Received: 19 May 2022 | Accepted: 7 October 2022

DOI: 10.1111/2041-210X.14038

Methods in Ecology and Evolution

APPLICATION

TiPS: Rapidly simulating trajectories and phylogenies from compartmental models

Gonché Danesh[1] | Emma Saulnier[1] | Olivier Gascuel[2] | Marc Choisy[3,4] | Samuel Alizon[1,5]

[1]MIVEGEC, CNRS, IRD, Université de Montpellier, Montpellie, France

[2]ISYEB, UMR 7205, MNHN, CNRS, SU, UA, Paris, France

[3]Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, UK

[4]Oxford University Clinical Research Unit, Ho Chi Minh City, Vietnam

[5]Center for Interdisciplinary Research in Biology (CIRB), College de France, CNRS, INSERM, Université PSL, Paris, France

Correspondence
Gonché Danesh
Email: gonche.danesh@gmail.com

Funding information
Fondation pour la Recherche Médicale, Grant/Award Number: ECO20170637560

Handling Editor: Samantha Price

Abstract

1. Stochastic population dynamics simulations are essential for many ecological and epidemiological studies to generate time series and genealogies that capture the relatedness between individuals. Many software packages allow one to simulate phylogenetic trees but these tend to suffer from one or two major limitations. First, the underlying population dynamics model is often simplistic (e.g. constant population size or exponential growth). Second, the software packages are not appropriate to simulate a large number of trees.
2. We introduce TiPS, an R package to generate trajectories and phylogenetic trees associated with a compartmental model. Trajectories are simulated using Gillespie's exact or approximate stochastic simulation algorithm, or a newly proposed mixed version of the two. Phylogenetic trees are simulated from a trajectory under a backwards-in-time approach (i.e. coalescent). TiPS is based on the Rcpp package, allowing to combine the flexibility of R for model definition and the speed of C++ for simulations execution. The model is defined in R with a set of reactions, which allow capturing heterogeneity in life cycles or any sort of population structure. TiPS converts the model into C++ code and compiles it into a simulator that is interfaced in R via a function.
3. TiPS is flexible, easy-to-use and available on the CRAN at https://cran.r-project.org/package=TiPS. Plus, benchmarking analyses show that TiPS is faster than existing packages.
4. This package is particularly well suited for population genetics and phylodynamics studies that need to generate a large number of phylogenies used for population dynamics studies.

KEYWORDS
compartmental models, phylogenies, population dynamics, R package, stochastic simulations

1 | INTRODUCTION

Stochastic simulations of population dynamics are routinely used in ecology and epidemiology to generate trajectories (i.e. time series of population sizes) and genealogies that capture the relatedness between individuals (Keeling & Rohani, 2008; Lenormand et al., 2009; Otto & Day, 2007). The increasing amount of genetic data is fuelling interest in linking population dynamics and genealogies because the former can leave footprints in individuals' genomes (Frost et al., 2015; Grenfell et al., 2004; Volz et al., 2013). Such phylodynamics studies involve computer-intensive methods that can require the simulation of many trajectories and genealogies (Gascuel et al., 2015; Ratmann et al., 2012; Saulnier et al., 2017).

A common method to simulate population dynamics trajectories is Gillespie's exact stochastic simulation algorithm (SSA) (Gillespie, 1976), which is rooted in probability theory (Kurtz, 1970). In the R software environment, it is implemented in packages such as GillespieSSA (Pineda-Krch, 2008), adaptivetau (Johnson, 2014) or epimdr (Bjørnstad, 2018). The computational speed of these software packages is facilitated by the fact that they do not keep track of the history of the process, that is, the trajectory. Conversely, in this same environment, geiger (Pennell et al., 2014), phytools (Revell, 2012), ape (Paradis & Schliep, 2019), and TreeSim can simulate phylogenies. However, the underlying model is often very simple, for example birth-death model. Also in R, some packages such as nosoi (Lequime et al., 2020) allow the user to implement detailed agent-based models but their computational time is slow and the outputs are difficult to compare to classical compartmental models based on differential equations.

However, some software packages can simulate both trajectories and genealogies. In R, rcolgem, which was updated to phydynR (Volz, 2012), combines an Euler-Maruyama integration and the structured coalescent to allow the user to rapidly simulate phylogenies from any compartmental model. This will be our main reference in the following in terms of accuracy and computational speed. Another exception is the software package MASTER (Vaughan & Drummond, 2013) in the BEAST2 platform (Bouckaert et al., 2014). Although MASTER is a useful tool to simulate both time series and genealogies, the specification of the model of interest in the XML language is not as intuitive as the packages of the R environment. Furthermore, although MASTER is one of the fastest options to simulate a few phylogenies (because it does not need to compile the code), its execution time quickly becomes limiting when simulating thousands (or millions) of phylogenies.

We introduce TiPS, a flexible and easy-to-use R package to rapidly simulate population trajectories and phylogenies using a backwards-in-time, that is coalescent, process with either predefined sampling dates or a stochastic sampling scheme. We also introduce a new approximate version of the Gillespie algorithm to increase the calculation speed. A brief benchmarking analysis shows that TiPS is faster than adaptivetau to simulate trajectories, especially for large populations (Figure 3a). It is also at least one order of magnitude faster than phydynR to simulate phylogenies (Figure 3b).

