Commented data archive to study of Samsoondar, S. and Knappertsbusch, M. (2024, in prep). *"Combined archival and morphometric research for evolution in Globorotalia archeomenardii-praemenardii-menardii lineage from Oligocene-Miocene Cipero Formation type- and co-type localities, Trinidad, SE Caribbean"*

Samsoondar, S. and Knappertsbusch, M.

Naturhistorisches Museum Basel, Augustinergasse 2, 4001-Basel, Switzerland

Version: [15 February 2024]

Size: 1.21 GB

Introduction

This archive contains the primary raw data and intermediate results to morphometric investigations about the evolution of menardiform globorotalids (planktonic foraminifera) from the Cipero Formation, Cipero Coast and surrounding areas, SW Trinidad, Caribbean Sea. The goal of this study was to evaluate the feasibility of quantitative and morphometric investigations about planktonic foraminiferal evolution from historical samples, that once were collected from iconic Oligocene-Miocene foraminiferal zonal type and co-type localities during the 1950ies in southwestern Trinidad and now are curated along the micropaleontological collections of the Natural History Museum in Basel (NMB), Switzerland. This study extends previous similar investigations by the second author about shell evolution of Late Miocene through Quaternary *Globorotalia menardii* to early Miocene forms from the tropical West Atlantic by including ancestral forms like *G. archeomenardii* and *G. praemenardii* from Trinidad, where these forms once have been erected.

The present Trinidad data is thus complimentary to previous such sets about menardiform globorotalid morphometric studies around the world. The archive is organized in three subdirectories, [IMAGES], [DATA], and [PROGRAMS], that can be downloaded as 7-zip archives. Images were collected using the AMOR imaging and orientation automat AMOR (version 3.28, see Knappertsbusch et al., 2009; Knappertsbusch and Eisenecker, 2022) and morphometric parameters were extracted from shell outlines in keel view using existing software developed for studying the morphological evolution of *Globorotalia menardii* and related forms (see Knappertsbusch, 2007, 2015a,b, 2016a,b, 2022, 2023). Measurements include the spiral height (δX) and axial length (δY), the area (Ar) of the silhouette in keel view, radii (R_{up}, R_{lo},) of osculating circles at the upper and lower keel, keel angles (Φ_1, Φ_2) in the upper and lower keel region, spiral and umbilical convexities (A, B), and convexity ratios (A/B) from 644 specimens between 15.75 Ma to 8.9 Ma. Next to measurements from the *G. menardii, G. praemenardii and G. archeomenardii*, the present archive includes also measurements from *G. scitula and G. praescitula* supposed to be ancestral forms to the *G. menardii* plexus, but not explicitly analyzed in the publication.

Reduced final results and statistical tables in Samsoondar and Knappertsbusch (2024a, in prep.) are provided as electronic supplement directly with the published article and are not fully duplicated in this commented archive. A review about deep archival investigations on Trinidad material held at NMB, especially within the legacy of Hans G. Kugler, and details about the numerical age determination of samples in the present archive are given in an accompanying publication by Samsoondar and Knappertsbusch (2024b, in prep).

The data are structured into three main directories, that are described in the sections further below. A mirror of the archive is also stored on the internal media server of the Natural History Museum Basel, Switzerland. Samples, measured slides, residues, notes about the age determination and laboratory journals to this study are deposited in the micropaleontological reference collections at the same institution.

Data containers:

The three main directories to this archive are the following ones:

[IMAGES]

[DATA]

[PROGRAMS]

[IMAGES]**:**
The main subdirectory [IMAGES] contains the raw images (in Tiff format), intermediate formats (Black and While images, in Tiff format) and the binary images (black and white images in raw format), from which the outline coordinates were derived. Images are first arranged in subdirectories per sample, therein sorted by slides per sample (a,b,c,…). Within slides, images are arranged in the above described formats:

[Tiff files] contains images and files directly from AMOR.

[B_W_Tiff] are the binary images (Black and White) in Tiff format after processing with ImageJ.

[Raw files] contain the binary images (Black and White) in Raw format (of Adobe Photoshop), that are fed to the outline extraction program. The Raw format was generated from [B_W_Tiff] by conversion using ImageJ.

