Received: 20 April 2016 | Revised: 5 April 2017 | Accepted: 25 April 2017
DOI: 10.1111/mec.14169

WILEY MOLECULAR ECOLOGY

Genomic analysis of a cardinalfish with larval homing potential reveals genetic admixture in the Okinawa Islands

A. L. Gould | P. V. Dunlap

Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA

Correspondence
Alison L. Gould, Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA.
Email: algould@umich.edu

Abstract

Discrepancies between potential and observed dispersal distances of reef fish indicate the need for a better understanding of the influence of larval behaviour on recruitment and dispersal. Population genetic studies can provide insight on the degree to which populations are connected, and the development of restriction site-associated sequencing (RAD-Seq) methods has made such studies of nonmodel organisms more accessible. We applied double-digest RAD-Seq methods to test for population differentiation in the coral reef-dwelling cardinalfish, *Siphamia tubifer*, which based on behavioural studies, have the potential to use navigational cues to return to natal reefs. Analysis of 11,836 SNPs from fish collected at coral reefs in Okinawa, Japan, from eleven locations over 3 years reveals little genetic differentiation between groups of *S. tubifer* at spatial scales from 2 to 140 km and between years at one location: pairwise F_{ST} values were between 0.0116 and 0.0214. These results suggest that the Kuroshio Current largely influences larval dispersal in the region, and in contrast to expectations based on studies of other cardinalfishes, there is no evidence of population structure for *S. tubifer* at the spatial scales examined. However, analyses of outlier loci putatively under selection reveal patterns of temporal differentiation that indicate high population turnover and variable larval supply from divergent source populations between years. These findings highlight the need for more studies of fishes across various geographic regions that also examine temporal patterns of genetic differentiation to better understand the potential connections between early life-history traits and connectivity of reef fish populations.

KEYWORDS
bioluminescent, population genetics, RAD-Seq, *Siphamia tubifer*, SNPs

1 | INTRODUCTION

The degree to which coral reef fish populations are connected, which has large consequences for their persistence and resilience to disturbances, has been disputed for decades. For most reef fishes, dispersal occurs during a planktonic larval phase that lasts from a few days to months, creating the potential for large dispersal distances (Barlow, 1981; Shulman & Bermingham, 1995), but the challenges associated with tracking the direct movement of relatively small larvae in a dynamic, fluid environment have limited our knowledge of the connectivity patterns of most reef fishes (Jones et al., 2009). Recently, however, the rapid advancement of genetic tools has facilitated the study of reef fish populations, revealing the wide variation of spatial scales at which marine populations exhibit connectivity (e.g., Gerlach et al., 2007; Horne et al., 2008; Jones, Planes, & Thorrold, 2005; van der Meer et al., 2012; Planes, 2002; Purcell et al., 2006; Taylor & Hellberg, 2003; Terry, Bucciarelli, & Bernardi, 2000). Most coral reef fish metapopulations lie somewhere in the middle of the spectrum between having relatively closed (low connectivity and high self-recruitment) to open (high connectivity

 © 2017 John Wiley & Sons Ltd

and no self-recruitment) populations (Cowen et al., 2000; Jones et al., 2009; Mora & Sale, 2002). Nonetheless, a surprising number of studies have provided evidence that reef fish populations are more closed than expected and exhibit a significant degree of local recruitment (Almany et al., 2007; Bode, Bode, & Armsworth, 2006; Cowen, Paris, & Srinivasan, 2006; Cowen et al., 2000; Jones et al., 1999, 2005; Swearer et al., 1999, 2002) and genetic differentiation at relatively small spatial scales (e.g., Gerlach et al., 2007; Planes, 1993; Planes, Parroni, & Chauvet, 1998; Taylor & Hellberg, 2003).

In response to the growing evidence of somewhat restricted gene flow among reef fish populations, the links between early life-history traits, larval behaviour, and patterns of larval recruitment and dispersal have been examined. Contrary to expectations, levels of genetic differentiation of marine fishes are generally not correlated with pelagic larval duration (PLD) or spawning mode (Bowen et al., 2006; Galarza et al., 2009; Jones et al., 2009; Shulman, 1998), and genetic structure in reef fish populations is often observed at smaller scales than predicted by larval dispersal alone (e.g., Taylor & Hellberg, 2003; Rocha et al., 2005; but see Riginos, Buckley, Blomberg, & Treml, 2014). These discrepancies can often be explained by larval fish behaviour, through which larvae employ some control over their dispersal trajectories (Kingsford et al., 2002; Leis, Hay, & Trnski, 2006; Leis & McCormick, 2002; Montgomery et al., 2006; Sponaugle et al., 2002), including larval swimming ability (Fisher et al., 2005), vertical migration (Paris & Cowen, 2004), and the use of navigational cues, such as chemical and acoustic cues, in the environment (Atema, 2012; Dixson et al., 2008; Kingsford et al., 2002; Leis, Siebeck, & Dixson, 2011; Leis et al., 2003; Paris et al., 2013; Simpson et al., 2008). Studies of larval fish behaviour suggest that larvae are not simply passive particles in the plankton, but instead can actively orient and navigate to settlement sites. Therefore, incorporating larval behaviour into dispersal models can dramatically alter projected population connectivity patterns (e.g., Paris & Cowen, 2004; Cowen et al., 2006; Sale et al., 2005; Leis, 2007; Paris, Chérubin, & Cowen, 2007; Staaterman, Paris, & Helgers, 2012). Nevertheless, few studies have linked larval traits and behaviours to gene flow and levels of population genetic differentiation, and it remains unknown how conserved such links may be within fish families and geographic regions.

Cardinalfishes (family Apogonidae) are known for their homing ability from relatively large distances and for their fidelity to particular daytime resting sites (Døving et al., 2006; Gould, Harii, & Dunlap, 2014; Kolm et al., 2005; Marnane, 2000; Rueger, Gardiner, & Jones, 2014). There are only a few population genetic studies of cardinalfishes to date, all of which have indicated genetic differentiation at relatively small spatial scales (Bernardi & Vagelli, 2004; Gerlach et al., 2007; Gotoh et al., 2009; Hoffman et al., 2005; Vagelli, Burford, & Bernardi, 2008), although the majority of studies examined the same focal species, the Banggai cardinalfish *Pterapogon kauderni*, which lacks a planktonic larval stage (Bernardi & Vagelli, 2004; Hoffman et al., 2005; Vagelli et al., 2008). One study of the cardinalfish, *Ostorhinchus doederleini*, which has a PLD of 16–27 days, connected larval behaviour to self-recruitment and genetic differentiation between populations only a few kilometres apart. Specifically, the observed population genetic structure was linked to olfactory preferences exhibited by *O. doederleini* larvae for their home reef water over the other nearby reefs examined (Gerlach et al., 2007). Due to the lack of knowledge of other cardinalfishes, however, it remains unknown whether the results of this study are specific to the *O. doederleini* and whether other cardinalfishes with planktonic larvae also exhibit the same degree of fine-scale genetic structure associated with larval homing.

To better understand the link between larval behaviour and gene flow in cardinalfishes, we examined the population genetic structure of the sea urchin cardinalfish *Siphamia tubifer*, for which homing and olfactory preferences have been previously described (Gould, Harii, & Dunlap, 2015; Gould et al., 2014). Similar to other cardinalfishes, *S. tubifer* adults and juveniles exhibit fidelity to a home site and return to a home reef from displacement distances of at least 2 km (Gould et al., 2014). As described for *O. doederleini* (Gerlach et al., 2007), newly settled *S. tubifer* also exhibit a preference for the olfactory cues of their home reef to that of a foreign reef (Gould et al., 2015). These findings suggest the potential for *S. tubifer* larvae to use olfaction to recognize and return to their natal reef. Additionally, like *O. doederleini*, *S. tubifer* are short-lived (typically <200 days), but have a slightly longer PLD of up to 30 days (Gould, Dougan, Koeningbauer, & Dunlap, 2016; Kingsford et al., 2014). Despite these similarities, *S. tubifer* is distinct among cardinalfishes in that it is symbiotically bioluminescent, hosting a dense population of the luminous bacterium *Photobacterium mandapamensis* in an abdominal light organ and using the bacterially-emitted light while foraging at night (Dunlap & Nakamura, 2011; Iwai, 1958). The host fish acquires its luminous symbiont from the environment during larval development; however, the timing and location of symbiont acquisition in the wild remain unknown (Dunlap et al., 2012). Defining levels of genetic struture in *S. tubifer* at various geographic scales and over time can provide insight on population connectivity and the location of symbiont acquisition by developing larvae.

We applied double-digest restriction site-associated sequencing (ddRAD-Seq) methods to test for genetic differentiation among groups of *S. tubifer* collected over a 3-year period from various locations in Okinawa, Japan. We used ddRAD-Seq methods as they do not require an extensive marker discovery process and enable the development of thousands of genomic markers without prior genetic data for the focal species (Davey & Blaxter, 2010; Seeb et al., 2011), and they can be used to infer genetic differentiation with fine-scale precision (Bradbury et al., 2015; Coates et al., 2009; Luikart et al., 2003; Willing, Dreyer, & Van Oosterhout, 2012). Our specific aims were to test for patterns of population genetic differentiation between groups of *S. tubifer* collected at various spatial scales and to test the stability of *S. tubifer* populations by examining temporal patterns of genetic divergence at a particular reef. If the link between larval olfactory preferences, homing potential and self-recruitment are conserved across similar cardinalfish species, we predicted that similar to *O. doederleini*, populations of *S. tubifer* would

have significant population genetic structure between sampling sites but less genetic differentiation at the same reef over time.

