negative rate or TNR), while the latter mimics the concept of sensitivity (aka as true positive rate or TPR). Following the approach of Kalaev et al., we process each alignment graph separately and apply a threshold of 0.05 to extract enriched GO terms. Table 2 summarizes the computed statistics.

In terms of overall sensitivity (for each pair of species in the alignment graph), mat3_auction shows superior performance, except in aligning human-versus-fly, where mat3_greedy bests the mat3_auction method. It is notable here that sensitivity is not comparable among different tables, since the total number of enriched terms is a function of evolutionary distance among species pairs that are being aligned. A pair of species that have diverged more recently are more probable to have common pathways. However, for a fixed pair of species, which define a unique functional space over the alignment graph, we can use sensitivity to compare different methods.

![Conserved components identified by the mat3-auction method](image)

Figure 6: Conserved components identified by the mat3-auction method. Each component is coherently annotated with specific branch of the biological process. (a) Peroxisome organization. (b) RNA splicing. (c) Histone modification. (d) Ribosome biogenesis.

In terms of average specificity, we observe more diversity among different methods. Surprisingly, mat3_greedy is the top-ranked method in 4 out of 6 experiments, except in aligning yeast-versus-bacterium and bacterium-versus-human, for which iso_hungarian and mat3-auction perform better respectively. These results suggest that using the well-known Hungarian algorithm, which is a maximum weighted bipartite matching algorithm, does not necessarily enhance the biological significance of the results. One possible explanation is the over-fitting problem occurring in this case. The objective function here is defined based on the pairwise similarity of aligned nodes in different graphs. Recalling that these similarity scores are initially computed using heuristic methods, they contain significant noise and over-fitting a