



NETWORK-BASED ANALYSIS OF PKU DISEASE-ASSOCIATED GENES USING WEB-BASED BIOINFORMATICS TOOLS

Suparna Deepak* and Bhumi Dere

Department of Biotechnology,

Pillai College of Arts, Commerce & Science (Empowered Autonomous),

New Panvel, Navi Mumbai, Maharashtra

*Corresponding author E-mail: suparnadeepak@mes.ac.in

Received: 12 December 2025

Revised: 12 January 2026

Accepted: 18 February 2026

Published: 27 February 2026

DOI: <https://doi.org/10.5281/zenodo.18898393>

Abstract:

Phenylketonuria (PKU) is an inherited metabolic disorder mainly due to mutations in the PAH gene, leading to the impaired metabolism of phenylalanine to tyrosine and the buildup of neurotoxic compounds. This study used a network-based bioinformatics approach to analyze genes, protein interactions, and biological pathways associated with PKU. The study used web-based tools to analyze gene-disease associations, protein interactions, functional enrichment, and pathway analysis. The genes that were identified as key to PKU include PAH, QDPR, PTS, GCH1, SPR, PCBD1, TH, TPH1, MAOB, SOD1, CAT, CPS1, OTC, and SLC6A19. These genes are involved in amino acid metabolism, tetrahydrobiopterin (BH₄) synthesis, neurotransmitter metabolism, oxidative stress regulation, and nitrogen metabolism. The results showed that PKU is a multi-pathway metabolic disorder and not an enzyme deficiency disorder. The results of this study show that PKU is a multi-pathway metabolic disorder and not an enzyme deficiency disorder. The bioinformatics approach used in this study provides a systems-level understanding of PKU pathogenesis.

Keywords: Phenylketonuria (PKU), PAH Gene, Phenylalanine Metabolism, Tetrahydrobiopterin (BH₄) Metabolism, Protein-Protein Interaction (PPI) Network.

Introduction

Phenylketonuria (PKU) is a traditional autosomal recessive congenital disorder of amino acid metabolism due to predominantly pathogenic variants in the phenylalanine hydroxylase (PAH) gene, which catalyzes the conversion of phenylalanine to tyrosine (1, 2). Reduced PAH activity causes the accumulation of phenylalanine and neurotoxicity, leading to intellectual disability, developmental delay, seizures, and behavioral problems if left untreated (3, 4). PKU has been shown to have high genetic heterogeneity, with more than a thousand variants of the PAH gene reported, resulting in a wide spectrum of phenotypes (5, 6, 7). Moreover, mutations in

tetrahydrobiopterin (BH₄) metabolism genes such as PTS, GCH1, QDPR, and PCBD1 have been shown to cause PKU-like phenotypes, thus underlining the role of cofactor-dependent pathways (8, 9, 10).

Conventional genetic methods such as mutation analysis and genotype-phenotype correlation analysis have enhanced diagnostic and therapeutic capabilities but have not provided much information regarding the role of molecular interactions in the variability of the disease phenotype (11, 12, 13). Evidence is emerging that modifier genes and pathway-level interactions are involved in the phenotypic variability of PKU (13, 14, 15).

Network-based systems biology tools allow the analysis of genes in the context of a network of molecular interactions, making it easier to identify functional modules and disease-related subnetworks (16, 17). Disease genes are often found to be grouped in particular biological pathways, and network analysis has successfully identified new functional associations in both Mendelian and complex diseases (18-22).

Web-based bioinformatics tools such as DisGeNET, STRING, GeneMANIA, and Cytoscape facilitate the identification and network analysis of disease-related genes (23-27). Thus, in this study, the network-based bioinformatics tools are used to investigate the PKU-related genes and attempt to clarify their functional associations and offer a systems-level perspective of the molecular structure of PKU.

Methodology

1. Gene identification

The first step in our study was to identify genes associated with PKU. For this, DisGeNET was used, which is a comprehensive, curated platform integrating information on human gene disease associations (23). We retrieved genes reported to be linked with PKU as well as related disorders involving tetrahydrobiopterin (BH₄) metabolism, such as mutations in PTS, GCH1, QDPR, and PCBD1. Each gene was selected based on high-confidence evidence, including peer-reviewed publications and curated database entries. This ensured that only genes with established or strongly suggested relevance to PKU were included for subsequent network and functional analyses.

