MDSC 308

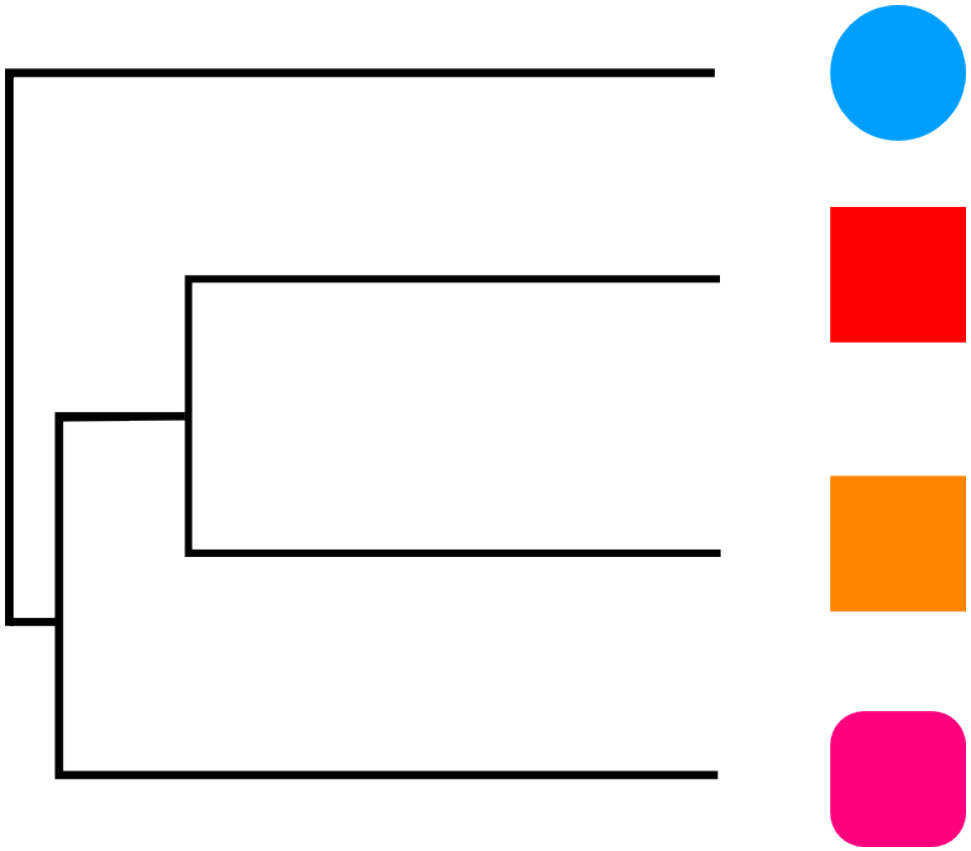
Phylogenetics

@i_kryukov
ivan-kryukov.github.io

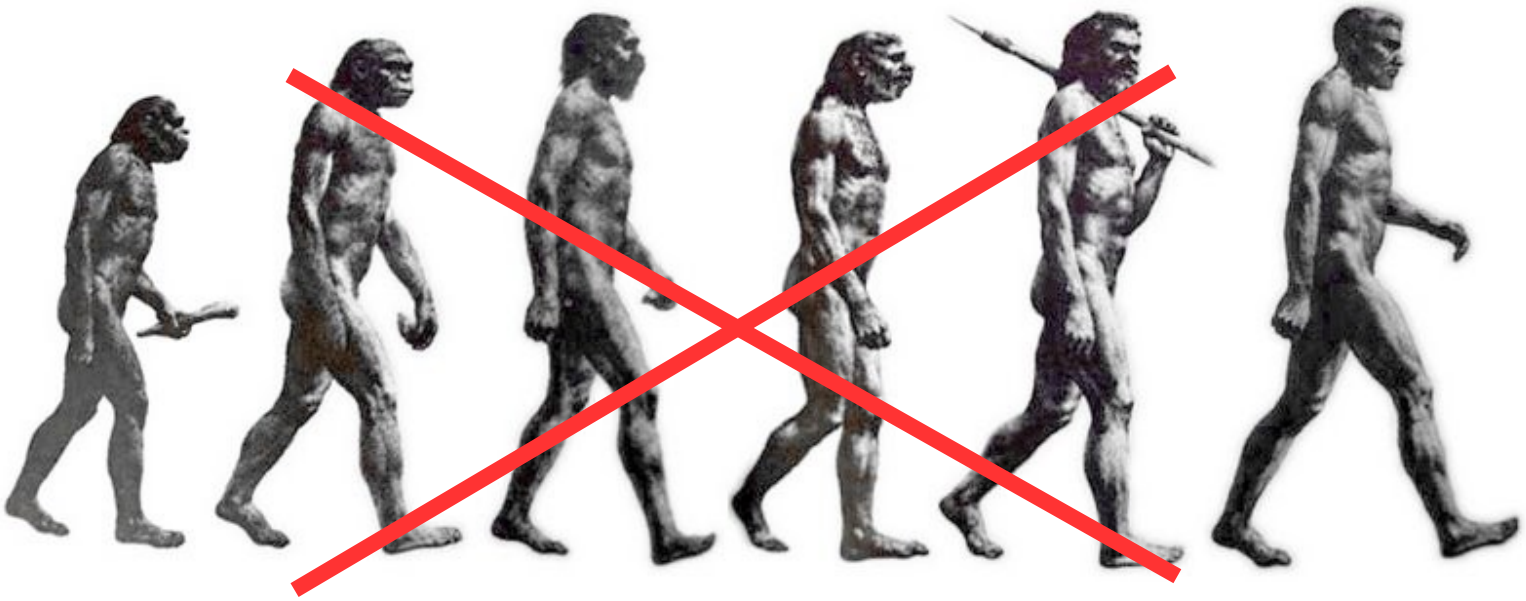
“Nothing in Biology Makes Sense Except in the Light of Evolution”

Theodosius Dobzhansky, 1973

Comparing things

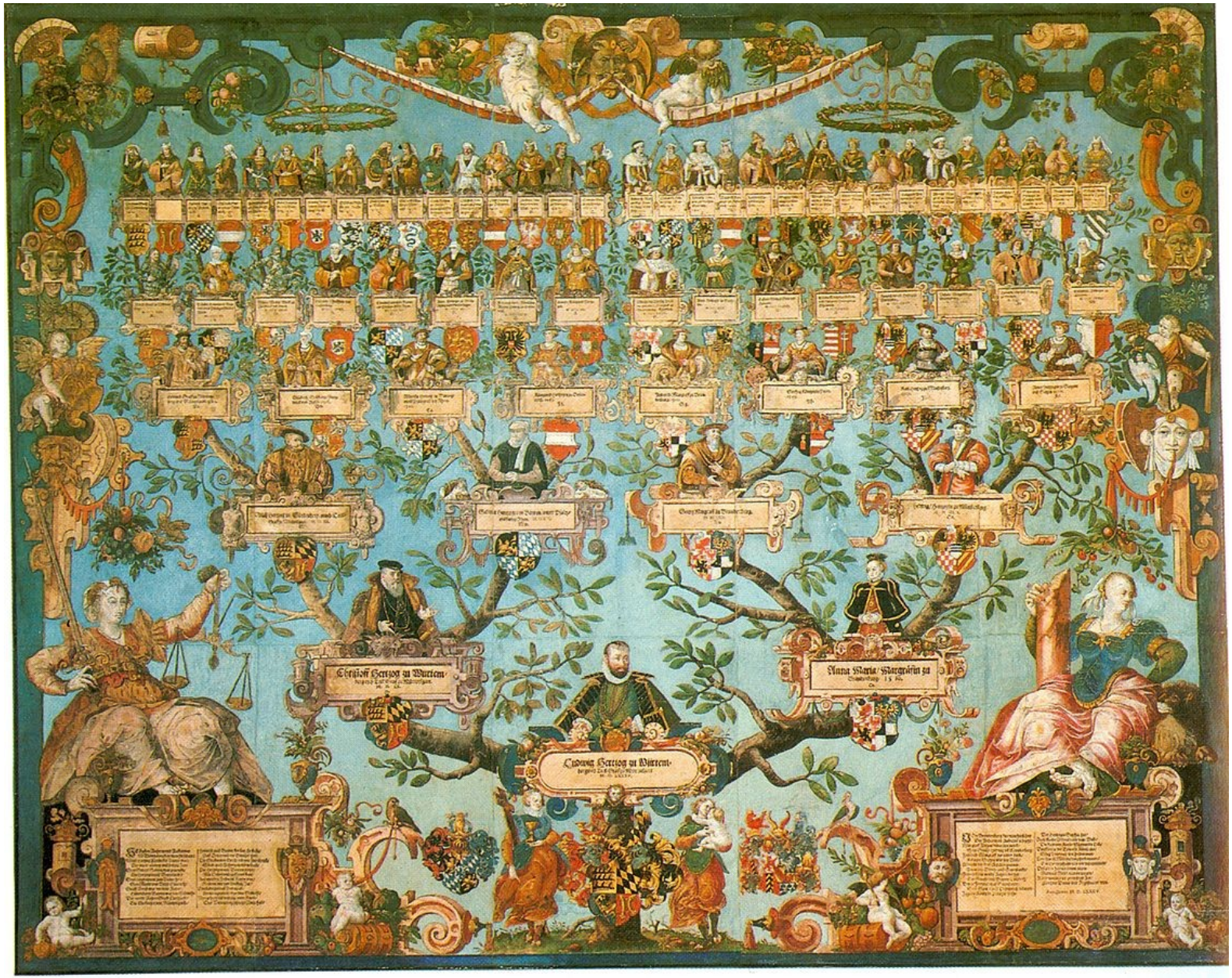


Evolution?



