Genes govern metabolism—Enzymes define pathways and metabolic relationships

Jiada Zhan[1,2], Zachery R. Jarrell[2], Melissa L. Kemp[3], Jaclyn Weinburg[2], Young-Mi Go[1,2], Greg S. Martin[2], Dean P. Jones[1,2]

1: Nutrition and Health Sciences, Laney Graduate School, Emory University, Atlanta, Georgia, USA

2: Division of Pulmonary, Allergy, Critical Care and Sleep Medicine, Department of Medicine, Emory University, Atlanta, Georgia, USA

3: The Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology & Emory University, Atlanta, Georgia, USA

Abstract

Gene centric pathway mapping tools, widely used to interpret untargeted LCMS metabolomics data, may underperform because a single metabolite can generate multiple spectral features, inflating false positive rates. Classic enzymology, which established metabolite flow long before gene sequencing, offers experimentally validated precursor product relationships that could overcome these ambiguities. We evaluated whether enzymology defined precursor product correlations are consistently detectable in human plasma LCMS data and explored their potential to enhance pathway analysis and metabolite identification. Using a high resolution LCMS platform, we detected amino acids and carnitine related metabolites in one individual sampled eight times over five years and in 50 adults sampled 6 to 8 times each. Spearman correlations were calculated for longitudinal and cross sectional data. In the single participant repeated measures, strong positive correlations were observed for every direct precursor product pair except the branch point metabolite palmitoylcarnitine. The longitudinal analysis reproduced these patterns, and the same relationships were retained when analysis was cross sectional. Despite contributions from multiple organ systems, plasma thus preserved core enzymatic relationships. Precursor product proportionality, a fundamental principle of enzymology, is readily detectable in large scale LCMS datasets and remains robust across longitudinal and cross sectional designs. Applying these correlations to metabolomics workflows can improve pathway

analysis, help metabolite identification, and reveal how genetic variations, diets, therapeutic drugs, and environmental exposures jointly impact metabolic pathways.

Recently, five biochemical and analytical challenges have been identified that can cause gene-centric tools to underperform when applied to metabolomics data[1]. Untargeted liquid chromatography mass spectrometry (LC-MS) metabolomics-based pathway analysis often relies on pathway-mapping tools[2,3], but ambiguities in mass spectral peak assignment – multiple peaks arising from the same metabolite – can generate high false positive rates[4,5]. Most metabolic pathways were established through enzymology before gene sequencing became available, and these studies offer direct, experimentally validated maps of metabolite flow. Thus, enzymology offers a simple but powerful way forward to enhance pathway analysis.

Enzymology established that metabolites flow from precursors to products and form dynamic linear, cyclical, and branched systems with regulation occurring at committed steps and branch points[6] (**Fig. 1a)**. Biochemical analyses showed that under steady-state conditions, the concentrations of metabolites often vary proportionately[6,7] (see legend, **Fig. 1)**. This principle was demonstrated in the pathway-level identification of xenobiotics and xenobiotic metabolites in blood[9] and endogenous metabolites in plant tissue[7]. Here, we show examples of precursor-product correlations observed in LC-MS of amino acids (glutamine, glutamate, ornithine, citrulline, and arginine) and carnitine-related metabolites (butyrobetaine, carnitine, propionylcarnitine, O-acetylcarnitine, palmitoylcarnitine) in one person's plasma sampled eight times across five years (**Fig. 1b**)[9]. Among these examples, the lack of proportional change of palmitoylcarnitine with carnitine illustrates that positive correlation is not always present, especially for metabolites at branch points and regulated sites.

To test whether positive precursor-product correlations exist at a population level, we analyzed LC-MS plasma data from 50 individuals, each with six to eight visits[9]. Consistent with the data found in the individual (**Fig. 1b**), amino acids and carnitine-related metabolites exhibit positive correlations across the 50 individuals (**Fig. 1c**), except for the lack of proportional variation of palmitoylcarnitine with carnitine. Unlike these longitudinal comparisons, however, most human metabolomics studies use a cross-sectional design with only single LC-MS snapshots of individuals. To determine whether enzymology-defined precursor-product correlations are present in a cross-sectional sampling, we restricted comparisons of the 50 individuals to include only baseline samples. Results show that the same relationships are present (**Fig. 1d**).

