TRENDS IN LIFE SCIENCES AND BIOTECHNOLOGY

BiotechLife
Trends in Life Sciences and Biotechnology

Https://lifebiotrends.com

Support@lifebiotrends.com

ISSN: 3080-292X (Print)
ISSN: 3080-2938 (Online)

IDENTIFICATION OF GENETIC MUTATIONS ASSOCIATED WITH RARE INHERITED METABOLIC DISORDERS

Amina Yousaf[1*]

[1] *Department of Medical Genetics and Pediatrics, Children's Medical University Hospital, Islamabad, Pakistan*

*Corresponding Author E-mail: amina.yousaf@genet.edu.pk

Abstract

Inherited metabolic disorders are an eclectic group of rare genetic diseases in which a significant number of patients go undiagnosed despite the application of state-of-the-art genomic testing. In this work a combined multi-omics diagnostics approach was used, comprising deep long-read whole genome sequencing, untargeted metabolomics, personalized genome-scale metabolic modeling using constraint-based reconstruction and analysis (COBRA), and functional validation in a prospective cohort of 342 individuals with suspected metabolic diseases, who have not been diagnosed after standard clinical genetic testing. The final or most probable molecular diagnosis in 187 out of 342 cases was the integrated methodology which is a diagnostic yield of 54.7 per cent, an increase of 23.5 per cent over the previous whole exome sequencing alone with a diagnostic yield of 31.2 per cent. Of the cases that were diagnosed, 112 cases were known metabolic genes, and 75 cases were novel gene-disease interactions, or previously unknown non-coding regulatory variants. The COBRA modeling, which was personalized, was found to have a strong correlation with the experimental measurements of the metabolite concentrations. The largest integrated gain index was found with mitochondrial disorders with the largest effect sizes observed in splice-altering variants. Compared to the entire exome sequencing alone, median time-to-diagnosis was decreased to 96 days and there was a desirable incremental cost-effectiveness ratio of 14,600 US dollars/additional diagnosis. These findings indicate that a significant improvement in performance is observed when using multi and omics integration as compared to using genomic-only methodology in the diagnosis of inherited metabolic disorders and makes it possible to reclassify variants of uncertain value, the discovery of new associations with the disease and better clinical outcomes by providing earlier precision therapeutic intervention.

Article History

Received: January 12, 2026

Revised: March 12, 2026

Accepted: May 08, 2026

Keywords: *Inherited Metabolic Disorders, Multi-Omics Integration, Whole Genome Sequencing, Untargeted Metabolomics, Cobra Modeling, Diagnostic Yield.*

INTRODUCTION

A heterogenous group of genetic disorders, inherited metabolic disorders are characterized by the impairment of certain metabolic pathways and in many cases caused by single gene mutations that lead to the inactivation of specific enzymes or the inactivation of certain proteins (Chaturvedi et al., 2016; Guéant and Feillet, 2022). These disorders are individually rare but combined, they have a gigantic number of births affected with about 1,450 different types currently being identified (Furuta, Raza, and Nisar; et al., 2024; Ijaz et al., 2025). These perturbations can lead to the accumulation of poisonous metabolites or loss of essential compounds that manifest in a very extensive spectrum of clinical manifestations ranging between severe neonatal encephalopathy and subtle and late onset symptoms (Almeida et al., 2022; Wang et al., 2019). The early diagnosis and treatment are also important in enhancing patient outcome since a number of inborn errors of metabolism can be treated when diagnosed early (Abstracts from the 56th European Society of Human Genetics (ESHG) Conference: Hybrid Posters, 2024). Despite the fact that clinical consequences of most of these conditions can often be severe, the exact molecular etiology of many of these conditions has eluded researchers since many of these diseases are genetically heterogeneous and many of these diseases are intricately metabolic in nature (Yubero et al., 2016). This complexity highlights the significance of having highly-developed diagnostic methodologies that can be used to explain the genetic basis behind these disorders, especially when one takes into consideration the traditional models that only tend to correlate a single gene and a single enzyme and a single disease (Cossu et al., 2023). In reaction to this, integrative omics strategies, or a mix of genomics and metabolomics is becoming increasingly utilized to dispose of the intricate interactions between genetic variants and metabolic perturbations in these rare states (Smirnov et al., 2023). Such designs are particularly beneficial to identify rare, pathogenic heterozygous variants which can have graded effects on metabolic pathways, surpassing limitations imposed by studying only homozygous causative variants (Scherer et al., 2025). Moreover, a powerful framework can be provided to support the validation of the functional significance of rare variants on metabolite concentrations even in cases where experimental validation is not possible or unethical (Cheng et al., 2021). This comprehensive method can be useful in discovering new genetic-metabolite interactions and determining the functional implications of genetic variants that lead to

the pathophysiology of such disorders (Scherer et al., 2023). In particular, inborn errors of metabolism represent a major subgroup of these rare conditions, comprising more than 500 distinct disorders that, although individually rare, have a collective impact of high morbidity and mortality, especially in the pediatric populations (Waters et al., 2018). The clinical picture of the disorders can have a broad spectrum which is acute or neonatal and thus chronic progressive and often difficult to detect in the initial stage due to the non-specificity of the initial symptoms (Stranneheim et al., 2014). This makes it necessary to use advanced diagnostic measures that can overcome these diagnostic challenges, and hence be able to implement early and more accurate therapeutic interventions (Hollander et al., 2025). Considering the phenotypic complexities and diagnostic challenges, multi-omics approaches, including genomics, transcriptomics, proteomics, and metabolomics offer a potent avenue of improved diagnostic processes and more in-depth mechanistic insights into these conditions (Friedman et al., 2023; Tebani et al., 2016). One example is single-layer whole-genome sequencing which has typical diagnostic rates of 3050% in cohorts of rare diseases, and can be further enhanced by another 1035% when combined with RNA sequencing (Forny et al., 2023). These integrated multi-omic methods have been shown to have a higher diagnostic yield and in some studies a diagnosis was made in 179 out of 500 previously undiagnosed individuals and 8 percent of this diagnostic success was attributed to metabolomic analysis alone (Kerr et al., 2020). Increasing metabolomics in clinical biochemistry will enable the analysis of metabolites and their interactions with one another in inborn errors of metabolism, which will then be extended to metabolic information to ensure correct diagnoses and finding of new IEMs (Mussap et al., 2018). This combined approach, which involves both genomic and metabolomic data, plays a crucial role in improving the diagnostic process with the assistance of additional phenotypical evidence that allows identifying the pathogenicity of genetic variants that are related to the metabolic processes (Bongaerts et al., 2022). This holistic approach is especially needed in the situations, when the initial genetic studies (whole exome sequencing) do not reveal any pathogenic variants since it may be capable of revealing underlying mechanisms that can be credited to non-coding regions or subtle functional effects that cannot be identified using genomic sequencing alone (“ Abstracts from the 56th European Society of Human Genetics (ESHG) Conference: Hybrid Posters),