The file Rapid_processing_AMOR_images.pdf provides details about methods for rapid image processing are described. For the present preparation of images the *option d* was applied.

[DATA]:

The main subdirectory [DATA] contains results of data extraction in subdirectories named after each respective program application. In this way, subdirectories are arranged in sequence of work flow. Program codes and examples for input and output are provided in the third main subdirectory called [PROGRAMS].

Note, that programs were written for application on an old iMac computer under Mac OS 9.2 (!), where file names do not have file extensions. On order to maintain readability under Windows OS all input and output files provided here were tagged with the respective file extensions (.txt for text; .raw for raw image files were necessary). This means, that if provided source codes are recompiled to executable applications on Macintosh or on PC, file extensions need to be removed first for running on a Mac, or alternatively, the code to individual programs needs to be slightly adjusted to include file extensions if prepared for running under Windows OS.

DATA subdirectory [00_MagCorr2]:

Contains corrections of magnifications, that are obtained from imaging with AMOR. This step was necessary in order to account for minor differences between magnifications returned by the AMOR 2.8 software (magnifications listed in List_of_files) and the actual physical magnification at the microscope (magnifications listed in List_of_files_corr). Format: The first 5 characters are the input file name of an image (example 0101r). The four characters after the comma indicate the magnification. The magnifications given in List_of_files_corr are used as input for the Trace_AMOR2_batch program. Input files to MagCorr2 are List_of_files, output files are List_of_files_corr. Files are first sorted by sample, then by slide (a,b,c,…) within a sample.

DATA Subdirectory [01_Rename5]:

Program Rename5 converts the 19 character long string (including the .tif extension) of a tif image obtained with AMOR on a PC (Windows) into a 5 character long string (without extension) for the name under Mac OS 9.2. Example: JS2514125aK0101.tif is converted into 0101r (in raw format).

In "JS2514125aK0101.tif" characters 1-4 denote the sample designation, characters 5-9 indicate size fraction, character 10 indicates the slide, character 11 (K) stands for keel view, characters 12-15 indicate the specimen in the slide with characters 12-13 being the field number (01 through 60) and characters 14-15 indicating the number of specimens per field (always 01). Characters 16-19 are reserved for the file extension (".tif"). In the converted filename only the specimen is indicated (characters 1-4) and the character r at position 5 indicates, that this is an image in raw format of Adobe Photoshop.

Images in the xx01r format are then fed to the Trace_AMOR2_batch program for extraction of outline coordinates on the iMac.

Files are first sorted by sample, then by slide (a,b,c,…) within a sample.

DATA subdirectory [02_Trace_AMOR2_batch]:

Program Trace_AMOR2_batch calculates the outline coordinates from each specimen. Input are the images in format xx01r, and the List_of_files_corr with corrected magnifications. Output are one text file "…._T" per specimen with the cartesian X,Y coordinates of outline points in micrometers. The string "_T" stands for "traced file". In the next step traced files are then visually checked for completeness using program XYPlot2.

Files are first sorted by sample, then by slide (a,b,c,…) within a sample.

DATA subdirectory [03_XYPlot2]: By running program XYPlot2 on the traced (_T) files, each outline is visually checked on the computer monitor whether the outline was extracted correctly and whether it is a closed outline. This subdirectory contains the checked outline (_T) files. If an error occurred the image of the particle was checked and re-processed if necessary and then re-traced.

DATA subdirectory [04_Sort_Species_nCh_Coil]:

At this level of analysis all traced and checked outlines (_T files) per sample were re-organised and *manually* re-sorted into groups of species, separated by number of chambers in the final whorl, and by the coiling direction (sinistral or dextral). This a-priori resorting was unavoidable for further comparison of outlines between specimens and across samples and fascilitates exploration for morphological and evolutionary trends.

The categorization was done visually under the microscope. Species determination was done using the taxonomic literature indicated in the article, and species names (always 7 characters long; a blank was replaced by "_") were encoded according to the file 00a_Name_conventions given in the DATA container. The number of chambers in the final whorl occurs as nCh, the coiling direction is abbreviated by sin or by dex.

