Introgression between wild and selectively-bred Sydney rock oysters (*Saccostrea glomerata*)

Jessica Ann Thompson

Department of Biological Sciences

Faculty of Science and Engineering

Macquarie University

Supervisors:

Professor David Raftos & Associate Professor Adam Stow

9th October 2015

Word count: 7631

This thesis is written in the form of a manuscript for submission to Marine & Freshwater Research, with the exceptions that the Introduction and Discussion sections are extended and figures are integrated into the text. Author instructions for Marine & Freshwater Research are appended.

Declaration:

I wish to acknowledge the following assistance in the research detailed in this report:

David Raftos, Adam Stow, Melanie Bishop, Paolo Momigliano and Elayna Truszewski for assistance with experimental design and analyses. David Raftos, Emma Thompson, Maria Vozzo, Michael Dove, Wayne O'Connor, NSW Department of Primary Industries, and Peter David for assistance with collection of samples. Diversity Arrays Pty Ltd for DNA extractions, next generation genotyping and SNP identification. Daniel Powell and Abigail Elizur (University of the Sunshine Coast) for assistance with characterising loci under selection. David Raftos, Adam Stow, Liam O'Hare and Peterson Coates for comments on a draft of this manuscript.

All other research described in this report is my own original work.

JESSICA ANN THOMPSON

9TH OCTOBER 2015

Introgression between wild and selectively-bred Sydney rock oysters (*Saccostrea glomerata*)

Abstract

Oysters provide vital ecosystem services for coastal ecosystems worldwide. Sydney rock oysters (*Saccostrea glomerata*) are among the most important native Australian oysters. In 1990, a selective breeding program for faster growth rates was implemented for *S. glomerata*. This has since expanded to include breeding for resistance to winter mortality and QX disease. Selectively-bred oysters are currently farmed in estuaries containing wild oyster populations, providing the opportunity for interbreeding that could alter the genetic variability and structure of the wild populations. Here, we use next-generation genotype-by-sequencing to investigate the genetic structures of wild populations and the selectively-bred B2 line of *S. glomerata*. These data are used to test for genetic introgression between the populations at two sites in the Georges River, an estuary in Sydney where selectively-bred oysters have been farmed since 1990. Strong genetic partitioning was identified between wild and selectively-bred populations and there was no evidence of sustained gene flow in the form of introgression. Contrary to our expectations, we found significantly higher levels of genetic diversity and heterozygosity in the selectively-bred population than the wild population. These results potentially reflect the impact of population bottlenecks and are relevant to the resilience of this species to environmental change.

Additional Key Words

Saccostrea glomerata; aquaculture; introgression

Introduction

Oysters are environmentally and economically important organisms (Schrobback *et al.* 2014). They form the structural and trophic foundation of coastal environments worldwide (Beck *et al.* 2011). Oyster reefs often provide the only hard substrate in ecosystems predominantly covered by soft sediment and are responsible for reducing wave energy and coastal erosion (Grabowski *et al.* 2007). They also sustain significant levels of biodiversity by providing refugia for a wide range of organisms (Coen *et al.* 2007; Beck *et al.* 2011). As filter feeders, oysters exert top-down control of phytoplankton blooms to prevent eutrophication, as well as improving overall water quality (Grabowski *et al.* 2007). They also link benthic and pelagic environments through nutrient cycling and biodeposition (Newell 2004; Coen *et al.* 2007). These processes facilitate increased levels of productivity within the ecosystem, as the reduction of turbidity via filter feeding promotes the growth of benthic aquatic plants (Newell 2004; Coen *et al.* 2007; Grabowski *et al.* 2007). In addition, oysters act as carbon sinks by sequestering carbonates into their shells (Grabowski *et al.* 2007).

Coastal oyster reefs were historically extensive and abundant. However, more than 85% of oyster reefs have been lost globally in recent times (Beck *et al.* 2011). This loss is due largely to the degradation of water quality and excessive harvesting of oysters (Grabowski *et al.* 2007). Many remaining oyster reefs are deteriorated to the point that they can no longer provide vital ecosystem services (Beck *et al.* 2011; Diggles 2013). It is likely that any remnant oyster populations have also been depleted of genes for fast growth and resilience. Historically, human exploitation of natural oyster beds has selectively harvested the largest, presumably fastest growing individuals (Gaffney 2006). As a result, some evidence suggests that wild oyster populations have decreased resilience to environmental stressors (Lenihan *et al.* 1999; Butt *et al.* 2006) and are more susceptible to disease (Lenihan *et al.* 1999). This depletion of natural resilience in wild oyster populations may have increasingly detrimental effects in the face of global climate change and increasing human development in coastal areas. By corollary, the decline of natural oyster populations and the associated loss of essential ecosystem services from coastal estuaries could have increasingly severe environmental and ecological impacts.

In Australia, Sydney rock oysters (*Saccostrea glomerata*; Gould 1980) are the most important native oyster in terms of ecoengineering. Although *S. glomerata* are most abundant along the coastlines of New South Wales (NSW) and Queensland, their natural range is extensive, covering the length of the eastern Australian coast from Victoria through NSW to Queensland, and also around the northern coast as far as Western Australia (Nell 2001). In line with global losses, wild Sydney rock oyster populations in Australia have been severely degraded, such that they are currently listed as functionally extinct (Beck *et al.* 2011). These remnant populations may also be increasingly susceptible to changing environmental conditions. For instance, CO_2-driven decreases in pH leading to ocean acidifcation are known to negatively affect the fertilisation and embryonic development of Sydney rock oyster larvae (Parker *et al.* 2009). This negative impact increases in severity when combined with suboptimal temperatures (Parker *et al.* 2009). Similar environmental stressors have also been found to increase the susceptibility of Sydney rock oysters to disease. Changes in temperature and salinity stimulate the production of the stress hormone noradrenaline, which inhibits immunological activity and increases disease susceptibility (Butt *et al.* 2006; Aladaileh *et al.* 2008). Slower growth rates of wild Sydney rock oyster populations compared to invasive species may also affect ecological competition. Wild Sydney rock oyster beds are often overgrown by exotic Pacific oysters (*Crassostrea gigas*), which grow twice as fast as the native species (Nell 1993; Krassoi *et al.* 2008; Bishop *et al.* 2010; Schrobback *et al.* 2014).

Due to their historical abundance, Sydney rock oysters are also one of the most economically important bivalve aquaculture species in Australia. They have been farmed commercially since 1870 and currently comprise over half of the edible oysters produced in Australia each year (Banks *et al.* 2006; O'Connor and Dove 2009). Sydney rock oyster production currently exceeds AUD$100 million annually (Schrobback *et al.* 2014). Operational aspects of the Sydney rock oyster aquaculture industry have changed substantially over recent decades to deal with these limitations of remnant, wild Sydney rock oyster populations. The spat (juvenile oysters) farmed in many growing areas are now bred in hatcheries rather than caught by natural recruitment within the growing area (Nell 2001). The development of hatchery technology has allowed for the efficient management of the industry and the enhancement of selective breeding programs (Nell 2001). Selective breeding of Sydney rock oysters has been undertaken by the NSW

Department of Primary Industries (DPI) since 1990 (Nell and Hand 2003). It has been based on mass selection, whereby individuals that display desired characteristics are identified, pooled together and artificially spawned to produce successive generations. The program initially aimed to increase the performance and growth rates of Sydney rock oysters. Hence, the largest individuals in each generation were selected for use in subsequent breeding events (Nell and Hand 2003). In 1997, the program was expanded to include selection for disease resistance, due to major outbreaks of two diseases (winter mortality and QX disease) that heavily impacted commercial production of Sydney rock oysters (Nell and Hand 2003; Simonian *et al.* 2009). Subsequently, only the survivors of disease outbreaks were used for mass selection.

The genetic relationships between farmed Sydney rock oysters and their wild conspecifics are likely to have been substantially altered as a result of these altered industry practices. Increased reliance on hatchery-produced spat means that new recruits for oyster crops are no longer sourced locally. Consequently, there is often a disjunct between the provenance of farmed oysters and the wild oyster population in many growing areas. It is likely that the differences in the genetic constitution between farmed and wild oysters are continuing to increase because most hatchery-produced spat are now the product of breeding programs, which have substantially altered many traits among selectively-bred oysters. Sydney rock oysters selected over five generations for fast growth reach market size a full year earlier than non-selected controls (Nell and Perkins 2005). Similarly, oysters in the 'B2' selective breeding line, which has been selected for dual resistance against QX and winter mortality, appear to have developed functional resistance to these diseases. B2 oysters exposed to both diseases currently exhibit mortality rates that are within the range of expected background mortality in the absence of disease (Nell and Perkins 2006). Coincidentally, B2 oysters are also more tolerant of ocean acidification (decreased pH) than non-selected Sydney rock oysters (Parker *et al.* 2011; Parker *et al.* 2012). This suggests that the selective breeding program has successfully replenished or enhanced desirable genetic traits within farmed Sydney rock oyster populations. However, it is also likely to have resulted in genetic differentiation between wild and selectively-bred Sydney rock oysters.

Differences between wild and selectively-bred Sydney rock oysters are clearly evident at the phenotypic level. However, there is little current understanding of the underlying genetic basis for the observed differences in growth, disease resistance and environmental resilience. Such detailed genetic information is crucial for two reasons. First, it will help to identify genetic markers that can be used in new marker-assisted selection programs, comparable to those being developed in Australia for the silver-lipped pearl oyster (*Pinctada maxima*; Jones *et al.* 2013). Second, comprehensive genetic data can be used to assess potential impacts that the farming of selectively-bred oysters may have on the wild oyster populations through gene flow between populations. Oysters reproduce by non-selective broadcast spawning (Parker *et al.* 2009). So, the coexistence of wild and selectively-bred oysters in farming locations provides the opportunity for interbreeding. Such interbreeding has the potential to alter the genetic variability and structure of wild Sydney rock oyster populations, with potentially beneficial or detrimental outcomes (Parker *et al.* 2012).

The procedure of selective breeding may result in unfavourable genetic characteristics that may be detrimental if introgression was to occur. Selective breeding programs generally have small founder population sizes and increased potential for inbreeding. This can decrease levels of genetic variation and fitness for traits other than those under active selection (English *et al.* 2001; Green *et al.* 2009). If gene flow is occurring between wild and farmed oysters, reduced genetic variation within the farmed selectively-bred population could be transferred into the wild population with widespread consequences at the ecosystem, community and population level (Hughes *et al.* 2008). Although no evidence of significant inbreeding or reduced genetic variation has been found among the selectively-bred lines of Sydney rock oysters (English *et al.* 2001; Simonian *et al.* 2009), there is the risk that it may occur as the breeding program progresses.

