Author Manuscript

1
2 MS. ALISON L. GOULD (Orcid ID : 0000-0002-2358-303X)
3
4
5 Article type : Original Article
6
7
8 **Genomic analysis of a cardinalfish with larval homing potential reveals genetic admixture**
9 **in the Okinawa Islands**
10
11 A. L. Gould[1] and P.V. Dunlap[1]
12
13 [1] *Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI*
14 *48109, U.S.A.*
15
16
17
18 **Corresponding author:**
19 Alison L. Gould
20 142 Life Sciences Addition #3200
21 Berkeley, CA 94720-3200
22 Phone: 1-415-350-5708
23 E-mail: algould@umich.edu
24
25
26 **Keywords**: RAD-Seq, SNPs, population genetics, *Siphamia tubifer*, bioluminescent
27
28 **Running Title:** Population genomics of *Siphamia tubifer*

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/mec.14169

This article is protected by copyright. All rights reserved

29

30 **Abstract**

31

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

51

52 **Introduction**

53

54 The degree to which coral reef fish populations are connected, which has large
55 consequences for their persistence and resilience to disturbances, has been disputed for decades.
56 For most reef fishes, dispersal occurs during a planktonic larval phase that lasts from a few days
57 to months, creating the potential for large dispersal distances (Barlow 1981; Shulman &
58 Bermingham 1995), but the challenges associated with tracking the direct movement of relatively

Author Manuscript

This article is protected by copyright. All rights reserved

59 small larvae in a dynamic, fluid environment has limited our knowledge of the connectivity
60 patterns of most reef fishes (Jones *et al.* 2009). Recently, however, the rapid advancement of
61 genetic tools has facilitated the study of reef fish populations, revealing the wide variation of
62 spatial scales at which marine populations exhibit connectivity (e.g. Terry *et al.* 2000; Planes
63 2002; Jones *et al.* 2005; Taylor & Hellberg 2003; Purcell *et al.* 2006; Gerlach *et al.* 2007; Horne
64 *et al.* 2008; van der Meer *et al.* 2012). Most coral reef fish metapopulations lie somewhere in the
65 middle of the spectrum between having relatively closed (low connectivity and high self-
66 recruitment) to open (high connectivity and no self-recruitment) populations (Cowen *et al.* 2000;
67 Mora & Sale 2002; Jones *et al.* 2009). Nonetheless, a surprising number of studies have provided
68 evidence that reef fish populations are more closed than expected and exhibit a significant degree
69 of local recruitment (Jones *et al.* 1999, 2005; Swearer *et al.* 1999, 2002; Cowen *et al.* 2000, 2006;
70 Bode *et al.* 2006; Almany *et al.* 2007) and genetic differentiation at relatively small spatial scales
71 (e.g. Planes 1993; Planes *et al.* 1998; Taylor & Hellberg 2003; Gerlach *et al.* 2007).
72 In response to the growing evidence of somewhat restricted gene flow among reef fish
73 populations, the links between early life history traits, larval behavior, and patterns of larval
74 recruitment and dispersal have been examined. Contrary to expectations, levels of genetic
75 differentiation of marine fishes are generally not necessarily correlated with pelagic larval
76 duration (PLD) or spawning mode (Shulman 1998; Bowen *et al.* 2006; Galarza *et al*. 2009; Jones
77 *et al.* 2009), and genetic structure in reef fish populations is often observed at smaller scales than
78 predicted by larval dispersal alone (e.g. Taylor & Hellberg 2003; Rocha *et al*. 2005; but see
79 Riginos *et al* 2014). These discrepancies can often be explained by larval fish behavior, through
80 which larvae employ some control over their dispersal trajectories (Kingsford *et al.* 2002; Leis &
81 McCormick 2002; Sponaugle *et al.* 2002; Leis *et al.* 2006; Montgomery *et al.* 2006), including
82 larval swimming ability (Fisher *et al.* 2005), vertical migration (Paris & Cowen 2004), and the
83 use of navigational cues, e.g., chemical and acoustic cues in the environment (Kingsford *et al.*
84 2002; Leis *et al.* 2003, 2011; Dixson *et al.* 2008; Simpson *et al.* 2008; Atema 2012; Paris *et al.*
85 2013). Studies of larval fish behavior suggest that larvae are not simply passive particles in the
86 plankton, but instead can actively orient and navigate to settlement sites. Therefore, incorporating
87 larval behavior into dispersal models can dramatically alter projected population connectivity
88 patterns (e.g. Paris & Cowen 2004; Cowen *et al.* 2006; Sale *et al.* 2005; Leis 2007; Paris *et al.*

Author Manuscript

This article is protected by copyright. All rights reserved

89 2007; Staaterman *et al*. 2012). Nevertheless, few studies have linked larval traits and behaviors to
90 gene flow and levels of population genetic differentiation, and it remains unknown how
91 conserved such links may be within fish families and geographic regions.

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

108 To better understand the link between larval behavior and gene flow in cardinalfishes, we
109 examined the population genetic structure of the sea urchin cardinalfish *Siphamia tubifer*, for
110 which homing and olfactory preferences have been previously described (Gould *et al.* 2014,
111 2015). Similar to other cardinalfishes, *S. tubifer* adults and juveniles exhibit fidelity to a home
112 site and return to a home reef from displacement distances of at least two kilometers (Gould *et al.*
113 2014). As described for *O. doederleini* (Gerlach *et al.* 2007), newly settled *S. tubifer* also exhibit
114 a preference for the olfactory cues of their home reef to that of a foreign reef (Gould *et al.* 2015).
115 These findings suggest the potential for *S. tubifer* larvae to use olfaction to recognize and return
116 to their natal reef. Additionally, *S. tubifer* are short-lived (typically <200 days) like *O.*
117 *doederleini*, but with a slightly longer PLD of up to 30 days (Kingsford *et al.* 2014; Gould *et al.*
118 2016). Despite these similarities, *S. tubifer* is distinct among cardinalfishes in that it is

This article is protected by copyright. All rights reserved

Author Manuscript

119 symbiotically bioluminescent, hosting a dense population of the luminous bacterium
120 *Photobacterium mandapamensis* in an abdominal light organ and using the bacterially-emitted
121 light while foraging at night (Iwai 1958; Dunlap & Nakamura 2011). The host fish acquires its
122 luminous symbiont from the environment during larval development, however the timing and
123 location of symbiont acquisition in the wild remain unknown (Dunlap *et al.* 2012). Describing
124 patterns of gene flow in *S. tubifer* at various geographic scales and over time can provide insight
125 on population connectivity and the location of symbiont acquisition by developing larvae.

126 We applied double digest restriction site-associated sequencing (ddRAD-Seq) methods to
127 look for evidence of genetic differentiation among groups of *S. tubifer* collected over a three-year
128 period from various locations in Okinawa, Japan. We used ddRAD-Seq methods as they do not
129 require an extensive marker discovery process and enable the development of thousands of
130 genomic markers without prior genetic data for the focal species (Davey & Blaxter 2010; Seeb *et*
131 *al.* 2011), and they can be used to infer genetic differentiation with fine-scale precision (Luikart
132 2003; Coates *et al.* 2009; Willing *et al.* 2012; Bradbury *et al.* 2015). Our specific aims were to
133 test for patterns of population genetic differentiation between groups of *S. tubifer* collected at
134 various spatial scales and to test the stability of *S. tubifer* populations by examining temporal
135 patterns of genetic divergence at a particular reef. If the link between larval olfactory preferences,
136 homing potential, and self-recruitment are conserved across similar cardinalfish species, we
137 predicted that similar to *O. doederleini,* populations of *S. tubifer* would have significant
138 population genetic structure between sampling sites but less genetic differentiation at the same
139 reef over time.

140

141 **Materials and Methods**

142

143 *Sampling*

144

145 A total of 295 *Siphamia tubifer* were collected from different locations (approximately 20
146 individuals per location) over three years in the Okinawa Islands, Japan. Ten of these locations
147 were sampled during the summer of 2013, three of which were again sampled in 2014, and one
148 location was sampled in three consecutive years, 2012, 2013, and 2014 (Fig. 1). Fish of various

Author Manuscript

This article is protected by copyright. All rights reserved

149 sizes associated with several different host sea urchins and from a broad sampling area were
150 collected captured at each location (Table 1). Upon collection, fish were immediately euthanized
151 and placed on ice**.** The intact light organ of each fish was then aseptically dissected and
152 individually preserved in RNAlater®, and the remainder of the fish specimen was stored in 98%
153 ethanol at -20°C.
154
155 *DNA extraction and library preparation*
156
157 Genomic DNA was extracted from intact, preserved light organs, which are comprised of
158 fish tissue and the symbiotic population of luminous bacteria, using QIAGEN DNeasy Blood and
159 Tissue Kits and following the manufacturer's protocol. A total of six ddRAD-Seq DNA libraries
160 were constructed, each from the genomic DNA of up to 50 *S. tubifer* light organs, following a
161 modified combination of the methods described in Parchman *et al*. (2012) and Peterson *et al*.
162 (2012). For each library, approximately 200 ng of genomic DNA from each light organ was
163 digested with the high-fidelity restriction enzymes *MseI* and *EcoRI* at 37°C for three hours. A
164 standardized concentration per library of each digestion product was then ligated to a uniquely
165 barcoded Illumina adaptor at the *EcoRI* cut site and an Illumina adaptor at the *MseI* cut site. The
166 ligation products were individually amplified with the Illumina Illpcr1 and Illpcr2 primers in two
167 20 μl PCR reactions per sample with a 98°C start, 12 cycles of 98°C for 20 seconds, 65°C for 30
168 seconds, 72°C for 40 seconds, followed by 10 minutes at 72°C, after which, the PCR products
169 from all samples within a library were pooled and concentrated to an approximate volume of 150
170 μl. Samples were purified with Agencourt AMPure XP magnetic beads following standard
171 protocols after the digestion, ligation, and PCR steps. The pooled, purified PCR products were
172 then size-selected between 300-400 bp on a Pippin Prep (Sage Science) machine, and the size-
173 selected DNA libraries were each sequenced in one lane on the Illumina HiSeq2000 platform
174 (San Diego, CA) at the Center for Applied Genomics, Toronto, ON, Canada, to generate 100 bp,
175 single-end sequence reads.
176
177 *Sequence analysis and processing*
178

