



Report on a Workshop:

Developing DiatomBase

20-22 March 2017

Prague, Czech Republic

Table of Contents

1. Background.....	3
2. Aims of the Workshop.....	4
3. Workshop Agenda.....	4
4. Participants.....	5
5. Outcomes.....	6
6. Products.....	8

1. Background

There has been a long history on the development of catalogues of diatom names, beginning in the early 19th century with lists of several hundred names. Published catalogues of diatom names occurred throughout the latter half of the 19th century, culminating in the 8-volume set of Van Landingham's Catalogue, which contained over 44,000 names.

In addition to these published resources, large compendia of unpublished resources were developed and maintained. Paul Silva who attempted to track the names of all algae, developed the *Index Nominum Algarum* (INA) which contained over 220,000 names of algae, including about 20,000 diatom names. The "New Species File" of diatom names was developed by Ruth Patrick at the Academy of Natural Sciences of Philadelphia, tracking names of diatoms after 1932 (the date of the last volume of a catalogue compiled by Mills, "Mills Index"). These resources remained in analog formats until recently.

The Catalogue of Diatom Names (Fourtanier and Kociolek 2011) was built as an on-line resource, integrating and cross-checking the above resources, and including verification of information across 12,500 references. Over 64,000 names of diatom taxa were included in the Catalogue of Diatom Names, which lists the names, describers, publication and homotypic synonyms for the names. It was last updated in 2011. *AlgaeBase* (Guiry & Guiry 2017) is an on-line source of names for all algae, as well as other information, and it has been the basis for the 45,000 diatom names for WoRMS; *AlgaeBase* was developed in part from the INA. *AlgaeBase* is a privately developed product.

In response to concerns about the current version of *AlgaeBase* in terms of diatom names, and about prospects about its future, the team at VLIZ contacted Kociolek about the transfer of the Catalogue of Diatom Names to VLIZ, the updating of the Catalogue to 2017, and a model of community participation and support that would ensure its continued updating, relevance and continuation. VLIZ could provide significant technical help, and provide long-term support in a variety of ways, just as it has done for a number of other groups with similar missions and aims.

With the concurrence of Dr. E. Fourtanier and the California Academy of Sciences, the Catalogue of Diatom Names database was transferred to VLIZ and integrated into the Aphia database structure. Ensuing discussions about ways to update the database, and to leverage the work of *AlgaeBase* and continue some interaction with that project were begun. In addition, other projects such as the Registry of Algal Names, New Taxon File at the Academy of Natural Sciences (an on-line continuation of Dr. Patrick's New Species File), among others, saw possibilities for collaboration. DiatomBase was also being envisioned to include more than names, but to be a resources for ecology, identification, and literature. A group of 20 editors, plus Kociolek serving as the inaugural Editor-in-Chief, from 14 countries, were identified and agreed to serve to help identify, resolve move the project forward. To introduce them to the project, discuss with them the short- and long-term needs and goals for the project and to give them practical hands-on experience working with editing and entering

names in the new product, a workshop was held. The meeting convened in Prague, in March 2017 ahead of the Central European Diatom Meeting. The meeting was convened at the Kaiserštejnský Palác - <http://kaiserstejnsky-palac.cz/en/contact.html> in Prague.

2. Aims for the workshop

A series of practical aims for the workshop were identified. They included:

- *Introducing the editors to the history, potential and challenges of developing and sustaining a project like DiatomBase.
- *Determining the priorities for the DiatomBase project (in the short-, mid- and long-terms).
- *Discussing and agreeing to the different roles and responsibilities for editors. In addition, a timeline was to be set up, to help set goals and move things forward.
- *All participants would learn how to use the online editing interface and make progress in editing their areas (taxonomic, geographical) of responsibility. During the focused hands-on editing sessions, help would be provided to ensure the priorities are tackled and that all editors know how to add and change information to DiatomBase.

