

# Worldwide border interceptions provide a window into human-mediated global insect movement

REBECCA M. TURNER <sup>1,16</sup> ECKEHARD G. BROCKERHOFF <sup>2</sup> CLEO BERTELSMEIER <sup>3</sup> RACHAEL E. BLAKE <sup>4</sup>  
BARNEY CATON <sup>5</sup> ALEX JAMES <sup>6,7</sup> ALAN MACLEOD <sup>8</sup> HELEN F. NAHRUNG <sup>9</sup> STEPHEN M. PAWSON <sup>1,10</sup>  
MICHAEL J. PLANK <sup>6,7</sup> DEEPA S. PURESWARAN <sup>11</sup> HANNO SEEBENS <sup>12</sup>  
TAKEHIKO YAMANAKA <sup>13</sup> AND ANDREW M. LIEBHOLD <sup>14,15</sup>

<sup>1</sup>Scion (New Zealand Forest Research Institute), P.O. Box 29237, Christchurch 8440 New Zealand

<sup>2</sup>Swiss Federal Research Institute WSL, Birmensdorf 8903 Switzerland

<sup>3</sup>Department of Ecology and Evolution, University of Lausanne, Lausanne 1015 Switzerland

<sup>4</sup>National Socio-Environmental Synthesis Center, Annapolis, Maryland 21401 USA

<sup>5</sup>U.S. Department of Agriculture, Raleigh, North Carolina 27606 USA

<sup>6</sup>Te Pūnaha Matatini, a New Zealand Centre of Research Excellence, Auckland 1142 New Zealand

<sup>7</sup>School of Mathematics and Statistics, University of Canterbury, Christchurch 8041 New Zealand

<sup>8</sup>Department for Environment, Food and Rural Affairs, York YO41 1LZ UK

<sup>9</sup>Forest Research Institute, University of the Sunshine Coast, Brisbane, Queensland 4102 Australia

<sup>10</sup>School of Forestry, University of Canterbury, Christchurch 8041 New Zealand

<sup>11</sup>Laurentian Forestry Centre, Canadian Forest Service, Quebec, Quebec G1V 4C7 Canada

<sup>12</sup>Senckenberg Biodiversity and Climate Research Centre, Frankfurt 60325 Germany

<sup>13</sup>Research Center for Agricultural Information Technology, NARO, Tokyo 3058604 Japan

<sup>14</sup>U.S. Department of Agriculture Forest Service Northern Research Station, Morgantown, West Virginia 26505 USA

<sup>15</sup>Faculty of Forestry and Wood Sciences, Czech University of Life Sciences Prague, Praha 6—Suchbát 165 00 Czech Republic

*Citation:* Turner, R. M., E. G. Brockerhoff, C. Bertelsmeier, R. E. Blake, B. Caton, A. James, A. MacLeod, H. F. Nahrung, S. M. Pawson, M. J. Plank, D. S. Pureswaran, H. Seebens, T. Yamanaka, and A. M. Liebhold. 2021. Worldwide border interceptions provide a window into human-mediated global insect movement. *Ecological Applications* 31(7):e02412. 10.1002/eap.2412

*Abstract.* As part of national biosecurity programs, cargo imports, passenger baggage, and international mail are inspected at ports of entry to verify compliance with phytosanitary regulations and to intercept potentially damaging nonnative species to prevent their introduction. Detection of organisms during inspections may also provide crucial information about the species composition and relative arrival rates in invasion pathways that can inform the implementation of other biosecurity practices such as quarantines and surveillance. In most regions, insects are the main taxonomic group encountered during inspections. We gathered insect interception data from nine world regions collected from 1995 to 2019 to compare the composition of species arriving at ports in these regions. Collectively, 8,716 insect species were intercepted in these regions over the last 25 yr, with the combined international data set comprising 1,899,573 interception events, of which 863,972 were identified to species level. Rarefaction analysis indicated that interceptions comprise only a small fraction of species present in invasion pathways. Despite differences in inspection methodologies, as well as differences in the composition of import source regions and imported commodities, we found strong positive correlations in species interception frequencies between regions, particularly within the Hemiptera and Thysanoptera. There were also significant differences in species frequencies among insects intercepted in different regions. Nevertheless, integrating interception data among multiple regions would be valuable for estimating invasion risks for insect species with high likelihoods of introduction as well as for identifying rare but potentially damaging species.

*Key words:* biological invasions; biosecurity; border inspection; invasive species; species composition; species pools.

## INTRODUCTION

International trade and travel have grown rapidly over the past century. This has resulted in increases in the unintentional movement of small organisms that may

contaminate goods moved across borders (Hulme 2009, Essl et al. 2011). Many of these contaminating organisms are insects, some of which have caused serious detrimental economic, ecological, and human-health impacts after they have arrived and established in nonnative habitats (Kenis et al. 2009, Bradshaw et al. 2016). As part of national biosecurity programs, inspection officers aim to reduce the global and domestic spread of unwanted organisms. They conduct inspections of arriving goods

Manuscript received 2 June 2020; revised 4 February 2021; accepted 3 March 2021; final version received 9 July 2021.  
Corresponding Editor: Davina L. Saccaggi.

<sup>16</sup>E-mail: Rebecca.Turner@scionresearch.com

and personal baggage at ports of entry (land borders, ports, and transitional facilities) (Saccaggi et al. 2016, Black and Bartlett 2020). However, the overwhelming volume of trade and travel means that it is usually only practical to inspect a small fraction of arriving material.

Insects are the most species-rich group of animals in the world, and that diversity is reflected in the large number of insects detected at ports of entry (McCullough et al. 2006). Insect life history traits often include cryptic habits as well as dormant and minute life stages that make detecting the full range of arriving insects a challenging task for border security agents. True arrival rates of potentially invading species are usually difficult to estimate. We use the term “arrival” in this paper to refer to arrival at a point of entry to a region. However, it does not automatically follow that an arriving insect will enter the region as, for example, phytosanitary treatments that kill, inactivate or remove pests, may be applied to shipments upon arrival (Hennessey et al. 2014). In addition, most nonnative species that enter will not establish successfully, and of those that do establish the majority are relatively innocuous in the invaded region (Brockerhoff and Liebhold 2017). Several studies have explored the relationship between arrival rates (propagule pressure) and establishment of nonnative species (e.g., Lockwood et al. 2005, Brockerhoff et al. 2006, Brockerhoff et al. 2014). Hence, information on insect contamination obtained from inspections performed at points of entry may serve as a useful proxy for arrival rates that could be used to evaluate invasion risk even though they may represent a small sample of the full cohort of globally traveling insects.

The concept of a species pool as a set of species that can potentially colonize and establish in a new community has been used to explain species richness in different ecological contexts and spatial scales (Zobel et al. 1998, Bennett et al. 2016). A source pool of potentially invading insect species would include species from source regions that have a high probability of being transported to new regions and, under favorable conditions, could establish at the destination (Seebens et al. 2018). In the context of the species arriving in a country, this global source pool includes species that have yet to establish as well as species that have already established in the destination country. As a simplification, insect species arriving at ports of entry may simply be “drawn” from a common global source pool, leading to a positive correlation between species abundance in the global source pool and border interception frequencies of arriving species. We refer to this model as the Global Model. Under this model, we would expect border interceptions of individual species to be correlated among different regions.

Under an alternative model, we could expect that the source pools of arriving insects show regional differences in that the composition of interceptions may also vary among regions due to differences in trading patterns and inspection practices. Interception frequencies are a

function of both the true arrival rates and the probability of (1) being found during inspections (which depends on how these are carried out) and (2) being recorded and entered into a database. We expect both of these factors to contribute to regional differences.

Ideally, interception information from multiple regions would be integrated to build a more complete picture of the species composition of insects in invasion pathways to identify invasion risk of individual species and higher taxa better. This could facilitate better prediction of species or families of insects that are likely to arrive and establish before they are detected in all individual regions, thus informing the process of pest risk analysis (MacLeod 2015). Such predictions would improve the readiness of regions with fewer resources available to monitor pathways and, in turn, could collectively slow the global spread of invading insects.

Here, our objective is to compare the differences and similarities in the composition of insect border interceptions among the following nine countries or world regions: New Zealand, Australia, South Korea, Japan, Canada, mainland United States, Hawaii, the United Kingdom and the region of the European and Mediterranean Plant Protection Organization (EPPO). For this we used available data collected at national ports of entry between 1995 and 2019 (details in Appendix S1: Table S1). Our goals here are (1) to identify similarities and differences among regions at the species level and summarized by higher taxa and (2) to use the Global Model as a reference to investigate the extent to which interceptions in individual regions deviate from each other. With this information, operational integration of international interception data sets may be uniquely valuable for increasing each region’s awareness of the range of insect species likely to arrive.

## DATA AND METHODS

### *Data acquisition and cleaning*

We analyzed interception data consisting of records of insects detected during inspections of international air and sea cargo, mail, vessels, and passenger baggage at ports of entry. Unwanted organisms discovered during these inspections were destroyed, so interceptions do not directly represent introductions of a species into a region but can be considered proxies for unobserved arrival of species. Port-of-entry interception data were provided by various government agencies around the world (see overview in Table 1). The sources and associated literature are listed in Appendix S1: Table S1, and data summaries in Data S1. In most regions, inspectors do not record the number of individuals discovered, but instead, they simply report species as being present in an inspected shipment or in passenger luggage. Hence, one interception represents a single detected arrival event for that species regardless of how many individuals were associated with that particular interception. Most of the data sets

TABLE 1. International port-of-entry interception summary statistics including all Insecta.

Region	Years	Interceptions	% identified to species level	Number of species	% of species with only one interception	Yearly interceptions per 2018 import value (in US\$ billions)	Yearly interceptions per 2018 international passenger arrivals (millions)
New Zealand	2000–2017	71,588 (3,977/yr)	47	1477	47	86†	11†
Australia	2003–2016	142,851 (10,204/yr)	43	1974	48	35	222
South Korea	1996–2014	131,061 (6,898/yr)	87	1040	33	12†	5†
Japan	1997–2017	355,422 (16,925/yr)	70	1153	25	21†	5†
Canada	1997–2019	13,640 (593/yr)	44	1259	54	1†	0†
USA mainland	1998–2018	1,140,279 (54,299/yr)	29	5641	45	12	273
USA Hawaii	1998–2018	17,261 (822/yr)	33	299	43	45	62
EPPO	1995–2010	14,718 (920/yr)	67	309	48	NA	NA
UK	2010–2018	12,753 (1,417/yr)	34	300	40	2	0

*Notes:* Data have been collected spanning the last two decades where possible. The year range is inclusive of the listed start and end year. Interception data may include detections from air and sea cargo, maritime vessels, international passenger baggage, and international mail/courier parcels. NA indicates not applicable. See Methods section and Appendix S1: Table S1 for notes about individual-region data sets.

†Estimates only as full data was not available, see Methods and Appendix S1 for further details.

did not contain information on whether interceptions originated from inspections of the passenger, cargo, or other pathways, so statistics on the average number of yearly interceptions per import value or international arrivals are based on the approximate percentage of interceptions that can be attributed to each pathway as listed in Appendix S1: Table S1. Further details on the sources of uncertainty are described in Appendix S1.

