This is the author-created version of the following work:


Access to this file is available from:
https://researchonline.jcu.edu.au/65778/

© 2020 John Wiley & Sons Ltd. The Author Accepted Manuscript is available from ResearchOnline@JCU from 21 December 2021. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

Please refer to the original source for the final version of this work:
http://doi.org/10.1111/mec.15780
Global drivers of recent diversification in a marine species complex

Catarina N.S. Silva¹,*, Nicholas P. Murphy², James J. Bell³, Bridget S. Green⁴, Guy Duhamel⁵, Andrew C. Cockcroft⁶, Cristián E. Hernández⁷,⁸, Jan M. Strugnell¹,²

¹ Centre for Sustainable Tropical Fisheries and Aquaculture and College of Science and Engineering, James Cook University, Townsville, Qld 4810, Australia
² Department of Ecology, Environment & Evolution, La Trobe University, Melbourne, Vic 3086, Australia
³ School of Biological Sciences, Victoria University of Wellington, Wellington, 6140, New Zealand
⁴ Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, TAS 7001, Australia
⁵ Département Adaptations du Vivant, BOREA, MNHN, Paris, 75005, France
⁶ Department of Agriculture, Forestry and Fisheries, Cape Town, 8012, South Africa
⁷ Departamento de Zoología, Facultad de Ciencias Naturales y Oceanográficas, Universidad de Concepción, Casilla 160C, Concepción, Chile
⁸ Universidad Católica de Santa María, Arequipa, Perú

*Corresponding author: catari.bio@gmail.com, +61 0434 685 673

Running title: Global drivers of diversification in lobsters

Keywords: demographic inference, environmental association, genomics, Jasus spp., lobsters, speciation-with-gene-flow
Abstract
Investigating historical gene flow in species complexes can indicate how environmental and reproductive barriers shape genome divergence during speciation. The processes influencing species diversification under environmental change remain one of the central focal points of evolutionary biology, particularly for marine organisms with high dispersal potential. We investigated genome-wide divergence, introgression patterns and inferred demographic history between species pairs of all six extant rock lobster species (*Jasus* spp.), which have a long larval duration of up to two years and have populated continental shelf and seamount habitats around the globe at approximately 40°S. Genetic differentiation patterns reflected geographic isolation and the environment (i.e. habitat structure). Eastern Pacific species (*J. caveorum* and *J. frontalis*) were geographically more distant and genetically more differentiated from the remaining four species. Species associated with continental shelf habitats shared a common ancestry, but are geographically distant from one another. Similarly, species associated with island/seamount habitats in the Atlantic and Indian Oceans shared a common ancestry, but are also geographically distant. Benthic temperature was the environmental variable that explained most of the genetic differentiation ($F_{ST}$), while controlling for the effects of geographic distance. Eastern Pacific species retained a signal of strict isolation following ancient migration, whereas species pairs from Australia and Africa, and seamounts in the Indian and Atlantic oceans, included events of introgression after secondary contact. Our results reveal important effects of habitat and demographic processes on the recent divergence of species within the genus *Jasus*, providing one of the first empirical studies of genome-wide drivers of diversification that incorporates all extant species in a marine genus with long pelagic larval duration.

Introduction

The discrete categorization of speciation modes as sympatric, allopatric or parapatric is now considered to be overly simplistic (Butlin *et al.* 2008). Several events (or modes of speciation) can influence the biogeographic states of populations at different time periods during divergence, and as a result, the speciation process is now generally considered to be gradual and reticulated (Smadja & Butlin 2011; Feder *et al.* 2014).
al. 2012). However, the processes responsible for influencing species diversification are still poorly understood and remain one of the central focal points of ecology and evolutionary biology (Arendt et al. 2016).

Reconstructing the diversification history of marine species complexes can be challenging (e.g. Palero et al. 2009; Momigliano et al. 2017) as many have weak genetic differentiation (Ovenden 2013). For marine species, it is often difficult to determine whether weak genetic differentiation is actually present (e.g. as a result of the potential for long distance dispersal) or masked by large population sizes (Lowe & Allendorf 2010). In addition, marine species with long distance dispersal can quickly fill available niches, leaving fewer opportunities for in situ cladogenesis (Pinheiro et al. 2017). As a result, only a few studies have estimated demographic history from genomic data in marine species (e.g. Crow et al. 2010; Le Moan et al. 2016; Momigliano et al. 2017; Souissi et al. 2018; Titus et al. 2019).

Changes in the distribution of marine species resulting from historical climatic variation have been an important driver of diversification across taxa (Davis et al. 2016). Climatic fluctuations during the late Pleistocene, in particular, resulted in periods of isolation intercalated by contact and gene flow between lineages (Hewitt 2000). These glaciatic events dramatically transformed available habitat causing major shifts in species distribution ranges and shaped the genetic structure of many marine species worldwide. For example, Pleistocene glaciations shaped contemporary genetic structure of the abalone Haliotis asinina in the Indo-Pacific (Benardine Jeffrey et al. 2007), the American lobster, Homarus americanus, along the northeastern coast of North America (Kennington et al. 2009), and the octopod Pareledone turqueti in the Southern Ocean (Strugnell et al. 2012). Pleistocene glaciations were also responsible for the divergence of species complexes, such as the Damselfishes Pomacentrus coelestis (Sorenson et al. 2014) and the capelin Mallotus villosus (Dodson et al. 2007; Cayuela et al. 2020). Sequential glacial and interglacial periods have then further shaped the divergence history of species as a result of periods of isolation intercalated by gene flow (Weigelt et al. 2016). A better understanding of the species-specific historical context of divergence is therefore needed to estimate the actual timing and role of gene flow during speciation. Understanding how historical climatic fluctuations shaped species divergence provides clues on how species might respond to
future environmental changes, which is vital for effective conservation and management plans (Olivieri et al. 2016).

Advances in next-generation sequencing (NGS) now provide the opportunity to investigate genome-wide patterns of differentiation along the speciation continuum, allowing the better detection of changes as two lineages diverge from one another on the path to reproductive isolation (Feder et al. 2012). In particular, these methods provide effective tools for species with no reference genomes (Catchen et al. 2017), which is the case for many marine species including rock lobsters (Silva et al. 2019). This technology has also allowed the integration of genomic and environmental data which can be used for testing the hypothesis that selection is more efficient than drift in opposing the homogenizing effects of migration (Manel & Holderegger 2013). In addition, this approach can also detect candidate markers underlying adaptation to local environments for species with moderate to long distance dispersal potential (e.g. Benestan et al. 2016; Sandoval-Castillo et al. 2018). This robust approach is particularly useful in the marine environment where isolation and speciation is increasingly found to be associated with selection/local adaptation (Rocha et al. 2005; Momigliano et al. 2017).

