



Cite this article: Gaither MR, Bowen BW, Toonen RJ. 2013 Population structure in the native range predicts the spread of introduced marine species. *Proc R Soc B* 280: 20130409. <http://dx.doi.org/10.1098/rspb.2013.0409>

Received: 15 February 2013

Accepted: 26 March 2013

Subject Areas:

ecology, genetics, molecular biology

Keywords:

alien species, invasiveness, introduction success, meta-analysis, genetic structure

Author for correspondence:

Michelle R. Gaither

e-mail: michellergaither@gmail.com

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2013.0409> or via <http://rspb.royalsocietypublishing.org>.

Population structure in the native range predicts the spread of introduced marine species

Michelle R. Gaither^{1,2}, Brian W. Bowen² and Robert J. Toonen²

¹Section of Ichthyology, California Academy of Sciences, 55 Music Concourse Drive, San Francisco, CA 94118, USA

²Hawai'i Institute of Marine Biology, University of Hawai'i at Mānoa, PO Box 1346, Kāne'ohe, HI 96744, USA

Forecasting invasion success remains a fundamental challenge in invasion biology. The effort to identify universal characteristics that predict which species become invasive has faltered in part because of the diversity of taxa and systems considered. Here, we use an alternative approach focused on the spread stage of invasions. F_{ST} , a measure of alternative fixation of alleles, is a common proxy for realized dispersal among natural populations, summarizing the combined influences of life history, behaviour, habitat requirements, population size, history and ecology. We test the hypothesis that population structure in the native range (F_{ST}) is negatively correlated with the geographical extent of spread of marine species in an introduced range. An analysis of the available data (29 species, nine phyla) revealed a significant negative correlation ($R^2 = 0.245-0.464$) between F_{ST} and the extent of spread of non-native species. Mode F_{ST} among pairwise comparisons between populations in the native range demonstrated the highest predictive power ($R^2 = 0.464$, $p < 0.001$). There was significant improvement when marker type was considered, with mtDNA datasets providing the strongest relationship ($n = 21$, $R^2 = 0.333-0.516$). This study shows that F_{ST} can be used to make qualitative predictions concerning the geographical extent to which a non-native marine species will spread once established in a new area.

1. Introduction

The rate of species introductions has increased dramatically in modern times correlating with human population growth, increased international trade and advances in transportation [1–3]. While most introduced species never become established, those that persist can have serious economic impacts [4,5], consequences to human health [6] and pose a threat to biodiversity and ecosystem function [7–9]. Wilcove *et al.* [10] estimated that 42 per cent of endangered species in the United States are under direct threat from invasive species: an estimate that is much higher in rare taxa [11]. Consequently, biological invasions are regarded as one of the greatest modern threats to global biodiversity [12].

Forecasting which species will become invasive and which ecosystems are most vulnerable is of great scientific and practical interest [3]. Despite considerable efforts, the identification of universal characteristics that predict the success of invasive species remains elusive [13,14]. However, failure to find generalized predictive traits is not surprising given the diversity of taxa and ecosystems subject to invasion. Furthermore, characteristics important to invasion success are likely to vary among the different stages of invasion. The accumulating evidence indicates that taxon specific traits such as reproductive strategy, growth rate, environmental tolerances and diet specificity [13,15,16], combine with introduction dynamics such as habitat match and propagule pressure [17–19], to produce successful invaders [13,14]. Distilling this complexity down to even a few metrics that predict invasion success across taxa would

be valuable to ecologists and managers working to control introduced species. One possible metric is F_{ST} , a common measure of population structure based on the alternate fixation of alleles between populations (reviewed by Holsinger & Weir [20] and Bird *et al.* [21]). Because only individuals that survive dispersal events, find suitable habitat and successfully reproduce are contributing to population gene pools, F_{ST} is a potential proxy for realized dispersal (but see [22–24]). While F -statistics have been used to estimate genetic differentiation and to infer migration rates among native populations (i.e. [25]), they have not previously been used to predict the outcome of introduction events.

(a) The role of F_{ST} in predicting invasion success: a Hawaiian case study

The relationship between F_{ST} and invasion success first came to our attention while studying introduced fishes in Hawai'i. During the 1950s three fishes: the bluespinner, *Lutjanus kasmira*, the blacktail snapper, *Lutjanus fulvus* and the peacock hind, *Cephalopholis argus* were deliberately introduced into Hawaiian waters [26–28]. These three species were introduced during the same time period and in roughly equal numbers ($n = 2204$ – 3175), yet they demonstrated contrasting patterns of success [28]. *Lutjanus kasmira*, with the widest distribution in Hawai'i, demonstrates little genetic structure (low F_{ST}) across nearly 20 000 km of its natural range (figure 1; [29,30]). By contrast, *L. fulvus*, with the smallest Hawaiian distribution of the three species, showed significant population structure at all geographical scales [30]. *Cephalopholis argus* demonstrated an intermediate pattern [31]. Given the myriad of factors that influence invader success, we were surprised by the relationship between population structure in the native range and the extent of spread demonstrated by these introduced fishes. This finding prompted the question of whether this relationship is broadly applicable to marine invaders. Here, we present an analysis of the available data, across 29 species and nine phyla, to determine whether there is a significant correlation between F_{ST} , as a summary statistic surrogate for realized dispersal, and the extent of spread (invasiveness) in introduced marine species.

