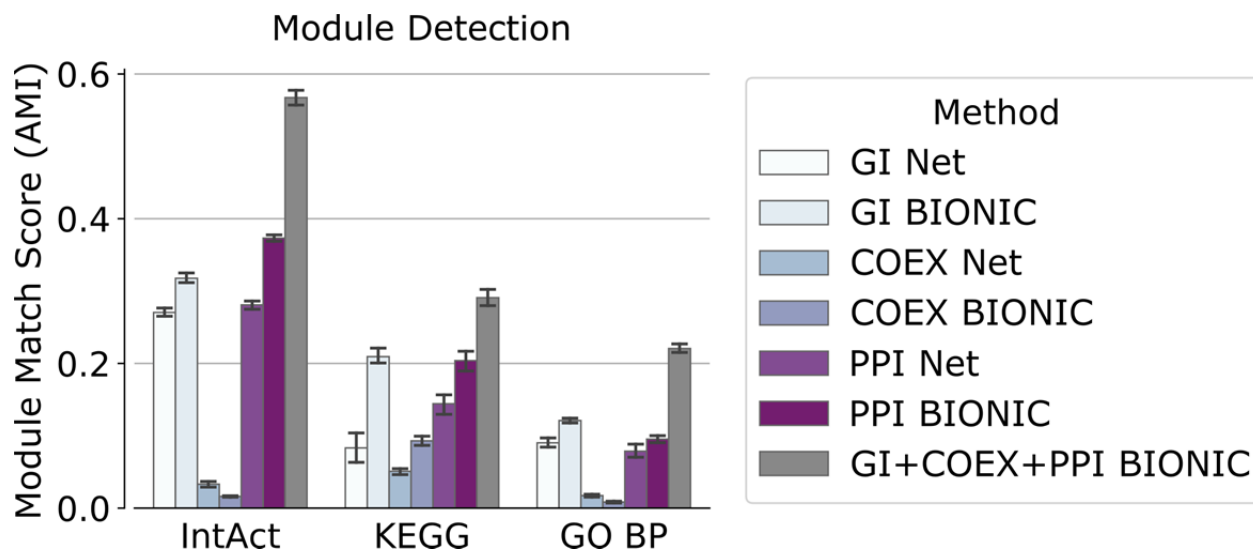
