Chemotherapeutic in biliary tract cancer: Application of network analysis

Edwin Kagereki

2022-02-25

Biliary tract cancer (BTC) or Cholangiocarcinoma (CC)

Cluster of highly heterogeneous & aggressive malignant tumours that can arise at any point of the biliary tree.

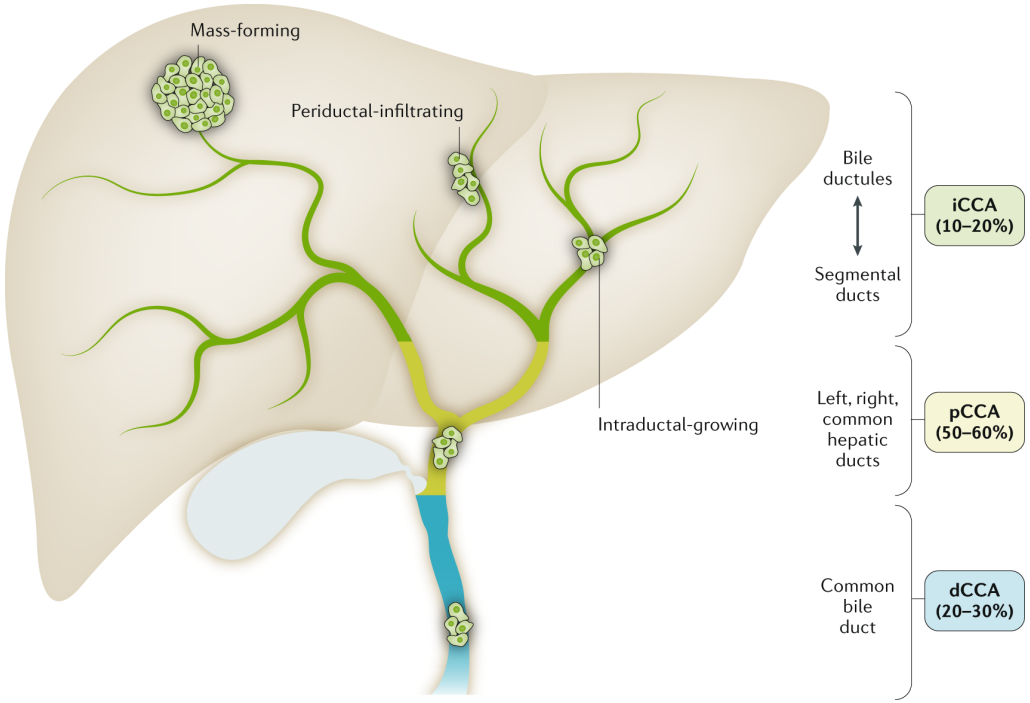


Fig 1.Anatomic classification of BTC. Reproduced from Banales et al [1]

Highlights:

- Low survival rates (5-year survival rate for metastatic disease being only 2%).
- Prevalence increasing globally (accounts for ~15% of all primary liver cancers and ~3% of gastrointestinal malignancies).
- High heterogeneity at the genomic, epigenetic and molecular levels severely compromises the efficacy of the available therapies.
- Limited treatment options.

Overview of BTC gene landscape

WHole genone sequencing studies have improved the understanding of the causal mechanisms in CCA, highlighting the genomic complexity in prevalent oncogenic modules affecting BTC.

1. DNA damage and genomic instability (TP53, CDKN2A, CCND1, ATM, ROBO2, BRCA1 and BRAC2);
2. MYC amplification; epigenetic regulation including NADPH metabolism (IDH1 and IDH2),
3. de-ubiquitination (BAP1), SWI–SNF complex (PBRM1, ARID1A, ARID1B, ARID2, SMARCA2, SMARCA4 and SMARCAD1)
4. histone (de-)methylation (MLL2, MML3, KMT2C, KDM4A, KDM5D, KDM6A and KDM6B);
5. kinase signalling (KRAS, ERBB1–3, BRAF, PIK3CA, PTEN, STK11, SMAD4 and FGFR1–3);
6. immune dysregulation (JAK–STAT3 signalling); FGFR2 and PRKCA–PRKCB fusions;
7. WNT–CTNNB1 pathway (APC); Hippo signalling (NF2, SAV1 deletion); METLL13 amplifications;

Mutational signature analysis in BTC

There are marked differences in the genomic features depending on the anatomical location and risk factors.

- Some examples of the differences in the mutations include:[1,2].

> Small bile duct iCCA can be characterized by isocitrate dehydrogenase (IDH1, IDH2) mutations or fibroblast growth factor receptor 2 (FGFR2) fusions.
>
> By contrast, large bile duct iCCA, similar to pCCA and dCCA, shows a high frequency of mutations in KRAS and/or TP53 genes.

- Growing evidence demonstrates that distinct cells of origin within an organ can give rise to different sub-types of cancer, typically tissue-specific stem and progenitor cells
- These findings could be useful to establish treatment and diagnostic strategies for BTCs based on genetic profile.

Therapeutic and prognostic importance of genetic signature.

