

La synthèse des connaissances :  
une introduction aux méta-analyses et  
revues systématiques  
**- Extraction des données quantitatives -**

Mercredi 5/10/22 - Montpellier

Dakis-Yaoba Ouédraogo – PatriNat (OFB-CNRS-MNHN)

[dakis-yaoba.ouedraogo@mnhn.fr](mailto:dakis-yaoba.ouedraogo@mnhn.fr)



# Extraction des données quantitatives

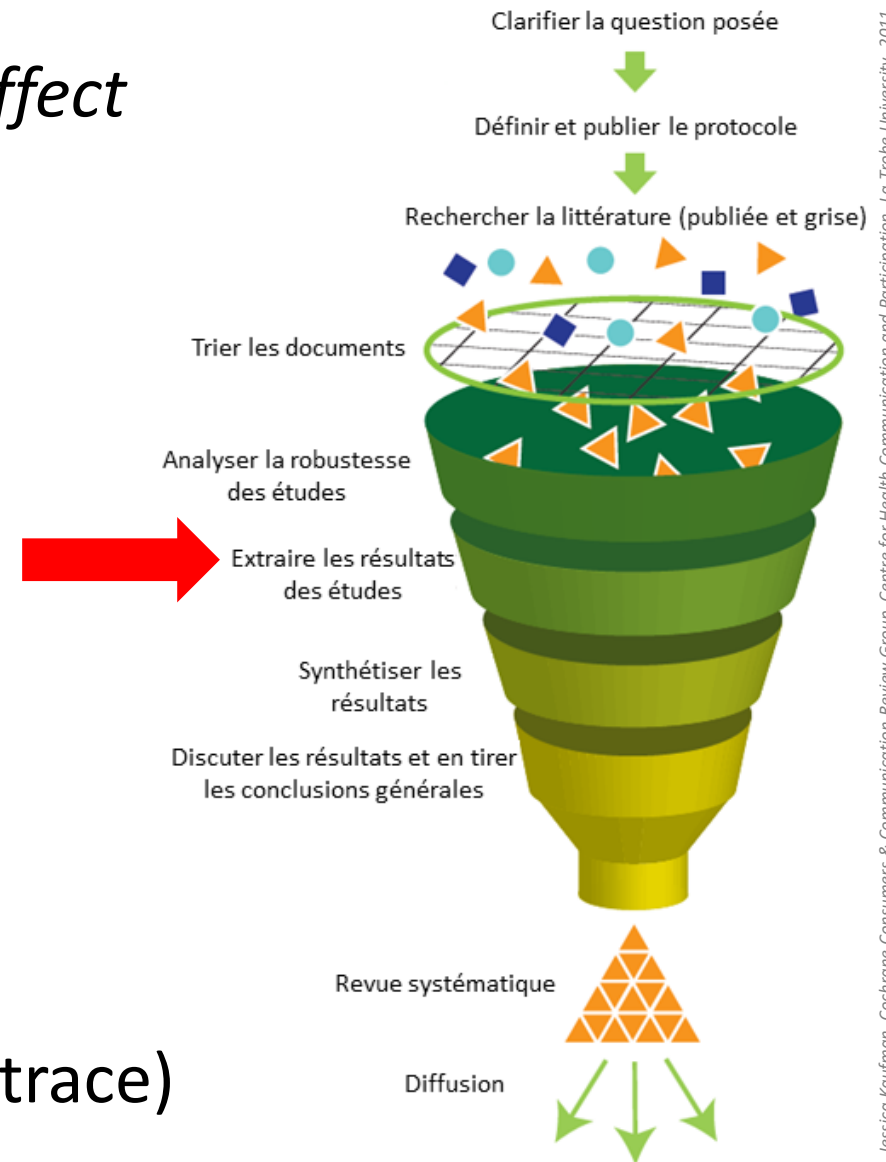
Extraction des données nécessaires au calcul des *effect sizes* (moyenne, effectif, sd/se/IC 95%)

+ extraction des variables pouvant expliquer l'hétérogénéité des effect sizes (*effect modifiers*)

Extraction à partir

- texte
- tableau
- figure
- supp. mat.

+ des calculs peuvent être nécessaires (garder une trace)



# Extraction des données quantitatives

! Attention !

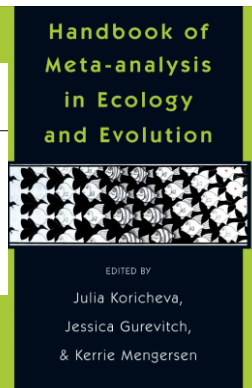
L'extraction des données prend du temps : bien définir la grille d'extraction, les *effect modifiers* à extraire

Importance de **tester** la grille d'extraction sur un échantillon d'articles pour vérifier l'adéquation avec le contenu des études

Documenter le travail, les décisions (transparence, répétabilité)

Décider quoi faire en cas d'**information manquante**  
(« missing data », contacter les auteurs, imputation)

13  
Recovering Missing or Partial Data from Studies:  
A Survey of Conversions and Imputations  
for Meta-analysis  
Marc J. Lajeunesse



# Consistency check

S'assurer de l'objectivité / robustesse de l'extraction :

- extraction des données de chaque étude effectuée **indépendamment** par 2 personnes
- si plusieurs personnes se partagent le travail, **vérifier la cohérence de l'extraction** entre les personnes avant le début du travail sur un échantillon (discuter les désaccords)
- si 1 seule personne, faire vérifier un échantillon de l'extraction par quelqu'un au début du travail (discuter les désaccords)