Not all interceptions were recorded at the species level. Depending on the region, a varying number of interceptions were recorded at the species, genus, tribe, family, or order level. However, for the analyses reported here only interceptions recorded at the species level were included, except for the overview statistics in Table 1. Over time and among regions, there has been variation in the nomenclature used to record names of species and higher taxa. A common taxonomic system was required to compare the data sets from all regions and years. No one single database exists for all insect names and their synonyms, so we used a multistep process based on the Global Biodiversity Information Facility (GBIF) backbone taxonomy to standardize taxonomic names. We chose GBIF (GBIF Secretariat 2020) as a base taxonomic database as it has good coverage of insect names and convenient access via R packages *rgbif* (Chamberlain et al. 2017) and *taxize* (Chamberlain and Szöcs 2013). Occasionally, the taxonomic names in GBIF are not the most up-to-date; however, we prioritized standardizing to unique taxonomic names over assigning the most up-to-date names. A detailed summary of the taxonomic resolution method and the R code is provided in Data S2. Despite our application of this largely automated and objective system, some synonyms may still be present among the tens of thousands of taxonomic names in the data set, although this is likely to be minimal.

#### *Analysis of taxonomic proportional representation*

To examine overall taxonomic trends, species interception counts from all nine regions were summed by species, family, and order levels. To evaluate whether intercepted species were a representative sample of world fauna, we compared the proportion of intercepted species in each order with comparable proportions from the entire world fauna (Zhang 2011). To test if the interception frequencies deviated from the null hypothesis that the proportions in each order are the same as for the globally described species, we calculated 95% confidence intervals around the expected number of intercepted species using the binomial distribution (standard binomial parameter for the probability of success,  $P = \text{number of interceptions}/\text{number of globally described species}$ , or  $P = \text{number of intercepted species}/\text{number of globally described species}$ ). Statistical analyses were conducted in the R language (R Development Core Team 2019).

#### *Analyses of species richness*

There was considerable variation among regions in the total numbers of interceptions, and this affects the comparability of numbers of species intercepted in each region. We therefore used rarefaction to standardize sample sizes to compare species richness of equal-sized samples (Chiarucci et al. 2008). Interception records are thought to represent only a small fraction of the total pool of insects arriving in each region (Kenis et al. 2007). Given the highly skewed distribution of species interception frequencies (Liebhold et al. 2017), it is likely that many species that arrive infrequently were never detected during inspections (Brockerhoff et al. 2014, Turner et al. 2020). Species richness was examined using sized-based rarefaction curves and sample coverage-based rarefaction curves using the *iNext*

package (Hsieh et al. 2016). Here, sample coverage refers to the proportion of the assemblage of arriving insects, that is, considering their relative abundances, which are represented by species in the interception sample sensu Good (1953, 2000) and Chao and Chiu (2016). We used abundance-based methods rather than incidence-based methods as used by Work et al. (2005), as interceptions in our data sets were consistently available as interception events but not consistently recorded with an associated sampling unit such as the container inspected. Work et al. (2005) note several of the limitations associated with using interception abundance data because of inspection and recording strategies. Therefore, we present the abundance-based rarefaction curves and explore the broad patterns while acknowledging the influence of interception probability. Chao1 species richness estimates were calculated using the package's `ChaoRichness()` function on the intercepted species abundance data.

### *Correlation analyses*

Pearson correlation coefficients in numbers of interceptions per species between regions were calculated to evaluate similarities among regions. Separate correlations among regions were calculated for each of the six most commonly intercepted orders, as well as for the most commonly intercepted families within those orders. We used Pearson's correlations of  $\log(x + 1)$  transformed interception frequencies to reduce the effect of extreme outliers, which might occur in some regions because of targeting of particular pathways, origins, or taxonomic groups. Following this transformation, the relationships between different regions' interception frequencies were roughly linear. However, to test the significance of correlations, we used Spearman's rank correlation coefficients because they are not sensitive to the highly skewed distribution of counts present in our species count data.

To evaluate the extent to which the compositions of interceptions in each region differ, we tested the null hypothesis that all regions were randomly sampled from a single pool of globally transported insect species; we compared observed correlations between pairs of regions with expected correlations generated under a null hypothesis of random draws from a single finite set of species with a set of relative frequencies (the Global Model). The pairwise correlations are analogous to a similarity matrix but, because each country has a different sample size, the correlations would be expected to deviate from 1 even if sampling with the same method from exactly the same pool of arriving insects (as in the Global Model). For an example of correlations expected under the Global Model for a range of sample sizes see Appendix S2: Fig. S1. The deviation from 1 is due to differences in sampling size and sampling with replacement. Observed correlations are likely to be less than the Global Model correlations because of both actual

differences in insect arrival rates (e.g., due to differences in imported goods and import sources), and differences in inspection methods and recording practices.

Expected correlations under the Global Model were calculated using a simulation method as follows for each taxonomic group: the Global Model is that all regions are randomly sampling from a single pool using equivalent methods, but at different sampling intensities; hence all interceptions could be combined into a pooled data set giving the total number of interceptions per species. For each region, a random sample (with replacement) was then drawn from this common pool of insects weighted by their frequency in the combined pool. The size of the random samples was set to the same number as the observed number of interceptions in each region. Note that the resulting simulated data set of interception frequency per species by region, just like the observed dataset, included zeros for species in the total pool which were not intercepted in the particular region. Correlations were calculated between the simulated data sets for each pair of regions. This process was replicated 1,000 times to generate samples from which the mean and standard deviation of the Fisher  $r$  to  $z$  transformed correlation coefficients for each regional pair could be calculated. Taking sample sizes into account, as we have done here, is important, because smaller sample sizes result in smaller correlations expected under the Model (as more of the data set are zeros, and see Appendix S2: Fig. S1 for an example of expected correlations relative to sample size). Fisher's  $r$  to  $z$  transformation was applied as the sampling distribution of correlation coefficients is not normal. To visualize the deviation of the observed correlation coefficients from the Global Model correlations, we calculated  $z$ -scores for each region pair. The  $z$ -score was calculated as the difference between the Fisher  $r$  to  $z$  transformed observed correlation and the mean of the Fisher  $r$  to  $z$  transformed correlations simulated under the Global Model, divided by the standard deviation. Observed correlations outside the range of 95% of the correlations simulated under the Global Model were identified based on the mean and standard deviations calculated from the Fisher's  $r$  to  $z$  transformed simulated correlation coefficients. A Bonferroni-type correction was applied to account for the 36 comparisons made per order (i.e., the pairwise comparison of all nine regions).

## RESULTS

### *Species interception frequencies and regional variation*

Across all regions, 863,972 of the 1,899,573 interceptions were identified to species level, with a total of 8,716 different insect species intercepted. Of those, 56% were intercepted more than once, 21% were intercepted more than 10 times, and 6% (563 species) were intercepted more than 100 times. Over half of all interceptions were of 82 species in common between at least eight (out of

nine) regions; 26 species, representing 21% of all interceptions, were intercepted in all regions. In contrast, a total of 75% of species were intercepted only in a single region, representing just 6% of all interceptions.

There was a strong correlation between the log-transformed total frequency of interceptions per species and the number of regions in which a species was intercepted ( $R = 0.75$ ,  $P < 0.001$ ; Fig. 1). However, many species were intercepted very frequently (in the hundreds to thousands) at ports of entry in only one region; the most frequently intercepted species intercepted in only a single country were *Frankliniella auripes* (Thysanoptera: Thripidae; 2,163 interceptions in mainland United States, arriving from South America), *Balanogastrius kolae* (Coleoptera: Curculionidae; 1,476 interceptions in mainland United States, mainly with baggage from Africa), *Brachypeplus rubidus* (Coleoptera: Nitidulidae; 1,208 interceptions in South Korea), and *Amrasca*

*biguttula* (Hemiptera: Cicadellidae; 1,031 interceptions in Japan, on vegetables from Southeast Asia). Despite the large diversity of insect species intercepted, the 2% most frequently intercepted species (174 species) accounted for 81% of all interceptions identified to species level. Across all regions the most frequently intercepted species were the gray pineapple mealybug *Dysmicoccus neobrevipes* (Hemiptera: Pseudococcidae; 49,905), the western flower thrips *Frankliniella occidentalis* (Thysanoptera: Thripidae; 43,809), and the onion thrips *Thrips tabaci* (Thysanoptera: Thripidae; 42,057). Thysanoptera were overrepresented in the 174 most frequently intercepted species compared to the full range of intercepted species (Pearson's one-sided chi-squared test,  $P = 0.003$ ; Fig. 2).

The overdispersed distribution of species interception frequencies was also seen in each of the individual regions' data sets, with a large proportion of species in

