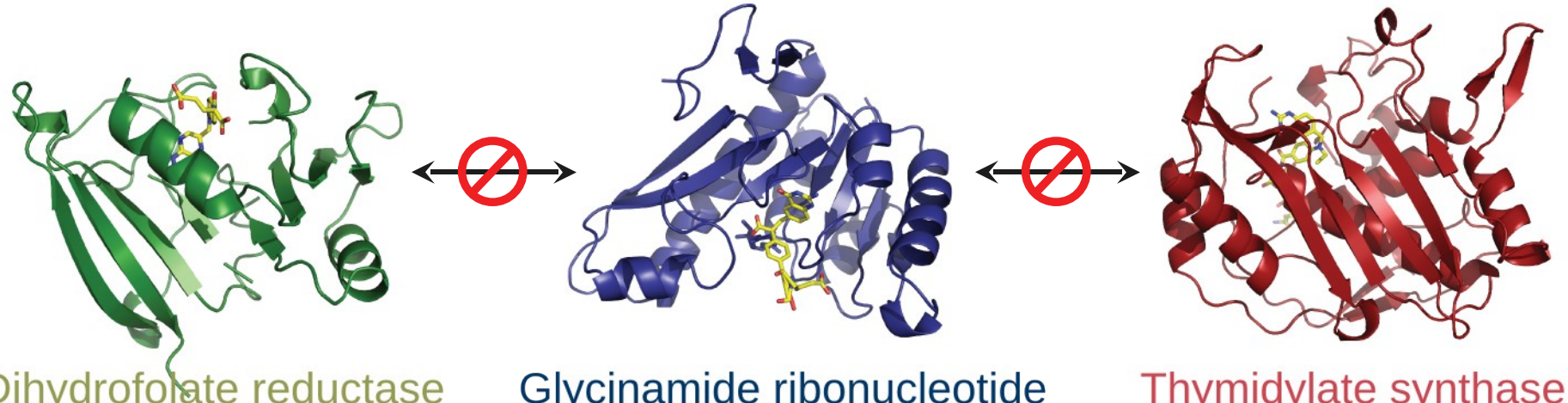
Can we quantify the relationship between drug target using their ligands?

Jérôme Hert

Antifolates receptors have no significant sequence or structural similarity

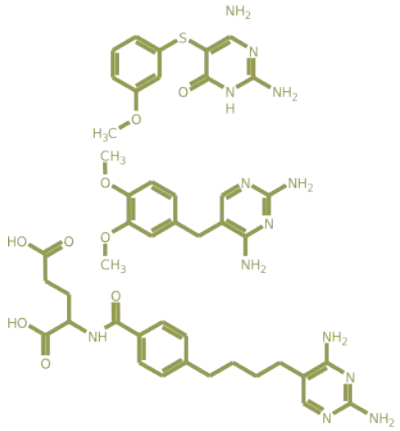


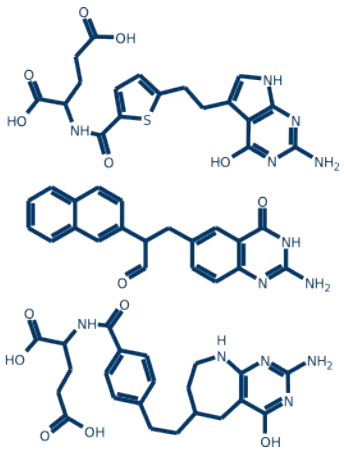
Dihydrofolate reductase (DHFR)

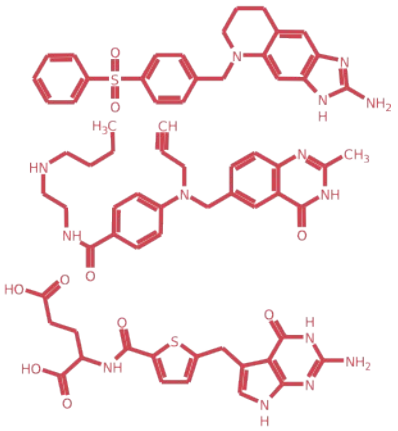
Glycinamide ribonucleotide formyltransferase (GART)

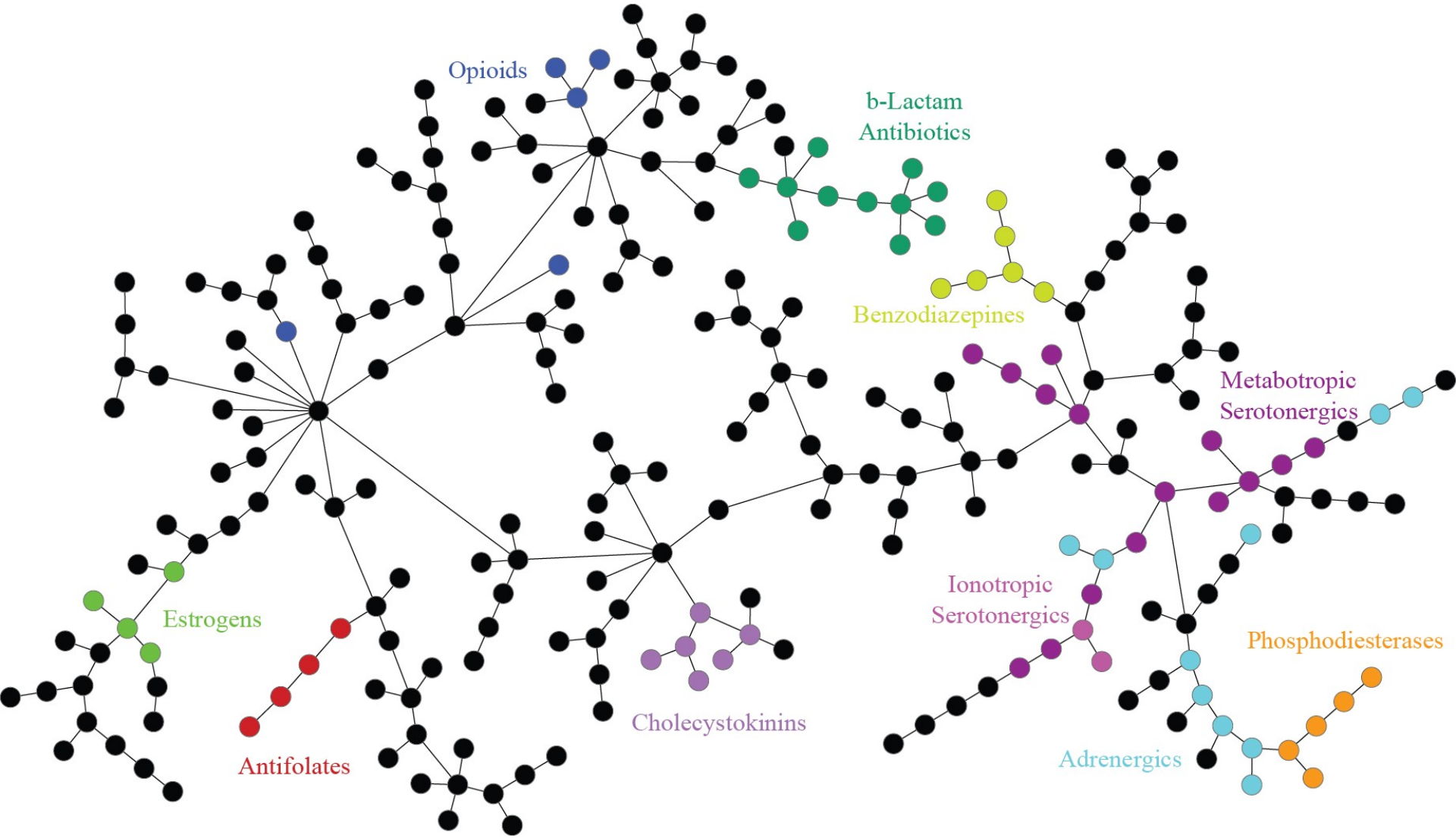
Thymidylate synthase (TS)

