

M2 Internship

Comparative analysis of genome-scale
metabolic networks among brown algae

Pauline Hamon-Giraud

Supervisors : Anne Siegel, Jeanne Got & Gabriel Markov



CNRS UPMC
Station Biologique
Roscoff



Summary

Objectives

1. Generate metabolic networks
 - a. AuCoMe tool
 - b. Generated Networks
 - c. Results
 - d. Accessibility of results
 - e. Perspectives
2. Study the evolution of gene loss in *Laminarionema elsbetiae*
 - a. Finding the lost genes
 - b. Results
 - c. Interpretation
 - d. Perspectives
3. Interactions between the endophyte *L. elsbetiae* and its host *S. latissima*

Objectives

1. Generate metabolic networks of brown algae from Phaeoexplorer data
2. Study the evolution of gene loss in *Laminarionema elsbetiae* (endophytic lifestyle)
3. Identify interactions of metabolic functions between the endophyte *L. elsbetiae* and its host *Saccharina latissima*

1. Generate metabolic networks

Why ?

- Identify the metabolic reactions / pathways of species
- Compare metabolic networks between species allow to reconstruct trees based on metabolic capability

How ?

- From annotated genome files : Phoexplorer + Public
- Thanks to the AuCoMe metabolic networks reconstruction tool

a. AuCoMe tool

Objective : Automatically reconstruct metabolic networks and homogenize the annotations of the input files

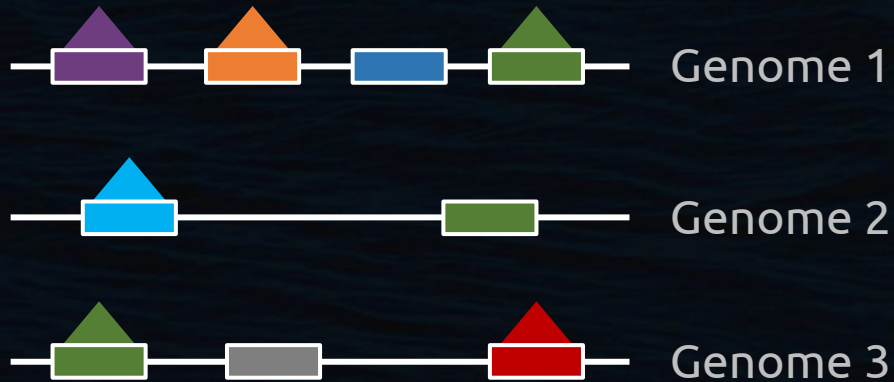
—> overcome the heterogeneity of the annotations according to the organisms

Main steps :

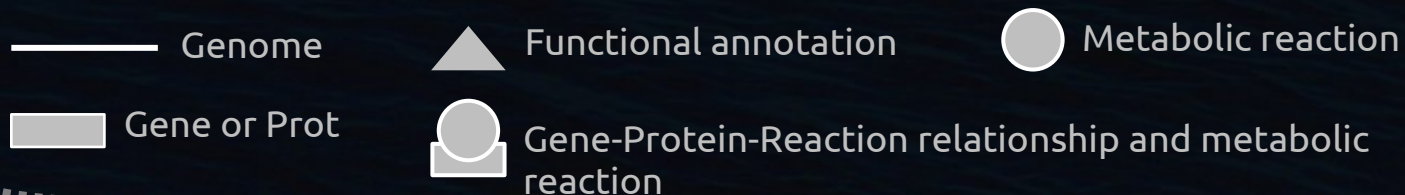
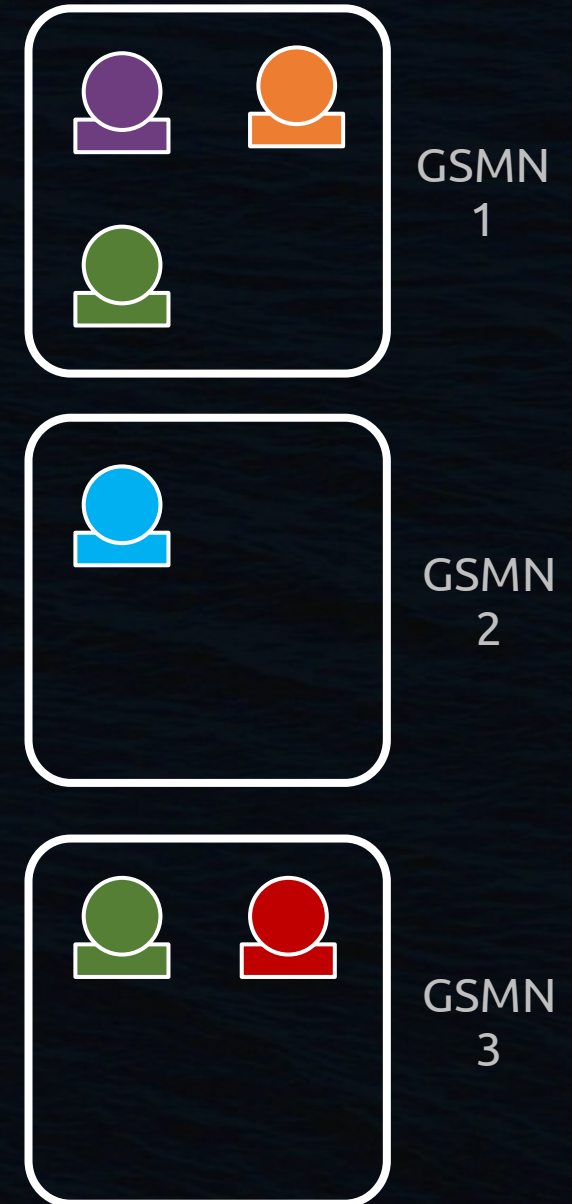
1. Reconstruction of networks from file annotations (PathwayTools)
2. Gene propagation by orthology (Orthofinder)
3. Search for potential missed annotations during initial annotation of files
4. Adding spontaneous reactions to networks
5. Processing the results : clustering (Pvclust), tables

a. AuCoMe tool

1. Draft reconstruction

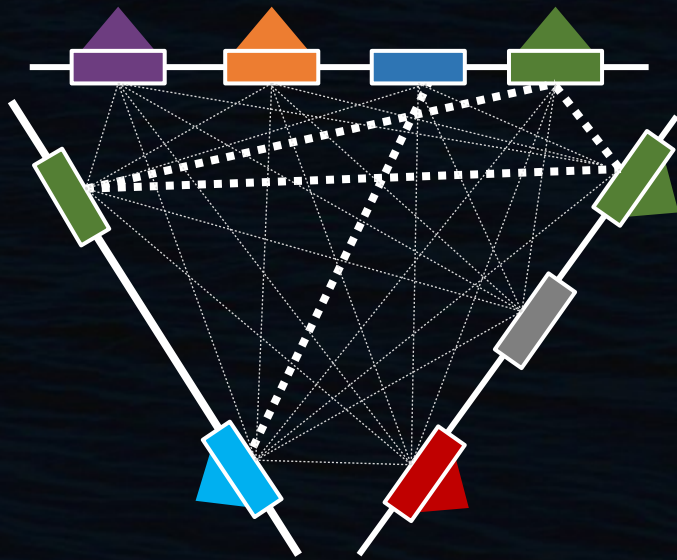


GSMN = Genome-Scale
Metabolic Networks



a. AuCoMe tool

2. Orthology propagation



OrthoFinder

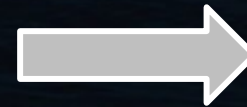
Orthogroups (OG)