2 | MATERIALS AND METHODS

2.1 | Sampling

A total of 295 *Siphamia tubifer* were collected from different locations (approximately 20 individuals per location) over 3 years in the Okinawa Islands, Japan. Ten of these locations were sampled during the summer of 2013, three of which were again sampled in 2014, and one location was sampled in three consecutive years, 2012, 2013 and 2014 (Figure 1). Fish of various sizes associated with several different host sea urchins and from a broad sampling area were collected at each location (Table 1). Upon collection, fish were immediately euthanized and placed on ice. The intact light organ of each fish was then aseptically dissected and individually preserved in RNAlater®, and the remainder of the fish specimen was stored in 98% ethanol at −20°C.

2.2 | DNA extraction and library preparation

Genomic DNA was extracted from intact, preserved light organs, which are comprised of fish tissue and the symbiotic population of luminous bacteria, using QIAGEN DNeasy Blood and Tissue Kits and following the manufacturer's protocol. A total of six ddRAD-Seq DNA libraries were constructed, each from the genomic DNA of up to 50 *S. tubifer* light organs, following a modified combination of the methods described in Parchman et al. (2012) and Peterson et al. (2012). For each library, approximately 200 ng of genomic DNA from each light organ was digested with the high-fidelity restriction enzymes *MseI* and *EcoRI* at 37°C for 3 hr. A standardized concentration per library of each digestion product was then ligated to a uniquely barcoded Illumina adaptor at the *EcoRI* cut site and an Illumina adaptor at the *MseI* cut site. The ligation products were individually amplified with the Illumina Illpcr1 and Illpcr2 primers in two 20 μl PCRs per sample with a 98°C start, 12 cycles of 98°C for 20 s, 65°C for 30 s, 72°C for 40 s, followed by 10 min at 72°C, after which, the PCR products from all samples within a library were pooled and concentrated to an approximate volume of 150 μl. Samples were purified with Agencourt AMPure XP magnetic beads following standard protocols after the digestion, ligation and PCR steps. The pooled, purified PCR products were then size-selected between 300 and 400 bp on a Pippin Prep (Sage Science) machine, and the size-selected DNA libraries were each sequenced in one lane on the Illumina HiSeq2000 platform (San Diego, CA) at the Center for Applied Genomics, Toronto, ON, Canada, to generate 100-bp, single-end sequence reads.

2.3 | Sequence analysis and processing

Raw sequence reads were quality-filtered and processed primarily using the program STACKS version 1.35 (Catchen et al., 2011, 2013). Raw reads were demultiplexed, trimmed to 90 bp and quality-filtered for a Phred score of 33 or higher using the *process_radtags* command in STACKS. To distinguish sequence reads that belonged to the host fish from those of the bacterial symbiont, we used the '*very_sensitive*' command in BOWTIE2 version 2.2.0 (Langmead & Salzberg, 2012) to filter all reads against the reference genome of *Photobacterium mandapamensis* (Urbanczyk et al., 2011). To ensure that all bacterial reads were removed, we also filtered all remaining reads against the genomes of *Escherichia coli* K12 (Durfee et al., 2008) and *Vibrio campbellii* (Lin et al., 2010) in the same manner. Sequence reads that did not align to the bacterial genomes were assigned as fish (*S. tubifer*) sequences.

Fish sequence reads were processed and assembled de novo to call single nucleotide polymorphisms (SNPs) using the STACKS pipeline with the *rxstacks* correction step. We first ran the DENOVO_MAP program with the parameters −m 3, −M 2 and −n 3, optimized to prevent over- and under-merging of homologous loci. These parameters

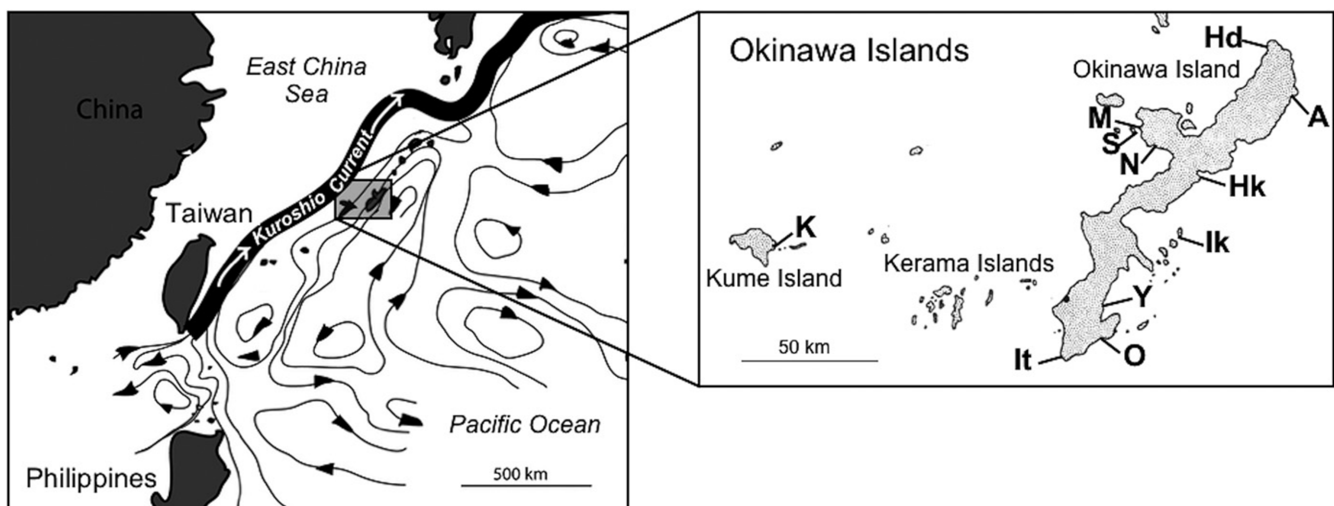

FIGURE 1 Collection sites for *Siphamia tubifer* in the Okinawa Islands, Japan, from 2012 to 2014, and the general current patterns in the region

TABLE 1 The locations and years in which *Siphamia tubifer* were collected in Okinawa, Japan. The range and mean standard lengths (SL) of fish specimens from each sampling site are listed as well as the numbers of individuals collected (N_i) and included in the final genomic dataset after quality filtering (N_f)

ID	Site	Latitude	Longitude	Year	SL (mean); mm	N_i	N_f
S	Sesoko	26.6354	127.8658	2012	15.0–31.0 (22.9)	17	17
				2013	22.0–38.5 (28.5)	18	18
				2014	15.0–38.0 (28.4)	22	21
M	Motobu	26.6558	127.8803	2013	21.0–35.5 (29.5)	26	20
N	Nago	26.6037	127.9324	2013	18.5–42.5 (29.5)	24	21
Hd	Hedo	26.8488	128.2525	2013	17.5–37.5 (26.3)	17	17
It	Itoman	26.0952	127.6585	2013	23.0–36.5 (27.9)	15	14
				2014	13.5–20.0 (16.8)	27	27
O	Ou	26.1280	127.7690	2013	16.5–25.0 (20.1)	16	16
Y	Yonabaru	26.2030	127.7712	2013	21.0–38.5 (28.7)	16	16
Ik	Ikei	26.3935	127.9886	2013	11.5–31.0 (17.3)	16	15
				2014	13.0–30.5 (21.5)	22	22
Hk	Henoko	26.5346	128.0461	2013	14.5–27.5 (19.6)	17	17
A	Ada	26.7420	128.3211	2013	23.0–34.5 (28.5)	16	15
K	Kume	26.3516	126.8201	2014	15.5–41.5 (27.9)	26	24

were previously used to increase the number of loci but to minimize genotyping error for the *de_novo* assembly of other RAD data sets (Mastretta-Yanes et al., 2015). We then implemented the *rxstacks* correction step using a bounded model (−bound_high 0.1) in which excess haplotypes were pruned, and loci for which 25% of individuals had a confounded match in the catalog (−conf_lim 0.25) or an average log likelihood less than (−10.0) were removed. After running this correction step, the CSTACKS (−n 3) and SSTACKS programs were re-applied to produce the final set of RAD tags across all individuals in the study.

2.4 | Population analysis

The mean depth of coverage per locus across all individuals was determined from the STACKS output files, and individuals with a mean coverage less than 10 across all loci were excluded from the analyses. Population summary statistics were computed with the *populations* program in STACKS for loci present in at least 10 populations and in 70% of individuals per population. Population differentiation was evaluated with pairwise F_{ST} values calculated in the *populations* program in STACKS for loci present in all populations and in 70% of individuals per population with a minor allele frequency greater that 5%. Pairwise AMOVA F_{ST} values were also computed in GENODIVE (Meirmans & Van Tienderen, 2004) with 1,000 bootstrap resampling steps, producing *p*-values for all pairwise population comparisons. Isolation by distance (IBD) was assessed by conducting Mantel tests with the Isolation by Distance Web Service (Jensen, Bohonak, & Kelley, 2005) on pairwise F_{ST} values and the natural logarithm of the shortest distances over water between sites; one-sided *p*-values were calculated by randomizing the data 30,000 times. Using the HIERFSTAT program (Goudet, 2005) in R version 3.1.1 (R Core Team 2014), the per locus F_{ST} values were calculated and compared for the following data sets: individuals collected in 2013, 2014 and in three consecutive years from the Sesoko (S) site. Loci in each data set were then split into 10% quantiles based on their estimated F_{ST} values and analysed independently as outliers.

To visualize genetic structure, principal components analyses (PCAs) were implemented on the loci present in at least 70% of individuals per population and in at least 10 populations with a minor allele frequency greater that 5% with the *dudi.pca* function in the ADEGENET version 1.4.2 package (Jombart, 2008; Jombart & Ahmed, 2011) in R (R Core Team 2014). The small percentage of missing data values per locus (an average of 4.1 ± 4.4%, *SD*) were replaced with the mean value across the entire data set at that locus. Analyses of molecular variance (AMOVA) (Excoffier, Smouse, & Quattro, 1992) were carried out in GENODIVE to test for genetic differences between populations and region (populations on the east and west coast of Okinawa). Both PCAs and AMOVAs were performed separately on the 2013 and 2014 populations as well as on the 3-year data set from Sesoko (S) (Figure 1).

