

<https://doi.org/10.1038/s42003-025-07540-y>

A phylogenetic epidemiology approach to predicting the establishment of multi-host plant pests

Shannon Colleen Lynch^{1,2}✉, Edeli Reyes-Gonzalez², Emily L. Bossard², Karen S. Alarcon², Natalie L. R. Love^{3,4}, Allan D. Hollander⁵, Beatriz E. Nobua-Behrman⁶ & Gregory S. Gilbert¹

Forecasting emergent pest spread is paramount to mitigating their impacts. For host-specialized pests, epidemiological models of spread through a single host population are well developed. However, most pests attack multiple host species; the challenge is predicting which communities are most vulnerable to infestation. Here, we develop a phylogenetically-informed approach to predict establishment of emergent multi-host pests across heterogeneous landscapes. We model a beetle-pathogen symbiotic complex on trees, introduced from Southeast Asia to California. The *phyloEpi* model for likelihood of establishment was predicted from the phylogenetic composition of woody species in the invaded community and the influence of temperature on beetle reproduction. Plant communities dominated by close relatives of known epidemiologically critical hosts were four times more likely to become infested than communities with more distantly related species. Where microclimate favored beetle reproduction, pest establishment was greater than expected based only on species composition. We applied this *phyloEpi* model to predict infestation risk in California using weather data and complete tree inventories from 9262 1-km² grids in 170 cities. Regions in the state predicted with low likelihood of infestation were confirmed by independent monitoring. Analysts can adapt these phylogenetic ecology tools to predict spread of any multi-host pest in novel habitats.

Most pathogen and insect pests attack multiple host species^{1–3}, but the epidemiological models traditionally used to predict the spread of emergent pests generally assume density- or frequency-dependent dynamics expected for host specialists^{4,5}. Predicting which communities are most vulnerable to emergent multi-host (or polyphagous) pests is an urgent management priority because they can cause large-scale transformations of naïve ecosystems⁶ and incur significant costs^{7–10}. An epidemic develops when a virulent pathogen (and any needed vectors) encounters a population of susceptible hosts where environmental conditions are favorable. For multi-host pests, the susceptible host community can include numerous species that vary in abundance and in their competence to support pest reproduction^{11–14}. Rather than the density of a single host species, it is the collective abundance of such alternative host species in a local community and their ability to support pest reproduction that determine the probability of infestation and local spread¹⁵. Decision-makers faced with emerging infectious diseases caused by multi-host pests require robust risk models

that apply knowledge of host range, reproductive potential, and environmental requirements to plant communities of concern. Unfortunately, the very novelty of emergent pests usually means that they are not at equilibrium with their environment, and therefore, empirical data on the susceptibility and competence of local host species are usually incomplete.

Evolutionary tradeoffs in traits that confer host susceptibility or in the ability of enemies to attack hosts produce a phylogenetic signal for host range, where closely related plant species are more likely to share pests¹⁶. Such a phylogenetic signal has been well documented for plant-pathogen and plant-insect interactions, including for insect-pathogen pest complexes^{9,12,17}. In addition, there is a predictive phylogenetic signal of the relative impacts of pests on their hosts. For example, phylogenetic relatedness has proved capable of predicting mortality caused by multi-host animal parasites¹⁸; host impacts for a variety of herbivores and pathogens¹⁹; severity of native and non-native pests on North American tree species²⁰; and the amount of herbivory-induced leaf damage on oak species (*Quercus* spp.)^{21–23}. Because

¹Department of Environmental Studies, University of California Santa Cruz, Santa Cruz, CA, USA. ²Department of Plant Pathology, University of California Davis, Davis, CA, USA. ³Biological Sciences Department, California Polytechnic State University, San Luis Obispo, CA, USA. ⁴Negaunee Institute for Plant Conservation Science and Action, Chicago Botanic Garden, Glencoe, IL, USA. ⁵Institute of the Environment, University of California Davis, Davis, CA, USA. ⁶University of California Cooperative Extension, Irvine, CA, USA. ✉e-mail: sclynch@ucdavis.edu

spread from introduced sites is an intrinsically spatial process, standard models of spread of invasive pests focus on spatial processes such as dispersal constraints as predictors of invasion²⁴. In addition to dispersal abilities, realized dispersal of pests depends on the spatial availability of communities that include susceptible and competent hosts in the landscape. Given that the establishment of multi-host pests is more complex than those limited to a single host species, spatial evaluation of risk should first focus on identifying those communities that are most likely to be vulnerable based on the combination of biotic composition and abiotic conditions. Previous work has shown that phylogenetic signal in host range can be used to predict the likelihood that resident pathogens from a local plant community will spill-over onto a novel host, based on previous knowledge of competent hosts¹⁵. Therefore, before introducing spatial processes into complex epidemiological models, we ask the reciprocal question here: can we use phylogenetic dimensions of pest host ranges to predict the likelihood that a novel multi-host pest will infest particular plant communities (Fig. 1)? We further ask how community phylogenetic structure compares in predictive importance to other factors such as environmental conditions. Once suitable habitat is identified, models of spatial spread, combined with surveillance, can be effective in describing pest spread across the landscape.

In this study, we develop and test a phylogenetically-informed approach to modeling the risk of pest establishment in wild and managed

plant communities (Fig. 1). We use monitoring data from an emergent epiphytic of *Fusarium* dieback–invasive shothole borers (FD-ISHB), caused by an ambrosia beetle–fungal pathogen complex that was apparently introduced to Southern California in the early 2000s from Southeast Asia²⁵. The complex is a symbiotic relationship in which the ISHB beetles (*Euwallacea fornicatus* and *E. kuroshio*) depend on the *Fusarium* fungus (*F. euwallacea* and *F. kuroshium*) as a source of food, and the fungal pathogen depends on the beetles for transmission to a new host²⁶. FD-ISHB has killed hundreds of thousands of apparently healthy trees in urban and wildland forests^{27,28}, costing the state over \$20 million²⁹. New FD-ISHB introductions have emerged in South Africa³⁰, where it has expanded from urban to wildland forests in eight of its nine provinces. Phylogenetic dispersion analysis on a comprehensive set of host range data³¹ shows that the strength of the phylogenetic signal is progressively more pronounced for more severely affected host species¹³; Supplementary Table 1). That is, the range of 18 host species killed by the pest-pathogen complex (i.e., killed-competent hosts) is phylogenetically narrower than the 59 additional species on which it can reproduce (i.e., competent hosts). Most recently, FD-ISHB has emerged in Western Australia³², where species-level empirical data on susceptibility and competence are not yet available. Phylogenetic models could be used in such situations to predict complex novel pest occurrences in different local ecosystems, serving as a critical risk assessment tool that facilitates rapid responses and forestalls damaging invasions.

We used our evolutionary understanding of the FD-ISHB host range¹³ and data from 15,000 trees in 207 0.25-ha monitoring plots to build an explanatory phylogenetic epidemiology (*phyloEpi*) model of the likelihood of pest establishment based on the phylogenetic structure of FD-ISHB host range and tree species composition in the plots (wpS; Fig. 1). Because generation time for the beetle vector is a function of temperature^{33,34}, we further examined how infestation risk based on tree community composition was modified by local microclimate. To test whether these explanatory models could be used as predictive tools to prioritize actions in response to emergent diseases, we applied the model to predict infestation risk in 170 cities across California using complete urban street tree inventories³⁵. We then tested the effectiveness of the model to predict FD-ISHB establishment across those cities using observational data that were not included in model development.

Results

The probability that the emergent disease complex *Fusarium* dieback–invasive shothole borers (FD-ISHB) establishes in a plot is a function of the phylogenetic composition of a local plant community. Plant communities with abundant close relatives of killed-competent host species are up to four times more likely to be infested than those with distantly related hosts (Fig. 2). Logistic regression revealed that neither local tree abundance nor species richness alone were predictive of beetle establishment (i.e., for total tree density $P = 0.373$; for basal area in a plot $P = 0.838$; $P = 0.079$ for species richness). Instead, the estimated probability that a site was infested increased significantly with density-based wpS ($P < 0.001$; Hosmer–Lemeshow Goodness-of-Fit Chi-squared = 1.78, $P = 0.672$; Fig. 2) and basal area-based wpS ($P < 0.001$; Hosmer–Lemeshow Goodness-of-Fit Chi-squared = 1.78, $P = 0.672$).