2. Material and methods

(a) Literature search

We conducted a Web of Science search using the following Boolean combination (phylogeography, genetic, molecular), and (invasive, introduced, alien species) to identify invasive species for which genetic surveys in the native range have been conducted. The initial literature search was conducted in August 2011. The search resulted in 125 species of potential interest (see the electronic supplementary material, table S1). For each of these 125 taxa, an additional search based on the species scientific name was conducted. For inclusion in our study, we set the following criteria: (i) the species must be primarily marine or estuarine (diadromous species were not considered), (ii) a genetic survey of at least three native locations must be available, and (iii) the extent of spread in the introduced region must be documented. The majority of species were omitted from consideration owing to the lack of published population genetic data. Suspected species complexes and species with ambiguous native ranges were omitted from the dataset (i.e. *Botryllus schlosseri*, *Ciona intestinalis* and *Styela clava*). For each case study, we




	levels of genetic structure	range in Hawai'i
	low very low levels of genetic structure across 20 000 km	2500 km
	moderate no genetic structure in the Pacific or Western Indian Oceans but high levels of structure between oceans	1200 km
	high high levels of genetic structure at all geographical scales	600 km

Figure 1. Genetic structure in the native range for each of the three fishes introduced to Hawai'i. Level of genetic structure is described and the extent of spread within the introduced range is provided.

recorded marker type and global and pairwise F_{ST} values as reported in the published literature. Because demographic data are rarely available for marine species we recorded expected heterozygosity (H_E) as a proxy for native population size. In studies where Mantel tests were conducted to test for isolation by distance (IBD) the slope of the regression was recorded. The extent of native range surveyed (km) was calculated as the shortest straight-line overwater distance between the two most distant sample sites. If desired F_{ST} values were not reported, haplotype/allele frequency files were created from the data published in tables in conjunction with sequences downloaded from GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>). When sufficient data were not available, authors were contacted and the missing data were requested. From these reconstructed datasets, H_E was calculated and global and pairwise F_{ST} [32] values between sample locations were estimated using the 'compute pairwise differences' option and 20 000 permutations in ARLEQUIN [33].

In cases such as the European green crab, *Carcinus maenas*, where multiple surveys of the native range were conducted using the same genetic marker [34,35], we used the study that reported the most complete dataset and/or covered the greatest geographical area (for *C. maenas* data from Darling *et al.* [34]). For two species (*C. maenas* and the sea walnut, *Mnemiopsis leidyi*), data from two marker types were included (see the electronic supplementary material, table S2). Several species have been introduced to multiple regions around the globe (i.e. *Microcosmus squamiger*, *Caprella mutica*, *Nematostella vectensis* and *Mya arenaria*). In these cases, we evaluated the extent of spread in the region with the earliest date of detection. The edible alga wakame (*Undaria pinnatifida*) is cultured in its native Asian range and has since been spread around the globe. We analysed the introduction to Argentina because this was the earliest introduction harbouring haplotypes from natural populations in the native range.

(b) Defining the extent of spread

The maximum extent of spread (MES) in the introduced range was defined as the shortest straight-line overwater distance (km) between the two furthest points in the introduced range. This is a highly conservative estimate of range expansion and does not account for multiple introduction events, secondary introductions or human-mediated dispersal within the introduced range. Disjunct distributions of alien species are common and likely reflect multiple or secondary introductions. To account for multiple introductions, we measured a second metric: the continuous extent of spread (CES). CES is defined as the geographical distance over

which there is suitable habitat and no gaps in distribution of greater than 100 km. For example, the Caribbean barnacle *Chthamalus proteus*, was first recorded on the Hawaiian island of O'ahu in 1995. This species has since been recorded throughout the Main Hawaiian Islands and at Midway Atoll over 2000 km northwest of O'ahu. *Chthamalus proteus* has not been detected at intermediate locations within the archipelago despite suitable habitat. MES for this species was, therefore, recorded as 2489 km (the distance between Hawai'i Island and Midway Atoll), whereas CES is recorded as 529 km (the distance between the islands of Hawai'i and Kaua'i) to reflect the possibility that *C. proteus* was secondarily transported to Midway Atoll on the hull of a boat or on fishing gear, rather than naturally jumping the intervening 1960 km.

(c) Relationship between F_{ST} and extent of spread

Using the pairwise F_{ST} values between populations in the native range, we calculated mean, median and mode values. In preliminary analyses, minimum and maximum F_{ST} were considered but were later omitted because these measures of genetic structure are heavily influenced by sampling scale and do not correlate well with other measures of F_{ST} . All data were log-transformed prior to analyses including F_{ST} values (global, mean, median and mode) and geographical distances (km). Because values of F_{ST} can be zero or negative, we added 1 to each value prior to transformation [$\ln(F_{ST} + 1)$]. Likewise, because the geographical scale over which genetic surveys are conducted (distance between two most distant sample sites) can influence the magnitude of genetic structure, we standardized global F_{ST} values by calculating F_{ST} per kilometre [$\ln(\text{global } F_{ST} + 1)/\ln(\text{km})$]. We used ordinary least-squares (OLS) regression analyses to determine if F_{ST} values (global F_{ST} , global F_{ST}/km , and mean, median and mode F_{ST} values for the pairwise comparisons) are a good predictor of the extent of spread in the introduced range. Simulation studies indicate that OLS regression analyses are preferred if the purpose of the study is not to estimate the parameters of a functional relationship, but instead to simply forecast values of the response variable for given explanatory variables (reviewed in [36]). Genetic structuring of marine taxa is often correlated with geographical distance (IBD). Therefore, in the subset of the studies in which IBD statistics were reported, we tested for a correlation between the IBD slope and extent of spread. To evaluate the influence of alternative variables on invasive success, we used the generalized linear model (GLM) with extent of spread as the response variable, F_{ST} as the explanatory variable, and marker type and H_E (of native populations) as covariates. Statistical analyses were conducted using SPSS v. 17.0 (IBM, Armonk, New York) except the GLM that was calculated using JMP Pro v. 10.0 (SAS, Cary, North Carolina). Plots of observed versus predicted residuals did not reveal any patterns that would bias the results or interpretation of the regression model (data not shown).

