Global port survey quantifies commercial shipping's effect on biodiversity

Authors

Jose Andrés^{1,2§}, Paul Czechowski^{1,3§}, Erin Grey^{4,5*§}, Mandana Saebi⁶, Kara Andres^{1,2}, Christopher Brown⁶, Nitesh Chawla⁷, James J. Corbett⁸, Rein Brys⁹, Phillip Cassey¹⁰, Nancy Correa¹¹, Marty R. Deveney¹², Scott P. Egan¹³, Joshua P. Fisher¹⁴, Rian vanden Hooff¹⁵, Charles R. Knapp¹⁶, Sandric Chee Yew Leong¹⁷, Brian J. Neilson¹⁸, Esteban M. Paolucci¹⁹, Michael E. Pfrender²⁰, Meredith R. Pochardt²¹, Thomas A.A. Prowse¹⁰, Steven S. Rumrill²², Chris Scianni²³, Francisco Sylvester²⁴, Mario N. Tamburri²⁵, Thomas W. Therriault²⁶, Darren C.J. Yeo²⁷, David Lodge^{1,2}

Affiliations

¹ Cornell University, Department of Ecology and Evolutionary Biology, 215 Tower Road, Ithaca, NY 14853, United States of America.

² Cornell University, Cornell Atkinson Center for Sustainability, 340 Tower Road, Ithaca, NY 14853, United States of America.

³ University of Otago, Department of Anatomy, 270 Great King Street, Dunedin, OTA 9016, New Zealand.

⁴ University of Maine, School of Biology and Ecology and Maine Center for Genetics in the Environment, 100 Murray Hall, Orono, ME 04469, United States of America.

⁵ Governors State University, Division of Science, Mathematics and Technology, 1 University Parkway, University Park, IL 60484, United States of America.

⁶ Golden Bear Research Center, California State University Maritime Academy. 200 Maritime Academy Drive, Vallejo, CA 94590, United States of America.

⁷ University of Notre Dame, Center for Network and Data Science (CNDS), Notre Dame, IN 46556, United States of America.

⁸ University of Delaware, College of Earth, Ocean, and Environment, 305 Robinson Hall, Newark, DE 19716, United States of America.

⁹ Research Institute for Nature and Forest, Gaverstraat 4, Geraardsbergen 9500, Belgium.

¹⁰ School of Biological Sciences, University of Adelaide, SA 5005, Australia.

¹¹ Servicio de Hidrografía Naval (Ministerio de Defensa), Av. Montes de Oca 2124, Buenos Aires, Argentina and Escuela de Ciencias del Mar, Sede Educativa Universitaria, Facultad de la Armada, UNDEF, Av. Antártida Argentina 425, Buenos Aires, Argentina.

¹² SARDI Aquatic Science and Marine Innovation SA, South Australian Research and Development Institute, 2 Hamra Avenue, West Beach SA 5024, Australia.

¹³ Rice University, Department of BioSciences, 6100 Main Street, Houston, Texas, 77005, United States of America.

¹⁴ United States Fish and Wildlife Service, Pacific Islands Fish and Wildlife Office, Honolulu, Hawaii 96850, United States of America.

¹⁵ Oregon Department of Environmental Quality, 700 NE Multnomah Street, Portland OR 97232, United States of America.

¹⁶ Daniel P. Haerther Center for Conservation and Research, John G. Shedd Aquarium, 1200 S. Lake Shore Dr., Chicago, IL 60605, United States of America.

¹⁷ St. John's Island National Marine Laboratory, Tropical Marine Science Institute, National University of Singapore, 18 Kent Ridge Road, Singapore 119227, Singapore.

¹⁸ State of Hawaii Division of Aquatic Resources, Honolulu, Hawaii, United States of America.

¹⁹ Museo Argentino de Ciencias Naturales "Bernardino Rivadavia"-CONICET, Av. Angel Gallardo 470, Buenos Aires C1405DJR, Argentina.

²⁰ Department of Biological Sciences and Environmental Change Initiative, University of Notre Dame, Notre Dame IN 14556 United States of America.

²¹ M. Rose Consulting, 321 3rd Ave. Haines, AK 99827. United States of America.

²² Marine Resources Program, Oregon Department of Fish and Wildlife, 2040 SE Marine Science Drive, Newport, OR 97365, United States of America.

²³ California State Lands Commission, Marine Invasive Species Program, 301 E. Ocean Bld., Ste. 350, Long Beach, CA 90802, United States of America.

²⁴ Instituto para el Estudio de la Biodiversidad de Invertebrados, Facultad de Ciencias Naturales, Universidad Nacional de Salta, Av. Bolivia 5150, A4408FVY Salta, Argentina; and Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Salta, Argentina.

²⁵ Chesapeake Biological Laboratory, University of Maryland Center for Environmental Science, 146 Williams Street, Solomons, MD, United States of America.

²⁶ Fisheries and Oceans Canada, Pacific Biological Station, 3190 Hammond Bay Road, Nanaimo, BC, V9T 6N7, Canada.

²⁷ Department of Biological Sciences, National University of Singapore, 16 Science Drive, Singapore 117558, Republic of Singapore; Lee Kong Chian Natural History Museum, National University of Singapore, 2 Conservatory Drive, Singapore 117377, Republic of Singapore.

*Correspondence to Erin K. Grey, erin.grey@maine.edu

^{\$}These authors contributed equally, and names are listed alphabetically.

Abstract

Spread of nonindigenous organisms by shipping is one of the largest threats to coastal ecosystems. Limited monitoring and understanding of this phenomenon currently hinder development of effective prevention policies. Surveying ports in North America, South America, Europe, Southeast Asia, and Australia we explored environmental DNA community profiles evident of ship-born species spread. We found that community similarities between ports increased with the number of ship voyages, particularly if the ports had similar environments, and when indirect stepping-stone connections were considered. We also found 57 known non-indigenous taxa, some in hitherto unreported locations. We demonstrate the usefulness of eDNA-based tools for global biodiversity surveys, and highlight that shipping homogenizes biodiversity in predictable that could inform policy and management.

