Molecular Phylogenetics and Evolution 109 (2017) 415–420

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Short Communication

Genome-wide SNPs resolve a key conflict between sequence and allozyme data to confirm another threatened candidate species of river blackfishes (Teleostei: Percichthyidae: *Gadopsis*)

Peter J. Unmack [a], Jonathan Sandoval-Castillo [b], Michael P. Hammer [c], Mark Adams [d,e], Tarmo A. Raadik [f], Luciano B. Beheregaray [b,*]

[a] *Institute for Applied Ecology, University of Canberra, ACT 2601, Australia*
[b] *Molecular Ecology Laboratory, School of Biological Sciences, Flinders University, Adelaide, SA 5001, Australia*
[c] *Natural Sciences, Museum and Art Gallery of the Northern Territory, Darwin, NT 0801, Australia*
[d] *Australian Centre for Evolutionary Biology and Biodiversity, University of Adelaide, SA 5005, Australia*
[e] *Evolutionary Biology Unit, South Australian Museum, Adelaide, SA 5000, Australia*
[f] *Applied Aquatic Ecology Section, Arthur Rylah Institute for Environmental Research, Department of Environment, Land, Water & Planning, 123 Brown Street, Heidelberg, VIC 3084, Australia*

ARTICLE INFO

Article history:
Received 21 November 2016
Revised 13 February 2017
Accepted 19 February 2017
Available online 22 February 2017

Keywords:
Gadopsis marmoratus
Freshwater fish
Australia
ddRAD-seq
Conservation genomics
Phylogenomics

ABSTRACT

Conflicting results from different molecular datasets have long confounded our ability to characterise species boundaries. Here we use genome-wide SNP data and an expanded allozyme dataset to resolve conflicting systematic hypotheses on an enigmatic group of fishes (*Gadopsis*, river blackfishes, Percichthyidae) restricted to southeastern Australia. Previous work based on three sets of molecular markers: mtDNA, nuclear intron DNA and 51 allozyme loci was unable to clearly resolve the status of a putative fifth candidate species (SWV) within *Gadopsis marmoratus*. Resolving the taxonomic status of candidate species SWV is particularly critical as based on IUCN criteria this taxon would be considered Critically Endangered. After all filtering steps we retained a subset of 10,862 putatively unlinked SNP loci for population genetic and phylogenomic analyses. Analyses of SNP loci based on maximum likelihood, fastSTRUCTURE and DAPC were all consistent with the previous and updated allozyme results supporting the validity of the candidate *Gadopsis* species SWV. Immediate conservation actions should focus on preventing take by anglers, protection of water resources to sustain perennial reaches and drought refuge pools, and aquatic and riparian habitat protection and improvement. In addition, a formal morphological taxonomic review of the genus *Gadopsis* is urgently required.

1. Introduction

Conflicting results from different molecular datasets have long confounded our ability to characterise species boundaries. There are many situations where single genes suggest phylogenetic relationships that are different to the true species tree. In some cases this may be a result of insufficient data or the use of inappropriate genes for the question being addressed. Introgression is a common phenomenon that exists with mitochondrial DNA (mtDNA), but introgression can also impact relationships obtained from nuclear DNA (Funk and Omland, 2003). In addition, ancestral lineage sorting, selection and other processes can lead to single genes having different histories to the species tree (Doyle, 1992; Brower et al., 1996). Over time our ability to collect molecular data from different marker types for obtaining species phylogenies has increased from allozymes (~50 loci, but difficult to use for phylogeny) and mtDNA restriction fragment analysis and sequence data (single locus), to sequencing one or a few nuclear genes (Beheregaray, 2008). Today next-generation sequencing (NGS) has massively increased the number of loci obtained and the breadth of the genome and research questions being examined (Garrick et al., 2015), including convenient NGS techniques such as anchored phylogenomics (hundreds of loci, e.g., Lemmon et al., 2012) and genotype by sequencing (thousands of loci, e.g. Peterson et al., 2012). While these larger datasets still have issues that can lead to incorrect relationships being inferred, the sheer number of loci and their distribution across the genome, combined with careful analysis, minimises the likelihood that different demographic histories and noisy data at different loci will drive the results (Jeffroy et al.,

* Corresponding author.
E-mail address: Luciano.Beheregaray@flinders.edu.au (L.B. Beheregaray).

http://dx.doi.org/10.1016/j.ympev.2017.02.013
1055-7903/

2006; Garrick et al., 2015). Here we provide an example of how genome-wide data resolves conflicting results from allozymes, mtDNA and nuclear DNA sequences in a genus of freshwater fishes.

