

Supplementary material

Main Text: “Meta-analysis on metabolites involved in bioenergetic pathways reveals a pseudohypoxic state in Down syndrome”

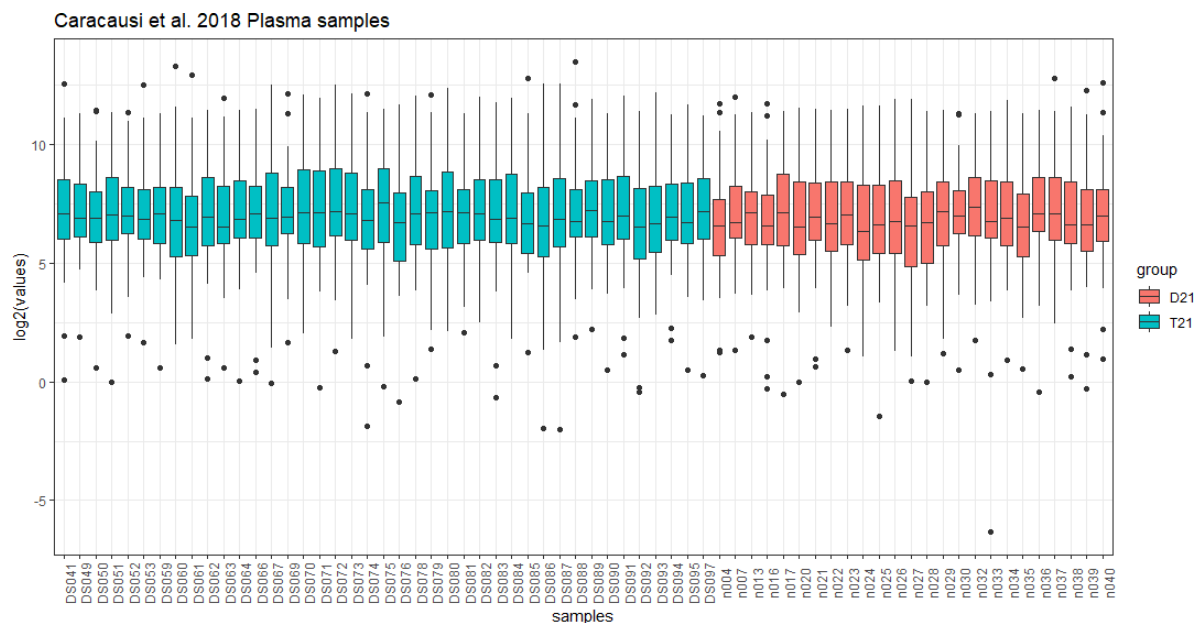
Laszlo Pecze, Elisa B. Randi and Csaba Szabo

1. Log2 transformation and normalization on raw datasets

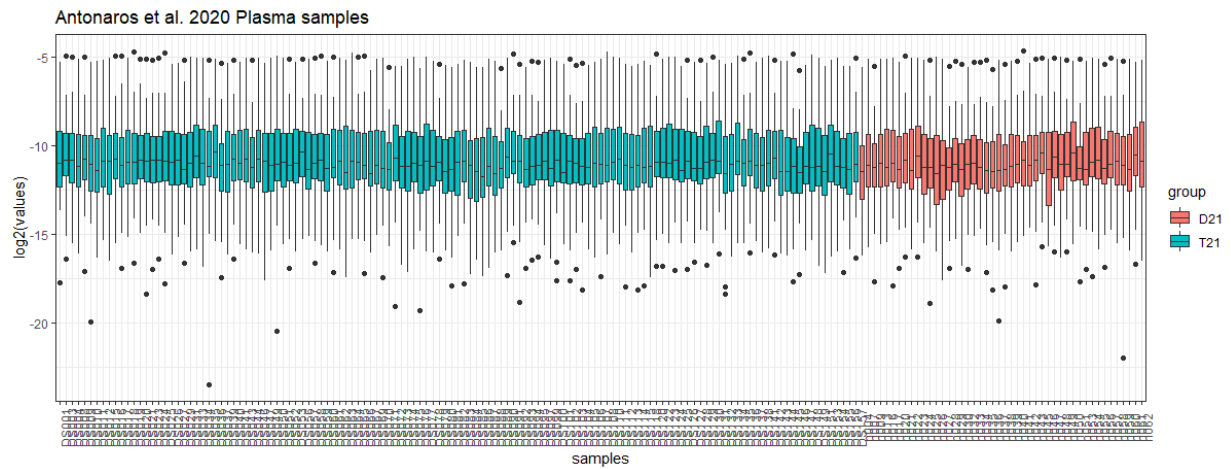
Caracausi et al. (2018) and Antonaros et al. (2020) used the unnormalized raw data for the evaluation of metabolic profile in plasma from DS and control children. Although there are no significant variations between samples (Supplementary Figs. 1-2), it cannot be excluded that a normalization procedure would have a beneficial effect. Nevertheless, we left it unnormalized for our meta-analysis. Caracausi et al. also reported metabolite levels in urine from DS and control children. They normalized the raw data according to PQN method (Dieterle et al., 2006). After normalization data show only small variability between samples (Supplementary Fig. 3).

Powers et al. (2019) used the removeBatchEffect function from R limma package (Ritchie et al., 2015) to normalize the raw data. The result of the normalization is shown in Supplementary Fig.4.

