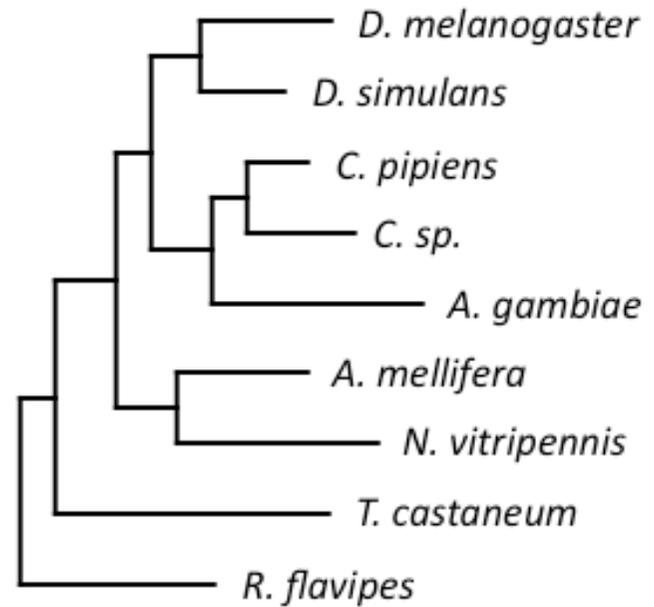
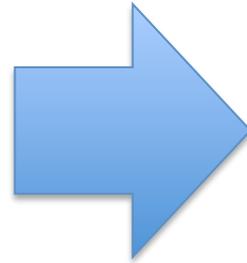
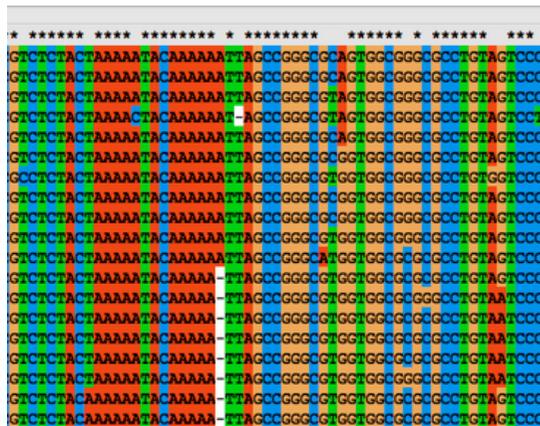
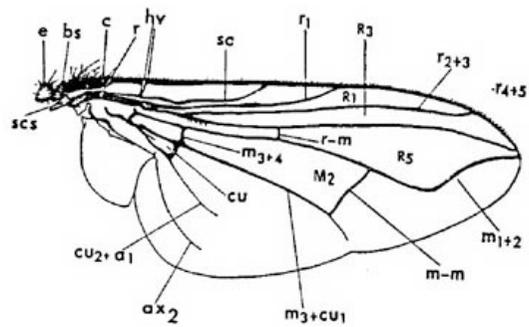


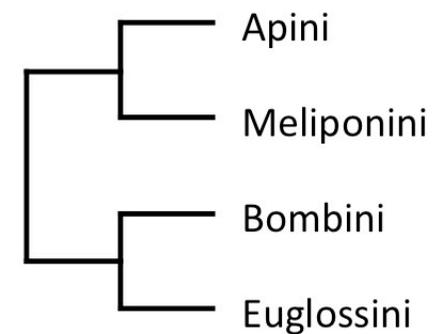
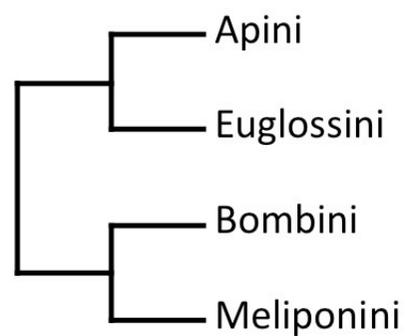
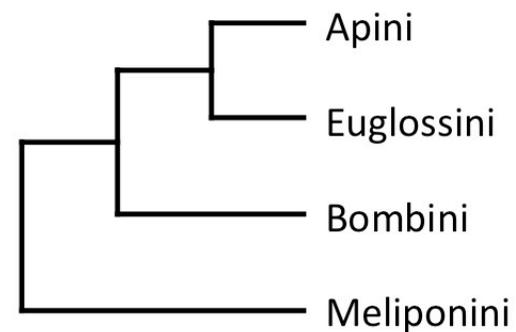
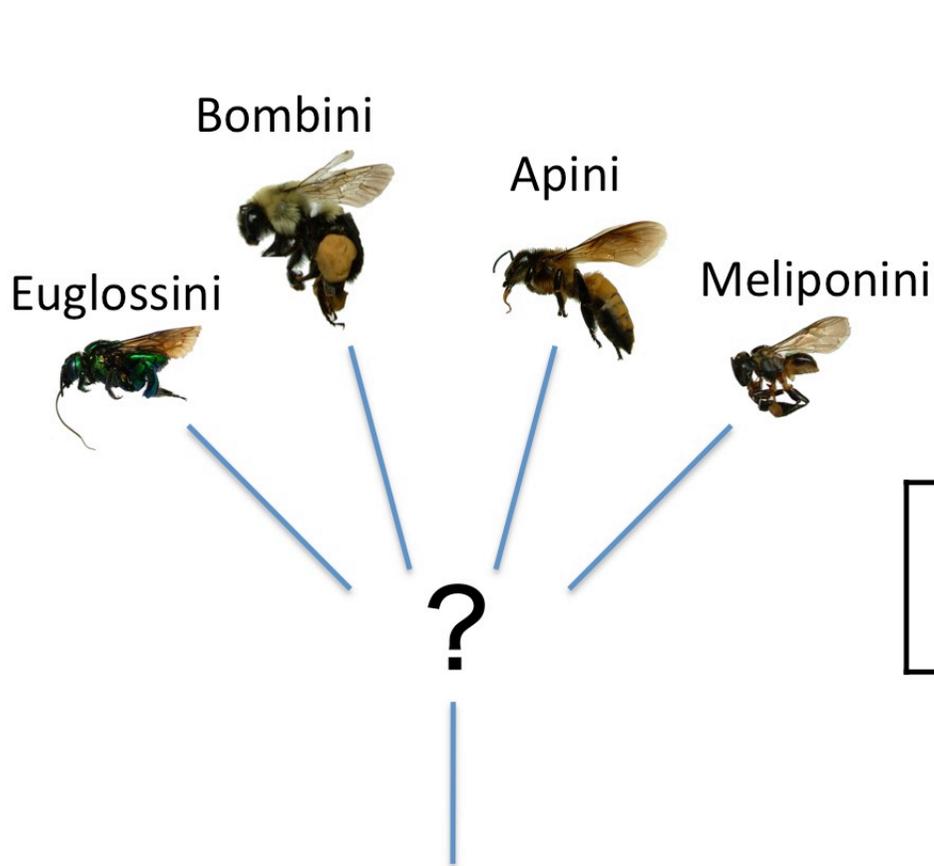
「系統解析」

