



Integrated Multi-Omics

amPD-1 Effects in DIO and WT Mice – Serum

Integrated Metabolomics and Proteomics



Table of Contents

<i>Project Summary</i>	2
<i>Experimental Methods</i>	3
Statistical analysis	3
<i>Results</i>	4
Unsupervised analyses	4
Integrated pathway analysis	7
Integrated network analysis	9
<i>Interpretation</i>	12
<i>Appendix</i>	16

Project Summary

Sample description:

Forty mouse serum samples were received. Twenty of the samples were from DIO mice bearing MC-38 subcutaneous tumors and twenty of the samples were from WT mice with the same tumor model. Ten mice from both the DIO and WT groups were treated with amPD-1 the remaining ten in each group were untreated.

Goal:

To identify and characterize the metabolic differences between the serum of tumor-bearing DIO and WT and to determine the differential metabolic responses to amPD-1 treatment between WT and DIO mice.

Analysis summary:

The processed SomaScan proteomic profiles were combined with the processed metabolomic profiles. Unsupervised analysis was performed to assess separation in the merged data across the sample groups. From the merged data, an integrated pathway analysis was performed as well as a network analysis that combined knowledge-driven connections between analytes and data-driven correlations between metabolites and proteins to reveal modules of dysregulation across the sample groups.

Conclusions:

Integration of the metabolomics and proteomics data provided greater separation of the sample groups than either approach individually, demonstrating the complementary measurements of protein and metabolite abundance. Metabolomics data provided separation between DIO and WT samples, while the proteomics data led to separation of amPD-1 treatment groups. Integrated pathway analysis revealed dysregulation of Warburg effect analytes that were increased in DIO mice relative to WT and decreased with amPD-1 treatment in DIO mice. Integrated network analysis of the metabolomics and proteomics data revealed modules of lipid and pentose phosphate metabolism having differential activity between WT and DIO mice and after amPD-1 treatment in WT and DIO mice. In particular, correlated levels of ribose 5-phosphate and transketolase indicated reduced activity of the non-oxidative branch of the pentose phosphate pathway with amPD-1 treatment in DIO mice. This pathway was not affected by amPD-1 treatment in WT mice. Future experiments into tumor tissue metabolomic and proteomic signatures would provide more mechanistic insight into the source of the metabolic alterations characterized in the serum.

Experimental Methods

Statistical analysis

For the unsupervised analysis (principle components analyses and clustering), the processed metabolic and proteomic profiles (both $\log_2$ transformed) were concatenated. When filtering based on statistical significance the corrected p-value from the one-way ANOVA performed on each datatype individually was used with a $q < 0.05$ cutoff.

The joint over-representation analysis was performed with a Fisher's Exact test that compares the expected number of analytes (proteins + metabolites) found to be statistically significant in each pathway with the number of significant analytes found in each pathway. For the over-representation analysis, a p-value cutoff of 0.05 was used. To consider a pathway for enrichment, at least five analytes must have been measured that belong to a given pathway. PathBank was used as the pathway database.

To integrate the metabolomic and proteomic profiles, the Pearson correlation between each metabolite/protein pair was computed. Additionally, connections between metabolites and proteins from the Human Metabolome Database were used to define knowledge driven interactions between proteins and metabolites. Further details on the network integration are provided in the Results.

Results

Unsupervised analyses

To examine the global structure of the integrated dataset, the metabolomic and proteomic profiles (after log2 transformation) were concatenated for each sample. In total, the merged profile amounted to 6,418 proteins and 6,974 metabolites measured in each sample. Principal components analysis (PCA) was applied to visualize the merged profiles, revealing strong separation between WT and DIO samples (Figure 1). Untreated WT samples are shown in blue, treated WT samples are shown in green, untreated DIO samples are shown in purple, and treated DIO samples are shown in red.

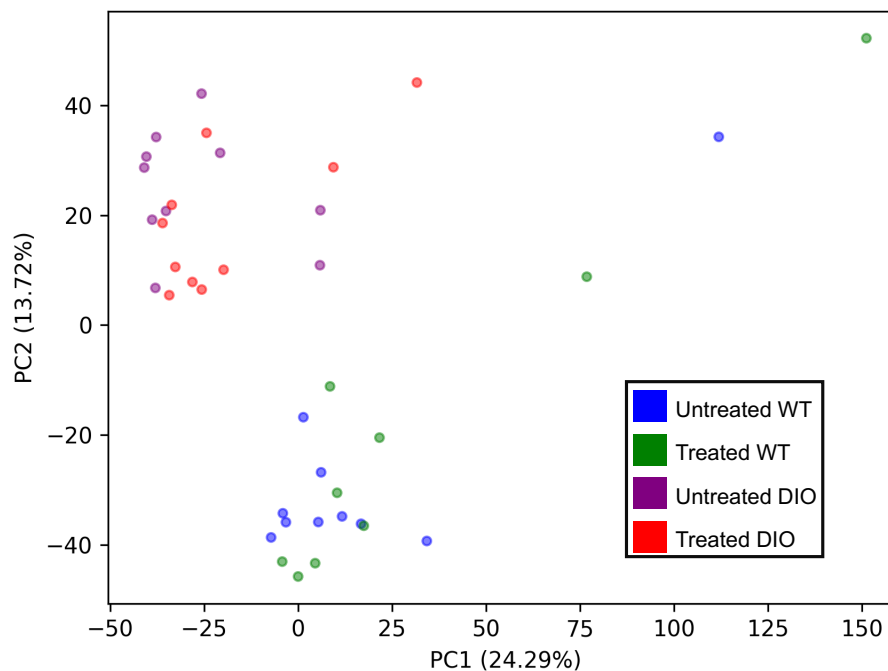


Figure 1: Principal components analysis of serum multi-omics data. The merged profiles for each sample are visualized in the scatter plot above. Each dot represents a sample. Untreated WT samples are shown in blue, treated WT samples are shown in green, untreated DIO samples are shown in purple, and treated DIO samples are shown in red.

Individual analyte (metabolites and proteins) levels are displayed as a heatmap in Figure 2. Samples (columns) are colored as in Figure 1. The rows of the heatmap indicate the data type of the analyte (white = metabolite, black = protein). Broad increases in metabolite levels drive the separation between DIO and WT mice.

