Posted on 22 Sep 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.166384820.08397260/v1 — This a preprint. Version of Record available at https://doi.org/10.1007/s00239-023-10106-3

PAReTT: a Python package for the Automated Retrieval and management of divergence time data from the Time Tree resource

Louis-Stéphane Le Clercq[1,2], J. Paul Grobler[1], Antoinette Kotze[1,2], and Desiré Dalton[3]

[1]University of the Free State
[2]South African National Biodiversity Institute
[3]Teesside University

July 27, 2023

Abstract

Evolutionary processes such as speciation happens gradually over time making such processes time-dependant. Many studies conducted over the past two decades have aimed at providing accurate, fossil-calibrated, estimates of the divergence times of both extant and extinct species in most lineages of the tree of life, including fish, amphibians, reptiles, birds, and mammals. Data from more than 4 000 of these studies are now publicly available from a central time tree resource and provide opportunities of retrieving divergence times, evolutionary timelines, and time trees in various formats to enhance scientific investigations of evolution. There is, however, still limited functionality when studying large lists of species that would require the batch retrieval of data. To overcome this, a PYTHON package called Python Automated Retrieval of Time Tree data, abbreviated as PAReTT, was created to facilitate the interaction with the time tree resource when working with species lists. This package was recently used in a meta-analysis of candidate genes to study migration genetics and was able to successfully retrieve data for forty or more species to illustrate the relationship between divergence times and genetic data. The PAReTT package is freely available for download from GitHub to implement in PYTHON or as a pre-compiled Windows executable, with extensive documentation on the package available on the PAReTT GitHub wiki pages on dependencies, installation, and implementation of the various functions.

PAReTT: a Python package for the Automated Retrieval and management of divergence time data from the Time Tree resource

Le Clercq, L.S[1,2,*], Kotzé, A[1,2]., Grobler, J.P[2]., Dalton, D.L.[3]

[1]*South African National Biodiversity Institute, Pretoria, 0001, South Africa.*

[2]*Department of Genetics, University of the Free State, Bloemfontein 9300, South Africa.*

[3]*School of Health and Life Sciences, Teesside University, Middlesbrough, TS1 3BA, United Kingdom.*

*Corresponding Author: Louis-Stéphane Le Clercq (leclercq.l.s@gmail.com).

(ORCiD: 0000-0002-8713-8920; WOS Researcher ID: AAO-9057-2021)

Short title: Python Automated Retrieval of Time Tree data.

Abstract

Evolutionary processes such as speciation happens gradually over time making such processes time-dependant. Many studies conducted over the past two decades have aimed at providing accurate, fossil-calibrated, estimates of the divergence times of both extant and extinct species in most lineages of the tree of life, including fish, amphibians, reptiles, birds, and mammals. Data from more than 4 000 of these studies

Posted on 22 Sep 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.166384820.08397260/v1 — This a preprint. Version of Record available at https://doi.org/10.1007/s00239-023-10106-3

are now publicly available from a central time tree resource and provide opportunities of retrieving divergence times, evolutionary timelines, and time trees in various formats to enhance scientific investigations of evolution. There is, however, still limited functionality when studying large lists of species that would require the batch retrieval of data. To overcome this, a PYTHON package called Python Automated Retrieval of Time Tree data, abbreviated as PAReTT, was created to facilitate the interaction with the time tree resource when working with species lists. This package was recently used in a meta-analysis of candidate genes to study migration genetics and was able to successfully retrieve data for forty or more species to illustrate the relationship between divergence times and genetic data. The PAReTT package is freely available for download from GitHub to implement in PYTHON or as a pre-compiled Windows executable, with extensive documentation on the package available on the PAReTT GitHub wiki pages on dependencies, installation, and implementation of the various functions.

Keywords: PAReTT, PYTHON, Time tree, Divergence time, Timelines, Diversification rate.

INTRODUCTION

Evolutionary processes are linked to time, be it diversification within a lineage which may lead to the emergence of a new species, or via subtle molecular changes over several generations steadily driving phenotypic variation (Wagner, 2018; Francisco Henao Diaz *et al.* , 2019). For example, some primary divisions between entire taxonomic orders of birds happened approximately 75 million years ago while more recent divisions between sub-species occurred as recently as 1 million years ago (Prum *et al.* , 2015). Many evolutionary processes are studied in the context of ecological and geographic processes that shape the landscape within which selection, adaptation, and extinction would have taken place and the paleogeography (Scotese, 2016; Müller *et al.* , 2018), which incorporates continental drift and known major periods of glaciation, can only be factored into the evolutionary history of a species if the time periods for speciation are known (**Figure 1**). Furthermore, accurate estimates of diversification rates within species are dependent on comparing the temporal range within which lineages, species, and subspecies are formed (Jetz *et al.* , 2012). It is therefore crucial that studies on evolutionary processes are contextualized within the relevant time frames they happened.

