Journal Title Here, 2026, pp. 1–17

doi: DOI HERE
Advance Access Publication Date: Day Month Year
Paper

PAPER

baltic: the Backronymed Adaptable Lightweight Tree vIsualisation Code

Barney I. Potter,[1] Karthik Gangavarapu,[2] Sidney M. Bell,[3] María Fernanda Torres Jimenez[4] and Gytis Dudas[5,*]

[1]Yale School of Public Health, Yale University, New Haven, Connecticut, USA, [2]Department of Translational Medicine, Scripps Research Institute, San Diego, California, USA, [3]Chan Zuckerberg Initiative, Redwood City, California, USA, [4]Department of Zoology, Institute of Biosciences, Life Sciences Center, Vilnius University, Vilnius, Lithuania and [5]Department of Eukaryote Gene Engineering, Institute of Biotechnology, Life Sciences Center, Vilnius University, Vilnius, Lithuania

[*]Corresponding author. gytis.dudas@gmc.vu.lt

FOR PUBLISHER ONLY Received on Date Month Year; revised on Date Month Year; accepted on Date Month Year

Abstract

For ten years, baltic (Backronymed Adaptable Lightweight Tree vIsualization Code) has been used to make annotated phylogeny figures in molecular epidemiology, including during the West African Ebola epidemic, the Zika epidemic in the Americas, and the SARS-CoV-2 pandemic, as well as other fields. At its core, baltic is a Python library used for the efficient parsing, traversal, manipulation, and visualisation of phylogenetic trees. It is a small library with few dependencies that reads tree formats common in phylodynamic analyses and gives the user leverage to interact with a lightweight tree data structure to produce publication-ready figures with matplotlib. We present baltic v1.0, its first formally released and documented version. baltic reads and writes BEAST Nexus, Newick, and Nextstrain/Auspice JSON, and can process large BEAST posterior tree files in parallel to extract user-defined posterior statistics. The same objects are used for tree manipulation and for plotting in a single script. New to this release: a set of rooting methods (midpoint rooting, rerooting on any branch, and root-to-tip regression); support for reticulate evolution, with reassortment and recombination edges; and composite figures that combine a tree with other data, such as Müller plots, skygrid plots, tanglegrams, and plots connecting trees to maps. The release includes a documentation site with an API reference, tutorials, and a matplotlib-style gallery of worked examples.

Key words: phylogenetics, phylodynamics, data visualisation, tree manipulation

Introduction

Modern phylogenies are (too) big

The scale of modern sequence data has transformed phylogenetics from a field limited by data availability into one increasingly constrained by interpretation. For example, since the beginning of the COVID-19 pandemic, over 15 million SARS-CoV-2 genomes have been sequenced and shared on sequence databases. Phylogenetic trees remain the central abstraction for reasoning about evolutionary relationships and population processes, but they are also an inherently difficult visual medium. At their core, modern phylogenetic visualisations are at best pseudo-two-dimensional: typically one of two visualisation dimensions is sacrificed to ensure taxa do not overlap and to convey the nested hierarchy of relationships between taxa, while the other visualisation dimension is commonly used to depict the amount of genetic or temporal distance on branches of the tree. In combination with increasingly more taxa and metadata, trees rapidly become dense, label-heavy, and uninterpretable at first glance. This is further exacerbated by sophisticated analyses that include phylogenetic uncertainty, ancestral state reconstruction, novel data structures (*e.g.* multitype trees, reassortment networks), as well as new formats (*e.g.* `Nextstrain`'s `Auspice` JSON). We aim to address these shortcomings via novel visualisation techniques which we implemented in an updated Python software package. Here we present `baltic`, the **b**acronymed **a**daptable **l**ightweight **t**ree **v**isualisation **c**ode, a Python ($\geq$3.5) module designed for parsing and manipulating phylogenetic trees which leverages `matplotlib` ($\geq$ 2.0.0) [1] as the visualisation engine. `baltic` focuses on user friendliness, thoughtful static data visualisation and modern phylogenetic analysis outputs. While `baltic` has seen some use in the past [2–181], here we present the latest version 1.0 with an improved API and newly implemented features.

baltic philosophy and gallery

When writing `baltic`, we were heavily inspired by `matplotlib`, which `baltic` uses for rendering. Like `matplotlib`, `baltic` providers users leverage and flexibility to construct complex visualisations out of simpler elements. For example, one of the most

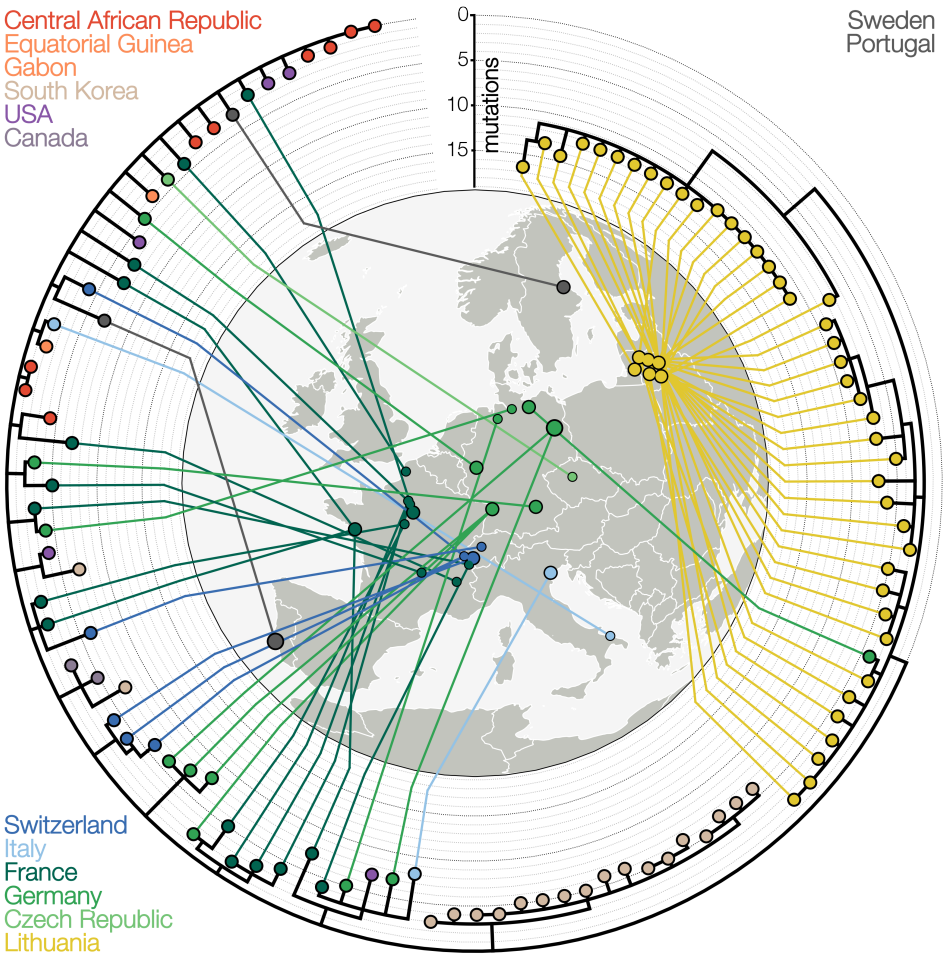


Fig. 1. An inward-facing circular tree around a map of Europe adapted from [6] showing a maximum likelihood tree of SARS-CoV-2 lineage B.1.620, first recognised in Lithuania, and connected to its sampling locations and coloured according to country.

argument-rich functions in `matplotlib`'s code base is used for drawing box-and-whiskers plots, arising out of the need to give users aesthetic control over each sub-element of the visualisation. While `baltic` also implements some field-specific advanced visualisations as single functions (*e.g.* exploded trees or state-collapsed trees), our goal is to strike a balance between utility and functional simplicity. To this end, we follow `matplotlib`'s example and rather than trying to foresee every possible `baltic` use case, we built a gallery of examples using `baltic` to showcase how various `baltic` functions can be used together to produce more advanced visualisations (*e.g.* Fig. 1). This gives users a solid starting point and complete artistic and study-specific control over the users' visualisations.

Implementation(s)

baltic core

`baltic` parses phylogenetic trees supplied in Newick, Nexus, or `Nextstrain/Auspice` JSON format [182] into a common in-memory representation. For Newick- and Nexus-derived strings, `make_tree` performs a single left-to-right pass over the tree string with a cursor index, incrementally matching tokens for parentheses (new internal nodes), tip labels (both plain and `BEAST`-style integer-encoded taxa), branch lengths, and annotation comments (*e.g.* `[&posterior=1.0,...]`), which are parsed into a per-branch `traits` dictionary; reticulate edges denoting recombination, reassortment, or conversion events are recognized via `#`-prefixed labels and linked to their destination once both ends have been seen. The `baltic.io` module provides format-specific wrapper functions: `load_newick` and `load_nexus` locate the tree string within a file (the latter also resolving Nexus `Translate` blocks that map numeric tip codes to full taxon names) and can extract tip sampling dates from taxon labels via a user-supplied regular expression, while `load_JSON` walks an `Auspice` v2 JSON tree recursively (`make_tree_JSON`), remapping `Nextstrain`'s `node_attrs`/`branch_attrs` fields onto the same internal attributes used by `baltic` Newick/Nexus parsers so that all three input formats converge on one tree object.

Every parsed tree has a `treeType` of either `"divergence"` or `"time"`, which governs how branch lengths are interpreted: in a divergence tree, lengths represent substitutions (or another divergence unit) and a pre-order traversal (`traverse_tree`) accumulates them into a `height` measured from the root, whereas in a time tree the same accumulated height is additionally mapped onto calendar time (`absoluteTime`) using either a most recent sampling date (`set_absolute_time`) or dates parsed directly from tip labels. Structurally, `baltic` stores each branch as an instance

of a class hierarchy rooted in the `BranchLike` superclass, which holds attributes common to every branch (incoming `length`, `height`, `absoluteTime`, `parent`, a `traits` dictionary of parsed annotations, and plotting coordinates). `BranchLike` is subclassed into `Leaf` (terminal taxa), `Node` (internal ancestors, which additionally track their `children` and the set of descendant tip names), `Reticulation` (non-vertical edges representing reticulate events, linked to a `target` branch elsewhere in the tree), and `Clade` (a placeholder representing a collapsed monophyletic subtree for compact display).

These branch objects are held collectively by a `Tree` object, which stores a flat list of all `BranchLike` objects (`Objects`), a reference to the `root`, the cursor (`curNode`) used during parsing, and exposes traversal, sorting, statistics, and plotting methods that operate uniformly across the branch-object hierarchy regardless of input format.

Python functions in baltic

Much of `baltic`'s flexibility comes from its canonical use of, and default design for Python functions as arguments (particularly Python's `lambda` functions). Most of `baltic`'s API interfaces, be they for selecting taxa for plotting or describing a color mapping, expect functions (typically taking just the argument `k`, a `baltic.branchLike` object) as their argument. The generous use of this Python feature allows users great flexibility in the specific implementations they choose, allowing both very simple functions that only return a single value on any input (*e.g.* a colour function that only returns "black"), or the expansion of more complex mappings into complex functions.

baltic format interoperability

`baltic` supports several file formats commonly used in phylogenetics, and though newick and nexus formats have been standard for many decades, their full range of flexibility has not always been supported by downstream software. For example, the comment format of nexus files is widely employed within the `BEAST` [183] ecosystem to encode information like branch-specific molecular clock rates and ancestral state reconstruction but this is not always supported by all phylogenetic tree parsers. Extensions beyond strictly clonal trees like ancestral recombination graphs, reassortment networks and other reticulate data structures are accommodated by even fewer parsers. Though newick and nexus file formats remain the *de facto* community standard, there are competing formats. For example, the `Auspice` JSON format developed and used in the `Nextstrain` ecosystem is often employed in genomic epidemiology (*i.e.* public health) contexts.

`baltic` currently supports newick, nexus, and `Auspice` JSON file formats as inputs, as well as the extended `newick` format as seen in the `BEAST2` [184] ecosystem and implemented in `IcyTree` [185]. Of these, the ability to handle the `Auspice` JSON format is of particular note, as few phylogenetic parsers are currently able to handle it. To the best of our knowledge these are `treeio` [186] for R, and `Nextstrain`'s `Auspice` with `taxonium` [187] for javascript-based web visualisation but nothing, to our knowledge, for Python. Likewise, parsing reticulate data structures remains rare, with `PhyloNetworks.jl` [188] (Julia) and `IcyTree` (javascript/web) being the few examples, and nothing available, to our knowledge, in Python currently.

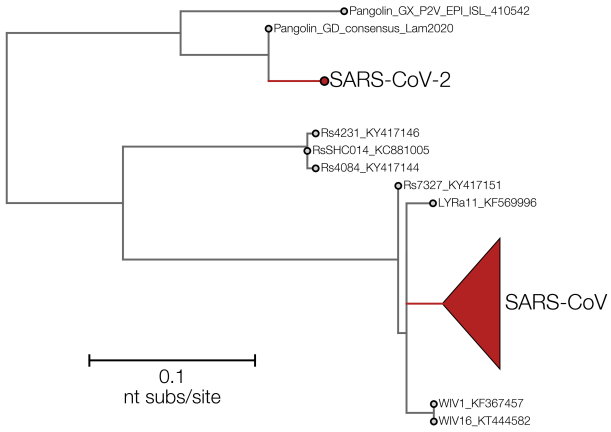


Fig. 2. A maximum likelihood phylogenetic tree of betacoronaviruses from [189]. SARS-CoV (collapsed clade) and SARS-CoV-2 are highlighted in red, with extra vertical space added around both using `baltic`'s `padNodes` feature.

Standard visualisations

Users of `baltic` have access to all the typical phylogenetic visualisation approaches. For large subtrees, `baltic` offers the possibility of collapsing clades into narrower triangles to conserve space. For cases where some internal branches are not needed, *e.g.* polytomies resolved into arbitrary bifurcations (either with zero or miniscule branch length) or low statistical support, `baltic` allows users to collapse these according to custom filtering functions. If a user is more interested in specific clades, `baltic` offers the ability to extract subtrees descended from any node. `baltic` also allows users to reroot (substitution space, *i.e.* `treeType='divergence'`) phylogenetic trees, and implements root-to-tip regressions for measurably evolving populations. Examples of subtree extraction, collapsing clades, rerooting, and padding nodes (see "Padded nodes and tanglegrams" section) are shown in Fig. 2.