3. Results

(a) Description of the dataset

Searching the literature resulted in 32 cases across 29 species that met our selection criteria: 10 molluscs, four fishes, four crustaceans, three tunicates, two vascular plants and one each of cnidarian, echinoderm, ctenophore, annelid, sponge and an alga (see the electronic supplementary material, table S2). The majority of the studies were based on mitochondrial DNA (21 of 32), including cytochrome oxidase I ($n = 16$; one of these is a concatenated COI/ND2 dataset), cytochrome *b* ($n = 3$), control region ($n = 1$) and intergenic spacers ($n = 1$). Nuclear markers were used in eleven studies

and included: microsatellites ($n = 6$), allozymes ($n = 2$), internal transcribed spacers ($n = 2$) and amplified fragment length polymorphism ($n = 1$). Species included in this study are native to regions across the globe including Europe, Asia, America, Australia, South America and Oceania. Introduced regions were also geographically diverse, and included the above continents plus southern Africa. The recording of the green crab, *C. maenas*, in North America in 1817 makes this the earliest introduction in the dataset, whereas the most recent event involved the detection of the sponge *Crambe crambe* in the Canary Islands in 1995. The number of native populations surveyed per species ranged from three to 33 (mean = 9.66). Sample sizes per location within the 29 studies ranged from nine to 70 individuals (mean = 25.69) (see the electronic supplementary material, table S2).

Genetic surveys were conducted over a wide range of geographical distances. The shortest distance was 50 km between five populations of the Japanese oyster drill, *Ocenebrellus inornatus*. The largest native range surveyed was for the bluestripe snapper, *L. kasmira*, for which 10 populations were sampled across nearly 20 000 km. Global F_{ST} values were obtained for 26 of 32 studies and ranged from 0 for the soft-shelled clam *M. arenaria* to 0.906 for the alga *U. pinnatifida*. After standardizing for the geographical range over which the genetic survey was conducted, F_{ST}/km ranged from 0 to 0.002 with the starlet sea anemone *N. vectensism* demonstrating the highest value. Pairwise F_{ST} values were obtained for 30 of 32 studies. From these pairwise values we calculated mean F_{ST} (range = 0.002–0.943), median (range = –0.018 to 0.969) and mode (range = –0.100 to 1.00) (see the electronic supplementary material, table S2). In four studies, although pairwise F_{ST} values were available, the number of populations surveyed was low (3 or 4) and precluded the calculation of mode F_{ST} . MES varied widely among species ranging from 42 km for the tunicate *Pyura praeputialis* to 4344 and 4497 km for the ctenophore *M. leidyi* and the bivalve *M. arenaria*, respectively. To take secondary introductions into account, we calculated CES for each species which resulted in a narrower range from 8 km for the alga *U. pinnatifida* to 2583 km for the bluestripe snapper *L. kasmira*.

(b) Relationship between F_{ST} and extent of spread

Weersing & Toonen [37] found that global F_{ST} was poorly correlated with geographical study scale in marine organisms, explaining only 2 per cent of the variance among 149 studies. Likewise, we found no correlation between global F_{ST} and study scale in our smaller dataset ($n = 30$, $R^2 = 0.004$, $p = 0.754$): a finding that may result from the inclusion of different marker types [38]. Similar to other studies, we found that correcting for geographical scale resulted in a slightly higher correlation between F_{ST} and the extent of spread (Global F_{ST}/km ; table 1; [38,39]). Contrary to expectations, we found no correlation between F_{ST} and H_E .

We found no correlation between genetic structure in the native range and MES. However, when we corrected for secondary introductions and possible human-mediated dispersal within the introduced range, we detected a significant negative correlation between genetic structure and CES. Regardless of the F_{ST} value used (global, mean, median or mode), we found significant correlations between population structure in the native range and CES (table 1). Global F_{ST} proved to be the weakest predictor of spread ($R^2 = 0.245$, $p = 0.010$) with a

Table 1. F_{ST} as a predictor of extent of spread of introduced species. Number of studies (n), regression coefficient (R^2), and corresponding goodness of fit F -statistic and p -values are reported. R^2 values in italics are significant ($\alpha = 0.05$). Regression equations are reported for significant correlations. MES, maximum extent of spread; and CES, continuous extent of spread.