Over the past few decades, a plethora of molecular studies have been published using variable methods from fossil calibrated Bayesian inference (Rannala and Yang, 2003; Kumar and Hedges, 2016) to comparable relative time approaches (Yang and Yoder, 2003; Tamura *et al.* , 2012) to establish the timeline for the emergence and diversification of most species including many mammals (Nyakatura and Bininda-Emonds, 2012; Springer, Murphy and Roca, 2018), reptiles (Tucker *et al.* , 2017), and birds (Barker *et al.* , 2015; Prum *et al.* , 2015)–both living and extinct. These studies have greatly advanced our understanding of evolutionary processes within the context of ecological changes and the time constraints that they occur in (Scholl and Wiens, 2016), and has helped clarify many of the questions we have with regards to the taxonomy and phylogeny of species; which have frequently been at odds with each other (Sangster, 2014; Springer, Murphy and Roca, 2018).

The result is a compendium of 4 075 (or more) studies that has culminated in a central "Time Tree" resource (Hedges, Dudley and Kumar, 2006), that collects and compiles divergence time estimates and time trees from published and peer-reviewed studies. From this resource, estimates of divergence times and related timelines and time trees are available online (Kumar *et al.* , 2017), including an incorporated version in MEGA (Mello, 2018) and a mobile phone app (Kumar and Hedges, 2011). Collectively this provides access to accurate diverge time estimates for use in calibrating phylogenetic trees according to time, comparing clades in phylogenetic trees to know clades of shared common ancestry, as well as comparing genetic distance between species to their temporal or evolutionary distance. There is, however, still limited functionality in retrieving divergence times from the resource when dealing with species lists rather than individual pairs of species. As evolutionary studies frequently focus on multiple species, and even multiple lineages, at a time this presents a significant roadblock towards the streamlining of the integration of divergence time data into larger studies.

Previous attempts at automation to facilitate batch retrieval provided limited utilities and were poorly

Posted on 22 Sep 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.166384820.08397260/v1 — This a preprint. Version of Record available at https://doi.org/10.1007/s00239-023-10106-3

maintained, resulting in the removal of the package from the CRAN repository. The Time tree resource has continued to develop and expand and the need for such capabilities is eminent in the ever-expanding field of evolutionary biology. To this effect, we have endeavoured to create an easily accessible and freely available resource to retrieve relevant data on evolutionary histories from the Time tree site for the seamless integration of divergence time data in molecular studies. PAReTT, short for Python-Automated Retrieval of Time Tree data, is a menu driven and user-friendly PYTHON package to automate interaction with the Time Tree resource for retrieving batch data with lists of species, freely available on GitHub or as a stand-alone Windows executable.

METHODS

2.1 Implementation

PAReTT (version 1.0.1) was scripted in the Spyder 5 IDE in pure PYTHON and is compatible with versions 3.6 and upward. A full list of dependencies is provided on the GitHub wiki, along with details for download and installation. This includes the use of several well-established PYTHON based libraries such as NumPy (version 1.20.1) and pandas (version 1.2.4) for ease of input and high order moulding of data structures with relative ease (McKinney, 2010; Harris *et al.* , 2020), as well as Bio (version 1.3.9) for handling trees in the newick format. PAReTT further uses the headless browser functionality implemented in Splinter (version 0.17.0) with the Selenium (version 4.1.5) extensions for the Firefox browser to submit user specified web data to the Time Tree website and retrieve the relevant results. Results are printed in real-time to the shell, while list results are first stored to a 'dataframe' object which is written to user specified output file. Some functionality had also been provided to validate data for any errors that may have occurred and preview basic tree files. The script was benchmarked to test for time and memory consumption of individual functions using memory-profiler (version 0.16.0).

2.2 Input and Output files

PAReTT uses two primary forms of input that are specified when the user is prompted to provide the name of the input file. The first is a basic text file, indicated as ".txt", while the second is a standard comma separated value file, indicated as ".csv". While it is preferred to specify the full name of the file (e.g., "Species.txt"), PAReTT was scripted with checkpoints to ensure the proper format of input files even when only a name is given and should read the input file as long as they are in the current working directory. If the files are not stored in the current working directory the full file path is preferred.

Output files differ based on the specific data being retrieved. For example, timelines are retrieved as pictures in the JPEG format while time trees for both taxa as well as specific lists of species are retrieved in the Newick format, which is readable by most phylogenetic software programs including FigTree (Rambaut, 2017) and MEGA (Kumar *et al.* , 2018); these trees can also be viewed directly in PAReTT. Any species in the provided list for which the divergence time could not be resolved in the tree or were substituted by a similar species are stored as a table for review. Divergence time data for pairs are printed on screen and divergence times, based on the mean adjusted time, for species lists are stored in 'dataframe' objects and provide the option to store the output as a vectorized three column matrix in a comma separated value file. Output files are stored in the active working directory.

2.3 Usage

The main menu presents the initial options to verify the availability of data for a species, determine divergence times, a timeline, a time tree, validate the data, or print the citation for Time Tree, as summarized in **Figure 2** . The data availability option brings up a submenu to specify if availability should be verified for a single species or species list. When checking a single species, the availability is printed to the screen as the species name followed by either 'Available' or 'Not Available'. This function was designed to ensure data availability prior to further data retrieval steps. For a list of species, the species name and availability are printed on screen and stored to export results in a CSV file.

