Predicting the spread of marine species introduced by global shipping

Hanno Seebens[a,b,c,1], Nicole Schwartz[a], Peter J. Schupp[a], and Bernd Blasius[a]

[a]Institute for Chemistry and Biology of the Marine Environment, University of Oldenburg, 26111 Oldenburg, Germany; [b]Department of Botany and Biodiversity Research, University of Vienna, 1030 Vienna, Austria; and [c]Senckenberg Biodiversity and Climate Research Centre, 60325 Frankfurt, Germany

The human-mediated translocation of species poses a distinct threat to nature, human health, and economy. Although existing models calculate the invasion probability of any species, frameworks for species-specific forecasts are still missing. Here, we developed a model approach using global ship movements and environmental conditions to simulate the successive global spread of marine alien species that allows predicting the identity of those species likely to arrive next in a given habitat. In a first step, we simulated the historical stepping-stone spreading dynamics of 40 marine alien species and compared predicted and observed alien species ranges. With an accuracy of 77%, the model correctly predicted the presence/absence of an alien species in an ecoregion. Spreading dynamics followed a common pattern with an initial invasion of most suitable habitats worldwide and a subsequent spread into neighboring habitats. In a second step, we used the reported distribution of 97 marine algal species with a known invasion history, and six species causing harmful algal blooms, to determine the ecoregions most likely to be invaded next under climate warming. Cluster analysis revealed that species can be classified according to three characteristic spreading profiles: emerging species, high-risk species, and widespread species. For the North Sea, the model predictions could be confirmed because two of the predicted high-risk species have recently invaded the North Sea. This study highlights that even simple models considering only shipping intensities and habitat matches are able to correctly predict the identity of the next invading marine species.

alien species | model predictions | species identity | marine ecoregion | climate change

The number of alien species transported by human assistance has increased rapidly during the last decades with serious consequences for native flora and fauna (1–3). These biological invasions are considered to be one of the major drivers of biodiversity changes (4–6). Once an unwanted alien species has naturalized in the new environment, it is nearly impossible to eradicate the species, and thus the mitigation of further introduction is the most efficient way of combating biological invasions (6, 7). However, a targeted monitoring and an efficient adaptive management requires knowledge about spreading dynamics of the next potential invaders and thus about the distribution of species, their invasiveness, and the likelihood of new introductions. Although all of these topics have been analyzed on their own, the potential to predict the spreading of alien species while combining these components remains to be tested.

A large amount of recent introductions can be attributed to the intensified global trade and transport as many species were accidentally or deliberately translocated through the exchange of commodities or the movements of transportation means (8, 9). The amount of exchanged commodities and the intensity of global traffic have therefore been found to be a good predictor to model the global spread of alien species (10–12). In most cases, predictions of alien species introductions are difficult to assess as model results could not be validated (i.e., quantitatively assessed using observed data) thoroughly due to the paucity of high-quality distributional data of alien species. Without any model validation, however, it is nearly impossible to assess the quality and the reliability of model predictions, which hampers the application of models for the management of alien species. In recent years, appropriate high-

quality data have been made accessible by various online databases, but testing model predictions with these data has still been lacking.

Model frameworks to predict the likelihood of new invasions have already been developed (10, 11, 13). However, these were not able to predict the identity of new invaders, but only the likelihood that any new species arrives from a certain source region on Earth. Here, we combined such a model, a slightly modified version of the vector-based model of marine invasion adopted from ref. 10, with datasets about the global distribution of marine alien species, which enabled us to predict the identity of the next species to arrive in a given local habitat. The model is a statistical model that describes how the probability a given species successfully invades a specific location depends on the shipping traffic and the environmental differences (temperature and salinity) between locations.

In a first step, to test the accuracy of the model, we used native ranges of 40 marine species from various taxonomic groups, ranging from algae to fish, as initial condition. For each species, we simulated the global spread outside its native range and compared the predicted alien range at each simulation time step with the observed one. This procedure allows the assessment of the quality of model predictions, although the degree of expansion distinctly varied among species, as some species are already widespread, whereas others occupy only a few alien regions either because they just started to spread or there are only a limited number of suitable habitats available.

In a second step, we used the reported distribution of 97 marine algal species with a known invasion history and six harmful algal species obtained from AlgaeBase (www.algaebase.org) and determined the species-specific invasion probabilities for each marine ecoregion not occupied by that species. Algae are particularly well suited for such an analysis because they are easily translocated by the exchange of ballast water, and especially invasive seaweeds are of global concerns because over 400 introductions have been reported

Significance

Predicting the arrival of alien species remains a big challenge, which is assumed to be a consequence of the complexity of the invasion process. Here, we demonstrate that spreading of alien marine species can be predicted by a simple model using only global shipping intensities, environmental variables, and species occurrence data. We provide species lists of the next potentially invading species in a local habitat or species causing harmful algal blooms with their associated probability of invasion. This will help to improve mitigation strategies to reduce the further introduction of alien species. Although this study focuses on marine algae, the model approach can be easily adopted to other taxonomic groups and their respective drivers of invasion.
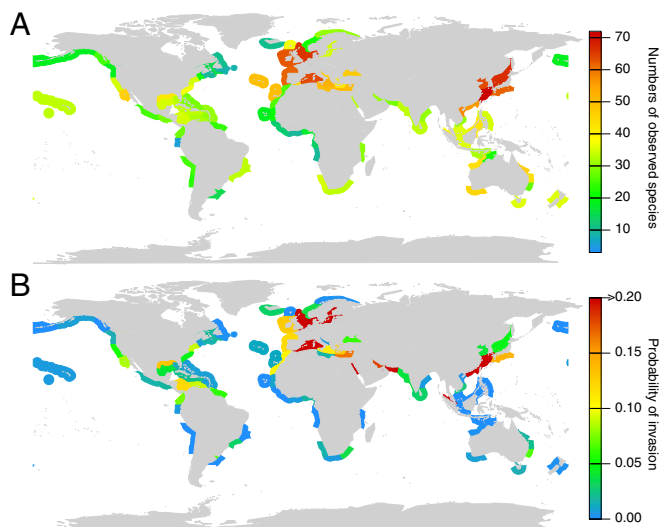
Fig. 1. Observed cumulative distribution (native plus alien) of 97 marine algal species with an invasion history (*A*) and predicted invasion probabilities for further spread of these species to formerly unoccupied marine ecoregions. Colors indicate the number of reported algal species in *A* and the predicted probability of invasion in *B*, respectively.