Gadopsis (river blackfishes) are an enigmatic group of fishes restricted to rivers and streams of southeastern Australia. In the southern portion of their distribution *Gadopsis* have been recorded up to 5.4 kg and over 625 mm as the largest freshwater species present after angullid eels. Their size, eating qualities and easy catchability make them an attractive angling species, although large fish over 300 mm are rare today. In the northern portion of their range the species is smaller, rarely reaching 350 mm. *Gadopsis* are probably long lived (5–15 years), have slow growth, low fecundity, and limited home range (Jackson, 1978; Khan et al., 2004). Such ecological attributes (further summarised in Hammer et al., 2014) have predisposed most *Gadopsis* species to long-term ongoing declines across their range and the chances of recolonisation after local extirpation events are limited (Lean et al., 2016). Today they are often restricted to smaller fragmented habitats, whereas formerly they would have been widespread throughout most river systems across their range. Such a pattern of decline is common to many freshwater fish species in southeastern Australia and elsewhere (Hammer et al., 2013).

Hammer et al. (2014) conducted detailed range-wide examination of *Gadopsis* based on three sets of molecular markers: mtDNA (cytochrome *b* [cyt*b*]), nuclear DNA (two introns of S7) and 51 allozyme loci. This study clearly supported separating *Gadopsis marmoratus* s.l. into four candidate species, with a fifth candidate species (SWV) having strong support from allozyme data, but no support from mtDNA or nuclear sequence datasets. The partially supported fifth candidate species, SWV, has experienced the greatest decline in what was already one of the taxa with the narrowest range within *Gadopsis* (Hammer et al., 2014). Consequently, resolving the genomic conflict with the SWV taxon has key conservation outcomes along with the systematic implications. We generated genome-wide SNP data using double digest restriction site-associated sequencing (ddRAD-seq) and also expanded the allozyme dataset of Hammer et al. (2014) for the southern lineage of *Gadopsis* (consisting of the candidate species SEV, SBA and SWV) to resolve the conflict between previous molecular datasets to determine whether SWV is supported as a distinct candidate species.

2. Materials and methods

2.1. Taxon sampling and further allozyme profiling

Individuals included in Hammer et al. (2014) were subsampled to cover the breadth of diversity and geographic extent of the three candidate species (SEV, SBA, SWV; Fig. 1). A total of 45 individuals from 31 sites (n = 1–2 per site; Table 1, Fig. 1) were sampled for genome-wide variation.

Field surveys were undertaken to locate and sample additional *Gadopsis* populations in the region harbouring SWV and the three candidate species that surround it (SBA, NGW, NMD; Fig. 1). The new SWV collections (2 sites; n = 12; Table 1), plus selected existing or new sites representing all relevant candidate species (8 sites, n = 12; Table 1), expanded our published allozyme study (n = 163 *G. marmoratus* complex for 51 putative loci; Hammer et al., 2014). Allozyme profiles were generated and analysed as described in Hammer et al. (2014).

2.2. ddRAD library preparation and sequencing

Genomic DNA was extracted using a salting out protocol (Sunnucks and Hale, 1996), primarily from muscle tissue preserved in liquid nitrogen in the field, held in −80 °C freezers and transferred to the lab in 96% ethanol. The ddRAD libraries were constructed 'in house' (Molecular Ecology Lab at Flinders University) using 300 ng of DNA per sample and restriction enzymes SbfI-HF and MseI (New England Biolabs) using the protocol of Peterson et al. (2012) with modifications detailed in Brauer et al. (2016). Pools of 48 individual genomic libraries were combined and sequenced in one lane of a HiSeq 2000 Illumina (100 bp paired-end reads) platform at the McGill University and Genome Quebec Innovation Centre.

2.3. Read filtering and assembly

Raw sequences were demultiplexed and trimmed using the 'process_radtags.pl' function in Stacks 1.19 (Catchen et al., 2013). We initially recovered reads with up to two errors in the individual barcode and up to two errors in the restriction site sequence. All remaining reads were then trimmed to 80 bp by removing the barcode, the restriction site, and the last 8 bp. Demultiplexed reads were then processed using pyRAD 3.0 (Eaton, 2014). First, bases with a Phred quality score <30 were replaced with N, and sequences having >four Ns (5%) were discarded. Filtered reads were clustered by an 80% threshold similarity, then clusters with >10× coverage per individual, <10% missing data per locus, and <0.6 observed heterozygosity were considered ddRAD loci and used to generate a concatenated sequence matrix for phylogenomic analysis. Ambiguity codes were used to represent heterozygous SNPs.