2. Protein-Protein Interaction (PPI) network construction

To explore interactions among PKU-associated genes, protein-protein interaction networks were constructed. STRING v11 (24) and GeneMANIA (25) were used to generate PPI networks. STRING provides predicted and known protein interactions derived from experimental data, text mining, co-expression, and database annotations. GeneMANIA complements this by predicting functional associations, co-localization, and pathway co-membership. High-confidence interactions were selected to reduce false positives. The networks were imported into Cytoscape v3.9.1 (26), which enabled detailed visualization and topological analysis. Network metrics, including degree, betweenness centrality, and clustering coefficient, were calculated to identify hub proteins and highly connected nodes that may play critical roles in PKU pathology.

3. Functional enrichment analysis

Functional characterization of the identified genes was performed using Enrichr, a widely used gene set enrichment analysis platform (27, 28). Enrichment analyses were conducted for biological processes, molecular functions, and cellular components based on Gene Ontology (GO) terms, as well as pathway enrichment using KEGG and Reactome databases. Significantly enriched terms were selected based on adjusted p-values (<0.05) to ensure statistical reliability. This step provided insights into the biological processes and pathways most relevant

to PKU-associated genes, such as amino acid metabolism, neurotransmitter biosynthesis, and BH₄-related pathways.

4. Pathway mapping

To further understand the biological context, the curated gene set was mapped onto metabolic and signaling pathways using the Reactome Pathway Browser (29). Reactome provides high-quality, manually curated pathways for human biological processes. Key pathways involved in phenylalanine metabolism, BH₄ biosynthesis, and associated signaling cascades were highlighted. This mapping enabled identification of functional clusters and potential pathway-level interactions that could influence disease severity or phenotype variability in PKU patients.

5. Network Integration and Visualization

Finally, all network and enrichment data were integrated and visualized using Cytoscape. Functional modules, hub genes, and pathway interconnections were highlighted to provide a comprehensive, systems-level perspective of PKU-associated genes. The visualization allowed easy identification of key regulatory nodes, potentially novel modifier genes, and clusters of interacting proteins that could serve as targets for further research or therapeutic intervention. By combining PPI analysis, functional enrichment, and pathway mapping, this approach provides a holistic understanding of the molecular landscape of PKU.

Result

1. DisGeNET

Table 1: DisGeNET-based gene–disease association table showing identified genes related to PKU

Gene Symbol	Gene Description	Score
PAH	phenylalanine hydroxylase	1.0
QDPR	quinoa dihydropteridine reductase	0.8
PTS	6-pyruvoyltetrahydropterin synthase	0.7
LOC126861615	CDK7 strongly-dependent group 2 enhancer GRCh37_chr12:103244689-103245888	0.4
NSUN2	NOP2/Sun RNA methyltransferase 2	0.4
TSPAN1	tetraspanin 1	0.4
POMGNT1	protein O-linked mannose N-acetylglucosaminyltransferase 1 (beta 1,2-)	0.4
COL1A1	collagen type I alpha 1 chain	0.4
PI4KA	phosphatidylinositol 4-kinase alpha	0.4
CAT	catalase	0.3
TH	tyrosine hydroxylase	0.25
LRIT1	leucine rich repeat, Ig-like and transmembrane domains 1	0.25
TTR	transthyretin	0.25
SOD1	superoxide dismutase 1	0.25
TPH1	tryptophan hydroxylase 1	0.25
SPR	sepiapterin reductase	0.25
SLC6A19	solute carrier family 6 member 19	0.25

ALB	albumin	0.25
LAT	linker for activation of T cells	0.2
STIN2-VNTR	serotonin transporter intronic VNTR enhancer	0.2
OTC	ornithine transcarbamylase	0.2
CPS1	carbamoyl-phosphate synthase 1	0.2
PRL	prolactin	0.2
GCH1	GTP cyclohydrolase 1	0.2
APP	amyloid beta precursor protein	0.2
PCBD1	pterin-4 alpha-carbinolamine dehydratase 1	0.2
DNAJC12	DnaJ heat shock protein family (Hsp40) member C12	0.2
CTSZ	cathepsin Z	0.2
MAOB	monoamine oxidase B	0.2
SLC25A13	solute carrier family 25 member 13	0.2

