Biol Invasions
https://doi.org/10.1007/s10530-019-01992-0

ORIGINAL PAPER

The genomics of invasion: characterization of red lionfish (*Pterois volitans*) populations from the native and introduced ranges

M. O. Burford Reiskind · E. M. X. Reed · A. Elias · J. J. Giacomini · A. F. McNear · J. Nieuwsma · G. A. Parker · R. B. Roberts · R. E. Rossi · C. N. Stephenson · J. L. Stevens · B. E. Williams

Received: 12 June 2018 / Accepted: 11 April 2019

Abstract Invasive species are one of the greatest threats to global biodiversity and ecosystem health, and population genetics provides promising tools for understanding the evolutionary process of successful invaders. The well-documented introduction of the red lionfish (*Pterois volitans*) to the western Atlantic, Gulf of Mexico and Caribbean has decimated native fauna due to the invader's voracious predation and growth rate. We tested whether our samples were within the region of the source of invasion into the Atlantic and Caribbean and investigated whether hybridization in the native or introduced range was responsible for the success of this invasive species. We used a reduced representation sequencing method to generate over 50,000 single nucleotide polymorphisms and sequence data from two mitochondrial DNA genes to analyze the population patterns. We found one location in the southeastern Pacific that was genetically similar to one location the southwestern Atlantic and evidence of the subsequent spread south to the Gulf of Mexico and the Caribbean, which supports previous findings. Within the native range, we found genetic divergences commensurate with distinct species and evidence of hybridization. We found limited structure within the introduced range and no evidence of *Pterois miles* or hybrids within this range. Finally, we found signatures of selection between the native and introduced range that may be a result of the introduction. Overall, this study of the red lionfish showed the pattern of introduction and suggested a deeper sampling of genomic data from individuals within the native range may reveal hybridization between species as the source of the aggressive invasion.

Keywords Invasion · Hybridization · Red lionfish · *Pterois volitans* · *Pterois miles* · ddRADseq

After first and second authors, author names were arranged in alphabetical order.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10530-019-01992-0) contains supplementary material, which is available to authorized users.

M. O. Burford Reiskind (✉) · E. M. X. Reed · J. J. Giacomini · A. F. McNear · R. E. Rossi · J. L. Stevens
Department of Applied Ecology, North Carolina State University, Campus Box 7617, Raleigh, NC 27695, USA
e-mail: mbreiski@ncsu.edu

A. Elias · J. Nieuwsma · G. A. Parker · R. B. Roberts · C. N. Stephenson · B. E. Williams
Department of Biological Sciences, North Carolina State University, Raleigh, NC 27695, USA

G. A. Parker · R. B. Roberts
WM Keck Center for Behavioral Biology, North Carolina State University, Raleigh, NC 27695, USA

Published online: 13 April 2019

Introduction

Invasive species present a major threat to biodiversity and ecosystem health (Shiganova 1998; Grosholz et al. 2000), second only to habitat loss (Walker and Steffen 1997; Allendorf and Lundquist 2003). While gaining information on how and when species become invasive is critical, most studies focus on documenting presence and spread of invasive species, not on the mechanism of invasion. The genetic characteristics and evolution of invasive species, which reflects how invasive species arrive, overcome small population dynamics, and rapidly adjust to novel environments in the introduced range, enhances our ability to predict whether an introduced species establishes and spreads. With the increase in accessibility of genomic information, we can generate accurate estimates of population parameters and make more informed conservation decisions about the management of both established and introduced species. One of the genetic paradoxes of introduced species is that small founding populations should have low genetic diversity and low capacity to adapt, yet many introduced species have greater genetic variation than native species, are better competitors, and/or benefit from predator release (Lee 2002; Kolbe et al. 2004; Frankham 2005; Allendorf et al. 2013). A second genetic paradox specific to native species receiving the invasive species is that despite better adaptation to their home environment they are not able to outcompete the introduced species that evolved under different selective regimes (Allendorf et al. 2013). To resolve these questions using a genomic approach, we focus on the genetic characteristics of the red lionfish (*Pterois volitans*), an invasive species with a large impact on native fauna.

Lionfish were first reported in Florida in 1985, likely because of releases from the marine ornamental aquarium trade. Since this first report, lionfish spread north along the Atlantic coast of the USA, as far as New England, south into the Gulf of Mexico and Caribbean Sea, and extending to the east coast of Brazil (Courtenay 1995; Schofield 2009; Morris and Akins 2009; Ferreira et al. 2015). Early reports suggest that two species native to the Indo-Pacific made up the invasive populations, *P. volitans* and *Pterois miles* (devil fire fish), with the former dominating the introduced range (Schofield 2009). The invasion of predatory lionfish in the Western Atlantic has had substantial negative effects on native species, such as groupers, snapper and smaller forage fish (Albins and Hixon 2008) and are an important concern to coastal fisheries managers (Whitfield et al. 2002).

The lionfish is an excellent invasive species due to its ability to inhabit a wide variety of habitats, its high fecundity, its popularity in the aquarium trade, and its release from natural predators (Cure et al. 2014). As with many invasive species, lionfish are found at much higher densities on reefs in their introduced range, while they are relatively sparse and difficult to detect within their native range (Kulbicki et al. 2012). Researchers classify lionfish as a voracious predator, consuming a wide variety of fish species and often reducing species richness on reefs throughout its introduced range (Albins 2015; Layman and Allgeier 2012). Lionfish also possess many additional traits that make them exceptional invaders, including defensive venomous spines, cryptic morphology, high growth and reproductive rates, and larval dispersion life-history strategy (Jud and Layman 2012). While the seasonality of lionfish reproduction is unknown in their native range, within the introduced range lionfish spawn every 4 days, year-round, producing around two million planktonic eggs per year, which can travel along the ocean currents, covering very large distances (Morris and Akins 2009; Ahrenholz and Morris 2010). Larval connectivity models suggest that dispersal of lionfish occurs over great distances, from the Bahamas to the New England coast, during the pelagic larval phase (Cowen et al. 2006). An escape from predation paired with such a high rate of reproduction could be key to the success of the lionfish invasion in the Atlantic.

A current hypothesis is that the invasion of *P. volitans* originated from a western Indonesia population (Hamner et al. 2007). Previous genetic analyses found strong founder effects in the introduced range, supported by lower haplotype diversity in the western Atlantic and high haplotype diversity in native ranges (Freshwater et al. 2009; Hamner et al. 2007; Toledo-Hernandez 2014). A study using both population genetic and demographic models to simulate the spread of the invasive lionfish estimated the initial population sizes of 96–272 colonists to the Atlantic (Selwyn et al. 2017), which would also support the strong founder effects in the introduced range. Further studies concluded that there is less genetic diversity in the southern U.S. populations than the northern populations, even though there is no known

geographical separation between them, which alludes to the possibility of multiple invasive events (Butterfield et al. 2015). Butterfield et al. (2015) also found no evidence of *P. miles* in their sampling regions, which questions previous observations of *P. miles* in collections or suggests the potential for hybridization between the two.

