MOLECULAR ECOLOGY

Molecular Ecology (2016) **25**, 3384–3396 doi: 10.1111/mec.13682

Genomic signatures of rapid adaptive evolution in the bluespotted cornetfish, a Mediterranean Lessepsian invader

GIACOMO BERNARDI,* ERNESTO AZZURRO,† DANIEL GOLANI‡ and MICHAEL RYAN MILLER§

*Department of Ecology and Evolutionary Biology, University of California Santa Cruz, 100 Shaffer Road, Santa Cruz, CA 95060, USA, †ISPRA, Institute for Environmental Protection and Research, Sts Livorno Piazzale dei Marmi 2, 57123 Livorno, Italy, ‡Department of Evolution, Systematics and Ecology, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel, §Department of Animal Science, University of California, One Shields Avenue, Davis, CA 95616, USA

Abstract

Biological invasions are increasingly creating ecological and economical problems both on land and in aquatic environments. For over a century, the Mediterranean Sea has steadily been invaded by Indian Ocean/Red Sea species (called Lessepsian invaders) via the Suez Canal, with a current estimate of ~450 species. The bluespotted cornetfish, *Fistularia commersonii*, considered a 'Lessepsian sprinter', entered the Mediterranean in 2000 and by 2007 had spread through the entire basin from Israel to Spain. The situation is unique and interesting both because of its unprecedented rapidity and by the fact that it took this species *c.* 130 years to immigrate into the Mediterranean. Using genome scans, with restriction site-associated DNA (RAD) sequencing, we evaluated neutral and selected genomic regions for Mediterranean vs. Red Sea cornetfish individuals. We found that few fixed neutral changes were detectable among populations. However, almost half of the genes associated with the 47 outlier loci (potentially under selection) were related to disease resistance and osmoregulation. Due to the short time elapsed from the beginning of the invasion to our sampling, we interpret these changes as signatures of rapid adaptation that may be explained by several mechanisms including preadaptation and strong local selection. Such genomic regions are therefore good candidates to further study their role in invasion success.

Keywords: biological invasions, *Fistularia commersonii*, population genomics, preadaptation

Received 14 July 2015; revision received 29 March 2016; accepted 14 April 2016

Introduction

While evidence to the contrary does exist (Davis 2009; Schlaeppfer *et al.* 2011; Davis *et al.* 2015), most biological invasions result in significant negative ecological and economic impacts (Simberloff & Rejmánek 2011; Vitule *et al.* 2012). It is therefore essential to understand the factors that make invaders successful (Carlton & Geller 1993; Sax *et al.* 2005; Sala *et al.* 2011; Chown *et al.* 2015). Invasive species that move across biogeographical barriers and are exposed to severe environmental changes may adapt rapidly to new environments under the pressure of novel selective forces (Chown *et al.* 2015; Dlugosch *et al.* 2015). Changes in the selective regime acting on the invader may include both increased selection for adapted genotypes and a relaxation for defence, because of the absence of co-evolved natural enemies and pathogens (Hänfling & Kollmann 2002). These evolutionary changes in invasive populations happen over contemporary timescales and are increasingly seen as a unique opportunity to explore ecological and evolutionary processes (Lee 2002; Lawson Handley *et al.* 2011). Specifically, the identification of genes that underlie these transformations is becoming one of the most fundamental themes in

Correspondence: Giacomo Bernardi, Fax: +1 831 459 3383; E-mail: bernardi@ucsc.edu

contemporary evolutionary biology, as it gives, for example, insights on how organisms will respond to global environmental changes (Shimada *et al.* 2011; Wang *et al.* 2012).

The invasion process involves several phases that may occur sequentially or concurrently including introduction, naturalization and range expansion, with selective forces operating during each of these phases (Sakai *et al.* 2001; Lee 2002; Wares *et al.* 2005). A number of examples demonstrate a clear link between invasive success and specific adaptations (Phillips *et al.* 2006; Keller & Taylor 2008). In other cases, however, it is not always clear whether local adaptation allowed or was a consequence of the invasion (Estoup & Guillemaud 2010).

Successful invasions have traditionally been seen as the result of few individuals migrating into a new habitat, where they can thrive and quickly expand for being released from competition, predation, disease and parasitic load (Sakai *et al.* 2001; Allendorf & Lundquist 2003; Richardson 2011). Due to founding events, the genetic expectations are signatures of bottlenecks at neutral sites and lowered fitness and adaptability at selected sites. Yet, contrary to this expectation, many successful biological invasions, terrestrial or marine, do not show evidence of population bottlenecks (Holland 2000; Kolbe *et al.* 2004; Golani *et al.* 2007; Dlugosch & Parker 2008; Geller *et al.* 2010), with yet some notable exceptions (Kalinowski *et al.* 2010).

The opening of the Suez Canal in 1869 started a process of marine bioinvasions from the Indian Ocean/Red Sea into the Mediterranean with broad ecological impacts (Por 1978, 2010; Golani *et al.* 2007; Sala *et al.* 2011; Galil *et al.* 2015a). This process, commonly termed Lessepsian migration, includes today ~450 species including over 90 recorded species of fish (Fricke *et al.* 2012; Golani *et al.* 2013; Galil *et al.* 2015b). For nearly a century, the migration of Lessepsian species into the Mediterranean Sea was lessened by natural barriers. Indeed, a strong salinity barrier needed to be breached by migrating individuals. The Red Sea displays a salinity gradient from south (35‰) to north (40.5‰), peaking in Suez Bay at around 42.5‰ (Edwards 1987). Further north, the high salinity of the bitter lakes, within the Suez Canal is encountered. While the opening of the canal lowered the salinity of the bitter lakes, from ~70‰ to about 46.8‰, it still remains an impediment to migration (Edwards 1987). Finally, the low salinity created by the large freshwater plume of the Nile river outflow at the mouth of the canal, contributed to play a role as a barrier to the dispersal of Red Sea organisms (Por 1971; Galil 2007; Golani 2010). Yet, the construction of the Aswan Dam in the 1960s, that lowered the amount of freshwater outflow at the Nile Delta, and the progressive dilution of the bitter lakes, as mentioned above, have significantly dampened those obstacles. Meanwhile, the natural gradient of salinity in the Mediterranean, which varies from a low level in the western basin (36.8‰) to a higher level in the east (38.7‰) (Emeis *et al.* 2000), has increased also in part due to the flow reduction of the Nile River (Nof 1979; Borghini *et al.* 2014).

Nevertheless, differences in salinity levels remain one of the main environmental barrier between the Red Sea and the Mediterranean (Golani 2010). Indeed, adaptability to salinity gradients is considered the main predictor of invasions success for Lessepsian migrants (Belmaker *et al.* 2013; Parravicini *et al.* 2015). It is therefore possible for selective pressure to strongly affect the outcome of invasion success at this stage. So far, molecular studies on Lessepsian immigrants have highlighted similar levels of genetic diversity between exotic and native populations, with little or no evidence of genetic bottlenecks (Golani & Ritte 1999; Bucciarelli *et al.* 2002; Hassan *et al.* 2003; Hassan & Bonhomme 2005; Azzurro *et al.* 2006; Bernardi *et al.* 2010). For these species, data showed that colonization had occurred by a large number of individuals, by multiple colonization events, or a combination of both, which is consistent with predictive models of invasion success (Blackburn *et al.* 2015).