This article is protected by copyright. All rights reserved

Author Manuscript

179 Raw sequence reads were quality filtered and processed primarily using the program
180 *Stacks* v. 1.35 (Catchen *et al*. 2011, 2013). Raw reads were demultiplexed, trimmed to 90 bp, and
181 quality filtered for a Phred score of 33 or higher using the *process_radtags* command in *Stacks*.
182 To distinguish sequence reads that belonged to the host fish from those of the bacterial symbiont,
183 we used the '*very_sensitive*' command in *Bowtie2* v. 2.2.0 (Langmead & Salzberg 2012) to filter
184 all reads against the reference genome of *Photobacterium mandapamensis* (Urbanczyk *et al*.
185 2011). To ensure that all bacterial reads were removed, we also filtered all reads against the
186 genomes of *Escherichia coli* K12 (Durfee *et al*. 2008) and *Vibrio campbellii* (Lin *et al*. 2010) in
187 the same manner. Sequence reads that did not align to the bacterial genomes were assigned as fish
188 (*S. tubifer*) sequences.
189 Fish sequence reads were processed and assembled *de novo* to call single nucleotide
190 polymorphisms (SNPs) using the *Stacks* pipeline with the *rxstacks* correction step. We first ran
191 the *denovo_map* program with the parameters -m 3, -M 2, and -n 3, optimized to prevent over-
192 and under- merging of homologous loci. These parameters were previously recommended to
193 increase the number of loci but minimize genotyping error for the *de_novo* assembly of other
194 RAD datasets (Mastrett-Yanes *et al.* 2015). We then implemented the *rxstacks* correction step
195 using a bounded model (--bound_high 0.1) in which excess haplotypes were pruned, and loci for
196 which 25% of individuals had a confounded match in the catalog (--conf_lim 0.25) or an average
197 log likelihood less than (-10.0) were removed. After running this correction step, the *cstacks* (-n
198 3) and *sstacks* programs were re-applied to produce the final set of RAD tags across all
199 individuals in the study.
200
201 *Population analysis*
202
203 The mean depth of coverage per locus across all individuals was determined from the
204 *Stacks* output files, and individuals with a mean coverage less than 10 across all loci were
205 excluded from the analyses. Population summary statistics were computed with the *populations*
206 program in *Stacks* for loci present in at least 10 populations and in 70% of individuals per
207 population. Population differentiation was evaluated with pairwise F_{ST} values calculated in the
208 *populations* program in *Stacks* for loci present in all populations and in 70% of individuals per

This article is protected by copyright. All rights reserved
Author Manuscript

209 population with a minor allele frequency greater that 5%. Pairwise AMOVA F_{ST} values were also
210 computed in *GenoDive* (Meirmans & Van Tienderen 2004) with 1,000 bootstrap resampling
211 steps, producing *P*-values for all pairwise population comparisons. Isolation by distance (IBD)
212 was assessed by conducting Mantel tests with the Isolation by Distance Web Service (Jensen *et*
213 *al*. 2005) on pairwise F_{ST} values and the natural logarithm of the shortest distances over water
214 between sites; one-sided *P*-values were calculated by randomizing the data 30,000 times. Using
215 the *hierfstat* program (Goudet 2005) in R v. 3.1.1 (R Core Team 2014), the per locus F_{ST} values
216 were calculated and compared for the following datasets: individuals collected in 2013, 2014, and
217 in three consecutive years from the Sesoko (S) site. Loci in each dataset were then split into 10%
218 quantiles based on their estimated F_{ST} values and analyzed independently as outliers.
219 To visualize genetic structure, principal components analyses (PCAs) were implemented
220 on the loci present in at least 70% of individuals per population and in at least 10 populations
221 with a minor allele frequency greater that 5% with the *dudi.pca* function in the *adegenet* v. 1.4.2
222 package (Jombart 2008; Jombart & Ahmed 2011) in R (R Core Team 2014). The small
223 percentage of missing data values per locus (an average of 4.1 ± 4.4%, S.D.) were replaced with
224 the mean value across the entire dataset at that locus. Analyses of molecular variance (AMOVA)
225 (Excoffier *et al*. 1992) were carried out in *GenoDive* to test for genetic differences between
226 populations and region (populations on the east and west coast of Okinawa). Both PCAs and
227 AMOVAs were performed separately on the 2013 and 2014 populations as well as on the three-
228 year dataset from Sesoko (S) (Fig. 1).
229 Complimentary clustering analyses were also performed with the program *STRUCTURE*
230 v. 2.3.4 (Pritchard *et al*. 2000) using the output data files from *Stacks* comprised of only the first
231 SNP per locus (to eliminate any SNPs that are linked within the same RAD site from the
232 analysis) for loci present in at least 70% of all individuals and in all populations for each dataset.
233 Group assignments in *STRUCTURE* were made using the admixture model with 100,000 burn-in
234 steps and 100,000 MCMC iterations for each number of pre-determined genotypic groups (K).
235 Analyses were repeated ten times for each value of K. For the 2014 and Sesoko datasets, K was
236 set from 1 to one more than the total number of populations sampled (5 and 4, respectively), and
237 for the 2013 dataset, K was set from 1 to 10 (the total number of sites sampled). Probable K
238 values were inferred by examining the change in the posterior probability of the log likelihood

Author Manuscript

This article is protected by copyright. All rights reserved

239 across all K values (ln P(X|K)) and by applying the Evanno Δ*K* method (Evanno *et al*. 2005) with
240 *STRUCTURE HARVESTER* (Earl 2012). All *STRUCTURE* results were visualized with the
241 program *DISTRUCT* (Rosenberg 2004).
242
243 *Outlier analysis*
244
245 An additional outlier approach was used to identify sets of loci with significantly higher
246 F_{ST} values than expected under a neutral model of selection for each dataset (2013, 2014, and
247 Sesoko) with the program *LOSITAN* (Antao *et al.* 2008). A subset of one SNP per locus present
248 in all populations and in 70% of individuals within a population with a minor allele frequency
249 greater that 5% were examined using the following parameters in *LOSITAN*: 50,000 simulations,
250 a confidence interval of 0.99, a false discovery rate of 0.1, and a subsample size of 30, using the
251 "neutral means F_{ST}" and "force means F_{ST}" options, which iteratively identify and remove F_{ST}
252 outliers when calculating the global F_{ST} distribution, turned on. Based on their probabilities of
253 being under selection, loci were then classified into the following categories: under divergent
254 selection ($P > 99\%$), neutral ($1\% < P < 90\%$), or under balancing selection ($P < 0.1\%$); all
255 remaining loci were conservatively considered to be unclassified. To visualize whether putative
256 loci under selection show any patterns concordant with sampling site or time, PCAs were
257 performed independently on each subset of outlier loci as previously described. We also
258 compared the outlier loci identified in *LOSITAN* to sets of loci within various per locus F_{ST}
259 quantiles (90-100%, 80-90%, and 70-80%).
260
261 *Ethics statement*
262
263 The protocols used here for the capture and handling of fish specimens were approved by
264 the University of Michigan's University Committee for the Use and Care of Animals
265 (PRO00004825), and they followed the requirements outlined in the University of the Ryukyus'
266 Guide for Care and Use of Laboratory Animals (Dobutsu Jikken Kisoku, version 19.6.26).
267
268 **Results**

Author Manuscript

This article is protected by copyright. All rights reserved

269

270 *Sequence analysis and processing*

271

272 The six ddRAD libraries each produced high quality sequence data with sufficient depth
273 of coverage across most individuals for population-level genetic analyses. Thirteen individuals
274 that had fewer than 800,000 remaining reads after quality filtering (Fig. S1) were discarded from
275 the analysis. On average, 87.05 ± 2.58% of all reads were retained from each library, with an
276 average of 9.72 ± 2.29% and 3.23 ± 2.51% of reads additionally discarded for having ambiguous
277 barcodes or RADtags or for low quality, respectively (Table S1). The final dataset consisted of
278 11,836 loci across 280 individuals and contained low percentages of missing data. On average,
279 4.1% of data was missing per locus (min = 0%, max = 22.1%) and 4.1% of data was missing per
280 individual (min = 0%, max = 13.4%).

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

293

294 *Population summary statistics*

295

296 For RAD sites that were polymorphic in at least one population (Table 2), the average
297 major allele frequency (P) and observed heterozygosity (H_{obs}) across all populations ranged from
298 0.9593 to 0.9607 and 0.0485 to 0.0519, respectively. When sites that were fixed across all

This article is protected by copyright. All rights reserved

Author Manuscript

299 populations were included, P increased up to 0.9983 and H_{obs} values all decreased to 0.0022
300 (Table S2). Across variant RAD1sites only, the percentages of polymorphic loci in populations
301 sampled in 2013 were between 30.92% and 36.09%, whereas the percentages of polymorphic loci
302 in populations sampled in 2014 were slightly higher, ranging from 36.79% to 41.75% (Table 2).
303 Levels of overall nucleotide diversity (π) across all sites, fixed and variant, were similar for all
304 groups of fish sampled, varying between 0.0026 and 0.0028. Average measures of Wright's
305 inbreeding coefficient (F_{IS}) calculated for all variant sites ranged from 0.0407 to 0.0567 across all
306 populations (Table 2). These F_{IS} values are all slightly positive indicating a deficiency in
307 heterozygotes within each population.
308
309 *Analysis of genetic differentiation*
310
311 Population genetic analyses of the identified 11,836 SNPs revealed no significant genetic
312 structure between sampling sites or between years. In 2013, pairwise F_{ST} values were consistently
313 low between sites, ranging between 0.0157 and 0.0214, with only a few significant values (Table
314 3), most of which were for comparisons with fish from the Motobu (M) site (Fig. 1). Pairwise F_{ST}
315 values for the 2014 samples were similarly low and non-significant, ranging from 0.0116 to
316 0.0139 (Table S3). An analysis of temporal genetic differentiation at the Sesoko (S) site (Fig. 1)
317 over three consecutive years also revealed no significant differentiation over time (F_{ST} values
318 ranged from 0.0158 to 0.0177, Table 4). Similarly the other two sites that were sampled in
319 consecutive years, Itoman (It) and Ikei (Ik) (Fig. 1), had low, non-significant F_{ST} values (It: F_{ST} =
320 0.0151, Ik: F_{ST} = 0.0165) between years. Per locus F_{ST} values for each dataset were all low with
321 seemingly normal distributions around zero (Fig. S2) and are summarized in Table S4 including
322 the values used to parse loci into various quantile groups.
323 Analyses of Isolation by distance revealed no significant relationship between the low
324 levels of population differentiation observed and geographic distance between locations within
325 either sampling year (2013: $F_{1,43} = 0.369$, $R^2 = 0.0085$, $P = 0.547$; 2014: $F_{1,4} = 3.372$, $R^2 = 0.457$
326 $P = 0.140$; Fig. S3). Similarly, results from the AMOVA indicate that all of the observed genetic
327 variation is attributed to variation within (F_{IT}) and among (F_{IS}) individuals and none is attributed