Robust OG →
GPR propagation



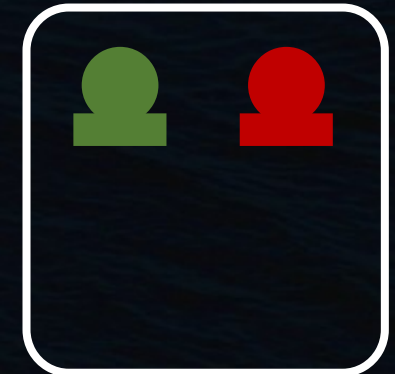
Non-robust OG →
no GPR propagation



GSMN
1



GSMN
2



GSMN
3

— Genome



Functional annotation



Metabolic reaction

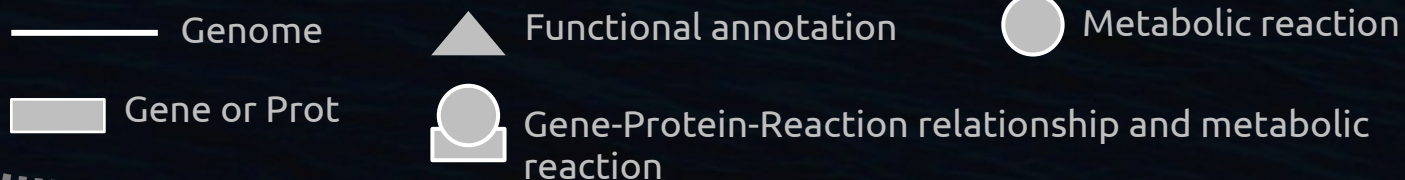
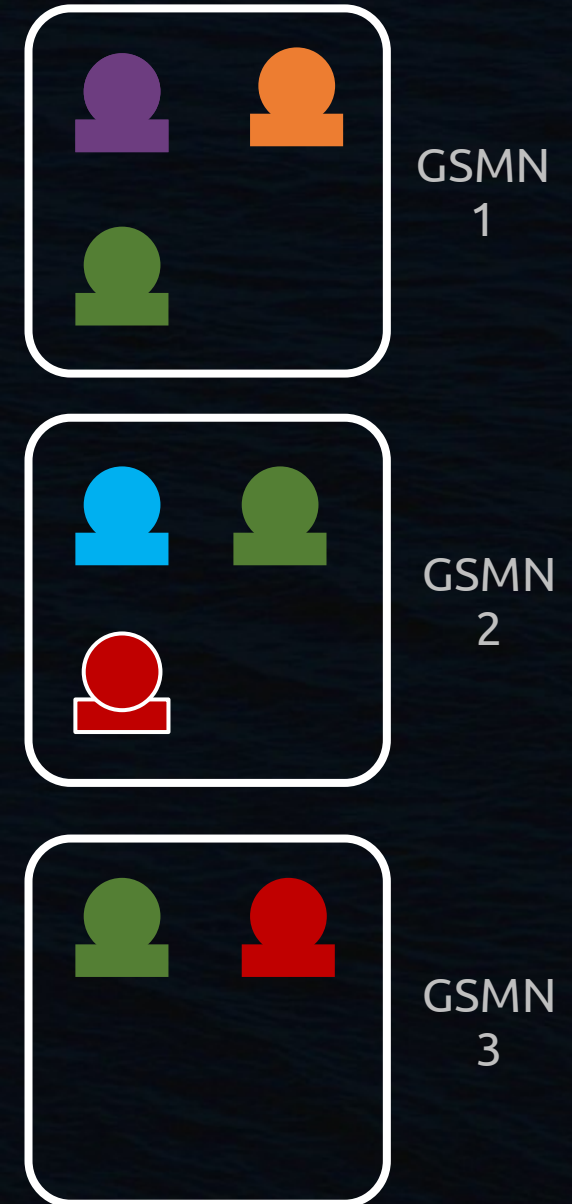
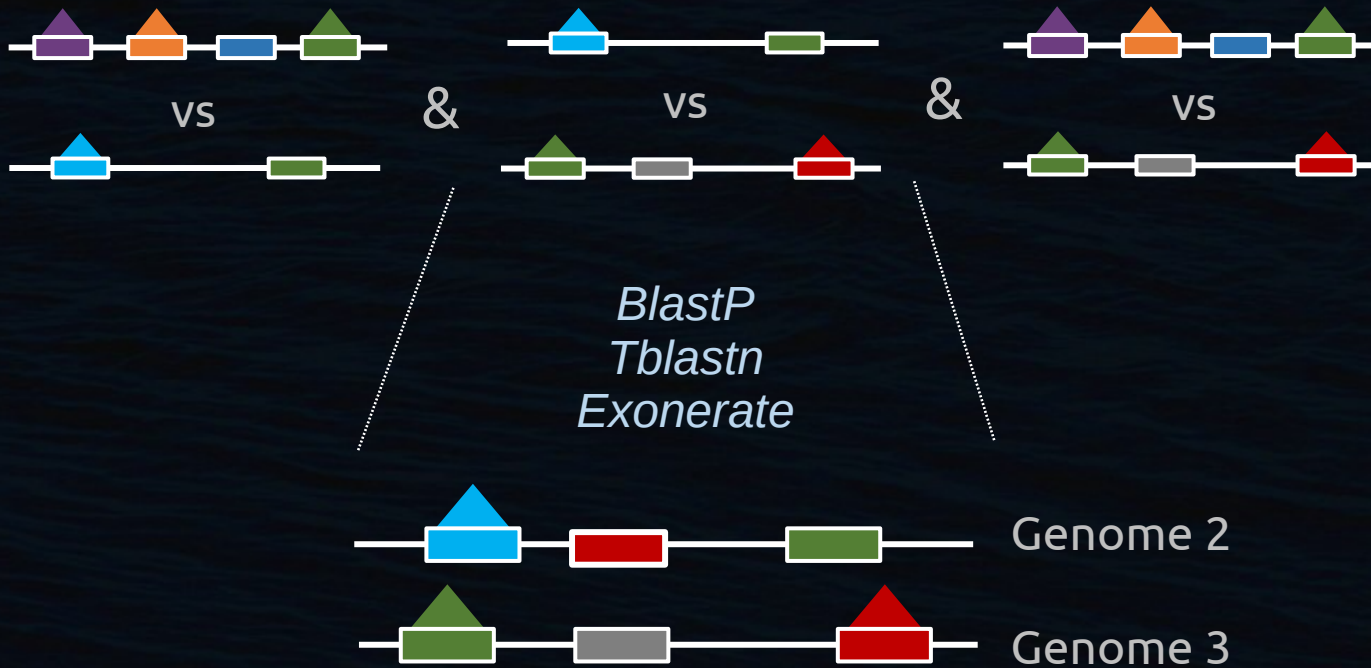
■ Gene or Prot