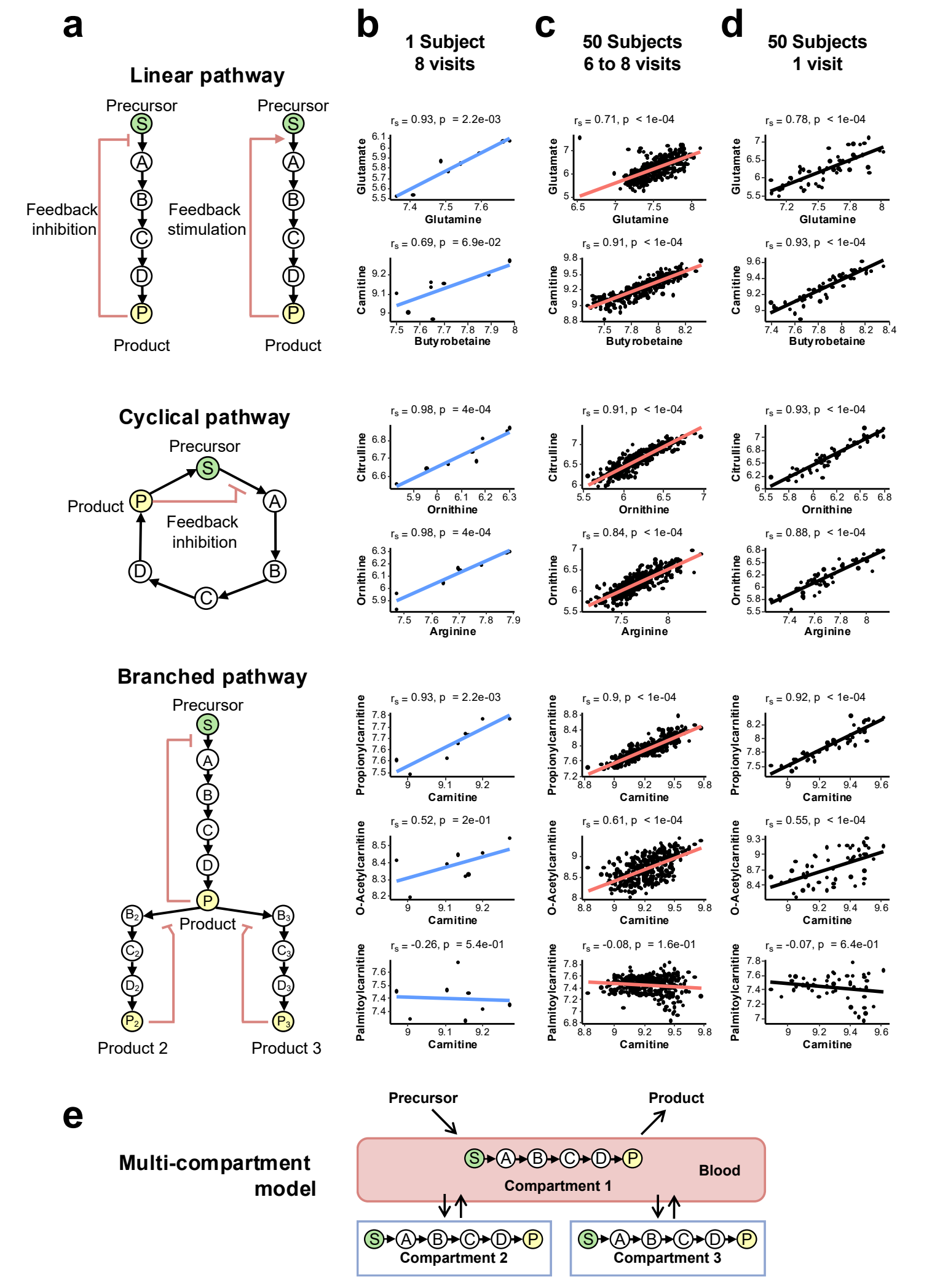
As metabolomics moves beyond static pathway maps, enzymology-defined precursor-product correlations provide a mechanistically grounded view of metabolic pathways. Even though most metabolism occurs in organ systems other than blood[1], the plasma data (**Fig. 1b, 1c, 1d**) show that precursor-product relationships are preserved despite variable contributions of organ systems (**Fig. 1e**). Incorporating these principles into metabolomics can improve pathway analysis, help metabolite identification, and reveal how genetic variations, diets, therapeutic drugs, and environmental exposures jointly impact metabolic pathways. As with gene-centric tools, one must be cautious about enzyme-based relationships[6]; some precursors and products do not show proportional behavior, and some distantly related metabolites have strong correlations. Despite these limitations, the results show that precursor-product relationships are highly preserved under different conditions and provide a valuable resource for LC-MS-based metabolomics research.

