Check for updates

Received: 13 May 2025 | Accepted: 10 February 2026

DOI: 10.1111/2041-210x.70278

Methods in Ecology and Evolution

APPLICATION

e3SIM: Epidemiological-ecological-evolutionary simulation framework for genomic epidemiology

Peiyu Xu[1] | Shenni Liang[2,3] | Andrew Hahn[2] | Vivian Zhao[2] | Wai Tung 'Jack' Lo[4] | Benjamin C. Haller[4] | Benjamin Sobkowiak[5] | Melanie H. Chitwood[5] | Caroline Colijn[6] | Ted Cohen[5] | Kyu Y. Rhee[7] | Philipp W. Messer[4] | Martin T. Wells[8] | Andrew G. Clark[1,4] | Jaehee Kim[4]

[1]Department of Molecular Biology & Genetics, Cornell University, Ithaca, New York, USA; [2]Department of Computer Science, Cornell University, Ithaca, New York, USA; [3]Tri-Institutional Computational Biology & Medicine PhD Program, Weill Cornell Medicine, New York, New York, USA; [4]Department of Computational Biology, Cornell University, Ithaca, New York, USA; [5]Department of Epidemiology of Microbial Disease, Yale School of Public Health, New Haven, Connecticut, USA; [6]Department of Mathematics, Simon Fraser University, Burnaby, British Columbia, Canada; [7]Department of Medicine, Weill Cornell Medicine, New York, New York, USA and [8]Department of Statistics and Data Science, Cornell University, Ithaca, New York, USA

Correspondence
Jaehee Kim
Email: jaehee.kim@cornell.edu

Funding information
National Institute of General Medical Sciences, Grant/Award Number: R01HG012473, R35GM152242 and R35GM156957; National Institute of Allergy and Infectious Diseases, Grant/Award Number: P01AI159402

Handling Editor: Matthew Silk

Abstract

1. Infectious disease dynamics result from the complex interplay of epidemiological, ecological and evolutionary (epi-eco-evo) processes. Accurately modelling these coupled processes is crucial for understanding pathogen spread and informing public health strategies. However, existing genomic epidemiology simulators typically assume conditional independence among these processes: generating transmission trees independently of pathogen evolution, and then superimposing neutral mutations onto fixed genealogies without ecological feedback. This simplification fails to capture how pathogen evolution dynamically reshapes epidemic trajectories.
2. We introduce e3SIM, an open-source, agent-based, forward-time simulator for macOS and Linux that explicitly integrates pathogen transmission dynamics, molecular evolution and environmental factors. e3SIM incorporates configurable compartmental models, user-defined host contact networks, customizable pathogen genetic architectures and optional eco-evolutionary features (e.g. within-host dynamics, multi-strain infections). This integration enables realistic modelling of pathogen spread and evolution. Key features include modularity, flexible epidemiological and population-genetic modelling, time-varying environmental factors and a user-friendly graphical interface.
3. We demonstrated e3SIM's capabilities by simulating SARS-CoV-2 and *Mycobacterium tuberculosis* outbreaks. e3SIM captured the emergence and

Peiyu Xu and Shenni Liang contributed equally to this work.

Andrew G. Clark and Jaehee Kim jointly supervised this work.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2026 The Author(s). *Methods in Ecology and Evolution* published by John Wiley & Sons Ltd on behalf of British Ecological Society.

spread of drug-resistant variants under sequential treatments, highlighting how pathogen evolution and environmental variations dynamically reshape epidemic trajectories. We also illustrated how interactions between pathogen transmissibility and host population structures, particularly those involving socially active "superspreaders", strongly influence pathogen lineage expansion and transmission clusters. Runtime profiling demonstrated computational efficiency and scalability.

4. e3SIM provides a powerful tool for simulating infectious disease dynamics through the explicit integration of epi-eco-evo processes, substantially enhancing realism and predictive accuracy in genomic epidemiology. Its modular, user-friendly design supports broad applications across diverse host-pathogen systems, enabling rigorous exploration of scenarios critical to public health.

KEYWORDS
agent-based simulation, epi-eco-evo coupling, genetic epidemiology, phylodynamics, population genetics

1 | INTRODUCTION

Infectious disease dynamics involve complex interactions among epidemiological, ecological and evolutionary (epi-eco-evo) processes (Grenfell et al., 2004; Schmid-Hempel, 2021). For instance, pathogen mutations that modify epidemiological traits (e.g. transmissibility (Volz et al., 2021) or drug resistance (Zur Wiesch et al., 2011)) affect epidemic progression and intervention efficacy in real time, while shifts in host demography, seasonality and control measures alter the selective landscape that drives further genetic change. To accurately represent these complexities, an integrated computational approach is essential for modelling and analysing genetic, ecological and epidemiological data (Baele et al., 2017).

Traditional analytical methods are often insufficient for capturing these complexities (Iranzo & Pérez-González, 2021; Willem et al., 2017), necessitating the development of more advanced simulation tools. Simulations provide a controlled framework for exploring analytically intractable models, rigorously evaluating statistical inference methods and generating realistic synthetic datasets when empirical ground truth is unavailable (Adrion et al., 2020; Hoban et al., 2012). Such datasets are particularly critical for hypothesis testing, validating computational algorithms and training deep-learning models that are increasingly applied to genomic epidemiology (Kraemer et al., 2025).

However, many existing genomic epidemiology simulators assume conditional independence among epi-eco-evo processes (Table 1): they first simulate an outbreak as a transmission tree independently of pathogen evolution and, if required, subsequently superimpose mutations onto this fixed genealogy to generate pathogen sequence data (Figure 1a). This two-step approach implicitly treats mutations as selectively neutral and ignores feedback from time-dependent ecological drivers (e.g. seasonality, interventions), thus precluding genetic changes from dynamically influencing ongoing disease spread (Figure 1b). Consequently, these models fail to capture scenarios such as highly transmissible variants that accelerate epidemics, emergent drug resistance that reshapes transmission networks or environmental fluctuations that alter pathogen demography and selection. This absence of dynamic coupling reduces realism and predictive power, limiting phylodynamic inference and method validation.