The bluespotted cornetfish, *Fistularia commersonii* (Actinopterygii, Fistulariidae), is considered to be one of the most effective invasive species of the Mediterranean Sea (Azzurro *et al.* 2012). Cornetfish are oviparous broadcast spawners with a long pelagic larval duration (46 days) and a large size at recruitment (60 mm) (Watson & Sandknop 1996; Lo-Yat 2002; Stier *et al.* 2014; Jackson *et al.* 2015b). These life-history characteristics contribute to a widespread distribution from the Tropical Eastern Pacific in Mexico to the Indian Ocean off the east coast of Africa and the Red Sea (Jackson *et al.* 2015b). Contrary to what was observed for other Lessepsian fishes, studies based on mitochondrial markers highlighted a lowered genetic diversity in the exotic cornetfish (Golani *et al.* 2007; Sanna *et al.* 2011, 2015; Jackson *et al.* 2015b). Yet, additional mitochondrial markers and one nuclear marker (Rhodopsin) showed, in Mediterranean populations, the presence of several mitochondrial haplotypes (but restricted to only two main clades) and importantly did not find a decrease in diversity at the nuclear marker level (Tenggardjaja *et al.* 2014). Taken together, these data suggested that Red Sea and Mediterranean populations have similar levels of genetic diversity.

In <7 years, starting in 2000, *F. commersonii* spread over the entire Mediterranean, which is further and faster than any other Mediterranean invader, being thus dubbed a 'Lessepsian sprinter' (Golani 2000; Karachle

et al. 2004; Azzurro *et al.* 2012; Tenggardjaja *et al.* 2014). We therefore considered that this species had great potential to show strong genomic signatures at both neutral and selected sites. The relatively small genome (1.8 Gb) of the bluespotted cornetfish (Ojima & Yamamoto 1990) makes it amenable to genome scanning approaches, providing enough power to tease out neutral and selected scenarios.

In this study, we present data based on restriction site-associated DNA (RAD) sequencing of *F. commersonii* from natural populations in the Red Sea and from the Mediterranean to understand the genomic characteristics that accompanied the recent invasion. During this process, we identified candidate loci under selection with the aim of understanding the direction of key adaptive changes necessary for a successful invasion.

Materials and methods

Sample collections

Samples of adult (sometimes co-occurring) specimens were collected by spear or obtained from local fishermen at eight localities: in the Red Sea (n = 40) from Egypt, Marsa Alam (n = 11) and Israel, Eilat (n = 29); and in the Mediterranean (n = 31) from Israel, Ashdod (n = 3), Jaffa (n = 1) and Haifa (n = 11); Greece, Rhodes (n = 10); and Italy, Lampedusa (n = 5) and (n = 1) Scilla (Fig. 1). These are the same individuals used in previous studies, where two mitochondrial and one nuclear marker were sequenced (Golani *et al.* 2007; Tenggardjaja *et al.* 2014; Jackson *et al.* 2015b). All samples were collected between 2003 and 2005, which is 3–5 years after the very beginning of the invasion (Golani 2000). Due to the lack of population structure within each region (i.e. Red Sea and Mediterranean), population analyses were carried out by comparing Red Sea and Mediterranean regions without separating each subpopulation within a region. Hereafter, we refer to our samples as clustered in two populations, one from the Red Sea and one from the Mediterranean.

Genotyping and data analysis for single nucleotide polymorphisms

Genomic DNA was extracted from tissues (n = 71) collected from two regions, the Red Sea (n = 40) and the Mediterranean (n = 31). We constructed RAD libraries using a variation of the original protocol with restriction enzyme SbfI (Miller *et al.* 2007, 2012; Baird *et al.* 2008). Initial genomic DNA concentrations for each individual were 400 ng. We physically sheared libraries on a Covaris S2 sonicator with an intensity of 5, duty cycle of 10%, cycles/burst of 200 and a cycle time of 30s. The final PCR amplification step was carried out in 50 μL reaction volumes with 18 amplification cycles. For all size selection and purification steps, we used Ampure XP beads (Agencourt). Samples used in this study were sequenced in one library containing 96 individually barcoded samples that was sequenced in a

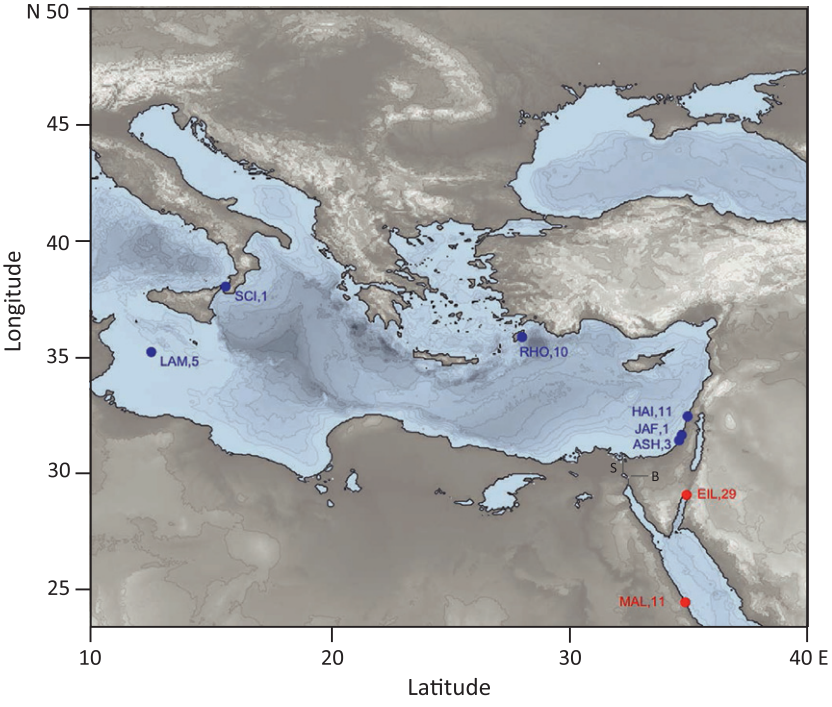

Fig. 1 Collection localities for genetic samples of *Fistularia commersonii*. Red Sea localities in red: Eilat (EIL), Israel; Marsa Alam (MAL), Egypt. Mediterranean localities in blue: Scilla (SCI) and Lampedusa (LAM), Italy; Rhodes (RHO), Greece; Haifa (HAI), Jaffa (JAF) and Ashdod (ASH), Israel. Numbers indicate sample sizes. Bitter lakes and Suez Canal are marked as B and S.

single lane on an Illumina HiSeq 2000 at the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley.

Single nucleotide polymorphism (SNP) discovery and genotyping were performed using modified Perl scripts (Miller *et al.* 2012) and using the software package STACKS version 1.29 (Catchen *et al.* 2011, 2013). All sequenced fragments were first trimmed from the 3′ end to a length of 92 bp. Low-quality reads with a probability of sequencing error >0.10% (Phred score = 33) were then filtered out. Reads without an exact match to the 6-bp barcode and 6-bp SbfI restriction site were also filtered out. For all remaining fragments, the combined 12-bp sequence (barcode plus restriction site) was then removed. Final filtered reads (80 bp) were then utilized in a population genomic analysis executed in STACKS. Putative SNPs were selected that met the following criteria: minimum depth coverage of 8×, present in at least 80% of individuals from the Red Sea and Mediterranean pooled populations. When present, only a single SNP was considered for each stack (i.e. for stacks with multiple SNPs, the first encountered SNP was the only one included in the subsequent analyses). Henceforth, a filtered 80-bp sequence used in the subsequent analyses is called a locus.

Genetic diversity and population structure

We used outputs of the population script of STACKS to create different types of infiles. Genetic diversity of each group was estimated using Arlequin (Excoffier & Lischer 2010). Fixation indexes and population structure between Red Sea and Mediterranean populations were first estimated with pairwise AMOVA Fst's directly from the output of the population script of STACKS (Weir 1996). We also estimated pairwise Fst using Arlequin (Excoffier & Lischer 2010) after modifying output stacks files using PGD SPIDER (Lischer & Excoffier 2012). Finally, we used a Bayesian approach implemented in STRUCTURE 2.2 (Pritchard *et al.* 2000; Falush *et al.* 2007) to analyse genetic clusters using structure files created by the population script of STACKS. Ten replicate runs were performed with K ranging from one to five for 100 000 iterations with a burn-in of 10 000, under the admixture model, with prior population information included to assist clustering. The most likely number of clusters in the data set was identified using the Evanno method and visualized in STRUCTURE HARVESTER (Pritchard *et al.* 2000; Evanno *et al.* 2005; Earl & VonHoldt 2012). Population structure was estimated when considering only outlier loci (a most stringent set with 47 loci, and a more relaxed set with 200 loci) and then considering only presumed neutral loci after removal of outlier loci. In this case, we used a conservative approach that removed all loci that might be under selection, with one set comprising all loci minus 200 outliers (12 598 loci), and the other set comprised all loci minus 1076 outliers (12 598 loci). Outlier loci were detected following the process outlined below.

