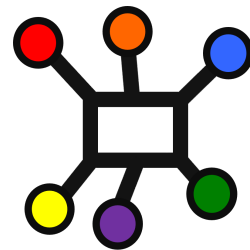


Phylogenetic Trees & Networks

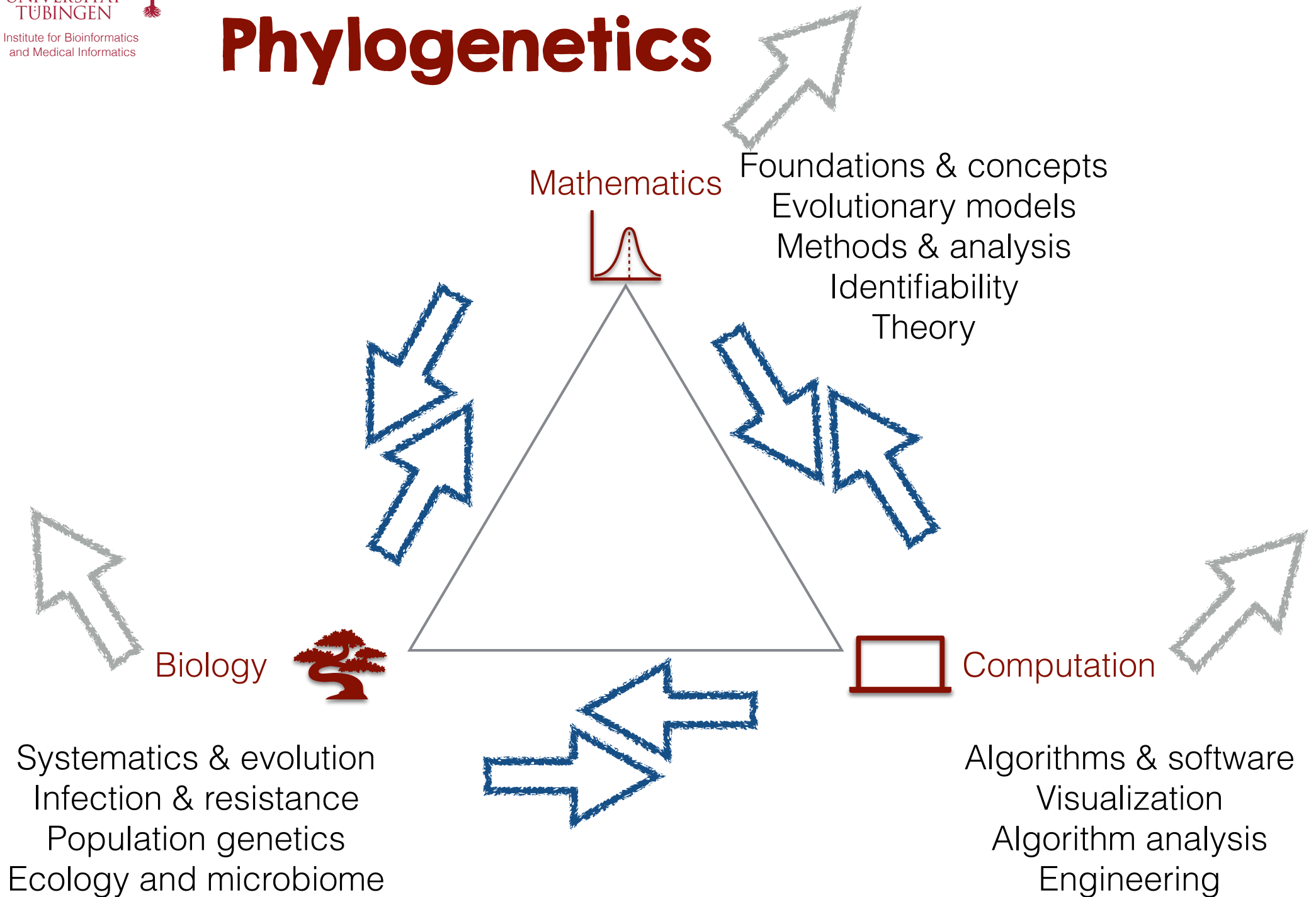
using SplitsTree6



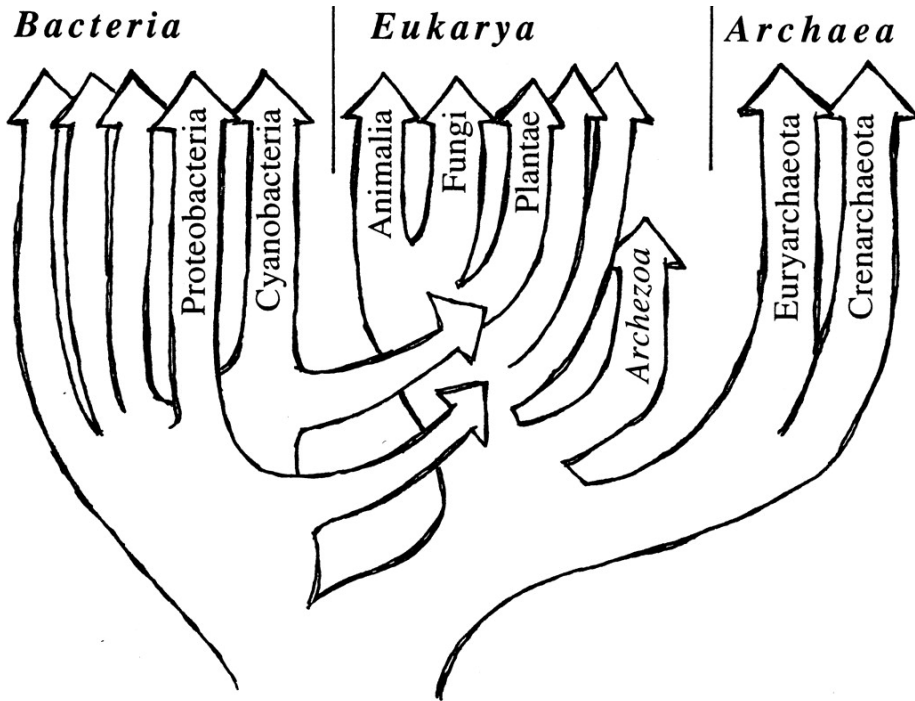
Daniel Huson



Phylogenetics



Phylogenetic trees

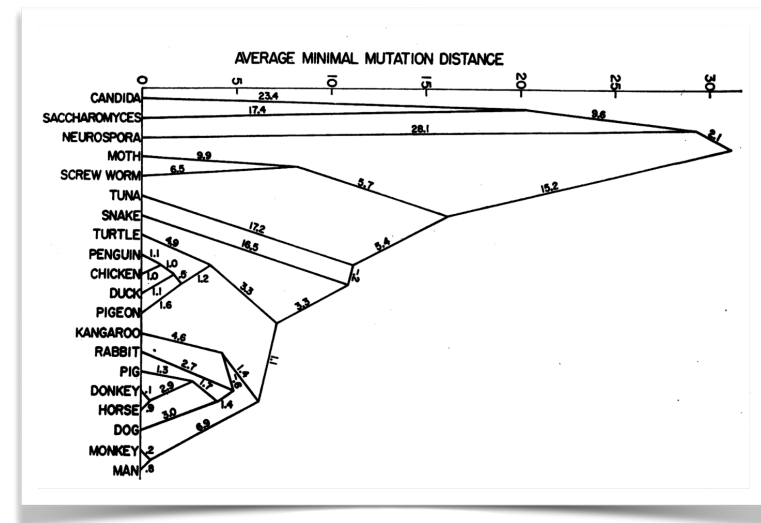
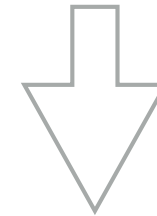


W. Ford Doolittle, Science 1999

Protein	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
1		1	13	15	15	13	11	14	15	15	16	16	17	29	29	30	33	64	62	68	Man (22)
2	1		12	15	14	12	11	13	15	14	15	15	16	28	29	29	32	63	61	67	Monkey (<i>Macacus mulatta</i>) (23)
3	13	12		9	8	6	7	8	13	13	13	14	15	26	27	27	30	61	59	65	Dog (24)
4	17	16	10		1	5	10	11	15	15	16	16	17	29	29	30	33	64	62	68	Horse (25)
5	16	15	8	1		4	9	10	14	14	15	15	16	28	28	29	32	63	61	67	Donkey (26)
6	13	12	4	5	4		7	8	13	12	13	13	14	26	27	27	30	61	59	65	Pig (27)
7	12	11	6	11	10	6		7	11	11	12	12	13	24	25	25	29	60	57	63	Rabbit (30)
8	12	13	7	11	12	7	7		13	13	14	14	15	27	27	28	31	62	60	66	Kangaroo (<i>Canopus canguru</i>) (28)
9	17	16	12	16	15	13	10	14		3	3	3	4	8	8	8	16	59	65	65	Pekin duck (29)
10	16	15	12	16	15	13	8	14	3		4	4	8	26	27	27	30	61	59	65	Pigeon (29)
11	18	17	14	16	15	13	11	15	3	4	2	9	27	27	28	31	62	60	66	Chicken (17)	
12	18	17	14	17	16	14	11	13	3	4	2	9	27	27	28	31	62	60	66	King penguin (<i>Aptenodytes patagonica</i>) (29)	
13	19	18	13	16	15	13	11	14	7	8	8	8	28	29	29	32	63	61	67	Snapping turtle (<i>Chelydra serpentina</i>) (31)	
14	20	21	30	32	31	30	25	30	24	24	28	28	30	33	34	37	68	66	72	Rattlesnake (<i>Crotalus adamanteus</i>) (32)	
15	31	32	29	27	26	25	26	27	26	27	26	27	27	38	35	38	69	67	73	Tuna (33)	
16	33	32	24	24	25	26	23	26	25	26	26	28	30	40	34	16	59	56	63	Screwworm fly (<i>Haematobia irritans</i>) (29)	
17	36	35	28	33	32	31	29	31	29	30	31	30	33	41	41	16	62	60	66	Moth (<i>Samia cynthia</i>) (34)	
18	63	62	64	64	64	64	62	66	61	59	61	62	65	61	72	58	59	56	62	<i>Neurospora (crassa)</i> (35)	
19	56	57	61	60	59	59	59	58	62	62	62	61	64	61	66	63	60	57	41	41	<i>Saccharomyces (oviformis) iso-1</i> (36)
20	66	65	66	68	67	67	67	68	66	66	66	65	67	69	69	65	61	61	41	41	<i>Candida (krusei)</i> (37)

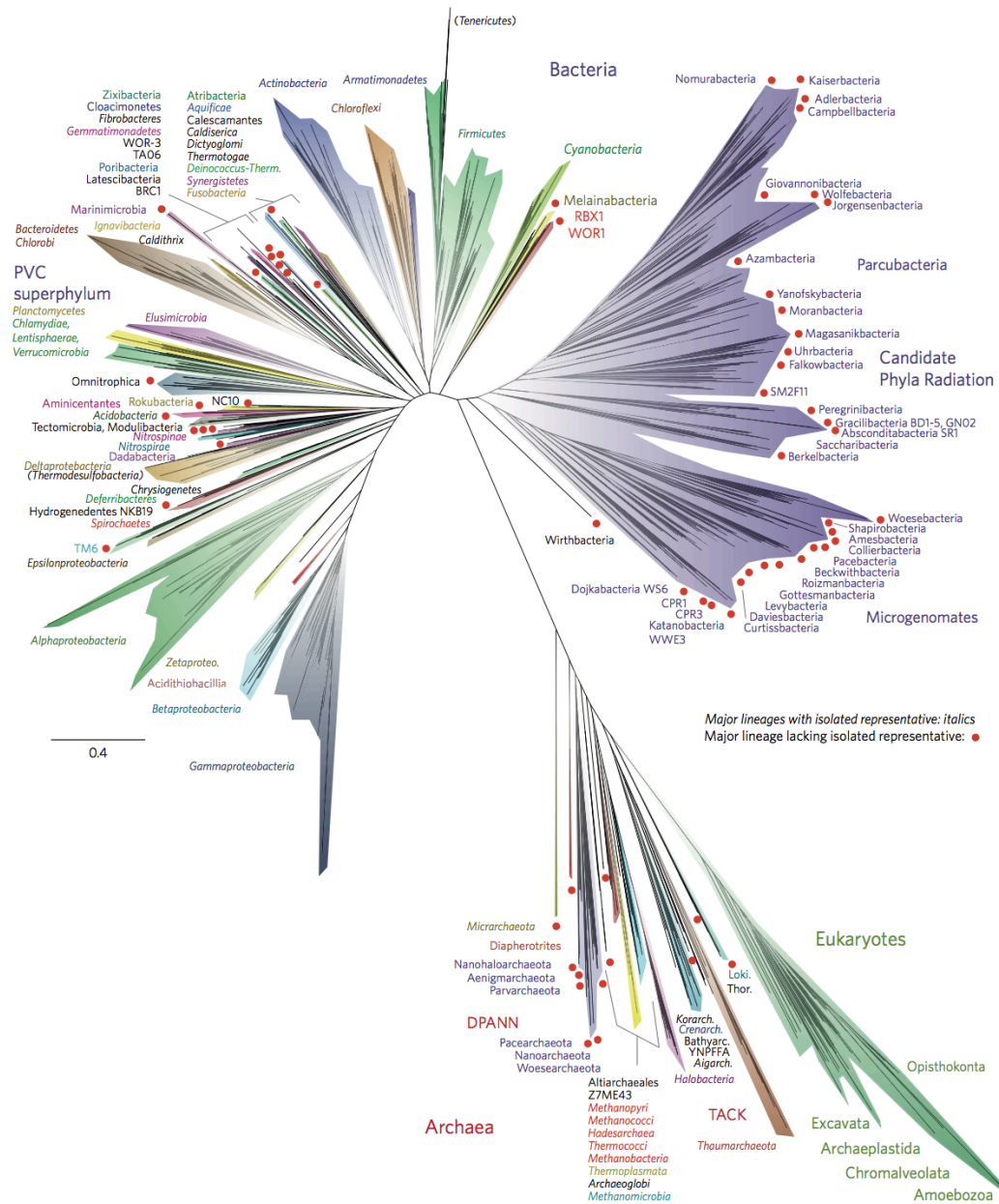
20 JANUARY 1967

281



Fitch & Margoliash, 1967

Phylogenetic trees

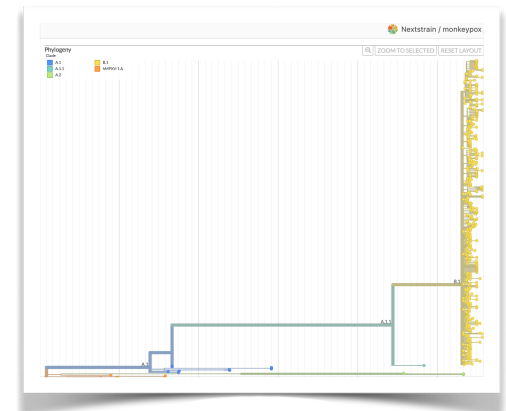


92 bacterial phyla, 26 archaeal phyla, and all eukaryotic phyla, built from 16 ribosomal proteins, Hug et al, 2016

Phylogenetic trees

Ever larger trees in applications:

- ▶ GTDB (Parks et al. 2020):
 - ▶ Reference taxonomy for archaea and bacteria
 - ▶ ~32,000 genomes, based on ANI and marker genes
- ▶ Silva (Quast et al. 2013)
 - ▶ 16S reference tree
 - ▶ 9.5 million sequences
- ▶ Nextstrain (Hadfield et al, 2018)
 - ▶ “Real-time tracking of pathogen evolution”



Phylogenetic trees

- ▶ Genome-scale data on populations for related species
- ▶ Phylogenetics meets population genetics

