



A global analysis of tree pests and emerging pest threats

Andrew V. Gougherty^{a,1} and T. Jonathan Davies^{a,b}

Edited by Daniel Simberloff, University of Tennessee at Knoxville, Knoxville, TN; received July 26, 2021; accepted February 7, 2022

Tree pests affect millions of hectares of natural and managed land annually, but we often lack a strong understanding of the factors limiting pest distributions and the drivers that facilitate the expansion of pests outside their hosts' native ranges. Here, we use hierarchical Bayesian regression models to identify the key determinants of pest distributions from a global dataset of >310,000 pest presences/absences across 206 countries and an additional >120,000 pest occurrences outside the native host ranges to validate the model. Our results show there are strong, generalizable controls on pest ranges, including effects of host richness and phylogeny, geography, and climate. Remarkably, our model fit to pest distributions in native host ranges was able to predict pest presences outside the host native range with ~79% accuracy. Our work has important implications for predicting regions that may be vulnerable to future pest invasions and understanding the accumulation of pests outside the native ranges of their hosts.

nonnative species | tree diseases and insects | species ranges | biogeography

Tree pathogens and insect pests, hereafter pests, play important roles in natural and managed systems, contributing to carbon cycling and forest regeneration, and may even contribute to shaping patterns of global tree diversity (1). However, some pests, especially those that are introduced outside their native range, can have severe impacts on agricultural crops and natural tree populations, resulting in billions of dollars of economic and ecological damage annually (2). Impacts of many pests, native and invasive, are likely to be exacerbated by climate change, which may allow additional pest life cycles per annum and provide more hospitable climates for pests for a longer part of the growing season (3, 4). Despite their immense economic and ecological importance, we lack a strong understanding of what limits the current global geographic extent of most tree pests. Constructing robust models that allow us to project pest geographic ranges through space and time is thus a critical step toward predicting future pest invasion and maintaining the ecological and economic integrity of natural and managed systems.

Modeling pest ranges, however, is more challenging than modeling the ranges of their hosts. Unlike free-living organisms, biotrophic pests require one or more hosts to persist in a region (5), but the ability of a pest to infect its hosts, reproduce, or spread may be limited by additional biotic and abiotic factors. For instance, regions where hosts are rare or distantly related to the surrounding vegetation could present a type of biotic resistance to pest establishment (6), especially as pest host ranges are often phylogenetically circumscribed (7–10). Like other organisms, pest distributions may also be limited by climate and geography, either directly or indirectly via the distribution of hosts. In addition, unlike their tree hosts, which tend to be large in size with visually identifiable characteristics, many pests are particularly vulnerable to undersampling, especially those that require specialty expertise or technology to accurately identify and diagnose. Our current knowledge of pest occurrences is thus likely skewed toward developed countries, which have resources to invest in pest surveillance and reporting (11). However, wealthy countries with high levels of imports from abroad may also tend to accumulate more pests, due to increased propagule pressure and opportunity for pest establishment (12). To accurately model pest distributions we must therefore integrate information on pest–host associations, host phylogeny, community structure, climate, and socioeconomic factors linked to detection probability and introduction pathways within a common statistical framework. Parameterizing such a model demands massive amounts of pest and host data, which until recently has not been readily available.

Here, we assembled a global dataset of >310,000 pest observations (presences and inferred absences) across 206 countries, to model pest distributions in their native host ranges (Fig. 1). We then use this model to predict pest observations in their nonnative host range using a separate set of >120,000 pest occurrences. We show that after accounting for sampling effects, our model can predict pest presences and absences within and outside the host native range with high accuracy, suggesting that the biotic,

Significance

The introduction of trees outside their native ranges has greatly expanded the potential ranges of their pathogens and insect pests, which risk spilling over and impacting native flora. However, we often lack a strong understanding of the host, climatic, and geographic factors that allow pests to establish outside their hosts' native ranges. Using global datasets of pest occurrences and the native and nonnative ranges of tree hosts, we show there are strong generalizable trends controlling pest occurrences and can predict the occurrence of pests outside their hosts' native ranges with >75% accuracy. Our modeling framework offers a powerful tool to identify future invasive pest species and the ecological mechanisms controlling the accumulation of pests outside their hosts' native ranges.

Author affiliations: ^aDepartment of Botany, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada; and ^bDepartment of Forest and Conservation Sciences, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada

Author contributions: A.V.G. and T.J.D. designed research; A.V.G. performed research; A.V.G. analyzed data; and A.V.G. and T.J.D. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2022 the Author(s). Published by PNAS. This article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

¹To whom correspondence may be addressed. Email: andy.gougherty@gmail.com.

This article contains supporting information online at <http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2113298119/-/DCSupplemental>.

Published March 21, 2022.