Here, we introduce e3SIM, an open-source, agent-based, forward-time simulation framework for macOS and Linux, built on `SLiM` (Haller & Messer, 2023), explicitly coupling pathogen genetic evolution with epidemiological and ecological processes in real time. Unlike existing simulators, e3SIM allows eco-evolutionary changes to continuously shape epidemic trajectories while ecological factors simultaneously drive evolutionary dynamics. e3SIM supports (i) configurable SEIRS (susceptible-exposed–infectious–recovered–susceptible) compartmental models, (ii) user-defined host contact networks (from random to empirical) and (iii) customizable pathogen genetic architectures, with optional within-host evolution, latent infection and multi-strain infection. Environmental fluctuations can be modelled through temporally varying ecological and epidemiological parameters, and phenotypic effects of *de novo* mutations propagate immediately into subsequent epidemiological events. Its modular design also supports neutral or partially coupled scenarios, alternative demographic models and diverse sampling schemes. A graphical user interface (GUI) further enhances accessibility.

2 | SOFTWARE ARCHITECTURE AND WORKFLOW

e3SIM comprises three components: (1) pre-simulation modules for configuration, (2) a main module for simulation execution and (3) post-simulation modules for analysis and visualization. The workflow begins with the sequential execution of pre-simulation modules (Figure 2) to configure the main module. To streamline this process,

2041210x, 2026, 4, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70278 by Cornell University Library, Wiley Online Library on [13/04/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Comparison of e3SIM with existing genomic epidemiology simulators across selected features relevant to this study.

Simulator name	Compartmental model	Epi-eco-evo dynamic coupling	Host population	Within-host evolution	Time-varying parameters
e3SIM [this paper]	**SEIRS User-defined**	**True**	**Contact network**	**True**	**True**
FAVITES (Moshiri et al., 2018)	User-defined	False	Contact network	True	False
`nosoi` (Lequime et al., 2020)	SIR	False	Structured	False	True
opqua (Cárdenas et al., 2022)	SIRDS	True	Structured	True	True
SANTA-SIM (Jariani et al., 2019)	—	True	—	False	True
SEEDY (Worby & Read, 2015)	SIR	False	Contact network	True	False
TiPS (Danesh et al., 2023)	User-defined	False	Structured	False	True
VGsim (Shchur et al., 2022)	SIS	True	Structured	False	True

Note: '—' indicates models without explicit host populations. Bold highlight that this is the method being presented in this paper. Compartmental states: S, susceptible; E, exposed; I, infectious; R, recovered; D, deceased.

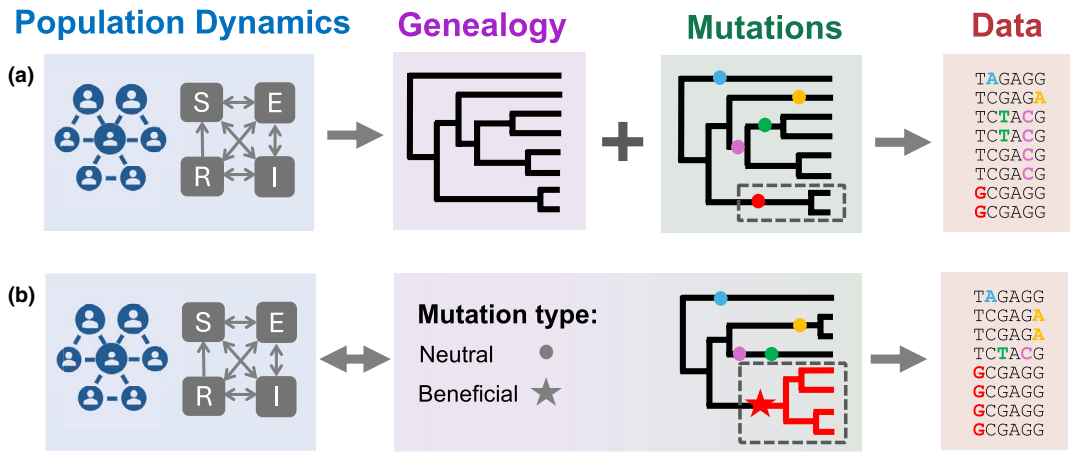

FIGURE 1 Schematic overview of epi-eco-evo coupling in genomic epidemiology simulators. (a) Conventional frameworks simulate transmission trees independently of pathogen evolution, subsequently superimposing mutations onto fixed genealogies. This two-step approach implicitly assumes selective neutrality of mutations, ignoring real-time feedback among genetic changes, epidemic dynamics and ecological factors. (b) e3SIM explicitly couples epi-eco-evo processes. Mutations immediately affect pathogen epidemiological traits, dynamically influencing epidemic spread, while ecological factors concurrently shape selective pressures and evolutionary dynamics. This integrated framework realistically models infectious disease dynamics.

a GUI is available, complementing command-line interaction. The main module employs `SLiM` 4.3 (Haller & Messer, 2023) as the back end for epi-eco-evo simulations (Figure 3a–c). Post-simulation modules generate visualizations and metadata summarizing results (Figure 3d).