川北 篤

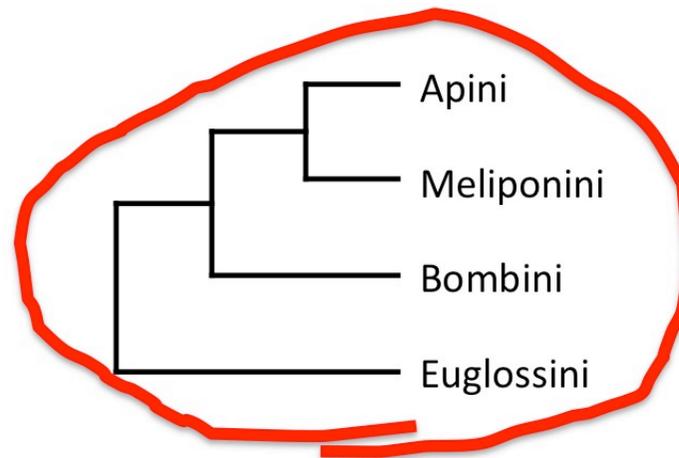


今日の話の内容

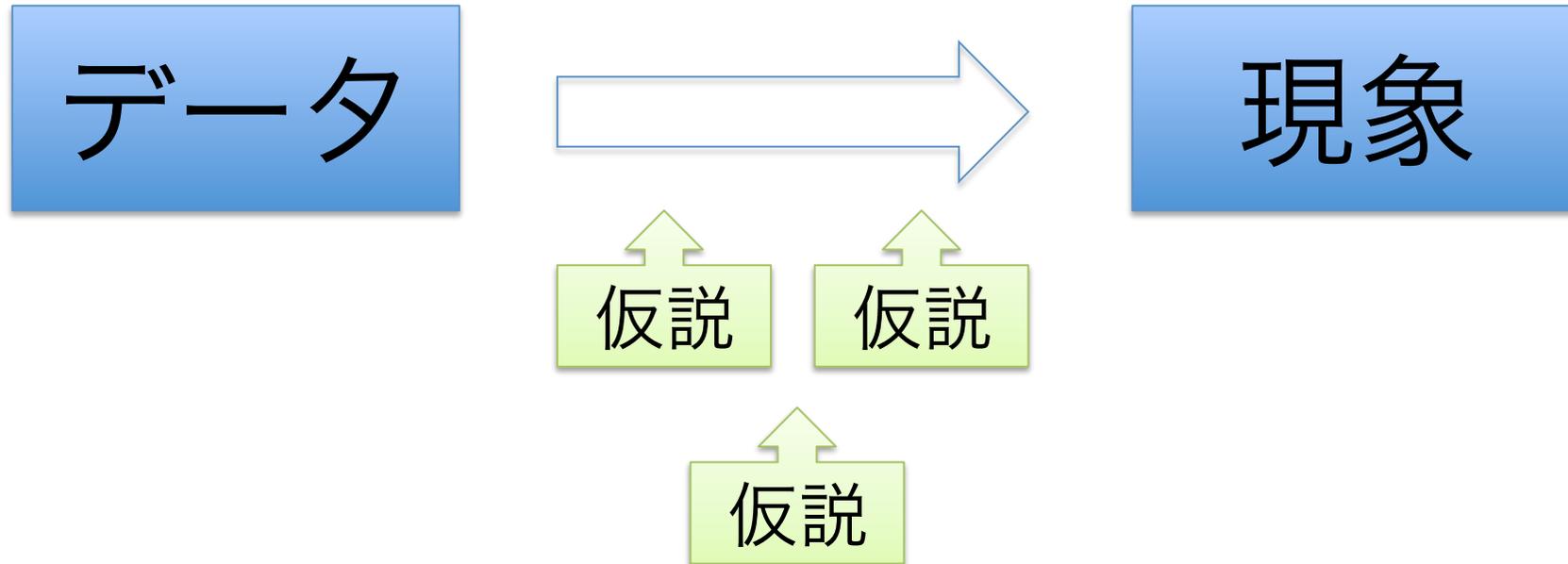
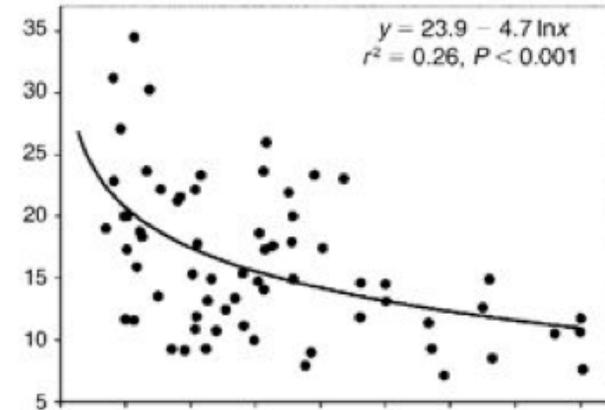
1. 系統解析とは何か？
2. どのように系統樹を求めるか？
3. 系統樹を用いたさまざまな解析
4. 種を単位とした統計の問題点
5. 系統的独立対比を用いた解析



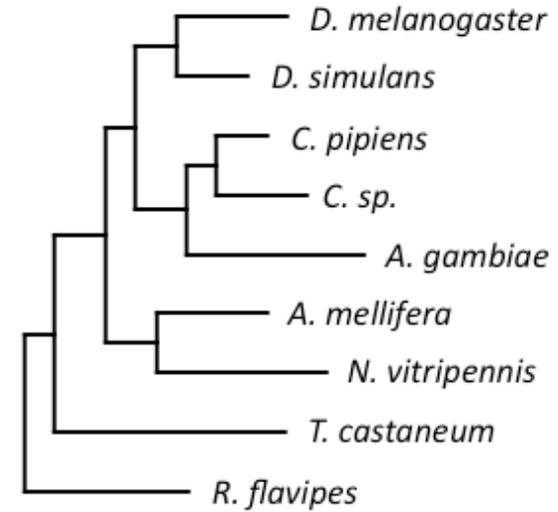
“真の系統樹”