Complimentary clustering analyses were also performed with the program STRUCTURE version 2.3.4 (Pritchard, Stephens, & Donnelly, 2000) using the output data files from STACKS comprised of only the first SNP per locus (to eliminate any SNPs that are linked within the same RAD site from the analysis) for loci present in at least 70% of all individuals and in all populations for each data set. Group assignments in STRUCTURE were made using the admixture model with 100,000 burn-in steps and 100,000 MCMC iterations for each number of predetermined genotypic groups (*K*). Analyses were repeated ten times for each value of *K*. For the 2014 and Sesoko data sets, *K* was set from 1 to one plus the total number of populations sampled (5 and 4, respectively), and for the 2013 data set, *K* was set from 1 to 10 (the total number of sites sampled). Probable *K* values were inferred by examining the change in the posterior probability of the log likelihood across all *K* values (ln $P(X|K)$) and by applying the Evanno Δ*K* method (Evanno, Regnaut, & Goudet, 2005) with STRUCTURE HARVESTER (Earl, 2012). All STRUCTURE results were visualized with the program DISTRUCT (Rosenberg, 2004).

2.5 | Outlier analysis

An additional outlier approach was used to identify sets of loci with significantly higher F_{ST} values than expected under a neutral model of selection for each data set (2013, 2014 and Sesoko) with the program LOSITAN (Antao, Lopes, Lopes, Beja-Pereira, & Luikart, 2008). A subset of one SNP per locus present in all populations and in 70% of individuals within a population with a minor allele frequency greater that 5% were examined using the following parameters in LOSITAN: 50,000 simulations, a confidence interval of 0.99, a false discovery rate of 0.1 and a subsample size of 30, using the “neutral means F_{ST}” and “force means F_{ST}” options, which iteratively identify and remove F_{ST} outliers when calculating the global F_{ST} distribution. Based on their probabilities of being under selection, loci were then classified into the following categories: under divergent selection ($p > 99\%$), neutral ($1\% < p < 90\%$), or under balancing selection ($p < .1\%$); all remaining loci were conservatively considered to be unclassified. To visualize whether loci putatively under selection show any patterns concordant with sampling site or time, PCAs were performed independently on each subset of outlier loci as previously described. We also compared the outlier loci identified in LOSITAN to the sets of loci within various per locus F_{ST} quantiles (90%–100%, 80%–90% and 70%–80%).

2.6 | Ethics statement

The protocols used here for the capture and handling of fish specimens were approved by the University of Michigan’s University Committee for the Use and Care of Animals (PRO00004825), and they followed the requirements outlined in the University of the Ryukyus’ Guide for Care and Use of Laboratory Animals (Dobutsu Jikken Kisoku, version 19.6.26).

3 | RESULTS

3.1 | Sequence analysis and processing

The six ddRAD libraries each produced high-quality sequence data with sufficient depth of coverage across most individuals for population-level genetic analyses. Thirteen individuals that had fewer than 800,000 remaining reads after quality filtering (Fig. S1) were discarded from the analysis. On average, 87.05 ± 2.58% of all reads were retained from each library, with an average of 9.72 ± 2.29% and 3.23 ± 2.51% of reads additionally discarded for having ambiguous barcodes or RADtags or for low quality, respectively (Table S1). The final data set consisted of 11,836 loci and contained low percentages of missing data. On average, 4.1% of data was missing per locus (min = 0%, max = 22.1%) and 4.1% of data was missing per individual (min = 0%, max = 13.4%).

Prior to genetic analysis, sequence reads that aligned to the genome of the light organ symbiont, *P. mandapamensis* or to the other examined bacterial genomes, were removed from the data set. The average per cent of quality-filtered reads per individual that aligned to the *P. mandapamensis* genome was 26.9 ± 9.1% (Fig. S1) and of the reads that did not align to *P. mandapamensis*, the average per cent per individual that aligned to the *V. campbelli* or the *E. coli* genomes were 0.293 ± 0.155% and 0.003 ± 0.155%, respectively. These bacterial sequence reads were removed from the data set, and the remaining reads were assigned as *S. tubifer* sequences, resulting in an average of 72.7 ± 9.3% of total reads per individual assigned as *S. tubifer* (Fig. S1). After de novo assembly of the *S. tubifer* sequence reads across all remaining individuals, the mean depth of coverage per locus was 21.6 ± 8.3. One individual with mean sequence coverage per locus below 10 was additionally removed from further analyses (Fig. S1), resulting in a total of 280 individuals in the final data set (Table 1).

3.2 | Population summary statistics

For RAD sites that were polymorphic in at least one population (Table 2), the average major allele frequency (P) and observed heterozygosity (H_{obs}) across all populations ranged from 0.9593 to 0.9607 and 0.0485 to 0.0519, respectively. When sites that were fixed across all populations were included, P increased up to 0.9983 and H_{obs} values all decreased to 0.0022 (Table S2). Across variant RAD sites only, the percentages of polymorphic loci in populations sampled in 2013 were between 30.92% and 36.09%, whereas the percentages of polymorphic loci in populations sampled in 2014 were slightly higher, ranging from 36.79% to 41.75% (Table 2). Levels of overall nucleotide diversity (π) across all sites, fixed and variant, were similar for all groups of fish sampled, varying between 0.0026 and 0.0028. Average measures of Wright’s inbreeding coefficient (F_{IS}) calculated for all variant sites ranged from 0.0407 to 0.0567 across all populations (Table 2). These F_{IS} values are all slightly positive indicating a deficiency in heterozygotes within each population.

3.3 | Analysis of genetic differentiation

Population genetic analyses of the 11,836 SNPs revealed no significant genetic structure between sampling sites or between years. In 2013, pairwise F_{ST} values were consistently low between sites, ranging between 0.0157 and 0.0214, with only a few significant values (Table 3), most of which were for comparisons with fish from the Motobu (M) site (Figure 1). Pairwise F_{ST} values for the 2014 samples were similarly low and nonsignificant, ranging from 0.0116 to 0.0139 (Table S3). An analysis of temporal genetic differentiation at the Sesoko (S) site (Figure 1) over three consecutive years also revealed no significant differentiation over time (F_{ST} values ranged from 0.0158 to 0.0177, Table 4). Similarly, the other two sites that were sampled in consecutive years, Itoman (It) and Ikei (Ik) (Figure 1), had low, nonsignificant F_{ST} values (It: F_{ST} = 0.0151, Ik: F_{ST} = 0.0165) between years. Per locus F_{ST} values for each data set were all low with seemingly normal distributions around zero (Fig. S2) and are summarized in Table S4 including the values used to parse loci into various quantile groups.

TABLE 2 Population genetic summary statistics calculated for each group of *Siphamia tubifer* sampled using only nucleotide positions that are polymorphic in at least one population. Populations are indicated by the abbreviation of their sampling location and year collected. Statistics listed are the average number of individuals analysed at each locus (*N*), the total number of nucleotide positions in the dataset (Sites), the number of unique variable sites in each population (Private), the per cent of polymorphic sites (% Poly), the average frequency of the major allele (P), the average per locus observed heterozygosity (H_{obs}), the average nucleotide diversity (π) and Wright's average inbreeding coefficient (F_{IS}). All statistics were calculated in STACKS

Population	N	Sites	Private	% Poly	P	H_{obs}	π	F_{IS}
A-13	14.3	109,236	2,474	32.38	0.9600	0.0517	0.0631	0.0416
Hd-13	16.0	109,504	2,663	33.96	0.9598	0.0512	0.0630	0.0452
Hk-13	16.0	107,956	2,627	34.03	0.9597	0.0514	0.0632	0.0455
Ik-13	14.1	108,401	2,284	31.71	0.9602	0.0509	0.0625	0.0419
Ik-14	20.7	106,997	3,274	37.55	0.9601	0.0505	0.0623	0.0502
It-13	13.1	109,264	2,120	30.92	0.9598	0.0516	0.0632	0.0407
It-14	25.6	110,648	4,581	41.75	0.9593	0.0519	0.0635	0.0545
K-14	22.4	98,587	3,210	38.43	0.9595	0.0501	0.0631	0.0567
M-13	18.5	82,512	2,166	34.92	0.9607	0.0485	0.0615	0.0534
N-13	19.6	98,854	2,692	36.09	0.9601	0.0492	0.0624	0.0550
O-13	15.0	106,753	2,382	32.67	0.9603	0.0507	0.0624	0.0439
S-12	15.5	97,232	2,070	32.56	0.9602	0.0495	0.0622	0.0473
S-13	16.8	108,356	2,846	34.69	0.9598	0.0510	0.0630	0.0473
S-14	19.7	107,381	3,086	36.79	0.9599	0.0505	0.0627	0.0511
Y-13	15.1	108,186	2,539	33.07	0.9603	0.0514	0.0624	0.0418

TABLE 3 Pairwise F_{ST} values (top diagonal) and the shortest distance (km) through water (bottom diagonal) between groups of *Siphamia tubifer* collected in 2013. F_{ST} values in bold are significant at $p < .05$

2013	A	Hd	Hk	Ik	It	M	N	O	S	Y
A	—	0.0187	0.0191	0.0197	0.0207	**0.0182**	0.0169	0.0192	0.0183	0.0194
Hd	22.4	—	0.0179	0.0191	0.0198	**0.0172**	0.0161	0.0184	0.0170	0.0183
Hk	38.2	59.7	—	0.0192	0.0198	**0.0173**	0.0163	0.0184	0.0176	0.0184
Ik	51.4	73.3	17.7	—	0.0214	0.0183	0.0172	0.0193	0.0185	0.0194
It	104.5	117.3	71.5	54.8	—	**0.0189**	0.0174	0.0205	0.0192	0.0199
M	68.0	47.8	105.3	119.5	72.4	—	**0.0157**	0.0176	0.0169	0.0176
N	78.2	58.0	115.5	129.3	80.8	10.4	—	0.0167	0.0157	**0.0168**
O	89.0	111.3	58.6	38.1	15.8	84.4	83.5	—	0.0181	0.0186
S	70.4	50.1	107.7	121.9	68.2	2.8	8.6	82.1	—	0.0175
Y	83.6	105.0	51.8	33.4	32.1	100.5	99.8	16.8	98.4	—

Analyses of isolation by distance revealed no significant relationship between the low levels of population differentiation observed and geographic distance between locations within either sampling year (2013: $F_{1,43} = 0.369$, $R^2 = .0085$, $p = .547$; 2014: $F_{1,4} = 3.372$, $R^2 = .457$ $p = .140$; Fig. S3). Similarly, results from the AMOVA indicate that all of the observed genetic variation is attributed to variation within (F_{IT}) and among (F_{IS}) individuals and none is attributed to differences between populations or between the east and west coast regions of Okinawa Island (Table 5).