Introduction

Harmful nonindigenous species transported through ballast water or biofouling of commercial ships are among the most severe threats to biodiversity and ecosystem services of marine coastal environments (Bax et al., 2003; J. Carlton, 1996; J. T. Carlton & Geller, 1993). Studies of marine traffic patterns and vessel characteristics have helped to develop more robust risks analyses and to identify global hotspots for invasion (Keller et al., 2011; Lodge et al., 2016; Seebens et al., 2013; Tidbury et al., 2016). However, lack of standardized biodiversity data at large spatial, temporal, and taxonomic scales has limited testing of this risk analyses. Cost-effective technologies suitable for quantifying species spread along individual shipping routes are required to accurately categorize shipping voyages and pathways by their risk of spreading nonindigenous species.

DNA analysis of environmental samples is an emerging method that has proven to be accurate and cost-effective for biomonitoring across a wide range of taxa and a variety of habitats (Thomsen & Willerslev, 2015) including commercial ports (Grey et al., 2018). Such analysis of environmental DNA (eDNA) also offers opportunities for quantifying global ecological impacts and evaluating risk models. Here we use a global environmental DNA metabarcoding survey to evaluate four models of ship-borne eukaryotic species spread risk. We show that eDNA analysis of port water samples, together with global data on environmental conditions and shipping traffic, enables route-specific risk analyses of ship-borne species spread (Porter & Hajibabaei, 2018; Saebi, Xu, Kaplan, et al., 2020; Thomsen & Willerslev, 2015; Xu et al., 2016). By implementing a standardized and efficient eDNA port survey method (Grey et al., 2018) we identified biological similarities between 19 ports and related them to four different risk estimates of ship-mediated species transport, while concurrently considering the effects of environmental similarity (Keller et al., 2011) and biogeography (Costello et al., 2017) on global community compositions. Extending this approach to further ports would allow for evaluation and refinement of risk analyses that could better inform policy and management of ship-born nonindigenous species spread.

Our four alternative shipping-risk estimates derived from long-term ship traffic data consider two different aspects of the species spread risk along routes in the global shipping network. The first aspect considered route spread risk as either function of the number of ships traveling between two ports (“traffic risk”) or as a function describing the amount of ballast water (“ballast risk”) transported between ports, based on correlations between ballast discharge volume and probability with ship size, ship type, and voyage time (Saebi, Xu, Grey, et al., 2020; Saebi, Xu, Kaplan, et al., 2020). The second aspect considered the importance of either direct port-to-port voyages (“direct connection risk”) or indirect connections through intermediary ports (“stepping-stone connection risk”). Stepping-stone spread occurs when species are transported from the origin port to an intermediary port, and then picked up by another ship that transports them to the destination port. In this scenario, it is possible to have species spread between two ports even without direct ship traffic between them. While stepping-stone transport of nonindigenous species has been documented at regional scales (Apte et al., 2000), its relative importance at the global scale is unknown.

To evaluate our risk models, we related eDNA-derived biological community profiles, estimates of shipping risk, environmental similarity and shared biogeography in linear models that also

considered all two-way interactions of these factors, and random effects related to each port-to-port shipping connection. We then used backward AIC selection to identify and rank the best model for each risk metric. We expected increased shipping (expressed through shipping-risk estimates) to increase biological similarities between connected ports, while keeping in mind that environmentally similar ports and ports within the same biogeographic region would be biologically similar also in the absence of shipping, thereby providing a backdrop to ship-mediated species spread.

Results

Nineteen of 21 port that we sampled yielded high quality sequence data (Fig. 1a, Table S1). Port locations spanned seven of the 30 marine biogeographic realms defined by Costello et al. (Costello et al., 2017), with the majority situated in the North Pacific, Caribbean and Gulf of Mexico, and Northeast Atlantic realms (Figure 1a). In total we obtained 11,975 eukaryotic Amplicon Sequence Variants (ASVs, (Callahan et al., 2017)) representing both planktonic and benthic species (Fig. S1). Port-to-port Unifrac distances (Lozupone et al., 2011), which are phylogenetic measures of community dissimilarities, ranged from 0.49 to 0.94 (mean: 0.79 ± 0.07) (Figure 1b). Port-to-port values of our four shipping risk estimates were similar each other, but dissimilar to the patterns observed for port-to-port environmental distances (Figure 1b).

Regression modeling results supports our prediction that shipping effect on biodiversity between ports can be detected amongst a backdrop of environmental similarity and, to a lesser extent, biogeographic patterns. All AIC-selected models identified significant effects of environmental similarity (temperature and salinity (Keller et al., 2011)) on biological similarity (Unifrac distance) between ports (Table 1) and connected port pairs with similar environments tended to have smaller Unifrac distances (i.e. more similar biological communities; Fig. S2). Furthermore, ports within the same biogeographical realm (Costello et al., 2017) had the most similar biological communities (Fig. S3) although this effect is only statistically significant in the direct risk models.

Consistent with our predictions, shipping increased the biological similarity between ports. All but one of the AIC-selected models included shipping as an explanatory variable (Table 1). Contrary to our predictions, models considering traffic risk and ballast risk had similar AIC values, suggesting that considering the complexities of ballast discharge beyond the number of voyages added little explanatory power. Models considering stepping-stone connection risk ranked higher than models considering direct connection risks, highlighting the potential need to consider indirect spread through intermediary ports when estimating route-based spread risks.

Most informative to explain species spread along shipping routes was a relatively simple model considering only environmental similarity, stepping-stone traffic connections between ports, and interactions of these variables (Table 1, first row, with lowest AIC). Further exploration of this model revealed that shipping risk increased biological community similarity, particularly on routes whose origin and destination are environmentally similar (Fig. 2). This observation is consistent with the expected pattern of ship-borne species spread, as organisms are unlikely to survive in vastly different environments, regardless of how many times they are introduced.

Overall, our analysis revealed that shipping homogenizes biodiversity among ports when evaluated against a backdrop of high eukaryotic biodiversity shaped primarily by environmental similarity. Across all models, environmental similarity (expressed through temperature and salinity) had the largest effect on UniFrac distances, as we found coefficients associated with shipping risk variables to be much smaller than those associated with environmental similarity. The smallest coefficients were those of the biogeographic variables (Table S2). Our modelling results are robust, as analogous analyses adding two more ports at a shallower sequencing depth yielded similar results, except with biogeographic realms emerging as an additional explanatory variable (Table S3).