WRONG

Kingdom of Wurttemberg



Trees

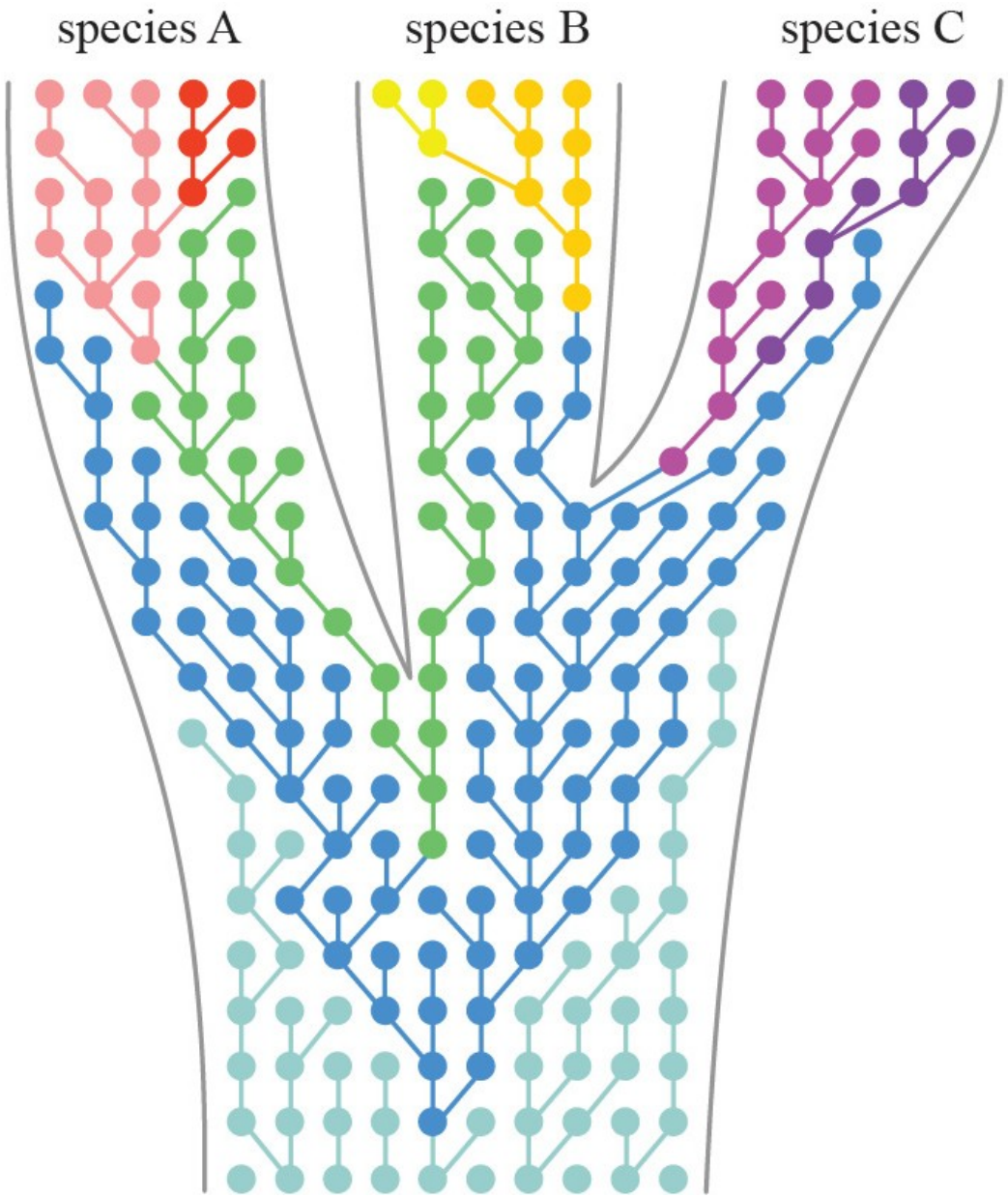
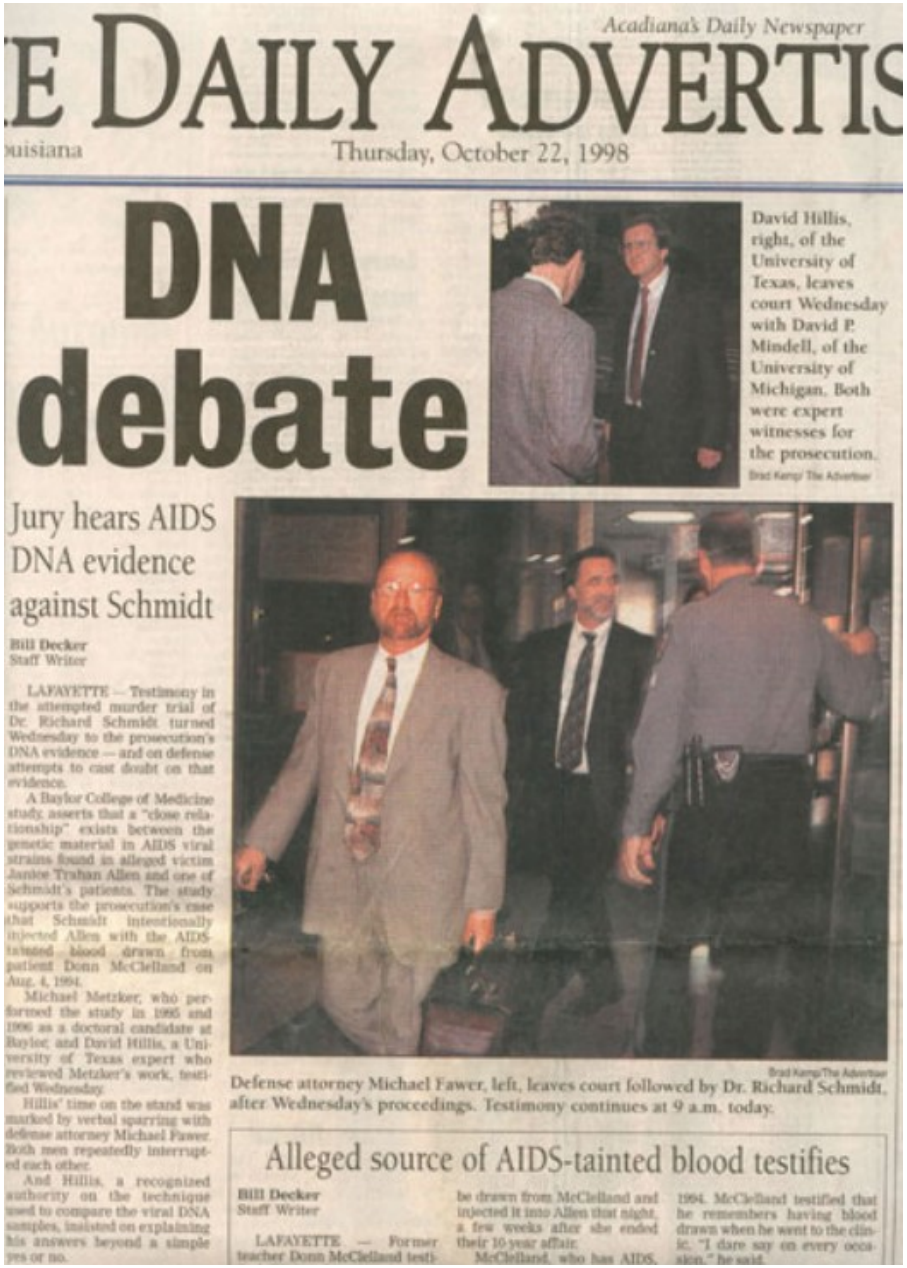


Figure author: Fredrick Leliaert

Tracing HIV infections



Acadiana's Daily Newspaper

E DAILY ADVERTIS

uisiana — Thursday, October 22, 1998

DNA debate

David Hillis, right, of the University of Texas, leaves court Wednesday with David P. Mindell, of the University of Michigan. Both were expert witnesses for the prosecution.

Brad Kemp/ The Advertiser

Jury hears AIDS DNA evidence against Schmidt

Bill Decker
Staff Writer

LAFAYETTE — Testimony in the attempted murder trial of Dr. Richard Schmidt turned Wednesday to the prosecution's DNA evidence — and on defense attempts to cast doubt on that evidence.

A Baylor College of Medicine study asserts that a "close relationship" exists between the genetic material in AIDS viral strains found in alleged victim Janice Trahan Allen and one of Schmidt's patients. The study supports the prosecution's case that Schmidt intentionally injected Allen with the AIDS-tainted blood drawn from patient Donn McClelland on Aug. 4, 1994.

Michael Metzker, who performed the study in 1995 and 1996 as a doctoral candidate at Baylor, and David Hillis, a University of Texas expert who reviewed Metzker's work, testified Wednesday.

Hillis' time on the stand was marked by verbal sparring with defense attorney Michael Fawer. Both men repeatedly interrupted each other.

And Hillis, a recognized authority on the technique used to compare the viral DNA samples, insisted on explaining his answers beyond a simple yes or no.