Principal components analyses of all 11,836 SNPs confirmed the lack of genetic structure, with no apparent clustering of individuals by location (Figure 2). The first two PC axes for the 2013 analysis each described less than 1% of the variation in the data. These values rose slightly for the 2014 data, accounting for a combined (PC1 and PC2) total of 2.57% of the total genetic variation (Figure 2). The PCA carried out on the temporal data set from the Sesoko site also indicates a lack of structure over time (Figure 2), and it is consistent with the low, nonsignificant pairwise F_{ST} values between groups of fish collected at that site in subsequent years (Table 3). Similarly, clustering analyses in STRUCTURE revealed that a $K = 1$ is most likely for the 2013, 2014 and Sesoko data sets when all loci were considered (Tables S5–S7). In the

TABLE 4 Pairwise F_{ST} values between groups of *Siphamia tubifer* collected at the Sesoko site in three consecutive years

Sesoko	2012	2013	2014
2012	—	0.0177	0.0161
2013		—	0.0158

TABLE 5 Analysis of molecular variance (AMOVA) of groups of *Siphamia tubifer* collected in 2013. Populations were grouped into regions determined by collection sites on the east or west coast of Okinawa

Source of variation	Nested in	% Variance	F-statistic	*p*-value
Within individual	—	86.4	F_{IT}	—
Among individual	Population	13.6	F_{IS}	.001
Among population	Region	0.0	F_{SC}	.245
Among region	—	0.0	F_{CT}	.012

case where $K = 1$, the Evanno ΔK method of detecting the true value of K is ineffective (Evanno et al., 2005). We therefore examined the mean posterior probabilities for each K value in both data sets; the log likelihood was highest when $K = 1$ for all data sets (Tables S5-S7). These results provide further evidence that one panmictic population of *S. tubifer* is present in the region, including Kume Island (K) 100 km to the west of Okinawa Island (Figure 1).

3.4 | Outlier analysis

The LOSITAN outlier analyses identified a small percentage of loci (0.7%–2.8%) within each data set that were classified as being under divergent selection (Figure 3a–c, Table 6). The percentages of shared loci under selection between the 2013 and 2014 LOSITAN data sets were 2% and 8%, respectively, and the number of shared outliers between the 2013 and 2014 data sets and the Sesoko data set were 12% and 6%. The numbers of shared loci under divergent selection between sampling years within the Sesoko data set were higher, ranging from 12% to 26%. Principal components analyses of these adaptive loci only revealed a slight signature of genetic differentiation between individuals collected at the broader geographic range sampled in 2014, but less so between individuals collected from reefs around Okinawa Island in 2013 (Figure 3d,e). Interestingly, there was more apparent differentiation at these putative loci under selection between individuals collected in different years at Sesoko (S) (Figure 3f).

Many of the loci classified as being under divergent selection were also present in the 90%–100% F_{ST} quantile data sets; 85%, 91% and 84% of the loci classified as being under divergent selection in LOSITAN were also in the 2013, 2014 and Sesoko F_{ST} outlier data sets, respectively. Additional STRUCTURE analyses of subsets of SNPs in each 10% F_{ST} quantile (Table S4) confirmed a lack of genetic structure for groups of fish collected from sites around Okinawa Island in 2013, even when examining only the loci with the highest levels of differentiation (90%–100% F_{ST} quantile) observed across all individuals (Fig. S4a). A weak signature of genetic similarity that matched sampling location was evident for the most highly differentiated subset of loci (90%–100% F_{ST} quantile) for the 2014 data set, and both the highest ln $P(K)$ and a rise in ΔK were evident when $K = 4$ (Table S5). However, this pattern of genetic structure was not apparent for the subset of loci in the 80%–90% F_{ST} quantile for the same populations (Fig. S4b). The most highly differentiated subset of loci (90%–100% F_{ST} quantile) in the Sesoko data set revealed three genetic clusters that matched sampling year (Fig. S4c) and corroborate with the temporal pattern of differentiation observed with the adaptive set of loci previously identified. Moreover, principal components analyses of the LOSITAN outliers produced comparable results to analyses of the 90%–100% F_{ST} outliers for each data set (Fig. S5) with the exception of the 2013 data set; the analysis of LOSITAN outliers revealed a weak pattern by sampling location that was not as evident for the 90%–100% F_{ST} outlier data set from 2013.

4 | DISCUSSION

The observed discrepancies between potential and actual dispersal distances of reef fish larvae (e.g., Bowen et al., 2006; Galarza et al.,

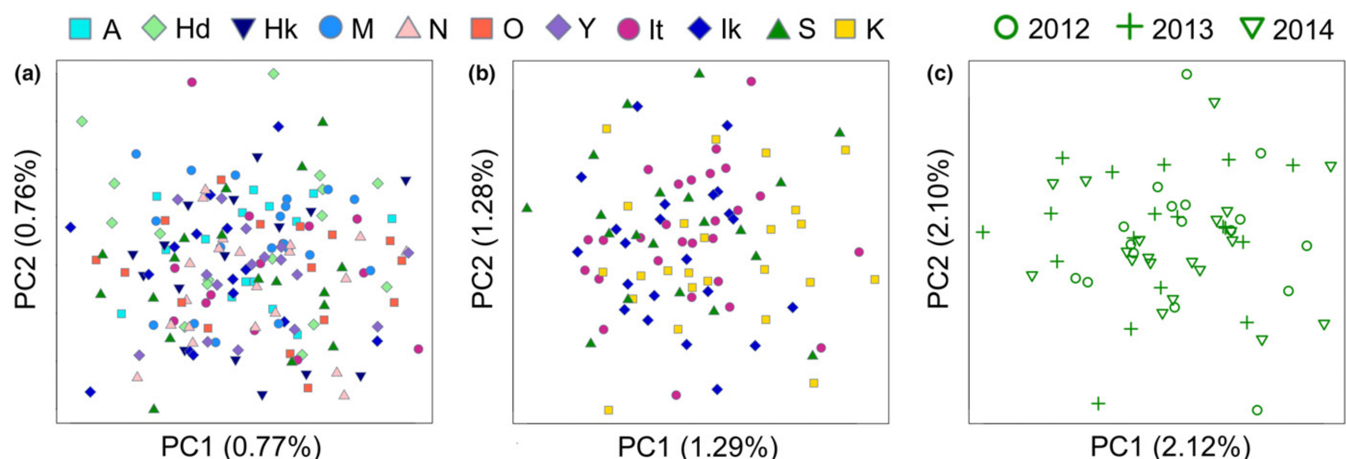

FIGURE 2 Principal components analyses of genetic differentiation among *Siphamia tubifer* collected in (a) 2013, (b) 2014 and (c) from the Sesoko site collected in three consecutive years. Points represent individuals along the PC1 and PC2 axes of genetic variation across 11,836 SNPs with the amount of variation explained by each axis in parentheses. Distinct colours and shapes indicate the sampling locations [Colour figure can be viewed at wileyonlinelibrary.com]