The timeline and time tree options will bring up a similar submenu that provides the option to retrieve

Posted on 22 Sep 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.166384820.08397260/v1 — This a preprint. Version of Record available at https://doi.org/10.1007/s00239-023-10106-3

information for an individual species, individual taxon, or a list of species. As before, individual species will result in a single output. The timeline retrieval function will retrieve an image in the JPEG format that illustrates the evolutionary history for the species highlighting the major time points where a kingdom, order, class, genus, and species first emerged. As an example, a timeline was retrieved for the Lazuli bunting, *Passerina amoena* [Say, 1823]. Input given as a list will retrieve individual images for each species in the list. The time tree option has similar functionality except that input can either be a larger taxonomic group, such as family or genus containing several congenic species, or a list of species as this output provides a time calibrated tree displaying both the interrelatedness of species as well as the timescale along which the branches diverged. As an example, two time trees were generated for the genus *Catharus* and *Malurus* respectively. The individual diversification rate (r) was calculated for each genus in R (R Core Team, 2020) using the *geiger* (version 2.0) package (Harmon *et al.* , 2008) with the Magallon-Sanderson equation (Magallón and Sanderson, 2001).

The Divergence time submenu provides the option to check for the divergence time between two species e.g., the Neotropical Swainson's thrush, *Catharus ustulatus* [Nuttall, 1840], and the Australasian Superb fairywren, *Malurus cyaneus* [Ellis, 1782]. The result will be printed to the screen as the name of 'Taxon a' and 'Taxon b' followed by the mean divergence time in millions of years ago (MYA) e.g., 35 MYA. The list option takes a text file with a list of species names and iterates through the list to determine the divergence time between each species in the list. The results are stored in a 'dataframe' object which can be exported as a CSV file in the form of a three column vectorized matrix with the names for 'Taxa 1' in the first column, the names of 'Taxa 2' in the second column, and the divergence times between them in the third column. Additional functionality is provided in the main menu to validate data. This option can be used to check output files for missing values, as well as retrieve and replace such values, in case a server error occurred.

RESULTS

Examples of timelines and time tree data retrieved from the Time tree resource are illustrated in **Figure 3** and **Figure 4** . The timeline was retrieved for the Neotropical bird the Lazuli bunting. The left panel indicates the major geologic timescales in Eons and Eras while the right indicates the main divergence times when specific taxa emerged. This includes the current Phanerozoic Eon as well as the subdivisions from the Paleozoic Era, which lasted until approximately 250 MYA, and the Mesozoic which lasted until 65 MYA. This period was marked by the emergence of the first birds (class: Aves) approximately 110 MYA. The specific species, a bunting in the family Cardinalidae, emerged from a common ancestor shared with other cardinals in the Cenozoic, while species of the genus *Passerina* first emerging 4 MYA and subsequently diverged to form species and subspecies.

Time trees were generated for two genera of birds that each contain several congenic species (11-12 species) but have different evolutionary histories in terms of the divergence times within their genus. The first represents the genus *Catharus* for which 12 species are included in the time tree. These species diverged over a period spanning approximately 4.73 million years and have an absolute diversification rate of 0.38. By comparison, the second tree for the genus *Malurus* includes a similar number of species, however, these species diverged over a period of approximately 9 million years and have a diversification rate ($r = 0.19$) of nearly half the rate observed in *Catharus* .

PAReTT was furthermore used to retrieve data availability and divergence times for a list of forty bird species, as well as a time tree for seventy-six species, in a recent review and meta-analysis on clock genes as candidate genes in migration studies, published in *Biological Reviews* . This study found a significant relationship between divergence times and the observed genetic distance in two candidate genes, illustrating high heritability of genotypes within lineages and a lack of selection. This illustrates the significance in studying divergence time data in relation to genetic data in molecular studies. The results from the benchmarking tests, performed in tandem, showed an average time for submitting data and retrieving the result ranged from fifteen to thirty seconds and memory consumption remained low, ranging from <100 to 400 Megabytes (**Supplementary table 1**).

Posted on 22 Sep 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.166384820.08397260/v1 — This a preprint. Version of Record available at https://doi.org/10.1007/s00239-023-10106-3

DISCUSSION

Time is an important variable in the study of evolutionary processes as it relates both to large scale geographic remodelling that occurred over millions of years and the timelines for the emergence and diversification of taxonomic and phylogenetic lineages. Time is also a key factor in determining the diversification rate within lineages which may be vastly different even between families in the same class and order. The Time tree resource provides relevant data that can easily be incorporated into molecular studies tracking diversification rates as well as studies comparing the heritability of specific genotypes as ancestrally inherited or currently undergoing active selection. The resource does, however, still lack options to facilitate batch retrieval for lists of species. Here, we illustrate the use of a newly scripted PYTHON package called PAReTT that provides an added layer of functionality to the Time Tree resource by enabling the batch retrieval of data for lists of species in a format that can enhance molecular and ecological studies.

