PHILOSOPHICAL TRANSACTIONS
OF THE ROYAL SOCIETY B
BIOLOGICAL SCIENCES

The reconstruction of invasion histories with genomic data in light of differing levels of anthropogenic transport

Journal:	*Philosophical Transactions B*
Manuscript ID	RSTB-2021-0023.R2
Article Type:	Research
Date Submitted by the Author:	n/a
Complete List of Authors:	Hudson, Jamie; University of Southampton, Ocean & Earth Science Bourne, Steven; University of Southampton, Ocean & Earth Science Seebens, Hanno; Biodiversitat und Klima Forschungszentrum Chapman, Mark; University of Southampton, Biological Sciences and Centre for Underutilised Crops Rius, Marc; University of Southampton, Ocean & Earth Science; CEAB, CSIC; University of Johannesburg
Issue Code (this should have already been entered and appear below the blue box, but please contact the Editorial Office if it is not present):	RANGES
Subject:	Ecology < BIOLOGY, Evolution < BIOLOGY, Genomics < BIOLOGY
Keywords:	Biological invasions, non-indigenous species, population connectivity, population genomics, genetic diversity, invasion routes

Author-supplied statements

Relevant information will appear here if provided.

Ethics

Does your article include research that required ethical approval or permits?:
This article does not present research with ethical considerations

Statement (if applicable):
CUST_IF_YES_ETHICS :No data available.

Data

It is a condition of publication that data, code and materials supporting your paper are made publicly available. Does your paper present new data?:
Yes

Statement (if applicable):
The data for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB46018 (https://www.ebi.ac.uk/ena/browser/view/PRJEB46018).

Accompanying scripts are available at https://doi.org/10.5281/zenodo.5046438

Conflict of interest

I/We declare we have no competing interests

Statement (if applicable):
CUST_STATE_CONFLICT :No data available.

The reconstruction of invasion histories with genomic data in light of differing levels of anthropogenic transport

Hudson J[1,*], Bourne SD[1], Seebens H[2], Chapman MA[3] , Rius M[1,4,5]

[1] School of Ocean and Earth Science, University of Southampton, National Oceanography Centre, European Way, Southampton, SO14 3ZH, United Kingdom.

[2] Senckenberg Biodiversity and Climate Research Centre, Senckenberganlage 25, 60325 Frankfurt, Germany.

[3] Biological Sciences, University of Southampton, Life Sciences Building 85, Highfield Campus, Southampton, SO17 1BJ, United Kingdom.

[4] Centre for Advanced Studies of Blanes (CEAB), CSIC, Accés a la Cala Sant Francesc 14, 17300 Blanes, Spain.

[5] Centre for Ecological Genomics and Wildlife Conservation, Department of Zoology, University of Johannesburg, Auckland Park, 2006, South Africa.

* Corresponding author: J.Hudson@soton.ac.uk

Abstract

Unravelling the history of range shifts is key for understanding past, current, and future species distributions. Anthropogenic transport of species alters natural dispersal patterns and directly affects population connectivity. Studies have suggested that high levels of anthropogenic transport homogenise patterns of genetic differentiation and blur colonisation pathways. However, empirical evidence of these effects remains elusive. We compared two range-shifting species *(Microcosmus squamiger* and *Ciona robusta)* to examine how anthropogenic transport affects our ability to reconstruct colonisation pathways using genomic data. We first investigated shipping networks from 1750 onwards, cross-referencing these with regions where the species have records to infer how each species has potentially been affected by different levels of anthropogenic transport. We then genotyped thousands of single nucleotide polymorphisms from 280 *M. squamiger* and 190 *C. robusta* individuals collected across their extensive species' ranges and reconstructed colonisation pathways. Differing levels of anthropogenic transport did not preclude the elucidation of population structure, though specific inferences of colonisation pathways were difficult to discern in some of the considered scenario sets. We conclude that genomic data in combination with information of underlying introduction drivers provide key insights into the historic spread of range-shifting species.

Keywords: Biological invasions, genetic diversity, invasion routes, non-indigenous species, population connectivity, population genomics

Introduction

The ever-increasing rate of globalisation of trade is intensifying the anthropogenic transport of species [1,2], leading to introductions of many species to regions away from their native ranges. As non-indigenous species (NIS) cause major impacts on ecological communities around the world, understanding the underlying mechanisms facilitating NIS' spread is fundamental for biodiversity conservation and management [3]. One way of studying NIS' spread is through identifying genetic patterns across different spatial scales [4–6]. Such studies have suggested that anthropogenic transport geographically reshuffles genotypes [7–10], and/or causes regional or global genetic homogenisation [11–14]. Because unravelling colonisation pathways is key for understanding NIS' spread [15] and for planning mitigation strategies [16], understanding how anthropogenic transport of species may dampen our ability to reconstruct invasion routes is fundamental.

Anthropogenic transport of species, by definition, increases population connectivity across species' ranges. The genetic composition of colonising populations can be affected by numerous different processes and scenarios. For example, genetic bottlenecks and founder effects in recent colonisations may lead to population structure across the species range [17]. Conversely, genetic homogenisation among populations may be expected if local adaptation within introduced ranges is weak, or if high levels of gene flow (through frequent introductions) persists [18]. Furthermore, the timing and magnitude of anthropogenic transport may affect population structure. For example, "recolonisations" of introduced genotypes back to the native range may result in reduced genetic structure throughout the species range. A similar pattern of homogenisation could also occur due to variation in effective population size, N_e. Previous work has found a positive correlation between N_e in the introduced range and time-since-invasion [19]. Large N_e would prevent genetic drift, slowing divergence between populations, even in the absence of ongoing gene flow via continuing introductions. Conversely, an ancient invader might be expected to develop strong population structure throughout its range if local adaptation of introduced populations has evolved and / or if reduced gene flow has led to genetic drift. Another mechanism enhancing population structure may be through multiple introductions of genotypes from genetically divergent source populations, increasing the propensity for intraspecific genetic admixture [20]. Changes in transportation routes of species, in the absence of natural population connectivity, can also lead to a subset of introduced populations becoming disconnected from other populations, resulting in a rapid change in allele frequencies [21] or a reduction in genetic diversity due to drift [22].

High throughput sequencing (HTS) enables scientists to obtain a substantial genomic coverage and capture patterns of genome-wide variation [23] and this HTS offers significantly higher resolution of

fine-scale gene flow than studies analysing a few loci [24]. HTS has been used to reconstruct invasion histories [16,25,26], inferring the presence of multiple and sequential introductions [27,28], as well as revealing the presence of genetic admixture that may have fitness consequences on colonising populations [25,29]. In addition, studies of neutral loci have analysed population genomic patterns of NIS in both introduced and native ranges [30,31], identified secondary contacts [32] and detected genetic bottlenecks [33]. However, no study using HTS has to date tested how anthropogenic transport of species affects our ability to infer colonisation pathways of NIS [26].

Here we used a comparative approach to compare the effects of different levels of anthropogenic transport on the reconstruction of introduction pathways using HTS data. For this, we studied two biologically similar sessile marine NIS that have widespread distributions but have presumably been affected by different levels of anthropogenic transport. Both species belong to the class Ascidiacea (phylum Chordata) and have limited natural dispersal capabilities with the duration of motile early life-history stages being only a few days [34,35]. Ascidiacea species are amongst the most prolific groups of invasive species on the planet [36], often causing negative economic impacts on important human activities [37]. We first analysed historical inter-regional shipping to detect patterns of anthropogenic transport among the regions where the study species were present. We then sequenced samples collected from across the ranges of the study species to explore range-wide connectivity patterns. Finally, we inferred the most likely colonisation pathways using Bayesian methods and determined the putative impact of anthropogenic transport on our ability to reconstruct invasion routes.

Materials and Methods

Study species and field sample collection

We studied two ascidian species, *Microcosmus squamiger* (Michaelsen, 1927) and *Ciona robusta* (Hoshino & Tokioka, 1967) for which species records suggest differing levels of anthropogenic transport (Table S1). Briefly, *M. squamiger* is native to Australia [38,39] and was first reported outside of its native range in the mid-20th century in the Mediterranean Sea and South Africa [39,40]. *Ciona robusta* is putatively native to the Northwest Pacific [41] and has been recorded in the Mediterranean Sea from the 19th century [42], followed by records in South Africa [43], northeast Pacific [44], Australia [45,46], New Zealand [47] and Hong Kong [48] throughout the 20th century, and the south coast of England [49] since the early 21st century. Both species' population genetics have previously been studied using a relatively small number of genetic markers [31,41,42], and thus no study to date has reconstructed the invasion routes of these NIS using genome-wide tools.