all markers							mtDNA						
genetic structure	n	R^2	goodness of fit F -statistic	p -value	regression equation	n	R^2	goodness of fit F -statistic	p -value	regression equation			
MES													
global F_{ST}	26	0.011	0.255	0.618		19	0.036	0.635	0.436				
global F_{ST}/km	26	0.017	0.422	0.522		19	0.045	0.804	0.382				
mean F_{ST}	30	0.001	0.026	0.873		21	0.016	0.302	0.589				
median F_{ST}	26	0.001	0.017	0.896		21	0.011	0.220	0.644				
mode F_{ST}	26	0.001	0.032	0.859		19	0.010	0.169	0.686				
CES													
global F_{ST}	26	<i>0.245</i>	7.801	0.010	$Y = -2.62x + 6.68$	19	0.333	8.497	0.010	$Y = -3.29x + 7.13$			
global F_{ST}/km	26	<i>0.288</i>	9.728	0.005	$Y = -29.10x + 6.83$	19	0.392	10.975	0.004	$Y = -37.88x + 7.34$			
mean F_{ST}	30	<i>0.294</i>	11.672	0.002	$Y = -4.78x + 6.77$	21	0.387	12.013	0.003	$Y = -5.58x + 7.05$			
median F_{ST}	26	<i>0.394</i>	18.201	<0.001	$Y = -4.85x + 6.69$	21	0.492	18.381	<0.001	$Y = -5.28x + 6.83$			
mode F_{ST}	26	<i>0.464</i>	20.814	<0.001	$Y = -5.25x + 6.47$	19	0.516	18.158	0.001	$Y = -5.25x + 6.53$			

Table 2. Results from generalized linear model used to evaluate the influence of alternative variables on invasive success. Extent of spread is the response variable, F_{ST} is the explanatory variable and marker type and H_E (of native populations) are covariates. p -values in italics are significant ($\alpha = 0.05$).

factor	d.f.	χ^2	p -value
CES			
marker type	1	0.075	0.784
global F_{ST}/km	1	6.626	<i>0.010</i>
mean F_{ST}	1	7.636	<i>0.006</i>
median F_{ST}	1	11.607	<i><0.001</i>
mode F_{ST}	1	16.289	<i><0.001</i>
H_E	1	0.519	0.471

slight improvement in predictive power when global F_{ST} was corrected for geographical scale of the study (global F_{ST}/km , $R^2 = 0.288$, $p = 0.005$). The best predictor of CES was the mode of the pairwise F_{ST} values ($R^2 = 0.464$, $p < 0.001$). While not as strongly correlated with spread as mode F_{ST} , both mean ($R^2 = 0.294$, $p = 0.002$) and median ($R^2 = 0.394$, $p < 0.001$) pairwise F_{ST} values were better predictors than global F_{ST} .

We did not find a significant effect for marker type, taxon or H_E on the extent of spread (table 2). However, previous studies have shown that marker type is a significant covariate when modelling global F_{ST} [37–39]. For this reason, we ran regression analyses on the mitochondrial datasets ($n = 21$) to determine if the correlation between F_{ST} and extent of spread improved. In all cases correlations were stronger when analyses were restricted to the mtDNA datasets (table 1 and figure 2). The hierarchy of predictive power between the different measures of F_{ST} was consistent between the datasets with global F_{ST} , providing the lowest predictive power ($R^2 = 0.333$, $p = 0.010$) and the mode pairwise F_{ST} having the greatest predictive power ($R^2 = 0.516$, $p < 0.001$).

Tests for IBD among native populations were available for only 10 of the 32 studies. We found no significant correlation between IBD slope and MES ($R^2 = 0.044$, $p = 0.560$). We detected a larger R^2 value for the IBD slope versus CES comparison ($R^2 = 0.190$; $p = 0.208$), however, the power to detect a significant correlation was limited owing to sample size.

4. Discussion

Identifying general characteristics that allow ecologists and managers to predict invaders has proved to be elusive. Here, we offer a metric with considerable power to predict the extent of spread in marine alien species. F_{ST} , a measure of alternate fixation of alleles, is a common proxy for realized dispersal among natural populations. Our analysis reveals a strong negative correlation between genetic structure in the native range and the extent of spread of invasive marine species across a diversity of taxa ($R^2 = 0.245$ – 0.464). An even stronger correlation was detected when marker type was taken into account ($R^2 = 0.333$ – 0.516). While there are insufficient data to determine if F_{ST} can predict the success of initial invasions, our results indicate that this metric is

