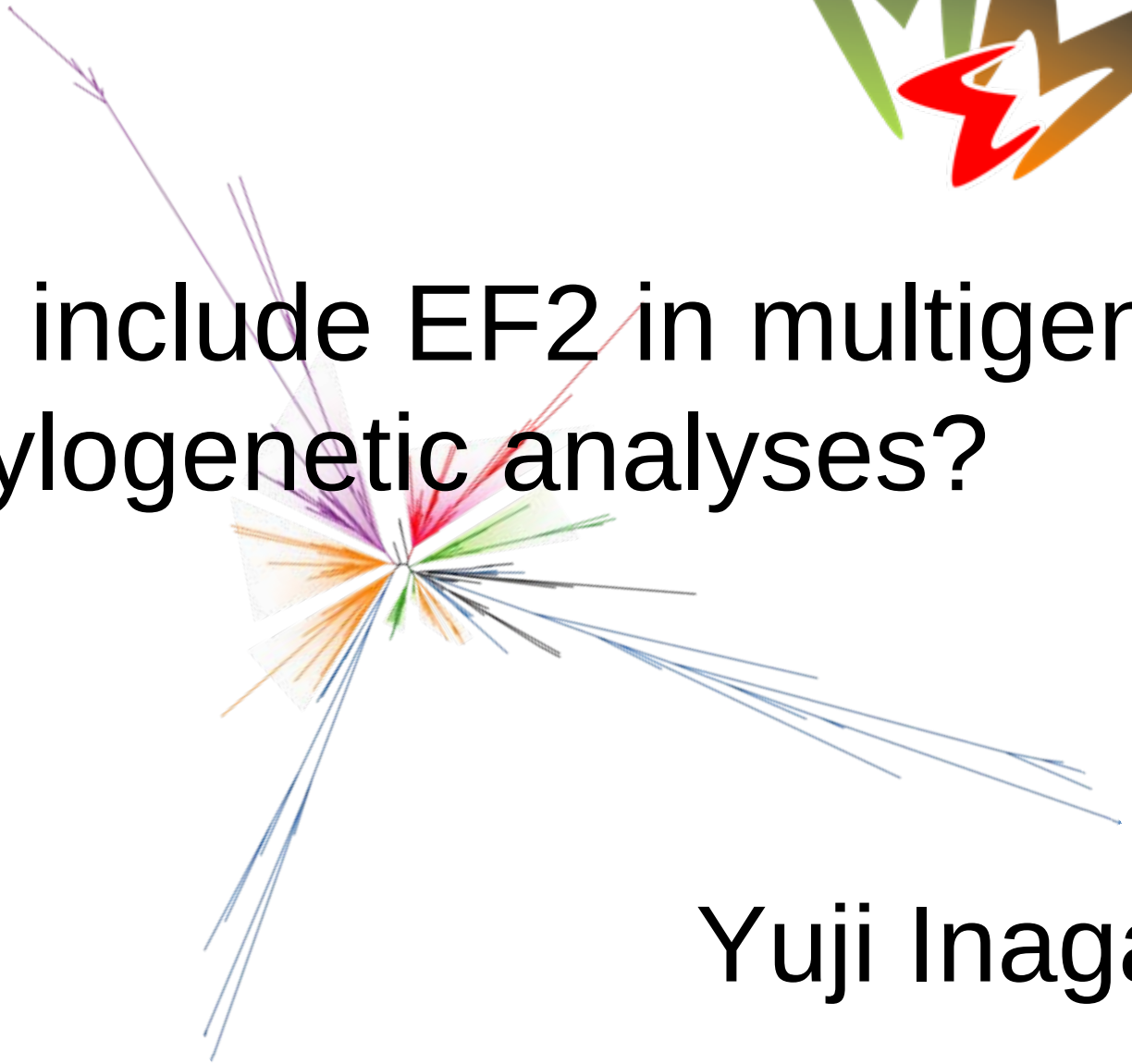
MOL EVOL MICROBES
UNIV.TSUKUBA

Can we include EF2 in multigene phylogenetic analyses?