worldwide (14). Furthermore, seaweeds deeply shape marine ecosystems and they can have strong detrimental ecological and economic impacts (14, 15). Using our model, we identified the likely hot spots of future invasions among 90 marine ecoregions of the world and algal species with the highest probability to arrive next.

Results

Comparing observed and predicted alien ranges of 40 marine species reveals that the model correctly predicted the presence/absence of the species in an originally alien ecoregion in 77% of all cases (median of all 40 spread simulations; Fig. S1). This value is robust to interspecific variation (e.g., using only 50% of the species reduced the number of correctly predicted ecoregions only to a median of 75%; Fig. S2), model parameterization, model structure, and variation of shipping intensity (*SI Text* and Table S1). Compared with observational data, the model overpredicts the number of alien species per ecoregion (Fig. S1). Predicted alien species numbers were distinctly higher for Southern Africa and East Asia, which are less studied regions, and thus higher alien species numbers can be expected, but also for the US West Coast.

The simulations of the global spreading dynamics of the 40 marine species were used to analyze common patterns of the global spread. For each simulated invasion step of each species, we determined the great circle distance from the centroids of the newly invaded ecoregion to the nearest already occupied ecoregion. This shows that long-distance jumps are very common among the very first invasion steps (Fig. S3). For 44% of all species spreads, the longest jump to a new ecoregion was found among the first three invasion steps with 10,376 km on average. That is, spreading dynamics of these species follow a common pattern: first, the suitable habitats worldwide were invaded irrespective of the distance from where the species originated, and subsequently the species spread to neighboring habitats. This results in a mean geographic distance between observed native and alien ranges of around 10,200 km. Alien species numbers weighted by their number of native and alien regions decreased to shorter and longer distances (Fig. S4). A similar although more complex pattern was also found in ref. 10, which indicates that the models applied in both studies revealed similar results (*Materials and Methods*).

After validation of model predictions, we used the distribution of 97 algal species with a known invasion history obtained from AlgaeBase as initial conditions to determine the species-specific

invasion probabilities to unoccupied ecoregions. These species mostly live in temperate to subtropical marine ecoregions in Europe and East Asia (Fig. 1*A*), corresponding to the known global hot spots of algal diversity (16, 17). The predicted invasion probabilities combined for all of these species reflect the global shipping intensity as marine ecoregions with intense ship traffic such as East Asian Seas or Northern European Seas have a high invasion probability (Fig. 1*B*). The East Asian Seas and the Northern European Seas are characterized by similar environmental conditions and are well connected by ships, which increases the likelihood of exchanges between both regions. Indeed, the reported native and alien ranges show that species have been mutually exchanged between both regions as many alien species in the Northern European Seas originate from East Asia (*n* = 11 of 19) and vice versa (4 of 16). Our model calculations revealed rankings of high-risk algal species for each ecoregion, which is exemplarily shown for the North Sea (Table S2). Two of the algal species of the top 10 high-risk species for the North Sea are known from the literature to have naturalized in that ecoregion in recent years [*Prorocentrum minimum* (18) and *Polysiphonia harveyi* (19)]. These species are not listed for the North Sea in AlgaeBase probably due to delays in data acquisition. That is, based on data provided by AlgaeBase, the model predicts a high probability that these species will enter the North Sea, which could indeed be confirmed by recent studies.

Predicted climate warming will likely modify the similarity of environmental conditions of marine ecoregions and thus will affect the probabilities of invasion. Using predicted mean sea surface temperatures during 2040–2060 revealed a latitudinal gradient in the change of invasion probabilities (Fig. 2): in general, the model predicts a decrease in invasion probabilities in the tropics, and an increase in temperate regions due to climate change. Highest increases in invasion probabilities can be expected for the Northeast Pacific and the Baltic Sea, whereas highest declines arise for ecoregions with high shipping intensity in the tropics such as ecoregions with access to the Panama Canal, the Persian Gulf, the Strait of Malacca, and South China.

The invasion probabilities distinctly vary among algal species, which—in our model—mainly depends on the species-specific global distribution of algae and their occupancy of ecoregions with high shipping intensity (thereby enhancing the chance of further spread). We found that the total invasion probability $P_{tot}(Inv)$ (i.e., the probability for a species to invade any unoccupied ecoregion worldwide) is a hump-shaped function of the initial number of occupied marine ecoregions (Fig. 3*A*). This can be explained by two contrasting mechanisms: increasing the number of occupied ecoregions results (*i*) in an increase in propagule pressure to unoccupied ecoregions, thereby increasing the likelihood of further spread, and (*ii*) in a reduction of the number of potential new habitats, which simultaneously decreases the chance of further spread. Combining both relationships results in the described hump-shaped curve.
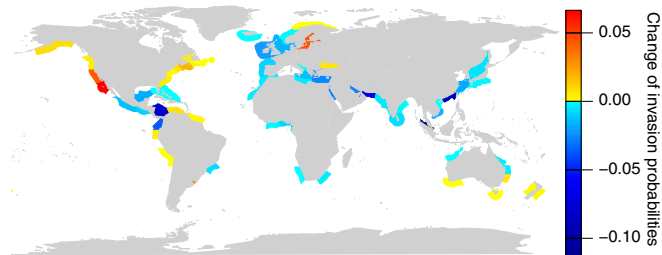

Fig. 2. Predicted changes in invasion probabilities due to climate warming. To simulate climate warming, invasion probabilities for 97 algal species with a known invasion history were calculated using water temperatures during 2040–2060 for the recipient ecoregions. The resulting invasion probabilities were compared with those calculated using current water temperatures, and the deviations are indicated by colors.