Brad Kemp/The Advertiser

Defense attorney Michael Fawer, left, leaves court followed by Dr. Richard Schmidt, after Wednesday's proceedings. Testimony continues at 9 a.m. today.

Alleged source of AIDS-tainted blood testifies

Bill Decker
Staff Writer

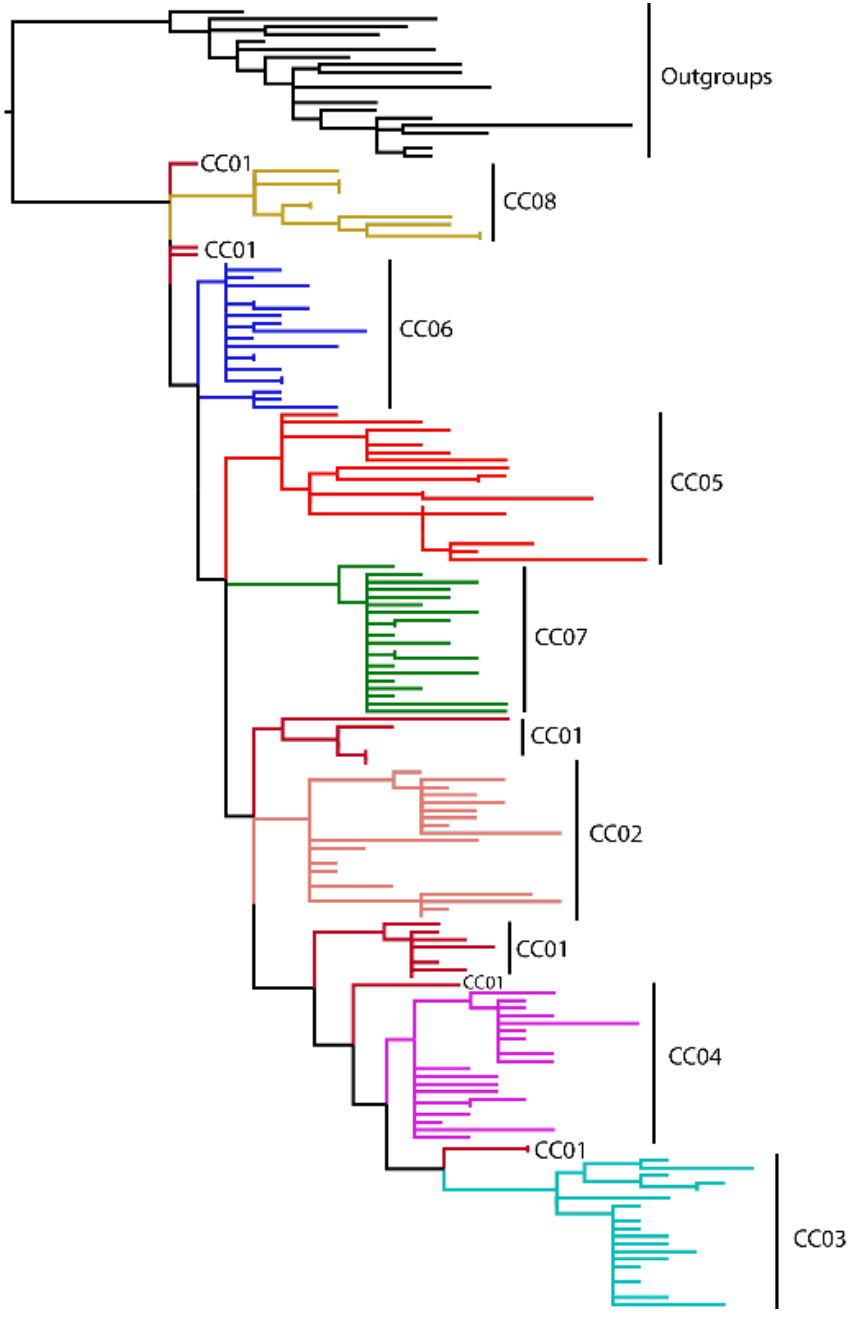
LAFAYETTE — Former teacher Donn McClelland testi-

be drawn from McClelland and injected it into Allen that night, a few weeks after she ended their 10-year affair.

McClelland, who has AIDS,

1994. McClelland testified that he remembers having blood drawn when he went to the clinic. "I dare say on every occasion," he said.

Tracing HIV infections



Tracing HIV infections

Acadiana's Daily Newspaper

E DAILY ADVERTIS

ıisiana Saturday, October 24, 1998

Schmidt guilty

Doctor faces 50-year jail sentence

Bill Decker
Staff Writer

LAFAYETTE — A Lafayette Parish jury found Dr. Richard Schmidt guilty late Friday of attempted second-degree murder, closing a bizarre and precedent-setting trial.

Schmidt, 50, was accused of intentionally injecting the AIDS virus into nurse Janice Trahan Allen of Lafayette in 1984, after she broke off their 10-year affair. Allen is now HIV-positive.

Waiting for the verdict just after 8:40 p.m. Friday, Allen, 36, clutched the hand of her husband, Jerry Allen. When the clerk read the verdict, Allen began sobbing loudly.

Prosecutor Keith Stutes

Schmidt stood as the verdict was read, showing no emotion. But as the discussion turned to whether his $500,000 bond would be continued pending sentencing, Schmidt sat and put his head in his hand. Moments later, he turned to embrace his wife, Barbara, who sobbed loudly.

Allen later fainted and was taken from the courtroom in a wheelchair.

"I didn't think this is what the jury should do," defense attorney Michael Fawer said after the trial. "I still don't. I thought we made it real clear that this case was fraught with reasonable doubt."

Fawer said he plans to appeal.

The nine-woman, three-man jury voted 10-2 to convict Schmidt after about four hours of deliberation. The jury rejected the lesser-

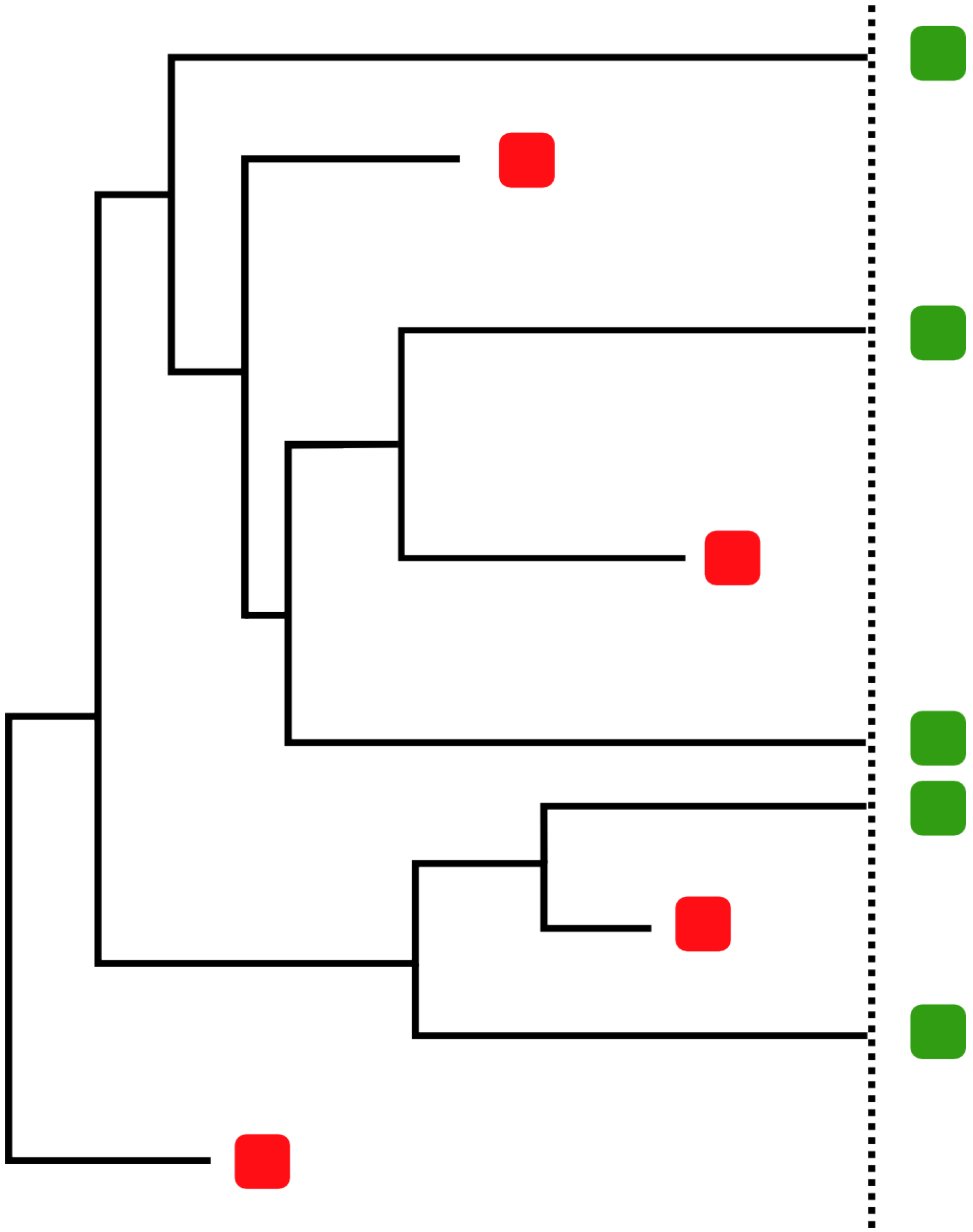
PC Paine/The Advertiser

Dr. Richard Schmidt, center, leaves the Lafayette Parish Courthouse Friday night with his despondent wife, Barbara, after being convicted of attempted second-degree murder. Schmidt is accompanied by courthouse security and defense attorney Gerald Block, left.