Improvements in methodology have further enabled the use of genome-wide polymorphism data to infer complex demographic histories and the relative influence of gene flow and historical processes on the genomic landscape. For example, in the marine environment this approach has been used in the European anchovy Engraulis encrasicolus (Le Moan et al. 2016), the Atlantic Salmon Salmo salar (Rougemont & Bernatchez 2018), and the corkscrew sea anemone Bartholomea annulate (Titus et al. 2019). One increasingly popular approach is demographic inference based on the computation of a joint allele frequency spectrum (JAFS) from genetic polymorphism data (Gutenkunst et al. 2009; Excoffier et al. 2013). This approach allows an estimation of several demographic parameters such as population sizes, migration rates and time intervals since specific events using a composite likelihood. Therefore, the role of historical events in the diversification and speciation of marine species can now be more accurately determined.

Rock lobsters (Jasus spp.) are a useful model to study the role of historical climatic variations and gene flow on divergence. The six extant Jasus lobster species (J. caveorum, J. edwardsii, J. frontalii, J. lalandii, J. paulensis and J. tristani) are distributed in a narrow latitudinal band (~25° to 47°; Fig. 1a) in the Southern
Hemisphere (Booth 2006) from 0 to 600 m (Holthuis 1991; Duhamel personal communication). These animals have a long pelagic larval duration (PLD; up to two years for *J. edwardsii*), with the potential for extensive dispersal (Bradford *et al.* 2015). Despite such a long PLD, all species have a restricted latitudinal distribution; for example, *J. caveorum* is only known from a single seamount in the eastern South Pacific Ocean (Webber & Booth 1995). Phylogenetic relationships between *Jasus* species have been investigated with a limited number of mtDNA markers (Brasher *et al.* 1992; Ovenden *et al.* 1997). Ovenden *et al.* (1997) identified a clade containing *J. edwardsii, J. lalandii* and *J. frontalii*, however, the relative branching order was not resolved by analysis of sequence variation in the cytochrome c oxidase subunit I (COI) and the 16S ribosomal RNA sequences. In addition, the species *J. tristani* and *J. paulensis*, which occur in islands and seamounts off the southern Atlantic and Indian Oceans, respectively, were hypothesized to have come into secondary contact during past glacial periods, resulting in low levels of mtDNA differentiation (Ovenden *et al.* 1997; Groeneveld *et al.* 2012). At the species level, population genetic studies have demonstrated a general pattern of low, yet often significant, differentiation (Ovenden *et al.* 1992; Matthee *et al.* 2007; Porobić *et al.* 2013; Thomas & Bell 2013; Villacorta-Rath *et al.* 2016). Post-settlement selection and chaotic genetic patchiness, also described as a shifting, ephemeral genetic pattern, has also been observed in *J. edwardsii*, highlighting the uncertainties in predicting connectivity between populations of highly dispersive marine organisms (Villacorta-Rath *et al.* 2018).

Although a few studies suggest a recent divergence between *Jasus* lineages (e.g. divergence within the *J. edwardsii* clade and within the *J. tristani/J. paulensis* clade may be 0.5 million years; Pollock 1990; Ovenden *et al.* 1997), relatively little attention has focused on investigating diversification processes in *Jasus* lobsters. Here we investigate speciation processes in all the extant lobster species of the genus *Jasus*. Given the potential for long distance dispersal, we expect that gene flow (i.e. secondary contact after divergence) between species are relatively common. Therefore, this study aims to determine if admixture/introgression has occurred between species, to investigate the genetic patterns associated with habitat structure (continental shelf or seamount/island) and to infer the demographic history of *Jasus* spp. pairs using genome-wide single nucleotide polymorphisms (SNP).
Methods

Sampling, DNA extractions and sequencing

Tissue samples of *Jasus* spp. were collected between 1995 and 2017 from 17 locations throughout the entire genus’ range (Fig. 1a). A total of 375 samples were collected from 17 populations in total: *J. caveorum* (n=1), *J. edwardsii* (n=5), *J. frontalis* (n=1), *J. lalandii* (n=5), *J. paulensis* (n=2) and *J. tristani* (n=3). Tissue was stored in 70% ethanol before processing. Total genomic DNA of *J. caveorum* historic samples (i.e. collected in 1995) was extracted using the QIAamp DNA Micro Kit (Qiagen) according to the manufacturer’s instruction. The remaining tissue samples were extracted using NucleoMag® Tissue (Macherey-Nagel) following to the manufacturer’s instructions.

Library preparation and sequencing was conducted by Diversity Arrays Technology, Canberra, Australia and followed standard protocols of DArTseq™ genotyping technology (Kilian et al. 2012). Briefly, approximately 100 ng (2 µL) of each sample was digested with the restriction enzymes PstI and SphI, and unique barcode sequences simultaneously ligated onto the ends of each resulting fragment as per Kilian et al. 2012. The PstI-compatible adapter included an Illumina flow-cell attachment sequence, a primer sequence and unique barcode, with the reverse SphI-compatible adaptor contained in the flow-cell attachment region. Size selection was performed using a competitive PCR, where longer fragments and those without both cut sites were excluded. A minimum of 15% random technical sequencing replicates were included for downstream quality control. Each sample with fragments containing both PstI and SphI cut sites was amplified in independent PCR reactions using the following conditions: 94°C for 1 min then 30 cycles of 94°C for 20 s, 58°C for 30 s, 72°C for 45 s, and 72°C for 7 min. Samples were checked visually on an agarose gel to ensure complete digestion and uniform range of fragment sizes. Using approximately 10 µL of each sample, samples were sequenced on a single flow-cell lane on the Illumina HiSeq2500 for 77 cycles.

De novo assembly and variant calling
Libraries were demultiplexed and reads were filtered for overall quality (–c, –q and –r options) using `process_radtags` in STACKS v.2.0b9 (Catchen et al. 2013). The Stacks pipeline `denovo_map.pl` was executed to run each of the Stacks modules individually (`ustacks`, `cstacks`, `sstacks` and `populations`). To optimise the de novo assembly we tested a range of parameters, including $m$ (minimum stack depth) of 3, 5 and 10 and $M$ (distance allowed between stacks) and $n$ (distance allowed between catalog loci) from 1 to 9, as recommended by Rochette & Catchen (2017) and Paris et al. (2017). The parameter test showed that $M=3$ provided a balance between obtaining true polymorphism and introducing sequencing error (i.e. the number of widely shared loci plateaued at about $M=3$) and that $M=3$ was sufficient to stabilize the proportions of loci with 1–5 SNPs (Fig. S1). The high coverage with the value of $m=3$ (64x) and consistent results with $m=3$ to $m=10$ imply a true biological signal (Fig. S1). As $m=3$ also performs well for a broad range of data sets (Paris et al. 2017; Rochette & Catchen 2017) we retained $m=3$ for the main analysis. The formation of loci was allowed with a maximum of two nucleotides between stacks ($M = 3$) and a minimum stack depth of three ($m = 3$) among reads for accurate calling (`ustacks` module). Reads were aligned de novo with each other, and a catalogue of putative RAD tags was created (`cstacks` module). Putative loci were searched against the catalog (`sstacks` module) and further filtering was then conducted in the `populations` module.