A recent phylogenetic reanalysis of closely related lionfish species (*P. volitans*, *P. miles*, *P. russelii* and *P. lunulata*), using sequence data from one mtDNA gene and two nuclear introns indicated that within the native range *P. lunulata* and *P. russelii* were not evolutionary distinct lineages and hybridized with *P. miles* in the Indian Ocean to produce *P. volitans* (Wilcox et al. 2017). The phylogenetic placement of the species followed geographic origin with *P. miles* from the Indian Ocean and the *P. lunulata*/*P. russelii* lineage from the eastern Indian Ocean to the western Pacific. In this reanalysis, *P. volitans* mixes with either of these species groups depending on location and mtDNA or nuclear gene analysis. This recent study has motivated a more comprehensive sampling of the genome to understand the dynamics of hybridization. Whether it is asymmetric (female of one species to male of another) or bidirectional, or whether the putative hybrid *P. volitans* backcrosses with either parental lineage, the question remains whether this hybrid is invasive not only in the Atlantic, but also throughout the Indo-Pacific.

Understanding genetic composition of invasive lionfish populations can shed light on possible source populations, help track dispersal, and identify genetic loci and related adaptive traits under selection. Founder events typically lead to significant bottleneck pressures that can reduce genetic diversity and potentially slow population growth (reviewed in Bouzat 2010). In contrast, multiple introductions, hybrid vigor, and selection for newly adaptive alleles may reduce bottleneck pressures and facilitate further expansion (Fitzpatrick and Shaffer 2007). The use of genomic information to identify population structure and outlier loci under selection improves accuracy of analyses, such as effective population size, migration rate, and population divergence (Allendorf et al. 2010). One major issue with studies on the genetic structure of small population dynamics is a lack of power to detect genetic differences or the signature of rapid evolution when you have relatively few markers. Reduced representation sequencing provides a method to increase the number and resolution of markers in the genome, allowing researchers to parse out neutral and adaptive locations in the genome and address a variety of questions including those related to invasive species and conservation (Gaither et al. 2015; Bernardi et al. 2016; Burford Reiskind et al. 2016; 2018).

Here, we address the population genetic structure of red lionfish within both the native and introduced ranges, and investigate whether there are regions of the genome under selection by conducting an outlier loci analysis between samples of *P. volitans* from its native and introduced range. The goal is to understand potential source population of introduction, geographic genetic structure, and signatures of genetic diversity and selection in the genome. In addition, we address any genotype by environment association of individuals in the native or introduced ranges of putative *P. volitans*. Given recent evidence of the origin of *P. volitans* as a hybrid species between *P. miles* and a distinct lineage (*P. russelii* and *P. lunulata*; Wilcox et al. 2017), we also address how much of the genome of species morphologically identified as *P. volitans* in the invasive range is shared with *P. miles* sampled within its native range. Genomic approaches have enhanced our ability to address these questions, increasing the power to detect important characteristics that allow a particular species to be a highly effective invader. The results of this study provide tools for active management of the successful red lionfish invasion, with the ultimate goal of limiting its impact on native fauna.

Methods

Sampling

We sampled red lionfish (*P. volitans*) from five locations in its invasive and two from its native range (Fig. 1, Table 1). These locations included Bahamas (n = 8), Gulf of Mexico (n = 13), Belize (n = 18), Honduras (n = 11), Panama (n = 12), Philippines (n = 9), and Taiwan (n = 8). We also had samples of *P. miles* individuals from Sri Lanka (n = 2), and *P. antennata* individuals from the Yale Museum (n = 2) for interspecies comparisons. The five sites from the Southern Atlantic and Caribbean were from the invasion range and the two sites from the eastern Pacific (Taiwan and Philippines) were from its native

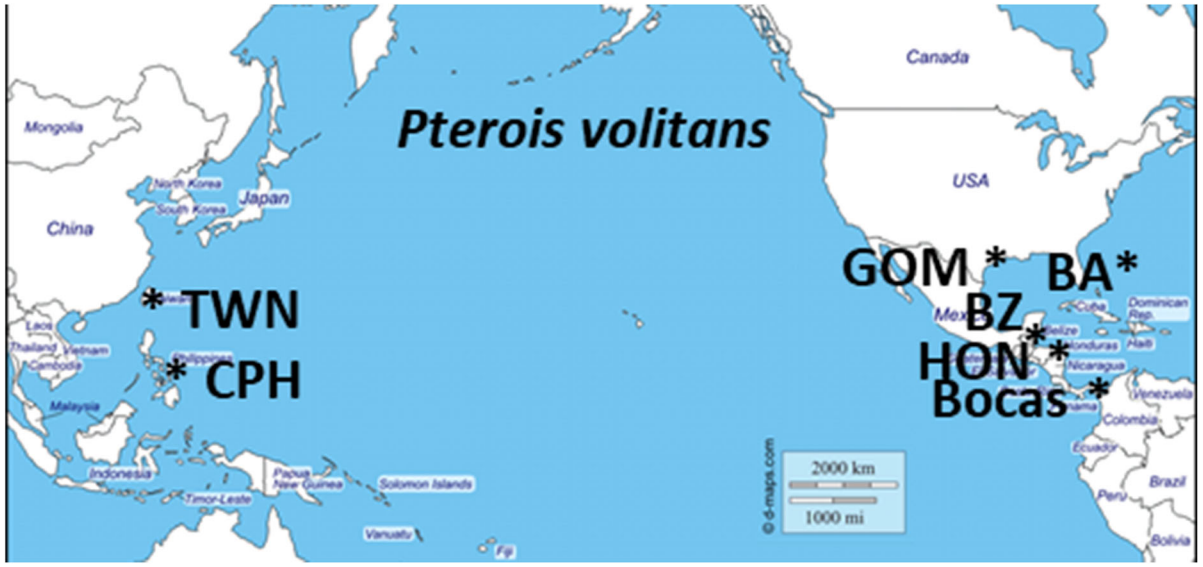

Fig. 1 Geographic locations of *P. volitans* Lionfish samples from the native and introduced ranges (2 and 5 locations, respectively). Not shown is the *P. miles* sample, from Sri Lanka, and the *P. antennata* sample, from the Yale museum. Here *TWN* Taiwan, *CPH* Philippines, *GOM* Gulf of Mexico, *BA* Bahamas, *BZ* Belize, *HON* Honduras, *Bocas* Panama

Table 1 Samples of lionfish from within and outside of its native range

Species	Code	Location	Number
P. volitans	BA	Bahamas	8
P. volitans	GOM	Gulf of Mexico	13
P. volitans	BZ	Belize	18
P. volitans	HON	Honduras	11
P. volitans	Bocas	Panama	12
P. volitans	**CPH**	**Philippines**	**9**
P. volitans	**TWN**	**Taiwan**	**8**
P. miles	**SL**	**Sri Lanka**	**2**
P. antennata	YL	Yale Museum	2

Those sites that are within the native range are bolded for all species

range. We preserved fin clips in EtOH and stored them in a 4 °C refrigerator.