We sampled individuals from both the native and introduced ranges of the study species (Fig. 1, Tables S3/S4). Sampling sites were chosen to maximize distributional coverage and to include geographic areas that were not covered in previous genetic studies [31,42]. Specifically, we made a concentrated effort to sample regions where little sampling was conducted in previous studies (e.g. [42]), such as Australasia or South Africa (Fig. 1). At each site, we collected 20-30 individuals by hand from ropes and marina buoys / pontoons, or from artificial rocky substrata using SCUBA. We enforced a spacing of a few tens of centimetres among sampled individuals to minimize the collection of closely related individuals. We then dissected a piece of the mantle (muscle tissue) from each individual and immediately fixed the tissue samples in >99% ethanol. Samples were then transported to the laboratory where they were stored at -80°C until DNA extraction.

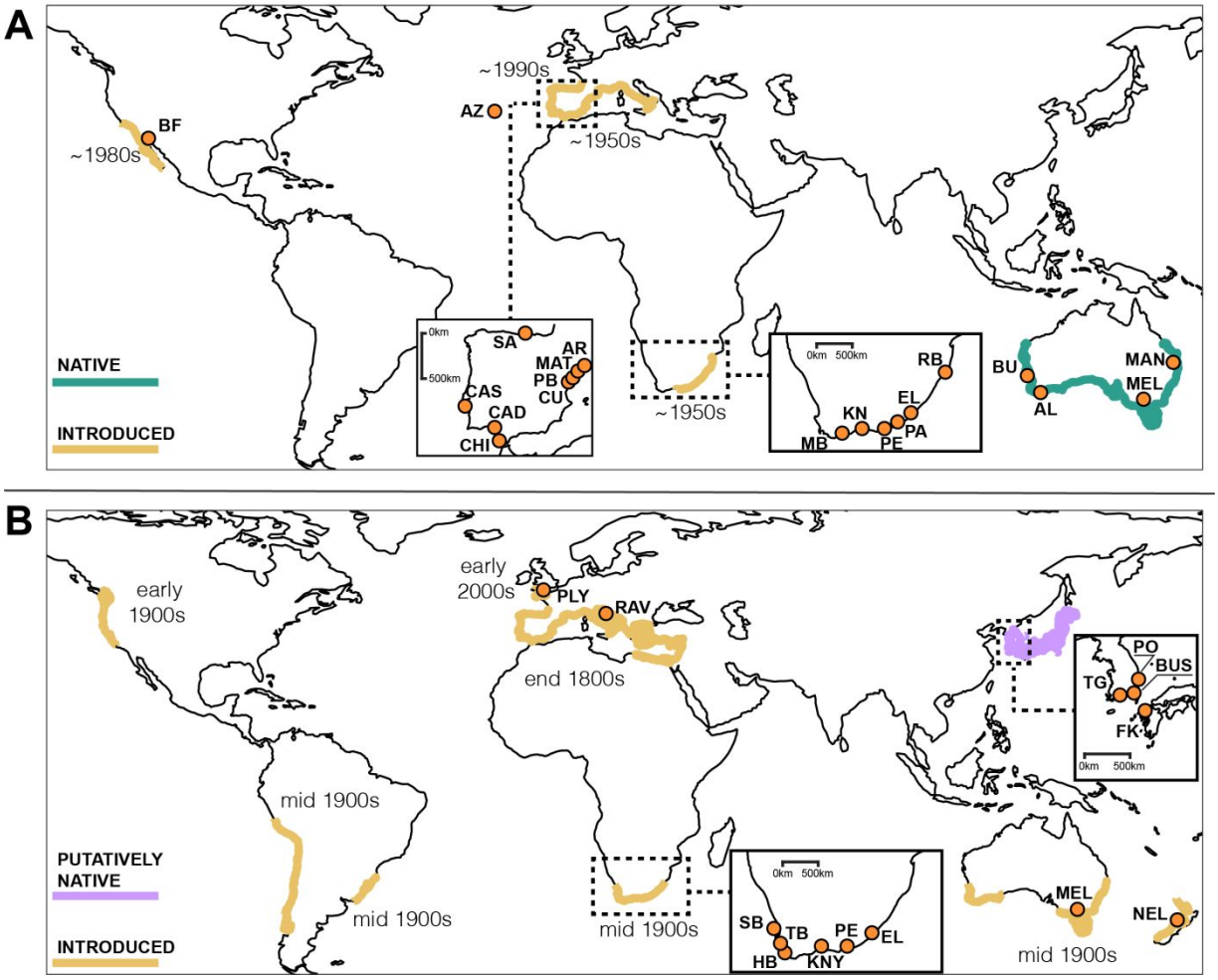

Fig. 1. Sampling sites and ranges of **(A)** *Microcosmus squamiger* (boxes show enlarged Iberian and South African sites) and **(B)** *Ciona robusta* (boxes shows enlarged South African, Iberian, and

northwest Pacific sites). Coloured areas show status of their ranges and years next to each region when each species was first recorded as introduced. Orange dots indicate sampling sites (see Table S2 for full details of these sites). Site abbreviations are as follows: A) BU = Bunbury, AL = Albany, MEL = Melbourne, BF = Bahía Falsa, AZ = Azores, SA = Santander, CA = Cascais, CAD = Cádiz, CHI = Chiclana, CU = Cubelles, PB = Port Barcelona, MAT = Mataró, AR = Arenys de Mar, MB = Mossel Bay, KNY = Knysna, PE = Port Elizabeth, PA = Port Alfred, EL = East London, RB = Richards Bay; B) FK = Fukuoka, BUS = Busan, PO = Pohang, TG = Tongyeong, NEL = Nelson, MEL = Melbourne, KNY = Knysna, PE = Port Elizabeth, EL = East London, SB = Saldanha Bay, TB = Table Bay, HB = Hout Bay, RAV = Ravenna, PLY = Plymouth.

Historical shipping data

We obtained historical shipping data from global regions across the study species' ranges. These data came from two independent datasets that spanned two sequential time periods: The Climatological Database for the World's Oceans (CLIWOC, 1750 - 1850, http://webs.ucm.es/info/cliwoc/) and the International Comprehensive Ocean-Atmosphere Data Set (ICOADS, 1865 - 2014, http://icoads.noaa.gov/). The CLIWOC data set draws from digitised British, Dutch, French and Spanish ships' logbooks, with a focus on ships sailing in the Atlantic and the Western Indo-Pacific. The ICOADS data derives from various sources worldwide (http://icoads.noaa.gov/). Both data sets were originally constructed to reconstruct historical ocean and atmospheric conditions, and not shipping dynamics. As a result, they do not include all shipping activity, but give a good representation of general shipping dynamics at that time.

Both datasets provided ship location dates and geographic details during their travel, enabling the reconstruction of individual ship trajectories and shipping intensities. The CLIWOC dataset provided additional information about anchor points, which can be interpreted as port calls of that ship. The ICOADS dataset did not provide information about anchor points, and it was thus necessary to infer port calls from ship trajectories. To determine actual port calls, we calculated the shortest distance of each ship coordinate to a list of 1620 ports obtained from the World Port Index 26th Edition (https://opendata-esri-de.opendata.arcgis.com/datasets/world-port-index). We only considered large ports (i.e., not recreational marinas which are mostly recent developments) that we could assume persisted over the past 250 years. Geographic details of ship locations were only provided once a day and no records were available when a ship stayed in the actual port. We therefore considered a port call if a ship sailed within 10 km distance from a port. We checked individual ship trajectories and used different distances to test the sensitivity of the reconstruction of shipping routes. In total we obtained 7,238 individual ship movements from the CLIWOC data set and

210,423 ship movements from ICOADS. For both data sets the temporal and spatial coverage was not always consistent and thus data were only analysed on coarse temporal (50-year intervals) and spatial (regional) scales. To visualize historical shipping data, we created chord diagrams using the R package 'circlize' [50], to show the number of direct ship travels between regions where the study species occur for each 50-year period between 1750 - 2000.

DNA extraction and genotyping

Total genomic DNA was extracted from all tissue samples using the ReliaPrep™ gDNA Tissue Miniprep System (Promega, Madison, Wisconsin, USA). DNA was sent for sequencing at Cornell Genomics Diversity Facility (Cornell University, Ithaca, NY, USA). The restriction enzymes *Pst*I, *Eco*T221 and *Ape*KI were trialled to identify the one that created suitable libraries (fragments <500 bp, presence of non-repetitive DNA), and thus *Pst*I was used for *M. squamiger*, and *Eco*T221 for *C. robusta*. Genotyping was performed using the genotyping-by-sequencing protocol [51], and took place on an Illumina HiSeq 2500, using single-end 100 bp reads.