2 | MATERIALS AND METHODS

2.1 | Package overview

TiPS generates two types of stochastic simulation output: population dynamics trajectories and phylogenies. These are obtained using a continuous-time individual-based model defined in R as a system of reactions. The model is first transcribed in C++ and then compiled, before being linked back to a simulation function in R thanks to the Rcpp package (Eddelbuettel & Francois, 2011). The general structure of the pipeline is illustrated in Figure 1.

The TiPS package is available on the CRAN at https://cran.r-project.org/web/packages/TiPS/index.html and is maintained on GitLab at https://gitlab.in2p3.fr/ete/tips/.

2.2 | Model description

TiPS simulates trajectories from a user-specified compartmental model. These models divide the population (animals, cells, *etc*) into distinct categories (geographic, clinical state, *etc*) or so-called compartments in which the sub-population behaves uniformly. In these models, the population may progress between the different compartments.

Here, we illustrate the use of TiPS with the SIR epidemiological model, where individuals can have three clinical states: susceptible (*S*), infected (*I*) and removed (*R*) (Keeling & Rohani, 2008). The model can be captured with a system of two individual-based reactions:

$$S \xrightarrow{\beta SI} I, \tag{1a}$$

$$I \xrightarrow{\gamma I} R, \tag{1b}$$

where β and γ are the transmission and recovery rates. The rate of occurrence of each reaction is indicated above the transition arrow and the corresponding population-based system of ordinary differential equations (ODE) is shown in the top-right panel of Figure 2.

A simulating function is generated from the individual-based reactions of the model of interest. These are entered by the user as a string vector (top left box of Figure 1).

2.3 | Simulating trajectories

The simulation function takes as arguments a named numeric vector that contains the initial number of individuals in each compartment, a named list with the parameter values, a vector of the time limits of the simulation, and the type of algorithm to use for the simulation. Users can also enter a vector of breakpoints, which allows parameter values to vary over time. These breakpoints are indicated in the time limits vector, and the corresponding parameter values are ordered chronologically in the named list of parameter values.