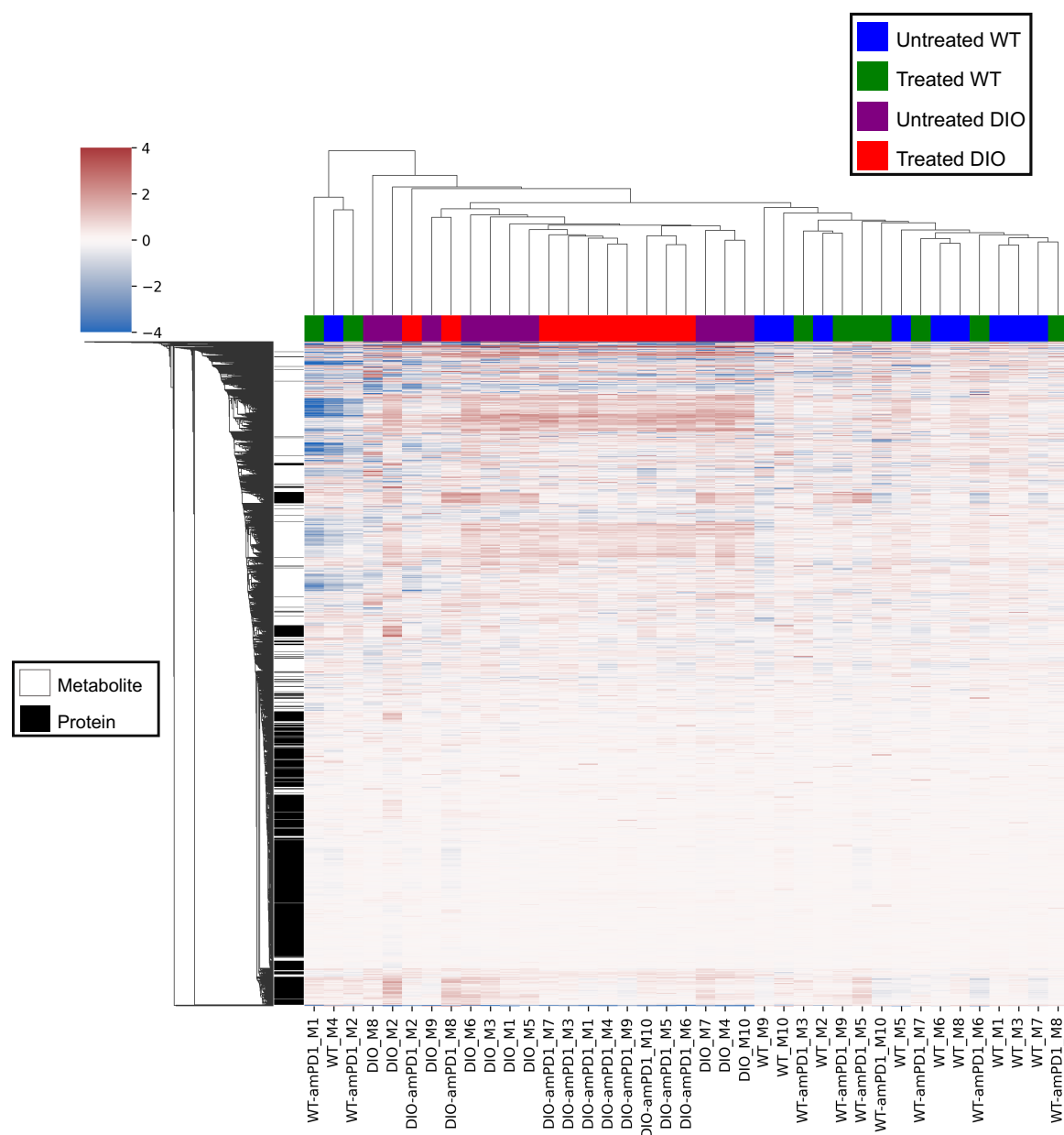


Figure 2: Heatmap of serum multi-omics data. The merged profiles for each sample are visualized in the heatmap above. Each column represents a sample, and each row represents an analyte. Columns are colored according to experimental group. Untreated WT samples are shown in blue, treated WT samples are shown in green, untreated DIO samples are shown in purple, and treated DIO samples are shown in red. The rows are colored by the data type of the analyte (white = metabolite, black = protein). The color of each cell indicates the log₂(fc) relative to the mean untreated WT level of each analyte.

Next, only analytes with a statistically significant (one-way ANOVA, corrected p-value < 0.05) alteration across at least one of the sample groups were kept (1,536 metabolites and 673 proteins). The PCA and heatmap visualizations considering only the statistically significant analytes are shown in Figure 3 and Figure 4, respectively. When considering only these compounds, separation between DIO treated and untreated samples can now be seen. The heatmap in Figure 4 clearly shows the metabolite alterations driving separation between the DIO and WT groups, while protein alterations are primarily responsible for the separation between treated and untreated DIO samples. No separation between treated and untreated WT samples is observed.

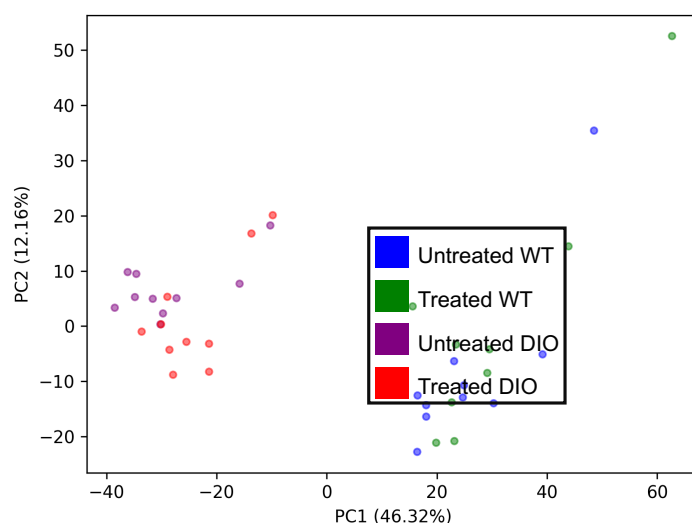


Figure 3: Principal components analysis of serum statistically significant analytes. The merged profiles (considering only analytes that showed a statistically significant alteration in at least one of the samples groups) for each sample are visualized in the scatter plot above. Each dot represents a sample. Untreated WT samples are shown in blue, treated WT samples are shown in green, untreated DIO samples are shown in purple, and treated DIO samples are shown in red.

