

Module Explorer

The Shiny application here described provides a very simple tool for exploring the full collection of multi-gene modules detected with the method described in: "Computational identification of co-evolving multi-gene modules in microbial biosynthetic gene clusters".

Getting Started

Despite being deployed as a web-application, the Module Explorer is meant to be run locally. Therefore it requires the installation of some software in order to run.

Prerequisites

Prior installation of the Module Explorer, it is necessary to download and install [R](#) version 3.4.4 or later. R is a programming language and free software environment. If you are unfamiliar with the R syntax and commands, it is recommended to read the [introduction to R](#) document first.

The download and installation of [RStudio](#) is also necessary. RStudio is a free and open-source integrated development environment for R.

In order to fully exploit the Module Explorer capabilities, [Firefox](#) the web browser should be installed in your computer and should be set as default browser.

The Module Explorer depends on a number of other R packages, which need to be installed. This can be simply done by copying and pasting the following into the RStudio GUI console

```
install.packages(c("shiny", "DT", "markdown"))
```

Installing

From the RStudio GUI console, you can navigate to the folder where you wish to install the Module Explorer using the following command:

```
setwd("/path/to/chosen/folder/")
```

Next, you can download the Module explorer software by copying and pasting the following:

```
download.file(url="https://github.com/francescodc87/Modules-explorer/archive/master.zip", destfile="Modules-explorer-master.zip")
```

Once downloaded, the file should be unzipped and deleted directly from the RStudio GUI console with the following commands:

```
unzip("Modules-explorer-master.zip")
file.remove("Modules-explorer-master.zip")
file.rename(from = "Modules-explorer-master/", to="Module_Explorer/")
```

Some of the files needed are too big to be hosted by GitHub. The [Zenodo](#) research data repository is used to store these additional data. The user can manually download the additional files from <https://zenodo.org/record/1434902/>, extract the .zip file and move it into the Module_Explorer/ folder. It is necessary that the www folder extracted from the .zip file is merged with the www folder already present within the Module_Explorer folder. Alternatively, the additional files can be simply downloaded by copying and pasting the following:

```
download.file(url="https://zenodo.org/record/1434902/files/www.zip?download=1", destfile = "Module_Explorer/www.zip")
unzip("Module_Explorer/www.zip", exdir="Module_Explorer/", unzip=getOption("unzip"), junkpaths = TRUE)
file.remove("Module_Explorer/www.zip")
```

The zip file downloaded is 3.6 GB in size; therefore the download and the unzipping will take several minutes.

Once the download is terminated, the Module Explorer web application can be launched by copying and pasting the following:

```
shiny::runApp("Module_Explorer")
```

If everything works correctly, after few moments the initial page of the Module Explorer should appear on your screen as shown.