Data processing

We processed data from each species independently using the same bioinformatics pipeline. Briefly, sequence data were first passed through FastQC [52] to investigate read quality. After successfully passing quality checks, the GBS reads were assembled de novo using ipyrad v.0.7.30 [53] using parameters recommended for single-end GBS data (http://ipyrad.readthedocs.io/). We then conducted read assembly, SNP filtering, and loci selection (see full description in the Supplementary Materials).

Genomic summary statistics, population structure and differentiation

Within-population indices of genetic diversity (expected heterozygosity [H_E], observed heterozygosity [H_O], and the inbreeding coefficient [F_{IS}]) were calculated using the "diversity v.1.9.90" package [54] within R [55]. To provide a graphical representation of between-site genetic differentiation, and to test for population structure within our datasets, we used two genetic clustering methods. Firstly, we used the "adegenet v.2.1.3" package [56] in R to perform a Principal Component Analysis (PCA) using the function *dudi.pca*. Secondly, we used the software ADMIXTURE v.1.3 [57] to group individuals into one of *K* putative clusters, using a maximum-likelihood estimation. For both species, the number of tested clusters ranged from 1 to *n*, where *n* = the

number of sites individuals were sampled from. The R package "hierfstat v.0.5-7" [58] was used to calculate genomic differentiation, as inferred through pairwise-population values of F_{ST}.

Combining genomic indices and shipping data

For each period of shipping data available we assessed the correlation between the number of shipping events (hereafter referred to as shipping intensity) and genomic differentiation. We grouped our study sites into regions corresponding to the spatial scale of our shipping data analysis, and calculated mean F_{ST} values of sites amongst these regions, before performing a Spearman's rank correlation between shipping intensity and F_{ST} in R using the package "ggpubr v.0.4.0" [59].

Reconstructing colonisation pathways

We used DIYABC Random Forest v.1.0 [60], which utilises Approximate Bayesian Computation to evaluate different evolutionary scenarios, to infer colonisation pathways. For all scenarios, training sets were generated using 2,000 simulations per model. Note that supervised machine learning methods such as random forest (RF) use all simulations to learn the mapping of data to models, and subsequently a smaller training set is required compared to ABC methods [60]. Current knowledge of the study species' global distribution and historical species records (Fig. 1) were used to inform model construction. In addition, the results of the PCAs and population differentiation were used to pool genomically similar geographical sites and guide the building of the models (for a detailed description of the model sets see the Supplementary Materials). We identified the most likely scenario of each set using the 'RF analysis' module of DIYABC-RF (see full details in the Supplementary Materials).

Results

Historical inter-regional shipping patterns

We found a clear pattern of increasing complexity and magnitude of global shipping over recent time (Fig. 2). Indeed, the total number of shipping events was small initially but showed a sharp increase from the beginning of the 20th century, with the period between the years 1750 and 1800 containing 155 events, 1801-1850 containing 88 events, 1851-1900 containing 68 events, 1901-1950 containing 1010 events, and 1951-2000 containing 1624 events. Among the regions of interest for this study, most intense shipping was consistently recorded in the northeast Atlantic, representing around 40% of shipping between 1750 and 2000 (Fig. 2F). South Africa was also a major shipping donor/recipient particularly before 1850 and was involved in minor shipping trade with the

northwest Pacific prior to the 1800s. Shipping within the *M. squamiger* native range (i.e., Australia), whilst being present at low intensity in the 18th century, intensified from the mid 19th century onwards. Mediterranean shipping steadily increased from 1750, representing 20% of shipping traffic from the 1950s onwards. These shipping data indicate that from the 1750s, shipping was prevalent among regions across the range of *C. robusta* (Australia, Mediterranean Sea, North-West Pacific, North-East Pacific, and South Africa; Figs. 2B and 2C). Thus, the combined used of historical shipping data and the species records (Table S1) suggested a longer history of anthropogenic transport in *C. robusta* compared to *M. squamiger*.

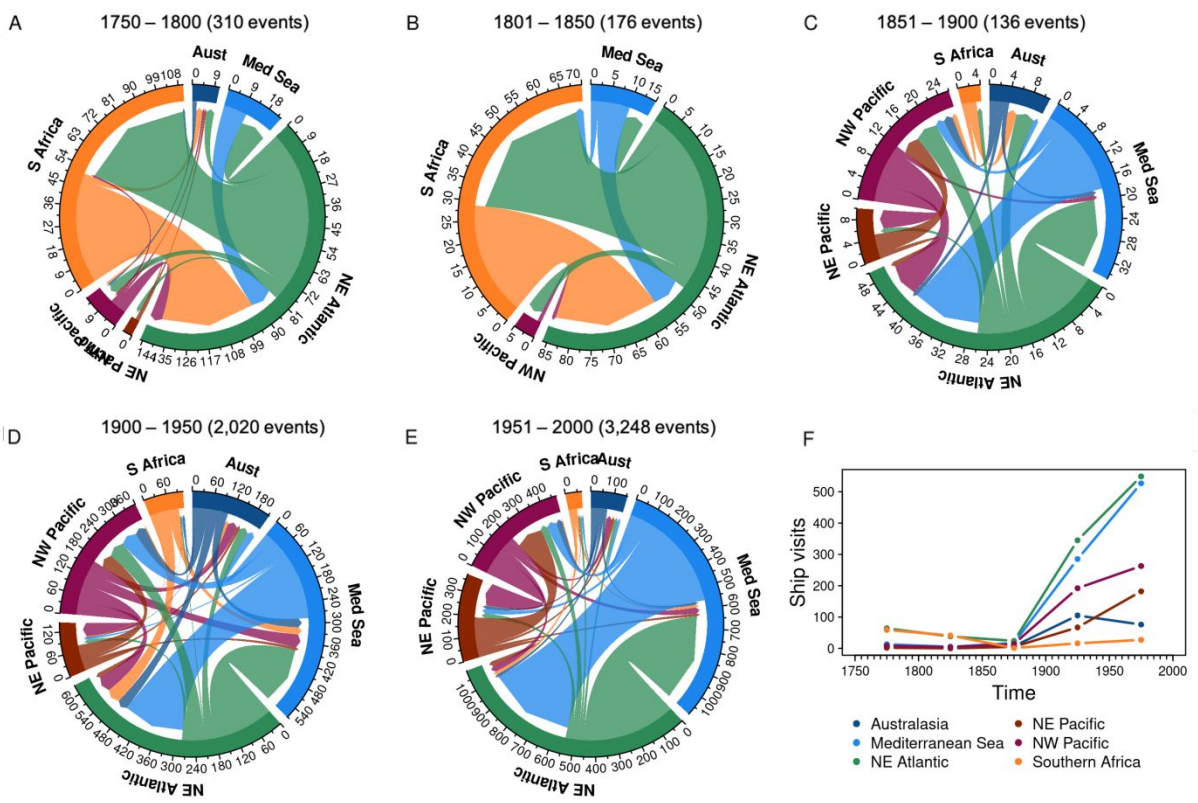

Fig. 2. Temporal development of the global shipping network from 1750 - 2000, considering the regions where the study species can be found. (A - E) Chord diagrams showing the number of ship travel events between marine regions over ~50-year intervals. The arrows at the end of the flows represent incoming travel to that region. Each region is colour assigned and represented by a circular segment proportional to the respective shipping intensity. (F) Temporal development of the total number of ship visits at ports for each region.

Genotyping of neutral single nucleotide polymorphisms

We genotyped 365 *M. squamiger* and 214 *C. robusta* individuals from across their species ranges. Of these, 280 *M. squamiger* and 190 *C. robusta* successfully passed our sequencing QC (Tables S3 and S4). Following our filtering protocol, we retained 2115 and 3227 SNPs for *M. squamiger* and *C.*

robusta respectively. We then identified putatively non-neutral SNPs using BAYESCAN and pcadapt and removed those that were presented in either method, leaving a dataset of 1994 SNPs and 3139 SNPs for *M. squamiger* and *C. robusta* respectively.

Genomic summary statistics

For *M. squamiger*, expected and observed heterozygosities were consistent across the range of (native range mean H_E = 0.111 and mean H_O = 0.064; introduced range mean H_E = 0.114 and mean H_O = 0.065; Fig. S1 and Table S5) and the mean number of private alleles per site was greater in the native range (mean = 31.3 private alleles per site) than the introduced range (mean = 6.2 private alleles per site). For *C. robusta*, H_E was higher in the putative native ranges (mean H_E = 0.191) than in the introduced range (mean H_E = 0.148; Fig. S2 and Table S6), however for Ravenna (the Mediterranean site) in the introduced range, H_E was higher than all other sites (0.241). The number of private alleles across the range showed the opposite pattern to H_E, with sites within the native range having fewer private alleles (mean = 10.0) than the introduced range (mean = 35.6). All sites, for both species, exhibited positive F_{IS} values (for values of genetic diversity indices for each site see Tables S5 and S6).