In contrast, gene flow could also transfer beneficial alleles for fast growth rates, disease resistance and environmental resilience from the selectively-bred farmed populations into local wild populations (Parker *et al.* 2012). Such beneficial outcomes have been observed in other species. For instance, the introgression of genes from domesticated goats into the

closely related Alpine ibex (*Capra ibex*) substantially improved the genetic variability and immunocompetence of the wild ibex population (Grossen *et al.* 2014). In the case of Sydney rock oysters, the introgression of beneficial genes could help to future-proof wild populations against QX and winter mortality disease, ocean acidification and other factors associated with environmental change. Faster growth rates may also allow wild Sydney rock oysters to more effectively compete with introduced Pacific oysters.

At present, the majority of studies examining the effects of cultured or genetically-altered organisms on wild populations are focused on terrestrial ecosystems, mainly involving crop species (Stewart *et al.* 2003; Warwick *et al.* 2009). Even within the literature for aquatic species, the majority of studies investigate farmed fish species, not molluscs (Utter and Epifanio 2002; Seamons *et al.* 2012; Heino *et al.* 2015). As such, very little research has assessed the possible deleterious impacts of hatchery-produced or selectively-bred oysters on the genetic structures of local wild oyster populations. Similarly, few studies have assessed the potentially beneficial use of selectively-bred oysters in ecological restoration projects, even though hatchery-produced spat are already being used to boost population sizes in areas where wild oysters have declined (Kingsley-Smith *et al.* 2009). Some oyster restoration projects in the USA already use hatchery-produced, selectively-bred Eastern oysters (*Crassostrea virginica*), although their impact on the wild population is largely unknown (Carlsson *et al.* 2008).

To rectify the lack of detailed population genetic data on *S. glomerata*, the current study uses next-generation genotype-by-sequencing to investigate the genetic structures of wild oyster populations and the selectively-bred B2 line. These data are used to test the hypothesis that gene flow resulting in genetic introgression is occurring between the farmed and wild populations of *S. glomerata* in a single coastal estuary where selectively-bred oysters have been farmed since 1990 (Nell *et al.* 2000).

Materials and Methods

Collection of oyster samples

Sydney rock oysters (*S. glomerata*) were collected from two sites in the Georges River, New South Wales; Woolooware Bay (34°02'14.2"S 151°08'51.5"E) and Quibray Bay (34°01'29.7"S 151°10'50.3"E; Fig. 1). Both sites experience QX disease outbreaks each year. Woolooware Bay is more heavily affected by QX than Quibray Bay, and Quibray Bay also experiences outbreaks of winter mortality (Nell 2006). Wild spat caught within the estuary have been farmed in the Georges River since the 1880s. Selectively-bred oysters, including the B2 line and its progenitors, have been grown at both sites since 1990 (Nell *et al.* 2000). Approximately 85% of the oysters currently farmed at the two sites are selectively bred B2 oysters, with the remainder wild caught within the Georges River estuary (Dr Mike Dove, NSW DPI, personal communication).

Three groups of Sydney rock oysters (comprising 12 oysters per group) were collected within each of the two bays. The three groups from each bay were; (i) wild oysters collected from a 10 – 20 m stretch of shoreline approximately 1 km from oyster farming leases (designated wild-caught, WC); (ii) oysters collected from sticks or pylons supporting oyster farming racks (designated overcatch, OC); (iii) selectively-bred, 5th generation B2 oysters from within the farming racks (designated B2). The B2 oysters were kindly supplied by NSW DPI.

In addition to oysters from the Georges River, two reference groups were also provided by NSW DPI. These were hatchery-reared 5th generation B2 (designated hatchery B2, HB2, n=11) and "control" oysters (designated hatchery controls, HC, n=10) grown in Port Stephens, New South Wales (32°44'12.5"S 152°03'18.9"E). The Port Stephens estuary is approximately 170km north of the Georges River estuary and has never experienced outbreaks of QX disease or winter mortality (Nell 2006). The hatchery B2 oysters were of the same stock as those collected in the Georges River (Dr Mike Dove, NSW DPI, personal communication). The hatchery controls were 5th generation progeny of the wild oysters used as the founder population of the B2 selective breeding line. However, these HC oysters have never been subjected to selective breeding.

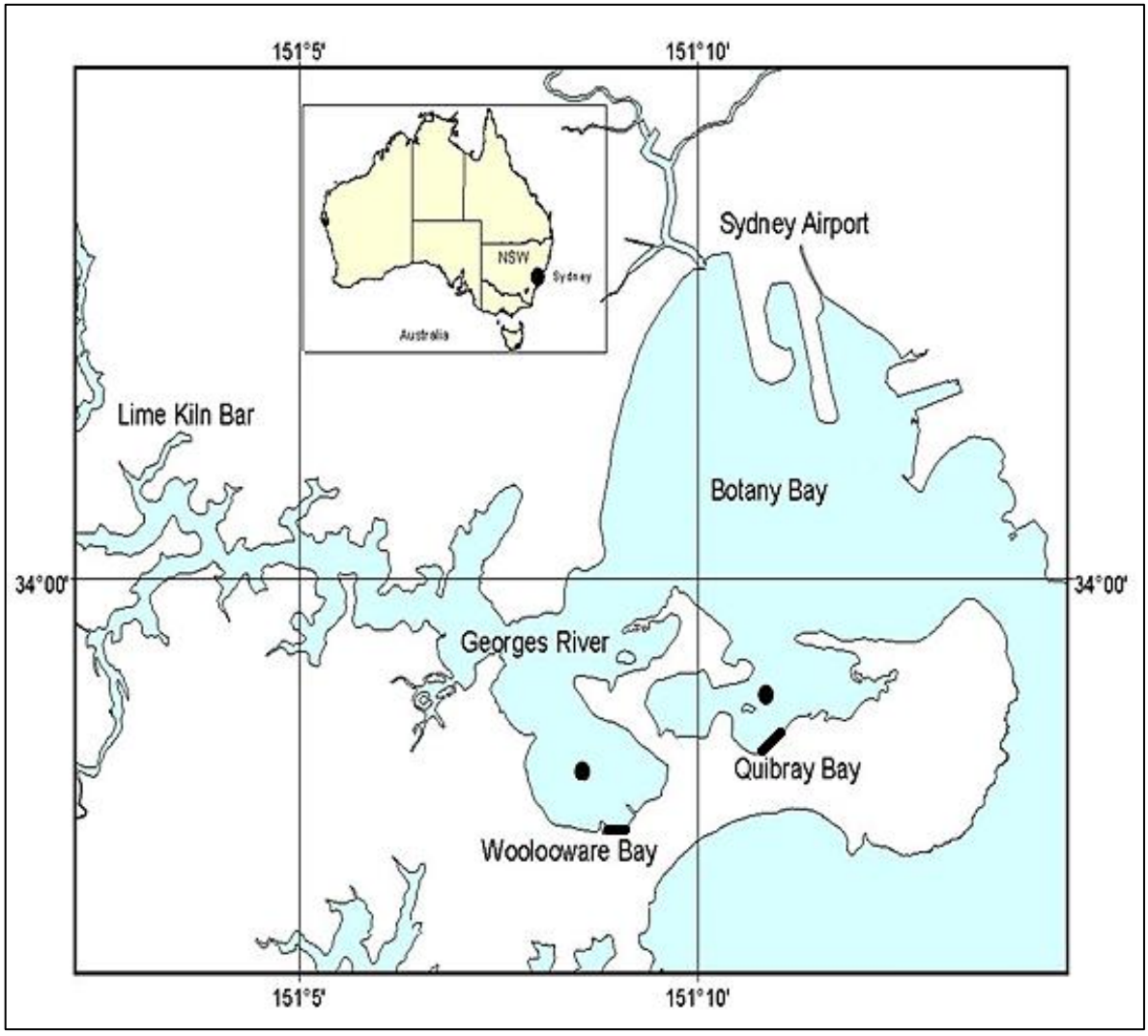


Fig. 1. Map of the two sampling sites in the Georges River estuary, New South Wales: Quibray Bay and Woolooware Bay. Circles indicate the location of oyster leases within each bay, where B2 and overcatch oysters were collected. Dark lines indicate corresponding areas of shoreline where wild-caught oysters were collected (Map courtesy of NSW DPI).

Tissue samples from all oysters (n=93) were collected in February 2015. The oysters were shucked and internal morphological features were used to confirm that they were Sydney rock oysters, rather than Pacific oysters (*C. gigas*). All of the oysters sampled had denticles indicative of *S. glomerata* around the lip of their shell (Thomson 1954). After shucking, 5mm x 15mm portions of gill tissue were extracted, immediately placed in 95% ethanol and stored at 4^{o}C. The tissues were further dissected into 4mm x 4mm portions that were stored in 70% ethanol for transport and subsequent molecular analyses.

Next-generation nucleotide sequencing, bioinformatics and genotyping

Tissue subsections from all 93 oysters were processed for sequencing and genotyping at Diversity Arrays Technology (DArT) Pty. Ltd. (Canberra, Australia). DNA extractions, sample preparation and sequencing followed the DArTSeq™ protocol to identify single nucleotide polymorphisms (SNPs). The DArTSeq™ protocol is a genotyping-by-sequencing approach using DArT markers (Jaccoud *et al.* 2001; Luikart *et al.* 2003) and Illumina sequencing platforms (Sansaloni *et al.* 2011). It facilitates the identification and genotyping of thousands of SNPs that are evenly distributed throughout the genome of the target species (Petroli *et al.* 2012). Detailed descriptions of DArT marker technologies for SNP discovery and genotyping are provided in Jaccoud *et al.* (2001) and Sansaloni *et al.* (2011).

In brief, DNA was extracted using GenCatch™ Blood and Tissue Genomic Mini Prep Kits (Epoch Biolabs) in accordance with the manufacturer's instructions. Purified DNA was held in a 1x solution of MultiCore™ restriction enzyme buffer (Promega). Aliquots of all DNA samples were electrophoresed on 0.8% agarose gels pre-stained with GelRed™ (Biotium Inc.; Huang *et al.* 2010) to confirm that they contained high molecular weight DNA. Each DNA sample (100 ng) was then digested using two restriction enzymes (*PstI* and *SphI*) and ligated to adapters specific to these enzymes. The *PstI* adapter contained an Illumina flow cell attachment sequence, a sequencing primer and a barcode sequence unique for each individual sample.

