Do Phylogenetic Tree Viewers correctly display Support Values?
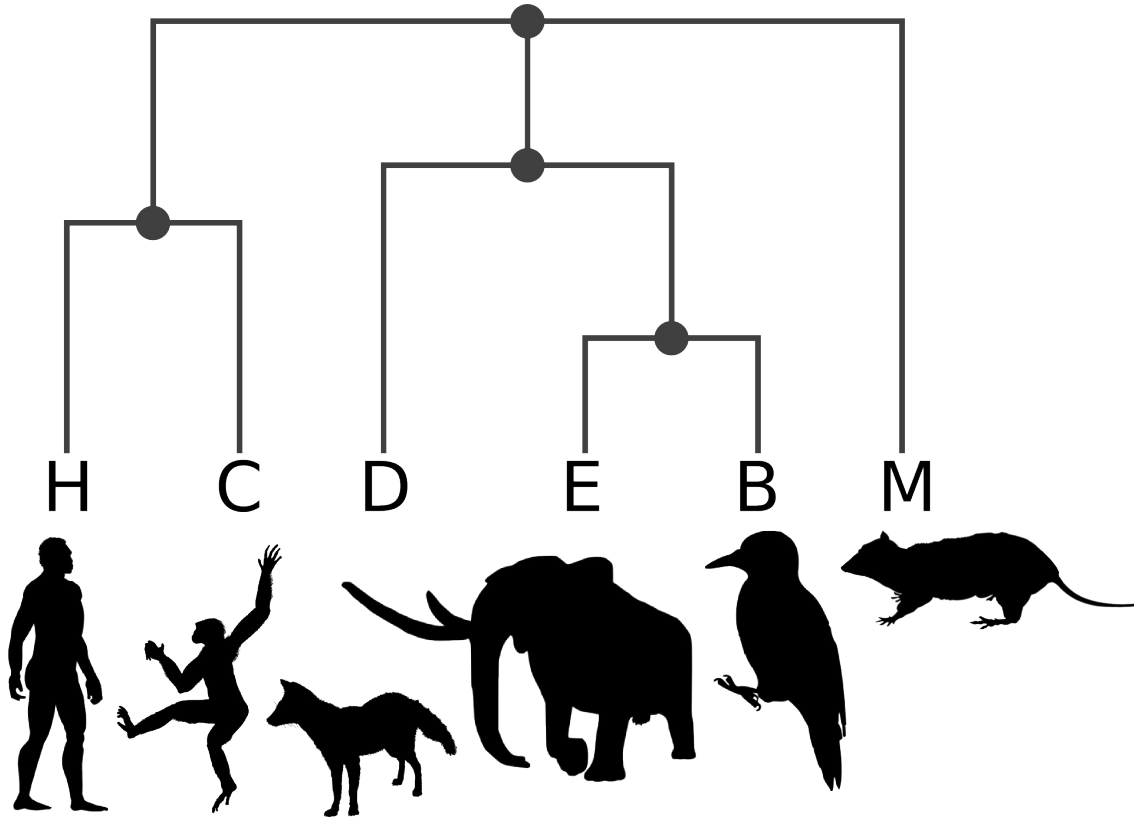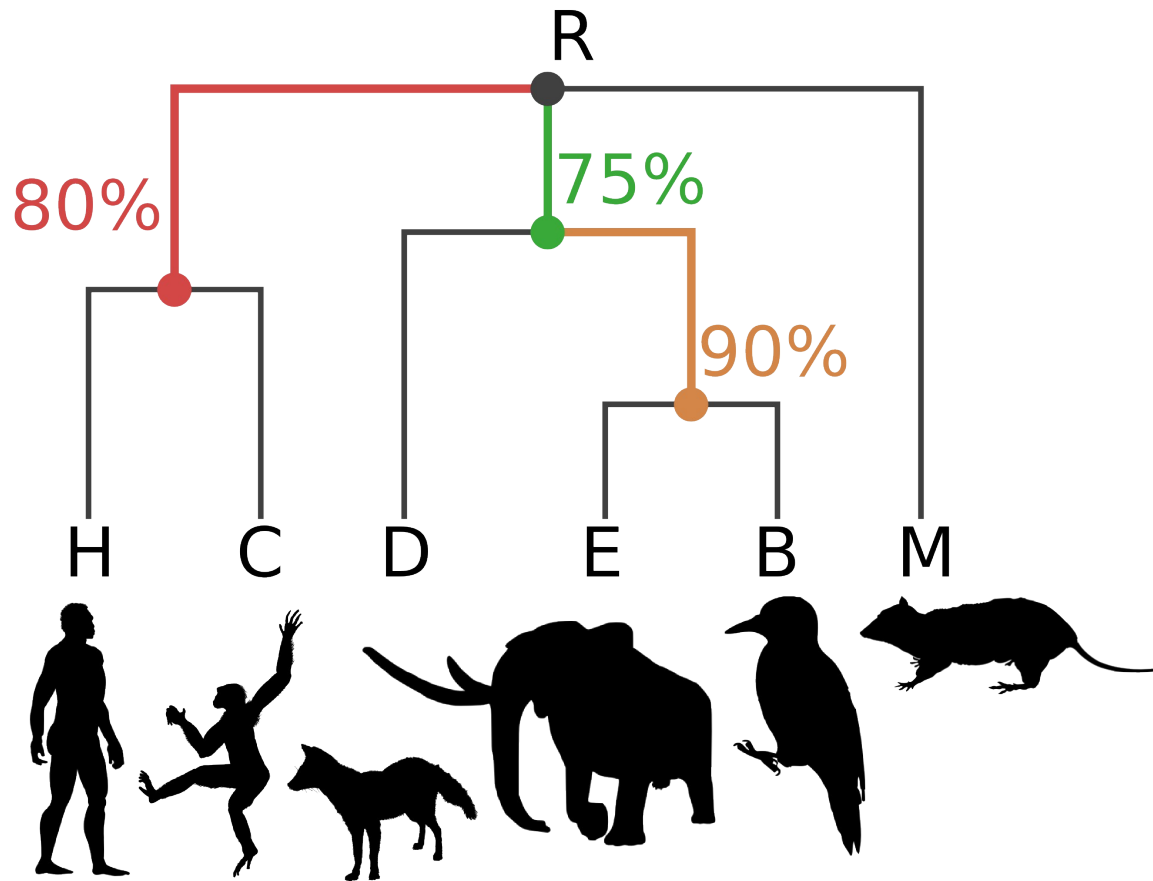
Lucas Czech

