

Computational Biology

Lecturers:
Tanja Stadler, Carsten Magnus & Tim Vaughan

Teaching Assistants:
Jūlija Pečerska, Jérémie Sciré,
Sarah Nadeau & Marc Manceau

Computational Evolution
Department of Biosystems Science and Engineering

HS 2019

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

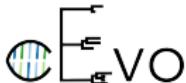
Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References



Parsimony & Maximum Likelihood: Discussion questions

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

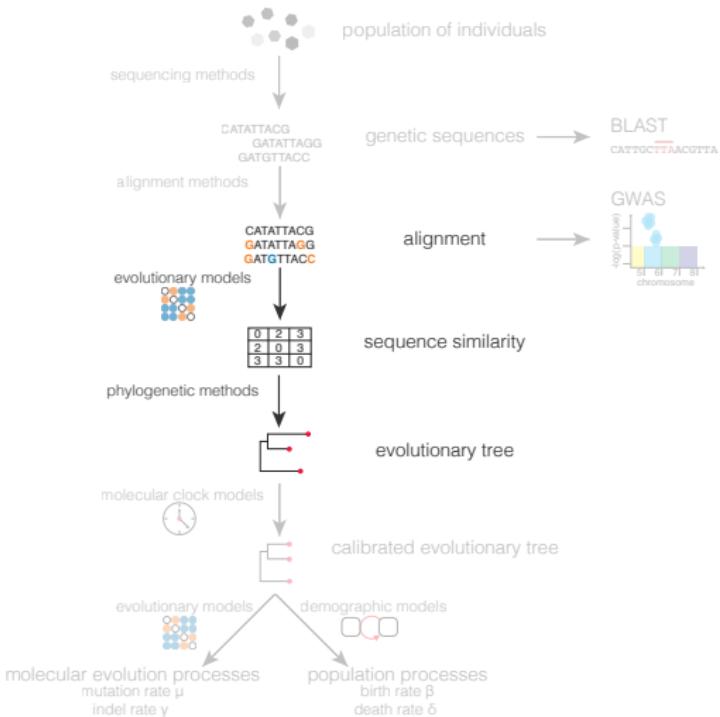
HIV origin

HIV transmission trees

References

- ② Consider the Fitch algorithm. Do you obtain all most parsimonious ancestral sequences when choosing the different nucleotides in the curly brackets?
- ② Does the maximum likelihood tree reconstruction method return estimates for the internal sequences? Give a reason for your answer.
- ② Does the Fitch algorithm return the parsimony score for any phylogenetic tree and any sequence alignment? Or are there situations when the Fitch algorithm does not return the smallest number of mutations required?

Overview



c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References

Overview

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References

- ▶ Searching tree space
- ▶ On model testing & confidence in estimates
- ▶ Phylogenetics in HIV research

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References

Searching tree space.

How to search tree space for the maximum likelihood tree?

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References

- ▶ We need to propose different unrooted trees: NNI, SPR, and TBR moves (next slides)
- ▶ We need to propose different branch lengths: multiply each branch length by some factor
- ▶ We can use “hill-climbing” strategies to find the optimum

NNI move

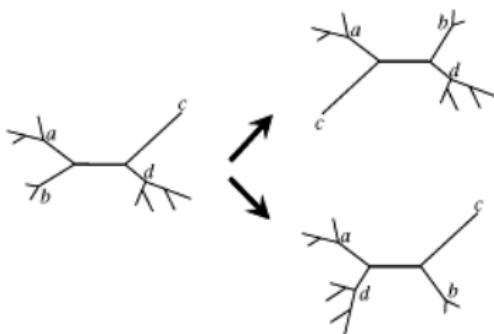


Figure adapted from [Yang, 2006]

Fig. 3.12: “The nearest-neighbour interchange (NNI) algorithm. Each internal branch in the tree connects four subtrees or nearest neighbours (a, b, c, d). Interchanging a subtree on one side of the branch with another on the other side constitutes an NNI. Two such rearrangements are possible for each internal branch.”

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References

SPR and TBR move

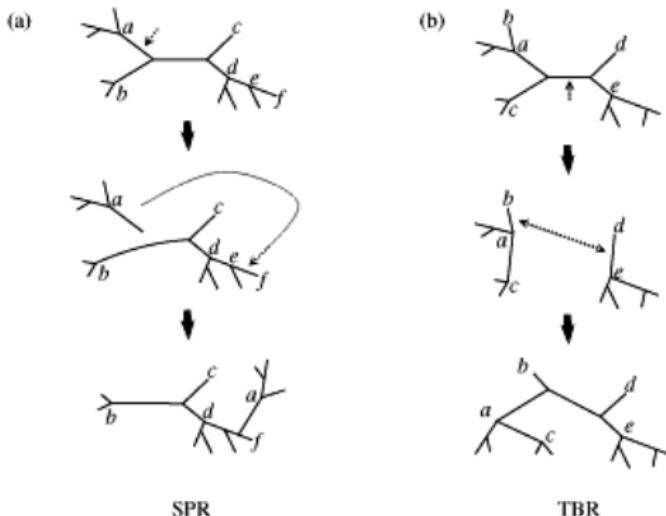


Figure adapted from [Yang, 2006]

