MOLECULAR ECOLOGY

Molecular Ecology (2013) **22**, 1820–1835 doi: 10.1111/mec.12227

# 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)

R. A. COLEMAN,*† A. R. WEEKS‡ and A. A. HOFFMANN*

*Victorian Centre for Aquatic Pollution Identification and Management, Bio21 Institute, Department of Zoology, University of Melbourne, Parkville,Vic.,3010, Australia, †Melbourne Water Corporation, 990 La Trobe Street, Docklands, Vic.,3008, Australia, ‡Bio21 Institute, Department of Genetics, University of Melbourne, Parkville,Vic.,3010, Australia

## Abstract

**Genetic markers are widely used to define and manage populations of threatened species based on the notion that populations with unique lineages of mtDNA and well-differentiated nuclear marker frequencies should be treated separately. However, a danger of this approach is that genetic uniqueness might be emphasized at the cost of genetic diversity, which is essential for adaptation and is potentially boosted by mixing geographically separate populations. Here, we re-explore the issue of defining management units, focussing on a detailed study of *Galaxiella pusilla*, a small freshwater fish of national conservation significance in Australia. Using a combination of microsatellite and mitochondrial markers, 51 populations across the species range were surveyed for genetic structure and diversity. We found an inverse relationship between genetic differentiation and genetic diversity, highlighting a long-term risk of deliberate isolation of *G. pusilla* populations based on protection of unique lineages. Instead, we adopt a method for identifying genetic management units that takes into consideration both uniqueness and genetic variation. This produced a management framework to guide future translocation and re-introduction efforts for *G. pusilla*, which contrasted to the framework based on a more traditional approach that may overlook important genetic variation in populations.**

*Keywords*: evolutionary significant unit, freshwater fish, *Galaxiella pusilla*, genetic diversity, management unit, threatened species, translocations

*Received 18 September 2012; revision received 13 December 2012; accepted 18 December 2012*

## Introduction

Threatened species conservation programmes often focus on ecological threats and habitat protection or restoration; however, genetic data can assist conservation strategies for threatened species by providing an understanding of levels of genetic diversity, distinctiveness due to an independent evolutionary history and population interconnectivity. To assist management planning, molecular data have often been used to separate populations into 'conservation units'. Various definitions for deriving conservation management boundaries using molecular and ecological data have been proposed (e.g. Ryder 1986; Moritz 1994; Moran 2002; Hey *et al.* 2003; Palsbøll *et al.* 2007; Funk *et al.* 2012), including broad-scale evolutionary significant units (ESUs) that focus on long-term management issues such as defining priorities and setting strategies, and finer-scale management units (MUs) that focus on short-term conservation actions by defining demographically independent units (Moritz 1994; Palsbøll *et al.* 2007).

The delineation of conservation units is useful for preserving unique genetic lineages that may be important

Correspondence: Rhys Coleman, Fax: +61 39679 7499; E-mail: rhys.coleman@melbournewater.com.au

from an evolutionary perspective, but consideration must also be given to the protection and augmentation of genetic diversity needed for adaptation (Bonin *et al.* 2007; Funk *et al.* 2012). Inbreeding and loss of genetic diversity in both captive and wild populations increase extinction risk, through a reduction in fitness resulting from inbreeding depression and through a reduced potential for adaptation to changing conditions (Moran 2002; Frankham 2003; Willi *et al.* 2006). For threatened species, where there has been substantial fragmentation and population decline, there is an increasing urgency to preserve genetic diversity for adaptation particularly with emerging threats such as climate change (Weeks *et al.* 2011). Conservation strategies should ensure that enough genetic diversity remains to enable survival of a species in the short term and allow diversification in the future (Moritz 1999).

At present, guidelines for the management of threatened species often emphasize genetic uniqueness rather than genetic diversity (Funk *et al.* 2012). For instance, Australian guidelines for the translocation of organisms such as freshwater fish (Bureau of Rural Sciences 1999) focus only on the maintenance of genetic uniqueness and increasing the size of local populations. Although efforts to maintain large population sizes would be expected to help preserve genetic diversity, once genetic diversity is lost (e.g. small threatened populations), merely increasing the size of the population will not restore genetic diversity. Instead, an approach that does not consider the genetic health of populations may lead to a loss of adaptability in populations—particularly where there is an overemphasis on the conservation of remnant populations that have reduced viability because of inbreeding depression, and when populations face rapidly changing environments requiring adaptive responses (Moritz 1999).

One approach that balances genetic diversity against uniqueness has been developed by Caballero & Toro (2002) for the conservation of breeding lineages. In their approach, genetic uniqueness is balanced against genetic diversity within populations through assessing the relative contribution of specific populations to diversity across the entire species range. Therefore, a small but unique population lacking in diversity might be ranked lower in conservation importance than a less distinctive population with high diversity. This approach is likely to be particularly useful if there has been an erosion of genetic diversity from random processes within a series of populations; for instance, genetically unique populations may arise from drift (bottlenecks, small populations) but lack diversity while populations less affected by drift and containing high diversity may have more conservation value (Barker 2001).

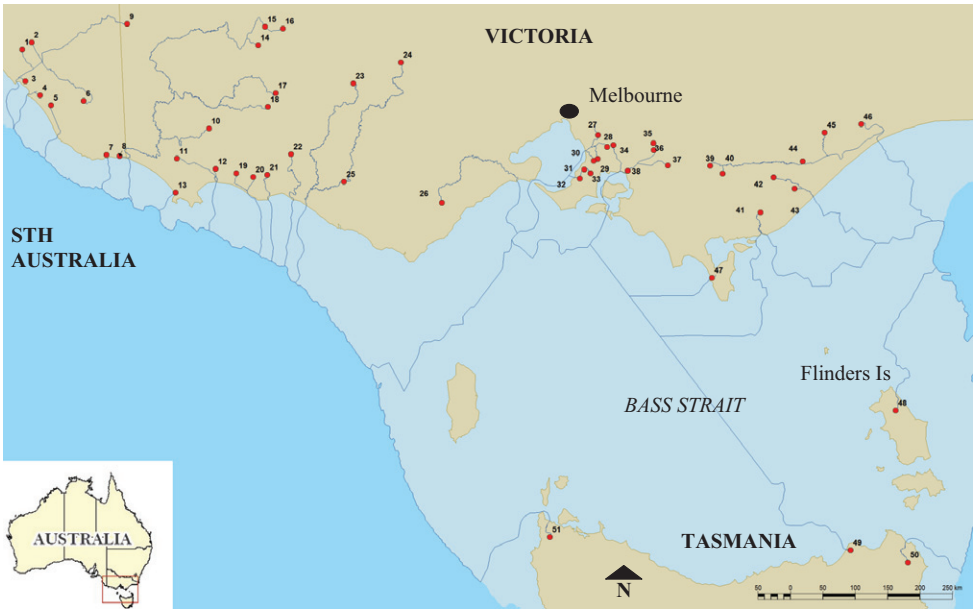
Here, we consider this approach and others for the management of the dwarf galaxias, *Galaxiella pusilla*, a small (usually <40 mm length) freshwater-dependent fish endemic to southeastern Australia, including southeast South Australia, southern Victoria and northern Tasmania. This species is found in shallow, still or slow flowing water, with dense submerged and/or emergent aquatic vegetation. *Galaxiella pusilla* is of national conservation significance, with formal listing as a threatened species. The 2008 IUCN Red List of Threatened Species classes *G. pusilla* as 'vulnerable' (Wager 1996). Habitat loss, water pollution and interaction with alien fishes (especially *Gambusia holbrooki*) are key threats (e.g. Koster 2003; Saddlier *et al.* 2010). A lack of diadromy and tendency to occur in freshwater habitats hydrologically isolated for extended periods also contribute to vulnerability, especially during times of drought when habitat may dry out and opportunities for dispersal and recruitment are reduced (Coleman *et al.* 2010).

An understanding of the genetic structure and diversity of dwarf galaxias populations is a priority within the dwarf galaxias National Recovery Plan (Saddlier *et al.* 2010). Recently, Coleman *et al.* (2010) isolated 11 microsatellite loci [along with mitochondrial cytochrome *c* oxidase subunit I (COI) sequences] to describe the broad-scale genetic structure of dwarf galaxias, as well as identifying substantial genetic differentiation between populations from western (South Australia and Victoria west of, and including, the Otway Ranges) and an eastern (Victoria east of the Otway Ranges and Tasmania) geographic regions. Using multiple mitochondrial markers with various rates of mutation [COI, cytochrome *b* (CytB), 16S rDNA (16S)] and the microsatellite loci isolated by Coleman *et al.* (2010), the present study provides a comprehensive analysis of *G. pusilla* across its entire range to determine genetic structure and diversity for guiding protection and translocation efforts. We discuss priority populations for conservation and highlight the importance of considering levels of genetic variation in populations as well as genetic uniqueness.

## Materials and methods

### *Sample collection and DNA extraction*

Caudal fin clips (<1/3 of the total fin area) were collected from up to 30 individuals from 50 populations across southeastern Australia (Fig. 1) following the method proposed by Coleman *et al.* (2010). No contemporary samples could be collected from the Harcus River site (population 51) in northwest Tasmania, so historical samples from 10 individuals caught in 2005 were used. Sampling locations were chosen to represent




* Where populations are: 1=Bray Drain, 2=Reedy Creek Wilmont Drain, 3=Death Hole Outlet Drain, 4=Bevilaqua Drain, 5=Lake Bonney Drain, 6=The Claypans,7=Ewans Ponds, 8=Pick's Swamp, 9=Mosquito Creek, 10=Crawford River, 11=Fitzroy River, 12=Darlot Creek Tributary, 13=Bridgewater Lakes, 14=Tea Tree Creek, 15=Glenelg River, 16=Mount Rosea Creek, 17=Wannon River, 18=Boonawah Creek, 19=Eumeralla River, 20=Shaw River, 21=Moyne River, 22=Spring Creek, 23=Fiery Creek, 24=Mt Emu Creek Upper, 25=Mt Emu Creek Lower, 26=Gosling Creek, 27=Tributary of Dandenong Creek, 28=Golf Links Road Drain, 29=Eastern Contour Drain Tributary, 30=Boggy Creek, 31=Balcombe Creek, 32=Tuerong Creek, 33=Blackscamp Wetland, 34=Cardinia Creek Tributary, 35=Dingo Creek Tributary, 36=Cannibal Creek, 37=King Parrot Creek, 38=Yallock Creek, 39=Moe Contour Drain, 40=Morwell River, 41=Deep Creek, 42=Merriman Creek, 43=Monkey Creek, 44=Flooding Creek, 45=Perry River, 46=Cobblers Creek, 47=Darby River , 48=Flinders Island, 49=Big Waterhouse Lake, 50=Icena Creek, 51=Harcus River

**Fig. 1** Location of sites in southeastern Australia where *Galaxiella pusilla* were collected. Waterway alignments represent minimum river distance pathways, including palaeodrainages beyond the current coastline mapped by Harris *et al.* (2005).

the diversity of habitats and geographic areas across the entire species range. Although many of these populations were from single locations along river systems, multiple populations were collected from some systems to examine gene flow along rivers systems as well as between adjacent systems.