Yuji Inagaki

Translation elongation factor 2 (EF2)

- Catalyzes translocation of aa-tRNA from A site to P site in ribosomes
  - Highly conserved
- EF2 is a major “maker”
  - Monophyly of red algae and green plants (Moreira, Le Guyader & Philippe 2000)
  - Stiller, Riley & Hall. (2001) suggested that red and green genes were recombined
- Here we tested
  - Multi-gene data w/ and w/o EF2
  - “Recombination” hypothesis by Stiller et al.

Model & genes

- ML method under the unlinked WAG + Γ model
  - Among-site rate variation
  - Variation of evolutionary patterns across genes
- 24 nuclear-encoded protein genes
  - EF-1α, EF2, IleRS, ValRS, RPS14, RPS15a, RPL5, RPL8, RPL10a, RPB1, cytosolic-Hsp70, ER-Hsp70, mitochondrial-Hsp70, cytosolic-Hsp90, Cpn60, CCTα, CCTδ, CCTγ, CCTζ, Actin, α-Tubulin (Tba), β-Tubulin (Tbb), vATPA, vATPB
- >10,000 aa positions

Eukaryotic groups considered

- Assumed the monophyly of the 7 groups
  - Opisthokonts
  - Amoebozoa
  - Green plants
  - Red algae
  - Euglenozoa + Heterolobosea
  - Alveolates + stramenopiles
  - Diplomonads + parabasalids
- Exhaustively searched for the ML tree from 945 test trees
- Bootstrap analysis