Fig. 3.13: “(a) Branch swapping by subtree pruning and regrafting (SPR). A subtree (for example, the one represented by node a) is pruned, and then reattached to a different location on the tree. (b) Branch swapping by tree bisection and reconnection (TBR). The tree is broken into two subtrees by cutting an internal branch. Two branches, one from each subtree, are then chosen and rejoined to form a new tree.”

- c07- Maximum likelihood method & testing
- Searching tree space
- Model testing
- Which model fits best?
- Uncertainty
- Phylogenetics in HIV
- HIV origin
- HIV transmission trees
- References

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References

Model testing.

Model testing

- ▶ Which evolutionary models are appropriate for our data?
 - ▶ likelihood ratio test
 - ▶ AIC
- ▶ How confident can we be in the inferred parameters, i.e. how confident can we be in the phylogeny and the substitution rates?
 - ▶ confidence assessed via likelihood ratios
 - ▶ confidence assessed via bootstrapping
- ▶ All methods above, with the exception of bootstrapping, require a maximum likelihood method

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References

Reminder from Lecture 03

- ▶ We used the hypergeometric distribution as a null model H_0 and asked if the **p-value** lies below the **significance level** α . In that case, the null model H_0 was rejected.
- ▶ We did not have an alternative model. However, we could approximately evaluate $P(X|\theta)$ for all possible data X and fixed parameter θ under the null model (in fact, we used the χ^2 distribution to approximate the probability of the data, namely the contingency table).
- ▶ Now we want to ask if we reject model H_0 in favour of model H_1 . This is advantageous in the phylogenetic setting, as it is not straightforward to evaluate $P(X|\theta)$ (where X is now an alignment) for all possible alignments.

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References

Likelihood ratios

- ▶ Assume the data evolved under a model H_0 .
- ▶ Assume a model H_1 within which H_0 is nested.
- ▶ For the given data, let the maximum likelihood parameter estimate under model H_0 be $\hat{\theta}_0$ and under H_1 be $\hat{\theta}_1$.
- ▶ Now under some “mild” conditions (in particular large amount of data),

$$2(\log L(\hat{\theta}_1) - \log L(\hat{\theta}_0)) \sim \chi_{df}^2$$

where df is the degree of freedom

- ▶ The degree of freedom in the χ_{df}^2 distribution is the difference between the number of parameters in the general and in the nested model. Be careful though when the special model parameter is at the parameter boundary (e.g. 0 or ∞), then the degree of freedom loss is typically only 0.5 (for more details see e.g. [Self and Liang, 1987]).
- ▶ $2(\log L(\hat{\theta}_1) - \log L(\hat{\theta}_0)) = 2 \log \left(\frac{L(\hat{\theta}_1)}{L(\hat{\theta}_0)} \right)$ explains the name “likelihood ratio”

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

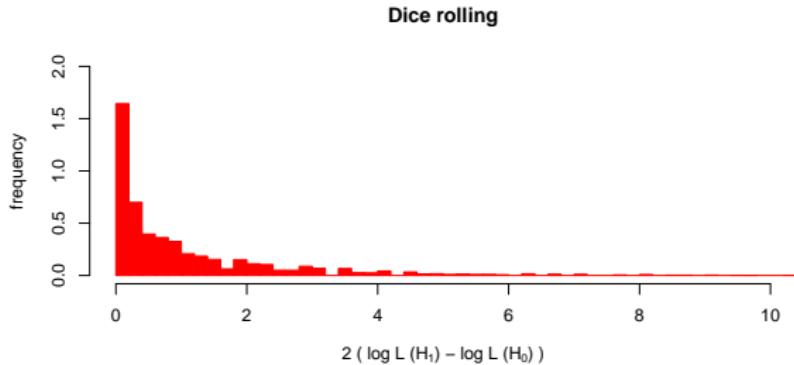
Phylogenetics in HIV

HIV origin

HIV transmission trees

References

Example: Die rolling



c07- Maximum likelihood method & testing
 Searching tree space
 Model testing
 Which model fits best?
 Uncertainty
 Phylogenetics in HIV
 HIV origin
 HIV transmission trees
 References

