Diatom metabarcoding for biomonitoring : 3rd part

F. Rimet

Reference library, proof of concept, impact of biovolumes

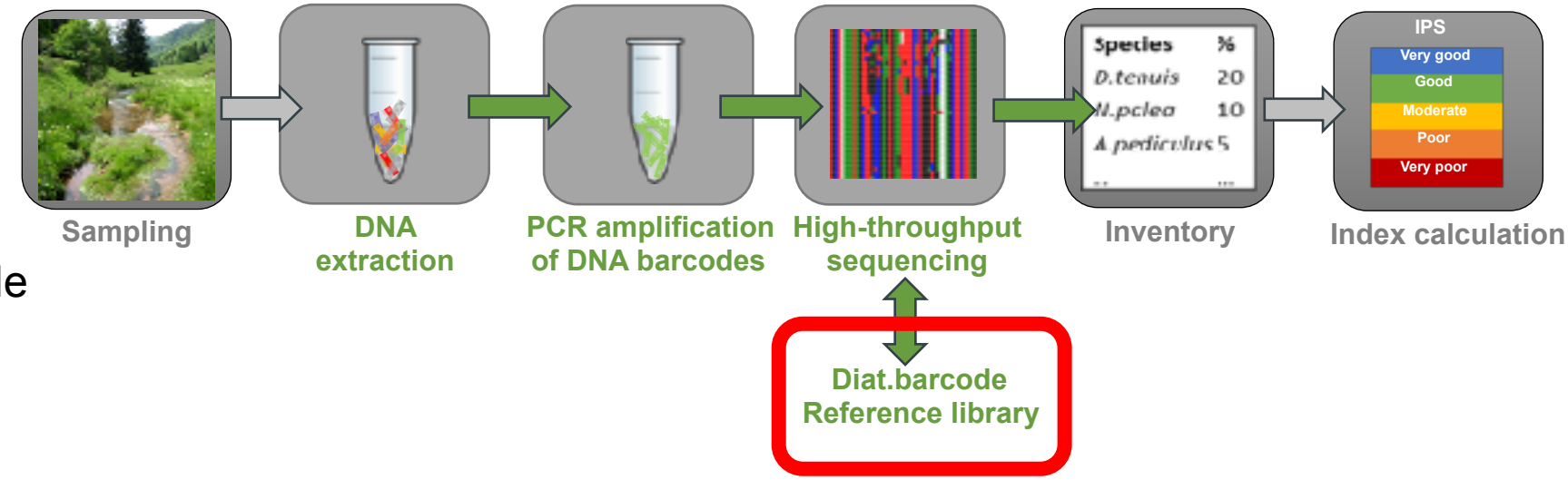
BIOLAWEB

Schedule

1- reference library 

2- 1st proofs of concept

3- Impact of biovolumes



❖**Schedule**

❖Presentation of Diat.barcode

❖Curation procedure

- ❖ Procedure
- ❖ Who? When?
- ❖ Standardisation

❖Completness of Diat.barcode

- ❖ In mainland France
- ❖ Particular cases

❖Divergences between classical and integrative taxonomy

- ❖ Divergence at specific level
- ❖ Divergence at deeper nodes

_01

Presentation of Diat.barcode

- Why a reference barcoding library for diatoms?
- Website presentation

Why maintaining a reference library for diatoms?

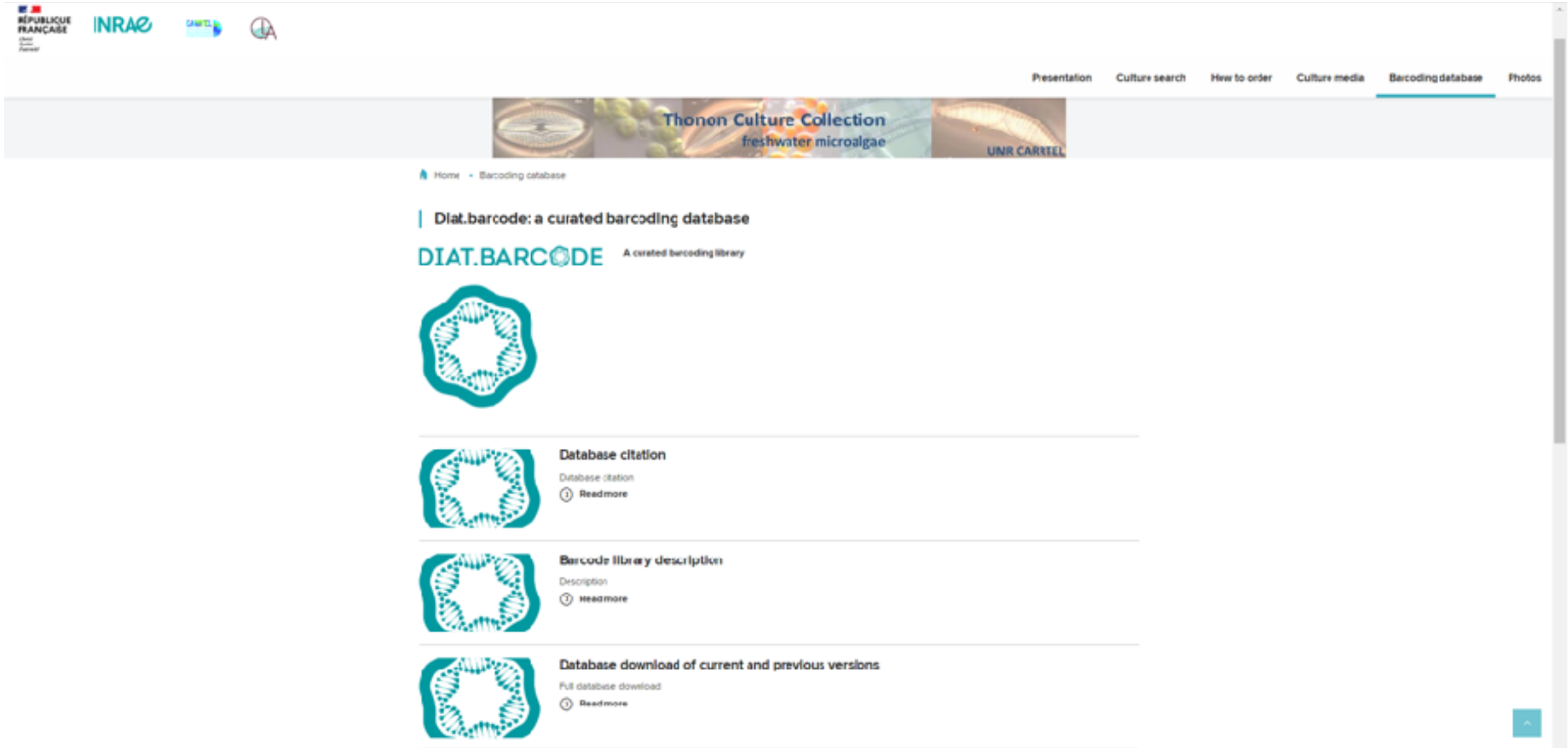
- **Some reference libraries exist:**
  - **Generalists libraries:**
    - **PR2 :** Guillou, L. et al. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. Nucleic Acids Res. 41, D597–D604 (2013).
    - **SILVA :** Quast, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 41, D590–D596 (2013).
    - **BOLD:** Barcode of life database
  - **Specialists libraries:**
    - **Phytool: phytoplankton :** Canino, A., A. Bouchez, C. Laplace-Treyture, I. Domaizon, & F. Rimet, 2021. Phytool, a ShinyApp to homogenise taxonomy of freshwater microalgae from DNA barcodes and microscopic observations. Metabarcoding and Metagenomics Pensoft Publishers 5: e74096. 16S, 23S
    - **PFR2: planktonic foraminifers:** Morard, R. et al. PFR2: a curated database of planktonic foraminifera 18S ribosomal DNA as a resource for studies of plankton ecology, biogeography and evolution. Mol. Ecol. Res. 15, 1472–1485 (2015).
    - **EukRef-Ciliophora: ciliates :** Boscaro, V. et al. EukRef-Ciliophora: A manually curated, phylogeny-based database of small subunit rRNA gene sequences of ciliates. Env. Microbiol. 20, 2218–2230 (2018).
    - **Dinoref: Dinophyta:** Mordret, S. et al. dinoref: A curated dinoflagellate (Dinophyceae) reference database for the 18S rRNA gene. Mol. Ecol. Res. 18, 974–987 (2018)

Why maintaining a reference library for diatoms?

- For diatoms:
  - SILVA, Initiative EukRef and PR2 -> can be used for 18s
  - Diat.barcode:
    - Traceability of data
    - Photos
    - If no photos: reference to the publication or the website where photos can be found
    - Demanding curation procedure (more than the above mentioned libraries)
    - Many unpublished data are made available (TCC, UK-barcoding project …)

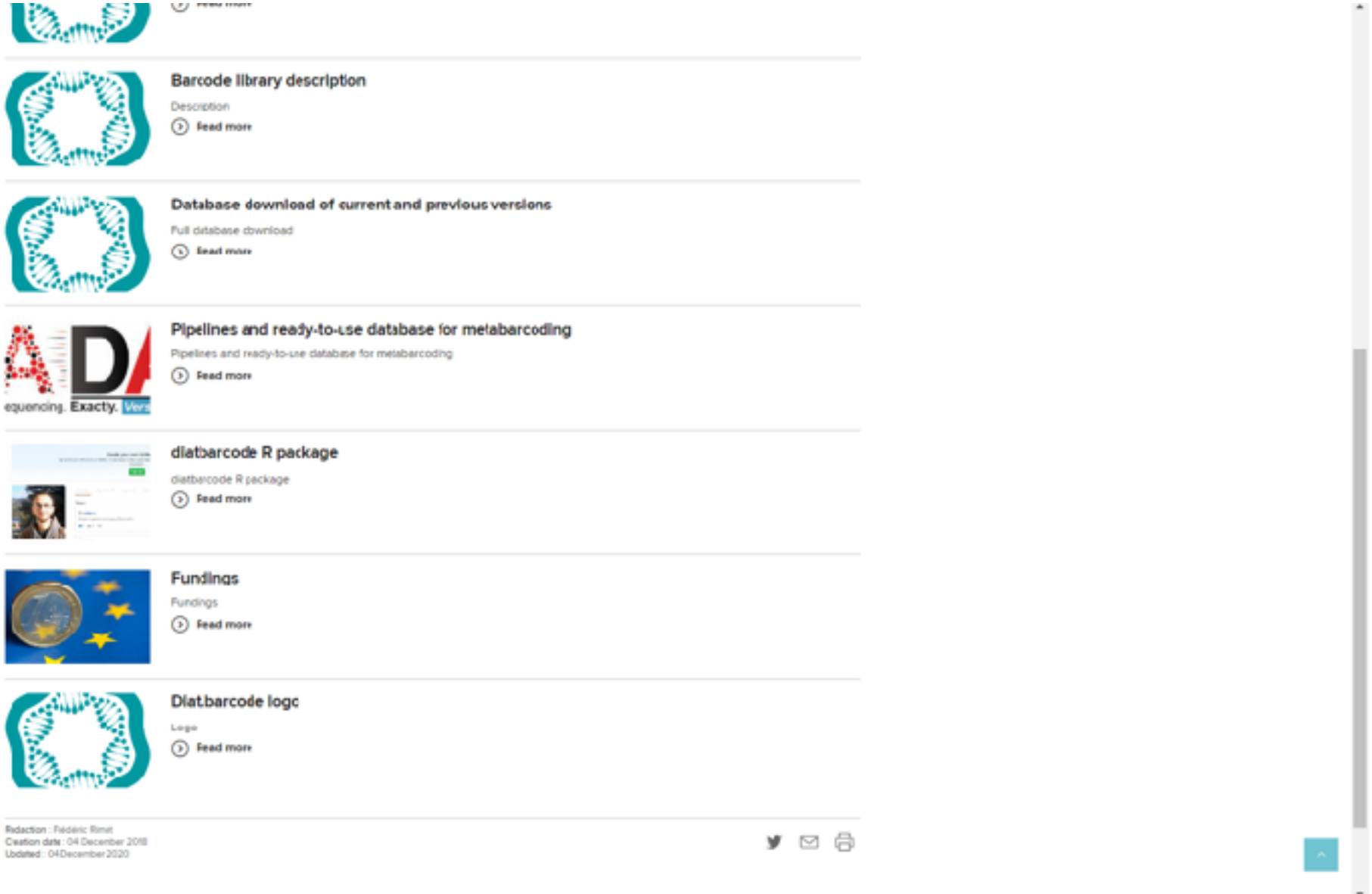
Diat.barcode website

https://www6.inrae.fr/carrtel-collection_eng/Barcoding-database



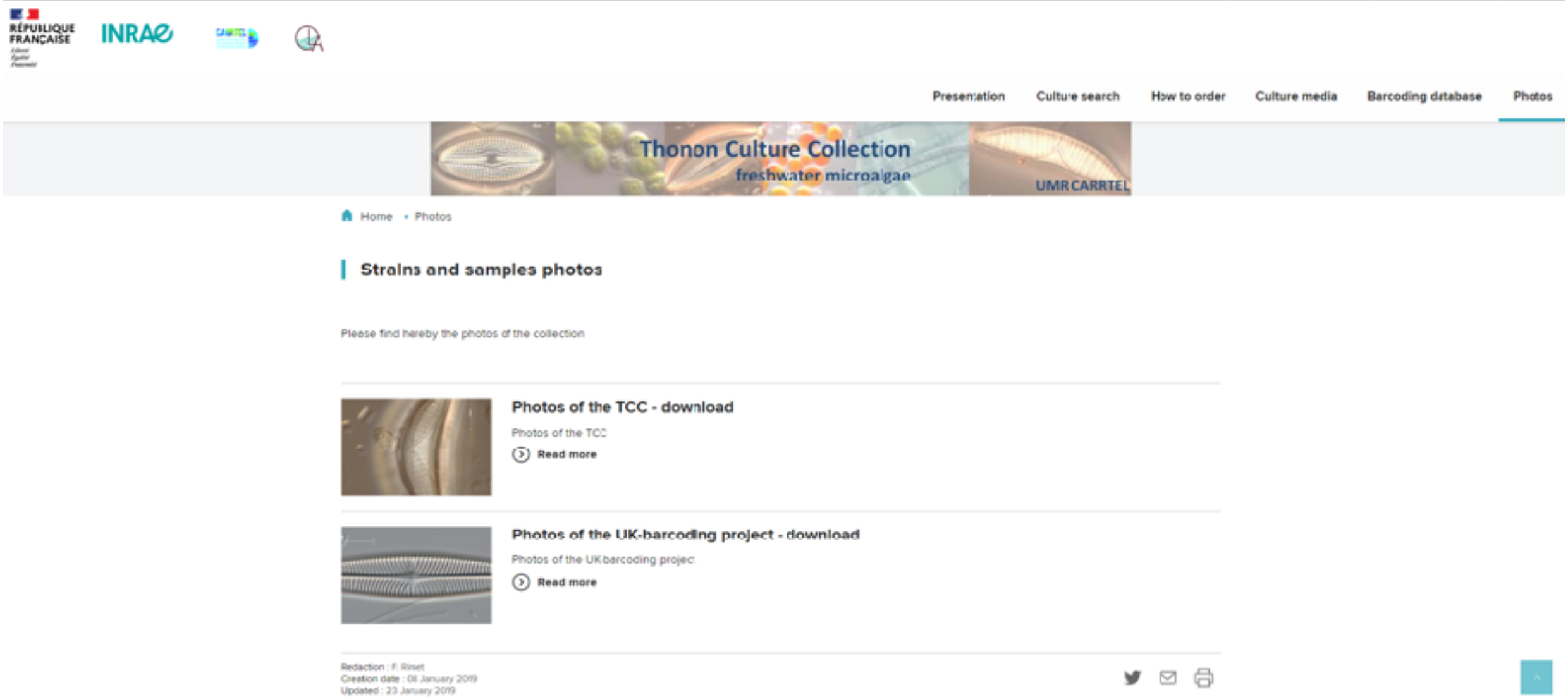
Diat.barcode website

https://www6.inrae.fr/carrtel-collection_eng/Barcoding-database



Diat.barcode website

❖All photos of the strains are downloadable



Diat.barcode download

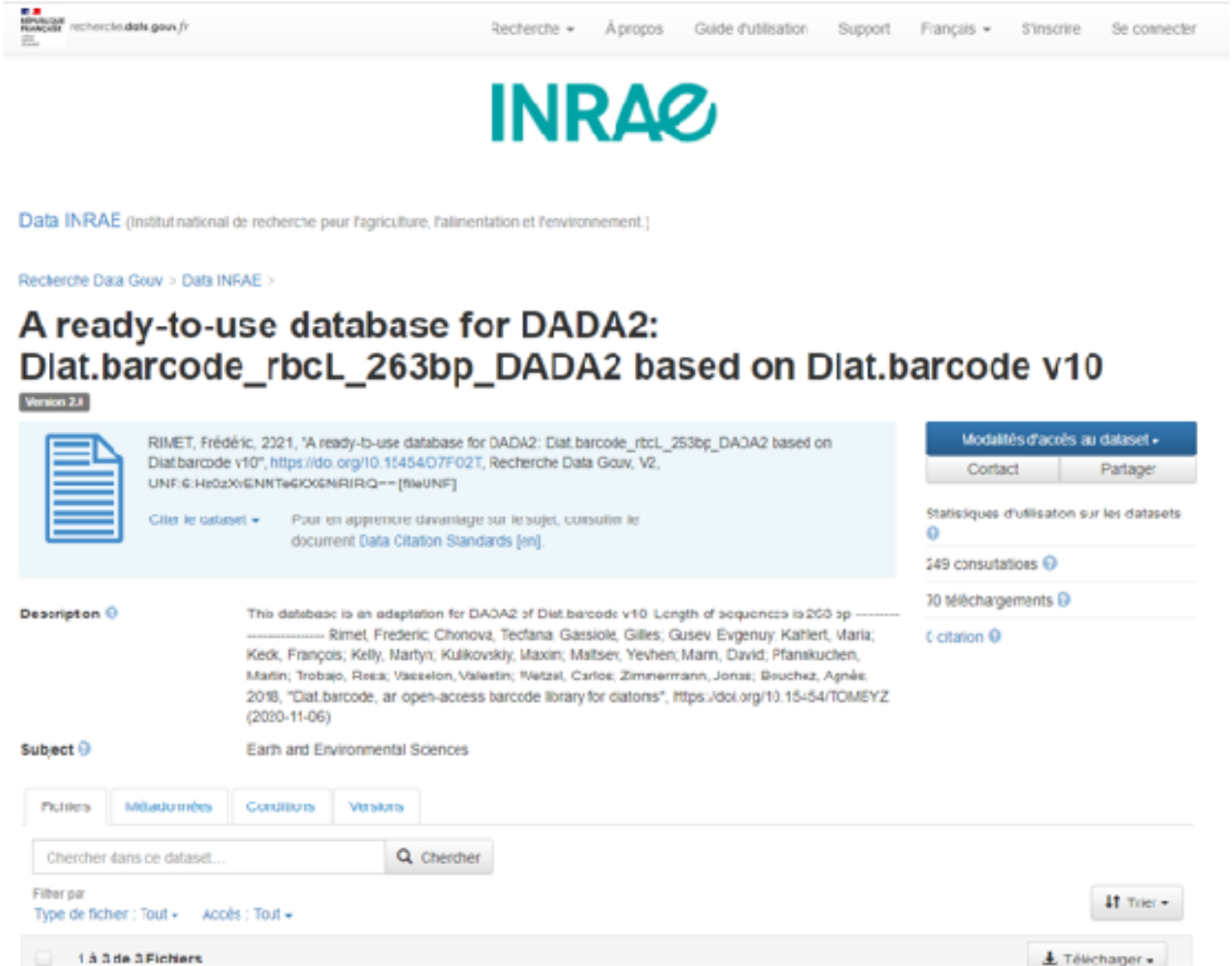
❖open access on the french research dataserve platform