- Liquid biopsy detection of cfDNA and miRNAs of mutated genes is a routine clinical diagnostic procedure and/or prognostic indicator of BTC.
- Immununotherapy with either chemotherapy or targeted therapy or loco regional therapies such as transarterial chemoembolization (TACE), cryotherapy, radiofrequency ablation (RFA), and radiotherapy is done based on the genetic mutations.
- Relevant genes and proteins involved in chemo-resistance therefore need to be ruled out

Problem statement

Different anatomical sites and genetic aberrations contribute to BTC heterogeneity.These BTC sub types have different prognosis and response to therapy.
This analysis seeks to identify the crucial genes in BTC and thereafter identify potential linkages with site specific chemotherapeutic compounds.

Objectives

Identify potential site-stecific therapeutic opportunities.

1. What are the most essential genes associated with BTC?
2. Which therapeutic compound can potentially be used to treat BTC based on the shared genes?
3. Which are the potential site specific therapeutic compounds based on the shared genes?

Methodology

1. Social network analysis
    - Network level measures:
        - Size
    - Node-level measures(Centrality)
        - Degree centrality.
        - Betweenness centrality.
        - Closeness centrality.
    - Link prediction: Common neighbors.
2. Tools: Neo4J and R.

Metagraph

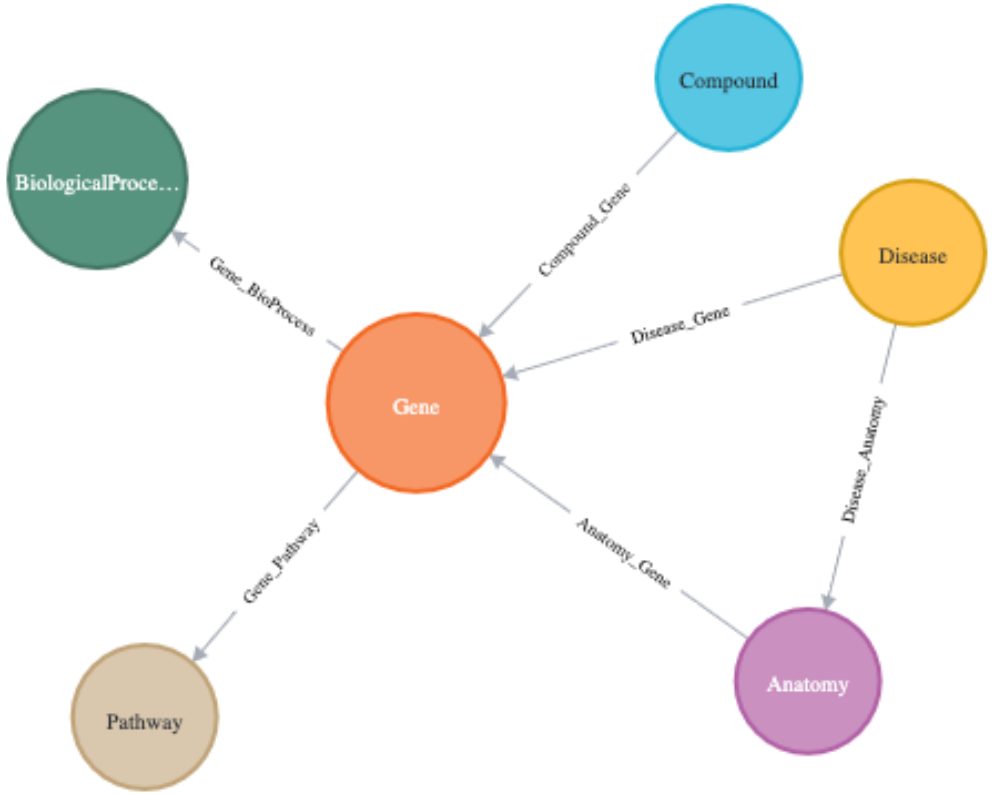


Fig 1.Metagraph showing the types of nodes used to build the network and the types of links defined to connect the nodes. Reproduced as provided by ALI DAWOOD

Dataset

Subset of Hetionet data centered on Biliary Tract Cancer.

Hetionet is an integrative network of biomedical knowledge combining information from 29 public databases. The network combines over 50 years of biomedical information into a single resource, consisting of 47,031 nodes (11 types) and 2,250,197 relationships (24 types)[3].

	Dataset	Triples	Node1	Relationship	Node2
1	compound_disease	0	Compound	TREATS_CtD	Disease
2	gene_bioprocess	11187	Gene	PARTICIPATES_GpBP	BiologicalProcess
3	disease_gene	42	Disease	ASSOCIATES_DaG	Gene
4	gene_pathway	8089	Gene	PARTICIPATES_GpPW	Pathway
5	disease_anatomy	16	Disease	LOCALIZES_DlA	Anatomy
6	compound_gene	8596	Compound	BINDS_CbG	Gene
7	gene_anatomy	13695	Anatomy	EXPRESSES_AeG	Gene