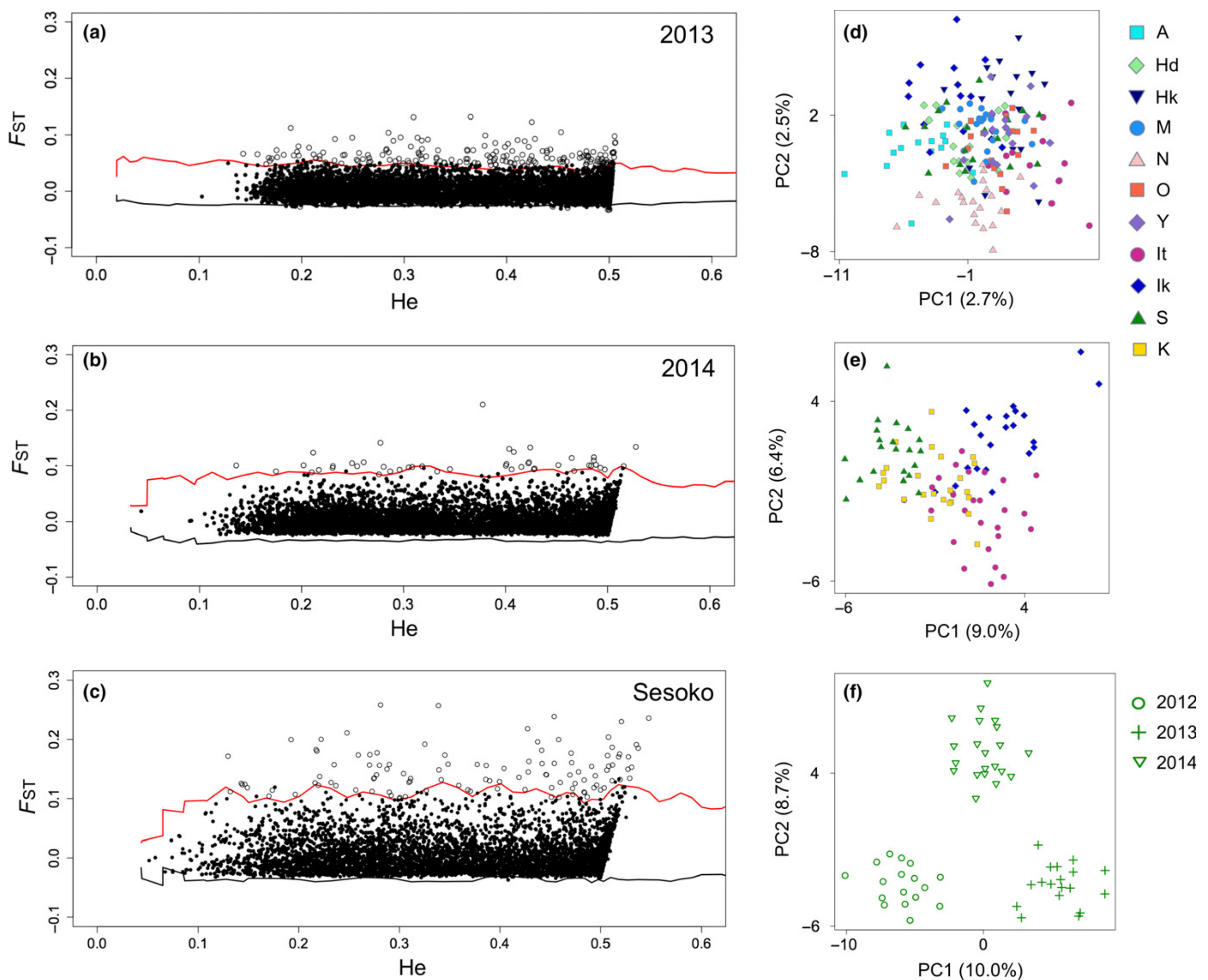

FIGURE 3 Results of the F_{ST} outlier tests in LOSITAN for 6,379 SNPs in (a) 2013, (b) 2014 and (c) at the Sesoko (S) site over three consecutive years (2012–2014) and corresponding principal components analyses of genetic differentiation for the outlier loci identified (d–f). Loci above the red line (open circles) are candidates of being under divergent selection above a 99% probability. Loci with a probability between 1% and 90% of being under selection are considered to be neutral and those below the bottom line with <0.1% probability are classified as being under balancing selection. All remaining loci were conservatively considered to be unclassified (neither neutral or under selection) [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 6 Summary of classification of SNPs from the LOSITAN outlier F_{ST} analysis of 6,379 loci for each dataset listed. Loci were classified based on the probabilities of being under selection indicated in parentheses. Remaining loci were conservatively considered to be unclassified

Data set	Divergent (>99%)	Neutral (1%–90%)	Balancing (<0.1%)	Unclassified
2013	180 (2.8%)	4,776 (74.9%)	107 (1.7%)	1,316 (20.6%)
2014	46 (0.7%)	5,922 (92.8%)	0 (0%)	411 (6.4%)
Sesoko	117 (1.8%)	5,456 (85.5%)	14 (0.22%)	792 (12.4%)

2009; Jones et al., 2009; Rocha et al., 2005; Taylor & Hellberg, 2003) highlight the need for a better understanding of the influence of larval fish behaviour on the connectivity of reef fish populations.

To narrow this gap in knowledge, general links between larval behaviour and gene flow across diverse groups of fishes and geographic regions are required. We applied RAD-Seq methods to test for genetic differentiation between populations of *S. tubifer*, as documented for another cardinalfish species with similar life-history traits and behaviour (Gerlach et al., 2007; Kingsford et al., 2014). Despite the potential of RAD-Seq methods to detect fine-scale genetic structure, and in contrast to studies of other cardinalfishes, we found little evidence of genetic differentiation between groups of *S. tubifer* at spatial scales up to 140 km. This lack of differentiation was particularly striking in that collection sites were located on both the sides of Okinawa Island and included a site, Kume Island, well separated to the west of the main island (Figure 1). Analyses of adaptive outlier loci, however, revealed a shallow signature of genetic divergence

between sampling sites, particularly at the larger spatial scales examined, but this signature was evident only for a small number of the most highly differentiated loci. The overall pattern of genetic admixture observed from the analysis of all identified SNPs indicates a significant amount of gene flow among groups of *S. tubifer* in the Okinawa Islands, presumably due to larval dispersal and mixing by strong ocean currents in the region (Figure 1). Although these results do not exclude the possibility of natal homing, they demonstrate that genetic exchange between the sampling sites occurs frequently enough to maintain genetic homogeneity in the region.

The typical dispersal distances of reef fish larvae are on the order of 50–100 km, with some local retention, and populations of fishes with high mortality rates tend to be subsidized with larvae from greater distances (Cowen et al., 2006). Consistent with this model, *S. tubifer* is apparently subject to high mortality rates (Gould et al., 2014, 2016) and may therefore depend on larval subsidies from other sources for population persistence. *Siphamia tubifer* also has a highly specialized habitat requirement; groups of *S. tubifer* closely associate with the sea urchins *Diadema setosum* and *Echinothrix calamaris* during the daytime, seeking shelter among the urchins' long spines (Eibl-Eibesfeldt, 1961; Gould et al., 2014; Lachner, 1955; Tamura, 1982). The distribution of diademid urchins in reef habitats can be patchy, partially due to variation in sediments between sites (Dumas et al., 2007; Nishihira et al., 1991). Therefore, the availability of suitable settlement habitat for *S. tubifer* larvae is also likely to be variable between reefs. In addition, *S. tubifer* acquires its species-specific luminous bacterium from the environment during larval development (Dunlap et al., 2012). These two factors, habitat availability and the ecological dynamics of symbiont acquisition, which are presently undefined, could limit the recruitment success of *S. tubifer* and thereby contribute substantially to the admixture observed. The amount of gene flow required to maintain genetic connectivity over a large scale is on the order of only a few individuals per generation (Leis, 2002; Shulman, 1998; Shulman & Bermingham, 1995); therefore if few, far-dispersing *S. tubifer* larvae settle sporadically on reefs, there would be little potential for genetic divergence to accumulate between populations over time.

Corresponding with substantial gene flow in the region, the islands in Japan's Ryukyu Archipelago receive larval supply from other reef habitats in the south. In particular, small reef fish from the Philippines have an ecologically significant linkage potential to the Ryukyu Islands (Treml et al., 2015). The connectivity potential from the northern Philippines is driven by the strong ocean currents in the region, namely the Kuroshio Current, which originates off the coast of the Philippines and flows northward through the Ryukyu Archipelago (Figure 1). Previous studies in the region have established evidence of genetic homogeneity among the Philippine Islands and across the Ryukyu Islands along the Kuroshio Current for the crown-of-thorns sea star *Acanthaster planci* (Yasuda et al., 2009) and for the broadcast-spawning coral *Acropora digitifera* (Nakajima et al., 2010). Moreover, at smaller spatial scales within Okinawa, there is evidence of significant gene flow for several coral species (Nishikawa, 2008; Nishikawa, Katoh, & Sakai, 2003), although two distinct genetic clusters between the North and South regions of Okinawa Island were apparent for the scleractinian coral *Goniastrea aspera*, indicating a substantial number of locally produced recruits and the potential for self-recruitment in the region despite the strong influence of the Kuroshio Current (Nishikawa & Sakai, 2005).

In addition to spatial differentiation, we examined temporal genetic divergence in *S. tubifer* at one study site by collecting specimens over three consecutive years. The dynamics of temporal genetic structure may be even more informative than spatial dynamics in marine systems, but this issue has generally been overlooked for most marine populations (Hedgecock, Barber, & Edmands, 2007; Hellberg et al., 2002). Of the studies that have examined temporal structure, instances of temporal stability have been rare (Bernal-Ramírez et al., 2003; Larsson et al., 2010), whereas temporal genetic differentiation has been reported for several marine fishes (e.g., Klanten, Choat, & van Herwerden, 2007; Maes et al., 2006; Planes & Lenfant, 2002; Selkoe et al., 2006). Genetic differentiation over time at a location can result from selection, random genetic drift or from variable larval supply from different source populations (Hedgecock et al., 2007). When examining all SNPs, we observed low levels of differentiation between groups of *S. tubifer* collected from the same site over 3 years. However, when we considered only outlier loci putatively under selection, a clear pattern of temporal differentiation was revealed. To determine whether the observed differentiation at these adaptive loci might be attributed to cohesive cohort dispersal and settlement, we compared relatedness values between individuals collected in the same year with those collected in different years but saw no significant differences in relatedness between these groups of individuals (Fig. S6). We also compared the numbers of shared outlier loci that were divergent between consecutive sampling years and found that approximately one-fourth of the loci under selection are shared between the interannual comparisons. Therefore, the observed temporal patterns of structure at these outlier loci can likely be attributed to variable larval supply from different upstream source populations rather than solely due to genetic drift.