Population structure and differentiation

Genomic differentiation was high among native sites of *M. squamiger*, but low within the introduced range (Fig. 3A). The optimum number of clusters identified by ADMIXTURE was *K* = 2, with one cluster containing three native sites (BU, AL, and MAN) and the second cluster containing the native site MEL and all introduced sites. Due to the heuristic nature of ADMIXTURE, we also plotted *K* = 3 – 5, which recovered further potential structure within the introduced range, separating South African sites and the Eastern Pacific from those in the Atlantic and Mediterranean and blurring the initially inferred structure (Fig. S3A). The PCA identified four main clusters, each corresponding to one of the four Australian sites (AL, BU, MEL and MAN) with individuals from all introduced sites clustered with those individuals from Melbourne. The first axis of the PCA recovered groupings matching the ADMIXTURE result at *K* = 2 (Fig. S4A). This close relationship between MEL and the introduced sites was reinforced by the pairwise genetic differentiation F_{ST} values (Fig. S5A).

A greater separation of clusters was identified in *C. robusta* than *M. squamiger*, as seen in the both the higher optimal value of *K* in the ADMIXTURE analysis (Fig. 3B), and PCA (Fig. S4B). The ADMIXTURE analysis recovered three distinct genomic clusters, with one group containing the native range and NEL, a second group containing the European sites (RAV and PLY) and the western South Africa sites (SB, TB, and HB), with the third group containing the eastern South Africa sites (KNY, PE,

and EL) (Fig. 3B). Interestingly Australian site MEL contained individuals composed of all three clusters (Fig. 3B). Unlike *M. squamiger*, increasing values of *K* did not result in blurring of inferred structure (Fig. S3B). The PCA recovered a similar picture, however it instead recovered four clusters (Fig. S4B). Individuals collected from the northwest Pacific (i.e., the native range) once again clustered together, individuals from South Africa were found in two clusters, corresponding to either the east (KNY, PE, and EL) or west coast (SB, TB, and HB), and both the site within the English Channel (PLY) and the site within the Mediterranean Sea (RAV), clustered closely to the western South African cluster. However, the PCA recovered the site from New Zealand as a unique cluster (NEL), and genotypes from the Australian site (MEL) encompassed all clusters except the native range (Fig. S4B). Considering population differentiation (see Fig. S5B), northwest Pacific sites were genetically similar (average F_{ST} = 0.01), but strongly differentiated from other sites (average F_{ST} = 0.13).

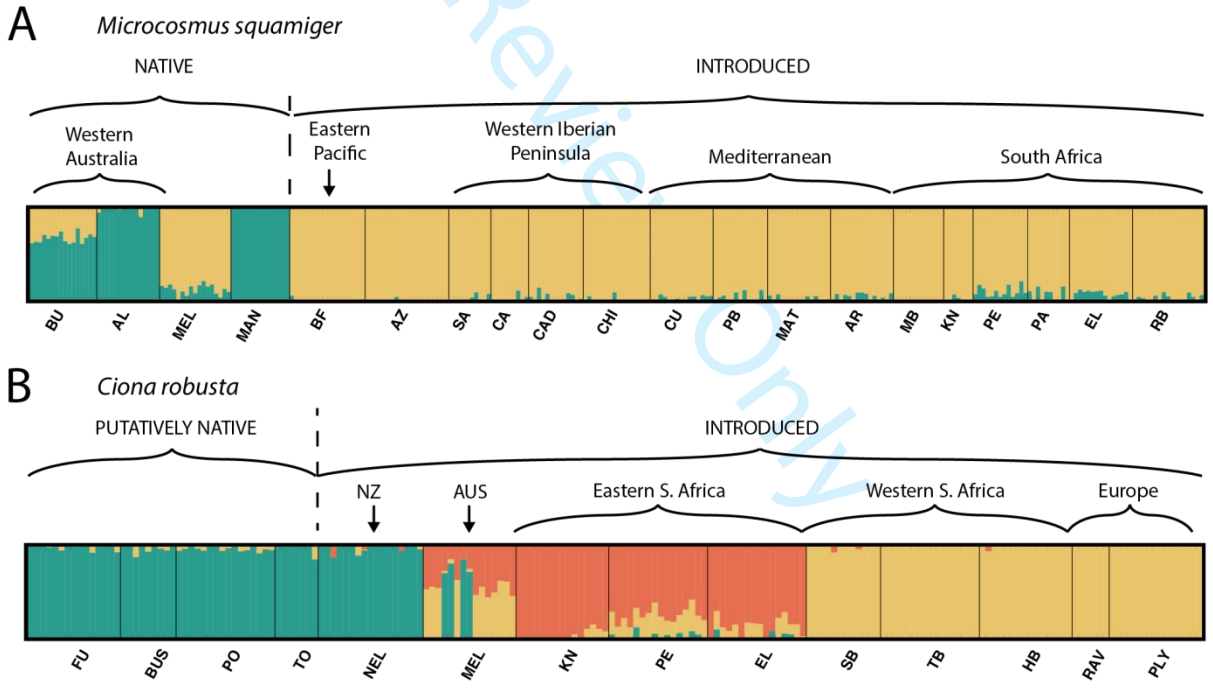

Fig. 3. Genomic clusters within (A) *Microcosmus squamiger* (optimal number of clusters = 2) and (B) *Ciona robusta* (optimal number of clusters = 3) as inferred by ADMIXTURE v.1.3. Population abbreviations are given in Fig. 1.

Regarding the correlation between historical shipping and genomic differentiation, values of F_{ST} were slightly negatively correlated with average shipping intensity between 1750-2000, though not significantly (Fig. S6), for both species.

Inference of colonisation routes

Preliminary analyses showed that 2,000 simulated datasets per model were suitable for inferring model choice by computing error metrics from both the entire training set and a subset. Likewise, evaluations for each DIYABC-RF run showed that the number of RF trees produced for each model set was sufficient (i.e., error rates stabilised with increasing number of trees). We thus tested a comprehensive variety of models for each species (Figs. S7 and S8).

For *M. squamiger,* DIYABC-RF was able to confidently identify a split between western and eastern Australian sites (Fig. S7.1 and Table S7, models 17 and 18), followed by admixture between the western site AL and an eastern site MEL. This admixture originated the other western site BU (Fig. 4A model 2; mean posterior probability = 0.601 - note the mean prior and mean posterior error rates for the chosen model were high [0.476 and 0.400 respectively, Table S8]). Strong evidence of admixture between MEL and BU (Fig. 4A.3) was also found. Though the final colonisation to the introduced range was unresolved, a consensus of potentially suitable models included a split between SA and MED (see the mean number of votes and standard deviations per model in Table S7 and posterior probabilities and error rates in Table S8), with these two populations being a bridgehead for the BF and ATL populations respectively.

Regarding *C. robusta,* the DIYABC-RF found that NWP split initially from an unsampled population, with MED (Fig. 4B.1) and ESA (Fig. 4B.2) also being sourced from unsampled populations. WSA was found to be sourced from the MED group (Fig. 4B.3), AUS was identified to be a result of admixture between the east and west coasts of South Africa (Fig. 4B.4), and NZ was recovered to be a secondary introduction from NWP (Fig. 4B.5). The most recent grouping, EC, was identified as sourced from MED (Fig. 4B). See Table S9 for posterior probabilities and error rates, and Table S10 for the mean number of votes and standard deviations per model.