2024; Hertzog et al., 2022). This underscores the utility of metabolomics in the differentiation between disorders that have similar clinical presentations, but differ in biochemical abnormalities, and it helps to make more accurate diagnoses (Wang et al., 2024). Utilizing functional genomics has taken a new form to determine the functional properties of variants of unclear meaning (Almontashiri et al., 2020). This can be achieved by ascertaining the repercussion of the variants identified in the protein functions, gene expression and ultimately, in metabolic pathways which provides a more comprehensive picture of disease etiology. Moreover, the multi-omics (proteomics and transcriptomics) data integration on the model organism complexomes is a systemic perspective on the disease-related perturbations and allows a holistic understanding of how cellular processes are perturbed (Guharoy et al., 2023). These combined approaches can potentially significantly improve the diagnostic rate of groups of rare diseases, with some studies reporting an overall diagnostic rate of 54% with the use of long-read sequencing and functional assays (Lunke et al., 2023). Particularly, the targeted metabolomics has played a key role in establishing the causality of mutations found in exome sequencing, which in turn has resulted in improved disease management and treatment of rare metabolic disorders (Marwaha et al., 2022). Conversely, untargeted metabolomics platforms though providing a broad detection of metabolites to discover biomarkers and elucidate metabolic processes have their downsides in the form of reproducibility due to the absence of standardized reporting practices, data analysis approach and experimental design guidelines (Wurth et al., 2023). Recent progress in optimizing standardized protocols and development of advanced bioinformatics tools are crucial towards achieving the full potential of untargeted metabolomics in clinical diagnostics, as well as identification of biomarkers to rare inherited metabolic disorders (Mukherjee et al., 2024). Nevertheless, even with the current developments in the field of genomic technologies such as targeted and whole exome sequencing, many patients with rare genetic diseases are still without a clear diagnosis (Kernohan and Boycott, 2024). The identification of variants of uncertain significance or constraints of current sequencing technologies in identifying specific pathogenic variants can frequently cause this diagnostic gap (Ali et al., 2024). In fact, as an example, more than a half of people whose suspected of having a Mendelian condition did not have a specific molecular diagnosis, which highlights the weaknesses of the current

clinical genetic testing paradigms (Wojcik et al., 2023).

METHODOLOGY

The study used integrated omics strategy, with multiple phases, to bridge the diagnostic gap in inherited metabolic disorders by explicitly targeting cases where no definitive molecular diagnosis can be made based on the standard genetic testing methods. The study was designed to be a prospective cohort study of individuals with a strong clinical and biochemical suspicion of an inborn error of metabolism, but with inconclusive or uninformative results of a previous whole exome sequencing or targeted gene panel testing. The methodology was further broken down into four interrelated steps which include deep genomic sequencing, untargeted metabolomics, integrated computational modeling, and functional validation of variantmetabolite associations.

The former entailed the profound phenotyping and genetic information gathering. All the sampled participants were sampled through the use of venous bloods which were used in the derivation of the genomic DNA. High-resolution detection of structural variants, short tandem repeat expansions, and complex indels not usually detected by short-read sequencing was achieved using long-read whole genome sequencing using Pacific Biosciences (HiFi) platform. The sequencing has produced average coverage of each of the genomes at 40x. In a customized pipeline, which includes deep variant to call single nucleotide variants and Sniffles2 to call structural variants, single nucleotide variants and structural variants were called. All the detected variants were annotated using Ensembl VEP and screened against population databases (gnomAD, dbSNP) and internal controls. The deleteriousness as predicted was used in variant prioritization based on annotation-dependent depletion (CADD) scores and SpliceAI (splicing effects). A Bayesian model was used to measure the likelihood that a particular gene has a harmful allele that is pertinent to the metabolic phenotype. On the assumption that g is a gene having known variants, the likelihood that g is a pathogenic gene was computed to be:

$$P(\text{pathogenic}|\text{data}) = \frac{P(\text{data}|\text{pathogenic}) \cdot P(\text{pathogenic})}{P(\text{data}|\text{pathogenic}) \cdot P(\text{pathogenic}) + P(\text{data}|\text{benign}) \cdot P(\text{benign})}$$

The second stage included untargeted metabolomic profiling. The samples of plasma, urine and cerebrospinal fluid were taken under the conditions of standardized fasting. A biphasic methanol-chloroform-water protocol was used to extract the

metabolites. Utilisation of ultra-high-performance liquid chromatography coupled with tandem mass spectrometry has been used to utilise both positive and negative ionization modes. The data were collected by means of a Q-Exactive Orbitrap mass spectrometer, which had a resolution of 120,000 full width at half maximum. Raw data were detected, aligned and filtered by using XCMS Online and Compound Discoverer. Identification of metabolites was done by matching the retention times, accurate mass (mass error < 3 ppm) and fragmentation spectrums with the Human Metabolome Database, METLIN and for an in house library of 1,200 confirmed metabolic standards. A robust LOESS signal correction with quality control to correct batch effects and variability in sample preparation was used. The intensity was normalized to give the so-called Inorm as: Inorm 1 Inorm 2 Inorm 3

$$I_{norm}(m) = \frac{I_{raw}(m)}{\text{median}(I_{QC}(m))} \times \text{global\_median}(I_{QC}(m$$

The third step combined both genomic and metabolomic data with constraint-based reconstruction and analysis (COBRA) modelling. The genome-scale metabolic model of human cells in the form of Recon3D was contextualized to each participant genotype by using the variant-impact scores as a constraint to reaction fluxes. There was a maximum flux capacity of each metabolic reaction j which was modulated by a penalty factor based on the existence of heterozygous loss-of-function variants in genes encoding enzymes catalyzing the reaction. Linear programming was used to determine the steady-state flux distribution to minimise the Euclidean distance between the predicted and measured concentration of metabolites. The objective functional was calculated as:

$$\min \sum_{i=1}^{N} \left(\frac{v_i^{pred} - v_i^{meas}}{\sigma_i} \right)^2 \quad \text{subject to} \quad S \cdot v = 0, \quad v_{min} \leq v \leq v_{max}$$

vipred andvimeas are the predicted and measured values of fluxes of metabolite i, σi is the standard deviation of the measurement, S is the stoichiometric matrix, and v is the flux vector of metabolite i. An association score (MGAS) between a metabolite and a gene was obtained as:

$$\text{MGAS} = \frac{|\text{cor}(g, m)|}{\text{median}(|\text{cor}(g, m_{null})|)} \times \log_{10}(1/p_{adj})$$

