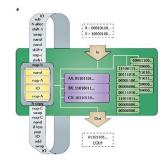
Phylogenies: how and why to track them in artificial life

ALife 2023 Sapporo, Japan Alexander Lalejini Emily Dolson Matthew Andres Moreno

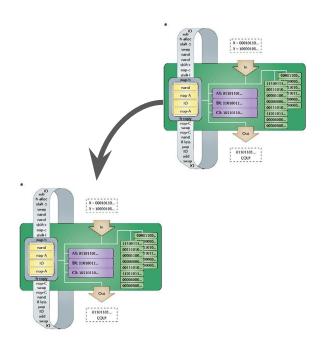
Agenda

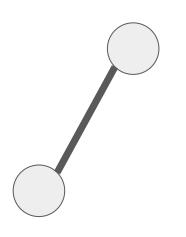
- 1. What is a phylogeny and what can it tell you?
- 2. Phylogenetic metrics & visualizations
- 3. Interactive demo
 - a. Centralized phylogenetic tracking
 - b. Decentralized phylogenetic inference
 - c. Visualizations



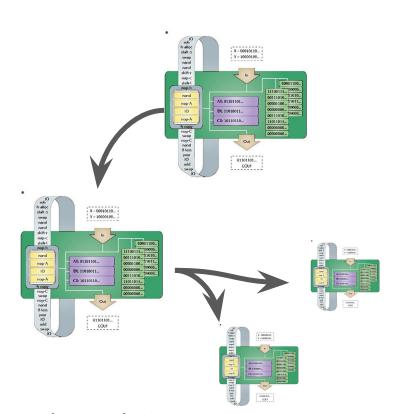


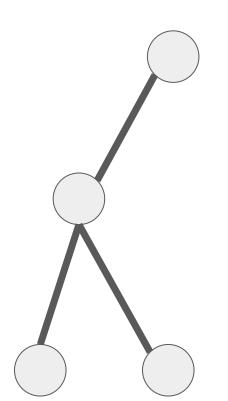
vertex:
 "taxonomic unit"



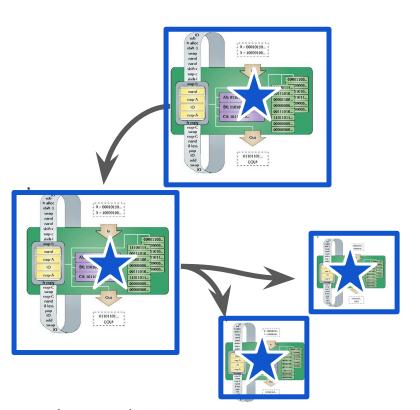


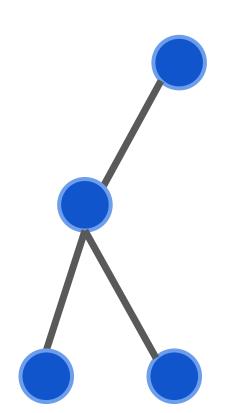
vertex:
 "taxonomic unit"



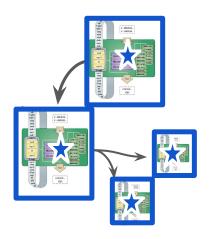


vertex:
 "taxonomic unit"



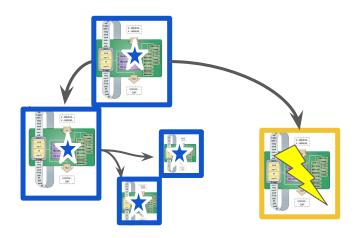


vertex:
"taxonomic unit"

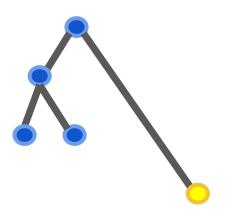


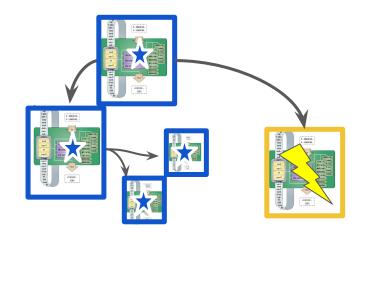


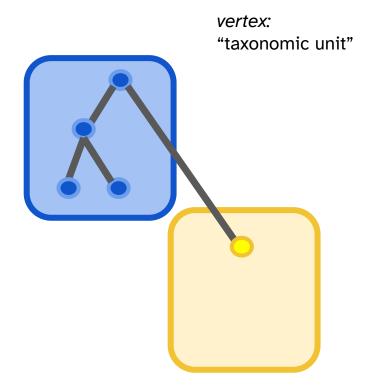
vertex:
 "taxonomic unit"

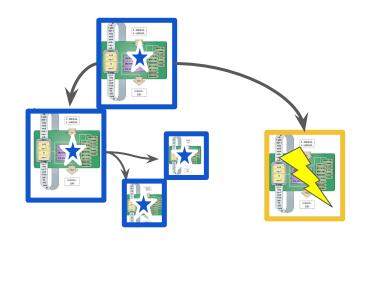


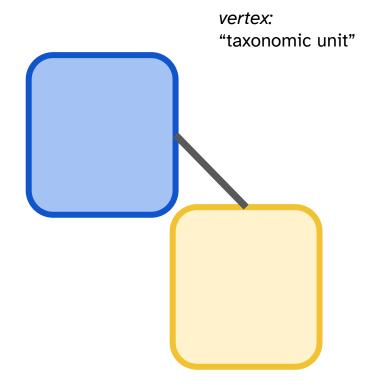
vertex:
"taxonomic unit"

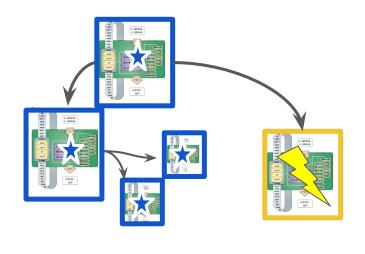




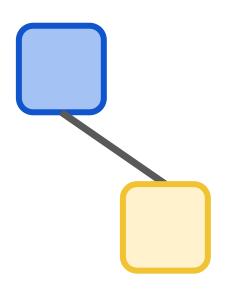


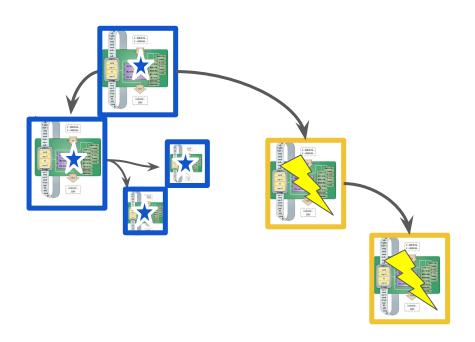




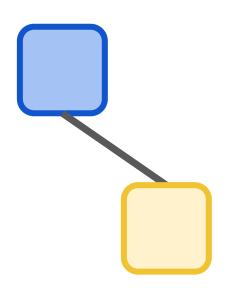


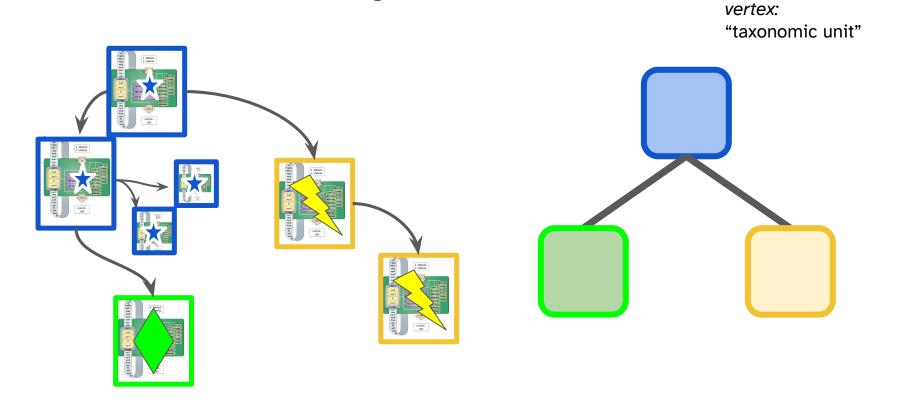
vertex:
"taxonomic unit"

