

# Prediction-based evaluation of community detection methods for maize gene expression data

*This manuscript ([permalink](#)) was automatically generated from [thferebee/PhenoEval-Coexpress-manubot@095ac3d](#) on August 28, 2020.*

## Authors

---

- **Taylor Ferebee**

 [XXXX-XXXX-XXXX-XXXX](#) ·  [thferebee](#) ·  [Ferebee2PhD](#)

Department of Computational Biology, Cornell University · Funded by Grant XXXXXXXX

- **Jane Roe**

 [XXXX-XXXX-XXXX-XXXX](#) ·  [janeroe](#)

Department of Something, University of Whatever; Department of Whatever, University of Something

# Abstract

---

## Visual Abstract

---

## Introduction

---

Since the early days of microarray experiments, much of biology has focused on inferring gene function on a global scale through the creation of gene coexpression networks. Gene co-expression networks (GCNs) are an intuitive way to understand the relationships between genes with similar gene expression profiles. In these networks, nodes represent genes and the connection between genes (edges) provide a visual

These relations are often found through mathematical methods such as Pearson correlation, mutual information

## Materials and Methods

---

### Community Detection Methods

**Hierarchical Clustering on a Topological Overlap Matrix** One of the most common clustering methods, agglomerative hierarchical clustering (implemented in WGCNA R package) attempts to create a hierarchy of clusters by starting with individual clusters and building to one cluster with all data. To begin the procedure, each gene is in its own cluster and a similarity (distance) matrix is considered. If two clusters (genes) are similar (close in distance), these two genes merge. If two clusters are not close, their clusters do not merge. Once again, we consider a similarity matrix between the newly formed clusters. This repeats until all clusters have merged into a single cluster. Most often, hierarchical clustering is presented in a dendrogram form which tracks the merges of the clusters.

**Fast Greedy** Fast greedy community detection uses topological features of a network and an agglomerative algorithm for detecting community structure. The fast greedy algorithm utilizes a network's modularity. The modularity measures the relationship between the density of links within communities compared with the rest of the network. In other words, a particularly modular network (as biological networks are assumed to be) have communities that are very connected within the community, but only have a few edges to other communities.

## Results

---

## Discussion

---

# References

---