# Biodiversity knowledge synthesis: an introduction to meta-analyses and systematic reviews
# - Quantitative data extraction -

4/10/23 - Montpellier

Dakis-Yaoba Ouédraogo (PatriNat)
dakis-yaoba.ouedraogo@mnhn.fr

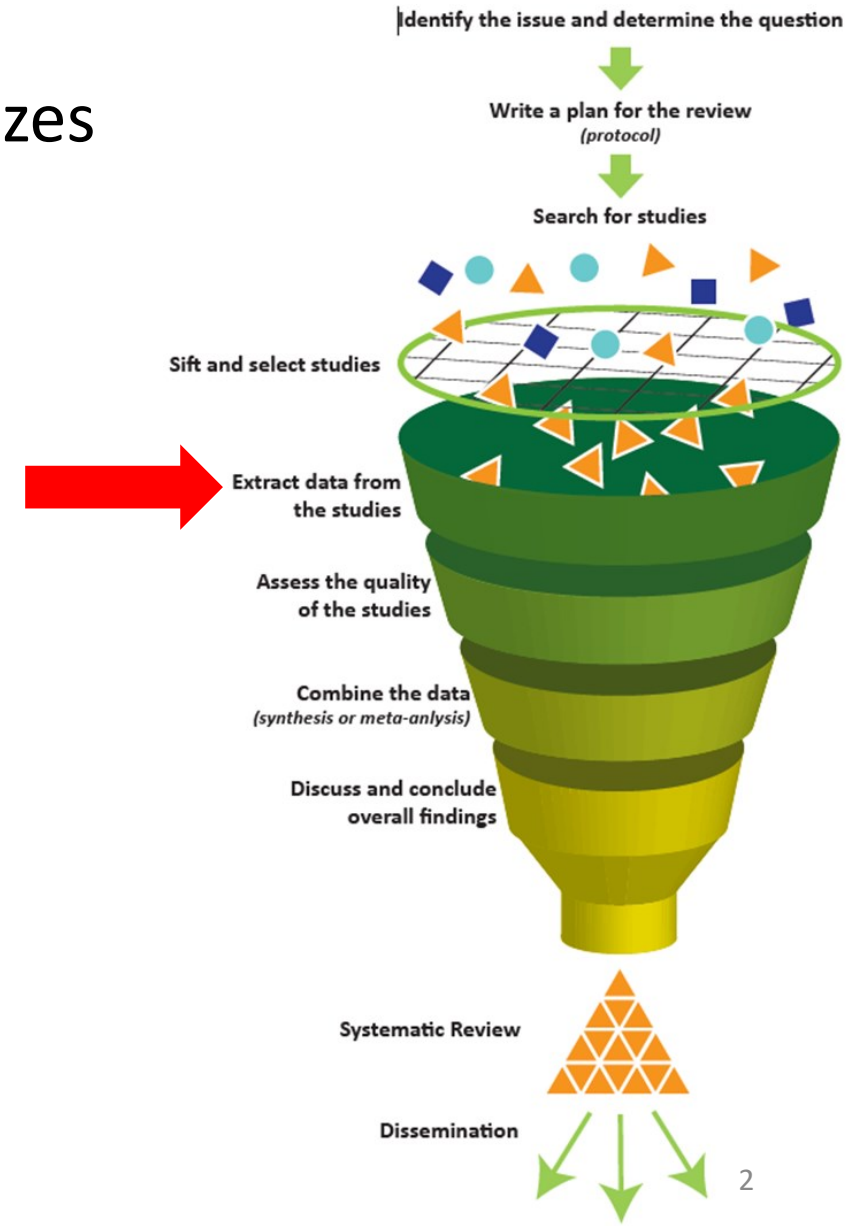
# Extraction of quantitative data

Extraction of the data needed to calculate effect sizes (e.g. mean, sample size, sd/se/95% CI)

+ extraction of variables that could explain the heterogeneity of effect sizes (*effect modifiers*)

Extraction from
- text
- table
- figure
- supp. mat.

+ calculations may be needed (keep a record)



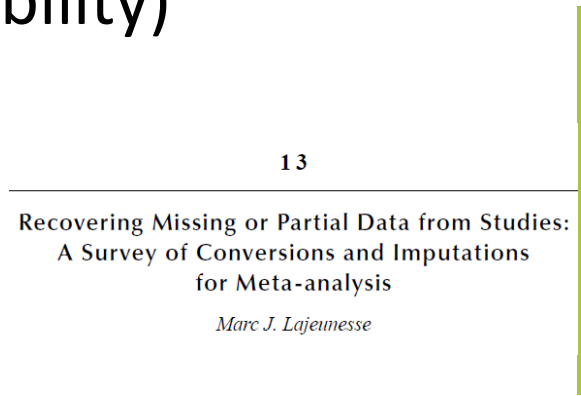
# Extraction of quantitative data

! Warning !
Data extraction is time-consuming: clearly define the extraction grid and the effect modifiers to be extracted

Importance of **testing** the extraction sheet on a sample of articles to check that it matches the content of the studies

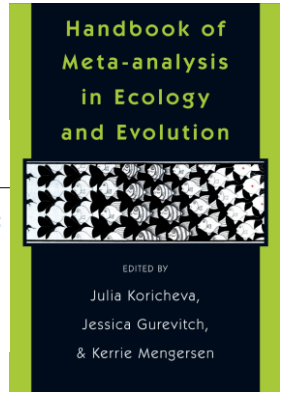
Document the work/decisions (transparency, repeatability)

Decide what to do in case of **missing information** ("missing data", contact the authors, imputation)


13

Recovering Missing or Partial Data from Studies: A Survey of Conversions and Imputations for Meta-analysis

*Marc J. Lajeunesse*



# Consistency check

To be sure that data extraction is objective / robust:

- Data extraction of each study carried out **independently** by 2 people

- if several people share the work, **check the consistency of data extraction** between people on a sample before starting the actual extraction (discuss any disagreements)