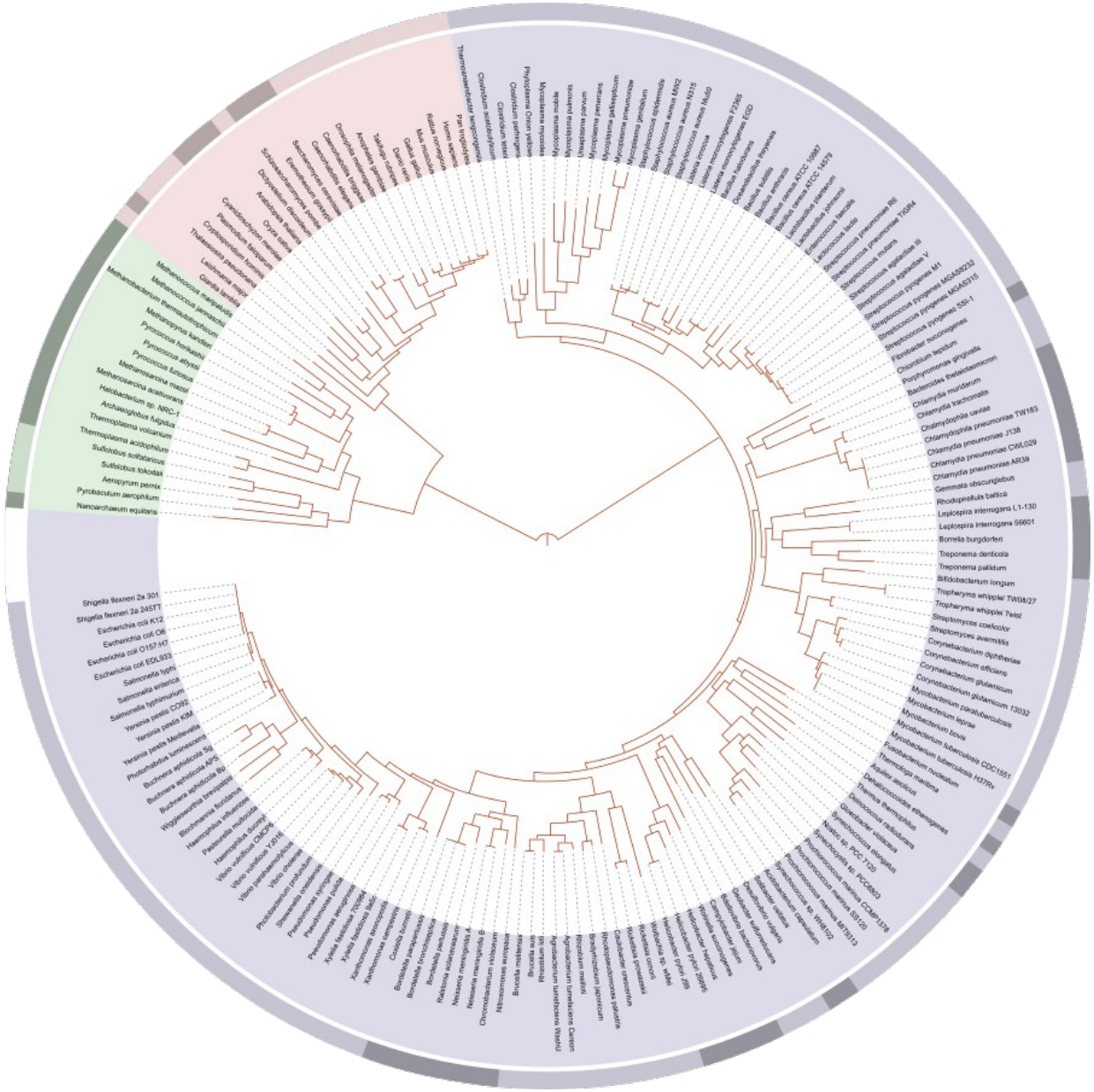
Claps, sobs, fainting spell greet verdict