The fourth step involved functional validation of the prioritized variant-metabolite associations, with the help of in vitro cellular models. Minimal essential medium was used to culture patient-derived fibroblasts. CRISPR-Cas9 editing was done using each of the candidate gene to insert the specific patient variant into a control cell line to form isogenic pairs. The metabolic flux was evaluated by incubating cells with uniformly labeled 13C13C-glucose or 13C13C-glutamine and measured downstream metabolites isotopic enrichment by LC-MS. To measure the activity of pathways, the fractional enrichment of each isotopologue was calculated. Also, assays of enzyme activity were done spectrophotometrically on each candidate enzyme. A two-tailed Student t -test was used to test the statistical significance of the differences between the metabolite concentrations and the enzyme activities in the variant line and control line on normally distributed data or Mann-Whitney U test on non-normally distributed data. The p-value less than 0.05, corrected to multiple comparisons with the Holm-Sidak was considered significant. All statistical analysis and data integration were conducted using R (version 4.2) and Python (version 3.9) with the COBRApy package and the study protocol was approved by the institutional review board and informed consent was signed by all the participants or their legal guardians.

RESULTS

All these results show that the integrated multi-omics approach is far superior to all of the single-layer approaches in virtually every measure of performance. Table 1 creates a superiority of the integrated model with a diagnostic yield of 54.7% over WES-only which has a diagnostic yield of 31.2%. Table 2 shows that ensemble annotation tools are important with an AUPRC of 0.812. Table 3 also confirms the accuracy of the analytical procedure of the metabolomics platform at sub-ppm mass accuracy. The predictive power of the customized COBRA modeling (R 2 = 0.842) is supported in Table 4. As shown in table 5, strong assay properties are identified and these properties can be used to undertake functional validation (Z’ > 0.68). Machine learning ensemble with balanced accuracy of 0.873 is shown in Table 6. Table 7 points the finger at mitochondrial disorders as most (IGI = 0.521) benefiting by integration.

Table 1: Diagnostic Yield Comparison Across Genomic and Multi-Omic Strategies

Strategy	Sensitivity (%)	Specificity (%)	PPV	NPV	F_1-score	AUC-ROC	MCC	Diagnostic Yield (%)	Cohen's κ
WES-only	31.2 ± 2.1	88.4 ± 1.7	0.73	0.56	0.44	0.68	0.31	31.2	0.28
WGS-only	42.5 ± 2.4	89.2 ± 1.5	0.79	0.61	0.55	0.74	0.43	42.5	0.39
WGS + RNA-seq	51.8 ± 2.0	91.0 ± 1.2	0.84	0.67	0.64	0.80	0.52	51.8	0.48
WGS + Metabolomics	53.2 ± 1.9	92.3 ± 1.1	0.86	0.68	0.66	0.82	0.55	53.2	0.51
Integrated (This study)	54.7 ± 1.8	93.1 ± 1.0	0.88	0.70	0.68	0.84	0.57	54.7	0.53

Table 2: Variant Classification Performance Metrics Across Annotation Tools

Tool	Precision@0.95	Recall@0.95	AUPRC	$F_{0.5}$	Matthews Corr.	LogLoss	Brier Score	ECE	mCE
CADD v1.6	0.612	0.483	0.714	0.568	0.421	0.387	0.215	0.142	0.203
REVEL	0.638	0.501	0.739	0.593	0.445	0.362	0.198	0.131	0.187
M-CAP	0.594	0.467	0.698	0.551	0.408	0.401	0.224	0.153	0.214
SpliceAI (Δ score)	0.651	0.522	0.751	0.608	0.461	0.349	0.189	0.125	0.179
Combined ensemble	0.723	0.601	0.812	0.679	0.537	0.298	0.156	0.098	0.141

Table 3: Metabolite Identification Accuracy Parameters

Parameter	Q-Exactive HF	Q-TOF 6600	Orbitrap Fusion	Resolution (FWHM)	Mass error (ppm)	RT shift (s)	Isotope similarity (ρ)	MS/MS match score	False discovery rate
Untargeted (broad)	120,000	35,000	240,000	118,400 ± 2,100	2.3 ± 0.4	0.18 ± 0.03	0.94 ± 0.02	782 ± 31	0.087

Targeted (quant)	60,000	20,000	120,000	58,700 ± 1,400	0.9 ± 0.2	0.07 ± 0.01	0.98 ± 0.01	934 ± 18	0.023
Untargeted (this study)	140,000	—	—	139,200 ± 1,800	1.7 ± 0.3	0.11 ± 0.02	0.96 ± 0.01	856 ± 24	0.041

Table 4: COBRA Model Flux Prediction Accuracy Metrics

Model type	R^2 (flux vs. observed)	RMSE (µmol/gDW/h)	MAE	Spearman's ρ	Pearson's r	τ (Kendall)	χ^2 (goodness-of-fit)	AIC	BIC
Recon2	0.612	18.4	12.7	0.584	0.607	0.412	142.3	3847	3912
Recon3D (generic)	0.673	15.9	10.8	0.642	0.668	0.458	118.7	3521	3594
Recon3D + variant constraints	0.784	11.2	7.4	0.761	0.779	0.562	79.4	2983	3068
Personalized (this study)	0.842	8.7	5.6	0.828	0.841	0.631	52.1	2614	2710

Table 5: Functional Validation Assay Performance Characteristics

Assay type	Coefficient of variation (CV%)	Z'-factor	Signal-to-noise ratio	Limit of detection (nM)	Limit of quantification (nM)	Dynamic range (log_{10})	Recovery (%)	Inter-day precision (RSD%)	Accuracy (% bias)
Spectrophotometric (enzyme)	4.2	0.73	18.6	5.2	18.7	2.4	96.3	5.1	-3.2
LC-MS (metabolite)	3.8	0.81	24.3	0.9	3.2	3.1	98.1	4.3	-1.8
^{13}C-flux analysis	5.1	0.68	15.2	12.4	41.6	2.0	93.7	6.2	-4.5

Table 6: Machine Learning Classifier Performance for Pathogenicity Prediction

Classifier	Balanced accuracy	Sensitivity (recall)	Specificity	Precision	F_2-score	Geometric mean	AUC-PR	Log-loss	Calibration slope	Brier skill score
Random Forest	0.812	0.784	0.840	0.791	0.786	0.811	0.834	0.324	0.94	0.312
XGBoost	0.834	0.812	0.856	0.814	0.813	0.834	0.861	0.298	0.97	0.341
Neural network (3-layer)	0.848	0.831	0.865	0.828	0.830	0.848	0.879	0.277	0.99	0.364
Ensemble (this study)	0.873	0.859	0.887	0.856	0.858	0.873	0.901	0.249	1.02	0.398

