

Microtomographic investigation of a large corpus of cichlids

This manuscript ([permalink](#)) was automatically generated from habi/EAWAG-manuscript@7894b29 on December 2, 2022.

Authors

- **David Haberthür**

 [0000-0003-3388-9187](#) ·  [habi](#) ·  [habi](#)

Institute of Anatomy, University of Bern, Switzerland

- **Mikki Law**

None

- **Kassandra Ford**

 [0000-0002-4413-2633](#) ·  [kassthefish](#) ·  [kassthefish](#)

Department of Biological Sciences, George Washington University; EAWAG Aquatic Research Institute, Switzerland;
Institute of Ecology & Evolution, University of Bern, Switzerland

- **Marcel Häsler**

None

- **Ole Seehausen**

 [0000-0001-6598-1434](#) ·  [OleSeehausen](#)

Aquatic Ecology and Evolution, Institute of Ecology and Evolution, University of Bern, CH-3012 Bern, Switzerland;
Department of Fish Ecology and Evolution, EAWAG, Swiss Federal Institute for Aquatic Science and Technology, CH-6047 Kastanienbaum, Switzerland

- **Ruslan Hlushchuk**

 [0000-0002-6722-8996](#) ·  [RuslanHlushchuk](#)

Institute of Anatomy, University of Bern, Switzerland

Abstract

A large corpus of Cichlids from Lake Victoria in Africa spanning a size range of 6 to 20 cm was nondestructively imaged using micro-computed tomography. We describe our method to efficiently obtain three-dimensional tomographic data sets of the oral and pharyngeal jaws and the whole skull of these fishes for accurately describing their morphology. The tomographic data we acquired (8.8 TB projection images) was reconstructed into 1.4 TB of three-dimensional images used for extracting the relevant features of interest. Herein we present our method and an outlook on two projects analyzing the acquired data; a morphological description of the oral and pharyngeal jaws of the fishes as well as a principal component analysis of landmark features on the fish skulls.

Introduction

History

Cichlid fish in African lakes are a powerful model systems in speciation and adaptive evolutionary radiation research [1,2]. The functional decoupling of their oral and pharyngeal jaws is hypothesized to be a factor in making cichlids unusually versatile in their feeding and making it possible for making them able to adapt to a wide range of environmental factors. The hypothesis is that the fusion of the lower pharyngeal jaws makes them powerful food processing tools, and this in turn releases the oral jaws from functional constraint. The oral jaws no longer need to process prey and can therefore specialize on prey capture.

The group of Ole Seehausen has studied the evolutionary diversification of cichlid fish radiations in the African Great Lakes since a long time [2,3]. We aim to better understand the functional anatomy of the skulls and jaws in these fishes in order to test the functional decoupling and other hypotheses about what may facilitate exceptionally high rates of morphological evolution.

The corpus of cichlid fish available is extremely valuable, hence a nondestructive imaging method is paramount for studying these samples. Since micro-computed tomography can be regarded as nondestructive method for biological samples, it is a very well suited method for investigating the oral and pharyngeal jaws as well as the skull features of the fished presented in this study.

Micro-computed tomography

X-ray microtomography is a valuable tool to gain insights into the inner structure of very diverse samples, namely for specimens related to research done in the biomedical sciences. Namely in the ‘fish sciences’, X-ray microtomography has been employed as a method of choice to nondestructively assess the morphology of various samples [4]¹

Depending on the structures of interest, biomedical samples are often tomographically scanned after the tissue/sample has been stained with a contrast agent, most often employing contrast agents containing heavy metals. Since the structures of interest for the two studies we touch upon in this manuscript (cichlid teeth and skulls) display large enough contrast to the surrounding tissue we did not stain our samples prior to the tomographic imaging presented here.

Materials and Methods

Sample procurement and preparation

The fishes were kept in 75% Ethanol for long-term storage in the EAWAG fish library. They were delivered to the Institute of Anatomy for micro-CT imaging sorted into several batches by approximately equal length.

