

Announcements:

1. Next week's lecture and exercise are cancelled
2. New office hours: Thursdays 14:00-15:00
3. Submissions
4. Late submission

Exercise No.2

Question 4

Homozygotic selection - it is advantageous not to be heterozygous

	AA	Aa	aa
	w_{11}	w_{12}	w_{22}
Weights	$1+s$	1	$1+s$

$$\Delta q = \frac{pq[p(w_{12} - w_{11}) + q(w_{22} - w_{12})]}{p^2 w_{11} + 2pqw_{12} + q^2 w_{22}}$$

$$\Delta q = \frac{pqs(q-p)}{1 + p^2 s + q^2 s}$$

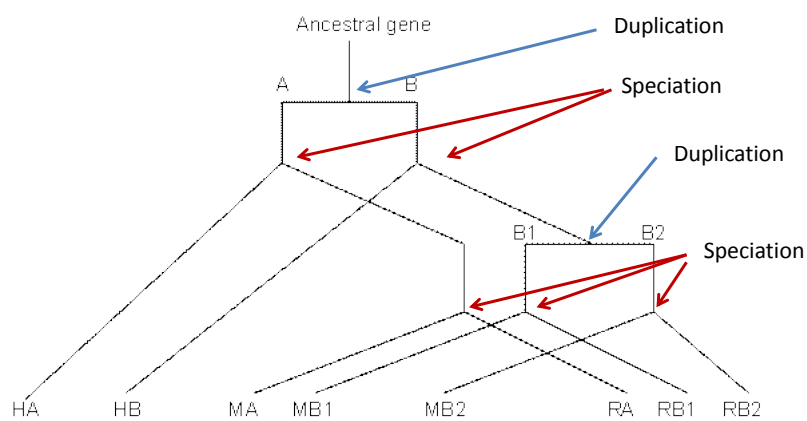
If s is small denominator can be approximated to be 1

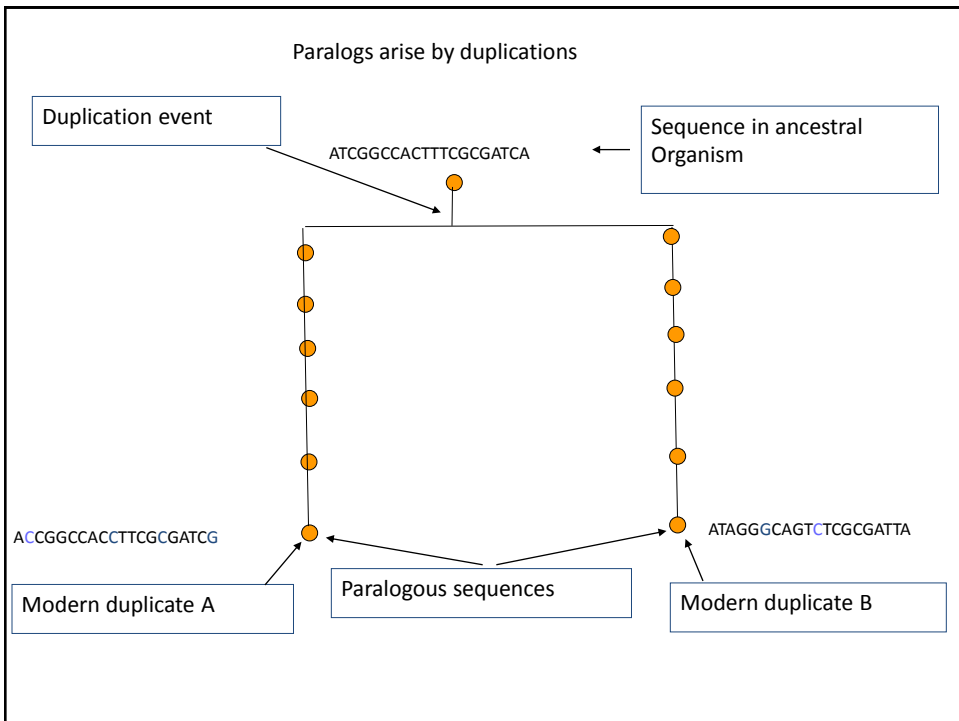
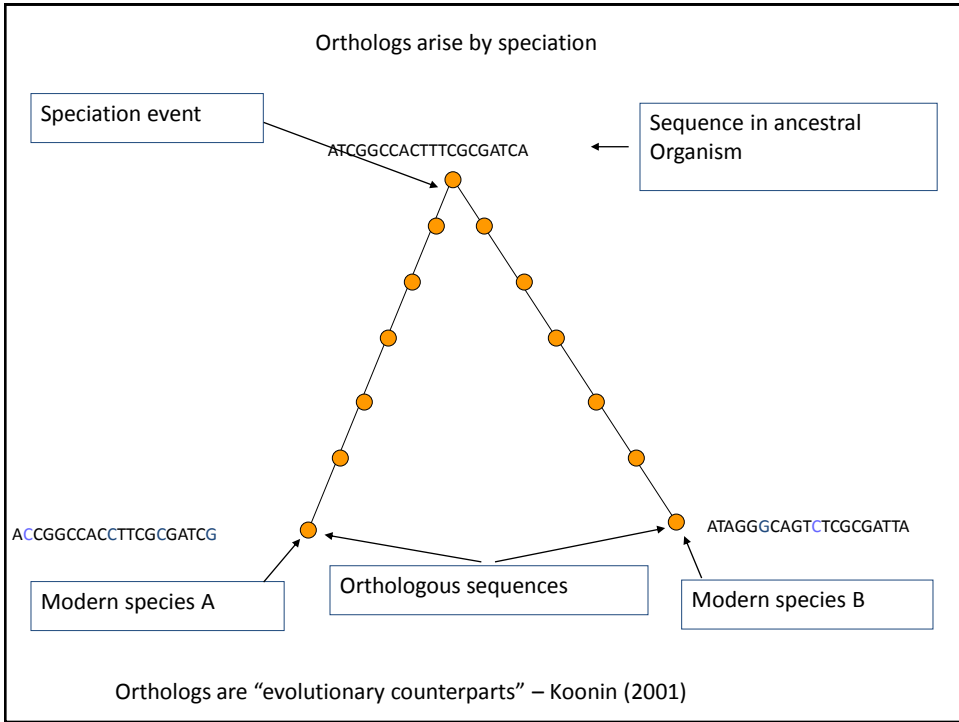
$$\Delta q = pqs(q-p)$$