# Exemples de grille d'extraction

Echelle **cas d'étude** : ex. plusieurs concentrations d'une même exposition

IDdata	ID_map	author	...	taxon	Population_descri	Life_stage	Type_system	Tempera	pH
ScreenTA_9680	880	Cantin, N.E. ...		<i>Acropora tenuis</i>	Colonies	Adult	500 L outdoor tank	27.5	NA
ScreenTA_9680	880	Cantin, N.E. ...		<i>Acropora tenuis</i>	Colonies	Adult	500 L outdoor tank	27.5	NA
ScreenTA_9680	884	Cantin, N.E. ...		<i>Acropora valida</i>	Colonies	Adult	500 L outdoor tank	27.5	NA
ScreenTA_9680	884	Cantin, N.E. ...		<i>Acropora valida</i>	Colonies	Adult	500 L outdoor tank	27.5	NA
ScreenTA_9680	889	Cantin, N.E. ...		<i>Pocillopora damicornis</i>	Colonies	Adult	500 L outdoor tank	27.5	NA
ScreenTA_9680	889	Cantin, N.E. ...		<i>Pocillopora damicornis</i>	Colonies	Adult	500 L outdoor tank	27.5	NA

Treatment_description	Control_description	Solvent	Concentration_nom	Concentration_eff	Duration	Measured_variable	Time_after
Diuron	Unfiltered oceanic seawater	No	1 µg/L	0.91 µg/L	53 days	Symbiodinium density / total protein	NA
Diuron	Unfiltered oceanic seawater	No	10 µg/L	8.8 µg/L	53 days	Symbiodinium density / total protein	NA
Diuron	Unfiltered oceanic seawater	No	1 µg/L	0.91 µg/L	90 days	Symbiodinium density / total protein	NA
Diuron	Unfiltered oceanic seawater	No	10 µg/L	8.8 µg/L	90 days	Symbiodinium density / total protein	NA
Diuron	Unfiltered oceanic seawater	No	1 µg/L	0.91 µg/L	67 days	Symbiodinium density / total protein	NA
Diuron	Unfiltered oceanic seawater	No	10 µg/L	8.8 µg/L	67 days	Symbiodinium density / total protein	NA

## Exemple de grille d'extraction

Treatment_description	Control_description	Solvent	Concentration_nom	Concentration_eff	Duration	Measured_variable	Time_after
Diuron	Unfiltered oceanic seawater	No	1 µg/L	0.91 µg/L	53 days	Symbiodinium density / total protein	NA
Diuron	Unfiltered oceanic seawater	No	10 µg/L	8.8 µg/L	53 days	Symbiodinium density / total protein	NA
Diuron	Unfiltered oceanic seawater	No	1 µg/L	0.91 µg/L	90 days	Symbiodinium density / total protein	NA
Diuron	Unfiltered oceanic seawater	No	10 µg/L	8.8 µg/L	90 days	Symbiodinium density / total protein	NA
Diuron	Unfiltered oceanic seawater	No	1 µg/L	0.91 µg/L	67 days	Symbiodinium density / total protein	NA
Diuron	Unfiltered oceanic seawater	No	10 µg/L	8.8 µg/L	67 days	Symbiodinium density / total protein	NA

Metaanalyse_data	unit	ID_experiment	ID_case	ID_common_control	N_c	Mean_c	Type_variation_c	Variation_c	N_t	Mean_t	Type_variation_t	Variation_t
OK (Fig3, SE, n=6)	x 10^6 / mg protein	1	3	1	6	4.2	sd	1.714642819	6	3.5142857	sd	0.979795897
OK (Fig3, SE, n=6)	x 10^6 / mg protein	1	4	1	6	4.2	sd	1.714642819	6	3.6	sd	0.524890659
OK (Fig3, SE, n=6)	x 10^6 / mg protein	2	7	2	6	0.928571	sd	0.454905237	6	1.4142857	sd	0.979795897
OK (Fig3, SE, n=6)	x 10^6 / mg protein	2	8	2	6	0.928571	sd	0.454905237	6	1.3142857	sd	0.699854212
OK (Fig3, SE, n=6)	x 10^6 / mg protein	3	11	3	6	1.714285	sd	0.699854212	6	2.2285714	sd	0.699854212
OK (Fig3, SE, n=6)	x 10^6 / mg protein	3	12	3	6	1.714285	sd	0.699854212	6	0.9142857	sd	0.244948974

[illegible]

# Outils d'extraction des figures

<https://automeris.io/WebPlotDigitizer/>

WebPlotDigitizer

Web based tool to extract data from plots, images, and maps

Home

Blog

Tutorials

Citation

Privacy

File

Help

Zoom

100%

Fit

Image

Axes

Datasets

Measurements

Choose Plot Type

2D (X-Y) Plot

2D Bar Plot

Polar Diagram

Ternary Diagram

Map With Scale Bar

Image

Align Axes

Cancel

Align Bar Chart Axes

P<sub>1</sub>

P<sub>2</sub>

0

2

4

6

8

10

1

2

3

4

5

6

7

8

Click on two known points (P<sub>1</sub>, P<sub>2</sub>) on the continuous axes along the bars

Proceed

Version 4.5 Released (August 15, 2021)

[\[ Release Notes \]](#)

It is often necessary to reverse engineer images of data visualizations to extract the underlying numerical data. WebPlotDigitizer is a semi-automated tool that makes this process extremely easy.