2.1 | Pre-simulation modules: Input file generation

The four pre-simulation modules (Figure 2) run sequentially, generating inputs for the main module. `NetworkGenerator` constructs a host contact network (hosts as nodes, contacts as edges) using standard random network models (e.g. Erdős–Rényi, Barabási–Albert, random partition) or user-provided adjacency lists. `SeedGenerator` initializes pathogen genomes ('seeds') using forward-time `SLiM` simulations under neutral Wright–Fisher or network-based epidemiological dynamics; alternatively, users can supply custom seed genomes. `GeneticEffectGenerator` assigns causal genomic regions and effect sizes for quantitative traits (e.g. transmissibility, drug resistance) from user-provided genomic annotations, assuming an additive genetic architecture. Lastly, `HostSeedMatcher` pairs pathogen seeds with hosts via random, degree-ranked or percentile-based methods, enabling flexible modelling of outbreak initiation scenarios. Table S1 summarizes inputs and outputs for each pre-simulation module; further details are in Text S1 and S2.

2.2 | Main module: Epi-eco-evo simulation

After the pre-simulation modules, e3SIM executes `OutbreakSimulator` using a `JSON` configuration file. It employs `SLiM` as its back end,

2041210x, 2026, 4, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70278 by Cornell University Library, Wiley Online Library on [13/04/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

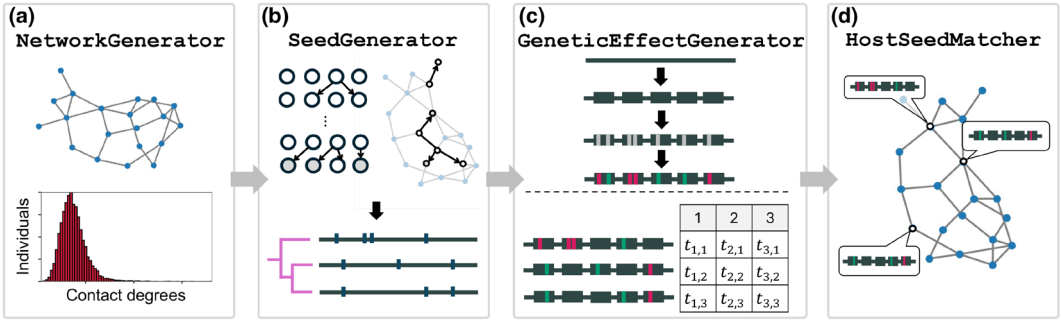

FIGURE 2 Pre-simulation modules. (a) `NetworkGenerator`. Constructs a host contact network. Histogram shows the contact-degree distribution. (b) `SeedGenerator`. Generates initial pathogen genomes ('seeds') and their genealogies using either the Wright–Fisher (top left) or a network-based epidemiological model (top right). Horizontal lines represent pathogen genomes; vertical ticks indicate mutations (bottom). (c) `GeneticEffectGenerator`. Defines causal sites and effect sizes for traits (grey: causal; magenta: positive; teal: negative). Trait values $t_{i,k}$ are sums of mutation effects for trait i and seed k. (d) `HostSeedMatcher`. Assigns seed pathogens to hosts to initiate outbreaks. All modules support user-defined inputs.

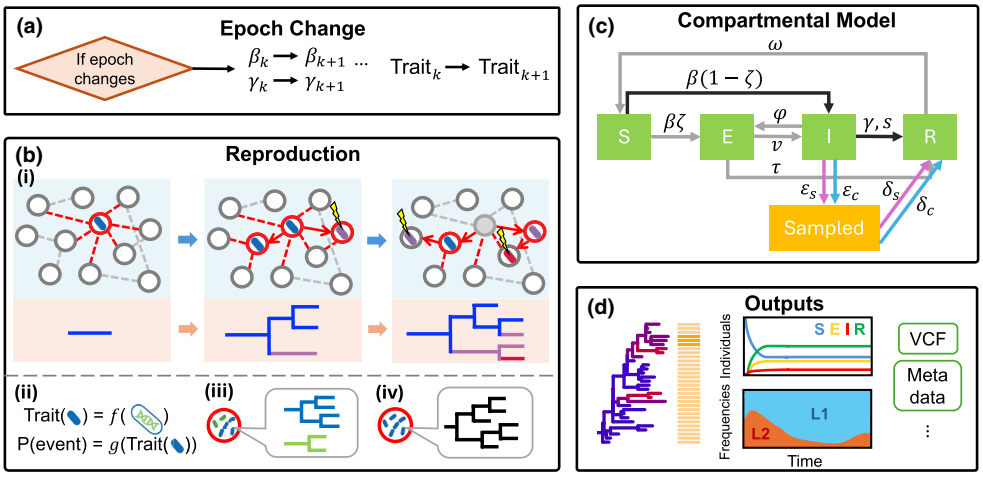

FIGURE 3 Main simulation scheme for e3SIM. (a) Epoch change. Epidemiological parameters (e.g. transmission probability β_k, recovery probability γ_k; panel c) and activation of genetic architectures for pathogen traits update from epoch k to $k+1$, modelling temporal environmental and epidemiological variation. (b) Reproduction. Pathogen transmission occurs on a host contact network based on trait-dependent transmission probabilities (panel i). Mutations (lightning symbols) arise stochastically during reproduction. Trait values are computed from pathogen mutation profiles and genetic architectures (panel ii). Superinfection (panel iii) and within-host replication (panel iv) occur if user-enabled. (c) Compartmental model and sampling. Hosts transition probabilistically between compartments (black lines: mandatory; grey lines: optional, user-specified). Purple and blue lines denote sequential and concerted sampling, respectively. Epidemiological parameters are defined in Table S3. (d) Outputs. Simulation outputs include genealogies (left; branches coloured by trait values), host compartment size trajectories (middle top), lineage proportions (middle bottom) and sampled pathogen mutation profiles and metadata (right).

generating an Eidos script from modular code blocks assembled per user specifications. These blocks implement the user-defined epi-eco-evo dynamics, scheduling functions at specified simulation points. Multiple replicates with identical configurations are supported.