The abundance-weighted probability of FD-ISHB establishment for the 99 observed infested plots (median wpS = 0.80) was left skewed (skewness = -0.77; Fig. 3a and S3a), while wpS values for the 108 non-infested plots (median = 0.50) were right-skewed (skewness = 0.82; Fig. 3b and Supplementary Fig. 3b). Of the total number of plots, 75.8% of infested and only 42.6% of non-infested plots had wpS values greater than 0.60. Predictive discriminant analysis on density-based wpS estimates was able to correctly classify 74.5% of the 47 infested plots (95% CI 70.2–80.9%), and 61.5% (95% CI 61.5–69.2%) of the 39 non-infested plots. For basal area-based wpS estimates, infested plots were correctly classified 76.6% (72.3–87.2%) and non-infested plots were correctly classified 64.1% (64.1–66.7%). As expected, analysis of distance-based autocorrelation in infestation patterns suggests fine-scale patterns of spatial autocorrelation

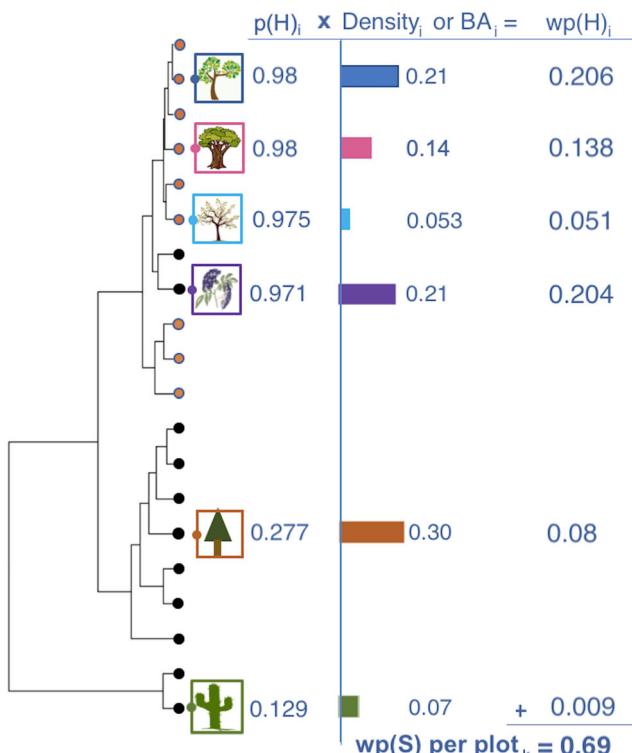


Fig. 1 | Procedure used to estimate probability that a plot would become infested by a multi-host pest. Scheme depicting the procedure used to estimate probability that a plot (k) would become infested by a multi-host pest, based on the interaction between phylogenetic structure and abundance (host density or basal area) of woody species in a plot. Orange tips on the phylogeny represent species that are killed-competent hosts. Here, the probability that a tree species i is a host ($p(H)_i$) is a function of the phylogenetic distance of species i to each of 18 known killed-competent hosts j . The probability that tree species i shares the pathogen with killed-competent host j ($p(S)_{ij} = \text{antilogit}(3.4 - 3.7 \times \log_{10}[\text{PD}_{ij} + 1])$) is based on the pairwise phylogenetic distance PD_{ij} . The overall probability that tree species i is a host ($p(H)_i = 1 - \prod[1 - p(S)_{ij}]$) is then the complement of the product of probabilities that tree species i is not susceptible to each of the 18 killed-competent host species j . This $p(H)_i$ is weighted by their relative abundances (density or basal area) for each species i within each plot k ($wp(H)_i$). The overall estimate of plot susceptibility ($wp(S)_k$) is the sum of the weighted host probabilities within each plot k .

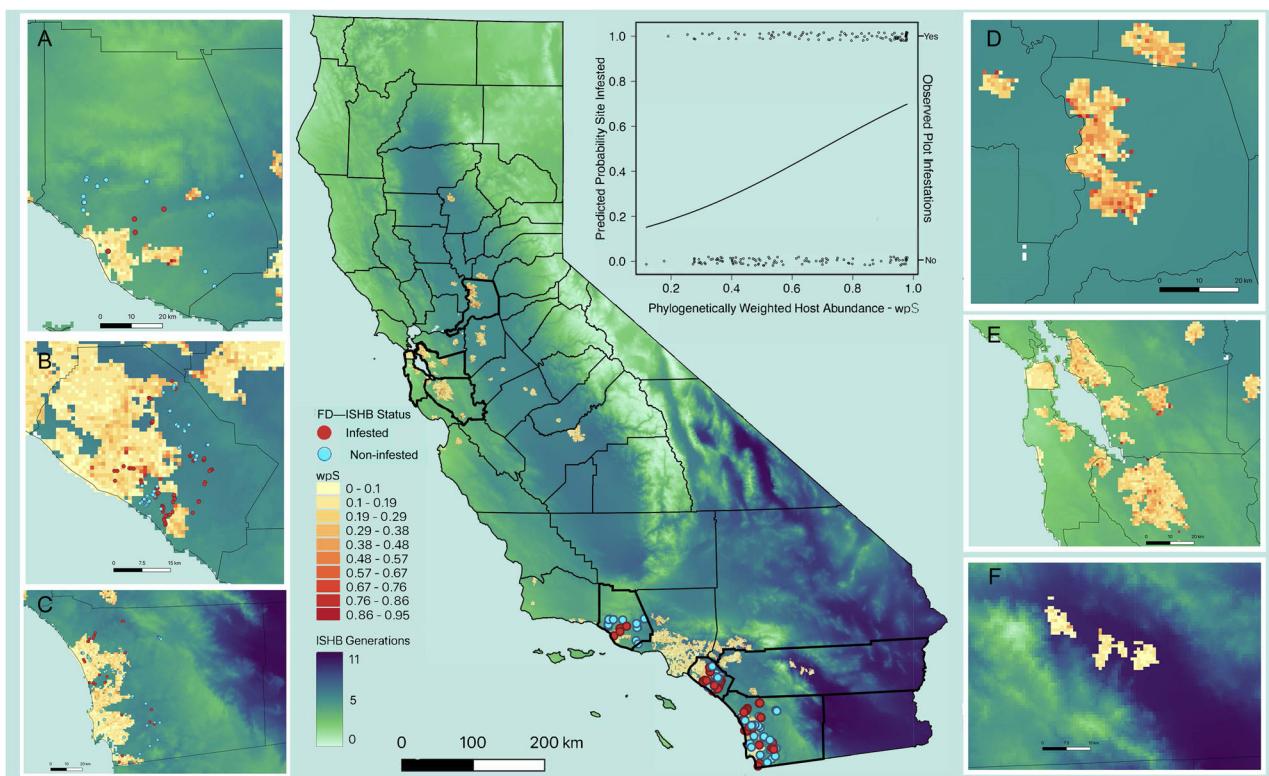


Fig. 2 | Risk map estimating plot or grid susceptibility to Fusarium dieback – invasive shothole borers (FD-ISHB). Risk map (1-km resolution) estimating plot or grid susceptibility to Fusarium dieback – invasive shothole borers (FD-ISHB) based on host species composition (yellow to red) and microclimate impact on beetle generations (light green to dark blue). The susceptibility model was developed using observational data from 207 monitoring plots (red and cyan dots) in Ventura (A), Orange (B), and San Diego (C) Counties (left panel) and applied to urban forests across the state (e.g., Sacramento County (D), Bay Area (E), and Palm Springs (F) in the right panel). Plot establishment was most strongly predicted by density-based phylogenetically weighted host abundance (wpS) of the local forest community

($p(S)_k = -2.15 + 3.26 \times wpS$; $P < 0.001$; AIC = 231; Nagelkerke $R^2 = 0.15$; Hosmer–Lemeshow Goodness-of-Fit $P = 0.672$). Grid clusters exhibit host composition-based susceptibility estimates (wpS) for urban forests with complete street tree inventories ranging from the least susceptible in yellow to the most susceptible in red. A degree-day model estimated the annual number of beetle generations that could occur within each grid; light green grids support fewer beetle generations than grids in dark blue. The current geographic infested range was determined based on regions of the state in which monitoring and surveillance had previously detected FD-ISHB.

expected in invasion processes (Fig S7). Analysis of DHARMA residuals of the model showed no significant pattern that deviates from a uniform distribution, suggesting the models fit the data well (Fig S8a).