Gene-Protein-Reaction relationship and metabolic
reaction

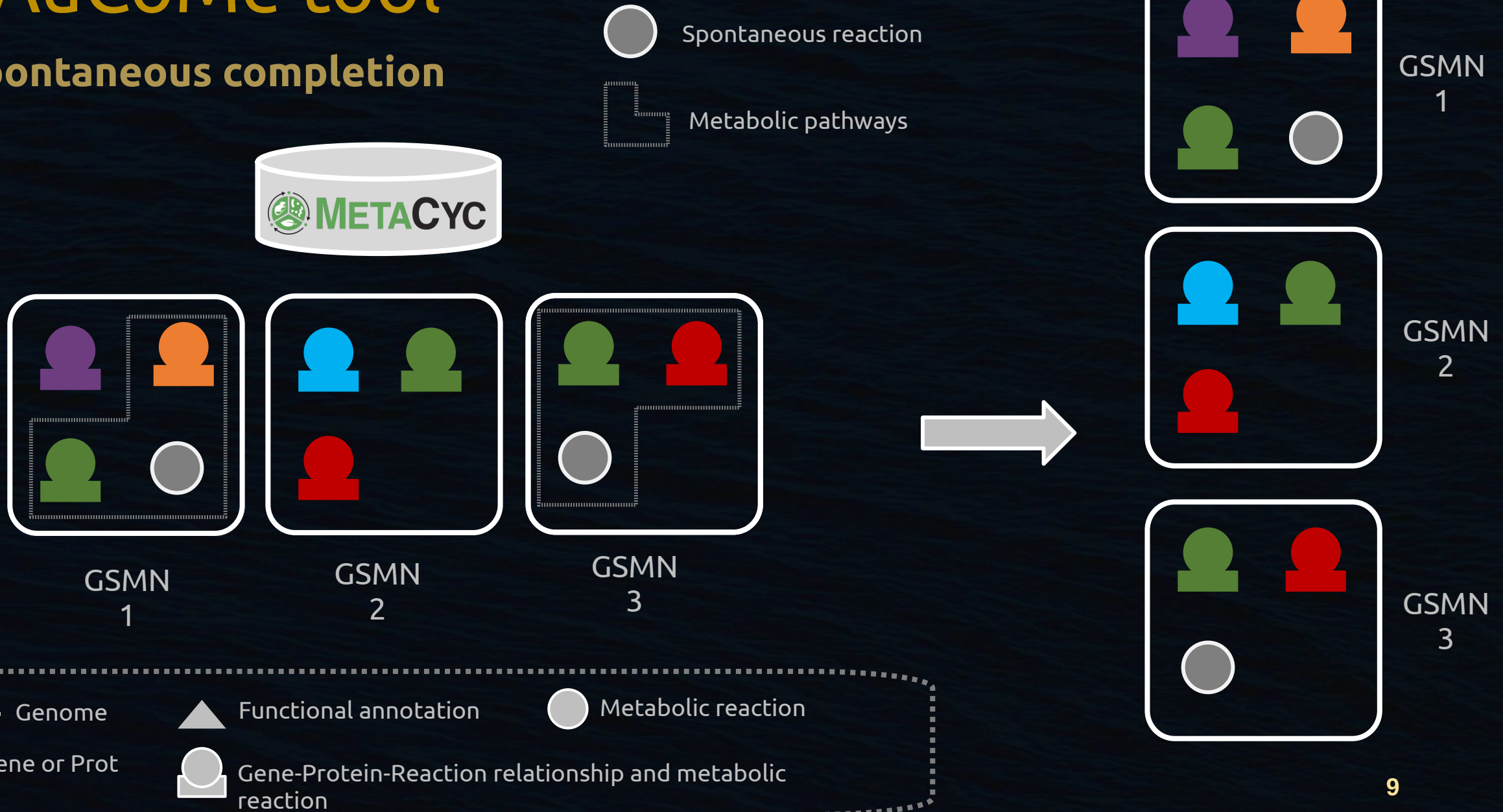
a. AuCoMe tool

3. Structural verification



a. AuCoMe tool

4. Spontaneous completion



b. Generated Networks

- Networks generated for a dataset including long read assemblies + Laminarionema.E : Dataset 01

Phaeoexplorer
Public
● Long read
Diatoms
Brown algae

18 Brown algae :

Public :

- *Undaria pinnatifida* (Kr2015)
- *Cladosiphon okamuranus*
- *Ectocarpus species 7*
- *Ectocarpus subulatus*
- *Nemacystus decipiens*
- *Saccharina japonica*

Phaeoexplorer :

- *Ectocarpus crouaniorum* (m) ●
- *Laminarionema elsbetiae*
- *Desmarestia herbacea* (m) ●
- *Porterinema fluviatile* ●
- *Scytosiphon promiscuus* (m) ●
- *Ectocarpus siliciosus* (m) ●
- *Chordaria linearis* ●
- *Pleurocladia lacustris* ●
- *Ectocarpus fasciculatus* (m) ●
- *Dictyota dichotoma* (m) ●
- *Fucus serratus* (m) ●
- *Saccharina latissima* (f) ●

6 Outgroups :

Public :

- *Fistulifera solaris*
- *Fragilariopsis cylindrus*
- *Phaeodactylum tricornutum*
- *Thalassiosira pseudonana*
- *Nannochloropsis gaditana*