Exploded trees

During the 2013–2015 West African Ebola virus epidemic, Philippe Lemey (KU Leuven) developed a visualisation for Ebola virus trees annotated with reconstructed geographic regions [10]. The visualisation depicted the time span between the arrival of an Ebola virus lineage to a new region and the sampling of that introduction's last descendant in the same region, ignoring lineages that migrated elsewhere. This approach conveys information about the number, size (in terms of sequenced genomes) and duration of each sub-outbreak. This idea was implemented in `baltic` using conditional tree traversals - each branch whose parent is in a different discrete state initiates a tree traversal conditioned on each child branch remaining in the new location, with subtrees resulting in tip-less subtrees being discarded. The resulting state-specific subtrees are grouped together and displayed, often with the pre-introduction state visually indicated with colour (Fig. 3A). This rearrangement of the full tree into blocks based on discrete traits allows the impact of each trait state's effect on the fate of introduced lineages to be immediately interpretable. Andrew Rambaut (University of Edinburgh) coined the term "exploded trees" for this visualisation and it has been used to define "transmission lineages" [190] during the COVID-19 pandemic and later implemented in the `Nextstrain` ecosystem under the "Explode tree by" option.

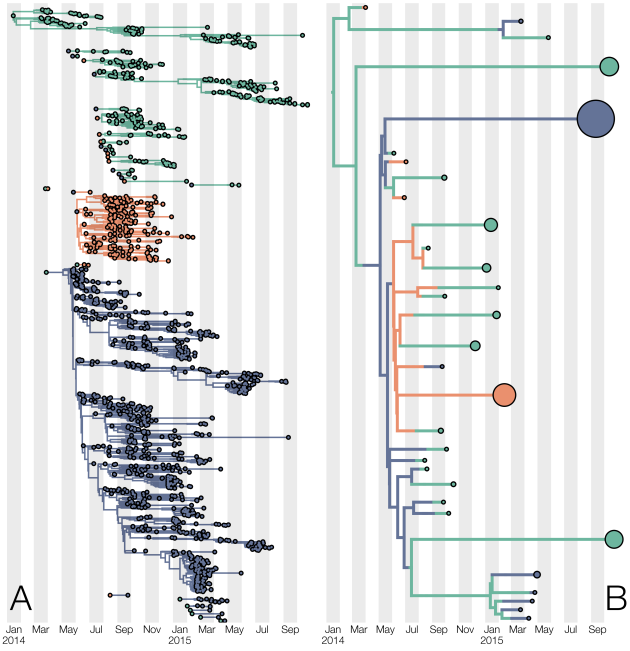


Fig. 3. A) An "exploded tree" visualisation derived from a molecular clock tree of 1610 Ebola virus genomes from the 2013-2015 West African Ebola virus epidemic from [10]. Each subtree begins with an international migration between the three most affected countries during the epidemic (Guinea in green, Sierra Leone in blue, Liberia in red) and tracks lineages for as long as they stay in the country of introduction. **B**) A "state-collapsed" visualisation derived from the same original tree as A). Rather than displaying entire subtrees (A), it preserves the phylogenetic position of the earliest tip descended from an international introduction, whose "tip" date is adjusted to indicate when the last sampled descendant of the introduction was sampled and circle size is scaled according to how many tips in total descend from the introduction. As before, Guinea is shown in green, Sierra Leone in blue, and Liberia in red.

State-collapsed trees

While the exploded tree visualisation merely rearranges an existing tree topology by sacrificing branches that have state changes and retaining relationships within the same state, there is a more compact approach. Initially developed around 2017 for the visualisation of the 2013–2015 West African Ebola virus epidemic and never used in a publication, the method involves partitioning the tree into stretches of continuous evolution under a given trait state following a migration. This is done by traversing the tree from the root with a heritable unique label which switches to a new label if a child branch switches to a new trait state. As a result, these assigned labels partition the tree into groups of related branches that represent continuous evolution in the new trait state. Then, for each unique label the earliest sampled tip is chosen to represent the entirety of that label, its height is adjusted to correspond to the last sampled tip with the same label, the number of tips with that label is represented by the circle size of the representative tip, and all non-representative tips are removed from the tree while preserving their phylogenetic relationships (see next sections). We call this visualisation "state-collapsed trees" (Fig. 3B), as it yields a compact visualisation of the history of trait transitions with the final number of tips equal to the number of transition events. Due to the pruning procedure state-collapsed trees yield multitype trees (with node objects that have a single descendant branch) that retain information about the timing (for time trees) of transition events as well. State-collapsed trees are conceptually close to phylotype maps from `PhyloType` [191], and inter-cluster transition trees from `EvoLaps2` [192] but yield phylogenetic data structures (multitype trees) that can still be imported by common phylogenetic software like `FigTree` [193], though trees with reconstructed ancestral states at internal nodes are arguably the most impactful use case.

Reduced trees, rerooting, and processing posterior distributions

Like many other phylogenetic parsing, analysis, and visualisation packages, `baltic` allows users to reroot non-molecular clock trees (translated from `BioPython`'s Phylo module), drop individual tips, and depict trees in circular and unrooted configurations. Owing to its genomic epidemiology roots, `baltic` also allows users to automatically infer or manually assign tip dates in absolute time as well as their uncertainty (*e.g.* when only the sample year but not full date of collection is known), and infer the root of the tree that optimises (via sum of squares, r^2 or correlation coefficient) a root-to-tip regression. Similarly, `baltic` can process posterior sets of trees in `BEAST` nexus format to extract various posterior statistics of interest, *e.g.* common ancestor dates (TMRCAs) or discrete trait probabilities along a single path between a tip and the root. Posterior tree set handling is done via a submodule of `baltic` called `samogitia` where users can use custom functions for processing each posterior realisation of a tree and return processed outputs to a standard Bayesian phylogenetic trace/log file.

Padded nodes and tanglegrams

Given the ever-growing size of phylogenetic trees, `baltic` natively implements explicit padding around specified branches. This is similar to `ggtree`'s [194] (R) `scaleClade` function, though in the `baltic` implementation users provide a dictionary (`padNodes`) where keys are branch objects and values are

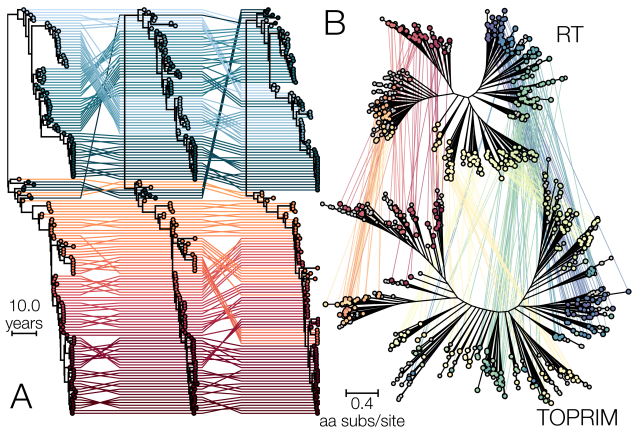


Fig. 4. A) A tangled chain of influenza B virus PB1, PB2 and HA segment time trees (in that order from left to right) from [18], down-sampled three-fold and with coloured lines connecting sequences derived from the same genome across all trees. Blue lines indicate Yamagata lineage sequences (in HA tree), red lines are Victoria lineage (in HA tree) with different shades indicating the y-axis position of each tip in the PB1 tree. **B**) An unrooted tanglegram of reverse transcriptase (RT) and topoisomerase-primase (TOPRIM) domains from retron types I-C, XI, and XII connecting domains derived from the same protein from [150]. Tip circles and connecting lines are coloured according traversal order from blue to red.

`ints` or `floats` that specify the amount of vertical space to be added around the specified branch (in `baltic`, each tip receives 1 unit of vertical space around it by default). `baltic` also supports tangled chain visualisations, where two or more rooted trees are displayed in series with tips representing the same sample in different trees are connected with lines to highlight phylogenetic incongruence (Fig. 4A). However, `baltic` provides the flexibility to easily extend this visualisation type to unrooted trees by providing users access to tree plotting coordinates (Fig. 4B). We also provide the option to "untangle" (reduce connecting line crossing) trees prior to plotting tangled chains, though we believe it to be misleading since trees can have tips positioned at the same y-axis coordinates (and thus have non-intersecting tanglegram lines) whilst not being phylogenetically congruent.

Müller plots

Müller plots are a unique and uncommon visualisation technique that preserves phylogenetic information and depicts absolute or relative genotype frequencies through time [196]. It was first used to intuitively explain the concepts of Müller's ratchet, clonal interference and the advantages of recombination in generating fit genotypes. Typically, the x-axis of Müller plots conveys the passage of time, the y-axis shows genotype frequencies and the hierarchical relationships between genotypes is depicted as descendant genotype frequencies emerging from the middle of the parental genotype frequency. There are relatively few software packages for doing Müller plots, with `ggmuller`, `MullerPlot`, and `EvoFreq` [197] being notable examples in the R ecosystem, and `pyfish` [198] and `muller` in the Python ecosystem. The `Augur` (Python) package, the back-end of `Nextstrain` analyses, is a unique example, where clade frequencies for individual phylogenetic trees are computed via nested logistic smoothing of temporal observations (tips in the tree). While `Augur` infers the frequency trajectories that are the basic building block of Müller plots, `Auspice`, the front-end part of `Nextstrain` doesn't display them as Müller plots in the strictest sense. In `Auspice`'s version of Müller plots simpler stacked area plots. In `baltic` we offer users an option for plotting frequencies such that the starting y-axis coordinates of descendant genotypes begin in the middle of the parental genotype frequency, thus preserving information about phylogenetic relationships. This stylistic choice is made via the boolean `Muller` argument of `baltic`'s `plot_Muller` function (`True` for strict Müller diagrams with descendant frequencies emerging from the ancestor frequency and `False` for `Auspice`-style plots) in `baltic`.

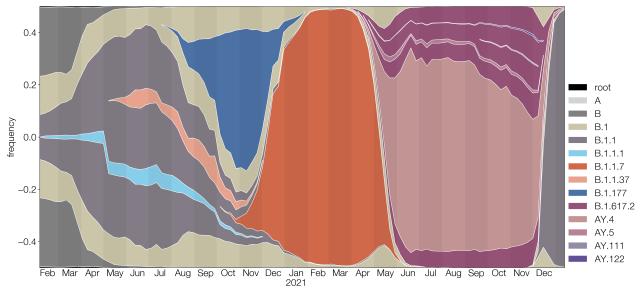


Fig. 5. A Müller plot of SARS-CoV-2 derived from more than 1 million genomes using data from [195]. The Müller plot depicts both the relative frequencies of multiple pre-specified SARS-CoV-2 lineages (indicated by colour, legend to the right) and abstractly depicts their phylogenetic relationships derived from pangolin lineage labels.

In `baltic`, we have reused `Augur`'s code to compute clade frequencies but modified it so observations (dates of tips) can be assigned directly to nodes in the tree. By doing this, we enable the user to compute and depict Müller plots with millions of observations while maintaining a minimal underlying phylogenetic data structure. The typical use case we envision is taking an abstract lineage tree (*e.g.* pango nomenclature for SARS-CoV-2 [199]) and annotating its branches with sample collection dates of each lineage. An example of this is shown in Fig. 5, which depicts the relationships and frequencies of over 1 million SARS-CoV-2 genomes from UK's SARS-CoV-2 genomic surveillance programme.

Conclusion

`baltic` is a lightweight and flexible Python tool for parsing, analysing, manipulating and visualising phylogenetic trees. By exposing much of the underlying (*e.g.* traits) and derived (*e.g.* plotting coordinates) attributes of the phylogenetic data structure, we give users the freedom to analyse, manipulate and process phylogenetic trees, and build their own publication-ready static figures. To help new users we provide a useful starting point with extensive tutorials and gallery. We encourage first time users to provide feedback that could improve `baltic`'s usability and advanced users to request support for non-standard phylogenetic data structures or visualisations.

Acknowledgments

Over the years, many people contributed to `baltic` directly or indirectly by regular use and inadvertent testing. In particular, we'd like to thank Anderson Brito, Verity Hill, Gage Moreno, Louise Moncla, Edyth Parker, Miguel Paredes, Jesse Bloom, John Huddleston, Nicola Müller, Áine O'Toole, Jonathan Pekar, Allison Black, Ifeanyi Omah, Matthew Scotch, Raphaëlle Klitting, Alvin X. Han, Praneeth Gangavarapu, and many others. A special thanks goes to Andrew Rambaut and Trevor Bedford for support, advice, ideas, and dedication to the craft of data visualisation in phylogenetics.

GD acknowledges funding from the Research Council of Lithuania, Lithuania under the EMBO Installation Grant programme under grant No 5305.

Materials and Methods

Data availability

Figure 1 is a simplified tree from [6]. GISAID accessions from a study of SARS-CoV-2 lineage B.1.620 [6] were recovered from a publicly available tree file (`https://github.com/evogytis/B.1.620-in-Europe/blob/main/data/trees/latest_continent_mcc.tre`). All accessions were downloaded from GISAID (for acknowledgments see EPI_SET `https://epicov.org/epi3/epi_set/260918gh`), aligned using `NextClade` CLI [200], hand-picked for quality, and the maximum likelihood phylogenetic tree recovered using `RAxML` v.8.2.11 [201] under a GTR [202] + Γ_4 [203] model.

Figure 2 is derived from a study into ACE2 binding of coronaviruses [189]. The tree used was publicly available at `https://github.com/jbloomlab/SARSr-CoV_homolog_survey/blob/52db008d423a17d622950d6311fd89392cdc5e1d/RBD_ASR/RBD_rm_RaTG13/ASR/tree.newick.txt`. Prior to visualisation, the clade descending from the common ancestor of SARS-CoV

and SARS-CoV-2 was extracted using the `subtree` method of the `baltic.Tree` class, midpoint rooted and the same `subtree` method called again on the common ancestor of SARS-CoV and SARS-CoV-2.

Figure 3A is a reproduction of Figure 4 from [10], and both it and Figure 3B are derived from a publicly available tree at `https://github.com/ebov/space-time/blob/master/Data/Makona_1610_cds_ig.GLM.MCC.tree`.

Figure 4A depicts three of eight segment trees of influenza B virus from a study into reassortment dynamics between Victoria and Yamagata lineages of this virus [18]. The trees are available at `https://github.com/evogytis/fluB/blob/master/data/mcc%20trees/InfB_PB1t_ALLs1.mcc.tre`, `https://github.com/evogytis/fluB/blob/master/data/mcc%20trees/InfB_PB2t_ALLs1.mcc.tre`, and `https://github.com/evogytis/fluB/blob/master/data/mcc%20trees/InfB_HAt_ALLs1.mcc.tre`.