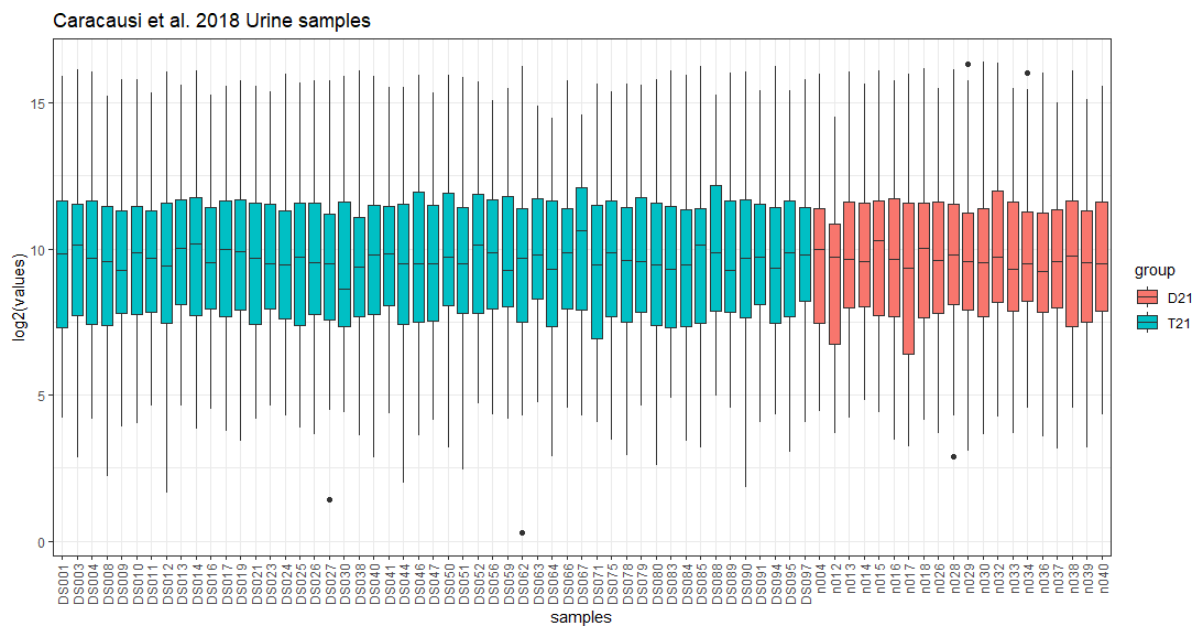
Culp-Hill et al. (2017) provided two datasets one is from targeted experiment and one from untargeted one. The targeted dataset does not show significant variation between samples. See Supplementary Fig. 5. However, there are significant variations between samples on the nontargeted metabolic result. See Supplementary Fig 6. Using a Wilcoxon non-parametric test on the average of the samples we have found that there is a significant difference between DS and control. (p -value < 0.0001). This means that, if the data are used without normalization, each component will show significantly higher values for DS group. Unfortunately, this is the case in the published article. Therefore, we used variance stabilization normalization (Huber et al., 2002) to implement these data into our meta-analysis. (Supplementary Fig. 7).



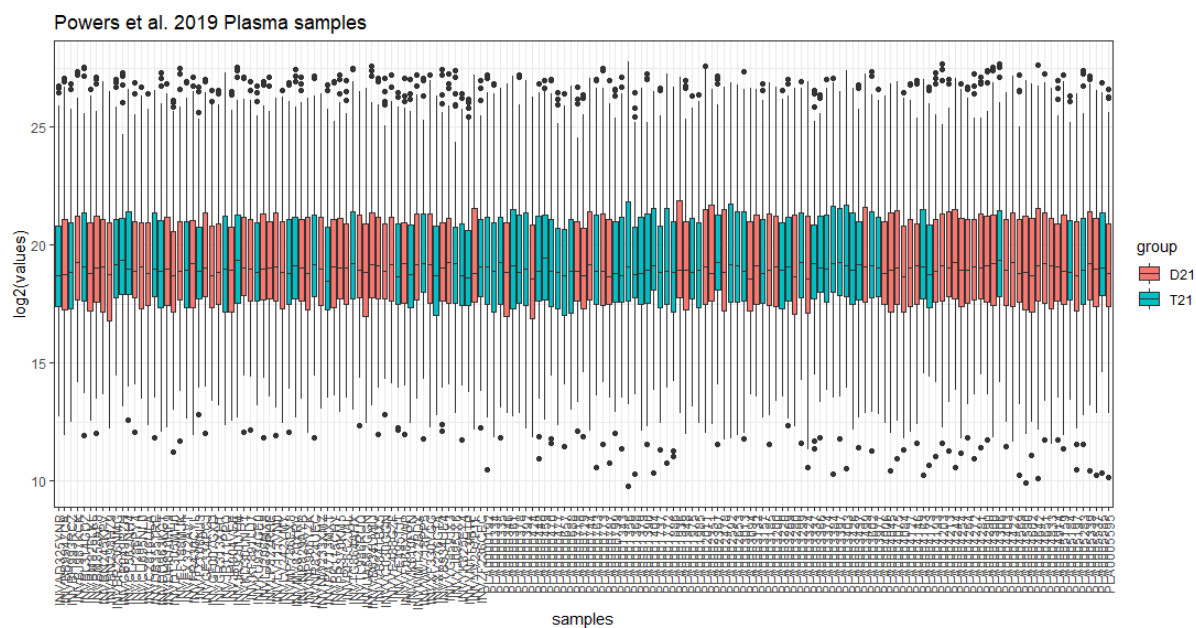
Supplementary Fig.1. Log2 transformation and normalization on raw datasets. The boxplot shows the distribution of plasma samples after log2 transformation. Data are from Caracausi et al., 2018.



