

## Yale University EliScholar – A Digital Platform for Scholarly Publishing at Yale

---

Yale Day of Data

Day of Data 2013

---

# Detecting modules in multiplex networks – an application for integrating expression profiles across multiple species

Koon-Kiu Yan

Yale University, [koon-kiu.yan@yale.edu](mailto:koon-kiu.yan@yale.edu)

Daifeng Wang

Yale University

Joel Rozowsky

Yale University

Henry Zheng


Yale University

Baikang Pei

Yale University

*See next page for additional authors*

Follow this and additional works at: <http://elischolar.library.yale.edu/dayofdata>

 Part of the [Bioinformatics Commons](#), [Computational Biology Commons](#), and the [Genomics Commons](#)

---

Koon-Kiu Yan, Daifeng Wang, Joel Rozowsky, Henry Zheng, Baikang Pei, and Mark Gerstein, "Detecting modules in multiplex networks – an application for integrating expression profiles across multiple species" (September 19, 2013). *Yale Day of Data*. Paper 6. <http://elischolar.library.yale.edu/dayofdata/2013/Posters/6>

This Event is brought to you for free and open access by EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Yale Day of Data by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact [elischolar@yale.edu](mailto:elischolar@yale.edu).

---

**Presenter/Creator Information**

Koon-Kiu Yan, Daifeng Wang, Joel Rozowsky, Henry Zheng, Baikang Pei, and Mark Gerstein

# Detecting modules in multiplex networks

## – an application for integrating expression profiles across multiple species

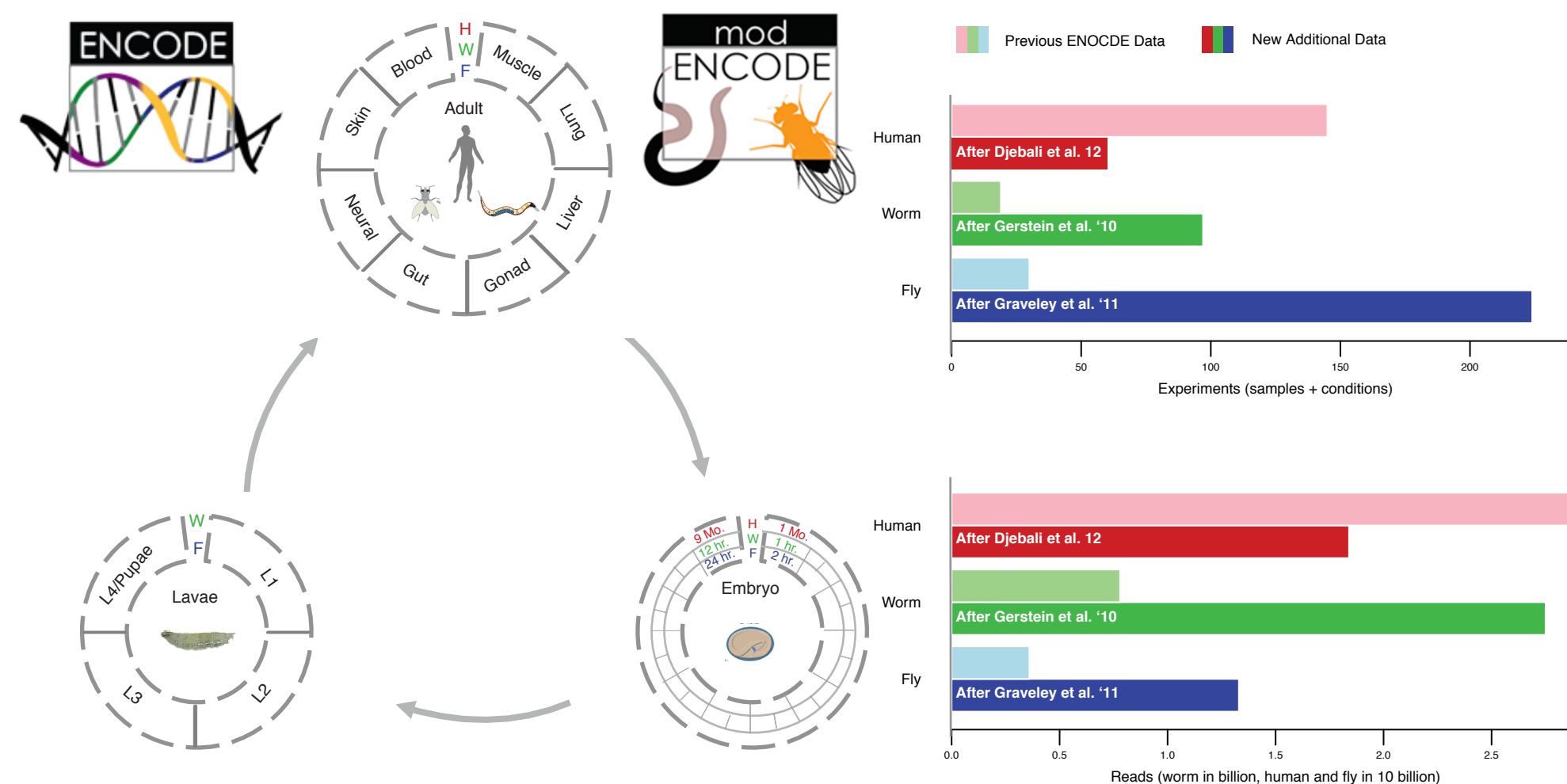
Koon-Kiu Yan<sup>\*1,2</sup>, Daifeng Wang<sup>\*1,2</sup>, Joel Rozowsky<sup>1,2</sup>, Henry Zheng<sup>1,2</sup>, Baikang Pei<sup>1,2</sup>, Mark Gerstein<sup>1,2,3</sup>