Heidelberg Institute for Theoretical Studies

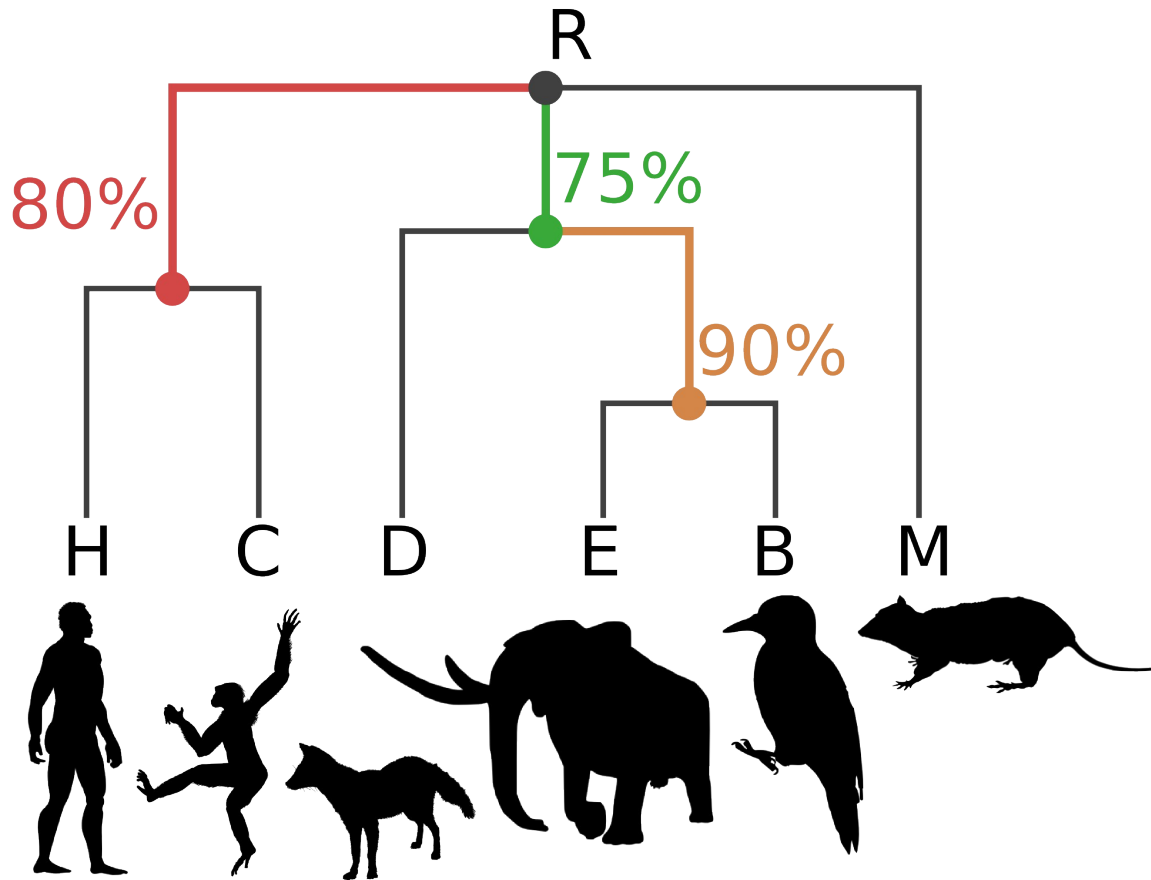Germany

Presentation at the Hbio 2016, 2016-11-19

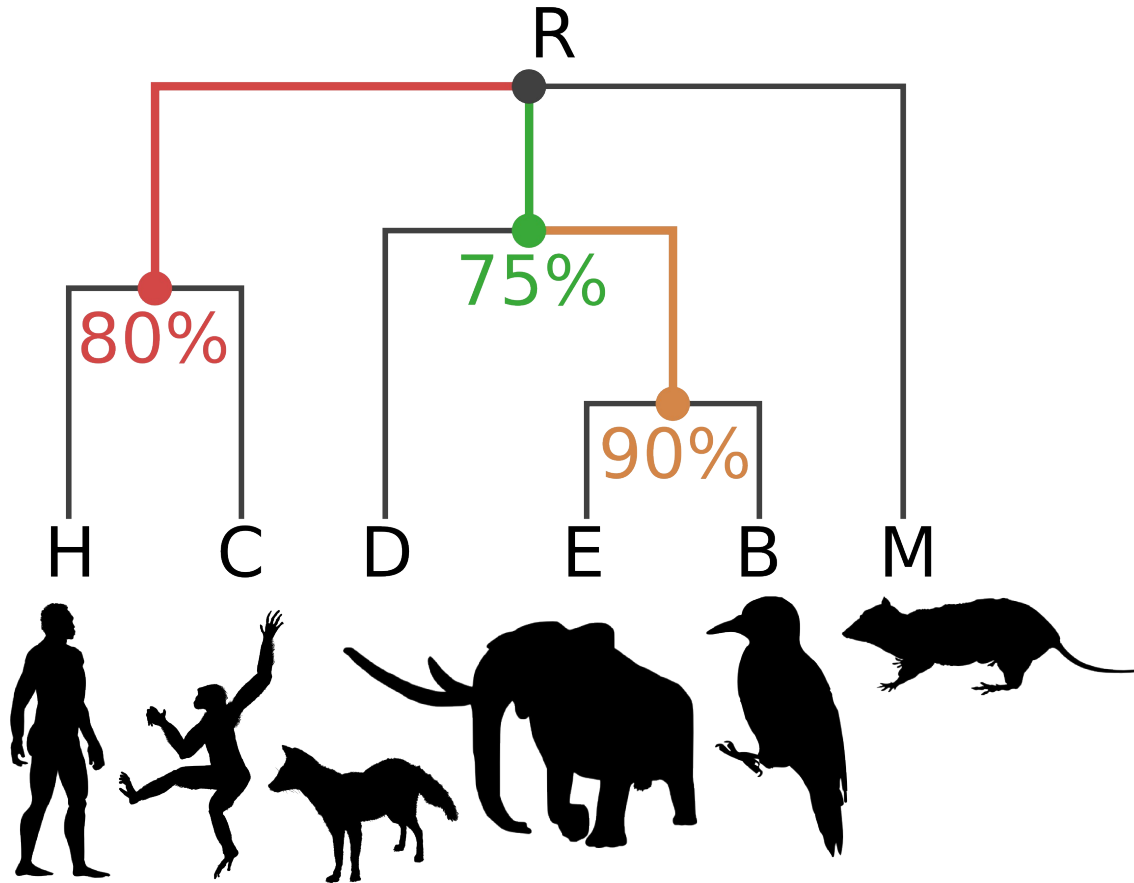# Phylogenetic Tree

Support Values

Newick Format
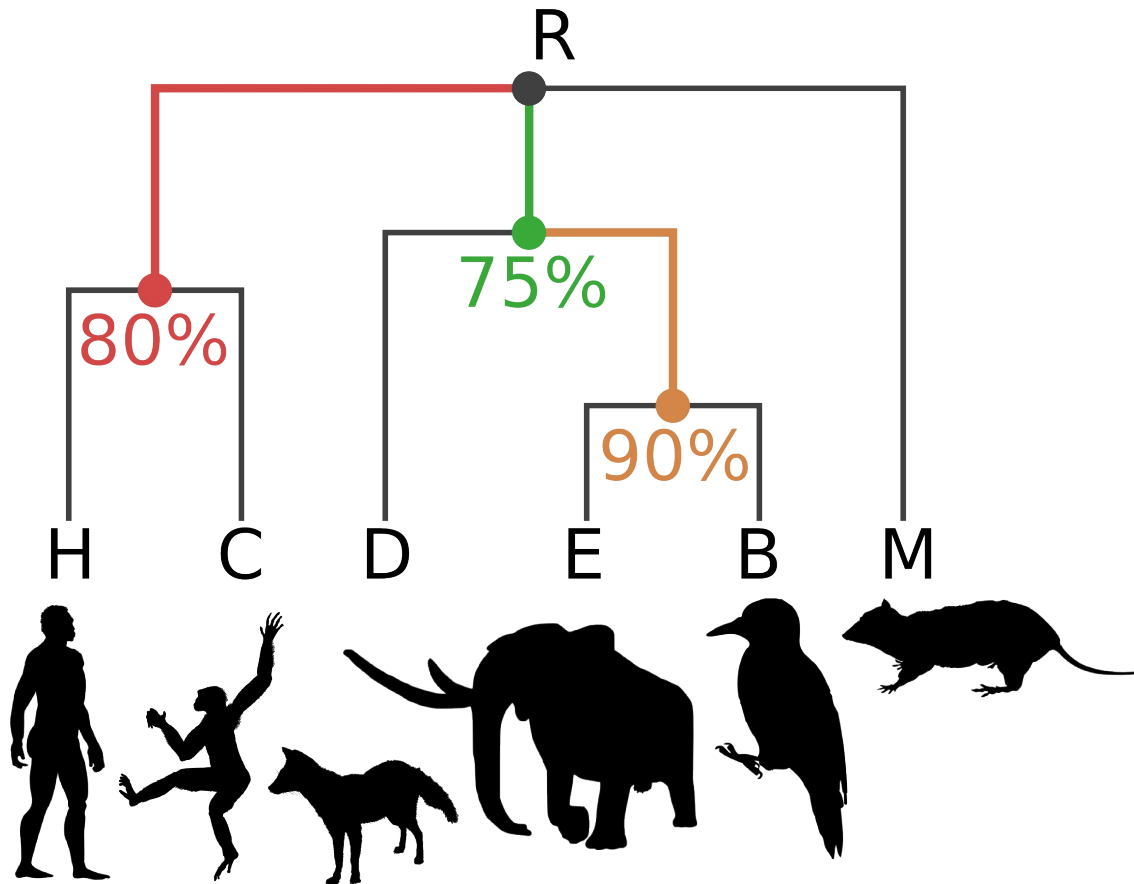

((H, C)0.8, (D, (E, B)0.9)0.75, M)R

Newick Format

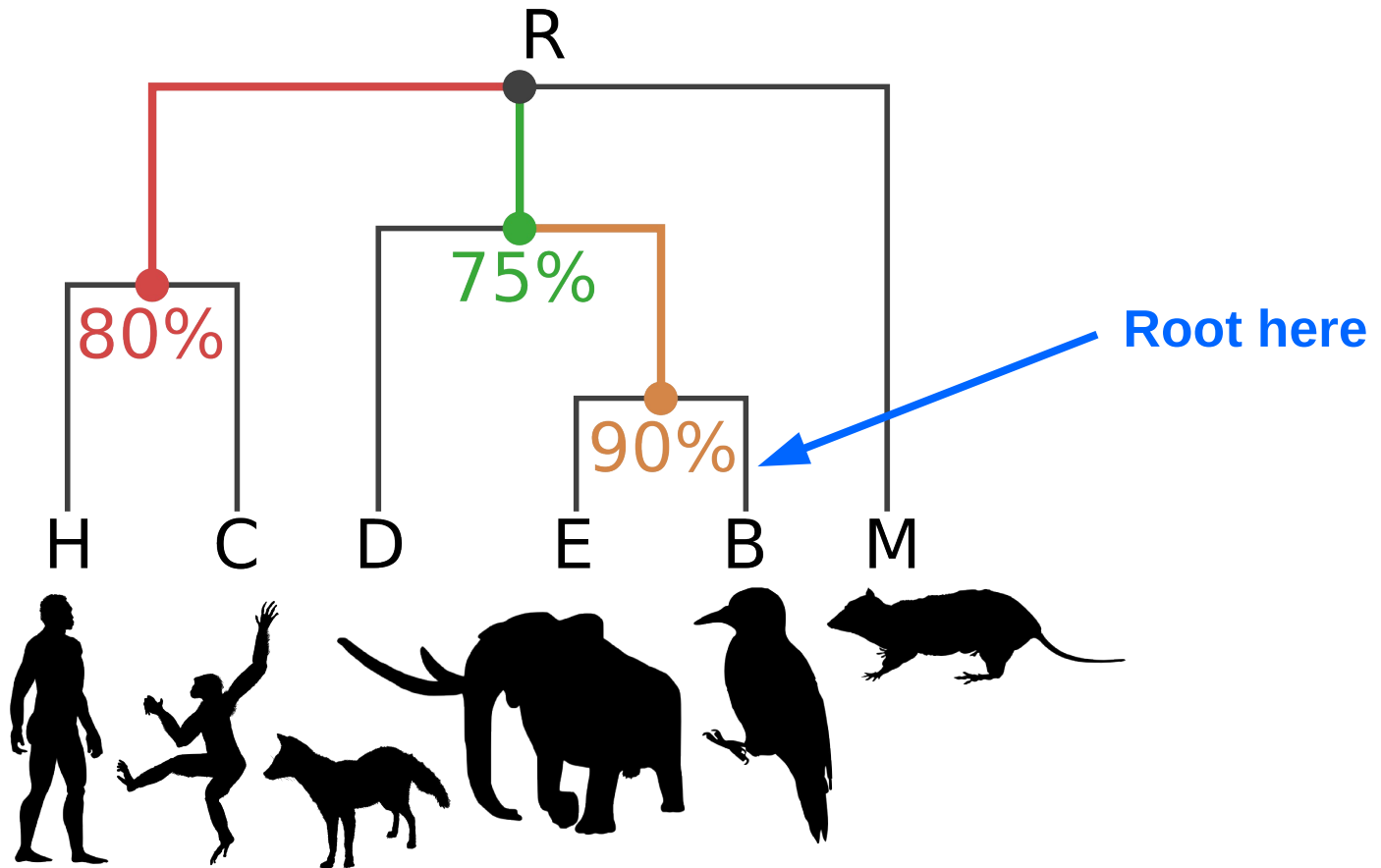$(\ (H, C)0.8, \ (D, \ (E, B)0.9 \ )0.75, \ M \ )R$