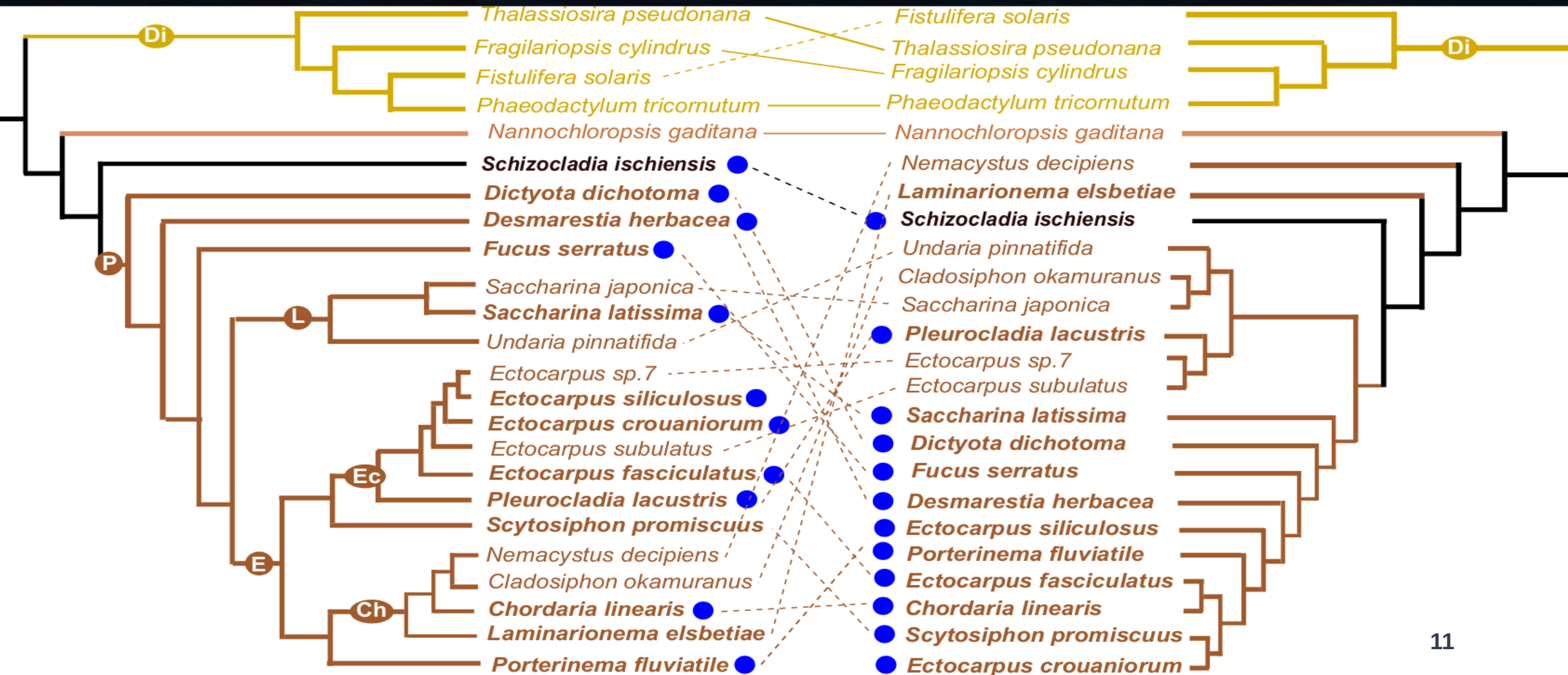
Phaeoexplorer

- *Schizocladia ischiensis* ●

c. Results : Dataset 01

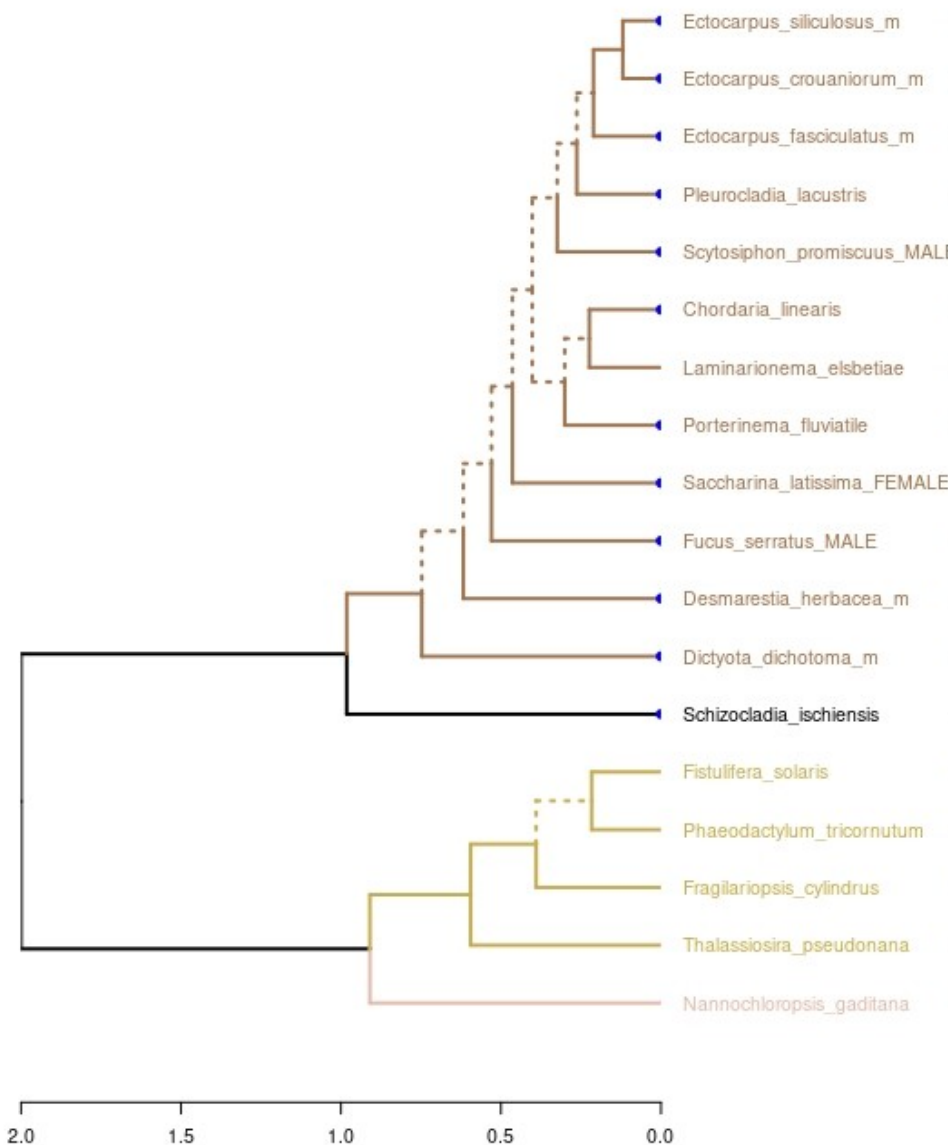
Original phylogeny

Results Aucome dataset 01

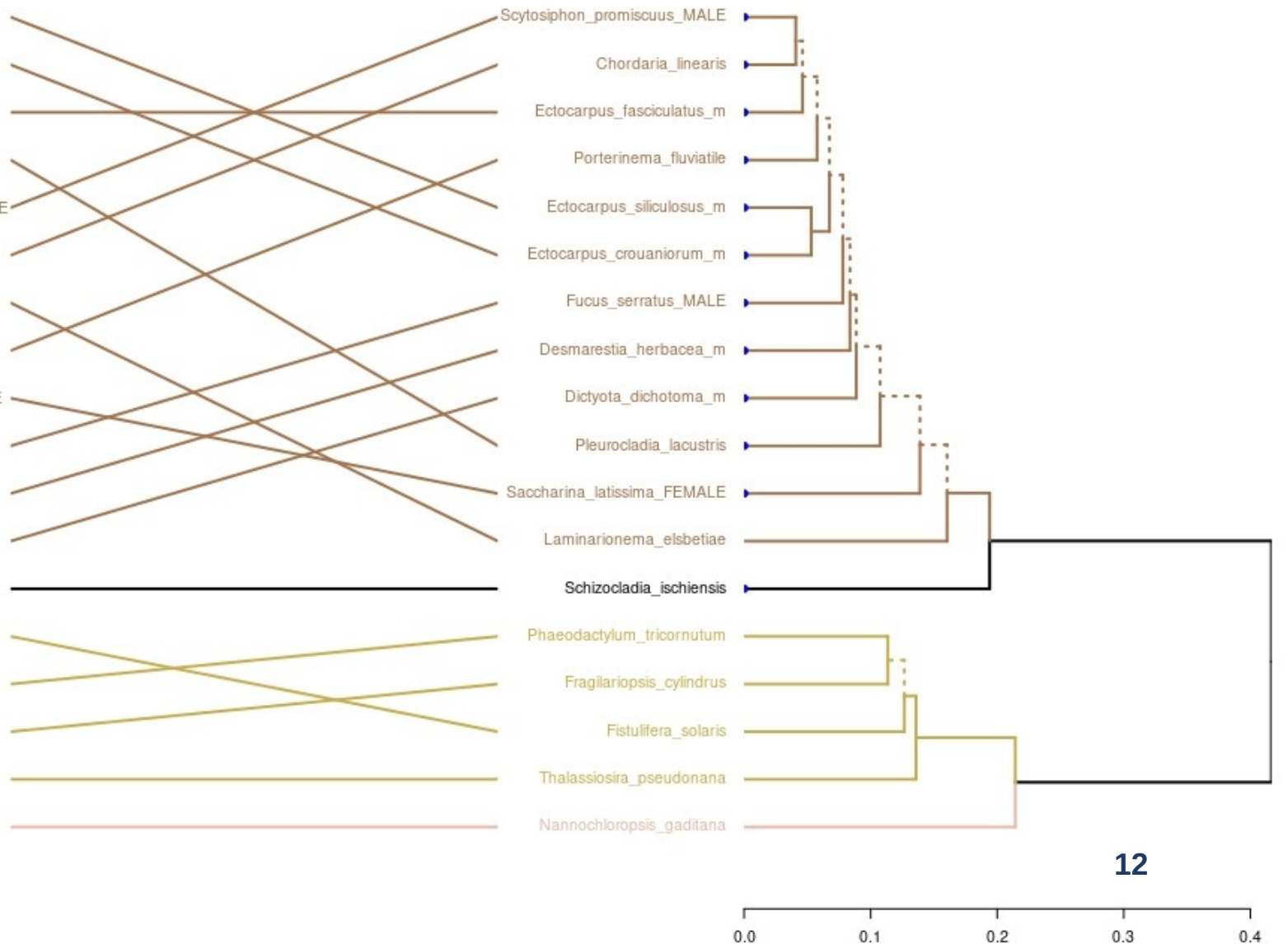


c. Results : Dataset 03

Original phylogeny



Run 03
Metabolic Dendrogram



c. Results : Dataset 03 -> Interpretation

How to get closer to the phylogeny of species ?

