**Title**

Temporal genetic patterns of diversity and structure evidence chaotic genetic patchiness in a spiny lobster

**Authors**

Cecilia Villacorta-Rath[1], Carla A. Souza[2], Nicholas P. Murphy[2], Bridget S. Green[1], Caleb Gardner[1], Jan M. Strugnell[3,2]

**Author affiliations**

1. Institute for Marine and Antarctic Studies, University of Tasmania, Tasmania 7001, Australia

2. Department of Ecology, Environment and Evolution, La Trobe University, Victoria 3086, Australia

3. Centre for Sustainable Tropical Fisheries and Aquaculture, James Cook University, Queensland 4811, Australia

**Keywords**

Chaotic genetic patchiness, post-settlement selection, population structure, genetic diversity, lobster

**Corresponding author**

Cecilia Villacorta-Rath; Address: Institute for Marine and Antarctic Studies, University of Tasmania, Tasmania 7001, Australia; E-mail: Cecilia.Villacorta@utas.edu.au

**Running title**

Chaotic genetic patchiness in a spiny lobster

**Abstract**

Population structure of many marine organisms is spatially patchy and varies within and between years, a phenomenon defined as chaotic genetic patchiness. This results from the combination of planktonic larval dispersal and environmental stochasticity. Additionally, in species with bi-partite life, post-settlement selection can magnify these genetic differences. The high fecundity (up to 500,000 eggs annually) and protracted larval duration (12-24 months) and dispersal of the southern rock lobster, *Jasus edwardsii*, make it a good test species for chaotic genetic patchiness and selection during early benthic life. Here we used double digest restriction-site associated DNA sequencing (ddRADseq) to investigate chaotic genetic patchiness and post-settlement selection in this species. We assessed differences in genetic structure and diversity of recently settled pueruli across four settlement years and between two sites in southeast Australia separated by approximately 1,000 km. Post-settlement selection was investigated by identifying loci under putative positive selection between recently settled pueruli and post-pueruli and quantifying differences in the magnitude and strength of the selection at each year and site. Genetic differences within and among sites through time in neutral SNP markers indicated chaotic genetic patchiness. Recently settled pueruli at the southernmost site exhibited lower genetic diversity during years of low pueruli catches, further supporting this hypothesis. Finally, analyses of outlier SNPs detected fluctuations in the magnitude and strength of the markers putatively under positive selection over space and time. One locus under putative positive selection was consistent at both locations during the same years, suggesting the existence of weak post-settlement selection.

## Introduction

Population genetic structure of benthic marine invertebrates is often shaped by larval dispersal and natural mortality during early life history (Li & Hedgecock, 1998). The level of genetic connectivity achieved through larval dispersal is critical to our understanding of population dynamics and management of marine species (Cowen, Lwiza, Sponaugle, Paris, & Olson, 2000). However, dispersal does not always happen in a consistent pattern, as it depends on a suite of factors including ocean advection (Cowen et al., 2000), time of spawning (Hendry, Berg, & Quinn, 1999), individual reproductive success (Hedgecock, 1994), pelagic larval duration (Palumbi, 1994), larval behavior (Raimondi & Keough, 1990) and natural selection (Cowen et al., 2000). Variability in the interaction of all those factors mean that ephemeral fine-scale genetic heterogeneity or chaotic genetic patchiness could occur rather than the homogenous genetic diversity that would be assumed to with species that have high reproductive output and long pelagic larval durations (Johnson & Black, 1982).

Selective processes taking place before and after settlement can also lead to temporal genetic variation in the allele frequencies of a population (Johnson & Wernham, 1999). In species with long pelagic larval duration (PLD) and high genetic exchange, post-settlement selection can be large enough to impact recruitment (Butler IV & Herrnkind, 1997; Palma, Wahle, & Steneck, 1998). Knowledge of post-settlement selection and its relationship to recruitment variation is especially important in commercial species (Butler IV & Herrnkind, 1997). The southern rock lobster, *Jasus edwardsii*, represents a valuable fishery in Australia and New Zealand. It occupies a wide geographic range, approximately between 120-180°E and 30-50°S, and sub populations are exposed to different temperature regimes along this latitudinal gradient (Hinojosa et al., 2017). Studies assessing population structure of *J. edwardsii* have distinguished large-scale neutral and adaptive differences between

populations in Australia and New Zealand (Morgan, Green, Murphy, & Strugnell, 2013; Thomas & Bell, 2013; Villacorta-Rath et al., 2016). Although the sampling design of those studies did not allow structure in adult populations to be distinguished at a regional level, larval transport simulations suggest that there is genetic exchange between populations in Australia (Bruce et al., 2007).

The reproductive biology and bi-partite life cycle of the southern rock lobster makes it a model species to investigate chaotic genetic patchiness and post-settlement selection. *J. edwardsii* is a long lived species, with a life span of more than 30 years (Godwin, Frusher, Montgomery, & Ovenden, 2011). Adults reach sexual maturity between two and seven years of age, depending on the latitude (Annala, McKoy, Booth, & Pike, 1980) and they reproduce once a year during the austral winter months (George, 2005). As is typical for species exhibiting genetic patchiness, *J. edwardsii* is highly fecund, each female can produce up to half a million eggs per year (Green, Gardner, & Kennedy, 2009). Shortly after hatching, phyllosoma larvae are transported offshore and can spend between 12-24 months in this phase before metamorphosing into a post-larva, known as puerulus (Booth, 1994). Phyllosoma experience up to 98% mortality during their protracted pelagic larval duration (PLD) (Lesser, 1978). This is not uncommon in marine invertebrates, and in the case of a sea urchin, although 2% of survivors were found to be sufficient to replenish a population, they carry a smaller representation of the overall genetic diversity of the adult population (Flowers, Schroeter, & Burton, 2002). Once pueruli settle they remain sedentary and short distance migration is restricted to nocturnal foraging or retreating to protected areas during moulting (George, 2005).

Settlement of *J. edwardsii* pueruli in southeastern Australia is highest between the months of July and February and peaks during winter months (Gardner, Frusher, Kennedy, &

99 Cawthorn, 2001). Puerulus settlement has been monitored in southeastern Australia for the
100 past four decades using permanent puerulus collectors (Linnane et al., 2010). In South
101 Australia, Tasmania and New Zealand, a puerulus index based on puerulus catches in
102 collectors is highly correlated to future commercial catches 3-7 years after settlement
103 (Gardner et al., 2001; Linnane et al., 2010). High variability in year-to-year settlement
104 abundance of *J. edwardsii* and the subsequent fluctuating commercial catch rates have been
105 described and attributed not only to variation in adult reproductive output (Kennington,
106 Melville-Smith, & Berry, 2013), but to fluctuating environmental factors prior to settlement.
107 Due to the wide geographical range of *J. edwardsii*, the strength of the environmental factors
108 driving settlement abundance differs from region to region (Hinojosa et al., 2017). Although
109 interannual variation in settlement abundance has been widely studied for *J. edwardsii*
110 (Booth, 1994; Hinojosa et al., 2017; Linnane et al., 2014), no study to date has assessed
111 whether year-to-year fluctuations are also observed in the genetic structure and diversity of
112 pueruli and to what extent natural selection during early post-settlement could affect
113 population structure.
114 We hypothesized that the combined effect of varying reproductive output between regions
115 and interannual changes in oceanic current patterns leading to varying patterns of settlement
116 would produce spatial and temporal genetic heterogeneity in *J. edwardsii* pueruli and early
117 juveniles. The present study investigates the spatio-temporal genetic variation of newly
118 settled *J. edwardsii* pueruli by assessing genetic structure among four year-classes from two
119 sites in Australia approximately 1,000 km apart from each other. We also studied possible
120 drivers of selection after settlement by assessing the level of population structure between
121 recently settled pueruli and pueruli that have recently moulted into juveniles. The aims of this
122 study were to: (1) determine whether there is interannual variability in the genetic structure

123 and diversity of recently settled individuals; and (2) to explore the strength and interannual
124 variability of the directional selection acting during settlement.
125
126 **Materials and methods**
127 *Sampling design*
128 Lobster pueruli were caught in crevice collectors (Booth & Tarring, 1986) deployed in
129 South Australia and Tasmania. Pueruli settling into two sites in South Australia, Cape Jaffa
130 and Kingston South East, 19 km apart, were pooled for the analysis and from now on are
131 referred to as "Cape Jaffa". Pueruli settling into two sites in Tasmania, Bicheno shallow and
132 Bicheno deep, separated by less than 1 km, were pooled for the analysis and are referred to as
133 "Bicheno" (Fig. 1, Table 1). Pueruli were collected monthly from 24 collectors at each region
134 from July to November between years 2009 and 2013, which correspond to periods of highest
135 settlement. Collection of pueruli was carried out by a diver placing a fishnet bag around each
136 collector and attaching a rope with a buoy to each bag. Collectors were hauled into a boat,
137 where they were cleaned (Gardner et al., 2001) and pueruli were immediately stored in 90%
138 ethanol. Recently settled pueruli were used to assess chaotic genetic patchiness at settlement.
139 However, due to color change after alcohol preservation it was not possible to distinguish the
140 first two puerulus stages. Stage 1 pueruli are recently settled individuals with transparent
141 bodies and no visible digestive tract, whereas stage 2 pueruli are those that settled within the
142 last two days and have a digestive track but no body pigmentation (Booth, 1979). For the
143 purposes of this study, stages 1 and 2 were considered together as recently settled individuals
144 and denominated "stage 1-2". The other group of pueruli considered in the present study was
145 the "post-puerulus" stage. Post-pueruli have undergone one moult, are fully pigmented and
146 where the pleopods have reduced in size as they are no longer used for forward swimming