- Works with a wide variety of charts (XY, bar, polar, ternary, maps etc.)
- Automatic extraction algorithms make it easy to extract a large number of data points
- Free to use, opensource and cross-platform (web and desktop)
- Used in hundreds of published works by thousands of users
- Also useful for measuring distances or angles between various features
- More to come soon...

WebPlotDigitizer - Copyright 2010-2019 Ankit Rohatgi

File Help

Image Axes Datasets Measurements

Symbiodinium spp. density (x10<sup>6</sup>) mg<sup>-1</sup> protein

control

Choose Plot Type

- ☐ 2D (X-Y) Plot
- ☒ 2D Bar Plot
- ☐ Polar Diagram
- ☐ Ternary Diagram
- ☐ Map With Scale Bar
- ☐ Image

Align Axes Cancel

Align Bar Chart Axes

Click on two known points (P<sub>1</sub>, P<sub>2</sub>) on the continuous axes along the bars

Proceed

WebPlotDigitizer - Copyright 2010-2019 Ankit Rohatgi

File Help

Image Axes Datasets Measurements

Symbiodinium spp. density (x10<sup>6</sup>) mg<sup>-1</sup> protein

control

Bar Chart Calibration

Enter the values at the two points selected on the continuous axes along the bars

Point 1 Point 2 Log Scale

0 5

☐ Rotated axes (not exactly vertical or horizontal)

OK

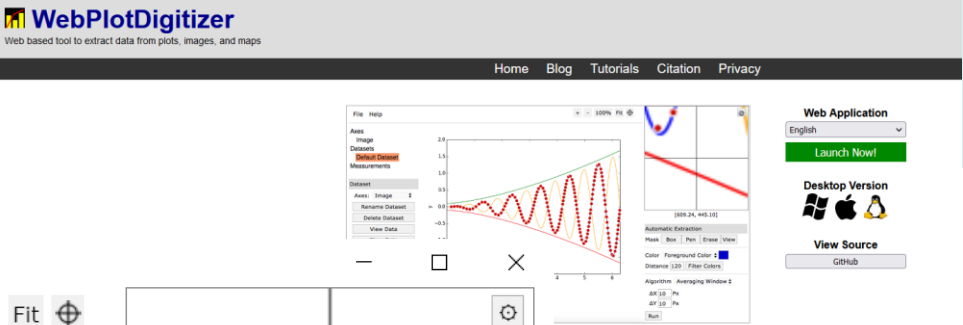
Axes Calibration

Click points to select and use cursor keys to adjust positions. Use Shift+Arrow for faster movement. Click complete when finished.

Complete!

A. tenuis A. valida P. damicornis

# Outils d'extraction des figures



WebPlotDigitizer - Copyright 2010-2019 Ankit Rohatgi

File Help

Image

Axes

Bar

Datasets

Default Dataset

Measurements

Dataset

Axes: Bar

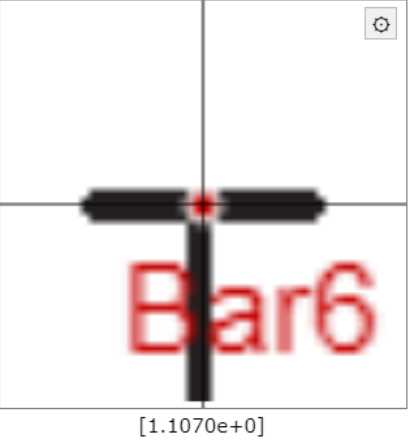
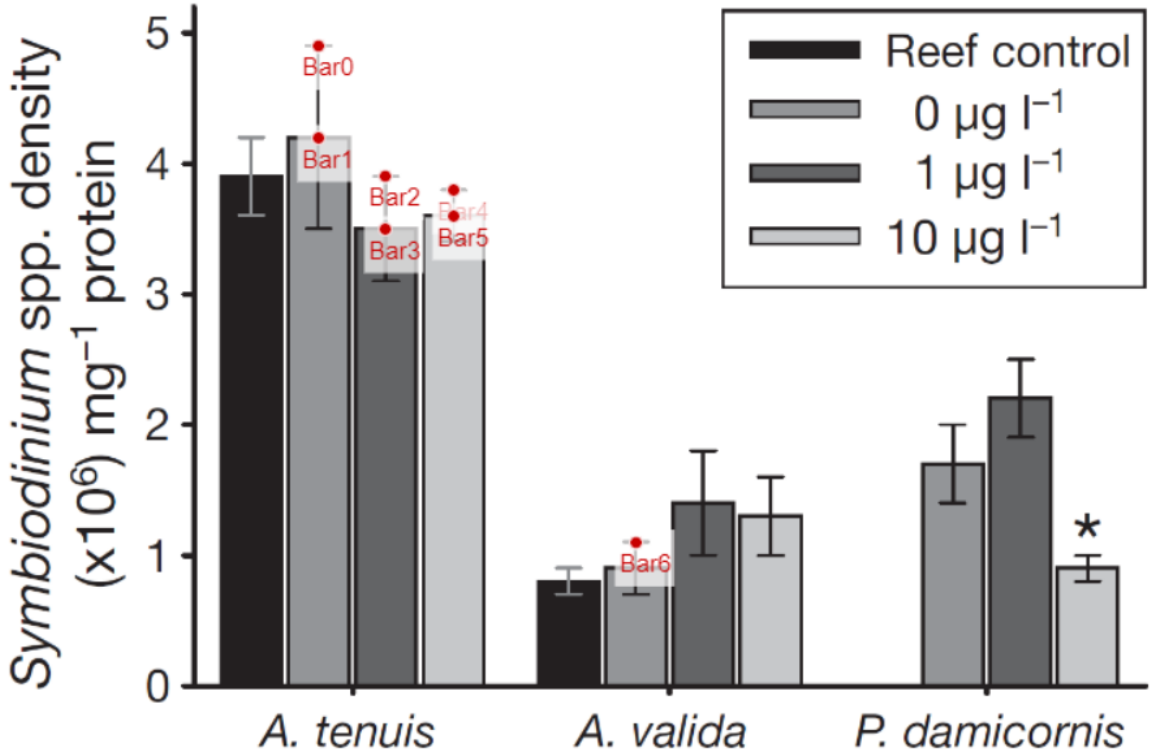
Rename Dataset