To explore fine scale genetic structure, we selected one SNP per ddRAD locus; if multiple SNPs were present in a locus, the one with the least missing data across taxa was selected.

2.4. Phylogenetic analyses

We reconstructed phylogenetic relationships using the concatenated sequence matrix and a maximum likelihood approach implemented in RAxML 8.0 (Stamatakis, 2014). We used the GTR + gamma model of sequence evolution, applying the rapid bootstrap algorithm with 1000 replicates, and performed a full search for best scoring maximum likelihood tree. The final tree was midpoint rooted as is it was not possible to provide an unbiased outgroup. Within Percichthyidae the sister group to *Gadopsis* remains unclear and it is evolutionarily very distant (Near et al., 2012). Within *Gadopsis* the only phylogenetic information is based on cyt*b* and S7 (Hammer et al., 2014), however, using any of the other *Gadopsis* lineages as outgroups, would have to assume the original topology was correct.

2.5. Hierarchical structure

Population structure was examined using the SNP data set with model-free and model-based approaches. First, we performed Discriminant Analysis of Principal Components (DAPC) using the R package Adegenet 1.4 (Jombart and Ahmed, 2011). We used K-mean and Bayesian Information Criterion to determine the best supported number of clusters with more than two samples and the minimum number of clusters necessary to explain the genetic variation in our data. The major advantage of DAPC is its lack of assumptions about evolutionary genetic models (Jombart and Ahmed, 2011). We also used a Bayesian clustering algorithm implemented in the software fastSTRUCTURE (Raj et al., 2014). We ran five replicates for each K value from 1 to 15, using the simple prior model. Then the number of clusters that best fit our data was determined using the utility 'chooseK.py' included in fastSTRUCTURE. This software defines the number of clusters minimizing departures from Hardy-Weinberg equilibrium and

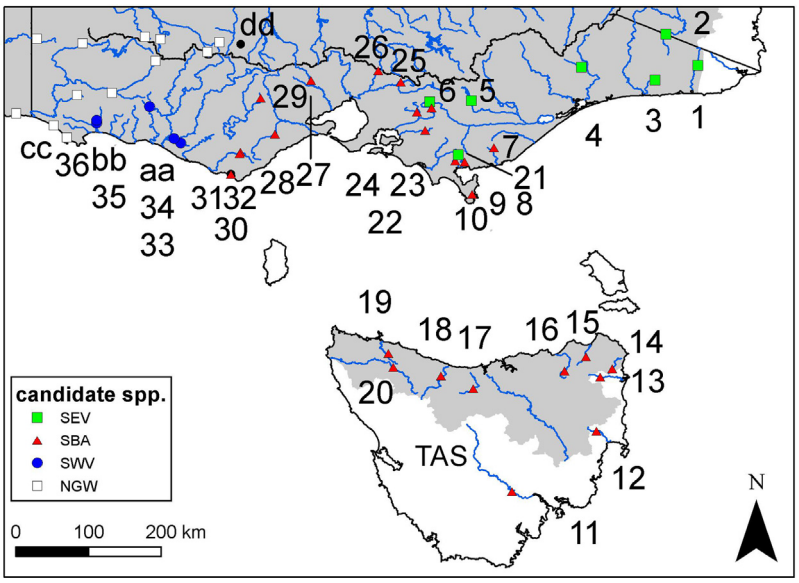

Fig. 1. Locality data for all *Gadopsis* samples using the same numbering scheme as Hammer et al. (2014), plus four new sites (referenced using letters). Note that not all sites shown were included in the current study. Refer to Table 1 for the corresponding site details. The shaded area identifies the known distribution of *Gadopsis marmoratus*. The symbol shape and colour refers to each candidate species.

Table 1
Locality data for all *Gadopsis* populations examined. Site refers to the localities shown in Fig. 1 and Hammer et al. (2014), sites with letters are newly added populations in the allozyme study. Station codes can be used to track references to genetic material deposited in the South Australian Museum and morphological samples deposited in the Australian, South Australian and Victorian museum collections. State abbreviations include NSW = New South Wales, TAS = Tasmania, VIC = Victoria, and sample sizes for ddRAD (N) and allozymes (N2) are shown for this study. Latitude and longitude are provided in decimal degrees. Candidate species (cand spp) represents the name given to each taxon identified in this study and Hammer et al., 2014.