- if only 1 person, have someone to check a sample of the extracted data at the start of the work (discuss any disagreements)

# Example of extraction sheet

A study = a taxon × an exposure × an outcome

**Case study level**: ex. several concentrations of a chemical

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

# Example of extraction sheet

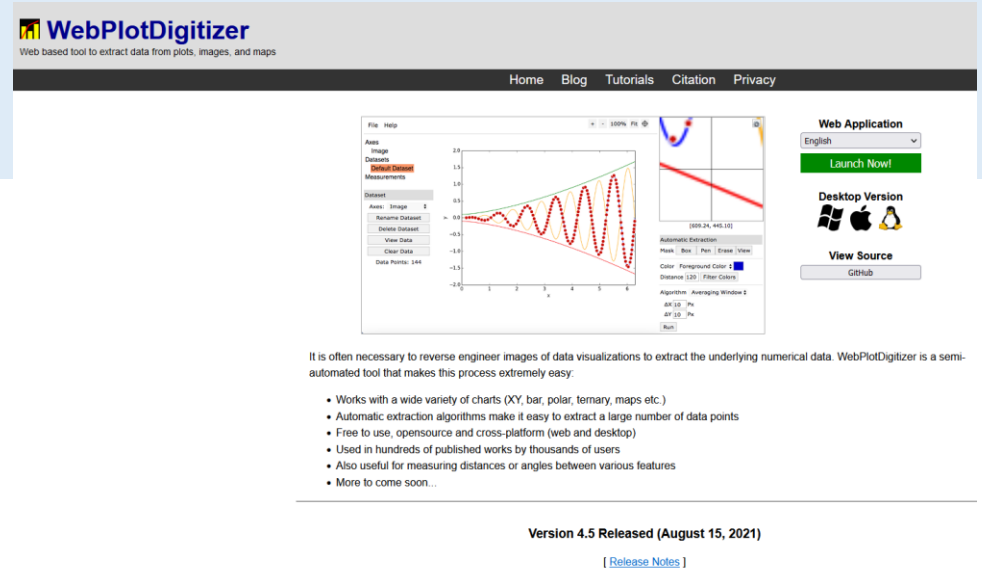
| Treatment_description | Control_description | Solvent | Concentration_nom | Concentration_eff | Duration | Measured_variable | Time_after |
|---|---|---|---|---|---|---|---|
| Diuron | Unfiltered oceanic seawater | No | 1 µg/L | 0.91 µg/L | 53 days | Symbiodinium density / total protein | NA |
| Diuron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

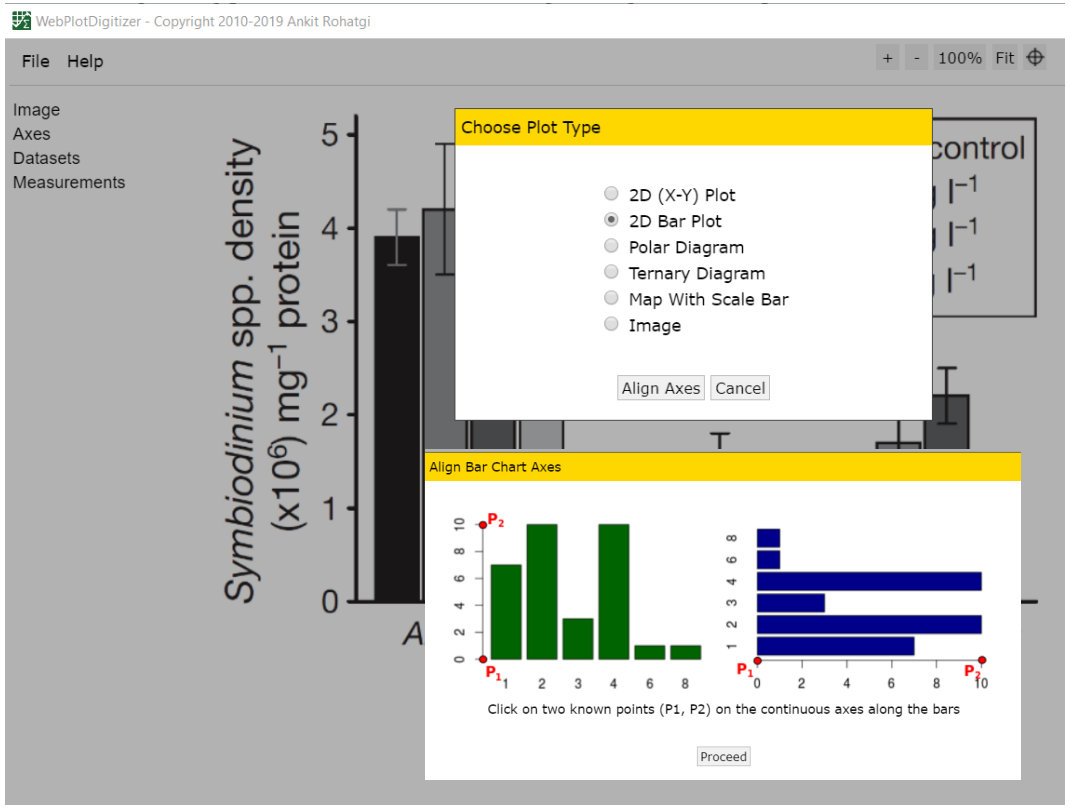
| Metaanalyse_data | unit | ID_experiment | ID_case | ID_common_control | N_c | Mean_c | Type_variation_c | Variation_c | N_t | Mean_t | Type_variation_t | Variation_t |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| OK (Fig3, SE, n=6) | x 10^6 / mg protein | 1 | 3 | 1 | 6 | 4.2 | sd | 1.714642819 | 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 |