Retained SNPs were present in at least 70% of samples within each species, were detected in all species, had a rare allele frequency of at least 2% (to minimize sequencing errors and exclude singletons; Linck & Battey 2019) and had no more than 2 alleles detected. Potential paralogs were excluded by removing markers with heterozygosity > 0.50 within samples and analyses were restricted to one random SNP per locus (using `--write_random_snp`). These filtering steps aimed to exclude as many SNPs as was possible with genotyping errors and missing data.

**Genetic diversity and population structure**

Allelic richness, pairwise $F_{ST}$ values and respective p-values were estimated using `hierfstat` package in R (Goudet 2005). The R package `adegenet` was used to estimate observed and expected heterozygosity and for discriminant analyses of principal components (DAPC) and membership probability plots (Jombart 2008).
DAPC was used on individual genotypes in a multivariate analysis to determine the best number of genetic clusters (K) to retain by running the function `find.clusters()`. Five clusters, five discriminant functions and two principal components (PC) were retained. Inbreeding coefficients were estimated using GenoDive v3 (Meirmans & Van Tienderen 2004). Outlier analyses were conducted in BayeScan to look for signatures of selection. Prior odds were set to 100 to minimize chances of false positives with 5,000 pilot runs, followed by 100,000 iterations (5,000 samples, a thinning interval of 10, and a burn-in of 50,000).

Environmental data collection and analyses

Initially, 13 environmental variables were obtained from Bio-Oracle (Assis et al. 2018; Table S1). Only uncorrelated variables (r<0.6) were retained in order to avoid testing strongly interdependent models and to effectively estimate the relative importance of different factors. This resulted in seven layers (surface and benthic temperature mean, surface salinity, surface and benthic current velocity, benthic iron and surface phytoplankton). Linear mixed-effects models were used to examine the association of geographic distance (estimated as the shortest path distance in the ocean) with patterns of genetic differentiation (measured as pairwise F_{ST} values), using the R package lme4 (Bates et al. 2015). Species was incorporated as a random effect to control for specific demographic histories. We tested associations for all (17) populations and for 15 populations only (i.e. removed the Pacific species *J. frontalis* and *J. caveorum*) as only one location per species was sampled.

Redundancy analysis (RDA; Forester et al. 2018) was used to investigate genotype-environment associations using the R package vegan (Oksanen et al. 2019). RDA is a two-step analysis in which genetic and environmental data are first analysed using multivariate linear regression, producing a matrix of fitted values (i.e. constrained axes) and then a principal component analysis (PCA) of the fitted values produces canonical axes (i.e. unconstrained axes), which are linear combinations of the predictors. Significance was assessed using a permutation test (999 permutations) for redundancy analysis using the function `anova.cca()`.

Relationships among lineages
The program TREEMIX v1.12 (Pickrell & Pritchard 2012) was used to further investigate historical relationships among lineages. A maximum-likelihood (ML) phylogeny was first inferred and then single migration events between branches were sequentially added to determine whether migration/admixture events improve the likelihood fit. To formally test for admixture between *Jasus* spp., the three-population test (Reich *et al.* 2009) included with TREEMIX was used. In this test, the $f_3$ ($X; A,B$) statistic is negative when a population $X$ does not form a simple tree with populations $A$ and $B$, but rather may be a mixture of $A$ and $B$.

As composite likelihood cannot be used directly for formal tests for significance, confidence in individual migration events was estimated using a resampling approach. Therefore “significant” values indicate that the hypothesized migration event significantly improves the fit to the data.

**Demographic modelling**

Previous analysis suggests evidence of admixture between species pairs, and so we tested several hypothesis of divergence modes, aiming to identify speciation events through time, for each closely related pair of species: *J. caveorum* - *J. frontalis*, *J. edwardsii* - *J. lalandii* and *J. tristani* - *J. paulensis*. The species pairs were selected based on their genetic and morphological relationships (Holthuis & Sivertsen 1967; George & Kensler 1970; Brasher *et al.* 1992; Ovenden *et al.* 1997; Groeneveld *et al.* 2012; this study). A set of simple scenarios (i.e. Divergence in Strict Isolation and Isolation-with-Migration) was chosen and complexity was added to these models to take into account specific aspects of divergence between these species. Repeated periods of secondary contact interrupted by isolation were tested (i.e. divergence with Ancient Migration and divergence with Secondary Contact) as species have a long larvae duration and therefore gene flow during divergence is likely. Finally, recent expansion (prefix ‘ex’) was also tested to allow for demographic events following the last glacial maximum (LGM). For each pair, 12 models were built representing alternative modes of divergence considering possible scenarios (Fig S4): (SI) Strict Isolation; (Slex) Strict Isolation with a recent expansion/contraction event; (IM) Isolation-with-Migration; (IMex) Isolation-with-Migration with a recent expansion/contraction event; (AM) Ancient Migration with an ancient gene flow event but recent isolation;
Ancient Migration with an ancient gene flow event but recent isolation and with a recent expansion/contraction event; 
Secondary Contact with a recent gene flow event after past isolation and with a recent expansion/contraction event; 
Repeated Ancient Migration with two ancient gene flow events but recent isolation; 
Repeated Ancient Migration with two ancient gene flow events but recent isolation and with a recent expansion/contraction event; 
Secondary Contact with two recent gene flow events after past isolation; 
Secondary Contact with two recent gene flow events after past isolation and with a recent expansion/contraction event. All models were implemented allowing for asymmetric migration rates (m12, m21).

Demographic inference was performed using the diffusion approximation method implemented in the software \(\delta a\delta i\) (Gutenkunst et al. 2009). The function \(vcf2dadi\) in the R package radiator (Gosselin 2017) was used to create \(\delta a\delta i\) SNP input files. We used the folded joint site frequency spectrum (JSFS) for model selection because the closest out-group (\(Sagmariasus verreauxi\)) was too distant (diverged around 11 Mya; Ovenden et al. 1997), which resulted in a highly reduced number of orientable polymorphisms. To address the impact of missing data in generating JSFS we have projected all spectra down to half of the samples per population (per species pair), as recommended by Gutenkunst et al. (2009). As there was a constant reduction in the number of segregating sites (i.e. the projection did not maximise the number of segregating sites; Table S3) we have used all samples for the inference of the final JSFS.