Sampling was conducted between May 2009 and May 2010, with fish captured using a combination of dip nets and nonbaited fine mesh bait traps, targeting shallow (<1 m) vegetated margins of habitats. Dip netting was usually sufficient to capture 30 individuals; if not, bait traps were set overnight and placed on substrate with part of the trap above water. In addition to the Harcus River site, two other sites required historical samples (Shaw River and Cobblers Creek samples collected in 2007–2008) to increase sample sizes to at least 20 per population. Samples were preserved in 0.5-mL microcentrifuge tubes filled will 100% ethanol and stored in the laboratory at −20 °C. DNA extractions from fin clips for PCRs were performed according to the method of Coleman *et al.* (2010).

### *Microsatellite PCR*

Four universal primers labelled with unique fluorescent tags were combined with multiple fluorophores to co-amplify microsatellite loci in multiplex PCRs following Blacket *et al.* (2012). Ten microsatellite loci from Coleman *et al.* (2010) with primers redesigned for GPW15 and GPW18 were used to increase the length of products, so that all 10 loci could be combined into two PCR multiplexes, that is, multiplex 1 = GPE04, GPE05, GPE08, GPE19, GPW15A, GPW18B and multiplex 2 = GPE12, GPE13, GPW01, GPW02. The new primers for locus GPW15A (without the forward primer fluorescent tag and reverse primer 'pigtail') were F: ATGCCTGA-ACTAGACCTTCC and R: AGGACAGATATAAAA GGTCCAC, and the new primers for locus GPW18B (without the forward primer fluorescent tag and reverse primer 'pigtail') GPW18B were F: GGCGCATTACAG-CAAAT and R: TCTAGCATACACTTACACACAC. Reaction mixes and PCR conditions were the same as in Coleman *et al.* (2010) with fragment analyses undertaken at AGRF laboratories (Melbourne, Australia) on an Applied Biosystems (Foster City, CA, USA) ABI3730 DNA analyser with a LIZ-500 size standard.

### *Microsatellite data analysis*

Allele sizes were assessed and scored with GENEMAPPER Version 4.0 (Applied Biosystems). Mean observed heterozygosity, mean expected heterozygosity and mean number of alleles per locus were calculated with GENALEX Version 6.3 (Peakall & Smouse 2006). Allelic richness

averaged over loci, *F* statistics ($F_{IS}$ and $F_{ST}$) and tests for linkage disequilibrium based on a log-likelihood ratio test were calculated with FSTAT Version 2.9.3 (Goudet 2001). An alternative measure of genetic differentiation ($D_{est}$) that more accurately accounts for differences in allelic diversity than $F_{ST}$ was calculated using SMOGD Version 1.2.5 (Jost 2008; Crawford 2010). Five hundred bootstrap replicates were used to estimate 95% confidence intervals for $D_{est}$ values. However, because $D_{est}$ values did not substantially alter the outcomes of the analyses (compared to $F_{ST}$), they have not been presented. To determine the genetic differentiation of a given population relative to all other populations in each region (i.e. 'west' and 'east' genetically distinct regions described by Coleman *et al.* (2010), also see results where west = populations 1–26 and east = populations 27–51), population-specific $F_{ST}$ were estimated using the hierarchical Bayesian framework in GESTE Version 2.0 and default settings of sample size = 10 000, thinning interval = 20, 10 pilot runs of 5000 length and an additional burn in = 50 000 (Foll & Gaggiotti 2006). Two-sided *P* values after 9999 permutations computed in FSTAT tested for differences in heterozygosity and allelic richness between the east and west regions. Deviations from Hardy–Weinberg equilibrium (HWE) were determined by exact tests and significance determined through permutation in GDA Version 1.1 (Lewis & Zaykin 2001). An estimate of the null allele frequency for each locus across all populations was performed in GENELAND (Guillot *et al.* 2005), and evidence for null alleles and large allele dropout for each locus within each population were assessed using MICRO-CHECKER Version 2.2.3 (Van Oosterhout *et al.* 2004).

To identify levels of genetic subdivision, a hierarchical analysis of molecular variance (AMOVA) was performed within ARLEQUIN Version 3.1 (Excoffier *et al.* 2006). Microsatellite data were partitioned to enable a comparison of variation among regions, among populations within regions and within populations. Pairwise $F_{ST}$ values were calculated by ARLEQUIN and used as the distance measure with significance determined by permutation (10 000 replicates).

To test for isolation by distance, Slatkin's (1995) linearized $F_{ST}$ transformation ($F_{ST}/1 - F_{ST}$) was regressed onto the natural log of geographic distance (Rousset 1997). The significance of this relationship was determined with a Mantel test (10 000 permutations) in POPTOOLS Version 3.1 (Hood 2002). A pairwise minimum 'river distance' matrix between sites was also calculated in MAPINFO PROFESSIONAL Version 10.0.3 using contemporary drainage lines at a scale of 1:250 000 (Geoscience Australia) and a palaeo-drainage map for the southeast continental margin of Australia (Harris *et al.* 2005). The natural log of river distance was also regressed against Slatkin's linearized $F_{ST}$. Partial Mantel tests with linearized $F_{ST}$ values, geographic distances and river distances across all populations and each region separately were performed using IBDWS Version 3.23 (Jensen *et al.* 2005). Regression analysis was performed with SPSS Version 20. A stronger relationship between genetic differentiation and 'river distance' is expected to indicate that historical dispersal within this species has been largely confined to movement along existing river alignments (including associated palaeo-drainages during periods of much lower sea levels), while a stronger relationship between genetic differentiation and 'geographic distance' is expected to indicate the importance of overland dispersal as well (e.g. connections between river systems or swamps during large floods, significant events that change drainage boundaries such as lava flows or mass channel erosion).

Bayesian clustering was applied to examine population genetic structure within regions without allocating individuals to populations prior to analysis; GENELAND (Guillot *et al.* 2005) was used for this purpose. Three runs of 100 000 iterations examined, with a *K* maximum = 25, thinning = 100, burn in = 200 and allele frequencies uncorrelated.

To determine whether populations were likely to have experienced a recent bottleneck (>100 generations), we manually calculated *M* ratios (mean ratio of the number of alleles to the range in allele size for each polymorphic locus) for each population (Garza & Williamson 2001). CRITICAL M (Garza & Williamson 2001) was then run to estimate the significance of *M* ratios based on the calculation of $M_c$. CRITICAL M parameters employed included the fraction of mutations larger than single steps = 0.1 ($p_S$ = 90%), average size of non-one-step mutations ($\Delta_g$) = 3.5 and prebottleneck effective population size ($N_e$) = 5000 as recommended by Garza & Williamson (2001). The mutation rate used for CRITICAL M was $5.56 \times 10^{-4}$ per locus per generation derived from a fish (*Cyprinus carpio*) (Yue *et al.* 2007), similar to the recommended default mutation rate of $5.0 \times 10^{-4}$ per locus per generation, but considered more appropriate for this study. To visualize the genetic structure of individuals over multiple loci, a multivariate factorial correspondence analysis was undertaken with GENETIX Version 4.05 (Belkhir *et al.* 2004) and the two factors that explained most of the differentiation were plotted. All tests involving multiple comparisons were corrected at the table-wide $\alpha' = 0.05$ level (Sokal & Rohlf 1995).

### *mtDNA PCR*

Ten individuals from each of the 51 populations were used for mtDNA sequencing of 16S rDNA (16S), COI and cytochrome *b* (CytB) fragments. Historical samples

from Coleman *et al.* (2010) were used for mtDNA sequencing where possible. Reactions were adjusted to a final volume of 30 μL with sterile $H_2O$ and contained 1× polymerase reaction buffer, 2 mM $MgCl_2$, 0.2 mM dNTPs, 0.03 U Immolase DNA polymerase (Bioline, Alexandria, Australia), 0.5 μM of each primer and 5 μL of template DNA. Primers for amplifying a 595-bp region of the COI gene were FishF1 and FishR1 as described in Ward *et al.* (2005) and PCR conditions were from Coleman *et al.* (2010). CytB primers for a 1083-bp fragment were PreLA (CCAGGACTAATGGCTTGAAAAA) and GPThr27 (TCTTCGGATTACAAAACCG) with an initial 10-min denaturing step at 95 °C and then 35 cycles of denaturation at 94 °C (30 s), annealing at 48 °C (45 s) and extension at 72 °C (90 s), with a final extension step of 72 °C for 5 min preceding an indefinite hold period at 4 °C (P. Unmack pers. comm.). The universal 16S primers 16Sar (LR-N-13398) and 16Sbr (LR-J-12887) (Palumbi 1996) amplified a 482-bp fragment with the PCR conditions being an initial 7-min denaturing step at 95 °C and then 40 cycles of denaturation at 95 °C (20 s), annealing at 53 °C (45 s) and extension at 72 °C (30 s), with a final indefinite hold period at 4 °C. PCR amplification was carried out with an Eppendorf Mastercycler (Eppendorf, Hamburg, Germany), and PCR products were sent to Macrogen laboratories (Seoul, Korea) for sequencing on an ABI3730 XL DNA sequencer (Applied Biosystems). All products were sequenced in both directions, and primers from the initial PCR were used for sequencing.

### *mtDNA data analysis*

Forward and reverse sequences were aligned and manually edited in SEQUENCHER Version 4.6 (Genecodes, Ann Arbor, MI, USA). Consensus sequences were then imported into MEGA Version 4 (Tamura *et al.* 2007) for multiple alignments with CLUSTALW. Alignments were performed with the default parameters. Haplotypes for the three mtDNA markers individually, and when combined, were determined with GENALEX.