DNA extraction, sequencing, and genomic library building

We extracted genomic DNA from fin clips using a Qiagen DNA Extraction Kit (Qiagen Inc., Valencia, CA, USA) following the manufacturer protocol. We quantified DNA concentrations using a Qubit fluorometer (Qubit 2.0; Invitrogen, Carlsbad, CA USA). We built ddRADseq libraries using the enzyme pairs *SphI* and *MluCI* and following the protocol and method outlined in Burford Reiskind et al. (2016). We constructed two libraries, containing 83 individual lionfish samples. We conducted single-end sequencing of 100 bp fragments of the libraries on the Illumina HiSeq2500 following specifications at North Carolina State University Genomic Sequencing Laboratory (Raleigh, NC, USA). In addition, we sequenced a subsample of the collected individuals to confirm genetic identification of individuals in comparison to previous phylogenetic work on the lionfish (Kochzius et al. 2003). We subsampled 35 individuals of the 83 using primer pairs for two regions. First, we sequenced control region (broadly and in the Dloop portion of the control region) (Meyer et al. 1994; Palumbi 1996; Freshwater et al. 2009; Betancur-R et al. 2011) following PCR protocol outlined in Butterfield et al. (2015) and Freshwater et al. (2009). Second, we sequenced cytochrome *b* (Kocher et al. 1989; Meyer et al. 1990; Hamner et al. 2007) following the PCR protocol outlined in Kochzius et al. (2003). We Sanger sequenced these individuals using an Applied BioSystems 3730xl Capillary sequencer following the specifications at the North Carolina State University

Genomic Sequencing Laboratory (Raleigh, NC, USA).

Sequence analysis

To confirm species identification, we sequenced a 680 bp region of the mtDNA Control Region using LionA-H and LionB-L primers (Freshwater et al. 2009) and for two individuals, one from Sri Lanka (*P. miles*) and one from Taiwan (*P. volitans*), we used the primers LPROF (Meyer et al. 1994) and 12SAAR-H (Palumbi 1996). We also amplified a 421 bp region the mtDNA cytochrome *b* gene using L14724 (Meyer et al. 1990) and H5149 (Kocher et al. 1989). We sequenced *P. volitans* individuals from Taiwan (n = 8), Philippines (n = 8), Bahamas (n = 8), Panama (n = 2), Belize (n = 2), Gulf of Mexico (n = 2), and Honduras (n = 2). We also sequenced *P. miles* individuals from Sri Lanka (n = 2). We aligned sequences using Geneious v.10.2.3 (http://www.geneious.com, Kearse et al. 2012), and compared those sequences to submitted sequences on GenBank using the forward primer, reverse primer, consensus sequence of forward and reverse, and with known sequences of *P. volitans*.

Bioinformatic pipeline

The Illumina platform automatically de-multiplexed the two indices of our ddRADseq libraries into separate FASTQ files. We used FASTQC (Babraham Bioinformatics; http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) to check the quality of the reads, using a high phred score criterion, prior to processing the barcodes as outlined in Burford Reiskind et al. (2016). We then ran the *process_radtags* script to filter and de-multiplex our variable length barcodes in STACKS v.1.24 (Catchen et al. 2011). We trimmed the reads to 90 basepairs, to make all read lengths identical in length as required by the STACKS platform. For SNP detection, we ran the denovo pipeline (*denovo.pl*) available in STACKS. We ran all runs through the denovo pipeline with the following parameters: $m = 3$ (minimum stack depth), $M = 2$ (mismatches allowed between loci within an individual), and $n = 2$ (mismatches allowed between loci when combining them in a catalog) (Catchen et al. 2011). We then used population pipeline (*populations*) in STACKS with parameters as follows: minimum number of stacks per individual at a locus ($m = 5$), number of populations loci present in ($p = 2$), proportion of individuals within a population that have these loci ($r = 0.75$), and appropriate output files for downstream analyses (following Burford Reiskind et al. 2016). While STACKS pipeline provides the possibility to create datasets in various formats, we used the PLINK format (PLINK v1.19 http://pngu.mgh.harvard.edu/purcell/plink/) as it is considered versatile for large NextGen sequenced data. We used the program PGDSpider v.2.1.1.0 (Lischer and Excoffier 2012; http://www.cmpg.unibe.ch/software/PGDSpider/) to transform the PLINK dataset in various input file formats required by the software we utilized: GENEPOP v.4.2 (Rousset 2008), LOSITAN (Antao et al. 2008), STRUCTURE v.2.3.4 (Pritchard et al. 2000; Hubisz et al. 2009), and DAPC analysis implemented in ADEGENET (Jombart 2008). We filtered the data set for minimum allele frequencies (MAF) in PLINK (–maf 0.01), removing all loci with a minimum allele frequency of 0.01 or lower and removed loci with more than 25% missing data (–geno 0.25) which would interfere in some of the downstream analyses. We measured genetic diversity (H_E), inbreeding coefficient (F_{IS}), and genetic differentiation (pairwise F_{ST} and pairwise exact test (MCMC parameters: 20 000 dememorization, 500 batches, 10 000 iterations per batch)) in GENEPOP. We ran STRUCTURE with a 10 000 burnins, 10 000 MCMC replicates, a K ranging from 1 to 7, and with 10 iterations per K for all locations, all locations except for outgroups (Yale and Sri Lanka), and for the invasive range locations of *P. volitans*, using a random number seed. We used Structure Harvester (Earl and vonHoldt 2012) to determine the likelihood of number of clusters and significance among sample locations. We also ran STRUCTURE using similar parameters for the entire data set, which included the two samples of *P. miles* and *P. antennata* from Sri Lanka and from the Yale Museum, respectively. In addition to STRUCTURE, we conducted a Discriminant Analysis of Principal Components (DAPC) in R using the package ADEGENET on the same groups of populations to further investigate genetic differentiation. We first cross-validated the data sets using 95 replicates to determine the optimal number of principal components (PCs) to evaluate genetic structure, to avoid overfitting these data.

To compare the native and invasive range for genomic regions of putative directional selection, we

conducted two outlier analyses between the most closely related locations in the native and introduced ranges. Using closely related locations for comparison eliminates a greater number of false positive outlier loci that are characteristic of geographic distance or drift and not from selection (see Burford Reiskind et al. 2018). In this analysis, we identified Taiwan (native range) and Bahamas (closely related to Taiwan) for outlier loci comparison. We looked for evidence of selective sweeps from an outlier analysis using LOSITAN with 10 reps of 1 000 000 simulations, which we did manually, and a DAPC analysis. To determine the number of outlier loci using DAPC, we first cross-validated the data set using 95 replicates and used the average linkage clustering method to set a threshold for outlier loci. We applied false discovery rate (FDR) correction factor of the p value of 0.05 based on reported high false positive loci rates in LOSITAN's main algorithm FDIST2 (Beaumont and Nichols 1996). For outlier loci, we first confirmed the number of outliers found for both analysis program, and then we compared both programs for loci that overlapped. For the outlier loci, we uploaded the FASTA file containing those reads to the NCBI GenBank and used the BLASTn tool to find similar sequences in the database. We used the default BLASTn option, which looks for highly similar sequences and selected those sequences that had a positive hit with a high percentage of similarity (Table S2).

