl; SM sphingomyelin; PI phosphatidylinositol; AC acylcarnitine; LPC lysoPC; LPE lysoPE; OA organic acid; Carb. carbohydrate; Alc. alcohol; AAD amino acid derivative; AA amino acid; Nhc nitrogen-heterocycle.

## Low replicate variance in non-fasted sampling is reproducible regardless of age and sex

Similar to humans, Nile rats can develop diet-induced diabetes throughout a large range of ages. To explore if the low replicate variance in non-fasted sampling is affected by age and sex, we performed a similar analysis to study non-fasted replicate plasma sampling in a mature cohort of male and female Nile rats. To select animals developing diabetes, we collected weekly plasma samples and weekly RBG measurements from 20 euglycemic Nile rats starting at 24 weeks old and took samples from the first 11 Nile rats (5 males and 6 females) that developed non-fasted hyperglycemia (**Supplemental Figure 4A**). Subsequent OGTT of these 11 mature Nile rats revealed OGTT glucAUC values (**Supplemental Figure 4B**) similar to values from the previous 10 young male Nile rats. Similar to previously described %RSD measurements (Figure 3), we assessed %RSD on triplicate non-fasted plasma samples across all lipids and polar metabolites for this mature cohort, and plotted their metabolite %RSD distributions (**Figure 6A**). In this mature cohort, every Nile rat displayed a lower median metabolite %RSD compared to the 22.2% median %RSD of non-fasted plasma samples in the previous young male cohort. Aggregating %RSDs within the 3 age and sex groups reveals a statistically significant difference in median %RSD between mature females and mature males (mature males median %RSD=18.4%, mature females median %RSD=16.9%, *p<10-7*). There was also a statistically significant difference in median %RSD in mature males versus young males (*p<10-49*) and mature females versus young males (*p<10-109*). These data support that the low replicate variance in non-fasted plasma sampling is also found in mature Nile rats of both sexes.

Sex differences in type 2 diabetes are well-recognized but poorly studied[37,38](https://www.zotero.org/google-docs/?FvGokQ). Here, we compare plasma metabolite abundances between sexes in our mature cohort. We plot log2 fold change of mean abundance between males and females across all annotated metabolites (**Figure 6B**). With the exception of SM d42:6, many sphingolipids including SM d44:1 and SM d38:1 are more abundant in females compared to males. This is consistent with human data that shows a similar trend wheremost sphingomyelins are more abundant in females[39,40](https://www.zotero.org/google-docs/?Ha060K). Notably, lipids with the greatest fold change contain polyunsaturated 22:6 fatty acyl, for example TG 22:6\_22:6\_22:6 (TG 66:18), TG 18:0\_22:6\_22:6 (TG 62:12) and PC 18:0\_22:6 (PC 40:6). In summary, the metabolic profiles of the male and female Nile rats during progression of diabetes are highly disparate.

****

**Figure 6 – Plasma metabolomics on non-fasted mature male and female Nile rats displaying signs of progressive glucose intolerance shows similar to lower median %RSD compared to young males across triplicate plasma measurements.** (**A**) A cohort of 6 female and 5 male Nile rats underwent triplicate plasma sampling over 3 weeks in the non-fasted state. All 11 Nile rats measured showed lower median plasma metabolite %RSD compared to the median %RSD measurements from all 10 young male Nile rats. (**B**) Volcano plot of log2 fold change differences between male and female 30 week Nile rats for all identified lipids. Polyunsaturated TGs and sphingolipids tend to have higher abundance in females. Abbreviations: %RSD percent relative standard deviation; TG triacylglycerol; PC phosphatidylcholine; SM sphingomyelin.

# Discussion

In this study, we used LC-MS to generate the metabolic profile of Nile rats with a spectrum of glucose tolerance spanning euglycemic to overt diabetes and assessed the differences between fasted and non-fasted sampling. In our primary cohort of juvenile male Nile rats, we showed that metabolite measurements in non-fasted samples were more reproducible with lower replicate variance per animal compared to fasted samples. In these rats, non-fasted metabolite measurements were also better than fasted measurements for predicting 12 week glucose tolerance in young male Nile rats. Next we assessed metabolites in the context of glucose tolerance, where we found 66 metabolites highly associated with OGTT glucAUC using a combined approach with multivariable elastic net and individual metabolite linear models. These include isoleucine, betaine, PC 18:0\_20:3 (38:3), SM d39:1, TG 16:0\_16:0\_16:0 (48:0), TG 16:0\_16:0\_18:0 (50:0) and TG 56:6 that were also found to be significant type 2 diabetes biomarkers in humans. Our findings support non-fasted blood sampling for metabolomics with stronger associations to escalating impaired glucose tolerance and disproved a long-held assumption of higher replicate variance from *ad libitum* feeding; we anticipate these data will crucially inform on future studies in this valuable Nile rat model for diabetes research.

In the past, metabolomics studies in humans have utilized plasma or serum sampled under fasted state[41](https://www.zotero.org/google-docs/?8R2Icd). However, a growing number of studies are promoting the use of non-fasted or postprandial sampling for metabolic studies[42,43](https://www.zotero.org/google-docs/?bisB46). Compared to these studies, our data provides similar conclusions, both in terms of biomarkers[36](https://www.zotero.org/google-docs/?4giitu) and higher reproducibility of non-fasted vs. fasted sampling[20,44](https://www.zotero.org/google-docs/?kIrqfA). Metabolite biomarkers found to agree between our work and a human meta-analysis of incident type 2 diabetes biomarkers are metabolites like isoleucine[45,46](https://www.zotero.org/google-docs/?0iCrBa), betaine[47](https://www.zotero.org/google-docs/?8EVGti), TGs in general[48](https://www.zotero.org/google-docs/?BL1FQf), and sphingomyelins[49](https://www.zotero.org/google-docs/?52FCk4). In addition, we find biomarkers that were not included in the larger human meta-analysis, but have been found in other human diabetes studies, including polyunsaturated lipids SM d40:7, PC 38:7, and PC 40:6[50](https://www.zotero.org/google-docs/?ZXQ0w7). We recognize that fasted sampling is also useful, and contains orthogonal information (*vide infra*). However, the benefits of non-fasted sampling in animal model studies outweighs fasted sampling, namely the lighter workload in managing animals, lower rates of complications due to fasting, and improved reproducibility.

This work is the first plasma metabolomics and lipidomics in the Nile rat species with several key benefits. First, we performed separate lipidomics and polar metabolomics to provide broader coverage of the diverse molecules present in plasma, from hydrophobic triacylglycerols to hydrophilic carbohydrates. Second, we used small amounts of plasma at just 5 microliters for each sample. With such small quantities, this opens up avenues for further analysis of plasma in smaller Nile rats, such as in weanlings, enabling the monitoring of plasma metabolites at even younger ages in populations that show early progression of diabetes. In addition, we used two separate machine learning approaches. The first method uses multivariate regression with regularization to determine a subset of metabolites that work together to predict OGTT glucAUC, while the second approach evaluates each metabolite’s association with OGTT glucAUC. In both approaches we used regression instead of categorical classification; while this approach is uncommon, the methods used here could benefit other diabetes studies that measure continuous variables such as blood glucose, insulin AUC, Homeostatic Model Assessment for Insulin Resistance (HOMA-IR), or hemoglobin A1C (HbA1c). Finally, we performed this study using a study design where each Nile rat underwent replicate sampling under both sampling conditions. This enabled greater statistical power in assessing metabolite replicate variance by using paired statistics between metabolites.