Further analysis indicates that the effect of temperature on infestation of plots depends on the community context: warmer places that can support more generations of beetles allow infestation of sites that are otherwise unfavorable. Logistic regression analysis detected a significant negative interaction between the effects of density-based wpS and the number of beetle generations on FD-ISHB establishment ($P < 0.001$; Nagelkerke $R^2 = 0.15$; Hosmer–Lemeshow Goodness of Fit Chi-squared = 5.76, $P = 0.98$). In other words, plots expected to be non-susceptible based on species composition (density-based wpS) had a greater chance of FD-ISHB establishment where local microclimate could support more beetle generations (Supplementary Table 2). Analysis of DHARMA residuals of the model showed no significant pattern that deviates from a uniform distribution, suggesting the models fit the data well (Supplementary Fig 8b). However, the relative importance of this interaction was minor compared to the effect of phylogenetic community composition alone (Supplementary Table 2). This pattern emerged largely because microclimates are highly similar throughout the current infested range (ANOVA, $P = 0.698$), so that model development was necessarily based on locations that favor more beetle generations (Fig. 4). In addition, the interaction was not statistically significant in analyses based on basal area-based wpS estimates ($P = 0.22$). Taken together, we see that infestation was only possible where temperatures favored an adequate number of beetle generations; even where density-based wpS indicated a likely infestation, areas with fewer than five beetle

generations remained non-infested. Defining the precise shape of the density-based wpS \times beetle generations interaction will require validation across a broader range of host composition and microclimate conditions. Hence, our final *phyloEpi* model treats the effects of phylogenetic community composition ($p(S)_k = -2.15 + 3.26 \times wpS$) and microclimate on beetle reproduction (generations = Cumulative degree days/ K_{ISHB}) as separate, but complementary factors, in a conceptual evaluation of site susceptibility to FD-ISHB establishment in California's forests (Fig. 2).

To test how well our model could predict susceptibility of a forest site to infestation, we used *phyloEpi* to predict the likely susceptibility of 832 urban forest grids from across California; these grids had independent FD-ISHB monitoring data and were outside our study area, so they were not used in model development (monitored grid points on Fig. 4). Most of the grids (87%) had wpS values < 0.3 and fewer than five expected beetle generations, suggesting a low probability of infestation (Fig. 4 and Supplementary Fig. 5). The *phyloEpi* model predicted observed infestation in the grids (Fig. 5; $P = 0.019$). Grids in which disease was observed generally had a greater abundance of close relatives to killed-competent hosts (Fig. 5), suggesting that phylogenetic community composition of a local plant community strongly drives pest establishment. *PhyloEpi* estimates of wpS > 0.3 were associated with 699 (24%) of the 2956 grids located outside the observed infested range, suggesting the potential for FD-ISHB establishment over a large geographic extent (e.g., Fig. 2d-f). For 32 of the 36 confirmed non-infested statewide monitored grids that were outside the zone of infestation (regions of the state outside of those areas in which monitoring and surveillance had previously detected FD-ISHB), our models forecasted a low

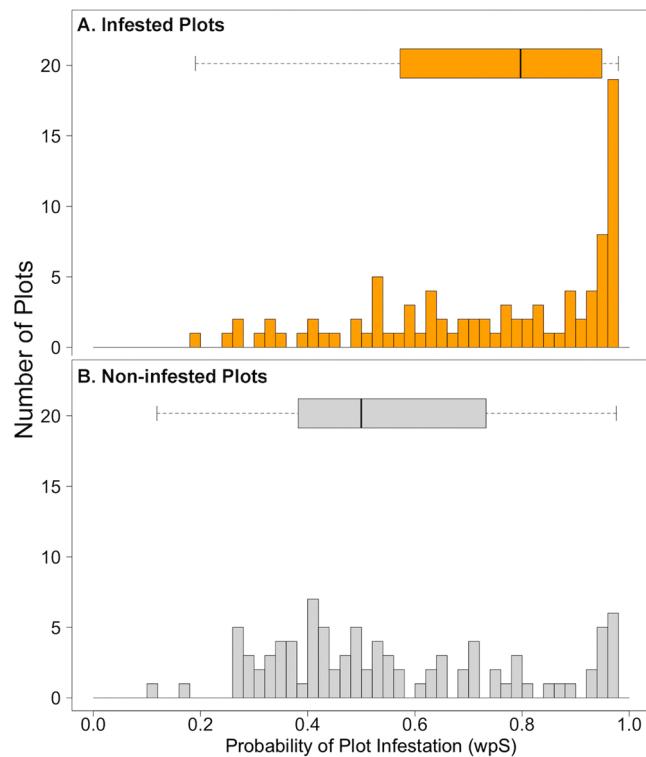


Fig. 3 | Number of observed infested and non-infested plots across abundance-weighted estimates of Fusarium dieback – invasive shothole borers (FD-ISHB) establishment in California. Number of observed (A) infested ($n = 99$) and (B) non-infested ($n = 108$) plots across abundance-weighted (host density) estimates of Fusarium dieback – invasive shothole borers (FD-ISHB) establishment in California. The predicted susceptibility of each plot is based on the local composition of woody species and their phylogenetic distance to known killed-competent host species. Whereas observed infested plots have a high predicted probability of being infested (left-skew; -0.77), non-infested plots are expected to have a low infestation probability (right-skew; 0.82).

probability of infestation, confirming the model's predictive power across a broader range of conditions. Strikingly, the majority of monitoring locations located outside the known infested range—grids that were not included in model development—are within the ranges of wpS and generation estimates expected for non-infested sites (black squares in Fig. 4 and Supplementary Fig. 5). These patterns of model classification suggests that the relative impact of phylogenetic community composition and microclimate are important drivers for FD-ISHB establishment.

Discussion

Decision-makers require predictive analytical tools that are robust and broadly applicable to effectively manage emergent pathogens and insect pests that attack multiple host species. Current approaches that are based on the underlying assumption that pests have narrow host ranges do not adequately address the epidemiological complexity of multi-host pests across heterogeneous plant communities. As a necessary first step to developing advanced spatial models to predict pest spread in novel habitats, we address that challenge by using the phylogenetic composition and microclimate conditions of local plant communities to predict the likelihood of multi-host pests invading those communities. The *phyloEpi* model at the core of our phylogenetic approach provides a powerful predictive tool that accounts for alternative host species driving the spread of multi-host pests across heterogeneous landscapes. As expected, plant communities with abundant species that are close relatives of known highly susceptible hosts were up to four times more likely to be infested than communities composed of distantly related hosts. The model did a better job predicting which plots would be infested (correctly classifying 72%) versus which would be non-

infested (61.5% correct), most likely because currently non-infested plots could eventually become infested as the epidemic proceeds. In addition, plots deemed unfavorable for infestation based on phylogenetically-weighted species abundance were more likely to suffer FD-ISHB establishment where local microclimate could support more beetle generations; this suggests that microclimate mediates the importance of community structure on disease establishment.

We applied the *phyloEpi* model to predict infestation risk in California's urban forests using complete street tree inventories from 9262 1-km² grids in 170 cities statewide, and monitoring and weather data sets independent from data used for model development. Our model indicated that nearly one quarter of the 2956 1-km² grid cells that were located outside the FD-ISHB infested range were likely to be highly susceptible to invasion. That there were no observed infestations in most of the low probability grids (32 of the 36 monitored sites) confirmed the model's predictive power across a broader range of conditions using data sources independent of those used to parameterize the *phyloEpi* model. Together, our results demonstrate that independent of landscape and spatial factors, the establishment of novel or emergent multi-host pests through complex landscapes can be predicted through host evolutionary relationships and their abundance in a local plant community, even when extensive empirical data about susceptibility of local plant species is not yet available.

In addition to improved predictions of multi-host establishment, our approach allows us to model expected impacts. Most pest risk analyses report the relative likelihood of a species' entry or establishment in an area, without addressing potential impacts (refs. 24,36; but see ref. 37). The model in this study predicts the probability of FD-ISHB establishment, but also incorporates two important indicators of potential damage (e.g., mortality, dieback): (1) phylogenetic signal in FD-ISHB host range, which is strongest on more severely impacted host species¹³; and (2) microclimate effects on beetle generations, which amplify impacts on hosts. Locations with high susceptibility estimates (high wpS) are therefore more likely to experience greater tree mortality. Accounting for pest impacts through a phylogenetic epidemiology approach thus fulfills the need to improve risk maps beyond potential pest distribution to also identify areas where they will likely cause the most harm²⁴.

Our *phyloEpi* model could be applied to existing spatio-temporal models of pests spread (e.g., dispersal-kernel, PoPs, process-based, network) and improve their utility by integrating pest impacts^{38–42}. Spread of invasive pests is inherently a spatial process, and dispersal kernels can be used to the model potential range of spread from data on the current distribution of an invasive pest³⁸. The susceptible habitats identified by the phylogenetic model could be overlaid with such spatial models to identify those areas most immediately threatened with infestation. Such an extension requires data from established monitoring programs on currently infested areas.