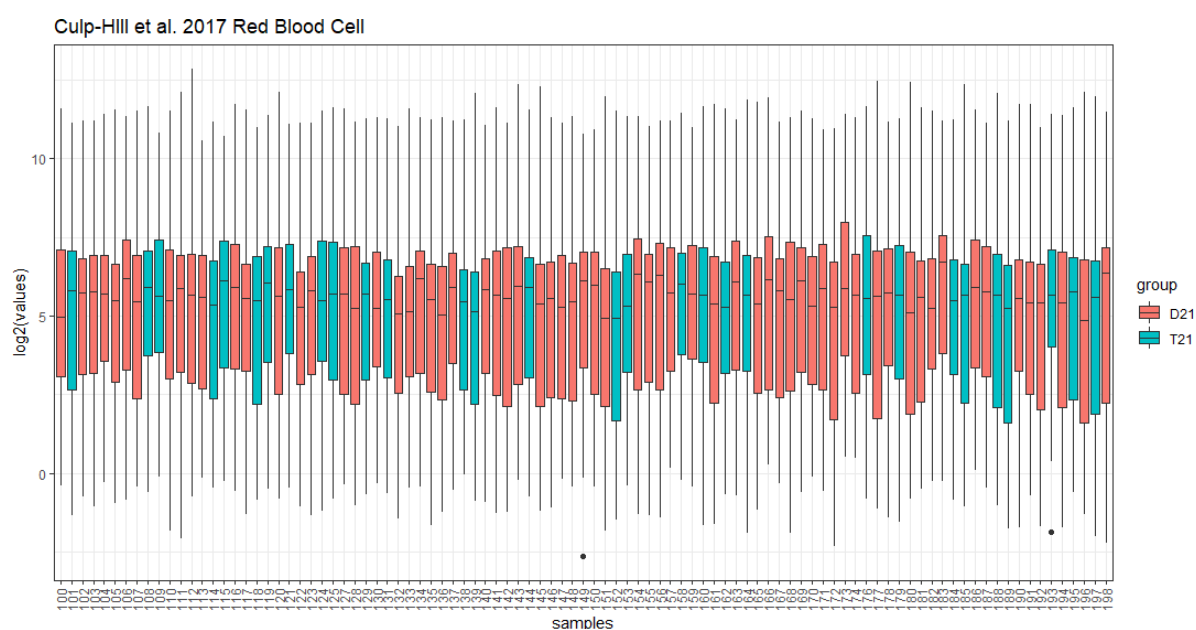
Supplementary Fig.2. Log2 transformation and normalization on raw datasets. The boxplot shows the distribution of plasma samples after log2 transformation. Data are from Antonaros et al., 2020.



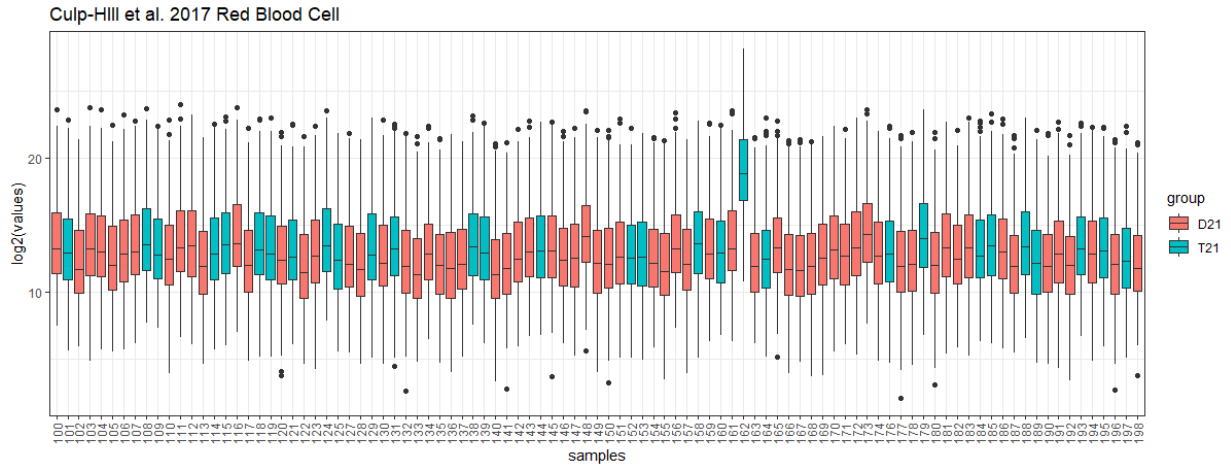
Supplementary Fig. 3. Boxplot showing the distribution of plasma samples after log2 transformation. The authors (Caracausi et al., 2018) used Probabilistic Quotient Normalization (PQN).



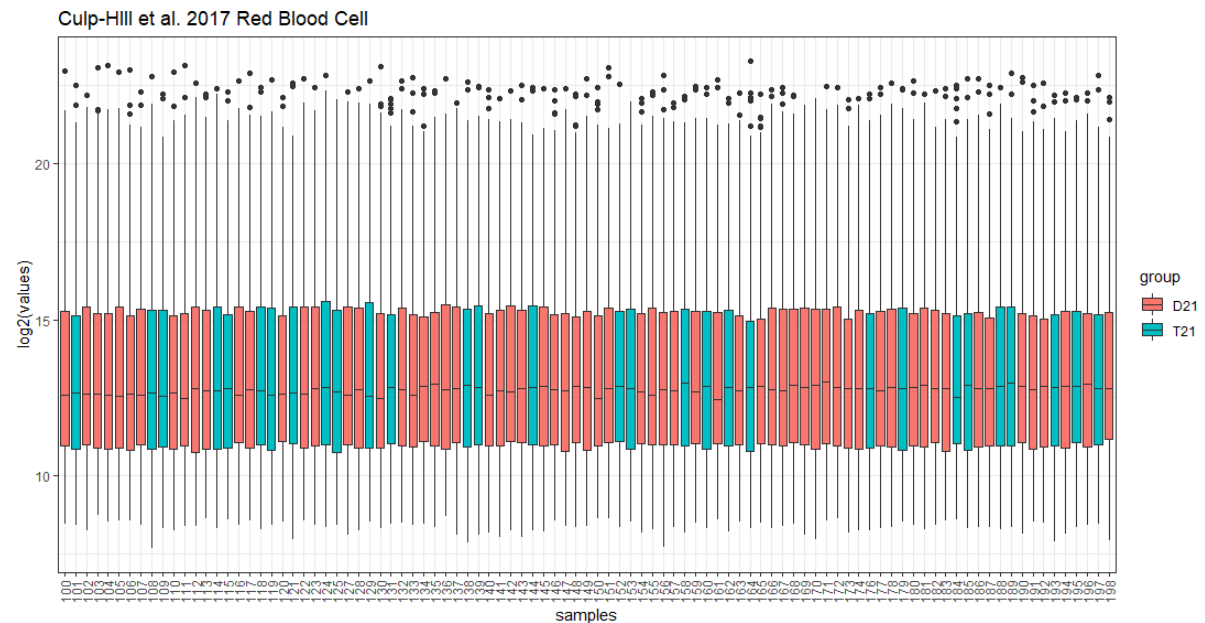
Supplementary Fig.4. Boxplot showing the distribution of plasma samples after log2 transformation. Normalized and log2 transformed samples from the study of Powers et al., 2019.



