MAFFT

For fasta format: mafft --6merpair --addfragments sars-cov-2.fas reference_sars_cov_2.fas > out_aln.fas

For clustal format mafft --clustalout --6merpair --addfragments sars-cov-2.fas reference_sars_cov_2.fas >

For identifying number of conserved nucleotides: grep -o "*" out_aln | wc -l

IQTREE

iqtree -s sars-cov-2_aln.fas -B 1000

TempEst

Eile Edit Help				-		
Best-fitting root Sample Datest Tree Root-to-tin Residuals Node density						
Function: heuristic residual mean squared v Shingke back in the Note being Residuals in due being						
More sect tip 1 2997E 2	Importunates Parse Dates Clear Dates 10	aces specified as Tears V pince some time in	rule pasc 🔹			
Coefficient of variati 0.3355	Name	Date	Precision	Height		
Stdev 43238F-4	1ins 25617c-25619c EPI ISL 11017424 Jamaica 2	-	-	0.0		
Variance 1.8695E-7	EPI ISL 13392375 Ghana 2021-08-10	-	-	0.0		
	EPI_ISL_13439948_Hong_Kong_2021-10-07	-	-	0.0		
	EPI_ISL_8187166_Colombia_2021-02-19	-	-	0.0		
	EPI ISL 8564574 Australia 2021-12-31	-	-	0.0		
	EPI_ISL_402124_2019-12-30	-	-	0.0		
	EPI_ISL_12644861_South_Africa_2021-06-14		-	0.0		
	EPI_ISL_6660906_Canada_2021-10-15	-	-	0.0		
	EPI_ISL_6705120_Argentina_2021-10-18		-	0.0		
	EPI_ISL_7225062_India_2021-05-11		•	0.0		
	EPI_ISL_8217102_Germany_2021-11-10		*	0.0		
	1ins_29890a-29901a_EPI_ISL_1111130_Peru_2021	* · · · · · · · · · · · · · · · · · · ·	*	0.0		
	EPI_ISL_1494722_Australia_2021-04-03		*	0.0		
	1ins_29734a-29746a_EPI_ISL_2891522_USA_2021	* · · · · · · · · · · · · · · · · · · ·	-	0.0		
	EPI_ISL_3824488_Switzerland_2021-08-16	-	-	0.0		
	EPI_ISL_4543939_Sint_Maarten_2021-09-06	-	-	0.0		
	EPI_ISL_10318197_Denmark_2022-02-15	-	-	0.0		
	EPI_ISL_11528229_Switzerland_2022-03-18	-	-	0.0		
	EPI_ISL_11680139_Bahamas_2021-12-15	-	-	0.0		
	EPI_ISL_7834572_Botswana_2021-12-01	-	-	0.0		
	1ins_22155g-22163a_EPI_ISL_9419101_India_2022	-	-	0.0		
	EPI_ISL_11449404_France_2021-07-17	÷	·	0.0		
	EPI_ISL_13025973_Malaysia_2021-05-26	•	·	0.0		
	EPI_ISL_2811956_New_Zealand_2021-06-15	*	*	0.0		
	EPI_ISL_5802461_Brazil_2021-03-07	*	*	0.0		
	EPI_ISL_12388429_Canada_2021-05-27	•	*	0.0		
	1ins_21931t-21933c_EPI_ISL_1962952_Netherland	-	-	0.0		
	1ins_21989t-21991c_EPI_ISL_4458156_Puerto_Ric	-	•	0.0		
	1ins_21989t-21991c_EPI_ISL_8482926_Mongolia_2	*	-	0.0		
	1ins_21937t-21939c_EPI_ISL_4984908_Colombia_2	*	-	0.0		
	1ins_21967t-21972c_EPI_ISL_8178768_Morocco_2	-	-	0.0		
	1ins_28263a-28266a_EPI_ISL_1197002_Italy_2021	-	-	0.0		
	1ins_28200a-28203a_EPI ISL_2655/13_Canada_20	-	-	0.0		
	1IDS_282158-282188_EPI_ISL_4080896_Brazil_2021	-	-	0.0		
	1INS_28224a-28227a_EPI_ISL_792683_Japan_2021	-	-	0.0		
	EPI_IDL_2493026_Angola_2021-03-22	-	-	0.0		
				~		
<				}		
Tree loaded, 36 taxa, contemporaneous tips						