先週までの統計解析



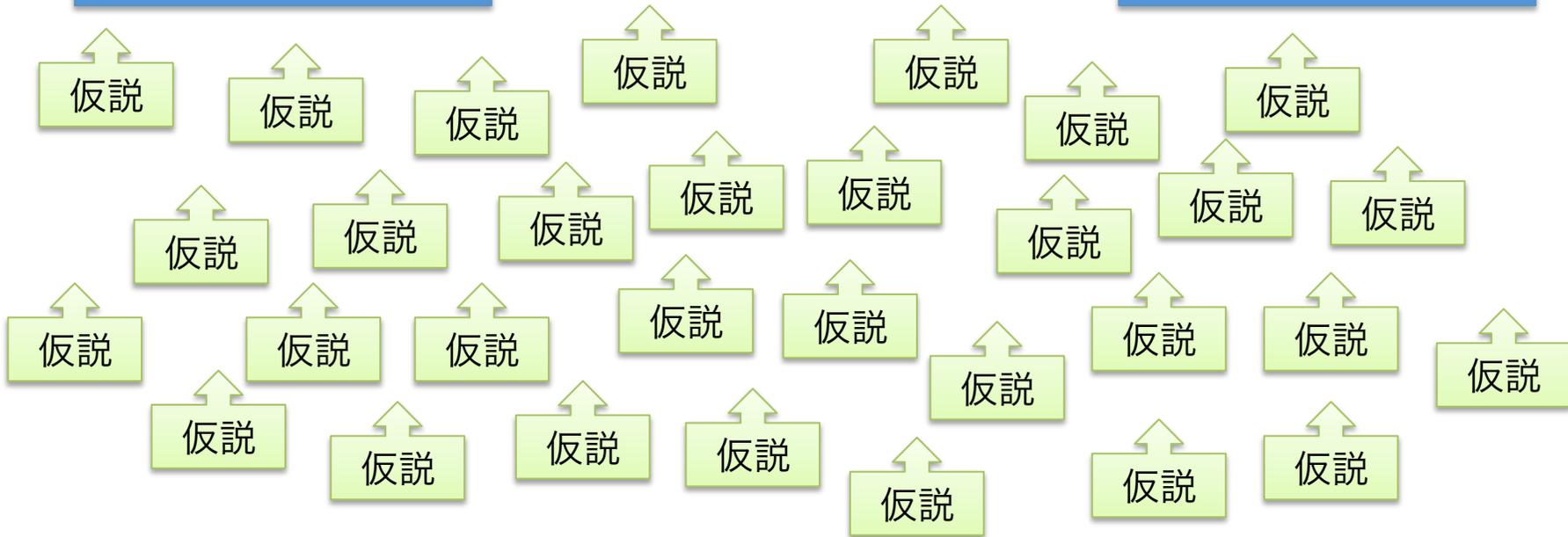
系統解析は？



データ



現象



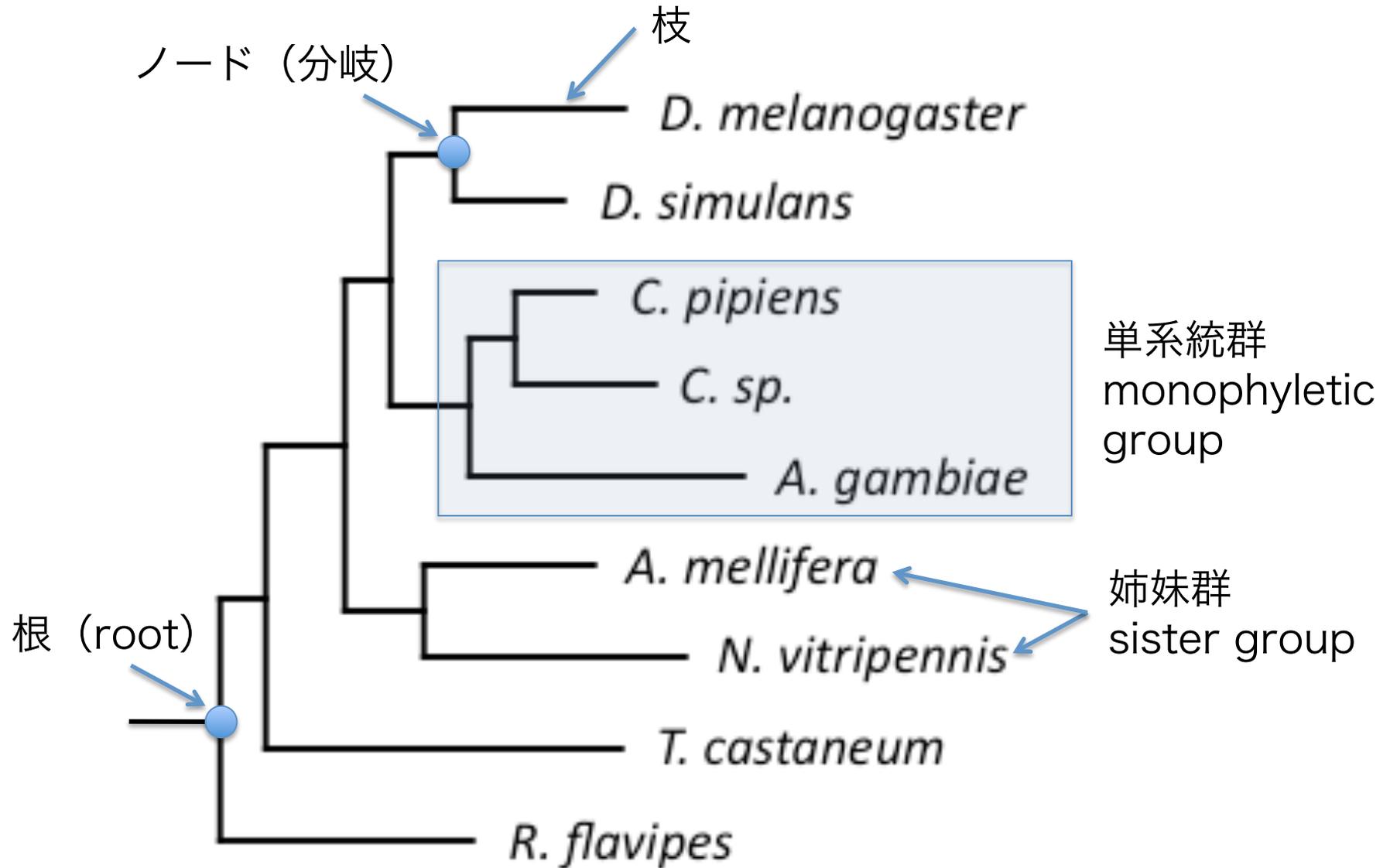
系統樹の本数

Table 3.1: The number of rooted, bifurcating, labeled trees for n species, for various values of n . The numbers for more than 20 species are approximate.

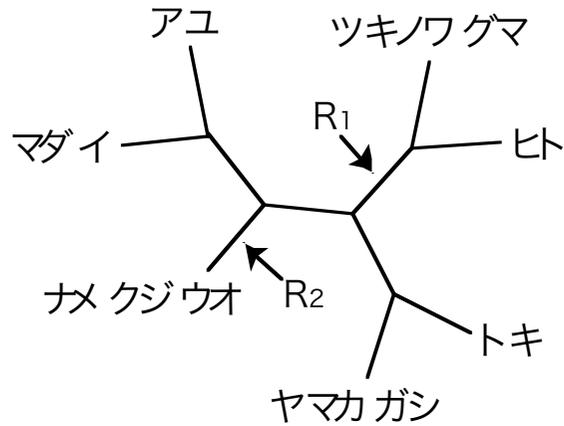
| Species | Number of trees |
|---------|-------------------------------|
| 1 | 1 |
| 2 | 1 |
| 3 | 3 |
| 4 | 15 |
| 5 | 105 |
| 6 | 945 |
| 7 | 10,395 |
| 8 | 135,135 |
| 9 | 2,027,025 |
| 10 | 34,459,425 |
| 11 | 654,729,075 |
| 12 | 13,749,310,575 |
| 13 | 316,234,143,225 |
| 14 | 7,905,853,580,625 |
| 15 | 213,458,046,676,875 |
| 16 | 6,190,283,353,629,375 |
| 17 | 191,898,783,962,510,625 |
| 18 | 6,332,659,870,762,850,625 |
| 19 | 221,643,095,476,699,771,875 |
| 20 | 8,200,794,532,637,891,559,375 |
| 30 | 4.9518×10^{38} |
| 40 | 1.00985×10^{57} |
| 50 | 2.75292×10^{76} |

