

Phylogenetics

December 2023



Gustav Klimt 1905 "Tree of Life"



Notebook around July 1837

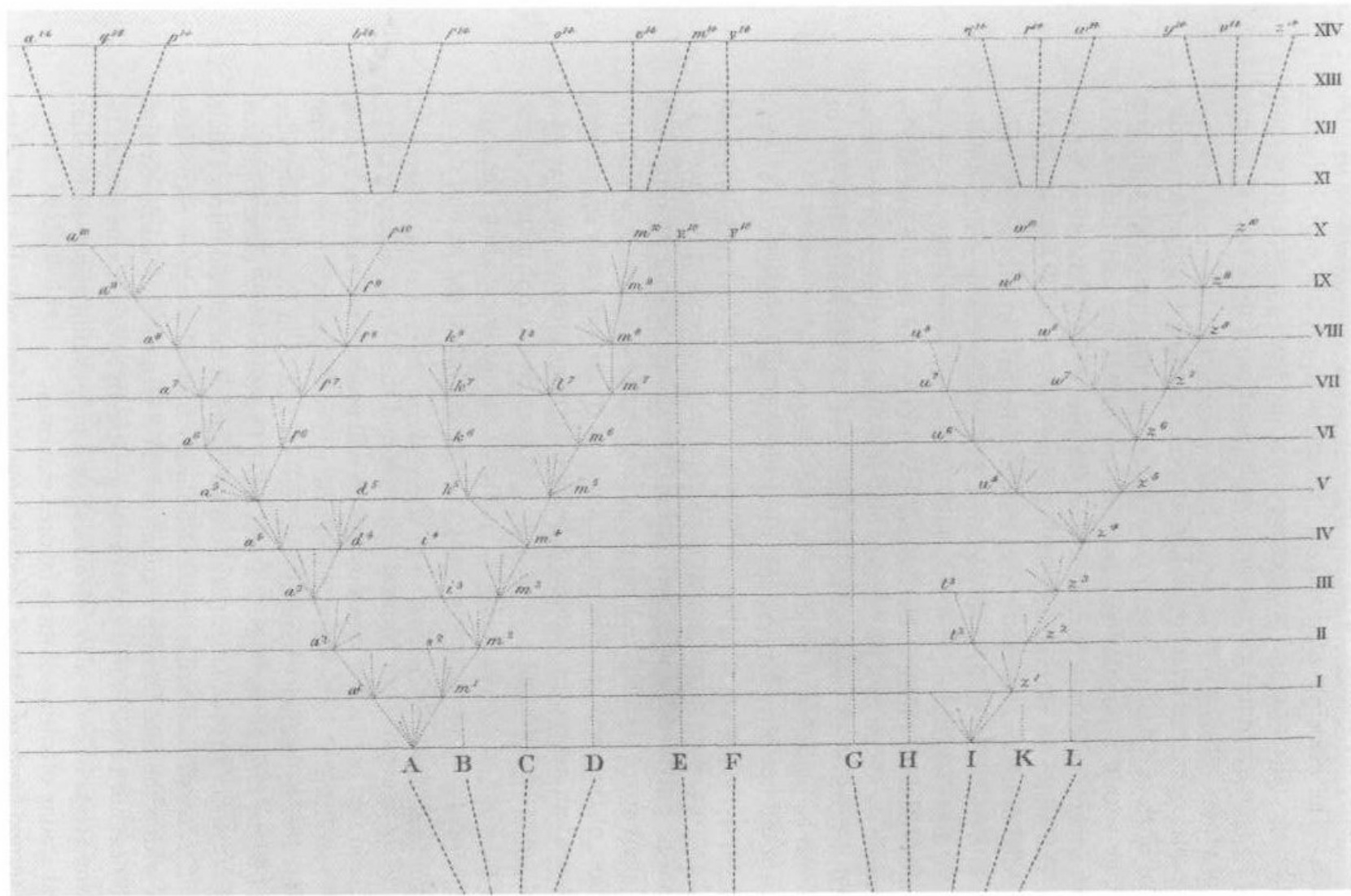
36

I think

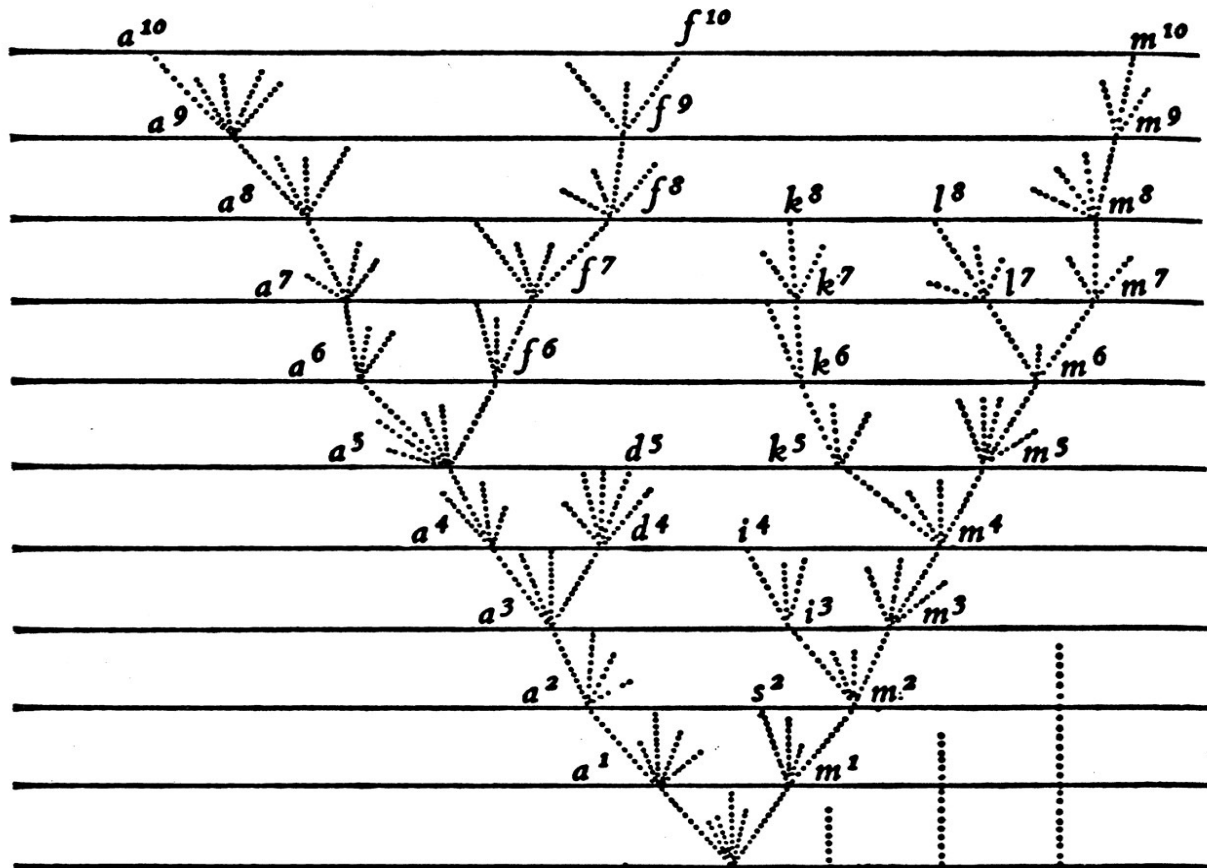
Can never be shown in
 form living or now
 as they are now
 Do not think it is
 as the same (as is) - species
 as the same.

Then between A & B. various
 sort of relation. C & B. The
 first gradation, B & D
 rather greater distinction
 Then genus would be
 formed. - bearing relation





Detail from only figure in origin of species



Darwin 1859

TABLEAU

Servant à montrer l'origine des différens animaux.

Vers.

Infusoires.
Polypes.
Radiaires.

Annelides.
Cirripèdes.
Mollusques.

Insectes.
Arachnides.
Crustacés.

Poissons.
Reptiles.

Oiseaux.

Monotrèmes.

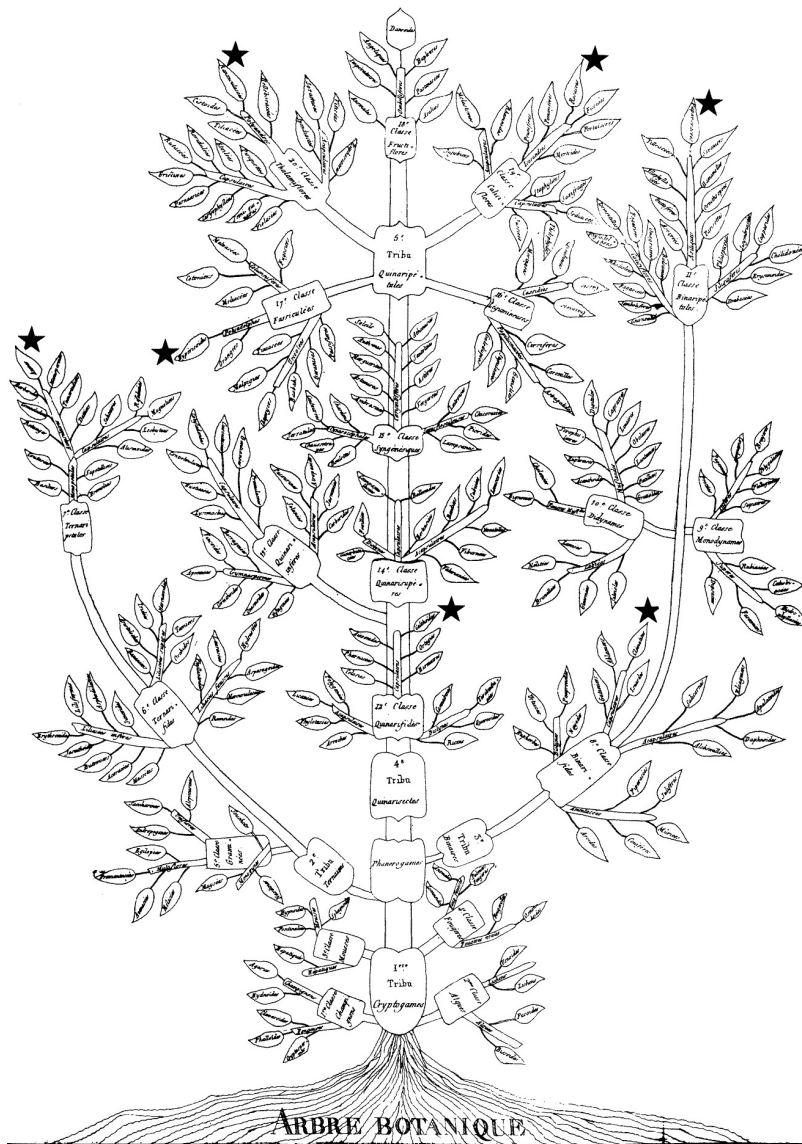
M. Amphibies.

M. Cétacés.

M. Ongulés.

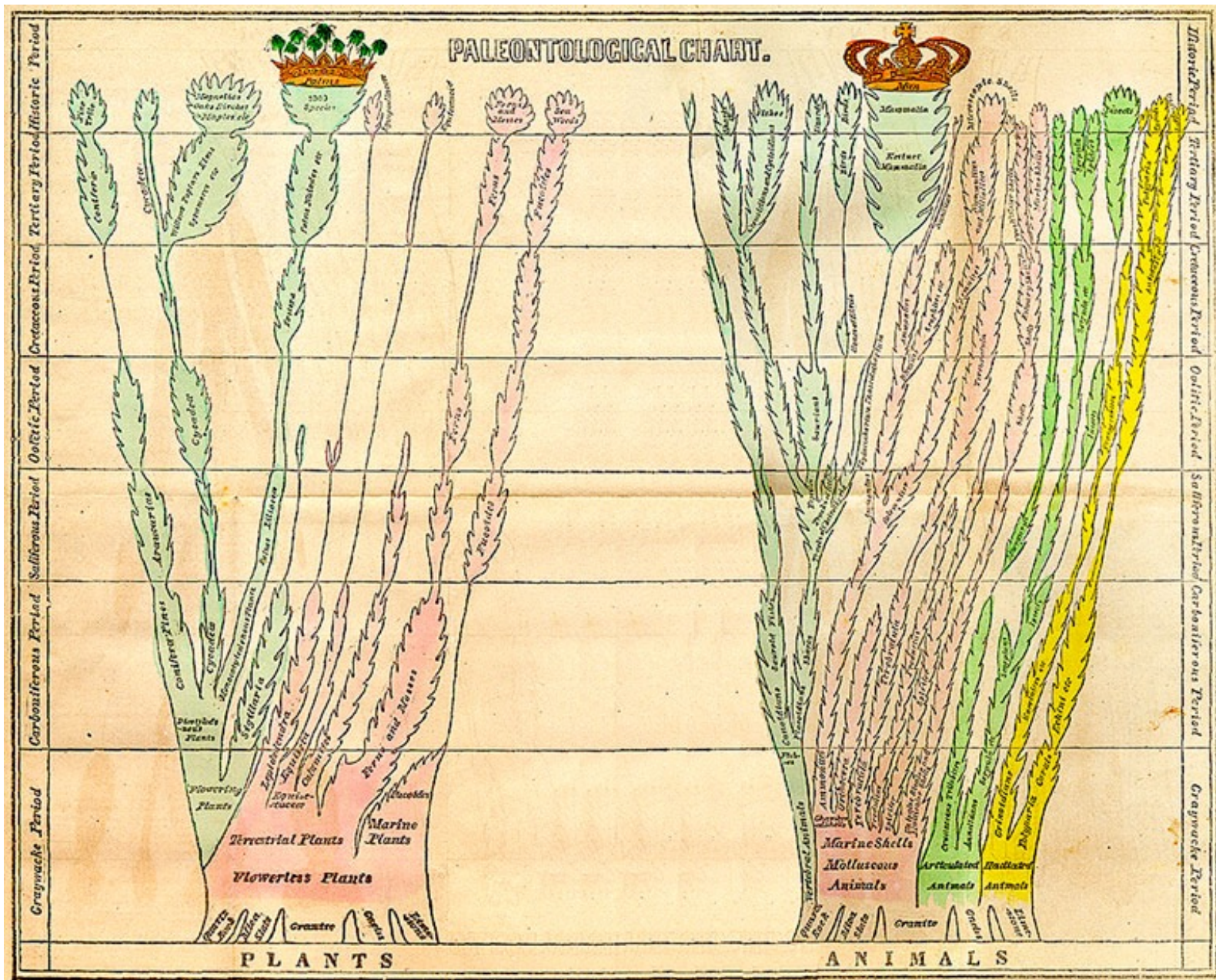
M. Onguiculés.

Cette série d'animaux commençant par deux



Augustin Augier 1801

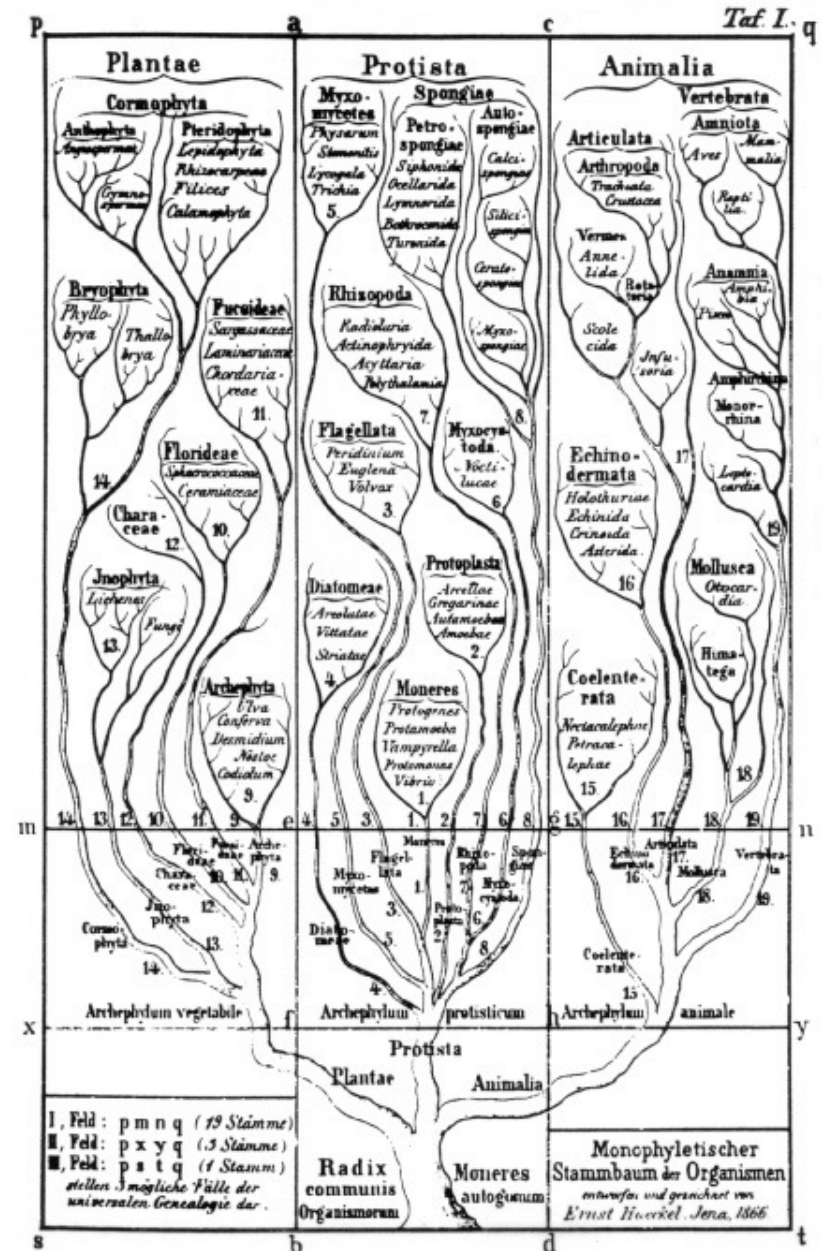
Jean-Baptiste Lamarck 1809





Ernst Haeckel MD.PhD. (1834-1919)