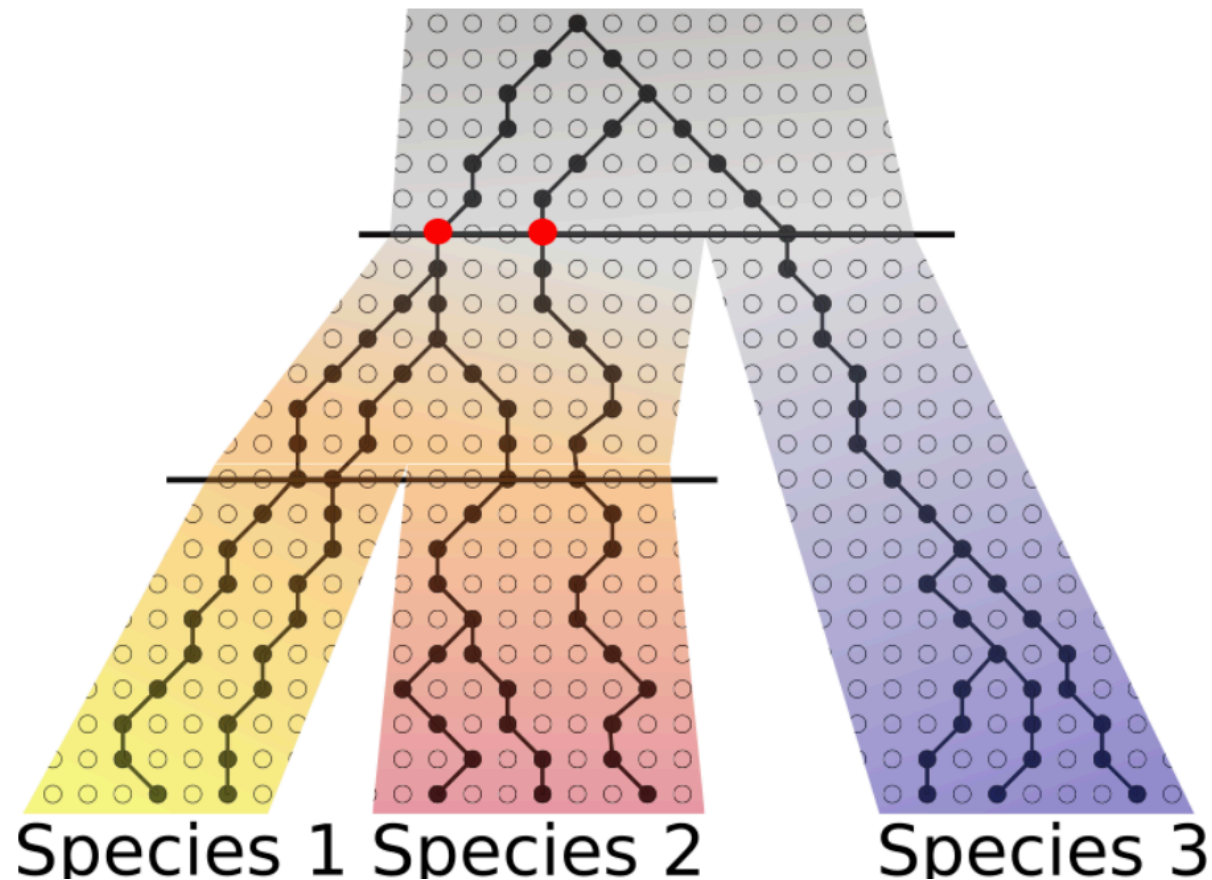
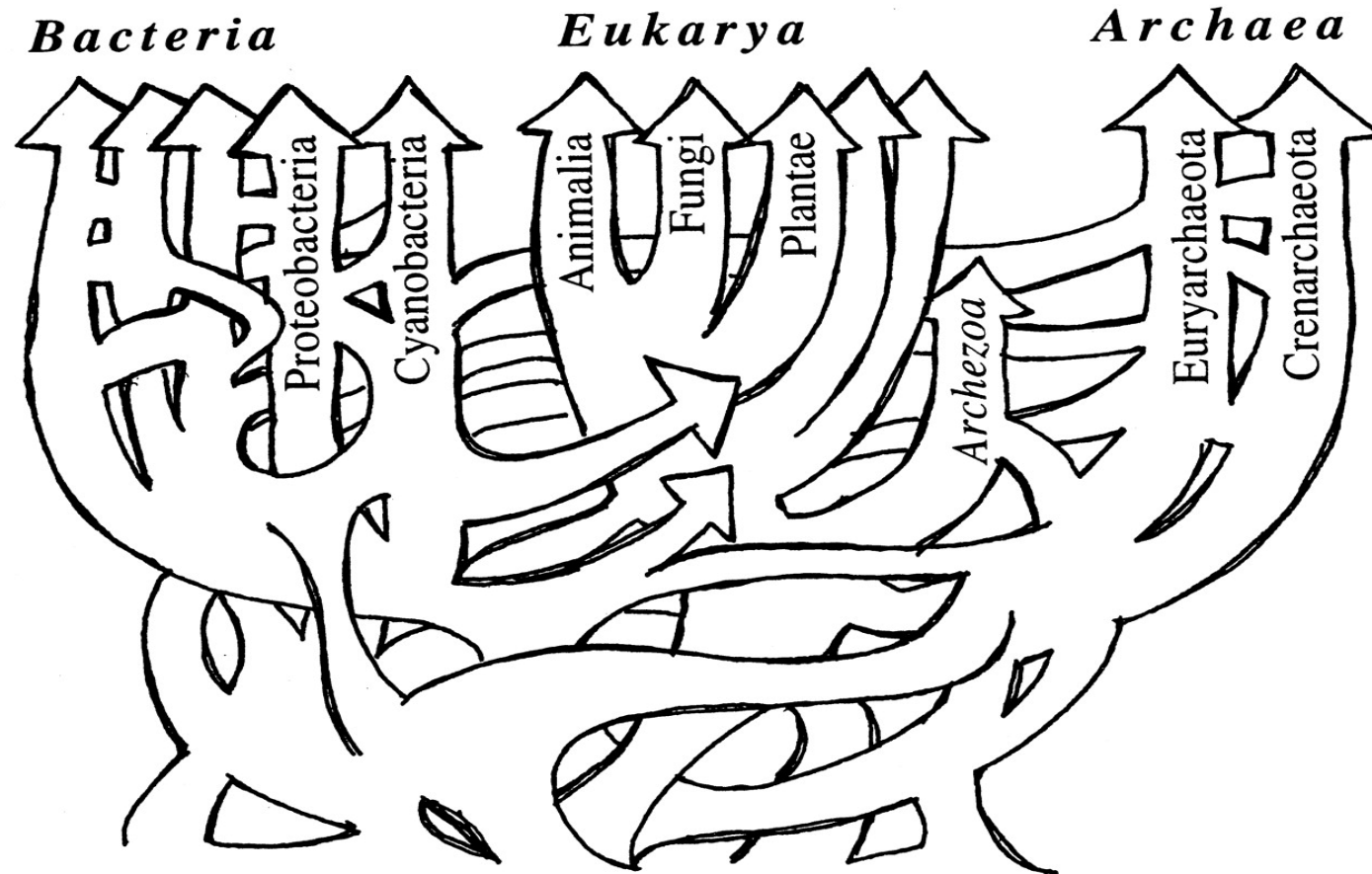


Image source: <https://www.beast2.org/>

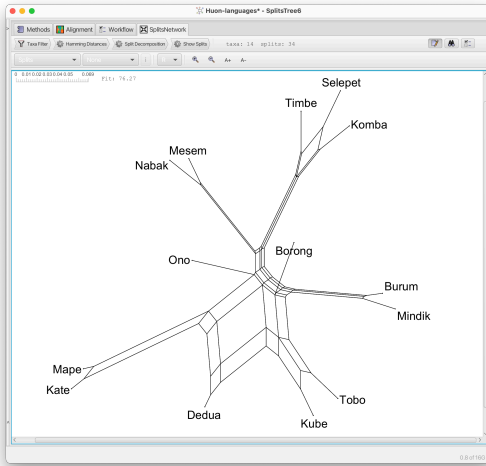
Phylogenetic networks



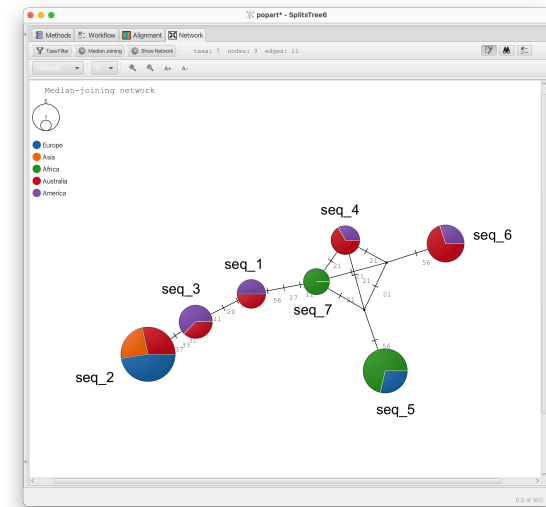
W. Ford Doolittle, Science 1999

Phylogenetic networks

► Unrooted phylogenetic networks

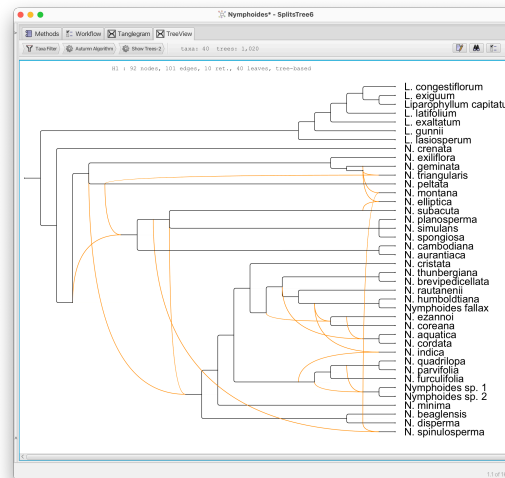


Splits networks
 Split decomposition method
 (Bandelt and Dress, 1992)



Haplotype networks
 Median-joining method
 (Bandelt, Forster & Röhl, 1999)

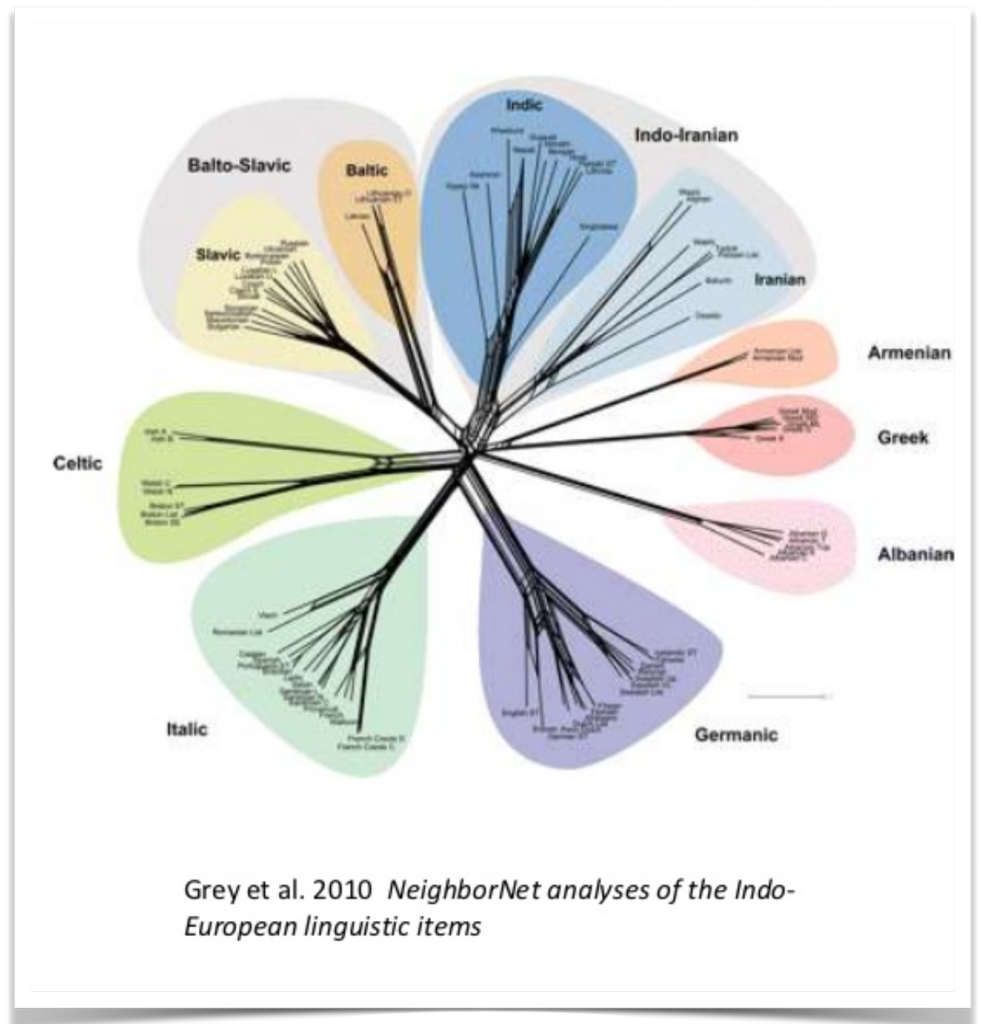
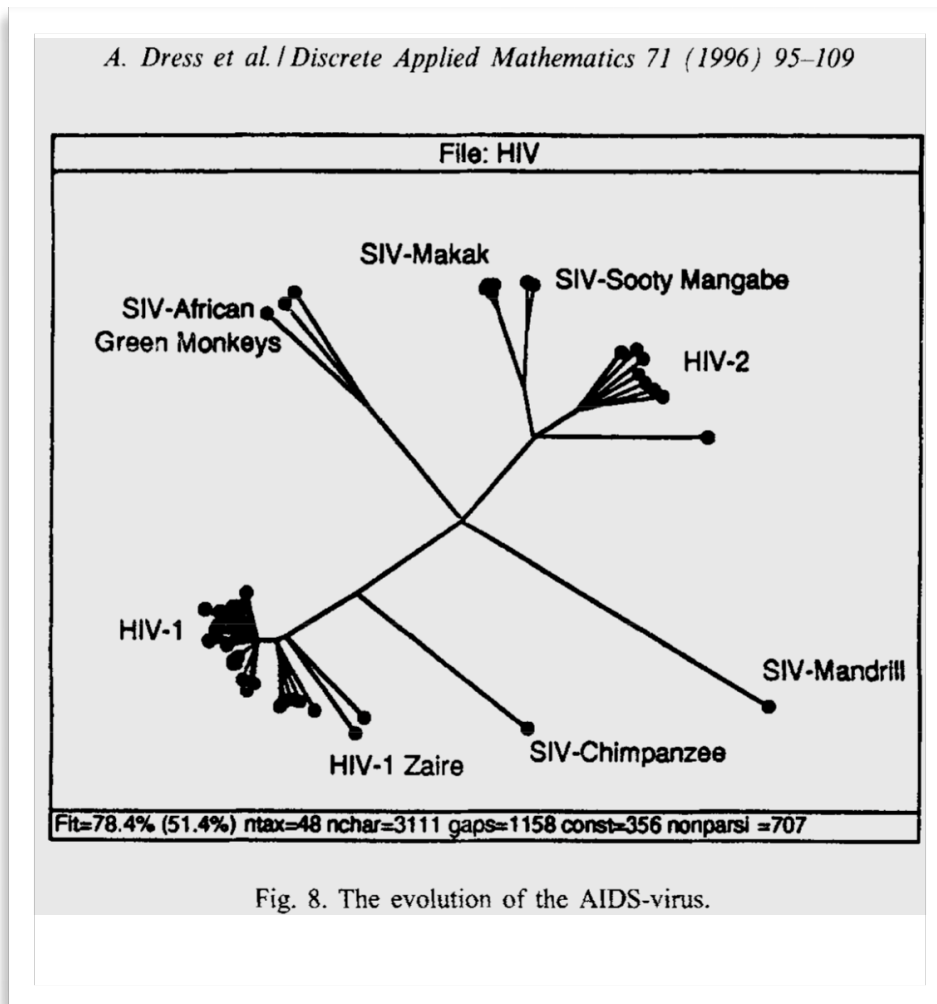
► Rooted phylogenetic networks