Laura Ayo
Staff Writer

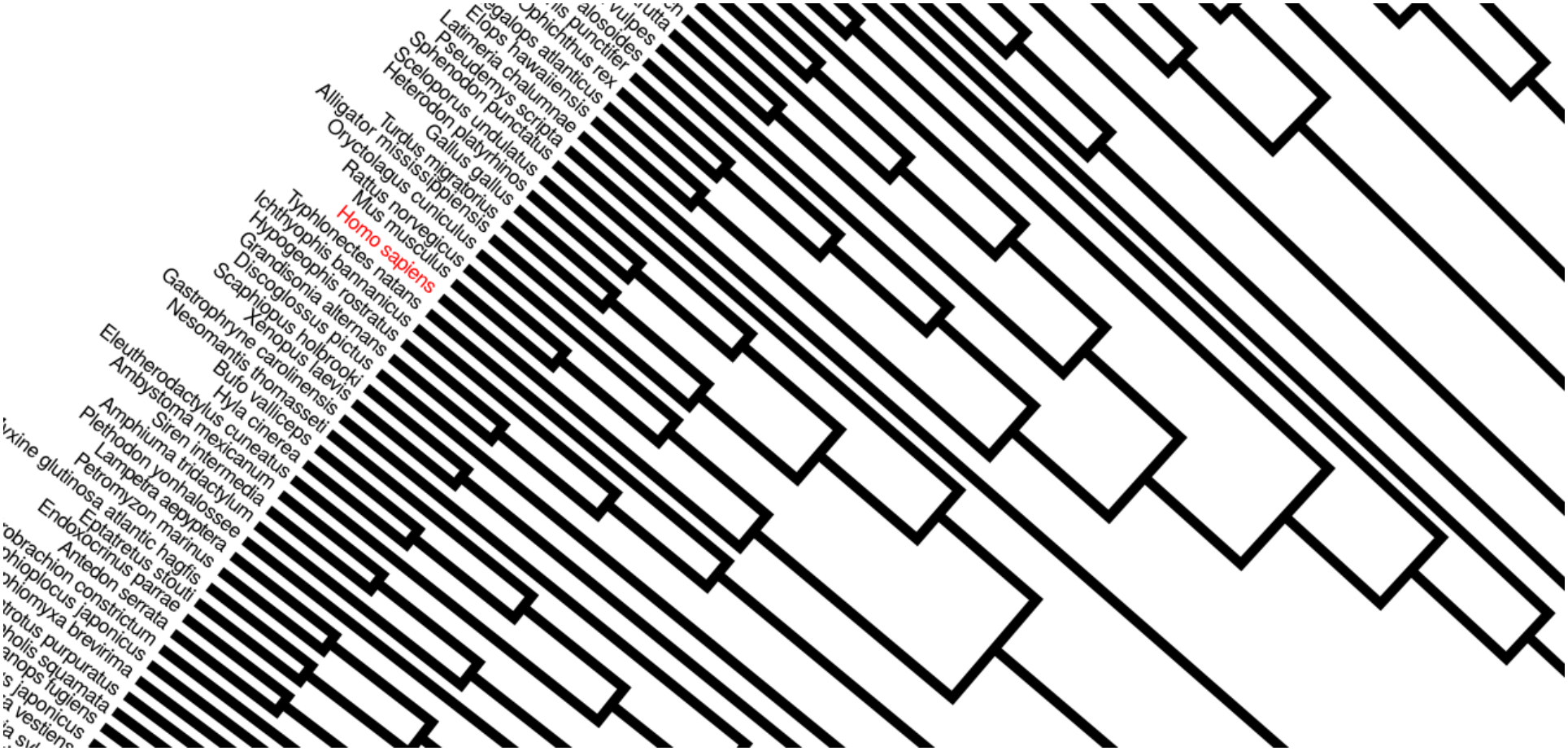
Extinction



Tree of life


Thermosediminibacter leitringianus
Clostridium acetobutylicum
Clostridium perfringens
Clostridium tetani
Physcomitrella patens
Mycoplasma mobile
Mycoplasma pneumoniae
Mycoplasma genitalium
Mycoplasma pulmonis
Ureaplasma parvum
Mycoplasma penetrans
Mycoplasma gallisepticum
Mycoplasma pneumoniae
Mycoplasma genitalium
Staphylococcus epidermidis
Staphylococcus aureus MW2
Staphylococcus aureus N315
Staphylococcus aureus Mu50
Listeria innocua
Listeria monocytogenes F2365
Listeria monocytogenes EGD
Oceanobacillus iheyensis
Bacillus halodurans
Bacillus subtilis
Bacillus anthracis
Bacillus cereus ATCC 10987
Bacillus cereus ATCC 14579
Lactobacillus plantarum
Lactobacillus johnsonii
Enterococcus faecalis
Lactococcus lactis
Streptococcus pneumoniae R6
Streptococcus pneumoniae TIGR4
Streptococcus mutans
Streptococcus agalactiae III
Streptococcus agalactiae V
Streptococcus pyogenes M1
Streptococcus pyogenes MGAS8232
Streptococcus pyogenes MGAS315
Streptococcus pyogenes SSI-1
Streptococcus sanguinis
Finegoldia magna
Clostridium tetani
Clostridium perfringens
Porphyromonas gingivalis
Bacteroides thetaiotaomicron
Chlamydia muridarum
Chlamydia trachomatis
Chlamydophila caviae
Chlamydophila pneumoniae TW183
Chlamydia pneumoniae J138
Chlamydia pneumoniae CWL029
Chlamydia pneumoniae AR39
Gemmata obscuriglobus
Rhodopirellula baltica
Leptospira interrogans L1-130
Leptospira interrogans 56601
Borrelia burgdorferi
Treponema denticola
Treponema pallidum
Bifidobacterium longum
Tropheryma whipplei TW08/27
Tropheryma whipplei Twist
Streptomyces coelicolor
Streptomyces avermitilis
Corynebacterium diphtheriae
Corynebacterium efficiens
Corynebacterium glutamicum
Corynebacterium glutamicum 13032
Mycobacterium paratuberculosis
Mycobacterium leprae
Mycobacterium bovis
Mycobacterium tuberculosis CDC1551
Mycobacterium tuberculosis H37Rv
Fusobacterium nucleatum
Thermotoga maritima
Aquifex aeolicus
Deinococcus radiodurans
Thermus thermophilus
Gloeobacter violaceus
Synechocystis sp. PCC6803
Nostoc sp. PCC 7120
Prochlorococcus marinus CCMP1378
Prochlorococcus marinus MIT9313
Prochlorococcus marinus SS120
Synechococcus sp. WH8102
Acidobacterium capsulatum
Solibacter usitatus
Desulfovibrio vulgaris
Desulfovibrio desulfuricans
Geobacter sulfurreducens
Geobacter metallireducens
Wolinella succinogenes
Campylobacter jejuni
Helicobacter hepaticus
Helicobacter pylori 26695
Helicobacter pylori J99
Wolbachia sp. wMel
Rickettsia prowazekii
Rickettsia conorii
Caulobacter crescentus
Rhodopseudomonas palustris
Bradyrhizobium japonicum
Rhizobium meliloti
Agrobacterium tumefaciens Cereon
Agrobacterium tumefaciens WashU
Brucella suis
Brucella melitensis
Nitrosomonas europaea
Chromobacterium violaceum
Neisseria meningitidis B
Neisseria meningitidis A
Ralstonia solanacearum
Bordetella pertussis
Bordetella bronchiseptica
Bordetella parapertussis
Clostridium botulinum
Xylella fastidiosa 9a5c
Xylella fastidiosa Temecula1
Xanthomonas campestris
Xanthomonas axonopodis
Pseudomonas syringae
Pseudomonas putida
Pseudomonas aeruginosa
Pseudomonas fluorescens
Shewanella oneidensis
Photobacterium profundum
Vibrio cholerae
Vibrio parahaemolyticus
Vibrio vulnificus YJ016
Vibrio vulnificus CMCP6
Haemophilus somnus
Pasteurella multocida
Haemophilus influenzae
Wigglesworthia brevipalpis
Blochmannia floridanus
Buchnera aphidicola Bp
Buchnera aphidicola APS
Photorhabdus luminescens
Yersinia pestis Medievalis
Yersinia pestis KIM
Yersinia pestis CO92
Salmonella typhimurium
Salmonella typhi
Salmonella enterica
Escherichia coli EDL933
Escherichia coli O157:H7
Escherichia coli O6
Escherichia coli K12
Shigella flexneri 2a 2457T
Shigella flexneri 2a 301
Nanoarchaeum equitans
Pyrobaculum aerophilum
Aeropyrum pernix
Sulfolobus tokodaii
Sulfolobus solfataricus
Thermoplasma acidophilum
Thermoplasma volcanium
Archaeoglobus fulgidus
Halobacterium sp. NRC-1
Haloarcula marismortui
Methanosarcina mazei
Pyrococcus abyssi
Pyrococcus furiosus
Pyrococcus horikoshii
Methanopyrus kandleri
Methanobacterium thermautotrophicum
Methanococcus jannaschii
Methanococcus maripaludis
Leishmania major
Giardia lamblia
Thalassiosira pseudonana
Cryptosporidium hominis
Plasmodium falciparum
Cryptosporidium parvum
Oryza sativa
Arabidopsis thaliana
Dictyostelium discoideum
Schizosaccharomyces pombe
Saccharomyces cerevisiae
Encephalitozoon cuniculi
Caenorhabditis elegans
Drosophila melanogaster
Anopheles gambiae
Takifugu rubripes
Danio rerio
Gallus gallus
Mus musculus
Rattus norvegicus
Homo sapiens
Pan troglodytes

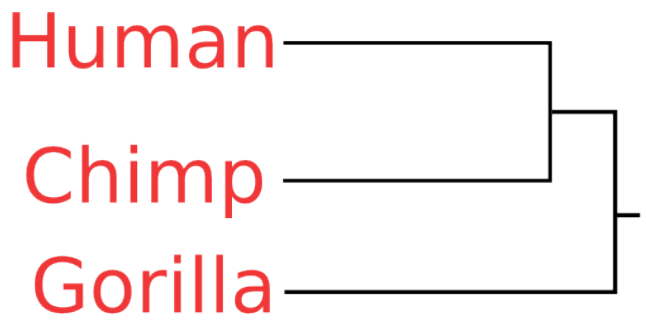
Tree of life

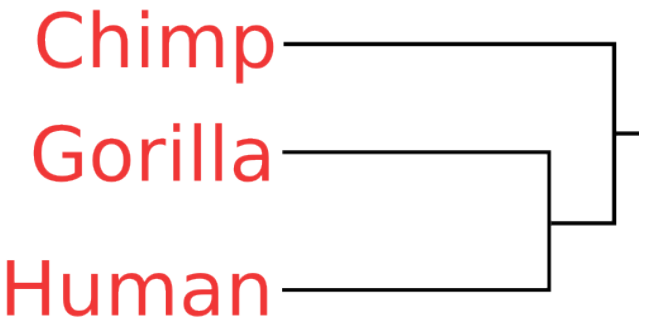


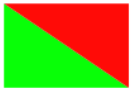
Mobiles

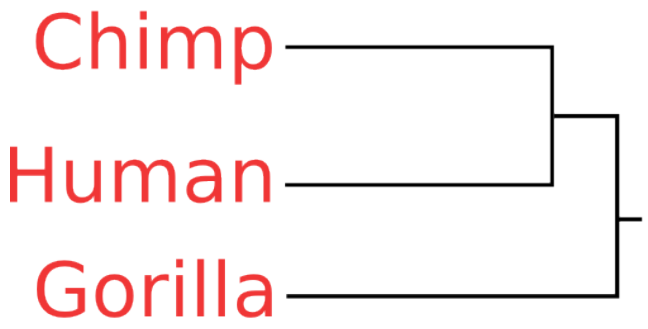
Mobiles

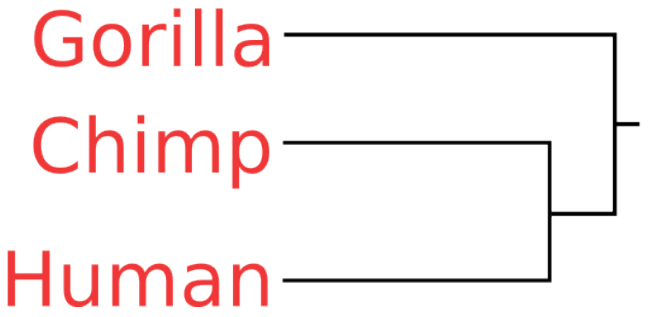
- There is no linear order
- Only branchings matter