DATA subdirectory [05_Sprep53]:

Program Sprep53 prepares the traced (_T) outlines for further extraction of morphometric parameters from the coordinates. The re-arrangement of the file structure from subdirectory 04 (e.g., species/nCh/coiling) is hereby maintained: Input to Spre53 is the list with the traced (_T) files and the traced files themselves. For each outline an interpolated outline (designated by "_INT" is calculated with a constant number of points. In the present study outlines were inter- or extrapolated to 250 points, as was done in all previous studies of Knappertsbusch. Cartesian coordinates are given in micrometers, with the gravity centers of each outline moved into the origin of the coordinate system). The "_POL" files contain the outline points in polar coordinates (RHO, in micrometers, PHI in radians, but not used in the present study). The output file MEASUREMENTS lists basic informations per specimen (Sample age in Ma, the original number of points in the _T file, Spiral height (δX); axial length (δY), both in micrometers, and the keel view area in square millimeters, see Figure_parameters and Figure 3 in the article). The file Errors_during_Sprep alerts if an outline is not closed, or if the number of outline points in the input file (_T) is less than the desired number of interpolated coordinates (e.g., 250). If this was the case the user has the option to exclude or include that respective outline (in the present study none of the outlines were excluded).

DATA subdirectory [06_KeelWidth13.0]:

Program KeelWidth13.0 derives additional morphometric parameters from each outline, and lists them in the "*Res*ult *K*eel *W*idth" file RESKW_[sample]_[Species]_[nCh]_[coil] in the same file hierarchy as in DATA subdirectories [04_Sort_Species_nCh_Coil] and [05_Spre53]. For this step, the constant Q was set to 0.05, and option 1 for output of convexity A and B is applied. The constant Q describes the height of the keel regions as a percentage of the total axial length in keel view. Input files are MEASUREMENTS, and the _T, _INT_, and _POL files obtained in DADA subdirectory 05. Output files are _LoKR and _UpKR files describing the upper and lower keel regions at 95% and 5% of axial length, and file List_KR_circles with the center coordinates and radii of upper and lower osculating circles in those keel regions. In the files RESKW_[sample]_[Species]_[nCh]_[coil] all distance measurements are given in micrometers, areas in square millimeter, angular measurements (PHI1, PHI2, PHI3)

in degrees (respectively Φ_1, Φ_2, and Φ_3 in Figure_parameters, see also Figure 3 in the article), and age is given in Ma.

DATA subdirectory [07_DexFlip]:

In order to compare geometric measurements of sinsitral tests with dextral ones in keel view, the pairs of coiling dependent parameters were exchanged against each other if specimens showed dextral coiling. This "DexFlip" operation is nothing else than vertically mirroring the image for coiling dependent parameters, and is done with program DexFlip, see furter explanations in Knappertsbusch (2016a and 2022) for further explanations).

Input file is a file that lists the RESKW_[sample]_[Species]_[nCh]_[coil] files to be dexflipped and the RESKW_[sample]_[Species]_[nCh]_[coil] files themselves. Output is a file Composed_Dflip that lists the morphometric parameters, where the coiling dependent variables of dextrally coiled specimens are mirrored, so that from this moment onwards *ALL* specimens conform to sinistral coiling. In the present analysis option 1 is applied, because convexities A and B are included. Files are sorted in the same file structure as in [04_Sort_Species_nCh_Coil] through [06_KeelWidth13.0].

Data subdirectory [08_SplitwStan2]:

Splitweighting was a necessary correction of specimen frequencies in order to deal with unequal splits between size fractions within a sample and for calculation of statistics like means or medians. This correction is done by using program SplitwStan2 (see also section about split-weighting in the article). For example: Only 1 specimen was found in the entire split (1/1 split) of the 250-1000 microns size fraction, while exceedingly many specimens were found in the 125-250 microns and 63-125 microns size fractions. In such a case, for example only 1/4 was used from the 125-250 microns and 63-125 microns fractions.