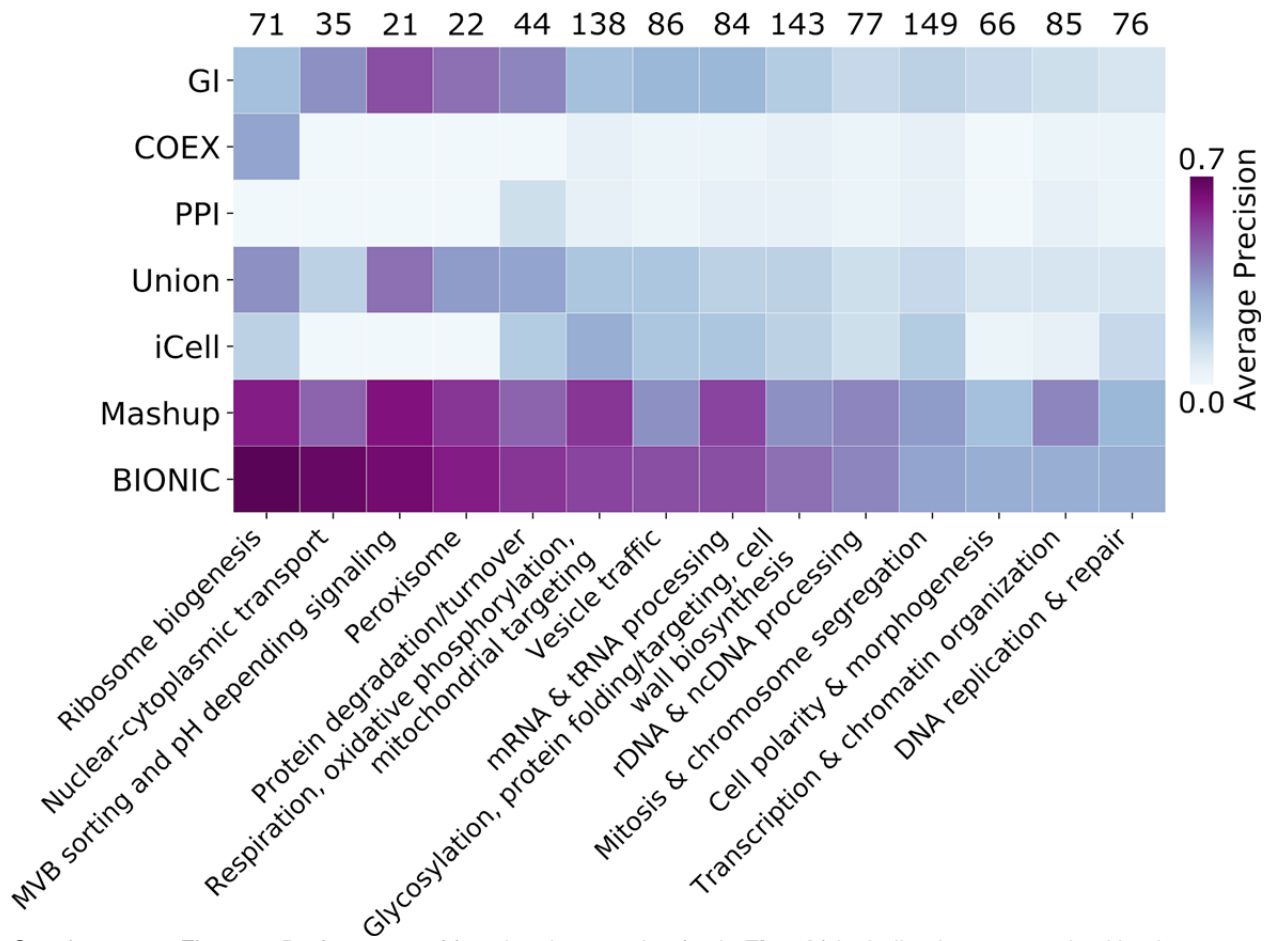


