

Modularity Based Scoring Subnetwork (MoBaS)

Marzieh Ayati and Mehmet Koyuturk

MoBaS is a software tool for finding the disease associated subnetworks in protein-protein interaction (PPI) networks. A disease associated subnetwork is a subset of gene products that are heavily connected to each other in the protein-protein interaction network and are significantly associated with the disease when considered together.

MoBaS uses a scoring method that is parameter-free and aims to score subnetworks by assessing the disease association of pairwise interactions and incorporating the statistical significance of network connectivity and disease association.

To run MoBas, three input files are required.

1. SNP Association
2. Mapping SNP to Gene
3. Protein-protein interaction network

Modularity Based Scoring Subnetwork (MoBaS)

Marzieh Ayati and Mehmet Koyuturk

SNP Association

This file can be produced by PLINK.

SNP Association File is a file that represents the p-value of each SNP with the disease computed by PLINK [1]. (.assoc)

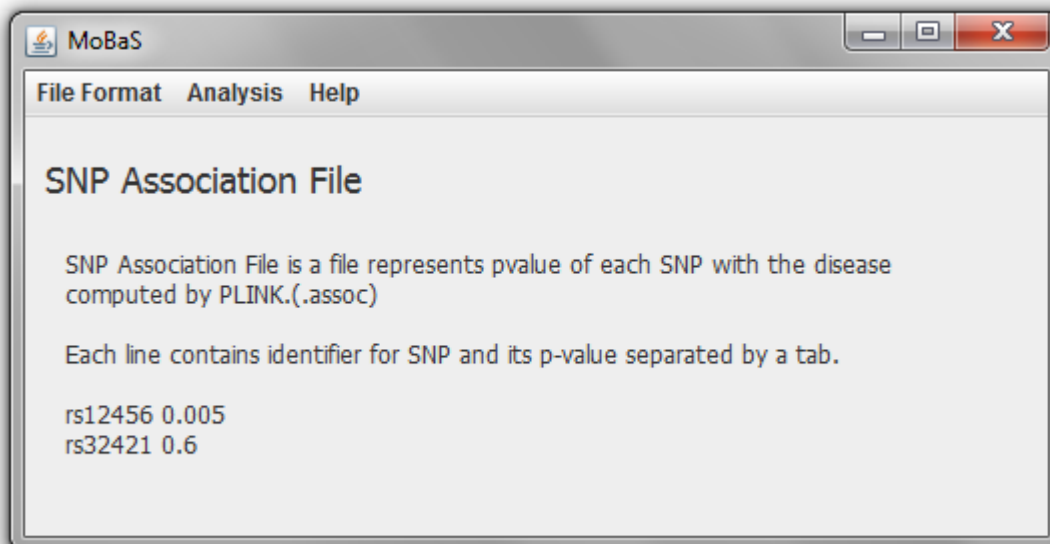
Each line contains the identifier for the SNP and its p-value separated by a tab.

Here is a sample of .assoc file as an input

```
rs12456 0.005
```

```
rs32421 0.06
```

Users can check the files in the Association folder. These files have been produced for Psoriasis disease.



Modularity Based Scoring Subnetwork (MoBaS)

Marzieh Ayati and Mehmet Koyuturk

Mapping SNP to Gene

Mapping File provides mapping file to gene symbol.