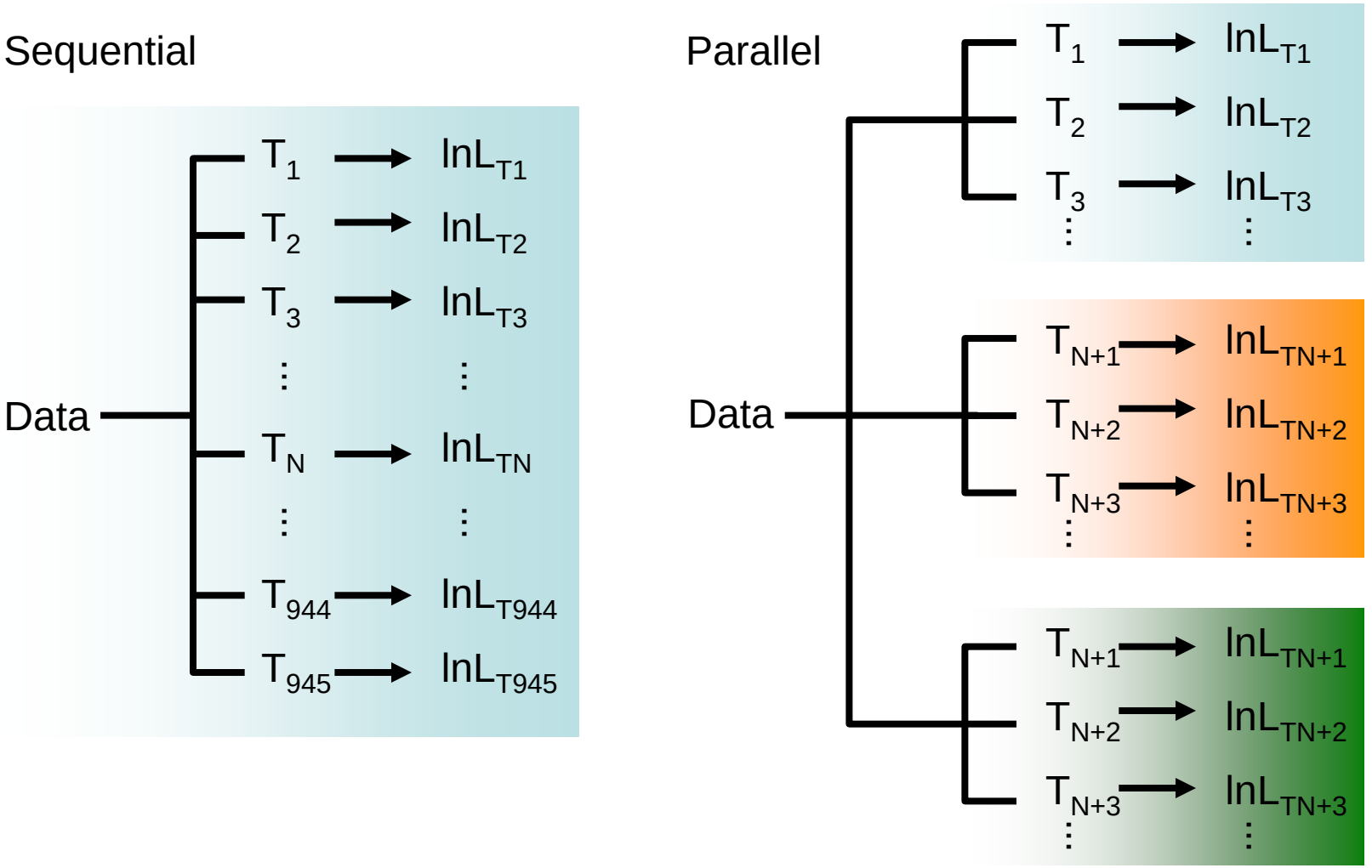
Exhaustive search for the ML tree

- Calculated the lnL of 945 trees for each gene
  - Parameters optimized for each gene

	T_1	T_2	T_3	・・・	T_N	・・・	T_{944}	T_{945}
Data 1	lnL	lnL	lnL	・・・	lnL	・・・	lnL	lnL
Data 2	lnL	lnL	lnL	・・・	lnL	・・・	lnL	lnL
Data 3	lnL	lnL	lnL	・・・	lnL	・・・	lnL	lnL
⋮	⋮	⋮	⋮		⋮		⋮	⋮
Data 10	lnL	lnL	lnL	・・・	lnL	・・・	lnL	lnL
⋮	⋮	⋮	⋮		⋮		⋮	⋮
Data 23	lnL	lnL	lnL	・・・	lnL	・・・	lnL	lnL
Data 24	lnL	lnL	lnL	・・・	lnL	・・・	lnL	lnL
Total	lnL_1	lnL_2	lnL_3	・・・	**lnL_N**	・・・	lnL_{944}	lnL_{945}