A strength of our analyses is the incorporation of taxon-agnostic eDNA profiles from a vast range of taxa from many samples, thereby including information from hitherto undescribed or cryptic species, and overcoming limitations of eDNA studies that attempt low-level taxonomic classification using binomial nomenclature. Nonetheless, species identification with binomial nomenclature is still a critical component of invasion risk management. Accordingly, to further explore overlapping distribution of species among sampled ports, we focused on the subset of ASVs with taxonomic annotations matching nominal species as included in the World Register of Introduced Marine Species (WRiMS) (Ahyong et al., n.d.).

We identified 273 ASVs of which 79 (28.9%), matched equally well to more than one nominal species, indicating possible reference database errors or a lack of taxonomic resolution. The remaining 194 taxonomically unambiguous ASVs represented 11 classes and 57 unique nominal species (Fig. 3), with the solitary tunicate *Ciona savignyi* (41 ASVs) and reef-forming polychaete *Ficopomatus enigmaticus* (42 ASVs) comprising over 40% of all reads for these ASVs. Two of these 57 species were found in non-native areas where they had not been previously reported (Table S4) - *Pseudocalanus elongatus* (Copepoda) at Coos Bay, and *Botrylloides leachii* (Ascidiacea) at Richmond and Oakland. One species, *Oithona davisae*, was found slightly outside of its reported native range in Singapore. All other species were found within their previously reported ranges.

Discussion

Our work contextualizes previous, more spatially limited studies on nonindigenous and invasive taxa in ports, including traditional and eDNA approaches (Borrell et al., 2017; Verna et al., 2016). We provide *in situ* evidence of the relative effects of shipping and the environment on biological homogenization at the global scale. We also reveal the importance of stepping-stone connections for species spread through the global shipping network, which highlights the need to consider the whole shipping network to predict species spread (Saebi, Xu, Kaplan, et al., 2020; Xu et al., 2016). While it is generally accepted that differences among ship types result in unequal risks of species transfer (Davidson et al., 2018; Drake & Lodge, 2004), including ship and voyage parameters influencing ballast discharge in our models had no significant effect on our ability to explain the biological similarities between ports. Although counterintuitive at first glance, it is possible that ship-borne species spread from biofouling or other vectors such as aquaculture obscures the ballast discharge effect in the studied ports (Ojaveer et al., 2018). Evaluation of models with more detailed shipping-risk estimates including biofouling-specific parameters such as vessel speeds and port residence times are needed to determine if this is the case.

Collectively, our results confirm that species spread risks differ among ports in predictable ways, and that relatively simple models yield informative predictions for invasion management. Our most informative model used a shipping risk estimate based on voyage counts between ports, such counts are routinely collected and readily available. The observed significant interaction between shipping and environment suggests that high volume shipping routes connecting environmentally similar ports are the highest risk routes, especially when passing through ports serving as hubs to many other ports, such as, but not limited to, Singapore. Prioritizing these risky routes with targeted management or biomonitoring would, in our opinion, be a more efficient way to reduce global ship-borne species spread.

The lack of a consistently strong signal of biogeographic realms in our results is puzzling, and could be caused by current biogeographic delineations being biased either taxonomically, spatially, or temporally and/or by ports being distinct habitats that are generally quite similar around the world (Bulleri & Aioldi, 2005; Forrest et al., 2013; Lambert & Lambert, 2003; Megina et al., 2016; Piola & Johnston, 2008). It could also be that the limited temporal scope of our study (one sampling event per point) does not reflect β -diversity patterns over longer time-scales (e.g., months to years). Further investigations including ports from more bioregions and other coastal habitats sampled throughout the year are needed to explore such non-exclusive hypotheses.

We demonstrate that eDNA analysis is useful for biomonitoring at a global scale (Campbell et al., 2007). Environmental DNA analysis yields robust β -diversity estimates suitable for hypotheses testing related to the many factors that influence observed biodiversity patterns. For species monitoring, the current drawbacks of eDNA applications are their inaccuracy in assigning taxonomy to some sequences when reference data is deficient, and their susceptibility to biased abundance estimates (Balvočiūtė & Huson, 2017; Kelly et al., 2019). Regardless, we discovered sequence variants highly similar to many known nonindigenous species and derived important insight on their global distributions, which could be further explored with colour-based DNA indicator reactions (Notomi et al., 2015) or enrichment (Wilcox et al., 2018) targeting specific species. Likewise, further exploitation of eDNA techniques may soon enable population genetic inquiries (Sigsgaard et al., 2017) and possibly, abundance estimates (Spear et al., 2021), also in across global ports.

Overall, our work suggests that route-specific management may be a feasible and more effective approach to manage ship-based species invasion risks than then current global “one-size-fits-all” management (Molnar et al., 2008; Wang et al., 2020). While the effect of biogeography and the most appropriate shipping risk metric remain uncertain, shipping clearly homogenizes biodiversity in predictable ways. We show that relatively simply obtained and readily available shipping risk and environmental similarity variables can be used to prioritize route-based management. Future port eDNA surveys including other bioregions and ports experiencing changes in environment and shipping could further refine understanding of ship-born species spread and empower efficient policy development.

Materials and Methods

Water sample collection.