Figure 4B is a reproduction of Supplementary Figure 1 from [150], showing the phylogenetic trees of reverse transcriptase (RT) and topoisomerase-primase (TOPRIM) domains of retron Eco2 and its relatives. Original tree files and annotations are available from `https://data.mendeley.com/datasets/v9b289v6ct/1`.

Figure 5 is derived from data on SARS-CoV-2 lineage dynamics in the UK [195] and is based on raw data available at `https://github.com/qinqin-yu/sars-cov-2_genetic_drift/blob/main/data/metadata/cog_england_2022-01-16_metadata.csv`.

All data and code to reproduce figures provided here are available at `https://github.com/phylo-baltic/baltic-v1-manuscript`. The code for `baltic` is available at `https://github.com/evogytis/baltic`.

User resources

We provide documentation for `baltic` at `https://baltic.readthedocs.io/en/latest/index.html`. Recognising that `baltic`'s strengths lie in combining multiple features in non-trivial ways, we also provide a gallery and associated code for advanced visualisations at `https://phylo-baltic.github.io/baltic-gallery/`.

Large language model use disclosure

At various points during the development of baltic v1.0 we used Codex (ChatGPT 5.4, 5.5, and 5.6 (High)) and Claude to fix bugs, optimise and implement new functions, and preparation of documentation.

Table 1: Highlights of `baltic`'s functions and features with comparable functionality in other phylogenetic software.

Function or feature	Brief description	Comparable functionality in other packages
Phylogenetic tree format support		
`Nextstrain`'s `Auspice` JSON v2	A format produced by `Nextstrain`'s `Augur` module with rich annotations.	`Nextstrain`'s `Auspice`; `taxonium`; `treeio`.
Cardona-style extended newick format [204]	A phylogenetic network format featuring reticulate edges marked connected via "`#H<number>`" tags in the tree string.	`IcyTree`; `ape::plot.evonet` [205]; `phangorn::plot.networx` [206]; `tanggle` [207]; `PhyloNetworks.jl`.
Subsetting and topology manipulation		
`Tree.subtree`	Deep-copy a subtree descending from a chosen branch, optionally retaining its stem and applying a traversal predicate.	`ape::extract.clade`; `Bio.Phylo.Tree.from_clade` [208]; `DendroPy` [209] `extract_tree`; `ETE` [210] subtree copying or detachment.
`Tree.reduce_tree`	Return the minimal topology connecting a specified set of retained tips, while suppressing unnecessary internal nodes.	`ape::keep.tip/drop.tip`; `ETE prune`; `DendroPy retain_taxa_with_labels` or `extract_tree_with_taxa`; repeated `Bio.Phylo.prune`.
`state_collapse_tree`	Compress adjacent branches that remain in the same inferred state, producing a reduced "phylotype" or state-transition tree.	PastML compressed ancestral-scenario trees are the closest conceptual equivalent; stochastic character maps from `phytools::make.simmap` [211] provide related state-labelled trees without the same compression.
Rooting, ordering, and tree-set processing		
`root_by_regression`;	Choose the continuous root position that optimises root-to-tip regression; supports uncertain sampling dates and parallel evaluation.	Closest equivalents are `TreeTime`'s [212] least-squares clock rooting and `TempEst`'s [213] root-to-tip optimisation. General-purpose tree libraries usually provide midpoint or outgroup rooting only.
`posterior_tree_iterator`; `process_posterior_trees`	Stream trees after burn-in from a posterior `.trees` file and apply a worker function in parallel.	`DendroPy.Tree.yield_from_files`; `Bio.Phylo.parse`; `BEAST LogCombiner/TreeAnnotator` workflows; `ape::read.nexus` followed by `parallel`.
`trace_lineage_trait_worker`; `tmrca_worker`; `tree_length_worker`	Posterior-tree workers for tracing ancestral trait states, calculating TMRCA distributions, and recording total tree length provided for `process_posterior_trees` API reference.	Comparable analyses can be assembled using `treeio/tidytree` over `BEAST` samples, `DendroPy` tree iterators, or `BEAST/Tracer` summaries; `baltic` supplies ready-made streaming workers.
Core tree drawing and annotation		
`padNodes`	Adds configurable empty space on both sides of selected subtrees or individual branches. In circular and unrooted layouts, this becomes additional angular separation around the subtree.	Closest to `ggtree::scaleClade`, although `scaleClade` rescales spacing within a clade rather than adding margins around it. `ETE`'s `branch_vertical_margin` is a global analogue; `ape::plot.phylo` requires manual coordinate manipulation.
`plot_exploded_tree`	Display trait-delimited subtrees as separated components while retaining their time or divergence coordinates.	Closest alternatives use `ggtree::groupClade` with facets or separate `ape`/`ggtree` panels; `Nextstrain`'s `Auspice` "Explode Tree by" functionality is directly inspired by the `baltic` implementation.
`plot_node_piechart`; `plot_node_bar`; `plot_node_treemap`	Display discrete ancestral-state probabilities at a node as a pie, stacked bar, or treemap.	`ape::nodelabels(pie=...)`; `phytools::nodelabels`; `ggtree::geom_inset` with node-pie or custom ggplot objects. Treemap glyphs have few direct phylogenetic equivalents.
Composite and phylodynamic visualisations		

Continued on next page

Table 1 — continued

Function or feature	**Brief description**	**Comparable functionality in other packages**
`tree_frequencies`; `plot_Muller`	Estimate nested clade frequencies through time and plot them as a Müller-style stacked-frequency diagram.	Directly ported from `Nextstrain`'s `Augur` frequency estimation, with visualisation handled by `Nexstrain`'s `Auspice`; `ggmuller` draws Müller plots but is not intrinsically tree-aware.
`connect_tree_to_map`	Draw lines from plotted tree tips to their geographic coordinates on a Cartopy map, with optional shoulder routing.	`phytools::phylo.to.map`; compositions of `ggtree`, `sf`, and map layers; `Nextstrain Auspice`'s linked tree and map panels.
`plot_tangled_chain`	Plot a sequence of trees with lines connecting corresponding tips between each adjacent pair.	`phytools::cophylo` and `ape::cophyloplot` handle two trees; `ggtree::ggdoubletree` provides paired-tree support. `baltic` generalises the idea to a chain of trees.

References

1. J. D. Hunter, "Matplotlib: A 2D Graphics Environment," *Computing in Science Engineering*, vol. 9, no. 3, pp. 90–95, May 2007.
2. Y. Park *et al.*, "Epidemiological inference for emerging viruses using segregating sites," *Nature Communications*, vol. 14, no. 1, p. 3105, May 2023, number: 1. [Online]. Available: https://www.nature.com/articles/s41467-023-38809-7
3. I. B. Olawoye *et al.*, "Emergence and spread of two SARS-CoV-2 variants of interest in Nigeria," *Nature Communications*, vol. 14, no. 1, p. 811, Feb. 2023, number: 1. [Online]. Available: https://www.nature.com/articles/s41467-023-36449-5
4. E. Normandin *et al.*, "High-depth sequencing characterization of viral dynamics across tissues in fatal COVID-19 reveals compartmentalized infection," *Nature Communications*, vol. 14, no. 1, p. 574, Feb. 2023, number: 1. [Online]. Available: https://www.nature.com/articles/s41467-022-34256-y
5. A. Tsitsiklis *et al.*, "Lower respiratory tract infections in children requiring mechanical ventilation: a multicentre prospective surveillance study incorporating airway metagenomics," *The Lancet Microbe*, vol. 0, no. 0, Mar. 2022. [Online]. Available: https://www.thelancet.com/journals/lanmic/article/PIIS2666-5247(21)00304-9/fulltext
6. G. Dudas *et al.*, "Emergence and spread of SARS-CoV-2 lineage B.1.620 with variant of concern-like mutations and deletions," *Nature Communications*, vol. 12, no. 1, p. 5769, Oct. 2021, bandiera_abtest: a Cc_license_type: cc_by Cg_type: Nature Research Journals Number: 1 Primary_atype: Research Subject_term: Molecular evolution;Pathogens;Phylogenetics;Viral infection Subject_term_id: molecular-evolution;pathogens;phylogenetics;viral-infection. [Online]. Available: https://www.nature.com/articles/s41467-021-26055-8
7. S. K. Gire *et al.*, "Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak," *Science*, vol. 345, no. 6202, pp. 1369–1372, Sep. 2014. [Online]. Available: https://www.science.org/doi/full/10.1126/science.1259657
8. J. Quick *et al.*, "Real-time, portable genome sequencing for Ebola surveillance," *Nature*, vol. 530, no. 7589, pp. 228–232, Feb. 2016, bandiera_abtest: a Cg_type: Nature Research Journals Number: 7589 Primary_atype: Research Subject_term: Microbial genetics;Microbiology;Molecular evolution;Next-generation sequencing;Viral infection Subject_term_id: microbial-genetics;microbiology;molecular-evolution;next-generation-sequencing;viral-infection. [Online]. Available: https://www.nature.com/articles/nature16996
9. M. W. Carroll *et al.*, "Temporal and spatial analysis of the 2014–2015 Ebola virus outbreak in West Africa," *Nature*, vol. 524, no. 7563, pp. 97–101, Aug. 2015, bandiera_abtest: a Cc_license_type: cc_y Cg_type: Nature Research Journals Number: 7563 Primary_atype: Research Subject_term: Ebola virus Subject_term_id: ebola-virus. [Online]. Available: https://www.nature.com/articles/nature14594
10. G. Dudas *et al.*, "Virus genomes reveal factors that spread and sustained the Ebola epidemic," *Nature*, vol. 544, no. 7650, pp. 309–315, Apr. 2017, bandiera_abtest: a Cg_type: Nature Research Journals Number: 7650 Primary_atype: Research Subject_term: Epidemiology;Phylogenetics;Statistical methods;Viral infection Subject_term_id: epidemiology;phylogenetics;statistical-methods;viral-infection. [Online]. Available: https://www.nature.com/articles/nature22040
11. D. J. Park *et al.*, "Ebola Virus Epidemiology, Transmission, and Evolution during Seven Months in Sierra Leone," *Cell*, vol. 161, no. 7, pp. 1516–1526, Jun. 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S009286741500690X
12. N. D. Grubaugh *et al.*, "Genomic epidemiology reveals multiple introductions of Zika virus into the United States," *Nature*, vol. 546, no. 7658, pp. 401–405, Jun. 2017, bandiera_abtest: a Cg_type: Nature Research Journals Number: 7658 Primary_atype: Research Subject_term: Viral epidemiology;Viral genetics Subject_term_id: viral-epidemiology;viral-genetics. [Online]. Available: https://www.nature.com/articles/nature22400
13. E. C. Holmes *et al.*, "The evolution of Ebola virus: Insights from the 2013–2016 epidemic," *Nature*, vol. 538, no. 7624, pp. 193–200, Oct. 2016, bandiera_abtest: a Cg_type: Nature Research Journals Number: 7624 Primary_atype: Reviews Subject_term: Ebola virus;Evolutionary biology;Molecular evolution Subject_term_id: ebola-virus;evolutionary-biology;molecular-evolution. [Online]. Available: https://www.nature.com/articles/nature19790
14. G. Dudas *et al.*, "MERS-CoV spillover at the camel-human interface," *eLife*, vol. 7, p. e31257, Jan. 2018. [Online]. Available: https://doi.org/10.7554/eLife.31257
15. D. K. W. Chu *et al.*, "MERS coronaviruses from camels in Africa exhibit region-dependent genetic diversity," *Proceedings of the National Academy of Sciences*, vol. 115, no. 12, pp. 3144–3149, Mar. 2018. [Online]. Available: https://www.pnas.org/content/115/12/3144
16. N. S. Lewis *et al.*, "The global antigenic diversity of swine influenza A viruses," *eLife*, vol. 5, p. e12217, Apr. 2016. [Online]. Available: https://doi.org/10.7554/eLife.12217
17. J. T. Ladner *et al.*, "Evolution and Spread of Ebola Virus in Liberia, 2014–2015," *Cell Host & Microbe*, vol. 18, no. 6, pp. 659–669, Dec. 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S193131281500462X
18. G. Dudas *et al.*, "Reassortment between Influenza B Lineages and the Emergence of a Coadapted PB1–PB2–HA Gene Complex," *Molecular Biology and Evolution*, vol. 32, no. 1, pp. 162–172, Jan. 2015. [Online]. Available: https://doi.org/10.1093/molbev/msu287
19. G. Dudas *et al.*, "MERS-CoV recombination: implications about the reservoir and potential for adaptation," *Virus Evolution*, vol. 2, no. 1, Jan. 2016. [Online]. Available: https://doi.org/10.1093/ve/vev023
20. G. Dudas *et al.*, "The ability of single genes vs full genomes to resolve time and space in outbreak analysis," *BMC Evolutionary Biology*, vol. 19, no. 1, p. 232, Dec. 2019. [Online]. Available: https://doi.org/10.1186/s12862-019-1567-0
21. S. L. M. Whitmer *et al.*, "Preliminary Evaluation of the Effect of Investigational Ebola Virus Disease Treatments on Viral Genome Sequences," *The Journal of Infectious Diseases*, vol. 214, no. suppl_3, pp. S333–S341, Oct. 2016. [Online]. Available: https://doi.org/10.1093/infdis/jiw177
22. N. F. Müller *et al.*, "Inferring time-dependent migration and coalescence patterns from genetic sequence and predictor data in structured populations," *Virus*

Evolution, vol. 5, no. 2, Jul. 2019. [Online]. Available: https://doi.org/10.1093/ve/vez030