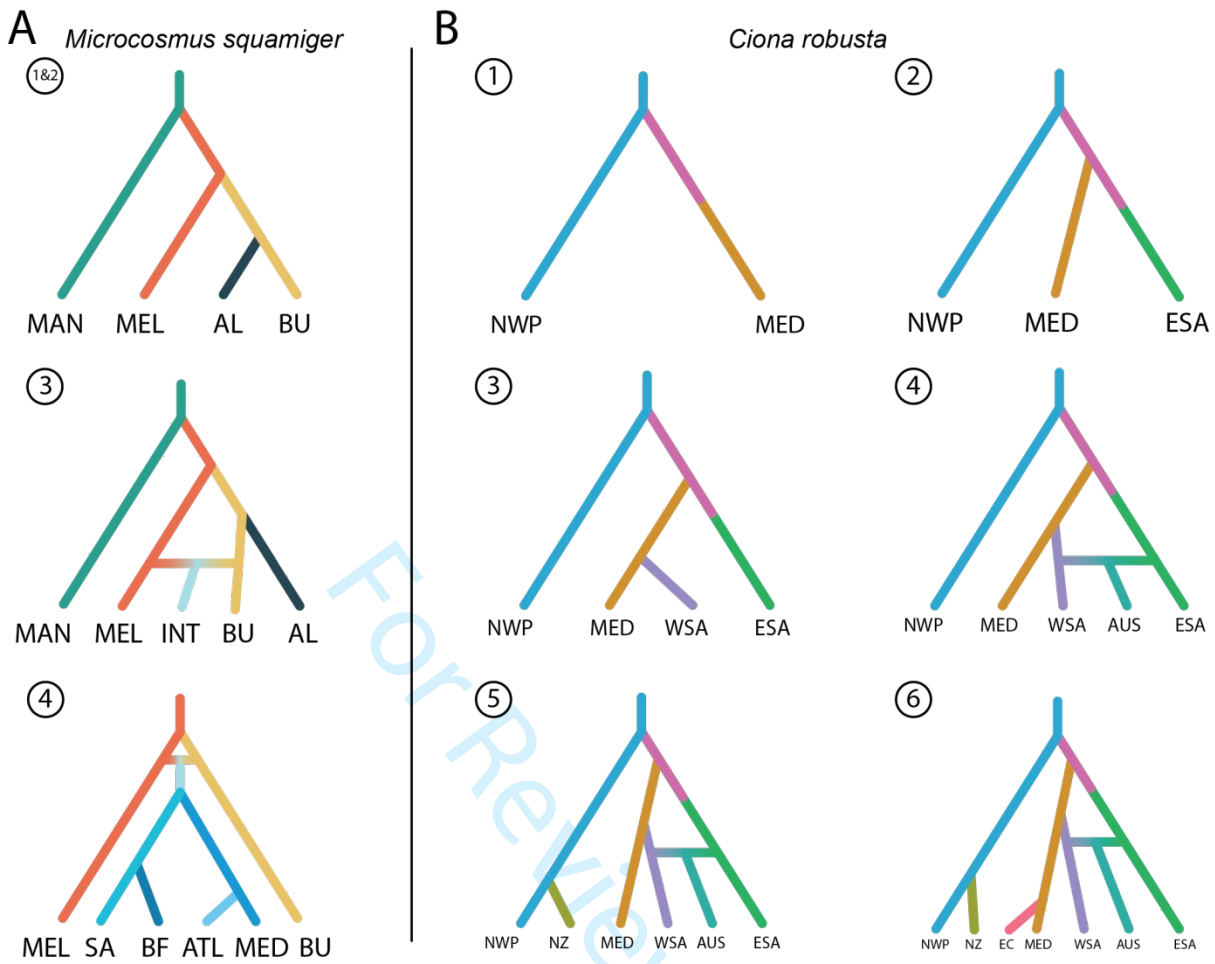

Fig. 4. Models of invasion routes identified as most likely using approximate Bayesian computation implemented in DIYABC-RF v.1.0 for the study species (*Microcosmus squamiger* and *Ciona robusta*). Progression through tree is backwards in time, so labelled terminal branches are present day. Numbers in circles indicate scenario set, as in Methods (but see Figs. S7 and S8 for visual representation of all models), with each set increasing in complexity. Labels are as follows: A) MAN = Manly, MEL = Melbourne, AL = Albany, BU = Bunbury, INT = all introduced sites, SA = grouped sites from South Africa (Mossel Bay, Knysna, Port Elizabeth, Port Alfred, East London, and Richards Bay), MED = grouped sites from the Mediterranean (Cubelles, Port Barcelona, Mataró, and Arenys de Mar), ATL = grouped sites from the Atlantic (Azores, Santander, Cascais, Cádiz, and Chiclana), and BF = Bahía Falsa; B) NWP = grouped sites from North-West Pacific (Fukuoka, Busan, Pohang, Tongyeong), MED = Ravenna, ESA = grouped sites from Eastern South Africa (Knysna, Port Elizabeth, East London), WSA = grouped sites from Western South Africa (Saldanha Bay, Table Bay, Hout Bay), AUS = Melbourne, NZ = Nelson, EC = Plymouth.

Discussion

Our results provide evidence that putatively varying levels of anthropogenic transport do not preclude our ability to recover patterns of population structure across species ranges that have undergone complex introduction histories. Our ability to unambiguously identify specific

colonisation pathways of NIS remained limited in some cases, though general features of population structure and invasion routes could be identified. Such signals would not be expected if anthropogenic transport consistently eroded the geographic distribution of genotypes across the species' ranges, and effectively homogenised genomic divergence across the species' ranges. Additionally, we showed that differing histories of anthropogenic transport can provide a suitable explanation for observed genomic differences between native and introduced ranges.

Historical patterns of shipping intensity and connectivity

Our temporal analysis of historical shipping networks showed a clear pattern of increasing complexity and intensity of shipping with time [61]. In addition, the results supported our initial assumption that the two studied species have been affected by different levels of anthropogenic transport. Both shipping data and species records suggested that *C. robusta* was subject to anthropogenic transport earlier than *M. squamiger*, providing more opportunities to be redistributed from its original range and a greater time to differentiate from the source populations. For example, the putative native range of *C. robusta* (the northwest Pacific), was an important region for shipping throughout all the time periods studied, becoming a sizeable contributor to shipping from the mid-19th century (Fig. 2). The observed patterns of historical shipping suggest that *C. robusta* was initially transported during a time with lower shipping intensity and connectivity amongst distant regions. Regarding the native range of *M. squamiger* (i.e. Australia), it appeared in our initial time period (1750 - 1800) but was not present again until 1854 - 1900 (Fig. 2), suggesting that in the early 19th century Australasia may not have been an important source or recipient of global shipping from or to the other study regions. By the time *M. squamiger* was being transported, shipping patterns were complex and thus one source population could spread quickly throughout the introduced range, possibly through a stepping-stone dispersal, which could explain the inferred high levels of population connectivity across much of the introduced range of *M. squamiger*. This may have subsequently led to an increased likelihood of repeated introductions [22]. Such a situation could have occurred when the Suez Canal opened in 1869, rapidly reducing the importance of South Africa as a transportation hub, as seen in the reduction of shipping intensity in the region between 1851-1900 (Fig. 2C). Footprints of founder effects, such as the reduction in genetic diversity observed in some *C. robusta* populations, could be explained by introductions of few individuals into the introduced range (as in [62]).

A fundamental assumption made in interpreting our results was the close association between NIS' introductions and shipping intensity. It would be unreasonable to assume every ship included in our data set of shipping intensity would lead to an introduction, and our data cannot resolve the

magnitude of ongoing, recurrent introductions. However, a higher intensity of ship traffic increases the likelihood of individuals being transported along a certain route and makes it therefore more likely that individuals colonise new sites [61,63]. Indeed, ascidians have been identified in ~6% the ballast waters of ships sailing from the western Pacific to eastern Pacific coastlines [64], and over time such a percentage will likely lead to high levels of propagule pressure. Our analyses including shipping dynamics were limited by the availability of historical data. Shipping data were obtained from two independent data sets spanning two different time periods (i.e., before and after 1850), which differ in their spatial coverage and comprehensiveness. While the early data set (CLIWIC) has a stronger focus on the Atlantic region, the latter (ICOADS) provides a more comprehensive global coverage, which explains the abrupt changes of shipping dynamics among time periods. Despite these caveats, the datasets gave a good representation of the overall development of the shipping network [65].

Genomic patterns within native ranges

Species' native ranges are expected to show a clear population structure [66] as the accumulation of mutations [67], genetic drift [68], and / or development of geographic barriers [69] increase population differentiation and the frequency of private-alleles over time [70]. Our analyses recovered separate genomic clusters within the native range of *M. squamiger*, with the number of clusters ranging between 2-3 depending on the analysis. Additionally, the number of private alleles present within sites within the native range was approximately six times greater than those found in the introduced range (Table S5). In contrast, the putative native range of *C. robusta* showed limited population structure. This could be due to high levels of gene flow within the native range [71], high effective population size [72], or insufficient sampling. Indeed, it is known that *C. robusta* can be found further east along the coast of Japan than the sampling conducted here [42]. Despite this, the sites from the NWP in the present study portray a similar picture to that from Bouchemousse et al. [42], that is, the NWP sites are similar to each other, though are reasonably genomically diverse too. Further sampling across the NWP would provide clarification as to whether the genomic homogeneity present in the native range is due to the genomic homogenisation of populations within the native range through anthropogenic transport [73].