Genetic structure between haplotypes was assessed by constructing phylogenetic trees using two contrasting methods: Bayesian inference (BI) and maximum parsimony. Sequence lengths to construct the 16S, COI and CytB trees were 482 bp, 595 and 1083 bp, respectively. A congener, the black-stripe minnow *Galaxiella nigrostriata*, and the common galaxias *Galaxias maculatus* (Galaxiidae) were included as outgroups (GenBank accession nos were NC008448 and NC004594, respectively). Genetic differentiation between haplotypes was also examined by calculating pairwise genetic distances under a Kimura 2-parameter model within MEGA.

MODELTEST Version 3.7 (Posada & Crandall 1998) identified the optimum nucleotide substitution model for each of the data sets that were then used for the construction of BI phylogenetic trees within MRBAYES Version 3.1.2 (Ronquist & Huelsenbeck 2003). The best-fit models of evolution for the Bayesian analyses were 16S TVM+I, COI TrN+I+G and for CytB GTR+I+G, and TIM+I+G when the three sequences were combined into a single analysis. Further methods for BI analyses can be found in the Supporting information.

Similar to Coleman *et al.* (2010), demographic history was explored using DNASP version 4.50.3 (Rozas *et al.* 2003). Signatures of recent bottleneck events were examined by assessing deviations from neutrality with Tajima's *D* statistic (Tajima 1989). Evidence for population range expansion was assessed with mismatch analysis (Rogers *et al.* 1996) of haplotypes with the raggedness index determined using coalescent simulations with 1000 replicates (Harpending *et al.* 1993). Evidence for population growth was also assessed using coalescent simulations with 1000 replicates to calculate Fu and Li's $F^*$ and $D^*$ (Fu & Li 1993), and Fu's $F_s$ (Fu 1997) statistics.

### *Genetic diversity and uniqueness*

To assess the status of each population according to both genetic diversity and genetic uniqueness, we undertook two types of analyses. First, we determined whether there was a negative relationship between genetic differentiation and genetic diversity using linear regression for the east and west regions independently (given the substantial genetic differentiation between regions). Expected heterozygosity and mean number of alleles (genetic diversity) were regressed against the mean population-specific $F_{ST}$ estimates for each population (genetic uniqueness), as we expected that these $F_{ST}$ values would be driving the relationship with genetic diversity. In addition to population-specific $F_{ST}$, we regressed the mean pairwise $F_{ST}$ for each population (based on each pairwise comparison for populations within the same region only)—giving similar, but slightly weaker, relationships that are not presented. Similar to the microsatellite data, mtDNA genetic uniqueness (mean Kimura 2-parameter distance measure for populations, averaging across haplotypes where there were multiple haplotypes within populations), were also regressed against expected heterozygosity and mean number of alleles. Second, we ran METAPOP Version 2.0 (Pérez-Figueroa *et al.* 2009) that considers both genetic uniqueness and genetic diversity by assessing the relative contribution of specific populations to diversity across all populations sampled using the nuclear data. This approach provides a statistically

robust way of identifying which populations contribute most to the overall diversity based on both factors.

## Results

### *Microsatellite statistics*

When each population was analysed separately, MICROCHECKER found no large allele dropouts, while the incidence of null alleles for each locus was generally low, ranging from being absent in all populations (GPE12) to present in 12% of the 51 populations (GPE08). Similarly, GENELAND estimates of the null allele frequency for individual loci across all populations within regions were low, with a mean of 0.060 (SE 0.007) in the west and 0.080 (0.005) in the east; when both regions were considered, the incidence of null alleles was <10% for all loci. In tests of genotypic linkage disequilibrium, we found no significant associations between pairs of loci across all populations after correcting for multiple comparisons.

Genetic diversity varied substantially among populations (Tables 1 and 2). Mean number of alleles and allelic richness ranged from $a = 1.20$ and $r = 1.18$ (Flinders Island) to $a = 8.60$ and $r = 6.06$ (Darlot Creek). Both measures of diversity tended to be lower in the east region; the mean number of alleles was 5.72 (SE 0.38) and allelic richness was 4.34 (0.23) in the west region, compared to values of 2.51 (0.16) and 2.14 (0.10) in the east. Expected heterozygosity ($H_E$) ranged from 0.03 (Flinders Island) to 0.62 (Darlot Creek). Across all populations in the west region, mean $H_E$ was 0.50 (SE 0.02), while in the east region it was 0.25 (0.02). Differences in allelic richness

**Table 1** Statistics for *Galaxiella pusilla* populations in the west region screened with 10 microsatellite loci. The number of individuals genotyped for each population is indicated ($n$). Mean values are indicated for number of alleles ($a$), allelic richness ($r$) where $n > 10$, observed ($H_O$) and expected ($H_E$) heterozygosities, multilocus estimates of $F_{IS}$, Hardy–Weinberg equilibrium $P$ values, GENELAND (GL) groupings, $M$ ratios, mean population-specific $F_{ST}$ and corresponding 95% highest probability density interval (HPDI)

| Pop No. | Pop Code | Waterbody | $n$ | $a$ | $r$ | $H_O$ | $H_E$ | $F_{IS}$* | HW-$P$† | GL $K = 14$ | $M$‡ | $F_{ST}$ | 95% HPDI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | BRD | Bray Drain | 30 | 6.8 | 5.10 | 0.517 | 0.526 | 0.034 | 0.149 | 1 | 0.764 | 0.117 | 0.080–0.158 |
| 2 | RCW | Reedy Creek Wilmont Drain | 30 | 7.6 | 5.40 | 0.527 | 0.547 | 0.055 | 0.041 | 1 | 0.755 | 0.084 | 0.055–0.114 |
| 3 | DEH | Death Hole Outlet Drain | 30 | 6.9 | 4.99 | 0.476 | 0.498 | 0.060 | 0.026 | 1 | 0.837 | 0.110 | 0.076–0.149 |
| 4 | BEV | Bevilaqua Drain | 30 | 8.2 | 5.91 | 0.523 | 0.566 | 0.093 | 0.001 | 1 | 0.745 | 0.064 | 0.041–0.089 |
| 5 | LBD | Lake Bonney Drain | 30 | 7.8 | 5.37 | 0.557 | 0.541 | −0.011 | 0.635 | 1 | 0.821 | 0.083 | 0.055–0.111 |
| 6 | CLP | The Claypans | 30 | 3.1 | 2.97 | 0.433 | 0.422 | −0.011 | 0.609 | 2 | 0.676 | 0.426 | 0.316–0.547 |
| 7 | EWP | Ewans Ponds | 30 | 4.9 | 3.94 | 0.483 | 0.496 | 0.043 | 0.140 | 3 | 0.830 | 0.225 | 0.160–0.299 |
| 8 | PIS | Pick's Swamp | 30 | 5.1 | 3.62 | 0.370 | 0.415 | 0.126 | 0.001 | 3 | 0.792 | 0.233 | 0.166–0.301 |
| 9 | MOS | Mosquito Creek | 30 | 4.5 | 3.74 | 0.514 | 0.518 | 0.023 | 0.275 | 4 | 0.688 | 0.242 | 0.169–0.321 |
| 10 | CRR | Crawford River | 30 | 6.9 | 5.12 | 0.582 | 0.587 | 0.025 | 0.255 | 5 | 0.849 | 0.133 | 0.093–0.174 |
| 11 | FIT | Fitzroy River | 30 | 7.2 | 4.93 | 0.460 | 0.492 | 0.082 | 0.009 | 6 | 0.803 | 0.131 | 0.093–0.171 |
| 12 | DLT | Darlot Creek Tributary | 30 | 8.6 | 6.06 | 0.628 | 0.616 | −0.002 | 0.543 | 5 | 0.858 | 0.064 | 0.042–0.085 |
| 13 | BRL | Bridgewater Lakes | 30 | 5.3 | 3.76 | 0.413 | 0.441 | 0.079 | 0.038 | 7 | 0.816 | 0.228 | 0.168–0.298 |
| 14 | TTC | Tea Tree Creek | 30 | 8.4 | 5.87 | 0.543 | 0.573 | 0.069 | 0.010 | 8 | 0.817 | 0.063 | 0.040–0.085 |
| 15 | GLR | Glenelg River | 30 | 7.6 | 5.51 | 0.516 | 0.582 | **0.130** | **<0.001** | 8 | 0.886 | 0.096 | 0.067–0.130 |
| 16 | MRC | Mount Rosea Creek | 30 | 6.0 | 4.64 | 0.497 | 0.535 | 0.088 | 0.007 | 8 | 0.695 | 0.162 | 0.112–0.212 |
| 17 | WAR | Wannon River | 30 | 5.7 | 4.51 | 0.500 | 0.521 | 0.057 | 0.059 | 9 | 0.727 | 0.188 | 0.136–0.248 |
| 18 | BOO | Boonawah Creek | 30 | 2.9 | 2.72 | 0.470 | 0.456 | −0.014 | 0.607 | 10 | 0.755 | 0.488 | 0.368–0.606 |
| 19 | EUR | Eumeralla River | 30 | 7.4 | 5.39 | 0.557 | 0.577 | 0.052 | 0.049 | 11 | 0.834 | 0.094 | 0.064–0.128 |
| 20 | SHR | Shaw River | 20 | 4.7 | 4.15 | 0.505 | 0.528 | 0.068 | 0.074 | 11 | 0.770 | 0.226 | 0.156–0.302 |
| 21 | MOY | Moyne River | 30 | 5.1 | 4.13 | 0.543 | 0.561 | 0.048 | 0.091 | 11 | 0.810 | 0.223 | 0.158–0.292 |
| 22 | SPR | Spring Creek | 30 | 5.5 | 4.34 | 0.520 | 0.548 | 0.069 | 0.034 | 11 | 0.848 | 0.206 | 0.149–0.269 |
| 23 | FIC | Fiery Creek | 30 | 3.0 | 2.66 | 0.423 | 0.396 | −0.053 | 0.851 | 12 | 0.711 | 0.458 | 0.339–0.572 |
| 24 | MEU | Mt Emu Creek Upper | 30 | 2.4 | 2.19 | 0.340 | 0.315 | −0.063 | 0.846 | 13 | **0.610** | 0.610 | 0.488–0.731 |
| 25 | MEL | Mt Emu Creek Lower | 30 | 5.3 | 4.19 | 0.503 | 0.516 | 0.042 | 0.127 | 13 | 0.815 | 0.237 | 0.173–0.306 |
| 26 | GOC | Gosling Creek | 30 | 1.7 | 1.55 | 0.173 | 0.172 | 0.006 | 0.721 | 14 | **0.658** | 0.743 | 0.615–0.855 |

*Significant $F_{IS}$ $P$ values after corrections for multiple comparisons are in bold.
†Significant Hardy–Weinberg equilibrium $P$ values after corrections for multiple comparisons are in bold.
‡Significant $M$ values (i.e. <critical $M$ or '$M_c$' of 0.675) are in bold.