147 (Booth, 1979). The time between stages 1-2 to post-puerulus is between 20-24 days in winter
148 (Booth & Kittaka, 1994) and post-pueruli are approximately 18 months old (Booth, 1979). In
149 order to be able to obtain appropriate sample sizes within each year, puerulus settling between
150 July and November at each sampling year were grouped together and denominated as a
151 “puerulus year". Year 2011 exhibited low puerulus catches in collectors, especially at
152 Bicheno. Only 56 pueruli collected from Bicheno and 119 from Cape Jaffa in 2011 and due to
153 the low abundance of stage 1-2 pueruli, year 2011was excluded from the analysis. Full
154 information on puerulus collector catches across the 5-year period is provided in Fig. S1.
155
156 *DNA extraction and ddRADseq library preparation*
157 DNA was extracted from tissue from the horns and legs of each individual using a DNeasy
158 Blood and Tissue kit (Qiagen) and DNA concentration was quantified on a Qubit® 2.0
159 Fluorometer (Life Technologies). DNA integrity was determined through gel electrophoresis
160 to verify high molecular weight DNA in all samples (>1,000 base pairs [bp] or higher).
161 ddRADseq libraries were prepared following the protocol described in Villacorta-Rath et
162 al. (2016) using a gel size selection of 400-600 bp in order to maximize the overlapping
163 region among libraries. One interlibrary technical replicate and between three and four
164 intralibrary technical replicates were included in each ddRADseq library to test for batch
165 effects (Mastretta-Yanes et al., 2015). Electrophoretic assays were performed on all
166 ddRADseq libraries by AGRF using an Agilent BioAnalyser (Agilent Technologies) to
167 determine their exact molecular weight range and library concentration. A total of 330
168 samples were sequenced across 5.5 lanes at the Australian Genome Research Facility (AGRF)
169 on the Illumina HiSeq 2500 platform using a 100 bp single end kit.
170

*Analyses of raw sequencing data and reference catalog building*

An initial quality check of raw indexed data was performed using FastQC v.0.10.1. Data was then demultiplexed using the "process_radtags" protocol from Stacks v.1.29 (Catchen et al., 2011) and hard trimmed to 75 bp (-t 75) to ensure that Phred Quality Score (Q Score) of all reads was above 30. Demultiplexed libraries were filtered for bacterial and viral content using Kraken-gcc v.0.10.4 (Wood & Salzberg, 2014).

*Jasus edwardsii* possesses a large genome, with a C-value of 5.01 (Deiana et al., 1999) and publicly available pipelines do not perform very well given the amount of repetitive regions contained within it. In our experience, these pipelines result in a reference catalog composed of large amount of paralogous loci. The rad-loci pipeline employed here (and within Villacorta-Rath et al., 2016) was developed in order to cluster raw reads using a series of steps that filter out alleles based on a percentage identity. This process maximizes the removal of paralogous sequences. Raw reads filtered through Kraken-gcc v.0.10.4 were initially merged and clustered allowing a 4% mismatch (3 bp) between alleles. We allowed each cluster to have a minimum depth of 305 reads. Given that the total number of sequenced samples was 330, this step assumes that most individual samples were represented with at least one read in each cluster. Subsequently, clusters that had less than two and more than 16 different alleles were discarded. The minimum number of two was chosen, as we wanted to obtain bi-allelic data. The maximum number of 16 was based on the fact that for biallelic data, there are 16 possible differences for reads that are divergent in 4 bp. During this step, paralogs and repetitive regions of the genome were discarded. A second round of clustering of the remaining clusters was performed at a 96% identity and a subsequent filtering of clusters that were not composed of a minimum of two and maximum of 16 alleles. Finally, samples

were mapped back to the identified loci and individuals with more than 30% missing data were removed from downstream analyses.

*Read alignment, variant calling and neutral SNP characterization*

Individual filtered reads were aligned to the reference catalog using bwa-intel v.0.7.12 (Langmead & Salzberg, 2012) using the BWA-MEM algorithm. The final alignment was output in the sequence alignment map (SAM) format. SAM files were then converted into their binary form (BAM), sorted and indexed. Subsequently variant calling of the sorted alignments was performed through the Genome Analysis Toolkit (GATK) v.3.3_0 (McKenna et al., 2010) using the Haplotype Caller option. This option outputs a gVCF file with raw, unfiltered SNPs for each sample. All gVCF files were then combined into a single gVCF file. Finally, the gVCF file was transformed into a genotype file (VCF). A further correction of the VCF file obtained from GATK was performed to ensure the accuracy of the reference and alternate allele calls and to filter out false positives. In the absence of a reference genome, the correction was made by calculating the ratio between the highest quality score over depth ("QD" in VCF file) and lowest QD. If one sample at a specific position had a ratio threshold of 10, which corresponds to a 10% error on a Phred scale, it was substituted by missing data for that SNP.

SNP filtering was performed in vcftools-gcc v.0.1.13 to ensure that only bi-allelic data was present (--min-alleles 2, --max-alleles 2), remove SNPs that were potentially in linkage disequilibrium (--min-r2 0.2), discard SNPs with a minor allele frequency (MAF) of less than 5% (--maf 0.05) and ensure that the minimum SNP depth was 5 (--minDP 5). The maximum amount of missing data for each locus was set to 25% (--max-missing 0.75) and the maximum amount of missing data for each individual was 25%. Individuals with more than the missing

data threshold were removed from subsequent analyses. Finally, only one SNP per locus (--thin 75) was retained.

Neutral loci identification was performed in LOSITAN (Antao, Lopes, Lopes, Beja-Pereira, & Luikart, 2008) using 100,000 simulations, a confidence interval of 0.95, and a false discovery rate of 0.1 (Jacobsen et al., 2014). LOSITAN uses an $F_{ST}$-outlier approach that identifies loci are outlier candidates when they exhibit too high or too low $F_{ST}$ compared to neutral expectations (Antao et al., 2008).

In order to minimize false positives in the panel of neutral loci an additional characterization was performed using the R package OutFLANK v.0.1 (Whitlock & Lotterhos, 2015). OutFLANK infers the $F_{ST}$ distribution of neutral loci by trimming loci contained in the tails of the distribution, which are likely to be under balancing and positive selection. Therefore the program does not assume a specific population's demographic history and is less prone to false positives than other programs (Whitlock & Lotterhos, 2015). The proportion of loci trimmed from both tails of the $F_{ST}$ distribution was set to 5% (LeftTrimFraction = 0.05, RightTrimFraction = 0.05), the minimum heterozygosity required before including calculations from a locus was set to 0 (Hmin = 0) and the false discovery rate was 0.1 (qthreshold = 0.1). Only loci shared between LOSITAN and OutFLANK were considered in the neutral SNPs panel.

*Interannual variability in genetic structure of puerulus*

To assess interannual variability in genetic structure only recently settled individuals (stage 1-2 pueruli) were analyzed. Global $F_{ST}$ values and confidence intervals for the neutral SNP panel were estimated using the R package mmod v.1.3.2 (Winter, 2012). Additionally, pairwise $F_{ST}$ values between sampling sites and years as well as confidence intervals were

242 calculated in the R package hierfstat v.0.04-22 (Goudet, 2005). A false discovery rate
243 correction (FDR) was applied to calculated p-values using the function "p.adjust" of the R
244 package stats v.3.2.4 (R Core Team, 2016).
245 Analyses of Molecular Variance (AMOVA) (Excoffier, Smouse, & Quattro, 1992) were
246 performed in order to quantify the variance explained by (1) puerulus year, (2) sampling site
247 and (3) year*site using the R packages poppr v.2.1.1 (Kamvar, Tabima, & Grünwald, 2014)
248 and Ade4 v.1.7-4 (Dray & Dufour, 2007). Genetic distances were corrected with the
249 "quasieuclidean" method. The number of permutations to test for significance was set to
250 9999. Finally, Discriminant analysis of principal components (DAPC) was used to determine
251 the number of possible genetic clusters based on allele frequency data of both sites across four
252 sampling years. DAPC was performed and results were plotted using the R package adegenet
253 v.1.4-2 (Jombart, 2008). The number of clusters present in the dataset was determined using
254 the K-means approach implemented in the function "find.clusters" (Jombart, Devillard, &
255 Balloux, 2010). The optimal number of clusters was the one with the lowest associated BIC
256 (Jombart et al., 2010).
257
258 *Temporal variation in genetic diversity*
259 Genetic diversity was measured by calculating individual heterozygosity for stage 1-2
260 puerulus only, in order to obtain the best representation of the genetic diversity in the
261 population before post-settlement mortality could influence diversity. To ensure that
262 heterozygosity of all individuals was measured on the same scale, the standardized
263 heterozygosity per individual (sh) was used. This metric is the proportion of heterozygous
264 loci over the mean heterozygosity across all markers (Coltman, Pilkington, Smith, &
265 Pemberton, 1999). Sh of stage 1-2 puerulus for the neutral SNP panel was calculated using

the R package Rhh v.1.0.1 (Alho, Välimäki, & Merilä, 2010). To test for differences in the minimum and maximum values of the sh distributions between sampling years, a Mann-Whitney test was performed (Hart, 2001). A Bonferroni correction was applied to the significance level since each sampling year was tested against the other three sampling years and therefore $\alpha = 0.05/3$.