`OutbreakSimulator` comprises three modules: epoch change, pathogen reproduction and compartmental model (Figure 3a–c). A time unit in e3SIM is a 'tick', and simulations are organized into discrete intervals ('epochs'). Epochs allow dynamic changes in epidemiological and evolutionary parameters and the activation of genetic architectures underlying specific pathogen traits at user-defined times (Figure 3a), enabling efficient modelling of temporal environmental variations or interventions.

The pathogen reproduction module models host-to-host transmission, constrained by network topology and pathogen trait values (e.g. transmissibility, drug resistance). Transmission events assume a single pathogen bottleneck, with optional superinfection (multiple simultaneous infections) or within-host replication up to a user-defined within-host capacity. Stochastic mutations occur according to user-defined substitution models and immediately update pathogen traits that influence epidemiological outcomes (Figure 3b), thereby dynamically coupling pathogen evolution with epidemiological and ecological processes.

The compartmental model employs a flexible SEIRS framework with trait-dependent state transition probabilities, supporting diverse

2041210x, 2026, 4, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70278 by Cornell University Library, Wiley Online Library on [13/04/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

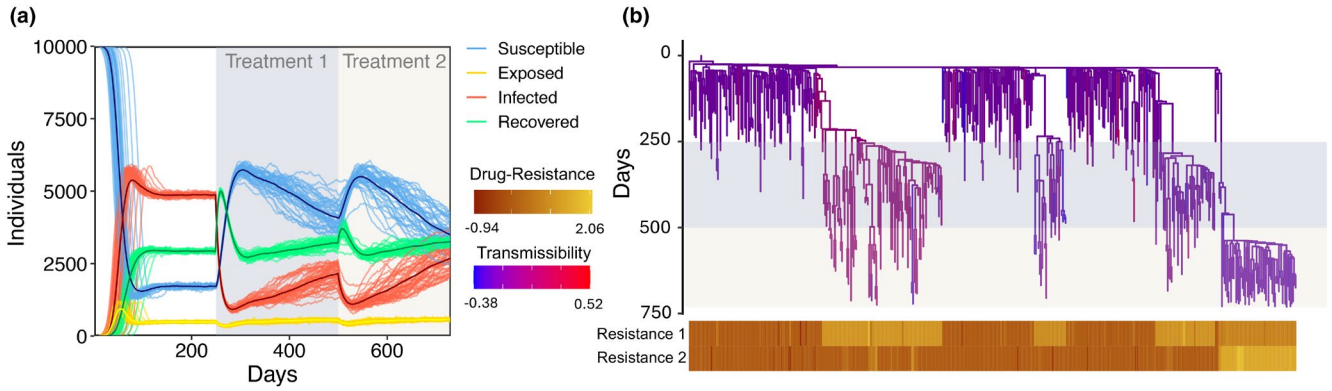

FIGURE 4 Simulated SARS-CoV-2 epi-eco-evo dynamics under sequential treatments. (a) SEIR trajectories across 43 replicates (bold line: replicate mean). Treatments begin on days 250 and 500. (b) Example sampled genealogy from one replicate. Branches are coloured by transmissibility (blue: lowest, red: highest); shaded backgrounds denote treatment epochs. Bottom heatmap shows resistance to treatments 1 and 2. Genealogies from other replicates are shown in Figure S2, with treatment-specific resistance colouring in Figures S3 and S4.

epidemiological scenarios, from simple SIR (susceptible–infectious–recovered) to complex SEIRS structures (Figure 3c). This trait-driven approach captures heterogeneous disease dynamics, including latency, reinfection and loss of immunity. The simulation supports sequential and concerted sampling schemes, common in epidemiological surveillance. Further model details are described in Text S3, with relevant configurations summarized in Tables S2 and S3.

2.3 | Post-simulation: Analysis and visualization

Upon completion of simulations, `OutbreakSimulator` executes post-simulation analyses (Figure 3d). e3SIM outputs mutation profiles (`VCF` or `FASTA`) and metadata for sampled pathogen sequences, from which users can infer genetic-distance–based or user-defined time-scaled phylogenies. A tree-sequence file (Haller et al., 2019) is generated and processed into genealogies (`NWK` format) of sampled progeny per seed. Detailed infection logs ('who infected whom'), integrated with genealogies, enable reconstruction of transmission histories and directly represent transmission trees when superinfection and within-host replication are disabled. e3SIM visualizes the resulting genealogies using the `ggtree` R package (Yu et al., 2017), with branches colour-coded by pathogen traits or seed ID. A seed tree is combined with simulated progeny genealogies to form a complete genealogy of sampled pathogens. e3SIM also plots lineage distributions and host compartment size trajectories.

2.4 | Graphical user interface

The primary means of interaction with e3SIM is its GUI, designed for users who prefer visual interaction over command-line tools. Developed using Python's `tkinter` framework, the GUI ensures a consistent user experience across macOS and Linux. It guides users step-by-step through simulation configuration and execution, with intuitive sections (Figure S1) for general parameters, evolutionary and epidemiological model settings, seed population setup and network model definition.

3 | SIMULATION DEMONSTRATIONS AND COMPUTATIONAL PERFORMANCE

3.1 | Simulation examples

3.1.1 | Emergence of new variants and epi-eco-evo dynamics of fast-evolving pathogens

We simulated a SARS-CoV-2 outbreak in 10,000 hosts on a Barabási–Albert contact network, demonstrating e3SIM's unique capacity to model dynamic coupling among epidemic dynamics, pathogen evolution and adaptive responses to changing environments. We specified genetic architectures for transmissibility and two drug-resistance traits, sequentially introducing two population-level drug treatments (Figure 4) to evaluate their impacts on outbreak dynamics, the emergence of drug resistance and long-term evolutionary consequences. Before treatment (days 0–249), epidemic dynamics reached equilibrium (Figure 4a), dominated by high-transmissibility pathogens (Figure 4b). The first treatment (days 250–499) initially caused a rapid decline in infected individuals, but subsequent de novo resistance mutations to the first drug (Figure S3) gradually reduced treatment efficacy, leading to a resurgence of infected and exposed populations. Even after introducing the second treatment (days 500–729), the selective advantage of strains resistant to the second drug (Figure S4), together with their elevated transmissibility (Figure S5), drove further epidemic resurgence.