Siphamia tubifer is short-lived, with an expected longevity of less than 200 days (Gould et al., 2016); therefore, population turnover at a reef might occur regularly, especially given the potential variability in larval supply consistent with large dispersal distances and high larval mortality rates associated with reef fishes. Several divergent *S. tubifer* populations upstream of the Okinawa Islands, such as in the Philippines, could therefore supply larvae that variably settle on available urchin habitat and establish temporary populations at a given reef site. These populations might subsequently be vulnerable to population bottlenecks due to high adult mortality rates (Gould et al., 2016) along with natural disturbance events, such as typhoons, which are known to cause a decline in the numbers of some reef fishes, especially for demersal fish with specific habitat requirements (Halford, Cheal, Ryan, & Williams, 2004; Harmelin-Vivien, 1994) like *S. tubifer*. Typhoon occurrence in the study region has increased in recent years to an average of approximately six typhoons per year (Tu, Chou, & Chu, 2009). Additionally, potential seasonality in the breeding pattern of *S. tubifer* could enhance the likelihood of population turnover at a reef

site. Many fish species have seasonal peaks in reproduction at higher latitude reefs associated with temperature fluctuations (Munday, Jones, Pratchett, & Williams, 2008), and seasonal reproduction has been documented for other reef-associated fishes in Okinawa (Kuwamura, Yogo, & Nakashima, 1994). Although breeding seasonality in *S. tubifer* has not been described, any seasonal differences in reproduction could also contribute to the observed temporal patterns of differentiation for the subset of adaptive loci identified.

Overall our study supports the importance of examining the degree of genetic differentiation within various geographic regions for a range of fish species as well as species-specific patterns across groups of reef fishes to better understand the relationships between life-history traits, larval behaviour and gene flow. We highlight the effectiveness of RAD-Seq methods, which have recently been applied to examine the genomics of other coral reef fishes (Gaither et al., 2015; Picq, McMillan, & Puebla, 2016; Puebla, Bermingham, & McMillan, 2014; Saenz-Agudelo et al., 2015; Stockwell et al., 2016), to study such nonmodel organisms. Our results show that in contrast to other cardinalfish species (Bernardi & Vagelli, 2004; Gerlach et al., 2007; Gotoh et al., 2009; Hoffman et al., 2005; Vagelli et al., 2008), *S. tubifer* exhibits genetic admixture over a 140-km region despite the species' demonstrated homing abilities and olfactory preferences for a home reef (Gould et al., 2015). These findings suggest that strong ocean currents combined with a month-long pelagic larval phase promote dispersal and gene flow in the region. We also found evidence for temporal genetic differentiation at a small number of loci putatively under selection, suggesting adaptive variation in the source populations that supply *S. tubifer* larvae to the Okinawa Islands. Future investigations *S. tubifer* populations at varying locations across this species' broad, Indo-Pacific distribution will provide insight on whether various source populations are undergoing divergent selection and whether the scale of genetic admixture observed here is region-specific.

ACKNOWLEDGEMENTS

We gratefully acknowledge Saki Harii and the staff of Sesoko Station, University of the Ryukyus for logistical assistance. We thank Katherine Dougan and Maggie Grundler, University of Michigan (UM), for technical and field assistance, and we thank members of Lacey Knowles' Lab (UM), especially Rob Massatti, for guidance and assistance with the RAD protocols and data analysis. We also thank Raquel Rivadeneira and the UM Genomic Diversity Lab for technical services and the Center for Applied Genomics for their sequencing services. Funding for this project was provided by National Science Foundation grant DEB-1405286, by a UM Rackham Graduate School Susan Lipschutz Award and by a UM Museum of Zoology Ichthyology Award.

DATA ACCESSIBILITY

Illumina RAD-tag sequences are accessible at NCBI SRA Accession No. SRP105806 (Biosample Accession Nos: SAMN06857385–SAMN06857664). The STACKS consensus sequences, SNP genotype calls and STRUCTURE input files are accessible at Dryad (https://doi.org/10.5061/dryad.5n882).

AUTHOR CONTRIBUTIONS

A.L.G. conceived of and designed the study, performed laboratory work and analysed the data. A.L.G. and P.V.D. performed fieldwork and wrote the article.

REFERENCES

Almany, G. R., Berumen, M. L., Thorrold, S. R., Planes, S., & Jones, G. P. (2007). Local replenishment of coral reef fish populations in a marine reserve. *Science*, *316*, 742–744.

Antao, T., Lopes, A., Lopes, R. J., Beja-Pereira, A., & Luikart, G. (2008). LOSITAN: A workbench to detect molecular adaptation based on a F_{ST}-outlier method. *BMC Bioinformatics*, *9*, 323.

Atema, J. (2012). Aquatic odor dispersal fields: Opportunities and limits of detection, communication and navigation. In C. Bronmark, & L.-A. Hansson (Eds.), *Chemical ecology in aquatic ecosystems* (pp. 1–18). Oxford: Oxford University Press.

Barlow, G. W. (1981). Patterns of parental investment, dispersal and size among coral-reef fishes. *Environmental Biology of Fishes*, *6*, 65–85.

Bernal-Ramírez, J. H., Adcock, G. J., Hauser, L., Carvalho, G. R., & Smith, P. J. (2003). Temporal stability of genetic population structure in the New Zealand snapper, *Pagrus auratus*, and relationship to coastal currents. *Marine Biology*, *142*, 567–574.

Bernardi, G., & Vagelli, A. (2004). Population structure in Banggai cardinalfish, *Pterapogon kauderni*, a coral reef species lacking a pelagic larval phase. *Marine Biology*, *145*, 803–810.

Bode, M., Bode, L., & Armsworth, P. R. (2006). Larval dispersal reveals regional sources and sinks in the Great Barrier Reef. *Marine Ecology Progress Series*, *308*, 17–25.

Bowen, B. W., Bass, A. L., Muss, A., Carlin, J., & Robertson, D. R. (2006). Phylogeography of two Atlantic squirrelfishes (Family Holocentridae): Exploring links between pelagic larval duration and population connectivity. *Marine Biology*, *149*, 899–913.

Bradbury, I. R., Hamilton, L. C., Dempson, B., Robertson, M. J., Bourret, V., Bernatchez, L., & Verspoor, E. (2015). Transatlantic secondary contact in Atlantic Salmon, comparing microsatellites, a single nucleotide polymorphism array and restriction-site associated DNA sequencing for the resolution of complex spatial structure. *Molecular Ecology*, *24*, 5130–5144.

Catchen, J. M., Amores, A., Hohenlohe, P. A., Cresko, W., & Postlethwait, J. H. (2011). Stacks: Building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics*, *1*, 171–182.

Catchen, J. M., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*, 3124–3140.

Coates, B. S., Sumerford, D. V., Miller, N. J., Kim, K. S., Sappington, T. W., Siegfried, B. D., & Lewis, L. C. (2009). Comparative performance of single nucleotide polymorphism and microsatellite markers for population genetic analysis. *Journal of Heredity*, *100*, 556–564.

Cowen, R. K., Lwiza, K. M., Sponaugle, S., Paris, C. B., & Olson, D. B. (2000). Connectivity of marine populations: Open or closed? *Science*, *287*, 857–859.

Cowen, R. K., Paris, C. B., & Srinivasan, A. (2006). Scaling of connectivity in marine populations. *Science*, *311*, 522–527.

Davey, J. W., & Blaxter, M. L. (2010). RADSeq: Next-generation population genetics. *Briefings in Functional Genomics*, *9*, 416–423.

Dixson, D. L., Jones, G. P., Munday, P. L., Planes, S., Pratchett, M. S., Srinivasan, M., . . . Thorrold, S. R. (2008). Coral reef fish smell leaves

to find island homes. *Proceedings of the Royal Society of London B: Biological Sciences*, *275*, 2831–2839.

Døving, K. B., Stabell, O. B., Östlund-Nilsson, S., & Fisher, R. (2006). Site fidelity and homing in tropical coral reef cardinalfish: Are they using olfactory cues? *Chemical Senses*, *31*, 265–272.

Dumas, P., Kulbicki, M., Chifflet, S., Fichez, R., & Ferraris, J. (2007). Environmental factors influencing urchin spatial distributions on disturbed coral reefs (New Caledonia, South Pacific). *Journal of Experimental Marine Biology and Ecology*, *344*, 88–100.

Dunlap, P. V., Gould, A. L., Wittenrich, M. L., & Nakamura, M. (2012). Symbiosis initiation in the bacterially luminous sea urchin cardinalfish *Siphamia versicolor*. *Journal of Fish Biology*, *81*, 1340–1356.

Dunlap, P. V., & Nakamura, M. (2011). Functional morphology of the luminescence system of *Siphamia versicolor* (Perciformes: Apogonidae), a bacterially luminous coral reef fish. *Journal of Morphology*, *272*, 897–909.

Durfee, T., Nelson, R., Baldwin, S., Plunkett, G., Burland, V., Mau, B., … Gibbs, R. A. (2008). The complete genome sequence of *Escherichia coli* DH10B: Insights into the biology of a laboratory workhorse. *Journal of Bacteriology*, *190*, 2597–2606.

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

Eibl-Eibesfeldt, I. (1961). Eine Symbiose zwischen Fischen (*Siphamia versicolor*) und Seeigeln. *Zeitschrift fur Tierpsychologie*, *18*, 56–59.

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., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, *131*, 479–491.

Fisher, R., Leis, J. M., Clark, D. L., & Wilson, S. K. (2005). Critical swimming speeds of late-stage coral reef fish larvae: Variation within species, among species and between locations. *Marine Biology*, *147*, 1201–1212.

Gaither, M. R., Bernal, M. A., Coleman, R. R., Bowen, B. W., Jones, S. A., Simison, W. B., & Rocha, L. A. (2015). Genomic signatures of geographic isolation and natural selection in coral reef fishes. *Molecular Ecology*, *24*, 1543–1557.

