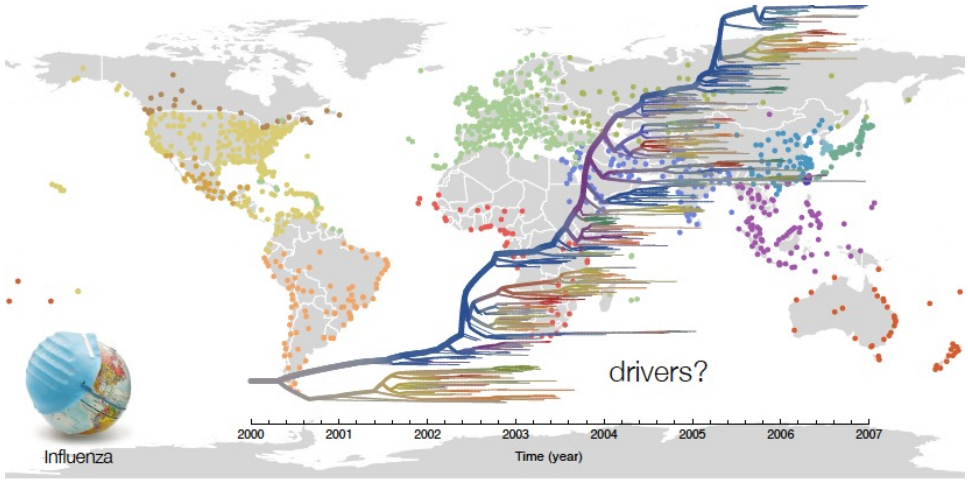
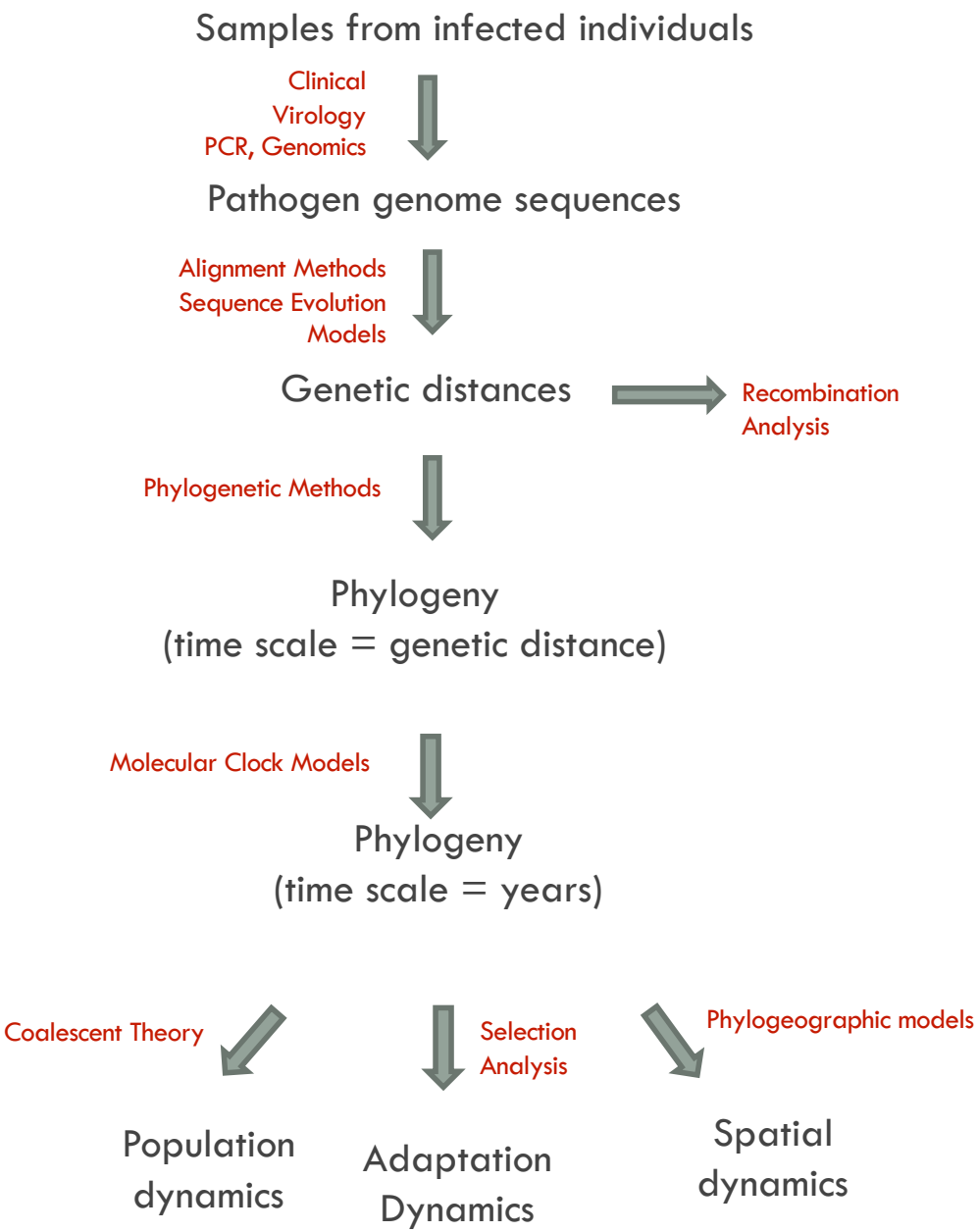
INTRODUCTION TO PHYLODYNAMICS

Marta Giovanetti
Oswaldo Cruz Foundation



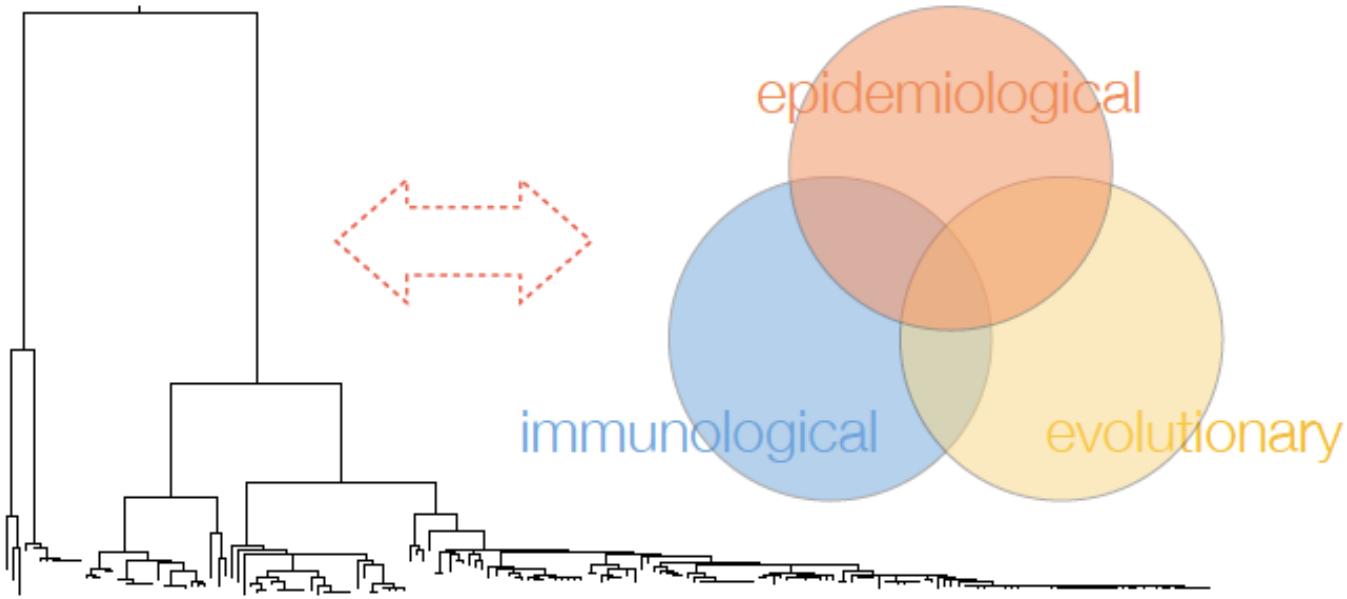

Samples from infected individuals
Clinical
Virology
PCR, Genomics
Pathogen genome sequences
Alignment Methods
Sequence Evolution
Models
Genetic distances
Recombination
Analysis
Phylogenetic Methods
Phylogeny
(time scale = genetic distance)
Molecular Clock Models
Phylogeny
(time scale = years)
Coalescent Theory
Selection
Analysis
Phylogeographic models
Population
dynamics
Adaptation
Dynamics
Spatial
dynamics

1. Definition of phylodynamics
2. Source phylodynamic variation
3. Phylodynamic techniques
4. Estimating Temporal Signal
5. Real-time phylodynamics

1. **Definition of phylodynamics**
2. Source phylodynamic variation
3. Phylodynamic techniques
4. Estimating Temporal Signal
5. Real-time phylodynamics

What is phylodynamics?

Viral phylodynamics is defined as the study of how epidemiological, immunological, and evolutionary processes act and potentially interact to shape viral phylogenies.

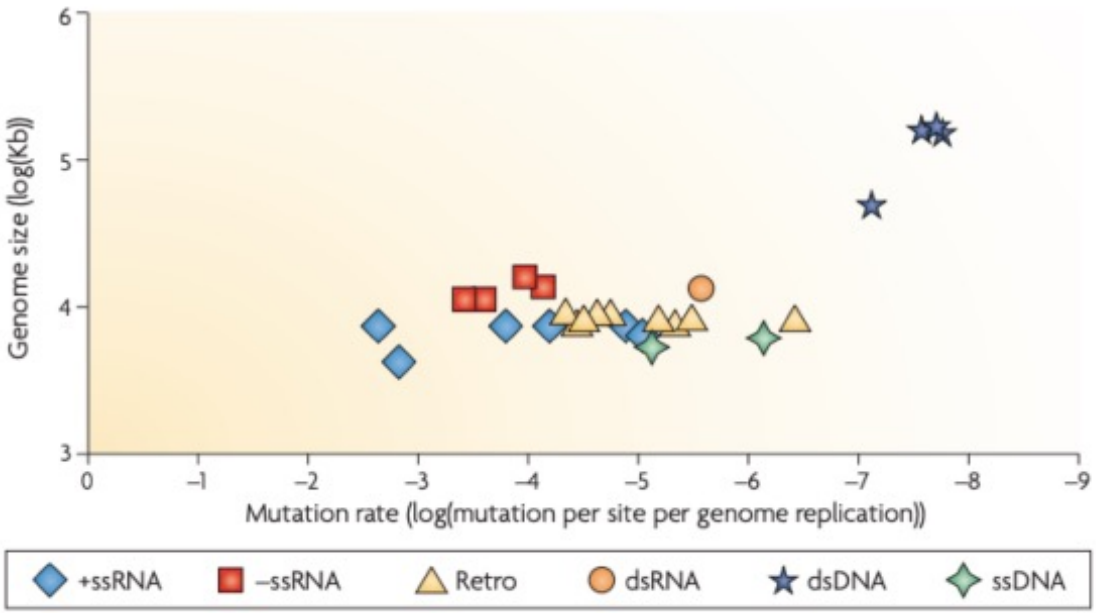


Greenfell et al. Science 2004

Why phylodynamics: viral dynamics

Epidemiological and evolutionary processes occur on a similar time scale.

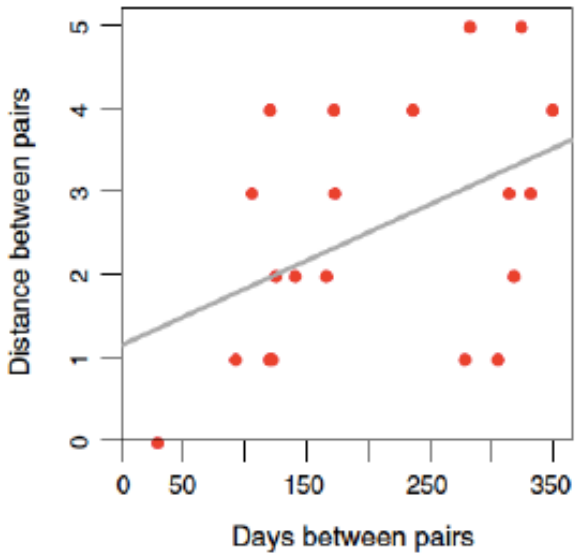
- High mutation rates
- Large population sizes
- Short generation times

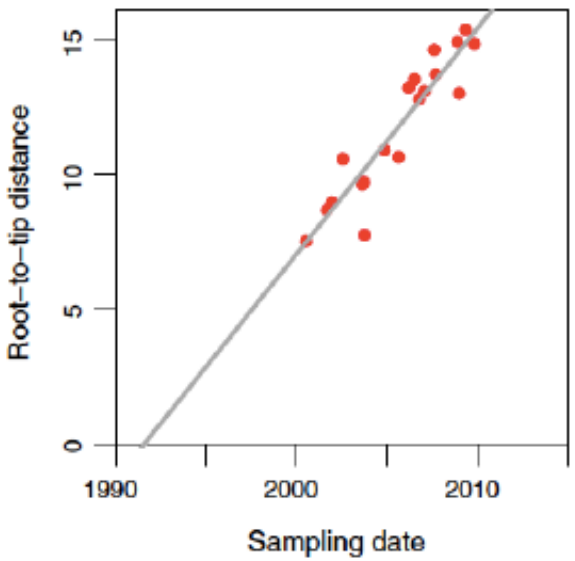


Duffy et al, 2008

Nucleotide substitutions / site / year	DNA viruses		RNA viruses
	$10^{-5} – 10^{-8}$	vs	$10^{-4} – 10^{-2}$

"RNA viruses are unique in that their ecological and evolutionary dynamics occur on the same timescale an therefore potentially interact"

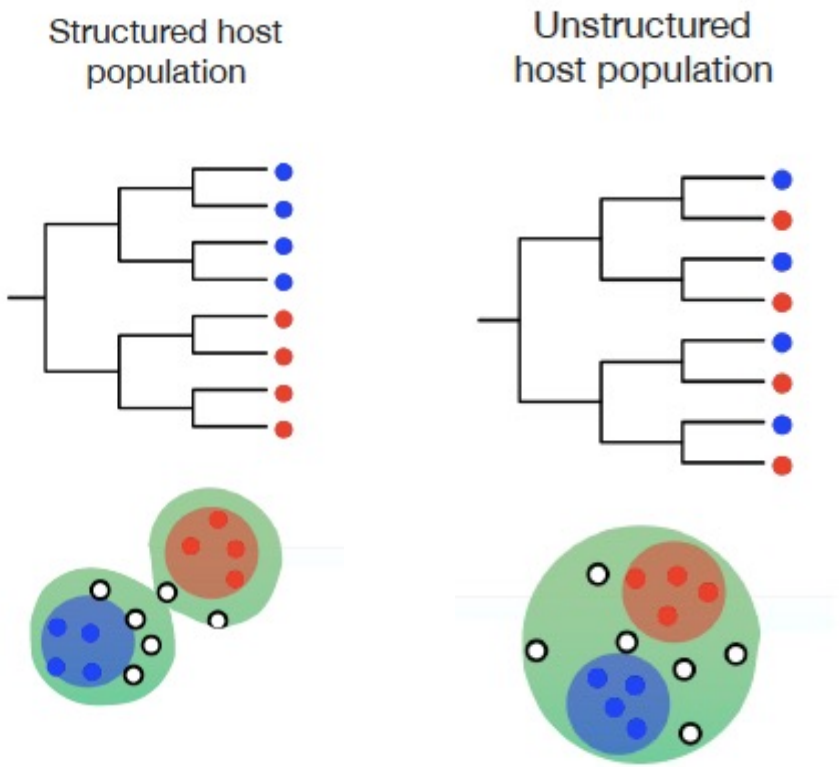




Temporal and spatial processes can be recovered from nucleotide sequences using phylodynamics models

1. Definition of phylodynamics
2. **Source phylodynamic variation**
3. Phylodynamic techniques
4. Estimating Temporal Signal
5. Real-time phylodynamics

1. Clustering of taxa on a viral phylogeny will be affected by host population structure

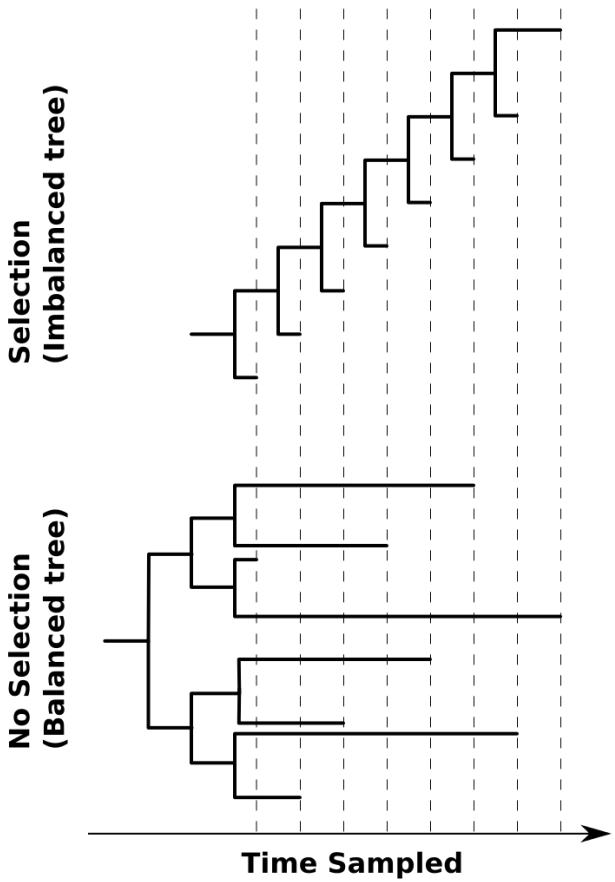


Viral Phylodynamics wiki (Volz, Koelle and Bedford, 2013)

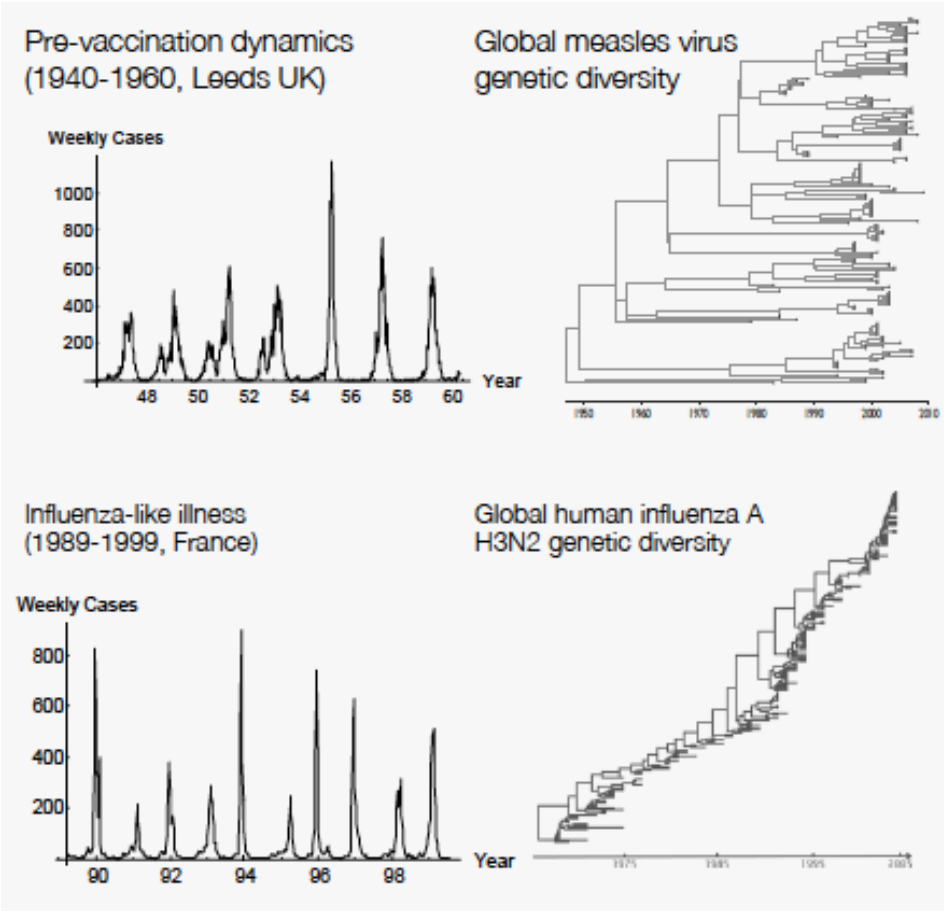
Population structure can exists at different levels, for example:

- Sampling dates:
  - Rambaut et al. 2017
- Geographic regions:
  - Gray R et al. AIDS 2012
  - Faria et al. Science 2014
- Body compartments:
  - Salemi et al. J Viral 2005
  - Jossefson et al. PNAS 2013
- Co-receptor usage:
  - Edo-Matas et al. MBE 2012
- Host species:
  - Streicker et al. Science 2010
  - Hill et al. www.virological.org
  - Trovao et al. MBE 2015
- HIV subtypes:
  - Ward et al. J Viral 2013

2. Tree balance will be affected by selection, most notably immune escape

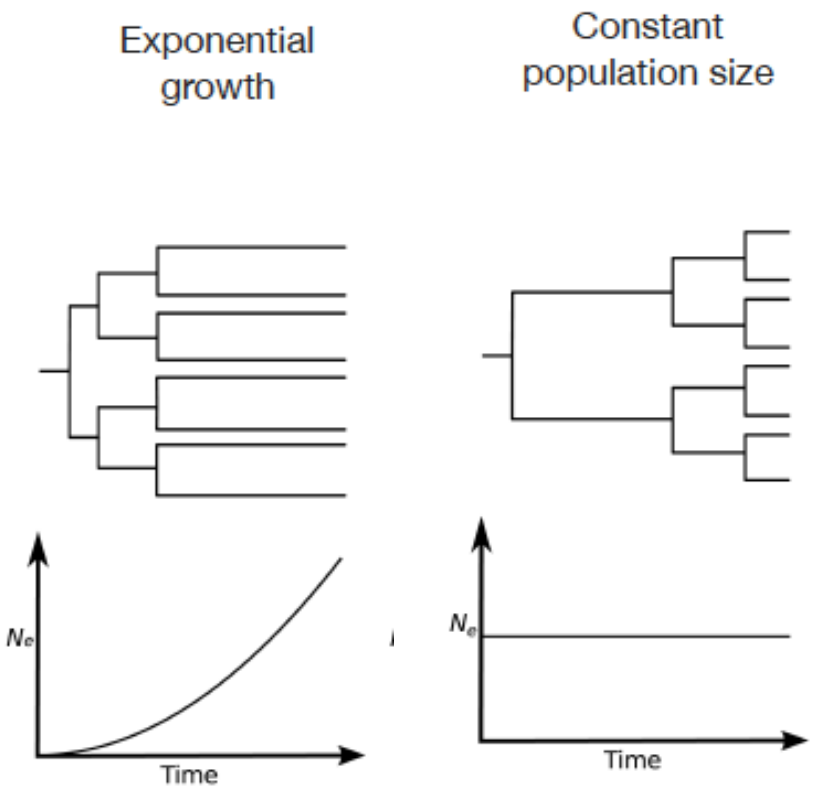


Viral Phylodynamics wiki (Volz, Koelle and Bedford, 2013)

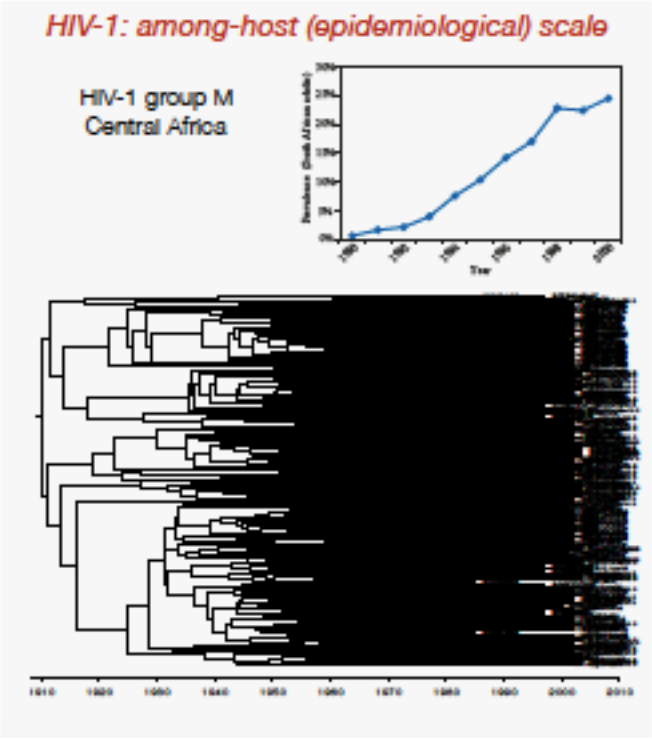


Adapted from Greenfell et al. Science 2004

3. The relative lengths of internal versus external branches will be affected by changes in viral population size over time

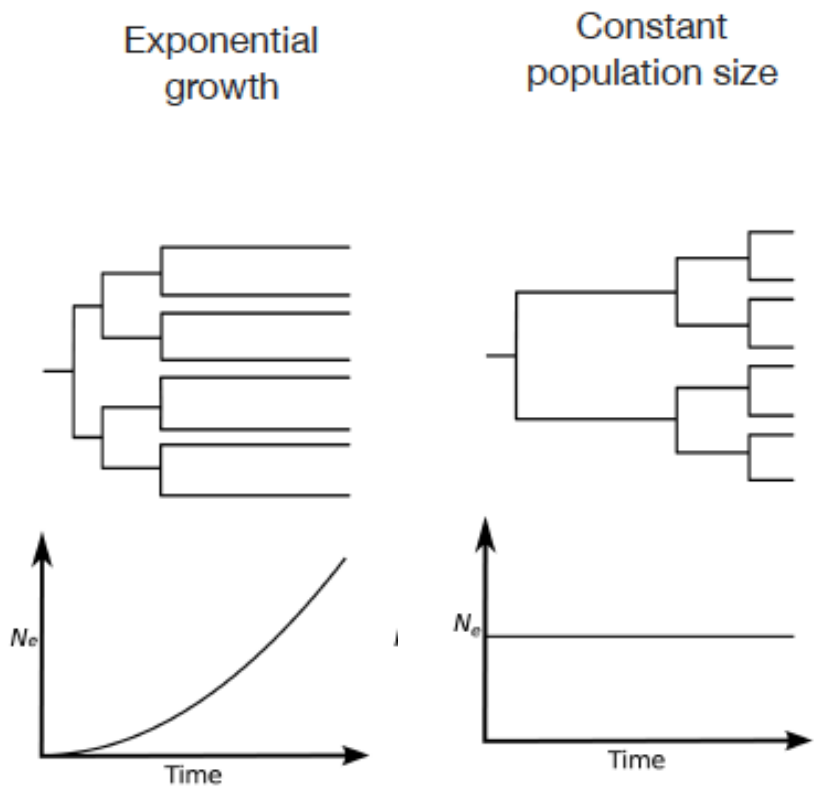


Viral Phylodynamics wiki (Volz, Koelle and Bedford, 2013)

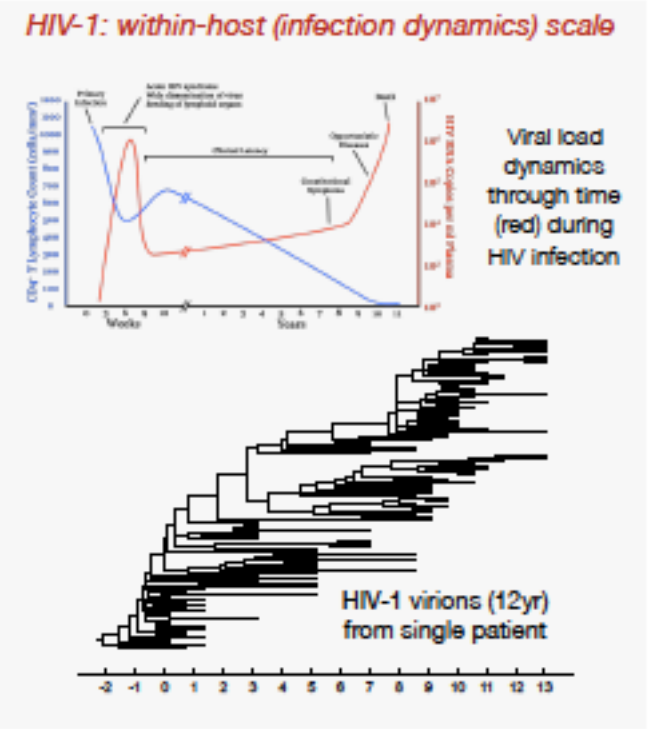


Adapted from Faria et al. Science 2014

3. The relative lengths of internal versus external branches will be affected by changes in viral population size over time



Viral Phylodynamics wiki (Volz, Koelle and Bedford, 2013)



Lemey et al. AIDS Rev 2004

Fundamental phylodynamic questions

How genetically diverse is a pathogen population?

How do pathogen genomes change through time?

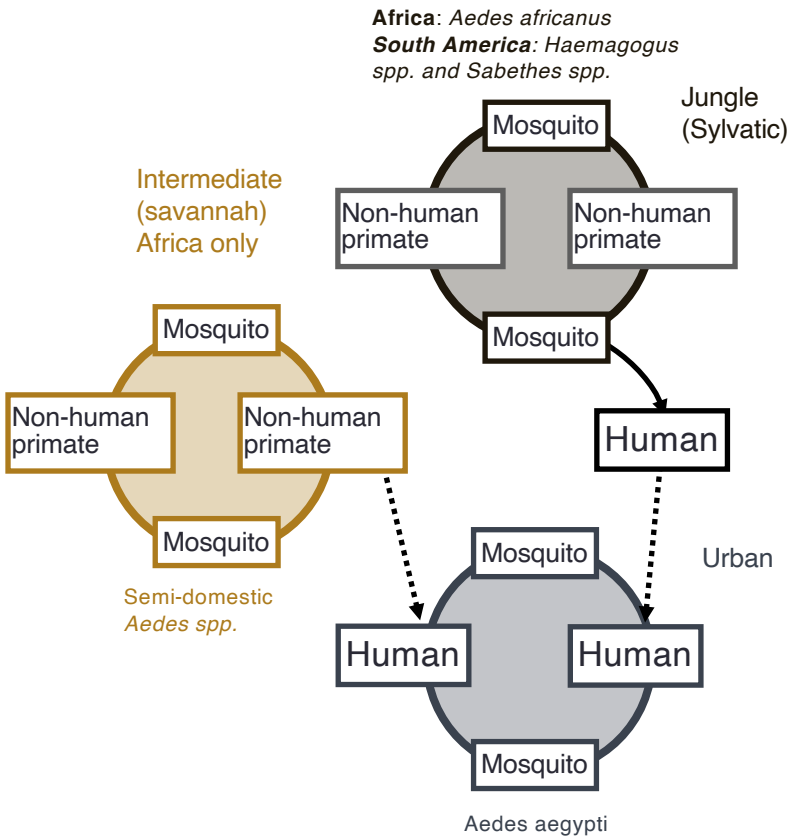
How does pathogen genetic diversity vary through space?

What processes and/or events determine these changes?

What are the effects of pathogen genetic diversity on virulence, transmissibility, resistance to treatment, etc

Phylodynamic inference issues addressed with genomic surveillance

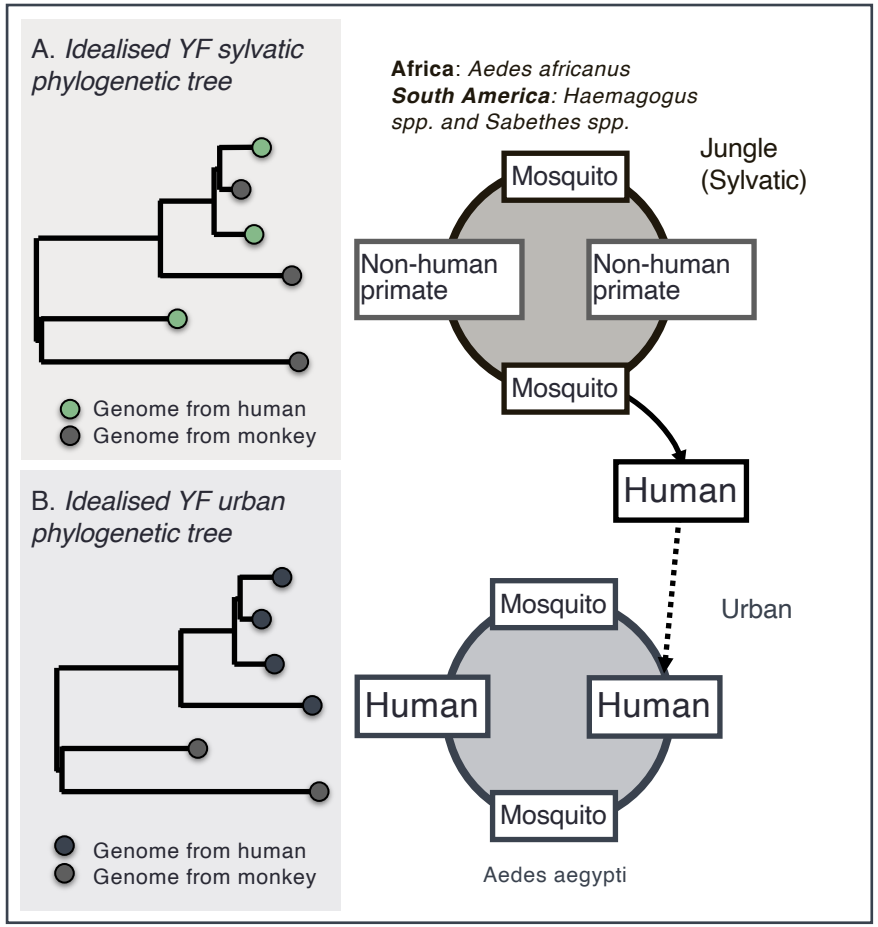
1. Are the strains associated with certain transmission routes?
2. How fast is it transmitting and in what direction is it spreading?
3. When did this epidemic start and where did it come from?
4. Are hosts A, B and C epidemiologically linked?
5. Can we predict future virus transmission?



https://www.cdc.gov/yellowfever/transmission

Phylodynamic inference issues addressed with genomic surveillance

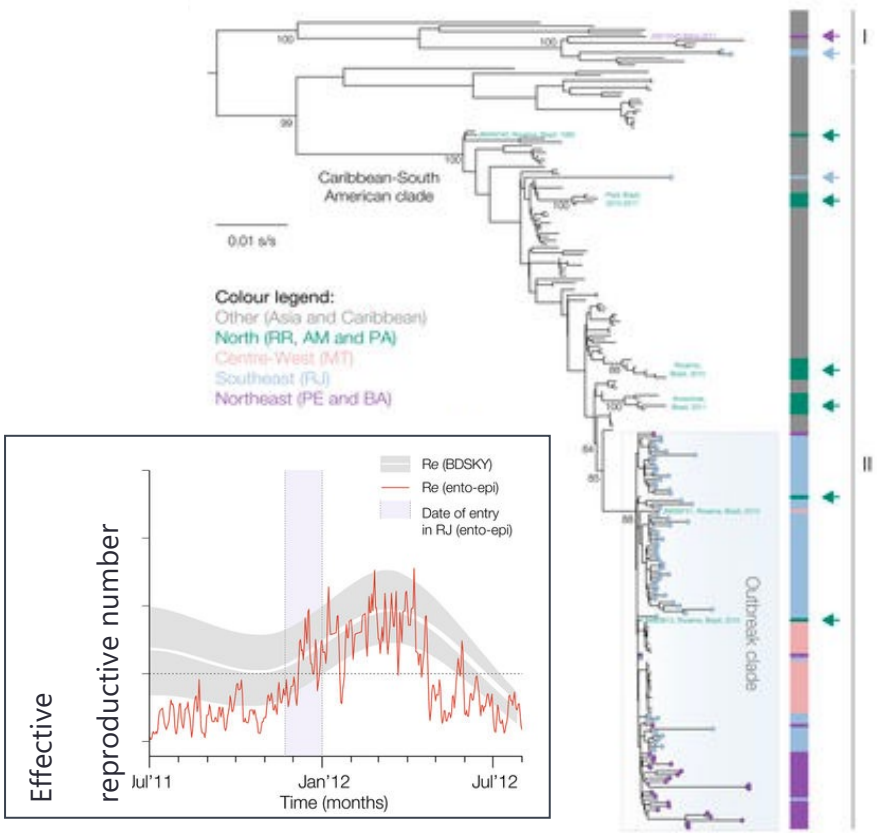
1. Are the strains associated with certain transmission routes?
2. How fast is it transmitting and in what direction is it spreading?
3. When did this epidemic start and where did it come from?
4. Are hosts A, B and C epidemiologically linked?
5. Can we predict future virus transmission?



https://www.cdc.gov/yellowfever/transmission

Phylodynamic inference issues addressed with genomic surveillance

1. Are the strains associated with certain transmission routes?
2. How fast is it transmitting and in what direction is it spreading?
3. When did this epidemic start and where did it come from?
4. Are hosts A, B and C epidemiologically linked?
5. Can we predict future virus transmission?



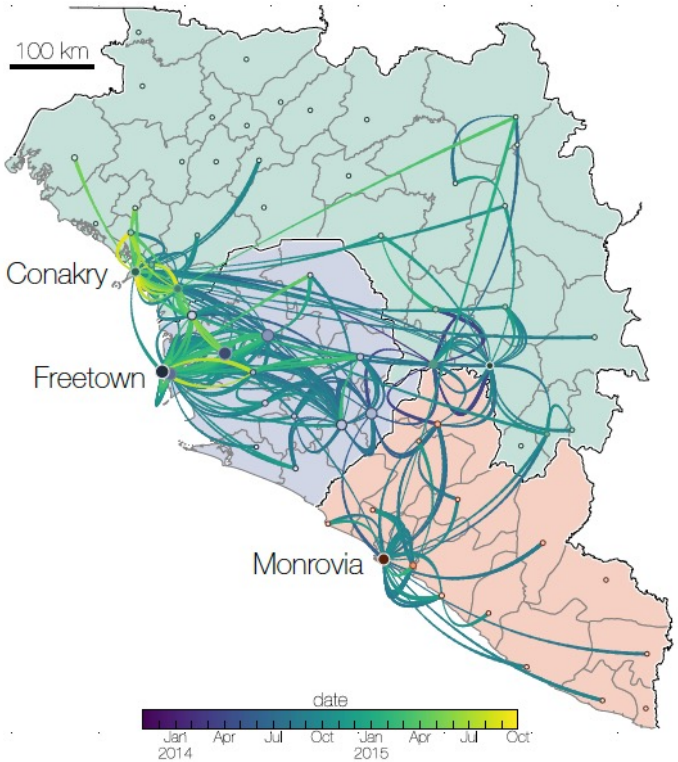
is the [expected number] of cases directly generated by one case in a population where all individuals are [susceptible] to infection.

Faria *et al. Scientific Reports* 7:15216 (2017)

Phylodynamic inference issues addressed with genomic surveillance

1. Are the strains associated with certain transmission routes?
2. How fast is it transmitting and in what direction is it spreading?
3. When did this epidemic start and where did it come from?
4. Are hosts A, B and C epidemiologically linked?
5. Can we predict future virus transmission?

