MOLECULAR ECOLOGY

Molecular Ecology (2016) **25**, 5093–5113 doi: 10.1111/mec.13830

Riverscape genomics of a threatened fish across a hydroclimatically heterogeneous river basin

CHRIS J. BRAUER,* MICHAEL P. HAMMER† and LUCIANO B. BEHEREGARAY*

**Molecular Ecology Laboratory, School of Biological Sciences, Flinders University, Adelaide, SA 5042, Australia, †Natural Sciences, Museum and Art Gallery of the Northern Territory, Darwin, NT 0801, Australia*

Abstract

Understanding how natural selection generates and maintains adaptive genetic diversity in heterogeneous environments is key to predicting the evolutionary response of populations to rapid environmental change. Detecting selection in complex spatial environments remains challenging, especially for threatened species where the effects of strong genetic drift may overwhelm signatures of selection. We carried out a basin-wide riverscape genomic analysis in the threatened southern pygmy perch (*Nannoperca australis*), an ecological specialist with low dispersal potential. High-resolution environmental data and 5162 high-quality filtered SNPs were used to clarify spatial population structure and to assess footprints of selection associated with a steep hydroclimatic gradient and with human disturbance across the naturally and anthropogenically fragmented Murray–Darling Basin (Australia). Our approach included F_{ST} outlier tests to define neutral loci, and a combination of spatially explicit genotype–environment association analyses to identify candidate adaptive loci while controlling for the effects of landscape structure and shared population history. We found low levels of genetic diversity and strong neutral population structure consistent with expectations based on spatial stream hierarchy and life history. In contrast, variables related to precipitation and temperature appeared as the most important environmental surrogates for putatively adaptive genetic variation at both regional and local scales. Human disturbance also influenced the variation in candidate loci for adaptation, but only at a local scale. Our study contributes to understanding of adaptive evolution along naturally and anthropogenically fragmented ecosystems. It also offers a tangible example of the potential contributions of landscape genomics for informing in situ and ex situ conservation management of biodiversity.

Keywords: Australia, climate change, conservation genetics, ddRAD-seq, landscape genomics, Murray–Darling Basin, *Nannoperca australis*

Received 15 June 2016; revision received 15 August 2016; accepted 23 August 2016

Introduction

The effects of human development and recent climate change on our natural environment are pervasive, and the threat these selective forces pose to global biodiversity is increasing (Vitousek *et al.* 1997; Walther *et al.* 2002; Thuiller 2007). Populations faced with environmental change can respond through range shifts, acclimation through phenotypic plasticity or by genetic evolutionary adaptation to their new environment (Bellard *et al.* 2012; Pauls *et al.* 2013). If one or some combination of these processes does not occur, populations and potentially entire species face an increased risk of extinction (Quintero & Wiens 2013). Range shifts in response to climate change have already been observed for some species (Davis & Shaw 2001; Cahill *et al.* 2012), but this can be problematic for those where the opportunity for dispersal is naturally limited or constrained by recent habitat fragmentation (Dawson *et al.* 2011). Phenotypic responses to environmental change have also been observed (Charmantier *et al.* 2008; Hendry

Correspondence: Luciano B. Beheregaray, Fax: 61(8) 82013015; E-mail: luciano.beheregaray@flinders.edu.au

et al. 2008), but there are costs and limits associated with plasticity (DeWitt *et al.* 1998) that make it unlikely to provide long-term solutions for many populations (Gienapp *et al.* 2008). This leads to several questions concerning the capacity of species to persist in situ and to adapt to altered environmental conditions. The currently rapid rate of climate and environmental change suggests that evolutionary adaptation will need to rely heavily on standing genetic variation (Barrett & Schluter 2008). However, for threatened species, it is unclear whether enough variation exists at adaptively important loci to facilitate an evolutionary response. Thus, it is important to ask the following questions: 'How are threatened populations locally adapted, what are the important environmental factors contributing to local adaptation, and how is adaptive genetic variation spatially distributed and maintained?'

Landscape genomics (LG) provides an ideal framework for addressing questions in ecology and evolution, which have become particularly relevant in threatened biotas that are both naturally and anthropogenically fragmented. This rapidly growing research field combines information about environmental heterogeneity and genomewide data of individuals sampled across the landscape to identify the spatial patterns of neutral and adaptive variation (Manel *et al.* 2010; Sork & Waits 2010). Although there has been recent debate over the practical application of genomics to conservation (Garner *et al.* 2015; Shafer *et al.* 2015a,b), LG has increasingly been applied to threatened and nonmodel species (Cooke *et al.* 2012a,b, 2014; Limborg *et al.* 2012; Bourret *et al.* 2013; Moore *et al.* 2014; Steane *et al.* 2014; Hand *et al.* 2015; Hecht *et al.* 2015; Laporte *et al.* 2015; Funk *et al.* 2016). In the context of conservation, the large number of markers generated by next-generation sequencing (NGS) can improve the resolution of demographic inferences (Luikart *et al.* 2003; Allendorf *et al.* 2010), but it is the identification of genomic regions under selection that has held great promise for increasing our understanding of the potential vulnerability or resilience of biodiversity to environmental change (Allendorf *et al.* 2010; Frankham 2010). There are, however, several characteristics common to many threatened species, such as small effective population sizes, population isolation, repeated local extinction–recolonization cycles and inbreeding, which may affect our ability to distinguish signals of selection from other confounding effects. These factors therefore need to be considered when selecting an analytical framework for LG studies of natural populations (Schoville *et al.* 2012).

The most common methods used to detect selection in LG studies are based on population genetics theory and the assumption that demographic processes such as migration and genetic drift should affect the genome uniformly, while selection should act on specific regions or loci (Lewontin & Krakauer 1973). Known as F_{ST} outlier tests, they have become a standard feature of most LG studies. These tests, however, assume specific demographic models and may not be robust to violations imposed by nonequilibrium demographic scenarios (Lotterhos & Whitlock 2014; Whitlock & Lotterhos 2015). Nevertheless, aside from considering outliers as candidates for selection, these tests offer an effective solution for creating a large neutral data set for improving inferences about population structure and demographic history (Luikart *et al.* 2003; Allendorf *et al.* 2010).

Genotype–environment association (GEA) approaches are an alternative strategy for detecting the signal of local adaptation by testing for direct associations between allele frequencies and environmental parameters. These methods are generally free from the constraints of simple demographic models. They can be used to test specific hypotheses related to environmental heterogeneity, including the possibility that it shapes polygenic adaptation in natural populations (Lasky *et al.* 2012; Bourret *et al.* 2014; Hecht *et al.* 2015). Moreover, GEA approaches generally also incorporate means to account for the effects of shared population history and can separate geographic and environmental effects (Joost *et al.* 2007; Coop *et al.* 2010; Frichot *et al.* 2013; Guillot *et al.* 2014; Rellstab *et al.* 2015). This is particularly important for complex spatial environments such as dendritic river networks where physical landscape structure can greatly affect the patterns of genetic variation (Hughes *et al.* 2009; Fourcade *et al.* 2013; Thomaz *et al.* 2016).

Here, we use a framework that capitalizes on a high-resolution environmental data set and on powerful LG approaches to assess footprints of selection in a threatened species found across a hydroclimatically heterogeneous and anthropogenically modified ecosystem. Our study system, the southern pygmy perch *Nannoperca australis*, is a small-bodied freshwater fish (<85 mm) endemic to southeastern Australia, including the Murray–Darling Basin (MDB) (Unmack *et al.* 2013). This ecological specialist is normally associated with streams and wetlands, sheltered microhabitats and aquatic macrophyte cover, is relatively short-lived (3–6 years; reaches maturity within 1 year), has large demersal eggs and limited dispersal ability (Lintermans 2007; Wedderburn *et al.* 2012). Climate in this region has been increasingly dry, but highly variable throughout the late Holocene (*c.* 3000 years), with conditions characterized by a steep gradient of aridity from east to west with higher, and more consistent rainfall and lower temperatures in the southeast highlands and drier, semi-arid conditions in the western lowlands (Donders *et al.* 2007; Pittock & Finlayson 2011).

Importantly, recent studies of wild populations of *N. australis* demonstrated that hydroclimatic-related factors, and in particular the variation in predictability of flow typical for many Australian rivers (Kennard *et al.* 2010), influence individual fitness and drive the predictive patterns of local adaptation in key reproductive traits and life history strategy (Morrongiello *et al.* 2010, 2013). In addition, female reproductive investment in egg and clutch size in *N. australis* varies predictability among populations along gradients of stream flow (Morrongiello *et al.* 2012). Instead of merely reflecting spatial phenotypic plasticity, these findings also support bet-hedging as a co-evolved adaptive strategy in *N. australis*, a view consistent with increasing theoretical and empirical evidence about the consequences of female investment in the evolution of life histories (Olofsson *et al.* 2009; Morrongiello *et al.* 2012).

Within the MDB, studies based on allozymes, mitochondrial DNA (mtDNA) and microsatellites showed that *N. australis* has very shallow basinwide phylogeographic divergence, but strong contemporary population structure shaped by the hierarchical drainage network (Unmack *et al.* 2013; Cole *et al.* 2016). In fact, coalescent analyses based on microsatellites suggest that isolation and demographic decline observed for some *N. australis* populations is associated with modification and fragmentation of the MDB that postdates the recent European settlement in Australia (Attard *et al.* 2016; Cole *et al.* 2016). Genetic evidence thus indicates that the metapopulation structure of *N. australis* does not reflect deeply historic isolation across its range in the MDB. Similar phylogeographic patterns have also been reported for many other widespread MDB fishes (e.g. Faulks *et al.* 2010a; Unmack *et al.* 2013) and are likely a result of the much greater flow of ancestral MDB rivers during the Pleistocene and the large inland lakes that inundated its lower reaches (Pels 1964).

The overall biogeographic scenario indicates that historically, populations across the MDB were likely larger and more connected despite highly variable natural environmental conditions. This suggests that although the effects of drift have probably recently intensified, the signal of local adaptation is unlikely to have been completely eroded, and that appropriate LG frameworks have the potential to address questions about hydroclimatic adaptation in *N. australis*. Since European settlement, the MDB has suffered from extensive development (e.g. wetland reclamation), river regulation, construction of thousands of barriers to fish passage and the introduction of exotic species (Lintermans 2007; Balcombe *et al.* 2011; Laurance *et al.* 2011). These factors have likely synergistically contributed to the widespread decline of *N. australis* populations and to its current listing as endangered in two Australian states (Hammer *et al.* 2013; Cole *et al.* 2016). An unprecedented severe and prolonged drought between 1997 and 2010 caused catastrophic loss of habitat and local extinction for some *N. australis* populations, particularly in the Lower Murray (Wedderburn *et al.* 2012; Hammer *et al.* 2013). In response to the decline, several conservation breeding and restoration programmes were initiated (Hammer *et al.* 2013; Attard *et al.* 2016) and additional translocations among populations of wild fish have been proposed. Given the ongoing conservation management of *N. australis*, and that climate change is expected to negatively impact its populations even further in future (Perry & Bond 2009; Balcombe *et al.* 2011; Morrongiello *et al.* 2011), it is important to understand how extant populations are adapted to local environmental conditions. More broadly, our study system also provides an opportunity for asking whether recent human-driven selection has impacted the genome of extant populations.

In this study, we test the core hypothesis that the steep hydroclimatic gradient across the MDB has contributed to adaptive genetic divergence of *N. australis* populations. This is based on the premise that the natural flow regime modulates many abiotic and biotic processes (Poff *et al.* 1997), such as habitat connectivity, physical disturbance, resource availability and ecological interactions, which have direct implications for shaping the genetic architecture of widespread aquatic species. We also explore the features of our study system to address three questions that have broad implications to conservation and ecological genomics. First, can LG be used to distinguish signals of selection from other confounding effects (e.g. strong drift) in a threatened, poor dispersive species? Second, can selection due to human disturbance be distinguished from selection due to natural environmental heterogeneity? Third, can GEA approaches detect genomic footprints of polygenic adaptation due to hydroclimatic heterogeneity? To answer these questions, we employ a combination of recently developed spatially explicit GEA approaches within a riverscape genomics framework that integrates environmental and genomewide data sets. These approaches are used to test for associations between population allele frequencies and a suite of environmental variables describing the variation in climate, hydrology and human disturbance while controlling for the effects of landscape structure and shared population history. We also discuss our results in the context of ongoing conservation efforts and the utility of genomics for guiding proactive conservation strategies such as translocations for genetic rescue, and for increasing the adaptive potential of populations in the face of ongoing climate change.

Methods

Sampling

Samples of *Nannoperca australis* were collected from the wild between 2000 and 2013 using netting, box trapping or electrofishing. They were preserved either as frozen specimens or fin clips in 99% ethanol and curated at the South Australian Museum, Adelaide. Initially, 550 individuals were sampled from 38 locations. A smaller sample was then selected to include all known populations in the MDB previously identified with allozyme, mtDNA and microsatellites (Unmack *et al.* 2013; Cole *et al.* 2016) while accommodating for unsatisfactory DNA quality for genomic analysis obtained from some individuals. This resulted in a final, high-quality data set of 263 individuals sampled from 25 locations and encompassing 13 catchments across the entire current MDB distribution of *N. australis* (Fig. 1; Table 1).