<sup>1</sup>Program in Computational Biology and Bioinformatics, <sup>2</sup>Department of Molecular Biophysics and Biochemistry, <sup>3</sup>Department of Computer Science, Yale University

### Abstract

Multiplex network, a set of networks linked through interconnected layers, is a useful mathematical framework for data integration. Here, we present a general method to detect modules in multiplex networks and apply it in a specific biological context: to simultaneously cluster the genome-wide expression profiles of *C. elegans* and *D. melanogaster* generated by the ENCODE and modENCODE consortia. The method revealed modules that are fundamentally cross-species and can either be conserved or species-specific. In general, the method could be applied in various contexts like the integration of different social networks.

### Motivation: Integrating Big Data in Genomics



### A spin model to detect modules in a multiplex network

Modules are densely connected regions in a network.

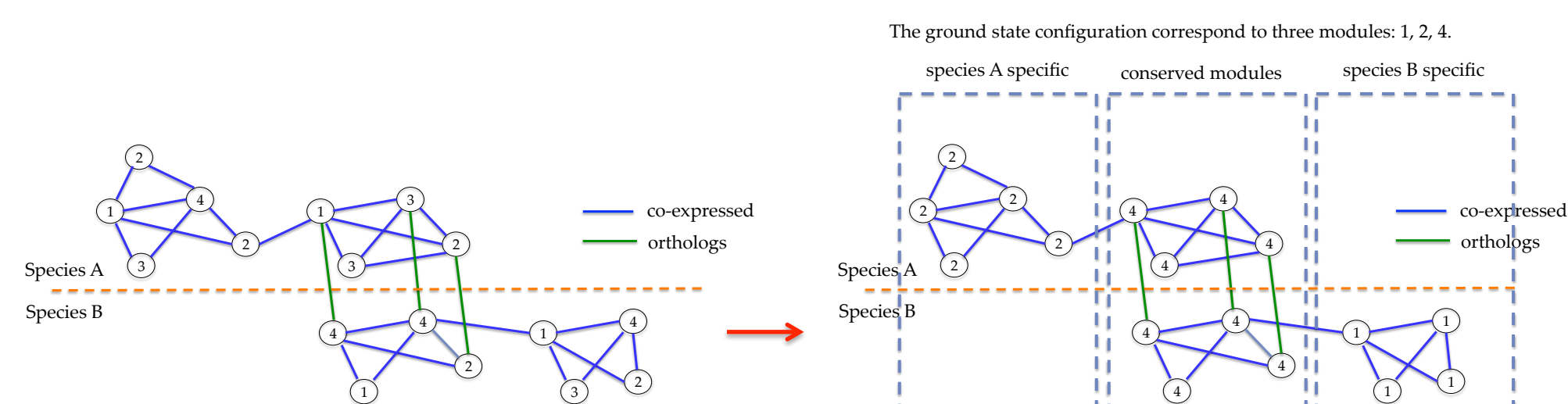
Every node  $i$  is assigned with a spin value  $\sigma_i$  (labels of modules:  $1, 2, \dots, q$ ).

