Nova Hedwigia, Beiheft 136, 325–331
Stuttgart, May 2010

Article

The Giraffe Pipe database project: A web-based database for siliceous microfossils from a freshwater Eocene waterbody

Peter A. Siver[1], Jeffry M. Pelczar[1], Anne Marie Lott[1], Andrzej Pisera[2]

[1]Department of Botany, Connecticut College, New London, CT, USA 06320
[2]Institute of Paleobiology, Polish Academy of Sciences, Warszawa, Poland

With 2 figures

Abstract: We are investigating the siliceous microfossil remains of an ancient Middle Eocene maar lake situated in the Northwest Territories, Canada. A drilled core from the site yielded over 50 meters of lacustrine sediments containing exquisitely preserved microfossils. As part of this project we plan to survey and inventory all siliceous microfossils found in the Giraffe Pipe core and make the results available through publications and the Web. The Web component of the project will include a searchable image database, the goals of which are to: 1) disseminate information on Eocene freshwater ecosystems, 2) foster the development of ideas on the biodiversity, paleoecology and evolution of siliceous organisms, 3) aid identification of specimens and, 4) distribute information among members of our research team. The database will include chrysophyte scales, bristles, cysts, diatoms, testate amoebae, sponge spicules, heliozoan scales and a category of unknown taxa. The project has been developed using FileMaker Pro ver. 9, a versatile relational database software package that includes a module for distribution of data via the Web. The database, now available for use by the scientific community, currently contains 2500 records, including thumbnail and larger-sized images of each specimen. Users can view all records or conduct custom searches in order to extract specific information. In the future, we would like to add information on modern analogs, make links to other databases, provide a discussion forum page and expand the project to include results from other Cenozoic sites. Comments and suggestions for improvement are welcomed.

Key words: scaled chrysophytes, database, diatoms, Eocene, Giraffe Pipe

Introduction

Investigations of Middle Eocene sediments representing freshwater environments contained within the Giraffe Pipe kimberlite crater have proved invaluable for the study of scaled chrysophytes (Siver & Wolfe 2005a, b) and diatoms (Wolfe et al. 2006, Siver & Wolfe 2007). The abundant numbers of microfossils contained within the Giraffe Pipe core are exquisitely preserved, span roughly eight million years of geologic history and should eventually provide significant insights concerning evolution of a number of protistian lineages, numerous descriptions of new taxa and processes concerning the development of freshwater ecosystems. The Giraffe Pipe waterbody,

www.borntraeger-cramer.de
DOI: 10.1127/1438-9134/2010/0136-0325
1438-9134/2010/0136–0325 $ 1.75

situated near the Arctic Circle in northern Canada at coordinates 64º 44' N and 109º 45' W (Wolfe et al. 2006), existed during a very warm period in the history of the Earth known as the Cenozoic hothouse (Zachos et al. 2001, 2008). As such, this waterbody provides a likely ancient analog for how Arctic freshwater environments may respond to future increases in atmospheric greenhouse gases and subsequent warming events.

As part of a long-term biotic inventory project funded, in part, by the National Science Foundation we are actively involved in describing and cataloging the siliceous microfossils contained within the Giraffe Pipe core. Due to the substantial amount of lacustrine sediment contained within the core, coupled with the impressive numbers of microfossils, we anticipate that the project will take a number of years to complete. Currently, our survey includes chrysophyte cysts, scales and bristles, diatoms, testate amoebae scales (e. g. *Euglypha*) and sponge spicules (megasclerids, microsclerids and gemmusclerids) (Fig. 1). Time permitting, we also hope to include remains of non-siliceous algae (e. g. *Botrycoccus*), heliozoans as well as unknown microfossil forms (Fig. 1).

One objective of the Giraffe Pipe project is to develop and actively maintain a Web database of all microfossil forms uncovered within the core that can be searched using various combinations of characteristics and variables. As of June 2008, the database, referred to as the "Giraffe Pipe Database", has become available online for use by the scientific community. The purpose of this