Molecular methods and bioinformatics

DNA was extracted using the Qiagen DNeasy Blood and Tissue Kit according to the manufacturer's protocol. DNA integrity was assessed by gel electrophoresis, and purity was measured using a NanoDrop 1000 spectrophotometer (Thermo Scientific).

Double-digest restriction site-associated DNA sequencing libraries were prepared following a protocol modified from Peterson *et al.* (2012). Libraries were multiplexed with 48 samples randomly assigned to each of six Illumina lanes and sequenced on a HiSeq2000 platform as paired-end, 100-bp reads. Raw sequences were demultiplexed using the *process_radtags* component of STACKS v.1.04 (Catchen *et al.* 2011) before de novo assembly of a reference catalogue and genotyping was performed with *dDocent.FB* v.1.2 (Puritz *et al.* 2014). Details about library preparation and bioinformatics are provided in Appendix S1 (Supporting information).

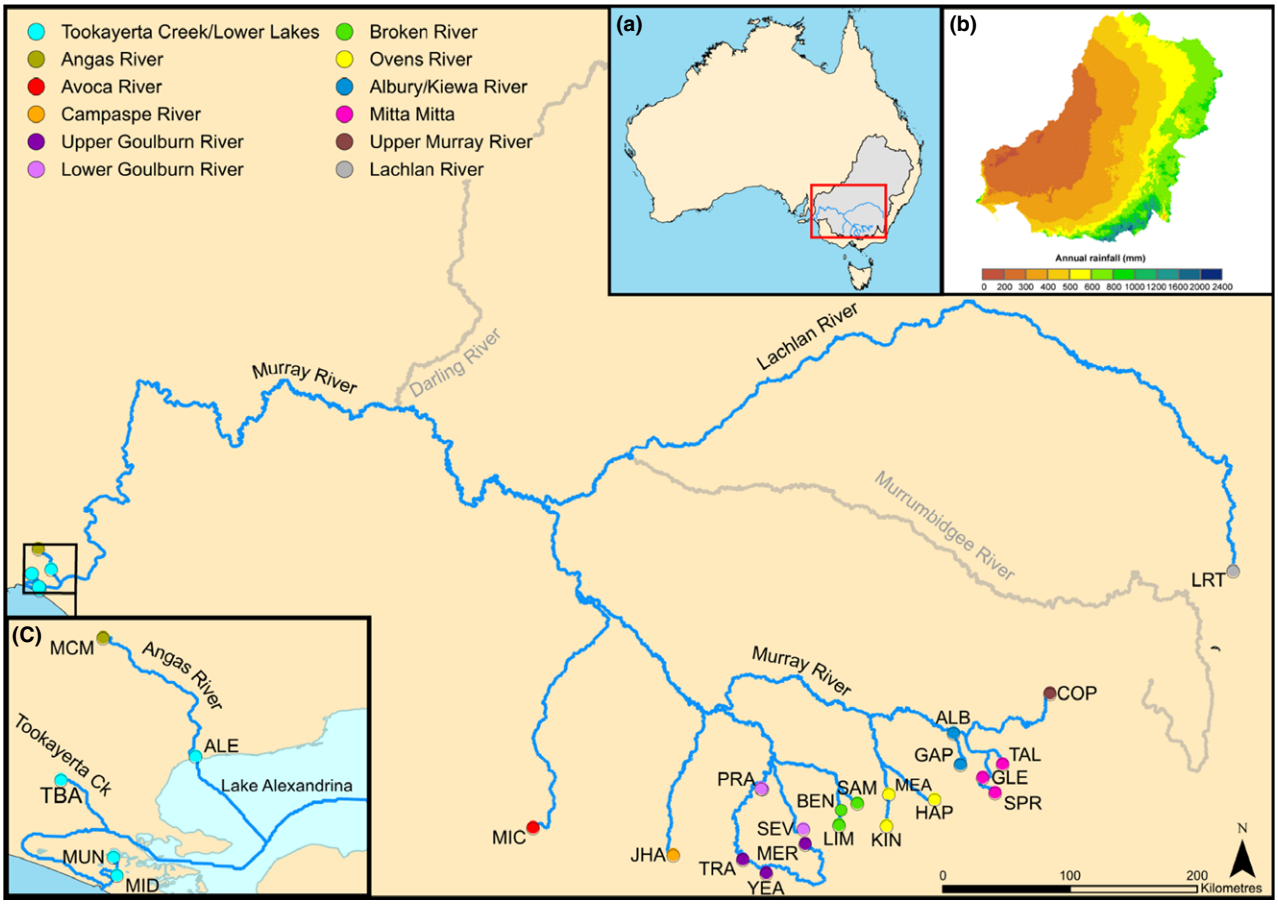

Fig. 1 *Nannoperca australis* sampling locations covering the entire current distribution of the species in the Murray–Darling Basin (MDB). Sites are colour-coded by catchment. Inset (a) shows the location of the MDB (shaded area), and inset (b) depicts the gradient of average annual rainfall across the basin (reproduced from Chiew *et al.* 2008). Inset (c) shows the Lower Murray sampling locations. The historical distribution of the species was essentially continuous within the MDB from the Lower Murray to the upper reaches of the Murray, Murrumbidgee and Lachlan rivers (but excluding the Darling River system), although local abundance likely varied substantially across that range (Llewellyn 1974).

Table 1 Information about localities and sample sizes genotyped for *Nannoperca australis* from the Murray–Darling Basin. Lowland wetland sites referred to as Lower Murray in the text are indicated in boldface

Catchment	Site	Location	*N*	Latitude	Longitude
Tookayerta (TOO)	**TBA**	**Tookayerta Ck, Black Swamp**	**7**	**−35.428**	**138.834**
Lower Lakes (LMR)	**ALE**	**Turvey's Drain, L. Alexandrina**	**10**	**−35.395**	**139.008**
	MID	**Mundoo Is., L. Alexandrina**	**7**	**−35.549**	**138.915**
	MUN	**Drain off Mundoo Channel**	**6**	**−35.520**	**138.904**
Angas (ANG)	MCM	Middle Ck	9	−35.250	138.887
Avoca (AVO)	MIC	Trib to Middle Ck, Warrenmang	11	−37.028	143.338
Campaspe (CAM)	JHA	Jews Harp Ck, Sidonia	12	−37.139	144.578
Upper Goulburn (UGO)	MER	Merton Ck	17	−36.981	145.727
	TRA	Trawool Ck	10	−37.135	145.193
	YEA	Yea R., Yea	8	−37.213	145.414
Lower Goulburn (LGO)	PRA	Pranjip Ck	9	−36.623	145.309
	SEV	Trib to Seven Creeks	11	−36.875	145.701
Broken (BRO)	BEN	Swanpool Ck, Swanpool	10	−36.723	146.022
	SAM	Sam Ck	10	−36.661	146.152
	LIM	Unnamed Ck, Lima South	18	−36.826	146.008
Ovens (OVE)	KIN	King R., Cheshunt	16	−36.795	146.424
	HAP	Happy Valley Ck	9	−36.579	146.824
	MEA	Meadow Ck, Moyhu	8	−36.573	146.423
Kiewa (KIE)	GAP	Gap Ck, Kergunyah	12	−36.317	147.022
Albury (ALB)	ALB	Murray R. lagoon, Albury	12	−36.098	146.928
Mitta Mitta (MIT)	SPR	Spring Ck	10	−36.499	147.349
	GLE	Glencoe Ck	10	−36.393	147.221
	TAL	Tallangatta Ck	7	−36.281	147.382
Upper Murray (COP)	COP	Coppabella Ck	16	−35.746	147.729
Lachlan (LAC)	LRT	Blakney Ck	8	−34.736	149.180

dDocent combines several existing software packages into a single pipeline designed specifically for paired-end RAD data; that is, it takes advantage of both forward and reverse reads for SNP discovery. The resulting variant call file (VCF) was filtered to retain only variants present in at least 70% of individuals and in 70% of populations. Complex variants (multinucleotide polymorphisms and composite insertions and substitutions) were decomposed into SNP and indel representation following Puritz *et al.* (2014), retaining only one biallelic SNP per locus with a minimum minor allele frequency (MAF) of 0.05. A further six filtering steps were performed to remove SNPs likely to be the result of sequencing errors, paralogs, multicopy loci or artefacts of library preparation (Table 2; Appendix S1, Supporting information).

Detecting neutral and outlier loci

Loci not conforming to neutral expectations were detected using a Bayesian approach with BAYESCAN v.2.1 (Foll & Gaggiotti 2008), and the coalescent-based FDIST method (Beaumont & Nichols 1996) in ARLEQUIN v.3.5 (Excoffier & Lischer 2010). BAYESCAN was run for 100 000 iterations using prior odds of 10 000, and loci significantly different from zero and with a q-value < 0.1 [false discovery rate (FDR) of 10%] were considered outliers. ARLEQUIN was run with 50 000 simulations of 13 groups, each with 100 demes, and P-values were corrected for multiple testing using the *p.adjust* function in R (R Core Team 2015). The hierarchical island model was specified (Excoffier *et al.* 2009), as it allows for the assumption of lower migration among catchments than among sampling sites within catchments. Loci significantly outside the neutral distribution at a FDR of 10% were considered as outliers.

The remaining, putatively neutral SNPs were tested for departure from Hardy–Weinberg equilibrium (HWE) in GENODIVE v.2.0b27 (Meirmans & Van Tienderen 2004). Significance was tested using 10 000 random permutations and loci were subsequently removed if found to depart from HWE at a FDR of 10% in more than 50% of sampling locations.

Genetic diversity, N_e and population structure

Expected heterozygosity (H_E) and observed heterozygosity (H_O) were calculated in GENODIVE for both the neutral and the candidate loci. Percentage of polymorphic loci was calculated in GENALEX v.6.5 (Peakall &

Table 2 The number of variant sites retained after each filtering step for *Nannoperca australis* from the Murray–Darling Basin. Number of filtered loci used for downstream analyses indicated in boldface. Detailed descriptions of each filtering step are included in Appendix S1 (Supporting information)

Step	SNP count
Raw SNP catalogue	2 589 251
Genotyped in	
50% of individuals, base quality ≥30, minor allele count of 3	243 334
>70% of individuals and >70% of populations	112 557
Biallelic only	85 647
Single SNP per locus, MAF > 0.05	24 315
Sequencing errors, paralogs, multicopy loci and artefacts of library preparation	
1 Allele balance	20 828
2 Read orientation	12 878
3 Mapping quality	10 251
4 Paired reads	8876
5 Read quality	6905
6 Read depth	**5162**
Outlier detection	
BayeScan outliers	643
Arlequin outliers	697
Outliers identified in at least one method	873
Putatively neutral	4289
Putatively neutral in HWE	**3443**

HWE, Hardy–Weinberg equilibrium; MAF, minor allele frequency.

Smouse 2012). To evaluate whether populations have experienced recent genetic bottlenecks, we used BOTTLENECK 1.2.02 (Piry *et al.* 1999). BOTTLENECK was run using the infinite alleles model and a Wilcoxon signed-rank test implemented using the *wilcox.test* function in R was used to test for significant heterozygosity excess compared to expectations under mutation–drift equilibrium. We estimated the effective population size (N_e) using the linkage disequilibrium (LD) method in NEESTIMATOR 2.01 (Do *et al.* 2014). This is based on the assumption that LD at independently segregating loci in a finite population is a function of drift, and performs particularly well with a large number of loci in situations where population sizes are expected to be small. NEESTIMATOR was run assuming random mating and using a P_{crit} value of 0.075 following guidelines for small sample sizes (Waples & Do 2010). A Wilcoxon signed-rank test was used to test for the differences in N_e estimates between Lower and Upper Murray regions.

Population-specific F_{ST} was estimated for each sampling site for both the neutral and candidate loci using the method of Weir & Hill (2002) calculated with the R package HIERFSTAT (Goudet 2005). Population-specific F_{ST} estimates local population divergence from the whole metapopulation considering the variation in the strength of genetic drift among demes due to the differences in effective population size (Foll & Gaggiotti 2006).