Size of the network

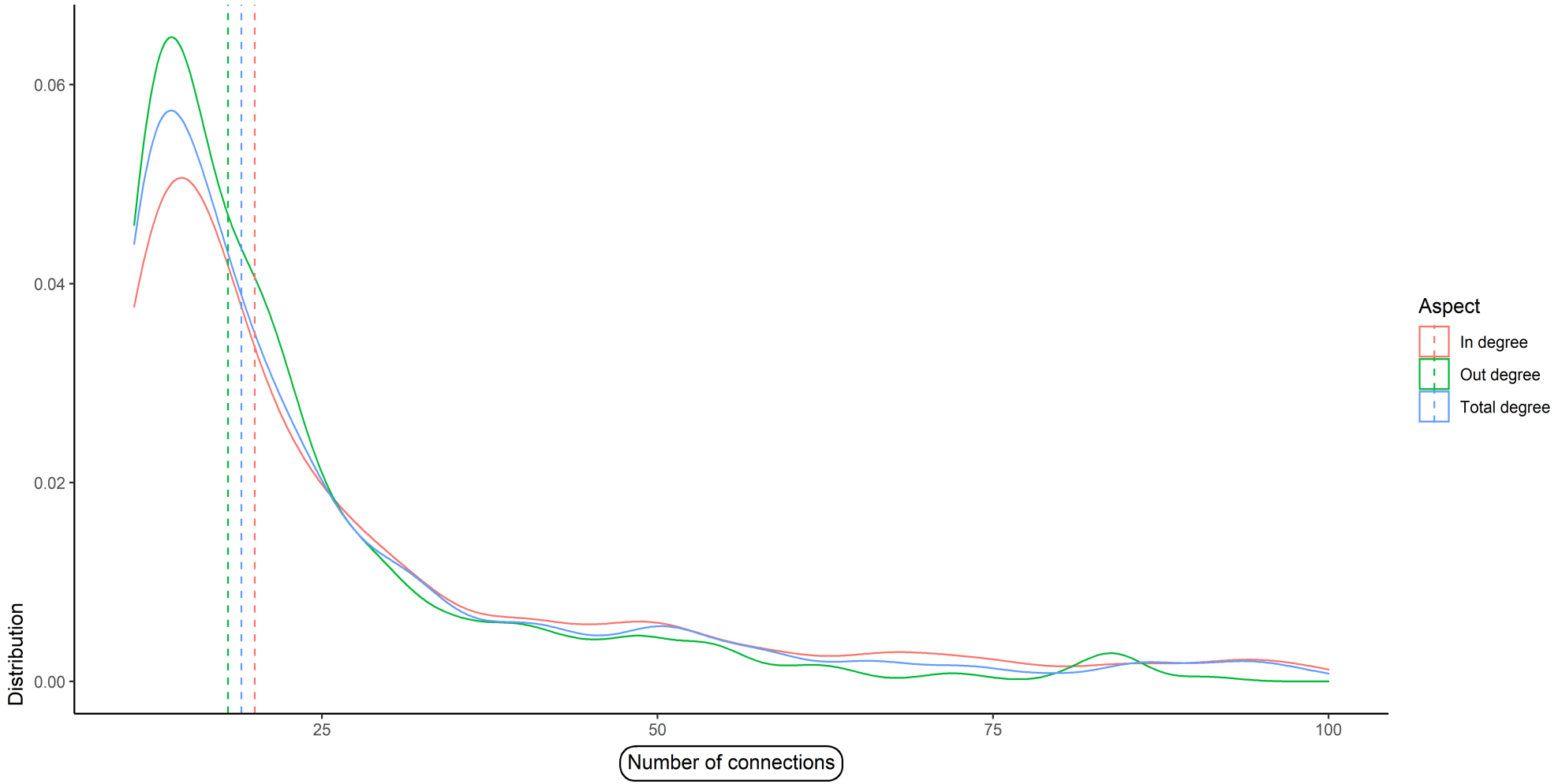
This was a directed graph with a total of 17535 nodes and 94950 edges.

Measure	Value
Gene nodes	13730
Compound nodes	1313
Disease nodes	1
Biological Process nodes	1026
Pathway nodes	1449

Measure	Value
Genes and Diseases	42
Compounds and Diseases	0
Disease and Anatomy	16
Anatomy and Gene	13695
Gene and Biological Process	11187
Gene and Pathway	13376