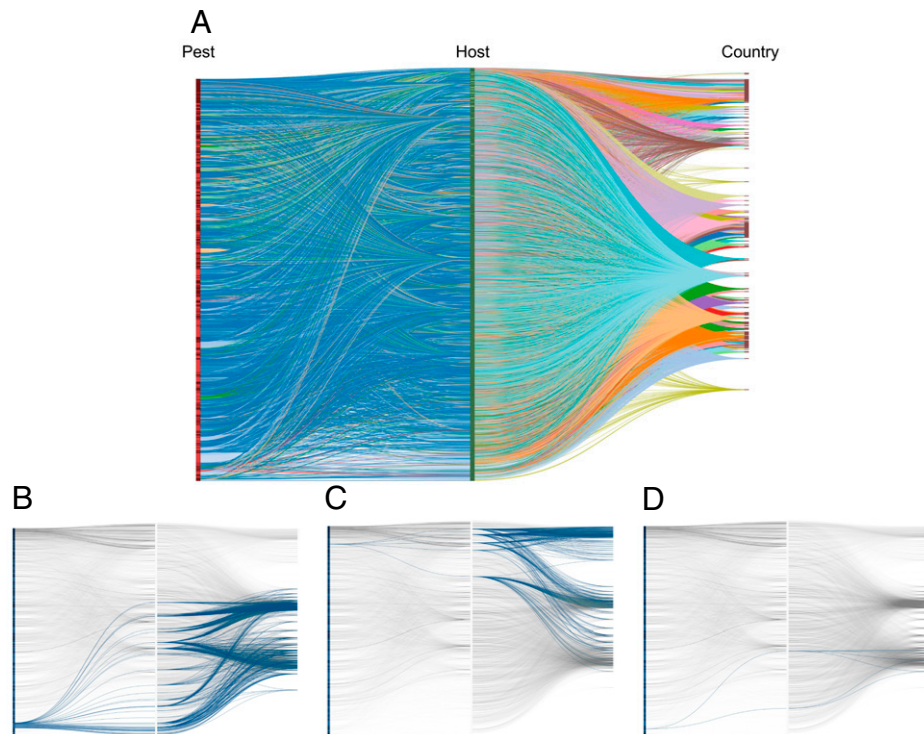


Fig. 1. Sankey diagram of pests, hosts, and countries used in our Bayesian regression of pest \times country occurrences. (A) Pests (Far Left) are linked to their hosts (Center), which are linked to their native countries (Far Right). Pest–host links are colored by pest types (i.e., arthropods, nematodes, molluscs, fungi, bacteria, virus, chromist, protists, and unknown microorganisms), and host–country links are colored by country. Smaller panels, below, are subnetworks for (B) *Austropuccinia psidii*, myrtle rust; (C) *Agrilus planipennis*, emerald ash borer; and (D) *Phoracantha semipunctata*, eucalyptus longhorned borer, where gray lines represent the underlying network in A, and blue lines are observed pest–host associations and host native ranges. Note that, despite having similar narrow host ranges (e.g., C and D), the native distribution of hosts varies dramatically.

abiotic, and socioeconomic constraints on pest occurrences are similar within and outside their native host ranges. Our statistical framework represents a major advance to understanding the global distribution of tree pests and allows us to make projections on future pest invasion risks.

Results

Pest Occurrences within and outside the Native Host Range.

Individual pests tended to occur in a relatively few number of countries where their host genera are native ($\sim 13\%$ of countries), and occurrences were most common in population-rich and/or geographically large countries, such as the United States (1,365 observed pest occurrences), India (1,074), China (1,070), Italy (855), and Japan (737). Countries with the largest number of inferred absences (i.e., host trees are present, but pests have not been reported) were typically geographically small and biased toward southern Asia, including Nepal (2,862 inferred absences), Bhutan (2,773), Myanmar (2,771), Vietnam (2,612), and Laos (2,557).

On average, host genera tended to have more pests in their native range than in their nonnative range (for the 438 host genera with pests in their native and nonnative range, there was a mean 15.76% fewer pests in the nonnative range). This was true even for hosts grown widely outside their native ranges for agriculture, forestry, or as ornamentals, and/or considered widely invasive (Fig. 2). *Pinus*, for instance, native to the Northern Hemisphere but grown extensively in the Southern Hemisphere as a forestry tree (and sometimes an aggressive invasive) (13), has hundreds more pests in its native range than its nonnative range. The same is true for other widely grown or invasive trees, including *Malus*, *Populus*, and *Prunus* (Fig. 2). Nonetheless, we observed a substantial

number of hosts ($\sim 22\%$) with a greater number of pests in their nonnative ranges than their native ranges. A notable example is *Eucalyptus*, which includes the most widely grown plantation tree species, native to Australia and the surrounding islands, but also grown in over 100 countries around the globe.

Drivers of Pest Occurrences in Host Native Ranges.

Our hierarchical Bayesian regression of pest occurrences in host native ranges revealed tree pests are most likely to occur in countries adjacent to other countries where the pest has been reported. Pests also tended to occur in countries with a greater richness of their hosts, especially when these host trees have a close phylogenetic affinity to the rest of the native tree flora within countries (Fig. 3 and *SI Appendix*, Table S1). Climate was an important predictor of pest occurrences, but effects tended to be more muted. In general, across their hosts' native ranges pests were more likely to occur in countries that were on average warmer and wetter. Countries with high research output (measured as the number of citable documents), also tended to have more pests, most likely due to their having greater access to resources for pest surveillance, diagnosis, and reporting. However, number of citable documents also covaried closely with volume of international trade, and in a model where the number of citable documents was replaced with the value of imported goods (these two variables were strongly correlated across countries, when log transformed $r = 0.84$), countries with greater trade had an increased probability of pest presences.

Arthropods had narrower geographic ranges than other pest types (except protists and those of unknown etiology), while molluscs tended to have the widest ranges. Per capita gross domestic product (GDP) and country area had little effect on pest occurrences. Similarly, mean human footprint (a metric of