This article is protected by copyright. All rights reserved

328 to differences between populations or between the east and west coast regions of Okinawa Island
329 (Table 5).
330 Principal components analyses of all 11,836 SNPs confirmed the lack of genetic structure,
331 with no apparent clustering of individuals by location (Fig. 2). The first two PC axes for the 2013
332 analysis each described less than 1% of the variation in the data. These values rose slightly for the
333 2014 data, accounting for a combined (PC1 and PC2) total of 2.57% of the total genetic variation
334 (Fig. 2). The PCA carried out on the temporal dataset from the Sesoko site also indicates a lack of
335 structure over time (Fig. 2), and it is consistent with the low, non-significant pairwise F_{ST} values
336 between groups of fish collected at that site in subsequent years (Table 3). Similarly, clustering
337 analyses in *STRUCTURE* revealed that a K = 1 is most likely for the 2013, 2014, and Sesoko
338 datasets when all loci were considered (Tables S5-7). In the case where K = 1, the Evanno ΔK
339 method of detecting the true value of K is ineffective (Evanno *et al.* 2005). We therefore
340 examined the mean posterior probabilities for each K value in both datasets; the log likelihood
341 was highest when K = 1 for all datasets (Tables S5-7). These results provide further evidence that
342 one panmictic population of *S. tubifer* is present in the Okinawa Islands, including Kume Island
343 (K) one hundred kilometers to the west (Fig. 1).
344
345 *Outlier analysis*
346
347 The *LOSITAN* outlier analyses identified a small percentage of loci (0.7% - 2.8%) within
348 each dataset that were classified as being under divergent selection (Fig. 3a-c, Table 6). The
349 percentages of shared loci under selection between the 2013 and 2014 *LOSITAN* datasets were
350 2% and 8%, respectively, and the number of shared outliers between the 2013 and 2014 datasets
351 and the Sesoko dataset were 12% and 6%. The numbers of shared loci under divergent selection
352 between sampling years within the Sesoko dataset was higher, ranging from 12% to 26%.
353 Principal components analyses of these adaptive loci only revealed a slight signature of genetic
354 differentiation between individuals collected at the broader geographic range sampled in 2014,
355 but less so between individuals collected from reefs around Okinawa Island in 2013 (Fig. 3d,e).
356 Interestingly, there was more apparent differentiation at these putative loci under selection
357 between individuals collected in different years at Sesoko (S) (Fig. 3f).

Author Manuscript

This article is protected by copyright. All rights reserved

358 Many of the loci classified as being under divergent selection were also present in the 90-
359 100% F_{ST} quantile datasets; 85%, 91%, and 84% of the loci classified under as being under
360 divergent selection in *LOSITAN* were also in the 2013, 2014, and Sesoko F_{ST} outlier datasets,
361 respectively. Additional *STRUCTURE* analyses of subsets of SNPs in each 10% F_{ST} quantile
362 (Table S4) confirmed a lack of genetic structure for groups of fish collected from sites around
363 Okinawa Island in 2013, even when examining only the loci with the highest levels of
364 differentiation (90-100% F_{ST} quantile) observed across all individuals (Fig. S4a). A weak
365 signature of genetic clustering that matched sampling location was evident for the most highly
366 differentiated subset of loci (90-100% F_{ST} quantile) for the 2014 dataset, and both the highest ln
367 P(K) and a rise in ΔK were evident when K = 4 (Table S5). However, this pattern of genetic
368 clustering was not apparent for the subset of loci in the 80-90% F_{ST} quantile for the same
369 populations (Fig. S4b). The most highly differentiated subset of loci (90-100% F_{ST} quantile) in
370 the Sesoko dataset revealed three genetic clusters that matched sampling year (Fig. S4c) and
371 corroborate with the temporal pattern of differentiation observed with the adaptive set of loci
372 previously identified. Moreover, principal components analyses of the *LOSITAN* outliers
373 produced comparable results to analyses of the 90-100% F_{ST} outliers for each dataset (Fig. S5)
374 with the exception of the 2013 dataset; the analysis of *LOSITAN* outliers revealed a weak pattern
375 by sampling location that was not as evident for the 90-100% F_{ST} outlier dataset from 2013.
376
377
378 **Discussion**
379
380 The observed discrepancies between potential and actual dispersal distances of reef fish
381 larvae (e.g. Taylor & Hellberg 2003; Rocha *et al*. 2005; Bowen *et al.* 2006; Galarza *et al*. 2009;
382 Jones *et al.* 2009) highlight the need for a better understanding of the influence of larval fish
383 behavior on the connectivity of reef fish populations. To narrow this gap in knowledge, general
384 links between larval behavior and gene flow across diverse groups of fishes and geographic
385 regions are required. We applied RAD-Seq methods to test for genetic differentiation between
386 populations of *S. tubifer*, as documented for another cardinalfish species with similar life history
387 traits and behavior (Gerlach *et al.* 2007, Kingsford *et al.* 2014). Despite the potential of RAD-Seq

Author Manuscript

This article is protected by copyright. All rights reserved

388 methods to detect fine-scale genetic structure, and in contrast to other genetic studies of
389 cardinalfishes, we found little evidence of genetic differentiation between groups of *S. tubifer* at
390 spatial scales up to 140 kilometers. This lack of differentiation was particularly striking in that
391 collection sites included both the east and west sides of Okinawa Island and a site, Kume Island,
392 well separated to the west of Okinawa Island (Fig. 1). Analyses of adaptive outlier loci, however,
393 revealed a shallow signature of genetic divergence between sampling sites, particularly at the
394 larger spatial scales examined, but this signature was evident only for a small number of the most
395 highly differentiated loci. The overall pattern of genetic admixture observed from the analysis of
396 all identified SNPs indicates a significant amount of gene flow among groups of *S. tubifer* in the
397 Okinawa Islands, presumably due to larval dispersal and mixing by strong ocean currents.
398 Although these results cannot exclude the possibility of natal homing, they demonstrate that
399 genetic exchange between populations occurs frequently enough to maintain genetic homogeneity
400 in the region.

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

Author Manuscript

This article is protected by copyright. All rights reserved

418 Leis 2002), therefore if few, far-dispersing *S. tubifer* larvae settle sporadically on reefs, there
419 would be little potential for genetic divergence to accumulate between populations over time.
420 Corresponding with substantial gene flow in the region, the islands in Japan's Ryukyu
421 Archipelago receive larval supply from other reef habitats in the south. In particular, small reef
422 fish from the Philippines have an ecologically significant linkage potential to the Ryukyu Islands
423 (Treml *et al.* 2015). The connectivity potential from the northern Phillipines is driven by the
424 strong ocean currents in the region, namely the Kuroshio Current, which originates off the coast
425 of the Philippines and flows northward through the Ryukyu Archipelago (Fig. 1). Previous
426 studies in the region have established evidence of genetic homogeneity among the Philippine
427 Islands and across the Ryukyu Islands along the Kuroshio Current for the crown-of-thorns sea
428 star *Acanthaster planci* (Yasuda *et al.* 2009) and for the broadcast-spawning coral *Acropora*
429 *digitifera* (Nakajima *et al.* 2010). Moreover, at smaller spatial scales within Okinawa, there is
430 evidence of significant gene flow for several coral species (Nishikawa *et al.* 2003; Nishikawa
431 2008), although two distinct genetic clusters between the North and South regions of Okinawa
432 Island were apparent for the scleractinian coral *Goniastrea aspera*, indicating a substantial
433 number of locally-produced recruits and the potential for self-recruitment in the region despite the
434 strong influence of the Kuroshio Current (Nishikawa & Sakai 2005).
435 In addition to spatial differentiation, we examined temporal genetic divergence in *S.*
436 *tubifer* at one study site by collecting specimens over three consecutive years. The dynamics of
437 temporal genetic structure may be even more informative than spatial dynamics in marine
438 systems, but this issue has generally been overlooked for most marine populations (Hellberg *et al.*
439 2002; Hedgecock *et al.* 2007). Of the studies that have examined temporal structure, instances of
440 temporal stability have been rare (Bernal-Ramírez *et al.* 2003; Larsson *et al.* 2010), whereas
441 temporal genetic differentiation has been reported for several marine fishes (e.g. Planes &
442 Lenfant 2002; Maes *et al.* 2006; Selkoe *et al.* 2006; Klanten *et al.* 2007). Genetic differentiation
443 over time at a location can result from selection, random genetic drift, or from variable larval
444 supply from different source populations (Hedgecock *et al.* 2007). When examining all SNPs, we
445 observed low levels of differentiation between groups of *S. tubifer* collected from the same site
446 over three years. However, when we considered only outlier loci putatively under selection, a
447 clear pattern of temporal differentiation was revealed. To determine whether the observed

Author Manuscript

This article is protected by copyright. All rights reserved

448 differentiation at these adaptive loci might be attributed to cohesive cohort dispersal and
449 settlement, we compared relatedness values between individuals collected in the same year with
450 those collected in different years; however, we saw no significant differences in relatedness
451 between these groups of individuals (Fig. S6). We also compared the numbers of shared outlier
452 loci that were divergent between consecutive sampling years and found that approximately one-
453 fourth of the loci under selection are shared between the inter-annual comparisons. Therefore, the
454 observed temporal patterns of structure at these outlier loci can likely be attributed to variable
455 larval supply from different upstream source populations rather than solely due to genetic drift.

456 *Siphamia tubifer* is short-lived, with an expected longevity of less than 200 days (Gould *et*
457 *al.* 2016); therefore population turnover at a reef might occur regularly, especially given the
458 potential variability in larval supply consistent with large dispersal distances and high larval
459 mortality rates associated with reef fishes. Several divergent *S. tubifer* populations upstream of
460 the Okinawa Islands, such as in the Philippines, could therefore supply larvae that variably settle
461 on available urchin habitat and establish temporary populations at a given reef site. These
462 populations might subsequently be vulnerable to bottlenecks due to high adult mortality rates
463 (Gould *et al.* 2016) along with natural disturbance events, such as typhoons, which are known to
464 cause a decline in the numbers of some reef fishes, especially for demersal fish with specific
465 habitat requirements (Harmelin-Vivien 1994, Halford *et al.* 2004) like *S. tubifer*. Typhoon
466 occurrence in the study region has increased in recent years to an average of approximately six
467 typhoons per year (Tu *et al.* 2009). Additionally, potential seasonality in the breeding pattern of
468 *S. tubifer* could enhance the likelihood of population turnover at a reef site. Many fish species
469 have seasonal peaks in reproduction at higher latitude reefs associated with temperature
470 fluctuations (Munday *et al.* 2008), and seasonal reproduction has been documented for other reef-
471 associated fishes in Okinawa (Kuwamura *et al.* 1994). Although breeding seasonality in *S. tubifer*
472 has not been described, any seasonal differences in reproduction could also contribute to the
473 observed temporal patterns of differentiation for the subset of adaptive loci identified.