All versions are available from v1 to v11

https://entrepot.recherche.data.gouv.fr/dataset.xhtml?persistentId=doi:10.15454/TOMBYZ

Adaptation of Diat.barcode for metabarcoding

❖open access on the french research dataserve platform

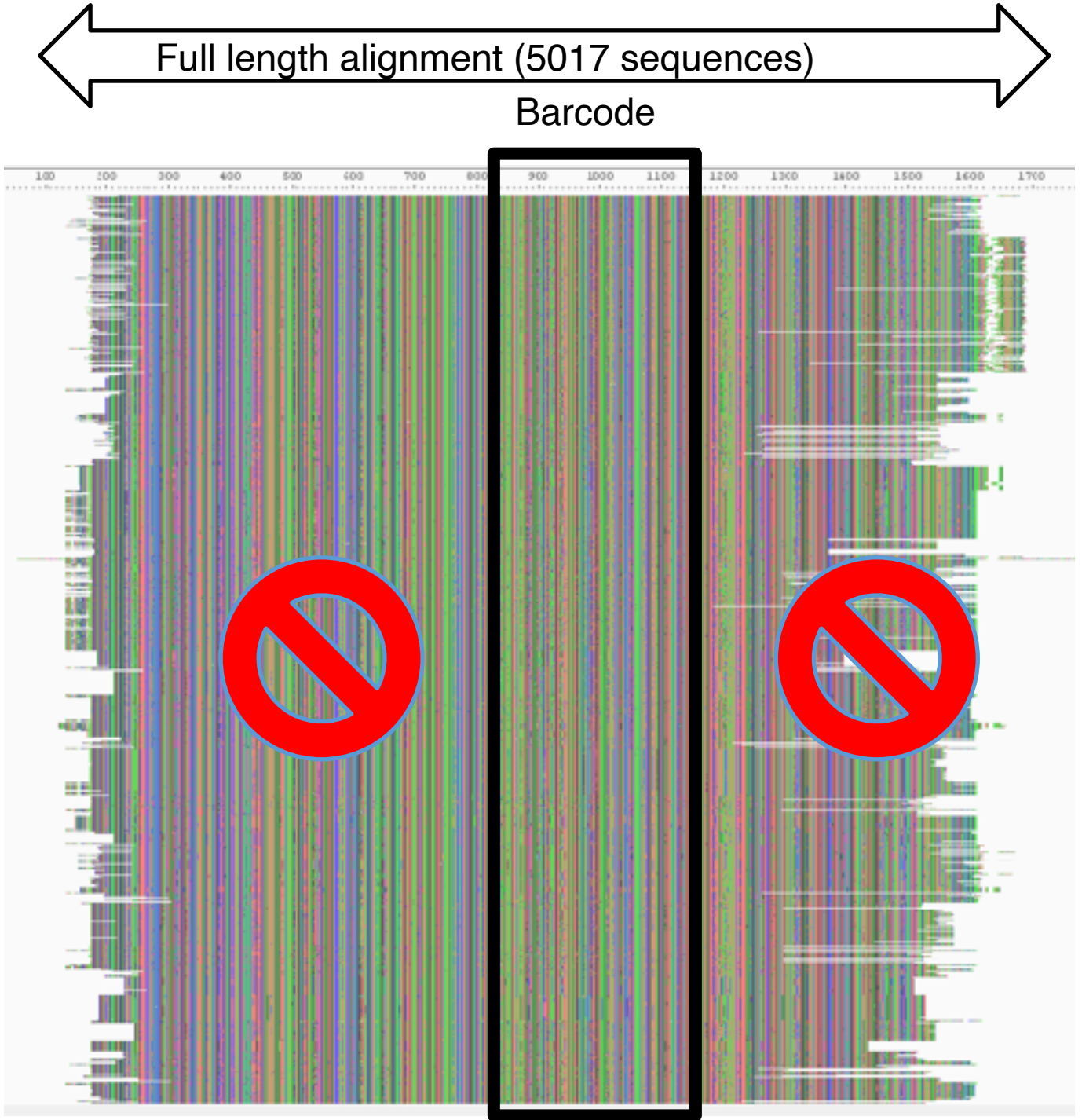


For DADA2:
https://entrepot.recherche.data.gouv.fr/dataset.xhtml?persistentId=doi:10.15454/D7FO2T

For MOTHUR:
https://entrepot.recherche.data.gouv.fr/dataset.xhtml?persistentId=doi:10.15454/V53JZV

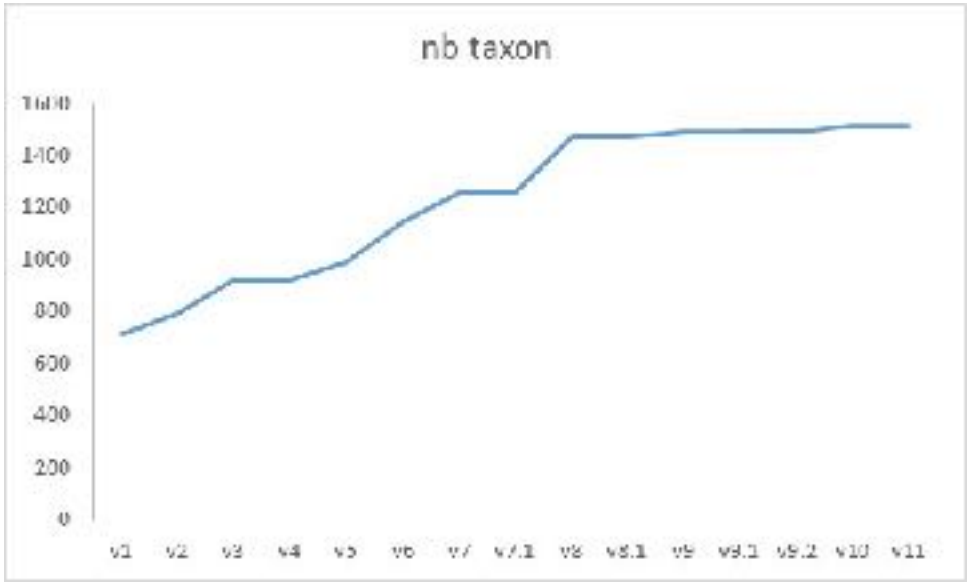
Why do we need an adaptation?

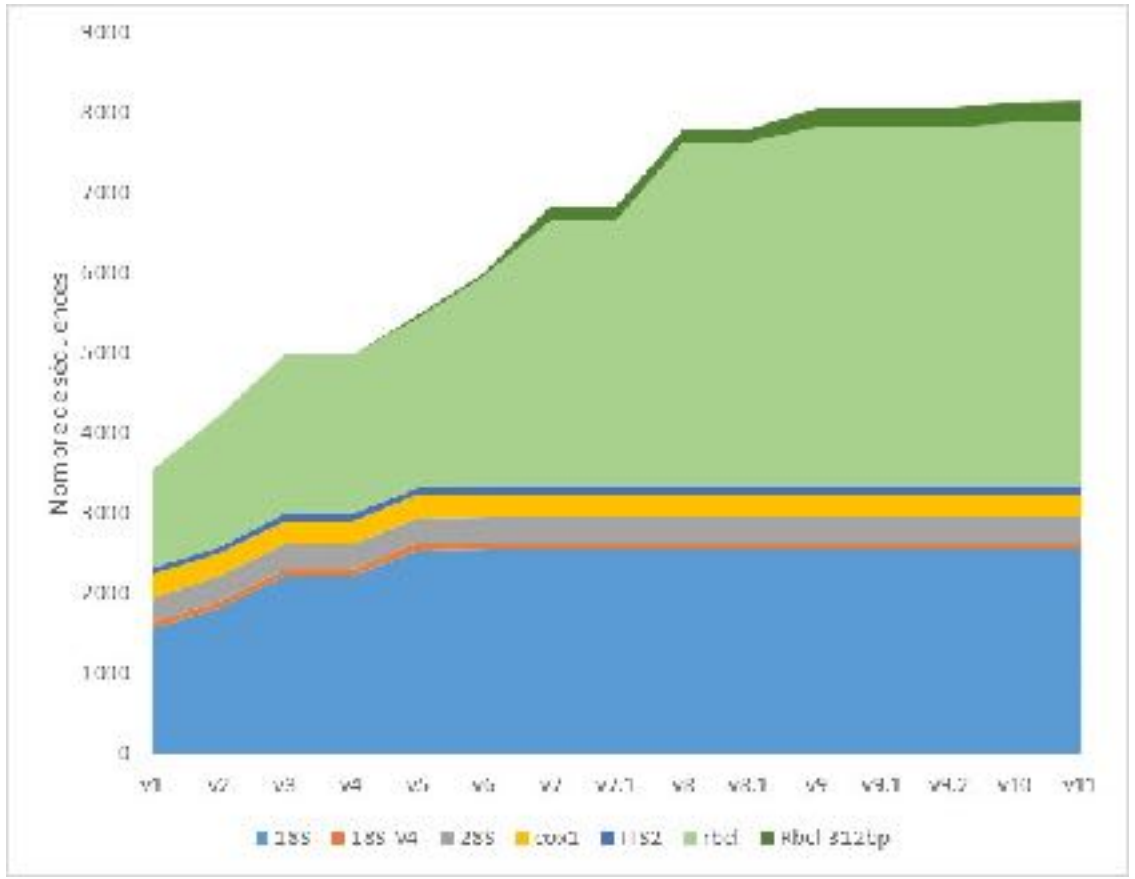
> 1500 bp -> 263 bp
> Some taxa (mostly varieties) which were different with full length rbcL, become identical with 263 bp barcode
> Need to adapt the taxonomy again

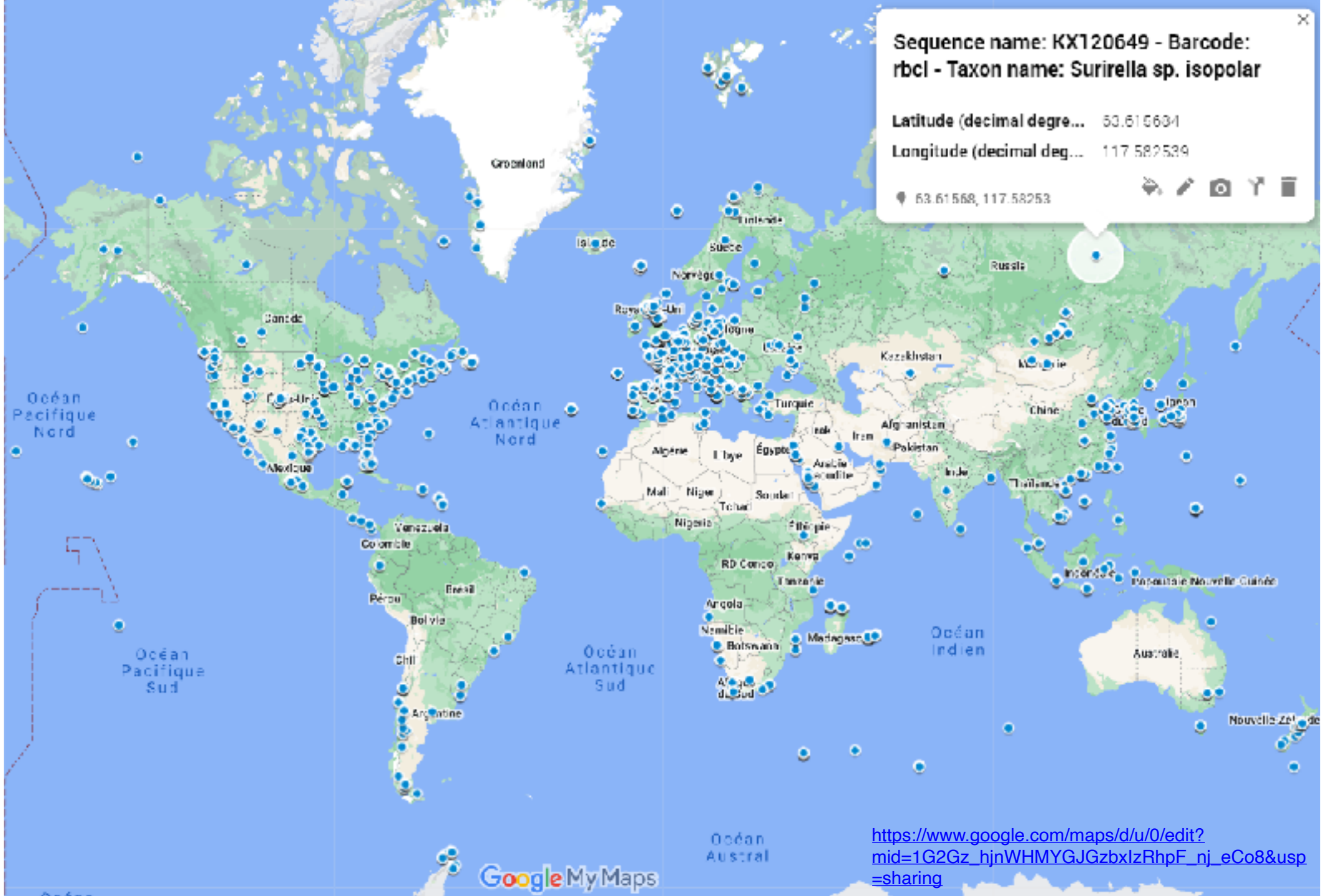


Diat.barcode

- V11 content :
  - 1512 taxa
  - 5017 rbcL sequences
  - 18S is not integrated since v5







https://www.google.com/maps/d/u/0/edit?mid=1G2Gz_hjnWHMYGJGzbxIzRhpF_nj_eCo8&usp=sharing

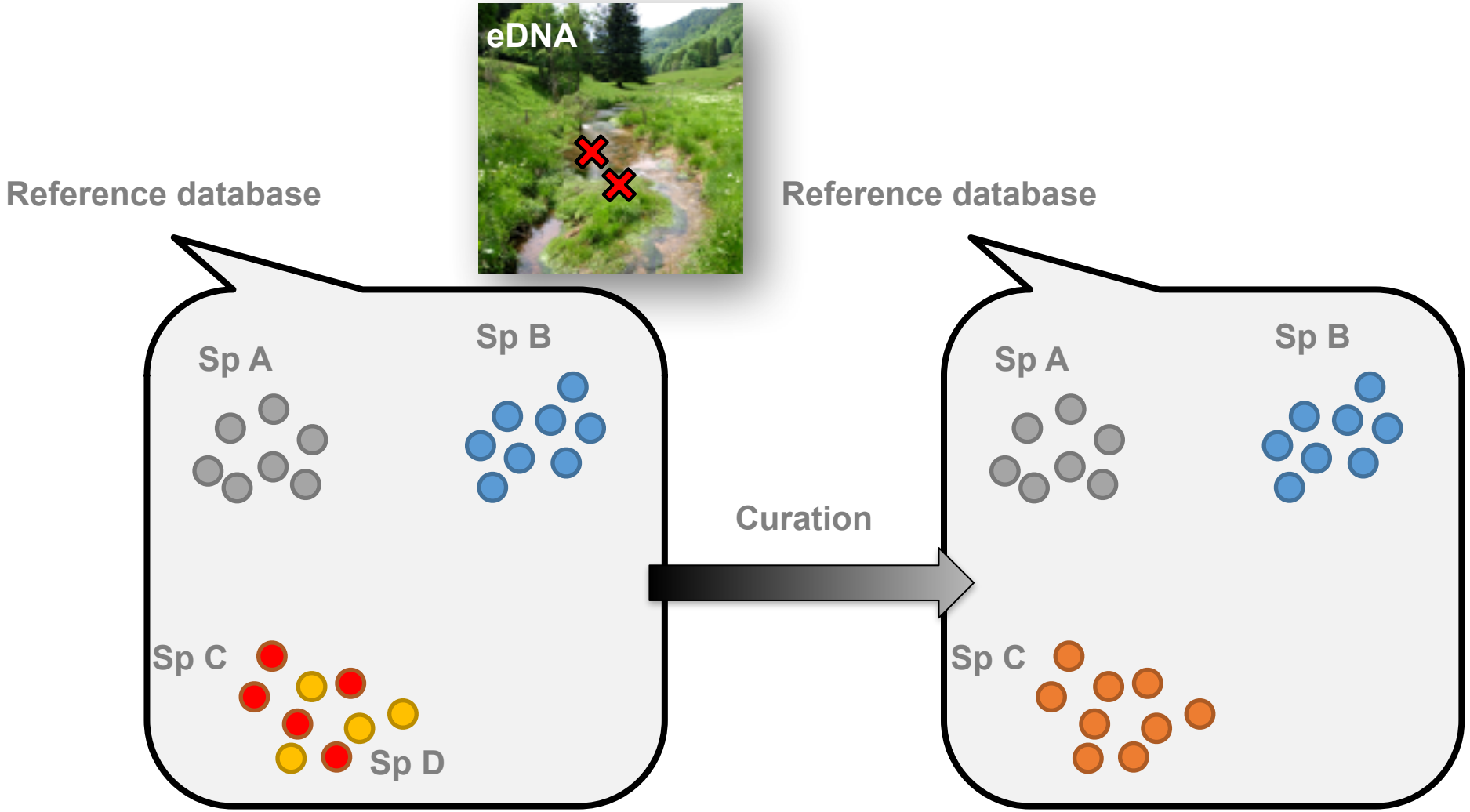
_02

Curation procedure

- **Curation**

Curation procedure

- Curation addresses a practical problem of taxonomic assignation



Curation procedure

❖Objective of the curation?