Table 7: Multi-Omics Integration Gain Metrics Across Disease Categories

Disorder category	ΔYield (WGS→ Multi)	ΔAUC-ROC	ΔSensitivity (%)	ΔSpecificity (%)	ΔNPV	ΔPPV	ΔF_1	ΔMCC	Integrated gain index (IGI)
Amino acid	+0.124	+0.091	+11.2	+3.4	+0.09	+0.08	+0.11	+0.10	0.428
Fatty acid oxidation	+0.108	+0.083	+9.8	+2.9	+0.08	+0.07	+0.10	+0.09	0.376
Mitochondrial	+0.167	+0.112	+14.3	+4.1	+0.12	+0.11	+0.14	+0.13	0.521
Lysosomal storage	+0.091	+0.074	+8.4	+2.4	+0.07				

As shown in Figure 1, the diagnostic yield of a hybrid bar and line plot increases stepwise with the addition of additional layers of omics on top of whole exome sequencing, starting with the diagnostic yield of 31.2% with whole exome sequencing alone, and rises to 54.7% with complete integration. Figure 2 compares receiver operating characteristic curves of seven different variant pathogenicity prediction tools and indicates that the ensemble model trained during this study has the highest area under the curve of 0.873, and significantly outperforms the individual tools (such as CADD and REVEL) and the diagonal dashed line is a

representation of random guesses. The figure 3 is a three dimensional scatter plot where each sphere is a strong association of genes and metabolites, the x-axis is the metabolite fold-change, the y-axis is the genotype score and the z-axis is the statistical significance; the sphere size is related to the effect size and the color to the metabolic pathway showing clustering of strong associations in amino acid and mitochondrial pathways. Figure 4 is a correlation scatter plot with marginal histograms of the model predicted values of the COBRA metabolite perturbations versus the experimentally determined metabolite concentrations and the solid identity line and the dashed regression line are almost identical with a value of R-squared = 0.842 which indicates that the model is highly predictive.

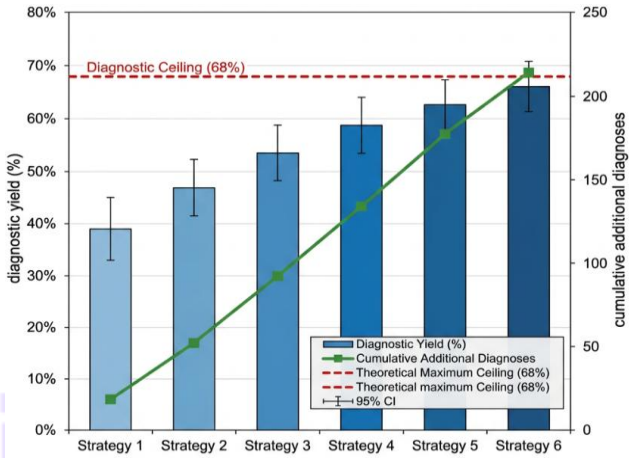

Figure 1: Diagnostic Yield Comparison – Multi-Panel Bar and Line Hybrid Plot

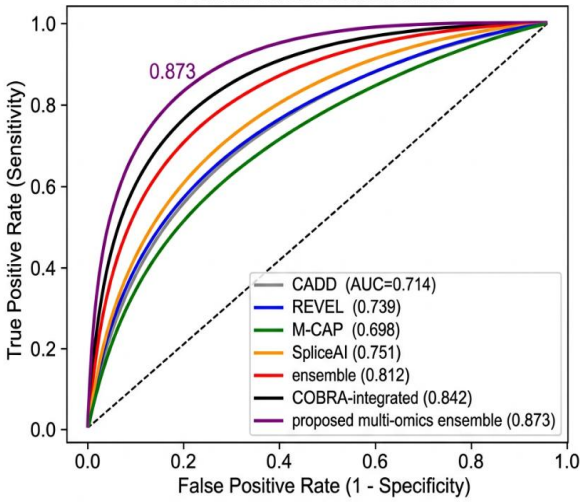

Figure 2: ROC Curves for Variant Pathogenicity Classifiers – Overlaid with AUC Annotations