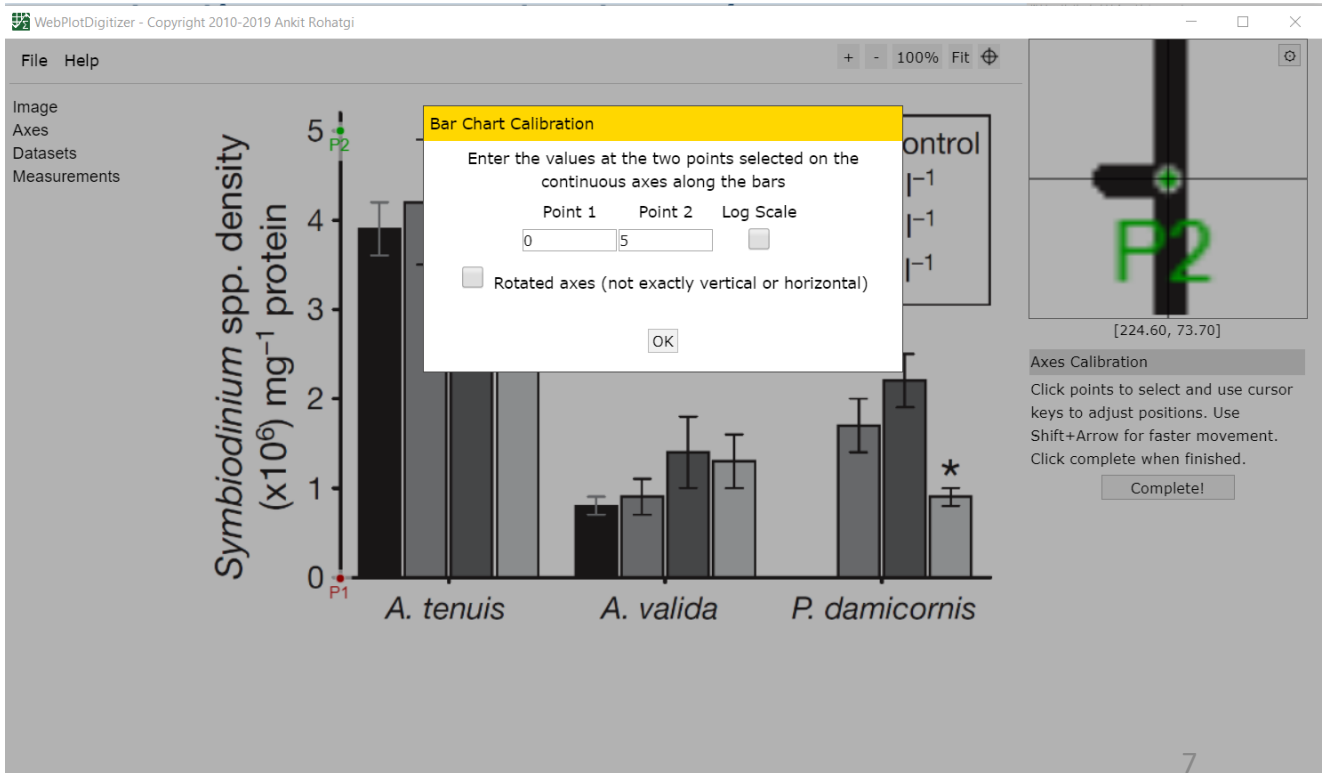
| Method_extraction | Source | Comment_extract | Name_data_extractio |
|---|---|---|---|
| Figure | Figure 3 | NA | DYO |
| Figure | Figure 3 | NA | DYO |
| Figure | Figure 3 | NA | DYO |
| Figure | Figure 3 | NA | DYO |
| Figure | Figure 3 | NA | DYO |
| Figure | Figure 3 | NA | DYO |