Finally, given that chaotic genetic patchiness can have an effect on the effective population size (Ne), Ne fluctuations were quantified over time at each site. Ne of stage 1-2 pueruli at each sampling year and site was estimated using the NeEstimator software v.2.01 (Do et al., 2014) under the temporal model based on allele frequencies of the neutral SNP panel.

*Evidence of natural selection at settlement*

Post-settlement selection between stage 1-2 pueruli and post-pueruli was investigated separately for each year and site. To account for potential genetic structure between sites and years an initial outlier SNP identification was carried out for each stage separately as "control" tests and any loci identified here were removed from subsequent analyses. Eight groups were considered for each stage at the control tests (four years and two sites).

SNP characterization was performed using LOSITAN (Antao et al., 2008) and OutFLANK (Whitlock & Lotterhos, 2015) with the same set of parameters used for the neutral loci characterization. Only SNPs characterized as being under putative positive selection by both approaches were included in downstream analyses.

The number and magnitude (mean $F_{ST}$ values) of SNPs under putative positive selection was analyzed. Differences in the average magnitude of the selection through time were determined through a Kruskal-Wallis test and post-hoc differences were examined using

Mann-Whitney tests. A Bonferroni correction was applied to the significance level at $\alpha = 0.05/5$.

Finally, to determine if the loci containing SNPs putatively under positive selection were contained in protein coding regions, sequences were BLASTed against the complete *Homarus americanus* transcriptome (McGrath, Vollmer, Kaluziak, & Ayers, 2016) and a *Jasus edwardsii* transcriptome database (SRA Bioproject accession number: PRJNA386609) using BLAST+ v.2.2.29. Queries with statistically significant e-values ($E < 10^{-6}$) and more than 84% identity were considered as valid alignments. Transcriptome sequences that provided significant alignments were annotated using the Trinotate pipeline (https://**trinotate**.github.io/) to determine if they aligned with any known protein domain.

## Results

*Sequencing data, SNP filtering and neutral SNP characterization*

HiSeq sequencing of 330 pueruli yielded an average of 2.7(±1.03) million reads per sample. A total of 64 samples were discarded due to high percentage of missing data after read alignment, SNP calling and filtering, and 266 samples with sufficient data for downstream analyses remained. The reference catalog built through the rad-loci pipeline produced 5,488 loci. After read alignment, variant calling and filtering for false positive SNPs, we obtained a total of 6,285 variable sites. The final number of SNPs after filtering for MAF, minimum site depth, LD, missing data and retaining one SNP per locus was 900. SNP characterization in LOSITAN identified 558 neutral SNPs, 305 SNPs under putative balancing selection and 37 SNPs under putative positive selection. OutFLANK characterized 888 neutral SNPs, 12 SNPs under putative positive selection and no SNPs under putative

312 balancing selection. The final neutral SNP panel contained 558 loci that were shared between
313 both SNP characterization methods.
314 A total of nine intralibrary and three interlibrary technical replicates remained after
315 removal of individuals with more than the missing data threshold. A Principal Components
316 Analysis (PCA) was performed to visualize the spatial distribution of replicated samples
317 using the genotypes of all loci before outlier identification (Villacorta-Rath et al., 2016) (Fig.
318 S2). The PCA plot showed each intralibrary technical replicate pair from both sites distributed
319 close to each other. The three interlibrary technical replicates were also distributed closely in
320 the PCA plot. This suggests that sequencing samples across multiple lanes did not introduce a
321 large bias in the catalog building process.
322
323 *Interannual variability in genetic structure of recently settled individuals*
324 No genetic differentiation was detected between recently settled pueruli (stage 1-2) of all
325 years and sites when analyzing the neutral SNP panel (global $F_{ST}$ = -0.0042, n.s.) and
326 similarly there was no evidence for population structure found through DAPC (Fig. S3).
327 Although Figure S3 shows two differentiated density distributions, the Bayesian information
328 criteria (BIC) indicated the existence of one cluster and therefore we cannot conclude that
329 there is population structure in the data. However, pairwise $F_{ST}$ values revealed genetic
330 differences amongst most of the collection years in Cape Jaffa, except between years 2009
331 and 2012 and years 2012 and 2013, and amongst all years in Bicheno (Table 2). Significant
332 genetic differences were also observed between Cape Jaffa and Bicheno during most years,
333 except during 2010 and 2013, where no genetic differentiation between sites was found,
334 suggesting a possible common population of origin for those years. A lack of genetic
335 differentiation was also seen between years 2012 in Cape Jaffa and 2009 in Bicheno. The lack

336 of a consistent pattern in genetic differentiation across years suggests that patterns of genetic
337 structure in pueruli are ephemeral as proposed by the chaotic genetic patchiness hypothesis. In
338 support of the differences in pairwise $F_{ST}$ values, allele frequencies of the neutral SNP panel
339 were significantly different between Cape Jaffa and Bicheno, but the degree of differentiation
340 between sites was extremely low ($\Phi = 0.1\%$) (AMOVA, Table 3). The interaction of sampling
341 site and puerulus year was also significant and explained a small amount of the variation
342 between samples ($\Phi = 0.2\%$).
343
344 *Temporal variation in genetic diversity*
345 Genetic diversity of recently settled pueruli estimated using the neutral SNP panel did not
346 differ significantly among years or locations (Fig. 2) despite the large year-to-year differences
347 in puerulus catches at both sites (Fig. S1). However, the maximum and minimum values of
348 genetic diversity were variable between years and sites. The minimum values of standardized
349 heterozygosity during years of low puerulus catches were lower than those during years of
350 high puerulus catches at Bicheno but not at Cape Jaffa. This could suggest that years of higher
351 catch rates at Bicheno also exhibited higher genetic diversity possibly due to the existence of
352 more "winners" contributing to population replenishment at those years.
353 Finally, Ne estimates for each year as well as the confidence intervals were "infinite"
354 (results not reported) and therefore we could not quantify Ne variation over time and space.
355 Genetic identity between sites, developmental stages and between years and the interaction
356 of all factors differed significantly at both sites (Table 4). The differences in allele frequencies
357 between sites, stages and years could be indicative of post-settlement mortality and suggest
358 that this is not constant geographically nor temporally. Pairwise $F_{ST}$ values supported this idea
359 revealing significant differences between developmental stages across almost all years at both

sites (Table S1). Post-pueruli from Cape Jaffa collected in 2010 were the exception, as they showed greater genetic similarity to stage 1-2 puerulus from Bicheno in 2009 and Cape Jaffa in 2012.

*Evidence of natural selection at settlement*

Following the control tests, two loci potentially confounded for selection for year and/or site were identified and removed from the analysis of post settlement selection. Both sites exhibited a varying number of SNPs under putative positive selection between stage 1-2 and post-pueruli across all years. A single common locus was shared between Cape Jaffa and Bicheno during years 2009 and 2012, supporting common, but weak post-settlement selection over multiple sites/years. All other loci identified as under selection were unique to site and year. Contrary to predictions of post-settlement selection, genetic diversity did not differ significantly.

Settlement at Cape Jaffa in 2009 exhibited the largest number of outlier SNPs (n = 10) whilst year 2013 exhibited only one outlier locus (Fig. 3a). The magnitude of the positive selection at Cape Jaffa also differed significantly across three sampling years (Kruskal-Wallis chi-squared = 8.7829, df = 2, p-value = 0.01238). Since year 2013 only exhibited one locus under putative positive selection it was excluded from the analysis of variance. Significant differences were detected between years 2009 and 2012 (Fig. 3a).