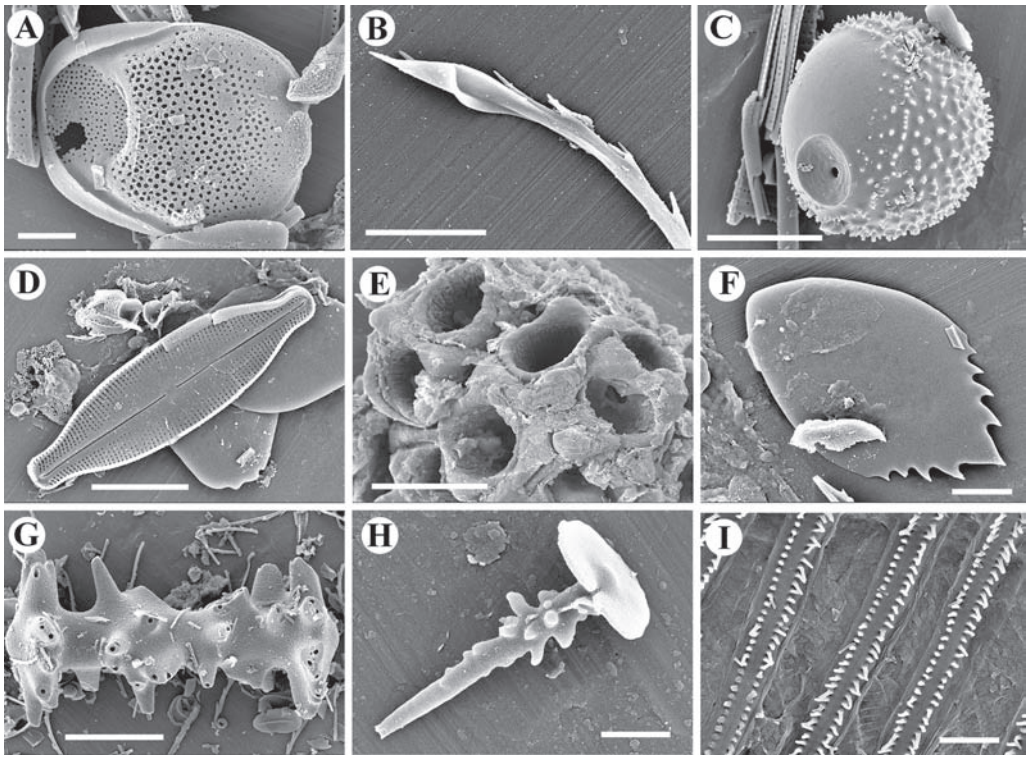

Fig. 1. Types of microfossils found in the Giraffe Pipe core that will be included in the database project. A) Chrysophyte scale, possibly from a *Mallomonas* cell. B) Chrysophyte bristle. C) Chrysophyte cyst. D) Valve of a raphe-bearing diatom. E) Remains of what is believed to be a *Botryococcus* colony. F) Scale from a testate amoebae related to *Euglypha*. G.) Sponge gemmusclerid. H) Heliozoan spine scale. I) Piece of an unknown organism, probably an invertebrate. Scale bars = 1 µm (H), 2 µm (A–B, F, I), 5 µm (C–E) and 10 µm (G).

communication is to describe the development and design of the database, its contents, a set of instructions on how to use the site and future plans to upgrade and improve the facility. Further, we hope this communication will result in suggestions and advice from users on how the site could be improved in order to enhance our dissemination of information on Eocene ecosystems.

Goals of the database

An overarching goal of the database is to provide a readily available means of distributing images and associated data for Giraffe Pipe microfossils to the scientific community. Although anyone will be able to access the database, we anticipate that it will be used mostly by phycologists, protistologists, paleontologists, geologists and environmental scientists, especially those interested in biological diversity, evolutionary biology, development and subsequent ontogeny of ancient freshwater ecosystems and responses to previous hot house climates. More specifically, the database will: 1) aid in dissemination of data among the team of scientists working on the Giraffe Pipe core, 2) facilitate identification of individual microfossils, 3) enhance development and discussion of hypotheses and, 4) provide a means of distributing additional information in support of peer-reviewed publications. For example, although the database has just recently been placed online, it has already been instrumental in selecting samples for further study of the diatom genus *Aulacoseira*.

Details of the database

The database was constructed and is maintained using the relational database software package FileMaker Pro ver. 9. FileMaker Pro is a versatile program that supports customized development of layout pages used for selection of search criteria and for the display of results. Multiple result layouts can be linked to a given search. For example, current searches are programmed to deliver initial results using a "List" format where a few details of each record are displayed on successive lines down a page. The user can scroll down the list and observe thumbnail images of each record. Currently, each record is further linked to three additional result pages accessed by clicking on function buttons contained within the "List" layout (more details are provided below). One option, referred to as "View Image Only," allows the user to view larger-sized images of each microfossil. A second option, the "View Full Record" layout, will display all information about the record, including such data as where it was found in the core, size, magnification, assigned name and eventually, links to other data (from other sources). The third option, the "View Summary Statistics" page, provides basic summary statistics for the set of records uncovered in the search, including the minimum, maximum and mean sizes of the microfossil and depths where the organisms were found. The standard deviation is also provided. Users can readily move from one type of result page to another through the use of scripted buttons.