Supplementary Fig. 5. Targeted metabolomic result from Culp-Hill study (Culp-Hill et al., 2017).



Supplementary Fig. 6. Untargeted metabolomic result from Culp-Hill study (Culp-Hill et al., 2017) showing significant technical variations.



Supplementary Fig.7. Untargeted metabolomic result from Culp-Hill study (Culp-Hill et al., 2017). We applied variance stabilization normalization to reduce technical variations.

2. Impact of fasting state on metabolic levels

Caracausi and colleagues (Caracausi et al., 2018) provided information on the fasting state of subjects i.e breakfast has been taken before sample collection or not. In order to know the impact of fasting on the levels of metabolites in plasma, we performed a multivariate linear regression, where the amount of a given metabolite is the dependent variable while Karyotypes (T21/D21), Gender (Male/Female), Age and Fasting (Yes/No) are the independent variable. If the regression coefficient statistically differs from zero, i.e. there is an effect, the regression coefficient was put into a table with its 95% confidence intervals. The result is shown in the Supplementary Table 1. We can conclude that fasting has an impact on the levels of glucose and the ketone bodies, but it has no significant effect on other metabolites. We also checked the impact of fasting on metabolite levels using the same multivariate regression model. Data are shown on Supplementary Table 2. In general, we can conclude that accounting for sex, age, and fasting state did not significantly affect the main result.

Compound	KaryotypeT21 β [95% CI]	FastingYes β [95% CI]	Age β [95% CI]	SexM β [95% CI]
Leucine	NS	NS	0.01 [0, 0.02]	0.17 [0.01, 0.32]
Isoleucine	NS	NS	0.01 [0, 0.03]	NS
Valine	NS	NS	NS	NS
Lactate + Threonine	NS	NS	-0.02 [-0.04, 0]	NS
Alanine	NS	NS	NS	NS
Acetate	0.56 [0.15, 0.96]	NS	NS	NS
Pyruvate	0.24 [0.04, 0.44]	NS	-0.02 [-0.03, -0.01]	NS
Acetone	0.5 [0.08, 0.92]	0.61 [0.14, 1.08]	NS	NS
Glutamine	NS	NS	NS	NS
Citrate	NS	NS	NS	NS
Glycine	NS	NS	NS	NS
Creatine	NS	NS	-0.05 [-0.08, -0.02]	NS
Creatinine	NS	NS	0.06 [0.04, 0.07]	NS
Lactate	NS	NS	-0.02 [-0.04, 0]	NS
Glucose	NS	-0.1 [-0.2, -0.01]	NS	NS
Mannose	NS	NS	NS	NS
Tyrosine	-0.21 [-0.41, 0]	NS	NS	NS
Histidine	NS	NS	NS	NS
Phenylalanine	NS	NS	NS	NS
Formate	0.47 [0.18, 0.76]	NS	NS	NS
Fumarate	NS	NS	NS	NS
Threonine	NS	NS	NS	NS
Glycerol	0.9 [0.19, 1.62]	NS	NS	NS
Lysine	NS	NS	NS	NS
Acetatoacetate	NS	NS	NS	NS
Methionine	NS	NS	NS	NS
Unk1	NS	0.28 [0.05, 0.51]	NS	NS
Unk2	NS	NS	NS	NS
Unk3	NS	NS	NS	NS
3-hydroxybutyrate	NS	0.87 [0.15, 1.58]	-0.05 [-0.09, -0.01]	NS
Succinate	0.6 [0.34, 0.86]	NS	NS	NS
2-hydroxybutyrate	NS	NS	NS	NS
Glutamate	NS	NS	-0.08 [-0.14, -0.01]	NS

Supplementary Table 1. Significant regression coefficients from a multiple linear regression model applied on each metabolite found in plasma. The raw data are from Caracausi's study (Caracausi et al., 2018).

Compound	KaryotypeT21 β [95% CI]	Fasting Yes β [95% CI]	Age β [95% CI]	SexM β [95% CI]
Fumarate	NS	NS	-0.01 [-0.03, 0]	NS
DMG	NS	NS	NS	NS
Allantoin	NS	NS	NS	NS
Acetone	0.45 [0.07, 0.82]	NS	NS	NS
Isoleucine	NS	NS	NS	NS
Leucine	NS	NS	NS	NS
Valine	NS	NS	NS	NS
Alanine	NS	NS	NS	NS
Asparagine	NS	NS	NS	NS
Creatine + Creatinine	NS	NS	0.04 [0.02, 0.05]	NS
Citrate	NS	NS	NS	-0.59 [-1.03, -0.15]
Glycine	-0.48 [-0.94, -0.02]	NS	NS	NS
Ethanolamine	NS	NS	0.01 [0, 0.03]	NS
4-hydroxyphenylacetate	NS	NS	NS	NS
3-hydroxyisovalericacid	NS	NS	NS	NS
Dimethylamine	NS	NS	NS	NS
Trimethylamine-N-oxide	0.79 [0.3, 1.27]	NS	NS	NS
Taurine	NS	NS	NS	NS
Creatine	NS	NS	-0.18 [-0.25, -0.1]	NS
Glycolate	NS	NS	NS	NS
Creatinine	NS	NS	0.05 [0.04, 0.07]	NS
Tyrosine	0.21 [0.01, 0.42]	NS	NS	NS
Hippurate	NS	NS	NS	NS
Formate	-0.53 [-1.01, -0.04]	NS	-0.04 [-0.07, -0.01]	NS
Trigonelline	NS	NS	NS	NS
1-methylnicotinamide	NS	NS	NS	NS
Phenylacetylglutamine	0.44 [0.03, 0.85]	NS	NS	NS
2-hydroxyisobutyricacid	NS	NS	NS	NS
Glutamate + Glutamine	NS	0.08 [0, 0.17]	-0.01 [-0.02, -0.01]	NS
Lysine	-0.21 [-0.38, -0.03]	NS	NS	NS