**Table 2** Statistics for *Galaxiella pusilla* populations in the east region screened with 10 microsatellite loci. The number of individuals genotyped for each population is indicated (*n*). Mean values are indicated for number of alleles (*a*), allelic richness (*r*) where $n > 10$, observed ($H_O$) and expected ($H_E$) heterozygosities, multilocus estimates of $F_{IS}$, Hardy–Weinberg equilibrium *P* values, GENELAND (GL) groupings, *M* ratios, mean population-specific $F_{ST}$ and corresponding 95% highest probability density interval (HPDI)

| Pop No. | Pop Code | Waterbody | *n* | *a* | *r* | $H_O$ | $H_E$ | $F_{IS}$* | HW-*P*† | GL *K* = 12 | *M*‡ | $F_{ST}$ | 95% HPDI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 27 | TDN | Tributary of Dandenong Creek | 30 | 2.9 | 2.49 | 0.303 | 0.286 | −0.042 | 0.726 | 1 | 0.710 | 0.429 | 0.306–0.554 |
| 28 | GLD | Golf Links Road Drain | 30 | 2.8 | 2.33 | 0.280 | 0.284 | 0.031 | 0.310 | 1 | 0.788 | 0.452 | 0.331–0.584 |
| 29 | ECD | Eastern Contour Drain Tributary | 30 | 1.9 | 1.82 | 0.307 | 0.272 | −0.110 | 0.967 | 1 | 0.729 | 0.638 | 0.489–0.778 |
| 30 | BOC | Boggy Creek | 30 | 2.0 | 1.80 | 0.210 | 0.200 | −0.032 | 0.757 | 1 | 0.694 | 0.608 | 0.467–0.752 |
| 31 | BAC | Balcombe Creek | 30 | 2.9 | 2.62 | 0.380 | 0.367 | −0.017 | 0.625 | 2 | **0.608** | 0.379 | 0.261–0.509 |
| 32 | TUC | Tuerong Creek | 30 | 3.0 | 2.46 | 0.253 | 0.267 | 0.068 | 0.026 | 2 | **0.569** | 0.408 | 0.293–0.535 |
| 33 | BCW | Blackscamp Wetland | 30 | 1.5 | 1.43 | 0.180 | 0.162 | −0.095 | 0.215 | 2 | **0.610** | 0.757 | 0.623–0.884 |
| 34 | CAR | Cardinia Creek Tributary | 30 | 2.0 | 1.92 | 0.200 | 0.233 | 0.157 | 0.007 | 3 | **0.625** | 0.618 | 0.479–0.762 |
| 35 | DIN | Dingo Creek Tributary | 30 | 2.2 | 2.02 | 0.203 | 0.214 | 0.065 | 0.102 | 3 | 0.830 | 0.567 | 0.424–0.709 |
| 36 | CAN | Cannibal Creek | 30 | 4.1 | 2.91 | 0.313 | 0.293 | −0.051 | 0.866 | 3 | 0.762 | 0.267 | 0.180–0.355 |
| 37 | KPC | King Parrot Creek | 30 | 2.0 | 1.86 | 0.193 | 0.187 | −0.018 | 0.773 | 3 | 0.792 | 0.630 | 0.493–0.768 |
| 38 | YAC | Yallock Creek | 30 | 4.0 | 2.99 | 0.287 | 0.291 | 0.031 | 0.274 | 3 | 0.762 | 0.268 | 0.185–0.366 |
| 39 | MOE | Moe Contour Drain | 30 | 2.3 | 1.99 | 0.250 | 0.251 | 0.021 | 0.417 | 4 | **0.579** | 0.583 | 0.449–0.715 |
| 40 | MOR | Morwell River | 30 | 2.9 | 2.22 | 0.247 | 0.253 | 0.044 | 0.279 | 4 | 0.698 | 0.441 | 0.319–0.565 |
| 41 | DEE | Deep Creek | 30 | 2.1 | 1.90 | 0.263 | 0.235 | −0.104 | 0.790 | 5 | 0.732 | 0.672 | 0.551–0.796 |
| 42 | MER | Merriman Creek | 30 | 2.8 | 2.30 | 0.317 | 0.325 | 0.043 | 0.264 | 6 | **0.626** | 0.480 | 0.358–0.605 |
| 43 | MON | Monkey Creek | 30 | 3.9 | 3.16 | 0.333 | 0.420 | **0.223** | **<0.001** | 6 | **0.614** | 0.302 | 0.209–0.397 |
| 44 | FLC | Flooding Creek | 30 | 3.8 | 2.86 | 0.310 | 0.315 | 0.033 | 0.282 | 7 | **0.670** | 0.318 | 0.221–0.416 |
| 45 | PER | Perry River | 30 | 1.9 | 1.69 | 0.197 | 0.172 | −0.127 | 0.651 | 7 | **0.656** | 0.651 | 0.509–0.785 |
| 46 | COB | Cobblers Creek | 22 | 2.0 | 1.72 | 0.182 | 0.181 | 0.020 | 0.638 | 7 | **0.518** | 0.608 | 0.470–0.756 |
| 47 | DAR | Darby River | 30 | 1.7 | 1.59 | 0.227 | 0.224 | 0.006 | 0.397 | 8 | 0.827 | 0.719 | 0.590–0.850 |
| 48 | WIN | Flinders Island | 30 | 1.2 | 1.18 | 0.037 | 0.034 | −0.063 | 0.749 | 9 | **0.214** | 0.833 | 0.719–0.945 |
| 49 | BWL | Big Waterhouse Lake | 30 | 2.3 | 2.09 | 0.267 | 0.276 | 0.050 | 0.081 | 10 | 0.886 | 0.564 | 0.436–0.706 |
| 50 | ICC | Icena Creek | 30 | 2.4 | 2.09 | 0.270 | 0.272 | 0.023 | 0.335 | 11 | 0.726 | 0.567 | 0.437–0.700 |
| 51 | HAR | Harcus River | 10 | 2.1 | 2.10 | 0.240 | 0.212 | −0.080 | 0.695 | 12 | 0.895 | 0.558 | 0.409–0.709 |

*Significant $F_{IS}$ *P* values after corrections for multiple comparisons are in bold.
†Significant Hardy–Weinberg equilibrium *P* values after corrections for multiple comparisons are in bold.
‡Significant *M* values (i.e. <critical *M* or '$M_c$' of 0.675) are in bold.

and $H_E$ of the populations between the east and west regions were significant ($P < 0.001$) when populations were treated as data points.

Over all loci, $F_{IS}$ was significantly >0 for two populations after corrections for multiple comparisons: the Glenelg River in the west region ($F_{IS} = 0.130$) and Monkey Creek in the east (0.223). Deviations from HWE were significant for both these populations when considered across loci but not for any other population.

### *Population structure based on microsatellites*

The global estimate of $F_{ST}$ over all populations and loci was significantly different from zero and high ($F_{ST} = 0.507$, 95% confidence intervals 0.419–0.614), reflecting substantial genetic structure across the species range. Pairwise $F_{ST}$ estimates differed significantly in all population comparisons after correcting for multiple comparisons, except for between Lake Bonney Drain (LBD) and Bray Drain (BRD) and Reedy Creek Wilmont Drain (RCW) (Tables S1 and S2, Supporting information). Over all populations, pairwise $F_{ST}$ values ranged from 0.018 (Bray and Reedy Creek Wilmont Drains) to 0.896 (Flinders Island and Gosling Creek). In the west region, pairwise $F_{ST}$ values ranged from 0.018 (Bray and Reedy Creek Wilmont Drains) to 0.605 (Picks Swamp and Gosling Creek) with an overall $F_{ST}$ of 0.232 (95% confidence intervals 0.172–0.319). Within the east region, pairwise $F_{ST}$ values ranged from 0.032 (Yallock and Cannibal Creeks) to 0.811 (Flinders Island and Harcus River) with an overall $F_{ST}$ of 0.483 (95% confidence intervals 0.411–0.551).

A significant positive relationship was observed between genetic distance and geographic distance ($r = 0.521$, $P < 0.001$) and remained significant when populations within east and west regions were tested independently ($r = 0.426$, $P < 0.001$ and $r = 0.513$, $P < 0.001$, respectively). There was also a significant positive relationship between genetic distance and river distance ($r = 0.444$, $P < 0.001$) and when populations within west and east regions were tested independently ($r = 0.597$, $P < 0.001$ and $r = 0.352$, $P < 0.001$, respectively). Regression analysis identified a nonlinear relationship between genetic distance and geographic distance for all populations ($R^2 = 0.442$, $P < 0.001$) and for the west ($R^2 = 0.378$, $P < 0.001$) and east ($R^2 = 0.301$, $P < 0.001$) when considered separately. Similarly, there was a nonlinear relationship between genetic distance and river distance for all populations ($R^2 = 0.406$, $P < 0.001$) and for the west ($R^2 = 0.373$, $P < 0.001$) and east ($R^2 = 0.441$, $P < 0.001$) when considered separately. Geographic distance and river distance were strongly correlated (Pearson correlation 0.819, $P < 0.001$). The relationship for genetic distance over all populations was stronger with geographic distance ($r = 0.319$, $P < 0.001$) than with river distance ($r = 0.098$, $P = 0.094$). This was also the case in the east region [geographic distance, $r = 0.265$ ($P = 0.008$); river distance, $r = 0.066$ ($P = 0.254$)]. In the west, there was a stronger relationship between genetic distance and river distance ($r = 0.435$, $P = 0.001$) than geographic distance ($r = 0.270$, $P = 0.001$).