- ❖ To homogenize taxonomical names for a given phylogenetic clade

❖Why?

- ❖ The identification skills of the authors of the barcodes can be different
- ❖ Taxonomy evolves from a year to another
- ❖ Length and quality of sequences may not be suitable for correct taxonomic identification

R-syst: curation

❖Collaborative work since January 2018


F. Rimet
(France)


L. Kermarrec
(France)


M. Kahlert (Sweden)


D. Mann
(United Kingdom)


M. Kelly
(United Kingdom)


M. Pfankuchen
(Croatia)


J. Zimmermann
(Germany)

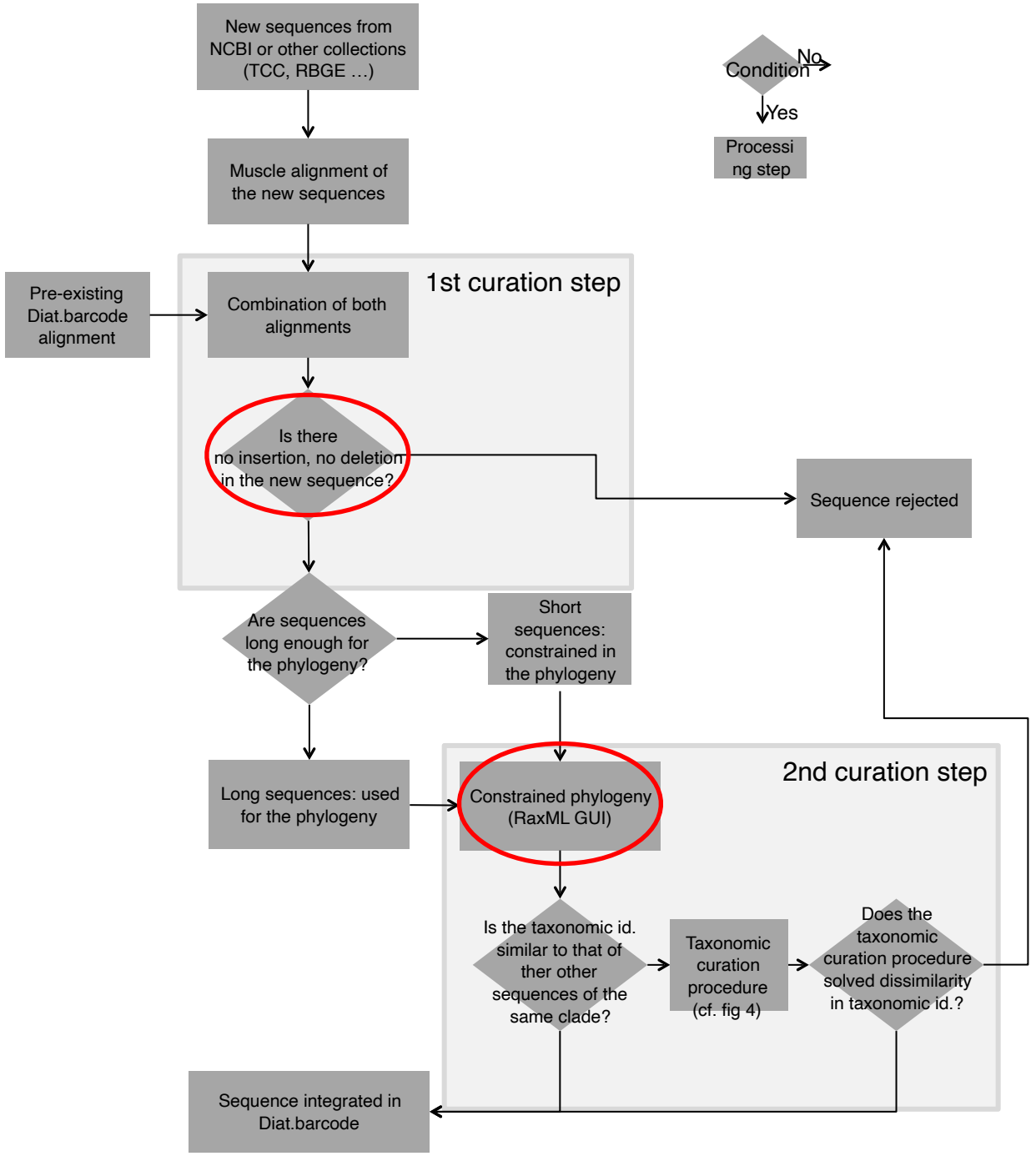

R. Trobajo
(Spain)

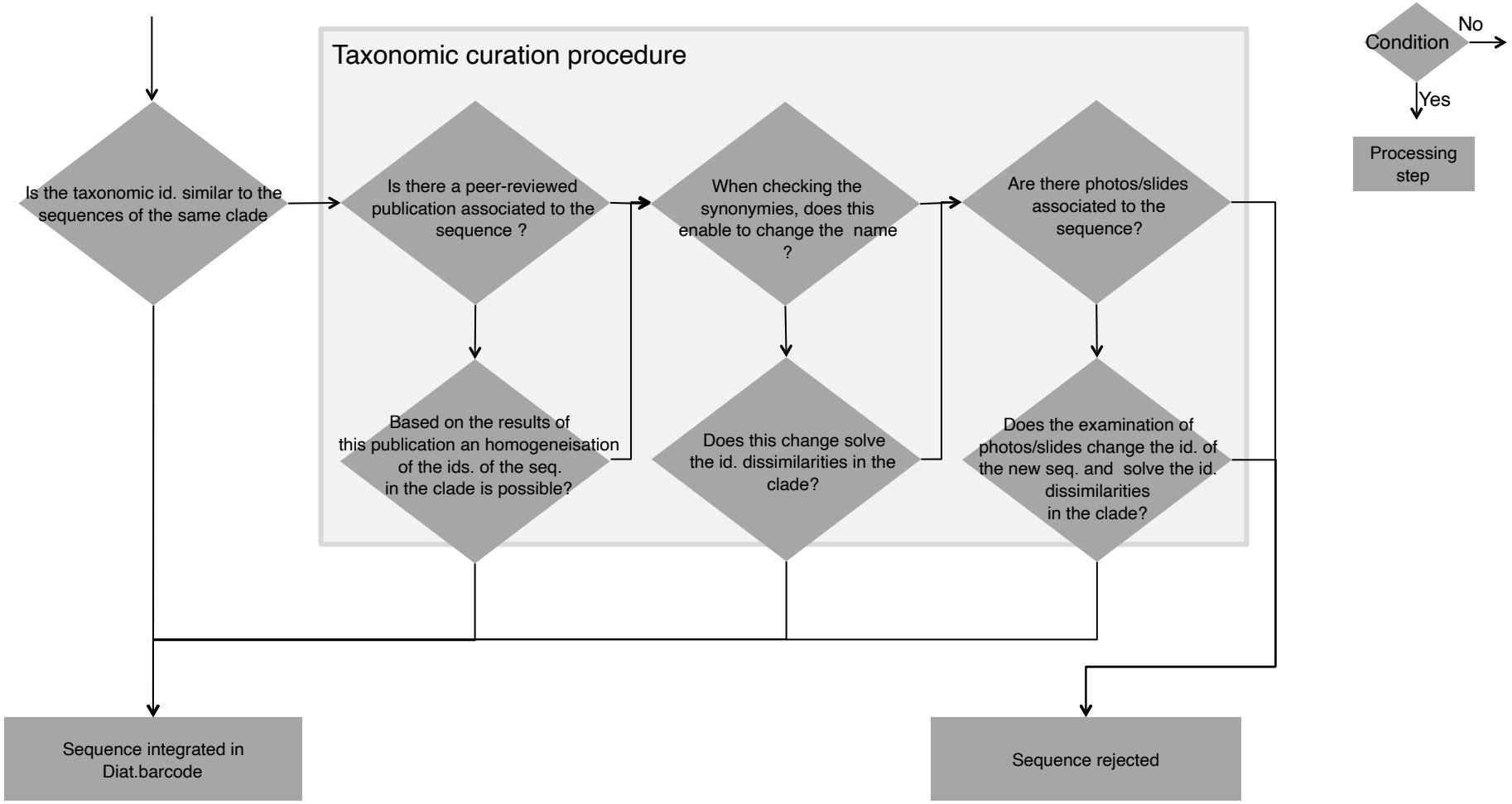

M. Kulikovskiy
(Russia)

❖Publications :

Rimet, Frederic; Chonova, Teofana; Gassiole, Gilles; Gusev, Evgenuy; Kahlert, Maria; Keck, François; Kelly, Martyn; Kulikovskiy, Maxim; Maltsev, Yevhen; Mann, David; Pfannkuchen, Martin; Trobajo, Rosa; Vasselon, Valentin; Wetzel, Carlos; Zimmermann, Jonas; Bouchez, Agnès, 2018, "Diat.barcode, an open-access barcode library for diatoms", https://doi.org/10.15454/TOMBYZ, Portail Data INRAE, V9

Rimet F., Gusev E., Kahlert M., Kelly M., Kulikovskiy M., Maltsev Y., Mann D., Pfannkuchen M., Trobajo R., Vasselon V., Zimmermann J., Bouchez A., 2019. Diat.barcode, an open-access curated barcode library for diatoms. Scientific Reports. https://www.nature.com/articles/s41598-019-51500-6


New sequences from NCBI or other collections (TCC, RBGE …)
Muscle alignment of the new sequences
Pre-existing Diat.barcode alignment
Combination of both alignments
1st curation step
Is there no insertion, no deletion in the new sequence?
Sequence rejected
Are sequences long enough for the phylogeny?
Short sequences: constrained in the phylogeny
Long sequences: used for the phylogeny
Constrained phylogeny (RaxML GUI)
2nd curation step
Is the taxonomic id. similar to that of ther other sequences of the same clade?
Taxonomic curation procedure (cf. fig 4)
Does the taxonomic curation procedure solved dissimilarity in taxonomic id.?
Sequence integrated in Diat.barcode
Condition
No
Yes
Processi ng step

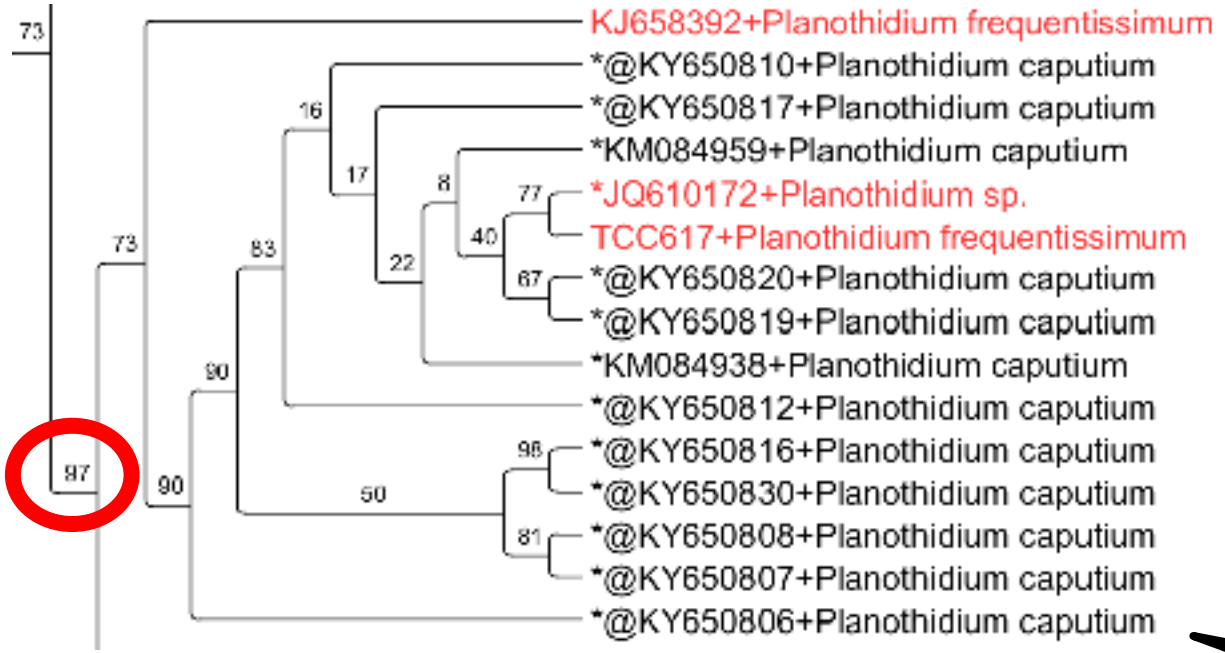

Taxonomic curation procedure
Is the taxonomic id. similar to the sequences of the same clade
Is there a peer-reviewed publication associated to the sequence ?
When checking the synonymies, does this enable to change the name ?
Are there photos/slides associated to the sequence?
Based on the results of this publication an homogeneisation of the ids. of the seq. in the clade is possible?
Does this change solve the id. dissimilarities in the clade?
Does the examination of photos/slides change the id. of the new seq. and solve the id. dissimilarities in the clade?
Sequence integrated in Diat.barcode
Sequence rejected
Condition
No
Yes
Processing step

R-syst: curation

❖**Exemple of curation done by J Zimmermann (in 2018)**

***P. caputium*: new species described in Zimmermann et al. 2014. PlosOne**

***P. caputium* > synonyme of *P. victorii* (transferred into *P. victorii* in v9)**



protistcentral.org

Publication

Diat.barcode

Publication + Algaterra

- Good bootstrap support = we can consider it as a robust species
- Availability of photos

Standardization: diatom libraries

- **General agreement:**
  - **Importance of metadata traceability !**
  - **Accessibility of data: public**
- **Storage of voucher specimens: in collections (Museums, Botanical gardens…), slides, strains, dry/frozen material**
- **Metadata:**
  - **Sampling: date, coordinates, country, habitat, depth, name of the person**
  - **Voucher: in which collection, voucher ID, duplicate, DNA voucher…**
  - **Molecular data: Sequencing technology, marker, primers, extraction method, PCR protocol, dates, name of the persons**
  - **Culture details: isolation date, name of isolator, culture medium, strain identifier**
  - **Taxonomic information: species name, name of the person who did the id, literature used**
- **Photos!**

Standardization: diatom libraries

- ❖Standardization: started in 2012
- ❖European standardization Committee: formally accepted in 4 June 2018 (after 6 years process), translated in French in the end of 2018
- ❖A Poulickova (CZ), D Mann (UK), M Kelly (UK), M Pfannkuchen (HR), M Kahlert (S), R Trobajo (SP), K Sabbe (B), J. Zimmermann (D), A Bouchez (FR), F Rimet (FR)
- ❖Neela ENKE (D)

EUROPEAN COMMITTEE FOR STANDARDIZATION
COMITÉ EUROPÉEN DE NORMALISATION
EUROPÄISCHES KOMITEE FÜR NORMUNG

TECHNICAL REPORT
RAPPORT TECHNIQUE
TECHNISCHER BERICHT

FINAL DRAFT
FprCEN/TR 17244

January 2018

ICS 13.060.70

English Version

Water quality - Technical report for the management of diatom barcodes

Standardization: all aquatic organisms

(Rimet et al. MBMG 2021)

MBMG
Metabarcoding & Metagenomics

Metabarcoding and Metagenomics 5: 1–31
DOI 10.3897/mbmg.5.58050

Methods

Metadata standards and practical guidelines for specimen and DNA curation when building barcode reference libraries for aquatic life

Frédéric Rimet[1], Eva Aylagas[2], Angel Borja[3,27], Agnès Bouchez[1], Alexis Canino[1], Christian Chauvin[4], Teofana Chonova[1], Fedor Ciampor Jr[5], Filipe O. Costa[6,7], Benoit J.D. Ferrari[8], Romain Gastineau[9], Chloé Goulon[1], Muriel Gugger[10], Maria Holzmann[11], Regine Jahn[12], Maria Kahlert[13], Wolf-Henning Kusber[12], Christophe Laplace-Treyture[1], Florian Leese[14], Frederik Leliaert[15], David G. Mann[16,17], Frédéric Marchand[18], Vona Méléder[19], Jan Pawlowski[11,20,21], Serena Rosoni[1], Sinziana Rivera[1], Rodolphe Rougerie[22], Magali Schweizer[23], Rosa Trobajo[19], Valentin Vasselon[24], Régis Vivien[8], Alexander Weigand[25], Andrzej Witkowski[9], Jonas Zimmermann[12], Torbjørn Ekrem[26]

DNAqua-Net

F. Leese
A. Bouchez
(leaders)

3.2 Categories of metadata

3.2.1 Biological material metadata

The metadata listed below give the obligatory and recommended items that ensure the traceability of the biological material used for DNA harvesting (see Section 2.1).