23. N. F. Müller *et al.*, "Bayesian inference of reassortment networks reveals fitness benefits of reassortment in human influenza viruses," *Proceedings of the National Academy of Sciences*, vol. 117, no. 29, pp. 17 104–17 111, Jul. 2020. [Online]. Available: https://www.pnas.org/content/117/29/17104
24. A. Rambaut *et al.*, "Comment on "Mutation rate and genotype variation of Ebola virus from Mali case sequences"," *Science*, vol. 353, no. 6300, pp. 658–658, Aug. 2016. [Online]. Available: https://www.science.org/doi/full/10.1126/science.aaf3823
25. J. Batson *et al.*, "Single mosquito metatranscriptomics identifies vectors, emerging pathogens and reservoirs in one assay," *eLife*, vol. 10, p. e68353, Apr. 2021. [Online]. Available: https://doi.org/10.7554/eLife.68353
26. N. Chazot *et al.*, "Conserved ancestral tropical niche but different continental histories explain the latitudinal diversity gradient in brush-footed butterflies," *Nature Communications*, vol. 12, no. 1, p. 5717, Sep. 2021. [Online]. Available: https://www.nature.com/articles/s41467-021-25906-8
27. C. A. Arze *et al.*, "Global genome analysis reveals a vast and dynamic anellovirus landscape within the human virome," *Cell Host & Microbe*, vol. 29, no. 8, pp. 1305–1315.e6, Aug. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1931312821003024
28. J.-P. Pirnay *et al.*, "Variant Analysis of SARS-CoV-2 Genomes from Belgian Military Personnel Engaged in Overseas Missions and Operations," *Viruses*, vol. 13, no. 7, p. 1359, Jul. 2021, number: 7. [Online]. Available: https://www.mdpi.com/1999-4915/13/7/1359
29. G. Dudas *et al.*, "Polymorphism of genetic ambigrams," *Virus Evolution*, vol. 7, no. 1, Jan. 2021. [Online]. Available: https://doi.org/10.1093/ve/veab038
30. B. A. Petros *et al.*, "Multimodal surveillance of SARS-CoV-2 at a university enables development of a robust outbreak response framework," *Med*, Sep. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666634022004044
31. N. F. Müller *et al.*, "Tracking the horizontal transfer of plasmids in Shigella sonnei and Shigella flexneri using phylogenetics," Oct. 2022, pages: 2022.10.27.514108 Section: New Results. [Online]. Available: https://www.biorxiv.org/content/10.1101/2022.10.27.514108v1
32. A. X. Han *et al.*, "Regional importation and asymmetric within-country spread of SARS-CoV-2 variants of concern in the Netherlands," *eLife*, vol. 11, p. e78770, Sep. 2022. [Online]. Available: https://doi.org/10.7554/eLife.78770
33. E. Parker *et al.*, "Regional connectivity drove bidirectional transmission of SARS-CoV-2 in the Middle East during travel restrictions," *Nature Communications*, vol. 13, no. 1, p. 4784, Aug. 2022, number: 1. [Online]. Available: https://www.nature.com/articles/s41467-022-32536-1
34. R. Xie *et al.*, "Genomic epidemiology of seasonal influenza circulation in China during prolonged border closure from 2020 to 2021," *Virus Evolution*, vol. 8, no. 2, p. veac062, Jul. 2022. [Online]. Available: https://doi.org/10.1093/ve/veac062
35. N. F. Müller *et al.*, "A Bayesian approach to infer recombination patterns in coronaviruses," *Nature Communications*, vol. 13, no. 1, p. 4186, Jul. 2022, number: 1. [Online]. Available: https://www.nature.com/articles/s41467-022-31749-8
36. J. Van Elslande *et al.*, "Two Separate Clusters of SARS-CoV-2 Delta Variant Infections in A Group of 41 Students Travelling from India: An Illustration of the Need for Rigorous Testing and Quarantine," *Viruses*, vol. 14, no. 6, p. 1198, Jun. 2022, number: 6. [Online]. Available: https://www.mdpi.com/1999-4915/14/6/1198
37. M. E. Petrone *et al.*, "Combining genomic and epidemiological data to compare the transmissibility of SARS-CoV-2 variants Alpha and Iota," *Communications Biology*, vol. 5, no. 1, pp. 1–10, May 2022, number: 1. [Online]. Available: https://www.nature.com/articles/s42003-022-03347-3
38. T. Gräf *et al.*, "Phylogenetic-based inference reveals distinct transmission dynamics of SARS-CoV-2 lineages Gamma and P.2 in Brazil," *iScience*, vol. 25, no. 4, p. 104156, Apr. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2589004222004266
39. G. K. Moreno *et al.*, "Distinct patterns of SARS-CoV-2 transmission in two nearby communities in Wisconsin, USA," medRxiv, Tech. Rep., Jul. 2020, type: article. [Online]. Available: https://www.medrxiv.org/content/10.1101/2020.07.09.20149104v1
40. A. Babiker *et al.*, "Unrecognized introductions of SARS-CoV-2 into the US state of Georgia shaped the early epidemic," *Virus Evolution*, vol. 8, no. 1, p. veac011, Jan. 2022. [Online]. Available: https://doi.org/10.1093/ve/veac011
41. K. M. Braun *et al.*, "Transmission of SARS-CoV-2 in domestic cats imposes a narrow bottleneck," *PLOS Pathogens*, vol. 17, no. 2, p. e1009373, Feb. 2021. [Online]. Available: https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1009373
42. M. J. Sanín *et al.*, "Volcanic events coincide with plant dispersal across the Northern Andes," *Global and Planetary Change*, vol. 210, p. 103757, Mar. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0921818122000248
43. L. J. Taylor *et al.*, "Redondovirus Diversity and Evolution on Global, Individual, and Molecular Scales," *Journal of Virology*, vol. 95, no. 21, pp. e00 817–21, Oct. 2021. [Online]. Available: https://journals.asm.org/doi/10.1128/JVI.00817-21
44. M. F. Torres Jimenez *et al.*, "Comparative phylogeography of an ant-plant mutualism: An encounter in the Andes," *Global and Planetary Change*, vol. 205, p. 103598, Oct. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0921818121001831
45. A. Russell *et al.*, "Spatiotemporal Analyses of 2 Co-Circulating SARS-CoV-2 Variants, New York State, USA," *Emerging Infectious Diseases*, vol. 28, no. 3, Mar. 2022. [Online]. Available: https://wwwnc.cdc.gov/eid/article/28/3/21-1972_article
46. T. MacCannell *et al.*, "Genomic Epidemiology and Transmission Dynamics of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) in Congregate Healthcare Facilities in Santa Clara County, California," *Clinical Infectious Diseases*, vol. 74, no. 5, pp. 829–835, Mar. 2022. [Online]. Available: https://doi.org/10.1093/cid/ciab553
47. E. Lasek-Nesselquist *et al.*, "A statewide analysis of SARS-CoV-2 transmission in New York," medRxiv, Tech. Rep., Feb. 2021, type: article. [Online]. Available: https://www.medrxiv.org/content/10.1101/2021.02.20.21251598v1

48. T. Mourier *et al.*, "SARS-CoV-2 genomes from Saudi Arabia implicate nucleocapsid mutations in host response and increased viral load," *Nature Communications*, vol. 13, no. 1, p. 601, Feb. 2022, number: 1. [Online]. Available: https://www.nature.com/articles/s41467-022-28287-8
49. S. Youk *et al.*, "Phylogenetic analysis, molecular changes, and adaptation to chickens of Mexican lineage H5N2 low-pathogenic avian influenza viruses from 1994 to 2019," *Transboundary and Emerging Diseases*, vol. n/a, no. n/a, Feb. 2022, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/tbed.14476. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/tbed.14476
50. M. F. T. Jiménez *et al.*, "Phylogenomics of the Palm Tribe Lepidocaryeae (Calamoideae: Arecaceae) and Description of a New Species of Mauritiella," *Systematic Botany*, vol. 46, no. 3, pp. 863–874, Oct. 2021.
51. U. Stolz *et al.*, "Joint Inference of Migration and Reassortment Patterns for Viruses with Segmented Genomes," *Molecular Biology and Evolution*, vol. 39, no. 1, p. msab342, Jan. 2022. [Online]. Available: https://doi.org/10.1093/molbev/msab342
52. N. F. Müller *et al.*, "Recombination patterns in coronaviruses," bioRxiv, Tech. Rep., Feb. 2022, section: New Results Type: article. [Online]. Available: https://www.biorxiv.org/content/10.1101/2021.04.28.441806v2
53. K. J. Siddle *et al.*, "Transmission from vaccinated individuals in a large SARS-CoV-2 Delta variant outbreak," *Cell*, vol. 185, no. 3, pp. 485–492.e10, Feb. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0092867421014902
54. D. Venkatesh *et al.*, "Avian Influenza Viruses in Wild Birds: Virus Evolution in a Multihost Ecosystem," *Journal of Virology*, vol. 92, no. 15, pp. e00 433–18, Aug. 2018. [Online]. Available: https://journals.asm.org/doi/10.1128/JVI.00433-18
55. K. M. Braun *et al.*, "Acute SARS-CoV-2 infections harbor limited within-host diversity and transmit via tight transmission bottlenecks," *PLOS Pathogens*, vol. 17, no. 8, p. e1009849, Aug. 2021. [Online]. Available: https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1009849
56. S. M. Bell *et al.*, "Modern-day SIV viral diversity generated by extensive recombination and cross-species transmission," *PLOS Pathogens*, vol. 13, no. 7, p. e1006466, Jul. 2017. [Online]. Available: https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1006466
57. J. M. Lee *et al.*, "Deep mutational scanning of hemagglutinin helps predict evolutionary fates of human H3N2 influenza variants," *Proceedings of the National Academy of Sciences*, vol. 115, no. 35, pp. E8276–E8285, Aug. 2018. [Online]. Available: https://www.pnas.org/doi/full/10.1073/pnas.1806133115
58. M. Wille *et al.*, "Where do all the subtypes go? Temporal dynamics of H8–H12 influenza A viruses in waterfowl," *Virus Evolution*, vol. 4, no. 2, p. vey025, Jul. 2018. [Online]. Available: https://doi.org/10.1093/ve/vey025
59. E. K. Dokubo *et al.*, "Persistence of Ebola virus after the end of widespread transmission in Liberia: an outbreak report," *The Lancet Infectious Diseases*, vol. 18, no. 9, pp. 1015–1024, Sep. 2018. [Online]. Available: https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(18)30417-1/fulltext
60. P. Jansen van Vuren *et al.*, "Phylodynamic Analysis of Ebola Virus Disease Transmission in Sierra Leone," *Viruses*, vol. 11, no. 1, p. 71, Jan. 2019, number: 1. [Online]. Available: https://www.mdpi.com/1999-4915/11/1/71
61. S. M. Bell *et al.*, "Dengue genetic divergence generates within-serotype antigenic variation, but serotypes dominate evolutionary dynamics," *eLife*, vol. 8, p. e42496, Aug. 2019. [Online]. Available: https://doi.org/10.7554/eLife.42496
62. M. J. Poen *et al.*, "Co-circulation of genetically distinct highly pathogenic avian influenza A clade 2.3.4.4 (H5N6) viruses in wild waterfowl and poultry in Europe and East Asia, 2017–18," *Virus Evolution*, vol. 5, no. 1, p. vez004, Jan. 2019. [Online]. Available: https://doi.org/10.1093/ve/vez004
63. T. M. Work *et al.*, "A novel host-adapted strain of Salmonella Typhimurium causes renal disease in olive ridley turtles (Lepidochelys olivacea) in the Pacific," *Scientific Reports*, vol. 9, no. 1, p. 9313, Jun. 2019, number: 1. [Online]. Available: https://www.nature.com/articles/s41598-019-45752-5
64. P. Mbala-Kingebeni *et al.*, "Medical countermeasures during the 2018 Ebola virus disease outbreak in the North Kivu and Ituri Provinces of the Democratic Republic of the Congo: a rapid genomic assessment," *The Lancet Infectious Diseases*, vol. 19, no. 6, pp. 648–657, Jun. 2019. [Online]. Available: https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(19)30118-5/fulltext
65. N. D. Grubaugh *et al.*, "Travel Surveillance and Genomics Uncover a Hidden Zika Outbreak during the Waning Epidemic," *Cell*, vol. 178, no. 5, pp. 1057–1071.e11, Aug. 2019. [Online]. Available: https://www.cell.com/cell/abstract/S0092-8674(19)30783-4
66. K. Theys *et al.*, "Advances in Visualization Tools for Phylogenomic and Phylodynamic Studies of Viral Diseases," *Frontiers in Public Health*, vol. 7, p. 208, Aug. 2019. [Online]. Available: https://www.frontiersin.org/article/10.3389/fpubh.2019.00208
67. M. R. Wiley *et al.*, "Lassa virus circulating in Liberia: a retrospective genomic characterisation," *The Lancet Infectious Diseases*, vol. 19, no. 12, pp. 1371–1378, Dec. 2019. [Online]. Available: https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(19)30486-4/fulltext
68. L. H. Moncla *et al.*, "Quantifying within-host diversity of H5N1 influenza viruses in humans and poultry in Cambodia," *PLOS Pathogens*, vol. 16, no. 1, p. e1008191, Jan. 2020. [Online]. Available: https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1008191
69. M. Suárez-Esquivel *et al.*, "Persistence of Brucella abortus lineages revealed by genomic characterization and phylodynamic analysis," *PLOS Neglected Tropical Diseases*, vol. 14, no. 4, p. e0008235, Apr. 2020. [Online]. Available: https://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0008235
70. D. Venkatesh *et al.*, "Regional Transmission and Reassortment of 2.3.4.4b Highly Pathogenic Avian Influenza (HPAI) Viruses in Bulgarian Poultry 2017/18," *Viruses*, vol. 12, no. 6, p. 605, Jun. 2020, number: 6. [Online]. Available: https://www.mdpi.com/1999-4915/12/6/605
71. B. B. Oude Munnink *et al.*, "Rapid SARS-CoV-2 whole-genome sequencing and analysis for informed

public health decision-making in the Netherlands," *Nature Medicine*, vol. 26, no. 9, pp. 1405–1410, Sep. 2020, number: 9. [Online]. Available: https://www.nature.com/articles/s41591-020-0997-y