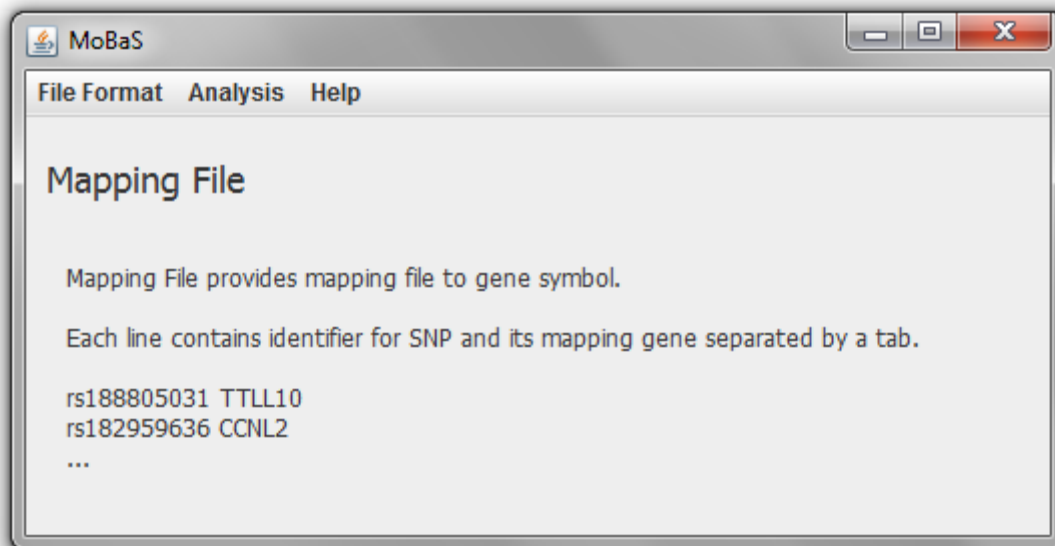
Each line contains identifier for SNP and its mapping gene separated by a tab.

Here is a sample for mapping file

```
rs188805031 TTLL10
```

```
rs182959636 CCNL2
```

Users can find a mapping file in SNP_Gene.mapping



Modularity Based Scoring Subnetwork (MoBaS)

Marzieh Ayati and Mehmet Koyuturk

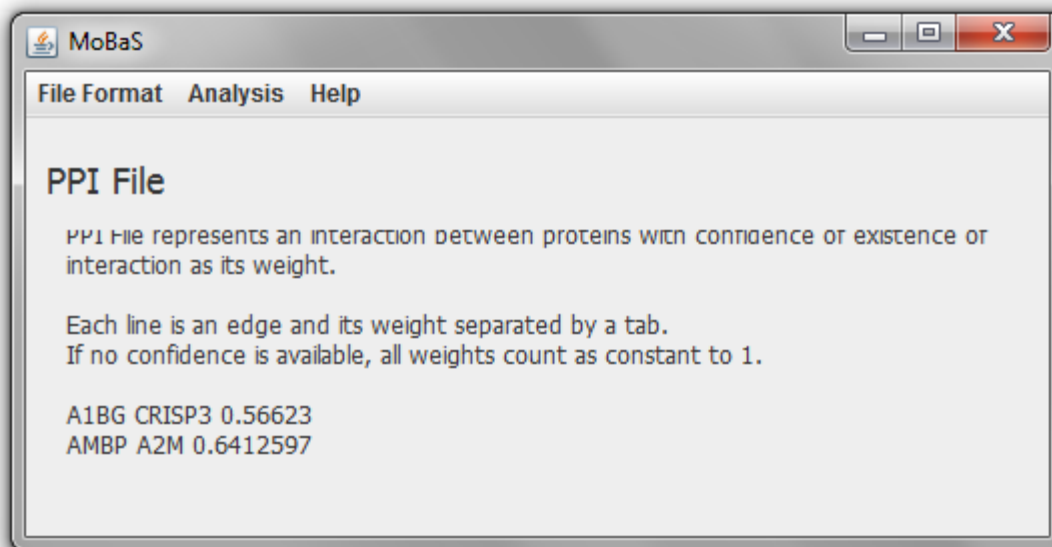
Protein-protein interaction network

PPI File represents an interaction between proteins with confidence of existence of interaction as its weight.

Each line is an edge and its weight separated by a tab.
If no confidence is available there, all weights count as constant to 1

```
A1BG CRISP3 0.56623  
AMBP A2M 0.6412597
```

ppi.txt is last HPRD PPI [2].



