Article

MOMMI-MP: A Comprehensive Database for Integrated Analysis of Metabolic and Microbiome Profiling of Mouse Pregnancy

Kaustubh Pachpor [1], Medha Priyadarshini [2], Derek Reiman [3], Brian Layden [2,4], and Yang Dai [1,5]

[1] Department of Biomedical Engineering, University of Illinois at Chicago, Chicago, USA; kpachp2@uic.edu
[2] Division of Endocrinology, Diabetes, and Metabolism, University of Illinois, Chicago, USA; mpriya2@uic.edu
[3] Toyota Technological Institute at Chicago, Chicago, USA;
[4] Jesse Brown Veterans Affair Medical Center, Chicago, USA; blayde1@uic.edu
[5] Center for Bioinformatics and Quantitative Biology, University of Illinois at Chicago, Chicago, USA; yangdai@uic.edu

Abstract: Pregnancy is a dynamic state with multiple metabolic changes occurring including insulin resistance. Gestational diabetes mellitus (GDM), a form of diabetes that appears during pregnancy, develops if metabolic aberrations occur, in particular, in normal pregnancy-induced insulin resistance. Multi-omics is a powerful approach for uncovering the mechanisms driving metabolic change in different physiologic and pathologic states. A recent study demonstrated that the gestational gut microbiome mediates pregnancy metabolic adaptations through effects on gut indoleamine-2,3 dioxygenase 1 activity and the production of kynurenine. Using the dataset generated from this highly controlled study, we performed a comprehensive analysis of the pregnancy-specific physiological and metabolic profiles, 16S rRNA microbiome, and plasma untargeted LC-MS metabolome data. To facilitate the utilization of these analysis results by other researchers, we developed MOMMI-MP, a database that provides an easy-to-use platform to browse and search differential abundant microbial taxa and metabolites, and to examine metabolic pathways. The datasets consist of data collected from 3 genetically diverse strains of mice (C57BL/6J, CD1, and NIH-Swiss) over 6 time points during the gestational (days 0, 10, 15, and 19 during gestation) and postpartum (days 3 and 20 after delivery) states, totaling 180 samples for each strain. The computational results are presented in various tables and plots, and organized in MOMMI-MP to empower exploratory analyses by other researchers. In conclusion, MOMMI-MP is a resource to facilitate the investigation of novel mechanisms governing metabolic changes during pregnancy.

Keywords: Gestational diabetes mellitus; Insulin resistance; Microbiome; Metabolome; Database; Differential abundance; Correlation; Pathways

1. Introduction

Gestational diabetes mellitus (GDM) is the most common metabolic disorder during pregnancy, affecting 2% to 10% of pregnancies in the United States [1]. Up to 60% of those who have GDM carry a risk of developing type 2 diabetes later in life [2]. GDM is associated with resistance to the action of insulin on glucose uptake and utilization [3]. Interestingly, a typical pregnancy does develop a degree of insulin resistance in the later states as a normal adaptive response [3].

Using the systems biology approach, we recently performed an extensive metabolic, gut microbial, and metabolomic characterization, which revealed a novel gut microbial-host metabolic pathway mediating gestational responses. Specifically, we demonstrated

 doi:10.20944/preprints202212.0378.v1

that the gestational gut microbiome mediates pregnancy metabolic adaptations through effects on the gut indoleamine-2,3 dioxygenase 1 activity and the production of kynurenine [4]. This multi-omics study collected pregnancy-specific physiological and metabolic response characteristics, gut microbiome, and plasma metabolome data on 3 genetically diverse strains of mice (C57BL/6J, CD1, and NIH-Swiss) during the gestational stage and postpartum pregnancy. Since the insulin resistance of normal pregnancy is multifactorial, the datasets obtained from this highly controlled study may provide additional insights into the interactions among metabolic response, gut microbes, and the metabolites mediating these effects. To facilitate the utilization of these datasets by other researchers, we developed MOMMI-MP, a database for providing comprehensive analysis results of the Multi-omics Metabolic & Microbiome Profiling of Mouse Pregnancy (Fig. 1).