Delete Dataset

View Data

Clear Data

Data Points: 7



Manual Extraction

Add Point (A) Adjust Point (S)

Delete Point (D) Edit Labels (E)

Automatic Extraction

Mask Box Pen Erase View

Color Foreground Color

Distance 120 Filter Colors

Algorithm Bar Extraction

$\Delta X$  30 Px

$\Delta Y$  10 Px

Run

a visualizations to extract the underlying numerical data. WebPlotDigitizer is a semi-  
r.  
(, ternary, maps etc.)  
extract a large number of data points  
(and desktop)  
is of users  
tween various features

1.4.5 Released (August 15, 2021)

[\[ Release Notes \]](#)



# Outils d'extraction des figures

<https://plotdigitizer.com/>

PlotDigitizer

Free Online App   Features   Download   Buy Now


All-in-One Tool to Extract Data from Graphs, Plots & Images

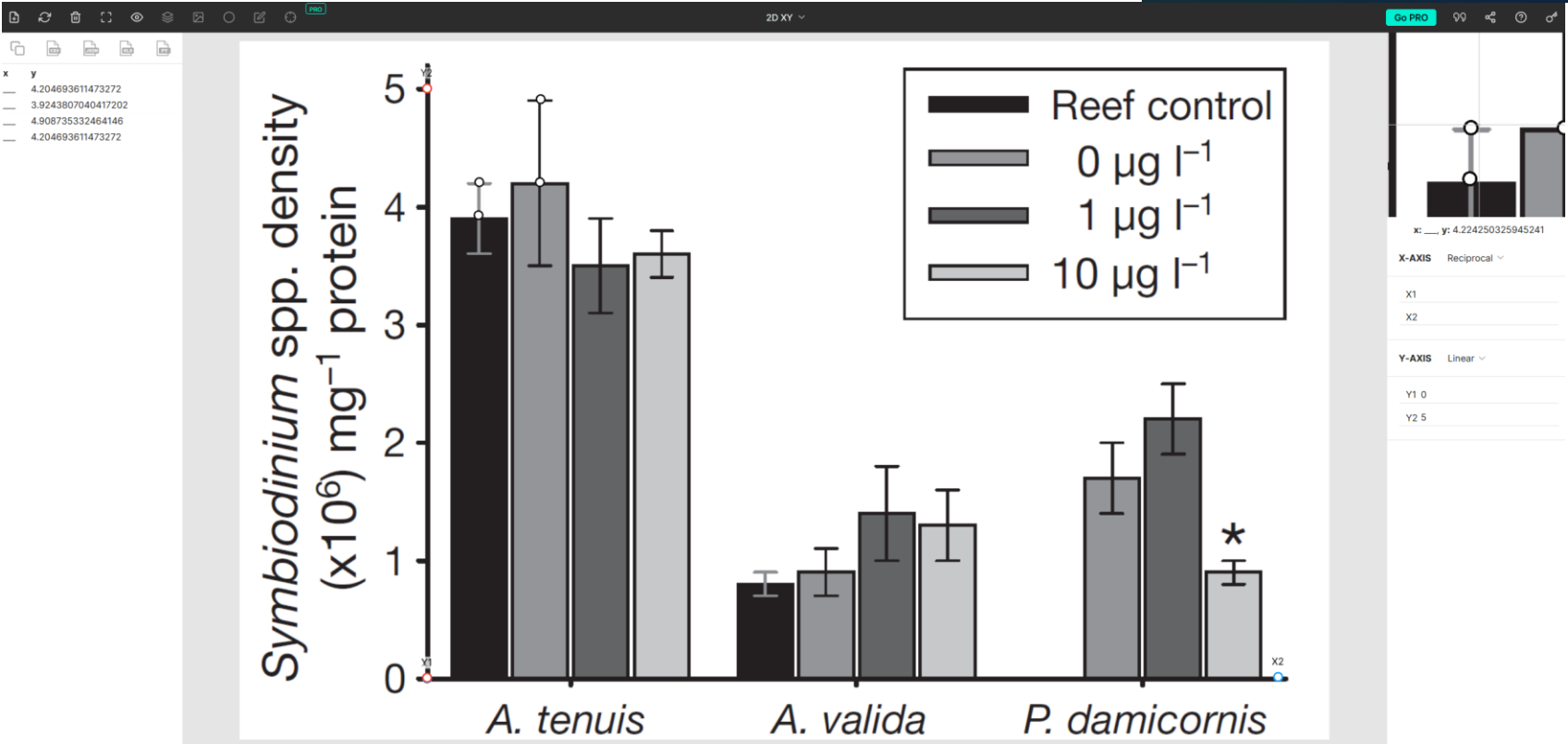
Plotdigitizer is an online data extraction tool that allows users to extract data from images in numerical format. In short, it reverse-engineers your visual graphs into numbers. The software comes with plenty of useful and time-saving features.