Since our focus was to develop an optimal method for reproducible plasma metabolite measurements in the Nile rat model, this study is limited in its ability to discover diabetes biomarkers due to the small study size and short sampling timeline. In addition, rodent chow is diabetogenic in Nile rats[4,28](https://www.zotero.org/google-docs/?bCYbpT) and this choice of diet skews the population toward higher glucose intolerance. Future studies to investigate the underlying mechanisms of glucose tolerance could be enhanced by incorporating food intake and additional metrics of diabetes; however these data were not captured here due to our priority on the metabolomics data. In summary, the method presented in this manuscript enables larger studies that could use metabolomics to explore diabetes progression, analyze the effects of different diets, and define the genetic and epigenetic contributions to diabetes in Nile rats.

The Nile rat model is highly valuable for mechanistic studies of type 2 diabetes, with a wide range of phenotypes and propensity to develop diet-induced diabetes on conventional rodent chow. Despite a modest cohort size, the metabolic biomarkers detected here in Nile rats show good agreement with human studies of type 2 diabetes. Importantly, we have strong evidence of low replicate variance in non-fasted sampling supporting the use of non-fasted sampling for future work. Lastly, the LC-MS metabolomics described here enables a broad coverage of metabolites and lipids in a very small volume of plasma. In conclusion, our method is highly suited to reveal complex metabolic changes occurring with progression towards overt diabetes.

# Methods

## Animal studies

All animal experiments were approved by the University of California (Protocol Number 893), Santa Barbara, Institutional Animal Care and Use Committee, and conducted in accord with the NIH Guide for the Care and Use of Laboratory Animals. The Nile rats were fed *ad libitum* on a regular rodent diet (Diet 5008; Newco Speciality, Rancho Cucamonga, CA, USA) [[29](https://www.nature.com/articles/s41374-019-0264-3#ref-CR29)], and housed in a 12-hour, 10am to 10pm, light cycle room. 3 cohorts of Nile rats were used: 10 male Nile rats had blood sampled at 8-10 weeks old (primary data set), 6 male and 7 female Nile rats had age range between 38-42 weeks old (12-hour RBG data set) and 5 males and 6 females had blood sampled at ages 26-34 weeks old (validation data set). To perform oral glucose tolerance test (OGTT), Nile rats were fasted for 16 hours from 6pm to 10am, and 2g of dextrose per kilogram body weight was introduced via oral gavage. Fasted plasma samples were collected around 10am to 11am and non-fasted samples were collected around 3 to 4pm, in the middle of the light-on duration. Blood collections were done under fasted and non-fasted conditions in triplicates spaced apart weekly. To minimize the effect of fasting on subsequent non-fasted samples, the Nile rat is allowed to recover for 3 days between the fasted sampling and the next non-fasting sampling. After the last collection, the rats were recovered for 2 weeks before OGTT. All plasma samples were stored at -80˚C. For the oral glucose tolerance test, the glucose area under curve (OGTT glucAUC) and insulin area under curve (OGTT insulin AUC) for each animal were calculated by trapezoidal integration of the corresponding blood glucose (mg/dL) or blood insulin (ng/dL) at measurement time points of 0, 10, 20, 30, 60 and 120 minutes. Animals used in this study were not subjected to any previous procedures and have not been genetically modified. Two animals from the validation data set were excluded because one had an unexpected weight drop and the other developed eye swelling and were euthanized for humane reasons.

## Glucose and insulin measurements

Blood glucose was measured by a Contour Next glucometer using blood from a tail prick. Plasma insulin was measured using the Ultrasensitive Mouse Insulin ELISA Kit [90080] by Crystal Chem according to standard protocol.

## Lipidomics and metabolomics sample preparation

Plasma samples were thawed once prior to the second thawing on ice for subsequent lipidomic and metabolomic sample preparation. Each sample of 5 µL of plasma was extracted with 500 µL 6:2:2 n-butanol:acetonitrile:water[51,52](https://www.zotero.org/google-docs/?tANmgc). Samples were sonicated in a chilled water bath (QSonica) at an amplitude of 30 for 5 minutes at 10 °C using time increments of 20 seconds on/10 seconds off. Samples were then vortexed for 10 s and then centrifuged at 14,000 x g for 2 min at 4 °C to precipitate the protein. 100 µL of extract was dried down in an amber autosampler vial with glass insert by a SpeedVac evaporator (Thermo Scientific). For lipidomics, each extract was resuspended in 50 µL 9:1 methanol:toluene. For metabolomics, each extract was resuspended in 25 µL 1:1 acetonitrile:water then analyzed on the mass spectrometer. For both LC-MS methods, run order of plasma samples was randomized to minimize confounding effects of instrument variance over time.

## HILIC-LC-MS Metabolomics

Sample analysis was performed on a ZIC-pHILIC HPLC column held at 50 °C (100 mm x 2.1 mm x 1.7 μm particle size; Millipore) using a Vanquish Binary Pump (150 μL/min flow rate; Thermo Scientific). Mobile phase A consisted of 10 mM ammonium acetate in ACN:H2O (10:90, v/v) containing 0.1% ammonium hydroxide. Mobile phase B consisted of 10 mM ammonium acetate in ACN:H2O (95:5, v/v) containing 0.1% ammonium hydroxide. Mobile phase B was initially held at 95% for 2 min and then decreased to 30% over 18 min. Mobile phase B was held for 6 min at 35%, then raised to 95% over 1 min. The column was re-equilibrated at 95% mobile phase B for 8 min. 2 µL of extract was injected by a Vanquish Split Sampler HT autosampler (Thermo Scientific).

The LC system was coupled to a Q Exactive-HF Orbitrap mass spectrometer through a heated electrospray ionization (HESI II) source (Thermo Scientific). Source conditions were as follow: HESI II and capillary temperature at 350 °C, sheath gas flow rate at 40 units, aux gas flow rate at 15 units, sweep gas flow rate at 1 units, spray voltage at |3.0 kV| for both positive and negative modes, and S-lens RF at 50.0 units. The MS was operated in a polarity switching mode acquiring positive and negative full MS and MS2 spectra (Top10) within the same injection. Acquisition parameters for full MS scans in both modes were 60,000 resolution, 1×10e6 automatic gain control (AGC) target, 100 ms ion accumulation time (max IT), and 70 to 900 m/z scan range. MS2 scans in both modes were then performed at 45,000 resolution, 1×10e5 AGC target, 100 ms max IT, 1.0 m/z isolation window, stepped normalized collision energy (NCE) at 20, 30, 40, and a 30.0 s dynamic exclusion.