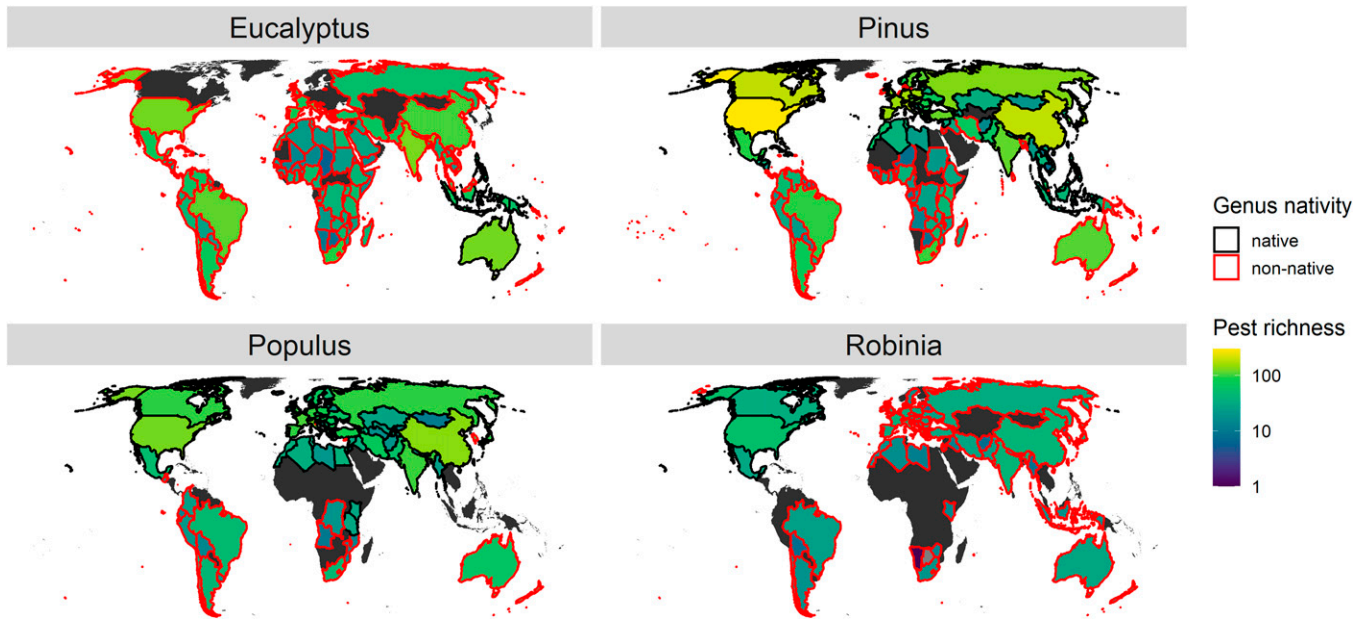


Fig. 2. Pest richness of *Eucalyptus*, *Pinus*, *Populus*, and *Robinia* within and outside their respective native ranges (\log_{10} scale). Gray areas are regions where either the host genus has not been documented in the Global Naturalized Alien Flora (40) or CABI Invasive Species Compendium (41) or no pests of the genus have been reported.

human pressure on the local environment) within countries was also not an important predictor of pest occurrence, but human footprint variability (i.e., countries with large variability in human pressure on the environment) had a small positive effect, that is, pests were more likely to occur in countries with a mixture of natural and human-impacted habitats.

Drivers of Pest Occurrence outside Their Native Host Ranges and Model Predictive Accuracy. After adjusting for the effect of sampling (by fixing the number of citable documents published in countries to the global maximum), our model was

able to correctly classify pest presences and absences within their native host range with high accuracy (true presence rate = 0.83, true absence rate = 0.81; *SI Appendix, Table S2*). Our model performed almost as well at predicting pest occurrences outside the native range of their hosts, i.e., occurrences not used in model fitting (true presence rate = 0.79, true absence rate = 0.80). When assessed within individual countries, classification error was low in nearly all countries examined (Fig. 4). Thresholding predictions to binary presence or absence can bias test statistics (14), but model performance was also strong when evaluating pest occurrences as a continuous probability

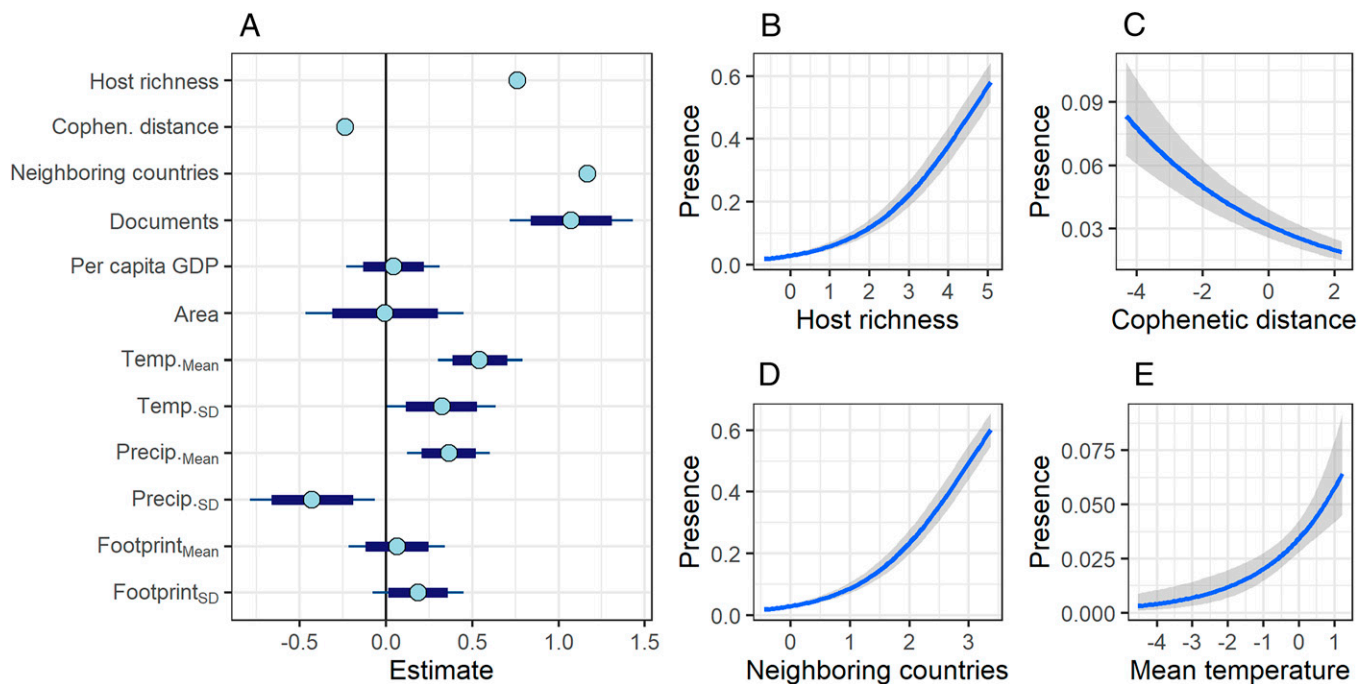


Fig. 3. (A) Posterior parameter distribution from a Bayesian regression of pest occurrences and the marginal effects of (B) host richness, (C) cophenetic distance between hosts and native tree flora, (D) proportion of neighboring countries with the pest, and (E) mean temperature. Variables were scaled (mean of 0 and SD of 1.0) and log transformed (except mean temperature). In A, points are means, bolder lines encompass 80% of values, and thinner lines encompass 95%. Categorical pest type variables are not shown. See also *SI Appendix, Table S1 and Fig. S1*.