Consequently, species with narrow or wide distributions have a comparatively low total invasion probability, whereas species with an intermediate distribution have the highest invasion probabilities. In addition to the number of occupied ecoregions, the interspecific variation of $P_{tot}(Inv)$ can be further explained by the maximum invasion probability $P_{max}(Inv)$ to any ecoregion (Fig. 3B).

To characterize the spreading potential of each species, we introduced an invasion threshold ε and counted the number of ecoregions with an invasion probability above this threshold. Note that a decrease of ε can be interpreted as a general increase in the probability of invasion due to e.g., elevated propagule pressure, interspecific variations in the ability to invade another ecoregion, or increased environmental match due to environmental changes. Decreasing ε (or increasing the invasion probability) leads to a higher number of ecoregions that potentially can be invaded by the species (Fig. 3C). The obtained species-specific invasion curves saturate when the species occupy all initially unoccupied ecoregions worldwide. For example, species with a wide initial distribution can only invade a few remaining unoccupied ecoregions, resulting in a low number of ecoregions at which the invasion curves saturate. Cluster analysis reveals three characteristic groups of invasion curves: the first group, which we call the "emerging species," has a narrow initial distribution (red dots in Fig. 3A and red lines in Fig. 3 C and D). A decrease in the invasion threshold ε results in a slow increase of the number of unoccupied ecoregions with an invasion threshold above ε. That is, the species have a low chance to invade another ecoregion. Species of the second group, called the "high-risk species," have an intermediate initial distribution (green dots in Fig. 3A). Their spreading potential increases comparatively steeply even at high invasion thresholds (green lines in Fig. 3 C and D), and thus these species have a high chance to invade other ecoregions. The third group of species, called the "widespread species," consists of species with a wide initial distribution (black dots in Fig. 3A). Their invasion curves increase with decreasing invasion threshold at a comparatively low rate similar to that of the emerging species (black lines in Fig. 3 C and D), but saturate at a low number of unoccupied ecoregions simply due to the low availability of ecoregion without the species.

Although the known invasion of a species in an ecoregion is the best predictor for the potential invasiveness of that species for other ecoregions (20, 21), it is interesting to assess the spreading potentials of species, which are not known for their invasiveness but may affect other species and potentially even ecosystems. Some algae produce toxic substances causing fish or shellfish poisoning, which can also be harmful for humans (22). Transportation of toxic algal cells or cysts in ballast water of ships is a likely explanation for harmful algal blooms in previously unaffected regions (22, 23). Although these species may not be known to naturalize outside their native region (i.e., to establish populations viable for years), they may produce harmful algal blooms in these regions at least during one season of favorable conditions. We therefore selected the distributions of six algal species from AlgaeBase responsible for either well-known human diseases such as ciguatera or various shellfish poisonings (24) and calculated invasion probabilities for each species. Most of these species have an intermediate to wide distribution and a comparatively high $P_{tot}(Inv)$ (Fig. 4). *Karenia brevis*, for example, causing neurotoxic shellfish poisoning, has a $P_{tot}(Inv)$ of 0.105, which is in the range of the highest invasion probabilities calculated for species with a known invasion history. The reason is that the *K. brevis* was found in an intermediate number of ecoregions with intense ship traffic (Fig. 4). According to our classification system for species with known invasion histories (Fig. 3), *Gambierdiscus toxicus* would fall in the category of emerging species, whereas *Protoperidinium crassipes*, *Dinophysis acuminata*, and *Karenia brevis* could be classified as high-risk species, and the remaining algae would be widespread species.

