# Open Tree of Life

University of California, Merced
ejmctavish@ucmerced.edu, Twitter:@snacktavish
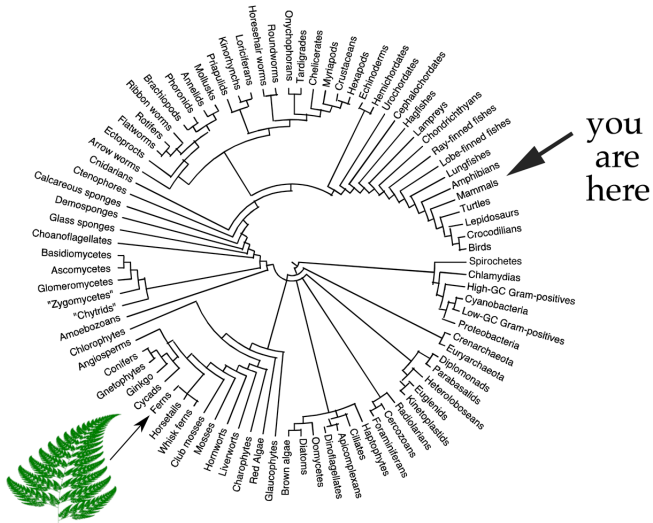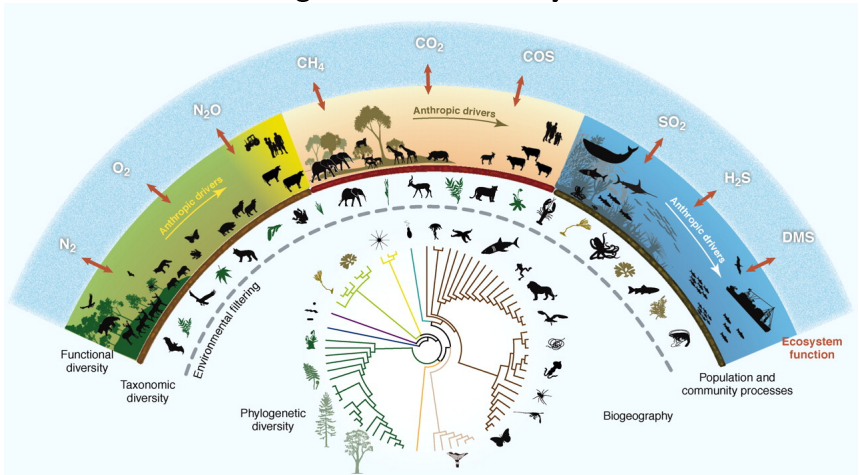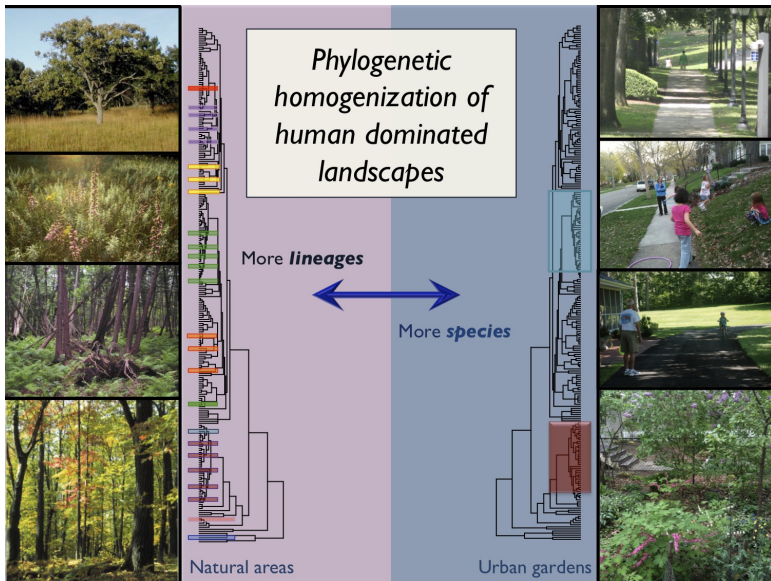
you
are
here

Image Ethan Hein

# Evolutionary context provides a framework for **understanding and conserving** global biodiversity



Naeem et al. (2012)

