

A Guide to Interpreting the Results of MetaDAG Analysis

Table of contents

1 Introduction

As an illustrative example for interpreting [metaDAG](#) results, we consider here the Eukaryotes test presented in Section 2.5. Namely, we consider all Eukaryotes from the KEGG database. [Experiment github](#).

First of all, results must be downloaded from:

Hash: [0a845f74-826e-3b46-aed9-e7ecf74db262](#)

URL: (https://bioinfo.uib.es/metadag/handleExperiment/result_0a845f74-826e-3b46-aed9-e7ecf74db262)

and saved in the folder:

“data/0a845f74-826e-3b46-aed9-e7ecf74db262”.

```
library(tidyverse)
library(igraph)
library(ComplexHeatmap)
library(viridis)
library(circlize)
library(plotly)
library(randomcoloR)
library(factoextra)
library(RColorBrewer)
library(kableExtra)
library(igraph)
library(GGally)
```

Metadag Tool

Web Page

Experiment id: ff15c187-62e7-37c2-96a7-c824f7eab671
Experiment name: Eukaryotes
Experiment creator: Pere Palmer

Home
 Experiment
 Results

Experiment definition

Give the values needed to perform the desired experiment.

General data

Here you can select the files containing the calculations made for your experiment.

MBB	<input checked="" type="checkbox"/>	similarities	<input checked="" type="checkbox"/>
mDAG	<input checked="" type="checkbox"/>		
Results	<input checked="" type="checkbox"/>		
Select All	<input checked="" type="checkbox"/>		

MetaDAG files

Here, you can select what components have to be added to the downloaded data. The amount of options can be variable depending on the experiment and its configurations.

ACGTCT 110101
CGGC 1100
GACCGT 001011

Universitat
de les Illes Balears

Departament de Ciències Matemàtiques i Informàtica

Metadag Tool

Web Page

Experiment id: ff15c187-62e7-37c2-96a7-c824f7eab671
Experiment name: Eukaryotes
Experiment creator: Pere Palmer

Home
 Experiment
 Results

Experiment definition

Give the values needed to perform the desired experiment.

General data

Here you can select the files containing the calculations made for your experiment.

MetaDAG files


Here, you can select what components have to be added to the downloaded data. The amount of options can be variable depending on the experiment and its configurations.

	type		format		representation		Select row
	core	pan image	svg data	graphml data	raw mDAG	Essential Reaction Compound	
Global	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Categories	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Individuals	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Select column	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

ACGTCT 110101
CGGC 1100
GACCGT 001011

Universitat
de les Illes Balears

Departament de Ciències Matemàtiques i Informàtica






Metadag Tool




Web Page

Experiment id: ff15c187-62e7-37c2-96a7-c824f7eab671

Experiment name: Eukaryotes

Experiment creator: Pere Palmer

 Home
  Experiments
  Results

Viewing the results


Once an experiment has been calculated it's possible to view the results. To do this, it's necessary to give the experiment unique identifier supplied in the email sent when the experiment was submitted. This unique identifier is also reported in response to the submission.




In the email sent, an url is given to access directly to the experiment results page.

But here, you can also access to the experiment results with the unique identifier.

Enter the experiment identification code *

ff15c187-62e7-37c2-96a7-c824f7eab671



Universitat

de València

Departament

de Ciències Matemàtiques i Informàtica

2 Load data

```
experiment=
  "0a845f74-826e-3b46-aed9-e7ecf74db262/"
path_exp=paste0("data/",experiment)
knitr::kable(data.frame(
  Directory_files_and_folders=dir(path_exp),
  Type=c(rep("Data file",2),
        rep("Directory",3),
        rep("Data file",6),
        "Directory")))
```

Directory_files_and_folders	Type
Different_MBB.csv	Data file
Different_mDAG.csv	Data file
Global	Directory
Groups	Directory
Individuals	Directory
Report.pdf	Data file
Results.csv	Data file
Similarities_MBB_MSAMethod.csv	Data file
Similarities_MBB_MunkresMethod.csv	Data file
Similarities_mDAG_MSAMethod.csv	Data file
Similarities_mDAG_MunkresMethod.csv	Data file
TaxonomyLevels	Directory

```
MBB=read_csv(paste0(path_exp,"Different_MBB.csv"),
  show_col_types = FALSE)
mDAG=read_csv(paste0(path_exp,"Different_mDAG.csv"),
  show_col_types = FALSE)
```