Climate conditions provide outer envelopes around potential geographic distributions of pests, given their strong influence on pest phenology, reproduction, dispersion, and overwintering survival⁴³. In this study, temperature constraints on ISHB development pointed to those regions where beetles could thrive, including locations where tree-level data are unavailable (Fig. 2). Our analysis detected that warmer places that can support more annual beetle generations could permit infestation of sites that would have been considered unfavorable based only on tree species composition. However, this predicted interaction between phylogenetic community composition and microclimate requires further validation across a broader range of host composition and microclimate conditions than where empirical data are currently available. We therefore treat these predicted patterns as tentative, but potentially useful, complementary guidance for decision makers as they evaluate where to monitor or apply appropriate preventative control measures in the immediate term. For example, some grid cells in the Central Valley and in the Bay Area have similar high probabilities of infestation risk based on host composition. Decision-makers might prioritize preventative actions in the Central Valley because those grids have more favorable temperatures for beetle development, which

Fig. 4 | Density-based wpS estimates (phylogenetic likelihood of infestation) of plot or grid susceptibility to Fusarium dieback – invasive shothole borers (FD-ISHB) establishment in California as a function of estimated number of temperature-dependent ISHB beetle generations supported in those locations. Data from 99 infested (pink filled circles) and 108 non-infested (yellow filled circles) plots were used to parameterize the model. The model was then applied to predict infestation risk in 9262 1-km² grids (gray dots) in 170 cities across California using complete urban forest inventories. A subset of 832 grids were independently monitored for FD-ISHB across its known infested (red or blue triangles) and non-infested (black squares) geographic range from 2012–2021 and used to test the model.

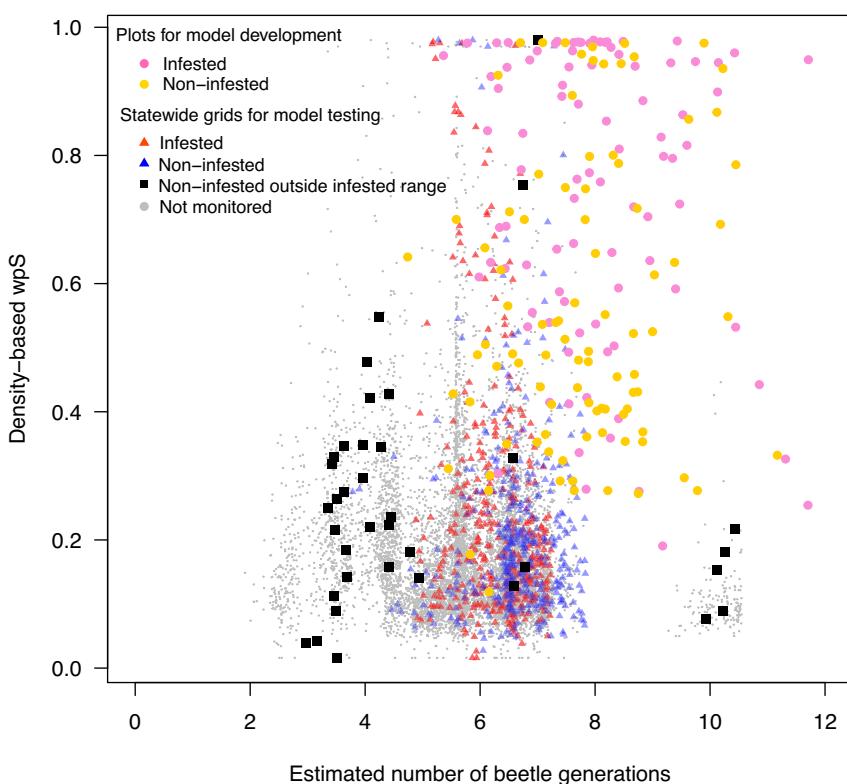
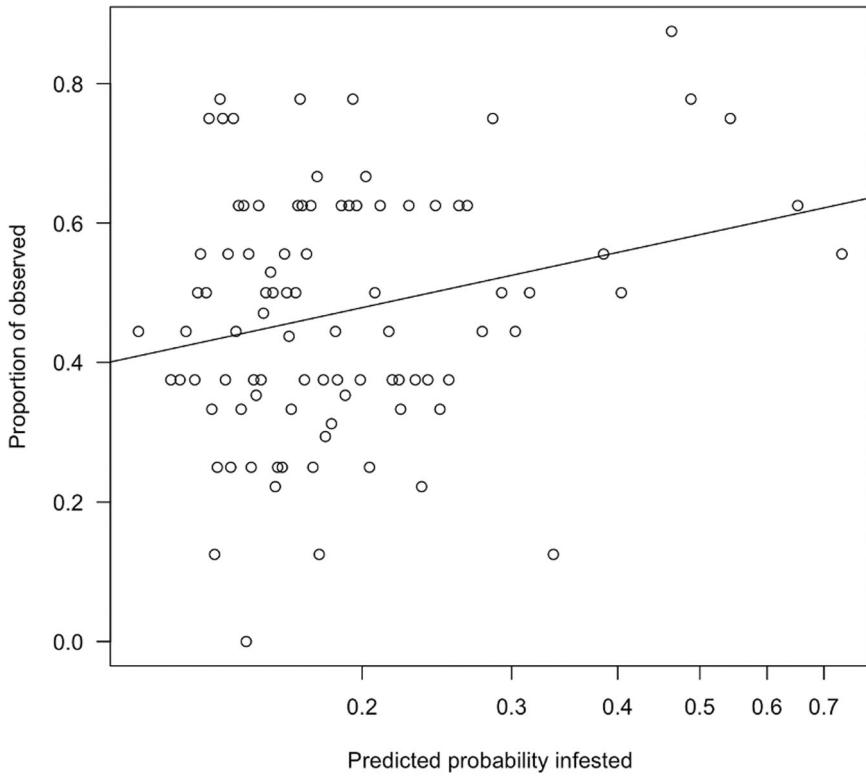


Fig. 5 | Proportion observed grids that were infested by Fusarium dieback—invade shothole borers against predicted probability of being infested. The linear model is significant ($\text{prop}(\text{Observed}) = 8.2 + 2.4 \times \log_{10}(\text{Predicted Prob})$, $P = 0.018$, $R^2 = 0.06$, residual standard error (RSE) = 0.1717). Estimates were based on bins of 0.01 ($n = 101$) based on density-based wpS alone.



amplifies local susceptibility (Fig. 2d, e). Similarly, monitoring or preventative actions could also be prioritized in Palm Springs where microclimate is highly favorable for beetle establishment even though host composition suggests a lower likelihood of infestation (Fig. 2f). This approach mirrors the DAMA protocol (document, assess, monitor, act)

under the Stockholm Paradigm⁴⁴ and lays the groundwork to predicting FD-ISHB establishment under future climate change scenarios.

While climate-based risk mapping systems for pest risk analysis (e.g., CLIMEX, BIOCLIM, GARP, and NAPPFAST) provide an important first pass to setting climate-envelope bounds on future spread, they do not give

land managers the resolution needed to make timely and costly site-specific decisions for host generalists^{43,45–48}. Our finding that the phylogenetic composition of local communities is a better indicator of the likelihood of FD-ISHB establishment than either microclimate or host abundance alone points to the need for more high-quality tree inventory data. Sample-based forest health monitoring initiatives have produced robust and statistically reliable national forest inventory data in many countries (www.afritron.org^{49–52}, including the hundreds of thousands of periodically resampled 672-m² plots in the USA Forest Inventory Analysis program⁴⁹). However, a sample density of one plot per 24.3 km² is inadequate to map tree species distributions for regional pest risk modeling, although species distribution imputation approaches (e.g., LEMMA), can be useful in filling in spatial gaps. In addition, urban forest inventories are underrepresented. The urban forest dataset used in this study includes data from publicly available municipal inventories and inventories from private arborist companies; while it represents the largest inventory of individual urban trees in the United States now publicly available³⁵, the data cover only a small fraction of potential areas for FD-ISHB establishment. Given that global estimates of biological invasions represent a sizable economic burden^{7–10}, anticipating where any current and future pest (e.g., spotted lantern fly) will establish and cause the most damage requires an investment in developing accessible and dependable ground-truthed host distribution datasets.

Importantly, the recent introduction of the pest means that it has not yet spread throughout the entire suitable range. Therefore, areas not currently infested include both those that are not suitable for the pest (i.e., Fig. 3b, wpS < 0.55) and suitable areas that have not yet been reached (i.e., Fig. 3b, wpS > 0.55). This means that our model predictions are both conservative and tentative, but the alternative would be to wait until the pest has reached a global equilibrium across the state (which may never happen). Our purpose here is to show how areas of high susceptibility to a multi-host pest can be identified early in the process, allowing targeted surveillance and strategic management actions. For this reason, such tentative predictions hold value now.