Phylogenetic homogenization of human dominated landscapes

More *lineages*

More *species*

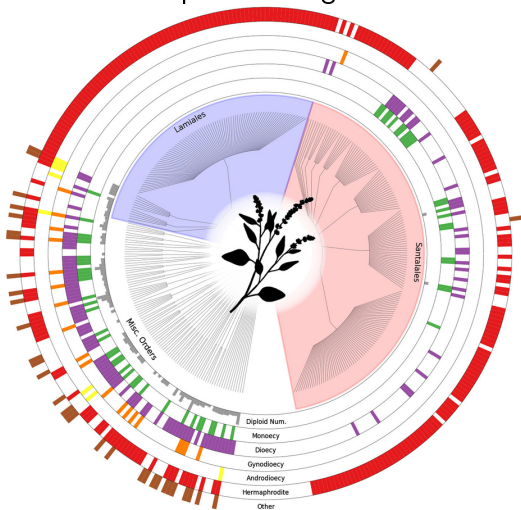Natural areas

Urban gardens

(example from the Cavender-Bares lab webpage)

Taxonomy is often used as a proxy for shared evolutionary history

# Evolutionary transitions in plant mating



Tree of Sex Consortium (2014)

**at best** taxonomy is a coarse representation of evolutionary history

**at worst** taxonomy is a mis-representation of evolutionary history

# **at worst** taxonomy is a mis-representation of evolutionary history

RESEARCH ARTICLE

## Traditional Taxonomic Groupings Mask Evolutionary History: A Molecular Phylogeny and New Classification of the Chromodorid Nudibranchs

Rebecca Fay Johnson, Terrence M. Gosliner



Johnson and Gosliner (2012)

*Why do we use taxonomy as a proxy for shared evolutionary history?*

*Why do we use taxonomy as a proxy for shared evolutionary history?*

**Phylogenies:**

  don't include all the species we are interested in,

  keep changing,

  are hard to access.

*Goal*: Build a tree of all life.

*Goal*: Build a tree of all life.

    Every named species

    Updated as new data becomes available

    Freely and easily accessible

*Goal*: Build a tree of all life.
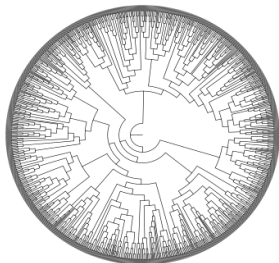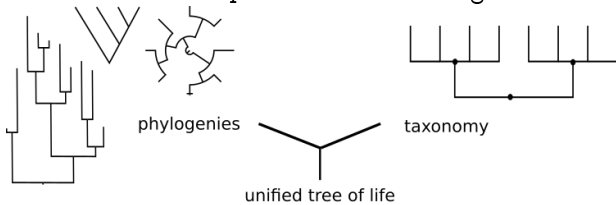
    Every named species
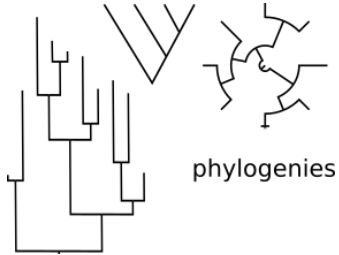
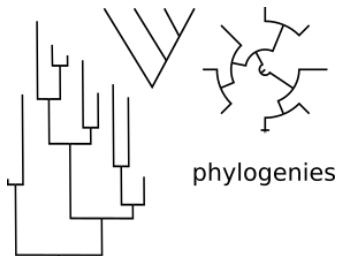    Updated as new data becomes available

    Freely and easily accessible

Available at `tree.opentreeoflife.org`

Major collaborative project, NSF AVATOL 2012-2016 to 11 PI's at 10 institutions; NSF ABI 2018-2022

# tree.opentreeoflife.org



phylogenies

taxonomy

unified tree of life

phylogenies

phylogenies

**Current synthetic tree** (as of Dec 2019)

- 1,216 community curated representative phylogenies

- 87,740 tips from phylogeny

- New draft (12.3) posted December 23, 2019, will be updated with more input trees every few months

Redelings and Holder (2017)

Need to build on existing phylogenetic information.

Need to build on existing phylogenetic information.