Samples were then purified with PCR clean up kits (Qiagen) and amplified by PCR with primers specific to both the adapter and barcode sequences. The PCR conditions included 1 minute initial denaturation at 94 C°, then 30 cycles of denaturation (20 sec, 94 °C), annealing (30 sec, 58 °C) and extension (45 sec, 72 °C), with a final extension of 7 min at 72 °C. Following PCR, all samples were pooled together in equal molar quantities. The samples were then diluted and denatured with NaOH prior to hybridisation to the flow cell. An Illumina HiSeq2500 single read platform was then used to sequence the library. This process used 77 cycles and produced reads of equal length (65 bp). To ensure reproducibility, technical replicates were created by carrying approximately 30 - 40% of the samples through a second run of the library preparation protocol and through subsequent downstream analyses (Donnellan *et al.* 2015).

Illumina HiSeq2500 software was used to convert the resulting raw sequence data to .fastq files. Reads from individual oysters were de-multiplexed using the individual-specific barcode sequence ligated to the samples. All reads were then subjected to quality control. This involved checking for contaminants using GenBank viral and bacterial sequences and an in-house DArT database (Diversity Arrays Technology Pty. Ltd., Canberra). Any reads with PHRED (Ewing *et al.* 1998) Q-scores < 25 were also removed. The remaining reads were assessed using the DArT proprietary pipeline DArTSoft14^{TM} (Diversity Arrays Technology, Canberra) to identify and call SNPs. This pipeline is comparable to the STACKS pipeline (Catchen *et al.* 2013). However DArTSoft14^{TM} calls the sequence clusters for all the pooled samples first, followed by the calling of the sequence clusters for each individual. Sequence clusters found to be monomorphic were removed. SNPs were then identified and filtered to further ensure quality. Any SNPs that did not occur in both homozygous and heterozygous forms were removed, along with SNPs with a read depth < 5 and reproducibility < 95%. Any locus with very high read depth was also removed so that only SNPs with an average ratio of read depth between alleles of 0.75 were retained.

The final SNP dataset provided by DArT was further filtered for missing data at both the individual and locus level. As a result, only individuals and loci with ≥ 95% available data were retained. The data were screened for allele coverage, with any SNPs displaying a read depth < 10 removed from the dataset (Lemay and Russello 2015). SNP's were also filtered for minor allele frequencies < 5%, as low frequency SNP's are known to create biases when analysing genetic signatures of selection (Roesti *et al.* 2012). Lastly, in cases where multiple polymorphisms were found within the same sequence length, only one SNP was retained and all other duplicates were removed in order to avoid bias due to physical linkage (Lemay and Russello 2015). There is no reason to suspect systematic bias as a result of these filtering steps, which are commonplace in population genetics analyses. The final filtered dataset will be made available online prior to publication.

After filtering, the dataset was converted into the genotype coding system appropriate for use in the Microsoft Excel add-in package GenAlEx version 6.5 (Peakall and Smouse 2006; Peakall and Smouse 2012). GenAlEx then was used for preliminary data exploration and for export of the data into the Arlequin project (.arp) format. All other required file types were created from .arp files using the file conversion program PGDSpider version 2.0.8.3 (Lischer and Excoffier 2012).

Detecting loci under selection

Fst outlier tests were used to detect loci that were potentially under selection. Although the exact methods for these tests vary between programs, the common process involves identifying loci that have Fst values outside of the range expected to occur through genetic drift. Low Fst values indicate balancing or negative selection, while high Fst values indicate disruptive or positive selection (Lewontin and Krakauer 1973). In the current study, we undertook Fst outlier tests in three independent programs.

Potential Fst outliers were first detected using FDIST tests in Lositan (Antao *et al.* 2008) implementing the methods of Beaumont and Nichols (1996). The options for 'force mean Fst' and 'neutral mean Fst' were selected, as recommended by Antao *et al.* (2008), and the false discovery rate was set at 0.05. All other settings were left at default (mutation model=infinite alleles; number of simulations = 50,000; confidence intervals = 95%). Lositan is simulation-based, so two independent sets of simulations were run in order to generate a consensus to identify markers putatively under selection.

In addition to Lositan, potential Fst outliers were detected using the Bayesian simulation method of Beaumont and Balding (2004) in BayeScan version 2.1 (Foll and Gaggiotti 2008). All the default parameters were used (sample size = 5000; thinning interval = 10; pilot runs = 20; pilot run length = 5000; additional burn in = 50,000; prior odds for neutral model = 10). 'SNP genotypes matrix data' was also selected. Outputs were analysed with all default parameters and false discovery rate of 0.05 (Foll and Gaggiotti 2008; Foll 2012), using the plot_bayescan function in R (R Development Core Team 2011).

Potential Fst outliers were also detected using the coalescent method of hierarchically structured populations in Arlequin version 3.5.2.2 (Excoffier and Lischer 2010), using a variation of the method of Beaumont and Nichols (1996). Default parameters were used (20,000 coalescent simulations, 10 groups and 100 demes). Outliers were identified using a threshold p-value of < 0.05 and the type of selection (positive vs negative) was determined based on Fst values. Loci with negative Fst values were deemed to be under negative selection, while loci with large, positive Fst values were deemed to be under positive selection.

A consensus list of outliers under putative selection was constructed using results from all three Fst outlier tests described above. Outliers that were identified as under selection across all three Fst outlier tests were removed from the original data to produce a dataset comprising only loci incorporating putatively selectively-neutral SNPs. This edited dataset of 1182 loci containing neutral SNPs was used for all subsequent analysis, unless stated otherwise. Please note that creating a more conservative list of neutral loci by removing all loci identified as under selection across all three outlier tests made no difference to the outcome of the results.

Population structure and introgression

Summary statistics for each sample group, as well as totals for the pooled wild and B2 groups, were calculated in GENEPOP version 4.3 (Rousset 2008). These included observed and expected heterozygosities and F-statistics calculated by the methods of Weir and Cockerham (1984), as well as deviations from Hardy-Weinberg equilibrium (HWE) determined by the exact tests method by Guo and Thompson (1992) using 1000 permutations. The loci were also screened for fixed loci or private alleles in the different sample groups. Where necessary, the significance of differences in allele frequencies between oyster populations was determined by Chi-squared analysis. Genetic diversity across sample groups was then assessed by principal components analysis (PCA) in the Adegenet package of R (Jombart 2008) using the multivariate method described by Jombart *et al.* (2010).

A discriminant analysis of principal components (DAPC) was also undertaken in Adegenet to determine the likely number of genetically distinct populations present within the dataset. This discriminant analysis included all oysters without any prior grouping information. The function find.clusters was used to analyse the data and determine *K*, the number of genetically distinct groups present within the dataset (Jombart 2013). A new dataset was then created based on the population groupings identified in the DAPC. Summary statistics and frequencies of private alleles were then re-calculated across the re-grouped dataset. In addition, individual multilocus heterozygosities were calculated for both the neutral and positive selection datasets. The resulting heterozygosity frequencies were compared across the newly identified genetic groups using a two-sample Kolmogorov-Smirnov test.

Results

SNP databases

Diversity Array Technology (DArT) genotyped 93 individuals to produce a dataset of 15,250 SNPs (average read depth 41.8). Further filtration steps narrowed this dataset to a final 1200 SNPs (average read depth 67.1).

Two independent simulations using Lositan identified a consensus of 88 loci under balancing selection and 64 loci under positive selection. In contrast, BayeScan detected 18 loci under positive selection and no loci under balancing selection, whilst Arlequin identified 122 loci under balancing selection and 122 loci under positive selection. The 18 loci identified as putatively under positive selection by BayeScan were the only loci identified as under selection by all three forms of outlier analysis. Hence, two final datasets were assembled. One contained the 18 loci under positive selection (mean Fst = 0.314) and the second comprised the remaining 1182 selectively neutral loci (mean Fst = 0.0452). Thirteen of the 18 outlier loci under positive selection could be identified based on their sequence similarities to known genes (see Supplementary Material, Table S1).

Population structure and introgression

Summary statistics for the eight sample groups are shown in Table 1. For these sample groups, expected heterozygosities (He) ranged from 0.2685 to 0.2850, while observed heterozygosities (Ho) ranged from 0.1952 to 0.2657 and Fis values ranged from 0.0677 to 0.2731. QWC had the lowest Ho (0.1952) and He (0.2685), as well as the highest Fis (0.2731). QB2 had the highest Ho (0.2657) and He (0.2850), as well as the lowest Fis (0.0677). Four of the eight sample groups (QOC, QWC, WWC and HC) deviated significantly ($p < 0.05$) from Hardy-Weinberg Equilibrium (HWE). Total Ho and He values for the pooled B2 and wild groups were similar, with Ho values of 0.2325 and 0.2199 and He values of 0.2865 and 0.2818, respectively. Fis values for the pooled B2 and wild groups were 0.1884 and 0.2195, respectively. Both of these pooled groups deviated significantly

393 from HWE (<0.001). Overall, there were 61 instances of fixed alleles across the eight
394 sample groups. These fixed alleles fell in 26 loci (24 neutral, 2 under positive selection),
395 with some alleles fixed in more than one sample group. No private alleles were found within
396 any single sample group.

397

398 **Table 1.** Summary statistics for each of the eight sample groups of oysters, including mean
399 observed (Ho) and expected (He) heterozygosities, Fis values and p-values testing the null
400 hypothesis that there was no deviation from Hardy–Weinberg equilibrium (HWE).

401

Sample Group	Group Code	N	Ho	He	Fis	HWE
Hatchery B2	HB2	11	0.2331	0.2826	0.1750	0.3837
Hatchery control	HC	10	0.2047	0.2769	0.2606	0.0008
Quibray Bay B2	QB2	9	0.2657	0.2850	0.0677	1
Quibray Bay overcatch	QOC	12	0.2053	0.2786	0.2633	< 0.001
Quibray Bay wild-caught	QWC	12	0.1952	0.2685	0.2731	< 0.001
Woolooware Bay B2	WB2	12	0.2390	0.2742	0.1283	0.5949
Woolooware Bay overcatch	WOC	12	0.2435	0.2786	0.1259	0.9968
Woolooware Bay wild-caught	WWC	12	0.2212	0.2750	0.1957	0.0182
Total B2	B2	32	0.2325	0.2865	0.1884	< 0.001
Total wild (overcatch, wild-caught and hatchery controls)	Wild	58	0.2199	0.2818	0.2195	< 0.001

402

403 Principal component analysis (PCA) of the selectively-neutral dataset identified two main
404 clusters of oysters (Fig. 2). One cluster comprised only oysters from the three B2 sample
405 groups (HB2, QB2 and WB2), whilst the other cluster incorporated all of the wild (QOC,
406 QWC, WOC and WWC) and hatchery control (HC) oysters. There was substantial overlap
407 between the sample groups within each cluster. There was no overlap between the two
408 clusters, except for three individuals from QB2 (oysters 28, 29 and 31) that fell into the
409 wild/hatchery control cluster. The B2 cluster exhibited greater principle component
410 separation (diversity) between individuals when compared to the cluster comprising the wild
411 and hatchery control oysters.