For subsequent calculation of statistics (means, medians, etc) it is therefore necessary to compensate for using different splits between size fractions of a sample. In the present example this was compensated by listing the measurements from the 250-1000 μm fraction (1/1 split) 1x (NSPL=1), while measurements from the smaller size fractions (1/4 split each) were listed 4x (NSPL=4) for importing the data into a statistics program. In addition, in our Cipero Coast study we decided to account for unequal sample splits across samples by up-scaling (standardizing) each sample to the entire (1/1) split. For example: If in sample X 1/1 split was imaged and analysed and in sample Y only 1/4 split was used, then the measurements from sample Y was listed 4x prior to importing the data into a statistis program.

In program SplitwStan the split number NSPL (first column of input file) is considered to represent the combined effect of splitweighting and standardisation.

Files are sorted in the same file structure as in [04_Sort_Species_nCh_Coil] through [07_DexFlip].

Data subdirectory [09_By_species_spw_comb]:

This step is a second change of the file structure from DATA subdirectory 04 through subdirectory 07: This change removes the previous arrangement of ResKW files of a species into nCh and coiling and recombines all "dexflipped" and "splitweighted" specimens of a species within a sample. This second resorting was again done manually by merging together the respective files within a sample. Files were then renamed according to the scheme Sample_species_Spw, where the string "Spw means, that data have been corrected for unequal split sizes. After this procedure files are now arranged only by sample, so that they can be fed to program Statistica 12 for calculation of basic statistics for morphometric parameters per sample.

Data subdirectory [10_Statistica12]:

In this subdirectory the files obtained in the previous step 09 are prepared for and entered in commercial Statistica12 from StatSoft. Files are arranged by sample. Input files into Statistica12 are given in subdirectory [Tabbed_input_files]. Results with basic statistics per sample and species are given in subdirectory [Stats_per_sample] in the format of statistica workbook files (file extension .stw). Readable text versions from these statistica workbook files are provided in the subsequent DATA subdirectory [11_Statistics_Integrated], and are also delivered as plane tabulated text files Statistics_archaeo.txt (for *G. archeomenardii*), Statistics_praemen.txt (for *G. praemenardii*), and Statistics_menardi.txt (for *G. menardii*) in the APPENDIX STATISTICS of the article.

DATA subdirectory [11_Statistics_Integrated]:

Here, the basic statistics for *G. archeomenardii*, *G. praemenardii*, *G. menardii* (and also data for *G. praescitula* and *G. scitula* but not discussed in here) are converted from statistica workbook (.stw) format to plane tabulated text format. These were input files for generating diagrams in the next following DATA subdirectory [12_Plots].

DATA subdirectory [12_Plots]:

In this data container input files (.txt) to plots (Figures 4 through 10 of the article) of mean keel view area (Ar), spiral height (δX), axial length (δY), spiral convexity (A), convexity ratio (A/B), radius lower osculating circle (R_{lo}), radius upper osculating circle (R_{up}) against time for *G. archeomenardii*, *G. premenardii*, and *G. menardii* for the present Trinidad samples are given here. With the data also the basic plots are provided in in post-script (.ps) and .pdf format. Scatter plots were generated using "prehistoric" (though flexible and very efficient) GricketGraph software under MacOs 9.2. From these the final illustrations in the article were composed using Adobe Illustrator. The subdirectory [DxDy] contains the data for generation of contoured frequency diagrams of δX versus δY per species and sample. Bivariate δX,δY measurements are given in text files with the general file name species_Age.txt. The sample age in Ma is encoded in the filename by 5 characters XX_XX, where the underscore "_" stands for a decimal point. Example: the file menardi_08_90.txt contains the δX versus δY measurements in micrometers for *G. menardii* in the sample at 08.90 Ma. Bivariate frequency matrices (=gridded files) of δX versus δY are given in files that contain the string "grid". Gridding from the δX,δY was done using program Grid2.2, at grid-cells of 100micrometers in δX direction and 50 micrometers

in δY direction (see explanations in the article, and code to Grid 2.2 in PROGRAMS subdirectory [14_Grid2.2]). Frequencies are given in specimens per grid-cell. Contour diagrams were generated importing z-value matrices into software Surface III on an iMac under MacOs 9.2. Contour diagrams were named according to the scheme Age_cont_intn, where intn means contour intervals=n. Example: 12_00Ma_cont_int10.ps in DATA subdirectory [12_Plots/DxDy/menardi] is the contoured diagram at 12.00 Ma with contour intervals=10, postscript file. File menardi_cipero_cont_int20.pdf in the same subdirectory is the stacked contour interval for *G. menardii* from all samples from the present study of Trinidad samples.