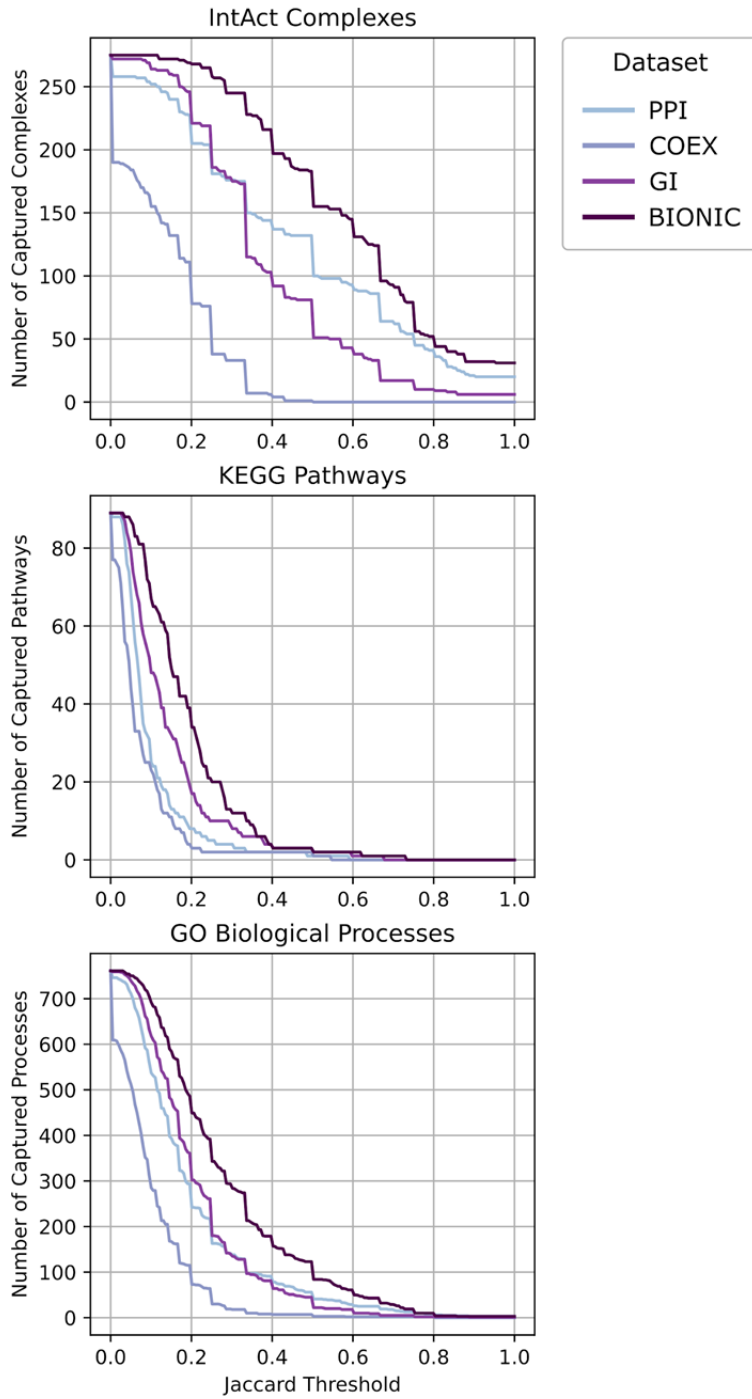
Supplementary Figures



Supplementary Figure 1. Running individual networks through BIONIC generally produces features (ending with “BIONIC”) with higher module detection performance than the original networks (ending with “Net”). Integrating these networks (denoted “GI+COEX+PPI BIONIC”) yields features that outperform the individual networks. BP = biological process, GI = genetic interaction, COEX = co-expression, PPI = protein-protein interaction. These are the same networks as Fig. 2. Error bars indicate the 95% confidence interval.