In an AMOVA, where populations were divided into the east and west regions, there was significant variation ($P < 0.001$) at all three (region, population, individual) hierarchical levels. Most variation was between regions (40.48%), with less variation among populations within regions (20.00%), and the remaining 39.52% within populations. Analysis of population structure using GENELAND indicated that the most likely values for $K$ were 14 in the west region (Table 1) and 12 in the east region (Table 2).

The results of the factorial correspondence analysis are shown in Fig. 2. When all samples were included in the analysis, there was a clear separation between the east and west regions along factor 1 (20.43% of variation). Factor 2 (7.27% of variation) also separates the Deep Creek samples from the other east region populations. The relationships between individuals within each region are presented in Fig. S1 (Supporting information). Within the west region, the Gosling Creek population is separated from the rest of the populations. Other groupings within the west are less clear. Within the east region, the Deep Creek population is separated from the others. There are two other main groups: (i) Gippsland and (ii) Port Phillip, Western Port, Wilsons Promontory, Flinders Island and Tasmania. Within these groups, there is some separation of Merriman and Monkey Creeks from the Gippsland group.

### *Population history based on microsatellites*

$M$ ratios ranged from 0.610 (Mt Emu Creek Upper) to 0.886 (Glenelg River) in the west region and 0.214 (Flinders Island) to 0.895 (Harcus River) in the east region. Based on a critical $M$ ratio ($M_c$) of 0.675, populations in the west region that are likely to have undergone recent bottlenecks are Mt Emu Creek Upper and Gosling Creek. $M$ ratios in the east region indicate that 12 of the 25 populations have undergone recent bottlenecks, namely Balcombe Creek, Tuerong Creek, Blackscamp Wetland, Cardina Creek Tributary, Moe Contour Drain, Morwell River, Merriman Creek, Monkey Creek, Flooding Creek, Perry River, Cobblers Creek and Flinders Island.

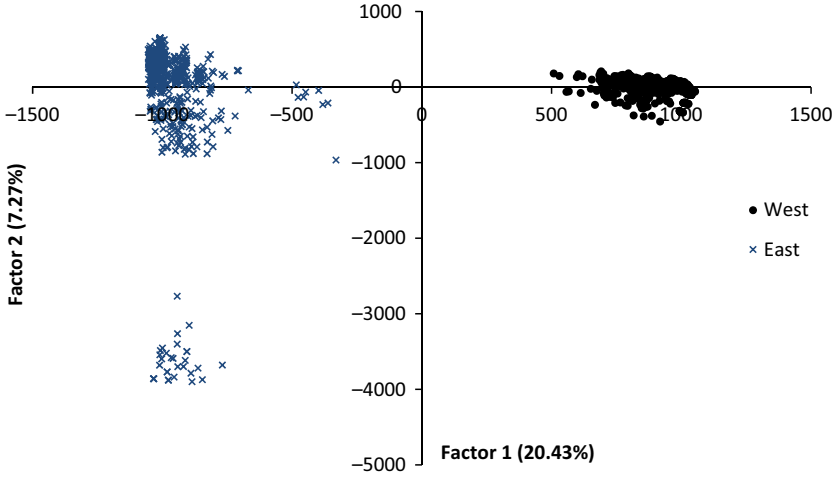



**Fig. 2** Two-dimensional plot showing the relationships among individuals of *Galaxiella pusilla* based on a multivariate factorial correspondence analysis of 10 loci from all populations. The first two factors are shown, with the percentage of variance explained in parenthesis.

### *mtDNA haplotype summary*

The west region contained 10 16S haplotypes, 36 COI haplotypes and 56 CytB haplotypes. The dominant 16S haplotype was haplotype 3 ($n$ = 186 or 72% of west individuals), absent only from Shaw, Fitzroy and Crawford Rivers, Death Hole Drain, Fiery Creek and Darlot Creek. There were two dominant COI haplotypes within the west, haplotype 6 ($n$ = 35 or 13% of west individuals) and haplotype 25 ($n$ = 48 or 18% of west individuals). Haplotype 6 was found in populations from South Australia (Bray Drain, The Claypans, Lake Bonney Drain and Reedy Creek Wilmont Drain), and haplotype 25 was found in populations from Boonawah and Spring Creeks, and Shaw, Moyne and Eumeralla Rivers. Four of the most dominant CytB haplotypes found in the west region were haplotype 43 ($n$ = 16 or 6%) from Bevilaqua Drain and Death Hole Drain in South Australia, haplotype 50 ($n$ = 21 or 8%) from the Glenelg River, Mt Rosea Creek and Ti Tree Creek in the Grampians, haplotype 56 ($n$ = 26 or 10%) from Boonawah Creek, Moyne River and Spring Creek in Western Victoria and haplotype 79 ($n$ = 14 or 5%) from Bray Drain, Lake Bonney Drain and Reedy Creek Wilmont Drain, also from South Australia.

The east region contained five 16S haplotypes, 15 COI haplotypes and 26 CytB haplotypes. The dominant 16S haplotype within the east was haplotype 13 ($n$ = 189 or 76% of east individuals), absent only from the Perry River, Tasmania, Flinders Island and Wilson Promontory. The dominant COI haplotypes within the east were haplotype 38 ($n$ = 118 or 47% of east individuals) and haplotype 50 ($n$ = 40 or 16%). Haplotype 38 consisted of populations from Port Phillip and Westernport areas and haplotype 50 consisted of the Gippsland populations: Cobblers Creek, Flooding Creek, Moe Main Drain and Morwell River. The dominant CytB haplotypes within the east were haplotype 2 ($n$ = 107 or 43%) from Port Phillip and Westernport populations and haplotype 10 ($n$ = 44 or 18%) from the Gippsland populations: Cobblers Creek, Flooding Creek, Moe Main Drain, Morwell River and Monkey Creek.

Kimura 2-parameter distances within populations from the west region varied from 0.002 to 0.011 (mean = 0.0049, SE = <0.001) for 16S, 0.002 to 0.033 (0.014, <0.001) for COI and <0.001 to 0.024 (0.014, <0.001) for CytB, while in the east region they varied from 0.002 to 0.006 (mean 0.0036, SE <0.001) for 16S, 0.002 to 0.014 (0.008, <0.001) for COI and 0.001 to 0.020 (0.010, <0.001) for CytB. Kimura 2-parameter distances between regions varied from 0.013 to 0.026 (mean = 0.0196, SE = 0.0004) for 16S, 0.073 to 0.096 (0.082, <0.001) for COI and 0.089 to 0.111 (0.101, <0.001) for CytB.

### *mtDNA genetic structure*

Results for the BI combined gene tree (Fig 3) show strong divergence between east and west regions. In the west region, the South Australian populations are clearly separated, as are the Grampians populations (including Mosquito Creek). Finer-scale groupings within the west include the Discovery Bay sites (Bridgewater Lakes, Pick's Swamp, Ewans Ponds), Port Fairy-Warrnambool (Eumeralla, Moyne, and Shaw Rivers, Boonawah and Spring Creeks) and Crawford-Fitzroy populations. Within the east region, there is a broad group of Tasmania and Wilsons Promontory populations, and a finer-scale grouping of eastern Tasmania with Flinders Island populations. Also in the east region, the Port Phillip and Western Port populations group within a broader one that includes the Gippsland populations. Individual BI trees for each of the 16S, COI and CytB genes are provided in Appendix S1 (Supporting information) and generally support the relationships found for the combined gene tree.

### *mtDNA demographic history*

Demographic history was explored using 16S, COI and CytB haplotypes (Table S3, Supporting information). Tajima's $D$ and Fu and Li's $F^*$ and $D^*$ testing for selection were nonsignificant for both the west and east region for all three mtDNA markers. Mismatch analyses for growing and declining populations produced significant raggedness index values for CytB in the west region only. Fu's $F_s$ were significant for all mtDNA markers in the west as well as COI and CytB in the east, indicating population growth in both regions. This measure has been demonstrated by Ramos-Onsins and Rozas (2002) to be more powerful than $D$, $F^*$ and $D^*$ for detecting growth.

### *Genetic diversity and uniqueness*

A comparison of genetic diversity and genetic uniqueness within populations for each region is provided in Figs 4 and 5 (microsatellites data) and Fig. S2 (Supporting information) (mtDNA data). When heterozygosity and mean number of alleles (genetic diversity) were regressed against the mean of population-specific $F_{ST}$ estimates for each population (genetic uniqueness), strong negative relationships were found in both the west (mean population-specific $F_{ST}$ and heterozygosity, $R^2$ = 0.814; mean population-specific $F_{ST}$ and mean number of alleles, $R^2$ = 0.883) and east (mean population-specific $F_{ST}$ and heterozygosity, $R^2$ = 0.644; mean population-specific $F_{ST}$ and mean number of alleles, $R^2$ = 0.953) regions independently. Using the CytB data,