PLOS | BIOLOGY

**Perspective**

# Lost Branches on the Tree of Life

**Bryan T. Drew[1]\*, Romina Gazis[2], Patricia Cabezas[3,4], Kristen S. Swithers[5], Jiabin Deng[1], Roseana Rodriguez[1], Laura A. Katz[5], Keith A. Crandall[4], David S. Hibbett[2], Douglas E. Soltis[1,6]**

**1** University of Florida, Gainesville, Florida, United States of America, **2** Clark University, Worcester, Massachusetts, United States of America, **3** Brigham Young University, Provo, Utah, United States of America, **4** George Washington University, Washington, DC, United States of America, **5** Smith College, Northampton, Massachusetts, United States of America, **6** Florida Museum of Natural History, Gainesville, Florida, United States of America

Drew et al. PLoS Biology 2013

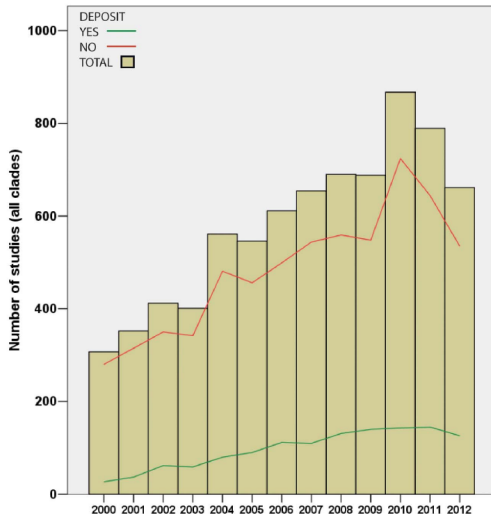# only 16% of phylogenies published 2000-2013 are digitally available



**Figure 1. Overview of total number of publications surveyed from animal, fungus, seed plant, microbial eukaryote, archaea, and bacteria literature (indicated in red), and the number of those publications that archived their trees and alignments in either Dryad or TreeBASE (indicated in green).**

**only 16% of phylogenies published 2000-2013 are digitally available** (Drew et al. PLoS Biology 2013)

**20% of phylogenies published 2013-2018** (McTavish et al. BioEssays 2018)

## Adding phylogenetic data

- Trees can be uploaded from any source, does not have to be own data.

- Easy to use browser based interface

- Track curation attribution by name or pseudonym

- Files are json representation of NeXML phylogenetic data format

- Data store is hosted publicly on GitHub

github.com/OpenTreeOfLife/phylesystem-1
McTavish et al. Bioinformatics 2015

## Input trees

- 260 individual curators of 4,451 uploaded studies
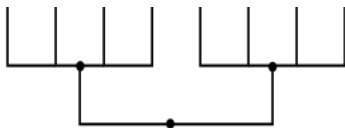- Easy to use browser based interface
- Trees can be uploaded from any source, does not have to be own data
- Files are json representation of NeXML phylogenetic data format
- Data store is hosted publicly on GitHub

github.com/OpenTreeOfLife/phylesystem-1

McTavish et al. (2015)

taxonomy

taxonomy

- 2.4 million named taxa
- Merges NCBI, Index fungorum, Silva, IRMNG, GBIF and other taxonomies
- Scaffold for combining ranked phylogenetic estimates
- New drafts released as inputs change

Rees and Cranston (2017)

# Open Tree taxonomy: **Tramea lacerata**

*The current taxonomy version is ott3.1 (click for more information). See the OTT documentation for an explanation of the taxon flags used below, e.g.,* `extinct`

### Taxon details

species **Tramea lacerata**     ncbi:126233 (gbif:1428475, irmng:10341497)     `sibling_higher` (OTT id 340907)
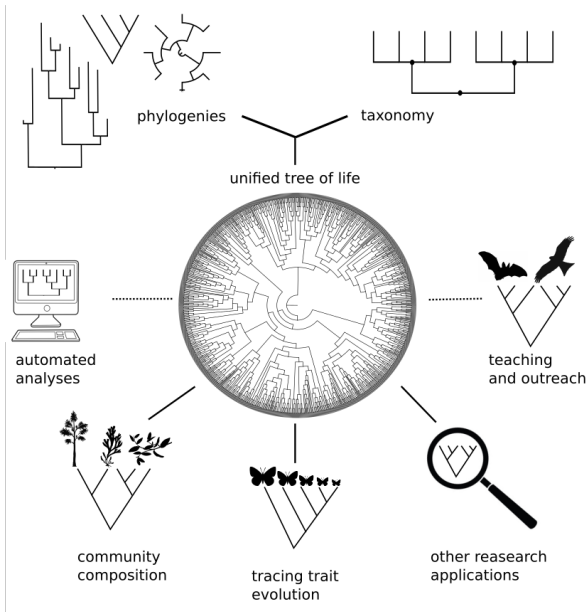
View this taxon in the current synthetic tree

