

# Supplementary Materials for “BMTK: a toolkit for determining modules in biological bipartite networks”

## 1 Description

**BMTK** (Bipartite Module detection ToolKit) is an online tool for determining modules in bipartite networks. It implements six methods programmed in different computing environments in a unified platform. These methods include *AdaptiveBrim* (Barber, 2007), *BCFinder* (Lehmann *et al.*, 2008), *BiLPA* (Li *et al.*, 2016), *BiSBM* (Daniel *et al.*, 2014), *LeadingEigenvector* (Newman, 2006), *LPBrim* (Liu and Murata, 2009) and *LocalBM* (Nepusz *et al.*, 2012). We also provide an online tutorial to help users to employ BMTK conveniently.

## 2 Application

We first apply BMTK onto a set of simulated data with sizes of  $100 \times 120$ ,  $250 \times 300$ ,  $500 \times 600$  and  $1000 \times 1200$  and diverse numbers of embedded modules (Table S1–S3).

Table S1: Summary of detected modules on a simulated network of size  $100 \times 120$  with 2 embedded modules.

Method	#Modules	Modularity	ARI	Time (seconds)
AdaptiveBrim	2	0.4146	1	0.0399
BiLPA	10	0.2531	0.3118	3.9659
BiSBM	2,2	N/A	N/A	0.227
LeadingEigenvector	2	0.4146	1	0.4574
LPBrim	2	0.4146	1	0.4234
LocalBM	2	0.4078	0.9640	1.7104

#Modules: the number of detected modules; Modularity: a quantitative index for module partition; ARI: adjust rand index for comparing modular results against an external criterion.

Table S2: Brief summary of detected modules for a simulated network of size  $250 \times 300$  with 5 embedded modules.

Method	#Modules	Modularity	ARI	Time (seconds)
AdaptiveBrim	5	0.5143	1	0.1671
BiLPA	27	0.2479	0.3899	40.5069
BiSBM	5,5	N/A	N/A	4.8120
LeadingEigenvector	5	0.5132	0.9954	4.9575
LPBrim	5	0.5143	1	2.8303
LocalBM	5	0.5048	0.9740	9.8748

We apply BMTK to an Ion Channel drug-target network consisting of 208 drugs, 201 targets and 1473 experimentally validated drug-target interactions (Yamanishi *et al.*, 2008). The degree

Table S3: Brief summary of detected modules for a simulated network of size  $500 \times 600$  with 10 embedded modules.

Method	#Modules	Modularity	ARI	Time (seconds)
AdaptiveBrim	6	0.3780	0.6134	0.4456
BiLPA	56	0.1922	0.3899	300.5944
BiSBM	10,10	N/A	N/A	173.8180
LeadingEigenvector	10	0.4248	1	55.6066
LPBrim	10	0.4248	1	8.2964
LocalBM	11	0.4162	0.9779	28.2682

distributions of drugs and targets are provided in Figure S1. We apply BMTK to this bipartite network, and detect a different number of drug-target modules using each method (Figures S2–S7). We collect the function information about these drugs and targets by means of KEGG API (<http://www.kegg.jp/kegg/rest/keggapi.html>) and analyze the functions of drugs and targets for each module. The details for the functional enrichment analysis on drugs and targets in each detected module can be found in Supplementary Tables S4–S10. The drugs and targets in each module usually show very similar functions. There is a drug-target module identified by almost all the methods. It is respectively the 8th, 8th, 10th/12th (drug/target module), 10th, 7th, 6th one by AdaptiveBrim, BiLPA, BiSBM, LeadingEigenvector, LPBrim, LocalBM (Supplementary Table S4–S10). It is closely related to serotonin, thus we named it as serotonin module. The serotonin module contains 10 drugs and six targets. Among them, seven drugs (i.e., D04370, D02041, D00678, D00677, D00456, D00451, and D00283) are all serotonin receptor agonists/antagonists, and five targets (i.e., hsa:9177, hsa:3359, hsa:285242, hsa:200909, and hsa:170572) all involved in serotonergic synapse and taste transduction pathways. For the drug D00563, it is a specific serotonergic antidepressant. D00513 (pindolol) is a kind of beta-adrenergic receptor agonists/antagonists and D00633 belongs to dopamine receptor agonists/antagonists. Adrenaline, dopamine and serotonin are important neurotransmitters. Thus, these drugs are all used to treat some neuropsychiatric diseases, such as depressant, anxiety, schizophrenic. For the target part, five targets (i.e., hsa:9177, hsa:3359, hsa:285242, hsa:200909 and hsa:170572) all involve in serotonergic synapse and taste transduction pathways, which are very related to the drug functions in this module. The last gene hsa:8001 does not locate in such signaling pathways, but it is still related to neural activity – neuroactive ligand-receptor interaction. Therefore, this drug-target module is closely related to serotonin. Besides, we are able to find some potential drug-target interactions from this module.