Many ligands are folic acid derivatives

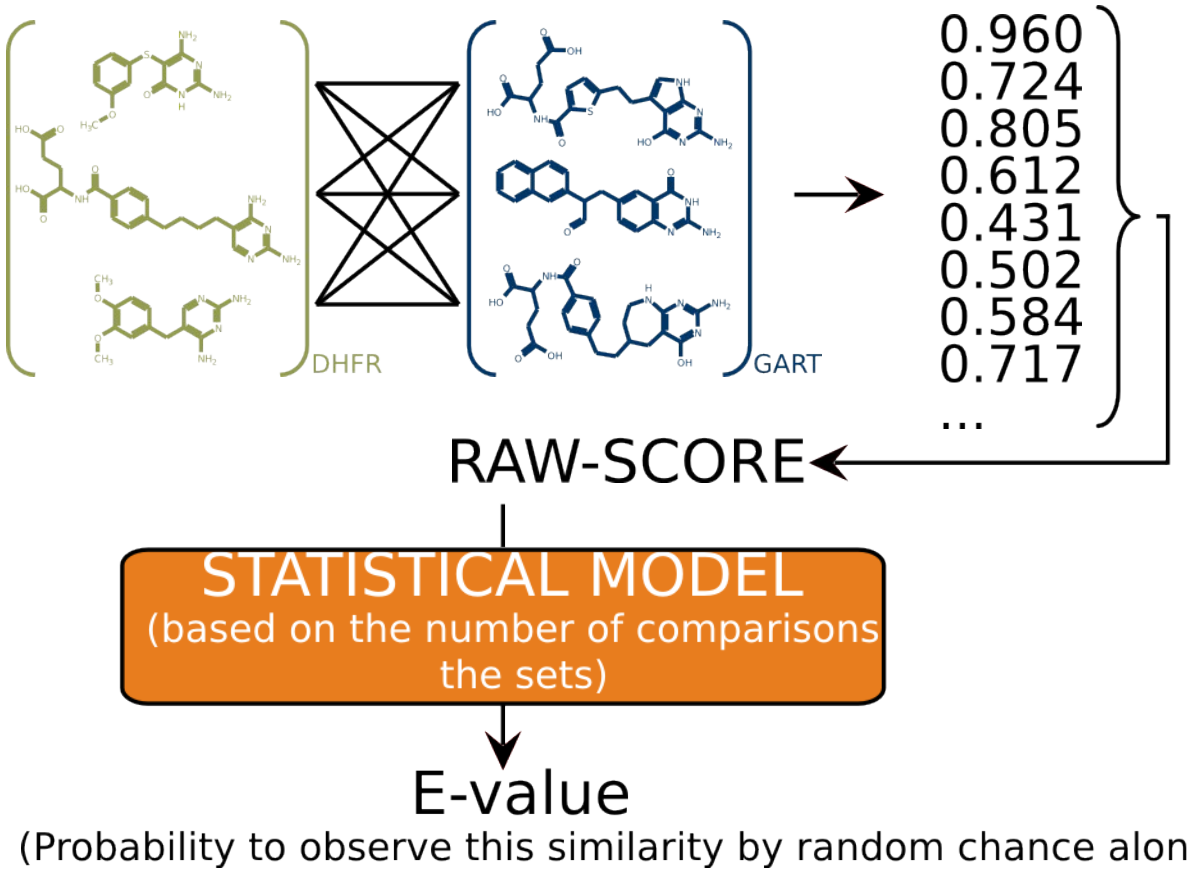






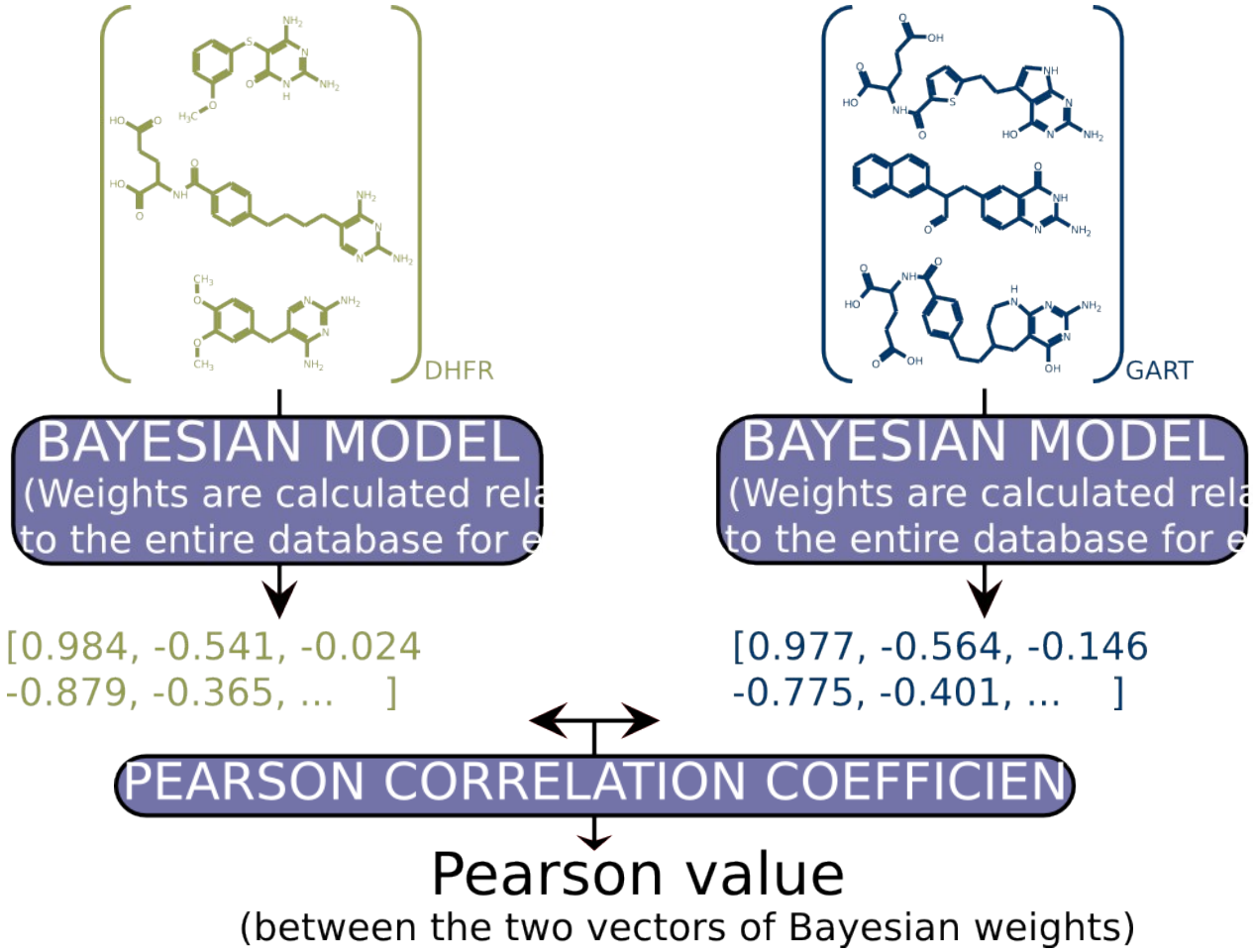

Opioids
b-Lactam
Antibiotics
Benzodiazepines
Metabotropic
Serotonergics
Ionotropic
Serotonergics
Phosphodiesterases
Adrenergics
Cholecystokinins
Estrogens
Antifolates

Similarity Ensemble Approach (SEA)



Keiser, M. J. *et al*. Relating Protein Pharmacology by their ligands. Nature Biotechnology 2007, 25, 197-206.