Population genetic structure was assessed using the neutral loci with a combination of frequency- and genotype-based methods. Pairwise F_{ST} (Weir & Cockerham 1984) was estimated among sampling sites using GENODIVE with significance assessed using 10 000 permutations. GENODIVE was also used to perform a hierarchical AMOVA based on F_{ST} (Weir & Cockerham 1984) among catchments, among sites within catchments and among individuals within sites using 10 000 permutations. In all cases, missing data were replaced with alleles drawn randomly from the overall allele frequency distribution. Bayesian clustering analysis of individual genotypes was then performed using FASTSTRUCTURE (Raj *et al.* 2014). This model-based method assumes that there are K populations and that population allele frequencies are in HWE. Individuals are assigned to one or more populations based on the probability of their genotypes belonging to each population. Ten independent runs for each value of K (1–25) were completed to ensure consistency and the most likely K was assessed by comparing the model complexity that maximized marginal likelihood across replicate runs. Isolation by distance (IBD) was assessed using multiple matrix regression with randomization (MMRR) following Wang (2013). We examined the relationship between matrices of pairwise population F_{ST} calculated in GENODIVE and pairwise population distances along the river network calculated with ARCMAP v.10.2 (ESRI 2012), and tested for significance using 10 000 random permutations.

Environmental data and interaction among variables

To characterize environmental conditions at each sampling site, we used the comprehensive Australian hydrological geospatial fabric (Geoscience Australia 2011; Stein *et al.* 2014), which links spatial data depicting surface water features to a set of environmental attributes describing natural and anthropogenic characteristics of waterways. These include summary statistics on climate, land use, topography and hydrological characteristics organized according to stream hierarchy to allow the assessment of environmental factors at multiple scales (i.e. stream vs. catchment level). Also included in the environmental attributes are series of river disturbance indices designed to evaluate the impact of human activities such as disturbance to the flow regime and the effect of land use on the health of freshwater ecosystems (Stein *et al.* 2002).

A subset of 40 candidate variables were selected based on those identified as important predictors of freshwater fish occurrence in southeastern Australia (Bond *et al.* 2011) along with others that have previously been identified as influencing genetic diversity of freshwater fishes in the MDB (Table S1, Supporting information). These variables were divided into five categories concerning the variation in temperature, precipitation, flow regime, human disturbance and topography. Within each category, variance inflation factor (VIF) analysis was used to exclude highly correlated variables in a stepwise fashion until all remaining variables were below a VIF threshold of 10 (Dyer *et al.* 2010). Principal components analyses (PCAs) were then performed for the remaining variables in each category. This was carried out using the *dudi.pca* function in the ADE4 R package (Dray & Dufour 2007) and principal components (PCs) with eigenvalues greater than one were retained (Yeomans & Golder 1982) as synthetic environmental variables in GEA analyses. The *dimdesc* function in the FACTOMINER R package (Lê *et al.* 2008) was used to identify individual variables significantly ($P < 0.05$) associated with each PC.

Signatures of selection at local and regional scales

Evidence for local selection was assessed using both univariate and multivariate GEA methods to identify the strong associations between allele frequencies and environmental variables. First, we used a spatially explicit generalized linear mixed-model approach implemented in *gINLAnd* (Guillot *et al.* 2014). This method generates two competing models for each locus: one in which the fixed effect of an environmental variable influences population allele frequencies and one where the environmental variable has no effect. Both models account for spatial genetic structure by including a random spatial effect based on geographic coordinates. Due to the dendritic nature of the MDB river system, modelling the spatial arrangement of sites using xy coordinates provides a distorted measure of the true biological distance among sites. To overcome this, the *cmdscale* base function in R was used to perform multidimensional scaling (MDS) on the matrix of pairwise river distances. The MDS returned new coordinates that better represent among-site river distances. Using these coordinates, parameters describing the spatial covariance structure of the allele frequency data were estimated using a subset of 500 randomly selected loci as recommended by Guillot *et al.* (2014). These spatial parameters were then used to control for spatial genetic structure in the final models. *gINLAnd* was run for each of the environmental PCs and log-Bayes factors were calculated for the two models for each locus and used to rank loci in terms of dependence of the allele frequencies on the environmental variable. Using a conservative interpretation of Bayes factors (Kass & Raftery 1995), loci with a log-Bayes factor >15 were considered strong candidates for selection.

Second, we used partial redundancy analysis (RDA) to assess the effect of environmental variation on the patterns of genomic diversity while also controlling for the effects of spatial genetic structure using parameters describing the spatial distribution of sampling sites. The MDS spatial coordinates were expressed as third-degree polynomials and subjected to a forward selection procedure based on the method of Meirmans (2015). To account for collinear explanatory variables in the RDA model, VIF analysis was again used, this time to identify environmental PCs strongly correlated with other PCs in the model. Initially, RDA was performed with all retained environmental PCs before using a backwards-stepwise selection procedure implemented in VEGAN (Oksanen *et al.* 2015) to select the final model. The final RDA evaluated this reduced environmental model. Significance of the model, as well as marginal significance of each environmental PC, was assessed by 1000 ANOVA permutations. The mean locus score across all loci was calculated for each of the first three RDA axes, and individual loci with a score greater than three standard deviations from the mean were considered candidates for selection (Forester *et al.* 2015).

Functional annotation and mode of selection

To examine gene ontology (GO) annotation terms associated with the SNP loci, BLAST2GO (Conesa *et al.* 2005) was used to perform a BLAST search and annotation of the flanking sequences for all 5162 SNPs against the NCBI nonredundant nucleotide database with the BLAST *e*-value threshold set to 1×10^{-3} and an annotation threshold *e*-value threshold of 1×10^{-6}. Fisher's exact test was then used with a FDR of 0.05 to assess the GEA candidate loci for the enrichment of any biological processes, molecular functions or cellular components compared with the whole SNP data set.

Finally, to better understand the type of selection likely to have generated the genomic footprints detected by either F_{ST} outlier tests or GEA analyses, we examined the distribution of F_{ST} values observed for each locus for (i) the whole data set, (ii) the outlier loci (identified by ARLEQUIN and BAYESCAN) and (iii) the GEA candidate loci (identified by *gINLAnd* and the RDA). Here, we expect a relatively low average and broad distribution of F_{ST} values for polygenic 'soft' selective sweeps because in this scenario adaptation is expected to proceed without major changes in allele frequency (Pritchard *et al.* 2010). On the other hand, much higher F_{ST}

values are expected for loci involved in 'hard' selective sweeps because alternate alleles in these loci should have approached or reached fixation (i.e. F_{ST} of 1) (Pritchard & Di Rienzo 2010; Messer & Petrov 2013). F_{ST} for each locus was calculated in ARLEQUIN, and the density distribution kernel for each data set was plotted in R.

Results

Genotyping, outlier detection and genomewide variation

A total of 1 602 903 910 forward and reverse sequence reads were generated with the Illumina platform (detailed sequencing statistics are listed in Table S2, Supporting information). After filtering the data with stringent criteria, 5162 SNPs were retained from 2 589 251 variant sites present in the VCF file produced by *dDocent* (Table 2). BAYESCAN identified 643 outlier loci, while ARLEQUIN identified 697, with 467 of these identified by both. Outliers from both methods were conservatively combined such that the 873 unique loci considered as outliers by either BAYESCAN or ARLEQUIN were excluded to create a neutral data set. After filtering the 4289 remaining loci for HWE, 3443 putatively neutral SNPs were retained for population structure and demographic analyses.

Genetic variation based on 3443 neutral SNPs was low with an average H_E of 0.161 (0.057–0.263), average H_O of 0.123 (0.043–0.206) and an average of 46.3% (19.0–71.7%) polymorphic loci (Table 3). There was, however, a striking contrast between regions with average H_E of 0.253 in the Lower Murray wetlands compared with 0.143 in the upper reaches (Table 3). Overall, genetic variation for the candidate loci was generally lower, but followed a pattern similar to the neutral data (Table 3). Population-specific F_{ST} estimated with both neutral and candidate SNPs was generally inversely proportional to genetic diversity, with the most highly differentiated populations also containing the least genetic variation (Table 3). Results from the

Table 3 Summary by sampling site of expected heterozygosity (H_E), observed heterozygosity (H_O), % polymorphic loci and population-specific F_{ST} (Weir & Hill 2002) based on 3443 neutral and 216 candidate adaptive SNPs for *Nannoperca* from the Murray–Darling Basin. Lowland wetland sites referred to as Lower Murray in the text are indicated in boldface

		H_E		H_O		% Polymorphic loci		F_{ST}	
Catchment	Site	Neutral	Adaptive	Neutral	Adaptive	Neutral	Adaptive	Neutral	Adaptive
TOO	**TBA**	**0.227**	**0.225**	**0.151**	**0.155**	**58.4**	**58.8**	**0.06**	**0.03**
LMR	**ALE**	**0.263**	**0.269**	**0.161**	**0.153**	**71.7**	**70.8**	**0.07**	**0.03**
	MID	**0.262**	**0.255**	**0.158**	**0.161**	**67.1**	**64.8**	**0.09**	**0.05**
	MUN	**0.260**	**0.271**	**0.167**	**0.178**	**61.6**	**63.0**	**0.03**	**0.02**
ANG	MCM	0.097	0.066	0.090	0.057	26.7	16.7	0.56	0.58
AVO	MIC	0.114	0.080	0.104	0.068	32.8	23.2	0.41	0.47
CAM	JHA	0.091	0.069	0.073	0.055	26.7	19.4	0.36	0.38
GOU	MER	0.075	0.041	0.062	0.035	30.4	19.9	0.47	0.52
	TRA	0.075	0.034	0.059	0.026	23.9	13.0	0.43	0.47
	YEA	0.087	0.049	0.066	0.027	24.4	14.8	0.36	0.43
	PRA	0.243	0.194	0.180	0.149	68.2	56.5	0.18	0.23
	SEV	0.218	0.173	0.170	0.125	59.7	48.6	0.12	0.17
BRO	BEN	0.236	0.203	0.191	0.153	68.5	57.4	0.16	0.20
	SAM	0.234	0.188	0.206	0.161	68.7	58.3	0.19	0.24
	LIM	0.118	0.105	0.094	0.075	39.7	38.0	0.34	0.36
OVE	KIN	0.104	0.091	0.077	0.068	36.2	34.3	0.30	0.29
	HAP	0.114	0.070	0.094	0.063	33.0	25.9	0.37	0.40
	MEA	0.158	0.182	0.129	0.137	43.6	44.4	0.25	0.14
KIE	GAP	0.168	0.102	0.145	0.094	51.8	35.2	0.30	0.39
ALB	ALB	0.226	0.140	0.182	0.106	66.9	43.5	0.30	0.47
MIT	SPR	0.152	0.087	0.119	0.066	48.1	32.9	0.26	0.35
	GLE	0.143	0.074	0.117	0.057	42.8	23.2	0.41	0.51
	TAL	0.164	0.092	0.135	0.068	46.7	27.3	0.48	0.58
COP	COP	0.133	0.100	0.111	0.079	39.9	32.4	0.30	0.35
LAC	LRT	0.057	0.040	0.043	0.031	19.0	16.7	0.67	0.71
Mean	Mean	0.161	0.128	0.123	0.094	46.3	37.6	0.30	0.34
Min	Min	0.057	0.034	0.043	0.026	19.0	13.0	0.03	0.02
Max	Max	0.263	0.271	0.206	0.178	71.7	70.8	0.67	0.71

BOTTLENECK tests for excess heterozygosity confirmed recent reductions in population size at all sites ($P < 1 \times 10^{-10}$) except MER ($P = 0.193$) and LRT ($P = 0.748$) (Table S3, Supporting information). Estimates of N_e were generally low (Table S3, Supporting information), but varied between Lower and Upper Murray regions with average estimate of 194.75 (190.9–198.6) for Lower Murray sites significantly higher ($P = 0.02$) than the upper reaches average estimate of 88.4 (13.7–305.4).

Population genetic structure

High levels of population genetic structure were evident between most demes of *Nannoperca australis*, with population pairwise comparisons of F_{ST} ranging from 0 to 0.79 (global $F_{ST} = 0.48$). All pairwise F_{ST} estimates were significant ($P < 0.003$) except between Lower Lakes sites MID and MUN ($F_{ST} = -0.002$, $P = 0.66$) (Table S4, Supporting information). Results of the MMRR indicated that river distance was not a good predictor of F_{ST} and that no significant pattern of IBD was apparent across the MDB (regression coefficient = 0.108, $P = 0.342$).

Based on F_{ST}, AMOVA calculated across all sites attributed 30.3% of the variation to the differences among catchments ($P < 0.001$), 10.7% to differences between sites within catchments ($P < 0.001$) and 13.5% among individuals within sites ($P < 0.001$) (Table S5, Supporting information). When calculated separately, the AMOVA for each of the catchments containing multiple sites suggests the differences in levels of within-catchment connectivity across the MDB (Table S5, Supporting information). Sites within the Lower Murray appear to be highly connected, suggesting that TBA, ALE, MID and MUN constitute a single population. This is in contrast to less connected Upper Murray catchments (Table S5, Supporting information).

Clustering analysis in FASTSTRUCTURE based on neutral SNPs identified 12 distinct populations that mostly correspond with the MDB catchment boundaries (Fig. 2), except for the following. In the Lower Murray, TBA (Tookayerta Catchment) grouped together with ALE, MID and MUN (Lower Lakes). Goulburn River (five sites) and Broken River (three sites) catchments were split into three groups: an Upper Goulburn cluster (MER/TRA/YEA), a distinct Broken River site (LIM) and an admixed group consisting of two Lower Goulburn and two Broken River sites (PRA/SEV/BEN/SAM). The site at Albury is most similar to the Kiewa River site (GAP), however also shares some affinity with sites further upstream from the Mitta Mitta catchment.