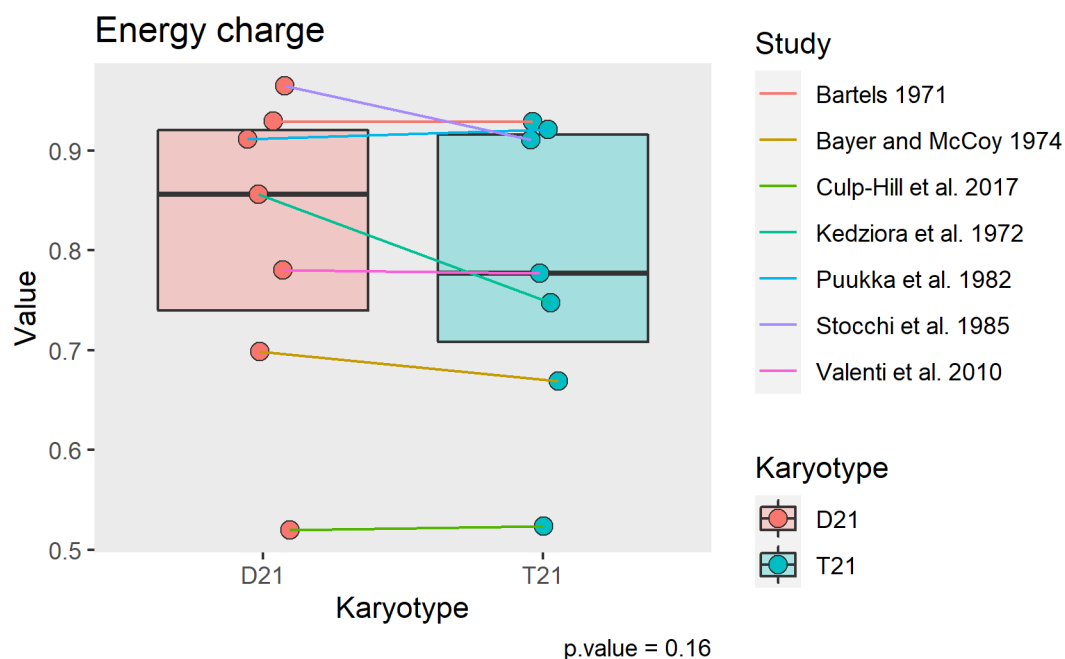
Supplementary Table 2. Significant regression coefficients from a multiple linear regression model applied on each metabolite found in urine. The raw data are from Caracausi's study (Caracausi et al., 2018).

3. Energy charge

The energy charge is related to ATP, ADP and AMP concentrations to account for the energy status in metabolism (Atkinson and Walton, 1967). Energy charge is defined as:

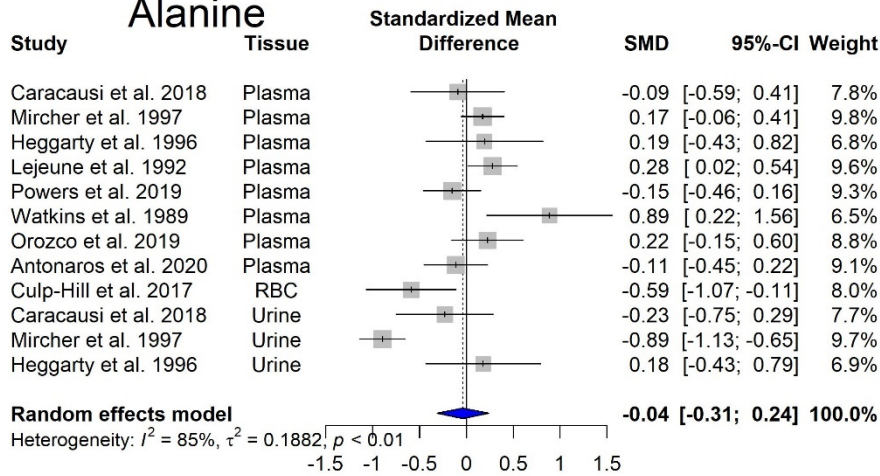
$$\text{Energy charge} = ([\text{ATP}] + 1/2[\text{ADP}]) / ([\text{ATP}] + [\text{ADP}] + [\text{AMP}])$$

The energy charge of most cells varies between 0.7 and 0.95 (Atkinson and Walton, 1967). The energy charge was calculated DS and CTR cells, if the study provides information for all the metabolites i.e. ATP, ADP, and AMP (Supplementary Fig. 8). Although the cells from Down syndrome subjects are less charged, the change did not reach the significance level at $p > 0.05$ (using a paired t-test).