Biological specimens and environmental samples

Obligatory metadata

1. Location of the sampling site
   a) Geographical coordinates: for example, expressed in decimal values in WGS84 or in a different, specified geographical positioning system.
   b) Country according to the ISO 3166 standard, accepted name of ocean or sea.
   c) Name of the locality.

Remarks:

For species of heritage interest, Red List species or endangered species, national or regional regulations might ask not to reveal precise location coordinates in order to protect their populations. These regulations should be followed and the exact locality information hidden.

In some cases, exact coordinates are not available (e.g. when older museum specimens are used; here a georeference of the locality plus an estimated uncertainty in metres can be added).

2. Date of sampling, preferably in ISO-format (YYYY-MM-DD).
3. Name of person who collected the specimen.
4. Photo(s) of the voucher specimen showing diagnostic features, including scale(s).
   a) Macroscopic organisms: whole specimen or specified parts important for morphological identification. For fish, photos should be taken of the left side of the specimen. For specimens in which the morphology can be altered during storage (e.g. dry storage of molluscs) or preservation method (e.g.

Recommended metadata

1. Environment (ecosystem) at sampling site (e.g. Lake, river, swamp, tidal flat, open sea, groundwater, hyporheos, mangrove, lagoon, estuary, deep sea, rocky shore, coral reef, etc.).
2. Substrate (rock, macrophyte, sediment, hot vent, interstitial, etc.).
3. Habitat (e.g. plankton, epipelon, epilithon, epipsammon, hyporheos, alluvial region, porous or karstic aquifer, sea floor, pelagic, benthic, intertidal, subtidal, etc.).
4. Sampling elevation (m a.s.l.).
5. Sampling depth (m).
6. Sampling device or sampling protocol.
7. Photos of the sampling site.
8. Environmental measurements: luminosity, pH, conductivity, salinity, temperature, sediment's grain size, organic matter content and redox potential.
9. Main ecological function(s) of the specimen (if known). For instance, already existing ecological classifications, such as FAPROTAX (Louca et al. 2016) and TaxaFun (Aßhauer et al. 2015) for bacteria, the classification of Reynolds et al. (2002) and Padisák et al. (2009) for phytoplankton and Rimet and Bouchez (2012) for diatoms, for macro-invertebrates the classification of Usseglio-Polatera et al. (2000) or for plants, the one of Katrge et al. (2011).
10. Photos should carry the name of the photographer and associated licence, preferably a Creative Commons Licence that allows usage by third parties.
11. The FAO fishery area where the sampling was done (for marine taxa).

Cultures

Obligatory

1. Metadata associated with the environmental sample which was used to establish the culture (see above section).

BIOLAWEB

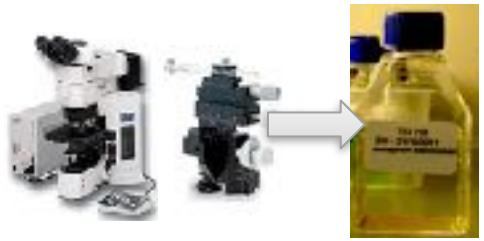
_03

Completeness of Diat.barcode

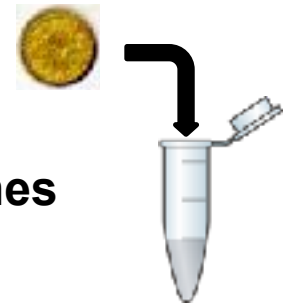
- How can we complete the library?
- Completeness in rivers of France
- Completeness in other habitats

How can we complete the library?

- Cell isolation, cultures:
  - This is the classical method
  - Long and tedious



- Single cell PCR (Hamilton et al. 2015):
  - Identification of living cell
  - Ok for big species, more difficult for small ones



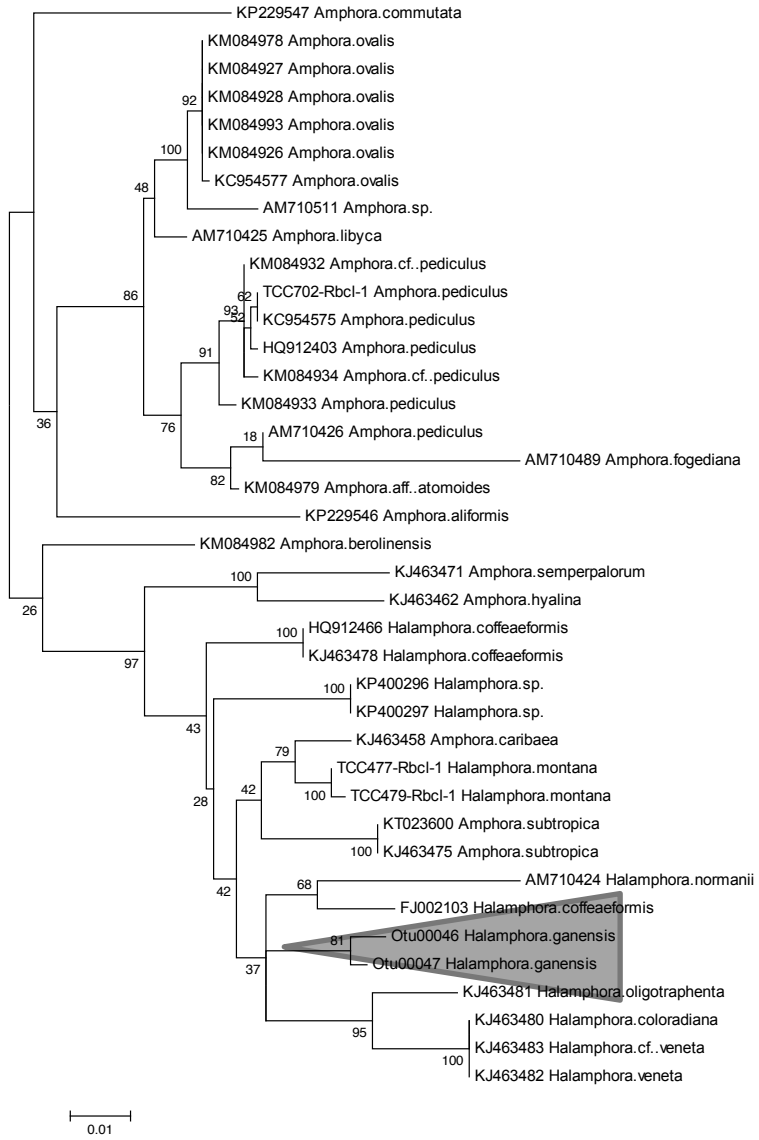
- Cloning-sequencing of natural samples (Khan-Bureau et al. 2016)
  - Long sequences of good quality
  - Expensive
  - Can be done for samples with low diversities

How can we complete the library?

❖Use of HTS sequences

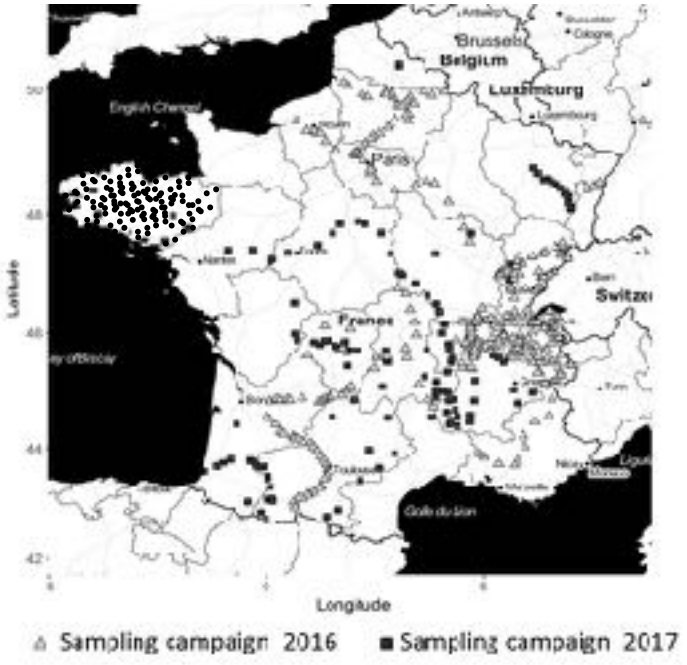
- ❖ Select samples with target species never sequenced
- ❖ Criteria: - frustule and sequences must be abundant
  - rbcl (coding): no indel, no stop codon
  - Phylogenetic checkout

❖ F. RIMET, N. ABARCA, A. BOUCHEZ, R. JAHN, M. KAHLERT, F. KECK, M.G. KELLY, D.G. MANN, A. PIUZ, R. TROBAJO, K. TAPOLCZAI, V. VASSELON, J. ZIMMERMANN. Fottea



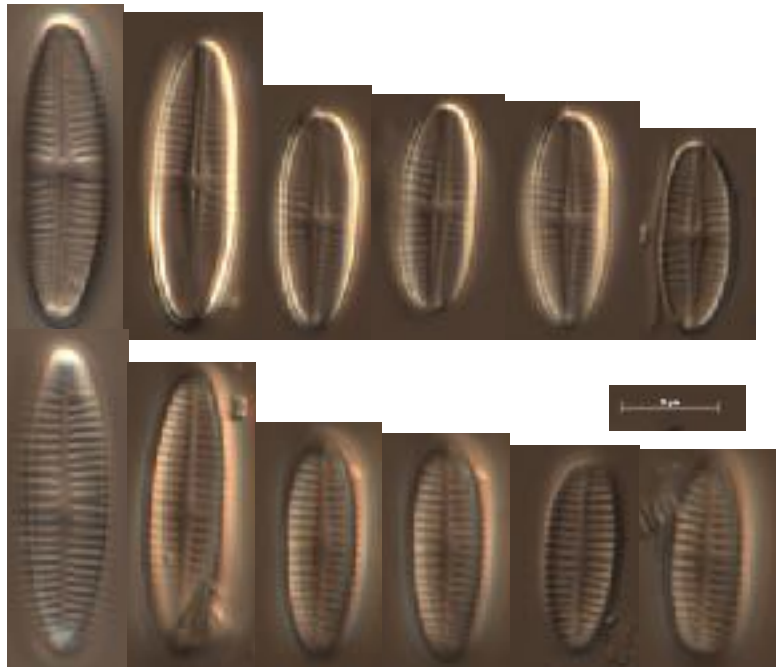
Résultats

- ❖Completion of Diat.barcode using environnemental samples
- ❖Several tens of abundant species in mainland France were added in Diat.barcode



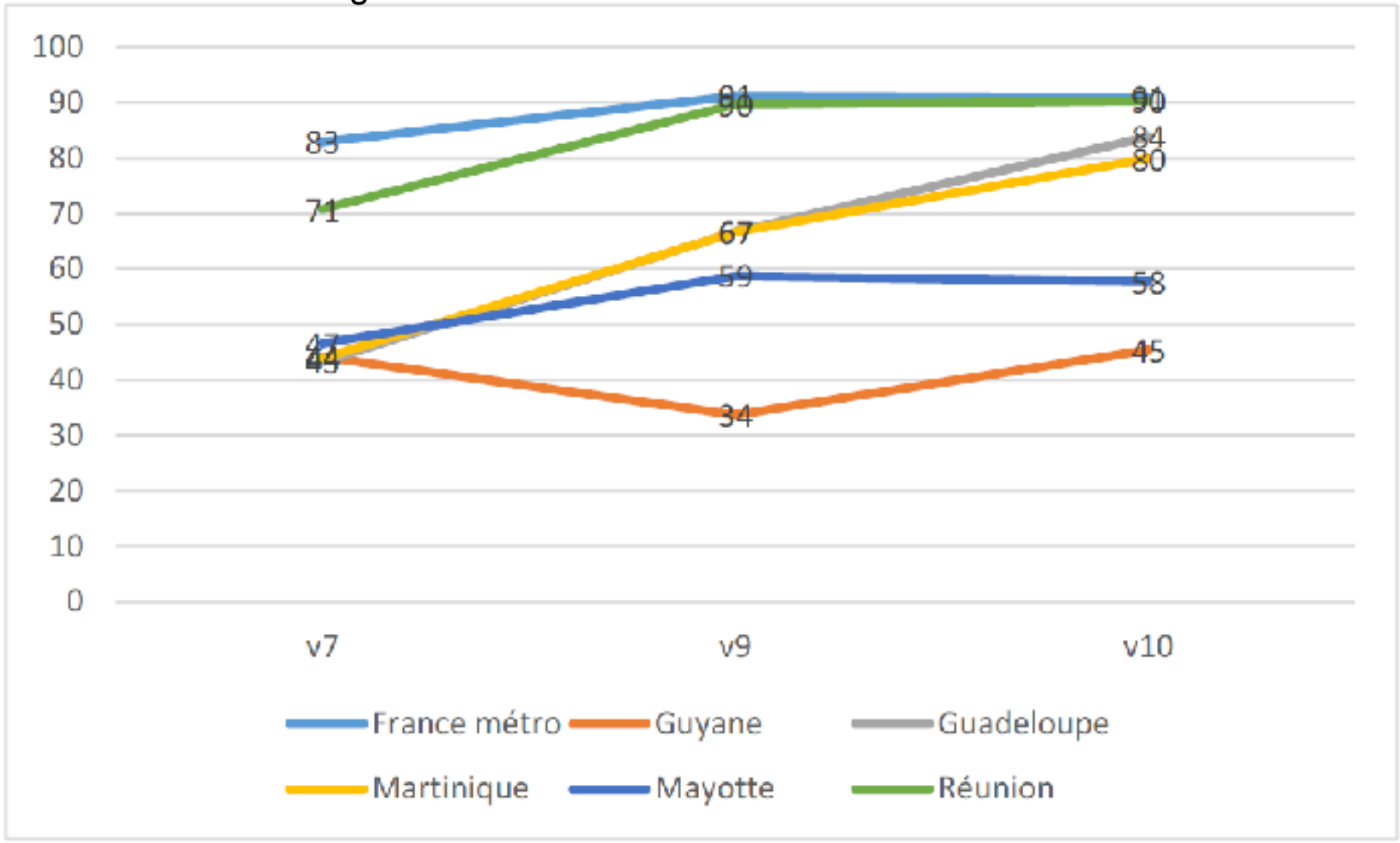
+ Bretagne 2019

Example of species added in Diat.barcode



Achnanthidium delmontii
TCC961, 03/05/2017
ADOUR - MAUBOURGET
Prélev : R. IMBERT
Identif : F. PERES

Proportion of reads assigned to a species level according to the three versions of Diat.barcode (v7, v9 and v10), for the 6 regions of interest.



Diat.barcode : completness?

However, for some environments, the library remains incomplete.