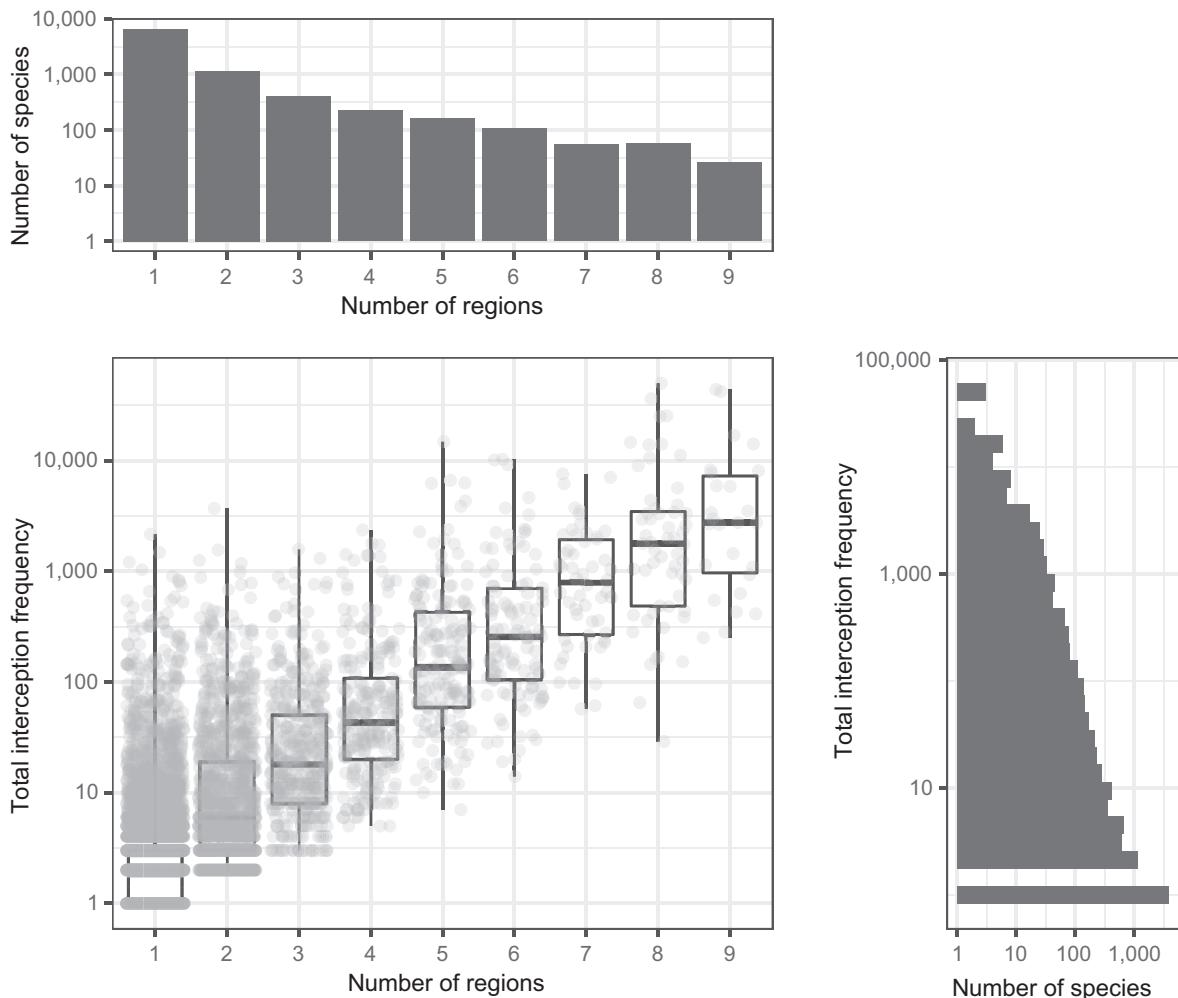


FIG. 1. Distribution of species' total interception frequencies by the number of regions in which they were intercepted. Interception frequencies, plotted on the log scale, of species identified to the species level from the combined international interception data set. Each gray point in the main panel represents a species. The top panel shows the total number of species intercepted in each region and the right panel shows the overall distribution of interception frequency across all regions and species.

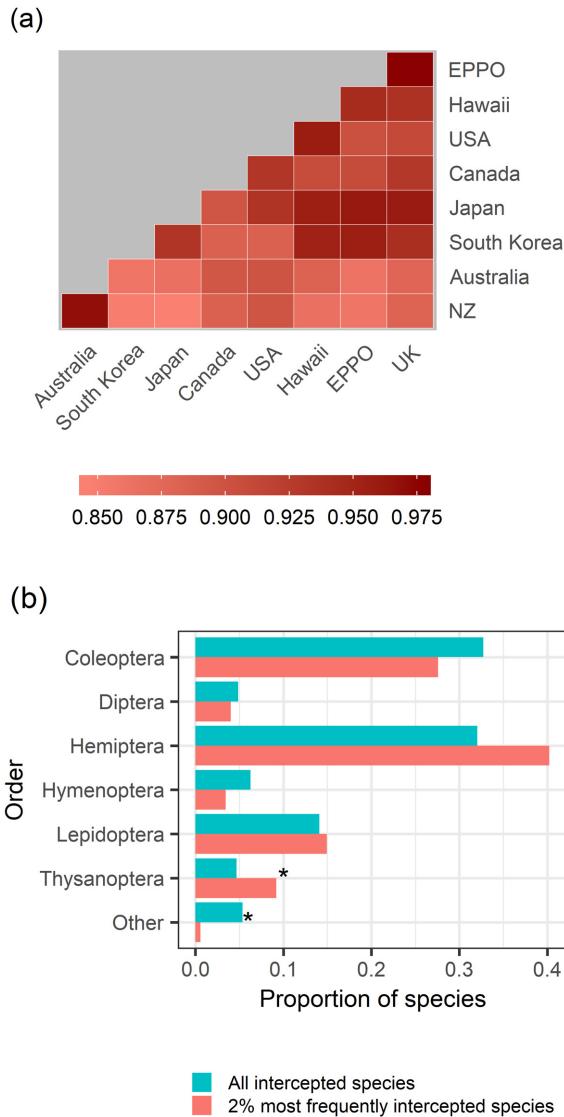


FIG. 2. Comparison of number of species by order. (a) Proportion of intercepted species in each insect order for all species and the 2% most frequently intercepted species from the combined data sets. EPPO, European and Mediterranean Plant Protection Organization; UK, United Kingdom; NZ, New Zealand. The proportion of species in each order are significantly different in the top 2% most frequently intercepted species compared to all intercepted species (Pearson's chi-squared test  $P$ -value = 0.001). The asterisks on the graph indicate which orders have significantly different proportions comparing all intercepted species with the most frequently intercepted species at the 95% confidence level, applying a Bonferroni correction for multiple comparisons. (b) Pearson's correlation matrix between regions in  $\log(x + 1)$  numbers of species per order.

each region having only one interception, hereafter referred to as singleton interceptions (Table 1). The mainland U.S. data set, which had the largest number of interceptions of all data sets, also had the largest number of species intercepted, comprising 65% of species intercepted from all regions (Fig. 3a). However, with only

around 50% of interceptions identified to the species level across all nine regions, it is likely that many other species were intercepted in each region but not recorded at the species level. Japan had a similar number of interceptions identified to the species level as the mainland United States (293,320 and 334,721, respectively) but many fewer species and a smaller percentage were singletons. In contrast, Canada, which had a similar number of intercepted species as Japan, had the highest percentage of singletons and no species intercepted >300 times. Consequently, the Canadian data set had the lowest average number of interceptions per species compared to all other regions.

As stated above, there were inconsistencies among regions with respect to the fraction of intercepted organisms that were identified to species. Biosecurity agencies in each region likely had different priorities for identifying and reporting different species. Furthermore, although data from most regions included records with higher taxonomic identification of interceptions lacking species identifications, interception data from South Korea did not include any data on interceptions that were not identified at the species level, and data from Japan did not include data on interceptions that were not identified to at least genus level. Such inconsistencies limit the use of these data in several ways and likely introduce biases in the use of data for characterizing the composition of arriving species.

The number of intercepted species (i.e., species richness) per order had only a weak linear relationship with the total number of interceptions identified to the species level by order, even on the log scale (Fig. 3b), whereas at the family level, the relationship between the number of species and the total number of interceptions was stronger (Fig. 3c). We further examined the proportion of interceptions in each order using both the number of species intercepted and the total number of species-level interceptions (Fig. 3c, d respectively). Across all regions, Thysanoptera had the most species-level interceptions per species on average, followed by Hemiptera (Fig. 3b). Thysanoptera and Hemiptera were overrepresented in interception data relative to global species richness based on both species-level interception counts and number of species (Pearson's chi-squared test,  $P$  values <0.001, Fig. 3d, e). However, some of these patterns varied among regions; for example, Hemiptera comprise a significantly larger fraction of interceptions in the United States compared to other regions (Appendix S2: Fig. S2).

A subset of regions included interception data at the level of order, family, or genus that were not identified to species level. For this subset of the interception data, the proportional representation of orders, for all interceptions, was compared to proportions of interceptions by order for interceptions that were identified to the species level (Appendix S2: Fig. S3). Analysis of variance was used to test whether the proportion of interceptions in an order within a region were dependent

on whether interception data included interceptions not recorded at the species level. Although there were significant differences in the proportional representation among orders ( $P < 0.001$ ), there was no significant interaction with whether data included interceptions not recorded at the species level ( $P = 0.6$ ). Hemiptera was, on average, the most frequently intercepted order regardless of whether interceptions not recorded at the species level were included (Appendix S2: Fig. S3). Diaspididae and Thysanoptera were among the most frequently intercepted families both for interceptions recorded at the species level and for all interceptions.

Less than 5% of the interceptions in most regions were from passenger baggage and even less were from international mail, with the predominant pathway being imported cargo. Yearly interception frequencies were normalized relative to the pathway size in each region using the estimated proportion of interceptions on the cargo or passenger baggage pathways. Although the mainland United States had the highest absolute yearly interception frequency, New Zealand, followed by Australia and Hawaii, had the highest yearly interception frequency expressed as a function of the value of imported cargo, consistent with the relative importance placed on biosecurity in these regions (Table 1). The U.S. mainland, Hawaii, and Australia, the only regions with more than 5% of their interceptions from the air passenger pathway, had the highest yearly interception frequency per international traveler arrivals. The statistics for South Korea and Japan are underestimates, because these countries only reported interceptions of organisms identified to the genus or species level.

#### *Species richness*

Size-based rarefaction curves of pooled interception frequencies (Fig. 4a) show the relative species richness of the six most frequently intercepted insect orders. When interception data were standardized for equal sample size, Coleoptera was the most species-rich order and Thysanoptera was the least species rich, consistent with the relative number of globally described species. The coverage-based rarefaction curves (Fig. 4b) indicate that Thysanoptera is the closest to complete coverage, followed by Hemiptera, while Hymenoptera lags far behind.

The Chao1 estimator for abundance data gives a biased estimate of species richness such that the estimates can be viewed as lower bounds for the species richness of the unobserved total arriving assemblage. Chao1 estimates of species richness (and estimated standard error) for pooled interceptions are 5,148 (182) Coleoptera, 853 (87) Diptera, 4,516 (143) Hemiptera, 965 (74) Hymenoptera, 2,540 (152) Lepidoptera, and 561 (39) Thysanoptera.

Differences in the species' rarefaction patterns among regions varied with insect order (Fig. 4c). For most orders, Canada appears to have a higher species richness

per sample size, but this may be an artifact of the data collection methodology driving the low average number of interceptions per species. For most orders, the U.S. mainland generally had a high species richness relative to sample size compared to other regions. Aside from Canada, Australia and New Zealand had high Diptera species richness relative to sample size. Australia and New Zealand also had high Hymenoptera species richness comparable with the mainland United States.

#### *Correlations of species frequencies among regions*

Correlations in species interception frequencies quantify the species-level similarity of the assemblages intercepted among regions. That is, correlations quantify the extent to which the same species are intercepted at different regions. Linear relationships, on the log scale, are reported using Pearson's correlations in Fig. 5, and rank similarities are reported using Spearman's correlations in Appendix S2: Fig. S4. Correlations were positive and significantly different from zero (tested using Spearman correlations) between most regions for most orders except for the Hymenoptera (Appendix S2: Fig. S4). Similarly, most correlations were significant for the Diaspididae, the Thripidae, and the Noctuidae, but most were not significant for the Tephritidae and the Formicidae (Appendix S2: Fig. S5). These trends may reflect sample size, because the regions and families with the most significant correlations were generally the orders and families with the greatest numbers of observations (Appendix S2: Fig. S4 and Appendix S2: Fig. S5; also see Appendix S2: Fig. S1).