412

Discriminant analysis of principal components (DAPC) also identified two distinct clusters of oysters (K = 2), based on allele frequencies in the selectively-neutral dataset (Fig. 3 A-C). The two clusters comprised 61 and 32 oysters respectively. These numbers corresponded to the sample groups containing all wild/hatchery control oysters (n = 58) and the sample groups containing the B2 oysters (n = 35), with the exception of the same three QB2 individuals that fell into the wild/hatchery control cluster in the PCA. These three individuals were deemed to be outliers likely derived from inadvertent mixing of populations during farming (see Discussion), and so were removed from subsequent analyses. Comparable results were returned in STRUCTURE (K=2), however these results are not shown here as our dataset did not meet all of the assumptions for this program.

Comparison of the two DAPC clusters (wild/hatchery controls and B2) identified 112 private alleles. Of these, 77 were found exclusively in the wild/hatchery control cluster (76 neutral loci, 1 locus under positive selection) and 35 were found exclusively in the B2 group (30 neutral loci, 5 loci under positive selection). This represented a significant difference in the frequencies of private alleles between the two clusters (X^2 = 8.0046; p = 0.0047). Three loci (all selectively-neutral) were fixed across all B2 individuals, whilst none were fixed across all wild/hatchery control individuals.

Significant differences in the frequencies of individual multilocus heterozygosities between the B2 and wild/hatchery control clusters were evident for both the neutral loci and the loci under positive selection (two-sample Kolmogorov-Smirnov tests; both $p < 0.001$). Mean individual multilocus heterozygosity across the neutral loci was higher in the B2 group (0.2439 ± 0.0513; mean ± SD) than in the wild/hatchery control cluster (0.2155 ± 0.0447; Fig. 4A). Loci under positive selection also had higher mean individual multilocus heterozygosity in the B2 cluster (0.2387 ± 0.1350) relative to the wild/hatchery control cluster (0.1048 ± 0.0724; Fig. 4B).

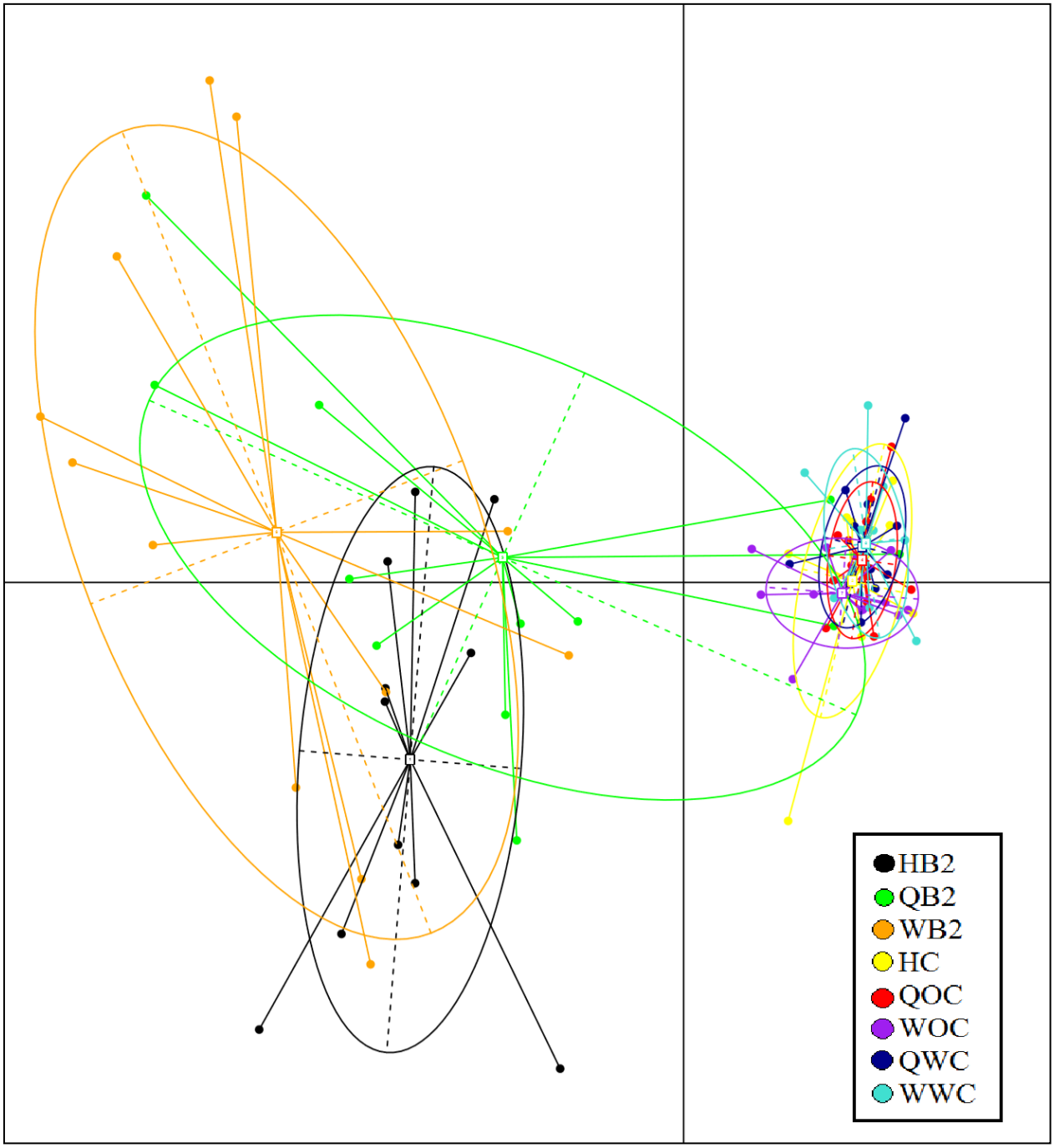
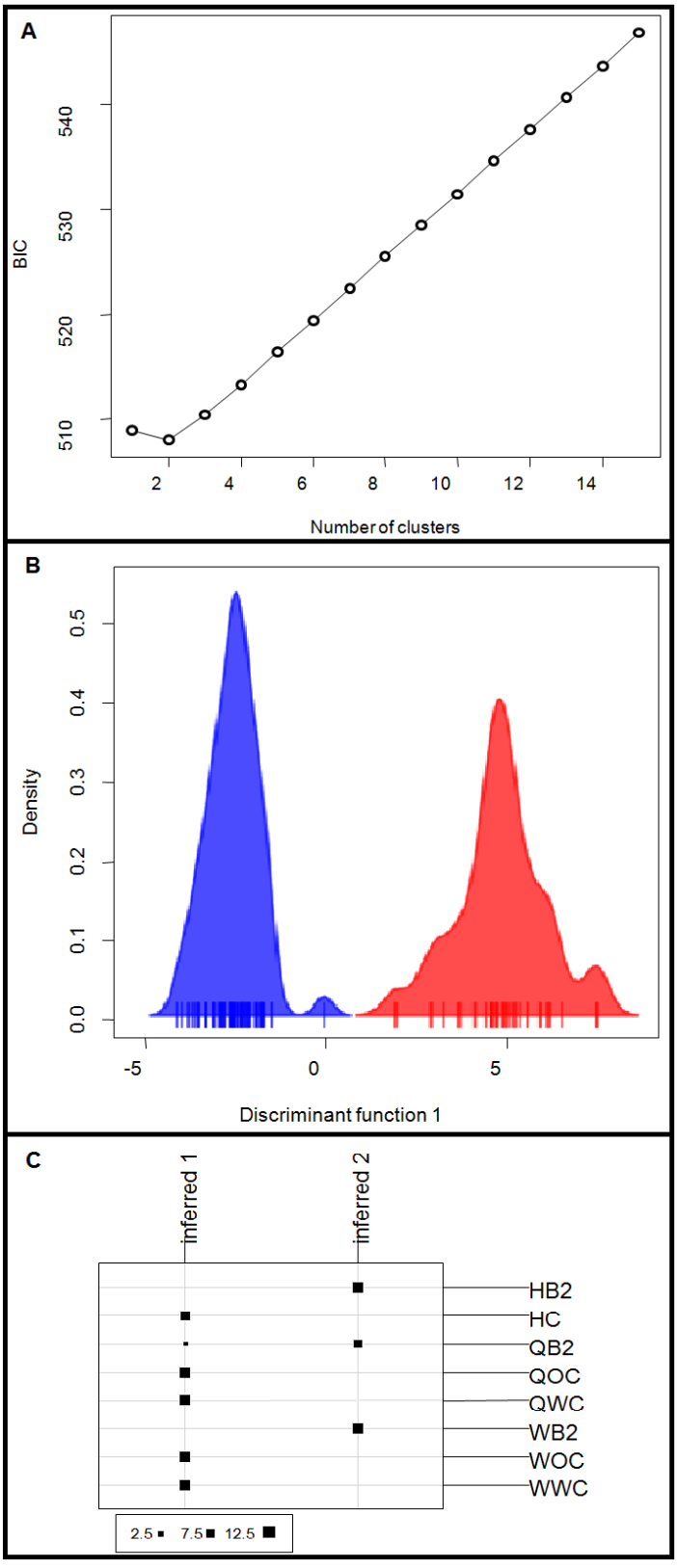


Fig. 2. Principal component analysis (PCA) of variation between individual oysters (n=93) based on 1182 selectively-neutral SNPs. Each point represents an individual oyster. Ovals represent 95% confidence intervals for each sample group (see Table 1 for a description of sample group codes). Principal component analysis of the full dataset (1182 selectively neutral loci + 18 loci under positive selection) yielded directly comparable results to this analysis of the neutral dataset (data not shown).



447

448 **Fig. 3.** Genetic clusters based on SNP data as determined by discriminant analysis of
449 principal components (DAPC). A: BIC values for possible numbers of clusters (K) ranging
450 from 1 to 15; B: density of the values of the first discriminant function across the two
451 inferred clusters; C: allocation of individuals from each sample group into each of the
452 inferred clusters (see Table 1 for a description of sample group codes).