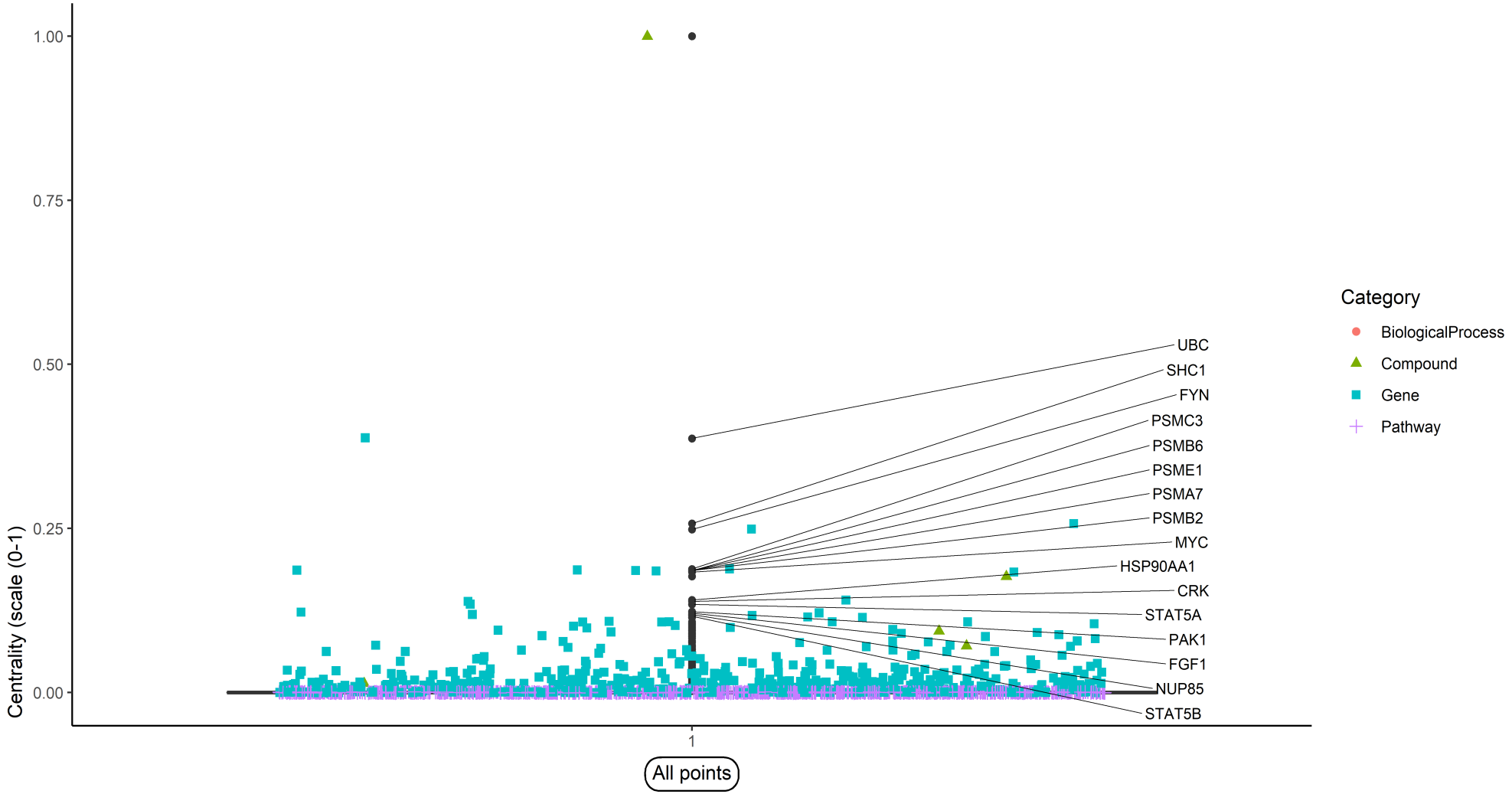
Density plot of degree centrality

The in-degree, out degree and total degree distribution for all the nodes. This plot excludes nodes with degrees below 10 and above 1000