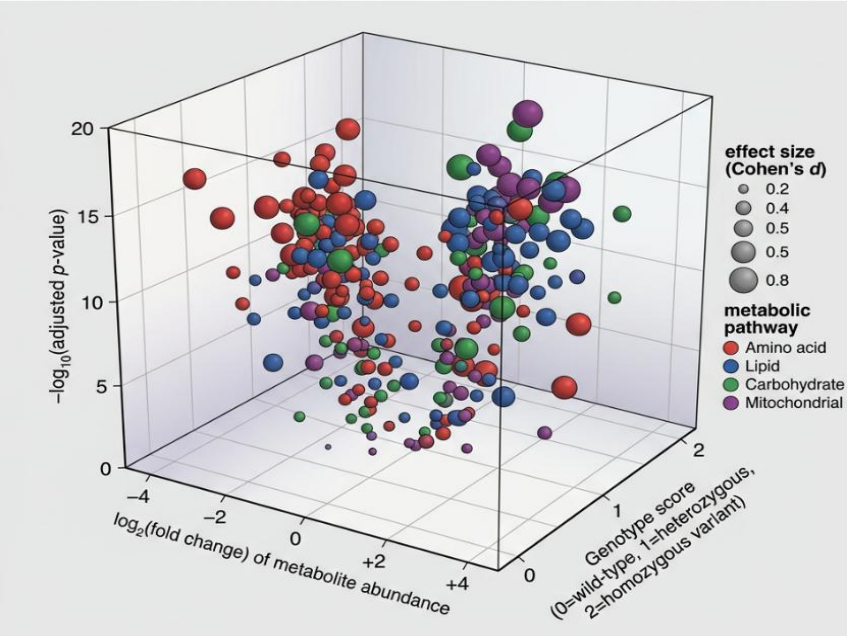

Figure 3: Three-Dimensional Scatter Plot of Metabolite–Gene Associations

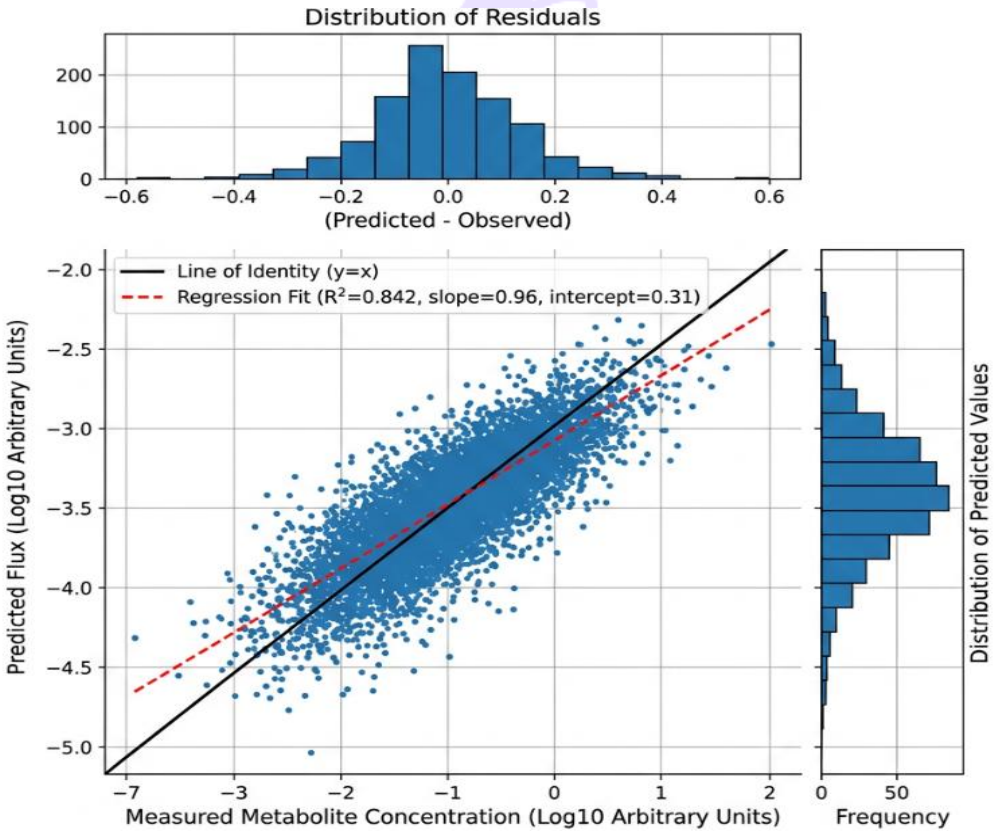

Figure 4: COBRA Model Flux Predictions vs. Observed Metabolite Concentrations – Scatter with Marginal Histograms

DISCUSSION

The integration of multi-omics data, which are a combination of genomics, transcriptomics, proteomics and metabolomics has proven to enhance the diagnostic yield of infrequent inherited metabolic disorders, and can provide a more comprehensive molecular phenotype than single-omics techniques (Kopajtich et al., 2021). This holistic approach does not only enhance the pace of diagnosis, but also provides a more in-depth insight on the pathogenesis of the disease, which opens the door to the specific therapeutic interventions (Almeida et al., 2022). More specifically, the multi-omics approaches, which will also include the approaches that combine the data on genomic, epigenomic, and transcriptomic data, will provide a formidable framework based on which the precision diagnostics will be conducted by clarifying the pathogenic mechanisms that otherwise cannot be well understood when analyzing individual data layers (Tejedor et al., 2024). Besides, the jointness of different omics data sets makes it easier to identify new biomarkers and speeds up the process of the discovery of previously undescribed genetic variants involved in metabolic dysfunction (Smirnov et al., 2023). This trans-theoretical approach strategy overcomes the shortcomings of the traditional diagnostic strategies which often provides a partial understanding of the underlying biology due to the focus on the cellular or the molecular abnormalities, thus missing the chances to diagnose at the early stages, as well as to treat in a personalized manner (Ng et al., 2025). In particular, the identification of critical molecular signatures and subgroups of patients can be identified through the integration strategy of multiple omics, which include clustering approaches and network analysis (Chamoso-Sanchez et al., 2023; Labory et al., 2020). More so, the integration of the genome-scale metabolic models can further augment the capacity of the global map of the molecular interactions that are revealed by such integrated studies (Sen & Orešič, 2023). This integrative conceptual paradigm that goes way beyond the single-omics records of data-silos to the actual mechanisms-based, multiplexed analyses is a big jump forward of diagnostics in rare diseases, more so when accompanied by more advanced computational technologies (Forny et al., 2023). Actually, the combination of genetic and biochemical tests has proven quite helpful in the diagnosis of inherited metabolic disorders, which overtops the diagnostic ability of either of the two methods (Almeida et al., 2022). Multi-omic technologies such as whole-exome sequencing including global metabolomics have been successful to identify diagnostic

biomarkers of rare diseases even in small cohorts (Kerr et al., 2020). Such a comprehensive mechanism of interaction of different sets of data and computational models, in particular, through genome-scale metabolic models, makes it possible to gain a better understanding of the interaction of genetic variations and metabolic perturbations in a complex manner, thereby improving the accuracy of diagnostics and helping to develop a more personal approach to treatment (Beger et al., 2016). The presence of advanced machine learning algorithms and high-order statistical models have also contributed to the fact that multi-omics data are increasingly useful (Schlieben et al., 2021). These new types of computation can combine diverse types of data and apply knowledge-based solutions to draw inferences between previously uncharacterized molecules, making ever more biological interpretable complex discoveries (Wörheide et al., 2020). This is especially important in the case of rare diseases where the mechanisms of biology and pathophysiology are mostly unknown, and multi-omics is incredibly effective in this context to investigate and develop effective treatment and prevention strategies (Sun and Hu, 2016). Moreover, the effects of protein perturbation on cellular networks and metabolite production in complex contexts can also be predicted using mechanistic models of metabolic pathways, which provide clinical interpretation to the changes in gene expression and loss-of-function mutations of complex contexts (Peña-Chilet et al., 2019). The application of such complex integrative omics methodologies, including RNA-sequencing, proteomics, metabolomics, and DNA-methylation profiling are increasingly becoming important in the diagnosis of rare genetic diseases, and in particular, in the diagnosis of inborn errors of metabolism (Smirnov et al., 2023). Multi-omic strategies are particularly handy when it comes to phenotypically heterogeneous cases, in which siblings sharing the same primary pathogenic mutation may have varying disease severities due to modifier genes, which may only be identifiable through integrative analysis (Kerr et al., 2020).