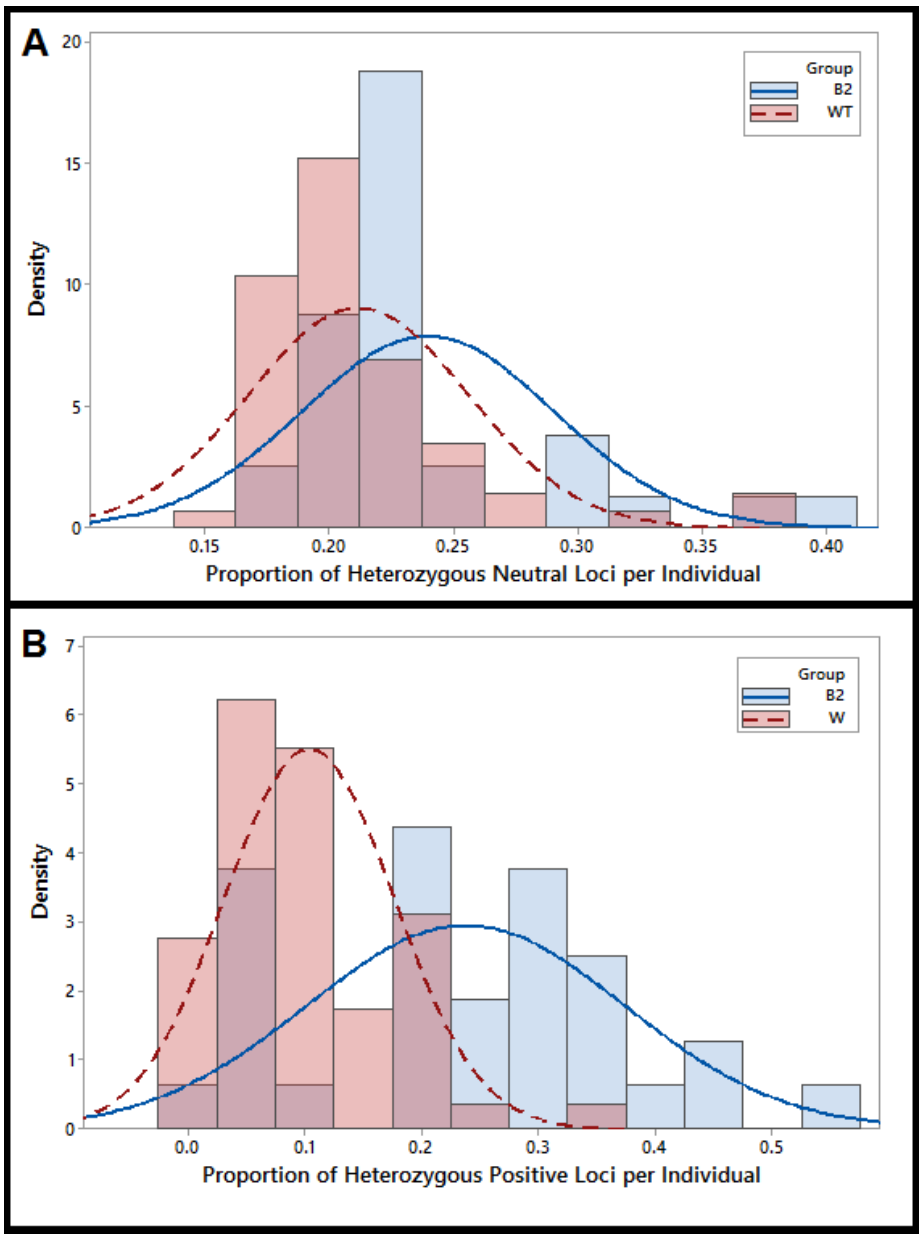


Fig. 4. Individual multilocus heterozygosities for selectively-bred (B2) and wild/hatchery control (W) oysters. Data are presented for A) neutral loci, and B) loci identified as under positive selection.

Discussion

The current study has shown that the selectively-bred B2 line of Sydney rock oysters is genetically distinct from wild Sydney rock oysters, and there is no evidence of introgression occurring between farmed B2 oysters and the wild population growing in the Georges River. Both discriminant analysis of principle components (DAPC) and principle components analysis (PCA) identified a clear distinction between the wild and farmed B2

sample groups, which formed two discrete clusters. The large number of private alleles (112) identified between the B2 oysters and the wild groups is also indicative of limited gene flow between the two clusters. All of these data suggest that stable genetic introgression between the farmed B2 oysters and the wild oyster population has not occurred in the Georges River. The only discrepancies in the segregation of the B2 and wild population were three oysters collected from the B2 farming racks at Quibray Bay (sample group QB2). These oysters appeared to have wild, rather than B2 genotypes. There are several explanations for the occurrence of these oysters (see below). However, none of these explanations are consistent with gene flow from the farmed B2 oysters into the wild population because the gene flow implied by the presence of these oysters is in the opposite direction (wild into B2).

There are a number of explanations for the lack of introgression seen in this study. Wild and selectively-bred Sydney rock oyster gametes may have low compatibility and/or hybrid spat may have low survival rates in comparison to wild spat. The selectively-bred line has been reared in a hatchery environment and so may have developed characteristics that are not suited to living in the wild. The lack of interbreeding between wild and selectively-bred Sydney rock oysters might also be the result of farming practices. Farmed oysters are generally sold once they reach a defined size class in order to return the most profit. Sydney rock oysters are often sold immediately after reaching sexual maturity, when they have become "fat" due to their swollen (gravid) reproductive organs. Therefore, it is likely that the majority of farmed oysters are sold before they have had the opportunity to reproduce within the estuary. Another reason for the lack of introgression may be the relative size of oyster populations in the Georges River. Anecdotal evidence (Wayne O'Connor, NSW DPI, personal communication) suggests that the number of wild oysters growing in the Georges River may be orders of magnitude greater than the farmed B2 population. In this case, the number of gametes from wild oysters may simply overwhelm those from the farmed population

It is also possible that introgression has occurred in the Georges River, but was not detected in this study. Larval dispersal patterns of Sydney rock oysters are poorly understood (Banks *et al.* 2006). It is not known how long larvae spend in the water column or how far they travel. The circulation patterns of the Georges River estuary are also not well documented,

so the direction that larvae travel and exchange patterns between bays remain unknown. Although our study accounted for potential larval dispersal by choosing two bays geographically separated by approximately 10 km of coastline, it is possible that larvae produced in these bays may have settled further afield. More extensive studies of Sydney rock oysters within other bays in this estuary will help to clarify whether the effects of introgression can be detected elsewhere, particularly if hybrid larvae are travelling further from the farming areas than expected.

It is important to note that the current study focussed on detecting sustained levels of introgression, where gene flow between populations leads to hybridisation and further backcrossing (Anderson and Hubricht 1938). Our experimental design, which is based on relatively low sample sizes from each population (n = 9-12 for each sample group at each sampling site) does not preclude the possibility of low level gene flow between the farmed and wild oyster populations. Such gene flow has been observed in other oyster species. For instance, Carlsson *et al.* (2008) found an extremely low level of hybridisation (0.6%) between wild and selectively-bred eastern oysters (*C. virginica*) in Chesapeake Bay, USA. This was despite the release of approximately 18.5 million selectively-bred spat within this region for population restoration projects. The study also reported no significant increase in the frequency of haplotypes known to be common in the selectively-bred oysters (35.9%) but rare in the wild oysters (1.2%; Carlsson *et al.* 2008). Similar or even lower levels of hybridisation could be occurring between the farmed B2 oysters and the wild population in the Georges River. However, in the incidence of such low levels of hybridisation, the sheer number of wild oysters in the Georges River combined with the relatively small number of samples collected in this study would have made it unlikely for us to detect such hybrids.

In addition to assessing introgression, the current study also allowed detailed comparisons of the genetic variation in the B2 and wild Sydney rock oysters. This provided further evidence for the substantial genetic differentiation between the two populations, resulting from selective breeding and hatchery rearing. These differences were best visualised by PCA, which revealed that the variability of neutral markers is far lower in wild oysters than in B2 oysters. Differential variability was also evident in the individual multilocus heterozygosities of B2 and wild oysters across both the selectively-neutral dataset and the loci under positive selection. In both cases, the B2 oysters were found to have significantly

higher levels of individual multilocus heterozygosity ($p < 0.001$). Such differences in individual multilocus heterozygosity between populations were contrary to expectations. The B2 oysters were expected to have lower individual multilocus heterozygosity than the wild oysters, due to both the selection pressures of the breeding program and the higher chance of inbreeding as a consequence of small parental population size.

Even though it is counterintuitive, the lower variability of wild oysters is in line with the findings of English *et al.* (2001). They reported that the second and third generations of *S. glomerata* breeding lines selected for fast growth had higher genetic variation (in terms of observed and expected heterozygosities) than oysters from their founder population. The oysters assessed by English *et al.* (2001) were the progenitors of the current *S. glomerata* breeding lines, including the B2 line used in the current study. Similarly, proteomic analyses by Simonian *et al.* (2009) found that the proteomes of wild Sydney rock oysters differed substantially among a range of selective breeding lines and that substantial heterogeneity had been maintained in these selective breeding lines, even for proteins that appeared to be under selection (Simonian *et al.* 2009). All of these findings contradict the literature for other oyster species. Studies for other oysters (*P. maxima*, Lind *et al.* 2009; *C. gigas*, Hedgecock and Sly 1990; *C. virginica*, Carlsson *et al.* 2006) have consistently found lower genetic diversity in hatchery-produced and/or selectively-bred oysters than in wild populations, after as little as one generation. This indicates that the genetic structure of wild Sydney rock oyster populations may differ substantially from other broadcast spawning bivalves.

We observed lower observed heterozygosity than expected under HWE for some groups of Sydney Rock Oysters. Comparison of observed versus expected heterozygosities identified significant departures from HWE in all of the wild oyster groups except WOC, resulting in high Fis values. In contrast, none of the B2 groups deviated significantly from HWE when assessed independently. However, the pooled totals for the wild and B2 groups both deviated significantly from HWE. The deviation from HWE evident among wild oysters could be a result of flow dynamics in the Georges River acting to disrupt gene flow and/or disease pressures. Mortality rates resulting from QX disease and winter mortality in the Georges River can exceed 90% per year among non-selected oysters (Peters and Raftos 2003). These levels of mortality may be sufficient to generate an identifiable genetic

bottleneck. Disease may also explain the significantly lower than expected heterozygosities of the hatchery control oyster group. The founding populations for the hatchery controls were collected from the Georges River, along with other rivers in northern NSW that are subject to severe annual outbreaks of QX disease (Nell 2006). There is also evidence for comparable bottlenecks in other oyster species. For instance, Hedgecock and Sly (1990) identified a heterozygote excess in the first generations of two hatchery-produced Pacific oysters (*C. gigas*), whilst there was a heterozygote deficit in the wild founder population. B2 oysters have been selected for resistance to QX and winter mortality and so suffer far lower levels of mortality (Nell and Perkins 2006). Hence, the B2 oysters may not have been subjected to a comparable bottleneck as their wild conspecifics.

