
Adding missing pieces to the gharial puzzle: new morphological and biostratigraphical framework to unravel the evolution of long-snouted crocodylians.

A Data Management Plan created using DMPonline.be

Creators: Sophie Boerman, Johan Vellekoop, n.n. n.n.

Affiliation: KU Leuven (KUL)

Funder: Fonds voor Wetenschappelijk Onderzoek - Research Foundation Flanders (FWO)

Template: FWO DMP (Flemish Standard DMP)

Grant number / URL: 11PM124N

ID: 206703

Start date: 01-11-2023

End date: 31-10-2027

Project abstract:

Among the crocodylians still alive are two species with long, narrow snouts: the Indian gharial and the Malayan or "false" gharial. These enigmatic species have spurred extensive debate over the last decades, as morphologists and molecular biologists could not agree whether the gharials are closely related or only distant cousins. With the implementation of new methods, some morphological studies now agree that the extant species are sister taxa. However, the rich fossil record of longsnouted crocodylians highlights that the two lineages split from each other far earlier than the molecular estimates suggest. As a result, both the divergence time and phylogenetic relationships of fossil taxa are still unknown, and the evolutionary history of the gharials is left unclear.

Belgian museum collections comprise an abundance of well-preserved fossil long-snouted crocodylians, providing an excellent archive to contribute towards solving the gharial problem. Nevertheless, few modern morphological studies have been performed on these specimens, and their stratigraphic age is often poorly constrained. Therefore, I propose a project to study the phylogenetic relationships of key fossil taxa through new morphological approaches and determine their age with dinoflagellate biostratigraphy on their sedimentary matrix. Current morphological datasets will be revised in this project, using the latest phylogenetic methods to construct a new and robust morphological framework.

Last modified: 30-04-2024

Adding missing pieces to the gharial puzzle: new morphological and biostratigraphical framework to unravel the evolution of long-snouted crocodylians.

FWO DMP (Flemish Standard DMP)

1. Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

				Only for digital data	Only for digital data	Only for digital data	Only for physical data
Dataset Name	Description	New or reused	Digital or Physical	Digital Data Type	Digital Data format	Digital data volume (MB/GB/TB)	Physical volume
		<i>Please choose from the following options:</i> <ul style="list-style-type: none"> • Generate new data • Reuse existing data 	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> • Digital • Physical 	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> • Observational • Experimental • Compiled/aggregated data • Simulation data • Software • Other • NA 	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> • .por, .xml, .tab, .csv, .pdf, .txt, .rtf, .dwg, .gml, ... • NA 	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> • <100MB • <1GB • <100GB • <1TB • <5TB • <10TB • <50TB • >50TB • NA 	
Macrofossil specimens	Specimens of crocodylians/eusuchians (skulls, bones, skeletons)	Reuse existing data	Physical				one new specimen, < 1 m3; about 10 reused specimens, = 1 m3
Photographs	Photographs of macrofossil specimens	Generate new data	Digital	Observational	.png .raw	<100MB each < 1 TB total	
Drawings	Anatomical drawings of macrofossil specimens	Generate new data	Digital	Observational	.png	<100MB each and <100MB in total	
Micro-CT data	Micro-CT scans of macrofossil specimens	Generate new data	Digital	Observational	.tif and .ply	<100GB each < 5 TB total	
Rock samples	Collected from the field or from rock matrix that is attached to macrofossil specimens	Generate new data AND reuse existing data	Physical				+/- 10 samples < 1 dm^3 in total

Microscope slides	Boxes with glass slides containing microfossils, e.g. thin sections, smear slides, or chemically extracted microfossils	Generate new data	Physical				<40 slides <1 dm ³ in total
Phylogenetic matrices	Matrices containing morphological or morphological + molecular character scorings for crocodylian / eusuchian taxa	Reuse existing data	Digital	Observational & Compiled/aggregated data	.xlsx .csv .txt	<100MB each and <100MB in total	

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

Macrofossil specimens: specimens from museum/university collections will be used. Most macrofossil specimens used in this project are curated by the Royal Belgian Institute of Natural Sciences (RBINS).

Rock samples: some rock samples will be collected from macrofossil specimens from museum/university collections in cases where rock matrix is preserved with these specimens.

Phylogenetic matrices: existing matrices will be used and updated, e.g. the matrix created by Brochu (1999; ; doi.org/10.1080/02724634.1999.10011201), the one by Rio & Mannion (2021; doi.org/10.7717/peerj.12094), or more recent adaptations of those matrices.

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? Describe these issues in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Will you process personal data? If so, briefly describe the kind of personal data you will use in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, ...)? If so, please comment per dataset or data type where appropriate.

- Yes

CT scan data and derived 3D models of macrofossil specimens might be interesting for commercial exploitation, e.g. the selling of 3D printed skeletons for aesthetic purposes. For this reason, these data will only be made available with restricted access (see below).