Site	Station code	Locality	Drainage	N	N2	Latitude	Longitude	Cand spp
1	TR02-24	Back Creek, Noorinbee North, VIC	E. Gippsland	2		−37.42944	149.20750	SEV
2	PU02-63	Delegate R, Delgate, NSW	Snowy	2		−37.04750	148.81167	SEV
3	PU02-64	Brodribb R, VIC	Snowy	1		−37.61278	148.67472	SEV
4	PU02-66	Haunted Stream, VIC	Tambo	1		−37.45278	147.76556	SEV
6	PU02-81	Latrobe R, Noojee, VIC	Latrobe	2		−37.88194	145.89250	SEV
7	PU02-97	Greig Ck, Yarrum, VIC	S. Gippsland	2		−38.44917	146.68361	SBA
8	PU02-101	Tin Mine Ck, VIC	S. Gippsland	1		−38.62500	146.32333	SBA
9	PU02-95	Deep Ck, Forster, VIC	S. Gippsland	2		−38.60806	146.20861	SBA
10	PU02-72	Blackfish Ck, Wilsons Prom, VIC	S. Gippsland	2		−39.02194	146.41778	SBA
11	TAS05-07	Styx R, Bushy Park, TAS	Derwent	1		−42.70139	146.90750	SBA
14	TAS03-11	Ansons R, Ansons Bay, TAS	East	1		−41.18028	148.14500	SBA
16	TAS03-20	Great Forester R, Scottsdale, TAS	Piper	1		−41.20750	147.55833	SBA
17	TAS03-03	Minnow R, Beulah, TAS	Mersey	1		−41.42861	146.43167	SBA
18	TAS03-04	Leven R, Gunns Plains, TAS	Smithton	1		−41.26944	146.03028	SBA
19	TAS03-27	Black R, Mawbanna, TAS	Smithton	1		−40.99139	145.37445	SBA
20	TAS03-23	Relapse Ck, TAS	Arthur	1		−41.16056	145.43583	SBA
21	PU02-78	Turtons Ck, VIC	S. Gippsland	1		−38.53444	146.24611	SEV
22	PU02-105	Minnieburn Ck, VIC	Bunyip	1		−38.23556	145.83445	SBA
23	PU02-82	Tarago R, VIC	Bunyip	2		−37.95694	145.91417	SBA
24	PU03-05	Diamond Ck, Tonimbuk, VIC	Bunyip	1	2	−38.00472	145.73222	SBA
25	TR02-268	Donnellys Ck, Healesville, VIC	Yarra	2		−37.63639	145.53333	SBA
26	TR02-16	Running Ck, Kinglake, VIC	Yarra	2		−37.49444	145.24972	SBA
27	TR02-210	Lerderderg R, VIC	Werribee	1		−37.61722	144.42222	SBA
28	PU02-85, PU02-108	Barwon R, Winchelsea, VIC	Barwon	1		−38.27917	143.97584	SBA
29	TR02-373	Kuruc-A-Ruc Ck, Dereel, VIC	Corangamite	2		−37.83750	143.79861	SBA
30	PU02-110	Ford R, VIC	Otway	1		−38.73528	143.42056	SBA
31	PU02-109	Loves Ck, Gellibrand, VIC	Otway	1	2	−38.50472	143.55056	SBA
32	FISH93:Gell	Gellibrand R, Gellibrand, VIC	Otway	1		−38.52000	143.53833	SBA
33	PU03-08/09	Brucknells Ck, Naringal East, VIC	Hopkins	3	2	−38.39222	142.80917	SWV
34	PU02-112, PU03-07	Mount Emu Ck, Panmure, VIC	Hopkins	2		−38.33667	142.72889	SWV
aa	TR15-177	Mustons Ck	Hopkins		6	−37.93401	142.42663	SWV
35	FISH93:Darl	Darlots Ck, VIC	Portland	1		−38.14694	141.77083	SWV
bb	TR15-003	Darlots Ck, VIC	Portland		6	−38.11959	141.78055	SWV
36	PU09-122	Bridgewater Lakes, VIC	Portland		2	−38.32075	141.40469	NGW
cc	PU09-128	Lake Monibeong	Glenelg		2	−38.13338	141.18576	NGW
dd	PU13-77	Bet Bet Ck	Loddon		2	−37.16508	143.552865	NMD