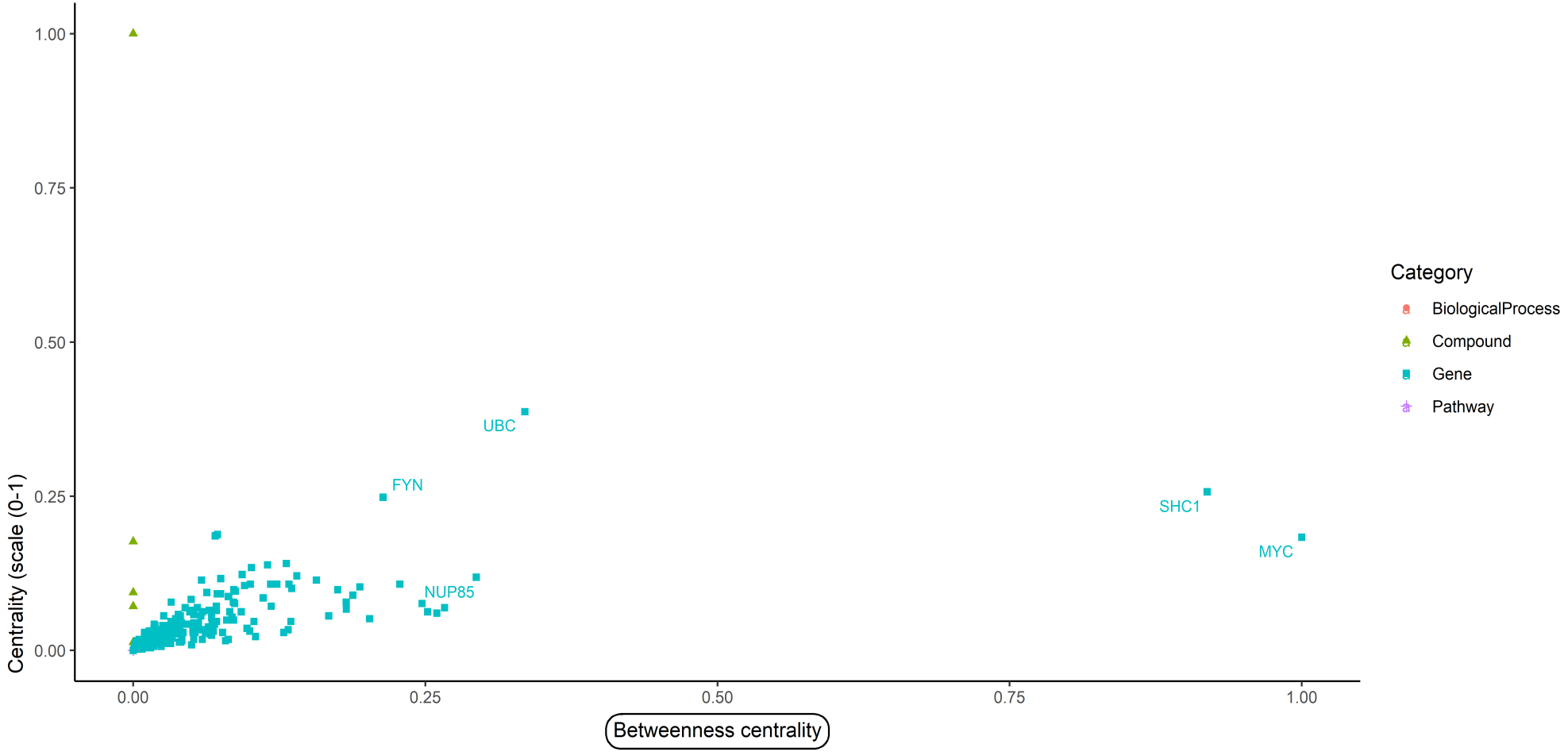
Centrality Measure

Only genes in the 10th-percentile are labeled.



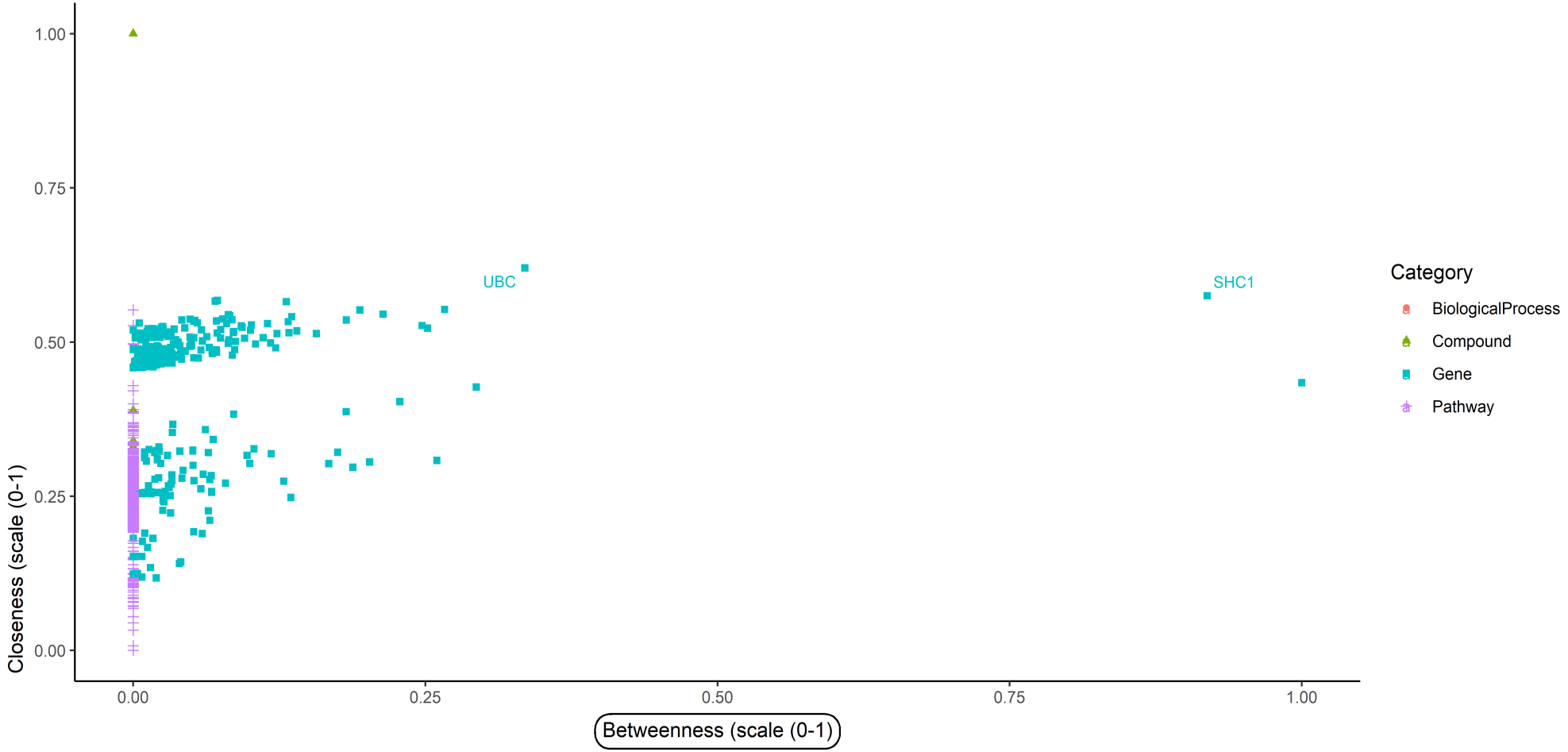
Influential connectors

Centrality and betweenness scatter plot to identify nodes that have a larger influence in the network. High centrality and betweenness shows many and influential linkages. ONly genes within the 10th-percentile in both are labeled