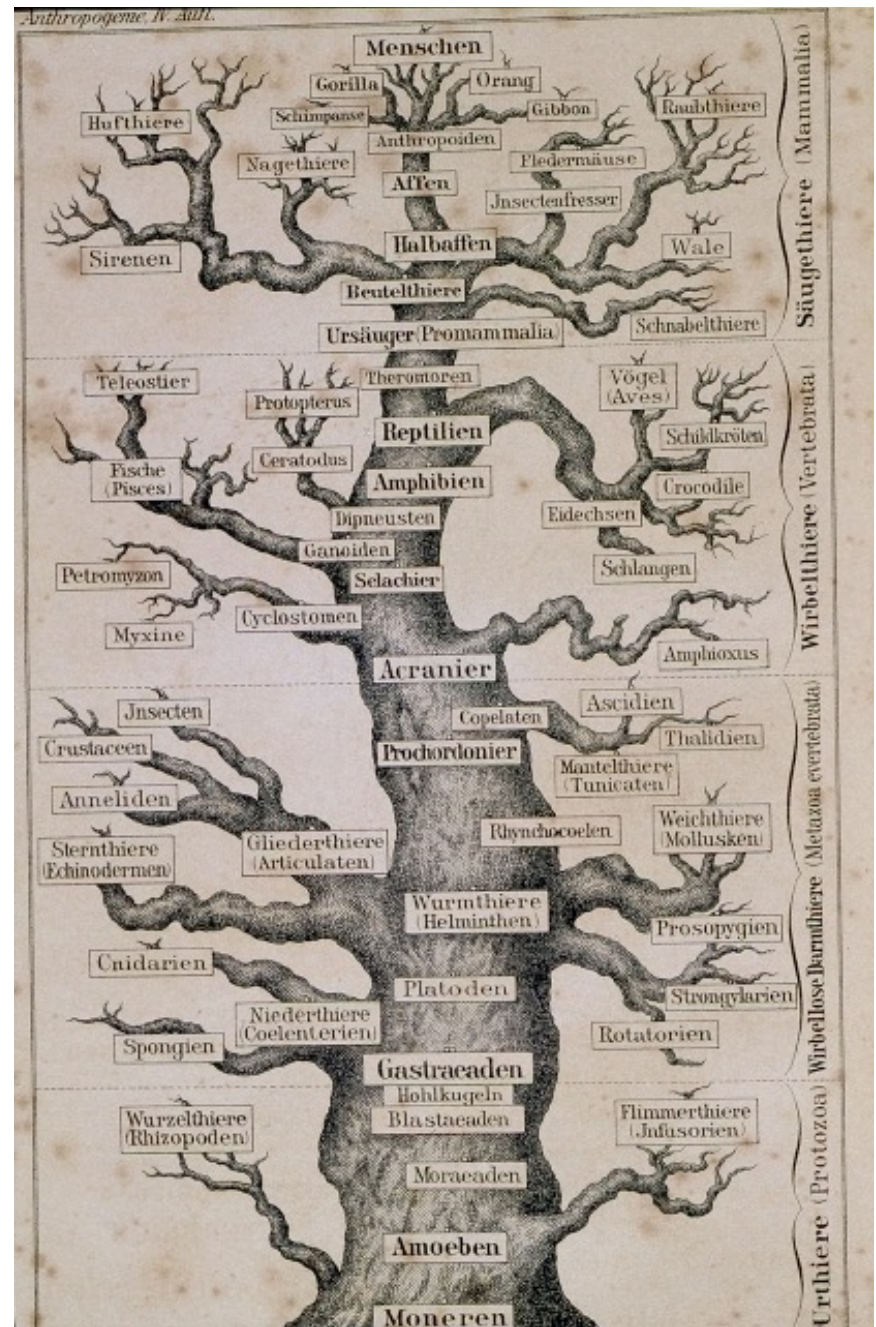
- “ontogeny recapitulates phylogeny”
- coined words “protista”, “phylum”, “ecology”, “phylogeny”
- TREE OF LIFE

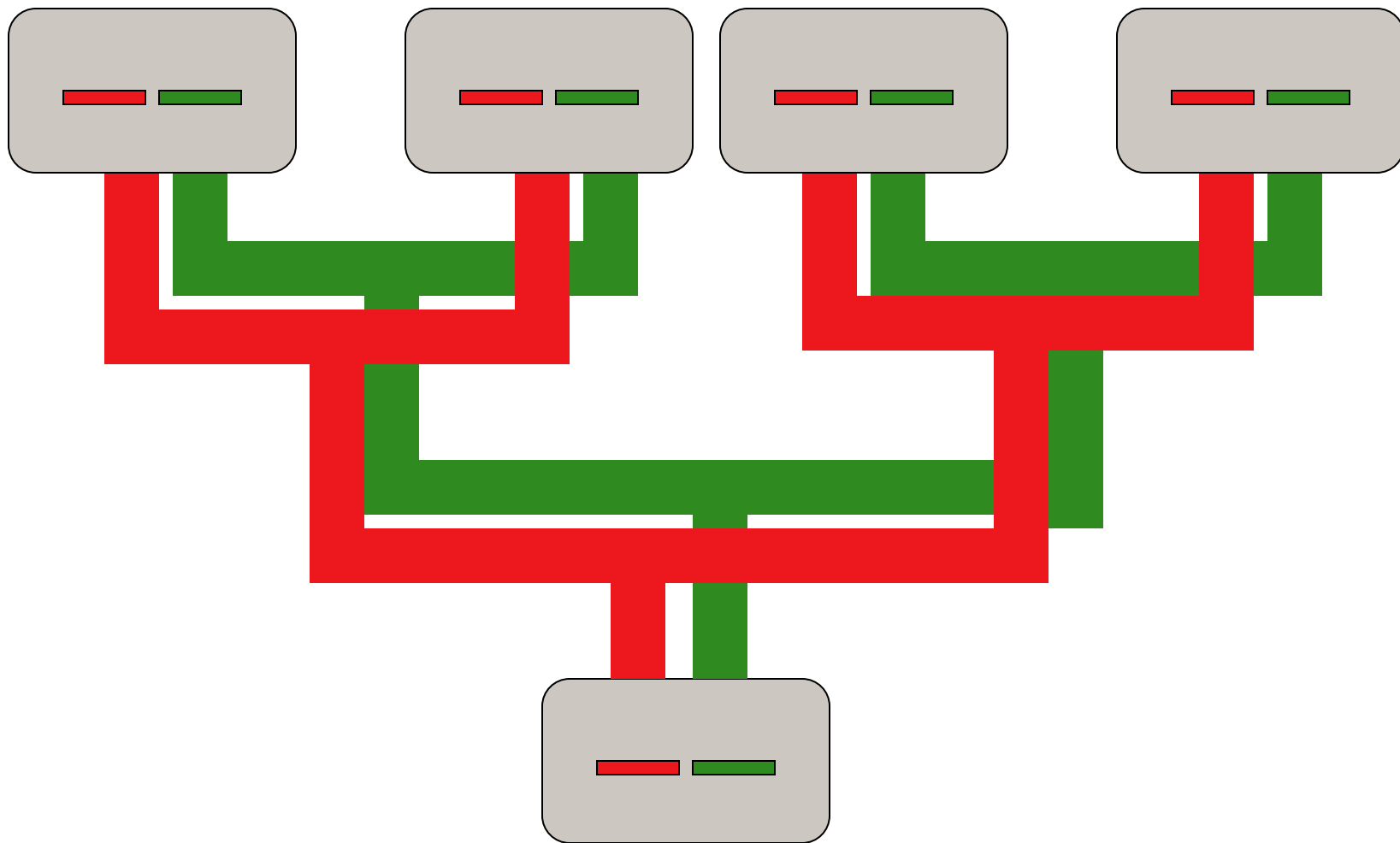


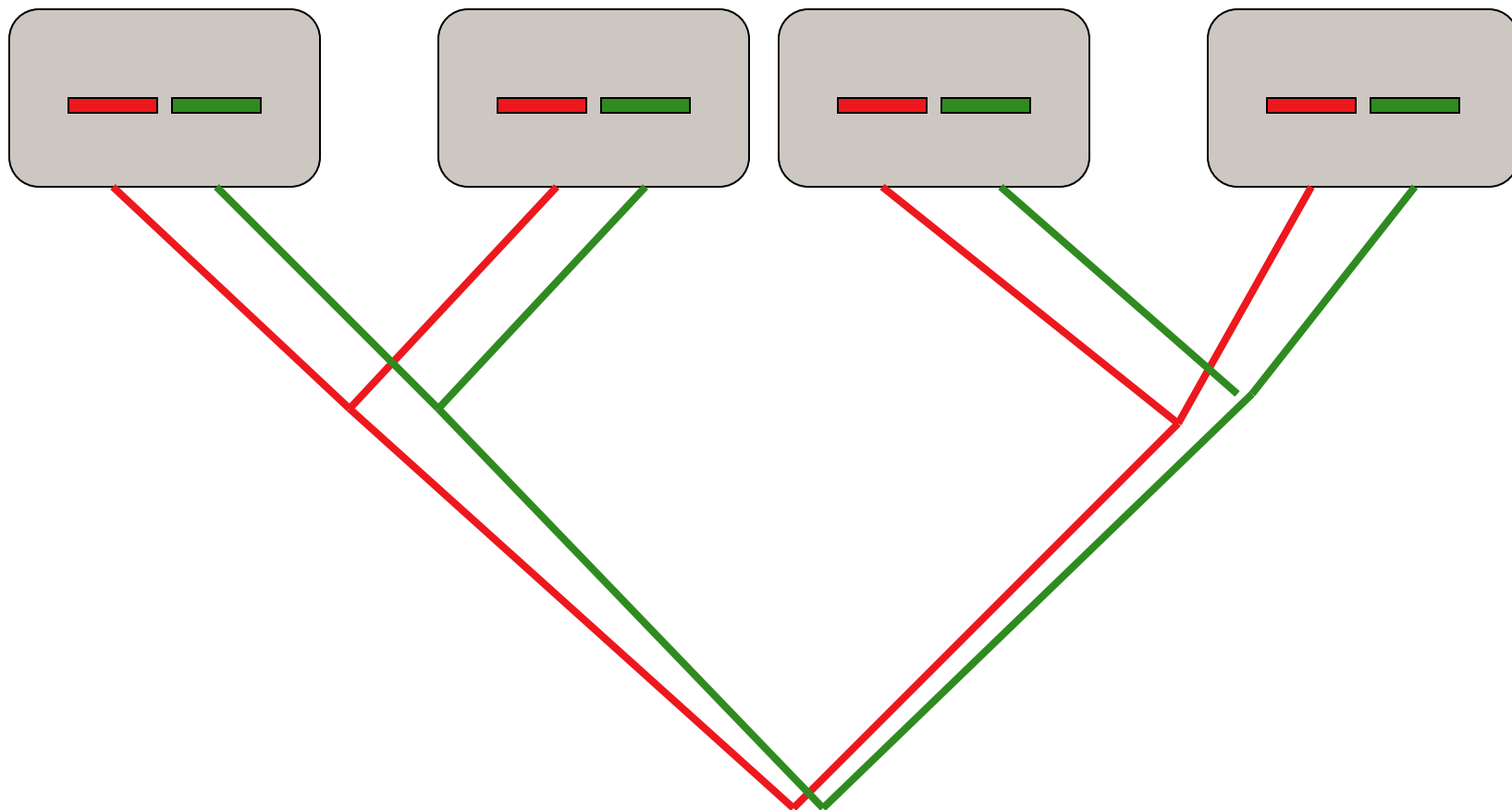


1866

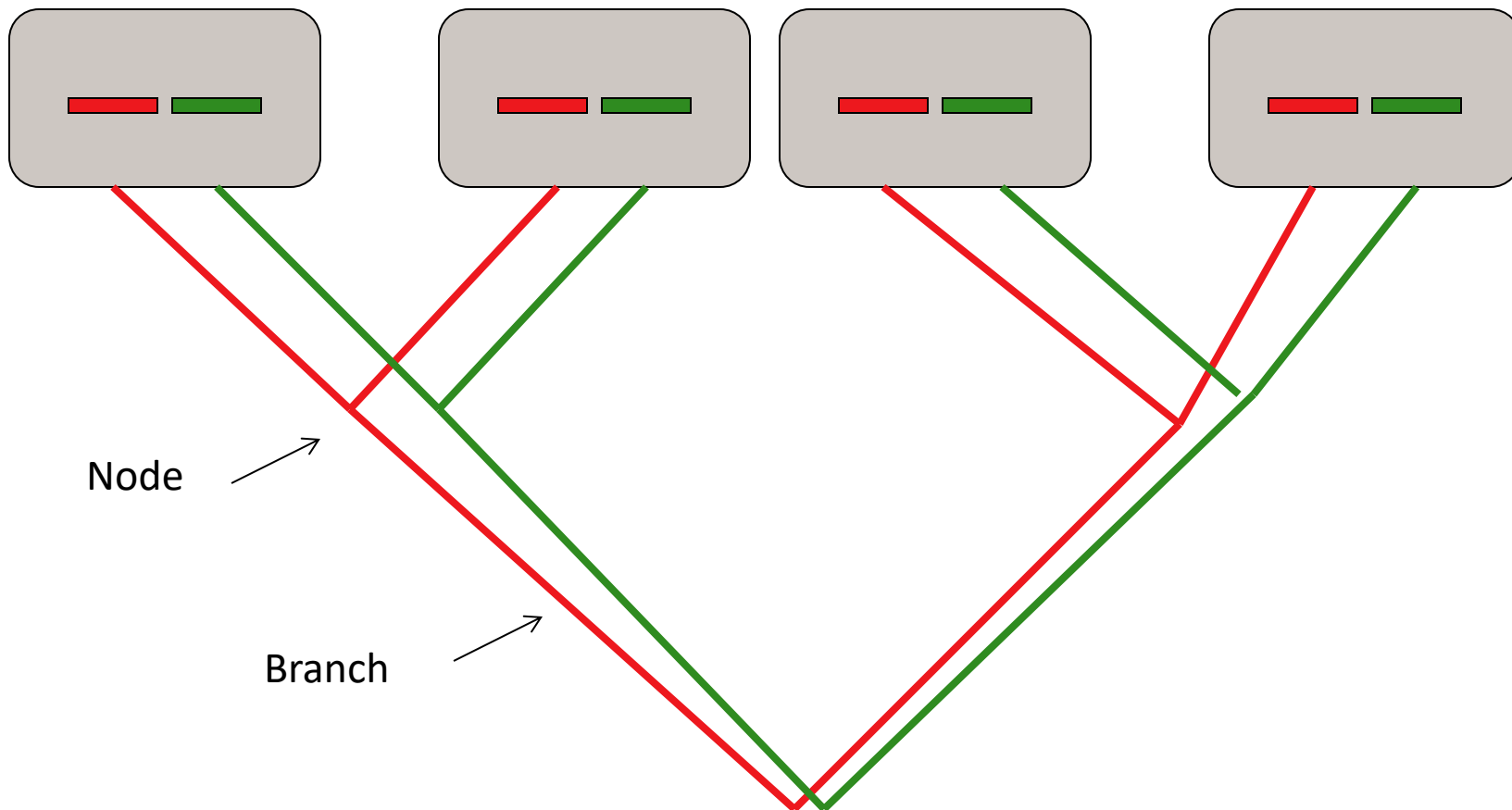
The Evolution of Man (1879)