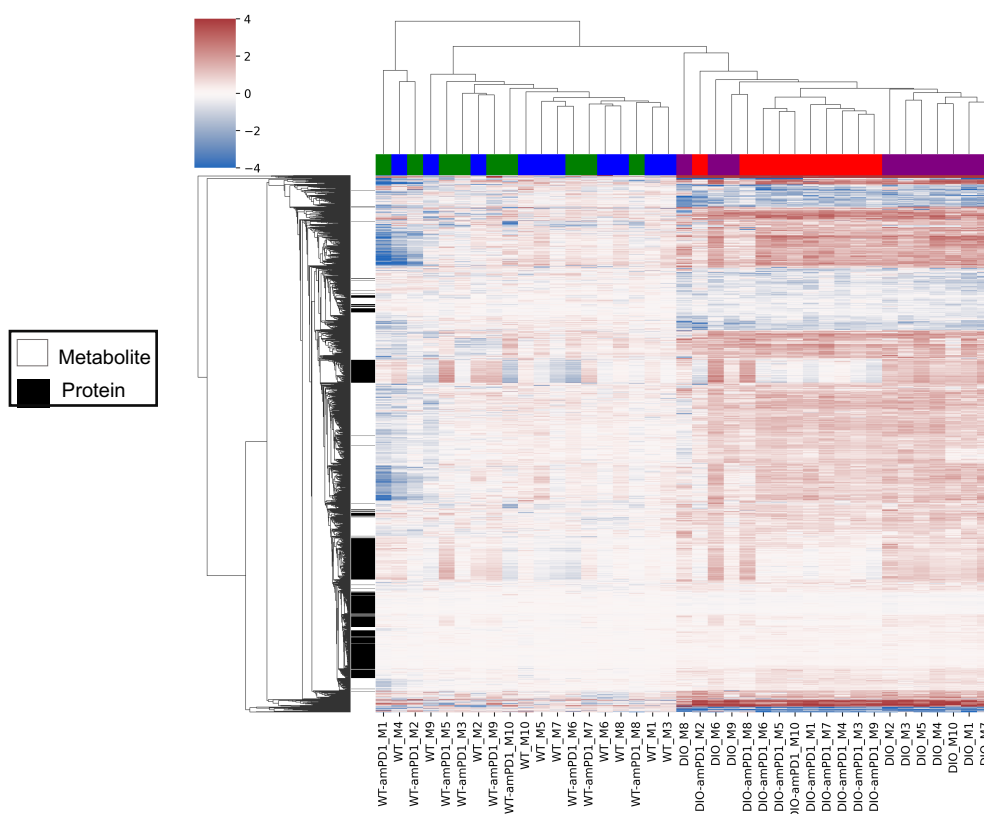


Figure 4: Heatmap of serum statistically significant multi-omics data. The merged profiles (considering only analytes that showed a statistically significant alteration in at least one of the samples groups) for each sample are visualized in the heatmap above. Each column represents a sample, and each row represents an analyte. Columns are colored according to experimental group. Untreated WT samples are shown in blue, treated WT samples are shown in green, untreated DIO samples are shown in purple, and treated DIO samples are shown in red. The rows are colored by the data type of the analyte (white = metabolite, black = protein). The color of each cell indicates the $\log_2(\text{fc})$ relative to the mean untreated WT level of each analyte.

Integrated pathway analysis

Next, the merged profiles were leveraged to perform an over-representation pathway analysis that integrates both metabolites and proteins into defined metabolic pathways from PathBank. This analysis looked for metabolic pathways that were overrepresented in the statistically altered proteins and metabolites. As a large fraction of the statistically significant metabolites were lipids or unidentified compounds, which are not included in this analysis, the pathway analysis results were largely driven by the proteomics data. First, we performed the analysis on analytes dysregulated between untreated WT and DIO mice (Figure 5). The analysis revealed a single enriched pathway, the Warburg effect, which was driven solely by ten statistically significant proteins.