Launch App

Buy Now

PlotDigitizer






# Outils d'extraction des figures : metaDigitise

Received: 13 July 2018 | Accepted: 12 October 2018

DOI: 10.1111/2041-210X.13118

## APPLICATION


Methods in Ecology and Evolution 

## Reproducible, flexible and high-throughput data extraction from primary literature: The **META**DIGITISE R package

Joel L. Pick  | Shinichi Nakagawa | Daniel W. A. Noble 



Install Packages

Install from:  [Configuring Repositories](#)

Repository (CRAN) ▼

Packages (separate multiple with space or comma):

metaDigitise

(+) permet d'enregistrer, de re-tracer et de modifier l'extraction des données

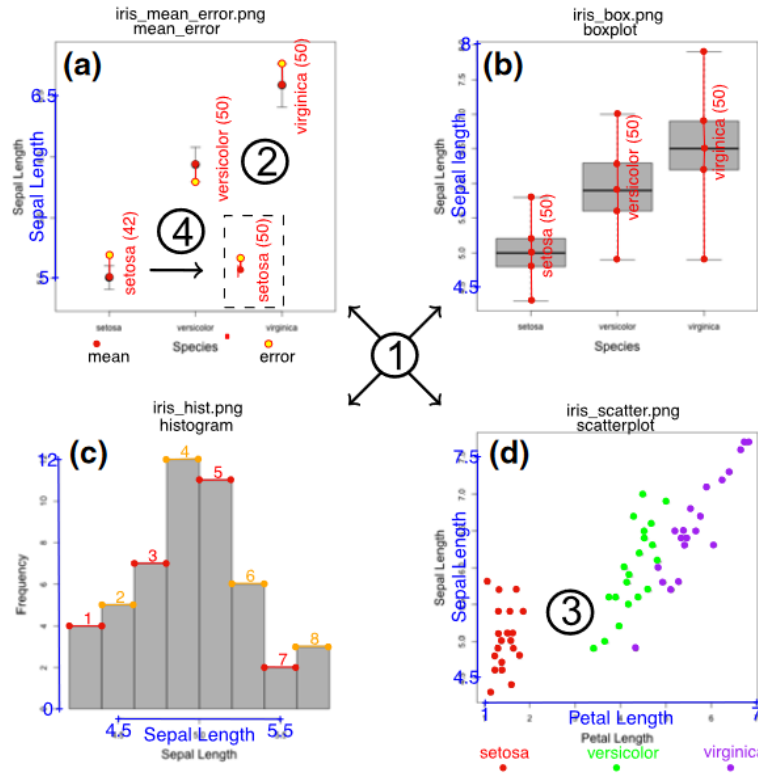
(-) pas de zoom

<https://cran.r-project.org/web/packages/metaDigitise/vignettes/metaDigitise.html>

# Outils d'extraction des figures : metaDigitise

## metaDigitise

### Plot interface



### Data output

filename	variable	group_id	mean	sd	n	r	plot_type
iris_box.png	Sepal length	setosa	5.01	0.317	50	NA	boxplot
iris_box.png	Sepal length	versicolor	5.93	0.497	50	NA	boxplot
iris_box.png	Sepal length	virginica	6.49	0.603	50	NA	boxplot
iris_hist.png	Sepal Length	setosa	4.95	0.364	50	NA	histogram
iris_mean_error.png	Sepal Length	setosa	5.01	0.680	50	NA	mean_error
iris_mean_error.png	Sepal Length	versicolor	5.94	1.025	50	NA	mean_error
iris_mean_error.png	Sepal Length	virginica	6.59	1.251	50	NA	mean_error
iris_scatter.png	Petal Length	setosa	1.44	0.215	20	0.109	scatterplot
iris_scatter.png	Sepal Length	setosa	5.03	0.427	20	0.109	scatterplot
iris_scatter.png	Petal Length	versicolor	4.29	0.415	20	0.786	scatterplot
iris_scatter.png	Sepal Length	versicolor	5.97	0.603	20	0.786	scatterplot
iris_scatter.png	Petal Length	virginica	5.66	0.668	20	0.932	scatterplot

## FUNCTIONALITY

### ① Different plot types

Capable of handling A) mean error plots, B) boxplots, C) histograms and D) scatterplots

### ② Entry of Metadata

Enter sample sizes variable and group names while digitising that are displayed on plot

### ③ Grouped Data

Enter as many groups as needed to capture descriptive statistics for sub-samples of data

### ④ Digitise, edit or replot digitisations

Simple user interface to guide user. Can digitise new images, edit digitisations or easily replot previous digitisations and metadata by cycling through images or choosing specific images

### ⑤ Summarising data

Get descriptive statistics automatically calculated for all plot types or use raw x,y data, if desired

### ⑥ Multiple image processing

Process as many images at once as needed and of varying types efficiently and quickly. New plots automatically plotted for digitisation

**Question** : estimer les effets des substances chimiques sur la performance de la photosynthèse (*maximum quantum yield,  $F_v/F_m$* ) des coraux constructeurs de récifs tropicaux