To test the hypothesis that all regions are randomly drawn from the same pool of arriving species and subsequently examine how the observations deviate from the Global Model, correlations between pairs of regions were compared to correlations expected under the Global Model. In general, correlations between regions were lower than expected under the Global Model (Fig. 5). In fact, among the six main orders and all the regional comparisons, only Thysanoptera interceptions in the U.K. and the EPPO data sets were more highly correlated than expected under the Global Model. The expected variation in the correlation coefficient declines with increasing sample size per number of species (Appendix S2: Fig. S1) so we show the difference between observed and expected correlations as  $z$ -scores (Figs. 5b, 6b). In general, deviations from the expected mean were smallest for the Thysanoptera and Hymenoptera. For Hymenoptera in particular, the observed correlations varied strongly between pairs of regions. However, this correlation can be explained in part by differences in sample size between regions as the  $z$ -scores show roughly similar deviations for all pairs from the simulated correlations. For example, correlations in Hymenoptera species interception counts between Japan and other regions are quite low. This may be because the Japanese data set contains no interceptions of the

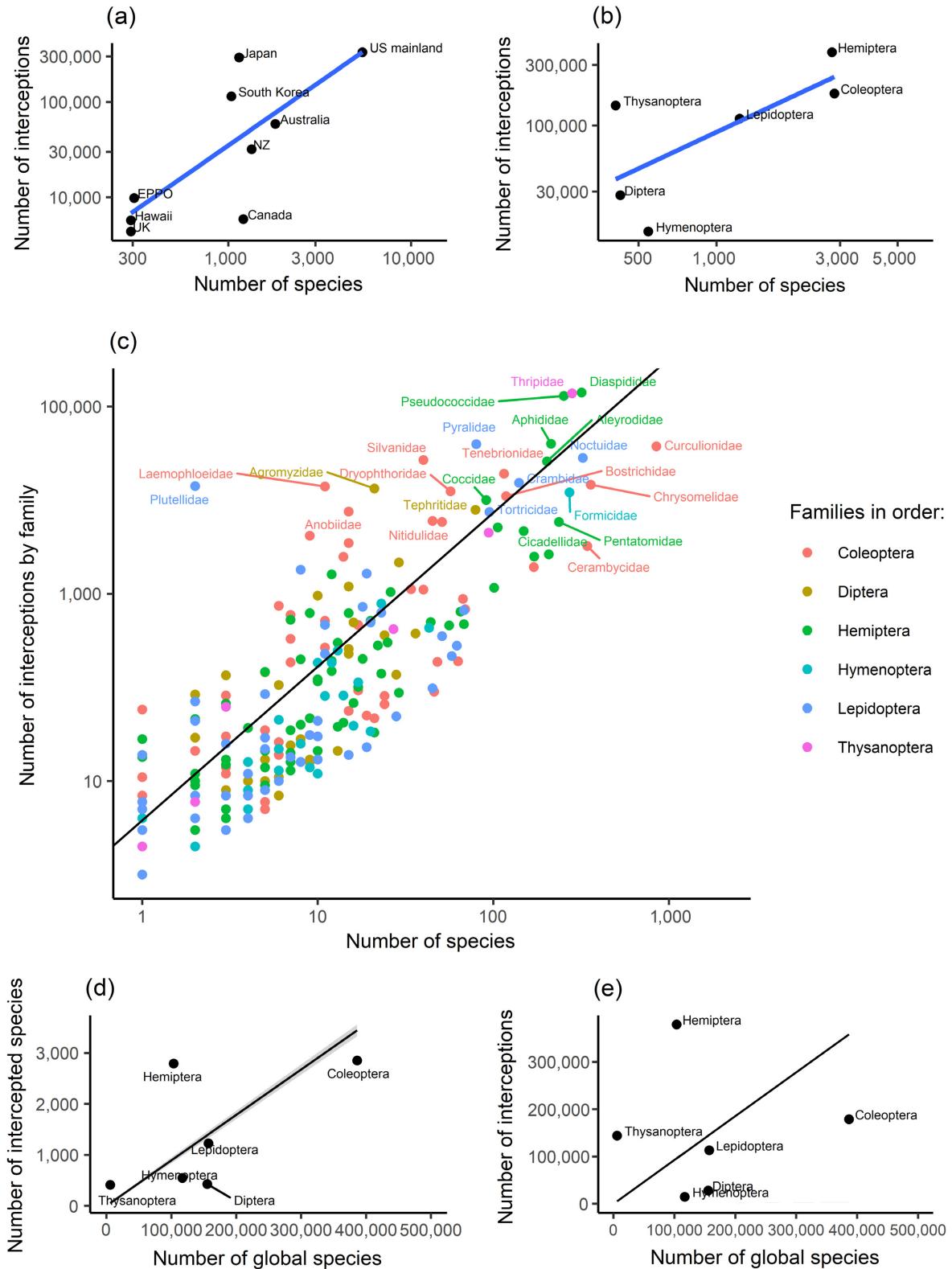


FIG. 3. Number of interceptions by number of species among the six insect orders: Coleoptera, Diptera, Hemiptera, Hymenoptera, Lepidoptera, and Thysanoptera, for interceptions identified to species level, and comparison of intercepted species with global species. (a) Grouped by region. EPPO, European and Mediterranean Plant Protection Organization; UK, United Kingdom; NZ,

(Fig. 3. *Continued*)

New Zealand. Blue line shows linear regression fit of logged data, adjusted  $R$ -squared: 0.52,  $P = 0.017$ . (b) Grouped by insect order. Blue line shows linear regression fit of logged data, adjusted  $R$ -squared: 0.38,  $P = 0.12$ . (c) Grouped by insect family. Families with more than 200 intercepted species or 5,900 interceptions are labeled. Black line shows linear regression fit of logged data, adjusted  $R$ -squared 0.78,  $P < 0.001$ . (d), (e) Global numbers for each order are the number of named species from Zhang (2011). Black line with gray surrounding region shows (d) the number of intercepted species or (e) number of interceptions expected per global species and the 95% confidence interval assuming that the interceptions would occur in the same proportions as in the global species.

globally most frequently intercepted Hymenoptera family—Formicidae (ants). The regions with the strongest correlations in Hymenoptera were those with the most intercepted Formicidae species.

Observed correlations were also compared to correlations generated by the Global Model for species in the most frequently intercepted families. The spread in correlation strengths was much greater at the family level, with some families showing very strong correlations between pairs of regions and others showing very weak correlations. The families shown in Fig. 6 are the families with the highest number of intercepted species for each of the six most common insect orders in the combined international data set. Correlations for a broader range of insect families can be found in Appendix S2: Fig. S6 and Appendix S2: Fig. S7.

## DISCUSSION

### *Similarities and differences*

Border interception records provide a unique window into the composition of insects arriving in areas outside their native ranges via global trade, travel, and international mail. Although there were considerable differences among regions in the composition of intercepted species (as evidenced by the unsurprising rejection of the null hypothesis that all regions sampled from the same global source pool), we also found some remarkable similarities in the composition of insect assemblages intercepted among different regions. Within most orders and families, species-level frequencies were significantly correlated (Figs. 5a, 6a, as well as Appendix S2: Fig. S4 and Appendix S2: Fig. S5), indicating that many species are arriving at various ports in roughly similar proportions. Similarly, the proportion of arriving species within each order was highly correlated between regions (Fig. 2b) suggesting that most of the same types of insects, as well as the same individual species, are arriving at different ports around the world.

Furthermore, we observed some similarity between the composition of intercepted species and the composition of the world's insect fauna (Fig. 3d). For example, the Coleoptera, the most species-rich insect order in the world, was also the most species-rich order among all interceptions. This pattern supports the notion that insects from various orders are in principle capable of entering invasion pathways.

Like the composition of interception data, the composition of established alien insects only broadly resembles

the world's insect diversity. Some taxa are over- or underrepresented in nonnative assemblages compared to native assemblages (Liebhold et al. 2016). For example, Liebhold et al. (2016) found that the Hemiptera and Thysanoptera are generally overrepresented in established alien species assemblages and hypothesized that the unexpectedly high species richness of these groups reflected both their high frequency in invasion pathways and their pronounced abilities to establish successfully. Our analysis here (see Fig. 3d, e) confirms that these insect orders, in particular, are overrepresented in pathways (see also Roques and Auger-Rozenberg 2006, Liebhold et al. 2012, Meurisse et al. 2019). The success of Thysanoptera as highly successful global invaders can be attributed in part to their small size, cryptic habits and frequent presence in plants transported with global trade and travel (Mound 2005, Morse and Hoddle 2006). Similarly, the overrepresentation of Hemiptera in alien assemblages can also be attributed to their association with live plant movement pathways (Liebhold et al. 2016, Brockerhoff and Liebhold 2017). In addition, some Thysanoptera and Hemiptera possess life-history traits such as parthenogenesis that facilitate successful establishment of newly arrived populations (Liebhold et al. 2016, Brockerhoff and Liebhold 2017).

Frequency distributions of intercepted species suggest that a large fraction of individual species is rarely intercepted. Indeed, 44 percent of all species were intercepted only once across all regions. On the other hand, there were a few species that were intercepted very frequently. The 2% most frequently intercepted species (174 species) accounted for 81% of all interceptions. Such a pattern of a highly skewed species frequency distribution is also a ubiquitous phenomenon in most naturally occurring communities (Magurran 2004). The presence of this pattern in interceptions (see also Liebhold et al. [2017], Turner et al. [2020]) suggests that the composition of invasion pathways can be considered a "sample" from natural communities. However, the abundance of species in invasion pathways may reflect both their abundance in source species pools as well as a biological propensity to become associated with commodities in trade (Gippet et al. 2019).

One of the consequences of this highly skewed distribution is that a small number of species are very abundant in invasion pathways. The three most frequently intercepted species, the gray pineapple mealybug (*Dysmicoccus neobrevipes*), the western flower thrips (*Frankliniella occidentalis*), and the onion thrips (*Thrips tabaci*) are all sap-feeding insects commonly associated