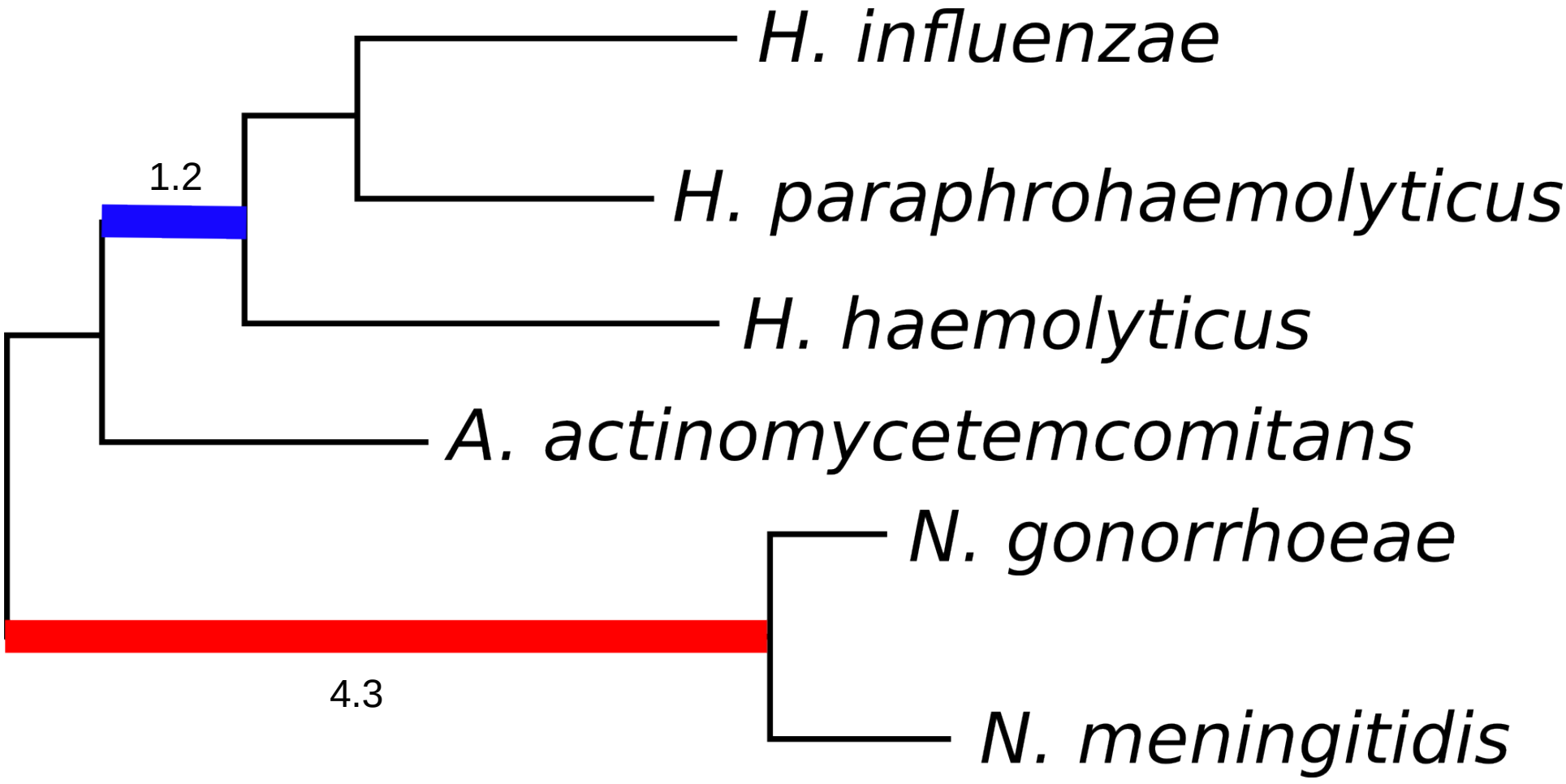




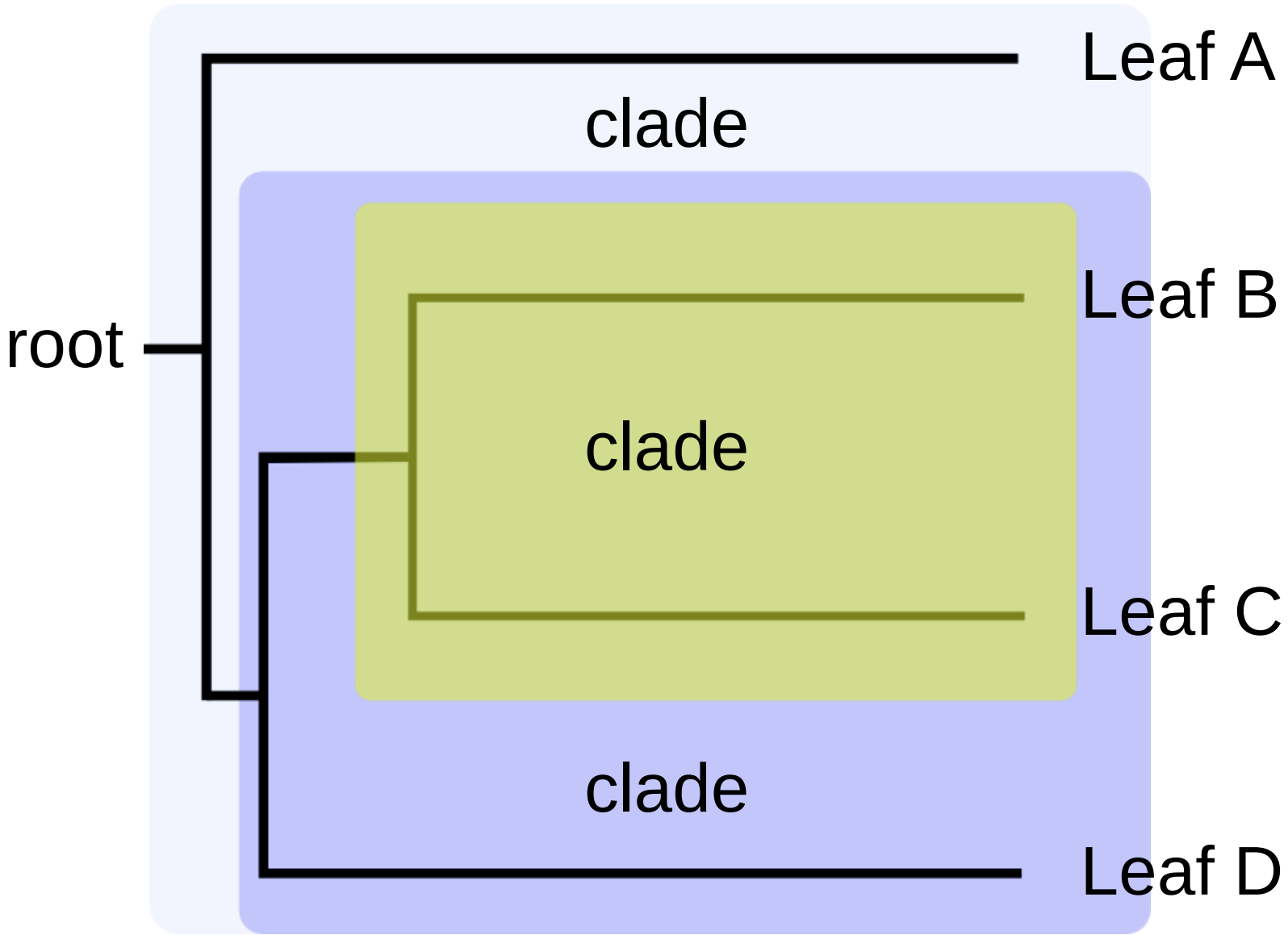


Branch length

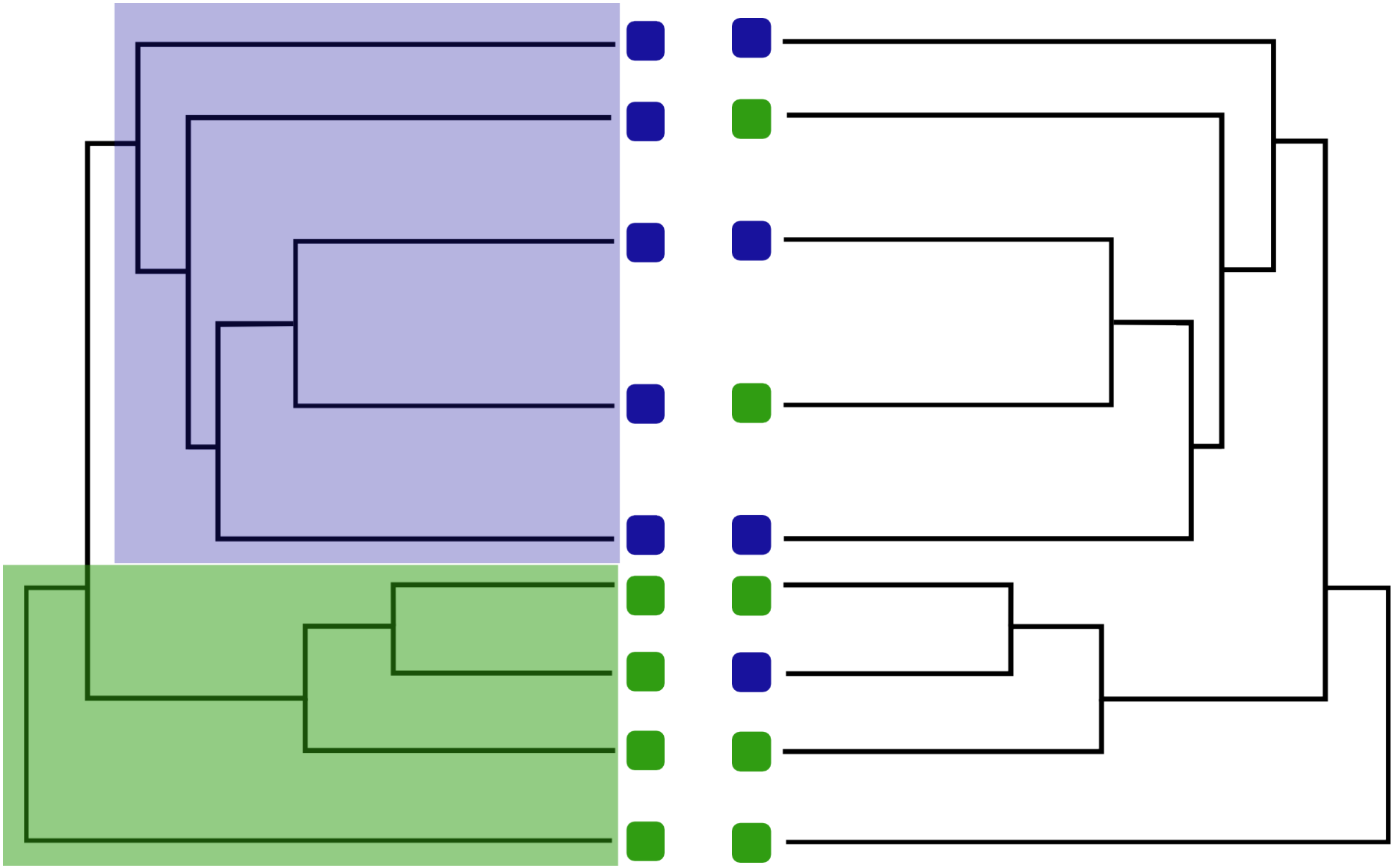
Distance = time * rate of substitutions



Jargon



Grouping



How many trees are there?