Echantillon de 3 articles

Une **étude** = combinaison **un taxon** × **une exposition** × **un outcome**

Un **cas d'étude** = un niveau de concentration-durée testé

**Effect size** = différence de moyenne standardisée

Si suivi au cours du temps, extraction de la durée d'exposition la plus longue

Extraire les données en utilisant le package metaDigitise

**Question** : estimer les effets des substances chimiques sur la performance de la photosynthèse (***maximum quantum yield, Fv/Fm***) des coraux constructeurs de récifs tropicaux

Une **étude** = combinaison **un taxon × une exposition × un outcome**

Un **cas d'étude** = un niveau de concentration-durée testé

**Effect size** = différence de moyenne standardisée

Si suivi au cours du temps, extraction de la durée d'exposition la plus longue

## 1 – Extraire les données

Installer et charger le package metaDigitise ; spécifier le répertoire de travail

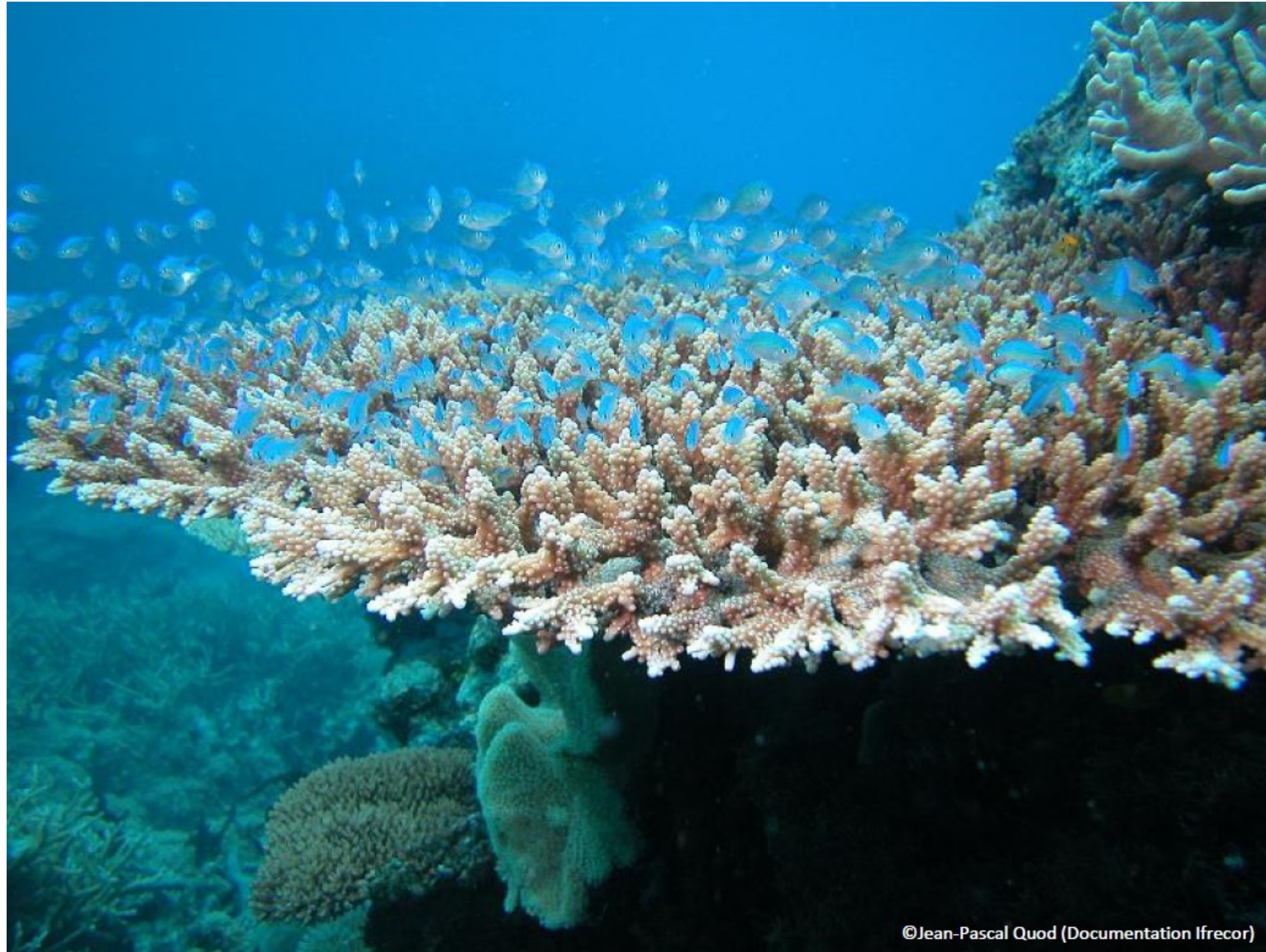
Préparer les figures avec les données à extraire (capture écran → .png), les mettre dans un dossier « figs » dans le répertoire de travail

```
> dat <- metaDigitise(dir = "./figs")
```