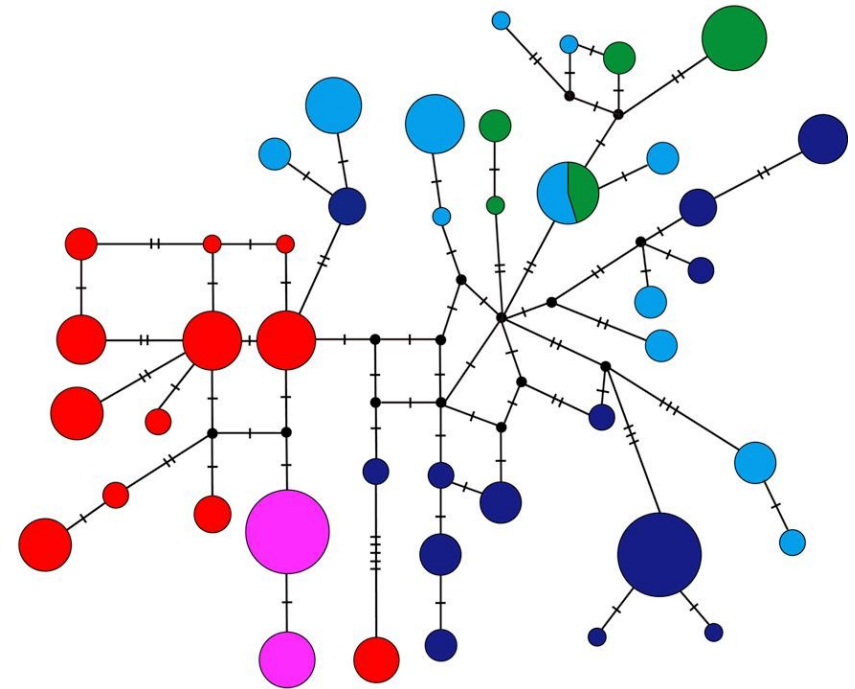
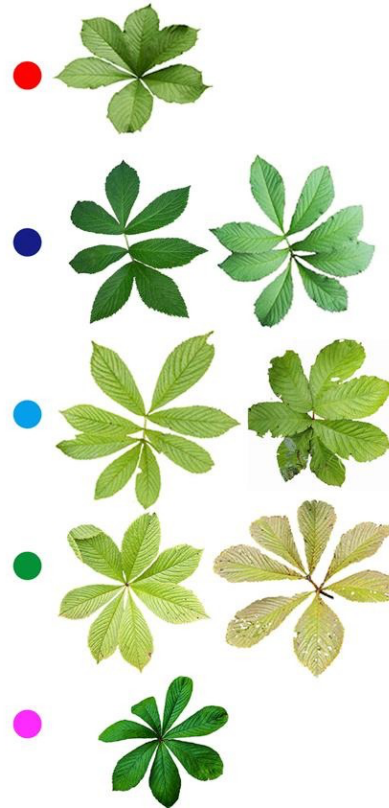
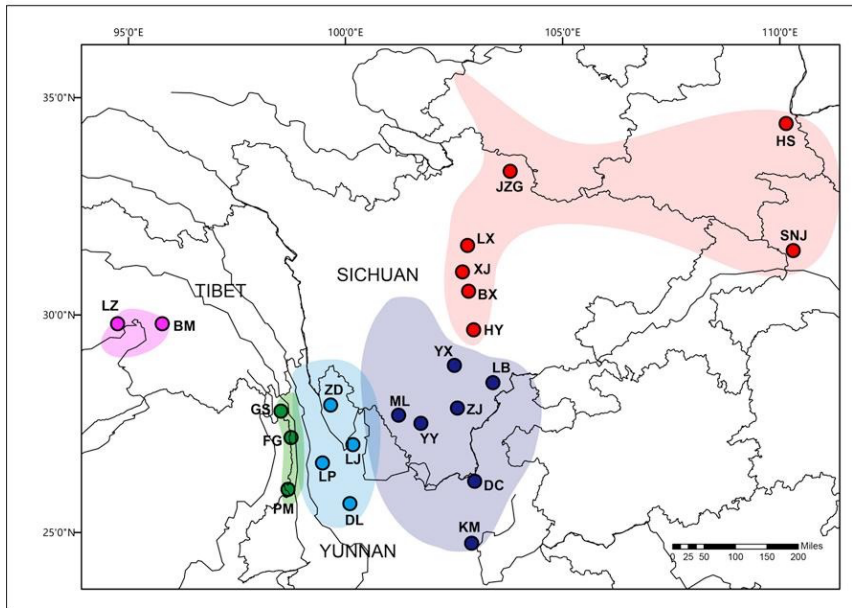
Rooted DAG

Unrooted phylogenetic networks

- Use phylogenetic networks when you can't assume that data fits on a tree



Haplotype networks



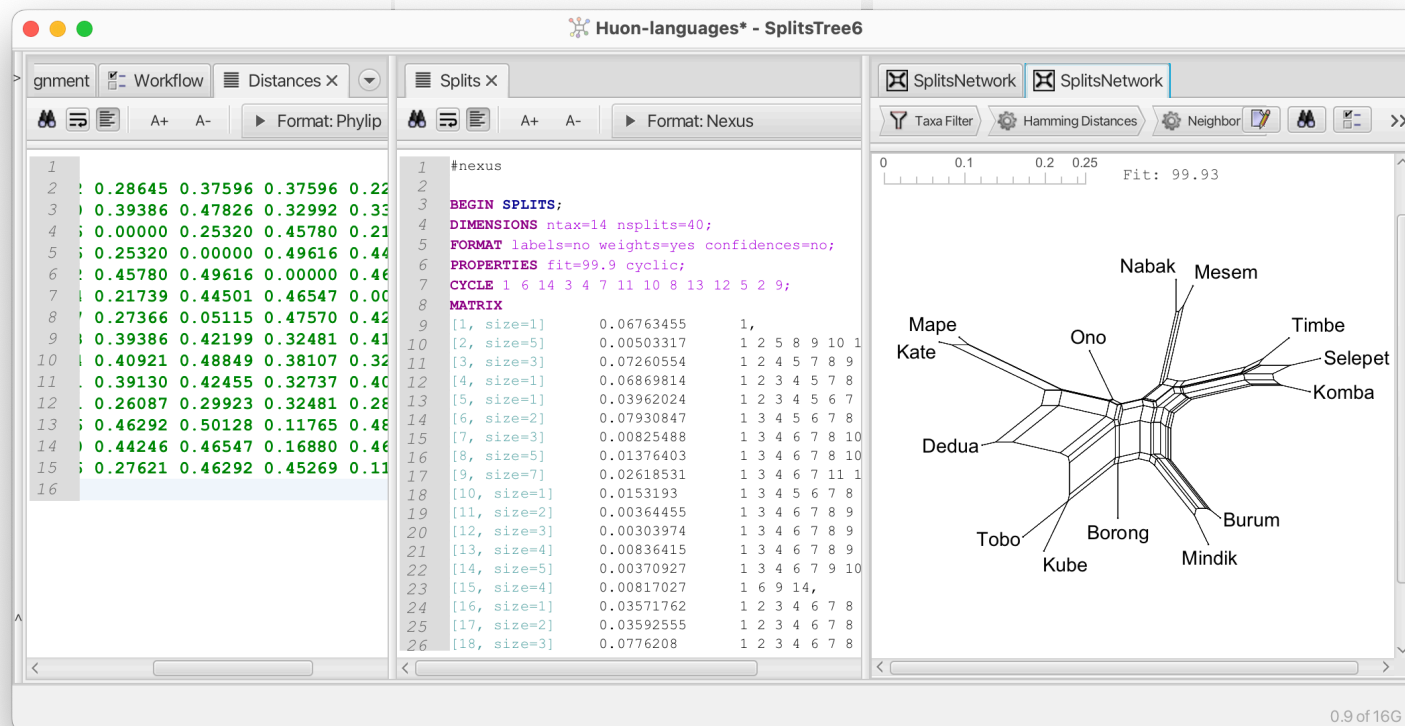
Haplotype network using
median-joining method

Ma, Sun & Sun., Historical introgression among the species of *Rodgersia* (Saxifragaceae) in mountainous forests of southwest China (2018)

Unrooted phylogenetic networks

Neighbor-Net (Bryant and Moulton, 2004)

- ▶ Input: distance matrix D on taxon set X
- ▶ Output: set of “circular” splits Σ on X
- ▶ Equal-angle algorithm computes network (Dress & H, 2004)

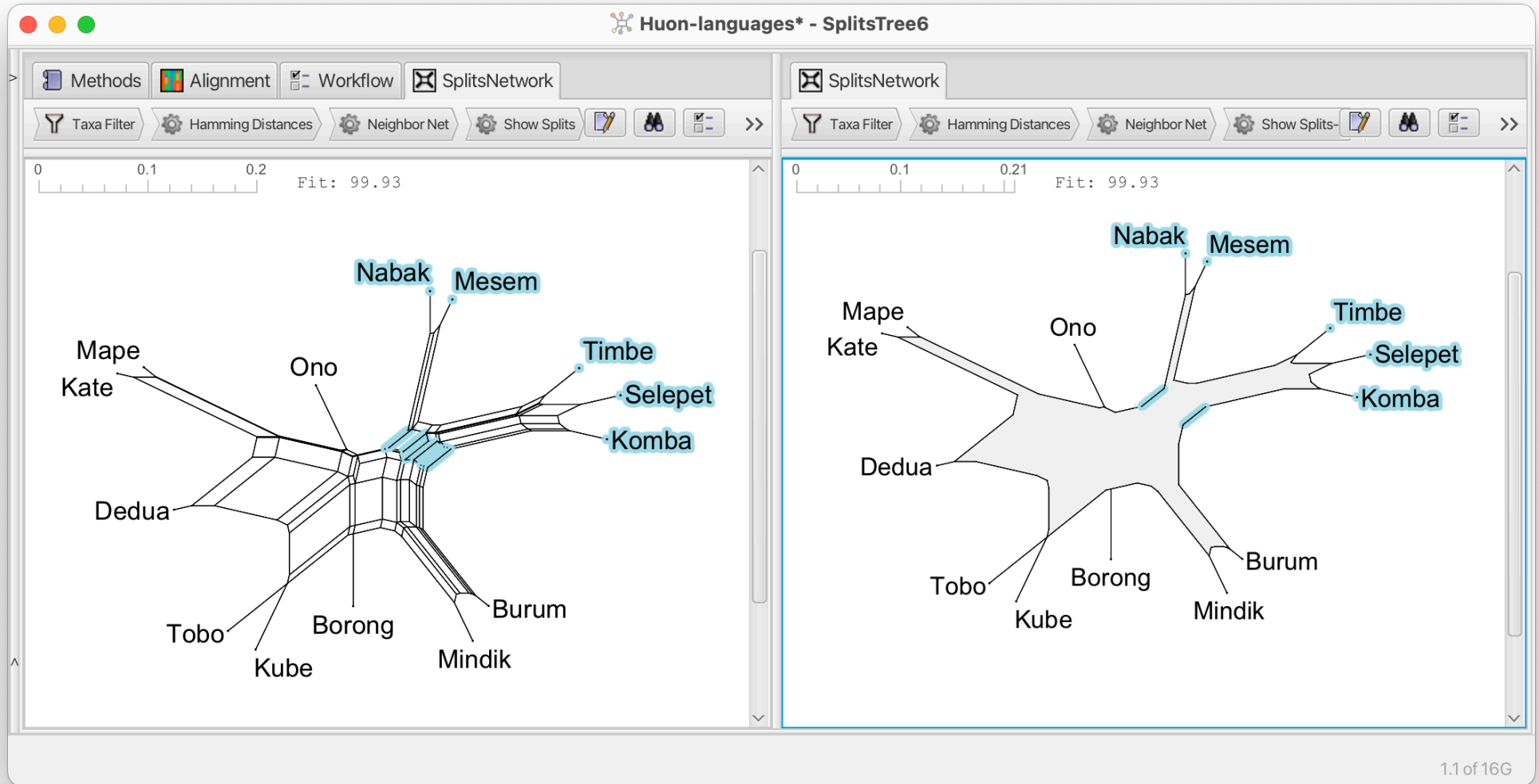


More work on Neighbor-Net

- ▶ Analysis and improvement (Levy & Pachter, 2011)
- ▶ Alternative using ILP (Lpnet, Gao & Grünewald, ?)
- ▶ Dave Bryant and H:
 - ▶ Speed-up weight optimization
 - ▶ Use “phylogenetic outline” to display
- ▶ Python implementation: Splits.py

Phylogenetic outlines

► Split network for circular splits:

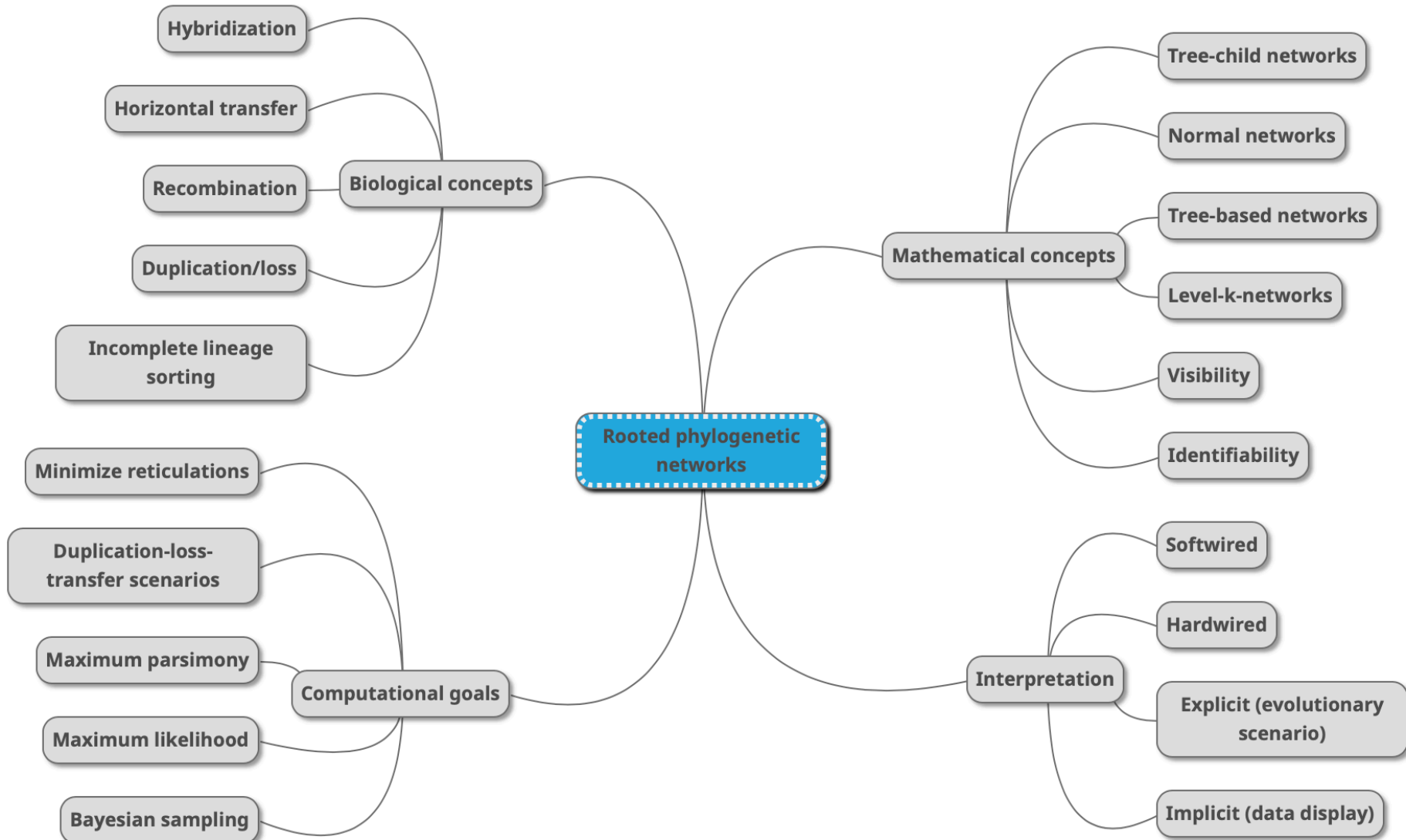


$$O(n^4)$$

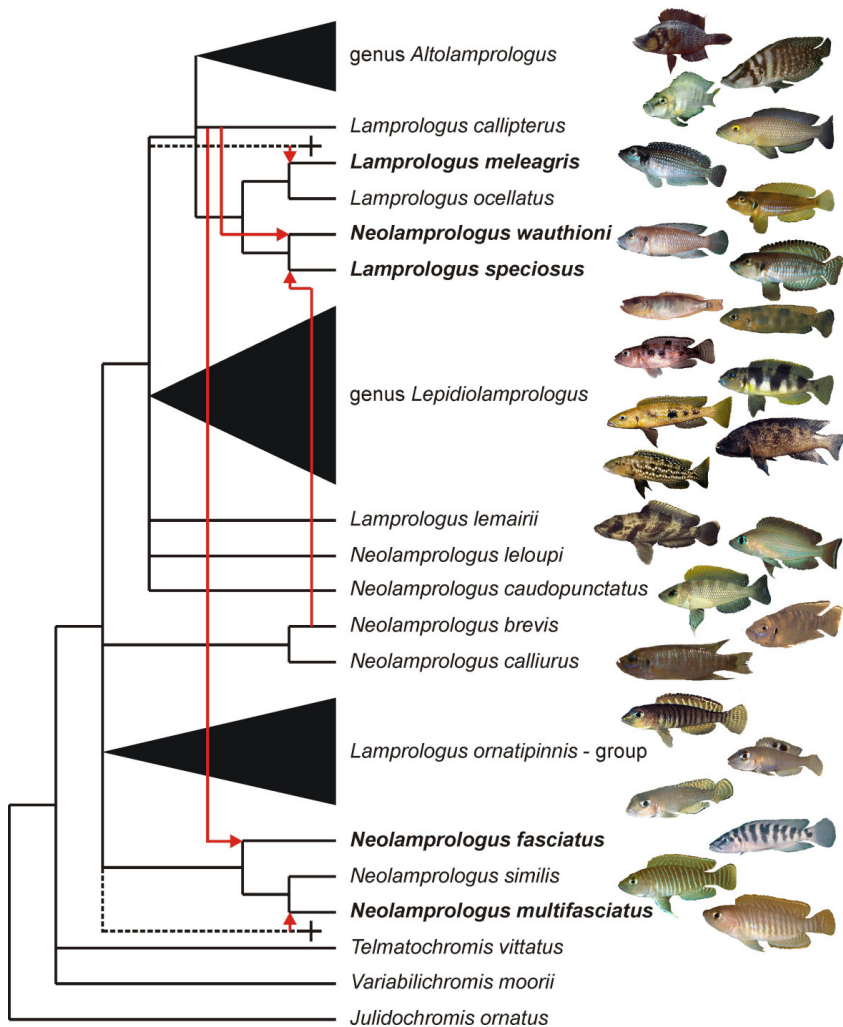
$$O(n^2) \text{ Bagci et al, 2021}$$

Rooted phylogenetic networks

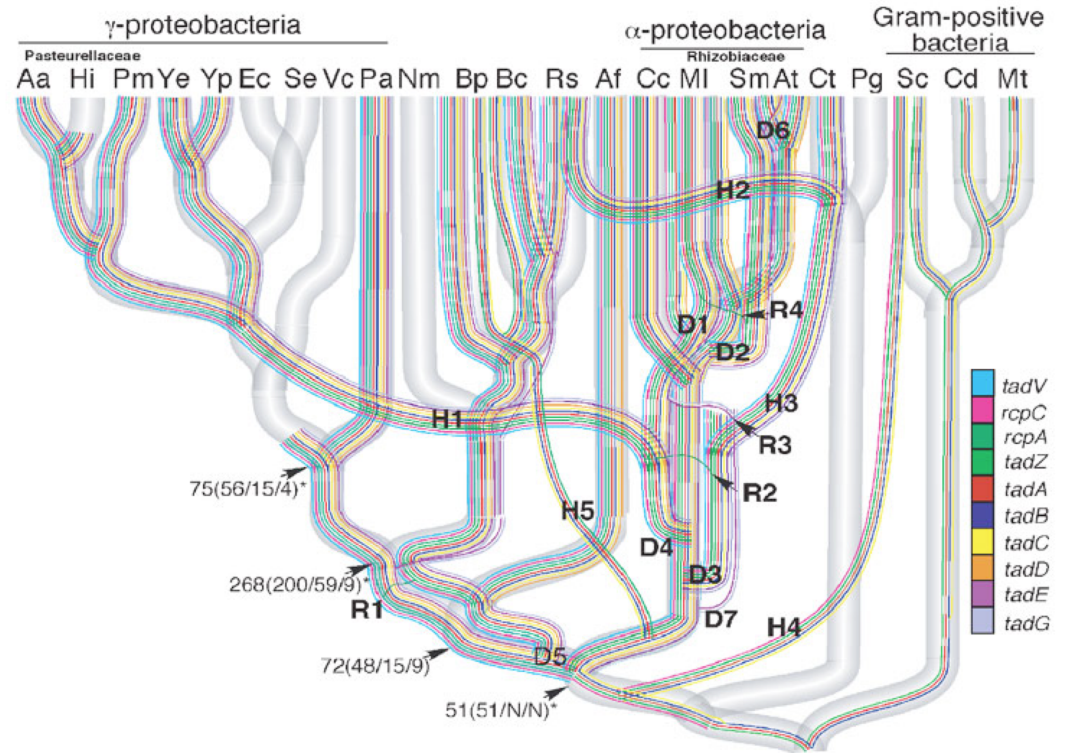
► Generalization of rooted phylogenetic trees



Rooted phylogenetic networks



Koblmuller et al, BMC Evol. Bio, (2007)



Planet et al, Nat. Gen. (2003)

Cluster networks (“hardwired”)

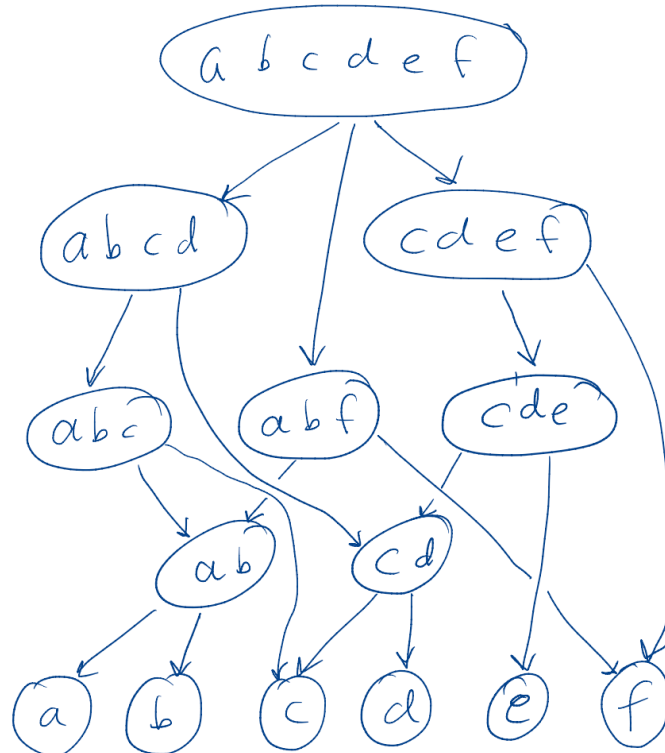
► Cluster network - based on Hasse diagram:

Clusters

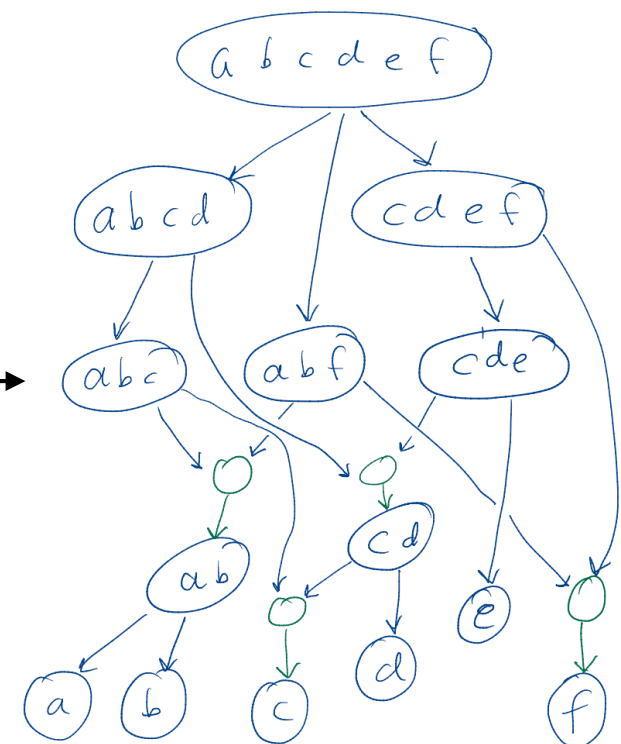
$\{a, b, c, d, e, f\}$,
 $\{a, b, c, d\}$,
 $\{c, d, e, f\}$,
 $\{a, b, c\}$,
 $\{a, b, f\}$,
 $\{c, d, e\}$,
 $\{a, b\}$,
 $\{c, d\}$,
 $\{a\}, \dots, \{f\}$.



Hasse diagram



Cluster network



Polynomial time, size

(Splits: network can have exponential size)

Hybridization networks (“softwired”)

► Hybridization network:

- Input: Rooted phylogenetic trees T_1 and T_2
- Output: Rooted network N displaying T_1 and T_2 ,
using minimum number of reticulations,
hybridization number h

Combinatorial problem

Bordewich and Semple (2007); van Iersel and Linz (2013); van Iersel et al. (2014) and more.

Hybridization networks

Autumn algorithm (H. and Linz, 2018):

- ▶ Input: two rooted trees T_1 and T_2
 - ▶ Multifurcations ok
 - ▶ Missing taxa ok
- ▶ Output:
 - ▶ hybridization number h and
 - ▶ all minimum networks

Hybridization networks

Syst. Biol. 58(6):612–628, 2009
© The Author(s) 2009. Published by Oxford University Press, on behalf of the Society of Systematic Biologists. All rights reserved.
For Permissions, please email: journals.permissions@oxfordjournals.org
DOI:10.1093/sysbio/syp068
Advance Access publication on October 21, 2009

Reticulation, Data Combination, and Inferring Evolutionary History: An Example from Danthoioideae (Poaceae)

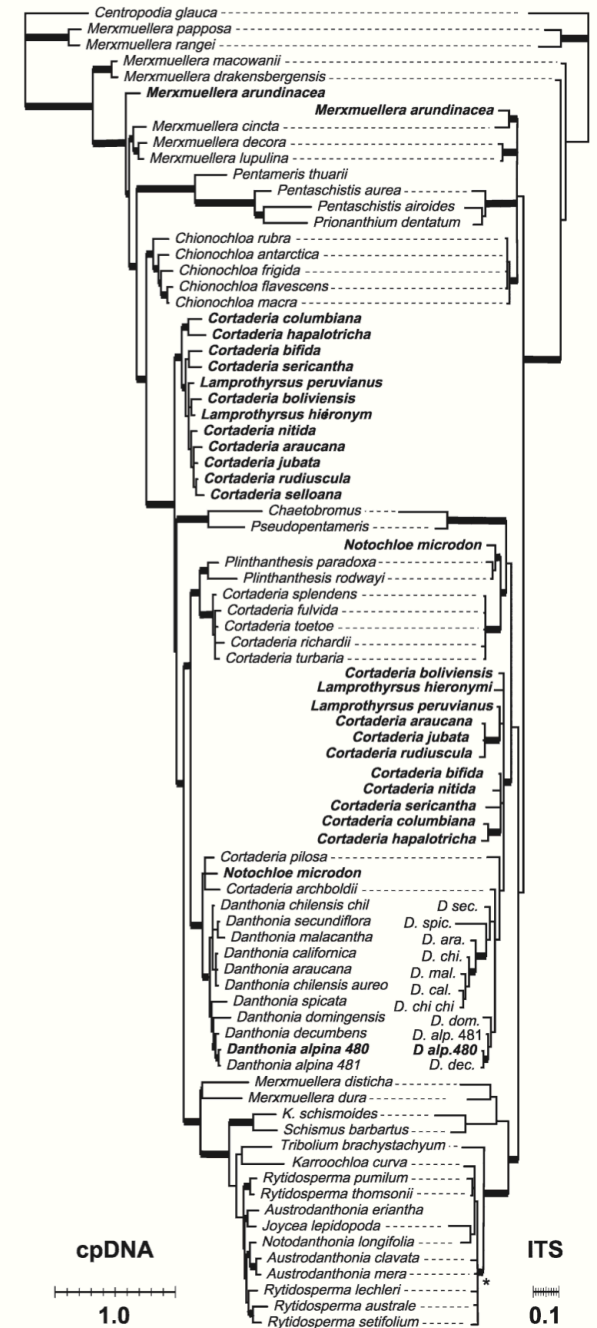
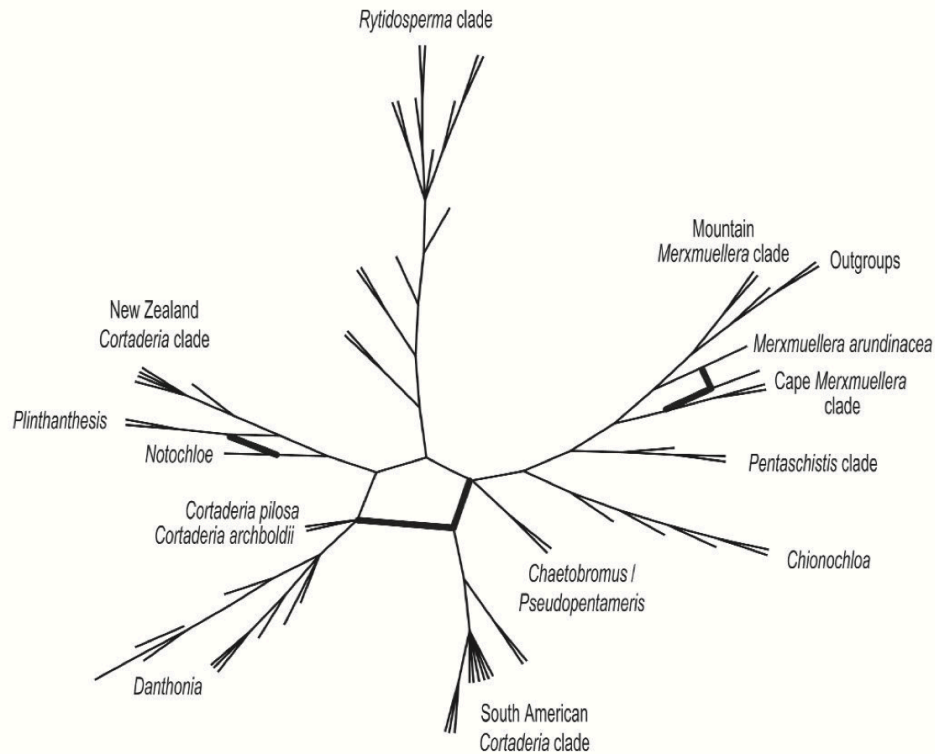
MICHAEL D. PIRIE^{1,2,*}, AELYS M. HUMPHREYS¹, NIGEL P. BARKER³, AND H. PETER LINDER¹

¹Institute for Systematic Botany, University of Zurich, CH-8008 Zurich, Switzerland;

²Present address: Department of Biochemistry, University of Stellenbosch, Private Bag X1, Matieland 7602, Stellenbosch, South Africa; and

³Molecular Ecology and Systematics Group, Department of Botany, Rhodes University, Grahamstown 6140, South Africa;

*Correspondence to be sent to: Department of Biochemistry, University of Stellenbosch, Private Bag X1, Matieland 7602, Stellenbosch, South Africa;
E-mail: mpirie@sun.ac.za.



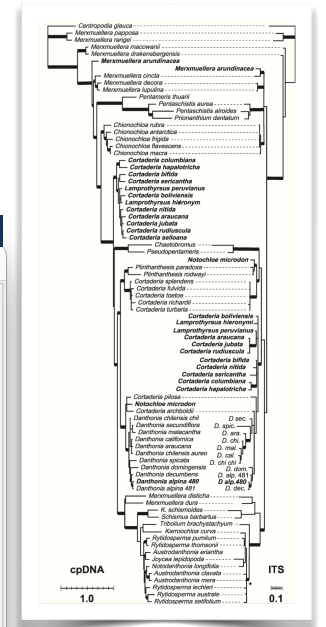
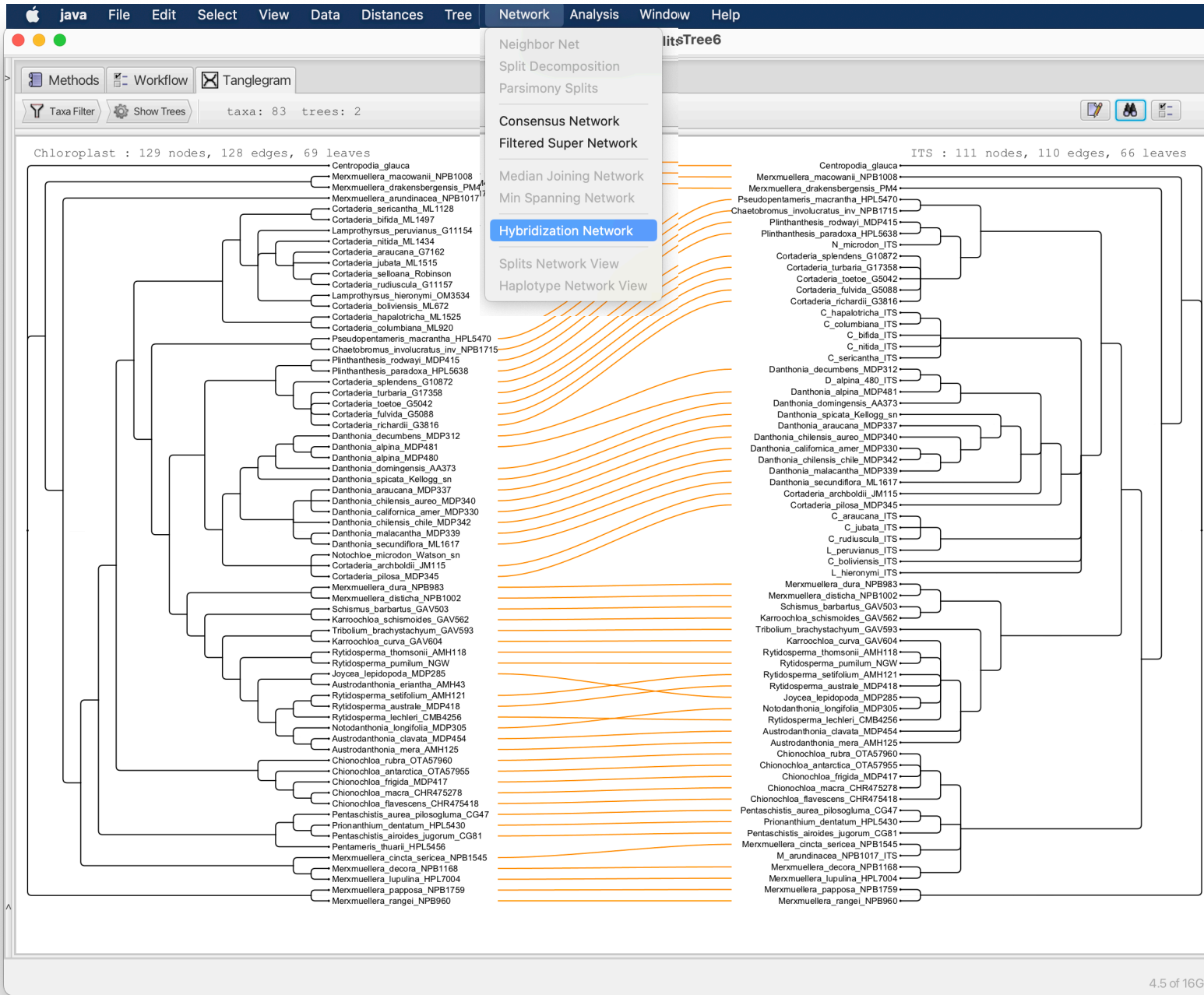
cpDNA