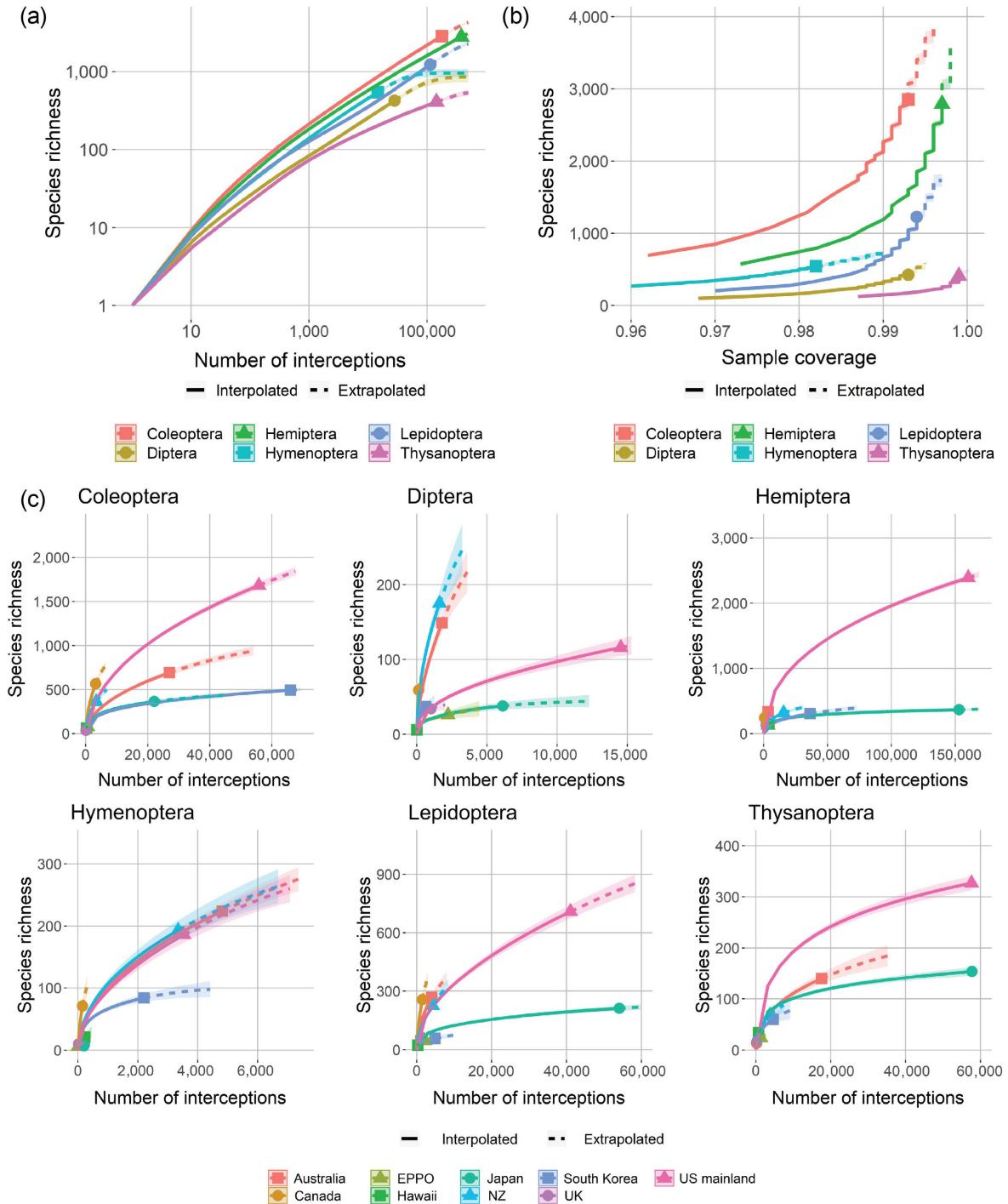


FIG. 4. Species richness rarefaction curves for each insect order. (a) Size-based rarefaction curves from pooled international interceptions by insect order (i.e., abundance data), shown on the log scale. (b) Coverage-based rarefaction from pooled international interceptions by insect order. (c) Size-based rarefaction curves by region and by order. Symbol locations show the observed values.

with widely traded plants and plant material, and all these have been successful invaders worldwide (e.g., Kirk and Terry [2003], Ji et al. [2010]). Such species can be considered “super-invaders” because they have a

frequent presence in pathways and have successfully invaded several world regions. Many of these very abundant species are common among the interceptions of all or most regions.

Despite the similarities among regions in the frequently intercepted species, about 75% of all species were intercepted in only one region (see Fig. 1). These included both species that were rarely intercepted (see “Taxonomic coverage”) and species that were intercepted quite often. The latter case appears to reflect specific bilateral trade patterns of products not widely traded globally or inspection targeting. Because of the large proportion of these region-specific interceptions, it appears that different regions are not always “sampling” the exact same global pool of insects that are being transported.

Even though there are many rarely intercepted species and species unique to a single region, there are some strong similarities in species interception frequencies among regions within certain orders (e.g., Thysanoptera; Fig. 4a). Correlations are particularly strong for some insect families. For example, the correlations for the Bostrichidae and Tenebrionidae are much stronger than for the Coleoptera as a whole (Appendix S2: Fig. S6 and Appendix S2: Fig. S7), possibly because of the high number of interceptions per species relative to other speciose Coleoptera families. Correlation analyses also indicate that assemblages in some regions are more similar to each other than others. For example, EPPO and U.K. species interception frequencies are consistently similar across the Diptera, Lepidoptera, and Thysanoptera (Fig. 5a); this similarity likely reflects the overlap in actual interception records, similarities in species targeted during interceptions, and similarity in trading partners. Exceptional similarities are also seen between Australia and New Zealand as well as between Japan and South Korea (Fig. 5a), again reflecting possible similarities in import source regions or inspection targets.

Despite the existence of substantial positive correlations in species interception frequencies among regions, our analysis revealed that these correlation strengths are mostly lower than expected under the Global Model that all regions are sampling from the same pool of arriving insects (Figs. 5, 6). The Thysanoptera generally had higher correlations relative to other orders, and the deviations from expectations under the Global Model were among the lowest for the six insect orders examined along with the Hymenoptera. The exceptionally high correlations of interception frequencies for the Thysanoptera may be explained by their generally high coverage in most data sets; rarefaction analysis (see “Taxonomic coverage”) indicated that interceptions were detecting a higher fraction of all species in pathways. Nevertheless, deviations from the Global Model for most groups indicate that inspections carried out in each region intercept slightly different sets of insect species. There are multiple factors that may contribute to these deviations.

#### *Sources of differences: real and apparent*

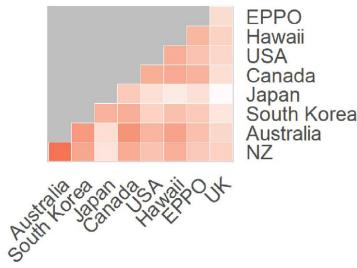
First, correlations of interception frequencies between regions would be lowered by the simple existence of

species with low interception probabilities (decreasing the chance a region will intercept the same species, in the same proportions) or variation in interception probabilities or underlying arrival rates between regions. Hence, given that the correlations for Thysanoptera are strong, this suggests that interception probabilities are reasonably high, which is supported by the sample coverage estimates, and the high number of interceptions to number of globally described species ratio (Fig. 3e; >20 interceptions per species on average). In addition, there is low variability in interception probabilities and arrival rates among Thysanoptera species between regions. However, it is also possible that all the regions included here have the same blind spots and are missing commonly arriving Thysanoptera that are hard to find or on alternative pathways. In contrast, the strong Pearson’s correlations between several regions for Hymenoptera, specifically Formicidae (Fig. 6), and their high interception frequencies relative to the other regions suggests that correlation with the remaining regions is weak partly because of low interception probabilities in these regions. Interceptions of Formicidae in countries such as Australia and New Zealand are high (among both species-level and all taxonomic-level interceptions), probably because of focused national biosecurity surveillance programs to improve management of risk posed by invasive ants in those countries (Craddock and Mattson 2014, Wylie et al. 2020).

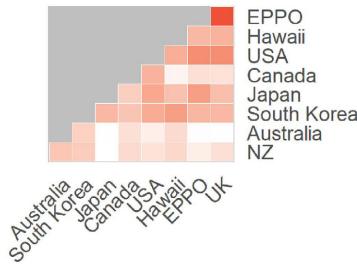
Another highly likely reason why different regions intercept slightly different sets of species is that there are biases in data collection that influence the selection of commodities or pathways that are most intensely inspected and hence which species are reported. On one hand, frequently arriving species become easy to identify, and are therefore more likely to be recorded. For example, 80% of Thysanoptera interceptions in Australia were comprised of two species. On the other hand, a region may prioritize recording information on uncommonly detected species that have not yet established and refrain from sending common species for identification or even refrain from recording common species, leading to their underrepresentation. The availability of taxonomic keys will also influence recording. Certain countries and regions maintain priority lists of regulated goods liable to convey species with quarantine status (Eschen et al. 2015) that are searched out and recorded preferentially or exclusively. The relatively small number of species in the interception records of EPPO and the United Kingdom may reflect the focused priority list of pest species and pathways, respectively, as well as the general low intensity of inspection due to generally low priority being given to biosecurity in these regions (Eschen et al. 2015). However, with the implementation of new biosecurity regulations in Europe (e.g., Anon [2016, 2019]), this may be changing.

In general, inconsistencies among countries with regard to inspection and reporting priorities prevent us from explaining deviations from the Global Model. It is

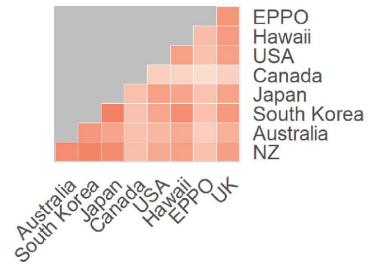
(a) Coleoptera: 2,852 species  
178,565 interceptions



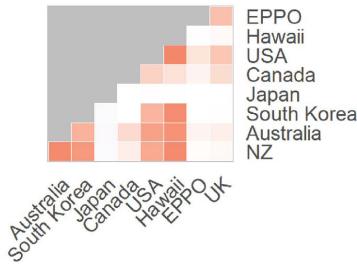
Diptera: 425 species  
28,101 interceptions



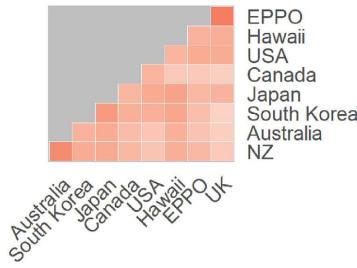
Hemiptera: 2,792 species  
379,007 interceptions



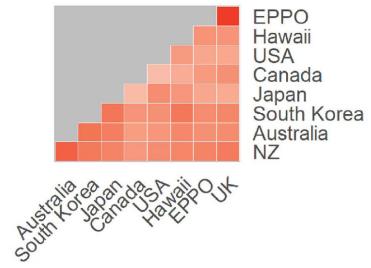
Hymenoptera: 545 species  
14,525 interceptions



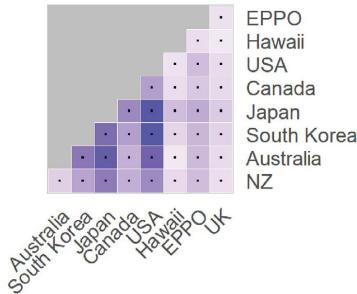
Lepidoptera: 1,227 species  
113,090 interceptions



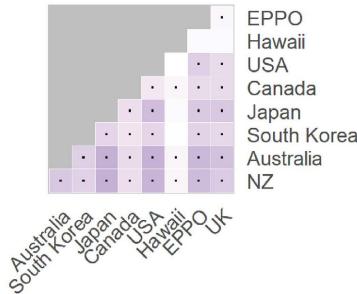
Thysanoptera: 408 species  
143,795 interceptions



