

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: 2026-10-20 03:00:20. Filters applied: disease name contains "Polycystic liver disease 2" (4 matching pairs). · Rows: 4 · Generated: 2026-10-20 05:58:53

Disease A	Disease B	Shared Genes	Corroborated Shared Genes	Shared Gene Names	Similarity Score	Overlap Coefficient	P-value	FDR Q-value	Shared Cluster
Coronary artery disease	Polycystic liver disease 2	1	1	SEC63 (3)	0.00087	1.00000	7.373e-2	7.485e-2	
Polycystic kidney disease	Polycystic liver disease 2	1	1	SEC63 (4)	0.01961	1.00000	3.228e-3	4.073e-3	
Polycystic liver disease	Polycystic liver disease 2	1	1	SEC63 (7)	0.04167	1.00000	1.485e-3	2.122e-3	
Polycystic liver disease 2	Tubulointerstitial kidney disease	1	1	SEC63 (3)	0.08333	1.00000	7.102e-4	1.174e-3	