DATA subdirectory [13_ODP_925B_all_menardii]:

The files in this folder are based from measurements from ODP Hole 925B, 0.00 Ma through 7.96 Ma. Data include δX versus δY measurements from the 4669 *G. menardii* specimens described in Knappertsbusch (2016a), plus those extra 1826 *G. menardii* specimens from the 9 supplemental samples from Hole 925B described in Knappertsbusch (2023). Together these are 6495 specimens of *G. menardii* from ODP Hole 925B, that serve as a comparison to the present measurements from the Trinidad materials. The gridded data and statistics to those 6495 specimens are given here in order to reproduce the contoured stack diagram of *G. menardii* (0-7.96 Ma) from Ceara Rise illustrated in Figure 12 of the present article.

DATA subdirectory [Conv Wade2011-Berggren1995]:

This container lists the conversions of numerical ages of planktonic foraminiferal datums from Wade et al. (2011) into the chronology of Berggren et al. (1995). The Berggren et al. (1995) biogeochronology was applied for the datation of Trinidad samples in the present study (see explanations in the article).

[PROGRAMS]:

The main subdirectory [PROGRAMS] contains the codes of fortran77 programs, that were used in the present study. The majority of these programs are documented in the MorphCol collection of Knappertsbusch (2015a), but are reproduced here again for simplicity reasons. The program DexFlip was developed in 2015 for the study of menardiform globorotalids at ODP Hole 925B and is documented in the data files with Knappertsbusch (2016a). The program Trace_AMOR2_batch was slightly modified in 2018 and is documented in the appendix to Knappertsbusch (2022). Program SplitwStan2 was written for the present investigation on Trinidad materials and is documented in the comment section of the code in PROGRAMS subdirectory [10_SplitwStan2]. During the present investigation the programs were applied in the following sequence (see also section "Summary of steps for morphometric analysis" in the article):

1.)   MagCorr2.out (corrects magnification readings from AMOR to magnification at microscope).
2.)   Rename5.out (renames raw files for input to Trace_AMOR2_batch.out).
3.)   Trace_AMOR2_batch.out (outline extraction from raw images).
4.)   XYPlot2.out (visual check for outlines).
5.)   Sort_Species_nCh_Coil (manually re-sorting files to species, chamber numbers, coiling).
6.)   Sprep53.out (interpolates and centers outlines).

7.) KeelWidth13.out (calculates morphometric parameters from outlines).
8.) XYPlot3.1out (plots outlines to screen visual check for osculating circles).
9.) DexFlip.out (mirrors coiling dependent parameters in dextral specimens).
10.) SplitwStan2.out (corrections for uneven splits from size fractions).
11.) *By_species_spw_comb* (manually resorting files to species only (remove nCh, Coil).
12.) Statistica12 (calculation of statistics for parameters per species and sample).
13.) Statistics_Integrated (manually integrate statistics per species, plot line graphs).
14.) Grid2.out (gridding spiral height (δX) vs axial lenth (δY) data for preparation of contour plots).

Acknowledgments

References

Berggren, W.A., Kent, D.V., Swisher, C.C. III, and Aubry, M.-P. (1995). A revised Cenozoic geochronology and chronostratigraphy. In: Berggren, W.A., Kent, D.V., Aubry, M.-P., Hardenbol, J. (eds.). Geochronology, Time Scales, and Global Stratigraphic Correlation. SEPM Special Publication #54, SEPM (Society for Sedimentary Geology), pp. 129-212.

Knappertsbusch, M. (2007). Morphological variability of *Globorotalia menardii* (planktonic foraminiferan) in two DSDP cores from the Caribbean Sea and the Eastern Equatorial Pacific. Carnets de Géologie / Notebooks on Geology, Brest, Article 2007/04 (CG2007_A04). URL: http://paleopolis.rediris.es/cg/CG2007_A04/index.html. For raw data see Knappertsbusch (2016b).