Newick Format

((H, C)0.8, (D, (E, B)0.9)0.75, M)R

Convention: Map values towards the root.

Newick Format

((H, C)0.8, (D, (E, B)0.9)0.75, M)R

Convention: Map values towards the root.

HITS
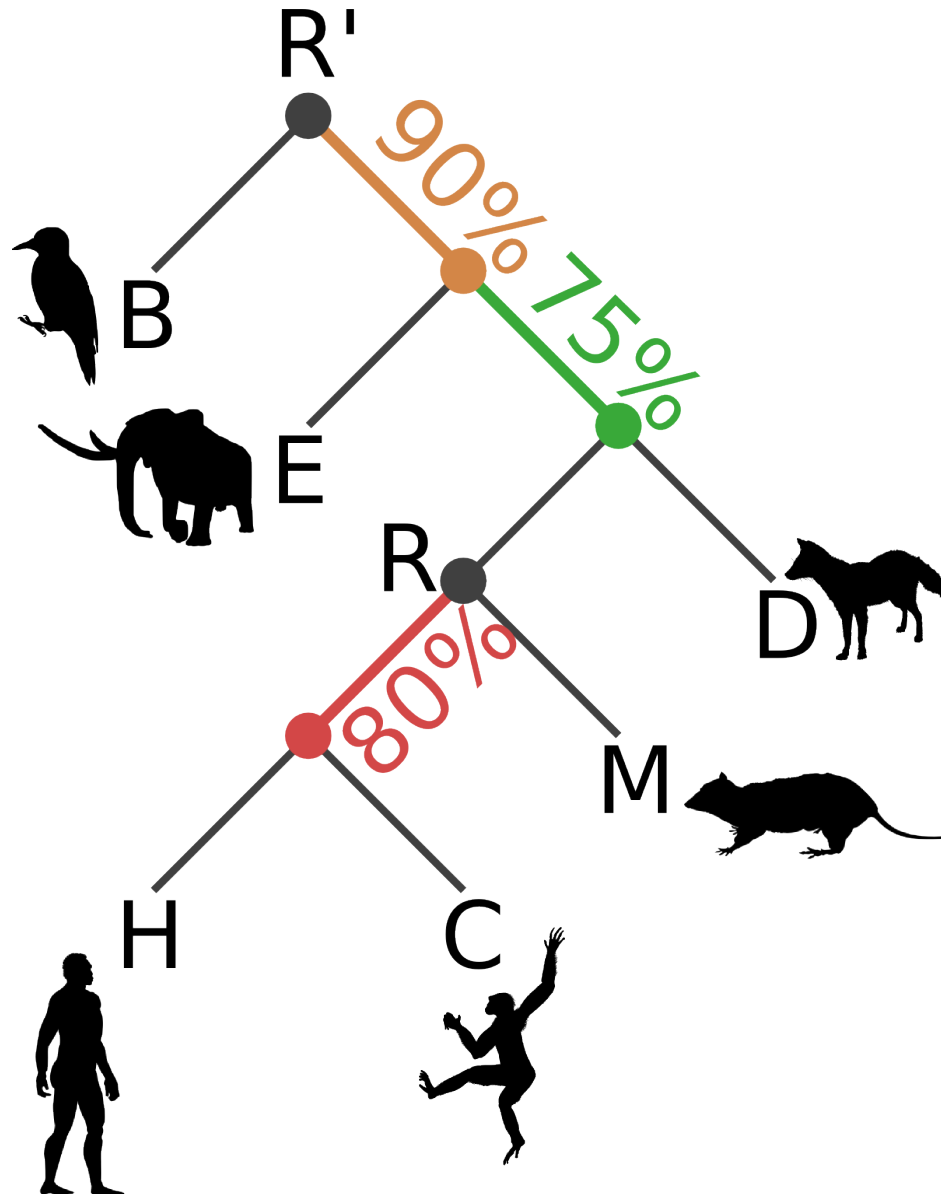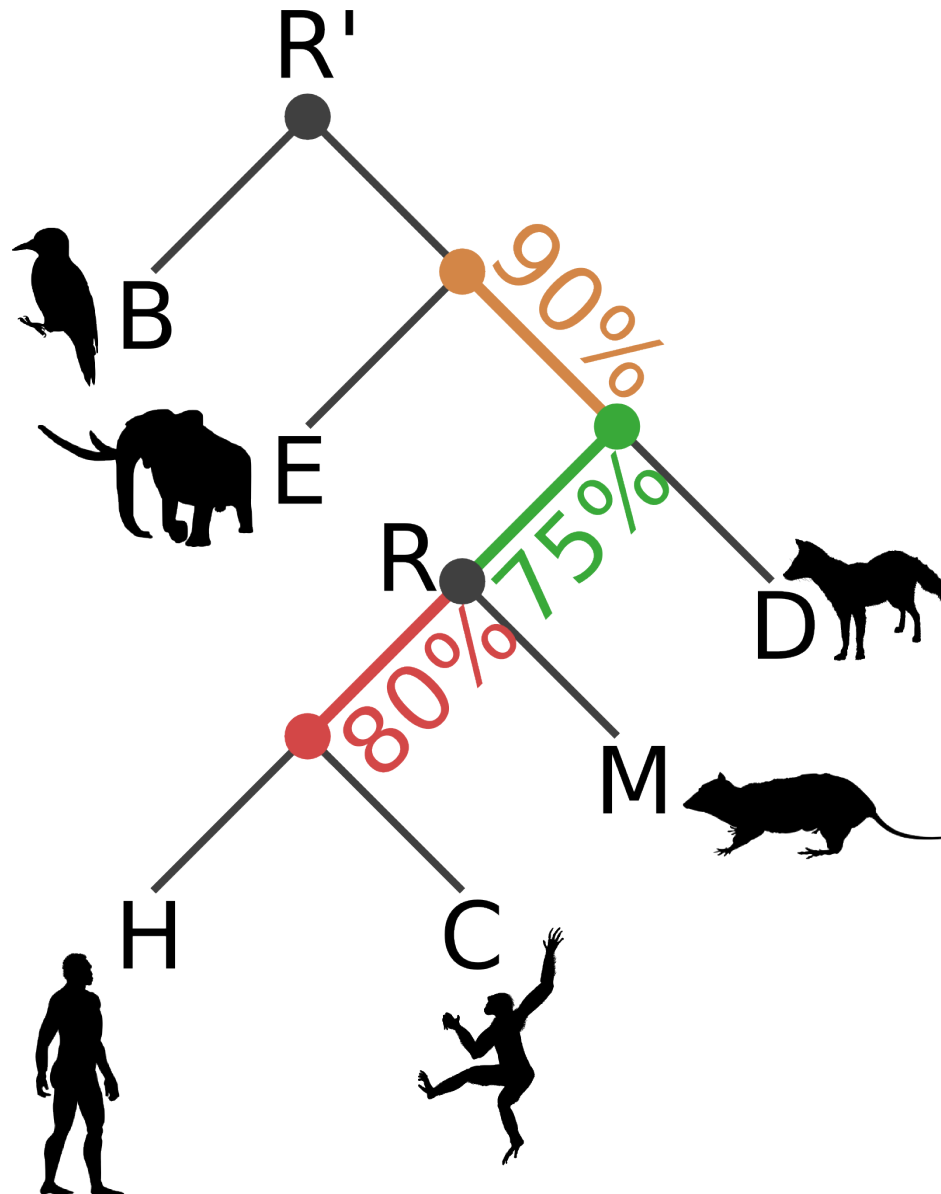
Newick Format with Comments
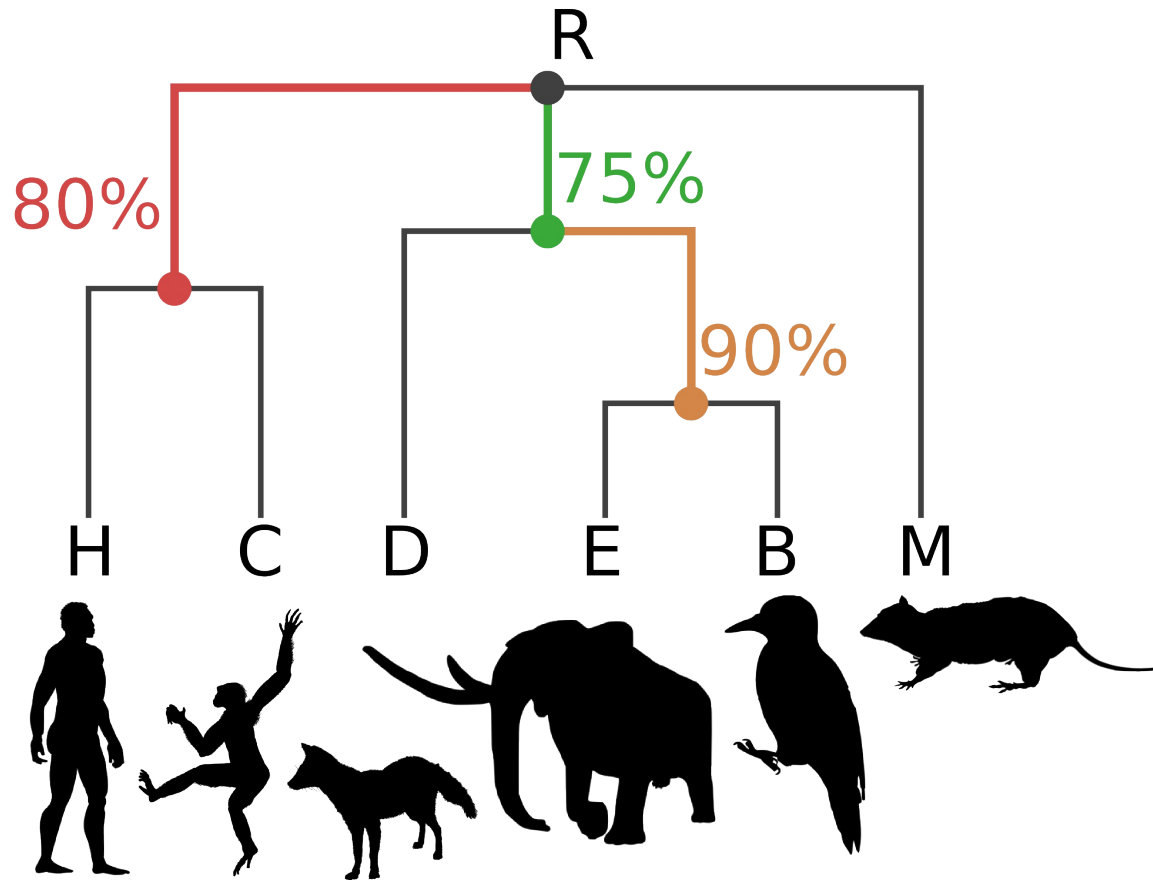
((H, C)[0.8], (D, (E, B)[0.9])[0.75], M)R