Micro-computed tomographic imaging

All samples were scanned on two of the three available high-resolution micro-CT machines of the Institute of Anatomy of the University of Bern in Switzerland, a SkyScan 1272 and a SkyScan 2214 (both Bruker microCT, Kontich, Belgium).

The fishes were sorted into ‘bins’ based on their physical size. We used a custom-made sample-holder to scan each of the fish in our machines. It was 3D-printed on a Form 2 desktop stereolithography printer (Formlabs, Somerville, Massachusetts, USA) and is freely available online [5] as part of a library of sample holders for tomographic scanning of biomedical samples [6]. The sample holder was custom-made for this project and is easily parametrized to the different width, height and length classes of the fishes we scanned.

In total, we acquired 362 tomographic scans of 129 different fishes. All the scanning parameters are collected in a table in the [Supplementary Materials](#), a generalized rundown is given below.

Since the fishes greatly varied in their length, the voxel sizes of each of the acquired datasets also varies greatly. We acquired datasets with (isometric) voxel sizes ranging from 3–50 µm.

Depending on the size of the specimen we set the x-ray source voltage to 50–80 kV and—depending on the voltage—to a current between 107 and 200 µA. Also depending on the size of the fishes, the x-ray spectrum was filtered either by an Aluminum filter of varying thickness (either 0.25, 0.5 or 1 mm) before digitization to projection images or recorded in an unfiltered way. In total we recorded 9.5 TB of projections images (*.?if files) for this project.

All the recorded projection images were subsequently reconstructed into a 3D stack of axial images spanning the regions of interest of each fish. All the specimens were scanned with their mouths facing downward in the sample holder and rotating along their long axis. We thus manually aligned each of the reconstructed datasets so that the lateral axis of the fish was horizontal in relation to the x and y direction of the reconstructed slices. We reconstructed the projection images with NRecon (Version 1.7.4.6, Bruker microCT, Kontich Belgium) with varying ring artifact and beam hardening correction values, depending on each fish (again, all relevant values are listed in the [Supplementary Materials](#)). In total, this resulted in 1.4 TB of reconstruction images (*.rec*.png files).

While performing the work, a subset of the data was always present on the production system, for working with it (see [Preparation for analysis](#) below). A small bash script [7] was used to generate redundant (archival) copies of the raw projection images and copy all the files to a shared network drive on the [Research Storage](#) infrastructure of the University of Bern, enabling easy collaboration on the data by all authorized persons at the same time.

Data analysis

We wrote a set of *Jupyter* [8] notebooks with *Python* code to work with the images and wrangle the acquired data. The notebooks were written at the start of the project, to be able to process new scans as soon as they were reconstructed. Re-runs of the notebook added newly scanned and reconstructed fishes to the analysis, facilitating an efficient quality check of the scans and batched processing of the data.

All analysis notebooks for this work are available online [9].

Preparation for analysis

The [main Jupyter notebook](#) for this manuscript dealt with reading all log and image files and preparing images for quality checking and further analysis. Briefly summarized the below process was implemented.

At first, *all* log files of *all* the data present in the processed folder were read into a into a *pandas* [10] dataframe. This already enabled us to extract the specimen name and scan, since we performed several scans for each specimen, i.e. a low resolution scan with large field of view for the whole head and one or several scans in high resolution focusing on the oral and pharyngeal jaws. From the log files we extracted the relevant values for double-checking the necessary parameters of each scan. All relevant values for each scan were also saved into the dataframe and saved out to the aforementioned table in the [Supplementary Materials](#) at the end of each run of the notebook.

After several ‘sanity checks’ of the data, we used *Dask* [11] to efficiently access the very large total amount of axial reconstructions for this project (in the end amounting to a total of nearly a million single images). On average, each of the tomographic datasets contains around 3000 slices, so the total amount of data is much too large to keep in memory. The use of the Dask library facilitated efficient access to the huge amount of data on disk.