Province-level Ebola virus migration in West Africa

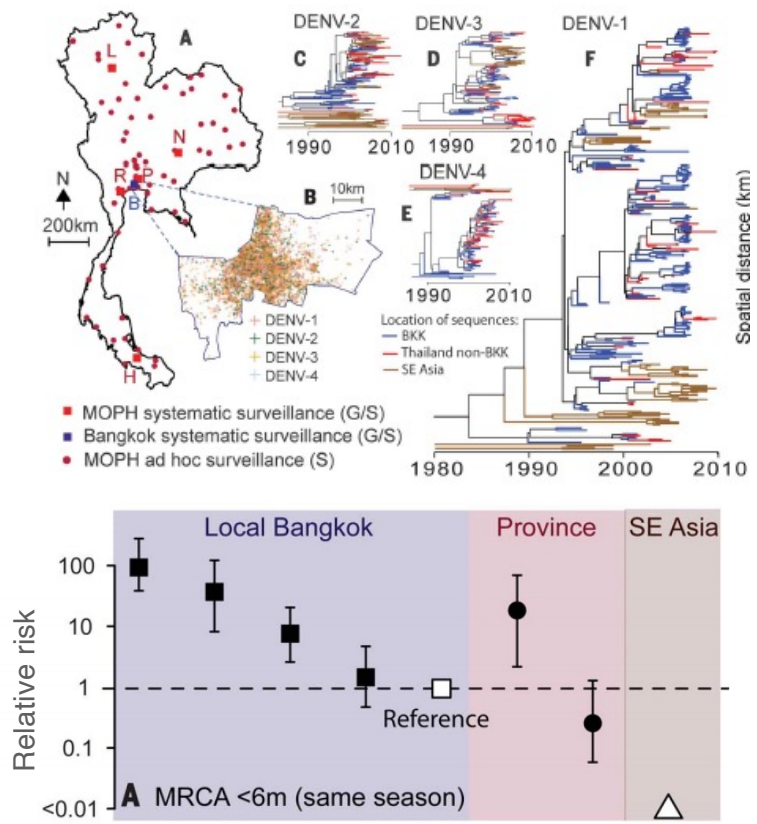


Dudas *et al*. *Nature* 544 309-315 (2017)

Phylodynamic inference issues addressed with genomic surveillance

1. Are the strains associated with certain transmission routes?
2. How fast is it transmitting and in what direction is it spreading?
3. When did this epidemic start and where did it come from?
4. Are hosts A, B and C epidemiologically linked?
5. Can we predict future virus transmission?

Sources of same-season transmission

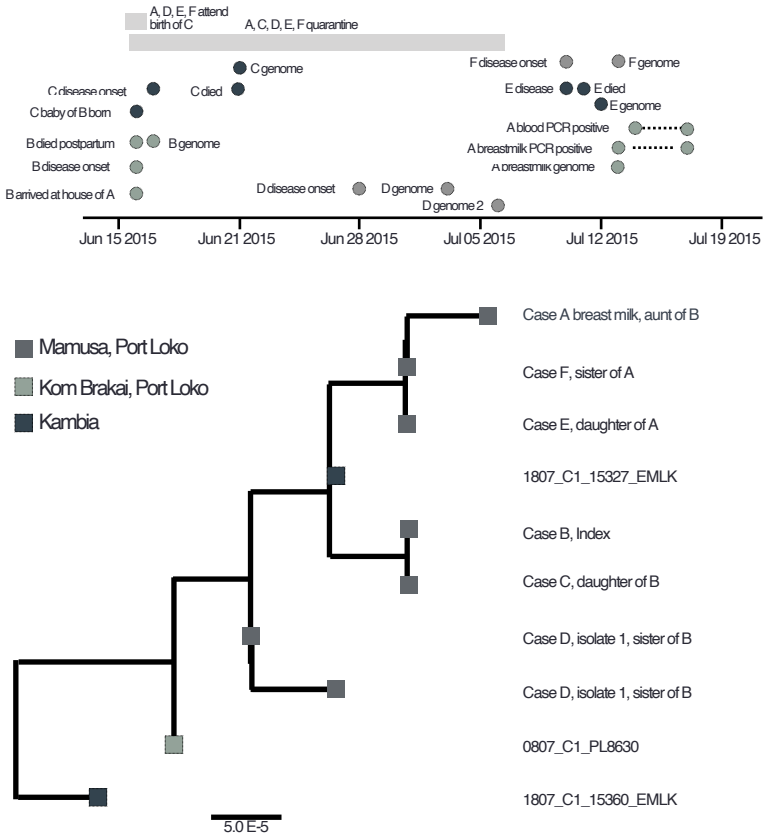


Salje *et al. Science* 355 1302-1306 (2017)

Phylodynamic inference issues addressed with genomic surveillance

1. Are the strains associated with certain transmission routes?
2. How fast is it transmitting and in what direction is it spreading?
3. When did this epidemic start and where did it come from?
4. Are hosts A, B and C epidemiologically linked?
5. Can we predict future virus transmission?

Breast milk transmission of Ebola virus

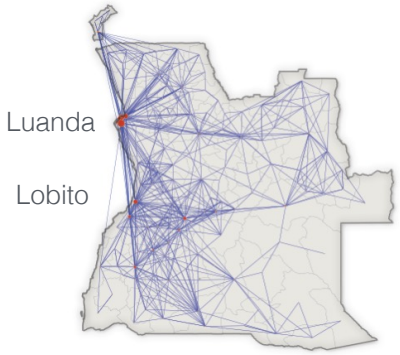


Arias *et al.* *Virus Evolution* 2 1-10 (2016)

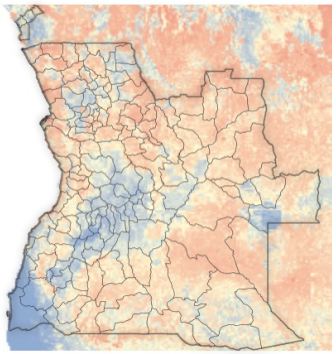
Phylodynamic inference issues addressed with genomic surveillance

1. Are the strains associated with certain transmission routes?
2. How fast is it transmitting and in what direction is it spreading?
3. When did this epidemic start and where did it come from?
4. Are hosts A, B and C epidemiologically linked?
5. Can we predict future virus transmission?

Human mobility patterns for Angolan districts (n=163)

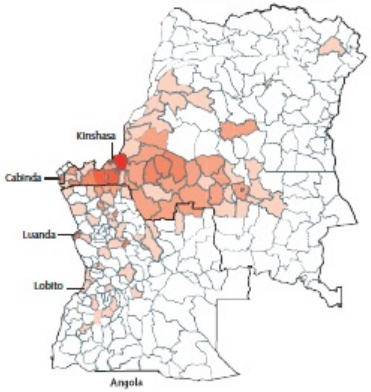


Aedes aegypti suitability



Predictions of YFV (4 week ahead)

Predictions of YFV (8 week ahead)



Kraemer *et al. Lancet Inf Dis* 17 330-38 (2017)

1. Definition of phylodynamics
2. Source phylodynamic variation
3. **Phylodynamic techniques**
4. Estimating Temporal Signal
5. Real-time phylodynamics

Most common phylodynamic techniques

Techniques	Softwares (e.g.)
Coalescent models	BEAST, CoaSim, Lamarc, Migrate, simcoal2, IcyTree
Phylogeography	BEAST, Seraphim, Migrate, MacClade, Hyphy
Selection Analysis	Hyphy, BEAST, Mega, PAML
Recombination Analysis	RDP, SplitsTree, SimPlot, VisRD

393 phylogeny packages and 54 free web servers at:
http://evolution.genetics.washington.edu/phylip/software.html

Bayesian

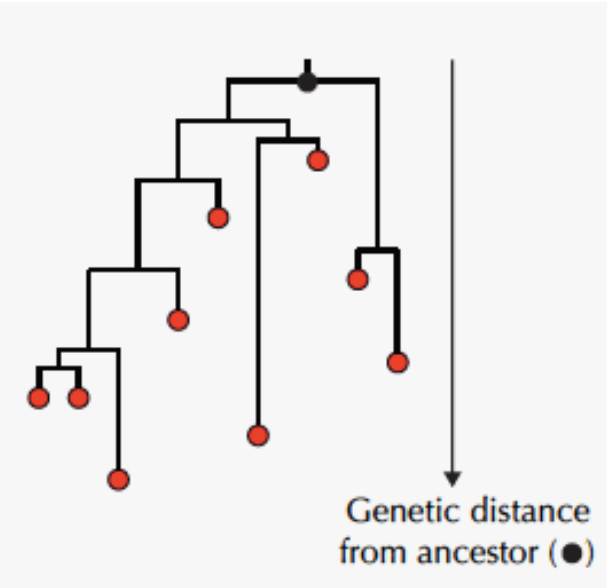
- Probabilistic method - reverse probability

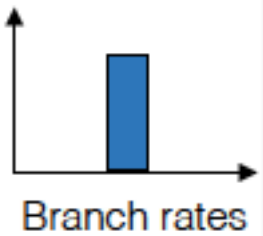
- In a Bayesian approach, the **posterior probability** distribution is obtained by multiplying the "likelihood" for the "**a priori probability**" of each tree according to the Bayes' theorem.

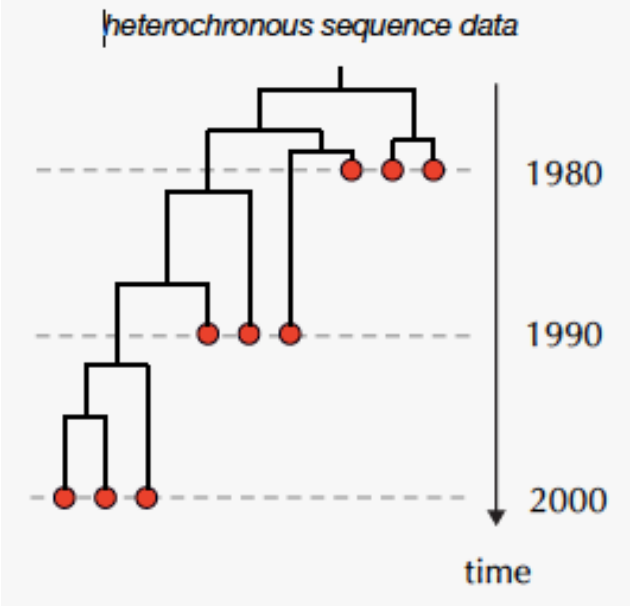
Phylodynamic tools
(1) The molecular clock

Strict molecular clock

Substitution rates: One rate rules them all



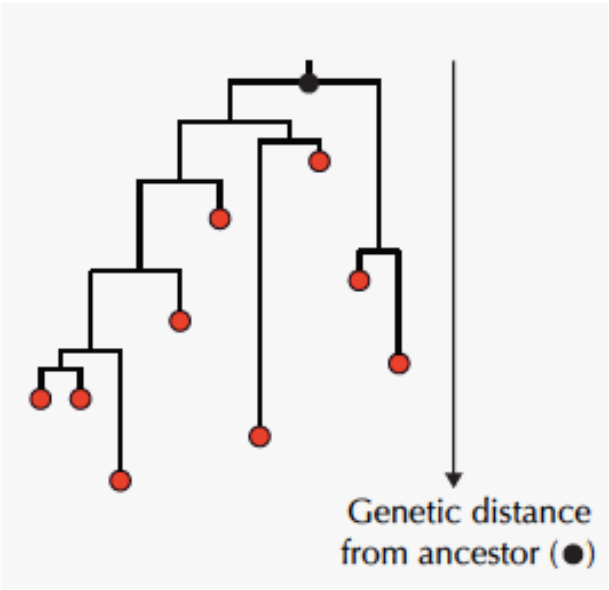


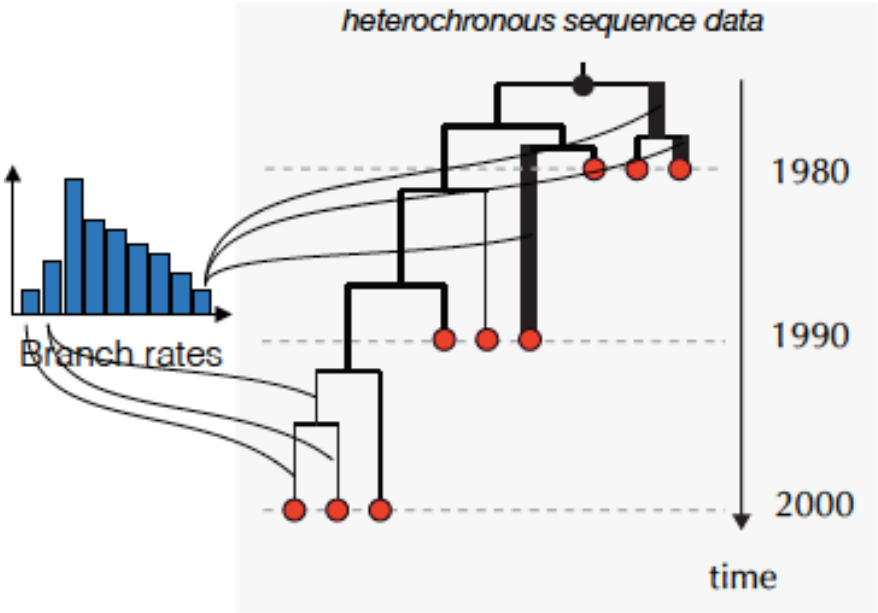


Phylodynamic tools
(1) The molecular clock

Relaxed molecular clock

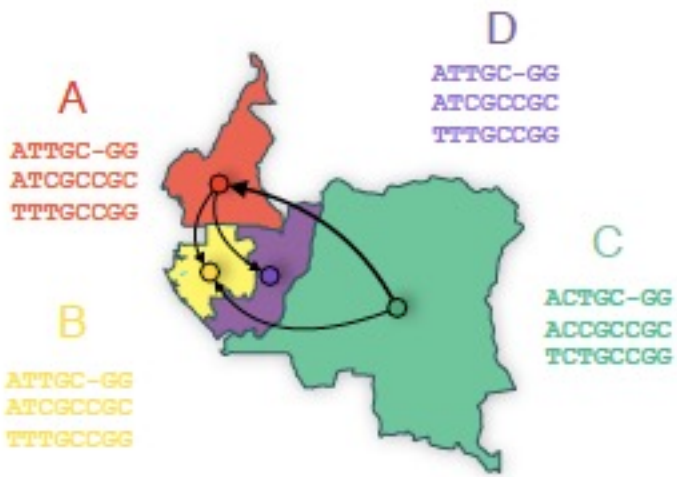
Substitution rates can vary among lineages according
to an underlying distribution (e.g. lognormal, exponential, gamma)

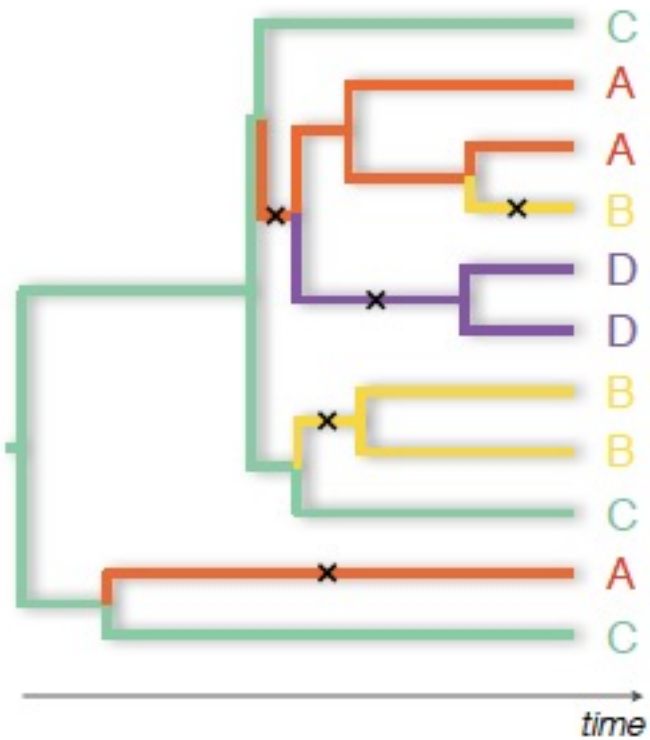




Drummond et al. PLoS Biol 2006

Phylodynamic tools
(2) Phylogeography

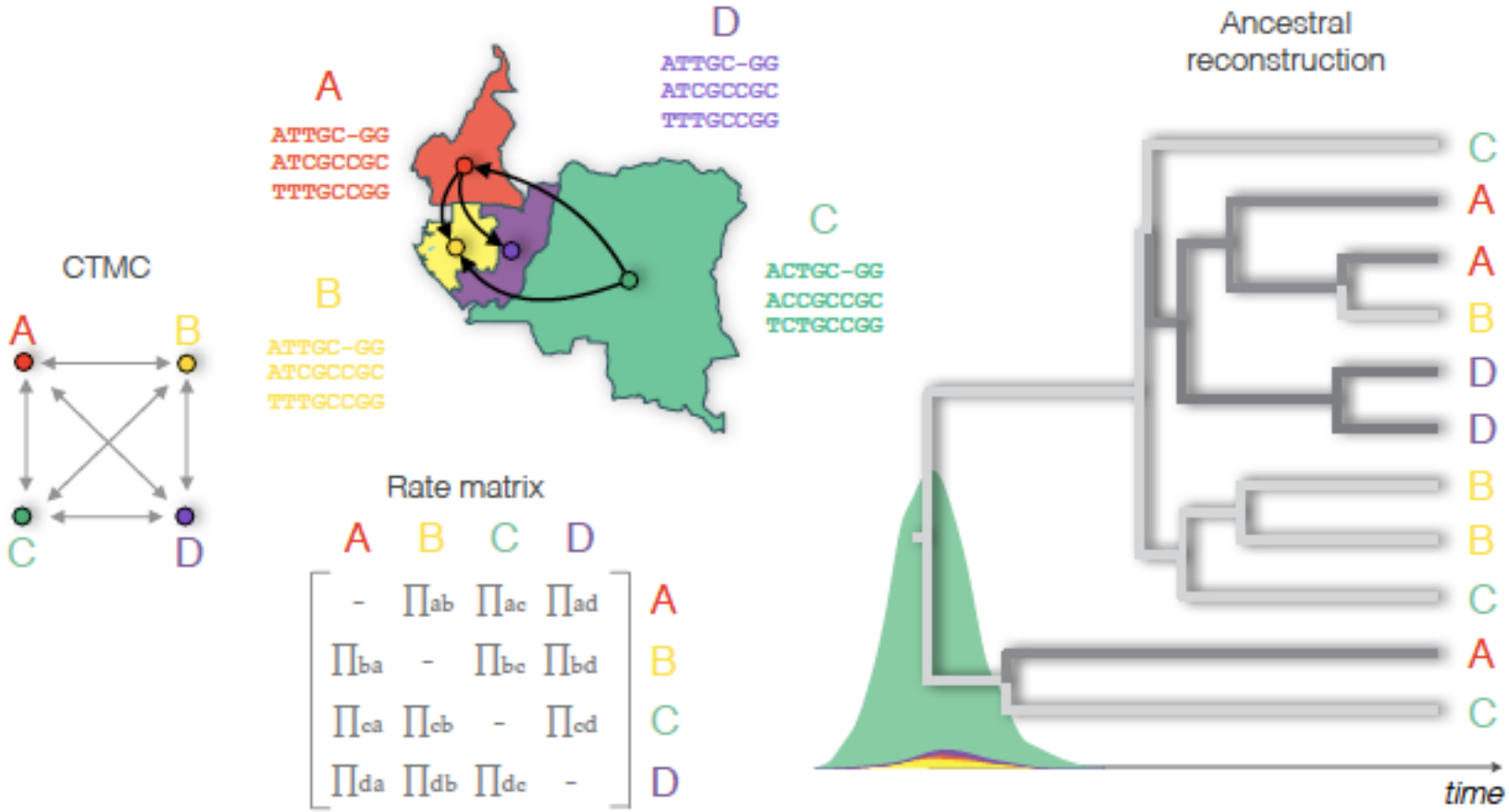




Definition of phylogeography
"a field of study concerned with the principles and processes governing the geographic distribution of genealogical lineages, especially those within and among closely related species" Avise, 2000

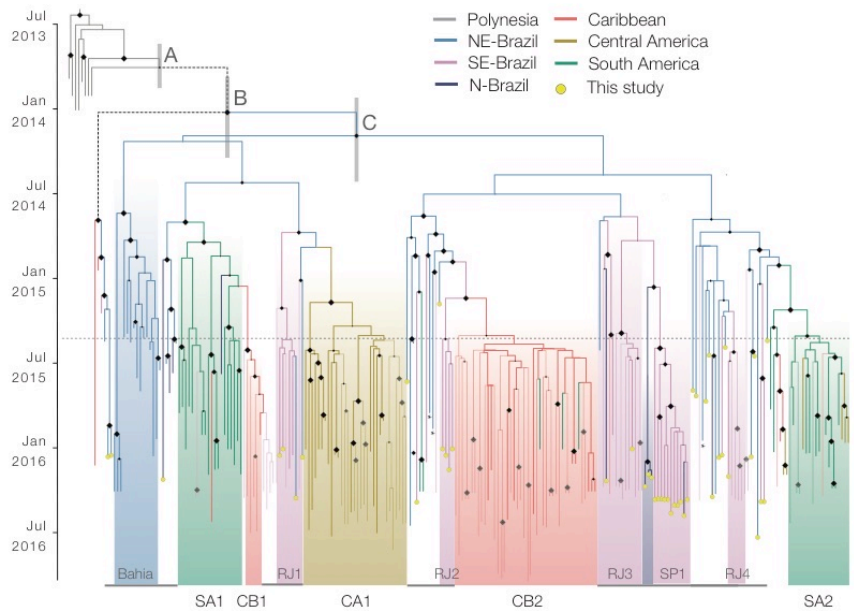
Lemey et al., PLoS Comp Biol, 2009, Talbi et al. PLoS Pathogens, 2011; Drummond et al.