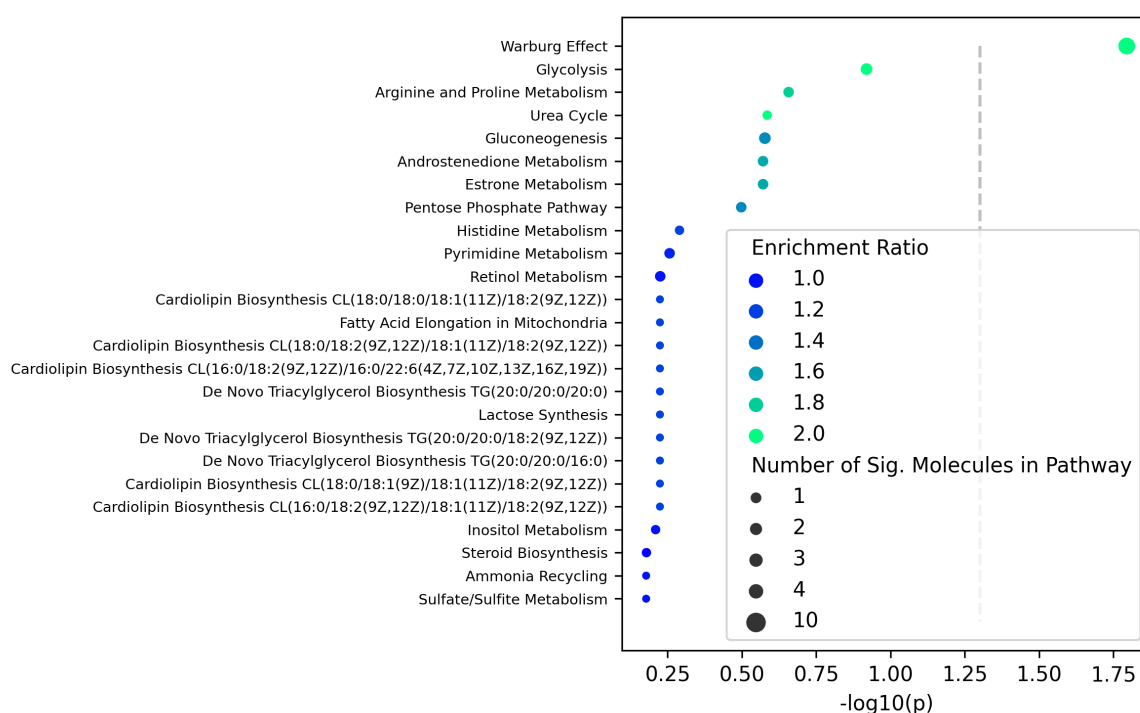


Figure 5: Pathway analysis of serum analytes showing statistically significant differences between untreated WT and DIO samples. The pathways found to be enriched between untreated DIO and WT samples (both untreated) is shown in the dot plot with the corresponding significance levels on the x-axis. The grey dashed line indicates the $p = 0.05$ significance threshold. The size of the dots indicates the number of analytes (proteins and metabolites) that were statistically different between the sample groups, and the color indicates the enrichment ratio (the number of statistically significant analytes observed in the pathway divided by what is expected by random chance).

Next, to interpret the differences in response of amPD-1 treatment, we again performed pathway analysis on the protein and metabolite alterations present after amPD-1 treatment in WT mice (Figure 6) and DIO mice (Figure 7). Notably, there was no overlap between the enriched pathways found in DIO and WT mice. In WT mice, valine leucine and isoleucine degradation was enriched. In DIO mice, the Warburg effect, glycolysis, and gluconeogenesis were enriched.

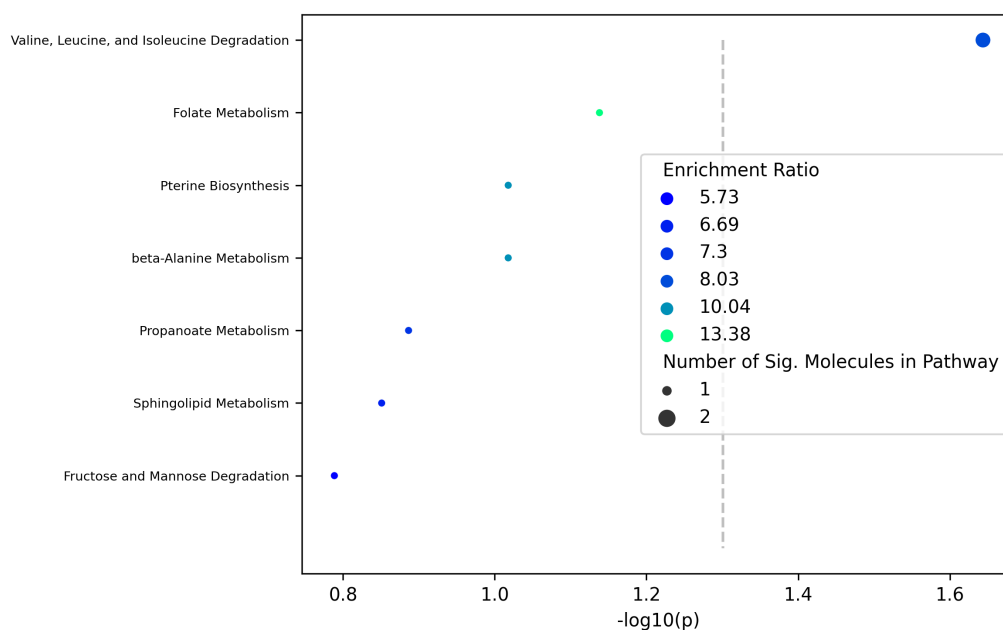


Figure 6: Pathway analysis of serum analytes showing statistically significant differences between untreated and treated WT samples. The pathways found to be enriched between untreated WT and treated WT samples is shown in the dot plot with the corresponding significance levels on the x-axis. The grey dashed line indicates the $p = 0.05$ significance threshold. The size of the dots indicates the number of analytes (proteins and metabolites) that were statistically different between the sample groups, and the color indicates the enrichment ratio (the number of statistically significant analytes observed in the pathway divided by what is expected by random chance).

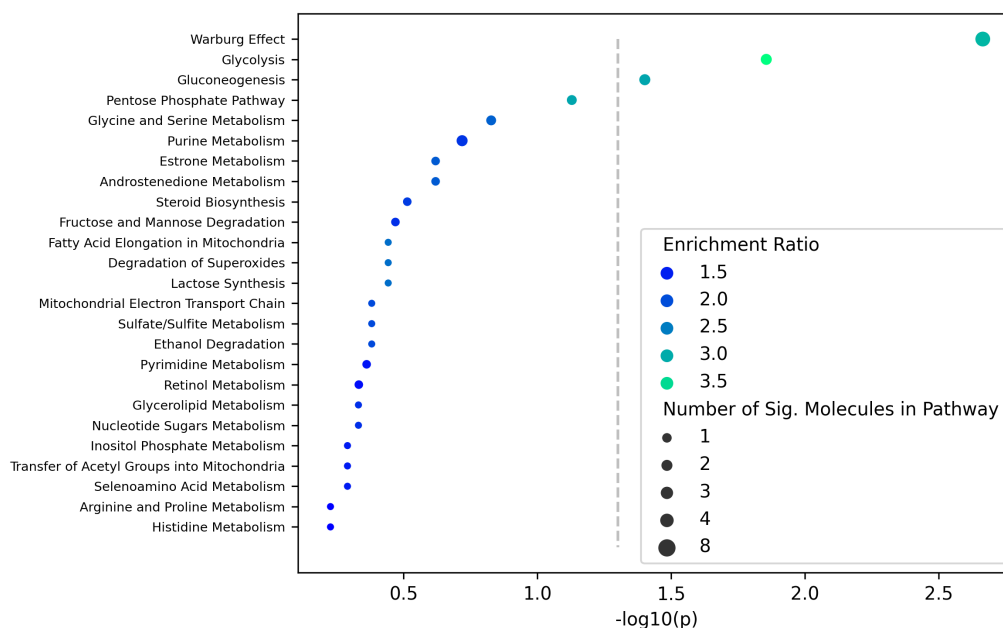


Figure 7: Pathway analysis of serum analytes showing statistically significant differences between untreated and treated DIO samples. The pathways found to be enriched between untreated DIO and treated DIO samples is shown in the dot plot with the corresponding significance levels on the x-axis. The grey dashed line indicates the $p = 0.05$ significance threshold. The size of the dots indicates the number of analytes (proteins and metabolites) that were statistically different between the sample groups, and the color indicates the enrichment ratio (the number of statistically significant analytes observed in the pathway divided by what is expected by random chance).