## RP-LC-MS Lipidomics

10 µL of sample extract was injected via Vanquish Split Sampler HT autosampler (Thermo Scientific) onto an ACQUITY CSH C18 column held at 50 °C (100 mm x 2.1 mm x 1.7 μm particle size; Waters) using a Vanquish Binary Pump (400 μL/min flow rate; Thermo Scientific). A reversed phase gradient length of 30 min was used to separate the lipids, using mobile phase A, consisting of 10 mM ammonium acetate in acetonitrile:water (70:30, v/v) containing 250 μL/L acetic acid, and mobile phase B, consisting of 10 mM ammonium acetate in isopropanol:acetonitrile (90:10, v/v) with the same additives. Mobile phase B was initially held at 2% for 2 min and then increased to 30% over 3 min. Mobile phase B was further increased to 50% over 1 min, then raised to 85% over 14 min, and finally raised to 99% over 1 min and held at 99% for 7 min. Mobile phase B was then decreased to 2% over 0.25 min, and the column was re-equilibrated with mobile phase B at 2% for 1.75 min before the next injection.

The LC system was coupled online to a Q Exactive HF Orbitrap mass spectrometer through a heated electrospray ionization (HESI II) source (Thermo Scientific). In both ionization modes, the HESI-II and capillary temperature, spray voltage, S-lens RF level, sheath gas, aux gas, and sweep gas were held at 300 °C, |3.5 kV|, 90.0 units, 25 units, 15 units, and 5 units, respectively. The MS was operated in a polarity switching mode acquiring positive and negative full MS and MS2 spectra (Top2) within the same injection. Acquisition parameters for full MS scans in both modes were 17,500 resolution, 1×106 automatic gain control (AGC) target, 100 ms ion accumulation time (max IT), and 200 to 1600 m/z scan range. Data-dependent MS2 scans in both modes were then performed at 17,500 resolution, 1×105 AGC target, 50 ms max IT, 1.0 m/z isolation window, stepped normalized collision energy (NCE) at 20, 30, 40, and a 10.0 s dynamic exclusion.

## LC-MS Data processing

RP–LC–MS raw lipidomics data were processed in Compound Discoverer 3.1 (Thermo Scientific) in conjunction with LipiDex[53](https://www.zotero.org/google-docs/?a3v7dq). In brief, MS1 scans from 100 Da to 5000 Da precursor mass as well as retention time of 0.4 min to 21 min were extracted and aligned, using alignment parameters as follows: 0.2 min retention time tolerance, 10-ppm mass, a minimum peak intensity of 5×105, a maximum peak width of 0.25 min, and a minimum signal-to-noise (S/N) ratio of 1.5, to form distinct chromatographic profiles, or compound groups. From the chromatographic features that were at least 3-fold greater in intensity than blanks, the consequent MS2 features were searched against an *in silico* generated lipid spectral library. Compounds were annotated only if the corresponding MS2 fulfilled the following requirements: a minimum lipid spectral purity of 75% from co-eluting isobaric lipids that elute within a 3.5 median absolute retention time deviation from each other, a minimum MS2 spectral match dot product of 500, a minimum MS2 spectral match reverse dot product of 700, and found within at least two processed files. For individual fatty acid substituents that could not be resolved, the identifications were generated with the sum of the fatty acid substituents. Features were removed from further consideration if the %RSD values from quality control replicates were greater than 30%.

HILIC-LC-MS raw metabolomics data were processed using the default workflow Untargeted Metabolomics using Online Databases, mzLogic and Molecular Networks in Compound Discoverer 3.3 (Thermo Scientific). Annotations for polar metabolites were derived from MS2 libraries using authentic standards, or from mzCloud library matching followed by manual validation of identifications using combined evidence from MS2 library matching score greater than 80, in addition to retention time and the presence of metabolite in databases of plasma metabolites[54](https://www.zotero.org/google-docs/?EyrKQA). Polar metabolite features were removed if %RSD of replicate quality control was >30%.

## Statistical analysis

Data processing was performed in Python 3.7 with the following packages and versions: statsmodels 0.13.2; shap 0.41.0; scikit-learn 1.0.2; scipy 1.7.3; pandas 1.3.5; numpy 1.21.6; networkx 2.6.3; matplotlib 3.5.2; matplotlib-venn 0.11.5; seaborn 0.11.2.

Principal component analysis (PCA) was calculated on all 60 plasma samples (10-week-old males), combining both annotated and unannotated LCMS chromatographic features from reversed phase lipidomics method and HILIC polar metabolomics method. Points represent samples and were plotted based on principal components 1 and 2. Samples were labeled according to the Nile rat label and whether the sample was fasted or non-fasted. In the process of PCA, the fasted 9-week sample from Nile rat A was found to lie within the non-fasted cluster. Further analysis of the metabolite profile revealed elevated amino acids with an outlier effect of > 2 standard deviations compared to other fasted samples, leading us to remove this sample from PCA visualization and discard this sample from further downstream analyses. The heatmap in Figure 2D was generated using python seaborn. Each column is one Nile rat’s annotated metabolite profile in either fasted or non-fasted conditions, averaged across triplicate sampling weeks. Rows are one annotated metabolite, and log2 fold change is given as the difference between the triplicate averaged log2 abundance and the mean log2 abundance of all fasted plasma samples. Rows were hierarchically clustered using method complete linkage with Euclidean (L2 norm) distance metric.

%RSD is calculated for each metabolite feature, within each sampling method (non-fasted and fasted), within each Nile rat, in both Nile rat cohorts (young males and mature males/females). The calculation uses metabolite log2 abundances to find the standard deviation of triplicate sampling across 3 weeks divided by the mean of these 3 values. Young male Nile rat A was excluded from %RSD calculations due to discarding outlier week of fasted sampling. Significance testing between young male non-fasted and fasted metabolite %RSDs was performed using Wilcoxon signed rank test on paired %RSD values among the metabolite groups using scipy wilcoxon function. Calculated p-values from Wilcoxon signed rank were corrected for false discovery rate by Bonferroni method using statsmodels multipletests function.

Multivariate machine learning models were trained using the associated sklearn method (LinearRegression, Lasso, Ridge, ElasticNet, PLSRegression, RandomForestRegressor). Cross validation was performed using sklearn cross\_validate with n\_repeats=200 and n\_splits=6, with random seed set identically for all 6 models ensuring the same training data. Competing models for non-fasted and fasted sampling were trained on all young male non-fasted and fasted plasma samples, and the median R2 value from all 1200 folds were presented. At each fold, the beta coefficient of each metabolite feature was recorded. Metabolite importance is calculated as the average beta coefficient across all 1200 folds divided by the maximum average beta coefficient of all metabolites. Normalized importances were then calculated by taking the absolute value of the importance to set each metabolite’s normalized importance value between 0 and 1. Individual metabolite linear regressions and R2 values of OGTT glucAUC vs. log2 abundance (used in Figure 4D) are calculated using all non-fasted and fasted plasma samples, with the dots on the plots representing the mean log2 abundance from each Nile rat.