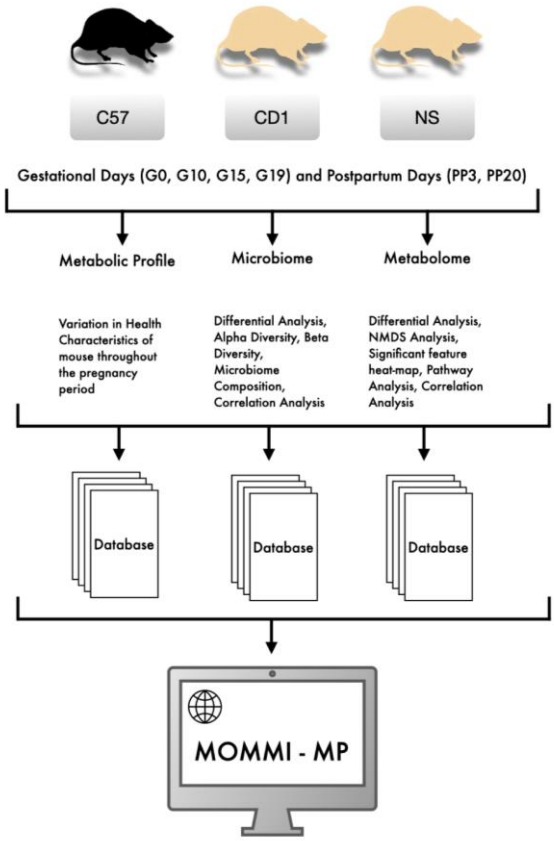


Figure 1. Structural outline of MOMMI-MP

2. Materials and Methods

2.1. Datasets and Pre-processing

The datasets consist of mouse metabolic health characteristics, microbiome, and metabolomic profiles measured at 4 time points during gestation (days G0, G10, G15, and G19) and 2 time points postpartum (days PP3, and PP20) from 3 mouse strains (C57BL/6J, CD1, and NIH-Swiss) with 10 mice at each time point per strain. The 16S rRNA gene amplicon sequencing and untargeted LC-MS were used to generate microbiome and metabolome profiles, respectively. Detailed experimental protocols for data collection and quantification procedures for microbiome and metabolomic features can be found in [4]. Briefly, the metabolic health characteristics are comprised of blood glucose and insulin levels, body weight, and change in weights of different types of adipose depots (subcutaneous and visceral) at the different time points of pregnancy. The microbiome data are provided with Operational Taxonomic Unit (OTU) abundance at 7

 doi:10.20944/preprints202212.0378.v1

taxonomic ranks (i.e., kingdom, phylum, class, order, family, genus, and species). The data are available at [5]. The untargeted LC-MS analysis generated metabolite feature abundance. The metabolomics data are available at [6].

The microbiome data consists of the OTU table as well as the amplicon sequence variant (ASV) table. The abundance obtained for an individual sample at each taxa level was used for further analysis. For the metabolome data, features that were present in less than a set threshold value (<20% of samples) were removed and scaled to remove unwanted variation caused by the features. Finally, data were log-transformed before the analysis. The summaries of the three data types are shown in Table 1 and Table 2

Table 1. Number of Samples and features for microbiome and metabolome data sets

Strain	# of Samples	# Metabolic Characteristics	# ASV	# Metabolite Features	
				Before filtering	After filtering
C57	60	12	1522	43682	4784
CD1	60	12	1522	43682	3340
NS	60	12	1522	43682	3137

Table 2. Number of Microbes obtained after pre-processing.

	C57BI6/J	CD1	NIH-Swiss
Class	20	20	20
Order	36	36	36
Family	60	60	60
Genus	92	92	92
Species	101	101	101

2.2. Alpha Diversity and Beta Diversity of Microbiome Data

Alpha diversity was computed by rarifying the dataset by the even depth method and functions such as 'ggplot', 'plot_richness', and 'estimate richness' in the phyloseq package [7]. 'Strain' attribute of the dataset was used to perform the analysis. Two methods, 'Observed' and 'Shannon', were used to display the results of the same. Beta diversity was obtained using two different methods, Principal Coordinate Analysis (PCoA) and Non-metric Multidimensional Scaling (NMDS). The dataset was rarefied and the binary Jaccard metric was used to perform the NMDS analysis, whereas the Bray-Curtis dissimilarity was used to plot the ordination visualization for the PCoA analysis.