Phylodynamic tools
(2) Phylogeography

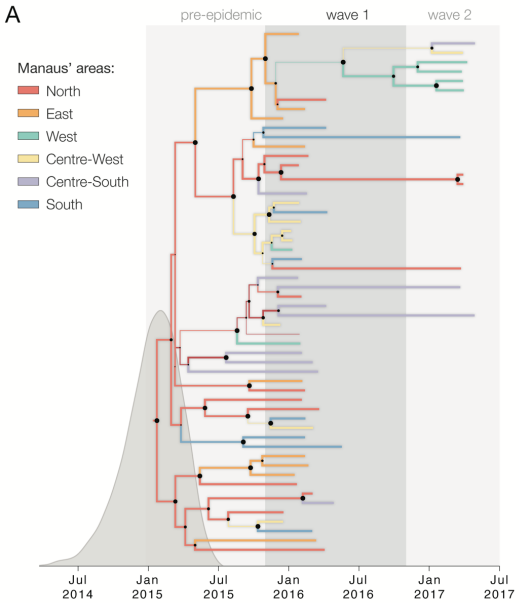


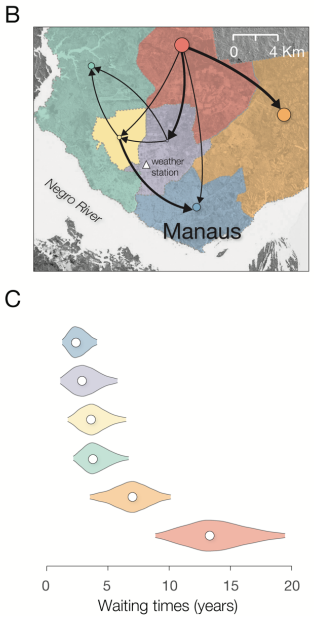
Lemey et al., PLoS Comp Biol, 2009, Talbi et al. PLoS Pathogens, 2011; Drummond et al. MBE 2012; O'Brien et al. MBE 2009; Minin & Suchard, J Math Biol 2008

Phylodynamic tools
(2) Phylogeography



Faria et al. Nature 2017, 10.1038/nature22401

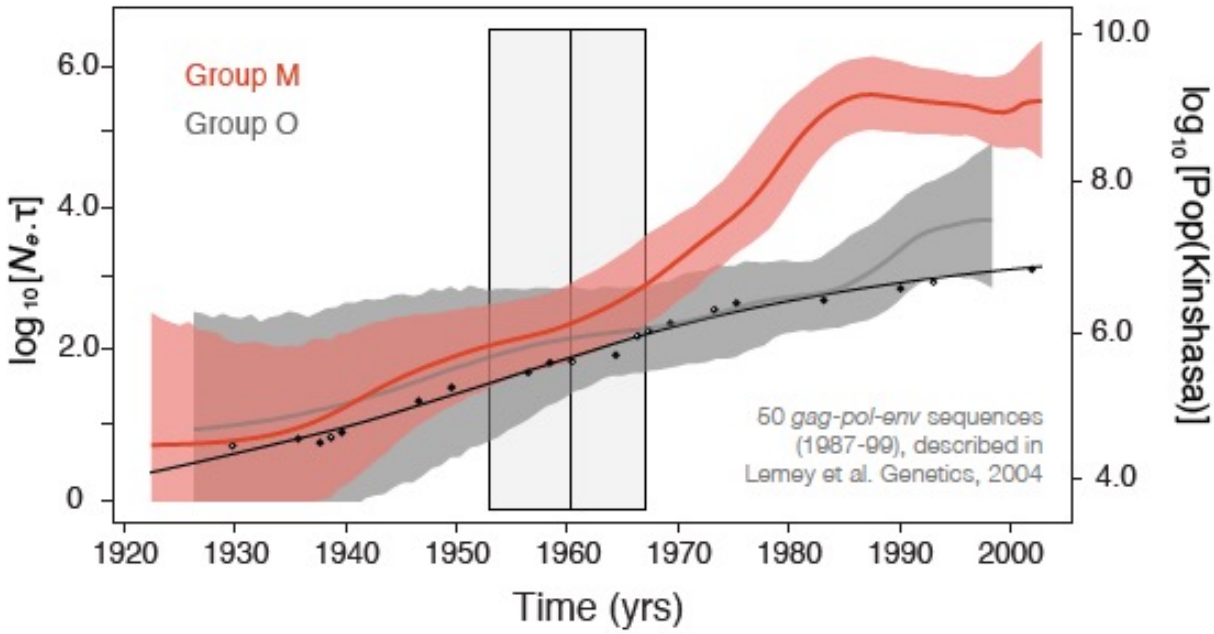




Giovanetti et al. *Cell Reports*, 2020

Phylodynamic tools
(3) Coalescent theory

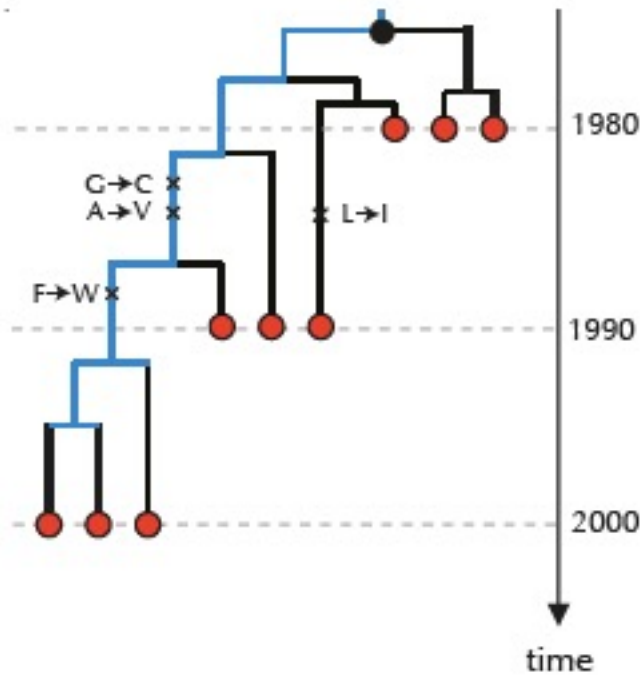
e.g. non-parametric coalescent method employed to estimate population dynamics of HIV-1 group M

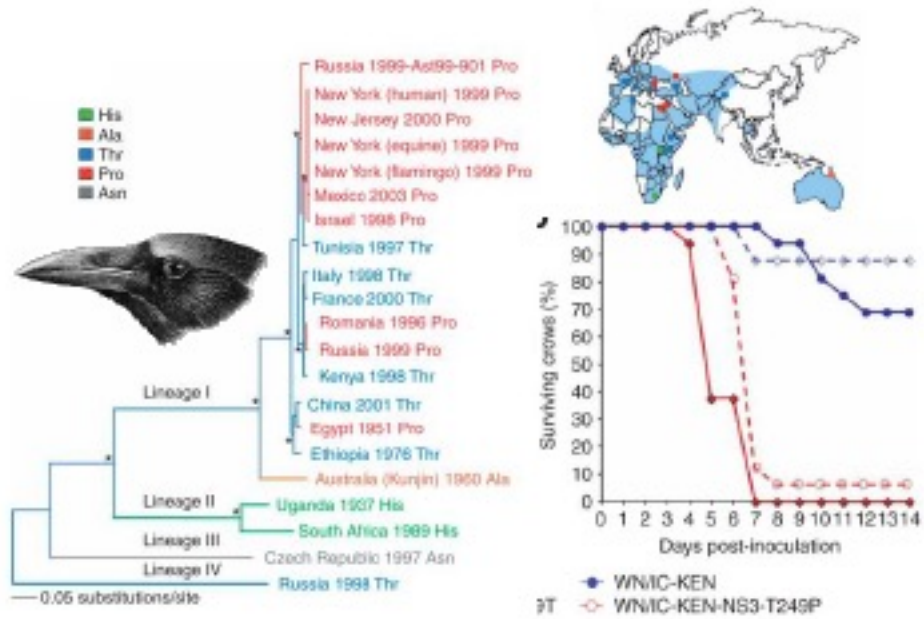


Faria et al. Science 2014

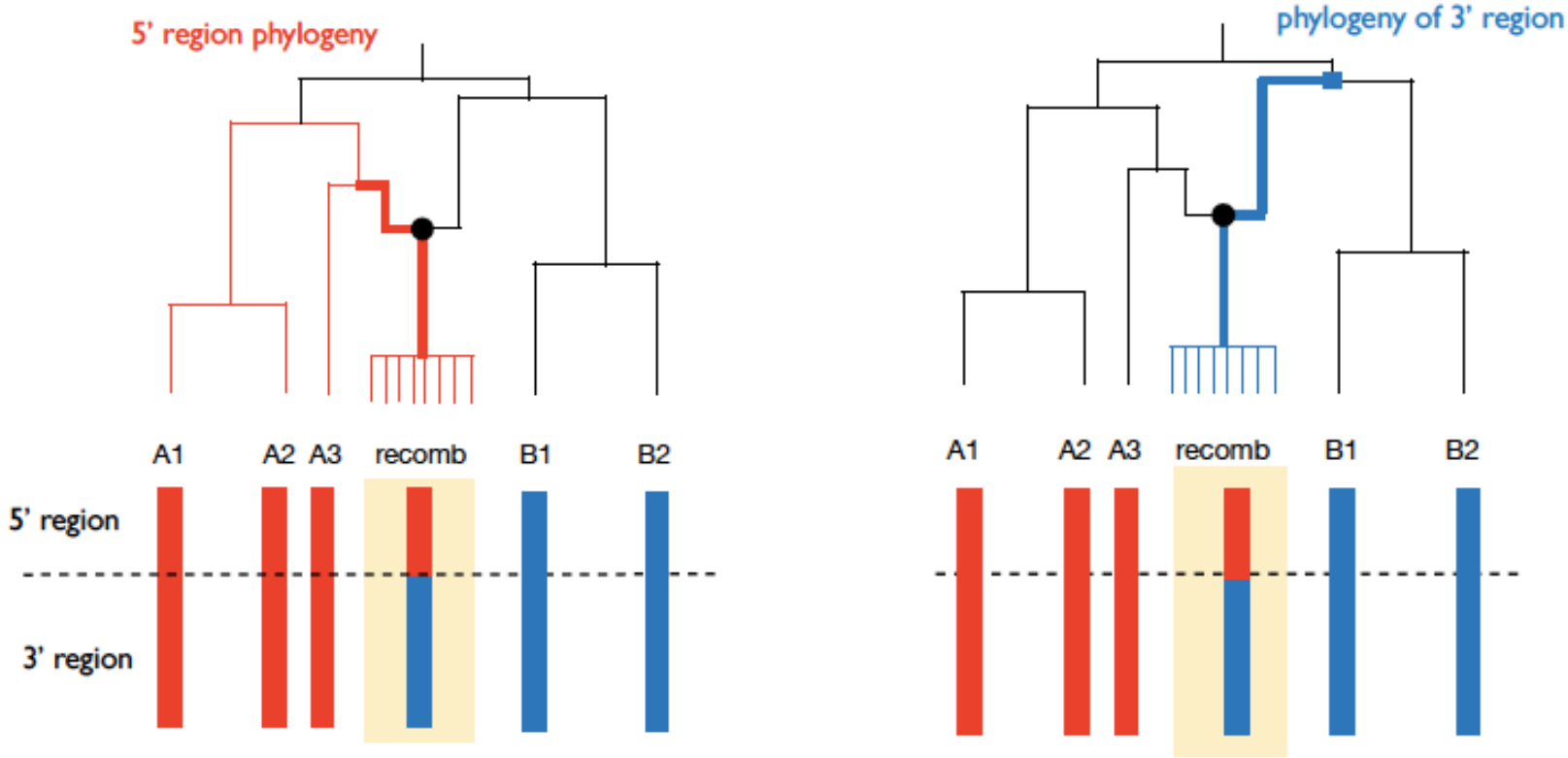
Phylodynamic tools
(4) Selection Analysis

e.g. Red lines and symbols represent viruses with the NS3-T249P substitution, and blue lines and symbols represent viruses with NS3-Thr249





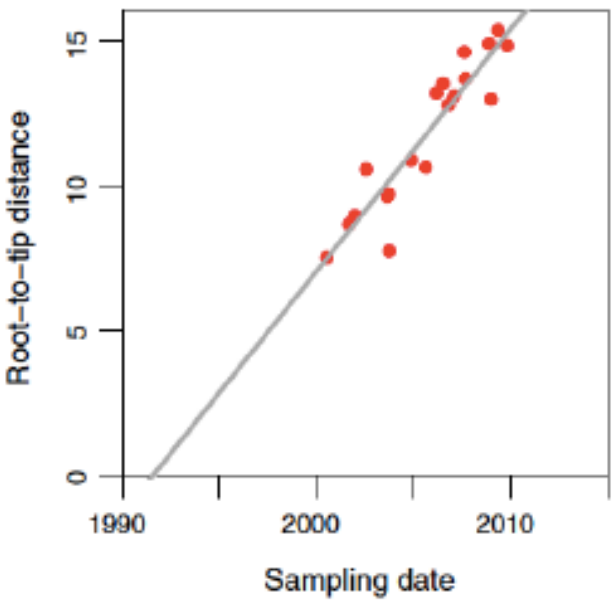
Phylodynamic tools
(5) Recombination analysis



1. Definition of phylodynamics
2. Source phylodynamic variation
3. Phylodynamic techniques
4. **Estimating Temporal Signal**
5. Real-time phylodynamics

Investigating temporal signal

- Pathogen genomes are sampled at different points in time
- Hence transmission history is estimated on a real time-scale (e.g. years)
- The ability to genetically distinguish sequences sampled at different times dependants on:
  - the rate of evolution of the gene / genome that is obtained
  - the length of time between samples
  - the sequence length of the gene/ genome that is obtained



Biek et al. TREE 2015

TempEst

It is able to examine the temporal signal of a data set and to look for problematic or erroneous sequences.

It just needs to have an imput which it will be the ML tree

TempEst - ice_viruses.fasta.treefile

Best-fitting root

Function: heuristic resid...

Dated Tips	
Date range	88
Slope (rate)	3.2905E-3
X-Intercept (T...	1874.4146
Correlation Co...	0.3504
R squared	0.1228
Residual Mean...	5.5861E-2

Sample Dates | Tree | Root-to-tip | Residuals | Node density

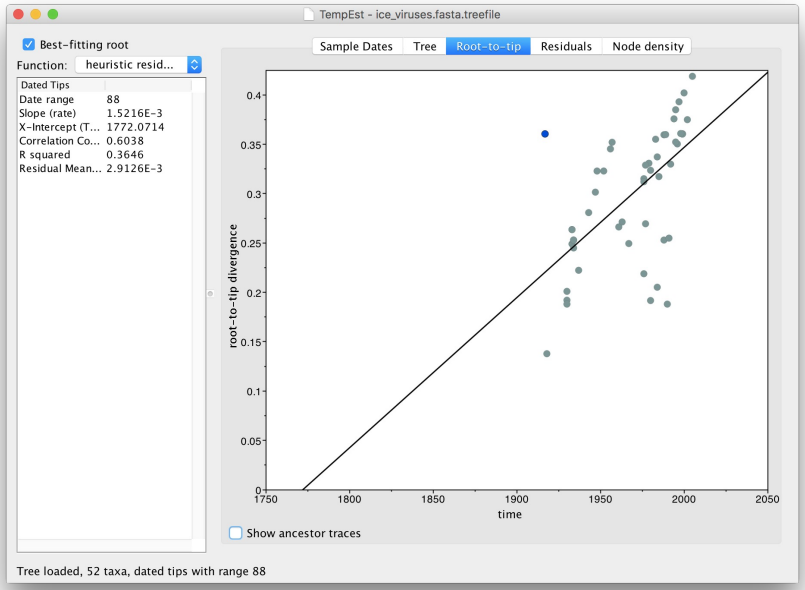
Import Dates | Parse Dates | Clear Dates | Dates specified as

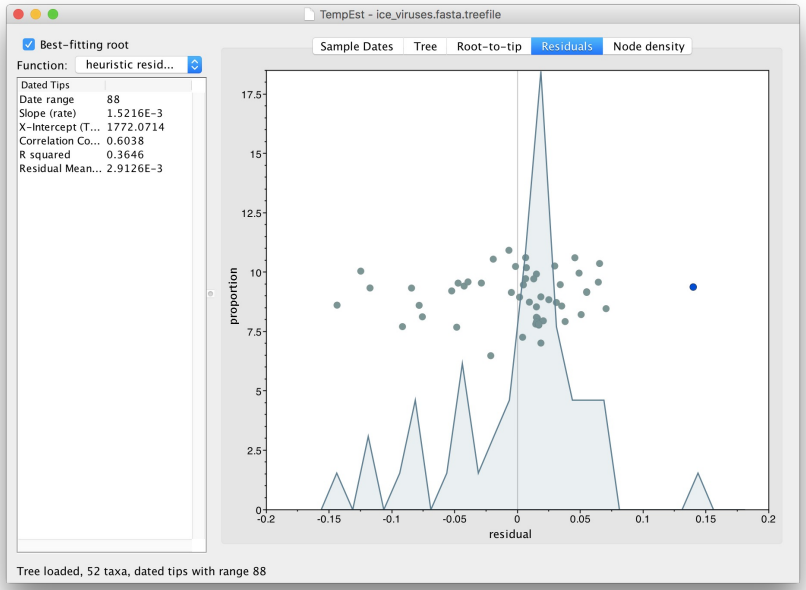
Name	Date	Precis...	Height
A.SouthCarolina.6.1988.L19025_1988	1988.0	0.0	17.0
A.Yamagata.32.1989.D31949_1989	1989.0	0.0	16.0
A.BuenosAires.T114.97.AF534026_1997	1997.0	0.0	8.0
A.Beijing.262.95.AY289928_1995	1995.0	0.0	10.0
A.Alaska.1173.00.AY029287_2000	2000.0	0.0	5.0
A.Finland.4.1984.L33491_1984	1984.0	0.0	21.0
A.Fiji.15899.83.AJ289702_1983	1983.0	0.0	22.0
A.Kiev.59.1979.M38353_1979	1979.0	0.0	26.0
A.FortMonmouth.1.1947.U02085_1947	1947.0	0.0	58.0
A.Meguro.1.1956.AB043485_1956	1956.0	0.0	49.0
A.Saga.2.1957.AB043486_1957	1957.0	0.0	48.0
A.Kojiya.1.1952.AB043482_1952	1952.0	0.0	53.0
A.BrevigMission.1.18.AF116575_1918	1918.0	0.0	87.0
A.goose.HongKong.8.1976.U46021_1976	1976.0	0.0	29.0
A.mallard.Tennessee.11464.85.AF091311_1985	1985.0	0.0	20.0
A.teal.Alberta.141.1992.CY004539_1992	1992.0	0.0	13.0
A.pintail.Ohio.25.1999.CY017725_1999	1999.0	0.0	6.0
A.mallard.Ohio.66.1999.CY016955_1999	1999.0	0.0	6.0
A.mallard.Ohio.56.1999.CY012824_1999	1999.0	0.0	6.0
A.Teal.Ohio.72.1999.CY017717_1999	1999.0	0.0	6.0
A.BrantGoose.1.1917.AY095226_1917	1917.0	0.0	88.0