Discussion

The mitigation of further introductions of alien species is a challenging task mostly due to the complexity of the invasion process and the lack of data. Models can help to improve mitigation strategies by the determination of hot spots of biological invasions, the determination of high-risk pathways, and the identification of species likely to arrive next. Several studies have tried to predict the likelihood of new invasions (10, 13, 25–28); however, the common paucity of model validations hampers the assessment of the quality of model predictions and their comparison between different approaches. Here, we provide a robust approach for the validation of colonization models, simulating the spread of alien species, which may serve as an example for other models and taxonomic groups. We show that our model correctly predicts the invasion status of species in ecoregions outside their native range in good agreement with observations (77% of the presence/absence of an alien species correctly predicted). This is particularly striking because of the simplicity of the model, considering only habitat matches and maritime traffic intensity and lacking seemingly important predictors like species-specific characteristics, biotic interactions, additional environmental variables, or historic shipping data. Indeed, the consideration of additional parameters in the model such as further environmental variables did not improve the model fit (Table S1). The quality of the fit ranges between those found in other modeling studies [e.g., $R^2 = 0.64$ for the global spread of terrestrial vascular plants (29) or an 87–94% accuracy for the introduction of fishes into the Great Lakes (27)], but other models were either more complex or restricted to much smaller geographic areas or a single taxonomic group. We are not aware of any study,
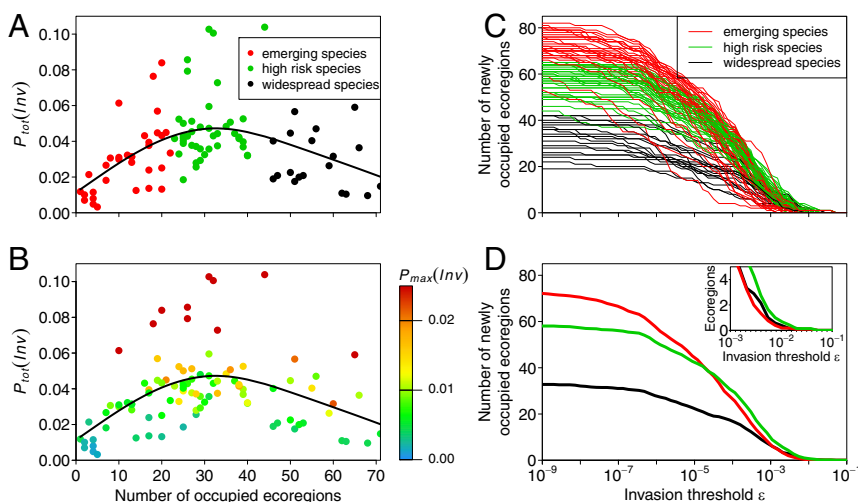

Fig. 3. Spreading potential of marine alien algae. (*A*) The total invasion probabilities $P_{tot}(Inv)$ of algal species follow a hump-shaped curve of the number of initially occupied marine ecoregions, with species of intermediate global distribution having the highest $P_{tot}(Inv)$. The colors denote the category of spreading potentials classified by cluster analysis (*C* and *D*). (*B*) The same data as in *A*, but the colors indicate the maximum invasion probability $P_{max}(Inv)$ from the occupied range of the species to any unoccupied ecoregion. (*C*) Invasion curves for each species estimated as the number of ecoregions outside the range, which the species initially occupied, above an invasion threshold ε. Cluster analysis of interspecific variations identified three groups of species with different invasion curves (red, emerging species; green, high-risk species; and black, widespread species). (*D*) The same as *C*, but showing the mean invasion curves for each cluster. The *Inset* indicates a zoom to high invasion threshold values.
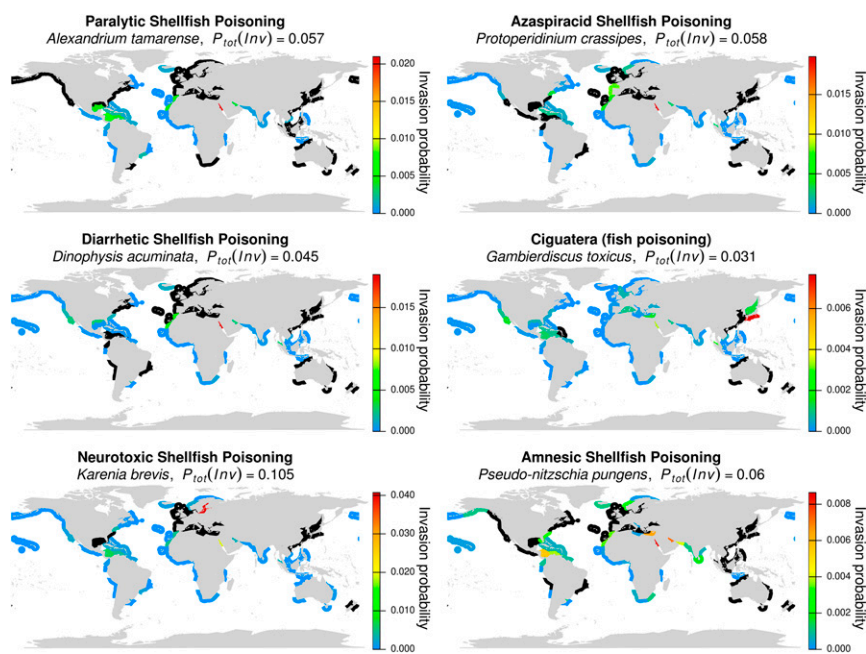
Fig. 4. Invasion probabilities for a selection of six algal species causing fish and shellfish poisonings potentially harmful to humans. The selected algae are known to produce toxic compounds causing the respective poisoning of fish or shellfish. Black areas denote the current distribution of species, whereas colors indicate the invasion probability to an unoccupied ecoregion. The total invasion probabilities $P_{tot}(Inv)$ expressed as the integration of all single invasion probabilities are given in the subheading of each panel.

simulating the global spread of marine species from various taxonomic groups simultaneously as done here, which renders a direct comparison difficult. Probably the best indication of the model's quality is the fact that two of the algal species predicted to be among the top 10 high-risk species for the North Sea are known to have naturalized in that ecoregion in recent years. That is, in two cases our species-specific predictions could be confirmed by recent samplings, which are not yet included in AlgaeBase. This also shows that it is possible to explain the spread of alien species to a large degree by simple vector-based invasion models. The sensitivity analysis (*SI Text*) highlights that our model predictions are robust to the choice of species for model validation, and to the variation in model structure, parameterization, and input datasets.