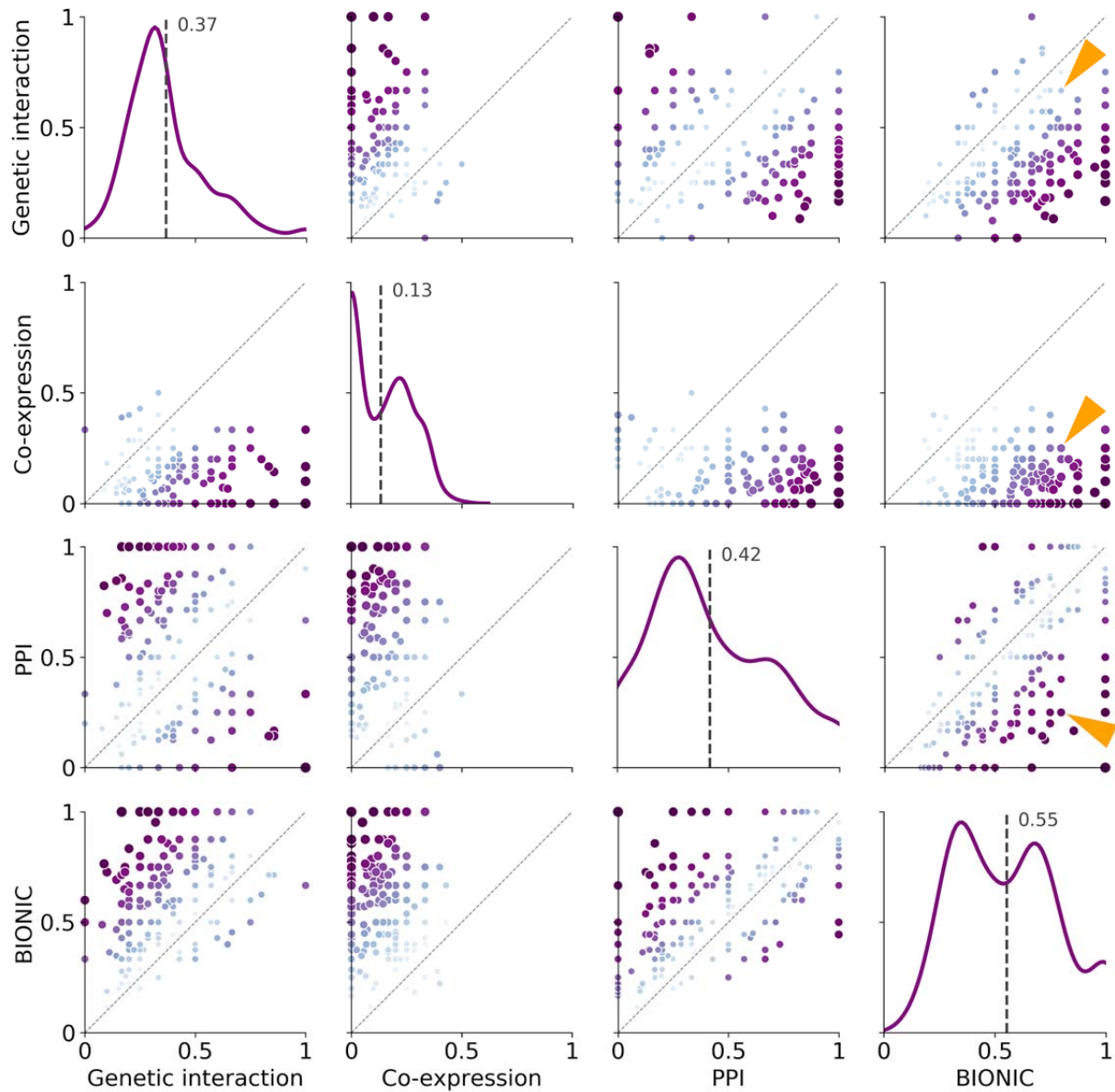
11
 12 Supplementary Figure 2. Performance of functional categories (as in Fig. 2b) including input networks. Numbers
 13 above each column represent the size of the functional category (in number of genes). GI = genetic interaction,
 14 COEX = co-expression, PPI = protein-protein interaction



15
16
17
18
19
20
21
22
23
24
25

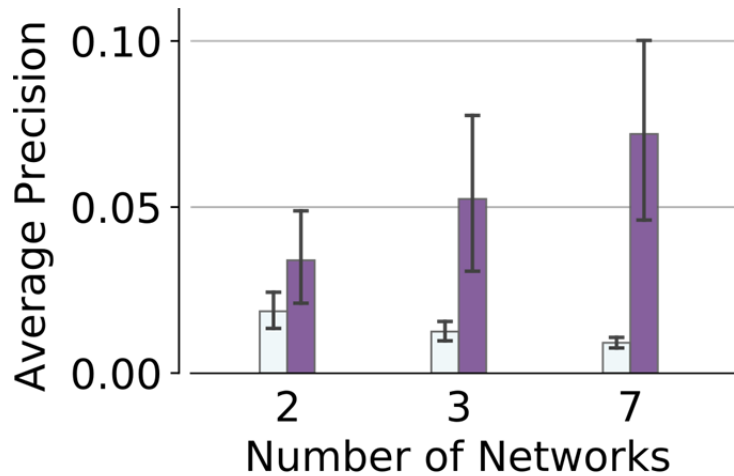
Supplementary Figure 3. Coverage of functional gene modules by individual networks and the BIONIC integration of these networks (denoted BIONIC), as determined by a parameter optimized module detection analysis where the clustering parameters were optimized for each module individually. The number of captured modules is reported for a range of overlap scores (Jaccard threshold). Higher threshold indicates greater correspondence between the clusters obtained from the dataset and their respective modules given by the standard. GI = genetic interaction, COEX = co-expression, PPI = protein-protein interaction. These are the same networks and BIONIC features as Fig. 2.