$$H = \sum_{i,j} (-W_{ij}^{(A)} + p_{ij}^{(A)}) \delta_{\sigma_i \sigma_j} + \sum_{i,j} (-W_{ij}^{(B)} + p_{ij}^{(B)}) \delta_{\sigma_i \sigma_j} - \kappa \sum_{(i,j) \in \text{Ortho}} \delta_{\sigma_i \sigma_j}$$

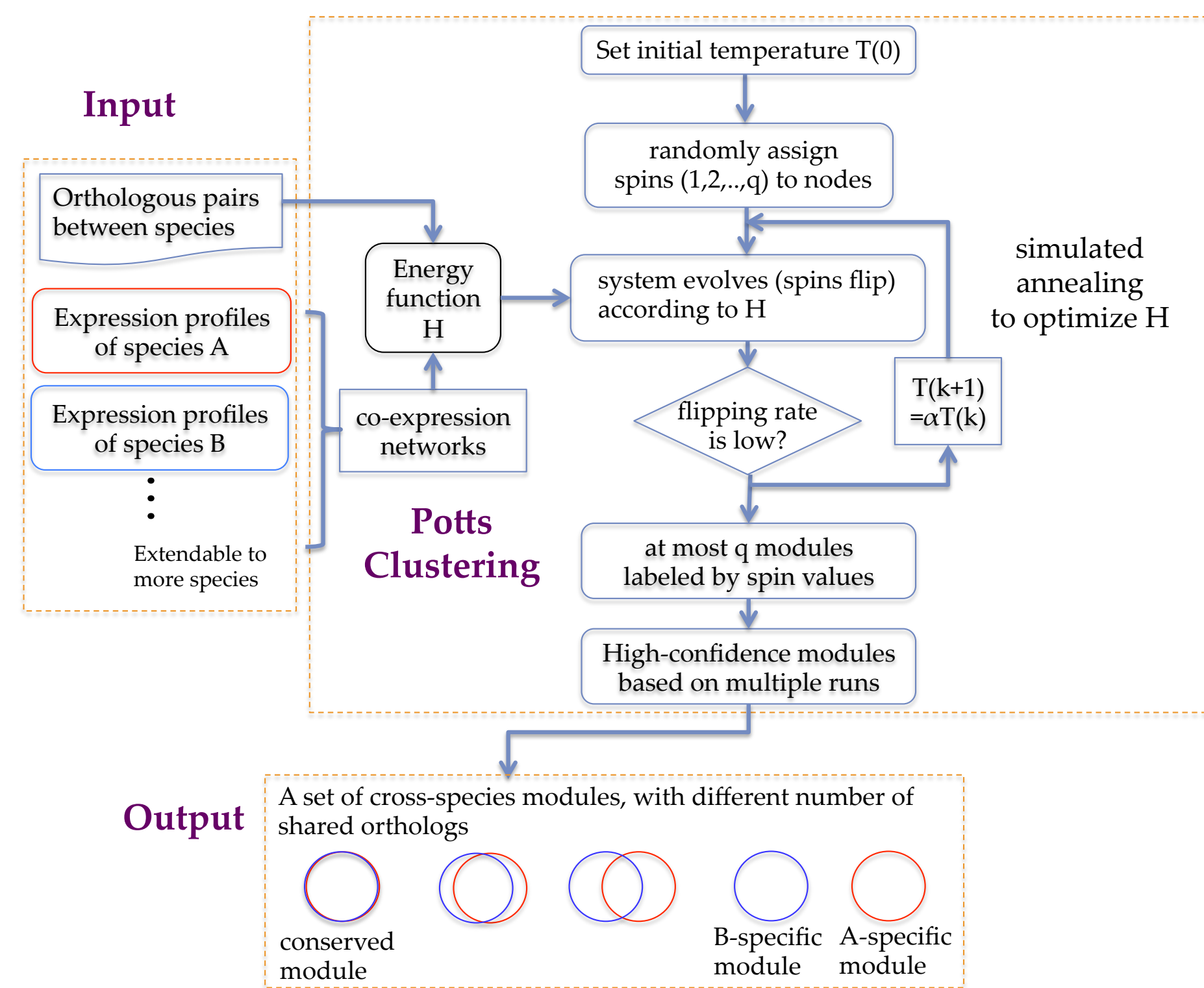
reward a co-expressed pair with the same value