At first, we extracted the central view of each of the three axial directions of the datasets (i.e. ‘anteroposterior’, ‘lateral’ and ‘dorsoventral’ view) and either saved those to disk or loaded them from disk if they were already generated in prior runs of the notebook. The notebook then also generated the maximum intensity projection (MIP) for each of the anatomical planes and either saved them to disk or loaded them from prior runs.

At the end of the notebook we performed a final sanity check on the MIP images. In this check we examined the mapping of the gray values of the raw projection images to gray values in the reconstructions, i.e. checked that no overexposed pixels are present in the MIP images. This is an an efficient way for double-checking the gray value mapping, since the MIP images have already been generated in prior steps of the notebook and contain the highest gray values present in all the reconstructed images of each scan.

Image processing

Extraction of oral and pharyngeal jaws, visualization of tomographic data

To extract the oral jaw (OJ) and pharyngeal jaw (PJ) of the fishes, we used [3DSlicer](#) (Version 4.11.20210226) [12] extended with the *SlicerMorph* tools [13] which aim to help biologists to work with 3D specimen data. The reconstructed PNG stacks were loaded into *ImageStacks*, depending on their size we reduced the image resolution (i.e. downscaled the images) for this first step. The three-dimensional volume was rendered via [VTK GPU Ray Casting](#). A custom-made volume property (created by Kassandra Ford) was used as an input to view the scans. Using toggles in the volume rendering, we defined regions of interest (ROIs) for both the OJs and PJs in each specimen.. These

ROIs were then extracted in their native resolution from the original dataset for further processing. Using the gray value thresholding function in 3DSlicers [Segment Editor](#) the teeth in both the oral and pharyngeal jaws were extracted. We used the *Scissor* and *Island* tools of the Segment Editor to isolate single regions.

Processed regions of interest were exported as NRRD [14] files. The three-dimensional visualizations of all regions of interest for each specimen were compiled into overview images (see Figure 1 for an example from the compilation document). In total we compiled overview of XXX specimens with full head morphology, oral jaw and lower pharyngeal jaw profiles.