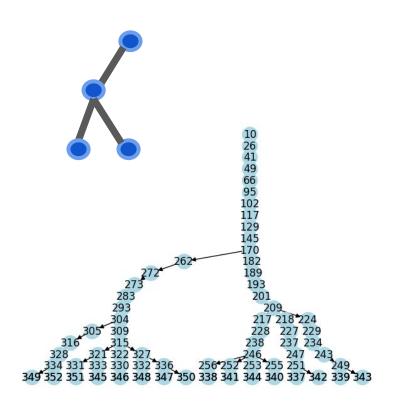


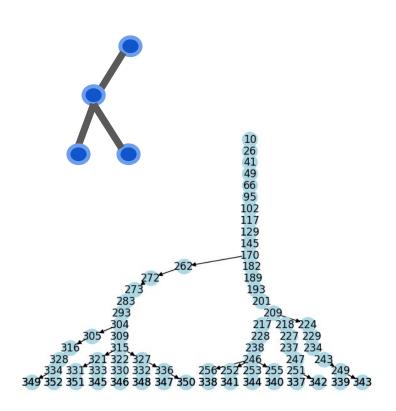


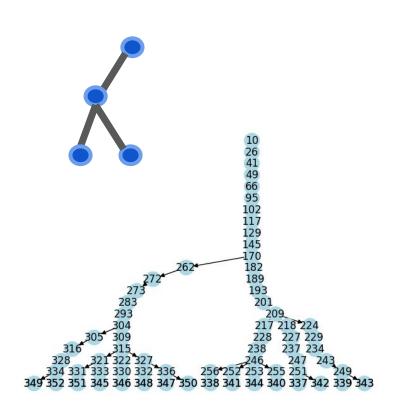
vertex:
"taxonomic unit"