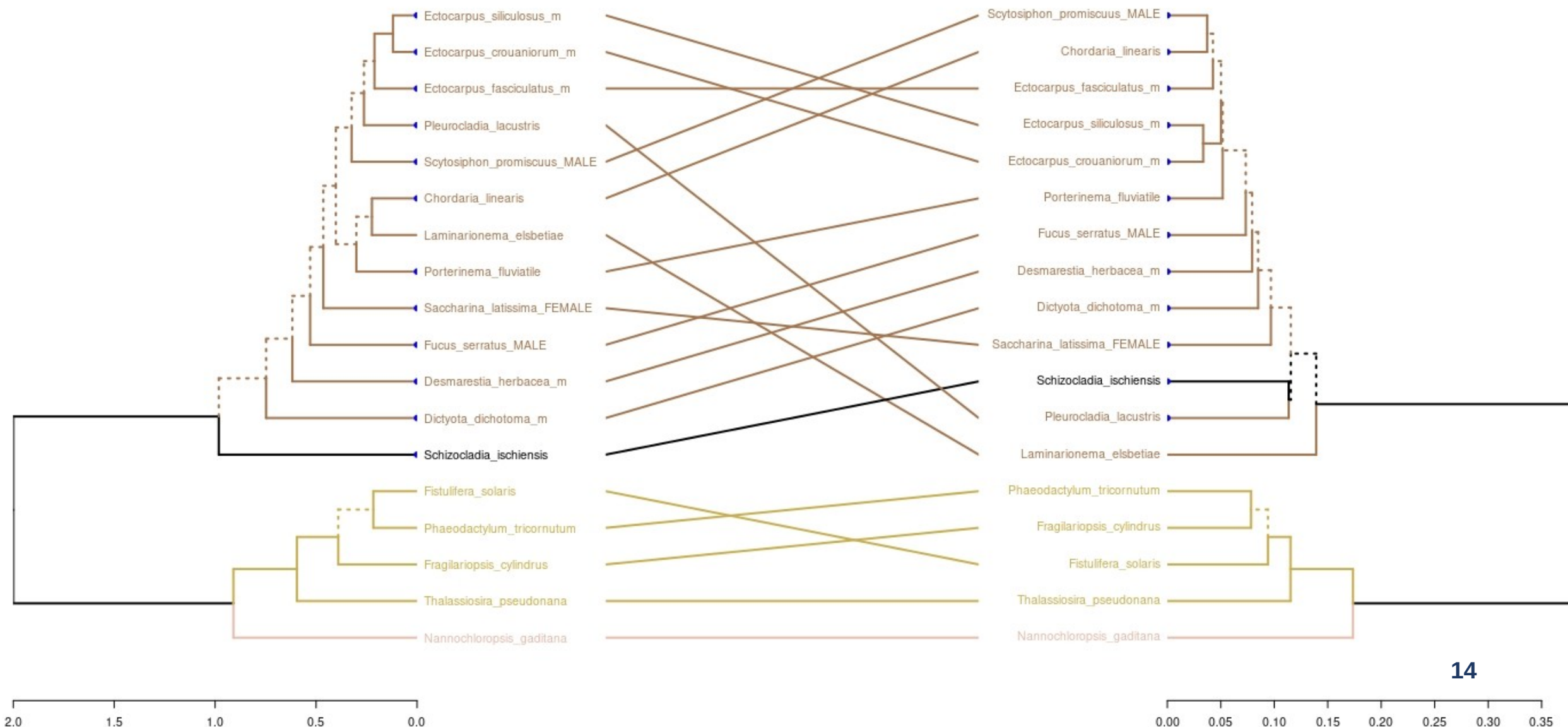
Filter the reactions on brown algae -> remove the ones that seem questionable

- **Idea 1** : Remove reactions present in only 2-3 brown algae unless they belong to basal species of the phylogeny
- **Idea 2** : Remove the reactions present in only 1-2 brown algae if they belong to the wrong species (too much at the base)