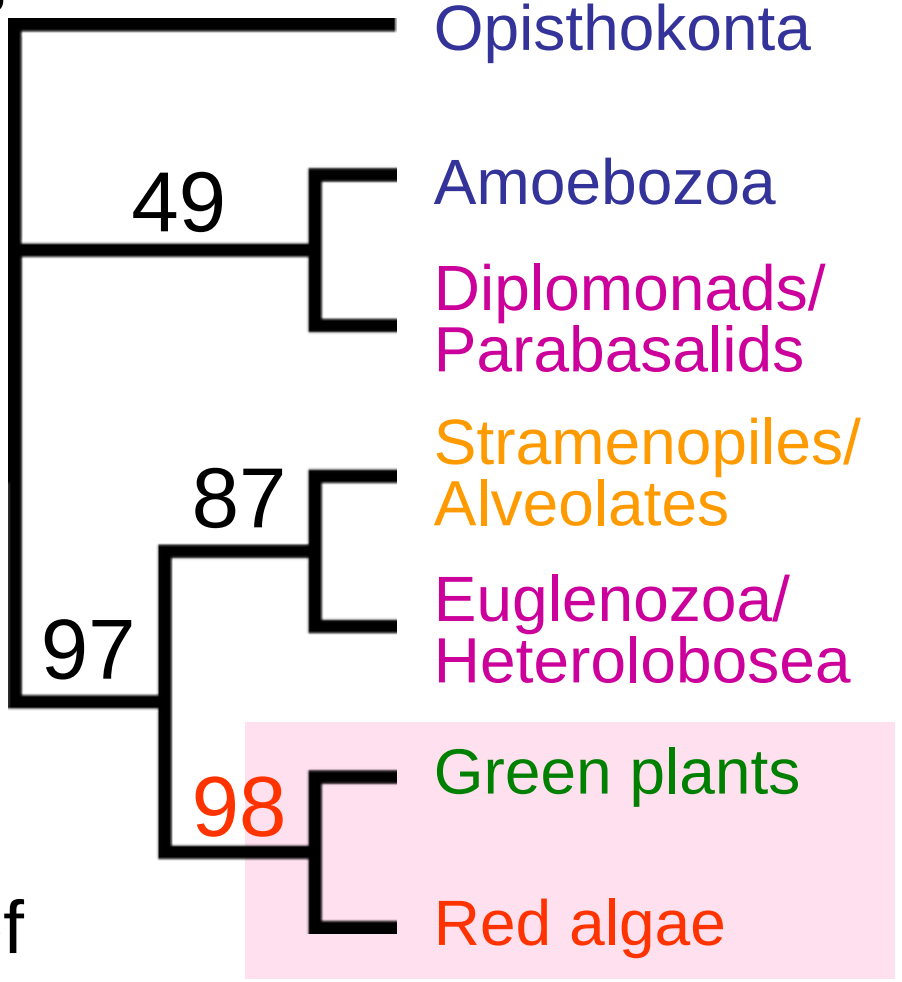
MPI-Puzzle – For exhaustive tree search

- Calculated lnL by a *modified* Tree-Puzzle



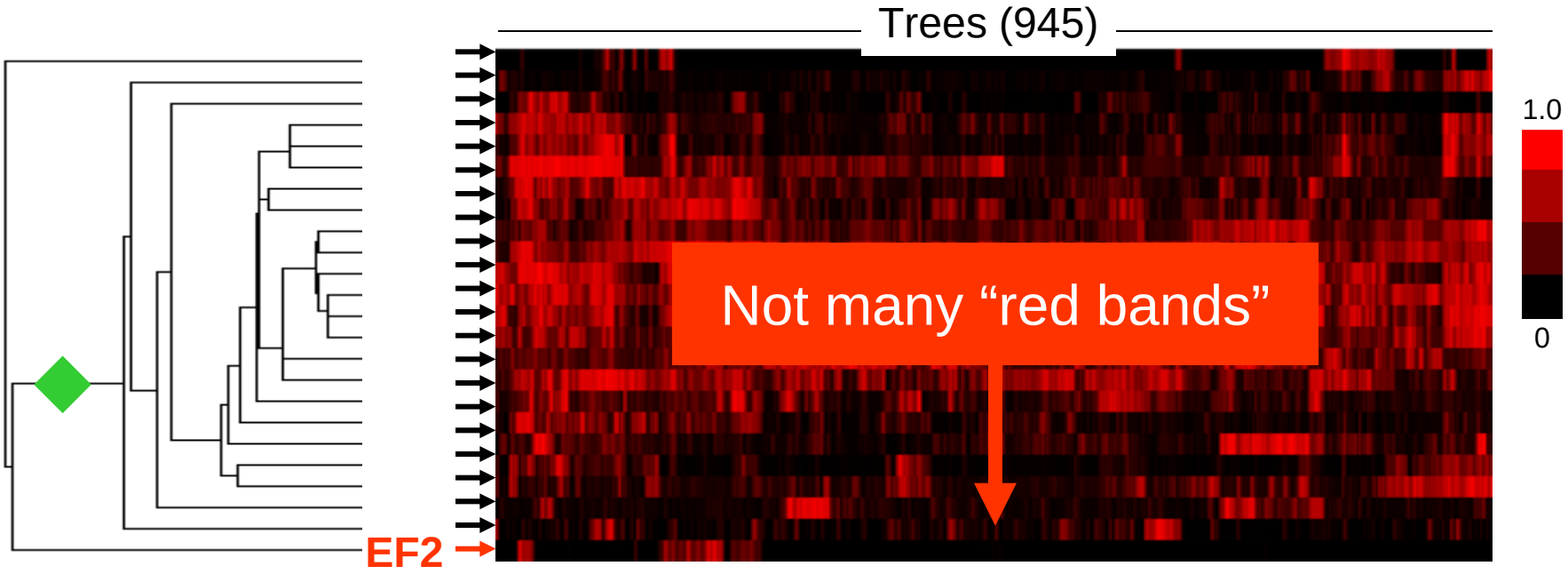
24-gene phylogeny

- Are strong supports true?
  - EF2 prefers R+G
- Are the signals uniform?
  - Single-gene JK