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements/ research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- No

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- No

2. Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g., in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, Codebook.tsv etc. where this information is recorded).

Accompanying information for each type of digital data will be included in the manuscript in which the data are published. This includes:

- The procedures used to analyse the phylogenetic data matrices
- Information on the macrofossil specimens figured in photographs, drawings, and CT scans (such as locality, age, and taxonomic determination of the specimens).

For the physical data, paper labels with the date of discovery, locality, age, and taxonomic determination are kept with each specimen.

Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify (where appropriate per dataset or data type) which metadata standard will be used. If not, please specify (where appropriate per dataset or data type) which metadata will be created to make the data easier to find and reuse.

- No

See above

3. Data storage & back-up during the research project

Where will the data be stored?

Macrofossils and microscope slides will remain in permanent storage of RBINS and KU Leuven respectively. Rock samples collected during the PhD will be stored in the permanent storage of KU Leuven, whereas those collected from museum-curated specimens will be stored at the institution from which they were obtained.

Digital data are saved on OneDrive cloud storage of KU Leuven (2 TB).

Raw CT data of RBINS macrofossils will be stored both on hard disks belonging to the RBINS and on the Belspo Long Term Preservation (LTP) platform.

How will the data be backed up?

During research, digital data (drawings, photographs, and phylogenetic matrices) are saved both on OneDrive cloud storage of KU Leuven and on a personal disk space folder (I:).

Raw CT data of RBINS macrofossils will be stored both on hard disks belonging to the RBINS and on the Belspo Long Term Preservation (LTP) platform.

It is not possible to provide back up for the physical data.

**Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely.
If no or insufficient storage or backup capacities are available, then explain how this will be taken care of.**

- Yes

There is sufficient storage for all physical data in the RBINS and KU Leuven for the macrofossil specimens and microscope slides, respectively.

There is enough space on the OneDrive storage provided by KU Leuven to store most of the digital data, with the exception of CT scans which will likely amount to more than 2 TB in total. These will be stored on hard disks belonging to the RBINS and on the Belspo Long Term Preservation (LTP) platform.

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

Physical data are kept in rooms with locked doors.

For all the online data, KU Leuven works with a 2-step verification in which only authorized personnel can access the data.

RBINS computers where the CT scan data are stored are password protected.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

No costs are expected for data storage and backup.

4. Data preservation after the end of the research project

Which data will be retained for at least five years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

All data will be preserved for at least 10 years after the end of the research, according to KU Leuven RDM policy.

Where will these data be archived (stored and curated for the long-term)?

After research, the digital data will be saved on KU Leuven SharePoint permanent (L drive, 10 TB) storage servers and RBINS storage servers, both system-encrypted with built-in versioning and back-up features. Raw CT data of RBINS macrofossils will be stored both on hard disks belonging to the RBINS and on the Belspo Long Term Preservation (LTP) platform.

Macrofossils and microscope slides will remain in permanent storage of RBINS and KU Leuven respectively.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

No costs are expected for data storage and backup.

5. Data sharing and reuse

Will the data (or part of the data) be made available for reuse after/during the project? In the comment section please explain per dataset or data type which data will be made available.

- Yes, in an Open Access repository
- Yes, in a restricted access repository (after approval, institutional access only, ...)

Digital data will be published online with Open Access, according to the FWO Open Access regulations.

CT scan data and physical data will be made available to researchers upon request (non-commercial purposes only).

If access is restricted, please specify who will be able to access the data and under what conditions.

CT scan data and physical data will be made available to researchers upon request (non-commercial purposes only).

For CT scan data and physical data belonging to the RBINS, researchers have to request access to the curator of the paleontological collections. The requester has to sign a data user agreement with the RBINS before the data will be made available to them.

For physical data stored at KU Leuven, access must be requested to Johan Vellekoop.

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)? Please explain in the comment section per dataset or data type where appropriate.

- No

Where will the data be made available? If already known, please provide a repository per dataset or data type.

Physical data will be made available upon request in the respective storage facilities at KU Leuven and RBINS.

Digital data will be uploaded in KU Leuven RDR.

CT data will be made available by the RBINS upon request by the curator of the paleontological collections.

When will the data be made available?

Upon acceptance of the publication of research results.

Which data usage licenses are you going to provide? If none, please explain why.

Any data derived from RBINS-curated specimens (photographs, drawings, CT scans) are protected under a CC BY-NC-ND 4.0 license.

Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, you have the option to provide it in the comment section.

- Yes

What are the expected costs for data sharing? How will these costs be covered?

No costs are expected for data sharing.

6. Responsibilities

Who will manage data documentation and metadata during the research project?

Sophie Boerman

Who will manage data storage and backup during the research project?

Sophie Boerman

Who will manage data preservation and sharing?

Sophie Boerman, Annelise Folie (curator of RBINS paleontological collections), Johan Vellekoop

Who will update and implement this DMP?

Sophie Boerman