2.3. Differential Abundance Analysis of Microbiome Data

 doi:10.20944/preprints202212.0378.v1

The package ANCOM-BC was used to perform the differential analysis [8]. It identifies taxa whose absolute abundances, per unit volume microbe, differ significantly throughout the pregnancy period. In this part of the analysis, the first gestational time point (G0) was set as the reference point, and the remaining time points were used individually along with the reference time point to perform the differential analysis. The result obtained was an adjusted p-value table that describes the significance of abundance difference between the comparison groups for individual taxa, where a threshold (0.05) was set to report taxa that are differentially abundant in at least one comparison group.

2.4. Differential Analysis of Metabolome Data

Samples at gestational time points for days 15 and 19 (G15 and G19) were grouped together and compared to all the remaining time points were grouped together. This grouping is based on the previous results that the metabolomic profiles at G15 and G19 distinguish those from the rest of the time points [4]. The PERMANOVA [9] analysis was performed on those two groups for every individual feature. The result obtained was an adjusted p-value table that describes the significance of separation between the groups for individual features. Metabolite features with adjusted p-values less than 0.05 are considered differentially abundant and are reported in the table.

2.5. Feature-Compound Mapping

First, the Mummichog algorithm implemented in the online tool MetaboAnalyst [10] was used to identify the KEGG pathways and the compound hits from the significantly abundant metabolite features identified from the differential analysis. The input to the tool was the adjusted p-value table obtained from the PERMANOVA analysis. The parameters were set in the following way: Mass tolerance was set to 10 ppm and retention time was set to 'seconds'. The tool's output includes a visual representation of the pathways along with multiple tables which gave the mapped compound IDs and significant hits for every pathway. Subsequently, the compound IDs along with the adjusted p-value table were used in the KEGG-Rest tool [11] to identify the compound names, which later were used to replace the unique metabolite feature IDs and filter out the significant features which are not linked to any compounds. The abundance table of the remaining significant features was then used to perform correlation analysis with the differentially abundant microbes.

2.6. Heatmap and NMDS Analysis

After obtaining the adjusted p-value table from the PERMANOVA analysis and mapping the features, a visualization was obtained using the heatmap to show the relationship between the time points and the obtained significant features. Also, NMDS analysis was performed using 'metaMDS' function from the 'vegan' library [12] to get the plots based on the time points for every mouse strain, which showed the grouping of multiple time points.

2.7. Correlation Analysis

The correlation analysis was performed, (1) between the significantly abundant microbes identified from ANCOM-BC and the health characteristics, (2) between those microbe pairs, (3) between the significantly abundant metabolites, (4) between the significantly abundant metabolites and the health characteristics, and (5) between the significantly abundant microbes and metabolites. In all correlation analyses, Pearson correlations were calculated.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 December 2022 doi:10.20944/preprints202212.0378.v1

2.8. Pathway Analysis

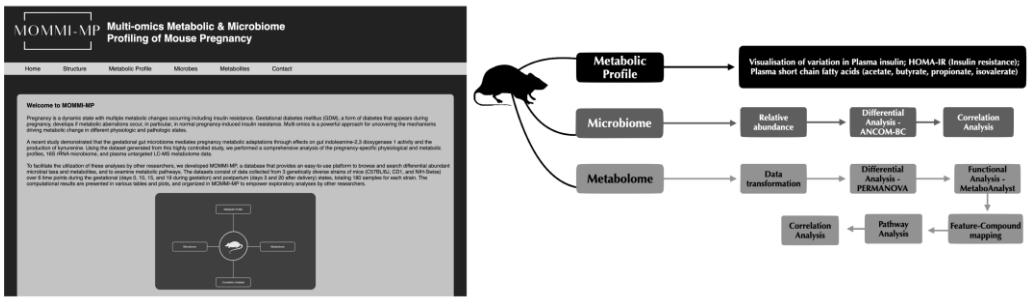
Pathway analysis was performed using the Mummichog algorithm in the MetaboAnalyst tool using the identified metabolites (i.e., compounds) obtained following the analysis procedure described in 2.6. The input consists of the compound names and their respective adjusted p-values. The pathway analysis gave a visual bar plot of the mapped features and the gamma values, which represent adjusted p-values from the enrichment analysis.

