

# Group 14



## PUTATIVE DISEASE GENE IDENTIFICATION AND DRUG REPURPOSING FOR OSTEOPOROSIS

Members:  
Emre Pelzer  
Syed Habibul Bashar

# INTRODUCTION

- Osteoporosis Overview:
  - Progressive bone disease with high fracture risk.
  - Current treatments manage symptoms but lack targeted solutions
- Objective:
  - Use network biology to identify disease genes and explore drug repurposing opportunities.

# METHODOLOGY

- Approach:
  - Reconstruct human interactome using BioGRID data.
  - Identify disease genes and prioritize using algorithms.
  - Perform enrichment analysis and link genes to drugs.

# DATA COLLECTION

- Protein-Protein Interaction Data:
  - Source: BioGRID.
  - Filter: Human-specific, physical interactions.
- Gene-Disease Association Data:
  - Source: Curated datasets.
  - Verified with HGNC database.

# INTERACTOME RECONSTRUCTION

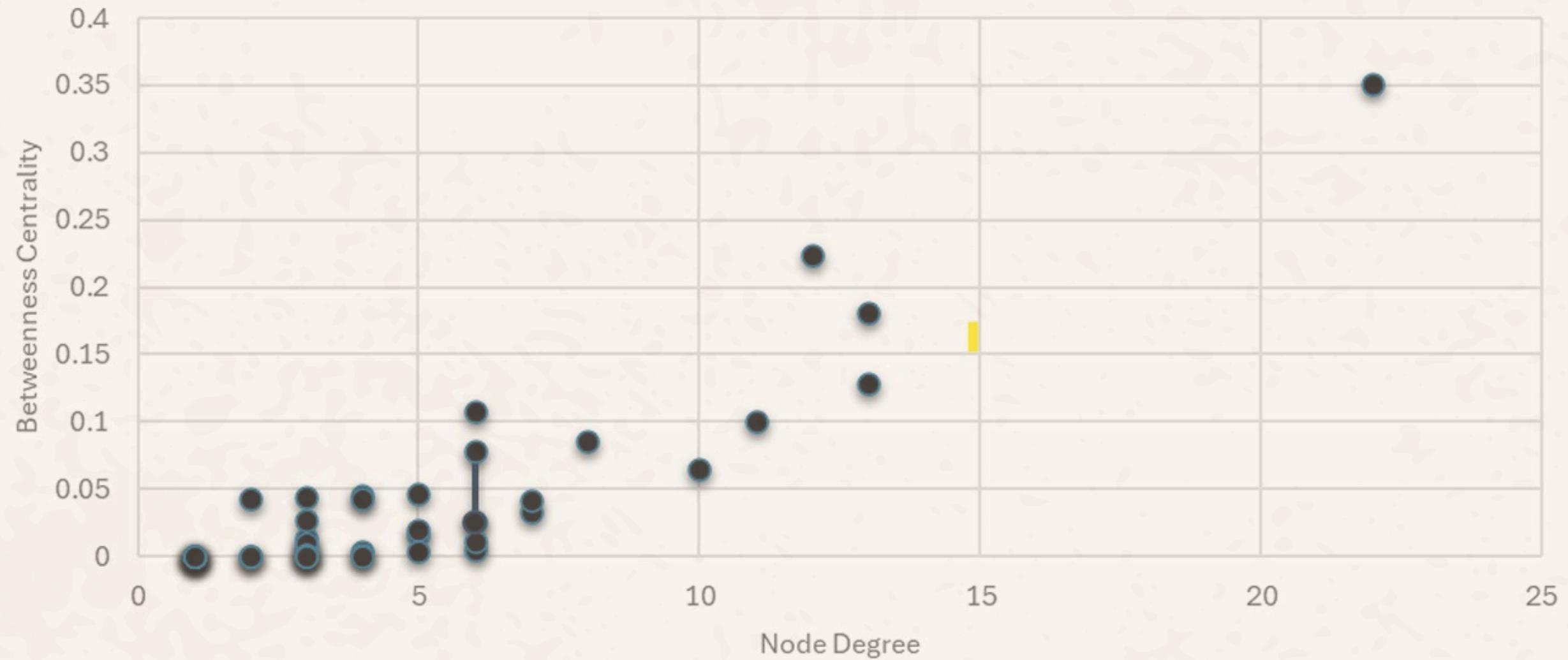
## Largest Connected Component (LCC):

- Disease interactome: 85 genes identified, 84 mapped.
- Organism interactor A & B = 9606
- Experimental System type : Physical.
- Removal of self loops.
- LCC with 47 nodes & 113 edges
- Metrics analyzed: Degree, Betweenness, Eigenvector, Closeness.

disease name	UMLS disease ID	MeSH disease class	number of associated genes	number of genes present in the interactome	LCC size of the disease interactome
Osteoporosis	C0029456	T047	85	84	47

# NODE DEGREE VS BETWEENNESS CENTRALITY

Node Degree Vs Betweenness Centrality



# ALGORITHM COMPARISON

Find genes related to a disease by looking at how they connect in a network of proteins

## DIAMOnD

Best for precision and relevance; focuses on genes closely linked in the network.

identifies disease clusters

## DiaBLE

Smaller universe size and less precision.

sacrifices specificity for inclusiveness

## Diffusion Based

Limited effectiveness, especially at higher diffusion times, as precision and recall drop to zero.

spread a signal from the known disease genes through the network

# ALGORITHM COMPARISON

- Precision: Proportion of identified genes that are true positives.
- Recall: Proportion of true positives identified from all possible positives.
- F1 Score: Balances precision and recall.