Another explanation for the low levels of observed heterozygosities in the wild oysters is low effective population size (Ne). Contrary to expectations for broadcast spawners, oysters are known to have relatively low Ne. For instance, Pacific oysters (*C. gigas*) have a ratio of Ne/N of less than 10^{-6} (Hedgecock *et al.* 1992). These low effective population sizes are thought to be due to unequal sex ratios and limited availability of spawners, as well as substantial differences in viability between families (Lind *et al.* 2009). In *C. gigas*, spat produced by mass spawning events have Ne values lower than the number of broodstock used and lower than the number of parents that were known to have contributed to the offspring (Li and Yu 2009). Additionally, oyster larvae generally have low levels of survival in the wild. This leads to relatively few survivors within a generation, often comprising the spawn of relatively few parents. These factors are also evident in the hatchery environment, where the differential survival of families can lead to the overrepresentation of some families within a generation. Lind *et al.* (2009) analysed the progeny of a mass spawning event of silver-lipped pearl oysters (*P. maxima*) and found parental contributions from 28 individuals. However, approximately 40% of the offspring were identified as full siblings. Low Ne values are also expected when successive generations are under conditions of strong selection, particularly so if these pressures are relatively recent. In the case of the Georges River, disease represents one such selection pressure (Nell 2001).

Further evidence for the occurrence of a population bottleneck in the Georges River can be found by comparing current levels of heterozygosities with that evident in other locations in NSW. For example, wild Sydney rock oysters in Jervis Bay, NSW, were reported to have

expected heterozygosities ranging from 0.608 to 0.936 and observed heterozygosities ranging from 0.571 to 0.885, across nine microsatellite loci (Banks *et al.* 2006). Additionally, none of the nine microsatellite loci showed any significant deviation from Hardy-Weinberg Equilibrium (HWE, $p > 0.1$). Jervis Bay is an area in southern NSW that has not experienced QX disease or winter mortality (Nell 2006). In contrast, Buroker *et al.* (1979) found levels of heterozygosity (Ho =0.184, He = 0.172) that are similar to those reported in the current study in Sydney rock oysters collected from the Georges River four decades ago. Such persistently low levels of heterozygosities in the Georges River could be attributed to disease pressures, since outbreaks of winter mortality have been known to occur in the area since 1920. However, Buroker *et al.* (1979) identified a heterozygote excess in the Georges River, while the current study identified a heterozygote deficit. This may be indicative of a more recent bottleneck related to QX disease, which first appeared in the Georges River in 1994 (Nell 2006).

The substantial genetic differentiation observed between the wild and selectively-bred (B2) populations may have been facilitated by the high mutation rates evident in oysters. Oysters are known to have exceptionally high levels of DNA polymorphisms, within the order of magnitude of the highest known for any animal (Sauvage *et al.* 2007). Sauvage *et al.* (2007) found that the average density of SNPs in the Pacific oyster (*C. gigas*) as one in 60bp in coding regions and one in 40bp in non-coding regions. This is tentatively corroborated by our dataset for Sydney rock oysters. A number of loci had to be removed from our analysis during the filtration stages due to the proximity of multiple SNPs within the reads. Individual 64bp sequence reads contained up to 9 unique SNPs. These high mutation rates could explain the relatively large number of private alleles identified in the current study, when comparing the wild and selectively-bred oyster groups.

In contrast to wild oysters, the substantial genetic diversity evident within the B2 line might explain the only anomaly in our data. Both DAPC and PCA identified three oysters from B2 farming racks at Quibray Bay that appeared to have wild genotypes. The genetic diversity of B2 oysters is such that some individuals may display similar genotypes to those of wild oysters by chance, accounting for these three unusual oysters. However, a more pragmatic explanation is that these three oysters were wild-caught oysters that were mistakenly placed in the B2 farming racks. Spat of both B2 and wild-caught oysters are produced in the Port

630 Stephens hatchery and subsequently farmed in the Georges River, where racks containing
631 each type of oyster are often processed side by side. While these farming practices provide
632 the opportunity for accidental mixing of the two types of oysters, it is more likely that cross-
633 contamination or hybridisation would occur in the hatchery where spat are much smaller
634 and harder to see (Wayne O'Connor, NSW DPI, personal communication).

635

636 In addition to providing detailed information on the population genetics of Sydney rock
637 oysters, the current study also identified loci under positive selection in selectively-bred
638 oysters. Extensive analysis of the loci under positive selection was beyond the scope of the
639 current study. However, it will be the focus of ongoing investigation. Such information
640 could greatly improve selective breeding programs through the implementation of marker-
641 assisted selection of the type being investigated for the silver-lipped pearl oyster (Jones *et*
642 *al.* 2013). Further research into genetic aspects of *S. glomerata*, beyond those explored here,
643 may also provide insights into the underlying processes responsible for the current findings.
644 Factors such as effective population size and more detailed analysis of population
645 bottlenecks should be investigated. Additional studies could assess multiple generations of
646 selectively-bred B2 oysters and wild oysters to determine whether the observed differences
647 between these populations are sustained across generations.

648

649 ## Acknowledgements

650

651 The author wishes to thank David Raftos, Adam Stow, Melanie Bishop, Paolo Momigliano
652 and Elayna Truszewski for assistance with experimental design and analyses; David Raftos,
653 Emma Thompson, Maria Vozzo, Michael Dove, Wayne O'Connor, NSW Department of
654 Primary Industries and Peter David for assistance with collection of samples; Daniel Powell
655 and Abigail Elizur for assistance with characterizing loci under selection; David Raftos,
656 Adam Stow, Liam O'Hare and Peterson Coates for comments on a draft of this manuscript.
657 This research was funded by Macquarie University and the Australian Research Council.

References

Aladaileh, S., Nair, S. V., and Raftos, D. A. (2008). Effects of noradrenaline on immunological activity in Sydney rock oysters. *Developmental & Comparative Immunology* **32**, 627-636.

Anderson, E., and Hubricht, L. (1938). Hybridization in Tradescantia. III. The Evidence for Introgressive Hybridization. *American Journal of Botany* **25**, 396-402.

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

Banks, S., Piggott, M., Raftos, D., and Beheregaray, L. (2006). Microsatellite markers for the Sydney rock oyster, *Saccostrea glomerata*, a commercially important bivalve in southeastern Australia. *Molecular Ecology Notes* **6**, 856-858.

Beaumont, M. A., and Balding, D. J. (2004). Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology* **13**, 969-980.

Beaumont, M. A., and Nichols, R. A. (1996). Evaluating Loci for Use in the Genetic Analysis of Population Structure. *Proceedings: Biological Sciences* **263**, 1619-1626.

Beck, M. W., Brumbaugh, R. D., Airoldi, L., Carranza, A., Coen, L. D., Crawford, C., Defeo, O., Edgar, G. J., Hancock, B., and Kay, M. C. (2011). Oyster reefs at risk and recommendations for conservation, restoration, and management. *Bioscience* **61**, 107-116.

Bishop, M. J., Krassoi, F. R., McPherson, R. G., Brown, K. R., Summerhayes, S. A., Wilkie, E. M., and O'Connor, W. A. (2010). Change in wild-oyster assemblages of Port Stephens, NSW, Australia, since commencement of non-native Pacific oyster (*Crassostrea gigas*) aquaculture. *Marine & Freshwater Research* **61**, 714-723.

Buroker, N. E., Hershberger, W., and Chew, K. (1979). Population genetics of the family Ostreidae. I. Intraspecific studies of *Crassostrea gigas* and *Saccostrea commercialis*. *Marine Biology* **54**, 157-169.

Butt, D., Shaddick, K., and Raftos, D. (2006). The effect of low salinity on phenoloxidase activity in the Sydney rock oyster, *Saccostrea glomerata*. *Aquaculture* **251**, 159-166.

Carlsson, J., Carnegie, R. B., Cordes, J. F., Hare, M. P., Leggett, A. T., and Reece, K. S. (2008). Evaluating recruitment contribution of a selectively bred aquaculture line of the oyster, *Crassostrea virginica* used in restoration efforts. *Journal of Shellfish Research* **27**, 1117-1124.

Carlsson, J., Morrison, C. L., and Reece, K. S. (2006). Wild and aquaculture populations of the eastern oyster compared using microsatellites. *Journal of Heredity* **97**, 595-598.

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

Coen, L. D., Brumbaugh, R. D., Bushek, D., Grizzle, R., Luckenbach, M. W., Posey, M. H., Powers, S. P., and Tolley, S. (2007). Ecosystem services related to oyster restoration. *Marine Ecology Progress Series* **341**, 303-307.

Diggles, B. K. (2013). Historical epidemiology indicates water quality decline drives loss of oyster (*Saccostrea glomerata*) reefs in Moreton Bay, Australia. *New Zealand Journal of Marine and Freshwater Research* **47**, 561-581.

Donnellan, S. C., Foster, R., Junge, C., Huveneers, C., Rogers, P., Kilian, A., and Bertozzi, T. (2015). Fiddling with the proof: the Magpie Fiddler Ray is a colour pattern variant of the common Southern Fiddler Ray (Rhinobatidae: Trygonorrhina). *Zootaxa* **3981**, 367-384.

English, L., Nell, J., Maguire, G., and Ward, R. (2001). Allozyme variation in three generations of selection for whole weight in Sydney rock oysters (*Saccostrea glomerata*). *Aquaculture* **193**, 213-225.

Ewing, B., Hillier, L., Wendl, M. C., and Green, P. (1998). Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. *Genome research* **8**, 175-185.

Excoffier, L., and Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**, 564-567.

Foll, M. (2012). Bayescan v2. 1 user manual. *Ecology* **20**, 1450-1462.

Foll, M., and Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* **180**, 977-993.

Gaffney, P. M. (2006). The role of genetics in shellfish restoration. *Aquatic Living Resources* **19**, 277-282.

Grabowski, J. H., Peterson, C. H., Cuddington, K., Byers, J., Wilson, W., and Hastings, A. (2007). Restoring oyster reefs to recover ecosystem services. In Ecosystem engineers: plants to protists. pp. 281-298. (Elsevier Science: The Netherlands)

Green, T. J., Dixon, T. J., Devic, E., Adlard, R. D., and Barnes, A. C. (2009). Differential expression of genes encoding anti-oxidant enzymes in Sydney rock oysters, *Saccostrea glomerata* (Gould) selected for disease resistance. *Fish & Shellfish Immunology* **26**, 799-810.