Results

Mitochondrial sequence data

We amplified 461 bps of cytochrome *b* and 680 bps of the Dloop part of the control region for 26 of the 35 samples. The results suggested that most individual's sequences aligned with *P. volitans*, with some notable exceptions (Table S1). The two samples of *P. miles* from Sri Lanka (LF_SL8 and LF_SL9) aligned to *P. volitans* at control region and with *P. mombasae* at cytochrome *b* (only 87% sequence identity with *P. volitans* and *P. miles*) suggesting these species share SNPs in the sequence regions in their mtDNA (Table S1). One individual from Taiwan that was morphologically identified as *P. volitans* also carried both *P. mombasae* characteristic SNPs at cytochrome *b*, and *P. volitans* at the control region (Table S1).

Nuclear genome data

After filtering the raw data output from STACKs for the 83 individuals, we found 57,742 loci, which reduced to 52,742 loci after we filtered out loci with allele frequencies lower than 0.01 (PLINK maf = 0.01). After we ran the population pipeline and excluded *P. miles* (Sri Lanka) and *P. antennata* (Yale) samples, there were 79 putative *P. volitans* individuals with 44,190 loci post PLINK filtering of monomorphic loci. For the invasive range only, there were 62 individuals with 18,240 polymorphic loci after filtering out loci with allele frequencies lower than 0.01.

Genetic diversity

Genetic diversity was similar between invasive and native populations of *P. volitans*, with the exception of Taiwan, which had almost two times higher H_E than samples from invasive range. Of the invasive populations, the Bahamas had slightly higher diversity than the other samples in this range. Estimates of inbreeding based on F_{IS} were higher in the native populations, especially in the Taiwan individuals (F_{IS}= 0.47). All invasive populations had relatively low F_{IS}, despite the likelihood of founder effects. Both Sri Lanka *P. miles* individuals and *P. antennata* showed low diversity and high F_{IS}, which was inconclusive given the low sample size (n = 2; Table 2).

Genetic differentiation and structure

When we compared *P. volitans* individuals from the introduced and native range using pairwise F_{ST} and the exact test in GENEPOP, we found significant pairwise differences, except for the Bahamas which was not significantly different from Taiwan (F_{ST}= 0.04). We found no evidence of genetic differentiation among the populations of the introduced range. However, we found significant genetic differentiation between the two native populations (Taiwan and Philippines, F_{ST}= 0.47; Table 3), such that Taiwan was less genetically divergent from population in the introduced range than it was to *P. volitans* samples in the Philippines.

Table 2 Inbreeding coefficient F_{IS} for the 9 lionfish populations

Species	Population	Observed heterozygosity	Expected heterozygosity	F_{IS}
P. volitans	Belize (BZ)	0.077	0.079	0.028
P. volitans	Bahamas (BA)	0.078	0.081	0.041
P. volitans	Gulf of Mexico (GOM)	0.074	0.078	0.052
P. volitans	Panama (Bocas)	0.075	0.079	0.052
P. volitans	Honduras (HON)	0.073	0.080	0.080
P. volitans	**Philippines (CPH)**	**0.089**	**0.102**	**0.122**
P. volitans	**Taiwan (TWN)**	**0.074**	**0.141**	**0.478**
P. miles	**Sri Lanka (SL)**	**0.040**	**0.055**	**0.270**
P. antennata	Yale Library	0.037	0.0419	0.117

Those sites that are within the native range are bolded for all species

Table 3 Pairwise F_{ST} for the 9 lionfish populations

Code	BZ	BA	GOM	Bocas	HON	CPH	TWN	SL
Bahamas	0.006							
G. Mexico	0.006	0.000						
Panama	0.001	0.001	0.002					
Honduras	0.004	0.003	0.004	0.000				
Philippines	**0.588***	**0.570***	**0.584***	**0.581***	**0.580***			
Taiwan	**0.081***	0.048	**0.073***	**0.065***	**0.065***	**0.479***		
Sri Lanka (*P. miles*)	**0.869***	**0.868***	**0.870***	**0.869***	**0.868***	**0.848***	**0.734***	
Yale Library (*P. antennata*)	**0.861***	**0.866***	**0.864***	**0.864***	**0.867***	**0.848***	**0.748***	**0.825***

We include measures between species (*P. volitans* vs. *P. miles* and *P. antennata*) for reference. * represents a significant difference according to the pairwise exact test (all significant population pairs were denoted as "highly significant"). Colors correspond to the STRUCTURE population clusters identified in Fig. 2

Among the eight sampled populations including *P. miles* from Sri Lanka, *P. volitans* from Taiwan and the Philippines, and all locations from the introduced range (excluding *P. antennata*), results of STRUCTURE harvester reported three distinct clusters ($\Delta K = 3$). The STRUCTURE results showed that the two Sri Lanka individuals, putatively identified as *P. miles*, and one individual from Taiwan belonged to the same cluster. The Philippines was a separate cluster and Taiwan clustered with the invaded range, except for the one individual described above (Fig. 2).

When we removed *P. miles* and reran STRUCTURE with only the *P. volitans*, the results from STRUCTURE harvester reported two clusters ($\Delta K = 2$). As in the eight-population analysis, the Philippines population was genetically distinct, and the introduced range was genetically homogenous. STRUCTURE assigned the Taiwan individual that had previously assigned to the Sri Lanka population to the Philippines cluster, and the remaining Taiwan individuals grouped with the invasive range (Fig. 3a). The DAPC cluster analysis supported this pattern using 10 PCs (mean success = 0.618, RMSE = 0.396). The pattern reported in the DAPC analysis showed the Philippines population as the most distinct from all other locations, with Taiwan also distinct but much closer to the five introduced populations, which all clustered together (Fig. 3b).