DATA AVAILABILITY

The custom Python script for PAReTT version 1.0.1 is available for download for installation from source code on GitHub (https://github.com/LSLeClercq/PAReTT), or as a Windows executable and includes example files used for testing.

ACKNOWLEDGEMENTS

This work is based on the research supported wholly/in part by the National Research Foundation of South Africa (Grant Numbers: 112062).

AUTHOR CONTRIBUTIONS

LSLC scripted the PAReTT program, created the images and wrote the draft manuscripts. DLD, AK, and JPG provided research support and edited the final draft.

REFERENCES

Barker, K. *et al.* (2015) 'New insights into New World biogeography: An integrated view from the phylogeny of blackbirds, cardinals, sparrows, tanagers, warblers, and allies', *Auk* , 132(2), pp. 333–348. doi:10.1642/AUK-14-110.1.

Francisco Henao Diaz, L. *et al.* (2019) 'Macroevolutionary diversification rates show time dependency', *Proceedings of the National Academy of Sciences of the United States of America* , 116(15), pp. 7403–7408. doi:10.1073/pnas.1818058116.

Harmon, L.J. *et al.* (2008) 'GEIGER: investigating evolutionary radiations', *Bioinformatics* , 24(1), pp. 129–131. doi:10.1093/BIOINFORMATICS/BTM538.

Harris, C.R. *et al.* (2020) 'Array programming with NumPy', *Nature* . Nature Publishing Group, pp. 357–362. doi:10.1038/s41586-020-2649-2.

Hedges, S.B., Dudley, J. and Kumar, S. (2006) 'TimeTree: A public knowledge-base of divergence times among organisms', *Bioinformatics* , 22(23), pp. 2971–2972. doi:10.1093/bioinformatics/btl505.

Jetz, W. *et al.* (2012) 'The global diversity of birds in space and time', *Nature* , 491(7424), pp. 444–448. doi:10.1038/nature11631.

Kumar, S. *et al.* (2017) 'TimeTree: A Resource for Timelines, Timetrees, and Divergence Times', *Molecular Biology and Evolution* , 34(7), pp. 1812–1819. doi:10.1093/MOLBEV/MSX116.

Kumar, S. *et al.* (2018) 'MEGA X: Molecular evolutionary genetics analysis across computing platforms', *Molecular Biology and Evolution* , 35(6), pp. 1547–1549. doi:10.1093/molbev/msy096.

Kumar, S. and Hedges, S.B. (2011) 'Timetree2: Species divergence times on the iPhone', *Bioinformatics* , 27(14), pp. 2023–2024. doi:10.1093/bioinformatics/btr315.

Posted on 22 Sep 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.166384820.08397260/v1 — This a preprint. Version of Record available at https://doi.org/10.1007/s00239-023-10106-3

Kumar, S. and Hedges, S.B. (2016) 'Advances in time estimation methods for molecular data', *Molecular Biology and Evolution* , 33(4), pp. 863–869. doi:10.1093/molbev/msw026.

Magallón, S. and Sanderson, M.J. (2001) 'Absolute diversification rates in angiosperm clades', *Evolution* , 55(9), pp. 1762–1780. doi:10.1111/j.0014-3820.2001.tb00826.x.

McKinney, W. (2010) 'Data Structures for Statistical Computing in Python', in *Proceedings of the 9th Python in Science Conference* , pp. 56–61. doi:10.25080/majora-92bf1922-00a.

Mello, B. (2018) 'Estimating timetrees with MEGA and the timetree resource', *Molecular Biology and Evolution* , 35(9), pp. 2334–2342. doi:10.1093/molbev/msy133.

Müller, R.D. *et al.* (2018) 'GPlates: Building a Virtual Earth Through Deep Time', *Geochemistry, Geophysics, Geosystems* , 19(7), pp. 2243–2261. doi:10.1029/2018GC007584.

Nyakatura, K. and Bininda-Emonds, O.R.P. (2012) 'Updating the evolutionary history of Carnivora (Mammalia): A new species-level supertree complete with divergence time estimates', *BMC Biology* , 10. doi:10.1186/1741-7007-10-12.

Prum, R.O. *et al.* (2015) 'A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing', *Nature* , 526(7574), pp. 569–573. doi:10.1038/nature15697.

R Core Team (2020) 'R: A Language and Environment for Statisitical Computing'. Vienna: R Foundation for Statistical Computing. Available at: https://www.r-project.org/.

Rambaut, A. (2017) 'FigTree-version 1.4.3, a graphical viewer of phylogenetic trees'. Computer program distributed by the author. Available at: http://tree.bio.ed.ac.uk/software/figtree.

Rannala, B. and Yang, Z. (2003) 'Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci', *Genetics* , 164(4), pp. 1645–1656. doi:10.1093/genetics/164.4.1645.