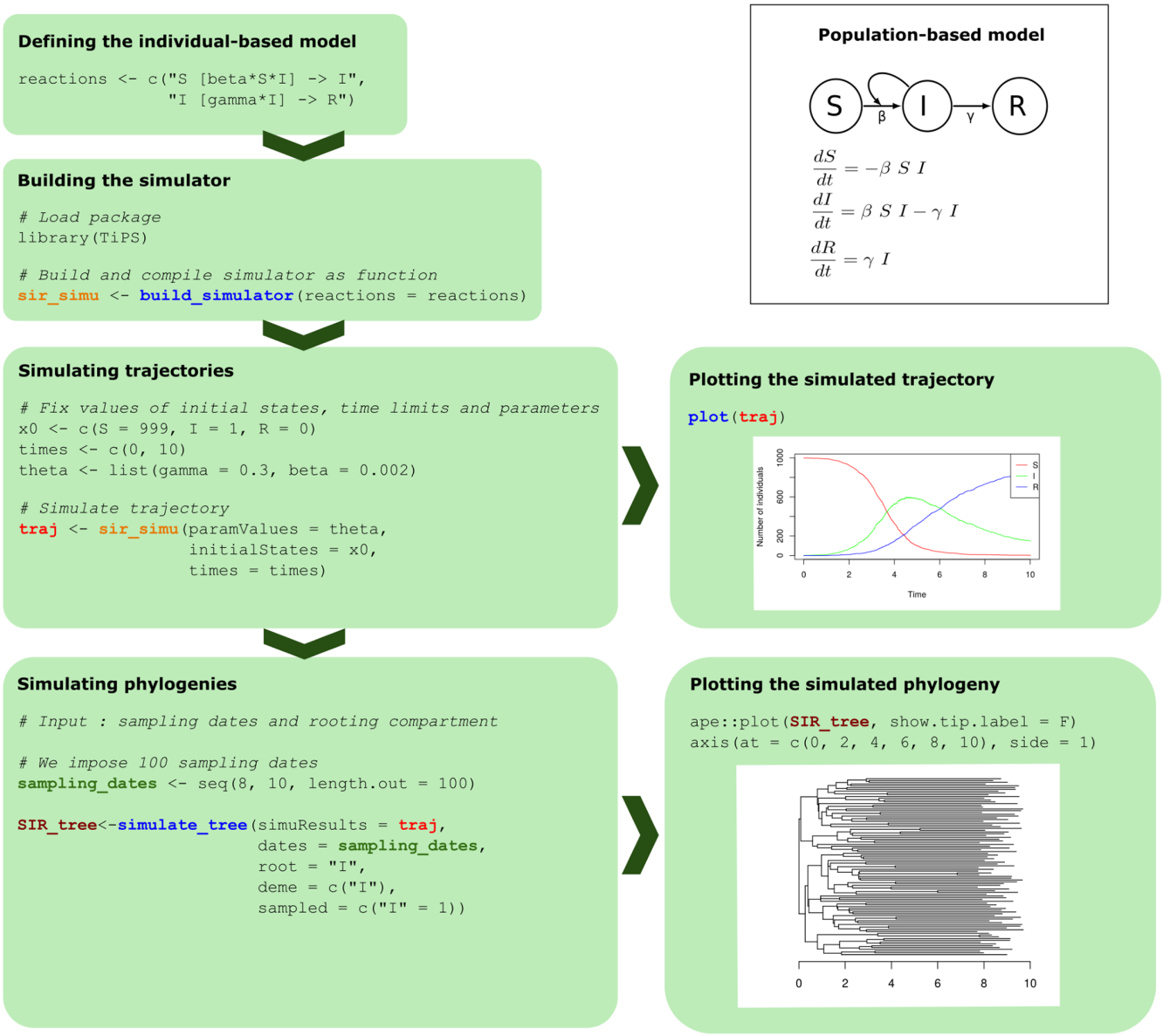

FIGURE 1 General structure of the TiPS pipeline. The equations and outputs correspond to the *SIR* epidemiological model (Keeling & Rohani, 2008). The functions of the R package are in blue. The simulator of trajectories, which is built as a function, is in orange. The variable *traj*, in red, is the output trajectory of class *simutraj*. It can be plotted using our *plot* method. TiPS used the simulated trajectory and 100 sampling dates that we generated (variable *sampling_dates* in green) to simulate a sampled phylogeny. The output simulated phylogeny is a *phylo* class object from the ape R package (Paradis & Schliep, 2019), which can be used for plotting.

Three simulation algorithms are implemented in the Rcpp simulating function:

1. Gillespie's Direct Algorithm (GDA, the default option) (Gillespie, 1976) is an exact algorithm that simulates the time until the next event δ_t by assuming that waiting times are exponentially distributed. A limitation is that its computational complexity scales linearly with the number of simulated events.
2. Gillespie's Tau-Leap Algorithm (GTA) (Gillespie, 2001) is an approximate algorithm that introduces a fixed time step τ during which the number of events of each type is assumed to be Poisson distributed. This algorithm is limited in terms of computation time if the time step is small compared to the rate at which events occur.
3. The Mixed Simulation Algorithm (MSA) is a new algorithm that switches from GDA to GTA depending on the respective values of δ_t and τ. The algorithm switches from GDA to GTA if 10 successive estimations of δ_t are shorter than $\tau / 10$. The algorithm switches back to GDA if the total number of realised events is less than the number of possible events. The MSA shares similarities with the slow-scale stochastic simulation algorithm (Cao et al., 2005) or the adaptive explicit-implicit tau-leaping method for an optimised tau-leap selection (Cao et al., 2007).

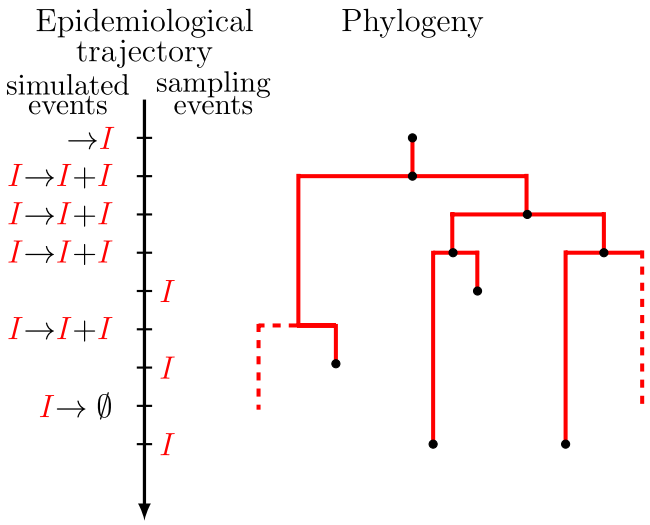

FIGURE 2 Tree simulation for a SIR model. The trajectory shows the series of epidemiological events and sampling events. The phylogeny represents the epidemiological history of the individuals carrying sampled viruses in solid lines. The rest of the history is shown in dashed lines. In this representation, at each transmission event (branching), the donor's pathogen lineage is deviated to the right side and the recipient's (i.e. newly infected individual) pathogen lineage to the left side.

The output of a trajectory simulation is a named list containing the simulated trajectory and some details about the model and the simulation, such as the reactions of the model, the parameter values, the time range, the algorithm used to perform simulations, the time step in case the algorithm is the GTA or the MSA, and, for reproducibility purposes, the random seed used to initialise a pseudorandom number generator. If specified by the user, instead of storing the trajectory as a Rcpp data frame, TiPS can write the simulated trajectory (i.e. at each time step, the time, the size of each compartment, the reaction and the number of reactions simulated) directly into an tab delimited output file.

2.4 | Simulating phylogenies

A phylogeny is the representation of the evolutionary history and relationships between genes, organisms, or groups of organisms. The root of the tree represents the ancestor of all lineages and the leaves represent the most recent descendants of that ancestor. TiPS simulates binary phylogenies rooted in time.