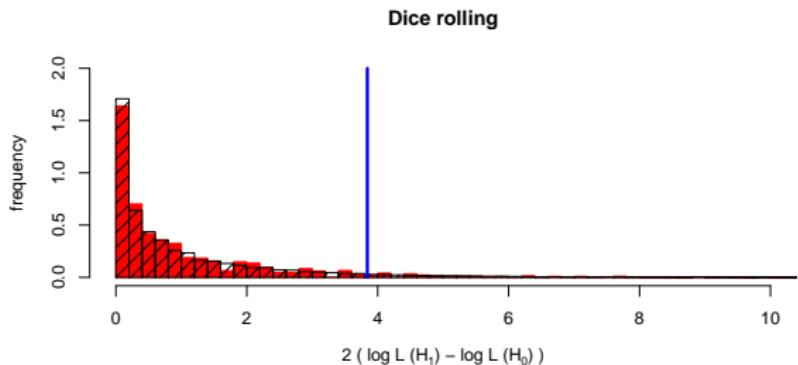
- We roll a fair (i.e. $\theta_0 = 1/6$ for probability of rolling a 6) die $n = 1000$ times and record how often we throw a 6, denote this with k . The value $\hat{\theta}_1 = k/n$ is the maximum likelihood estimate for the probability of rolling a 6 (lecture 4).
- We calculate

$$l = 2(\log L(\hat{\theta}_1 = k/n) - \log L(\hat{\theta}_0 = 1/6)).$$

Recall that $L(\theta) = \binom{n}{k} \theta^k (1-\theta)^{n-k}$.

- This is done 10000 times; we show the histogram for l (red).

Example: Die rolling



- c07- Maximum likelihood method & testing
- Searching tree space
- Model testing
- Which model fits best?
- Uncertainty
- Phylogenetics in HIV
- HIV origin
- HIV transmission trees
- References

- ▶ we additionally plot the $\chi^2_{df=1}$ distribution (black) - you can observe that this distribution is approximating $2(\log L(\hat{\theta}_1 = k/n) - \log L(\hat{\theta}_0 = 1/6))$
- ▶ we draw a blue line at the .95 quantile of the $\chi^2_{df=1}$ distribution (i.e. only 5 % of the values are to the right of the blue line)

Statistical testing

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

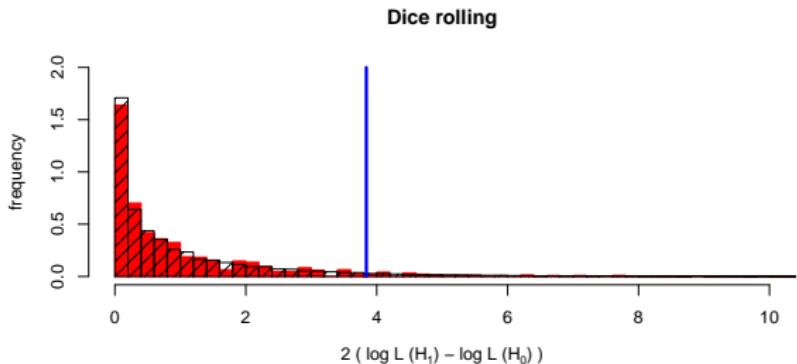
Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References



- ▶ null model, $H_0 : \hat{\theta}_0 = 1/6$
- ▶ general model, $H_1 \in (0, 1)$
- ▶ maximum likelihood estimate $\hat{\theta}_1 = k/n$ (lecture 4)
- ▶ we reject the null model at $\alpha = 0.05$ iff
 $2(\log L(\hat{\theta}_1 = k/n) - \log L(\hat{\theta}_0 = 1/6)) > \chi^2_{1,5\%} = 3.84$

Likelihood ratio test

- ▶ consider two models: H_1 general model parameterised in θ_1 ,
 H_0 nested model parameterised in θ_0
- ▶ derive the likelihood function for both models and the maximum likelihood estimators $\hat{\theta}_0$, $\hat{\theta}_1$ for a given data set
- ▶ calculate $2(\log L(\hat{\theta}_1) - \log L(\hat{\theta}_0))$
- ▶ reject the null model if $2(\log L(\hat{\theta}_1) - \log L(\hat{\theta}_0))$ is in the α tail of the χ^2_{df}
 - ▶ most often $\alpha = 0.05$
 - ▶ α is called the significance level (see also Lecture 3)
 - ▶ if the null model was the true model, we would falsely reject it in a proportion α of tests.
 - ▶ if the null model was the false model, then we expect the null model to have a much lower likelihood than the true model, and thus we would accept the null model only in a very low proportion of tests.

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References

Connections to Lecture 3

- ▶ In lecture 3, we considered $P(X|\theta)$ under the null model, and asked if θ is an extreme outcome. Thus, for the die experiment, we checked if the number of times a 6 was thrown is in the 5% tail of the binomial distribution.
- ▶ Here we compare nested models, i.e. when a simple model (H_0) is obtained by restricting parameters in the general model (H_1). For our die, the simple model had no free parameter, but in general $\hat{\theta}_0$ may be the maximum likelihood estimate of a simple model with free parameters.
- ▶ While in lecture 3 we assessed the overall fit of a model, here we only assess the fit of H_0 relative to H_1 . In other words, even though H_1 might be a very bad model, we may reject H_0 in favor of H_1 since H_0 is even much worse.