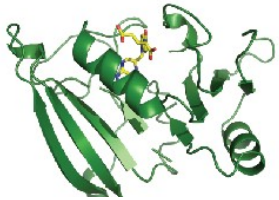
Comparison of Bayesian models

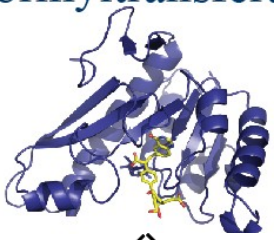


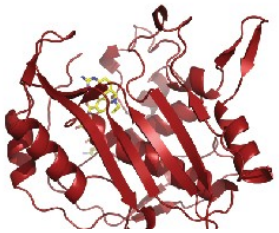
Dihydrofolate reductase

Glycinamide ribonucleotide formyltransferase

Thymidylate synthase







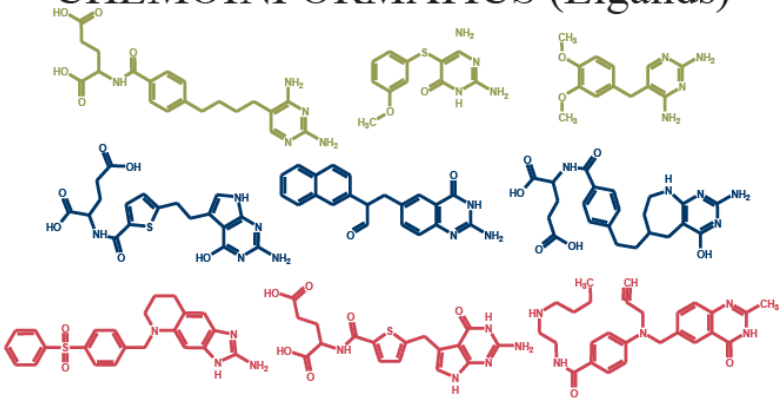
BIOINFORMATICS (Sequences)

```
VGSLNCIVAVSQNMGIGKNGDLPWPPLRNEFRYFQRMTTTSSVEGKQNLVIMGKKTWFSI
PEKNRPLKGRINLVLSRELKEPPQGAHFLSRSLDDALKLTEQPELANKVDMVWIVGGSSV
YKEAMNHPGHLKLFVTRIMQDFESDTFFPEIDLEKYKLLPEYPGVLSDVQEEKGIKY...

MAARVLIIGSGGREHTLAWKLAQSHHVKQVLVAPGNAGTACSEKISNTAISISDHTALAQ
FCKEKKIEFVVVGPEAPLAAGIVGNLRSAGVQCFGPTAEAAQLESSKRFAKEFMDRHGIP
TAQWKAFTKPEEACSFILSADFPALVVKASGLAAGKGVIVAKSKEEACKAVQEIMQE...

PVAGSELPRRPLPPAAQERDAEPRPPHGELQYLGQIQHILRCGVRKDDRTGTGTLSVFGM
QARYSLRDEFPLLTTKRVFWKGVLEELLWFIKGSTNAKELSSKGVKIWDANGSRDFLDSL
GFSTREEGDLGPVYGFQWRHFGAEYRDMESDYSGQGVDQLQRVIDTIKTNPDDRRIIMCA
WNPRDLPLMALPPCHALCQFYVVNSELSCQLYQRSGDMGLGVPFNIASYALLTYMIA...
```

CHEMOINFORMATICS (Ligands)

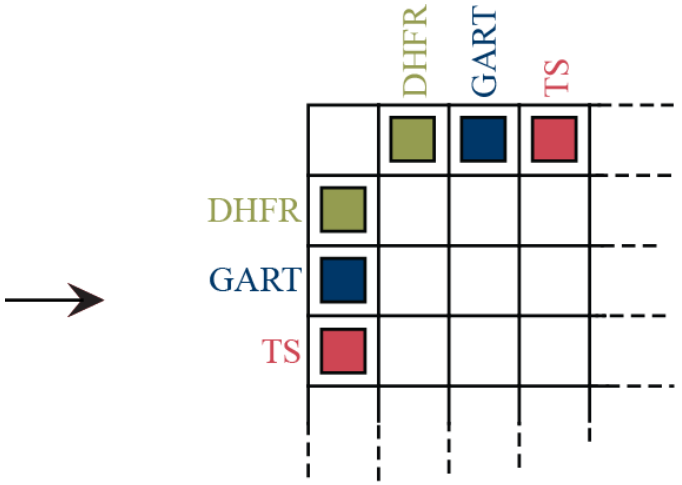


PSI-Blast

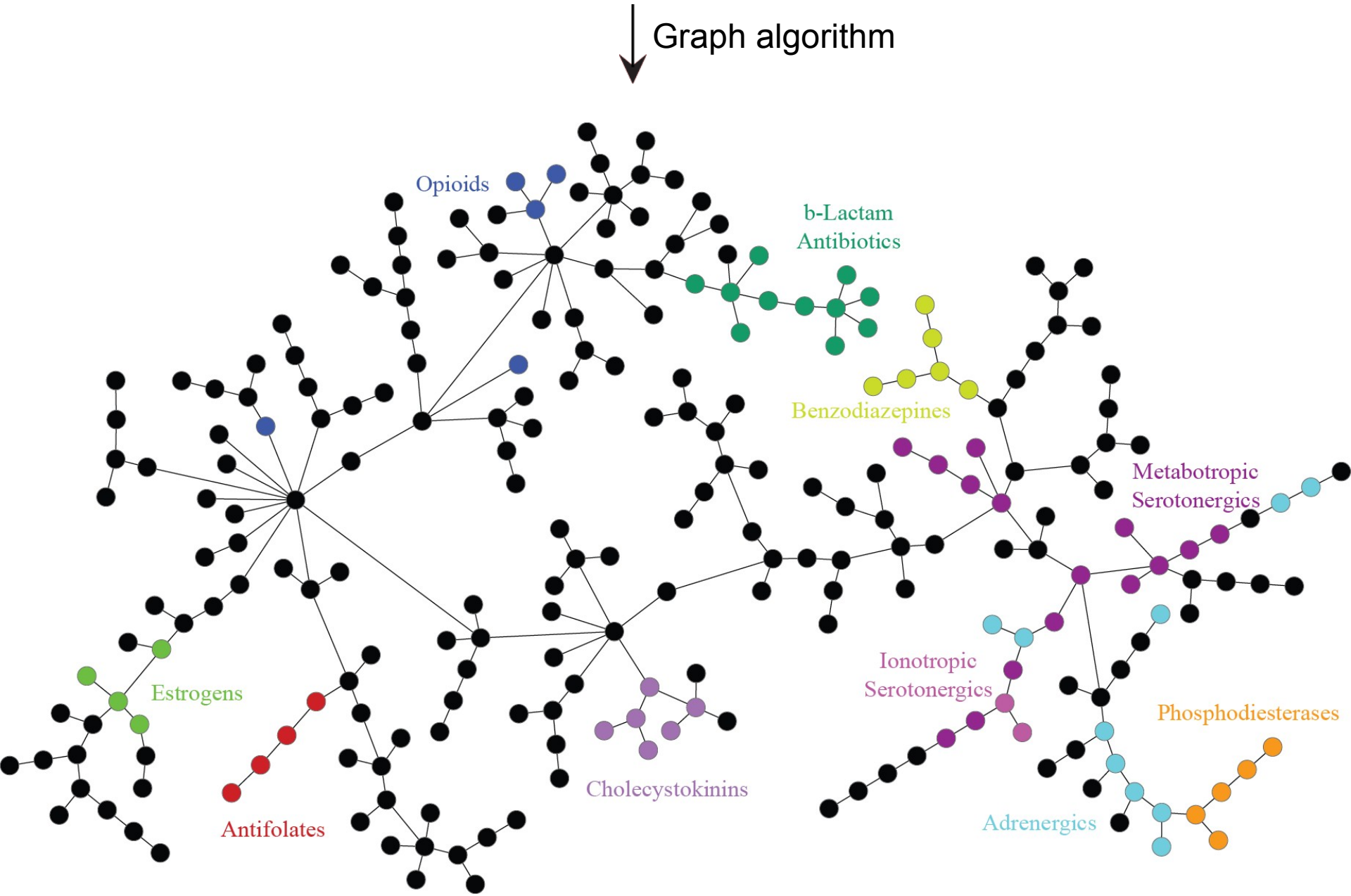
SEA or BAY

Rank	Ligand-set annotation	SEA E-value	PSI-Blast E-value
1	GART	1.3E-264	118.0
2	Thymidylate Synthetase	8.5E-142	53.0
3	Folylpolyglutamate Synthetase	7.5E-132	171.0
4	Uridine Phosphorylase	29471.9	1596.0
5	Dehydropeptidase-I	30607.4	233.0
6	Antiestrogen	30874.3	16.0
7	Calpain	30875.3	155.0
8	Cyclooxygenase-1	30875.6	95.0
9	Proteoglycanase	30876.0	10474.0

Most similar ligand-sets of DHFR according to SEA



Matrix of all pair-wise protein similarities



MDDR database – SEA - ECFP_4 fingerprints

How similar are chemo- and bio-informatics networks?

Are the chemoinformatics networks robust?

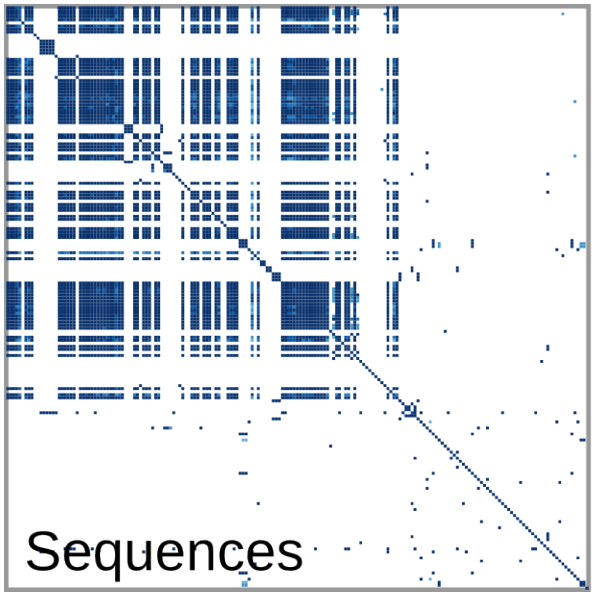
Are the chemoinformatics networks relevant?