Modes of invasion and kinship

Alternative scenarios of invasions may be tested using different approaches. One powerful approach is based on approximate Bayesian computation (ABC) (Cornuet *et al.* 2008). While this approach has been applied successfully in the past (Guillemaud *et al.* 2010; Jackson *et al.* 2015a), our data set is not amenable to this approach due to the fact that our samples were collected only a few years after introduction and therefore potentially being a mix of first-, second- and third-generation invaders, thus violating the conditions of use of ABC approaches.

Considering that there is a potential for very few (related or unrelated) individuals to have entered the Mediterranean (as evidenced by the presence of few mitochondrial haplotypes), their descendants might be closely related. To test for potential relatedness in the Mediterranean samples, we calculated kinship coefficients (Loiselle *et al.* 1995) for each pair of individuals using Genodive (Meirmans & Van Tienderen 2004). These coefficients are based on the probability of identity of two alleles for each pair of homologous genes compared between each pair of individuals. Kinship was estimated with respect to the allele frequencies for the full data set, so these coefficients provide an index of relative relatedness between each pair of individuals based on default settings in Genodive. To estimate different degrees of relatedness, we used the Loiselle *et al.*'s (1995) coancestry coefficients (full-sib = 0.25, half-sib = 0.125) to generate the following bins: 'nearly identical' ($0.57 > k > 0.375$), 'full-sib' ($0.374 > k > 0.1875$), 'half-sib' ($0.1874 > k > 0.09375$) and 'quarter-sib' ($0.09374 > k > 0.047$).

Bottlenecks

In general, tests to detect genetic bottlenecks rely on departures from allelic frequencies expected under Hardy–Weinberg equilibrium (Cornuet & Luikart 1997; Garza & Williamson 2001). Unfortunately, these test are often prone to major biases that depend on the power of detection due to the small number of samples or loci (Zachariah Peery *et al.* 2012). For Lessepsian migrants, the situation is more complex, because populations are not in equilibrium and in some cases have not gone through several generations of inbreeding. In the case of *Fistularia commersonii*, samples used in this study

were collected only a few years after the Mediterranean invasion, thus precluding the use of common bottleneck tests because of the violation of most assumptions of those tests. Functional bottlenecks, however, were investigated using genetic diversity, population structure and kinship estimates, as described above.

Fst outliers

To identify SNPs showing evidence of selection, we identified Fst outliers, which is a commonly used approach, but not without a number of potential pitfalls (Bierne *et al.* 2011, 2013; Fourcade *et al.* 2013; Lotterhos & Whitlock 2015). To identify Fst outliers, we used three different approaches. First, we estimated outliers by looking at loci that were three standard deviations above the average AMOVA Fst, as obtained directly from the STACKS population output (Weir 1996). We also carried out outlier scans with BAYESCAN version 2.1 (Foll & Gaggiotti 2008). We tested results with prior odds set at different levels, from 8 to 10 000 corresponding to a belief that the neutral model is x times more likely than the model with invoking selection at any given locus (our data set contains ~10 000 loci). A locus was considered to be an outlier if the posterior odds favouring a model with selection were above a threshold set by the software to ensure a false discovery rate of 20%. As a third approach, we used a modified Fdist method (Beaumont & Nichols 1996) implemented in the Arlequin loci under selection option (Excoffier *et al.* 2009; Excoffier & Lischer 2010). From preliminary analyses, this produced results that are similar, but apparently more conservative than the LOSITAN workbench (Antao *et al.* 2008) with 500 000 simulations and a 99.5% confidence interval, as seen also by others (Narum & Hess 2011; Tsumura *et al.* 2014). To graphically compare the results of the three approaches, we created a Venn diagram using the R package VennDiagram by Hanbo Chen and Paul C. Boutros.

Finally, all outlier loci were compared to GenBank entries with BLAST, where *E*-values of 0.001 and below were kept and recorded (probability of obtaining the same result by chance <0.001). When found as matching sequences, protein coding genes were classified using KEGG assignments (Ogata *et al.* 1999; Kanehisa *et al.* 2008; http://www.genome.jp/kegg/).

Results

Single nucleotide polymorphisms

Restriction site-associated DNA tag libraries were created by individually barcoding 71 *Fistularia commersonii* individuals from the Red Sea (40 individuals) and the Mediterranean (31 individuals) (Fig. 1). Sequences with localities and dates of collections for each individual were deposited in GenBank (Accession no. PRJNA310355). One lane of sequencing yielded more than 220 million reads. Quality filtering of the raw data set left a remaining 9.1 million reads. For each individual, we identified an average of 44 694 stacks, where each stack is comprised of filtered reads representing a potential locus (Fig. S1, Supporting information). After specifying a depth of coverage of no <8× and SNP presence in at least 80% of all individuals using the populations script in STACKS (considering one pooled population in the Red Sea and one pooled population in the Mediterranean), we identified a total of 31 923 SNPs within the RAD tag sequences, corresponding to 13 674 stacks (Fig. S1, Supporting information). In order to avoid biases due to linkage, we only used a single SNP for the subsequent analyses. Therefore, all analyses described below, unless stated, are based on 13 674 SNPs that are on 13 674 separate loci.

Population genetic diversity

We compared the genetic diversity of the Red Sea and Mediterranean populations (Table 1). While the sample numbers are slightly different between these two regions, we found that the number of polymorphic loci in the Red Sea and Mediterranean populations (8580 and 8431, respectively), the average heterozygosity (0.13226 and 0.14379, respectively) and a measure of genetic diversity, theta, (0.11650 and 0.12096, respectively) were very similar in the two populations (Table 1).

Kinship

Early studies on *F. commersonii* suggested that very few individuals (as few as two females) might have originally entered the Mediterranean (Golani *et al.* 2007). Therefore, we assumed that sampled individuals might potentially be related, especially in the Mediterranean populations. However, we did not find Mediterranean samples to be more related to each other than Red Sea samples. Indeed, we found that in the Red Sea, 41 pairwise comparisons revealed potential relatedness, while 18 of the pairwise comparisons of Mediterranean individuals suggested relatedness. Due to the difference in total numbers of pairwise comparisons (5.2% and 3.9% of the comparisons, respectively), these are nearly identical results (Table 2), indicating that Mediterranean individuals are no more related to each other than Red Sea ones. In addition, as Mediterranean samples cluster into two different mitochondrial clades (as described in

Table 1 Genetic diversity of *Fistularia commersonii* populations (usable loci represent those loci where the locus is present in more than 80% of individuals). Individuals collected in the Mediterranean were further divided into their mitochondrial clades, Clade 1 and Clade 2

	Red Sea	Mediterranean	Clade1	Clade2
Number of individuals	40	31	4	27
Number of loci	13 674	13 674	13 672	13 672
Usable Loci	10 709	11 147	5061	10 796
Polymorphic loci	8580	8431	1598	7622
Mean heterozygosity	0.1323	0.1438	0.3780	0.1537
Genetic diversity	0.0792	0.0847	0.1819	0.1192

Table 2 Kinship assessment among 71 individuals of *Fistularia commersonii* based on 13 674 loci. A number of pairwise comparisons (and their percentage) are given for putative nearly identical, full-sibs, half-sibs and quarter-sib individuals (see Materials and methods). Total number of pairwise comparisons is 2485. The number of pairwise comparisons among 40 Red Sea individuals is 780, among 31 Mediterranean individuals 465, between Red Sea and Mediterranean individuals 1240

	Red Sea (%)	Mediterranean (%)	Red Sea/ Med (%)
Nearly identical	1 (0.1)	1 (0.2)	0 (0.0)
Full-sib	0 (0.0)	0 (0.0)	2 (0.2)
Half-sib	10 (1.3)	4 (0.9)	5 (0.4)
Quarter-sib	30 (3.8)	13 (2.8)	47 (3.8)
Total related individuals	41 (5.2)	18 (3.9)	54 (4.4)

Golani *et al.* 2007), we checked whether Mediterranean individuals that belonged to the same mitochondrial clade were more related to each other than expected by chance and they were not (not shown).