$$N_r = \prod_{k=2}^{T} (2k-3)$$

$$N_u = \prod_{k=3}^{T} (2k-5)$$

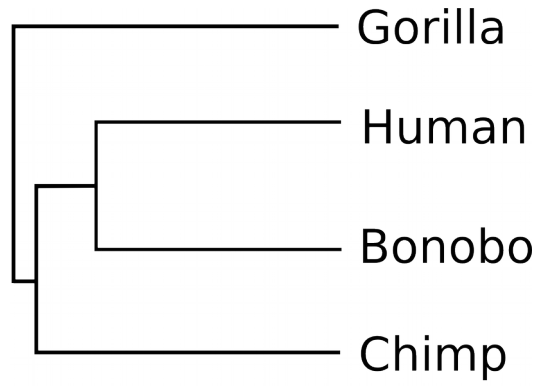
T	Unrooted	Rooted
10	2,027,025	3×10^{7}
15	7×10^{12}	2×10^{8}
20	2×10^{20}	8×10^{21}
100	2×10^{182}	2×10^{184}

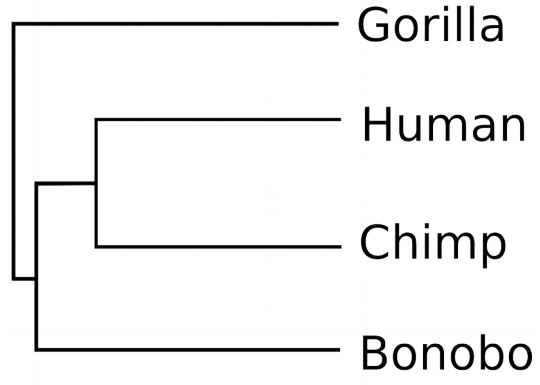
Ways to build a tree

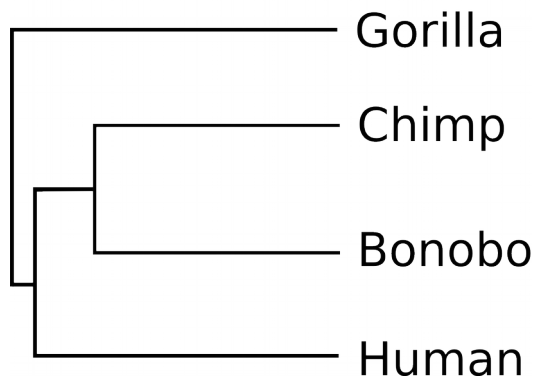
- Parsimony
  - Minimize number of mutations
  - Simple, but inaccurate
- Maximum Likelihood (most common)
  - Maximize probability of observation given model
  - Harder, but accurate
  - Does not provide confidence intervals
- Bayesian inference
  - Choose best model that maximizes probability given data
  - Hardest, very accurate
  - Explores all reasonable hypotheses

Parsimony

H: ACTGCATG
C: ACGGCCTC
G: CCGGCATG
B: ACTGCCTC







Parsimony

```
H: ACTGCATG
C: ACGGCCTC
G: CCGGCATG
B: ACTGCCTC
```

1

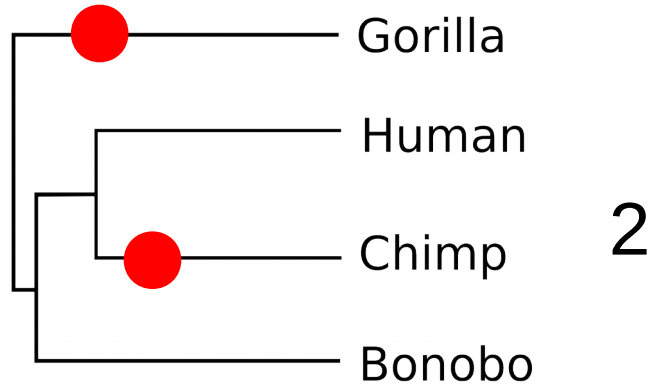
1

1

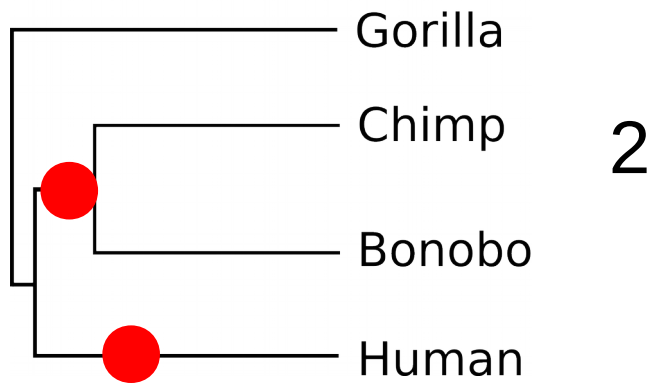
Parsimony

```
H: ACTGCATG
C: ACGGCCTC
G: CCGGCATG
B: ACTGCCTC
```

1



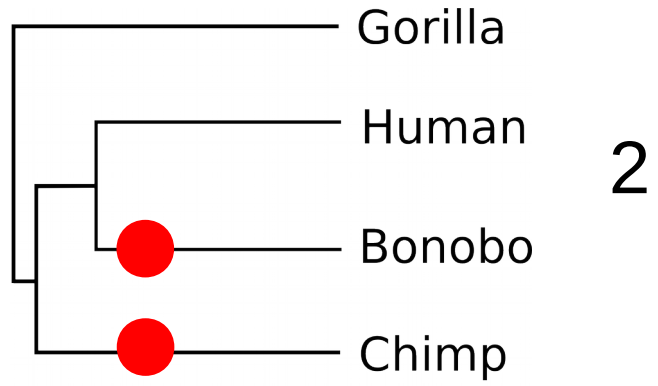
2



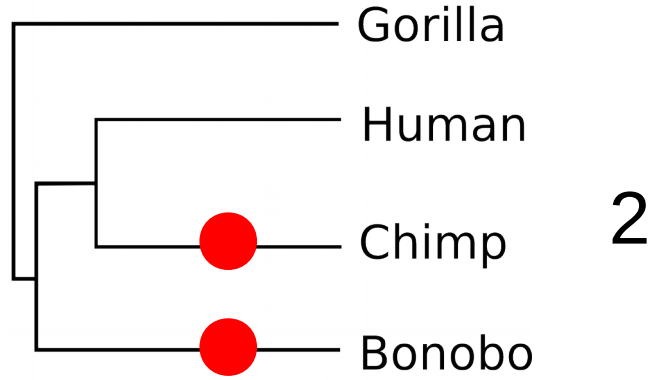
2

Parsimony

```
H: ACTGCATG
C: ACGGCCTC
G: CCGGCATG
B: ACTGCCTC
```



2

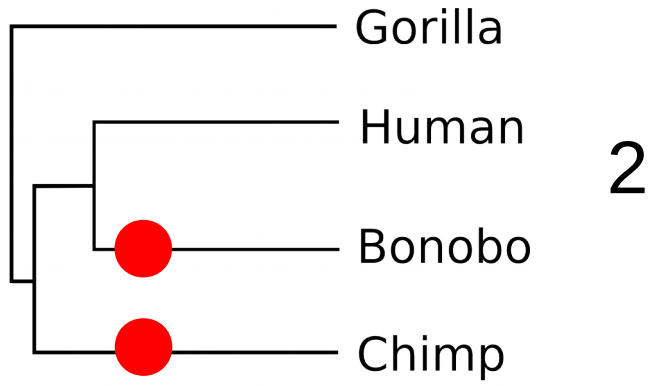


2

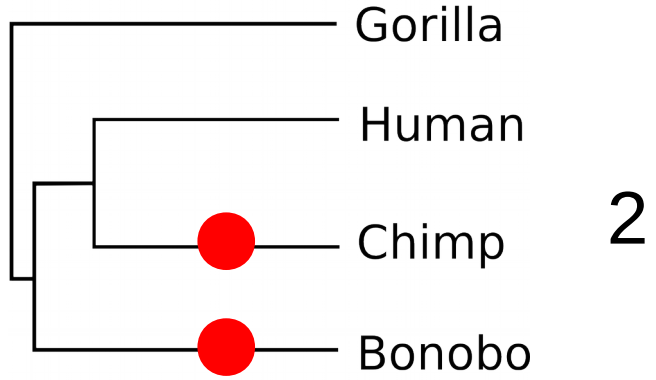
1

Parsimony

H: ACTGCATG
C: ACGGCCTC
G: CCGGCATG
B: ACTGCCTC



2

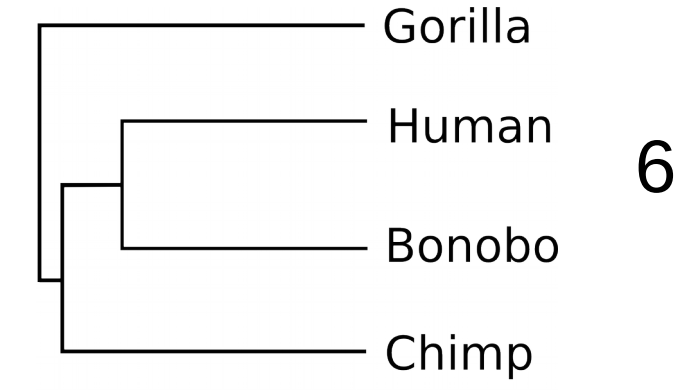


2

1

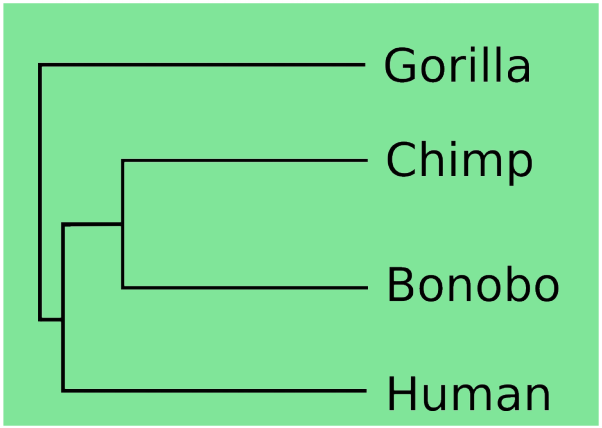
Parsimony

```
H: ACTGCATG
C: ACGGCCTC
G: CCGGCATG
B: ACTGCCTC
```