3.1.2 | Effects of superspreaders and contact network on epidemic dynamics

To demonstrate e3SIM's ability to model dynamic interactions between host population structure and pathogen genetic traits, we

2041210x, 2026, 4, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70278 by Cornell University Library, Wiley Online Library on [13/04/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

simulated three epidemic scenarios with *Mtb* (Melsew et al., 2019) using five seed strains with varying transmissibility (Figure 5a): (1) random seed-host matching in an Erdős–Rényi network; (2) matching high-transmissibility seeds to socially active (high-degree) hosts and low-transmissibility seeds to less active (low-degree) hosts in a Barabási–Albert network (Figure 5b); and (3) the reverse scenario, matching low-transmissibility seeds to high-degree hosts and high-transmissibility seeds to low-degree hosts in a Barabási–Albert network (Figure 5c). Transmission dynamics depended strongly on host contact structure. High-transmissibility strains generally produced more progeny (Figure 5d), particularly when infecting socially active hosts (Figure 5e). However, even low-transmissibility strains could yield substantial numbers of progeny when infecting socially active hosts (Figure 5f), complicating inferences of pathogen transmissibility from observed cluster sizes (Tupper et al., 2022). Network topology and seeding strategy further influenced compartment sizes and equilibrium dynamics (Figure 5g–i; Figure S6).

3.2 | Runtime profiling

The runtime of `OutbreakSimulator` depends on various parameters influencing expected outbreak size, including the number of seeds and traits, genome length, mutation rate and epidemiological