Taxon



Node

Branch

Root

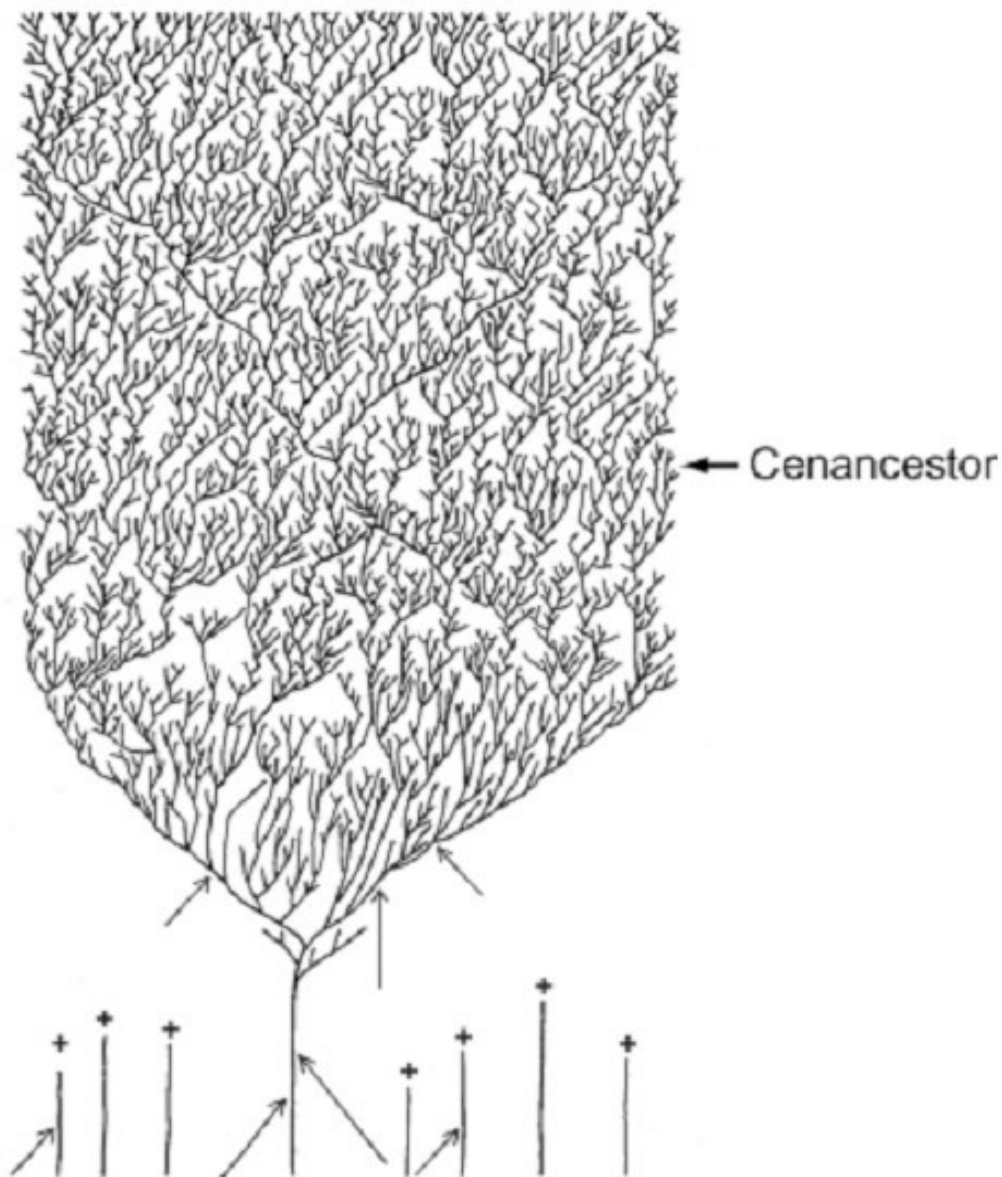
Present Day

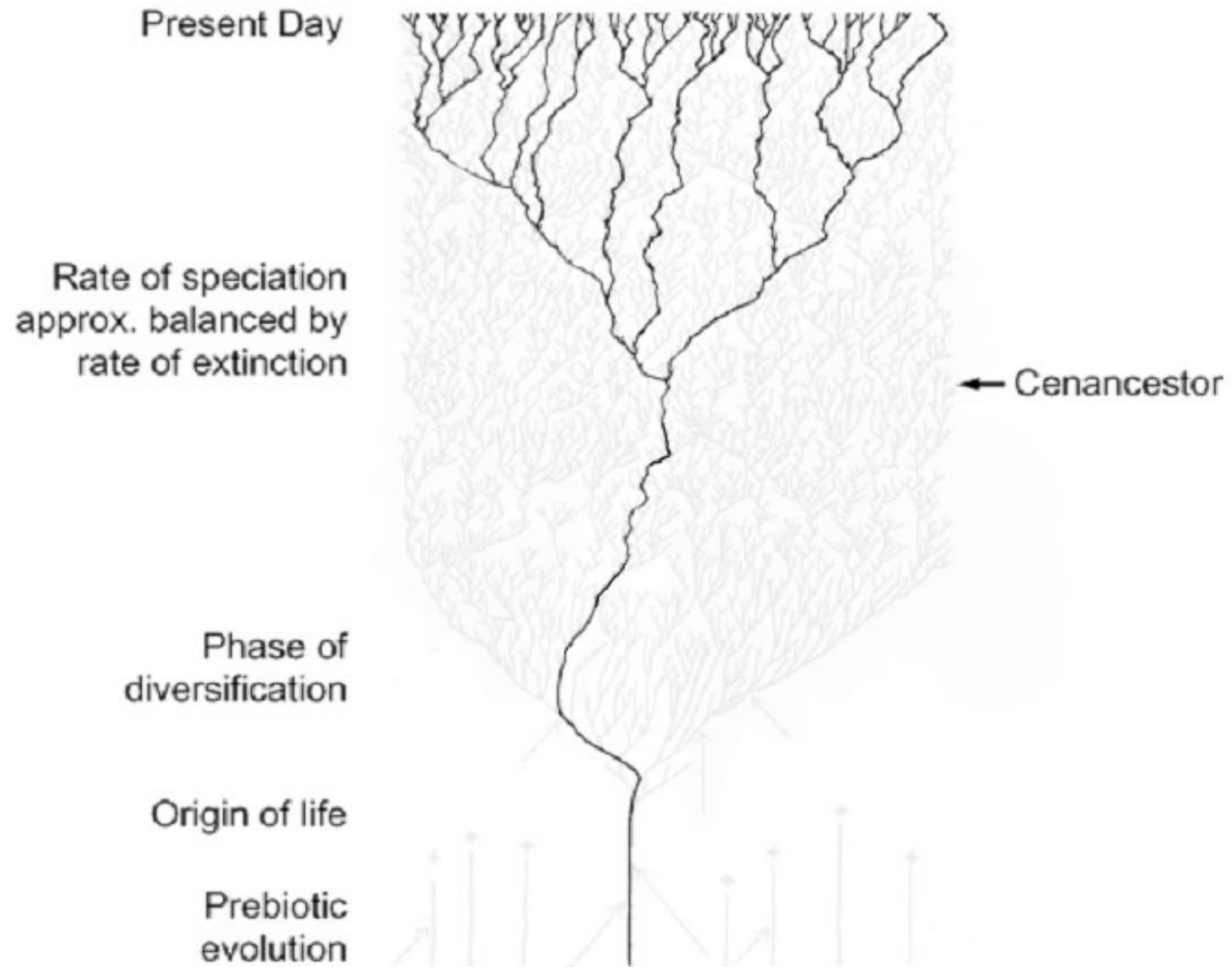
Rate of speciation
approx. balanced by
rate of extinction

Phase of
diversification

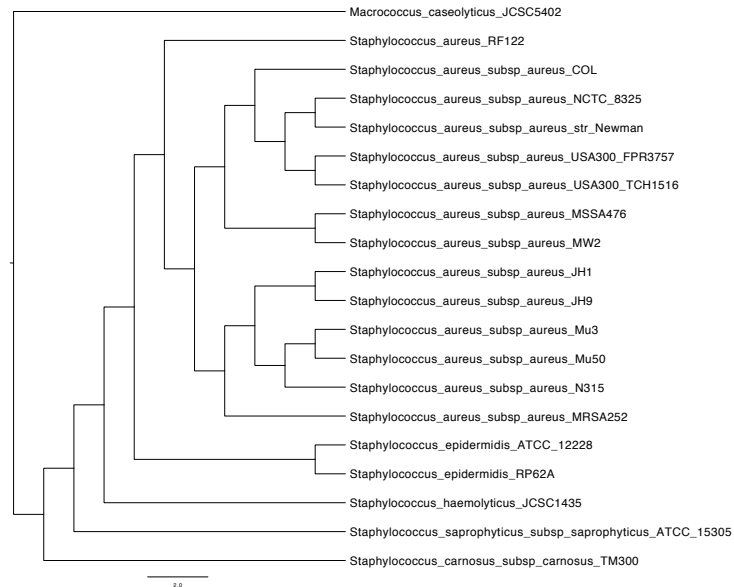
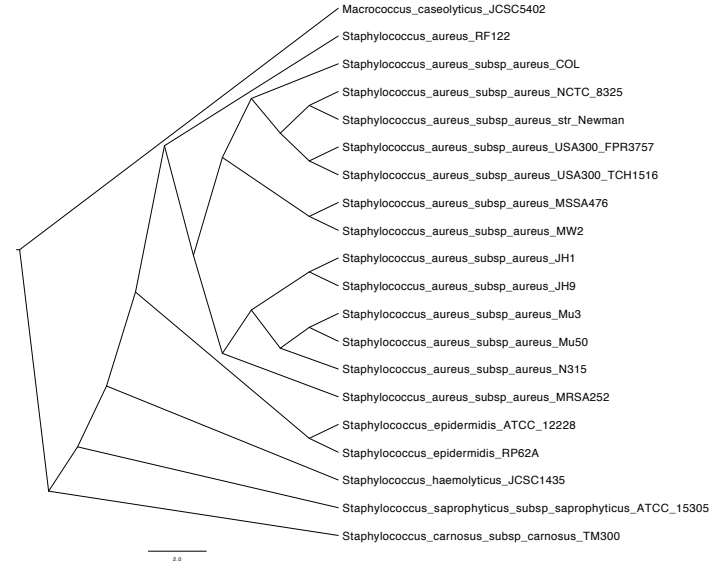
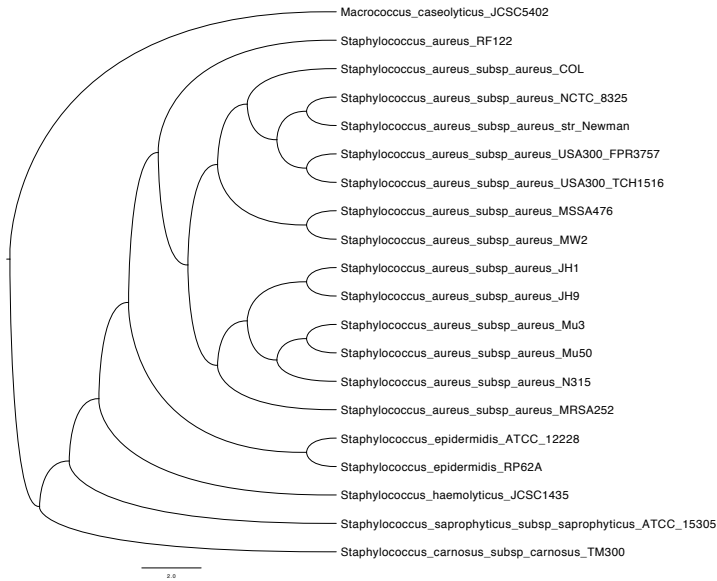
Origin of life

Prebiotic
evolution





Look Different but Same Topologies

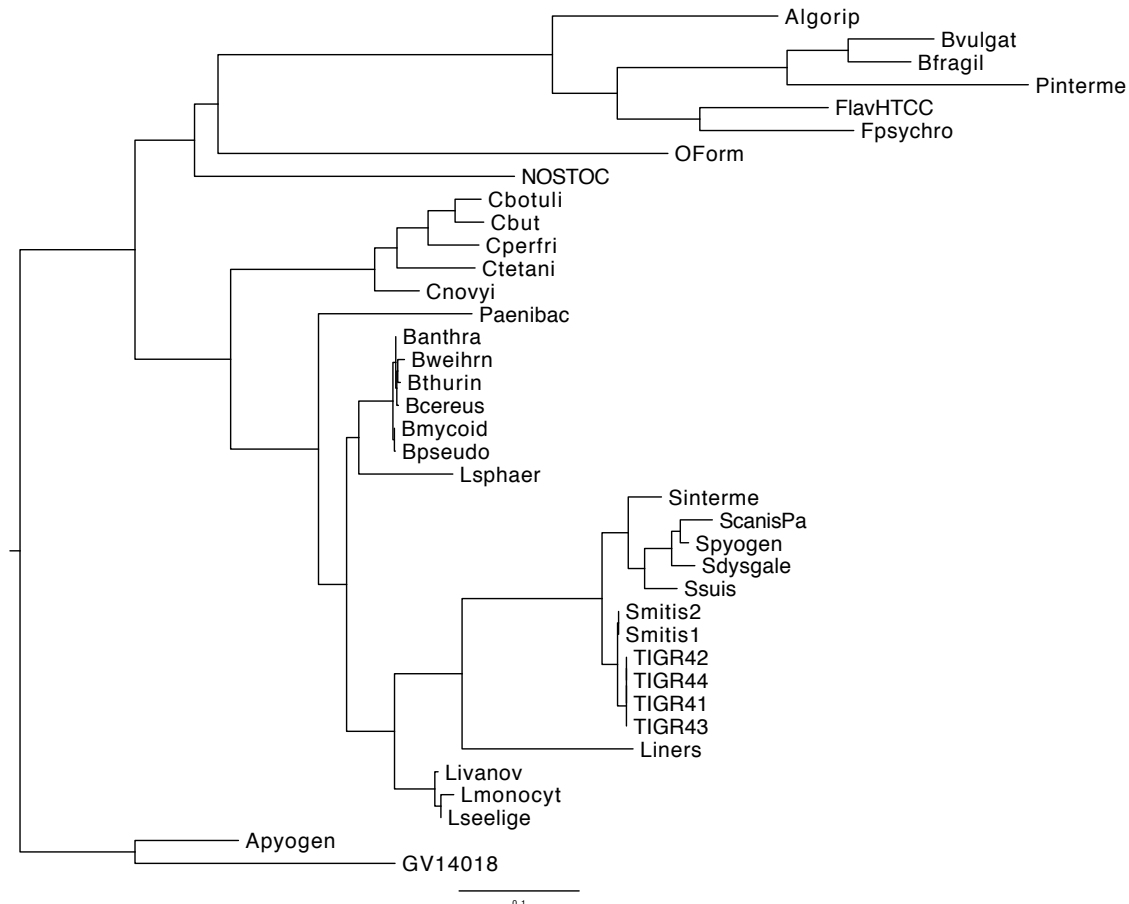


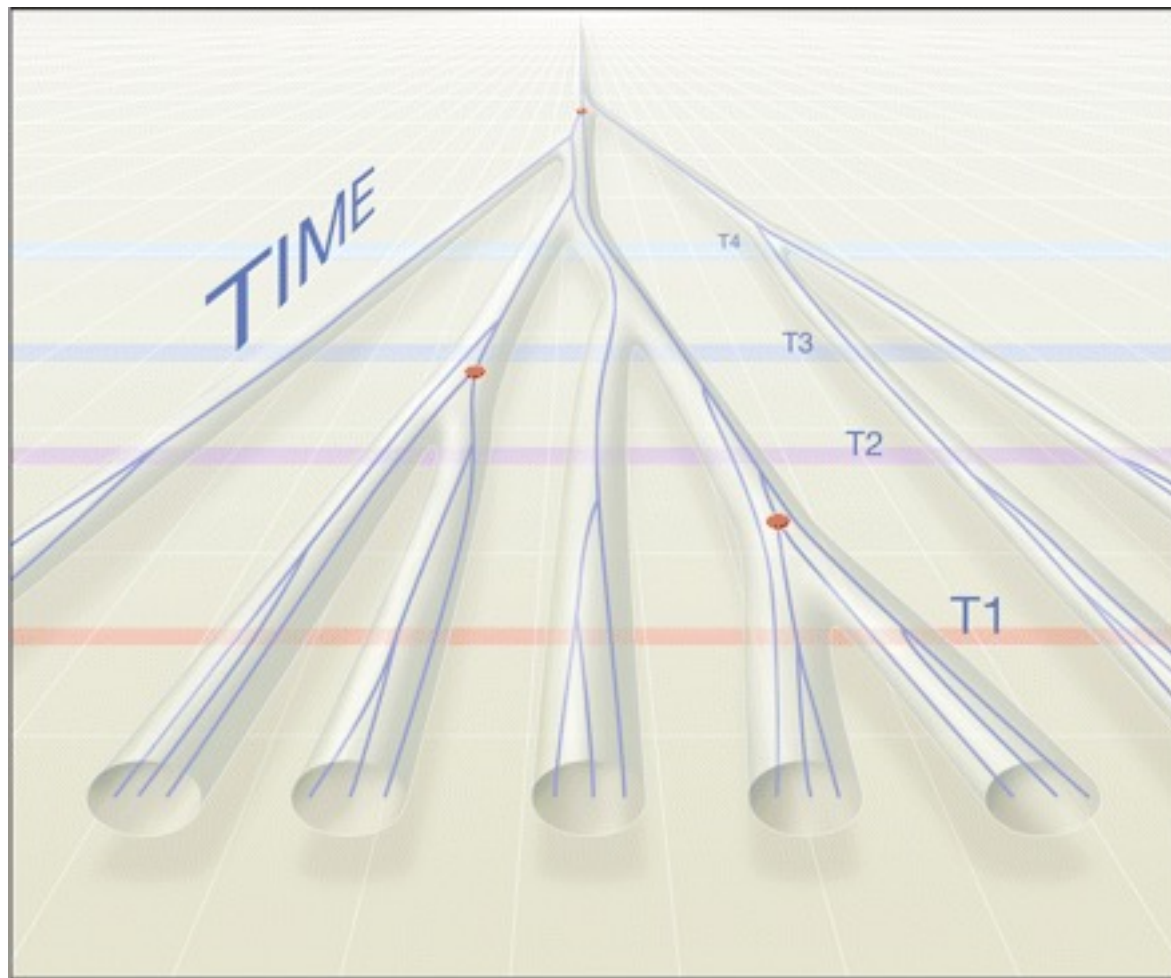
Time



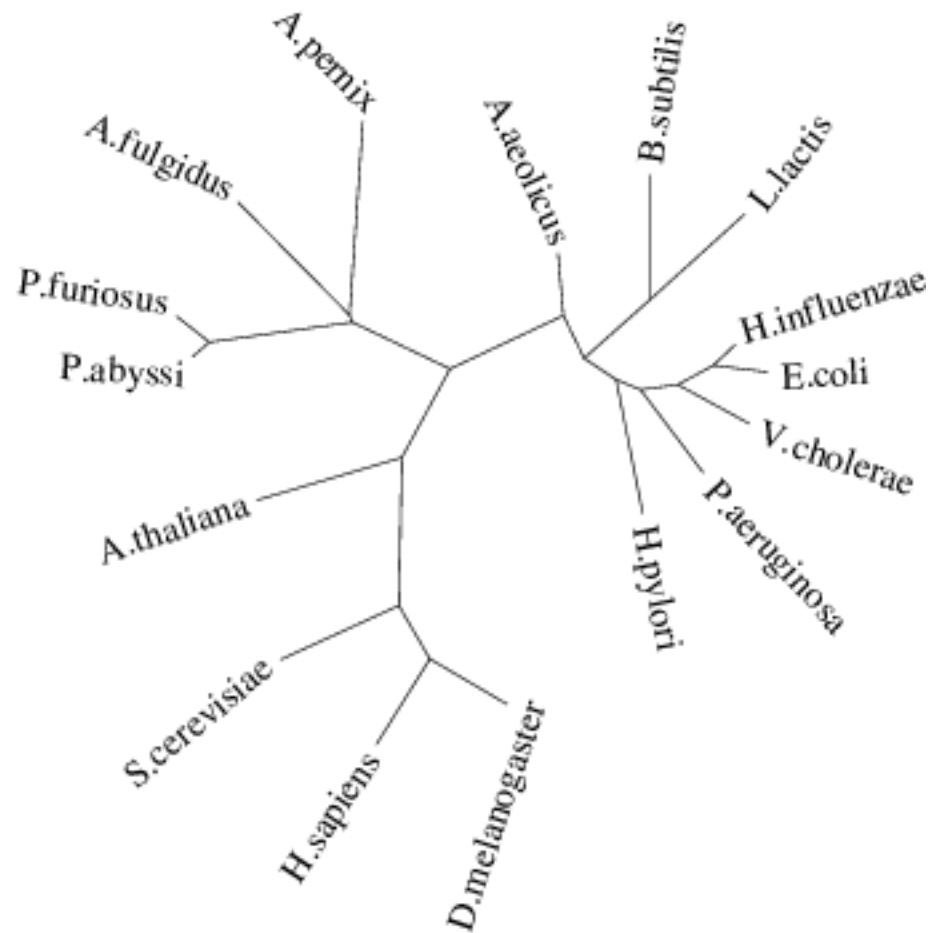
What about branch length:

1. Branch length can signify many different things (time, difference, inferred changes)
2. General rule: longer branch length indicates **more evolutionary change**