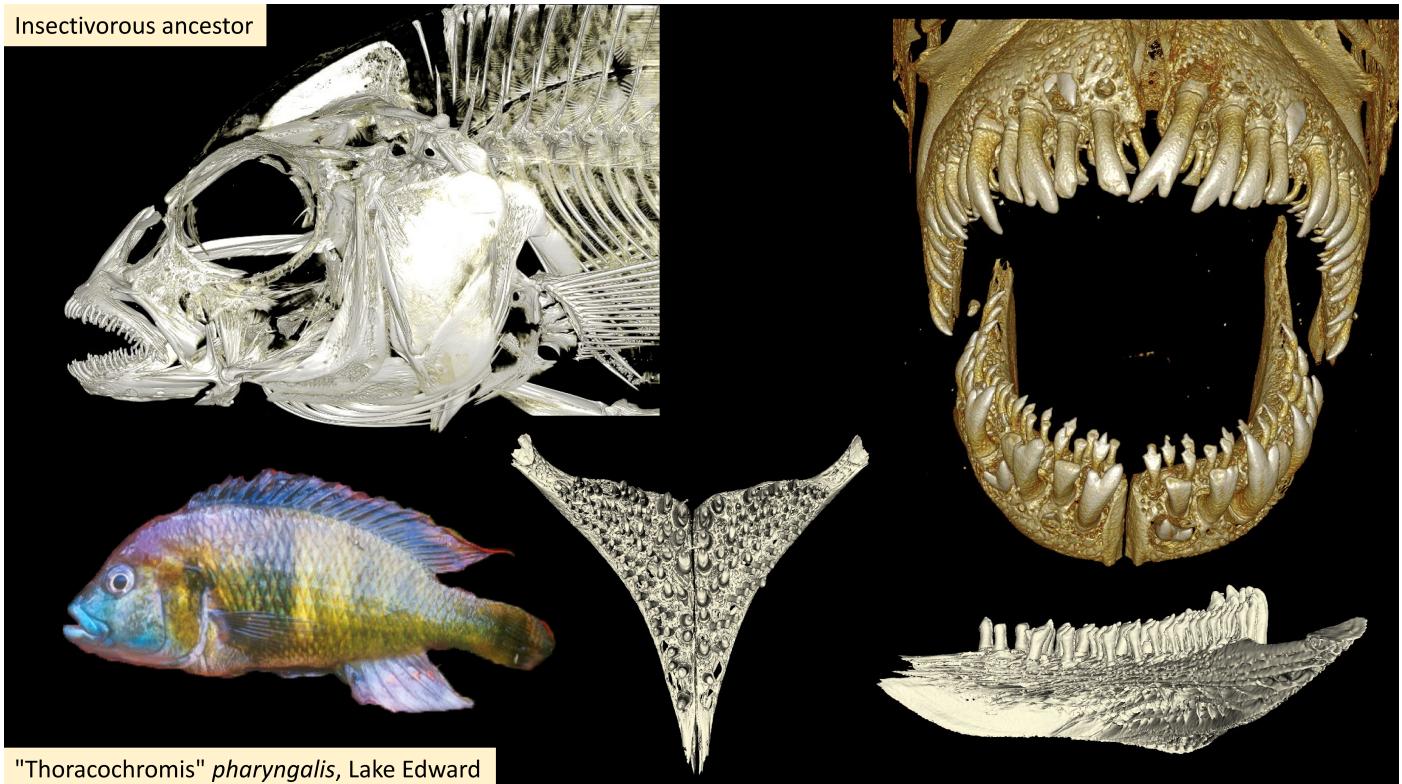


Figure 1: Overview of data from 'KAT-13, Lake Edward, "Thoracochromis" pharyngalis (pharyngeal mollusc crusher - shrimp)'

Principal component analysis of skull landmarks

Current studies are using 3D geometric morphometrics to compare the morphological shape of these scanned cichlids using statistical analysis. A homologous landmark scheme was produced [15,16], and landmarks were placed on each specimen using 3DSlicer. To examine differences in shape across the species sampled, we performed a Generalized Procrustes Superimposition on the landmark data to remove the effects of location, size, and rotation from the analysis using the *geomorph* package in *R* (Version 4.2.1) with *RStudio* (Version 2022.07.2+576) [17,18,19,20,21]. This process brings all specimens to a common origin, scales the landmarks to a unit centroid size, and rotates specimens to reduce distances between landmark sets. A principal component analysis was then performed on the superimposed landmark data to visualize the major axes of shape change across sampled species. We then used phylogenetic information to identify instances of repeated evolution of trophic adaptations in cichlids.

Automatic extraction of otoliths

The so-called otoliths in the fish heads structures made up of mostly calcium carbonate and are located in the head of fishes. Due to their composition they show up nicely on the X-ray images we acquired. We devised a very robust image processing method to automatically detect the location of

the otoliths in the heads of the fishes and extract them from the original data. The whole method is implemented in its own [Jupyter notebook](#) (part of the aforementioned analysis repository [9]) and is briefly explained below.

Since we took great care to scan the fishes parallel to their *anteroposterior* direction and reconstructed the tomographic datasets parallel to the *lateral* and *dorsoventral* direction of the fishes we could use this ‘preparation’ for automatically extracting the otoliths in scans of the full heads of the fishes. By extracting both the peaks and the peak widths of the gray values along both the horizontal and vertical direction of the MIP (generated above) we robustly detected the position of the otoliths in the datasets. The robust detection is supported by suppressing a small, configurable part of each region, i.e. the front and back, top and bottom or the flanks.

Figure 2 shows the visualization of the process. The colored horizontal and vertical bars in each of the directional MIPs denote the found peak location of the two appropriate values. The white bars show the mean of the two detected positions, which was used for extracting the otoliths from the original datasets. Making use of the `dask` library facilitated efficient access to all the data on disk and writing out small, cropped copies of the datasets around the otolith positions.