Tree loaded, 52 taxa, dated tips with range 88

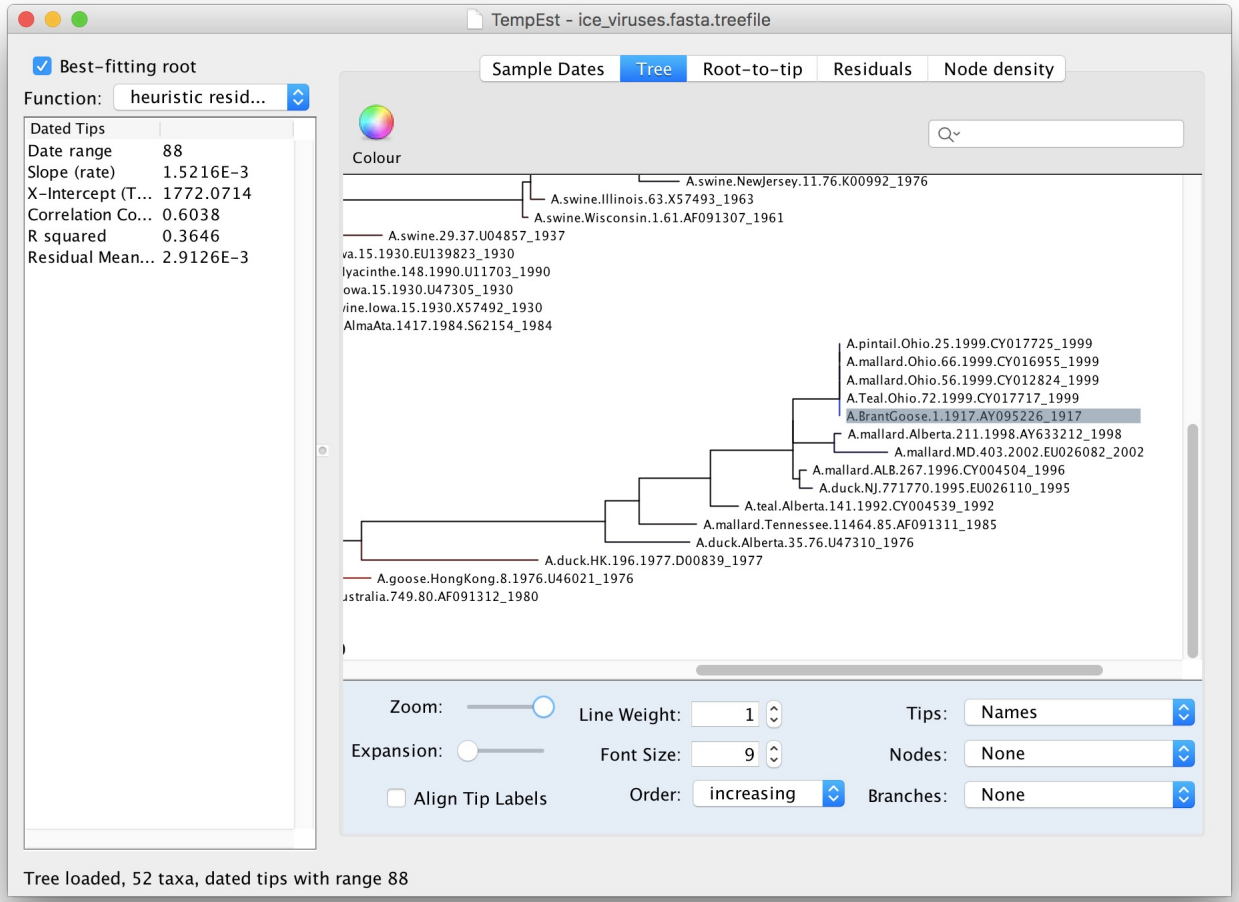
Finding and interpreting problematic sequences

Switch to the Root-to-tip panel. Look for the point furthest from the line in the top left hand quadrant. If you click and drag your pointer over the point it will be highlighted in a blue colour:

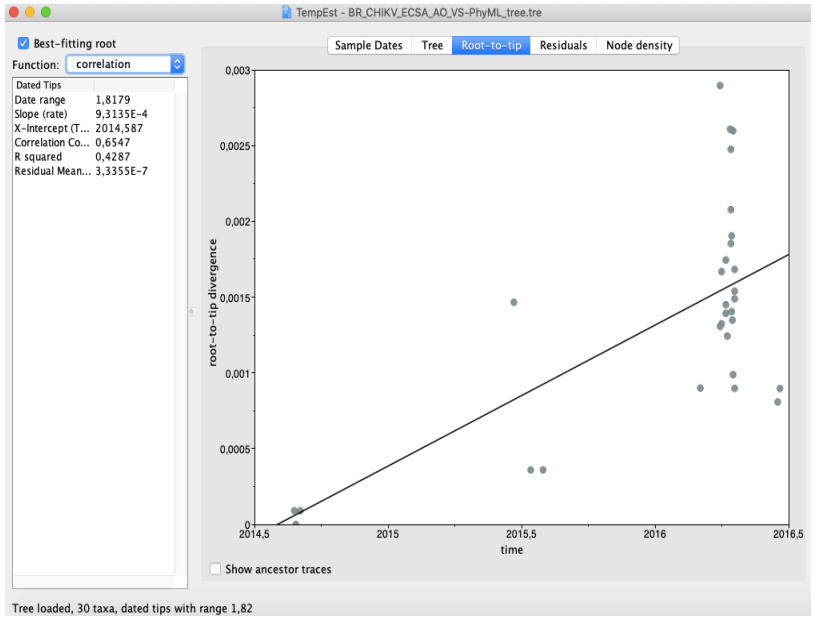




Is looking such an outlier.



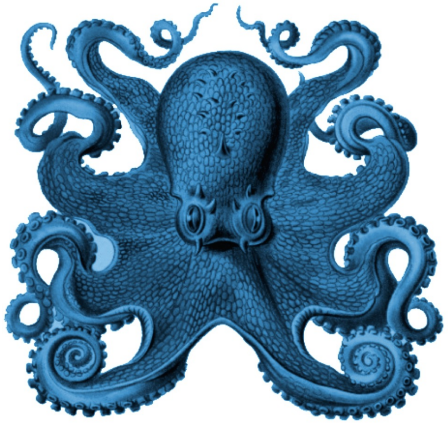
To check for a correlation we will need to choose the «Best-fitting-root» and the «correlation option»



the correlation we were observed was not the correct one as the root was arbitrarily set by the phylogeny reconstruction software and thus divergence from root was meaningless.

Does our dataset has a temporal singal ?

Dated Tips	
Date range	88
Slope (rate)	l.5216E-3
X-Intercept (TMRCA)	1772.0714
Correlation Cofficient	0.6038
R squared	0.3646
Resid uai Mean Squared	2.9126E-3



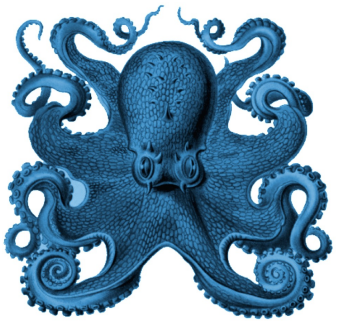
What is BEAST ?

BEAST is a Bayesian MCMC analysis package

Unlike MrBayes we have concentrated on rooted trees with timescale

Numerous models are included and can be used in many different combinations

What can we estimate?



Population genetics:
Population sizes
Growth/decline in population
Bottlenecks/transition points
Gene trees/species trees

Rates of evolution:
Dates of MRCAs
Times of bottlenecks/ transitions
Rates for genes
Rates for lineages

Phylogenetic:
Founder-effects
Epidemiological tracking
Phylogeography
Trait evolution
Virus transmission dynamics

Of what does BEAST comprise?

Setting up the model in BEAUti…

BEAUti

BEAUti v1.8.1.app

Utility program for creating BEAST input files

Accepts alignment files as input

imports FASTA format

Allows ages of samples to be input

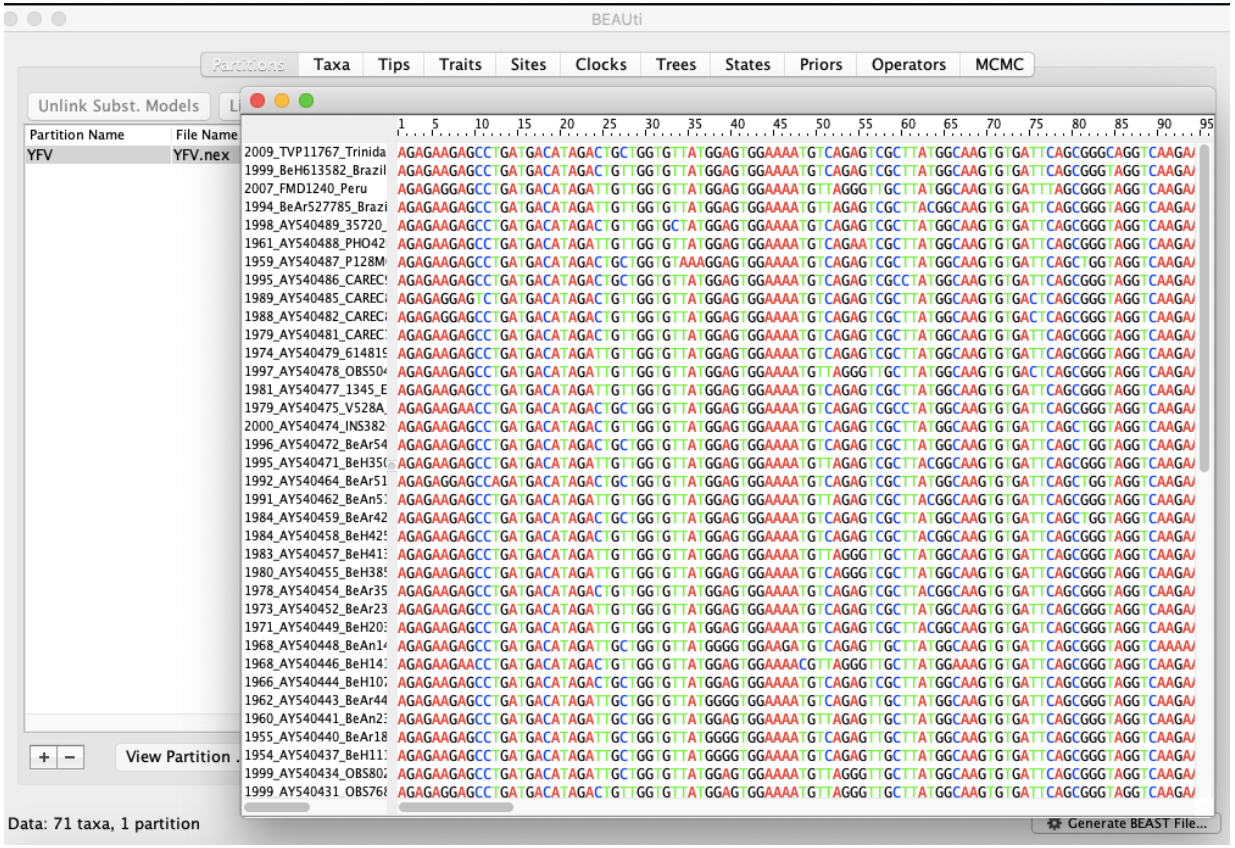
Allows selection of strict and relaxed clocks

Allows selection of nucleotide substitution models

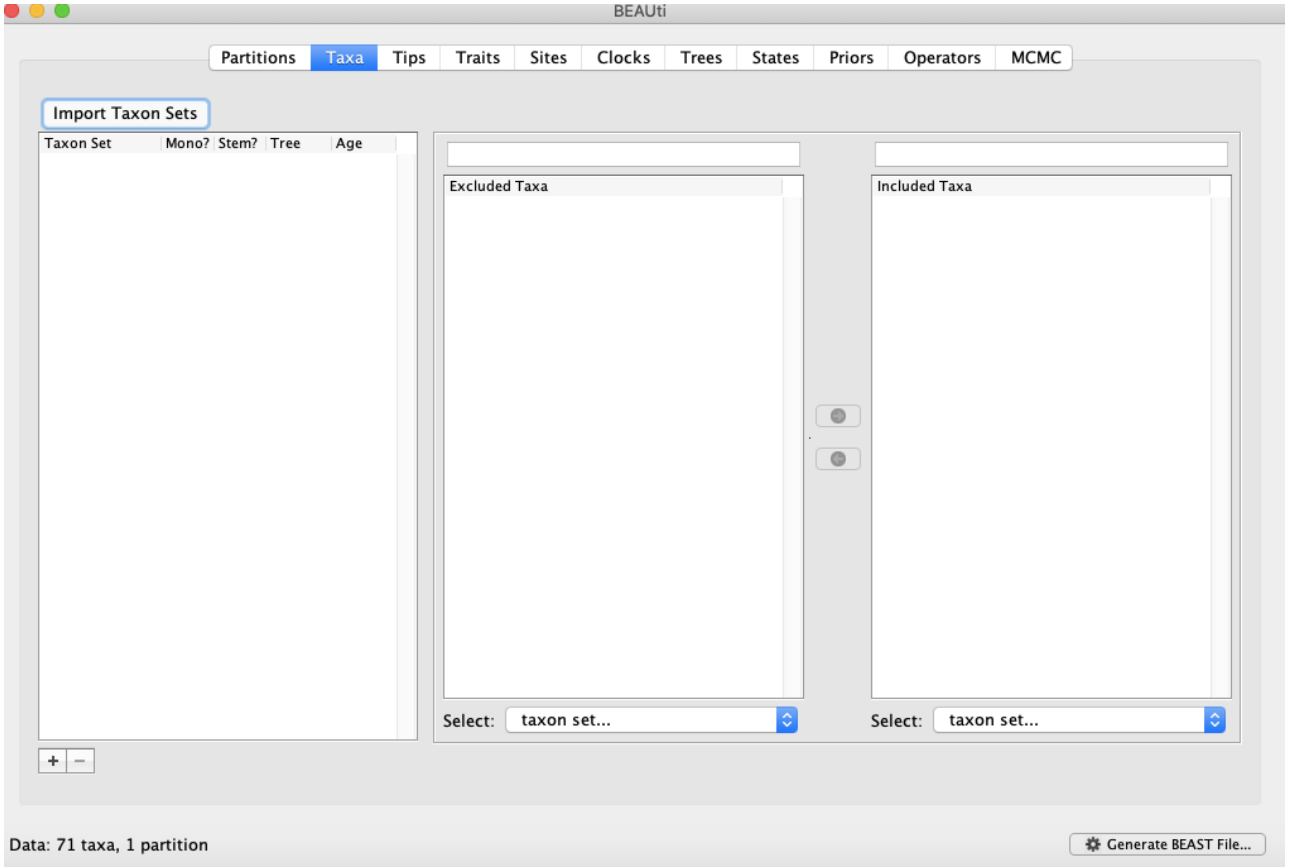
Allows manipulation of priors, operators, and MCMC chain settings

Create a BEAST XML file

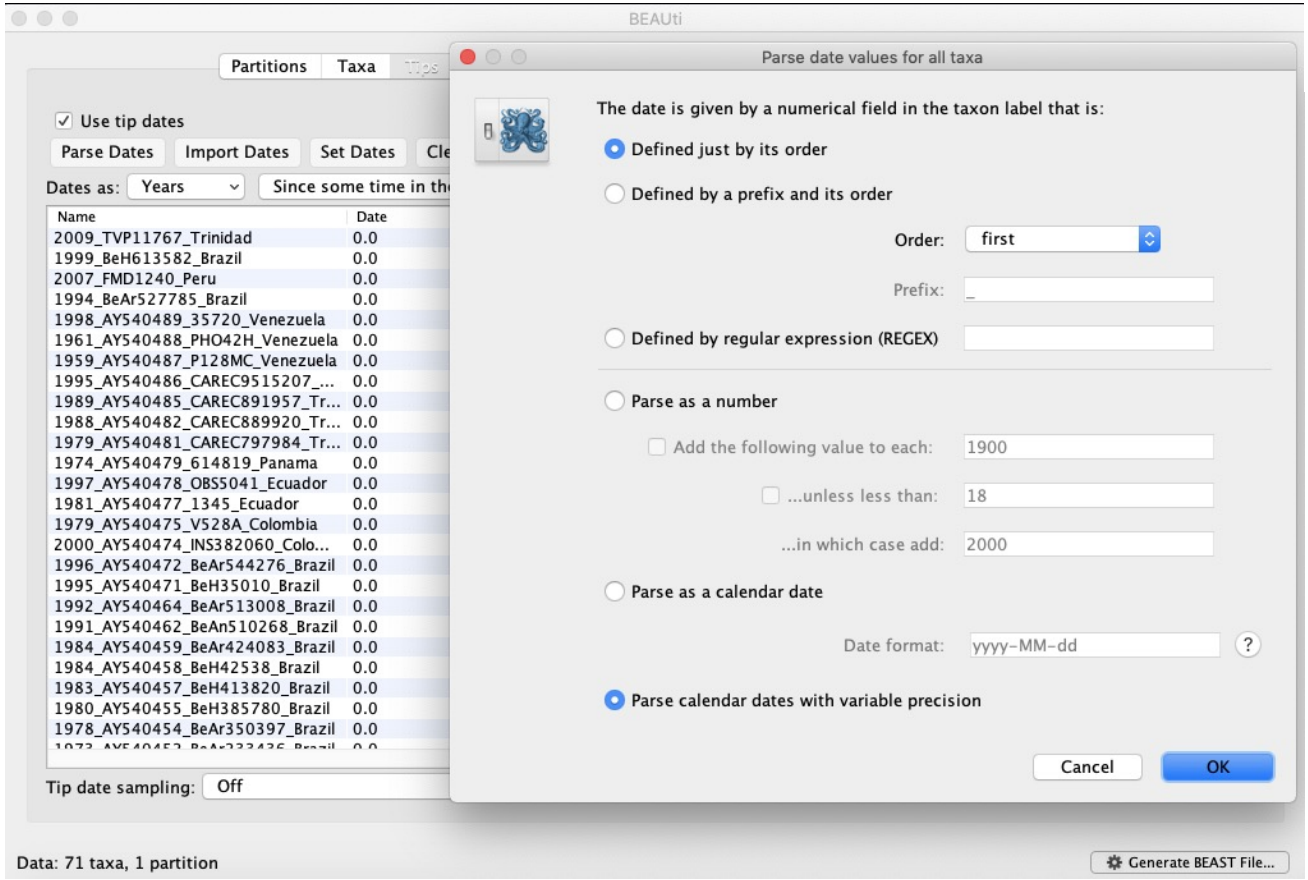
BEAUti


BEAUti
Partitions
Taxa
Tips
Traits
Sites
Clocks
Trees
States
Priors
Operators
MCMC
Unlink Subst. Models
Partition Name
File Name
YFV
YFV.nex
View Partition
Data: 71 taxa, 1 partition
Generate BEAST File...