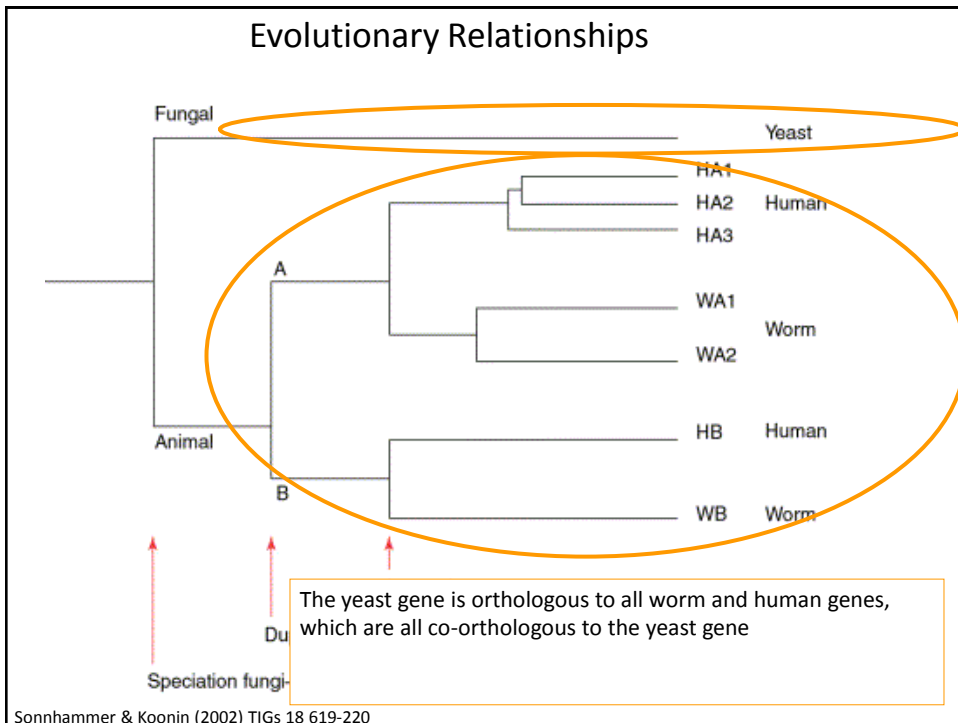
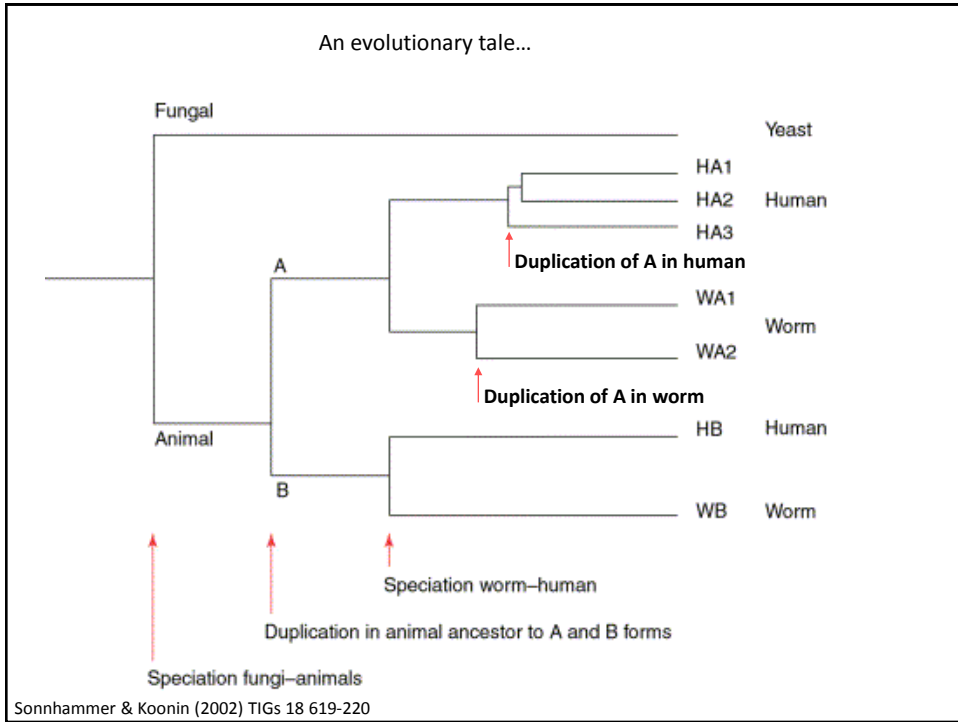
Exercise No.3