2.4.1 | Infectious disease transmission models

In the context of infectious diseases, under models such as the SIR model, when simulating phylogenies, TiPS traces back the epidemiological history of a sampled pathogen infection, where a forward-in-time transmission event is represented as a coalescence between two lineages under a backward-in-time process, and an end of infection (e.g. caused by death, treatment, or sampling) is represented by a leaf in the phylogeny.

In the SIR model, the pathogen is only present in infected individuals that belong to the *I* compartment, referred to as 'deme' individuals in the deme compartment. The other compartments (i.e. *S* and *R*) are referred to as 'non deme' compartments.

There are four types of reactions in epidemiological models in TiPS:

1. Transmission: this reaction corresponds to the generation of a new deme individual (Equation 1 in the SIR model);
2. Removal: this corresponds to the removal of a deme individual from its deme compartment or the displacement of a deme individual from its deme compartment to a non-deme compartment caused, for example due to treatment or death (Equation 1b the SIR model);
3. Migration: this corresponds to the displacement of a deme individual from its deme compartment to another deme compartment (e.g. from exposed to infectious following the reaction $E \rightarrow I$ in a SEIR model);
4. Sampling: this corresponds to the sampling of deme individuals (*I* in the SIR model) and leads to the end of the infectious period (e.g. through quarantine or change of behaviour).

Note that not all deme individuals can be sampled. For example, in an SEIR model where the *E* and *I* individuals are demes, one can consider that only infectious *I* individuals can be sampled. In other models, the sampled individuals can be only hospitalised individuals, or those in a chronic phase of a disease.

TiPS uses a coalescent approach (Kingman, 1982) to simulate phylogenetic trees based on trajectories, which correspond to a list of dated events (or 'reactions') and sampling dates (e.g. based on observed data). Each node in the simulated phylogenetic tree is associated with a state, that is, a compartment name, and its height, that is, its distance to the root.

TiPS can simulate the phylogeny of the entire trajectory or that of a sampled phylogeny if sampling dates are provided. These dates can either be provided by the user as a vector (bottom left panel of Figure 1) or generated at random during trajectory simulation by adding a sampling reaction in the model. Since parameter values can vary over time, TiPS can reproduce temporal variations in sampling intensity. If sampled individuals can belong to more than one deme compartment, the user can choose between defining the state associated with each sampling event or defining a proportion of sampling events associated with each state. In the last case, TiPS randomly associates a state with each sampling date.

After this preprocessing, the sampling dates are organised as a named list containing a vector of decimal dates (with a column named 'Date') and a vector containing the reactions indicating the state of individuals to sample (a column named 'Reaction'). TiPS then

TABLE 1 Description of demes and reactions for an ecological and an epidemiological model. 'ODEs' stands for ordinary differential equations.

Model	Description	ODEs	Deme compartments	Non-deme compartments	Individual-based reactions	Tree deme individual-based reactions	Tree reaction type
N_1N_2	Logistic growth in two patches model: Two population from the same species That live in different patches (1 and 2) That are linked by migration events.	$\frac{dN_1}{dt} = r_1N_1\left(1-\frac{N_1}{K_1}\right) + \mu(N_2 - N_1)$ $\frac{dN_2}{dt} = r_2N_2\left(1-\frac{N_2}{K_2}\right) + \mu(N_1 - N_2)$	N_1, N_2	-	$\emptyset \xrightarrow{r_1N_1(1-N_1/K_1)} N_1$ $\emptyset \xrightarrow{r_2N_2(1-N_2/K_2)} N_2$ $N_1 \xrightarrow{\mu N_1} N_2$ $N_2 \xrightarrow{\mu N_2} N_1$ $N_1 \xrightarrow{dN_1} \emptyset$ $N_2 \xrightarrow{dN_2} \emptyset$	$N_1 \to N_1 + N_1$ $N_2 \to N_2 + N_2$ $N_1 \to N_2$ $N_2 \to N_1$ $N_1 \to \emptyset$ $N_2 \to \emptyset$	New deme generation New deme generation Migration Migration Removal Removal
SEIR	Epidemiological model: S: Susceptible E: Exposed I: Infectious R: Removed	$\frac{dS}{dt} = -\beta SI$ $\frac{dE}{dt} = \beta SI - \sigma E$ $\frac{dI}{dt} = \sigma E - \gamma I$ $\frac{dR}{dt} = \gamma I$	E, I	S, R	$S \xrightarrow{\beta SI} E$ $E \xrightarrow{\sigma E} I$ $I \xrightarrow{\gamma I} R$	$I \to I + E$ $E \to I$ $I \to \emptyset$	New deme generation Migration Removal

incorporates these pieces of information into the recorded trajectory (which also contains dates and reactions) in chronological order.

The simulation of a phylogeny (sampled or not) starts from the most recent sampling date (or the most recent death event) and progresses through the simulated trajectory backward-in-time.