In total, 12 models were tested per species pair, fitted with the observed joint site frequency spectrum (SFS) using 20 replicate runs per model and the best model was retained (Fig. S3). The Akaike information criterion (AIC) was used to perform comparisons among models (Sakamoto et al. 1986).

To compare among nested models of increasing complexity and address over-parametrization issues we used the comparative framework of Tine et al. (2014) by penalizing models which contain more parameters. For each species pair, a score was estimated for each model using:

\[
\text{Score} = \frac{(\Delta_{\text{max}} - \Delta_{\text{AIC}})}{\Delta_{\text{max}}} \quad (1)
\]
where, $\Delta_{\text{max}}$ corresponds to the difference in AIC between the worst and the best performing model ($\Delta_{\text{max}} = \text{AIC}_{\text{max}} - \text{AIC}_{\text{min}}$) and $\Delta\text{AIC}_i = \text{AIC}_i - \text{AIC}_{\text{min}}$. Therefore, the worst model has a score of 0 and the best model has a score of 1. To evaluate the relative probabilities of the different models within each species pair, Akaike weights ($W_{\text{AIC}}$) were also calculated following:

$$W_{\text{AIC}} = \sum_{i=1}^{R} \frac{e^{-\frac{1}{2} \Delta\text{AIC}_i}}{\sum_{i=1}^{R} e^{-\frac{1}{2} \Delta\text{AIC}_i}^2}$$

where $R$ corresponds to the total number of models considered ($R=12$).

Demographic parameters were converted into indicative biologically units, given the missing crucial information about mutation rate per generation in Jasus spp. The ancestral effective population size ($N_{\text{ref}}$) before split for each species pair was estimated following:

$$N_{\text{ref}} = \frac{\theta}{4L\mu}$$

with $\theta$ being the optimal multiplicative scaling factor, $\mu$ the mutation rate (fixed at $8\times10^{-8}$ mutations per site per generation; Obbard et al. 2012) and $L$ the effective length of the genome explored:

$$L = \frac{zy73}{x}$$

where $x$ is the number of SNPs originally detected from $y$ RAD-tags of 73 bp present in the initial data set, and $z$ the number of SNPs retained, following Rougeux et al. (2017). Estimated units in $2N_{\text{ref}}$ were converted to years assuming a generation time of 10 years (Pecl et al. 2009). Estimated migration rates were divided by $2N_{\text{ref}}$ to obtain the proportion of migrants in every generation.

Results

Genetic diversity and population structure
Sequencing yielded a total of 1,501,921,855 quality-trimmed sequencing reads, providing an average depth of coverage per individual over all SNPs of 58.9x. After applying the different filtering steps, 2,596 SNPs common to all species were retained for subsequent analyses, which had an average depth of coverage of 64x (mean coverage per individual over all SNPs). The lowest values of observed heterozygosity, expected heterozygosity, and allelic richness were observed for J. caveorum and J. lalandii had the highest inbreeding coefficients (0.178). The highest values of allelic richness were observed for J. lalandii (Table 1). The highest pairwise FST values were observed for J. tristani – J. caveorum and J. paulensis – J. caveorum (FST = 0.463 and FST = 0.436, respectively, p < 0.05), while the lowest values were observed for J. tristani – J. paulensis (FST = 0.022, p < 0.01; Table 2).

### Table 1
Summary statistics of genetic diversity per species using 2,596 SNPs. H0: observed heterozygosity, HE: expected heterozygosity, FIS: inbreeding coefficient, AR: allelic richness

<table>
<thead>
<tr>
<th>Species</th>
<th>Sample size</th>
<th>H0</th>
<th>HE</th>
<th>FIS</th>
<th>AR</th>
</tr>
</thead>
<tbody>
<tr>
<td>J. caveorum</td>
<td>11</td>
<td>0.012</td>
<td>0.012</td>
<td>0.094</td>
<td>1.04</td>
</tr>
<tr>
<td>J. frontalis</td>
<td>53</td>
<td>0.064</td>
<td>0.065</td>
<td>0.026</td>
<td>1.32</td>
</tr>
<tr>
<td>J. tristani</td>
<td>41</td>
<td>0.092</td>
<td>0.104</td>
<td>0.127</td>
<td>1.31</td>
</tr>
<tr>
<td>J. lalandii</td>
<td>129</td>
<td>0.086</td>
<td>0.104</td>
<td>0.178</td>
<td>1.60</td>
</tr>
<tr>
<td>J. paulensis</td>
<td>49</td>
<td>0.087</td>
<td>0.103</td>
<td>0.159</td>
<td>1.31</td>
</tr>
<tr>
<td>J. edwardsii</td>
<td>92</td>
<td>0.084</td>
<td>0.100</td>
<td>0.166</td>
<td>1.34</td>
</tr>
</tbody>
</table>

### Table 2
Pairwise FST values (below diagonal) and corresponding p-values (above diagonal) estimated using hierfstat package in R.

<table>
<thead>
<tr>
<th></th>
<th>J. caveorum</th>
<th>J. frontalis</th>
<th>J. tristani</th>
<th>J. lalandii</th>
<th>J. paulensis</th>
<th>J. edwardsii</th>
</tr>
</thead>
<tbody>
<tr>
<td>J. caveorum</td>
<td>0</td>
<td>0.010</td>
<td>0.013</td>
<td>0.010</td>
<td>0.011</td>
<td>0.012</td>
</tr>
<tr>
<td>J. frontalis</td>
<td>0.081</td>
<td>0</td>
<td>0.007</td>
<td>0.005</td>
<td>0.006</td>
<td>0.007</td>
</tr>
<tr>
<td>J. tristani</td>
<td>0.463</td>
<td>0.206</td>
<td>0</td>
<td>0.007</td>
<td>0.008</td>
<td>0.009</td>
</tr>
<tr>
<td>J. lalandii</td>
<td>0.305</td>
<td>0.137</td>
<td>0.387</td>
<td>0</td>
<td>0.006</td>
<td>0.007</td>
</tr>
<tr>
<td>J. paulensis</td>
<td>0.436</td>
<td>0.229</td>
<td>0.022</td>
<td>0.408</td>
<td>0</td>
<td>0.007</td>
</tr>
<tr>
<td>J. edwardsii</td>
<td>0.413</td>
<td>0.106</td>
<td>0.441</td>
<td>0.230</td>
<td>0.452</td>
<td>0</td>
</tr>
</tbody>
</table>
No signatures of selection were detected by the outlier detection analyses (Fig. S2). Lobster species were grouped into three main clusters by discriminant analyses of principal components when using 2 PCs (52.3% variation) (Fig. 1b). There was evidence of admixture, in particular between *J. paulensis* - *J. tristani* in the membership probability plot, the DAPC results and pairwise $F_{ST}$ values (Fig. 1a, b). The first DAPC axis (LD1) explained 29.9% of the variation and highlighted the divergence between habitat structure (i.e. *J. edwardsii* and *J. iodandii* vs. remaining species; Fig. S3a), while the second DAPC axis (LD2), which explained 22.4% of the variation, showed three main clusters and highlighted the differences between *J. paulensis* and *J. tristani* and the remaining species (Fig. S3b).
Genotype-environment associations

Genetic variation (FST) between rock lobster species was significantly correlated with benthic temperature (p<0.001 for 17 populations and p<0.01 for 15 populations) and benthic current velocity.
(p<0.001), when controlling for the effects of geographic distance. Models including geographic distance as
the only predictor of genetic variation were also significant (p<0.05 for 17 populations and p<0.001 for 15
populations; Table 3).