1.0

ITS

0.1

Autumn algorithm

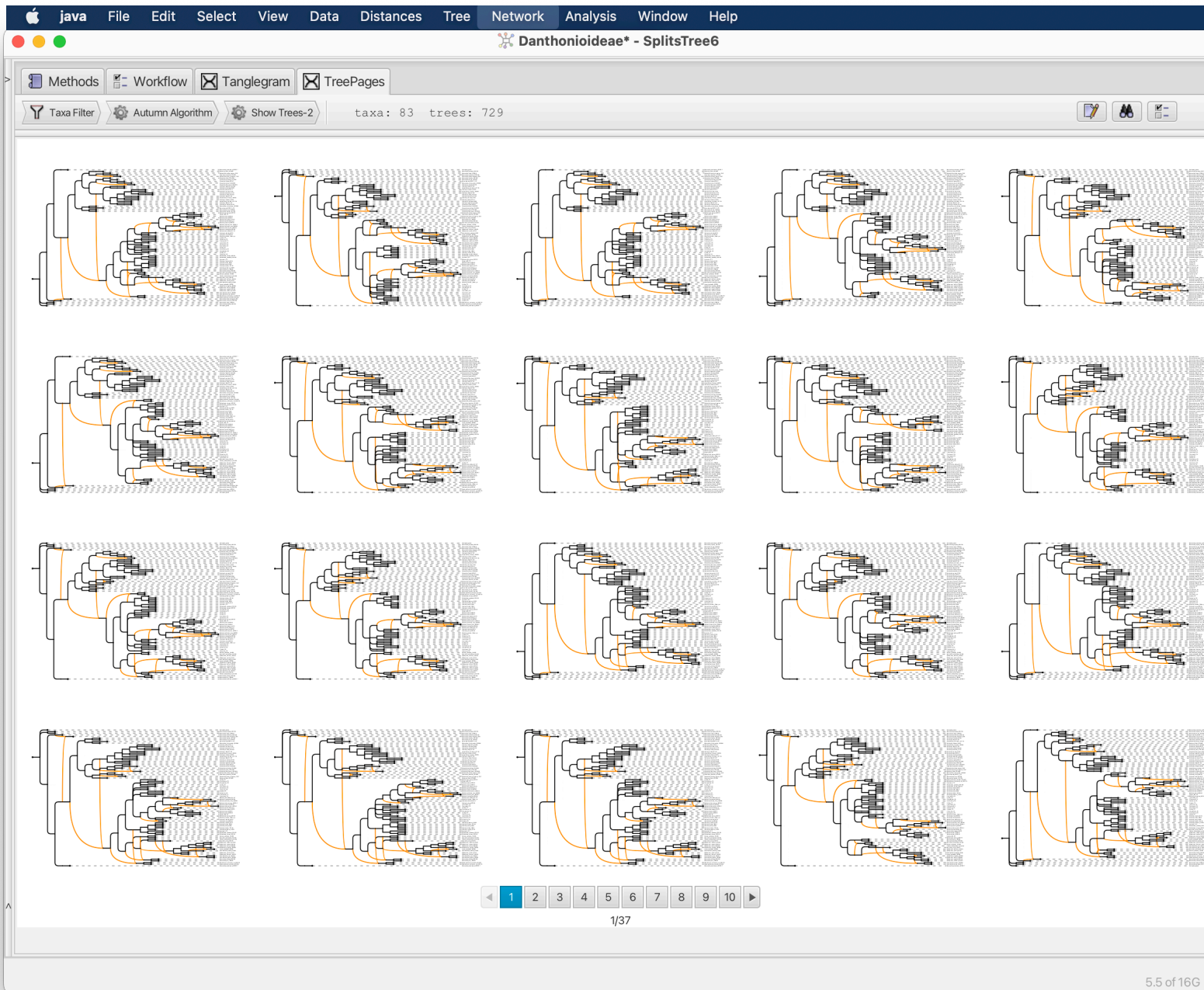


Autumn algorithm



2-3 seconds

Autumn algorithm



$h=13$

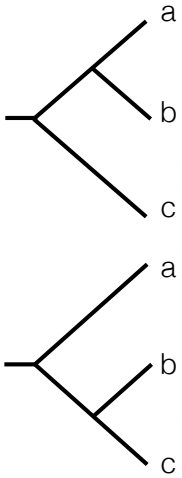
Networks:

729

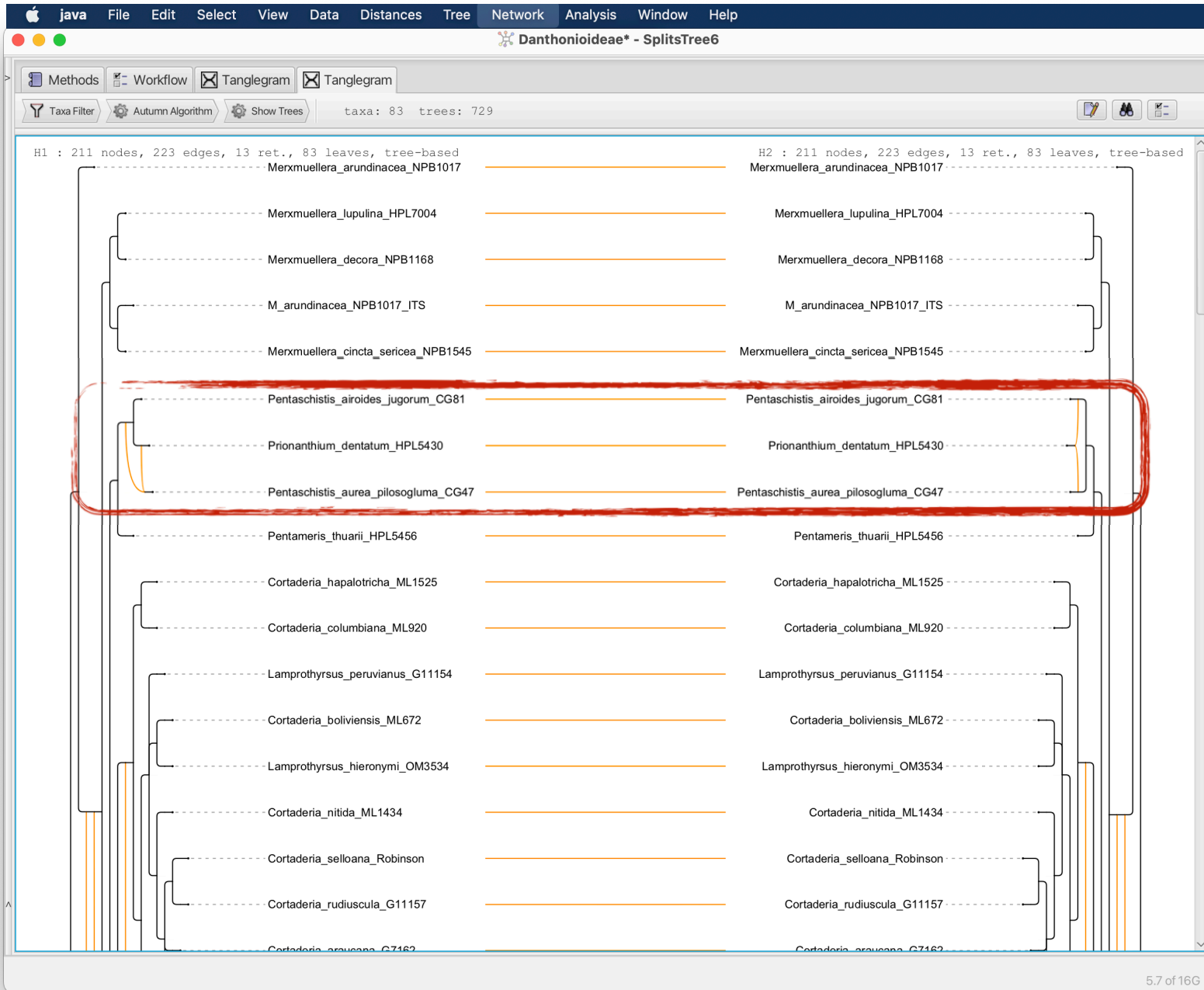
$=3^6$

?

Autumn algorithm



3
ways



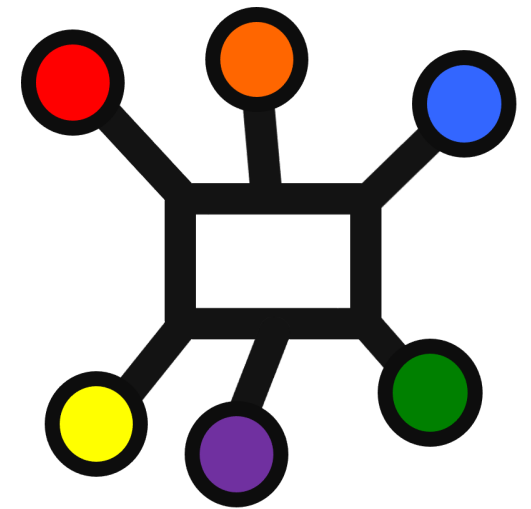
Networks:

729

= 3⁶

?

SplitsTree6



New interactive program for:

- ▶ Unrooted phylogenetic networks
- ▶ Phylogenetic trees
- ▶ Rooted phylogenetic networks

- ▶ JavaFX, based on SplitsTree4 (2006) and Dendroscope3 (2012)

Example: New Guinea languages

RESEARCH ARTICLE

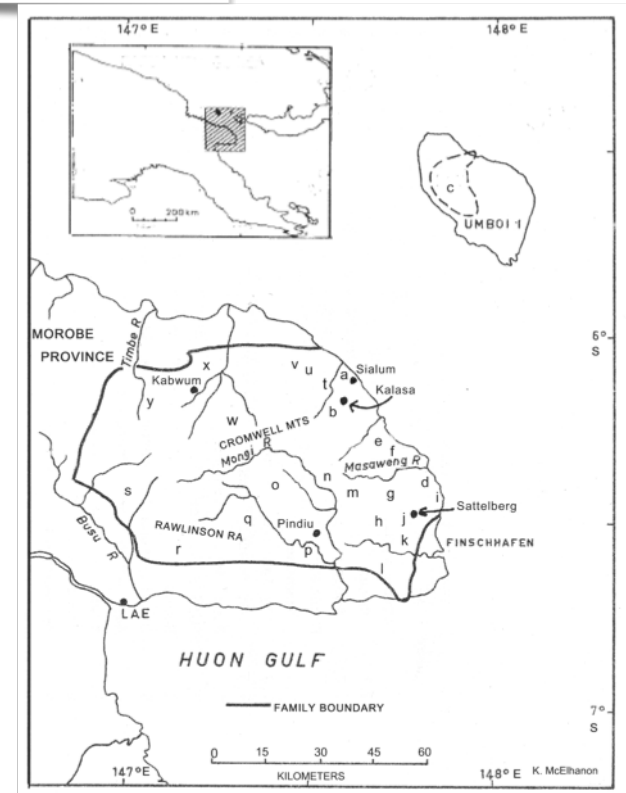
PLOS ONE | DOI:10.1371/journal.pone.0141563 October 27, 2015

TransNewGuinea.org: An Online Database of New Guinea Languages

Simon J. Greenhill^{1,2,3*}

1 ARC Centre of Excellence for the Dynamics of Language, Australian National University, Canberra, Australia, 2 ANU College of Asia and the Pacific, Australian National University, Canberra, Australia, 3 Max Planck Institute for the Science of Human History, Jena, Germany