Felsenstein (2004)
Inferring Phylogenies

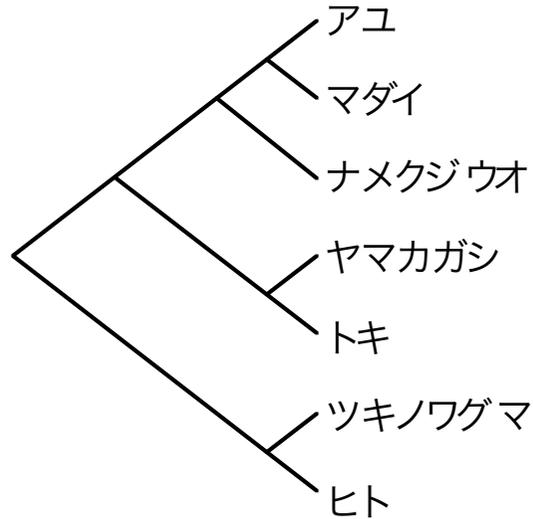
系統解析で使う用語



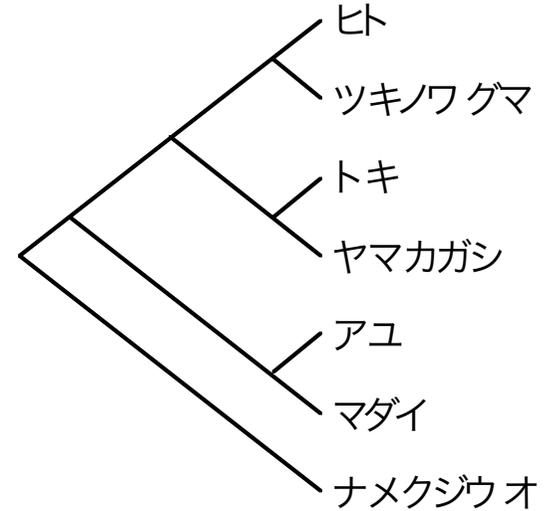
無根系統樹



R₁に根をつけた場合



R₂に根をつけた場合

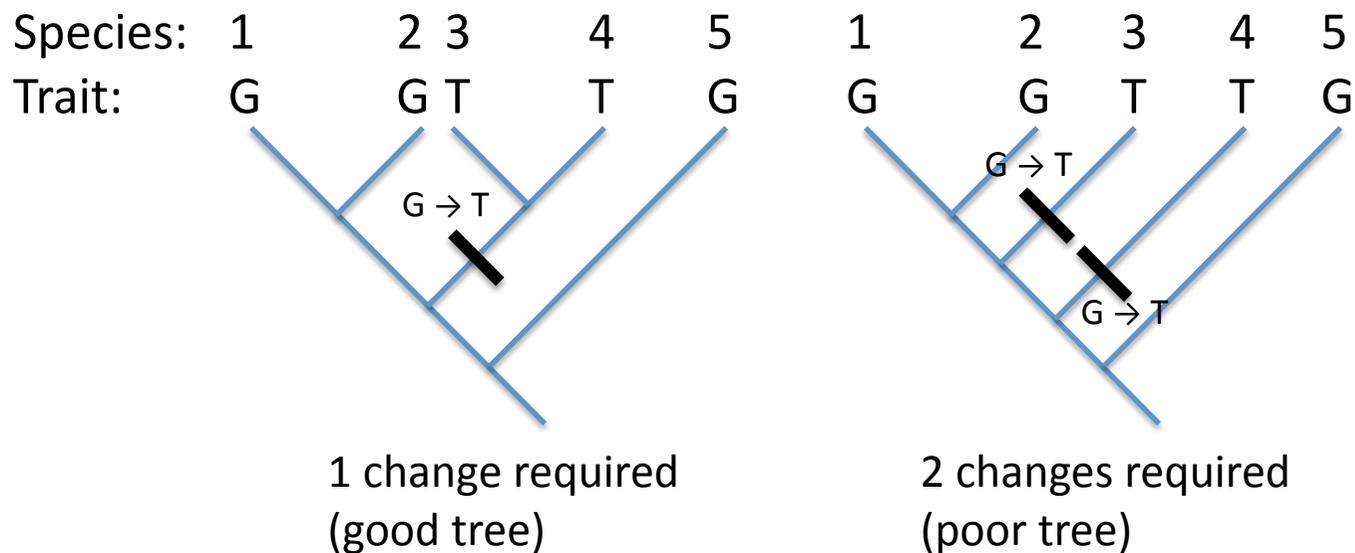


解析の対象となる分類群に含まれない生物を「外群」として解析に加え、事後的に根の位置を決める

系統樹の推定法

1. 最節約法 Maximum parsimony
2. 近隣結合法 Neighbor-joining
3. 最尤法 Maximum likelihood
4. (ベイズ法 Bayesian method)

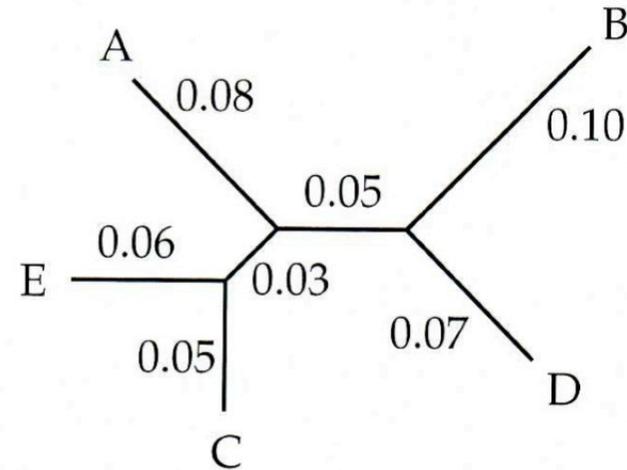
最節約法



- 原理が単純（形質を選ばない）
- 計算が速い
- 複数解が存在する
- 間違えることがある

近隣結合法 (距離法)

| | A | B | C | D | E |
|---|------|------|------|------|------|
| A | 0 | 0.23 | 0.16 | 0.20 | 0.17 |
| B | 0.23 | 0 | 0.23 | 0.17 | 0.24 |
| C | 0.16 | 0.23 | 0 | 0.15 | 0.11 |
| D | 0.20 | 0.17 | 0.15 | 0 | 0.21 |
| E | 0.17 | 0.24 | 0.11 | 0.21 | 0 |

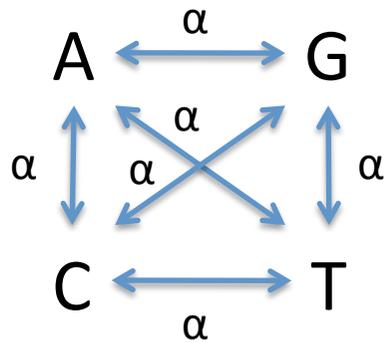


Felsenstein (2004)
Inferring Phylogenies

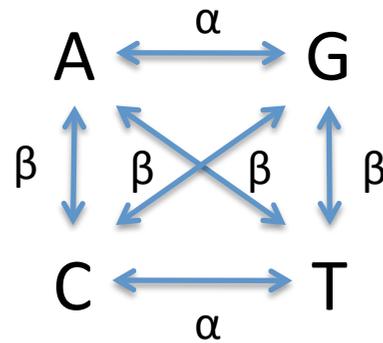
- 計算が速い
- 単一解
- 距離に変換する分、情報量が減る
- 距離の求め方に結果が依存