Settlement in year 2012 exhibited the largest number of SNPs under putative positive selection at Bicheno (n = 9), whereas year 2009 exhibited only one SNP under putative positive selection (Fig. 3b). The magnitude of the positive selection at this site differed significantly across three sampling years (Kruskal-Wallis chi-squared =8.8676, df = 2, p-value = 0.01187). Year 2009 was not included in the analysis of variance because there was

384 only one SNP characterized as being under putative positive selection. Therefore the new
385 level of α after the Bonferroni correction was 0.05/3. Post-hoc tests indicated that mean $F_{ST}$
386 values differed significantly between years 2012 and 2010 (Fig. 3b). The level of observed
387 heterozygosity (Ho) remained constant between stage 1-2 and post-pueruli at both sites (Fig.
388 S4), supporting the pattern of weak selection.
389 Finally, transcriptome sequences from the optical nerve and green gland of *J. edwardsii*
390 exhibited significant hits with three loci containing SNPs under putative positive selection.
391 Additionally, no significant hits between the panel of SNPs under putative positive selection
392 and the *Homarus americanus* transcriptome were found.
393 .
394 **Discussion**
395 Our results support chaotic genetic patchiness in *J. edwardsii*. We demonstrate variable
396 patterns of genetic divergence between recently settled individuals in two sampling sites
397 separated by approximately 1,000 km and across four years. Genetic differences in neutral
398 SNPs within and among sites through time as well as differences in the minimum and
399 maximum values of genetic diversity between years of high and low puerulus catches were
400 indicative of chaotic genetic patchiness. There was some evidence of post-settlement selection
401 which exhibited temporal and spatial variation in the magnitude and strength of selection
402 following settlement.
403
404 *Chaotic genetic patchiness*
405 The significant genetic differentiation between stage 1-2 pueruli across sampling sites
406 and years found in the present study point to the existence of chaotic genetic patchiness.
407 Genetic patchiness is a widespread phenomenon occurring in marine invertebrates, including

408 spiny lobsters (Iacchei et al., 2013; Johnson & Wernham, 1999; Kennington, Berry, Groth,
409 Johnson, & Melville-Smith, 2013; Thompson, Hanley, & Johnson, 1996; Truelove et al.,
410 2017). For example, temporal variation in allele frequencies of two allozymes revealed
411 ephemeral structure in the adult population of the western rock lobster, *Panulirus cygnus* (
412 Thompson et al., 1996). Since *P. cygnus* constitutes a single panmictic stock along its whole
413 distribution, the authors attributed the temporal discrepancies to variation in recruitment
414 among the study sites (Thompson et al., 1996). Similarly, unexpected patterns of genetic
415 heterogeneity were found the adult population of the California spiny lobster, *Panulirus*
416 *interruptus*, despite no evidence of isolation-by-distance throughout its geographic range
417 (Iacchei et al., 2013). This genetic structure was attributed to localized recruitment of related
418 individuals (Iacchei et al., 2013).
419 The significant interaction between site and year (AMOVA neutral SNPs, Table 3) found
420 in the present study and the significantly different $F_{ST}$ values between years within both sites
421 found herein further support the existence of chaotic genetic patchiness. In the case of *J.*
422 *edwardsii*, while examination of the adult populations generally concludes panmixia within
423 Australia (Morgan et al., 2013; Thomas & Bell, 2013; Villacorta-Rath et al., 2016), our
424 examination including cohorts of pueruli over time and space blurs this conclusion. Recently
425 settled individuals (stage 1-2 pueruli) analyzed in the present study exhibited a non-
426 significant global level of genetic differentiation across sites and years ($F_{ST}$ = -0.0042),
427 suggesting that there is gene flow between South Australia and Tasmania. However pairwise
428 $F_{ST}$ values differed significantly both geographically and temporally. This incongruence is
429 indicative of genetic patchiness. Similar trends in lack of global population differentiation but
430 fine-scale genetic heterogeneity has been reported in *P. cygnus* (Kennington et al., 2013). The
431 authors hypothesized that even in a panmictic population, such as that of *P. cygnus*, genetic

heterogeneity can arise from sweepstakes reproductive success (SRS) or pre-settlement mortality (Kennington et al., 2013).

Studies assessing chaotic genetic patchiness during early life stages of lobsters have also identified ephemeral genetic structure. Johnson & Wernham (1999) sampled recently settled *P. cygnus* in two sites 350 km apart during three consecutive recruitment seasons. Allele frequencies of one allozyme marker differed significantly between sites only during one recruitment season. Given that the pueruli settling into each of the two sites at a given recruitment season were genetically homogenous, the authors concluded that SRS was not a driver of genetic patchiness. In the present study, pueruli also showed genetic homogeneity within sites at each year (Appendix S2), suggesting that the observed ephemeral population structure is possibly not a result of differential reproduction. However, the significant differences observed in the minimum and maximum values of the standardized individual heterozygosity between years of high and low puerulus catches at Bicheno could be due to the existence of a higher number of genetically different groups contributing to the next generation of settlers.

Estimates of Ne and the Ne/N ratio across years could help further in determining the cause of chaotic genetic patchiness. One possible cause of genetic patchiness in recently settled individuals is differential adult reproduction (Johnson & Black, 1982). In the presence of SRS, Ne is much smaller than the census size of a population, given that a minority of individuals reproduces successfully (Hedgecock, 1994). Therefore large variations in Ne through time could be indicative of SRS. Unfortunately, we were unable to determine if effective population size fluctuates through time due to the infinite values of Ne obtained, probably due to the low sample size per site and year. However, given the lack of significant differences in the median levels of genetic diversity found across the four-year sampling

456 period of the present study, it is unlikely that we would have been able to distinguish
457 differences in Ne between populations in the presence of larger sample sizes.
458 Ocean advection and retention of larvae can generate unexpected genetic patchiness in
459 highly dispersive organisms. A biophysical model of larval dispersal explained a large
460 percentage of the variation in population structure between populations of *Panulirus argus*
461 across the Caribbean (Truelove et al., 2017). The authors identified an oceanographic
462 boundary dividing the region into two provinces: a southern province characterized by
463 recirculation and a northern province influenced by offshore currents. These two distinct
464 oceanic features created genetic patchiness along *P. argus'* geographic range (Truelove et al.,
465 2017). In the case of *J. edwardsii*, patterns of settlement abundance have also been attributed
466 to oceanographic features that vary in strength and importance among regions (Hinojosa et
467 al., 2017). An oceanographic model simulating larval transport of *J. edwardsii* suggested that
468 there is a high degree of self-recruitment in areas such as Cape Jaffa (Bruce et al., 2007).
469 Conversely, East Tasmania (Bicheno) receives settlers that originated from South Australia
470 and Tasmania but that fluctuations in ocean advection create year-to-year variability in the
471 source of pueruli (Bruce et al., 2007). In the present study, the different environmental factors
472 influencing larval transport in South Australia and Tasmania could explain the clear structure
473 between sites whereas the year-to-year variability in oceanic features could be behind the
474 genetic structure between years at each site.
475
476 *Post-settlement selection*
477 We identified some evidence of weak positive selection acting shortly after settlement of
478 *J. edwardsii* that is consistent across site and year, however the majority of the results suggest
479 variable selection through space and time. Predation during early post-settlement of spiny

480 lobsters is extremely high and is considered a potential bottleneck to recruitment (Smith &
481 Herrnkind, 1992). At the same time, predation risk in recently settled clawed and spiny
482 lobsters is highly dependent on shelter availability (Butler IV, Herrnkind, & Hunt, 1997). For
483 example, survival of the Caribbean spiny lobster during the "algal-phase" was positively
484 related to settlement habitat suitability (Butler IV et al., 1997). The common locus between
485 sites and years found herein could be linked to a trait that makes individuals more vulnerable
486 to predation or affect their ability to settle into suitable habitat. However, we should interpret
487 these results with caution since there was only one SNP under putative positive selection
488 shared between both sites and we were not able to ascribe a function to this SNP.
489 The majority of evidence for post settlement selection results from different panels of
490 putative positively selected SNPs identified for each year and site. This suggests that either
491 the selective forces acting during early post-settlement of *J. edwardsii* differ across the
492 geographic range of the species and from year to year, or that different loci are responding to
493 the same selection pressure. Differences in selection pressures would be the most likely
494 scenario given that climate-ocean forces driving settlement in *J. edwardsii* are also variable
495 across the species distribution (Hinojosa et al., 2017). As mentioned above, predation is the
496 main source of mortality after settlement. Since the vulnerability to post-settlement predation
497 decreases with individual size (Butler IV & Herrnkind, 1997; Palma et al., 1998), variability
498 in factors affecting growth across sampling sites and years may play a role in the divergence
499 in magnitude and strength of the selection found herein. Water temperature has been
500 described as the main abiotic factor determining intermoult interval and moult increments in
501 the early benthic phase of Caribbean spiny lobsters (Forcucci, Butler IV, & Hunt, 1994).
502 Although that study sampled only one site, lobster growth was determined during summer
503 and winter months. The temperature gradient between seasons in Forcucci et al.'s study was

504 of approximately 10°C, which is higher than the temperature gradient between sampling sites
505 of the present study. However, given that *J. edwardsii* exhibits high phenotypic plasticity,
506 slight variation in water temperature can have large impact on growth and size-at-age
507 (Chandrapavan, Gardner, & Green, 2010).