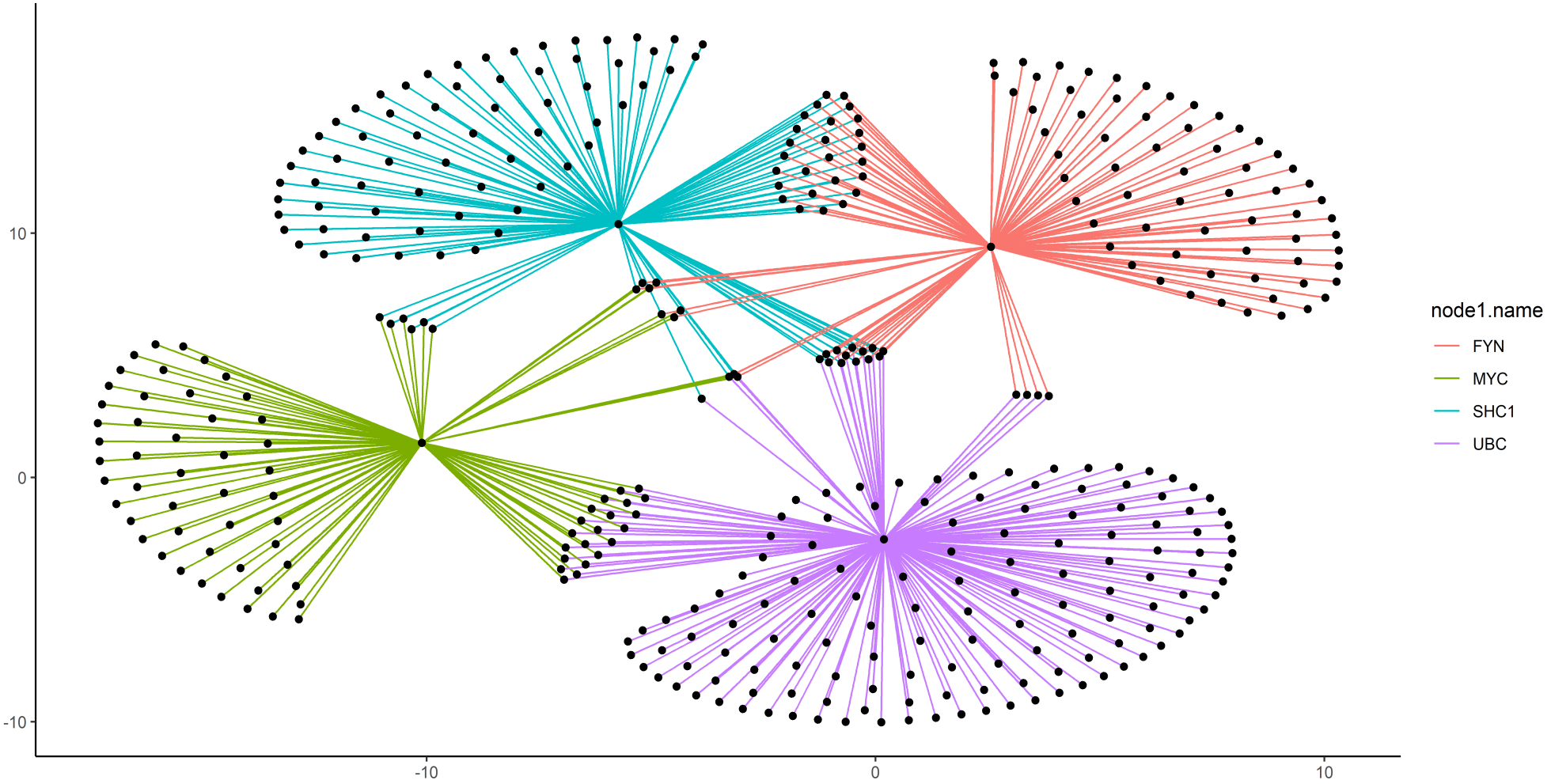
Information flow influencers

Closeness centrality and betweenness centrality scatter plot to identify nodes that have a larger influence in the network. High closeness betweenness support distribution of information effectively throughout the network.