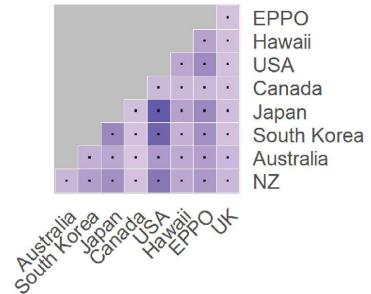
(b) Coleoptera



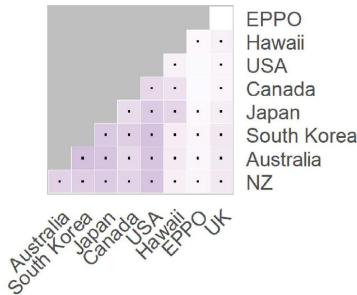
Diptera



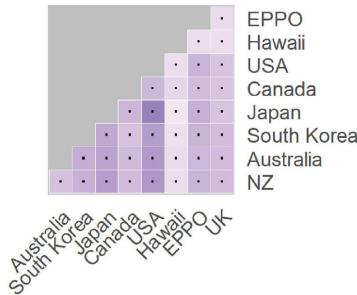
Hemiptera



Hymenoptera



Lepidoptera



Thysanoptera

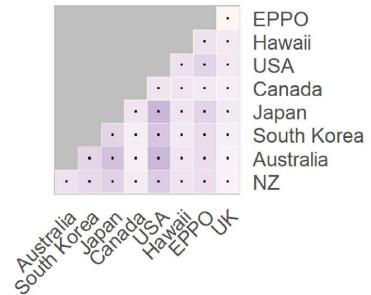


FIG. 5. Pearson's correlations of species interception frequencies between regions within orders. Eppo, European and Mediterranean Plant Protection Organization; UK, United Kingdom; NZ, New Zealand. Showing deviation from the Global Model of arriving species. (a) Correlation of number of interceptions for each species (using  $\log(x + 1)$  transformation) between regions by order. (b)  $z$ -scores showing difference between the observed Fisher's  $r$  to  $z$  transformed correlation and the Fisher's  $r$  to  $z$  transformed correlation expected under the Global Model. Negative (blue) values show that observed correlations were less than expected. A black square indicates if the observed correlation was outside the range of 95% of the correlation distribution simulated under the Global Model (using a Bonferroni-type correction to account for the 36 pairwise comparisons made per order).

clear that the regions from which we sourced data vary with regard to inspection and reporting priorities, but we were unable to obtain consistent information documenting those differences. This lack of information constrains the use of interception data for quantifying invasion risk at a global level.

Although we assume countries follow international standards for sampling imported consignments (FAO 2008), we agree with the recommendation of Saccaggi et al. (2016) that the value of interception as a monitoring tool would be greatly enhanced if biosecurity agencies provided documentation of their inspection sampling strategies, including records of inspection effort, to assist in the interpretation of these complex interception data sets (Kenis et al. 2007, EFSA Panel on Plant Health 2019). Furthermore, port inspections based on a statistically based sampling program are much more powerful for the utilization of data to estimate true approach rates and other properties (Moffitt et al. 2010). The International Plant Protection Convention has adopted two International Standards, ISPM 23 "Guidelines for Inspection" and "ISPM 31 "Methodologies for Sampling of Consignments" that guide countries in implementing border inspection (International Plant Protection Convention 2005, 2009) that specify methods that countries can adopt for implementing statistically based inspections. Many countries are adopting statistically based risk-based inspection methods (Griffin 2017) and such statistically based inspection procedures will provide much greater power in comparing and combining interception records from different countries in order to quantify risk better.

Another reason explaining differences in the composition of intercepted species among regions is variation in the source of imports. Several studies have shown that the sources (i.e., countries) of imported goods influence the composition of established species (Meyerson and Mooney 2007, Banks et al. 2015, Paini et al. 2016, Chapman et al. 2017, Liebhold et al. 2017, Turbelin et al. 2017, Lichtenberg and Olson 2018). Not surprisingly, the proportion of interceptions originating from different world regions differs among the intercepting regions. This is reflected, for example, in the differences among the interception data of South Korea (see Lee et al. [2016]), the United States (see McCullough et al. [2006]), and Europe (see Roques and Auger-Rozenberg [2006]). The higher-than-expected number of species that are highly intercepted in only one region observed here aligns with the conclusion that different regions are not

receiving identical pools of arriving species. Unfortunately, data on the source country, pathway (e.g., passenger baggage, international mail, vessel, air, or sea cargo), commodity (e.g., fresh produce, inanimate object, etc.) of material inspected was not available for all the data sets analyzed here; thus, we were unable to evaluate the influence of these factors. However, it is highly likely that the composition of arriving insects is strongly influenced by the composition of trade and international travelers arriving in that region. This is consistent with the findings of Work et al. (2005), who applied rarefaction analysis to examine relationships among the number of source countries, cargo pathways, and commodities on arrival rates of nonnative insects in the United States between 1997 and 2001. They found that there were higher rates of interception of new species on pathways with a wider variety of commodities. On the other hand, the relationship between arrival rates and the number of source countries was ambiguous. Clearly, the role of import origins and types of import commodities on the composition of insects in pathways warrants further investigation.

An interesting result of our study is that Hawaii ranks high in terms of the number of interceptions from both the cargo and passenger pathways (Table 1). The emphasis placed on inspection of both cargo imports and passengers arriving in Hawaii is understandable because Hawaii has been impacted particularly severely by many invasive insect species. As of 2006, Hawaii had at least 2,651 established nonnative insect species with up to 500 species causing widespread economic or ecological damage (Nishida 2002, Kraus and Duffy 2010). New Zealand, which ranks first in terms of interceptions per import value, is known to have a particularly rigorous biosecurity system, perhaps the strictest in the world (e.g., Eschen et al. [2015]). That New Zealand ranks lower (fourth) in terms of interceptions per passenger arrivals seems surprising as New Zealand also operates very strict inspections of passenger baggage. However, New Zealand has an effective system to discourage arriving passengers from carrying fruit and other high-risk items and provides amnesty quarantine bins; this may be the reason for the relatively smaller number of interceptions with passengers.

#### *Taxonomic coverage*

Rarefaction analysis indicates that we are probably not close to complete coverage of the total number of

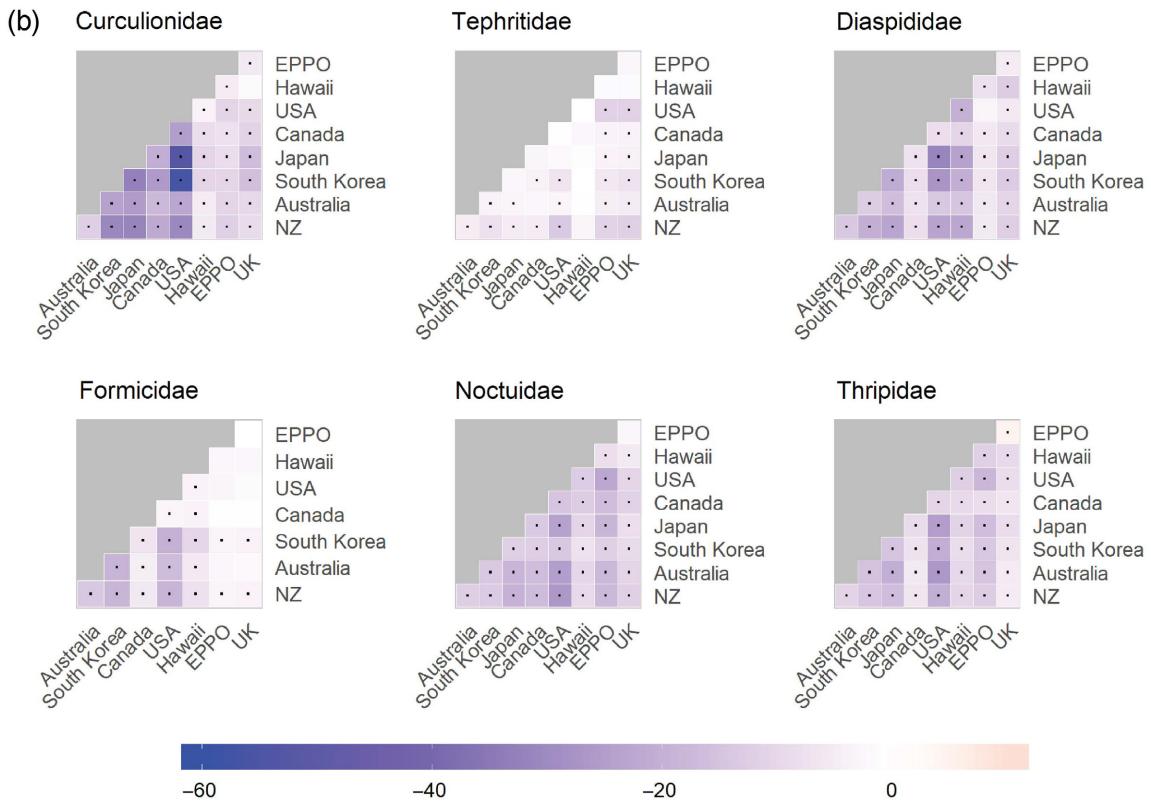
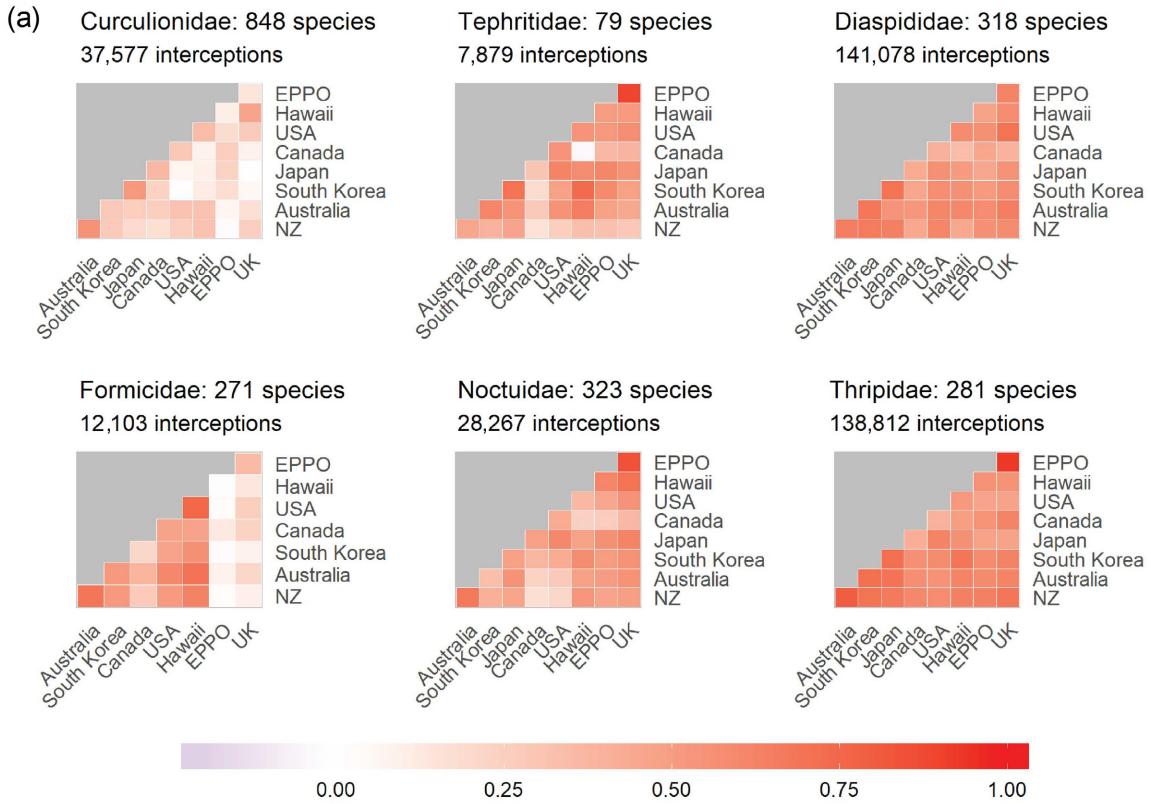