Sangster, G. (2014) 'The application of species criteria in avian taxonomy and its implications for the debate over species concepts',*Biological Reviews* , 89(1), pp. 199–214. doi:10.1111/BRV.12051.

Scholl, J.P. and Wiens, J.J. (2016) 'Diversification rates and species richness across the Tree of Life', *Proceedings of the Royal Society B: Biological Sciences* , 283(1838). doi:10.1098/RSPB.2016.1334.

Scotese, C.R. (2016) 'PALEOMAP PaleoAtlas for GPlates and the PaleoData Plotter Program'. PALEOMAP Project. Available at: http://www.earthbyte.org/paleomap--%0Apaleoatlas--for--gplates/.

Springer, M.S., Murphy, W.J. and Roca, A.L. (2018) 'Appropriate fossil calibrations and tree constraints uphold the Mesozoic divergence of solenodons from other extant mammals', *Molecular Phylogenetics and Evolution* , 121, pp. 158–165. doi:10.1016/j.ympev.2018.01.007.

Tamura, K. *et al.* (2012) 'Estimating divergence times in large molecular phylogenies', *Proceedings of the National Academy of Sciences of the United States of America* , 109(47), pp. 19333–19338. doi:10.1073/pnas.1213199109.

Tucker, D.B. *et al.* (2017) 'Genomic timetree and historical biogeography of Caribbean island ameiva lizards (Pholidoscelis: Teiidae)', *Ecology and Evolution* , 7(17), pp. 7080–7090. doi:10.1002/ece3.3157.

Wagner, C.E. (2018) 'Improbable Big Birds Darwin's finches prove a mechanism for the rapid formation of new species', *Science* . American Association for the Advancement of Science, pp. 157–159. doi:10.1126/science.aar4796.

Yang, Z. and Yoder, A.D. (2003) 'Comparison of Likelihood and Bayesian Methods for Estimating Divergence Times Using Multiple Gene Loci and Calibration Points, with Application to a Radiation of Cute-Looking Mouse Lemur Species', *Systematic Biology* , 52(5), pp. 705–716. doi:10.1080/10635150390235557.

Figures

Posted on 22 Sep 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.166384820.08397260/v1 — This a preprint. Version of Record available at https://doi.org/10.1007/s00239-023-10106-3

Figure 1: Paleogeographic reconstructions of Earth for the past 70 million years. a) Positions of the continents approximately 70 million years ago during the Upper Cretaceous before West Africa had merged with the main continent and when India was still an island. b) Continents approximately 50 million years ago during the Palaeocene after the African continent formed but before the Americas were connected and shortly before India merged with Asia or the polar caps formed. c) Geography of Earth during the Eocene, 30 million years ago, by which time most continents had formed but central America did not connect the North and South yet and much of Europe was still under water. d) Modern day geography of Earth in the current Holocene. (image created in BioRender.com)

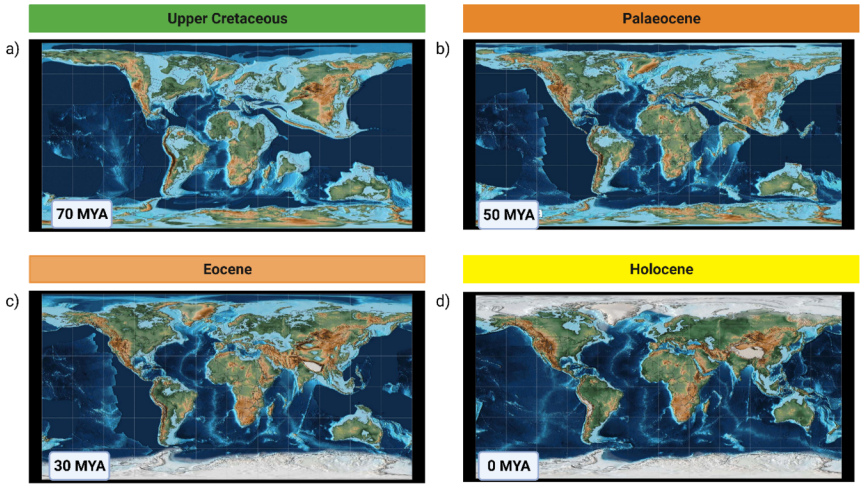


Figure 2: Graphical summary of menu options and sub-menu's for PAReTT. The main menu options are: *, to verify the availability of data; a or b, to get divergence times (pair or batch); c, get a timeline; d, get a time tree; e, print the Time Tree citation; f, validate data; or q, to exit . The data availability, timeline, and time tree options bring up a sub-menu, indicated by yellow arrows, to retrieve information for an individual species/taxon or a list. Output generated for lists are exported as a table (csv), images (jpeg), or trees (newick). The validation option brings up the choice of finding or replacing missing values for divergence times data (options a or b)as well as to view the tree topology (option c), for output files. (image created in BioRender.com)

Posted on 22 Sep 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.166384820.08397260/v1 — This a preprint. Version of Record available at https://doi.org/10.1007/s00239-023-10106-3