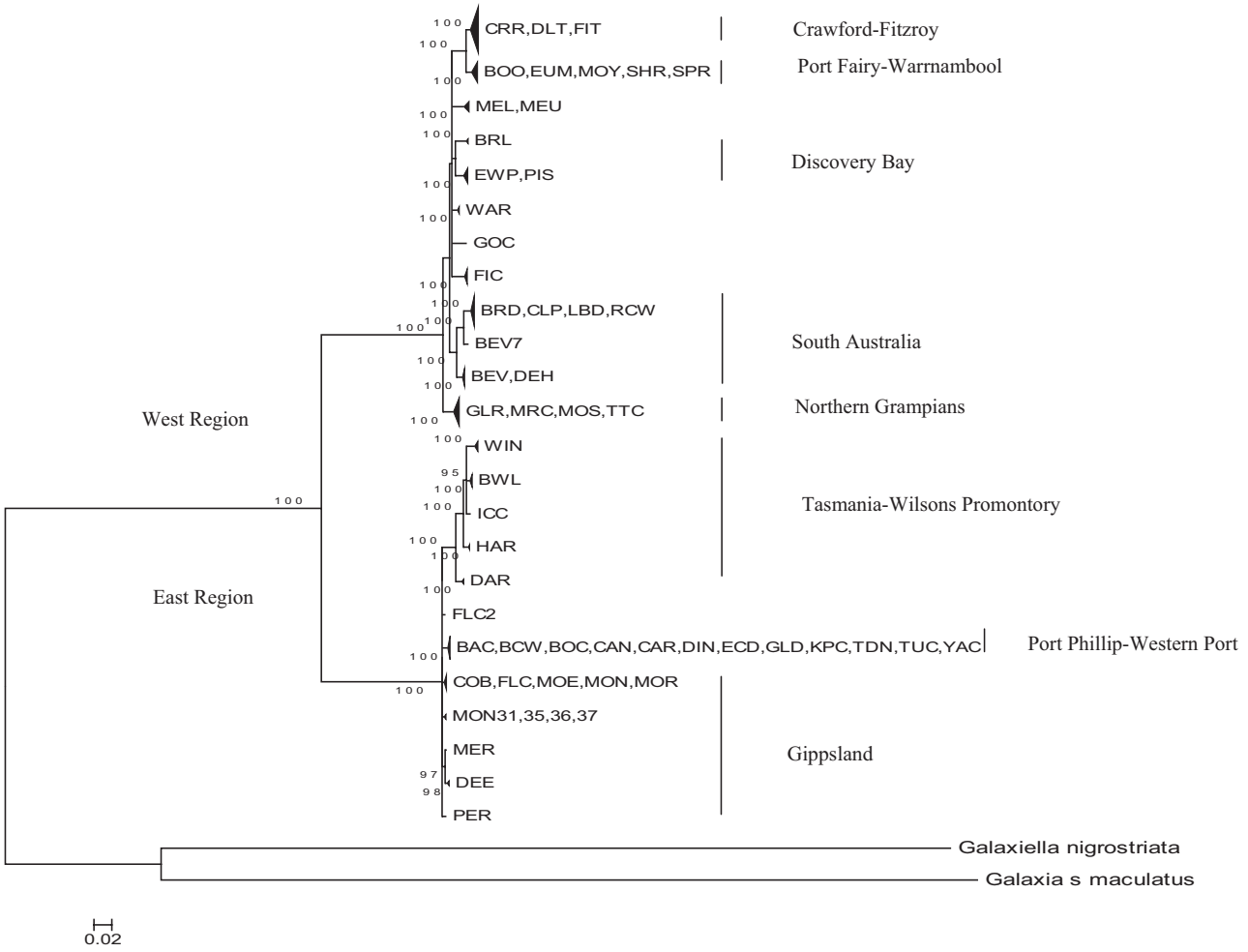



**Fig. 3** Bayesian inference phylogenetic tree for *Galaxiella pusilla* haplotypes generated from combined 16S rDNA (16S), cytochrome *c* oxidase subunit I (COI) and cytochrome *b* (CytB) fragments of 482, 595 and 1083 bp respectively. The black-stripe minnow, *Galaxiella nigrostriata*, and common galaxias, *Galaxias maculatus*, were outgroups. Branches with <0.95 Bayesian posterior probability support are collapsed.

the relationship between genetic uniqueness (mean Kimura 2-parameter distance measure for populations, averaging across haplotypes where there were multiple haplotypes within populations) and genetic diversity (heterozygosity and mean number of alleles) was weak in both regions—although the most unique populations such as Gosling Creek (GOC) and Mt Emu Creek Upper (MEU) in the west and Flinders Island (WIN) in the east had the lowest diversity.

When considering genetic uniqueness and genetic diversity by assessing the relative contribution of specific populations to diversity across all populations estimated through the METAPOP program, the above pattern was also evident. The regression of +/− allelic diversity (genetic diversity) against the mean of population-specific $F_{ST}$ estimates for each population (genetic uniqueness) resulted in a strong negative relationship for both the west ($R^2 = 0.831$) and east ($R^2 = 0.827$) regions. The same pattern was evident when +/− gene diversity between populations (genetic uniqueness) from METAPOP was regressed against +/− gene diversity within populations (genetic diversity) also estimated from METAPOP for both the west ($R^2 = 0.873$) and east ($R^2 = 0.209$) regions.

## Discussion

### *Conserving genetic diversity in Galaxiella pusilla*

An understanding of genetic structure and diversity for threatened species is integral for conservation planning. In particular, when combined with information on local adaptation (Fraser *et al.* 2011; Funk *et al.* 2012), it helps to guide management activities to protect or restore

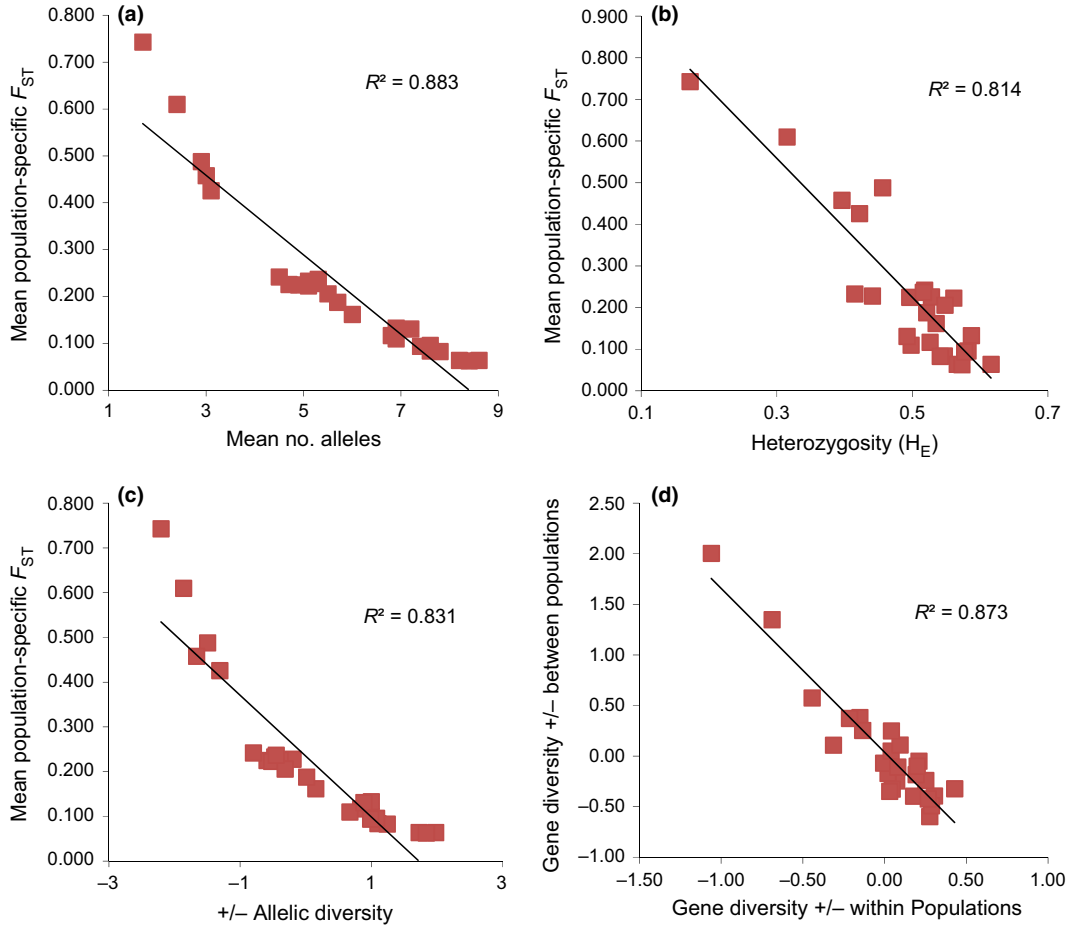



**Fig. 4** Two-dimensional plot comparing genetic diversity to genetic uniqueness for the west region. (a) Mean number of alleles, (b) heterozygosity and (c) +/− allelic diversity of each population to total allelic diversity across all populations, regressed against the mean population-specific $F_{ST}$ estimates for each population. (d) Contribution of +/− gene diversity between populations regressed against +/− gene diversity within populations.

gene flow for the purpose of fitness, adaptability and resilience to future change. Population genetic data are often used to establish management boundaries or 'conservation units' (Funk *et al.* 2012) to preserve genetic uniqueness and manage demographically independent units. Here, we show that there are some limitations with an approach that does not consider genetic diversity and over emphasizes genetic uniqueness—particularly where 'uniqueness' represents populations that have lost genetic variation rather than possessing unique alleles.

In this species, there was a strong negative relationship between genetic diversity and genetic uniqueness, showing that genetically distinct populations had lower genetic diversity and suggesting that conservation efforts focussed on preserving unique populations might decrease overall genetic diversity. Based on mtDNA monophyletic groupings, two evolutionary significant units or 'ESUs' would be identified (further supported by large microsatellite-based genetic distances), one in the west and one in the east. In addition, based on the microsatellite data (supported in some cases by the mtDNA groupings), multiple management units or 'MUs' would be identified within both ESUs. Based on GENELAND outputs, there might be 14 MUs in the west and 12 MUs in the east.

While the west and east regions are clearly unique based on the mtDNA and nuclear data and need to be managed separately, it appears that uniqueness within each region is driven largely by a loss of diversity and changes in allele frequencies rather than the presence of unique alleles. Importantly, populations with lower genetic diversity tended to contain a subset of alleles found in more genetically diverse populations. This suggests that changes in allele frequencies and reductions in genetic diversity have arisen from genetic drift, reflecting population bottlenecks/founder events and perhaps persistently small population size. In such cases, allocating low genetic diversity populations to separate MUs may not preserve genetic diversity within