Environmental data and interactions among variables

Calculating pairwise distance among sampling sites with the revised MDS coordinates improved the correlation with along-river distances ($R^2 = 0.97$) compared with the original spatial coordinates ($R^2 = 0.87$). The revised MDS coordinates were then substituted for downstream analyses requiring spatial coordinates.

Following VIF analyses, 19 variables representing all five environmental categories were retained for the environmental PCAs (see Table S1, Supporting information for the explanation of variable codes). These are four temperature variables (STRCOLDMTHMIN, STRDRYQTEMP, CATDRYQTEMP and STRWETQTEM), three precipitation variables (CATDRYQRAIN, STRWETQRAIN and CATCOLDQRAIN), two disturbance factors (CDI and FRDI), five flow variables (RUNCVMAXMTH, RUNPERENIALITY, RUNANNMEAN, SUBEROSIVITY and CATEROSIVITY) and five topographic variables (STRAHLER, SUBELEMAX, CATELEMEAN, VALLEYSLOPE and CATSLOPE).

The first two components of each PCA for the temperature, flow and topographic variables scored eigenvalues >1 and explained 75.3%, 83.1% and 83.3% of the total variance, respectively. Just one component each for the precipitation and human disturbance PCAs scored an eigenvalue >1, and thus, all individual variables rather than PCs for these categories were used for downstream analyses (Table S6, Supporting information). The PCA plots for temperature and precipitation depict the climatic gradient across the MDB (Fig. 3), with sites from the Lower Murray experiencing higher winter temperatures and lower rainfall than headwater sites such as those in the Ovens River and Mitta Mitta catchments. Measures of flow variation (erosivity and perenniality) dominate the first PC of the flow PCA and demonstrate that Lower Murray sites experience lower variation in flow than in headwater locations (Fig. S2a, Supporting information). The second flow PC explains the differences in average annual flow that tend to reflect accumulated increases in total flow from headwaters to the main channel rather than the variation in climate (Fig. S1a, Supporting information). The topographic PCA was most influenced by elevation and Strahler stream order, which describes intrinsic physical factors related to each site's position in the river network (Fig. S1b, Supporting information). In contrast to the other categories, there was no evidence of any regional spatial pattern for the human disturbance PCA, confirming that human disturbance mostly affects conditions at a local scale (Fig. S1c, Supporting information). This resulted in a final list of 11 environmental variables (six PCs and five individual precipitation and disturbance variables) that describe the variation in the

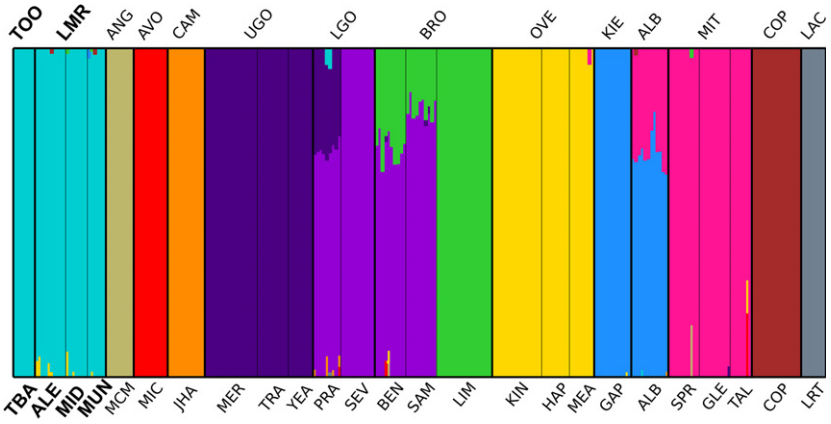

Fig. 2 Admixture plot based on 3443 'neutral' SNPs for *Nannoperca australis* from the Murray–Darling Basin (MDB) depicting $K = 13$ clusters determined by maximum marginal likelihood using FASTSTRUCTURE. Codes above and below the plot refer to catchment and sampling site, respectively (Table 1). Lowland wetland sites referred to as Lower Murray in the text are indicated in boldface.

environment across the MDB, and the correlations between individual variables and the PCs are described in Table S6 (Supporting information).

Signatures of selection at local and regional scales

gINLAnd provided strong evidence (log-Bayes factor >15) for the associations between allele frequencies of 178 unique loci and the 11 environmental variables identified based on PCAs, as above (Table S7, Supporting information). Candidate loci were identified by *gINLAnd* for all environmental variables, with precipitation-related variables associated with 85 loci, temperature (53 loci), flow (39 loci) and topography (26 loci). Variables associated with high numbers of loci included CATCOLDQRAIN (74), Temp2 (41), Flow1 (35) and catchment disturbance index (CDI) (35) (Fig. 4). As expected, there was also a high degree of overlap with 95 loci identified as candidates associated with more than one variable. Human disturbance variables describe mostly local-scale disturbance and were associated with 41 loci, of which 22 were not associated with any other variables.

The RDA triplot summarizes the first two axes of the final model and indicates that temperature, rainfall and topography are the major environmental factors influencing genetic variation of 42 candidate adaptive loci (Fig. 5). Winter rainfall (CATCOLDQRAIN) and summer temperature (Temp2) were the most influential predictive variables in the model suggesting that the climate is the major factor driving selection across the region. Following VIF analyses, Temp1, Temp2, CATCOLDQRAIN, Topo1 and Topo2 were retained as predictive variables in the final model, along with two spatial conditional variables. The RDA was globally significant ($P = 0.007$) with environmental variation explaining 23.83% of the total genetic variation after accounting for spatial structure (30.07% of total genetic variation). Assessment of the marginal significance of each explanatory variable revealed that Temp2, CATCOLDQRAIN, Topo1 and Topo2 were each significant predictors of allele frequencies ($P < 0.05$). The first three RDA axes explained 85.41% of the variation (33.89%, 28.88% and 22.65%, respectively) and individual locus scores for 42 loci (9, 17 and 16 for each of the three RDA axes) were further than three standard deviations from the mean (Fig. S2, Supporting information) and were considered as candidate loci potentially under selection. Triplots including RDA3 are presented in Fig. S3 (Supporting information).

Functional annotation and mode of selection

The GEA analyses together identified 216 unique candidate adaptive loci (178 *gINLAnd*, 42 RDA with four loci identified by both methods). BLAST2GO recorded blast hits for 1289 (*e*-value $< 1 \times 10^{-3}$) of the 5162 loci—of which 638 could be annotated and were assigned 885 GO terms (*e*-value $< 1 \times 10^{-6}$)—and 49 blast hits for the 216 candidate loci—of which 24 were annotated and were assigned 138 GO terms. Enrichment analysis did not find any GO terms significantly (FDR of 0.05) under- or over-represented in the candidate adaptive data set compared with all loci. Blast results and annotations are available on Dryad (see Data accessibility).

The distribution of F_{ST} values for the entire SNP data set is broad and extends across the entire theoretical range (i.e. from 0 to 1). This contrasts with the narrow distribution for the F_{ST} outlier loci, which include several values close to 1. On the other hand, the distribution of F_{ST} values for the GEA candidate loci is not only broader than the F_{ST} outliers, but also peaks at much lower F_{ST} values (Fig. 6). The average single locus F_{ST} for all 5162 loci was 0.461 (0–1), compared with 0.826 (0.695–1) for the 177 F_{ST} outliers and 0.634 (0.356–1) for the 216 GEA candidate loci.

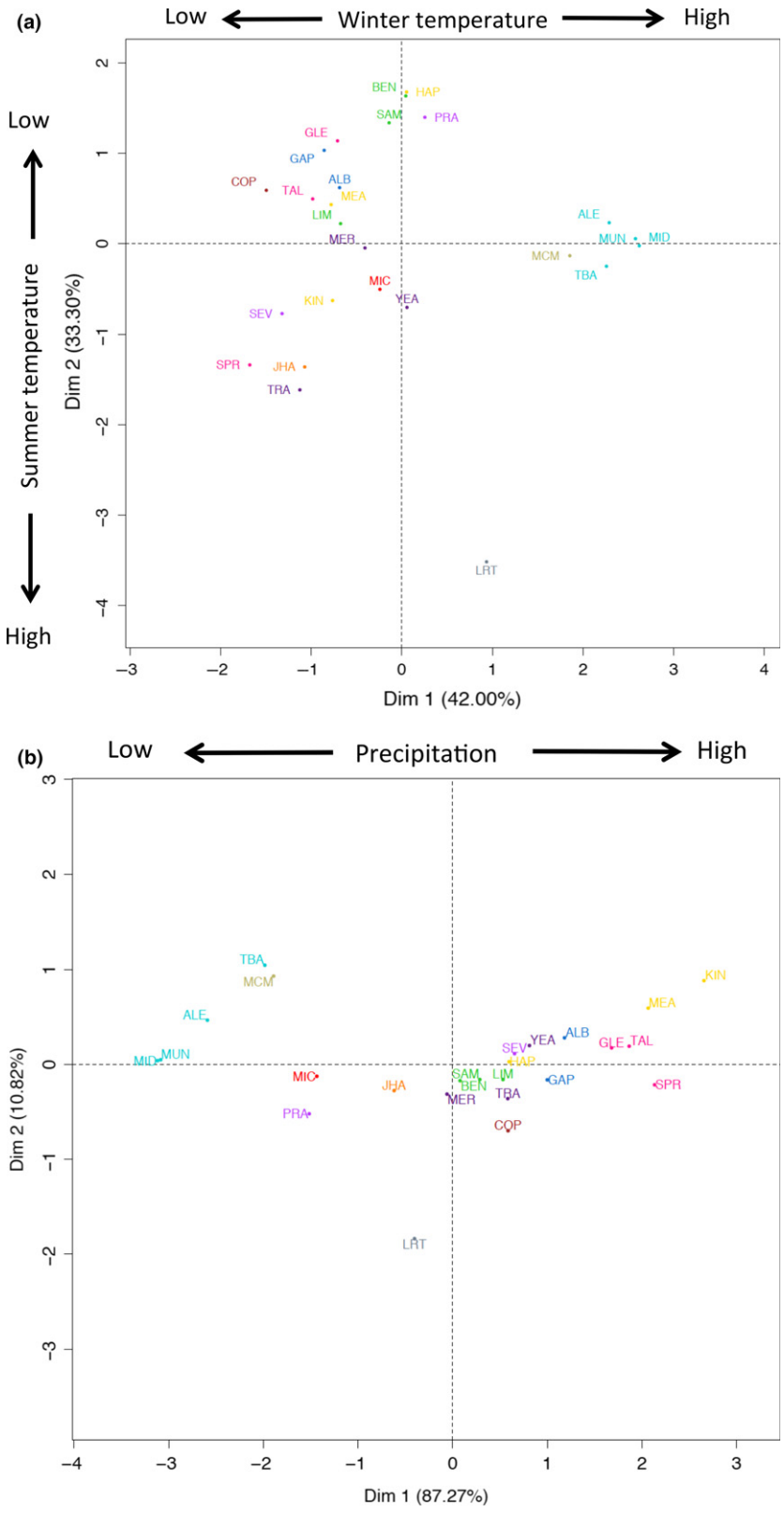

Fig. 3 Environmental principal components analyses (PCAs) of the Murray–Darling Basin (MDB) describing the relationship between each *Nannoperca australis* sampling location based on variables related to (a) temperature and (b) precipitation. Site names are colour-coded based on the colours used in Fig. 1. Annotations above and to the left of plots describe which variables contribute the most to each axis. Environmental PCAs based on variables related to flow, human disturbance and topography are in Fig. S1 (Supporting information).