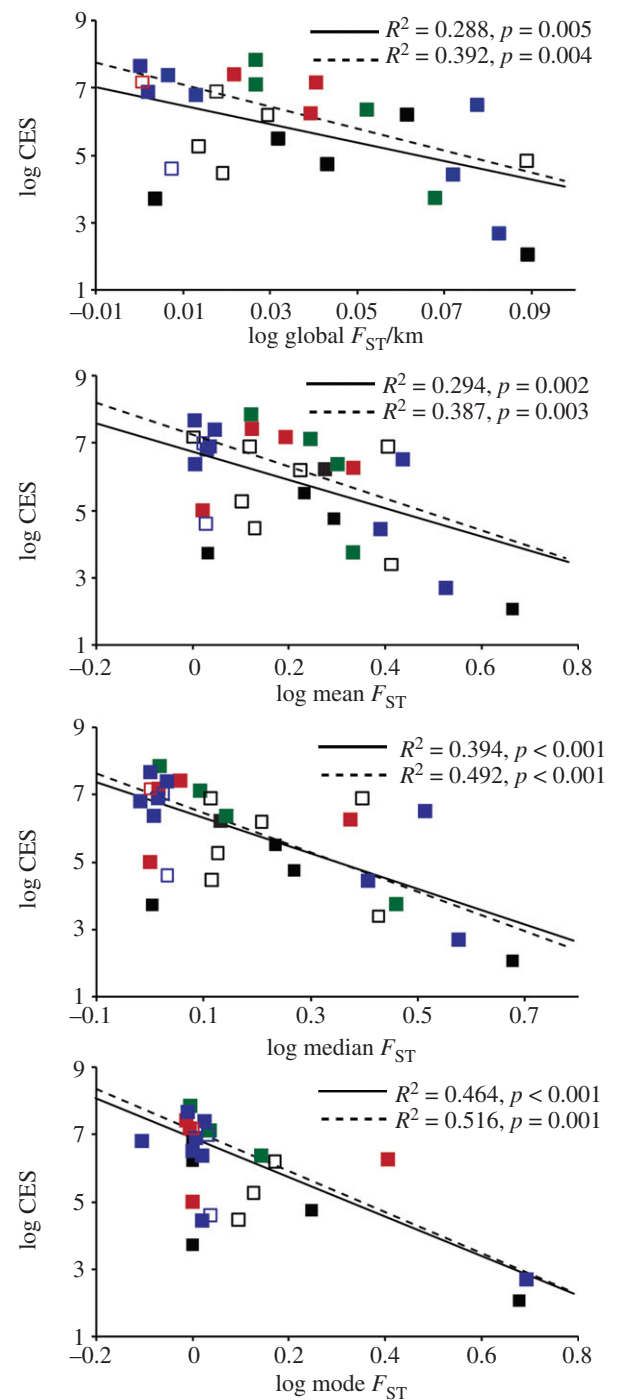


Figure 2. Population genetic structure (F_{ST}) versus continuous extent of spread (CES) across nine marine phyla (molluscs, blue; crustaceans, red; fishes, green). Regression lines are plotted for the entire dataset (solid lines) and for just the mtDNA dataset (dashed lines). R^2 values and corresponding p -values are shown.

suitable to qualitatively predict the extent of spread once an invasive species is established.

F_{ST} is not a simple measure of population differentiation, but instead is influenced by population size and genetic diversity (heterozygosity). There is a predicted relationship between population size, H_E and F_{ST} [40–42]. Under equilibrium conditions, large effective population sizes are expected to produce high genetic diversity, and low population structure (F_{ST}). Consequently, our correlation between F_{ST} and invasive spread may indicate that species that sustain large genetically diverse populations make better invaders [43]. Direct population size estimates are not available for most marine

species; however, if we assume drift/mutation equilibrium, heterozygosity can be a proxy for population size. We did not find the expected negative correlation between F_{ST} and H_E in our dataset nor was there a significant relationship between extent of spread and H_E , indicating that native population size is not driving the correlations described here.

We restricted our study to marine species because of the strikingly different modes of dispersal between land and sea. Most marine organisms have a biphasic life cycle in which adults are benthic and largely sedentary. Dispersal, over even short distances (1–10 km), is achieved primarily during a pelagic phase (eggs and larvae); the length of which varies greatly among taxa, from 0 days to more than a year in a few taxa (spiny lobsters and eels). As a result, marine larvae can be transported 100's or even 1000's of km before settlement. However, the length of the pelagic larval phase is just one factor effecting dispersal in marine organisms. Timing of spawning, local oceanographic conditions, larval swimming and sensory ability and habitat requirements all play important roles in determining effective dispersal in marine organisms. These myriad of factors probably explain why a consistent correlation between the length of the pelagic phase and F_{ST} has not been found except over small geographical scales [37,39]. In the terrestrial realm, perhaps only wind dispersed seeds have a comparable mode of dispersal. Even in fresh water taxa, many of which also have pelagic larvae, dispersal is confined to the lakes and rivers of discrete drainage systems (i.e. zebra mussel in North America) that have no clear analogy in the sea.

(a) F_{ST} as a proxy for realized dispersal

Because F_{ST} is a summary statistic, significant structure among populations may be a result of differences in effective population size, demographic or colonization history, migration, or some combination of these factors, especially for populations that may not have reached migration drift equilibrium, and thus direct interpretation of population structure in the context of gene flow can sometimes be problematic (reviewed by Hart & Marko [22], Lowe & Allendorf [23] and Marko & Hart [24]). We argue that the use of a summary statistic in this approach is warranted as it represents the cumulative effects of all factors on the genetic structure of the species, and despite the potential confounding variables, we find a significant relationship between F_{ST} and CES. Our conclusion, which is somewhat intuitive, is that species that show little or no population structure in their native range (e.g. effective dispersers, habitat or diet generalists, good competitors and broad environmental tolerances) also tend to become the most widespread invaders.

(b) Natural realized dispersal and invasiveness

Our data indicate that up to half of the variation in the extent of spread of marine invasive species can be explained by native population structure as measured by F_{ST} . Global F_{ST} had the lowest predictive power, explaining 25 per cent of the variation which increased to 39–46% using either median or mode F_{ST} . While our dataset was not sufficient to test for the effect of marker type on estimates of genetic structure, measures of F_{ST} from mtDNA are generally higher than from nDNA, complicating direct comparisons of F_{ST} among marker types ([37–39], but see [44]). Here, we detected higher correlations across estimates of F_{ST} (with as

much as 52% of the variance explained) when only the mtDNA datasets were considered.