# Tools for extracting data from figures

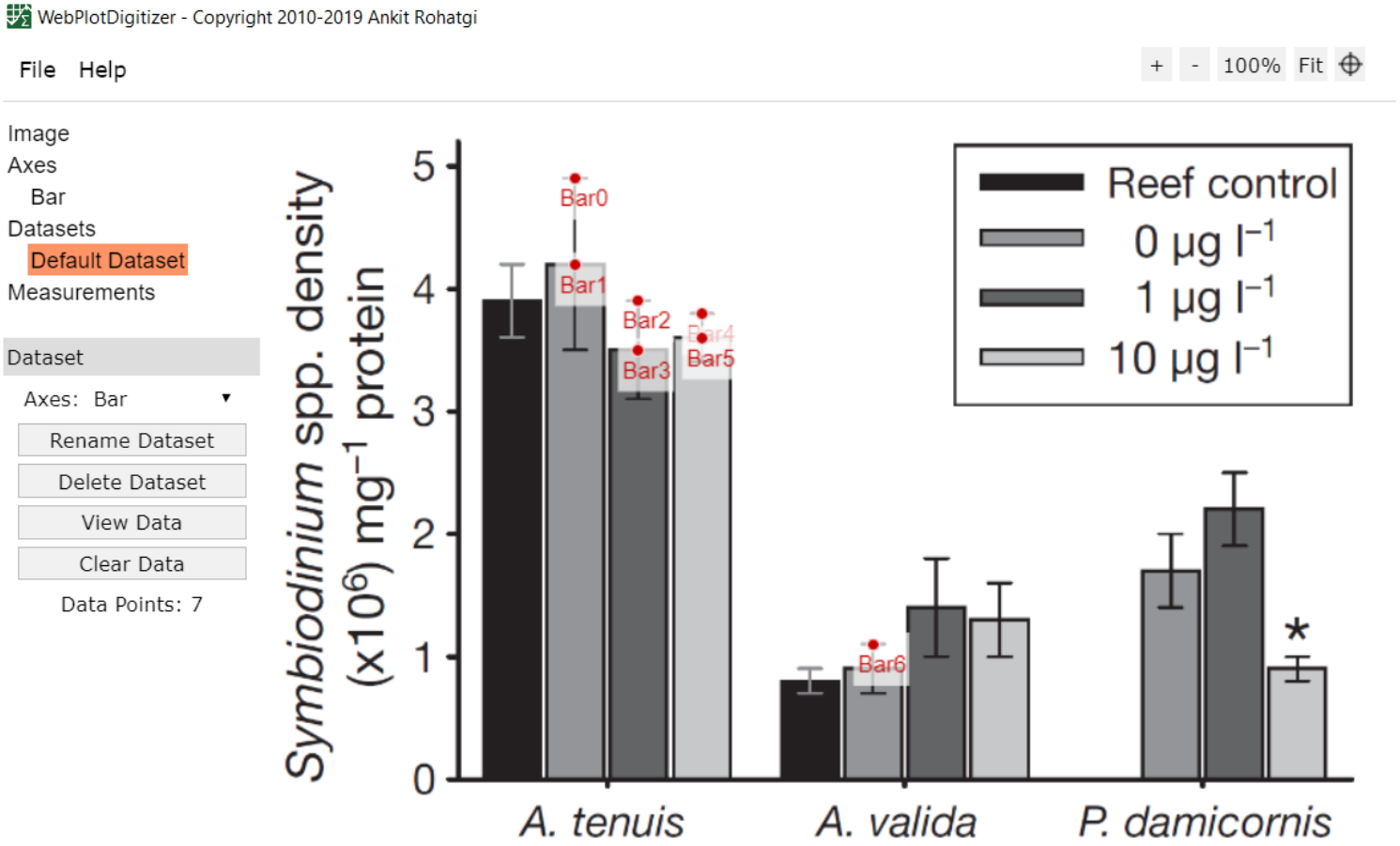
https://automeris.io/WebPlotDigitizer/

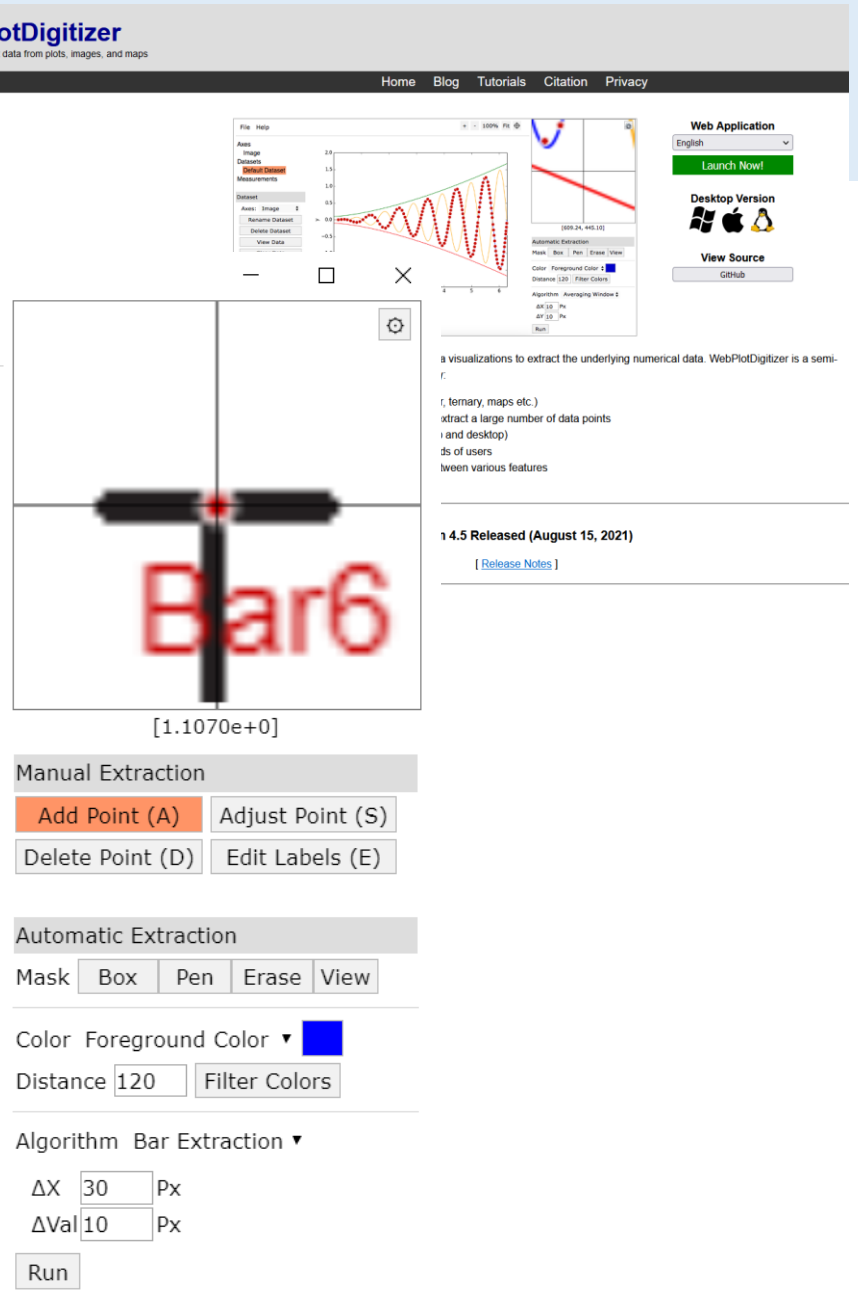


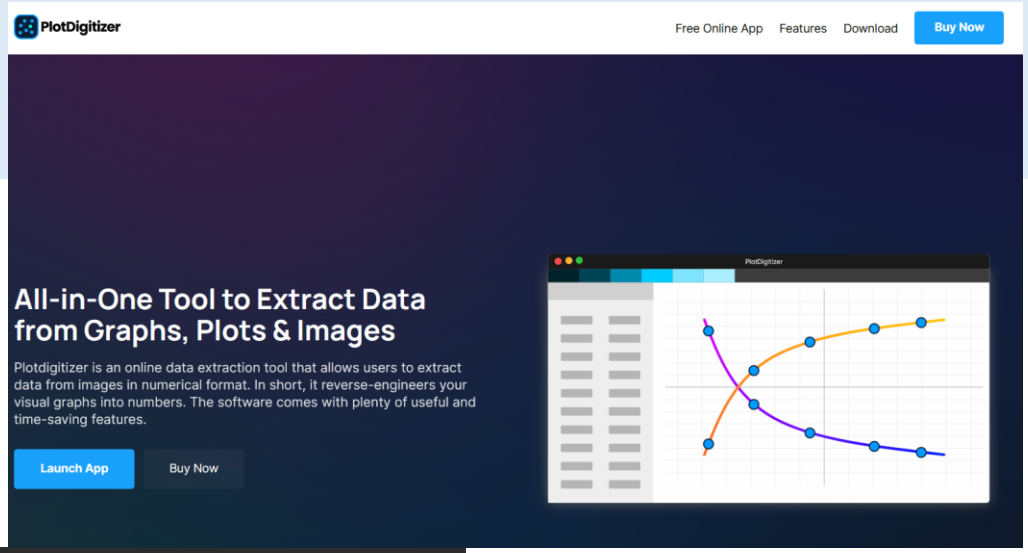
# Tools for extracting data from figures