Mayotte Island

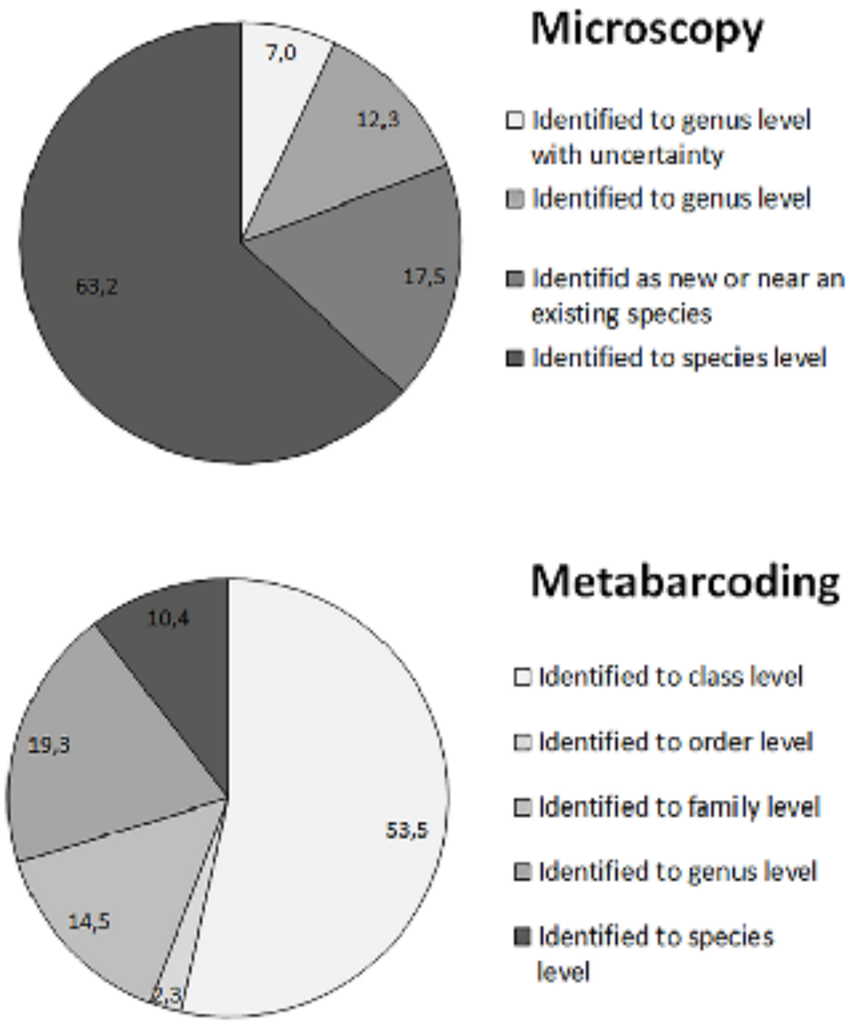


Fig 2. Identification levels reached by microscopy and by metabarcoding. Percentages of the identified taxa through microscopy and of the OTU identified using R-Syst::diatom v4 (version of 16-05-2015) are given in the pie charts, respectively.

https://doi.org/10.1371/journal.pone.[illegible].g002

_04

Divergences between classical and integrative taxonomy

- Divergence at specific level
- Divergence at deeper nodes

Divergences between classical and integrative taxonomy

- **Classical taxonomy:**
  - Based on a single criteria
  - For diatoms: frustule morphology
  - Problem: phenotypique plasticity >> poses problems for the delimitation of species

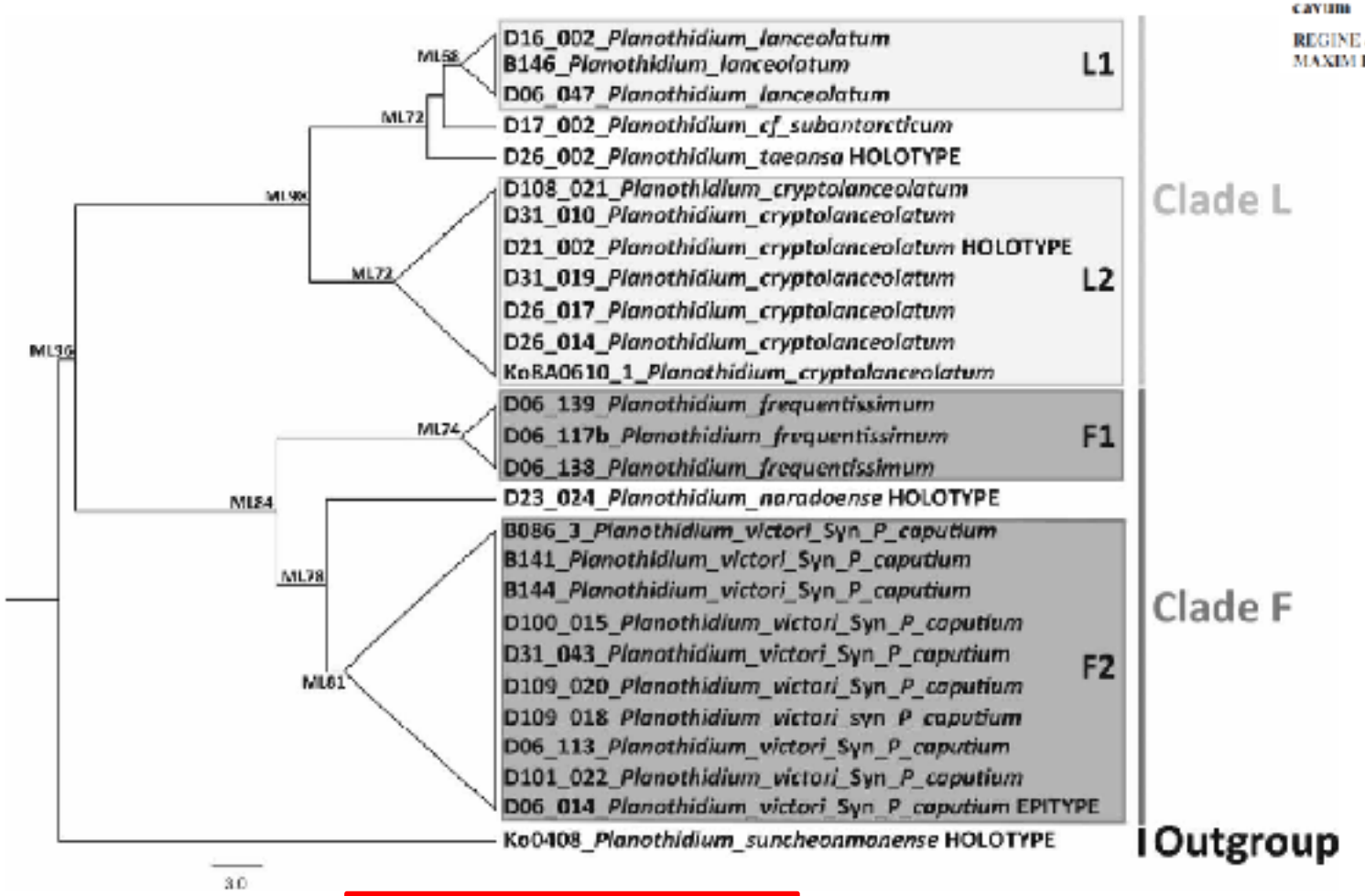
- **Integrative taxonomy**
  - Dayrat, B. (2005): Towards integrative taxonomy. – Biol. J. Linn. Soc. 85: 407–415.
  - Species description/delimitation should be based on several criteria:
    - Morphological
    - Molecular
    - Ecological
    - …

Divergences between classical and integrative taxonomy

❖Integrative taxonomy: example with *Planothidium*

Planothidium lanceolatum and *Planothidium frequentissimum* reinvestigated with molecular methods and morphology: four new species and the taxonomic importance of the sinus and cavum

REGINE JAHN[1*], NÚRIA ABARCA[1], BIRGIT GEMEINHOLZER[2], DEMETRIO MORA[1], OLIVER SKIBBE[1], MAXIM KULIKOVSKIY[3], EVGENY GUSEV[3], WOLF-HENNING KUSBER[1] & JONAS ZIMMERMANN[4]



Morphological criteria

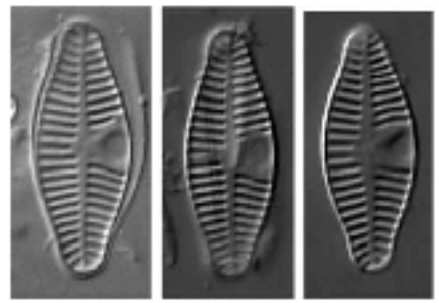
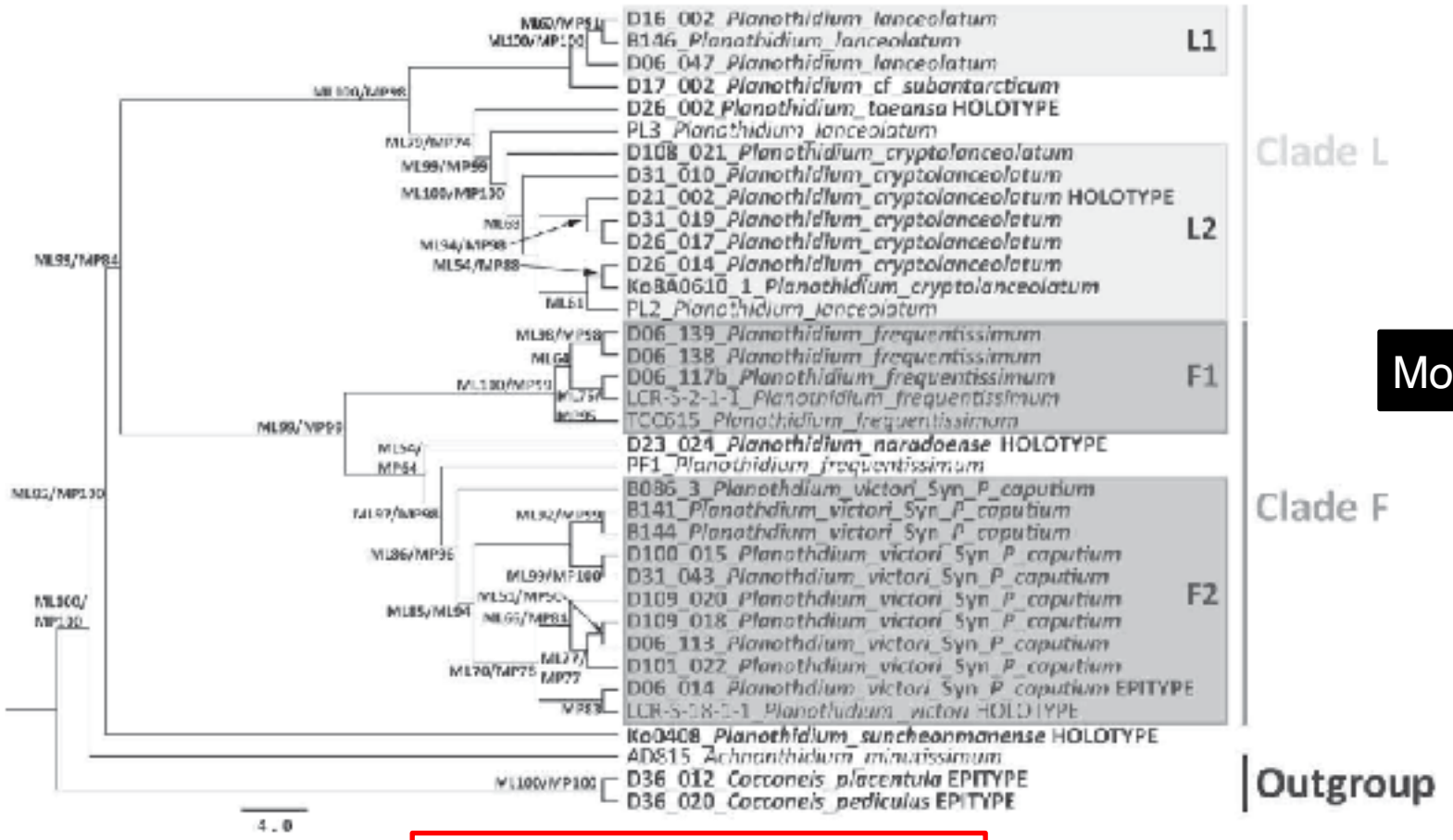


Fig. 2. Strict consensus tree of the ML analysis of the morphological character matrix (Table 3) with bootstrap statistics (> 50).

Divergences between classical and integrative taxonomy

❖Integrative taxonomy: example with *Planothidium*



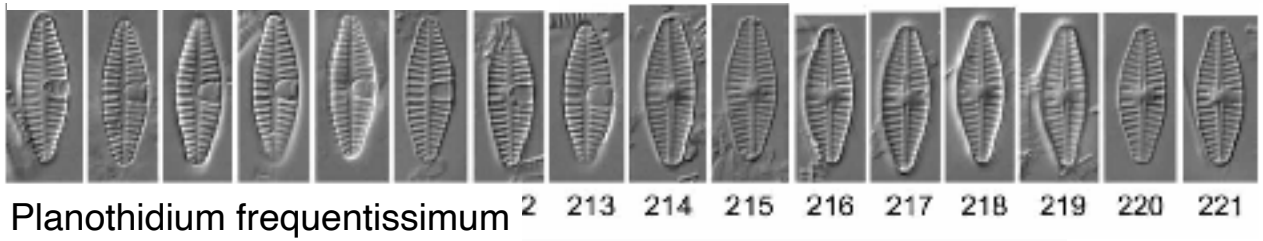
Molecular criteria

Fig. 1. Concatenated strict consensus tree of the combined dataset of the *rbc*L and 18S molecular markers with bootstrap statistics (> 50) for ML (LB) and MP (PB). Bold strains cultured by the authors.

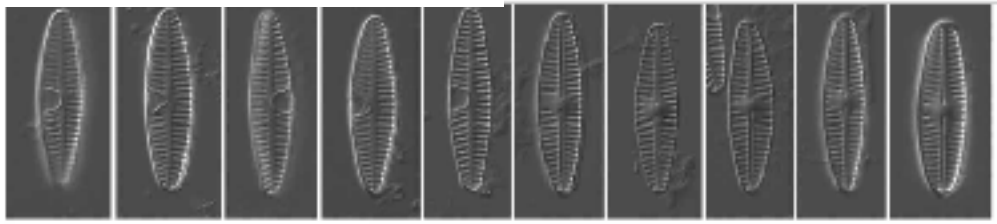
BIOLAWEB

Combination of morphological + molecular criteria to delimit species

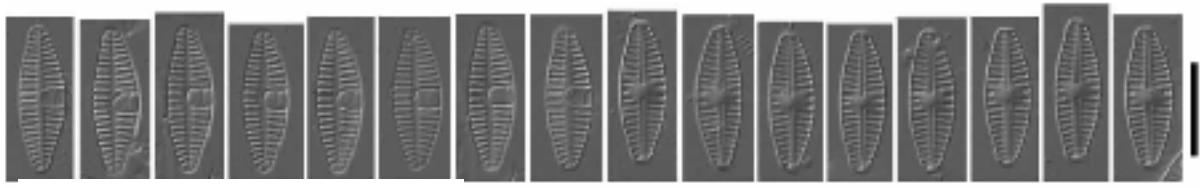
Planothidium cryptolanceolatum



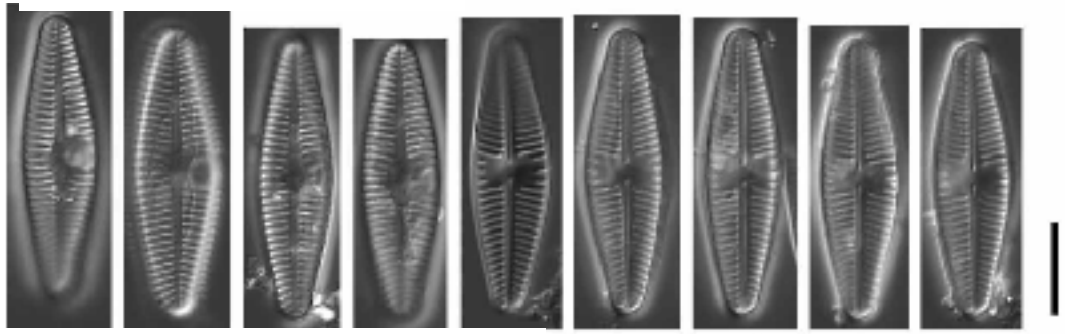
Planothidium frequentissimum



Planothidium victorii

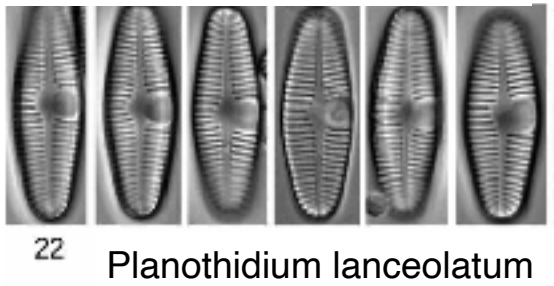


Planothidium naradoense

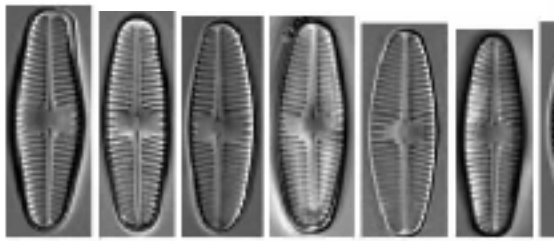


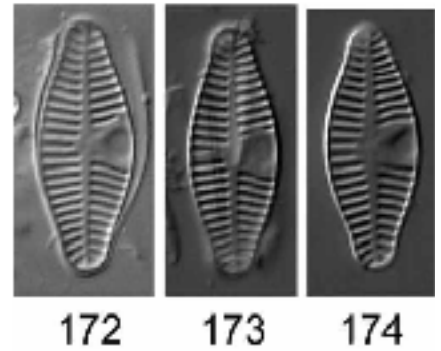
Planothidium cf. subantarcticum


Planothidium suncheonmanense

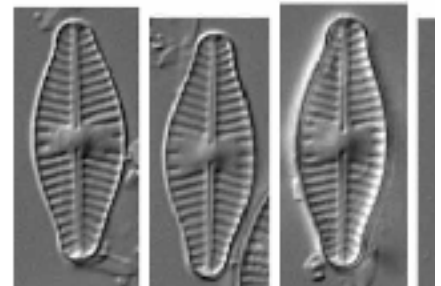


Planothidium lanceolatum





Planothidium taeansa



Divergences between classical and integrative taxonomy

❖Integrative taxonomy : example with

39 cultures of *Gomphonema parvulum*:

- ❖ sequenced (ITS, rbcL, cox1),
- ❖ morphology (LM, SEM), morphometry

❖4 clades

❖Biogeographic distribution limited (tropics/temperate)

❖Criteria discriminating these clades are not classically used

Does the Cosmopolitan Diatom *Gomphonema parvulum* (Kützing) Kützing Have a Biogeography?