The phylogenetic model here is applicable across systems and can be used as a robust tool to quickly determine the spread of novel pests, even as a first approximation with limited information. At a minimum, it requires a list of known hosts for a pest, to which evolutionary distances to local plant species can be calculated (i.e., p(H), Fig. 1). Because most emergent pests have been introduced from elsewhere, the local host range may not yet be known when decisions on monitoring or actions must be made. In such cases, known hosts from the geographic range of origin can be used for those calculations, with the known-host list augmented and refined as further information on locally susceptible hosts becomes available.

We developed this phylogenetic epidemiology approach to predicting multi-host pest spread out of an urgent management need for accessible decision-making tools that capture the biological realism of pests that are not host specialists^{53,54}. Both range shifts driven by climate change and accidental introductions of these pests into areas outside their place of origin have resulted in novel species interactions that cause irreversible ecosystem changes with consequent ecological, social, and economic harm^{55–57}. Such predictive tools are crucial for developing effective response policies because multi-host pests are more prevalent than host specialists^{2,3,6} and they are responsible for a large proportion of the most threatening emerging infectious diseases to plants, humans, and wildlife^{1,53,58}. Phylogenetic signal in host range is pervasive across different groups of microbial and insect pests and can be used to predict generalist pest host ranges¹⁵. Here, we demonstrate that by extension, the phylogenetic composition of communities is a key predictor for the establishment of a multi-host pest. The phylogenetic ecology tools used in this study can be readily adapted to predict the establishment of any emergent multi-host pest in novel habitats. To calculate phylogenetic community composition and implement this straight forward approach, analysts can generate informative risk estimates if they have a list of known host species, plant species distribution and abundance data (i.e., density or basal area), and a host phylogeny generated using readily available tools^{59–62}. Environmental data that are widely available (e.g.,

PRISM, CHELSA, WORLDCLIM) can then be incorporated to account for key temperature or moisture constraints on pest performance, if known. A critical next step is to integrate landscape composition and configuration factors into these phylogenetically-informed predictive models and to identify the relative importance of local and landscape features driving multi-host pest spread.

Methods

Site selection and plot monitoring

We established a network of 207 monitoring plots (99 infested and 108 non-infested) to measure the effect of the phylogenetic structure of tree communities and environmental conditions on the likelihood that FD-ISHB will establish in a plot, and to capture variation in pest-pathogen spread over time and space. Plots were considered infested when at least one tree within $a \geq 0.25$ ha area was infested by the pest-pathogen complex. Plots were established between July–November 2017 in riparian habitats and oak woodlands in the coastal regions of Southern California, from Ventura County in the north (34.457666, -119.292731) to San Diego County in the south (32.555394, -117.088553) (Fig. 2 and Supplementary Data 1). Plot locations cover the range of environmental conditions in which the beetle species have been observed through extensive surveys by trained experts representing the University of California (UC) Riverside, Santa Cruz, and Davis; UC Cooperative Extension; Orange, San Diego, Los Angeles and Ventura County Agriculture; USDA Forest Service, Forest Health Protection; California Department of Forestry and Fire Protection; Disney; the Huntington Library Art Collections and Botanical Gardens; and the Los Angeles County Arboretum and Botanic Gardens (<https://ucanr.edu/sites/pshb/pest-overview/ishb-fd-distribution-in-california/>). Study sites were located within distinct stream courses or canyons and were chosen based on the presence of suitable tree hosts.

We established one to three monitoring plots within each site, with a minimum of 200 m between plots; vegetation composition often changed noticeably within a few hundred meters. The locations of the monitoring plots within a study site were chosen based on vegetation composition and without regard to presence or absence of the pest (Supplementary Data 1). Plot sizes were variable (0.25–2.75 ha; median = 0.27 ha) to account for variation in tree density among plant communities; plots were a minimum 50 × 50 m but were extended in low-density stands to ensure each plot included at least 50 geo-referenced trees (13, Supplementary Table 1). Plots varied in species composition and phylogenetic distances from the 77 known competent host tree species.

In each plot, we recorded attributes on every tree ≥ 5 cm diameter at standard height (d.s.h., measured at 1.37 m), and included the species and infestation status based on the presence of beetle entry holes. All trees were censused first in 2017 and then a second time in 2018. We measured hourly temperature (°C) and relative humidity (%) using iButton Hygrochron data loggers (Maxim Integrated, San Jose, CA, USA).

All together, the monitoring plots covered 7.6 ha in Ventura County, 50 ha in Orange County, and 22 ha in San Diego County (Supplementary Data 1), with 15,000 trees examined (Supplementary Table 1). Specifically, we established plots in the Oxnard Plain and Santa Clara River Valley, the Western Transverse Ranges, and the western Santa Monica Mountains in Ventura County; the Santa Ana Mountains of the Central Region and the San Joaquin hills, upper Newport Bay, and the Santa Ana River wetlands of the Coastal Region of Orange County (Aerial Information Systems 2015); and the San Luis Rey, Carlsbad, San Marcos, San Dieguito, Penaquitos, San Diego, Sweetwater, Otay, and Tijuana watersheds of the Peninsular Range in San Diego County. The network includes regions with multiple, independent pest introductions, beginning with San Diego (2014) and most recently Ventura County (2016).

Riparian habitats included both ephemerally and perennially flooded stream courses. The stream courses varied in structure, with sandy or gravelly alluvium deposited in canyons that ranged from narrow to fairly broad⁶³. In addition to canyons, riparian habitat in Ventura County occurred in large valley floors of the Oxnard Plain and Santa Clara River

Valley. Riparian forest physiognomy varied from woodland stands with moderately large canopy gaps in the upper tree stratum to forest stands where tree canopies overlapped between individuals so that canopy cover exceeded 100%⁶⁴. Study sites in non-riparian coast live oak woodlands occurred along coastal foothills on north-facing slopes and shaded ravines⁶⁴.

Stands consisted of a diversity of evergreen and winter-deciduous tree species belonging to the California broadleaf forest and woodland and southwestern North American riparian woodland, forest, and wash scrub vegetation groups (Supplementary Figs. 1 and 2)^{65–67}. Combined across all sites, there were a total of 50 tree species from 25 families; 27 species were native to California (Supplementary Table 1). Common species found across most sites included western sycamore (*Platanus racemosa*), arroyo willow (*Salix lasiolepis*), red willow (*S. laevigata*), black willow (*S. gooddingii*), cottonwood (*Populus fremontii*), white alder (*Alnus rhombifolia*), and coast live oak (*Quercus agrifolia*) (Supplementary Table 1). Tree species composition varied from north to south due to the presence or absence of less dominant tree species (Supplementary Table 1). Importance values (IV) were highest for coast live oak (range 27–81%), followed by western sycamore (27–78%) and arroyo willow (25–72%) (Supplementary Table 1). Relative density of coast live oak ranged from 5 to 23%, 19–41% for arroyo willow, and 3–11% for western sycamore (Supplementary Table 1). All stands had a dense to sparse understory of shrubs including elderberry (*Sambucus nigra*) and poison oak (*Toxicodendron diversilobum*); the understory of riparian forests and woodlands also included coyote bush (*Baccharis pilularis*) and mule fat (*B. salicina*) while non-riparian oak woodlands also had toyon (*Heteromeles arbutifolia*) and California sagebrush (*Artemesia californica*).

PhyloEpi model development

We tested how well the observed patterns of plot infestation could be predicted by the phylogenetic composition of the local tree community (Fig. 1). Adapting the *PhyloSusceptibility* model from Parker et al.¹⁵, we first calculated susceptibility (*S*) to the FD-ISHB pest-pathogen complex for tree species *i* as a function of its phylogenetic distance *PD* from a killed-competent host *j* as $\text{logit}(S)_{ij} = 3.38 - 3.68 \times \log_{10}(PD_{ij} + 1)$. The logistic regression coefficients in this equation ($\beta_0 = 3.38$, $\beta_1 = -3.68$) were generated using a resampling approach for the FD-ISHB pest-pathogen complex. Specifically, we randomly selected one killed-competent host species (source) and then used the *PD* from source to each plant species (target) in a logistic regression, where the response variable was 1 (susceptible) or 0 (not susceptible) and *PD* is twice the time to the most recent common ancestor in Myr. Pairwise phylogenetic distances among all tree and shrub species in California were calculated using a dated ultrametric phylogenetic tree developed by Lynch et al.¹³ and the cophenetic function in the R package Picante v. 1.2–0⁶⁶. Logistic regression was repeated for 1000 total runs, with new random selections of source host species for each run. We calculated the median intercept and slope coefficients and the 95% confidence interval across all 1000 runs. If the 95% confidence interval of a coefficient did not include zero, it was considered significant.