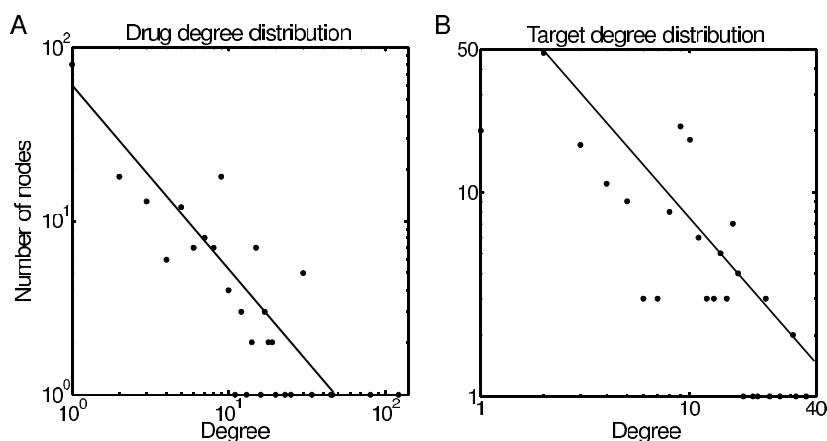


Figure S1: The degree distributions of (A) drugs and (B) targets in the Ion Channel drug-target bipartite network.

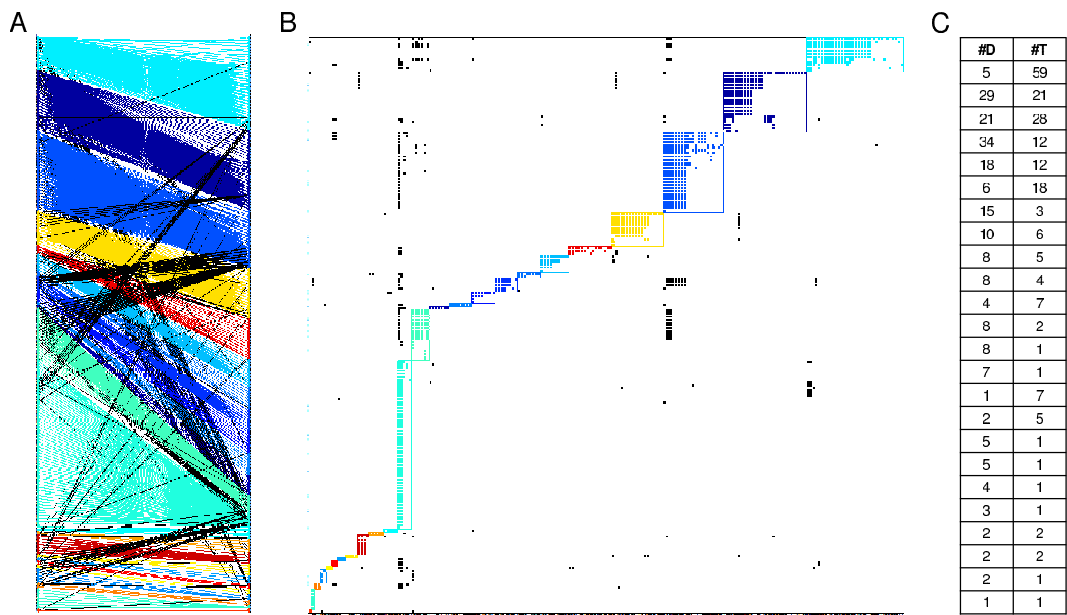


Figure S2: The bipartite modules computed by AdaptiveBrim with two types of visualization: (A) The drug-target bipartite network with 24 identified drug-target modules marked in different colors; (B) Reordered adjacency matrix of the input bipartite network marked in different colors. (C) The numbers of drugs (#D) and targets (#T) for each module are shown.