All seven environmental variables explained 18% of the variation in rock lobster species (p<0.001)
when using the constrained ordination in RDA analyses. All values of the variance inflation factors were below
five, indicating that multicollinearity among the predictor variables is not inflating the model. The first five
constrained axes were significant in explaining the genetic variation between species (each explaining 53.3%,
25.3%, 12%, 5.2% and 2.2%, respectively; p<0.001; Fig. 2). Genetic variation of *J. caveorum* and *J. frontalis* was
associated with higher surface temperature, while *J. paulensis* and *J. tristani* were associated with lower
surface temperature. *J. edwardsii* and *J. lalandii* were associated with higher benthic temperature, benthic
current velocity and benthic iron. Finally, *J. lalandii* was associated with higher surface phytoplankton, while
*J. edwardsii* was associated with higher surface current velocity (Fig. 2).

Table 3 Summary of the mixed effects model analyses for all species (17 populations) and for *J. tristani, J.
*lalandii, J. paulensis* and *J. edwardsii* only (15 populations) using *F*$_{ST}$ as a measure of genetic differentiation
(dependent variable) and species as a random effect. GeoDist: geographic distance (km); SurfTemp and
BenTemp: Surface and benthic temperature (°C); SurfSal: Surface salinity (PSS); SurfCurren and BenCurren:
Surface and benthic current velocity (m.s$^{-1}$); BenIron: Benthic dissolved iron (mmol.m$^{-3}$); SurfPhyto: surface
phytoplankton (mmol.m$^{-3}$).

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>All species, 17 populations</th>
<th>Four species, 15 populations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>REML</td>
<td>p-value</td>
</tr>
<tr>
<td>GeoDist</td>
<td>131.9</td>
<td>&lt;0.050</td>
</tr>
<tr>
<td>GeoDist + SurfTemp</td>
<td>137.7</td>
<td>0.336</td>
</tr>
<tr>
<td>GeoDist + BenTemp</td>
<td>110.9</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>GeoDist + SurfSal</td>
<td>133.9</td>
<td>0.124</td>
</tr>
<tr>
<td>GeoDist + SurfCurren</td>
<td>130.2</td>
<td>0.251</td>
</tr>
<tr>
<td>GeoDist + BenCurren</td>
<td>117.7</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>GeoDist + BenIron</td>
<td>125.6</td>
<td>0.921</td>
</tr>
<tr>
<td>GeoDist + SurfPhyto</td>
<td>137.0</td>
<td>0.097</td>
</tr>
</tbody>
</table>
Fig. 2. Ordination plot of redundancy analysis (RDA) of Jasus spp. The vectors are the environmental predictors (see Table 3 for a detailed description).

Relationships among lineages

Results from TREEMIX identified three ancestral events of admixture (Fig. 1c). However, from the three-population test of admixture, only two $f_3$ values were negative with associated Z-scores $<-0.6$, indicating evidence that J. tristani does not form a simple tree with J. paulensis, J. lalandii and J. edwardsii, but rather may be a mixture of these (Table S2, Supporting information). Therefore, the three-population test supported the ancestral event of admixture detected by TREEMIX from the most recent common ancestor (MRCA) of J. lalandii and J. edwardsii to J. tristani. The genetic relationships among species inferred by TREEMIX revealed similar patterns to the genetic differentiation analyses, clearly separating species pairs J. lalandii – J. edwardsii, J. paulensis – J. tristani and J. caveorum – J. frontalis (Fig. 1c).

Demographic modelling

The models for secondary contact (SC) and repeated secondary contact (PSC) provided better fits to the data with good predictions of the joint site frequency spectrum (JSFS) asymmetry for the J. paulensis – J. tristani and J. edwardsii – J. lalandii pairs while ancient migration (AM), repeated ancient migration (PAM) and
strict isolation (SI) had better support for the *J. caveorum*–*J. frontalis* pair (Table S4, Fig. S4). Incorporating population expansion events improved the fit of PSC models for all species pairs but there was not a clear pattern for PAM models. In contrast, the strict isolation (SI) and ancient migration (AM, PAM) models were weakly supported for the *J. paulensis*–*J. tristani* and *J. edwardsii*–*J. lalandii* pairs while the secondary contact models (SC, PSC) were weakly supported for the *J. caveorum*–*J. frontalis* pair (Table S4, Fig. S3).

Asymmetries in gene flow with ratios of $m_{21}/m_{12}$ indicated a stronger migration from population two to population one in all species pairs, and the lower proportion of migrants was observed for the *J. edwardsii*–*J. lalandii* pair (Fig. 3, Table 4). Detailed results for demographic inferences are provided in Fig. 3, Table 4, Fig. S4, Fig. S5, Fig. S6 and Table S4 (supporting information).