The results from STRUCTURE harvester reported three clusters ($\Delta K = 3$) for the introduced range alone. It assigned all individuals to 2–3 clusters, with all three appearing in each sampling location (Supplemental

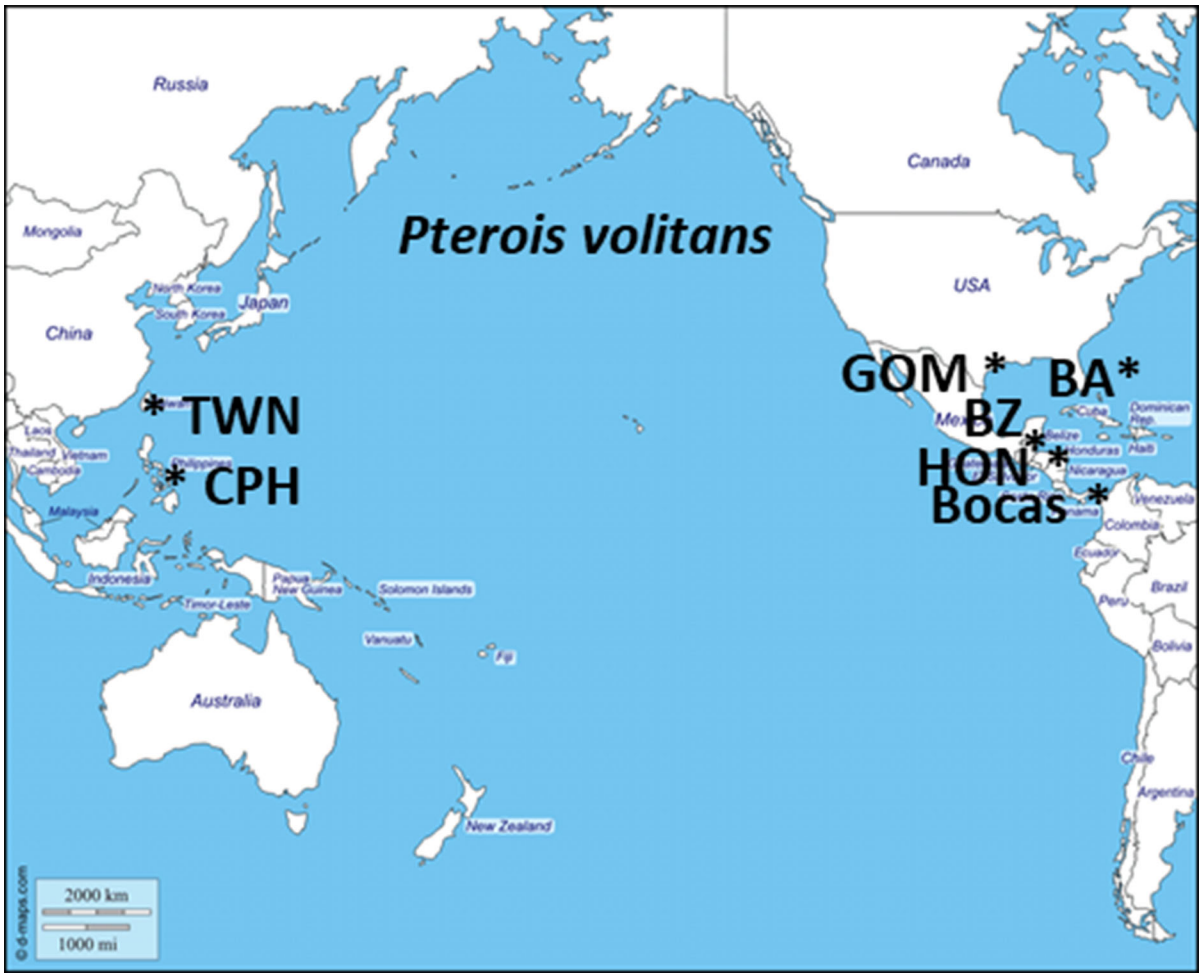

Fig. 2 Estimated population structure from STRUCTURE analysis for eight lionfish populations (excluding *P. antennata*) with a $K = 3$ using 52,742 polymorphic loci. Each color corresponds to a genetic cluster and each individual is represented by a vertical bar. Structure harvester supports a $K = 3$ Here *BZ* Belize, *BA* Bahamas, *GOM* Gulf of Mexico, *Bocas* Panama, *HON* Honduras, *CPH* Philippines, *TWN*Taiwan, and *SL* Sri Lanka. BZ, BA, GOM, Bocas, and HON are from the Lionfish invasive range, CPH and TWN are Lionfish from the native range identified in the field as *P. volitans,* and SR are Lionfish from the native range identified as *P. miles*. All invasive populations and TWN (except for one individual) were assigned to one cluster, SL and one TWN individual were assigned to a second, and CPH individuals comprised the third

Fig S1a). The DAPC cluster analysis using 40 PCs (mean success = 0.454, RMSE = 0.571) showed genetic structure among the remaining populations, with both the Gulf of Mexico and Belize populations distinct from the other two (Panama and Honduras; Supplemental Fig S1b). The limited number of PCs retained in this method prevented overfitting these data. We did not find any signature of isolation by distance (Pairwise $F_{ST}/(1 - F_{ST})$ versus Euclidian distance and measured distance using dominant water current direction, Mantel test $P = 0.642$ and $P = 0.758$ respectively, data not shown).

Outlier loci analysis

Comparing putative *P. volitans* from the Taiwan and Bahamas populations, we identified 1,866 outlier loci in LOSITAN, and 71 outlier loci across 67 sequences using DAPC in the ADEGENET R package. We also found that all 67 sequences found using DAPC were also outlier loci in LOSITAN. Using BLASTn, all 67 loci mapped to genomes of other fish species with varying degrees of similarity (Table S2), many mapped unannotated regions of chromosomes, and a few mapped to protein-coding regions. For example,

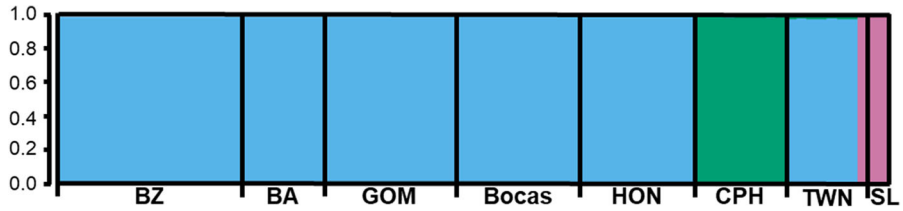

Fig. 3 Graphical representation of the seven putative *P. volitans* populations using 44,190 polymorphic loci, where the first five populations are from the invasive range (*BZ* Belize, *BA* Bahamas, *GOM* Gulf of Mexico, *Bocas* Panama, *HON* Honduras) and the last two are from the native range (*CPH* Philippines, *TWN* Taiwan). **a** Visual output from STRUCTURE analysis with a $K = 2$ and $K = 3$. Structure harvester supported a $K = 2$, and we included the representation for three clusters to highlight the putatively misidentified individual from Taiwan (TWN). All invasive populations and TWN except for the one individual were assigned to the same cluster. CPH individuals were all assigned to a second cluster, to which the TWN individual was also assigned when $K = 2$. When $K = 3$, the one TWN individual was unique from all others. **b** Visual output using Discriminant analysis of principal components (DAPC) analysis with 10 principal components, chosen to avoid overfitting the data, and cross-validation of 95 iterations. The DAPC analysis shows that CPH is genetically distinct from the invasive and TWN populations

one outlier locus mapped to the adenosine receptor A1 (*adoral*) in multiple fish species, which researchers found expressed differentially in fish at different depths (Murray and Siebenaller 1987; Wakisaka et al. 2017). In addition, this gene is involved in anxiety, boldness, and behavior in zebra fish (Maximino et al. 2011). We also found an outlier locus near the Gnrh2 gene, which is involved in reproduction, brain function, and suppression of food intake in fish (Feng et al. 2018; Nishiguchi et al. 2012, respectively).