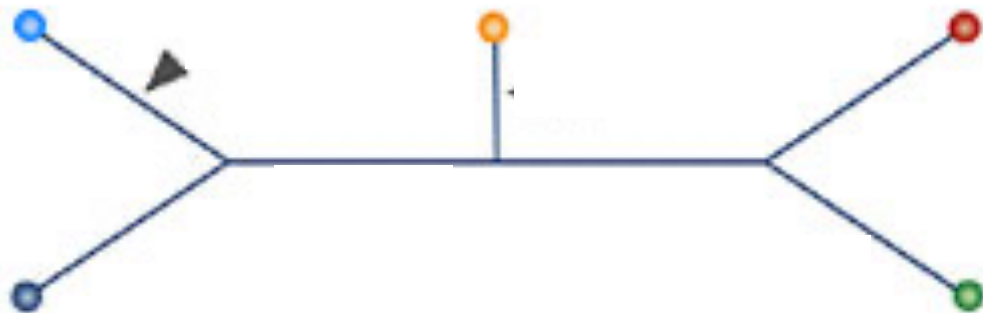


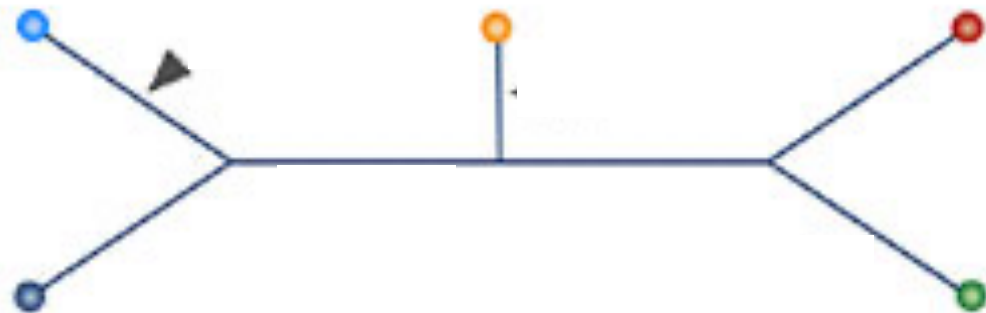


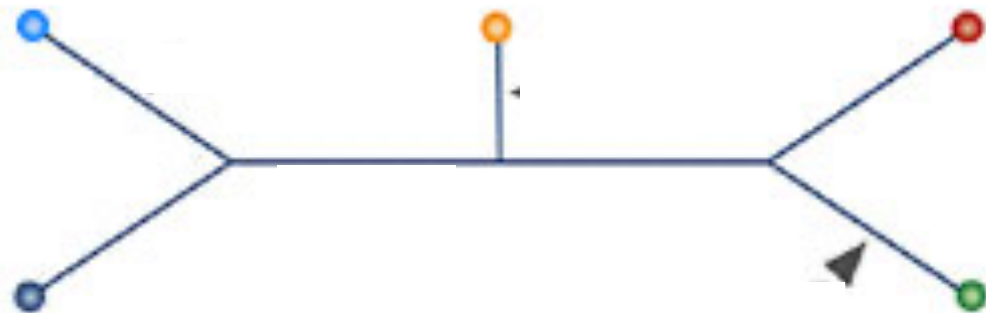
UNROOTED TREES

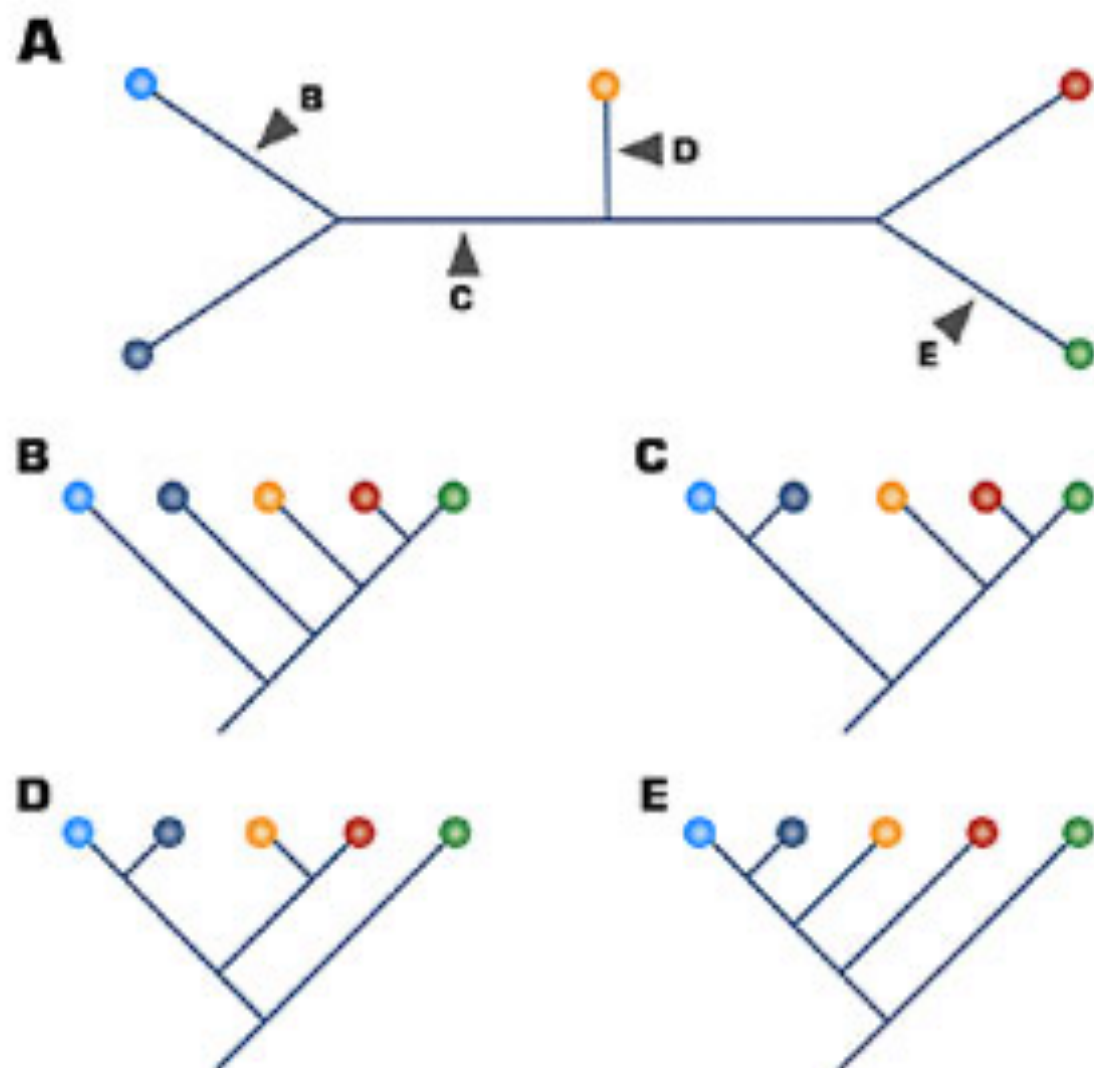


A



A**B**

A**B****C****D****E**



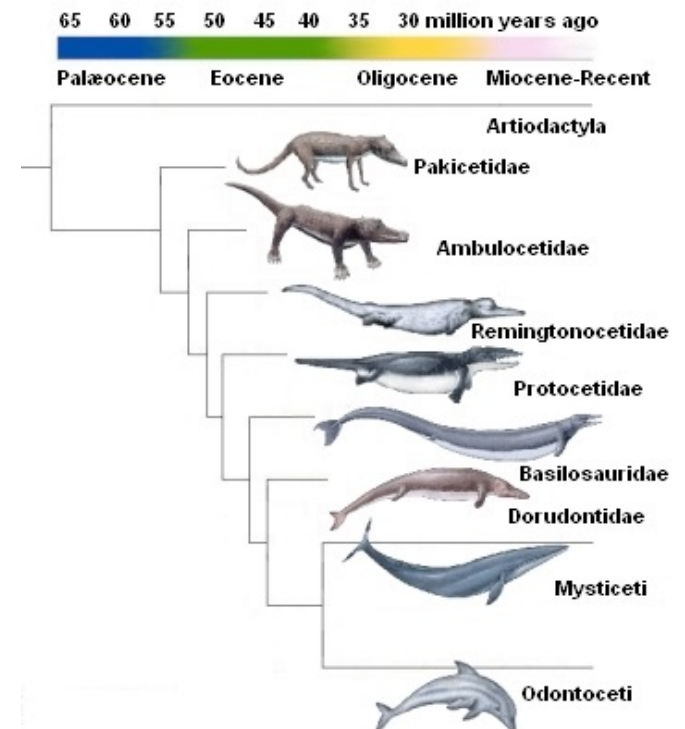
How do you decide how to root a phylogeny?

1. Use an **outgroup**:

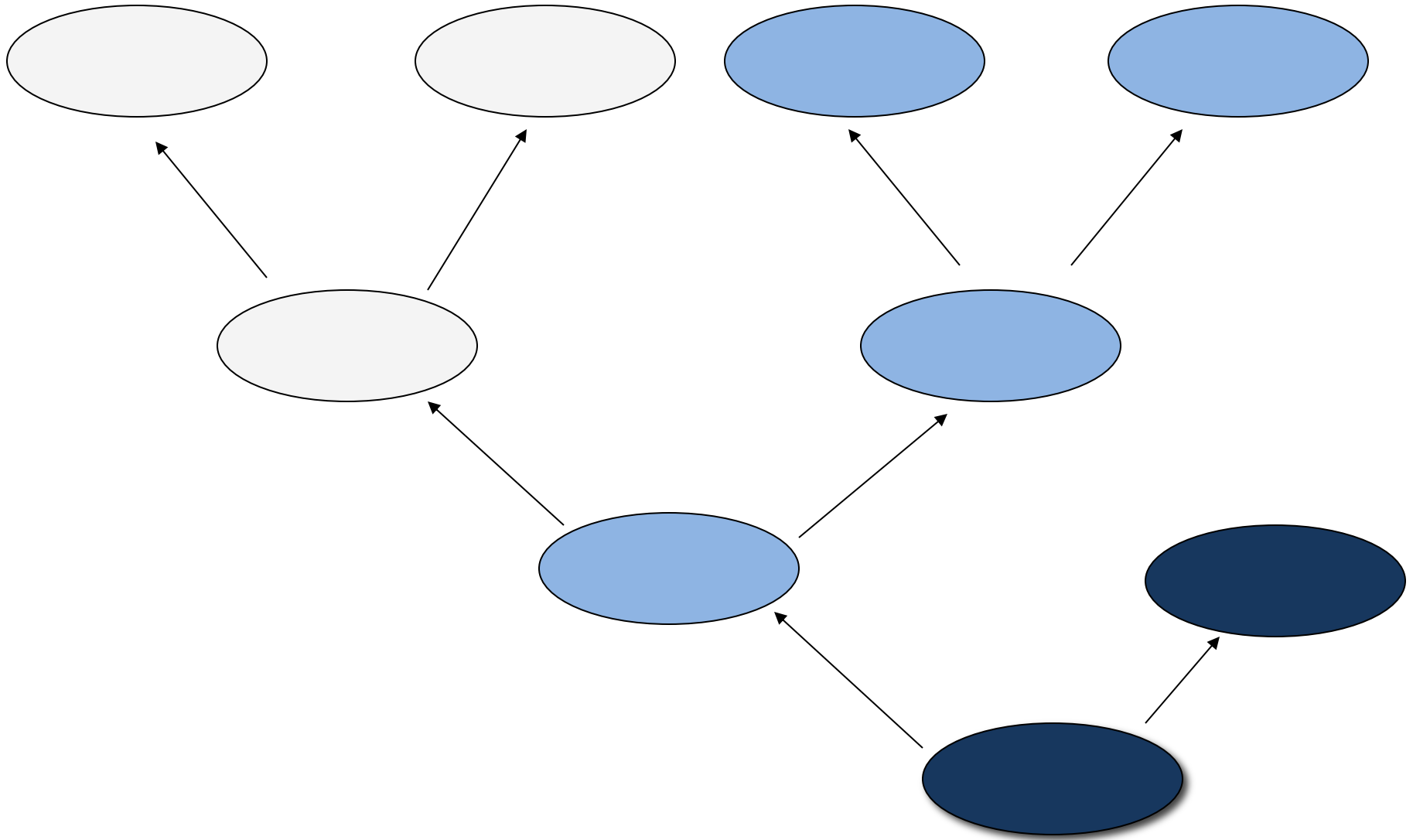
- a. The best outgroup is a taxon that is a close relative that you are pretty sure is not in the ingroup.
- b. The more outgroups the better.
- c. The branch of the outgroup becomes the root.

2. **Character polarity**

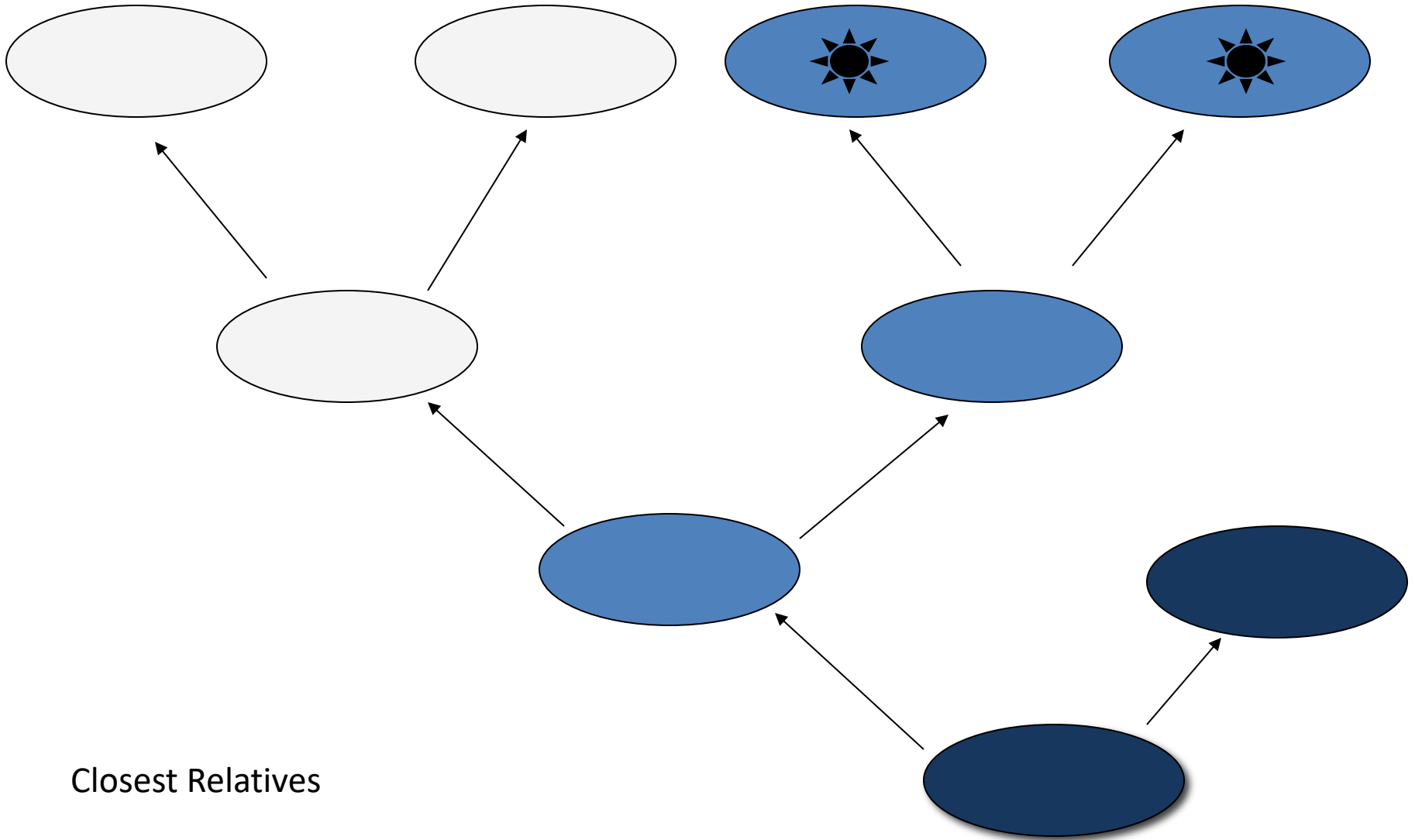
- a. Polarity is the direction of evolutionary change (ie., forelimb evolves into wing)
- b. Make sure that the topology has all of the assumed polarities going in the right direction.



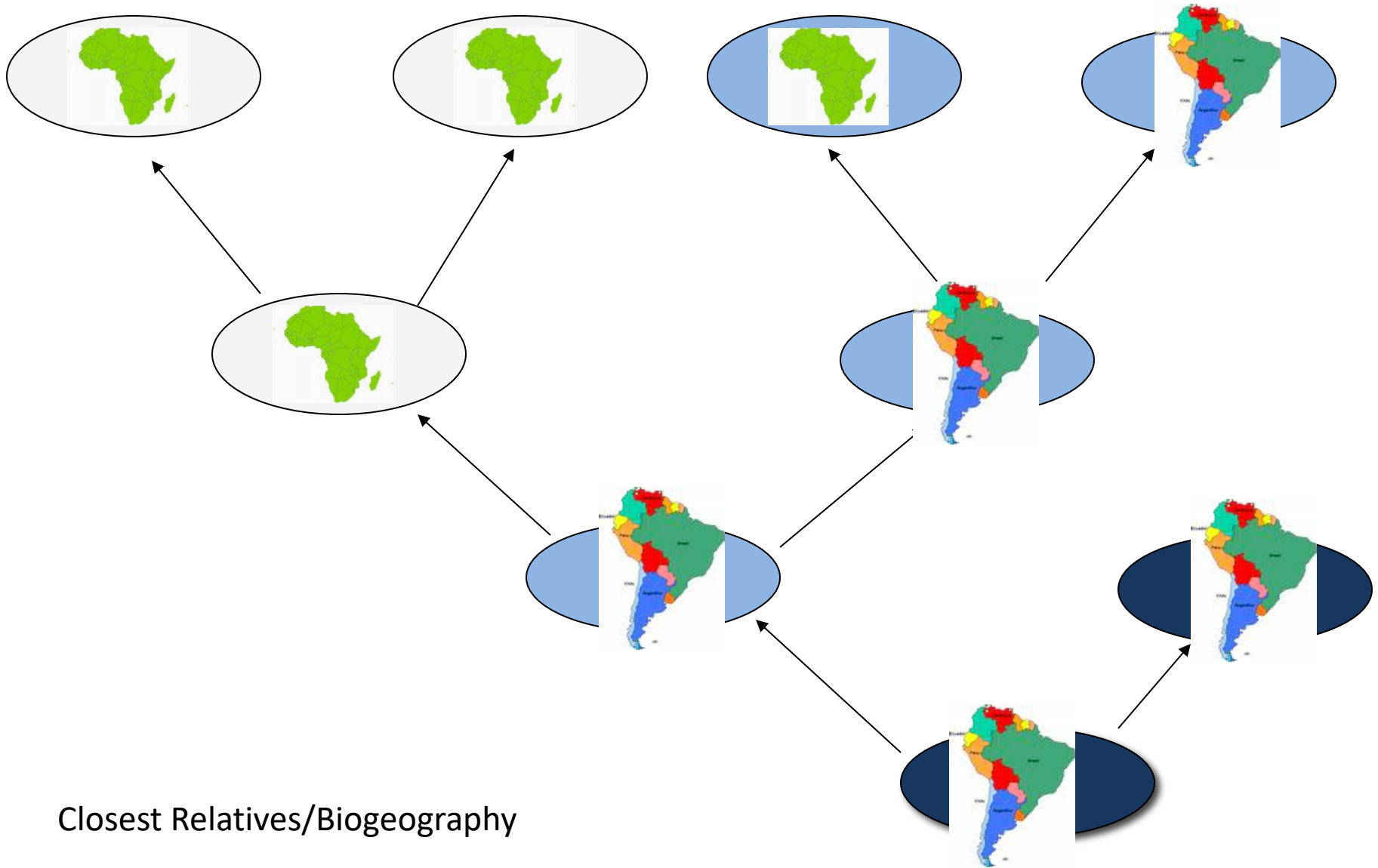
What info is in a tree?