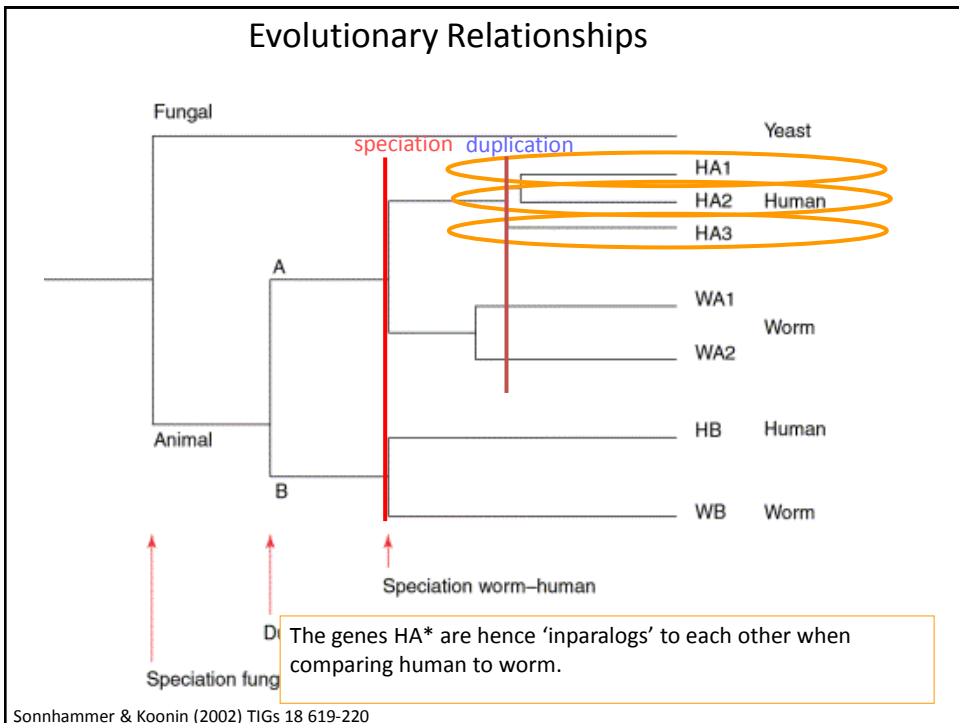
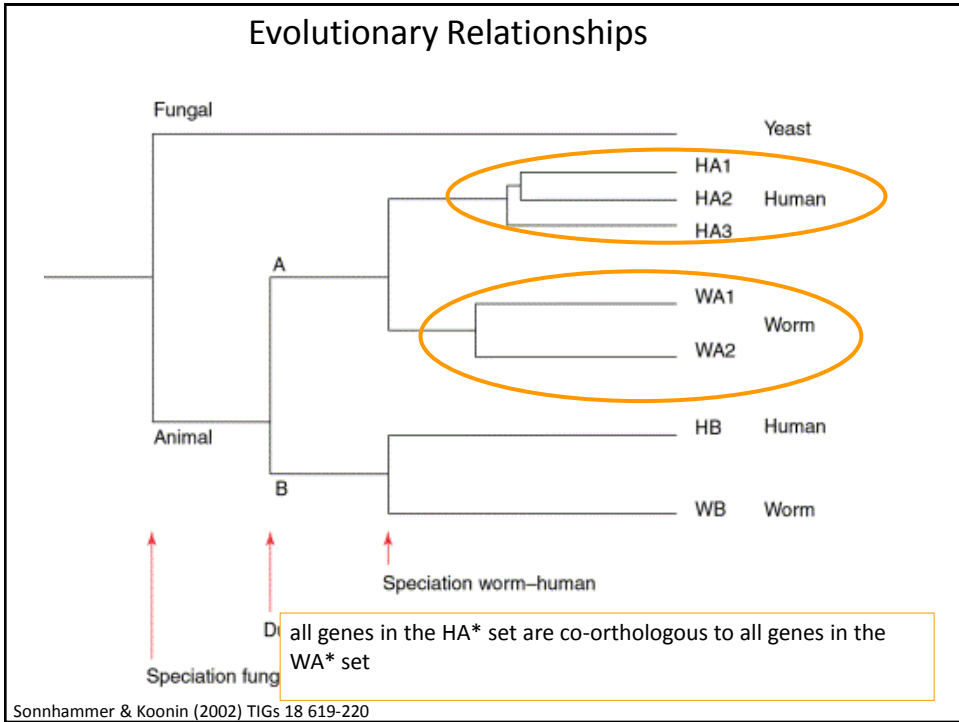
Problem 1: *Warm-up.* What's wrong with the following sentence "Proteins A and B are 90% homologous"?