3. Agenda

Monday 20 March 2017 (day 1)

- Welcome [Pat Kociolek]
- Practical information
- Round table [all]
- History, background, status and future of the Catalogue of Diatom Names as DiatomBase [Pat Kociolek]
- General introduction to VLIZ, Aphia & WoRMS [Leen Vandepitte]
- General introduction to VLIZ, Aphia & WoRMS (continued) [Leen Vandepitte]
- WoRMS online editing [Leen Vandepitte]
 - o How to edit online
 - o Overview of existing tools & functionalities
 - o Alternatives to online editing, through Excel templates
 - o Priorities: giving examples on how to deal with each priority type
- WoRMS online editing (continued) [Leen Vandepitte]
- Priorities & responsibilities for DiatomBase, including examples
- Priorities & responsibilities for DiatomBase (continued)
- Hands-on online data editing [all]

Tuesday 21 March 2017 (day 2)

- Restatement of outcomes of day 1
- Priorities for entry & consistency in editing – the online editing manual

- Hands-on online data editing (continued) [all]
- WoRMS online editing (continued)
- Hands-on online data editing [all]
- Hands-on online data editing [all, continued]
- Possibilities for DiatomBase
 - o Brainstorming
 - o Evaluation
 - o Fundraising opportunities

Wednesday 22 March 2017 (day 3)
--

- Restatement of outcomes of day 2
- Hands-on online data editing [all, continued]
- Moving forward
- Sum up tasks, duties & responsibilities
- Short-term, mid-term & future prospects for DiatomBase

4. Participants: *The following is the team of editors for DiatomBase, including their affiliations. All were invited to the Workshop. Names with an “*” indicate those editors who attended.*

***Kociolek, J.P.:** Museum of Natural History and Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, CO 80309 USA

***Balasubramanian, K.:** Biodiversity and Paleobiology Group, Agharkar Research Institute, Pune 411004, India

Blanco, S: Laboratorio de Diatomología / Diatom Lab, The Institute of the Environment, La Serna, 58 - 24007 León, Spain

Coste, M.: Irstea-EABX-CARMA, Unité de Recherche – Ecosystèmes Aquatiques et Changements Globaux (EABX) Equipe Contaminants Anthropiques et Réponses des Milieux Aquatiques (CARMA), 50 Avenue de Verdun, 33610 Cestas, France

***Ector, L.:** Luxembourg Institute of Science and Technology, 5 avenue des Haurs-Fourneaux, L-4362, Esch-sur-Alzette, Luxembourg

***Liu, Y.:** College of Life Science and Technology, Harbin Normal University, Harbin, 150025, China

***Kulikovskiy, M.:** Department of Water Plants Taxonomy and Geography, I.D. Papanin Institute for Biology of Inland Waters, Russian Academy of Sciences, 152742 Yaroslavl, Nekouz, Borok; Russia

Lundholm, N.: Natural History Museum of Denmark, University of Copenhagen, Sølvgade 83S, 1307 Copenhagen K, Denmark

***Ludwig, T.:** Department of Botany, Universidade Federal do Paraná, Brazil

***Potapova, M.:** Academy of Natural Sciences of Drexel University, 19th & The Parkway, Philadelphia, PA 19104, USA

***Rimet, F.:** NRA - UMR Carrtel, 75 av. de Corzent - BP 511, FR-74203 Thonon les Bains cedex, France

***Sabbe, K.:** Protistology & Aquatic Ecology Lab, Ghent University, Krijgslaan 281-S8, 9000 Ghent, Belgium

***Sala, S.:** División Ficología, Facultad de Ciencias Naturales y Museo, Universidad Nacional de La Plata, Paseo del Bosque s/n. 1900, La Plata, Argentina

***Sar, E.:** División Ficología, Facultad de Ciencias Naturales y Museo, Universidad Nacional de La Plata, Paseo del Bosque s/n. 1900, La Plata, Argentina

***Taylor, J.:** North-West University, Private Bag X 6001, Potchefstroom 2520, South Africa

***Van de Vijver, B.:** Botanic Garden Meise, Department of Bryophyta & Thallophyta, Nieuwelaan 38, B-1860, Belgium and Department of Biology-ECOBIE, University of Antwerp, Universiteitsplein 1, B-2610 Wilrijk, Belgium

***Wetzel, C.E.:** Luxembourg Institute of Science and Technology, 5 avenue des Haurs-Fourneaux, L-4362, Esch-sur-Alzette, Luxembourg

***Williams, D.M.:** Department of Life Sciences, the Natural History Museum, Cromwell Road, London, SW7 5BD, U. K.

***Witkowski, A.:** Palaeoceanology Unit, University of Szczecin, Mickiewicza 18, PL-70 – 383 Szczecin, Poland

***Witkowski, J.:** Department of Geology and Paleogeography, University of Szczecin, Mickiewicza 18, PL-70 – 383 Szczecin, Poland

In addition, other participants included **Wolf-Henning Kusber** (from Botanical Garden Botanical Museum, Berlin, Germany) and **Leen Vandepitte, Wim Decock, and Bart Vanhoorne** from VLIZ.