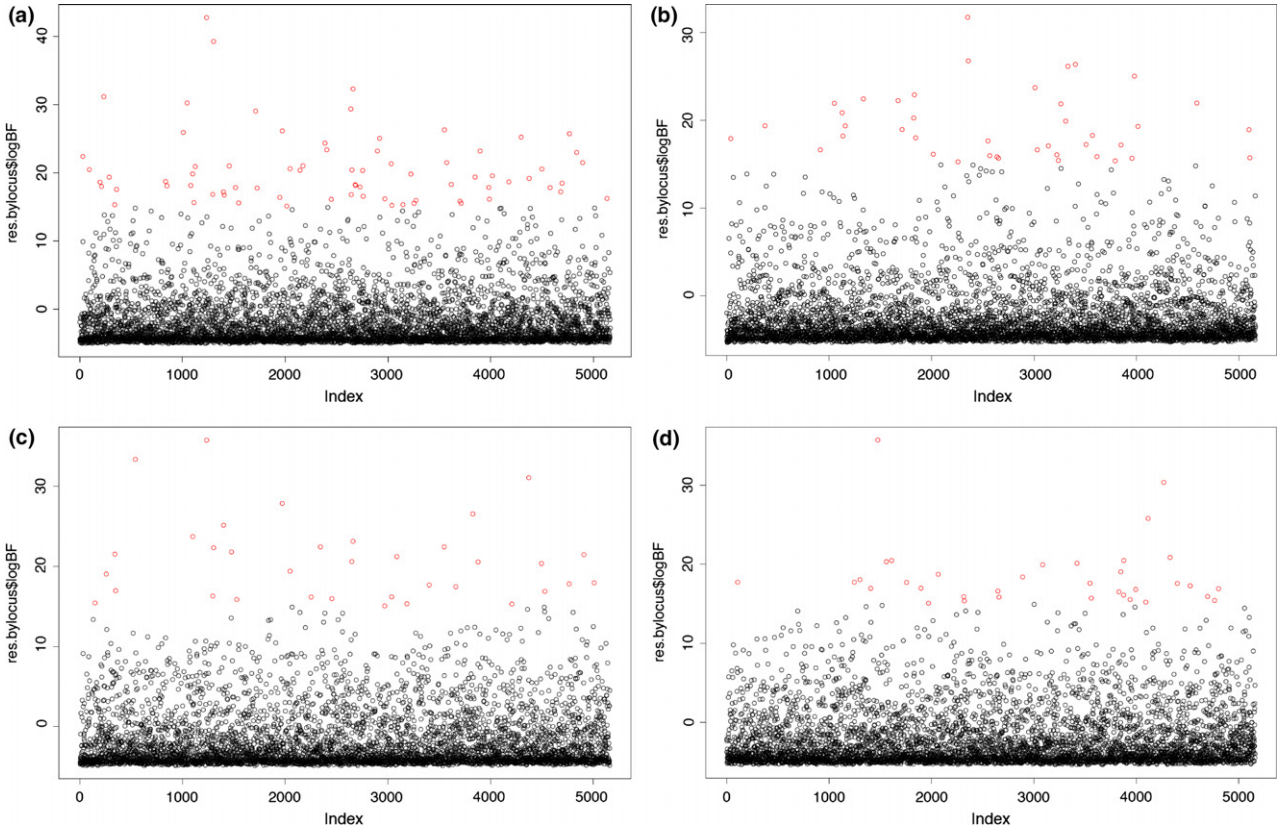

Fig. 4 log-Bayes factor scores for each of 5162 SNPs for their association with environmental variables. (a) Catchment average coldest quarter rainfall (CATCOLDQRAIN), (b) average summer temperature (Temp2 PC), (c) stream flow variation (Flow1 PC) and (d) catchment disturbance index (CDI). Loci with a log-Bayes factor >15 are highlighted in red and were considered as candidates for selection. PC, principal component.

Discussion

The rapid rate of climate change and other anthropogenic threats suggests that evolutionary adaptation will be required for many species to persist into future (Stockwell *et al.* 2003; Mergeay & Santamaria 2012; Losos *et al.* 2013). However, in order to gauge the potential of species to adapt to environmental change, we need to first understand how the environment shapes intraspecific variation across the genome. Here, replicate populations of a threatened and poor dispersive Australian freshwater fish sampled across a steep hydroclimatic gradient were examined using 5162 high-quality SNPs and compared with high-resolution environmental data in a riverscape genomics framework. Overall, strong population structure associated with the hierarchical river network and low genetic variation was identified with putatively neutral SNPs. This is consistent with the findings from studies based on other classes of selectively neutral markers (Attard *et al.* 2016; Cole *et al.* 2016), confirming that drift is a major evolutionary process shaping genetic diversity in this system.

On the other hand, evidence is also provided for a marked pattern of hydroclimatic-driven selection, with temperature and precipitation emerging as the most important of several environmental parameters influencing the allele frequencies of 216 candidate adaptive loci, both at regional (basinwide) and at local (catchment) scales. Human disturbance also influenced putatively adaptive variation, but for a smaller number of candidate loci and only at a local scale. In addition, despite strong drift and geographic isolation, adaptive divergence among populations appears to be explained by a pattern of nonallelic fixation consistent with a genomic footprint of polygenic adaptation. This represents the first riverscape genomics study of an Australian fish and, as such, makes an original contribution to our understanding of adaptation across large freshwater ecosystems—a topic dominated by studies of Northern Hemisphere fishes, in particular salmonids (e.g. Bourret *et al.* 2014; Hecht *et al.* 2015) and sticklebacks (e.g. Raeymaekers *et al.* 2014; Ferchaud & Hansen 2016). More broadly, our results highlight the utility of spatially explicit GEA methods for elucidating the

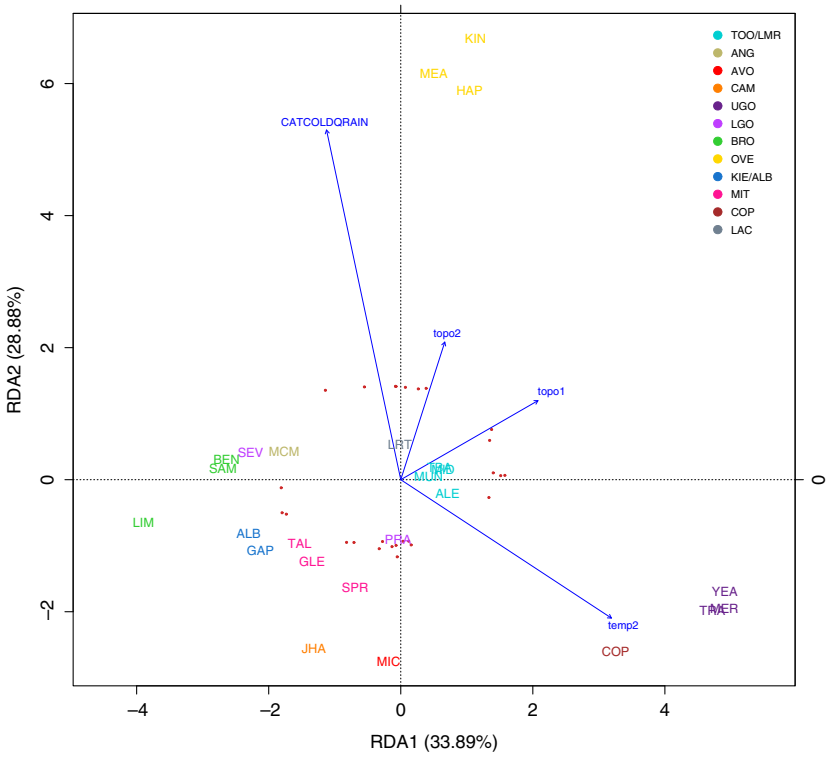

Fig. 5 Triplot summarizing the first two axes of the partial redundancy analysis (RDA). Sampling sites are colour-coded according to Fig. 1 and depict each site's position in the environmental model. Significant environmental factors ($P < 0.05$) are represented as blue vectors where the direction of the arrowhead indicates high values (e.g. site KIN receives the high rainfall, while site MIC receives low rainfall). The length of each vector represents the magnitude of their contribution to the model, and the angle between each vector represents the correlation between variables. Allele frequency vectors for individual SNPs significantly associated with the model have been rescaled to the same ordination space and are indicated by red markers. Their position depicts the direction of allele frequency variation in relation to the environmental model. Plots including RDA3 are in Fig. S3 (Supporting information).

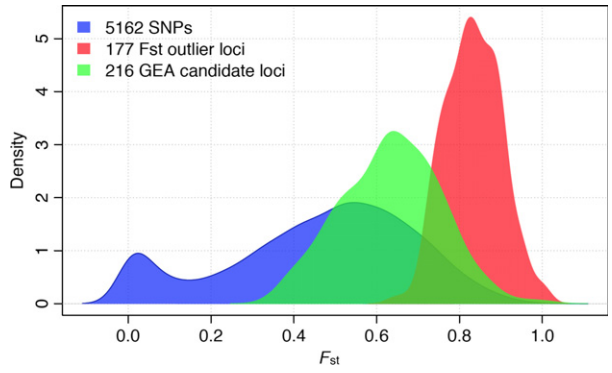

Fig. 6 Density distribution of F_{ST} values for all 5162 SNPs (blue), 177 F_{ST} outlier loci (red) and 216 candidate loci identified using genotype–environment association analyses (green).

signal of selection in spatially complex and anthropogenically modified ecosystems and for informing conservation management of endangered biodiversity.

Boom–bust cycles and dendritic landscapes influence genomewide variation

Understanding the relationship between landscape heterogeneity, environmental variability and population genetic diversity in river basins is an important topic in ecology and evolution because these are among the most diverse and yet most threatened ecosystems (Palmer *et al.* 2008; Strayer & Dudgeon 2010). Levels of genomewide variation in *Nannoperca australis* (Table 3) are lower than those reported in other population genomic studies of freshwater fishes (Matala *et al.* 2014; Skovrind *et al.* 2016). This is unsurprising and consistent with low levels of microsatellite DNA variation reported for *N. australis* (Cook *et al.* 2007; Attard *et al.* 2016; Cole *et al.* 2016), and more broadly for other MDB fishes (Faulks *et al.* 2010b, 2011; Brauer *et al.* 2013; Coleman *et al.* 2013; Sasaki *et al.* 2016), which exhibit generally very low genetic variation compared with non-Australian freshwater fishes (DeWoody & Avise 2000). This emerging paradigm is likely due to the naturally variable hydroclimatic environment of the MDB and of several other Australian river systems (Kennard *et al.* 2010) that result in frequent cycles of population booms and busts. These cycles cause fluctuations in population size that produce bottlenecks and affect the spatial patterns of gene flow in Australian freshwater fishes (Huey *et al.* 2008; Faulks *et al.* 2010b). Natural boom–bust cycles, likely in combination with recent and localized human disturbance (e.g. Attard *et al.* 2016), have no doubt contributed to the patterns observed here, as evident in the signal of genetic bottlenecks and small N_e inferred for populations across the basin.

Despite this general pattern, genomewide variation in *N. australis* was markedly reduced in the upper

compared to the lower reaches of the MDB. The Lower Murray is composed of a large system of linked wetlands and lakes, whereas the upper reaches of the MDB consist of small, often disconnected rivers and creeks (e.g. Hammer *et al.* 2013). As expected based on landscape configuration, N_e estimates obtained for *N. australis* from the lower MDB were significantly larger ($P = 0.02$) compared with estimates for the smaller upper MDB waterways. In addition, historical demographic analyses indicate that Lower Murray *N. australis* maintained relatively stable N_e until before European settlement (Attard *et al.* 2016), followed by very recent bottlenecks and near local extirpation (Hammer *et al.* 2013; Attard *et al.* 2016). Theoretical (Morrissey & de Kerckhove 2009; Paz-Vinas & Blanchet 2015; Thomaz *et al.* 2016) and empirical studies (Crispo *et al.* 2006; Barson *et al.* 2009; Osborne *et al.* 2014) of the effects of landscape structure on genetic variation of fishes suggest that not only intrinsic physical landscape properties but also asymmetrical downstream migration generates higher variation downstream compared with headwater populations. For *N. australis*, we detected strong and hierarchical population structure (i.e. differentiation was much greater among than within catchments) and no migration between most catchments (Fig. 2), consistent with nil contemporary microsatellite-based gene flow observed in a larger sample ($n = 578$) (Cole *et al.* 2016). Nonetheless, a more contiguous metapopulation occupying and dispersing along the Murray River corridor prior to European Settlement, enhanced in wetter periods over evolutionary timescales (Unmack *et al.* 2013; Cole *et al.* 2016), has also probably contributed to higher genomewide variation and lower population-specific F_{ST} observed in the Lower Murray (i.e. the latter is, on average, the most similar to all other populations in the basin; Fig. S1, Supporting information).

Detecting the signal of selection across a large and heterogeneous river basin

Environmental variability and instability likely exacerbate the effects of drift for *N. australis*, yet average hydroclimatic conditions vary substantially among catchments across the MDB. In this case, natural selection is also expected to contribute to population divergence, especially when gene flow among populations is restricted (Willi *et al.* 2007; Blanquart *et al.* 2012; Harrisson *et al.* 2014). Detecting the signal of selection in complex river networks, however, is particularly challenging and inferences can be misleading if based on approaches using inappropriate null models (Fourcade *et al.* 2013; Thomaz *et al.* 2016). Through the use of spatially explicit GEA methods, we aimed to disentangle the signal of adaptive variation responding to the environment from the strong spatial pattern of neutral genetic variation. The RDA confirmed that spatial population structure was responsible for the patterns of genomewide diversity (30.07% of total genetic variation); however, temperature, precipitation and topography were also important factors accounting for a large amount of the residual variation (23.83% of the total). This is reinforced by the *gINLAnd* results, where precipitation and temperature variables were associated with the majority of candidate adaptive loci (106 and 58, respectively, of 178). Our suggestion of adaptive population divergence is further strengthened by the fact that the loci identified by the GEA methods are responding in parallel with the environment across a number of demographically independent populations. This also builds on studies showing that local adaptation of traits related to reproductive fitness in *N. australis* varies predictably along the gradients of variability in water flow (Morrongiello *et al.* 2010, 2012, 2013), and supports the hypothesis that hydroclimatic selection has driven adaptive genetic differentiation of populations. These results also add to a growing body of evidence that climate is a major factor contributing to adaptive divergence among freshwater fish populations. For example, Bourret *et al.* (2013) found that climate and geology were associated with adaptive divergence of Atlantic salmon (*Salmo salar*) populations, and Hecht *et al.* (2015) identified precipitation and temperature as significant factors shaping adaptive variation of Chinook salmon (*Oncorhynchus tshawytscha*).