Each of the four types of reactions previously mentioned can result in a modification in the simulated tree. A forward-in-time transmission event can be represented by a coalescence between two lineages under a backwards-in-time process (or a branching under a forward-in-time process). A sampling event interrupting the transmission of the pathogen is represented as an external node (or leaf) in the phylogeny, and a death event, if observed, is also represented as a leaf. A migration event of an individual from one deme compartment to another will change and update the state of its corresponding lineage.

In some cases, especially with the tau-leap method, more than one event may occur on the same date (e.g. three new transmission events). To determine the number of events that lead to a change in the phylogeny (e.g. a coalescence), we draw a number from a hypergeometric distribution, which is appropriate since it describes the number of successful events (k) when drawing n times (without replacement) from a total population of size N that contains K elements corresponding to a 'success'. For example, when a transmission event in the trajectory is encountered in the phylogeny simulation, the algorithm draws, using the hypergeometric distribution, the number of sampled child lineages and the number of sampled parent lineages to coalesce into one sampled parent lineage. If, at time t, the number of sampled individuals (i.e. sampled nodes) is smaller than the number of simulated transmission events (n) in the trajectory, then the number of coalescence events between two sampled lineages is smaller than n. Therefore, all transmission events at time t may not lead to an observed coalescence in the tree. Similarly, when a migration event is encountered, the algorithm draws the number of sampled lineages associated with a change in state. Figure 2 illustrates how the tree is simulated from the trajectory and further details can be found in Appendix A.

The output simulated phylogeny is an R object of class *phylo* as defined in the ape package (Paradis & Schliep, 2019). The simulated phylogenetic tree can be written in an output file in Newick (by default) or Nexus format if asked and specified by the user.

An illustration of how to simulate a phylogeny can be found in the 'Simulating phylogenies' box in Figure 1 and in the Appendix A for a logistic growth ecological model.

2.4.2 | Ecological population dynamics models

The use of TiPS to simulate phylogenetic trees can also be applied to ecological population dynamics models. As in epidemiological models, there are four types of reactions: generation of new demes, migrations, removals, and samplings.

For example, TiPS can simulate the demographic history and the underlying phylogeny of two populations (N_1 and N_2) belonging to

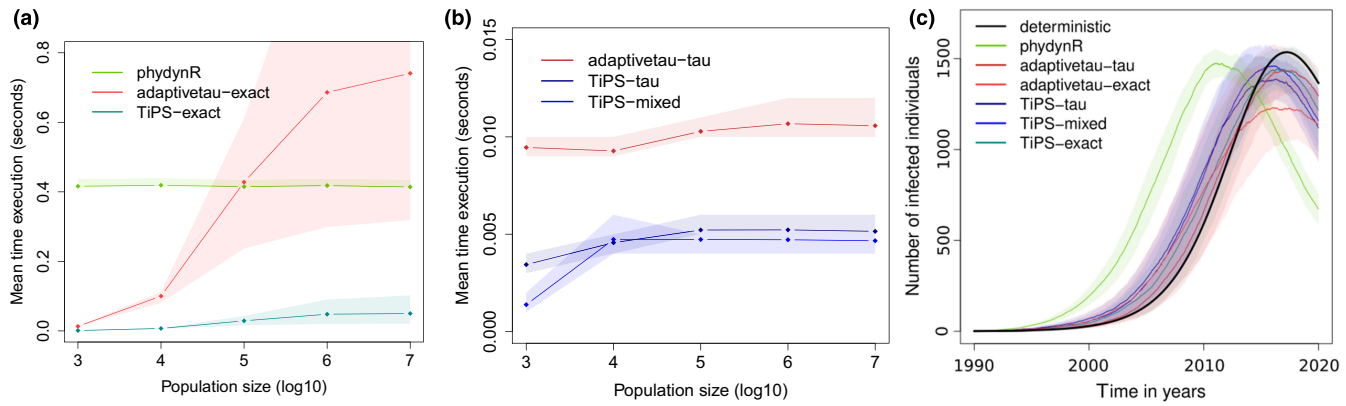

FIGURE 3 Benchmark analyses of trajectories simulations. (a) Median computation speed and 50% confidence intervals (CI) for one simulation using GDA or EMI methods, (b) median computation speed and 50% CI for one simulation using approximating GDA methods and (c) mean resulting trajectories for prevalence time series and the 90% CI. GDA stands for Gillespie's direct method and EMI for Euler-Maruyama integration. In total, 10,000 trajectories simulations were performed under an epidemiological SIR model, for five initial population sizes N, varying from 10^3 to 10^7 and with parameter values $\mathscr{R}_0 = 2, \gamma = 1/3$ and $\beta = \frac{\mathscr{R}_0\gamma}{N}$.

the same species living in different patches under a logistic growth assumption. In such a case, both compartments are deme compartments. The generation of a new deme individual (or 'birth') in a compartment depends on the growth rate and on the population size, and can be represented as coalescence between two lineages in a backward-in-time process. Migration corresponds to the actual migration of a deme individual from one patch to the other and leads to an update of the state of the corresponding lineage in the tree. A removal event corresponds to the death and, if observed, is represented as a leaf. Finally, sampling events correspond to an observation process that interrupts the biological process (e.g. birth) and are also represented as leaves in the tree. These events can correspond to the sterilisation or capture of the sampled individuals.