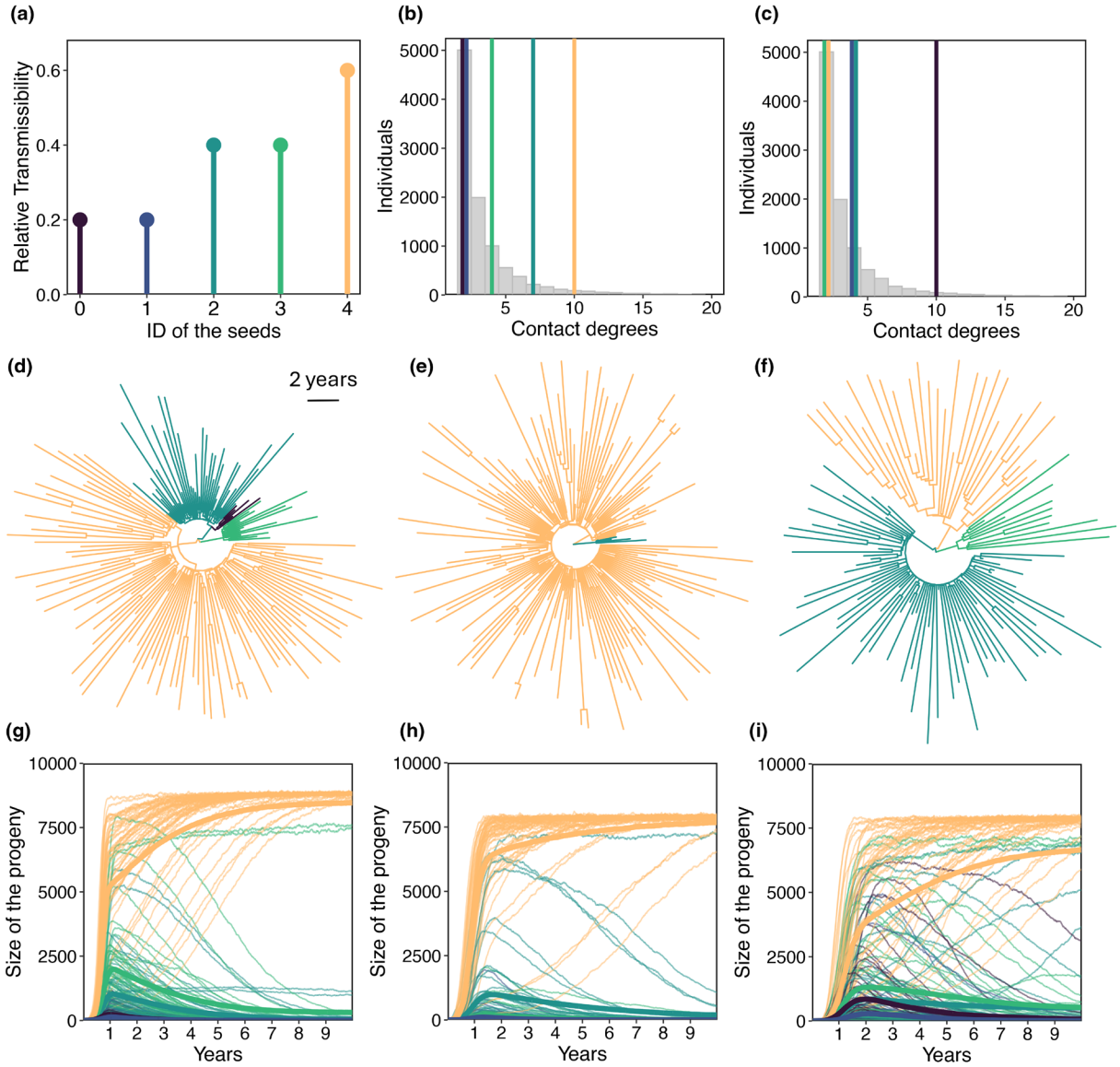

FIGURE 5 Effects of host population structure and pathogen genetic traits on *Mtb* transmission dynamics. (a) Relative seed transmissibility (IDs 0–4). Colours denote seed IDs across panels. (b, c) Truncated contact-degree distributions (degree ≤ 20; grey histograms) for seed-host matching scenario 2 (b, assortative: high-transmissibility → high-degree; low-transmissibility → low-degree) and scenario 3 (c, disassortative: low → high; high → low). Vertical lines indicate seed-matched contact degrees. (d–f) Example genealogies of sampled pathogens coloured by seed transmissibility for scenarios 1–3: random (d); assortative (e); disassortative (f). (g–i) Progeny-size trajectories per seed for scenarios 1–3 (50 replicates per scenario); bold lines indicate means. Corresponding SEIR compartment trajectories appear in Figure S6.

2041210x, 2026, 4, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70278 by Cornell University Library, Wiley Online Library on [13/04/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

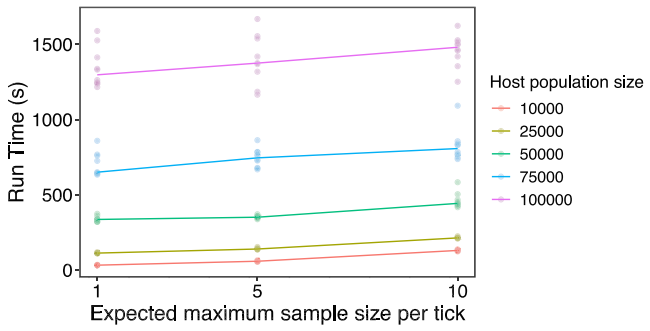

FIGURE 6 Runtime profiling on a personal computer (single core). Dots show replicate runtimes; lines indicate the median runtime across successful replicates (out of 10) per parameter set. The *x*-axis shows the expected maximum samples per tick (host population size × sequential sampling probability), set to 1, 5 or 10 across population sizes (Table S6).

parameters. We profiled runtime by varying host population size, sampling rate and base transmission probability (β). Simulations were run on a MacBook Pro (Apple M2 Pro CPU, 32 GB RAM). Runtime increased substantially with host population size (Figure 6; Figure S7A). For 100,000 hosts, simulations completed in ~30 min and reached equilibrium (~60% infected) after ~100 ticks. Increasing the expected maximum sample size per tick from 1 to 10 increased runtime for small populations (4.0-fold at 10,000 hosts) but had a minor effect for larger populations (~9% at 100,000 hosts). Base transmission probability (β) strongly influenced outbreak size, and thus runtime, as confirmed by additional profiling on Linux (Figure S7B; parameters in Table S5).

4 | DISCUSSION

We introduced e3SIM, an epi-eco-evo simulation framework designed for studying how the eco-evolutionary dynamics of pathogens impact disease transmission through the explicit coupling of these processes. With its modular architecture, e3SIM supports diverse epidemiological and population-genetic complexities, integrating compartmental models, host contact networks and quantitative trait models for pathogen traits. Additionally, e3SIM features a user-friendly GUI for broad accessibility. Simulations involving SARS-CoV-2 and *Mtb* demonstrate e3SIM's unique capability to realistically model pathogen molecular evolution jointly with epidemiological dynamics under environmental drivers.

Many simulators offer specialized features absent in e3SIM (e.g. TiPS's exact sample matching—Danesh et al., 2023; `nosoi`'s continuous-space transmission—Lequime et al., 2020), but our focus is on the efficient, fully integrated coupling of epi-eco-evo dynamics in a unified simulation framework (Table 1). While Cárdenas et al., 2022 introduced several advanced features, their approach is restricted to the SIRDS (susceptible–infectious–recovered–deceased–susceptible) compartmental model, thus excluding key epidemiological complexities such as latency (e.g. in *Mtb*), and it lacks molecular evolution models. Moreover, its assumption of uniform contact rates within structured populations restricts accurate modelling of individual-level heterogeneity. Another limitation of Cárdenas et al., 2022 is its inability to generate true pathogen genealogies, an essential component for phylodynamic studies in genomic epidemiology (Featherstone et al., 2022; Guinat et al., 2021; Ingle et al., 2021). Its complexity and lack of a user-friendly interface further limit practical usability. e3SIM addresses these limitations by comprehensively integrating these features within a modular and user-friendly framework.

While e3SIM is a powerful tool, it has several limitations. First, it does not model recombination, a crucial evolutionary mechanism in pathogens (Awadalla, 2003) such as HIV and hepatitis C, potentially constraining accurate representation of genetic diversity and evolutionary dynamics in non-clonal pathogens. e3SIM also does not explicitly model host-pathogen co-evolution, an important evolutionary driver in both host and pathogen (Woolhouse et al., 2002). However, these limitations could be addressed by leveraging `SLiM`'s recombination and multispecies features (Haller & Messer, 2023) and tree-sequence recording for phylogenetic network or ancestral recombination graph output with `tskit` (Wong et al., 2024). Furthermore, the current implementation employs a static contact network, limiting realistic representation of dynamic human interactions. In real-world scenarios, contact patterns can vary temporally due to social behaviours, public health interventions and seasonal variations (Bansal et al., 2010). To our knowledge, no existing genomic epidemiology simulators explicitly incorporate dynamic contact networks. Future enhancements of e3SIM could integrate dynamic networks and stochastic models of temporal contact formation and dissolution to achieve more realistic host interactions. Lastly, the integration of individual-based contact networks, genome-based pathogen evolution and dynamic epi-eco-evo coupling makes e3SIM computationally better suited to local rather than global-scale epidemic simulations.