474 Overall our study supports the importance of examining the degree of genetic
475 differentiation within various geographic regions for a range of fish species as well as species-
476 specific patterns across groups of reef fishes to better understand the relationships between life
477 history traits, larval behavior, and gene flow. We highlight the effectiveness of RAD-Seq

Author Manuscript

This article is protected by copyright. All rights reserved

478 methods, which have recently been applied to examine the genomics of other coral reef fishes
479 (Puebla *et al.* 2014; Gaither *et al.* 2015; Saenz-Agudelo *et al.* 2015; Stockwell *et al.* 2016; Picq *et*
480 *al.* 2016), to study such non-model organisms. Our results show that in contrast to other
481 cardinalfish species (Bernardi & Vagelli 2004; Hoffman *et al.* 2005; Gerlach *et al.* 2007; Vagelli
482 *et al.* 2008; Gotoh *et al.* 2009), *S. tubifer* exhibits genetic admixture over a 100 kilometer region
483 despite the species' demonstrated homing abilities and olfactory preferences for a home reef
484 (Gould *et al.* 2015). These findings suggest that strong ocean currents combined with a month-
485 long pelagic larval phase promote dispersal and gene flow of *S. tubifer* in the region. We also
486 found evidence for temporal genetic differentiation at a small number of putative loci under
487 selection, suggesting adaptive variation in the source populations that supply *S. tubifer* larvae to
488 Okinawa. Future investigations *S. tubifer* populations at varying locations across this species'
489 broad, Indo-Pacific distribution will provide insight on whether various source populations are
490 undergoing divergent selection and whether the scale of genetic admixture observed here is
491 region-specific.
492
493
494 **Acknowledgements**
495
496 We gratefully acknowledge Saki Harii and the staff of Sesoko Station, University of the Ryukyus
497 for logistical assistance. We thank Katherine Dougan and Maggie Grundler, University of
498 Michigan (UM), for technical and field assistance, and we thank members of Lacey Knowles'
499 Lab (UM), especially Rob Massatti, for guidance and assistance with the RAD protocols and data
500 analysis. We also thank Raquel Rivadeneira and the UM Genomic Diversity Lab for technical
501 services. Funding for this project was provided by National Science Foundation grant DEB-
502 1405286, by a UM Rackham Graduate School Susan Lipschutz Award, and by a UM Museum of
503 Zoology Ichthyology Award.
504 **References**
505
506 Almany GR, Berumen ML, Thorrold SR *et al.* (2007) Local replenishment of coral reef fish
507 populations in a marine reserve. *Science*, **316**, 742-744.

Author Manuscript

This article is protected by copyright. All rights reserved

508 Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G (2008) LOSITAN: a workbench to detect
509 molecular adaptation based on a F_{ST}-outlier method. BMC Bioinformatics, **9**, 323.
510 Atema J (2012) Aquatic odor dispersal fields: opportunities and limits of detection,
511 communication and navigation. In: Chemical ecology in aquatic ecosystems (eds Bronmark
512 C, Hansson L-A), pp. 1-18. Oxford University Press, Oxford.
513 Barlow GW (1981) Patterns of parental investment, dispersal and size among coral-reef fishes.
514 *Environmental Biology of Fishes*. **6**, 65-85.
515 Bernal-Ramírez JH, Adcock GJ, Hauser L *et al.* (2003) Temporal stability of genetic population
516 structure in the New Zealand snapper, *Pagrus auratus*, and relationship to coastal currents.
517 *Marine Biology*, **142**, 567-574.
518 Bernardi G, Vagelli A (2004) Population structure in Banggai cardinalfish, *Pterapogon kauderni*,
519 a coral reef species lacking a pelagic larval phase. *Marine Biology*, **145**, 803-810.
520 Bode M, Bode L, Armsworth PR (2006) Larval dispersal reveals regional sources and sinks in the
521 Great Barrier Reef. *Marine Ecology Progress Series*, **308**, 17-25.
522 Bowen BW, Bass AL, Muss A *et al.* (2006) Phylogeography of two Atlantic squirrelfishes
523 (Family Holocentridae): exploring links between pelagic larval duration and population
524 connectivity. *Marine Biology*, **149**, 899-913.
525 Bradbury IR, Hamilton LC, Dempson B *et al.* (2015) Transatlantic secondary contact in Atlantic
526 Salmon, comparing microsatellites, a single nucleotide polymorphism array and
527 restriction- site associated DNA sequencing for the resolution of complex spatial structure.
528 *Molecular Ecology*, **24,** 5130-5144.
529 Catchen JM, Amores A, Hohenlohe PA *et al.* (2011) Stacks: building and genotyping loci de
530 novo from short-read sequences. *G3: Genes, Genomes, Genetics*, **1**, 171-182.
531 Catchen JM, Hohenlohe PA, Bassham S *et al.* (2013). Stacks: an analysis tool set for population
532 genomics. *Molecular Ecology*, **22**, 3124-3140.
533 Coates BS, Sumerford DV, Miller NJ *et al.* (2009) Comparative performance of single nucleotide
534 polymorphism and microsatellite markers for population genetic analysis. *Journal of*
535 *Heredity*, **100**, 556-564.
536 Cowen RK, Lwiza KM, Sponaugle S *et al.* (2000) Connectivity of marine populations: open or
537 closed?. *Science*, **287**, 857-859.

This article is protected by copyright. All rights reserved

Author Manuscript

538 Cowen RK, Paris CB, Srinivasan A (2006) Scaling of connectivity in marine populations.
539 *Science*, **311**, 522-527.
540 Davey JW, Blaxter ML (2010) RADSeq: next-generation population genetics. *Briefings in*
541 *Functional Genomics*, **9**, 416-423.
542 Dixson DL, Jones GP, Munday PL *et al.* (2008). Coral reef fish smell leaves to find island homes.
543 *Proceedings of the Royal Society of London B: Biological Sciences*, **275**, 2831-2839.
544 Døving KB, Stabell OB, Östlund-Nilsson S *et al.* (2006) Site fidelity and homing in tropical coral
545 reef cardinalfish: are they using olfactory cues?. *Chemical Senses*, **31**, 265-272.
546 Dumas P, Kulbicki M, Chifflet S *et al.* (2007). Environmental factors influencing urchin spatial
547 distributions on disturbed coral reefs (New Caledonia, South Pacific). *Journal of*
548 *Experimental Marine Biology and Ecology*, **344**, 88-100.
549 Dunlap PV, Nakamura M (2011) Functional morphology of the luminescence system of
550 *Siphamia versicolor* (Perciformes: Apogonidae), a bacterially luminous coral reef fish.
551 *Journal of Morphology*, **272**, 897-909.
552 Dunlap PV, Gould AL, Wittenrich ML *et al.* (2012) Symbiosis initiation in the bacterially
553 luminous sea urchin cardinalfish *Siphamia versicolor*. *Journal of Fish Biology*, **81**, 1340-
554 1356.
555 Durfee T, Nelson R, Baldwin S *et al.* (2008) The complete genome sequence of *Escherichia coli*
556 DH10B: insights into the biology of a laboratory workhorse. *Journal of Bacteriology*, **190**,
557 2597-2606.
558 Earl DA (2012) STRUCTURE HARVESTER: a website and program for visualizing
559 STRUCTURE output and implementing the Evanno method. *Conservation Genetics*
560 *Resources*, **4**, 359-361.
561 Eibl-Eibesfeldt I (1961) Eine Symbiose zwischen Fischen (*Siphamia versicolor*) und Seeigeln.
562 Zeitschrift fur Tierpsychologie, **18**, 56–59.
563 Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the
564 software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611-2620.
565 Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric
566 distances among DNA haplotypes: application to human mitochondrial DNA restriction data.
567 *Genetics*, **131**, 479-491.

Author Manuscript

This article is protected by copyright. All rights reserved

568 Fisher R, Leis JM, Clark DL *et al.* (2005) Critical swimming speeds of late-stage coral reef fish
569 larvae: variation within species, among species and between locations. *Marine Biology*, **147**,
570 1201-1212.
571 Gaither MR, Bernal MA, Coleman RR *et al.* (2015) Genomic signatures of geographic isolation
572 and natural selection in coral reef fishes. *Molecular Ecology*, **24**, 1543-1557.
573 Galarza JA, Carreras-Carbonell J, Macpherson E *et al.* (2009). The influence of oceanographic
574 fronts and early-life-history traits on connectivity among littoral fish species. *Proceedings of*
575 *the National Academy of Sciences*, **106**, 1473-1478.
576 Gerlach G, Atema J, Kingsford MJ *et al.* (2007) Smelling home can prevent dispersal of reef fish
577 larvae. *Proceedings of the National Academy of Sciences*, **104**, 858-863.
578 Gotoh RO, Sekimoto H, Chiba SN *et al.* (2009) Peripatric differentiation among adjacent marine
579 lake and lagoon populations of a coastal fish, *Sphaeramia orbicularis* (Apogonidae,
580 Perciformes, Teleostei). *Genes & Genetic Systems*, **84**, 287-295.
581 Goudet J (2005) Hierfstat, a package for R to compute and test hierarchical F- statistics.
582 *Molecular Ecology Notes*, **5**, 184-186.
583 Gould AL, Harii S, Dunlap PV (2014) Host preference, site fidelity, and homing behavior of the
584 symbiotically luminous cardinalfish, *Siphamia tubifer* (Perciformes: Apogonidae). *Marine*
585 *Biology*, **161**, 2897-2907.
586 Gould AL, Harii S, Dunlap PV (2015) Cues from the reef: olfactory preferences of a
587 symbiotically luminous cardinalfish. *Coral Reefs*, **34**, 673-677.
588 Gould AL, Dougan KE, Koeningbauer S, Dunlap PV (2016) Life history of a symbiotically
589 luminous cardinalfish. *Journal of Fish Biology*, **89**, 1359-1377.
590 Halford A, Cheal AJ, Ryan D, Williams DM (2004) Resilience to large-scale disturbance in coral
591 and fish assemblages on the Great Barrier Reef, Australia. *Ecology*, **85**, 1892–1905.
592 Harmelin-Vivien M (1994) The Effects of Storms and Cyclones on Coral Reefs: A
593 Review. *Journal of Coastal Research,* 211-231.
594 Hedgecock D, Barber PH, Edmands S (2007) Genetic approaches tomeasuring connectivity.
595 *Oceanography,* **20**, 70–79.
596 Hellberg ME, Burton RS, Neigel JE *et al.* (2002) Genetic assessment of connectivity among
597 marine populations. *Bulletin of Marine Science*, **70**, 273-290.

Author Manuscript

This article is protected by copyright. All rights reserved

598 Hoffman EA, Kolm N, Berglund A *et al*. (2005) Genetic structure in the coral‐ reef‐ associated Banggai cardinalfish, *Pterapogon kauderni*. *Molecular Ecology*, **14**, 1367-1375.