While our findings provide a new measure of invasiveness, caution is indicated in the interpretation of F_{ST} values. Our review of the published literature resulted in a moderately sized dataset (29 species) of which the two data points with the highest F_{ST} values and correspondingly low CES had a large impact on the relationships revealed here (brown algae, *U. pinnatifida*; horn snail, *Batillaria attramentaria*). The former is anchored to substrate and the latter is believed to have a very brief pelagic larval stage. Omitting these two data points results in significant relationship only for median F_{ST} ($R^2 = 0.177$, $p = 0.026$). Examining the data plots (figure 2), it becomes clear that much of the variation in the dataset is at moderate values of F_{ST} and CES, indicating that at these values F_{ST} has less predictive power. This leads us to conclude that F_{ST} can be used to make only qualitative predictions concerning the extent of spread of invasive species [45]. Future studies may provide additional data points to fill in the high end of the F_{ST} spectrum and clarify the pattern.

The amount of time that has elapsed since introduction will influence the extent of spread. In our dataset, there is a 12-fold difference in the number of years since the earliest and most recent introductions (the green crab, *C. maenas*, to North America in 1817 and the sponge *C. crambe* to the Canary Islands in 1995). However, correcting for this variation is not straightforward. First, the time disparity between first record of occurrence and the actual date of introduction can be vast. Some alien species go undetected for decades, and survey efforts vary considerably across geographical regions [3]. Second, rates of spread vary across taxa and even through time [46]. Therefore, we were not surprised to find no correlation between population structure and invasiveness when we corrected for time since introduction (extent of spread in km per year).

Given the numerous sources of variance in the dataset, the variety of factors that determine invasion success, and the diversity of taxa and systems examined, our finding that F_{ST} explains up to 52 per cent of the variance is remarkable. We suspect that the predictive power of F_{ST} would increase if analyses could be conducted at the level of individual phyla. However, there is a paucity of population genetic data on alien species. Of the 125 candidate species considered only 23 per cent had genetic data from the native range, and the taxonomic group with the greatest coverage (molluscs, $n = 10$) was insufficient for a robust analysis.

(c) Continuous versus total extent of spread

The power of F_{ST} to predict the geographical spread of alien species is confounded by secondary introductions or human-mediated dispersal within the introduced range [34]. For example, fouling organisms such as tunicates, sponges and oysters can easily be translocated between harbours on boat hulls or fishing gear [47], whereas the larvae of some species can survive in the ballast water of ships [48–50]. The same mechanisms that promote long range introductions can also facilitate spread within the new range. The result is often disjoint distributions in the non-native range that circumvent suitable habitat: a scenario that is less likely with innate (natural) dispersal. Our data indicate that human-mediated secondary introductions are an important means of spread for many alien species.

(d) Terrestrial and freshwater systems

Several reviews have attempted to identify characteristics that predict invader success. However, none have attempted to correlate realized dispersal with invasiveness. Kolar & Lodge [13] review the plant and animal literature, and examine 68 species-level characteristics many of which influence dispersal such as reproductive mode, dispersal mechanisms in plants, fecundity and length of juvenile period. Of those characteristics only reproductive mode in plants was predictive of invasive status (plants with vegetative reproduction were more likely to spread and become abundant). Hayes & Barry [14] examined a larger set of characteristics (115 across seven biological groups) and found that only climate/habitat match was significantly associated with exotic range size across biological groups but not across studies within groups. Here, we add a new quantifiable quality to the array of invasive characteristics.

(e) Conclusions and applications

Using F -statistics to predict the outcome of marine introductions is a novel approach that shows considerable promise. F_{ST} as a surrogate for realized dispersal incorporates many of the species-level characteristics that are known to influence invader success: reproductive strategy, habitat specificity and ecology [13]. While our findings show that up to 52 per cent of the variance in the spread of marine invaders can be explained by values of F_{ST} , our data do not address the

important question of whether a species is likely to become established. Instead, F_{ST} would be most useful to wildlife managers when incorporated into specific risk assessment models with success and failure trees for each stage of introduction. In this context, F_{ST} could be used to determine which species, once established, are likely to become widespread, providing wildlife officers with a stronger scientific foundation for setting management priorities.

This study was supported by University of Hawai'i Sea Grant no. NA05OAR4171048 (B.W.B.), National Science Foundation (grant no. OIA0554657, OCE-0929031, and OCE-1260169 to R.J.T. and B.W.B.) and NOAA National Marine Sanctuaries Program (MOA grant no. 2005-008/66882 to R.J.T.). We thank the three co-trustees of the Papahānaumokuākea Marine National Monument: NOAA National Marine Sanctuaries, US Fish and Wildlife Service, and the State of Hawai'i. The views expressed herein are those of the authors and do not necessarily reflect the views of NOAA or any of their sub-agencies. Special thanks to Gail Ashton, Keith Bayha, April Blakeslee, Michael Blum, Sören Bolte, James Carlton, Janie Civile, James Coyer, John Darling, Sharyn Goldstien, Deniz Haydar, Brenden Holland, Shigeru Kojima, Susanna Lopez-Legentil, Angela Mead, Joe Roman, Marc Ruis, Evangelina Schwindt, Benoit Simon-Bouhet, Erik Sotka, Peter Teske, Frédérique Viard and Robert Ward for generously providing datasets and/or constructive comments. We thank editor Gary Carvalho, reviewer Robin Waples, and four anonymous reviewers for comments that greatly improved the manuscript. This is contribution no. 1549 from the Hawai'i Institute of Marine Biology, no. 8916 from the School of Ocean and Earth Science and Technology, and UNIH-SEAGRANT-JC-10-35 from the University of Hawai'i Sea Grant Program.