508 Processes occurring prior to settlement can also influence post-settlement mortality in
509 marine invertebrates (Palma et al., 1998). For example, inadequate food intake during the
510 larval phase can reduce juvenile growth rates (Pechenik & Tyrell, 2015). Since plankton
511 abundance fluctuates spatially and temporarily (Jeffs, Nichols, Mooney, Phillips, & Phleger,
512 2004; Wang et al., 2015), condition of *J. edwardsii* phyllosoma could potentially be very
513 variable throughout its prolonged PLD. Latent effects of larval development on subsequent
514 stages have been reported for a wide range of invertebrates (Pechenik, 2006). Consequently,
515 phyllosoma that developed under sub-optimal feeding conditions could experience reduced
516 growth after settlement and therefore be more vulnerable to mortality during this period
517 (Pechenik, 2006).

518 Interestingly, the years of highest puerulus catches at both sites also exhibited the largest
519 number of SNPs under putative positive selection. This could suggest that *J. edwardsii*
520 experiences density-dependent mortality at settlement, however this hypothesis is highly
521 unlikely. Recently settled spiny lobster pueruli exhibit solitary living in macroalgae (Marx &
522 Herrnkind, 1985) however their survival and growth is influenced by habitat complexity
523 rather than conspecific density (Butler IV et al., 1997). The existence of more loci under
524 putative positive selection during years of higher puerulus catches at each site (year 2009 at
525 Cape Jaffa and year 2012 at Bicheno) could be due to the larger sample sizes available for
526 those years. Nevertheless, years 2009 and 2012 at both sites had very similar sample sizes and
527 yet there were large differences in the number of loci under putative positive selection.

528 Therefore the variability in the loci could indicate that there is varying selection pressure
529 among years of higher and lower puerulus catches.
530 Finally, the constant pattern of observed heterozygosity between stage 1-2 and post-pueruli
531 could mean that neither homozygote nor heterozygote genotypes are favored by the weak
532 selection found herein. Another possible explanation for the lack of differences in Ho among
533 developmental stages is that the post-pueruli analyzed in the present study do not comprise
534 the same genetic cohort as recently settled pueruli. If multi-cohorts were sampled (stage 1-2
535 pueruli being genetically different cohort from post-pueruli), comparisons of genetic diversity
536 between stages would give biased results. However, the sampling design of this study
537 prevents us from being able to answer this question and therefore a fine-scale study assessing
538 how the genetic identity of settlers arriving in consecutive pulses within a settlement season
539 fluctuates could clarify this subject.
540
541 **Conclusions**
542 The significant genetic differences at neutral loci between some of the sampling years are
543 consistent with the existence of chaotic genetic patchiness in *J. edwardsii*. In the presence of
544 high self-recruitment and high egg production in South Australia, suggested by an earlier
545 oceanographic model, population structure at the northernmost site could be mainly driven by
546 ocean advection influencing the number of settlers arriving at this site every year. At the
547 southernmost site, structure could be primarily driven by fluctuations in the amount of source
548 populations producing successful settlers every year.
549 A weak positive post-settlement selection was identified, however, selective forces acting
550 during early post-settlement of *J. edwardsii* differ by site and time, providing further evidence
551 of chaotic genetic patchiness. The possibility of having sampled multiple cohorts could also

be driving the observed pattern of fluctuating positive selection between developmental stages. In the former case, differential mortality across a geographic range can be due to environmental factors and also genotype-related fitness of recently settled individuals.

**Acknowledgements**

Funding for this research was provided by an Australian Research Council Linkage Project grant (Project No. LP120200164) and Australian Research Council Discovery Project grant (Project No. DP150101491). We would like to thank Peter Hawthorne (South Australian Research and Development Institute, SARDI) for puerulus collections and data on settlement abundance from South Australia. We thank Ivan Hinojosa, Ruari Colquhoun, David Fallon, Kylie Cahill, Graeme Ewing and Sarah Pyke for puerulus collections in Tasmania. Thanks for laboratory assistance to Brooke Whitelaw and Daniel Laird. Special thanks to Karen J. Miller who contributed to the original project idea. We thank two anonymous reviewers who helped improving this manuscript.

**References**

Alho, J. S., Välimäki, K., & Merilä, J. (2010). Rhh: an R extension for estimating multilocus heterozygosity and heterozygosity-heterozygosity correlation: Computer program note. *Molecular Ecology Resources*, *10*(4), 720–722. doi:10.1111/j.1755-0998.2010.02830.x

Annala, J. H., McKoy, J. L., Booth, J. D., & Pike, R. B. (1980). Size at the onset of sexual maturity in female *Jasus edwardsii* (Decapoda: Palinuridae) in New Zealand. *New Zealand Journal of Marine and Freshwater Research*, *14*(3), 217–227. doi:10.1080/00288330.1980.9515864

576 Antao, T., Lopes, A., Lopes, R. J., Beja-Pereira, A., & Luikart, G. (2008). LOSITAN: A
577 workbench to detect molecular adaptation based on a Fst-outlier method. *BMC*
578 *Bioinformatics*, *9*(1), 323. doi:10.1186/1471-2105-9-323
579 Booth, J. D. (1979). Settlement of the rock lobster, *Jasus edwardsii* (Decapoda:
580 Palinuridae), at Castlepoint, New Zealand. *New Zealand Journal of Marine and*
581 *Freshwater Research*, *13*(3), 395–406. doi:10.1080/00288330.1979.9515816
582 Booth, J. D. (1994). *Jasus edwardsii* larval recruitment off the east coast of New Zealand.
583 *Crustaceana*, *66*(3), 295–317.
584 Booth, J. D., & Kittaka, J. (1994). Growout of juvenile spiny lobster. In B. F. Phillips, J. S.
585 Cobb & J. Kittaka (Eds.), *Spiny lobster management* (pp. 429-239). Oxford:
586 Blackwell Scientific Publications.
587 Booth, J. D., & Tarring, S. C. (1986). Settlement of the red rock lobster, *Jasus edwardsii* ,
588 near Gisborne, New Zealand. *New Zealand Journal of Marine and Freshwater*
589 *Research*, *20*(2), 291–297. doi:10.1080/00288330.1986.9516150
590 Bruce, B. D., Griffin, D. A., Bradford, R. W., CSIRO (Australia), Marine Research, &
591 Fisheries Research & Development Corporation (Australia). (2007). *Larval*
592 *transport and recruitment processes of southern rock lobster*. Hobart: CSIRO
593 Marine Research.
594 Butler IV, M. J., & Herrnkind, W. F. (1997). A test of recruitment limitation and the
595 potential for artificial enhancement of spiny lobster (*Panulirus argus*)
596 populations in Florida. *Canadian Journal of Fisheries and Aquatic Sciences*, *54*,
597 452–463.

598 Butler IV, M. J., Herrnkind, W. F., & Hunt, J. H. (1997). Factors affecting the recruitment of
599 juvenile Caribbean spiny lobsters dwelling in macroalgae. *Bulletin of Marine*
600 *Science*, *61*(1), 3–19.
601 Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., Postlethwait, J. H., & De Koning, D.-J.
602 (2011). Stacks: Building and genotyping loci de novo from short-read sequences.
603 *Genes|Genomes|Genetics*, *1*(3), 171–182. doi:10.1534/g3.111.000240
604 Chandrapavan, A., Gardner, C., & Green, B. S. (2010). Growth rate of adult rock lobsters
605 *Jasus edwardsii* increased through translocation. *Fisheries Research*, *105*(3), 244–
606 247. doi:10.1016/j.fishres.2010.03.014
607 Coltman, D. W., Pilkington, J. G., Smith, J. A., & Pemberton, J. M. (1999). Parasite-mediated
608 selection against inbred soay sheep in a free-living, island population. *Evolution*,
609 *53*(4), 1259. doi:10.2307/2640828
610 Cowen, R. K., Lwiza, K. M. M., Sponaugle, S., Paris, C. B., & Olson, D. B. (2000).
611 Connectivity of marine populations: Open or closed? *Science, 287*, 857–859.
612 doi:10.1126/science.287.5454.857
613 Deiana, A. M., Cau, A., Coluccia, E., Cannas, R., Milia, A., Salvadori, S., & Libertini, A.
614 (1999). Genome size and AT-DNA content in thirteen species of decapoda. In R.
615 Schram & J.C. von Vaupel Klein (Eds.), *Crustaceans and the biodiversity crisis* (pp.
616 981–985). Leiden, The Netherlands: Koninklijke Brill NV.
617 Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., & Ovenden, J. R. (2014).
618 NeEstimator v2: Re-implementation of software for the estimation of
619 contemporary effective population size ($N_e$) from genetic data. *Molecular Ecology*
620 *Resources*, *14*(1), 209–214. doi:10.1111/1755-0998.12157