The extracted otoliths are then prepared for further analysis and display. The aforementioned Jupyter notebook also offers a simple three-dimensional visualization through an integrated visualization library [22].

Results

Extraction of oral and pharyngeal jaws, visualization of tomographic data

Principal component analysis of skull landmarks

Automatic extraction of otoliths

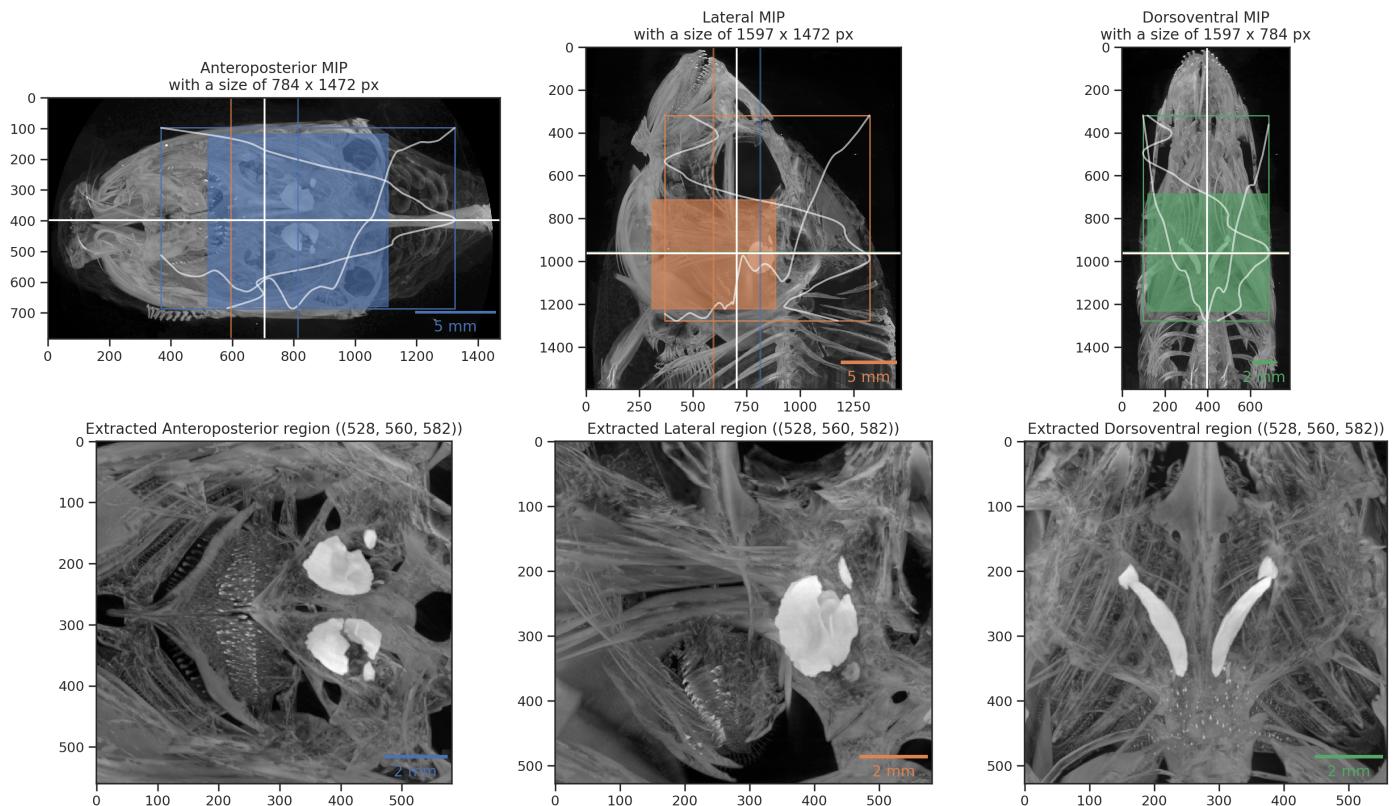


Figure 2: Visualized result of automatic otolith extraction.