Table 1 illustrates demes and reaction types for specific ecological and epidemiological models.

2.5 | Benchmarking

To evaluate TiPS's performances, we designed a benchmarking analysis on both modules of the software package, that is, the trajectory and the phylogeny simulators, comparing with existing R packages (adaptivetau and phydynR). Table S1 summarises the main features of the approaches used.

We first evaluated the computational speed and the accuracy of the trajectories (i.e. populating dynamics). For five initial population sizes and three different R packages, we simulated 10,000 trajectories of the epidemiological Susceptible-Infected-Recovered (SIR) model and measured the execution time.

We then compared the time to simulate phylogenies under an SIR model. We varied the sampling proportions to obtain target phylogenies of various sizes (10, 100, 500, 1000, and 1500 leaves). We simulated 1000 phylogenies under each sampling scheme using phydynR and our package.

Furthermore, assuming a more detailed epidemiological model with two host types, as described in (Danesh et al., 2021), we simulated an epidemiological trajectory using the tau-leap algorithm and a complete phylogeny, such that each end of infection event corresponded to a leaf in the tree. The simulation generated a full transmission chain corresponding to a phylogeny with 154,507 leaves. From this complete phylogeny, we generated 10 subtrees by randomly sampling and keeping 1000 leaves for each subtree. We then used TiPS and phydynR to simulate 1000 phylogenies with each package under the same epidemiological model with the same parameter values. Note that the 1000 dates of each target subtree were imposed when simulating these phylogenies using a backwards-in-time approach. To compare the 2000 simulated phylogenies with the target one, we computed 60 summary statistics for each of them using the methods described in (Saulnier et al., 2017).

3 | RESULTS

In Figure 3a, we show the median execution time for one trajectory simulation and the 50% interquartile envelope. TiPS and adaptivetau rely on Gillespie's direct method (GDA), whereas phydynR uses the Euler-Maruyama integration (EMI). As expected, the population size, and hence the number of events per unit of time, increases the execution time for GDA-based packages but not for the EMI-based package. However, TiPS remains faster than the other two software packages for large populations (10^7 individuals). In Figure 3b, we perform the same simulations using approximations of the Gillespie algorithm with fixed time steps. The computational speed of this GTA implemented in TiPS and adaptivetau is comparable and much faster than the GDA. Our new MSA algorithm outperforms existing methods and improves computational speed compared to our GTA, especially for small population sizes.

In Figure 3c, we show the deterministic trajectory and, for each algorithm used, the mean simulated trajectory and its 90% confidence envelope. Among the packages studied, TiPS (in blue) is the one that yields the trajectories that are the closest to the deterministic prediction (in black). Note that there is a temporal shift for all the stochastic simulations, with a more rapid increase in population size compared to the deterministic model. This comes from the fact that stochasticity tends to favours trajectories that spread faster than average because they are less likely to go extinct. This known effect in population genetics models has also been described in epidemiology (Hartfield & Alizon, 2014).

We then analyse the median execution time to simulate phylogenies for each tool and each sampling scheme (Figure 4a). TiPS's median simulation time is several orders of magnitude faster than that of the other method, with a more pronounced advantage for large phylogenies.

Regarding the accuracy of the simulations, we see in Figure 4b that the cross, which indicates the projection of the summary statistics from the target phylogeny, is contained in the envelope containing 90% of the phylogenies simulated using TiPS but not of those simulated using phydynR, and that for three different target phylogenies used represented by distinct colours. We observed the same results for seven other target phylogenies used. This means that the target phylogeny cannot be distinguished from phylogenies simulated using the same model with the same parameters using our package. The discrepancy observed for phydynR could originate from the trajectory, which strongly differs from the deterministic prediction (Figure 3c). Furthermore, TiPS is favoured in this analysis because the same algorithm (Gillespie's Tau-Leap) was used to simulate the target phylogeny and the 1000 sub-trees. Figure S7 shows the distributions of summary statistics computed from the phylogenies simulated using TiPS and phydynR for each analysis using a different target tree.

4 | DISCUSSION

We developed a flexible R package to rapidly simulate trajectories and phylogenies from compartmental models. Its structure allows the user to include several sources of heterogeneity between individuals or between populations, for example different life stages or metapopulation structures. The simulation of the phylogeny relies on the trajectory and involves a coalescent approach.

Our benchmarking analyses show that TiPS is comparable to or outperforms existing R packages in terms of speed when generating numerous trajectories or phylogenies. The same is true for the accuracy of the simulation outputs.

The first asset of this software package is its flexibility since it can readily be used for any compartmental model. Another asset is its computation speed as it can simulate trajectories in a matter of milliseconds on a regular desktop computer. These properties have already been used for infection phylodynamics studies involving approximate Bayesian computing (ABC) (Danesh et al., 2021) or to illustrate the effect of superspreading events (Alizon, 2021).