Extraire les données dans un fichier Excel

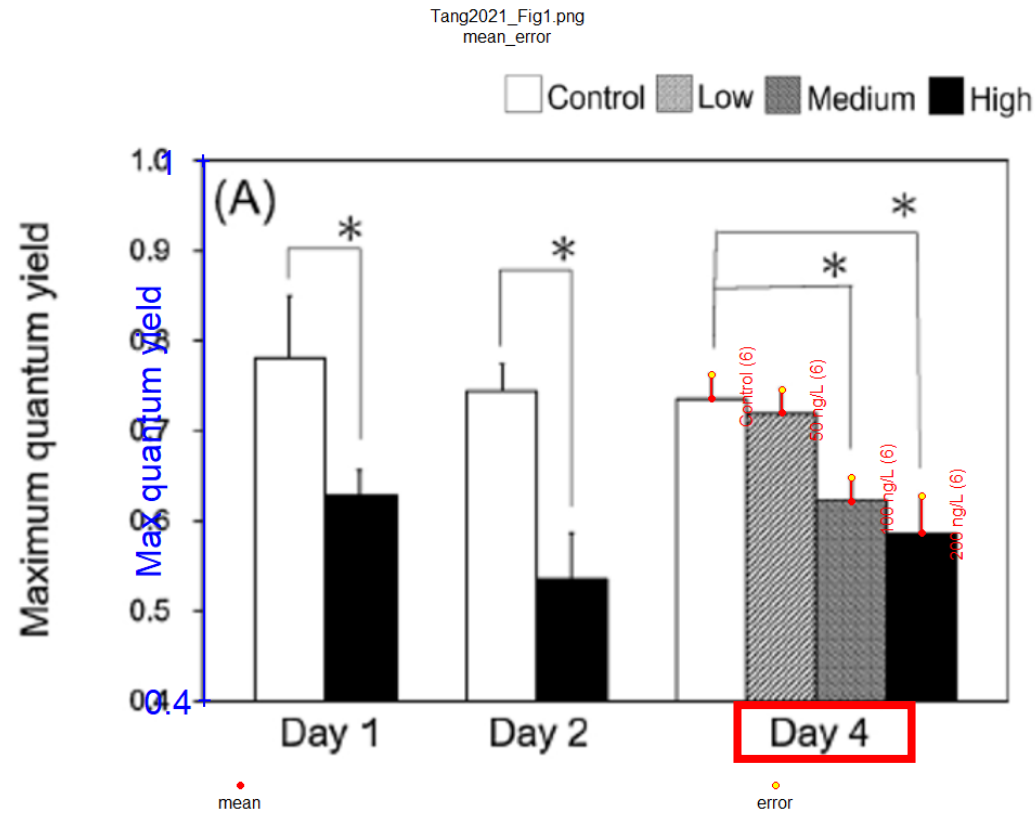
## 2 – Discussion





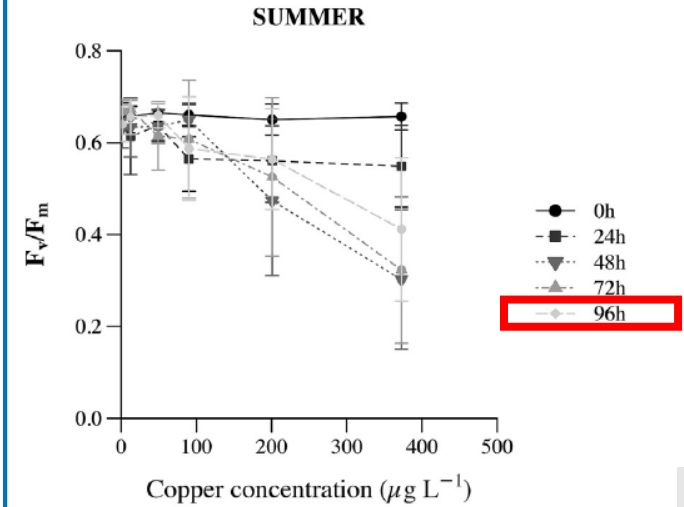
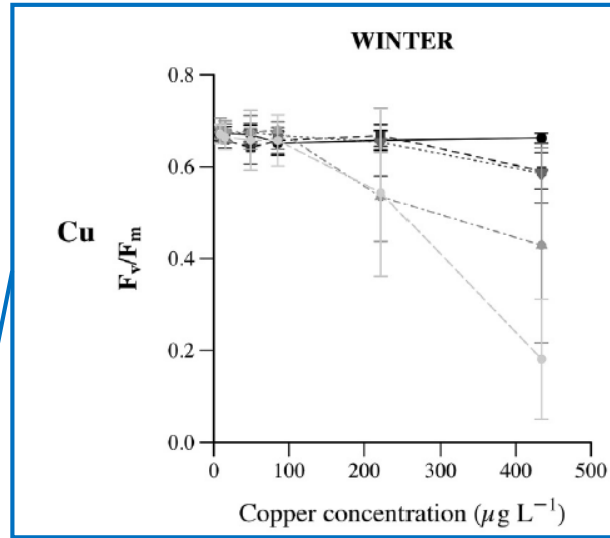
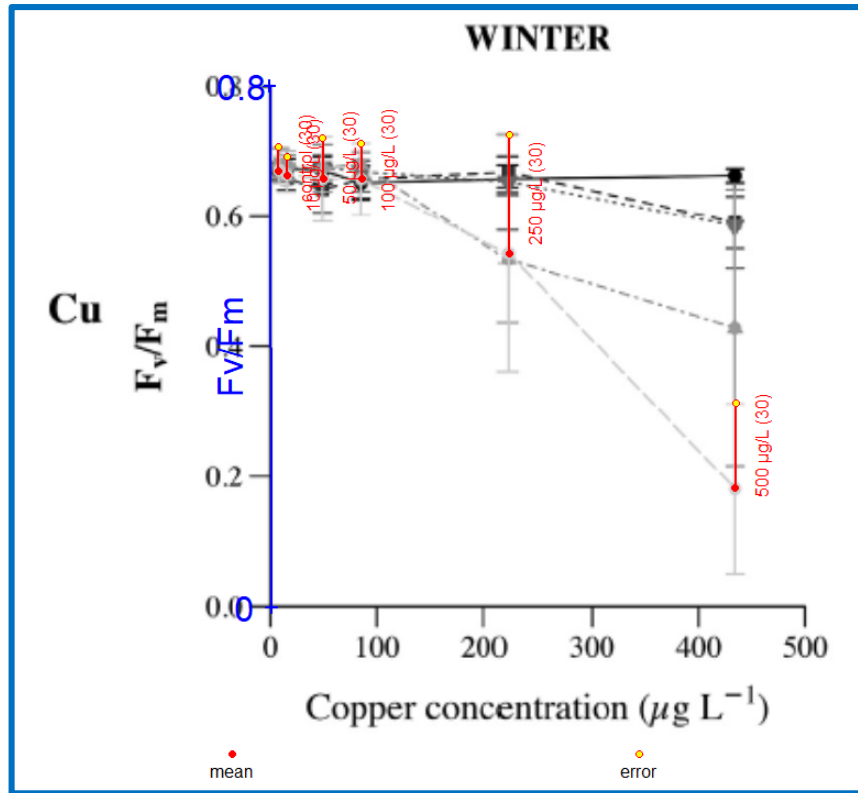
## Tang et al. 2021