punish a non co-expressed pair with the same value

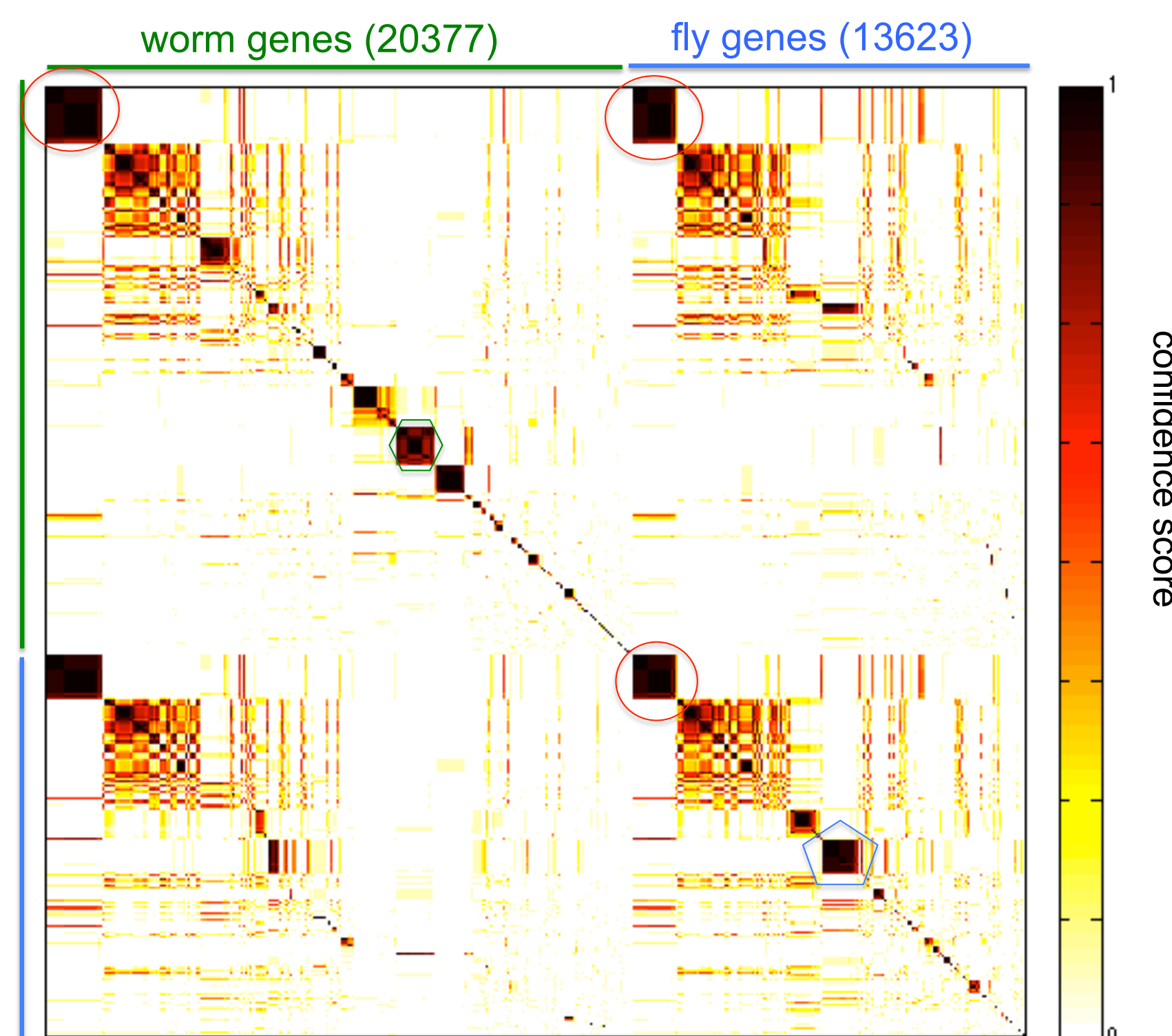
reward an orthologous pair with the same value