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References

Model testing errors

	H_0 true	H_0 false
reject H_0	Type I error	Correct
accept H_0	Correct	Type II error

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References

- ▶ Accuracy = $1 - (\text{Type I error})$
 - ▶ The type I error is the significance level, and thus the accuracy is controlled by setting α .
- ▶ Power = $1 - (\text{Type II error})$
 - ▶ Power can generally only be assessed via simulating under the general model H_1 and assessing the number of times that H_0 is accepted.

Example: Die rolling - Statistical testing errors

- ▶ simulation of 10 000 die rolling experiments with different θ_1 s where we report the fraction of 6 among 1 000 die rollings
- ▶ H_0 : the probability to obtain a 6 is $\theta_0 = 1/6$; we test with significance level $\alpha = 0.05$, i.e. accuracy is 0.95
- ▶ simulations with different θ_1 s, tested under the same H_0 :
 1. $\theta_1 = 1/6$
 H_0 is rejected in 5.1% of the 10 000 experiments. This simply highlights again that we chose an accuracy of 0.95.
 2. $\theta_1 = 1/5$
 H_0 is rejected in 77.66% of the 10 000 experiments. Thus the power is estimated to be 0.78.
 3. $\theta_1 = 1/2$
 H_0 is rejected in all of the 10 000 experiments. Thus the power is estimated to be 1.
- ▶ The power increases with an increasing difference of the true model and the null model H_0 !

c07- Maximum likelihood method & testing
Searching tree space
Model testing
Which model fits best?
Uncertainty
Phylogenetics in HIV
HIV origin
HIV transmission trees
References

Testing non-nested models

Akaike Information Criterion (AIC) for testing non-nested models:

$$\text{AIC} = -2 \log L_i(\hat{\theta}_i) + 2p_i$$

where p_i is the number of parameters and L_i the likelihood function of model i

Workflow:

- ▶ calculate the AIC for each model
- ▶ choose the model with the lowest AIC
- ▶ rationale: AIC aims to pick the model with the smallest expected Kullback-Leibler distance to the true model

Rule of thumb for multiple model comparison:

- ▶ models having AIC within 1-2 of the minimum: substantial support, should receive consideration in inference
- ▶ models having AIC within 4-7 of the minimum: considerably less support
- ▶ models having $\text{AIC} > 10$ above the minimum: essentially no support

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References

Example: maximum likelihood tree on 12 plant species

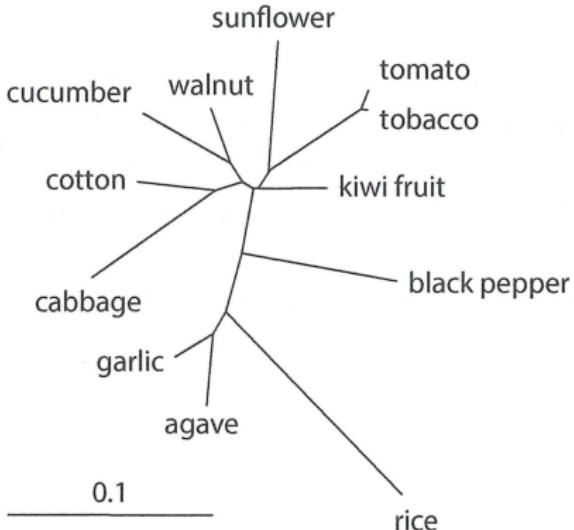


Figure adapted from [Yang, 2014]

Maximum likelihood tree (plastid rbcL genes) from 12 plant species, estimated under HKY85+ Γ_5 .

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References

Which evolutionary model fits best?

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References

Model	p	ℓ	MLEs
JC69	21	-6,262.01	
K80	22	-6,113.86	$\hat{\kappa} = 3.561$
HKY85	25	-6,101.76	$\hat{\kappa} = 3.620$
JC69 + Γ_5	22	-5,937.80	$\hat{\alpha} = 0.182$
K80 + Γ_5	23	-5,775.40	$\hat{\kappa} = 4.191, \hat{\alpha} = 0.175$
HKY85 + Γ_5	26	-5,764.26	$\hat{\kappa} = 4.296, \hat{\alpha} = 0.175$

Figure adapted from [Yang, 2014]

Log likelihood values when maximizing the likelihood over substitution model parameters and branch lengths (for a fixed topology on the 12 species).

- ▶ Can we apply likelihood ratio tests?
- ▶ Can we apply AIC?