Throughout the text, the term individual metabolite linear models is used, which is defined as the regression model given by equation 1.

Significance testing was performed each of the 3 terms in equation 1 using likelihood ratio test in statsmodels ols function. p-values for each of the 3 terms were corrected across all metabolites using Benjamini-Hochberg false discovery rate correction using statsmodels fdrcorrection. Resulting q-values were significant at a value of less than 0.05. Each metabolite also underwent linear regression of OGTT glucAUC versus log2 abundance within each sampling method, and resulting p-values for the effect size (regression slope value; used in Figure 5B) were also corrected for multiple testing using the Benjamini-Hochberg method, with q-values significant at less than 0.05.

# Data Availability

All mass spectrometry files are available in the public repository MassIVE under accession number MSV000091033.

## Code Availability

Code for analysis and figures is contained in GitHub repository at <https://github.com/benton-anderson/nile_rat_multiomics>

# Acknowledgements

This study was supported by The Garland Initiative funded by William K. Bowes Jr., awarded to J.A.T and H.T. Foundation and the National Institutes for Health, National Institute of General Medical Sciences (P41 GM108538) to J.J.C. B.J.A. acknowledges the support of Computation and Informatics in Biology and Medicine NIH T15LM007359 training grant. We thank Jenna Lukes for assistance with Nile rat handling.

## Conflict of Interest

J.J.C. is a consultant for Thermo Fisher Scientific.

# References

[1. International Diabetes Federation. *IDF Diabetes Atlas*.](https://www.zotero.org/google-docs/?gH1g4l)

[2. Suliman, S. M., Shumake, S. A. & Jackson, W. B. Food Preference in the Nile Rat Arvicanthis niloticus. *Trop. Pest Manag.* **30**, 151–158 (1984).](https://www.zotero.org/google-docs/?gH1g4l)

[3. Senzota, R. B. M. A Case of Rodent—Ungulate Resource Partitioning. *J. Mammal.* **64**, 326–329 (1983).](https://www.zotero.org/google-docs/?gH1g4l)

[4. Toh, H., Thomson, J. A. & Jiang, P. Maternal High-Fiber Diet Protects Offspring against Type 2 Diabetes. *Nutrients* **13**, E94 (2020).](https://www.zotero.org/google-docs/?gH1g4l)

[5. King, A. & Bowe, J. Animal models for diabetes: Understanding the pathogenesis and finding new treatments. *Biochem. Pharmacol.* **99**, 1–10 (2016).](https://www.zotero.org/google-docs/?gH1g4l)

[6. Singh, J. *et al.* Characterization of the Nile Grass Rat as a Unique Model for Type 2 Diabetic Polyneuropathy. *J. Neuropathol. Exp. Neurol.* **77**, 469–478 (2018).](https://www.zotero.org/google-docs/?gH1g4l)

[7. Ranaei Pirmardan, E., Barakat, A., Zhang, Y., Naseri, M. & Hafezi-Moghadam, A. Diabetic cataract in the Nile grass rat: A longitudinal phenotypic study of pathology formation. *FASEB J.* **35**, e21593 (2021).](https://www.zotero.org/google-docs/?gH1g4l)

[8. Schneider, J., Kuny, S., Beker, D., Sauvé, Y. & Lemieux, H. Cardiovascular sexual dimorphism in a diet-induced type 2 diabetes rodent model, the Nile rat (Arvicanthis niloticus). *PLOS ONE* **13**, e0208987 (2018).](https://www.zotero.org/google-docs/?gH1g4l)

[9. Toh, H. *et al.* Vascular changes in diabetic retinopathy—a longitudinal study in the Nile rat. *Lab. Invest.* **99**, 1547–1560 (2019).](https://www.zotero.org/google-docs/?gH1g4l)

[10. Rabiu, S. & Rose, R. K. A quantitative study of diet in three species of rodents in natural and irrigated savanna fields. **42**, 55–70 (1997).](https://www.zotero.org/google-docs/?gH1g4l)

[11. Subramaniam, A., Park, B., Raphael, D., Landstrom, M. & Hayes, K. C. Dietary Carbohydrate as Glycemic Load, Not Fat, Coupled with Genetic Permissiveness Favoring Rapid Growth and Extra Calories, Dictate Metabolic Syndrome and Diabetes Induction in Nile Rats (Arvicanthis niloticus). *Nutrients* **14**, 3064 (2022).](https://www.zotero.org/google-docs/?gH1g4l)

[12. Blanchong, J. A. & Smale, L. Temporal Patterns of Activity of the Unstriped Nile Rat, Arvicanthis Niloticus. *J. Mammal.* **81**, 595–599 (2000).](https://www.zotero.org/google-docs/?gH1g4l)

[13. Toh, H. *et al.* A haplotype-resolved genome assembly of the Nile rat facilitates exploration of the genetic basis of diabetes. *BMC Biol.* **20**, 245 (2022).](https://www.zotero.org/google-docs/?gH1g4l)

[14. Abrahamson, M. J. Optimal Glycemic Control in Type 2 Diabetes Mellitus: Fasting and Postprandial Glucose in Context. *Arch. Intern. Med.* **164**, 486–491 (2004).](https://www.zotero.org/google-docs/?gH1g4l)

[15. Zeevi, D. *et al.* Personalized Nutrition by Prediction of Glycemic Responses. *Cell* **163**, 1079–1094 (2015).](https://www.zotero.org/google-docs/?gH1g4l)

[16. Cavalot, F. *et al.* Postprandial Blood Glucose Predicts Cardiovascular Events and All-Cause Mortality in Type 2 Diabetes in a 14-Year Follow-Up: Lessons from the San Luigi Gonzaga Diabetes Study. *Diabetes Care* **34**, 2237–2243 (2011).](https://www.zotero.org/google-docs/?gH1g4l)

[17. Bolsinger, J., Landstrom, M., Pronczuk, A., Auerbach, A. & Hayes, K. C. Low glycemic load diets protect against metabolic syndrome and Type 2 diabetes mellitus in the male Nile rat. *J. Nutr. Biochem.* **42**, 134–148 (2017).](https://www.zotero.org/google-docs/?gH1g4l)

[18. Subramaniam, A., Landstrom, M. & Hayes, K. C. Genetic Permissiveness and Dietary Glycemic Load Interact to Predict Type-II Diabetes in the Nile rat (Arvicanthis niloticus). *Nutrients* **11**, 1538 (2019).](https://www.zotero.org/google-docs/?gH1g4l)

[19. Subramaniam, A., Landstrom, M., Luu, A. & Hayes, K. C. The Nile Rat (Arvicanthis niloticus) as a Superior Carbohydrate-Sensitive Model for Type 2 Diabetes Mellitus (T2DM). *Nutrients* **10**, 235 (2018).](https://www.zotero.org/google-docs/?gH1g4l)

[20. Li-Gao, R. *et al.* Assessment of reproducibility and biological variability of fasting and postprandial plasma metabolite concentrations using 1H NMR spectroscopy. *PLOS ONE* **14**, e0218549 (2019).](https://www.zotero.org/google-docs/?gH1g4l)