5. Outcomes

1. Activities over the recent past were reviewed, analyzed and discussed. Included among these outcomes were

*Catalogue of Diatom Names transferred to VLIZ - Done

*Integrate Catalogue into APHIA database - More or Less Done!

*Meeting of Editorial Board and Editors in Belgium @ CEDM/Prague-March 2017!

*Preparing for Launch of DiatomBase after IPC-Szczecin: We are on track to have the website go live in 1 week after this meeting (pending feedback from the presentation).

2. Priorities for the Short and Long-Term were identified and discussed. Among them were

1. Check discrepancies between AlgaeBase & DiatomBase

- Spelling variations
 - Differences in authority and/or publication date
2. Check pre-2011 names from AlgaeBase which are not in CAS
 - Misspellings?
 - Actually missing from CAS?
 3. Check post-2011 names from AlgaeBase which are not in CAS.
 4. Get accepted names right
 - Who has decided on ‘accepted’ status? Is this correct?
 - Back up statements with literature references
 5. Transfer “uncertain – unassessed” to accepted/unaccepted
 6. Add new names (post-2011, incl. original descriptions)
 7. Complete environment flags (cfr. inheritance => assign on higher level)
 8. Agree on Classification System we will adopt for the Project

Numbers 1, 2, 3 and 8 on this list were identified as the top priorities, that were felt to be realistic.

VLIZ representatives identified numbers 5 and 7 as top priorities in the short-term, especially as they relate to WoRMS. This sparked a wide-ranging discussion, especially as it related to “accepted/unaccepted” names. As this editorial board understood the concept, it was suggested that rather than arbitrary decisions by editorial board members as assigning these terms to names, that there needs to be identified documentation for decisions relating to these designations. Such documentation might be the most recent revisionary studies. However, such studies in diatoms are wanting for many groups, and determinations of what are accepted names would become original research. It was recognized that while a priority for VLIZ, this was not a part of the Catalogue of Diatom Names (only the tying together of homotypic synonyms); no determinations of what names were “accepted” was attempted. Names indicated as accepted in AlgaeBase had no documentation. This will require a major undertaking for the project, and will need to be done on a group-by-group basis.

It was determined that communications will be an important aspect of this work. Venues for presentations about DiatomBase included:

Central European Diatom Meeting, Prague, Czech Republic, 23-25 March, 2017

International Phycological Congress, Szczecin, Poland, 13-20 August, 2017

South American Phycological Congress in Columbia, 5-10 November, 2017

A presentation about the project would be distributed to editorial board members for their use at national and regional meetings.

6. Products

Several products were identified as outcomes from the Workshops. These included:

The aforementioned presentation that Editorial Board members can use for presentations. One was developed for the CEDM in Prague, and updated and refined. At this writing, the presentation will be made at the meeting in Szczecin, and then distributed to editorial board members.

The DiatomBase Database is of course the main product of the work. For an August 2017 debut, this includes:

*Design of the page, including a color scheme and images

*Write-Ups of topics related to diatoms and the DiatomBase website

*Some discussion has ensued as to whether DiatomBase should have DOI assigned to it. Several arguments have been in favor of adopting DOI's for DiatomBase. These include:

- a. The DOI system can serve to identify and locate a specific version of a dataset
- b. Datasets to which a DOI is assigned should stay persistently accessible.
- c. Although a DOI is being assigned to a dynamic database, claims based on the data can always be checked as the date of consultation needs to be added to the citation and can be traced back to the archived versions of WoRMS.
- d. Minimum metadata - including the names of the data author(s) – needs be associated with the DOI during registration. Having the data authors associated with the DOI will assure that they can receive the appropriate credit and recognition for their work.
- e. DataCite will store the associated metadata and can disseminate it to other initiatives like the Data Citation Index (Thomson Reuters), which will link the datasets to publications citing them.
- f. By assigning a DOI to the dataset, the data publisher attests that the dataset answers to certain quality standards and contains all the necessary metadata to make the dataset usable by other scientists. This will encourage other scientists to use it for their analyses and then cite your dataset in their publications.