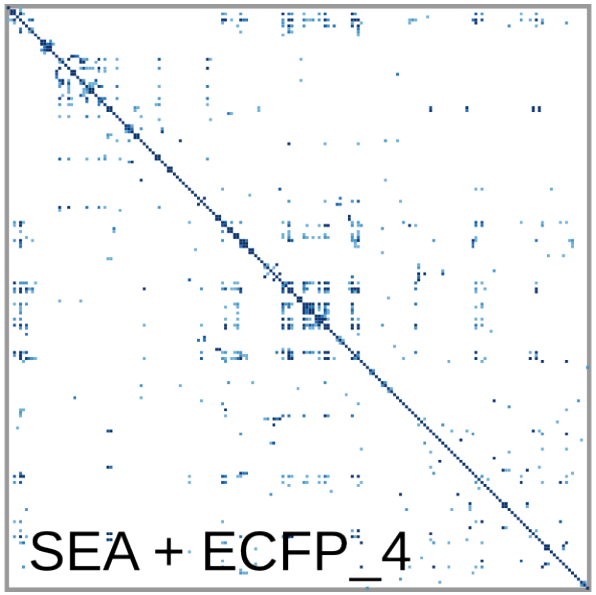
Chemogenomics databases:

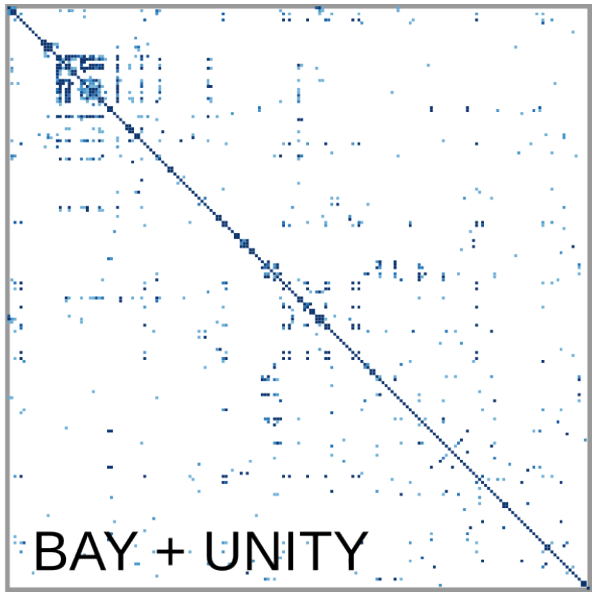
- MDL Drug Data Report (MDDR)
- World of Molecular Bioactivity (WOMBAT)