[21. Sitnikov, D. G., Monnin, C. S. & Vuckovic, D. Systematic Assessment of Seven Solvent and Solid-Phase Extraction Methods for Metabolomics Analysis of Human Plasma by LC-MS. *Sci. Rep.* **6**, 38885 (2016).](https://www.zotero.org/google-docs/?gH1g4l)

[22. Townsend, M. K. *et al.* Reproducibility of metabolomic profiles among men and women in two large cohort studies. *Clin. Chem.* **59**, 10.1373/clinchem.2012.199133 (2013).](https://www.zotero.org/google-docs/?gH1g4l)

[23. Wang, Y. *et al.* Reproducibility of non-fasting plasma metabolomics measurements across processing delays. *Metabolomics Off. J. Metabolomic Soc.* **14**, 129 (2018).](https://www.zotero.org/google-docs/?gH1g4l)

[24. Harada, S. *et al.* Reliability of plasma polar metabolite concentrations in a large-scale cohort study using capillary electrophoresis-mass spectrometry. *PloS One* **13**, e0191230 (2018).](https://www.zotero.org/google-docs/?gH1g4l)

[25. Breier, M. *et al.* Targeted metabolomics identifies reliable and stable metabolites in human serum and plasma samples. *PloS One* **9**, e89728 (2014).](https://www.zotero.org/google-docs/?gH1g4l)

[26. Linke, V. *et al.* A large-scale genome–lipid association map guides lipid identification. *Nat. Metab.* **2**, 1149–1162 (2020).](https://www.zotero.org/google-docs/?gH1g4l)

[27. Drogan, D. *et al.* Untargeted Metabolic Profiling Identifies Altered Serum Metabolites of Type 2 Diabetes Mellitus in a Prospective, Nested Case Control Study. *Clin. Chem.* **61**, 487–497 (2015).](https://www.zotero.org/google-docs/?gH1g4l)

[28. Chaabo, F., Pronczuk, A., Maslova, E. & Hayes, K. Nutritional correlates and dynamics of diabetes in the Nile rat (Arvicanthis niloticus): a novel model for diet-induced type 2 diabetes and the metabolic syndrome. *Nutr. Metab.* **7**, 29 (2010).](https://www.zotero.org/google-docs/?gH1g4l)

[29. Hayashi, T. *et al.* Patterns of Insulin Concentration During the OGTT Predict the Risk of Type 2 Diabetes in Japanese Americans. *Diabetes Care* **36**, 1229–1235 (2013).](https://www.zotero.org/google-docs/?gH1g4l)

[30. Fahy, E., Sud, M., Cotter, D. & Subramaniam, S. LIPID MAPS online tools for lipid research. *Nucleic Acids Res.* **35**, W606–W612 (2007).](https://www.zotero.org/google-docs/?gH1g4l)

[31. Liebisch, G. *et al.* Update on LIPID MAPS classification, nomenclature, and shorthand notation for MS-derived lipid structures. *J. Lipid Res.* **61**, 1539–1555 (2020).](https://www.zotero.org/google-docs/?gH1g4l)

[32. M. Parsons, H., R. Ekman, D., W. Collette, T. & R. Viant, M. Spectral relative standard deviation: a practical benchmark in metabolomics. *Analyst* **134**, 478–485 (2009).](https://www.zotero.org/google-docs/?gH1g4l)

[33. *Applied linear statistical models*. (McGraw-Hill, 2005).](https://www.zotero.org/google-docs/?gH1g4l)

[34. Boulesteix, A.-L., Janitza, S., Kruppa, J. & König, I. R. Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. *WIREs Data Min. Knowl. Discov.* **2**, 493–507 (2012).](https://www.zotero.org/google-docs/?gH1g4l)

[35. Cook, R. D. & Forzani, L. Partial least squares prediction in high-dimensional regression. *Ann. Stat.* **47**, 884–908 (2019).](https://www.zotero.org/google-docs/?gH1g4l)

[36. Morze, J. *et al.* Metabolomics and Type 2 Diabetes Risk: An Updated Systematic Review and Meta-analysis of Prospective Cohort Studies. *Diabetes Care* **45**, 1013–1024 (2022).](https://www.zotero.org/google-docs/?gH1g4l)

[37. Brennan, L. & Gibbons, H. Sex matters: a focus on the impact of biological sex on metabolomic profiles and dietary interventions. *Proc. Nutr. Soc.* **79**, 205–209 (2020).](https://www.zotero.org/google-docs/?gH1g4l)

[38. Pettersson, U. S., Waldén, T. B., Carlsson, P.-O., Jansson, L. & Phillipson, M. Female Mice are Protected against High-Fat Diet Induced Metabolic Syndrome and Increase the Regulatory T Cell Population in Adipose Tissue. *PLOS ONE* **7**, e46057 (2012).](https://www.zotero.org/google-docs/?gH1g4l)

[39. Tabassum, R. *et al.* Lipidome‐ and Genome‐Wide Study to Understand Sex Differences in Circulatory Lipids. *J. Am. Heart Assoc.* **11**, e027103 (2022).](https://www.zotero.org/google-docs/?gH1g4l)

[40. Beyene, H. B. *et al.* High-coverage plasma lipidomics reveals novel sex-specific lipidomic fingerprints of age and BMI: Evidence from two large population cohort studies. *PLoS Biol.* **18**, e3000870 (2020).](https://www.zotero.org/google-docs/?gH1g4l)

[41. Park, J.-E., Lim, H. R., Kim, J. W. & Shin, K.-H. Metabolite changes in risk of type 2 diabetes mellitus in cohort studies: A systematic review and meta-analysis. *Diabetes Res. Clin. Pract.* **140**, 216–227 (2018).](https://www.zotero.org/google-docs/?gH1g4l)

[42. Mora, S. Nonfasting for Routine Lipid Testing: From Evidence to Action. *JAMA Intern. Med.* **176**, 1005–1006 (2016).](https://www.zotero.org/google-docs/?gH1g4l)

[43. Farukhi, Z. & Mora, S. Is it time to abandon fasting for routine lipid testing? *Cleve. Clin. J. Med.* **84**, 919–922 (2017).](https://www.zotero.org/google-docs/?gH1g4l)

[44. Agueusop, I., Musholt, P. B., Klaus, B., Hightower, K. & Kannt, A. Short-term variability of the human serum metabolome depending on nutritional and metabolic health status. *Sci. Rep.* **10**, 16310 (2020).](https://www.zotero.org/google-docs/?gH1g4l)

[45. Giesbertz, P. & Daniel, H. Branched-chain amino acids as biomarkers in diabetes. *Curr. Opin. Clin. Nutr. Metab. Care* **19**, 48 (2016).](https://www.zotero.org/google-docs/?gH1g4l)

[46. Bloomgarden, Z. Diabetes and branched-chain amino acids: What is the link? *J. Diabetes* **10**, 350–352 (2018).](https://www.zotero.org/google-docs/?gH1g4l)