Fixation index and population Structure

As expected, we did not find evidence of population structure between the Red Sea and the Mediterranean (due to the recent invasion event of unrelated individuals). Calculated Fst between those populations based on all 13 674 loci was 0.00011 ($P > 0.25$) using Arlequin. Values of AMOVA Fst calculated by STACKS varied between 0 and 0.0943 with an average of 0.0086 (Fig. S2, Supporting information). When using only neutral loci (see below for loci under selection), the lack of population structure was also illustrated by a STRUCTURE plot (Fig. 2). There, the Evanno method identified 2 or 3 clusters ($K = 2$, or $K = 3$) depending on the data set (Fig. S3, Supporting information). However, a strict adherence to K values may be problematic in some cases, as recently discussed (Meirmans 2015). Here, we interpret these results as consistent with a lack of population structure.

Evidence of selection – local adaptation

For the following analyses, we first checked SNP positions within each locus to ensure that they were not clustered, or at the beginning or the end of the sequence, and thus prone to result from sequencing errors (Davey *et al.* 2013). For the loci described below, SNP position within each 80-bp locus was on average at position 30.7 (±24.0), indicating that they were not clustered at the very beginning or the very end of the stack and unlikely to be vulnerable to sequencing error. Their average position being before the middle 40-bp mark was expected because we restricted our analyses to the first encountered SNP as mentioned in the Materials and methods section.

We used three approaches to address evidence of selection using outlier loci. Using the STACKS output of AMOVA Fst values, we found that 200 loci were three standard deviations above average (Fig. S2, Supporting information). Arlequin and Bayescan identified 147 and 89 loci under directional selection, respectively. The nine loci identified by STACKS as having the highest AMOVA Fst, where also among the highest ranked in both Arlequin and Bayescan (Fig. 3). Overall, 47 loci were identified as being under directional selection by all three methods (Fig. S4, Supporting information). Finally, Arlequin identified 22 loci as being under balancing selection (Fig. 3a); however, these loci were recovered by neither Bayescan nor STACKS.

A STRUCTURE analysis based on the 47 and 200 outlier loci, as defined by STACKS, identified, as expected, that the Red Sea and Mediterranean populations formed two distinct genetic clusters (based on the Evanno Method, Figs 2 and S3, Supporting information). However, the two clusters were not genetically homogeneous, as shown by several individuals in each cluster that were partially assigned to the other genetic cluster (Fig. 2).

Outlier loci GenBank matches

We compared the most extensive list of outliers (200 loci from the STACKS output) to GenBank entries to

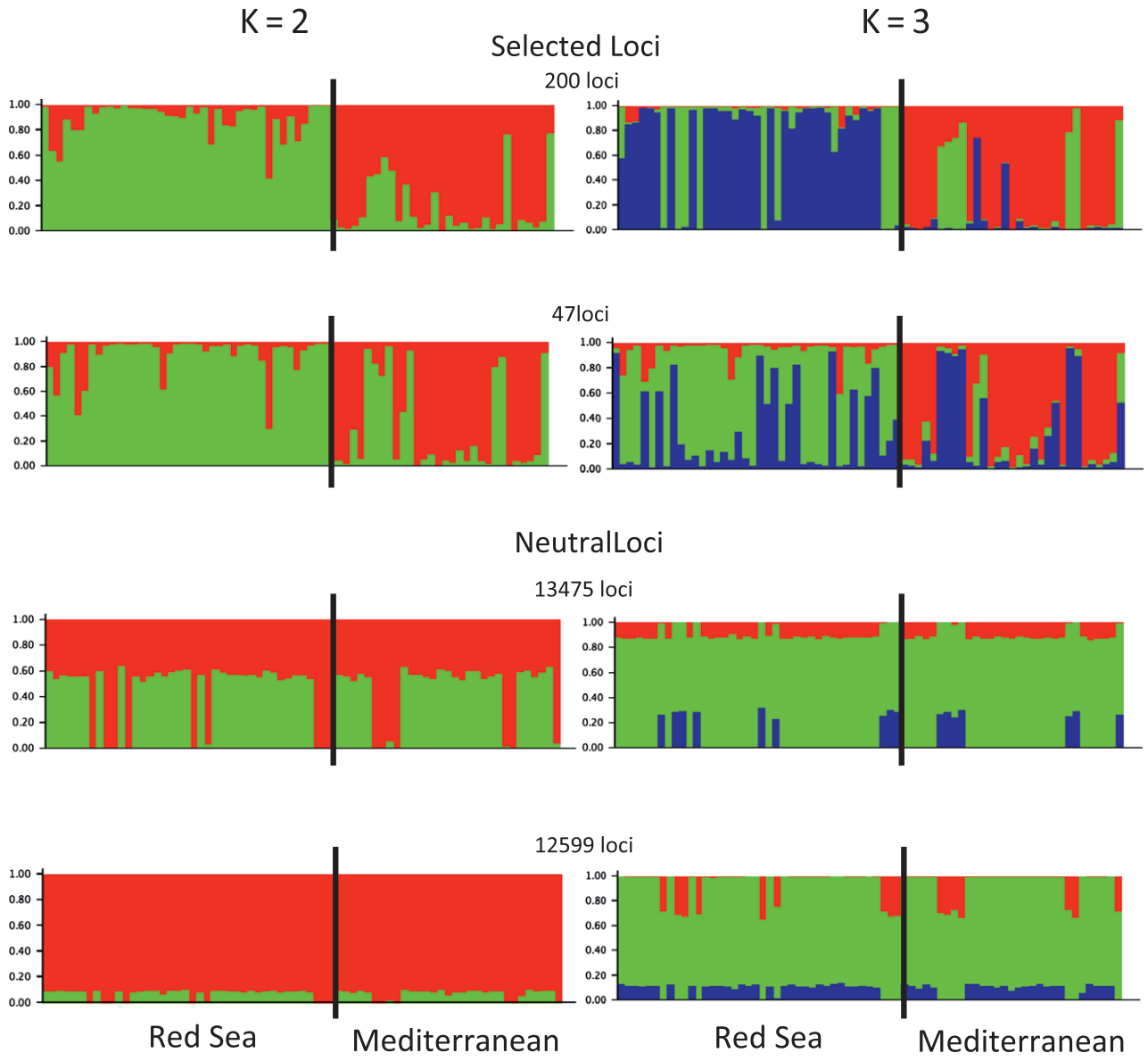

Fig. 2 Bayesian population assignment test using the program STRUCTURE based on loci presumed under selection (upper panel) and presumed neutral (lower panel). Number of loci used are above the panels and results are presented in Figure S3 (Supporting information). Results are presented with two clusters (K = 2, left column) and three clusters (K = 3, right column). Black vertical lines represent the limit between Red Sea (left) and Mediterranean (right) individuals.