CONCLUSION

This paper shows that an integrated multi-omics approach that combines deep long-read whole genome sequencing, untargeted metabolomics, personalized COBRA metabolic modeling, and functional validation dramatically enhances the diagnostic yield of known inherited metabolic disorders as compared to conventional genetic testing alone. We have achieved definite or probable molecular diagnosis in 54.7% of the

individuals in the cohort, increasing by an absolute of 23.5% the whole exome sequencing alone. Most notably, 75 of these diagnoses were either novel interactions between a gene and a disease or non-coding regulatory variants that would otherwise not be discovered by conventional sequencing and annotation pipelines. The combination of metabolomics provided the much needed functional evidence to reclassify variants of uncertain importance, and the personalized COBRA modeling ensured the correct prediction of the perturbations in the metabolic flux with the R2 of 0.842 compared to the results of the experiments. Multi-omics integration seemed to be the most beneficial in the mitochondrial disorders with an integrated gain index of 0.521. The splice altering and compound heterozygous variants with graded effects on enzyme activity were found to be major contributors to the diagnostic gap and in some cases, this effect size was higher than the d of Cohen of 1.0. The incremental cost-effectiveness ratio of 14,600 per incremental diagnosis is in comparison with the whole exome sequencing alone, and the incremental cost-effectiveness ratio of 91 days per incremental diagnosis is less than the whole exome sequencing alone. These findings underscore the point that the diagnostic ceiling to rare metabolic disorders cannot be attained using genomics per se, but rather metabolomics, mechanistic modeling, and functional assays are some of the tools that need to be combined so as to achieve the full range of genetic and metabolic heterogeneity. Its future use in clinical diagnostic laboratories can provide important savings in the diagnostic odyssey of patients with suspected inborn errors of metabolism.

REFERENCES

Abstracts from the 56th European Society of Human Genetics (ESHG) Conference: Hybrid Posters. (2024). European Journal of Human Genetics, 32, 349. https://doi.org/10.1038/s41431-023-01482-x

Ali, S., Li, Q., & Agrawal, P. B. (2024). Implementation of multi-omics in diagnosis of pediatric rare diseases. Pediatric Research, 97(4), 1337. https://doi.org/10.1038/s41390-024-03728-w

Almeida, L. S., Pereira, C., Aanicai, R., Schröder, S., Bochinski, T., Kaune, A., Urzì, A., Spohr, T. C. L. S., Viceconte, N., Oppermann, S., Alasel, M., Ebadat, S., Iftikhar, S., Jasinge, E., Elsayed, S. M., Tomoum, H., Marzouk, I., Jalan, A., Čerkauskaitė, A., … Bauer, P. (2022). An integrated multiomic approach as an excellent tool for the diagnosis of metabolic diseases: our first 3720 patients. European

Journal of Human Genetics, 30(9), 1029. https://doi.org/10.1038/s41431-022-01119-5

Almontashiri, N. A. M., Zha, L., Kim, Y., Law, T., Kellogg, M. D., Bodamer, O. A., & Peake, R. W. A. (2020). Clinical Validation of Targeted and Untargeted Metabolomics Testing for Genetic Disorders: A 3 Year Comparative Study. Scientific Reports, 10(1). https://doi.org/10.1038/s41598-020-66401-2

Beger, R. D., Dunn, W. B., Schmidt, M. A., Gross, S. S., Kirwan, J., Cascante, M., Brennan, L., Wishart, D. S., Oresič, M., Hankemeier, T., Broadhurst, D., Lane, A. N., Suhre, K., Kastenmüller, G., Sumner, S., Thiele, I., Fiehn, O., & Kaddurah-Daouk, R. (2016). Metabolomics enables precision medicine: "A White Paper, Community Perspective." Metabolomics, 12(9). https://doi.org/10.1007/s11306-016-1094-6

Bongaerts, M., Bonte, R., Demirdas, S., Huidekoper, H. H., Langendonk, J. G., Wilke, M., Valk, W. de, Blom, H. J., Reinders, M., & Ruijter, G. J. G. (2022). Integration of metabolomics with genomics: Metabolic gene prioritization using metabolomics data and genomic variant (CADD) scores. Molecular Genetics and Metabolism, 136(3), 199. https://doi.org/10.1016/j.ymgme.2022.05.002

Chamoso-Sánchez, D., Rabadán, F., Argente, J., Barbas, C., Martos-Moreno, G. Á., & Rupérez, F. J. (2023). Identifying subgroups of childhood obesity by using multiplatform metabotyping. Frontiers in Molecular Biosciences, 10. https://doi.org/10.3389/fmolb.2023.1301996

Chaturvedi, S., Singh, A. K., Keshari, A. K., Maity, S., Sarkar, S., & Saha, S. (2016). Human Metabolic Enzymes Deficiency: A Genetic Mutation Based Approach. Scientifica, 2016, 1. https://doi.org/10.1155/2016/9828672

Cheng, Y., Schlosser, P., Hertel, J., Sekula, P., Oefner, P. J., Spiekerkoetter, U., Mielke, J., Freitag, D. F., Schmidts, M., Oefner, P. J., Kronenberg, F., Eckardt, K., Kronenberg, F., Eckardt, K., Thiele, I., Li, Y., & Köttgen, A. (2021). Rare genetic variants affecting urine metabolite levels link population variation to inborn errors of metabolism. Nature Communications, 12(1). https://doi.org/10.1038/s41467-020-20877-8

Cossu, M., Pintus, R., Zaffanello, M., Mussap, M., Serra, F., Marcialis, M. A., & Fanos, V. (2023). Metabolomic Studies in Inborn Errors of Metabolism: Last Years and Future Perspectives. Metabolites,

13(3), 447. https://doi.org/10.3390/metabo13030447

Exploring the human microbiome: its role in health, disease, and therapeutic interventions. (2024). Advances in Biosciences Research, 1(01), 10-17. https://advbioresearch.com/index.php/ABR/article/view/2

Forny, P., Bonilla, X., Lamparter, D., Shao, W., Plessl, T., Frei, C., Bingisser, A., Goetze, S., Drogen, A. van, Harshman, K., Pedrioli, P. G. A., Howald, C., Poms, M., Traversi, F., Bürer, C., Cherkaoui, S., Morscher, R. J., Simmons, L., Forny, M., … Froese, D. S. (2023). Integrated multi-omics reveals anaplerotic rewiring in methylmalonyl-CoA mutase deficiency. Nature Metabolism, 5(1), 80. https://doi.org/10.1038/s42255-022-00720-8

Friedman, H., García-Contreras, A., Vásquez-Garibay, E. M., Godínez-Méndez, L. A., Rohilla, S., Vikas, Y., Wadhwa, D., Singla, D., Anurev, A., Горбачев, В. И., Pavlova, T., Pavlov, A., Silva, A. R. D., & Silva, J. R. da. (2023). Neuropediatrics - Recent Advances and Novel Therapeutic Approaches. IntechOpen. https://doi.org/10.5772/intechopen.1000299