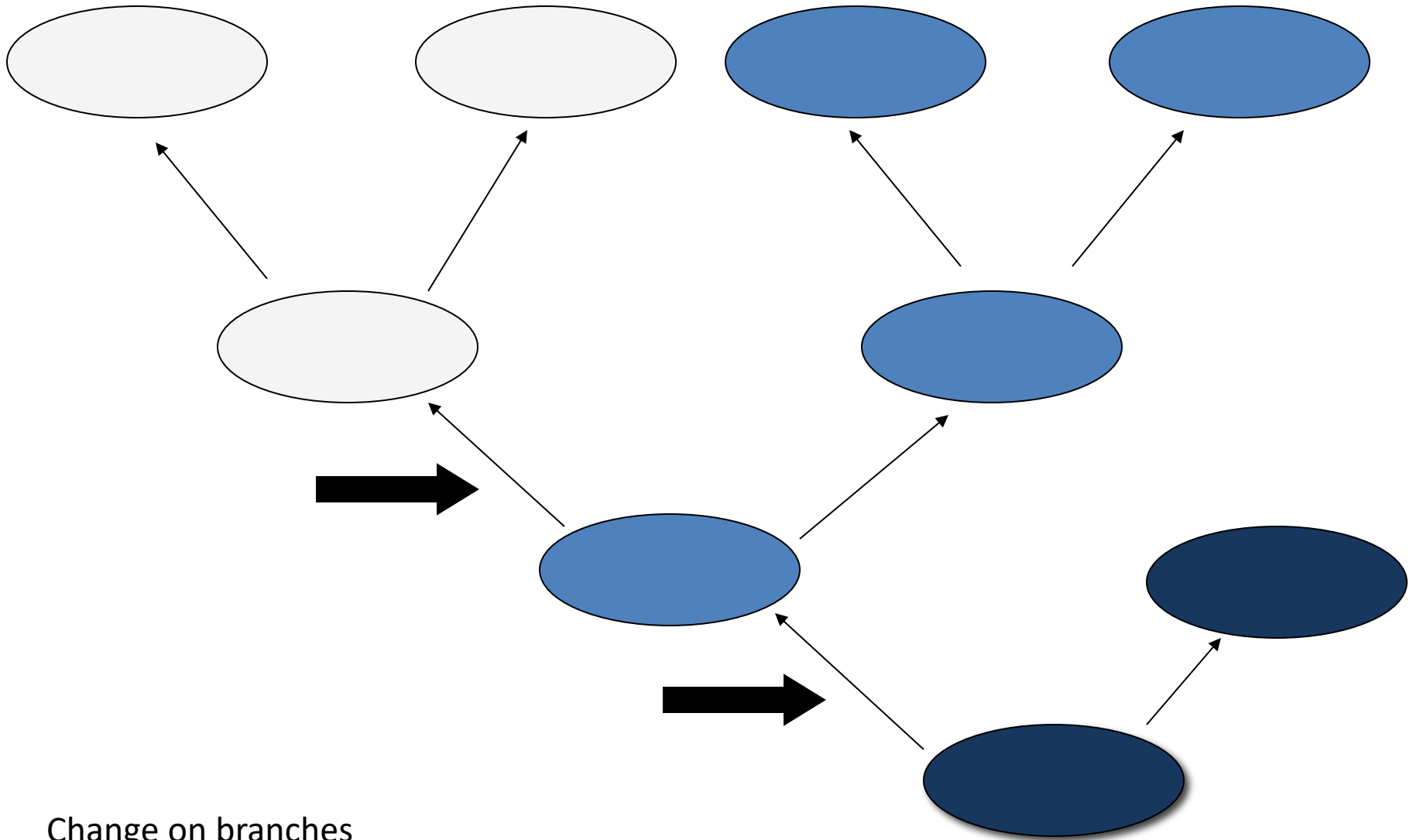
What info is in a tree?



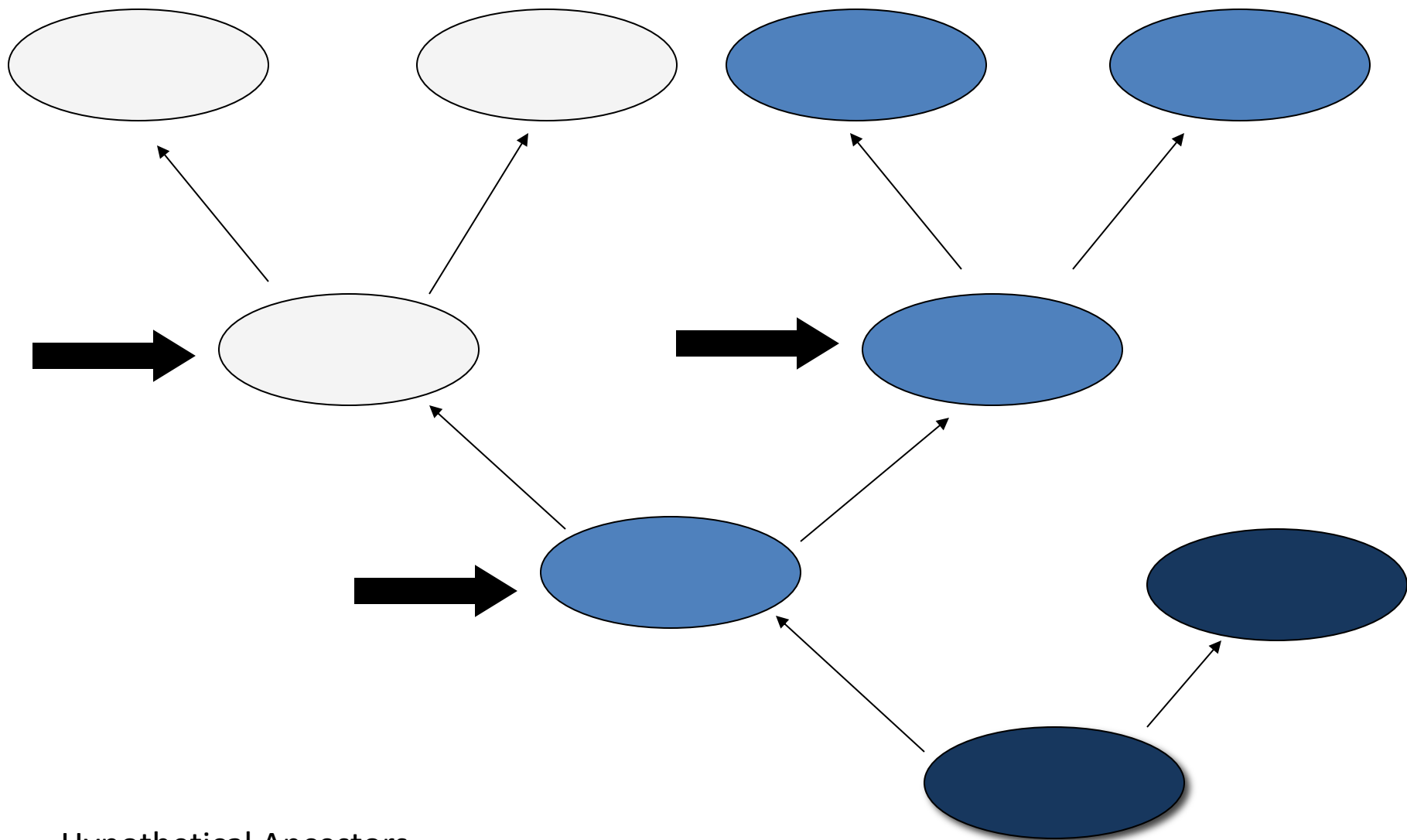
What info is in a tree?



What info is in a tree?

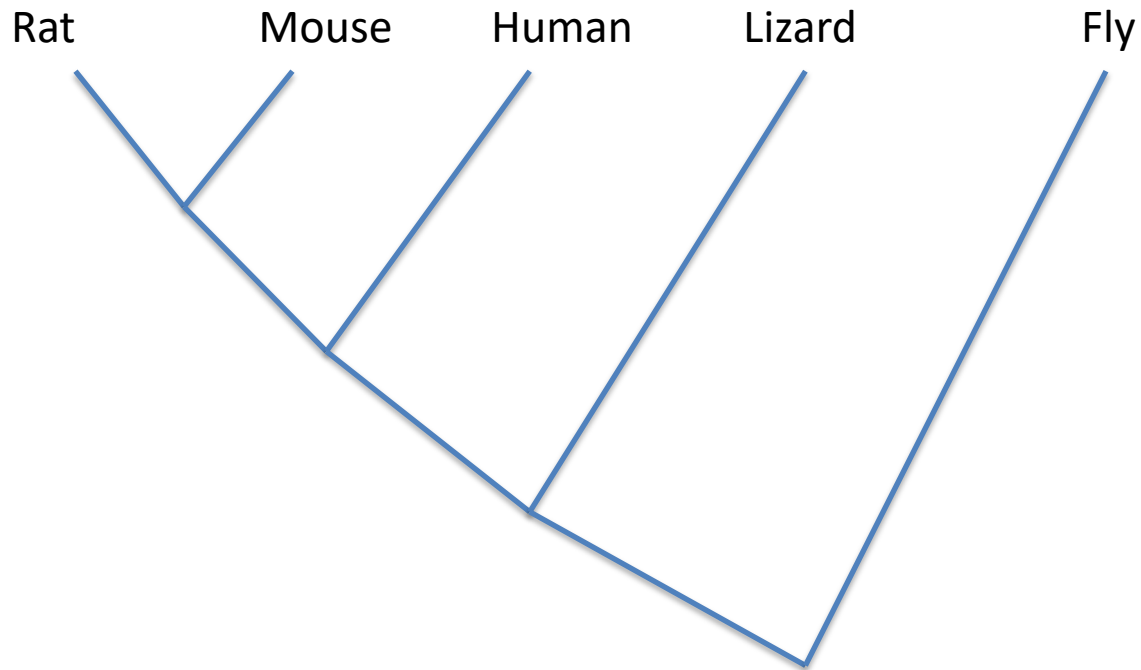


Change on branches
Evolutionary Direction/Polarity

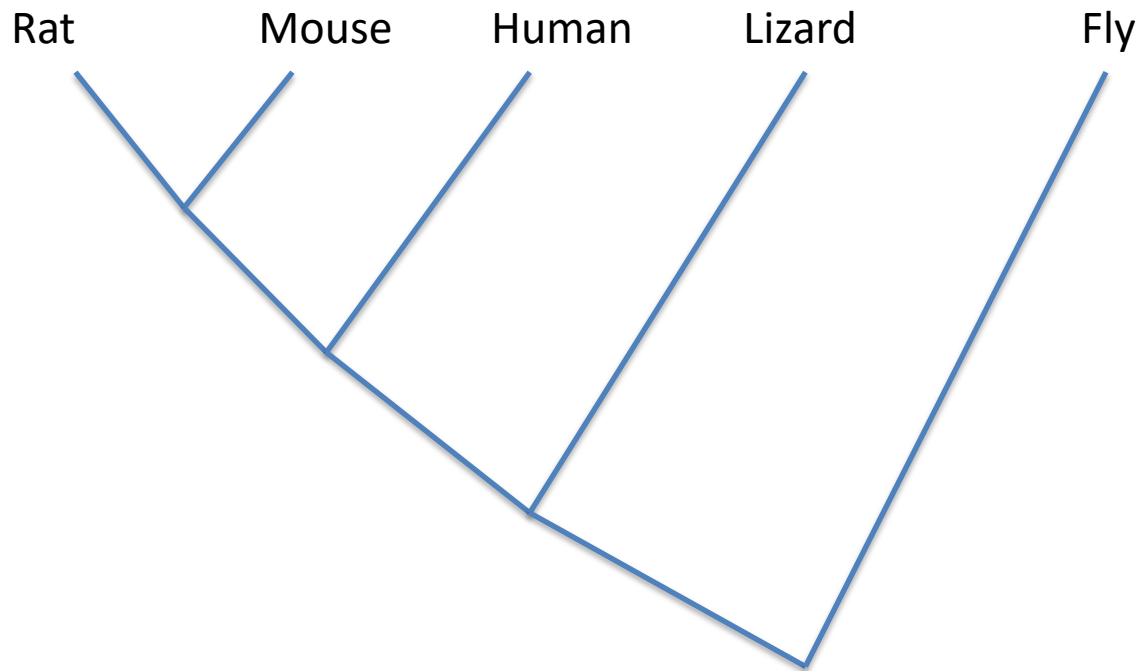


Hypothetical Ancestors

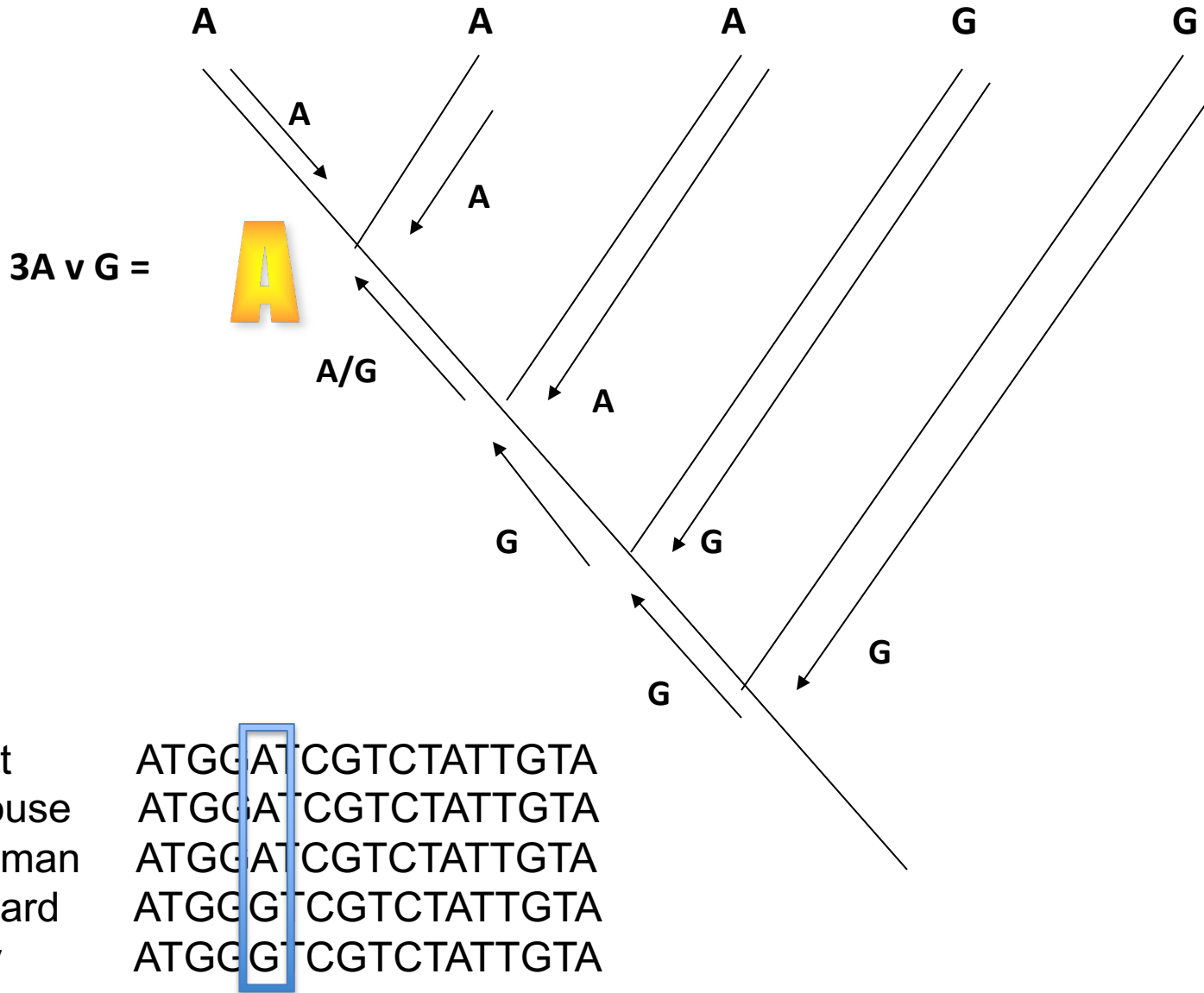
| | |
|--------|-------------------|
| Rat | ATGGATCGTCTATTGTA |
| Mouse | ATGGATCGTCTATTGTA |
| Human | ATGGATCGTCTATTGTA |
| Lizard | ATGGGTCGTCTATTGTA |
| Fly | ATGGGTCGTCTATTGTA |