identify potential matches with previously annotated sequences (Table S1, Supporting information). We found that 42 sequences (21%) had GenBank matches with *E*-values of <0.001 (0.001 to 9.e-^{24}). All 42 sequences showed highest similarities with fish sequences (Table S1, Supporting information). Of those, 32 (76%) were associated with known protein coding regions. This is strikingly (and statistically significantly) higher than the theoretical random chance expectation of matching a protein coding sequence, which, given the genome size of *F. commersonii* and number of protein coding genes, is ~0.1%, providing further evidence that the outlier loci are likely to be functionally significant. In addition, eight loci matched a unique gene cluster originally sequenced in the ricefish, *Orizias dancena* (Map7 cluster) (Table S1, Supporting information). These loci matched different regions of the sequence cluster, potentially indicating a large linked region that is under selection. Genes in that cluster have several functions, most notably linked with osmoregulation. In addition, protein analyses following KEGG protocols placed the outlier loci into few

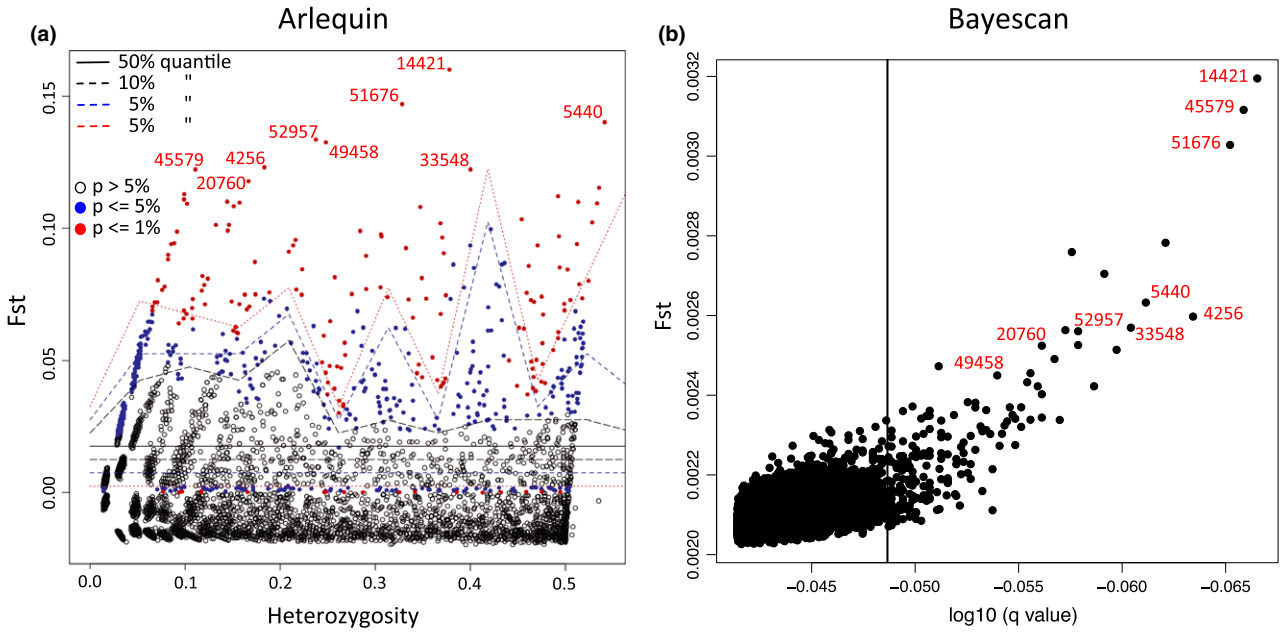

Fig. 3 Identification of Fst outlier loci. We used a modified Fdist method implemented by Arlequin (left panel a), red dots show outlier loci, and the Bayescan likelihood method (right panel b) where outlier loci are to the right of the vertical bar. Numbered loci are the nine outlier loci with highest Fst value estimated by STACKS.

categories, the largest portion of them (43%) being in the category of disease-related genes (Fig. S5, Supporting information).

Discussion

The precise mode of invasion for Lessepsian migrants is poorly known, which is in great part due to the paucity of studies in the Gulf of Suez and the Suez Canal itself (Belal & Ghobashy 2012). However, it seems that, at least for some fishes, the active dispersal of adult individuals (as opposed to larval dispersal) was responsible for the introduction of these species, as shown by parasitological evidence (Merella *et al.* 2010).

The bluespotted cornetfish, *Fistularia commersonii*, invaded the Mediterranean with extraordinary speed and range (Azzurro *et al.* 2012). Being very conspicuous and looking nothing like any native Mediterranean species, it is likely that the established presence of *F. commersonii* in 2000 is an accurate estimate of its invasion time (Golani 2000). With relatively small population sizes and few mitochondrial haplotypes found in Mediterranean populations as described in early studies (Golani *et al.* 2007; Sanna *et al.* 2011, 2015; Tenggardjaja *et al.* 2014), we expected strong signatures of genetic bottlenecks, potential for neutral drift, and also selection considering the important environmental differences between the Red Sea and the Mediterranean.

Neutral changes

For species with large distributions, such as the bluespotted cornetfish, the Red Sea is a peripheral area that is expected to harbour little genetic diversity compared to the larger Indian and Pacific Oceans. Indeed, peripheral populations of coral reef fishes experience decreased diversity in the Pacific (Gaither *et al.* 2015) and even more so in the Red Sea as for the threadfin butterflyfish, *Chaetodon auriga* (DiBattista *et al.* 2015). Yet, for the bluespotted cornetfish, Red Sea populations were shown to exhibit genetic diversity that was similarly high to their Indian Ocean and Pacific counterparts (Jackson *et al.* 2015b). In this study, we found that the genetic diversity of Red Sea and Mediterranean populations was nearly identical. While lowered genetic diversity in the Mediterranean has been the prediction for Lessepsian bioinvaders since the earliest studies (Golani & Ritte 1999; Bonhomme *et al.* 2003; Hassan *et al.* 2003), we found that *F. commersonii* samples taken only a few years after the recorded date of invasion are more likely to represent a random subsample of the native populations, rather than the outcome of a population bottleneck. In that respect, there should be no expectation of difference between native and introduced populations, as it would take several generations for these to be apparent. Results are also consistent with our findings on relatedness values, which were no higher in the Mediterranean than in the Red Sea.

Finally, STRUCTURE assignments based on neutral loci showed no population structure between Red Sea and Mediterranean.

Evidence of selection

Our data revealed strong evidence of selection, where all methods found a number of outlier loci, with 47 loci found in common with three different approaches. These results could also be visualized with a STRUCTURE analysis based on outlier loci, where strong fixation was seen between the Red Sea and the Mediterranean.

The genome of the bluespotted cornetfish is not yet available, so it is not possible to map the position of the outlier loci; however, the finding of several outlier loci in the same (and potentially syntenic) genomic region of *Orizias*, and the finding of more than expected loci associated with protein coding genes, is consistent with the idea that natural selection plays an important role in the exotic population of *F. commersonii*. Importantly, the use of outlier loci in a STRUCTURE analysis reveals more than the obvious expected population structure. Several individuals in both native and invasive population did not fully match the expected genetic assignment, indicating that loci under selection are not yet fully fixed in the invasive population.

Loci under selection

Almost half (43%) of the protein coding genes associated with loci under selection were related to immune response. This is noteworthy because the success of invading species is often assumed to be related with their capabilities of escaping common native diseases. In addition, the fact that several outlier loci match a gene complex in *Orizias*, with members that may play a role in osmoregulation is also tantalizing, because the historical role of salinity barriers during the introduction through the Suez Canal is the presence of brackish water in the bitter lakes, and freshwater plumes related to the Nile river, at the mouth of the canal (Golani 2010).

The critical importance of salinity for *F. commersonii* has also been demonstrated recently by species distribution modelling in the exotic range (Azzurro *et al.* 2012). Thus, the capacity of this species to cope with different salinity levels, especially with low salinity (Azzurro *et al.* 2012), was expected to be under strong selective pressure. As highlighted in other aquatic invaders, footprints of this selection can be tracked in loci involved in osmoregulatory functions (Shimada *et al.* 2011) but also result in realized adaptations (Lee & Petersen 2002).