621 Dray, S., & Dufour, A.-B. (2007). The ade4 package: Implementing the duality diagram for
622 ecologists. *Journal of Statistical Software*, *22*(4), 1–20.
623 Excoffier, L., Smouse, P. E., & Quattro, J. M. (1992). Analysis of molecular variance
624 inferred from metric distances among DNA haplotypes: Application to human
625 mitochondrial DNA restriction data. *Genetics*, *131*, 479–491.
626 Flowers, J. M., Schroeter, S. C., & Burton, R. S. (2002). The recruitment sweepstakes has
627 many winners: Genetic evidence from the sea urchin *Strongylocentrotus*
628 *purpuratus*. *Evolution*, *56*(7), 1445–1453.
629 Forcucci, D., Butler IV, M. J., & Hunt, J. H. (1994). Population dynamics of juvenile
630 Caribbean spoiny lobster, *Panulirus argus*, in Florida Bay, Florida. *Bulletin of*
631 *Marine Science*, *54*(3), 805–818.
632 Gardner, C., Frusher, S. D., Kennedy, R. B., & Cawthorn, A. (2001). Relationship between
633 settlement of southern rock lobster pueruli, Jasus edwardsii, and recruitment to
634 the fishery in Tasmania, Australia. *Marine and Freshwater Research*, *52*(8), 1271–
635 1275.
636 George, R. W. (2005). Review: Evolution of life cycles, including migration, in spiny
637 lobsters (Palinuridae). *New Zealand Journal of Marine and Freshwater Research*,
638 *39*(3), 503–514. doi:10.1080/00288330.2005.9517329
639 Godwin, R. M., Frusher, S., Montgomery, S. S., & Ovenden, J. (2011). Telomere length
640 analysis in crustacean species: *Metapenaeus macleayi, Sagmariasus verreauxi, and*
641 *Jasus edwardsii*. *ICES Journal of Marine Science*, *68*(10), 2053–2058.
642 doi:10.1093/icesjms/fsr144
643 Goudet, J. (2005). Hierfstat, a package for R to compute and test hierarchical F-statistics.
644 *Molecular Ecology Notes*, *5*(1), 184–186.

645 Green, B. S., Gardner, C., & Kennedy, R. B. (2009). Generalised linear modelling of
646 fecundity at length in southern rock lobsters, *Jasus edwardsii*. *Marine Biology*,
647 *156*(9), 1941–1947. doi:10.1007/s00227-009-1226-x
648 Hart, A. (2001). Mann-Whitney test is not just a test of medians: Differences in spread
649 can be important. *BMJ, 323*, 391–393. doi:10.1136/bmj.323.7309.391
650 Hedgecock, D. (1994). Does variance in reproductive success limit effective population
651 sizes of marine organisms? In A. Beaumont (Ed.), *Genetics and evolution of*
652 *aquatic organisms* (pp. 122–134). London: Chapman & Hall.
653 Hendry, A. P., Berg, O. K., & Quinn, T. P. (1999). Condition dependence and adaptation-
654 by-time: Breeding date, life history, and energy allocation within a population of
655 Salmon. *Oikos*, *85*(3), 499. doi:10.2307/3546699
656 Hinojosa, I. A., Gardner, C., Green, B. S., Jeffs, A., Leon, R., & Linnane, A. (2017). Differing
657 environmental drivers of settlement across the range of southern rock lobster (
658 *Jasus edwardsii* ) suggest resilience of the fishery to climate change. *Fisheries*
659 *Oceanography*, *26*(1), 49–64. doi:10.1111/fog.12185
660 Iacchei, M., Ben-Horin, T., Selkoe, K. A., Bird, C. E., García-Rodríguez, F. J., & Toonen, R. J.
661 (2013). Combined analyses of kinship and $F_{ST}$ suggest potential drivers of chaotic
662 genetic patchiness in high gene-flow populations. *Molecular Ecology*, *22*(13),
663 3476–3494. doi:10.1111/mec.12341
664 Jacobsen, M. W., Pujolar, J. M., Bernatchez, L., Munch, K., Jian, J., Niu, Y., & Hansen, M. M.
665 (2014). Genomic footprints of speciation in Atlantic eels ( *Anguilla anguilla* and *A.*
666 *rostrata* ). *Molecular Ecology*, *23*(19), 4785–4798. doi:10.1111/mec.12896
667 Jeffs, A. G., Nichols, P. D., Mooney, B. D., Phillips, K. L., & Phleger, C. F. (2004). Identifying
668 potential prey of the pelagic larvae of the spiny lobster *Jasus edwardsii* using

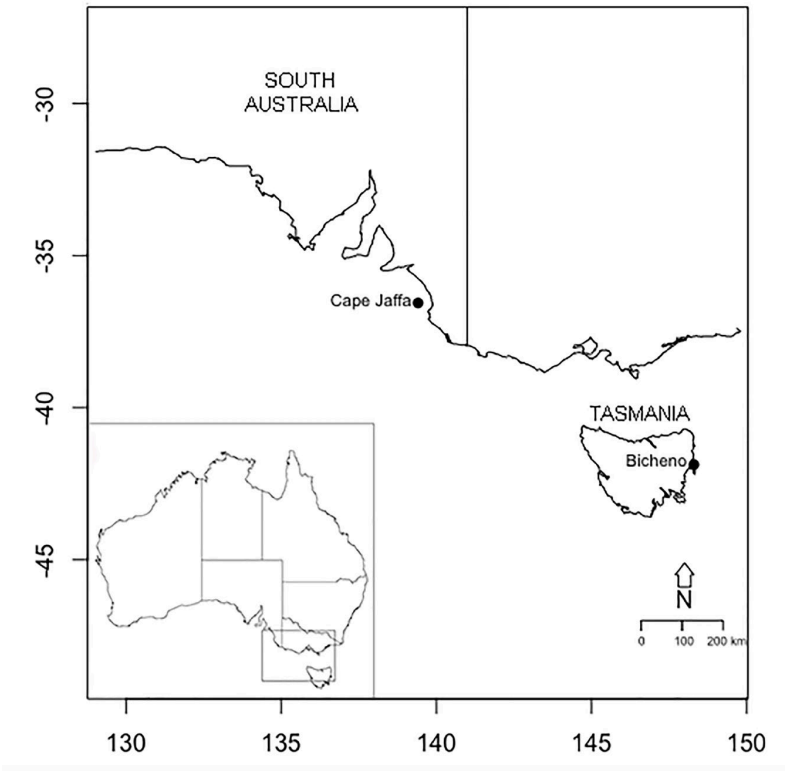
669 signature lipids. *Comparative Biochemistry and Physiology Part B: Biochemistry*
670 *and Molecular Biology*, *137*(4), 487–507. doi:10.1016/j.cbpc.2004.02.003
671 Johnson, M. S., & Black, R. (1982). Chaotic genetic patchiness in an intertidal limpet,
672 *Siphonaria* sp. *Marine Biology*, *70*(2), 157–164.
673 Johnson, M. S., & Wernham, J. (1999). Temporal variation of recruits as a basis of
674 ephemeral genetic heterogeneity in the western rock lobster *Panulirus cygnus*.
675 *Marine Biology*, *135*(1), 133–139.
676 Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic
677 markers. *Bioinformatics*, *24*(11), 1403–1405.
678 doi:10.1093/bioinformatics/btn129
679 Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal
680 components: A new method for the analysis of genetically structured
681 populations. *BMC Genetics*, *11*(1), 1.
682 Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: An R package for genetic
683 analysis of populations with clonal, partially clonal, and/or sexual reproduction.
684 *PeerJ*, *2*, e281. doi:10.7717/peerj.281
685 Kennington, W., Berry, O., Groth, D., Johnson, M., & Melville-Smith, R. (2013). Spatial
686 scales of genetic patchiness in the western rock lobster *Panulirus cygnus*. *Marine*
687 *Ecology Progress Series*, *486*, 213–221. doi:10.3354/meps10360
688 Kennington, W. J., Melville-Smith, R., & Berry, O. (2013). Genetics of the wild and captive
689 lobster populations. In B. F. Phillips (Ed.), *Lobsters: biology, management,*
690 *aquaculture and fisheries* (pp. 36-63). Oxford: Blackwell Publishing.
691 Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2.
692 *Nature Methods*, *9*(4), 357–359. doi:10.1038/nmeth.1923

693 Lesser, J. H. R. (1978). Phyllosoma larvae of *Jasus edwardsii* (Hutton) (Crustacea:
694 Decapoda: Palinuridae) and their distribution off the east coast of the North
695 Island, New Zealand. *New Zealand Journal of Marine and Freshwater Research*,
696 *12*(4), 357–370. doi:10.1080/00288330.1978.9515763
697 Li, G., & Hedgecock, D. (1998). Genetic heterogeneity, detected by PCR-SSCP, among
698 samples of larval Pacific oysters (*Crassostrea gigas*) supports the hypothesis of
699 large variance in reproductive success. *Canadian Journal of Fisheries and Aquatic*
700 *Sciences*, *55*(4), 1025–1033.
701 Linnane, A., Gardner, C., Hobday, D., Punt, A., McGarvey, R., Feenstra, J., … Green, B.
702 (2010). Evidence of large-scale spatial declines in recruitment patterns of
703 southern rock lobster *Jasus edwardsii*, across south-eastern Australia. *Fisheries*
704 *Research*, *105*(3), 163–171. doi:10.1016/j.fishres.2010.04.001
705 Linnane, A., McGarvey, R., Gardner, C., Walker, T. I., Matthews, J., Green, B., & Punt, A. E.
706 (2014). Large-scale patterns in puerulus settlement and links to fishery
707 recruitment in the southern rock lobster (*Jasus edwardsii*), across south-eastern
708 Australia. *ICES Journal of Marine Science*, *71*(3), 528–536.
709 doi:10.1093/icesjms/fst176
710 Marx, J. M., & Herrnkind, W. F. (1985). Macroalgae (Rhodophyta: *Laurencia* spp.) as
711 habitat for young juvenile spiny lobsters, *Panulirus argus*. *Bulletin of Marine*
712 *Science*, *36*(3), 423–431.
713 Mastretta-Yanes, A., Arrigo, N., Alvarez, N., Jorgensen, T. H., Piñero, D., & Emerson, B. C.
714 (2015). Restriction site-associated DNA sequencing, genotyping error estimation
715 and *de novo* assembly optimization for population genetic inference. *Molecular*
716 *Ecology Resources*, *15*(1), 28–41. doi:10.1111/1755-0998.12291