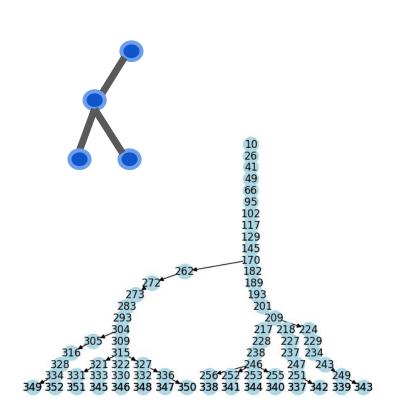




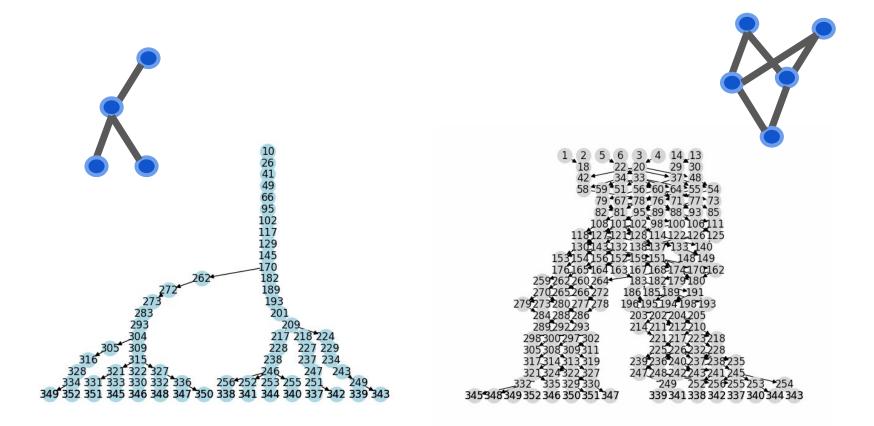


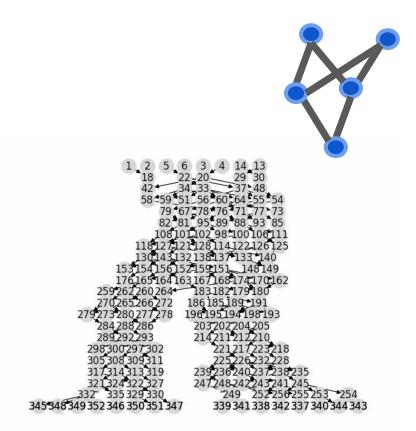


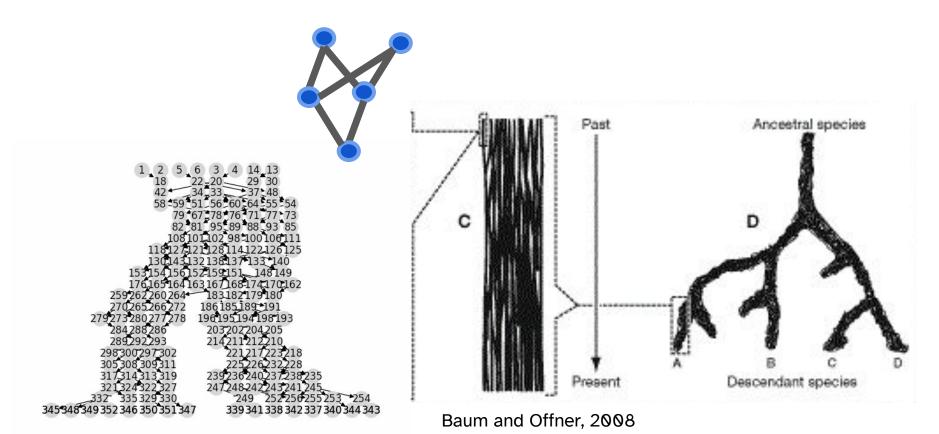




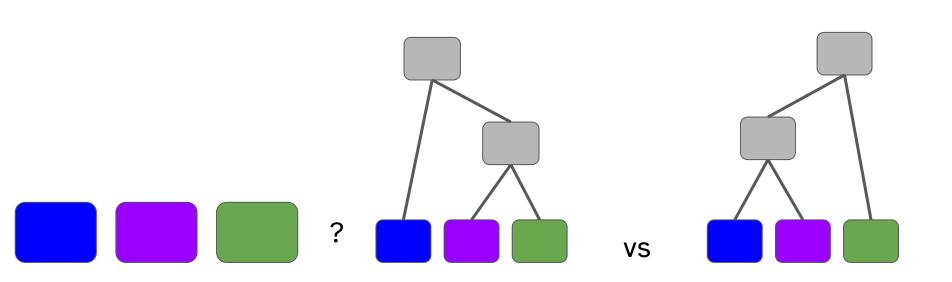


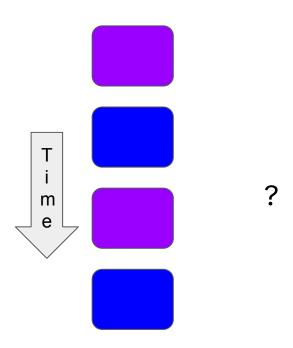


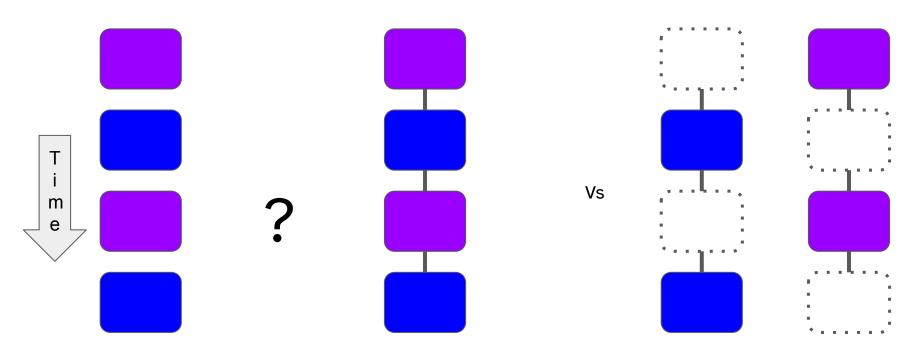


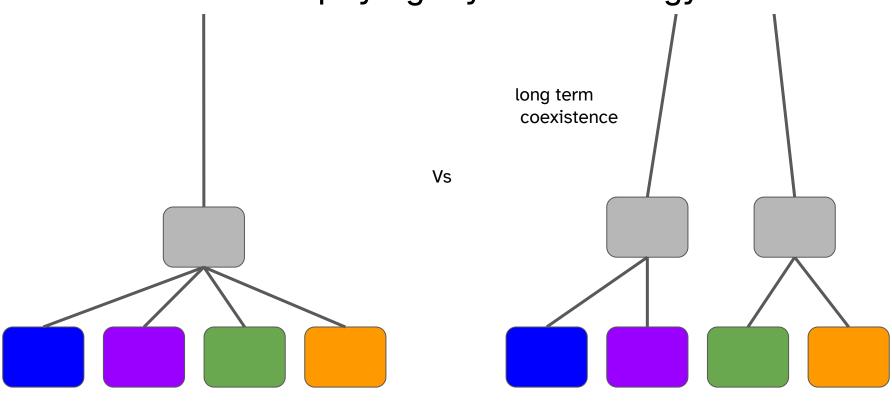


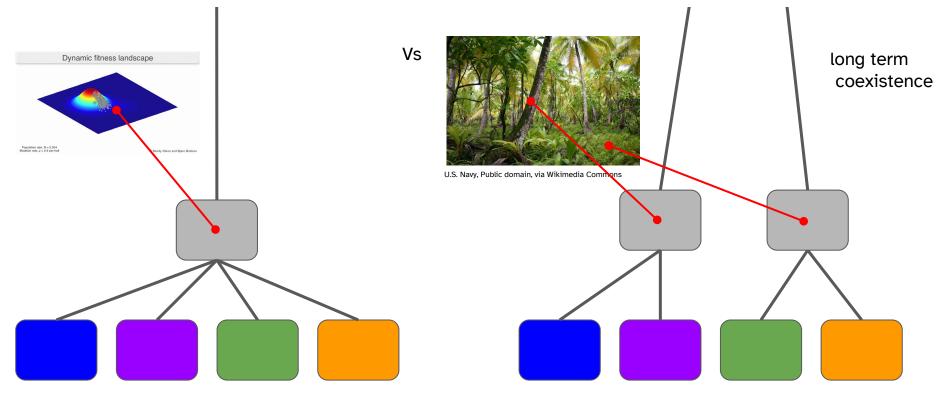


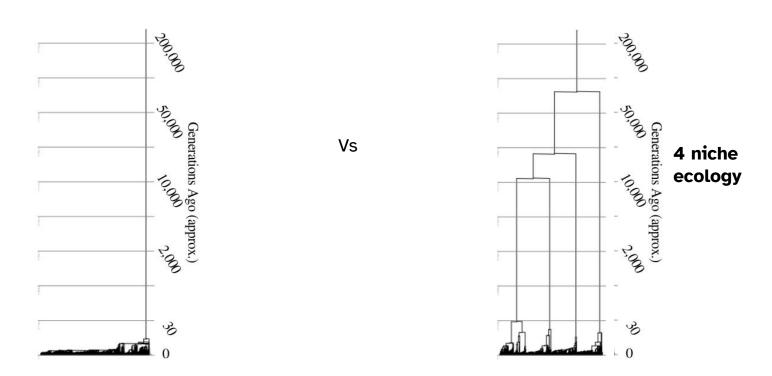






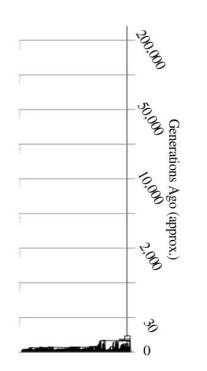




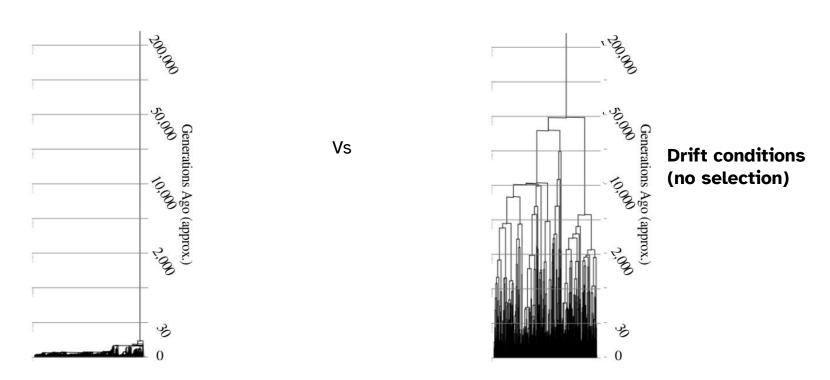


Intro: what does a phylogeny tell?