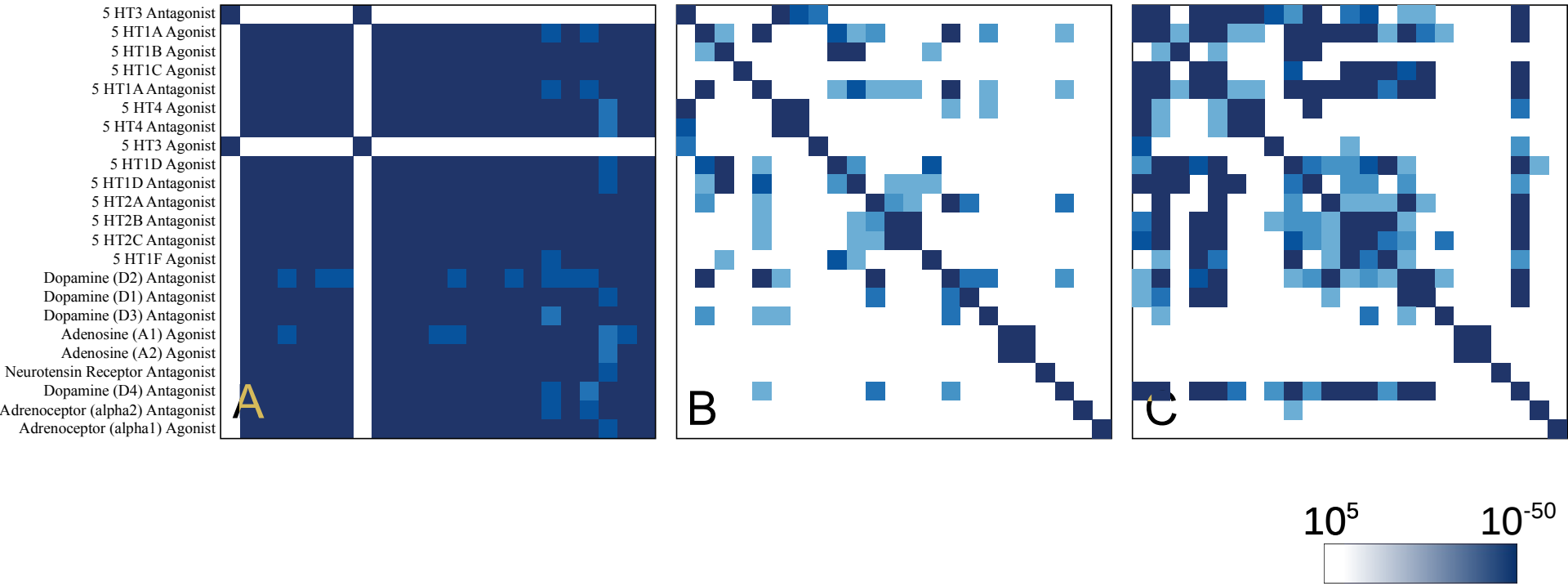
Descriptors

- Daylight, Unity
- MDL Keys
- ECFP_4, FCFP_4
- CATS
- FEPOPS

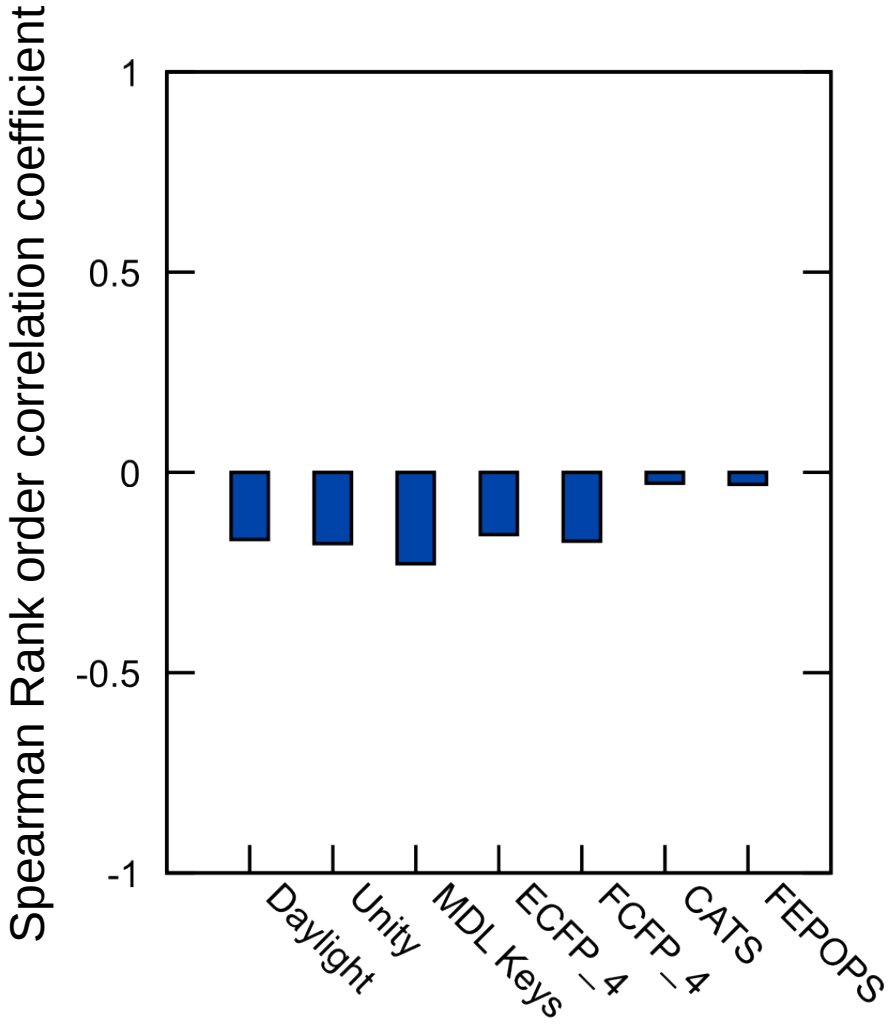

Sequences


SEA + ECFP_4

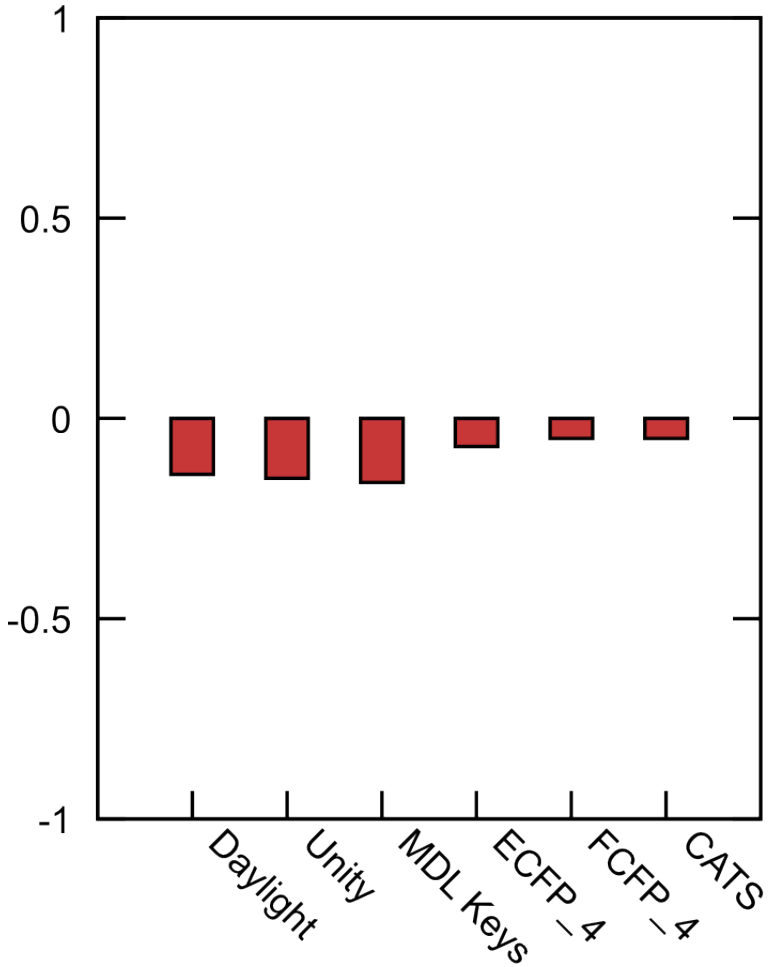

BAY + UNITY


5 HT3 Antagonist
5 HT1A Agonist
5 HT1B Agonist
5 HT1C Agonist
5 HT1A Antagonist
5 HT4 Agonist
5 HT4 Antagonist
5 HT3 Agonist
5 HT1D Agonist
5 HT1D Antagonist
5 HT2A Antagonist
5 HT2B Antagonist
5 HT2C Antagonist
5 HT1F Agonist
Dopamine (D2) Antagonist
Dopamine (D1) Antagonist
Dopamine (D3) Antagonist
Adenosine (A1) Agonist
Adenosine (A2) Agonist
Neurotensin Receptor Antagonist
Dopamine (D4) Antagonist
Adrenoceptor (alpha2) Antagonist
Adrenoceptor (alpha1) Agonist
A
B
C
10^{5}
10^{-50}