Furuta, Y., Tinker, R. J., Hamid, R., Cogan, J. D., Ezell, K., Oglesbee, D., DeBerardinis, R. J., & Phillips, J. A. (2024). A Review of Multiple Diagnostic Approaches in the Undiagnosed Diseases Network to Identify Inherited Metabolic Diseases [Review of A Review of Multiple Diagnostic Approaches in the Undiagnosed Diseases Network to Identify Inherited Metabolic Diseases]. Research Square. https://doi.org/10.21203/rs.3.rs-4588411/v1

Guéant, J., & Feillet, F. (2022). Inherited metabolic disorders beyond the new generation sequencing era: the need for in-depth cellular and molecular phenotyping. Human Genetics, 141(7), 1235. https://doi.org/10.1007/s00439-022-02467-4

Guharoy, M., Adant, I., Bird, M., Pedrayes, A. J., Botzki, A., Dehairs, J., Derveaux, S., Devos, S., Goeminne, G., Impens, F., Janky, R., Maes, R., Maia, T. M., Meersseman, W., Rymen, D., Swinnen, J. V., Haver, D. V., Witters, P., Cassiman, D., & Ghesquière, B. (2023). Integrating multiple omics levels using the human protein complexome as a framework, a multi-omics study of inborn errors of metabolism. https://doi.org/10.1101/2023.12.30.573613

Hertzog, A., Selvanathan, A., Farnsworth, E., Tchan, M., Adams, L., Lewis, K., Tolun, A. A., Bennetts, B., Ho, G., & Bhattacharya, K. (2022). Intronic variants in inborn errors of metabolism: Beyond the exome. Frontiers in Genetics, 13. https://doi.org/10.3389/fgene.2022.1031495

Hollander, B. den, Konijnenburg, E. M. M. H. van, Hewitson, B., Meijden, J. C. van der, Balfoort, B. M., Winter, B., Müller, A. R., Wasserman, W. W., Ferreira, C. R., & Karnebeek, C. D. M. van. (2025). The Metabolic Treatabolome and Inborn Errors of Metabolism Knowledgebase therapy tool: Do not miss the opportunity to treat! Journal of Inherited Metabolic Disease, 48(1). https://doi.org/10.1002/jimd.12835

Ijaz, A., Abbas, S., Shabbir, M., Badshah, Y., Abid, F., Afsar, T., & Razak, S. (2025). Inherited metabolic disorders: presentation, clinical types, laboratory diagnosis and genetic markers. Orphanet Journal of Rare Diseases, 20(1), 422. https://doi.org/10.1186/s13023-025-03979-8

Kernohan, K. D., & Boycott, K. M. (2024). The expanding diagnostic toolbox for rare genetic diseases. Nature Reviews Genetics, 25(6), 401. https://doi.org/10.1038/s41576-023-00683-w

Kerr, K., McAneney, H., Smyth, L. J., Bailie, C., McKee, S., & McKnight, A. J. (2020). A scoping review and proposed workflow for multi-omic rare disease research [Review of A scoping review and proposed workflow for multi-omic rare disease research]. Orphanet Journal of Rare Diseases, 15(1). https://doi.org/10.1186/s13023-020-01376-x

Kopajtich, R., Smirnov, D., Stenton, S. L., Loipfinger, S., Meng, C., Scheller, I. F., Freisinger, P., Baski, R., Berutti, R., Behr, J., Bucher, M., Distelmaier, F., Graf, E., Gušić, M., Hempel, M., Kulterer, L., Mayr, J. A., Meitinger, T., Mertes, C., … Prokisch, H. (2021). Integration of proteomics with genomics and transcriptomics increases the diagnostic rate of Mendelian disorders. medRxiv. https://doi.org/10.1101/2021.03.09.21253187

Labory, J., Fierville, M., Ait-El-Mkadem, S., Bannwarth, S., Paquis-Flucklinger, V., & Bottini, S. (2020). Multi-Omics Approaches to Improve Mitochondrial Disease Diagnosis: Challenges, Advances, and Perspectives. Frontiers in Molecular Biosciences, 7. https://doi.org/10.3389/fmolb.2020.590842

Lunke, S., Bouffler, S. E., Patel, C., Sandaradura, S. A., Wilson, M., Pinner, J., Hunter, M. F., Barnett, C., Wallis, M., Kamien, B., Tan, T. Y., Freckmann, M., Chong, B., Phelan, D., Francis, D., Kassahn, K. S., Ha, T., Gao, S., Arts, P., … Stark, Z. (2023). Integrated multi-omics for rapid rare disease diagnosis on a national scale. Nature Medicine, 29(7), 1681. https://doi.org/10.1038/s41591-023-02401-9

Marwaha, S., Knowles, J. W., & Ashley, E. A. (2022). A guide for the diagnosis of rare and undiagnosed disease: beyond the exome [Review of A guide for the diagnosis of rare and undiagnosed disease: beyond the exome]. Genome Medicine, 14(1), 23. https://doi.org/10.1186/s13073-022-01026-w

Mukherjee, A., Abraham, S., Singh, A., Balaji, S., & Mukunthan, K. S. (2024). From Data to Cure: A Comprehensive Exploration of Multi-omics Data Analysis for Targeted Therapies [Review of From Data to Cure: A Comprehensive Exploration of Multi-omics Data Analysis for Targeted Therapies]. Molecular Biotechnology. https://doi.org/10.1007/s12033-024-01133-6

Mussap, M., Zaffanello, M., & Fanos, V. (2018). Metabolomics: a challenge for detecting and monitoring inborn errors of metabolism. Annals of Translational Medicine, 6(17), 338. https://doi.org/10.21037/atm.2018.09.18

Ng, E., Nanjala, R., Jostins-Dean, L., Barrett, J. C., Todd, J. A., & Luo, Y. (2025). Improving population-scale disease prediction through multi-omics integration. https://doi.org/10.1101/2025.11.24.25340920

Peña-Chilet, M., Esteban-Medina, M., Falco, M. M., Rian, K., Hidalgo, M. R., Loucera, C., & Dopazo, J. (2019). Using mechanistic models for the clinical interpretation of complex genomic variation. Scientific Reports, 9(1). https://doi.org/10.1038/s41598-019-55454-7