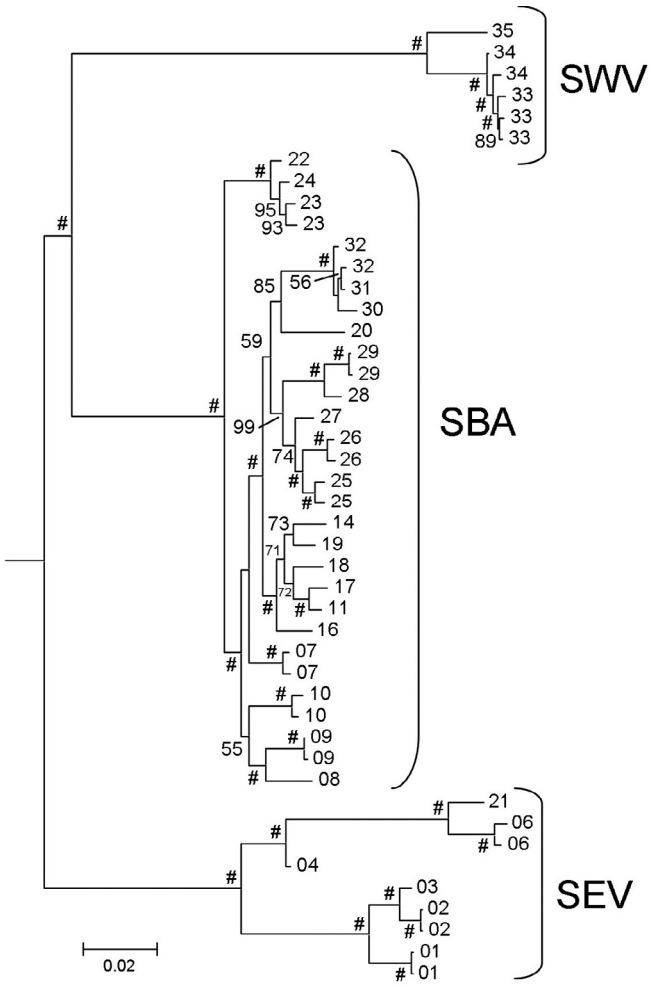

Fig. 2. Maximum likelihood tree for concatenated SNP loci from 45 *Gadopsis* individuals from three candidate species based on 1,541,113 bp per OTU. The major lineage branches are labelled according to candidate species. Each OTU code is based on the population number described in Table 1 and Fig. 1. The geographic distribution of clades is shown in Fig. 1. Bootstrap values shown are based on 1000 replicates, with a # representing nodes with a value of 100.

maximizing linkage equilibrium. Since it uses a variational Bayesian inference to determine the ancestry proportion in the model, it is computationally more efficient that other Bayesian clustering algorithms (Raj et al., 2014).

3. Results

3.1. Expanded allozyme analyses

Each of the 24 additional *Gadopsis* genotyped in this study displayed the allozyme profile expected for the predicted candidate species (Supplementary Fig. 1 and Table 1), as based on their geographic location (Fig 1). Accordingly, all previously-documented fixed allozyme differences remain diagnostic for the five candidate taxa in *G. marmoratus* (Table 3 in Hammer et al., 2014).

3.2. Sequence characteristics

We obtained >177 million raw reads and >17 billion bp of sequence data. Seventy five percent of the raw sequenced reads passed quality filtering, with an average of 4.6 million reads retained per individual. After alignment of filtered RAD sequences, a total of 20,224 SNPs were retained in >90% of all individuals. After all filtering steps we retained a subset 10,862 putatively unlinked SNPs (data files were deposited in Dryad, doi: 10.5061/dryad.9d12p) for analyses of population genetic structure. The final concatenated DNA sequence dataset yielded 1,541,113 bp per OTU for phylogenomic analysis.

3.3. Phylogenomic distinctiveness among recently diverged lineages

The ML phylogenetic reconstruction based on the concatenated sequence data identified three reciprocally monophyletic clades, each corresponding to the three candidate species (SWV, SBA, SEV) supported by 100% bootstrap (Fig. 2). One clade consisted of individuals from SEV, with a deep split separating eastern and western populations. SBA and SWV formed well separated sister clades (Fig. 2).