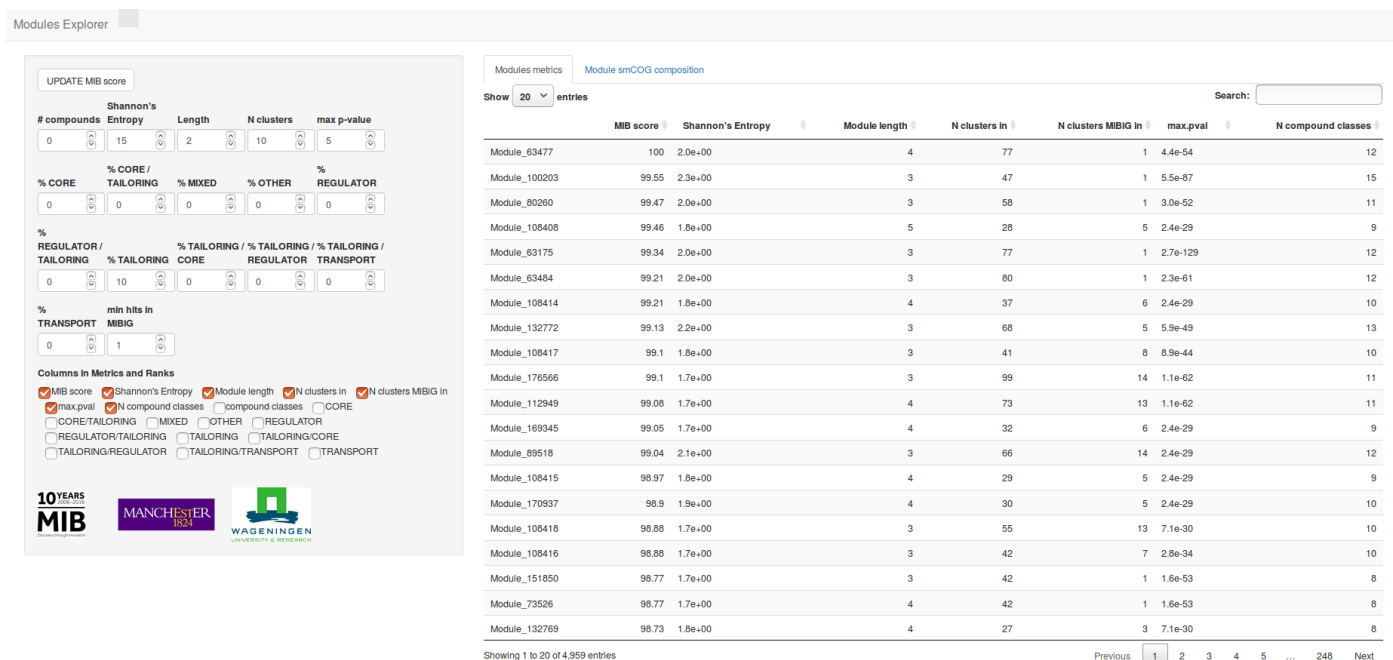


Figure 1: initial page of Module Explorer web application.

Module explorer usage

The table on the left side of Figure 1 contains all the detected modules ranked according to the MIB score computed using the weights described in the manuscript. As shown in Figure 2, it is possible to modify the weights according to the user needs.

UPDATE MIB score

Shannon's

compounds Entropy Length N clusters max p-value

0 5 1 0 0

% CORE % CORE / % MIXED % OTHER % REGULATOR

-1 -1 -1 5 -1

% REGULATOR / % TAILORING % TAILORING / % TAILORING / % TAILORING /

TAILORING % TAILORING CORE REGULATOR TRANSPORT

-1 15 -1 -1 -1

% min hits in

TRANSPORT MIBIG

-1 2

Columns in Metrics and Ranks

☒ MIB score ☒ Shannon's Entropy ☒ Module length ☒ N clusters in ☒ N clusters MIBIG in

☒ max.pval ☒ N compound classes ☐ compound classes ☐ CORE

☐ CORE/TAILOING ☐ MIXED ☐ OTHER ☐ REGULATOR

☐ REGULATOR/TAILOING ☐ TAILORING ☐ TAILORING/CORE

☐ TAILORING/REGULATOR ☐ TAILORING/TRANSPORT ☐ TRANSPORT

10 YEARS MIB 2006-2016

MANCHESTER 1824

WAGENINGEN UNIVERSITY & RESEARCH

Figure 2: MIB score weights.

After modifying the parameters shown in Figure 2, it is necessary to click on the UPDATE MIB score button in order to update the shown table.

It is also possible, as shown in Figure 3, to select any module from the table by simply clicking on the related row.