2.1 Load metadata

Organisms are sorted by Kingdom, Phylum and Class:

```
path_exp
```

```
[1] "data/0a845f74-826e-3b46-aed9-e7ecf74db262/"
```

```
Results=read_csv(paste0(path_exp,"Results.csv"))
#rename MetaDaG variables
names(Results)[c(1,2,3,4,5)]=c("Organism","Categories","Groups","mDAG_Id","Full_Name")
taxo=Results %>% select(Organism:Full_Name)
meta_taxo=taxo %>% separate(Categories,into=c("Kingdom","Phylum","Class"))
index=which(is.na(meta_taxo$Class))
meta_taxo$Class[index]=paste(meta_taxo$Phylum[index])
rm(taxo)
aux=table(meta_taxo$Kingdom)
Freq_Kingdom=tibble(Kingdom=names(aux),Freq_Kingdom=aux)
aux=table(meta_taxo$Phylum)
Freq_Phylum=tibble(Phylum=names(aux),Freq_Phylum=aux)
aux=table(meta_taxo$Class)
Freq_Class=tibble(Class=names(aux),Freq_Class=aux)

meta_taxo = meta_taxo %>%
  left_join(Freq_Kingdom) %>%
  left_join(Freq_Phylum) %>%
  left_join(Freq_Class)
meta_taxo = meta_taxo %>%
  arrange(desc(Freq_Kingdom),
           desc(Freq_Phylum),
           desc(Freq_Class))

#arrange metaxo by frequencies of kingdom phylum and class

knitr::kable(head(meta_taxo))
```

Organism	Kingdom	Phylum	Class	Groups	mDAG_Id	Full_Name	Freq_Kingdom	Freq_Phylum	Freq_Class
aamp	Animalia	Vertebrata	Mammalia	Cluster 1	0313	Arvicola amphibius (Eurasian water vole)	535	331	139
afz	Animalia	Vertebrata	Mammalia	Cluster 1	0143	Antechinus flavipes (yellow-footed antechinus)	535	331	139
ajm	Animalia	Vertebrata	Mammalia	Cluster 1	0221	Artibeus jamaicensis (Jamaican fruit-eating bat)	535	331	139

Organism	Kingdom	Phylum	Class	Groups	mDAG	Full_Name	Freq_Kingdom	Freq_Phylum	Freq_Class
aju	Animalia	Vertebrata	Mammalia	Cluster0224	1	Acinonyx jubatus (cheetah)	535	331	139
aml	Animalia	Vertebrata	Mammalia	Cluster0279	1	Ailuropoda melanoleuca (giant panda)	535	331	139
anu	Animalia	Vertebrata	Mammalia	Cluster0310	1	Arvicanthis niloticus (African grass rat)	535	331	139

```
table(meta_taxo$Kingdom) %>% kable %>%
  kable_styling("striped", full_width = F, position="left") %>%
  scroll_box(width = "400px", height = "200px")
```

Var1	Freq
Animals	535
Fungi	154
Plants	139
Protists	56

```
table(meta_taxo$Phylum, meta_taxo$Kingdom) %>% kable %>%
  kable_styling("striped", full_width = F, position="left") %>%
  scroll_box(width = "500px", height = "500px")
```

	Animals	Fungi	Plants	Protists
Alveolates	0	0	0	25
Amoebozoa	0	0	0	7
Annelids	1	0	0	0
Arthropods	158	0	0	0
Ascomycetes	0	113	0	0
Basal	0	0	2	0
Basidiomycetes	0	36	0	0
Brachiopoda	1	0	0	0
Cephalochordates	2	0	0	0
Choanoflagellates	0	0	0	2
Cnidarians	10	0	0	0
Cryptomonads	0	0	0	1
Echinoderms	3	0	0	0
Eudicots	0	0	98	0
Euglenozoa	0	0	0	9
Ferns	0	0	1	0