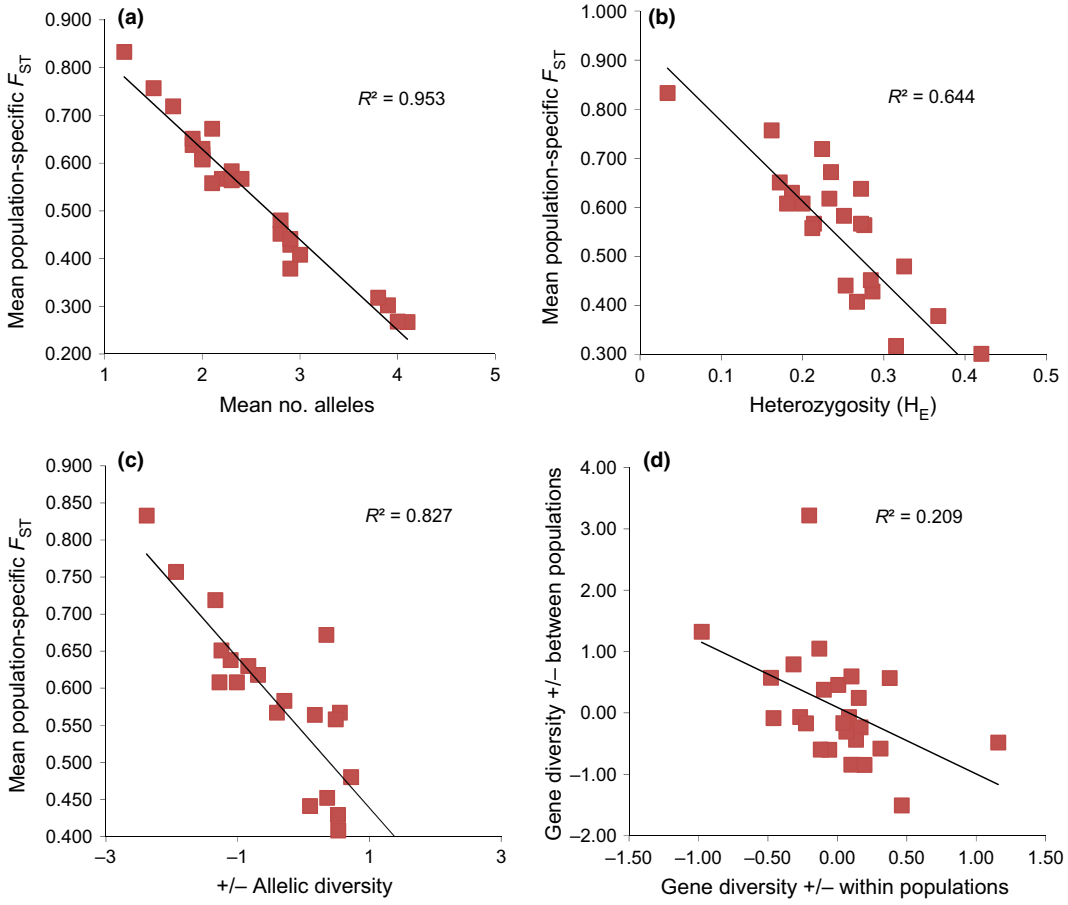



**Fig. 5** Two-dimensional plot comparing genetic diversity to uniqueness for the east region. (a) Mean number of alleles, (b) heterozygosity, (c) +/− allelic diversity of each population to total allelic diversity across all populations, regressed against the mean population-specific $F_{ST}$ estimates for each population. (d) Contribution of +/− gene diversity between populations regressed against +/− gene diversity within populations. See Table 1 for site codes.

the species as a whole. This could lead to poor investments of conservation resources, inappropriately imposed isolation of populations to preserve 'uniqueness', and in the long term a potential increased risk of local extinction. Examples of such populations based on GENELAND-derived MUs are Gosling Creek and Fiery Creek in the west, or Flinders Island and Darby River in the east region. In fact, analyses using METAPOP aimed at maximizing the conservation of genetic variation across breeding lineages (treating each population as a lineage) indicated that genetically diverse populations such as Darlot Creek in the west and Monkey Creek in the east region were a priority for conservation, and not low diversity populations such as Gosling Creek, Fiery Creek, Flinders Island and Darby River. In terms of genetic management priorities for *G. pusilla*, consideration should therefore be given to the extent that populations contribute to total diversity in the species, unless there is other and independent evidence for local adaptation.

Management units defined by genetic divergence will often be appropriate for delineating conservation units and managing threatened species. However, the *G. pusilla* situation illustrates that there can be a trade-off between genetic diversity and genetic uniqueness, and this is likely to be the case in other threatened freshwater fish and faunal groups. In such cases, it is worth checking if unique populations carry a subset of alleles from other more diverse populations and also use approaches such as outlined in Caballero & Toro (2002) to prioritize populations for maintaining genetic variation across a species. Without consideration of genetic diversity, there is a risk of developing management strategies that involve the continued isolation of 'unique populations' that perpetuate the progressive loss of genetic diversity and thereby potentially increase extinction risk as a consequence of small population size (Willi *et al.* 2006). An approach for prioritizing genetic conservation measures for species like *G. pusilla* that incorporates both genetic uniqueness and diversity

is expected to result in better long-term outcomes, particularly if mtDNA variation is not correlated with nuclear-based uniqueness (as in *G. pusilla*). In addition, conservation units can be prioritized on environmental factors and local adaptation as such information becomes available, as in the case of salmonid fish where local adaptation is common (Fraser *et al.* 2011) and where unique genotypes under selection have been identified (e.g. Miller *et al.* 2012).

Populations of *G. pusilla* with a comparatively high diversity are also likely to be good candidates as source populations for translocations to sites with low diversity that require genetic intervention (especially where there is little potential for natural gene flow such as between the Gosling Creek population (GOC) in the west and the Flinders Island population (WIN) in the east). Given substantial genetic differentiation between the west and east regions, intermixing between regions is not recommended. Within the regions, selection of source populations for translocations might be based on populations in close geographic proximity with high genetic diversity (Weeks *et al.* 2011) or populations geographically distant but exposed to similar environmental conditions (e.g. habitat, rainfall, altitude). For example, potential donor populations for the Flinders Island population could include one of the more genetically diverse Tasmanian populations such as Big Waterhouse Lake (BWL) or a diverse lowland coastal swamp site in eastern Victoria such as Yallock Creek (YAC). Potential donor populations for the Gosling Creek site could include Mt Emu Creek lower (MEL) based on diversity and proximity or Darlot Creek (DLT) based on diversity and habitat characteristics.

Low genetic diversity in *G. pusilla* populations may reflect dispersal, likely to involve long distance movement of a few fish during floods between previously disconnected habitats (e.g. wetlands, swamps, billabongs, stream pools). Genetic drift in very dry periods may further deplete diversity because habitats dry and contract into refuge pools with a much lower carrying capacity. In Australia, several other species of freshwater fish may be prone to similar forces: these include golden perch (e.g. Faulks *et al.* 2010), eel-tailed catfish (e.g. Huey *et al.* 2006) and southern pygmy perch (e.g. Cook *et al.* 2007). Similarly, overseas fish species that are prone to founder events and/or bottlenecks due to drought or anthropogenic changes could also show a similar genetic pattern.

### *Patterns of diversity across Galaxiella pusilla*

Genetic structure across the species range is largely consistent with an 'isolation by distance' model, with a significant relationship between genetic distance and geographic distance, as well as (palaeo-) river distance. Given the relative strength of these relationships for both geographic and river distances, gene flow is likely to have been facilitated by both overland events (e.g. floods, geological/fluvial geomorphological changes in drainage boundaries) and connectivity of drainage networks (contemporary and palaeochannels). The relatively lower genetic diversity and population differentiation in the east region may be indicative of historical bottlenecks or more recent range expansions/population growth. The availability of refuge habitats in the west (e.g. expansive swamps and numerous volcanic lakes) may have helped preserve genetic diversity during periods of extreme aridity (e.g. last glacial maxima) and could account for the contrasting interregional genetic patterns (Coleman *et al.* 2010). *M* ratios indicate a much greater prevalence of recent bottlenecks in the east region than the west region; however, we failed to find consistent evidence of longer-term range expansion or bottlenecks involving multiple populations despite using multiple mtDNA markers and extensive sampling. Although it is difficult to differentiate between the influence of human impacts and palaeo-biogeographic history, because differences in genetic diversity between west and east populations appear to be greater for most of the west region, it is more likely that regional differences in genetic diversity are associated with palaeo-biogeographic history than anthropogenic disturbance given there are varying levels of human disturbance throughout both regions.

For many populations, genetic similarity can be attributed to dispersal throughout contemporary drainage systems (e.g. several South Australian populations), ancient drainage systems (e.g. palaeodrainage network that connected Darby River, Flinders Island and Tasmanian mainland populations or expanded Lake Corangamite system that once connected the headwaters of the Barwon River system and the Mt Emu Creek) or previously connected systems (e.g. progressive draining of the Koo Wee Rup swamplands in the late 1800s–early 1900s that disconnected Cardinia Creek, Bunyip River and Yallock Creek populations). Connections between geographically close populations during wet periods (e.g. large floods, swampland connections) may also be important for gene flow such as between Port Phillip and Western Port waterways (e.g. headwaters of the Hallam Valley, Boggy Creek or Balcombe Creek connecting with Cardinia Creek or Watsons Creek). Events that altered river alignments are expected to have facilitated expansion and interconnection of dwarf galaxias populations, particularly within South Australia and Western Victoria. For example, the distribution of dwarf galaxias in South Australia overlaps with historical alignments of the lower Glenelg River and the intermittent formation of

expansive swamps described by Boutakoff (1963). Lava flows at various times were also responsible for diverting the Crawford River into the Glenelg River (formerly joined the Fitzroy River), diverting the upper reaches of the Wannon River into the Glenelg River (formerly flowed via the Fitzroy-Darlot valleys to the sea) and separating Darlot Creek and Eumeralla River (formerly joined southwest of Mount Eccles) (Boutakoff 1963; Douglas & Ferguson 1976).

## Concluding remarks

Translocations are becoming more popular as an approach for conserving biodiversity, particularly for populations of threatened species that have low numbers of individuals and are highly fragmented (Weeks *et al.* 2011). Here, we have identified a strong relationship between genetic diversity and genetic uniqueness that appears to be driving genetic structure within two ESUs of *G. pusilla*, a pattern that may exist more widely in freshwater fish and other species whose population structure is affected by genetic drift. Threatened species may be particularly prone to this phenomenon due to their small population size. In this case, it is important to conserve genetic diversity generally rather than 'uniqueness' alone. From a genetic perspective, priority should be given to preserving populations that together account for the greatest proportion of the total gene pool, thereby increasing the potential for future adaptation and persistence. If there is evidence for local adaptation through breeding experiments or genetic signatures of selection, this information can also be used in prioritizing populations (Funk *et al*. 2012). Also, populations with a greater genetic diversity are likely to be suitable candidates for sourcing individuals when translocating into other populations at risk from inbreeding depression and local extinction. To increase the chances of species recovery, it is also essential that ecological issues (e.g. habitat degradation and interactions with other species) are effectively integrated with genetic management strategies (Moran 2002).