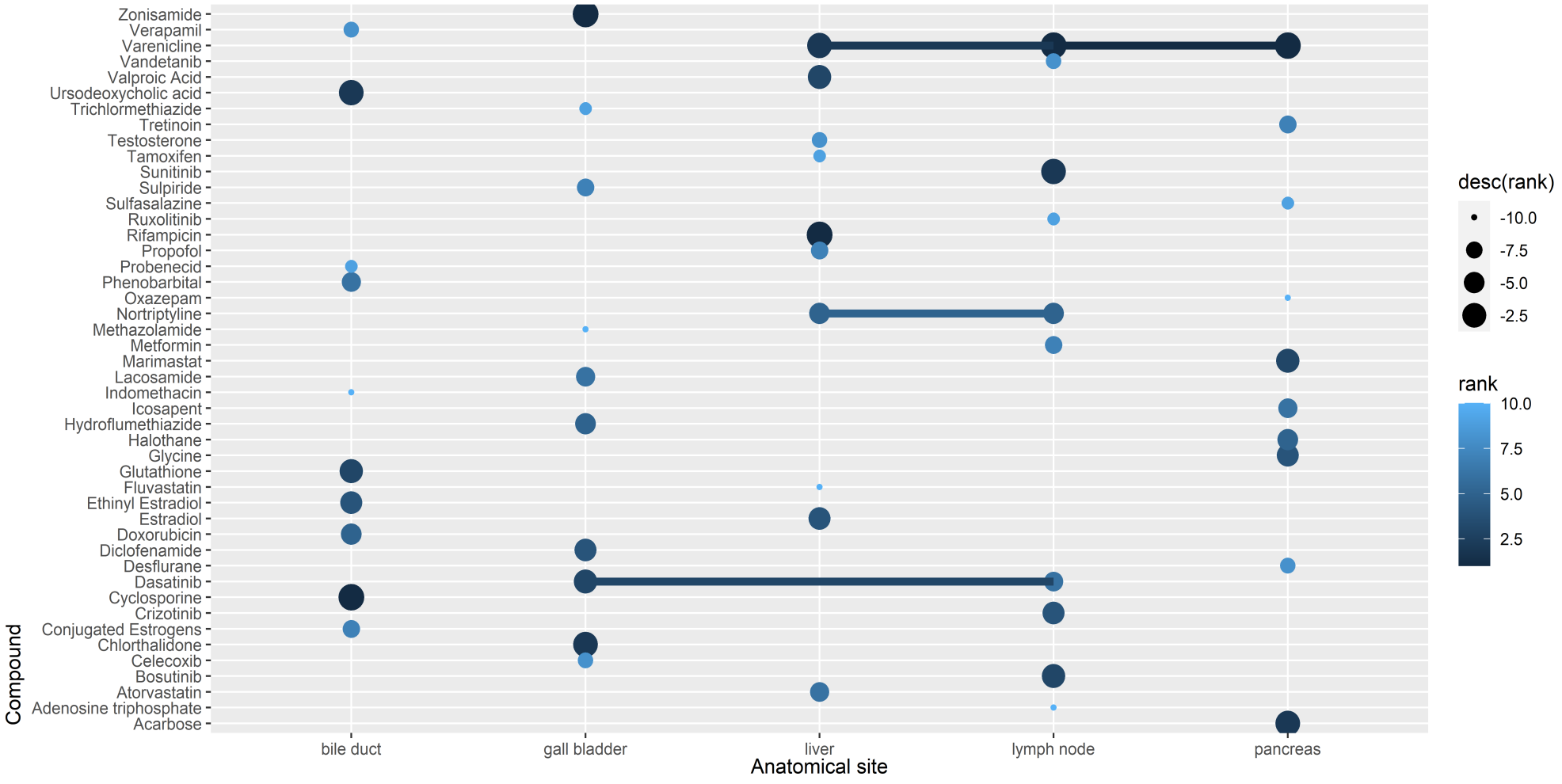
Subset of the network of the top influential genes

Visualization of the network formed by the four genes ranked to have the highest degree centrality, betweenness and closeness centrality.



Potential linkages between the compound and site based on shared genes.

Potential linkages with highest common neighbor scores (46 out of 90)