- ▶ Example: 14 languages of Huon peninsula
- ▶ 391 cognate word characters

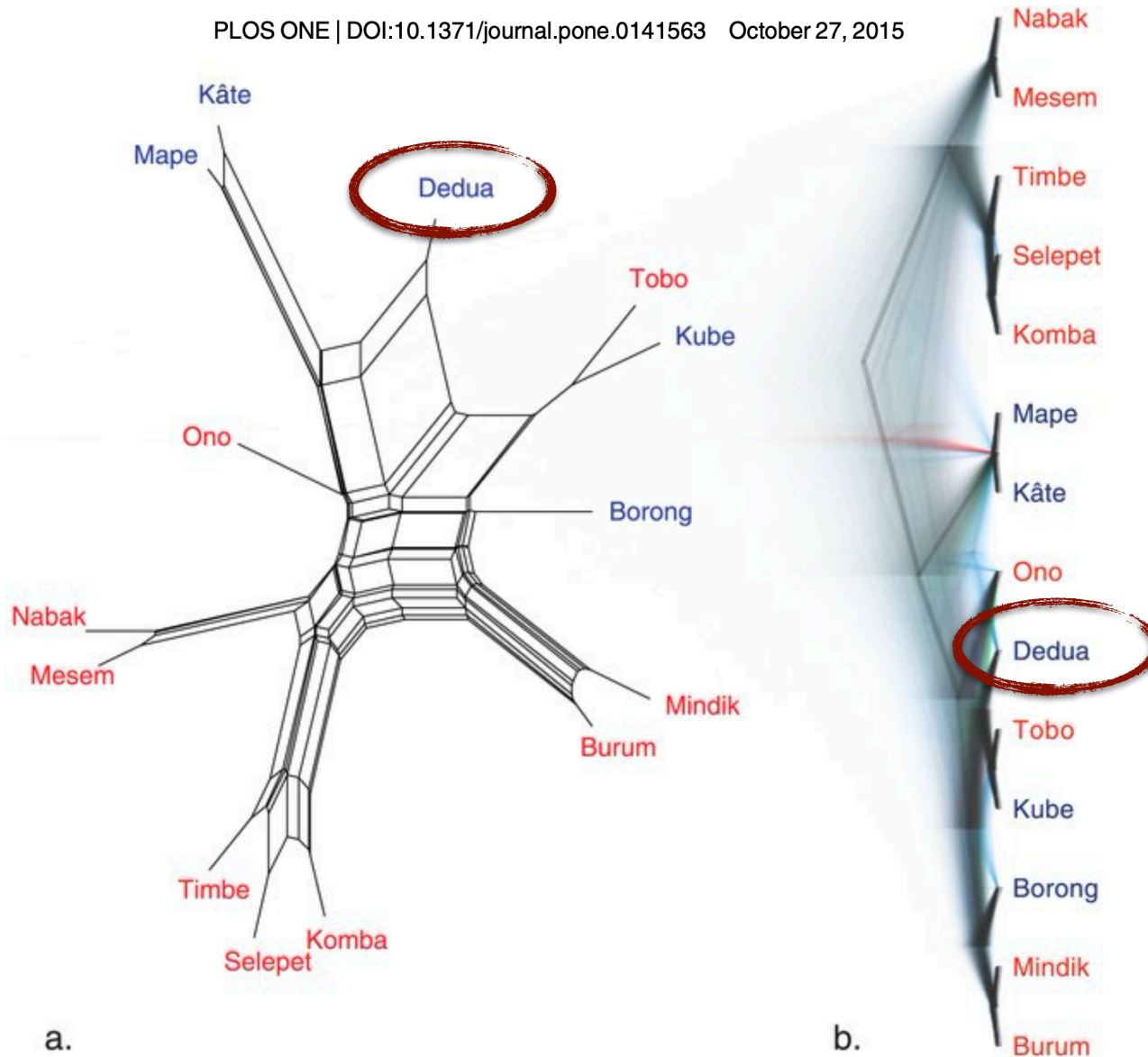


Edgar Suter, PhD thesis 2018

New Guinea languages

PLOS ONE | DOI:10.1371/journal.pone.0141563 October 27, 2015

SplitsTree4



Beast2
 Bouckaert et al, 2014

DensiTree
 Bouckaert, 2010

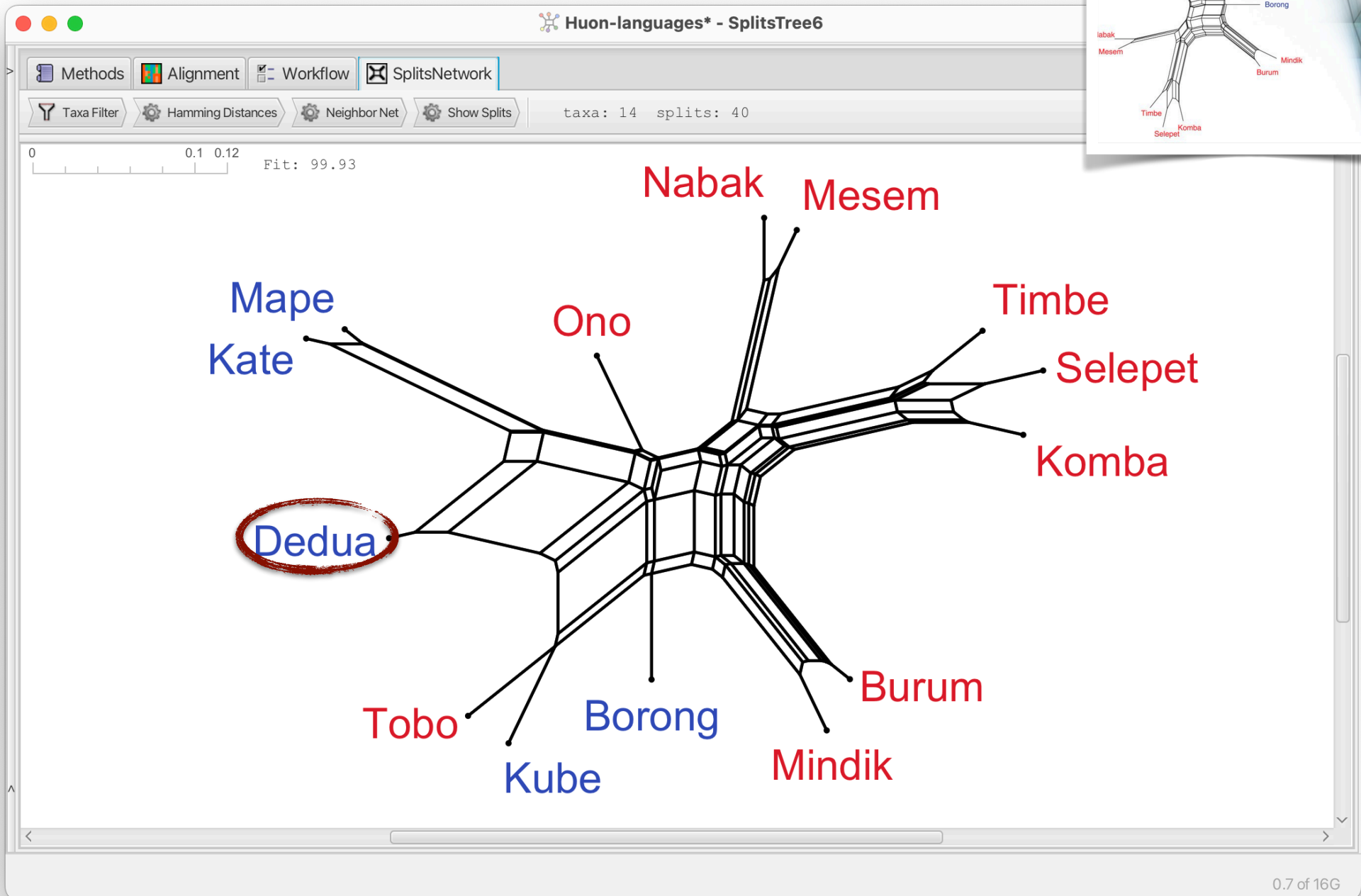
Colors:
 Blue: Western
 Red: Eastern

a.

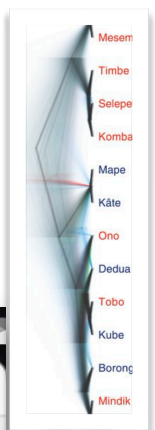
b.

Greenhill, 2015

Splits network



Bayesian analysis



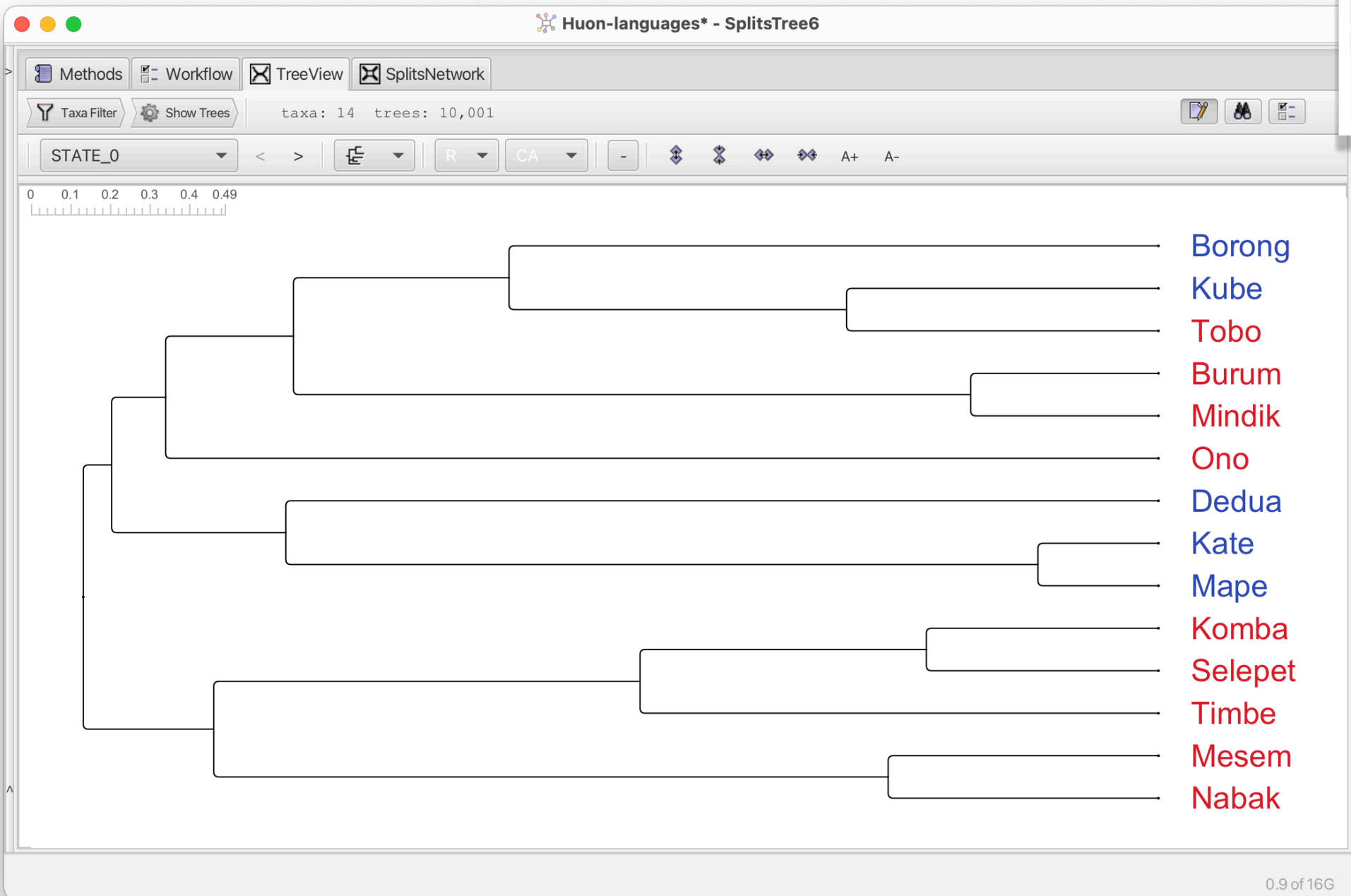
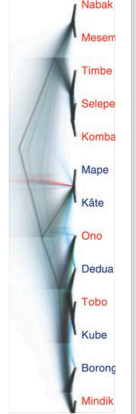
Operator	Tuning Parameter
ScaleOperator(YuleBirthRateScaler.t:Huon-languages)	0.2913
ScaleOperator(YuleModelTreeScaler.t:Huon-languages)	0.7468
ScaleOperator(YuleModelTreeRootScaler.t:Huon-languages)	0.7287
Uniform(YuleModelUniformOperator.t:Huon-languages)	-
SubtreeSlide(YuleModelSubtreeSlide.t:Huon-languages)	0.5693
Exchange(YuleModelNarrow.t:Huon-languages)	-
Exchange(YuleModelWide.t:Huon-languages)	-
WilsonBalding(YuleModelWilsonBalding.t:Huon-languages)	-

Tuning: The value of the operator's tuning parameter, or '-' if t
 #accept: The total number of times a proposal by this operator has
 #reject: The total number of times a proposal by this operator has
 Pr(m): The probability this operator is chosen in a step of the
 Pr(accEnd likelihood: -2427.4555013835156
 |m): The acceptance probability (#accept as a fraction of the total pr
 Total calculation time: 1834.743 seconds

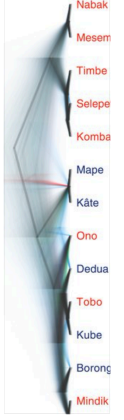
10,001 trees

Bouckaert et al, Beast 2, 2014

Tree view



Tree pages



Methods Workflow TreePages SplitsNetwork

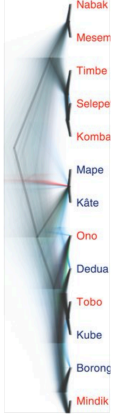
Taxa Filter Settings taxa: 14 trees: 10,001

TreeView
TreePages
Tanglegram
DensiTree
Text

7/501

0.9 of 16G

Tree pages



Huon-languages* - SplitsTree6

Methods Workflow TreePages SplitsNetwork

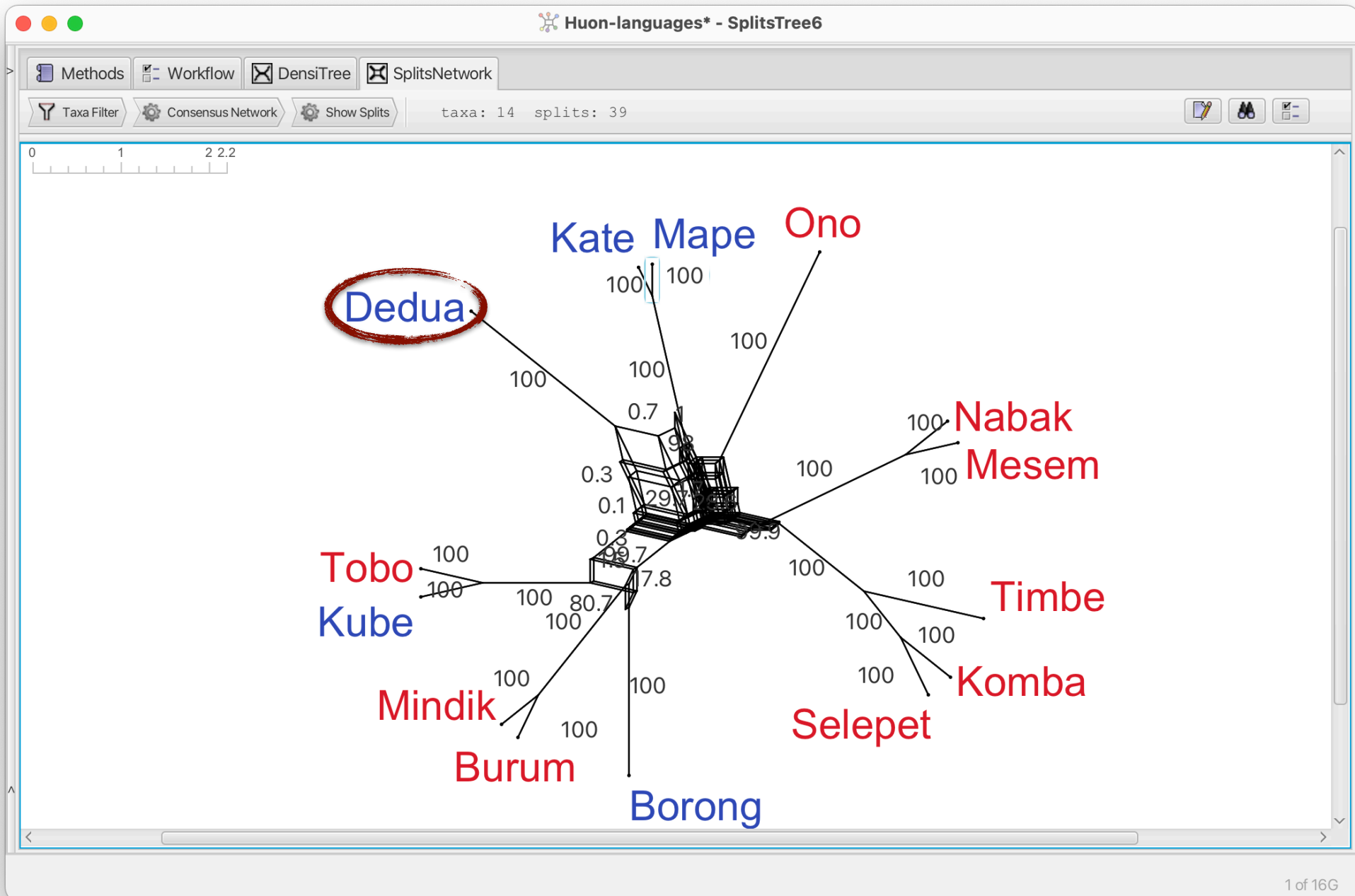
Taxa Filter Show Trees taxa: 14 trees: 10,001

1 2 3 4 5 6 7 8 9 10

1/501

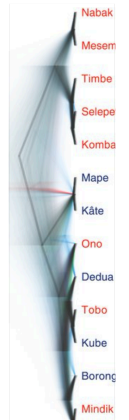
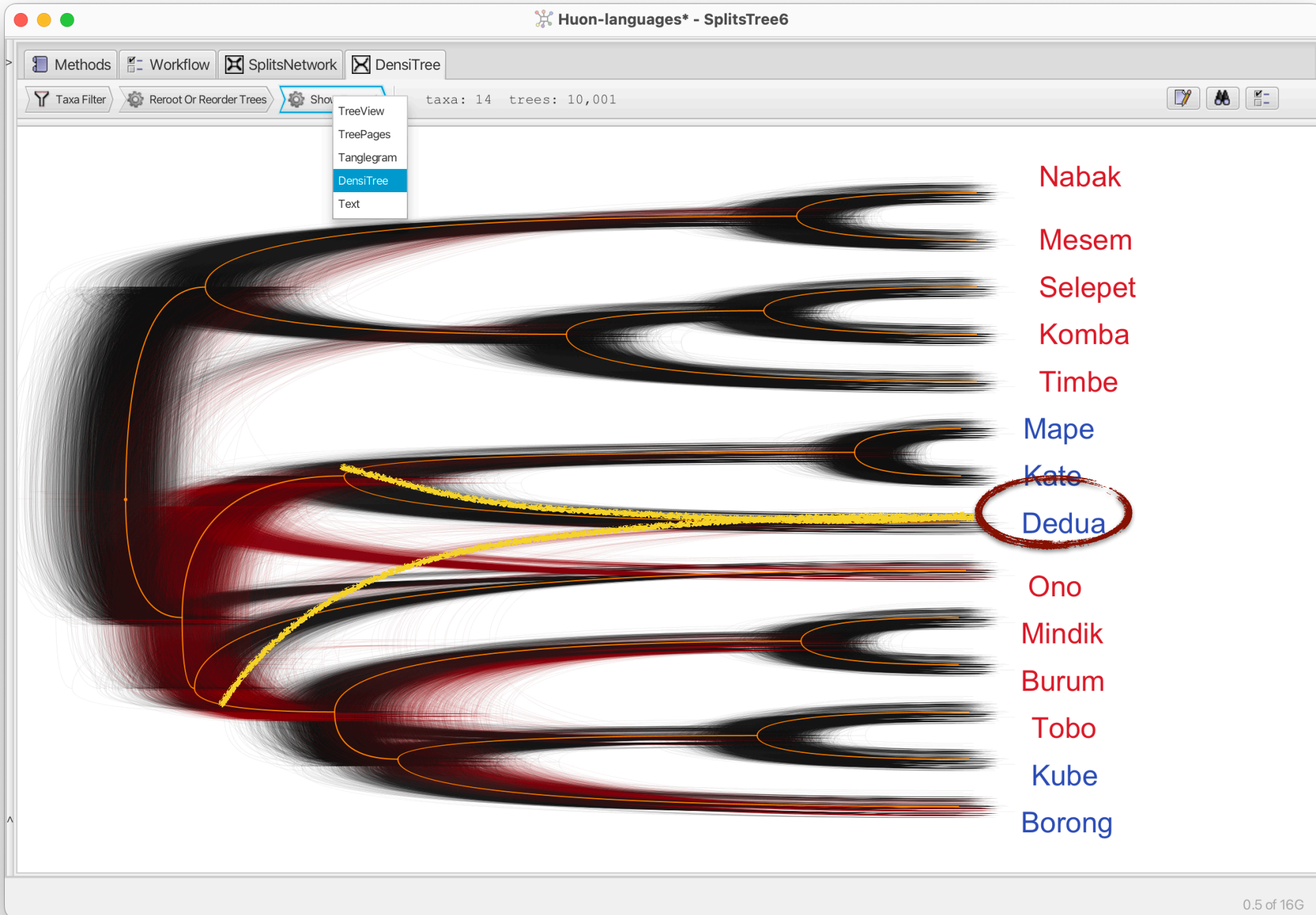
1.2 of 16G