Tree Viewers

Viewer	Citations	Nodes	Branches	Default	Correct?
Archaeopteryx	268	x	x	nodes	x
ATV	288		x	branches	x
Dendroscope	1,348	x	x*	dialog*	x*
ETE (GUI)	238	x	x	branches	x
EvolView	105		x	branches	
FigTree	2,362	x	x	both	
iTOL	1,879	x	x	input dependent	x
PhyloWidget	113	x		nodes	
TreeView	10,570		x	branches	x
T-REX	285		x	branches	(x)
Total	**17,446**				

Bioinformatics Toolkits

Toolkit	Citations	Nodes	Branches	Default	Correct?
APE	3,915	x	x*	nodes	x*
BioPerl	1,410	x	x	nodes	x
BioPython	797	x		nodes	
Dendropy	525	x	x*	nodes	(x)
ETE (API)	238	x	x	branches	x
Geneious	1,689	x	(x)	nodes	
Mega	69,134		x	branches	x
Mesquite	5,616	x	x	nodes	x
Newick Utilities	31	x	x*	nodes	x*
Pycogent/scikit-bio	148		x	branches	x
Total	**83,275**				

HITS

EvolView

EvolView
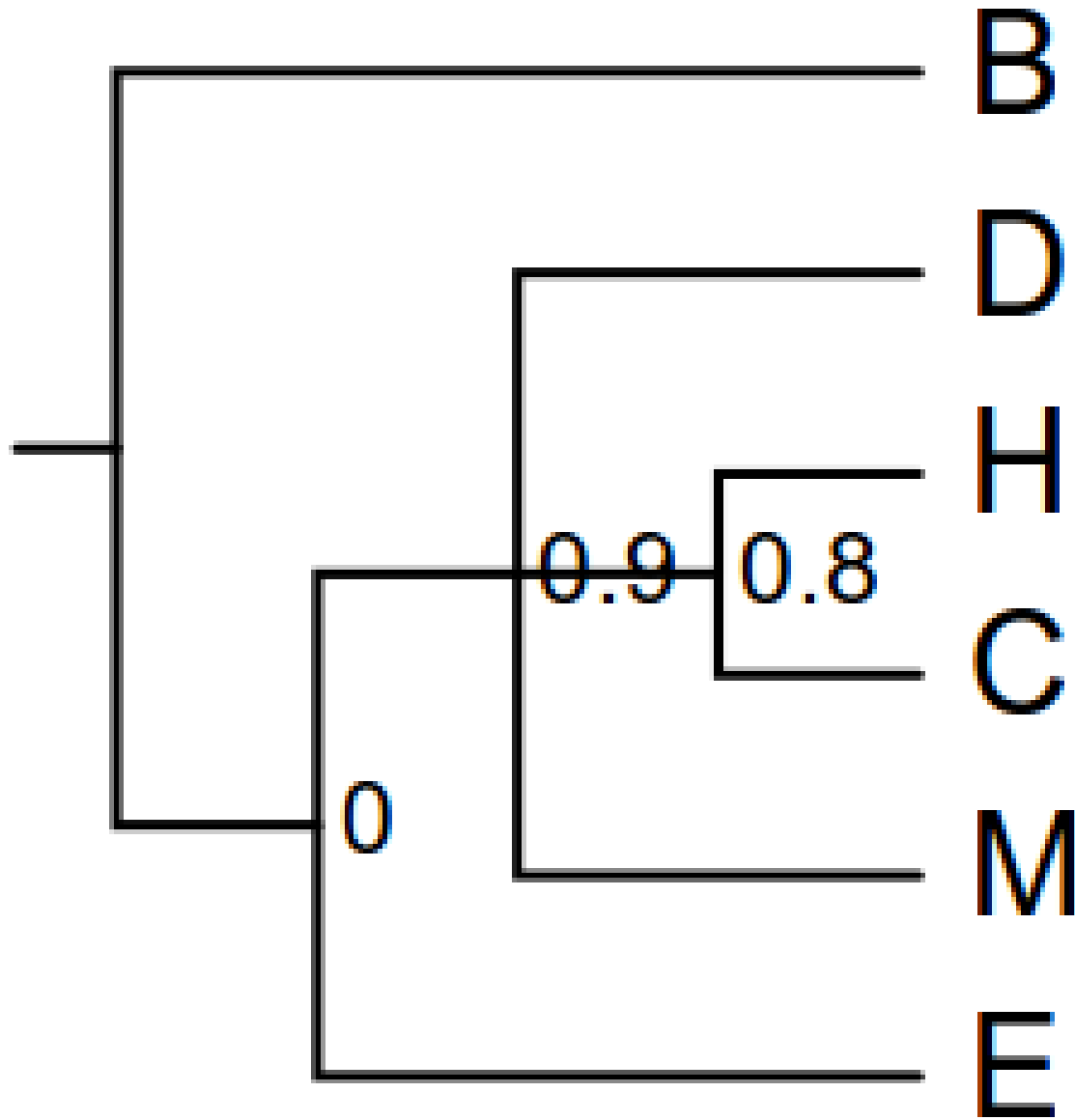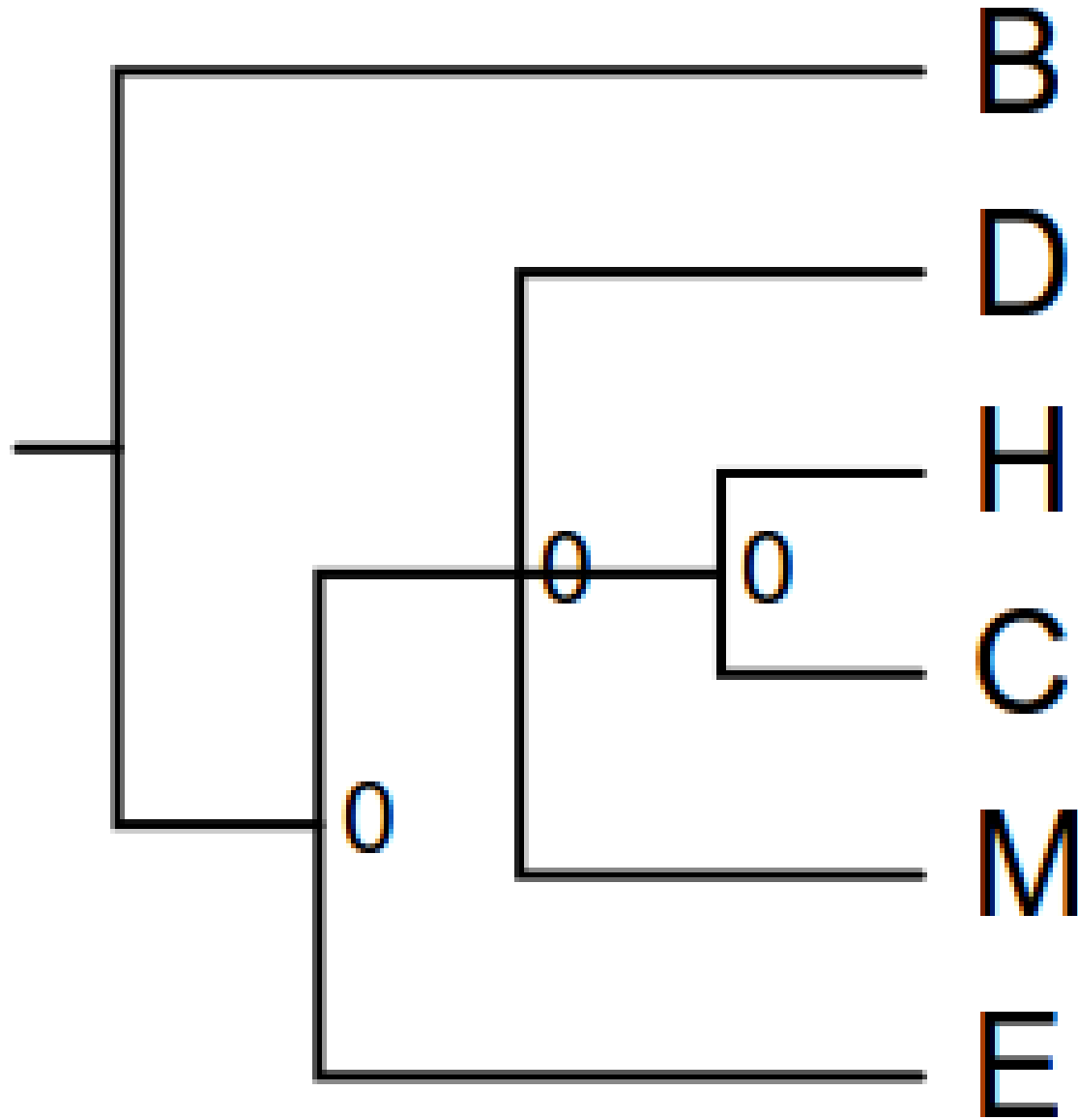
Dendroscope