Factors for a successful invasion

The tremendous success of *F. commersonii* and other Lessepsian fish species in the Mediterranean Sea has recently been explained as the result of a low direct competition with resident species, in agreement with Darwin's naturalization hypothesis (Azzurro *et al.* 2014). These evolutionary adjustments may be achieved by two different processes: preadaptation and post-invasion selection.

In preadaptation situations, genetic changes that will increase fitness in the new environment occur by chance and in the absence of a selective background before the invasion. Once in the new environment, only those individuals that carry changes that are adaptive in the new selective background are successful. It is possible that some individuals were adapted to variable salinity in further regions such as the Gulf of Aden, the Persian Gulf or the Arabian Sea (Bower *et al.* 2000) and were able to cross the Suez Canal at very low frequency. The presence of Mediterranean individuals that cluster with Red Sea individuals in the STRUCTURE plots may be a signature of such a scenario. The very low chance of immigration of preadapted individuals or individuals deriving from distant populations is consistent with the long lag time of more than 130 years for the successful colonization of the Mediterranean Sea as well as the failed attempts of colonization before the beginning of the invasion (Bariche *et al.* 2014). As for other invasive organisms, it is likely that the invasion began in earnest only when adapted genotypes reached their new selective environment (Wares *et al.* 2005; Bossdorf *et al.* 2008).

In situations where post-invasion selection is prevalent, selective pressure in the new environment is strong and acts on large numbers of individuals, retaining only individuals that carry the selected traits (Elton 1958; Baker & Stebbins 1965; Baker 1974; Bossdorf *et al.* 2008; Estoup & Guillemaud 2010). However, considering the life history of *F. commersonii*, no more than three generations might have elapsed since the very beginning of the invasion (in 2000) to the time of the study (2002–2005), and the hypothesis of in situ evolution seems unlikely. These two not mutually exclusive explanations fully represent the original and apparently contrasting ideas of Darwin on the success of exotic species (Darwin 1859).

Conclusion

Anthropogenic tampering of the environment in the region surrounding the Suez Canal resulted in major ecological changes. The opening of the canal in 1869, the construction of the Aswan dam in the 1960s, the

increased agriculture demands on Nile waters and global climate change with its associated increases in temperature and evaporation rates have all contributed to enormous changes in salinity levels in the Mediterranean and Red Sea. What was once a somewhat impenetrable barrier to marine organisms has become less of an obstacle for many. Yet, differences in salinity are still present and need to be dealt with by potential Lessepsian migrants. Here, we used genomic approaches to suggest the possibility that natural selection on osmoregulatory regions in the bluespotted cornetfish may have contributed to its highly successful invasion of the Mediterranean via the Suez Canal. Therefore, the predicted increase in temperature and salinity levels in the Mediterranean due to climate change (Borghini *et al*. 2014) and the recent doubling in size of the Suez Canal (Galil *et al*. 2015a), which would together lower the natural salinity barrier between the Mediterranean and the Red Sea, are likely to result in a major increase in the rate of invasion of the Mediterranean by Lessepsian migrants.

Acknowledgements

We would like to thank Francois Bonhomme, Gary Longo, Moises Bernal, Iria Fernandez, Eva Salas and Daniel Simberloff for discussions and Patrick Meirmans and Laurent Excoffier for help with their programs. We would like to thank Stephen Hauskins for his invaluable IT help at UCSC.

References

Allendorf FW, Lundquist LL (2003) Introduction: population biology, evolution, and control of invasive species. *Conservation Biology*, **17**, 24–30.

Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G (2008) LOSITAN: a workbench to detect molecular adaptation based on a Fst-outlier method. *BMC Bioinformatics*, **9**, 323.

Azzurro E, Golani D, Bucciarelli G, Bernardi G (2006) Genetics of the early stages of invasion of the Lessepsian rabbitfish *Siganus luridus*. *Journal of Experimental Marine Biology and Ecology*, **333**, 190–201.

Azzurro E, Soto S, Garofalo G, Maynou F (2012) *Fistularia commersonii* in the Mediterranean Sea: invasion history and distribution modeling based on presence-only records. *Biological Invasions*, **15**, 977–990.

Azzurro E, Tuset VM, Lombarte A *et al*. (2014) External morphology explains the success of biological invasions (M Vila, Ed,). *Ecology Letters*, **17**, 1455–1463.

Baird NA, Etter PD, Atwood TS, *et al*. (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, **3**, e3376.

Baker H (1974) The evolution of weeds. *Annual Review of Ecology and Systematics*, **5**, 1–24.

Baker HG, Stebbins GL (1965) *The Genetics of Colonizing Species*. Academic Press, New York, New York.

Bariche M, Kazanjian G, Azzurro E (2014) A lag of 25 years: evidence from an old capture of *Fistularia commersonii* Rüppell, 1838 from Lebanon (Mediterranean Sea). *Journal of Applied Ichthyology*, **30**, 535–536.

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society B: Biological Sciences*, **263**, 1619–1626.

Belal AAM, Ghobashy AFA (2012) Settlement behaviour and description of the lessepsian immigrant of the serpulid polychaete *Pomatoleios kraussii* in the Suez Bay. *The Egyptian Journal of Aquatic Research*, **38**, 23–30.

Belmaker J, Parravicini V, Kulbicki M (2013) Ecological traits and environmental affinity explain Red Sea fish introduction into the Mediterranean. *Global Change Biology*, **19**, 1373–1382.

Bernardi G, Golani D, Azzurro E (2010) The genetics of Lessepsian bioinvasions. In: *Fish Invasions of the Mediterranean Sea: Change and Renewal* (eds Golani D, Applebaum-Golani B), pp. 71–84. Pensoft Publishers, Sofia-Moscow.

Bierne N, Welch J, Loire E, Bonhomme F, David P (2011) The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Molecular Ecology*, **20**, 2044–2072.

Bierne N, Roze D, Welch J (2013) Pervasive selection or is it . . . ? Why are Fst outliers sometimes so frequent? *Molecular Ecology*, **22**, 2061–2064.

Blackburn TM, Lockwood JL, Cassey P (2015) The influence of numbers on invasion success. *Molecular Ecology*, **24**, 1942–1953.

Bonhomme F, Baranes AVI, Golani D, Harmelin-Vivien M (2003) Lack of mitochondrial differentiation between Red Sea and Mediterranean populations of the Lessepsian rabbitfish, *Siganus rivulatus* (Perciformes: Siganidae)*. *Scientia Marina*, **67**, 215–217.

Borghini M, Bryden H, Schroeder K, Sparnocchia S, Vetrano A (2014) The Mediterranean is becoming saltier. *Ocean Science*, **10**, 693–700.

Bossdorf O, Lipowsky A, Prati D (2008) Selection of preadapted populations allowed *Senecio inaequidens* to invade Central Europe. *Diversity and Distributions*, **14**, 676–685.

Bower AS, Hunt HD, Price JF (2000) Character and dynamics of the Red Sea and Persian Gulf outflows. *Journal of Geophysical Research*, **105**, 6387–6414.

Bucciarelli G, Golani D, Bernardi G (2002) Genetic cryptic species as biological invaders: the case of a Lessepsian fish migrant, the hardyhead silverside *Atherinomorus lacunosus*. *Journal of Experimental Marine Biology and Ecology*, **273**, 143–149.

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

Catchen JM, Amores A, Hohenlohe PA *et al*. (2011) Stacks: building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics*, **1**, 171–182.

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. *Molecular Ecology*, **22**, 3124–3140.

Chown SL, Hodgins KA, Griffin PC *et al*. (2015) Biological invasions, climate change and genomics. *Evolutionary Applications*, **8**, 23–46.

Cornuet JM, Luikart G (1997) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, **144**, 2001–2014.