Consensus network

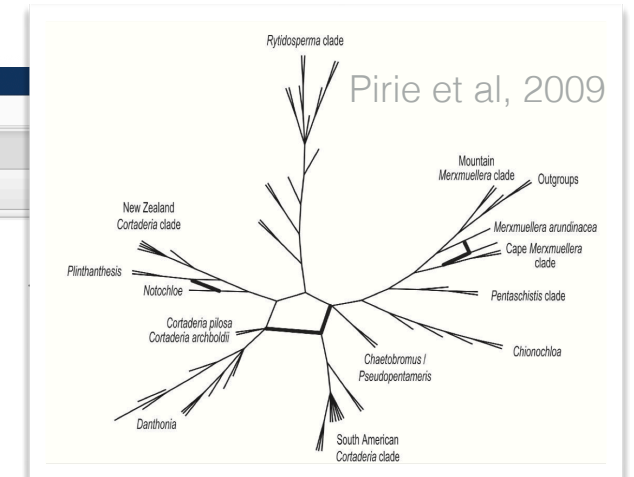
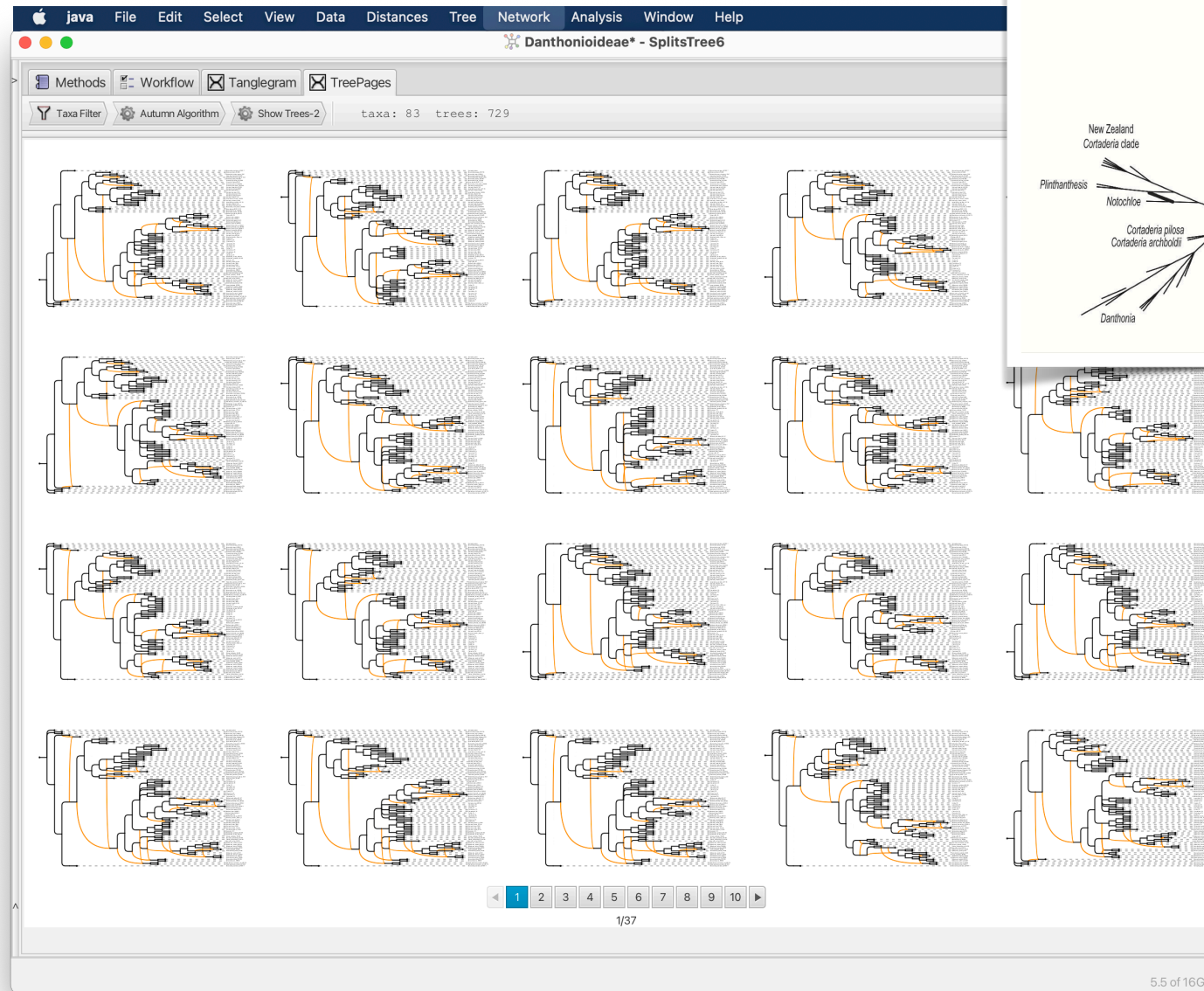


Holland and Moulton, 2003

Densi-tree



Many rooted networks



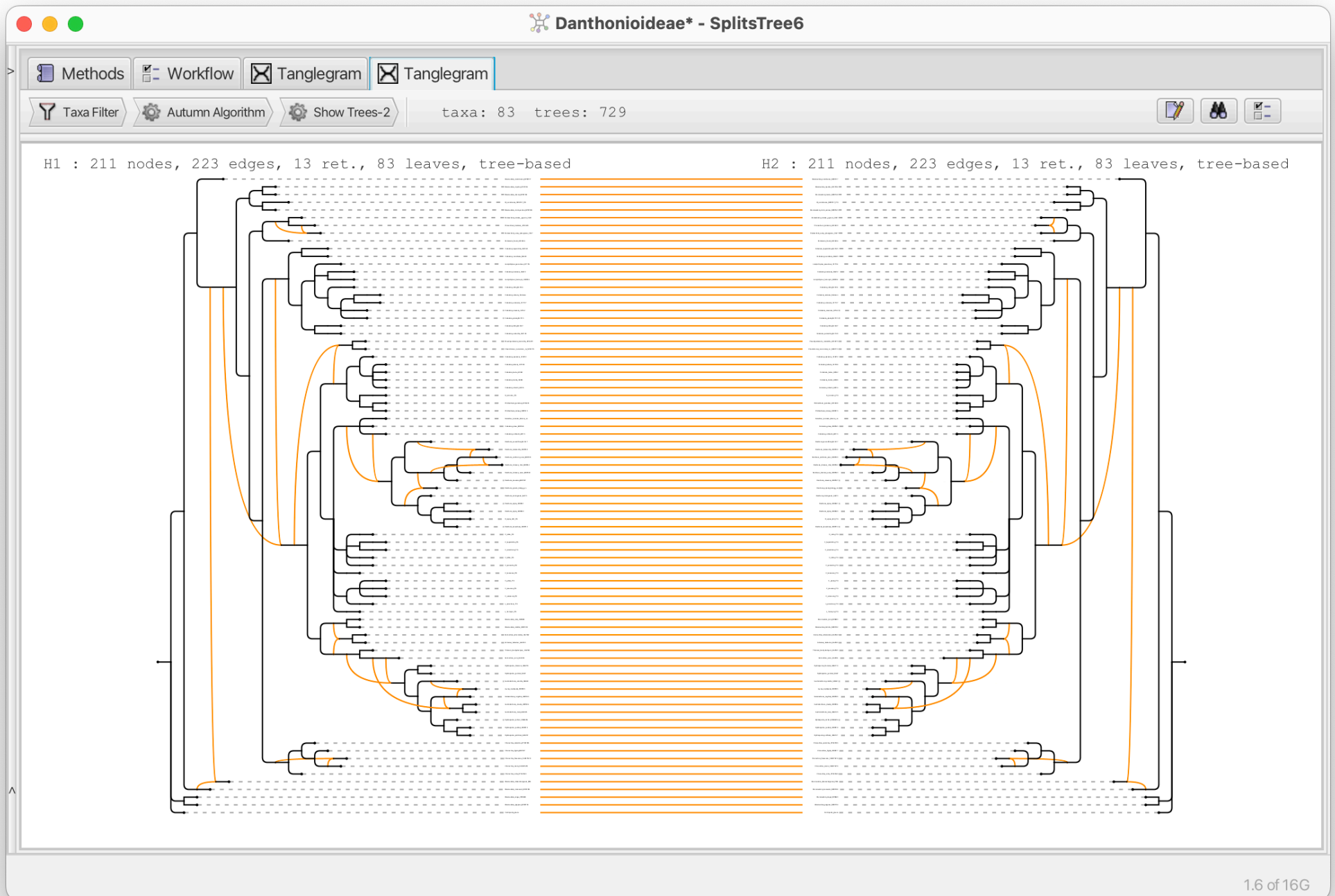
$h=13$

Networks:
729

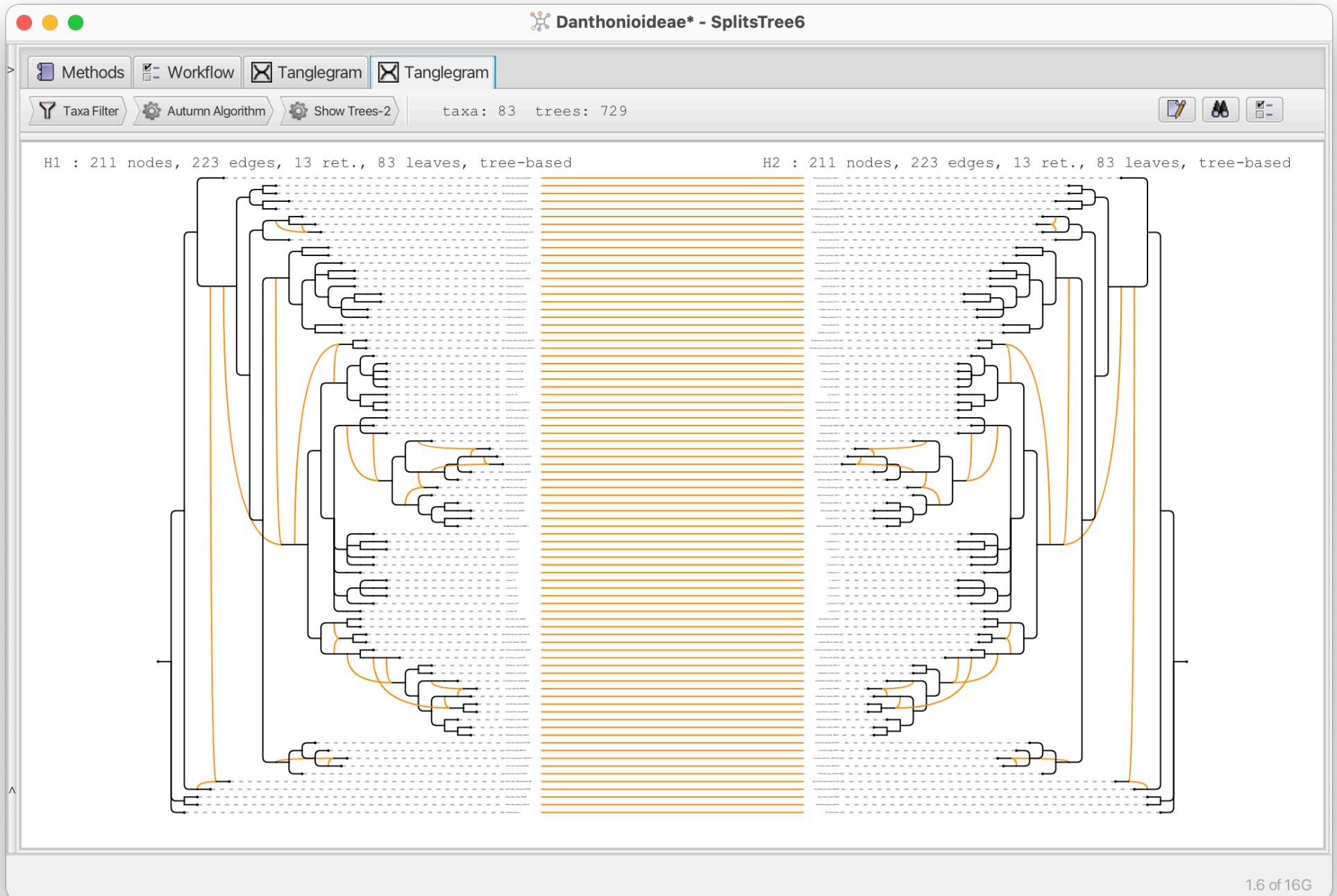
► How to summarize?

► Densi-network?