DisGeNET analysis identified PAH as the most strongly associated gene with Phenylketonuria (score = 1.0), confirming its central role in PKU pathogenesis. Other high-confidence genes such as QDPR, PTS, SPR, GCH1, and PCBD1 were also identified, highlighting the involvement of the tetrahydrobiopterin (BH₄) metabolic pathway. Additionally, genes related to amino acid metabolism, oxidative stress, and neurotransmitter synthesis, including TH, TPH1, SOD1, CAT, and MAOB, were observed. The presence of genes involved in liver function and transport processes, such as ALB, SLC6A19, and CPS1, suggests a broader metabolic impact of PKU. Overall, the DisGeNET results demonstrate that PKU is a multigenic metabolic disorder, involving not only phenylalanine 6 but also oxidative stress regulation and neurotransmitter biosynthesis.

2. STRING

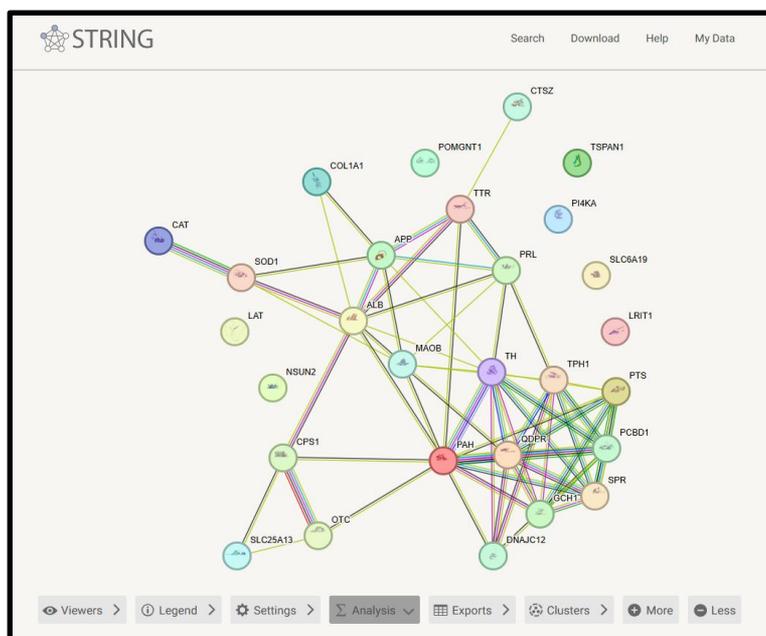


Figure 1: STRING Protein-Protein Interaction (PPI) Network of PKU-Associated Genes

Showed a strongly interconnected protein network centered on PAH, highlighting its key role in PKU. Enrichment analysis revealed significant involvement of phenylalanine metabolism, tetrahydrobiopterin biosynthesis, and neurotransmitter-related pathways, confirming that PKU arises from coordinated metabolic network disruptions.