Galarza, J. A., Carreras-Carbonell, J., Macpherson, E., Pascual, M., Roques, S., Turner, G. F., & Rico, C. (2009). The influence of oceanographic fronts and early-life-history traits on connectivity among littoral fish species. *Proceedings of the National Academy of Sciences*, *106*, 1473–1478.

Gerlach, G., Atema, J., Kingsford, M. J., Black, K. P., & Miller-Sims, V. (2007). Smelling home can prevent dispersal of reef fish larvae. *Proceedings of the National Academy of Sciences*, *104*, 858–863.

Gotoh, R. O., Sekimoto, H., Chiba, S. N., & Hanzawa, N. (2009). Peripatric differentiation among adjacent marine lake and lagoon populations of a coastal fish, *Sphaeramia orbicularis* (Apogonidae, Perciformes, Teleostei). *Genes & Genetic Systems*, *84*, 287–295.

Goudet, J. (2005). Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, *5*, 184–186.

Gould, A. L., Dougan, K. E., Koeningbauer, S., & Dunlap, P. V. (2016). Life history of a symbiotically luminous cardinalfish. *Journal of Fish Biology*, *89*, 1359–1377.

Gould, A. L., Harii, S., & Dunlap, P. V. (2014). Host preference, site fidelity, and homing behavior of the symbiotically luminous cardinalfish, *Siphamia tubifer* (Perciformes: Apogonidae). *Marine Biology*, *161*, 2897–2907.

Gould, A. L., Harii, S., & Dunlap, P. V. (2015). Cues from the reef: Olfactory preferences of a symbiotically luminous cardinalfish. *Coral Reefs*, *34*, 673–677.

Halford, A., Cheal, A. J., Ryan, D., & Williams, D. M. (2004). Resilience to large-scale disturbance in coral and fish assemblages on the Great Barrier Reef, Australia. *Ecology*, *85*, 1892–1905.

Harmelin-Vivien, M. (1994). The effects of storms and cyclones on coral reefs: A review. *Journal of Coastal Research*, *12*, 211–231.

Hedgecock, D., Barber, P. H., & Edmands, S. (2007). Genetic approaches to measuring connectivity. *Oceanography*, *20*, 70–79.

Hellberg, M. E., Burton, R. S., Neigel, J. E., & Palumbi, S. R. (2002). Genetic assessment of connectivity among marine populations. *Bulletin of Marine Science*, *70*, 273–290.

Hoffman, E. A., Kolm, N., Berglund, A., Arguello, J. R., & Jones, A. G. (2005). Genetic structure in the coral-reef-associated Banggai cardinalfish, Pterapogon kauderni. *Molecular Ecology*, *14*, 1367–1375.

Horne, J. B., van Herwerden, L., Choat, J. H., & Robertson, D. R. (2008). High population connectivity across the Indo-Pacific: Congruent lack of phylogeographic structure in three reef fish congeners. *Molecular Phylogenetics and Evolution*, *49*, 629–638.

Iwai, T. (1958). A study of the luminous organ of the apogonid fish *Siphamia versicolor* (Smith and Radcliffe). *Journal of the Washington Academy of Sciences*, *42*, 267–270.

Jensen, J. L., Bohonak, A. J., & Kelley, S. T. (2005). Isolation by distance, web service. *BMC Genetics*, *6*, 13.

Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*, 1403–1405.

Jombart, T., & Ahmed, I. (2011). *adegenet 1.3-1*: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*, 3070–3071.

Jones, G. P., Almany, G. R., Russ, G. R., Sale, P. F., Steneck, R. S., Van Oppen, M. J. H., & Willis, B. L. (2009). Larval retention and connectivity among populations of corals and reef fishes: History, advances and challenges. *Coral Reefs*, *28*, 307–325.

Jones, G. P., Milicich, M. J., Emslie, M. J., & Lunow, C. (1999). Self-recruitment in a coral reef fish population. *Nature*, *402*, 802–804.

Jones, G. P., Planes, S., & Thorrold, S. R. (2005). Coral reef fish larvae settle close to home. *Current Biology*, *15*, 1314–1318.

Kingsford, M. J., Finn, M. D., O'Callaghan, M. D., Atema, J., & Gerlach, G. (2014). Planktonic larval duration, age and growth of *Ostorhinchus doederleini* (Pisces: Apogonidae) on the southern Great Barrier Reef, Australia. *Marine Biology*, *161*, 245–259.

Kingsford, M. J., Leis, J. M., Shanks, A., Lindeman, K. C., Morgan, S. G., & Pineda, J. (2002). Sensory environments, larval abilities and local self-recruitment. *Bulletin of Marine Science*, *70*, 309–340.

Klanten, O. S., Choat, J. H., & van Herwerden, L. (2007). Extreme genetic diversity and temporal rather than spatial partitioning in a widely distributed coral reef fish. *Marine Biology*, *150*, 659–670.

Kolm, N., Hoffman, E. A., Olsson, J., Berglund, A., & Jones, A. G. (2005). Group stability and homing behavior but no kin group structures in a coral reef fish. *Behavioral Ecology*, *16*, 521–527.

Kuwamura, T., Yogo, Y., & Nakashima, Y. (1994). Population dynamics of goby Paragobiodon echinocephalus and host coral Stylophora pistillata. *Marine Ecology Progress Series*, *103*, 17–23.

Lachner, E. A. (1955). Inquilinism and a new record for *Paramia bipunctata*, a new cardinal fish from the Red Sea. *Copeia*, *1955*, 53–55.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie2. *Nature Methods*, *9*, 357–359.

Larsson, L. C., Laikre, L., André, C., Dahlgren, T. G., & Ryman, N. (2010). Temporally stable genetic structure of heavily exploited Atlantic herring (*Clupea harengus*) in Swedish waters. *Heredity*, *104*, 40–51.

Leis, J. M. (2002). Pacific coral-reef fishes: The implications of behaviour and ecology of larvae for biodiversity and conservation, and a reassessment of the open population paradigm. *Environmental Biology of Fishes*, *65*, 199–208.

Leis, J. M. (2007). Behaviour as input for modelling dispersal of fish larvae: Behaviour, biogeography, hydrodynamics, ontogeny, physiology

and phylogeny meet hydrography. *Marine Ecology Progress Series*, *347*, 185–193.

Leis, J. M., Carson-Ewart, B. M., Hay, A. C., & Cato, D. H. (2003). Coral-reef sounds enable nocturnal navigation by some reef-fish larvae in some places and at some times. *Journal of Fish Biology*, *63*, 724–737.

Leis, J. M., Hay, A. C., & Trnski, T. (2006). In situ ontogeny of behaviour in pelagic larvae of three temperate, marine, demersal fishes. *Marine Biology*, *148*, 655–669.

Leis, J. M., & McCormick, M. I. (2002). The biology, behavior, and ecology of the pelagic, larval stage of coral reef fishes. In P. F. Sale (Ed.), *Coral reef fishes: Dynamics and diversity in a complex ecosystem* (pp. 171–199). New York: Academic Press.

Leis, J. M., Siebeck, U., & Dixson, D. L. (2011). How Nemo finds home: The neuroecology of dispersal and of population connectivity in larvae of marine fishes. *Integrative and Comparative Biology*, *51*, 826–843.

Lin, B., Wang, Z., Malanoski, A. P., O'grady, E. A., Wimpee, C. F., Vuddhakul, V., ... Vora, G. J. (2010). Comparative genomic analyses identify the *Vibrio harveyi* genome sequenced strains BAA-1116 and HY01 as *Vibrio campbellii*. *Environmental Microbiology Reports*, *2*, 81–89.

Luikart, G., England, P. R., Tallmon, D., Jordan, S., & Taberlet, P. (2003). The power and promise of population genomics: From genotyping to genome typing. *Nature Reviews Genetics*, *4*, 981–994.

Maes, G. E., Pujolar, J. M., Hellemans, B., & Volckaert, F. A. (2006). Evidence for isolation by time in the European eel (*Anguilla anguilla* L.). *Molecular Ecology*, *15*, 2095–2107.

Marnane, M. J. (2000). Site fidelity and homing behaviour in coral reef cardinalfishes. *Journal of Fish Biology*, *57*, 1590–1600.

Mastretta-Yanes, A., Arrigo, N., Alvarez, N., Jorgensen, T. H., Piñero, D., & Emerson, B. C. (2015). Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Molecular Ecology Resources*, *15*, 28–41.

van der Meer, M. H., Hobbs, J. P. A., Jones, G. P., & Van Herwerden, L. (2012). Genetic connectivity among and self-replenishment within island populations of a restricted range subtropical reef fish. *PLoS ONE*, *7*, e49660.

Meirmans, P. G., & Van Tienderen, P. H. (2004). GENOTYPE and GENODIVE: Two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, *4*, 792–794.

Montgomery, J. C., Jeffs, A., Simpson, S. D., Meekan, M., & Tindle, C. (2006). Sound as an orientation cue for the pelagic larvae of reef fishes and decapod crustaceans. *Advances in Marine Biology*, *51*, 143–196.

Mora, C., & Sale, P. F. (2002). Are populations of coral reef fish open or closed? *Trends in Ecology & Evolution*, *17*, 422–428.

Munday, P. L., Jones, G. P., Pratchett, M. S., & Williams, A. J. (2008). Climate change and the future for coral reef fishes. *Fish and Fisheries*, *9*, 261–285.

Nakajima, Y., Nishikawa, A., Iguchi, A., & Sakai, K. (2010). Gene flow and genetic diversity of a broadcast-spawning coral in northern peripheral populations. *PLoS ONE*, *5*, e11149.