We collected and analyzed five water samples in each of 21 ports across seven biogeographic realms (3, 7, 11, 13, 17, 24, and 26 as defined in Costello *et al.* (Costello et al., 2017) including Australia, China, Oceania, Europe and the Americas (Fig.1). To control for diurnal, tidal, and seasonal variation, all samples were collected during daylight hours at slack high tide after at least a 12-hour period of no rain during the summer or early fall (June-November) (Table S4). Sampling followed a protocol developed specifically for ports and published in Grey *et al.* (Grey et al., 2018). Briefly, each sample consisted of 250 mL surface water, which was stored at 4°C for a maximum of 12 hours after collection. Within that time, water was filtered through a 0.45 µm cellulose-nitrate membrane, immediately immersed in 700 µL Longmire's buffer and stored at room temperature for up to two weeks, and then at -20°C until extraction. If 250 mL could not be filtered through a single membrane, as was the case at some ports with high turbidity and algae, then additional membranes were used and stored in the same microtube. For each site, a collection "blank" sample was taken by filtering 250 mL of bottled or tap water, which was subsequently treated identically to other samples.

eDNA extraction

DNA extractions for the samples and collection blanks (see below) taken at the ports of Adelaide, Chicago, and Singapore were done using phenol-chloroform in a dedicated PCR-free lab at the University of Notre Dame as described by Grey *et al.* (Grey et al., 2018). Total DNA from all other samples and collection blanks were extracted in an eDNA facility at Cornell University (UV light for 8 h d⁻¹, HEPA filtered air under positive pressure, personnel wearing full body suits, face- shields, and breathing masks) using a Blood and Tissue DNA extraction kit (Qiagen, Hilden, DE-NW) and an eDNA optimized protocol (Spens et al., 2017).

Library design, amplification and sequencing

18S rRNA amplicon libraries were prepared using two primers targeting the eukaryotic V4 region (18S_574F and 18S_952R) that were previously evaluated by Hadziavdic et al. (Hadziavdic et al., 2014). For 19 ports (all but Adelaide, Chicago, and Singapore), we followed a single-step amplification approach using a single set of primers including Illumina adapters and a 12-bp barcode unique to each sample (Fig. S4).

Threefold PCR amplifications were carried out for all samples to account for amplification bias. Each reaction of 25 µL total volume contained 2 µL template DNA, 1.5 mM MgCl₂, 1× Colourless GoTaq Flexi Buffer (Promega, Fitchburg, US-WI), 0.25 mM of each dNTP, 0.2 µM of forward and reverse primer and 1.25 units GoTaq Hot Start DNA Polymerase (Promega). After initial denaturation at 95°C for 2:00 min, we performed 40 amplification cycles (95°C for 45 s, 49°C for 1 min, 72°C for 1 min), and a final elongation at 72°C for 5 min. Amplicon triplicates were pooled, visualized on 2% agarose gels, and then purified using Mag-Bind Total Pure NGS beads (Omega Bio-Tek, Norcross, US-GA). DNA concentration for all samples was quantified using PicoGreen reagent (Thermo Fisher Scientific, Waltham, US-MA) and a Spectramax M2 multi-detection microplate reader (Molecular Devices, San Jose, US-CA). Libraries were prepared by combining samples in equimolar ratios, and pair-sequenced (2 x 250) on an Illumina MiSeq platform (Fig. S5). Libraries for the ports of Adelaide, Chicago, and Singapore were previously generated by some of the authors using the protocols described in (Grey et al., 2018).

Blanks and Positive controls.

Alongside port water DNA libraries, we generated libraries for three different blank types (collection, extraction, and amplification controls) and three positive controls: One containing genomic DNA of *Danio rerio*, and two equimolar mixtures of either 4 freshwater fishes (*Notropis topeka*, *Noturus taylori*, *Umbra limi*, *Thoburnia atripinnis*) or 6 marine fishes (*Pseudanthias dispar*, *Ecsenius bicolor*, *Macropharyngodon negrosensis*, *Centropyge bispinosa*, *Salarius fasciatus*, *Amphiprion ocellary*; Fig. S5). Reference sequences for all of the above species were generated using genomic DNA and Sanger sequencing on an ABI 3730xl. Sequencing failed for *Macropharyngodon negrosensis*.

Sequence data processing

Amplicon Sequence Variants (ASVs) were generated from combined 18S amplicon sequence data with Qiime v2 2019-04 (Bolyen et al., 2019), after using BBtools to include sequence data from (Grey et al., 2018) into this work. To do so, reads were initially imported, paired, and filtered using Qiime's import functions with default parameters. Read pairs were then subjected an additional round of adapter trimming using Cutadapt v1.18 (Martin, 2011), denoised using DADA2 (v1.6.0)(Callahan et al., 2016) and an Expected Error value of 9 (Edgar & Flyvbjerg, 2015).(Edgar & Flyvbjerg, 2015) Prior to merging, reads were trimmed to 220 nucleotides. The resulting unique Amplicon Sequence Variants (ASVs) were then filtered using Qiime's taxonomic classifier Vsearch (Rognes et al., 2016) to obtain eukaryotic ASVs as follows: After appending our positive control reference sequences to the Qiime release of the SILVA database v132 (Pruesse et al., 2007), we first ran a separate *in silico* experiment to optimize the assignment settings of Vsearch. Specifically, we optimized the classifier to unequivocally identify the species of the positive controls and to yield maximum correspondence to the analogous taxonomy assignments obtained using the Blast algorithm (Camacho et al., 2009) and the nt/nr NCBI database (Benson et al., 2018). We then used the optimized Vsearch classifier and the SILVA database v132 to identify and retain all Eukaryotic ASVs (*-p-per-identity*= 0.875, *-p-min-consensus*= 0.5, *-p-query-cov*= 0.9).

To account for contamination in the field and lab, we inspected and subtracted ASVs present in the negative controls (see blanks indicated in Fig. S5) from all subsequent datasets. Sequence run read outputs and read counts at each bioinformatics stage are shown in Fig. S6.

Eukaryotic ASV rarefaction and port sampling effort

To account for different read depths across samples, we generated rarefaction curves using Qiime to determine the minimum number of ASVs describing the eukaryotic diversity found among the sampled ports. To do so, we rarefied all samples at depths between 1 to 55,000 sequences in steps of 1,000, repeating each step 4 times. Based on the resulting accumulation curves we initially determined 49,900 ASVs/port sample to be sufficient. Because choosing specific plateauing values is to some extent arbitrary, we also consider a second, lower, value of 37,900, thereby allowing inspecting of two more ports (Fig. S7).