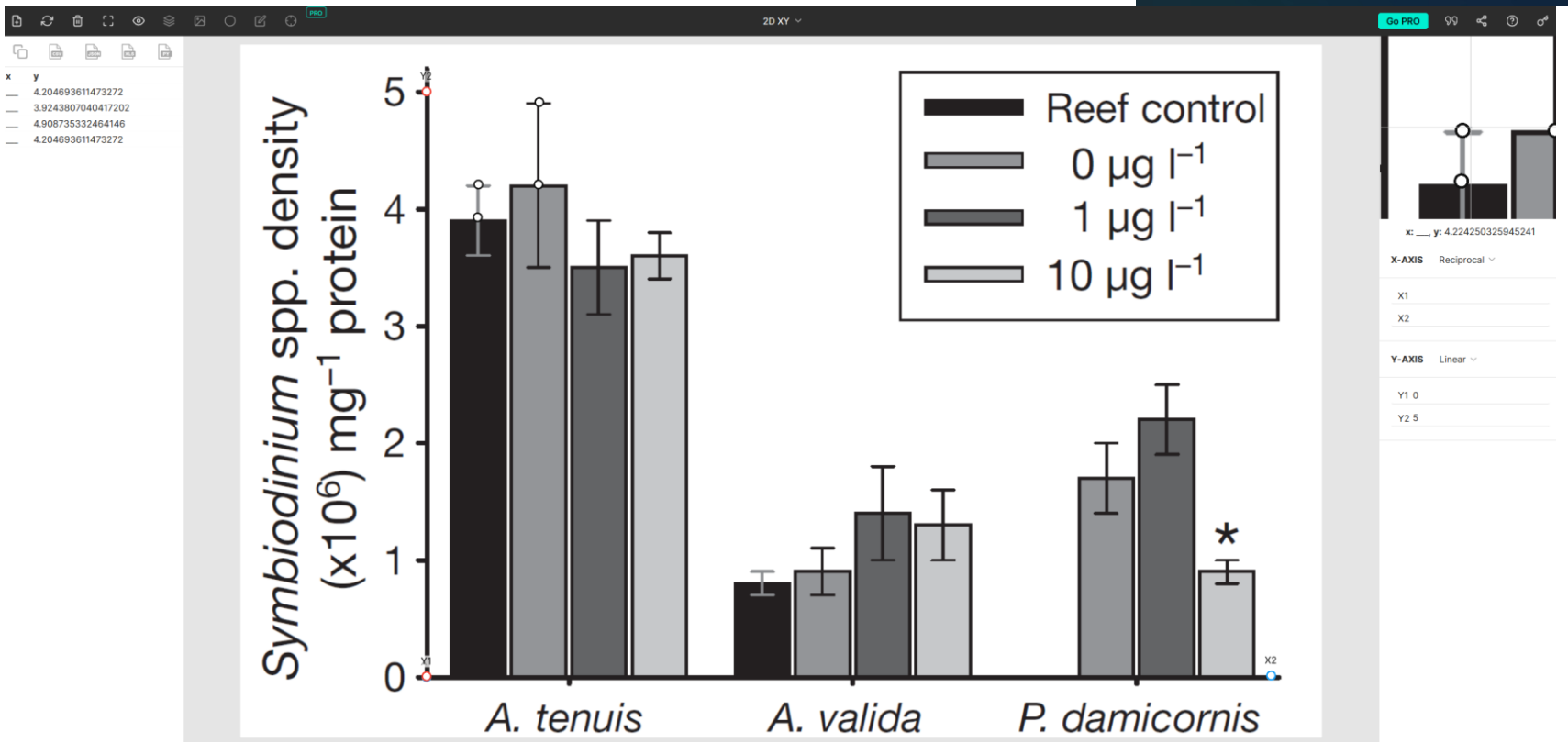


# Tools for extracting data from figures

https://plotdigitizer.com/



# Tools for extracting data from figures: metaDigitise

Received: 13 July 2018 | Accepted: 12 October 2018

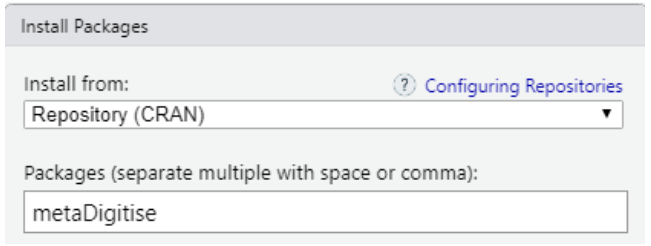
DOI: 10.1111/2041-210X.13118

APPLICATION

Methods in Ecology and Evolution | BRITISH ECOLOGICAL SOCIETY

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

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

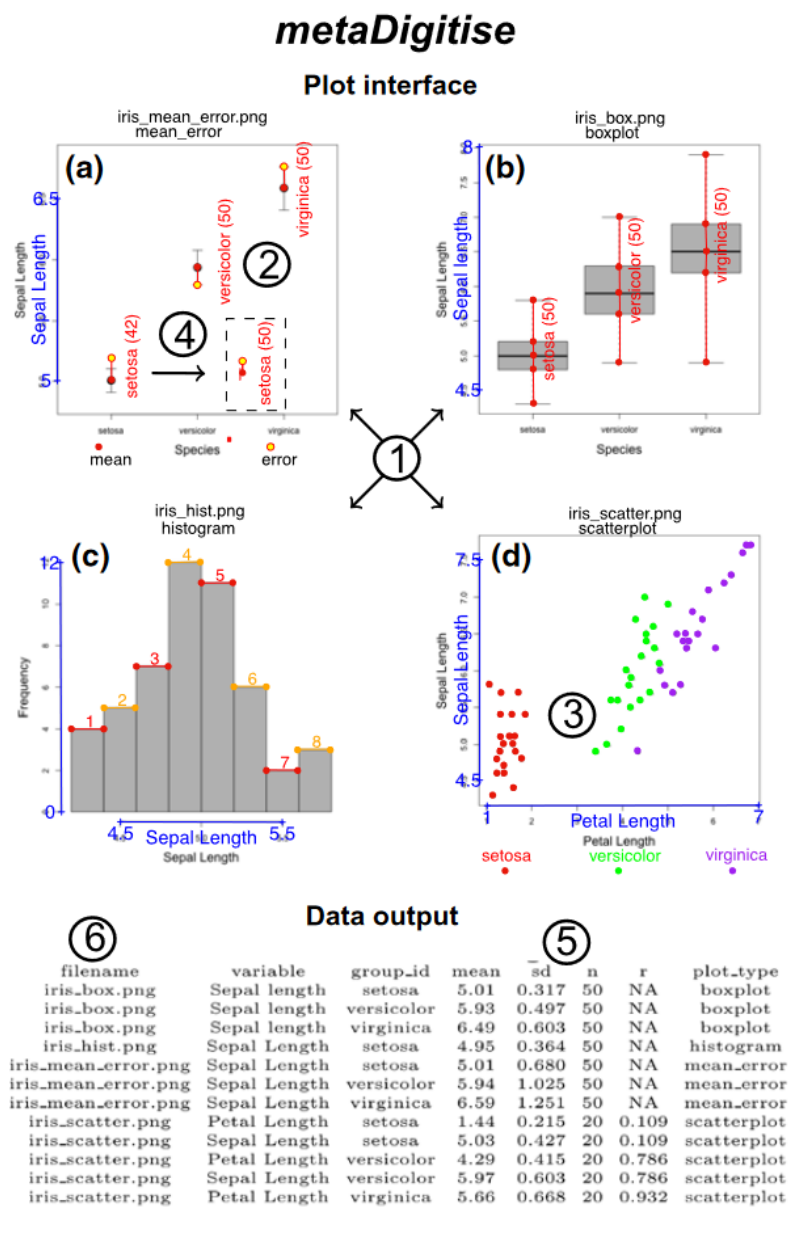



(+) possible to save, trace and modify data extraction

(-) no zoom

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

# Tools for extracting data from figures: metaDigitise




## 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

# TD

**Question**: estimate the effects of chemicals on the photosynthetic performance **(*maximum quantum yield*, Fv/Fm)** of tropical reef-building corals

Sample of 3 articles