Implementations

- ▶ ModelTest [Posada and Crandall, 1998] and its new implementation jModelTest: you can test dozens of models on a fixed tree or with estimating the tree.
- ▶ Given you want to test Jukes-Cantor against the more general GTR model. Can you do a likelihood ratio test?
 - ▶ Yes, given you perform the test on the same tree with fixed branch lengths.
 - ▶ No, if you perform the test on different trees (e.g. on the maximum likelihood tree for Jukes-Cantor and GTR, which may be different). Each tree is a different parameter, thus the full models are not nested. You need to employ the AIC.

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

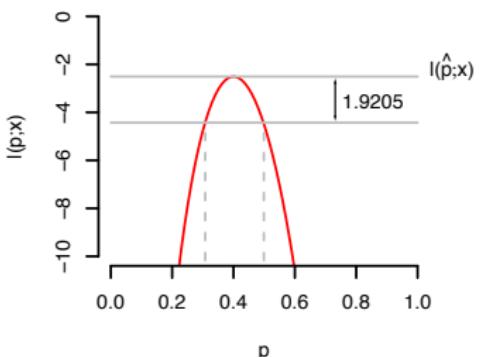
References

Confidence intervals

Each parameter value which is not rejected based on the likelihood ratio test at the 0.05 level is within the 95 % confidence interval.

In lecture 4 we learned:

- ▶ determine the value of the log-likelihood function in $\hat{\theta}$, $l(\hat{\theta}; x)$
- ▶ subtract $0.5\chi^2_{k,5\%}$ i.e. calculate $l(\hat{\theta}; x) - 0.5\chi^2_{k,5\%}$.
- ▶ determine those θ values for which $l(\theta; x) = l(\hat{\theta}; x) - 0.5\chi^2_{k,5\%}$



Same strategy can be used to calculate confidence intervals for the evolutionary parameters given a fixed tree.

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References

More confidence intervals

How confident are we in the maximum likelihood estimate for rolling a 6 in our die experiment?

1. We calculate the confidence interval as explained on last slide. (However, for complex objects like tree topologies, this is not possible.)
2. We can do more experiments. The 95 % interval of the experiment outcomes is obtained by ignoring the smallest 2.5 % and largest 2.5 % outcomes, and then considering the minimum and maximum outcome. However, for many questions, it is not possible to do more experiments (e.g. we cannot repeat plant speciation).
3. We can mimic more experiments by “bootstrapping”. Bootstrapping refers to tests relying on random sampling with replacement. For an additional die experiment, we sample 1000 results of the die roll from our initial 1000 die rolls (with replacement; otherwise we get the original result back). If we had enough die rolls initially, then the bootstrap results are the same as rolling the die again, and we obtain the 95 % interval as in case 2 (redoing experiments).