For the 97 marine algal species considered in this study, the highest probability of invasion arises in Asian and European ecoregions (Fig. 1*B*), mostly along the major shipping routes. The Asian Seas were also identified as invasion hot spots in a previous analysis of marine invasion, thereby ignoring observed species distributions (10). In contrast to our study, the invasion probabilities of Northern European Seas were very low in ref. 10, which was explained by the, on average, low environmental similarity to most other ecoregions worldwide. The reason for the elevated invasion probabilities found in this study is that many algal species predicted to be at high risk in Northern European Seas are native to East Asian waters, which provide similar environmental conditions to Northern European Seas. As both regions are highly interconnected due to intense ship traffic, the probability to invade from one region to the other is also high, resulting in the described mutual exchange of species between these regions. This shows that, although the overall patterns of both studies were similar, observational data are important to further improve the predictions of invasion hot spots.

Our model predicts that climate warming will lead to reduced invasion probabilities in the tropics and elevated ones in temperate regions, particularly in North America (Fig. 2). The reason is that many species used in this study live in temperate to subtropical ecoregions and were assumed to be adapted to water temperatures of those regions. If water temperatures increase, these species will find appropriate environmental conditions at higher latitudes than inhabited now, whereas environmental conditions will get less suitable in tropical regions. The highest increase in invasion probability arises for the Northeast Pacific. For this ecoregion, the by-far most important donor area constitutes the Northwest Pacific

(10), and an elevation of water temperatures in the Northeast Pacific will increase the environmental match of both regions, resulting in increased invasion probabilities. Indeed, a recent finding of the Asian macroalga *Sargassum horneri* at the West Coast of Mexico and the United States is likely is a consequence of this increased environmental match (30), which further supports our model predictions.

Our model predicts the total invasion probability of a species to be a hump-shaped function of the initially occupied number of ecoregions (Fig. 3 *A* and *B*), the reason being that species with an intermediate global distribution pose a large propagule pressure to a comparatively large number of unoccupied ecoregions. This pattern corresponds to the relationship of the colonization rate on the number of occupied patches of classical metacommunities models (31). In these models, the colonization rate of new patches has to be zero if either none or all patches are occupied by a species and positive in between, which indicates the close relationship between both types of models.

The reliable prediction of invasion dynamics remains one of the biggest challenges in invasion ecology. We here demonstrate an approach for a robust validation of global invasion models, which allows the assessment of the quality of model predictions. This study highlights that the combination of invasion models with observational data can essentially improve the predictions of invasion probabilities. The model applied here can be easily adopted to simulate the spread of other taxonomic groups (see ref. 29 for an example of terrestrial vascular plants). Many online databases emerged in recent years. Notable examples are the Delivering Alien Invasive Species Inventories for Europe (DAISIE) (www.europe-aliens.org), the Invasive Species Compendium by the Centre for Agriculture and Bioscience International (CABI), or the upcoming Global Register of Invasive Species (GRIS) by the Invasive Species Specialists Group (www.issg.org), all of which provide native and alien ranges of numerous species. It is therefore the logical consequence to combine invasion models with available distributional data to improve the predictability of invasion dynamics as shown here. The prediction of future invasions is a prerequisite of efficient mitigation strategies, and the determination of the species identity enables a targeted monitoring of potential high-risk species.

Materials and Methods

Data. As input variables, the model requires data about the arrival and departure dates of single ships moving between ports, ship type, ship size, ballast tank volumes, and environmental data of ports. The ship and route-specific data (i.e., ship size, ship type, and arrival and departure dates) were obtained from a large dataset of ship movements and ship characteristics provided by Lloyd's Register Fairplay (www.ihs.com), consisting of nearly 3 million ship voyages of 32,511 ships between 1,469 ports during 2007–2008 (10), and ballast tank volume data taken from the American Bureau of Shipping (32). The shipping intensity was assumed to be constant in time (but see the discussion about the reliability of using recent ship movement data to predict historic spreading dynamics in *SI Text*). Sea surface temperatures and nutrient concentrations (nitrate, phosphate, and silicate) were obtained from the World Ocean Atlas (WOA) (www.nodc.noaa.gov), providing 50-y averages at 1° spatial resolution. Surface salinity data of ports were calculated from port-specific data of water densities provided by Lloyd's Register Fairplay (www.sea-web.com/portguide.html) for most of the ports (69%). Water densities were recalculated to salinities using temperatures taken from the WOA. For the remaining ports, salinities were taken from the WOA. Future sea surface temperatures were obtained from the Coupled Model Intercomparison Project (CMIP5) (cmip-pcmdi.llnl.gov/cmip5/) providing predictions of sea surface temperatures using an Intergovernmental Panel on Climate Change scenario of intermediate greenhouse gas concentration (RCP6) (33). The average sea surface temperatures were calculated for each marine ecoregion for 2040–2060.

Model. We applied a vector-based model of marine invasion to calculate the probability of invasion by global shipping between 1,469 major ports worldwide (10). In the original form, the model consisted of three independent probabilities: the probability to be alien, P(Alien); the probability of introduction, P(Intro); and the probability of establishment, P(Estab). Here, however, when we combine the model with observational data of species distributions, the term P(Alien), constituting a theoretical estimator for biogeographical dissimilarity is not required anymore. We therefore removed P(Alien) from the model.

According to ref. 10, the probability of introduction,

$$P_r(Intro) = \left(1 - e^{-\lambda B_r}\right) e^{-\Delta t_r}, \qquad [1]$$

is a function of exchanged ballast water volume B_r on route r between donor port i and recipient port j and travel time Δt. $P_r(Intro)$ is obtained for each single ship on a certain route r taken from the ship movement dataset. B_r is calculated as $B_r = zW_r(1 - zW_r/V_r)^{\delta_r}$, with W_r being the ship type- and ship size-specific amount of released ballast water in cubic meters obtained from regression fits shown in figure S1 in ref. 10, V_r denoting the ship size-specific mean volume of ballast tanks of a ship, δ_r representing the number of stopover ports on route r, and z being the fraction of nonzero releases depending on the type and the size of a ship. Note that in ref. 10, z was falsely described as the fraction of zero releases, although it has to be the fraction of nonzero releases as stated here.