717 McGrath, L. L., Vollmer, S. V., Kaluziak, S. T., & Ayers, J. (2016). De novo transcriptome
718 assembly for the lobster *Homarus americanus* and characterization of differential
719 gene expression across nervous system tissues. *BMC Genomics*, *17*(1).
720 doi:10.1186/s12864-016-2373-3
721 McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., … DePristo,
722 M. A. (2010). The Genome Analysis Toolkit: A MapReduce framework for
723 analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–
724 1303. doi:10.1101/gr.107524.110
725 Morgan, E. M. J., Green, B. S., Murphy, N. P., & Strugnell, J. M. (2013). Investigation of
726 genetic structure between deep and shallow populations of the southern rock
727 lobster, *Jasus edwardsii* in Tasmania, Australia. *PLoS ONE*, *8*(10), e77978.
728 doi:10.1371/journal.pone.0077978
729 Palma, A. T., Wahle, R. A., & Steneck, R. S. (1998). Different early post-settlement
730 strategies between American lobsters *Homarus americanus* and rock crabs
731 *Cancer irroratus* in the Gulf of Maine. *Marine Ecology Progress Series*, *162*, 215–
732 225.
733 Palumbi, S. R. (1994). Genetic divergence, reproductive isolation, and marine speciation.
734 *Annual Review of Ecology and Systematics*, 547–572.
735 Pechenik, J. A. (2006). Larval experience and latent effects - Metamorphosis is not a new
736 beginning. *Integrative and Comparative Biology*, *46*(3), 323–333.
737 doi:10.1093/icb/icj028
738 Pechenik, J. A., & Tyrell, A. S. (2015). Larval diet alters larval growth rates and post-
739 metamorphic performance in the marine gastropod *Crepidula fornicata*. *Marine*
740 *Biology*, *162*(8), 1597–1610. doi:10.1007/s00227-015-2696-7

741 R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna,
742 Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-
743 project.org/
744 Raimondi, P. T., & Keough, M. J. (1990). Behavioural variability in marine larvae.
745 *Australian Journal of Ecology*, *15*(4), 427–437.
746 Smith, K. N., & Herrnkind, W. F. (1992). Predation on early juvenile spiny lobsters
747 *Panulirus argus* (Latreille): Influence of size and shelter. *Journal of Experimental*
748 *Marine Biology and Ecology*, *157*, 3–18.
749 Thomas, L., & Bell, J. J. (2013). Testing the consistency of connectivity patterns for a
750 widely dispersing marine species. *Heredity*, *111*, 345–354.
751 doi:10.1038/hdy.2013.58
752 Thompson, A. P., Hanley, J. R., & Johnson, M. S. (1996). Genetic structure of the western
753 rock lobster, *Panulirus cygnus*, with the benefit of hindsight. *Marine and*
754 *Freshwater Research*, *47*(7), 889–896.
755 Truelove, N. K., Kough, A. S., Behringer, D. C., Paris, C. B., Box, S. J., Preziosi, R. F., & Butler,
756 M. J. (2017). Biophysical connectivity explains population genetic structure in a
757 highly dispersive marine species. *Coral Reefs, 36*(1), 233-244.
758 doi:10.1007/s00338-016-1516-y
759 Villacorta-Rath, C., Ilyushkina, I., Strugnell, J. M., Green, B. S., Murphy, N. P., Doyle, S. R., …
760 Bell, J. J. (2016). Outlier SNPs enable food traceability of the southern rock
761 lobster, *Jasus edwardsii*. *Marine Biology*, *163*(11). doi:10.1007/s00227-016-3000-
762 1
763 Wang, M., O'Rorke, R., Waite, A. M., Beckley, L. E., Thompson, P., & Jeffs, A. G. (2015).
764 Condition of larvae of western rock lobster (*Panulirus cygnus*) in cyclonic and

765 anticyclonic eddies of the Leeuwin Current off Western Australia. *Marine and*
766 *Freshwater Research*, *66*(12), 1158. doi:10.1071/MF14121
767 Whitlock, M. C., & Lotterhos, K. E. (2015). Reliable detection of loci responsible for local
768 adaptation: Inference of a null model through trimming the distribution of FST.
769 *The American Naturalist*, *186*(S1), S24–S36. doi:10.1086/682949
770 Winter, D. J. (2012). MMOD : An R library for the calculation of population differentiation
771 statistics. *Molecular Ecology Resources*, *12*(6), 1158–1160. doi:10.1111/j.1755-
772 0998.2012.03174.x
773 Wood, D. E., & Salzberg, S. L. (2014). Kraken: Ultrafast metagenomic sequence
774 classification using exact alignments. *Genome Biology*, *15*(3), 1.
775
776
777
778 **Data accessibility**
779 Genotype (VCF) file and reference loci sequences: Dryad doi: 10.5061/dryad.6gh40
780
781 **Author contributions**
782 B.S.G., J.M.S. and C.V.R. conceived the sampling design; C.V.R. performed part of the field
783 sampling; C.V.R. and C.A.S. conducted laboratory work; C.V.R. conducted analyses and
784 wrote the manuscript; N.P.M., J.M.S., C.A.S., C.G. and B.S.G. contributed to versions of the
785 manuscript.




786

787 **FIGURE 1** Puerulus collector sites at Cape Jaffa (South Australia) and Bicheno (Tasmania)
788

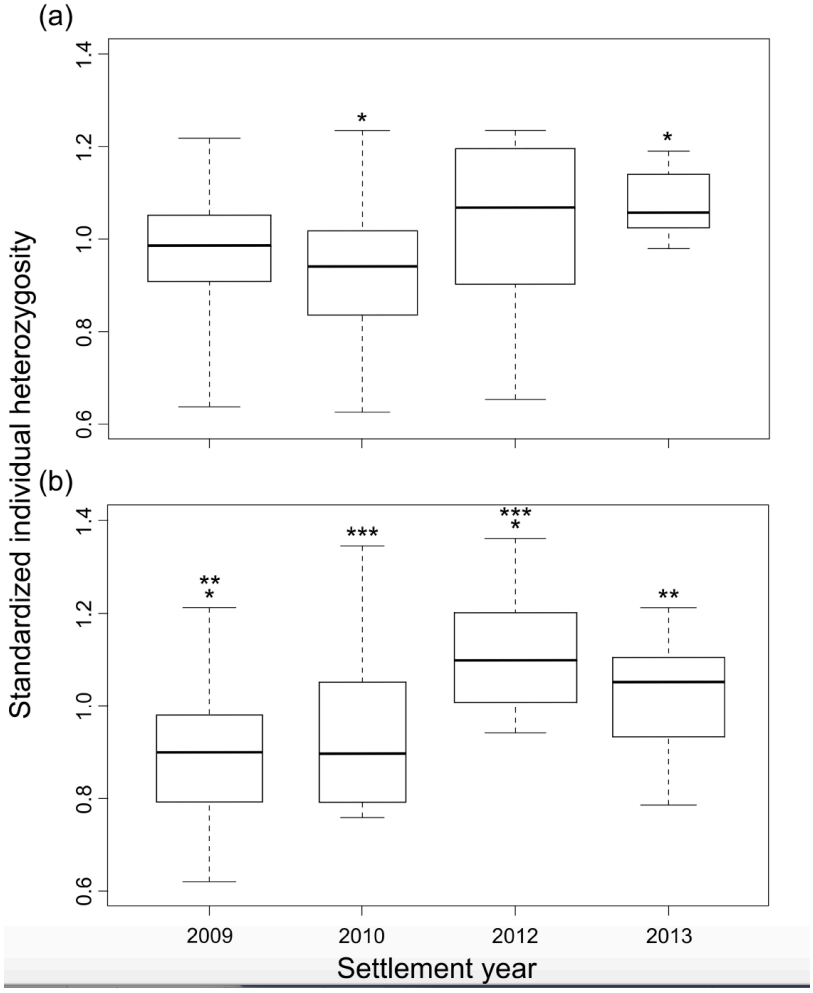



**FIGURE 2** Median standardized individual heterozygosity of stage 1-2 pueruli per settlement year using neutral SNPs for (a) Cape Jaffa and (b) Bicheno. Error bars represent the minimum and maximum observed values of the standardized individual heterozygosity. (*, **, ***) Denote significant differences between years