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

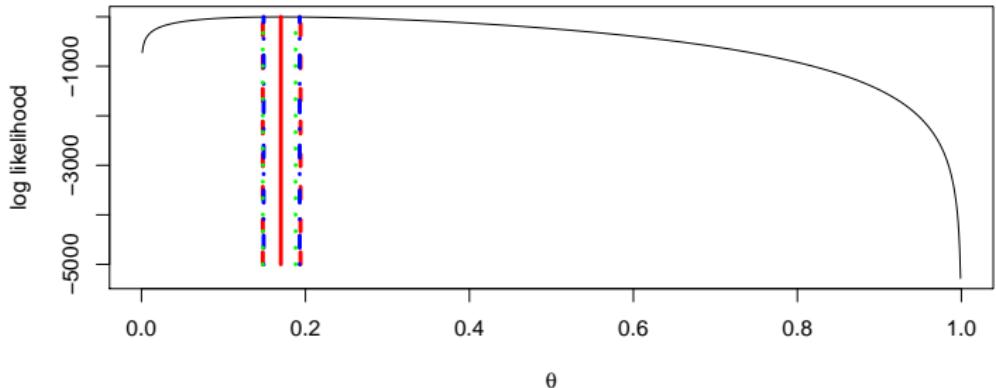
HIV origin

HIV transmission trees

References

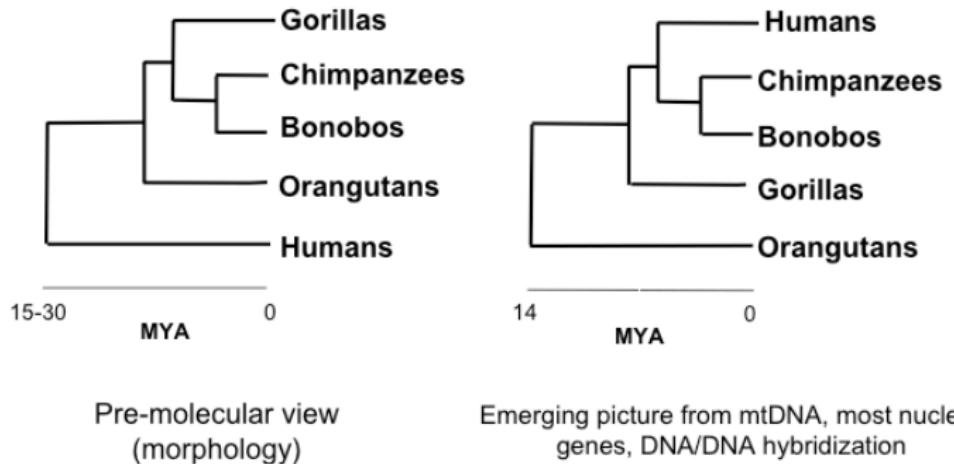
Confidence intervals for parameters

c07- Maximum likelihood method & testing
Searching tree space
Model testing
Which model fits best?
Uncertainty
Phylogenetics in HIV
HIV origin
HIV transmission trees
References



We show the log likelihood function of one experiment in solid black, the maximum likelihood estimate $\hat{\theta}_1 = k/n$ in a solid red line, and the 95 % confidence interval with dashed lines (case 1). 95 % interval obtained from another 100 experiments are in green (case 2), 95 % bootstrap confidence intervals are in blue (case 3).

Assessing the confidence in a tree



Does genetic data reject the morphological tree?

- c07- Maximum likelihood method & testing
- Searching tree space
- Model testing
- Which model fits best?
- Uncertainty
- Phylogenetics in HIV
- HIV origin
- HIV transmission trees
- References

Bootstrapping

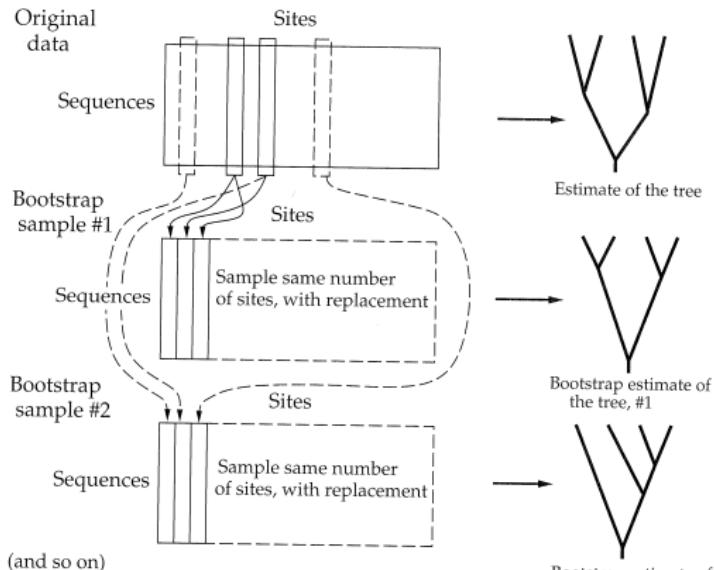


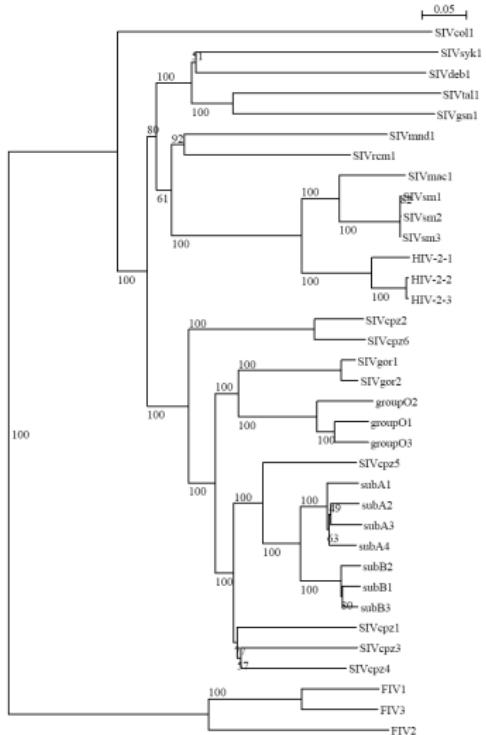
Figure adapted from [Felsenstein, 2004]

- c07- Maximum likelihood method & testing
- Searching tree space
- Model testing
- Which model fits best?
- Uncertainty
- Phylogenetics in HIV
- HIV origin
- HIV transmission trees
- References

Bootstrapping for phylogenies based on an alignment sequences with length m [Felsenstein, 1985]:

- ▶ sample m sites at random with replacement
- ▶ infer a phylogeny based on the new data
- ▶ repeat this procedure many times

Bootstrapping for an HIV/SIV tree



Each node in the tree is labelled with the number of bootstrap trees containing a node with the same descendant tips.