Our database takes advantage of many of the different types of fields available in FileMaker Pro, including ones for text, numbers, calculations, containers and summary functions. In addition, a layout can include what is referred to as a "web portal." A potential problem with large databases containing numerous images is that the size of the database can become overwhelming, resulting in a large amount of time necessary to complete a search. This is especially a concern if the database contains thousands of images. A "web portal" simply contains an address or link to, for example, an image. The images used in the database can be stored on a hard drive outside of the database, but linked to it by the "web portal" address. As a result of the use of "web portals", the issues regarding the size of the database and the time needed to complete a search are largely overcome. All images associated with the Giraffe Pipe database are maintained outside of the database and accessed through the use of "web portals."

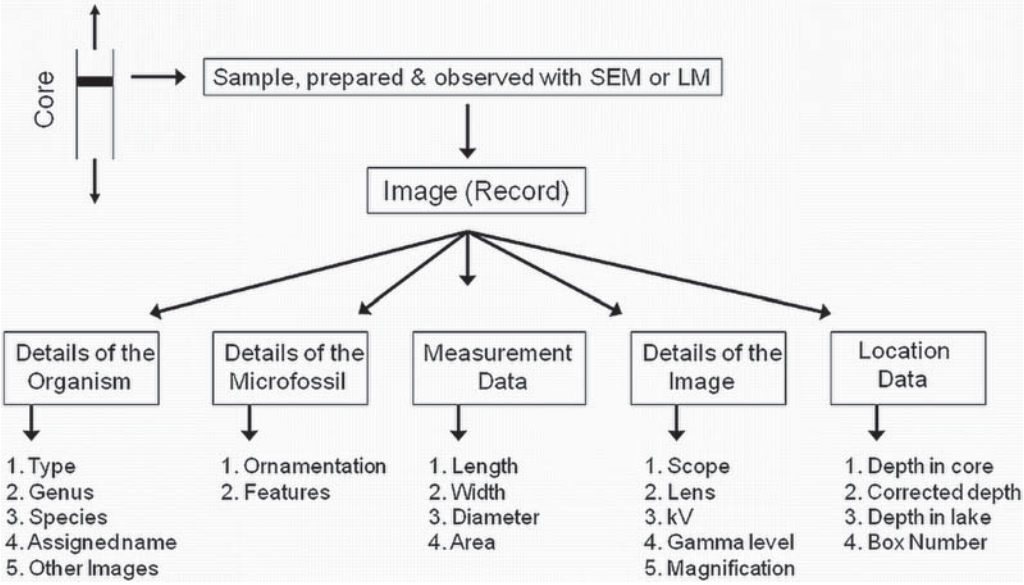

Fig. 2. Flow diagram of the five sets of information associated with each image record.

Information about the Giraffe Pipe project and access to the database is available through the website, http://silicasecchidisk.conncoll.edu. The Silicasecchidisk is a website initiated ten years ago by P.A. Siver as a means of distributing results from numerous research projects on freshwater ecosystems carried out at Connecticut College. Initially, the Silicasecchidisk contained physical, chemical and biological data on lakes within the state of Connecticut. The site has since grown to include data on hundreds of waterbodies along the east coast of North America, inventories and image libraries of scaled chrysophytes and diatoms observed in the study lakes, a section on newly described species and a reference database for the study of chrysophytes. In addition, the site has a suite of educational tools for the study of algae, including a non-hierarchal key linked to web pages containing high-end images and digital movies (Shayler & Siver 2006). The site is linked to multiple large searchable databases, including the one for Giraffe Pipe.

Contents of the Giraffe Pipe database

Image micrographs form the basis of a record. Each record contains five sets of information, including details on the organism, morphometric information, elements about the microfossil image, location observed within the core and information on how the image was obtained (Fig. 2). In addition to a web portal for the image (stored in jpeg form), the type of microfossil (e. g. a chrysophyte scale), genus, species, assigned name (e. g. GPCyst001) and links to other images of the same specimen will be recorded. Morphometric data will also be added for each microfossil image including length, width, diameter (if appropriate) and potentially other measurements such as area, rib density and striae density (for diatoms). Measurements will be provided only for whole specimens of individual microfossils so that summary statistics will be most meaningful. Details on the type of microfossil are also provided so that more customized searches can be performed by

the user. For example, whether a scale has a dome, V-rib, wing, spine or secondary siliceous layer is noted. Likewise, the type of collar and surface ornamentation for each cyst will be recorded.

