Contents of the *MegaFD\_publication* folder

Last updated: 24.07.2023

# MegaFD\_data:

This is the data folder from where I source and load data in my scripts.

It contains trait data, spatial data and RData files for data that were too time-consuming to compute every time I run the code as I moved forward with the analysis.

Note: the original shapefiles of species distribution from IUCN I used to generate the presence-absence matrix can be found in the *marine\_megafauna\_FD (fixing mapping issue)* folder: <https://drive.google.com/drive/folders/1tKrCEFuK3PW8GdNapFpZLUezEtIzZA-v?usp=drive_link>

# MegaFD\_functions:

This is folder contains the scripts with the functions I source for computing indices (e.g. FRic, FUSE, Redundancy by Ricotta et al. etc).

# thesis\_figures/ IUCN&Aqua\_combo\_figures:

This folder contains all the plots and maps I have been generating over the past months. Some of them are redundant. I am now working on a script where I gather all the plots and maps we’re thinking of including in the manuscript, so that we can organise the visuals in a much effective way.

# Mega\_FD\_markdown

IUCN\_species\_presabs\_matrix.Rmd:

* Inspects the spatial data sourced from IUCN.
* Detials the data I have used to generate the presence absence matrix for the species with data from IUCN.
* Generates the original presence absence matrix I’ve later used in combination with the Aquamaps one (see the next two scripts).

## presabs\_matrix\_IUCN\_Aqua.Rmd:

* Generates presence-absence matrices based on the species for which IUCN and Aquamaps have data for.
* Presence-absence matrices for Aquamaps’ species are generated for different thresholds of probability of occurrence per grid cell.

## matching\_IUCN\_Aqua.Rmd:

* Lists the species from Aquamaps that complement the IUCN spatial dataset.
* Compares maps of SR and FRic based on the shared species pool and assesses the most appropriate probability-of-occurrence threshold to choose for complementing Aquamaps with IUCN data.
* Generates a new data file with IUCN spatial data complemented with Aquamaps.

# Mega\_FD\_markdown/IUCN&Aqua\_combined:

This folder gathers all the Rmd files with the analyses conducted on the combined spatial dataset from IUCN and Aquamaps.

01.raw\_trait\_data\_inspection.R:

* Inspects the file with raw trait data collected for the 334 species of marine megafauna.
* Checks what traits have been selected for the analysis, how their values have been coded, and inspects what species are missing trait data.

## 02.bis.data\_curation.Rmd:

* Curates the trait data, computing the mean trait value per species from all the imputations.
* Sets the global species pool based only on those species that occur in grid cells with more than 4 species (i.e. one unit bigger than the chosen dimension of the functional space).

## 03.bis.distance\_matrix.Rmd:

* Computes the trait distance matrix.

## 04.bis.quality\_functional\_space.Rmd:

* Assesses the most parsimonious dimension to chose to build the functional space.

## 06.bis.biodiversity\_metrics.Rmd:

* Computes FRic and SR per grid.

P.s. I called the script “biodiversity metrics” because I thought of including FUn, FSp etc in there too but then changed my mind and forgot to rename the file.

## 07.bis.FRic\_SR\_relationship.Rmd:

* Models the relationship between FRic and SR with a LOESS curve.
* Plots the residuals from the LOESS model.

## 08.bis.FRed.Rmd:

* Computes Uniqueness and Redundancy according to the methodology suggested by Ricotta et al. (2016).

## 09.bis.FUSE\_FUn\_FSp\_FUS\_maps.Rmd:

* Plots the species richness of the top 10, 15, 20, and 25% species for FUSE, FUn, and FSp.

P.s. As previously discussed, this analysis and visuals shouldn’t be part of the manuscript anymore.

## 10.bis.FUn.Rmd:

* Computes functional uniqueness per grid cell as:
  + mean distance of a species from its five nearest neighbors from the the global species pool (i.e. global FUn per grid).
  + mean distance of a species from the five nearest neighbours oly from the pool of species occurring its the same cell (i.e. local FUn per grid).

## 12.bis.FD\_per\_clade.Rmd:

* Computes and maps FRic, SR, FUn (both global and local), and Ricotta’s redundancy per clade (mammals, bony fish, sharks + rays, reptiles + aves + molluscs).

## 13.bis.FRic\_SR\_spatial\_overlap.Rmd:

* Identifies the hotspots of FRic and SR.
* Assesses the spatial overlap between hotspots of FRic and SR.
* Plots FRic and SR hotspots together with Redundancy coldspots.
* Performs S.E.S analysis to test whether the FRic scores of the FRic hotspots are higher than expected based on their species richness.

## 16.bis.bioregionalisation\_marine\_waters.Rmd:

* Shows how coastal and pelagic waters are divided in marine provinces according to Spalding et al. (2007, 2012).

## FRic\_hotspots\_per\_province\_top10FUSE\_spp.R:

* Clusters FRic hotspots per marine province according to Spalding et al. (2007).
* Identifies the top 10 FUSE species within each province.
* Assess how similar the lists of top 10 FUSE species are across provinces with Jaccard (dis)similarity index.
* Maps the top 10 FUSE species in the global functional space per each province.

## FRed\_coldspots\_per\_province\_top10FUSE\_spp.R:

* Clusters Redundancy coldpsots per marine province according to Spalding et al. (2007, 2012).
* Identifies the top 10 FUSE species within each coldspot province.
* Assess how similar the lists of top 10 FUSE species are across provinces with Jaccard (dis)similarity index.
* Maps the top 10 FUSE species in the global functional space per each province.

manuscript\_figures.Rmd:

* This is the code used to format maps and graphs and group them in multi-figure panels.

Contents of *marine\_megafauna\_FD (fixing mapping issue)* folder

Last updated: 24.07.2023

This folder contains the up-to-date R scripts I’ve used to regenerate the higher resolution (0.5 degree) presence-absence matrix based on IUCN spatial data (*megafauna\_IUCN\_presabs\_0.5res.txt*, which can be found here: <https://drive.google.com/file/d/1qjFScymu-pqRsQJDCVNSgri5dZeGjeoW/view?usp=drive_link>).

The error I’ve encountered is explained in the script *02.IUCN\_pres\_abs\_matrix\_to\_fix.R*.

The script *test\_pres\_abs\_0.5res\_a\_cornous.R* tested the code for regenerating the higher resolution presence-absence matrix with a subset of the spatial data - I thought it worked but I was wrong.