3. Results

The MOMMI-MP database is composed of results from 1) statistical and bioinformatic analyses within each of the metabolic health profiles, microbiomes, and metabolomes, 2) correlation analyses among metabolic health characteristics, microbes, and metabolites, and 3) metabolic pathway enrichment analyses. The database is designed to allow easy browsing and query of differentially abundant microbial taxa and metabolites; and the affected metabolic pathways.

3.1. MOMMI-MP and Structure

The results were organized on an online server in the form of a database website MOMMI-MP (URL) (Fig. 2a). The website's home screen has a visual representation of the basic structure of the database. Users can explore the results of individual data types by clicking the menu bar at the top of the website as well. The "Structure" connects to a page that presents the flowchart of analysis performed in the database (Fig. 2b). Other tabs, such as "Metabolic Profiles, "Microbiome" and "Metabolome", include links to the respective analysis results.



(a) **(b)**

Figure 2. MOMMI-MP Assembly. **(a)** MOMMI-MP homepage. The figure shows the pictorial representation of the basic blocks of the platform. **(b)** Flow of MOMMI-MP. Every block in the flowchart represents a method/operation which was performed on the respective data set.

3.2. Metabolic Profile

This webpage consists of a visual representation of the 3 mouse strains. Here, users can click on any mouse strain to access the plots for individual characteristics in the metabolic profiles. Fig. 3a-3c show the boxplots for subcutaneous fat at each time point for the three mouse strains as an example.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 December 2022 doi:10.20944/preprints202212.0378.v1

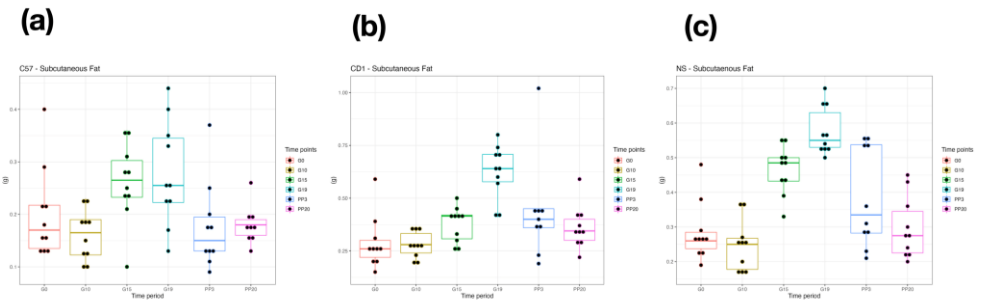


Figure 3. Metabolic Profile. The plots represent the variation in subcutaneous fat (one of the mouse health characteristics) in all three strains. **(a)** C57BI6/J; **(b)** CD1; **(c)** NIH-Swiss.

3.3. Microbes

The 'Microbe' webpage contains links to various web pages of analysis results. For example, users can examine the Alpha diversity (Fig. 4) and Beta diversity results (Fig. 5). They can also view microbiome compositions for all the taxonomical levels for each strain. Fig. 6a-6c show the microbiome compositions of C57BL/6J at Class, Family, and Genus levels, respectively.