For each of these two rarefied datasets, we then used a bootstrapping approach to determine the minimum number of samples per port that result in stable estimates of Unifrac distances between port pairs. Our bootstrapping algorithm initially divided the full Unifrac distance matrix (including all samples and ports) into unique port-pair matrices. From each of these port matrices, we then generated a series of bootstrapped squared, symmetric matrices of varying dimensions ranging from 1x1 to 7x7 samples/port (n=10,000/dimension). For each bootstrapped matrix a mean distance value was calculated. Finally, the median absolute deviation (MAD for

each port pair and sampling dimension) was calculated and plotted (Fig. S8 and S9). Based on the pattern of MAD decline, we determined that, for both rarefied datasets, 5 samples were sufficient to provide robust estimates of Unifrac distances between port pairs.

Given the above results, we generated two different datasets by randomly selecting five sufficiently covered samples per port and rarefying them to a depth of either 49,900 or 37,900 eukaryotic ASVs. The former dataset consists of 10,279,400 and excludes the ports of Nanaimo and Vancouver while the latter one comprises 8,262,200 eukaryotic ASVs and included all sampled ports.

Estimation of Unifrac distances between port pairs

For each of the two rarefied datasets, we calculated the Unifrac distances between ports using Qiime. To do so, we first obtained reference trees by aligning eukaryote ASV's using MAFFT (Katoh & Standley, 2013) with automatic algorithm and parameter selection. This initial alignment was then post-processed to retain only consenting columns supported by at least 50% of ASV's and allowing only columns with < 10% gaps across the entire alignment. Finally, reference trees were obtained using Fasttree (Price et al., 2010) with the default settings, and unweighted Unifrac distances were calculated among all port pairs.

Modelling of traffic- and ballast-associated spread risks

NIS (non-indigenous species) spread via shipping activity depends on multiple environmental and geographical factors, as well as complex interactions in the global shipping network. For example, many ships do not discharge their entire ballast water in their first port of call. As a result, the remaining ballast water discharged by the ship at the intermediate stops poses the risk of NIS transfer from the original port to the intermediate stops. To account for more such patterns of NIS spread, we followed the approach of (Saebi, Xu, Grey, et al., 2020; Saebi, Xu, Kaplan, et al., 2020) to obtain 4 different traffic- and ballast- associated risks. We used eight years of shipping data from 1997 to 2018 obtained from Lloyd's List Intelligence, an Informa Group Company (LLI, New York, NY, USA). We calculated the risk of spread based either on direct traffic, stepping-stone traffic, direct ballast, or stepping-stone ballast. In the two ballast network models, the spread risk was obtained by first calculating the introduction probability for all ships travelling from port i to port j . For each ship v travelling during time Δt_{ij}^v we applied:

$$P(intro)_{ij}^v = \rho^v (1 - e^{-\lambda D_{ij}^v}) e^{-\mu \Delta t_{ij}^v}$$

Where D_{ij}^v , ρ^v , λ are the amount of ballast water discharged at the destination, the efficacy of ballast water management for the route, and the species introduction potential per volume of discharge.

In the stepping-stone network models, for each port pair, we included an indirect risk of spread by calculating a Jaccard coefficient by dividing the number of ports directly connected to both ports by the total number of direct connections to either port in the pair.

Statistical modelling

A critical first step in understanding the role of shipping traffic in the homogenization of marine biota is the selection of an appropriate set of models covering reasonable hypotheses. Marine species distributions are known to be influenced by temperature, salinity, and biogeography (e.g. (Costello et al., 2017; Keller et al., 2011)). Therefore, we considered the additive model:

$$D_U = \alpha + \beta_1 Env + \beta_2 Eco + \beta_3 Tra + \beta_4 (Env.Eco) + \beta_5 (Env.Tra) + \beta_6 (Eco.Tra) + \mu(PortO)\gamma + \mu(PortD)\gamma + \varepsilon$$

Where D_U is the Unifrac distance between port pairs, the ports of origin (*PortO*) and destination (*PortD*) are random effects, shipping traffic (*Tra*) and environmental similarity (*Env*) are fixed factors, and the biogeographic distance between port pairs (*Eco*) is a binary fixed effect that reflects if the ports of origin and destination are in the same marine realm as defined by Costello *et al.* (Costello *et al.*, 2017) or not. We ran 4 different models using voyage frequency, with or without consideration of “stepping-stone connection risk” estimates, which were calculated as Jaccard indices for each port pair. The environmental similarity between ports was estimated as described elsewhere (Keller *et al.*, 2011). These complex models consider all additive effects and the potential effect of all the two-way interactions. To avoid over- or under-fitting, we then used a backward AIC variable selection approach. Statistical modeling and model selection were done using centered and scaled response variables and the *lme4* package (Bates *et al.*, 2015) with the R programming language (R Core Development Team, 2019).

Marine invasive species identification

Taxonomic assignment of ASVs was performed using the Blast algorithm version 2.9 (Camacho *et al.*, 2009), the April 2019 NCBI nr/nt reference database (Benson *et al.*, 2018) (excluding subject sequences solely identified as “environmental sample”), and a minimum e-value of 10^{-5} . For each query-subject pair, we kept the top 5 high-scoring alignments (>98.5% percent identity) and excluded “ambiguous” assignments where more than one species matched equally well to the query sequence. Unambiguously identified species were then matched to the World Register of Introduced Marine Species, WRiMS (Ahyong *et al.*, n.d.) and the species ports’ presence data was contrasted with the known distributions of these species according to the World Register of Marine Species, WoRMS (WoRMS Editorial Board, 2019) (Table S2). R packages *phyloseq* (McMurdie & Holmes, 2013) and *tidyverse* (Wickham, 2017) were used to import and ASV data and Blast results. We used *taxonomizr* to obtain taxonomic information of the best Blast hits.