AUTHOR CONTRIBUTIONS

Jaehee Kim conceived the ideas; Peiyu Xu, Andrew G. Clark and Jaehee Kim designed the methodology; Peiyu Xu, Shenni Liang, Andrew Hahn, Vivian Zhao, Wai Tung 'Jack' Lo and Benjamin C. Haller developed the software; Peiyu Xu and Shenni Liang collected and analysed the data; Jaehee Kim led the writing of the manuscript. All authors contributed to the drafts and gave final approval for publication. Andrew G. Clark and Jaehee Kim supervised the study.

ACKNOWLEDGEMENTS

PX, SL, AH, VZ, WTJL, BS, MHC, CC, TC, KYR, MTW and AGC were supported by the National Institutes of Health Grant P01AI159402. PX, WTL and JK were supported by the National Institute of General Medical Sciences of the National Institutes of Health under award R35GM156957. BCH and PWM were supported by the National Institutes of Health under awards R01HG012473 and R35GM152242.

2041210x, 2026, 4, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70278 by Cornell University Library, Wiley Online Library on [13/04/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

CONFLICT OF INTEREST STATEMENT

All authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data and source code for e3SIM are available via Zenodo at https://doi.org/10.5281/zenodo.18489185 (Xu et al., 2026). The Zenodo record includes the archived e3SIM release, the user manual and the scripts and data required to reproduce the simulation examples and runtime profiling. The source code repository is hosted on GitHub at https://github.com/EpiEvoSoftware/e3SIM.

ORCID

Jaehee Kim https://orcid.org/0000-0002-5210-2004

REFERENCES

Adrion, J. R., Cole, C. B., Dukler, N., Galloway, J. G., Gladstein, A. L., Gower, G., Kyriazis, C. C., Ragsdale, A. P., Tsambos, G., Baumdicker, F., Carlson, J., Cartwright, R. A., Durvasula, A., Gronau, I., Kim, B. Y., McKenzie, P., Messer, P. W., Noskova, E., Ortega-Del Vecchyo, D., … Kern, A. D. (2020). A community-maintained standard library of population genetic models. *eLife*, *9*, e54967. https://doi.org/10.7554/eLife.54967

Awadalla, P. (2003). The evolutionary genomics of pathogen recombination. *Nature Reviews Genetics*, *4*(1), 50–60. https://doi.org/10.1038/nrg964

Baele, G., Suchard, M. A., Rambaut, A., & Lemey, P. (2017). Emerging concepts of data integration in pathogen phylodynamics. *Systematic Biology*, *66*(1), e47–e65. https://doi.org/10.1093/sysbio/syw054

Bansal, S., Read, J., Pourbohloul, B., & Meyers, L. A. (2010). The dynamic nature of contact networks in infectious disease epidemiology. *Journal of Biological Dynamics*, *4*(5), 478–489. https://doi.org/10.1080/17513758.2010.503376

Cárdenas, P., Corredor, V., & Santos-Vega, M. (2022). Genomic epidemiological models describe pathogen evolution across fitness valleys. *Science Advances*, *8*(28), eabo0173. https://doi.org/10.1126/sciadv.abo0173

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*(2), 487–495. https://doi.org/10.1111/2041-210X.14038

Featherstone, L. A., Zhang, J. M., Vaughan, T. G., & Duchene, S. (2022). Epidemiological inference from pathogen genomes: A review of phylodynamic models and applications. *Virus Evolution*, *8*(1), veac045. https://doi.org/10.1093/ve/veac045

Grenfell, B. T., Pybus, O. G., Gog, J. R., Wood, J. L., Daly, J. M., Mumford, J. A., & Holmes, E. C. (2004). Unifying the epidemiological and evolutionary dynamics of pathogens. *Science*, *303*(5656), 327–332. https://doi.org/10.1126/science.1090727

Guinat, C., Vergne, T., Kocher, A., Chakraborty, D., Paul, M. C., Ducatez, M., & Stadler, T. (2021). What can phylodynamics bring to animal health research? *Trends in Ecology & Evolution*, *36*(9), 837–847. https://doi.org/10.1016/j.tree.2021.04.013

Haller, B. C., Galloway, J., Kelleher, J., Messer, P. W., & Ralph, P. L. (2019). Tree-sequence recording in SLiM opens new horizons for forward-time simulation of whole genomes. *Molecular Ecology Resources*, *19*(2), 552–566. https://doi.org/10.1111/1755-0998.12968

Haller, B. C., & Messer, P. W. (2023). SLiM 4: Multispecies eco-evolutionary modeling. *The American Naturalist*, *201*(5), E127–E139. https://doi.org/10.1086/723601

Hoban, S., Bertorelle, G., & Gaggiotti, O. E. (2012). Computer simulations: tools for population and evolutionary genetics. *Nature Reviews Genetics*, *13*(2), 110–122. https://doi.org/10.1038/nrg3130

Ingle, D. J., Howden, B. P., & Duchene, S. (2021). Development of phylodynamic methods for bacterial pathogens. *Trends in Microbiology*, *29*(9), 788–797. https://doi.org/10.1016/j.tim.2021.02.008

Iranzo, V., & Pérez-González, S. (2021). Epidemiological models and COVID-19: A comparative view. *History and Philosophy of the Life Sciences*, *43*(3), 104. https://doi.org/10.1007/s40656-021-00457-9

Jariani, A., Warth, C., Deforche, K., Libin, P., Drummond, A. J., Rambaut, A., Matsen, F. A., IV, & Theys, K. (2019). SANTA-SIM: simulating viral sequence evolution dynamics under selection and recombination. *Virus Evolution*, *5*(1), vez003. https://doi.org/10.1093/ve/vez003