600 Hogan JD, Thiessen RJ, Heath DD (2010) Variability in connectivity indicated by chaotic genetic patchiness within and among populations of a marine fish. *Marine Ecology Progress Series*, **417**, 263-275.

603 Horne JB, van Herwerden L, Choat JH *et al.* (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.

606 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.

608 Jensen JL, Bohonak AJ, Kelley ST (2005) Isolation by distance, web service. *BMC Genetics*, **6**, 13.

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

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

614 Jones GP, Milicich MJ, Emslie MJ *et al.* (1999) Self-recruitment in a coral reef fish population. *Nature*, **402**, 802-804.

616 Jones GP, Planes S, Thorrold SR (2005) Coral reef fish larvae settle close to home. *Current Biology*, **15**, 1314-1318.

618 Jones GP, Almany GR, Russ GR *et al.* (2009) Larval retention and connectivity among populations of corals and reef fishes: history, advances and challenges. *Coral Reefs*, **28**, 307-325.

621 Kingsford MJ, Leis JM, Shanks A *et al.* (2002) Sensory environments, larval abilities and local self-recruitment. *Bulletin of Marine Science*, **70**, 309–340.

623 Kingsford MJ, Finn MD, O'Callaghan MD *et al.* (2014) Planktonic larval duration, age and growth of *Ostorhinchus doederleini* (Pisces: Apogonidae) on the southern Great Barrier Reef, Australia. *Marine Biology*, **161**, 245-259.

626 Klanten OS, Choat JH, 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.

Author Manuscript

This article is protected by copyright. All rights reserved

628 Kolm N, Hoffman EA, Olsson J *et al.* (2005) Group stability and homing behavior but no kin
629 group structures in a coral reef fish. *Behavioral Ecology*, **16**, 521-527.
630 Kuwamura T, Yogo Y, Nakashima Y (1994) Population dynamics of goby Paragobiodon
631 echinocephalus and host coral Stylophora pistillata. *Marine Ecology Progress Series*, **103**, 17-
632 23.
633 Lachner EA (1955) Inquilinism and a new record for *Paramia bipunctata*, a new cardinal fish
634 from the Red Sea. *Copeia* **1955**, 53–55.
635 Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie2. *Nature Methods*, **9**,
636 357-359.
637 Larsson LC, Laikre L, André C *et al.* (2010) Temporally stable genetic structure of heavily
638 exploited Atlantic herring (*Clupea harengus*) in Swedish waters. *Heredity*, **104**, 40-51.
639 Leis JM (2002) Pacific coral-reef fishes: the implications of behaviour and ecology of larvae for
640 biodiversity and conservation, and a reassessment of the open population paradigm.
641 *Environmental Biology of Fishes*, **65**, 199-208.
642 Leis JM (2007) Behaviour as input for modelling dispersal of fish larvae: behaviour,
643 biogeography, hydrodynamics, ontogeny, physiology and phylogeny meet hydrography.
644 *Marine Ecology Progress Series*, **347**, 185-193.
645 Leis JM, McCormick MI (2002) The biology, behavior, and ecology of the pelagic, larval stage
646 of coral reef fishes. In: *Coral reef fishes: dynamics and diversity in a complex ecosystem* (ed
647 Sale, PF), pp. 171-199. Academic Press, New York.
648 Leis JM, Carson‐Ewart BM, Hay AC *et al.* (2003) Coral‐reef sounds enable nocturnal
649 navigation by some reef‐fish larvae in some places and at some times. *Journal of Fish*
650 *Biology*, **63**, 724-737.
651 Leis JM, Hay AC, Trnski T (2006) In situ ontogeny of behaviour in pelagic larvae of three
652 temperate, marine, demersal fishes. *Marine Biology*, **148,** 655-669.
653 Leis JM, Siebeck U, Dixson DL (2011) How Nemo finds home: the neuroecology of dispersal
654 and of population connectivity in larvae of marine fishes. *Integrative and Comparative*
655 *Biology*, **51**, 826–843.

Author Manuscript

This article is protected by copyright. All rights reserved

656 Lin B, Wang Z, Malanoski AP *et al.* (2010) Comparative genomic analyses identify the Vibrio
657 harveyi genome sequenced strains BAA‐ 1116 and HY01 as *Vibrio campbellii*.
658 *Environmental Microbiology Reports*, **2,** 81-89.
659 Luikart G, England PR, Tallmon D *et al.* (2003) The power and promise of population genomics:
660 from genotyping to genome typing. *Nature Reviews Genetics*, **4**, 981–994.
661 Maes GE, Pujolar JM, Hellemans B *et al.* (2006) Evidence for isolation by time in the European
662 eel (*Anguilla anguilla* L.). *Molecular Ecology*, **15**, 2095-2107.
663 Marnane MJ (2000) Site fidelity and homing behaviour in coral reef cardinalfishes. *Journal of*
664 *Fish Biology*, **57**, 1590-1600.
665 Mastretta‐ Yanes A, Arrigo N, Alvarez N *et al.* (2015) Restriction site‐ associated DNA
666 sequencing, genotyping error estimation and *de novo* assembly optimization for population
667 genetic inference. *Molecular Ecology Resources*, **15,** 28-41.
668 Meirmans PG, Van Tienderen PH (2004) GENOTYPE and GENODIVE: two programs for the
669 analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, **4**, 792-794.
670 Montgomery JC, Jeffs A, Simpson SD *et al.* (2006) Sound as an orientation cue for the pelagic
671 larvae of reef fishes and decapod crustaceans. *Advances in Marine Biology*, **51**, 143-196.
672 Mora C, Sale PF (2002) Are populations of coral reef fish open or closed?. *Trends in Ecology &*
673 *Evolution*, **17**, 422-428.
674 Nakajima Y, Nishikawa A, Iguchi A *et al.* (2010) Gene flow and genetic diversity of a broadcast-
675 spawning coral in northern peripheral populations. *PLoS ONE*, **5**, e11149.
676 Nishihira M, Sato Y, Arakaki Y *et al.* (1991) Ecological distribution and habitat preference of
677 four types of the sea urchin *Echinometra mathaei* on the Okinawan coral reefs. In: *Biology of*
678 *Echinodermata: Proceedings of the 7th International Echinoderm Conference* (eds
679 Yanagisawa T, Yasumasu I, Oguro C *et al.*), pp. 91–104. Balkema, Rotterdam.
680 Nishikawa A (2008) Degree and pattern of gene flow in several scleractinian corals in the
681 Ryukyu Archipelago, Southern Japan. *Pacific Science*, **62**, 413-421.
682 Nishikawa A, Sakai K (2005) Genetic connectivity of the scleractinian coral *Goniastrea aspera*
683 around the Okinawa Islands. *Coral Reefs*, **24**, 318-323.

Author Manuscript

This article is protected by copyright. All rights reserved

684 Nishikawa A, Katoh M, Sakai K (2003) Larval settlement rates and gene flow of broadcast-
685 spawning (*Acropora tenuis*) and planula-brooding (*Stylophora pistillata*) corals. *Marine*
686 *Ecology Progress Series*, **256**, 87-97.
687 Parchman TL, Gompert Z, Mudge J *et al.* (2012) Genome‐ wide association genetics of an
688 adaptive trait in lodgepole pine. *Molecular Ecology*, **21**, 2991-3005.
689 Paris CB, Cowen RK (2004) Direct evidence of a biophysical retention mechanism for coral reef
690 fish larvae. *Limnology and Oceanography*, **49**, 1964-1979.
691 Paris CB, Atema J, Irisson JO *et al.* (2013) Reef odor: a wake up call for navigation in reef fish
692 larvae. *PloS ONE*, **8**, e72808.
693 Peterson BK, Weber JN, Kay EH *et al.* (2012) Double digest RADseq: an inexpensive method for
694 de novo SNP discovery and genotyping in model and non-model species. *PloS ONE*, **7**,
695 e37135.
696 Pew J, Muir PH, Wang J, Frasier TR (2015) related: an R package for analysing pairwise
697 relatedness from codominant molecular markers. *Molecular Ecology Resources*, **15,** 557-561.
698 Picq S, McMillan WO, Puebla O (2016) Population genomics of local adaptation versus
699 speciation in coral reef fishes (*Hypoplectrus spp*, Serranidae). *Ecology and Evolution*, **6**,
700 2109–2124.
701 Planes S (1993) *Acanthurus triostegus* in French Polynesia. *Marine Ecology Progress Series*, **98**,
702 237-246.
703 Planes S, Parroni M, Chauvet C (1998) Evidence of limited gene flow in three species of coral
704 reef fishes in the lagoon of New Caledonia. *Marine Biology*, **130**, 361-368.
705 Planes S (2002) Biogeography and larval dispersal inferred from population genetic analysis. In:
706 *Coral reef fishes: dynamics and diversity in a complex system.* (ed. Sale PF), pp. 201-220.
707 Academic Press, New York.
708 Planes S, Lenfant P (2002) Temporal change in the genetic structure between and within cohorts
709 of a marine fish, *Diplodus sargus*, induced by a large variance in individual reproductive
710 success. *Molecular Ecology*, **11**, 1515-1524.
711 Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus
712 genotype data. *Genetics*, **155**, 945-959.

Author Manuscript

This article is protected by copyright. All rights reserved

713 Puebla O, Bermingham E, McMillan WO (2014) Genomic atolls of differentiation in coral reef
714 fishes (*Hypoplectrus* spp., Serranidae). *Molecular Ecology*, **23**, 5291-5303.
715 Purcell JF, Cowen RK, Hughes CR *et al.* (2006) Weak genetic structure indicates strong dispersal
716 limits: a tale of two coral reef fish. *Proceedings of the Royal Society of London. Series B,*
717 *Biological Sciences*, **273**, 1483-1490.
718 R Core Team (2014) R: A language and environment for statistical computing. R Foundation for
719 Statistical Computing, Vienna, Austria.
720 Riginos C, Buckley YM, Blomberg SP, Treml EA (2014) Dispersal capacity predicts both
721 population genetic structure and species richness in reef fishes. *The American Naturalist*, **184**,
722 52-64.
723 Rocha LA, Robertson DR, Roman J *et al*. (2005) Ecological speciation in tropical reef fishes.
724 *Proceedings of the Royal Society of London Series B,* **272**, 573–579.
725 Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure.
726 *Molecular Ecology Notes*, **4**, 137-138.
727 Rueger T, Gardiner NM, Jones GP (2014) Relationships between pair formation, site fidelity and
728 sex in a coral reef cardinalfish. *Behavioural Processes*, **107**, 119-126.
729 Saenz- Agudelo P, Dibattista JD, Piatek MJ *et al.* (2015) Seascape genetics along environmental
730 gradients in the Arabian Peninsula: insights from ddRAD sequencing of anemonefishes.
731 *Molecular Ecology*, **24**, 6241-6255.
732 Sale PF (2004) Connectivity, recruitment variation, and the structure of reef fish communities.
733 *Integrative and Comparative Biology*, **44**, 390-399.
734 Sale PF, Cowen RK, Danilowicz BS *et al.* (2005) Critical science gaps impede use of no-take
735 fishery reserves. *Trends in Ecology & Evolution*, **20**, 74-80.
736 Seeb JE, Carvalho G, Hauser L *et al.* (2011) Single- nucleotide polymorphism (SNP) discovery
737 and applications of SNP genotyping in nonmodel organisms. *Molecular Ecology Resources*,
738 **11**, 1-8.
739 Selkoe KA, Gaines SD, Caselle JE *et al.* (2006) Current shifts and kin aggregation explain
740 genetic patchiness in fish recruits. *Ecology*, **87**, 3082-3094.
741 Shulman MJ (1998) What can population genetics tell us about dispersal and biogeographic
742 history of coral- reef fishes?. *Australian Journal of Ecology*, **23**, 216-225.