<u>Figure Legend</u>

Figure 1. Precursor-product relationships in human plasma. (**a**). Metabolism occurs in linear (left, top), cyclic (left, middle), and branched pathways (left, bottom). For sequential enzyme reactions in an open system at steady state, rates of precursor (S) input, formation of intermediate products (A, B, C, D), and formation and elimination of end product (P) are equivalent. With the same reaction rates for all steps, steady-state concentrations of intermediates [A, B, C, D] are related by the respective pseudo-first order rate constants (k_j): $k_1[S] = k_2[A] = k_3[B] = k_4[C] = k_5[D]$. Because regulation typically occurs only at branch points or at sites committing a hydrocarbon to a specific fate, this means that proportional relationships of precursors and products are expected. (**b**) Proportional concentrations for precursor and product metabolites in different pathways are evident in eight repeated blood samplings from an individual collected over five years (blue). Examples are provided for amino acids and carnitine-related metabolites, measured by liquid chromatography mass spectrometry (LC-MS) of plasma extracts (See Reference 10 for methodological details[10]). All the metabolites on the x-axis are the precursors, and all the metabolites on the y-axis are the products. Glutamine, glutamate, butyrobetaine, and carnitine are considered part of a linear pathway, ornithine, citrulline, and arginine are considered part of a cyclical pathway (urea cycle), and carnitine, propionylcarnitine, O-acetylcarnitine, and palmitoylcarnitine are considered part of a branched pathway (carnitine shuttle). Sample collection involved overnight fasting for the first blood draw, but not for the others. Note that an example for a precursor-product pair without proportional concentrations, palmitoylcarnitine and carnitine (bottom row), is provided. (**c**) Proportional concentrations for the same precursor and product pairs as in Panel b for 50 individuals with multiple repeated blood samplings across five years (red). (**d**) Cross-sectional analysis of baseline samples for the 50 individuals in Panel c (black) shows that precursor-product correlations are preserved for the same precursor and product pairs. (**e**) In multicompartment models, most metabolism occurs in organ compartments rather than the blood. Despite this, the data in Panels **b** to **d** show that proportional relationships of precursors and products are observed in blood plasma samples.

References

1. Lee, K. S., Su, X. & Huan, T. Metabolites are not genes — avoiding the misuse of pathway analysis in metabolomics. *Nat. Metab.* 1–4 (2025).
2. Li, S. *et al.* Predicting Network Activity from High Throughput Metabolomics. *PLOS Comput. Biol.* **9**, e1003123 (2013).
3. Pang, Z. *et al.* MetaboAnalyst 6.0: towards a unified platform for metabolomics data processing, analysis and interpretation. *Nucleic Acids Res.* gkae253 (2024).
4. Giera, M., Aisporna, A., Uritboonthai, W. & Siuzdak, G. The hidden impact of in-source fragmentation in metabolic and chemical mass spectrometry data interpretation. *Nat. Metab.* **6**, 1647–1648 (2024).
5. Domingo-Almenara, X., Montenegro-Burke, J. R., Benton, H. P. & Siuzdak, G. Annotation: A Computational Solution for Streamlining Metabolomics Analysis. *Anal. Chem.* **90**, 480–489 (2018).
6. Steuer, R., Kurths, J., Fiehn, O. & Weckwerth, W. Observing and interpreting correlations in metabolomic networks. *Bioinforma. Oxf. Engl.* **19**, 1019–1026 (2003).
7. Rose, I. A. & Warms, J. V. Control of glycolysis in the human red blood cell. *J. Biol. Chem.* **241**, 4848–4854 (1966).
8. Liu, K. H. *et al.* Large scale enzyme based xenobiotic identification for exposomics. *Nat. Commun.* **12**, 5418 (2021).
9. Tabassum, R. *et al.* A Longitudinal Study of Health Improvement in the Atlanta CHDWB Wellness Cohort. *J. Pers. Med.* **4**, 489–507 (2014).
10. Liu, K. H. *et al.* Reference Standardization for Quantification and Harmonization of Large-Scale Metabolomics. *Anal. Chem.* **92**, 8836–8844 (2020).


a
Linear pathway
Precursor
Feedback inhibition
Feedback stimulation
Product
Cyclical pathway
Precursor
Product
Feedback inhibition
Branched pathway
Precursor
Product
Product 2
Product 3
b
1 Subject
8 visits
c
50 Subjects
6 to 8 visits
d
50 Subjects
1 visit
Glutamate
Glutamine
Carnitine
Butyrobetaine
Citrulline
Ornithine
Arginine
Propionylcarnitine
O-Acetylcarnitine
Palmitoylcarnitine
e
Multi-compartment model
Precursor
Product
Blood
Compartment 1
Compartment 2
Compartment 3