Genomic patterns within introduced ranges

Whilst genetic bottlenecks are often mentioned in the literature of biological invasions [74], it is becoming increasingly appreciated that introduced populations do not regularly undergo a significant reduction in genomic diversity [20]. Multiple introductions [75], high gene flow [76], and / or genetic admixture [77] often overcome any reduction in genetic diversity associated with

bottlenecks. We did not find evidence of a reduction in genomic diversity between the native and introduced range of *M. squamiger*, possibly either due to increased propagule pressure owing to intense anthropogenic transport, or genetic admixture between native sites (see results of the DIYABC analyses). The extensive introduced range of *M. squamiger* was highly homogenous, both in terms of population structure and genomic diversity patterns. Global genomic homogeneity within the introduced range could be the result of the introduction of genotypes from a single source population from the native range [78] or high levels of population connectivity within the introduced range due to intense anthropogenic transport [79] promoting stepping-stone dispersal. In contrast, we found evidence of population structure within the introduced range of *C. robusta*. Population structure within introduced ranges has been found in other ascidians [80], and can be attributed to multiple independent introduction events [62]. The observed population structure in *C. robusta* was present at differing spatial scales. For example, geographically distant regions such as Europe and western South Africa were genomically homogenous, supporting previous results found by Zhan et al. [79]. Historical records of *C. robusta* identify the ascidian as being present along the western coast of South Africa since the 1950s [81]. Whether the observed similarity in genomic makeup between these two regions is a result of ongoing anthropogenic transport, or the result of high N_e supressing the effects of genetic drift, remains unknown, though we recovered a drop in H_E in western South Africa sites compared to those found in eastern South Africa (Fig. S2). It is unclear whether the limited natural dispersal potential of ascidians, coupled with their affinity to inhabit artificial environments (i.e. marinas, ports, harbours), has an effect on N_e. However previous work on the congener *C. savignyi* showed a large effective population size as inferred in San Francisco Bay [82]. On a regional scale, we found clear structure along the South African coastline. Strong regional differentiation in South Africa could be due to demographic processes or introductions from multiple independent source populations. Regarding genomic diversity, we observed a decrease from the putative native range to western South Africa populations, which may provide evidence for demographic processes contributing to genomic differentiation. As *C. robusta* has been present along the western coast of South Africa since at least the 1950s [81], it is unlikely that the low levels of genomic diversity is the result of a recent introduction. Finally, the DIYABC RF analyses identified different introduction sources for both the eastern and western coasts of South Africa. Taken together, *C. robusta* displays population structure in South Africa likely due to existing marine biogeographic provinces, demographic processes and / or independent introductions.

Reconstructing invasion routes

The species with the shorter history of anthropogenic transport, *M. squamiger,* showed limited confidence in the reconstruction of invasion routes, with only one scenario set having a prior error rate of <45% (Table S8). In accordance with previous work using microsatellite and DNA sequence data [31], we found strong evidence that *M. squamiger* is native to Australia. Furthermore, we found evidence that the genomic homogeneity of the introduced range of *M. squamiger* resulted from a single-source introduction from an unsampled site comprising individuals from either Melbourne or from admixture between Melbourne and Bunbury sites, with subsequent stepping-stone dispersal. Such a signature of high homogeneity across the introduced range has been observed in other marine organisms. For example, genetic homogeneity has been identified within the introduced range of the invasive lionfish (*Pterois volitans*) with the conclusion that gene flow can quickly erode previous signals of genetic divergence [13].

Whilst we found evidence of population structure between introduced populations of *M. squamiger* and the native range outside Melbourne, we could not discount the possibility of introduced alleles re-entering the native range. This is suggested by the discord between the clustering and the DIYABC-RF analyses, with the former indicating that Melbourne was the sole source. Further evidence for Melbourne being the source of all the introduced populations came from the fact that the lowest number of private alleles across all native sites were found in Melbourne. We know from historical data that Melbourne and Bunbury opened as ports from the 1850s onwards [83,84], and just over a century later *M. squamiger* individuals were found in California [85] and the Mediterranean Sea [39]. This was reinforced by our shipping history data, which showed that Australia only started increasing its shipping activity from the 1850s, and indeed only became a significant global contributor after the 1900s. This further indicates that over the 20th century, *M. squamiger* colonized distant regions around the globe, demonstrating how rapidly anthropogenic transport can facilitate the establishment and spread of NIS.

Poorly documented species records from the literature posed a challenge for guiding our analyses of the colonisation history of *C. robusta*. Whilst the prior error rates of the scenario sets were lower (i.e. higher confidence in model choice) than those for *M. squamiger*, they still ranged between 14 – 36% (Table S9). This in part may be the reason why our DIYABC RF analyses were unable to identify the source of the Mediterranean and eastern South Africa sites (both coming from an unsampled population). In turn, we were able to find evidence for multiple introductions and potential admixture (e.g., Fig. 4) events promoting the expansion of the species. A previous genetic study of *C. robusta* also sampled a large part of the species range [42] and found that in line with previous work that the northwest Pacific is the putative native range; although an introduced status of *C. robusta* in

the northwest Pacific could not be disproved based on their evidence, consistent with the results presented here.

Until recently, little has been known regarding the effects of anthropogenic transport on genetic patterns across species ranges, but a growing number of studies are unravelling invasion routes despite an intensification of anthropogenic transport in recent decades / centuries (Fig. 2). For example, Manni *et al.* [14] were able to accurately define the source populations of the Japanese Asian tiger mosquito (*Aedes albopictus*) despite exhibiting chaotic propagule dispersion associated with trans-continental anthropogenic transport. Similarly, Lesieur *et al.* [86] found that despite a complex invasion history and long-distance dispersal owing to anthropogenic transport of species, the invasion pathway of the Western conifer seed bug (*Leptoglossus occidentalis*) could still be tracked. Our genomic results showed that invasion routes of NIS with high historical anthropogenic transport can be studied with similar confidence as NIS with both shorter residence times in the introduced range, and lower levels of anthropogenic transport. We therefore conclude that although considering anthropogenic transport remains important, it does not preclude inference with genomic data, providing that sampling is of sufficient geographic breadth.

With anthropogenic transport of species being a major factor dictating the distribution of many range shifting species [4,7,87], it is essential to consider artificial connectivity pathways amongst populations to plan both management and mitigation actions [88]. Specifically, knowledge of source/s of prolific range-shifting populations may aid planning management actions such as vector/NIS eradication. Our study results unravelled how anthropogenic transport changes the geographic distribution of genetic lineages, as well as provided applied knowledge particularly relevant to stakeholders with an interest in mitigating the effects of NIS.

Acknowledgments

We thank Chul-Woong Oh, Hitoshi Sawada and Dustin Marshall (and their research groups) for logistical support during sample collection in Korea, Japan and Australia respectively. We thank Marco Abbiati, Aibin Zhan, Federica Costantini for providing additional samples from Europe. We thank Joshua Murray, Caitriona Hanley and Jessica Adams for assistance with DNA extraction. HS acknowledges support through the 2017-2018 Belmont Forum and BiodivERsA joint call for research proposals, under the BiodivScen ERA-Net COFUND programme, and with the funding organisations German Federal Ministry of Education and Research (BMBF; grant 01LC1807A). SDB was supported by the Natural Environment Research Council (grant number N3/L002531/1).This work was funded

484 by the Adventure in Research Grant AAIR15 from the University of Southampton awarded to MR and
485 MAC.

References

1. Hulme PE. 2021 Unwelcome exchange: International trade as a direct and indirect driver of biological invasions worldwide. *One Earth* **4**, 666–679. (doi:10.1016/j.oneear.2021.04.015)
2. Hulme PE. 2017 Climate change and biological invasions: evidence, expectations, and response options: Alien species and climate change in Great Britain. *Biol. Rev.* **92**, 1297–1313. (doi:10.1111/brv.12282)
3. Katsanevakis S, Zenetos A, Belchior C, Cardoso AC. 2013 Invading European Seas: Assessing pathways of introduction of marine aliens. *Ocean Coast. Manag.* **76**, 64–74. (doi:10.1016/j.ocecoaman.2013.02.024)
4. David AA, Loveday BR. 2018 The role of cryptic dispersal in shaping connectivity patterns of marine populations in a changing world. *J. Mar. Biolog. Assoc. U.K.* **98**, 647–655.
5. Sherman CDH, Lotterhos KE, Richardson MF, Tepolt CK, Rollins LA, Palumbi SR, Miller AD. 2016 What are we missing about marine invasions? Filling in the gaps with evolutionary genomics. *Mar. Biol.* **163**, 198.
6. Viard F, David P, Darling JA. 2016 Marine invasions enter the genomic era: three lessons from the past, and the way forward. *Curr. Zool.* **62**, 629–642. (doi:10.1093/cz/zow053)
7. Blakeslee A, McKenzie C, Darling J, Byers J, Pringle J, Roman J. 2010 A hitchhiker's guide to the Maritimes: anthropogenic transport facilitates long-distance dispersal of an invasive marine crab to Newfoundland. *Divers. Distrib.* **16**, 879–891.
8. Crooks J, Suarez A. 2006 Hyperconnectivity, invasive species, and the breakdown of barriers to dispersal. In *Connectivity Conservation* (eds K Crooks, M Sanjayan), pp. 451–478. Cambridge University Press.
9. David AA. 2018 Reconsidering Panmixia: The Erosion of Phylogeographic Barriers Due to Anthropogenic Transport and the Incorporation of Biophysical Models as a Solution. *Front. Mar. Sci.* **5**. (doi:10.3389/fmars.2018.00280)
10. Lippens C *et al.* 2017 Genetic structure and invasion history of the house mouse (*Mus musculus domesticus*) in Senegal, West Africa: a legacy of colonial and contemporary times. *Heredity* **119**, 64–75.
11. Christie MR, Knowles LL. 2015 Habitat corridors facilitate genetic resilience irrespective of species dispersal abilities or population sizes. *Evol. Appl.* **8**, 454–463.
12. David AA, Matthee CA, Loveday BR, Simon CA. 2016 Predicting the Dispersal Potential of an Invasive Polychaete Pest along a Complex Coastal Biome. *Integr. Comp. Biol.* **56**, 600–610.
13. Pérez-Portela R *et al.* 2018 Genetic homogeneity of the invasive lionfish across the Northwestern Atlantic and the Gulf of Mexico based on Single Nucleotide Polymorphisms. *Sci. Rep.* **8**, 5062.
14. Manni M *et al.* 2017 Genetic evidence for a worldwide chaotic dispersion pattern of the arbovirus vector, *Aedes albopictus*. *PLOS Negl. Trop. Dis.* **11**, 0005332.