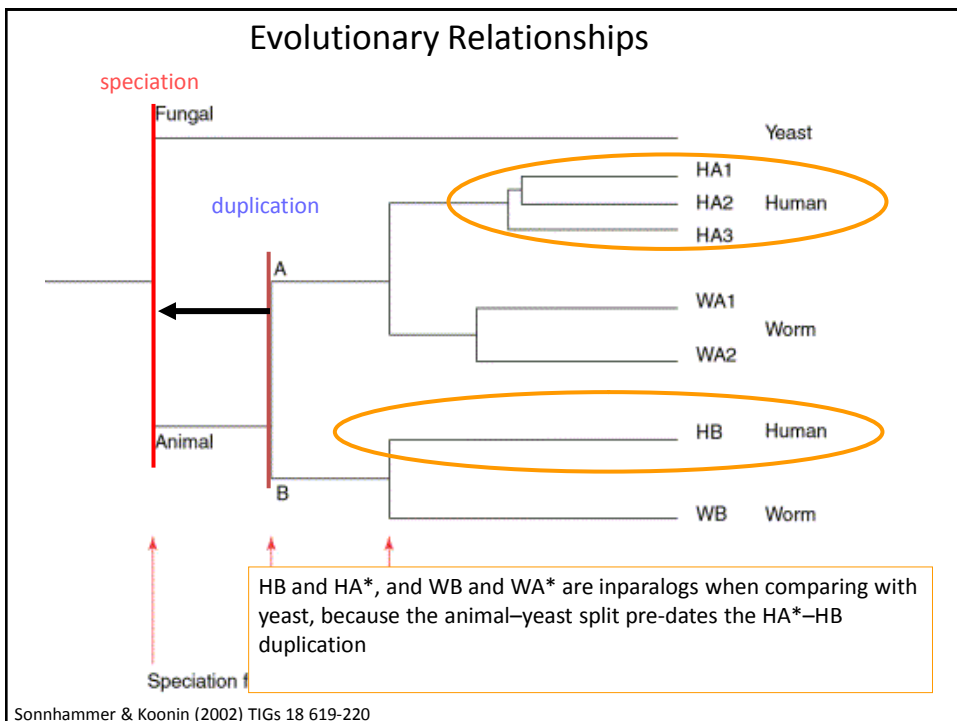
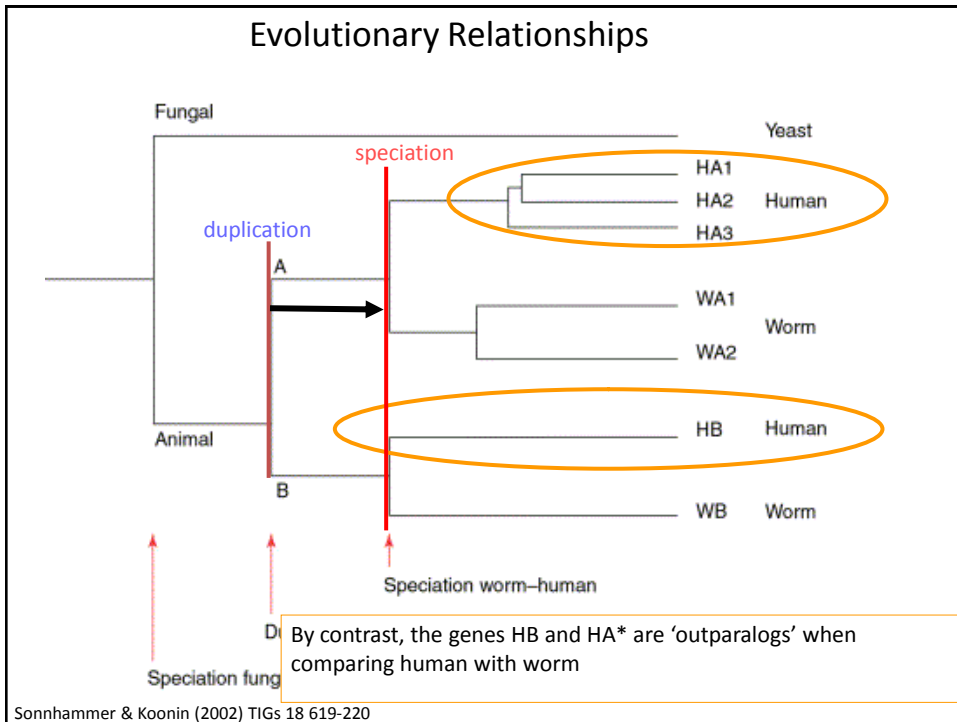
Problem 2:











Testing the Molecular Clock

Problem 3:



<http://www.megasoftware.net/>

Recommendation: Download this free software and learn to use it!

Problem 3: Correcting for multiple hits

For this problem, you will need to download MEGA4: <http://www.megasoftware.net/> and open the *Drosophila*_Adh.meg file (this is the Adh region for 11 flies).

Here is a tree of relationships among the flies (we learn what this means in the 5th lecture).



Calculate the sequence distance between the *D. melanogaster* and *D. pseudoobscura* according to:

1. The fraction of bases different between them
2. The Jukes-Cantor corrected distance

Also, explain why the Jukes-Cantor distance is higher or lower.

MEGA 4.0.1

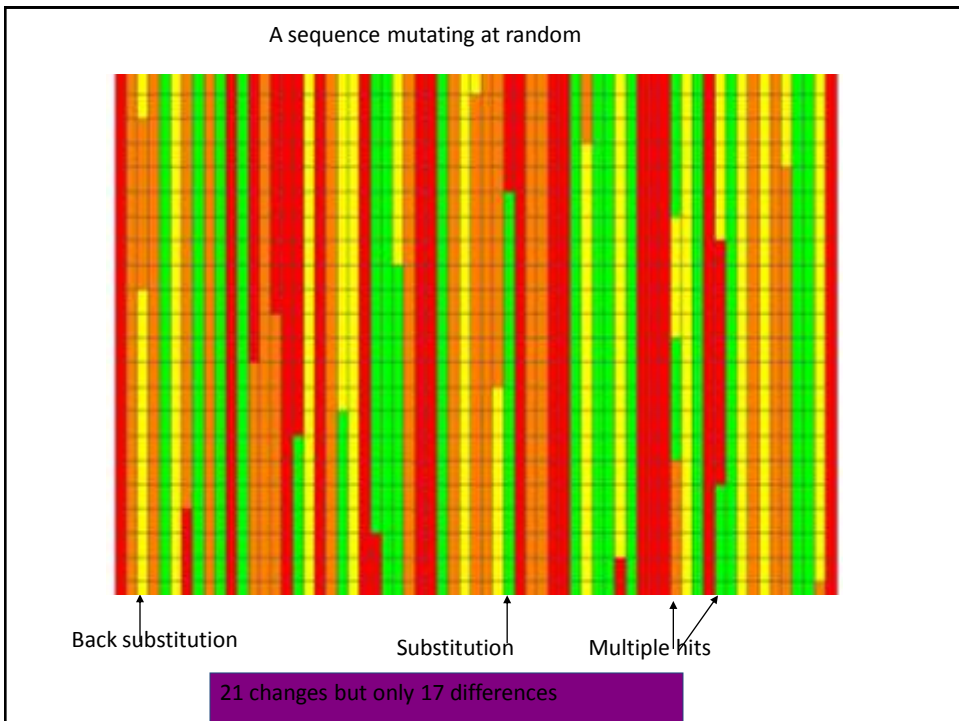
File Data Sources Phylogeny Partition Selection Alignment Window Help