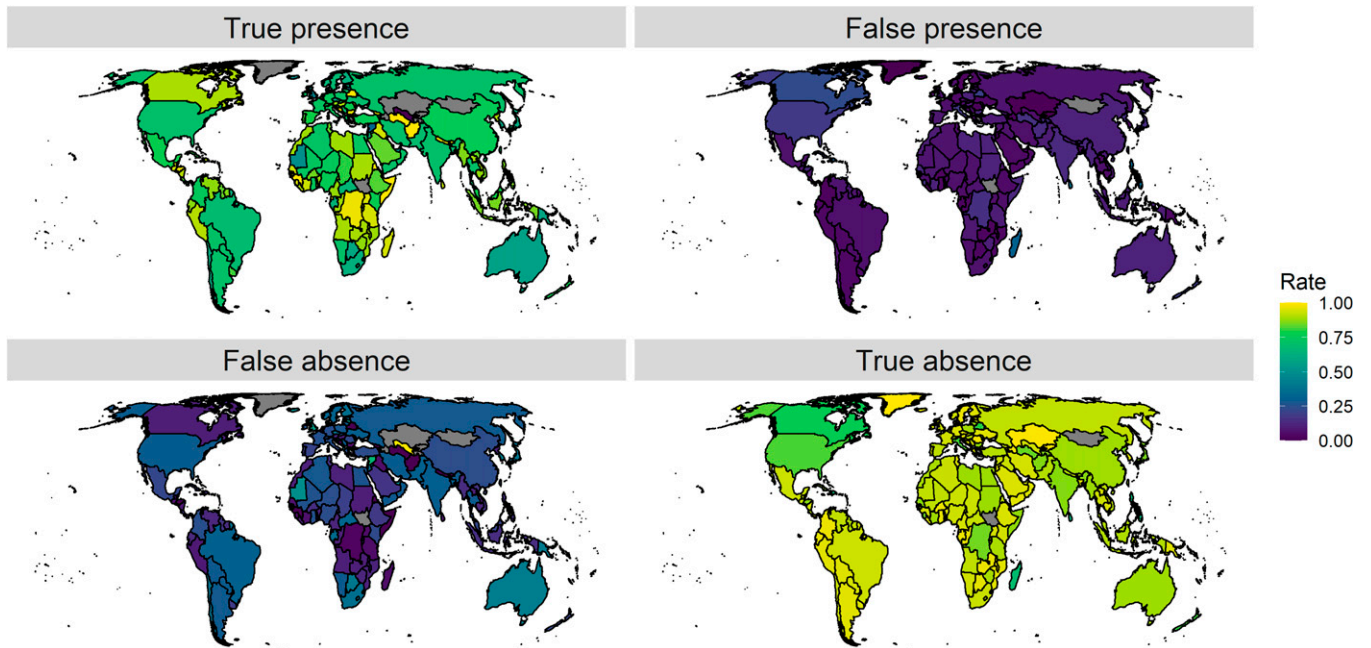


Fig. 4. Geographic pattern of validation statistics for predicting pests outside their native host ranges. Statistics were calculated based on the presence and inferred absence of 3,080 tree pests outside their hosts' native range. True presences are pest occurrences that were correctly predicted present within countries by the model; false presences are pest occurrences that were incorrectly predicted as present; false absences are pest occurrences that were incorrectly predicted as absent; and true absences are absences that were correctly predicted as absent. The false presence rate is particularly important as it indicates the regions where the model predicts the greatest relative increase in pest occurrences through new documentations or invasions. Countries colored gray have either no recorded nonnative plants or no pests of nonnative plants.

(area under the receiver operating curve [AUC] = 0.88, Boyce index = 1.0).

While overall predictive accuracy of our model was high, ~20% of pests documented in countries outside the native range of their hosts were misclassified (i.e., false absences). Many of the misclassified pests (>35%) were known in only one or two countries in the host native ranges (i.e., two or fewer presences were used in model training, for these pests), despite, in some cases, hosts being widely distributed (mean \pm SD native host range = 68.4 ± 46.3 countries). Imbalance between presences and absences in the native host range likely results in a low predicted probability of occurrence regardless of other modeled predictors. Although the false absence rate varied by country, there were no strong geographic trends, suggesting we are not systematically misclassifying pests geographically.

Our primary goal was to predict pest presences (as opposed to absences); nonetheless, our model also predicted absences with reasonably high accuracy. The marginally higher classification error for absences than presences (~20% of pest absences in our dataset were predicted as presences), however, should not necessarily be regarded as a failure of the model. Globally, pests are not at equilibrium with their hosts, and the vast majority of pests occur in only a small portion of their host native ranges. Furthermore, tree pests continue to establish in new regions across the globe (15, 16), a process that is expected to persist into the foreseeable future (17). Thus, some apparent misclassifications may represent regions where hosts are present; the tree community is sufficiently closely related to known hosts; climate is suitable; and the pest is present in nearby countries, but the pest has either not yet been introduced or reported.

Discussion

By analyzing a database of over a quarter million occurrence records of >3,000 tree pests, we were able to predict with high accuracy the occurrence of tree pests within and outside their

hosts' native ranges. Our results indicate that within the native range of their hosts, pests tend to occur in warmer and wetter countries with a large number of known host trees, especially when these hosts are closely related to native flora. Our ability to identify the factors limiting pest host ranges, and accurately predict pest occurrences outside the native host range, illustrates the potential of our modeling framework to identify pests most likely to invade new regions (and the regions most susceptible to future pest invasions).

Factors Controlling Pest Occurrences. Within the native ranges of their hosts, pests tended to occur in countries with a large number of their hosts, especially when those hosts were closely related to the native tree flora of countries. The link between pest occurrences and host richness and evolutionary affinity could have multiple explanations. In regions where pests are native, pests may accumulate hosts over evolutionary time as new pest variants emerge capable of overcoming the defenses of host plants. If host defenses are phylogenetically conserved, this might also structure the expansion of pest host ranges to closely related tree hosts. Alternatively, pests may accumulate hosts as pests spread to new regions within their host native ranges and come into contact with previously isolated, but competent novel host species, perhaps close relatives of pests' coevolved hosts. If the latter mechanism is common, shifts in pest ranges in response to global change is likely to facilitate the further expansion of pest host breadths. The link between pest occurrences and host richness could also result from pests invading regions where (many) hosts are already present. For example, generalist invasive pests, such as spongy moth (*Lymantria dispar dispar*) and winter moth (*Operophtera brumata*) (both native to Eurasia), have accumulated hundreds of novel hosts in their invaded North American range, likely due to these pests being preadapted to the large number of related tree taxa North America shares with Eurasia. This sort

of phylogenetic facilitation likely contributes to the spread of nonnative pests to regions where coevolved host species are absent.

