

GeDiPNet Shared-Gene Disease Pairs

Similarity: Jaccard index on shared curated genes (disease_gdp). Significance: upper-tail hypergeometric test ($P(X \geq k)$), Benjamini-Hochberg FDR-corrected across all tested pairs. Corroborated Shared Genes: count of shared genes backed by ≥ 2 distinct disease_gdp.main_source values combined across both diseases (e.g. GWAS-backed for one, ClinVar-backed for the other still counts); Shared Gene Names lists each gene's own source count in parentheses. Overlap Coefficient: shared genes / MIN(each disease's own total gene count) -- complements the Jaccard-based Similarity Score by showing how much of the SMALLER disease's entire known gene-set is covered by this pair, which Jaccard (dividing by the union) can understate. Shared Cluster: cluster id if both diseases fall in the same label-propagation cluster (scripts/build_disease_clusters.php), blank otherwise. See scripts/build_related.php on gedipnet.bicnirrh.res.in. Data as of: 2026-10-09 03:00:13. Filters applied: disease name contains "Hyperprolinemia" (5 matching pairs). · Rows: 5 · Generated: 2026-10-09 00:53:55

Disease A	Disease B	Shared Genes	Corroborated Shared Genes	Shared Gene Names	Similarity Score	Overlap Coefficient	P-value	FDR Q-value	Shared Cluster
Amino acid metabolism disorder	Hyperprolinemia	1	1	PRODH (6)	0.03571	0.50000	3.374e-3	4.372e-3	74
Hyperprolinemia	Nervous system disease	1	1	PRODH (6)	0.01429	0.50000	8.813e-3	1.016e-2	
Eating disorder	Hyperprolinemia	1	1	ALDH4A1 (7)	0.01351	0.50000	9.330e-3	1.063e-2	
Anorexia nervosa	Hyperprolinemia	1	1	ALDH4A1 (7)	0.00395	0.50000	3.234e-2	3.429e-2	
Hyperprolinemia	Nonalcoholic fatty liver disease	1	1	ALDH4A1 (7)	0.00314	0.50000	4.062e-2	4.250e-2	