  - More systematic way to compare signal in each gene
  - By using *P* values of SH tests



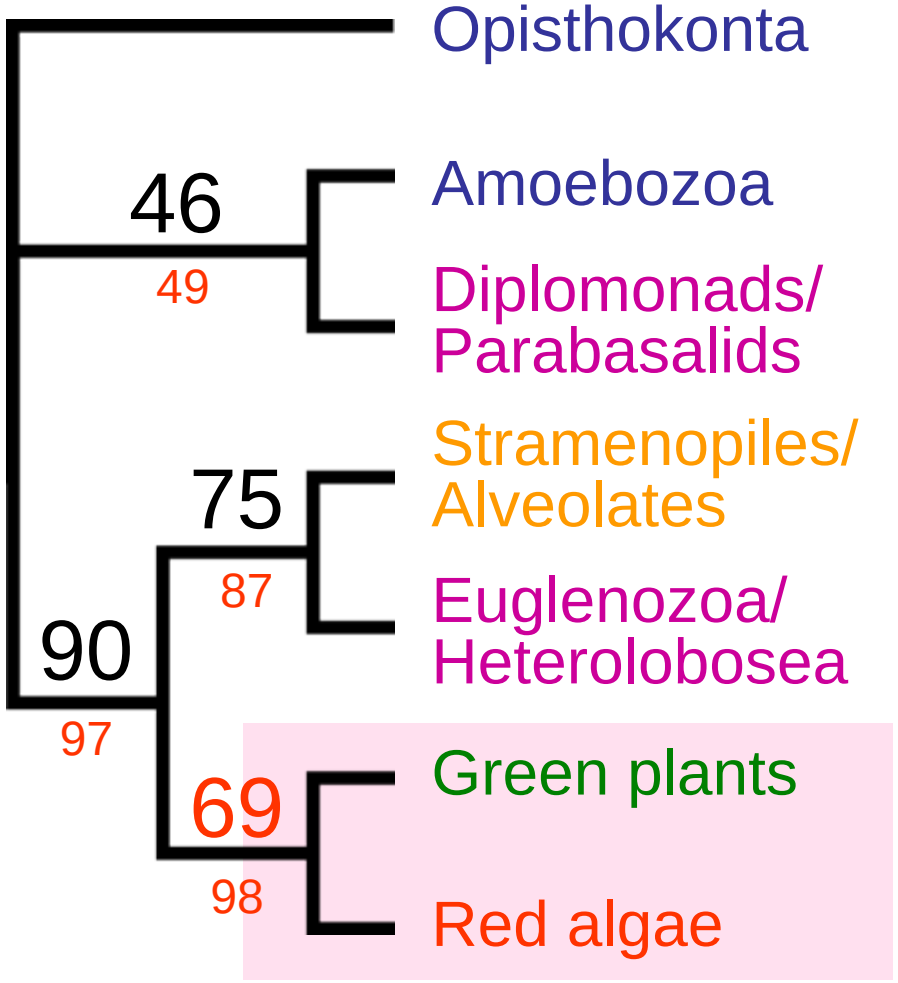
P value heat map

- "Gene-specific signal" was represented by *P* values of SH tests
  - *P* values are color-coded
  - EF2: 109 trees passed
  - All trees w/ "R+G" passed