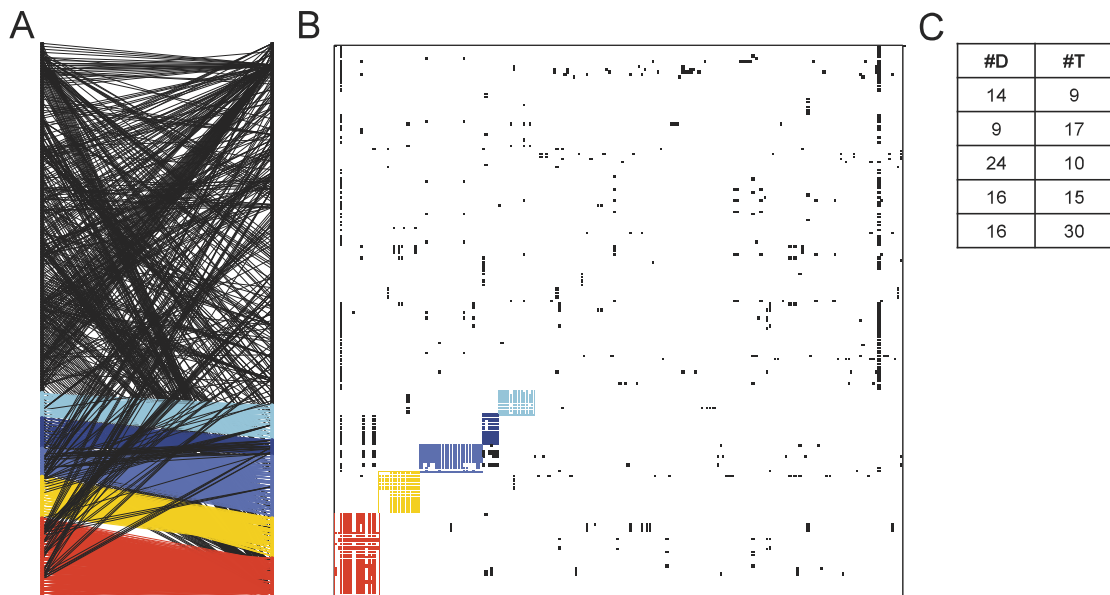


Figure S3: The bipartite modules computed by BCFinder.

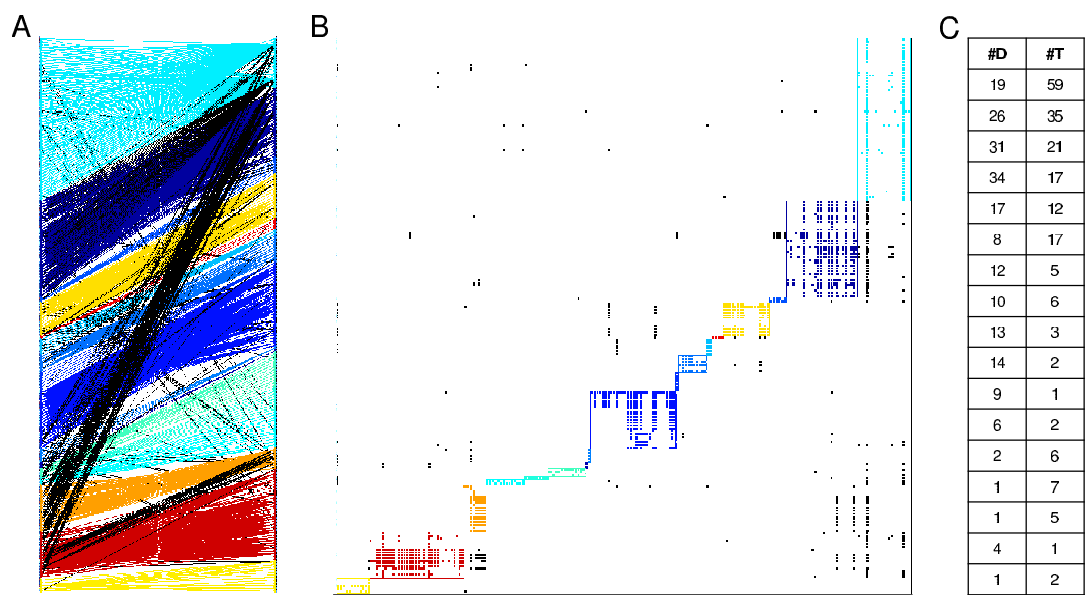


Figure S4: The bipartite modules computed by BiLPA.