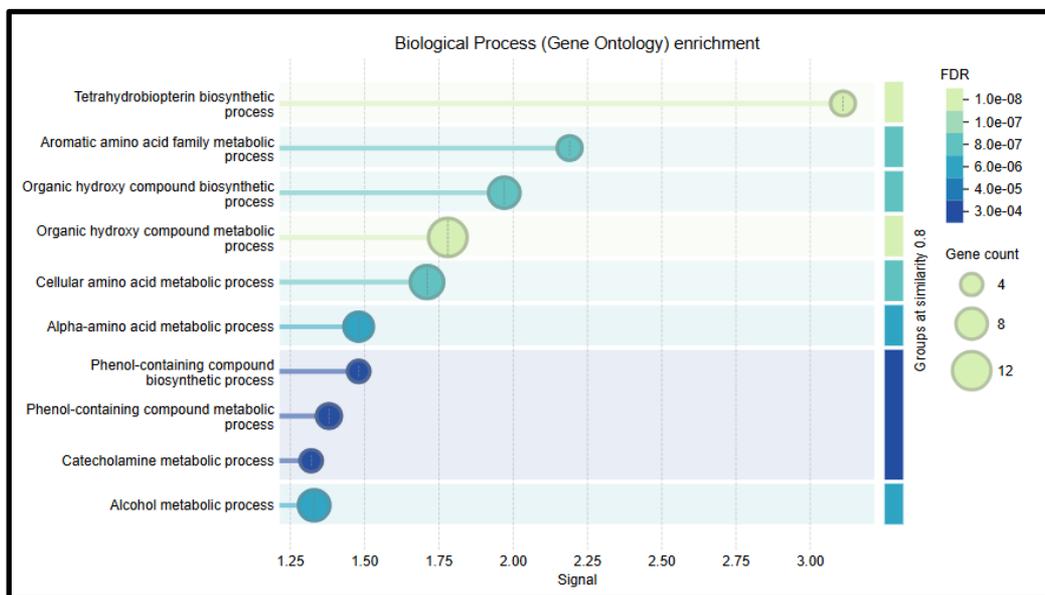


Figure 2: Gene Ontology (GO) Biological Process enrichment bubble plot

The most strongly enriched process is tetrahydrobiopterin (BH₄) biosynthesis, indicating a major role of cofactor metabolism. Other highly enriched terms include aromatic amino acid metabolism, cellular amino acid metabolism, and alpha-amino acid metabolic processes, highlighting the central involvement of amino acid biochemical pathways. Enrichment of phenol-containing compound metabolism, catecholamine metabolism, and organic hydroxy compound metabolism suggests participation in neurotransmitter synthesis and related metabolic reactions. The presence of alcohol and hydroxy compound metabolic processes reflects broader biochemical detoxification and intermediary metabolism.

3. GeneMANIA

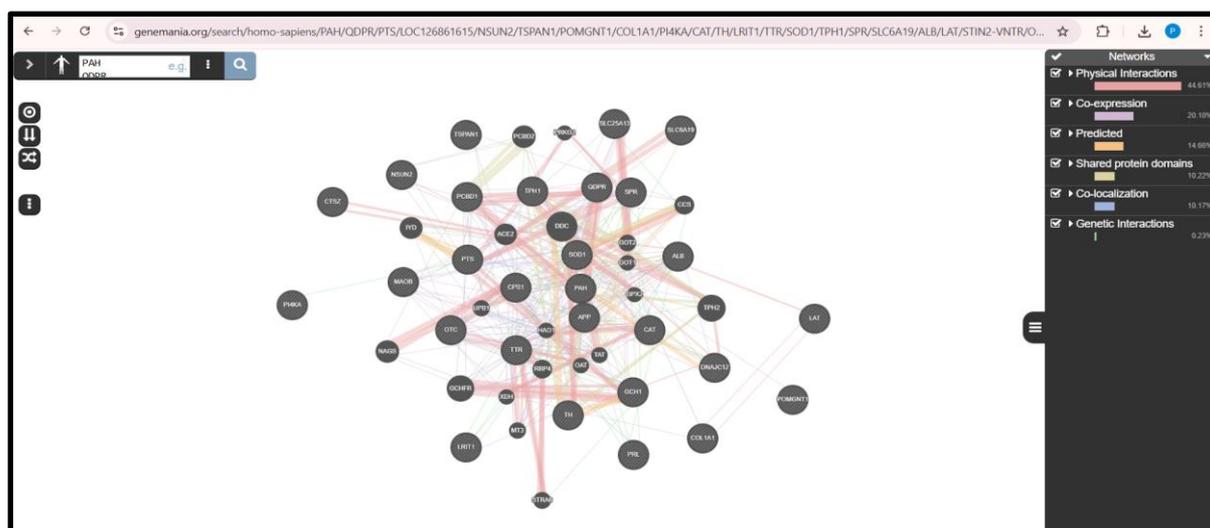


Figure 3: GeneMANIA Functional Interaction Network of PKU Genes

The network reveals a highly interconnected hub, indicating that the genes are working together as a group rather than individually. Metabolic and enzyme genes like PAH, QDPR, PTS, GCH1, and PCBD1 are found to be centrally located, emphasizing their role in cofactor metabolism and amino acid metabolism. Oxidative stress and cell protection genes like CAT and SOD1 are also found to be highly interconnected, emphasizing their regulatory role. Neurotransmitter genes like TH, TPH1, and MAOB are found to have interaction relationships, emphasizing their coordinated regulation. Transport and metabolic regulators like SLC6A19, SLC25A13, CPS1, and OTC also play a role in the functional network, emphasizing their role in cellular metabolism and biochemical homeostasis.