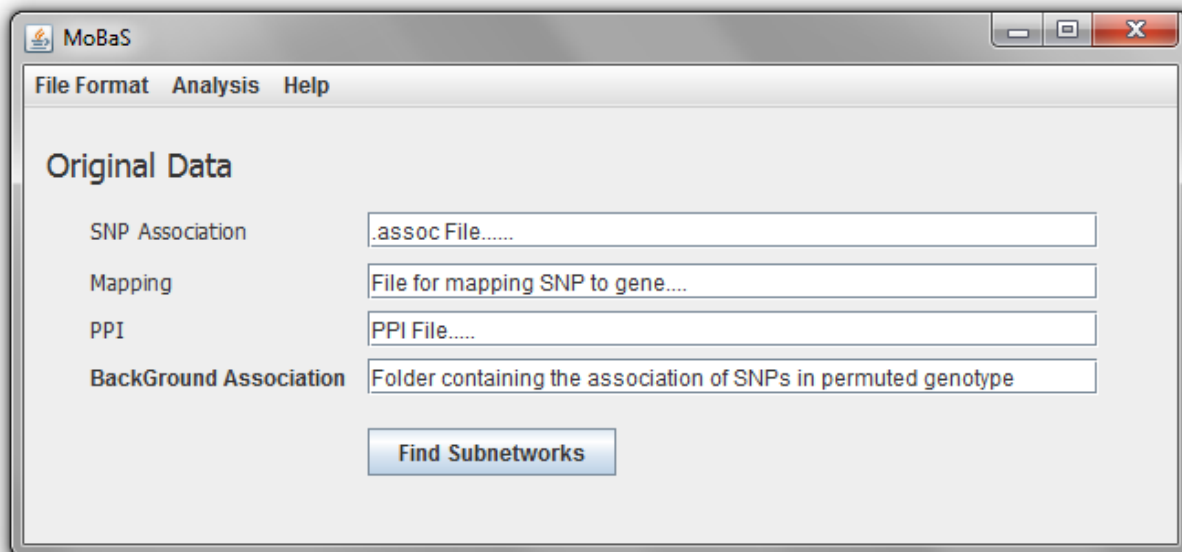
Modularity Based Scoring Subnetwork (MoBaS)

Marzieh Ayati and Mehmet Koyuturk

Identifying modules

After preparing the input files, you can run the MoBaS. However, the user has to also permute the genotypes and provide the mean of p-values of each SNP over n permutation (n=100 in our experiment).

MoBaS creates an output file that each line contains a score and the genes within each module.

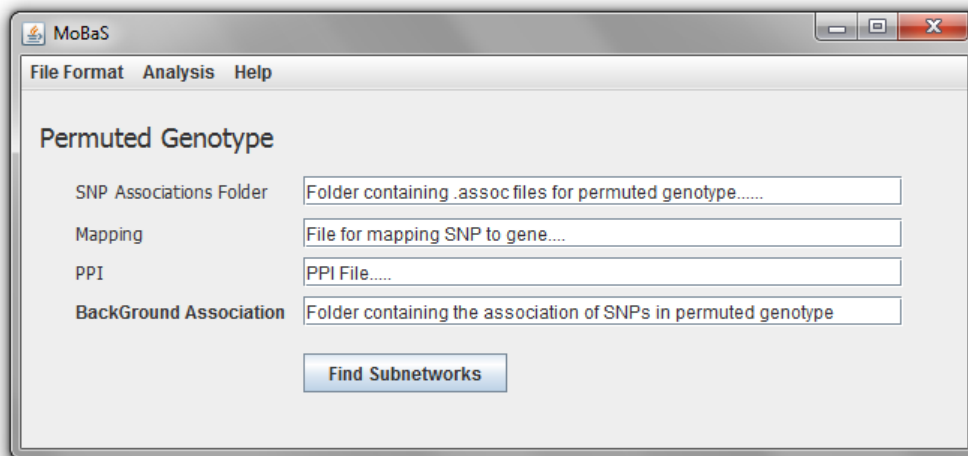


Modularity Based Scoring Subnetwork (MoBaS)