Discussion

While many studies have found evidence of hybridization between native and introduced species in invaded ranges (Nolte et al. 2005; Muhlfeld et al. 2009; Schierenbeck and Ellstrand 2009), we did not find evidence for hybridization between *P. miles* and *P. volitans* or between *P. volitans* and any other putative lionfish species in the introduced range. However, we did find evidence of a mismatch between the mtDNA and nuclear DNA within individuals in the native range that suggest the potential for a hybrid species to form in the native range and then invade. The hybridization within the native range prior to introduction could help explain the high population growth and reproduction rates observed in the introduced range (Albins 2015; Layman and Allgeier 2012; Jud and Layman 2012; Morris and Akins 2009; Ahrenholz and Morris 2010, but see Benkwitt et al. 2017). Previous studies on invasive species suggest that successful introductions may stem from admixture of multiple introductions (Kolbe et al. 2004; Frankham 2005), but our results support a single region for the source and a single introduction to the invaded range. While our sampling is limited in the native region and cannot therefore point to the exact source populations, this result suggests rapid spread and likely evolution of this species in the introduced range. We found significant genetic structure both within the native range and between the native and introduced ranges of *P. volitans*, yet within the native range, we found genetic divergences between putative *P. volitans* samples commensurate with distinct species. Finally, we found outlier loci between Taiwan and the Bahamas suggesting putative genomic regions of selection that may contain genes that support the invasion. Collectively, the results of this study suggest that a genomic approach will provide an important contribution to understanding how this invasive species has been so successful and help further clarify potential interspecific hybridization in the native range.

This study confirmed the previous findings that the southern coastline of North America was likely the initial source for the spread of the *P. volitans* throughout the introduced range. While the exact location of introduction may not be the Bahamas, we hypothesize that it is nearby and/or receives migrants from it. This was supported by the slightly higher genetic diversity found in this location and the lack of genetic differentiation between this location and the native range (i.e. Taiwan). In addition, the analysis

from DAPC showed that the Bahamas population clusters separately from the rest of the samples in the introduced range and suggests a more recent invasion at the location outside of the Bahamas. The remaining locations were relatively genetically homogenous, indicating high levels of gene flow, which matches the structure results from this study and findings by Pérez-Portela et al. (2018). Patterns of genetic diversity and population relatedness also suggest that the spread from the initial population likely moved in a southward direction from the northern Gulf of Mexico to the southern Caribbean. Cowen et al's (2006) definitions of connectivity barriers in the region supported our finding of low-level genetic structure between and within the northwestern and southwestern regions of our sampling in the invasive range. While passive dispersal limits connectivity among regions in the invasive range (Cowen et al. 2006), the length of the pelagic duration of the lionfish may increase connectivity. In addition, aprevious study by Butterfield et al. (2015) using microsatellite loci also found both lower genetic diversity and a significant break in genetic structure between northern samples (Bahamas, North Carolina and Bermuda) compared to sample locations throughout the Caribbean. We also found support for this specific genetic break between the southwestern Atlantic and the Gulf of Mexico/Caribbean regions at least with one of our analyses, and we did so with a small number of individuals per location, supporting the increased power of the genomic approach.

While we did not find any pure *P. miles* or *P. miles*/*P. volitans* hybrids in the introduced range, we did find support for hybridization and/or misclassification *P. miles* or *P. volitans* individuals in the native range. While our sampling of *P. miles* was limited, these results suggest that species classification based on morphological characters is not clear and may support a hybrid origin for *P. volitans*. For example, the degree of genetic divergence between the *P. volitans* from the Philippines and Taiwan was greater than that between Taiwan and populations in the introduced range, and close to what we would expect across species boundaries. In addition, one individual from Taiwan (TWN_10_12) that did not group with other Taiwan individuals in the cluster analysis and carried *P. volitans* sequence at the control region gene, suggesting this was a hybrid between another species and a female *P. volitans*. While Wilcox et al. (2017) hypothesized *P. volitans* resulted from a hybridization event between *P. miles* from the Indian Ocean and *P. lunulata/russelii* from the Pacific Ocean, it is unclear which two species produced the hybrid here given what we found. Wilcox et al. (2017) reported that *P. russelli* dominated the sample collection from this region, but we cannot confirm whether this individual was *P. russelli*, only that it was not *P. volitans* or *P. miles* in the nuclear DNA, only potentially *P. volitans* in the mtDNA. Our genomic approach could elucidate the pattern of hybridization for *P. volitans* throughout its native range and further reveal whether part of the success of *P. volitans* in introduced regions is because of hybrid vigor. While our study identifies Taiwan as the source population for the *P. volitans* invasion, further analysis throughout this region, with a larger sample size, and among the native populations across large geographic distances in the Indo-Pacific will determine whether this exact location or one nearby was the source.

Comparing the populations of Taiwan and Bahamas provided the first indication of putative genomic regions for directional selection between the native and introduced range. Without an annotated genome for the lionfish, we cannot identify the exact genes, nor can we assign the cause given the potential bottleneck between source and introduction, but we have identified candidate regions for further investigation. Moreover, genes identified through comparisons to other fish genomes showed a variety of functions that suggest these genes may important difference between individuals from the native and introduced range. These genes have a variety of functions including boldness and anxiety behaviors, reproduction, and food intake suppression in fish. Further sampling between the most closely related population in the introduced and native range will help to clarify which of these genomic regions has allowed the lionfish to be a successful invader.

Given the low sampling of individuals and locations, we are encouraged that our approach provided information on genetic differentiation within and among the two regions and are confident that this method will further resolve the pattern of invasion in the introduced range for this and other invasive species. Moreover, with additional sampling of the native range, the regions near the first introduction, and locations at the boundaries of the introduced range, the outlier loci approach could identify genomic regions under directional selection that may

clarify not only how this species evolved after introduction, but also provide information on the potential for divergent evolution in the introduced range, as previous research has attempted to do with similar genomic approaches in other marine fish species (Bernardi et al. 2017; Gaither et al. 2015). Demographic data suggest that in the introduced range this species has multiple broods a year and outcompetes many native reef fishes by either interference or exploitative competition. In the future, outlier loci may reveal genomic changes related to these important phenotypic differences in reproductive output between the two. In addition, the ability of *P. volitans* to invade to the northern regions of the western Atlantic may also show genomic regions related to cold tolerance and habitat differences that are different from that in the native range. To manage this and other invasive species effectively requires a better understanding of not only the pattern of introduction and spread, but also the mechanism that allows these species to become invasive.