Cornuet JM, Santos F, Beaumont MA *et al.* (2008) Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics*, **24**, 2713–2719.

Darwin C (1859) *The Origin of Species*. J. Murray, London.

Davey JW, Cezard T, Fuentes-Utrilla P *et al.* (2013) Special features of RAD Sequencing data: implications for genotyping. *Molecular Ecology*, **22**, 3151–3164.

Davis MA (2009) *Invasion Biology*. Oxford University Press, Oxford, UK.

Davis MA, Anderson MD, Bock-Brownstein L *et al.* (2015) Little evidence of native and non-native species influencing one another's abundance and distribution in the herb layer of an oak woodland. *Journal of Vegetation Science*, **26**, 1005–1012.

DiBattista JD, Waldrop E, Rocha LA *et al.* (2015) Blinded by the bright: a lack of congruence between colour morphs, phylogeography and taxonomy for a cosmopolitan Indo-Pacific butterflyfish, *Chaetodon auriga*. *Journal of Biogeography*, **42**, 1919–1929.

Dlugosch KM, Parker IM (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, **17**, 431–449.

Dlugosch KM, Anderson SR, Braasch J, Cang FA, Gillette HD (2015) The devil is in the details: genetic variation in introduced populations and its contributions to invasion. *Molecular Ecology*, **24**, 2095–2111.

Earl DA, VonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.

Edwards F (1987) Climate and oceanography. In: *Red Sea* (eds Edwrads A, Head S), pp. 46–69. Pergamon Press, Oxford.

Elton C (1958) *The Ecology of Invasions by Animals and Plants*. Methuen, London, UK.

Emeis KC, Struck U, Schulz HM *et al.* (2000) Temperature and salinity variations of Mediterranean Sea surface waters over the last 16,000 years from records of planktonic stable oxygen isotopes and alkenone unsaturation ratios. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **158**, 259–280.

Estoup A, Guillemaud T (2010) Reconstructing routes of invasion using genetic data: why, how and so what? *Molecular Ecology*, **19**, 4113–4130.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically structured population. *Heredity*, **103**, 285–298.

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, **7**, 574–578.

Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics*, **180**, 977–993.

Fourcade Y, Chaput-Bardy A, Secondi J, Fleurant C, Lemaire C (2013) Is local selection so widespread in river organisms? Fractal geometry of river networks leads to high bias in outlier detection. *Molecular Ecology*, **22**, 2065–2073.

Fricke R, Golani D, Appelbaum-Golani B (2012) First record of the Indian Ocean anchovy *Stolephorus insularis* Hardenberg, 1933 (Clupeiformes: Engraulidae) in the Mediterranean. *BioInvasions Records*, **1**, 303–306.

Galil BS (2007) Loss or gain? Invasive aliens and biodiversity in the Mediterranean Sea. *Marine Pollution Bulletin*, **55**, 314–322.

Gaither MR, Bernal MA, Coleman RR *et al.* (2015) Genomic signatures of geographic isolation and natural selection in coral reef fishes. *Molecular Ecology*, **24**, 1543–1557.

Galil BS, Boero F, Campbell ML *et al.* (2015a) "Double trouble": the expansion of the Suez Canal and marine bioinvasions in the Mediterranean Sea. *Biological Invasions*, **17**, 973–976.

Galil B, Boero F, Fraschetti S *et al.* (2015b) The enlargement of the Suez Canal and introduction of non-indigenous species to the Mediterranean Sea. *Association for the Science of Limnology and Oceanography*, **24**, 43–45.

Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology*, **10**, 305–318.

Geller J, Darling J, Carlton J (2010) Genetic perspectives on marine biological invasions. *Annual Review of Marine Science*, **2**, 367–393.

Golani D (2000) First record of the bluespotted cornetfish from the Mediterranean Sea. *Journal of Fish Biology*, **56**, 1545–1547.

Golani D (2010) Colonization of the Mediterranean by Red Sea fishes via the Suez Canal - Lessepsian migration. In: *Fish Invasions of the Mediterranean Sea: Change and Renewal* (eds Golani D, Appelbaum-Golani B), pp. 145–188. Pensoft Publishers, Sofia.

Golani D, Ritte U (1999) Genetic relationship in goatfishes (Mullidae : Perciformes) of the Red Sea and the Mediterranean, with remarks on Suez Canal migrants*. *Scientia Marina*, **63**, 129–135.

Golani D, Azzurro E, Corsini-Foka M *et al.* (2007) Genetic bottlenecks and successful biological invasions: the case of a recent Lessepsian migrant. *Biology Letters*, **3**, 541–545.

Golani D, Fricke R, Appelbaum-Golani B (2013) First record of the Senegalese sole, *Solea senegalensis* (Actinopterygii: Pleuronectiformes: Soleidae) from the Mediterranean coast of Israel. *Acta Ichthyologica Et Piscatoria*, **43**, 319–321.

Guillemaud T, Beaumont MA, Ciosi M, Cornuet J-M, Estoup A (2010) Inferring introduction routes of invasive species using approximate Bayesian computation on microsatellite data. *Heredity*, **104**, 88–99.

Hänfling B, Kollmann J (2002) An evolutionary perspective of biological invasions. *Trends in Ecology & Evolution*, **17**, 545–546.

Hassan M, Bonhomme F (2005) No reduction in neutral variability of mitochondrial and nuclear genes for a Lessepsian migrant, *Upeneus moluccensis*. *Journal of Fish Biology*, **66**, 865–870.

Hassan M, Harmelin-Vivien M, Bonhomme F (2003) Lessepsian invasion without bottleneck : example of two rabbitfish species (*Siganus rivulatus* and *Siganus luridus*). *Journal of Experimental Marine Biology and Ecology*, **291**, 219–232.

Holland BS (2000) Genetics of marine bioinvasions. *Hydrobiologia*, **420**, 63–71.

Jackson H, Strubbe D, Tollington S *et al.* (2015a) Ancestral origins and invasion pathways in a globally invasive bird correlate with climate and influences from bird trade. *Molecular Ecology*, **24**, 4269–4285.

Jackson AM, Tenggardjaja K, Perez G *et al.* (2015b) Phylogeography of the bluespotted cornetfish, *Fistularia commersonii*: a predictor to bioinvasion success? *Marine Ecology*, **36**, 887–896.

Kalinowski ST, Muhlfeld CC, Guy CS, Cox B (2010) Founding population size of an aquatic invasive species. *Conservation Genetics*, **11**, 2049–2053.

Kanehisa M, Araki M, Goto S *et al.* (2008) KEGG for linking genomes to life and the environment. *Nucleic Acids Research*, **36**, 480–484.

Karachle PK, Triantaphyllidis C, Stergiou KI (2004) Bluespotted cornetfish, *Fistularia commersonii* Ruppell, 1838: a Lessepsian sprinter. *Acta Ichthyologica et Piscatoria*, **34**, 103–108.

Keller SR, Taylor DR (2008) History, chance and adaptation during biological invasion: separating stochastic phenotypic evolution from response to selection. *Ecology Letters*, **11**, 852–866.

Kolbe JJ, Glor RE, Rodriguez Schettino L *et al.* (2004) Genetic variation increases during biological invasion by a Cuban lizard. *Nature*, **431**, 177–181.

Lawson Handley LJ, Estoup A, Evans DM *et al.* (2011) Ecological genetics of invasive alien species. *BioControl*, **56**, 409–428.

Lee CE (2002) Evolutionary genetics of invasive species. *Trends in Ecology & Evolution*, **17**, 386–391.

Lee CE, Petersen CH (2002) Genotype-by-environment interaction for salinity tolerance in the freshwater-invading copepod *Eurytemora affinis*. *Physiological and Biochemical Zoology*, **75**, 335–344.

Lischer HEL, Excoffier L (2012) PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, **28**, 298–299.