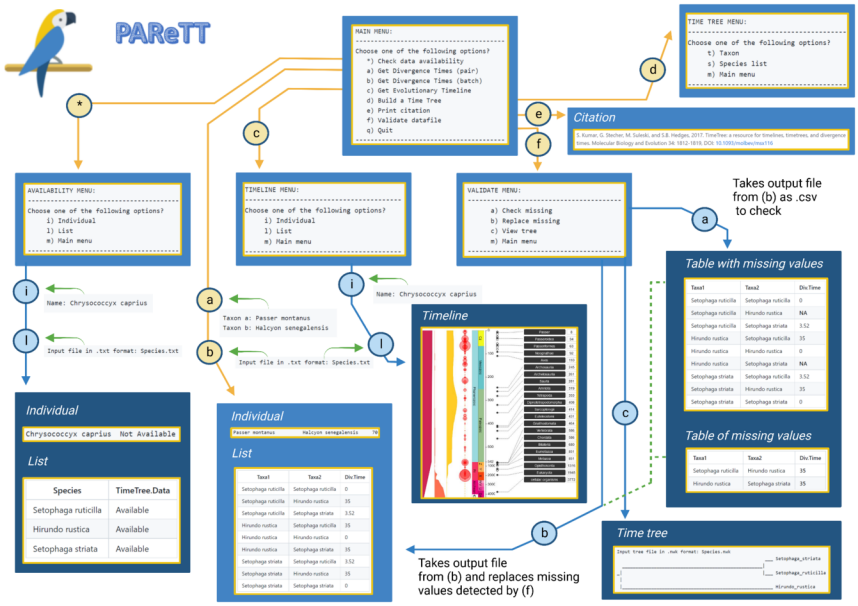


Figure 3: Example of an evolutionary timeline retrieved using PAReTT for the Lazuli bunting. The left panel indicates the major geologic timescale of the past 2000 million years including the prevailing Phanerozoic Eon as well as the subdivisions by era Paleozoic which lasted until approximately 250 MYA when the Mesozoic started which lasted until 65 MYA. As is illustrated, this period was marked by the emergence of the first birds in the class Aves approximately 110 MYA. The species currently recognized in the family Cardinalidae emerged in the Cenozoic era with most species of the genus *Passerina* first emerging 4 MYA.

Posted on 22 Sep 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.166384820.08397260/v1 — This a preprint. Version of Record available at https://doi.org/10.1007/s00239-023-10106-3