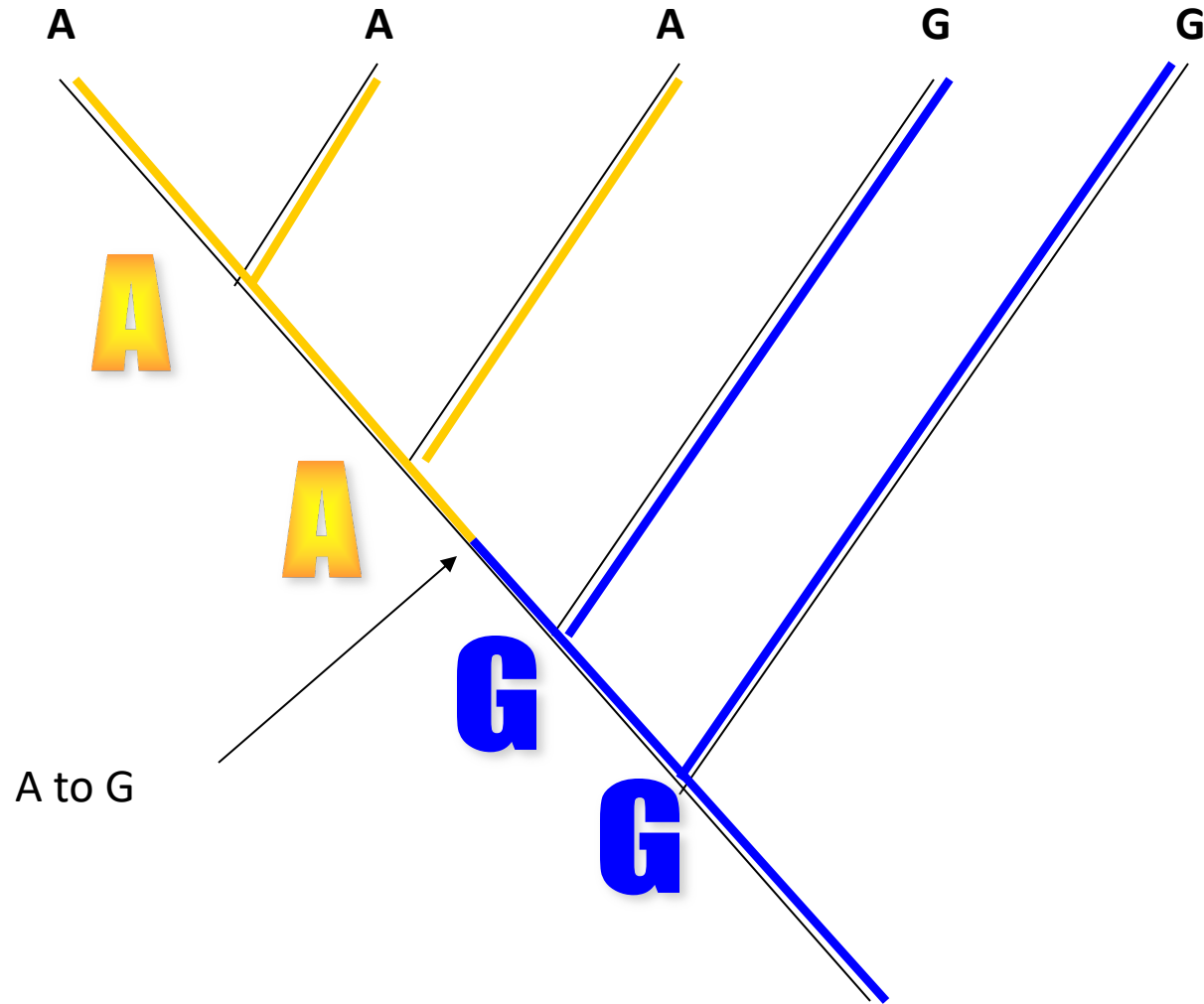
| | |
|--------|-------------------|
| Rat | ATGGATCGTCTATTGTA |
| Mouse | ATGGATCGTCTATTGTA |
| Human | ATGGATCGTCTATTGTA |
| Lizard | ATGGGTCGTCTATTGTA |
| Fly | ATGGGTCGTCTATTGTA |



FITCH OPTIMIZATION



FITCH OPTIMIZATION



Disclaimer: There are other more complicated statistical methods for this using maximum likelihood and Bayesian analysis but I did not think it was necessary to make you integrate!

How to make a tree from scratch starting with this sequence:

Human ACTCCAATGATAAATTTATAGTGGTGGAGCGCAATAAGAAATCACTGACAACTTCAC

Find Homologs and Align

| | |
|----------|--|
| Human | ACTCCAATGATAAATTTATAGTGGTGGAGCGCAATAAGAAATCACTGACAACCTTCAC |
| Chimp | ATTCAAATGGTAAATTTGTAGTAGTGGAACGCGAGAAAAAATCACTTACAACGTCAC |
| Chimp2 | AAATCATTCTTGAGACATTATGGGATAATGGGTTTGGTAAAACAATTCATGCACAGG |
| Gorilla | ATTCAAATGGTAAATTTGTAGTAGTGGAACGCGAGAAAAAATCACTTACAACCTTCAC |
| Gorilla2 | AAATCATTCTTGAGACATTATGGGATAATGGGTTTGGTAAAACAATTCATGCACAGG |
| Gibbon | AAACTCCAAATAAATATGTTGTAGTAAAACGAGAAAAGAAAAGTGTTTCAGATTCTA |
| Siamang | ATATTGAGGATAAATTCATTGTTGTAGAACGTACAAAAAAATCTATAAATACAACCTC |
| Human2 | AAAATCCAGATAAATATATAGTAATAAAGCGTGAAAAGAAAAGCATATCAGATTCAA |
| Lucy | AGACTGGTAATAAATTTATAGTTGTAGAACGTCAAAAAAGATCCCTTACAACATCAC |
| Macaca | AAAATCCAGATAAGTTCATTGTGGTAAAACGTGAGAAGAAGAGTATTTTCAGATTCCA |

Find Homologs and Align

| | |
|----------|--|
| Human | ACTCCAATGATAAATTTATAGTGGTGGAGCGCAATAAGAAATCACTGACAACCTTCAC |
| Chimp | ATTCAAATGGTAAATTTGTAGTAGTGGAACGCGAGAAAAAATCACTTACAACGTCAC |
| Chimp2 | AAATCATTCTTGAGACATTATGGGATAATGGGTTTGGTAAAACAATTCATGCACAGG |
| Gorilla | ATTCAAATGGTAAATTTGTAGTAGTGGAACGCGAGAAAAAATCACTTACAACCTTCAC |
| Gorilla2 | AAATCATTCTTGAGACATTATGGGATAATGGGTTTGGTAAAACAATTCATGCACAGG |
| Gibbon | AAACTCCAAATAAATATGTTGTAGTAAAACGAGAAAAGAAAAGTGTTTCAGATTCTA |
| Siamang | ATATTGAGGATAAATTCATTGTTGTAGAACGTACAAAAAAATCTATAAATACAACCTC |
| Human2 | AAAATCCAGATAAATATATAGTAATAAAGCGTGAAAAGAAAAGCATATCAGATTCAA |
| Lucy | AGACTGGTAATAAATTTATAGTTGTAGAACGTCAAAAAAGATCCCTTACAACATCAC |
| Macaca | AAAATCCAGATAAGTTCATTGTGGTAAAACGTGAGAAGAAGAGTATTTTCAGATTCCA |

Align Homologs

| | |
|----------|--|
| Human | ACTCCAATGATAAATTTATAGTGGTGGAGCGCAATAAGAAATCACTGACAACCTTCAC |
| Chimp | ATTCAAATGGTAAATTTGTAGTAGTGGAACGCGAGAAAAAATCACTTACAACGTCAC |
| Chimp2 | AAATCATTCTGAGACATTATGGGATAATGGGTTTGGTAAACAATTCATGCACAGG |
| Gorilla | ATTCAAATGGTAAATTTGTAGTAGTGGAACGCGAGAAAAAATCACTTACAACCTTCAC |
| Gorilla2 | AAATCATTCTGAGACATTATGGGATAATGGGTTTGGTAAACAATTCATGCACAGG |
| Gibbon | AAACTCCAAATAAATATGTTGTAGTAAACGAGAAAAGAAAAGTGTTCAGATTCTA |
| Siamang | ATATTGAGGATAAATTCATTGTTGTAGAACGTACAAAAAATCTATAAATACAACCTC |
| Human2 | AAAATCCAGATAAATATATAGTAATAAAGCGTGAAAAGAAAAGCATATCAGATTCAA |
| Lucy | AGACTGGTAATAAATTTATAGTTGTAGAACGTCAAAAAAGATCCCTTACAACATCAC |
| Macaca | AAAATCCAGATAAGTTCATTGTGGTAAACGTGAGAAGAAGAGTATTCAGATTCCA |



Find the best tree: that optimizes “some function”

Align Homologs

```
Human  ACTCCAATGATAAATTTATAGTGGTGGAGCGCAATAAGAAATCACTGACAACCTTCAC
Chimp  ATTCAAATGGTAAATTTGTAGTAGTGGAACGCGAGAAAAAATCACTTACAACGTCAC
Chimp2 AAATCATTCTGAGACATTATGGGATAATGGGTTTGGTAAAACAATTCATGCACAGG
Gorilla ATTCAAATGGTAAATTTGTAGTAGTGGAACGCGAGAAAAAATCACTTACAACCTTCAC
Gorilla2 AAATCATTCTGAGACATTATGGGATAATGGGTTTGGTAAAACAATTCATGCACAGG
Gibbon  AAACTCCAAATAAATATGTTGTAGTAAAACGAGAAAAAGAAAAGTGTTCAGATTCTA
Siamang ATATTGAGGATAAATTCATTGTTGTAGAACGTACAAAAAATCTATAAATACAACCTC
Human2  AAAATCCAGATAAATATATAGTAATAAAGCGTGAAAAGAAAAGCATATCAGATTCAA
Lucy    AGACTGGTAATAAATTTATAGTTGTAGAACGTCAAAAAAGATCCCTTACAACATCAC
Macaca  AAAATCCAGATAAGTTCATTGTGGTAAAACGTGAGAAGAAGAGTATTCAGATTCCA
```



Find the best tree: that optimizes “some function”

How many trees do I need to look through?

FOR UNROOTED TREES:

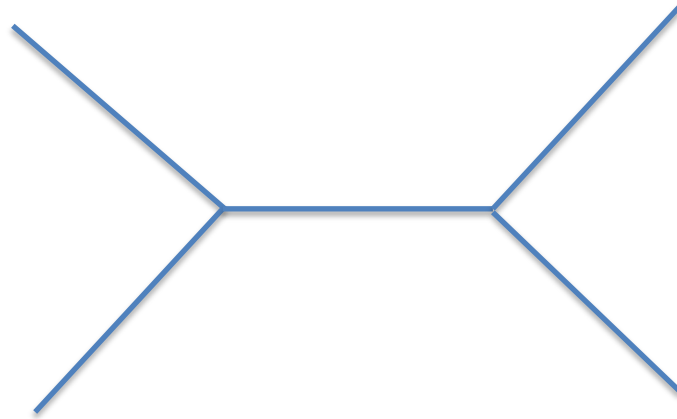
$$(2n-5)!/[2n-3(n-3)!]$$

FOR UNROOTED TREES:

$$(2n-3)!/[2n-2(n-2)!]$$

How tree building actually works.

A AAAAT
B AATTA
C TTTTA
D AATAT



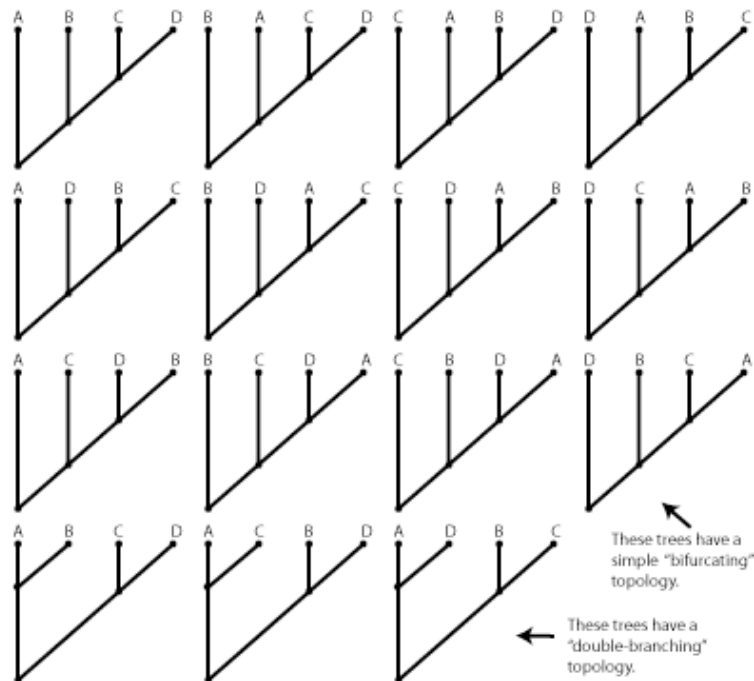
Number of Taxa Number of Rooted Trees

| | |
|----|---------------------------------|
| 3 | 3 |
| 4 | 15 |
| 5 | 105 |
| 6 | 945 |
| 7 | 10395 |
| 8 | 135135 |
| 9 | 2027025 |
| 10 | 34459425 |
| 11 | 654729075 |
| 12 | 13749310575 |
| 13 | 316234143225 |
| 14 | 7905853580625 |
| 15 | 213458046676875 |
| 16 | 6190283353629375 |
| 17 | 191898783962510625 |
| 18 | 6332659870762850625 |
| 19 | 221643095476699771875 |
| 20 | 8200794532637891559375 |
| 21 | 319830986772877770815625 |
| 22 | 13113070457687988603440625 |
| 23 | 563862029680583509947946875 |
| 24 | 25373791335626257947657609375 |
| 25 | 1192568192774434123539907640625 |

Align Homologs

| | |
|---------|--|
| Human | ACTCCAATGATAAATTTATAGTGGTGGAGCGCAATAAGAAATCACTGACAACCTTCAC |
| Chimp | ATTCAAATGGTAAATTTGTAGTAGTGGAACGCGAGAAAAAATCACTTACAACGTCAC |
| Gorilla | ATTCAAATGGTAAATTTGTAGTAGTGGAACGCGAGAAAAAATCACTTACAACCTTCAC |
| Orang | AAATCATTCTGAGACATTATGGGATAATGGGTTTGGTAAACAATTCATGCACAGG |

4 taxa = 15 trees

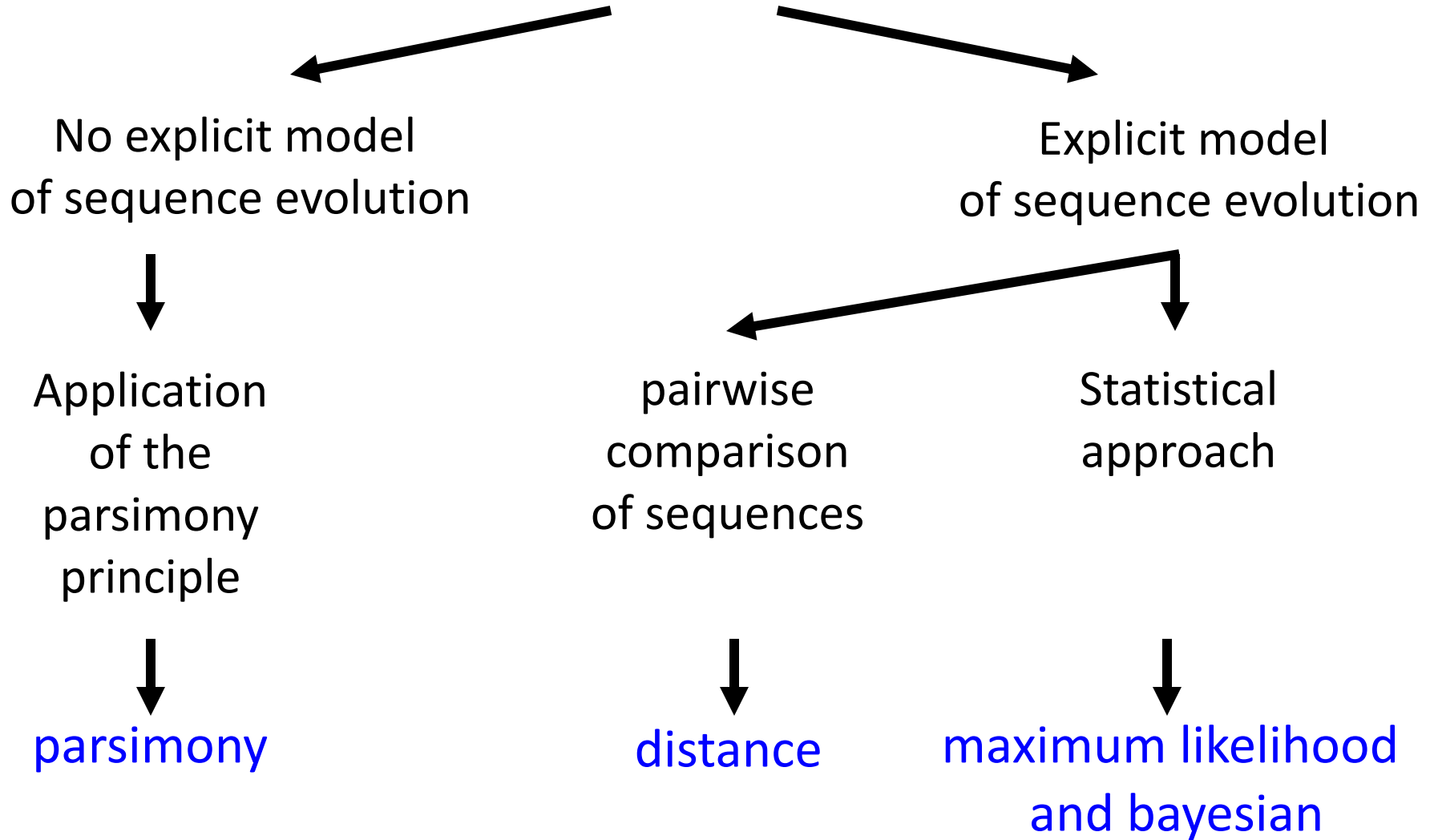


Deciding which tree to pick:

Optimality Criteria

- **Maximum Parsimony** (what is the simplest explanation that accounts for all the data)
- **Maximum Likelihood** (what is the highest likelihood of observing the data given a tree)
- **Bayesian** (what is the highest probability that the tree is correct given the data)
- **Distance/Similarity** (minimize the distance or group the most similar things together)