Integrated network analysis

A primary advantage of having paired metabolomics and proteomics data is that knowledge and data-driven networks can be constructed to link metabolite alterations with protein alterations, revealing modules of dysregulation across the sample groups. To use such an approach, first a data-driven network linking metabolites and proteins must be constructed by correlating the levels of each metabolite against each protein. Figure 8 shows the correlation of each protein ($n=673$) and each metabolite ($n=1,536$) that showed a statistically significant alteration (one-way ANOVA) between at least two of the sample groups. The vast majority of metabolites showed at least a moderate correlation with several proteins. Over half of the proteins, on the other hand, only showed a substantial correlation with a few metabolites.

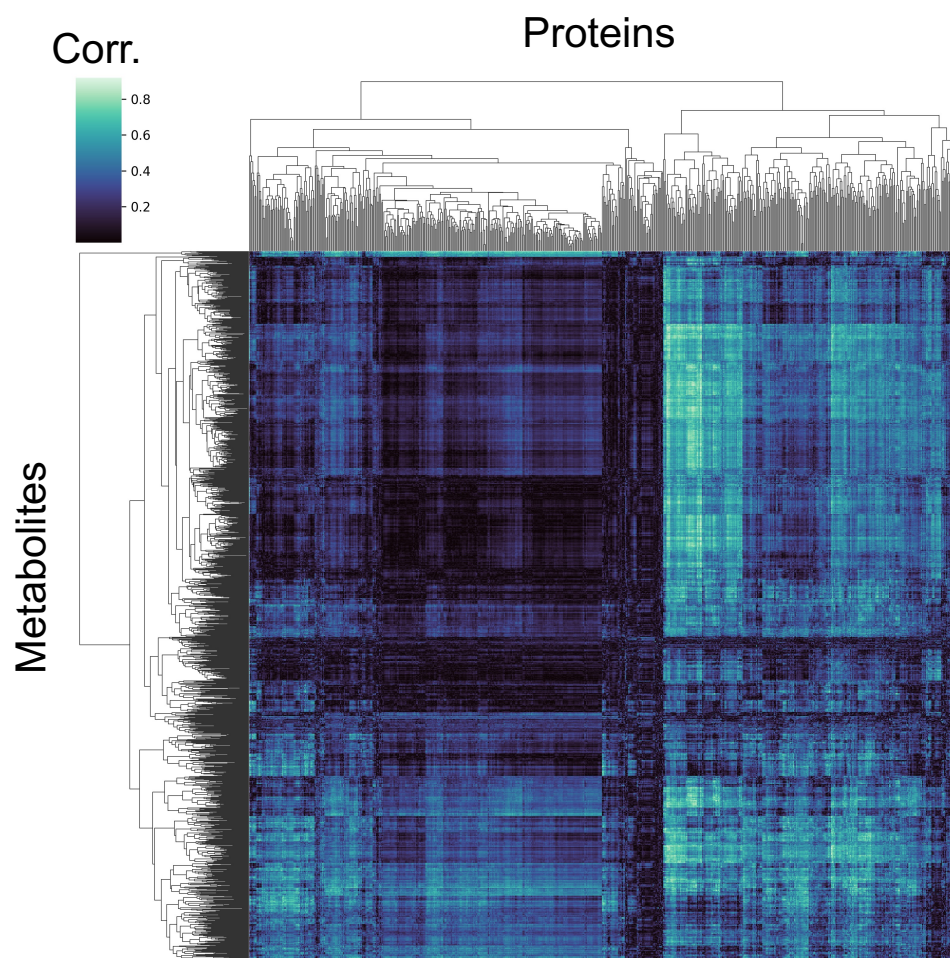


Figure 8: Data driven protein-metabolite network. The Pearson correlation between each metabolite and protein that showed statistically significant alterations between at least two of the sample groups. The rows represent metabolites, and the columns represent proteins.

To incorporate existing information about connections between metabolites and proteins (e.g., when a metabolite is the reactant/product of a metabolic reaction catalyzed by a particular protein), we also gathered known connections between metabolites and proteins measured in this study, forming a knowledge-driven network. To integrate the data-driven and knowledge-driven networks, we define an edge weight between a metabolite and protein to be the Pearson correlation between the two analytes plus 1.0 if there is known relationship between the protein and metabolite. As there are tens-of-thousands of metabolite/protein pairs, we need to prune this network by eliminating edges with low weights to improve the interpretability of the network and enable visualization of the specific amolecular changes across the dataset. Accordingly, we removed edges with a weight less than 0.9. This cutoff was chosen based on the distribution of edge weights and plotting the log10 transform of the number of edges with weights above a certain threshold and finding the elbow point (Figure 8). After pruning, we are left with an interpretable set of tightly connected metabolites and proteins. Only metabolites and proteins that show at least one connection above this threshold remain in the integrated network.

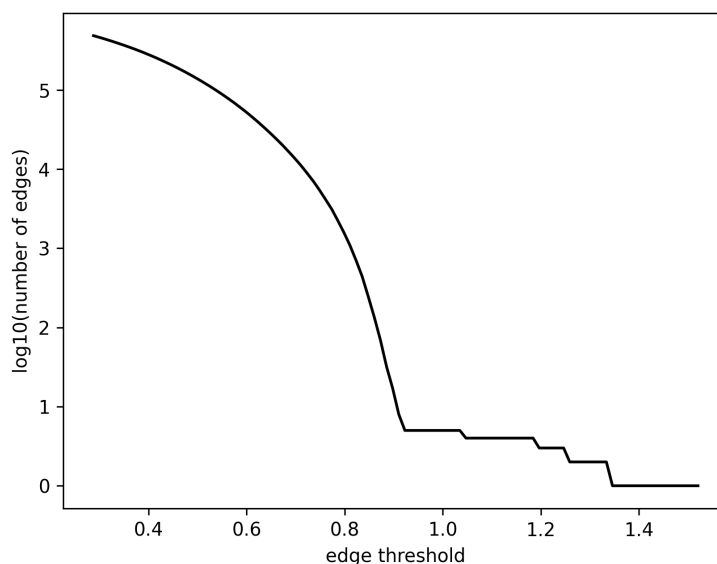


Figure 9: Determining edge weight threshold for integrated network. The number of edges (connections between metabolites and proteins) with a weight above a particular threshold is plotted as a function of threshold. The elbow point represents the cutoff used to prune with network edges.