A **study** = a taxon × an exposure × an outcome
A **case study** = one tested concentration-duration
**Effect size** = standardized mean difference

<u>If monitoring over time, extraction of the longest duration of exposure</u>

Extract data using the metaDigitise package

# TD

**Question**: estimate the effects of chemicals on the photosynthetic performance **(*maximum quantum yield*, Fv/Fm)** of tropical reef-building corals
Sample of 3 articles
A **study** = a taxon × an exposure × an outcome
A **case study** = one tested concentration-duration
**Effect size** = standardized mean difference
<u>If monitoring over time, extraction of the longest duration of exposure</u>

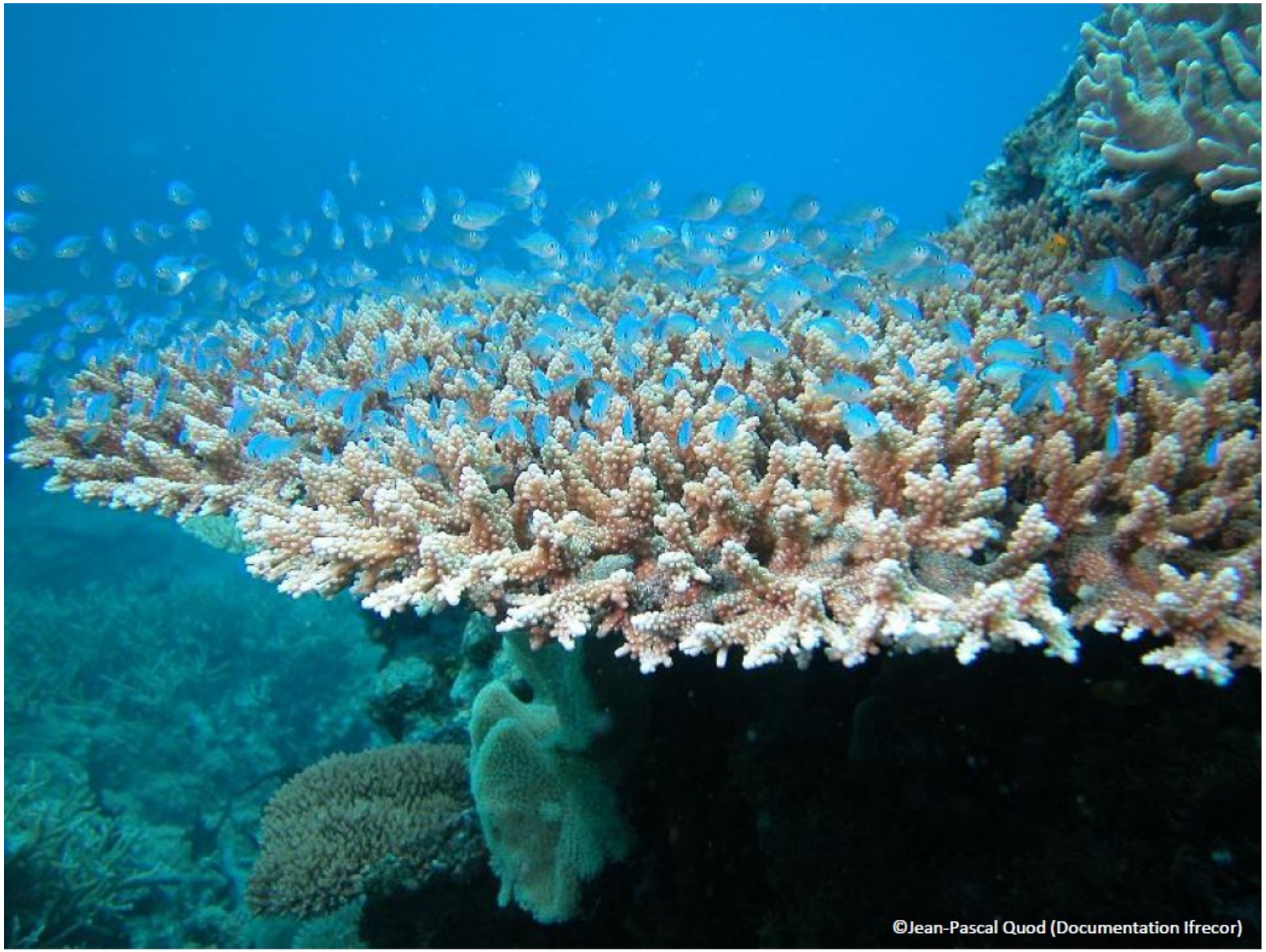
1 – Extract data
Install et load metaDigitise package; set working directory
Prepare the figure files (screen capture → .png), put them in a "figs" folder in the working directory