Powerful GEA methods have also recently shown promise in detecting polygenic adaptation in natural populations (Lasky *et al.* 2012; Bourret *et al.* 2014; Hecht *et al.* 2015). Empirical and modelling studies suggest that local adaptation to environmental change may predominantly arise through polygenic 'soft' selective sweeps (Hermisson & Pennings 2005; Pritchard & Di Rienzo 2010). This mechanism involves relatively small changes in allele frequencies at a large number of loci underlying the trait under selection. On the other hand, genome scans based on F_{ST} outlier tests are primarily designed to identify 'hard' selective sweeps that lead to fixation or near fixation of alternate alleles (Messer & Petrov 2013). For our data set, the distribution of F_{ST} values for the vast majority of GEA loci is inconsistent with alternate fixation of candidate adaptive alleles (Fig. 5; mean F_{ST} of 0.634) and instead supports recent views that adaptation of complex quantitative traits probably takes place by simultaneous selection acting on variants at many loci of small effects (Pritchard *et al.* 2010). Rapid adaptation to environmental change due to polygenic selection is possible if sufficient standing genetic variation exists in the population (Pritchard

et al. 2010; Crisci *et al.* 2016), underscoring the potential benefits of incorporating GEA methods into conservation studies of adaptation (Le Corre & Kremer 2012).

Are small fragmented populations subjected to more divergent selection?

As the field of LG is evolving, so too is the idea that fragmentation not only reduces habitat size and quality, but also increases environmental variation within, and among, habitat fragments (Wood *et al.* 2014). An emerging paradigm challenges the classical view of population genetic theory and predicts that natural selection can promote adaptive diversity even in small populations where strong drift is expected to constrain adaptive evolution (Koskinen *et al.* 2002; Fraser *et al.* 2014; Wood *et al.* 2016). Our results support this idea and the hypothesis that, despite reduced genetic diversity due to drift in the small and fragmented populations typically found in catchments from the upper MDB, heterogeneous selection pressure is also driving local adaptive divergence in response to increased environmental variation resulting from decreasing local habitat size. The environmental analysis shows that many upper catchments appear to harbour unique and divergent habitats, especially in regard to precipitation (Fig. 3a) and to flow and human disturbance (Fig. S2, Supporting information). This is in part supported by the RDA results, where several upper MDB catchment populations showed the most divergent GEA profiles across the basin (e.g. Ovens and Upper Goulburn rivers). Accordingly, fragmentation and habitat quality are known to impose divergent selection that alters microgeographic adaptation in isolated populations (Willi & Hoffmann 2012). An alternative, but not mutually exclusive, view is that habitat complexity, rather than size only, might have impacted on evolutionary persistence of these small fragmented populations. Here, fragmentation would lead to random subsampling of habitat types, with population persistence at small N_e dependent on the quality of the habitat sampled in the resulting occupied fragment (Fraser *et al.* 2014). For instance, adaptive differentiation in fragmented brook trout (*Salvelinus fontinalis*) populations was greater among small than among large populations, with very small populations still very much affected by natural selection (Fraser *et al.* 2014). Whether this could lead to more variable evolutionary responses (and even perhaps enhance persistence) in fragmented *N. australis* is a hypothesis that remains to be assessed.

Regardless of the mechanisms, the small and fragmented upper MDB populations have comparatively little standing variation at hydroclimatically selected loci compared with the significantly larger Lower Murray populations (Table 3). This highlights that these populations may have reduced adaptive potential to respond to the rapid climate change. Conversely, in the Lower Murray, the combination of higher diversity and low population-specific F_{ST} at candidate loci [i.e. low adaptive divergence (Funk *et al.* 2012)] indicates that these populations may be a reservoir of adaptive variation for the species and reinforces the critical nature of ongoing conservation efforts in this region (e.g. Bice *et al.* 2012; Hammer *et al.* 2013; Attard *et al.* 2016). The adaptive sink hypothesis is also consistent with recent findings of unexpectedly high levels of variation and strong positive selection at the MHC IIB gene of Lower Murray *N. australis* (Bracamonte *et al.* 2015).

Implications and recommendations for conservation

Despite ongoing conservation efforts including habitat restoration, environmental water allocation, captive breeding and re-introductions (Bice *et al.* 2012; Hammer *et al.* 2013; Pearce 2015; Attard *et al.* 2016), *N. australis* remains endangered or threatened across the MDB. In fact, during sampling for this study, we observed at least 10 populations that are now extirpated due to the loss of habitat associated with river regulation and drought. To promote the long-term persistence of remnant populations, conservation efforts need to be proactive and should focus on maintaining natural habitat and restoring evolutionary processes to avoid further loss of genetic diversity and to increase resilience to environmental change (Crook *et al.* 2010; Morrongiello *et al.* 2011; Hammer *et al.* 2013). In this sense, carefully considered translocations provide an attractive option for conservation management of small and fragmented populations (Sgrò *et al.* 2011; Weeks *et al.* 2011; Frankham *et al.* 2014; Frankham 2015). Genetic rescue (Tallmon *et al.* 2004) [here we also include the closely related concept of genetic restoration proposed by Hedrick (2005)] can occur when translocations are used to restore gene flow between recently isolated populations. This can thereby reduce the genetic consequences associated with small population sizes such as inbreeding depression, reduced genetic variation and genetic load (Weeks *et al.* 2011; Whiteley *et al.* 2015). For *N. australis*, translocations among populations within catchments would replicate natural evolutionary and demographic processes by restoring connectivity among recently isolated but historically connected demes. This could be achieved by translocating several individuals to provide ~20% gene flow initially, followed by a small number of migrants per generation thereafter (Hedrick 1995; Lopez *et al.* 2009). It has been suggested that outbreeding depression could lead to reduced fitness in target populations (Edmands 2007); however, Frankham

(2015) argued that the risk of outbreeding depression has likely been overemphasized in the literature (see also Weeks *et al.* 2016). Populations of *N. australis* have clearly been historically connected at the catchment level, and given the observed historic and ongoing declines, we argue that their risk of extirpation due to inbreeding depression, loss of genetic diversity and stochastic demographic events outweighs risks posed by outbreeding depression.

Where species inhabit a wide range of environments, the potential also exists to select source populations based on information from LG and predictions of future environmental conditions to build evolutionary resilience to future environmental change (Aitken & Whitlock 2013). In addition to genetic rescue, our findings provide the opportunity to also consider a second strategy, and use translocations to introduce new alleles that may increase the potential for populations to adapt to environmental change (Sgrò *et al.* 2011; Aitken & Whitlock 2013). Populations of *N. australis* in the Lower Murray experience hotter and drier conditions than elsewhere in the MDB (Fig. 4), and our results suggest that these populations are locally adapted. Increasing aridity and climate variability are predicted for the whole MDB in future (Kershaw *et al.* 2003; Morrongiello *et al.* 2011; Davis *et al.* 2015), and we propose that translocations within catchments could be additionally supplemented with a small number of individuals harbouring adaptive genetic variation from other populations. In this case, the Lower Murray would provide an ideal source population due to their higher genetic diversity and low average neutral, and potentially adaptive, divergence.

Conclusions

Understanding the evolutionary potential of populations to respond to rapid climate change demands knowledge of how environmental factors contribute to local adaptation of populations. The recent transition from landscape genetics to LG has already yielded strong evidence for the role of climate in shaping the patterns of intraspecific genetic variation. Inferring selection in complex spatial environments, however, remains challenging. Our riverscape genomic approach used spatially explicit GEA methods to control for the effects of landscape structure and shared population history. It showed that hydroclimatic conditions influence the population genetic architecture of *Nannoperca australis* in the MDB. We revealed precipitation and temperature as the most important of several environmental parameters influencing adaptive genetic variation, both at local and at regional scales. Human disturbance also influenced putatively adaptive variation, but only at a local scale. The 216 candidate loci we identified provide a basis for further work exploring the functional significance of genomic regions involved in local adaptation to hydroclimatic heterogeneity. Recently, there has been a call for genomic approaches currently used to address questions in ecology and evolution to move beyond the realm of academic research and contribute more to solving the practical issues of conservation biology (Shafer *et al.* 2015a). Our work is an initial step towards that goal and will inspire further debate and research into how knowledge of adaptive genetic variation may best be incorporated into species conservation.

Acknowledgements

We thank the many people who helped with fieldwork, especially Peter Unmack and Mark Adams for curating samples. Minami Sasaki and Jonathan Sandoval-Castillo provided valuable assistance with laboratory and bioinformatics, respectively. We also thank Vincent Bourret for comments on the manuscript. Collections were obtained under permits from various state fisheries agencies, and research is under Flinders University Animal Welfare Committee approval E313. Financial support was provided by the Australian Research Council via a Future Fellowship Project to Luciano Beheregaray (FT130101068) and the AJ & IM Naylon PhD Scholarship to Chris Brauer.

References

Aitken SN, Whitlock MC (2013) Assisted gene flow to facilitate local adaptation to climate change. *Annual Review of Ecology, Evolution, and Systematics*, **44**, 367–388.

Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. *Nature Reviews Genetics*, **11**, 697–709.

Attard C, Möller L, Sasaki M *et al.* (2016) A novel holistic framework for genetic-based captive-breeding and reintroduction programs. *Conservation Biology*, **30**, 1060–1069.

Balcombe SR, Sheldon F, Capon SJ *et al.* (2011) Climate-change threats to native fish in degraded rivers and floodplains of the Murray-Darling Basin, Australia. *Marine and Freshwater Research*, **62**, 1099–1114.

Barrett RDH, Schluter D (2008) Adaptation from standing genetic variation. *Trends in Ecology & Evolution*, **23**, 38–44.

Barson N, Cable J, Van Oosterhout C (2009) Population genetic analysis of microsatellite variation of guppies (*Poecilia reticulata*) in Trinidad and Tobago: evidence for a dynamic source–sink metapopulation structure, founder events and population bottlenecks. *Journal of Evolutionary Biology*, **22**, 485–497.

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London B: Biological Sciences*, **263**, 1619–1626.

Bellard C, Bertelsmeier C, Leadley P, Thuiller W, Courchamp F (2012) Impacts of climate change on the future of biodiversity. *Ecology Letters*, **15**, 365–377.

Bice C, Whiterod N, Wilson P, Zampatti B, Hammer M (2012) *The Critical Fish Habitat Project: Reintroductions of Threatened Fish Species in the Coorong, Lower Lakes and Murray Mouth Region in 2011/12*. SARDI Aquatic Sciences, Adelaide, Australia.

Blanquart F, Gandon S, Nuismer S (2012) The effects of migration and drift on local adaptation to a heterogeneous environment. *Journal of Evolutionary Biology*, **25**, 1351–1363.

Bond N, Thomson J, Reich P, Stein J (2011) Using species distribution models to infer potential climate change-induced range shifts of freshwater fish in south-eastern Australia. *Marine and Freshwater Research*, **62**, 1043–1061.

Bourret V, Dionne M, Kent MP, Lien S, Bernatchez L (2013) Landscape genomics in Atlantic salmon (*Salmo salar*): searching for gene–environment interactions driving local adaptation. *Evolution*, **67**, 3469–3487.

Bourret V, Dionne M, Bernatchez L (2014) Detecting genotypic changes associated with selective mortality at sea in Atlantic salmon: polygenic multilocus analysis surpasses genome scan. *Molecular Ecology*, **23**, 4444–4457.

Bracamonte SE, Smith S, Hammer M *et al.* (2015) Characterization of MHC class IIB for four endangered Australian freshwater fishes obtained from ecologically divergent populations. *Fish & Shellfish Immunology*, **46**, 468–476.

Brauer CJ, Unmack PJ, Hammer MP, Adams M, Beheregaray LB (2013) Catchment-scale conservation units identified for the threatened Yarra pygmy perch (*Nannoperca obscura*) in highly modified river systems. *PLoS ONE*, **8**, e82953.

Cahill AE, Aiello-Lammens ME, Fisher-Reid MC *et al.* (2012) How does climate change cause extinction? *Proceedings of the Royal Society of London B: Biological Sciences*, **280**, 1–10.