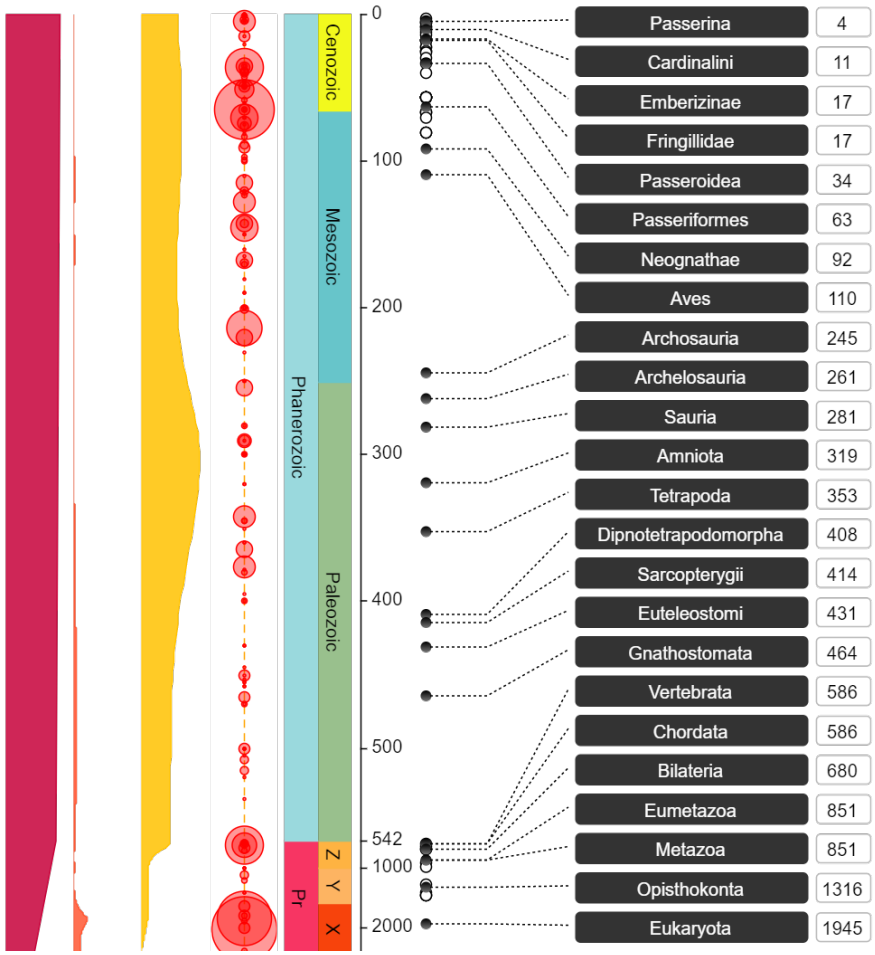
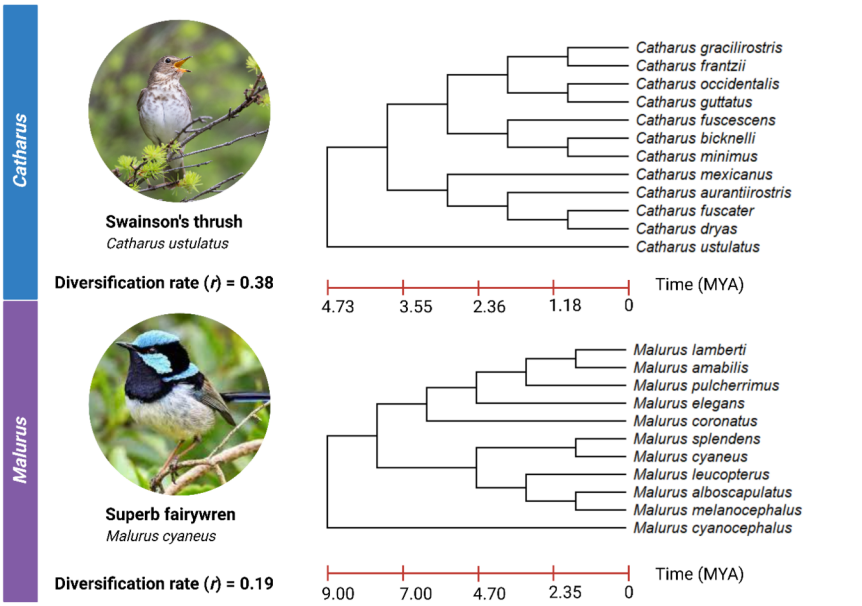
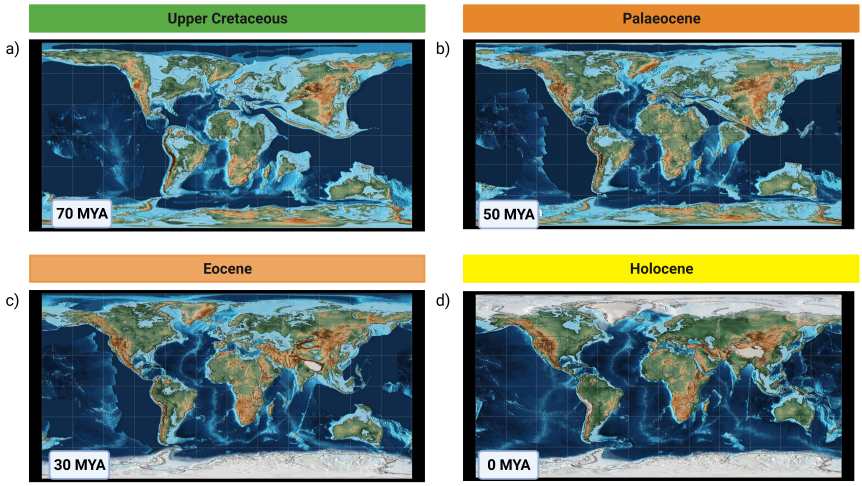
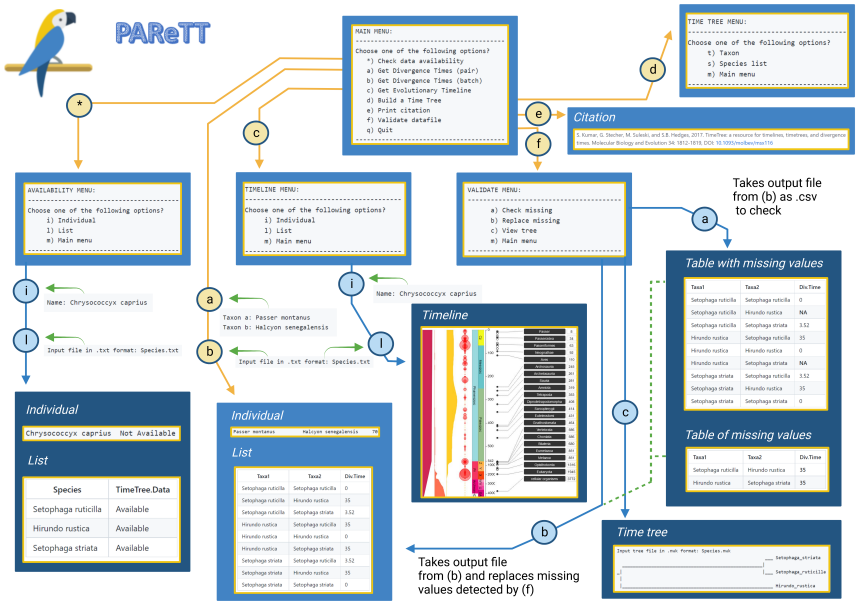