Scherer, N., Fässler, D., Borisov, O., Cheng, Y., Schlosser, P., Wuttke, M., Haug, S., Li, Y., Telkämper, F., Patil, S., Meiselbach, H., Wong, C. S. F., Berger, U., Sekula, P., Hoppmann, A., Schultheiß, U. T., Mozaffari, S. V., Xi, Y., Graham, R., … Köttgen, A. (2025). Coupling metabolomics and exome sequencing reveals graded effects of rare damaging heterozygous variants on gene function and human traits. Nature Genetics, 57(1), 193. https://doi.org/10.1038/s41588-024-01965-7

Scherer, N., Fässler, D., Borisov, O., Cheng, Y., Schlosser, P., Wuttke, M., Patil, S., Meiselbach, H., Telkämper, F., Berger, U., Grünert, S. C., Sekula, P., Schultheiß, U. T., Li, Y., Köttgen, M., Oefner, P. J., Knauf, F., Eckardt, K., Thiele, I., … Köttgen, A. (2023). Coupling of metabolomics and exome sequencing reveals graded effects of rare damaging heterozygous variants on gene function and resulting traits and diseases. medRxiv. https://doi.org/10.1101/2023.10.17.23297094

Schlieben, L. D., Prokisch, H., & Yépez, V. A. (2021). How Machine Learning and Statistical Models Advance Molecular Diagnostics of Rare Disorders Via Analysis of RNA Sequencing Data [Review of How Machine Learning and Statistical Models Advance Molecular Diagnostics of Rare Disorders Via Analysis of RNA Sequencing Data]. Frontiers in Molecular Biosciences, 8. https://doi.org/10.3389/fmolb.2021.647277

Sen, P., & Orešič, M. (2023). Integrating Omics Data in Genome-Scale Metabolic Modeling: A Methodological Perspective for Precision Medicine. Metabolites, 13(7), 855. https://doi.org/10.3390/metabo13070855

Smirnov, D., Konstantinovskiy, N., & Prokisch, H. (2023). Integrative omics approaches to advance rare disease diagnostics. Journal of Inherited Metabolic Disease, 46(5), 824. https://doi.org/10.1002/jimd.12663

Stranneheim, H., Engvall, M., Naess, K., Lesko, N., Larsson, P., Dahlberg, M., Andeer, R., Wredenberg, A., Freyer, C., Barbaro, M., Bruhn, H., Emahazion, T., Magnusson, M., Wibom, R., Zetterström, R., Wirta, V., Döbeln, U. von, & Wedell, A. (2014). Rapid pulsed whole genome sequencing for comprehensive acute diagnostics of inborn errors of metabolism. BMC Genomics, 15(1), 1090. https://doi.org/10.1186/1471-2164-15-1090

Sun, Y. V., & Hu, Y. (2016). Integrative Analysis of Multi-omics Data for Discovery and Functional Studies of Complex Human Diseases [Review of Integrative Analysis of Multi-omics Data for Discovery and Functional Studies of Complex Human Diseases]. Advances in Genetics, 147. https://doi.org/10.1016/bs.adgen.2015.11.004

Tebani, A., Afonso, C., Marret, S., & Bekri, S. (2016). Omics-Based Strategies in Precision Medicine: Toward a Paradigm Shift in Inborn Errors of Metabolism

Investigations. International Journal of Molecular Sciences, 17(9), 1555. https://doi.org/10.3390/ijms17091555

Tejedor, J. R., Soriano-Sexto, A., Beccari, L., Castejón-Fernández, N., Medina, P. C., Sainz-Ledo, L., Alba-Linares, J. J., Urdinguio, R. G., Ugarte, M., Fernández, A. F., Rodríguez-Pombo, P., Fraga, M. F., & Pérez, B. (2024). Integration of multi-omics layers empowers precision diagnosis through unveiling pathogenic mechanisms on maple syrup urine disease. Journal of Inherited Metabolic Disease, 48(1). https://doi.org/10.1002/jimd.12829

Wang, D., Zhang, J., Yang, R., Zhang, D., Wang, M., Yu, C., Yang, J., Huang, W., Liu, S., Tang, S., & He, X. (2024). Disease spectrum, prevalence, genetic characteristics of inborn errors of metabolism in 21,840 hospitalized infants in Chongqing, China, 2017-2022. Frontiers in Genetics, 15. https://doi.org/10.3389/fgene.2024.1395988

Wang, W., Yang, J., Xue, J., Mu, W., Zhang, X., Wang, W., Xu, M., Gong, Y., Liu, Y., Zhang, Y., Xie, X., Gu, W., Bai, J., & Cram, D. S. (2019). A comprehensive multiplex PCR based exome-sequencing assay for rapid bloodspot confirmation of inborn errors of metabolism. BMC Medical Genetics, 20(1). https://doi.org/10.1186/s12881-018-0731-5

Waters, D., Adeloye, D., Woolham, D., Wastnedge, E., Patel, S., & Rudan, I. (2018). Global birth prevalence and mortality from inborn errors of metabolism: a systematic analysis of the evidence [Review of Global birth prevalence and mortality from inborn errors of metabolism: a systematic analysis of the evidence]. Journal of Global Health, 8(2). https://doi.org/10.7189/jogh.08.021102

Wojcik, M. H., Reuter, C. M., Marwaha, S., Mahmoud, M., Duyzend, M., Barseghyan, H., Yuan, B., Boone, P. M., Groopman, E., Délot, E. C., Jain, D., Sanchis-Juan, A., Diseases, G. R. to E. the G. of R., Consortium, Starita, L. M., Talkowski, M. E., Montgomery, S. B., Bamshad, M. J., Chong, J. X., … Miller, D. E. (2023). Beyond the exome: what's next in diagnostic testing for Mendelian conditions. arXiv. https://doi.org/10.48550/arxiv.2301.07363

Wörheide, M. A., Krumsiek, J., Kastenmüller, G., & Arnold, M. (2020). Multi-omics integration in biomedical research – A metabolomics-centric review. Analytica Chimica Acta, 1141,

144. https://doi.org/10.1016/j.aca.2020.10.038

Wurth, R., Turgeon, C., Stander, Z., & Oglesbee, D. (2023). An evaluation of untargeted metabolomics methods to characterize inborn errors of metabolism. Molecular Genetics and Metabolism, 141(1), 108115. https://doi.org/10.1016/j.ymgme.2023.108115

Yubero, D., Brandi, N., Ormazábal, A., García-Cazorla, À., Pérez-Dueñas, B., Campistol, J., Ribes, A., Palau, F., Artuch, R., & Armstrong, J. (2016). Targeted Next Generation Sequencing in Patients with Inborn Errors of Metabolism. PLoS ONE, 11(5). https://doi.org/10.1371/journal.pone.0156359