c. Results : Dataset 03 -> Idea 1

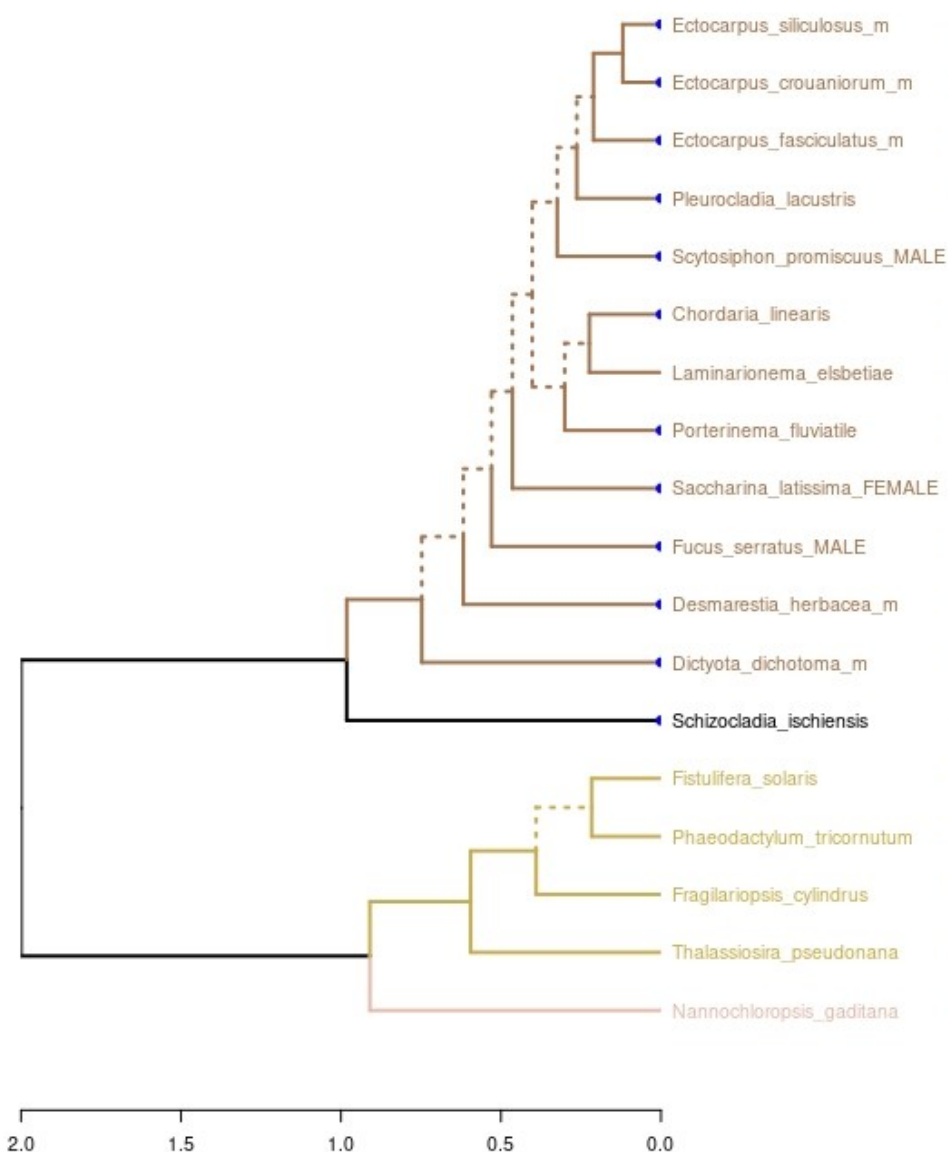
Original phylogeny

Run 03 : cut 3 basal
Metabolic Dendrogram

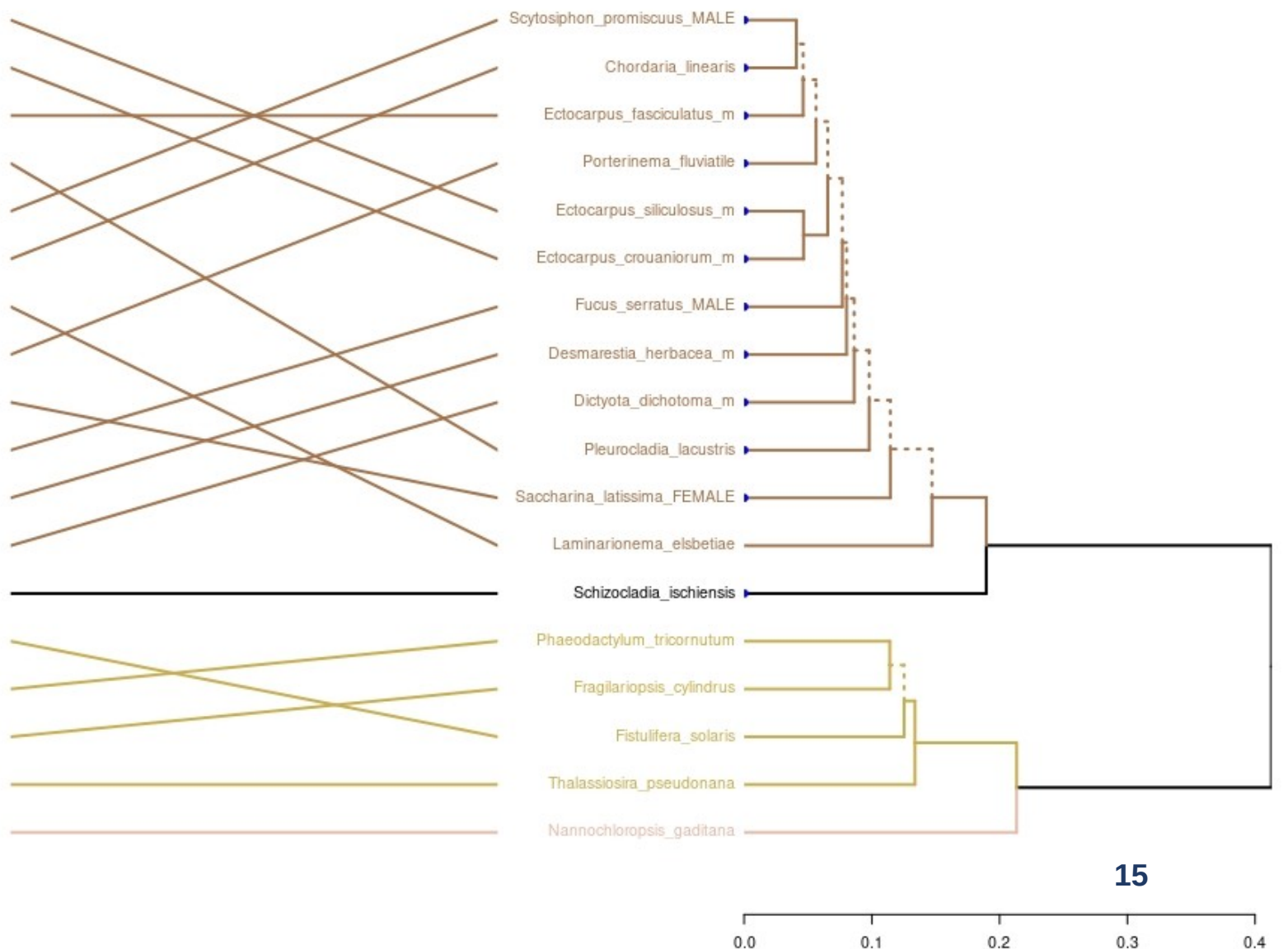


c. Results : Dataset 03 -> Idea 2

Original phylogeny



Run 03 : cut 2 misplaced Metabolic Dendrogram



c. Results : Dataset 03 -> #Reations



d. Accessibility of results

- Data storage in wikis
 - Availability of file results in an easy human-readable online page
 - For each organism, list of genes, reactions, pathways and metabolites
 - Information on the origin of the integration of reactions : orthology or annotation



Main Page

description [\[edit\]](#)

Automatic reconstruction with [AuReMe](#) [\[edit\]](#)

Model summary: [summary](#)

Download **AuReMe** Input/Output [LINK OR MEDIA data]

The automatic reconstruction of *MODEL_NAME* results to a Genome scale **Model** containing 3230 reactions, 3610 metabolites, 6239 genes and 1483 pathways. This GeM was obtained based on the following sources:

- Based on orthology data:
 - Creation of a global metabolic network containing 2932 reactions
- Based on annotation data:
 - Creation of a metabolic network containing 2306 reactions
- Based on expertise:
 - 6 reaction(s) added
- List of tools used:
 - Tool: [Orthofinder](#)
 - Tool: [PathwayTools](#)

File:Venn.png

Collaborative curation [\[edit\]](#)

- Suggest reactions to add or remove:
 - Download this [form](#)
- Suggest new reactions to create and add:
 - Download this [form](#)
- **Follow the examples given in the form(s) to correctly share your suggestions**
- Send the filled form(s) to: CONTACT_MAIL

Expert input : possibility to suggest modifications of reactions

Main page

[SIMV-ASK](#)
[workflow command history](#)
[Random page](#)
[Files](#)

[Metabolic network components](#)

[organism](#)
[reaction](#)
[gene](#)
[pathway](#)
[metabolite](#)

[Reconstruction categories](#)

[annotation manual](#)
[orthology](#)

[Reconstruction tools](#)

[curation](#)
[orthofinder](#)
[pathwaytools](#)

[Tools](#)

[What links here](#)
[Related changes](#)
[Special pages](#)
[Printable version](#)
[Permanent link](#)
[Page information](#)
[Browse properties](#)

Category [Discussion](#)[Read](#) [Edit](#) [View history](#)

Category:Organism

[?](#) [Help](#)

	↕ Nb reaction associated ↕	Nb gene associated ↕	Nb pathway associated ↕
Ectocarpus siliculosus m	3362	8307	1506
Schizocladia ischiensis	3331	8107	1533
Scytosiphon promiscuus MALE	3323	8261	1515
Ectocarpus crouaniorum m	3321	8249	1508
Chordaria linearis	3306	7940	1511
Porterinema fluviatile	3299	7593	1513
Ectocarpus fasciculatus m	3296	8555	1499
Desmarestia herbacea m	3274	7869	1507
Fucus serratus MALE	3271	9622	1490
Saccharina latissima FEMALE	3249	7991	1504
Nemacystus decipiens	3240	5462	1520
Pleurocladia lacustris	3236	7515	1502
Saccharina japonica	3230	6239	1483
Dictyota dichotoma m	3225	8089	1495
Cladosiphon okamuranus	3194	5565	1473
Laminarionema elsbetiae	3172	7543	1490
Undaria pinnatifida Kr2015	3152	6186	1466
Ectocarpus subulatus	3144	6900	1470
Fragilariopsis cylindrus	3076	3692	1472
Fistulifera solaris	2974	5882	1433
Phaeodactylum tricornutum	2956	3164	1463
Nannochloropsis gaditana	2953	3337	1461
Thalassiosira pseudonana	2924	3383	1450

Overview for each organism :
number of genes / reactions /
pathways

This category currently contains no pages or media.

[Main page](#)
[SMW-Ask](#)
[workflow command](#)
[history](#)
[Random page](#)
[Files](#)

[Metabolic network](#)
[components](#)

[organism](#)
[reaction](#)
[gene](#)
[pathway](#)
[metabolite](#)

[Reconstruction](#)
[categories](#)

[annotation](#)
[manual](#)
[orthology](#)

[Reconstruction tools](#)

[curation](#)
[orthofinder](#)
[pathwaytools](#)

[Tools](#)

[What links here](#)
[Related changes](#)
[Special pages](#)
[Printable version](#)
[Permanent link](#)
[Page information](#)
[Browse properties](#)



Main page
SMW-Ask
workflow command
history
Random page
Files

Metabolic network
components

organism
reaction
gene
pathway
metabolite

Reconstruction
categories

annotation
manual
orthology

Reconstruction tools

curation
orthofinder
pathwaytools

Tools

What links here
Related changes
Special pages
Printable version
Permanent link
Page information
Browse properties

Page

Discussion

Read

Edit

View history

Search metabolic_network



Schizocladia ischiensis

Contents [hide]

- 1 Organism Schizocladia_ischiensis
- 2 Summary
- 3 Reaction(s) associated
- 4 Gene(s) associated
- 5 Pathway(s) associated

Details for a specific organism :
list of genes / reactions /
pathways

Organism Schizocladia_ischiensis [edit]

Summary [edit]

- 3331 reaction(s)
- 8107 gene(s)
- 1533 pathway(s)

Reaction(s) associated [edit]

[Expand]

Gene(s) associated [edit]

[Expand]

Pathway(s) associated [edit]

[Expand]

Category: Organism



Ex : Reactions informations

Category

[Discussion](#)[Read](#)[Edit](#)[View history](#)

Category:Reaction

[Help](#)[Main page](#)
[SMW-Ask](#)
[workflow command history](#)
[Random page](#)
[Files](#)[Metabolic network components](#)[organism](#)
[reaction](#)
[gene](#)
[pathway](#)
[metabolite](#)[Reconstruction categories](#)[annotation](#)
[manual](#)
[orthology](#)[Reconstruction tools](#)[curation](#)
[orthofinder](#)
[pathwaytools](#)[Tools](#)[What links here](#)
[Related changes](#)
[Special pages](#)
[Printable version](#)
[Permanent link](#)
[Page information](#)
[Browse properties](#)

	Common-name	Ec-number	Nb gene associated	Nb pathway associated	Nb reconstruction source	Reconstruction category	Reconstruction tool	Reconstruction comment
+-NEOMENTHOL-DEHYDROGENASE-RXN		Ec-1.1.1.208	1	1		Orthology	Orthofinder	N.a
1-ACYLGLYCEROL-3-P-ACYLTRANSFER-RXN	Nil 1-acylglycerol-3-phosphate o-acyltransferase	Ec-2.3.1.51	10	2		Orthology Annotation	Orthofinder Pathwaytools	N.a
1-PHOSPHATIDYLINOSITOL-3-KINASE-RXN	Nil 1-phosphatidylinositol-3-kinase	Ec-2.7.1.137	20	1		Orthology Annotation	Orthofinder Pathwaytools	N.a
1-PHOSPHATIDYLINOSITOL-KINASE-RXN	Nil 1-phosphatidylinositol 4-kinase	Ec-2.7.1.67	15	2		Orthology Annotation	Orthofinder Pathwaytools	N.a
1.1.1.145-RXN		Ec-1.1.1.145	2	0		Orthology	Orthofinder	N.a
1.1.1.178-RXN	Nil	Ec-1.1.1.178	1	2		Orthology Annotation	Orthofinder Pathwaytools	N.a
1.1.1.188-RXN	Prostaglandin d2 11-ketoreductase	Ec-1.1.1.188	1	0		Orthology Annotation	Orthofinder Pathwaytools	N.a
1.1.1.197-RXN		Ec-1.1.1.197	4	0		Orthology	Orthofinder	N.a
1.1.1.215-RXN		Ec-1.1.1.215	1	1		Orthology	Orthofinder	N.a
1.1.1.220-RXN	Nil	Ec-1.1.1.220	1	2		Orthology Annotation	Orthofinder Pathwaytools	N.a
1.1.1.264-RXN	Nil	Ec-1.1.1.264	1	1		Annotation	Pathwaytools	N.a
1.1.1.271-RXN	Nil	Ec-1.1.1.271	1	0		Orthology Annotation	Orthofinder Pathwaytools	N.a
1.1.1.272-RXN	Nil	Ec-1.1.1.272	1	0		Orthology Annotation	Orthofinder Pathwaytools	N.a