The best supported model for the *J. paulensis*–*J. tristani* pair was PSCex (Table S4). Within this model, total divergence time between species was approximately $25,826 \pm 6,286$ years ago (Table 4). The period without contact was approximately 5.4 times longer than the period with secondary contact. The best supported model for the *J. edwardsii*–*J. lalandii* pair was PSCex (Table S4). Total divergence time between *J. edwardsii* and *J. lalandii* was approximately $38,460 \pm 12,242$ years ago (Table 4). The period without contact was approximately 38.8 times longer than the period with secondary contact. Finally, the best supported model for the *J. caveorum*–*J. frontalis* pair was PAMex (Table S4). Within this model, total divergence time between species was approximately $28,709 \pm 12,674$ years (Table 4) and the period without contact was approximately 8.1 times longer than the period with ancient migration. Therefore, divergence times with errors overlap across the three species pairs and was estimated to be between 19,540 and 32,112 years for *J. paulensis*–*J. tristani*, 26,218 and 50,702 years for *J. edwardsii*–*J. lalandii* and 16,035 and 41,383 years for *J. caveorum*–*J. frontalis*. 
Fig. 3. Representation of the best demographic model for each species pair; *J. caveorum* – *J. frontalis*: ancient migration with two periods of ancient gene flow and recent population contraction (PAMex); *J. lalandii* – *J. edwardsii* and *J. tristani* – *J. paulensis*: secondary contact with two periods of contact and recent population expansion/contraction (PSCex). Asymmetric migration rates ($m_{21}$ and $m_{12}$) are represented by the arrows with higher rates of migration from population two to population one for all species pairs (thicker lines in arrows). Width of the boxes represent sizes of the ancestral population ($N_{ref}$), population sizes before expansion/contraction ($N_1$, $N_2$) and population sizes after expansion/contraction ($N_{1e}$, $N_{2e}$). $T_s$ is the time of divergence in strict isolation, $T_{SC/AM}$ the time of secondary contact or ancient migration and $T_e$ the time of expansion.
Table 4. Parameters estimates for the best model of each species pair with standard deviation. *J. paulensis* – *J. tristani*: secondary contact with two periods of contact and recent population expansion (PSCex); *J. edwardsii* – *J. lalandii*: secondary contact and recent population expansion (SCex); *J. caveorum* – *J. frontalis*: ancient migration with two periods of ancient gene flow and recent population contraction (PAMex).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>K</td>
<td>PSCex</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>$N_{ref}$</td>
<td>43.96</td>
<td>85.64</td>
<td>64.19</td>
<td></td>
</tr>
<tr>
<td>$N_1$</td>
<td>3.50 ± 0.84</td>
<td>1.50 ± 0.25</td>
<td>1.30 ± 0.35</td>
<td></td>
</tr>
<tr>
<td>$N_2$</td>
<td>49.67 ± 19.38</td>
<td>41.01 ± 15.61</td>
<td>19.19 ± 9.09</td>
<td></td>
</tr>
<tr>
<td>$N_{xe}$</td>
<td>0.67 ± 0.16</td>
<td>2.44 ± 0.41</td>
<td>0.05 ± 0.24</td>
<td></td>
</tr>
<tr>
<td>$N_{xe}$</td>
<td>2.52 ± 0.53</td>
<td>1.28 ± 0.34</td>
<td>0.20 ± 0.23</td>
<td></td>
</tr>
<tr>
<td>$m_{12}$</td>
<td>11.00 ± 3.77</td>
<td>1.77 ± 0.23</td>
<td>7.71 ± 2.33</td>
<td></td>
</tr>
<tr>
<td>$m_{21}$</td>
<td>2.59 ± 2.22</td>
<td>0.09 ± 0.16</td>
<td>1.99 ± 0.36</td>
<td></td>
</tr>
<tr>
<td>$T_S$</td>
<td>11.94 ± 2.67</td>
<td>10.66 ± 3.37</td>
<td>9.74 ± 4.48</td>
<td></td>
</tr>
<tr>
<td>$T_{SC/AM}$</td>
<td>2.20 ± 0.70</td>
<td>0.27 ± 0.10</td>
<td>1.20 ± 0.19</td>
<td></td>
</tr>
<tr>
<td>$T_e$</td>
<td>0.51 ± 0.20</td>
<td>0.29 ± 0.11</td>
<td>0.24 ± 0.27</td>
<td></td>
</tr>
<tr>
<td>$T_{total}$</td>
<td>29.29 ± 7.15</td>
<td>22.46 ± 7.15</td>
<td>22.36 ± 9.87</td>
<td></td>
</tr>
<tr>
<td>$*m_{12}$</td>
<td>0.12 ± 0.04</td>
<td>0.01 ± 0.001</td>
<td>0.06 ± 0.02</td>
<td></td>
</tr>
<tr>
<td>$*m_{21}$</td>
<td>0.03 ± 0.03</td>
<td>0.00 ± 0.00</td>
<td>0.01 ± 0.00</td>
<td></td>
</tr>
<tr>
<td>$*T_S$</td>
<td>10,524 ± 2,346</td>
<td>18,258 ± 5,769</td>
<td>12,507 ± 5,746</td>
<td></td>
</tr>
<tr>
<td>$*T_{SC/AM}$</td>
<td>1,941 ± 618</td>
<td>470 ± 167</td>
<td>1,541 ± 239</td>
<td></td>
</tr>
<tr>
<td>$*T_e$</td>
<td>447 ± 179</td>
<td>503 ± 185</td>
<td>306 ± 352</td>
<td></td>
</tr>
<tr>
<td>$*T_{total}$</td>
<td>25,826 ± 6,286</td>
<td>38,460 ± 12,242</td>
<td>28,709 ± 12,674</td>
<td></td>
</tr>
</tbody>
</table>

K: The number of free parameters in the model
$N_{ref}$: The effective size of the ancestral population before the split
$N_1$: The effective size of population 1 before expansion in units of $2N_{ref}$ generations
$N_2$: The effective size of population 2 before expansion in units of $2N_{ref}$ generations
$N_{xe}$: The effective size of population 1 after expansion in units of $2N_{ref}$ generations
$N_{xe}$: The effective size of population 2 after expansion in units of $2N_{ref}$ generations
$m_{12}$: The neutral movement of genes from population 2 to population 1 in units of $2N_{ref}$ generations
$m_{21}$: The neutral movement of genes from population 1 to population 2 in units of $2N_{ref}$ generations
$T_S$: The time of divergence in strict isolation in units of $2N_{ref}$ generations
$T_{SC/AM}$: The time of secondary contact/ancient migration in units of $2N_{ref}$ generations
$T_e$: The time of expansion in units of $2N_{ref}$ generations
$T_{total}$: The total time since the split in units of $2N_{ref}$ generations
$*m_{12}$: The proportion of migrants per generation from population 2 to population 1
$*m_{21}$: The proportion of migrants per generation from population 1 to population 1
$*T_S$: The time of divergence in strict isolation in units of numbers of years
$*T_{SC/AM}$: The time of secondary contact/ancient migration in units of numbers of years
$*T_e$: The time of expansion in units of numbers of years
$*T_{total}$: The total time since the split in units of numbers of years
Discussion

Here we investigated genome-wide divergence and introgression patterns in all extant species of rock lobsters (*Jasus* spp.) for the first time. Genetic differentiation patterns revealed the effects of geographical isolation and the environment (i.e. habitat structure), with benthic temperature being the environmental variable that explained most of the genetic differentiation ($F_{ST}$) while controlling for the effects of geographic distance. Closely related species pairs were identified and for all three species pairs we detected recent divergence and multiple introgression events (gene flow) since first divergence.