Location data, including the core box number, precise level within each box, depth within the core (uncorrected), corrected vertical depth and estimated depth within the lake phase, is also provided. Users are encouraged to review the information on the Silicasecchidisk concerning how the Giraffe Pipe core was taken, stored and sampled. Briefly, the core was drilled at a 47 degree angle, cut into 1.5 m sections and stored within boxes each of which contains three 1.5 m sections (i. e. 4.5 m of sediment per box). The uncorrected depth refers to the depth along the original core (i. e. drilled at the 47 degree angle). The corrected depth refers to the true vertical distance below ground level. Lastly, the depth of the sample within the original lake is also provided and from an ecological point of view is probably the most meaningful. Higher box numbers reflect deeper sections of the core and each length of core within a given box is numbered "1", "2" or "3" where length "1" represents the younger section and "3" the oldest section. Lastly, the distance in cm along each length of core is provided. Thus, a location of "11-3-31" represents a sample from box 11 that is 31 cm down along length of core #"3".

Lastly, details of how the micrograph was taken are also added to each record, including the person who took the image, whether it represents a light (LM) or scanning electron micrograph (SEM), magnification and microscope model. For LM images, the type of optics (e. g. brightfield, phase contrast, differential interference contrast or reflected interference contrast) and lens used to obtain the image are also stored. For SEM images, the gamma level, kV and length of the scale bar are provided.

Currently, there are over 2500 records in the database, including 403 cysts, 701 chrysophyte scales, 158 bristles, 641 diatoms, 126 testate amoebae, 161 sponges, 51 images of fractured mudstone showing *in situ* microfossil remains and 325 other taxa (e. g. heliozoans, *Botryococcus*, pollen grains, invertebrate remains and unknowns). Now that the database structure is in place, images taken with LM and from other researchers will soon be added. In addition, images taken for specific projects, such as geomorphic morphometric studies of testate amoebae and chrysophyte scales, will eventually be added to the database. As a result, we anticipate that the number of records will double within the next year and most likely triple soon thereafter.

Performing a Search & Result pages

Access the Giraffe Pipe project homepage using the URL http://silicasecchidisk.conncoll.edu/Giraffe_Pipe_homepage.html and click on the "Image Data Base" button. This brings you to the sign-in page. Select the "Guest Account" option and click on the "login" button. The next page, entitled "Giraffe Pipe Image Database", provides a short set of instructions on performing a search, a "Begin Search" button to open the "Search Form" and a button to view all records. The "Search Form" provides the fields on which searches can be performed, along with instructions explaining each variable. Numerous options are available and the user can run a search by filling in one or more fields. A search can be as simple as selecting chrysophyte cysts or more complex, for example chrysophyte cysts over 20 µm and found at water depths less than 4 m. Most of the text fields provide pop-up menus for easy selection of parameters. In some cases, for example the field where the core box number can be selected, more information is given on the project website.

The most important instruction warns the user to navigate within the database software when moving between results pages and the search form. As with any relational database software with search and find capabilities, navigation needs to be confined within the software environment. That is, using the "Back" function of the browser should be avoided as its use can result in incomplete searches. For example, assume a search is performed that yields a subset of records, X. If the browser "Back" key is then used to return to the "Search Form" any subsequent searches will only be performed on subset X. This may be fine if the user wishes to refine the original suite of

search parameters to yield a further subset of records. However, if the user wants to begin a totally new search or one that would yield a larger or different set of records than subset X navigating back to the "Search Form" needs to be done within the software package so that the whole set of records can be initially reselected. Using the "Begin a New Search" button automatically selects all records and places the user back within the "Search Form." Beginning a new search, or returning back to the Silicasecchidisk home page, are available options on all results and search form pages.

As noted above, the results of a search are displayed on four different layouts. The initial layout presented includes a "List" of the records matching the search criteria, each with a thumbnail image. A larger view of the image, the "View Image Only", or all information contained within a record, the "View Full Record", can be accessed from any record within the "List" view and the user can flip through the records in consecutive order from either of these pages. Additionally, a layout summarizing results of all records retrieved from the search, the "View Summary Statistics" page, can be accessed via a button along the top of the "List" view.