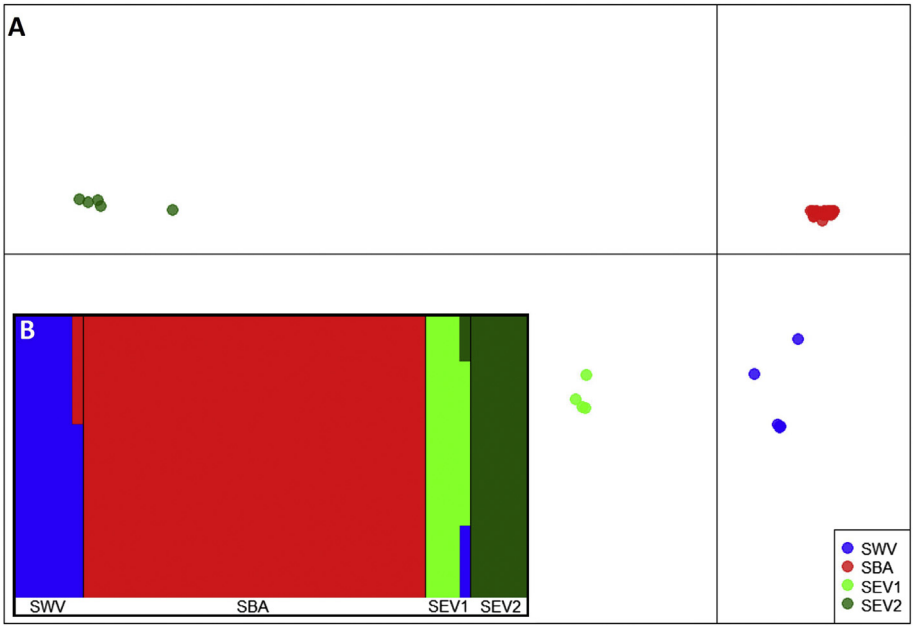

Fig. 3. (A) Discriminant analysis of principal components and (B) fastSTRUCTURE plot based on 10,862 putatively unlinked SNPs for three *Gadopsis* candidate species.

3.4. Population differentiation

Patterns of population structure identified with the model-based fastSTRUCTURE and the model-free DAPC were both largely consistent with each other. FastSTRUCTURE and DAPC both identified four main population clusters that corresponded to the three candidate species (SWV, SBA and SEV), with SEV containing two clusters (Fig. 3). The membership for individuals from each of these four groups exactly matched their assignments to SWV, SBA, and the two groups within SEV identified with the phylogenomic dataset. The two exceptions relate to putatively admixed individuals (one in SWV and the other in SEV1) that showed low ancestry to other cluster populations in the FastSTRUCTURE plots.

4. Discussion

4.1. Clarifying incongruence

Our findings clearly demonstrate that the phylogenomic and the SNP data support the previous and updated allozyme results relative to the number and distribution of candidate *Gadopsis* species. While broadly congruent with each other, the mitochondrial and nuclear intron sequences presented in Hammer et al. (2014) appear as misleading relative to species boundaries based on results from nuclear loci across the broader genome. It seems most likely these misleading results relate to locus specific characteristics that are different to the overall patterns of species diversification, possibly due to introgression between SWV and SBA, lineage sorting or similar selective regimes.

Other discrepancies between these datasets exist. SNPs and phylogenomic data clearly group all populations in SEV together (Figs. 1–2), whereas mitochondrial data had the westernmost populations (6 and 21) as the first branching lineage (Hammer et al., 2014), rather than being monophyletic with remaining SEV populations. In addition, population 4 grouped with populations 6 and 21 in the phylogenomic and SNP data (Figs. 1–2), which is a better fit based on riverine connectivity, whereas in the mitochondrial dataset population 4 was closely related to those further east (Hammer et al., 2014).

The allozyme dataset indicated two distinct phylogeographic lineages within SWV ($SWV^{\#1}$ and $SWV^{\#2}$, with three fixed differences; Supplementary Material) were not particularly distinctive in the whole genome analyses. This highlights that even allozymes with a moderate number of loci sampled (~50) can still be prone to locus specific effects that are different from the massively larger number of SNP loci sampled.

4.2. Utility of SNP loci

Approaches using SNPs in phylogenetic and phylogeographic contexts are becoming more common and provide considerable advantages over previous methods (e.g., mtDNA and/or small numbers of nuclear loci) that have limitations (Doyle, 1992; Brower et al., 1996). Despite limitations, datasets based on only a few loci are often surprisingly accurate, but difficult to interpret when obvious incongruence arises as would be expected when only a few loci are available for comparison. The use of SNP loci is relatively new, and some issues may exist (e.g., gene tree/species tree, inability to partition data to apply different models of sequence evolution, potentially high numbers of missing loci, Rivers et al., 2016). However, they appear to provide robust phylogenetic hypotheses based on their often similar results to previous studies (e.g., Leaché et al., 2015), along with evidence suggesting concatenation of many SNP loci may not be strongly effected by gene tree/species tree issues (Rivers et al., 2016).