Select "Parse Dates" tab

For example, identifiers are in following format: where "_" is used as a seperator and last field is date of isolation EPI_ISL_402124_2019-12-30

	Data				
ISL 1	Extract dates from taxon labels				
a_2021					
	The date is given by a numerical field in the taxon label that is:				
ia_202	Defined iust by its order				
-30 Africa					
a 2021	O Defined by a prefix and its order				
ina_20 021-05	Order: last				
1y_202	Prefix:				
_ISL_1					
ISL_21	O Defined by regular expression (REGEX)				
tland_: aarten	O Parse as a number				
ark_20	Add the following with a to each 4000				
erland	Add the rollowing value to each: 1900				
nas_20 ana_20	unless less than: 16				
ISL_9	in which case add: 2000				
e_2021 sia_201					
ealand	O Parse as a calendar date				
2021-0					
Ja_202	Date format: yyyy-MM-dd ?				
ISL_84	Parse calendar dates with variable precision				
_ISL_49					
ISL 20	Cancel OK OK				
ISL 40					

Check the "**Best-fitting root**" option (top left) and view the graph in "**Root-to-tip**" tab to understand clock behaviour

The value of "Slope(Rate)" is to be noted and used for BEAST run



BeauTi From File menu, load the alignment file (fasta format) using "Import data" option

								_
<u>Eile Edit H</u> elp								
Partitions Taxa Tips Tra	Partitions Taxa Tips Traits Stees Clocks Trees States Priors Operators MCMC							
Unlink Subst. Models	Link Subst. Models I Unlink C	lock Models Link Clock Mode	els i Unlink Trees Link T	Trees				
Partition Name	File Name	Taxa	Sites	Data Type	Site Model	Clock Model	Partition Tree	î
sars-cov-2_aln	sars-cov-2_aln.has	36	29891	nucleotide	default	erault	default	
								- 1
								- 1
								- 1
								- 1
								- 1
								- 1
								- 1
								- 1
								- 1
								- 1
								- 1
								- 1
								- 1
								- 1
								- 1
								- 1
								- 1
								- 1
								I
<								, ``
+ - View Partition	Create partition from tr	ait						
Data: 36 taxa, 1 partition							🏘 Generate B	EAST File

From "Tips" tab \rightarrow Select "Use Tip dates" \rightarrow Parse dates

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: last
Prefix:
O Defined by regular expression (REGEX)
O Parse as a number
Add the following value to each: 1900
unless than: 16
in which case add: 2000
O Parse as a calendar date
Date format: yyyy-MM-dd ?
• Parse calendar dates with variable precision
😵 Cancel 🛛 😪 OK

From "Sites tab" \rightarrow Substitution Model: GTR; Base frequencies: Estimated; Site Heterogenity model: Gamma + Invariant Sites

Ele Edt Help	
Partitions Taxa Tips Traits Sites Clocks Trees States Priors Operators MCMC	
Substitution M. Nucleotide Substitution Model - default	
default Substitution Model: CTR 💌	
Base frequencies: Estimated v	
Site Heterogeneity Model: [Camma + Invariant Sites] 🗸	
Number of Gamma Categories: 4 💌	
Partition into codon positions: Off	
Link/Unlink parameters Image: Substitution rate parameters across codon positions Image: Ima	
Clone Settings	
Data: 36 taxa, 1 partition	Generate BEAST File

From "Clocks" tab \rightarrow Clock Type: Uncorrelated relaxed clock; Distribution: Lognormal