Bio- and Chemoinformatics networks are different

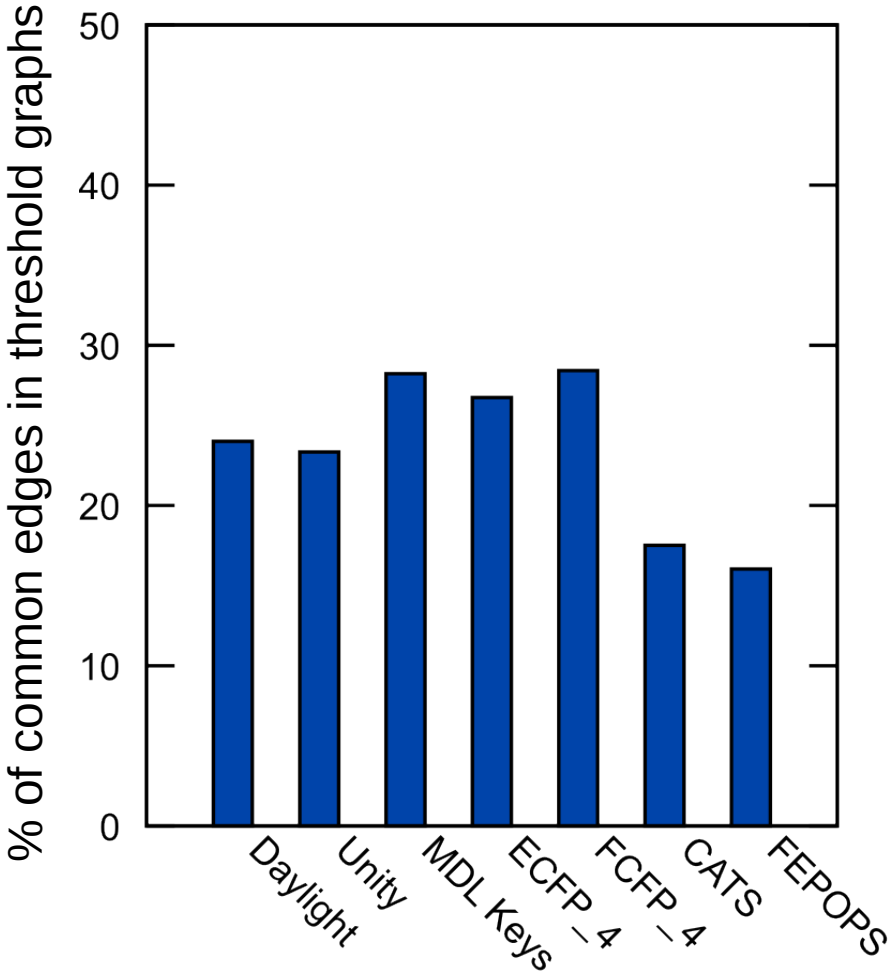


Sequence vs SEA

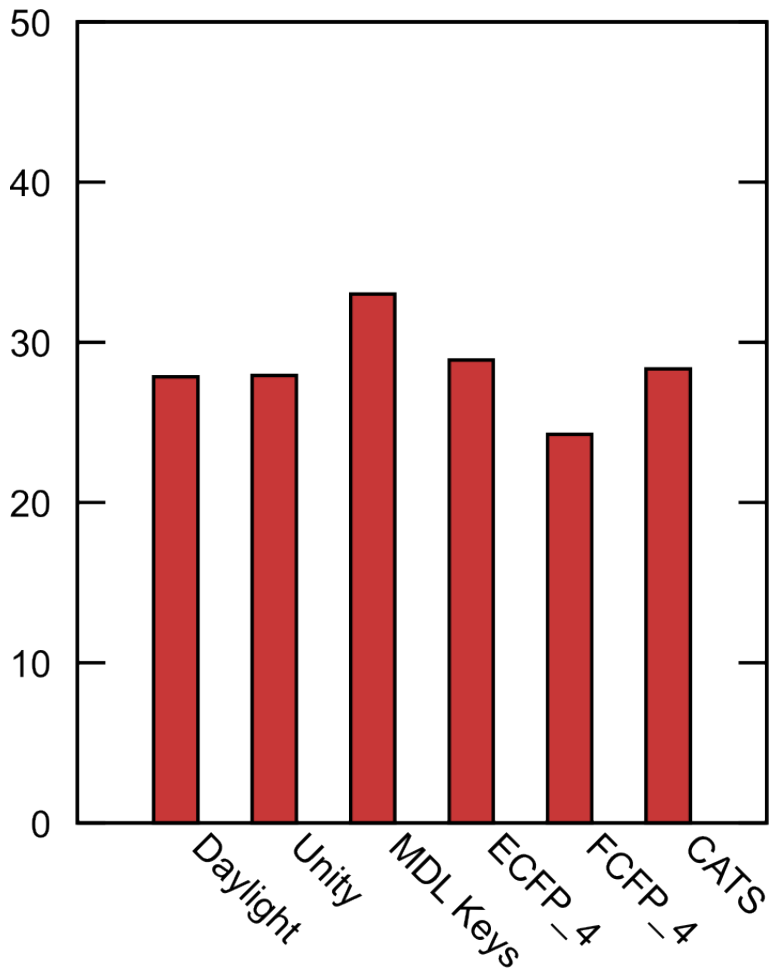


Sequence vs Bayes

Bio- and chemoinformatics networks are different

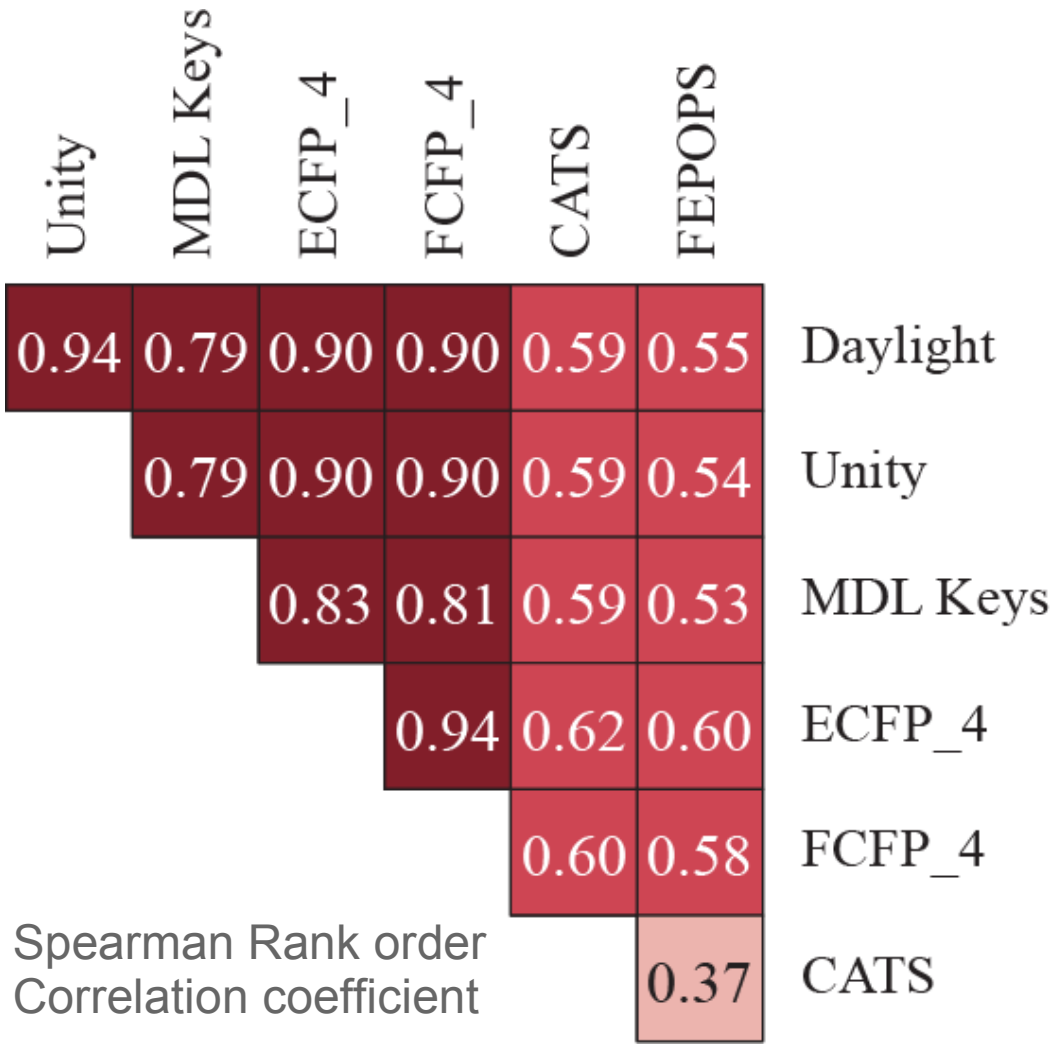


Sequence vs SEA

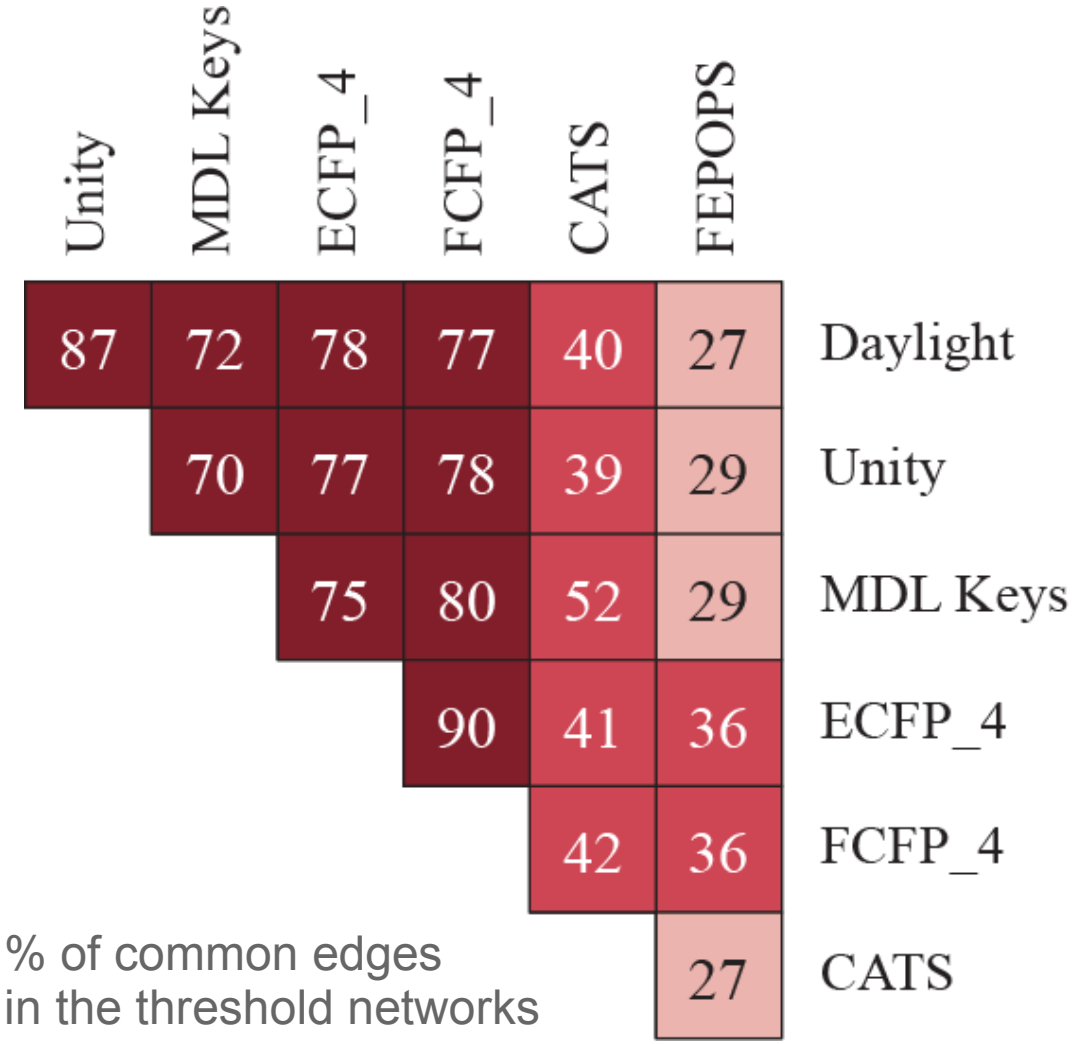


Sequence vs Bayes

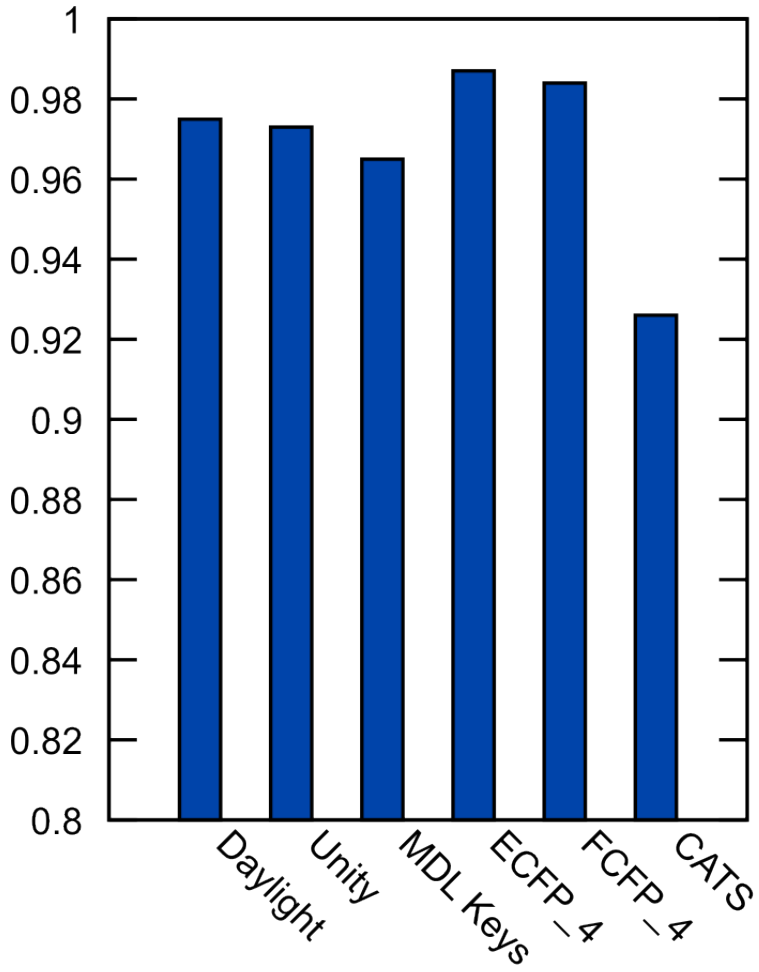
Chemoinformatics networks are robust



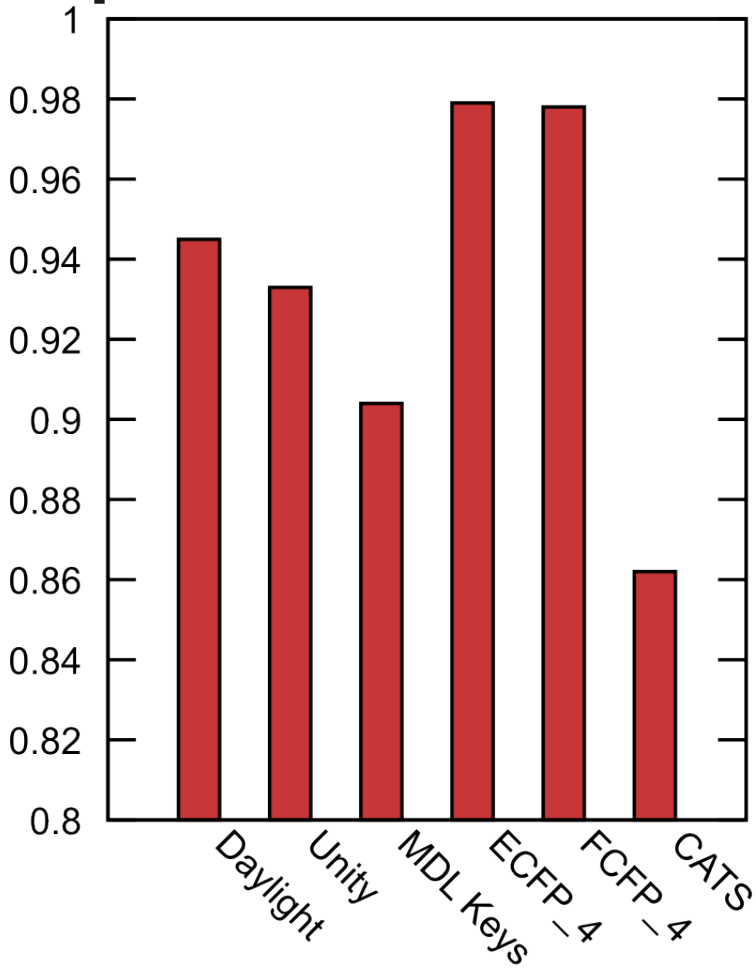
Chemoinformatics networks are robust



Average AUC-ROC in a 10-fold validation experiment