4. Enrichr

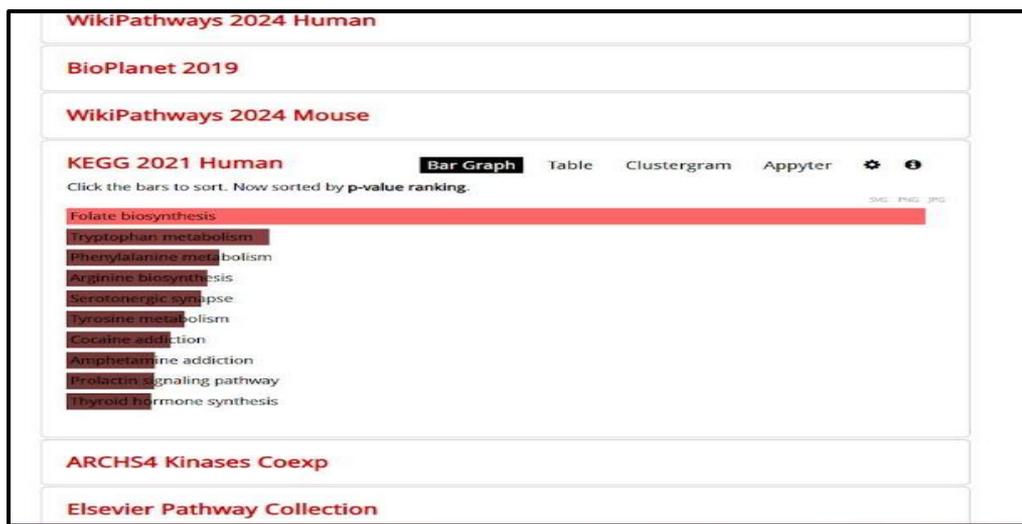


Figure 4: Enrichr Transcription Factor and Regulatory Enrichment Analysis

The figure depicts KEGG 2021 Human pathway enrichment analysis (bar graph representation) of the significantly enriched metabolic and signaling pathways related to the uploaded gene set. The most significantly enriched pathway is Folate biosynthesis, suggesting a strong link to one-carbon metabolism and cofactor biosynthesis. Other significantly enriched pathways include Tryptophan metabolism, Phenylalanine metabolism, Tyrosine metabolism, and Arginine biosynthesis, suggesting a broad link to amino acid metabolism. Neurotransmitter-related pathways such as Serotonergic synapse are also significantly enriched, suggesting a link to neurochemical signaling. Other significantly enriched pathways include Prolactin signaling pathway and Thyroid hormone synthesis, suggesting a possible link to hormonal regulation. The presence of pathways such as Cocaine addiction and Amphetamine addiction suggests a link to dopaminergic and serotonergic neurotransmission mechanisms rather than addiction to these substances.

5. REACTOME

Reactome pathway analysis revealed strong enrichment of phenylalanine metabolism, amino acid catabolism, and tetrahydrobiopterin (BH₄) biosynthesis and recycling pathways, confirming the core biochemical basis of phenylketonuria (PKU). Key genes including PAH, GCH1, PTS, QDPR, SPR, and PCBD1 were mapped to these pathways, highlighting their roles in phenylalanine hydroxylation and cofactor regeneration. Disruption of these processes leads to impaired conversion of phenylalanine to tyrosine and subsequent phenylalanine accumulation, a hallmark of PKU.