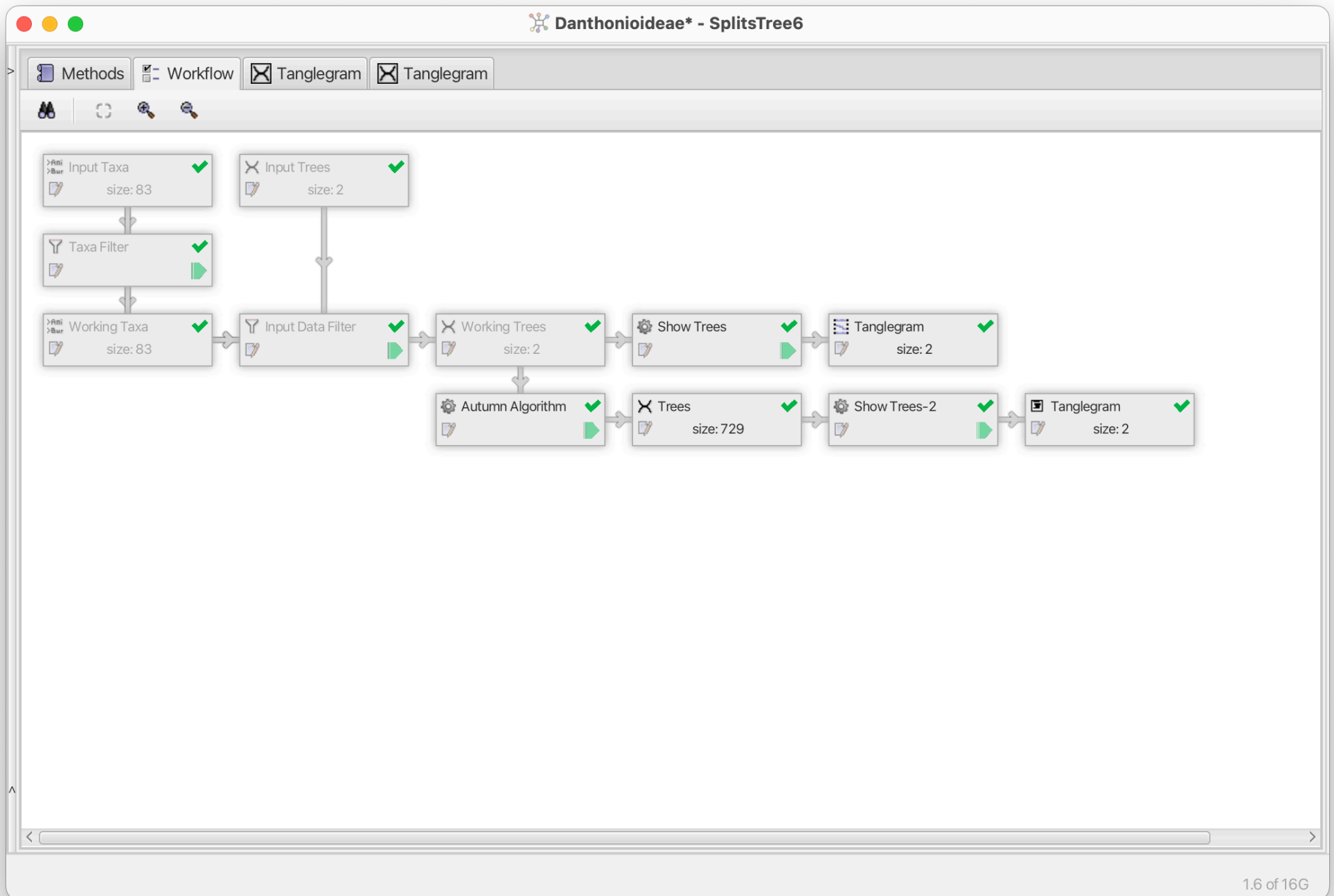
SplitsTree6 uses a workflow



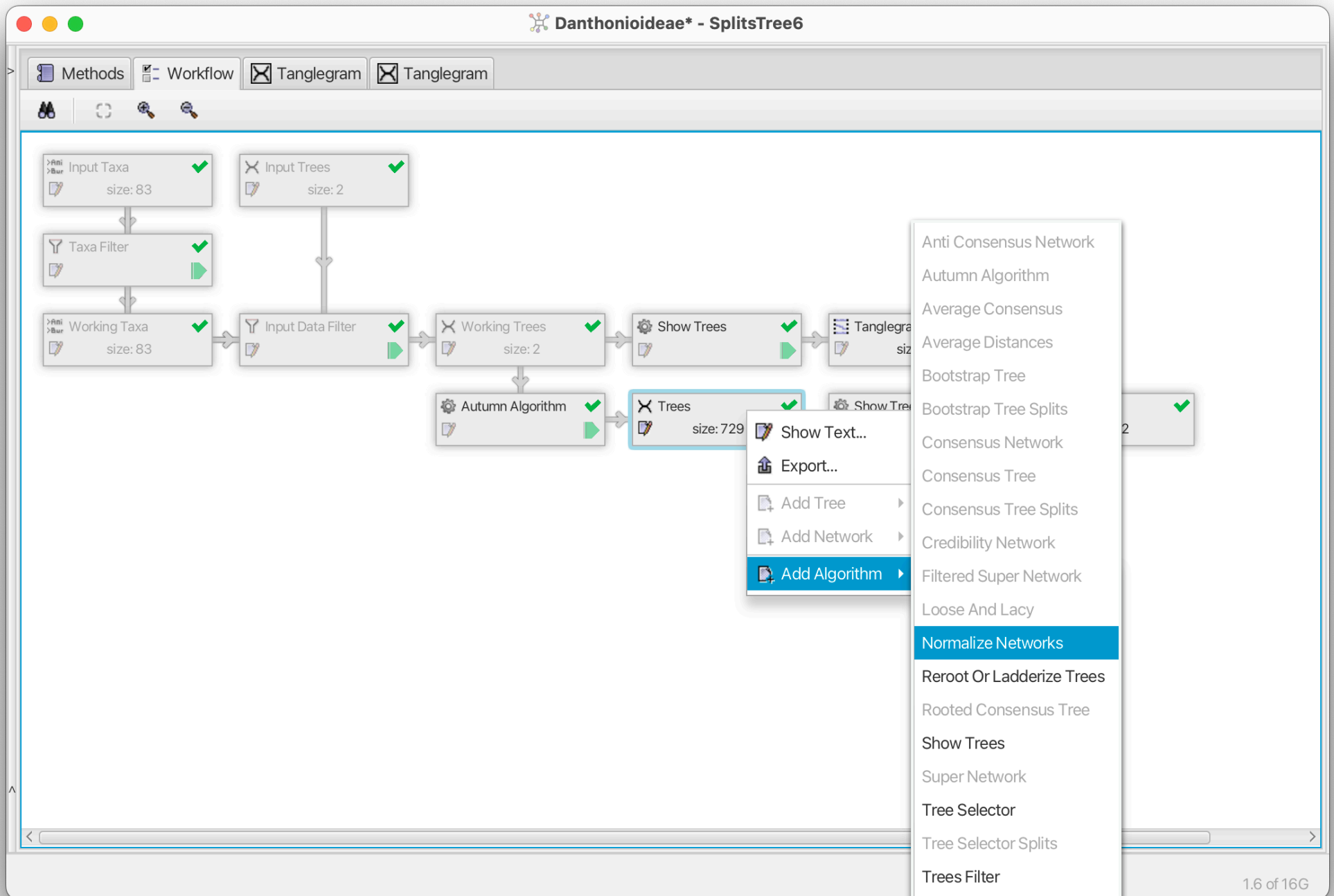
Workflow



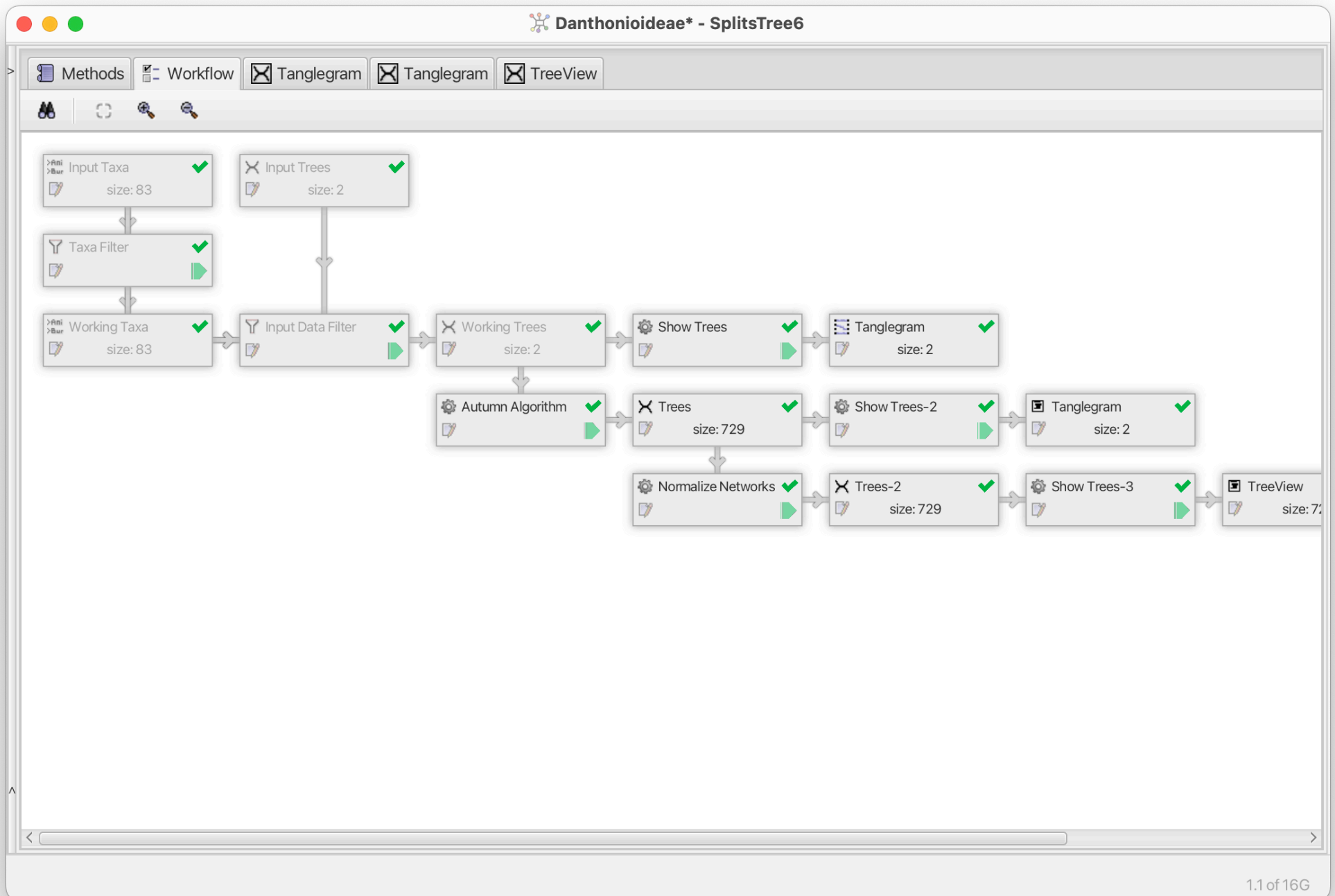
Workflow



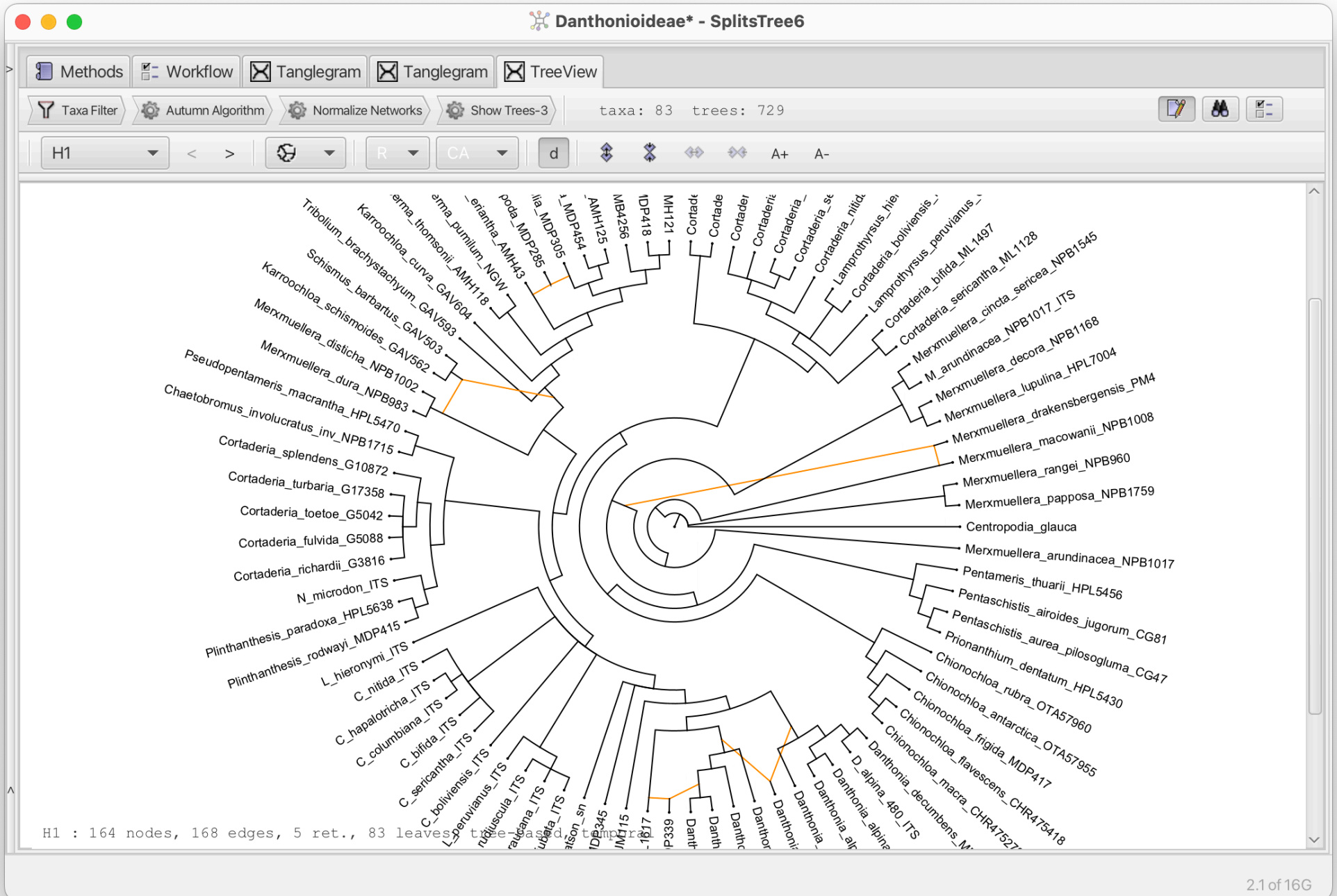
Workflow



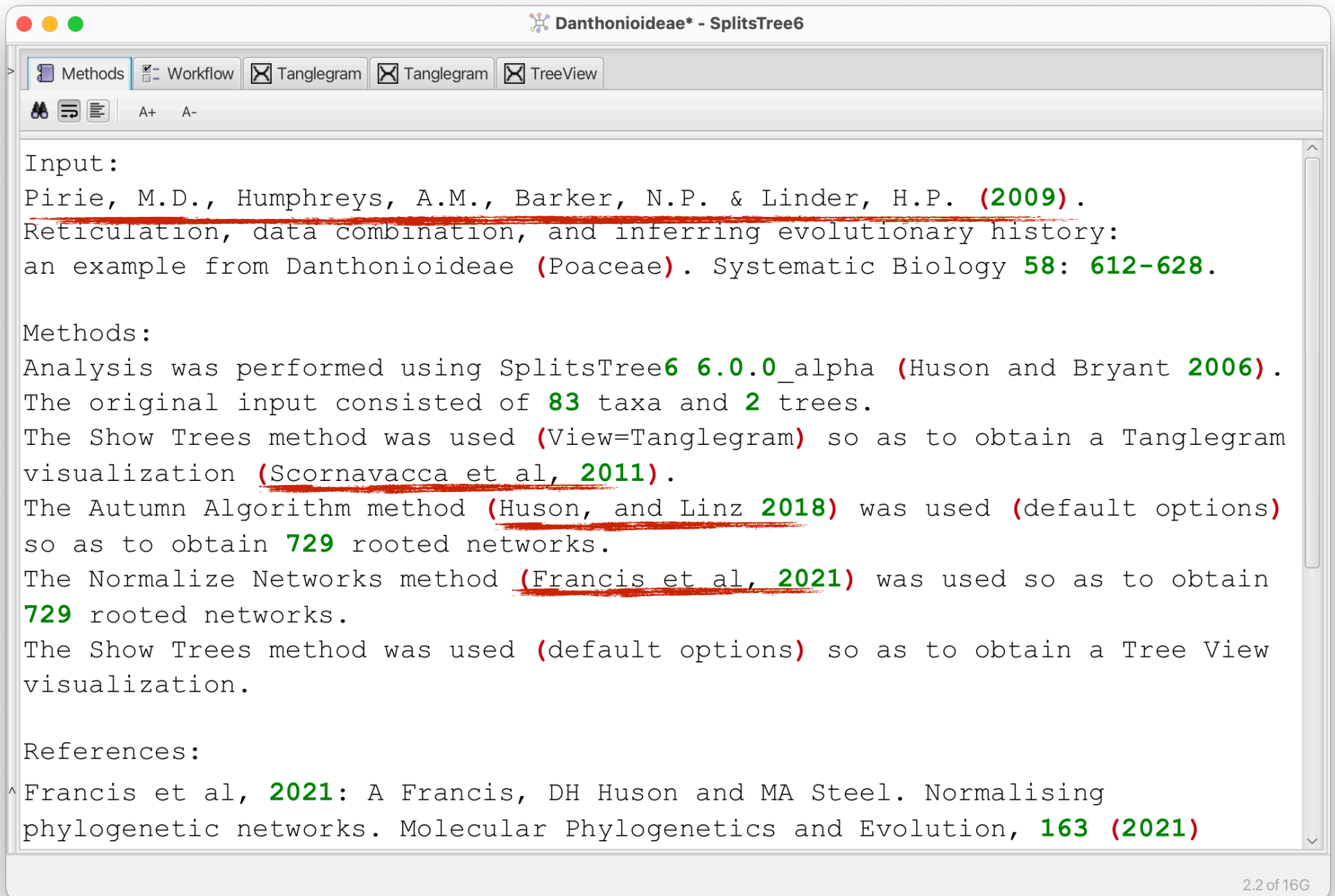
Workflow



Workflow



Workflow and methods



The screenshot shows the SplitsTree6 software interface. The title bar reads "Danthonioideae* - SplitsTree6". The menu bar includes "Methods", "Workflow", "Tanglegram", "Tanglegram", and "TreeView". Below the menu bar are icons for a tree, a list, and a document, along with "A+" and "A-" zoom controls. The main text area contains the following content:

Input:
Pirie, M.D., Humphreys, A.M., Barker, N.P. & Linder, H.P. (2009).
Reticulation, data combination, and inferring evolutionary history:
an example from Danthonioideae (Poaceae). Systematic Biology **58: 612–628.**

Methods:
Analysis was performed using SplitsTree6 **6.0.0_alpha** (Huson and Bryant **2006**).
The original input consisted of **83** taxa and **2** trees.
The Show Trees method was used (View=Tanglegram) so as to obtain a Tanglegram
visualization (Scornavacca et al, 2011).
The Autumn Algorithm method (Huson, and Linz 2018) was used (default options)
so as to obtain **729** rooted networks.
The Normalize Networks method (Francis et al, 2021) was used so as to obtain
729 rooted networks.
The Show Trees method was used (default options) so as to obtain a Tree View
visualization.

References:
Francis et al, **2021**: A Francis, DH Huson and MA Steel. Normalising
phylogenetic networks. Molecular Phylogenetics and Evolution, **163 (2021)**

2.2 of 16G

Singapore 5-29 Sep 2023



Workshops on:

- ▶ Foundations of Phylogenetic Networks
- ▶ Algorithms and Software in Phylogenetic Analysis
- ▶ Evolution of Viruses

The screenshot shows a web browser window displaying the event page for "Mathematics of Evolution-Phylogenetic Trees and Networks" at the IMS Auditorium, National University of Singapore, from September 5 to 29, 2023. The page includes navigation links, contact information for the Institute for Mathematical Sciences, and social media icons for Facebook, Twitter, LinkedIn, and YouTube.

Mathematics of Evolution-Phylogenetic
Trees and Networks
(05 Sep 2023–29 Sep 2023)

Venue
IMS Auditorium

Home / Events / Mathematics of Evolution-Phylogenetic Trees and Networks

Institute for Mathematical Sciences
National University of Singapore

3 Prince George's Park
Singapore 118402
+65 6516 1897
ims-enquiry@nus.edu.sg

f t in YouTube



Thank you!

Joint work with Dave Bryant

Some programming by Daria Evseeva

`https://software-ab.informatik.uni-tuebingen.de/
download/splitstree6`