Tree-Building Methods

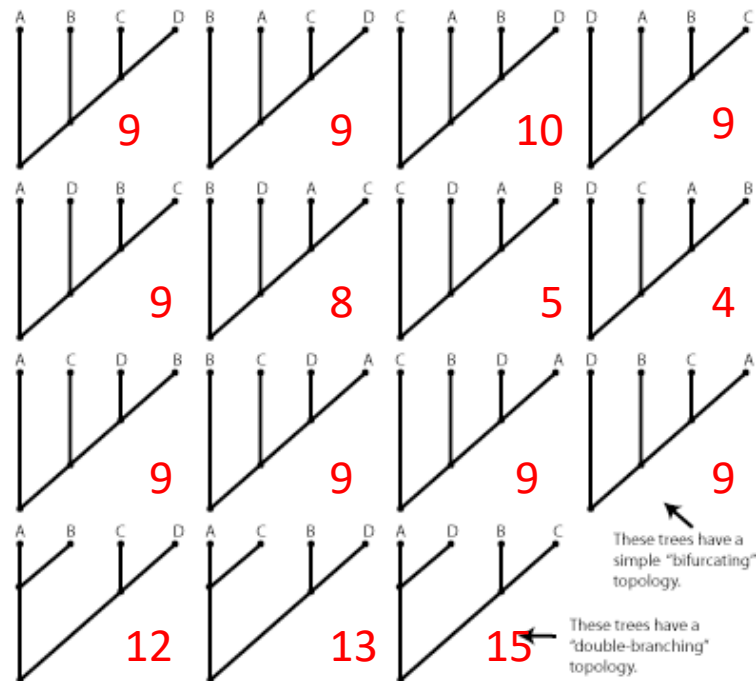


Align Homologs

| | |
|------------|--|
| A. Human | ACTCCAATGATAAATTTATAGTGGTGGAGCGCAATAAGAAATCACTGACAACCTTCAC |
| B. Chimp | ATTCAAATGGTAAATTTGTAGTAGTGGAACGCGAGAAAAAATCACTTACAACGTCAC |
| C. Gorilla | ATTCAAATGGTAAATTTGTAGTAGTGGAACGCGAGAAAAAATCACTTACAACCTTCAC |
| D. Orang | AAATCATTTCTGAGACATTATGGGATAATGGGTTTGGTAAAACAATTCATGCACAGG |

4 taxa = 15 trees

Steps in red

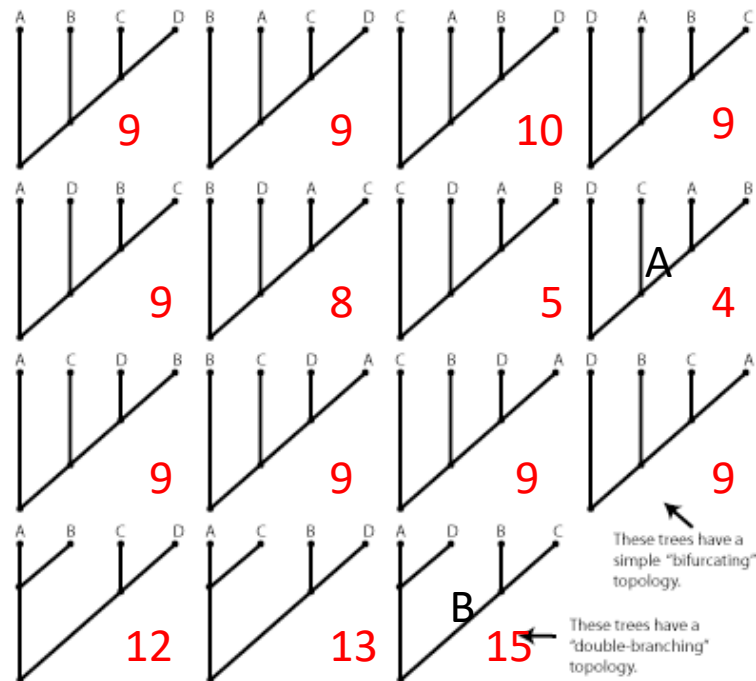


Align Homologs

| | |
|------------|--|
| A. Human | ACTCCAATGATAAATTTATAGTGGTGGAGCGCAATAAGAAATCACTGACAACCTTCAC |
| B. Chimp | ATTCAAATGGTAAATTTGTAGTAGTGGAACGCGAGAAAAAATCACTTACAACGTCAC |
| C. Gorilla | ATTCAAATGGTAAATTTGTAGTAGTGGAACGCGAGAAAAAATCACTTACAACCTTCAC |
| D. Orang | AAATCATTTCTGAGACATTATGGGATAATGGGTTTGGTAAAACAATTCATGCACAGG |

4 taxa = 15 trees

Steps in red

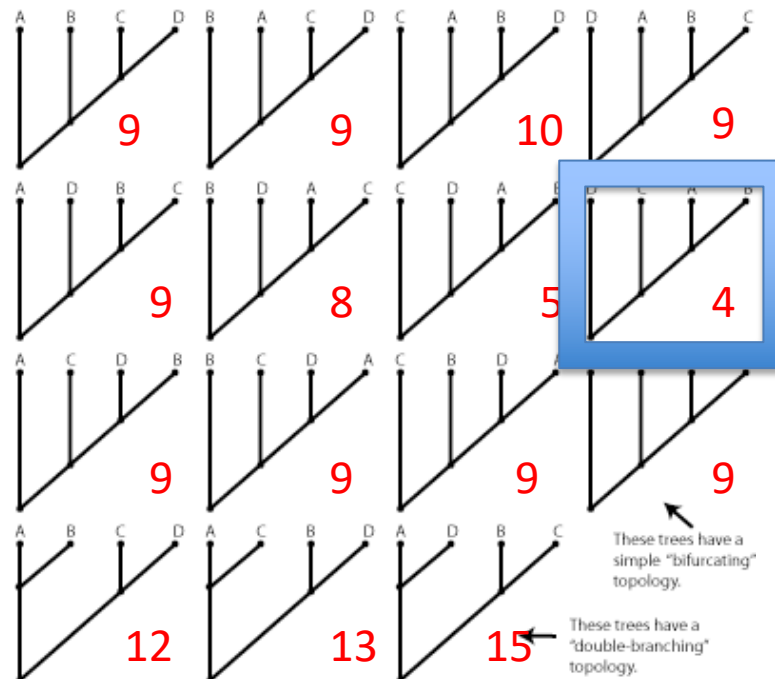


Align Homologs

A. Human ACTCCAATGATAAATTTATAGTGGTGGAGCGCAATAAGAAATCACTGACAACCTTCAC
 B. Chimp ATTCAAATGGTAAATTTGTAGTAGTGGAACGCGAGAAAAAATCACTTACAACGTCAC
 C. Gorilla ATTCAAATGGTAAATTTGTAGTAGTGGAACGCGAGAAAAAATCACTTACAACCTTCAC
 D. Orang AAATCATTCTGAGACATTATGGGATAATGGGTTTGGTAAAACAATTCATGCACAGG

4 taxa = 15 trees

Steps in red

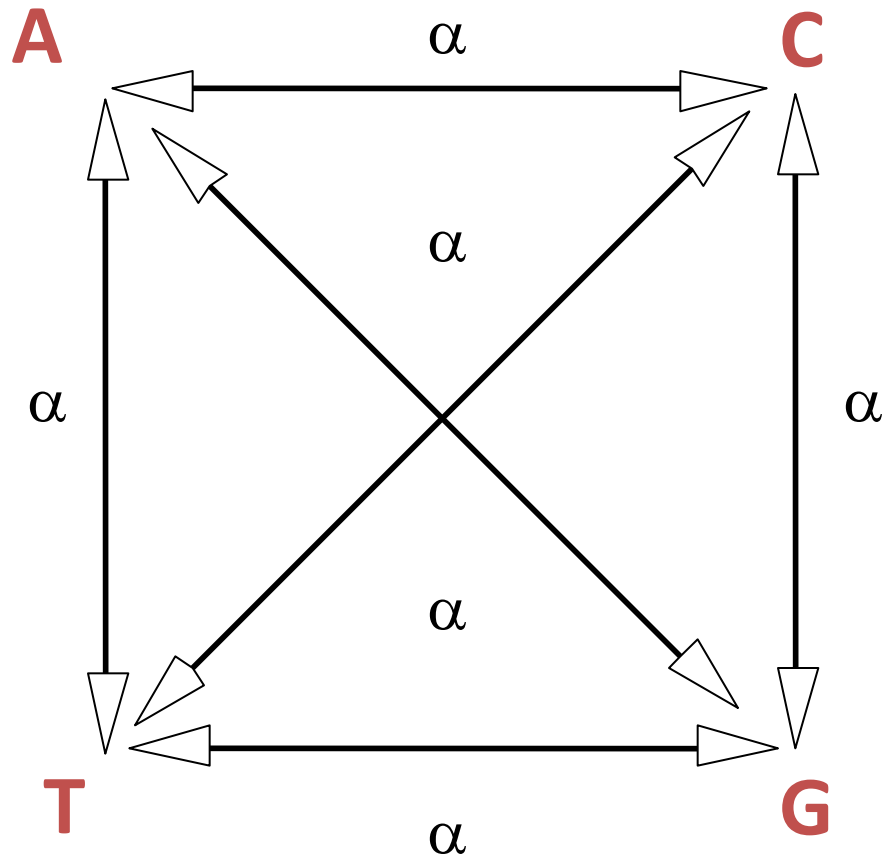


Estimating Genetic Distance

| | | | | | |
|--------|-------------|------------|------------|------------|--------------|
| SIVcpz | ATGGGTGCGA | GAGCGTCAGT | TCTAACAGGG | GGAAAATTAG | ATCGCTGGGA |
| HIV-1 | ATGGGTGCGA | GAGCGTCAGT | ATTAAGCGGG | GGAGAATTAG | ATCGATGGGA |
| SIVcpz | AAAAGTTCGG | CTTAGGCCCG | GGGGAAGAAA | AAGATATATG | ATGAAACATT |
| HIV-1 | AAAAATTTCGG | TTAAGGCCAG | GGGGAAAGAA | AAATATAAA | TTAAAACATA |
| SIVcpz | TAGTATGGGC | AAGCAGGGAG | CTGGAAAGAT | TCGCATGTGA | CCCCGGGCTA |
| HIV-1 | TAGTATGGGC | AAGCAGGGAG | CTAGAACGAT | TCGCAGTTAA | TCC TGG CCTG |
| SIVcpz | ATGGAAAGTA | AGGAAGGATG | TACTAAATTG | TTACAACAAT | TAGAGCCAGC |
| HIV-1 | TTAGAAACAT | CAGAAGGCTG | TAGACAAATA | CTGGGACAGC | TACAACCATC |
| SIVcpz | TCTCAAAACA | GGCTCAGAAG | GACTGCGGTC | CTTGTTTAAC | ACTCTGGCAG |
| HIV-1 | CCTTCAGACA | GGATCAGAAG | AACTTAGATC | ATTATATAAT | ACAGTAGCAA |
| SIVcpz | TACTGTGGTG | CATACATAGT | GACATCACTG | TAGAAGACAC | ACAGAAAGCT |
| HIV-1 | CCCTCTATTG | TGTGCATCAA | AGGATAGAGA | TAAAGACAC | CAAGGAAGCT |
| SIVcpz | CTAGAACAGC | TAAAGCGGCA | TCATGGAGAA | CAACAGAGCA | AAACTGAAAG |
| HIV-1 | TTAGACAAGA | TAGAG--GAA | -----GAGCA | AAACAAGT | AA---GAAAA |
| SIVcpz | TAACTCAGGA | AGCCGTGAAG | GGGGAGCCAG | TCAAGGCGCT | AGTGCCTCTG |
| HIV-1 | AAGCACAGCA | AGC-----AG | CAGCTGACA- | -CAGGACAC- | AG--CAGC-- |
| SIVcpz | CTGGCATTAG | TGGAAATTAC | | | |
| HIV-1 | CAGG--TCAG | CCAAAATTAC | | | |

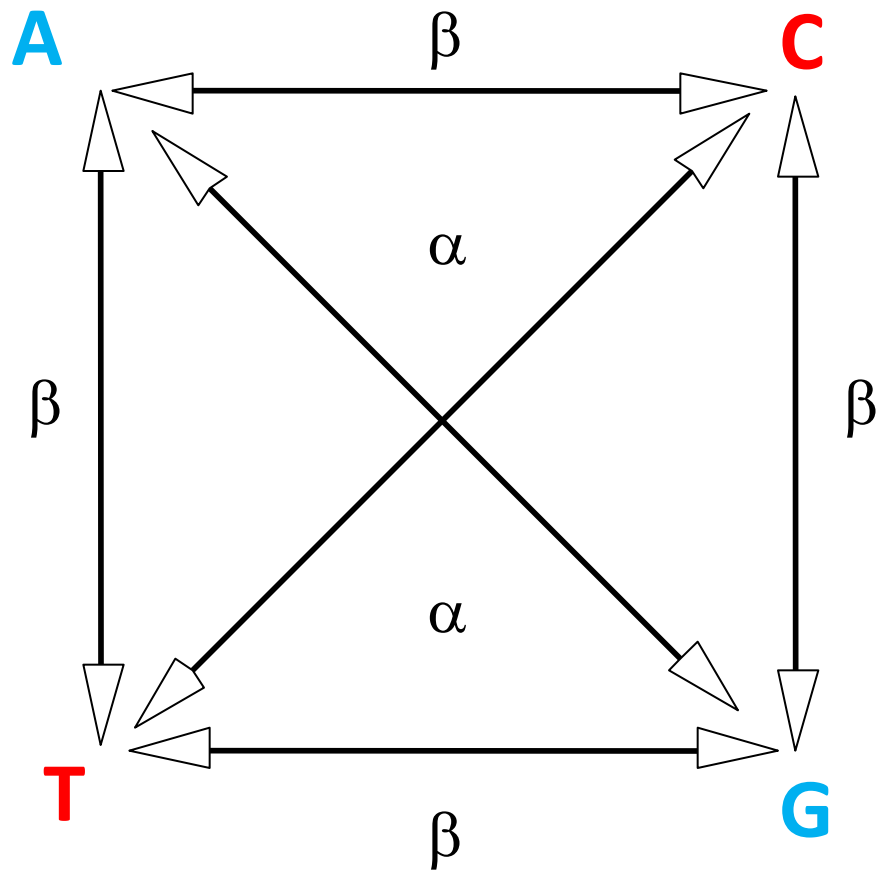
Models of DNA Substitution

- Models of DNA sequence evolution are required to recover the missing information through correcting for multiple substitutions.
 - i. The probability of substitution between bases (e.g. A to C, C to T...)
 - ii. The probability of substitution along a sequence (different sites/regions evolve at different rates)



All substitutions occur at the same rate (α)

Is this model too simple for real data?



Transitions (α) and transversions (β) occur at a different rate

Models of DNA Substitution

*Simplest
(few parameters)*

1. Base frequencies are equal and all substitutions are equally likely
(Jukes-Cantor)



2. Base frequencies are equal but transitions and transversions occur at different rates
(Kimura 2-parameter)



3. Unequal base frequencies and transitions and transversions occur at different rates
(Hasegawa-Kishino-Yano)



*Most complex
(many parameters)*

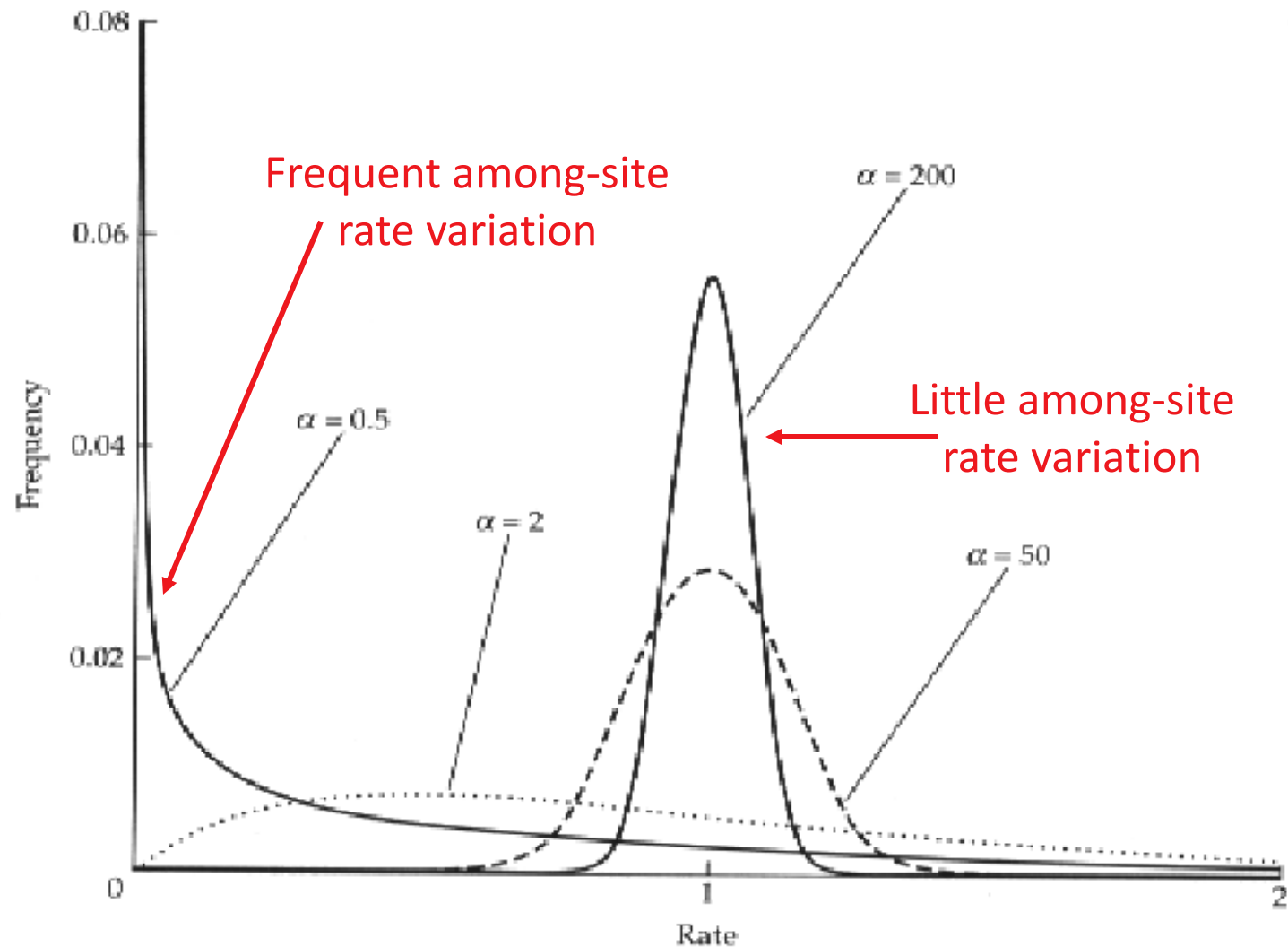
4. Unequal base frequencies and all substitution types occur at different rates
(General Reversible Model)