References

- Ahyong, S., Costello, M. J., Galil, B. S., Gollasch, S., Hutchings, P., Katsanevakis, S., Lejeusne, C., Marchini, A., Occhipinti, A., Pagad, S., Poore, G. C. B., Rius, M., Robinson, T. B., Sterrer, W., Turon, X., Willan, R. C., & Zhan, A. (n.d.). *World Register of Introduced Marine Species (WRiMS)*. <https://doi.org/10.14284/347>
- Apte, S., Holland, B. S., Godwin, L. S., & Gardner, J. P. A. (2000). Jumping ship: A stepping stone event mediating transfer of a non-indigenous species via a potentially unsuitable

- environment. *Biological Invasions*, 2(1), 75–79.
- <https://doi.org/10.1023/A:1010024818644>
- Balvočiūtė, M., & Huson, D. H. (2017). SILVA, RDP, Greengenes, NCBI and OTT — how do these taxonomies compare? *BMC Genomics*, 18(S2), 114.
- <https://doi.org/10.1186/s12864-017-3501-4>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1). <https://doi.org/10.18637/jss.v067.i01>
- Bax, N., Williamson, A., Agüero, M., Gonzalez, E., & Geeves, W. (2003). Marine invasive alien species: A threat to global biodiversity. *Marine Policy*, 27(4), 313–323.
- [https://doi.org/10.1016/S0308-597X\(03\)00041-1](https://doi.org/10.1016/S0308-597X(03)00041-1)
- Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Ostell, J., Pruitt, K. D., & Sayers, E. W. (2018). GenBank. *Nucleic Acids Research*, 46(D1), D41–D47.
- <https://doi.org/10.1093/nar/gkx1094>
- Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., ... Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*, 37(8), 852–857. <https://doi.org/10.1038/s41587-019-0209-9>
- Borrell, Y. J., Miralles, L., Do Huu, H., Mohammed-Geba, K., & Garcia-Vazquez, E. (2017). DNA in a bottle—Rapid metabarcoding survey for early alerts of invasive species in ports. *PLOS ONE*, 12(9), e0183347. <https://doi.org/10.1371/journal.pone.0183347>

- 429 Bulleri, F., & Airoidi, L. (2005). Artificial marine structures facilitate the spread of a non-
430 indigenous green alga, *Codium fragile* ssp. *Tomentosoides*, in the north Adriatic Sea.
431 *Journal of Applied Ecology*, 42(6), 1063–1072. [https://doi.org/10.1111/j.1365-](https://doi.org/10.1111/j.1365-2664.2005.01096.x)
432 2664.2005.01096.x
- 433 Callahan, B. J., McMurdie, P. J., & Holmes, S. P. (2017). Exact sequence variants should replace
434 operational taxonomic units in marker-gene data analysis. *The ISME Journal*, 11(12),
435 113597. <https://doi.org/10.1038/ismej.2017.119>
- 436 Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P.
437 (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature*
438 *Methods*, 13(7), 581–583. <https://doi.org/10.1038/nmeth.3869>
- 439 Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T.
440 L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, 10(1), 421.
441 <https://doi.org/10.1186/1471-2105-10-421>
- 442 Campbell, M. L., Gould, B., & Hewitt, C. L. (2007). Survey evaluations to assess marine
443 bioinvasions. *Marine Pollution Bulletin*, 55(7–9), 360–378.
444 <https://doi.org/10.1016/j.marpolbul.2007.01.015>
- 445 Carlton, J. (1996). Marine Bioinvasions: The Alteration of Marine Ecosystems by
446 Nonindigenous Species. *Oceanography*, 9(1), 36–43.
447 <https://doi.org/10.5670/oceanog.1996.25>
- 448 Carlton, J. T., & Geller, J. B. (1993). Ecological Roulette: The Global Transport of
449 Nonindigenous Marine Organisms. *Science*, 261(5117), 78–82.
450 <https://doi.org/10.1126/science.261.5117.78>

- Costello, M. J., Tsai, P., Wong, P. S., Cheung, A. K. L., Basher, Z., & Chaudhary, C. (2017). Marine biogeographic realms and species endemicity. *Nature Communications*, 8(1), 1057. <https://doi.org/10.1038/s41467-017-01121-2>
- Davidson, I. C., Scianni, C., Minton, M. S., & Ruiz, G. M. (2018). A history of ship specialization and consequences for marine invasions, management and policy. *Journal of Applied Ecology*, 55(4), 1799–1811. <https://doi.org/10.1111/1365-2664.13114>
- Drake, J. M., & Lodge, D. M. (2004). Global hot spots of biological invasions: Evaluating options for ballast–water management. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271(1539), 575–580. <https://doi.org/10/bn4946>
- Edgar, R. C., & Flyvbjerg, H. (2015). Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics*, 31(21), 3476–3482. <https://doi.org/10.1093/bioinformatics/btv401>
- Forrest, B. M., Fletcher, L. M., Atalah, J., Piola, R. F., & Hopkins, G. A. (2013). Predation limits spread of *Didemnum vexillum* into natural habitats from refuges on anthropogenic structures. *PLoS ONE*, 8(12), 1–12. <https://doi.org/10.1371/journal.pone.0082229>
- Grey, E. K., Bernatchez, L., Cassey, P., Deiner, K., Deveney, M., Howland, K. L., Lacoursière-Roussel, A., Leong, S. C. Y., Li, Y., Olds, B., Pfrender, M. E., Prowse, T. A. A., Renshaw, M. A., & Lodge, D. M. (2018). Effects of sampling effort on biodiversity patterns estimated from environmental DNA metabarcoding surveys. *Scientific Reports*, 8(1), 2–11. <https://doi.org/10.1038/s41598-018-27048-2>