Grossen, C., Keller, L., Biebach, I., The International Goat Genome Consortium, and Croll, D. (2014). Introgression from Domestic Goat Generated Variation at the Major Histocompatibility Complex of Alpine Ibex. *PLoS Genetics* **10**, e1004438.

Guo, S. W., and Thompson, E. A. (1992). Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* **48**, 361-372.

Hedgecock, D., Chow, V., and Waples, R. S. (1992). Effective population numbers of shellfish broodstocks estimated from temporal variance in allelic frequencies. *Aquaculture* **108**, 215-232.

Hedgecock, D., and Sly, F. (1990). Genetic drift and effective population sizes of hatchery-propagated stocks of the Pacific oyster, *Crassostrea gigas*. *Aquaculture* **88**, 21-38.

Heino, M., Svåsand, T., Wennevik, V., and Glover, K. A. (2015). Genetic introgression of farmed salmon in native populations: quantifying the relative influence of population size and frequency of escapees. *Aquaculture Environment Interactions* **6**, 185-190.

Huang, Q., Baum, L., and Fu, W.-L. (2010). Simple and Practical Staining of DNA with GelRed. *Clinical Laboratory* **20**, 149-152.

Hughes, A. R., Inouye, B. D., Johnson, M. T., Underwood, N., and Vellend, M. (2008). Ecological consequences of genetic diversity. *Ecology Letters* **11**, 609-623.

Jaccoud, D., Peng, K., Feinstein, D., and Kilian, A. (2001). Diversity arrays: a solid state technology for sequence information independent genotyping. *Nucleic Acids Research* **29**, e25.

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

Jombart, T. (2013). 'A tutorial for Discriminant Analysis of Principal Components (DAPC) using adegenet 1.4-0.' Available at https://github.com/thibautjombart/adegenet/wiki [Accessed 14 August 2015].

Jombart, T., Devillard, S., and Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* **11**, 94.

Jones, D. B., Jerry, D. R., Khatkar, M. S., Raadsma, H. W., and Zenger, K. R. (2013). A high-density SNP genetic linkage map for the silver-lipped pearl oyster, *Pinctada maxima*: a valuable resource for gene localisation and marker-assisted selection. *BMC Genomics* **14**, 810.

Kingsley-Smith, P. R., Harwell, H. D., Kellogg, M. L., Allen, S. M., Allen Jr, S. K., Meritt, D. W., Paynter Jr, K. T., and Luckenbach, M. W. (2009). Survival and growth of triploid *Crassostrea virginica* (Gmelin, 1791) and *C. ariakensis* (Fujita, 1913) in bottom environments of Chesapeake Bay: implications for an introduction. *Journal of Shellfish Research* **28**, 169-184.

Krassoi, F. R., Brown, K. R., Bishop, M. J., Kelaher, B. P., and Summerhayes, S. (2008). Condition-specific competition allows coexistence of competitively superior exotic oysters with native oysters. *Journal of Animal Ecology* **77**, 5-15.

Lemay, M. A., and Russello, M. A. (2015). Genetic evidence for ecological divergence in kokanee salmon. *Molecular Ecology* **24**, 798-811.

Lenihan, H. S., Micheli, F., Shelton, S. W., and Peterson, C. H. (1999). The influence of multiple environmental stressors on susceptibility to parasites: an experimental determination with oysters. *Limnology and Oceanography* **44**, 910-924.

Lewontin, R., and Krakauer, J. (1973). Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics* **74**, 175-195.

Li, R., and Yu, R. (2009). Parentage determination and effective population size estimation in mass spawning Pacific oyster, Crassostrea gigas, based on microsatellite analysis. *Journal of the World Aquaculture Society* **40**, 667-677.

Lind, C. E., Evans, B. S., Knauer, J., Taylor, J. J., and Jerry, D. R. (2009). Decreased genetic diversity and a reduced effective population size in cultured silver-lipped pearl oysters (*Pinctada maxima*). *Aquaculture* **286**, 12-19.

Lischer, H., and Excoffier, L. (2012). PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* **28**, 298-299.

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

Nell, J. A. (1993). Farming the Sydney rock oyster (*Saccostrea commercialis*) in Australia. *Reviews in Fisheries Science* **1**, 97-120.

Nell, J. A. (2001). The History of Oyster Farming in Australia. *Marine Fisheries Review* **63**, 14-25.

Nell, J. A. (2006). Manual for mass selection of Sydney rock oysters for fast growth and disease resistance. NSW Department of Primary Industries, Fisheries Research Report Series No. 13, New South Wales.

Nell, J. A., and Hand, R. E. (2003). Evaluation of the progeny of second-generation Sydney rock oyster *Saccostrea glomerata* (Gould, 1850) breeding lines for resistance to QX disease *Marteilia sydneyi*. *Aquaculture* **228**, 27-35.

Nell, J. A., and Perkins, B. (2005). Evaluation of progeny of fourth generation Sydney rock oyster *Saccostrea glomerata* (Gould, 1850) breeding lines. *Aquaculture Research* **36**, 753-757.

Nell, J. A., and Perkins, B. (2006). Evaluation of the progeny of third-generation Sydney rock oyster *Saccostrea glomerata* (Gould, 1850) breeding lines for resistance to QX disease *Marteilia sydneyi* and winter mortality *Bonamia roughleyi*. *Aquaculture Research* **37**, 693-700.

Nell, J. A., Smith, I. R., and McPhee, C. C. (2000). The Sydney rock oyster *Saccostrea glomerata* (Gould 1850) breeding programme: progress and goals. *Aquaculture Research* **31**, 45-49.

Newell, R. I. (2004). Ecosystem influences of natural and cultivated populations of suspension-feeding bivalve molluscs: a review. *Journal of Shellfish Research* **23**, 51-62.

O'Connor, W. A., and Dove, M. C. (2009). The changing face of oyster culture in New South Wales, Australia. *Journal of Shellfish Research* **28**, 803-811.

Parker, L. M., Ross, P. M., and O'Connor, W. A. (2009). The effect of ocean acidification and temperature on the fertilization and embryonic development of the Sydney rock oyster *Saccostrea glomerata* (Gould 1850). *Global Change Biology* **15**, 2123-2136.

Parker, L. M., Ross, P. M., O'Connor, W. A., Borysko, L., Raftos, D. A., and Pörtner, H.-O. (2012). Adult exposure influences offspring response to ocean acidification in oysters. *Global Change Biology* **18**, 82-92.

Parker, L. M., Ross, P. M., and O'Connor, W. A. (2011). Populations of the Sydney rock oyster, *Saccostrea glomerata*, vary in response to ocean acidification. *Marine Biology* **158**, 689-697.

Peakall, R., and Smouse, P. E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**, 288-295.

Peakall, R., and Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* **28**, 2537-2539.

Peters, R., and Raftos, D. A. (2003). The role of phenoloxidase suppression in QX disease outbreaks among Sydney rock oysters (*Saccostrea glomerata*). *Aquaculture* **223**, 29-39.

Petroli, C. D., Sansaloni, C. P., Carling, J., Steane, D. A., Vaillancourt, R. E., Myburg, A. A., da Silva Jr, O. B., Pappas Jr, G. J., Kilian, A., and Grattapaglia, D. (2012). Genomic

characterization of DArT markers based on high-density linkage analysis and physical mapping to the Eucalyptus genome. *PLoS One* **7**, e44684.

R Development Core Team (2011). 'R: a Language and Environment for Statistical Computing.' (R Foundation for Statistical Computing: Vienna.)

Roesti, M., Salzburger, W., and Berner, D. (2012). Uninformative polymorphisms bias genome scans for signatures of selection. *BMC Evolutionary Biology* **12**, 94-100.

Rousset, F. (2008). GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* **8**, 103-106.

Sansaloni, C., Petroli, C., Jaccoud, D., Carling, J., Detering, F., Grattapaglia, D., and Kilian, A. (2011). Diversity Arrays Technology (DArT) and next-generation sequencing combined: genome-wide, high throughput, highly informative genotyping for molecular breeding of Eucalyptus. *BMC Proceedings* **5**, 54-55.

Sauvage, C., Bierne, N., Lapegue, S., and Boudry, P. (2007). Single nucleotide polymorphisms and their relationship to codon usage bias in the Pacific oyster *Crassostrea gigas*. *Gene* **406**, 13-22.

Schrobback, P., Pascoe, S., and Coglan, L. (2014). Impacts of introduced aquaculture species on markets for native marine aquaculture products: the case of edible oysters in Australia. *Aquaculture Economics & Management* **18**, 248-272.

Seamons, T. R., Hauser, L., Naish, K. A., and Quinn, T. P. (2012). Can interbreeding of wild and artificially propagated animals be prevented by using broodstock selected for a divergent life history? *Evolutionary Applications* **5**, 705-719.

Simonian, M., Nair, S. V., Nell, J. A., and Raftos, D. A. (2009). Proteomic clues to the identification of QX disease-resistance biomarkers in selectively bred Sydney rock oysters, *Saccostrea glomerata*. *Journal of Proteomics* **73**, 209-217.

Stewart, C. N., Halfhill, M. D., and Warwick, S. I. (2003). Transgene introgression from genetically modified crops to their wild relatives. *Nature Reviews Genetics* **4**, 806-817.

Thomson, J. (1954). The genera of oysters and the Australian species. *Marine & Freshwater Research* **5**, 132-168.

Utter, F., and Epifanio, J. (2002). Marine aquaculture: genetic potentialities and pitfalls. *Reviews in Fish Biology and Fisheries* **12**, 59-77.

Warwick, S. I., Beckie, H. J., and Hall, L. M. (2009). Gene flow, invasiveness, and ecological impact of genetically modified crops. *Annals of the New York Academy of Sciences* **1168**, 72-99.

928
929 Weir, B. S., and Cockerham, C. C. (1984). Estimating F-statistics for the analysis of
930 population structure. *Evolution*, 1358-1370.

Supplementary Material

Table S1. Characterisation of loci under positive selction. Short read sequences for loci containing SNPs under positive selection were searched against a custom database of *S. glomerata* genomic DNA sequence contigs. *Local e-value* and *Greatest % identify* refer to these comparions between short reads containing SNPs and the larger contigs. Sequences for the matching contigs were then Blasted against NCBI's non-redundant nucleotide sequence database. *E-value*, Abbrviated *description* and *Function group* refer to the results from those Blast searches. Locus ID numbers, the position of the SNP within the short reads and the Fst values for the individual loci are also shown.