Drivers of speciation

In our study, the Eastern Pacific species pair (*J. caveorum* and *J. frontalis*) were genetically more differentiated from the remaining species. Therefore, geographic distance and reduced ocean current velocity in the Southern Pacific might have driven the differentiation between the Eastern Pacific and the remaining *Jasus* species. On the other hand, the main driver of differentiation between the other species pairs, *J. tristani* - *J. paulensis* and *J. lalandii* - *J. edwardsii*, was habitat structure. Species that were associated with the same habitat structure (continental shelf or seamount/island) were genetically more closely related to each other than to species from the alternate habitat, despite these habitat types being geographically interspersed. For example, *J. edwardsii* and *J. lalandii* (associated with continental shelf habitat) were genetically more similar to each other ($F_{ST}$=0.230), as were *J. tristani* and *J. paulensis* (from island/seamount habitat; $F_{ST}$=0.022). Yet *J. lalandii* is geographically much closer to *J. tristani* ($F_{ST}$=0.387) and *J. paulensis* ($F_{ST}$=0.408) than it is to *J. edwardsii* (Fig. 1). It is important to note that these results must be taken with caution as sample size is necessarily limited. Although connectivity is possible between these habitats given the high larval dispersal potential (indeed, *J. lalandii* larvae have been found in the southwest Indian Ocean as far east as Amsterdam Island, adjacent to the *J. paulensis* habitat (Booth & Ovenden 2000)), species appear to be adapted to local environmental conditions.

Our results suggest that benthic temperature is associated with the differentiation between species from island/seamount and continental shelf habitats. Benthic temperature explained most of the genetic
differentiation ($F_{ST}$) while controlling for the effects of geographic distance, and *J. edwardsii* and *J. lalandii* were associated with higher mean benthic temperatures. Temperature is important for regulating the rate of embryological development in lobsters (Phillips 2013) and thermal adaptation has been detected in American lobster *Homarus americanus* (Benestan et al. 2016). Therefore, sea temperature could limit reproduction between populations of *Jasus* spp. adapted to local benthic temperatures.

Other factors may have also played a role in isolating locally adapted populations/species despite potential for long distance dispersal. Pollock (1990) noted that genetic differentiation between *Jasus* species may have been caused by interruptions of gene flow between populations as a result of topographic forcing of ocean currents by shoaling ridges, rises and island/seamount chains. For example, larvae may be retained within permanent eddies found offshore (Chiswell & Booth 1999), which may also be limiting dispersal between habitats (continental shelf and islands/seamounts). In addition, despite *Jasus* larvae being weak horizontal swimmers, they can migrate vertically in the water column (Booth 1994) and therefore can use this to their advantage to avoid being transported offshore or towards unsuitable habitat. Differential settlement trends between species can also play a role in isolating locally adapted populations/species. It has been shown that during the post-larval or puerulus stage, *J. edwardsii* are able to recognize environmental cues such as chemical, acoustic and substrate cues, and increase settlement success in suitable habitats (Stanley et al. 2015; Hinojosa et al. 2016, 2018). Therefore, a combination of factors can drive divergence between *Jasus* spp. regardless of geographic location.

Divergence times within species pairs estimated by the best fitting demographic models in this study overlap across all species comparisons (from 38,460± 12,242 to 25,826 ± 6,286 years ago), suggesting that global changes in environmental conditions might have driven initial divergence across all species pairs. Such a scenario closely matches the recent radiations inferred for the synonymous genera of perciform fishes *Nemadactylus* and *Acantholatris*, which overlap in distribution with *Jasus* spp. and also have a long pelagic larval stage of 7-12 months (Burridge 1999). Therefore, we hypothesise that global changes in environmental conditions have similarly shaped the divergence patterns of these marine species complexes.
Decreasing global mean sea level and a change in the Southern Ocean (SO) barotropic stream function could have contributed to the co-occurring patterns of divergence among species pairs. The SO barotropic stream function is a measure of the strength of the Antarctic Circumpolar current (ACC), which drives currents in a predominately clockwise direction around Antarctica linking the major basins of the world. A significant change in the direction of the SO barotropic stream function occurred around 28,500 years ago (Fogwill et al. 2015; Fig 5e and f therein), which coincides with the expansion of Antarctic sea ice and the divergence period (within species pairs) estimated in our study (Fig. 4).

Fig. 4. Changes in global sea level (light green, adapted from Huybrechts 2002) and global mean proportion of available habitat in the benthic photic zone, from 0 to 65 metres (dark green, Schaaf 1996) relative to present day. 1) Antarctic sea ice expands around 28,500 years ago (Fogwill et al. 2015); 2) The end of the last glacial maximum (around 19,000 years ago; Clark et al. 2009); 3) West Antarctic Ice Sheet deglaciation (around 14,500 years ago; Clark et al. 2009). Divergence period corresponds to the initial divergence period between species pairs estimated from the demographic inference in this study.

Secondary contact

In our study, the best fitting demographic models showed that for all species pairs the periods of isolation were longer than the periods of gene flow after divergence occurred. Past periods of isolation were
estimated to be between 5- and 39-fold longer than periods of gene flow for the *J. paulensis* – *J. tristani* pair and *J. edwardsii* – *J. lalandii*, respectively. The genetic structure observed reflects these demographic history characteristics, as the *J. paulensis* – *J. tristani* pair showed more distinct signatures of admixture than the *J. edwardsii* – *J. lalandii* pair. The best fitting demographic models also showed repeated periods of isolation followed by secondary contact for all species pairs. These patterns likely result from the impact of Quaternary climatic oscillations on *Jasus* populations, which are known to have promoted retreats towards lower latitudes during glacial periods (e.g. Kenchington et al. 2009; Le Moan et al. 2016; Jenkins et al. 2019) and thereby created new opportunities for isolation/contact during transition to interglacial periods.

During the last glacial maxima (LGM) about 19,000-22,000 years ago, temperature and sea levels reached minimum values (Clark et al. 2009). Low temperatures likely drove a northward shift in the distribution of species as they tracked their thermal optima. A more northerly distribution of *J. lalandii* and *J. edwardsii* than the present day, for example, could result in a reduction in connectivity between these species as larval dispersal would be less influenced by ocean currents flowing eastwards (e.g. Southern Ocean Subtropical Front; Gersonde et al. 2003). In our study, demographic models showed predominant eastward migration from *J. tristani* to *J. paulensis* and from *J. lalandii* to *J. edwardsii* and a predominant westward migration from *J. frontalis* to *J. caveorum*. Although migration estimates are averaged across all events of secondary contact, these estimates reflect the predominant direction of connectivity between the *J. paulensis* – *J. tristani* and *J. edwardsii* – *J. lalandii* pairs, both influenced by the Southern Ocean Subtropical Front and the *J. frontalis* – *J. caveorum* pair, influenced by the Humboldt current (along the western coast of South America).