最尤法

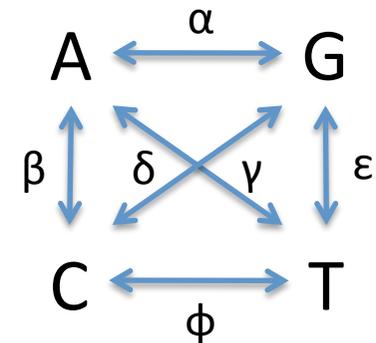
与えられた塩基置換モデルのもとで
データを最もよく説明する系統樹を選ぶ



Jukes-Cantor model



Kimura's 2 parameter
model

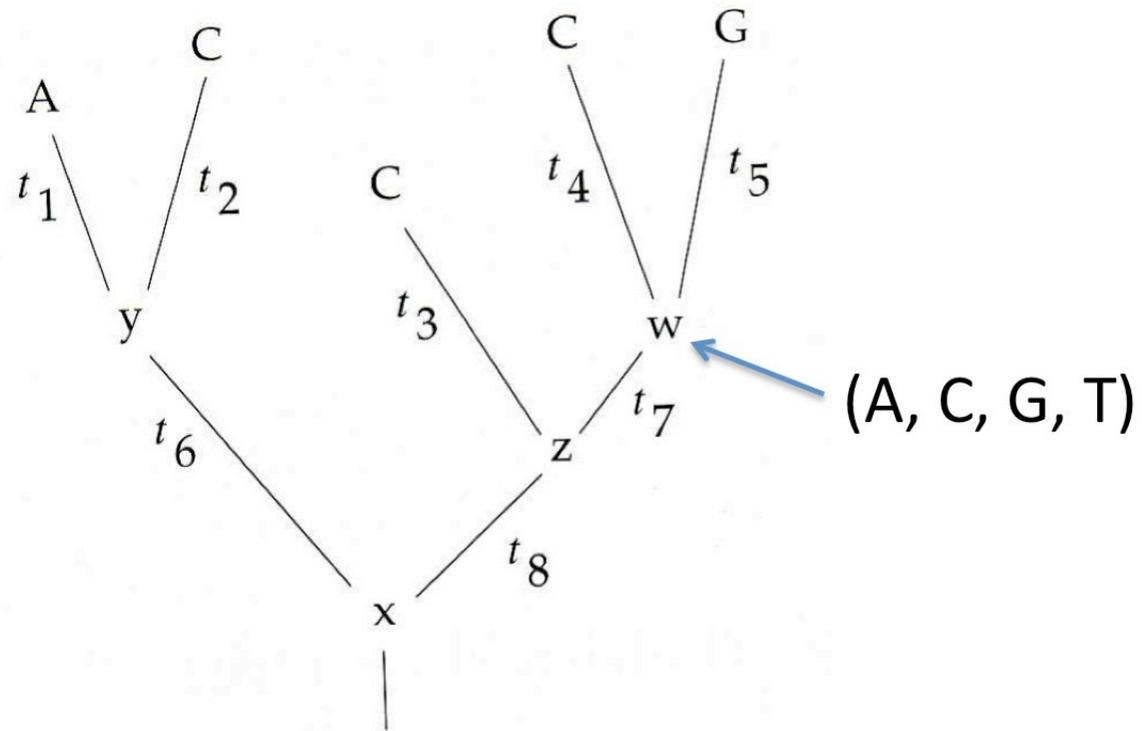


General-time-reversible
model

データを最もよく説明する塩基置換モデルを選ぶ
AICなどを使う

最尤法

尤度 = モデルのもとでデータが実現する確率



推定の信頼性を評価する ブートストラップ解析

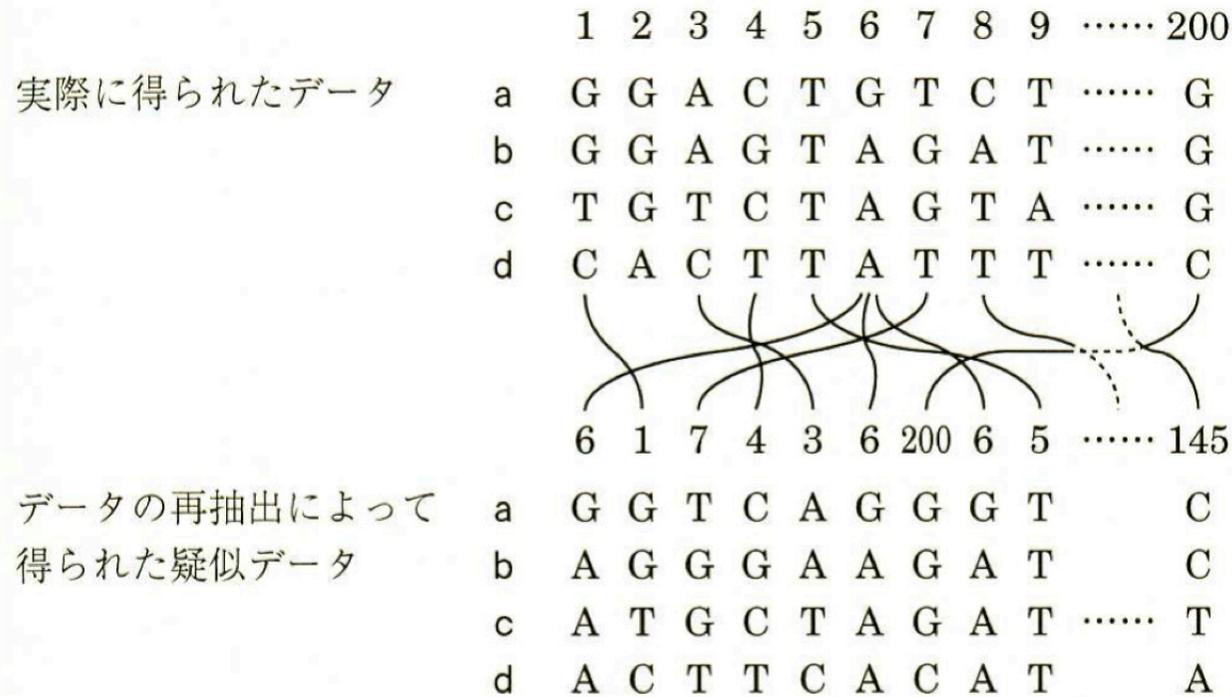


図 3-11 a~d は OTU を, A, C, G, T は塩基を示す

上島 (2004)