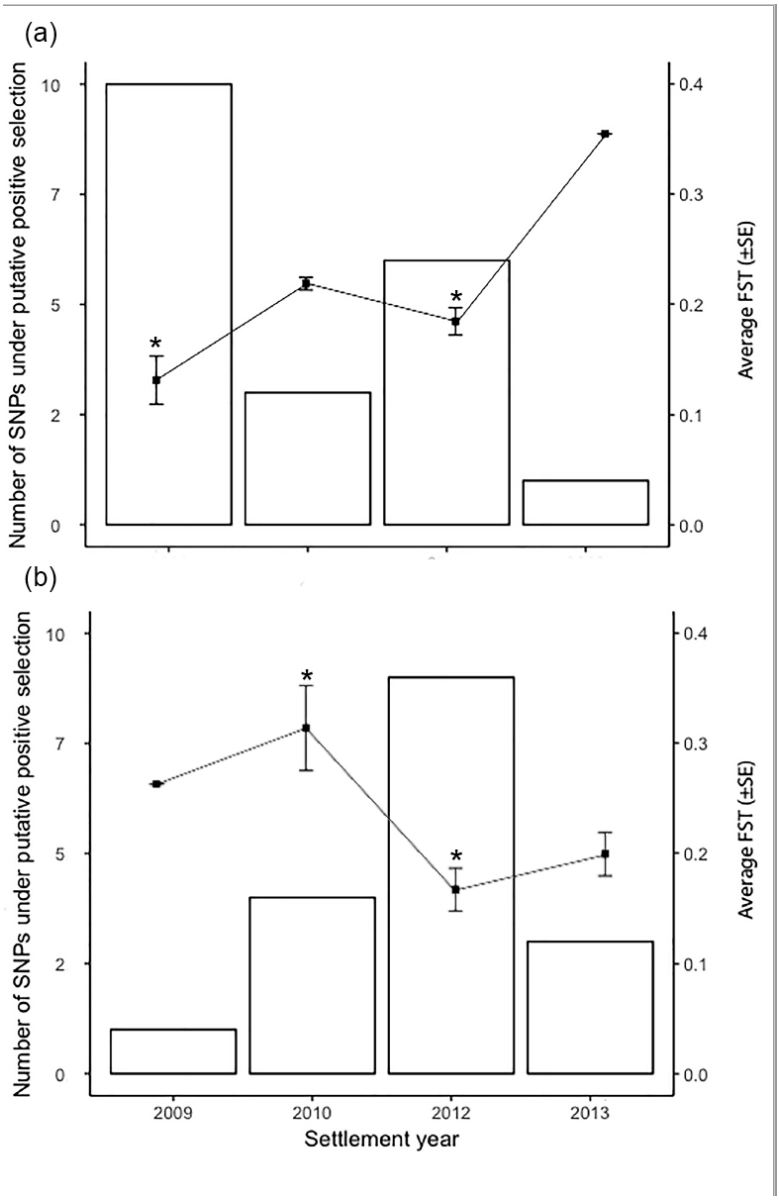



**FIGURE 3** Number of SNPs under putative positive selection and average $F_{ST}$ values per settlement year at (a) Cape Jaffa and (b) Bicheno. The bars represent to the number of SNPs under putative positive selection and the black squares the corresponding average $F_{ST}$ values (±SE). (*) Denotes significant differences in average $F_{ST}$ values between years 2009 and 2012 at Cape Jaffa and years 2010 and 2012 at Bicheno

**TABLE 1** Number of individuals sequenced (n) per year of collection and stage at each collector site

| Collector site | State | Latitude | Longitude | Puerulus year | $n_{Stage\ 1\text{-}2}$ | $n_{Post\text{-}puerulus}$ |
|---|---|---|---|---|---|---|
| Cape Jaffa | South Australia | 36°33'S | 139°24'E | 2009 | 20 | 20 |
| | | | | 2010 | 16 | 18 |
| | | | | 2012 | 19 | 20 |
| | | | | 2013 | 9 | 16 |
| Bicheno | Tasmania | 41°52'S | 148°18'E | 2009 | 18 | 18 |
| | | | | 2010 | 12 | 10 |
| | | | | 2012 | 24 | 15 |
| | | | | 2013 | 20 | 11 |

**TABLE 2** Nei's pairwise $F_{ST}$ values between stage 1-2 puerulus from Cape Jaffa (CJ) and Bicheno (BIC) across year (2009, 2010, 2012, 2013) using neutral SNPs. Values in bold are significant at an $\alpha$=0.05 after a FDR correction

| | CJ 2009 | CJ 2010 | CJ 2012 | CJ 2013 | BIC 2009 | BIC 2010 | BIC 2012 |
|---|---|---|---|---|---|---|---|
| CJ 2010 | **0.0176** | | | | | | |
| CJ 2012 | 0.0028 | **0.0195** | | | | | |
| CJ 2013 | **0.0073** | **0.0178** | 0.0040 | | | | |
| BIC 2009 | **0.0094** | **0.0083** | 0.0043 | **0.0141** | | | |
| BIC 2010 | **0.0274** | -0.0008 | **0.0236** | **0.0240** | **0.0107** | | |
| BIC 2012 | **0.0120** | **0.0149** | **0.0062** | **0.0133** | **0.0153** | **0.0175** | |
| BIC 2013 | **0.0094** | **0.0138** | **0.0051** | 0.0047 | **0.0066** | **0.0139** | **0.0138** |

**TABLE 3** Results of the AMOVA using genetic distance as a function of: puerulus year, site and the interaction of year and site for the neutral SNP panel

| Test | Variance component | Variance | % total | *P* | Φ |
|---|---|---|---|---|---|
| **Neutral SNPs** | | | | | |
| Site | Within samples | 91.543 | 136.139 | 1 | -0.361 |
| | Between samples within site | -24.383 | -36.261 | 1 | -0.363 |
| | Between site | 0.082 | 0.123 | **0.004** | 0.001 |
| Year*Site | Between site within year | 0.14 | 0.209 | **0.005** | 0.002 |

**TABLE 4** Results of the AMOVA using genetic distance as a function of: stage, puerulus year, site and the interaction of stage, year and site for the neutral SNP panel

| Test | Variance component | Variance | % total | *P* | Φ |
|---|---|---|---|---|---|
| Stage | Within samples | 93.985 | 134.702 | 1 | -0.347 |
| | Between samples within stage | -24.251 | -34.757 | 1 | -0.348 |
| | Between stage | 0.038 | 0.055 | **0.01** | 0.001 |
| Year | Within samples | 93.985 | 134.714 | 1 | -0.347 |
| | Between samples within year | -24.27 | -34.788 | 1 | -0.348 |
| | Between year | 0.051 | 0.074 | **0.01** | 0.001 |
| Site | Within samples | 0.065 | 0.093 | 1 | -0.347 |
| | Between samples within site | -24.264 | -34.77 | 1 | -0.348 |
| | Between site | 93.985 | 134.677 | **0.01** | 0.001 |
| Stage*Year | Between year within stage | 0.073 | 0.105 | **0.01** | 0.001 |
| Stage*Site | Between site within stage | 0.104 | 0.149 | **0.01** | 0.001 |
| Year*Site | Between site within year | 0.099 | 0.141 | **0.01** | 0.001 |

**TABLE S1** Nei's pairwise $F_{ST}$ values between stage 1-2 and post-pueruli from Cape Jaffa (CJ) and Bicheno (BIC) using neutral SNPs. Values in bold are significant at an $\alpha$=0.05

| | | POST-PUERULUS | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | CJ2009 | CJ2010 | CJ2012 | CJ2013 | BIC2009 | BIC2010 | BIC2012 | BIC2013 |
| STAGE 1-2 | CJ2009 | **0.0056** | **0.0036** | **0.0162** | **0.0135** | **0.0160** | **0.0158** | **0.0107** | **0.0149** |
| | CJ2010 | **0.0255** | **0.0101** | **0.0256** | **0.0190** | **0.0287** | **0.0177** | **0.0223** | **0.0103** |
| | CJ2012 | **0.0072** | 0.0035 | **0.0110** | **0.0109** | **0.0115** | **0.0144** | **0.0118** | **0.0116** |
| | CJ2013 | **0.0101** | **0.0070** | **0.0062** | **0.0079** | **0.0091** | **0.0110** | **0.0119** | **0.0102** |
| | BIC2009 | **0.0156** | 0.0024 | **0.0175** | **0.0110** | **0.0200** | **0.0114** | **0.0069** | **0.0065** |
| | BIC2010 | **0.0339** | **0.0143** | **0.0267** | **0.0192** | **0.0307** | **0.0219** | **0.0272** | **0.0073** |
| | BIC2012 | **0.0124** | **0.0109** | **0.0180** | **0.0178** | **0.0138** | **0.0179** | **0.0206** | **0.0180** |
| | BIC2013 | **0.0142** | **0.0065** | **0.0106** | **0.0069** | **0.0148** | **0.0132** | **0.0072** | **0.0047** |