[47. Walford, G. A. *et al.* Metabolite Profiles of Diabetes Incidence and Intervention Response in the Diabetes Prevention Program. *Diabetes* **65**, 1424–1433 (2016).](https://www.zotero.org/google-docs/?gH1g4l)

[48. Ginsberg, H. N., Zhang, Y.-L. & Hernandez-Ono, A. Regulation of Plasma Triglycerides in Insulin Resistance and Diabetes. *Arch. Med. Res.* **36**, 232–240 (2005).](https://www.zotero.org/google-docs/?gH1g4l)

[49. Chew, W. S. *et al.* Large-scale lipidomics identifies associations between plasma sphingolipids and T2DM incidence. *JCI Insight* **4**, (2019).](https://www.zotero.org/google-docs/?gH1g4l)

[50. Xuan, Q. *et al.* Serum lipidomics profiles reveal potential lipid markers for prediabetes and type 2 diabetes in patients from multiple communities. *Front. Endocrinol.* **13**, 966823 (2022).](https://www.zotero.org/google-docs/?gH1g4l)

[51. Overmyer, K. A. *et al.* Large-Scale Multi-omic Analysis of COVID-19 Severity. *Cell Syst.* **12**, 23-40.e7 (2021).](https://www.zotero.org/google-docs/?gH1g4l)

[52. Muehlbauer, L. K. *et al.* Rapid Multi-Omics Sample Preparation for Mass Spectrometry. *Anal. Chem.* (2023) doi:10.1021/acs.analchem.2c02042.](https://www.zotero.org/google-docs/?gH1g4l)

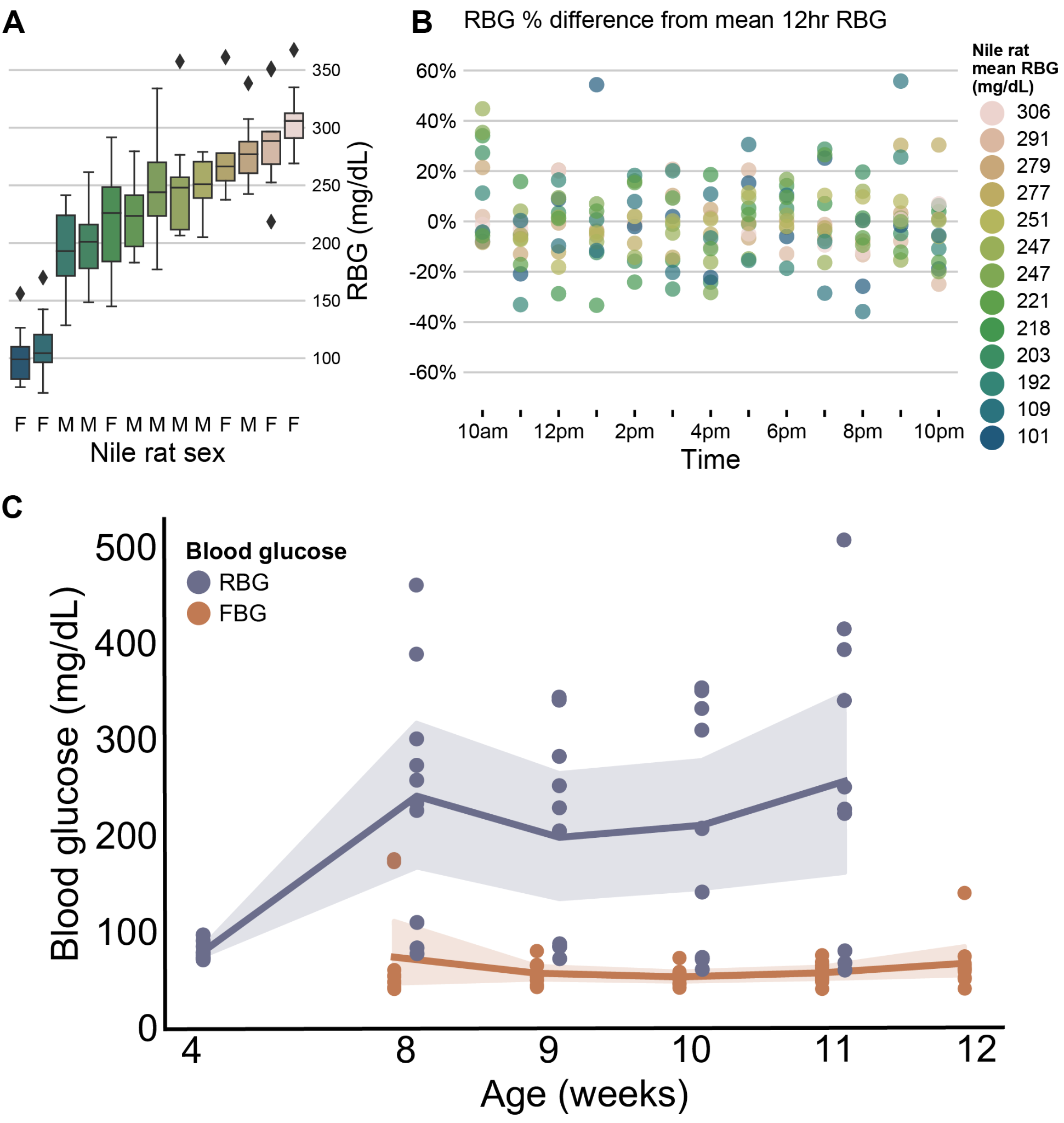
[53. Hutchins, P. D., Russell, J. D. & Coon, J. J. LipiDex: An Integrated Software Package for High-Confidence Lipid Identification. *Cell Syst.* **6**, 621-625.e5 (2018).](https://www.zotero.org/google-docs/?gH1g4l)

[54. Wishart, D. S. *et al.* HMDB 5.0: the Human Metabolome Database for 2022. *Nucleic Acids Res.* **50**, D622–D631 (2022).](https://www.zotero.org/google-docs/?gH1g4l)

[55. Thirty Years of Research on the Dawn Phenomenon: Lessons to Optimize Blood Glucose Control in Diabetes | Diabetes Care | American Diabetes Association. https://diabetesjournals.org/care/article/36/12/3860/33148/Thirty-Years-of-Research-on-the-Dawn-Phenomenon.](https://www.zotero.org/google-docs/?gH1g4l)