72. R. S. Sikkema *et al.*, "COVID-19 in health-care workers in three hospitals in the south of the Netherlands: a cross-sectional study," *The Lancet Infectious Diseases*, vol. 20, no. 11, pp. 1273–1280, Nov. 2020. [Online]. Available: https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(20)30527-2/fulltext
73. H. A. Pawestri *et al.*, "Genetic and antigenic characterization of influenza A/H5N1 viruses isolated from patients in Indonesia, 2008–2015," *Virus Genes*, vol. 56, no. 4, pp. 417–429, Aug. 2020. [Online]. Available: https://doi.org/10.1007/s11262-020-01765-1
74. A. F. Brito *et al.*, "Lying in wait: the resurgence of dengue virus after the Zika epidemic in Brazil," *Nature Communications*, vol. 12, no. 1, p. 2619, May 2021, number: 1. [Online]. Available: https://www.nature.com/articles/s41467-021-22921-7
75. E. Alm *et al.*, "Geographical and temporal distribution of SARS-CoV-2 clades in the WHO European Region, January to June 2020," *Eurosurveillance*, vol. 25, no. 32, p. 2001410, Aug. 2020. [Online]. Available: https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2020.25.32.2001410
76. L. H. Moncla *et al.*, "Repeated introductions and intensive community transmission fueled a mumps virus outbreak in Washington State," *eLife*, vol. 10, p. e66448, Apr. 2021. [Online]. Available: https://doi.org/10.7554/eLife.66448
77. J. T. Ladner *et al.*, "An Early Pandemic Analysis of SARS-CoV-2 Population Structure and Dynamics in Arizona," *mBio*, vol. 11, no. 5, pp. e02 107–20, Sep. 2020. [Online]. Available: https://journals.asm.org/doi/10.1128/mBio.02107-20
78. D. Miller *et al.*, "Full genome viral sequences inform patterns of SARS-CoV-2 spread into and within Israel," *Nature Communications*, vol. 11, no. 1, p. 5518, Nov. 2020, number: 1. [Online]. Available: https://www.nature.com/articles/s41467-020-19248-0
79. G. K. Moreno *et al.*, "Revealing fine-scale spatiotemporal differences in SARS-CoV-2 introduction and spread," *Nature Communications*, vol. 11, no. 1, p. 5558, Nov. 2020, number: 1. [Online]. Available: https://www.nature.com/articles/s41467-020-19346-z
80. E. B. Hodcroft *et al.*, "Emergence in late 2020 of multiple lineages of SARS-CoV-2 Spike protein variants affecting amino acid position 677," medRxiv, Tech. Rep., Feb. 2021, type: article. [Online]. Available: https://www.medrxiv.org/content/10.1101/2021.02.12.21251658v3
81. E. C. Thomson *et al.*, "Circulating SARS-CoV-2 spike N439K variants maintain fitness while evading antibody-mediated immunity," *Cell*, vol. 184, no. 5, pp. 1171–1187.e20, Mar. 2021. [Online]. Available: https://www.cell.com/cell/abstract/S0092-8674(21)00080-5
82. M. Zeller *et al.*, "Emergence of an early SARS-CoV-2 epidemic in the United States," *Cell*, vol. 184, no. 19, pp. 4939–4952.e15, Sep. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0092867421008898
83. T. Alpert *et al.*, "Early introductions and transmission of SARS-CoV-2 variant B.1.1.7 in the United States," *Cell*, vol. 184, no. 10, pp. 2595–2604.e13, May 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0092867421004347
84. N. Muller *et al.*, "Severe Acute Respiratory Syndrome Coronavirus 2 Outbreak Related to a Nightclub, Germany, 2020," *Emerging Infectious Diseases*, vol. 27, no. 2, pp. 645–648, Feb. 2021. [Online]. Available: https://wwwnc.cdc.gov/eid/article/27/2/20-4443_article
85. N. L. Washington *et al.*, "Emergence and rapid transmission of SARS-CoV-2 B.1.1.7 in the United States," *Cell*, vol. 184, no. 10, pp. 2587–2594.e7, May 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0092867421003834
86. A. M. Hassan *et al.*, "Ongoing evolution of Middle East Respiratory Syndrome Coronavirus, Kingdom of Saudi Arabia, 2023-2024," Sep. 2024, pages: 2024.09.12.612455 Section: New Results. [Online]. Available: https://www.biorxiv.org/content/10.1101/2024.09.12.612455v1
87. J. Zeng *et al.*, "Spatiotemporal genotype replacement of H5N8 avian influenza viruses contributed to H5N1 emergence in 2021/2022 panzootic," *Journal of Virology*, vol. 98, no. 3, pp. e01 401–23, Feb. 2024. [Online]. Available: https://journals.asm.org/doi/10.1128/jvi.01401-23
88. B. Dadonaite *et al.*, "Deep mutational scanning of H5 hemagglutinin to inform influenza virus surveillance," *PLOS Biology*, vol. 22, no. 11, p. e3002916, Nov. 2024. [Online]. Available: https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3002916
89. A. N. Loes *et al.*, "High-throughput sequencing-based neutralization assay reveals how repeated vaccinations impact titers to recent human H1N1 influenza strains," *Journal of Virology*, vol. 0, no. 0, pp. e00 689–24, Sep. 2024. [Online]. Available: https://journals.asm.org/doi/10.1128/jvi.00689-24
90. A. Crits-Christoph *et al.*, "Genetic tracing of market wildlife and viruses at the epicenter of the COVID-19 pandemic," *Cell*, vol. 187, no. 19, pp. 5468–5482.e11, Sep. 2024. [Online]. Available: https://www.cell.com/cell/abstract/S0092-8674(24)00901-2
91. V. Hill *et al.*, "A new lineage nomenclature to aid genomic surveillance of dengue virus," *PLOS Biology*, vol. 22, no. 9, p. e3002834, Sep. 2024. [Online]. Available: https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3002834
92. D. M. Maloney *et al.*, "ARTIC RSV amplicon sequencing reveals global RSV genotype dynamics." Aug. 2024, pages: 2024.08.23.609324 Section: New Results. [Online]. Available: https://www.biorxiv.org/content/10.1101/2024.08.23.609324v1
93. S. L. M. Whitmer *et al.*, "Inference of Nipah virus evolution, 1999–2015," *Virus Evolution*, vol. 7, no. 1, p. veaa062, Jan. 2021. [Online]. Available: https://doi.org/10.1093/ve/veaa062
94. B. Dadonaite *et al.*, "Full-spike deep mutational scanning helps predict the evolutionary success of SARS-CoV-2 clades," Nov. 2023, pages: 2023.11.13.566961 Section: New Results. [Online]. Available: https://www.biorxiv.org/content/10.1101/2023.11.13.566961v1
95. C. B. F. Vogels *et al.*, "DengueSeq: a pan-serotype whole genome amplicon sequencing protocol for dengue virus," *BMC Genomics*, vol. 25, no. 1, p. 433, May 2024. [Online]. Available: https://doi.org/10.1186/s12864-024-10350-x
96. J.-Y. Kim *et al.*, "Genomic epidemiology of highly pathogenic avian influenza A (H5N1) virus in wild birds in South Korea during 2021–2022: Changes in viral epidemic patterns," *Virus Evolution*, vol. 10,

no. 1, p. veae014, Jan. 2024. [Online]. Available: https://doi.org/10.1093/ve/veae014

97. A. M. Casto *et al.*, "SARS-CoV-2 diversity and transmission on a university campus across two academic years during the pandemic," Mar. 2024, pages: 2024.02.29.24303285. [Online]. Available: https://www.medrxiv.org/content/10.1101/2024.02.29.24303285v1
98. K. Gangavarapu *et al.*, "Many-core algorithms for high-dimensional gradients on phylogenetic trees," *Bioinformatics*, vol. 40, no. 2, p. btae030, Feb. 2024. [Online]. Available: https://doi.org/10.1093/bioinformatics/btae030
99. B. I. Potter *et al.*, "Contemporary and historical human migration patterns shape hepatitis B virus diversity," *Virus Evolution*, p. veae009, Feb. 2024. [Online]. Available: https://doi.org/10.1093/ve/veae009
100. R. Liang *et al.*, "Epistasis in the receptor-binding domain of contemporary H3N2 viruses that reverted to bind sialylated di-LacNAc repeats," *Cell Reports*, vol. 44, no. 8, p. 116007, Aug. 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2211124725007788
101. E. Parker *et al.*, "Genotypic diversity and dynamic nomenclature of Parechovirus A," bioRxiv, Tech. Rep., Aug. 2020, section: New Results Type: article. [Online]. Available: https://www.biorxiv.org/content/10.1101/2020.08.14.251231v1
102. E. Parker *et al.*, "Genomics reveals zoonotic and sustained human mpox spread in West Africa," *Nature*, pp. 1–9, May 2025. [Online]. Available: https://www.nature.com/articles/s41586-025-09128-2
103. C. Kikawa *et al.*, "High-throughput neutralization measurements correlate strongly with evolutionary success of human influenza strains," *eLife*, vol. 14, Jun. 2025. [Online]. Available: https://elifesciences.org/reviewed-preprints/106811
104. L. M. Mendoza *et al.*, "Antigenic Characterization of H1N1 Influenza Viruses That Circulated During the 2019–2020 Season in Philadelphia, Pennsylvania," *Influenza and Other Respiratory Viruses*, vol. 19, no. 6, p. e70104, 2025, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/irv.70104. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/irv.70104
105. D. Jorgensen *et al.*, "Evolution and transmission dynamics of wild poliovirus in Pakistan and Afghanistan (2012-2023)," *Nature Communications*, vol. 16, no. 1, p. 5170, Jun. 2025. [Online]. Available: https://www.nature.com/articles/s41467-025-60432-x
106. K. Pham *et al.*, "Large-Scale Genomic Analysis of SARS-CoV-2 Omicron BA.5 Emergence, United States," *Emerging Infectious Diseases*, vol. 31, no. 13, pp. 45–56, May 2025.
107. G. K. Moreno *et al.*, "Geospatial and demographic patterns of SARS-CoV-2 spread in Massachusetts from over 130,000 genomes," Apr. 2025, pages: 2025.04.04.25324273. [Online]. Available: https://www.medrxiv.org/content/10.1101/2025.04.04.25324273v1
108. J. E. Pekar *et al.*, "Transmission dynamics of the 2022 mpox epidemic in New York City," *Nature Medicine*, pp. 1–10, Mar. 2025. [Online]. Available: https://www.nature.com/articles/s41591-025-03526-9
109. B. B. Larsen *et al.*, "Functional and antigenic landscape of the Nipah virus receptor-binding protein," *Cell*, vol. 0, no. 0, Mar. 2025. [Online]. Available: https://www.cell.com/cell/abstract/S0092-8674(25)00257-0
110. B. R. Wasik *et al.*, "The evolution and epidemiology of H3N2 canine influenza virus after 20 years in dogs," *Epidemiology & Infection*, vol. 153, p. e47, Jan. 2025. [Online]. Available: https://www.cambridge.org/core/journals/epidemiology-and-infection/article/evolution-and-epidemiology-of-h3n2-canine-influenza-virus-after-20-yea 0743D5B6FC771D5CD4087B4CF50EAE73
111. J. Y. Siegers *et al.*, "Emergence of a Novel Reassortant Clade 2.3.2.1c Avian Influenza A/H5N1 Virus Associated with Human Cases in Cambodia," Feb. 2025, pages: 2024.11.04.24313747. [Online]. Available: https://www.medrxiv.org/content/10.1101/2024.11.04.24313747v2
112. J. E. Pekar *et al.*, "The molecular epidemiology of multiple zoonotic origins of SARS-CoV-2," *Science*, vol. 377, no. 6609, pp. 960–966, Aug. 2022. [Online]. Available: https://www.science.org/doi/full/10.1126/science.abp8337
113. M. I. Paredes *et al.*, "Underdetected dispersal and extensive local transmission drove the 2022 mpox epidemic," *Cell*, vol. 187, no. 6, pp. 1374–1386.e13, Mar. 2024. [Online]. Available: https://www.cell.com/cell/abstract/S0092-8674(24)00124-7
114. N. L. Matteson *et al.*, "Genomic surveillance reveals dynamic shifts in the connectivity of COVID-19 epidemics," *Cell*, vol. 186, no. 26, pp. 5690–5704.e20, Dec. 2023. [Online]. Available: https://www.cell.com/cell/abstract/S0092-8674(23)01278-3
115. Y. Butera *et al.*, "Genomic and transmission dynamics of the 2024 Marburg Virus Outbreak in Rwanda," *Nature Medicine*, pp. 1–1, Dec. 2024. [Online]. Available: https://www.nature.com/articles/s41591-024-03459-9
116. H. S. Bouzidi *et al.*, "Generation and evaluation of protease inhibitor-resistant SARS-CoV-2 strains," *Antiviral Research*, p. 105814, Jan. 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0166354224000226
117. H. N. Oltean *et al.*, "Changing genomic epidemiology of COVID-19 in long-term care facilities during the 2020–2022 pandemic, Washington State," *BMC Public Health*, vol. 24, no. 1, p. 182, Jan. 2024. [Online]. Available: https://doi.org/10.1186/s12889-023-17461-2
118. A. V. Demirev *et al.*, "Exploring the Genetic Diversity and Molecular Evolution of Seoul and Hantaan Orthohantaviruses," *Viruses*, vol. 16, no. 1, p. 105, Jan. 2024, number: 1. [Online]. Available: https://www.mdpi.com/1999-4915/16/1/105
119. M. Farjo *et al.*, "Within-host evolutionary dynamics and tissue compartmentalization during acute SARS-CoV-2 infection," *Journal of Virology*, vol. 0, no. 0, pp. e01 618–23, Jan. 2024. [Online]. Available: https://journals.asm.org/doi/10.1128/jvi.01618-23
120. D. A. Rasmussen *et al.*, "Espalier: Efficient Tree Reconciliation and Ancestral Recombination Graphs Reconstruction Using Maximum Agreement Forests," *Systematic Biology*, vol. 72, no. 5, pp. 1154–1170, Sep. 2023. [Online]. Available: https://doi.org/10.1093/sysbio/syad040
121. Á. O'Toole *et al.*, "APOBEC3 deaminase editing in mpox virus as evidence for sustained human transmission since at least 2016," *Science*, vol. 382, no. 6670, pp. 595–600, Nov. 2023. [Online]. Available: https://www.science.org/doi/10.1126/science.adg8116
122. G. Dudas *et al.*, "Accumulated metagenomic studies reveal recent migration, whole genome evolution, and

undiscovered diversity of orthomyxoviruses," *Journal of Virology*, vol. 97, no. 10, pp. e01 056–23, Oct. 2023. [Online]. Available: https://journals.asm.org/doi/10.1128/jvi.01056-23