Activities	🛃 dr-app-beauti-BeautiApp 🔻		20 Aug 08:42 •	₩ • •
			BEAUti	- = 8
	Elle Edit Help			
• >_	Partitions Taxa Tips Traits Sites Clocks Trees States Priors Operators MCMC			
	Clock Model CLOck Mobel: "Dei Jolic Gefault	Clock Type:	Uncorrelated relaxed clock	
•			Rambaut (2006) PLoS Biology 4, e88.	
		Relaxed Distribution:	Lognormal	
			Use continuous quantile parameterization.	
• 🔤				
Â				
?				
•				
· 🏠				
	Clone Settings			
	Data: 36 taxa, 1 partition			& Generate BEAST File

From "Priors tab" \rightarrow Select "**ucld.mean**" and use the following values; repeat same for "u**cld.stdev**"

			BEAUti	-	e X
Eile Edit Help				Prior for Parameter ucld mean	
Partitions Taxa Tips Traits Sites Clocks Tre	ees States Priors Operators MCMC				
Use classic priors/operators				select phot distribution for deta.mean	
Parameter	Prior	Bound	Description	Prior Distribution: Normal	
gtr.rates	* Dirichlet [1,1]	[[0, ∞]	GTR transition rates parameter		
Frequencies aloba	* Dirichlet [1,1]	[0, ∞]	base frequencies	leitial values 7.500/E 4	_
plnv	* Uniform [0, 1], initial=0.5	[0, 1]	proportion of invariant sites parameter	Initiat value. 7.3626E*4	_
ucld.mean	* Approx. Reference Prior, in	[0, ∞]	uncorrelated lognormal relaxed clock mean	Mean: 7 56265-4	
treeModel rootHeight	* Exponential [0.333333], Init	[0, ∞] [2.213699 ∞]	uncorrelated lognormal relaxed clock stdev	Mean. 7.30200-4	
constant.popSize	* 1/x, initial=1	[0, ∞]	coalescent population size parameter	Stday 0.1	
				Sidev. U.I	
				Truncate to:	
				Upper: +INF 00	
				0 -0.2 -0.1 -2.77555706-17 0.1 0.2	
				Quantile: 2.5%: -0.195	
				Madian: 7.5630-4	
				97.5%: 0.197	
Link parameters together Link paramet	ers into a hierarchical model I Unlink p	arameters			
* Marked parameters currently have a defaul	lt prior distribution. You should check that	these are approp	riate.	Cancel 🛇 OK	
Data: 36 taxa, 1 partition				& Generate BEAST	File

From "Priors tab" \rightarrow Select "treeModel.rootHeight" and set "Prior Distribution" to **uniform**

Activities	🛃 dr-app-beauti-BeautiApp 🔻			20 Aug 08:50 •	♣ � Ů ▾
-				BEAUti	8
· 🕑 🛛	ile <u>E</u> dit <u>H</u> elp				
	Partitions Tava Tins Traits Sites Clocks To	ees States Priors Operators MCMC			
• >-		ces states mais operators meme			
	Use classic priors/operators	Delas	Deved	Description	
	otrates	*Dirichlet [1 1]		CTR transition rates parameter	
	frequencies	*Dirichlet [1,1]	[0, ∞]	base frequencies	
	alpha	*Exponential [0.5], initial=0.5	[0, ∞]	gamma shape parameter	
•	ucld mean	Vormal [7, 5626E-4, 0, 1], initial=0.5	[[0, 1] [[0, m]	proportion of invariant sites parameter upcorrelated lognormal relaxed clock mean	
	ucld.stdev	Normal [7.5626E-4, 0.1], initi	[0, ∞]	uncorrelated lognormal relaxed clock stdev	
	treeModel.rootHeight	Uniform [2.213699, 1E100]	[2.213699,	1E100] root height of the tree	
	constant.popsize	. *1/x, initial=1	_[[0, ∞]	coalescent population size parameter	
				Prior for Parameter treeModel.rootHeight	
				Select prior distribution for treeModel.rootHeight	
• 🔛					
-				Prior Distribution: Uniform	
				opper: 1.0e100	
(?)				Lower 2136986301368324	
				Lowel. 2100000100024	
				😵 Cancel 💽 OK	
	Link parameters together Link parame	ters into a hierarchical model I Unlink pa	rameters		
	* Marked parameters currently have a defai	It prior distribution. You should check that	hese are an	unconclate	
	marked parameters contently have a derac	ac provolativo do a lo di si lo di di check triat	incacial e ap	propriots.	
	Data: 36 taxa, 1 partition			0	Generate BEAST File