References

- Carlton JT, Geller JB. 1993 Ecological roulette: the global transport of nonindigenous marine organisms. *Science* **261**, 78–82. (doi:10.1126/science.261.5117.78)
- Cowie RH. 1998 Patterns of introduction of non-indigenous non-marine snails and slugs in the Hawaiian Islands. *Biodiv. Conserv.* **7**, 349–368. (doi:10.1023/A:1008881712635)
- Mack RN, Simberloff D, Lonsdale WM, Evans H, Clout M, Bazzaz FA. 2000 Biotic invasions: causes, epidemiology, global consequences, and control. *Ecol. Appl.* **10**, 689–710. (doi:10.1890/1051-0761(2000)010[0689:BICEGC]2.0.CO;2)
- Pimentel D, Lach L, Zuniga R, Morrison D. 2000 Environmental and economic costs of nonindigenous species in the United States. *BioScience* **50**, 53–65. (doi:10.1641/0006-3568(2000)050[0053:EAECN]2.3.CO;2)
- Pimentel D, Zuniga R, Morrison D. 2005 Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecol. Econ.* **52**, 273–288. (doi:10.1016/j.ecolecon.2004.10.002)
- Tatem AJ, Hay SI, Rogers DJ. 2006 Global traffic and disease vector dispersal. *Proc. Natl Acad. Sci. USA* **103**, 6242–6247. (doi:10.1073/pnas.0508391103)
- Williamson M, Fitter A. 1996 The varying success of invaders. *Ecology* **77**, 1661–666. (doi:10.2307/2265769)
- Cox GW. 1999 *Alien species in North America and Hawaii: impacts on natural ecosystems*. Washington, DC: Island Press.
- Simberloff D. 2005 Non-native species do threaten the natural environment. *J. Agri. Environ. Ethics* **18**, 595–607. (doi:10.1007/s10806-005-2851-0)
- Wicove DS, Rothstein D, Dubow J, Phillips A, Losos E. 1998 Quantifying threats to imperiled species in the United States. *BioScience* **48**, 607–615. (doi:10.2307/1313420)
- Armstrong S. 1995 Rare plants protect Cape's water supplies. *New Sci.* **11**, 8.
- Elton CS. 1958 *The ecology of invasions by animals and plants*. London, UK: Methuen.
- Kolar CS, Lodge DM. 2001 Progress in invasion biology predicting invaders. *Trends Ecol. Evol.* **16**, 199–204. (doi:10.1016/S0169-5347(01)02101-2)
- Hayes KR, Barry SC. 2008 Are there any consistent predictors of invasion success? *Biol. Invasions* **10**, 483–506. (doi:10.1007/s10530-007-9146-5)
- Rejmanek M, Richardson DM. 1996 What attribute make some plant species more invasive? *Ecology* **77**, 1655–1661. (doi:10.2307/2265768)
- Reichard SH, Hamilton CW. 1997 Predicting invasions of woody plants introduced into North America. *Conserv. Biol.* **11**, 193–203. (doi:10.1046/j.1523-1739.1997.95473.x)
- Roman J, Darling J. 2007 Paradox lost: genetic variation and the success of aquatic invasions. *Trends Ecol. Evol.* **22**, 454–464. (doi:10.1016/j.tree.2007.07.002)
- 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)
- Holland BS. 2000 Genetics of marine bioinvasions. *Hydrobiologia* **420**, 63–71. (doi:10.1023/A:1003929519809)
- Holsinger KE, Weir BS. 2009 Genetics in geographically structured populations: defining, estimating, and interpreting F_{ST} . *Nat. Rev. Genet.* **10**, 639–650. (doi:10.1038/nrg2611)
- Bird CE, Karl SA, Smouse P, Toonen RJ. 2011 Detecting and measuring genetic differentiation. In *Crustacean issues: phylogeography and population genetics in crustacea* (eds S Koenemann, C Held, C Schubart), pp. 31–55. Boca Raton, FL: CRC Press.
- Hart MW, Marko PB. 2010 It's about time: divergence, demography, and the evolution of developmental modes in marine invertebrates. *Int. Comp. Biol.* **50**, 643–661. (doi:10.1093/icb/icq068)
- Lowe WH, Allendorf FW. 2010 What can genetics tell us about population connectivity? *Mol. Ecol.* **19**, 3038–3051. (doi:10.1111/j.1365-294X.2010.04688.x)
- Marko PB, Hart MW. 2011 Retrospective coalescent methods and the reconstruction of metapopulation histories in the sea. *Evol. Ecol.* **26**, 291–315. (doi:10.1007/s10682-011-9467-9)
- Waples R. 1998 Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *J. Heredity* **89**, 438–450. (doi:10.1093/jhered/89.5.438)
- Oda DK, Parrish JD. 1982 Ecology of commercial snappers and groupers introduced to Hawaiian