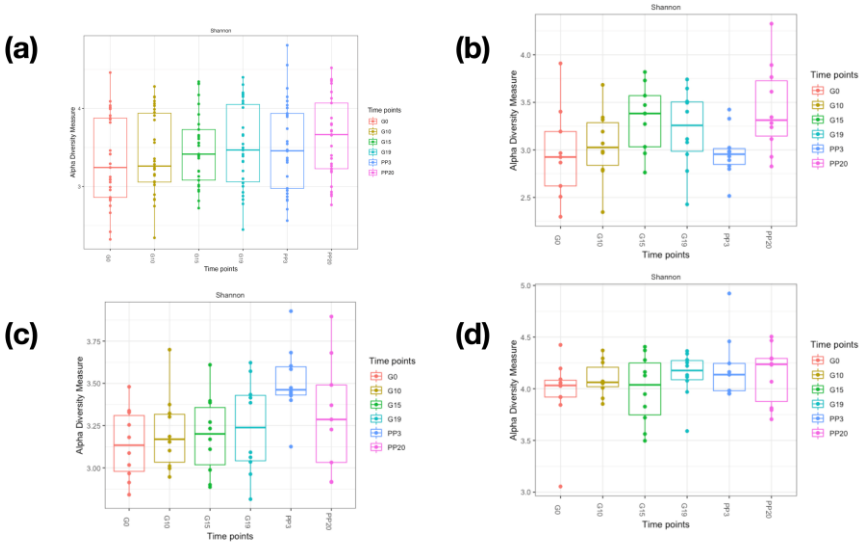
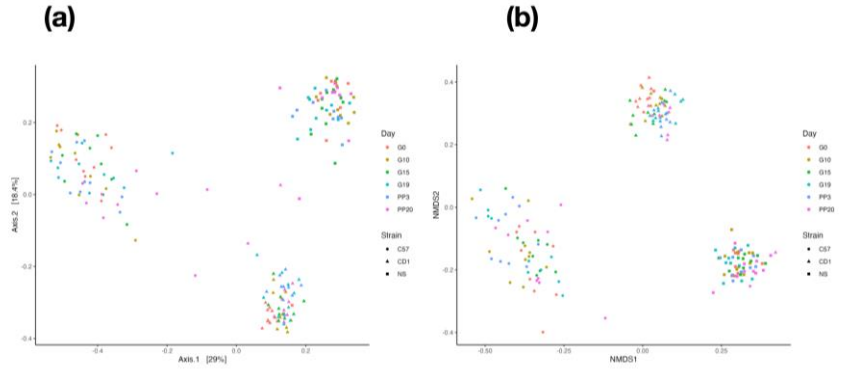


Figure 4. Alpha (Shannon) diversity. **(a)** All Strains; **(b)** C57BL/6J; **(c)** CD1; **(d)** NIH-Swiss.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 December 2022 doi:10.20944/preprints202212.0378.v1



(a) (b)

Figure 5. Beta Diversity: **(a)** PCoA (all strains); **(b)** NMDS (all strains).

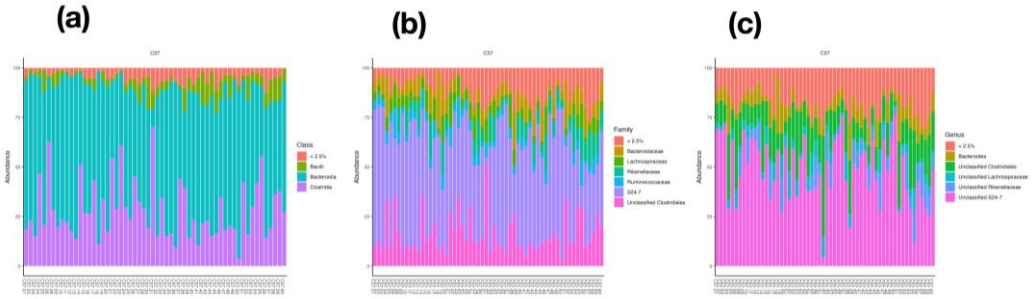


Figure 6. Microbiome composition for different taxonomical levels. **(a)** Class; **(b)** Family; **(c)** Genus.

The ANCOM-BC analysis was performed at each taxonomic level of Class, Order, Family, Genus, and Species. The numbers of identified differentially abundant taxa are summarized in Table 3.

Table 3. The number of taxa obtained from ANCOM-BC for each taxonomic level for each strain.

	Class	Order	Family	Genus	Species
C57	9	10	14	6	6
CD1	8	10	17	5	7
NS	4	6	8	6	8

A table of the adjusted p-values can be viewed for the identified microbes from another webpage. Users can interact with the table by sorting it or searching for any specific microbe. Each microbe in the table is linked to the respective boxplots of abundance (Fig. 7). The webpage also consists of a link redirecting the users to another webpage providing results of correlation analysis of microbe-microbe, microbe-health characteristic, and microbe-metabolite, which will be described later.