Knappertsbusch, M., Binggeli, D., Herzig, A., Schmutz, L., Stapfer, S., Schneider, C., Eisenecker, J., and Widmer, L. (2009). AMOR - A new system for automated imaging of microfossils for morphometric analyses. Palaeontologia Electronica, Vol. 12, Issue 2; 2T: 20 p. URL: http://palaeo-electronica.org/2009_2/165/index.html

Knappertsbusch, M., (2015a). MorphCol 2004-2013. A collection of Fortran 77 programs for geometric morphometry. Unpublished Technical Report, 1st Edition. Naturhistorisches Museum Basel, Augustinergasse 2, 4001-Basel, Switzerland. http://doi.pangaea.de/10.1594/PANGAEA.848773

Knappertsbusch (2015b). Raw-data to morphometric investigations about the Neogene planktonic foraminifer *Globorotalia menardii* and related forms from ODP Hole 154-925B (Ceara Rise, western tropical Atlantic). http://doi.pangaea.de/10.1594/PANGAEA.855898

Knappertsbusch, M. (2016a). Evolutionary prospection in the Neogene planktic foraminifer *Globorotalia menardii* and related forms from ODP Hole 925B (Ceara Rise, western tropical Atlantic): evidence for gradual evolution superimposed by long distance dispersal ? Swiss Journal of Palaeontology, vol. 135 (2), pp. 205-248. https://dx.doi.org/10.1007/s13358-016-0113-6

Knappertsbusch, M., (2016b). Raw data to Knappertsbusch, M. (2007). Morphological variability of *Globorotalia menardii* (planktonic foraminiferan) in two DSDP cores from the Caribbean Sea and the Eastern Equatorial Pacific. File archive for download (1.48 GB). https://doi.pangaea.de/10.1594/PANGAEA.863580

Knappertsbusch, M. (2022). Morphological evolution of menardiform globorotalids at Western Pacific Warm Pool ODP Hole 806C (Ontong-Java Plateau). Revue de Micropaléontologie 74 (2022) 100608. https://doi.org/10.1016/j.revmic.2022.100608. The approved manuscript and data files to this publication is also accessible from https://zeonodo.org/doi.org/10.5281/zenodo.803885

Knappertsbusch, M. and Eisenecker, J. (2022). Towards a Fleet of Robots for Orientation, Imaging, and Morphometric Analyses of Planktonic Foraminifera. Frontiers in Marine Science. Technology and Code. Published 08 March 2022. https://doi.org/10.3389/fmars.2022.798002

Knappertsbusch, M. (2023). Morphological cladogenesis and terminal dwarfing in extinct Late Miocene through Pliocene menardiform globorotalids: New complementary data to *"Evolutionary prospection in the Neogene planktic foraminifer Globorotalia menardii and related forms from ODP Hole 925B (Céara Rise, western tropical Atlantic):*

evidence for gradual evolution superimposed by long distance dispersal ? Swiss J. Palaeontology, 135:205-248". Data archive and report published on https://doi.org/10.5281/zenodo.7812050

Samsoondar and Knappertsbusch (2024a, in prep.). Combined archival and morphometric research to document shell evolution in *Globorotalia archeomenardii-praemenardii-menardii* lineage at iconic Oligocene-Miocene Cipero Coast type section, Trinidad and Tobago, SE Caribbean. Manuscript in prep.

Samsoondar and Knappertsbusch (2024b, in prep). Archival research for the reapraisal of the classical Oligocene to Miocene Cipero Coast type section, Cipero Formation, Trinidad, SE Caribbean, for integrated biogeochronology. Manuscript in prep.

Wade, B.S., Pearson, P.N., Berggren, W.A., Pälike, H. (2011). Review and revision of Cenozoic tropical planktonic foraminiferal biostratigraphy and calibration to the geomagnetic polarity and astronomical time scale. Earth-Science Reviews, vol. 104, pp. 111-142. DOI: 10.1016/j.earscirev.2010.09.003