6

7



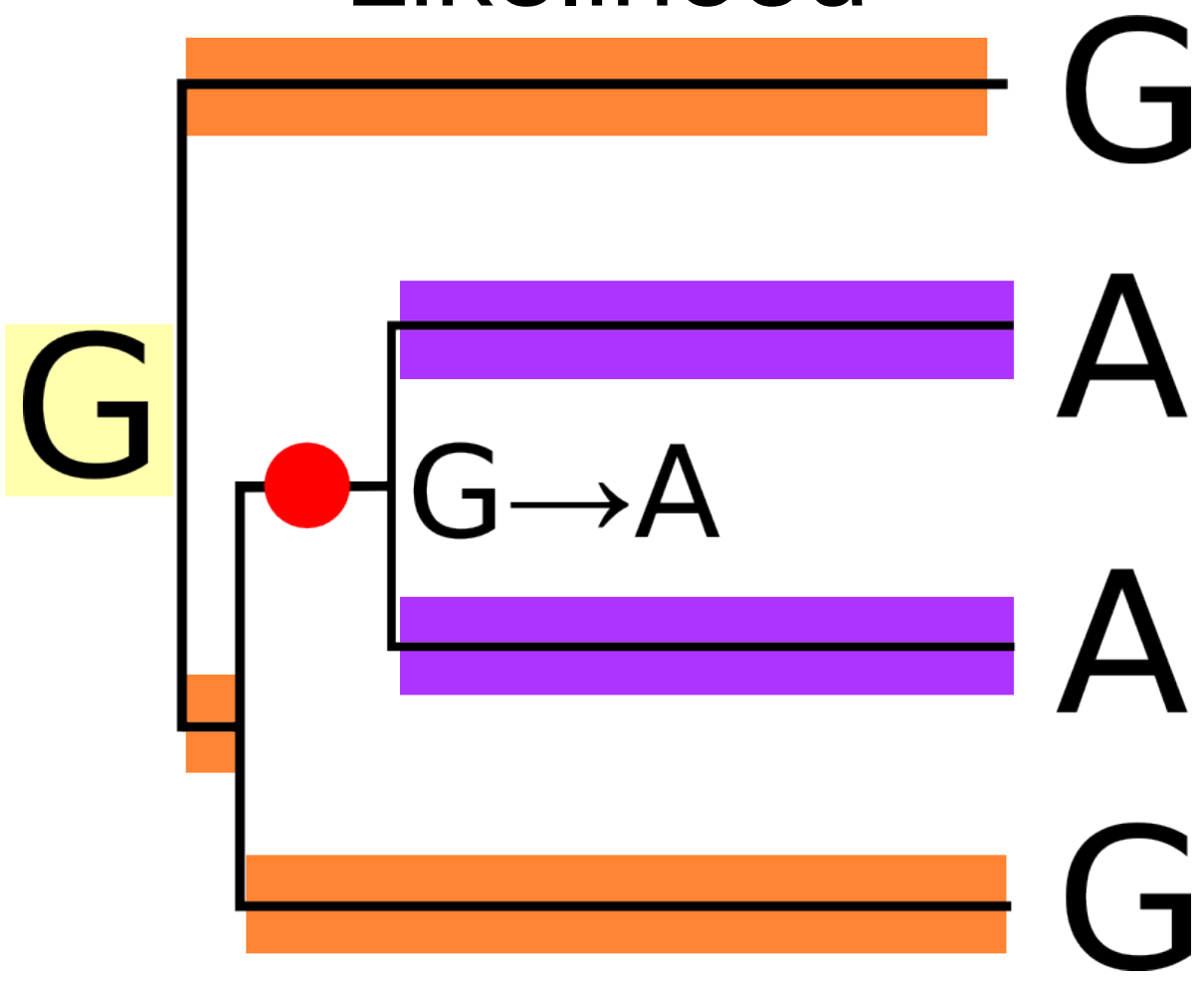
5

Likelihood

- Calculate the probability of each event according to a model
- Need to do this for all substitution histories
- Kimura 2 parameter model
  - Transition rate *a*
  - Transversion rate *b*

	A	C	T	G
A		b	a	b
C	b		b	a
T	a	b		b
G	b	a	b	

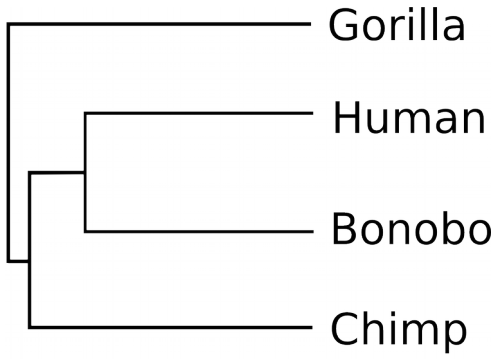
Likelihood

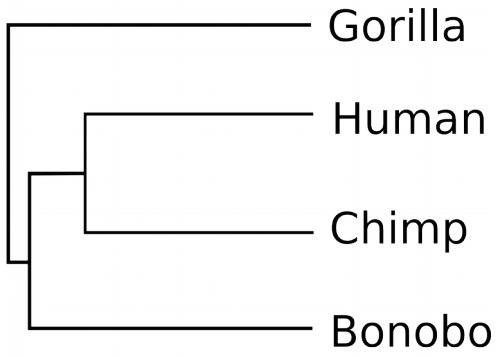


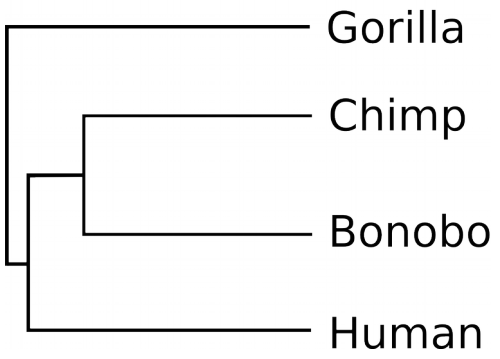
$$P(T) = P(G) \times P(G \to G) \times P(G \to G) \times P(G \to G) \times P(G \to A) \times P(A \to A) \times P(A \to A)$$

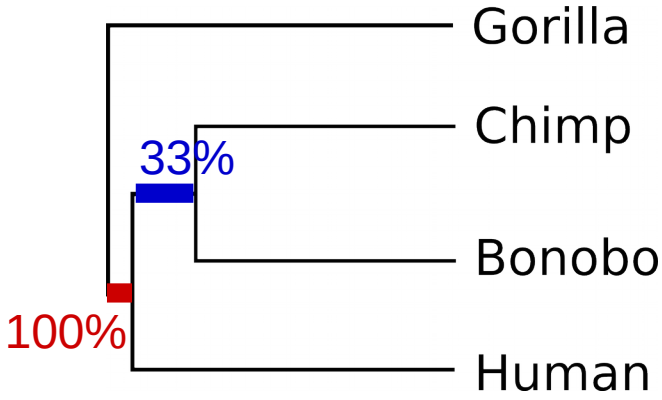
Bootstrap

```
H: ACGCGTAT
C: ACGCCGCT
G: CCGCGGAT
B: ACGCCTCT

H: ATGGCTCA
C: AGCGCTCC
G: CGGGCTCA
B: ATCGCTCC

H: ACTGCATG
C: ACGGCCTC
G: CCGGCATG
B: ACTGCCTC
```









References

- March of progress http://en.wikipedia.org/wiki/March_of_Progress
- Genealogy of Ludwig Herzog von Wurttemberg http://commons.wikimedia.org/wiki/File:900-158_Ahnentafel_Herzog_Ludwig.jpg#mediaviewer/File:900-158_Ahnentafel_Herzog_Ludwig.jpg
- Mobile ad: http://weedecor.com/wp-content/uploads/2013/05/safari_baby_mobile.jpg
- Phylogenetic trees on wikipedia: http://en.wikipedia.org/wiki/Phylogenetic_tree
- Fredrick Leliaert on species trees: http://frederikleliaert.wordpress.com/green-algae/dna-based-species-delimitation-in-algae/
- MAFFT: http://mbe.oxfordjournals.org/content/30/4/772
- Phylogeny.fr: http://www.ncbi.nlm.nih.gov/pubmed/18424797