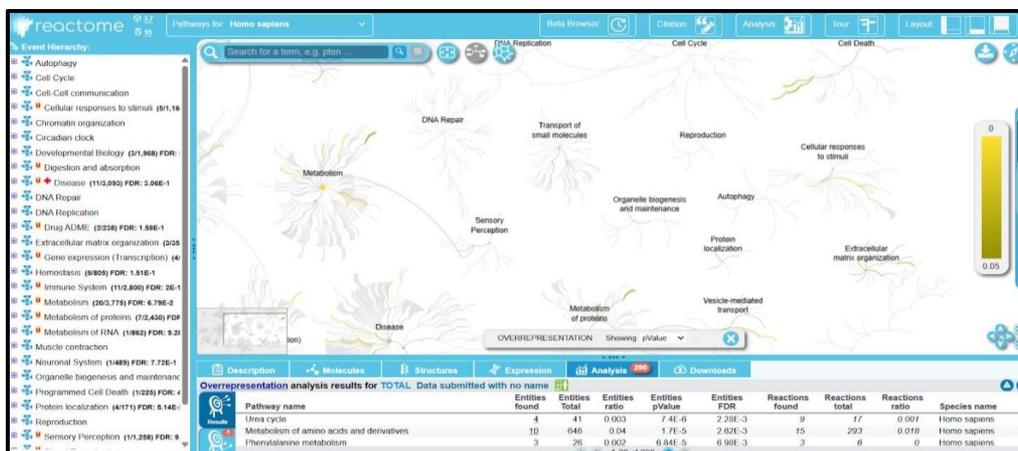


Figure 5: Reactome pathway enrichment analysis of PKU-associated genes.

The analysis also showed enrichment of neurotransmitter synthesis pathways involving genes such as TH, TPH1, and MAOB, explaining the neurological and cognitive manifestations observed in untreated patients. Additionally, pathways related to nitrogen metabolism and the urea cycle (CPS1, OTC) suggest secondary metabolic disturbances associated with altered amino acid homeostasis.

Discussion

Phenylketonuria (PKU) is a classical example of an inborn error of metabolism in which disruption of a single enzymatic step leads to widespread metabolic and neurological consequences. In the present bioinformatics-based study, pathway enrichment and network analysis highlighted the central role of phenylalanine metabolism and its associated cofactor-dependent reactions, providing molecular support to the established biochemical understanding of PKU. The Reactome pathway analysis demonstrated a strong enrichment of phenylalanine and tyrosine metabolic pathways, primarily driven by the involvement of the PAH gene. This finding is consistent with earlier genetic and biochemical studies reporting that loss of phenylalanine hydroxylase activity results in the accumulation of phenylalanine and reduced tyrosine synthesis (2). The localization of PAH at a critical metabolic junction reinforces its role as the principal determinant of disease severity and treatment responsiveness.

In addition to PAH, genes associated with tetrahydrobiopterin (BH₄) biosynthesis and regeneration (GCH1, PTS, QDPR, and SPR) were significantly enriched. Previous studies have shown that deficiencies in BH₄ metabolism can mimic classical PKU by functionally reducing PAH activity (1). The enrichment of these pathways in the present analysis supports the concept that PKU is not exclusively an enzyme deficiency but may also arise from impaired cofactor availability. This finding is in agreement with bioinformatics-based analyses that emphasize the importance of cofactor pathways in metabolic disorders (30).

Furthermore, pathways related to neurotransmitter biosynthesis, particularly dopamine and serotonin metabolism, were identified through the involvement of genes such as TH, TPH1, and MAOB. Reduced levels of tyrosine and BH₄ are known to limit monoamine neurotransmitter production, which explains the neurological and cognitive impairments observed in untreated PKU patients (31). Similar associations between amino acid imbalance and neurotransmitter dysregulation have been reported in pathway-based studies using KEGG and Reactome databases (32). The enrichment of nitrogen metabolism and urea cycle-associated genes, including CPS1 and OTC, suggests secondary metabolic stress resulting from disrupted amino acid homeostasis. Although

PKU is not classified as a primary urea cycle disorder, elevated phenylalanine levels can indirectly influence nitrogen balance, a phenomenon also reported in metabolomic and systems biology studies (33). This highlights the interconnected nature of metabolic pathways and the systemic impact of a single enzymatic defect. When compared with earlier studies that focused solely on gene mutation analysis or clinical biomarkers, the present bioinformatics approach offers a broader systems-level perspective. Unlike mutation-centric studies, pathway enrichment analysis integrates multiple genes and reactions, allowing the identification of both primary and secondary metabolic effects. Similar advantages of pathway-based analyses have been reported in studies utilizing KEGG, STRING, and Reactome for inherited metabolic disorders (34, 35).

The findings of this study align well with existing literature while providing additional insights into the multi-pathway involvement in PKU. The integration of bioinformatics tools enables a comprehensive understanding of disease mechanisms, supports genotype-phenotype correlations, and may assist in identifying novel therapeutic targets beyond conventional dietary management.