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References

Overview of maximum likelihood inference

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References

1. infer a maximum likelihood tree
 - ▶ employ Felsenstein's pruning algorithm for each tree & branch lengths
 - ▶ choose the tree with branch lengths maximising the likelihood
 - ▶ do this for each substitution model and calculate its AIC
2. determine the substitution model & tree with highest support using AIC
3. determine the confidence interval for the substitution model parameters based on the likelihood ratios
4. determine the confidence in your maximum likelihood tree using bootstrap

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References

Maximum likelihood analysis in HIV research.

How / why did the virus jump from simians to humans: Polio hypothesis

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

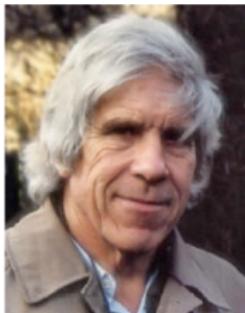
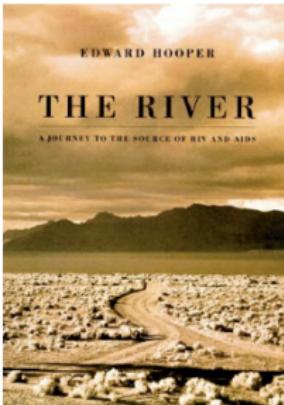
Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References



- ▶ Polio vaccines found in the freezer did not show any signs of contamination, thus the hypothesis was rejected

How / why did the virus jump from simians to humans: Hunter hypothesis

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References

SIV prevalence and first HIV outbreak



Bush meat market

- ▶ Multiple introductions into the human population are expected to be observed under this hypothesis
- ▶ Now, this hypothesis is supported by most scientists!

HIV in Switzerland

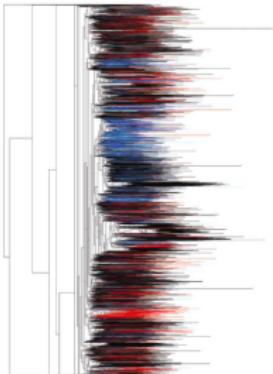


Figure adapted from [Kouyos et al., 2010]

- ▶ 5700 HIV sequences from the Swiss epidemic (Kouyos et al., JID, 2010)
- ▶ Investigates mixing of different population groups (highlighted by different colours)
- ▶ Tips are coloured according to transmission group
 - ▶ Blue: Swiss intravenous drug users
 - ▶ Red: Swiss men having sex with men
 - ▶ Cyan: Swiss heterosexuals
 - ▶ Black: Non-swiss

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References

An HIV criminal case



Figure adapted from [Metzker et al., 2002]

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References

- ▶ Tips in squares are HIV samples coming from a single infected individual (Metzker et al., PNAS, 2002)
- ▶ Large square: suspect
Small square: potential victim

Florida dentist

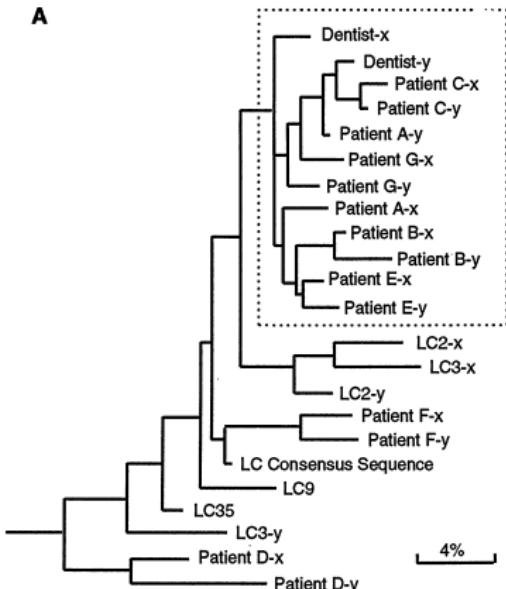


Figure adapted from [Ou et al., 1992]

An HIV-positive Florida dentist infected 6 of his patients. More background reading from the [NY Times](#) and the [LA Times](#).

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References

Nurses in Libya

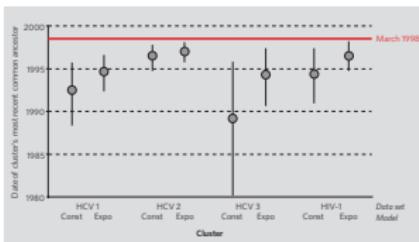
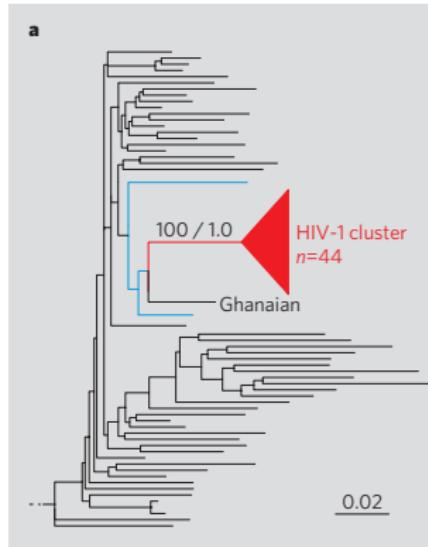


Figure adapted from [de Oliveira et al., 2006]

Six Bulgarian nurses were accused of having infected more than 400 children in a Libyan hospital. The displayed tree is a maximum likelihood tree, the dates are obtained from a Bayesian analysis (lecture 11-12). An informative summary of the case can be found on [wikipedia](#).