After pruning, 13 metabolites and 11 proteins remain in the network with 15 edges connecting these analytes. The network is shown in Figure 10. Of note is that several of the unknown lipid compounds are correlated with measured proteins, such as Group XIIIB secretory phospholipase A2-like protein, providing biological context for these unidentified compounds. The levels of the network analytes in each sample are shown in the heatmap given in Figure 11.

Interpretation

From the unsupervised analyses of the merged profiles (metabolomics + proteomics), we see greater separation than with either datatype alone. This result demonstrates the complementary nature of two datasets. The broad increases in many metabolites (primarily lipids) drives the separation between DIO and WT mice. The changes in protein abundance after amPD-1 treatment, on the other hand, drive the separation between treated and untreated DIO mice. Notably, even with the merged profile, little separation is seen for treated and untreated WT mice, underscoring the differential response to treatment between DIO and WT mice. The results of the integrated pathway analysis were largely driven by the proteomic profiles as many of the statistically significant metabolites were unidentified and thus not mappable to pathways. When comparing WT and DIO mice (both untreated), we see an elevation in proteins of the Warburg effect pathway. After treatment with amPD-1 in DIO mice, we see a decrease in these analytes.

To provide a more mechanistic understanding of the differences between WT and DIO mice and their respective response to amPD-1 treatment, we performed an integrated network analysis that defined connections between metabolites and proteins showing significant differences between the four sample groups. The resulting network represents a mechanistic view of metabolic changes observed in this study.

First, to visualize the differences between WT and DIO mice, the network is annotated with the mean log2 fold change in DIO mice relative to the WT mice (Figure 12). This visualization clearly shows the increase in many of the lipid species and associated proteins as well as changes in the pentose phosphate pathway enzyme, transketolase, and metabolite, ribose 5-phosphate.

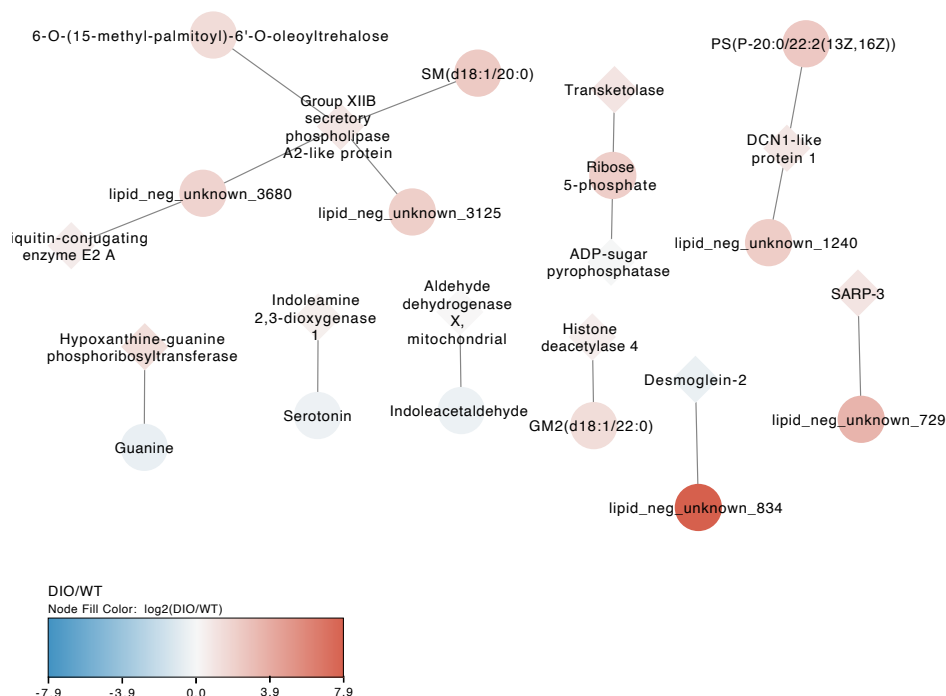


Figure 12: DIO versus WT metabolic alterations. The metabolites (circles) and proteins (diamonds) in the integrated network are shown in the above network plot. The nodes (analytes) are colored based on the mean log2 fold change of the DIO mice relative to the WT samples (both untreated).

Next, to visualize the differences between WT untreated and treated mice, the network was re-annotated with the mean log2 fold change in treated WT mice relative to the untreated WT mice (Figure 13). This visualization clearly shows a decrease in many of lipid species as well as decreases in serotonin and indoleacetaldehyde. Notably, there are also slight increases in pentose phosphate pathway analytes. As discussed above, the changes seen in WT samples is relatively small with log2 fold-changes less than 1.0.