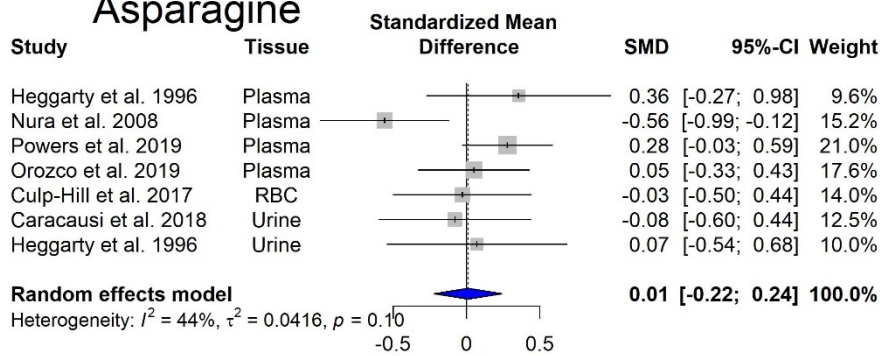


Supplementary Fig. 8. Energy charge values in DS cells vs. control.

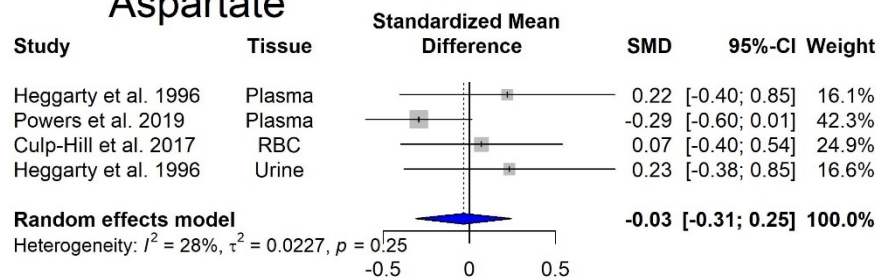
Alanine



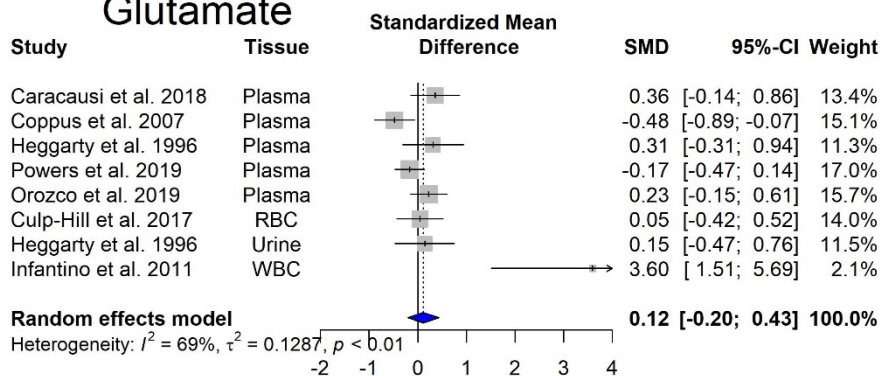
Asparagine



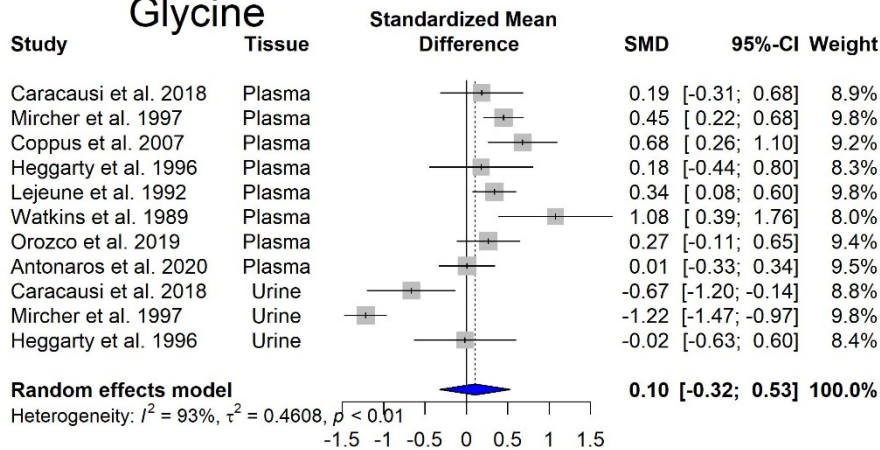