The probability of establishment,

$$P_{ij}(Estab) = \alpha e^{-\frac{1}{2}\left[\left(\frac{\Delta T_{ij}}{\sigma_T}\right) + \left(\frac{\Delta S_{ij}}{\sigma_S}\right)\right]}, \qquad [2]$$

was modeled as a Gaussian function of differences of water temperature ΔT_{ij} and salinity ΔS_{ij} of donor port i and recipient port j standardized by the width of the respective ecological niche σ_T and σ_S. α represents the basic probability of establishment. To analyze the influence of climate warming on model predictions, $P_{ij}(Estab)$ was calculated using future sea surface temperature (2040–2060; *Data*) for the recipient regions.

The probability of invasion from port i to port j is given by the complement of species failing to invade on all ship routes r_{ij} connecting both ports,

$$P_{ij}(Inv) = 1 - \prod_{r_{ij}}\left[1 - P_r(Intro)P_{ij}(Estab)\right]. \qquad [3]$$

The invasion probabilities between ports were aggregated accordingly to obtain the invasion probabilities $P_j(Inv)$ from the ecoregions occupied by the respective species to each unoccupied ecoregion j. The total invasion probability was calculated for each species as $P_{tot}(Inv) = 1 - \prod_j[1 - P_j(Inv)]$ and the maximum probability of invasion as $P_{max}(Inv) = \max(P_j(Inv))$.

Although some modifications of the model structure and parameterizations slightly improved the accuracy of model predictions (*SI Text* and Table S1), we adopted the basic model version and the parameterization provided by ref. 10 for consistency. A more detailed description of the model, its parameterization, the input variables, and tests of the robustness of this approach are provided in ref. 10.

Model Validation. For model validation (i.e., the quantitative assessment of the quality of model predictions using observed data), we compiled data of native and alien distributions of marine alien species from the CABI Invasive Species Compendium (www.cabi.org/isc/). In a first step, we used the implemented search engine of CABI with the search term "(HAB marine) AND ballast water" to filter for species in marine habitats likely to be transported by ballast water. We manually removed species not fitting to this category as, e.g., freshwater species with a high salinity tolerance were included as well, with the exception of *Dreissena polymorpha* as this species is known to be transported by open-sea vessels (34). We also removed regions with uncertain invasion status of a species, which resulted in a total of 40 species with known native and alien distributions from various taxonomic groups ranging from algae to fish. The following species were used for model validation: *Acanthogobius flavimanus, Acentrogobius pflaumii, Ascidiella aspersa, Asterias amurensis, Austrominius modestus, Caprella mutica, Carcinus maenas, Caulerpa racemosa var. cylindracea, Charybdis hellerii, Ciona intestinalis, Ciona savignyi, Crassostrea virginica, Crepidula fornicata, Diplosoma listerianum, Dreissena polymorpha, Ensis directus, Gracilaria salicornia, Grateloupia turuturu, Hemigrapsus sanguineus, Hemigrapsus takanoi, Hemimysis anomala, Littorina littorea, Marenzelleria neglecta, Microcosmus squamiger, Mnemiopsis leidyi, Musculista senhousia, Mytilus galloprovincialis, Palaemon elegans, Palaemon macrodactylus, Phyllorhiza punctata, Polyandrocarpa zorritensis, Pseudochattonella verruculosa, Pterois volitans, Rapana venosa, Rhithropanopeus harrisii, Schizoporella errata, Spartina alterniflora, Styela plicata, Ulva pertusa,* and *Ulva reticulate*. CABI provides distributions usually on a country basis, but in some cases also subnational units such as states, provinces, or islands are provided. We translated the distributional data provided by CABI into marine ecoregions, thereby assuming that if a species was listed for a geographic unit bordering or located in a certain ecoregion, the species can be found in the whole ecoregion including the ports. This may result in an overprediction of the actual distribution of species, but given the highly patchy sampling coverage of the world's coastlines, simulations can only be done on a coarse geographic resolution, which necessitates such simplifications. A marine ecoregion represents an area of similar environmental conditions and species assemblages, and thus it seems likely that a species can also be found in other parts of that ecoregion. We adopted the classification of "marine ecoregions of the world" by Spalding et al. (35) but considered only those ecoregions that have at least one port listed in our shipping database, resulting in a total of 90 marine ecoregions. To assess the reliability of using recent ship movement data for the prediction of historic spreading dynamics of alien species, we compiled a list of the years of first record of the alien species in a region (Table S3). Inevitably, the year of first record is not provided for all species and all regions due to the lack of data.

For each species, we calculated the mean temperature and salinity requirements from their known native distribution by taking the average of all native ecoregion means of temperature and salinity. The obtained species-specific ecological niche was used as input variables T_i and S_i in Eq. **2**. Note that i now represents the geographic area used to calculate the ecological niche of a species. As initial condition, we set all ports where the species is considered to be native as "occupied" and all other ports to "unoccupied." We then applied the model to calculate the invasion probability to all unoccupied ports. The port with the highest invasion probability was identified and set to occupied. This was repeated until a threshold of a very low invasion probability was reached, which corresponds to a hypothetical nearly worldwide coverage of the species. The sequence of invaded ports was transformed to the sequence of invaded ecoregions, thereby removing duplicated ecoregions. For each time step of the modeled spread, we compared the predicted alien distribution of the species with the observed alien distribution obtained from CABI and calculated the goodness-of-fit of predicted and observed alien distributions. The goodness-of-fit was estimated by the number of correctly predicted invaded ecoregions (true positives) divided by the sum of false predictions (false positives plus false negatives). At the time step of the largest goodness-of-fit value, the deviation between predicted and observed distribution was lowest, and the respective distribution was selected as the best fit. This procedure was repeated for each of the 40 species and the median of the resulting 40 goodness-of-fit values was taken as the overall goodness-of-fit of the model.