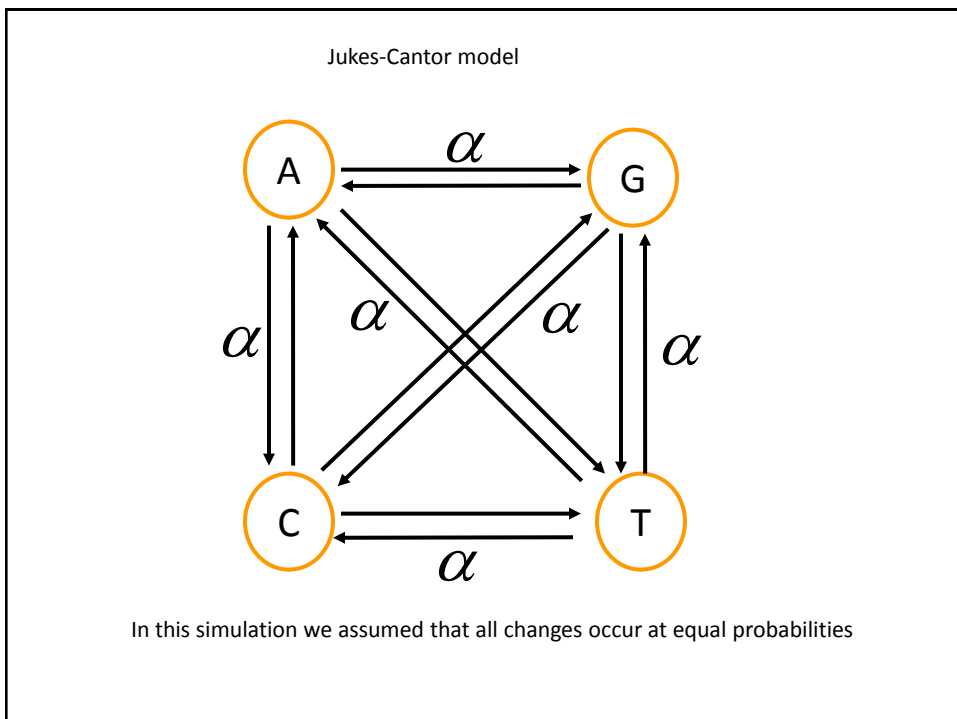
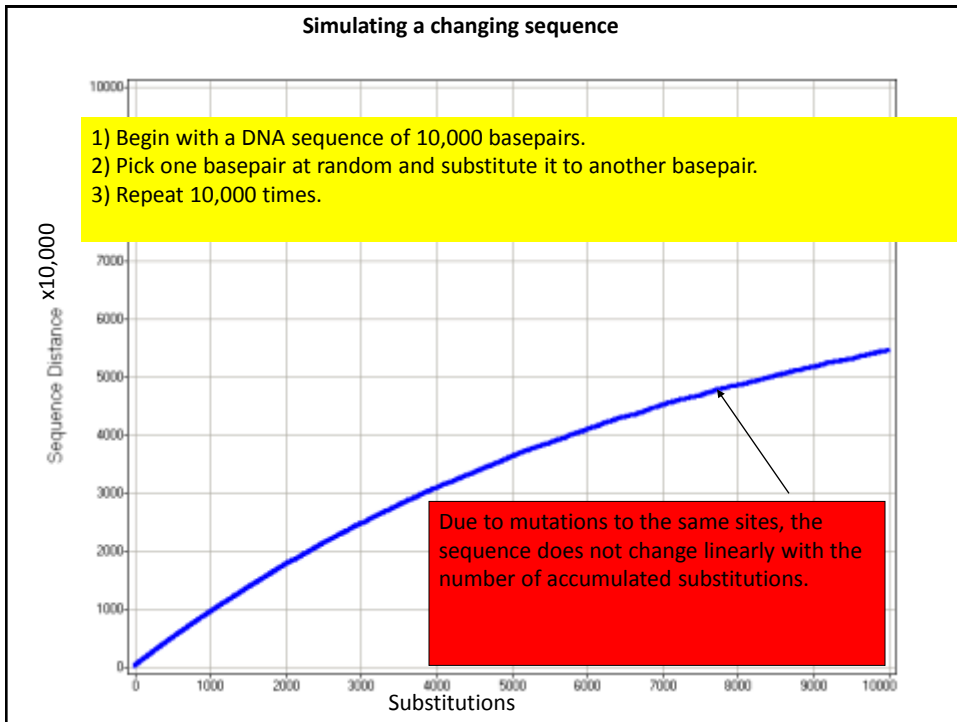
MEGA 4.0.1 (Beta 3): View Sequence Data