Algorithm	Precision (mean ± SD)	Recall (mean ± SD)	F1-Score (mean ± SD)
DIAMOnD	0.0200 ± 0.0126	0.0625 ± 0.0395	0.0303 ± 0.0192
DiABLE	0.0160 ± 0.0196	0.0500 ± 0.0612	0.0242 ± 0.0297
Diffusion_t=0.002	0.0080 ± 0.0098	0.0250 ± 0.0306	0.0121 ± 0.0148
Diffusion_t=0.005	0.0080 ± 0.0098	0.0250 ± 0.0306	0.0121 ± 0.0148
Diffusion_t=0.01	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000

DIAMOnD effectively identifies relevant genes while minimizing false positives

# PROCONSUL VS DIAMOND

Objective: Compare gene prioritization methods for disease modules.

## ProConSuL:

- Focused on small datasets where we want precise clusters
- Limited scalability to large PPI networks.
- Finds precise but smaller clusters.

## DIAMOnD:

- Designed for large, complex networks.
- Expands modules by adding genes based on network connections.
- High precision and relevance for large datasets.

ProConSuL_node	DIAMOnD_node
AGR2	AGR2
BAP1	TIMP2
VCP	CDK2
CDK2	CLIC4
CLIC4	PRDX6
PRDX6	CLIC1
CLIC1	BAP1
HSPA8	VCP
EZH2	DSTN
UBE2M	MYOC
DSTN	FN1
PRDX2	TKT
SOD1	CFL1
CFL1	UBE2M
IQGAP1	SOD1
ISG15	PRDX2
U2AF2	NME2
ACO2	LDHA
MCM2	EZH2
TKT	ACO2

# PUTATIVE GENE IDENTIFICATION

- Objective: Identify genes associated with osteoporosis using DIAMOnD

- Methods:

- 84 known disease genes (seed genes) in PPI network, constructed by BioGRID.
- DIAMOnD prioritized 100 putative genes based statistically, on network proximity to known disease genes.
- Pathway enrichment (GO and KEGG validated biological relevance).

- Key Result:

- Top pathways: Bone remodeling, cellular organization.
- Example pathway: KEGG Bone Remodeling.

biological process  
molecular function  
cellular component  
System-wide biological pathways

Curated pathways

Table 4: Obtained data from EnrichR (adj. p-value < 0.05)

Category	Total Original Terms	Total Putative Terms	Number of Overlapping Terms	Percentage Overlap (Original)	Percentage Overlap (Putative)	Unique Terms in Original	Unique Terms in Putative
GO-BP	387	133	29	7.49	21.80	358	104
GO-MF	27	36	5	18.52	13.89	22	31
GO-CC	22	16	9	40.91	56.25	13	7
Reactome	240	174	72	30.00	41.38	168	102
KEGG	74	28	12	16.22	42.86	62	16

# DRUG REPURPOSING

Objective:

Link prioritized genes to approved drugs for osteoporosis treatment.

Methods:

Top 20 genes mapped using DGIdb.

Validation:

Cross-checked drugs in clinical trials.

Key Results:

8 genes linked to approved drugs.

gene_name	associated_drugs	drug_count
CDK2	['LOVASTATIN', 'PACLITAXEL', 'RESVERATROL', 'DAUNORUBICIN LIPOSOMAL', 'CARBOPLATIN', 'ERIBULIN MESYLATE', 'ACETAMINOPHEN', 'RALTITREXED', 'DEXAMETHASONE']	9
BAP1	['PANOBINOSTAT', 'VORINOSTAT', 'SUNITINIB', 'EVEROLIMUS', 'VALPROIC ACID', 'OLAPARIB']	6
CFL1	['CLOTRIMAZOLE', 'CINNARIZINE', 'CLOFIBRATE', 'FENOFIBRATE MICRONIZED', 'SERTRALINE HYDROCHLORIDE']	5
NME2	['LAMIVUDINE', 'TENOFOVIR', 'ZIDOVUDINE', 'PROGESTERONE', 'ADEFOVIR DIPIVOXIL']	5
SOD1	['TETRACYCLINE', 'DOXYCYCLINE ANHYDROUS', 'OXYTETRACYCLINE ANHYDROUS']	4
EZH2	['TAZEMETOSTAT', 'DABRAFENIB', 'TAZEMETOSTAT HYDROBROMIDE']	3
FN1	['OCRIPLASMIN', 'DACARBAZINE']	2
VCP	['HEXACHLOROPHENE']	1

Tenofovir and Lamivudine, both currently under clinical trials for osteoporosis.

This demonstrates the potential for repurposing existing treatments

**THANK YOU!**