How to interpret internal node numbers

Internal nodes have labels (example: '0.8'),
how should they be interpreted?

Interpret as node labels ▼
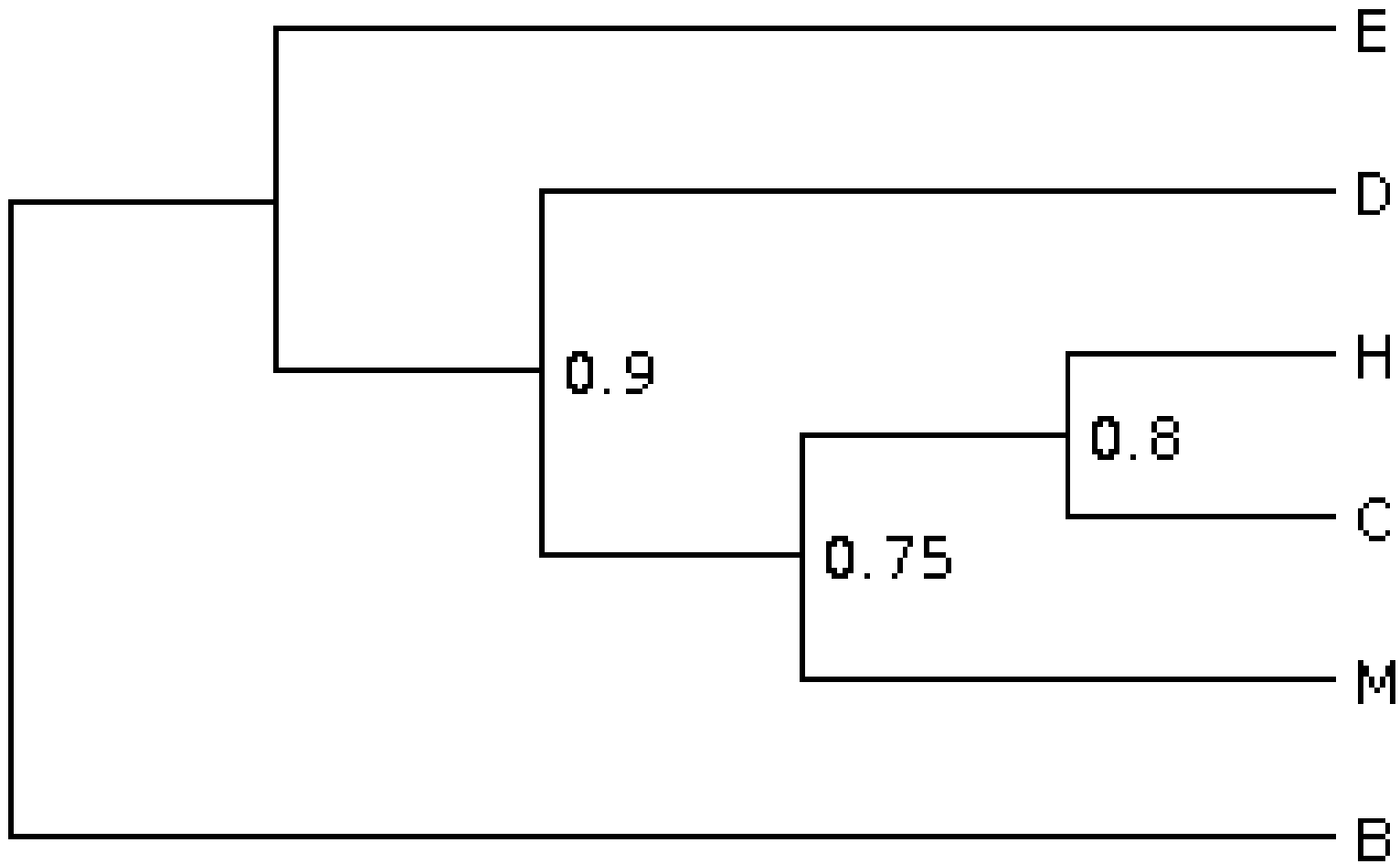
Interpret as node labels
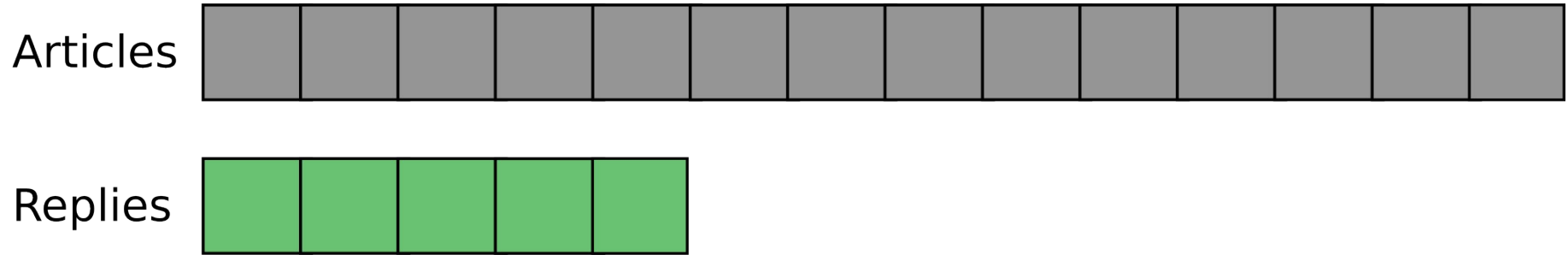Interpret as edge labels (such as bootstrap values)
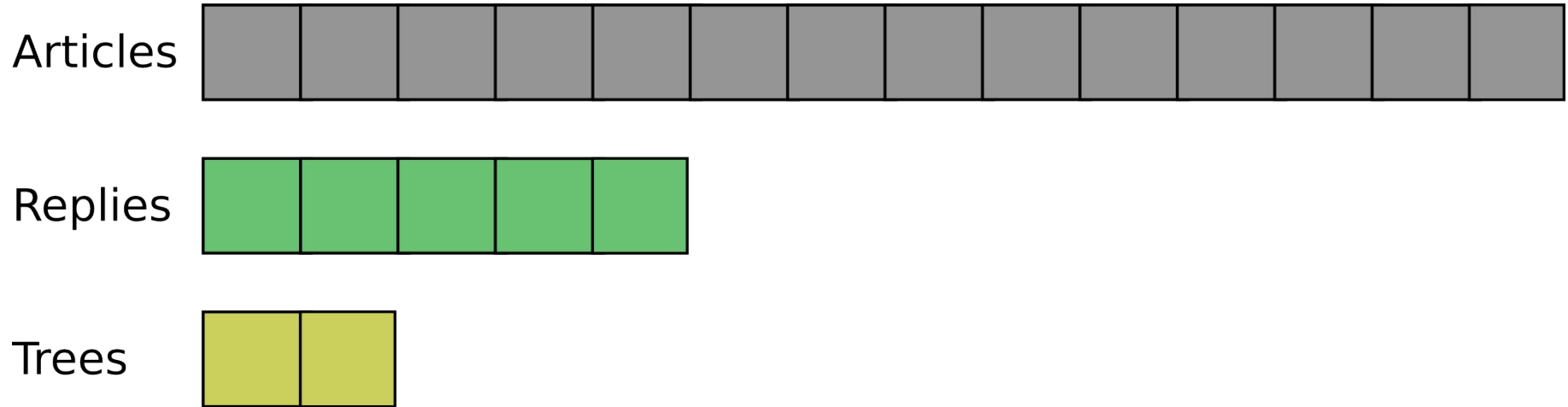Delete