Although we were unable to unambiguously separate the effects of sampling from trade in our models (as both tend to be high in wealthy countries), these variables represent distinct, but not mutually exclusive, factors shaping pest occurrences. Our index of citable documents provides a measure of detection probability and reporting, but it may also capture additional socioeconomic variables correlated with pest introduction. Trade has also frequently been found to have a strong effect on the accumulation of nonnative pests, especially agricultural imports from climatically similar regions (12). While the effect of trade on nonnative pests is relatively well established, it is less clear whether trade should have as direct an impact on the distribution of pests in their native host ranges (used in model fitting), which includes regions where hosts and pests have historically interacted. This dependency could explain why our sampling variable (the number of citable documents) had a stronger effect on pest presences. While we were able to predict pest occurrences outside their hosts' native ranges with high accuracy with both documents and trade, future work that allowed us to distinguish between the effects of trade and sampling could improve our understanding of the importance of dispersal in limiting global pest distributions.

Accurately predicting pest occurrences outside the hosts' native range helps elucidate the ecological dynamics that follow plant invasions and intentional introductions. For instance, many trees that are purposefully grown outside their native range are able to escape the pests of their native ranges that limit their growth or production—the enemy release hypothesis (18). Our results provide some support for enemy release, with most hosts experiencing lower pest pressure in their introduced ranges. However, enemy release tends to be ephemeral as pests from the native range can eventually track their hosts to new regions, and native pests may expand their host breadth to include the novel nonnative host (19). For instance, *Eucalyptus* has accumulated more pests, on aggregate, outside its native range than are documented in its native distribution. While the greater number of *Eucalyptus* pests in the nonnative range could be explained by undersampling in the native range, the center of *Eucalyptus* native diversity is Australia, one of the wealthiest countries globally, and likely far better sampled than many countries where *Eucalyptus* has been introduced. More likely, exporting *Eucalyptus* around the globe has exposed it to many additional pests not present in the native range and for which it may have no evolved resistance (20). Thus, while few individual countries had more *Eucalyptus* pests than Australia, the total pest load is greater outside the native range. (Although it is worth noting that many herbivores/pathogens with minimal impact on *Eucalyptus* in the native range may not be considered pests, per se.) Our finding indicates that there are many pests that could pose a threat to natural and managed *Eucalyptus* forests if introduced into Australia or other parts of its native range (21).

An important future challenge is to elucidate the geographic and environmental influences that facilitate host jumps during invasions, although data distinguishing novel or recent hosts from historic or coevolved hosts is sparse for most pests. Further parsing of pest distributions to differentiate between introduced versus historic geographic ranges could help to identify not only where pests may be likely to establish in the future, but which hosts may be most at risk from yet-to-be

established pests. While identifying the mechanisms that allow both geographic and host jumps could further improve our understanding of the regions and hosts most at risk from nonnative pests, our model nevertheless provides strong inferences on the principal drivers of pest spread, as illustrated by our ability to predict pest occurrences outside host native ranges.

Identifying the pests most likely to track their hosts outside the host native range can help to protect not only high-value nonnative plants (agricultural, ornamental, and forestry species), but also, by extension, native plants that could be at risk from pest spillover. Many of the most destructive nonnative pests of native North American trees were introduced on nonnative live plants imported from abroad that subsequently jumped to native species (e.g., balsam and hemlock woolly adelgids [*Adelges piceae* and *A. tsugae*] and beech scale [*Cryptococcus fagisuga*] (22)). One notable example, *Cryphonectria parasitica* (causal agent of chestnut blight), was first introduced into North America likely on Japanese chestnut (*Castanea crenata*) nursery stock and spilled over to native trees, leading to the effective extirpation of mature American chestnuts (*C. dentata*) in North America. Pests tracking their hosts to the nonnative range can also be detrimental when nonnative hosts are economically valuable. Pitch canker of pines (caused by *Fusarium circinatum*), for instance, likely native to Mexico, has been exported to many of the pine-growing regions of the world where it has caused economic losses in nonnative pine plantations and nurseries (23, 24).

Identifying pests before their introduction, however, remains challenging as many destructive invasive pests may not be damaging within their native ranges. Indeed, many nonnative pests are, at least initially, cryptogenic. Swiss needle cast (caused by *Phaeocryptopus gaeumannii*), for instance, first described in the early 20th century in Germany and Switzerland in planted Douglas fir plantations (*Pseudotsuga menziesii*), was not known to cause disease in its native North American range until the 1970s (25), decades after its description in Europe. The false presences identified in our model clearly illustrate many pests have suitable hosts (native and nonnative) and climate in countries where they have not yet been reported. These “misclassifications” may thus provide a useful guide for identifying pests that could become invasive in the future, or alternatively, could already be present although are as yet unreported, perhaps because current impacts are low.

While we are able to highlight pests with potential to become invasive, and regions vulnerable to future invasions, the successful establishment of a pest in a new region is likely mediated by multiple additional factors that we were not able to quantify in our model. Pest establishment and spread, for instance, could be facilitated by the lack of natural enemies in the pests' nonnative range. Enemy release has commonly been found for nonnative plants (26), but its importance for nonnative pests is poorly understood and underdocumented (27). Numerous case studies have been described in the literature (sometimes inferred from biocontrol efforts) (28), although it is unclear whether this is a general trend among nonnative pests. The presence or absence of closely related pests may also affect pest establishment. If niches are phylogenetically conserved, for instance, nonnative pests may be more likely to establish in regions with native congeners; conversely, if closely related taxa are strong competitors, pests may be excluded from regions with close relatives (i.e., Darwin's naturalization conundrum). While this conundrum in plants can be partially explained by spatial scale (29), it remains unclear whether tree pests follow similar trends and, if they do, at what spatial scale they operate.

Last, we note that future global change is likely to shift the regions available for pest colonization, and in some cases, could affect pests' host breadths. Recent climate change has resulted in plant pests shifting poleward at a rate of, on average, 2.7 km/y (30), and if pests and plants track climate asynchronously (31), this is likely to expose many pests to novel hosts. The recent climate-induced expansion of mountain pine beetles (*Dendroctonus ponderosae*) into northern Alberta, for instance, has exposed jack pine (*Pinus banksiana*) to beetle attack (32) and potentially provides a corridor to other susceptible hosts in eastern North America. Warmer climates may also be associated with more favorable environmental conditions for pests, additional pest lifecycles, and greater host stress, each of which could increase the likelihood and impacts of pest establishments. Likewise the continued introduction of trees to new regions (accidentally or intentionally) similarly expands the potential range of many pests as host composition is the primary driver of tree pest composition globally (33). Future work integrating effects of global change on pest distributions could help refine projections of invasion risks over the next few decades.