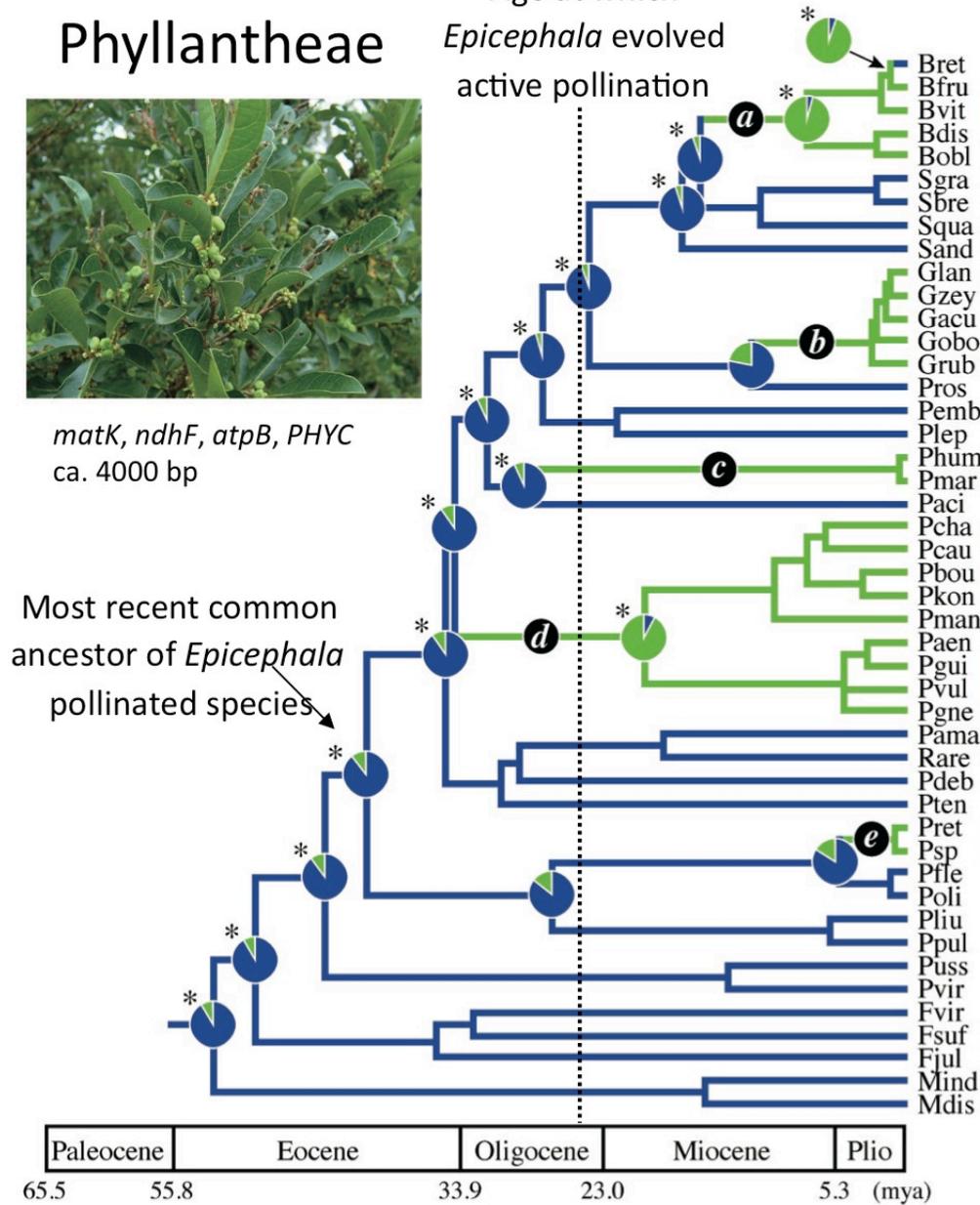
データがどれほど一貫して仮説を支持しているかの指標

Phyllanthaceae



matK, ndhF, atpB, PHYC
ca. 4000 bp

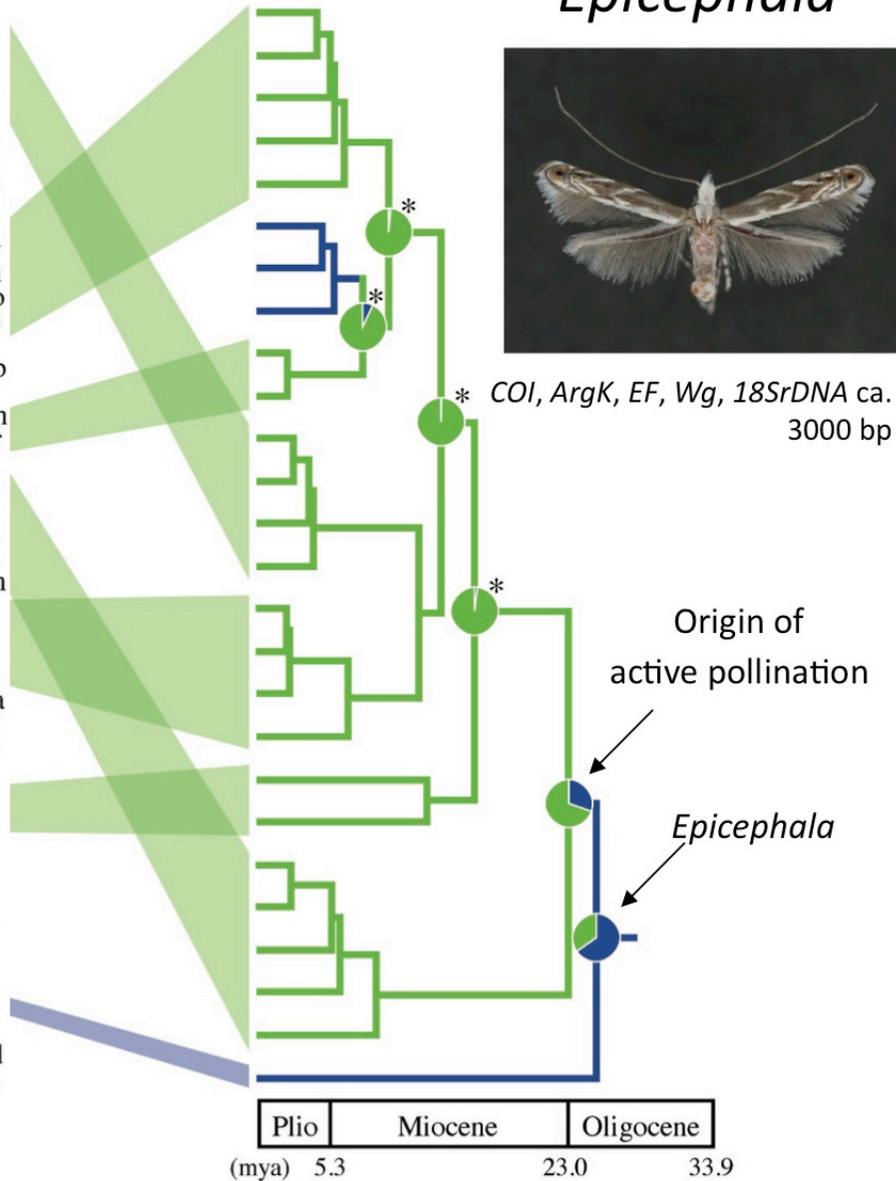
Age at which
Epicephala evolved
active pollination



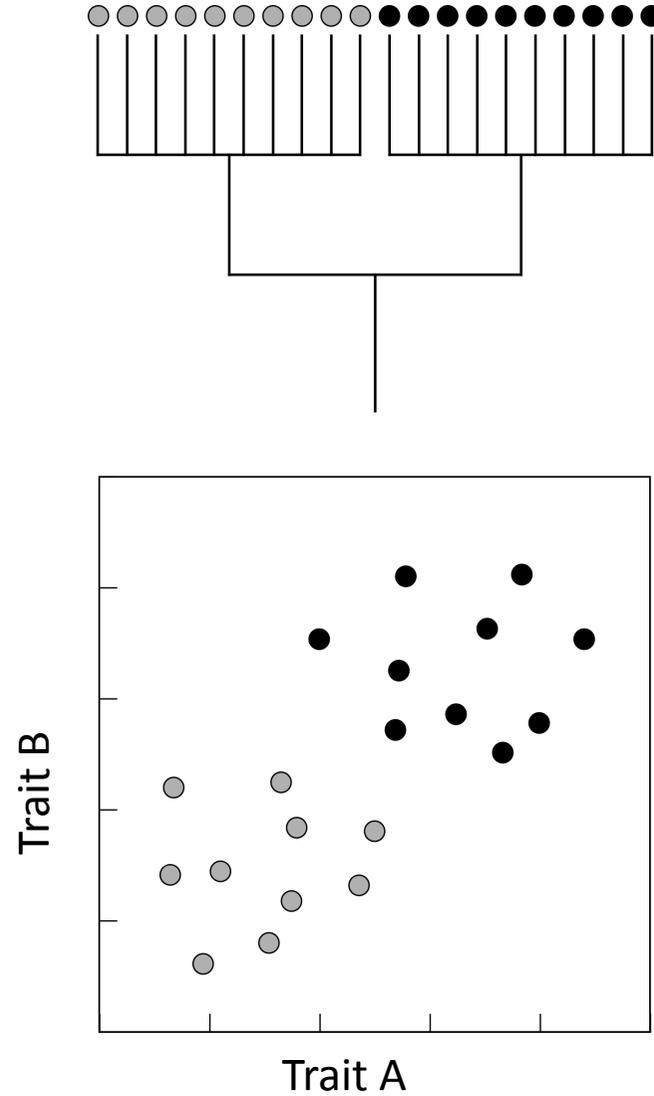
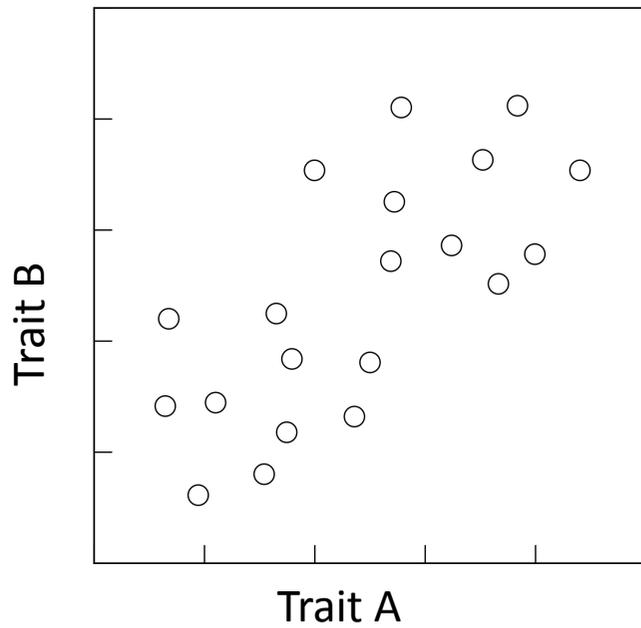
Epicephala