Intro: what does a phylogeny tell? Selection strength



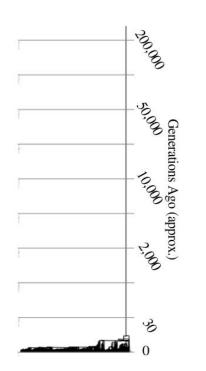
Intro: what does a phylogeny tell? Selection strength



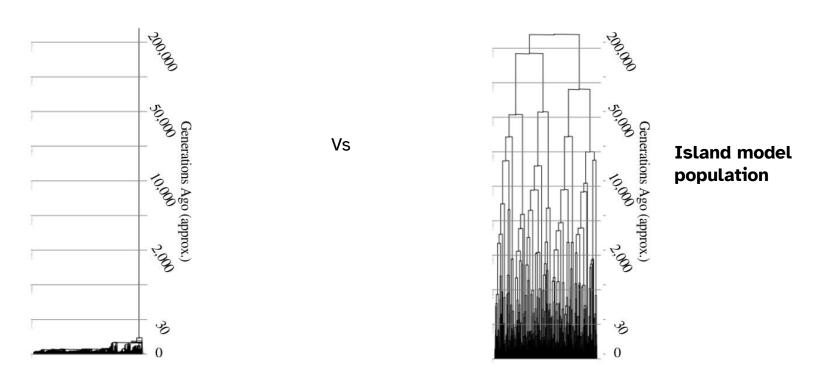
Moreno et al., 2023

Intro: what does a phylogeny tell? Spatial structure

Intro: what does a phylogeny tell? Spatial structure

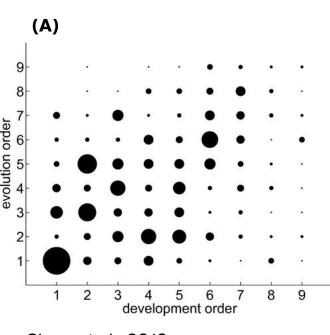


Intro: what does a phylogeny tell? Spatial structure

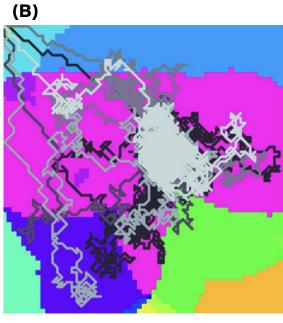


Moreno et al., 2023

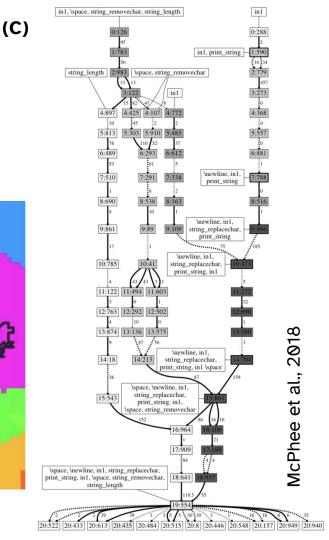
Example Applications of Phylogenetics in Digital Evolution



Clune et al., 2012



Dolson and Ofria, 2017

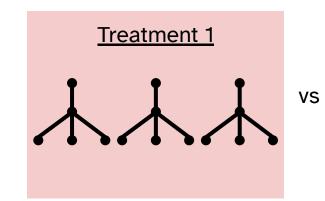


Quantifying phylogenies (& lineages)

Often we want to quantify different characteristics of evolutionary histories

Summarize evolutionary histories across many replicates

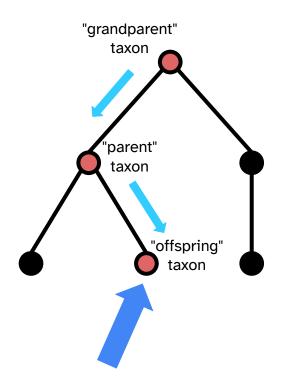
Compare the characteristics of evolutionary histories across experimental treatments



Treatment 2

Quantifying the history of a single taxon

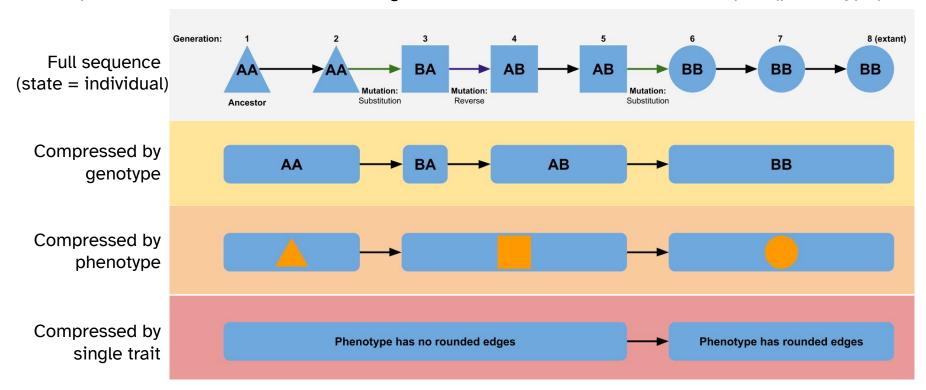
I.e., quantifying lineages



What can we measure along a single lineage?

Ancestry can be abstracted into sequences of states

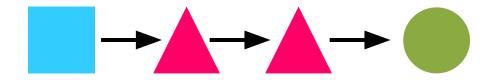
Example: individuals have 2-character genomes that are translated into shapes (phenotype)



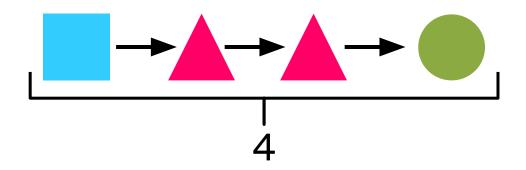
Ancestry can be abstracted into sequences of states

Example: individuals have 2-character genomes that are translated into state This is also true for phylogenies! Taxa ("states") may represent individuals, phenotypes, genotypes, etc!

Summarizing state sequences (lineages)



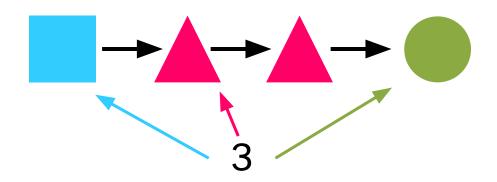
Summarizing state sequences (lineages): **Length**



Meaning of a measurement depends on what states represent.

- States = individuals: length represents generations
- States = phenotypes: length represents number of times the phenotype changed along the lineage

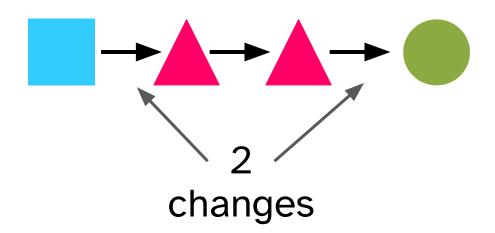
Summarizing state sequences (lineages): **Unique states**



Meaning of a measurement depends on what states represent.

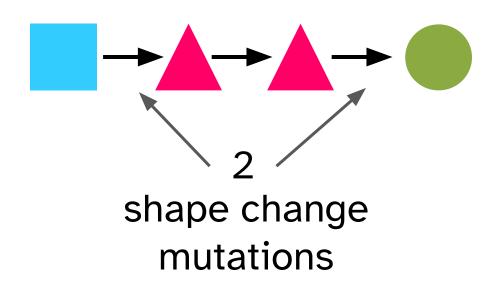
 States = phenotypes: how many different phenotypes are represented in an individual's ancestry?