Nishihira, M., Sato, Y., Arakaki, Y., & Tsuchiya, M. (1991). Ecological distribution and habitat preference of four types of the sea urchin *Echinometra mathaei* on the Okinawan coral reefs. In T. Yanagisawa, I. Yasumasu, C. Oguro, et al. (Eds.), *Biology of Echinodermata: Proceedings of the 7th International Echinoderm Conference* (pp. 91–104). Rotterdam: Balkema.

Nishikawa, A. (2008). Degree and pattern of gene flow in several scleractinian corals in the Ryukyu Archipelago, Southern Japan. *Pacific Science*, *62*, 413–421.

Nishikawa, A., Katoh, M., & Sakai, K. (2003). Larval settlement rates and gene flow of broadcast-spawning (*Acropora tenuis*) and planula-brooding (*Stylophora pistillata*) corals. *Marine Ecology Progress Series*, *256*, 87–97.

Nishikawa, A., & Sakai, K. (2005). Genetic connectivity of the scleractinian coral *Goniastrea aspera* around the Okinawa Islands. *Coral Reefs*, *24*, 318–323.

Parchman, T. L., Gompert, Z., Mudge, J., Schilkey, F. D., Benkman, C. W., & Buerkle, C. (2012). Genome-wide association genetics of an adaptive trait in lodgepole pine. *Molecular Ecology*, *21*, 2991–3005.

Paris, C. B., Atema, J., Irisson, J. O., Kingsford, M., Gerlach, G., & Guigand, C. M. (2013). Reef odor: A wake up call for navigation in reef fish larvae. *PLoS ONE*, *8*, e72808.

Paris, C. B., Chérubin, L. M., & Cowen, R. K. (2007). Surfing, spinning, or diving from reef to reef: Effects on population connectivity. *Marine Ecology Progress Series*, *347*, 285–300.

Paris, C. B., & Cowen, R. K. (2004). Direct evidence of a biophysical retention mechanism for coral reef fish larvae. *Limnology and Oceanography*, *49*, 1964–1979.

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE*, *7*, e37135.

Pew, J., Muir, P. H., Wang, J., & Frasier, T. R. (2015). related: An R package for analysing pairwise relatedness from codominant molecular markers. *Molecular Ecology Resources*, *15*, 557–561.

Picq, S., McMillan, W. O., & Puebla, O. (2016). Population genomics of local adaptation versus speciation in coral reef fishes (*Hypoplectrus* spp., Serranidae). *Ecology and Evolution*, *6*, 2109–2124.

Planes, S. (1993). *Acanthurus triostegus* in French Polynesia. *Marine Ecology Progress Series*, *98*, 237–246.

Planes, S. (2002). Biogeography and larval dispersal inferred from population genetic analysis. In P. F. Sale (Ed.), *Coral reef fishes: Dynamics and diversity in a complex system* (pp. 201–220). New York: Academic Press.

Planes, S., & Lenfant, P. (2002). Temporal change in the genetic structure between and within cohorts of a marine fish, *Diplodus sargus*, induced by a large variance in individual reproductive success. *Molecular Ecology*, *11*, 1515–1524.

Planes, S., Parroni, M., & Chauvet, C. (1998). Evidence of limited gene flow in three species of coral reef fishes in the lagoon of New Caledonia. *Marine Biology*, *130*, 361–368.

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

Puebla, O., Bermingham, E., & McMillan, W. O. (2014). Genomic atolls of differentiation in coral reef fishes (*Hypoplectrus* spp., Serranidae). *Molecular Ecology*, *23*, 5291–5303.

Purcell, J. F., Cowen, R. K., Hughes, C. R., & Williams, D. A. (2006). Weak genetic structure indicates strong dispersal limits: A tale of two coral reef fish. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, *273*, 1483–1490.

R Core Team. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Riginos, C., Buckley, Y. M., Blomberg, S. P., & Treml, E. A. (2014). Dispersal capacity predicts both population genetic structure and species richness in reef fishes. *The American Naturalist*, *184*, 52–64.

Rocha, L. A., Robertson, D. R., Roman, J., & Bowen, B. W. (2005). Ecological speciation in tropical reef fishes. *Proceedings of the Royal Society of London Series B*, *272*, 573–579.

Rosenberg, N. A. (2004). DISTRUCT: A program for the graphical display of population structure. *Molecular Ecology Notes*, *4*, 137–138.

Rueger, T., Gardiner, N. M., & Jones, G. P. (2014). Relationships between pair formation, site fidelity and sex in a coral reef cardinalfish. *Behavioural Processes*, *107*, 119–126.

Saenz-Agudelo, P., Dibattista, J. D., Piatek, M. J., Gaither, M. R., Harrison, H. B., Nanninga, G. B., & Berumen, M. L. (2015). Seascape genetics along environmental gradients in the Arabian Peninsula: Insights from ddRAD sequencing of anemonefishes. *Molecular Ecology*, *24*, 6241–6255.

Sale, P. F., Cowen, R. K., Danilowicz, B. S., Jones, G. P., Kritzer, J. P., Lindeman, K. C., ... Steneck, R. S. (2005). Critical science gaps impede use of no-take fishery reserves. *Trends in Ecology & Evolution*, *20*, 74–80.

Seeb, J. E., Carvalho, G., Hauser, L., Naish, K., Roberts, S., & Seeb, L. W. (2011). Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. *Molecular Ecology Resources*, *11*, 1–8.

Selkoe, K. A., Gaines, S. D., Caselle, J. E., & Warner, R. R. (2006). Current shifts and kin aggregation explain genetic patchiness in fish recruits. *Ecology*, *87*, 3082–3094.

Shulman, M. J. (1998). What can population genetics tell us about dispersal and biogeographic history of coral-reef fishes? *Australian Journal of Ecology*, *23*, 216–225.

Shulman, M. J., & Bermingham, E. (1995). Early life histories, ocean currents, and the population genetics of Caribbean reef fishes. *Evolution*, *49*, 897–910.

Simpson, S. D., Meekan, M. G., Jeffs, A., Montgomery, J. C., & McCauley, R. D. (2008). Settlement-stage coral reef fishes prefer the higher frequency invertebrate-generated audible component of reef noise. *Animal Behavior*, *75*, 1861–1868.

Sponaugle, S., Cowen, R. K., Shanks, A., Morgan, S. G., Leis, J. M., Pineda, J., ... Munro, J. L. (2002). Predicting self-recruitment in marine populations: Biophysical correlates and mechanisms. *Bulletin of Marine Science*, *70*, 341–375.

Staaterman, E., Paris, C. B., & Helgers, J. (2012). Orientation behavior in fish larvae: A missing piece to Hjort's critical period hypothesis. *Journal of Theoretical Biology*, *304*, 188–196.

Stockwell, B. L., Larson, W. A., Waples, R. K., Seeb, L. W., & Carpenter, K. E. (2016). The application of genomics to inform conservation of a functionally important reef fish (*Scarus niger*) in the Philippines. *Conservation Genetics*, *17*, 239–249.

Swearer, S. E., Caselle, J. E., Lea, D. W., & Warner, R. R. (1999). Larval retention and recruitment in an island population of a coral-reef fish. *Nature*, *402*, 799–802.

Swearer, S. E., Shima, J. S., Hellberg, M. E., Thorrold, S. R., Jones, G. P., Robertson, D. R., ... Warner, R. R. (2002). Evidence of self-recruitment in demersal marine populations. *Bulletin of Marine Science*, *70*, 251–271.

Tamura, R. (1982). Experimental observations on the association between the cardinalfish (*Siphamia versicolor*) and the sea urchin (*Diadema setosum*). *Galaxea*, *1*, 1–10.

Taylor, M. S., & Hellberg, M. E. (2003). Genetic evidence for local retention of pelagic larvae in a Caribbean reef fish. *Science*, *299*, 107–109.

Terry, A., Bucciarelli, G., & Bernardi, G. (2000). Restricted gene flow and incipient speciation in disjunct Pacific Ocean and Sea of Cortez populations of a reef fish species, *Girella nigricans*. *Evolution*, *54*, 652–659.

Treml, E. A., Ford, J. R., Black, K. P., & Swearer, S. E. (2015). Identifying the key biophysical drivers, connectivity outcomes, and metapopulation consequences of larval dispersal in the sea. *Movement Ecology*, *3*, 1–16.

Tu, J. Y., Chou, C., & Chu, P. S. (2009). The abrupt shift of typhoon activity in the vicinity of Taiwan and its association with western North Pacific-East Asian climate change. *Journal of Climate*, *22*, 3617–3628.

Urbanczyk, H., Ogura, Y., Hendry, T. A., Gould, A. L., Kiwaki, N., Atkinson, J. T., ... Dunlap, P. V. (2011). Genome sequence of *Photobacterium mandapamensis* strain svers. 1.1, the bioluminescent symbiont of the cardinal fish *Siphamia versicolor*. *Journal of Bacteriology*, *193*, 3144–3145.

Vagelli, A., Burford, M., & Bernardi, G. (2008). Fine scale dispersal in Banggai Cardinalfish, *Pterapogon kauderni*, a coral reef species lacking a pelagic larval phase. *Marine Genomics*, *1*, 129–134.

Willing, E. M., Dreyer, C., & Van Oosterhout, C. (2012). Estimates of genetic differentiation measured by F_{ST} do not necessarily require large sample sizes when using many SNP markers. *PLoS ONE*, *7*, e42649.

Yasuda, N., Nagai, S., Hamaguchi, M., Okaji, K., Gerard, K., & Nadaoka, K. (2009). Gene flow of *Acanthaster planci* (L.) in relation to ocean currents revealed by microsatellite analysis. *Molecular Ecology*, *18*, 1574–1590.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Gould AL, Dunlap PV. Genomic analysis of a cardinalfish with larval homing potential reveals genetic admixture in the Okinawa Islands. *Mol Ecol*. 2017;26:3870–3882. https://doi.org/10.1111/mec.14169