In the context of infectious diseases, when simulating phylogenetic trees with TiPS, we assume that the time of a transmission event is the same as the time of coalescence in the tree. However, this is a simplification since the time of coalescence in the phylogenetic tree should take place before transmission, during the infection of the 'donnor' host (Ypma et al., 2013). A variety of within-host evolutionary processes may further weaken this assumption. This limitation is common to virus phylodynamics studies and an active line of research (Volz et al., 2017).

Beyond epidemiology, this software package can be used more broadly to simulate population dynamics and associated genealogies. Some future extensions of TiPS will consist in introducing non-Markovian dynamics, simulating multifurcating phylogenies, and implementing other optimised stochastic simulation algorithms (Cao et al., 2005; Cao et al., 2007).

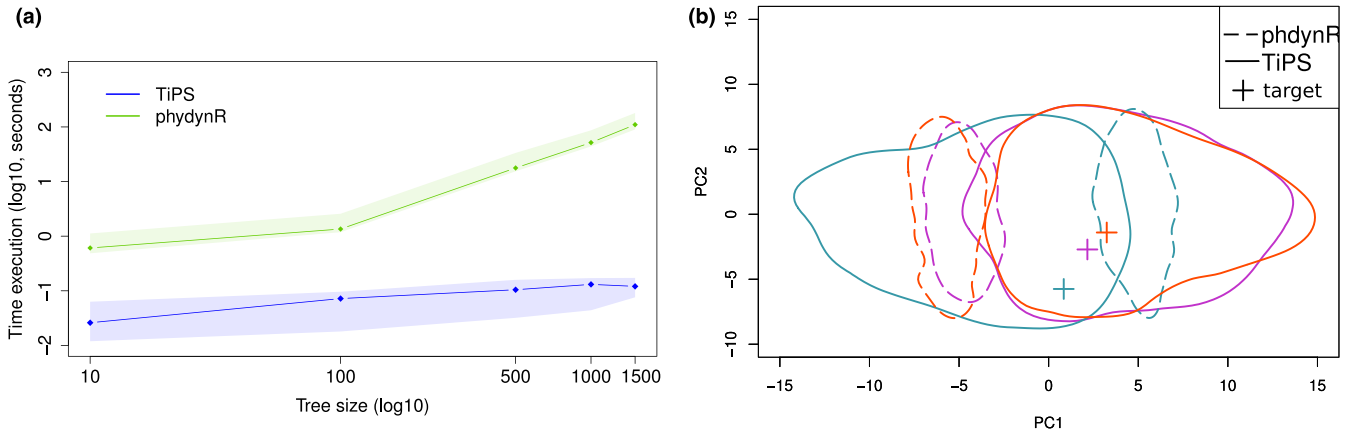

FIGURE 4 Benchmarking analysis of phylogenies simulations. (a) Median computation speed for simulating one phylogeny of a given size and (b) first two axes of a principal component analysis (PCA) of phylogenies summary statistics. In (a), the shaded area shows the 95% confidence intervals. In (b), the cross shows the target phylogeny. The three colours represent three analyses using different target sub-trees. The phylogenies summary statistics used for the PCA are the same as in (Saulnier et al., 2017).

AUTHOR CONTRIBUTIONS

All authors conceived the study. Gonché Danesh and Emma Saulnier wrote the code. Gonché Danesh conducted the analyses. Gonché Danesh and Samuel Alizon led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

ACKNOWLEDGEMENTS

Gonché Danesh was funded by the Fondation pour la Recherche Médicale (grant ECO20170637560). Gonché Danesh and Samuel Alizon acknowledge further support from the CNRS, the IRD and the itrop HPC (South Green Platform) at IRD montpellier, which provided HPC resources that contributed to the results reported here (https://bioinfo.ird.fr/).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/2041-210X.14038.

DATA AVAILABILITY STATEMENT

The scripts and data are available at Zenodo https://doi.org/10.5281/zenodo.7062897 (Danesh et al., 2022).

ORCID

Gonché Danesh https://orcid.org/0000-0001-8688-7666
Emma Saulnier https://orcid.org/0000-0002-8392-0624
Olivier Gascuel https://orcid.org/0000-0002-9412-9723
Marc Choisy https://orcid.org/0000-0002-5187-6390
Samuel Alizon https://orcid.org/0000-0002-0779-9543