Summarizing state sequences (lineages): **Volatility**



 E.g., if states = phenotypes, a changing environment might result in lineages with greater volatility

Summarizing state sequences (lineages): **Mutation accumulation**



Caveats for sexually reproducing populations

Sexual reproduction (or any form of horizontal gene transfer) makes ancestry messy!

• an individual's ancestral history is no longer a linear chain

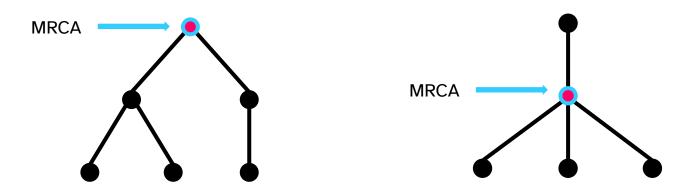
Strategies for dealing with sexual lineages:

- Build lineages based on sites in a genome
- Apply a lossy compression to reduce sexual lineages into linear sequences
- Average metrics over all possible ancestry paths



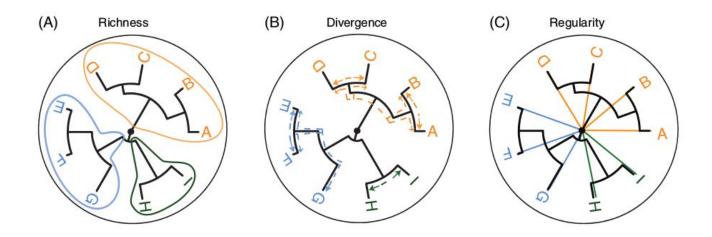
What ways can we quantify the differences between these two phylogenies?

depth of most-recent common ancestor (MRCA)



- Can measure relative to the extant population (leaves of the phylogeny) or the original ancestor (root of the phylogeny)
- A distant (from extant population) MRCA can indicate stable, long-term coexistence
- A recent MRCA indicates a recent selective sweep
- Frequency of MRCA changes can indicate strength of selection on population

The size and topology of a phylogeny are informative.



Richness metrics quantify the total amount of evolutionary history contained in a set of taxa.

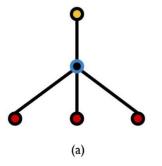
Divergence metrics quantify how distinct the taxa in the extant population are from each other.

Regularity metrics quantify how balanced the branches are in the phylogeny.

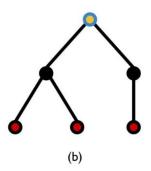
Phylogenetic richness

Richness metrics quantify the total amount of evolutionary history contained in a set of taxa.

- Phylogenetic diversity number of nodes in the minimum spanning tree from the MRCA to all extant taxa
- Sum of pairwise distances calculate pairwise distance between all taxa and sum them



Phylogenetic diversity = 4

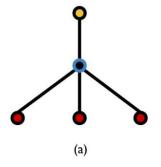


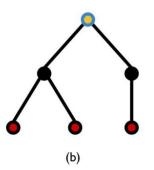
Phylogenetic diversity = 6

Phylogenetic divergence

Divergence metrics quantify how distinct the taxa in the extant population are from each other.

One option: average the pairwise distances across all taxa in the extant population





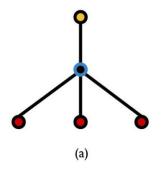
Mean pairwise dist. among extant = 2

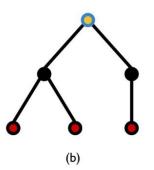
Mean pairwise dist. among extant = 3.33

Phylogenetic regularity

Regularity metrics quantify how balanced the branches are in the phylogeny.

 One option: calculate the variance of pairwise distances between all taxa in extant population





Variance pairwise dist. among extant = 0

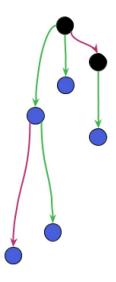
Variance pairwise dist. among extant = 1.33

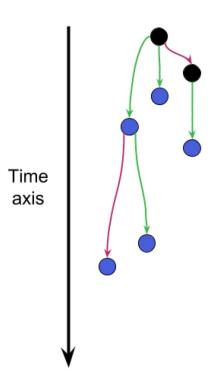
Bibliography

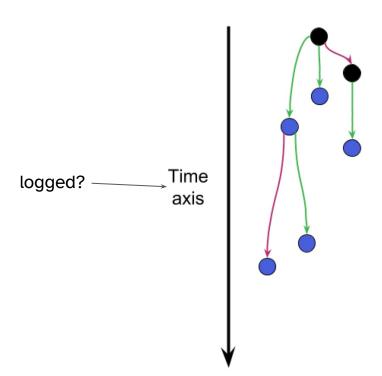
Dolson, E., Lalejini, A., Jorgensen, S. & Ofria, C. Interpreting the Tape of Life: Ancestry-Based Analyses Provide Insights and Intuition about Evolutionary Dynamics. Artificial Life 26, 58–79 (2020).

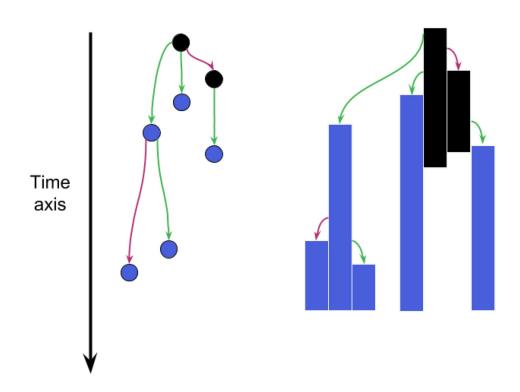
Tucker, C. M. et al. A guide to phylogenetic metrics for conservation, community ecology and macroecology: A guide to phylogenetic metrics for ecology. Biol Rev 92, 698–715 (2017).