4.3. Conservation status of SWV

Specific historical distribution data for *Gadopsis* SWV are absent from the scientific literature, however some information is contained in early newspaper reports. *Gadopsis* SWV was known to have been present upstream to at least the mid reaches of the Hopkins River Basin (HRB) (345 m in elevation) in the late 1800s (Flyrod, 1886; C.L.F., 1892; F.R., 1907). The taxon was also recorded as present in Brucknell Creek, a tributary of the lower Hopkins River downstream of Hopkins Falls, and the Merri River system (all HRB), as well as streams of the Portland Coast Basin (PCB), including the Eumeralla and Fitzroy rivers and Darlot Creek. Though not present by the 1970s (Tunbridge and Rogan, 1976), it is highly likely that the species was also found in the intervening streams (e.g. Shaw and Moyne rivers (all PCB)), as these also had similar aquatic habitat, though data are lacking. *Gadopsis* SWV has declined substantially since the late 1800s, with lower abundance evident by the early 1900s (F.R., 1907). Population abundance had been noted to increase slightly over wetter seasons, though large fish were absent (F.R. 1909). Probable causes for decline were overfishing, negative interactions with alien species, riparian vegetation loss, swamp reclamation and stream channelisation as forests were converted into pasture.

Declines appear to be ongoing. Recent sampling has confirmed only three small, geographically isolated populations: (1) lower reaches of Mount Emu Creek (pop. 34) and immediate adjoining area of Hopkins River, (2) a short section of the upper Muston Creek (pop. aa, also HRB), and (3) Darlot Creek (pop. 35 & bb, PCB). These reaches appear to be perennial, with aquatic habitat sustained during dry periods by groundwater outflow. Local extirpation is feared for the lower Fitzroy River (PCB) as a small number of adults were recorded in 2007, but the taxon could not be relocated during sampling in 2014–2015 (Raadik, unpublished data). This time period straddles the end of the 'Millennium Drought' (van Dijk et al., 2013) and extended dry conditions may have caused the extirpation of this population and others such as Merri River and Brucknell Creek (cf. Hammer et al. 2013).

Based on IUCN (2012) criteria this taxon should be considered Critically Endangered based on B2a,b(i–v), c(i–iv): extent of occurrence is estimated as 3170 km^2 but area of occurrence is estimated at 1.8 km^2. Immediate conservation actions should focus on preventing take by anglers, protection of water resources to sustain perennial reaches and drought refuge pools, and aquatic and riparian habitat protection and improvement such as fencing and physical habitat supplementation. A formal morphological taxonomic review of the genus *Gadopsis* is urgently required.

Acknowledgements

We thank the many people who helped with laboratory and field work and other aspects relative to obtaining specimens. All collections were obtained under permit from various state fisheries agencies and research was conducted under Flinders University Animal Welfare Committee approval E313. Financial support was provided by the Australian Research Council to L.B.B. (FT130101068).

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2017.02.013.

References

Beheregaray, L.B., 2008. Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. Mol. Ecol. 17, 3754–3774.