Fig. 1. Differences in photophysiological parameters of symbiotic algae in coral exposed to Irgarol levels of 50 (low), 100 (Medium) and 200 (High) ng/L. The marker "\*\*\*" indicates a significant difference between the Irgarol-treated and control groups (paired *t*-test,  $p < 0.05$ ,  $N = 6$ ). The error bar indicates the value of the standard deviation.



N = 6  
SD

## Hédouin et al. 2016



N = 30  
SD

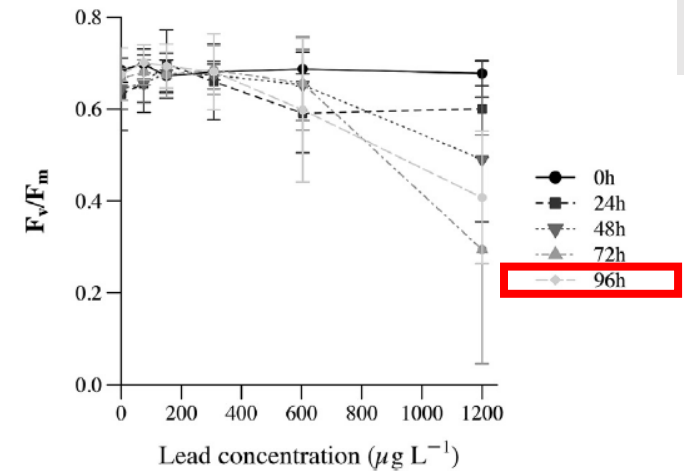
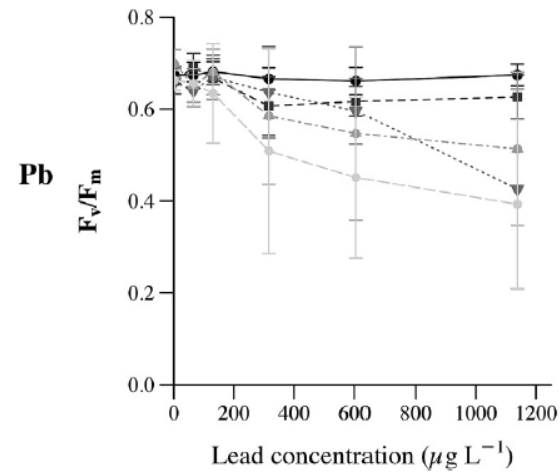


Fig. 8. Dark-adapted quantum yield ( $F_v/F_m$ , mean  $\pm$  standard deviation,  $n = 30$  nubbins per condition, -pooled data-) of *Pocillopora damicornis* corals exposed for 96 h to a range of dissolved Cu and Pb concentration in the summer and winter seasons.



## Kegler et al. 2015

Table 3. Summary of *P. verrucosa* responses.

	Dark respiration [mgO <sub>2</sub> h <sup>-1</sup> cm <sup>-2</sup> ]	Net photosynthesis [mgO <sub>2</sub> h <sup>-1</sup> cm <sup>-2</sup> ]	Gross photosynthesis [mgO <sub>2</sub> h <sup>-1</sup> cm <sup>-2</sup> ]	Maximum quantum yield 48 h [F <sub>v</sub> /F <sub>m</sub> ]	Maximum quantum yield 84 h [F <sub>v</sub> /F <sub>m</sub> ]	Tissue loss after 84 h [% loss]
Control	0.019 ± 0.005	0.008 ± 0.003	0.011 ± 0.003	0.71 ± 0.02	0.71 ± 0.02	-
High temperature	0.012 ± 0.003	0.003 ± 0.001	0.009 ± 0.003	0.74 ± 0.01	0.72 ± 0.01	-
Diesel	0.015 ± 0.001	0.006 ± 0.003	0.001 ± 0.003	0.71 ± 0.02	0.71 ± 0.02	-
LAS	-	-	-	0.73 ± 0.01	-	52.5 ± 30.15
Diesel + high temperature	0.023 ± 0.003	0.008 ± 0.002	0.014 ± 0.005	0.74 ± 0.01	0.71 ± 0.01	-
LAS + high temperature	-	-	-	0.63 ± 0.13	-	92.25 ± 7.26

N = 4  
SD