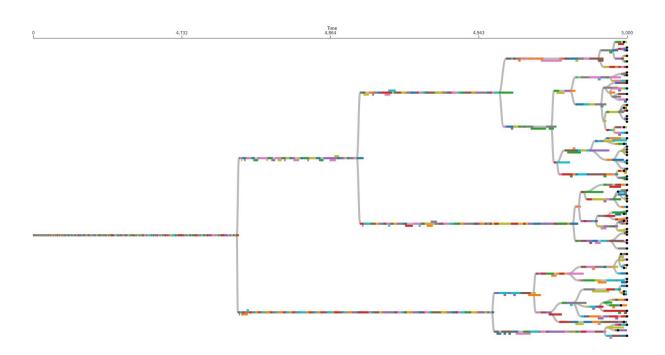
Phylogeny Visualizations





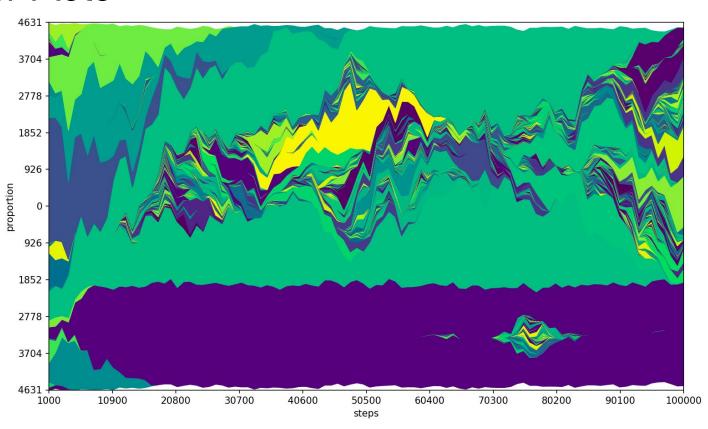




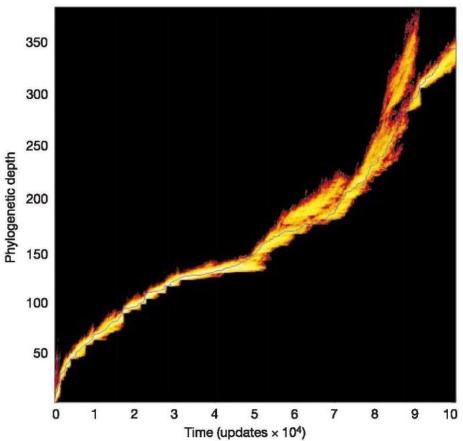




Muller Plots

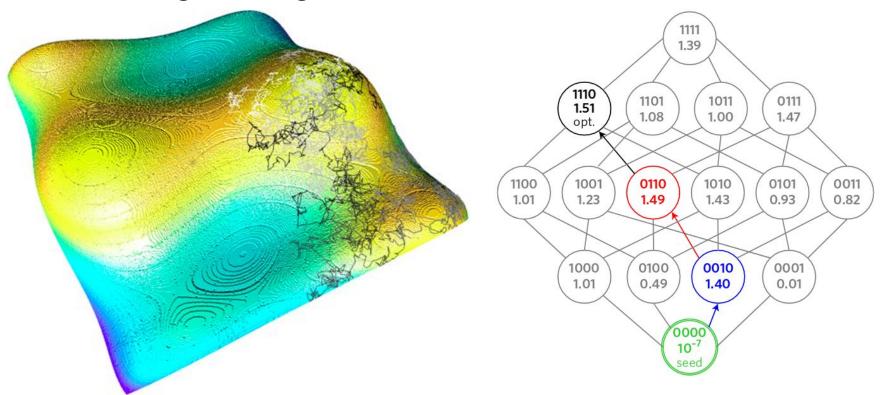


Flame graphs



Lenski et. al, 2003

Overlaying phylogenies

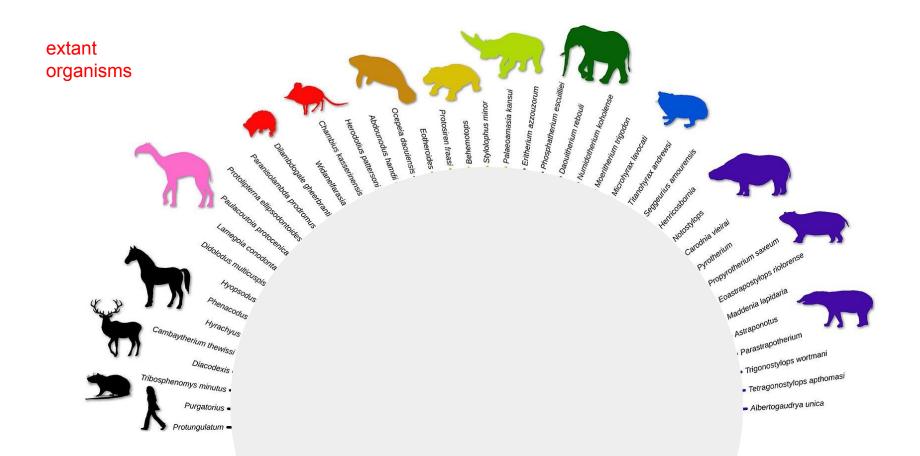


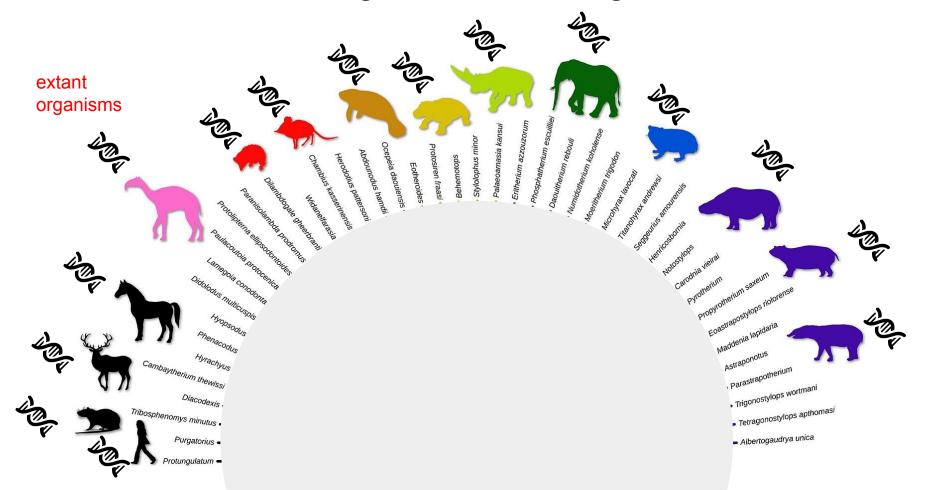
Ogbunugafor and Epstein, 2017

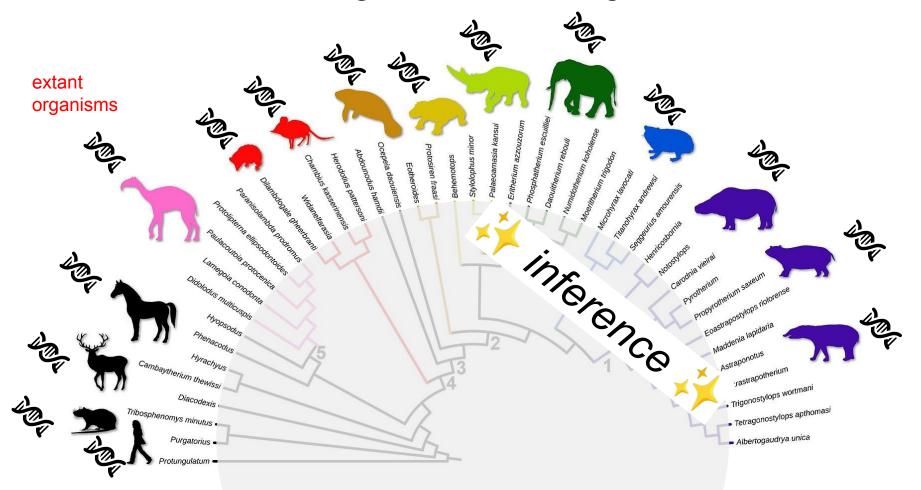
Centralized phylogeny tracking demo

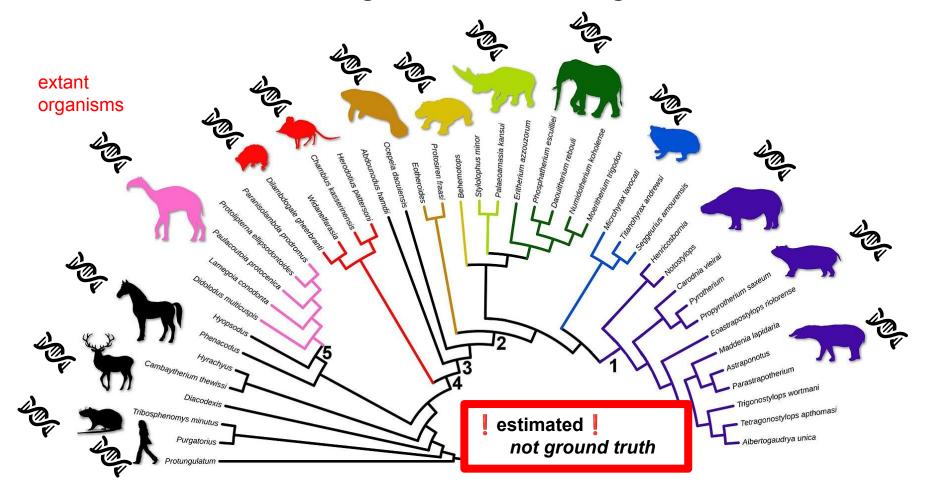


Carlos Jones/ORNL, CC BY 2.0 https://creativecommons.org/licenses/by/2.0, via Wikimedia Commons

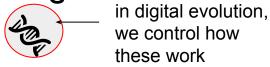








Decentralized Phylogenetic Tracking in digital evolution, we control how these work extant organisms parastrapotherium Trigonostylops wortmani Tetragonostylops apthomasi Tribosphenomys minutus . stimated ! Albertogaudrya unica not ground truth



Q: how to design a genome so phylogenetic reconstruction is easy & informative?

Decentralized Phylogenetic Tracking

Q: how to design a genome so phylogenetic reconstruction is easy & informative?

attach as neutral "annotation" to any digital genome

"hereditary stratigraphic column"

genetic material designed to facilitate phylogenetic reconstruction (Moreno et al., 2022)

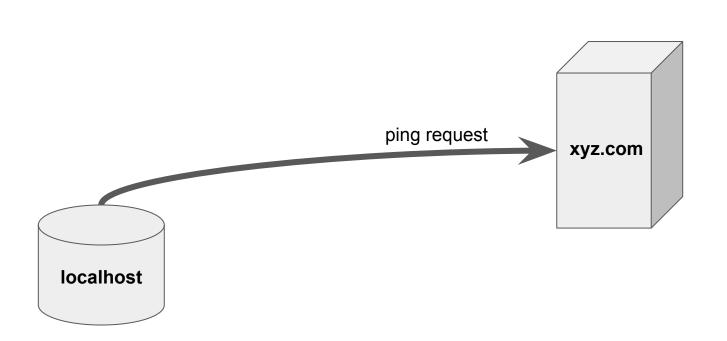
hstrat demo

(with "pinging" genomes)