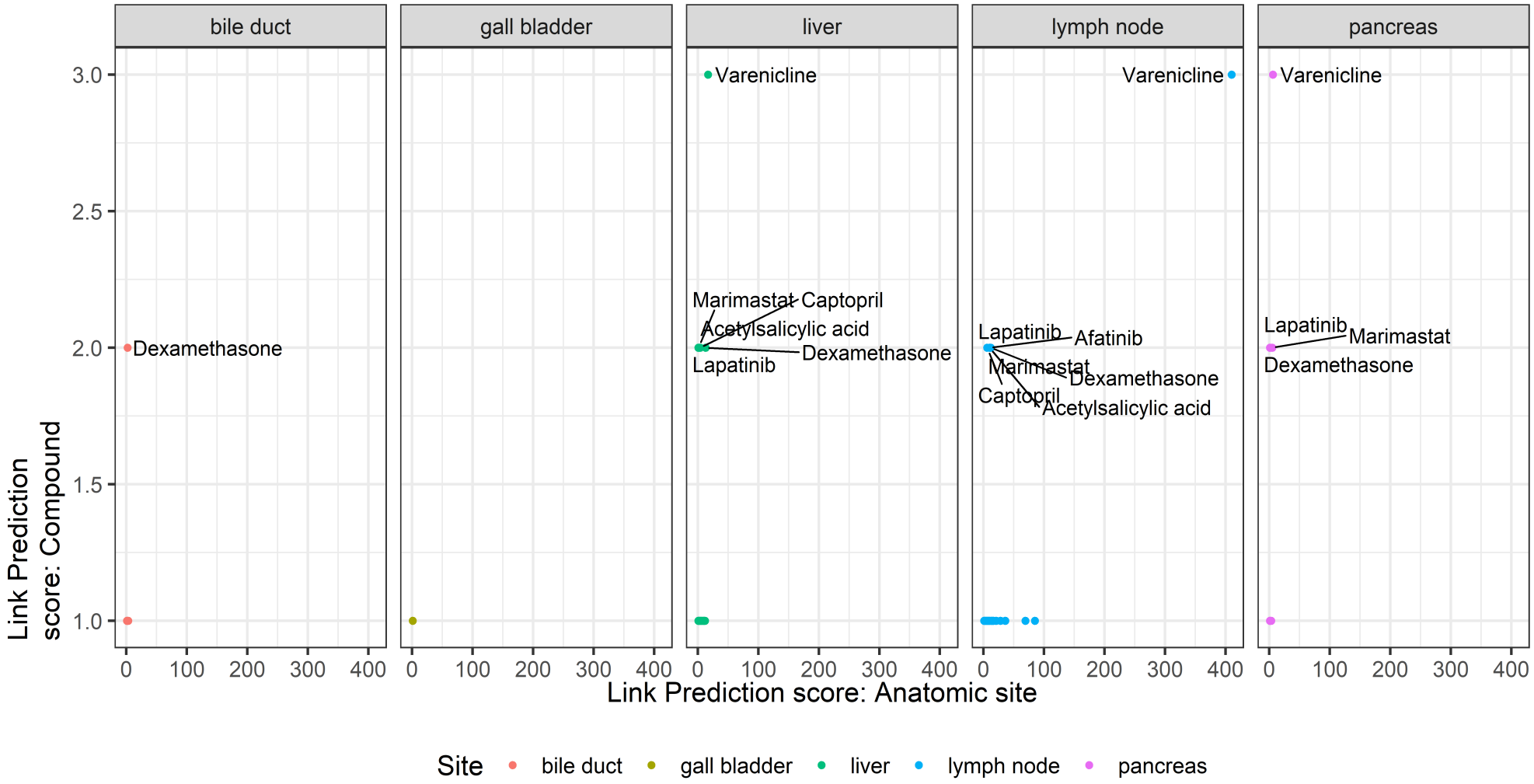


Link prediction between compounds and BTC based on shared genes.

The strength of the linkage is represented by the size of the word.

Predicting site specific compounds.

Correlation between link prediction scores of BTC-compound and Site-Compound



Genomic landscape of BTC

Influential genes

UBC

FYN

NUP85

SHC1

MYC

- Aberrations in three of these genes have been associated in the pathogenesis of BTC - MYC[1],UBC [1] FYN[4]
- NUP85 is a low tissue specific gene, but has been attributed to poor prognosis of gastrointestinal cancers[4].
- SHC1 is a also a low tissue specific gene, but has been attributed to poor prognosis of renal, lung and cervical cancers[4].

Drug target prediction by leveraging on gene network.

Varenicline: Used for smoking cessation.Has been associated with negative effects among cancer patients.

Dexamethasone: Associated with hepatic injury.

Afatinib: Some ongoing trials as an add on to chemo-therapy naive patients with advanced BTC.

Lapatinib: This drug has been shown to be effective in the BTC cases showing HER2/neu mutations.

Captopril: Indicated for hypertension but, there are trials for use in cancer.

Marimastat: Though not used for BTC, it is conventionally used for other GIT cancers.

Conclusion

- While the role of the three genes (MYC,UBC and FYN) is potentially associated with the pathology of BTC,the strong links to the two other genes (SHC1 and NUP85) maybe attributed to likelihood of metastasis from neighboring organs.

- Considering the compounds, the following can be deduced:
  Drugs indicated for use: Lapatinib.
  Potentially useful drugs: Afatinib, Captopril and Marimastat.
  Potentially harmful drugs: Varenicline and Dexamethasone.

References:

1. Banales, J.M., Marin, J.J.G., Lamarca, A. et al. Cholangiocarcinoma 2020: the next horizon in mechanisms and management. Nat Rev Gastroenterol Hepatol 17, 557–588 (2020). https://doi.org/10.1038/s41575-020-0310-z

2. Farshidfar, Farshad et al. “Integrative Genomic Analysis of Cholangiocarcinoma Identifies Distinct IDH-Mutant Molecular Profiles.” Cell reports vol. 18,11 (2017): 2780-2794. doi:10.1016/j.celrep.2017.02.033

3. Heterogeneous Network Edge Prediction: A Data Integration Approach to Prioritize Disease-Associated Genes Himmelstein DS, Baranzini SE PLOS Computational Biology (2015) DOI: 10.1371/journal.pcbi.1004259 · PMID: 26158728 · PMCID: PMC4497619

4. Zou S, Li J, Zhou H, Frech C, Jiang X, Chu JSC, et al. Mutational Landscape of Intrahepatic Cholangiocarcinoma. Nat Commun (2014) 5(1):5696. doi: 10.1038/ncomms6696.

5. Xia, R., Tang, H., Shen, J. et al. Prognostic value of a novel glycolysis-related gene expression signature for gastrointestinal cancer in the Asian population. Cancer Cell Int 21, 154 (2021). https://doi.org/10.1186/s12935-021-01857-4

Thanks!

Project repository:

em-kagereki/Graph-analysis