 doi:10.20944/preprints202212.0378.v1

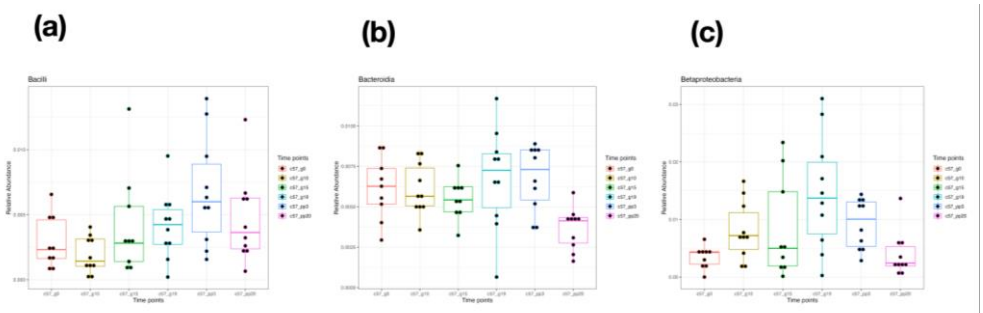


Figure 7. Differential Abundance Analysis for C57BL/6J. **(a)** Bacilli; **(b)** Bacteroidia; **(c)** Betaproteobacteria.

3.4. Metabolome

This webpage 'Metabolome' provides the analysis results obtained from the metabolome data. The grouping of mice was based on the metabolite features using the remaining features after the pre-processing (Fig. 8).

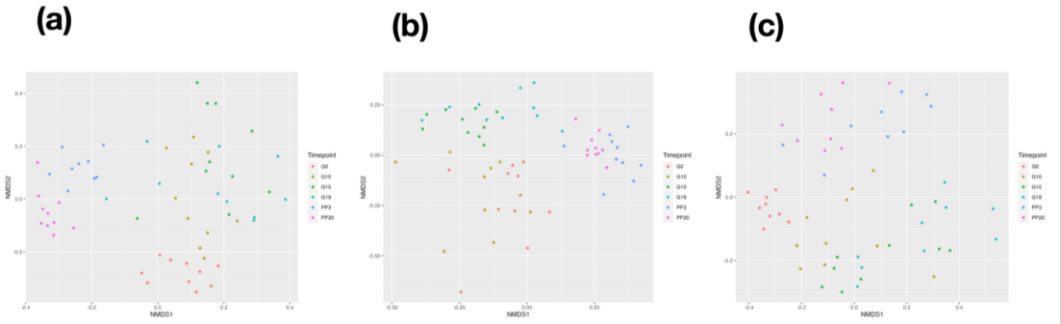


Figure 8. NMDS Analysis. **(a)** C57BL/6J; **(b)** CD1; **(c)** NIH-Swiss.

The heatmaps show the abundance of patterns of the significantly changed metabolite features identified from the PERMONOVA analysis (Fig. 9).

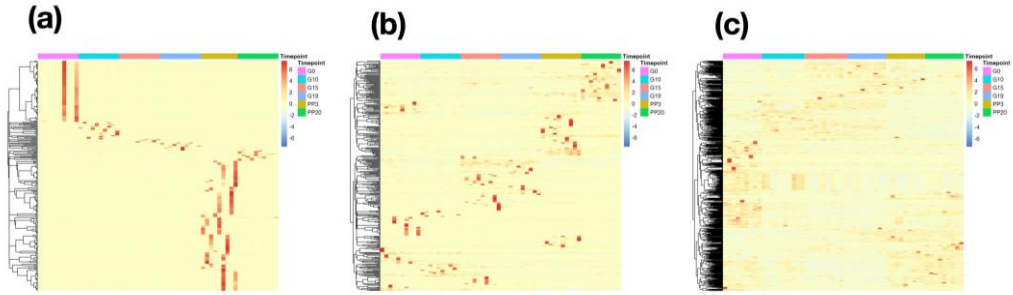
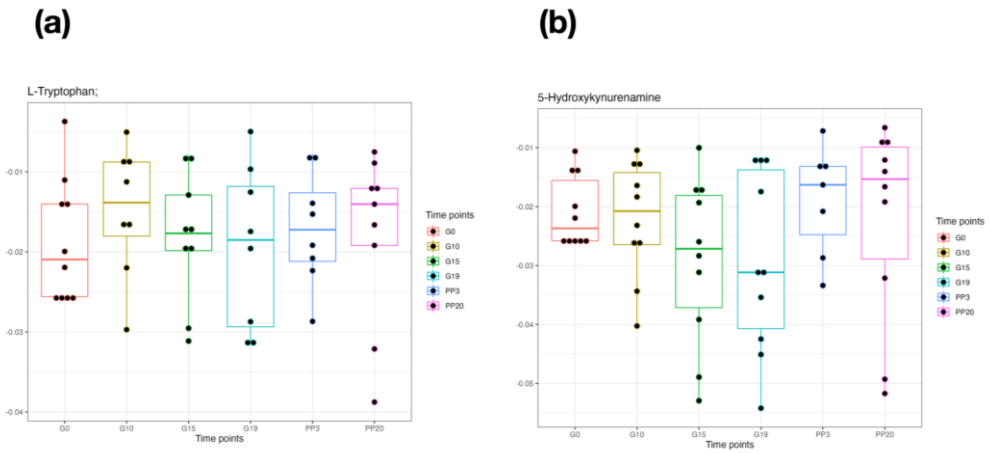