- Hadziavdic, K., Lekang, K., Lanzen, A., Jonassen, I., Thompson, E. M., & Troedsson, C. (2014). Characterization of the 18S rRNA Gene for Designing Universal Eukaryote Specific Primers. *PLoS ONE*, 9(2), e87624. <https://doi.org/10.1371/journal.pone.0087624>
- Katoh, K., & Standley, D. M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*, 30(4), 772–780. <https://doi.org/10.1093/molbev/mst010>
- Keller, R. P., Drake, J. M., Drew, M. B., & Lodge, D. M. (2011). Linking environmental conditions and ship movements to estimate invasive species transport across the global shipping network. *Diversity and Distributions*, 17(1), 93–102. <https://doi.org/10.1111/j.1472-4642.2010.00696.x>
- Kelly, R. P., Shelton, A. O., & Gallego, R. (2019). Understanding PCR Processes to Draw Meaningful Conclusions from Environmental DNA Studies. *Scientific Reports*, 9(1), 12133. <https://doi.org/10.1038/s41598-019-48546-x>
- Lambert, C., & Lambert, G. (2003). Persistence and differential distribution of nonindigenous ascidians in harbors of the Southern California Bight. *Marine Ecology Progress Series*, 259(1917), 145–161. <https://doi.org/10.3354/meps259145>
- Lodge, D. M., Simonin, P. W., Burgiel, S. W., Keller, R. P., Bossenbroek, J. M., Jerde, C. L., Kramer, A. M., Rutherford, E. S., Barnes, M. A., Wittmann, M. E., Chadderton, W. L., Apriesnig, J. L., Beletsky, D., Cooke, R. M., Drake, J. M., Egan, S. P., Finnoff, D. C., Gantz, C. A., Grey, E. K., ... Zhang, H. (2016). Risk Analysis and Bioeconomics of Invasive Species to Inform Policy and Management. *Annual Review of Environment and Resources*, 41(1), 453–488. <https://doi.org/10/gd589r>

- Lozupone, C., Lladser, M. E., Knights, D., Stombaugh, J., & Knight, R. (2011). UniFrac: An effective distance metric for microbial community comparison. *The ISME Journal*, 5(2), 169–172.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, 17(1), 10. <https://doi.org/10.14806/ej.17.1.200>
- McMurdie, P. J., & Holmes, S. (2013). phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE*, 8(4), e61217. <https://doi.org/10.1371/journal.pone.0061217>
- Megina, C., González-Duarte, M. M., & López-González, P. J. (2016). Benthic assemblages, biodiversity and invasiveness in marinas and commercial harbours: An investigation using a bioindicator group. *Biofouling*, 32(4), 465–475. <https://doi.org/10.1080/08927014.2016.1151500>
- Molnar, J. L., Gamboa, R. L., Revenga, C., & Spalding, M. D. (2008). Assessing the global threat of invasive species to marine biodiversity. *Frontiers in Ecology and the Environment*, 6(9), 485–492. <https://doi.org/10.1890/070064>
- Notomi, T., Mori, Y., Tomita, N., & Kanda, H. (2015). Loop-mediated isothermal amplification (LAMP): Principle, features, and future prospects. *Journal of Microbiology*, 53(1), 1–5. <https://doi.org/10.1007/s12275-015-4656-9>
- Ojaveer, H., Galil, B. S., Carlton, J. T., Alleway, H., Gouletquer, P., Lehtiniemi, M., Marchini, A., Miller, W., Occhipinti-Ambrogi, A., Peharda, M., Ruiz, G. M., Williams, S. L., & Zaiko, A. (2018). Historical baselines in marine bioinvasions: Implications for policy and

- management. *PLOS ONE*, 13(8), e0202383.
<https://doi.org/10.1371/journal.pone.0202383>
- Piola, R. F., & Johnston, E. L. (2008). Pollution reduces native diversity and increases invader dominance in marine hard-substrate communities. *Diversity and Distributions*, 14(2), 329–342. <https://doi.org/10.1111/j.1472-4642.2007.00430.x>
- Porter, T. M., & Hajibabaei, M. (2018). Scaling up: A guide to high-throughput genomic approaches for biodiversity analysis. *Molecular Ecology*, 27(2), 313–338.
<https://doi.org/10.1111/mec.14478>
- Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE*, 5(3), e9490.
<https://doi.org/10.1371/journal.pone.0009490>
- Pruesse, E., Quast, C., Knittel, K., Fuchs, B. M., Ludwig, W., Peplies, J., & Glockner, F. O. (2007). SILVA: A comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Research*, 35(21), 7188–7196. <https://doi.org/10.1093/nar/gkm864>
- R Core Development Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ*, 4, e2584. <https://doi.org/10.7717/peerj.2584>
- Saebi, M., Xu, J., Grey, E. K., Lodge, D. M., Corbett, J. J., & Chawla, N. (2020). Higher-order patterns of aquatic species spread through the global shipping network. *PLoS ONE*, 15(7 July), 1–24. <https://doi.org/10.1371/journal.pone.0220353>

- 536 Saeabi, M., Xu, J., Kaplan, L. M., Ribeiro, B., & Chawla, N. V. (2020). Efficient modeling of
537 higher-order dependencies in networks: From algorithm to application for anomaly
538 detection. *EPJ Data Science*, 9(1), 15. <https://doi.org/10.1140/epjds/s13688-020-00233-y>
- 539 Seebens, H., Gastner, M. T., & Blasius, B. (2013). The risk of marine bioinvasion caused by
540 global shipping. *Ecology Letters*, 16(6), 782–790. <https://doi.org/10/mf3>
- 541 Sigsgaard, E. E., Nielsen, I. B., Bach, S. S., Lorenzen, E. D., Robinson, D. P., Knudsen, S. W.,
542 Pedersen, M. W., Jaidah, M. Al, Orlando, L., Willerslev, E., Møller, P. R., Thomsen, P.
543 F., Al Jaidah, M., Orlando, L., Willerslev, E., Moller, P. R., & Thomsen, P. F. (2017).
544 Population characteristics of a large whale shark aggregation inferred from seawater
545 environmental DNA. *NATURE ECOLOGY & EVOLUTION*, 1(1), 0004.
546 <https://doi.org/10.1038/s41559-016-0004>
- 547 Spear, M. J., Embke, H. S., Krysan, P. J., & Vander Zanden, M. J. (2021). Application of eDNA
548 as a tool for assessing fish population abundance. *Environmental DNA*, 3(1), 83–91.
549 <https://doi.org/10.1002/edn3.94>
- 550 Spens, J., Evans, A. R., Halfmaerten, D., Knudsen, S. W., Sengupta, M. E., Mak, S. S. T. T.,
551 Sigsgaard, E. E., & Hellström, M. (2017). Comparison of capture and storage methods
552 for aqueous microbial <scp>eDNA</scp> using an optimized extraction protocol:
553 Advantage of enclosed filter. *Methods in Ecology and Evolution*, 8(5), 635–645.
554 <https://doi.org/10.1111/2041-210X.12683>
- 555 Thomsen, P. F., & Willerslev, E. (2015). Environmental DNA – An emerging tool in
556 conservation for monitoring past and present biodiversity. *Biological Conservation*, 183,
557 4–18. <https://doi.org/10.1016/j.biocon.2014.11.019>