HITS

Dendroscope

Published Data

Articles

Published Data

Articles

Replies

Published Data

Articles

Replies

Trees

Published Data

Articles

Replies

Trees

Errors

Published Data

T_thermophilus_HB27_NrdA@YP_006145
T_thermophilus_HB8_NrdA@YP_145448
T_aquaticusY51MC23_NrdA@ZP_03497649
S_ruber_13855_NrdA@YP_444285
H_lacusprofundi49239_NrdA@YP_002564381
Halophage_AAJ2005_NrdA@ABB77922
Env_halophage_1_NrdA@ABB77922
N_pharaonis_DSM_216_NrdA@YP_327711
H_mukohataei12286_NrdA@ZP_03875490
D_nodosus_VCS1703A_NrdA@YP_001209517
Psychromonas_CNPT3_NrdA2@ZP_01215780
P_irgensii_23P_NrdA@ZP_01117179
L_blandensisMED217_NrdA@EAQ51287
uid_eubacteriumSCB49_NrdA@ZP_01889446
F_johnsoniae_UW101_NrdA1@YP_001196392
C_sputigena_Capno_NrdA@ZP_03392007
Aeromonas_phage_31_NrdA@YP_238837
F_novicidaGA993548_NrdA@YP_002170840
F_novicida_U112_NrdA@YP_898623
F_tularensisWY963418_NrdA@YP_001121958
F_tula_t_SCHUS4_NrdA@YP_169563
F_tularensisLVS_NrdA@YP_513694
Grouper_iridovirus_NrdA@AAV91060
Singapore_g_iridov_NrdA@YP_164159
SoftshelTurtleVirus_NrdA@YP_002854273
Frog_virus_3_NrdA@YP_031616
A_tigrinum_st_v_NrdA@AAP33245
R_tigrina_ranavirus_NrdA@NP_571996
Lymphocystis_d_v_1_NrdA@NP_078756
Lymphocystis_v_Ch_NrdA@YP_025104
Invert_irid_v_6_NrdA@NP_149548
A_taeniorhynchus_IV_NrdA@YP_654637
S_alaskensis_RB2256_NrdA@YP_616819
S_wittichii_RW1_NrdA@YP_001262005
N_aromat_DSM12444_NrdA@YP_495556
E_litoralis_HTCC259_NrdA@YP_459216
Erythrobacter_NAP1_NrdA@ZP_01041166
ErythrobacterSD21_NrdA@ZP_01863359
R_rubrum_ATCC_11170_NrdA@YP_426532
B_japonicum_110_NrdA@NP_766647
B_elkanii_NrdA@CAQ57558

0.996000
1.000000
0.348000
0.897000
0.979000
0.993000
0.995000
0.991000
0.987000
0.536000

New Root

Published Data

Correct

Aeromonas_phage_31_NrdA@YP_238837

0.993000
F_johnsoniae_UW101_NrdA1@YP_001196392
C_sputigena_Capno_NrdA@ZP_03392007

0.979000
Psychromonas_CNPT3_NrdA2@ZP_01215780
P_irgensii_23P_NrdA@ZP_01117179
L_blandensisMED217_NrdA@EAQ51287
uid_eubacteriumSCB49_NrdA@ZP_01889446

0.897000
D_nodosus_VCS1703A_NrdA@YP_001209517
S_ruber_13855_NrdA@YP_444285

0.348000
H_lacusprofundi49239_NrdA@YP_002564381
Halophage_AAJ2005_NrdA@ABB77922
Env_halophage_1_NrdA@ABB77922

1.000000
N_pharaonis_DSM_216_NrdA@YP_327711
H_mukohataei12286_NrdA@ZP_03875490

T_aquaticusY51MC23_NrdA@ZP_03497649
T_thermophilus_HB27_NrdA@YP_006145
T_thermophilus_HB8_NrdA@YP_145448

F_novicidaGA993548_NrdA@YP_002170840
F_novicida_U112_NrdA@YP_898623
F_tularensisWY963418_NrdA@YP_001121958
F_tula_t_SCHUS4_NrdA@YP_169563
F_tularensisLVS_NrdA@YP_513694

0.995000
Grouper_iridovirus_NrdA@AAV91060
Singapore_g_iridov_NrdA@YP_164159
SoftshelTurtleVirus_NrdA@YP_002854273
Frog_virus_3_NrdA@YP_031616

0.991000
A_tigrinum_st_v_NrdA@AAP33245
R_tigrina_ranavirus_NrdA@NP_571996
Lymphocystis_d_v_1_NrdA@NP_078756
Lymphocystis_v_Ch_NrdA@YP_025104

Invert_irid_v_6_NrdA@NP_149548

1.000000
A_taeniorhynchus_IV_NrdA@YP_654637

0.987000
S_alaskensis_RB2256_NrdA@YP_616819
S_wittichii_RW1_NrdA@YP_001262005
N_aromat_DSM12444_NrdA@YP_495556
E_litoralis_HTCC259_NrdA@YP_459216
Erythrobacter_NAP1_NrdA@ZP_01041166
ErythrobacterSD21_NrdA@ZP_01863359

R_rubrum_ATCC_11170_NrdA@YP_426532
B_japonicum_110_NrdA@NP_766647
B_elkanii_NrdA@CAQ57558