Examples of searches

As noted above, simple and complex searches can be performed using the Giraffe Pipe database. Four sample searches are presented, each representing successively more restrictive search criteria. First, the simplest type of search would be one where the user fills in a single field, for example selecting "V-rib and Dome" from the "Chrysophyte Scale" field. This search will return all chrysophyte scales matching these characteristics, regardless of species. The statistics presented in the "View Summary Statistics" page will likewise summarize all of the records containing single whole scales. Second, selecting all records from depths in the lake greater than 20 m is done by entering ">20" in the "Depth in the Lake" field. This will yield records of all organisms in the database regardless of type from the deeper sections of the waterbody. Such a search may reveal the presence of specific types of organisms or species within a certain section of the core. However, in this case the summary statistics represent values for the entire assemblage of organisms and therefore may not be meaningful. Third, the sizes of specific organisms can be compared between different lake water depths. For example, the mean diameter of cyst type "GPCyst001" when the lake was less than 2 m deep vs. depths greater than 20 m can be compared by examination of the "View Summary Statistics" pages from two simple searches. In both cases "GPCyst001" is entered in the "Assigned Name" field. Values of "<2" and ">20" are entered into the "Depth in the Lake" field during successive search runs and results compared from the two searches. Fourth, in order to find the range in depth that large (>20 μm) smooth cysts have been recorded the user would select "Smooth, no secondary ornamentation" under the "Chrysophyte Cysts" field, "> 20" in the "Diameter" field, run the search and examine the "Minimum" and "Maximum" fields on the "View Summary Statistics" page. In conclusion, the numerous combinations of search criteria available within the database should significantly aid in the dissemination of information and results from the Giraffe Pipe project.

The future

Our initial plans are to significantly expand the number of records, add images taken with light microscopy and enter information from our collaborators. We further plan to link microfossil taxa to their closest living relative(s) and provide appropriate reference(s). Additionally, we would like to link records to images of the core. For the immediate future all of the records will represent samples from the Giraffe Pipe core. However, our long term plan is to merge information from other cores or sites representing the Eocene and eventually material from other geologic time

periods as well (e. g. Paleocene). We would like to develop a discussion forum, most likely in the form of a wiki, and link our database to other existing web-based resources such as Micro*scope (http://starcentral.mbl.edu/microscope/portal.php) and Stom@ocysts Wiki (http://www.stomatocysts.unibe.ch/wiki/Home). Lastly, we would eventually like to add software that can graph and display results of searches, such as a frequency distribution plot of diameters of cysts or scales.

Information Services (IS) at Connecticut College has pledged to maintain the Silicasecchidisk site, including the Giraffe Pipe project, indefinitely into the future. Currently, the site resides on a server that is maintained and systematically backed up by IS personnel. Please send comments and suggestions for improvement to P.A. Siver.

Acknowledgments

This research was supported by the National Science Foundation (NSF-DEB-0716606). We thank BHP Billiton Diamonds Inc. (Kelowna) and the Geological Survey of Canada (Calgary) for providing access to the Giraffe Pipe material. We also extend thanks to Frank Fulchiero (Connecticut College) for help with FileMaker Pro.

References

Shayler, H.A. & P.A. Siver (2006): Key to Freshwater Algae: A web-based tool to enhance understanding of microscopic biodiversity. – J. of Science Education and Technology **15**: 298–303.

Siver, P.A. & A.P. Wolfe (2005a): Eocene scaled chrysophytes with pronounced modern affinities. – Internat. J. of Plant Science **166**: 533–536.

Siver, P.A. & A.P. Wolfe (2005b): Scaled Chrysophytes in Middle Eocene lake sediments from Northwestern Canada, including descriptions of six new species. – Nova Hedwigia Beih. **128**: 295–308.

Siver, P.A. & A.P. Wolfe (2007): *Eunotia* spp. (Bacillariophyceae) from Middle Eocene lake sediments and comments on the origin of the diatom raphe. – Can. J. Bot. **85**: 83–90.

Wolfe, A.P., M.B. Edlund, A.R. Sweet & S. Creighton (2006): A first account of organelle preservation in Eocene nonmarine diatoms: observations and paleobiological implications. – Palaios **21**: 298–304.

Zachos, J.C., M. Pagani, L. Sloan, E. Thomas & K. Billups (2001): Trends, rhythms, and aberrations in global climate 65 Ma to present. – Science **292**: 686–693.

Zachos, J.C., G.R. Dickens & R.E. Zeebe (2008): An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. – Nature **451**: 279–283.