- 558 Tidbury, H. J., Taylor, N. G. H., Copp, G. H., Garnacho, E., & Stebbing, P. D. (2016). Predicting
559 and mapping the risk of introduction of marine non-indigenous species into Great Britain
560 and Ireland. *Biological Invasions*, 18(11), 3277–3292. [https://doi.org/10.1007/s10530-](https://doi.org/10.1007/s10530-016-1219-x)
561 016-1219-x
- 562 Verna, D. E., Harris, B. P., Holzer, K. K., & Minton, M. S. (2016). Ballast-borne marine invasive
563 species: Exploring the risk to coastal alaska, USA. *Management of Biological Invasions*,
564 7(2), 199–211. <https://doi.org/10.3391/mbi.2016.7.2.08>
- 565 Wang, Z., Nong, D., Countryman, A. M., Corbett, J. J., & Warziniack, T. (2020). Potential
566 impacts of ballast water regulations on international trade, shipping patterns, and the
567 global economy: An integrated transportation and economic modeling assessment.
568 *Journal of Environmental Management*, 275(May), 110892.
569 <https://doi.org/10.1016/j.jenvman.2020.110892>
- 570 Wickham, H. (2017). *tidyverse: Easily Install and Load the “Tidyverse”* (R package version
571 1.2.1).
- 572 Wilcox, T. M., Zarn, K. E., Piggott, M. P., Young, M. K., McKelvey, K. S., & Schwartz, M. K.
573 (2018). Capture enrichment of aquatic environmental DNA: A first proof of concept.
574 *Molecular Ecology Resources*, 18(6), 1392–1401. [https://doi.org/10.1111/1755-](https://doi.org/10.1111/1755-0998.12928)
575 0998.12928
- 576 WoRMS Editorial Board. (2019). *World Register of Marine Species*. World Register of Marine
577 Species. <https://doi.org/10.14284/170>
- 578 Xu, J., Wickramaratne, T. L., & Chawla, N. V. (2016). Representing higher-order dependencies
579 in networks. *Science Advances*, 2(5), e1600028. <https://doi.org/10.1126/sciadv.1600028>

Acknowledgments

For helpful discussions and/or logistic collaboration we thank Kimberly Howland, Anna Lacoursiere, Yiyuan Li, and Aibin Zhan, Demetrio Boltovskoy, Pedro Brandão, Juul Lorre, Dias F. Pinto, Kristy Deiner, Vanessa Hodes, Jill Remeysen, Lindsay Schaffner, Evangelina Schwindt, Peter Avis, and Giles Hooker.

Funding: This work was funded by NSF Coastal SEES Award 1748389 awarded to Lodge, Grey, Corbett, and Chawla.

Author contributions:

Conceptualization: EKG, MP, DML
 Methodology: PC, JA, EKG, DML, MS
 Investigation: PC, JA, KA, EKG, DML, MS
 Visualization: JA, KA, PC, EKG
 Supervision: JA, DML, NC
 Writing—original draft: PC, JA, EKG, DML
 Writing—review & editing: all authors

Competing interests: All authors declare that they have no competing interests.

Data availability: Further details on data and analyses are included in the Supplementary Material. All data, code, and materials used in the analyses will be available electronically via Zenodo [link TBD]. Sequences will be available through the NCBI Sequence Read Archive [Project # and link TBD]

Figures and Tables

Fig. 1. Sampling extent and variable summary. (A) Sampled ports, colored by biogeographic realms (Costello et al., 2017), included Adelaide (AU-ADL), Antwerp (BE-ANR), Baltimore (US-BAL), Coos Bay (US-COB), Ghent (BE-GNE), Haines (US-HNS), Honolulu (US-HNL), Houston (US-HOU), Long Beach (US-LGB), Miami (US-MIA), New Orleans (US-MSY), Oakland (US-OAK), Portland (US-PDX), Puerto Madryn (AR-PMY), Richmond (US-RCH), Rotterdam (NL-RTM), Singapore (SG-SIN), Wilmington (US-ILG), and Zeebrugge (BE-ZEE) **(B)** Summaries of network-like relationships between ports expressing biological similarity, environmental similarity, and different shipping-risk estimates. To allow for direct comparison, we inverted biological and environmental dissimilarities by subtracting each dissimilarity value from either the theoretical (1 for Unifrac distance) or observed (4.36 for environmental dissimilarity) maximum. Line color and width are scaled according to each metric, where thick black lines represent the highest values and thin light-grey lines the lowest values.

Fig. 2. Observed Relationship between biological similarity, environmental similarity, and stepping-stone connection risk. Biological similarity between ports is best explained by the interaction of stepping-stone vessel traffic and the environmental similarity between the ports of origin and destination (Table 1). Medium = Mean; Low = Mean - 1SD; High = Mean + 1SD. Bars represent the $\pm$ 95% CI.

Fig. 3. Evaluation of marine nonindigenous species. Count of ASVs and number of unambiguous nominal species (99% identity) matching to the WRiMS database of marine invasive species colored by taxonomic Class.

Table 1. Model summaries. Selected linear models including environmental similarity (ENV), traffic risk (SHP), and biogeographic realms (BGR). Ports of origin (ORG) and destination (DEST) were modelled as random factors. AIC: Akaike Information Criteria.

SHIP RISK METRIC	AIC-SELECTED MODEL	AIC
Stepping-stone traffic risk	UNIFRAC = ENV+ SHP + ENV*SHP + (ORG) + (DEST)	- 479.34
Stepping-stone ballast risk	UNIFRAC = ENV + SHP + ENV*SHP+ (ORG) + (DEST)	- 478.78
Direct ballast risk	UNIFRAC = ENV + BGR+ SHP + BGR*ENV+BGR*SHP + (ORG) + (DEST)	- 478.51
Direct traffic risk	UNIFRAC = ENV + BGR + ENV*BGR+ (ORG) + (DEST)	-476.9