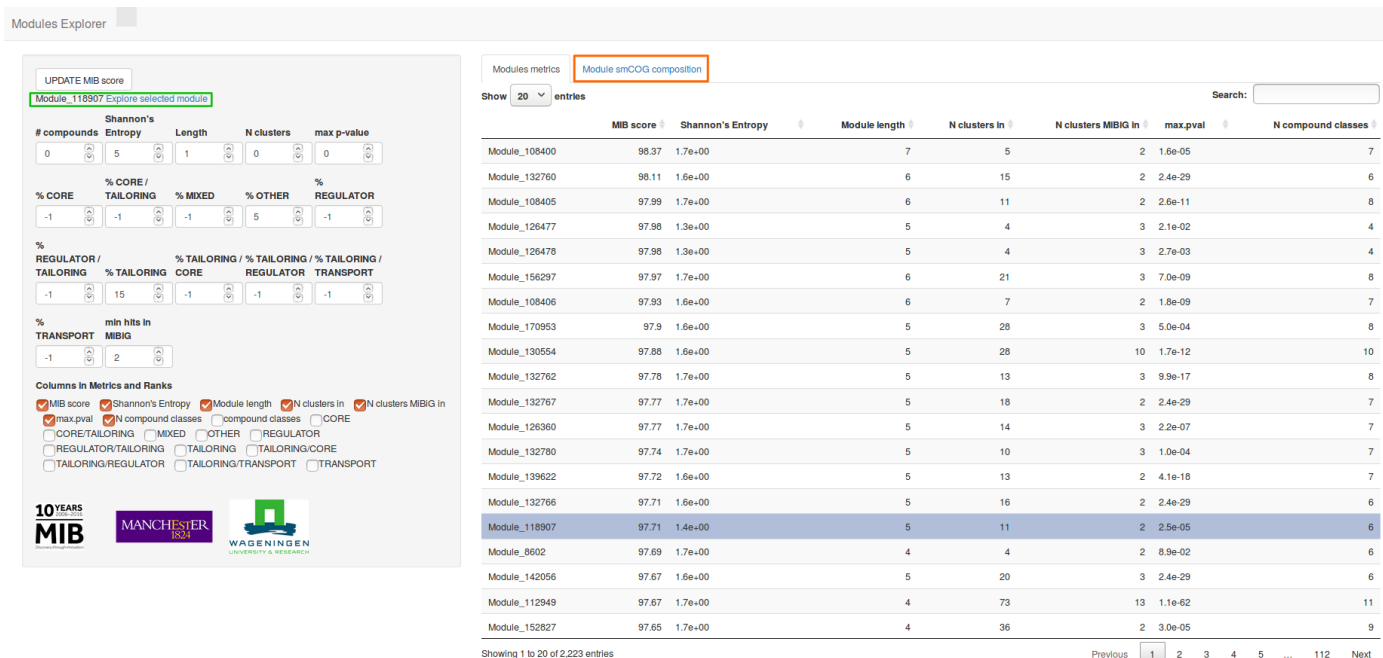


Figure 3: Selected module.

Once a module is selected, it is possible to click on the Module smCOG composition tab (highlighted in orange in Figure 3).

Modules metrics

Module smCOG composition

Show 10 entries

Search:

SMCOG10031	SMCOG10154	SMCOG10236	SMCOG10393	SMCOG11537
EAZ82985.1	EFQ56927.1	EFQ56945.1	EFQ56878.1	YP_001250706.1 methoxymalonyl_CoA_synthase
ZP_07720899.1 acyl-CoA_dehydrogenase	ZP_07726052.1 putative_3-hydroxybutyryl-CoA_dehydrogenase	ZP_07726070.1 conserved_hypothetical_protein	ZP_07726003.1 FkbH_domain_protein	YP_003619243.1 methoxymalonyl_CoA_synthase
EFP16644.1	ZP_08192355.1 3-hydroxyacyl-CoA_dehydrogenase_NAD-binding	EGH41898.1 hypothetical_protein	AAM54094.1 unknown	YP_095952.1 methoxymalonyl_CoA_synthase
ZP_07417577.1 acyl-CoA_dehydrogenase_fadE7	ZP_08192360.1 3-hydroxyacyl-CoA_dehydrogenase_NAD-binding	ZP_06496701.1 hypothetical_protein	YP_003100945.1 FkbH_like_protein	YP_124231.1 hypothetical_protein
EFP24781.1	EFK33383.1	YP_237384.1 hypothetical_protein	YP_003493870.1 putative_methoxymalonnate_biosynthesis_protein	YP_127243.1 hypothetical_protein
ZP_07425718.1 acyl-CoA_dehydrogenase_fadE7	ZP_07086591.1 possible_3-hydroxyacyl-CoA_dehydrogenase	EGH69552.1 hypothetical_protein	AAZ94395.1 methoxymalonnate_biosynthesis_protein	CBX00237.1 hypothetical_protein
EFQ56943.1	EGB52006.1 3-hydroxyacyl-CoA_dehydrogenase	EGH76910.1 hypothetical_protein	ZP_09983145.1 subfamily_IIIc_HAD-superfamily_phosphatase	ADU56327.1 31-O-demethyl-FK506_methyltransferase
ZP_07726068.1 putative_acyl-CoA_dehydrogenase	ZP_04534876.1 3-hydroxybutyryl-CoA_dehydrogenase	ACY01398.1 acyl_carrier_protein	ZP_09986934.1 subfamily_IIIc_HAD-superfamily_phosphatase	ZP_10073068.1 FkbM_family_methyltransferase
EFU49538.1 acyl-CoA_dehydrogenase_protein	ZP_07198201.1 putative_3-hydroxybutyryl-CoA_dehydrogenase	ACY01388.1 acyl_carrier_protein	ADC79624.1 BatE	ADX99529.1 FkbM
ZP_04534874.1 conserved_hypothetical_protein	ZP_07177405.1 putative_3-hydroxybutyryl-CoA_dehydrogenase	ZP_08240513.1 putative_acyl_carrier_protein	AAO06922.1 GdmH	CAA60470.1 methyltransferase

Showing 1 to 10 of 652 entries

Previous 1 2 3 4 5 ... 66 Next

Figure 4: smCOG composition of selected module.

As shown in Figure 4, this will show the smCOG composition of the selected module.