*All these models can be tested using the program jMODELTEST
(darwin.uvigo.es/software/jmodeltest.html)*

Models of DNA Substitution

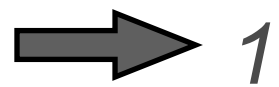
- i. The probability of substitution between bases
(e.g. A to C, C to T...)
- ii. The probability of substitution along a sequence
(different sites/regions evolve at different rates)

A Gamma Distribution Can be Used to Model Among-Site Rate Heterogeneity

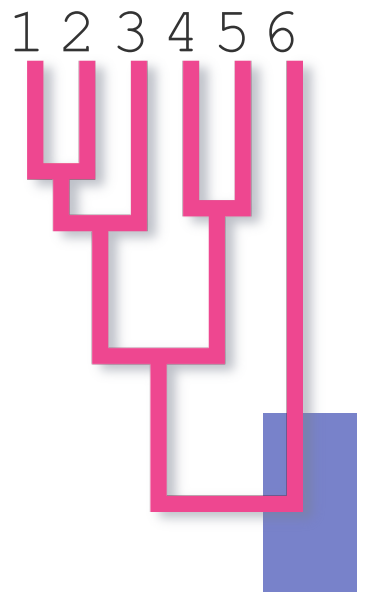
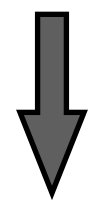
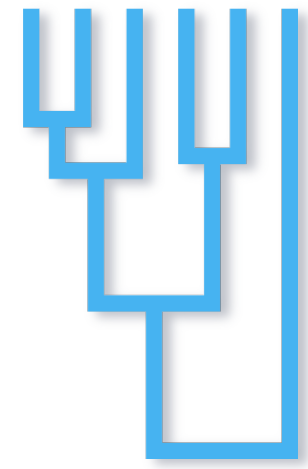
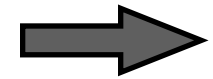


Non-Parametric Bootstrapping

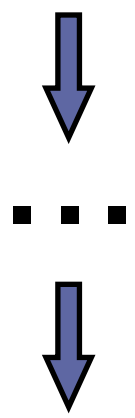
| | | | | | | |
|---|---|---|---|---|---|---|
| 1 | A | C | C | T | G | G |
| 2 | A | C | C | T | G | G |
| 3 | A | C | C | T | A | C |
| 4 | A | T | C | T | A | T |
| 5 | A | T | C | T | A | T |
| 6 | A | T | G | G | A | A |



| | | | | | | |
|---|---|---|---|---|---|---|
| 1 | G | G | G | A | T | C |
| 2 | G | G | G | A | T | C |
| 3 | A | A | C | A | T | C |
| 4 | A | A | T | A | T | T |
| 5 | A | A | T | A | T | T |
| 6 | A | A | A | A | G | T |

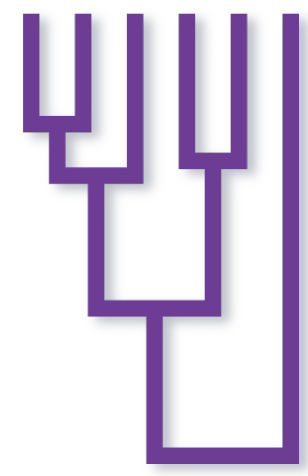
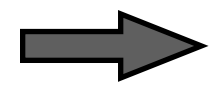


1000



Resample with replacement multiple times

| | | | | | | |
|---|---|---|---|---|---|---|
| 1 | T | G | C | C | A | G |
| 2 | T | G | C | C | A | G |
| 3 | T | C | C | C | A | A |
| 4 | T | T | C | C | A | A |
| 5 | T | T | C | C | A | A |
| 6 | G | A | G | G | A | A |



LARGE-SCALE PHYLOGENETIC ANALYSIS

1. Genome Concatenation or Whole Genome alignments
2. SNP calling with reference to a matrix

LARGE-SCALE STANDARD PHYLOGENETIC ANALYSIS

Concatenation

```
TAAATGTTACGGCACATGCTTTATTAATAATTTTTATAGTAATACCT
GCTTAATACAGCCCAAGGCTCTTCTAATAATATCTCGAGTCCCTCCCT
GCTTAATACAGCCCAAGGACTGCTTATGATGTTCTTGTAGTCTCCCT
GCTTAATACAGCCCAAGGCTTCTAATAATATTTTTGTAGTCTCCCA
GTTTAATACAGCCCAAGGCTACTTATGATGTTCTTGTAGTCTCCCT
GCTTATACAGCTCATGGGCTTCTAATAATATCTTGTAGTCTCCCT
GCTTAATACAGCTCATGGACTACTAATAATATTTTTGTAGTCTCCCT
GCTTATACAGCTCATGGCTTATTAATGATATTTTTGTAGTCTCCCT
GCTTATACAGCTCATGGCTTATTAATGATATTTTTGTAGTCTCCCT
GCTTATACAGCTCATGGACTTATTAATAATTTTTGTAGTAAACCT
```

Gene 1

```
TCCGACATAGCATTCCTCCAGAAATAATATAAGATTTGATTACTAC
GCCAGATATGGCTTTCCTCGATTAATAAATCAGGATTTGACTAATTC
ACCTGATATGGCTTTCCTCGAATAATAAATCAGGATTCGATTAATTC
CCCAACATAGCTTTCCTCCAGTATTAATAATCTTGGATTTGACTAGTTC
ACCTGACATGGCTTTCCTCCAGAAATAATAATCAGGATTCGACTAATTC
CCCCGATATAGCTTTCCTCCAGAAATAATAATCAGGATTTGACTAATTC
ACCAACATGGCTTTCCTCGAATAATAAATCAGGATTTGATTAATTC
TCCAACATGGCTTTCCTCCAGTATTAATAATCTTGGGTTCTGACTTATTC
TCCAACATGGCTTTCCTCCAGTATTAATAATCTTGGGTTCTGACTTATTC
TCCAACATAGCTTTCCTCCAGGATTAATAATAAGGATTTGACTACTTC
```

Gene 2

```
AGAAATTAAGCAACAGGAGCATTAATTTGAGATGATCAAAATTTAATG
TGAATTTGGCAACCTGGGACTCTTTTAGGTGATGACCAAACTATAAAT
GGAACTAGGACAAACAGGAACCTTTTAGGAGACGACCAAAATTAACAAT
AGAGCTAGGTCAACCTGGTACACTTTAGGAGATGACCAAAATTAACAAT
TGAATTTGGCAACAGGAACCTTTTAGGAGATGACCAAAATTAACAAT
TGAATTTGGCAACAGGACCTTTAGGAGATGACCAAAATTAACAAT
TGAATTTGGCAACAGGAACTCTTTAGGAGATGACCAAAATTAACAAT
TGAATTTGGCAACAGGAACTCTTTAGGAGATGACCAAAATTAACAAT
TGAATTTGGCAACAGGAACTCTTTAGGAGATGACCAAAATTAACAAT
TGAATTTGGCAACAGGATCTTTTAGGATGATCAAAATTAACAAT
```

Gene 3

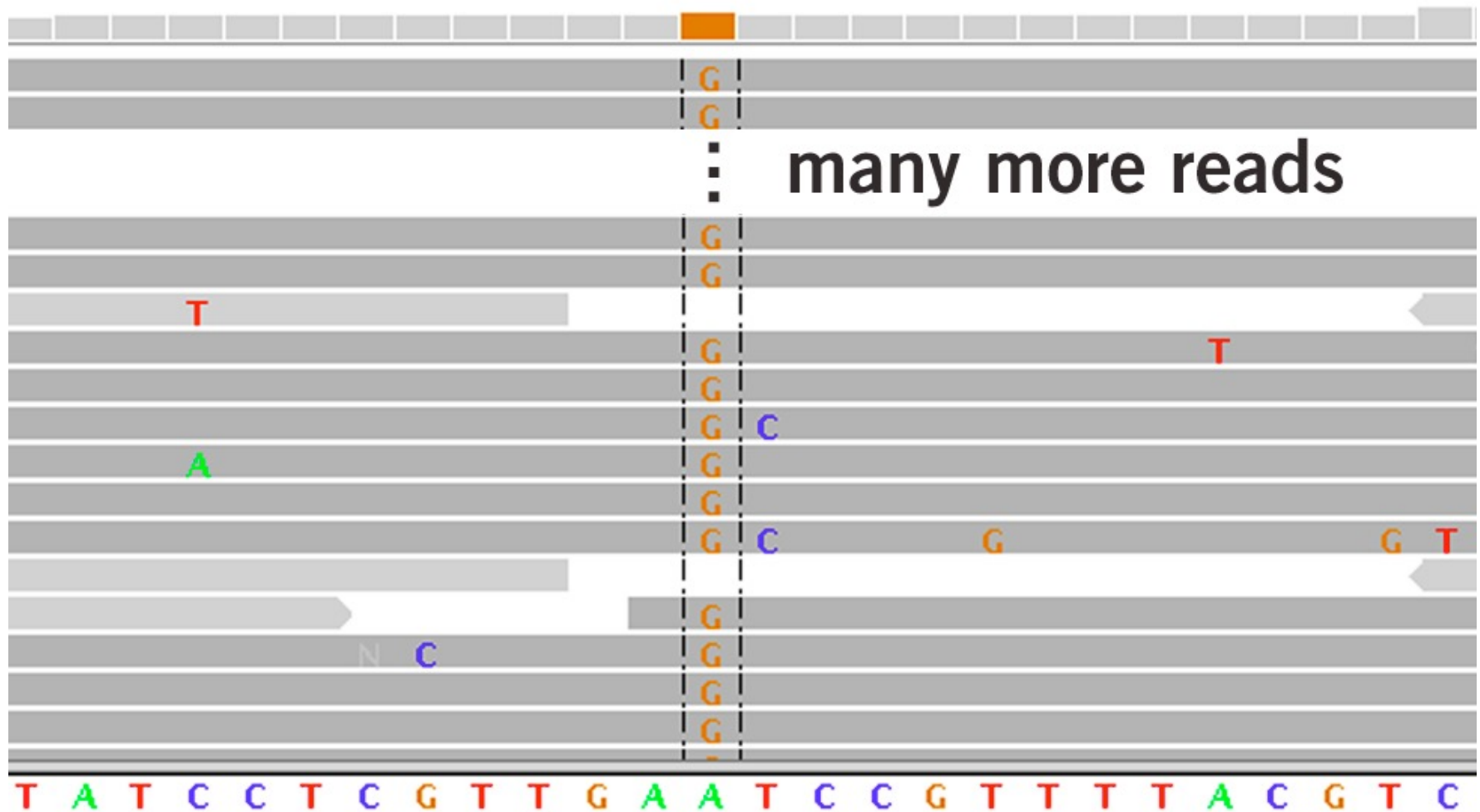
```
CTCCTGCTCTTCTTTATTATTAGTAAGAAATAGTTGAAAAAGGAGCT
CTCCCGCAGTATCTCTCTTGTAAATATCAGCTTTATGAAAAAGGTGCC
CTCCAGCAGTAAATCTTCTAGTAATATCTGCTTTATGAAAAAGGAGCT
CTCCAGCAGTAAATCTTCTAGTAATATCTGCTTTATGAAAAAGGTGCC
CTCCTGCTCTTCTTTATTATTAGTAAGAAATAGTTGAAAAAGGAGCT
CTCCCGCAGTATCTCTCTTGTAAATATCAGCTTTATGAAAAAGGTGCC
CTCCCGCAGTAAATCTTCTAGTAATATCTGCTTTATGAAAAAGGTGCC
CTCCCGCAGTAAATCTTCTAGTAATATCTGCTTTATGAAAAAGGTGCC
CTCCCGCAGTAAATCTTCTAGTAATATCTGCTTTATGAAAAAGGTGCC
CTCCCGCAGTAAATCTTCTAGTAATATCTGCTTTATGAAAAAGGTGCC
```

Gene 4

[illegible]

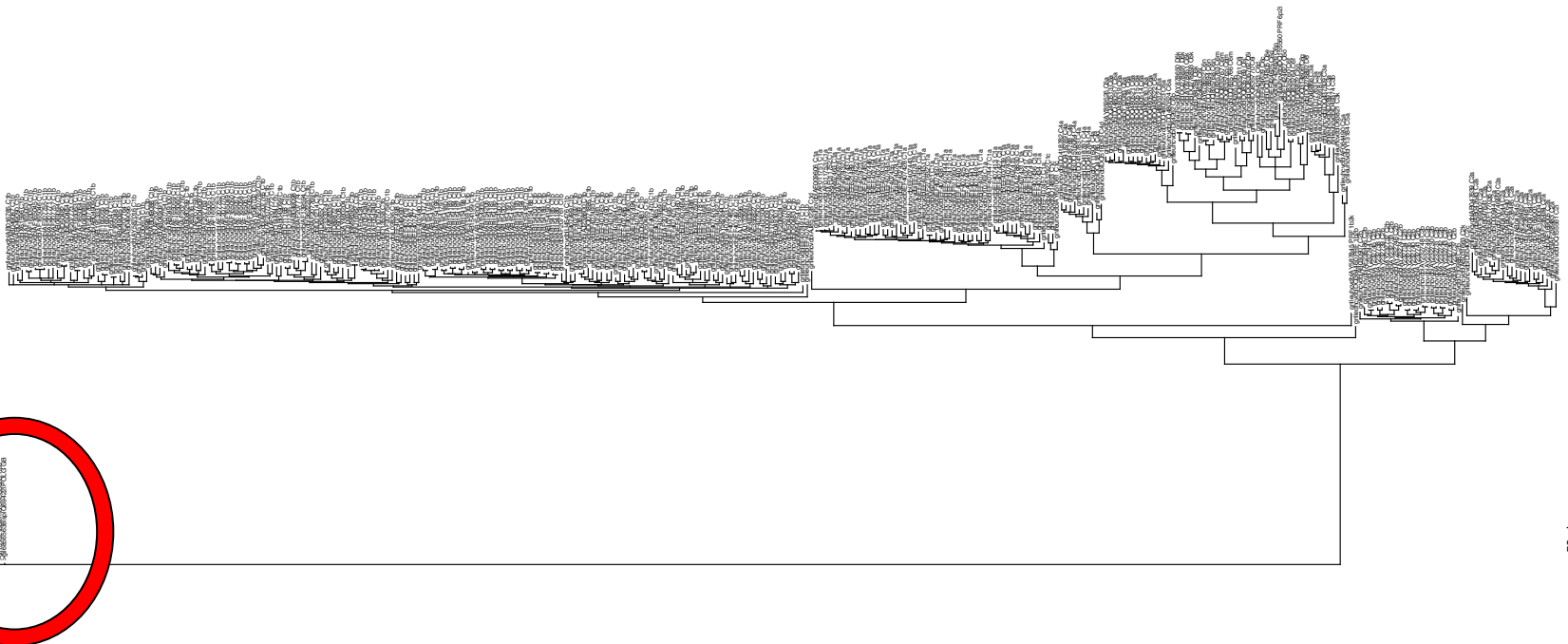
LARGE-SCALE PHYLOGENETIC ANALYSIS (Reference Based)

SNP Calling (final matrix has only invariant positions=danger
ascertainment bias)



Things to think about when you look at a tree

1. How is it rooted?
2. What do the branch lengths mean?
3. Do the branch lengths make sense?



AFSW3168-3
AFSW3168-2
AFSW3624-4
AFSW3481-4
AF1746
AF2391
AF1744
GCF 004011165.1 ASM401116v1
GCF 003309865.1 ASM330986v1
AFSW3538-3
AF2068
AF2070
AF2069
AFSW3937-4
AFSW3701-5
AFSW3120-1
AFSW3248-1
GCF 004012035.1 ASM401203v1
GCF 004681195.1 ASM468119v1

MRN Key:

10988
12371
15787
19135
19822
20004
25576
26019

RAxML Settings:

Model of Binary: GTRCAT (GTR + Optimization of substitution rates + Optimization of site-specific evolutionary rates which are categorized into numberOfCategories distinct rate categories for greater computational efficiency.

Bootstrapping Criteria: autoMRE (for the majority-rule tree based criteria)

BFGS: Disabled automatic usage of BFGS method to optimize GTR rates on unpartitioned DNA datasets

Ascertainment Bias Correction: Paul Lewis Method

Rapid Bootstrap Number Seed: 123

Parsimony Random Seed: 256

GCF 000636555.1 Stap aure 1110904178 V1
AFSW3700-3
GCF 000296595.1 ASM29659v1
GCF 000641355.1 Stap aure VET0105R V1
GCF 000642735.1 Stap aure VET0259R V1
GCF 003353195.1 ASM335319v1
GCF 003353135.1 ASM335313v1
GCF 003353145.1 ASM335314v1
GCF 002208475.1 ASM220847v1
hp20814-024
AF2137
GCF 002204575.1 ASM220457v1
GCF 001887075.1 ASM188707v1
GCF 002273545.1 ISU-924 v1.0
GCF 002275205.1 ISU-922 v1.0
AFSW3174-1
AFSW3340-1
AFSW3521-1
AFSW3672-4
AFSW3820-1
AFSW3330-1
AFSW3170-1
GCF 000239555.1 ASM23955v2
GCF 000443285.1 S94v1
GCF 000638635.1 Stap aure C2679 V1
GCF 002208395.1 ASM220839v1
GCF 001696365.1 ASM169636v1
GCF 002208455.1 ASM220845v1
GCF 000638795.1 Stap aure C3965 V1
GCF 000638615.1 Stap aure C2549 V1
GCF 003720975.1 ASM372097v1
GCF 000258685.1 ASM25868v1
GCF 000252405.1 ASM25240v2
AF2575
AF2573
AF2572
AF2574
GCF 000639575.1 Stap aure PA11 V1
GCF 000639595.1 Stap aure PA57 V1
GCF 000636895.1 Stap aure 122 V1
GCF 000638875.1 Stap aure C5086 V1
AFSPU3290-1
AFSW3325-1
GCF 002089035.2 ASM208903v2
AFSW3313-1
AFSW3877-1
AFSW3687-1
AFSW3544-5
AFSW3783-4
AFSW3806-4
AFSW3614-4
AFSPU3614-4
AFSW3008-2
AFSW3104-2
AFSW3474-3
AFSPU3296-1
AFSW3296-1