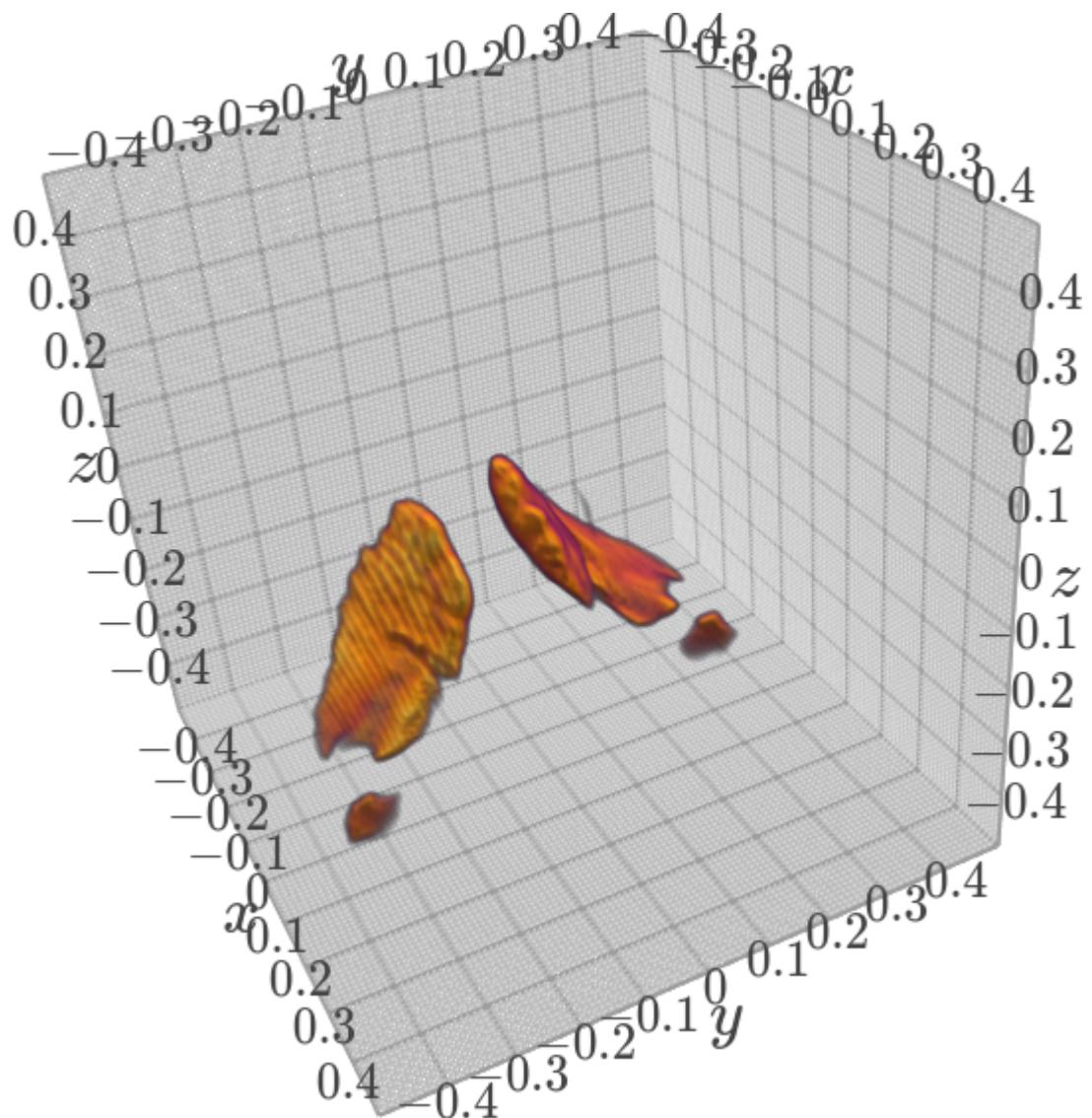


Figure 3: Result of automatic otolith extraction. Three-dimensional view of extracted otolith.