Nelida Abarca[1]*, Regine Jahn[1], Jonas Zimmermann[1,2], Neela Enke[1]

❖Answer: Yes (PlosOne 2014)

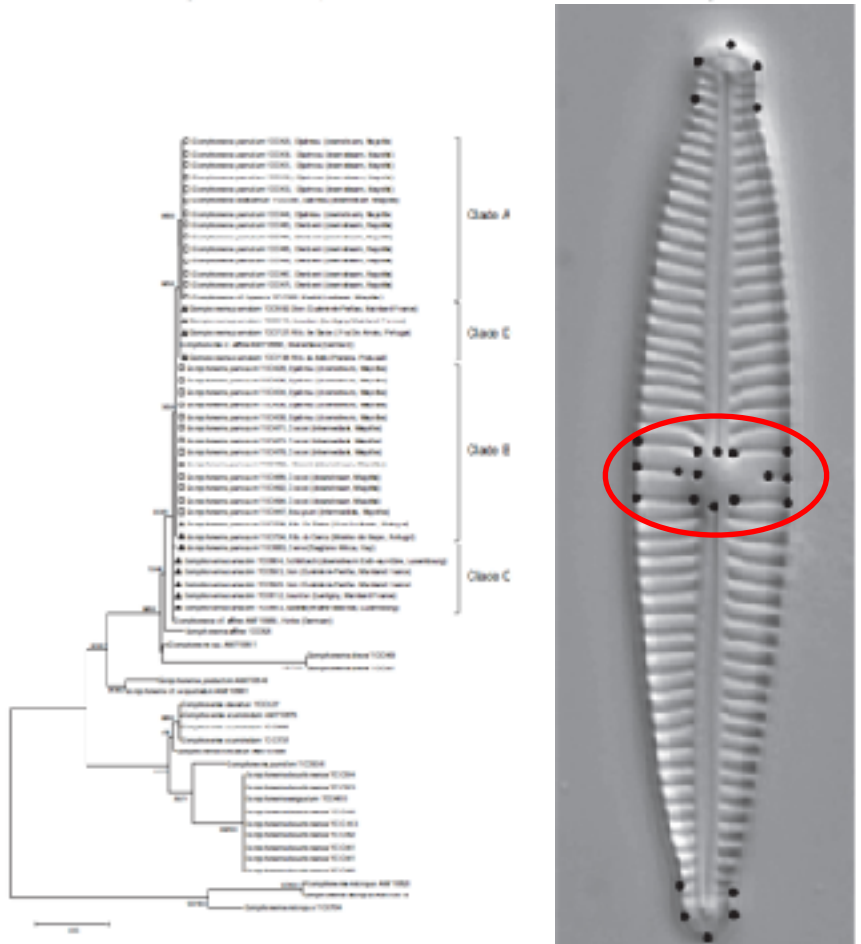
❖Formal description of new species almost impossible to reconized in LM

Protist

ORIGINAL PAPER

First Evidence of the Existence of Semi-Cryptic Species and of a Phylogeographic Structure in the *Gomphonema parvulum* (Kützing) Kützing Complex (Bacillariophyta)

Lenaig Kermarec, Agnès Bouchez, Frédéric Rimet, and Jean-François Humbert



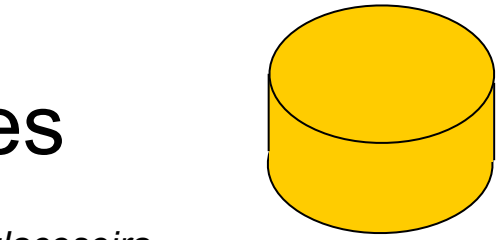
Divergences between classical and integrative taxonomy

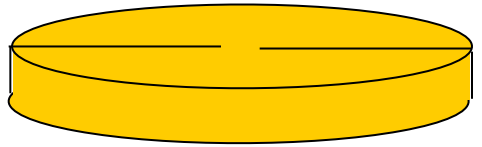
❖**Divergence at deeper nodes**

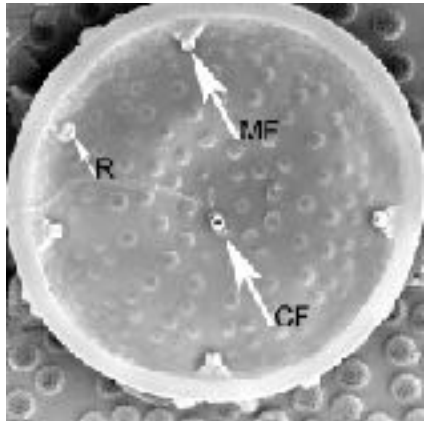
Less concerned in metabarcoding but important for our general back ground

Divergences between classical and integrative taxonomy

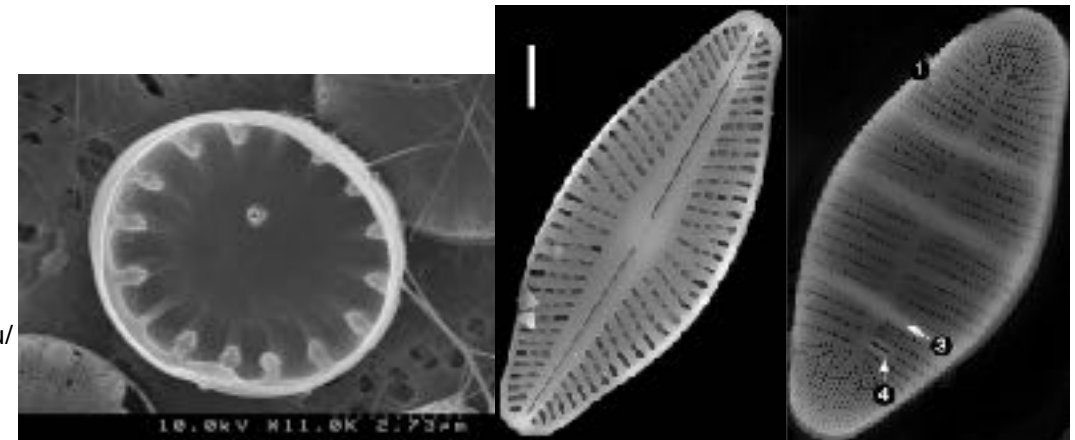
- Traditional classification : 4 classes
  - Coscinodiscophyceae : centrics cf. *Melosira, Aulacoseira*
  - Mediophyceae : centrics cf. *Cyclotella, Stephanodiscus*
  - Fragilapiophyceae : bilateral symmetry, no raphe
  - Bacillariophyceae : bilateral symmetry, raphe







Internal view of valve of Stephanodiscus sp. showing the rimoportula (R), central fultoportula (CF) and four marginal fultoportulae (MF). Each fultoportula has two satellite pores https://westerndiatoms.colorado.edu/glossary/term/Fultoportula



BIOLAWEB

- 207 cultures
- 6 genes:

1 nuclear, 4 chloroplastic, 1 mitochondria

Molecular Phylogenetics and Evolution 89 (2015) 28–36

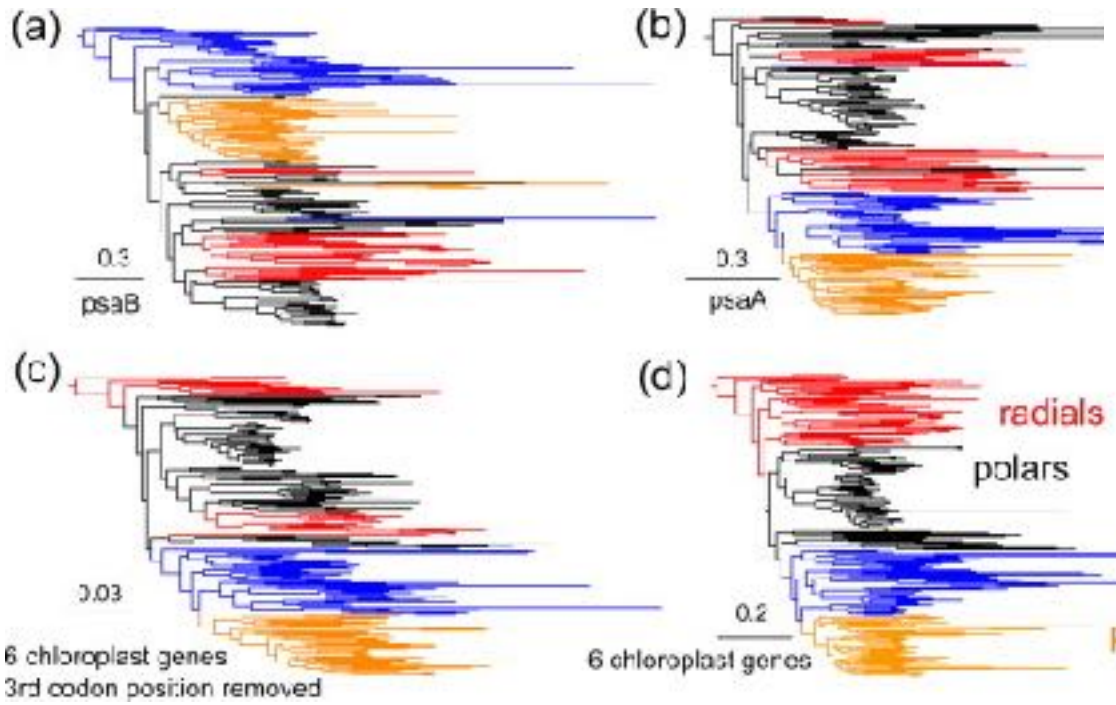
Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Dissecting signal and noise in diatom chloroplast protein encoding genes with phylogenetic information profiling

Edward C. Theriot [a], Matt P. Ashworth [b,*], Teofil Nakov [b,1], Elizabeth Ruck [c], Robert K. Jansen [a,d]



In addition to the nuclearly encoded small subunit of the rDNA (nSSU), and two chloroplast encoded genes, ribulose bisphosphate carboxylase large chain, (*rbcL*), and photosystem II CP43 protein gene (*psbC*) of Theriot et al. (2010, 2011), we sequenced the photosystem I p700 apoprotein A1 gene (*psaA*), the photosystem I p700 apoprotein A2 gene (*psaB*) and photosystem II protein D1 gene (*psbA*), and the ATP synthetase beta-subunit gene (*atpB*). DNA extraction and sequencing followed Alverson et al. (2007) and newly-developed primer sequences are available in Table S2.

Coscinodiscophyceae

Mediophyceae

Fragilariophyceae

Bacillariophyceae

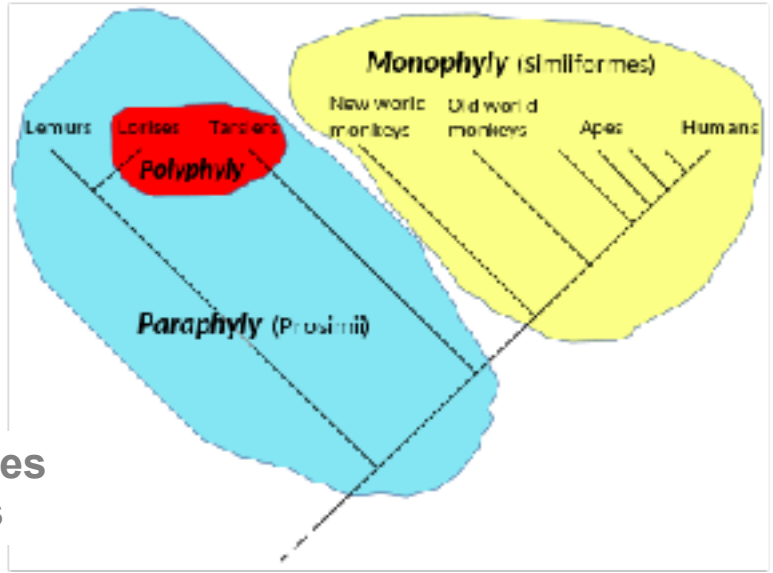
Fig. 2. Maximum likelihood trees of diatom structural group relationships from analysis of (a) the chloroplast encoded *psaA* gene, (b) *psaB* gene, (c) 1st and 2nd codon positions of all six chloroplast genes aggregated, and (d) all six chloroplast genes, all positions. Branch color reflects major structural groupings (red = radials, black = polars, blue = araphids, orange = Raphid pennates).

http://dx.doi.org/10.1016/j.ympev.2015.03.012

Divergences between classical and integrative taxonomy

❖The 4 diatom classes are paraphyletic : !! Systematic only accept monophyletic groups !!

Polyphyletic : group whose descendants do not have a common ancestor



Monophyletic : group whose descendants have a common ancestor

Paraphyletic: group that does not include all descendants

BIOLAWEB

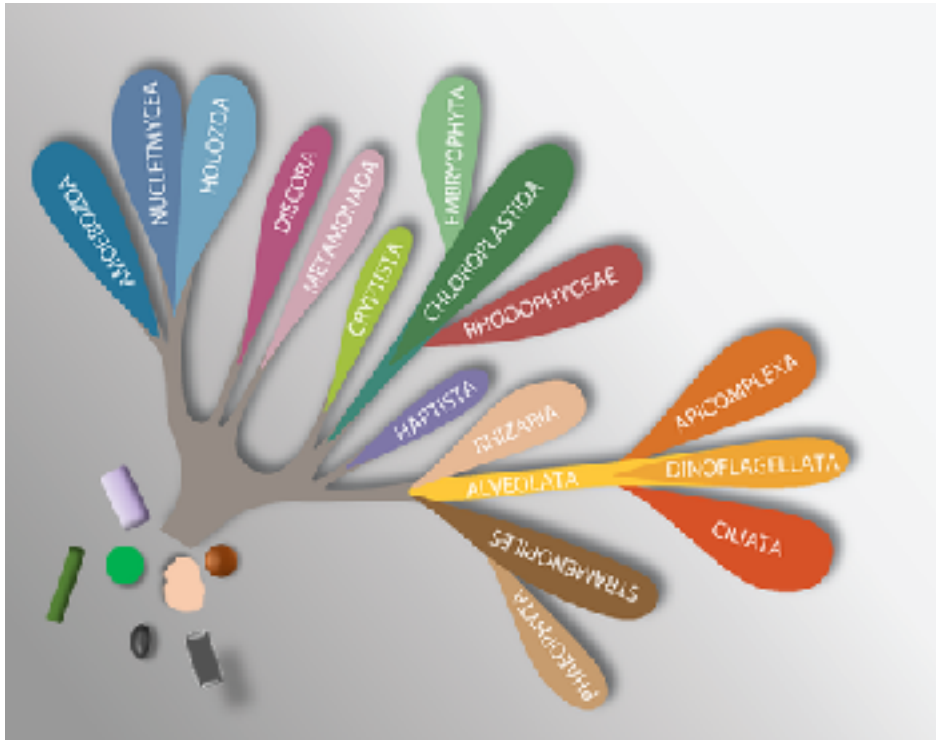
Journal of Eukaryotic Microbiology ISSN 1066-5234

ORIGINAL ARTICLE

Revisions to the Classification, Nomenclature, and Diversity of Eukaryotes

Sina M. Adl[a,*], David Bass[b,c], Christopher E. Lane[d], Julius Lukeš[e,f], Conrad L. Schoch[g], Alexey Smirnov[h], Sabine Agatha, Cedric Berney, Matthew W. Brown[i,j], Fabien Burki[k], Paco Cárdenas[l], Ivan Čepička[p], Lyudmila Chistyakova[p], Javier del Campo[q], Micah Dunthorn[r,s], Bente Edvardsen[t], Yana Eglit[u], Laure Guillou[v], Vladimir Hampl[w], Aaron A. Heiss[x], Mona Hoppenrath[y], Timothy Y. James[z], Anna Karnkowska[aa], Sergey Karpov[bb,cc], Eunsoo Kim[x], Martin Kolisko[dd], Alexander Kudryavtsev[b,ab], Daniel J.G. Lahr[ac], Enrique Lara[ad,ae], Line Le Gall[af], Denis H. Lynn[ag,ah], David G. Mann[ai], Ramon Massana[aj], Edward A.D. Mitchell[ak,al], Christine Morrow[d], Jong Soo Park[am], Jan W. Pawlowski[an], Martha J. Powell[ao], Daniel J. Richter[ap], Sonja Rueckert[aq], Lora Shadwick[ar], Satoshi Shimano[as], Frederick W. Spiegel[at], Guifré Torruella[au], Noha Youssef[av], Vasily Zlatogursky[b,aw] & Qianqian Zhang[ax]

Results of Theriot were integrated in the revision of Adl by Mann



Update of diatom taxonomy

Diat.barcode include « classical » and updated taxonomy

DIAT.BARCODE

Phylum	Subdivision	Class	Sub class	Genre
Diatomeae	Leptocylindrophytina	Leptocylindrophyceae		Leptocylindrus Tenuicylindrus
Diatomeae	Leptocylindrophytina	Corethrophyceae		Corethron
Diatomeae	Ellerbeckiophytina			Ellerbeckia
Diatomeae	Probosciophytina			Proboscia
Diatomeae	Melosirophytina			Aulacoseira Melosira Hyalodiscus Stephanopyxis Paralia Endictya
Diatomeae	Coscinodiscophytina			Actinoptychus Coscinodiscus Actinocyclus Asteromphalus Aulacodiscus Stellarima
Diatomeae	Rhizosoleniophytina			Guinardia Rhizosolenia Pseudosolenia
Diatomeae	Arachnoidiscophytina			Arachnoidiscus
Diatomeae	Bacillariophytina	Mediophyceae		
Diatomeae	Bacillariophytina	Mediophyceae	Chaetocerotophycidae	Hydrosera
Diatomeae	Bacillariophytina	Mediophyceae	Lithodesmiophycidae	Lithodesmium Lithodesmioides Helicotheca Bellerochea Ditylum
Diatomeae	Bacillariophytina	Mediophyceae	Thalassiosirophycidae	Thalassiosira Lindavia Cyclotella Stephanodiscus Cyclostephanos Discostella Bacteriosira Skeletonema Detonula
Diatomeae	Bacillariophytina	Mediophyceae	Cymatosirophycidae	Cymatosira Minutocellus Papiliocellulus Leyanella Extubocellulus Plagiogrammopsis Campylosira Brockmanniella Pierrecomperia
Diatomeae	Bacillariophytina	Mediophyceae	Odontellophycidae	Odontella Triceratium Cerataulus Pleurosira Pseudauliscus Amphitetras Trieres
Diatomeae	Bacillariophytina	Mediophyceae	Chrysanthemodiscophycidae	Chrysanthemodiscus Biddulphiopsis Trigonium Isthmia Lampriscus Stictocyclus Ardissonea Climacosphenia Toxarium
Diatomeae	Bacillariophytina	Biddulphiophyceae	Biddulphiophycidae	Biddulphia Attheya
Diatomeae	Bacillariophytina	Bacillariophyceae	Striatellaceae	Striatella Pseudostriatella
Diatomeae	Bacillariophytina	Bacillariophyceae	Urneidophycidae	Plagiogramma Dimeregramma Rhaphoneis Delphineis Psammoneis Bleakeleya
Diatomeae	Bacillariophytina	Bacillariophyceae	Fragilariophycidae	Fragilaria Synedra Tabellaria Asterionella Diatoma Tabularia Cyclophora Astrosyne Licmophora Rhabdonema Grammatophora Staurosira Thalassionema
Diatomeae	Bacillariophytina	Bacillariophyceae	Bacillariophycidae	Cylindrotheca Navicula Seminavis Haslea Stauroneis Pleurosigma Gyrosigma Achnanthidium Cocconeis Frustulia Diploneis Sellaphora Pinnularia Gomphonema Cymbella Didymosphenia Phaeodactylum Amphora Entomoneis Epithemia