- reefs. In *Proc. 4th Int. Coral Reef Symp.*, pp. 59–67, Manila.
27. Randall JE. 2007 *Reef and Shore Fishes of the Hawaiian Islands*. Honolulu, HI: Sea Grant College Program, University of Hawaii.
 28. Gaither MR, Toonen RJ, Bowen BW. 2012 Coming out of the starting blocks: extended lag time rearranges genetic diversity in introduced marine fishes of Hawai'i. *Proc. R. Soc. B* **279**, 3948–3957. (doi:10.1098/rspb.2012.1481)
 29. Gaither MR, Bowen BW, Toonen RJ, Planes S, Messmer V, Earle J, Robertson DR. 2010 Genetic consequences of introducing two allopatric lineages of Bluestripe Snapper (*Lutjanus kasmira*) to Hawaii. *Mol. Ecol.* **19**, 1107–1121. (doi:10.1111/j.1365-294X.2010.04535.x)
 30. Gaither MR, Toonen RJ, Robertson DR, Planes S, Bowen BW. 2010 Genetic evaluation of marine biogeographic barriers: perspectives from two widespread Indo-Pacific snappers (*Lutjanus kasmira* and *Lutjanus fulvus*). *J. Biogeogr.* **37**, 133–147. (doi:10.1111/j.1365-2699.2009.02188.x)
 31. Gaither MR, Bowen BW, Bordenave T, Rocha LA, Newman SJ, Gomez JA, van Herwerden L, Craig MT. 2011 Phylogeography of the reef fish *Cephalopholis argus* (Epinephelidae) indicates Pleistocene isolation across the Indo-Pacific Barrier with contemporary overlap in the Coral Triangle. *BMC Evol. Biol.* **11**, 189. (doi:10.1186/1471-2148-11-189)
 32. Weir BS, Cockerham CC. 1984 Estimating F -statistics for the analysis of population structure. *Evolution* **38**, 1358–1370. (doi:10.2307/2408641)
 33. Excoffier L, Lischer HEL. 2010 Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Res.* **10**, 564–567. (doi:10.1111/j.1755-0998.2010.02847.x)
 34. Darling JA, Bagley MJ, Roman J, Tepolt CK, Geller JB. 2008 Genetic patterns across multiple introductions of the globally invasive crab genus *Carcinus*. *Mol. Ecol.* **17**, 4992–5007. (doi:10.1111/j.1365-294X.2008.03978.x)
 35. Roman J, Palumbi SR. 2004 A global invader at home: population structure of the green crab, *Carcinus maenas*, in Europe. *Mol. Ecol.* **13**, 2891–2898. (doi:10.1111/j.1365-294X.2004.02255.x)
 36. Legendre P, Legendre L. 1998 *Numerical Ecology*. Amsterdam, The Netherlands: Elsevier Science BV.
 37. Weersing K, Toonen RJ. 2009 Population genetics, larval dispersal, and connectivity in marine systems. *Mar. Ecol. Prog. Ser.* **393**, 1–12. (doi:10.3354/meps08287)
 38. Riginos C, Douglas KE, Jin Y, Shanahan DF, Trembl EA. 2011 Effects of geography and life history traits on genetic differentiation in benthic marine fishes. *Ecography* **34**, 566–575. (doi:10.1111/j.1600-0587.2010.06511.x)
 39. Selkoe KA, Toonen RJ. 2011 Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. *Mar. Ecol. Prog. Ser.* **436**, 291–305. (doi:10.3354/meps09238)
 40. Slatkin M. 1985 Gene flow in natural populations. *Annu. Rev. Ecol. Syst.* **16**, 393–430. (doi:10.1146/annurev.ecolsys.16.1.393)
 41. Frankham R. 1996 Relationship of genetic variation to population size in wildlife. *Conserv. Biol.* **10**, 1500–1508. (doi:10.1046/j.1523-1739.1996.10061500.x)
 42. Neigel JE. 2002 Is F_{ST} obsolete? *Conserv. Genet.* **3**, 167–173. (doi:10.1023/A:1015213626922)
 43. Wares JP, Hughes AR, Grosberg RK. 2005 Mechanism that drive evolutionary change. In *Species invasions: insights into ecology, evolution, and biogeography* (eds FS Sax, JJ Stachowicz, SD Gaines), pp. 229–257. Sunderland, UK: Sinauer Associates, Inc.
 44. DiBattista JD, Craig MT, Rocha LA, Feldheim KA, Bowen BW. 2012 Phylogeography of the Indo-Pacific butterflyfishes, *Chaetodon meyeri* and *Chaetodon ornatissimus*: sister species reveal divergent evolutionary histories and discordant results from mtDNA and microsatellites. *J. Heredity* **103**, 617–629. (doi:10.1093/jhered/ess056)
 45. Fauby S, Barber PH. 2012 Theoretical limits to the correlation between pelagic larval duration and population genetic structure. *Mol. Ecol.* **21**, 3419–3432. (doi:10.1111/j.1365-294X.2012.05609.x)
 46. Grosholz ED. 1996 Contrasting rates of spread for introduced species in terrestrial and marine systems. *Ecology* **77**, 1680–1686. (doi:10.2307/2265773)
 47. Bax N, Williamson A, Aguero M, Gonzalez E, Geeves W. 2003 Marine invasive alien species: a threat to global biodiversity. *Mar. Policy* **27**, 313–323. (doi:10.1016/S0308-597X(03)00041-1)
 48. Medcof JC. 1975 Living marine animals in a ship's ballast water. *Proc. Natl. Shellfish Assoc.* **65**, 54–55.
 49. Carlton JT. 1985 Transoceanic and interoceanic dispersal of coastal marine organisms: the biology of ballast water. *Oceanogr. Mar. Biol. Annu. Rev.* **23**, 313–371.
 50. Gollasch S, Lenz J, Dammer M, Andres H-G. 2000 Survival of tropical ballast water organisms during a cruise from the Indian Ocean to the North Sea. *J. Plankton. Res.* **22**, 923–937. (doi:10.1093/plankt/22.5.923)