Brauer, C.J., Hammer, M.P., Beheregaray, L.B., 2016. Riverscape genomics of a threatened fish across a hydroclimatically heterogeneous river basin. Mol. Ecol. 25, 5093–5113.
Brower, A.V.Z., DeSalle, R., Vogler, A., 1996. Gene trees, species trees, and systematics: a cladistic perspective. Annu. Rev. Ecol. Syst. 27, 423–450.
Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., Cresko, W.A., 2013. Stacks: an analysis tool set for population genomics. Mol. Ecol. 22, 3124–3140.
C.F.L., 1892. Across the border: fishing at Emu Creek (VIC). Aust. Town Country J. 1 (October), 38.
Doyle, J.J., 1992. Gene trees and species trees: molecular systematics as one-character taxonomy. Syst. Bot. 17, 144–163.
Eaton, D.A., 2014. PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. Bioinformatics, btu121.
Flyrod, 1886. The blackfish. The Leader 27 (November), 22.
F.R., 1907. Eels and blackfish. The Australasian (March), 51.
F.R., 1909. Eels and blackfish. The Australasian (March), 51.
Funk, D.J., Omland, K.E., 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annu. Rev. Ecol. Evol. Syst. 34, 397–423.
Garrick, R.C., Bonatelli, I.A.S., Hyseni, C., Morales, A., Pelletier, T.A., Perez, M.F., Rice, E., Satler, J.D., Symula, R.E., Thomé, M.T.C., Carstens, B.C., 2015. The evolution of phylogeographic data sets. Mol. Ecol. 24, 1164–1171.
Hammer, M.P., Bice, C.M., Hall, A., Frears, A., Watt, A., Whiterod, N.S., Beheregaray, L. B., Harris, J.O., Zampatti, B.P., 2013. Freshwater fish conservation in the face of critical water shortages in the southern Murray-Darling Basin, Australia. Mar. Freshw. Res. 64, 807–821.
Hammer, M.P., Unmack, P.J., Adams, M., Raadik, T.A., Johnson, J.B., 2014. A multi-gene molecular assessment of cryptic biodiversity in the iconic freshwater blackfishes (Teleostei: Percichthyidae: *Gadopsis*) of southeastern Australia. Biol. J. Linn. Soc. 111, 521–540.
IUCN, 2012. IUCN Red List Categories and Criteria: Version 3.1. IUCN, Gland, Switzerland & Cambridge, UK. 32pp.
Jackson, P.D., 1978. Spawning and early development of the river blackfish, *Gadopsis marmoratus* Richardson (Gadopsiformes: Gadopsidae), in the McKenzie River, Victoria. Aust. J. Mar. Freshwater Res. 29, 293–298.
Jeffroy, O., Brinkmann, H., Delsuc, F., Philippe, H., 2006. Phylogenomics: the beginning of incongruence? Trends Genet. 22, 225–231.
Jombart, T., Ahmed, I., 2011. Adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27, 3070–3071.
Khan, M.T., Khan, T.A., Wilson, M.E., 2004. Habitat use and movement of river blackfish (*Gadopsis marmoratus* R.) in a highly modified Victorian stream, Australia. Ecol. Freshwater Fish 13, 285–293.
Leaché, A.D., Banbury, B.L., Felsenstein, J., de Oca, A.N.M., Stamatakis, A., 2015. Short tree, long tree, right tree, wrong tree: new acquisition bias corrections for inferring SNP phylogenies. Syst. Biol. 64, 1032–1047.
Lean, J., Hammer, M.P., Unmack, P.J., Adams, M., Beheregaray, L.B., 2016. Landscape genetics informs mesohabitat preference and conservation priorities for a surrogate indicator species in a highly fragmented river system. Heredity. http://dx.doi.org/10.1038/hdy.2016.111.
Lemmon, A.R., Emme, S.A., Lemmon, E.M., 2012. Anchored hybrid enrichment for massively high-throughput phylogenomics. Syst. Biol. 61. 727-244.
Near, T.J., Sandel, M., Kuhn, K.L., Unmack, P.J., Wainwright, P.C., Smith, W.L., 2012. Nuclear gene-inferred phylogenies resolve the relationships of the enigmatic Pygmy Sunfishes, *Elassoma* (Teleostei: Percomorpha). Mol. Phylogenet. Evol. 63, 388–395.
Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., Hoekstra, H.E., 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS ONE 7, e37135.
Raj, A., Stephens, M., Pritchard, J.K., 2014. fastSTRUCTURE: Variational inference of population structure in large SNP data sets. Genetics 197, 573–589.
Rivers, D.M., Darwell, C.T., Althoff, D.M., 2016. Phylogenetic analysis of RAD-seq data: examining the influence of gene genealogy conflict on analysis of concatenated data. Cladistics 32, 672–681. http://dx.doi.org/10.1111/cla.12149.
Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics, btu033.
Sunnucks, P., Hale, D.F., 1996. Numerous transposed sequences of mitochondrial cytochrome oxidase I–II in aphids of the genus Sitobion (Hemiptera: Aphididae). Mol. Biol. Evol. 13, 510–524.
Tunbridge, B.R., Rogan, P., 1976. A Guide to the Inland Angling Waters of Victoria. Ministry for Conservation, Victoria, p. 149.
van Dijk, A.I.J.M., Beck, H.E., Crosbie, R.S., de Jeu, R.A.M., Liu, Y.Y., Podger, G.M., Timbal, B., Viney, N.R., 2013. The Millennium Drought in southeast Australia (2001–2009): natural and human causes and implications for water resources, ecosystems, economy, and society. Water Resour. Res. 49, 1040–1057.