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

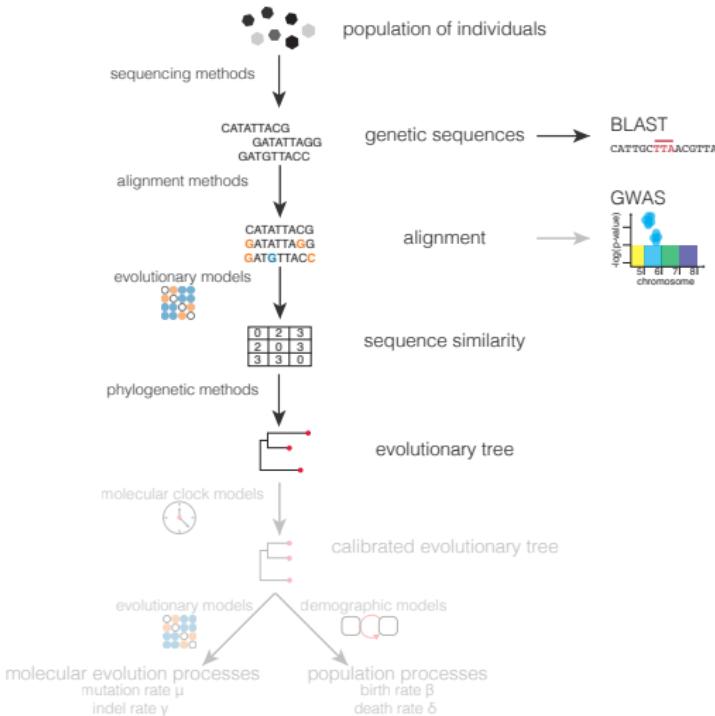
Phylogenetics in HIV

HIV origin

HIV transmission trees

References

Overview



c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References

Statistical testing: Questions

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References

- ① Is there a way to test how to best root a maximum likelihood tree?
- ② Can you use the bootstrapping ideas for assessing confidence in a UPGMA tree?
- ③ What is required to infer the direction of transmission from a phylogeny?

References |

- de Oliveira, T., Pybus, O. G., Rambaut, A., Salemi, M., Cassol, S., Ciccozzi, M., Rezza, G., Gattinara, G. C., D'Arrigo, R., Amicosante, M., et al. (2006). Molecular epidemiology: Hiv-1 and hcv sequences from libyan outbreak. *Nature*, 444(7121):836–837.
- Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, 39(4):783–791.
- Felsenstein, J. (2004). *Inferring phylogenies*. Sinauer associates Sunderland.
- Kouyos, R. D., Von Wyl, V., Yerly, S., Böni, J., Taffé, P., Shah, C., Börgisser, P., Klimkait, T., Weber, R., Hirscher, B., et al. (2010). Molecular epidemiology reveals long-term changes in hiv type 1 subtype b transmission in switzerland. *Journal of Infectious Diseases*, 201(10):1488–1497.
- Metzker, M. L., Mindell, D. P., Liu, X.-M., Ptak, R. G., Gibbs, R. A., and Hillis, D. M. (2002). Molecular evidence of hiv-1 transmission in a criminal case. *Proceedings of the National Academy of Sciences*, 99(22):14292–14297.
- Ou, C.-Y., Ciesielski, C. A., Myers, G., Bandea, C. I., Luo, C.-C., Korber, B. T., Mullins, J. I., Schochetman, G., Berkelman, R. L., Economou, A. N., et al. (1992). Molecular epidemiology of hiv transmission in a dental practice. *Science*, 256(5060):1165–1171.
- Posada, D. and Crandall, K. A. (1998). Modeltest: testing the model of dna substitution. *Bioinformatics*, 14(9):817–818.
- Self, S. G. and Liang, K.-Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association*, 82(398):605–610.
- Yang, Z. (2006). *Computational Molecular Evolution*. Oxford University Press.
- Yang, Z. (2014). *Molecular Evolution – A Statistical Approach*. Oxford University Press.

c07- Maximum likelihood method & testing

Searching tree space

Model testing

Which model fits best?

Uncertainty

Phylogenetics in HIV

HIV origin

HIV transmission trees

References