Here, the probability that tree species *i* is susceptible ($p(S)_i$) based on its phylogenetic distance from host *j* is $p(S) = \exp[\text{logit}(S)_{ij}] / [1 + \exp[\text{logit}(S)_{ij}]]$. The overall probability that tree species *i* is a host ($p(H)_i$) is the complement of the product of probabilities that tree species *i* does not share FD-ISHB with each of the 18 killed-competent host species *j*: $p(H)_i = 1 - \prod_{j=1}^{18} [1 - p(S)_{ij}]$. This $p(H)_i$ was weighted by their relative abundances (density or basal area; Supplementary Fig. 4) for each species *i* within each plot *k* as $wp(H)_i = p(H)_i * RA_i$. The overall estimate of plot susceptibility $wp(S)_k$ is the sum of the weighted host probabilities within each plot *k*: $wp(S)_k = \sum_{i=1}^n wp(H)_i$.

Estimates of plot susceptibility ($wp(S)_k$) were compared to observed FD-ISHB establishment in plots to evaluate predictable effects of alternative hosts on pest-pathogen establishment using logistic regression and predictive quadratic discriminant analysis (QDA) on our dataset in Orange and San Diego Counties, where the beetle has been established the longest. For

each analysis, those plot infestation data were randomly partitioned equally into “training” and “testing” plot data sets. We used the training data to find parameters for the discriminant model using the qda function in the MASS R package (v. 7.3–51.6), and then applied that model to the test data to predict the status of each plot (infested or not infested) based on discriminant scores. To account for classification variability on randomly partitioned data, this process was repeated 100 times for each analysis on different random sets of training and testing samples. Models at the 0.025, 0.5, and 0.975 quantile classification rates are reported as representative results.

Annual ISHB generation estimates

Sites with microclimatic conditions more often within the optimal temperature range for ISHB development are expected to support more generations of beetles. We used a degree-day model to estimate the annual number of beetle generations that could occur within each plot as an indicator for potential propagule pressure. Observed daily minimum (T_{\min}) and maximum (T_{\max}) air temperatures in each plot were used to calculate the cumulative number of degree-days (CDD) for each day. CDD were calculated using the sine wave model⁶⁷ with ISHB threshold temperatures for beetle emergence ($T_{\min} = 13^{\circ}\text{C}$; $T_{\max} = 32^{\circ}\text{C}$) determined from previous experimental work under laboratory conditions^{33,34}. To estimate the annual number of ISHB generations, the CDD accumulated in 2017 and 2018 was divided by the experimentally derived thermal constant *K* for ISHB (i.e., the number of degree-days required for complete ISHB development)^{33,34}. For subsequent logistical analysis of the likelihood of FD-ISHB establishment ($p(\text{FD-ISHB})$), we included the density-based *wpS* term, the number of expected beetle generations (*Gens*), and an interaction term: $p(\text{FD-ISHB}) = 8.76 \times wpS + 0.5 \times Gens - 0.8 \times (wpS \times Gens) - 8.06$.

PhyloEpi model applications

We applied the *phyloEpi* host composition-based susceptibility model (density-based *wpS*), which was developed and validated with our robust plot monitoring dataset, to predict urban forest susceptibility in California outside the geographic range used to train the models. We estimated susceptibility (density-based *wpS*) for 170 cities in California, which have complete street tree inventories in the California Urban Forest Inventory (CUFI)³⁵ and are distributed across broad geographic regions where the pest-pathogen is currently found (i.e., Southern California) and where it is not (e.g., Palm Springs and Northern California’s Central Valley and San Francisco bay area) (Fig. 2). Individual tree data from the inventory were aggregated to 1-km² grids across the extent of California. A total of 9262 grid cells containing 5,280,301 individual trees in 1037 species were constructed across the spatial extent of the CUFI data (Supplementary Data 2 and 3). Of the total, 2956 grids (32%) were outside the current geographic infested range. The likelihood of FD-ISHB establishment was then estimated using the total number of trees representing each species for each 1-km² grid.

We tested how well *phyloEpi* predicted FD-ISHB establishment within those grids where monitoring efforts outside of our study occurred. Monitoring activities were conducted at least once from 2012–2021 in 832 (9%) of the 9262 urban forest grids used to predict density-based *wpS* across California (Fig. 2). The majority of data were collected during visual tree assessments within artificial boundaries where researchers had permission to survey (e.g., county parks, riparian corridors, university campuses)⁶⁸. As such, the dataset is likely affected by some degree of spatial bias due to uneven sampling effort—a common obstacle in species distribution modeling^{69,70}. Yet, with 19,161 FD-ISHB presence (9202) and absence (9959) records, it represents the most spatially and temporally comprehensive dataset currently available. We compiled FD-ISHB occurrence data from trees surveyed by trained experts representing the University of California (UC) Riverside, Davis, and Santa Cruz; UC Agriculture and Natural Resources Cooperative Extension and Integrated Pest Management; and USDA Forest Service, Forest Health Protection. Point locations for individual trees were collected using hand-held GPS devices. Grids containing confirmed infested trees were assigned “infested”, and those with

only confirmed negative trees or traps outside the infested range were designated “non-infested”. A location was considered “outside the infested range” when it was not within an FD-ISHB infested county.

In addition to community-based estimates of FD-ISHB susceptibility statewide, we ran the aforementioned degree-day model to estimate the mean annual number of beetle generations across 20 years (2001–2020) within 1-km² grids across California. To estimate the annual number of ISHB generations, we ran the model over a data stack giving continuous gridded estimates of daily minimum (T_{\min}) and maximum (T_{\max}) air temperature as a statewide raster for each calendar year. For this data stack, we used Daymet data in NetCDF format (https://daac.ornl.gov/DAYMET/guides/Daymet_Daily_V4.html)^{71–73}. To improve computational performance given the large size of this data stack (46 GB, 365 days × 20 years × 552,000 pixels statewide), we exported the original R script to the Julia language⁷⁴. Annual estimates were averaged across 20 years for each grid to produce a single bioclimatic map of beetle generation estimates (Fig. 2). Thus, rather than extrapolating to areas with different environmental conditions, we worked within the envelope of environmental conditions used to calibrate the model.

Ultimately, we made wpS estimates ($(p(S))_k = -2.15 + 3.26 \times wpS$) across the state only within grids that had complete tree inventory data, and removed the interaction with the effect of temperature on beetle generations (Fig. 2, wpS). We overlaid those estimates onto a separate layer of estimates of the number of beetle generations that could occur within a grid (generations = Cumulative degree days/K_{ISHB}) using Daymet climate data that were available for each grid (Fig. 2, ISHB Generations). As such, we took a conservative approach in our final *phyloEpi* model statewide predictions by separating the effects of phylogenetic community composition and microclimate on beetle reproduction and treating them as separate but complementary factors in a conceptual evaluation of site susceptibility to FD-ISHB establishment in California’s forests.

Statistics and reproducibility

The details about study design and statistics performed in this study are given in the results and methods sections. All analyses were performed using R statistical framework, with functions from the Picante v. 1.2–0, Vegan v. 1.17–8, Hmisc v. 4.3.0, phytools v. 0.6, phangorn v. 2.5.5, Geiger v. 2.0.6.2, caper v. 1.0.1 and Stats v. 2.12.2 packages (<http://cran.r-project.org/>).

Data availability

All source data underlying the analyses and graphs presented in the main figures are available for download as Supplementary Data^{1–9} in this manuscript. Data and accompanying code used in this study can also be downloaded from the Zenodo repository (<https://doi.org/10.5281/zenodo.14600099>)⁷⁵. The 20 year-long raster stack of daily temperature minimum and maximum data to compute ISHB generation estimates within grids across California were obtained from Daymet (https://daac.ornl.gov/DAYMET/guides/Daymet_Annual_V4R1.html).

Code availability

The R (v.4.2.2) and Julia (v.1.5.0) codes used in this study are available at the Zenodo repository (<https://doi.org/10.5281/zenodo.14600099>)⁷⁵.