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

Charmantier A, McCleery RH, Cole LR *et al.* (2008) Adaptive phenotypic plasticity in response to climate change in a wild bird population. *Science*, **320**, 800–803.

Chiew F, Teng J, Kirono D *et al.* (2008) *Climate Data for Hydrologic Scenario Modelling Across the Murray-Darling Basin: A Report to the Australian Government from the CSIRO Murray-Darling Basin Sustainable Yields Project*. CSIRO, Canberra, Australian Capital Territory.

Cole T, Hammer M, Unmack P *et al.* (2016) Range-wide fragmentation in a threatened fish associated with post-European settlement modification in the Murray-Darling Basin, Australia. *Conservation Genetics*, doi: 10.1007/s10592-10016-10868-10598.

Coleman RA, Weeks AR, Hoffmann AA (2013) Balancing genetic uniqueness and genetic variation in determining conservation and translocation strategies: a comprehensive case study of threatened dwarf galaxias, *Galaxiella pusilla* (Mack) (Pisces: Galaxiidae). *Molecular Ecology*, **22**, 1820–1835.

Conesa A, Götz S, García-Gómez JM *et al.* (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, **21**, 3674–3676.

Cook BD, Bunn SE, Hughes JM (2007) Molecular genetic and stable isotope signatures reveal complementary patterns of population connectivity in the regionally vulnerable southern pygmy perch (*Nannoperca australis*). *Biological Conservation*, **138**, 60–72.

Cooke GM, Chao NL, Beheregaray LB (2012a) Divergent natural selection with gene flow along major environmental gradients in Amazonia: insights from genome scans, population genetics and phylogeography of the characin fish *Triportheus albus*. *Molecular Ecology*, **21**, 2410–2427.

Cooke GM, Chao NL, Beheregaray LB (2012b) Natural selection in the water: freshwater invasion and adaptation by water colour in the Amazonian pufferfish. *Journal of Evolutionary Biology*, **25**, 1305–1320.

Cooke GM, Landguth EL, Beheregaray LB (2014) Riverscape genetics identifies replicated ecological divergence across an Amazonian ecotone. *Evolution*, **68**, 1947–1960.

Coop G, Witonsky D, Di Rienzo A, Pritchard JK (2010) Using environmental correlations to identify loci underlying local adaptation. *Genetics*, **185**, 1411–1423.

Crisci JL, Dean MD, Ralph P (2016) Adaptation in isolated populations: when does it happen and when can we tell? *Molecular Ecology*, **25**, 3901–3911.

Crispo E, Bentzen P, Reznick DN, Kinnison MT, Hendry AP (2006) The relative influence of natural selection and geography on gene flow in guppies. *Molecular Ecology*, **15**, 49–62.

Crook DA, Reich P, Bond NR *et al.* (2010) Using biological information to support proactive strategies for managing freshwater fish during drought. *Marine and Freshwater Research*, **61**, 379–387.

Davis MB, Shaw RG (2001) Range shifts and adaptive responses to Quaternary climate change. *Science*, **292**, 673–679.

Davis J, O'Grady AP, Dale A *et al.* (2015) When trends intersect: the challenge of protecting freshwater ecosystems under multiple land use and hydrological intensification scenarios. *Science of the Total Environment*, **534**, 65–78.

Dawson TP, Jackson ST, House JI, Prentice IC, Mace GM (2011) Beyond predictions: biodiversity conservation in a changing climate. *Science*, **332**, 53–58.

DeWitt TJ, Sih A, Wilson DS (1998) Costs and limits of phenotypic plasticity. *Trends in Ecology & Evolution*, **13**, 77–81.

DeWoody JA, Avise JC (2000) Microsatellite variation in marine, freshwater and anadromous fishes compared with other animals. *Journal of Fish Biology*, **56**, 461–473.

Do C, Waples RS, Peel D *et al.* (2014) NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (*Ne*) from genetic data. *Molecular Ecology Resources*, **14**, 209–214.

Donders TH, Haberle SG, Hope G, Wagner F, Visscher H (2007) Pollen evidence for the transition of the Eastern Australian climate system from the post-glacial to the present-day ENSO mode. *Quaternary Science Reviews*, **26**, 1621–1637.

Dray S, Dufour A-B (2007) The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software*, **22**, 1–20.

Dyer RJ, Nason JD, Garrick RC (2010) Landscape modelling of gene flow: improved power using conditional genetic distance derived from the topology of population networks. *Molecular Ecology*, **19**, 3746–3759.

Edmands S (2007) Between a rock and a hard place: evaluating the relative risks of inbreeding and outbreeding for conservation and management. *Molecular Ecology*, **16**, 463–475.

ESRI (2012) *ArcMap 10.1*. Environmental Systems Research Institute, Inc., Redlands, California.

Excoffier L, Lischer HEL (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.

Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically structured population. *Heredity*, **103**, 285–298.

Faulks LK, Gilligan DM, Beheregaray LB (2010a) Evolution and maintenance of divergent lineages in an endangered freshwater fish, *Macquaria australasica*. *Conservation Genetics*, **11**, 921–934.

Faulks LK, Gilligan DM, Beheregaray LB (2010b) Islands of water in a sea of dry land: hydrological regime predicts genetic diversity and dispersal in a widespread fish from Australia's arid zone, the golden perch (*Macquaria ambigua*). *Molecular Ecology*, **19**, 4723–4737.

Faulks LK, Gilligan DM, Beheregaray LB (2011) The role of anthropogenic vs. natural in-stream structures in determining connectivity and genetic diversity in an endangered freshwater fish, Macquarie perch (*Macquaria australasica*). *Evolutionary Applications*, **4**, 589–601.

Ferchaud A-L, Hansen MM (2016) The impact of selection, gene flow and demographic history on heterogeneous genomic divergence: three-spine sticklebacks in divergent environments. *Molecular Ecology*, **25**, 238–259.

Foll M, Gaggiotti O (2006) Identifying the environmental factors that determine the genetic structure of populations. *Genetics*, **174**, 875–891.

Foll M, 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.

Forester BR, Jones MR, Joost S, Landguth EL, Lasky JR (2015) Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. *Molecular Ecology*, **25**, 104–120.

Fourcade Y, Chaput-Bardy A, Secondi J, Fleurant C, Lemaire C (2013) Is local selection so widespread in river organisms? Fractal geometry of river networks leads to high bias in outlier detection. *Molecular Ecology*, **22**, 2065–2073.

Frankham R (2010) Where are we in conservation genetics and where do we need to go? *Conservation Genetics*, **11**, 661–663.

Frankham R (2015) Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. *Molecular Ecology*, **24**, 2610–2618.

Frankham R, Bradshaw CJA, Brook BW (2014) Genetics in conservation management: revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biological Conservation*, **170**, 56–63.

Fraser DJ, Debes PV, Bernatchez L, Hutchings JA (2014) Population size, habitat fragmentation, and the nature of adaptive variation in a stream fish. *Proceedings of the Royal Society of London B: Biological Sciences*, **281**, 1–8.

Frichot E, Schoville SD, Bouchard G, François O (2013) Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution*, **30**, 1687–1699.

Funk WC, McKay JK, Hohenlohe PA, Allendorf FW (2012) Harnessing genomics for delineating conservation units. *Trends in Ecology & Evolution*, **27**, 489–496.

Funk WC, Lovich RE, Hohenlohe PA *et al.* (2016) Adaptive divergence despite strong genetic drift: genomic analysis of the evolutionary mechanisms causing genetic differentiation in the island fox (*Urocyon littoralis*). *Molecular Ecology*, **25**, 2176–2194.

Garner BA, Hand BK, Amish SJ *et al.* (2015) Genomics in conservation: case studies and bridging the gap between data and application. *Trends in Ecology & Evolution*, **31**, 81–83.

Geoscience Australia (2011) *National surface water information*. Available from http://www.ga.gov.au/topographic-mapping/national-surface-water-information.html

Gienapp P, Teplitsky C, Alho JS, Mills JA, Merila J (2008) Climate change and evolution: disentangling environmental and genetic responses. *Molecular Ecology*, **17**, 167–178.

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

Guillot G, Vitalis R, le Rouzic A, Gautier M (2014) Detecting correlation between allele frequencies and environmental variables as a signature of selection. A fast computational approach for genome-wide studies. *Spatial Statistics*, **8**, 145–155.

Hammer MP, Bice CM, Hall A *et al.* (2013) Freshwater fish conservation in the face of critical water shortages in the southern Murray-Darling Basin, Australia. *Marine and Freshwater Research*, **64**, 807–821.

Hand BK, Muhlfeld CC, Wade AA *et al.* (2015) Climate variables explain neutral and adaptive variation within salmonid metapopulations: the importance of replication in landscape genetics. *Molecular Ecology*, **25**, 689–705.

Harrisson KA, Pavlova A, Telonis-Scott M, Sunnucks P (2014) Using genomics to characterize evolutionary potential for conservation of wild populations. *Evolutionary Applications*, **7**, 1008–1025.

Hecht BC, Matala AP, Hess JE, Narum SR (2015) Environmental adaptation in Chinook salmon (*Oncorhynchus tshawytscha*) throughout their North American range. *Molecular Ecology*, **24**, 5573–5595.

Hedrick PW (1995) Gene flow and genetic restoration: the Florida panther as a case study. *Conservation Biology*, **9**, 996–1007.

Hedrick P (2005) 'Genetic restoration': a more comprehensive perspective than 'genetic rescue'. *Trends in Ecology & Evolution*, **20**, 109.

Hendry AP, Farrugia TJ, Kinnison MT (2008) Human influences on rates of phenotypic change in wild animal populations. *Molecular Ecology*, **17**, 20–29.

Hermisson J, Pennings PS (2005) Soft sweeps. *Genetics*, **169**, 2335–2352.

Huey JA, Baker AM, Hughes JM (2008) The effect of landscape processes upon gene flow and genetic diversity in an Australian freshwater fish, *Neosilurus hyrtlii*. *Freshwater Biology*, **53**, 1393–1408.

Hughes JM, Schmidt DJ, Finn DS (2009) Genes in streams: using DNA to understand the movement of freshwater fauna and their riverine habitat. *BioScience*, **59**, 573–583.

Joost S, Bonin A, Bruford MW *et al.* (2007) A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Molecular Ecology*, **16**, 3955–3969.

Kass RE, Raftery AE (1995) Bayes factors. *Journal of the American Statistical Association*, **90**, 773–795.

Kennard MJ, Pusey BJ, Olden JD *et al.* (2010) Classification of natural flow regimes in Australia to support environmental flow management. *Freshwater Biology*, **55**, 171–193.

Kershaw P, Moss P, Van Der Kaars S (2003) Causes and consequences of long-term climatic variability on the Australian continent. *Freshwater Biology*, **48**, 1274–1283.

Koskinen MT, Haugen TO, Primmer CR (2002) Contemporary fisherian life-history evolution in small salmonid populations. *Nature*, **419**, 826–830.

Laporte M, Pavey SA, Rougeux C *et al.* (2015) RAD sequencing reveals within-generation polygenic selection in response to anthropogenic organic and metal contamination in North Atlantic Eels. *Molecular Ecology*, **25**, 219–237.

Lasky JR, Des Marais DL, McKay J *et al.* (2012) Characterizing genomic variation of *Arabidopsis thaliana*: the roles of geography and climate. *Molecular Ecology*, **21**, 5512–5529.

Laurance WF, Dell B, Turton SM *et al.* (2011) The 10 Australian ecosystems most vulnerable to tipping points. *Biological Conservation*, **144**, 1472–1480.

Le Corre V, Kremer A (2012) The genetic differentiation at quantitative trait loci under local adaptation. *Molecular Ecology*, **21**, 1548–1566.

Lê S, Josse J, Husson F (2008) FactoMineR: an R package for multivariate analysis. *Journal of Statistical Software*, **25**, 1–18.

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

Limborg MT, Helyar SJ, De Bruyn M *et al.* (2012) Environmental selection on transcriptome-derived SNPs in a high gene flow marine fish, the Atlantic herring (*Clupea harengus*). *Molecular Ecology*, **21**, 3686–3703.

Lintermans M (2007) *Fishes of the Murray-Darling Basin: An Introductory Guide*. Murray-Darling Basin Commission, Canberra, Australian Capital Territory.

Llewellyn L (1974) Spawning, development and distribution of the southern pigmy perch *Nannoperca australis australis* Gunther from inland waters in eastern Australia. *Marine and Freshwater Research*, **25**, 121–149.

Lopez S, Rousset F, Shaw FH, Shaw RG, Ronce O (2009) Joint effects of inbreeding and local adaptation on the evolution of genetic load after fragmentation. *Conservation Biology*, **23**, 1618–1627.

Losos JB, Arnold SJ, Bejerano G *et al.* (2013) Evolutionary biology for the 21st century. *PLoS Biology*, **11**, e1001466.

Lotterhos KE, Whitlock MC (2014) Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Molecular Ecology*, **23**, 2178–2192.

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

Manel S, Joost S, Epperson BK *et al.* (2010) Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field. *Molecular Ecology*, **19**, 3760–3772.