FIG. 6. Pearson's correlations of species interception frequencies between regions within families. EPPO, European and Mediterranean Plant Protection Organization; UK, United Kingdom; NZ, New Zealand. Showing deviation from the Global Model of arriving species. (a) Correlation of number of interceptions for each species (using  $\log(x + 1)$  transformation) between regions by family, for the families with the most intercepted species in each of the orders from Fig. 5. Note that Japan is missing from the Formicidae panel, because the data set contained no Formicidae interceptions. (b)  $z$ -scores showing difference between the observed Fisher's  $r$  to  $z$  transformed correlation and the Fisher's  $r$  to  $z$  transformed correlation expected under the Global Model. Negative (blue) values show that observed correlations were less than expected. A black square indicates if the observed correlation was outside the range of 95% of the correlation distribution simulated under the Global Model (using a Bonferroni-type correction to account for the number of pairwise comparisons made per family).

arriving species for any of the orders (Fig. 4) though an exceptional situation exists for the Thysanoptera. Of all the orders, Thysanoptera has the lowest percentage of species with singleton interceptions in the combined data set (29% compared to 45%–52% for the remaining five main orders). Similarly, Thysanoptera has the highest observed sample coverage at 99.9%, and Hymenoptera the lowest observed coverage at 98.2% that fits with the trend seen in the correlation strengths. The sample coverage estimates take into account the relative abundance of species, hence the high estimates for coverage.

Given the high percentage of species in the interception data with singleton interceptions (44%) and the results of the rarefaction analysis, it is difficult to estimate the true total number of species moving globally. There are various methods for estimating species richness (e.g., Chao richness [Chao 1984, 1987]) but these are often lower bounds and hence tend to underestimate the true total number of species. Inspection-method biases will also limit the range of species intercepted. For example, strategies that favor inspection of certain pathways or species over others or recording procedures, which do not record species level identifications of some groups, may not sample from the full range of arriving species. Therefore, the true number of internationally moving insects is likely to be greater than the Chao1 estimates for the minimum species richness reported here, contributing to the deviation between observed and expected correlations.

Rarefaction analyses are strongly affected by the distribution of the interception frequencies in each region. For example, Fig. 4 shows Canada having greater diversity because the most frequently intercepted species in Canada are not that frequent relative to its rarely intercepted species (only 100s of interceptions) even considering the lower total interception frequency for all Insecta. Gotelli and Graves (1996) noted that greater evenness in the species distribution results in a steeper rarefaction curve, as is seen in the Canadian interception data. Japan and South Korea illustrate the opposite extreme, with their most frequently intercepted species recorded tens of thousands of times and thus rarefaction curves indicate relatively low diversity (Fig. 4). However, it is often difficult to discern whether such differences in species interception frequency distributions among regions reflect a difference in sampling methodology or true differences in rates of arriving insects.

Variation in coverage across insect taxonomic groups indicates that family-level differences in coverage should be considered if using interception frequency to predict arrival rates. For example, the spotted lanternfly, *Lycorma delicatula*, has only been intercepted once in the combined international data set—even though it is spreading rapidly on a global scale (Lee et al. 2019). However, this should not be surprising since the entire Fulgoridae family is infrequently intercepted relative to other families in Hemiptera, indicating that the limited interception of *Lycorma delicatula* was to be expected. The relative interception frequencies of different families may reflect differences in detectability as well as inspection effort and efficiency rather than actual differences in arrival rates. Furthermore, differences in inspection effort may be caused by a focus on particular pathways rather than particular families.

## CONCLUSIONS

We found that there is a remarkable similarity in the composition of insect species arriving in different world regions, with a small number of species frequently intercepted in most world regions and a large number of species intercepted infrequently in just one or a few regions. Despite similarities among regions, there are unique aspects to the composition of species intercepted in different regions. The differences may be due in part to differences in the composition of source countries or pathways for each region as well as differences in inspection methodology and inspector priorities. Analyses also demonstrate that a large fraction of species present in invasion pathways have not been detected during inspections. Many of these undetected species probably arrive at low rates but some still pose significant risk.

These results highlight the potential benefit that may come from sharing interception data among world regions. Given that inspection programs in most countries are only able to detect a small fraction of species present in import pathways, pooling of interception data from other regions may allow countries to identify the presence of potentially dangerous organisms in invasion pathways and adapt biosecurity practices to mitigate that risk.

## ACKNOWLEDGMENTS

RT's fellowship was funded by the Ministry of Business, Innovation and Employment (New Zealand's Biological

Heritage National Science Challenge, C09X1501) and Te Pūnaha Matatini. MJP and AJ acknowledge financial support from Te Pūnaha Matatini. SP and EB were also supported by the New Zealand Biological Heritage National Science Challenge. Additional funding was provided by U.S. Department of Agriculture Forest Service International Programs and grant EVA4.0, No. CZ.02.1.01/0.0/0.0/16\_019/0000803 financed by OP RDE. Support was also provided by the National Socio-Environmental Synthesis Center (SESYNC) under funding received from the National Science Foundation DBI-1639145. HS received funding by the BiodivERsA-Belmont Forum Project 'AlienScenarios' through the national funder BMBF (grant 01LC1807A). CB was funded by the Swiss National Science Foundation and the Programme de la Fondation Sandoz-Monique de Meuron pour la relève universitaire. The authors would like to thank the Ministry for Primary Industries, New Zealand, the Australian Chief Plant Protection Office, the United States Department of Agriculture, the Canadian Food Inspection Agency, and the Department for Environment, Food and Rural Affairs (Defra), United Kingdom, for providing data. The authors would also like to thank Cathy Sue Katsar for thoughtful comments on the manuscript and useful input from Robert Griffin. Authors' contributions: the paper was jointly conceived by all authors; data preparation was conducted by RT and RB; analyses were performed by RT and HN; RT wrote the first draft with considerable input from DP, EB, and AL; all authors participated in subsequent revision of the manuscript.

## LITERATURE CITED

- Anon. 2016. Regulation (EU) 2016/2031 of the European Parliament of the Council of 26 October 2016 on protective measures against pests of plants. Official Journal of the European Union L 317:4–104. 23.11.2016.
- Anon. 2019. Commission Implementing Regulation (EU) 2019/2072 of 28 November 2019 establishing uniform conditions for the implementation of Regulation (EU) 2016/2031 of the European Parliament and the Council, as regards protective measures against pests of plants, and repealing Commission Regulation (EC) No 690/2008 and amending Commission Implementing Regulation (EU) 2018/2019. Official Journal of the European Union L 319:1–279. 10.12.2019.
- Banks, N. C., D. R. Paini, K. L. Bayliss, and M. Hodda. 2015. The role of global trade and transport network topology in the human-mediated dispersal of alien species. *Ecology Letters* 18:188–199.
- Bennett, J. A., K. Riibak, E. Kook, Ü. Reier, R. Tamme, C. Guillermo Bueno, and M. Pärtel. 2016. Species pools, community completeness and invasion: disentangling diversity effects on the establishment of native and alien species. *Ecology Letters* 19:1496–1505.
- Black, R., and D. M. F. Bartlett. 2020. Biosecurity frameworks for cross-border movement of invasive alien species. *Environmental Science & Policy* 105:113–119.
- Bradshaw, C. J. A., B. Leroy, C. Bellard, D. Roiz, C. Albert, A. Fournier, M. Barbet-Massin, J.-M. Salles, F. Simard, and F. Courchamp. 2016. Massive yet grossly underestimated global costs of invasive insects. *Nature Communications* 7:1–8.
- Brockerhoff, E. G., J. Bain, M. Kimberley, and M. Knížek. 2006. Interception frequency of exotic bark and ambrosia beetles (Coleoptera: Scolytinae) and relationship with establishment in New Zealand and worldwide. *Canadian Journal of Forest Research* 36:289–298.
- Brockerhoff, E. G., M. Kimberley, A. M. Liebhold, R. A. Haack, and J. F. Cavey. 2014. Predicting how altering propagule pressure changes establishment rates of biological invaders across species pools. *Ecology* 95:594–601.
- Brockerhoff, E. G., and A. M. Liebhold. 2017. Ecology of forest insect invasions. *Biological Invasions* 19:3141–3159.
- Chamberlain, S., K. Ram, V. Barve, D. McGlenn, and M. S. Chamberlain. 2017. Package 'rgbif'. <https://cran.r-project.org/web/packages/rgbif/index.html>
- Chamberlain, S. A., and E. Szöcs. 2013. taxize: taxonomic search and retrieval in R. *F1000Research* 2:191.
- Chao, A. 1984. Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics* 11:265–270.
- Chao, A. 1987. Estimating the population size for capture–recapture data with unequal catchability. *Biometrics* 43: 783–791.
- Chao, A., and C. H. Chiu. 2016. Species richness: estimation and comparison. *Wiley StatsRef: Statistics Reference Online* 1–26. <http://dx.doi.org/10.1002/9781118445112.stat03432.pub2>
- Chapman, D., B. V. Purse, H. E. Roy, and J. M. Bullock. 2017. Global trade networks determine the distribution of invasive non-native species. *Global Ecology and Biogeography* 26:907–917.
- Chiarucci, A., G. Bacaro, D. Rocchini, and L. Fattorini. 2008. Discovering and rediscovering the sample-based rarefaction formula in the ecological literature. *Community Ecology* 9:121–123.
- Craddock, P., and L. Mattson. 2014. National invasive ant surveillance programme annual report. Contract no. 15740. MPI Technical Paper no. 2014/35. Ministry for Primary Industries, Wellington, New Zealand.
- EFSA Panel on Plant Health (PLH). 2019. Pest categorisation of *Spodoptera litura*. *EFSA Journal* 17:5765.
- Eschen, R., A. Roques, and A. Santini. 2015. Taxonomic dissimilarity in patterns of interception and establishment of alien arthropods, nematodes and pathogens affecting woody plants in Europe. *Diversity and Distributions* 21:36–45.
- Essl, F., et al. 2011. Socioeconomic legacy yields an invasion debt. *Proceedings of the National Academy of Sciences of the United States of America* 108:203–207.
- FAO 2008. Pages 24. International Standards for Phytosanitary Measures (ISPM) 31. Methodologies for sampling of consignments. IPPC, FAO, Rome, Italy.
- GBIF Secretariat. 2020. GBIF backbone taxonomy. Checklist dataset. <https://doi.org/10.15468/39omei> accessed via GBIF.org 2020-01-17
- Gippet, J. M. W., A. M. Liebhold, G. Fenn-Moltu, and C. Bertelsmeier. 2019. Human-mediated dispersal in insects. *Current Opinion in Insect Science* 35:96–102.
- Good, I. J. 1953. The population frequencies of species and the estimation of population parameters. *Biometrika* 40:237–264.
- Good, I. J. 2000. Turing's anticipation of empirical Bayes in connection with the cryptanalysis of the naval Enigma. *Journal of Statistical Computation and Simulation* 66:101–111.
- Gotelli, N. J., and G. R. Graves 1996. Null models in ecology. Smithsonian Institution Press, Washington, D.C., USA.
- Griffin, R. L. 2017. Introduction to the international symposium for risk-based sampling. Pages 6–11 in *Proceedings International Symposium for RiskBased Sampling*. North American Plant Protection Organization, Raleigh, North Carolina, USA.
- Hennessey, M. K., et al. 2014. Phytosanitary treatments. Pages 269–308 in G. Gordh and S. McKirdy, editors. *The handbook of plant biosecurity*. Springer Netherlands.
- Hsieh, T. C., K. H. Ma, and A. Chao. 2016. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods in Ecology and Evolution* 7:1451–1456.
- Hulme, P. E. 2009. Trade, transport and trouble: managing invasive species pathways in an era of globalization. *Journal of Applied Ecology* 46:10–18.