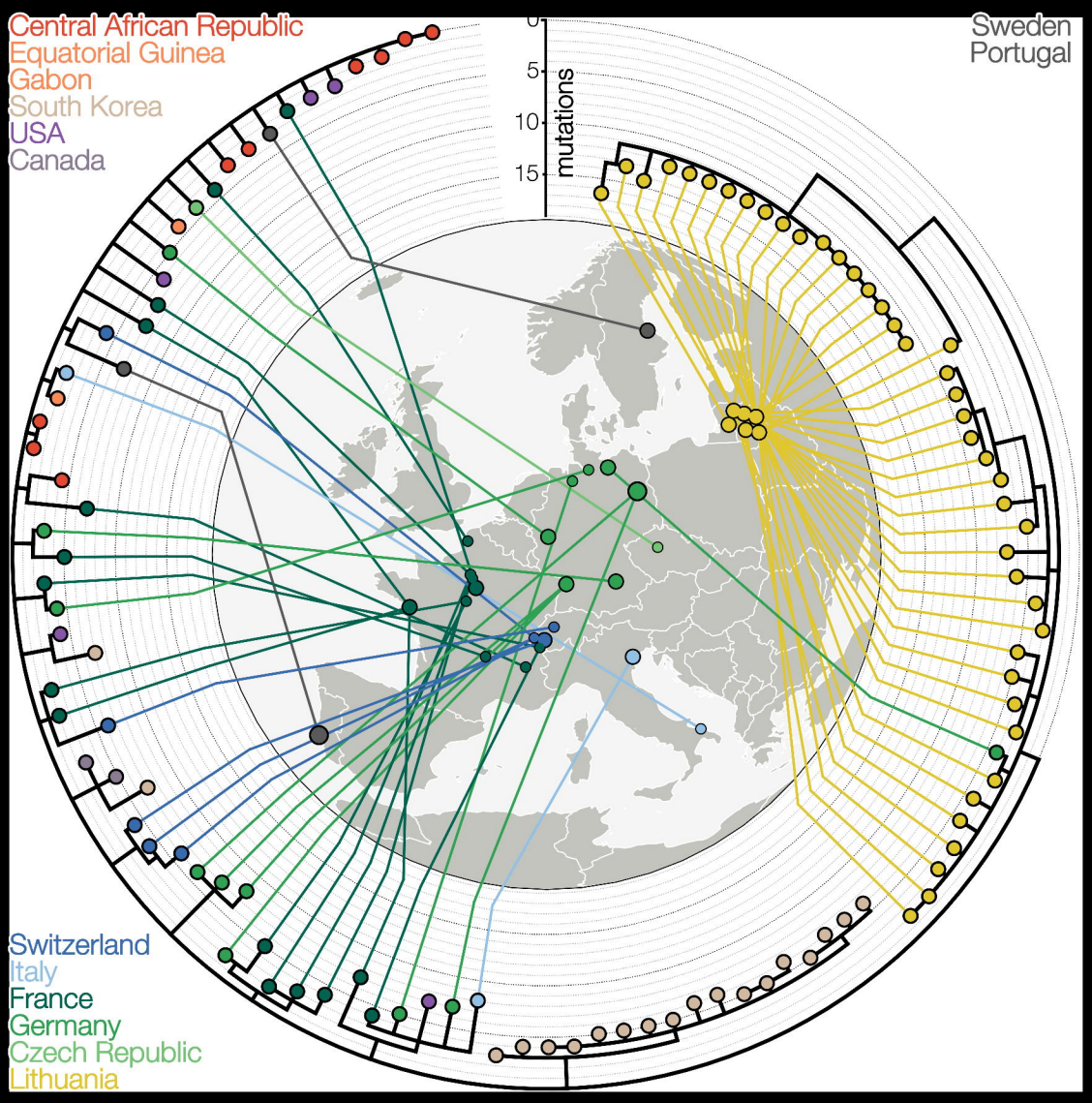
123. M. Scotch *et al.*, "Genomic epidemiology reveals the dominance of Hennepin County in the transmission of SARS-CoV-2 in Minnesota from 2020 to 2022," *mSphere*, vol. 0, no. 0, pp. e00 232–23, Oct. 2023. [Online]. Available: https://journals.asm.org/doi/10.1128/msphere.00232-23
124. R. Xie *et al.*, "The episodic resurgence of highly pathogenic avian influenza H5 virus," *Nature*, pp. 1–8, Oct. 2023. [Online]. Available: https://www.nature.com/articles/s41586-023-06631-2
125. J. Diray-Arce *et al.*, "Multi-omic longitudinal study reveals immune correlates of clinical course among hospitalized COVID-19 patients," *Cell Reports Medicine*, vol. 4, no. 6, Jun. 2023. [Online]. Available: https://www.cell.com/cell-reports-medicine/abstract/S2666-3791(23)00198-2
126. E. Akin *et al.*, "Emergence of an Antigenically Drifted and Reassorted Influenza B Virus at the end of the 2024-25 Influenza Season," Jul. 2025, iSSN: 2692-8205 Pages: 2025.07.24.666632 Section: New Results. [Online]. Available: https://www.biorxiv.org/content/10.1101/2025.07.24.666632v1
127. L. Lyu *et al.*, "Tracing SARS-CoV-2 clusters across local scales using genomic data," *Proceedings of the National Academy of Sciences*, vol. 122, no. 32, p. e2501435122, Aug. 2025, company: National Academy of Sciences Distributor: National Academy of Sciences Institution: National Academy of Sciences Label: National Academy of Sciences. [Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.2501435122
128. B. Jackson *et al.*, "Generation and transmission of interlineage recombinants in the SARS-CoV-2 pandemic," *Cell*, vol. 184, no. 20, pp. 5179–5188.e8, Sep. 2021. [Online]. Available: https://www.cell.com/cell/abstract/S0092-8674(21)00984-3
129. M. Kosuge *et al.*, "Landscape of evolutionary arms races between transposable elements and KRAB-ZFP family," *Scientific Reports*, vol. 14, no. 1, p. 23358, Oct. 2024. [Online]. Available: https://www.nature.com/articles/s41598-024-73752-7
130. M. Brusselmans *et al.*, "On the importance of assessing topological convergence in Bayesian phylogenetic inference," *Virus Evolution*, vol. 10, no. 1, p. veae081, Oct. 2024. [Online]. Available: https://doi.org/10.1093/ve/veae081
131. P. C. Dishuck *et al.*, "Structural variation, selection, and diversification of the *NPIP* gene family from the human pangenome," *Cell Genomics*, p. 100977, Aug. 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666979X25002332
132. C. N. Malaka *et al.*, "Genomic epidemiology of clade Ia monkeypox viruses circulating in the Central African Republic in 2022–24: a retrospective cross-sectional study," *The Lancet Microbe*, vol. 0, no. 0, Sep. 2025. [Online]. Available: https://www.thelancet.com/journals/lanmic/article/PIIS2666-5247(25)00101-6/fulltext
133. P. Gangavarapu *et al.*, "Cryptic spread of SARS-CoV-2 undermined targeted travel restrictions," Sep. 2025, iSSN: 3067-2007 Pages: 2025.08.29.25334701. [Online]. Available: https://www.medrxiv.org/content/10.1101/2025.08.29.25334701v1
134. G. Stoddard *et al.*, "Using genomic epidemiology of SARS-CoV-2 to support contact tracing and public health surveillance in rural Humboldt County, California," *BMC Public Health*, vol. 22, no. 1, p. 456, Mar. 2022. [Online]. Available: https://doi.org/10.1186/s12889-022-12790-0
135. S. Podenas *et al.*, "A New Genus and Two New Species of Short-Palped Crane Flies (Diptera: Limoniidae) from Central Asia," *Diversity*, vol. 15, no. 9, p. 988, Sep. 2023. [Online]. Available: https://www.mdpi.com/1424-2818/15/9/988
136. H. Tegally *et al.*, "Dispersal patterns and influence of air travel during the global expansion of SARS-CoV-2 variants of concern," *Cell*, vol. 186, no. 15, pp. 3277–3290.e16, Jul. 2023. [Online]. Available: https://www.cell.com/cell/abstract/S0092-8674(23)00641-4
137. P. Varilly *et al.*, "Delphy: scalable, near-real-time Bayesian phylogenetics for outbreaks," Mar. 2025, pages: 2025.03.25.645253 Section: New Results. [Online]. Available: https://www.biorxiv.org/content/10.1101/2025.03.25.645253v1
138. A. Sodeinde *et al.*, "Heterogeneity in inhibition of genetically diverse dengue virus strains by Wolbachia," Sep. 2025, iSSN: 2692-8205 Pages: 2025.09.18.677129 Section: New Results. [Online]. Available: https://www.biorxiv.org/content/10.1101/2025.09.18.677129v1
139. J. E. Pekar *et al.*, "The recency and geographical origins of the bat viruses ancestral to SARS-CoV and SARS-CoV-2," *Cell*, vol. 188, no. 12, pp. 3167–3183.e18, Jun. 2025. [Online]. Available: https://www.cell.com/cell/abstract/S0092-8674(25)00353-8
140. G. M. Fiol *et al.*, "Global evolutionary patterns of Yersinia pestis and its spread into Africa," Nov. 2024, pages: 2024.11.26.625443 Section: New Results. [Online]. Available: https://www.biorxiv.org/content/10.1101/2024.11.26.625443v1
141. H. G. Castelán-Sánchez *et al.*, "Phylogenetic discordance can substantially overestimate genomic reassortment in avian influenza virus," Oct. 2025, iSSN: 2692-8205 Pages: 2025.02.24.639875 Section: New Results. [Online]. Available: https://www.biorxiv.org/content/10.1101/2025.02.24.639875v2
142. L. Damodaran *et al.*, "Ecology and spread of the North American H5N1 epizootic," *Nature*, pp. 1–10, Nov. 2025. [Online]. Available: https://www.nature.com/articles/s41586-025-09737-x
143. N. F. Müller *et al.*, "Quantifying plasmid movement in drug-resistant Shigella species using phylodynamic inference," *PLOS Pathogens*, vol. 21, no. 12, p. e1013621, Dec. 2025. [Online]. Available: https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1013621
144. M. Scotch *et al.*, "Campus-based genomic surveillance uncovers early emergence of a future dominant A(H3N2) influenza clade," *iScience*, vol. 28, no. 12, Dec. 2025. [Online]. Available: https://www.cell.com/iscience/abstract/S2589-0042(25)02202-3
145. Á. O'Toole *et al.*, "Human outbreak detection and best practice MPXV analysis and interpretation with squirrel," *Virus Evolution*, vol. 12, no. 1, p. veaf095, Jan. 2026. [Online]. Available: https://doi.org/10.1093/ve/veaf095
146. R. Rivero *et al.*, "Genomic epidemiology of dengue virus 2 and 3 reveals repeated introductions and exportations of several lineages in Colombia," *Cell Reports*, vol. 45, no. 1, p. 116844, Jan. 2026. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S221112472501616X