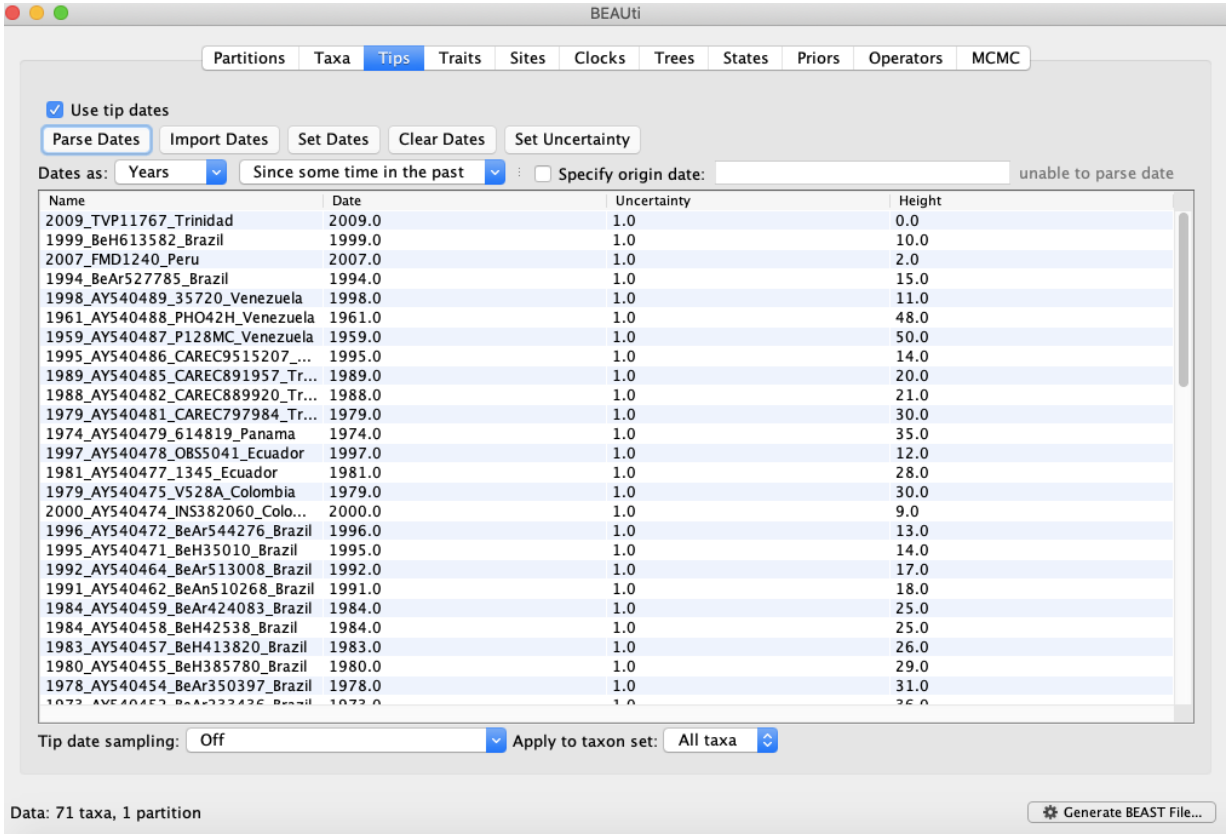
BEAUti


BEAUti
Partitions
Taxa
Tips
Traits
Sites
Clocks
Trees
States
Priors
Operators
MCMC
Import Taxon Sets
Taxon Set
Mono?
Stem?
Tree
Age
Excluded Taxa
Included Taxa
Select:
taxon set...
Select:
taxon set...
Data: 71 taxa, 1 partition
Generate BEAST File...

BEAUti

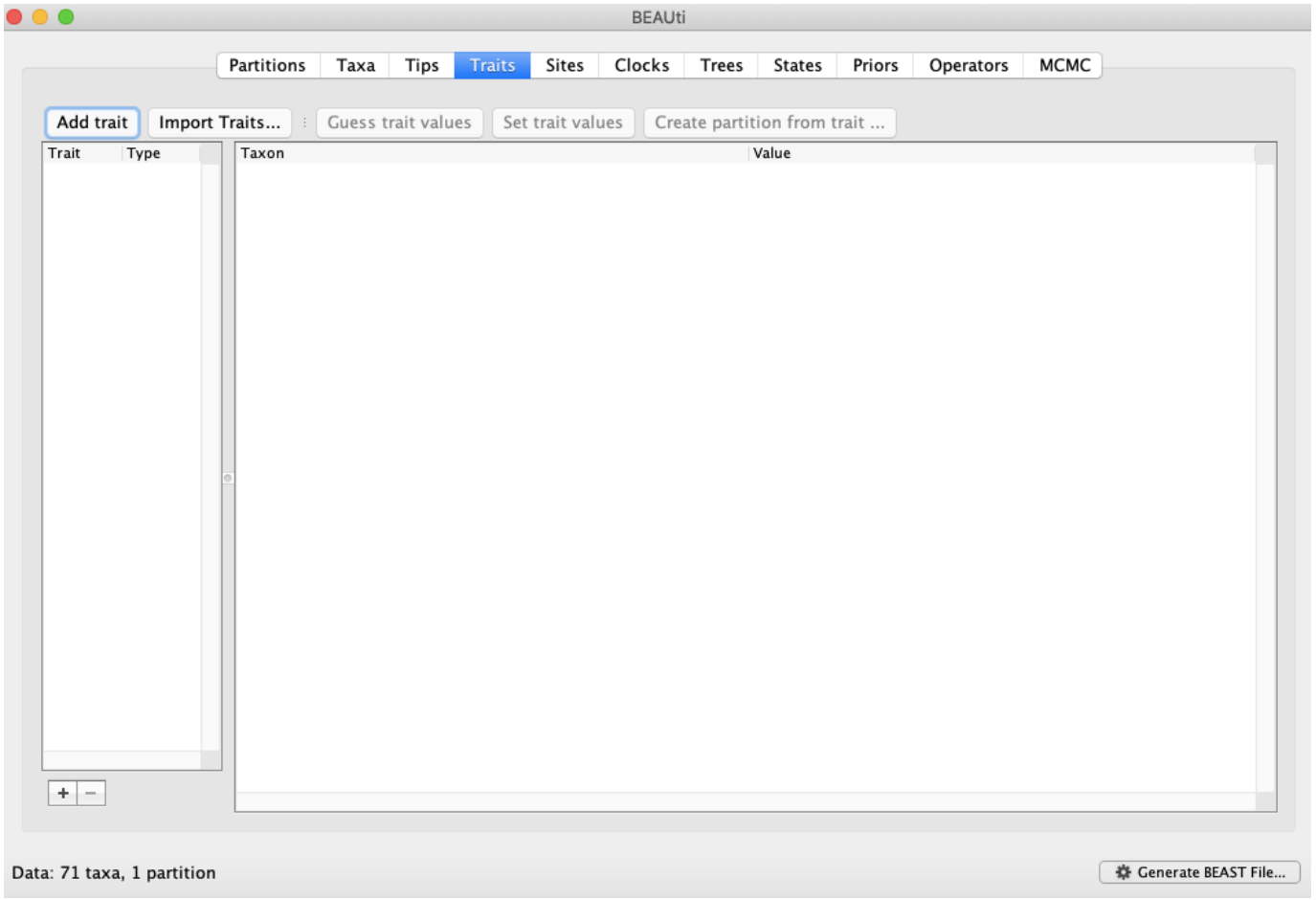

BEAUti
Partitions
Taxa
Tips
Use tip dates
Parse Dates
Import Dates
Set Dates
Dates as:
Years
Since some time in th
Name
Date
2009_TVP11767_Trinidad 0.0
1999_BeH613582_Brazil 0.0
2007_FMD1240_Peru 0.0
1994_BeAr527785_Brazil 0.0
1998_AY540489_35720_Venezuela 0.0
1961_AY540488_PHO42H_Venezuela 0.0
1959_AY540487_P128MC_Venezuela 0.0
1995_AY540486_CAREC9515207_... 0.0
1989_AY540485_CAREC891957_Tr... 0.0
1988_AY540482_CAREC889920_Tr... 0.0
1979_AY540481_CAREC797984_Tr... 0.0
1974_AY540479_614819_Panama 0.0
1997_AY540478_OBS5041_Ecuador 0.0
1981_AY540477_1345_Ecuador 0.0
1979_AY540475_V528A_Colombia 0.0
2000_AY540474_INS382060_Colo... 0.0
1996_AY540472_BeAr544276_Brazil 0.0
1995_AY540471_BeH35010_Brazil 0.0
1992_AY540464_BeAr513008_Brazil 0.0
1991_AY540462_BeAn510268_Brazil 0.0
1984_AY540459_BeAr424083_Brazil 0.0
1984_AY540458_BeH42538_Brazil 0.0
1983_AY540457_BeH413820_Brazil 0.0
1980_AY540455_BeH385780_Brazil 0.0
1978_AY540454_BeAr350397_Brazil 0.0
Tip date sampling:
Off
Parse date values for all taxa
The date is given by a numerical field in the taxon label that is:
Defined just by its order
Defined by a prefix and its order
Order:
first
Prefix:
_
Defined by regular expression (REGEX)
Parse as a number
Add the following value to each:
1900
...unless less than:
18
...in which case add:
2000
Parse as a calendar date
Date format:
yyyy-MM-dd
Parse calendar dates with variable precision
Cancel
OK
Data: 71 taxa, 1 partition
Generate BEAST File...

BEAUti



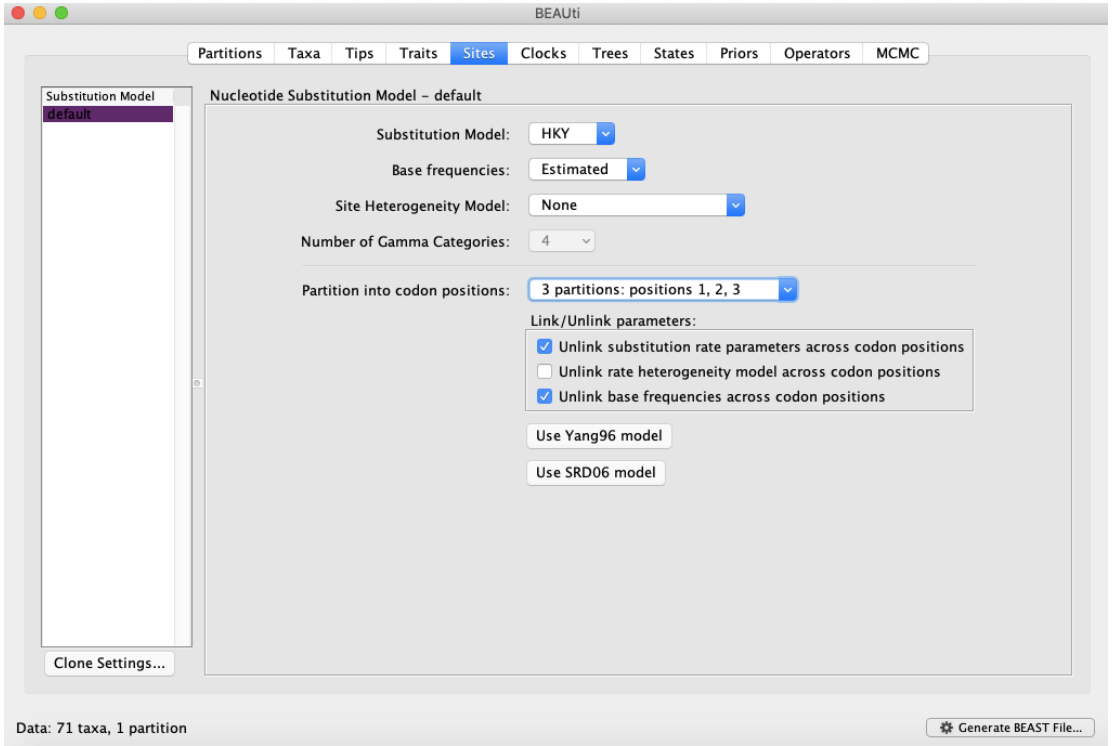
Name	Date	Uncertainty	Height
2009_TVP11767_Trinidad	2009.0	1.0	0.0
1999_BeH613582_Brazil	1999.0	1.0	10.0
2007_FMD1240_Peru	2007.0	1.0	2.0
1994_BeAr527785_Brazil	1994.0	1.0	15.0
1998_AY540489_35720_Venezuela	1998.0	1.0	11.0
1961_AY540488_PHO42H_Venezuela	1961.0	1.0	48.0
1959_AY540487_P128MC_Venezuela	1959.0	1.0	50.0
1995_AY540486_CAREC9515207_...	1995.0	1.0	14.0
1989_AY540485_CAREC891957_Tr...	1989.0	1.0	20.0
1988_AY540482_CAREC889920_Tr...	1988.0	1.0	21.0
1979_AY540481_CAREC797984_Tr...	1979.0	1.0	30.0
1974_AY540479_614819_Panama	1974.0	1.0	35.0
1997_AY540478_OBS5041_Ecuador	1997.0	1.0	12.0
1981_AY540477_1345_Ecuador	1981.0	1.0	28.0
1979_AY540475_V528A_Colombia	1979.0	1.0	30.0
2000_AY540474_INS382060_Colo...	2000.0	1.0	9.0
1996_AY540472_BeAr544276_Brazil	1996.0	1.0	13.0
1995_AY540471_BeH35010_Brazil	1995.0	1.0	14.0
1992_AY540464_BeAr513008_Brazil	1992.0	1.0	17.0
1991_AY540462_BeAn510268_Brazil	1991.0	1.0	18.0
1984_AY540459_BeAr424083_Brazil	1984.0	1.0	25.0
1984_AY540458_BeH42538_Brazil	1984.0	1.0	25.0
1983_AY540457_BeH413820_Brazil	1983.0	1.0	26.0
1980_AY540455_BeH385780_Brazil	1980.0	1.0	29.0
1978_AY540454_BeAr350397_Brazil	1978.0	1.0	31.0

BEAUti

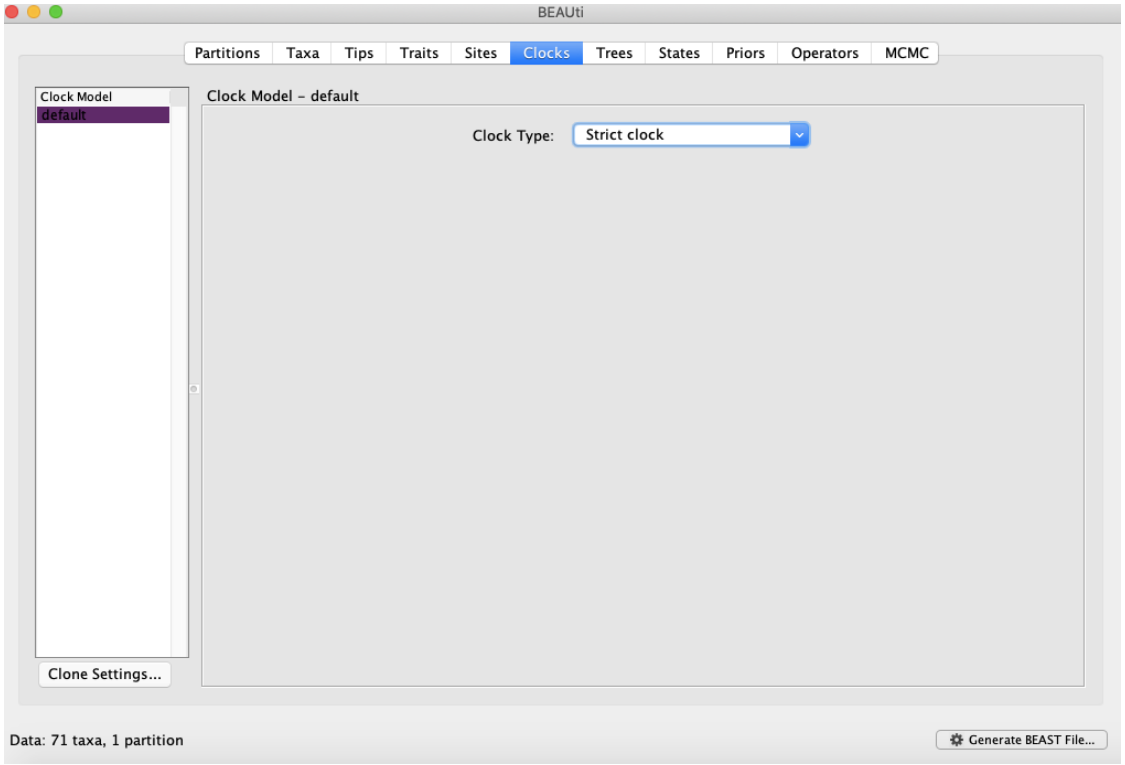

BEAUti
Partitions
Taxa
Tips
Traits
Sites
Clocks
Trees
States
Priors
Operators
MCMC
Add trait
Import Traits...
Guess trait values
Set trait values
Create partition from trait ...
Trait
Type
Taxon
Value
Data: 71 taxa, 1 partition
Generate BEAST File...

BEAUti

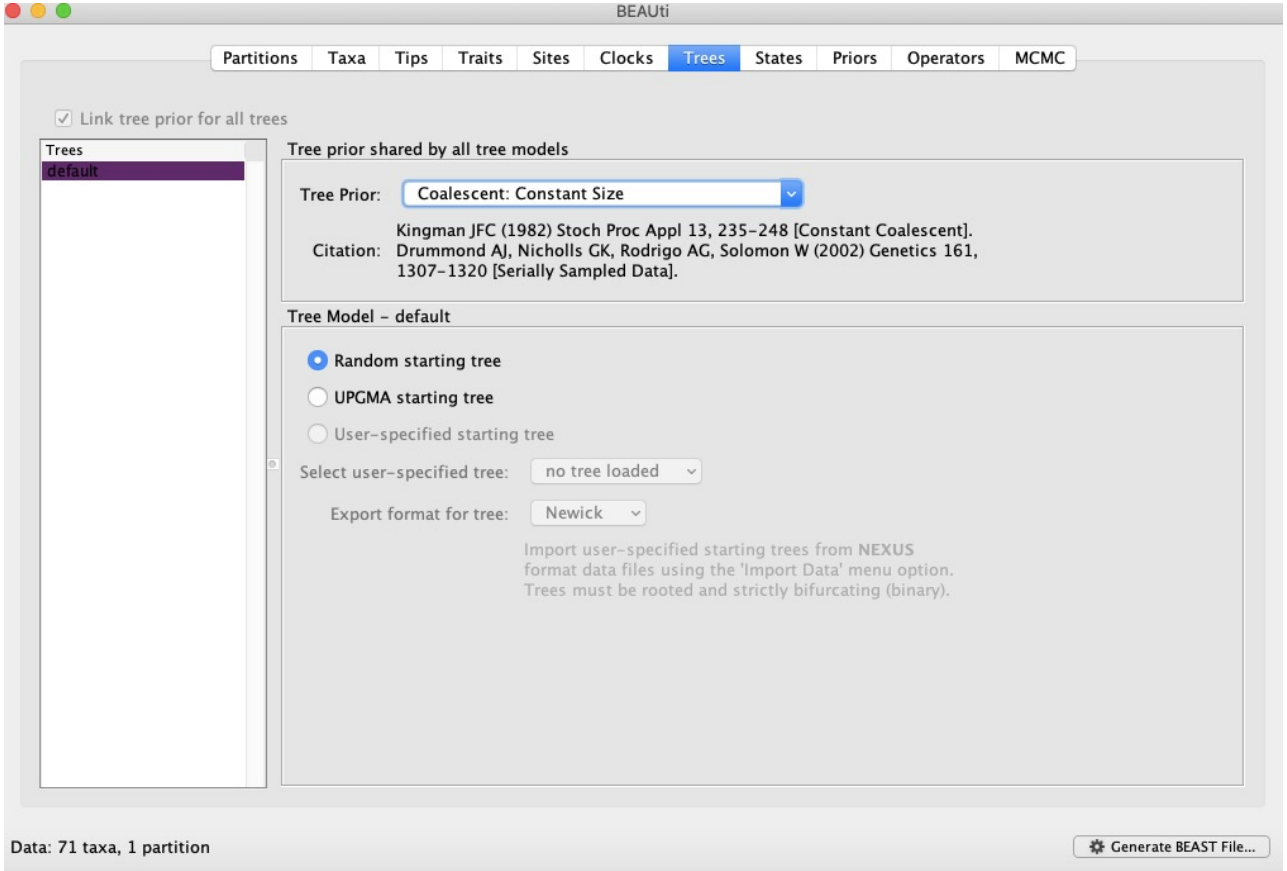
Selecting the Partition into codon positions option assumes that the data are aligned as codons. This option will then estimate a separate rate of substitution for each codon position, or for 1+2 versus 3, depending on the setting.