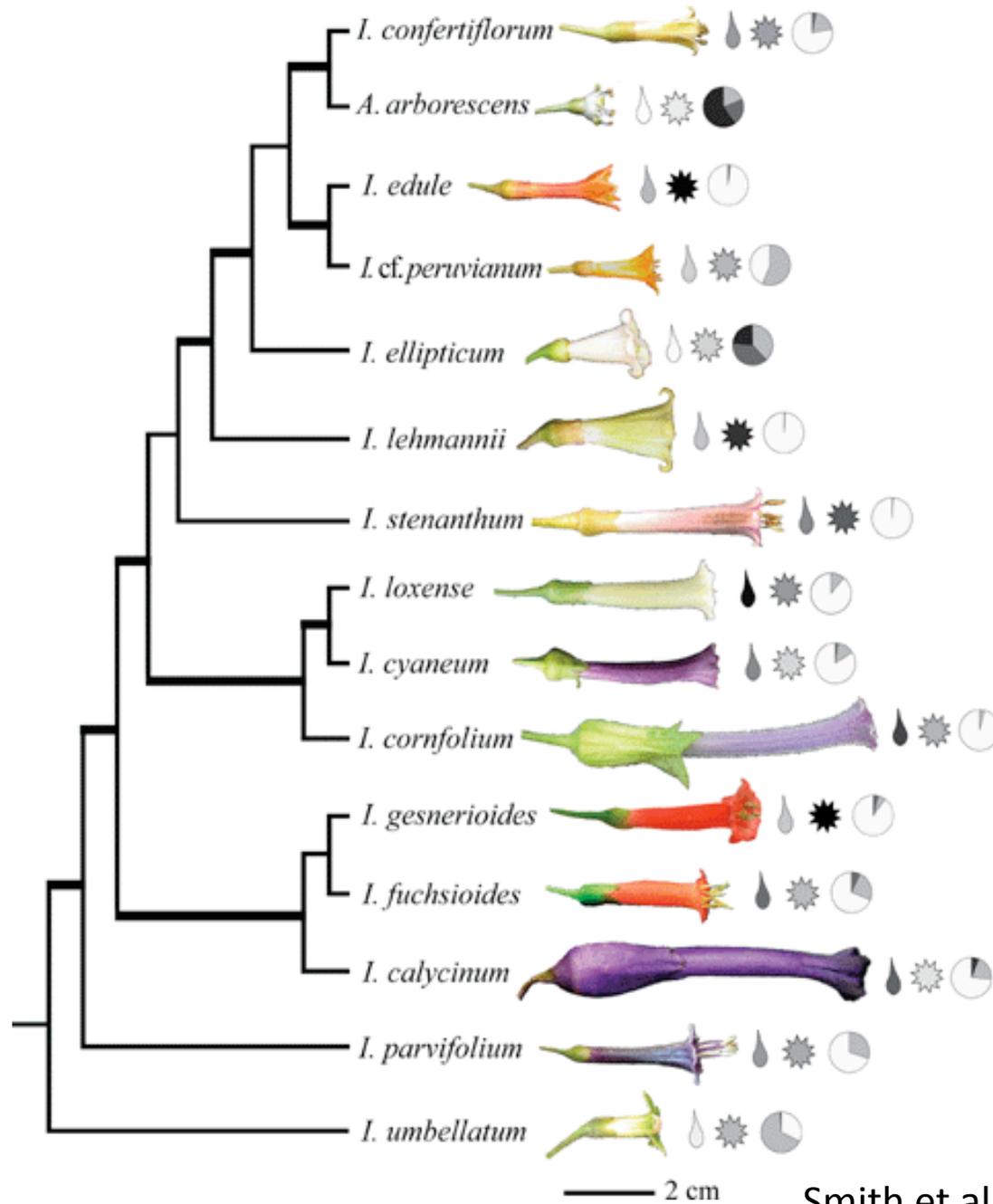


COI, ArgK, EF, Wg, 18SrDNA ca.
3000 bp



種を単位とした統計解析における問題

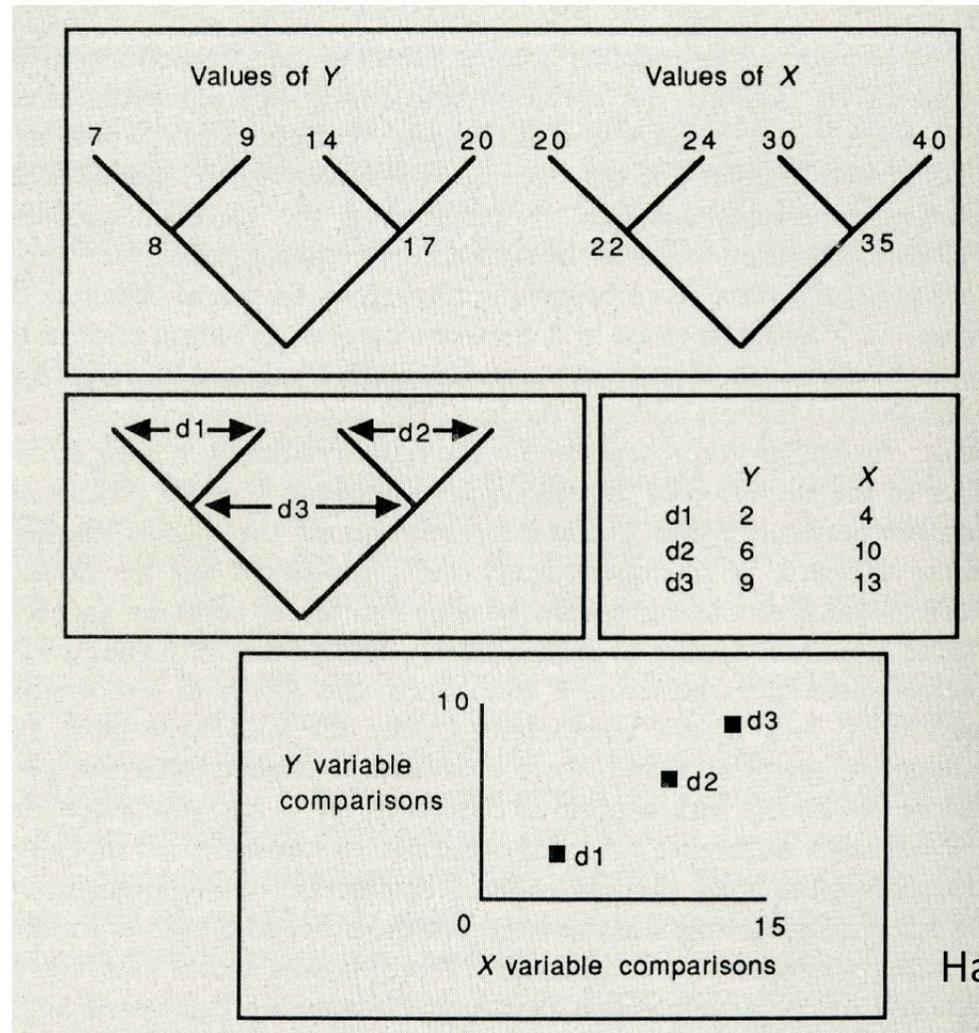




Smith et al. (2008) *Evolution*

系統的獨立對比

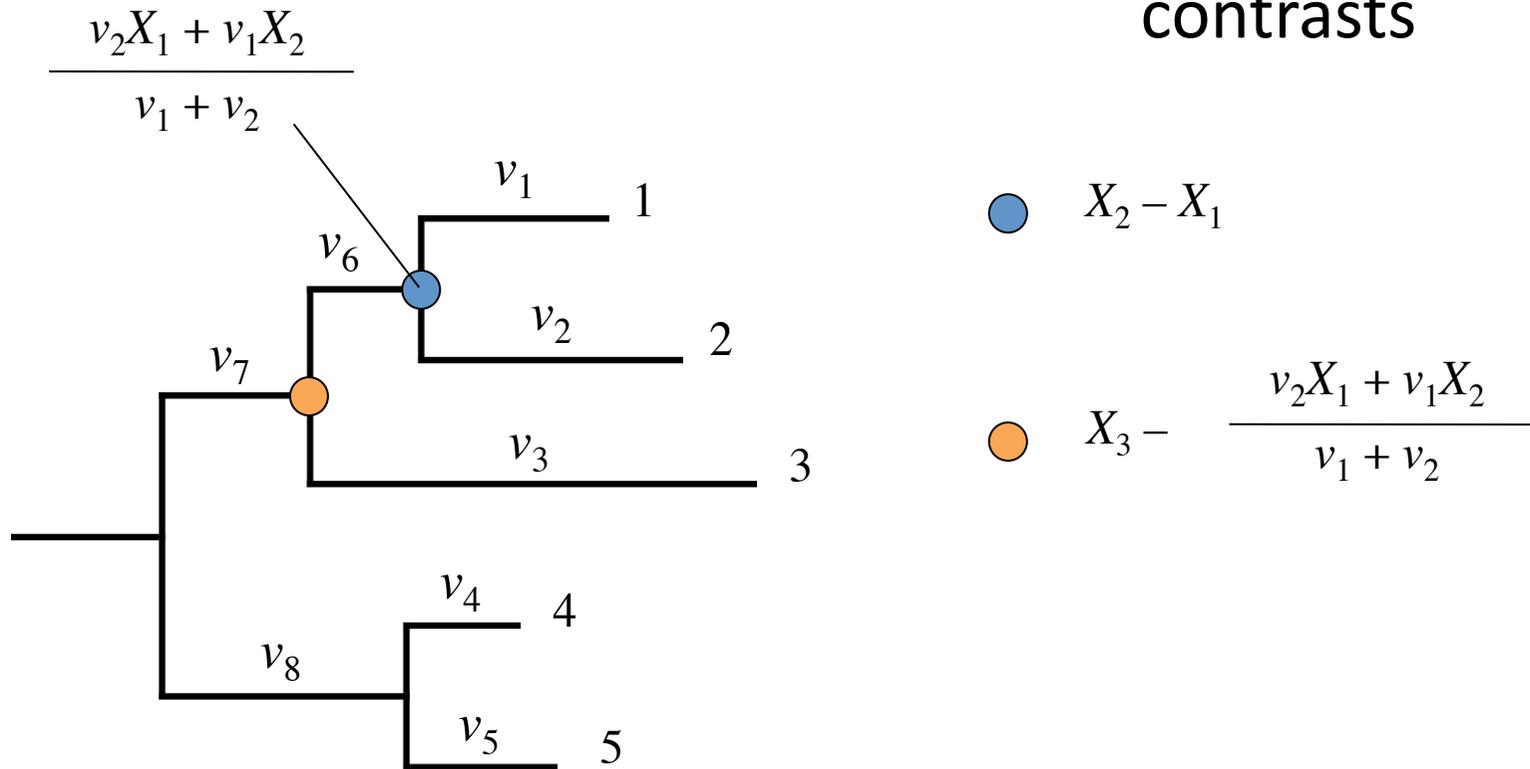
Phylogenetic independent contrast (Felsenstein, 1985)



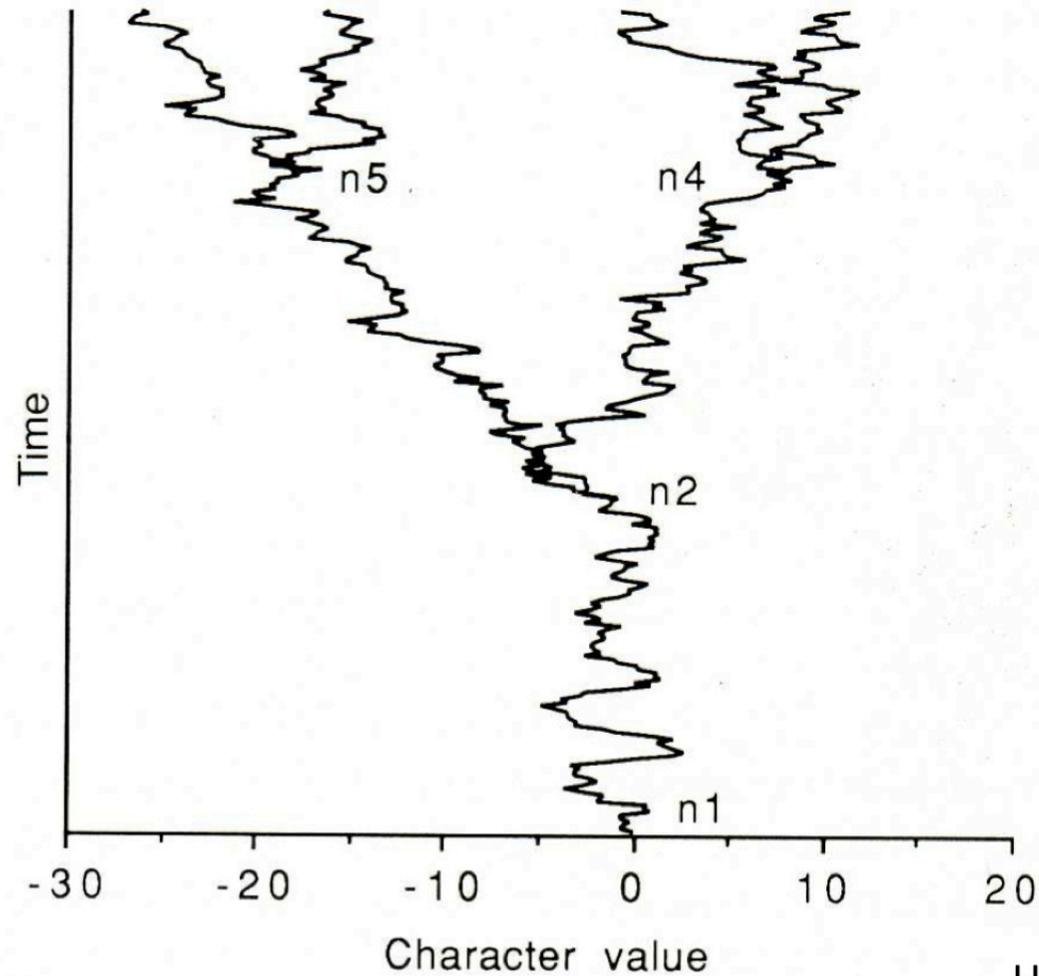
Harvey & Pagel (1991)

系統的獨立對比

Phylogenetic independent contrast



ブラウン運動モデル Brownian motion model



Harvey & Pagel (1991)

系統的 최소二乗法

Phylogenetic generalized least squares (PGLS) method

- 系統的独立対比の拡張版
- ブラウン運動モデルを補正し、系統の効果を調節する
- α parameter
 - 0のときブラウン運動モデルと同じ
 - 大きい値をとるほど系統の効果がなくなる

参考文献

- Felsenstein J (2004) *Inferring Phylogenies*. Sinauer Associates.
- Hall B (2007) *Phylogenetic Trees Made Easy: A How-to Manual, Third Edition*. Sinauer Associates.
- Harvey PH, Pagel MD (1991) *The Comparative Method in Evolutionary Biology*. Oxford University Press.
- Lemey P, Salemi M, Vandamme AM (2009) *The Phylogenetic Handbook: A Practical Approach to Phylogenetic Analysis and Hypothesis Testing, 2nd Edition*. Cambridge University Press.
- 三中信宏 (1997) 『生物系統学』 東京大学出版会.
- 奥山雄大・川北篤 (2008) 系統解析プロトコルー塩基配列から分子系統樹へ. 『共進化の生態学』 種生物学会編, 文一総合出版.
- 上島励 (1996) 系統樹をつくる. 『生物の種多様性』 岩槻邦男・馬渡峻輔編, 裳華房.