# 

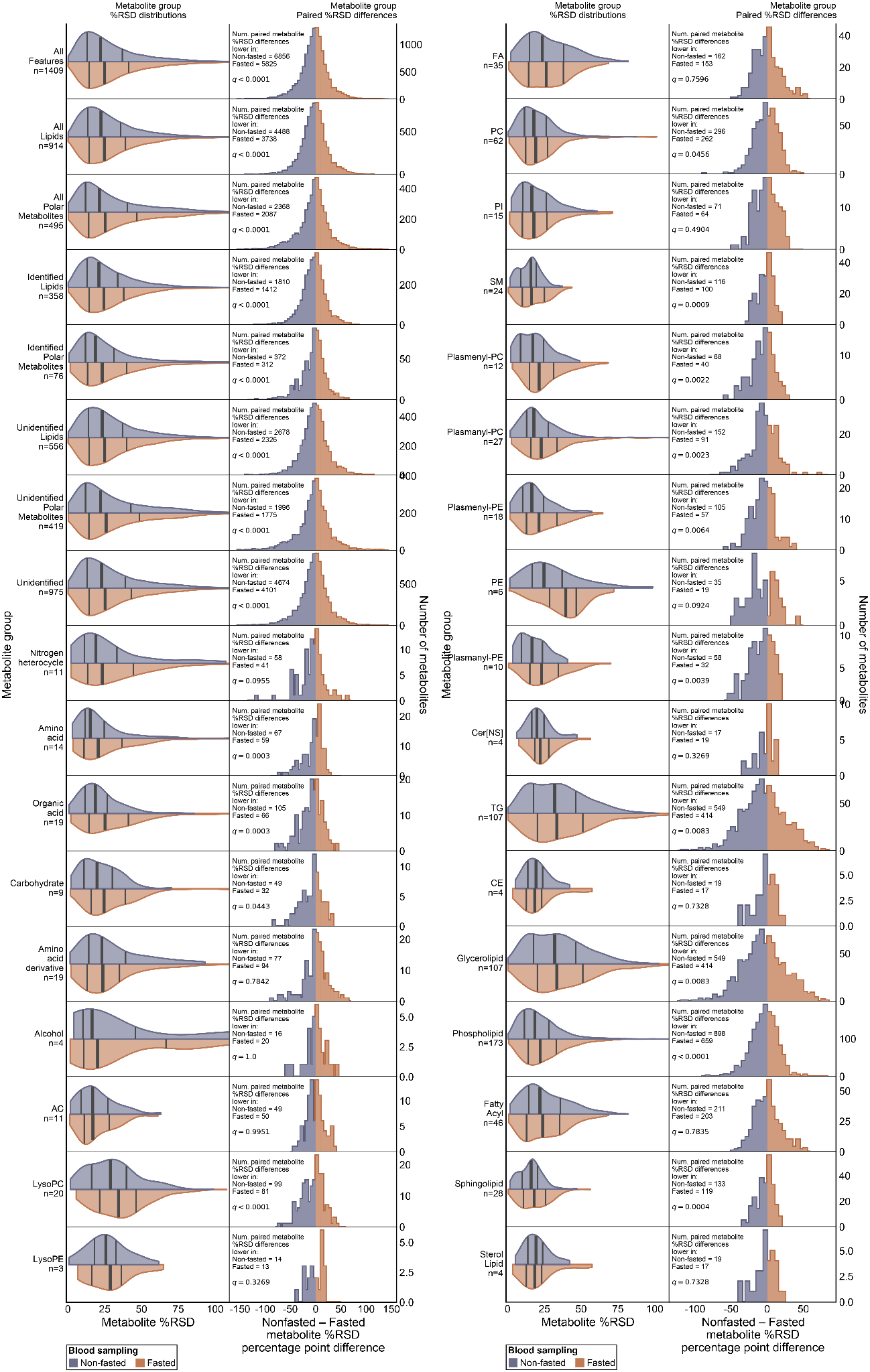
# Supplemental Figures



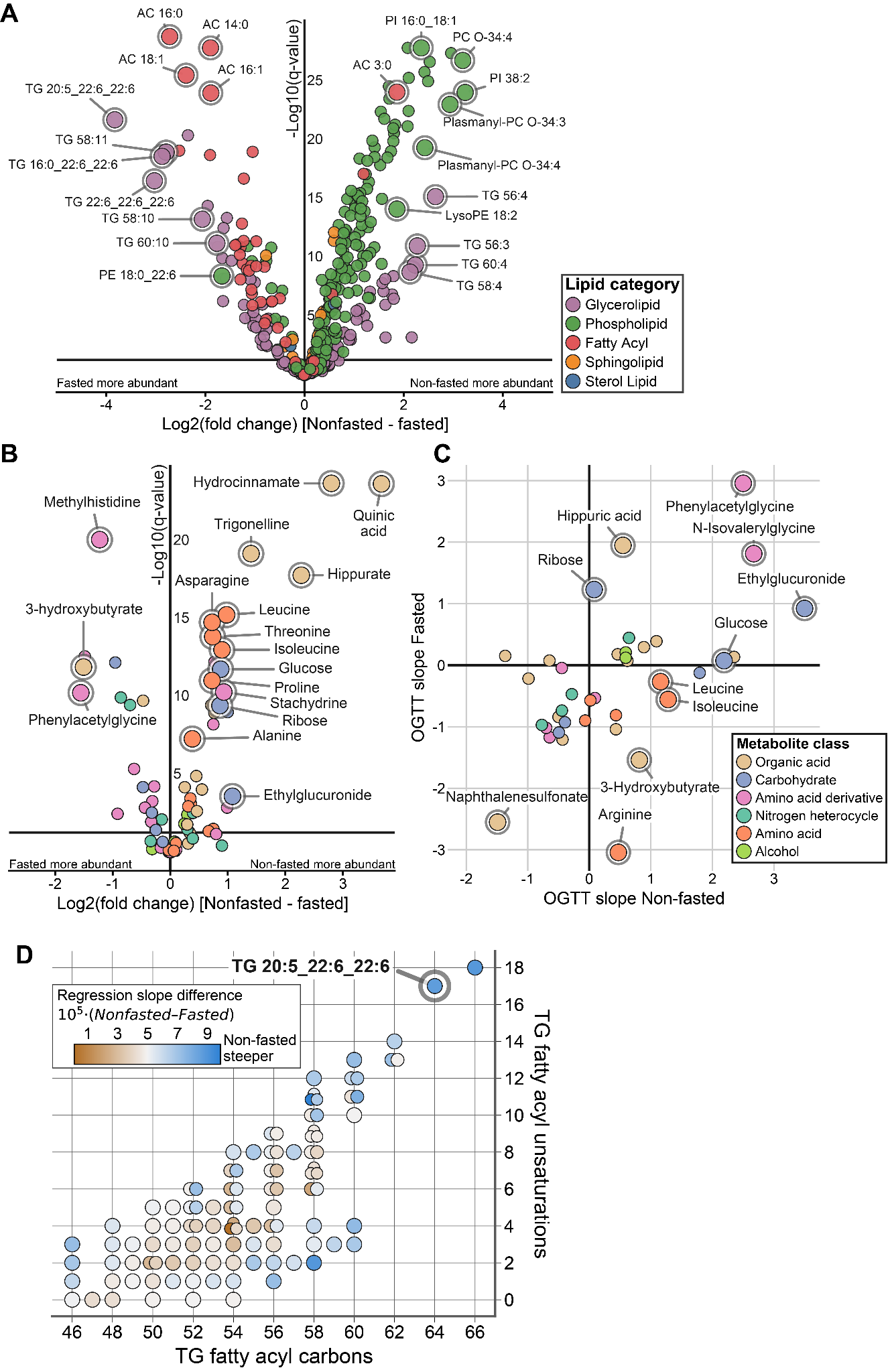
**Supplemental Figure 1 - Nile rat hourly RBG testing across active hours shows no significant difference in variance across mean RBGs, and no significant difference in average RBG throughout the day**

(**A**) Boxplots show distribution of random blood glucose measurements taken every 1 hour over 12 hours from 10 am to 10 pm in a cohort of 6 female and 7 male Nile rats (age 38-42 weeks). Significance testing using Levene’s test reveals no significant difference in the variance among all Nile rats’ RBG 12-hour distributions (*p=0.99*).