IntAct Complexes



26
27
28
29
30
31
32
33
34
35
36
37
38
39
40

Supplementary Figure 4. Known protein complexes (as defined by the IntAct standard) captured by individual networks and the BIONIC integration of these networks (denoted BIONIC). Hierarchical clustering was performed on the datasets and resulting clusters were compared to known IntAct complexes and scored for set overlap using the Jaccard score (ranging from 0 to 1). The clustering algorithm parameters were optimized for each module individually, unlike the analysis in Fig. 2 where the clustering parameters were optimized for the standard as a whole. Each point is a protein complex, as in Fig. 2c. The dashed line indicates instances where the given data sets achieve the same score for a given module. Histograms indicate the distribution of overlap (Jaccard) scores for the given dataset, and the labelled dashed line indicates the mean of this distribution. The individual modules shown here as well as for the KEGG Pathways and IntAct Complexes module standards can be found in **Supplementary Data File 4**. The SEC62-SEC63 complex is indicated by the arrows. GI = genetic interaction, COEX = co-expression, PPI = protein-protein interaction. This analysis uses the same networks and BiONIC features as Fig. 2.



Method Mashup SVD Mashup

Supplementary Figure 5. Comparison of Mashup with the author provided singular value decomposition approximation of Mashup (denoted Mashup SVD). Random sets of yeast co-expression networks were sampled, integrated and scored against the KEGG pathways co-annotation standard, analogous to Fig. 3a. Error bars indicate the 95% confidence interval.

Supplementary Data Files

Supplementary Data File 1

Integrated Network Details. Publication, gene count, edge count and experimental type for each yeast network and each human network used in Fig. 2 and 3. Rows in yellow indicate the yeast networks used in Fig. 2 integrations. <https://docs.google.com/spreadsheets/d/1iEiugyP4CBnmRmibVc0KRpMWODda5sQfyWpGA2oAEWw/edit?usp=sharing>

Supplementary Data File 2

Figure 2 Evaluation Standards Details. Gene count, co-annotation count, module count and class count details for each standard used in the Fig. 2 and 3 evaluations. https://docs.google.com/spreadsheets/d/1B-EsF1zVVa23ssol5p3hYSNUbTOx3q5x_8LEj0ocnIQ/edit?usp=sharing

Supplementary Data File 3

Figure 2 Module Detection Results. Overlap of standard-optimized clusters obtained from the Fig. 2 module detection analysis for networks and BIONIC features. Module standards are IntAct Complexes, KEGG Pathways and GO Biological Processes. <https://docs.google.com/spreadsheets/d/1TKXw-GqklvFR1YU8GBgYOHWJPn-CK4yqJe3WE--g8K4/edit?usp=sharing>

Supplementary Data File 4

Fig. S3, S4 Module Detection Results. Overlap of known per-module-optimized clusters obtained from the Fig. 2 input networks and integration methods, with IntAct Complex, KEGG Pathway and GO Biological Process modules. https://docs.google.com/spreadsheets/d/1UQ1zFnK2PY3ojaYixU2WELVmubS_6kM2-858-veZSBs/edit?usp=sharing

Supplementary Data File 5

Integrated BIONIC Features and Evaluation Standards. Learned BIONIC features from yeast networks integrated in Fig. 2, and standards used in all evaluations. <https://data.wanglab.ml/BIONIC/>