Figure 9. Heatmaps for significant metabolite features. **(a)** C57BL/6J; **(b)** CD1; **(c)** NIH-Swiss.

The webpage 'Differential Abundant Feature' consists of the obtained adjusted p-value table from the PERMANOVA analysis with features replaced by their corresponding compounds. Similar to the microbiome analysis, users can interact with the table by sorting or searching for any specific compound name. Also, there is a link embedded in each compound where users can view respective boxplots of the abundance over the time

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 December 2022 doi:10.20944/preprints202212.0378.v1

points (Fig. 10). The total numbers of the significant metabolite features obtained for all strains are given in Table 4.



(a) **(b)**

Figure 10. Differential Analysis for C57BL/6J. **(a)** L-Tryptophan; **(b)** 5-Hydroxykynurenamine.

Table 4. Numbers of differentially abundant metabolite features.

Strain	# Features	# Features mapped to known compounds
C57	4784	837
CD1	3340	645
NS	3137	529

Pathway analysis results after mapping the compounds to the features are uploaded on their respective webpage (Fig. 11). Also, correlation analysis between microbes and metabolites along with the correlation coefficients are displayed on an individual web page (Fig. 12).

 doi:10.20944/preprints202212.0378.v1

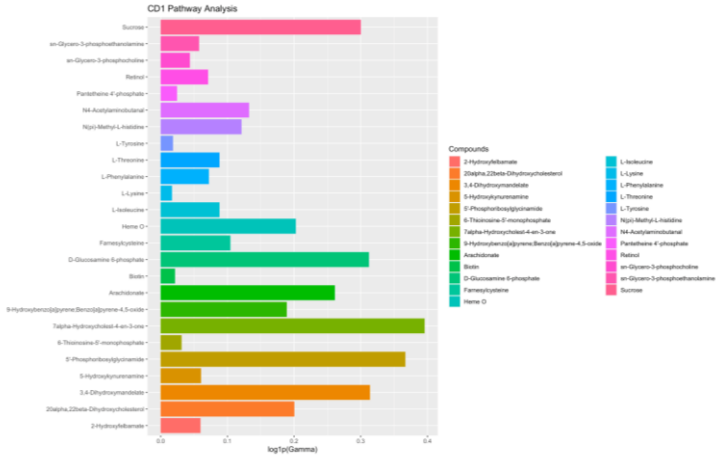


Figure 11. Pathway Analysis for C57BL/6J.

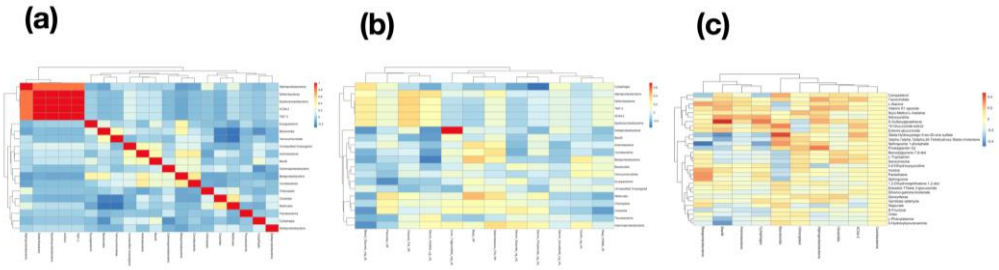


Figure 12. Correlation Analysis for C57BL/6J. **(a)** microbe-microbe; **(b)** microbe-metabolic characteristic; **(c)** metabolite-microbe.