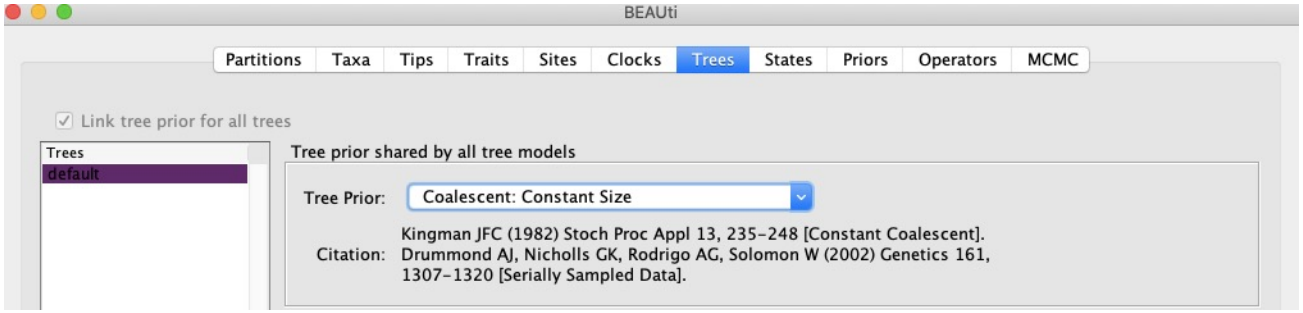
BEAUti


BEAUti
Partitions
Taxa
Tips
Traits
Sites
Clocks
Trees
States
Priors
Operators
MCMC
Clock Model
default
Clock Model – default
Clock Type:
Strict clock
Clone Settings...
Data: 71 taxa, 1 partition
Generate BEAST File...

BEAUti


BEAUti
Partitions
Taxa
Tips
Traits
Sites
Clocks
Trees
States
Priors
Operators
MCMC
Link tree prior for all trees
Trees
default
Tree prior shared by all tree models
Tree Prior:
Coalescent: Constant Size
Citation:
Kingman JFC (1982) Stoch Proc Appl 13, 235-248 [Constant Coalescent].
Drummond AJ, Nicholls GK, Rodrigo AG, Solomon W (2002) Genetics 161, 1307-1320 [Serially Sampled Data].
Tree Model - default
Random starting tree
UPGMA starting tree
User-specified starting tree
Select user-specified tree:
no tree loaded
Export format for tree:
Newick
Import user-specified starting trees from NEXUS format data files using the 'Import Data' menu option. Trees must be rooted and strictly bifurcating (binary).
Data: 71 taxa, 1 partition
Generate BEAST File...

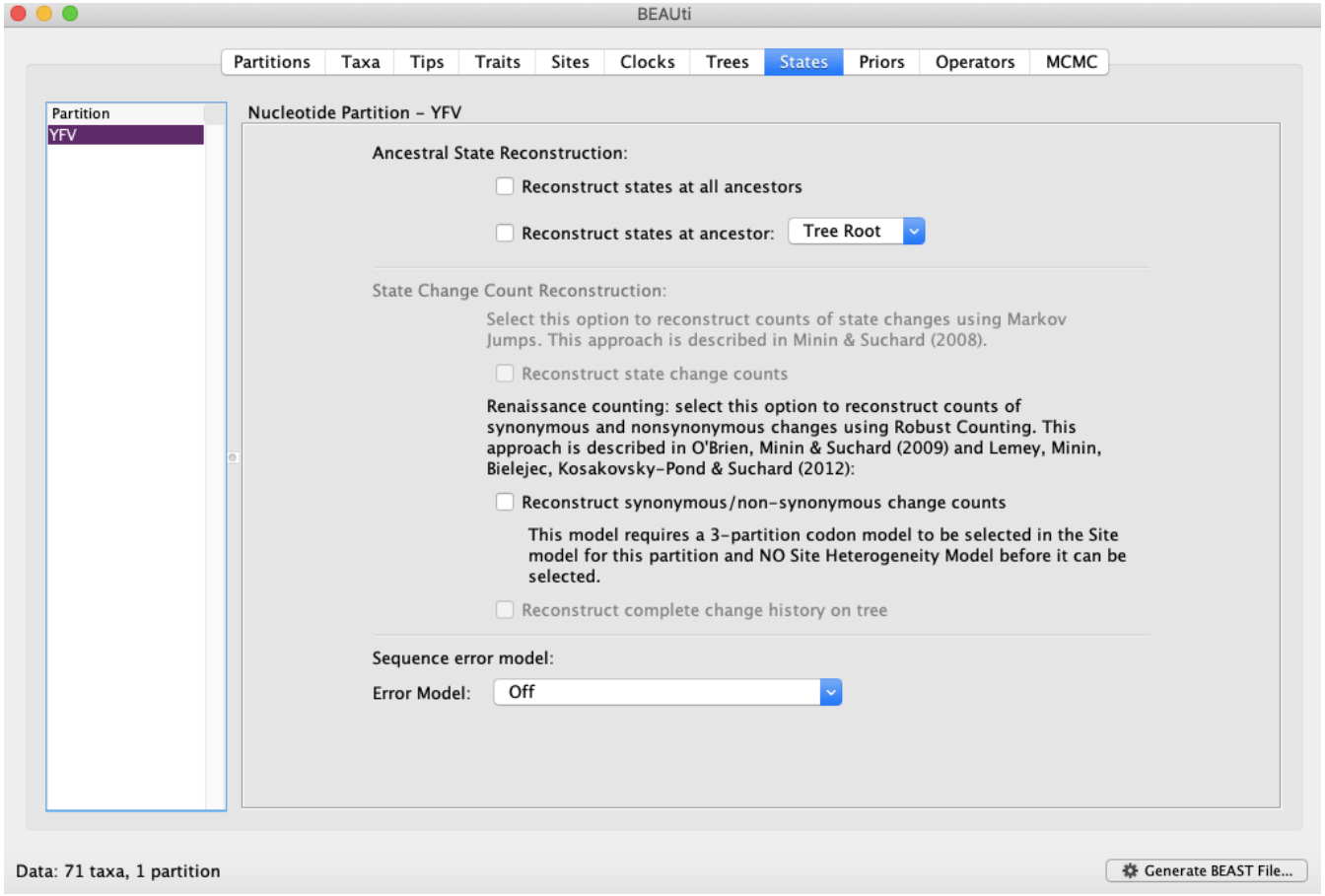
BEAUti



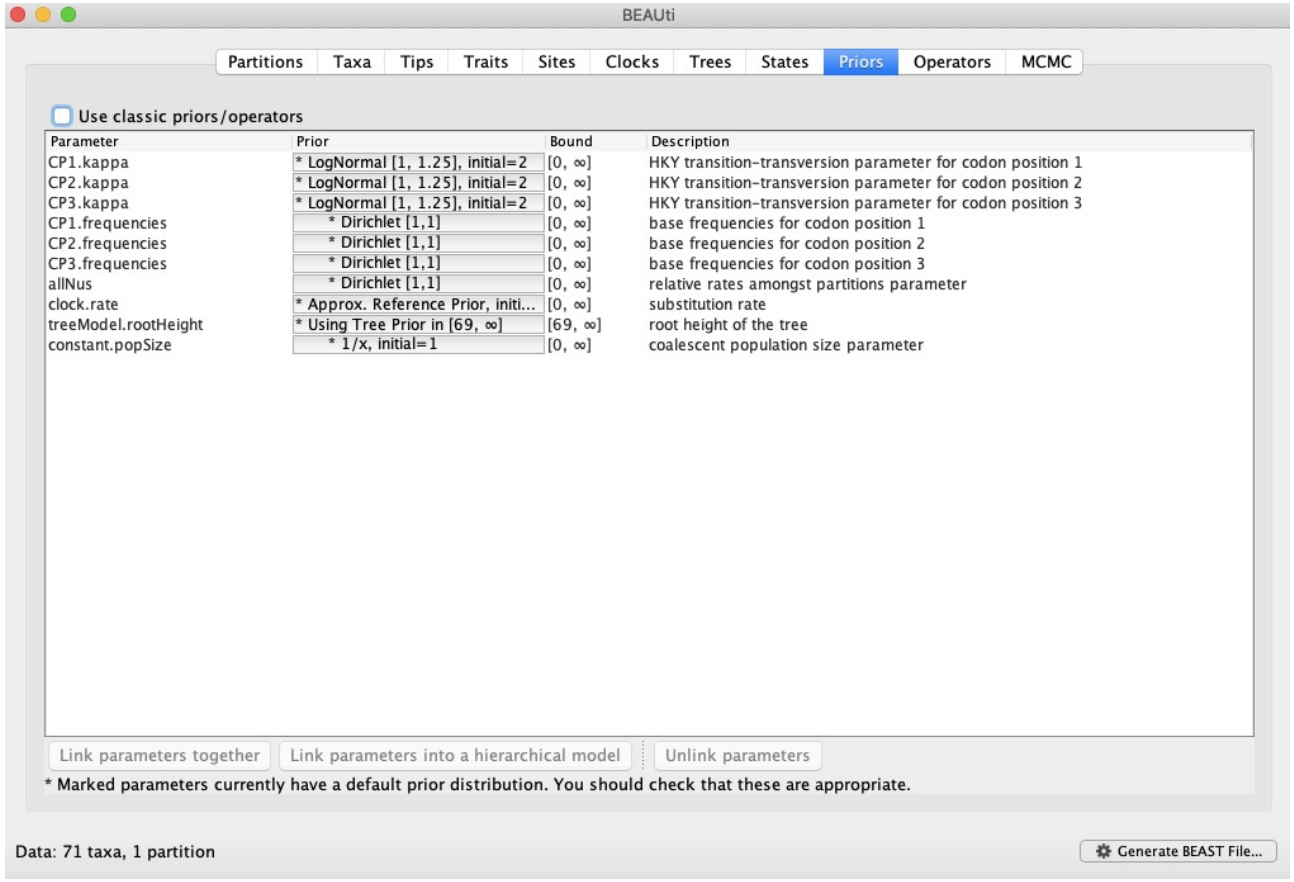
It offers a range of prior distributions to model population size changes through time (i.e., demography).

These are **coalescent** priors, where the effective population size \(N_e\) varies through time according to a certain function \(N_e(t) \).

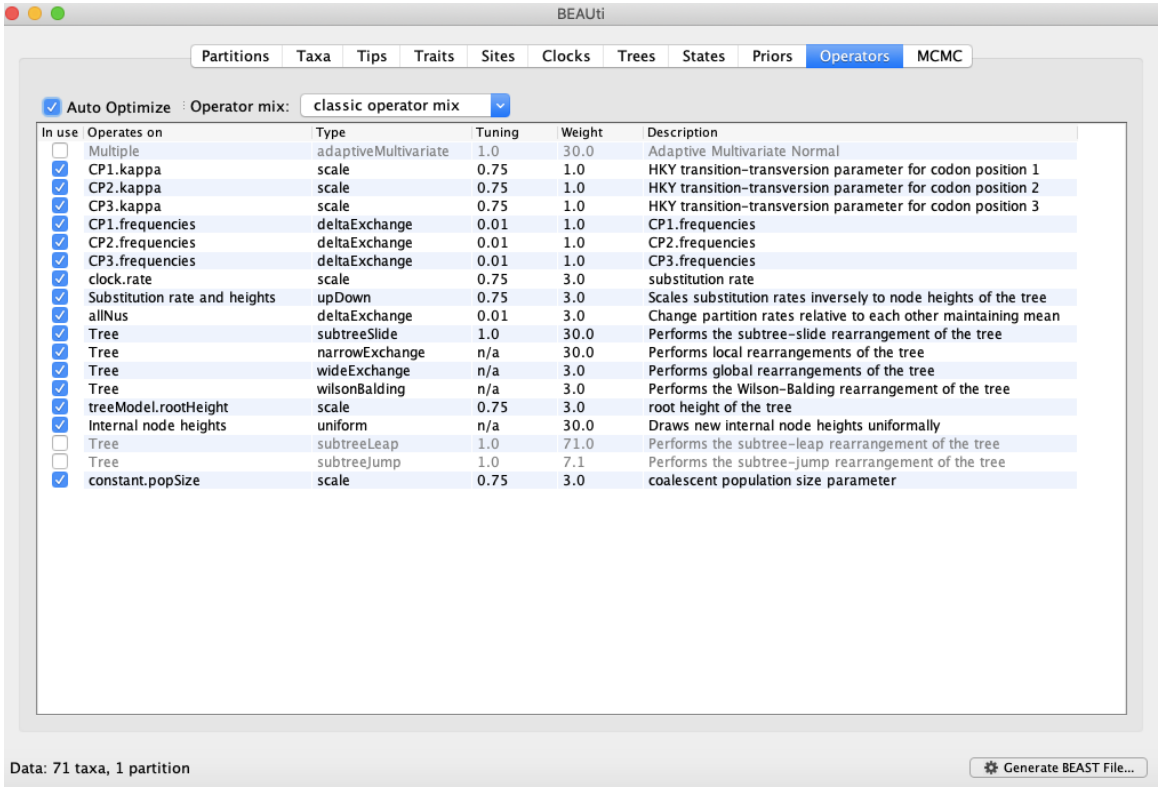
BEAUti


BEAUti
Partitions
Taxa
Tips
Traits
Sites
Clocks
Trees
States
Priors
Operators
MCMC
Partition
YFV
Nucleotide Partition – YFV
Ancestral State Reconstruction:
Reconstruct states at all ancestors
Reconstruct states at ancestor:
Tree Root
State Change Count Reconstruction:
Select this option to reconstruct counts of state changes using Markov Jumps. This approach is described in Minin & Suchard (2008).
Reconstruct state change counts
Renaissance counting: select this option to reconstruct counts of synonymous and nonsynonymous changes using Robust Counting. This approach is described in O'Brien, Minin & Suchard (2009) and Lemey, Minin, Bielejec, Kosakovsky-Pond & Suchard (2012):
Reconstruct synonymous/non-synonymous change counts
This model requires a 3-partition codon model to be selected in the Site model for this partition and NO Site Heterogeneity Model before it can be selected.
Reconstruct complete change history on tree
Sequence error model:
Error Model:
Off
Data: 71 taxa, 1 partition
Generate BEAST File...

BEAUti

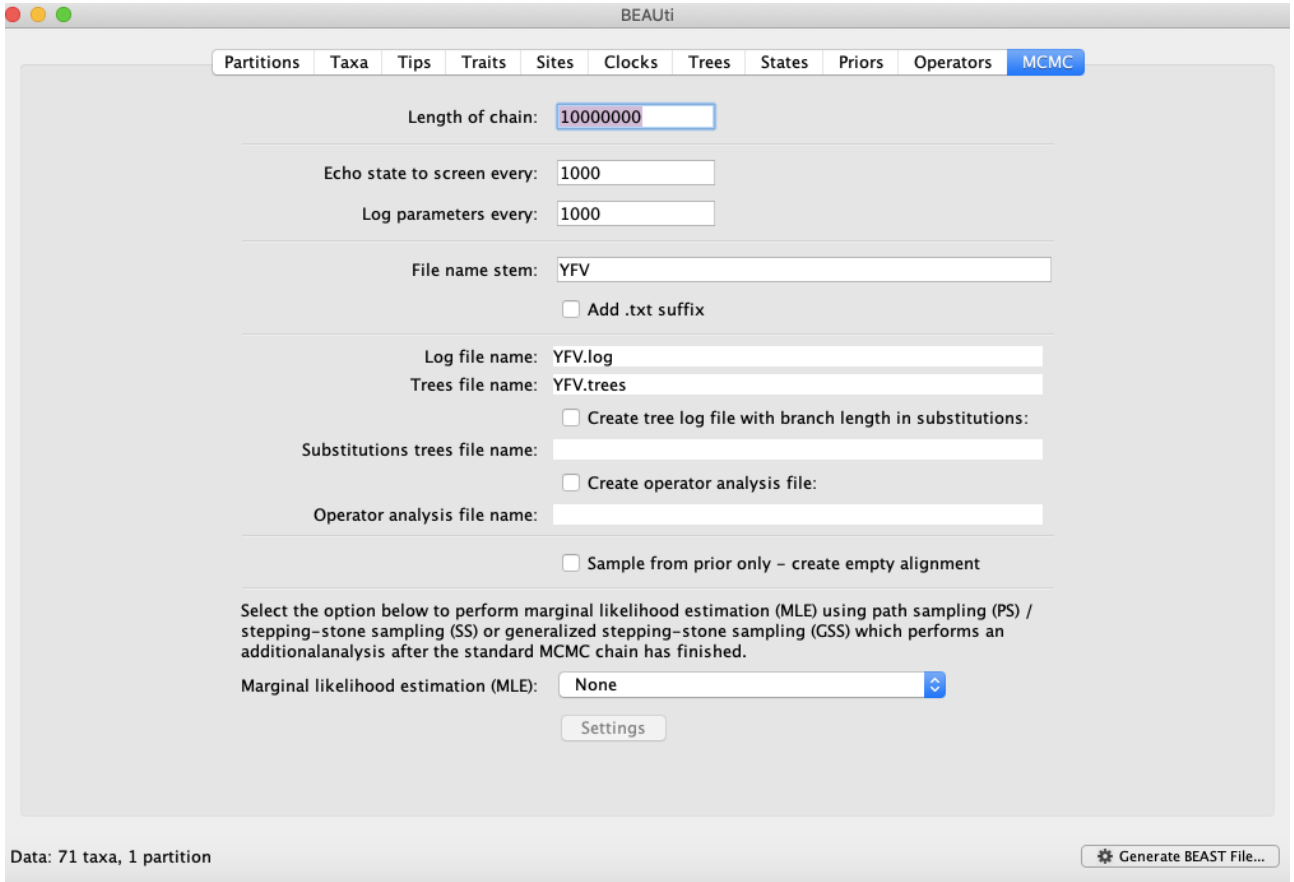

BEAUti
Partitions
Taxa
Tips
Traits
Sites
Clocks
Trees
States
Priors
Operators
MCMC
Use classic priors/operators
Parameter
Prior
Bound
Description
CP1.kappa
* LogNormal [1, 1.25], initial=2
[0, ∞]
HKY transition-transversion parameter for codon position 1
CP2.kappa
* LogNormal [1, 1.25], initial=2
[0, ∞]
HKY transition-transversion parameter for codon position 2
CP3.kappa
* LogNormal [1, 1.25], initial=2
[0, ∞]
HKY transition-transversion parameter for codon position 3
CP1.frequencies
* Dirichlet [1,1]
[0, ∞]
base frequencies for codon position 1
CP2.frequencies
* Dirichlet [1,1]
[0, ∞]
base frequencies for codon position 2
CP3.frequencies
* Dirichlet [1,1]
[0, ∞]
base frequencies for codon position 3
allNus
* Dirichlet [1,1]
[0, ∞]
relative rates amongst partitions parameter
clock.rate
* Approx. Reference Prior, initi...
[0, ∞]
substitution rate
treeModel.rootHeight
* Using Tree Prior in [69, ∞]
[69, ∞]
root height of the tree
constant.popSize
* 1/x, initial=1
[0, ∞]
coalescent population size parameter
Link parameters together
Link parameters into a hierarchical model
Unlink parameters
* Marked parameters currently have a default prior distribution. You should check that these are appropriate.
Data: 71 taxa, 1 partition
Generate BEAST File...