Locus ID	SNP position	Fst	Local e-value	Greatest identity %	e-value	Abbreviated description	Function group
200	20: G>A	0.17654	$3.13e^{-26}$	98.55	$3.9e^{-3}$	vwfa and cache domain-containing protein 1-like	Calcium homeostasis
214	38: A>T	0.19457	$3.13e^{-26}$	98.55	$6e^{-74}$	Spondin-1 like isoform [*C .gigas*]	Extracellular matrix protein
406	61: T>G	0.22739	$7.35e^{-28}$	100	$2.5e^{-150}$	Serine threonine-protein kinase nim1	Intracellular signal transduction
424	65: A>T	0.17356	$7.35e^{-28}$	100	-	-	-
457	24: C>G	0.18596	$7.35e^{-28}$	100	-	-	-
618	9: T>C	0.19556	$8.96e^{-27}$	100	0.24	Hypothetical protein CGI_10000944	Unknown
666	34: C>T	0.16616	$7.35e^{-28}$	100	$1.2e^{-24}$	Protein FLP	Unknown

Locus ID	SNP position	Fst	Local e-value	Greatest identity %	e-value	Abbreviated description	Function group
721	22: G>A	0.17809	$2.74e^{-14}$	91.89	0.011	Probable secreted beta-glucosidase ADG3 isoform x2	Carbohydrate metabolism
836	25: C>T	0.14316	$5.65e^{-23}$	95.71	$6e^{-53}$	B-box type zinc finger protein NCL-1-like [*C. gigas*]	Transcription factor
858	17: C>A	0.17908	$3.13e^{-26}$	98.55	$6.1e^{-32}$	Latent-transforming growth factor beta-binding protein 4	Intracellular signal transduction
899	63: G>A	0.19195	$7.35e^{-28}$	100	$1.4e^{-21}$	5-AMP-activated protein kinase subunit gamma-1-like isoform x7	Intracellular signal transduction
942	27: T>A	0.24146	$3.13e^{-26}$	98.55	$2e^{-11}$	Calcium homeostasis endoplasmic reticulum isoform x2	Calcium homeostasis
1060	15: A>C	0.19922	$7.35e^{-28}$	100	-	-	-
1097	15: T>C	0.17169	-	-	-	-	-
1119	27: A>C	0.20475	$7.35e^{-28}$	100	$1.7e^{-5}$	Hypothetical protein CGI_10011078	Unknown
1126	16: G>A	0.17951	$3.13e^{-26}$	98.55	$4.3e^{-101}$	Pogo transposable element with ZNF domain	Cell cycle
1174	8: A>T	0.19553	$3.13e^{-26}$	98.55	-	-	-
1194	16: T>C	0.16211	$7.35e^{-28}$	100	$3.1e^{-4}$	SICA antigen partial	Unknown

Appendix:

Instructions for Authors: Marine & Freshwater Research

Presentation

The work should be presented in clear and concise English. All text should be in Times New Roman, 12 point font, with double or 1.5-line spacing throughout, and with a margin of at least 3 cm on the left-hand side. **Every line of each page must be consecutively numbered in the left-hand margin, starting from 1 to the highest numbers needed as this greatly assists the referees.** All pages of the manuscript must be numbered consecutively, including those carrying references, tables and captions to illustrations, all of which are to be placed after the text. Follow the form of headings, tables and illustrations exemplified in recent issues of the Journal.

Format

Papers should usually be in the form Title, Abstract, Additional keywords, Introduction, Materials and methods, Results, Discussion, Acknowledgements and References. If authors choose to combine the Results and Discussion sections, they must also include a Conclusion to summarise their key findings.

The title should be concise and appropriately informative and should contain all keywords necessary to facilitate retrieval by online search engines. The abstract (< 200 words) should **open with a clear statement of the broad relevance of the work**, briefly summarise the aims and research approach, give the principal findings, and conclude by specifying the main implications of the results to aquatic science. Additional keywords not already in the title or abstract should be listed beneath the abstract. A running head (< 50 letter spaces) should be supplied for use at the top of the printed page.

The Introduction should set the global relevance of the work in the opening sentences. Text should only cover essential background literature and clearly indicate the reason for the work. This section should close with a paragraph specifying aims and, where appropriate, testable hypotheses.

In the Materials and methods, sufficient detail should be given to enable the work to be repeated. If a commercial product such as an analytical instrument is mentioned, supply its full model name and location of the manufacturer. Give complete citations and version numbers for computer software. Data analysis must be explained clearly, especially when complex models or novel statistical procedures are used.

Results should be stated concisely and without interpretation (although in complex studies, modest interpretation of some data may provide context helpful for understanding subsequent sections). Data presented should address aims and testable hypotheses raised in the Introduction. Use tables and figures to illustrate the key points but do not repeat their contents in detail.

The Discussion should explain the scientific significance of the results in context with the literature, clearly distinguishing factual results from speculation and interpretation. Avoid excessive use of references - more than three to support a claim is usually unnecessary. Limitations of methods should also be addressed where appropriate. Conclude the Discussion with a section on the implications of the findings. Footnotes should be used only when essential.

Acknowledgments, including funding information, should appear in a brief statement at the end of the body of the text.

References

We use the Harvard system. Cite references chronologically in the text by the author and date. Multiple references from the same year should be cited alphabetically. In the text, the names of two coauthors are linked by ´and´; for three or more, the first author´s name is followed by ´*et al.*´. In the reference list, include the full author list, article title and journal name (i.e. no abbreviations). Authors must ensure that they have permission to cite material as a personal communication and can provide unpublished data if required by a reviewer.

Examples:

Journal article
Prince, J. D., Sellers, T. L., Ford, W. B., and Talbot, S. R. (1988). Confirmation of a relationship between localised abundance of breeding stock and recruitment for *Haliotis rubra* Leach (Mollusca: Gastropoda). *Journal of Experimental Marine Biology and Ecology* **122**, 91-104.

Book chapter
Wolanski, E., Mazda, Y., and Ridd, P. (1992). Mangrove hydrodynamics. In ´Tropical Mangrove Ecosystems´. (Eds A. I. Robertson and D. M. Alongi.) pp. 43-62. (American Geophysical Union: Washington, DC.)

Book
Sokal, R. R., and Rohlf, F. J. (1981). ´Biometry. The Principles and Practice of Statistics in Biological Research.´ 2nd Edn. (W. H. Freeman: New York.)
Attiwill, P. M., and Adams, M. A. (Eds) (1996). ´Nutrition of Eucalypts.´ (CSIRO Publishing: Melbourne.)

Thesis
Silver, M. W. (1970). An experimental approach to the taxonomy of the genus *Enteromorpha* (L.). PhD Thesis, University of Liverpool.
Harrison, A. J. (1961). Annual reproductive cycles in the Tasmanian scallop *Notovola meridionalis*. BSc (Hons) Thesis, University of Tasmania, Hobart.

Report or Bulletin
Chippendale, G. M., and Wolf, L. (1981). The natural distribution of *Eucalyptus* in Australia. Australian National Parks and Wildlife Service, Special Publication No. 6, Canberra.

Conference Proceedings
Kawasu, T., Doi, K., Ohta, T., Shinohara, Y., and Ito, K. (1990). Transformation of eucalypts (*Eucalyptus saligna*) using electroporation. In ´Proceedings of the VIIth International Congress on Plant Tissue and Cell Culture, Florence, 12-17 June 1994´. pp. 64-68. (Amsterdam IAPTC: Amsterdam.)

Web-based material
Goudet, J. (2001). ´FSTAT, a Program to Estimate and Test Gene Diversities and Fixation Indices (Version 2.9.3).´ Available at http://www2.unil.ch/popgen/softwares/fstat.htm [accessed 15 November 2007].

Tables and figures

Tables must be numbered with Arabic numerals and have a self-explanatory title. A headnote containing material relevant to the whole table should start on a new line, as it will be set in a different font. Tables should be arranged with regard to the dimensions of the printed page (17.5 by 23 cm) and the number of columns kept to a minimum. Excessive subdivision of column headings is undesirable; use abbreviations that can then be expanded upon in the headnote. The first letter only of headings to rows and columns should be capitalized. The symbol for the unit of measurement should be placed in brackets beneath the column heading. Footnotes should be kept to a minimum and be reserved for specific items in columns. Horizontal rules should be inserted only above and below column headings and at the foot of the table. Vertical rules must not be used.

All figures must be referred to in the text (e.g., Fig. 1, Fig. 2a-d, Figs 1 and 2), and should be numbered consecutively in the order that they are cited within the paper. Electronic submission of figures is required. Photographs and line drawings should be of the highest quality and, if not created digitally, should be scanned at high-resolution: photographs at 300 dpi at final size, saved as .jpg files; hand-drawn line drawings at least 600 dpi at final size, saved as .tif files. Black-and-white photographs should be saved in greyscale format as .tif or Photoshop files. Labels must be applied electronically to the scanned images in Photoshop,

rather than scanning manually labelled figures. Colour figures and photographs must be submitted in CMYK format for printing purposes, not in RGB. Photographs and images must be of the highest quality, and trimmed squarely to exclude irrelevant features. When in a group, adjacent photographs must be separated by uniform spaces that will be 2 mm wide after reduction. A scale bar is desirable on micrographs and photographs lacking reference points. Important features to which attention has been drawn in the text should be indicated.

The lettering of figures must be in sans-serif type (Helvetica is ideal) with only the first letter of the first word of any proper names capitalised, and should not be in bold type. For letter size, the height of a lower-case ´x´ after reduction should be approximately 1.2 mm. Do not use the symbols ´+´ or ´×´ for data points. Grid marks should point inwards and legends to axes should state the quantity being measured and be followed by the appropriate units in parentheses. Thickness of lines on line diagrams at final size must be no less than 0.5 pt. Grouped figures should not exceed 17.5 cm by 23 cm. Colour graphics will be accepted, but the cost of production is borne by the author.

Supplementary material

In an effort to make best use of printed journal space, **Marine & Freshwater Research** strongly encourages authors to place supporting files such as additional tables, figures and raw data in ´Supplementary Material´, which is linked online to the paper when it is published electronically. Such material is not crucial to the paper's interpretation but would bolster claims, illustrate specific aspects of interest, or expand on a point in the text. There is no special format for Supplementary Material and it should be cited in the main text as ´..available as Supplementary Material...´ or ´(see Supplementary Material)´.

Data deposition

DNA sequences published in the Journal should be deposited in one of the following nucleotide sequence databases: EMBL, GenBank or DDBJ. An accession number for each sequence or sequence set must be included in the manuscript before publication. In addition, electronic copies of the data sets in nexus format should be supplied with the manuscript to aid the review process.