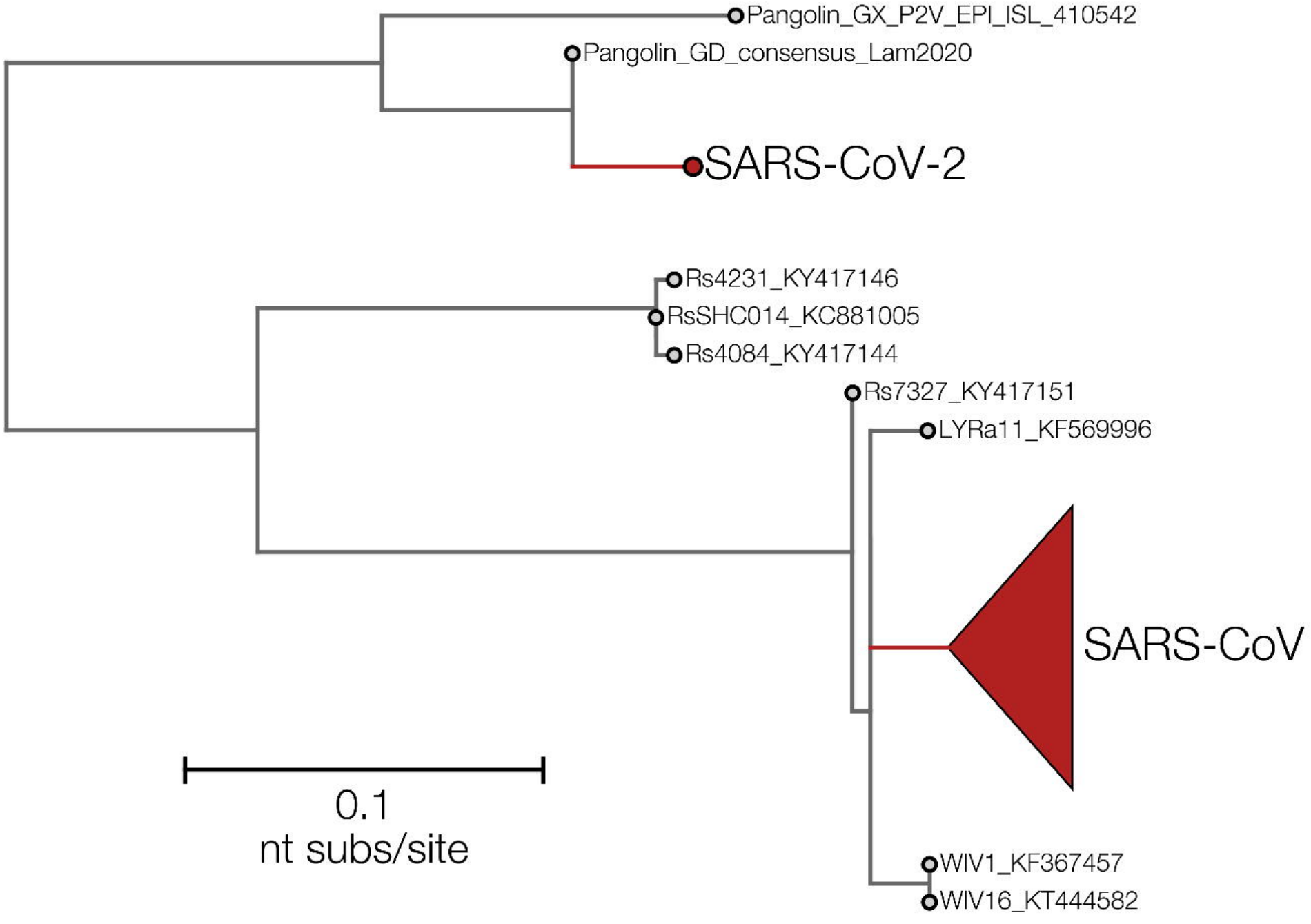
147. M. Chen-Germán *et al.*, "Detection of Oropouche and Punta Toro Virus Infections by Enhanced Surveillance, Panama, 2023–2024 - Volume 32, Number 1—January 2026 - Emerging Infectious Diseases journal - CDC," Jan. 2026. [Online]. Available: https://wwwnc.cdc.gov/eid/article/32/1/25-1224_article
148. S. Freed *et al.*, "Genomic characterization of a severe West Nile Virus transmission season using a single reaction amplicon sequencing approach," *PLOS Neglected Tropical Diseases*, vol. 20, no. 1, p. e0013931, Jan. 2026. [Online]. Available: https://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0013931
149. J. Weber *et al.*, "The receptor binding properties of H5Ny influenza A viruses have evolved to bind to avian-type mucin-like O-glycans," *PLOS Pathogens*, vol. 22, no. 1, p. e1013812, Jan. 2026. [Online]. Available: https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1013812
150. M. Jasnauskaitė *et al.*, "Structure and mechanism of antiphage retron Eco2," *Nature Structural & Molecular Biology*, vol. 33, no. 2, pp. 330–340, Feb. 2026. [Online]. Available: https://www.nature.com/articles/s41594-026-01754-2
151. G. Mboowa *et al.*, "Multicountry genomic analysis underscores regional cholera spread in Africa," *Nature Communications*, Feb. 2026. [Online]. Available: https://www.nature.com/articles/s41467-026-68642-7
152. M. Smičius *et al.*, "Anthropozoonotic spillovers reveal sustained long-term cryptic circulation of SARS-CoV-2 within and between Lithuanian mink farms," *Virus Evolution*, p. veag014, Mar. 2026. [Online]. Available: https://doi.org/10.1093/ve/veag014
153. M. Scotch *et al.*, "Influenza A(H3N2) subclade K (J.2.4.1) viruses associated with a surge at a university health clinic, Arizona, the United States, November to early December 2025," *Eurosurveillance*, vol. 31, no. 7, p. 2600111, Feb. 2026. [Online]. Available: https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2026.31.7.2600111
154. V. L. Sage *et al.*, "Fatal human H3N8 influenza virus has a moderate pandemic risk," *PLOS Pathogens*, vol. 22, no. 3, p. e1013586, Mar. 2026. [Online]. Available: https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1013586
155. M. H. Chitwood *et al.*, "The global phylogeography of rapidly expanding multidrug resistant Ural lineage 4.2 Mycobacterium tuberculosis," *Nature Communications*, Mar. 2026. [Online]. Available: https://www.nature.com/articles/s41467-026-71193-6
156. J. A. Hassard *et al.*, "Bovine H5N1 influenza viruses have adapted to more efficiently use receptors abundant in cattle," Apr. 2026, iSSN: 2692-8205 Pages: 2026.04.02.715584 Section: New Results. [Online]. Available: https://www.biorxiv.org/content/10.64898/2026.04.02.715584v1
157. J. E. Pekar *et al.*, "The emergence and molecular evolution of H5N1 influenza viruses in United States dairy cattle," Apr. 2026, iSSN: 2692-8205 Pages: 2026.03.30.713641 Section: New Results. [Online]. Available: https://www.biorxiv.org/content/10.64898/2026.03.30.713641v1
158. L. M. Huggins *et al.*, "Genomic analysis of Klebsiella aerogenes circulating in New Mexico," *Microbial Genomics*, vol. 12, no. 2, p. 001650, 2026. [Online]. Available: https://www.microbiologyresearch.org/content/journal/mgen/10.1099/mgen.0.001650
159. K. Gangavarapu *et al.*, "Tensor Cores Unlock Efficient and Lower-Energy Massive Parallelization on Phylogenetic Trees," *Systematic Biology*, p. syag017, Mar. 2026. [Online]. Available: https://doi.org/10.1093/sysbio/syag017
160. G. Baele *et al.*, "HIPSTR: highest independent posterior subtree reconstruction in TreeAnnotator X," *Bioinformatics*, vol. 41, no. 10, p. btaf488, Oct. 2025. [Online]. Available: https://doi.org/10.1093/bioinformatics/btaf488
161. G. K. Moreno *et al.*, "Genomic surveillance reveals age-structured SARS-CoV-2 transmission across demographics and settings," Mar. 2026, iSSN: 3067-2007 Pages: 2025.04.04.25324273. [Online]. Available: https://www.medrxiv.org/content/10.1101/2025.04.04.25324273v3
162. M. I. Paredes *et al.*, "Local-scale phylodynamics reveal differential community impact of SARS-CoV-2 in a metropolitan US county," *PLOS Pathogens*, vol. 20, no. 3, p. e1012117, Mar. 2024. [Online]. Available: https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1012117
163. L. Damodaran *et al.*, "Frequent seasonal reassortment between high and low path viruses drives the diversification of influenza A/H5N1," Apr. 2026, iSSN: 2692-8205 Pages: 2026.04.17.719307 Section: New Results. [Online]. Available: https://www.biorxiv.org/content/10.64898/2026.04.17.719307v1
164. M. I. Paredes *et al.*, "Viral introductions and return to baseline sexual behaviors maintain low-level mpox incidence in Los Angeles," *Nature Communications*, Apr. 2026. [Online]. Available: https://www.nature.com/articles/s41467-026-71993-w
165. E. Bourgikos *et al.*, "Evolutionary history of Jamestown Canyon virus reveals complex multi-vector ecology," *Current Biology*, vol. 0, no. 0, Apr. 2026. [Online]. Available: https://www.cell.com/current-biology/abstract/S0960-9822(26)00396-9
166. S. H. Li *et al.*, "Childhood immunological imprinting of cross-subtype antibodies targeting the hemagglutinin head domain of influenza viruses," *Cell Host & Microbe*, vol. 0, no. 0, Apr. 2026. [Online]. Available: https://www.cell.com/cell-host-microbe/abstract/S1931-3128(26)00160-5
167. A. K. O. Campbell *et al.*, "Genomic epidemiology of the 2025 mpox epidemic in Sierra Leone," *Nature Medicine*, pp. 1–10, May 2026. [Online]. Available: https://www.nature.com/articles/s41591-026-04385-8
168. L. A. Krasilnikova *et al.*, "Genomic epidemiology as a tool for understanding drivers of hepatitis A community outbreaks in Massachusetts and New Hampshire," May 2026, iSSN: 3067-2007 Pages: 2026.05.14.26352933. [Online]. Available: https://www.medrxiv.org/content/10.64898/2026.05.14.26352933v1
169. X. Li *et al.*, "Phylodynamic Analysis of the Global Dispersal and Spatiotemporal Dynamics of HIV-1 Subtype C," Jun. 2026, iSSN: 2692-8205 Pages: 2026.06.03.729861 Section: New Results. [Online]. Available: https://www.biorxiv.org/content/10.64898/2026.06.03.729861v1
170. L. d. Moraes *et al.*, "Better data, better trees: GenBank-GISAID deduplication and source-specific artifact masking in viral genomics," Jun. 2026, iSSN: 2692-8205 Pages: 2026.06.12.731931 Section: New Results. [Online]. Available: https://www.biorxiv.org/content/10.64898/2026.06.12.731931v1

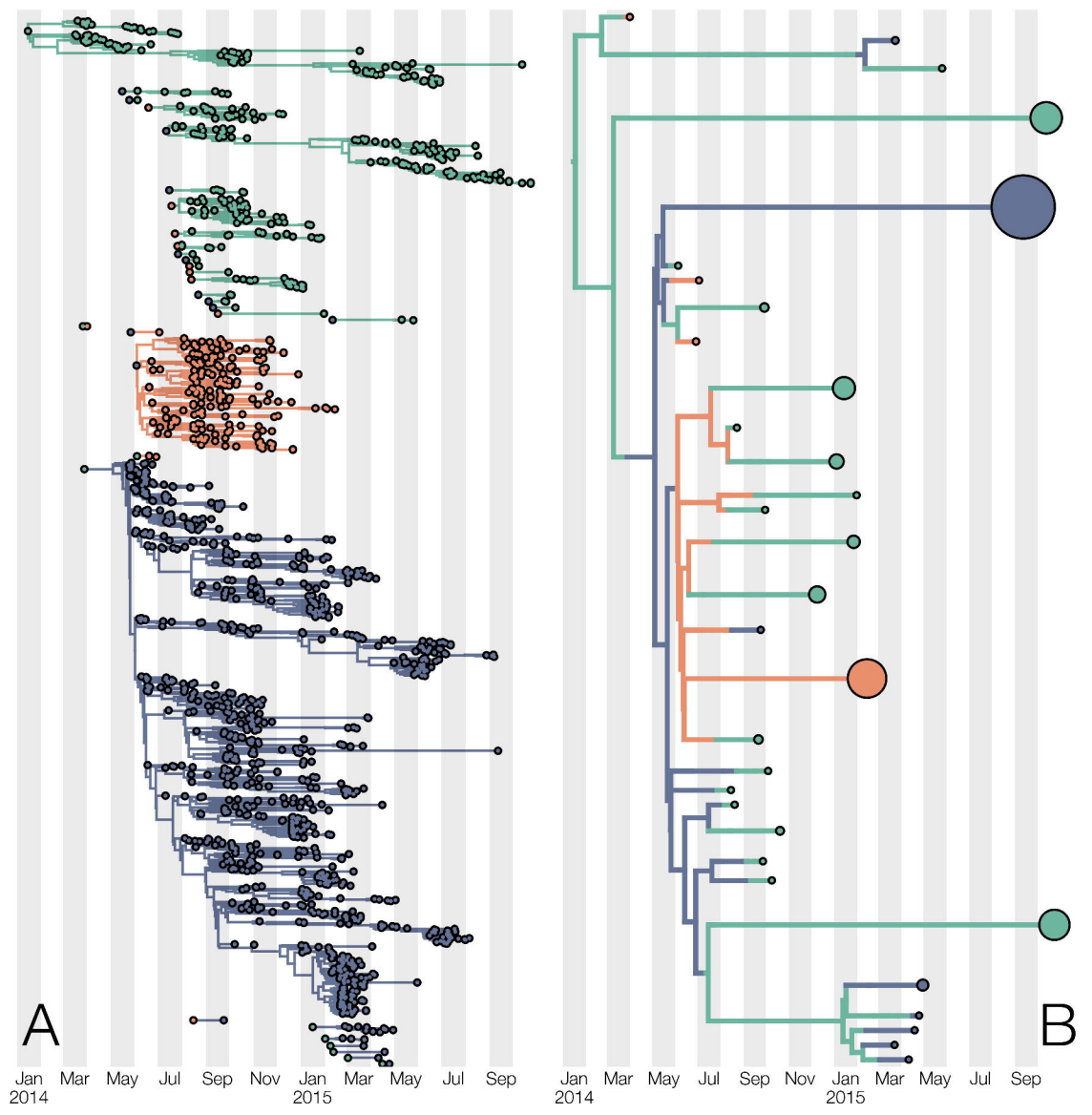
171. A. S. Jaeger *et al.*, "Within- and between-host dynamics of highly pathogenic avian influenza in domestic birds from Pennsylvania farms and live bird markets," Jul. 2026, iSSN: 2692-8205 Pages: 2026.07.17.739193 Section: New Results. [Online]. Available: https://www.biorxiv.org/content/10.64898/2026.07.17.739193v1
172. P. Slavin *et al.*, "The introduction and spread of plague in the Great Lakes region of East Africa, c.1700–1900: a palaeogenetic and social history," *Social History*, vol. 51, no. 2, pp. 107–140, Apr. 2026, _eprint: https://doi.org/10.1080/03071022.2026.2629160. [Online]. Available: https://doi.org/10.1080/03071022.2026.2629160
173. W. E. W. Schravesande *et al.*, "Asynchronous Viral Spread of Two Unrelated Viruses Determines Lettuce Big Vein Disease Symptom Development," *Phytopathology®*, vol. 116, no. 4, pp. 658–667, Apr. 2026. [Online]. Available: https://apsjournals.apsnet.org/doi/abs/10.1094/PHYTO-04-25-0155-R
174. L. Ulloa-Zepeda *et al.*, "Genomic Surveillance of Andes Virus Uncovers Hidden Diversity in Chile," Apr. 2026. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=6620250
175. J. Bahl *et al.*, "Standing HA phenotypic breadth shapes H5N1 cross-host potential," Jul. 2026, iSSN: 2693-5015. [Online]. Available: https://www.researchsquare.com/article/rs-10349790/v1
176. J. Moselen *et al.*, "Dengue virus in Solomon Islands 2023-2025: a whole genome surveillance study," Jul. 2026, iSSN: 3067-2007 Pages: 2026.06.30.26356797. [Online]. Available: https://www.medrxiv.org/content/10.64898/2026.06.30.26356797v1
177. Z. Pella *et al.*, "Detection and Genomic Characterization of Heartland and Bourbon Viruses in Amblyomma americanum ticks from Nebraska," Aug. 2026, iSSN: 3067-2007 Pages: 2026.08.06.26359924. [Online]. Available: https://www.medrxiv.org/content/10.64898/2026.08.06.26359924v1
178. C. A. L. Simonich *et al.*, "Complete definition of how antibodies used to prevent RSV are affected by F mutations," Jul. 2026, iSSN: 2692-8205 Pages: 2026.02.12.705519 Section: New Results. [Online]. Available: https://www.biorxiv.org/content/10.64898/2026.02.12.705519v2
179. S. Gunasekera *et al.*, "Phylogenetic network reconstruction reveals reassortment signatures at segment and genotype levels in human Rotavirus A," Aug. 2026, iSSN: 2692-8205 Pages: 2026.08.11.744215 Section: New Results. [Online]. Available: https://www.biorxiv.org/content/10.64898/2026.08.11.744215v1
180. Z. T. Thornton *et al.*, "antigen-prime: Simulating coupled genetic and antigenic evolution of influenza virus," Aug. 2026, iSSN: 2692-8205 Pages: 2026.01.23.701420 Section: New Results. [Online]. Available: https://www.biorxiv.org/content/10.64898/2026.01.23.701420v2
181. K. Kistler *et al.*, "Seasonal Influenza Viruses Show Distinct Adaptive Dynamics During Growth in Chicken Eggs," *Molecular Biology and Evolution*, vol. 42, no. 10, p. msaf227, Oct. 2025. [Online]. Available: https://doi.org/10.1093/molbev/msaf227
182. J. Hadfield *et al.*, "Nextstrain: real-time tracking of pathogen evolution," *Bioinformatics*, vol. 34, no. 23, pp. 4121–4123, Dec. 2018. [Online]. Available: https://doi.org/10.1093/bioinformatics/bty407
183. G. Baele *et al.*, "BEAST X for Bayesian phylogenetic, phylogeographic and phylodynamic inference," *Nature Methods*, vol. 22, no. 8, pp. 1653–1656, Aug. 2025. [Online]. Available: https://www.nature.com/articles/s41592-025-02751-x
184. R. Bouckaert *et al.*, "BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis," *PLOS Computational Biology*, vol. 15, no. 4, p. e1006650, Apr. 2019. [Online]. Available: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1006650
185. T. G. Vaughan, "IcyTree: rapid browser-based visualization for phylogenetic trees and networks," *Bioinformatics*, vol. 33, no. 15, pp. 2392–2394, Aug. 2017. [Online]. Available: https://doi.org/10.1093/bioinformatics/btx155
186. L.-G. Wang *et al.*, "Treeio: An R Package for Phylogenetic Tree Input and Output with Richly Annotated and Associated Data," *Molecular Biology and Evolution*, vol. 37, no. 2, pp. 599–603, Feb. 2020. [Online]. Available: https://doi.org/10.1093/molbev/msz240
187. T. Sanderson, "Taxonium, a web-based tool for exploring large phylogenetic trees," *eLife*, vol. 11, p. e82392, Nov. 2022. [Online]. Available: https://doi.org/10.7554/eLife.82392
188. C. Solís-Lemus *et al.*, "PhyloNetworks: A Package for Phylogenetic Networks," *Molecular Biology and Evolution*, vol. 34, no. 12, pp. 3292–3298, Dec. 2017. [Online]. Available: https://doi.org/10.1093/molbev/msx235
189. T. N. Starr *et al.*, "ACE2 binding is an ancestral and evolvable trait of sarbecoviruses," *Nature*, vol. 603, no. 7903, pp. 913–918, Mar. 2022. [Online]. Available: https://www.nature.com/articles/s41586-022-04464-z
190. L. du Plessis *et al.*, "Establishment and lineage dynamics of the SARS-CoV-2 epidemic in the UK," *Science*, vol. 371, no. 6530, pp. 708–712, Feb. 2021. [Online]. Available: https://www.science.org/doi/10.1126/science.abf2946
191. F. Chevenet *et al.*, "Searching for virus phylotypes," *Bioinformatics*, vol. 29, no. 5, pp. 561–570, Mar. 2013. [Online]. Available: https://doi.org/10.1093/bioinformatics/btt010
192. F. Chevenet *et al.*, "EvoLaps 2: Advanced phylogeographic visualization," *Virus Evolution*, vol. 10, no. 1, p. vead078, Jan. 2024. [Online]. Available: https://doi.org/10.1093/ve/vead078
193. A. Rambaut, "Figtree: A graphical viewer of phylogenetic trees," 2018, available at http://ed.ac.uk. [Online]. Available: http://ed.ac.uk
194. G. Yu *et al.*, "ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data," *Methods in Ecology and Evolution*, vol. 8, no. 1, pp. 28–36, 2017. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.12628
195. Q. Yu *et al.*, "Lineage frequency time series reveal elevated levels of genetic drift in SARS-CoV-2 transmission in England," *PLOS Pathogens*, vol. 20, no. 4, p. e1012090, Apr. 2024. [Online]. Available: https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1012090
196. H. J. Muller, "Some Genetic Aspects of Sex," *The American Naturalist*, vol. 66, no. 703, pp. 118–138, Mar. 1932. [Online]. Available: https://www.journals.uchicago.edu/doi/abs/10.1086/280418
197. C. D. Gatenbee *et al.*, "EvoFreq: visualization of the Evolutionary Frequencies of sequence and model data," *BMC Bioinformatics*, vol. 20, no. 1, p. 710,