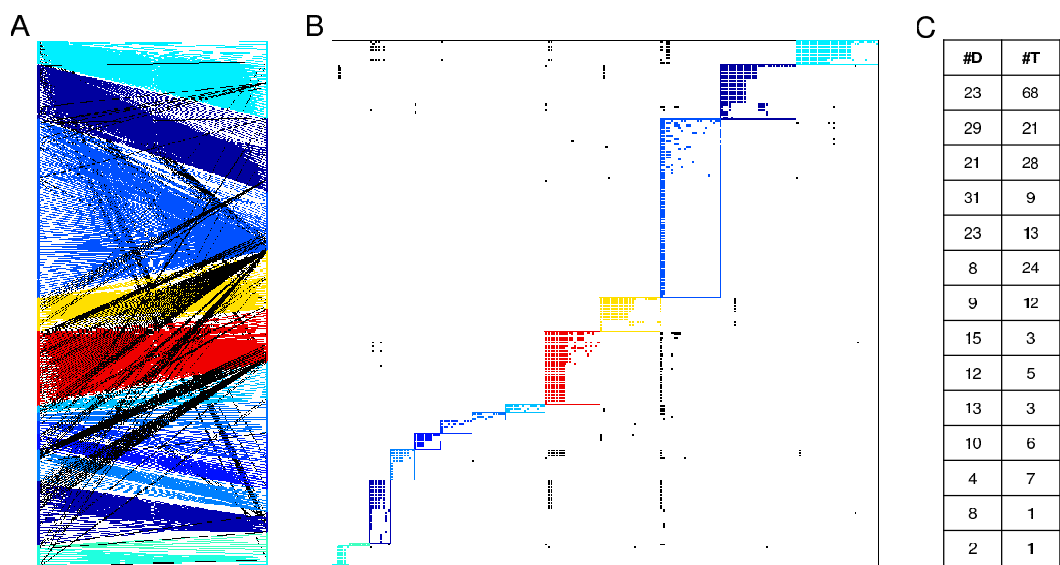


Figure S5: The bipartite modules computed by LeadingEigenvector.