Author Manuscript

This article is protected by copyright. All rights reserved

743 Shulman MJ, Bermingham E (1995) Early life histories, ocean currents, and the population
744 genetics of Caribbean reef fishes. *Evolution*, 897-910.
745 Simpson SD, Meekan MG, Jeffs A *et al.* (2008) Settlement-stage coral reef fishes prefer the
746 higher frequency invertebrate-generated audible component of reef noise. *Animal Behavior*,
747 **75**, 1861–1868.
748 Sponaugle S, Cowen RK, Shanks A *et al.* (2002) Predicting self-recruitment in marine
749 populations: biophysical correlates and mechanisms. *Bulletin of Marine Science*, **70**, 341-375.
750 Staaterman E, Paris CB, Helgers J (2012) Orientation behavior in fish larvae: a missing piece to
751 Hjort's critical period hypothesis. *Journal of Theoretical Biology*, **304**, 188-196.
752 Stockwell BL, Larson WA, Waples RK *et al.* (2016) The application of genomics to inform
753 conservation of a functionally important reef fish (*Scarus niger*) in the Philippines.
754 *Conservation Genetics*, **17**, 239-249.
755 Swearer SE, Caselle JE, Lea DW *et al.* (1999) Larval retention and recruitment in an island
756 population of a coral-reef fish. *Nature*, **402**, 799-802.
757 Swearer SE, Shima JS, Hellberg ME *et al.* (2002) Evidence of self-recruitment in demersal
758 marine populations. *Bulletin of Marine Science*, **70**, 251-271.
759 Tamura R (1982) Experimental observations on the association between the cardinalfish
760 (*Siphamia versicolor*) and the sea urchin (*Diadema setosum*). *Galaxea*, **1**, 1–10.
761 Taylor MS, Hellberg ME (2003) Genetic evidence for local retention of pelagic larvae in a
762 Caribbean reef fish. *Science*, **299**, 107-109.
763 Terry A, Bucciarelli G, Bernardi G (2000) Restricted gene flow and incipient speciation in
764 disjunct Pacific Ocean and Sea of Cortez populations of a reef fish species, *Girella nigricans*.
765 *Evolution*, **54**, 652-659.
766 Treml EA, Halpin PN, Urban DL *et al.* (2008) Modeling population connectivity by ocean
767 currents, a graph-theoretic approach for marine conservation. *Landscape Ecology*, **23**, 19-36.
768 Treml EA, Ford JR, Black KP *et al.* (2015). Identifying the key biophysical drivers, connectivity
769 outcomes, and metapopulation consequences of larval dispersal in the sea. *Movement*
770 *Ecology*, **3**, 1-16.

Author Manuscript

This article is protected by copyright. All rights reserved

771 Tu JY, Chou C, Chu PS (2009) The abrupt shift of typhoon activity in the vicinity of Taiwan and
772 its association with western North Pacific–East Asian climate change. *Journal of Climate*, **22**,
773 3617-3628.
774 Urbanczyk H, Ogura Y, Hendry TA *et al.* (2011) Genome sequence of *Photobacterium*
775 *mandapamensis* strain svers. 1.1, the bioluminescent symbiont of the cardinal fish *Siphamia*
776 *versicolor*. *Journal of Bacteriology*, **193**, 3144-3145.
777 Vagelli A, Burford M, Bernardi G (2008) Fine scale dispersal in Banggai Cardinalfish,
778 *Pterapogon kauderni*, a coral reef species lacking a pelagic larval phase. *Marine Genomics*, **1**,
779 129-134.
780 van der Meer MH, Hobbs JPA, Jones GP *et al.* (2012) Genetic connectivity among and self-
781 replenishment within island populations of a restricted range subtropical reef fish. *PloS One*,
782 **7**, e49660.
783 Willing EM, Dreyer C, Van Oosterhout C (2012) Estimates of genetic differentiation measured
784 by F_{ST} do not necessarily require large sample sizes when using many SNP markers. *PLoS*
785 *ONE*, **7**, e42649.
786 Warner RR, Cowen RK (2002) Local retention of production in marine populations: evidence,
787 mechanisms, and consequences. *Bulletin of Marine Science*, **70**, 245-249.
788 Yasuda N, Nagai S, Hamaguchi M *et al.* (2009) Gene flow of *Acanthaster planci* (L.) in relation
789 to ocean currents revealed by microsatellite analysis. *Molecular Ecology*, **18**, 1574-1590.
790
791
792 **Data Accessibility**
793
794 Illumina RAD-tag sequences are accessible at NCBI SRA accession number SRP105806
795 (Biosample accession numbers: SAMN06857385- SAMN06857664). The Stacks consensus
796 sequences, SNP genotype calls, and STRUCTURE input files are accessible at Dryad
797 (doi:10.5061/dryad.5n882).
798
799
800 **Author Contributions**

Author Manuscript

This article is protected by copyright. All rights reserved

801

802 A.L.G. conceived of and designed the study, performed laboratory work, and analyzed the data.

803 A.L.G. and P.V.D. performed fieldwork and wrote the article.

804

805

806 **Tables**

807

808 **Table 1** The locations and years in which *Siphamia tubifer* were collected in Okinawa, Japan.
809 The range and mean standard lengths (SL) of fish specimens from each sampling site are listed as
810 well as the numbers of individuals collected (N_i) and included in the final genomic dataset after
811 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

812

813

814 **Table 2** Population genetic summary statistics calculated for each group of *Siphamia tubifer*
815 sampled using only nucleotide positions that are polymorphic in at least one population. Statistics
816 listed are the average number of individuals analyzed at each locus (N), the total number of

Author Manuscript

This article is protected by copyright. All rights reserved

817 nucleotide positions in the dataset (Sites), the number of unique variable sites in each population
818 (Private), the percent of polymorphic sites (% Poly), the average frequency of the major allele
819 (P), the average per locus observed heterozygosity (H_{obs}), the average nucleotide diversity (π),
820 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

821

822 **Table 3** Pairwise F_{ST} values (top diagonal) and the shortest distance (km) through water (bottom
823 diagonal) between groups of *Siphamia tubifer* collected in 2013. F_{ST} values in bold are significant
824 at $P < 0.05$

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

Author Manuscript

This article is protected by copyright. All rights reserved

Y	83.6	105.0	51.8	33.4	32.1	100.5	99.8	16.8	98.4	-

825

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

Sesoko	2012	2013	2014
2012	-	0.0177	0.0161
2013		-	0.0158

828

829 **Table 5** Analysis of molecular variance (AMOVA) of groups of *Siphamia tubifer* collected in
830 2013. Populations were grouped into regions determined by collection sites on the east or west
831 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}	0.001
Among population	Region	0.0	F_{SC}	0.245
Among region	--	0.0	F_{CT}	0.012

832

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

Dataset	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%)

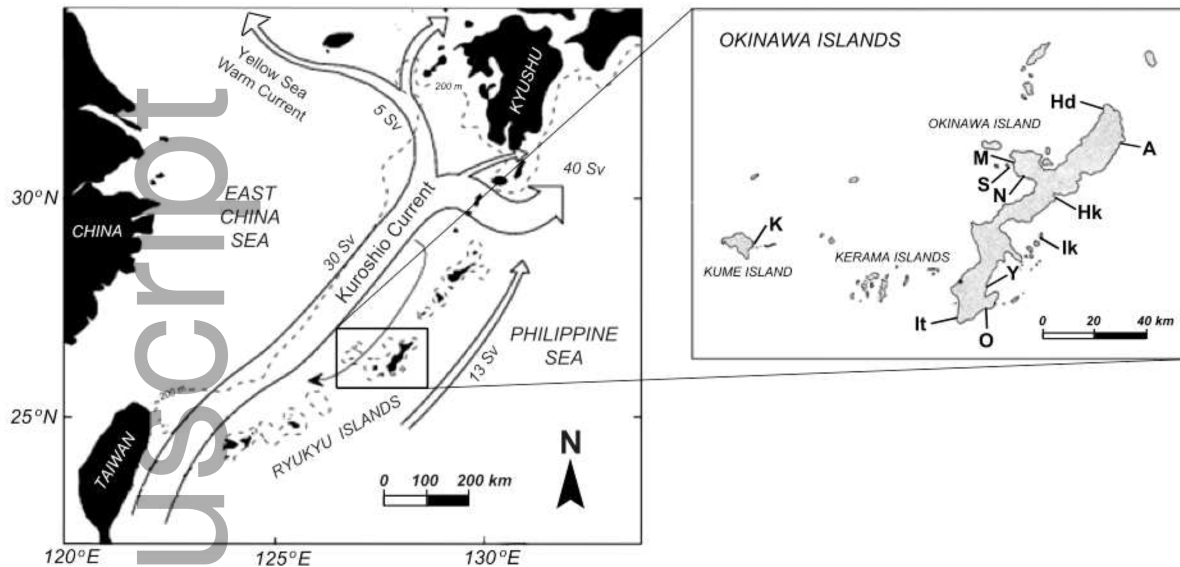
836

837

838 **Figures**

Author Manuscript

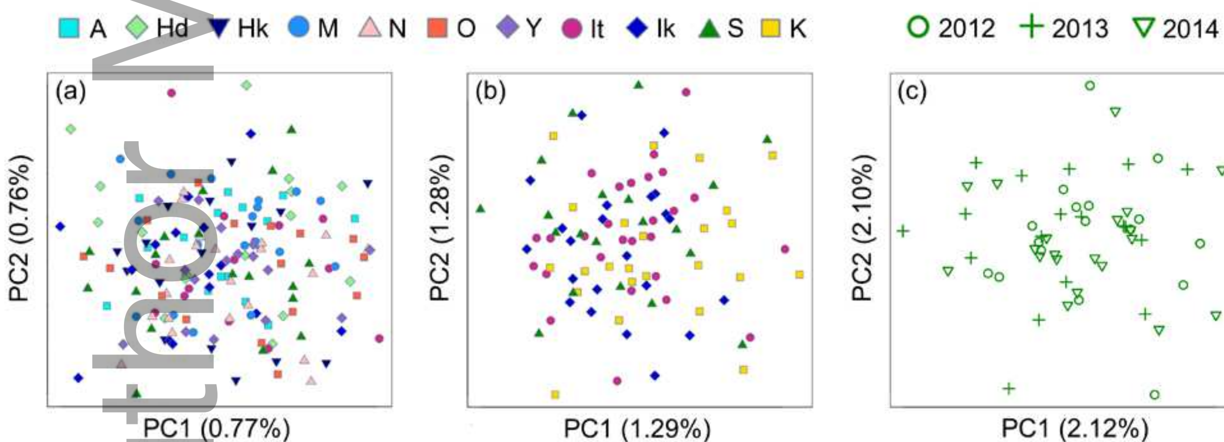
This article is protected by copyright. All rights reserved



839

840 **Fig. 1** Collection sites for *Siphamia tubifer* in Okinawa, Japan, from 2012 to 2014, and the
841 general current patterns in the region.