Flatworms	4	0	0	0
Green	0	0	11	0
Haptophyta	0	0	0	1
Hemichordates	1	0	0	0
Heterolobosea	0	0	0	1
Metamonada	0	0	0	2
Microsporidians	0	5	0	0
Mollusks	14	0	0	0
Monocots	0	0	23	0
Mosses	0	0	1	0
Nematodes	6	0	0	0
Placozoans	1	0	0	0
Poriferans	1	0	0	0
Red	0	0	3	0
Stramenopiles	0	0	0	8
Tunicates	2	0	0	0
Vertebrates	331	0	0	0

2.2 Table of MBBs

In this example MBB is a table with 5149 rows and 4122 columns. It displays, for every MBB, the selected groups (Kingdoms, families, etc.) to which it belongs.

```
#100
knitr::kable(MBB[1:20,1:10]) %>%
  scroll_box(width = "100%", height = "200px")
```

MBB Id	natural	#pathways	Protists	Fungi	Plants	Animals	Alveolates	Amoebozoa	Annelids
0	0	0	0	0	0	0	0	0	0
0.0	0	0	0	0	0	0	0	0	0
0.0.0	0	0	0	0	0	0	0	0	0
0.0.0.0	1	1	0	0	0	1	0	0	0
0.0.0.0.0	1	1	0	0	0	1	0	0	0
0.0.0.0.0.0	1	1	0	0	0	1	0	0	0
0.0.0.1	1	1	0	0	0	1	0	0	0
0.0.0.2	1	1	0	0	0	1	0	0	0
0.0.1	0	0	0	0	0	0	0	0	0
0.0.1.0	1	1	0	0	0	1	0	0	1

0.0.1.1	1	1	1	0	0	0	0	0	0
0.0.1.1.0	1	1	1	0	0	0	0	0	0
0.0.1.2	1	1	0	0	0	1	0	0	0
0.0.1.3	1	1	0	0	0	1	0	0	0
0.0.1.4	1	1	0	0	0	1	0	0	0
0.0.1.4.0	1	1	0	0	0	1	0	0	0
0.0.1.4.0.0	1	1	0	0	0	1	0	0	0
0.0.1.5	1	1	1	0	0	0	0	0	0
0.0.1.6	1	4	3	0	0	1	0	0	0
0.0.1.6.0	1	3	0	0	3	0	0	0	0

2.3 Table of m-DAGs

In this example `mDAG` is a table with 1132 rows and 5278 columns. It displays, for every m-DAG, the selected groups (Kingdoms, families, etc.) in which it belongs.

```
kable(mDAG[1:20,1:10]) %>% scroll_box(width = "100%", height = "200px")
```

mDAG Id	#Categories	Animals	Plants	Fungi	Protists	Alveolates	Amoebozoa	Annelids	Arthro
0001	3	1	0	0	0	0	0	0	
0002	2	0	0	1	0	0	0	0	
0003	2	1	0	0	0	0	0	0	
0004	3	1	0	0	0	0	0	0	
0005	3	1	0	0	0	0	0	0	
0006	3	0	1	0	0	0	0	0	
0007	2	0	1	0	0	0	0	0	
0008	3	0	1	0	0	0	0	0	
0009	3	0	1	0	0	0	0	0	
0010	3	1	0	0	0	0	0	0	
0011	3	1	0	0	0	0	0	0	
0012	3	0	0	0	1	0	0	0	
0013	3	1	0	0	0	0	0	0	
0014	3	0	0	0	1	1	0	0	
0015	2	0	0	1	0	0	0	0	
0016	3	0	0	0	1	0	1	0	
0017	3	1	0	0	0	0	0	0	
0018	3	1	0	0	0	0	0	0	
0019	3	1	0	0	0	0	0	0	

0020	3	1	0	0	0	0	0	0
------	---	---	---	---	---	---	---	---

```
dim(mDAG)
```

```
[1] 1132 5278
```

```
names(mDAG)[1:6]
```

```
[1] "mDAG Id"      "#Categories" "Animals"      "Plants"      "Fungi"
[6] "Protists"
```

```
head(names(mDAG)[7:(dim(mDAG)[2]-1150)])
```

```
[1] "Alveolates"      "Amoebozoa"      "Annelids"
[4] "Arthropods"      "Ascomycetes"     "Basal angiosperms"
```

```
# 28 to 1213 code MBB: 1 if MBB in mDAG 0
```

