

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: 2026-10-11 03:00:14. Filters applied: disease name contains "Bile duct neoplasms" (5 matching pairs). · Rows: 5 · Generated: 2026-10-11 14:48:31

Disease A	Disease B	Shared Genes	Corroborated Shared Genes	Shared Gene Names	Similarity Score	Overlap Coefficient	P-value	FDR Q-value	Shared Cluster
Bile duct neoplasms	Leiomyoma	1	0	EGFR (1)	0.09091	1.00000	6.494e-4	1.241e-3	
Bile duct neoplasms	Papilloma	1	0	EGFR (1)	0.07692	1.00000	7.793e-4	1.405e-3	
Bile duct neoplasms	Giant cell glioblastoma	1	1	EGFR (2)	0.07143	1.00000	8.443e-4	1.499e-3	
Bile duct neoplasms	Gliosarcoma	1	1	EGFR (2)	0.05882	1.00000	1.039e-3	1.739e-3	
Bile duct neoplasms	Cerebral arteriovenous malformations	1	1	EGFR (2)	0.04348	1.00000	1.429e-3	2.226e-3	