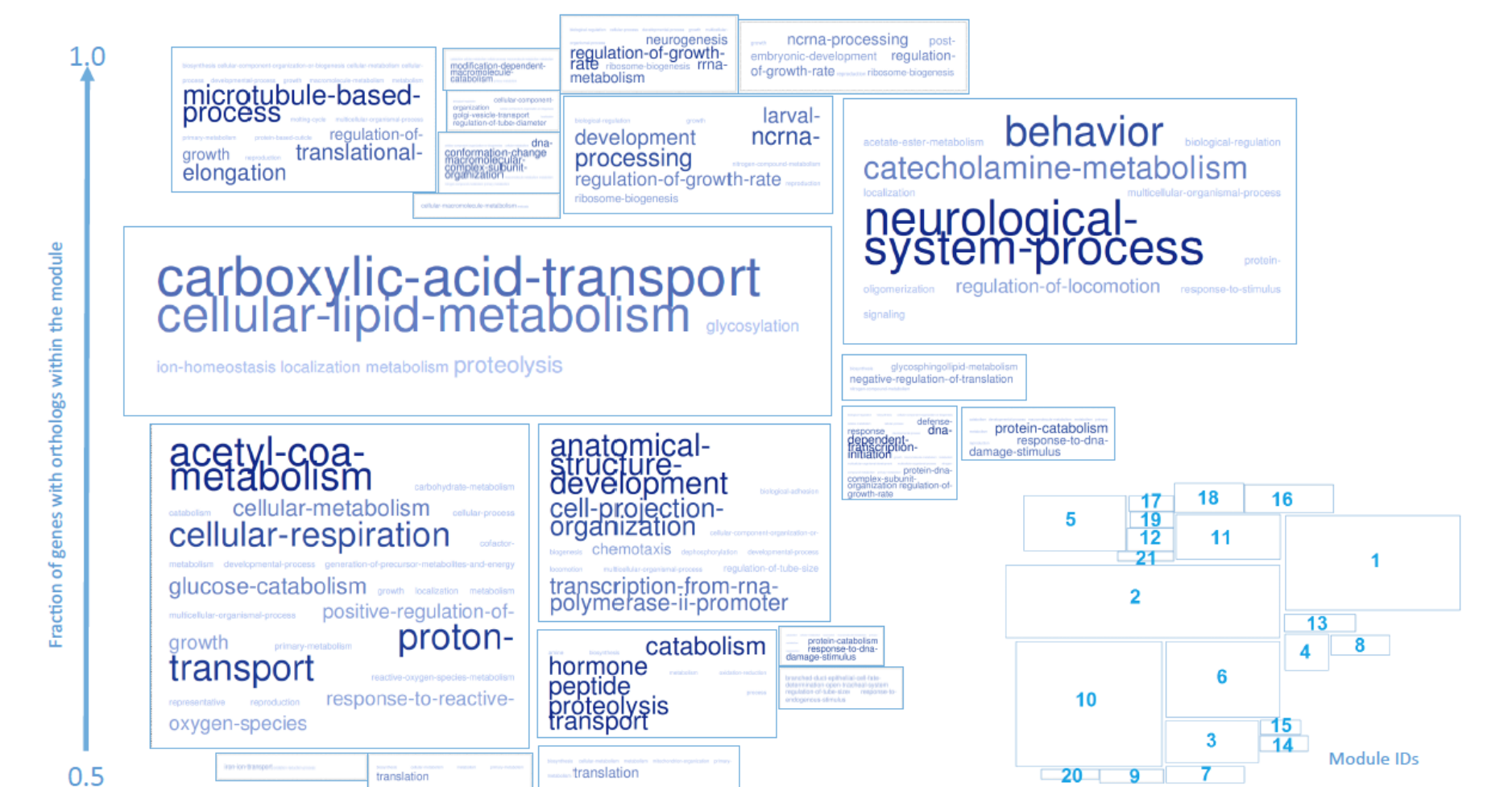
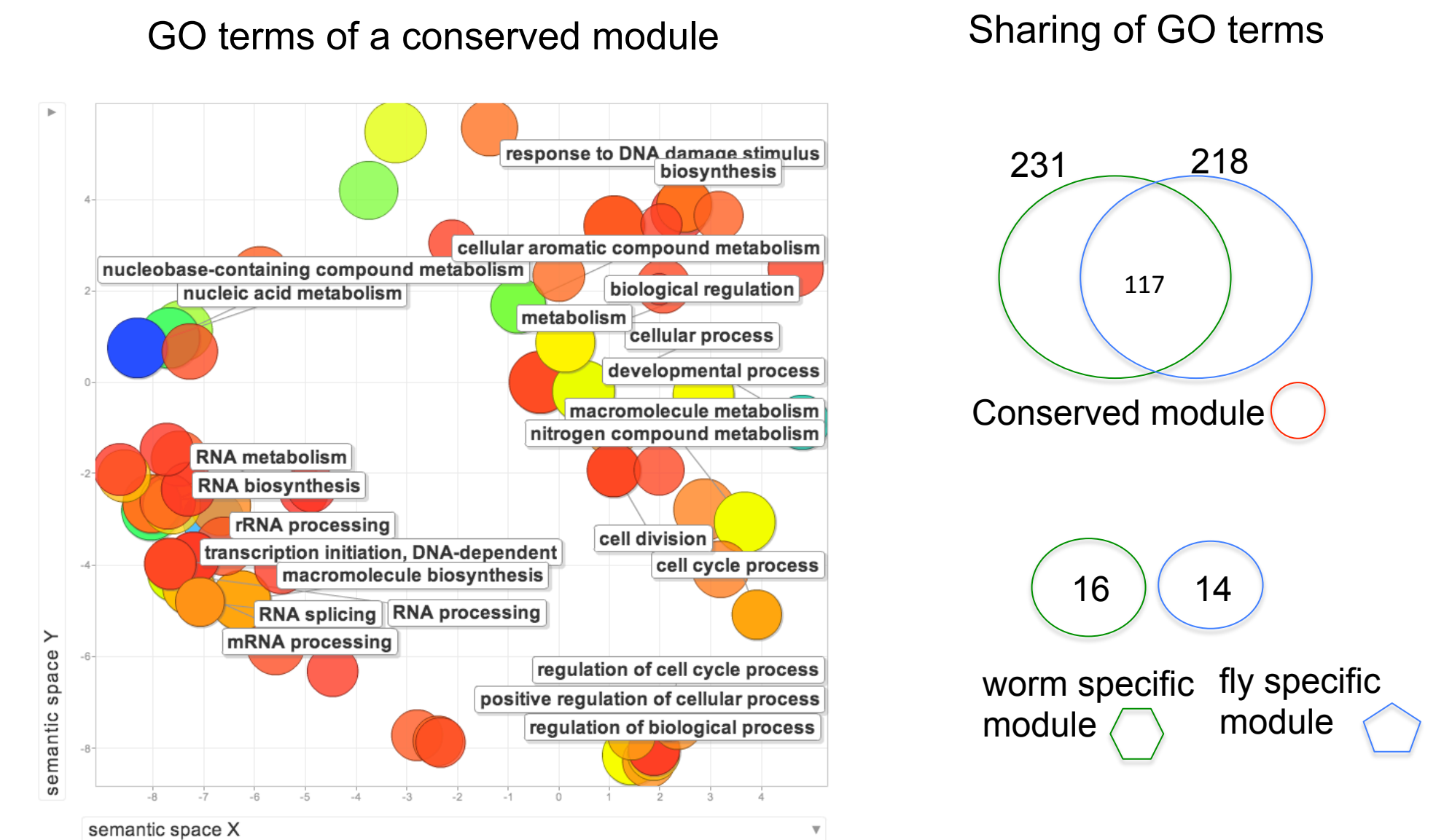
### Method Overview



### Simultaneous clustering of expression profiles



### Conserved modules and species-specific modules



### Further Application and Outlook

- Modules can be used to infer the functions of proteins as well as non-coding RNAs based on "guilt-by-association". ncRNAs from the worm and fly anchored to a module may potentially have analogous functions.
- Applying other optimization techniques to speed up the algorithm.
- Various social networks (Facebook and LinkedIn) can be integrated into a multiplex network. The method presented could be used to detect communities in a social setting.