Matala AP, Ackerman MW, Campbell MR, Narum SR (2014) Relative contributions of neutral and non-neutral genetic differentiation to inform conservation of steelhead trout across highly variable landscapes. *Evolutionary Applications*, **7**, 682–701.

Meirmans PG (2015) Seven common mistakes in population genetics and how to avoid them. *Molecular Ecology*, **24**, 3223–3231.

Meirmans PG, Van Tienderen PH (2004) GenoType and GenoDive: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, **4**, 792–794.

Mergeay J, Santamaria L (2012) Evolution and biodiversity: the evolutionary basis of biodiversity and its potential for adaptation to global change. *Evolutionary Applications*, **5**, 103–106.

Messer PW, Petrov DA (2013) Population genomics of rapid adaptation by soft selective sweeps. *Trends in Ecology & Evolution*, **28**, 659–669.

Moore J-S, Bourret V, Dionne M *et al.* (2014) Conservation genomics of anadromous Atlantic salmon across its North American range: outlier loci identify the same patterns of population structure as neutral loci. *Molecular Ecology*, **23**, 5680–5697.

Morrissey MB, de Kerckhove DT (2009) The maintenance of genetic variation due to asymmetric gene flow in dendritic metapopulations. *The American Naturalist*, **174**, 875–889.

Morrongiello JR, Bond NR, Crook DA, Wong BBM (2010) Nuptial coloration varies with ambient light environment in a freshwater fish. *Journal of Evolutionary Biology*, **23**, 2718–2725.

Morrongiello JR, Beatty SJ, Bennett JC *et al.* (2011) Climate change and its implications for Australia's freshwater fish. *Marine and Freshwater Research*, **62**, 1082–1098.

Morrongiello JR, Bond NR, Crook DA, Wong BBM (2012) Spatial variation in egg size and egg number reflects trade-offs and bet-hedging in a freshwater fish. *Journal of Animal Ecology*, **81**, 806–817.

Morrongiello JR, Bond NR, Crook DA, Wong B (2013) Intraspecific variation in the growth and survival of juvenile fish exposed to *Eucalyptus* leachate. *Ecology and Evolution*, **3**, 3855–3867.

Oksanen J, Blanchet FG, Kindt R *et al.* (2015) *Package 'vegan'*. Community ecology package, version, 2.3–0.

Olofsson H, Ripa J, Jonzén N (2009) Bet-hedging as an evolutionary game: the trade-off between egg size and number. *Proceedings of the Royal Society of London B: Biological Sciences*, **276**, 2963–2969.

Osborne MJ, Perkin JS, Gido KB, Turner TF (2014) Comparative riverscape genetics reveals reservoirs of genetic diversity for conservation and restoration of Great Plains fishes. *Molecular Ecology*, **23**, 5663–5679.

Palmer MA, Reidy Liermann CA, Nilsson C *et al.* (2008) Climate change and the world's river basins: anticipating management options. *Frontiers in Ecology and the Environment*, **6**, 81–89.

Pauls SU, Nowak C, Bálint M, Pfenninger M (2013) The impact of global climate change on genetic diversity within populations and species. *Molecular Ecology*, **22**, 925–946.

Paz-Vinas I, Blanchet S (2015) Dendritic connectivity shapes spatial patterns of genetic diversity: a simulation-based study. *Journal of Evolutionary Biology*, **28**, 986–994.

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

Pearce L (2015) *Surveys, Monitoring and Conservation Status of Southern Pygmy Perch (Nannoperca australis) within Blakney and Pudman Creeks*. NSW Department of Primary Industries, Albury, Australia.

Pels S (1964) The present and ancestral Murray River system. *Australian Geographical Studies*, **2**, 111–119.

Perry GLW, Bond NR (2009) Spatially explicit modeling of habitat dynamics and fish population persistence in an intermittent lowland stream. *Ecological Applications*, **19**, 731–746.

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

Piry S, Luikart G, Cornuet J-M (1999) BOTTLENECK: a program for detecting recent effective population size reductions from allele data frequencies. *Journal of Heredity*, **90**, 502–503.

Pittock J, Finlayson CM (2011) Australia's Murray-Darling Basin: freshwater ecosystem conservation options in an era of climate change. *Marine and Freshwater Research*, **62**, 232–243.

Poff NL, Allan JD, Bain MB *et al.* (1997) The natural flow regime. *BioScience*, **47**, 769–784.

Pritchard JK, Di Rienzo A (2010) Adaptation–not by sweeps alone. *Nature Reviews Genetics*, **11**, 665–667.

Pritchard JK, Pickrell JK, Coop G (2010) The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Current Biology*, **20**, 208–215.

Puritz JB, Hollenbeck CM, Gold JR (2014) dDocent: a RADseq, variant-calling pipeline designed for population genomics of non-model organisms. *PeerJ*, **2**, e431.

Quintero I, Wiens JJ (2013) Rates of projected climate change dramatically exceed past rates of climatic niche evolution among vertebrate species. *Ecology Letters*, **16**, 1095–1103.

R Core Team (2015) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Raeymaekers JAM, Konijnendijk N, Larmuseau MHD *et al.* (2014) A gene with major phenotypic effects as a target for selection vs. homogenizing gene flow. *Molecular Ecology*, **23**, 162–181.

Raj A, Stephens M, Pritchard JK (2014) fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics*, **197**, 573–589.

Rellstab C, Gugerli F, Eckert AJ, Hancock AM, Holderegger R (2015) A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, **24**, 4348–4370.

Sasaki M, Hammer MP, Unmack PJ, Adams M, Beheregaray LB (2016) Population genetics of a widely distributed small freshwater fish with varying conservation concerns: the southern purple-spotted gudgeon, *Mogurnda adspersa*. *Conservation Genetics*, **17**, 875–889.

Schoville SD, Bonin A, François O *et al.* (2012) Adaptive genetic variation on the landscape: methods and cases. *Annual Review of Ecology, Evolution, and Systematics*, **43**, 23–43.

Sgrò CM, Lowe AJ, Hoffmann AA (2011) Building evolutionary resilience for conserving biodiversity under climate change. *Evolutionary Applications*, **4**, 326–337.

Shafer ABA, Wolf JBW, Alves PC *et al.* (2015a) Genomics and the challenging translation into conservation practice. *Trends in Ecology & Evolution*, **30**, 78–87.

Shafer ABA, Wolf JBW, Alves PC *et al.* (2015b) Reply to Garner *et al.*. *Trends in Ecology & Evolution*, **31**, 83–84.

Skovrind M, Olsen MT, Vieira FG *et al.* (2016) Genomic population structure of freshwater-resident and anadromous ide (*Leuciscus idus*) in north-western Europe. *Ecology and Evolution*, **6**, 1064–1074.

Sork VL, Waits L (2010) Contributions of landscape genetics—approaches, insights, and future potential. *Molecular Ecology*, **19**, 3489–3495.

Steane DA, Potts BM, McLean E *et al.* (2014) Genome-wide scans detect adaptation to aridity in a widespread forest tree species. *Molecular Ecology*, **23**, 2500–2513.

Stein JL, Stein JA, Nix HA (2002) Spatial analysis of anthropogenic river disturbance at regional and continental scales: identifying the wild rivers of Australia. *Landscape and Urban Planning*, **60**, 1–25.

Stein J, Hutchinson M, Stein J (2014) A new stream and nested catchment framework for Australia. *Hydrology and Earth System Sciences*, **18**, 1917–1933.

Stockwell CA, Hendry AP, Kinnison MT (2003) Contemporary evolution meets conservation biology. *Trends in Ecology & Evolution*, **18**, 94–101.

Strayer DL, Dudgeon D (2010) Freshwater biodiversity conservation: recent progress and future challenges. *Journal of the North American Benthological Society*, **29**, 344–358.

Tallmon DA, Luikart G, Waples RS (2004) The alluring simplicity and complex reality of genetic rescue. *Trends in Ecology & Evolution*, **19**, 489–496.

Thomaz AT, Christie MR, Knowles LL (2016) The architecture of river networks can drive the evolutionary dynamics of aquatic populations. *Evolution*, **70**, 731–739.

Thuiller W (2007) Biodiversity: climate change and the ecologist. *Nature*, **448**, 550–552.

Unmack PJ, Hammer MP, Adams M, Johnson JB, Dowling TE (2013) The role of continental shelf width in determining freshwater phylogeographic patterns in south-eastern Australian pygmy perches (Teleostei: Percichthyidae). *Molecular Ecology*, **22**, 1683–1699.

Vitousek PM, Mooney HA, Lubchenco J, Melillo JM (1997) Human domination of earth's ecosystems. *Science*, **277**, 494–499.

Walther G-R, Post E, Convey P *et al.* (2002) Ecological responses to recent climate change. *Nature*, **416**, 389–395.

Wang IJ (2013) Examining the full effects of landscape heterogeneity on spatial genetic variation: a multiple matrix regression approach for quantifying geographic and ecological isolation. *Evolution*, **67**, 3403–3411.

Waples RS, Do C (2010) Linkage disequilibrium estimates of contemporary *Ne* using highly variable genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary Applications*, **3**, 244–262.

Wedderburn S, Hammer M, Bice C (2012) Shifts in small-bodied fish assemblages resulting from drought-induced water level recession in terminating lakes of the Murray-Darling Basin, Australia. *Hydrobiologia*, **691**, 35–46.

Weeks AR, Sgro CM, Young AG *et al.* (2011) Assessing the benefits and risks of translocations in changing environments: a genetic perspective. *Evolutionary Applications*, **4**, 709–725.

Weeks AR, Stoklosa J, Hoffmann AA (2016) Conservation of genetic uniqueness of populations may increase extinction likelihood of endangered species: the case of Australian mammals. *Frontiers in Zoology*, **13**, 1.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Weir BS, Hill W (2002) Estimating F-statistics. *Annual Review of Genetics*, **36**, 721–750.

Whiteley AR, Fitzpatrick SW, Funk WC, Tallmon DA (2015) Genetic rescue to the rescue. *Trends in Ecology & Evolution*, **30**, 42–49.

Whitlock MC, Lotterhos KE (2015) Reliable detection of loci responsible for local adaptation: inference of a null model through trimming the distribution of Fst. *The American Naturalist*, **186**, S24–S36.

Willi Y, Hoffmann AA (2012) Microgeographic adaptation linked to forest fragmentation and habitat quality in the tropical fruit fly *Drosophila birchii*. *Oikos*, **121**, 1627–1637.

Willi Y, Van Buskirk J, Schmid B, Fischer M (2007) Genetic isolation of fragmented populations is exacerbated by drift and selection. *Journal of Evolutionary Biology*, **20**, 534–542.

Wood JLA, Belmar-Lucero S, Hutchings JA, Fraser DJ (2014) Relationship of habitat variability to population size in a stream fish. *Ecological Applications*, **24**, 1085–1100.

Wood JLA, Yates MC, Fraser DJ (2016) Are heritability and selection related to population size in nature? Meta-analysis and conservation implications. *Evolutionary Applications*, **9**, 640–657.

Yeomans KA, Golder PA (1982) The Guttman-Kaiser criterion as a predictor of the number of common factors. *The Statistician*, **31**, 221–229.

L.B.B. designed the study, with input from C.J.B. C.J.B. generated and analysed the data with assistance from L.B.B. and M.P.H. C.J.B. and L.B.B. lead the writing of the manuscript, with input from M.P.H.

Data accessibility

Reference sequences, SNP genotypes, BLAST results, annotations, sample coordinates and environmental data used in analyses are available on Dryad: DRYAD entry doi:10.5061/dryad.3dp50.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Environmental PCAs describing the relationship between each *Nannoperca australis* sampling location based on variables related to (a) flow, (b) human disturbance and (c) topography.

Fig. S2 Locus scores for the first three axes (a, b, and c respectively) of the partial redundancy analysis.

Fig. S3 RDA triplots with (a) RDA1 vs. RDA3, and (b) RDA2 vs. RDA3.

Table S1 Details of environmental variables considered for landscape genomics analysis of *Nannoperca australis* from the Murray-Darling Basin (MDB).

Table S2 Illumina HiSeq2000 sequencing statistics for 263 *Nannoperca australis* ddRAD libraries.

Table S3 Effective population size estimates (N_E) and *P*-values from bottleneck test for excess heterozygosity.

Table S4 Pairwise population F_{ST} among *Nannoperca australis* sampling sites (F_{ST} below and *P* values above the diagonal).

Table S5 Hierarchical analysis of molecular variance (AMOVA) based on neutral (3443 SNPs) F_{ST} for all samples and those catchments containing multiple sampling sites for *Nannoperca australis* from the Murray-Darling Basin (MDB).

Table S6 Variation explained by retained environmental principal components and the correlation between the original variables and each component.

Table S7 Number of candidate loci identified for each environmental variable using gINLAnd (log-Bayes factor >15).

Appendix S1 Supplementary methods.