Summary. Humans have markedly modified the distributions of plant and pest species, and in doing so we have revealed that many species can greatly extend their distributions when geographic barriers that have historically precluded dispersal are removed. Plants have been continually introduced outside their native ranges—either inadvertently or intentionally as agricultural, ornamental, or forestry species—for millennia (34), and some have been of immense benefit to humankind. Many non-native invasive plants, however, not only have direct negative economic/ecological effects, but also act as pathways for the introduction of novel pests. Our global model of pest occurrences indicates host richness and phylogenetic composition are among the most important factors limiting pest distributions; however, we show that many pests are absent from countries with suitable hosts and climates. We suggest there is substantial risk of these pests expanding their ranges in the future, threatening biodiversity and food security. Our statistical framework allows us to predict these future pest emergence events with high accuracy. Nonetheless, predicting the occurrence of rare or endemic pests remains challenging and may ultimately require additional fine-scale information on pest and host ecologies and data on coevolved versus acquired hosts and historic versus invaded pest ranges.

Materials and Methods

Pest Geographic and Host Data. Geographic and host information for tree pests was extracted from the Centre for Agriculture and Bioscience International's Crop Protection Compendium (CABI CPC) (35). The CPC tracks host and geographic occurrences of thousands of plant pests globally and utilizes sources that cover over 50 languages. In our analyses, we considered any pests that had at least one tree host recorded (defined below). The CPC, following the United Nations Food and Agriculture Organization, defines pests as "(a)ny species, strain or biotype of plant, animal or pathogenic agent injurious to plants or plant products" (36). Thus, while some species in our dataset may not be considered pests, per se, in their entire distribution, utilizing all available geographic data helps to ensure we are capturing the full environmental and host conditions required for pest occurrences. Pests include arthropods, chromists, fungi, molluscs, nematodes, protists, viruses, and several pests of unknown etiology. The database was accessed with permission in December 2020.

Host Geographic Data. Hosts in the CPC are variously defined at the species and genus level. Because the true host range of pests is often poorly known at

the species level, we delimit our host associations at the genus level. To extract tree hosts, we matched the species hosts in the CPC against a list of all tree species globally, using the GlobalTreeSearch (GTS) (37), which maintains a list of tree species present in each country. Next, for hosts only resolved at the genus level in the CPC, we first determined which genera include only free standing, woody species. This step was necessary as some genera include multiple growth forms. As an example, hosts listed in the CPC as "*Quercus alba*" and "*Solanum lycocarpum*" would be included in our analyses (as "*Quercus*" and "*Solanum*," respectively), as both are recognized tree species in the GTS. Hosts listed as "*Quercus* sp." would also be included, as all *Quercus* species are woody and free standing, but hosts listed as "*Solanum* sp." would not, as this genus includes multiple growth forms and several widely grown (nontree) crop species. Growth form information was extracted from Taseki et al. (38) and Zanne et al. (39).

Tree geographic ranges were extracted from the GTS (native ranges), Global Naturalized Alien Flora database (GloNAF, nonnative ranges) (40), and CABI Invasive Species Compendium (ISC, nonnative ranges) (41). For each pest, we identified the geographic extent of its host genera (native and nonnative ranges) and the countries where the pest was recorded as present. We assumed hosts were nonnative in any countries listed in GloNAF or ISC that were not included in GTS. We used pest occurrences and inferred absences in the native range of the hosts to fit the model described below, and we then used this model to predict pests in the nonnative host range.

Predictor Variables. Pest occurrences in host native ranges were fit using a hierarchical Bayesian regression model, which allowed us to specify a large complex model and explore the variability in predicted outcomes. Predictors included three pest \times country variables (number of hosts, average phylogenetic distance between hosts and native tree flora, and proportion of neighboring countries that had the pest), nine country-level variables (research output, per capita GDP, country area, and mean and SD of temperature, precipitation, and human footprint within countries), and pest type as a factor (i.e., arthropod, chromist, fungus, mollusc, nematode, protist, virus, and unknown).

Number of tree hosts was calculated on a per-pest basis, as the number of host genera native to a country. Because pest hosts tend to be phylogenetically conserved and the full host range of pests may not be entirely known, we also calculated the phylogenetic distance between known hosts and native tree communities in countries. Phylogenetic distances were quantified using a comprehensive phylogenetic tree of vascular plants, which represents the mean cophenetic distance between known hosts and all tree genera within countries (from the GlobalTreeSearch). The phylogeny was constructed with V.PhyloMaker (42) and an updated phylogenetic topology from Zanne et al. (39). Genera in our dataset but not present in the phylogeny were added as a polytomy to the family node (42).

For each pest \times country observation, we calculated a pest-specific spatial effect as the proportion of neighboring countries that were also known to have the pest, regardless of the presence or nativity of tree hosts. We first identified all countries that shared a land border, using the `poly2nb` function in the `spdep` package (43). Next, for each pest \times country presence or absence, we used geographic information of pest occurrence to determine the proportion of countries that also contained the pest. This variable helps to account for spatial autocorrelation in individual pest ranges.

We included nine country-level variables. Research output, as a proxy for sampling effort, was approximated using the number of citable documents published within countries between 1996 and 2019, downloaded from Scimago Journal Rankings (44). Per capita GDP was downloaded from the CIA World Factbook (45) and supplemented with information from the World Bank. Data on temperature, precipitation, and human footprint were obtained from global raster files. Pixel values were extracted within country borders, and the mean and SD were calculated. Countries that only encompassed a single raster pixel were assigned a SD of 0. Temperature (`biodim1`: mean annual temperature) and precipitation (`bioclim12`: annual precipitation) were each downloaded from WorldClim (46). Human footprint measures the intensity of eight human pressure variables (built-up environments, population density, electric power infrastructure, crop lands, pasture lands, roads, railways, and navigable waterways) in 1-km² grid cells across the globe (47, 48). The human footprint raster was reprojected and resampled to match the projection and resolution of the climate data before extracting values. Across all observations, variables were not strongly

correlated—the maximum correlation between any pair of variables being 0.63 (country area and temperature variability).

We also explored fitting the value of imported goods as a metric of trade and propagule pressure within the model. However, we found that the value of imported goods within countries was strongly correlated with the number of citable documents ($r = 0.84$). Furthermore, we found when fitting the model, described below, with imports in place of documents, the model with citable documents had a greater ability to predict pest presences in the native and non-native host ranges. Thus, we present the model with citable documents in the main text, and the model with “value of imported goods” in *SI Appendix*. Because these variables (imports and citable documents) represent disparate mechanisms, we address both in *Discussion*.