842



843

844 **Fig. 2** Principal components analyses of genetic differentiation among *Siphamia tubifer* collected
845 in (a) 2013, (b) 2014, and (c) from the Sesoko site collected in three consecutive years. Points
846 represent individuals along the PC1 and PC2 axes of genetic variation across 11,836 SNPs with
847 the amount of variation explained by each axis in parentheses. Distinct colors and shapes indicate
848 the sampling locations.

Author Manuscript

This article is protected by copyright. All rights reserved

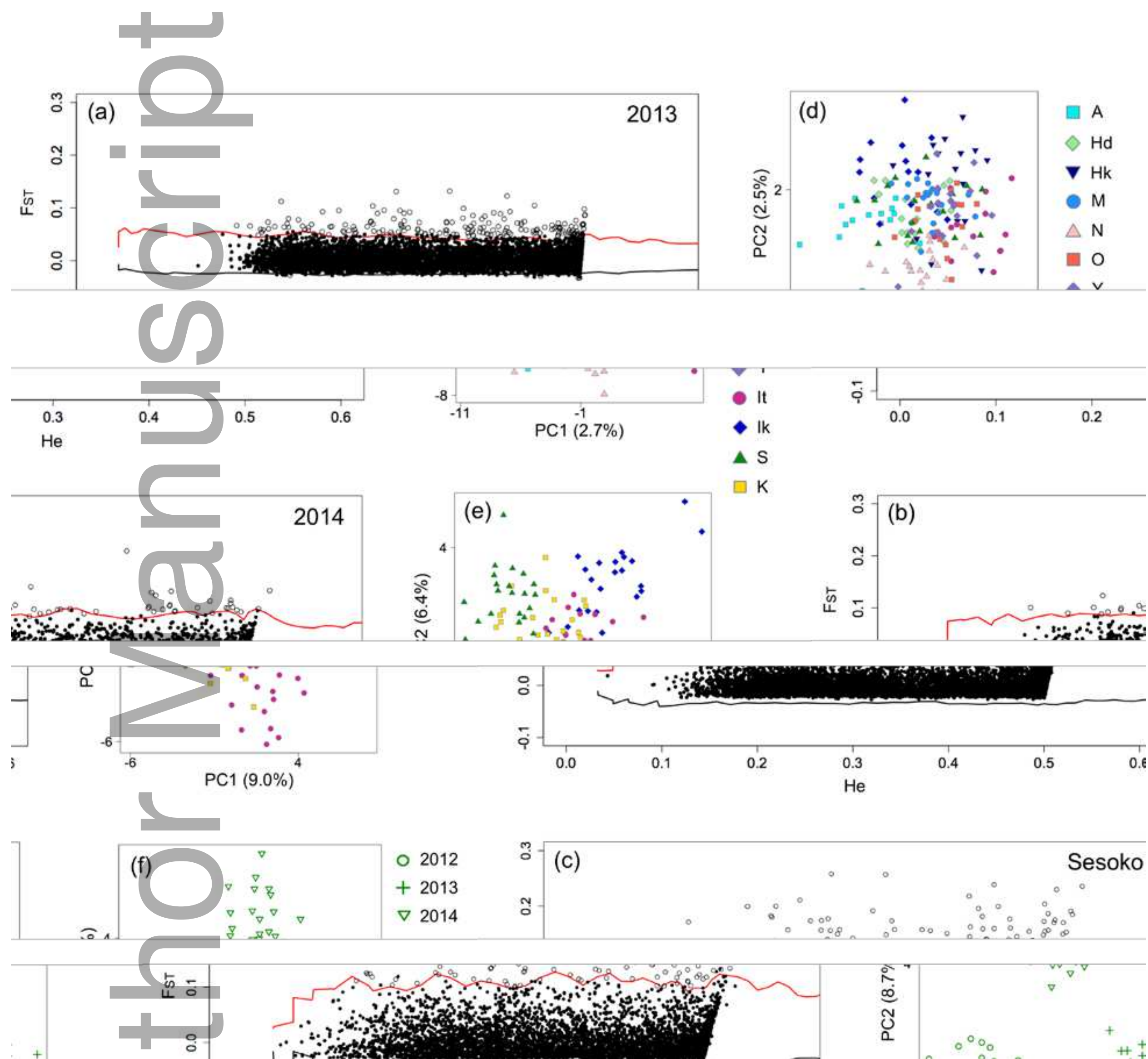
849

850

851

852

Author Manuscript



853

854 **Fig. 3** Results of the F_{ST} outlier tests in *LOSITAN* for 6,379 SNPs in (a) 2013, (b) 2014, and (c) at
855 the Sesoko (S) site over three consecutive years (2012-2014) and corresponding principal
856 components analyses of genetic differentiation for the outlier loci identified (d-f). Loci above the
857 red line (open circles) are candidates of being under divergent selection above a 99% probability.
858 Loci with a probability between 1% and 90% of being under selection are considered to be
859 neutral and those below the bottom line with <0.1% probability are classified as being under

This article is protected by copyright. All rights reserved

860 balancing selection. All remaining loci were conservatively considered to be unclassified (neither
861 neutral or under selection).

Author Manuscript

This article is protected by copyright. All rights reserved

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

Author Manuscript

This article is protected by copyright. All rights reserved

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

Author Manuscript

This article is protected by copyright. All rights reserved

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	-

Author Manuscript

This article is protected by copyright. All rights reserved

Sesoko	2012	2013	2014
2012	-	0.0177	0.0161
2013		-	0.0158

Author Manuscript

This article is protected by copyright. All rights reserved

Source of variation	Nested in	% Variance	F-statistic	P-value
Within individual	--	86.4	F_{IT}	--
Among individual	Population	13.6	F_{IS}	0.001
Among population	Region	0.0	F_{SC}	0.245
Among region	--	0.0	F_{CT}	0.012

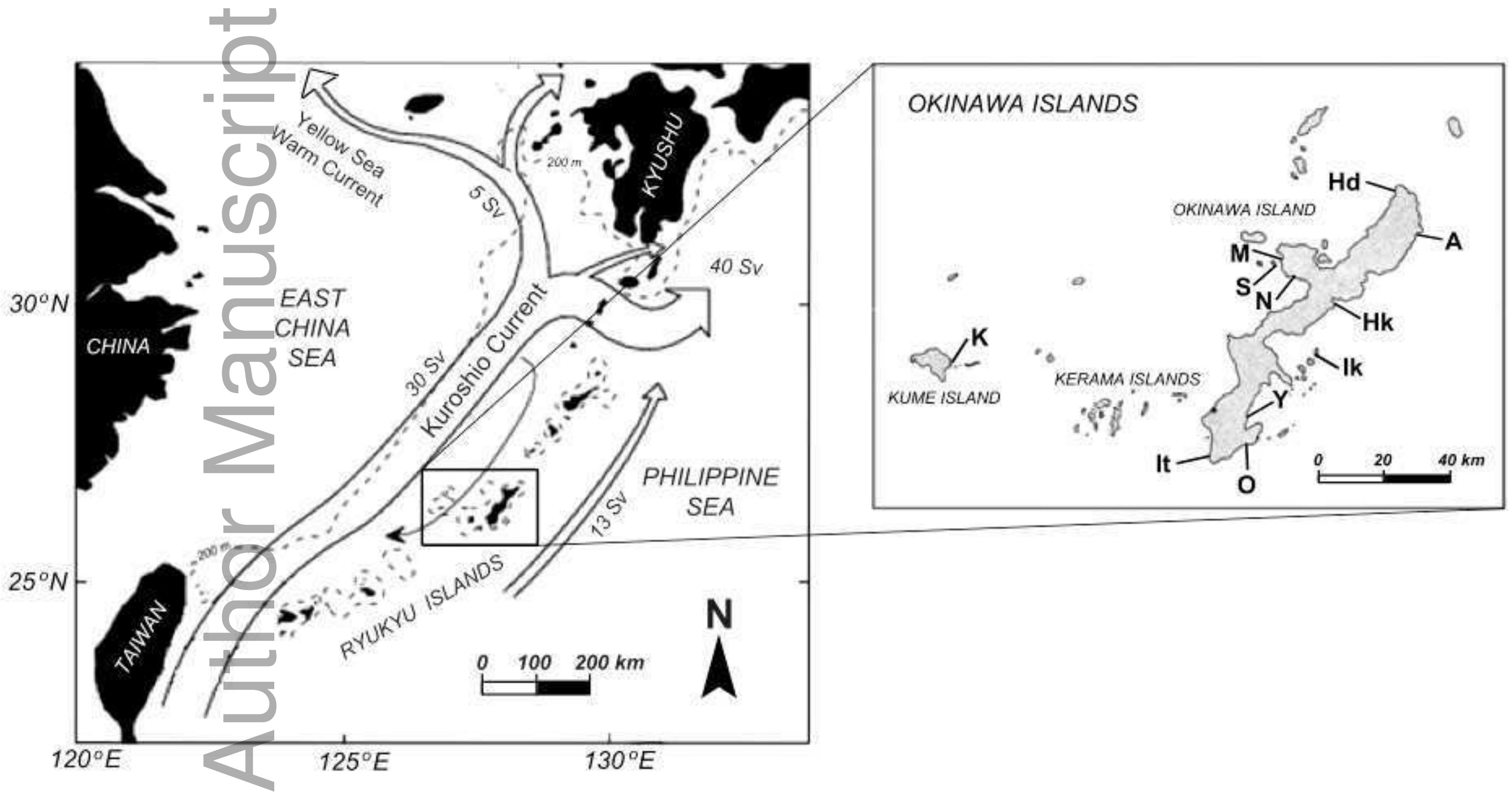
Author Manuscript

This article is protected by copyright. All rights reserved

Dataset	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%)