To test the robustness of our model predictions, we performed an extensive sensitivity analysis, thereby varying the simulation process, the model parameterization, and the distributional data used for model validation, and discuss the reliability of the ship movement data for our modeling purpose (*SI Text*).

Model Application. After model validation, we calculated invasion probabilities using the global distribution of marine algae taken from AlgaeBase (www.algaebase.org) to identify future invasion hot spots and high-risk species of marine algae. We used the species names listed as "accepted names" in the database. AlgaeBase provides detailed information for numerous algal species, including their current global distribution, but AlgaeBase does not provide the information whether a species is native or alien. We therefore adopted the finding of previous studies that the best forecasting tool to predict the invasion risk is whether the species has naturalized elsewhere (20), which is sometimes even taken as the only permanent predictor of invasion success (21). We screened the literature and other databases to compile a list of algal species with a known invasion history. We started with a list of algal species provided by Molnar et al. (36) and added other species from the literature and online databases and finally ended up with 97 algal species with a known invasion history. For these species, we selected their current distribution from AlgaeBase. For each species, the ecological niche represented by T_i and S_i was calculated as the mean temperatures and salinities of all occupied ecoregions. We standardized the geographic units to the level of a country, island, or marine ecoregion depending on the level of detail provided by AlgaeBase. For large countries (United States, Canada, Russia, and Australia), we distinguish subnational units such as provinces or states if this information was provided. For example, if a species was listed for Australia, we assigned all ecoregions surrounding Australia to this species, whereas if the species was found only at the island of Tasmania or at the coast of Queensland, we only selected the respective ecoregions. Assuming that the species is homogeneously distributed in the respective geographic unit (along a country coast or in a marine ecoregion), all ports in that area were set to occupied. For instance, if the species was found

in the North Sea, we set all North Sea ports to occupied, whereas if the species was listed for Germany, we selected the North Sea ports and Baltic Sea ports located in Germany as occupied ports. For each species, we then calculated the invasion probabilities to any unoccupied port.

Spreading Potentials of Species. The invasion probabilities obtained from the model applied to the AlgaeBase data were used to characterize the spreading potential of each species. This was done by introducing a threshold ε of invasion probability ranging from 10^{-9} to 10^{-1}. For each species, the ecoregions outside the species' native range with an invasion probability above ε were counted. This can be interpreted as those ecoregions that can be invaded by the species given a certain threshold value. While successively reducing ε, the number of these ecoregions with an invasion probability above ε increased, but at different rates depending on the species-specific spreading potentials resulting in species-specific invasion curves. To characterize the interspecific differences, we calculated the area bounded between all pairs of invasion curves and applied a hierarchical clustering algorithm using the open-source software package R (37). The distances between clusters were calculated with "Ward's minimum variance method" (38), which aims at finding compact spherical clusters. We tested other clustering algorithms as well, which, however, did not change the results significantly.