CONCLUSION

This bioinformatics analysis offers a pathway-level insight into the mechanisms of phenylketonuria (PKU) through gene enrichment and pathway analysis with Enrichr and Reactome. The data clearly support that the primary pathological mechanism is the disruption of phenylalanine metabolism, mainly because of PAH gene alterations. The enrichment of tetrahydrobiopterin (BH₄) biosynthesis and regeneration pathways, including genes such as GCH1, PTS, QDPR, and SPR, suggests that cofactor-dependent regulation is an important factor in the progression of the disease. The changes in neurotransmitter biosynthesis pathways further clarify the neurological symptoms caused by high levels of phenylalanine, while secondary enrichment of nitrogen metabolism and urea cycle pathways suggest that there is a disturbance in the whole metabolic system.

The results clearly show that PKU is a multi-pathway metabolic disorder, not a simple enzyme deficiency. This study illustrates the importance of bioinformatics analysis in correlating genetic information with functional pathways, which can be used to develop better understanding and treatment strategies for the disease.

References

1. Blau, N., van Spronsen, F. J., & Levy, H. L. (2010). Phenylketonuria. *The Lancet*, 376(9750), 1417–1427. [https://doi.org/10.1016/S0140-6736\(10\)60961-0](https://doi.org/10.1016/S0140-6736(10)60961-0)
2. Scriver, C. R., & Kaufman, S. (2001). Hyperphenylalaninemia: Phenylalanine hydroxylase deficiency. In *The Metabolic and Molecular Bases of Inherited Disease* (8th ed., pp. 1667–1724). McGraw-Hill.
3. Güttler, F. (1980). Hyperphenylalaninemia: Diagnosis and classification of the various types. *Acta Paediatrica Scandinavica*, 69(1), 1–5. <https://pubmed.ncbi.nlm.nih.gov/7006896/>
4. Weglage, J., Pietsch, M., Feldmann, R., Koch, H. G., & Zschocke, J. (2001). Normal clinical outcome in early treated phenylketonuria: Results of a long-term study. *Molecular Genetics and Metabolism*, 73(2), 146–152. <https://pubmed.ncbi.nlm.nih.gov/11462238/>
5. Blau, N., van Spronsen, F. J., & Levy, H. L. (2014). Phenylketonuria. *The Lancet*, 376(9750), 1417–1427. <https://pubmed.ncbi.nlm.nih.gov/24756728/>
6. Waters, D., Adelman, W., & MacDonald, A. (2018). Genotype-phenotype correlations in phenylketonuria: Implications for clinical management. *Molecular Genetics and Metabolism Reports*, 15, 14–20.
7. Zschocke, J. (2003). Phenylketonuria mutations in Europe. *Human Mutation*, 21(4), 345–356.