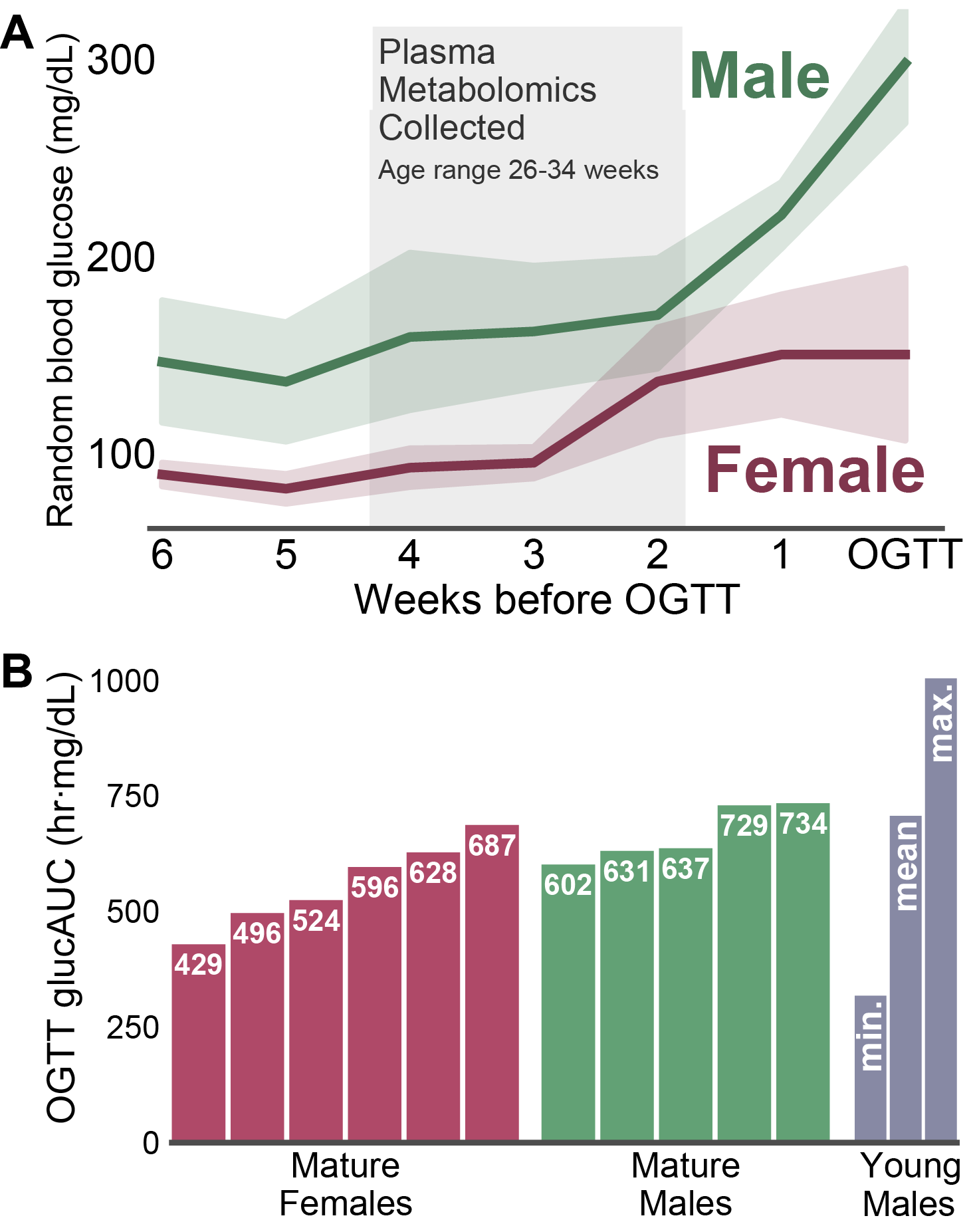
(**B**) Scatterplot of RBG measurements from all 13 Nile rats, grouped by hour of day. Dots are colored by Nile rat’s mean RBG value from the 12 hour period. RBG Percent difference from mean 12 hour RBG is calculated within each animal for all hourly measurements as [(hourRBG - meanRBG) / meanRBG] \* 100%. Significance testing using one-way ANOVA shows no significant difference among hourly RBGs (*p=0.14*). The appearance of increased RBGs at 10 am could be attributable to the dawn phenomenon[55](https://www.zotero.org/google-docs/?zHOS48) of heightened morning RBG values, which has been observed in humans. (**C**) Young male Nile rats in the main study cohort underwent random and fasted blood glucose (RBG and FBG) measurements at different ages. Shaded areas indicate 95% bootstrapped confidence intervals for each week’s blood glucose measurements.



**Supplemental Figure 2** - Distributions for %RSDs among levels of metabolite groupings are presented. Split violin plots give the distributions of non-fasted and fasted Nile rat plasma samples. Histograms show the percentage point difference between %RSD for each metabolite between the two methods. Significance testing on paired %RSD measurements was performed using Wilcoxon signed rank test. *p*-values were corrected using Benjamini-Hochberg false discovery correction. *q-*values below 0.05 were considered significant.

****

**Supplemental Figure 3 - Lipids and polar metabolites volcano plots. Fasted vs. non-fasted regression slope plot for polar metabolites.** (**A**) Volcano plot showing log2 fold change of non-fasted samples minus fasted samples for all annotated lipids, with select lipids highlighted. (**B**) Volcano plot showing log2 fold change of non-fasted samples minus fasted samples for all annotated polar metabolites, with select polar metabolites highlighted. (**C**) Plot of regression slopes in fasted versus non-fasted individual metabolite linear models. Metabolites with regression slope effect size *q*-value > 0.1 have been excluded. (**D**) Similar to Figure 5D, all annotated TGs are separated by fatty acyl carbon counts and unsaturations, with dots colored by the difference in OGTT glucAUC regression slope between non-fasted samples and fasted samples.

****

**Supplemental Figure 4 -** (**A**)A cohort of mature Nile rats were monitored weekly for signs of elevated random blood glucose, with concurrent non-fasted blood sampling. Triplicate plasma samples for %RSD measurement were selected in a 3 week range prior to elevated RBG. The average age at sampling was 30 weeks, with minimum sampling age 26 weeks and maximum sampling age 34. (**B**) Subsequent OGTT glucAUC testing on the cohort of mature Nile rats revealed a range of values that fall within the range of young males assessed earlier in the study, indicating a range of impaired glucose tolerances.