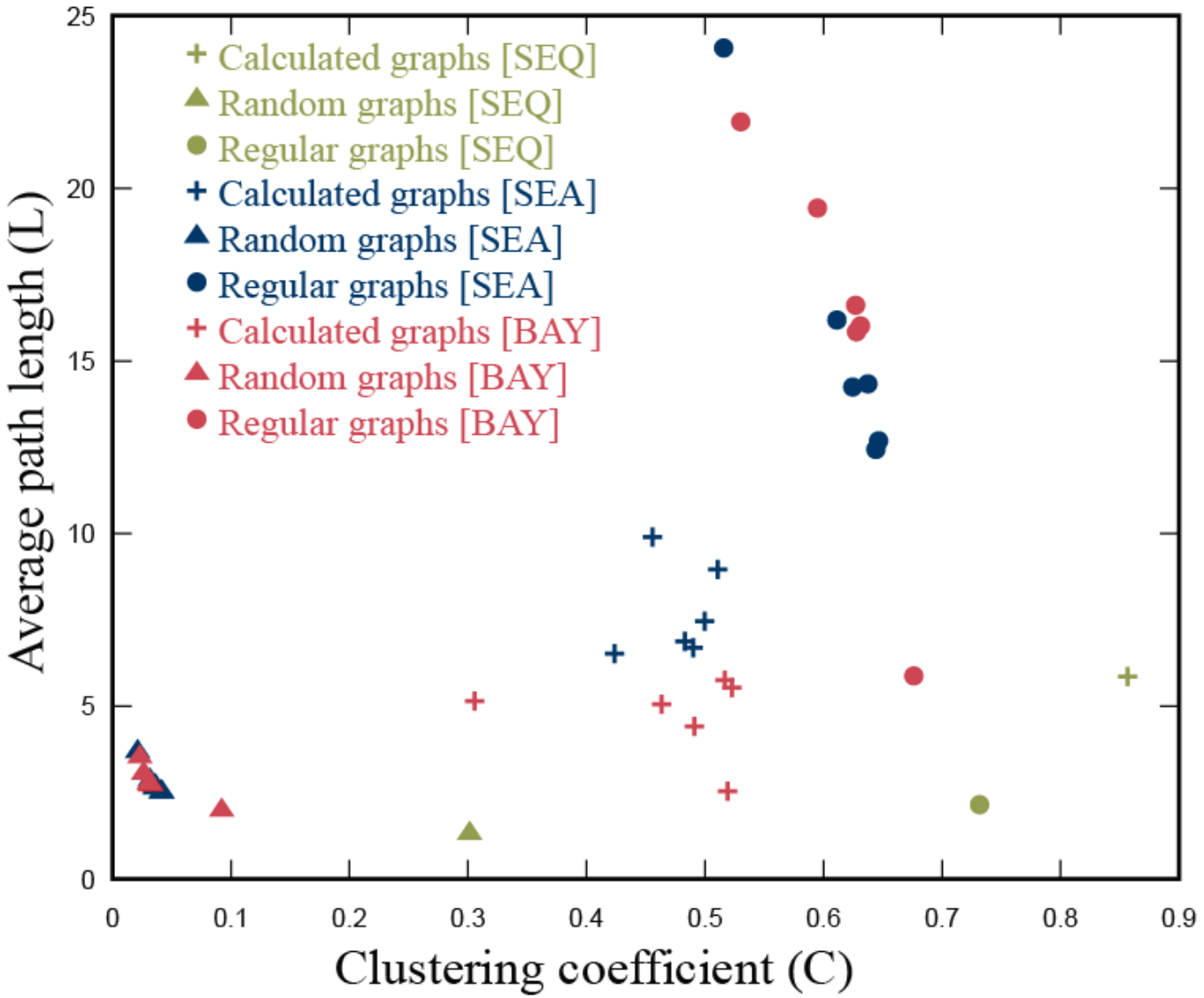


SEA

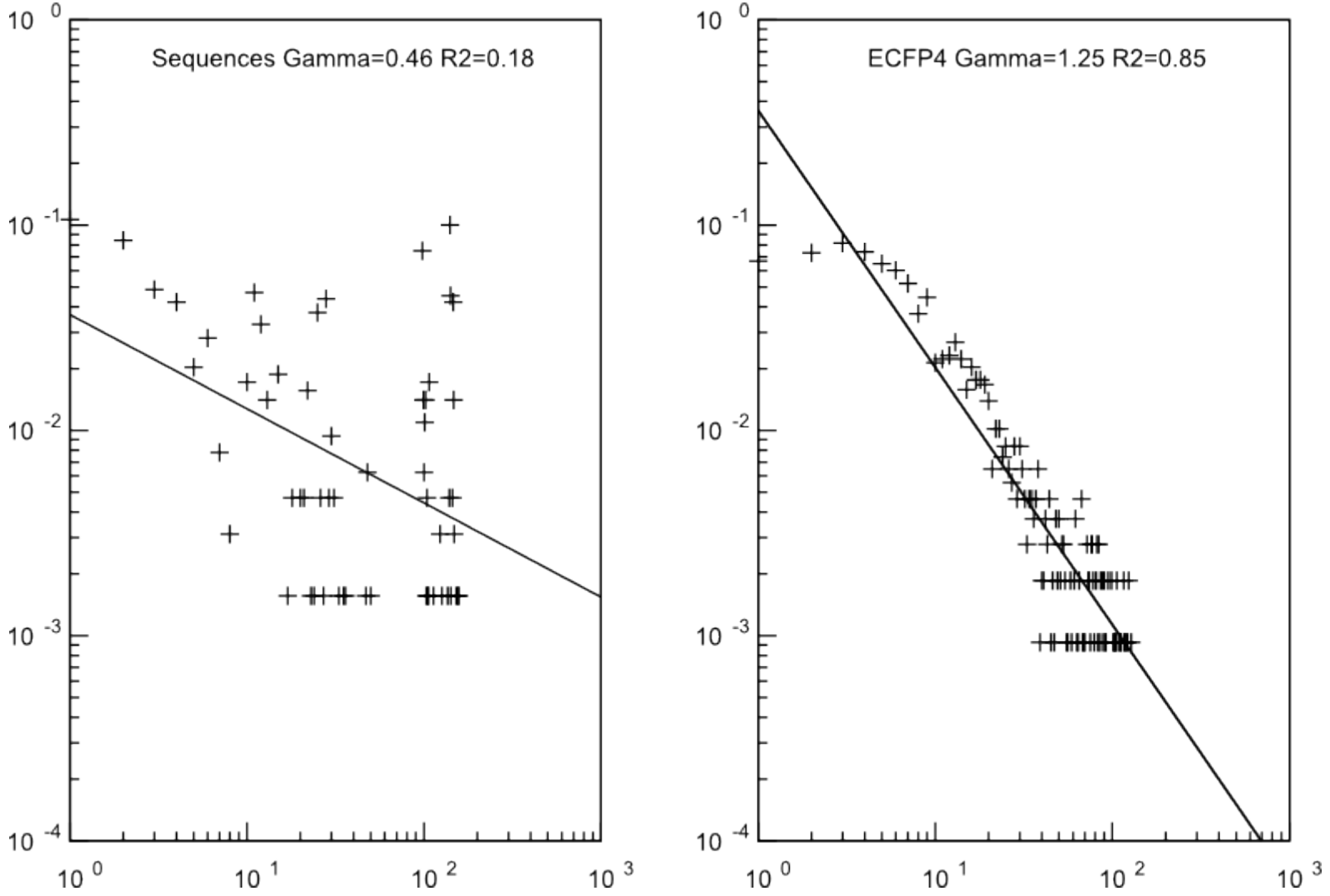


Bayesian Method

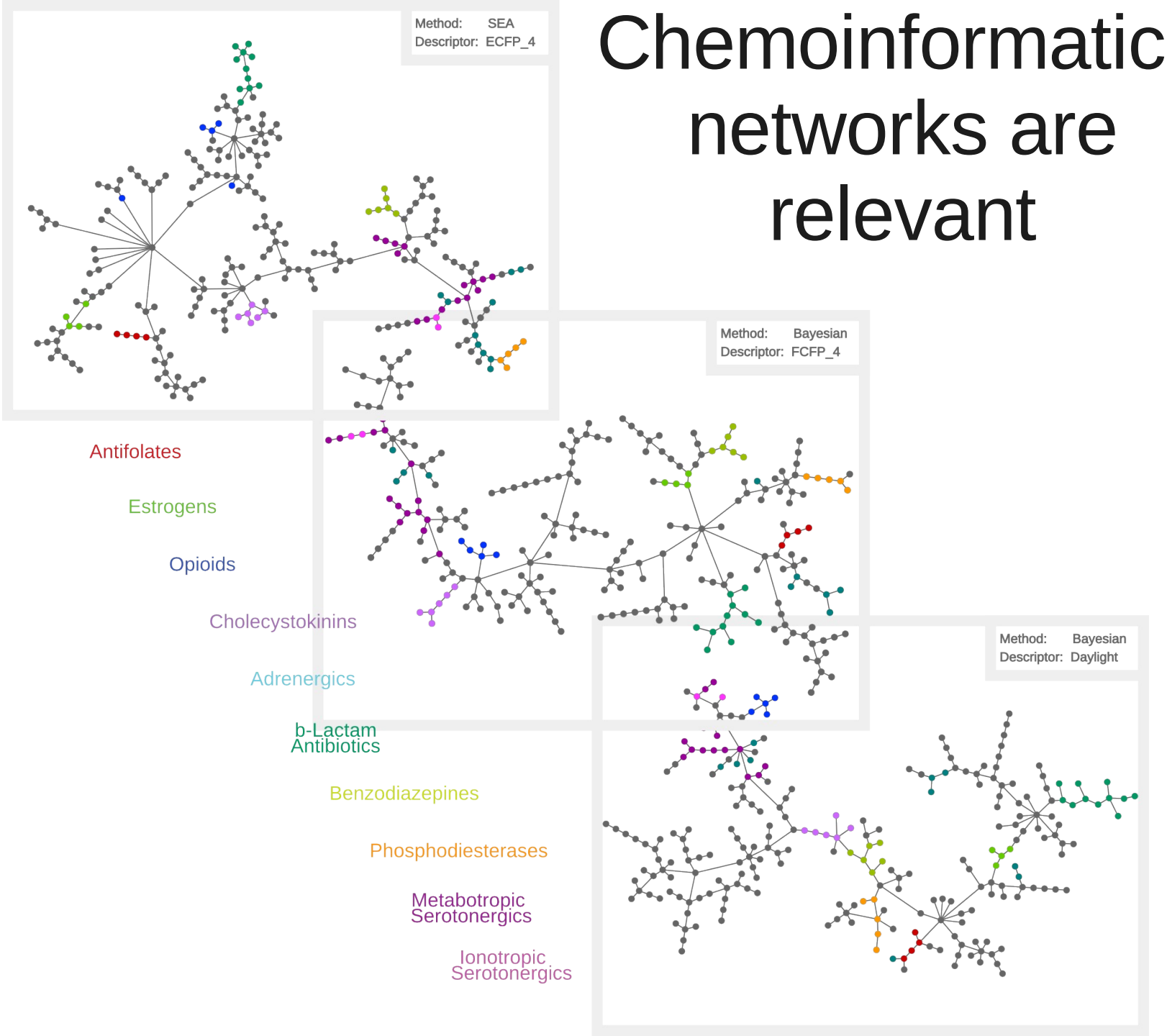
Chemoinformatics networks are "small-world"



Chemoinformatics networks are “broad-scale”

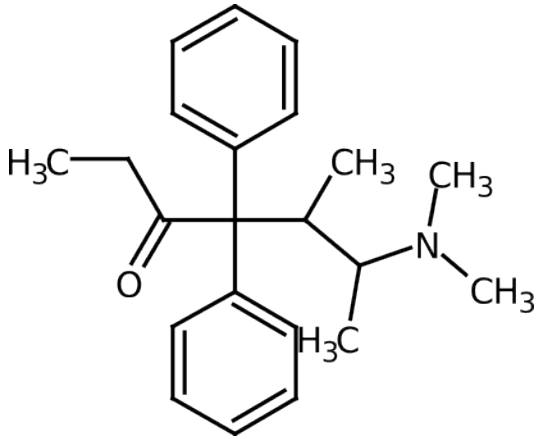


Chemoinformatics networks are relevant



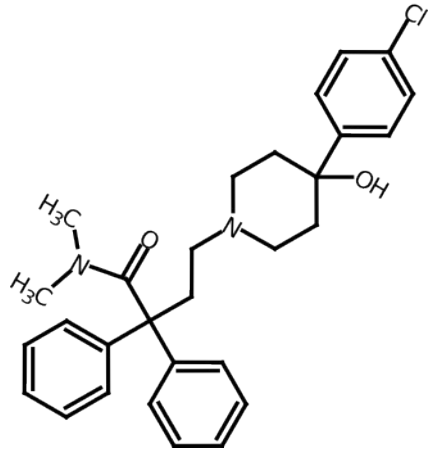
Leads to testable predictions

Methadone



- Known Opioid
- Predicted (and experimentally confirmed) Muscarinic M3 inhibitor

Loperamide



- Known Opioid
- Predicted (and experimentally confirmed) Neurokinin NK inhibitor

Take away!

- Bioinformatics and chemoinformatics networks of drug targets are different.
- Chemoinformatics networks are robust – Sets of ligands contain information about pharmacology of the protein.
- SEA and Bayesian method lead to testable predictions.

Acknowledgment

Brian K. Shoichet

John J. Irwin

http://zinc.docking.org

http://dud.docking.org

Michael J. Keiser

http://sea.docking.org

UCSF

University of California
San Francisco

MARIE CURIE ACTIONS