4. Conclusions

The MOMMP-MP is a database that provides comprehensive analysis results of the metabolic health characteristics, microbiome, and metabolome data obtained from 3 genetically distinct strains of mice (C57BL/6J, CD1, and NIH-Swiss) during the gestational and postpartum stages. MOMMP-MP is an easy-to-use platform that facilitates explorations of the analysis results.

5. Acknowledgement

BTL acknowledges support from NIH R01DK104927-01A1, NIH P30DK020595, and VA merit 1I01BX003382-01-A1.

References

1. Gregory EC, Ely DM. Trends and Characteristics in Gestational Diabetes: United States, 2016-2020. Natl Vital Stat Rep. 2022 Jul;71(3):1-15. PMID: 35877134.
2. Angueira AR, Ludvik AE, Reddy TE, Wicksteed B, Lowe WL Jr, Layden BT. New insights into gestational glucose metabolism: lessons learned from 21st century approaches. Diabetes. 2015 Feb;64(2):327-34. doi: 10.2337/db14-0877. PMID: 25614666; PMCID: PMC487679
3. Sonagra AD, Biradar SM, K D, Murthy D S J. Normal pregnancy- a state of insulin resistance. J Clin Diagn Res. 2014 Nov;8(11):CC01-3. doi: 10.7860/JCDR/2014/10068.5081. Epub 2014 Nov 20. PMID: 25584208; PMCID: PMC4290225.

4. Priyadarshini M, Navarro G, Reiman DJ, Sharma A, Xu K, Lednovich K, Manzella CR, Khan MW, Garcia MS, Allard S, Wicksteed B, Chlipala GE, Szynal B, Bernabe BP, Maki PM, Gill RK, Perdew GH, Gilbert J, Dai Y, Layden BT. Gestational Insulin Resistance Is Mediated by the Gut Microbiome-Indoleamine 2,3-Dioxygenase Axis. Gastroenterology. 2022;162(6):1675-89.e11. Epub 20220113. doi: 10.1053/j.gastro.2022.01.008. PubMed PMID: 35032499. PMCID: PMC9040389.
5. EMBL-ENA ID: PRJEB45047."Gestational gut microbiome-IDO1 axis mediates pregnancy insulin resistance". https://www.ebi.ac.uk/ena/browser/view/PRJEB45047
6. MetaboLights ID: MTBLS3598. EMBL-ENA ID: PRJEB45047. "Gestational gut microbiome-IDO1 axis mediates pregnancy insulin resistance". https://www.ebi.ac.uk/metabolights/search
7. Paul J. McMurdie and Susan Holmes. Phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data; 2013; PLoS ONE 8(4):e61217.
8. Lin, H.; Peddada, S. D. Analysis of Compositions of Microbiomes with Bias Correction. Nat. Commun. 2020, 11 (1), 3514. https://doi.org/10.1038/s41467-020-17041-7.
9. Anderson, M. J. Permutational Multivariate Analysis of Variance (PERMANOVA). In Wiley StatsRef: Statistics Reference Online; 2017; pp 1–15. https://doi.org/10.1002/9781118445112.stat07841.
10. Pang, Z., Zhou, G., Ewald, J. et al. Using MetaboAnalyst 5.0 for LC–HRMS spectra processing, multi-omics integration and covariate adjustment of global metabolomics data. Nat Protoc 17, 1735–1761 (2022). https://doi.org/10.1038/s41596-022-00710-w.
11. Dan Tenenbaum and Bioconductor Package Maintainer (2021). KEGGREST: Client-side REST access to the Kyoto Encyclopedia of Genes and Genomes (KEGG). R package version 1.32.0.
12. Jari Oksanen, Gavin L. Simpson, F. Guillaume Blanchet, Roeland Kindt, Pierre Legendre, Peter R. Minchin, R.B. O'Hara, Peter Solymos, M. Henry H. Stevens, Eduard Szoecs, Helene Wagner, Matt Barbour, Michael Bedward, Ben Bolker, Daniel Borcard, Gustavo Carv. https://CRAN.R-project.org/package=vegan.