Acknowledgements We thank M.R. Reiskind and R.R. Dunn for comments on an early version of this manuscript and several anonymous reviewers for their comments that improved this manuscript. The Wynne Innovation Grant from the CALS Dean's Enrichment Grant program at NCSU awarded to MO Burford Reiskind funded this research. We also thank both the Applied Ecology and Entomology and Plant Pathology Departments at NCSU for providing matching funds to further fund this collaborative research. This study was a project in MO Burford Reiskind's Conservation Genetics course at NCSU for undergraduate and graduate students.

Data Accessibility We provided supplemental material with the original submission for the online version of this manuscript. This will include the results of the Genbank blast search for outlier loci. Data from this manuscript will be available through Dr. Martha Burford Reiskind's DRYAD account. This will include post-STACK analysis input data files for PGDSpider from which subsequent input data files can be generated. The complete list of aligned sequences containing outlier loci. In addition, the raw sequence data generated in this study will be available upon request, as the data files far exceed the limits at DRYAD without addition payment.

References

Ahrenholz DW, Morris JA Jr (2010) Larval duration of the lionfish, *Pterois volitans* along the Bahamian Archipelago. Environ Biol Fish 88:305–309. https://doi.org/10.1007/s10641-010-9647-4

Albins MA (2015) Invasive Pacific lionfish *Pterois volitans* reduce abundance and species richness of native Bahamian coral-reef fishes. Mar Ecol Prog Ser 522:231–243. https://doi.org/10.3354/meps11159

Albins MA, Hixon MA (2008) Invasive Indo–Pacific lionfish *Pterois volitans* reduce recruitment of Atlantic coral-reef fishes. Mar Ecol Prog Ser 367:233–238. https://doi.org/10.3354/meps07620

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

Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. Nat Rev Genet 11:697–709. https://doi.org/10.1038/nrg2844

Allendorf FW, Luikart G, Aitkin SN (2013) Conservation and the genetics of populations, 2nd edn. Blackwell, London

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 Bioinform 9:323. https://doi.org/10.1186/1471-2105-9-323

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. Proc R Soc Lond B Biol Sci 263:1619–1626. https://doi.org/10.1098/rspb.1996.0237

Benkwitt CE, Albins MA, Buch KL, Ingeman KE, Kindinger TL, Pusack TJ, Stallings CD, Hixon MA (2017) Is the lionfish invasion waning? Evidence from The Bahamas. Coral Reefs 36:1255–1261. https://doi.org/10.1007/s00338-017-1620-7

Bernardi G, Azzurro E, Golani D, Miller MR (2016) Genomic signatures of rapid adaptive evolution in the bluespotted cornetfish, a Mediterranean Lessepsian invader. Mol Ecol 25:3384–3396. https://doi.org/10.1111/mec.13682

Betancu-R R, Hines A, Acero AP, Orti G, Wilbur AE, Freshwater DW (2011) Reconstructiong the lionfish invasion: insights into Greater Caribbean biogeography. J Biogeogr 38:11281–11293. https://doi.org/10.1111/j.1365-2699.2011.02496.x

Bouzat J (2010) Conservation genetics of population bottlenecks: the role of chance, selection, and history. Conserv Genet 11:463–478. https://doi.org/10.1007/s10592-010-0049-0

Burford Reiskind MO, Coyle K, Daniels HV et al (2016) Development of a universal double-digest RAD sequencing approach for a group of nonmodel, ecologically and economically important insect and fish taxa. Mol Ecol Resour 16:1303–1314. https://doi.org/10.1111/1755-0998.12527

Burford Reiskind MO, Labadie P, Bargielowski I, Lounibos LP, Reiskind MH (2018) Rapid evolution and the genomic consequences of selection against interspecific mating. Mol Ecol. https://doi.org/10.1111/mec.14821

Butterfield JSS, Díaz-Ferguson E, Silliman BR et al (2015) Wide-ranging phylogeographic structure of invasive red lionfish in the Western Atlantic and Greater Caribbean. Mar Biol 162:773–781. https://doi.org/10.1007/s00227-015-2623-y

Catchen JM, Amores A, Hohenlohe P et al (2011) Stacks: building and genotyping Loci De Novo from short-read sequences. G3 Genes Genom Genet 1:171–182. https://doi.org/10.1534/g3.111.000240

Courtenay WR (1995) Marine fish introductions in southeastern Florida. Am Fish Soc Introd Fish Newsl 14:2–3. https://doi.org/10.3391/ai.2009.4.3.5

Cowen RK, Paris CB, Srinivasan A (2006) Scaling of connectivity in marine populations. Science 311:522–527. https://doi.org/10.1126/science.1122039

Cure K, McIlwain JL, Hixon MA (2014) Habitat plasticity in native Pacific red lionfish *Pterois volitans* facilitates successful invasion of the Atlantic. Mar Ecol Prog Ser 506:243–253. https://doi.org/10.3354/meps10789

Earl DA, Vonholdt BA (2012) Structure Harvester: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361. https://doi.org/10.1007/s12686-011-9548-7

Feng K, Luo H, Hou M, Li Y, Chen J, Zhu Z, Hu W (2018) Alternative splicing of GnRH2 and GnRH2-associated peptide plays roles in gonadal differentiation in the rice field eel, *Monopterus albus*. Gen Comp Endocrinol. https://doi.org/10.1016/j.ygcen.2018.05.021

Ferreira CEL, Luiz OJ, Floeter SR, Lucena MB, Barbosa MC, Rocha CR, Rocha LA (2015) First record of invasive lionfish (*Pterois volitans*) for the Brazilian coast. PLoS ONE 10:1–5. https://doi.org/10.1371/journal.pone.0123002

Fitzpatrick BM, Shaffer HB (2007) Hybrid vigor between native and introduced salamanders raises new challenges for conservation. Proc Natl Acad Sci 104:15793–15798. https://doi.org/10.1073/pnas.0704791104

Frankham R (2005) Resolving the genetic paradox in invasive species. Heredity 94:385

Freshwater DW, Hines A, Parham S, Wilbur A, Sabaoun M, Woodhead J, Akins L, Purdy B, Whitfield PA, Paris CB (2009) Mitochondrial control region sequence analyses indicate dispersal from the US East Coast as the source of the invasive Indo-Pacific lionfish *Pterois volitans* in the Bahamas. Mar Biol 156:1213–1221. https://doi.org/10.1007/s00227-009-1163-8

Gaither MR, Bernal MA, Coleman RR, Bowen BW, Jones SA, Simison WB, Rocha LA (2015) Genomic signatures of geographic isolation and natural selection in coral reef fishes. Mol Ecol 24:1543–1557. https://doi.org/10.1111/mec.13129

Grosholz ED, Ruiz GM, Dean CA, Shirley KA, Maron JL, Connors PG (2000) The impacts of a nonindigenous marine predator in a California bay. Ecology 81:1206–1224. https://doi.org/10.1890/0012-9658(2000)081%5b1206:TIOANM%5d2.0.CO;2

Hamner RM, Freshwater DW, Whitfield PE (2007) Mitochondrial cytochrome *b* analysis reveals two invasive lionfish species with strong founder effects in the western Atlantic. J Fish Biol 71:214–222. https://doi.org/10.1111/j.1095-8649.2007.01575.x