2.4 Results Table

The **Results** table contains for every organism (row) the following information: its category (taxonomy), selected group, Full name, m-DAG id and all reactions name id with their corresponding enzyme. When a reaction is present in the corresponding m-DAG, the MBB to which it belongs is displayed in this column.

```
kable(Results[1:20,1:10])%>%
  row_spec(0, angle = 0) %>%
  scroll_box(width = "300%", height = "1000px")
```

Organism	Categories	Groups
aaf	Protists|Stramenopiles|Pelagophytes	MSA Cluster 3|MUN Cluster 3
aag	Animals|Arthropods|Insects	MSA Cluster 2|MUN Cluster 2
aalb	Animals|Arthropods|Insects	MSA Cluster 2|MUN Cluster 2
aali	Animals|Arthropods|Insects	MSA Cluster 2|MUN Cluster 2
aalt	Fungi|Ascomycetes|Dothideomycetes	Fungui and Algae|MSA Cluster 3|MUN Cluster 3
aam	Animals|Vertebrates|Birds	Cluster 1
aamp	Animals|Vertebrates|Mammals	Cluster 1
aang	Animals|Vertebrates|Fishes	Cluster 1
aara	Animals|Arthropods|Insects	MSA Cluster 2|MUN Cluster 2

abe	Fungi|Ascomycetes|Eurotiomycetes	Fungui and Algae|MSA Cluster 3&#
abp	Fungi|Basidiomycetes	Fungui and Algae|MSA Cluster 3&#
abv	Fungi|Basidiomycetes	Fungui and Algae|MSA Cluster 3&#
acan	Protists|Amoebozoa|Acanthamoeba	MSA Cluster 3|MUN Cluster 3
acar	Animals|Vertebrates|Birds	Cluster 1
acep	Animals|Arthropods|Insects	MSA Cluster 2|MUN Cluster 2
acer	Animals|Arthropods|Insects	MSA Cluster 2|MUN Cluster 2
achc	Animals|Vertebrates|Birds	Cluster 1
ache	Fungi|Ascomycetes|Eurotiomycetes	Fungui and Algae|MSA Cluster 3&#
achl	Animals|Vertebrates|Birds	Cluster 1
acoz	Animals|Arthropods|Insects	MSA Cluster 2|MUN Cluster 2

```
dim(Results)
```

```
[1] 1132 3998
```

```
names(Results)[1]# organisms kegg id class representant of mDAG
```

```
[1] "Organism"
```

```
names(Results)[2]# taxonomy separate by /
```

```
[1] "Categories"
```

```
names(Results)[3]# groups
```

```
[1] "Groups"
```

```
names(Results)[4]# mDAG_Id
```

```
[1] "mDAG_Id"
```

```
names(Results)[5]# Full name representant
```

```
[1] "Full_Name"
```

```
names(Results)[6:36]# columns 6 to 3998
```

```

[1] "R00005(3.5.1.54)" "R00009(1.11.1.6)" "R00010(3.2.1.28)"
[4] "R00014(1.2.4.1)" "R00014(4.1.1.1)" "R00021(1.4.7.1)"
[7] "R00022(3.2.1.52)" "R00024(4.1.1.39)" "R00028(3.2.1.20)"
[10] "R00031(1.10.3.1)" "R00032(1.13.11.63)" "R00036(4.2.1.24)"
[13] "R00045(1.10.3.1)" "R00066(2.5.1.9)" "R00068(1.10.3.3)"
[16] "R00073(1.14.99.1)" "R00075(2.5.1.43)" "R00078(1.16.3.1)"
[19] "R00084(2.5.1.61)" "R00086(3.6.1.15)" "R00086(3.6.1.5)"
[22] "R00087(3.6.1.9)" "R00093(1.4.1.14)" "R00095(1.6.5.4)"
[25] "R00100(1.6.2.2)" "R00102(3.2.2.5)" "R00102(3.2.2.6)"
[28] "R00103(3.6.1.22)" "R00103(3.6.1.9)" "R00104(2.7.1.23)"
[31] "R00111(1.14.13.39)"