Additionally, when selecting a module, a new link will appear below the UPDATE MIB score (highlighted in green). When clicked, this link will open a new tab showing the selected module mapped on all the biosynthetic gene clusters containing it, similarly to what is shown in Figure 5.

Module 118907 mapped to gene clusters

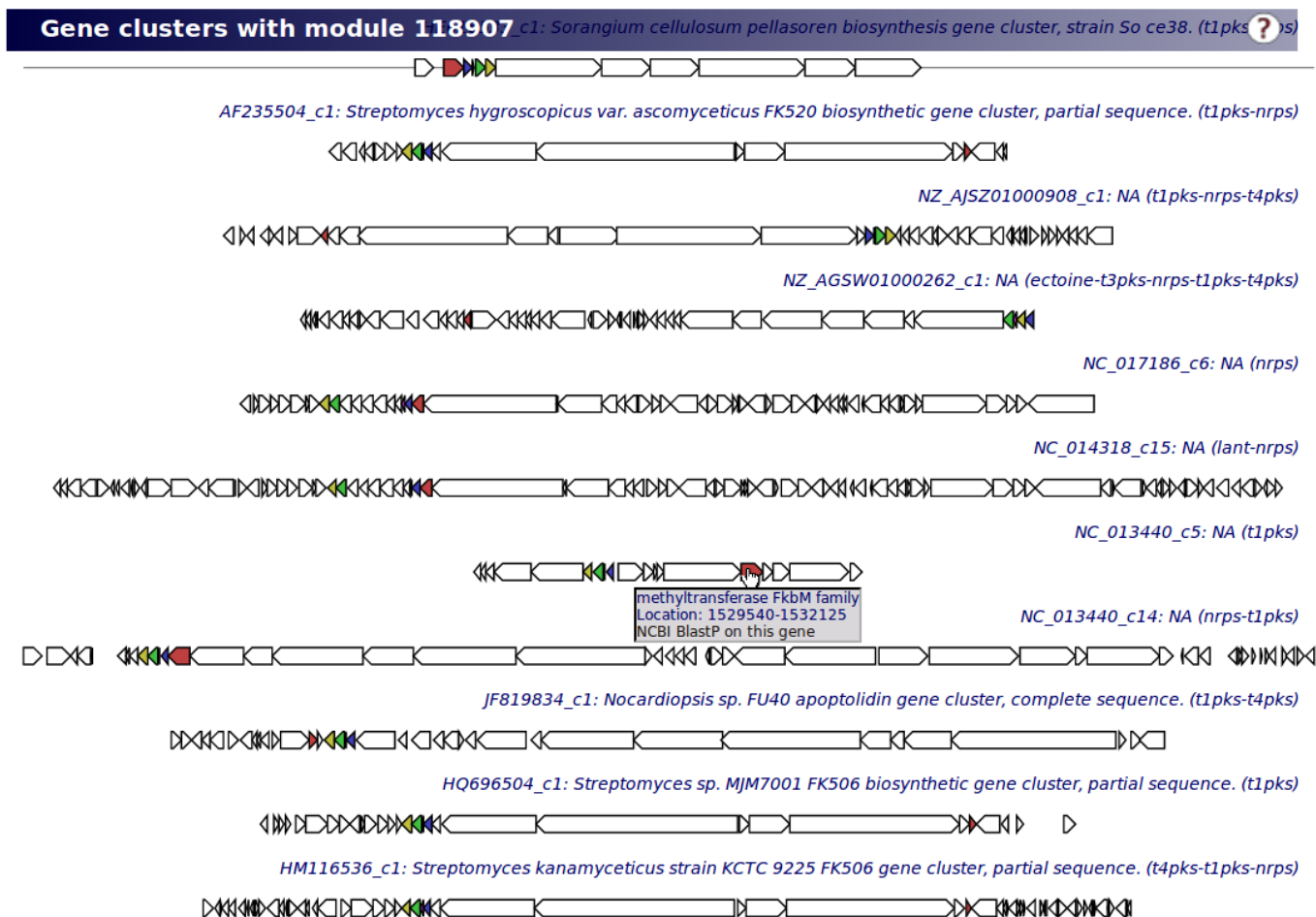


Figure 5: selected module mapped on Biosynthetic gene clusters.

It is possible to click on any gene shown in this page, this will open a pop-up window showing the gene annotation, its location and providing a link to the NCBI Blast result for this gene.

Built With

- [antiSMASH](#)
- [R](#)
- [RStudio](#)
- [Shiny](#)

Authors

- Francesco Del Carratore
- Konrad Zych
- Matthew Cummings
- Eriko Takano
- Marnix Medema
- Rainer Breitling