Discussion

The discussion of the results and the outlook to what we'll do in the future is going into this file here.

Acknowledgments

We thank the `manubot` project [23] for helping us write this manuscript collaboratively.

Supplementary Materials

Parameters of tomographic scans of all the fishes

The CSV file [ScanningDetails.csv](#) gives a tabular overview of all the (relevant) parameters of all the scans we performed. This file was generated with the [data processing notebook](#) and contains the data which is read from *all* the log files of *all* the scans we performed. A copy of each log file is available in a [folder in the data processing repository](#).

References

1. **Adaptive evolution and explosive speciation: the cichlid fish model**
Thomas D Kocher
Nature Reviews Genetics (2004-04-01) <https://doi.org/dkgf93>
DOI: [10.1038/nrg1316](https://doi.org/10.1038/nrg1316) · PMID: [15131652](#)
2. **African cichlid fish: a model system in adaptive radiation research**
Ole Seehausen
Proceedings of the Royal Society B: Biological Sciences (2006-05-09) <https://doi.org/frrhbq>
DOI: [10.1098/rspb.2006.3539](https://doi.org/10.1098/rspb.2006.3539) · PMID: [16846905](#) · PMCID: [PMC1635482](#)
3. **Process and pattern in cichlid radiations – inferences for understanding unusually high rates of evolutionary diversification**
Ole Seehausen
New Phytologist (2015-05-13) <https://doi.org/f7gczg>
DOI: [10.1111/nph.13450](https://doi.org/10.1111/nph.13450) · PMID: [25983053](#)
4. **CT Scans - #ScanAllFish**
Adam P Summers
(2015-09-29) <https://osf.io/ecmz4/>
5. **OpenSCAD models**
TomoGraphics
(2022-02-28)
<https://github.com/TomoGraphics/Hol3Drs/blob/dedca665a77fa782c103eed63b866b2cce30de9/STL/EAWAG.Fish.stl>
6. **TomoGraphics/Hol3Drs: A release**
David Haberthür
Zenodo (2019-03-08) <https://doi.org/gg9fxh>
DOI: [10.5281/zenodo.2587555](https://doi.org/10.5281/zenodo.2587555)
7. **Jupyter notebooks for image processing and data analysis of EAWAG Cichlids project**
David Haberthür
(2022-07-05)
<https://github.com/habi/EAWAG/blob/0a51a339bbe3134c635b78890e208e813e7b87ae/rsync-fishes.sh>
8. **Jupyter Notebooks – a publishing format for reproducible computational workflows**
Thomas Kluyver, Benjamin Ragan-Kelley, Fernando Pérez, Brian Granger, Matthias Bussonnier, Jonathan Frederic, Kyle Kelley, Jessica Hamrick, Jason Grout, Sylvain Corlay, ... Jupyter development team
IOS Press (2016) <https://eprints.soton.ac.uk/403913/>
DOI: [10.3233/978-1-61499-649-1-87](https://doi.org/10.3233/978-1-61499-649-1-87)
9. **habi/EAWAG: First release, for minting a Zenodo DOI**
David Haberthür
Zenodo (2022-07-05) <https://doi.org/gggdtp>
DOI: [10.5281/zenodo.6798632](https://doi.org/10.5281/zenodo.6798632)
10. **pandas-dev/pandas: Pandas**
The Pandas Development Team
Zenodo (2022-09-19) <https://doi.org/gqw4db>

DOI: [10.5281/zenodo.7093122](https://doi.org/10.5281/zenodo.7093122)