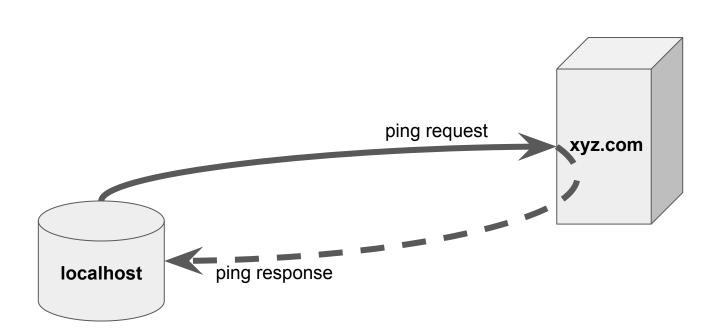
interactive Colab notebook @



hop th.ru

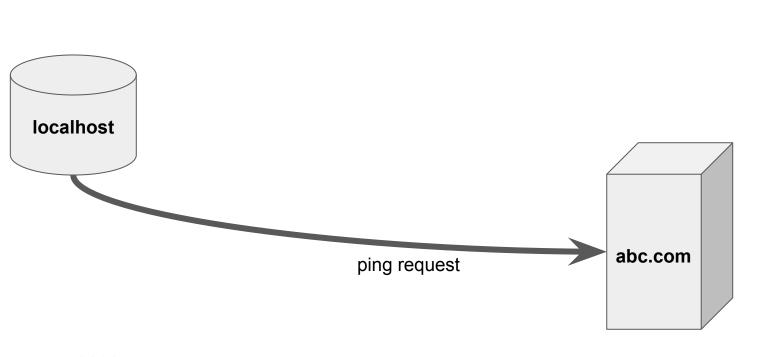




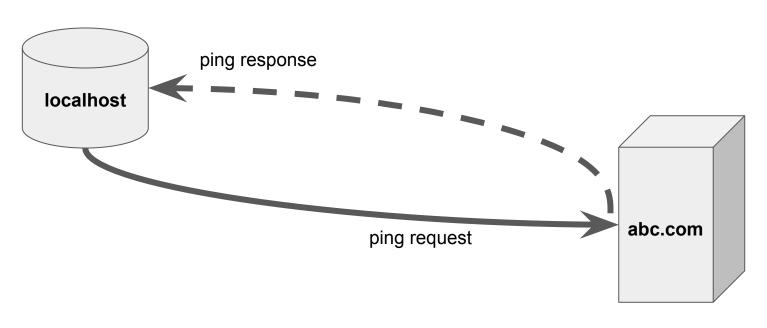




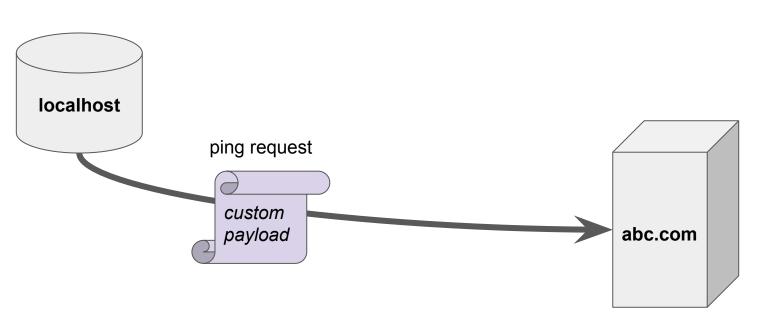




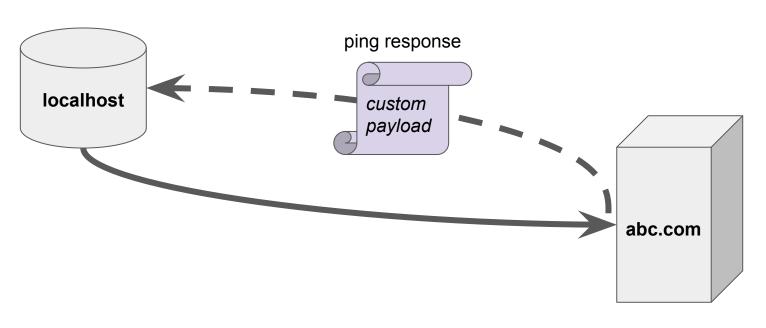








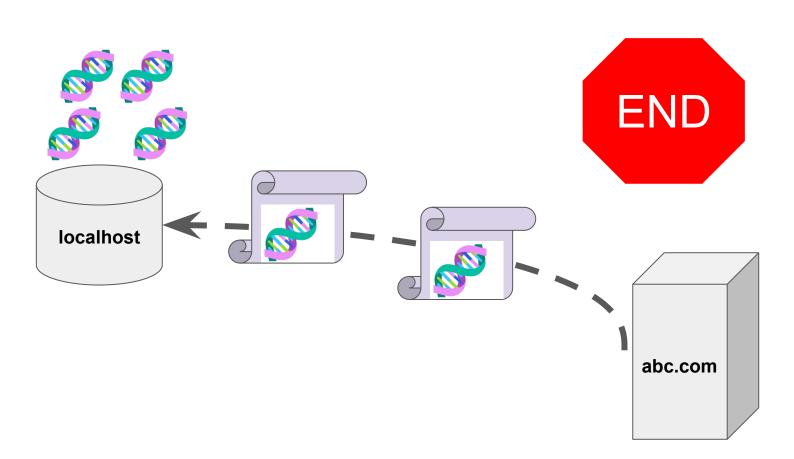




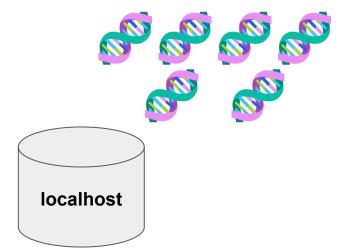




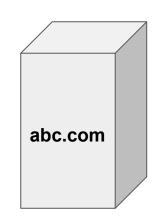


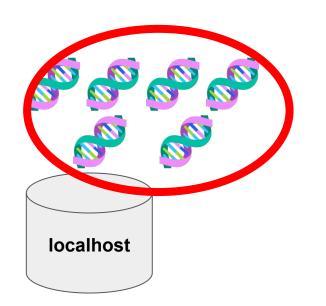




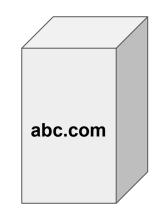


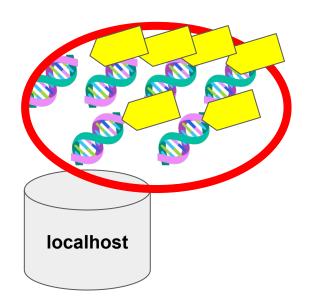




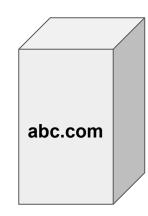


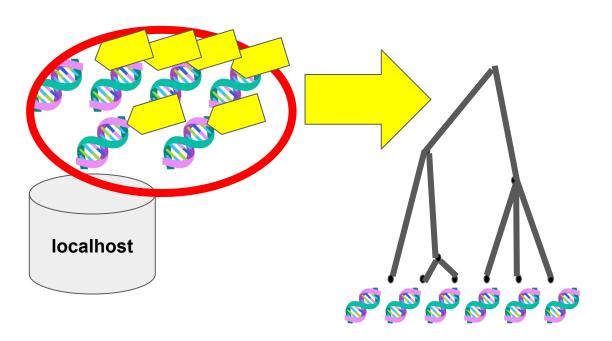


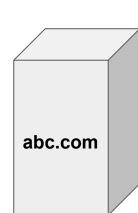














abc.com

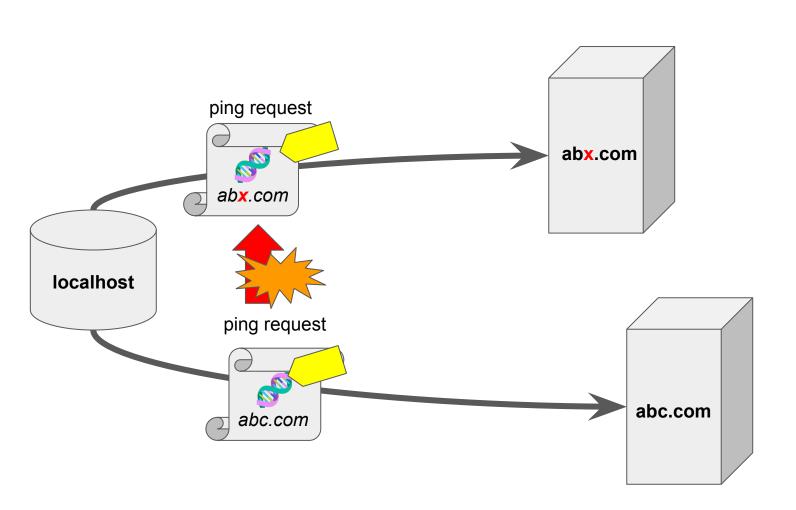
interactive Colab notebook @



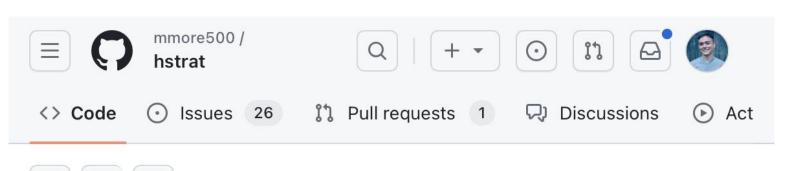
https://
hop
th.ru
/cg

abc.com









hstrat enables phylogenetic inference on distributed digital evolution populations

পাঁত View license

Code of conduct

☆ 4 stars
♀ 1 fork
⊙ 3 watching
- Activity

Public repository

interactive Colab notebook @



hopth.ru/cg

also,

shstrat Python library... **?

github.com /mmore500

/hstrat



Questions?

github.com/emilydolson/phylotrackpy

github.com/mmore500/hstrat

Bibliography