1. Simberloff D, et al. (2013) Impacts of biological invasions: What's what and the way forward. *Trends Ecol Evol* 28(1):58–66.
2. Pyšek P, et al. (2012) A global assessment of invasive plant impacts on resident species, communities and ecosystems: The interaction of impact measures, invading species' traits and environment. *Glob Change Biol* 18(5):1725–1737.
3. McGeoch MA, et al. (2010) Global indicators of biological invasion: Species numbers, biodiversity impact and policy responses. *Divers Distrib* 16(1):95–108.
4. Butchart SHM, et al. (2010) Global biodiversity: Indicators of recent declines. *Science* 328(5982):1164–1168.
5. Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-Being: Biodiversity Synthesis* (Island Press, Washington, DC).
6. Pyšek P, Richardson DM (2010) Invasive species, environmental change and management, and health. *Annu Rev Environ Resour* 35(1):25–55.
7. Puth LM, Post DM (2005) Studying invasion: Have we missed the boat? *Ecol Lett* 8(7):715–721.
8. Levine JM, D'Antonio CM (2003) Forecasting biological invasions with increasing international trade. *Conserv Biol* 17(1):322–326.
9. Hulme PE, et al. (2008) Grasping at the routes of biological invasions: A framework for integrating pathways into policy. *J Appl Ecol* 45(2):403–414.
10. Seebens H, Gastner MT, Blasius B (2013) The risk of marine bioinvasion caused by global shipping. *Ecol Lett* 16(6):782–790.
11. Keller RP, Drake JM, Drew MB, Lodge DM (2010) Linking environmental conditions and ship movements to estimate invasive species transport across the global shipping network. *Divers Distrib* 17(1):93–102.
12. Drake JM, Lodge DM (2004) Global hot spots of biological invasions: Evaluating options for ballast-water management. *Proc Biol Sci* 271(1539):575–580.
13. Jerde CL, Lewis MA (2007) Waiting for invasions: A framework for the arrival of nonindigenous species. *Am Nat* 170(1):1–9.
14. Williams SL, Smith JE (2007) A global review of the distribution, taxonomy, and impacts of introduced seaweeds. *Annu Rev Ecol Evol Syst* 38:327–359.
15. Wallentinus I, Nyberg CD (2007) Introduced marine organisms as habitat modifiers. *Mar Pollut Bull* 55(7-9):323–332.
16. Bolton JJ (1994) Global seaweed diversity: Patterns and anomalies. *Bot Mar* 37(3):241–246.
17. Kerswell AP (2006) Global biodiversity patterns of benthic marine algae. *Ecology* 87(10):2479–2488.
18. Hoppenrath M (2004) A revised checklist of planktonic diatoms and dinoflagellates from Helgoland (North Sea, German Bight). *Helgol Mar Res* 58(4):243–251.
19. Maggs CA, Stegenga H (1998) Red algal exotics on North Sea coasts. *Helgol Meersunters* 52(3-4):243–258.
20. Kulhanek SA, Ricciardi A, Leung B (2011) Is invasion history a useful tool for predicting the impacts of the world's worst aquatic invasive species? *Ecol Appl* 21(1):189–202.
21. Williamson M (1999) Invasions. *Ecography (Cop)* 22(1):5–12.
22. Hallegraeff GM (1993) A review of harmful algal blooms and their apparent global increase. *Phycologia* 32(2):79–99.
23. Van Dolah FM (2000) Marine algal toxins: Origins, health effects, and their increased occurrence. *Environ Health Perspect* 108(Suppl 1):133–141.
24. Wang D-Z (2008) Neurotoxins from marine dinoflagellates: A brief review. *Mar Drugs* 6(2):349–371.
25. Herborg LM, Jerde CL, Lodge DM, Ruiz GM, MacIsaac HJ (2007) Predicting invasion risk using measures of introduction effort and environmental niche models. *Ecol Appl* 17(3):663–674.
26. Paini DR, Yemshanov D (2012) Modelling the arrival of invasive organisms via the international marine shipping network: A Khapra beetle study. *PLoS One* 7(9):e44589.
27. Kolar CS, Lodge DM (2002) Ecological predictions and risk assessment for alien fishes in North America. *Science* 298(5596):1233–1236.
28. Gallien L, Munkemuller T, Albert CH, Boulangeat I, Thuiller W (2010) Predicting potential distributions of invasive species: Where to go from here? *Divers Distrib* 16(3):331–342.
29. Seebens H, et al. (2015) Global trade will accelerate plant invasions in emerging economies under climate change. *Glob Change Biol* 21(11):4128–4140.
30. Marks LM, et al. (2015) Range expansion of a non-native, invasive macroalga *Sargassum horneri* (Turner) C. Agardh, 1820 in the eastern Pacific. *BioInvasions Rec* 4(4):243–248.
31. Levins R (1969) Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bull Entomol Soc Am* 15:237–240.
32. American Bureau of Shipping (2011) *Ballast Water Treatment Advisory* (American Bureau of Shipping, Houston). Available at ww2.eagle.org/content/eagle/en.html. Accessed October 23, 2012.
33. Moss R, et al. (2008) *Towards New Scenarios for Analysis of Emissions, Climate Change, Impacts, and Response Strategies* (Intergovernmental Panel on Climate Change, Geneva).
34. Ricciardi I, MacIsaac HJ (2000) Recent mass invasion of the North American Great Lakes by Ponto-Caspian species. *Trends Ecol Evol* 15(2):62–65.
35. Spalding MD, et al. (2007) Marine ecoregions of the world: A bioregionalization of coastal and shelf areas. *Bioscience* 57(7):573–583.
36. Molnar JL, Gamboa RL, Revenga C, Spalding MD (2008) Assessing the global threat of invasive species to marine biodiversity. *Front Ecol Environ* 6(9):485–492.
37. R Core Team (2014) R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, Vienna), Version 3.2.3. Available at www.r-project.org.
38. Ward JH (1963) Hierarchical grouping to optimize an objective function. *J Am Stat Assoc* 58(301):236–244.
39. Hewitt CL, et al. (2004) Introduced and cryptogenic species in Port Phillip Bay, Victoria, Australia. *Mar Biol* 144(1):183–202.
40. DAISIE (2009) *Handbook of Alien Species in Europe* (Springer, Dordrecht, The Netherlands).
41. O'Flynn C, Kelly J, Lysaght L (2014) Ireland's invasive and non-native species—trends in introductions. *Natl Biodivers Data Cent Ser* 2:1–50.
42. Bungartz F, et al., eds (2009) *Charles Darwin Foundation Galapagos Species Checklist—Lista de Especies de Galápagos de la Fundación Charles Darwin*. Available at www.darwinfoundation.org/datazone/checklists/. Accessed November 9, 2015.
43. Williams SL (2007) Introduced species in seagrass ecosystems: Status and concerns. *J Exp Mar Biol Ecol* 350(1-2):89–110.
44. Çinar ME, Bilecenoğlu M, Öztürk B, Katagan T, Aysel V (2005) Alien species on the coasts of Turkey. *Mediterr Mar Sci* 6(2):119–146.
45. Carlton JT, Eldredge LG (2009) Marine bioinvasions of Hawai'i: The introduced and cryptogenic marine and estuarine animals and plants of the Hawaiian Archipelago. *Bish Museum Bull Cult Environ Stud* 4:1–203.
46. Coles SL, DeFelice RC, Eldredge LG, Carlton JT (1999) Historical and recent introductions of non-indigenous marine species into Pearl Harbor, Oahu, Hawaiian Islands. *Mar Biol* 135(1):147–158.
47. Roy HE, et al. (2012) *Non-Native Species in Great Britain: Establishment, Detection and Reporting to Inform Effective Decision Making* (Non-Native Species Secretariat, York, UK).
48. Wu S, et al. (2010) Insights of the latest naturalized flora of Taiwan: Change in the past eight years. *Taiwania* 55(2):139–159.
49. Xu H, et al. (2012) An inventory of invasive alien species in China. *NeoBiota* 15:1–26.