REFERENCES

Alizon, S. (2021). Superspreading genomes. *Science*, *371*(6529), 574–575.
Bjørnstad, O. N. (2018). *Epidemics. Models anding R*. Springer.
Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., Suchard, M. A., Rambaut, A., & Drummond, A. J. (2014). BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, *10*(4), 1–6.
Cao, Y., Gillespie, D. T., & Petzold, L. R. (2005). The slow-scale stochastic simulation algorithm. *The Journal of Chemical Physics*, *122*(1), 014116. https://doi.org/10.1063/1.1824902
Cao, Y., Gillespie, D. T., & Petzold, L. R. (2007). Adaptive explicit-implicit tau-leaping method with automatic tau selection. *Journal of Chemical Physics*, *126*(22), 1–10.
Danesh, G., Saulnier, E., Gascuel, O., Choisy, M., & Alizon, S. (2022). Data and scripts from: TiPS: Rapidly simulating trajectories and phylogenies from compartmental models. *Zenodo*, https://doi.org/10.5281/zenodo.7062897
Danesh, G., Virlogeux, V., Ramière, C., Charre, C., Cotte, L., & Alizon, S. (2021). Quantifying transmission dynamics of acute hepatitis C virus infections in a heterogeneous population using sequence data. *PLoS Pathogens*, *17*(9), e1009916.
Eddelbuettel, D., & Francois, R. (2011). Rcpp: Seamless R and C++ integration. *Journal of Statistical Software*, *40*(1), 1–18.
Frost, S. D., Pybus, O. G., Gog, J. R., Viboud, C., Bonhoeffer, S., & Bedford, T. (2015). Eight challenges in phylodynamic inference. *Epidemics*, *10*, 88–92.
Gascuel, F., Ferrière, R., Aguilée, R., & Lambert, A. (2015). How ecology and landscape dynamics shape phylogenetic trees. *Systematic Biology*, *64*(4), 590–607.
Gillespie, D. T. (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, *22*(4), 403–434.
Gillespie, D. T. (2001). Approximate accelerated stochastic simulation of chemically reacting systems. *Journal of Chemical Physics*, *115*(4), 1716–1733.
Grenfell, B. T., Pybus, O. G., Gog, J. R., Wood, J. L. N., Daly, J. M., Mumford, J. A., & Holmes, E. C. (2004). Unifying the epidemiological and evolutionary dynamics of pathogens. *Science*, *303*(5656), 327–332.
Hartfield, M., & Alizon, S. (2014). Epidemiological feedbacks affect evolutionary emergence of pathogens. *The American Naturalist*, *183*(4), E105–E117.
Johnson, P. (2014). *Adaptivetau: Efficient stochastic simulations in R*. R Package Version.
Keeling, M. J., & Rohani, P. (2008). *Modeling infectious diseases in humans and animals*. Princeton University Press.
Kingman, J. F. C. (1982). The coalescent. *Stochastic Processes and their Applications*, *13*(3), 235–248.
Kurtz, T. G. (1970). Solutions of ordinary differential equations as limits of pure jump markov processes. *Journal of Applied Probability*, *7*(1), 49–58.
Lenormand, T., Roze, D., & Rousset, F. (2009). Stochasticity in evolution. *Trends in Ecology & Evolution*, *24*(3), 157–165.
Lequime, S., Bastide, P., Dellicour, S., Lemey, P., & Baele, G. (2020). Nosoi: A stochastic agent-based transmission chain simulation framework in r. *Methods in Ecology and Evolution*, *11*(8), 1002–1007.
Otto, S. P., & Day, T. (2007). *A biologist's guide to mathematical modeling in ecology and evolution*. Princeton University Press.
Paradis, E., & Schliep, K. (2019). Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*, 526–528.
Pennell, M., Eastman, J., Slater, G., Brown, J., Uyeda, J., FitzJohn, R., Alfaro, M. E., & Harmon, L. J. (2014). Geiger v2.0: An expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics*, *30*, 2216–2218. https://doi.org/10.1093/bioinformatics/btu181
Pineda-Krch, M. (2008). GillespieSSA: Implementing the stochastic simulation algorithm in R. *Journal of Statistical Software*, *25*(12), 1–18.
Ratmann, O., Donker, G., Meijer, A., Fraser, C., & Koelle, K. (2012). Phylodynamic inference and model assessment with approximate bayesian computation: Influenza as a case study. *PLoS Computational Biology*, *8*(12), e1002835.
Revell, L. J. (2012). Phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, *3*, 217–223.
Saulnier, E., Gascuel, O., & Alizon, S. (2017). Inferring epidemiological parameters from phylogenies using regression-ABC: A comparative study. *PLoS Computational Biology*, *13*(3), e1005416.
Vaughan, T. G., & Drummond, A. J. (2013). A stochastic simulator of birth-death master equations with application to phylodynamics. *Molecular Biology and Evolution*, *30*(6), 1480–1493.
Volz, E. M. (2012). Complex population dynamics and the coalescent under neutrality. *Genetics*, *190*(1), 187–201.
Volz, E. M., Koelle, K., & Bedford, T. (2013). Viral phylodynamics. *PLoS Computational Biology*, *9*(3), e1002947.
Volz, E. M., Romero-Severson, E., & Leitner, T. (2017). Phylodynamic inference across epidemic scales. *Molecular Biology and Evolution*, *34*(5), 1276–1288.

Ypma, R. J. F., van Ballegooijen, W. M., & Wallinga, J. (2013). Relating phylogenetic trees to transmission trees of infectious disease outbreaks. *Genetics*, *195*(3), 1055–1062.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Danesh, G., Saulnier, E., Gascuel, O., Choisy, M., & Alizon, S. (2023). TiPS: Rapidly simulating trajectories and phylogenies from compartmental models. *Methods in Ecology and Evolution*, *14*, 487–495. https://doi.org/10.1111/2041-210X.14038