Adami, Christoph. (2006). Adami, C. Digital genetics: unravelling the genetic basis of evolution. Nature Rev. Genet. 7, 109-118. Nature reviews. Genetics. 7. 109-18. 10.1038/nrg1771.

Baum, David A., and Susan Offner. "Phylogenics & tree-thinking." The American Biology Teacher 70.4 (2008): 222-229.

Clune, J., Pennock, R. T., Ofria, C., & Lenski, R. E. (2012). Ontogeny tends to recapitulate phylogeny in digital organisms. The American Naturalist, 180(3), E54-E63.

Dolson, Emily, and Charles Ofria. "Spatial resource heterogeneity creates local hotspots of evolutionary potential." Artificial Life Conference Proceedings. Press, 2017.

Matthew Andres Moreno, Emily Dolson, and Santiago Rodriguez-Papa.. "Toward Phylogenetic Inference of Evolutionary Dynamics at Scale." Artificial Life Conference Proceedings (2023).

Moreno, M. A., Dolson, E., & Ofria, C. (2022). hstrat: a Python Package for phylogenetic inference on distributed digital evolution populations. *Journal of Open Source Software*, 7(80), 4866.

Murphy, T., VII. (2022, April 8). Harder Drive: Hard drives we didn't want or need. A Record of the Proceedings of SIGBOVIK 2022, 259–277. Pittsburgh, USA: ACH.

McPhee, N. F., Finzel, M. D., Casale, M. M., Helmuth, T., & Spector, L. (2018). A Detailed Analysis of a PushGP Run. Genetic Programming Theory and Practice XIV, 65.