```

reactions name id with its enzyme.

```

reactions=names(Results)[-c(1:5)]
reverse_reactions=stringr::str_detect(reactions,"rev")
reverse_reactions=table(reverse_reactions)
dimnames(reverse_reactions)$reverse_reactions=
  c("Non reverse reaction","Reverse reaction")
reverse_reactions %>% kable %>%
  kable_styling("striped", full_width = F, position="left")

```

reverse_reactions	Freq
Non reverse reaction	3399
Reverse reaction	594

3 Metabolic Graphs

```
gc()
```

	used	(Mb)	gc trigger	(Mb)	max used	(Mb)
Ncells	2771214	148.0	4247817	226.9	4247817	226.9
Vcells	4571320	34.9	10146329	77.5	8012275	61.2

```
load(file='metadag_work_space.RData')
gc()
```

	used	(Mb)	gc trigger	(Mb)	max used	(Mb)
Ncells	2870799	153.4	4247817	226.9	4247817	226.9
Vcells	38153618	291.1	49085484	374.5	38174318	291.3

We present here some analysis examples of the metabolic graphs generated in GraphML format.

3.1 Metabolic graphs for each organism

Read the individual metabolic graphs generated for Homo sapiens (KEGG id: hsa) in the directory(Individuals/hsa)

```
experiment=
  "0a845f74-826e-3b46-aed9-e7ecf74db262/"
path_exp=paste0("data/",experiment)
files_hsa=dir(paste0(path_exp,"Individuals/hsa"))
files_hsa
```

[1] "hsa_mDAG.graphml"	"hsa_mDAG.pdf"	"hsa_mDAG.svg"
[4] "hsa_mDAG_adj.csv"	"hsa_mDAG_biggerDAG.pdf"	"hsa_mDAG_biggerDAG.svg"
[7] "hsa_mDAG_n1.csv"	"hsa_mDAG_structure.csv"	"hsa_R_adj.csv"
[10] "hsa_R_n1.csv"	"hsa_RC.graphml"	"hsa_RC.pdf"
[13] "hsa_RC.svg"	"hsa_summary.txt"	

files_Individual_hsa	Description
hsa_mDAG.graphml	m-DAG GraphML format
hsa_mDAG.pdf	m-DAG pdf graphic
hsa_mDAG.svg	m-DAG svg graphic
hsa_mDAG_adj.csv	csv file with the adjacency matrix of the m-DAG
hsa_mDAG_biggerDAG.pdf	pdf graphic with the biggest conected componet of the m-DAG
hsa_mDAG_biggerDAG.svg	svg graphic with the biggest conected componet of the m-DAG
hsa_mDAG_nl.csv	csv file with the node (MBBs) labels of the m-DAG
hsa_mDAG_structures.csv	csv file with all connected components of the m-DAG
hsa_R_adj.csv	csv file with the adjacency matrix of the reaction graph
hsa_R_nl.csv	csv file with the node (reactions) labels of the reaction graph
hsa_RC.graphml	reaction graph GraphML format
hsa_RC.pdf	reaction graph pdf graphic
hsa_RC.svg	reaction graph svg graphic
hsa_summary.txt	text summary file with the number of MBBs, reactions, etc. in the previous graphs

3.2 Pan & core metabolic graphs

Pan and core metabolic graphs for every group were generated. For instance, one can read the pan and core metabolic graphs generated for the group Algae in the directory (Groups/Algae).

```
files_Algae=dir(paste0(path_exp,"Groups/Algae"))
files_Algae
```

```
[1] "core" "pan"
```

The global core reaction graph, which is the core of all the organisms' reaction graphs in this Eukaryotes test, is empty.

```
graph_core_RC=read.graph(
  paste0(path_exp,
    "Global/core/core_RC.graphml"),
  format = "graphml")
summary(graph_core_RC)
```

```
IGRAPH 1cff44a D--- 0 0 --
+ attr: color (v/c), label (v/c), id (v/c)
```