23-gene analyses – w/o EF2

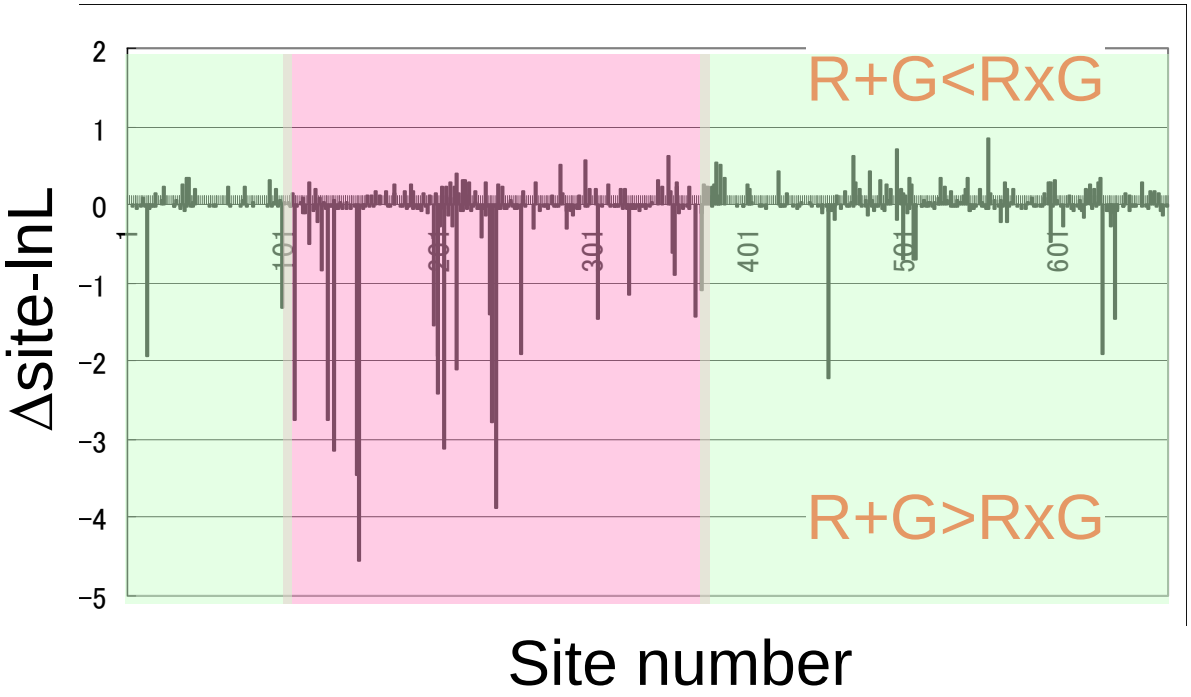
- R+G recovered
  - BP decreased
  - Other genes also prefer R+G
- EF2 signal
  - Is not incongruent with those of other genes
- OK to use EF2
- How about the GR?



EF2: gene recombination between Red & Green?

- Stiller, Riley & Hall. JME 2001 52:527-39
  - Red & Green EF2 were recombined
  - EF2 phylogeny: Artifact from the GR signal
- Test the “GR” hypothesis
  - R+G tree: the ML tree
  - RxG tree: the best tree with no R+G clade
- Calculated site-lnLs over the two trees
  - Sites that prefer the R+G tree over the other
  - Do the R+G sites pack in a short region?

EF2: gene recombination between Red & Green?



- R+G signal locates the non-pink area
- The ML tree of the green area: no R+G clade
  - A R+G tree: The 2nd best (Δ lnL = 0.2; P = 0.828)
- No strong reason to invoke the GR proposed