BIOLAWEB

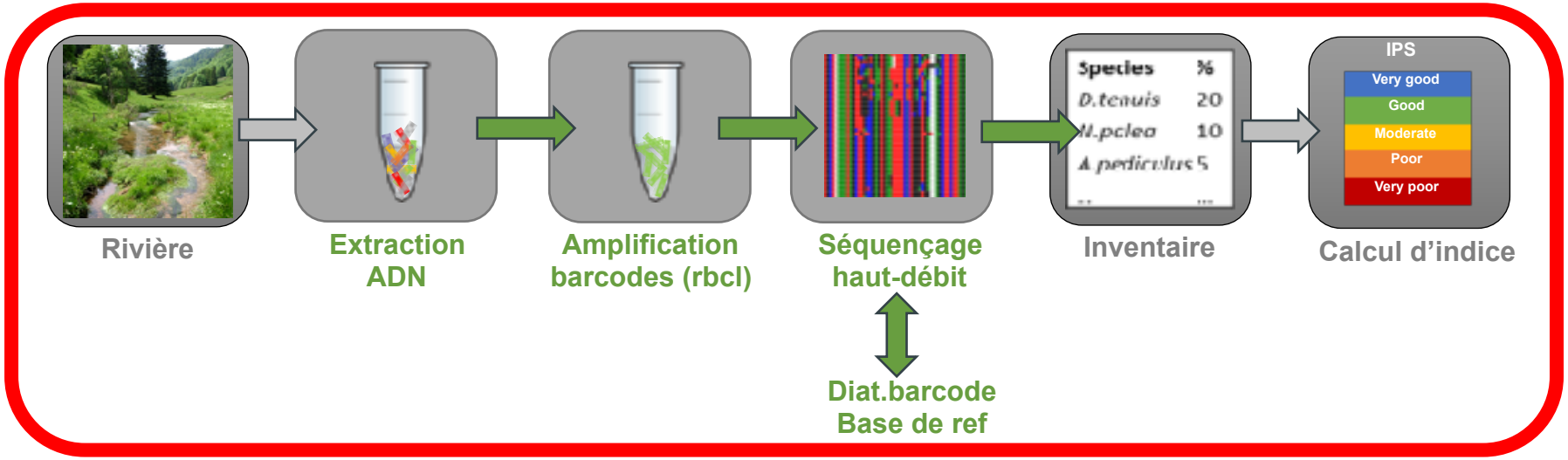
Schedule

1- reference library

2- 1st proofs of concept

3- Impact of biovolumes

1st proofs of concept



A next-generation sequencing approach to river biomonitoring using benthic diatoms

Lenaïg Kermarrec[1,2,3,7], Alain Franc[1,6,8], Frédéric Rimet[2,3,9], Philippe Chaumeil[4,6,10], Jean-Marc Frigerio[4,5,11], Jean-François Humbert[6,12], and Agnès Bouchez[2,3,13]

[1]Asconit Consultants, 3 boulevard Clairfont, 66350 Toulouges, France
[2]INRA, UMR CARRTEL, 75 avenue de Corzent, BP 511, 74203 Thonon-les-Bains cedex, France
[3]University of Savoie, UMR CARRTEL, 73370 Le Bourget du Lac, France
[4]INRA, UMR BioGeCo, 69 route d'Arcachon, 33612 Cestas cedex, France
[5]University of Bordeaux 1, UMR BioGeCo, 33400 Talence, France
[6]INRA, UMR BIOEMCO, site de l'ENS, 46 rue d'Ulm 75005 Paris France

CHICAGO JOURNALS

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

ELSEVIER

Research paper

Assessing ecological status with diatoms DNA metabarcoding: Scaling-up on a WFD monitoring network (Mayotte island, France)

Valentin Vasselon*, Frédéric Rimet, Kálmán Tapolczai, Agnès Bouchez

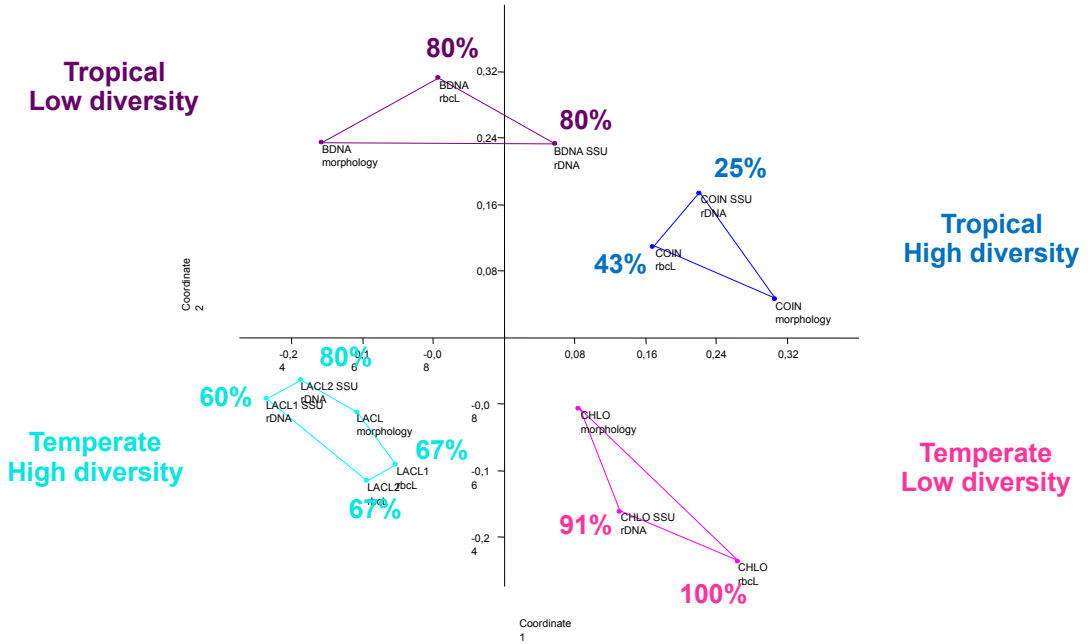
CARRTEL, INRA, Université de Savoie Mont-Blanc, 74200 Thonon-les-Bains, France

Funded by European Union

1st proofs of concept (1)

For the first proof of concept we wanted to compare the inventories (not the index values)

- **4 samples 2009 :**
  - **Tropical low diversity**
  - **Tropical high diversity**
  - **Temperate low diversity**
  - **Temperate high diversity**
- **Roche 454: 18s + rbcl**
- **Bioinformatic: developped a homemade program: Metamatch**
- **Good correspondence rbcl/18s/microscopy**



Kermarrec et al. Freshwater Science 2014

1st proofs of concept (2)

Funded by European U

1st proofs of concept (2)

Funded by European U

1st proofs of concept (2)

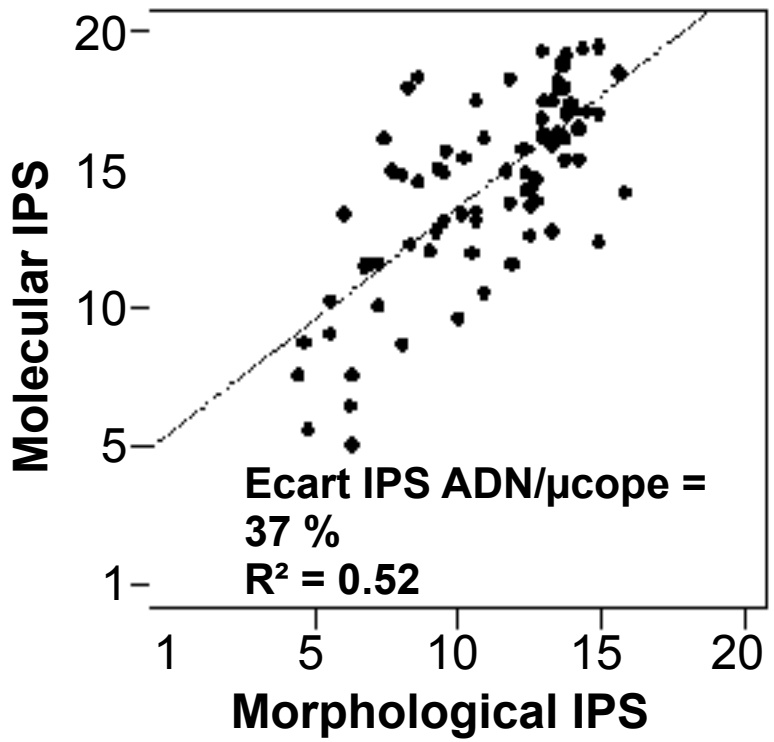
> 2011: Mayotte became a French department after a referendum

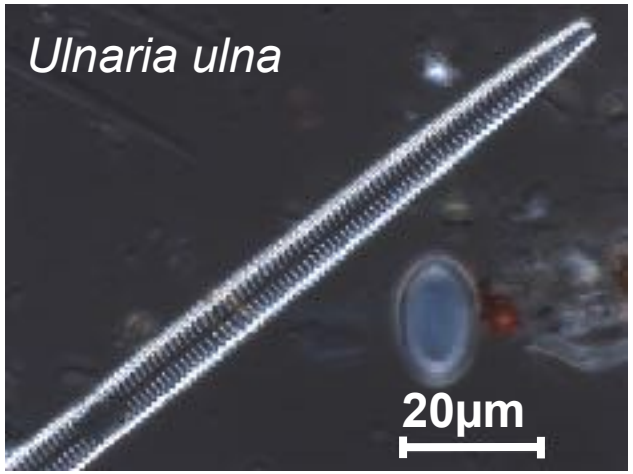
> Part of EU, and all directive must be applied, including the WFD

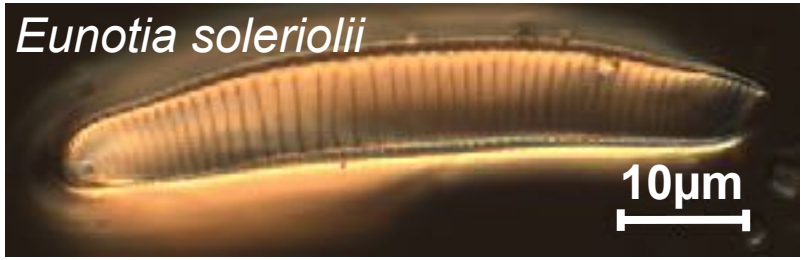
> Our team was in charge of developing its monitoring system and to test alternative methods, including DNA metabarcoding

Funded by European Union

Calculation of index values from microscope and DNA inventories



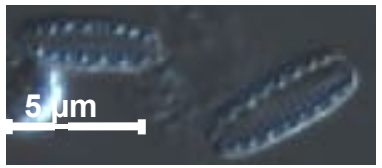




Dominant in DNA inventories

- Good correlation but…
- Values of IPS $_{DNA}$ > IPS $_{morpho}$
- Dominant taxa are different (*Eunotia* et *Ulnaria*)
- Systematically big species were dominant in DNA

Nitzschia inconspicua



Dominant in morphological inventories

Funded by European Union

BIOLAWEB

Schedule

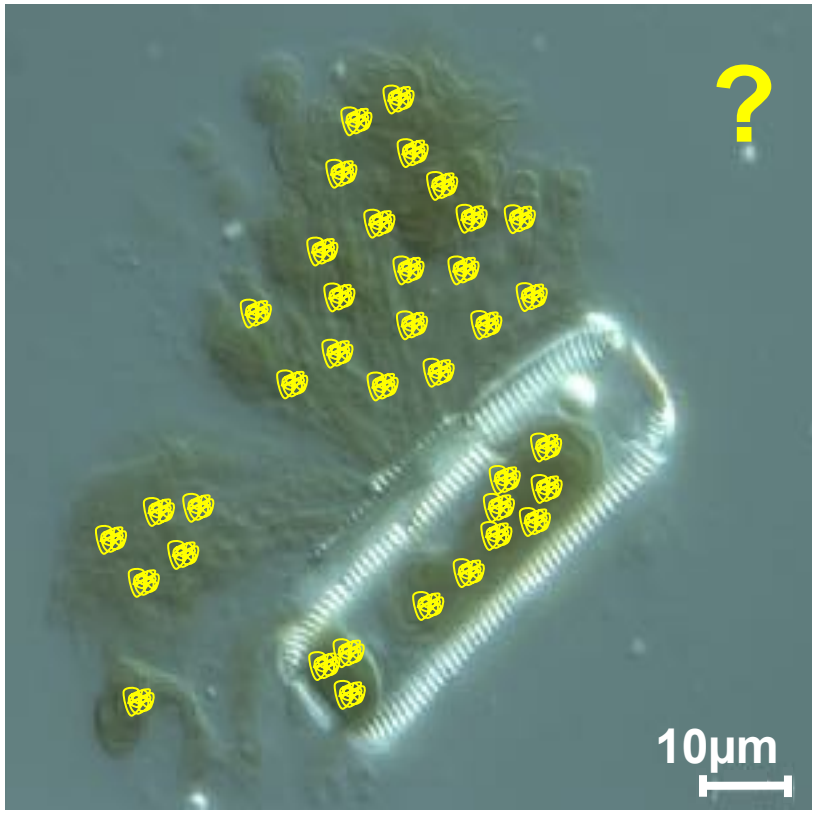
1- reference library

2- 1st proofs of concept

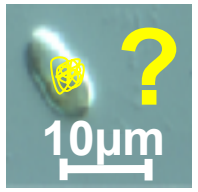
3- Impact of biovolumes

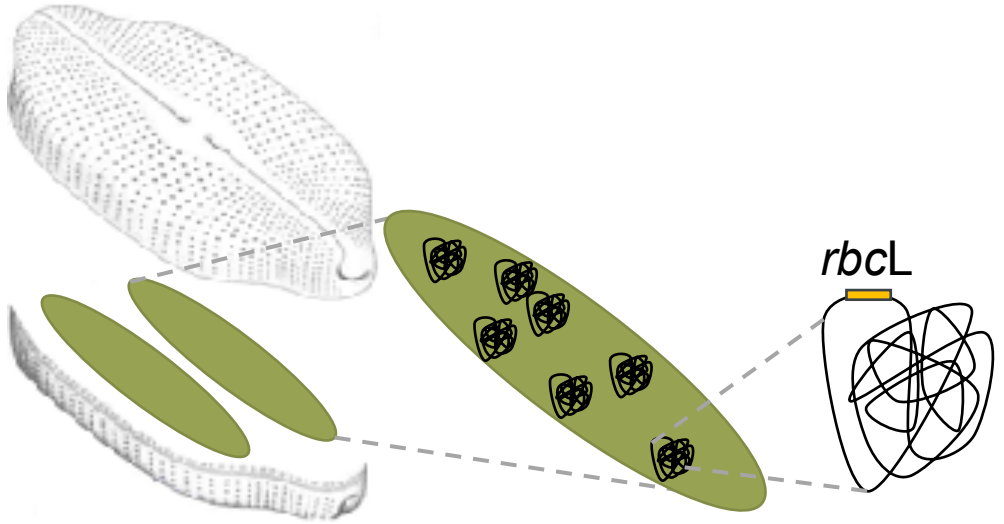
Link between gene copies and cell biovolume?

Pinnularia viridiformis (≈ 13,700 µm³)



Achnanthidium minutissimum (≈ 80 µm³)



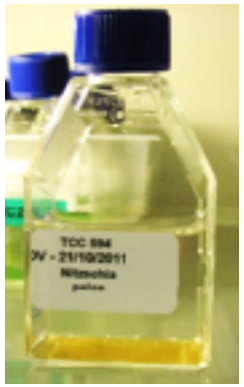


Chloroplast nb. X Genome nb. X *rbc*L nb.

Vasselon, V., Bouchez, A., Rimet, F., Jacquet, S., Trobajo, R., Corniquel, M., Tapolczai, K., Domaizon, I., 2018. Avoiding quantification bias in metabarcoding: Application of a cell biovolume correction factor in diatom molecular biomonitoring. Methods in Ecology and Evolution 9, 1060–1069. https://doi.org/10.1111/2041-210X.12960

Link between gene copies and cell biovolume?