Aspartate



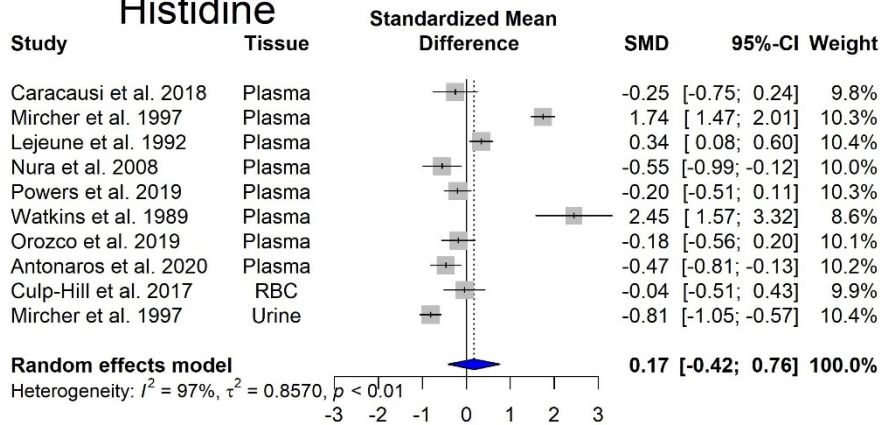
Glutamate



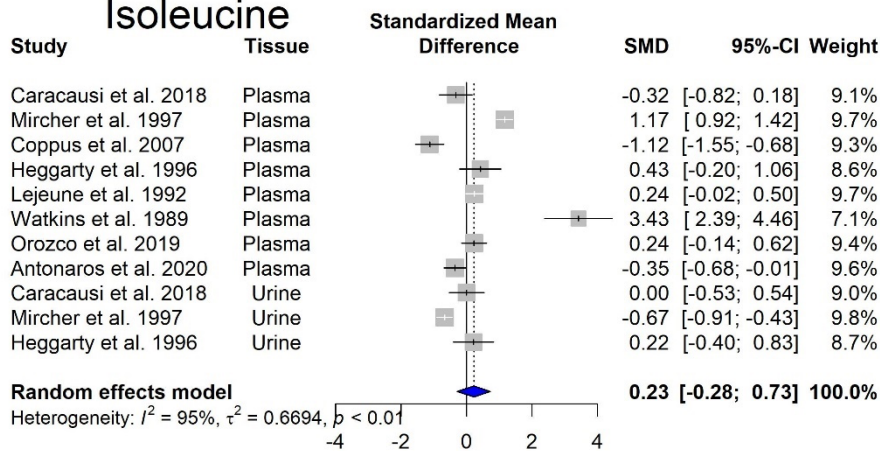
Glycine

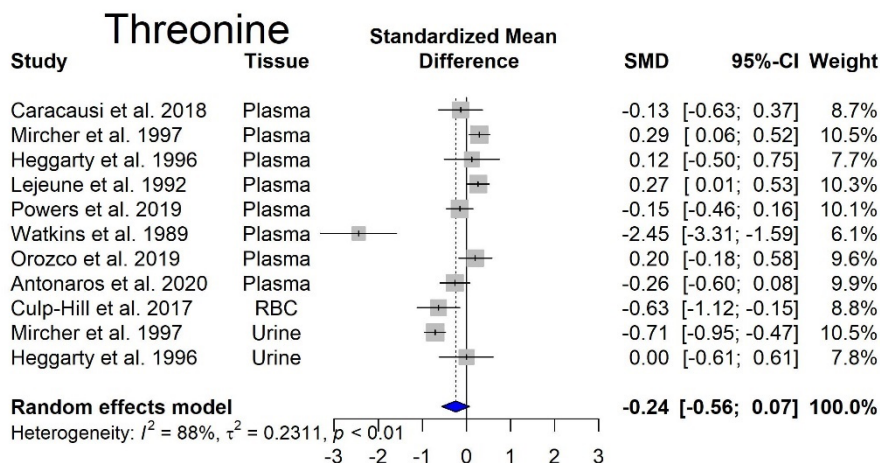
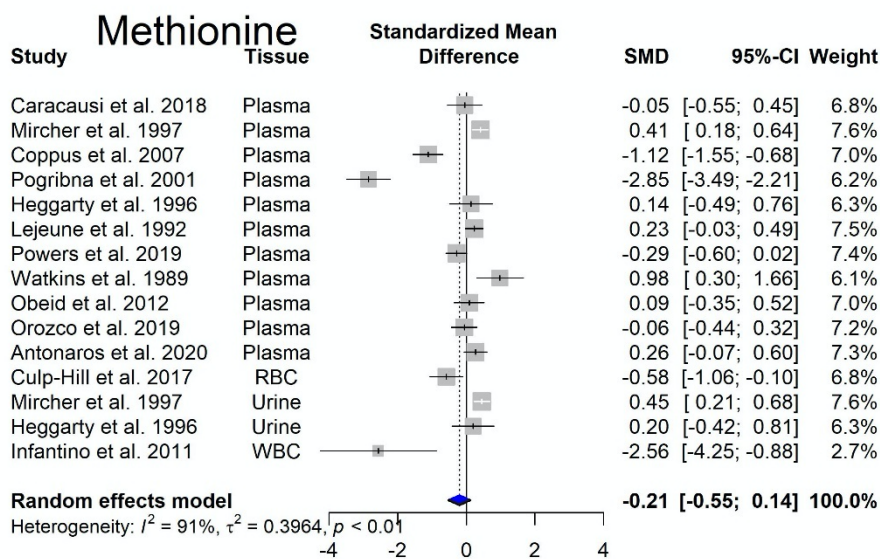
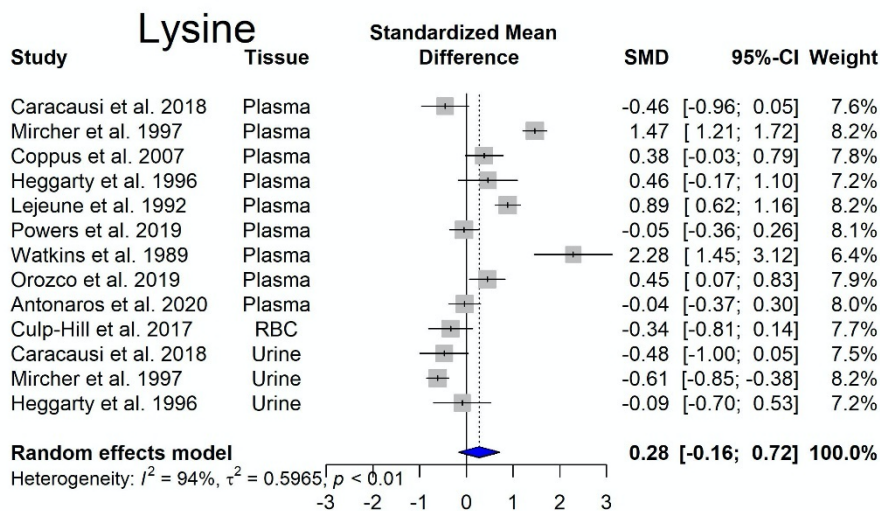