- International Plant Protection Convention. 2005. ISPM No. 23. Guidelines for inspection. United Nations Food and Agriculture Organization, Rome, Italy.
- International Plant Protection Convention. 2009. ISPM No 31. Methodologies for sampling of consignments. United Nations Food and Agriculture Organization, Rome, Italy.
- Ji, J.-Y., W. San-an, and S.-J. Suh. 2010. A list of mealybugs (Hemiptera: Pseudococcidae) intercepted at the Republic of Korea ports of entry on plants imported from China. *Korean Journal of Applied Entomology* 49:163–166.
- Kenis, M., M.-A. Auger-Rozenberg, A. Roques, L. Timms, C. Péré, M. J. W. Cock, J. Settele, S. Augustin, and C. Lopez-Vaamonde. 2009. Ecological effects of invasive alien insects. *Biological Invasions* 11:21–45.
- Kenis, M., W. Rabitsch, M. A. Auger-Rozenberg, and A. Roques. 2007. How can alien species inventories and interception data help us prevent insect invasions? *Bulletin of Entomological Research* 97:489–502.
- Kirk, W. D. J., and L. I. Terry. 2003. The spread of the western flower thrips *Frankliniella occidentalis* (Pergande). *Agricultural and Forest Entomology* 5:301–310.
- Kraus, F., and D. C. Duffy. 2010. A successful model from Hawaii for rapid response to invasive species. *Journal for Nature Conservation* 18:135–141.
- Lee, D.-H., Y.-L. Park, and T. C. Leskey. 2019. A review of biology and management of *Lycorma delicatula* (Hemiptera: Fulgoridae), an emerging global invasive species. *Journal of Asia-Pacific Entomology* 22:589–596.
- Lee, W., Y. Lee, S. Kim, J.-H. Lee, H. Lee, S. Lee, and K.-J. Hong. 2016. Current status of exotic insect pests in Korea: comparing border interception and incursion during 1996–2014. *Journal of Asia-Pacific Entomology* 19:1095–1101.
- Lichtenberg, E., and L. J. Olson. 2018. The fruit and vegetable import pathway for potential invasive pest arrivals. *PLoS One* 13:e0192280.
- Liebholt, A. M., E. G. Brockerhoff, L. J. Garrett, J. L. Parke, and K. O. Britton. 2012. Live plant imports: the major pathway for forest insect and pathogen invasions of the US. *Frontiers in Ecology and the Environment* 10:135–143.
- Liebholt, A. M., E. G. Brockerhoff, and M. Kimberley. 2017. Depletion of heterogeneous source species pools predicts future invasion rates. *Journal of Applied Ecology* 54:1968–1977.
- Liebholt, A. M., T. Yamanaka, A. Roques, S. Augustin, S. L. Chown, E. G. Brockerhoff, and P. Pyšek. 2016. Global compositional variation among native and non-native regional insect assemblages emphasizes the importance of pathways. *Biological Invasions* 18:893–905.
- Lockwood, J. L., P. Cassey, and T. Blackburn. 2005. The role of propagule pressure in explaining species invasions. *Trends in Ecology and Evolution* 20:223–228.
- MacLeod, A. 2015. The relationship between biosecurity surveillance and risk analysis. Pages 109–122 in F. Jarrad, S. Low Chow, and K. Mengersen, editors. *Biosecurity surveillance: quantitative approaches*. CABI, Wallingford, UK.
- Magurran, A. E. 2004. *Measuring biological diversity*. Blackwell Science Ltd., Carlton, Victoria, Australia.
- McCullough, D. G., T. T. Work, J. F. Cavey, A. M. Liebhold, and D. Marshall. 2006. Interceptions of nonindigenous plant pests at US ports of entry and border crossings over a 17-year period. *Biological Invasions* 8:611.
- Meurisse, N., D. Rassati, B. P. Hurley, E. G. Brockerhoff, and R. A. Haack. 2019. Common pathways by which non-native forest insects move internationally and domestically. *Journal of Pest Science* 92:13–27.
- Meyerson, L. A., and H. A. Mooney. 2007. Invasive alien species in an era of globalization. *Frontiers in Ecology and the Environment* 5:199–208.
- Moffitt, L. J., J. K. Stranlund, and C. D. Osteen. 2010. Securing the border from invasives: robust inspections under severe uncertainty. *Economics Research International* 2010:1–9.
- Morse, J. G., and M. S. Hoddle. 2006. Invasion biology of thrips. *Annual Review of Entomology* 51:67–89.
- Mound, L. A. 2005. Thysanoptera: diversity and interactions. *Annual Review of Entomology* 50:247–269.
- Nishida, G. M. 2002. *Hawaiian terrestrial arthropod checklist*. Fourth edition. Hawaii Biological Survey (Bishop Museum) Technical Report, Honolulu, Hawaii, USA.
- Paini, D. R., A. W. Sheppard, D. C. Cook, P. J. De Barro, S. P. Worner, and M. B. Thomas. 2016. Global threat to agriculture from invasive species. *Proceedings of the National Academy of Sciences of the United States of America* 113:7575–7579.
- R Development Core Team. 2019. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. [www.R-project.org](http://www.R-project.org)
- Roques, A., and M. A. Auger-Rozenberg. 2006. Tentative analysis of the interceptions of non-indigenous organisms in Europe during 1995–2004 I. *EPPO Bulletin* 36:490–496.
- Saccaggi, D. L., M. Karsten, M. P. Robertson, S. Kumschick, M. J. Somers, J. R. U. Wilson, and J. S. Terblanche. 2016. Methods and approaches for the management of arthropod border incursions. *Biological Invasions* 18:1057–1075.
- Seebens, H., et al. 2018. Global rise in emerging alien species results from increased accessibility of new source pools. *Proceedings of the National Academy of Sciences of the United States of America* 115:E2264–E2273.
- Turbelin, A. J., B. D. Malamud, and R. A. Francis. 2017. Mapping the global state of invasive alien species: patterns of invasion and policy responses. *Global Ecology and Biogeography* 26:78–92.
- Turner, R. M., M. J. Plank, E. G. Brockerhoff, S. Pawson, A. Liebhold, and A. James. 2020. Considering unseen arrivals in predictions of establishment risk based on border biosecurity interceptions. *Ecological Applications* 30:e02194.
- Work, T. T., D. G. McCullough, J. F. Cavey, and R. Komsa. 2005. Arrival rate of nonindigenous insect species into the United States through foreign trade. *Biological Invasions* 7:323.
- Wylie, R., C. C. S. Yang, and K. Tsuji. 2020. Invader at the gate: The status of red imported fire ant in Australia and Asia. *Ecological Research* 35:6–16.
- Zhang, Z.-Q. 2011. *Animal biodiversity: An outline of higher-level classification and survey of taxonomic richness*. Magnolia Press, Auckland, New Zealand.
- Zobel, M., E. van der Maarel, and C. Dupré. 1998. Species pool: the concept, its determination and significance for community restoration. *Applied Vegetation Science* 1:55–66.

## SUPPORTING INFORMATION

Additional supporting information may be found online at: <http://onlinelibrary.wiley.com/doi/10.1002/eap.2412/full>

## OPEN RESEARCH

Summary data are provided in the supplementary material Data S1 and additional meta data in Appendix S1. Example R code for taxonomic processing is available in Data S2. The data accessibility for individual regions or countries are as follows: New Zealand: Sensitive data collected by the Ministry for Primary Industries, New Zealand, are subject to data agreements. For details of the data and how to request access contact Michael Ormsby at the Ministry for Primary Industries or [info@mpi.govt.nz](mailto:info@mpi.govt.nz). Australia: Sensitive data collected by the Australian Government Department of Agriculture, Water and the Environment are subject to data agreements. Details of the data and how to request access are available from the Australian Chief Plant Protection Office ([ACPPO@awe.gov.au](mailto:ACPPO@awe.gov.au)). Japan: Data collected by Ministry of Agriculture, Forestry and Fisheries (MAFF) are available: <http://www.pps.go.jp/TokeiWWW/Pages/report/index.xhtml> or contact MAFF via their website: <https://www.maff.go.jp/e/index.html>. Canada: Sensitive data collected by the Canadian Food Inspection Agency are available upon request. For details of the data and how to request access contact the Plant Health Lab Services Unit ([cfia.plantdiagnostics-diagnosticsvegetaux.acia@inspection.gc.ca](mailto:cfia.plantdiagnostics-diagnosticsvegetaux.acia@inspection.gc.ca)). USA: Sensitive data collected by the U.S. Department of Agriculture, Animal and Plant Health Inspection Service (APHIS), are subject to data agreements. For details of the data and how to request access contact Barney Caton at the United States Department of Agriculture, USA or [Phytosanitary.Advanced.Analytics.Team@usda.gov](mailto:Phytosanitary.Advanced.Analytics.Team@usda.gov). EPPO region: Data compiled by the European and Mediterranean Plant Protection Organization (EPPO) Secretariat are available from the EPPO reporting service: [https://www.eppo.int/RESOURCES/eppo\\_publications/eppo\\_reporting\\_service](https://www.eppo.int/RESOURCES/eppo_publications/eppo_reporting_service). Issues from the EPPO global database containing information on interceptions can be found by filtering the cumulative index for the keyword “interceptions”. United Kingdom: Data collected by the Department for Environment, Food & Rural Affairs are available upon request. To request data collected by Department for Environment, Food & Rural Affairs, please e-mail [plantpestrisks@defra.gov.uk](mailto:plantpestrisks@defra.gov.uk).