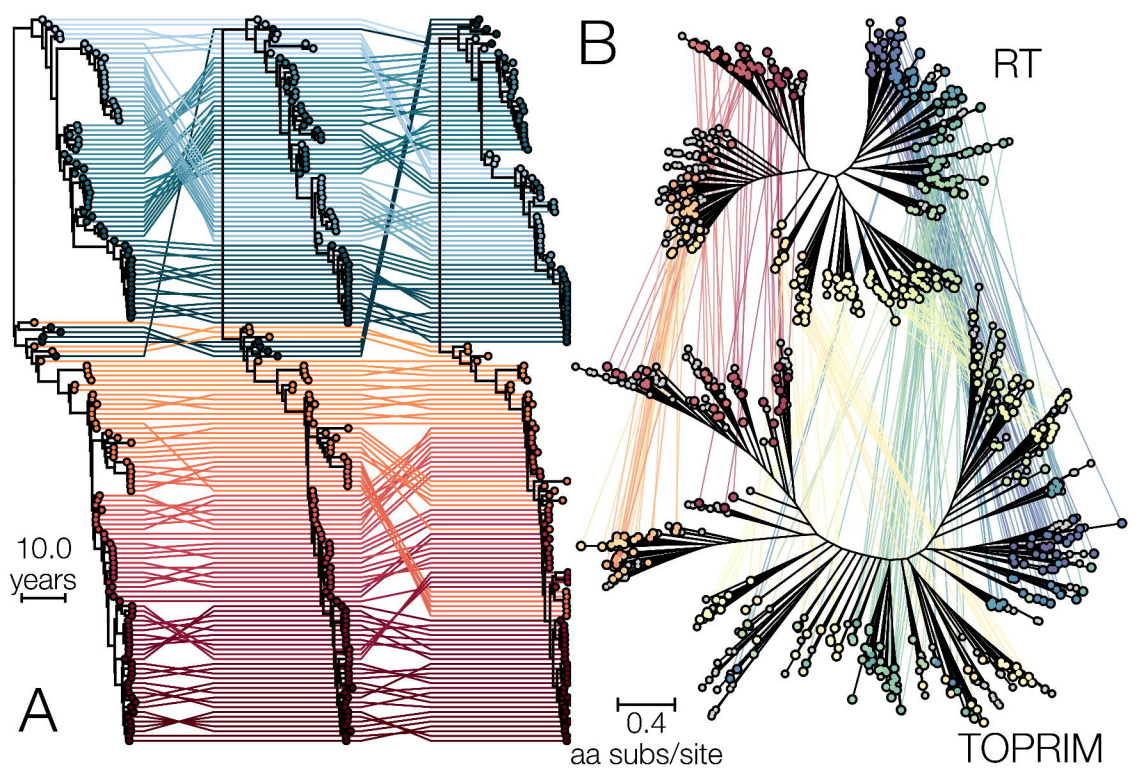
Dec. 2019. [Online]. Available: https://doi.org/10.1186/s12859-019-3173-y

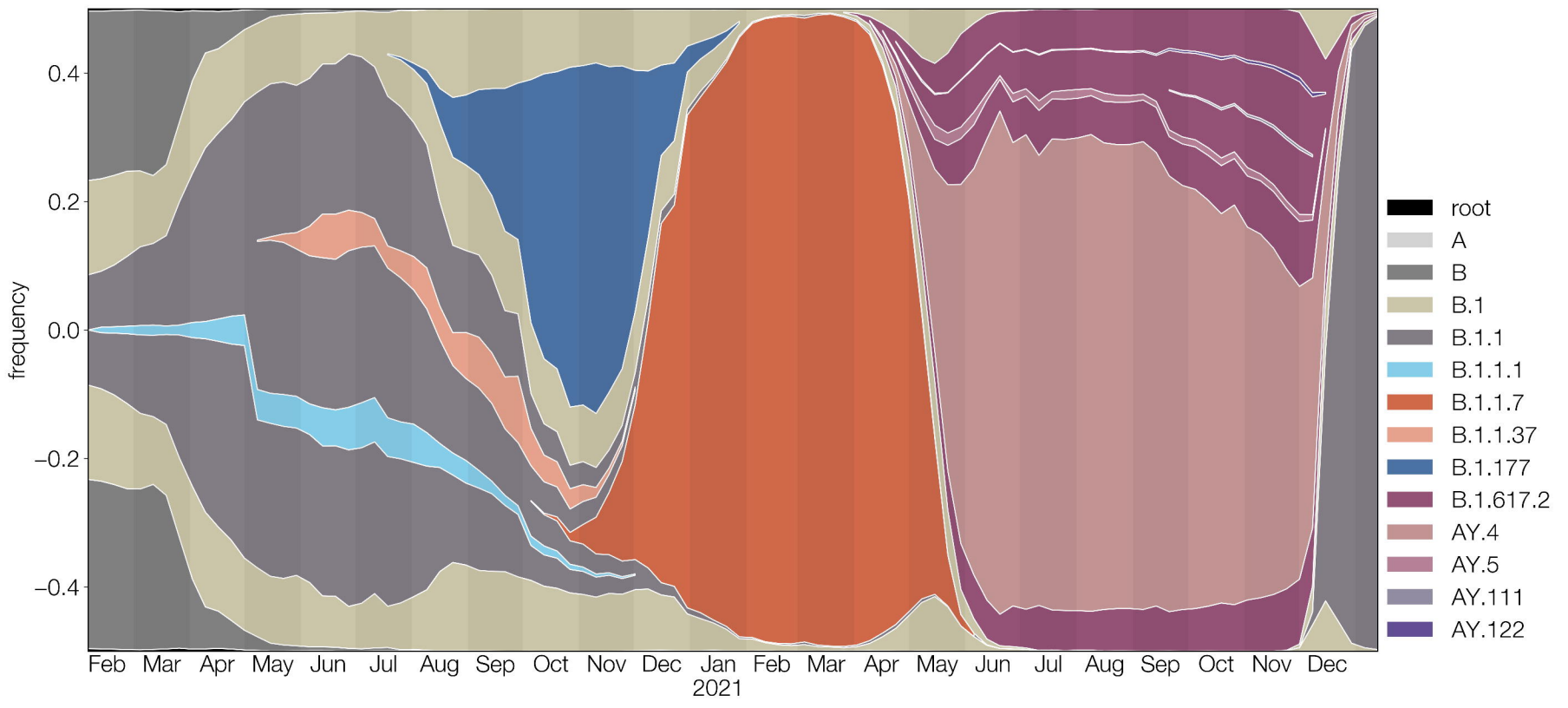
198. A. Streck *et al.*, "SMITH: spatially constrained stochastic model for simulation of intra-tumour heterogeneity," *Bioinformatics*, vol. 39, no. 3, p. btad102, Mar. 2023. [Online]. Available: https://doi.org/10.1093/bioinformatics/btad102
199. Á. O'Toole *et al.*, "Pango lineage designation and assignment using SARS-CoV-2 spike gene nucleotide sequences," *BMC Genomics*, vol. 23, no. 1, p. 121, Feb. 2022. [Online]. Available: https://doi.org/10.1186/s12864-022-08358-2
200. I. Aksamentov *et al.*, "Nextclade: clade assignment, mutation calling and quality control for viral genomes," *Journal of Open Source Software*, vol. 6, no. 67, p. 3773, Nov. 2021. [Online]. Available: https://joss.theoj.org/papers/10.21105/joss.03773
201. A. Stamatakis, "RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies," *Bioinformatics*, vol. 30, no. 9, pp. 1312–1313, May 2014. [Online]. Available: https://doi.org/10.1093/bioinformatics/btu033
202. S. Tavaré, "Some Probabilistic and Statistical Problems in the Analysis of DNA Sequences," *American Mathematical Society: Lectures on Mathematics in the Life Sciences*, vol. 17, pp. 57–86, 1986.
203. Z. Yang, "Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods," *Journal of Molecular Evolution*, vol. 39, no. 3, pp. 306–314, Sep. 1994. [Online]. Available: https://doi.org/10.1007/BF00160154
204. G. Cardona *et al.*, "Extended Newick: it is time for a standard representation of phylogenetic networks," *BMC Bioinformatics*, vol. 9, no. 1, p. 532, Dec. 2008. [Online]. Available: https://doi.org/10.1186/1471-2105-9-532
205. E. Paradis *et al.*, "ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R," *Bioinformatics*, vol. 35, no. 3, pp. 526–528, Feb. 2019. [Online]. Available: https://doi.org/10.1093/bioinformatics/bty633
206. K. Schliep *et al.*, "Intertwining phylogenetic trees and networks," *Methods in Ecology and Evolution*, vol. 8, no. 10, pp. 1212–1220, 2017. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.12760
207. K. Schliep *et al.*, "tanggle: An R package for the visualization of phylogenetic networks," *Applications in Plant Sciences*, vol. 14, no. 3, p. e70060, 2026. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/aps3.70060
208. P. J. A. Cock *et al.*, "Biopython: freely available Python tools for computational molecular biology and bioinformatics," *Bioinformatics*, vol. 25, no. 11, pp. 1422–1423, Jun. 2009. [Online]. Available: https://doi.org/10.1093/bioinformatics/btp163
209. M. A. Moreno *et al.*, "DendroPy 5: a mature Python library for phylogenetic computing," *Journal of Open Source Software*, vol. 9, no. 101, p. 6943, Sep. 2024. [Online]. Available: https://joss.theoj.org/papers/10.21105/joss.06943
210. J. Huerta-Cepas *et al.*, "ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data," *Molecular Biology and Evolution*, vol. 33, no. 6, pp. 1635–1638, Jun. 2016. [Online]. Available: https://doi.org/10.1093/molbev/msw046
211. L. J. Revell, "phytools: an R package for phylogenetic comparative biology (and other things)," *Methods in Ecology and Evolution*, vol. 3, no. 2, pp. 217–223, 2012. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.2041-210X.2011.00169.x
212. P. Sagulenko *et al.*, "TreeTime: Maximum-likelihood phylodynamic analysis," *Virus Evolution*, vol. 4, no. 1, p. vex042, Jan. 2018. [Online]. Available: https://doi.org/10.1093/ve/vex042
213. A. Rambaut *et al.*, "Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen)," *Virus Evolution*, vol. 2, no. 1, p. vew007, Jan. 2016. [Online]. Available: https://doi.org/10.1093/ve/vew007


Central African Republic
Equatorial Guinea
Gabon
South Korea
USA
Canada
0
5
10
15
mutations
Sweden
Portugal
Switzerland
Italy
France
Germany
Czech Republic
Lithuania


Pangolin_GX_P2V_EPI_ISL_410542
Pangolin_GD_consensus_Lam2020
SARS-CoV-2
Rs4231_KY417146
RsSHC014_KC881005
Rs4084_KY417144
Rs7327_KY417151
LYRa11_KF569996
SARS-CoV
WIV1_KF367457
WIV16_KT444582
0.1
nt subs/site


A
B
Jan 2014
Mar
May
Jul
Sep
Nov
Jan 2015
Mar
May
Jul
Sep


B
RT
10.0
years
A
0.4
aa subs/site
TOPRIM


0.4
0.2
0.0
−0.2
−0.4
frequency
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
2021
root
A
B
B.1
B.1.1
B.1.1.1
B.1.1.7
B.1.1.37
B.1.177
B.1.617.2
AY.4
AY.5
AY.111
AY.122