Model Fitting and Validation. Pest occurrences within their hosts’ native ranges were fit with a Bernoulli distribution and a logit link function in a hierarchical Bayesian regression framework. Pest identity and country were included as grouping variables with random intercepts. The model was defined as follows:

$$\begin{aligned} occurrence_{ij} &\sim \text{Bernoulli}(p_{ij}) \\ \text{logit}(p_{ij}) &= \alpha + \beta_1 \text{hosts}_{ij} + \beta_2 \text{cophen}_{ij} + \beta_3 \text{neighboring}_{ij} + \\ &\beta_4 \text{documents}_j + \beta_5 \text{gdp}_j + \beta_6 \text{area}_j + \\ &\beta_7 \text{tempMean}_j + \beta_8 \text{tempSD}_j + \beta_9 \text{precipMean}_j + \\ &\beta_{10} \text{precipSD}_j + \beta_{11} \text{footprintMean}_j + \beta_{12} \text{footprintSD}_j + \\ &\beta_{13} \text{pestType}_i + u_i + u_j, \end{aligned}$$

where $occurrence_{ij}$ is the presence or absence of pest i in country j , α is the intercept, $hosts_{ij}$ is the number of host genera of pest i in country j , $cophen_{ij}$ is the average cophenetic distance between hosts of pest i and all native trees in country j , $neighboring_{ij}$ is the proportion of neighboring countries of country j with pest i present, $documents_j$ are the number of documents published in country j , gdp_j is the per capita GDP of country j , $area_j$ is the geographic area of country j , $tempMean_j$, $tempSD_j$, $precipMean_j$, $precipSD_j$, $footprintMean_j$, $footprintSD_j$ are the mean and SD of temperature, precipitation, and human footprint in country j , $pestType_i$ is the pest type of pest i , and u_i and u_j are random effects (intercepts)

- J. A. LaManna *et al.*, Plant diversity increases with the strength of negative density dependence at the global scale. *Science* **356**, 1389–1392 (2017).
- D. Pimentel *et al.*, Economic and environmental threats of alien plant, animal, and microbe invasions. *Agric. Ecosyst. Environ.* **84**, 1–20 (2001).
- J. A. Logan, J. Régnière, J. A. Powell, Assessing the impacts of global warming on forest pest dynamics. *Front. Ecol. Environ.* **1**, 130–137 (2003).
- R. N. Sturrock *et al.*, Climate change and forest diseases. *Plant Pathol.* **60**, 133–149 (2011).
- R. Poulin, B. R. Krasnov, D. Moullot, D. W. Thielgts, The comparative ecology and biogeography of parasites. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **366**, 2379–2390 (2011).
- B. A. Roy *et al.*, Increasing forest loss worldwide from invasive pests requires new trade regulations. *Front. Ecol. Environ.* **12**, 457–465 (2014).
- G. S. Gilbert, C. O. Webb, Phylogenetic signal in plant pathogen-host range. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 4979–4983 (2007).
- V. Novotny *et al.*, Low host specificity of herbivorous insects in a tropical forest. *Nature* **416**, 841–844 (2002).
- I. M. Parker *et al.*, Phylogenetic structure and host abundance drive disease pressure in communities. *Nature* **520**, 542–544 (2015).
- I. S. Pearse, F. Altermatt, Predicting novel trophic interactions in a non-native world. *Ecol. Lett.* **16**, 1088–1094 (2013).
- D. P. Bebber, T. Holmes, D. Smith, S. J. Gurr, Economic and physical determinants of the global distributions of crop pests and pathogens. *New Phytol.* **202**, 901–910 (2014).
- D. Chapman, B. V. Purse, H. E. Roy, J. M. Bullock, Global trade networks determine the distribution of invasive non-native species. *Glob. Ecol. Biogeogr.* **26**, 907–917 (2017).
- D. M. Richardson, Forestry trees as invasive aliens. *Conserv. Biol.* **12**, 18–26 (1998).
- A. Jiménez-Valverde, J. M. Lobo, Threshold criteria for conversion of probability of species presence to either-or presence-absence. *Acta Oecol.* **31**, 361–369 (2007).
- J. E. Aukema *et al.*, Historical accumulation of nonindigenous forest pests in the continental United States. *Bioscience* **60**, 886–897 (2010).
- H. Seebens *et al.*, No saturation in the accumulation of alien species worldwide. *Nat. Commun.* **8**, 14435 (2017).
- H. Seebens *et al.*, Projecting the continental accumulation of alien species through to 2050. *Glob. Change Biol.* **27**, 970–982 (2020).
- H. Liu, P. Stiling, Testing the enemy release hypothesis: A review and meta-analysis. *Biol. Invasions* **8**, 1535–1545 (2006).
- S. L. Flory, K. Clay, Pathogen accumulation and long-term dynamics of plant invasions. *J. Ecol.* **101**, 607–613 (2013).
- T. D. Paine, M. J. Steinbauer, S. A. Lawson, Native and exotic pests of eucalyptus: A worldwide perspective. *Annu. Rev. Entomol.* **56**, 181–201 (2011).
- B. P. Hurlley *et al.*, Increasing numbers and intercontinental spread of invasive insects on eucalypts. *Biol. Invasions* **18**, 921–933 (2016).
- A. M. Liebhold, E. G. Brockerhoff, L. J. Garrett, J. L. Parke, K. O. Britton, Live plant imports: The major pathway for forest insect and pathogen invasions of the US. *Front. Ecol. Environ.* **10**, 135–143 (2012).

of pest and country. We used weakly informative priors for all beta coefficients (β_{1-12} , normal distribution with mean 0 and SD of 1). The model was fit with four chains, and 5,000 iterations per chain with the first 1,000 iterations used as warmup. We verified model convergence by ensuring that was near 1.0 and sufficient bulk and tail effective sample size. The model was fit in Stan (49) and called with the brms package (50) in R. In total, 314,802 pest \times country presences and absences were used in model fitting. Variables were scaled (mean of 0 and SD of 1.0) and log transformed (except mean temperature) before model fitting.