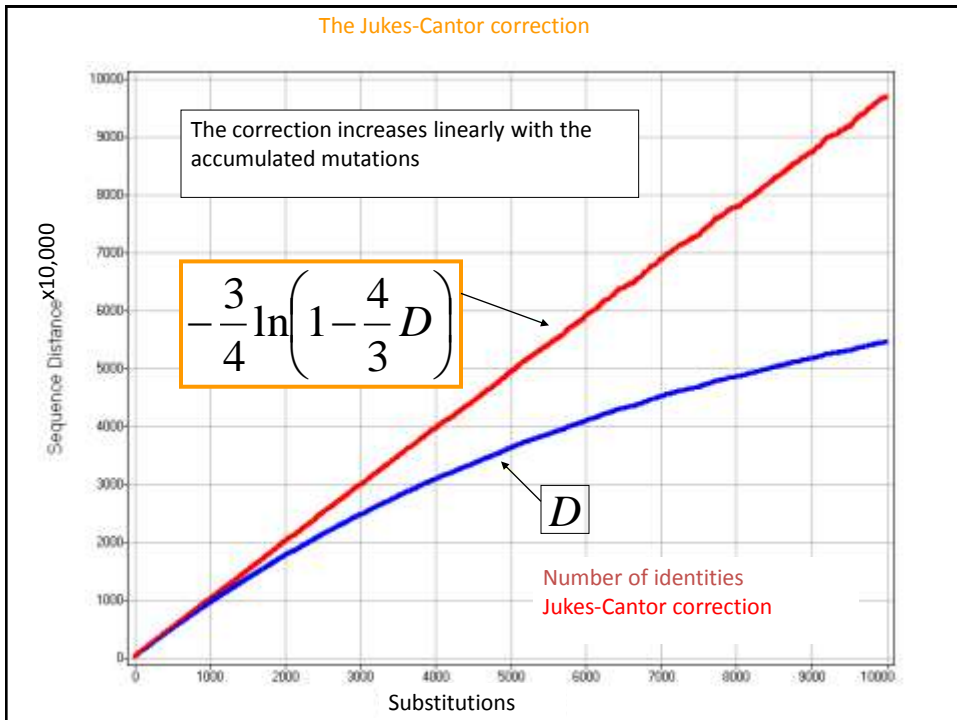
Data Display Highlight Statistics Help

	A	T	G	C	T	G	A	C	A	A	C	A	G	A	C	G	T	G	A	T	T	T	C	G	T	G	C	T	G	
1. D. melanogaster																														
2. D. pseudoobscura																														
3. D. obscurus																														
4. S. albivittata																														
5. D. crossifenus																														
6. D. mulleri																														
7. D. affinis																														
8. D. heteroneura																														
9. D. nigra																														
10. D. adactyla																														
11. D. nigra																														

1/762 Conserved: 572/762 Adj data (Adj)

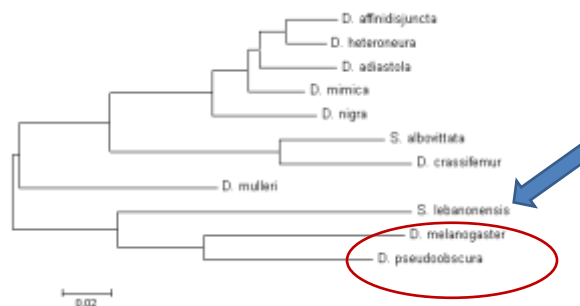






Problem 4: Relative rate test.

Again, use MEGA4 to answer the following question: Can you rule out the molecular clock in the *Adh* sequences of *D. melanogaster* and *D. pseudoobscura* since their divergence with *S. lebanonensis*? Hint: After opening the same file as in Problem 3: In the "Phylogeny" tab, Select "Relative rate tests" -> "Tajima's test"



Tajima's relative rate test

```

Computation 1: Tajima's relative rate test.
Gaps/missing data : Complete Deletion
Mode : Nucleotides
Data Info : (All)
Changes considered : All
Sequence A : Clibanarius vittatus
Sequence B : Coenobita sp.
Sequence C : Pagurus acadianus
Test type : Rooted on Seq C
No. of Sites : 327

m_111 = 272 (identical sites in all three sequences)
m_11k = 22 (divergent sites in all three sequences)
m_11j = 30 (unique differences in SeqA)
m_12j = 29 (unique differences in SeqB)
m_13j = 44 (unique differences in SeqC)
Chi-square= 0.02 with 1 degree of freedom (p=0.896)

```

P is not significant
so we cannot rule
out the molecular
clock

$$E(m_1) = E(m_2)$$

$$\chi^2 = (m_1 - m_2)^2 / (m_1 + m_2)$$

$$\chi^2 = (30 - 29)^2 / (30 + 29) = 0.0169$$

Table. Results from the Tajima test for 3 Sequences [1].

Configuration	Count
Identical sites in all three sequences (m_{111})	549
Divergent sites in all three sequences (m_{1jk})	14
Unique differences in Sequence A (m_{11j})	143
Unique differences in Sequence B (m_{12j})	32
Unique differences in Sequence C (m_{13j})	24

Note: The equality of evolutionary rate between *D. melanogaster* and *S. albovittata* is tested using *D. crassifemur* as an outgroup in Tajima's relative rate test in MEGA4 [1, 2]. The χ^2 test statistic was 70.41 ($P = 0.00000$ with 1 degree[s] of freedom). P -value less than 0.05 is often used to reject the null hypothesis of equal rates between lineages.