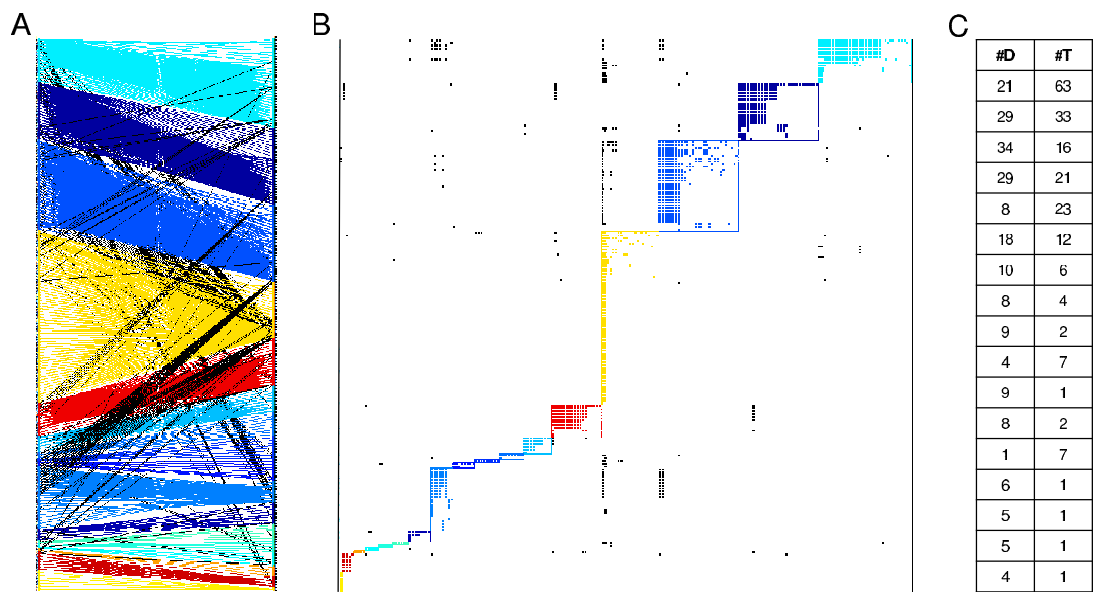


Figure S6: The bipartite modules computed by LPBrim.

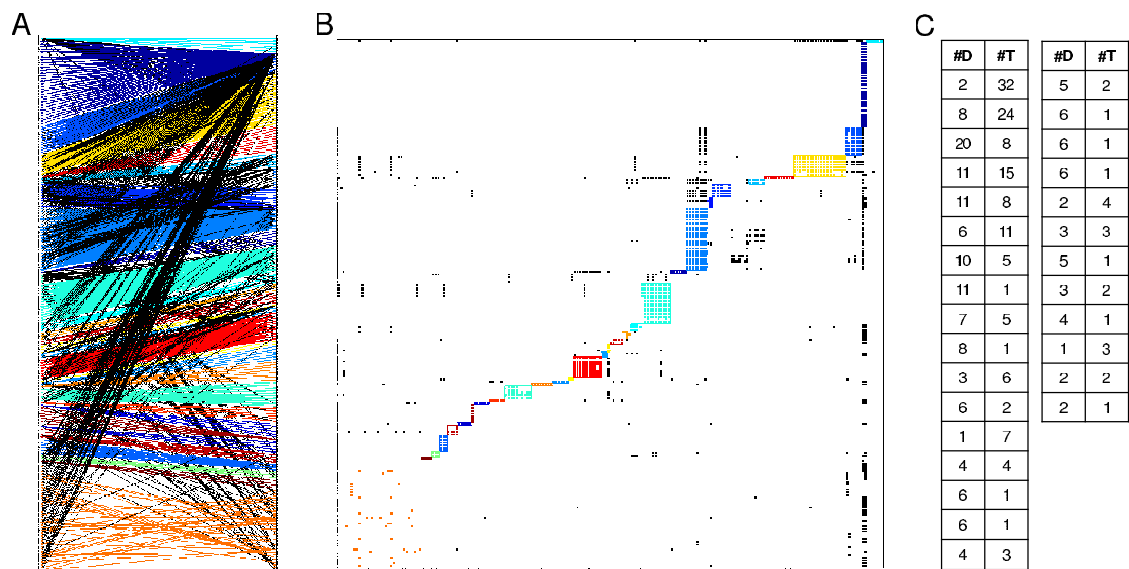


Figure S7: The bipartite modules computed by LocalBM.

## References

- Barber, M.J. (2007) Modularity and community detection in bipartite networks, *Phys. Rev. E*, **76**, 066102.
- Daniel, B.L. *et al.* (2014) Efficiently inferring community structure in bipartite networks, *Phys. Rev. E*, **90**, 012805.
- Lehmann, S. *et al.* (2008) Biclique communities, *Phys. Rev. E*, **78**, 016108.
- Li, Z.P. *et al.* (2016) Quantitative function and algorithm for community detection in bipartite networks, *Information Sciences*, **367-368**, 874-889.
- Liu, X. and Murata, T. (2009) Community detection in large-scale bipartite networks, *Web Intelligence and Intelligent Agent Technologies*, **1**, 50-57.
- Nepusz, T., Yu, H. and Paccanaro, A. (2012) Detecting overlapping protein complexes in protein-protein interaction networks. *Nature Meth.*, **9**, 471-472.
- Newman, M.E.J. (2006) Modularity and community structure in networks, *PNAS*, **23**, 8577-8582.
- Santo, F. (2010) Community detection in graphs, *Phy. Rep.*, **486**, 75-174.
- Yamanishi, Y. *et al.* (2008) Prediction of drug-target interaction networks from the integration of chemical and genomic spaces, *Bioinformatics*, **24**, 232-240.