### Lineage

**life** › **cellular organisms** › **Eukaryota** › **Opisthokonta** › **Holozoa** › **Metazoa** › **Eumetazoa** › **Bilateria** ›
**Protostomia** › **Ecdysozoa** › **Panarthropoda** › **Arthropoda** › **Mandibulata** › **Pancrustacea** › **Hexapoda** ›
**Insecta** › **Dicondylia** › **Pterygota** › **Palaeoptera** › **Odonata** › **Epiprocta** › **Anisoptera** › **Cavilabiata** ›
**Libellulidae** › **Tramea**

**Taxonomic versioning**

OpenTree Taxonomy (OTT) v.3.2 published December 29, 2019

Includes newest NCBI and GBIF updates, and curated amendments

phylogenies

taxonomy

unified tree of life

automated analyses

teaching and outreach

community composition

tracing trait evolution

other reasearch applications

McTavish et al. (2017)

Open Tree resources are available via a range of implementations

- Browser interface, `tree.opentreeoflife.org`
- Open Tree of Life API
-  and python-opentree python wrapper
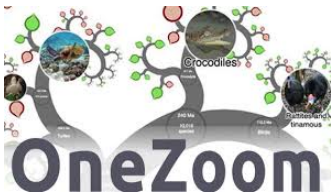- R Open Tree of Life (rotl) 

CC0 license provides fully open access for downstream re-usability

CC0 license provides fully open access for downstream re-usability

Open Tree provides the tree backend for:

**Conclusions**

Phylogeny not taxonomy, can and should be used to capture evolutionary relationships among taxa.

**Conclusions**

- Phylogeny not taxonomy, can and should be used to capture evolutionary relationships among taxa.

- Open Tree cross-links phylogenetic and taxonomic information across biodiversity informatics resources

**Conclusions**

- Phylogeny not taxonomy, can and should be used to capture evolutionary relationships among taxa.

- Open Tree cross-links phylogenetic and taxonomic information across biodiversity informatics resources

- A variety of tools and approaches provides wide access to Open Tree resources

**In progress!**

Automated updating with new genetic data

Branch lengths / Node ages

Private data stores

Custom synthesis

Holder, McTavish ABI
Cranston

**Workshop today:**

- Browser interface, `tree.opentreeoflife.org`
- Standardizing taxon names
- Accessing Open Tree data using rotl (R) and python-opentree (python); including:

  Getting existing trees for arbitrary sets of taxa
  Visualizing conflict between estimates

- Uploading and curating phylogenies

# Thank You

NSF ABI 1759846
Mark Holder
Karen Cranston
Jim Allman
Ben Redelings
Luna Luisa Sanchez Reyes



NSF AVATOL 1208809

AVATOL PI'S: Burleigh, Crandall, Cranston, Gude, Hibbett, Holder, Katz, Ree, Smith, Soltis, Williams

## Citations

Johnson, R. F. and Gosliner, T. M. (2012). Traditional Taxonomic Groupings Mask Evolutionary History: A Molecular Phylogeny and New Classification of the Chromodorid Nudibranchs. *PLOS ONE*, 7(4):e33479.

McTavish, E. J., Drew, B. T., Redelings, B., and Cranston, K. A. (2017). How and Why to Build a Unified Tree of Life. *BioEssays*, 39(11).

McTavish, E. J., Hinchliff, C. E., Allman, J. F., Brown, J. W., Cranston, K. A., Holder, M. T., Rees, J. A., and Smith, S. A. (2015). Phylesystem: a git-based data store for community-curated phylogenetic estimates. *Bioinformatics*, 31(17):2794–2800.

Naeem, S., Duffy, J. E., and Zavaleta, E. (2012). The Functions of Biological Diversity in an Age of Extinction. *Science*, 336(6087):1401–1406.

Redelings, B. D. and Holder, M. T. (2017). A supertree pipeline for summarizing phylogenetic and taxonomic information for millions of species. *PeerJ*, 5:e3058.

Rees, J. A. and Cranston, K. (2017). Automated assembly of a reference taxonomy for phylogenetic data synthesis. *Biodiversity Data Journal*, (5).

Tree of Sex Consortium (2014). Tree of Sex: A database of sexual systems. *Scientific Data*, 1:140015.