Ex : Reconstruction by orthology informations

Orthology

	Common-name	Ec-number	Reconstruction tool	Reconstruction source	Reconstruction comment	Nb gene associated	Nb pathway associated
+NEOMENTHOL-DEHYDROGENASE-RXN		Ec-1.1.1.208	Orthofinder	Cladosiphon okamuranus	N.a	1	1
1-ACYLGLYCEROL-3-P-ACYLTRANSFER-RXN	Nil 1-acylglycerol-3-phosphate o-acyltransferase	Ec-2.3.1.51	Orthofinder Pathwaytools	Cladosiphon okamuranus Genome	N.a	10	2
1-PHOSPHATIDYLINOSITOL-3-KINASE-RXN	Nil 1-phosphatidylinositol-3-kinase	Ec-2.7.1.137	Orthofinder Pathwaytools	Cladosiphon okamuranus Fistulifera solaris Undaria pinnatifida kr2015 Genome Ectocarpus subulatus Ectocarpus siliculosus	N.a	20	1
1-PHOSPHATIDYLINOSITOL-KINASE-RXN	Nil 1-phosphatidylinositol 4-kinase	Ec-2.7.1.67	Orthofinder Pathwaytools	Cladosiphon okamuranus Fistulifera solaris Undaria pinnatifida kr2015 Genome Ectocarpus subulatus Ectocarpus siliculosus Nannochloropsis gaditana Thalassiosira pseudonana	N.a	15	2
1.1.1.145-RXN		Ec-1.1.1.145	Orthofinder	Ectocarpus subulatus Ectocarpus siliculosus	N.a	2	0
			Orthofinder	Cladosiphon okamuranus			22

- Main page
- SMW-Ask
- workflow command history
- Random page
- Files
- Metabolic network components
- organism
- reaction
- gene
- pathway
- metabolite
- Reconstruction categories
- annotation
- manual
- orthology
- Reconstruction tools
- curation
- orthofinder
- pathwaytools
- Tools
- What links here
- Related changes
- Special pages
- Printable version
- Permanent link
- Page information
- Browse properties

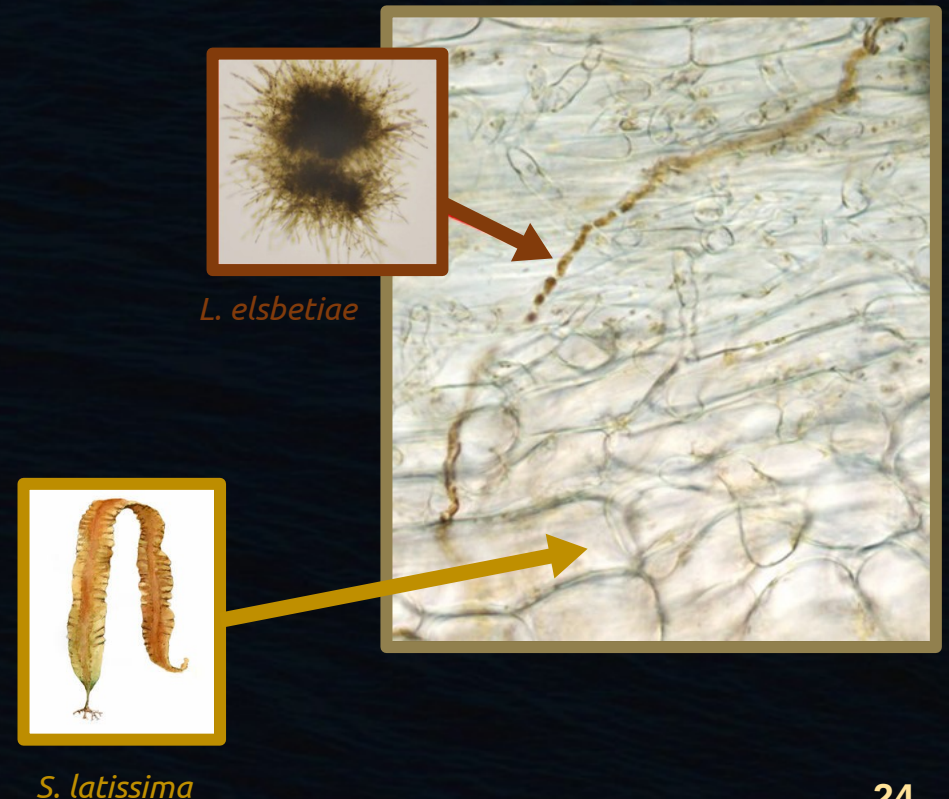
e. Perspectives

- Add more brown algae species :
 - species sequenced by Illumina in short read
 - soon available species sequenced in long read
 - *Ascophyllum nodulosum*
 - *Pylaiella littoralis*
 - *Heterosigma akashiwo*
- Understand the potential technical biases that would induce phylogenetic mismatches

2. Study the evolution of gene loss in *Laminarionema elsbetiae*

Laminarionema elsbetiae:

- Filamentous Brown Algae
- Order of Ectocarpales
- Endophyte of *Saccharina latissima*
 - Also found in *S. japonica* and rarely in *L. digitata*
- Different impacts (physiological and transcriptomic) depending on the host inhabited



a. Finding the lost genes

Simple method :

From a file indicating which reactions are present in which species :

- Select the reactions present in all species except one
- Select precisely those where it is the case for *L. elsbetiae*
- Find the intersection of these reactions between the different runs
- Define the genes / proteins associated with these reactions

b. Results

- **Dataset 02 (2022)** : ● Long read only + Diatoms
- **Dataset 03 (2022)** : ● Long read + *L. elsbetiae* + Diatoms
- **Dataset 40 (20 ??)** : 40 algae + 7
 - 12 Brown Algae
- **Dataset A0 (2020)** : 12 brown algae + 4 outgroups
- **Dataset A1 (2022)** : same as 0A with new datafiles
- **Dataset A2 (2022)** : same as 1A : species with better quality assembly & + 1 species in outgroup
- Quality expected : $A2 > A1 > A0$

b. Results

Intersection with A0 : 01 n 40 n A0

- **8** Reactions lost

Intersection with A1 : 01 n 40 n A1

- **16** Reactions lost

Intersection with A2 : 01 n 40 n A2

- **27** Reactions lost

c. Interpretation

These supposed lost genes could also be :

- Genes misassembled
- Genes not annotated

Some of these supposedly lost genes are found in the transcriptome

c. Perspectives

Inject transcriptome information into the data

The simple method could be improved :

- Is 1 the optimal number to chose of species not having the reaction ?
- Doesn't take into account the position of species in the phylogenic tree
- Doesn't interpret lost or gain

3. Interactions between the endophyte *L. elsbetiae* and its host *S. latissima*

Study perspective :

- Understand whether losses of metabolic function are compensated by the host
- Understand whether certain metabolic functions of the host explain its acclimatization to the infection of the endophyte
- Understand how the endophyte manage to inactivate host defenses