0.536000
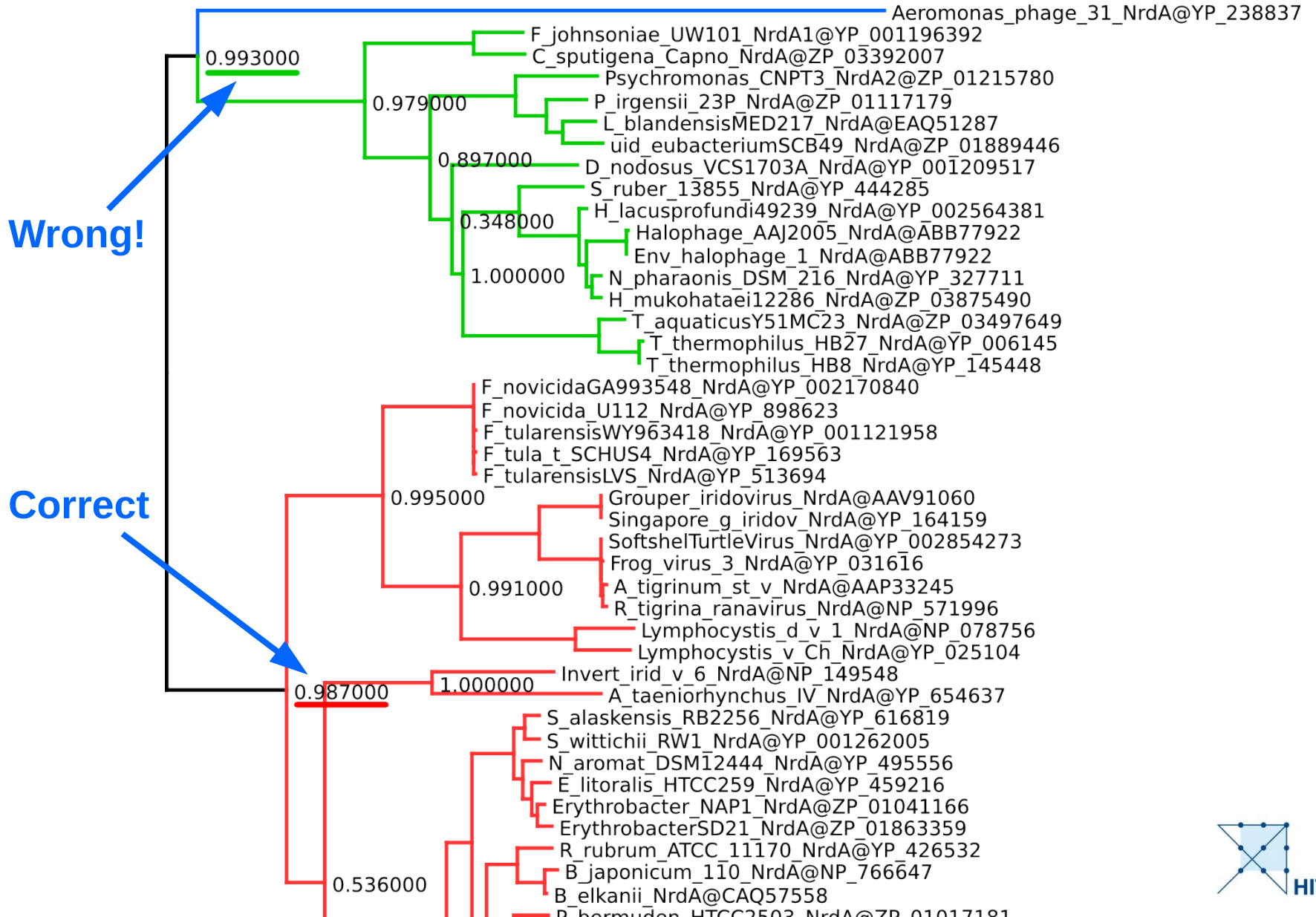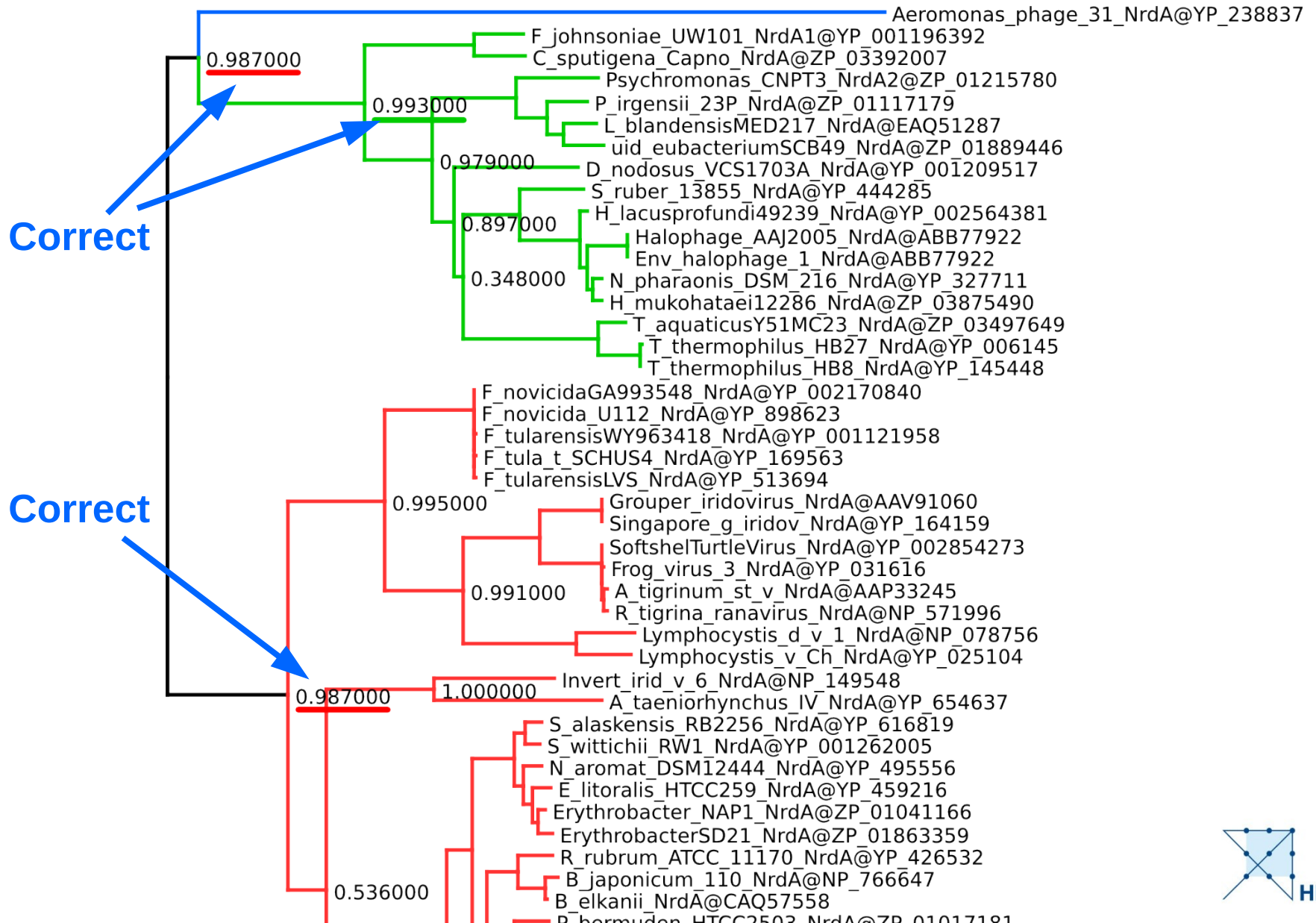R_bermuden_HTCC2503_NrdA@ZP_01017181

HITS

Published Data

Wrong!

Correct

Published Data

Do Phylogenetic Tree Viewers correctly display Support Values?

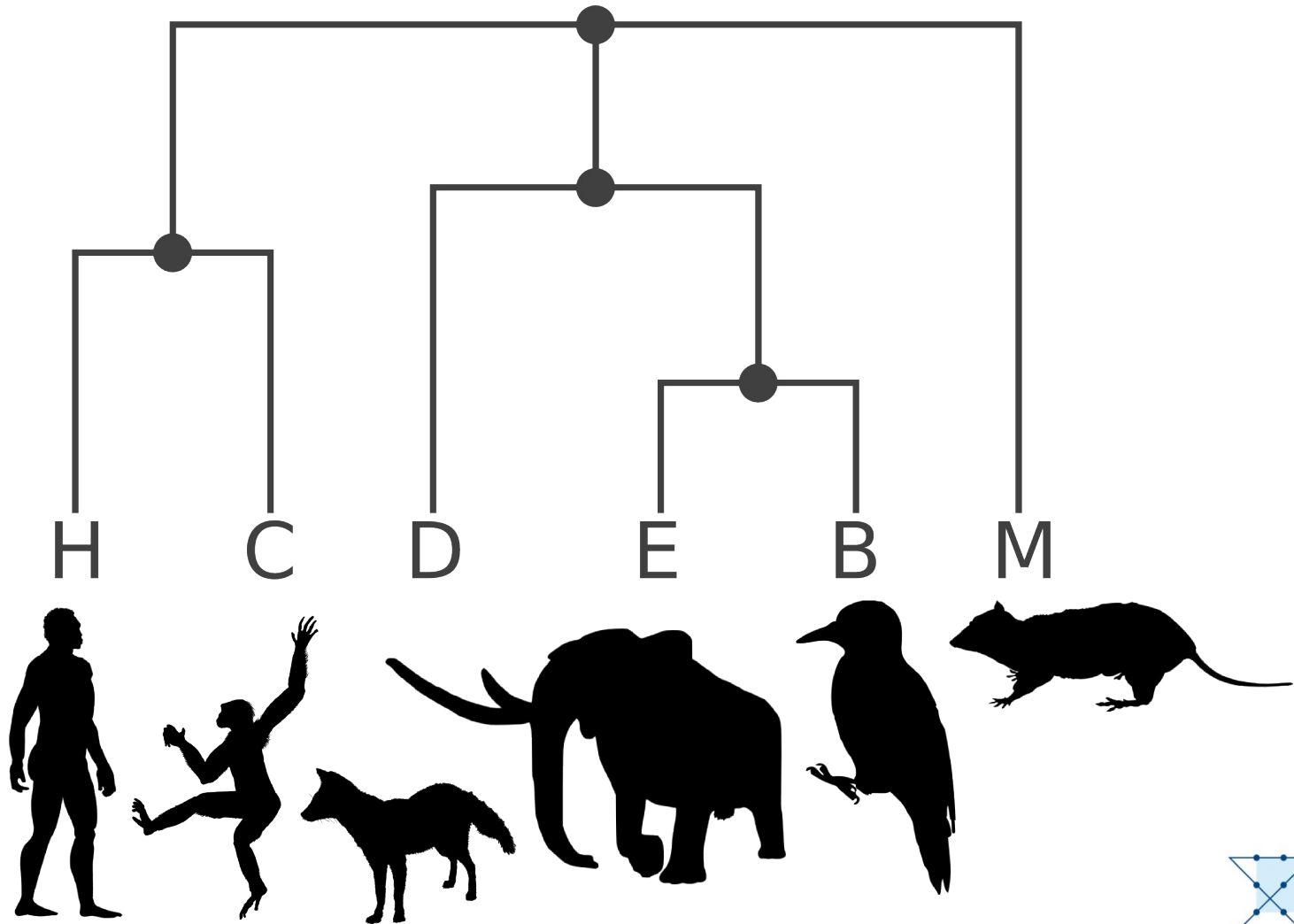# NO!