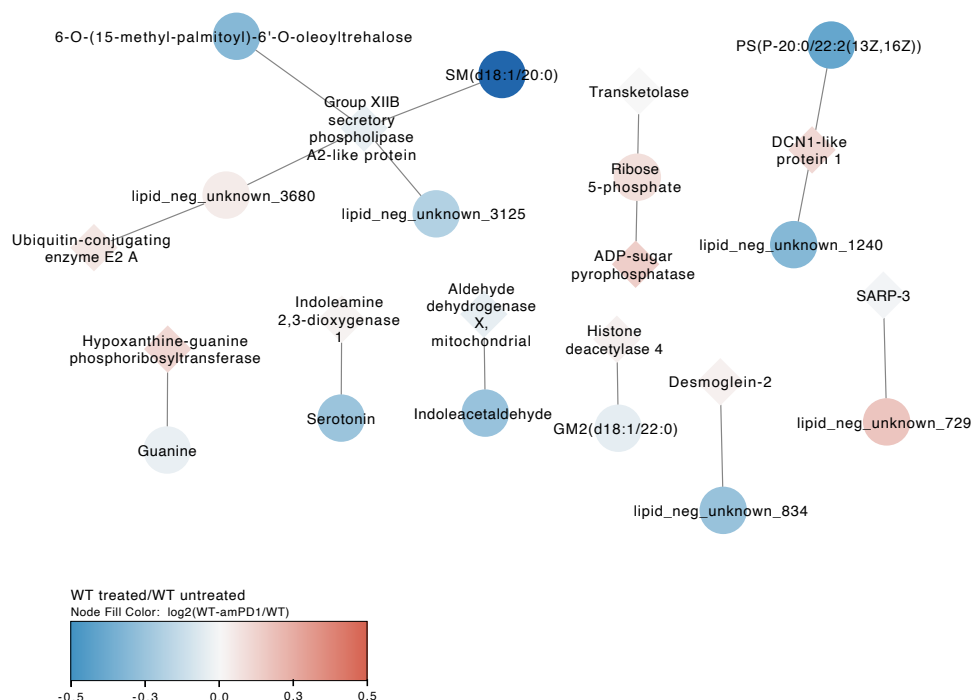


Figure 13: Treated WT versus untreated WT metabolic alterations. The metabolites (circles) and proteins (diamonds) in the integrated network are shown in the above network plot. The nodes (analytes) are colored based on the mean log2 fold change of the treated WT mice relative to the untreated WT samples.

Lastly, we visualized the differences between DIO untreated and treated mice by annotating the network with the mean log2 fold change in treated DIO mice relative to the untreated DIO mice (Figure 14). As shown, amPD-1 treatment in DIO mice had little similarity to the effect in WT mice. Notably, the most substantial dysregulation comes in the pentose phosphate analytes, especially transketolase and ribose 5-phosphate which show nearly four-fold lower abundance in treated DIO mice relative to untreated mice. Interestingly, we see little change to circulating lipid levels, serotonin, and indoleacetaldehyde which showed the largest change in WT mice.

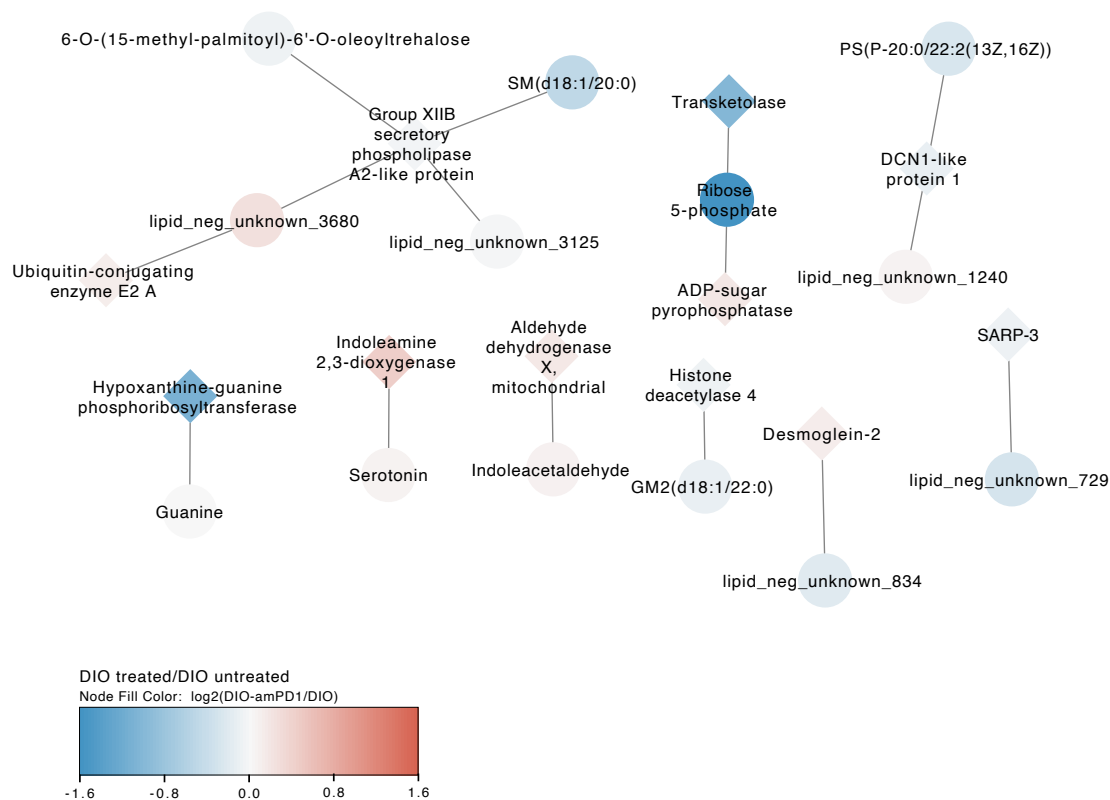


Figure 14: Treated DIO versus untreated DIO metabolic alterations. The metabolites (circles) and proteins (diamonds) in the integrated network are shown in the above network plot. The nodes (analytes) are colored based on the mean log₂ fold change of the treated DIO mice relative to the untreated DIO samples.

In total, we see there are many tightly associated metabolite/protein pairs that are dysregulated between WT and DIO samples and after amPD-1 treatment in DIO mice. Two modules in the network showing such a relationship are related to lipid metabolism (Figure 15) and the pentose phosphate pathway (Figure 16). For the lipids, we see two unknown lipids, a sphingomyelin, and a acyltrehalose that are tightly correlated with the abundance of a ubiquitin conjugating enzyme and a phospholipase. However, only the ubiquitin conjugating enzyme and the connected unknown lipid show a trend with amPD-1 treatment in DIO mice. For the pentose phosphate pathway module, there are two connections to the metabolite ribose 5-phosphate: transketolase and ADP-sugar pyrophosphatase, both of which directly catalyze reactions involving ribose 5-phosphate. However, only transketolase is correlated with ribose 5-phosphate abundance, and ADP-sugar pyrophosphatase does not show a trend with amPD-1 treatment in DIO mice.