Loiselle BA, Sork VL, Nason J, Graham C (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiacea). *American Journal of Botany*, **82**, 1420–1425.

Lotterhos KE, Whitlock MC (2015) The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology*, **24**, 1031–1046.

Lo-Yat A (2002) *Variabilite´ temporelle de la colonisation par les larves de poissons de l'atoll de Rangiroa (Tuamotu, Polyne´sie franc¸aise) et utilisation de l'outil "otolithe" de ces larves*. (Doctoral dissertation). Université de la Polynésie Francaise, Tahiti.

Meirmans PG (2015) Seven common mistakes in population genetics and how to avoid them. *Molecular Ecology*, **24**, 3223–3231.

Meirmans PG, Van Tienderen PH (2004) Genotype and genodive: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, **4**, 792–794.

Merella P, Casu M, Garippa G, Pais A (2010) Lessepsian fish migration: genetic bottlenecks and parasitological evidence. *Journal of Biogeography*, **37**, 975–978.

Miller MR, Dunham JP, Amores A, Cresko WA, Johnson EA (2007) Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research*, **17**, 240–248.

Miller MR, Brunelli JP, Wheeler PA *et al.* (2012) A conserved haplotype controls parallel adaptation in geographically distant salmonid populations. *Molecular Ecology*, **21**, 237–249.

Narum SR, Hess JE (2011) Comparison of Fst outlier tests for SNP loci under selection. *Molecular Ecology Resources*, **11**, 184–194.

Nof D (1979) On man-induced variations in the circulation of the Mediterranean Sea. *Tellus*, **31**, 558–564.

Ogata H, Goto S, Sato K *et al.* (1999) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, **27**, 29–34.

Ojima Y, Yamamoto K (1990) Cellular DNA contents of fishes determined by flow cytometry. *La Kromosomo II*, **57**, 1871–1888.

Parravicini V, Azzurro E, Kulbicki M, Belmaker J (2015) Niche shift can impair the ability to predict invasion risk in the marine realm: an illustration using Mediterranean fish invaders. *Ecology Letters*, **18**, 246–253.

Phillips BL, Brown GP, Webb JK, Shine R (2006) Invasion and the evolution of speed in toads. *Nature*, **439**, 803.

Por FD (1971) One hundred years of Suez Canal - a century of Lessepsian migration: retrospect and viewpoints. *Systematic Zoology*, **20**, 138–159.

Por FD (1978) *Lessepsian Migration – The Influx of Red Sea Biota into the Mediterranean by Way of the Suez Canal*. Springer-Verlag, Berlin.

Por FD (2010) The new Tethyan ichthyofauna of the Mediterranean - historical background and prospect. In: *Fish Invasions of the Mediterranean Sea: Change and Renewal* (eds Golani D, Appelbaum-Golani B), pp. 13–33. Pensoft Publishers, Sofia-Moscow.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Richardson DM (2011) *Fifty Years of Invasion Ecology: The Legacy of Charles Elton*. Wiley-Blackwell, Oxford.

Sakai AK, Allendorf FW, Holt JS *et al.* (2001) The population biology of invasive species. *Annual Review of Ecology and Systematics*, **32**, 305–332.

Sala E, Kizilkaya Z, Yildirim D, Ballesteros E (2011) Alien marine fishes deplete algal biomass in the Eastern Mediterranean. *PLoS ONE*, **6**, e17356.

Sanna D, Merella P, Lai T *et al.* (2011) Combined analysis of four mitochondrial regions allowed the detection of several matrilineal lineages of the lessepsian fish *Fistularia commersonii* in the Mediterranean Sea. *Journal of the Marine Biological Association of the United Kingdom*, **91**, 1289–1293.

Sanna D, Scarpa F, Lai T *et al.* (2015) *Fistularia commersonii*: walking through the Lessepsian paradox of mitochondrial DNA. *Italian Journal of Zoology*, **82**, 499–512.

Sax DF, Stachowicz JJ, Gaines SD (2005) *Species Invasions: Insights into Ecology, Evolution and Biogeography*. Sinauer Associates Inc., Sunderland, Massachusetts.

Schlaeppfer MA, Sax DF, Olden J (2011) The potential conservation value of non-native species. *Conservation Biology*, **25**, 428–437.

Shimada Y, Shikano T, Merilä J (2011) A high incidence of selection on physiologically important genes in the threespined stickleback, *Gasterosteus aculeatus*. *Molecular Biology and Evolution*, **28**, 181–193.

Simberloff D, Rejmánek M (2011) *Encyclopedia of Biological Invasions*. University of California Press, Berkeley, 765 pp.

Stier AC, Hein AM, Parravicini V, Kulbicki M (2014) Larval dispersal drives trophic structure across Pacific coral reefs. *Nature Communications*, **5**, 5575.

Tenggardjaja K, Jackson AM, Leon F *et al.* (2014) Genetics of a Lessepsian sprinter the bluespotted cornetfish, *Fistularia commersonii*. *Israel Journal of Ecology and Evolution*, **59**, 181–185.

Tsumura Y, Uchiyama K, Moriguchi Y *et al.* (2014) Genetic differentiation and evolutionary adaptation in *Cryptomeria japonica*. *G3: Genes, Genomes, Genetics*, **4**, 2389–2402.

Vitule JRS, Freire CA, Vazquez DP, Nuñez MA, Simberloff D (2012) Revisiting the potential conservation value of non-native species. *Conservation Biology*, **26**, 1153–1155.

Wang T, Chen G, Zan Q, Wang C, Su YJ (2012) AFLP genome scan to detect genetic structure and candidate loci under selection for local adaptation of the invasive weed *Mikania micrantha*. *PLoS ONE*, **7**, e41310.

Wares JP, Hughes AR, Grosberg RK (2005) Mechanisms that drive evolutionary change. Insights from species introductions and invasions. In: *Species Invasions: Insights in to Ecology, Evolution, and Biogeography* (eds Sax DF, Stachowicz JJ, Gaines SD), pp. 259–290. Sinauer Associates Inc., Sunderland, Massachusetts.

Watson W, Sandknop E (1996) Fistulariidae: cornetfishes. In: *The Early Stages of Fishes in the California Current Region* (ed. Moser H), pp. 718–723. California Cooperative Oceanic Fisheries Investigations (CalCOFI), Lawrence, Kansas.

Weir BS (1996) *Genetic Data Analysis II*. Sinauer Associates Inc., Sunderland, Massachusetts.

Zachariah Peery M, Kirby R, Reid BN *et al.* (2012) Reliability of genetic bottleneck tests for detecting recent population declines. *Molecular Ecology*, **21**, 3403–3418.

G.B., E.A., D.G. and M.R.M. conceived and designed the experiments; G.B. performed the experiments; G.B., E.A., D.G. and M.R.M. contributed materials and analytical tools; and G.B., E.A., D.G. and M.R.M. wrote the article.

Data accessibility

Data were submitted to GenBank, where they are freely available.

Data deposited in GenBank (Accession no. PRJNA310355).

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1. GenBank matches with outlier loci. Loci (42 out of 200 outlier loci) with GenBank matches with E values of <0.001 are represented. Columns correspond from left to right to Locus number, Locus ID, AMOVA Fst, GenBank Locus name, E-value and GenBank Accession number.

Figure S1. Histograms representing the number of stacks, polymorphic loci, and SNPs for each sampled individual.

Figure S2. Histogram representing the distribution of pairwise Fst's, with an arrow pointing at the value for 3 standard deviations.

Figure S3. Results of STRUCTURE HARVESTER based on the Evanno Method (see Materials and Methods). Best likelihood is highlighted in yellow. Four datasets were used, with presumed selected and neural loci. These results are presented graphically in Figure 2.

Figure S4. Venn diagram of the number of outlier loci identified by the three methods used in this study: STACKS, Arlequin, and Bayescan.

Figure S5. Functional groups of protein coding genes that match outlier loci using KEGG.