Figure 4: Comparison of time trees retrieved for two genera using PAReTT, one for several species in the *Catharus* genus of Neotropical thrushes and another of species in the *Malurus* genus of Australasian fairywrens. The *Catharus* genus contains 12 species, including Swainson's thrush, that diverged over a period of 4.74 million years with a diversification rate of 0.38. The*Malurus* genus contains 11 species, including the Superb fairywren, that diverged over a period of 9 million years with a diversification rate of 0.19. (image created in BioRender.com)

Posted on 22 Sep 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.166384820.08397260/v1 — This a preprint. Version of Record available at https://doi.org/10.1007/s00239-023-10106-3

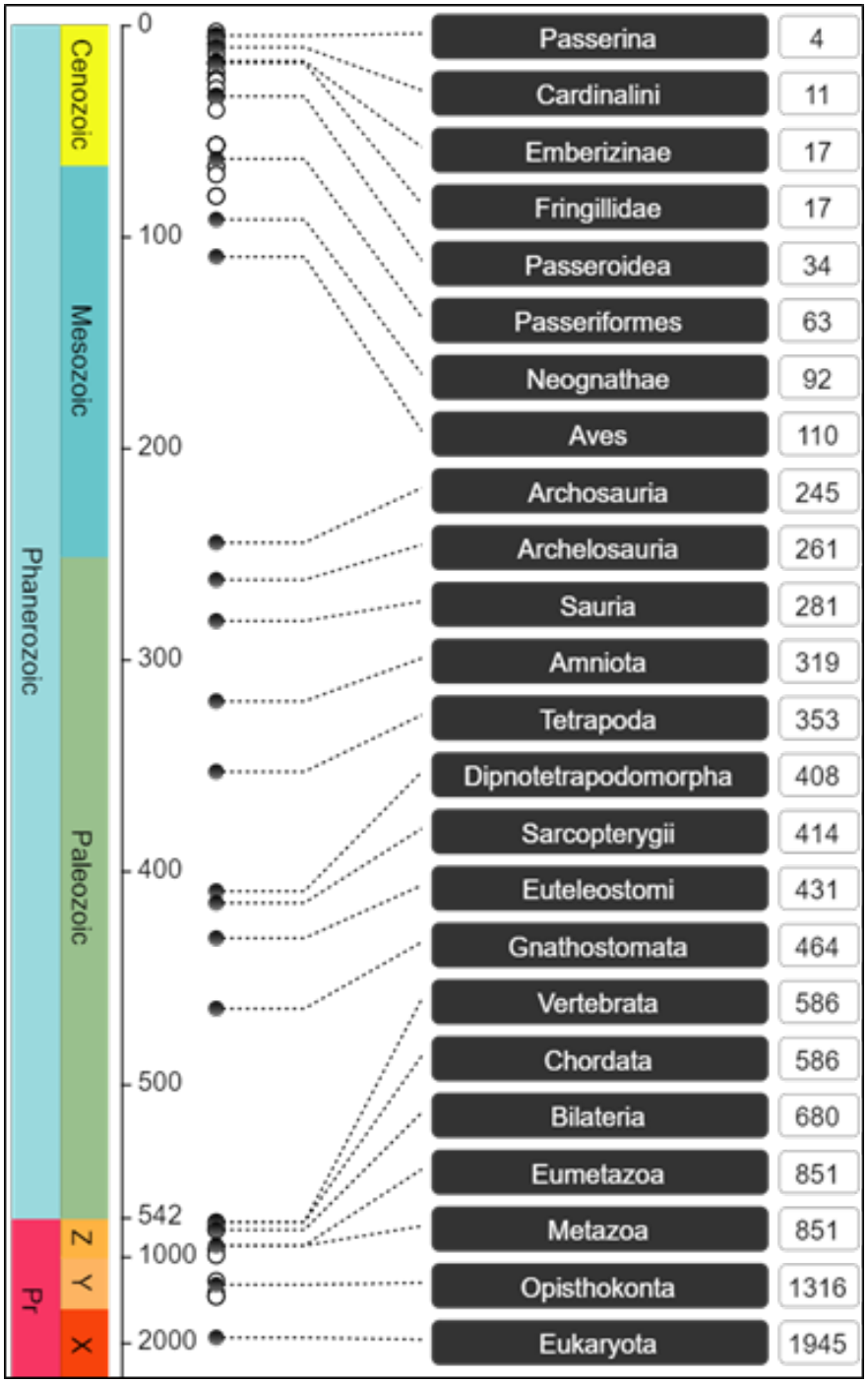






Posted on 22 Sep 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.166384820.08397260/v1 — This a preprint. Version of Record available at https://doi.org/10.1007/s00239-023-10106-3

Posted on 22 Sep 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.166384820.08397260/v1 — This a preprint. Version of Record available at https://doi.org/10.1007/s00239-023-10106-3



Posted on 22 Sep 2022 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.166384820.08397260/v1 — This a preprint. Version of Record available at https://doi.org/10.1007/s00239-023-10106-3