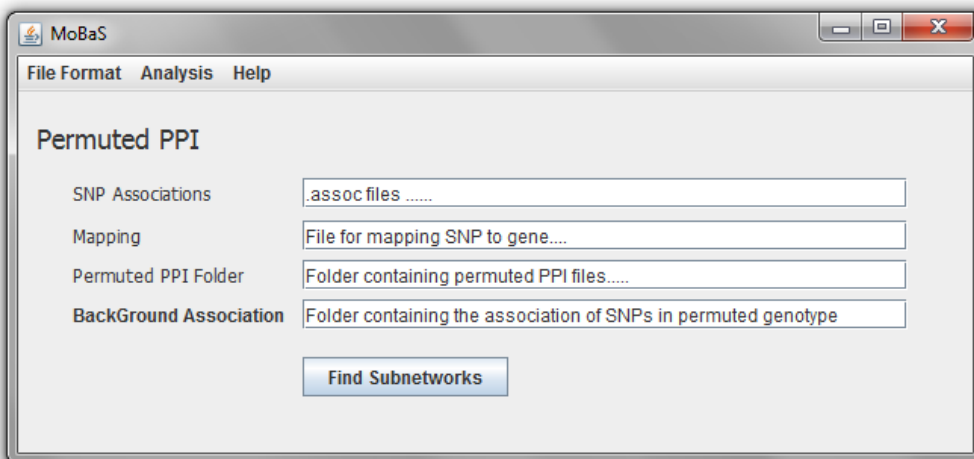
Marzieh Ayati and Mehmet Koyuturk

Statistical Analysis of Modules

If the users wants to compute the statistical analysis of identified modules and its q-value, they can provide permutation of genotype or PPI as an input and MoBaS compute the q-value of each identified modules in the original dataset.



The screenshot shows the MoBaS software interface for the "Permutated Genotype" analysis. The window title is "MoBaS" and the menu bar includes "File", "Format", "Analysis", and "Help". The main area is titled "Permutated Genotype" and contains four input fields with placeholder text: "SNP Associations Folder" (Folder containing .assoc files for permutated genotype.....), "Mapping" (File for mapping SNP to gene....), "PPI" (PPI File.....), and "BackGround Association" (Folder containing the association of SNPs in permutated genotype). A "Find Subnetworks" button is located at the bottom.



The screenshot shows the MoBaS software interface for the "Permutated PPI" analysis. The window title is "MoBaS" and the menu bar includes "File", "Format", "Analysis", and "Help". The main area is titled "Permutated PPI" and contains four input fields with placeholder text: "SNP Associations" (.assoc files), "Mapping" (File for mapping SNP to gene....), "Permutated PPI Folder" (Folder containing permutated PPI files.....), and "BackGround Association" (Folder containing the association of SNPs in permutated genotype). A "Find Subnetworks" button is located at the bottom.

Modularity Based Scoring Subnetwork (MoBaS)

Marzieh Ayati and Mehmet Koyuturk

Publications

-
- M. Ayati, S. Erten, Mark R. Chance, M. Koyuturk., "MOBAS: identification of disease-associated protein subnetworks using modularity-based scoring", *EURASIP Journal on Bioinformatics and Systems Biology*, 2015
- M. Ayati, S. Erten, M. Koyuturk. "What Do We learn from Network-based Analysis of GWAS?", In proceeding of EvoBio'14

References

[1] Purcell, Shaun, et al. "PLINK: a tool set for whole-genome association and population-based linkage analyses." *The American Journal of Human Genetics* 81.3 (2007): 559-575.

[2] Prasad, T. S. K. *et al.* (2009) Human Protein Reference Database - 2009 Update. *Nucleic Acids Research*. 37, D767-72