Low temperatures and sea levels during the LGM also resulted in changes in available habitat, which have possibly impacted species associated with continental shelf habitat (*J. lalandii* and *J. edwardsii*) and with seamount/islands (*J. tristani* and *J. paulensis*) differently. Schaaf (1996) used a theoretical approach to quantify the reduction/augmentation of the photic sea-bottom area (from 0 to 65 metres) during sea level fall/rise. During the LGM the global mean proportion of available habitat in the benthic photic zone was at its lowest (Fig. 4) and this shrinkage of benthic habitat was more pronounced in the continental margins than on
seamounts and oceanic islands (Schaaf, 1996; Fig. 5 therein). Changes in habitat area varied with the habitat and region, and were highly dependent on the amount/direction of sea level change (Schaaf 1996; Holland 2012). During this period, areas in ocean ridges that were previously too deep for lobsters to inhabit might have become suitable but, at the same time, shallow areas in the continental shelf and around islands/seamounts might have become exposed and unsuitable for lobsters.

Transition from the LGM to the Holocene precipitated further changes in the available shallow benthic habitat for lobsters. The deglaciation of the West Antarctic Ice Sheet which was the primary source of an abrupt rise in sea level around 14,500 years ago (Clark et al. 2009; Fig. 4) coincided with a major expansion of available shallow benthic habitat (global averages) at around 14,000 years ago (Schaaf 1996). A southward shift in the distribution of Jasus spp. while tracking their thermal optima would have increased connectivity between species resulting in the recent periods of secondary contact and admixture observed in the demographic models and population genetic structure analyses, respectively. A combination of fluctuating conditions in temperature, sea level and available habitat have likely resulted in alternating periods of isolation and gene flow that have shaped the speciation processes of Jasus lobsters, which might still be occurring, given the recent divergence times within species pairs.

Our study provides genome-wide evidence of admixture between J. paulensis - J. tristani. Although George & Kensler (1970) have noted that J. tristani and J. paulensis possess a significant difference in the abdominal sculpture index, previous genetic evidence (using the mitochondrial cytochrome oxidase I gene; Groeneveld et al. 2012) suggests that these species can be synonymized as J. paulensis. Our results demonstrate that, since initial divergence, J. tristani and J. paulensis spent 4.1 times longer in secondary contact than J. edwardsii – J. lalandii and 1.2 times longer than J. caveorum and J. frontalis. The Tristan da Cunha and Gough Islands (current distribution of J. tristani) and the Amsterdam and St. Paul Islands (current distribution of J. paulensis) have been grouped in the same zoogeographic province (called the West Wind Drift Islands Province) based on endemic fish fauna distribution (Collette & Parin 1991). Archipelagic level endemism is common in marine taxa (Paulay & Meyer 2002) and has been observed for other species (e.g.
snails of the genus *Echinolittorina*; Williams & Reid 2004). Therefore, archipelagic endemism and the long periods of gene flow may explain the close relationship between *J. tristani* and *J. paulensis*.

**Conclusion**

In highly dispersive marine taxa, interglacial recolonization of high-latitude habitats can occur rapidly (Hewitt 2000). Such patterns have been established for a range of Northern Hemisphere marine species (e.g. Marko 2004; Ledoux *et al.* 2018), but relatively little is known about the genetic effects of recent glaciations in the Southern Hemisphere (but see e.g. Fraser *et al.* 2009; Strugnell *et al.* 2012; Porobić *et al.* 2013). This study revealed genome-wide patterns of divergence and introgression in all extant species of a highly dispersive marine taxa for the first time. While results point to the important role of demographic and neutral processes of differentiation between species pairs, it also suggests a possible effect of selection in promoting genetic divergence between habitats. Future studies should address the role of adaptive processes to elucidate their relative contribution in shaping genome divergence and speciation of *Jasus* lobsters and to better understand how future environmental change will impact species distribution.

**Acknowledgments**

Funding for this research was provided by an Australian Research Council Discovery Project grant (Project No. DP150101491) awarded to J.M.S., J.J.B., B.S.G. and N.P.M. We would like to thank Gary Carlos (University of Tasmania), Colin Fry (University of Tasmania), Daniel Ierodiaconou (Deakin University), Kent Way, Andrew Kent, Geoff Liggins, Marcus Miller, Giles Ballinger, Darrel Sykes (DAFF), Rick Webber (Te Papa Museum), Jason How (Department of Fisheries, Western Australia) and Sadie Mills (NIWA) for field assistance and sample collection; T.A.A.F. (Terres Australes et Antarctiques Françaises), for their French fisheries observer service “COPEC”, the fishery observer Sophie Cascade on board the F.V. “AUSTRAL” and Charlotte Chazeau to have made available biological scientific samples and data of *Jasus paulensis* from catches in the Saint-Paul/Amsterdam French EEZ; the help of crew has also been appreciated. C.E.H was supported by FONDECYT grants #1170815 and #1201506. We would also like to thank Ira Cooke for the helpful discussions about...
demographic modelling, the journal Editor, Alan Le Moan and two other anonymous reviewers for the constructive comments.

Author contributions
All authors contributed insights about data analysis, interpretation of results and edited the final drafts of the manuscript. C.N.S.S. analysed the data. C.N.S.S., J.M.S., B.S.G. and N.P.M. conceived the study.

Data availability
Sequencing data, pipelines for de novo assembly, genetic structure, environmental association and demographic inference analyses are available at:

References


742 Rougemont Q, Bernardzhe L (2018) The demographic history of Atlantic salmon (Salmo salar) across its
distribution range reconstructed from approximate Bayesian computations*. *Evolution*, 72, 1261–1277.
toward inferring the divergence history of lake whitefish species pairs (Coregonus clupeaformis).
744 *Genome Biology and Evolution*, 9, 2057–2074.
adaptive divergence in a connected and commercially important mollusc, the greenlip abalone (Haliotis
747 Schaa A (1996) Sea level changes, continental shelf morphology, and global paleoecological constraints in
748 Silva CNS, Villacorta-Rath C, Woodings LN, Murphy NP, Green BS, Hartmann K, Gardner C, Bell JJ, Strugnell
JM (2019) Advancing our understanding of the connectivity, evolution and management of marine
749 Smadja CM, Butlin RK (2011) A framework for comparing processes of speciation in the presence of gene
coelestis* species complex (Pisces: Pomacentridae): historical biogeography and species boundaries.
751 Marine Biology, 161, 2495–2507.
of differential introgression between two divergent fish species (Solea spp.). *Heredity*, 121, 579–593.
754 Strugnell JM, Watts PC, Smith PJ, Allcock AL (2012) Persistent genetic signatures of historic climatic events in
755 Thomas L, Bell JJ (2013) Testing the consistency of connectivity patterns for a widely dispersing marine
Bonhomme F et al. (2014) European sea bass genome and its variation provide insights into adaptation
toward euryhalinity and speciation. *Nature Communications*, 5, 5770.
757 Titus BM, Blischak PD, Daly M (2019) Genomic signatures of sympatric speciation with historical and
758 Villacorta-Rath C, Ilyushkina I, Strugnell JM, Green BS, Murphy NP, Doyle SR, Hall NE, Robinson AJ, Bell JJ