We used the model fit above to predict the occurrence (presence/absence) of 3,800 pests outside the native range of their hosts (n observations = 122,533). To adjust for sampling effects (measured as the number of citable documents, see above), we set the number of citable documents in all countries to the highest level globally. This allowed us to predict occurrences as if all countries had research output (i.e., sampling) equivalent to that of the United States, which was the highest globally. We tested the predictive power of the model using point estimates of the posterior predicted distribution of pest occurrences (both medians and means) and several standard validation statistics. The median effectively thresholds point estimates to 0 or 1, while the mean is a continuous proportion. For median estimates, we calculated the true presence and absence rates and false presence and absence rates both globally (all pest occurrences in all countries) and within individual countries. For mean estimates, we calculated the area under the receiver operating curve (AUC) and the Boyce index (51). While AUC has been criticized as a performance metric of distribution models (52), it remains widely used and reported. The Boyce index quantifies whether an increasing proportion of presences occurs in areas of increasingly high suitability and is measured as the Spearman’s rank correlation.

Data Availability. Anonymized data have been deposited in Figshare (10.6084/m9.figshare.17912159) (53).

ACKNOWLEDGMENTS. This work was supported by a Natural Sciences and Engineering Research Council of Canada Discovery Grant (to T.J.D.).

- R. Drenkhan *et al.*, Global geographic distribution and host range of *Fusarium circinatum*, the causal agent of pine pitch canker. *Forests* **11**, 724 (2020).
- M. J. Wingfield *et al.*, Pitch canker caused by *Fusarium circinatum*—A growing threat to pine plantations and forests worldwide. *Australas. Plant Pathol.* **37**, 319–334 (2008).
- G. Řitřková *et al.*, Swiss needle cast in western Oregon Douglas-fir plantations: 20-year monitoring results. *Forests* **7**, 155 (2016).
- K. Meijer, M. Schilthuis, L. Beukeboom, C. Smit, A review and meta-analysis of the enemy release hypothesis in plant-herbivorous insect systems. *PeerJ* **4**, e2778 (2016).
- E. G. Brockerhoff, A. M. Liebhold, Ecology of forest insect invasions. *Biol. Invasions* **19**, 3141–3159 (2017).
- M. Kenis, B. P. Hurlley, A. E. Hajek, M. J. W. Cock, Classical biological control of insect pests of trees: Facts and figures. *Biol. Invasions* **19**, 3401–3417 (2017).
- D. S. Park, X. Feng, B. S. Maitner, K. C. Ernst, B. J. Enquist, Darwin’s naturalization conundrum can be explained by spatial scale. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 10904–10910 (2020).
- D. P. Bebber, M. A. Ramotowski, S. J. Gurr, Crop pests and pathogens move polewards in a warming world. *Nat. Clim. Change* **3**, 985 (2013).
- E. K. Meineke, C. C. Davis, T. J. Davies, Phenological sensitivity to temperature mediates herbivory. *Glob. Change Biol.* **27**, 2315–2327 (2021).
- C. I. Cullingham *et al.*, Mountain pine beetle host-range expansion threatens the boreal forest. *Mol. Ecol.* **20**, 2157–2171 (2011).
- A. V. Gougherty, T. J. Davies, Host phylogenetic diversity predicts the global extent and composition of tree pests. *Ecol. Lett.* **25**, 101–112 (2022).
- R. N. Mack, “Plant invasions: Early and continuing expressions of global change” in *Past and Future Rapid Environmental Changes*, B. Huntley, W. Cramer, A. V. Morgan, H. C. Prentice, J. R. M. Allen, Eds. (NATO ASI Series, Springer, 1997), pp. 205–216.
- CABI, *Crop Protection Compendium* (CAB International, Wallingford, UK, 2020).
- CABI, *Plant Pest* (CAB Int. Gloss, Wallingford, UK, 2009).
- E. Beech, M. Rivers, S. Oldfield, P. P. Smith, GlobalTreeSearch: The first complete global database of tree species and country distributions. *J. Sustain. For.* **36**, 454–489 (2017).
- G. M. Taseski *et al.*, A global growth-form database for 143,616 vascular plant species. *Ecology* **100**, e02614 (2019).
- A. E. Zanne *et al.*, Three keys to the radiation of angiosperms into freezing environments. *Nature* **506**, 89–92 (2014).
- M. van Kleunen *et al.*, The Global Naturalized Alien Flora (GloNAF) database. *Ecology* **100**, e02542 (2019).
- CABI, *Invasive Species Compendium* (CAB International, Wallingford, UK, 2020).
- Y. Jin, H. Qian, V. Phylomaker: An R package that can generate very large phylogenies for vascular plants. *Ecography* **42**, 1353–1359 (2019).
- R. S. Bivand, E. J. Pebesma, V. Gomez-Rubio, *Applied Spatial Data Analysis with R* (Springer, New York, NY, ed. 2, 2013).
- Scimago, Journal & Country Rank. <https://www.scimagojr.com/countryrank.php>. Accessed 28 November 2020.

45. Central Intelligence Agency, The World Factbook. <https://www.cia.gov/the-world-factbook/>. Accessed 24 April 2020.
46. S. E. Fick, R. J. Hijmans, WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).
47. O. Venter *et al.*, Global terrestrial Human Footprint maps for 1993 and 2009. *Sci. Data* **3**, 160067 (2016).
48. O. Venter *et al.*, *Last of the Wild Project, Version 3 (LWP-3): 2009 Human Footprint, 2018 Release* (NASA Socioeconomic Data and Applications Center (SEDAC), 2018).
49. B. Carpenter *et al.*, Stan: A probabilistic programming language. *J. Stat. Softw.* **76**, 1–32 (2017).
50. P.-C. Bürkner, brms: An R package for Bayesian multilevel models using Stan. *J. Stat. Softw.* **80**, 1–28 (2017).
51. A. H. Hirzel, G. Le Lay, V. Helfer, C. Randin, A. Guisan, Evaluating the ability of habitat suitability models to predict species presences. *Ecol. Modell.* **199**, 142–152 (2006).
52. J. M. Lobo, A. Jiménez-Valverde, R. Real, AUC: A misleading measure of the performance of predictive distribution models. *Glob. Ecol. Biogeogr.* **17**, 145–151 (2008).
53. A. V. Gougherty, T. J. Davies, Data from "A global analysis of tree pests and emerging pest threats." Figshare. https://figshare.com/articles/dataset/A_global_analysis_of_tree_pests_and_emerging_pest_threats/17912159. Deposited 13 January 2022.