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

Extract data into an Excel file

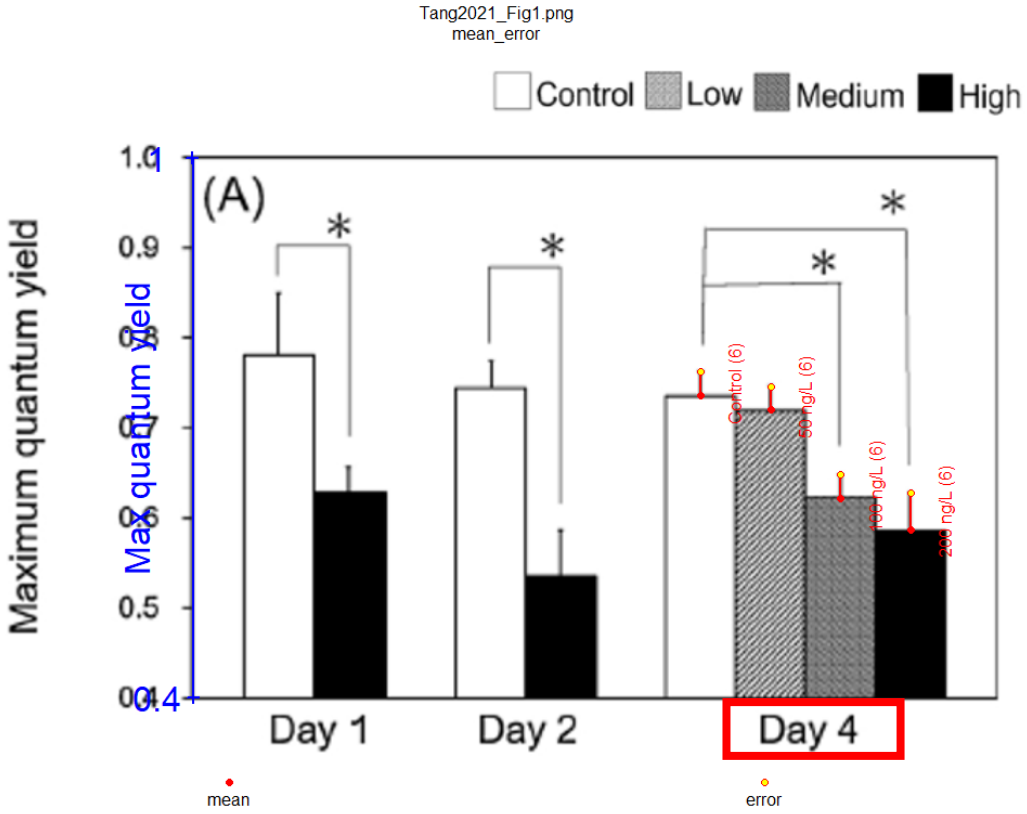
2 – Discussion

# TD



# Tang et al. 2021

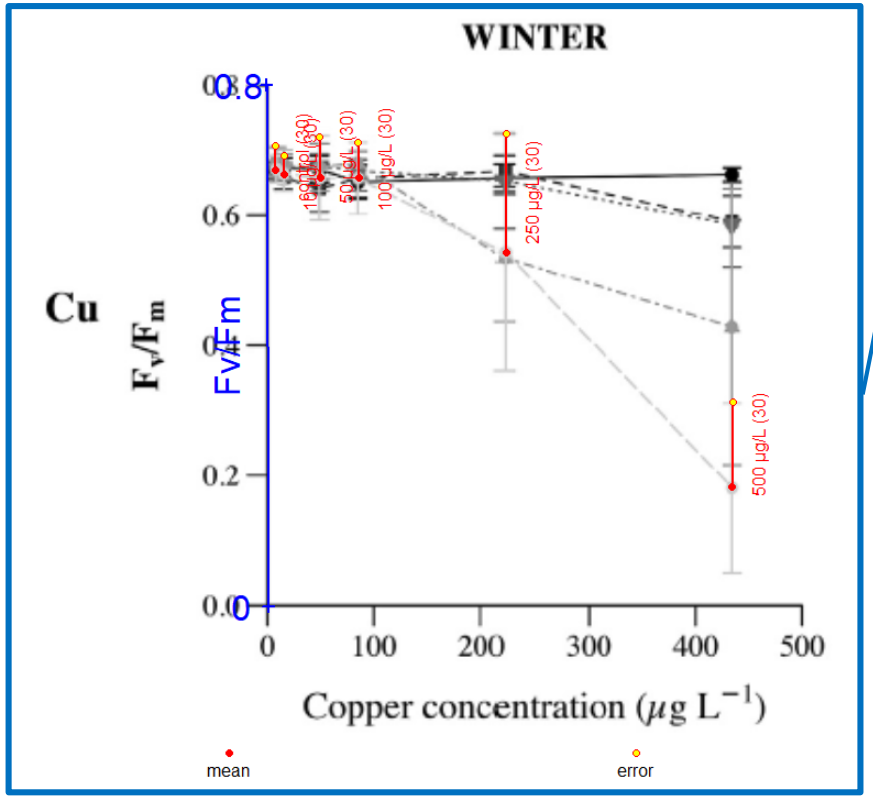
**Fig. 1.** Differences in photophysiological parameters of symbiotic algae in coral exposed to Irgarol levels of 50 (low), 100 (Medium) and 200 (High) ng/L. The marker "*" indicates a significant difference between the Irgarol-treated and control groups (paired *t*-test, $p < 0.05$, $N = 6$). The error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

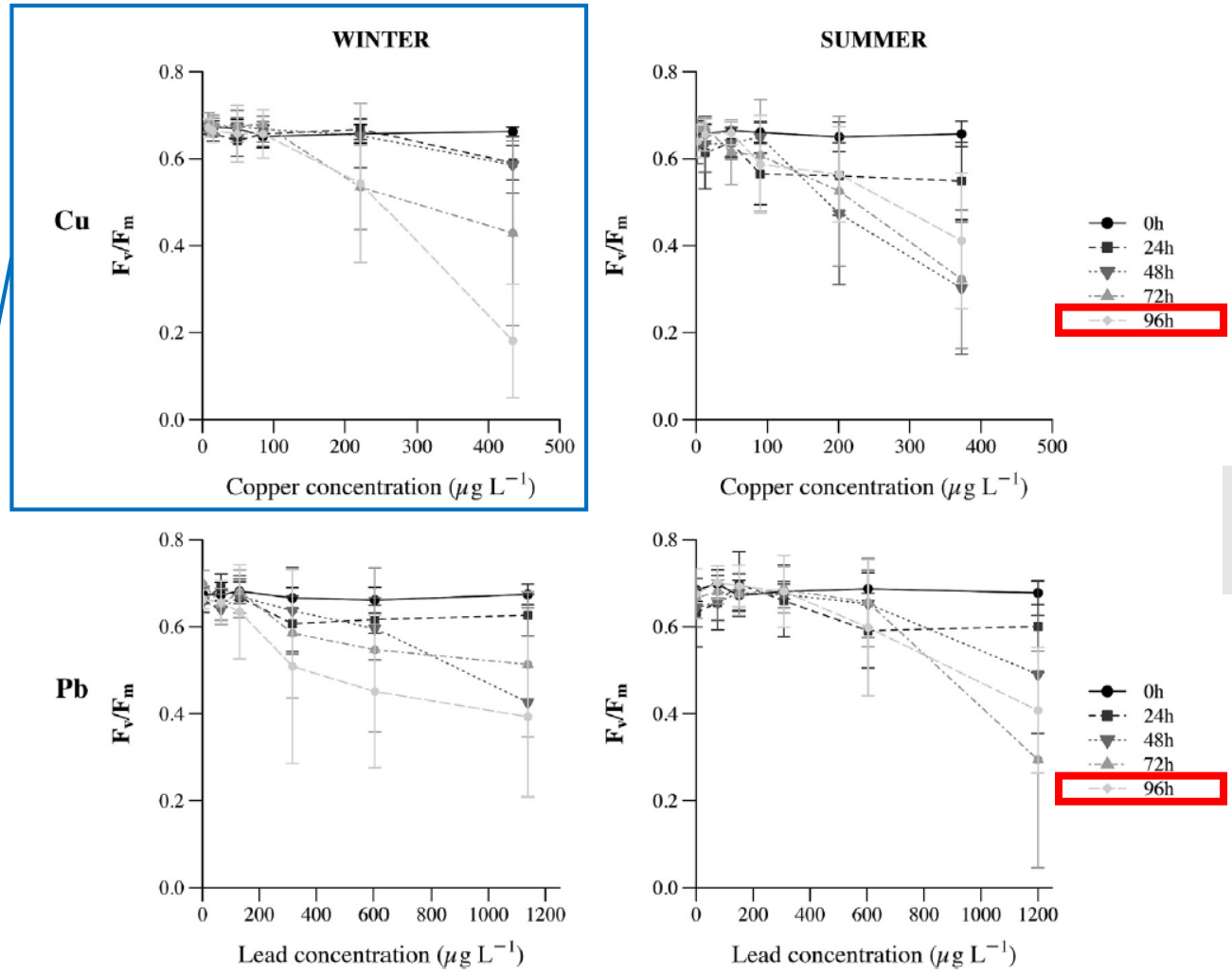



N = 6
SD

# TD

## Hédouin et al. 2016




N = 30
SD

**Fig. 8.** Dark-adapted quantum yield ($F_v/F_m$, mean ± 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.

# TD

## Kegler et al. 2015

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

| | Dark respiration [$mgO_2\ h^{-1}\ cm^{-2}$] | Net photosynthesis [$mgO_2\ h^{-1}\ cm^{-2}$] | Gross photosynthesis [$mgO_2\ h^{-1}\ cm^{-2}$] | Maximum quantum yield 48 h [$F_v/F_m$] | Maximum quantum yield 84 h [$F_v/F_m$] | 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