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. Mol Ecol Resour 9:1322–1332. https://doi.org/10.1111/j.1755-0998.2009.02591.x

Jombart T (2008) Adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jud ZR, Layman CA (2012) Site fidelity and movement patterns of invasive lionfish, *Pterois* spp., in a Florida estuary. J Exp Mar Biol Ecol 414–415:69–74. https://doi.org/10.1016/j.jembe.2012.01.015

Kearse M, Moir R, Wilson A et al (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649

Kocher TD, Thomas WK, Meyer A et al (1989) Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. Proc Natl Acad Sci 86:6196–6200. https://doi.org/10.1073/pnas.86.16.6196

Kochzius M, Soller R, Khalaf MA, Blohm D (2003) Molecular phylogeny of the lionfish genera *Dendrochirus* and *Pterois* (Scorpeanidae, Pteroinae) based on mitochondrial DNA sequences. Mol Phylogenetics Evol 28:396–403

Kolbe JJ, Glor RE, Schettino LR, Lara AC, Larson A, Losos JB (2004) Genetic variation increases during biological invasion by Cuban lizard. Nature 431:177–181

Kulbicki M, Beets J, Chabanet P, Cure K, Darling E, Floeter SR, Galzin R, Green A, Harmelin-Vivien M, Hixon M, Letourneur Y, de Loma TL, McClanahan T, McIlwain J, MouTham G, Myers R, O'Leary JK, Planes S, Vigliola L, Wantiez L (2012) Distributions of Indo–Pacific lionfishes *Pterois* spp. in their native ranges: implications for the Atlantic invasion. Mar Ecol Prog Ser 446:189–205. https://doi.org/10.3354/meps09442

Layman CA, Allgeier JE (2012) Characterizing trophic ecology of generalist consumers: a case study of the invasive lionfish in The Bahamas. Mar Ecol Prog Ser 448:131–141. https://doi.org/10.3354/meps09511

Lee CE (2002) Evolutionary genetics of invasive species. Trends Ecol Evol 17:386–391

Lischer HEL, Excoffier L (2012) PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. Bioinformatics 28:298–299. https://doi.org/10.1093/bioinformatics/btr642

Maximino C, Lima MG, Olivera KR et al (2011) Adenosine A1 but not A2, receptor blockade increases anxiety and arousal in zebrafish. Basic Clin Pharmacol Toxicol 109:203–207

Meyer A, Kocher TD, Basasibwaki P, Wilson AC (1990) Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. Nature 347:550–553

Meyer A, Morrissey JM, Schartl M (1994) Recurrent origin of a sexually selected trait in *Xiphophorus* fishes inferred from a molecular phylogeny. Nature 368:539–542

Morris JA, Akins JL (2009) Feeding ecology of invasive lionfish (*Pterois volitans*) in the Bahamian archipelago. Environ Biol Fish 86:389. https://doi.org/10.1007/s10641-009-9538-8

Muhlfeld CC, Kalinowski ST, McMahon TE, Taper ML et al (2009) Hybridization rapidly reduces fitness of a native trout in the wild. Biol Lett 5:328–331

Murray TF, Siebenaller JF (1987) Comparison of the binding properties of A1 adenosine receptors in brain membranes of two congeneric marine fishes living at different depths. J Comp Physiol B Biochem Syst Environ Physiol 157:267–277. https://doi.org/10.1007/BF00693353

Nishiguchi R, Azuma M, Yokobori E et al (2012) Gonadotropin-releasing hormone 2 suppresses food intake in the zerbrafish, *Danio rerio*. Front Endocrinol 3:1–6

Nolte AW, Freyhof J, Stemshorn KC, Tautz D (2005) An invasive lineage of sculpins, *Cottus* sp. (Pisces, Teleostei) in the Rhine with new habitat adaptations has originated from hybridization between old phylogeographic groups. Proc R Soc B 272:2379–2387

Palumbi SR (1996) Nucleic Acids II: the polymerase chain reaction. In: Hillis DM, Moritz C, Mable BK (eds) Molecular systematics, 2nd edn. Sinauer, Sunderland, pp 205–247

Pérez-Portela R, Buford A, Coffman B, Wedelich S, Davenport M, Fogg A, Swenarton MK, Coleman F, Johnston MA, Crawford DL, Oleksiak MF (2018) Genetic homogeneity of the invasive lionfish across the Northwestern Atlantic and the Gulf of Mexico based on single nucleotide polymorphisms. Sci Rep 8:5062. https://doi.org/10.1038/s41598-018-23339-w

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

Rousset F (2008) genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Mol Ecol Resour 8:103–106. https://doi.org/10.1111/j.1471-8286.2007.01931.x

Schierenbeck KA, Ellstrand NC (2009) Hybridization and the evolution of invasiveness in plants and other organisms. Biol Invasions 11:1093–1105

Schofield PJ (2009) Geographic extent and chronology of the invasion of non-native lionfish (*Pterois volitans* [Linnaeus 1758] and *P. miles* [Bennett 1828]) in the Western North Atlantic and Caribbean Sea. Aquat Invasions 4:473–479. https://doi.org/10.3391/ai.2009.4.3.5

Selwyn JD, Johnson JE, Downey-Wall AM, Bynum AM, Hamner RM, Hogan JD, Bird CE (2017) Simulations indicate that scores of lionfish (*Pterois volitans*) colonized the Atlantic Ocean. PeerJ 5:e3996. https://doi.org/10.7717/peerj.3996

Shiganova TA (1998) Invasion of the Black Sea by the Ctenophore *Mnemiopsis leidyi* and recent changes in pelagic community structure. Fish Oceanogr 7:305–310. https://doi.org/10.1046/j.1365-2419.1998.00080.x

Toledo-Hernandez C (2014) Population ecology and genetics of the invasive lionfish in Puerto Rico. Aquat Invasions 9:227–237. https://doi.org/10.3391/ai.2014.9.2.12

Wakisaka N, Miyasaka N, Koide T et al (2017) An adenosine receptor for olfaction in fish. Curr Biol 27:1437–1447.e4. https://doi.org/10.1016/j.cub.2017.04.014

Walker BH, Steffen W (1997) An overview of the implications of global change for natural and mangaged terrestrial ecosystems. Conserv Ecol 2(2). http://www.consecol.org/vol1/iss2/art2

Whitfield PE, Gardner T, Vives SP, Gilligan MR, Courtenay WR Jr, Ray GC, Hare JA (2002) Biological invasion of the Indo-Pacific lionfish *Pterois volitans* along the Atlantic coast of North America. Mar Ecol Prog Ser 235:289–297. https://doi.org/10.3354/meps235289

Wilcox CL, Motomura H, Matsunuma M, Bowen BW (2017) Phylogeography of lionfishes (*Pterois*) indicate taxonomic over splitting and hybrid origin of the invasive *Pterois volitans*. J Hered 109:162–175. https://doi.org/10.1093/jhered/esx056