BEAUti

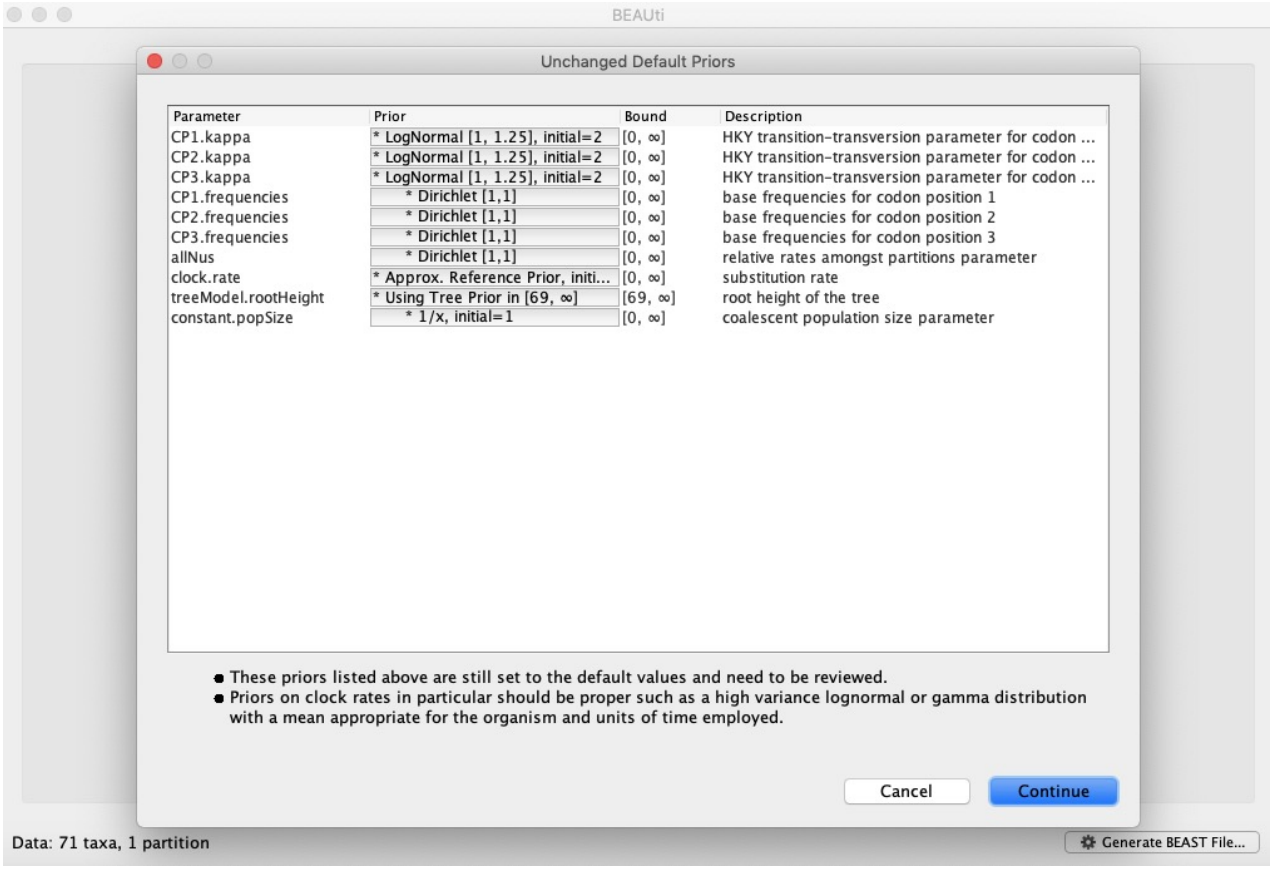


In use	Operates on	Type	Tuning	Weight	Description
☐	Multiple	adaptiveMultivariate	1.0	30.0	Adaptive Multivariate Normal
☑	CP1.kappa	scale	0.75	1.0	HKY transition-transversion parameter for codon position 1
☑	CP2.kappa	scale	0.75	1.0	HKY transition-transversion parameter for codon position 2
☑	CP3.kappa	scale	0.75	1.0	HKY transition-transversion parameter for codon position 3
☑	CP1.frequencies	deltaExchange	0.01	1.0	CP1.frequencies
☑	CP2.frequencies	deltaExchange	0.01	1.0	CP2.frequencies
☑	CP3.frequencies	deltaExchange	0.01	1.0	CP3.frequencies
☑	clock.rate	scale	0.75	3.0	substitution rate
☑	Substitution rate and heights	upDown	0.75	3.0	Scales substitution rates inversely to node heights of the tree
☑	allNus	deltaExchange	0.01	3.0	Change partition rates relative to each other maintaining mean
☑	Tree	subtreeSlide	1.0	30.0	Performs the subtree-slide rearrangement of the tree
☑	Tree	narrowExchange	n/a	30.0	Performs local rearrangements of the tree
☑	Tree	wideExchange	n/a	3.0	Performs global rearrangements of the tree
☑	Tree	wilsonBalding	n/a	3.0	Performs the Wilson-Balding rearrangement of the tree
☑	treeModel.rootHeight	scale	0.75	3.0	root height of the tree
☑	Internal node heights	uniform	n/a	30.0	Draws new internal node heights uniformally
☐	Tree	subtreeLeap	1.0	71.0	Performs the subtree-leap rearrangement of the tree
☐	Tree	subtreeJump	1.0	7.1	Performs the subtree-jump rearrangement of the tree
☑	constant.popSize	scale	0.75	3.0	coalescent population size parameter

BEAUti

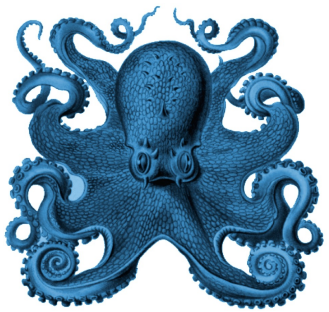

BEAUti
Partitions
Taxa
Tips
Traits
Sites
Clocks
Trees
States
Priors
Operators
MCMC
Length of chain:
10000000
Echo state to screen every:
1000
Log parameters every:
1000
File name stem:
YFV
Add .txt suffix
Log file name:
YFV.log
Trees file name:
YFV.trees
Create tree log file with branch length in substitutions:
Substitutions trees file name:
Create operator analysis file:
Operator analysis file name:
Sample from prior only - create empty alignment
Select the option below to perform marginal likelihood estimation (MLE) using path sampling (PS) / stepping-stone sampling (SS) or generalized stepping-stone sampling (GSS) which performs an additionalanalysis after the standard MCMC chain has finished.
Marginal likelihood estimation (MLE):
None
Settings
Data: 71 taxa, 1 partition
Generate BEAST File...

BEAUti

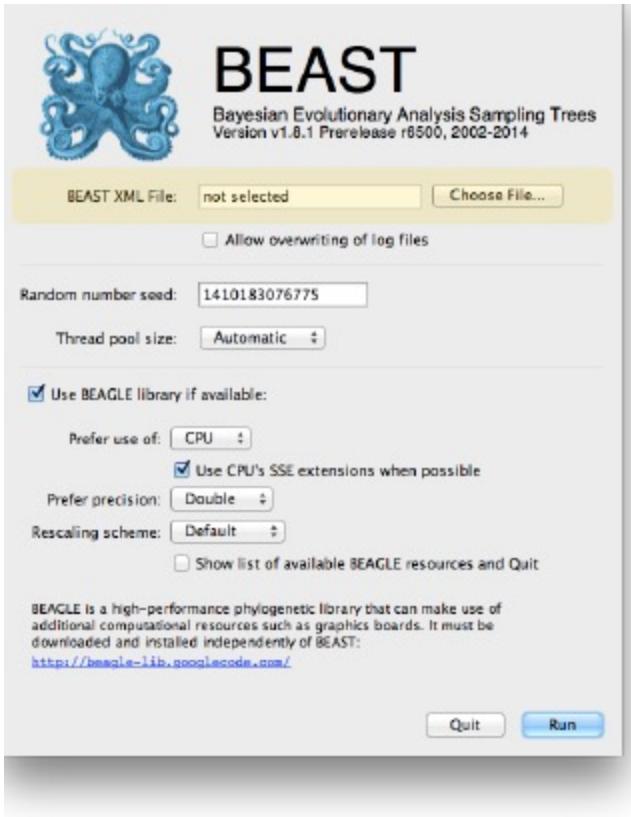

BEAUti
Unchanged Default Priors
Parameter | Prior | Bound | Description
CP1.kappa | * LogNormal [1, 1.25], initial=2 | [0, ∞] | HKY transition-transversion parameter for codon ...
CP2.kappa | * LogNormal [1, 1.25], initial=2 | [0, ∞] | HKY transition-transversion parameter for codon ...
CP3.kappa | * LogNormal [1, 1.25], initial=2 | [0, ∞] | HKY transition-transversion parameter for codon ...
CP1.frequencies | * Dirichlet [1,1] | [0, ∞] | base frequencies for codon position 1
CP2.frequencies | * Dirichlet [1,1] | [0, ∞] | base frequencies for codon position 2
CP3.frequencies | * Dirichlet [1,1] | [0, ∞] | base frequencies for codon position 3
allNus | * Dirichlet [1,1] | [0, ∞] | relative rates amongst partitions parameter
clock.rate | * Approx. Reference Prior, initi... | [0, ∞] | substitution rate
treeModel.rootHeight | * Using Tree Prior in [69, ∞] | [69, ∞] | root height of the tree
constant.popSize | * 1/x, initial=1 | [0, ∞] | coalescent population size parameter
These priors listed above are still set to the default values and need to be reviewed.
Priors on clock rates in particular should be proper such as a high variance lognormal or gamma distribution with a mean appropriate for the organism and units of time employed.
Cancel
Continue
Data: 71 taxa, 1 partition
Generate BEAST File...

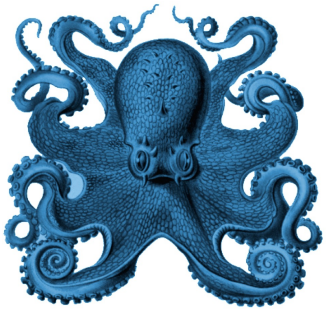
Running BEAST...

BEAST



Reads instructions from the xml file
Performs the analysis
go for a coffee while this happens
or lunch
or a vacation
Produces two files (with names specified in BEAUti):
[analysisResults].log
[analysisResults].trees





Input 'XML' file

Random number seed

BEAGLE settings

Zaire_chiro_partitioned.xml

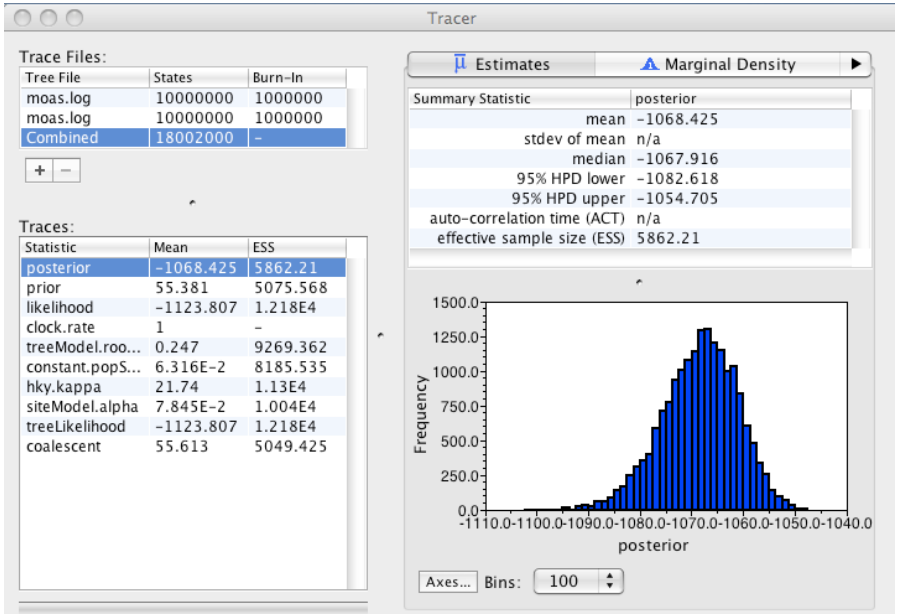
```
  Branch rate model used: strictClockBranchRates
  Using BEAGLE resource 0: CPU
    with instance flags:  PRECISION_DOUBLE COMPUTATION_SYNCH EIGEN_REAL SCALING_MANUAL SCALERS_RAW VECTOR_SSE THREADING_NONE PROCE
  Ignoring ambiguities in tree likelihood.
  With 225 unique site patterns.
  Using rescaling scheme : dynamic (rescaling every 100 evaluations)
Using BEAGLE TreeLikelihood
  Branch rate model used: strictClockBranchRates
  Using BEAGLE resource 0: CPU
    with instance flags:  PRECISION_DOUBLE COMPUTATION_SYNCH EIGEN_REAL SCALING_MANUAL SCALERS_RAW VECTOR_SSE THREADING_NONE PROCESSOR_CPU F
  Ignoring ambiguities in tree likelihood.
  With 350 unique site patterns.
  Using rescaling scheme : dynamic (rescaling every 100 evaluations)
Using BEAGLE TreeLikelihood
  Branch rate model used: strictClockBranchRates
  Using BEAGLE resource 0: CPU
    with instance flags:  PRECISION_DOUBLE COMPUTATION_SYNCH EIGEN_REAL SCALING_MANUAL SCALERS_RAW VECTOR_SSE THREADING_NONE PROCESSOR_CPU F
  Ignoring ambiguities in tree likelihood.
  With 33 unique site patterns.
  Using rescaling scheme : dynamic (rescaling every 100 evaluations)
Using BEAGLE TreeLikelihood
  Branch rate model used: strictClockBranchRates
  Using BEAGLE resource 0: CPU
    with instance flags:  PRECISION_DOUBLE COMPUTATION_SYNCH EIGEN_REAL SCALING_MANUAL SCALERS_RAW VECTOR_SSE THREADING_NONE PROCESSOR_CPU F
  Ignoring ambiguities in tree likelihood.
  With 343 unique site patterns.
  Using rescaling scheme : dynamic (rescaling every 100 evaluations)
Parameter weights for delta exchange are: 9486  4743    267     4465
Optimisation Schedule: log

 ---------------------------------
Creating ctmcScalePrior model.
        If you publish results using this prior, please reference:
                1. Ferreira and Suchard (2008) for the conditional reference prior on CTMC scale parameter prior;

 ---------------------------------
Creating ctmcScalePrior model.
        If you publish results using this prior, please reference:
                1. Ferreira and Suchard (2008) for the conditional reference prior on CTMC scale parameter prior;

 ---------------------------------
Creating ctmcScalePrior model.
        If you publish results using this prior, please reference:
                1. Ferreira and Suchard (2008) for the conditional reference prior on CTMC scale parameter prior;
Likelihood computation is using an auto sizing thread pool.
Creating the MCMC chain:
  chainLength=20000000
  autoOptimize=true
  autoOptimize delayed for 200000 steps
# BEAST v1.8.1 Prerelease r6500, r6500
# Generated Tue Sep 09 10:34:08 CEST 2014 [seed=1410183076775]
state   Posterior       Prior           Likelihood      rootHeight      cds.clock.rate  diagnostic.clock.rate   intergenic.clock.rate
0       -91363.8582     -5135.5341      -86228.3241     339.359         7E-4            7E-4            7E-4            -
2000    -62544.2432     -995.9626       -61548.2806     397.856         7.81006E-5      2.20371E-4      1.73146E-4      -
4000    -53809.6682     -952.8936       -52856.7746     351.161         5.14098E-5      1.84409E-4      1.13455E-4      -
6000    -48151.6389     -922.2231       -47229.4159     300.362         4.03324E-5      2.34534E-4      8.21841E-5      -
8000    -43713.2833     -869.4692       -42843.8141     536.732         3.09051E-5      5.18719E-4      5.53193E-5      -
10000   -40858.4087     -876.1709       -39982.2377     811.562         2.10911E-5      2.99179E-4      3.66545E-5      -
```

Summary of results and diagnostics

LogCombiner

Combines .log and .tree files from multiple (identical) runs

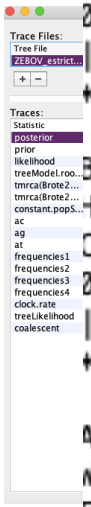
Also possible to combine log files in Tracer

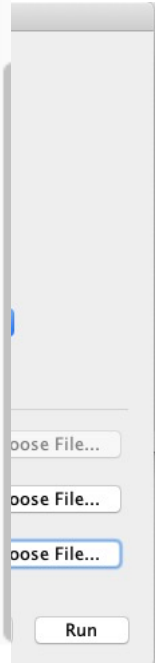


tre

Sumr
Make

```
0                   25                  50                  75                  100
|-------------------|-------------------|-------------------|-------------------|
*****

Total trees read: 1001
Ignoring first 1500000 states (75 trees).
Total unique clades: 10012

Finding maximum credibility tree...
Analyzing 926 trees...
0                   25                  50                  75                  100
|-------------------|-------------------|-------------------|-------------------|
*************************************************************************************

Best tree: STATE_17620000 (tree number 882)
Highest Log Clade Credibility: -116.97806096082365
Collecting node information...
0                   25                  50                  75                  100
|-------------------|-------------------|-------------------|-------------------|
*************************************************************************************

Annotating target tree...
Writing annotated tree....
Finished - Quit program to exit.
```





Summary of results and diagnostics

1. Definition of phylodynamics
2. Source phylodynamic variation
3. Phylodynamic techniques
4. Estimating Temporal Signal
5. **Real-time phylodynamics**

Towards real-time phylodynamics

the time is now

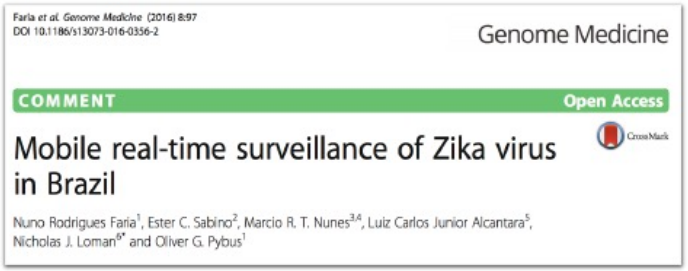
Gardy et al. Genome Biology (2015) 16:155
DOI 10.1186/s13059-015-0726-x

Genome Biology

COMMENT Open Access

Real-time digital pathogen surveillance — the time is now

Jennifer Gardy[1], Nicholas J. Loman[2*] and Andrew Rambaut[3]



Faria et al. Genome Medicine (2016) 8:97
DOI 10.1186/s13073-016-0356-2

Genome Medicine

COMMENT Open Access

Mobile real-time surveillance of Zika virus in Brazil

Nuno Rodrigues Faria[1], Ester C. Sabino[2], Marcio R. T. Nunes[3,4], Luiz Carlos Junior Alcantara[5], Nicholas J. Loman[6*] and Oliver G. Pybus[1]

Democratization of genome sequencing

Genome data can now be generated in real-time at a cost of <50$ per virus genome (Quick et al. Nature Protocols 2017)

Real-time virus phylodynamics is the future of virus epidemiology

Integration of genetic, epidemiological and ecological information provides critical information about the past, present and the future of viral epidemics

Let’s practice!