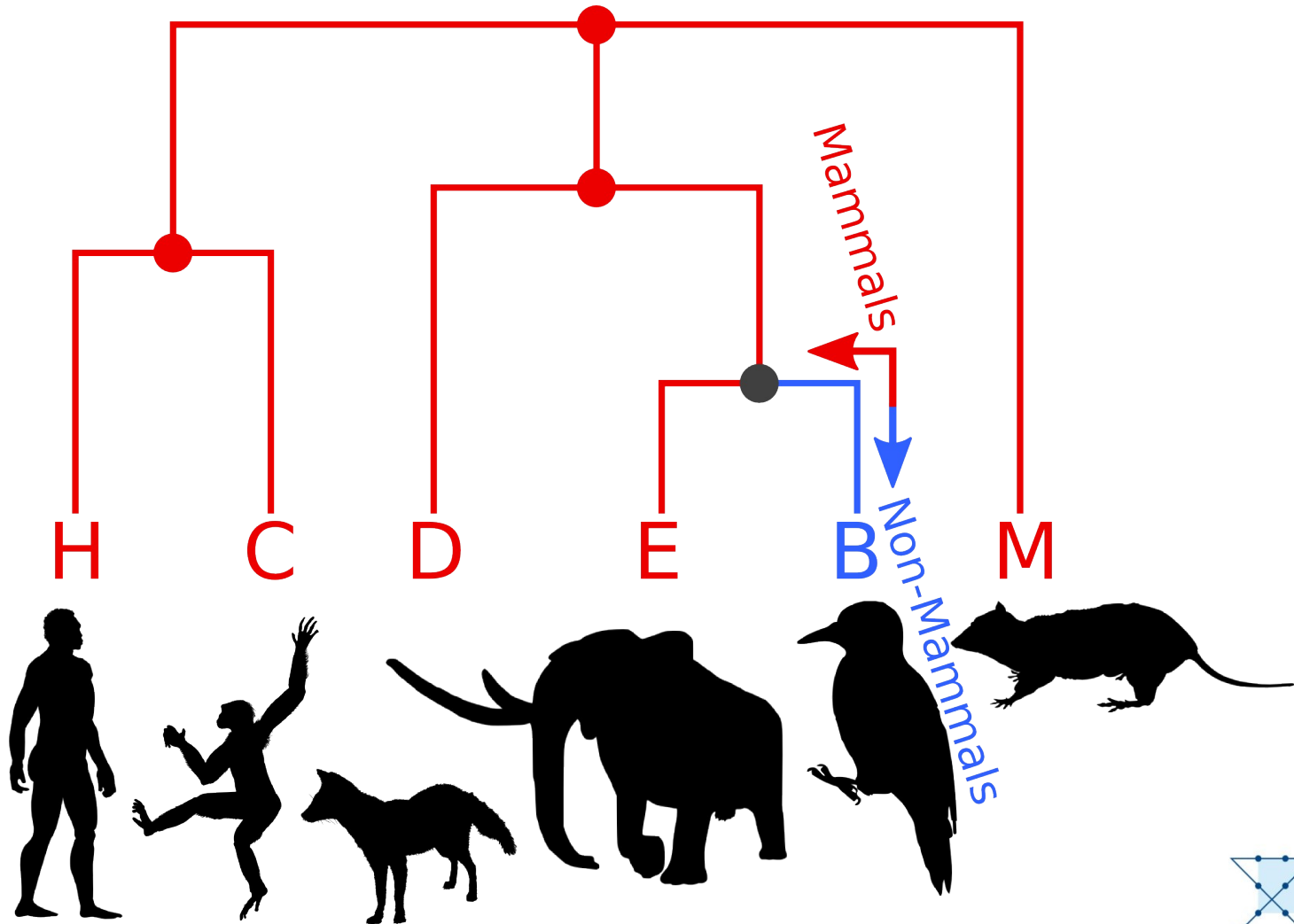# Do Phylogenetic Tree Viewers correctly display Support Values?
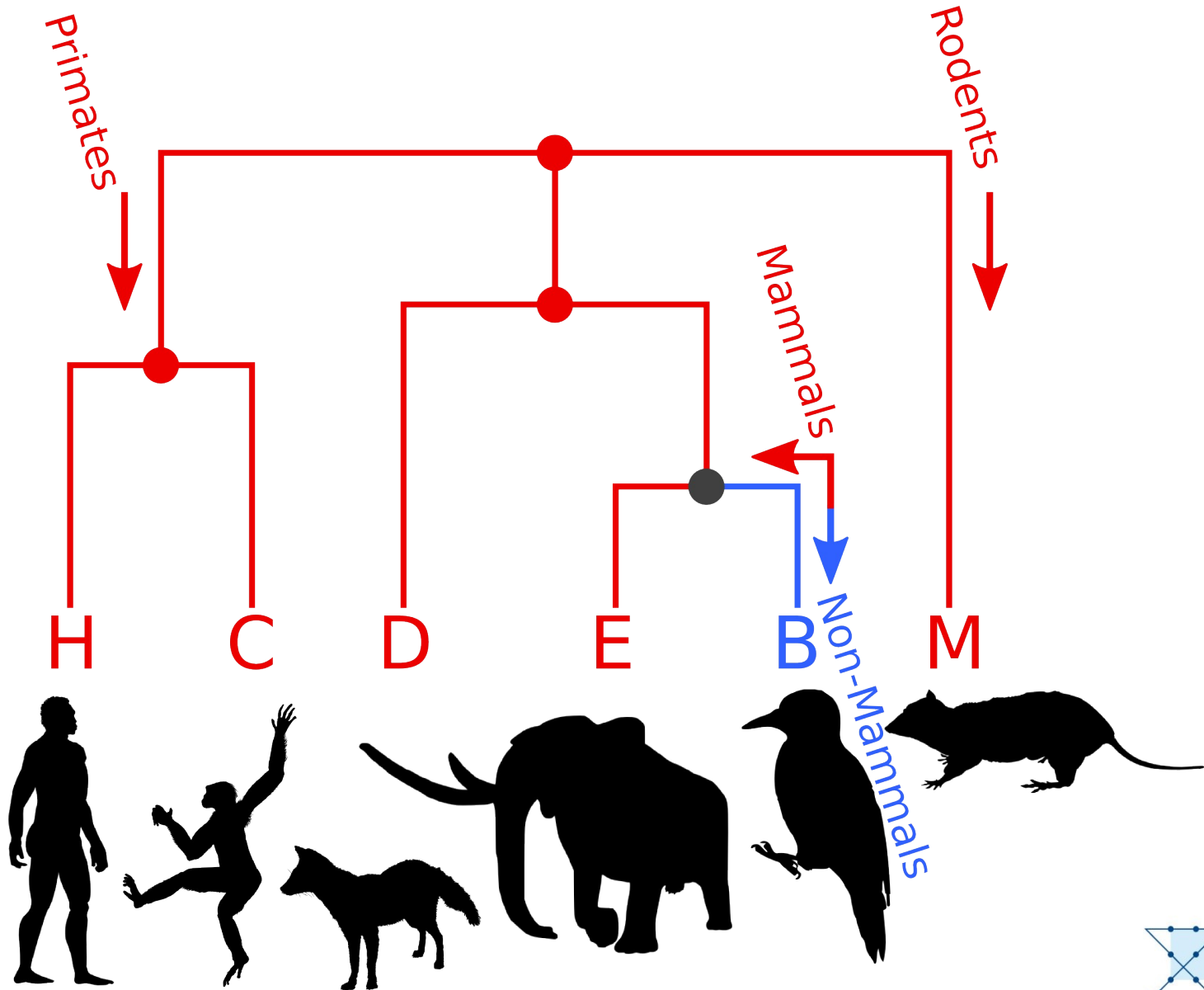
NO!

… but is that all?

HITS

Clade Names

Values Attached to Trees

- Node values: Species names

- Branch values: Support values

- Clade values: Clade names

- Node values: Species names

- Branch values: Support values

- Clade values: Clade names

Newick is hopelessly overburdened!

HITS

- Node values: Species names

- Branch values: Support values

- Clade values: Clade names

Newick is hopelessly overburdened!

- PhyloXML?

- Need for a new standard?

HITS

- Pay attention to ambiguous file formats.

HITS

- Pay attention to ambiguous file formats.

- If in doubt, let the user choose!

HITS

Suggestions for Developers

- Pay attention to ambiguous file formats.

- If in doubt, let the user choose!

- Keep in mind how re-rooting affects annotated values.

HITS

- Pay attention to ambiguous file formats.

- If in doubt, let the user choose!

- Keep in mind how re-rooting affects annotated values.

- Raise awareness for the issue.

HITS

- Pay attention to the options a tool offers.

HITS

- Pay attention to the options a tool offers.

- If available, use the option to set the desired interpretation.

HITS

- Pay attention to the options a tool offers.

- If available, use the option to set the desired interpretation.

- Be sure that re-rooting is a valid operation for your type of tree and its associated data.

HITS

- Pay attention to the options a tool offers.

- If available, use the option to set the desired interpretation.

- Be sure that re-rooting is a valid operation for your type of tree and its associated data.

- Double-check your results, maybe try other tools, or conduct a visual inspection.

HITS

Thanks & Acknowledgement

EMBL — European Molecular Biology Laboratory
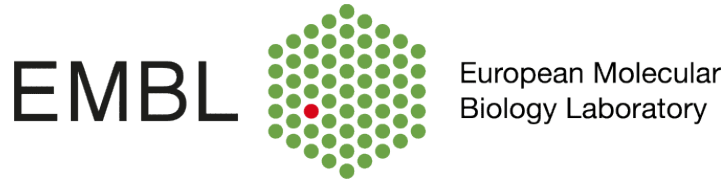
Klaus Tschira Stiftung gemeinnützige GmbH — KTS

Jaime Huerta-Cepas   Alexandros Stamatakis

HITS

Time for questions.