15. 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**, 189–202.

16. Estoup A, Guillemaud T. 2010 Reconstructing routes of invasion using genetic data: why, how and so what? *Mol. Ecol.* **19**, 4113–4130. (doi:10.1111/j.1365-294X.2010.04773.x)

17. Gortat T, Rutkowski R, Gryczynska A, Kozakiewicz A, Kozakiewicz M. 2017 The spatial genetic structure of the yellow-necked mouse in an urban environment – a recent invader vs. a closely related permanent inhabitant. *Urban Ecosyst.* **20**, 581–594. (doi:10.1007/s11252-016-0620-7)

18. Lenormand T. 2002 Gene flow and the limits to natural selection. *Trends Ecol. Evol.* **17**, 183–189. (doi:10.1016/S0169-5347(02)02497-7)

19. Braasch J, Barker BS, Dlugosch KM. 2019 Expansion history and environmental suitability shape effective population size in a plant invasion. *Mol. Ecol.* **28**, 2546–2558. (doi:10.1111/mec.15104)

20. Rius M, Turon X, Bernardi G, Volckaert FAM, Viard F. 2015 Marine invasion genetics: from spatio-temporal patterns to evolutionary outcomes. *Biol. Invasions* **17**, 869–885. (doi:10.1007/s10530-014-0792-0)

21. Marchini GL, Arredondo TM, Cruzan MB. 2018 Selective differentiation during the colonization and establishment of a newly invasive species. *J. Evol. Biol.* **31**, 1689–1703. (doi:10.1111/jeb.13369)

22. Dlugosch KM, Parker IM. 2008 Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Mol. Ecol.* **17**, 431–449. (doi:10.1111/j.1365-294X.2007.03538.x)

23. Rašić G, Filipović I, Weeks AR, Hoffmann AA. 2014 Genome-wide SNPs lead to strong signals of geographic structure and relatedness patterns in the major arbovirus vector, *Aedes aegypti*. *BMC Genom.* **15**, 275.

24. Low GW, Chattopadhyay B, Garg KM, Irestedt M, Ericson PGP, Yap G, Tang Q, Wu S, Rheindt FE. 2018 Urban landscape genomics identifies fine-scale gene flow patterns in an avian invasive. *Heredity* **120**, 138–153.

25. van Boheemen LA, Lombaert E, Nurkowski KA, Gauffre B, Rieseberg LH, Hodgins KA. 2017 Multiple introductions, admixture and bridgehead invasion characterize the introduction history of *Ambrosia artemisiifolia* in Europe and Australia. *Mol. Ecol.* **26**, 5421–5434.

26. Cristescu ME. 2015 Genetic reconstructions of invasion history. *Mol. Ecol.* **24**, 2212–2225. (doi:10.1111/mec.13117)

27. Blumenfeld AJ *et al.* 2021 Bridgehead effect and multiple introductions shape the global invasion history of a termite. *Commun. Biol.* **4**, 1–12. (doi:10.1038/s42003-021-01725-x)

28. Puckett EE, Magnussen E, Khlyap LA, Strand TM, Lundkvist Å, Munshi-South J. 2020 Genomic analyses reveal three independent introductions of the invasive brown rat (*Rattus norvegicus*) to the Faroe Islands. *Heredity* **124**, 15–27. (doi:10.1038/s41437-019-0255-6)

29. Fraimout A *et al.* 2017 Deciphering the Routes of invasion of *Drosophila suzukii* by Means of ABC Random Forest. *Mol. Biol. Evol.* **34**, 980–996.

30. Hudson J, Johannesson K, McQuaid CD, Rius M. 2020 Secondary contacts and genetic admixture shape colonization by an amphiatlantic epibenthic invertebrate. *Evol. Appl.* **13**, 600–612. (doi:10.1111/eva.12893)

31. Rius M, Turon X, Ordóñez V, Pascual M. 2012 Tracking invasion histories in the sea: facing complex scenarios using multilocus data. *PLOS ONE* **7**, e35815. (doi:10.1371/journal.pone.0035815)

32. Jeffery NW *et al.* 2018 Genomewide evidence of environmentally mediated secondary contact of European green crab (*Carcinus maenas*) lineages in eastern North America. *Evol. Appl.* **11**, 869–882. (doi:10.1111/eva.12601)

33. Cammen KM, Schultz TF, Bowen WD, Hammill MO, Puryear WB, Runstadler J, Wenzel FW, Wood SA, Kinnison M. 2018 Genomic signatures of population bottleneck and recovery in Northwest Atlantic pinnipeds. *Ecol. Evol.* **8**, 6599–6614. (doi:10.1002/ece3.4143)

34. Chiba S, Sasaki A, Nakayama A, Takamura K, Satoh N. 2004 Development of *Ciona intestinalis* juveniles (through 2nd ascidian stage). *Zool. Sci.* **21**, 285–298. (doi:10.2108/zsj.21.285)

35. Rius M, Turon X, Marshall DJ. 2009 Non-lethal effects of an invasive species in the marine environment: the importance of early life-history stages. *Oecologia* **159**, 873–882. (doi:10.1007/s00442-008-1256-y)

36. Zhan A, Briski E, Bock DG, Ghabooli S, MacIsaac HJ. 2015 Ascidians as models for studying invasion success. *Mar. Biol.* **162**, 2449–2470. (doi:10.1007/s00227-015-2734-5)

37. Palanisamy SK, Thomas OP, McCormack G. 2018 Bio-invasive ascidians in Ireland: a threat for the shellfish industry but also a source of high added value products. *Bioengineered* **9**, 55–60.

38. Rius M, Pascual M, Turon X. 2008 Phylogeography of the widespread marine invader *Microcosmus squamiger* (Ascidiacea) reveals high genetic diversity of introduced populations and non-independent colonizations. *Divers. Distrib.* **14**, 818–828. (doi:10.1111/j.1472-4642.2008.00485.x)

39. Turon X, Nishikawa T, Rius M. 2007 Spread of *Microcosmus squamiger* (Ascdiacea- Pyuridae) in the Mediterranean Sea and adjacent waters. *J. Exp. Mar. Biol. Ecol.* **342**, 185–188.

40. Ramos-Espla AA, Izquierdo A, Çinar ME. 2013 *Microcosmus exasperatus* (Ascidiacea: Pyuridae), current distribution in the Mediterranean Sea. *Mar. Biodivers. Rec.* **6**, e89.

41. Lin Y, Chen Y, Yi C, Fong JJ, Kim W, Rius M, Zhan A. 2017 Genetic signatures of natural selection in a model invasive ascidian. *Sci. Rep.* **7**, 44080.

42. Bouchemousse S, Bishop JDD, Viard F. 2016 Contrasting global genetic patterns in two biologically similar, widespread and invasive *Ciona* species (Tunicata, Ascidiacea). *Sci. Rep.* **6**, 24875. (doi:10.1038/srep24875)

43. Millar RH. 1955 On a collection of ascidians from South Africa. *Proc. Zool. Soc. Lond.* **125**, 169–221.

44. Rodholm AK. 1932 *Contribution to the Biology of the Tube-building Amphipod,* Corophium bonellii *(Milne-Edwards)*. University of California.

45. Kott P. 1952 The ascidians of Australia. I. Stolidobranchiata Lahille and Phlebobranchiata Lahille. *Mar. Freshw. Res.* **3**, 205–334.