Histidine

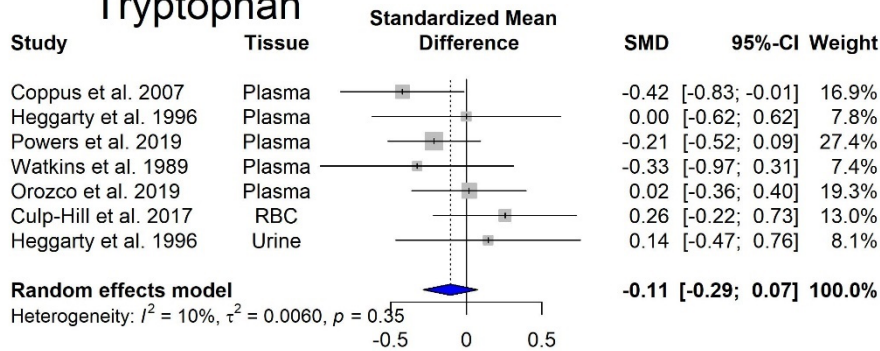


Isoleucine

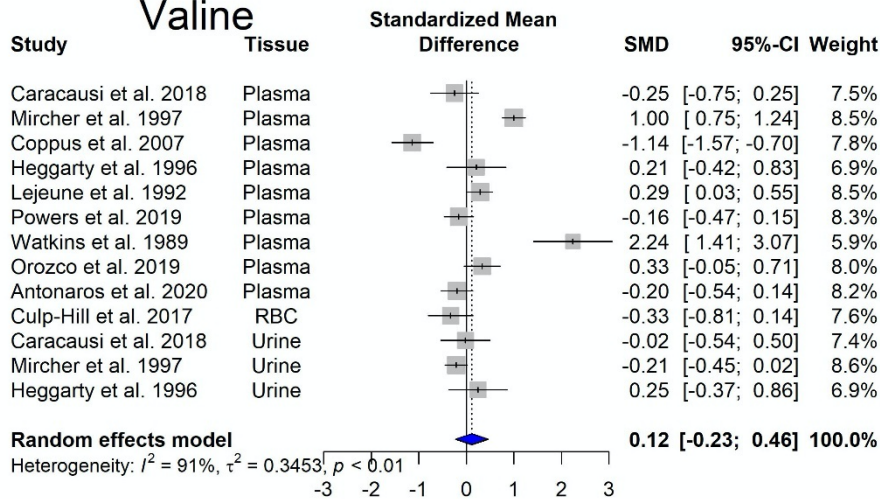




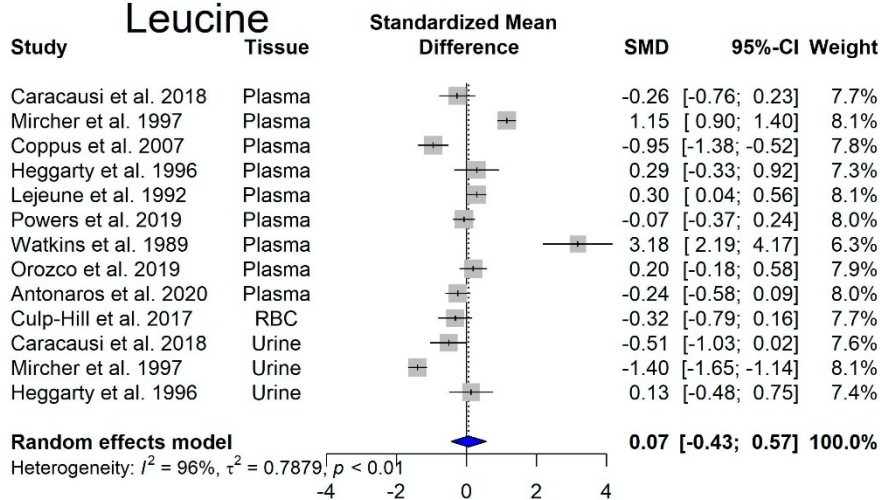
Tryptophan



Valine



Leucine



Supplementary Fig. 9. Meta-analysis of amino acids not included in the Main Text.

References

- Antonaros, F., Ghini, V., Pulina, F., Ramacieri, G., Cicchini, E., Mannini, E., Martelli, A., Feliciello, A., Lanfranchi, S., Onnivello, S., Vianello, R., Locatelli, C., Cocchi, G., Pelleri, M. C., Vitale, L., Strippoli, P., Luchinat, C., Turano, P., Piovesan, A., and Caracausi, M. (2020) Plasma metabolome and cognitive skills in Down syndrome. *Scientific Reports* **10**, 10491.
- Atkinson, D.E., Walton, G.M. (1967) Adenosine triphosphate conversion in metabolic regulation. Rat liver cleavage enzyme. *The Journal of Biological Chemistry* **242**, 3239-3241.
- Caracausi, M., Ghini, V., Locatelli, C., Mericio, M., Piovesan, A., Antonaros, F., Pelleri, M. C., Vitale, L., Vacca, R. A., and Bedetti, F. (2018) Plasma and urinary metabolomic profiles of Down syndrome correlate with alteration of mitochondrial metabolism. *Scientific reports* **8**, 1-16.
- Culp-Hill, R., Zheng, C., Reisz, J. A., Smith, K., Rachubinski, A., Nemkov, T., Butcher, E., Granrath, R., Hansen, K. C., and Espinosa, J. M. (2017) Red blood cell metabolism in Down syndrome: Hints on metabolic derangements in aging. *Blood advances* **1**, 2776-2780.
- Dieterle, F., Ross, A., Schlotterbeck, G., and Senn, H. (2006) Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. *Anal Chem* **78**, 4281-4290.
- Huber, W., von Heydebreck, A., Sültmann, H., Poustka, A., and Vingron, M. (2002) Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* **18 Suppl 1**, S96-104.
- Powers, R. K., Culp-Hill, R., Ludwig, M. P., Smith, K. P., Waugh, K. A., Minter, R., Tuttle, K. D., Lewis, H. C., Rachubinski, A. L., and Granrath, R. E. (2019) Trisomy 21 activates the kynurenine pathway via increased dosage of interferon receptors. *Nature communications* **10**, 1-11.
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., and Smyth, G. K. (2015) Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47.