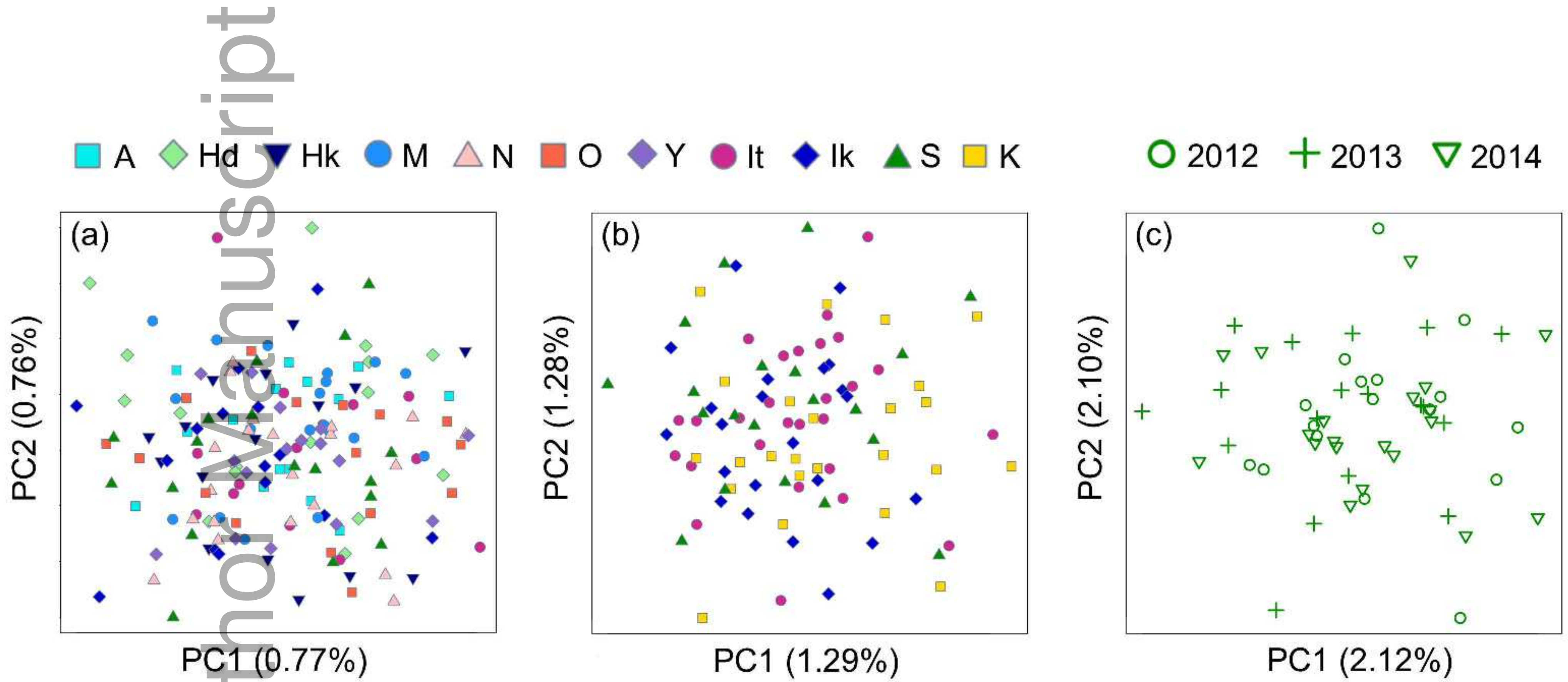
Author Manuscript

This article is protected by copyright. All rights reserved



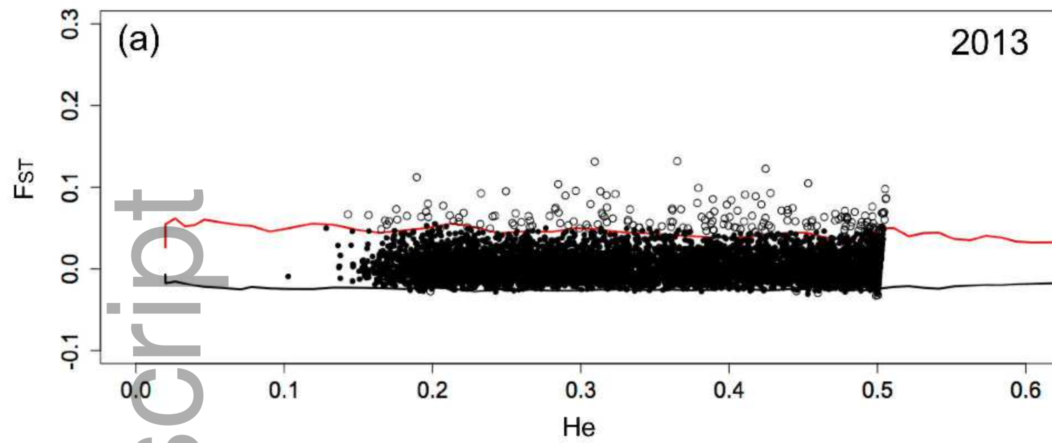
Author Manuscript

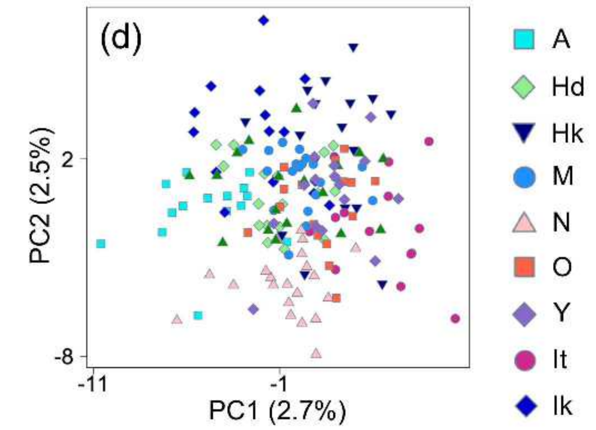
This article is protected by copyright. All rights reserved

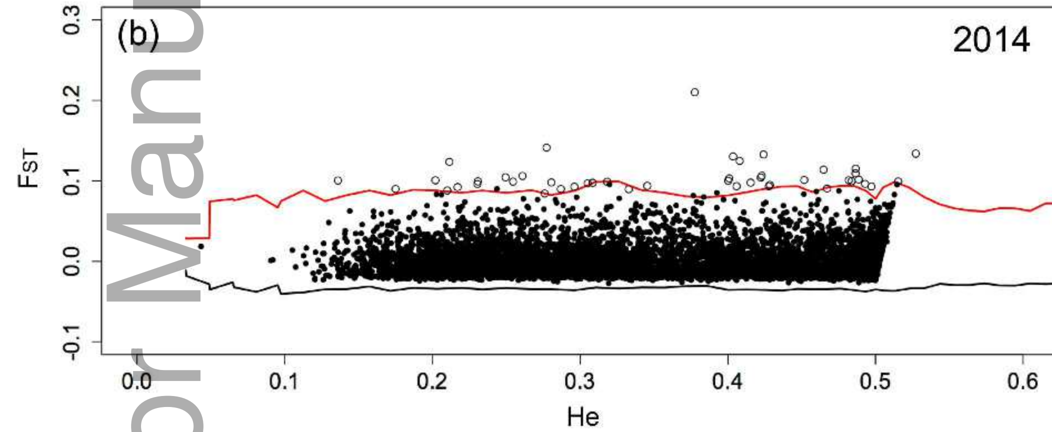


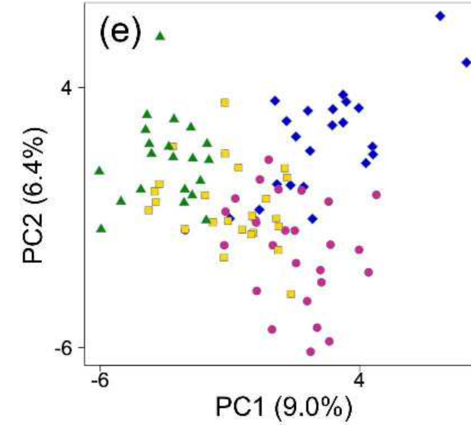
Author Manuscript

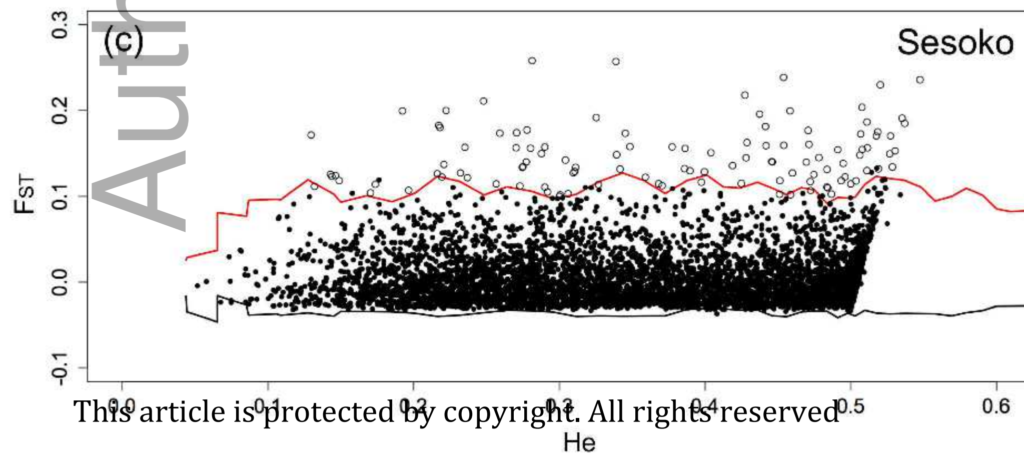
This article is protected by copyright. All rights reserved

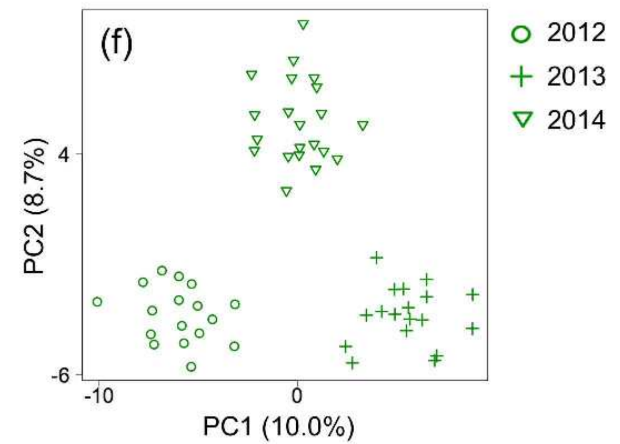












Author Manuscript

This article is protected by copyright. All rights reserved