		BEAULI		- * 😣
Eile Edit Help				
Partitions Taxa Tips Traits Sites Clocks Trees States Priors Operators MCM	1C			
	Length of chain:	1000000	_	
	Echo state to screen every:	1000		
	Log parameters every:	1000		
	File name stem:	sars-cov-2_aln		
		Add.txt suffix		
	Log file name:	sars-cov-2_aln.log	-	
	Trees file name:	sars-cov-2_aln.trees		
		Create tree log file with branch length in substitutions:		
	Substitutions trees file name:			
1		Create operator analysis file:		
	Operator analysis file name:	sars-cov-2_aln.ops		
		Sample from prior only - create empty alignment	-	
	Select the option below to perform m stepping-stone sampling (SS) or gene additionalanalysis after the standard I	arginal likelihood estimation (MLE) using path sampling (PS) / ralized stepping-stone sampling (GSS) which performs an MCMC chain has finished.		
	Marginal likelihood estimation (MLE):	None		
		Settings		
Data: 36 taxa, 1 partition				& Generate BEAST File

From "MCMC" tab, use default parameters \rightarrow Select "Generate BEAST File" (bottom right)

The output file is saved as ".xml" file

BEAST

Choose the ".xml" file generated using BeauTi, run BEASt with all default options

E	BEAST v1.10.4 Prerelease #bc6cbd9 8					
SE .	BEAST Bayesian Evolutionary Analysis Sampling Trees Version v1.10.4 Prerelease #bc6cbd9, 2002-2018 https://github.com/beast-dev/beast-mcmc/commit/bc6cbd9					
BEAST XML File:	sars-cov-2.xml Choose File					
	Allow overwriting of log files					
Random number seed:	1660982568611					
Thread pool size:	Automatic 💌					
The BEAGLE library is	required to run BEAST:					
Prefer use of: CPL						
	Jse CPU's SSE extensions when possible					
Prefer precision: Dou	apple 🔺					
Rescaling scheme: Def	ault 👻					
S	Show list of available BEAGLE resources and Quit					
BEAGLE is a high-performance phylogenetic library that can make use of additional computational resources such as graphics boards. It must be downloaded and installed independently of BEAST: http://github.com/beagle-dev/beagle-lib/						
Run Quit						

The output files generated are: .log and .trees

Tracer

Upload the log file generated using BEAST (select "+" symbol on left panel)



Treeannotator

Set "Burnin (as states)": 1000; posterior probability limit: 0.8 "Input Tree file": ".trees" file generated by BEAST

generate a blank out file

"Output File": Select the blank file generated earlier

TreeAnnotator v1.10.4 Prerelease #bc6cbd9 🛛 😵					
Specify the burnin as the	Specify the burnin as the number of states				
Burnin (as states):	1000				
⊖ Specify the burnin as the	e number of trees				
Burnin (as trees):					
Posterior probability limit:	0.8				
Target tree type:	ype: Maximum clade credibility tree 💌				
Node heights:	Median heights	-			
Target Tree File:	not selected	Choose File			
Input Tree File:	sars-cov-2.trees.txt	Choose File			
Output File:	out.tree	Choose File			
	Run Quit				

Figtree

Set the following options from Control panel in Figtree (left panel)

- 1) Time Scale \rightarrow Scale by factor \rightarrow Offset by: 2022 (date of isolation of latest entry)
- 2) Node labels \rightarrow Display \rightarrow Node Ages
- 3) Scale Axis \rightarrow Reverse Axis

Save the tree file