11. **Dask: Library for dynamic task scheduling**
Dask Development Team
(2016) <https://dask.org>
12. **3D Slicer as an image computing platform for the Quantitative Imaging Network**
Andriy Fedorov, Reinhard Beichel, Jayashree Kalpathy-Cramer, Julien Finet, Jean-Christophe Fillion-Robin, Sonia Pujol, Christian Bauer, Dominique Jennings, Fiona Fennessy, Milan Sonka, ... Ron Kikinis
Magnetic Resonance Imaging (2012-11) <https://doi.org/gfghgd>
DOI: [10.1016/j.mri.2012.05.001](https://doi.org/10.1016/j.mri.2012.05.001) · PMID: [22770690](https://pubmed.ncbi.nlm.nih.gov/22770690/) · PMCID: [PMC3466397](https://pubmed.ncbi.nlm.nih.gov/PMC3466397/)
13. **SlicerMorph: An open and extensible platform to retrieve, visualize and analyse 3D morphology**
Sara Rolfe, Steve Pieper, Arthur Porto, Kelly Diamond, Julie Winchester, Shan Shan, Henry Kirveslahti, Doug Boyer, Adam Summers, AMurat Maga
Methods in Ecology and Evolution (2021-07-22) <https://doi.org/gqtgv8>
DOI: [10.1111/2041-210x.13669](https://doi.org/10.1111/2041-210x.13669)
14. <https://w.wiki/5mBK>
15. **Mosaic Evolution of Craniofacial Morphologies in Ghost Electric Fishes (Gymnotiformes: Apterontidae)**
Kassandra L Ford, Maxwell J Bernt, Adam P Summers, James S Albert
Ichthyology & Herpetology (2022-05-27) <https://doi.org/gq3gtd>
DOI: [10.1643/i2021016](https://doi.org/10.1643/i2021016)
16. **Convergence is Only Skin Deep: Craniofacial Evolution in Electric Fishes from South America and Africa (Apterontidae and Mormyridae)**
Kassandra L Ford, Rose Peterson, Maxwell Bernt, James S Albert
Integrative Organismal Biology (2022) <https://doi.org/gq3gtc>
DOI: [10.1093/iob/obac022](https://doi.org/10.1093/iob/obac022) · PMID: [35976714](https://pubmed.ncbi.nlm.nih.gov/35976714/) · PMCID: [PMC9375771](https://pubmed.ncbi.nlm.nih.gov/PMC9375771/)
17. **RRPP: An <scp>r</scp> package for fitting linear models to high-dimensional data using residual randomization**
Michael L Collyer, Dean C Adams
Methods in Ecology and Evolution (2018-05-29) <https://doi.org/gdwp9q>
DOI: [10.1111/2041-210x.13029](https://doi.org/10.1111/2041-210x.13029)
18. **Geomorph: Software for geometric morphometric analyses. R package version 4.0.4**
DC Adams, ML Collyer, A Kaliontzopoulou, EK Baken
(2022) <https://cran.r-project.org/package=geomorph>
19. **RRPP: Linear model evaluation with randomized residuals in a permutation procedure, r package version 0.6.2.**
ML Collyer, DC Adams
(2021) <https://cran.r-project.org/package=RRPP>
20. **R: A language and environment for statistical computing**
R Core Team
R Foundation for Statistical Computing (2018) <https://www.R-project.org/>
21. **RStudio: Integrated development environment for r**
RStudio Team
RStudio, PBC. (2022) <https://www.rstudio.com/>

22. **K3D Jupyter**

K3D-tools

(2022-11-29) <https://github.com/K3D-tools/K3D-jupyter>

23. **Open collaborative writing with Manubot**

Daniel S Himmelstein, Vincent Rubinetti, David R Slochower, Dongbo Hu, Venkat S Malladi, Casey S Greene, Anthony Gitter

PLOS Computational Biology (2019-06-24) <https://doi.org/c7np>

DOI: [10.1371/journal.pcbi.1007128](https://doi.org/10.1371/journal.pcbi.1007128) · PMID: [31233491](#) · PMCID: [PMC6611653](#)

1. For which David made a tomographic scan of an adult zebrafish ages ago. ↵