- ❖8 diatom cultures of different species
- ❖Cultivated in triplicates
- ❖Evaluation at 7 different dates :
  - nb of rbcL copies with qPCR
  - cell concentration in the culture



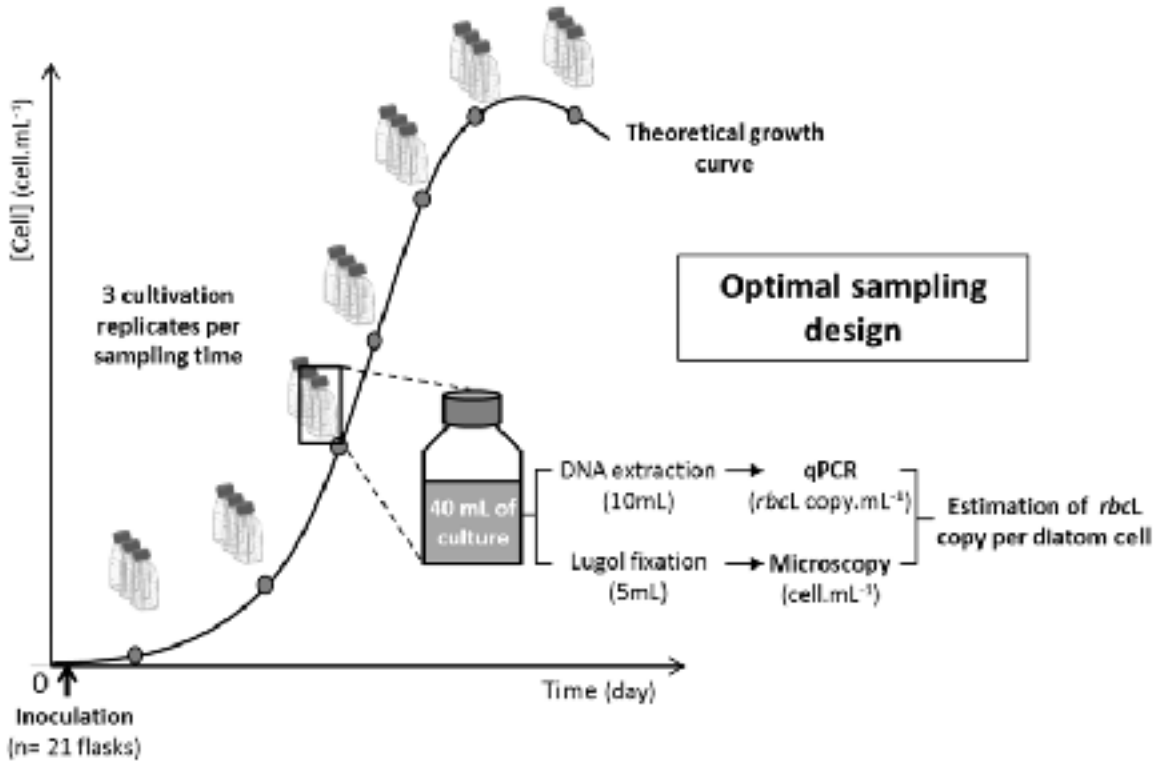
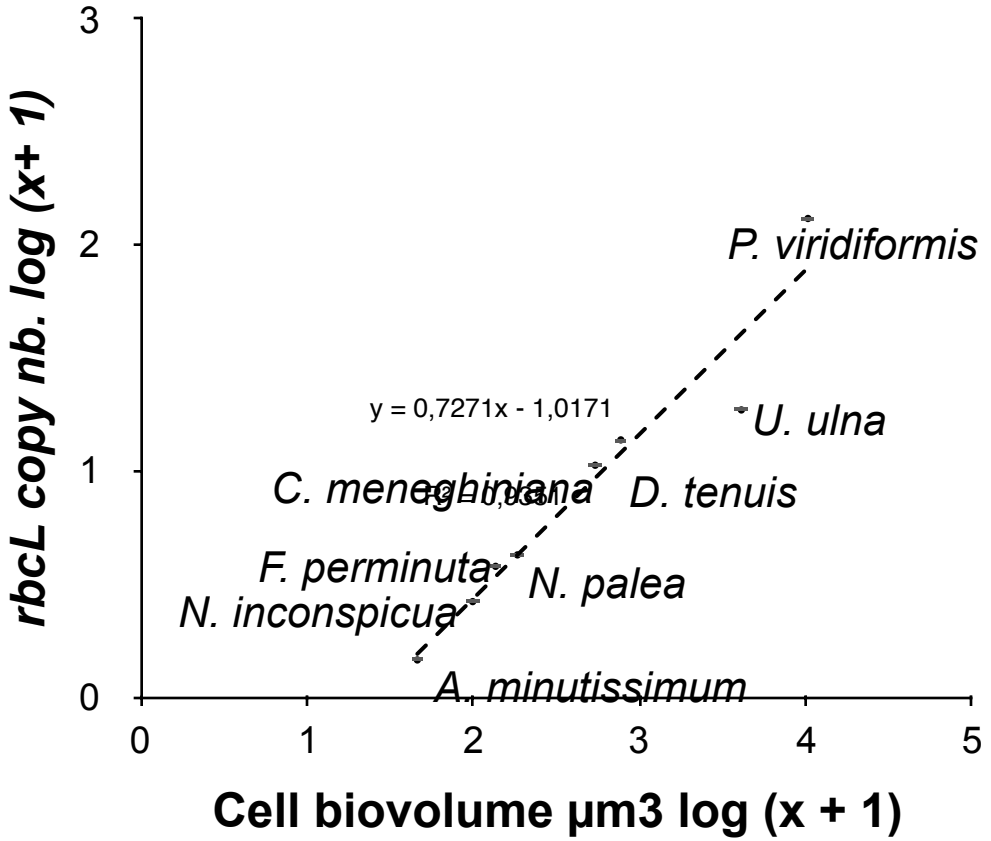


Figure 25 – Experimental design applied to the 8 diatom species. After the inoculation of 21 flasks containing 40mL of DV media, diatom culture growth was followed at 7 sampling time (from T0 to T6) and analysis was performed in triplicate (3 flasks per sampling time).

Link between gene copies and cell biovolume?

- There is a relation between cell biovolume and nb of copies
- It is a log correlation



Link between gene copies and cell biovolume?

❖Based on this correlation, we proposed a correction factor to transform DNA inventories to make them similar to microscope inventories

❖This enable to calculate indices based on the transformed DNA inventories

❖Each species has a specific correction factor based on its biovolume (available in Diat.barcode):

	Biovolume μm3	Correction factor CFv2
Achnanthidium minutissimum	76	2,18310673
Pinnularia acrosphaeria	5500	75,7149868

Let’s transform a DNA inventory into a microscopy-like inventory!

Let’s transform a DNA inventory into a microscopy-like inventory!

Taxon	Frustule %	Biovol. (µm3)	Sequences %			
Achnanthidium minutissimum	75	76	15			
Amphora pediculus	13	72	2			
Navicula cryptotenella	10	386	10			
Melosira varians	2	14515	73			

Let’s transform a DNA inventory into a micrscopy-like inventory!

DIAT.BARCODE

Taxon	Frustule %	Biovol. (µm3)	Sequences %	CFv2		
Achnanthidium minutissimum	75	76	15	2,18		
Amphora pediculus	13	72	2	2,13		
Navicula cryptotenella	10	386	10	5,62		
Melosira varians	2	14515	73	282,93		

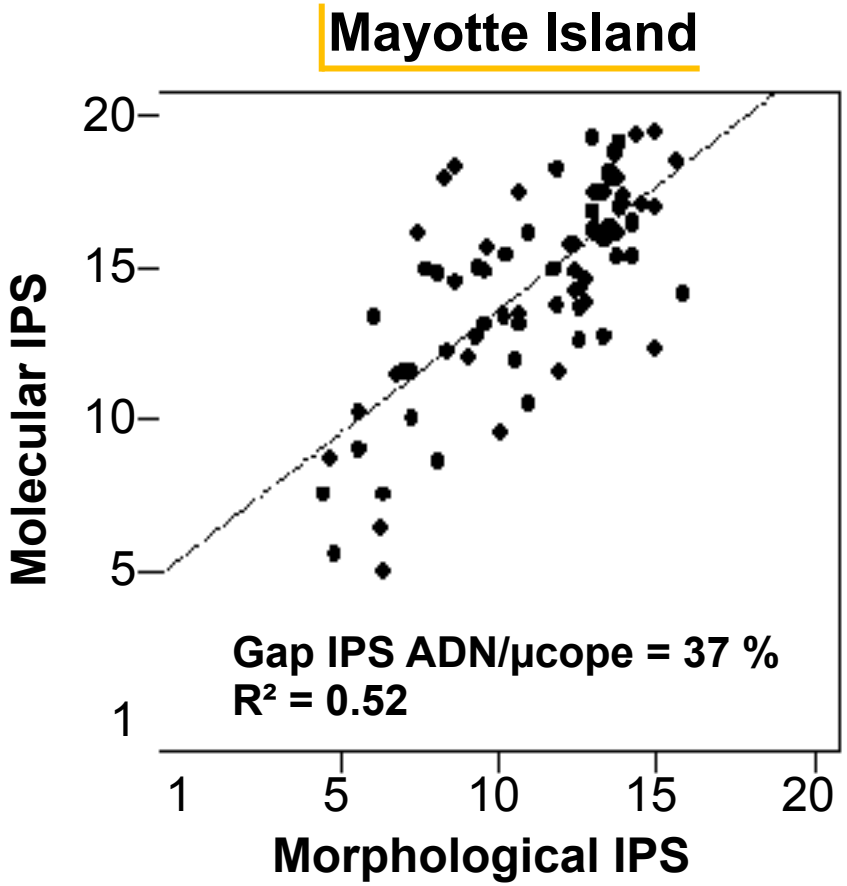
Funded by European Union

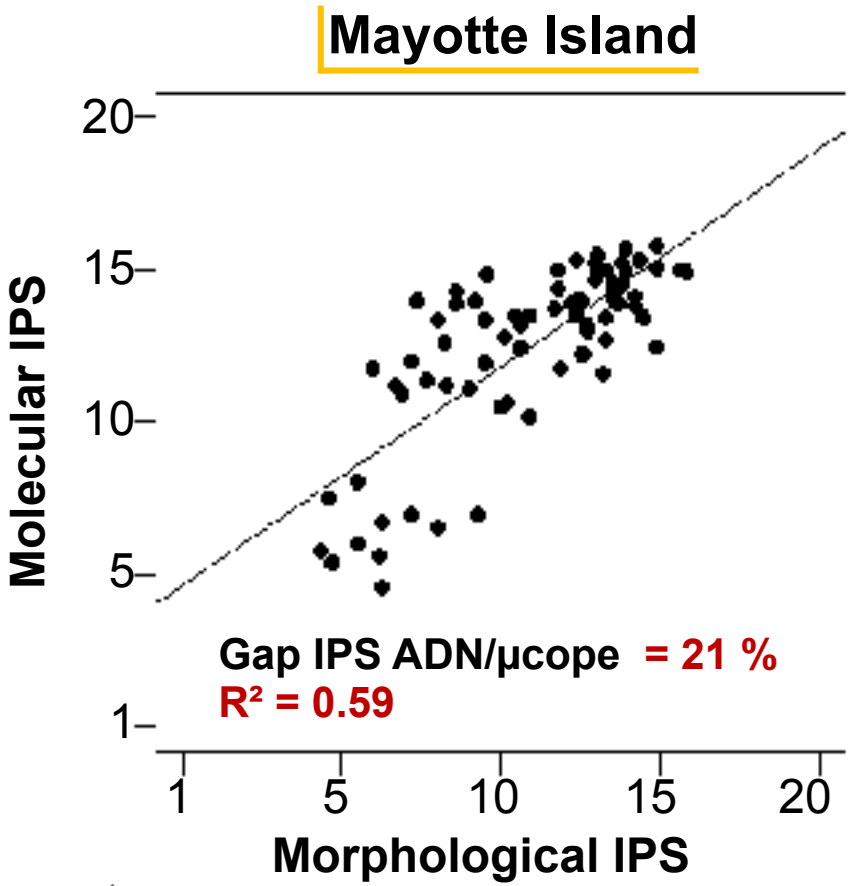
Let’s transform a DNA inventory into a micrscopy-like inventory!

DIAT.BARCODE

Taxon	Frustule %	Biovol. (µm3)	Sequences %	CFv2	Seq % / CF	Sequences transf. %
Achnanthidium minutissimum	75	76	15	2,18	6,87	70
Amphora pediculus	13	72	2	2,13	0,94	10
Navicula cryptotenella	10	386	10	5,62	1,78	18
Melosira varians	2	14515	73	282,93	0,26	3

Application on Mayotte rivers

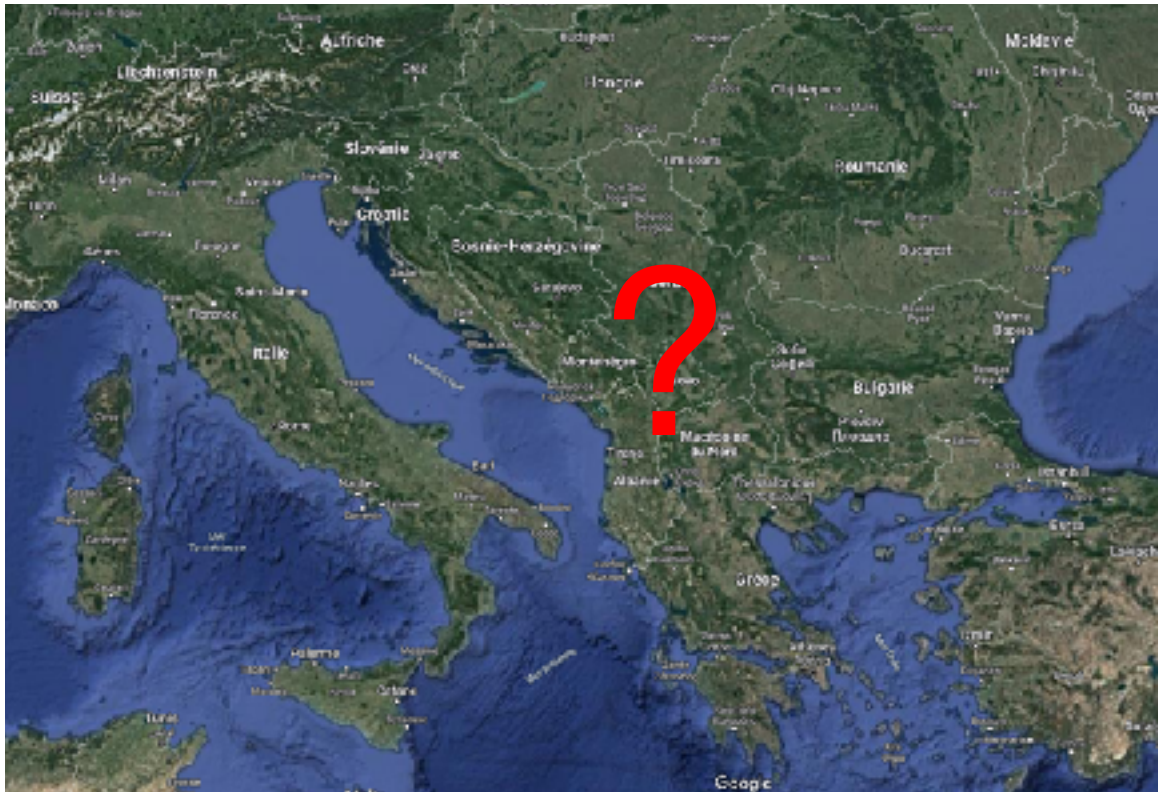


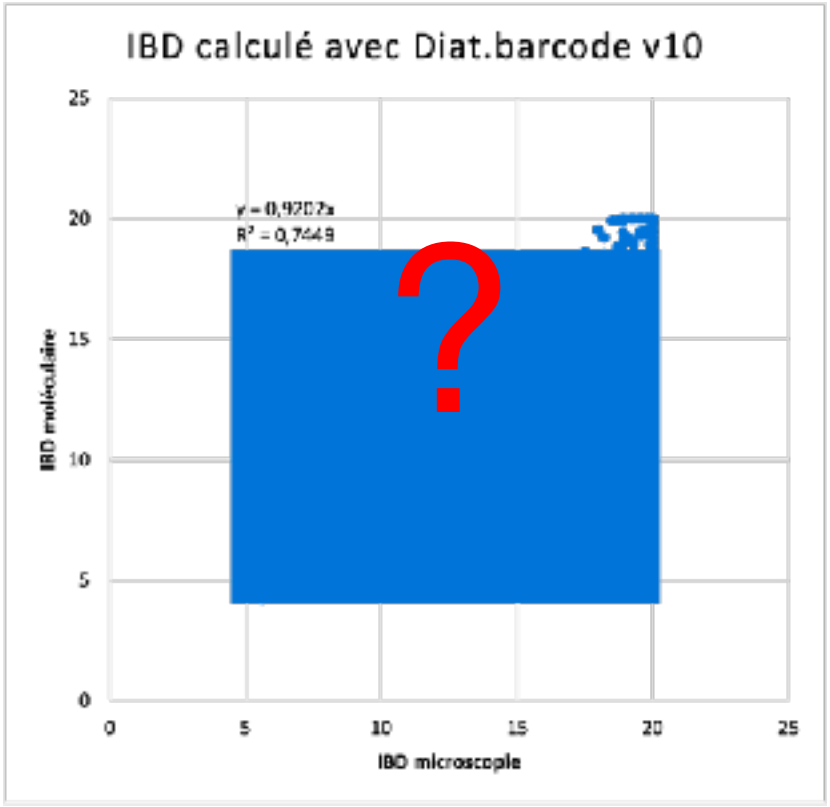


CFv1

Funded by European Union

Application to balkan region?





BIOLAWEB

Questions?