8. Blau, N., Thöny, B., Cotton, R. G., & Hyland, K. (2001). Disorders of tetrahydrobiopterin and related biogenic amines. *Lancet*, 358, 141–147.
9. Opladen, T., Blau, N., & Hoffmann, G. F. (2012). Tetrahydrobiopterin deficiency: Pathophysiology, diagnosis, and treatment. *Journal of Inherited Metabolic Disease*, 35, 637–647.
10. Thöny, B., Auerbach, G., & Blau, N. (2000). Tetrahydrobiopterin biosynthesis, regeneration and functions. *Biochemical Journal*, 347(1), 1–16. <https://doi.org/10.1042/0264-6021:3470001>
11. Guldberg, P., Rey, F., Zschocke, J., et al. (1998). A European multicenter study of phenylalanine hydroxylase deficiency. *American Journal of Human Genetics*, 63, 71–79.
12. Kayaalp, E., Treacy, E., Waters, P. J., et al. (1997). Human phenylalanine hydroxylase mutations and hyperphenylalaninemia phenotypes. *American Journal of Human Genetics*, 61, 1309–1317.
13. Danecka, M. K., Woidy, M., Zschocke, J., et al. (2015). Functional and structural consequences of PAH mutations. *Human Mutation*, 36, 343–352.
14. Pey, A. L., Desviat, L. R., Gámez, A., et al. (2007). Phenylketonuria: genotype–phenotype correlations. *Human Mutation*, 28, 457–464.
15. Scriver, C. R. (2007). The PAH gene, phenylketonuria, and genotype–phenotype relationships. *Human Mutation*, 28, 831–845.
16. Barabási, A. L., & Oltvai, Z. N. (2004). Network biology: Understanding the cell's functional organization. *Nature Reviews Genetics*, 5, 101–113.
17. Barabási, A. L., Gulbahce, N., & Loscalzo, J. (2011). Network medicine: A network-based approach to human disease. *Nature Reviews Genetics*, 12, 56–68.
18. Goh, K. I., Cusick, M. E., Valle, D., et al. (2007). The human disease network. *Proceedings of the National Academy of Sciences*, 104, 8685–8690.
19. Menche, J., Sharma, A., Kitsak, M., et al. (2015). Uncovering disease–disease relationships through the incomplete interactome. *Science*, 347(6224), 1257601.
20. Cowen, L., Ideker, T., Raphael, B. J., & Sharan, R. (2017). Network propagation: A universal amplifier of genetic associations. *Nature Reviews Genetics*, 18, 551–562.
21. Loscalzo, J., & Barabási, A. L. (2011). Systems biology and the future of medicine. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 3, 619–627.
22. Vidal, M., Cusick, M. E., & Barabási, A. L. (2011). Interactome networks and human disease. *Cell*, 144, 986–998.
23. Piñero, J., Saüch, J., Sanz, F., & Furlong, L. I. (2020). DisGeNET: a comprehensive platform integrating information on human disease-associated genes. *Nucleic Acids Research*, 48(D1), D845–D855. <https://doi.org/10.1093/nar/gkz1021>
24. Szklarczyk, D., Gable, A. L., Lyon, D., et al. (2021). The STRING database in 2021: customizable protein–protein networks and functional characterization. *Nucleic Acids Research*, 49(D1), D605–D612. <https://doi.org/10.1093/nar/gkaa1074>
25. Warde-Farley, D., Donaldson, S. L., Comes, O., et al. (2010). The GeneMANIA prediction server: Biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Research*, 38(suppl_2), W214–W220. <https://doi.org/10.1093/nar/gkq537>

26. Shannon, P., Markiel, A., Ozier, O., et al. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research*, 13(11), 2498–2504. <https://doi.org/10.1101/gr.1239303>
27. Bader, G. D., & Hogue, C. W. V. (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics*, 4, 2.
28. Kuleshov, M. V., Jones, M. R., Rouillard, A. D., et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Research*, 44(W1), W90–W97. <https://doi.org/10.1093/nar/gkw377>
29. Fabregat, A., Sidiropoulos, K., Viteri, G., et al. (2018). Reactome pathway analysis: a high-performance in-memory approach. *BMC Bioinformatics*, 19, 142. <https://doi.org/10.1186/s12859-018-2207-2>
30. Thöny, B., Blau, N., & Bélanger-Quintana, A. (2014). Tetrahydrobiopterin deficiency and treatment. *Journal of Inherited Metabolic Disease*, 37(5), 739–748. <https://doi.org/10.1007/s10545-014-9710-4>
31. de Groot, M. J., Hoeksma, M., Blau, N., Reijngoud, D. J., & van Spronsen, F. J. (2015). Pathogenesis of cognitive dysfunction in phenylketonuria. *Molecular Genetics and Metabolism*, 99, S86–S93.
32. Hanley, W. B., et al. (2018). Amino acid metabolism and neurotransmitter imbalance in phenylketonuria. *Metabolic Brain Disease*, 33, 1505–1515.
33. Vockley, J., Andersson, H. C., Antshel, K. M., et al. (2014). Phenylalanine hydroxylase deficiency: diagnosis and management guideline. *Genetics in Medicine*, 16, 188–200.
34. Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M., & Tanabe, M. (2021). KEGG: Integrating viruses and cellular organisms. *Nucleic Acids Research*, 49(D1), D545–D551.
35. Williams, R. A., Amiotte, C. D. S., Burnett, J. R., & Smith, C. (2008). Phenylketonuria: An inborn error of phenylalanine metabolism. *Clinical Biochemistry*, 41(12), 1024–1031. <https://pubmed.ncbi.nlm.nih.gov/18420015/>