46. Kott P. 1990 The Australian Ascidiacea Part 2, Aplousobranchia (1). *Mem. Queensl. Mus.* **29**, 1–266.

47. Brewin BI. 1950 Ascidians of New Zealand. Part IV. Ascidians in the vicinity of Christchurch. *Trans. Proc. R. Soc. N. Z.* **78**, 344–353.

48. Morton B. 1987 Recent marine introductions into Hong Kong. *Bull. Mar. Sci.* **41**, 503–513.

49. Bishop JD, Wood CA, Yunnie AL, Griffiths CA. 2015 Unheralded arrivals: non-native sessile invertebrates in marinas on the English coast. *Aquat. Invasions* **10**, 249–264.

50. Gu Z, Gu L, Eils R, Schlesner M, Brors B. 2014 circlize implements and enhances circular visualization in R. *Bioinformatics* **30**, 2811–2812.

51. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE. 2011 A robust, simple Genotyping-by-Sequencing (GBS) approach for high diversity species. *PLOS ONE* **6**, e19379. (doi:10.1371/journal.pone.0019379)

52. Andrews S. 2010 FastQC: a quality control tool for high throughput sequence data. *http://www.bioinformatics.babraham.ac.uk/projects/fastqc/*. See http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

53. Eaton DAR, Overcast I. 2020 ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics* **36**, 2592–2594. (doi:10.1093/bioinformatics/btz966)

54. Keenan K, McGinnity P, Cross TF, Crozier WW, Prodöhl PA. 2013 diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods Ecol. Evol.* **4**, 782–788. (doi:10.1111/2041-210X.12067)

55. R Core Team. 2016 *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

56. Jombart T. 2008 adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403–1405. (doi:10.1093/bioinformatics/btn129)

57. Alexander DH, Novembre J, Lange K. 2009 Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664. (doi:10.1101/gr.094052.109)

58. Goudet J. 2005 hierfstat, a package for R to compute and test hierarchical F-statistics. *Mol. Ecol. Notes* **5**, 184–186. (doi:10.1111/j.1471-8286.2004.00828.x)

59. Kassambara A. 2020 *ggpubr: 'ggplot2' Based Publication Ready Plots*.

60. Collin FD, Durif G, Raynal L, Lombaert E, Gautier M, Vitalis R, Marin JM, Estoup A. 2021 Extending Approximate Bayesian Computation with Supervised Machine Learning to infer demographic history from genetic polymorphisms using DIYABC Random Forest. *Mol. Ecol. Resour.* (doi:10.22541/au.159480722.26357192)

61. Seebens H, Gastner MT, Blasius B. 2013 The risk of marine bioinvasion caused by global shipping. *Ecol. Lett.* **16**, 782–790. (doi:10.1111/ele.12111)

62. Shirk RY, Hamrick JL, Zhang C, Qiang S. 2014 Patterns of genetic diversity reveal multiple introductions and recurrent founder effects during range expansion in invasive populations of *Geranium carolinianum* (Geraniaceae). *Heredity* **112**, 497–507.

63. Lockwood JL, Cassey P, Blackburn T. 2005 The role of propagule pressure in explaining species invasions. *Trends Ecol. Evol.* **20**, 223–228. (doi:10.1016/j.tree.2005.02.004)

64. Carlton JT, Geller JB. 1993 Ecological roulette: the global transport of nonindigenous marine organisms. *Science* **261**, 78–82.

65. García-Herrera R, Können GP, Wheeler DA, Prieto MR, Jones PD, Koek FB. 2005 CLIWOC: A Climatological Database for the World's Oceans 1750–1854. *Clim. Change* **73**, 1–12. (doi:10.1007/s10584-005-6952-6)

66. Zheng Y, Peng X, Liu G, Pan H, Dorn S, Chen M. 2013 High Genetic Diversity and Structured Populations of the Oriental Fruit Moth in Its Range of Origin. *PLOS ONE* **8**, 78476.

67. Fisher RA. 1922 Darwinian evolution of mutations. *The Eugenics Review* **14**, 31–34.

68. Song S, Dey DK, Holsinger KE. 2006 Differentiation among populations with migration, mutations and drift: implications for genetic inference. *Evolution* **60**, 1–12.

69. Rieseberg LH, Church SA, Morjan CL. 2004 Integration of populations and differentiation of species. *New Phytol.* **161**, 59–69.

70. Excoffier L, Foll M, Petit RJ. 2009 Genetic consequences of range expansions. *Annu. Rev. Ecol. Evol. Syst.* **40**, 481–501. (doi:10.1146/annurev.ecolsys.39.110707.173414)

71. Winkler M, Koch M, Hietz P. 2011 High gene flow in epiphytic ferns despite habitat loss and fragmentation. *Conserv. Genet.* **12**, 1411–1420.

72. Sui L, Zhang F, Wang X, Bossier P, Sorgeloos P, Hänfling B. 2009 Genetic diversity and population structure of the Chinese mitten crab *Eriocheir sinensis* in its native range. *Mar. Biol.* **156**, 1573–1583. (doi:10.1007/s00227-009-1193-2)

73. Hudson J, Viard F, Roby C, Rius M. 2016 Anthropogenic transport of species across native ranges: unpredictable genetic and evolutionary consequences. *Biol. Lett.* **12**, 20160620. (doi:10.1098/rsbl.2016.0620)

74. Tepolt CK, Palumbi SR. 2015 Transcriptome sequencing reveals both neutral and adaptive genome dynamics in a marine invader. *Mol. Ecol.* **24**, 4145–4158. (doi:10.1111/mec.13294)

75. Genton BJ, Shykoff JA, Giraud T. 2005 High genetic diversity in French invasive populations of common ragweed, *Ambrosia artemisiifolia*, as a result of multiple sources of introduction. *Mol. Ecol.* **14**, 4275–4285. (doi:https://doi.org/10.1111/j.1365-294X.2005.02750.x)

76. Simberloff D. 2009 The role of propagule pressure in biological invasions. *Annu. Rev. Ecol. Evol. Syst.* **40**, 81–102. (doi:10.1146/annurev.ecolsys.110308.120304)

77. Rius M, Darling JA. 2014 How important is intraspecific genetic admixture to the success of colonising populations? *Trends Ecol. Evol.* **29**, 233–242. (doi:10.1016/j.tree.2014.02.003)

78. Zhang Y-Y, Zhang D-Y, Barrett SCH. 2010 Genetic uniformity characterizes the invasive spread of water hyacinth (*Eichhornia crassipes*), a clonal aquatic plant. *Mol. Ecol.* **19**, 1774–1786.

79. Zhan A, Macisaac HJ, Cristescu ME. 2010 Invasion genetics of the *Ciona intestinalis* species complex: from regional endemism to global homogeneity. *Mol. Ecol.* **19**, 4678–4694. (doi:10.1111/j.1365-294X.2010.04837.x)

80. Casso M, Turon X, Pascual M. 2019 Single zooids, multiple loci: independent colonisations revealed by population genomics of a global invader. *Biol. Invasions* **21**, 3575–3592. (doi:10.1007/s10530-019-02069-8)

81. Millar RH. 1962 Further descriptions of South African ascidians. *Ann. S. Afr. Mus.* **46**, 113–221.

82. Small KS, Brudno M, Hill MM, Sidow A. 2007 Extreme genomic variation in a natural population. *PNAS* **104**, 5698–5703. (doi:10.1073/pnas.0700890104)

83. Corporation VP. 2016 *Timeline of the Port of Melbourne*. See http://www.vicports.vic.gov.au/about-us/port-history/Pages/timeline.aspx.

84. Coves G. 2010 *Concise Historical Timeline of Bunbury Port: Bunbury Port Authority*. See http://www.byport.com.au/history/hist_timeline.html.

85. Lambert CC, Lambert G. 1998 Non-indigenous ascidians in southern California harbors and marinas. *Mar. Biol.* **130**, 675–688. (doi:10.1007/s002270050289)

86. Lesieur V, Lombaert E, Guillemaud T, Courtial B, Strong W, Roques A, Auger-Rozenberg M-A. 2018 The rapid spread of *Leptoglossus occidentalis* in Europe: a bridgehead invasion. *J. Pest Sci.* **92**, 189–200.

87. Carlton JT. 1999 The scale and ecological consequences of biological invasions in the world's oceans. In *Invasive Species and Biodiversity Management* (eds OT Sandlund, PJ Schei, Å Viken), Dordrecht , The Netherlands: Kluwer Academic Publishers.

88. Carlton JT, Ruiz GM. 2005 Vector science and integrated vector management in bioinvasion ecology: conceptual frameworks. In *Invasive alien species: a new synthesis* (eds HA Mooney, RA Mack, JA McNeely, LE Neville, PJ Schei, JK Waage), pp. 36–58. Washington, DC: Island Press.