Received: 6 October 2023; Accepted: 13 January 2025;

Published online: 24 January 2025

References

- Cleaveland, S., Laurenson, M. K. & Taylor, L. H. Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **356**, 991–999 (2001).
- Gilbert, G. S., Magarey, R., Suiter, K. & Webb, C. O. Evolutionary tools for phytosanitary risk analysis: phylogenetic signal as a predictor of host range of plant pests and pathogens. *Evol. Appl.* **5**, 869–878 (2012).
- Woolhouse, M. E., Taylor, L. H. & Haydon, D. T. Population biology of multi-host pathogens. *Science* **292**, 1109–1112 (2001).
- Anderson, R. M. & May, R. M. Population biology of infectious diseases I. *Nature* **280**, 361 (1979).
- Burdon, J. J. & Chilvers, G. A. Host density as a factor in plant disease ecology. *Annu. Rev. Phytopathol.* **20**, 143–166 (1982).
- Young, H. S., Parker, I. M., Gilbert, G. S., Sofia Guerra, A. & Nunn, C. L. Introduced species, disease ecology, and biodiversity-disease relationships. *Trends Ecol. Evol.* **32**, 41–54 (2017).
- Aukema, J. E. et al. Economic impacts of non-native forest insects in the continental United States. *PLoS ONE* **6**, e24587 (2011).
- Diagne, C. et al. High and rising economic costs of biological invasions worldwide. *Nature* **592**, 571–576 (2021).
- Fantle-Lepczyk, J. E. et al. Economic costs of biological invasions in the United States. *Sci. Total Environ.* **806**, 151318 (2022).
- Seebens, H. et al. No saturation in the accumulation of alien species worldwide. *Nat. Commun.* **8**, 14435 (2017).
- Davidson, J. M., Wickland, A. C., Patterson, H. A., Falk, K. R. & Rizzo, D. M. Transmission of *Phytophthora ramorum* in mixed-evergreen forest in California. *Phytopathology* **95**, 587–596 (2007).
- Truitt, L. L., McArt, S. H., Vaughn, A. H. & Ellner, S. P. Trait-based modeling of multihost pathogen transmission: plant-pollinator networks. *Am. Nat.* **193**, E149–E167 (2019).
- Lynch, S. C., Eskalen, A. & Gilbert, G. S. Host evolutionary relationships explain tree mortality caused by a generalist pest-pathogen complex. *Evol. Appl.* **14**, 1083–1094 (2021).
- LoGiudice, K., Ostfeld, R. S., Schmidt, K. A. & Keesing, F. The ecology of infectious disease: effects of host diversity and community composition on Lyme disease risk. *Proc. Natl. Acad. Sci. USA* **100**, 567–571 (2003).
- Parker, I. M. et al. Phylogenetic structure and host abundance drive disease pressure in communities. *Nature* **520**, 542–544 (2015).
- Gilbert, G. S. & Webb, C. O. Phylogenetic signal in plant pathogen-host range. *Proc. Natl. Acad. Sci. USA* **104**, 4979–4983 (2007).
- Janik, K., Panassiti, B., Kerschbamer, C., Burmeister, J. & Trivellone, V. Phylogenetic triage and risk assessment: how to predict emerging phytoplasma diseases. *Biology* **12**, 732 (2023).
- Farrell, M. J. & Davies, T. J. Disease mortality in domesticated animals is predicted by host evolutionary relationships. *Proc. Natl. Acad. Sci. USA* **116**, 7911–7915 (2019).
- Gilbert, G. S., Briggs, H. M. & Magarey, R. The impact of plant enemies shows a phylogenetic signal. *PLoS ONE* **10**, e0123758 (2015).
- Potter, K. M., Escanferla, M. E., Jetton, R. M. & Man, G. Important insect and disease threats to United States tree species and geographic patterns of their potential impacts. *For. Trees Livelihoods* **10**, 304 (2019).
- Pearse, I. S. & Hipp, A. L. Phylogenetic and trait similarity to a native species predict herbivory on non-native oaks. *Proc. Natl. Acad. Sci. USA* **106**, 18097–18102 (2009).
- Pearse, I. S. & Hipp, A. L. Native plant diversity increases herbivory to non-natives. *Proc. Biol. Sci.* **281**, 20141841 (2014).
- Moreira, X. et al. Greater phylogenetic distance from native oaks predicts escape from insect leaf herbivores by non-native oak saplings. *Am. J. Bot.* **106**, 1202–1209 (2019).
- Venette, R. C. et al. Pest risk maps for invasive alien species: a roadmap for improvement. *Bioscience* **60**, 349–362 (2010).
- Stouthamer, R. et al. Tracing the origin of a cryptic invader: phylogeography of the *Euwallacea fornicatus* (Coleoptera: Curculionidae: Scolytinae) species complex. *Agric. For. Entomol.* **19**, 366–375 (2017).
- Freeman, S. et al. *Fusarium euwallaceae* sp. nov.—a symbiotic fungus of *Euwallacea* sp., an invasive ambrosia beetle in Israel and California. *Mycologia* **105**, 1595–1606 (2013).
- Boland, J. M. The impact of an invasive ambrosia beetle on the riparian habitats of the Tijuana River Valley, California. *PeerJ* **4**, e2141 (2016).