Kraemer, M. U. G., Tsui, J. L.-H., Chang, S. Y., Lytras, S., Khurana, M. P., Vanderslott, S., Bajaj, S., Scheidwasser, N., Curran-Sebastian, J. L., Semenova, E., Zhang, M., Unwin, H. J. T., Watson, O. J., Mills, C., Dasgupta, A., Ferretti, L., Scarpino, S. V., Koua, E., Morgan, O., … Bhatt, S. (2025). Artificial intelligence for modelling infectious disease epidemics. *Nature*, *638*(8051), 623–635. https://doi.org/10.1038/s41586-024-08564-w

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. https://doi.org/10.1111/2041-210X.13422

Melsew, Y. A., Gambhir, M., Cheng, A. C., McBryde, E. S., Denholm, J. T., Tay, E. L., & Trauer, J. M. (2019). The role of super-spreading events in *Mycobacterium tuberculosis* transmission: Evidence from contact tracing. *BMC Infectious Diseases*, *19*(1), 1–9. https://doi.org/10.1186/s12879-019-3870-1

Moshiri, N., Ragonnet-Cronin, M., Wertheim, J. O., & Mirarab, S. (2018). FAVITES: simultaneous simulation of transmission networks, phylogenetic trees and sequences. *Bioinformatics*, *35*(11), 1852–1861. https://doi.org/10.1093/bioinformatics/bty921

Schmid-Hempel, P. (2021). *Evolutionary parasitology: The integrated study of infections, immunology, ecology, and genetics* (2nd ed.). Oxford University Press.

Shchur, V., Spirin, V., Sirotkin, D., Burovski, E., De Maio, N., & Corbett-Detig, R. (2022). VGsim: Scalable viral genealogy simulator for global pandemic. *PLoS Computational Biology*, *18*(8), e1010409. https://doi.org/10.1371/journal.pcbi.1010409

Tupper, P., Pai, S., COVID Schools Canada, & Colijn, C. (2022). COVID-19 cluster size and transmission rates in schools from crowdsourced case reports. *eLife*, *11*, e76174. https://doi.org/10.7554/eLife.76174

Volz, E., Hill, V., McCrone, J. T., Price, A., Jorgensen, D., O'Toole, A., Southgate, J., Johnson, R., Jackson, B., Nascimento, F. F., Rey, S. M., Nicholls, S. M., Colquhoun, R. M., da Silva Filipe, A., Shepherd, J., Pascall, D. J., Shah, R., Jesudason, N., Li, K., … Pybus, O. G. (2021). Evaluating the effects of SARS-CoV-2 spike mutation D614G on transmissibility and pathogenicity. *Cell*, *184*(1), 64–75.e11. https://doi.org/10.1016/j.cell.2020.11.020

Willem, L., Verelst, F., Bilcke, J., Hens, N., & Beutels, P. (2017). Lessons from a decade of individual-based models for infectious disease transmission: A systematic review (2006–2015). *BMC Infectious Diseases*, *17*(1), 612. https://doi.org/10.1186/s12879-017-2699-8

Wong, Y., Ignatieva, A., Koskela, J., Gorjanc, G., Wohns, A. W., & Kelleher, J. (2024). A general and efficient representation of ancestral recombination graphs. *Genetics*, *228*(1), iyae100. https://doi.org/10.1093/genetics/iyae100

Woolhouse, M. E., Webster, J. P., Domingo, E., Charlesworth, B., & Levin, B. R. (2002). Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nature Genetics*, *32*(4), 569–577. https://doi.org/10.1038/ng1202-569

Worby, C. J., & Read, T. D. (2015). 'SEEDY' (simulation of evolutionary and epidemiological dynamics): An R package to follow accumulation of within-host mutation in pathogens. *PLoS One*, *10*(6), e0129745. https://doi.org/10.1371/journal.pone.0129745

2041210x, 2026, 4, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70278 by Cornell University Library, Wiley Online Library on [13/04/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Xu, P., Liang, S., Hahn, A., Zhao, V., Lo, W. T. J., Haller, B. C., Sobkowiak, B., Chitwood, M. H., Colijn, C., Cohen, T., Rhee, K. Y., Messer, P. W., Wells, M. T., Clark, A. G., & Kim, J. (2026). e3SIM software and demo files. *Zenodo*. https://doi.org/10.5281/zenodo.18489185

Yu, G., Smith, D. K., Zhu, H., Guan, Y., & Lam, T. T.-Y. (2017). ggtree: An R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution*, *8*(1), 28–36. https://doi.org/10.1111/2041-210X.12628

Zur Wiesch, P. A., Kouyos, R., Engelstädter, J., Regoes, R. R., & Bonhoeffer, S. (2011). Population biological principles of drug-resistance evolution in infectious diseases. *The Lancet Infectious Diseases*, *11*(3), 236–247. https://doi.org/10.1016/S1473-3099(10)70264-4

SUPPORTING INFORMATION

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

Appendix S1. Supporting Information for e3SIM, including additional computational and simulation details, supplementary tables and figures, and example configuration files.

How to cite this article: Xu, P., Liang, S., Hahn, A., Zhao, V., Lo, W. T. '., Haller, B. C., Sobkowiak, B., Chitwood, M. H., Colijn, C., Cohen, T., Rhee, K. Y., Messer, P. W., Wells, M. T., Clark, A. G., & Kim, J. (2026). e3SIM: Epidemiological-ecological-evolutionary simulation framework for genomic epidemiology. *Methods in Ecology and Evolution*, *17*, 1051–1059. https://doi.org/10.1111/2041-210x.70278

2041210x, 2026, 4, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210x.70278 by Cornell University Library, Wiley Online Library on [13/04/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License