Transketolase is a key enzyme in the non-oxidative phase of the pentose phosphate pathway. While the oxidate phase of the pentose phosphate pathway generates NADPH, the non-oxidative phase moves ribose 5-phosphate carbon back to glycolysis. The decrease of both ribose 5-phosphate and transketolase after amPD-1 treatment suggests that amPD-1 decreases the activity of the non-oxidative phase of the pentose phosphate pathway. However, as these data are from serum, the tissue where this change in activity occurred cannot be deduced. Many cancers have well known activation of the pentose phosphate pathway to buffer redox stress and support lipid and nucleotide synthesis. Activation of the non-oxidative branch generally indicates that the pentose phosphate pathway is primarily being used for producing NADPH as the ribose unit is being rerouted back to hexose phosphates rather than towards nucleotide biosynthesis.

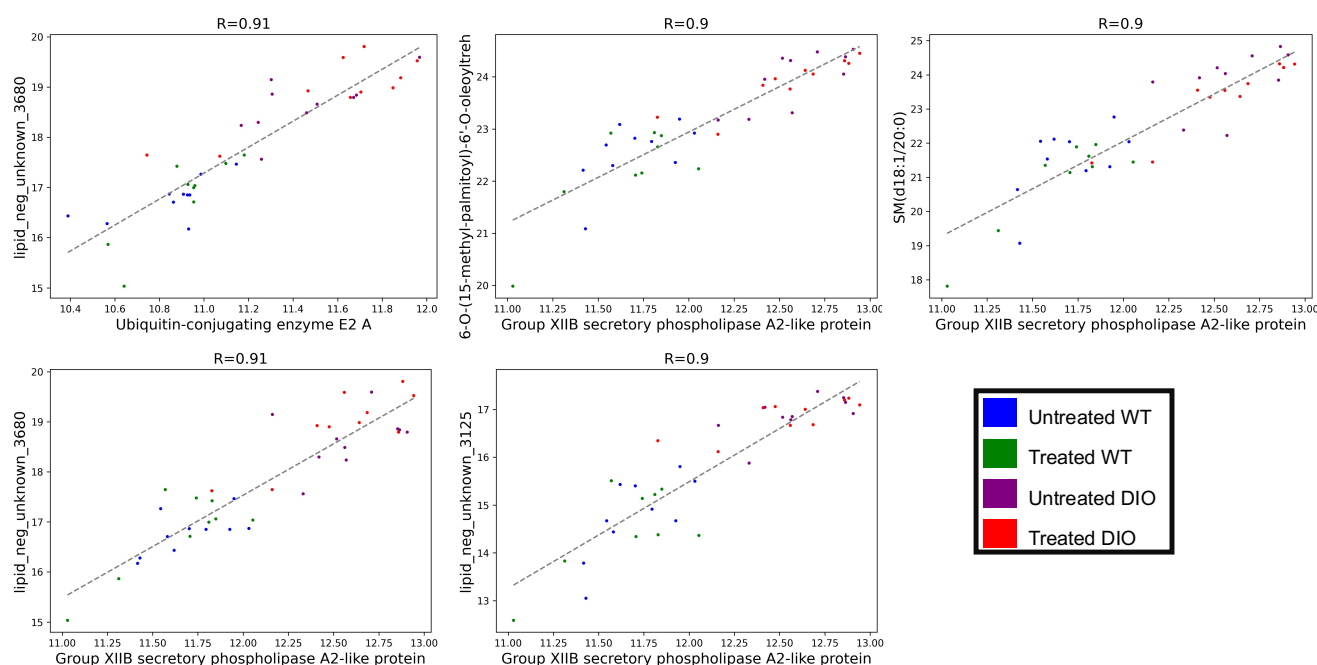


Figure 15: Dysregulated lipid metabolism module. The levels of dysregulated lipid proteins and metabolites are shown for each metabolite/lipid pair in the scatter plots above. The Pearson correlation of each pair is given by the R value. Each dot is a sample. Samples are colored according to the experimental group.

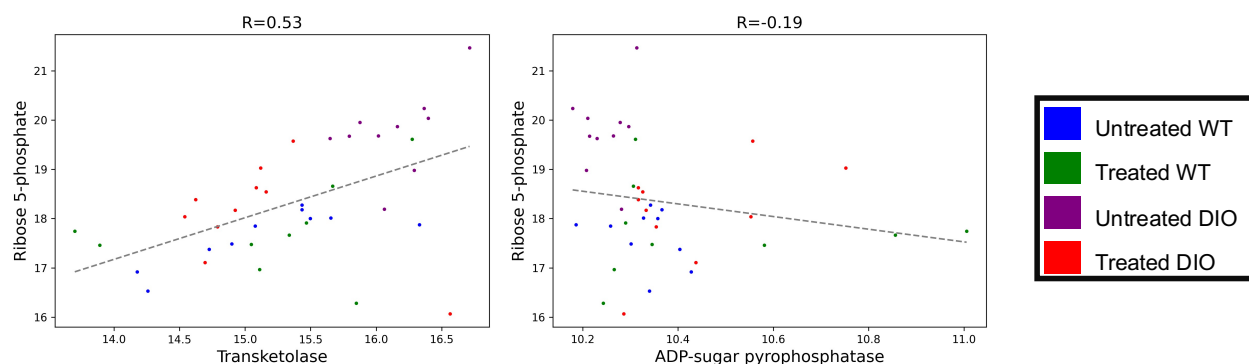


Figure 16: Dysregulated pentose phosphate pathway module. The levels of dysregulated pentose phosphate pathway proteins and metabolites are shown for each metabolite/lipid pair in the scatter plots above. The Pearson correlation of each pair is given by the R value. Each dot is a sample. Samples are colored according to the experimental group.

We recommend follow-on experiments analyzing tumor tissue to determine if the alterations in lipid metabolism and the pentose phosphate pathway are driven by changes in tumor metabolism or is a consequence of reduced tumor burden. Additionally, targeted approaches that use isotope tracers to quantitative the metabolic flux of the oxidative and non-oxidative branches of the pentose phosphate pathway would provide a direct read out of activity.

Appendix

Supplementary Files

1. SupplementaryTables.xlsx: excel file containing the detailed results of the statistical analyses.
2. Plots.zip: high-resolution images for all figures in this report as well as scatterplots showing the abundance of correlated protein and metabolites across all experimental conditions.