28. Mendel, Z. et al. The role of *Euwallacea* nr. *Fornicatus* (Coleoptera: Scolytinae) in the wilt syndrome of avocado trees in Israel. *Phytoparasitica* **45**, 341–359 (2017).
29. Lynch, S. C., Gilbert, G. S., & Eskalen, A. *Management and Monitoring of Fusarium Dieback – Invasive Shot Hole Borers Complex* (California Department of Fish and Wildlife, Final Report, 2019).
30. Paap, T., de Beer, Z. W., Migliorini, D., Nel, W. J. & Wingfield, M. J. The polyphagous shot hole borer (PSHB) and its fungal symbiont *Fusarium euwallaceae*: a new invasion in South Africa. *Australas. Plant Pathol.* **47**, 231–237 (2018).
31. Eskalen, A. et al. Host range of Fusarium dieback and its ambrosia beetle (Coleoptera: Scolytinae) vector in Southern California. *Plant Dis.* **97**, 938–951 (2013).
32. International Plant Protection Convention. *Euwallacea fornicatus* (polyphagous shot- hole borer, PSHB) in Western Australia. Pest Report AUS-105/1. <https://www.ippc.int/en/countries/australia/pestreports/> (2021).
33. Dodge, C. & Stouthamer, R. Effect of temperature on fecundity, development, and emergence of the invasive ambrosia beetle *Euwallacea kuroshio* (Coleoptera: Scolytinae). *Agric. For. Entomol.* **23**, 79–86 (2021).
34. Umeda, C. & Paine, T. Temperature can limit the invasion range of the ambrosia beetle *Euwallacea* nr. *fornicatus*. *Agric. For. Entomol.* **21**, 1–7 (2019).
35. Love, N. L. R. et al. Diversity and structure in California’s urban forest: what over six million data points tell us about one of the world’s largest urban forests. *Urban For. Urban Green.* **74**, 127679 (2022).
36. Srivastava, V., Roe, A. D., Keena, M. A., Hamelin, R. C. & Griess, V. C. Oh the places they’ll go: improving species distribution modelling for invasive forest pests in an uncertain world. *Biol. Invasions* **23**, 297–349 (2021).
37. Donatelli, M. et al. Modelling the impacts of pests and diseases on agricultural systems. *Agric. Syst.* **155**, 213–224 (2017).
38. Hudgins, E. J., Liebhold, A. M. & Leung, B. Predicting the spread of all invasive forest pests in the United States. *Ecol. Lett.* **20**, 426–435 (2017).
39. Jones, C. M. et al. Iteratively forecasting biological invasions with PoPS and a little help from our friends. *Front Ecol. Environ.* **19**, 411–418 (2021).
40. Jones, C. et al. Spotted lanternfly predicted to establish in California by 2033 without preventative management. *Commun. Biol.* **5**, 558–567 (2022).
41. Koch, F. H., Yemshanov, D., Haack, R. A. & Magarey, R. D. Using a network model to assess risk of forest pest spread via recreational travel. *PLoS ONE* **9**, e102105 (2014).
42. Yemshanov, D., Koch, F. H., Ben-Haim, Y. & Smith, W. D. Detection capacity, information gaps and the design of surveillance programs for invasive forest pests. *J. Environ. Manag.* **91**, 2535–2546 (2010).
43. Magarey, R. D. et al. NAPPFAST: an internet system for the weather-based mapping of plant pathogens. *Plant Dis.* **91**, 336–345 (2007).
44. Brooks, D. R., Hoberg, E. P., Boeger, W. A. *The Stockholm Paradigm: Climate Change and Emerging Disease*. <https://doi.org/10.7208/chicago/9780226632582.001.0001> (University of Chicago Press, 2019).
45. Baker, R. H. A. Predicting the limits to the potential distribution of alien crop pests. In *Invasive Arthropods in Agriculture* (eds. Hallman, G. J. & Schwalbe, C. P.) 208–241 (Science Publishers, Enfield, NH, 2002).
46. Elith, J. et al. Novel methods improve prediction of species’ distributions from occurrence data. *Ecography* **29**, 129–151 (2006).
47. Kriticos, D. et al. *CLIMEX Version 4: Exploring the Effects of Climate on Plants, Animals and Diseases* (CSIRO Publishing, Canberra Australia, 2016).
48. Worner, S. P. Predicting the invasive potential of exotic insect pests. In *Invasive Arthropods in Agriculture* (eds. Hallman, G. J. & Schwalbe, C. P.) 119–137 (Science Publishers, Enfield, NH., 2002).
49. Chirici, G. et al. National Forest Inventory contributions to forest biodiversity monitoring. *For. Sci.* **58**, 257–268 (2012).
50. Gillis, M. D., Omule, A. Y. & Brierley, T. Monitoring Canada’s forests: the national forest inventory. *For. Chron.* **81**, 214–221 (2005).
51. McRoberts, R. E., Bechtold, W. A., Patterson, P. L., Scott, C. T. & Reams, G. A. The enhanced forest inventory and analysis program of the USDA forest service: historical perspective and announcement of statistical documentation. *J. For.* **103**, 304–308 (2005).
52. Ramírez, C., Alberdi, I., Bahamondez, C., Veloso de Freitas, J. *National Forest Inventories of Latin America and the Caribbean* <https://doi.org/10.4060/cb7791en> (FAO, 2022, accessed 19 February 2023).
53. Fountain-Jones, N. M. et al. Towards an eco-phylogenetic framework for infectious disease ecology. *Biol. Rev. Camb. Philos. Soc.* **93**, 950–970 (2018).
54. Gougherty, A. V. & Davies, T. J. Towards a phylogenetic ecology of plant pests and pathogens. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **376**, 20200359 (2021).
55. Lovett, G. M., Canham, C. D., Arthur, M. A., Weathers, K. C. & Fitzhugh, R. D. Forest ecosystem responses to exotic pests and pathogens in Eastern North America. *Bioscience* **56**, 395–405 (2006).
56. Lovett, G. M. et al. Nonnative forest insects and pathogens in the United States: impacts and policy options. *Ecol. Appl.* **26**, 1437–1455 (2016).
57. Morales-Castilla, I. et al. Forecasting parasite sharing under climate change. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **376**, 20200360 (2021).
58. Daszak, P., Cunningham, A. A. & Hyatt, A. D. Emerging infectious diseases of wildlife—threats to biodiversity and human health. *Science* **287**, 443–449 (2000).
59. Jin, Y. & Qian, H. V. PhyloMaker: an R package that can generate very large phylogenies for vascular plants. *Ecography* **42**, 1353–1359 (2019).
60. Smith, S. A. & Brown, J. W. Constructing a broadly inclusive seed plant phylogeny. *Am. J. Bot.* **105**, 302–314 (2018).
61. Webb, C. O. & Donoghue, M. J. Phylomatic: tree assembly for applied phylogenetics. *Mol. Ecol. Notes* **5**, 181–183 (2005).
62. Zanne, A. E. et al. Three keys to the radiation of angiosperms into freezing environments. *Nature* **506**, 89–92 (2014).
63. Sands, A. (ed) *Riparian Forests in California: Their Ecology and Conservation: A Symposium* (Division of Agricultural Sciences, University of California, 1980).
64. USNVC (United States National Vegetation Classification) Database. Federal Geographic Data Committee, Vegetation Subcommittee. Washington D.C. <https://usnvc.org/> (2019).
65. Sawyer, J. O., Keeler-Wolf, T., & Evans, J. A. *Manual of California Vegetation* (California Native Plant Society Press, 2009).
66. Kembel, S. W. et al. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* **26**, 1463–1464 (2010).
67. Higley, L. G., Pedigo, L. P. & Ostlie, K. R. Degday: a program for calculating degree-days, and assumptions behind the degree-day approach. *Environ. Entomol.* **15**, 999–1016 (1986).
68. Umeda, C., Eskalen, A. & Paine, T. D. Polyphagous shot hole borer and fusarium dieback in California. In *Insects and Diseases of Mediterranean Forest Systems* (eds. Paine, T. D. & Lieutier, F.) 757–767 (Springer International Publishing, 2016).
69. Rushton, S. P., Ormerod, S. J. & Kerby, G. New paradigms for modelling species distributions? *J. Appl. Ecol.* **41**, 193–200 (2004).
70. Newbold, T. Applications and limitations of museum data for conservation and ecology, with particular attention to species distribution models. *Prog. Phys. Geogr.: Earth Environ.* **34**, 3–22 (2010).
71. Thornton, P. E., Running, S. W. & White, M. A. Generating surfaces of daily meteorological variables over large regions of complex terrain. *J. Hydrol.* **190**, 214–251 (1997).

72. Thornton, M. M. et al. *Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 4*. <https://doi.org/10.3334/ORNLDaac/1840> (ORNL DAAC, Oak Ridge, Tennessee, USA, 2020).
73. Thornton, P. E. et al. Gridded daily weather data for North America with comprehensive uncertainty quantification. *Sci. Data* **8** (2021).
74. Bezanson, J., Edelman, A., Karpinski, S. & Shah, V. B. Julia: a fresh approach to numerical computing. *SIAM Rev.* **59**, 65–98 (2017).
75. Lynch, S. C., Hollander, A. D. & Gilbert, G. S. Analysis of a phylogenetic epidemiology approach to predicting the establishment of multi-host plant pests [Data set]. Zenodo <https://doi.org/10.5281/zendodo.14600099> (2025).

Acknowledgements

This research was funded by the San Diego Association of Governments Land Management Grant (5004987 & 5005757); USDA APHIS Farm Bill (AP19PPQS&T00C242 & AP18PPQS&T00C162); California Department of Fish and Wildlife Local Assistance Grant (in partnership with the Natural Communities Coalition of Orange County; OC Parks; and Irvine Ranch Conservancy) (17-01-NCC); California Avocado Commission (65119 & 58245); California Department of Food and Agriculture Specialty Crop Block Grant (CDFASCB16051); The Nature Conservancy (P102283); The California Native Plant Society; Southern California Botanists; Sigma Xi; Sea and Sage Audubon Society; and Los Angeles Center for Urban Natural Resources and Sustainability. Special thanks to Milan Mitrovich (Natural Communities Coalition of Orange County), Kristine Preston and Yvonne Moore (San Diego Management and Monitoring Program), and Zack Principe (The Nature Conservancy), who did the heavy lifting to help get the FD-ISHB epidemiology and control research funded. We thank Akif Eskalen, Tom Atkinson, Richard Stouthamer, Paul Rugman-Jones, John Kabashima, Ben Faber, Hannah Walchak, Nathan Gregory and Jutta Burger for their help in facilitating site access and for insightful discussions throughout the course of this research. The following land managers and park rangers kindly provided us with access to their sites to conduct this research Richard Schaffer, Candice Hubert, Steve Aleshire, Brad Barker, Barbara Norton, Matthew Stegner, Steve Jax, Travis Martz, Jim Simkins; Lana Nguyen and Riley Pratt from Crystal Cove State Park; San Diego Park and Recreation Rangers including Nick Sloan, Daniel Meyers, Rusty Rodes, Mike Verberder, Lauren Raduc, Maureen Abare-Laudy, Kyle Icke, Patricia Heyden, Lauren Raduc, Ketsela Mengistu, Alejandro Santos, Gabrielle Widseth, John Barone, Kelly Hahs; California Department of Fish and Wildlife including Gabriel Penaflor, Hans Sin, and Tracie Nelson; Jim Park from Sycuan Reservation; Markus Spiegelberg from CNLM; Rush Abrams from Caltrans; Sara Allen from San Diego City Parks; Anna Huber from Conejo Open Space Conservation Agency; Theresa Lubin from Ventura County Parks and Recreation; Brian Stark from Ojai Valley Land Conservancy; Kat Selm from The Nature Conservancy in Ventura County. Special thanks to E. Aiello, S.

Barasi, A. Conover, J. Detka, S. Grove, E. Lai, N. Lustenhouwer, M. Melen, I. Parker, and S. Philpott for their critical comments on this manuscript.

Author contributions

S.C.L. and G.S.G. conceptualized the research and developed the methodology; S.C.L., E.R.-G., E.B., and K.S.A., performed the research; S.C.L. performed the data analysis with contributions from G.S.G.; N.L.R.L. compiled and provided urban street tree inventory data; A.D.H. generated the statewide raster for the degree-day model; B.N.B. collected, compiled, and provided statewide monitoring data.

Competing interests

The authors declare no competing interest.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42003-025-07540-y>.

Correspondence and requests for materials should be addressed to Shannon Colleen Lynch.

Peer review information *Communications Biology* thanks Chris Jones and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Hannes Schuler and David Favero.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025