## Acknowledgements

We thank Melbourne Water Corporation for funding this research. Thanks to Vin Pettigrove and Tarmo Raadik for general advice throughout the study, and thanks to three anonymous reviewers for detailed comments. For ongoing assistance on laboratory methods and analysis, thanks to Nancy Endersby and Adam Miller. Thanks to Frank Clancy, Vin Pettigrove, Edward Tsyrlin, Elise Furlan, Bryant Gagliardi and to Megan, James and Harry Coleman for field assistance. Thanks to Peter Unmack for cytochrome *b* primers and Michael Hammer for tissue samples from Tasmania. Thanks also to Anthony Bright for assistance with distance calculations. Research was carried out with Melbourne University animal ethics approval (AEC 0705391.2) and State research permits for threatened species, national parks and fisheries permits for South Australia (PIRSA 9902333, DEH M25833-1), Victoria (DPI RP916, DSE 10004233) and Tasmania (IFS 2008-7, TFA 08090). A.A.H. and A.R.W. were supported by fellowships from the Australian Research Council.

## References

Barker JSF (2001) Conservation and management of genetic diversity: a domestic animal perspective. *Canadian Journal of Forestry Research*, **31**, 588–595.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2004) *GENETIX 4.03, logiciel sous WindowsTM pour la génétique des populations*. Laboratoire Génome, Populations, Interactions, Université de Montpellier II, Montpellier, France.

Blacket MJ, Robin C, Good RT, Lee SF, Miller AD (2012) Universal primers for fluorescent labelling of 1 PCR fragments – an efficient and cost effective approach to genotyping by fluorescence. *Molecular Ecology Notes*, **12**, 456–463.

Bonin A, Nicole F, Pompanon F, Miaud C, Taberlet P (2007) Population Adaptive Index: a new method to help measure intraspecific genetic diversity and prioritize populations for conservation. *Conservation Biology*, **21**, 697–708.

Boutakoff N (1963) The geology and geomorphology of the Portland Area. *Memoirs of the Geological Survey of Victoria*, **22**, 172pp.

Bureau of Rural Sciences (1999) *National Policy for the Translocation of Live Aquatic Organisms-Issues, Principles and Guidelines for Implementation*. Fisheries and Forestry Science Division, Canberra, ACT.

Caballero A, Toro MA (2002) Analysis of genetic diversity for the management of conserved subdivided populations. *Conservation Genetics*, **3**, 289–299.

Coleman RA, Pettigrove V, Raadik TA, Hoffmann AA, Miller AD, Carew ME (2010) Microsatellite markers and mtDNA indicate two distinct groups in dwarf galaxias, *Galaxiella pusilla* (Mack) (Pisces: Galaxiidae), a threatened freshwater fish from south-eastern Australia. *Conservation Genetics*, **11**, 1911–1928.

Coleman RA, Weeks AR, Hoffmann AA (2012) Data from: 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). *Dryad Digital Repository*, doi: 10.5061/dryad.8566p.

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.

Crawford NG (2010) SMOGD: software for the measurement of genetic diversity. *Molecular Ecology Resources*, **10**, 556–557.

Douglas JG, Ferguson JA (eds) (1976) *Geology of Victoria*. Geological Society of Australia Special Publication No.5. Geological Society of Australia (Victorian Division), Melbourne, Australia.

Excoffier L, Laval G, Schneider S (2006) *Arlequin ver 3.1: An Integrated Software Package for Population Genetics Data Analysis*.

Computational and Molecular Population Genetics Laboratory, University of Berne, Switzerland.

Faulks LK, Gilligan DM, Beheregaray LB (2010) 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.

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

Frankham R (2003) Genetics and conservation biology. *Comptes Rendus Biologies*, **326**, S22–S29.

Fraser DJ, Weir LK, Bernatchez L, Hansen MM, Taylor EB (2011) Extent and scale of local adaptation in salmonid fishes: review and meta-analysis. *Heredity*, **106**, 404–420.

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.

Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. *Genetics*, **133**, 693–709.

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

Garza JC, Williamson EG (2001) Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology*, **10**, 305–318.

Goudet J (2001) FSTAT, A Program to Estimate and Test Gene Diversities and Fixation Indices (version 2.9.3). Available from: http://www.unil.ch/izea/softwares/fstat.html.

Guillot G, Mortier F, Estoup A (2005) GENELAND: a computer package for landscape genetics. *Molecular Ecology Notes*, **5**, 712–715.

Harpending HC, Sherry ST, Rogers AR, Stoneking M (1993) The genetic structure of ancient human populations. *Current Anthropology*, **34**, 483–496.

Harris P, Heap A, Passlow V *et al.* (2005) *Geomorphic Features of the Continental Margin of Australia*. Geoscience Australia, Record 2003/30, 142pp.

Hey J, Waples RS, Arnold ML, Butlin RK, Harrison RG (2003) Understanding and confronting species uncertainty in biology and conservation. *Trends in Ecology and Evolution*, **18**, 597–603.

Hood G (2002) *POPTOOLS*. CSIRO, Canberra.

Huey JA, Hughes JM, Baker AM (2006) Patterns of gene flow in two species of eel-tailed catfish, *Neosilurus hyrtlii* and *Porochilus argenteus* (Siluriformes: Plotosidae), in western Queensland's dryland rivers. *Biological Journal of the Linnean Society*, **87**, 457–467.

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

Jost L (2008) $G_{ST}$ and its relatives do not measure differentiation. *Molecular Ecology*, **17**, 4015–4026.

Koster W (2003) Threatened species of the world: *Galaxiella pusilla* (Mack 1936) (Galaxiidae). *Environmental Biology of Fishes*, **68**, 268.

Lewis PO, Zaykin D (2001) Genetic Data Analysis: A Computer Program for the Analysis of Allelic Data. Available from: http://lewis.eeb.uconn.edu/lewishome/software.html.

Miller MR, Brunelli JP, Wheeler PA *et al.* (2012) A conserved haplotype controls parallel adaptation in geographically distant salmonid populations. *Molecular Ecology*, **21**, 237–249.

Moran P (2002) Current conservation genetics: building an ecological approach to the synthesis of molecular and quantitative genetic methods. *Ecology of Freshwater Fish*, **11**, 30–55.

Moritz C (1994) Defining 'Evolutionary Significant Units' for conservation. *Trends in Ecology and Evolution*, **9**, 373–375.

Moritz C (1999) Conservation units and translocations: strategies for conserving evolutionary processes. *Hereditas*, **130**, 217–228.

Palsbøll PJ, Bérubé M, Allendorrf FW (2007) Identification of management units using population genetic data. *Trends in Ecology and Evolution*, **22**, 11–16.

Palumbi SR (1996) PCR and molecular systematics. In: *Molecular Systematics*, 2 edn. (eds Hillis D, Moritz C, Mable B), pp. 205–247. Sinauer Press, Sunderland, Massachusetts.

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

Pérez-Figueroa A, Saura M, Fernández J, Toro MA, Caballero A (2009) METAPOP—A software for the management and analysis of subdivided populations in conservation programs. *Conservation Genetics*, **10**, 1097–1099.

Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Ramos-Onsins SE, Rozas J (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, **19**, 2092–2100.

Rogers AR, Fraley AE, Bamshad MJ, Watkins WS, Jorde LB (1996) Mitochondrial mismatch analysis is insensitive to the mutational process. *Molecular Biology and Evolution*, **13**, 895–902.

Ronquist F, Huelsenbeck JP (2003) MRBAYES 3:Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.

Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, **145**, 1219–1228.

Rozas J, Sanchez-De I, Barrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.

Ryder OA (1986) Species conservation and systematics: the dilemma of subspecies. *Trends in Ecology and Evolution*, **1**, 9–10.

Saddlier S, Jackson J, Hammer M (2010) *National Recovery Plan for the Dwarf Galaxias Galaxiella pusilla*. Department of Sustainability and Environment, Melbourne.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139**, 457–462.

Sokal RR, Rohlf FJ (1995) *Biometry: The Principles and Practice of Statistics in Biological Research*. W.H. Freeman, New York.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, **24**, 1596–1599.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.

Wager R (1996) *Galaxiella pusilla*. In: *IUCN 2008*. 2008 IUCN Red List of Threatened Species. www.iucnredlist.org. Downloaded on 26 January 2009.

Ward RD, Tyler SZ, Innes BH, Last PR, Herbert PDN (2005) DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society Biological Sciences*, **1462**, 1847–1857.

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.

Willi Y, Van Buskirk J, Hoffmann AA (2006) Limits to the adaptive potential of small populations. *The Annual Review of Ecology, Evolution and Systematics*, **37**, 433–58.

Yue GH, David L, Orban L (2007) Mutation rate and pattern of microsatellites in common carp (*Cyprinus carpio* L.). *Genetica*, **129**, 329–331.

R.A.C. contributed to research design, performed the research, contributed to the analysis approach, analysed and interpreted the data, and wrote the article. A.R.W. contributed to the analysis approach, interpretation of results and writing of the article. A.A.H. contributed to the research design, analysis approach, interpretation of results and writing of the article.

## Data accessibility

mtDNA haplotype sequences: GenBank accession nos KC282473–KC282620.

Microsatellite data for each population, mtDNA haplotype sequences, population and sample information (Coleman *et al.* 2012): DRYAD entry doi: 10.5061/dryad.8566p.

## Supporting information

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

**Appendix S1** Bayesian inference phylogenetic trees for *Galaxiella pusilla* haplotypes generated from (a) 16S rDNA (16S), (b) cytochrome *c* oxidase subunit I (COI) and (c) cytochrome *b* (CytB) fragments of 482, 595, and 1083 bp in size, respectively.

**Table S1** $F_{ST}$ values for pairwise comparisons (below the diagonal) and geographic distance (km) (above the diagonal) between *Galaxiella pusilla* populations in the west region.

**Table S2** $F_{ST}$ values for pairwise comparisons (below the diagonal) and geographic distance (km) (above the diagonal) between *Galaxiella pusilla* populations in the east region.

**Table S3** Analysis of demographic history of *Galaxiella pusilla* populations, where populations are grouped according to east and west regions.

**Fig. S1** Two-dimensional plot showing the relationships among individuals of *Galaxiella pusilla* based on a multivariate factorial correspondence analysis of 10 loci from (a) the west region and (b) the east region.

**Fig. S2** Two-dimensional plot comparing genetic diversity (mean number of alleles and heterozygosity) to genetic uniqueness (mean cytochrome *b* Kimura 2-parameter distance) for each *Galaxiella pusilla* population within the east region (a) and (b) and the west region (c) and (d).